# Computational Insights into Mouse Placental Histology


Peter McCann Strain

St Anne's College

University of Oxford

A thesis submitted for the degree of

*Doctor of Philosophy*

Michaelmas 2023

# Acknowledgements

Undertaking this PhD has been the most challenging endeavor of my life, yet it has also spurred the most significant personal growth I've ever experienced. I am acutely aware that this achievement is not solely mine. I am lifted atop the shoulders of giants, without whom I would not be where I am today. In light of this, I would like to extend my deepest gratitude to some key individuals.

Firstly, my appreciation goes to my supervisors, Cecilia, Chris, Annie and Henrik. Even through challenging times, the lessons I've learned from each of you have been invaluable, contributing immensely to my growth. My gratitude extends to the MRC for their generous funding of my PhD. The opportunities for learning and growth provided by your support have been fundamental to my journey. And to Rob Gilbert, your commitment to students and your remarkable kindness are truly inspiring. I also extend my gratitude to my college community at St Annes and St Johns. I learned so much in the MCR and the sports team, and speaking to so many interesting people kept me sane and taught a boy who hadn't seen much beyond Glasgow a whole lot about the world.

I must express profound gratitude to my family, whose sacrifices have significantly shaped my journey. The path to where I am today can be traced back generations.The diligence of my grandparents on both sides cultivated an environment where my parents could learn vital life lessons, ultimately leading to me becoming the first in our family to attend university. The understanding and resilience they demonstrated highlighted education as a pathway out of poverty. This foresight laid the groundwork for my academic pursuits.

A special acknowledgement goes to my grandparents, whose unwavering emotional and financial support eased the challenges I encountered. My grandfather's pride in taking me to and from university, showcasing me around campus on my first undergraduate day, and my grandmother's inspirational tales of trekking miles to school and overcoming greater obstacles than I have faced for her education are memories that continually inspire me. Their sacrifices and teachings have instilled in me the determination to persevere.

To my parents: My father, who selflessly raised us, ferrying my siblings and me to countless sports and clubs, made untold sacrifices for our well-being. The foundation of my achievements rests on these sacrifices. And to my mother, who encouraged

us to voice our opinions from a young age, defying contrary advice from respected individuals. Her insistence on diligence in our studies and her motivational stories about the power of education assured us that we could achieve anything. This belief has been instrumental in shaping who I am. My debt to both of you is immeasurable.

Additionally, I extend my thanks to my sister and brother, whose constant challenges and inspiration have propelled me forward. Your unwavering support and encouragement have been invaluable. I owe you both immensely.

And finally to Amy Perkins for pushing me to the end and being a constant cheerleader.

.

# Abstract

This thesis represents a significant contribution to the field of placental biology, combining advanced computational techniques with traditional biological research, with a focus on computer vision methods. I utilised the Deciphering the Mechanisms of Developmental Disorders (DMDD) dataset as a factual basis, confirming and demonstrating the capabilities of computational paradigms in the detailed analysis of placental histology.

The research consists of four distinct phases, each aiming to provide in-depth information for the investigation. The first phase involved thorough annotation and preparation of the DMDD dataset to ensure that subsequent computational analyses were methodologically sound and valid. This phase formed the foundation for the rest of the study.

The second phase introduced an automated placental phenotyping framework that uses machine learning algorithms such as RetinaNet and InceptionResNetV2. This approach transcended species boundaries, encompassing both mouse and human placentae, and improved the categorisation of cellular and histological features with state of the art accuracy. This framework represented a significant shift in approach by eliminating the limitations of traditional manual labor and reducing interpretive subjectivity.

During the third phase of the study, the DMDD dataset was thoroughly analyzed using an automated pipeline. This analysis helped to identify unknown phenotypic markers, particularly in E14.5 mouse placentae. The newfound markers were carefully verified through scholarly corroboration and cross-referencing, and they played a crucial role in advancing knowledge of the complex interplay between placental and embryonic phenotypes in knockout mice. Machine learning models like XGBoost were used for further feature analysis to support the findings, adding an extra layer of empirical robustness.

In the last phase of this thesis, I explored the intricate world of unsupervised and self-supervised learning techniques. I made use of advanced computational approaches such as t-distributed Stochastic Neighbour Embedding (t-SNE), Uniform Manifold Approximation and Projection (UMAP), and SimCLR to gain a deeper understanding of the intraclass variance in the cellular structure of placental tissue. My investigations provided initial insights into this complex field and highlighted methodological guidelines, challenges, and opportunities for future research.

This thesis builds on existing literature as a catalyst for transforming thinking in the field. The study demonstrates how computational methods, specifically computer vision, can revolutionize our comprehension of placental biology and developmental studies. This research blends technological innovation with biological discovery to enhance our understanding of the placenta.

# Contents

# List of Figures

# List of Abbreviations

. . . . . ADAM - Adaptive Moment Estimation

. . . . . AI - Artificial Intelligence

. . . . . AMI - Adjusted Mutual Information

. . . . . ANOVA - Analysis of Variance

. . . . . ANNs - Artificial Neural Networks

. . . . . ARI - Adjusted Rand Index

. . . . . BYOL - Bootstrap Your Own Latent

. . . . . CEL - Cross-Entropy Loss

. . . . . CH Index - Calinski-Harabasz Index

. . . . . CHP - Chorionic Plate Cells

. . . . . CNN - Convolutional Neural Network

. . . . . CNVs - Copy Number Variations

. . . . . CoCo - Common Objects in Context

. . . . . CPUs - Central Processing Units

. . . . . CRISPR - Clustered Regularly Interspaced Short Palindromic Repeats

. . . . . CT - Computed Tomography

. . . . . CTB - Cytotrophoblasts

. . . . . DEC - Decidual Cells

. . . . . DMDD - Deciphering Mouse Developmental Disorders

. . . . . E14.5 - Embryonic Day 14.5

. . . . . E9.5 - Embryonic Day 9.5

. . . . . EM - Expectation-Maximization

. . . . . END - Endothelial Cells

. . . . . eEVT - Endovascular Extravillous Trophoblast

. . . . . ExM-dC - Extraembryonic Mesoderm-Derived Core

. . . . . FDR - False Discovery Rate

. . . . . FAIR - Facebook AI Research

. . . . . FET - Fetal Blood Cells

. . . . . GANs - Generative Adversarial Networks

. . . . . GIA - Trophoblast Giant Cells

. . . . . GlyT - Glycogen Trophoblast

. . . . . GLY - Glycogen Cells

. . . . . GPUs - Graphics Processing Units

. . . . . GMM - Gaussian Mixture Models

. . . . . GWAS - Genome-Wide Association Studies

. . . . . H & E - Hematoxylin and Eosin

. . . . . H0 - Null Hypothesis

. . . . . H1 - Alternative Hypothesis

. . . . . HDBSCAN - Hierarchical Density-Based Spatial Clustering of Applications with Noise

. . . . . HED - Hematoxylin and Eosin and DAB color spaces (assumed)

. . . . . HEM - Placental Phenotypes in Mutant Strains

. . . . . Het - Heterozygous

. . . . . HREM - High-Resolution Episcopic Microscopy

. . . . . ICM - Inner Cell Mass

. . . . . IKMC - International Knockout Mouse Consortium

. . . . . iEVT - Interstitial Extravillous Trophoblast

. . . . . IV - In Vitro

. . . . . KL - Kullback-Leibler

. . . . . LAB - Labyrinth Cells

. . . . . MSE - Mean Squared Error

. . . . . MAE - Mean Absolute Error

. . . . . mAP - Mean Average Precision

. . . . . MoCo - Momentum Contrast

. . . . . MRI - Magnetic Resonance Imaging

. . . . . NMF - Non-negative Matrix Factorisation

. . . . . P-TGCs - Parietal Trophoblast Giant Cells

. . . . . PFA - Paraformaldehyde

. . . . . PILLOW - Python Imaging Library (PIL) Fork

. . . . . P14 - Postnatal Day 14

. . . . . RAM - Random Access Memory

. . . . . ReLU - Rectified Linear Unit

. . . . . ResNet - Residual Network

. . . . . RGB - Red Green Blue

. . . . . RNA - Ribonucleic Acid

. . . . . ROI - Region of Interest

. . . . . S-TGC - Sinusoidal Trophoblast Giant Cells

. . . . . SPO - Spongiotrophoblast Cells

. . . . . SKimage - Scikit-image

. . . . . SIFT - Scale-Invariant Feature Transform

. . . . . SimCLR - Simultaneous Contrastive Learning of Representations

. . . . . SNPs - Single Nucleotide Polymorphisms

. . . . . SpA-TGC - Spiral Artery-Associated Trophoblast Giant Cells

. . . . . SpT - Spongiotrophoblast

. . . . . SSL - Self-Supervised Learning

. . . . . STB - Syncytiotrophoblasts

. . . . . SVM - Support Vector Machines

. . . . . SVD - Singular Value Decomposition

. . . . . SYN - Syncytiotrophoblast

. . . . . t-SNE - t-Distributed Stochastic Neighbor Embedding

. . . . . TTTS - Twin-to-Twin Transfusion Syndrome

. . . . . UMAP - Uniform Manifold Approximation and Projection

. . . . . VGG - Visual Geometry Group

. . . . . VOC - Visual Object Classes

. . . . . WSS - Within-Cluster Sum of Squares

. . . . . WT - Wild Type

. . . . . WSIs - Whole Slide Images

. . . . . XGBoost - Extreme Gradient Boosting

. . . . . YOLO - You Only Look Once

-+

*It makes sense that the placenta almost looks like a tree with many branches - a tree of life.*

# 1 Introduction

## Contents



## 1.1 Motivation

The placenta is a biological marvel vital in maternal and foetal health, facilitating nutrient transport, waste removal, and immunological shielding. Although it plays a

crucial role, the importance of the placenta has been overlooked in scientific literature. This thesis seeks to correct this neglect. The implications of the placenta's many functions extend far beyond academic discourse as they directly relate to severe conditions such as preeclampsia, gestational diabetes, and intrauterine growth restriction. These conditions pose immediate health risks and have far-reaching maternal and foetal health consequences. Thus, the pertinence of placental research transcends academic boundaries, emerging as a global public health imperative.

Our understanding of placental structure and function has traditionally relied on manual histological analyses. However, these conventional techniques have some limitations. For instance, sample preparation requires much manual labour, and the analysis is subjective. Additionally, the analysis tasks are time-consuming. In particular, the subjectivity inherent in these traditional methods imposes limitations on the speed and scale at which such analyses can be conducted which is crucial in clinical settings where rapid diagnosis and treatment are often imperative for patient well-being.

Traditional histological methods have limitations, so more efficient and objective analytical methodologies are needed. Machine learning and computer vision offer a promising solution. Such methods are capable of discerning patterns and structures in histological images that may remain imperceptible to human evaluators. These techniques can significantly augment histological analysis's accuracy and efficiency, thereby addressing some of the most pressing limitations of traditional methods.

The main focus of this thesis is to find a solution to this question: "How can computational techniques, particularly those based on machine learning and computer vision, be used to improve the precision, effectiveness, and comprehensiveness of analysing placental histology?" This inquiry lies at the intersection of developmental biology and computational science and will be the driving force behind the research goals of this thesis.

Solving this issue has the potential to significantly impact the way placental research is conducted. This change is not only necessary for academic purposes but also has implications for the health of mothers and babies, as placental diseases

are common worldwide. As a result, the thesis seeks to combine technological advancements with biological knowledge, creating a new approach for research and clinical use.

Contemporary epidemiological research has underscored an escalating trend in the prevalence of placental diseases, notably preeclampsia and intrauterine growth restriction. These conditions collectively afflict approximately 5–8% of all pregnancies globally. This statistic poses daunting challenges for public health infrastructure, especially considering the associated risks for maternal and foetal morbidity and mortality.

Historically, the placenta has been studied using histopathological examinations, clinical assays, and observational studies. While these methods have been valuable, they often need to provide a complete understanding of the complex biochemical and physiological processes involved in placental pathologies. Additionally, these traditional approaches can be time- and cost-intensive and rely on the subjective opinions of medical professionals, calling for more objective and scalable alternatives.

Using computational techniques in placental research is a promising area for new methods. Initial studies, including one by Ferlanio et al. [1], have used machine learning algorithms to analyse the placenta. However, these early efforts have faced challenges including those relating to scalability and accuracy. Additional research is needed to overcome these limitations.

To this end, an imperative exists for further development and refinement of computational technique. Machine learning and computer vision technologies offer substantial promise for advancing the field of placental research. However, these techniques have yet to be adequately tested on large, diverse datasets or optimised for performance across different placental tissues and conditions. Moreover, the computational requirements for implementing these complex algorithms can be prohibitively expensive, limiting their applicability in broader research and clinical settings.

## 1.2 Thesis Aims

### 1.2.1 Comprehensive Annotation and Preprocessing of the DMDD Placental Dataset

The initial objective of this thesis is to curate and structure the DMDD placental dataset for computational examination. This undertaking encompasses more than a cursory assessment, requiring a data cleansing procedure, annotation, and standardisation procedure.

### 1.2.2 Development of an End-to-End Efficient Phenotyping Pipeline for Placental Histology

The second aim centres on crafting an end-to-end automated phenotyping pipeline explicitly tailored to the requirements of placental histological analysis. This pipeline will incorporate state-of-the-art machine learning algorithms, addressing the limitations of manual analysis methods and efficient data processing. The objective is to produce a robust, scalable, and accurate system for identifying and classifying various placental features.

### 1.2.3 Comparative Evaluation of the DMDD Phenotyping Pipeline with the Hemberger Lab's Analyses

To achieve the third goal, I will comprehensively evaluate the DMDD phenotyping pipeline by comparing its performance metrics with those established by the Hemberger Lab. This assessment will help confirm the accuracy and dependability of the newly created pipeline and highlight any possible improvements that can be made.

### 1.2.4 Integrative Analysis of Developmental Genes and Correlation with Embryo Phenotypes

The fourth objective goes beyond just describing the physical traits of the placenta, instead aiming to uncover the relationship between developmental genes and the observed placental structures. Furthermore, this aim investigates possible connections between the placental and embryonic characteristics found in the

DMDD High-Resolution Episcopic Microscopy (HREM) embryo dataset. By taking this approach, I hope to enhance our knowledge of the genetic basis for placental development.

### 1.2.5 Advanced Exploration of Intraclass Variance Using Self-Supervised Learning Techniques

The final goal is to use unsupervised and self-supervised learning techniques to examine the differences in placental cellular morphology within the DMDD dataset. I aim not only to utilise these advanced methods but also to thoroughly evaluate their effectiveness in advancing the field of placental histology. This objective will explore the potential of self-supervised learning techniques to reveal new cellular morphology phenotypes.

## 1.3 Original Contribution

### 1.3.1 The DMDD Dataset

The contribution of this chapter resides in the annotation and preparation of the DMDD placental dataset, poised for computational analysis.

In this chapter, I have curated a dataset related to the placenta. Along with providing detailed annotations and labels for placental sections, I have also created training data for mouse nuclei and placental cells. The approach involves a high level of organsation and structure to ensure the dataset is practical and easily accessible to the broader research community. This will enable collaborative scholarly advancements in this field.

### 1.3.2 Development of an Automated Placental Phenotyping Method

The original contribution of this research chapter is the development of an automated placental phenotyping pipeline. This innovative procedure is a culmination of refined methodologies and advanced implementations, markedly enhancing the analysis of placental histology.

The pipeline builds on existing human models, particularly those developed by Ferlaino et al. The enhancements and refinements in this research address the inherent limitations of previous works and elevate the precision and reliability of placental analyses. The research delves deep into the intricacies of placental studies, systematically addressing the nuances and augmenting the methodologies to render them more versatile and universally applicable. This addresses the important issues of generalisability and efficiency inherent in the analytical pipeline of whole slide images.

This research chapter presents innovative mouse placental cellular localisation and classification models that significantly enhance the precision and breadth of mouse placental histology analysis. These models provide unprecedented detailed morphological insights, capturing the cellular composition of mouse placenta with high accuracy and generalizability. This advancement enables rapid computational analysis of mouse placental histology, yielding new and valuable insights.

The research chapter also stands out for its unique approach to analyzing many other placental data. I created a processing pipeline that efficiently manages the extensive histological data. This pipeline balances computational efficiency with analytical depth, providing time-saving solutions for large-scale analyses. Overall, it offers practical and effective management of the data.

### 1.3.3 Automated Analysis of the DMDD Placental Data

This research chapter centres on analyzing placental data from the DMDD. The goal is to replicate previous findings and expand our understanding of the placenta. One new contribution is a tissue segmentation approach that uses a modified concave hull algorithm. This approach is able to segment tissue regions, allowing for a more detailed examination of specific placental regions.

The automated method developed in the research chapter *Development of an Automated Placental Phenotyping Method* is applied to analyze histological placental data in the DMDD dataset. This method improves manual scoring processes, providing more accurate and reliable measurements that can be replicated.

The method allows for a more extensive exploration and quantification of different placental structures and features, which can provide valuable insights into phenotypic characteristics that were previously unknown.

This chapter describes the wild-type placenta, which will be a fundamental reference point for upcoming genetic research. This characterisation will enable future researchers to expand on this work and establish a standard for identifying any abnormalities in the placenta.

Moreover, I delve into the topic of sexual dimorphism in placental samples. This exploration helps to further research on the existing literature regarding sexual dimorphism and emphasises the significance of considering sex as a variable in placental research.

The chapter also has a comprehensive analysis pipeline for a detailed comparison of Wildtype (WT) and Knockout ()KO) placenta. It corroborates and extends the scope of existing research, revealing distortions and abnormalities in KO phenotypes and offering insights into new specific targets for further research. The in-depth phenotypic analysis and the extensiveness of figures and visuals provided showcase the sensitivity of the developed pipeline, contributing to the advancement of the understanding of placental diseases.

Furthermore, exploring the predictive capability of placental phenotypes in determining embryonic outcomes is innovative in research direction. Implementing innovative methodologies and feature analysis enhances the findings, providing insights into understanding congenital diseases and placental phenotypes and improving the accuracy of embryonic outcomes predictions.

This research chapter substantially contributes to the field through its multi-faceted and in-depth investigations, innovative methodologies, and detailed analyses. It advances the current knowledge base and paves the way for future explorations and innovations in placental studies.

### 1.3.4 Development of Unsupervised Approach for Detecting Variation in Cell Morphology

In this chapter, I explore the topic of cellular heterogeneity in placental cells and examine the use of automated methods for detecting it. I evaluate various computational methods and machine learning algorithms, focusing on unsupervised and self-supervised learning techniques, to study the morphological variations of placental cells. My approach is comprehensive, trialling a mix of supervised, unsupervised, and self-supervised learning models to detect intra-class variance. This innovative methodology has the potential to provide new insights into the understanding of placental disease.

The chapter evaluates the practical utility of the developed learning pipeline in real-world biological and medical scenarios, thus bridging the gap between theoretical development and practical application. It applies established methodologies to the specific context of cell morphology and placental cell morphological variance identification, offering a foundational comparison and evaluation for subsequent experimental work.

The SimCLR framework has been adapted to biological data innovatively in this chapter. This adaptation demonstrates the commitment to using novel applications and rigorous methods. A preliminary study is achieved by incorporating a range of model training and validation. This analysis's empirical and analytical insights enhance existing knowledge and suggest ways to improve the model's capabilities.

The chapter's detailed exploration of cell morphology detection methods constitutes a preliminary study in cell morphology differentiation and computational analysis. It contributes to biological and computational domains, pushing the boundaries of knowledge in cellular heterogeneity and its implications in pregnancy-related conditions.

*"The placenta: a silent sentinel, an unheralded guardian in the symphony of creation. It is the unsung hero of pregnancy's grand narrative, quietly performing its vital duties. With a stoicism borne of purpose, it weaves the delicate tapestry of life within its labyrinthine folds. This humble organ, often overlooked, holds within it the profound secret of life, cradling the nascent spark of existence with a tenderness that speaks of the profound miracles of nature."*

# 2

# The Placenta: A Review

## Contents

## 2.1 Introduction

The placenta, distinguishable as a discoid structure (Figure) 2.1) that forms the interphase between the mother and the foetus during pregnancy, has, throughout history, been treated with varying degrees of significance and reverence in different societies [2]. In Western cultures, the placenta is often medically perceived as an 'afterbirth', with its postpartum role limited to pathological analysis before incineration. By contrast, in various other cultures, the placenta assumes a ritualistic prominence. It may be accorded ceremonial funeral rites or seen as an alter ego to the infant, exemplifying the myriad perceptions of its role and importance [3].

In addition to the mystical properties, the placenta has been a subject of scientific interest since ancient times. The earliest known writings about the placenta came from Diogenes of Apollonia around 480 BC, who believed that the organ was a source of foetal nutrition. Aristotle built on Diogenes' observations, and Roman physician Galen synthesised this knowledge, proposing that the embryo was nourished by direct connection with maternal blood through the umbilical cord. Galen's theory was the predominant belief for about 1500 years until the Scottish anatomists and physicians, William and John Hunter, began their studies in the 1700s.

William and John Hunter were pioneers in understanding the complex nature of placental circulation. They used innovative techniques, such as injecting wax and coloured dyes into uterine arteries, to visualise the separate maternal and foetal circulations.

Erasmus Darwin extended the Hunter brothers' work by emphasising the placenta's primary role in oxygenating foetal blood. However, his conjecture that the amniotic fluid was the source of foetal nutrition was later refuted [4].

Technological developments, such as microscopy and serial sectioning, aided further advancements in understanding the placenta. Edwin Klebs introduced the technique of embedding tissues in paraffin for later sectioning in 1869. This, combined with the invention of the automatic rotary microtome by Minot, capable of cutting precise ultra-thin tissue sections, enabled detailed microscopic examination of the placenta [5].

Otto Grosser significantly contributed to the understanding of placental anatomy. He proposed a classification system for placental structure based on the number and types of tissue layers separating the maternal and foetal blood, which is still used today.

Throughout the 20th century, advanced technologies such as electron microscopes helped deepen our understanding of placental cells [6]. Today, we understand the placenta as an organ with four primary roles: metabolism, transport of gases, endocrine function, and immune response [7].

The placenta, despite the brief scientific overview detailed above, remains, one of the most understudied organs. This chapter seeks to demystify the placenta by exploring its importance, investigating the various methods by which it can be studied, and discussing its connection to adverse birth outcomes and congenital disease. It also underscores the current knowledge gaps in our understanding of this critical organ [8].

The chapter begins with a definition of the placenta, followed by a discussion on the evolution of the placenta and an exploration of its structural diversity across mammalian species. An updated report on the formation, function, and composition of the human placenta sets the stage for a more in-depth discussion on placental disease, specifically focusing on the genetic origins of the disease. The chapter concludes with a discussion on the use of model organisms in advancing our understanding of the placenta, with particular emphasis on the formation, function, and cellular composition of the mouse placenta and its differences from humans.

## 2.2 What is the placenta?

Defining the placenta, in a biological sense, poses a significant challenge, primarily due to its wide-ranging functions, including, but not limited to, immune response, gas transport, endocrine function, and metabolism. Its role in coordinating and regulating the mother and foetus's development further underscores its complexity [2]. Throughout gestation, the placenta evolves, assuming new functions as necessary, and finally orchestrates its demise upon completing its role.

**Figure 2.1:** *Top-down view of a fully formed human placenta, showcasing its sprawling network of blood vessels, nutrient-rich tissue, and the intricate connection to the umbilical cord. (Figure taken from https://www.rcplondon.ac.uk/file/12492/download)*

In his seminal monograph, Mossman simplified the definition of the placenta, describing it as 'an apposition or fusion of the foetal membranes to the uterine mucosa for physiological exchange' [5]. Building on Mossman's explanation, we can infer that the placenta comprises two essential components—foetal and maternal—resulting in an intimate connection between two living organisms. This connection is vital since deviations in placental function can significantly impact the pregnancy outcome and the offspring's lifelong health [2]. Consequently, an improved understanding of the placenta is crucial for maternal and infant health and mitigating health service costs related to treating short-term and congenital conditions.

## 2.3 How did the placenta evolve in mammals, and why is it structurally so diverse?

From an evolutionary perspective, the placenta's primary function is to facilitate the birth of a neonate capable of appropriate growth into a viable adult without necessitating extensive postpartum maternal care. Given the diversity among placental mammals, it is unsurprising that the development and morphology of the placenta differ markedly across taxa. This disparity suggests that different species have evolved diverse solutions in response to unique challenges. Interestingly, similar challenges have often been addressed differently across various species, indicating a rich evolutionary tapestry [9].

Viviparity, or the retention and growth of the fertilised egg within the maternal body until the offspring is capable of independent existence, has evolved independently across numerous animal groups. Under viviparity, the retained fertilised egg must either subsist on its reserves, typically as a yolk sac, or procure resources from the mother. This situation led to the emergence of the placenta as the defining feature of most mammalian species [9].

Three mammalian lineages—monotremes, marsupials, and eutherians—persist today, with eutherians distinguished from the others by their complex, well-developed placentas. As placental mammals, humans fall within the Eutherian lineage.

The broad spectrum of species within placental mammals, subject to diverse selective pressures, results in a broad range of placental types and significant variation in neonatal development. For example, carnivorous offspring are born blind and effectively helpless, whereas giraffe calves and horse foals can move competently within hours of birth [10]. This diversity denotes the fascinating complexity and adaptive capacity of placental evolution.

## 2.4 Placental Types and Their Multifaceted Characteristics in Mammalian Reproduction

The placenta, as discussed above, is an organ of paramount importance in mammalian reproductive biology. It serves a multitude of vital functions, including nutrient transport, waste elimination, and gas exchange. Understanding the intricate morphological features of the placenta allows for better appreciation of reproductive strategies across species, maternal and foetal health, and evolutionary biology. The morphology of the placenta can be dissected into five features: its geometrical shape, the nature of its interface with maternal tissues, the complexity of fetomaternal interdigitation, the interrelations of fetomaternal blood flow, and the neonatal to placental weight ratio [11]. Whilst the first three features have been the subject of more detailed study, the latter two remain less explored.

### 2.4.1 Placental Shape: Anatomical Forms and Functional Imperatives

The shape of the placenta provides an initial framework for classification and has broader functional implications related to an animal's physiological needs and lifestyle adaptations. As illustrated in Figure 2.2i, the placental morphologies can be primarily categorised into four types:

- **Diffuse Placenta** Characterised by an even distribution of villi or folds across the entire placental surface, this type is most commonly observed in ungulates like pigs and horses. The extensive surface area serves specific functional imperatives, such as facilitating effective gas and nutrient exchange. Moreover, this type of placenta is usually thin, allowing for a more expansive surface area [12].

- **Cotyledonary or Placentomal Placenta** This form is marked by chorionic villi organised into distinct, tuft-like structures, known as 'cotyledons.' The number of these cotyledons can vary considerably, ranging from as few as 5 in deer to as many as 150 in giraffes. While the cotyledonary placenta is

prevalent in ruminants and is traditionally viewed as an adaptation for efficient nutrient uptake, recent studies suggest its role is complex and influenced by a range of factors including maternal nutrition, placental morphology, and fetal demands. These factors collectively modulate the placenta's capacity for nutrient transport, often invoking structural and functional adaptations to optimize fetal development under varying nutritional conditions.

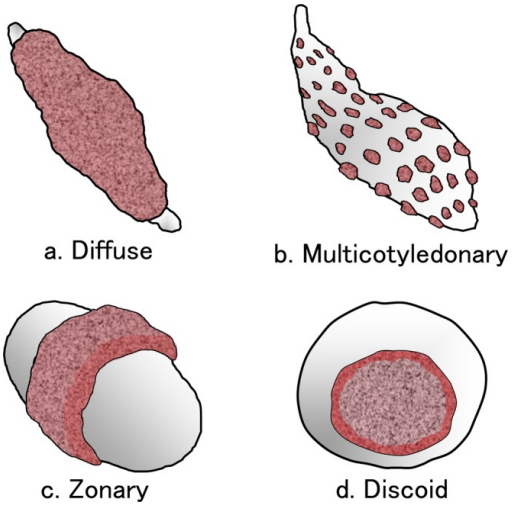
- **Zonary Placenta** In this configuration, complex folds are restricted to an equatorial band or pathway along the uterus, a feature commonly observed in carnivores. This arrangement provides a compromise between surface area for exchange and mechanical stability during gestation.

- **Discoid Placenta** Here, villi are confined to one or two disc-shaped regions. This is most frequently seen in rodents, insectivores, and anthropoids. The concentrated area allows for efficient material exchange but may also be associated with certain limitations, such as reduced surface area compared to other types [12].

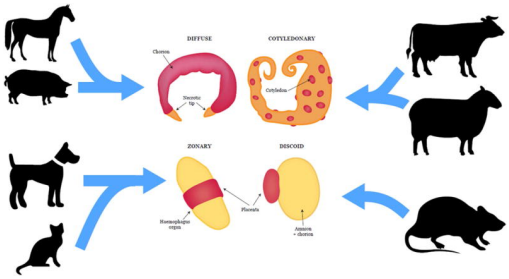
### 2.4.2 The Placental Interface

The placental interface, a term conceptualised by Grosser in 1909, signifies the critical zone where maternal and foetal tissues are in intimate contact, thereby serving as the point for physiological exchange. This interface undergoes dynamic changes throughout gestation, and these alterations have significant implications for both maternal and foetal health. As depicted in Figure 2.3, four primary interface types exist, each varying in the number of maternal tissue layers that are removed or retained during chorionic development:

- **Epitheliochorial Placenta** In this form, all maternal layers are retained, establishing direct contact between the uterine epithelium and the chorion. This is commonly seen in ungulates like pigs and horses. Although this multiple-layer structure can act as a barrier against pathogenic transmission, it may compromise the efficiency of nutrient and gas exchange.




**(i)** *Detailed comparative illustrations delineating the diverse morphological shapes of placentas across varying mammalian species, annotated for clarity.*




**(ii)** *A compendium of placental shapes, portrayed across a selection of mammalian species, emphasising their evolutionary divergence and the exceptions to the primary categories.*

**Figure 2.2:** *A comprehensive visual comparison of placental shapes across diverse mammalian species, annotated for enhanced understanding. (i) (a) Diffuse placenta, ubiquitously observed in ungulates, where the placenta is thinly distributed over an expansive uterine area. (b) Cotyledonary or 'multicotyledonary' placenta, prevalent in ruminants, characterised by distinct, tuft-like areas of chorionic villi, known as 'cotyledons'. (c) Zonary placenta, commonly observed in carnivores, distinguished by an equatorial band of complex folds. (d) Discoid placenta, a hallmark of rodents and anthropoids, where the placenta forms a discoidal structure. Figure adapted from Link and Link*

- **Syndesmochorial Placenta** Characterised by the removal of the uterine epithelium, this type allows for direct contact between maternal connective tissue and the chorion. Predominantly found in ruminants, this configuration tends to be more efficient for material exchange than the epitheliochorial type.
- **Endotheliochorial Placenta** In this category, both the maternal uterine epithelium and connective tissue are removed, leaving the maternal endothelial basement membrane to interface directly with the chorion. This type is typical in carnivores and permits an efficient transfer of nutrients and gases.
- **Haemochorial Placenta** This is the most invasive type, where all maternal tissue layers are removed to allow the chorion to directly contact circulating maternal blood. This form is most common in rodents, insectivores, and anthropoids, facilitating the most efficient form of material exchange [14].

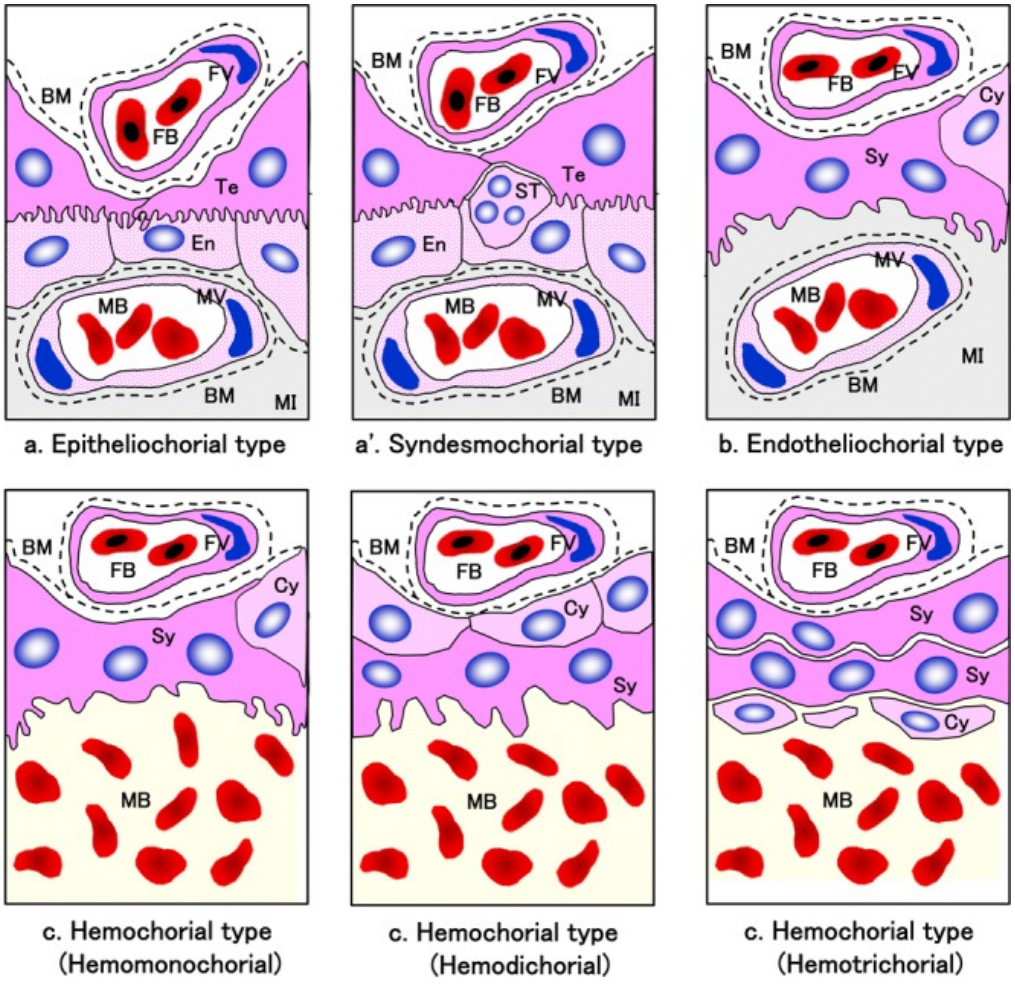



**Figure 2.3:** *Schematic representations elucidating the diverse types of placental interfaces, distinguished by their relationship between the chorion and the uterine wall, annotated for enhanced clarity. (a) Epitheliochorial placenta, a characteristic of ungulates, where the uterine epithelium remains intact and is in direct contact with the chorion. (a') Syndesmochorial placenta, typically observed in ruminants, where the uterine epithelium is eliminated, thereby placing the maternal connective tissue in direct contact with the chorion. (b) Endotheliochorial placenta, a feature of carnivores, where both the uterine epithelium and connective tissue are removed, enabling the maternal endothelial basement membrane to directly interface with the chorion. (c) Haemochorial placenta, a trait of rodents and anthropoids, where all maternal tissue layers are excised, rendering the chorion in direct contact with circulating maternal blood. Figure taken from - Link*

### 2.4.3 Foetomaternal Interdigitation

The degree of interdigitation between foetal and maternal tissues is a critical determinant of the available surface area for material exchange. Furthermore, the complexity of this interdigitation is often directly proportional to both the length and the surface area of the placenta, underscoring the importance of this feature in

placental functionality. Various types of interdigitation can be distinguished, each offering different levels of efficiency for nutrient and gas exchange:

- **Folded Placenta.** This type is marked by minimally branching, ridge-like folds of the chorion that fit into corresponding grooves in the uterine mucosa. Commonly found in pigs, this simple arrangement is effective for species with modest nutrient exchange requirements.

- **Lamellar Placenta.** Characterised by multiple branching ridges that form intricate systems of slender chorionic lamellae, oriented parallel to each other and separated by corresponding endometrial folds. This elaborate structure is typical in carnivores and is optimised for a more efficient exchange.

- **Trabecular Placenta.** This form features branching folds with leaf-like and finger-like villi branching off the main structure. Typically found in platyrrhine monkeys, this structure offers a moderately efficient surface for exchange.

- **Villous Placenta.** Distinguished by a tree-like branching pattern of the chorion, the villi of this type fit into corresponding endometrial crypts that are directly surrounded by maternal blood. This form, common in ruminants and higher primates, maximises the surface area, thereby facilitating highly efficient nutrient and gas exchange.

- **Labyrinthine Placenta.** Notable for its complex network of channels filled with maternal blood that surround the foetal capillaries, this is the most common form in rodents, lagomorphs, bats, and insectivores. The labyrinthine structure maximises the rate of material exchange between the maternal and foetal circulatory systems [13].

### 2.4.4 Foetomaternal Blood Flow and Neonatal/Placental Weight Ratio

The interrelationships of fetomaternal blood flow and the ratio of neonatal weight to placental weight although are also used to distinguish the placenta.

- **Fetomaternal Blood Flow Interrelationships:** The geometric architecture and spatial orientation of maternal and foetal blood vessels within the placenta have profound implications on the efficiency of material exchange. Specifically, the configuration of these vessels influences the direction and magnitude of blood flow, affecting both the oxygen and nutrient supply to the foetus. While this feature plays a vital role in placental function, it remains relatively understudied and warrants further academic scrutiny.

- **Neonatal/Placental Weight Ratio:** The neonatal/placental weight ratio provides a critical metric for assessing placental efficiency. This ratio reflects not only the gross morphological efficiency of the placenta in sustaining the fetus but also encapsulates finer aspects of placental function, such as the structural organization of the placenta and the density of nutrient and oxygen transporters that are vital for effective substrate exchange. These factors are excellent indicators of how well the placenta can support fetal growth. Research indicates that variations in this ratio may signal disruptions in placental blood flow or abnormalities in placental development, which are linked to fetal health outcomes. Hence, this ratio, beyond its simplicity, serves as a diagnostic tool that correlates closely with the physiological and pathological conditions of the placenta and foetus, providing insights into both current health and prognostic outcomes..

## 2.5 Formation and Maturation of the Human Placenta

The genesis of the human placenta is a fundamental process in pregnancy, commencing with the preimplantation embryo, known as the blastocyst. The blastocyst

differentiates into two distinct lineages: the inner cell mass (ICM), which will eventually form the embryo itself, and the outer trophoectoderm, which will develop into the placenta upon implantation into the maternal endometrium (around 6 - 7 days post-fertilisation).

The initial phase of placenta formation involves the active invasion of the trophoectoderm-derived syncytiotrophoblast through the epithelial surface and into the endometrium. This invasion process initiates significant tissue remodelling of the endometrium, transforming it into the decidua, a specialised uterine tissue that is indispensable for the successful maintenance of pregnancy and sustained embryonic growth.

The syncytiotrophoblast is an outer layer of fused cells (the syncytium) that directly interfaces with maternal blood. The cells of the underlying cytotrophoblast, which remains a layer of individual cells, undergo rapid proliferation, extending into the surrounding tissue to form primary villi. These villous structures consist of an outer layer of syncytiotrophoblasts and a core of cytotrophoblasts.

Consequently, the blastocyst is encased by three layers: the innermost chorionic plate, which remains in contact with the original blastocyst cavity; the developing villi that are separated by the intervillous space; and the outermost cytotrophoblast shell, which interfaces with the decidua. As the villous tree continues to expand rapidly, it forms foetal capillaries and initiates the creation of spaces for maternal-foetal exchange, marking the early formation of the placental structure [16].

The maturation of the placenta involves a series of intricate cellular processes. The intervillous space is structured and filled with maternal blood, while the villi spread out to increase the surface area available for exchange. In addition to the primary trophoblastic cell layers, the human placenta harbors a variety of other specialized cell types that perform crucial regulatory functions. These include Hofbauer cells, placental macrophages found within the chorionic villi, which are involved in immune regulation and cytokine production. Decidual cells, part of the maternal tissue interface, secrete essential growth factors such as Insulin-like Growth Factor and Vascular Endothelial Growth Factor, both critical for vascular

development and placental growth. Moreover, the placenta produces a wide array of hormones including human Chorionic Gonadotropin and Progesterone, which are pivotal for maintaining pregnancy and fostering a conducive environment for fetal development. These hormones and growth factors not only support the structural development of the foetus but also modulate the maternal physiological adaptations to pregnancy. By orchestrating these complex interactions, the placenta ensures the transfer of nutrients and gases is optimized to meet the growing demands of the fetus, thereby underlining the multifaceted nature of its endocrine and exocrine functions [17].

The uteroplacental unit, which is integral to pregnancy, comprises maternal-derived tissue from the endometrium and foetal tissue from the chorionic sac. The maternal aspect of the placenta is termed the basal plate, while the foetal aspect is known as the chorionic plate. This region accommodates the foetal chorionic blood vessels. The intervening space between these two regions, known as the intervillous space, houses the central functional units of the human placenta, the chorionic villi.

The chorionic villi, figure 2.4, are densely packed vascular projection structures of foetal tissue, enveloped by the chorion, and are in direct contact with maternal blood. The chorion comprises two cellular layers: the outer syncytiotrophoblast, which interfaces directly with maternal blood within the intervillous space, and the inner cytotrophoblast. The intervillous space itself is a large cavernous expanse into which the villi extend, thereby facilitating a highly efficient interface for maternal-foetal exchange.

This intricate and dynamic process of placenta formation and maturation underpins the successful outcome of pregnancy, providing the requisite interface for the exchange of nutrients, gases, and waste products between the maternal and foetal circulations. It also serves as a pivotal endocrine organ, producing a multitude of hormones and other signaling molecules that play essential roles in regulating maternal physiology, foetal development, and the timing of parturition. However, any disruption in these complex processes can lead to pregnancy complications

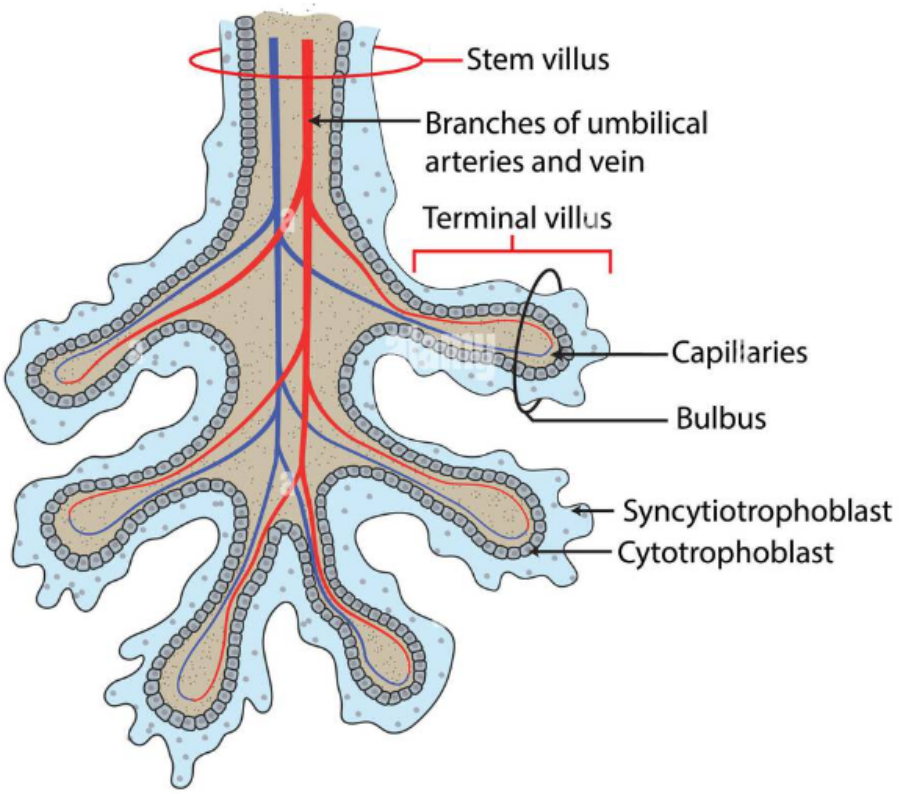



**Figure 2.4:** *Detailed illustration of the chorionic villus, a pivotal structure in placental function. The stem villus serves as the primary vascular conduit, connecting to branches of the umbilical arteries and vein. These vascular structures further extend into the terminal villus, where they form a network of capillaries. The capillaries are in intimate proximity to the bulbus, a specialised region facilitating gaseous and nutrient exchange. Two distinct layers of trophoblast cells envelop the villus: the outer syncytiotrophoblast and the inner cytotrophoblast. The syncytiotrophoblast is a multinucleated layer responsible for nutrient absorption and hormone secretion, whereas the cytotrophoblast acts as a reservoir for the syncytiotrophoblast. These layers play a critical role in mediating foetal-maternal interactions. Figure from - Link*

such as pre-eclampsia and interuterine growth restriction, highlighting the critical role of the placenta in human reproduction.

## 2.6 The Multifaceted Role of the Placenta

The placenta is an essential organ during gestation, fulfilling critical roles in metabolism, gas transport, endocrine production, and immune response. It nourishes the developing foetus with oxygen, water, carbohydrates, amino acids, lipids, vitamins, minerals, and other vital nutrients while eliminating carbon dioxide and waste materials. Moreover, it metabolises various substances, releasing their metabolic

products into maternal and foetal circulations. The placenta also safeguards the foetus against xenobiotic molecules, infections, and maternal diseases [2].

Additionally, it produces hormones that circulate in the maternal and foetal systems, influencing pregnancy, metabolism, foetal growth, parturition, and other physiological processes. Intriguingly, the placenta's functionalities alter during gestation and vary across species, although this discussion will exclusively focus on the human placenta [18].

The placenta's role in respiratory gas transportation is well-established. Oxygen rapidly diffuses from maternal to foetal blood within the intervillous space, while carbon dioxide transport occurs in the opposite direction. The diffusion rate across the placenta is influenced by several factors, including blood flow and the structural characteristics of the barrier itself. While blood flow to and from the exchange area is crucial for facilitating the transport of gases and nutrients, the thickness of the placental barrier also plays a vital role. A thinner barrier can enhance the efficiency of diffusion by reducing the distance over which substances must travel between maternal and fetal blood. This interaction between blood flow and barrier thickness ensures optimal conditions for the exchange of oxygen, carbon dioxide, and essential nutrients, accommodating the dynamic requirements of fetal development throughout gestation. Moreover, foetal haemoglobin exhibits a higher affinity for oxygen and a lower affinity for carbon dioxide than maternal haemoglobin, thus favouring oxygen transport to the foetus and carbon dioxide elimination to the mother [19].

The placenta metabolises diverse substrates to fuel foetal development. Placental metabolism is subject to species variation and adjusts to the escalating nutrient demands accompanying gestational progression. Its role in supplying water, carbohydrates, amino acids, lipids, vitamins, minerals, and other nutrients is facilitated by several cellular actions, which actively transport these substrates [20].

The placenta functions as an endocrine organ, playing a pivotal role in gestation success. Advances in technologies such as real-time PCR, immunohistochemistry, and mass spectrometry have significantly expanded our understanding of placental

endocrinology. Real-time PCR quantifies specific mRNA expressions, providing insights into the hormonal responses under various gestational conditions. Immunohistochemistry has been instrumental in mapping the spatial distribution of hormone receptors, offering clues about their roles in placental and fetal development. Mass spectrometry aids in the detailed identification and quantification of hormonal proteins and their metabolites, offering precise biochemical insights. Together, these technologies not only enhance our understanding of the placental functions but also contribute to the development of targeted therapies for managing pregnancy complications. Despite their profound impact, these methods come with challenges such as high costs and the need for specialized technical expertise, which may limit their widespread use in routine prenatal testing

The placenta forms an anatomical barrier protecting the developing foetus during pregnancy. However, it remains susceptible to breaches by infectious agents, potentially leading to pathogenic transmission from mother to foetus. The immune responses initiated by maternal and foetal cells at the decidual-placental interface play a crucial role in mitigating such threats [21].

## 2.7 Cellular Components of the Placenta

The placenta is a complex organ that has various cell types, each with specific functions (Figure 2.5).

Trophoblast cells primarily fulfil the placenta's fundamental functions. These cells were named by Ambrosius Arnold Willem Hubrecht in 1889 to describe cells that transport nutrients and create a protective barrier between the mother and the foetus. Hubrecht noted the highly invasive and corrosive nature of the trophoblast, which can be categorised into subtypes, including the syncytiotrophoblast, the cytotrophoblast, and the extravillous trophoblast [22] [23].

Syncytiotrophoblasts are the outer layer of cells covering the placental villi. They are involved in transferring nutrients and oxygen between the mother and foetus. These cells also produce hormones like Human chorionic gonadotropin

(hCG), Oestrogens, and Progesterone, which are important for supporting and regulate placental and foetal development [24].

Cytotrophoblasts play a critical role in modifying the structure of uterine blood vessels to ensure adequate blood supply to the placenta, a process essential for the proper development and function of the placenta throughout pregnancy. These cells migrate from the placenta into the maternal uterine tissue, specifically targeting the spiral arteries. Upon reaching these arteries, cytotrophoblasts adopt an endothelial phenotype, allowing them to integrate into the vessel walls.

This integration and subsequent transformation of the spiral arteries are crucial for reducing vascular resistance and increasing blood flow to the placenta. The cytotrophoblasts replace the muscular and elastic layers of the maternal spiral arteries with fibrinoid material, which dilates the arteries significantly. This dilation ensures that more maternal blood can reach the intervillous space of the placenta, facilitating enhanced exchange of oxygen and nutrients critical for fetal development. This remodeling process is also associated with a reduction in blood flow turbulence at the entry points, minimizing the risk of damage to the delicate placental tissue.

Failure in this vascular remodeling process, often due to inadequate cytotrophoblast function or integration, is linked to several pregnancy complications, including preeclampsia. In cases of preeclampsia, the incomplete transformation of the spiral arteries can lead to reduced placental perfusion, hypoxia, and subsequent fetal growth restrictions, illustrating the importance of cytotrophoblast-induced vascular changes in maintaining pregnancy health.[25].

Hofbauer cells, a category of human placental cells discovered by Hofbauer in 1903, serve as placental macrophages residing within the placental villous tree's connective tissue. These antigen-presenting cells are instrumental in maintaining host defence and also play an endocrine role, producing several hormones. Despite their relevance, Hofbauer cells have yet to be extensively studied, although their involvement in placental disease has been cited in several instances [26].

The human placenta comprises fourteen cell types and subtypes, each executing specific functions. A comprehensive discussion on each is beyond the scope of this

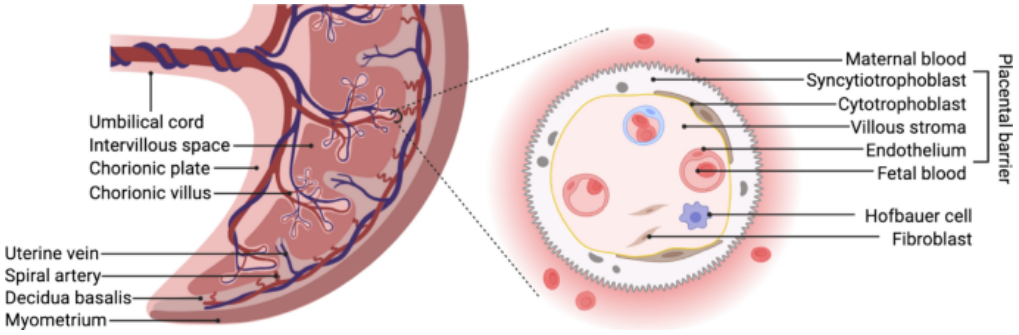



**Figure 2.5:** *Schematic representation of the human placental structure in term pregnancy. The placenta is anchored to the chorionic plate, which serves as the origin for both the umbilical cord and the chorionic villi. The intervillous space is filled with maternal blood, which enters via the remodeled maternal spiral arteries and leaves via the uterine veins. The first layer of the placental structure is the syncytiotrophoblast, which is formed by the fusion of cytotrophoblasts. While a complete layer of cytotrophoblasts is present in early pregnancy, these cells become sparse by term. Underneath the syncytiotrophoblast lies the villous stroma, which is composed of fibroblasts and placental macrophages, also known as Hofbauer cells. Within the villous stroma are foetal microvessels, which are lined with endothelial cells.Figure from Label*

thesis. However, it is indisputable that placental cells perform a broad spectrum of functions, with disturbances leading to placental diseases [27].

## 2.8 Placental Diseases and their Implications

Adverse birth outcomes pose significant financial burdens on healthcare systems and emotional challenges for affected families. The placenta, the most significant data source on the intrauterine environment and maternal-foetal health, holds considerable potential for diagnosing the origins of these adverse outcomes [28].

The most common adverse birth outcomes, including intrauterine growth restriction (IUGR), preterm birth, miscarriage, and preeclampsia, often originate in the placenta. IUGR, which impacts 15% of all births in the United States and Europe, is defined as the inability of the foetus to reach its genetically determined potential size. This condition is associated with impaired placental development, morphology, and structure, which hinder the organ's nutrient transport role [29]. Multiple factors contribute to IUGR, including uteroplacental blood flow, nutrient concentration gradients, the thickness of the exchange area, and metabolism [30] [31].

Intrauterine growth restriction (IUGR) poses immediate threats with high prenatal mortality rates and potentially lifelong consequences, such as elevated risks of developing metabolic and cardiovascular diseases later in life. For instance, in the Helsinki cohort – a large-scale study involving two groups of individuals born in Helsinki between 1924 and 1944, with extensive data collection from birth records, school checks, and health assessments – low placental and birth weights, as well as abnormal shapes, such as being more round, were correlated with hypertension and coronary disease in adulthood [32].

Preeclampsia, another significant pregnancy-related disease affecting 4-5% of pregnancies worldwide, imposes substantial burdens on maternal and foetal morbidity and mortality, contributing significantly to prematurity and maternal cardiovascular disease. This condition is characterised by new-onset hypertension, proteinuria, or other end-organ damage after 20 weeks of gestation. Eclampsia, the progression of preeclampsia, is defined by the development of tonic clonic seizures in a woman with preeclampsia. The pathogenesis of preeclampsia remains unclear, with various genetic, angiogenic, structural, and metabolic pathways implicated, including spiral artery remodelling, placental oxygenation, redox, immune tolerance at the maternal-foetal interface, and the balance of angiogenic and antiangiogenic factors. Delivery remains the best treatment option [33].

Impaired placental function significantly affects neurodevelopmental outcomes. Adverse neurodevelopmental outcomes in infants, including neonatal encephalopathy, bronchopulmonary dysplasia, and autism spectrum disorders, have all been linked to detecting placental lesions. It is increasingly clear that understanding the placenta and using it as a diagnostic tool is vital for the immediate outcome of the pregnancy and the long-term health of both mother and infant.

## 2.9 Genetic Origins of Placental Disease

The complex aetiology of placental diseases often involves a blend of genetic predispositions and environmental factors. This thesis, however, narrows its focus to the genetic underpinnings of these conditions. Advancements in genomic medicine,

exemplified by monumental projects like the Human Genome Project and Genome-Wide Association Studies (GWAS), have given unparalleled insights into the genetic architecture of a multitude of diseases, including those affecting the placenta [34].

Contrary to the widespread belief that genetic abnormalities in the placenta merely serve as a reflection of broader embryonic dysfunction, emerging evidence demonstrates these anomalies as independent drivers of disease. For instance, mouse models have shown that even embryonically lethal mutations can be mitigated by a genetically normal (wild-type) placenta, highlighting the autonomous influence of placental genetics [35].

Placental abruption, although often attributed to extrinsic factors like trauma or hypertension, has also been linked to genetic predispositions revealed through GWAS [36]. Similarly, preeclampsia has been found to have a genetic basis, particularly in familial aggregation studies [37]. Conditions like Intrauterine Growth Restriction, often a consequence of placental insufficiency, have been associated with specific genetic mutations affecting placental growth and function.

The scope of genetic influence also extends to gestational trophoblastic diseases, which range from benign hydatidiform moles to malignant gestational trophoblastic neoplasia (GTN) [38]. Even in recurrent pregnancy loss, which is often considered multifactorial, specific genetic aberrations affecting placental implantation and growth have been implicated [39]. Chorioamnionitis, traditionally viewed as an infectious disease, has also seen recent research pointing to genetic factors modulating susceptibility [40].

While Twin-to-Twin Transfusion Syndrome (TTTS) in monochorionic twin pregnancies is not directly caused by genetic factors, the underlying genetic setup does predispose to this condition [41]. The same can be said for Placenta Accreta Spectrum Disorders, where the placenta abnormally invades the uterine wall; although often linked to prior uterine surgery, genetic predispositions are currently under investigation [42].

Even chromosomal disorders like Trisomy 13, 18, and 21 have repercussions on placental morphology and function, leading to a host of complications [43].

Despite the enormous strides in understanding the genetic basis of placental diseases through advanced genomic techniques, much remains to be elucidated. Consultation with healthcare providers for accurate diagnosis and tailored management of these conditions remains of paramount importance.

The intricate landscape of placental diseases is increasingly being understood through the lens of genetics, thanks to advanced genomic techniques. However, the terrain remains largely uncharted, necessitating continued research to fully unravel the genetic intricacies underlying these conditions. Therefore, the role of model organisms becomes indispensable, especially as they provide an ethical and practical framework for investigating the genetic basis of diseases, including those of the placenta. Such studies can serve as a cornerstone for future therapies, risk assessments, and preventive measures, thereby making consultation with healthcare providers for precise diagnosis and individualised management ever more crucial.

## 2.10 Model Organisms for Understanding Gene Function

Model organisms which are non-human species utilised to study biological phenomena, offer invaluable insights into areas where human experimentation is unfeasible or unethical. Despite the Earth's immense biological diversity, many biological processes appear to be evolutionarily conserved.

Using animals in research is an ancient practice from ancient Greece and has been central to modern medical breakthroughs. Our understanding of heredity, development, physiology, and the underlying cellular and molecular processes owes much to studies with model organisms such as pea plants, Drosophila, yeast, zebrafish, and bacteria [44].

The laboratory mouse has been the premier model organism for the past century. Mice and humans share significant physiological and pathological traits, with many similarities evident in endocrine, immune, musculoskeletal, cardiovascular, and other internal organ systems. Factors such as cost-efficiency, ease of maintenance, straightforward breeding in captivity, and short lifespan (approximately two years

for a laboratory mouse), make mice ideal models for research. These advantages, coupled with the genomics revolution, have enriched the use of the mouse model, expanding our understanding of the relationship between genes and phenotypes [45].

Extensive tools for genetic manipulation, including techniques like CRISPR, have allowed for the engineering of genetic alterations such as knockouts, deletions, point mutations, and short insertions [46]. Knockout mice, wherein an artificial piece of DNA replaces an existing gene, can provide insights into gene function based on differences in the mice's behaviour or physiology. Large-scale application of this technique could offer a detailed understanding of the genome, as demonstrated by the International Mouse Phenotyping Consortium's work. They have established a high-throughput gene knockout and phenotyping pipeline to create a functional catalogue of the mammalian genome, systematically analysing every protein-coding gene [47].

## 2.11 Utilisation of Animal Models in Foetal Medicine

Research in human pregnancy is inherently fraught with ethical considerations, taking into account the wellbeing of both the mother and the unborn child [48]. As such, animal models have become invaluable tools for broadening our understanding of human gestation and the treatment of associated conditions. However, the extrapolation of findings from animal models to human pregnancy calls for careful consideration, particularly regarding three key aspects: variations in placentation, gestation periods, and litter sizes [49].

Gestation periods in animal models can be considered a double-edged sword. While brief gestation periods enhance experimental efficiency, extended gestations provide more opportunities to observe the effects of environmental and psychological influences. Furthermore, longer gestations allow for more complex foetal development and greater maternal demands.

In terms of litter size, animals with larger litters typically have smaller offspring, providing a unique comparative perspective. At the same time, larger litters produce higher sample sizes per gestation, enabling the generation of more robust datasets.

Finally, the variation in placentation across different species is a substantial consideration. The similarity of placental structure to humans is crucial for understanding drug transfer, biological mechanisms, and tracing variations in outcomes.

Various animal models, including sheep, guinea pigs, rats, mice, and non-human primates, have been utilised to enhance our understanding of human gestation. However, a perfect model for human gestation does not exist due to species-specific placental characteristics. As a result, researchers must select a model that most closely aligns with their research question. An basic overview of some models is given in Table 2.1

As disscussed in Section 2.10, murine models have garnered significant attention in pregnancy research, owing to their discoid-shaped placentas and haemochorial placental interface, which show noteworthy similarities to human placentation. Nevertheless, understanding the nuances that distinguish mouse and human placentas is crucial for the accurate extrapolation of experimental data. This section delves into a comparative analysis of mouse and human placentas, examining gestation length, litter size, and cellular and structural variations.

Mouse models offer several practical advantages, one of which is their short gestation period that typically lasts between 19-21 days. This abbreviated timeframe enables rapid data acquisition and the possibility of studying multiple generations within a short span [50]. However, it is imperative to note that mice produce altricial young, which are relatively underdeveloped at birth and continue to develop postnatally. In contrast, humans have a gestation period of approximately 259 days and give birth to more developed offspring [50]. This fundamental difference imposes limitations on the mouse model for studying the long-term effects and adaptations characteristic of human gestation.

Mice commonly give birth to litters comprising 6-8 pups, a feature that can be advantageous for data collection and for investigating developmental variations among littermates within the same intrauterine environment. However, this multi-foetal gestation fundamentally differs from the predominantly singleton pregnancies observed in humans. The multi-foetal environment in mice could potentially

| Model Organism | Placental Type | Advantages | Limitations |
|---|---|---|---|
| Mice | Haemochorial | Cost-effective, easy to handle. Genetically modifiable. Short gestation for rapid studies. | Differences from human placental structure and gestation. |
| Rats | Haemochorial | Larger size allows for easier surgical manipulations. Genetic and structural similarities to mice. | Different from human placenta. Increased costs compared to mice. |
| Rabbits | Haemochorial | Similar to humans in early gestation. Useful for studying early gestational events. | Higher costs, limited genetic tools, and complex husbandry. |
| Guinea Pigs | Haemomonochorial | Similar to human placenta. Used for oxygen transport studies. | Longer gestation, fewer offspring, and less convenient for high-throughput. |
| Sheep | Epitheliochorial | Large, accessible placenta for in utero studies. Suitable for foetal physiology. | Different from human placenta, high costs, and long gestation. |
| Pigs | Epitheliochorial | Physiological similarities to humans. Good for imaging studies. | Different from human placenta, high costs, and complex husbandry. |
| Non-Human Primates | Haemochorial | Closest to humans in structure and function. High relevance for translational studies. | Ethical concerns, extremely high costs, and complex husbandry. |
| In Vitro Models (IV) | N/A | Allows for controlled mechanistic studies. Options range from cell lines to organ-on-a-chip models. | Lack complexity of an entire organism. Limited in capturing in vivo function. |
| Computational Models | N/A | Useful for simulating placental function. Can be tailored for specific questions. | Require experimental validation. May oversimplify biological processes. |

**Table 2.1:** A comparison of model organisms commonly used in placental research, categorised by their placental type and focusing on the advantages and limitations of each. This table serves as guide for researchers, offering insights into the physiological relevance, ethical considerations, and practical constraints associated with each model. Abbreviations used are expanded for clarity.

introduce elements of competition for nutrients and affect placental development in ways rarely observed in human pregnancies.

Both mouse and human placentas exhibit a haemochorial placental interface, where syncytiotrophoblasts are in direct contact with maternal blood. This similarity extends to the development of highly branched villous structures, which arise through trophoblast differentiation and invasion. The foetal vasculature in both species originates from the embryo's allantoic mesoderm, while the maternal portion consists of maternal vasculature and uterine decidual cells. However, the extent to which these structures facilitate processes such as nutrient and gas exchange, as well as immune tolerance, may vary between the two species and warrants further investigation.

As illustrated in Figure 2.6, the tissue organisation in mouse and human placentas diverges significantly. The mouse placenta is composed of three distinct structural layers: the maternal decidua, the junctional zone, and the labyrinth. Each of these layers plays a unique role in supporting fetal development. The maternal decidua contains maternal endometrial cells that interface with the placenta. The junctional zone is crucial for hormone production and the regulation of nutrient transport to the fetus. Most importantly, the labyrinth layer, where the intensive exchange of gases and nutrients occurs, is characterized by its direct contact between maternal blood and fetal vessels. This is facilitated through a haemochorial barrier where a single layer of trophoblast cells—the syncytiotrophoblasts—lines the maternal blood spaces, allowing for efficient material transfer.

Fetomaternal indigitation refers to the interlocking structures formed by maternal and foetal tissues in the placenta, which are essential for efficient material exchange. In this context, the mouse placenta displays a labyrinthine structure, while the human counterpart exhibits a villous structure. Although both designs serve the purpose of facilitating material exchange between the mother and foetus, the structural differences could have implications for the efficiency and regulation of this exchange.

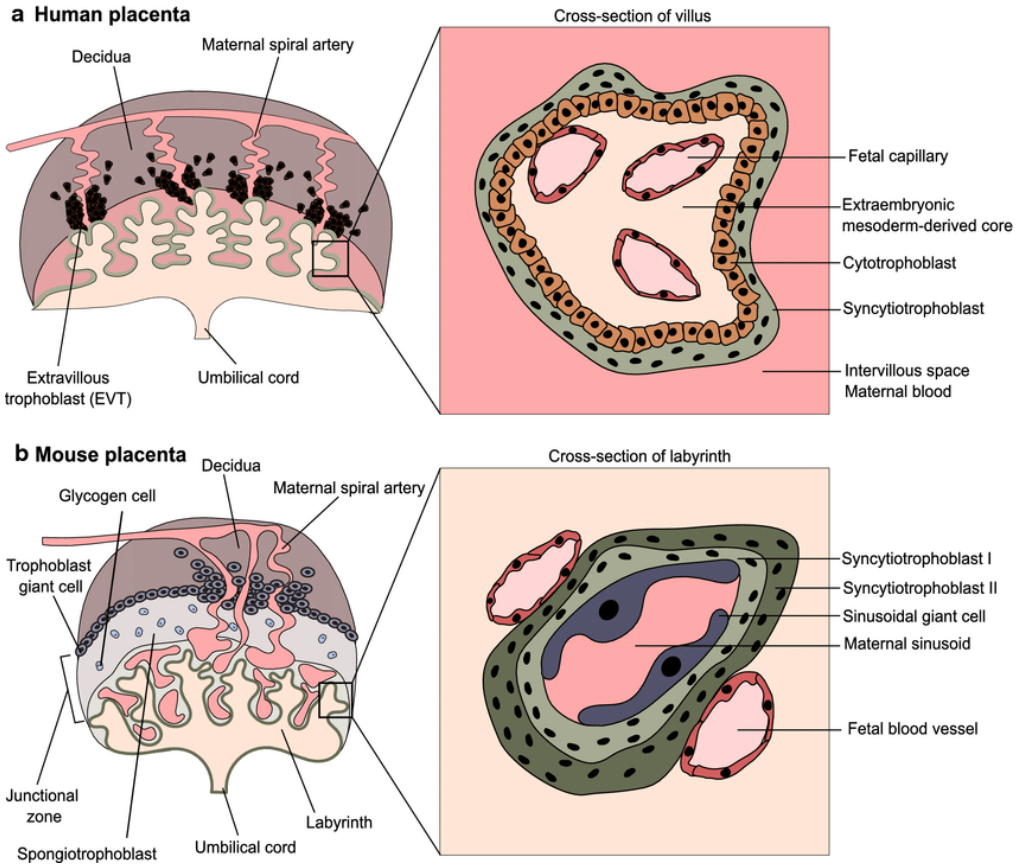



**Figure 2.6:** *(a) The principal unit of the human placenta is the villous tree. It encompasses an extraembryonic mesoderm-derived core (ExM-dC), which contains foetal capillaries. This ExM-dC is enveloped by a monolayer of cytotrophoblasts (CTB) and a multinucleated layer of syncytiotrophoblasts (STB), which is directly in contact with maternal blood. The inset presents a cross-sectional view through the villi, highlighting these layers. At the apex of the villi, cytotrophoblasts form a column. This column gives rise to two distinct types of extravillous trophoblasts: the interstitial extravillous trophoblast (iEVT), which invades the maternal decidua, and the endovascular extravillous trophoblast (eEVT), which invades the maternal spiral arteries, playing critical roles in ensuring adequate blood supply to the developing foetus. (b) The murine placenta is divided into two main zones: the junctional zone and the labyrinth zone. The junctional zone comprises spongiotrophoblasts and glycogen cells. Trophoblast giant cells, which invade the maternal decidua and remodel maternal arteries, are found in this zone and are considered to correspond to human EVT. The labyrinth zone is where the exchange of gases and nutrients occurs. Here, the foetal blood is separated from maternal blood sinusoids by foetal endothelial cells, two layers of STB (STB-I and STB-II), and sinusoidal giant cells. Functionally, the labyrinth is comparable to the human villi, serving as the site of material exchange between mother and foetus. Figure from Link*

| Feature/Aspect | Human Placenta | Mouse Placenta |
|---|---|---|
| Gestation Length | Approximately 259 days; gives rise to more developed, or precocial, offspring. Longer gestation allows for more complex developmental processes. | 19-21 days; produces altricial young, which are less developed and continue to mature postnatally. Shorter gestation facilitates rapid experimental turnover. |
| Litter Size | Typically one offspring per pregnancy (singleton). The intrauterine environment is geared towards the development of a single foetus. | Usually 6-8 pups per litter. The multi-foetal environment may lead to competition for nutrients and can influence placental development. |
| Placental Interface | Haemochorial; syncytiotrophoblasts come into direct contact with maternal blood. This allows for highly efficient nutrient and gas exchange. | Haemochorial; similar to humans, syncytiotrophoblasts are in direct contact with maternal blood. However, the efficiency and specifics of nutrient and gas exchange may vary. |
| Tissue Organisation | Single-layered (Haemomonochorial); consists mainly of syncytiotrophoblasts and cytotrophoblasts. Simpler structure but highly specialised for efficient material exchange. | Three-layered (Haemotrichorial); includes the maternal decidua layer, junctional zone, and labyrinth layer. Each layer consists of distinct cell types and structures, adding complexity to material exchange processes. |
| Nutrient and Waste Exchange | Highly efficient, facilitated by complex transporters and a rich vascular network. Designed for prolonged gestation and the needs of a single, precocial offspring. | Efficient but may involve different transporters and mechanisms. Adapted for shorter gestation and the needs of multiple, altricial offspring. |
| Environmental Adaptations | Limited data available, but potentially complex mechanisms for adapting to environmental stressors like nutrient deprivation or hypoxia. | More data available; commonly used for studying the impact of environmental factors on placental development and function. |
| Methodological Approaches | Limited by ethical considerations, usually involving non-invasive imaging and occasional biopsies. | Diverse methods possible due to fewer ethical constraints, allowing for more extensive histological and molecular analyses. |
| Temporal Dynamics | Complex and prolonged, allowing for intricate developmental processes and adaptations over time. | Simplified and rapid, making it easier to study multiple generations but may not capture all aspects of human placental dynamics. |

**Table 2.2:** A comprehensive comparison between human and mouse placentas, elaborating on gestation length, litter size, placental interface, tissue organisation, nutrient and waste exchange, environmental adaptations, methodological approaches, and temporal dynamics. Understanding these nuanced differences and similarities is vital for the judicious interpretation and application of research findings from mouse models to human pregnancy.

## 2.12 Cell Types in the Mature Mouse Placenta

The three distinct layers in the mature mouse placenta are constituted by various cell types and subtypes, which help distinguish the layers. In this context, the trophoblast is the most vital cell category, providing functional, endocrine, and structural properties. There are seven terminally differentiated trophoblast subtypes in mice, revealing the complexity of trophoblast cell subtypes.

Five terminally differentiated trophoblast subtypes in the mouse placenta are associated with maternal vasculature. They are classified based on location, developmental origin, and function [51]. Trophoblast giant cells are one of the most significant cells among these. Several subtypes include spiral artery-associated trophoblast giant cells (SpA-TGC), equivalent to extravillous cytotrophoblasts in humans [52]. Canal trophoblast giant cells are associated with the maternal canals that supply blood to the labyrinth layer, and sinusoidal trophoblast giant cells (S-TGC) occupy the maternal vascular space within the labyrinth. This unique attribute makes the mouse placenta haemochorial since foetally-derived trophoblast cells directly interact with maternal blood. Parietal trophoblast giant cells (P-TGCs) can be found at the junctional zone and decidua junction. These large cells establish the venous maternal blood spaces where the maternal vasculature exits the placenta. P-TGCs are distinguished by their large polytene nuclei and direct contact with maternal immune and decidual cells.

Other trophoblast-derived cells exist in the junctional zone, a largely avascular tissue region. These include glycogen trophoblast (GlyT) and spongiotrophoblast (SpT) cells. GlyT cells, which reside in clusters among spongiotrophoblast cells, are easily identifiable due to their large cytoplasm. As gestation progresses, GlyT cells accumulate glycogen and store it for later use in pregnancy. Meanwhile, SpT cells produce a wide range of hormones [53].

The labyrinth zone, the main site of nutrient and gas exchange in the mouse placenta, houses two syncytiotrophoblast cells (SynT layers I and II). The SynT-II cells are in direct contact with foetal endothelial cells and pericytes that line the foetal capillary network in the labyrinth. Syncytiotrophoblast cells in mice and

humans originate from the fusion of mononuclear progenitor cells. They express nutrient transporters on both apical and basal membranes, and the maternal side of the SynT cell in humans possesses protruding microvilli [54] [55].

Additional cell types in the mouse placenta include foetal blood cells, nucleated red blood cells vital for oxygen transport, and endothelial cells, essential for forming the vasculature in the placenta.

## 2.13 Evolution and Current Impact of Genetic Knockout Studies in Mouse Placental Research

The journey to understanding the intricacies of placental biology has been fascinating, enriched by numerous scientific discoveries and technological advancements. Notably, applying genetic knockout and transgenic technologies to mouse models has revolutionised foetal medicine. These techniques have afforded researchers an in-depth understanding of the complex interplay of genetic factors that govern placental function and foetal development.

The 1980s marked a significant era for biomedical research with the emergence of genetic manipulation techniques. Genetic engineering allowed researchers to modify the genome of organisms, effectively controlling gene expression and function.

In 1989 Martin Evans and colleagues generated the first "knockout" mouse. This pioneering work involved the inactivation or "knockout" of a specific gene to observe the effects of its absence on the organism's phenotype [56]. This research opened a new avenue for exploring the role of individual genes in intricate biological processes, such as those found in foetal development and placental function.

Rossant and colleagues were amongst the first to apply genetic knockout techniques to study placental biology in the 1990s. They investigated the role of the hprt gene in trophoblast development. Their study unveiled the critical impact of this gene on placental function, with the deficiency leading to abnormal placental development and subsequent foetal death [57].

This period also observed researchers combining genetic knockout techniques with the generation of transgenic mice, animals carrying artificially introduced genes. These models enabled more detailed investigations into the roles of specific genes in placental development and function. A seminal example of this research is the study conducted by Cross et al., which investigated the hand1 gene's role in trophoblast differentiation.

The dawn of the 21st century marked the advent of more advanced and precise genetic manipulation techniques in the realm of biomedical research, with the CRISPR/Cas9 system being the most notable. This groundbreaking technology has paved the way for researchers to conduct more efficacious gene editing, allowing for a meticulous investigation of gene function in placental development. For instance, Chuong et al. utilized the CRISPR/Cas9 system to investigate the role of endogenous retroviruses as species-specific enhancer elements in the placenta, which are pivotal for understanding the complexities of maternal-foetal nutrient exchange [58].

## 2.14 Current State and Future Prospects

Today, genetic knockout and transgenic mouse models remain pivotal in foetal medicine research. These models serve as invaluable platforms for mimicking human genetic disorders, elucidating the roles of specific genes in placental development and function, and piloting new therapeutic interventions for pregnancy complications.

Advancements in technology continue to enhance the complexity of these models, making them increasingly accurate representations of human physiology. Researchers are now equipped to explore more nuanced genetic interactions and their implications for placental function and pregnancy outcomes. This acquired knowledge is incrementally contributing to our understanding of both normal and pathological pregnancies, thereby positively impacting maternal and foetal health outcomes.

Looking forward, the ever-evolving arsenal of genetic manipulation techniques provides the potential for a future where we can manipulate the genetic landscape of organisms to answer increasingly complex research questions. In parallel with these genetic tools, emerging technologies such as artificial intelligence (AI), particularly

in the realm of computer vision, offer further opportunities to elevate our research capabilities. AI algorithms can process and analyse enormous quantities of complex, high-dimensional data at a speed and scale unattainable by human experts. This computational capability not only promises to accelerate the pace of discovery but may also reveal intricate patterns and insights that would otherwise remain elusive.

Specifically, in the domain of placental research, computer vision methods can revolutionise the way we interpret histopathological images. These techniques can be employed to rapidly and accurately analyse microscopic tissue sections, thereby aiding in the identification of subtle morphological changes associated with specific genetic alterations or pathological conditions. For example, automated image recognition algorithms can be trained to detect nuanced variations in cell structure, vascularisation, or the presence of abnormal tissue growths that may be indicative of certain placental diseases.

Furthermore, the integration of AI with cutting-edge genomic technologies, such as single-cell RNA sequencing, can offer unprecedented insights into the cellular composition and functional states of the placenta. An Advanced machine learning model for interpreting mouse placental histology would allow integration of large scale omics studies with observed phenotypes, to build models of development and disease

By amalgamating the capabilities of genetic manipulation and AI technologies, particularly computer vision, we anticipate a future where answers to complex biomedical questions can be uncovered at an unprecedented pace. These synergistic advances promise to deepen our understanding of the genetic basis of pregnancy disorders and to expedite the development of effective therapies for their amelioration.

*"The world is moving into a new age of AI, and as it does, we will see computer vision bring a new understanding to machines. No longer will machines only interact with the world through code and commands, but they will see the world as we do, in images and visualisations."*

— Demis Hassabis

# 3

# Computer Vision: A Review

## Contents

## 3.1 Introduction

The central research methodology employed in this thesis pivots around computer vision, an area of artificial intelligence (AI) that empowers computer systems to extract meaningful information from digital images, videos, and other visual inputs [59]. The idea of enabling machines to "see" and derive insights from visual data has intrigued scientists since the genesis of digital imagery.

The primary component of a digital image is the pixel, and in most computers, the RGB model orchestrates their assembly. This model employs three primary colour elements: red, green, and blue, each varying between 0 (absence of colour) and 255 (full saturation). The resultant array of colours used in image processing is stored as a three-dimensional matrix, with each layer symbolising the intensity of red, green, or blue pixels (Figure 3.1).

Starting as a niche scientific field in the 1950s, computer vision has proliferated dramatically since 2012. One of the significant driving factors behind its expansion is the evolution of deep learning methods, such as convolutional neural networks (CNNs), which have significantly revolutionised computer vision tasks such as image classification and object detection. This chapter thoroughly examines these developments and explores advances in unsupervised and self-supervised methods in computer vision. Subsequently, the application of these methods in medical science, particularly in histological analysis, will be discussed in detail.

## 3.2 The Historical Trajectory of Deep Learning and Computer Vision

### 3.2.1 The Pre-2010 Era of Computer Vision

Deep learning, a subset of machine learning that employs multilayered algorithms for data processing, has shared a co-evolutionary trajectory with computer vision since the 1950s. Moore's Law, initially posited by Gordon Moore in 1965, observed that

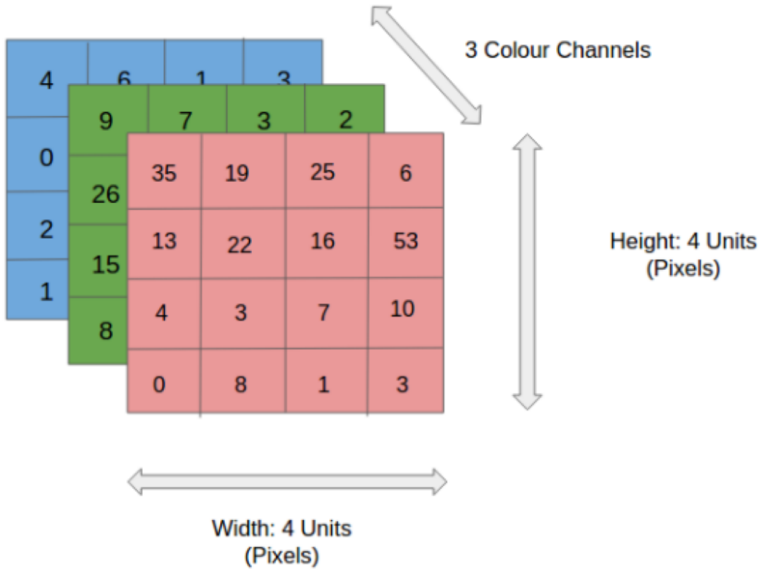



**Figure 3.1:** *Illustration of the RGB channels within a digital image. The image has been decomposed into its constitutive colour channels: Red (R), Green (G), and Blue (B). Each channel portrays intensity values of the respective colour within the image. The dimensions of the image are 4 pixels in width and 4 pixels in height, with each pixel housing a specific intensity value for each of the RGB channels. Figure from Label*

the number of transistors on a microchip doubles approximately every two years, leading to an exponential increase in computing power and efficiency. For many years, advancements in computing hardware as predicted by Moore's Law facilitated the incremental progress in machine learning and computer vision. However, it was not until the 2010s that deep learning techniques began to outpace the advancements predicted by Moore's Law, firmly establishing themselves as the gold standard in artificial intelligence.

Before the rise of deep learning and convolutional neural networks (CNNs) in the 2010s, computer vision research was diverse, exploring a range of methodologies. Early computational approaches in the field were heavily influenced by edge detection and geometric models, focusing on identifying points in a digital image where image brightness changes sharply. Another significant area was optical flow, concerned with determining the movement of objects between different frames of a video.

In the late 1990s and early 2000s, feature-based methods such as the Scale-Invariant Feature Transform (SIFT) by David Lowe in 1999 and the Viola-Jones

object detection framework became prominent. These methods focused on identifying specific features within an image for tasks like object recognition and robotic navigation.

The notion of artificial neural networks (ANNs) originated from attempts to replicate human brain function. Analogous to neurons firing in the human brain, ANNs involve signals passing through a system of connected nodes, triggering further nodes and propagating this process throughout the system (Figure 3.2). In 1943, McCulloch and Pitts [60] proposed a mathematical model that simplified this concept, and in 1958, Rosenblatt [61] expanded this model, developing the perceptron a fundamental unit of ANNs.

This artificial neuron could learn a function that maps input to output without explicitly implementing the function. This can be represented as shown in Figure 3.2ii. The neuron receives an input denoted as $x(i)$. The synapses harbour weight factors, represented as $w(i)$, which are applied to these inputs. Subsequently, the processed information traverses through the dendrites towards the cell body. At this juncture, a weighted sum of the inputs is calculated, and a bias term ($b$) is typically introduced. The activation function, originally a rudimentary step function in Rosenblatt's model, is applied to the resultant weighted sum to yield a binary output $\hat{y}$, which can then be compared with the ground truth label $y$.

The process of training this artificial neuron to accurately map inputs to outputs involves several key steps. These steps ensure that the neuron not only processes the inputs but also adjusts its internal parameters (weights and biases) to improve its prediction accuracy over time. The training involves the following steps:

1. Initialise the weight factors and biases with random numerical values. This random initialization kick-starts the learning process.

2. Compute the output $\hat{y}$ for each input sample $x(i)$. This involves the forward pass, where the neuron uses its current weights to predict an output.

3. For each sample, adapt the weights and biases in accordance with the correct label through the backpropagation algorithm. This step is crucial for learning,

as it involves adjusting the weights based on the error between the predicted output and the actual label.

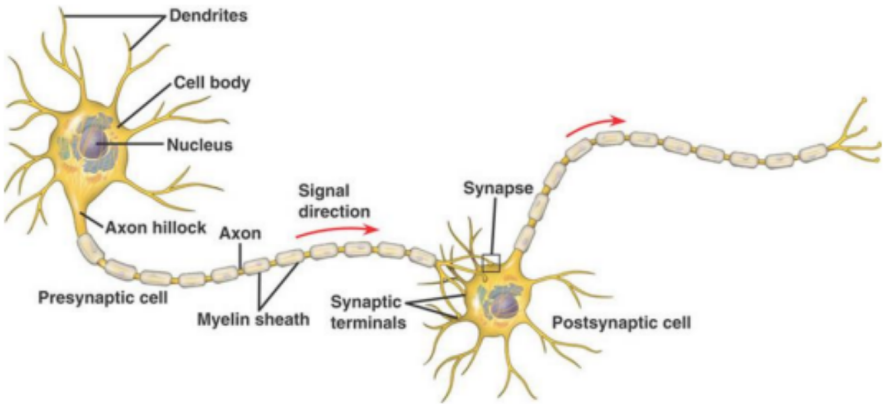
4. Iterate through Steps 2 and 3 until no errors are observed, or a predefined stopping criterion is met. This iterative process is essential for gradually refining the weights, leading to improved performance of the neuron in mapping inputs to their correct outputs.

Through these steps, the neuron evolves from a rudimentary function approximator to a more sophisticated model capable of making accurate predictions, exemplifying the fundamental mechanism of learning in artificial neural networks.

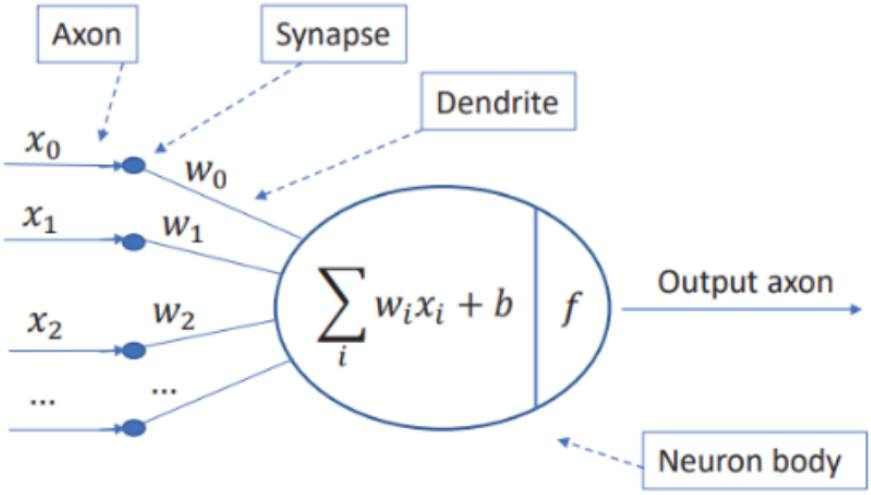
Hubel and Wiesel [62] conducted groundbreaking research into how visual processing occurs in the brain. They found that simple structures like edges always initiate visual processing, and more complex structures gradually build upon these initial signals. Their discovery laid the foundation for the layered approach adopted in deep learning.

In the same decade, Russell Kirsch and his team pioneered the concept of digital images by creating the first digital scan of an image, paving the way for computers to interpret and display images.

Henry J. Kelly [63] proposed the precursor to backpropagation in 1960, a learning algorithm for neural networks. Over the next few decades, this concept evolved significantly. Backpropagation, or "backward propagation of errors," is a method in machine learning for training artificial neural networks, crucial in optimising network weights through gradient descent. This process involves a forward pass where the network computes the output and the loss, which measures the discrepancy between the prediction and the actual result. Then, in the backward pass, the gradient of the loss function with respect to each weight. These gradients indicate the direction and magnitude to adjust the weights to minimise the loss. The weights are updated iteratively until the loss is minimised, allowing the model to learn from the data effectively. This methodology was brought into prominence when Rumelhart,




**(i)** Biological Neuron




**(ii)** Artificial Neuron

**Figure 3.2:** *(i) The diagram showcases a biological neuron, a complex cellular unit of the nervous system. It consists of dendrites (analogous to the weighted inputs $x_i$ in an artificial neuron) that receive incoming signals. These signals are then processed in the cell body which houses the nucleus. The axon hillock is where the decision to transmit an action potential (akin to the summation and activation function in an artificial neuron) is made. If the potential is deemed sufficient, it travels down the axon, a long structure enveloped by a myelin sheath, to the synaptic terminals, forming synapses with the postsynaptic cells to transfer the signal. (ii) The artificial neuron is a simplified computational model inspired by the functional properties of biological neurons. It processes multiple inputs $x_i$, each associated with a weight $w_i$. These weighted inputs are summed ($\sum_i w_i x_i + b$) within the neuron body, a process comparable to the integration of incoming signals within the cell body of a biological neuron. The sum is then passed through an activation function $f$, mimicking the generation of an action potential in biological neurons. The output is then sent to subsequent artificial neurons, reflecting the transmission of signals via synapses in biological systems. Despite its simplicity, the artificial neuron can mimic complex behaviours of biological neurons when used within larger networks. Figure adapted from Link*

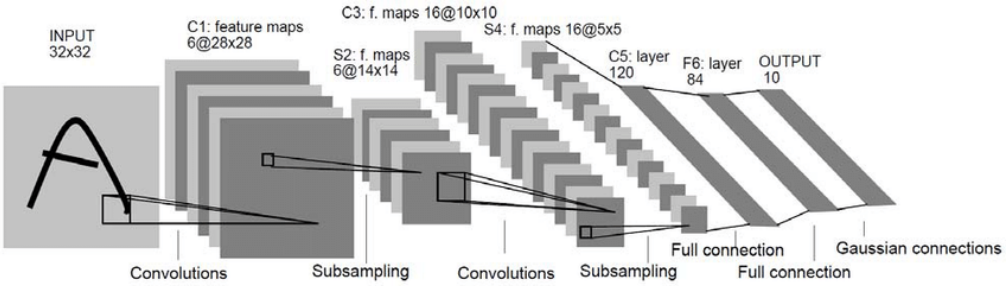



**Figure 3.3:** *Architecture of LeNet-5, a Convolutional Neural Network designed for digit recognition. The network comprises several layers each performing a specific function. The input layer accommodates a 32x32 pixel image. The first convolutional layer (C1) applies convolution operations to the input image producing 6 feature maps of size 28x28. These are then passed through the first subsampling (pooling) layer (S2), reducing the spatial dimensions to 14x14. The second convolutional layer (C3) further processes these pooled feature maps, generating 16 feature maps of size 10x10. The second subsampling layer (S4) applies pooling operations on these feature maps, resulting in 16 feature maps of size 5x5. A fully connected layer (C5) with 120 neurons then processes these pooled feature maps, followed by another fully connected layer (F6) with 84 neurons. Finally, a fully connected output layer with 10 neurons (corresponding to the 10 digit classes) produces the network's output. Each plane represents a feature map, i.e., a set of units whose weights are constrained to be identical. Figure from Link*

Williams, and Hinton demonstrated the utility of backpropagation in neural networks in 1985 [64], marking a significant advancement in the field of deep learning.

The summer of 1966 witnessed Seymour Papert [65] initiate an innovative project to solve machine vision problems. Although the initial expectations fell short, the endeavour underscored the complexity of the computer vision challenge. It marked a critical turning point, heralding the birth of computer vision as a scientific field.

In 1979, Kunihiko Fukushima [66] devised the first convolutional neural network (CNN), an architecture known as Neocognitron. This multilayered hierarchical artificial network incorporated several convolutional layers and weight vectors, also known as filters, and it could effectively recognise visual patterns. Later, in 1989, Yann LeCun [67] refined the CNN architecture, culminating in the development of LeNet-5 (Figure 3.3). This groundbreaking model combined convolutional networks with backpropagation to recognise handwritten numbers, thus providing a blueprint for modern CNNs [68].

The close of the 20th century saw significant advances in computational technology, particularly the development of graphic processing units (GPUs) that

accelerated image processing. With computational speeds surpassing Moore's law, neural networks began to perform more efficiently, marking the end of the so-called "AI winter." The term "AI winter" refers to periods in the history of artificial intelligence research where funding, interest, and optimism about the potential of AI significantly waned due to unmet expectations and technical challenges.

Meanwhile, the focus of computer vision research shifted towards feature-based object recognition. Lowe's seminal work on scale-invariant feature transform (SIFT) in 1999 [69] proposed a method for object recognition based on local features that were invariant to rotation, scale, and, to some extent, changes in illumination. As Lowe suggested, the SIFT features resemble the properties of neurons found in the inferior temporal cortex of primates which is the area involved in object recognition processes.

The advent of the 21st century brought an increasing demand for standardised performance measures in computer vision, leading to the development of benchmark datasets. The Pascal VOC [70] project, launched in 2006, served this purpose by offering a standardised dataset for object classification, complete with an annual competition to stimulate progress in the field. However, the 20-object categories needed to be more challenging for advanced models, leading to the creation of the more extensive ImageNet [71] dataset by Fei-Fei Li. ImageNet, containing 1000 object classes and 14 million images, set a more rigorous benchmark and has been instrumental in advancing the field of object detection and classification.

### 3.2.2 Deep Learning and Computer Vision in the 2010s

The 2010s marked a significant evolution in Deep Learning, particularly spurred by the substantial increase in GPU speeds. Graphics Processing Units (GPUs) facilitated more comprehensive Convolutional Neural Networks (CNNs) training from scratch. This section will chronicle the progression of CNNs in computer vision during this decade, categorically addressing two of the most influential computer vision tasks: image classification and object detection.

### CNNs for Image Classification

Image classification constitutes classifying an object within an image by a computer. For instance, if an image contains a cat, the computer will return a 'cat' label with an associated probability. The progression of image classification has been substantial since the 2010s, primarily attributed to the advancements and improvements of Convolutional Neural Networks (CNNs) for this purpose. The advancements brought by LeNet, AlexNet, VGG-16, Resnet, Inception, and MobileNet have had significant implications on image classification. This section provides an overview of the rapid advancements in CNNs within this field.

The first CNN applied for image classification was LeNet by LeCun et al. (1989) [67]. They pioneered the use of Stochastic Gradient Descent via backpropagation for training. While the relatively simple LeNet architecture, comprising two convolutional layers, a max pooling layer, and a fully connected layer, may not be suitable for contemporary complex tasks, it established the foundation for subsequent advancements in CNNs for image classification.

Despite the fundamental concepts for CNNs being formulated by Hinton, LeCun, and others during the 80s and 90s, it was in the 2010s, when computational power markedly improved, that the full potential of these networks could be realised. The first remarkable result from a CNN architecture emerged with the introduction of AlexNet in 2012 by Krizhevsky, Sutskever, and Hinton [72]. Winning the ImageNet competition, AlexNet, featuring five convolutional layers and a max pooling layer, recorded a top-5 error of 15.3%, surpassing the runner-up by more than 10.8 percentage points. This achievement spotlighted the importance of network depth in determining performance.

Subsequent research by Oxford University's Visual Geometry Group (VGG) delved into the impact of network depth, resulting in the development of the VGG-16 network [73]. The network consisted of 16 layers, with 12 being convolutional. The impressive performance of VGG-16, which achieved 92.7% accuracy across 1,000 classes in ImageNet, reinforced the notion that larger networks tend to yield better results. Despite the release of VGG-19 with an additional three layers, which only

slightly improved performance, the computer vision community generally favoured VGG-16 due to its more efficient computational resource requirements.

A significant milestone in CNNs came with the introduction of the Inceptionnet [74], also known as GoogleLeNet, developed by Szegedy et al. This model aimed to alleviate the computational burden of training CNNs by introducing 1x1 convolution for dimensionality reduction. It further enhanced performance by concurrently applying convolutions of multiple filter sizes, known as the inception block, thus accommodating for scale variations within images. Despite its depth, InceptionNet comprised only 2.4 million parameters, much fewer than VGG-16's 138 million, making it more computationally efficient and resulting in superior performance.

However, as networks increased in depth by stacking layers, problems related to vanishing and exploding gradients became more prominent. First identified in 2000, the vanishing gradient problem emerged when later network layers failed to continue learning the features established in the earlier layers. The primary cause was the insufficient learning signal reaching the earlier layers. Initial strategies to tackle vanishing gradients included employing different activation functions and gradient clipping. However, as networks deepened, these approaches proved inadequate, necessitating a more innovative solution.

ResNet [75], developed by He et al., marked a significant advancement in deep learning by addressing the vanishing gradient problem in very deep networks. ResNet's architecture, featuring up to 152 convolutional layers, was designed on the principle that a deeper network should, at minimum, maintain the performance of shallower counterparts. A key innovation in ResNet is the introduction of residual connections, which allow the network to default to the identity function at specific points. This means that the output of certain layers is the same as the input, ensuring that the performance learned in earlier layers is retained or improved upon. Residual connections effectively allow the network to 'skip' over some layers, thus simplifying the learning process and addressing the issue of vanishing gradients in deep networks. For a visual representation of how these residual connections function within ResNet, refer to Figure 3.4.

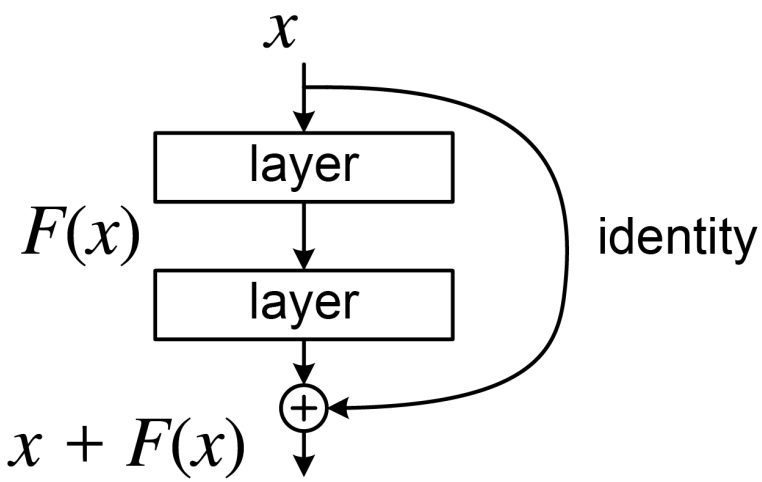



**Figure 3.4:** *Depiction of a Residual Block, a fundamental element of a Deep Residual Network. This diagram provides a clear visual illustration of the inherent architecture of a residual block, highlighting the role of the identity function and the transformation function $F(x)$. In deep networks, the training process becomes increasingly challenging with depth due to the problem of vanishing gradients, where back-propagated error signals can become exceedingly small, thereby significantly slowing, or even completely halting, the learning process. The residual block tackles this issue with its unique design: The input $x$ is passed through the transformation layers $F(x)$ and simultaneously bypassed via an identity function. The results are then added together to form $x + F(x)$, which becomes the output. In effect, the network is trained to learn the residual, or difference, between the input and the output of the transformation layers. This approach eases the learning process and enhances the performance of deep networks by facilitating the propagation of gradients all the way through to the earlier layers. It has proved to be a significant advancement in tackling tasks requiring recognition of hierarchical patterns within the data.*

Further advancements have been made with the introduction of Inception V2, V3, and V4. For this thesis, the crucial progression was the introduction of InceptionResnetv2 [76]. This method merged the structural benefits of the Inception block with its multiple-sized convolutional filters and ResNet's residual or residual connections. This combination mitigated performance degradation in deeper architectures while providing the performance and computational benefits of the Inception block.

### CNN's for Object Detection

The dawn of Convolutional Neural Networks (CNNs) ushered in significant strides in object detection, which necessitates localising objects within an image and their subsequent classification. This process traditionally entailed three stages: informative region selection, feature extraction, and classification.

Informative region selection addressed the varying appearances of objects within an image, along with their disparate sizes and aspect ratios. A proposed solution was to scan the entire image with a sliding window of different scales. Despite this method's relative effectiveness, it came with a heavy computational toll due to the high number of candidate windows. Most windows yielded a negative result, while a reduction in the number of windows led to insufficient region coverage.

The introduction of CNNs into the landscape offered two effective alternatives for object detection: two-stage object detection and single-shot object detection. The former proposes regions of interest (ROIs) that undergo subsequent prediction. The latter, as the name suggests, is a one-pass process.

R-CNN [77], a two-stage object detection method, surfaced in 2014. Instead of the previously inefficient exhaustive search methods, R-CNN proposed a selective search to extract about 2000 regions of interest, which were then passed through a CNN (Figure 3.5). This method notably reduced the number of locations under consideration and showed superior performance to preceding methods on the VOC2012 [70] dataset. However, an increase in algorithmic complexity and time consumption posed challenges.

Fast R-CNN [78], an advancement on R-CNN, replaced the method of feeding region proposals to the CNN with feeding the entire input image to the CNN to generate a feature map. This map is used for region proposal and then passed to the fully connected layers for class prediction. The culmination of this series was Faster R-CNN [79], which traded the selective search for a Region Proposal Network. Although Faster R-CNN's accuracy was similar to Fast R-CNN's, it was notably quicker.

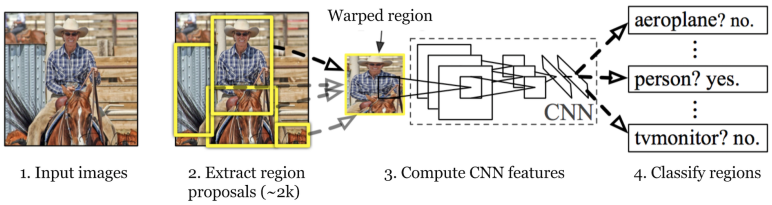



**Figure 3.5:** *This figure outlines the operation of an object detection system employing Region-based Convolutional Neural Networks (R-CNN). The R-CNN operates in two stages. In the first stage, the system takes an input image (1) and extracts a set of region proposals (2), which are potential bounding boxes that might contain objects. This extraction process, also referred to as region proposal, is conducted using a bottom-up approach, typically through a method such as Selective Search. Selective Search combines the strength of both exhaustive search and segmentation, generating possible object locations based on colour, texture, and other cues. In the second stage, each proposed region is passed to a large Convolutional Neural Network (CNN) (3), which computes a rich feature vector for each region. These feature vectors are then classified using class-specific linear Support Vector Machines (SVMs) (4), assigning a class label to each proposed region. This two-stage approach of region proposal followed by region classification allows the R-CNN to accurately detect and identify objects within an image. Figure from - Link*

The one-stage approach is an alternative for object detection. The approach executes all calculations simultaneously, using grids and anchor boxes. Among these methods are You Only Look Once [80] (YOLO) and Single Shot Detector [81] (SSD).

YOLO debuted in 2015, demonstrating the ability to perform object detection in real time. The method divides the image into grid cells and predicts bounding boxes and confidence scores for these boxes. YOLO V1 consists of 24 convolutional layers and a fully connected layer. YOLO V2 improved upon this by applying batch normalisation and anchor boxes. Anchor boxes are predefined bounding boxes of a specific width and height. These boxes aim to capture the dataset's scale and aspect ratio variation of certain object classes. YOLO-V3 significantly improved the mean average precision by increasing to 106 layers and changing the Loss.

SSD, introduced in 2015, outperformed YOLO slightly in terms of accuracy and speed. It utilised a VGG-16 as a feature extractor. However, one of the highest-performing one-stage detectors is RetinaNet [82], employing a Feature Pyramid Network and focal Loss to improve existing methodologies. This methodology forms the cornerstone of this thesis.

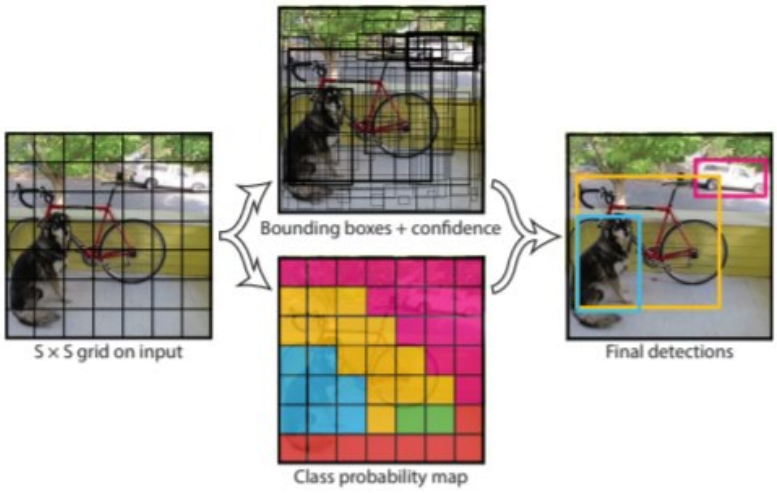



**Figure 3.6:** *This figure illustrates the operational paradigm of the YOLO (You Only Look Once) object detection system. Unlike traditional object detection systems that treat detection as a two-step process of region proposal followed by region classification, YOLO models detection as a single regression problem. This approach results in improved detection speed. The YOLO algorithm begins by dividing the input image into an $S \times S$ grid. Each grid cell is responsible for predicting $B$ bounding boxes along with the associated confidence scores for these boxes. The confidence score reflects how certain the model is that the box contains an object and how accurate it believes the box is. Additionally, each grid cell predicts $C$ class probabilities. These predictions are then encoded into an $S \times S \times (B \times 5 + C)$ tensor. For each bounding box, the model outputs five predictions: $x$, $y$, $w$, $h$, and confidence. The $(x, y)$ coordinates represent the centre of the box relative to the bounds of the grid cell. The width and height ($w$ and $h$) are predicted relative to the whole image. Finally, each grid cell also outputs a probability distribution over classes. This design of YOLO enables it to view the entire context of the object as well as the surrounding region, thereby making its predictions more globally informed. Figure from Link*

## 3.3 Computer Vision Applications in Medical Sciences

Computer vision has broad applications, from defence to autonomous vehicles and space exploration. However, its usage has been particularly prominent within the scientific community due to the ubiquity of imaging data, which has led researchers to incorporate computer vision methodologies into their investigations. In the realm of medical sciences, imaging modalities such as MRI, X-ray, ultrasound, and digital pathology, among others, are integral to routine practice [83]. The tasks to which computer vision is applied in this context span from basic functionalities, such as

noise filtering, contrast enhancement, and image sharpening, to more complex tasks, like 3D reconstruction, segmentation, and pattern recognition. All these techniques aid clinicians and researchers in accurately diagnosing and studying diseases [84].

### 3.3.1 Computer Vision in Radiology and Histopathology

With the advent of digitisation, radiology has pioneered the development and application of computer vision methods [85]. More recently, advances in capturing and managing whole slide images utilised in histopathology have resulted in an explosive increase in data volume. These advancements have facilitated rapid patient history reviews, remote consultations, and the introduction of computer-aided diagnosis [86].

Histopathology involves the microscopic examination of tissue to study the manifestations of disease. A fine tissue section, typically between $4 - 6\,\mu m$, is sliced using a microtome to generate a whole slide image. This tissue section is then mounted onto a glass slide and stained with a tissue staining solution, often a combination of hematoxylin and eosin (H&E). The slide can be immediately examined under a microscope or digitised for future reference.

Conventionally, a pathologist would study the slide under a microscope, analysing attributes such as tissue structures, cellular morphology, and colour to establish a disease diagnosis. Whilst the trained eye of a pathologist remains the gold standard, precise quantification is often laborious and subject to human error. This contributes to a reproducibility crisis in medical sciences wherein experiments cannot be consistently reproduced. Consequently, histopathology is a prime candidate for computer-aided diagnosis [87].

### 3.3.2 Challenges and Solutions in Pathological Analysis

For computational techniques to be effectively applied to pathological analysis, several challenges must be surmounted, including robust detection of nuclei and membranes, accurate classification of tissue objects, and coping with a high degree of stain variability, diverse nuclei sizes, shapes, and orientations, irregular or

non-existent cell outlines, overlapping nuclei, and significant variability in tissue morphology and background context.

The initial methodologies to address these issues relied heavily on colour-based image segmentation techniques such as thresholding [88]. This approach utilises features such as colour, texture, and intensity to segment images; for instance, if a pixel's intensity surpasses a certain threshold (Intensity X), it is designated black. Otherwise, it is white. However, these strategies are limited, including the need to iteratively trial and define parameter sets and handcraft features. Given the extreme variability in pathology images, no straightforward or universally applicable method exists for setting these parameters.

Deep learning mitigates the need to manually tune features, as it automatically learns these attributes through a cost function. As such, deep learning has ascended to the forefront of pathology analysis, addressing many of the aforementioned challenges. The advent of these machine-learning techniques in histopathology is elaborated in the following section.

## 3.4 Deep Learning in Histopathology

Artificial intelligence (AI) is rapidly advancing, making significant strides across various sectors worldwide. Medicine has not been exempt from this revolution, and the early results of AI implementation are promising. Disciplines that rely heavily on pattern recognition within data, especially those related to diagnostics, stand to gain immensely. These patterns could be evident within images or the broader context of a patient's health. However, evidence suggests that such diagnostic tasks' reproducibility is suboptimal [89].

Furthermore, the rise of precision medicine necessitates the analysis of increasingly large data sources [90]. Automated AI-based methods can address these challenges by detecting and quantifying patterns in medical data, thereby enhancing the efficiency and reproducibility of the diagnostic process while also increasing accuracy and precision [91].

Histopathology images serve as a crucial source of primary data and a rich reservoir of diagnostic information. Analysing these images demands the expertise of highly specialised pathologists, who meticulously assess gigapixel-scale images. These pathologists asseess gigiapixel scale images to diagnose and grade conditions such as cancer, inflammatory conditions, and congenital diseases based on the presence or absence of specific cell characteristics and disruptions in tissue architecture. However, the task has become untenable due to the escalating number of cases requiring extensive diagnosis and the need for more pathologists.

AI-based techniques for analysing tissue sections, termed computational pathology, have increasingly relied on deep neural networks. Initial applications of computational pathology endeavoured to implement computational features that mirrored biological processes or shapes. These generic feature banks quantified characteristics such as image texture, orientation, and contrast. Kather et al. demonstrated the success of this approach with a classifier that combined five different texture descriptors achieving 98.6% accuracy in distinguishing tumour and stroma cells in colorectal tissue sections [92].

However, computational pathology has seen a paradigm shift in recent years, moving almost entirely from feature engineering to deep learning. This transition has been driven by the advantage that deep learning offers in constructing algorithms through training rather than through the explicit programming of features, resulting in a powerful hierarchical feature representation that outperforms traditional methods.

The application of Convolutional Neural Networks (CNNs) in computational pathology began with the analysis of small patches from whole slide images, such as for mitosis counting [93]. Subsequently, methods that used entire whole slide images (WSIs) were developed for a variety of applications such as breast cancer segmentation, glaucoma classification, non-alcoholic fatty liver disease, assessment of renal transplant biopsies, and prostate cancer detection [94], [95].

As models grew more sophisticated, the development of publicly available datasets also proliferated, consistent with the trend in deep learning approaches. In

2016, the CAMELYON challenge was launched to develop computational pathology solutions for detecting breast cancer metastases in sentinel lymph node biopsies. This dataset, which comprised 1399 fully annotated WSIs from the lymph nodes of breast cancer patients, drew a range of innovative solutions from various sources, underscoring the importance of publicly available datasets in fostering innovation.

Nonetheless, there are significant obstacles to creating large, open-source datasets for computational pathology. While medical centres and laboratories can rapidly generate copious WSIs with variable staining, image quality, scanning characteristics, and tissue preparation, these datasets are only useful for current supervised deep-learning methods with annotations for training. The scarcity of annotated data can be attributed to several factors. Manual annotations required on a scale suitable for deep learning are time-consuming to generate and necessitate expert domain knowledge. This becomes especially daunting considering the wide variety of possible annotations, ranging from different cell types and subtypes to various tissue regions and disease states.

An alternative to the manual annotation of images is using clinical annotations stored by healthcare organisations. These often include disease progression reports and long-term outcomes and can provide a valuable resource. However, they present their own set of challenges. They require meticulous management by clinical data managers and must be handled with solid ethical considerations. Additionally, their usefulness may be limited; if a particular feature that has led to a certain conclusion is present in only a few cases, it may need to provide more data to train a robust model. Hence, scaling up manual annotations remains a necessary task.

One straightforward method of achieving this is to recruit more expert pathologists to annotate additional data, thus ensuring high-quality annotations. However, given the existing shortage of expert pathologists, this approach is costly and impractical. One study in 2020 adopted this solution, involving 12 senior pathologists who manually annotated over 2,000 WSIs of gastric cancer, with the agreement among all experts serving as the reference standard [96].

An alternative approach involves using individuals with varying skill levels to annotate data, with their work subsequently reviewed by a senior pathologist. This method has gained popularity, and platforms like Amazon Mechanical Turk have proven effective [97].

Regardless of the chosen approach, it is clear that the field of computational pathology and the broader medical science community stands to gain significantly from the application of computer vision and deep learning. These technologies offer the potential to improve the efficiency, accuracy, and reproducibility of diagnoses while also addressing the increasing workload on pathologists. The challenges that remain in this field primarily concern the collection and annotation of large, high-quality datasets for training deep learning models. However, the potential benefits make these challenges worth tackling.

## 3.5 Convolutional Neural Networks: A Technical Overview

Artificial Neural Networks (ANNs) are structured with input, hidden, and output layers, each containing numerous interconnected neurons. These neurons transmit information from one layer to the next, facilitating diverse applications such as natural language processing, speech recognition, and image processing [98].

### 3.5.1 Supervised Learning in Neural Networks

Machine learning algorithms primarily strive to learn the function that accurately maps input data to output data. One popular approach employed in computer vision is supervised learning. This training methodology uses labels corresponding to the input data to tune the inferred function. Consequently, the output produced has both a predicted and true label. The effectiveness of the prediction is then evaluated using a cost function, which calculates the discrepancy between these labels. Further tuning is carried out using gradient descent, incrementally adjusting the network's weights.

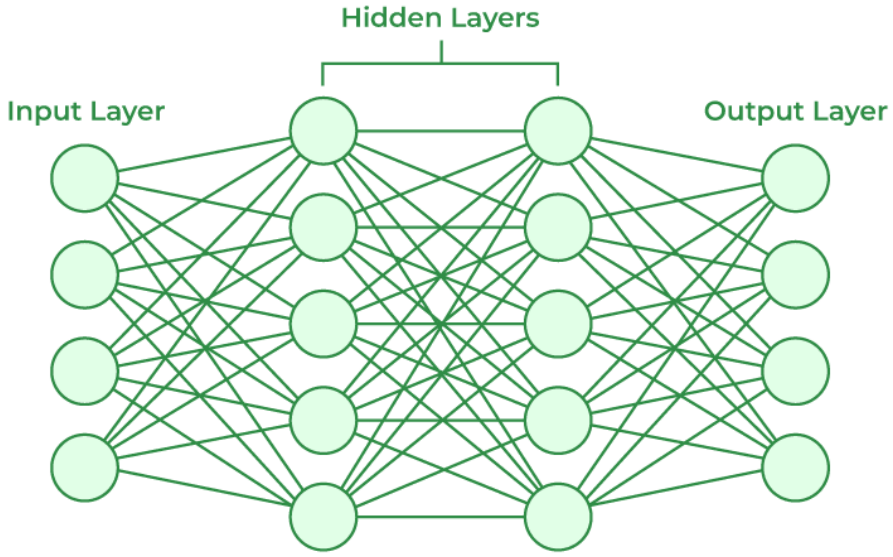



**Figure 3.7:** *This figure represents a typical architecture of a neural network, a computational model inspired by the human brain's interconnected network of neurons. The architecture comprises three distinct types of layers: an input layer, multiple hidden layers, and an output layer. The input layer, at the commencement of the network, receives the initial data for processing, which can be any form of quantifiable information, such as pixel values from an image. Following the input layer are one or more hidden layers, composed of a multitude of interconnected artificial neurons or 'nodes'. These nodes, capable of learning complex patterns and dependencies, transform the input data using a set of weights and biases, which are iteratively adjusted during the training phase of the network. Finally, the output layer, at the terminus of the network, delivers the final outcomes of the neural network's computations. These outcomes vary based on the specific task at hand, such as class labels in a classification task or continuous values in a regression task. Figure from Link*

### 3.5.2 Loss Functions and Gradient Descent

The neural network learning process relies heavily on loss functions and gradient descent. Loss functions measure the error between the true output ($y$) and the predicted output ($\hat{y}$) generated by the model. This error guides the model's learning, wherein an error of zero signifies that the predicted and true outputs match.

Gradient descent is an iterative optimisation algorithm that progressively updates the model's weights ($w(i)$) until the input function is accurately mapped. Starting with an initial (often random) value for $w(i)$, the aim is to find the lowest possible point in the loss landscape. This is achieved by calculating the loss gradient,

equivalent to the curve's derivative. A negative gradient indicates progress towards an optimal solution. The pace of this progress is determined by the learning rate, a parameter that dictates the magnitude of each step in the convex loss landscape.

The learning rate is a crucial parameter in the optimisation process. An excessively low learning rate might lead to slower convergence, while an overly high learning rate might result in divergence.

Choosing an appropriate loss function is critical, as it is the only criterion to evaluate model performance. Different loss functions may be utilised depending on the problem the model tries to solve. For instance, regression problems involve predicting continuous values and commonly use the Mean Absolute Error (MAE) or Mean Squared Error (MSE) loss functions. On the other hand, classification problems, which involve dividing datasets into distinct classes, often use Cross-Entropy Loss.

### 3.5.3 Regression and Classification Loss Functions

In regression problems, such as predicting stock prices or bounding boxes, Mean Absolute Error (MAE) is a commonly used loss function. This function computes the absolute difference between the actual and predicted values and sums these differences. MAE benefits datasets with outliers as it does not consider their direction.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - x_i|$$

However, the most widely used regression loss function is the Mean Squared Error (MSE). This function calculates the squared difference between the actual and predicted values and averages these squared differences. The MSE is particularly useful when variables are modelled into a Gaussian distribution.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Aside from MAE and MSE, other loss functions, such as Huber loss, can also be employed for regression tasks.

On the other hand, classification problems, which involve categorising data into distinct classes, require different loss functions. Classification tasks include email

spam detection or image categorisation into animal classes. Cross-Entropy Loss is the most common loss function used in classification tasks.

Cross-Entropy Loss measures the divergence between the predicted probability and the actual label. In simpler terms, the closer the output is to the actual value, the lower the Cross-Entropy Loss. This function is expressed as:

$$CE = -\sum_{i=1}^{n} t_i log(p_i)$$

Where $t_i$ represents the truth label, $p_i$ is the softmax probability for class i, and n is the number of classes. $t_i$ is 1 if the true class is i, or 0 otherwise.

The softmax function is key to the computation of Cross-Entropy Loss. This function converts logits into probabilities and is continuously differentiable, facilitating the calculation of the derivative of the loss function with respect to every weight in the neural network. This feature allows the model to adjust the weights to minimise the loss function, bringing the model output closer to the actual values.

### 3.5.4 Initialisation of Weights

Choosing initial weights and activation functions in neural networks is crucial in preventing issues such as vanishing/exploding gradient problems and ensuring optimal training speed [99]. The phenomenon of the vanishing gradient problem, initially discussed by Hochreiter et al. (1998), is often encountered when the network is initialised with small weight values. This issue arises because gradients propagated backwards during backpropagation to update the weights tend to vanish due to small weight values. As a result, the weight updates occur in minuscule increments, thereby slowing down the training process.

In contrast, the exploding gradients problem is the converse of the vanishing gradients problem and takes place when the weights are initialised with large values. In such instances, the gradients inflate excessively, resulting in weight updates that oscillate around the minimum of the loss function, causing instability during training.

Various methods are available for weight initialisation, including zero, one, random normal, random uniform, and Xavier initialisation. Initialising all weights

to zero or one results in identical weights and neuron activations, causing the model to behave similarly to a linear model, which is often undesirable for complex tasks. Random initialisation methods set weights to random numbers using a normal or uniform distribution, yet this approach can lead to vanishing and exploding gradients. Heuristic methods can be employed for deeper networks to initialise weights based on the choice of the nonlinear activation function. Notably, Xavier initialisation remains a popular approach in the research community due to its goal of equalising the variance across all layers [99].

### 3.5.5 Activation Functions

Activation functions in neural networks determine the output for a given input or set of inputs. Various activation functions exist, including Sigmoid, Tanh, Softmax, ReLU, and Leaky ReLU, and the selection of these can significantly influence model performance.

The Sigmoid function, a nonlinear activation function, maps real values to a range between 0 and 1. This function is beneficial for outputting probabilities in binary classification problems. However, its susceptibility to vanishing gradients in larger neural networks can hamper training.

Like the Sigmoid function, the Tanh function is nonlinear but maps real values to a range between -1 and 1. While the gradient of the Tanh function is stronger than that of the Sigmoid function, it is also prone to the vanishing gradient problem in deep networks.

The Rectified Linear Unit (ReLU) activation function, commonly used in convolutional neural networks, returns the maximum between an input value and zero. Despite its computational efficiency, ReLU can decrease the model's ability to fit data correctly as negative input values are not mapped, leading to the dying ReLU problem.

Leaky ReLU, a variant of the ReLU function, allows for a small negative value when the input is less than zero. This adjustment helps mitigate the issue

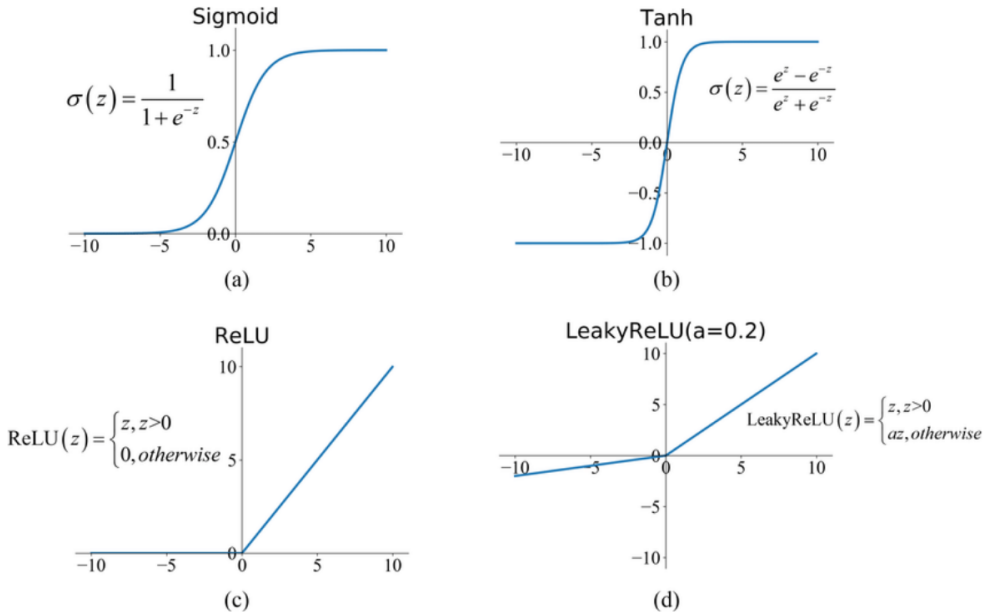



**Figure 3.8:** *This figure depicts four commonly used activation functions in neural networks. (a) The sigmoid function, often used in the output layer of binary classification tasks, maps any real-valued number into the range between 0 and 1. (b) The hyperbolic tangent function, or tanh, is a scaled version of the sigmoid function, mapping real-valued numbers into the range between -1 and 1. Both sigmoid and tanh functions are smooth, differentiable functions and are thus used when a probabilistic interation is needed. (c) The Rectified Linear Unit (ReLU) function, a popular choice for hidden layers in neural networks, gives an output equal to the input if it is positive, and zero otherwise. (d) The Leaky ReLU function, a variant of the ReLU, allows small negative values when the input is less than zero to mitigate the 'dying ReLU' problem, where a ReLU neuron becomes inactive and only outputs zero for any input. Here, we illustrate a Leaky ReLU with a slope of 0.2 for negative inputs. These activation functions are integral to the learning capability of neural networks, enabling them to model complex, non-linear relationships in the data.*

of dying ReLU by ensuring that neurons remain active and continue to learn during backpropagation.

Selecting an appropriate activation function is a key step in designing neural networks, requiring empirical testing and performance evaluation. Different problems may necessitate different activation functions—for instance, while Sigmoid functions may be suitable for binary classification problems, ReLU is often preferred for deep convolutional neural networks.

### 3.5.6 Regularisation Methods

Regularisation, a technique employed to mitigate overfitting in machine learning models, can be achieved through three main approaches: modifying the loss function, sampling, and adjusting the training algorithm.

Regularisation by modifying the loss function involves using L1 and L2 regularisation, both of which introduce a penalty term to the loss function. In L1 regularisation, parameters of most minor importance are decayed towards zero, effectively performing feature selection, while L2 regularisation decays weight slightly with each iteration but never reaches zero. L1 regularisation is particularly useful when the data contains many features, and there is a need to eliminate unimportant attributes, whereas L2 regularisation or weight decay is more often used in deep learning.

Modifying the sampling method offers another approach to overcome overfitting. Data augmentation, particularly popular in computer vision tasks, involves slight changes to the data until the model learns to focus on crucial features while ignoring noise. This can involve transformations such as flips, rotations, decolourising, adding noise, and jittering colour channels. A sampling-based regularisation approach also involves resampling the training data using k-fold validation and ensembling to determine the solution. In this method, the data is resampled k times, allowing each model to be trained on a different dataset, and a majority voting strategy is employed to make predictions.

A third approach to regularisation involves modifying the training algorithm. Dropout is a common technique that randomly sets a fraction of the input units to 0 at each update during training time, effectively reducing overfitting and providing a way to approximate model averaging [100]. This ensures that the network does not rely heavily on any single path to a solution, encouraging diverse features for finding solutions.

| Method | Description | Advantages | Disadvantages |
|---|---|---|---|
| L1 Regularisation | Adds absolute value of weights as penalty term. | Induces sparsity, enabling feature selection. | Loss of valuable information. |
| L2 Regularisation | Adds square of weights as penalty term. | Penalises large weights. | No feature selection. |
| Elastic Net | Combines L1 and L2 penalties. | Sparsity and weight smoothing. | Two hyperparameters. |
| Data Augmentation | Alters training data. | Increases data diversity. | Computationally expensive. |
| Dropout | Randomly sets input units to 0. | Reduces overfitting. | Increases training time. |
| Early Stopping | Halts training when performance stops improving. | Prevents overfitting. | May require validation set. |
| Batch Normalisation | Normalises each batch of data. | Accelerates training. | Adds complexity. |
| k-fold Cross-validation | Splits data into k subsets for validation. | Robust performance estimates. | Computationally expensive. |
| Ensembling | Uses multiple models for prediction. | Increases robustness. | Increases complexity. |
| Spectral Normalisation | Normalises the spectral norm of weight matrices. | Constrains model capacity. | May slow convergence. |
| Gradient Clipping | Limits the values of gradients. | Prevents exploding gradients. | May introduce bias. |

**Table 3.1:** Summary of regularisation methods in deep learning, including a brief description, advantages, and disadvantages.

### 3.5.7 Data-based Model Improvement

Model performance can also be enhanced by adjusting the dataset used for training. This can involve training the model, identifying areas of weak performance, and subsequently adding more data to these areas. For example, if a model demonstrates proficiency in classifying dogs but struggles with classifying cats, more images of cats may be added to the dataset to balance its performance.

A complementary strategy, known as hard negative mining, involves adding negative examples to the dataset. For instance, if the model incorrectly predicts a class when no object is present in an image, annotating the data to indicate the absence of any objects can help the model learn to differentiate between background and foreground. Another approach involves oversampling certain classes in the dataset. For example, suppose the dataset is heavily skewed towards one class (e.g., 70% dogs and 30% cats). In that case, duplicating images of the minority class can balance the class distribution and prevent overfitting the majority class.

Training a robust and high-performing neural network involves careful consideration and tuning of several aspects, including weight initialisation, activation function selection, regularisation techniques, and data-based model improvement. Each step is critical in ensuring the model learns effectively from the training data and generalises well to unseen data.

### 3.5.8 Convolutional Neural Networks

Convolutional Neural Networks (CNNs) are specialised variants of neural networks tailored explicitly for processing data with a grid-like topology, such as an image. The term "convolutional" derives from the mathematical operation "convolution," which replaces the standard matrix multiplication in at least one layer of these networks. CNNs have demonstrated tremendous efficacy in various practical applications due to their unique structure and operational mechanics.

### 3.5.9 The Convolution Operation

A convolution is a mathematical operation that measures the extent of overlap between two functions as one is shifted over the other, effectively combining the two functions. It is mathematically defined as:

$$(f * g)(t) = \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau$$

In convolutional networks, the function 'f' corresponds to the input, 'g' represents the kernel, and the result is the feature map. In machine learning contexts, the input usually manifests as a multidimensional array of data, such as an image with dimensions (w, h, c), and the kernel is another multidimensional array of parameters learned by the algorithm. These multidimensional arrays are commonly referred to as tensors.

Convolutional networks utilise three key concepts that improve machine learning: sparse interactions, parameter sharing, and equivariant representations. Sparse interactions, a distinctive characteristic of CNNs compared to traditional neural networks, result from using a kernel smaller than the input, enabling the detection of small but meaningful features such as edges and lines in images using significantly fewer parameters. This process decreases memory requirements, reduces computational load, and improves statistical efficiency, enhancing efficiency.

Parameter sharing refers to using the same parameter across multiple functions within a model. In traditional neural networks, each element of the weight matrix is used exactly once when computing the output layer. However, each kernel member is used at every input position in a CNN. This means the model learns a single set of parameters instead of a separate set for each location, further reducing storage requirements and necessary training data.

The specific form of parameter sharing in convolution imparts layers with the equivalence property to translation, meaning that if the input is translated or shifted, the output alters correspondingly. The effect of this property is demonstrated through the convolution of an image and a translation function, leading to the same result regardless of the order in which these operations are applied.

### 3.5.10 Layers of a Convolutional Neural Network

A typical layer in a convolutional network involves three stages. In the first stage, several convolutions are performed in parallel to yield a set of linear activations and biases. Each linear activation is then passed through a nonlinear activation function (often a Rectified Linear Unit, or ReLU) in stage two. Another layer commonly incorporated into CNN architectures is the fully connected layer, which receives output from the preceding layer and generates a feature vector in the final stages of the network that is then used to compute output probabilities.

The convolutional layer is responsible for extracting image features using multiple learned filters. The tunable hyperparameters for convolutional layers include the filter size, stride, and padding values. Each filter produces a feature map. Following one or more convolutional layers, the pooling layer reduces dimensionality by replacing the output of selected subregions with a summary statistic. Max pooling, which reports the maximum output within a rectangular neighbourhood, is the most commonly used pooling strategy. Pooling contributes to making the representation approximately invariant to small input translations.

The fully connected layer, usually placed in the final layers of the network, generates a feature vector for classification. Here, the relevant parameters to be set are the vector length and the activation function. The composition of these layers results in effective feature extraction, with early layers detecting small features such a, lines and edges, and later layers capturing composite shapes.

## 3.6 Terminology in Neural Networks

### 3.6.1 Parameters vs Hyperparameters

Parameters and hyperparameters represent critical components in constructing and optimising neural networks. They both play crucial roles in defining the model's behaviour and performance but are fundamentally different in nature and function.

Parameters are internal variables that the model learns through the training process. They represent the core representation learning capability of the model of

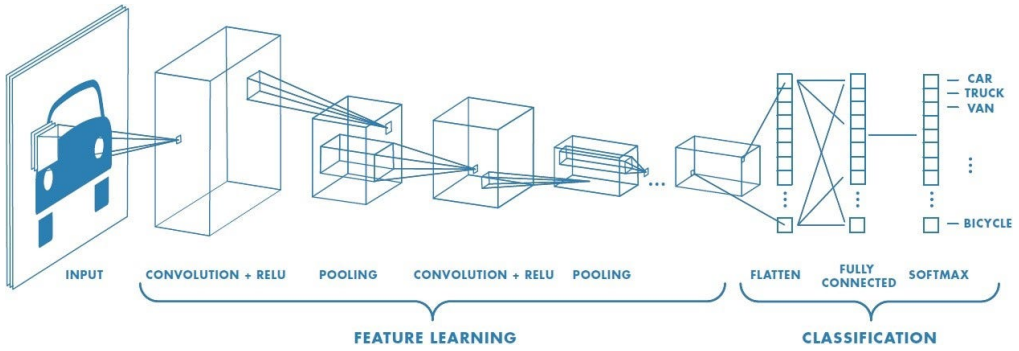



**Figure 3.9:** *This figure illustrates the architecture of a standard Convolutional Neural Network (CNN). The network commences with an input layer that accepts raw pixel data from an image. This is followed by a convolutional layer that uses a set of learnable filters to create feature maps, capturing local correlations in the image data. The convolution operation is followed by a ReLU (Rectified Linear Unit) activation function, introducing non-linearity into the model. Subsequently, a pooling or subsampling layer reduces the spatial dimensions of the feature maps whilst preserving their depth, thereby reducing computational complexity and enhancing translation invariance. Another convolutional layer, once again followed by a ReLU activation function, extracts higher-level features from the reduced feature maps. Another pooling layer then further subsamples the feature maps. The network then utilises a 'Flatten' operation to transform the 2D feature maps into a 1D feature vector, which is then inputted to a fully connected layer. The fully connected layer enables the network to learn non-linear combinations of high-level features as detected by the previous layers. Finally, a softmax function in the output layer provides a probabilistic distribution over the target classes, enabling the model to make a prediction. The initial layers (convolutional and pooling) can be seen as the 'feature-learning' part of the model, while the fully connected and softmax layers can be considered the 'classification' part. Figure from Link*

the model and directly contribute to its ability to predict outcomes based on input data. These include weights and biases in the network that are iteratively updated as the model seeks to minimise the loss function [101]. In a neural network, weights are the coefficients for each input, determining the influence of each input on the output. On the other hand, biases are constants added to the product of inputs and weights to adjust the output along the activation function.

In contrast, hyperparameters are external configurations set before the training process begins and guide the learning process of the model. Unlike parameters, hyperparameters are not learned and must be preset , for instance via cross validation of appropriate values. They include settings such as the learning rate, the number of epochs (complete passes through the entire dataset), the batch size (number of samples processed before the model is updated), the number of hidden layers and

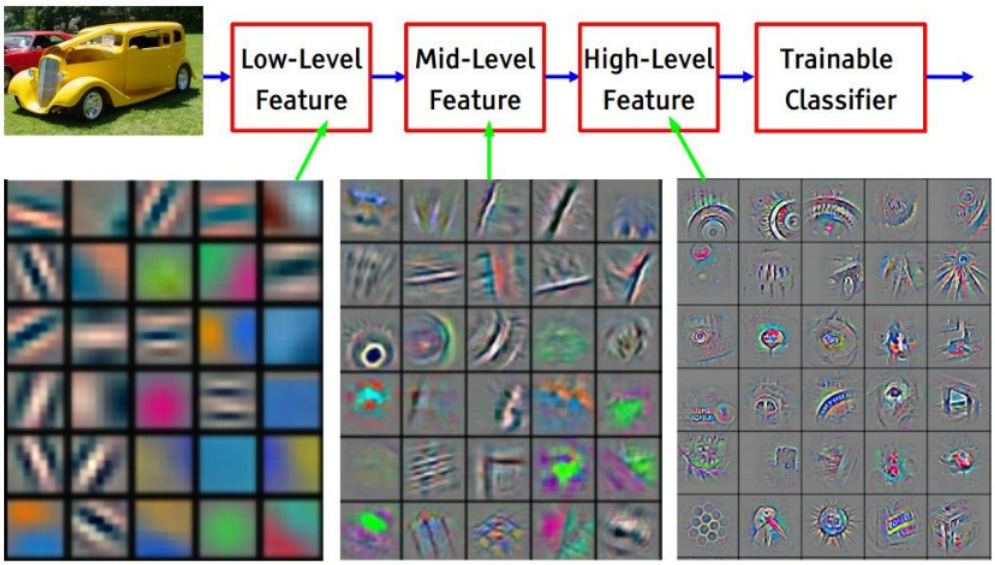



**Figure 3.10:** *This figure showcases the hierarchical feature extraction process within a Convolutional Neural Network (CNN) when processing an image of a car. The initial layers of the network (usually convolutional layers) extract low-level features such as edges and corners. These are fundamental visual elements that are relatively simple in nature but crucial for the understanding of more complex structures in the image. As we progress deeper into the network, mid-level features begin to emerge in the form of part-compositions of the low-level features, such as contours and basic object parts (e.g., wheels, windows). Further on, high-level features are abstracted from these mid-level features, which correspond to more complex and complete representations of the object (e.g., the car body or the entire car itself). The power of CNNs lies in their ability to automatically learn and abstract these features from raw pixel data, rather than relying on manually-engineered features, thereby enabling them to robustly handle a wide range of visual recognition tasks. Figure from Link*

neurons in the network, and regularisation parameters (to prevent overfitting) [102].

The learning rate is a critical hyperparameter that determines the step size during gradient descent in optimisation. A lower learning rate could make the learning process slower but more precise, while a larger learning rate could lead to faster learning but risk overshooting the optimal point.

Dropout is another hyperparameter frequently used in neural networks, especially in deep learning. This is a regularisation technique that prevents overfitting by randomly dropping out (i.e., temporarily removing) a fraction of neurons during training, forcing the model to spread out its learning and creating a more robust network.

### 3.6.2 Epochs vs Iterations

In machine learning, "iteration" and "epoch" represent two different but interconnected concepts pertaining to the process of training a model.

An "iteration" refers to the event of passing a single batch of inputs through the network. The "batch size," is a hyperparameter that determines the number of samples to work through before the model's internal parameters are updated.

An "epoch," on the other hand, is a complete pass through the entire training dataset. During an epoch, the neural network's weights are updated to reduce the loss function. The number of epochs is another hyperparameter and indicates how often the learning algorithm works through the entire training dataset.

Let us say we have a dataset of 100 images, and we have set the batch size to 10. Completing one epoch would take ten iterations (passing ten batches of 10 images each). The number of epochs and batch size are crucial hyperparameters that can influence the speed and effectiveness of the training process.

These terminologies form the backbone of understanding and optimising the training of neural networks. Mastery of these terms and concepts can enhance the effectiveness of model tuning and result in more efficient and accurate predictive models.

## 3.7 Summary and Conclusion

In this chapter, we have explored the extensive applications of computer vision and deep learning in the field of medical sciences, particularly in histopathology. This interdisciplinary intersection has significantly advanced the automated detection and diagnosis of diseases, promising increased efficiency and reduced human error. Despite the challenges posed by the complexity of pathological analysis and the high degree of variability in pathology images, deep learning methods, specifically Convolutional Neural Networks (CNNs), have emerged as powerful tools capable of addressing these challenges.

The intricacies of CNNs were discussed including their unique structure, operational mechanics, and the critical role of the convolution operation. These networks capitalise on several key concepts, such as sparse interactions, parameter sharing, and equivariant representations, to improve machine learning. We also reviewed the typical layer composition in a CNN, including convolutional, pooling, and fully connected layers, which collectively result in effective feature extraction.

To optimise the performance of these networks, we examined various aspects, such as weight initialisation, activation function selection, regularisation techniques, and data-based model improvement. We also elaborated on critical terminologies, including parameters, hyperparameters, iterations, and epochs, all of which underpin the understanding and optimisation of neural network training.

While challenges remain, particularly in relation to the collection and annotation of large, high-quality datasets for training deep learning models, the potential benefits to medical sciences are immense. The application of computer vision and deep learning technologies stands to enhance the efficiency, accuracy, and reproducibility of diagnoses, thereby aiding clinicians, pathologists, and researchers in their crucial work. As these techniques continue to evolve, we can anticipate their further integration into medical practice, ultimately contributing to improved patient outcomes and care.

*"Our genes are a template for our development, and our development is a dance between those genes and our environment."*

— Siddhartha Mukherjee

# 4 The DMDD Dataset

## Contents



## 4.1 Introduction

This chapter is the first result chapter of the thesis, where the preparation of the data is described. The primary dataset used in this study is derived from the Deciphering Mouse Developmental Disorders (DMDD) study. In this chapter, I will outline the key objectives of the project, the DMDD dataset, and some significant findings. I will then provide a detailed account of the DMDD placental dataset, which is the foundation of this thesis. I will discuss the laboratory preparation of the dataset thoroughly, followed by an examination of the work done to prepare it for computational analysis. This dataset will be publicly accessible for supplementary analysis or incorporation into a more comprehensive computational

pathology project.

## 4.2 The DMDD Project

Understanding gene function poses a significant challenge in genetics. Animal models, particularly those involving mice, have emerged as critical tools for discerning gene function, owing mainly to the advent of sophisticated gene-editing technologies. A key methodology for investigating gene function involves gene knockout experiments, a genetic technique that involves the inactivation or removal of a specific gene in an organism. Various techniques can be utilized to carry out knockouts, the most prominent being CRISPR-Cas9 [46].

Efforts to decipher the mouse genome have led to systematic knockout experiments that inactivate each of the 20,000 protein-coding genes, with the corresponding phenotypes noted. The International Knockout Mouse Consortium (IKMC) is a noteworthy initiative in this respect, seeking to construct a resource aimed at discerning the role of individual genes in maintaining normal physiological function in mice and identifying genes that might contribute to human disease [103].

Large-scale studies have found that approximately one-third of all targeted mouse knockouts result in embryonic or prenatal death [104]. Existing phenotyping pipelines, such as the one employed by the International Mouse Phenotyping Consortium, must be adequately equipped to phenotype the resultant embryonic lethal animals. These lines could offer invaluable insights to developmental biologists and significantly enhance our understanding of human developmental disorders.

With funding from the Wellcome Trust, the Deciphering Mechanisms of Developmental Disorders (DMDD) [105] was established to tackle the challenge of phenotyping knockout mutants that result in embryonic or prenatal death. Mouse lines were generated at the Wellcome Trust Sanger Institute and subjected to intensive imaging and detailed transcriptomic analysis. This facilitated the creation of a comprehensive public resource for standardised analysis to investigate the molecular and developmental pathways disturbed in mammalian development. The primary focus of the DMDD was on structural defects associated with lethality.

The phenotyping pipeline employed microscopy to detect morphological abnormalities associated with the gene knockout. The prenatal lethal lines were initially triaged based on the stage of development at which lethality occurred. The DMDD endeavoured to collect data at E14.5, as the significant period of organogenesis was complete by this stage. Embryos that could be harvested underwent extensive imaging using high-resolution episcopic microscopy (HREM). HREM is a volumetric data generation technique that yields stacks of inherently aligned digital images that mimic the appearance of eosin-stained histological sections, enabling the capture of relatively large volume specimens in histological quality.

The E14.5 embryos were comprehensively imaged with a spatial resolution of 2-3$\mu$m, which allows for phenotype detection with greater accuracy than with micro CT, optical projection tomography, or MRI. Lines not viable at E14.5 were harvested at E9.5 for imaging, and a complete transcriptome analysis was performed. The collected images were then systematically scrutinised for morphological defects by developmental biologists and anatomists. This analysis was conducted across littermates for each knockout and is available on the DMDD website (Figure 4.1).

## 4.3 DMDD Placenta Dataset

Historically, research on developmentally critical genes has been predominantly focused on the embryo, with extraembryonic tissues receiving considerably less attention. This has led to an underrepresentation of placental phenotypes in research databases. For instance, the Mouse Genome Informatics database identifies extraembryonic defects in 10% of embryonic lethal strains [107]. Recognising this gap, the DMDD sought to redress this imbalance by generating a comprehensive placental dataset. The goal was to gain a more precise understanding of the actual frequency of placental abnormalities.

The inception of this dataset was facilitated by pivotal studies highlighting the crucial role of extraembryonic tissue in development [35]. It is widely recognised that anomalies in placental function can drastically affect pregnancy outcomes, leading to severe complications such as premature birth and IUGR, the two principal causes

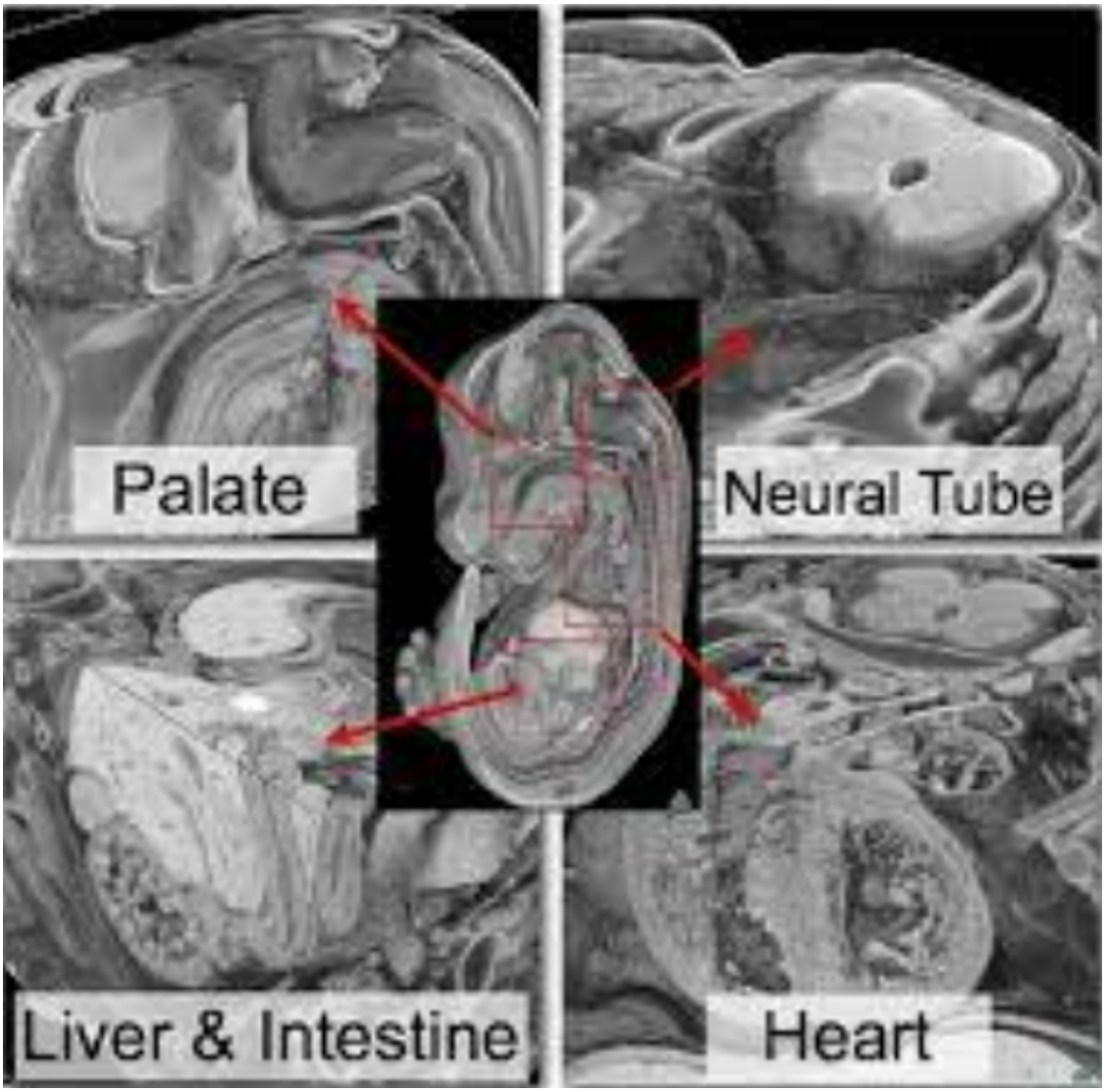



**Figure 4.1:** *A representative slice from High-Resolution Episcopic Microscopy (HREM) imaging of a mouse embryo, exhibiting discernible developmental disorders. The depicted anatomical regions illustrate defects in the formation of critical structures. The palate, a key component of craniofacial development, exhibits a noticeable deformity. Likewise, the neural tube, which serves as the precursor to the central nervous system, displays irregularities indicative of malformation. Abnormalities can also be observed in the liver and intestine, potentially signifying disruptions in the development of the digestive system. Moreover, the heart, vital for circulatory system development, also manifests clear signs of a developmental disorder. These anomalies underscore the potential of HREM as a powerful tool for investigating the aetiology and progression of developmental disorders in mouse models and the significant role of the DMDD project. Figure from [106]*

of infant mortality. Furthermore, preeclampsia, a disease primarily affecting the placenta, can result in severe or fatal outcomes for both mother and child [108] [33].

The literature underscores the significance of the intrauterine environment in adult health. Low birth-weight babies, for instance, are known to experience higher mortality rates due to heart disease later in life, along with elevated rates of metabolic disease, heightened risk of neurological dysfunction, and increased prevalence of other chronic diseases. This suggests that the placenta should be of considerable interest to the scientific community, not only for its potential for improving people's lives in both the short and long term but also for reducing the burden on health services [7].

The DMDD, therefore, sponsored a systematic analysis of 103 embryonic lethal knockouts. The Hemberger lab carried out this study. Most mouse lines were generated using the EUCOMM/KOMP conditional ready targeted ES cell resources, with some created via CRISPR-Cas9 gene deletion. These lines were produced at the Wellcome Trust Sanger Institute on a C57BL/6N genetic background. The term genetic background refers to the genetic makeup of all alleles, excluding the mutated gene of interest. Consistency in genetic background is desirable in a study as variation in other genetic material can significantly influence the phenotypes of the mutants.

Gene knockout lines were classified as lethal if no homozygous mutants were present among a minimum of 28 pups at P14 and subvisible if their proportions fell on or below 13% of total offspring from heterozygous intercrosses. After harvesting, placentas were fixed in 4% paraformaldehyde (PFA).

For histology analysis, at least three mutant and three wild-type placentas from at least two independent litters were processed. Wherever possible, placentas of female and male conceptuses were analyzed. The placentas' tissue appearance and cellular architecture were used to confirm the sample's viability for analysis. In the context of the DMDD placenta dataset, each 'sample' refers specifically to individual sections of placenta that have been processed and stained for histological examination. The term 'sample' encompasses both mutant and wild-type placental

tissues, each derived from separate placental entities. To ensure robust comparative analysis, multiple sections (samples) may be taken from a single placenta, depending on the experimental design and objectives. This approach allows for the assessment of variability and consistency across different areas of the same placenta, as well as between different placentas. After fixation, the placentas were sliced into consecutive 7um sections and processed for hematoxylin and eosin (H&E) staining using standard protocols. H&E staining, a technique that has been used for over a century, allows for identifying tissue types and morphological changes. Hematoxylin, which stains nucleic acids deep purple, and eosin, which stains proteins nonspecifically pink, provide a versatile stain that reveals a broad range of cytoplasmic, nuclear, and extracellular matrix features.

The H&E stained placental sections are sorted and sectioned through the sagittal midline chosen from imaging. This ensures that the same section of the placenta is being analysed in each sample. The sagittal midline at E9.5 is identified by the remnant of the uterine lumen and at E14.5 by the site of insertions of the umbilical cord. Imaging was carried out using a Hamamatsu slide scanner.

## 4.4 Preparing the DMDD Placental Data for Computational Analysis

The DMDD dataset, comprising images and an Excel spreadsheet with metadata for each sample, was received from the Hemberger lab. The images were in the NDPI format, indicative of a pyramid format. Image pyramids are multi-scale signal representations used to store large detailed images. A pyramid of images at different resolutions is constructed through repeated smoothing and subsampling, enabling only the area of interest in the image to be loaded into memory at any given time. Files stored in this manner allow analysis at the micro- and macroscopic scale with minimal computational memory requirements (Figure 4.2).

The initial step in preparing the data for computational analysis involved altering the naming conventions for all samples to make them computer-readable. The annotation of the placentas came subsequently. Given that each image contained,

on average, twelve image levels, the placentas were extracted at image level five, deemed optimal for whole placenta annotation. At this level, each pixel in the images represents 3.52 microns, achieved through the process of downsampling where the resolution is progressively reduced. Specifically, from the original resolution at Level 0, where each pixel corresponds to 0.11 microns, each subsequent level doubles the area each pixel covers by halving the image dimensions, thus increasing the pixel representation geometrically up to Level 5. This scaling ensures efficient data handling and consistency in feature analysis across images. The placentas were manually annotated by drawing a rectangle around each one and assigning a label linked to the sample metadata, using the VGG annotator. All details regarding the pyramid scale, metadata, and camera specifications were meticulously cataloged and stored in a low-memory Python object known as a pickle file, facilitating streamlined and efficient downstream analysis of the placentas. This methodological approach ensures that despite the reduced resolution inherent to Level 5 images, the annotations remain precise and comparable across the dataset.

Figure 4.2 provides a visualisation of the annotation process. The steps involve extracting a pyramid level suitable for annotation, using the VGG annotator to draw a tight bounding box around the placenta, linking the metadata to the appropriate bounding box, and saving all location data and metadata to a low-memory Python object.

A total of 8713 placentas from the DMDD dataset were annotated. After attaching metadata, the number of samples dropped to 7242 this was due to labels from the histology lab not being present that could attached meta data to a placenta slice. The data originated from 925 unique animals across 103 gene knockouts, with an approximately equal split between stages E9.5 and E14.5. Of the 3326 images that contained sex data, there was an approximately equal split between males and females. The data breakdown is shown in Table 4.1

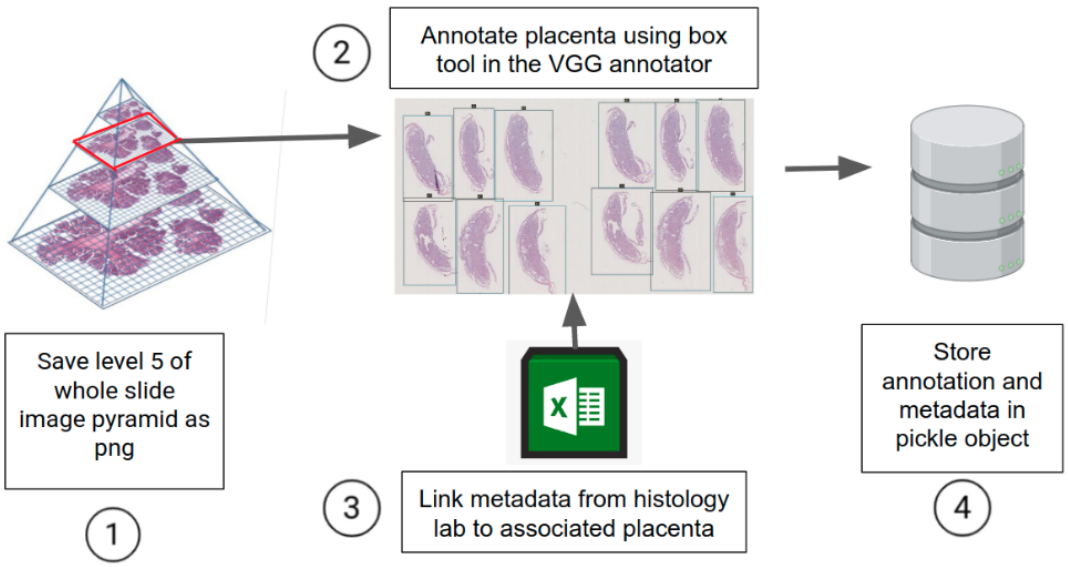



**Figure 4.2:** *Schematic representation of the procedure for annotating histology sections. Step 1 involves extracting a level from the image pyramid that offers sufficient resolution for accurate annotation. In Step 2, the VGG Image Annotator tool is utilised to delineate a tightly fitting bounding box around the placenta, thus marking the region of interest. Subsequently, in Step 3, the relevant metadata is linked to the identified bounding box, thereby associating the anatomical structure with pertinent contextual information. Step 4 involves storing all the location data and associated metadata in a lightweight Python object to facilitate efficient storage and retrieval. This methodical process ensures that each histological section is accurately annotated, thereby paving the way for subsequent analyses.*

## 4.5 Preparing Data for Model Training

Upon completing individual placenta annotation, the data was prepared for the training and inference of the deep learning methodologies. The specifics of these approaches are comprehensively discussed in Chapter 5. For the whole slide images to be utilized in the pipeline, they had to be segmented into smaller sub-images, commonly referred to as tiles. A tile size of 300x400 pixels at a resolution of 0.11 micrometres was chosen. All images were rescaled to this resolution using the camera metadata and a Python resize function.

Following the tiles' preparation, data preparation for nuclei detection began. The nuclei detection method is rooted in deep learning for object detection and necessitates bounding box annotation. Therefore, entire tiles were fully annotated with tight bounding boxes drawn around nuclei, as demonstrated in Figure 4.3.

**Table 4.1:** DMDD data breakdown. WT denotes Wild Type, Hom denotes Homozygous, and Het denotes Heterozygous. In this table, the term 'sample' refers to individual histological sections of placenta processed for analysis. Each sample may represent a different section of a single placenta or sections from multiple placentas, depending on the specific requirements of the study and the detailed parameters of tissue sampling

| **Data Type** | **Number of Samples** |
|---|---|
| Full annotated data set | 8713 |
| Data available post-processing | 8163 |
| Labelled data | 7242 |
| Unlabelled data | 921 |
| Unique animals | 925 |
| Unique genes | 103 |
| **Stage** | |
| E14.5 | 3262 |
| E9.5 | 3980 |
| **Genotype** | |
| WT | 3562 |
| Hom | 3609 |
| Het | 71 |
| **Sex** | |
| Female | 1560 |
| Male | 1766 |

However, due to the vast number of cells contained in whole slide images, no full slide was annotated, only sections of each slide. For the DMDD mouse nuclei detection, 29,401 nuclei were annotated for training and 5,880 nuclei for validation, spanning nuclei stages E14.5 and E9.5.

The exact number of unique placentas annotated could not be verified as this data was lost during the curation process. However, efforts were made to balance the representation of placentas during annotation. Annotations included nuclei from the subcategories E14.5 and E9.5 as listed in Tab 4.1, but specific numbers per subcategory are currently unconfirmed and require further data review.

Nuclei were annotated with a rigorous training set curation where each nucleus was annotated with a focus on its characteristic purple nuclei, typical of H&E-stained slides. This annotation helped train the deep learning model to accurately recognize and delineate nuclei, even in densely populated cell areas. Additional challenges

such as overlapping cells were addressed through methods like bootstrapping and hard negative mining. Bootstrapping involved the iterative addition of difficult or misclassified cases back into the training set, enhancing the model's learning from these complex scenarios. Hard negative mining was utilized to collect instances where the model initially failed, focusing on improving detection in areas of particular difficulty.

Once the training sets for nuclei detection were created, the next task was to create cell classification training sets. This involved defining the cell types visually identifiable in H&E-stained placental tissue. Eight cell types were determined for classification at the E14.5 stage:

1. Decidual Cells (DEC). This cell type encompasses all cells in the decidua of maternal origin.

2. Endothelial Cells (END). These cells are found around arteries, veins, and capillaries, primarily in the labyrinth.

3. Chorionic Plate Cells (CHP). This category comprises a general group of cells in the chorionic plate, consisting of mesoderm and trophoblast cells.

4. Foetal Blood Cells (FET). These are nucleated blood cells in the labyrinth.

5. Trophoblast Giant Cells (GIA). These large, distinct cells are mainly found between the decidua and junctional zone, but some can also be located in the labyrinth.

6. Glycogen Cells (GLY). As the name suggests, these cells, located in the junctional zone, store glycogen.

7. Labyrinth Cells (LAB). This group encompasses cells in the labyrinth not already mentioned, predominantly syncytiotrophoblast type 1 and 2.

8. Spongiotrophoblast Cells (SPO). These cells are located in the junctional zone.

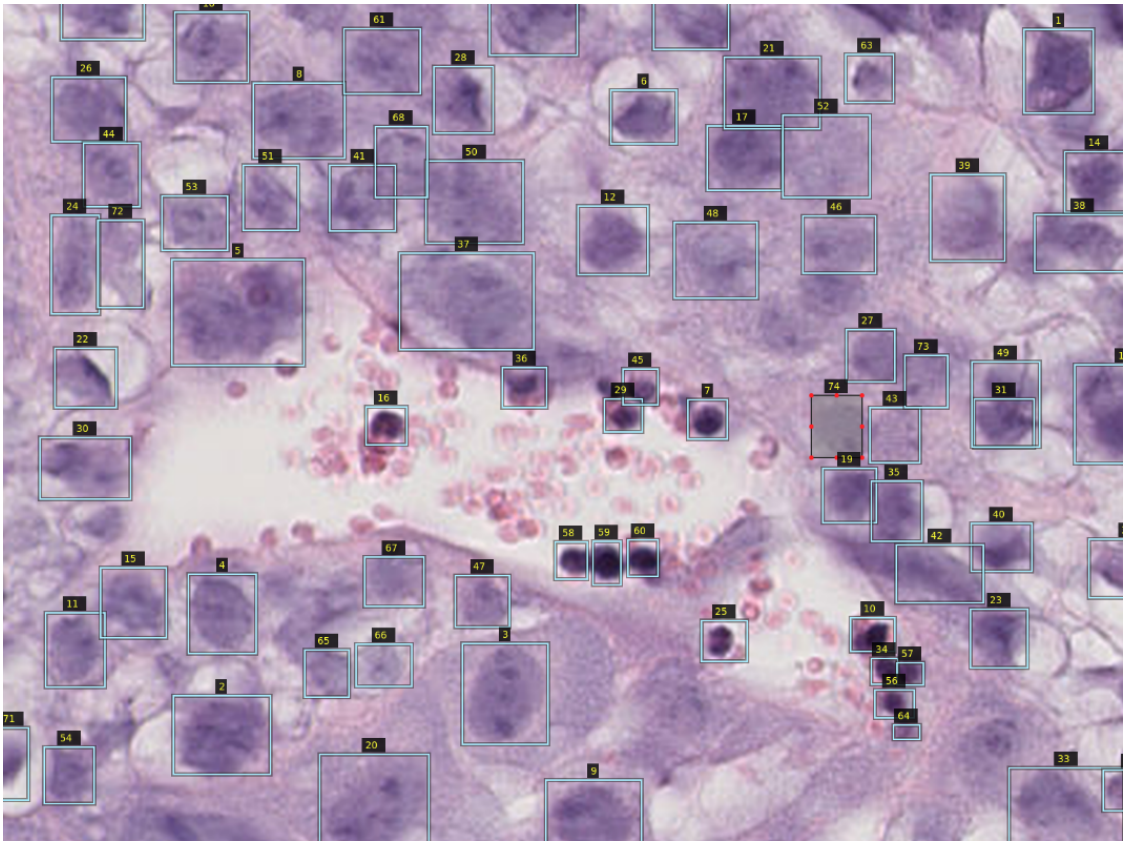


**Figure 4.3:** *This image demonstrates a fully annotated tile utilized to train the object detection model for placental nuclei identification. Each nucleus is marked with an encompassing bounding box, representative of the target output of the detector. This annotation process is crucial for the supervised learning approach adopted by the model, enabling it to recognize and localize nuclei in histological sections of the placenta with high precision. The image is 1200x1200 pixels and 1 pixel = 0.11 micron*

While these cell categories can be further subdivided into other types and subtypes, such detailed breakdowns are challenging to achieve through visual analysis alone. Advanced techniques such as RNA-seq and immunohistochemistry could provide more granular categorization if necessary. However, for the scope of this work, the classification into eight cell types is sufficiently detailed.

For the training dataset, a total of 34,401 cells were annotated, with an equal distribution across the eight categories for training. An additional 6,880 cells, again equally split across the categories, were annotated for validation. This dataset will also be made publicly available.

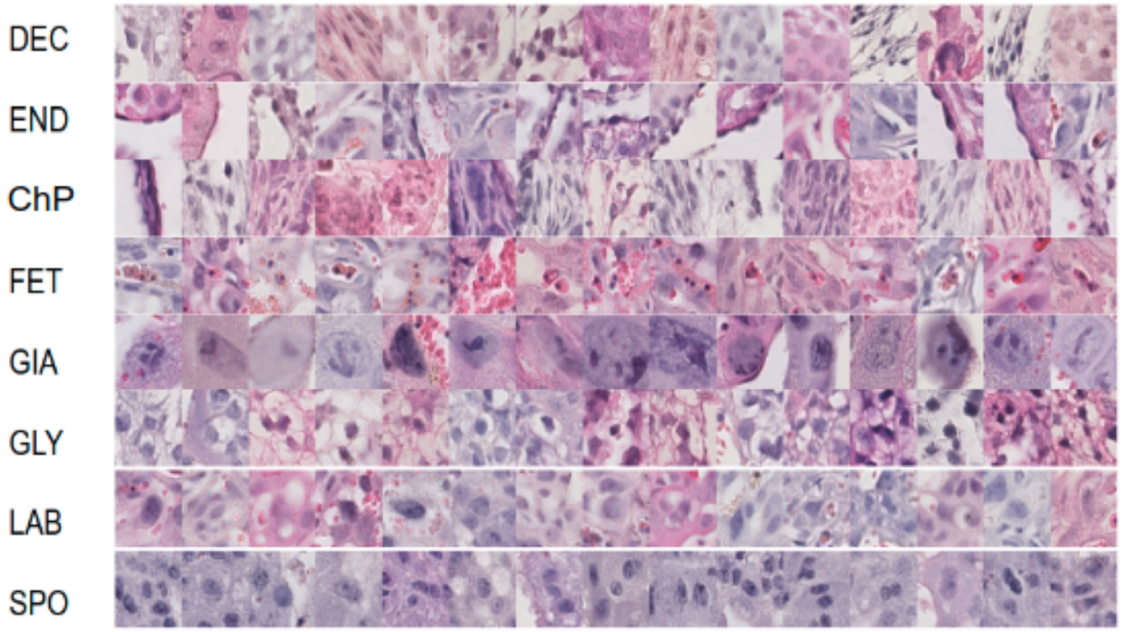



**Figure 4.4:** *A visual representation of the distinct cell types that have been annotated for the purpose of training the cell classification model. These include: (1) Decidual Cells (DEC), originating from the maternal decidua, (2) Endothelial Cells (END), primarily found surrounding the labyrinth's blood vessels, (3) Chorionic Plate Cells (CHP), comprising a mix of trophoblast and mesodermal cells located in the chorionic plate, (4) Foetal Blood Cells (FET), nucleated blood cells situated within the labyrinth, (5) Trophoblast Giant Cells (GIA), large, distinctive cells mainly found at the interface between the decidua and junctional zone, with a few in the labyrinth, (6) Glycogen Cells (GLY), glycogen-storing cells located in the junctional zone, (7) Labyrinth Cells (LAB), encompassing cells in the labyrinth not already categorised, primarily syncytiotrophoblast type 1 and 2 (8) Spongiotrophoblast Cells (SPO), situated in the junctional zone. Each cell type's annotation plays a critical role in the model's ability to accurately identify and classify cells within placental histological sections.*

## 4.6 Summary and Conclusion

This chapter has detailed the preparation of the Deciphering Mouse Developmental Disorders (DMDD) mouse placental dataset, a critical step in utilizing this resource for computational research on gene function and developmental disorders. The meticulous preparation and annotation of this dataset underscore its potential as a foundational tool for future studies, particularly in understanding the genetic basis of developmental processes and abnormalities in the mouse placenta.

The following key outcomes have been achieved:

1. Annotation of Placental Data: All placental samples within the dataset have

been systematically annotated. 2. Creation of Training Sets: Specialized training sets for nuclei detection and cell classification have been developed. These sets are crucial for the application of deep learning techniques in analyzing placental images. The training sets cover a variety of cell types ensuring that the computational models developed can robustly handle diverse data inputs.

In conclusion, the preparation of the DMDD dataset for computational analysis marks a significant advancement in placental biology research. By providing a richly annotated, publicly accessible resource, this work supports the broader scientific community in its efforts to decipher complex genetic interactions and their impacts on development. Future studies leveraging this dataset are poised to deepen our understanding of developmental disorders, potentially leading to innovations in both diagnosis and treatment. This chapter not only fulfills its objectives but also sets a high standard for data usability and collaborative research in the field of developmental genetics.

*"In creating AI, we're birthing a new form of life with unlimited potential for good or ill."*

— Max Tegmark (Life 3.0: Being Human in the Age of Artificial Intelligence)

# 5

# Development of an Automated Placental Phenotyping Method

## Contents

## 5.1 Introduction

Histology, the microscopic examination of plant or animal tissues, plays a pivotal role in diagnostic medicine and the study of biological processes such as cancer pathogenesis and embryogenesis. Routinely utilised in disease staging in clinical settings, histology provides a rich data source for analysis [109]. The advent of digital whole slide images, which facilitate tissue slide viewing on computers instead of conventional microscopes, has broadened the application of histological analysis in medicine [110]. However, the routine analysis of tissue sections remains a labour-intensive manual task performed by highly trained pathologists, resulting in high costs. Challenges such as a shortage of histopathologists and reproducibility issues in results could be significantly mitigated through automated histological analysis techniques.

Traditional automated tissue analysis techniques employed thresholding methods such as watershed and OTSU's methods [111] as well as other approaches that leveraged hand-crafted features like those available in Cellprofiler [112]. Since the introduction of Alexnet in 2012 [72], deep learning has advanced to the forefront of computer vision, becoming increasingly prevalent in automated histology analysis. Deep learning methods have surpassed human capabilities in various tasks, notably cancer diagnosis. Nevertheless, there are several tasks where automation remains limited.

A notable example of such a task is the analysis of placental histology tissue. Despite the placenta's crucial role in understanding developmental diseases, it remains surprisingly understudied (section 2). Large datasets of placental histology images are underused in research and as diagnostic tools, primarily due to a dearth of expert histopathologists and the high costs associated with the manual analysis of these samples. The application of automated deep learning methods could address this issue.

Ferlaino et al. [113] demonstrated that a placental study at the cellular level could be performed using an automated approach. The authors proposed using two deep convolutional neural networks, decoupling the cell localisation and classification tasks. Nuclei detection was performed using an object detector, followed by cell classification using an image classification method. They achieved a mean average precision (mAP) of 64% for nuclei detection and 89% classification accuracy across five cell classes in the human placenta, with an ensemble of three image classification networks reaching the same classification accuracy. Nuclei detection was performed using the object detection approach retinanet [114] and cell classification is carried out using an ensemble of image classification networks. Retinanet at the time of publication was the best performing object detection method in addition to being a one shot detector and the image classification networks were among the top performing.

Mean average precision is the benchmark performance metric in object detection and segmentation systems, with values ranging between 0 and 1. This metric integrates the two core tasks of object detection: object localisation and image classification. The object localisation score is calculated using the intersection over union, which evaluates the degree of overlap between a predicted bounding box and the ground truth, with a threshold set to determine success. Intersection values over union above 0.5 are commonly deemed as true positives. The subsequent task involves verifying the accuracy of the image classified within the bounding box.

Although Ferlaino et al.'s research marks a significant stride towards deep placental phenotyping, it has substantial limitations. The study does not incorporate

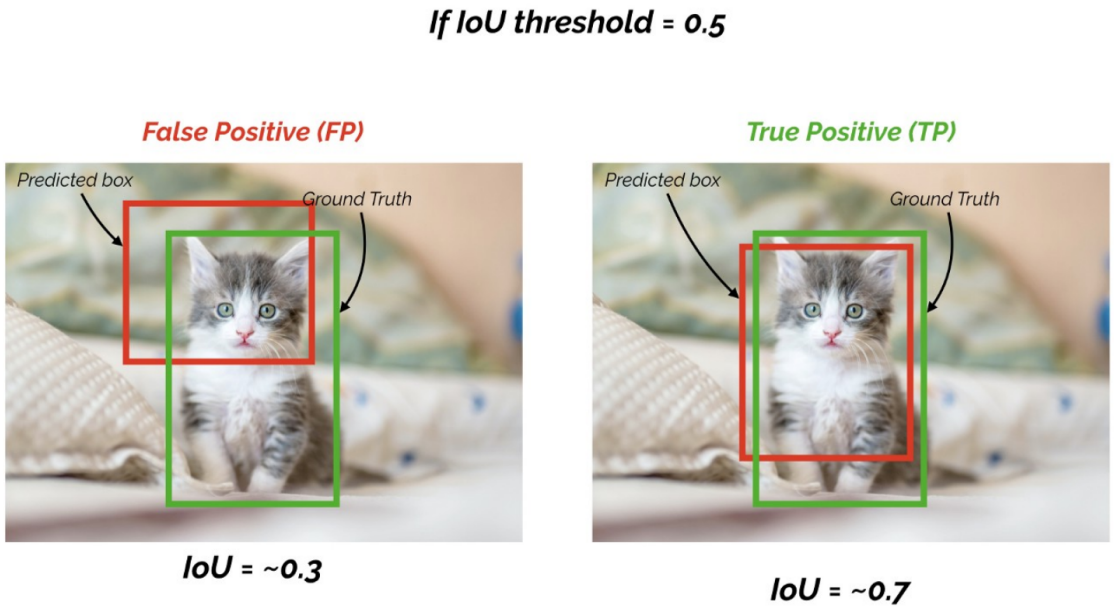



**Figure 5.1:** *This figure presents the Mean Average Precision (MAP) of a cat at differing Intersection over Union (IoU) thresholds, notably at IoU=0.3 and IoU=0.7. The IoU threshold is set at 0.5, serving as the determinant between true and false positives. The cat image with IoU=0.3 is a false positive, as its IoU value is lower than the threshold, suggesting the bounding box does not correctly encapsulate the cat. Conversely, the cat image with IoU=0.7 is a true positive, indicating a high degree of overlap between the predicted bounding box and the ground truth. This visual comparison underscores the importance of the IoU threshold in object detection tasks, as it significantly influences the evaluation of model performance. Figure from - towardsdatascience.com/map-mean-average-precision-might-confuse-you-5956f1bfa9e2*

whole slide image analysis, instead operating on image patches or tiles of 1600x1200 pixels, which may overlook edge cases. This method only permits the estimation of approximate cell ratios. Moreover, the study was conducted on a single placental sample with specific regions annotated, raising potential generalisability issues with the models. Given the variability in (H&E) stained tissue, it is essential to develop considerably more robust models. Lastly, the method was constructed in an early version of Keras, with minimal consideration for pipeline efficiency. Computational requirements and time efficiency are crucial factors for large-scale data deployment.

## 5.2 Aims of the Chapter

The primary purpose of this chapter is to elucidate the methodological framework adopted for the development of a comprehensive, automated placental phenotyping

procedure applicable to whole slide histology images.

### 5.2.1 Specific Objectives

The following objectives have been delineated:

1. Augment the pioneering human models established by Ferlanio et al., further extending their application to the realm of human placental studies.

2. Design and implement a mouse nuclei detection system, supplemented by a robust cell classification model, thereby enhancing the precision and reliability of the phenotyping process.

3. Devise an efficient strategy for processing whole slide images, thereby increasing the feasibility and practicality of large-scale histological examinations.

## 5.3 Core Methods

The methodology employed in this chapter for analysing whole slide placental images involves a two-step process: decoupled nuclei detection followed by cell classification. The one-stage object detection approach, RetinaNet, is used for nuclei detection, while the image classification architecture InceptionResNetV2 is used for cell classification.

### 5.3.1 RetinaNet

Object detection can be performed using either a two-stage or one-stage method (section 3). Two-stage approaches, such as R-CNN [77], generate a set of candidate regions in the first stage and then classify each foreground region as a class or background in the second stage. These two-stage approaches were previously considered state-of-the-art, consistently producing superior results compared to one-stage methods.

The comparatively poor performance of one-stage approaches was attributed to the foreground-background class imbalance, which overwhelmed the loss function. In two-stage detectors, candidate regions are reduced to around 2000 in the region

proposal step. In contrast, one-stage detectors have closer to 100,000 regions. To address the class imbalance problem, the authors of the RetinaNet paper [113] proposed an adjusted loss function. While object detection is typically conducted using cross-entropy loss, the authors introduced a scaling factor that decays to zero as the confidence in the correct class increases. This adjusted loss, termed focal loss, downweighs the effects of easy background examples and forces the model to focus more on challenging foreground examples.

The cross-entropy loss, denoted by $L(p_t)$, is given by:

$$L(p_t) = -\alpha_t log(p_t)$$

where $\alpha_t$ is the weighting factor, and $p_t$ is the class probability.

The proposed focal loss, denoted by $FL(p_t)$, is expressed as:

$$FL(p_t) = -\alpha_t(1 - p_t)^{\gamma} log(p_t)$$

where $\gamma$ is the focusing parameter.

The effects of focal loss are illustrated in Figure 5.2. This loss function allowed the model to outperform two-stage object detection approaches while maintaining the swift speed characteristic of one-stage detection. The employed architecture was a Resnet backbone with a feature pyramid network built on top. For object detection, a classification subnet and a bounding box regression subnet were used (Figure 5.3).

The feature pyramid network was developed in computer vision to address the common challenge of recognising objects at vastly different scales. Image pyramids had previously been employed in object detection, such as sliding window approaches where solutions were altered to identify objects at different scales. The feature pyramid network, however, leverages the hierarchical pyramidal nature of convolutional neural networks to tackle the scale problem. The feature maps at different spatial resolutions combine the low-resolution, semantically strong features with the high-resolution, semantically weak features. This results in a pyramid of features with rich semantics at all levels, enabling the network to perform well across all scales consistently.

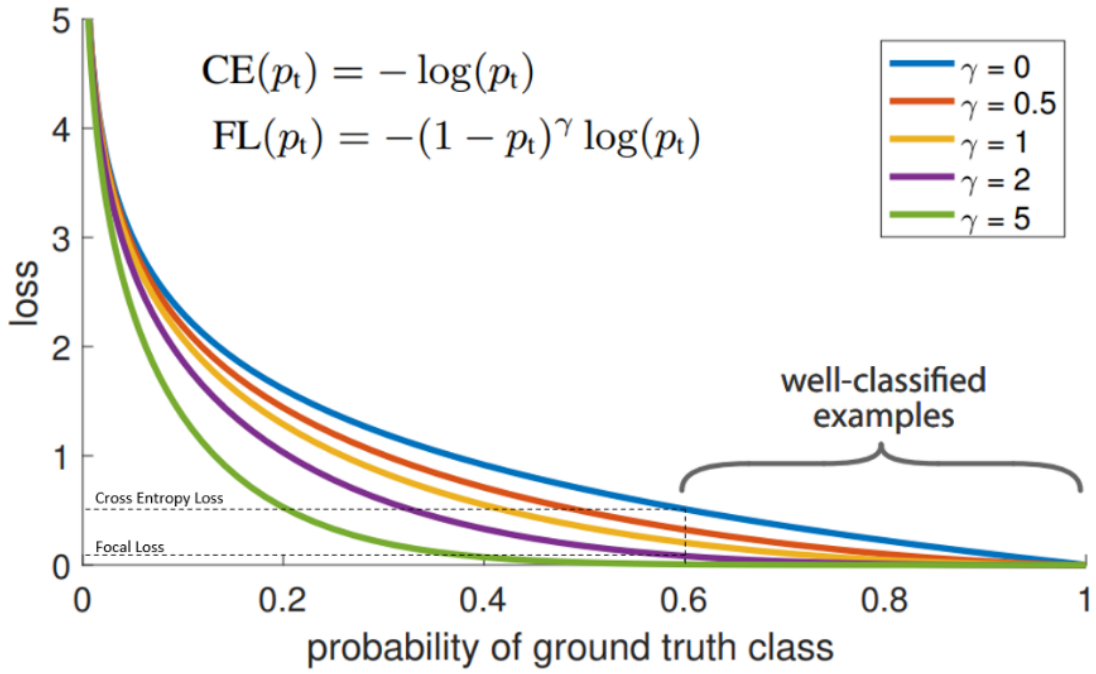



**Figure 5.2:** *This figure illustrates the behaviour of the Focal Loss function, a variant of cross-entropy loss designed to address class imbalance in object detection tasks. The y-axis represents the Focal Loss, while the x-axis denotes the ground truth probability. Various curves correspond to different focusing parameter $\gamma$ values, including 0, 0.5, 1, 2, and 5. As $\gamma$ increases, the function becomes more focused on hard, misclassified examples, reducing the loss contribution from easy, correctly classified instances. This visual comparison highlights the influence of $\gamma$ in shaping the loss function and underscores its role in managing class imbalance. Figure from - Link*

### 5.3.2 InceptionResNetV2

Over the past decade, deep learning has revolutionised the field of computer vision. Significant breakthroughs, such as the inception architecture and residual connections, have facilitated substantial performance improvements Figure 5.4. The InceptionResNetV2 architecture [76] endeavours to amalgamate these two techniques to create a wide and deep model, thus offering superior performance.

The InceptionResNetV2 model, comprising 164 layers, integrates the variable-sized convolutional filters of a single inceptionnets layer with the residual connections that default to the identity matrix, thereby mitigating degradation problems. The architecture of the model is depicted in Figure 5.4.

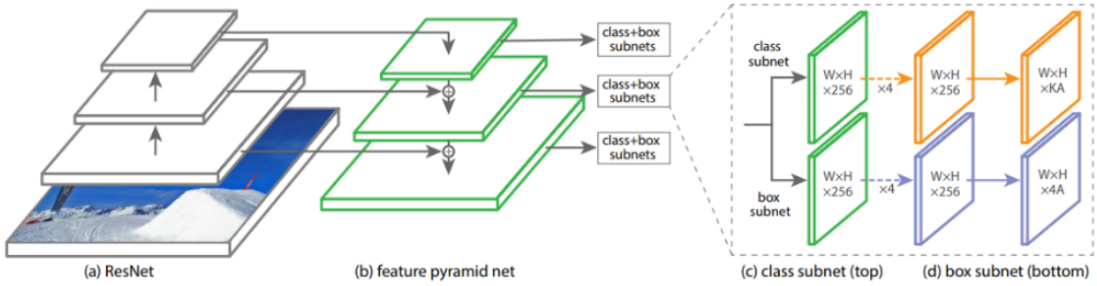



**Figure 5.3:** *This figure depicts the structure of the RetinaNet architecture, a state-of-the-art object detection model. The architecture comprises four principal components: a) ResNet, a convolutional neural network that serves as the backbone for feature extraction; b) Feature Pyramid Network (FPN), which uses the high-level semantic feature maps from ResNet to build a multi-scale pyramid of features; c) the Class Subnet, responsible for determining the object classes present within the detected bounding boxes; and d) the Box Subnet, tasked with regressing the object bounding boxes. This schematic representation illuminates the hierarchical and multi-scale nature of the RetinaNet architecture, underscoring its capacity for detecting objects across a wide range of scales and complexities. Link*

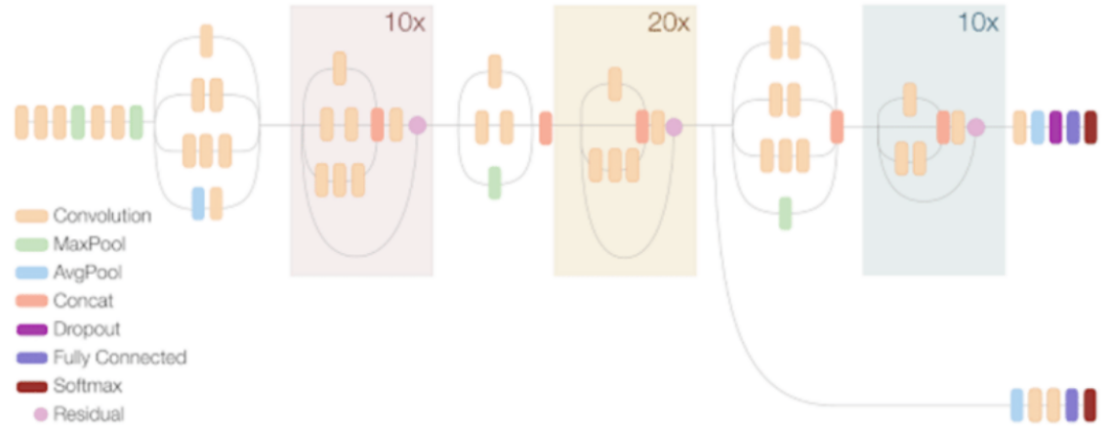



**Figure 5.4:** *This figure presents a schematic of the InceptionResNetV2 architecture, a convolutional neural network model that combines the strengths of the Inception networks and the ResNet model. The architecture incorporates several key components: convolutional layers for spatial feature extraction; max pooling and average pooling layers for feature down-sampling; concatenation operations to merge information from different layers; dropout for regularisation and prevention of overfitting; fully connected layers for high-level reasoning; softmax for output class probabilities; and residual connections to alleviate the vanishing gradient problem. This visual representation elucidates the complexity of the InceptionResNetV2 architecture, emphasising its capability to handle sophisticated pattern recognition tasks. Figure from - Link*

### 5.3.3 Original Contributions

This work presents several original contributions:

1. Development of an end-to-end whole slide analysis method tailored explicitly for placental analysis. This comprehensive approach enables the efficient and accurate examination of entire slides, enhancing the potential for nuanced insights into placental histology.

2. Creation of a mouse nuclei and cell classification model dedicated to localising and classifying mouse placental cells. This model represents a valuable tool for researchers studying mouse models of placental development and disease.

3. Refinement and significant improvement of a high-performing human nuclei and cell classification model. This updated model offers enhanced accuracy and reliability in identifying and classifying human placental cells, contributing to the depth and breadth of human placental research.

4. Assembly of a comprehensive annotated dataset for machine learning in the analysis of the DMDD dataset. This richly annotated dataset is a critical resource for training and validating machine learning models in the field of placental histology, thereby facilitating future research in this important area.

## 5.4 Methods

### 5.4.1 Method Overview

The methodology outlined in this chapter signifies a substantial evolution and personal contribution beyond the foundational work of Ferlaino et al. [1]. My approach not only maintains the core principles of their decoupled two-stage approach but also introduces significant enhancements that elevate both the results and the applicability of their study. To provide a clear understanding of the innovations and custom developments. Figure 5.5 illustrates the logical flow of the entire process is presented at the beginning of this section.

Central to the advancements presented in this chapter is the development of a sophisticated whole-slide image analysis method, which represents a significant evolution from the foundational work of Ferlaino et al. [1]. This new method adheres to the core principles of their decoupled two-stage approach but introduces substantial enhancements that significantly improve model performance and broaden the applicability across diverse datasets.

Key Enhancements

1. Framework Transition and Codebase Overhaul:

    - Framework Shift: A strategic shift from the Keras framework to the more advanced PyTorch framework marked the beginning of this overhaul. This transition was motivated by PyTorch's dynamic computational graph and robust community support, which enhance both the development flexibility and execution efficiency of deep learning models.
    - Codebase Reconstruction: The entire codebase was restructured to facilitate scalability and maintainability. This comprehensive redevelopment allows for quicker iterations during experimental setups and provides a modular architecture that can be easily adapted or extended for future research needs.

2. Model Performance and Generalizability:

    - Enhanced Accuracy and Efficiency: By integrating advanced deep learning techniques such as transfer learning, aggressive data augmentation, and a nuanced approach to model training, the analysis method now achieves superior accuracy. These improvements also extend the model's applicability to a broader range of histological image types and qualities.
    - Customized Models for Specific Studies: Models have been meticulously tailored to address specific characteristics of the mouse placenta. This customization is critical for studies in developmental biology, where detailed and accurate phenotypic assessments are essential.

3. Streamlined Analysis Process:

    - Abstraction Layers: Several layers of abstraction were introduced within the code to streamline the analysis process. This not only optimizes the workflow for the current project but also sets a foundation for adapting the pipeline to additional datasets without extensive reconfiguration.
    - Automated Image Processing: The pipeline now includes automated mechanisms for handling image variations and complexities, such as adjusting for different scales and resolutions found in whole-slide images, ensuring consistent output quality.

The proposed approach for whole-slide tissue analysis in this chapter encompasses five main phases (figure 5.5):

1. **Load in and process whole-slide histology images.** This stage involves identifying the placenta of interest in the whole slide image. Scale normalisation is then applied when passed to models. The image is then divided into smaller sub-images, known as tiles, which can be loaded into RAM and quickly passed through the models.

2. **Nuclei Detection.** Nuclei within tiles are detected using the one-stage object detection approach, RetinaNet. This model is trained using transfer learning, aggressive data augmentation, bootstrapping, and hard negative mining.

3. **Edge Case Detection Processing.** The edge cells of a tile are either missed or cut off as tiles are passed through the models. This issue is addressed by loading additional overlapping tiles and passing them to the nuclei detection model. Multi-detections of the same cell are removed using a KD-Tree clustering approach.

4. **Cell Classification.** In the cell classification stage, the input to the classifier consists of image segments centered around each detected nucleus. These segments are 100x100 pixel squares extracted from the whole-slide images.

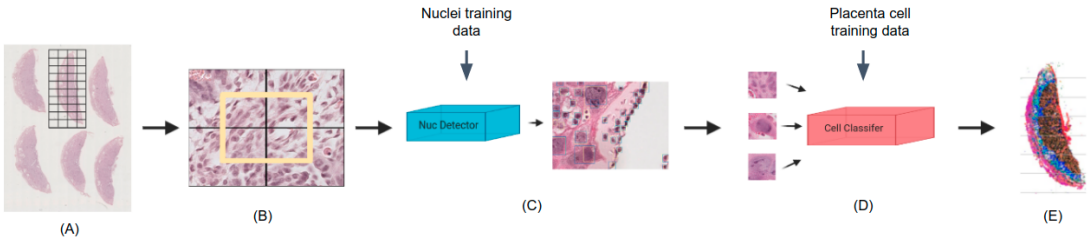



**Figure 5.5:** *This figure provides a detailed overview of the multi-stage process for comprehensive automated placental phenotyping using whole-slide histology images. The procedure involves: (1) loading and processing whole-slide images, involving the identification of the placenta, scale normalization, and tiling; (2) nuclei detection using a RetinaNet model, enhanced with techniques such as transfer learning, aggressive data augmentation, bootstrapping, and hard negative mining; (3) processing edge cases by using overlapping tiles and a KD-Tree clustering approach to address issues of missed or cut-off cells; (4) cell classification using a classifier built on an InceptionResNetv2 architecture, further improved with strategies including transfer learning, aggressive data augmentation, dropout, bootstrapping, and hard negative mining; and (5) global structure analysis, involving the reconstruction of tiles and the return of cell locations and classifications to their original positions for global phenotype analysis. The depicted pipeline underscores the intricacy and sophistication of the automated placental phenotyping process.*

Depending on the position of the nucleus and the density of the surrounding cellular material, these segments may capture a complete cell or just a part of it. This is also the case with each cell class having varying sizes. This variability is captured in the training data and is critical for training the classifier to accurately recognize different cellular configurations under varied conditions. Therefore the models performs well in capturing all of the cell classes as show in Figure 5.16.The cell classifier model is built using an InceptionResNetv2 architecture. Transfer learning, aggressive data augmentation, dropout, bootstrapping, and hard negative mining are all used to create a cell classifier.

5. **Global Structure Analysis.** The tiles are then reconstructed, and the cell locations and classifications are returned to their positions within the global structure. This allows global analysis to examine phenotypes such as cell ratios and density.

In the following few sections, I will detail the various phases and describe the approaches within each method step.

### 5.4.2 Load in and Process Whole-Slide Histology Images

#### Processing Whole-Slide Placental Images for Analysis

Histopathology images, such as those used in this research project, come as whole-slide images (WSIs). These images are obtained by digitising microscope slides at diagnostic resolution. Pathologists need to pan and zoom rapidly on these images to be adequately analysed. Therefore, the images are enormous and come in a pyramid format (figure 5.6). At the top of the pyramid are low-resolution images, which allow the viewer to see broader structural features. Higher-resolution images at the bottom of the pyramid allow the viewer to see cellular features.

Reading WSIs using standard image tools or libraries is challenging due to their large size and high-resolution format. WSIs exceed RAM, often occupying tens of gigabytes when uncompressed. Additionally, WSIs are typically multi-resolution; only a small amount of image data might be needed at a particular resolution.

There is no universal data format for WSIs, so each microscope vendor implements its own formats, libraries, and viewers. This generalises pipelines complex as WSIs are often collected using different microscopes. One of the aims of this research chapter is to create a pipeline for whole-slide analysis that is as generalisable as possible. This will allow the automated method to be widely deployed in research and clinical settings.

One library that allows for whole-slide image processing is OpenSlide[10]. This C library provides a simple interface to read the image. OpenSlide solves the problem of having a general solution for many microscopes. If the microscope file cannot be opened by OpenSlide, the next library the method developed in this chapter tries is Bioformats[11]. This is another extensive software tool for reading and writing image data. The combination of these open-source libraries allows a wide range of WSI files to be opened.

In a clinical context, the expeditious processing of information can be of critical importance. Equally, in research environments, accelerated data handling can contribute significantly to cost-effectiveness. Given the necessity of managing substantial data volumes in this pipeline, swift image processing emerges as a

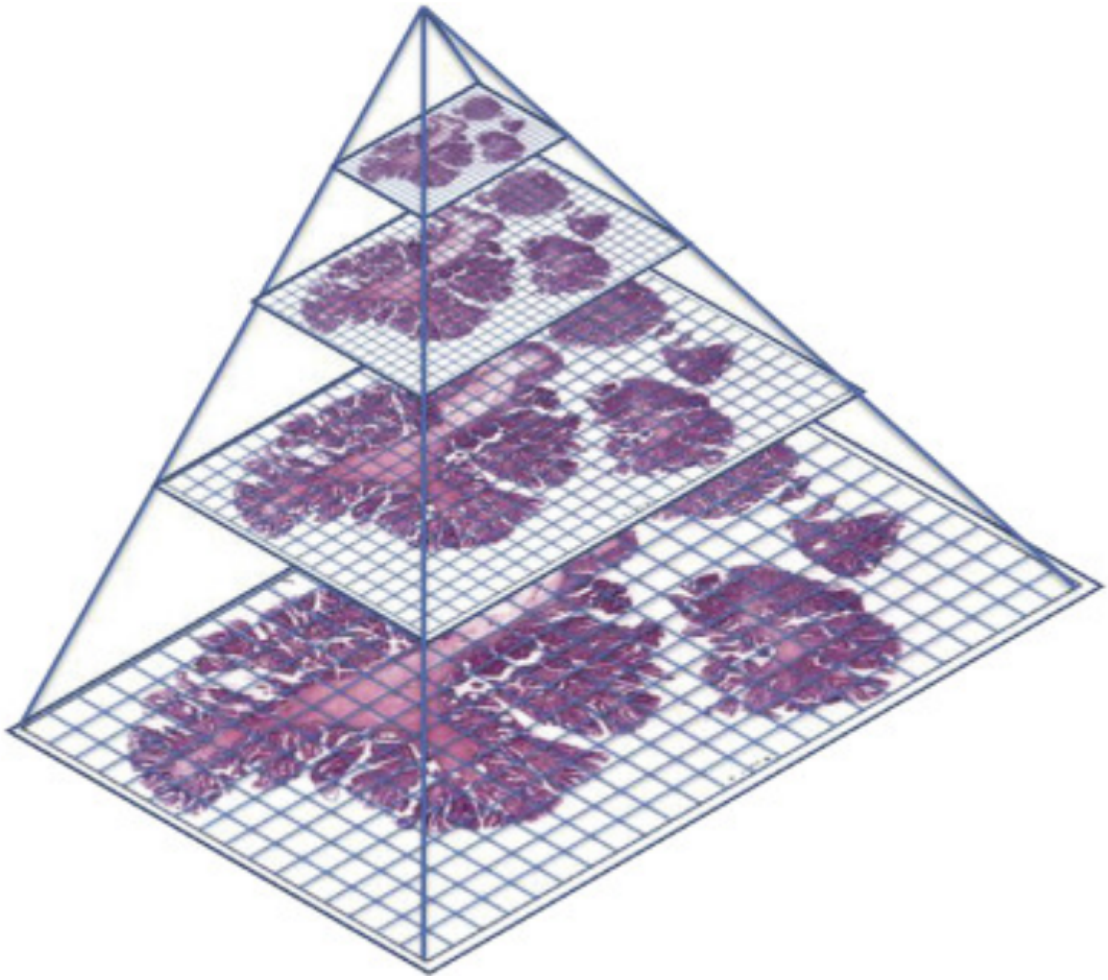


**Figure 5.6:** *This figure illustrates a whole slide image pyramid used for histology analysis, displaying four distinct levels of magnification. The pyramid provides an efficient representation of the image at varying scales, facilitating multi-scale analysis and interpretation. The topmost level offers a global, low-resolution overview of the histological sample, while subsequent levels progressively reveal finer details, with the bottommost level offering the highest magnification. This multi-resolution structure is instrumental in identifying structures of interest at varying scales and enables a comprehensive and detailed histological evaluation.*

major requirement. To address this need, we employ the libvips image processing library, renowned for its efficacy with large images. By retaining only the pixels currently under processing, libvips demonstrates a commendable economy of RAM usage. Additionally, its parallelised operation across multiple CPUs underscores its efficiency. A comparative analysis of the libvips library is presented in Table 5.1.

Another issue when processing large amounts of data is the processing and storage of data and results. It must be considered an easily understandable and

**Table 5.1:** Performance comparison of image opening and rescaling libraries

| | Mean Open WSI | Mean Read Random Region | Mean Rescale SKimage | Mean Rescale PILLOW | Mean Rescale PILLOW-SIMD |
|---|---|---|---|---|---|
| OpenSlide | 0.00959 | 0.00929 | 0.16467 | 0.00221 | 0.002404 |
| Libvips | 0.00117 | 0.00015 | 0.16399 | 0.0023 | 0.002749 |

*The table presents average values from 1000 test runs, with times recorded in seconds. (a) Mean Open WSI: Time required to open a full whole slide image, with libvips demonstrating significantly faster performance. (b) Mean Read Random Region: Time required to extract a random region of 1600x1200 pixels, again with libvips outperforming Openslide. (c) Mean Rescale: Performance comparison of image manipulation libraries in rescaling tasks from 400x300 pixels to 1600x1200 pixels. In this task, PILLOW emerges as the most efficient library, surpassing even PILLOW-SIMD, an optimised variant of PILLOW.*

accessible structure with a low memory footprint. Approximately 8000 placenta images and associated metadata exist in this research project. The results must save the x and y coordinates for every cell, the class label, etc. To contain all this information, the choice was made to work with pickle objects for each placenta. These low-memory containers will allow for an easy-to-work-with, structured, and low-memory solution to data management.

The steps outlined in this section ensured that the pipeline is built on good foundations for generalisable (OpenSlide) and fast (libvips) image processing and ensured data is saved in a low-memory structured format (pickle). These considerations are essential for big data solutions capable of deployment across various clinic and lab settings.

### 5.4.3 Tiling the Placental Images for Analysis

Now that the placental data is structured and easily opened and viewed, the challenge remains that the images at the highest resolution are too large to be loaded into RAM and passed through a machine-learning model. A standard solution for running inference on large images in computer vision is image tiling, which involves breaking a large image down into smaller sub-images or tiles.

The process of breaking down a large image into sub-images is straightforward. In this method, a size of 1600x1200 pixels is selected, and a sliding window iterates

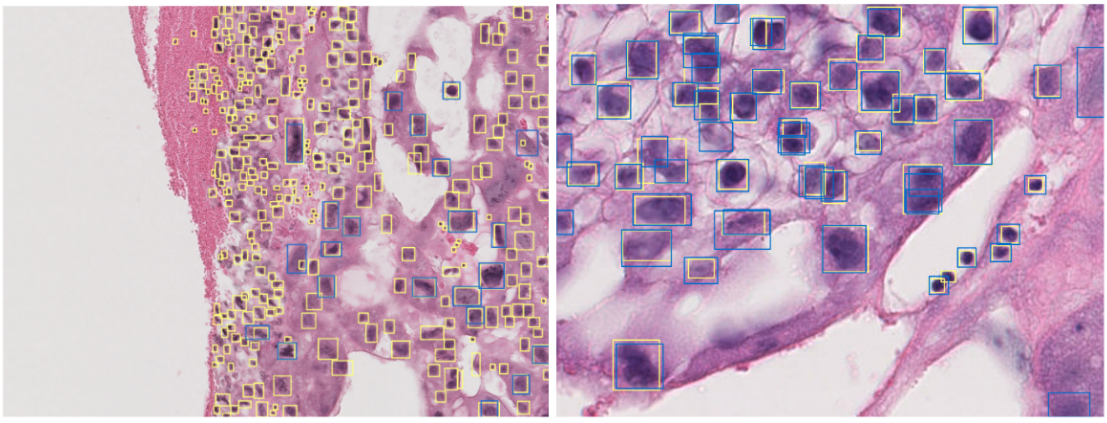


**Figure 5.7:** *This figure presents a comparative view of nuclei detection performance at disparate magnification levels. On the right, a zoomed-in image exhibits excellent nuclei detection, with the model accurately identifying and delineating individual nuclei. In contrast, the zoomed-out image on the left demonstrates a diminished performance in nuclei detection, with a significant proportion of nuclei inadequately detected. This comparison underscores the influence of image magnification on detection performance, highlighting the importance of appropriate image scaling and resolution in histological image analysis tasks.*

over the horizontal and vertical axes, outputting each image until the entire slide image has been covered. However, a key objective of this research is to create a fast, generalisable solution. One issue encountered when working with different datasets captured using different microscopes is image resolution. Images and objects at different scales are expected to result in diminished performance. Techniques such as feature pyramids can mitigate this issue, but wide variations in scale can still result in poor performance. An example of this is demonstrated in Figure 5.7 and discussed in the results section.

To overcome these scale issues, it is necessary to determine the resolution at which the microscope has captured the whole slide image and then rescale it to a predetermined size, in this method, 1600x1200 pixels. This ensures that the model will always see images of the same resolution. The resolution at which the whole slide image was captured is saved in the pickle files, ensuring streamlined processing. The rescaling factor is computed from the microscope resolution and the desired resolution and is then passed to an image-rescaling function. The efficiency of image processing is another crucial factor. Several image-rescaling functions were tested,

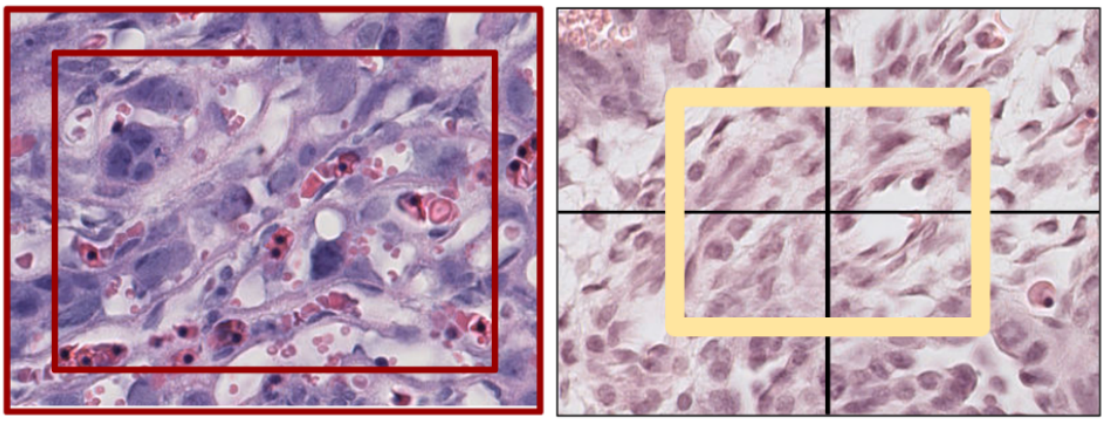


**Figure 5.8:** *This figure presents a two-part illustration of the process of handling edge cases in whole-slide histology images. The left image displays a placenta histology tile, where the edges representing the top 20% of the image, highlighted in red, demonstrate the risk of missing or cutting off cells during tiling. The right image shows the same tile, but with an additional layer of overlapping tiles covering the highlighted edges. These overlapping tiles are created with a 50% overlap, ensuring comprehensive coverage and accurate nuclei detection. The combination of overlapping tile creation and the subsequent use of a K-D Tree-based clustering approach to remove multiple detections of the same cell effectively addresses the challenge of edge cases in whole-slide image analysis.*

with the most efficient method coming as part of the Python image processing library Pillow, and is implemented in the end-to-end method (Table 5.1).

At this stage, the whole slide image is broken down into manageable chunks or tiles, which can be passed through the model. Images are rescaled so all are at the same resolution. This is done efficiently, ensuring continued processing speed. The image locations from tiling the WSI are saved in the pickle file, so this processing only has to be done once for further efficiency and reproducibility. Padding is added for images on the edge to ensure the model has a consistent input size.

### 5.4.4 Overcoming Cell Edge Cases in the Tiles

Working with large whole-slide images requires image tiling to meet computational requirements. However, this introduces a new challenge: cells at the edges of tiles can be missed. This issue is addressed by creating overlapping tiles passed to the model alongside the original tile. In this method, each overlapping tile is chosen to have a 50% overlap, ensuring maximum coverage for nuclei detection. The process of creating overlapping tiles is shown in Figure 5.8.

Now that overlapping cells have been captured and edge cells detected, overlapping tiles can result in multiple detections of the same cell, leading to inaccurate results. To overcome this problem, a clustering approach is implemented. Specifically, a K-D Tree-based approach is developed to remove multiple detections.

The KD-Tree, a nearest neighbour algorithm [115], linking nodes of the data, was developed by Jon Bentley in 1975. This method uses the KD-Tree approach to determine the nearest neighbours to nuclei. Conceptually, a cell can be modelled as a sphere of radius $r$, with a distance $d$ between the cell nucleus and the next cell nucleus. The aim is to find any detections within the cell area, find the centre point of these detections, and designate that as the detection location. To do this, the minimum distance $d$ between cells must be identified, and detections below that threshold are labelled as part of the same cell. The five nearest neighbours to every detection are found to calculate the threshold, and the distance to each nearest neighbour is plotted. As shown in Figure 5.11 (Results), plotting this as a histogram illustrates that close detections are high, decreasing before reaching a minimum and then increasing again. The distance of the minimum is the $r$-value distance that can be considered as the threshold for a single cell. A function is then built into the model that finds any detections below that threshold, their centre point, and designates that as the cell detection.

Through empirical analysis and extensive quality control, there are very few examples where a particularly large trophoblast giant cell might appear to have multiple nuclei spaced sufficiently apart to trigger multiple detections of the same cell. However, it's important to clarify that trophoblast giant cells typically undergo a process known as endoreduplication, resulting in a single, large polyploid nucleus rather than multiple distinct nuclei. In contrast, syncytiotrophoblasts, which line the placental villi, contain multiple nuclei due to their multinucleated structure. This distinction is crucial for accurate image analysis and nuclei detection. By understanding the unique cellular architectures of these trophoblast subtypes, our clustering algorithm, utilizing a KD-Tree-based approach, can more

accurately differentiate and identify single cells despite their size or the density of nuclear material."

### 5.4.5 Nuclei Detection Model Training

The task of nuclei detection is addressed by utilising an object detection strategy, specifically RetinaNet, a single-shot object detection approach developed by Facebook AI Research (FAIR) [82]. A RetinaNet model pre-trained on the COCO dataset was employed to commence this investigation.

Using a pre-trained model exemplifies a technique known in machine learning as transfer learning. Transfer learning is a research approach that involves leveraging knowledge acquired from one problem and applying it to a distinct but related problem [116]. This technique has proven particularly effective in deep learning applications related to computer vision, as images generally comprise low-level features such as lines, edges, and curves. Using transfer learning can significantly reduce training time as parameters in the early layers of the model are already proximate to their optimal state.

The feature extractor utilised for the RetinaNet model is based on a ResNet architecture. Experiments were conducted with various sizes of ResNet backbones, namely ResNet18, ResNet50, ResNet101, and ResNet152, with the number at the end denoting the number of layers. The findings from these experiments are discussed in the results section of this chapter.

The first stage of defining the model architecture involves generating the training and validation sets. The construction of these datasets is elaborated upon in section 4. 9,401 mouse nuclei were used for training, while the validation set consisted of 5,880 nuclei. All tiles were normalised for scale, considering the resolution at which the images were captured from the microscope file.

Data augmentation was applied to the training set images. Data augmentation is a technique that artificially inflates the quantity of data by creating new data points from existing ones. In computer vision, this involves making minor modifications to the images. This study applied image flips, rotations, random scaling, Gaussian

noise, blur, and stain augmentation at rates of 90% or higher to maximise the model's generalisability. The preference for stain augmentation over stain normalisation is discussed in the results.

Training and validation images were then passed to a DataLoader constructed using PyTorch, a machine-learning framework developed by FAIR. PyTorch enables fast tensor computation on GPUs and offers numerous helpful tools for simplifying the development of machine-learning models. The PyTorch DataLoader organises the data into batches. A batch represents the number of samples that the model processes before updating the model parameters via backpropagation. The batch size, a tunable hyperparameter, can impact both computational performance and model generalisation, and its effects are discussed in the introduction.

The training was conducted in two learning rate stages. The first 20 epochs were trained at a high learning rate to aid in finding the global minimum. The second stage employed a low learning rate until the validation achieved optimal mean average precision. Early stopping was utilised to maximise generalisability, with the results of these two training phases presented in the results. The Adam optimiser was used for gradient descent, which is discussed further in the introduction.

### 5.4.6 Nuclei Detection Inference

Given the considerable amount of data to be processed in nuclei detection inference, careful design considerations were imperative in processing the data through the pipeline to ensure coherent analysis.

The first aspect involved integrating the relevant image paths, placental coordinates, and metadata into the pipeline. Moreover, a structure that maintains all saved nuclei coordinates in a low-memory format was required. This was achieved using a low-memory Python object called a 'pickle' object. The pickle objects containing all the metadata were integrated, and nuclei detection was carried out. This was done in large batch sizes of 200 images to boost processing speed.

The nuclei bounding boxes were transformed into a centroid point at the centre of the bounding box around the nuclei. All of these locations were then saved. Multi-detections were removed using the described K-D tree approach. The clustered locations were saved and subsequently used for cell classification inputs.

## 5.5 Cell Classification Model Training

The challenge of cell classification is addressed using an image classification approach. These deep learning methods have become prominent in computer vision, as discussed in section 3. The specific model architecture utilised in this research is InceptionResNetV2.

The employed model was pre-trained on ImageNet, leveraging the widely adopted approach of transfer learning. As explained upon in section 3, this methodology offers the distinct advantage of significantly curtailing training time. Originally, the classification layer of the network comprised a vector of length 1001, corresponding to the number of classes in ImageNet. This was subsequently modified by excising the existing head and introducing a series of fully connected layers of decreasing length – 768, 384, and 192, respectively – culminating in an 8-length classification layer reflective of the eight distinct cell classes. Each layer in this new head is coupled with a dropout mechanism with a probability of 0.3. Dropout, a technique that mitigates overfitting by probabilistically setting weights to zero, is further disscussed in section 3.

The model was trained using a cell classification training set of 34,401 mouse cell images and a validation set comprising 6,880 mouse cell images. This dataset was randomly oversampled to ensure an even distribution across the eight classes, a technique that prevents the model from overfitting to a single class. This approach is further discussed in section 6. The training and validation sets were created using the method described in section 4, with cell labels validated by the Hemberger lab.

As with nuclei detection, it was crucial to ensure model generalisability. To this end, aggressive data augmentation was applied, including flips, rotations, stain augmentation, Gaussian blur, and random scaling.

The optimisation process was carried out using the ADAM optimiser. Although stochastic gradient descent and DFM were also tested, ADAM demonstrated superior performance. The loss function used was cross-entropy loss, the concept of which is discussed in section 3.

The model training was conducted in two phases. The first phase employed a learning rate of $1 \times 10^{-3}$ for 30 epochs, while the second phase used a learning rate of $1 \times 10^{-5}$. This learning rate was automatically decreased if the model went through ten epochs without a performance increase. This approach aimed to find the global minimum in the loss landscape with a high learning rate before moving into the global minimum in the second phase. A batch size of 150 was used for training, and early stopping was used to get the best performing model.

## 5.6 Cell Classification Inference

The pertinent pickle file is accessed, and the locations of nuclei are extracted for inference. Subsequently, a bounding box of 100x100 pixels is constructed around each nucleus, effectively encapsulating the corresponding cell. Experimental iterations were conducted using boxes ranging from 50x50 to 400x400 pixels, with the 100x100 pixel dimension emerging as the optimum. Cell images were subsequently passed through the model in batches of 200, with the resultant classifications recorded in the pickle file. Indexing was employed to ensure accurate association between each classification and its respective location.

The pickle file was additionally populated with false statements in instances where classifications were absent, thereby providing the algorithm with clear starting points in the event of process interruption. Through the implementation of such time-efficiency measures within the pipeline, it was possible to achieve inference within a timeframe of 12 minutes per placenta.

## 5.7 Results and Discussion

### 5.7.1 Investigation 1: Replicating the Ferlaino et al. results

**Aim:** The primary goal of this study was to validate and reproduce the findings presented in the Ferlaino et al. paper.

**Method:** Replication was achieved by employing the original study's published code and data, ensuring the fidelity of our methodological approach.

**Result:** Our results corroborated the original findings, with a mean average precision (mAP) of 65% for human nuclei detection and a classification accuracy of 89% across five human cell classes.

**Conclusion:** This investigation affirms the validity of the original study's results, successfully replicating its findings.

The foundation of this research is derived from the seminal work presented in the paper by Ferlaino et al. As such, the initiation of this investigation logically entailed the replication of the previously reported results. The methodology outlined in the aforementioned study leverages the capabilities of the Keras and TensorFlow libraries, providing a robust framework for constructing the pipeline. By utilising these tools, it was possible to reproduce the reported results with a mean average precision (mAP) of 65% in nuclei detection, and a 89% accuracy in classification. Crucially, the code and data from the original study were readily accessible and contained a random seed, thus facilitating a straightforward and precise replication of the results

### 5.7.2 Investigation 2: Translating human nuclei detection to mouse nuclei detection

**Aim 1:** The investigation sought to evaluate whether human nuclei detection models could be successfully applied to mouse nuclei, with the ultimate goal of developing a universal nuclei detector.

**Method 1:** The human nuclei detection models developed by Ferlaino et al. were directly applied to a sample set of annotated mouse placenta data. The composition of training and validation sets used for these experiments is detailed in Table 5.3, providing insights into the dataset's size and distribution across species.

**Result 1:** The application of human models yielded a modest mean average precision (mAP) of 1.85%, only slightly above the baseline result of 1.12% mAP obtained using the base CoCo pre-trained model. This indicated that the human models did not generalise well to the task of mouse nuclei detection. This is could be related to species differences or the fact the human placenta nuclei were captured at term where as the mouse at

stages E14,5 and E9.5

**Conclusion 1:** Direct application of human nuclei detection models to mouse nuclei detection tasks demonstrated limited efficacy, suggesting the need for model fine-tuning.

**Method 2:** The human nuclei detection models were fine-tuned with a small dataset of mouse nuclei, with the intention of adjusting the model's decision boundaries to accommodate for multi-species detection.

**Result 2:** The fine-tuned model performed poorly on a combined mouse and human validation set, indicating that it did not generalise well across species despite the increased dataset variability.

**Conclusion 2:** The endeavour to develop a multi-species nuclei detector proved to be a difficult challenge, with models demonstrating poor generalisability across different species.

**Future Work:** Future research could explore domain adaptation techniques to improve model generalisability across species. This approach could potentially enhance model performance in histopathology tasks, where staining and tissue variability often pose significant challenges.

The first step in localising and classifying cells in this decoupled approach involves localising the nuclei. This localisation of nuclei is achieved using the one-stage object detection approach, RetinaNet. The histological samples of the mouse and human placenta used in the DMDD and Ferlaino et al. papers are stained using Hematoxylin and Eosin. This staining method results in nuclei taking on a purplish-blue colour due to an oxidation reaction with the nuclear histones, while the extracellular matrix and cytoplasm are stained pink. A key question for this investigation was whether the models that detect human nuclei could be applied to detect mouse nuclei.

Annotation in computer vision is a particularly resource-intensive process. This is especially true in medical science, where expert knowledge is required for accurate annotation. The machine learning field has sought to overcome the challenges associated with the expensive annotation process by developing various techniques, such as self-supervised and unsupervised learning, which do not require labels for training. However, supervised learning remains the most effective method in

object detection. Applying transfer learning is one way to overcome the need for large amounts of expensive labels.

While Hematoxylin and Eosin staining results in nuclei taking on a purplish-blue color in both mouse and human placenta, subtle differences in the morphology and staining intensity may still exist due to species-specific cellular structures. Future studies will employ quantitative morphometric analysis to precisely characterize these differences, ensuring our detection models are robust across different biological contexts, a universal nuclei detector could be developed if this si sufficently characterised and addressed. This would save significant time in creating training data and add a dimension to the pipeline, allowing it to translate to other tissues easily.

In this investigation, the nuclei detection models developed for the human placenta by Ferlaino et al. were applied to a sample set of mouse placenta data. This dataset consisted of 1,156 mouse placenta nuclei, which were annotated. Applying the human model developed by Ferlaino et al., trained on 11,184 nuclei, yielded a 1.85% mAP. This was slightly above the baseline result of 1.12% mAP obtained using the base CoCo pre-trained model. This outcome indicates that while humans may intuitively distinguish between mouse and human nuclei, there is a significant challenge for supervised deep learning models to achieve similar levels of generalization across different species. Enhancing model adaptability to diverse biological datasets is crucial for improving their performance on varied tasks. This is widely reported in the literature, as supervised models perform well on tasks with few generic features but do not translate well to new or similar tasks.

**Fine-tuning the human models for mouse nuclei detection**

While the human models did not directly translate to the task of mouse nuclei detection, could this problem be overcome by fine-tuning the models on a small dataset of mouse nuclei? The idea here is that with a small amount of training, the decision boundaries in the model will be adjusted enough to address the issue, and a multi-species detector could be generated. This would save significant effort

**Table 5.2:** Results of multi-species nuclei detection experiments

| | Human nuclei validation | Mouse nuclei validation | Human + Mouse nuclei validation |
|---|---|---|---|
| Human nuclei training set | 65% mAP | 6% mAP | 32% mAP |
| Mouse nuclei training set | 4% mAP | 59% mAP | 25% mAP |
| Human + mouse nuclei training set | 12% mAP | 7% mAP | 18% mAP |

*The table presents the results of multi-species nuclei detection experiments. Each entry represents the mean average precision (mAP) obtained by applying the trained models to different validation sets. The results demonstrate the challenge of generalising models across different species, even with training on mixed-species datasets.*

in creating training data and add a dimension to the pipeline, allowing it to translate easily to other tissues.

The investigation into generating a multi-species nuclei detector began with the development of three models:

1. Train a new Retinanet human model using the dataset from Ferlaino et al., consisting of 11,184 human nuclei for 100 epochs.

2. Train a Retinanet pre-trained model based on mouse nuclei data. For this model, a dataset consisting of 2,258 mouse nuclei was annotated and trained the model for 100 epochs.

3. Train a Retinanet pre-trained model using mouse and human nuclei data for 100 epochs.

After all three models had converged, they were evaluated on three validation sets: one consisting of only human nuclei, one of only mouse nuclei, and one combining human and mouse nuclei. The results of this experiment are shown in Table 5.2.

The results show that when the human model is applied to human data, it performs well, and the same is true when the mouse nuclei detection model is applied to mouse data. However, neither model performed particularly well on the combined mouse and human validation set. This is likely because each model is specialised to one species, and it is widely reported in machine learning that models translate poorly to new tasks or similar ones. However, the limited genralisability

**Table 5.3:** Composition of training and validation sets

| | Training Set | Validation Set |
|---|---|---|
| **Human** | 11,184 nuclei | 876 nuclei |
| **Mouse** | 2,258 nuclei | 851 nuclei |
| **Human + Mouse (50/50)** | 4,516 nuclei | 1,702 nuclei |

*The table outlines the composition of the training and validation sets used in the multi-species nuclei detection experiments. These sets span human, mouse, and combined human and mouse nuclei.*

was noted. This factor will be considered and addressed in the next section of the chapter. As discussed, histopathology tissues and staining are highly variable; therefore, a model that does not translate well will not be effective in the long run.

Given the challenges observed with the current models' performance across different species, future investigations will focus on enhancing the adaptability and accuracy of our detection systems. We will explore domain adaptation techniques to better align the model trained on human data for effective use with mouse nuclei, potentially expanding this approach to other species as well. Furthermore, implementing advanced machine learning strategies such as semi-supervised and few-shot learning could significantly reduce the dependency on large labeled datasets. We also plan to refine our models to improve confidence in their predictions across varied histological samples. This will involve a thorough analysis of model confidence intervals and potentially the development of a multi-output model that can discriminate between human nuclei, mouse nuclei, and background. These efforts are aimed at creating a more versatile and reliable tool for histopathological analysis, capable of handling the inherent variability in staining and tissue structures.

The result for the model trained for nuclei detection using both training sets combined was the least reliable. This is surprising as it was assumed that having a more varied dataset at training would expand the decision boundary. However, it appears that training with a mixed-species dataset does not automatically improve model performance and might even undermine network confidence. This observation

challenges our assumptions and suggests a complexity in how different biological data interact within machine learning models.

A potential solution could be to develop a model that categorizes nuclei more distinctly—by outputting human nuclei, mouse nuclei, or background. This nuanced approach could potentially enhance model accuracy by allowing it to specialize its predictions based on clearer categories.

The composition of training and validation sets used for these experiments is detailed in Table 5.3

Further investigations will explore how domain adaptation techniques can be employed to improve these results. Adjusting models to perform effectively across different biological domains could reduce the need for extensive retraining and enhance generalizability. This approach will be particularly beneficial in histopathology, where variability in staining and tissue structure poses significant challenges. Incorporating advanced machine learning strategies, such as semi-supervised and few-shot learning, could also provide pathways to leverage limited data more effectively.

### 5.7.3 Investigation 3: Generating a high-performance mouse nuclei detector

**Aim:** The investigation aimed to enhance the generalisability of a mouse nuclei detector and improve its performance.

**Method:** The method involved redeveloping the original Ferlaino et al. model in PyTorch and enhancing its performance through extensive hyperparameter experimentation and aggressive data augmentation. A RetinaNet object detection architecture with a ResNet101 backbone was employed, and stain transformation techniques were applied to improve model robustness against stain variance.

**Result:** The newly developed method resulted in a significant improvement in mouse nuclei detection with an 88% mAP, and an enhancement in human nuclei detection to 86% mAP.

**Conclusion:** The investigation concluded that redeveloping the method in PyTorch, applying aggressive augmentation and stain transformation techniques, and conducting extensive hyperparameter experimentation significantly improved nuclei detection performance.

**Future Work:** Future research could explore additional augmentation techniques, such as random erasing and image mixing, to further improve model generalisability. The application of generative adversarial networks for style augmentation could be a promising avenue to explore.

The research conducted by Ferlaino et al. developed a method that could detect nuclei across a single slide with an mAP of 65% and perform cell classification across five human cell types with 89% accuracy. However, this method could have generalised better across tissue regions or to different samples. Moreover, the method only worked on single tiles, and to create a universal functional approach, a global whole slide tissue processing approach had to be developed.

To address these limitations, the method was redeveloped in the PyTorch library rather than its original format in Keras and TensorFlow. Keras and PyTorch are both powerful open-source machine-learning libraries. Keras is easier to use, while PyTorch offers faster performance, greater customisability, and more substantial community support and development. These advantages have made PyTorch the go-to machine learning library, making it a good choice for redeveloping the method. It was possible to significantly improve the results by doing so and addressing many underlying issues.

In this investigation, the training set used to train models consisted of a carefully curated collection of 29,401 annotated mouse nuclei images for training. These images were selected to represent a broad spectrum of staining variations and cellular conditions found in the mouse placenta, ensuring the model learns to recognize a diverse array of nuclei characteristics and is therefore as generalisable as possible. The validation set comprised 5,880 mouse nuclei images, which were used to evaluate the model's performance and ensure it could accurately generalize to new, unseen data. This segregation of data into distinct training and validation sets allows for a robust assessment of the model's predictive capabilities.

After implementing the code for nuclei detection, experiments varying the model backbone were conducted. The RetinaNet architecture consists of a model backbone for feature extraction, a Feature Pyramid Network for addressing issues related to object scale, and a classification and regression head for object detection (Figure

**Table 5.4:** Mean Average Precision of Different ResNet Architectures

| Model | Mean Average Precision |
|---|---|
| **ResNet18** | 76% |
| **ResNet50** | 81% |
| **ResNet101** | 84% |
| **ResNet152** | 85% |

*The table outlines the mean average precision of four different ResNet architectures.*

5.3). A standard and robust model backbone is a ResNet backbone. ResNets are discussed in section 3. These backbones come with 18, 50, 101, and 152 layers. Theoretically, more layers should result in better results; however, training and inference times will be longer as the number of parameters in the model increases. Another area for improvement with larger networks is overfitting the training dataset. To overcome this, a robust regularisation strategy must be implemented.

In the experiments, the best-performing model was the one with a ResNet152, as expected based on the literature. The results of the experiments are shown in Table 5.4. Despite the 152 models having the best performance, the gain between 152 and 101 was minimal. Therefore, the 101 models were implemented as good performance could be achieved with faster inference and training times. All model backbones used were pre-trained on ImageNet.

In the two-stage training process, experiments with learning rates were conducted. Initially, both high and low learning rates were tested, followed by trials with cyclical learning rates. The optimal learning rate was found to be $1 \times 10^{-5}$. Various optimisation algorithms were explored, including stochastic gradient descent, but the ADAM optimiser proved superior in performance.

To ensure excellent model generalisability, aggressive data augmentation was applied. These included flips, rotations, blur, and random scaling. One major issue with generalisability in histopathology data is variation in stain depth. The extent of that variation in the DMDD dataset is demonstrated in Figure 5.9. Some of this variability is combatted with how the training data was created. In developing the training dataset for the DMDD mouse model, it was ensured that data was labelled

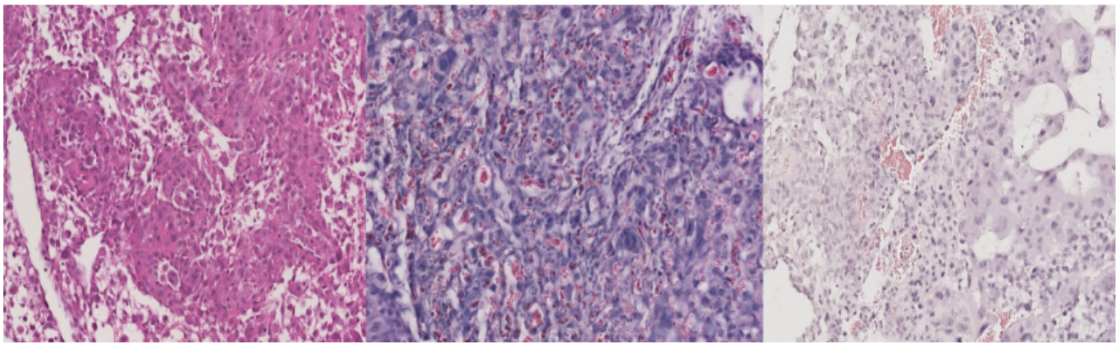


**Figure 5.9:** *This figure presents a visual demonstration of stain depth variation in the DMDD dataset, focusing on three representative histology tiles. Each tile represents a different stain depth, demonstrating the variance within the dataset. These variations in stain depth can significantly influence the subsequent image analysis and interpretation, underscoring the importance of factoring in such variations in the development of robust image analysis algorithms.*

from a large random sample and intentionally ensured the data had significant variation in stain depth. To ensure generalisability, however, a more comprehensive approach must be applied. To overcome stain variability, two main approaches exist: stain normalisation and stain transformation.

The most common approach used in automated histology analysis is stain normalisation. This approach minimises H&E stain variability across different digital scanners and laboratory protocols. Most methods involve the subjective selection of "reference" images and matrix decomposition approaches such as non-negative matrix factorisation (NMF). However, this makes scaling such approaches infeasible, particularly for whole slide image analyses.

In this approach to automated histology analysis, a "stain transformations' approach is implemented to be robust against stain variance. This approach converts RGB images into the HED colour scheme using colour deconvolution. Once in the HED space, images are randomly scaled by a factor uniformly sampled from the interval (0.5, 1.45). This scaling simulates either under or over-staining of the tissue, thus helping the model to be as generalisable as possible. The limits of the stain transformation were set via experiments to see the limits of the stain transformation. To do this, the variance was observed and then the

model was trained. The values used were the most extreme transformations, which still yielded good validation results.

The augmentation approaches used in this chapter successfully improves model performance and generalisability. However, there are other approaches which could be explored. Data Augmentation can be divided into five main categories: geometric transformations, colour space transformations, random erasing, kernel filters, and mixing images. Geometric transformations such as flips, crops, and rotations are employed in this work. Also, colour space transformations, such as random brightness and hue saturation, and kernel transformations, such as Gaussian blur, are used. Random erasing is an interesting avenue that has been successful in computer vision applications. This method removes regions of the images, encouraging the model to pay attention to the entire image rather than just a few features. This method is likely to produce great generalasbility due to the focus beiung on multiple cellular features. An effective augmentation technique is image mixing. CutMix, for example, takes patches from one image and inserts them into another. This encourages the model to rely less on specific parts of the object in the image. A more advanced image-mixing technique is style augmentation, which uses generative adversarial networks to mix images. This could create unlimited images, but it is unclear how useful this is since no new data is being added. It, however, would be a promising avenue for future investigation.

The best result for the mouse nuclei detection models after many iterations of training and experimentation with various hyperparameters was an mAP of 87% on the validation set. To achieve this result, bootstrapping was used where poor performance was viewed.

The best result for the human nuclei detection models was 88% mAP after many iterations of training and experimentation with hyperparameters. This was achieved using the same training and validation sets used in the paper. The result achieved using the approach developed significantly improved from the 65% mAP achieved in the paper. This is a testament to the time spent fixing all the bugs in the code base in the rewrite, additional machine learning methods

being applied, and more rigorous experimentation in training. The change from TensorFlow and Keras to PyTorch may have affected the results, but this will not have been significant. It is reasoned that RetinaNet has been implemented in both libraries, and the results are comparable.

## 5.8 Investigation 4: Generalisable Whole Slide Nuclei Detector for Mouse Placenta

**Aim:** The aim was to enhance the generalisability of the nuclei detection model across varying tissue samples and regions within the DMDD dataset.

**Method:** The method involved deploying the mouse nuclei detection model across fifty random tissue samples from the DMDD dataset. A quality control process was initiated involving extensive evaluation, bootstrapping, hard negative mining, and the removal of empty tiles to improve computational efficiency.

**Result:** The implementation of these enhancements led to a significant reduction in false positive detections and a more efficient detection process. The results were validated by successfully overlaying predicted nuclei locations onto whole slide images, demonstrating consistent performance.

**Conclusion:** The conclusion of this investigation was that the refined mouse nuclei detection model, complemented by a robust quality control process, can perform consistently across varying tissue samples and regions, making it an effective tool for the analysis of the DMDD dataset.

**Future Work:** Future work could explore the incorporation of additional negative examples in the training set or utilise more sophisticated techniques to further reduce false positive detections and enhance model performance.

Previous investigations detected high-performance nuclei in both mouse and human placenta. This component is a crucial part of the end-to-end whole slide placenta analysis pipeline. However, given the high variability of histology datasets, deep learning models often require enhancements for generalisability. The principal focus of this investigation was to deploy the analysis across the entire DMDD dataset and achieve consistent performance across varying tissue samples and tissue regions. To ensure this, extensive quality control was performed to confirm that the model and surrounding code would perform well across different samples.

In order to ensure the model performed well across the dataset, the mouse nuclei detection model was deployed across fifty random tissue samples from the DMDD dataset. The locations of the nuclei were reconstructed into a scatter plot to ensure even performance. This allowed for a comparison between the whole slide image and the predicted locations. A method was also developed to overlay the nuclei locations on the whole slide image. This facilitated a clear understanding of performance and allowed for reannotation and addition to the training set where necessary (bootstrapping).

After resolving some initial bugs with the placenta reconstructions from the tiles used in inference, one major non-code-based issue persisted. The model predicts nuclei in empty tiles uniformly (Figure 5.10). When the model was presented with an empty tile, it predicted three random boxes of variable size around the centre. The explanation for these predictions is that the training set contained no negative examples. Therefore, the model assumed every tile would have nuclei and predicted three bounding boxes at random. To overcome this issue, hard negative mining was performed. Hard negative mining involves taking a falsely detected object, explicitly creating a negative example out of that object, and adding that negative to the training set. When the model is retrained, it should perform better with this additional knowledge and not make as many false positive detections. Five hundred empty tiles across various placenta samples were added for the mouse nuclei detection model and the model was retrained. This significantly improved the false positive detections.

The whole slide image of the DMDD mouse placenta samples contains many empty tiles. Passing these through the model for inference requires substantial computational resources. In the method developed in this chapter, a computationally efficient big data approach was sought. Therefore, an attempt was made to remove empty tiles from being passed into the model in preprocessing. To achieve this objective, a method was developed to remove any white, black, or grey tiles. For all white, tiles with all pixels at 255 were removed. For all black, tiles with all pixels at 0 were removed. To remove the in-between tiles containing no tissue, a method

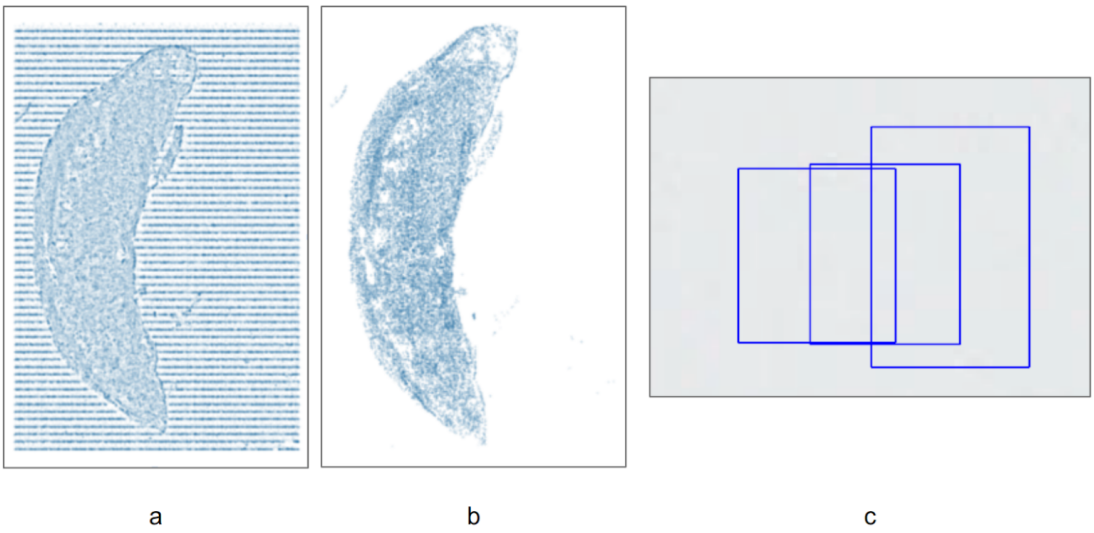



**Figure 5.10:** *This figure presents a three-part illustration of the progression in improving nuclei detection in placental images. Image a) depicts initial detections on black tiles, resulting in erroneous nuclei depictions in empty spaces. Image b) demonstrates the improvement achieved by preventing detection in empty tiles and implementing hard negative mining, which considerably reduces detections in empty tiles and provides a more accurate representation of nuclei in the reconstructed placenta. Finally, Image c) showcases the detections on the negative tiles with three bounding boxes of varying sizes located in the centre of the image.*

was developed to remove a tile where the ratio between the darkest and the lightest pixel was greater than 0.95. These approaches significantly reduced the number of tiles passed to the model while producing excellent results (Figure 5.10).

## 5.9 Investigation 5: Removing Multiple Detections of Identical Nuclei

**Aim:** The aim was to tackle the issue of duplicate nuclei detection resulting from the use of overlapping tiles, implemented to capture edge cases.

**Method:** A clustering approach, based on K-D tree method, was developed to consolidate multiple detections of the same nucleus into a single detection point. This method determined the centre of the detected nuclei within a specific radius, which represented the official location of the nucleus.

**Result:** In the DMDD mouse data, a radius of 11 pixels was identified as the optimal distance to aggregate multiple detections into a single nucleus, based on the analysis of nearest neighbours of every nucleus.

**Conclusion:** The K-D tree-based clustering approach effectively resolved

the issue of multiple detections, consolidating them into single nucleus representations. The optimal radius parameter, while dependent on specific data, was found to be approximately 11 pixels for the DMDD mouse data.

**Future Work:** Future studies could look into the refinement of the clustering radius in relation to varying cell sizes and different datasets. This would enhance the adaptability and accuracy of the approach.

The whole slide image is fragmented into smaller sub-images, or tiles, due to its substantial size. This fragmentation can result in the omission of cells or nuclei at the edge of each tile. When multiplied over every tile in the placental section, this can lead to a significant error in coverage. To mitigate this issue, overlapping tiles are created with 50% coverage. This approach ensures the capture of edge cases and provides the model with multiple opportunities to detect all the nuclei in the placenta. However, the use of overlapping tiles introduces an additional problem: multi-detections of identical nuclei.

To overcome the multi-detection issue, a clustering-based approach was developed to remove multiple nuclei detections within the same cell. This approach is based on the K-D tree clustering method. This method identifies the nearest neighbours within a distance of $x$, which translates to the distance between cells. The method then identifies the centre point of all the nuclei detections in the cell, which is saved as the location of the nucleus.

The challenge then becomes determining the value of distance $x$. For this purpose, the seven nearest neighbours of every nucleus were identified, and the number of nuclei at each distance was plotted, as shown in Figure 5.11. This figure illustrates that nuclei within the same cell, which would represent multiple detections of the same nucleus, will be very close together. There will then be a threshold that provides an approximate cell radius. In the DMDD E14.5 placental data, this cell radius occurs at 11 pixels, as shown in Figure 5.11. This radius was validated across multiple placental sections to ensure consistency across the dataset. Despite the varying cell sizes, it is a good approximation for clustering multiple detections into a single one.

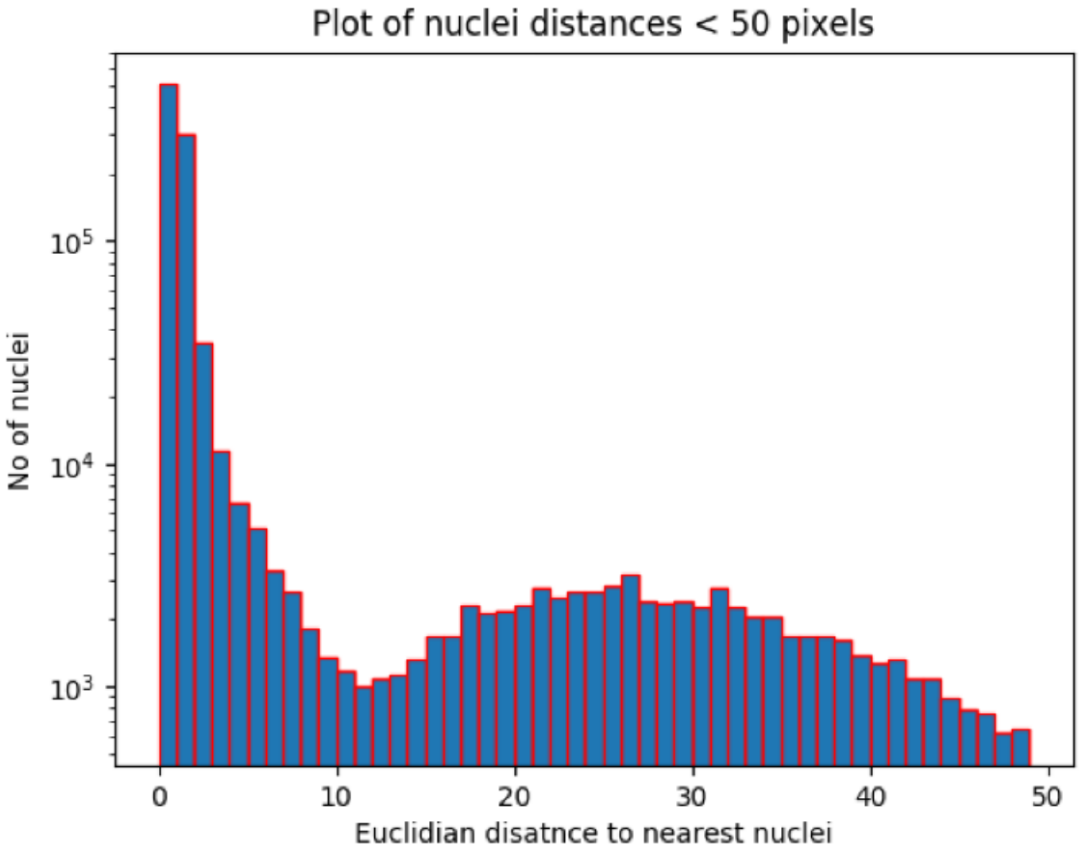



**Figure 5.11:** *This figure presents a crucial step in refining the nuclei detection process in placental images. The figure illustrates how the problem of multi-detections of the same nucleus, a consequence of overlapping tiles, is overcome using a K-D tree clustering approach. This approach identifies the nearest neighbours within a certain distance (denoted as x), which corresponds to the distance between cells. The figure also demonstrates how the optimal distance x is determined. The histogram shows the number of nearest neighbours at each distance, revealing a threshold distance that approximates a cell's radius. For the DMDD E14.5 placental data, this distance is found to be 11 pixels, as indicated in the figure. This calculated radius, verified across multiple placental sections, effectively clusters multiple detections into a single, accurate detection of a nucleus.*

In this study, post-processing was deemed unnecessary due to the effectiveness of our initial detection methods. We employed bootstrapping and hard negative mining techniques to enhance the accuracy of our nuclei detection algorithm. These methods proved highly effective in ensuring precise detection, and as a result, instances of multiple detections within the same cell were exceedingly rare during our quality control checks. This outcome is particularly significant given the inherent challenges associated with detecting nuclei in syncytiotrophoblasts (STB) and glycogen cells, where nuclei are either multiple or exceptionally small and faintly stained.

## 5.10 Investigation 6: Comprehensive Nuclei Detection in the DMDD Dataset

**Aim:** This investigation aimed to validate the performance of the developed nuclei detector across the DMDD dataset, focusing on its ability to accurately identify nuclei positions within various tissue samples and regions using detailed qualitative assessments.

**Method:** The validation process utilized two primary qualitative methods. First, the mouse nuclei detection model was applied to fifty random tissue samples, with a focus on detailed tile-level analysis to ensure accurate detection of nuclei within specific cells. Each tile was visually inspected to compare the model's predicted nuclei locations with actual nuclei positions within the cells. Second, for a holistic validation, scatter plots of detected nuclei locations for entire placental slides were generated and compared against corresponding whole slide images. This approach allowed for assessing the spatial accuracy of nuclei detection across the complete tissue sample.

**Result:** The tile-level inspections revealed a high accuracy in nuclei detection, corroborating the model's effectiveness in identifying and positioning nuclei within individual cells across diverse histological environments. The whole placenta scatter plot analyses, illustrated in Figures 5.12 and 5.13, demonstrated that the detected nuclei appropriately corresponded to actual histological features across various regions of the placenta, including the junctional zone and labyrinth. Challenges were identified in regions with dense nuclei packing, which were addressed through iterative model tuning.

**Conclusion:** The comprehensive qualitative assessment methods confirmed the effectiveness of the nuclei detection model, demonstrating its robustness and adaptability to the complex histological variability within the DMDD dataset. The model's capability to accurately reflect true nuclei locations both at the tile and whole placenta levels underscores its potential for broad histopathological applications.

**Future Work:** Future enhancements will focus on refining the detection algorithms to improve accuracy in challenging regions with densely packed nuclei and exploring advanced visualization techniques to further validate the spatial accuracy of nuclei detection.

This phase of evaluation used both detailed tile-level analysis and broader scatter plot comparisons to ensure the model's accuracy in nuclei detection. The method of overlaying predicted nuclei locations on whole slide images, combined with scatter plot analysis, provided a dual approach to verify and refine the model's performance.

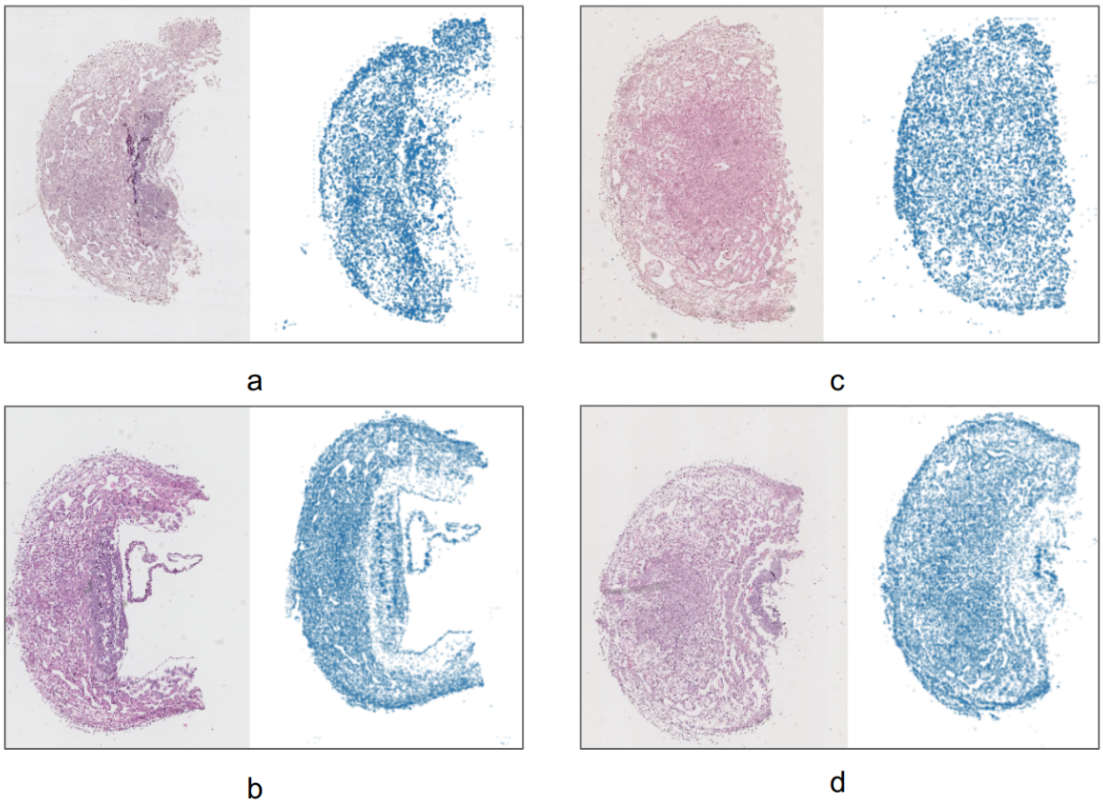



**Figure 5.12:** *This figure presents examples of nuclei detection on E9.5 placental samples, including both wild type and mutant samples. The overlays on the histology images directly compare actual placental histology with the nuclei detection results, highlighting the model's performance.*

These techniques were particularly effective in confirming the spatial accuracy of detected nuclei across different tissue samples, highlighting both successes and areas for improvement.

## 5.11 Investigation 7: Human Placenta Cell Classification Model

**Aim:** The focus of this investigation was to improve upon the cell classification model presented by Ferlaino et al., which distinguished among five human placenta cell types with an accuracy of 89%.

**Method:** The original model was transitioned from Keras to PyTorch for greater customizability and efficiency. Enhancements included the application of more aggressive data augmentation and the use of random oversampling to balance the training dataset.

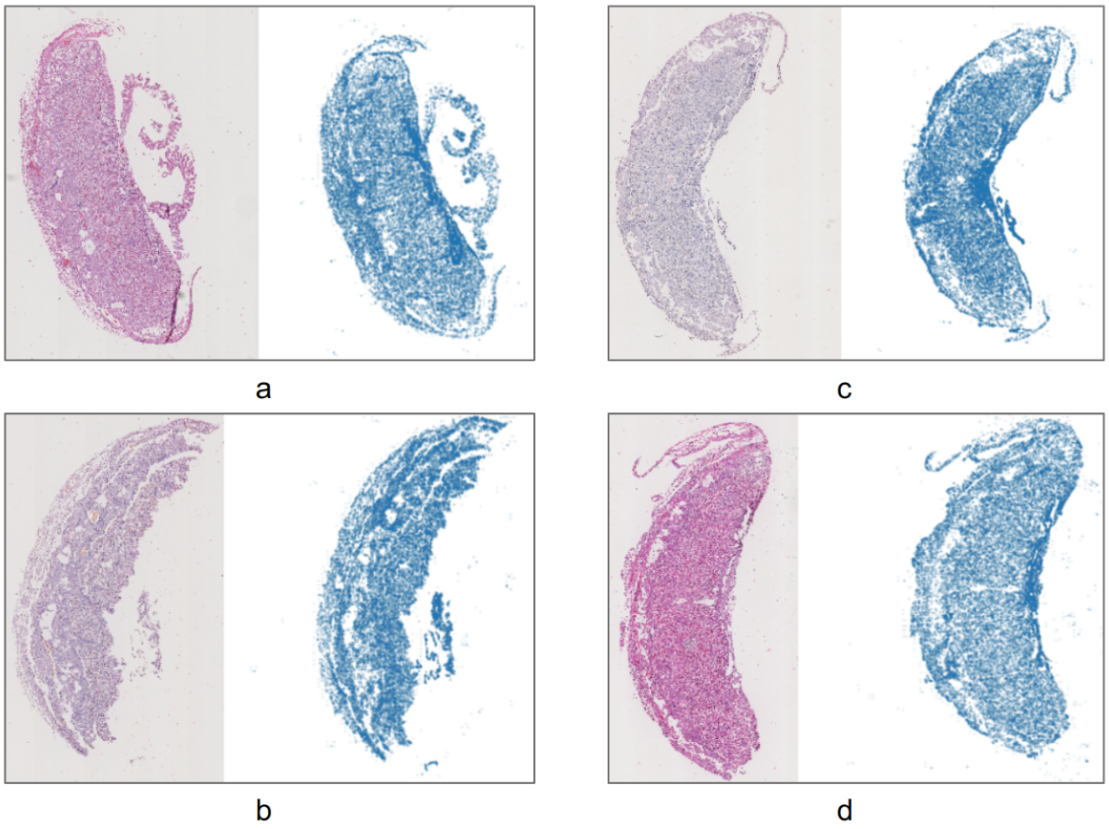



**Figure 5.13:** *This figure shows nuclei detection on E14.5 placental samples, encompassing both wild type and mutant variants. The illustrated overlays provide a comparative view of the actual histology and the nuclei detection outcomes.*

**Result:** Through these modifications, the classification accuracy was boosted to 90%, outperforming the original model. A detailed visualization of the model's performance across the different cell types is shown in Figure 5.14, which provides detailed insights into the model's performance across different cell types.

**Conclusion:** The redevelopment of the original model in PyTorch, complemented by enhanced data augmentation and oversampling, resulted in an improved cell classification model with superior accuracy. Figure 5.15 further illustrates the practical application of this model on a sample human tile, highlighting its operational effectiveness in identifying and classifying cell types within a complex histological context.

**Future Work:** Future investigations could delve into the implementation of more advanced machine learning techniques or the exploration of alternative architectures to further improve the model's performance.

In the initial phase of research into cell classification, the primary objective was to replicate and enhance the results obtained by Ferlaino et al., who achieved an

accuracy of 89% in classifying five human placenta cell types. This goal was pursued by transitioning the model from the Keras framework to PyTorch. Additionally, the methodology incorporated an increased level of data augmentation to provide the model with a more diverse range of training examples, thereby improving its ability to generalize. A crucial aspect of this enhancement was the implementation of random oversampling, which is detailed in the confusion matrix (Figure 5.14).

Random oversampling is a technique used to address imbalances in the training dataset. In the context of cell classification, certain cell types might be underrepresented in the available data. To mitigate this, random oversampling involves artificially increasing the presence of these underrepresented cell types in the training set. This is typically achieved by duplicating existing samples of these cell types or generating new, synthetic examples through various means. By employing this strategy, the model is exposed to a more balanced representation of all cell types, which helps in improving its accuracy and robustness in classification tasks. Consequently, these methodological advancements led to an improvement in the classification accuracy, raising it to 90%.

## 5.12 Investigation 8: Mouse Placenta Cell Classification Model

**Aim:** The aim was to develop a robust mouse cell classification model, capable of accurately distinguishing among the eight annotated cell types in the DMDD dataset, with a focus on the E14.5 developmental stage.

**Method:** Various image classification architectures were tested, including InceptionResnetV2, Xception, InceptionV4, and ResNext. The Inception-ResNetV2 model, which processed 100x100 pixel images centered around the nucleus, outperformed the others and was selected for further use.

**Result:** The InceptionResNetV2 model achieved an impressive classification accuracy of 90% across eight cell classes for the E14.5 placenta samples. As no annotations were available for E9.5, the discussion for this developmental stage is purely qualitative, focusing on the visual assessment and phenotypic descriptions observed in the E9.5 tissue samples.

**Conclusion:** A high-performing cell classification model was developed for the E14.5 placenta. The lack of annotated data for E9.5 necessitates

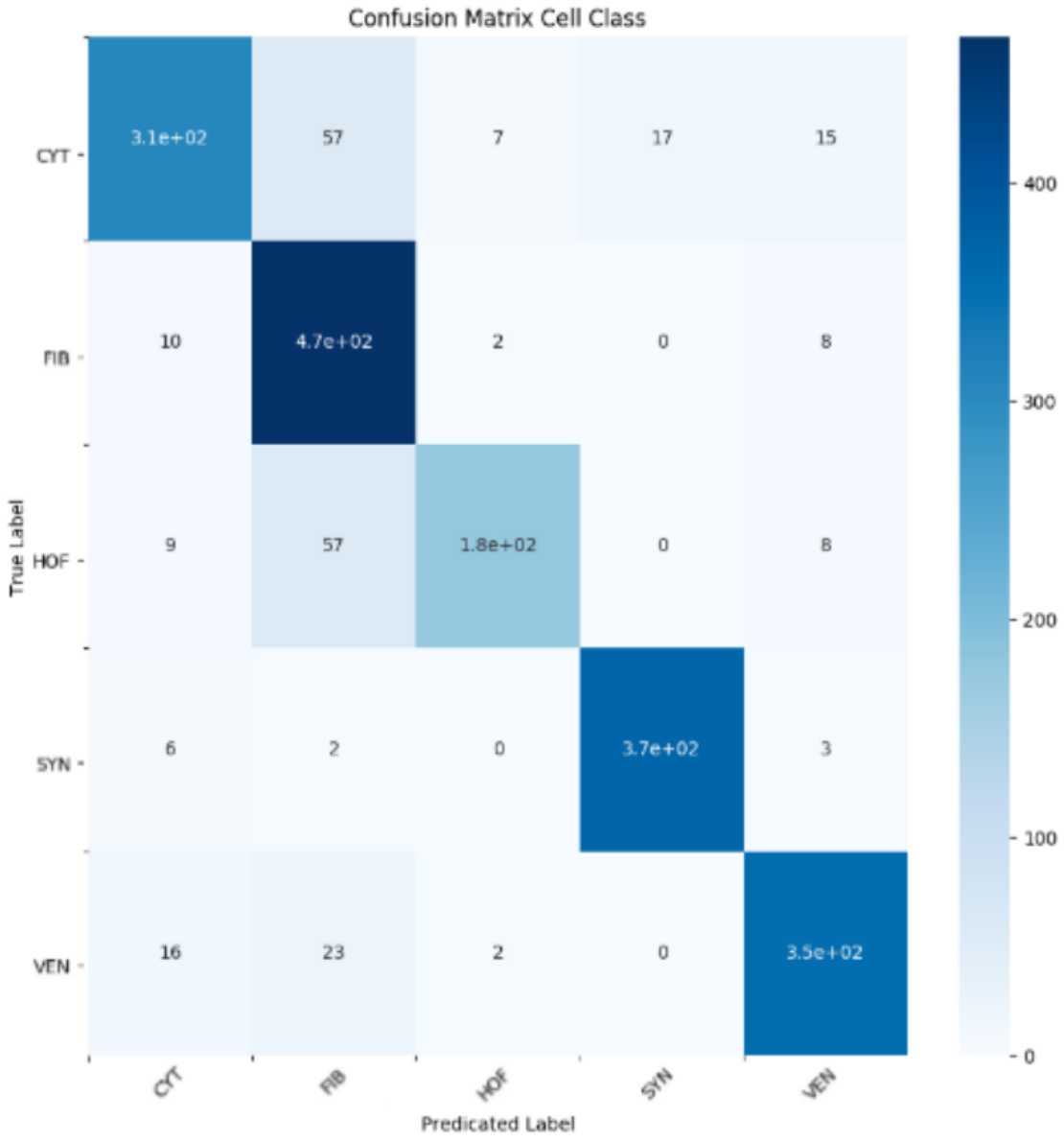



**Figure 5.14:** *This figure presents a confusion matrix for the classification of five human cell types: cytotrophoblast (CYT), fibroblast (FIB), Hofbauer (HOF), syncytiotrophoblast (SYN), and vascular (VAS). The matrix provides a visual representation of the performance of the classification model, highlighting the number of correct and incorrect predictions for each cell type.*

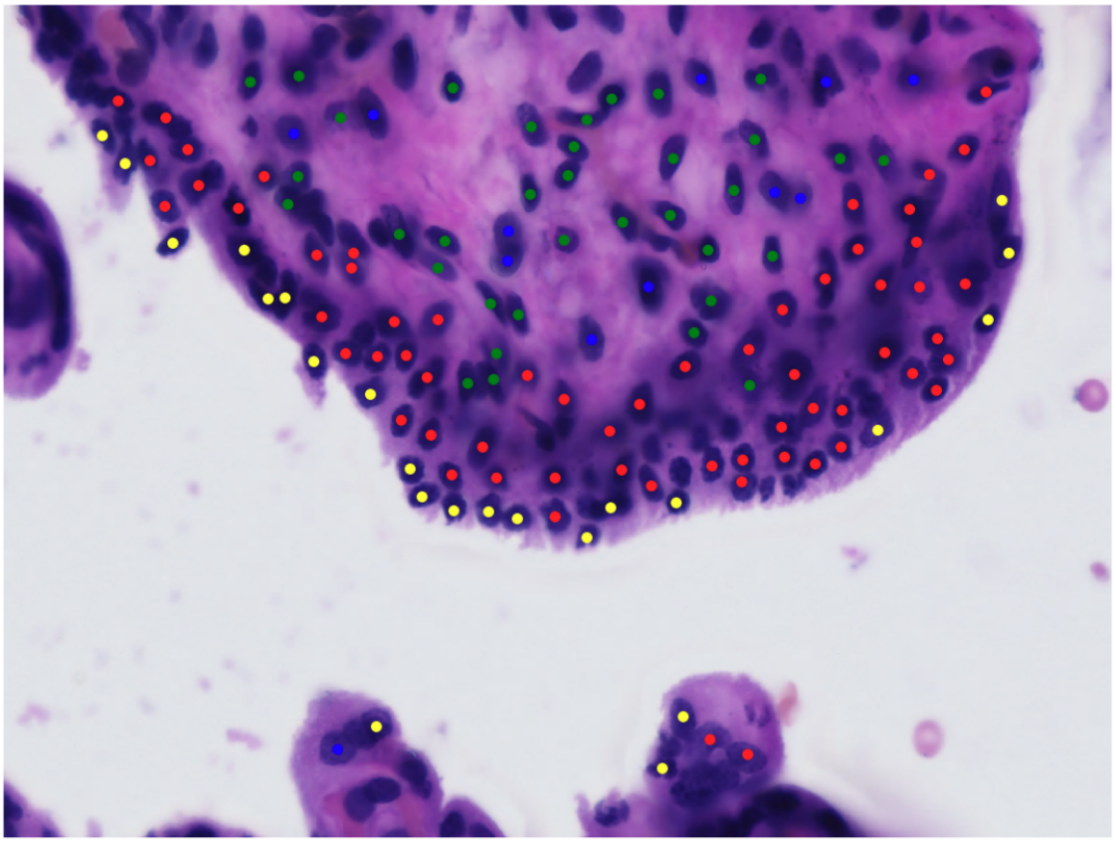


**Figure 5.15:** *This figure displays a sample human tile upon which the human model has been deployed. The different human cell types are colour-coded for clarity: cytotrophoblast (CYT) is highlighted in red, fibroblast (FIB) in green, Hofbauer (HOF) in blue, and syncytiotrophoblast (SYN) in yellow. This visual representation allows for an intuitive understanding of the cell type distribution within the sample tile.*

a qualitative approach to evaluating model performance at this stage, highlighting the need for annotated data in future studies.

**Future Work:** Future research could focus on expanding the annotated E9.5 cell class dataset to improve the model's performance and generalizability for the E9.5 tissue. Additionally, exploring the impact of varying developmental stages on classification accuracy could offer deeper insights into the model's adaptability.

The methodological approach adopted for classifying mouse cells in this study is grounded on the image classification methodologies examined in section 3. The experimentation phase involved four distinct architectures: InceptionResnetV2, Xception, InceptionV4, and ResNext. These models were selected due to their superior performance in the ImageNet Challenge, an influential competition in the field of image classification and computer vision. The high performance of these models

in the ImageNet Challenge suggested their robustness and reliability in classifying diverse categories of images, making them suitable candidates for this study.

Utilizing the capabilities of the PyTorch libraries, the training of these models was facilitated. Each model had been pretrained on the ImageNet dataset, providing a solid foundation for further refinement based on our specific requirements. The head of the model was subsequently excised, and three fully connected layers of length 768, 382, and 192 were appended in sequence. Dropout was applied in each layer with a probability of 0.3. Dropout is a technique used to prevent overfitting by randomly ignoring some neurons during the training process, thus forcing the network to learn features in a distributed way.

The final classification layer was tailored to be of length 8, reflecting the eight classes under consideration. The configuration of this layer was dictated by the specific requirements of our study, wherein we aimed to distinguish between eight distinct classes of cells.

The model training was executed in two distinct stages, a process designed to prevent overfitting and improve generalisation. In the first stage, a cycle of 20 epochs was run, during which only the final four layers were updated, using a relatively high learning rate of $1 \times 10^{-3}$. This stage allowed for rapid learning of the task-specific features in the newly added layers. Subsequently, all parameters were unfrozen for training at a reduced learning rate of $1 \times 10^{-5}$. This second stage allowed the entire model to fine-tune and converge to a solution that balanced the prelearned features and the task-specific features.

The learning rate was progressively decreased if no improvement in performance was observed within five epochs, a strategy known as learning rate decay. This technique is often employed to ensure the model converges to a minimum of the loss function. The selected loss function for this study was CrossEntropy, while the optimizer employed was ADAM, a popular and effective gradient descent optimization algorithm.

The input images were squares of 100x100 pixels, centered around the nucleus. To ensure that all input images conformed to the same scale, a standardization process, known as scale normalization, was applied.

Additionally, an aggressive data augmentation strategy was implemented, involving techniques such as rotation, scaling, and flipping of images. This approach, identical to the one adopted for nuclei detection, is a proven method to enhance the model's ability to generalize to unseen data and prevent overfitting

To prevent any class from being under-represented, random oversampling was also employed. This technique is particularly effective when dealing with imbalanced datasets and helps the model learn features from minority classes more effectively.

Training of the model was executed utilising 25 workers on a pair of Nvidia GPUs, with a batch size of 100 images. Each model underwent a training cycle of 400 epochs. After five iterations of each model and subsequent tuning of hyperparameters, the InceptionResNetV2 emerged as the most efficacious model, suggesting it had the optimal balance of complexity and predictive power for this task. Consequently, this model was selected for further progression.

Another facet of our experimentation involved the manipulation of the input image size. The dimensions of the input image dictate the degree of tissue context exposed to the model. The hypothesis driving this aspect of the study was that reducing the tissue context might help the model focus more on the individual cells, thereby improving its ability to distinguish between different cell types.

The results of these experiments, as depicted in Figure 5.16, demonstrated that the model's performance remained consistent even with an input size of 100x100 pixels. This finding suggests that the model was capable of classifying cells accurately based primarily on the cell features rather than the broader tissue context. Thus, 100x100 pixels was chosen as the preferred size for processing the entire DMDD dataset.

The final model selected for deployment was trained on 100x100 pixel cell boxes, achieving accuracy of 90% across the eight cell classes. A visual representation of the model's performance, the confusion matrix, is depicted in Figure 5.17. It

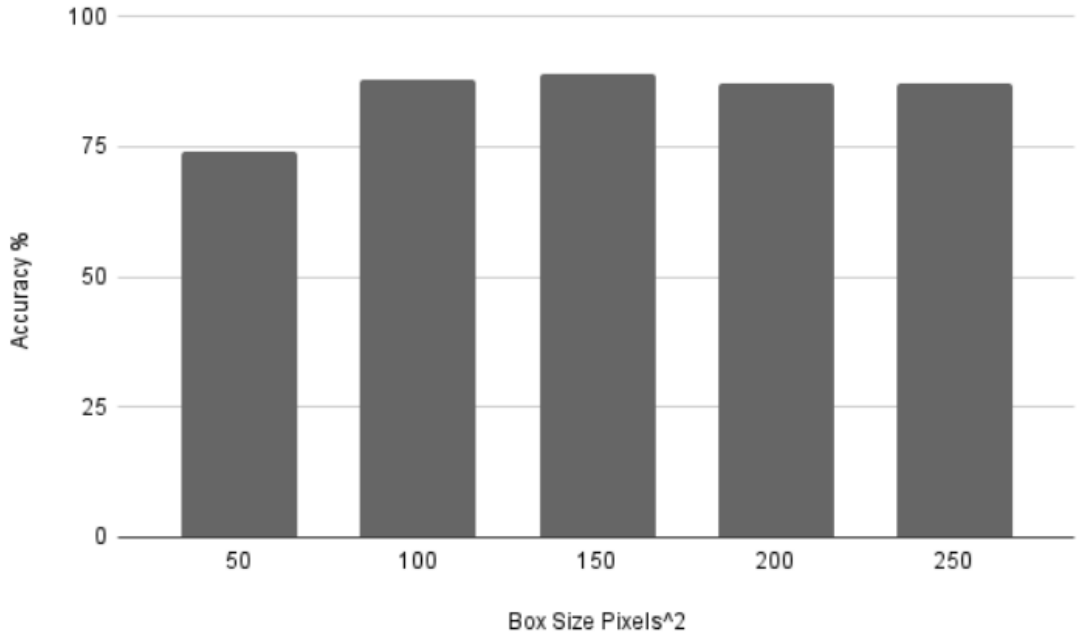



**Figure 5.16:** *This figure illustrates the relationship between the size of the square box encapsulating the cell (x-axis) and its corresponding impact on the model performance (y-axis). The size of the box is representative of the amount of tissue context available for the model. Larger boxes provide more tissue context, which could potentially aid in more accurate cell classification. Conversely, smaller boxes restrict the model's view to the cell only, limiting the surrounding tissue context. This relationship between box size and model performance is crucial in understanding the optimal context required for accurate cell classification.*

illustrates that decidual cells were frequently misclassified as chorionic plate cells, a result that can be attributed to their visual similarities.

While these misclassifications represent a limitation of the current study, potential solutions could include incorporating additional training data or further annotating cell classes to refine the granularity of the analysis. However, these issues did not significantly impact the overall results, suggesting that the model was robust to such misclassifications.

The overall effectiveness and generalizability of the model are demonstrated in Figure 5.18, where we observe cells accurately located within the expected regions at E14.5. This holds true across the majority of the dataset, attesting to the model's ability to classify cells in a variety of contexts accurately. Additionally, Figure 5.19 provides plots of cell type densities, serving as a reasonable estimate for

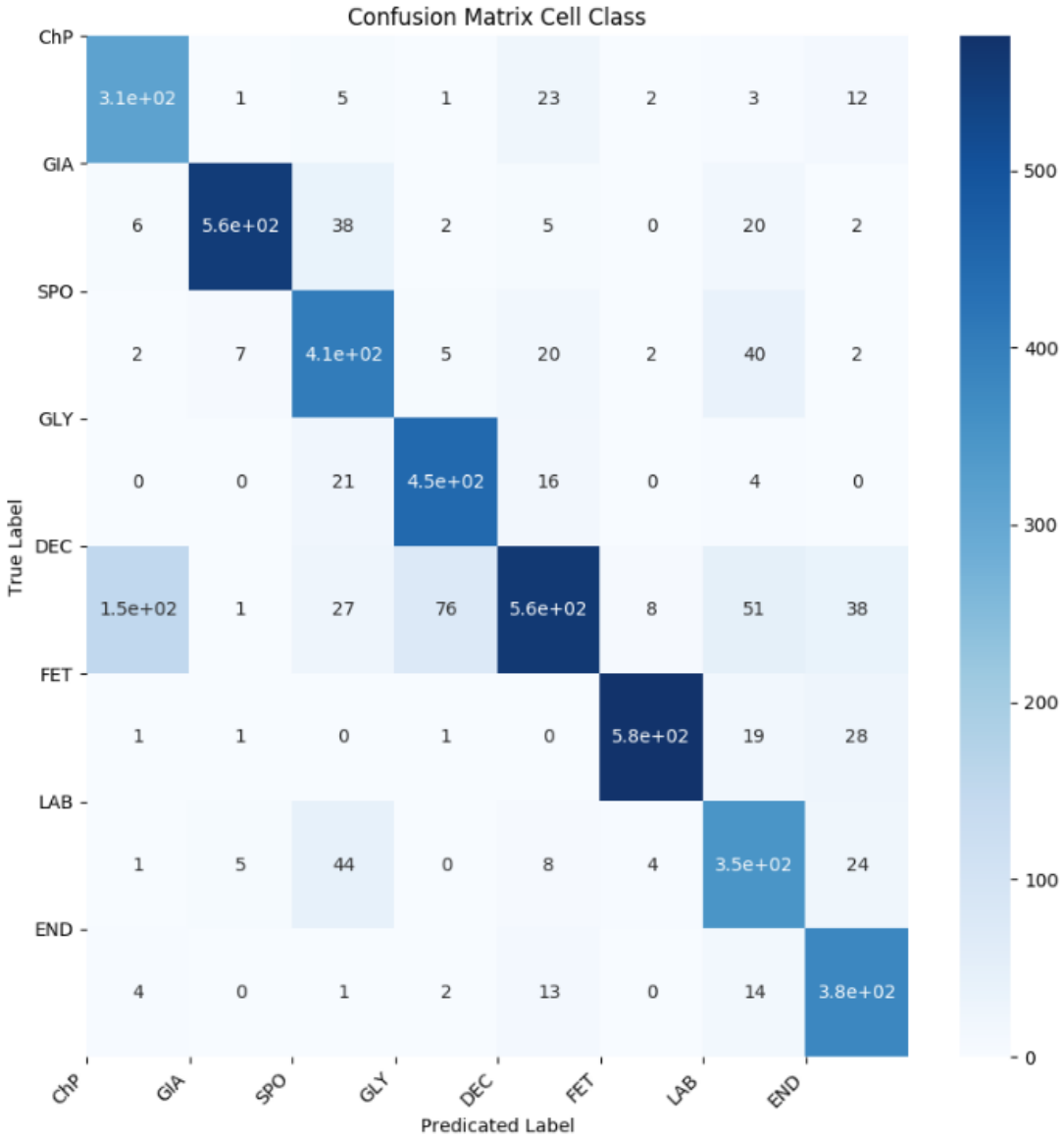



**Figure 5.17:** *This figure presents a confusion matrix for the Mouse Cell Classification Model. Each cell in the matrix represents the model's predictions (columns) against the true classes (rows) for the eight distinct placental cell types: Decidual Cells (DEC), Endothelial Cells (END), Chorionic Plate Cells (CHP), Foetal Blood Cells (FET), Trophoblast Giant Cells (GIA), Glycogen Cells (GLY), Labyrinth Cells (LAB), and Spongiotrophoblast Cells (SPO). The diagonal elements of the matrix represent instances where the predicted class matches the true class, providing insight into the model's performance.*

comparing spatial distributions in the absence of direct stereological data. While traditional stereology, as applied in studies like that of Coan et al., would offer a more direct comparison of structural complexities and density measures, the current methodological constraints and lack of access to comparable raw data necessitate the use of density plots derived from our classification model. These plots still provide valuable quantitative insights that parallel some aspects of stereological analysis, thus supporting the robustness of our computational approach in placental research.

Despite the high performance at E14.5, the model's efficacy diminished significantly at E9.5 (as shown in Figure 5.20). This discrepancy can be attributed to the insufficiency of annotated E9.5 cell class data. At E9.5, cells can exhibit significant differences due to their diverse roles as the placenta develops. Therefore, to achieve optimal performance, different models tailored to the specific developmental stages might be required.

The current study was unable to address this issue due to time constraints, which prevented the annotation and bootstrapping of a model specific to E9.5, thereby resulting in the observed reduced performance. However, it is expected that with dedicated effort towards annotating more E9.5 data, this limitation could be addressed in future studies.

This research represents a significant step towards the accurate classification of mouse cell types using image classification models. However, as with any study, there are areas for future improvement and expansion. For instance, future work could focus on refining the model's performance at different developmental stages or investigating other architectures or training strategies. Despite these areas for improvement, the models and methodologies employed in this study have proven to be robust and effective tools for cell classification.

## 5.13 Deployment Across the DMDD Dataset

The culmination of this research phase is the development of a new pipeline, representing a significant enhancement over the work of Ferlaino et al. This pipeline is not just an incremental upgrade but a comprehensive end-to-end research

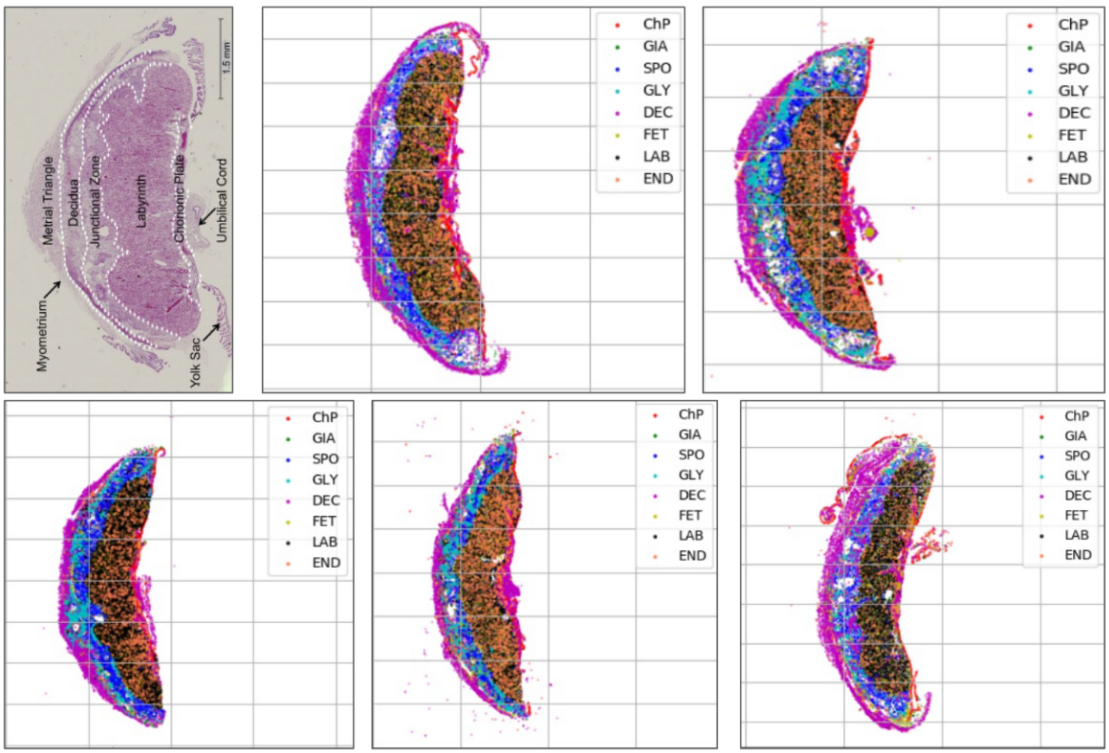



**Figure 5.18:** *This figure showcases the E14.5 cell classification across a whole slide image. The leftmost image is an annotated placental histology image, and the subsequent five images show the cell classes plotted on the whole slide image. The plots represent the Decidual Cells (DEC), Endothelial Cells (END), Chorionic Plate Cells (CHP), Foetal Blood Cells (FET), Trophoblast Giant Cells (GIA), Glycogen Cells (GLY), Labyrinth Cells (LAB), and Spongiotrophoblast Cells (SPO). Each image depicts a high degree of accuracy in cell classification, with the cells appearing precisely where they are expected.*

tool. It encompasses a full suite of functionalities including localisation of nuclei, cell classification, efficient storage of results, and streamlined processing — all integrated into a single, cohesive system. A notable achievement of this pipeline is its capability to perform inference on a complete placenta sample in an average of just 12 minutes, a testament to its efficiency and a substantial improvement from the previous methodologies.

What sets this pipeline apart is its design ethos, focusing on code efficiency, extensibility, and maintainability. These qualities ensure that the pipeline is not only effective for current research needs but is also poised for future expansions and adaptations. The addition of abstraction layers further enhances its generalisability, making it adaptable to a broad range of datasets beyond the scope of the current

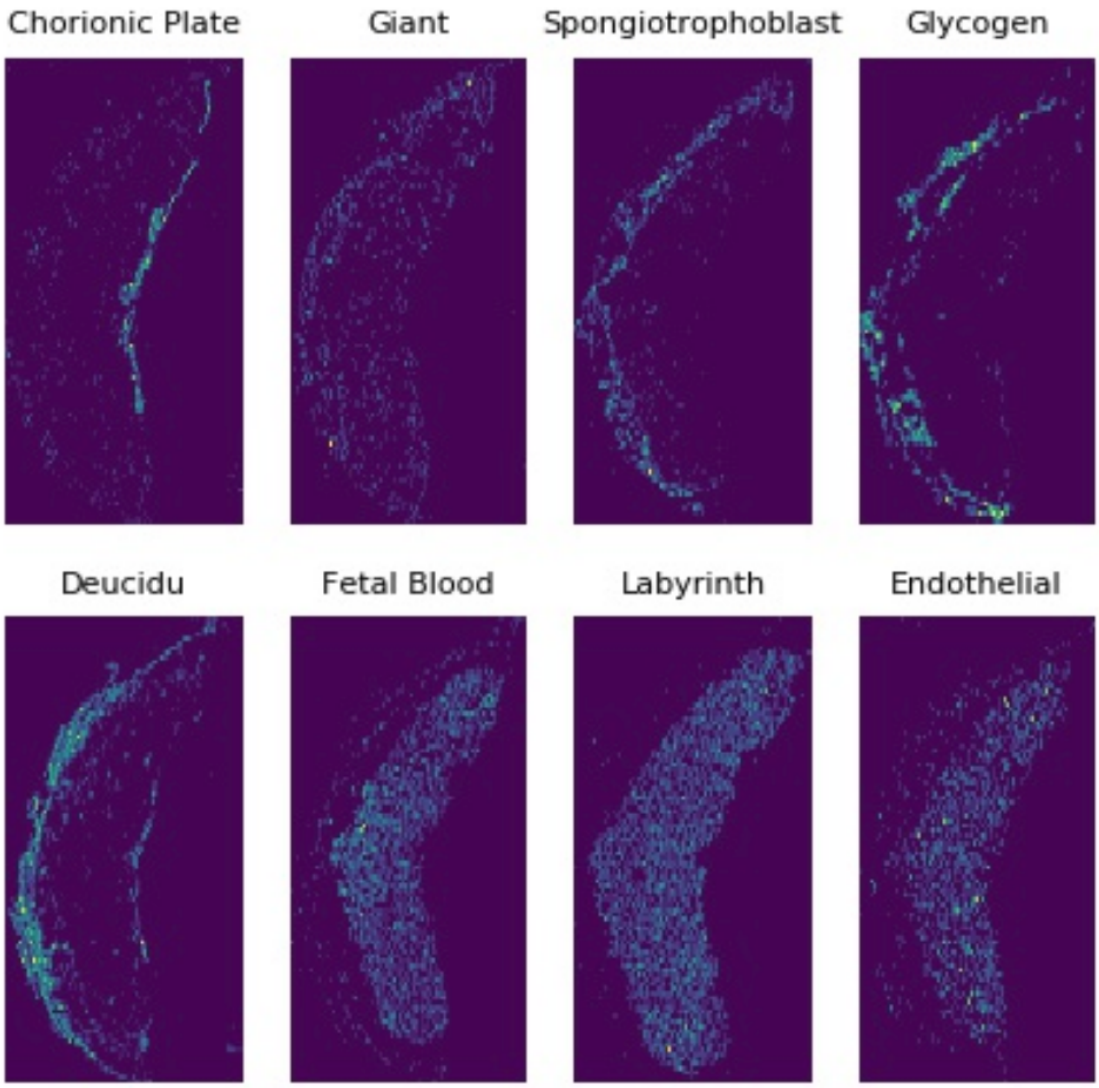



**Figure 5.19:** *This figure presents density plots for each of the eight mouse placental cell classes: Decidual Cells (DEC), Endothelial Cells (END), Chorionic Plate Cells (CHP), Foetal Blood Cells (FET), Trophoblast Giant Cells (GIA), Glycogen Cells (GLY), Labyrinth Cells (LAB), and Spongiotrophoblast Cells (SPO). These plots further reinforce the precision of the cell classification model, as the density distribution in each plot aligns with the expected locations of the respective cell classes within the placental tissue.*

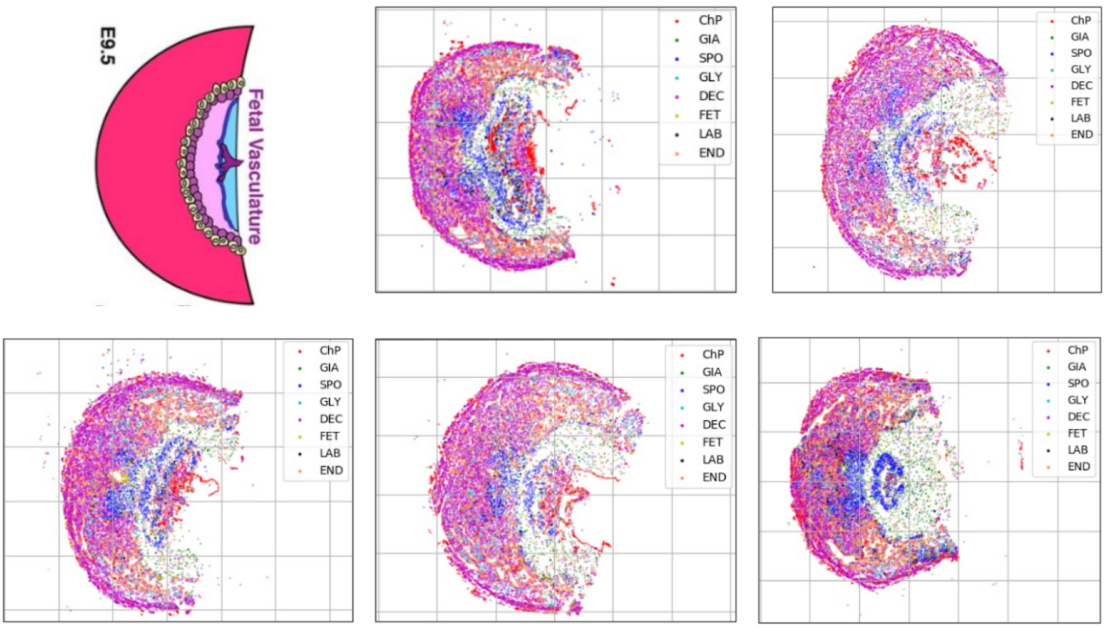



**Figure 5.20:** *The figure displays the results of cell classification across a whole slide image at the E9.5 stage. The annotated placental histology image and the five plots represent various cell classes. While the locations of the cells are reasonable, the classification accuracy at the E9.5 stage is somewhat lower than at the E14.5 stage. This discrepancy can be attributed to the smaller sample size of the training set annotated at E9.5, underscoring the impact of training data volume on model performance.*

DMDD dataset.

This pipeline, therefore, stands as a robust, versatile tool in the field of placental biology research. Its efficiency and comprehensive nature make it a valuable asset for current research, while its design principles ensure it is a foundation upon which future researchers can build. The detailed analysis of its performance and implications for large-scale data analysis will be the focus of the next chapter, highlighting the pipeline's role in advancing the field of developmental biology.

## 5.14 Summary and Conclusion

The overarching aim of this chapter was to construct a comprehensive automated placental phenotyping method for whole slide histology images. This aim was pursued through the attainment of several interlinked objectives:

- Enhancing the human cell classification models devised by Ferlaino et al. (2021) through the utilisation of advanced machine learning techniques and a

broader dataset.

- Developing a mouse nuclei detection and cell classification model that can accurately and efficiently identify and classify nuclei in mouse placental tissues.
- Establishing an efficient whole slide processing method capable of handling extensive datasets and producing accurate, repeatable results.

The datasets used in this research were derived from two primary sources: human and mouse placental cells. Despite both being placentas, they possess different cellular compositions and structures, necessitating unique training data for each. Detailed description of the datasets is provided in the text.

The methodology adopted in this study hinges on two primary stages: nuclei detection and cell classification. Both stages harness the power of advanced machine-learning techniques to deliver high-accuracy results. A detailed description of the models and techniques used is provided.

The codebase developed for this research ensured the efficient processing of whole slide images, allowing them to be dissected into smaller tiles and passed through the model for analysis. An integral feature of this codebase was the ability to identify and remove tiles containing no tissue, thereby saving valuable computational resources. The results were stored in low-memory objects for easy access and further analysis. An overlapping and clustering technique was developed and integrated into the codebase to account for edge cases arising from the tiling process.

This research endeavoured to develop a comprehensive automated whole slide placental phenotyping pipeline, a task that was accomplished. The resultant deep learning-based approach can localise and classify every cell in whole mouse and human placental tissues slide images. The pipeline is efficient and capable of deployment over large datasets, making it a valuable tool for extensive histological analyses. The slide plot illustrated in figure 5.18 clearly demonstrates the pipeline's effectiveness, showcasing comprehensive cell capture across various tissue regions.

Despite its effectiveness, the pipeline does face limitations, especially concerning its performance with E9.5 mouse placental tissues. This limitation arises primarily

due to cell morphological differences at different gestational stages. Annotating additional training data for E9.5 could significantly improve the model's performance for this specific gestational stage.

Future research directions involve expanding the number of tissue models to include the liver, pancreas, kidney, and other organs. This expansion would necessitate the annotation of more data to train new models, thereby offering a more comprehensive analysis. Furthermore, exploration into self-supervised pre-training and transfer learning could reduce the data required for accurate models, thereby increasing efficiency. In the long term, it would be interesting to investigate the potential for developing a self-supervised method that categorises cells automatically across tissue samples and species, thereby eliminating the need for supervised training data.

In addition, it would be advantageous to refine the results further by developing a graphical interface for ease of use in labs, coupled with a machine learning operations (MLOps) approach for real-time labelling refinement. This would provide scientists with a continuously improving tool tailored to their needs.

Ultimately, the research conducted in this chapter aims to contribute significantly to the field by providing a tool for automated placental phenotyping analysis. This tool can enable scientists engaged in embryo work to incorporate placental analysis, thus contributing to the exploration of this understudied organ. Furthermore, providing more detailed and repeatable quantitative results can assist those conducting placental work.

*"The stuff of life turned out to be not a quivering, glowing, wondrous gel but a contraption of tiny jigs, springs, hinges, rods, sheets, magnets, zippers, and trapdoors, assembled by a data tape whose information is copied, downloaded and scanned."*

— George Wald

# 6

# Automated Analysis of the DMDD Placental Data

## Contents

## 6.1 Introduction

Deciphering the genes crucial for normal embryogenesis is critical to understanding the process of embryo development. This understanding can potentially unravel the genetic foundations of developmental abnormalities and illuminate potential therapeutic avenues. This could, in turn, mitigate the substantial financial and emotional implications of congenital disorders [117].

Recent advancements in genetics and developmental biology have facilitated large-scale phenotyping initiatives, such as the International Mouse Phenotyping Consortium (IMPC) project. Studies under the umbrella of the IMPC project reveal that approximately 30% of mouse gene knockouts result in non-viable progeny. However, the majority of this research has focused on the effect of mutations on the embryo itself. Meanwhile, extraembryonic tissue has received relatively little attention, leading to an underrepresentation of placental phenotypes in the scholarly literature.

The scientific literature is replete with instances of placental abnormalities precipitating adverse birth outcomes, including intrauterine growth restriction, preeclampsia, and neurological disorders [118]. Experimental techniques such as tetraploid complementation or conditional gene ablation have demonstrated that the normal development of some embryos can be completely restored with a wild-type placenta. However, according to the Mouse Genome Informatics (MGI) database, 10% of embryonic-lethal strains manifest extraembryonic defects. A more comprehensive understanding of embryogenesis would benefit from an increased focus on the role of the placenta in development.

The Deciphering the Mechanisms of Developmental Disorders (DMDD) consortium represents a considerable effort to systematically probe the role of genes in development by examining both embryos and the placenta. High-Resolution Episcopic Microscopy (HREM) images were utilised to conduct detailed phenotypic

assessments of structural abnormalities in mutant embryos. Concurrently, the effect of each mutation on placental development was evaluated through the scoring of placental histology images, a task undertaken by the Hemberger laboratory [107].

The Hemberger group analysed 103 genetic knockout lines at two developmental stages (E9.5 and E14.5). Of these, 82 were classified as P14 lethal, as no surviving mutant offspring were recovered at that stage. The remaining 21 lines were categorised as sub-viable, as mutant pups constituted 13% or less of all offspring obtained, significantly lower than the 25% anticipated from heterozygous crosses. Upon histological examination of all placentas, it was discovered that less than 1% of wild-type placentas exhibited an abnormal phenotype, whereas 56 out of 82 (68%) of the P14 lethal strains displayed dysmorphologies in mutant placentas. Including the P14 sub-viable lines, the placental phenotype rate remained robust at 59%, substantially higher than the approximately 10% documented in the Mouse Genome Informatics (MGI) database.

A stage-dependent analysis of placental defects revealed that almost all lines that expired before E14.5 had placental abnormalities, compared to only 35% of viable lines beyond E14.5. This observation suggests that mutations resulting in embryonic lethality between E9.5 and E14.5 are likely associated with a defective placenta. Additionally, mutant E14.5 embryos from strains exhibiting a placental phenotype were shifted to an earlier developmental stage than those with normal placental development, underlining the placenta's critical role in nutrient supply.

The Hemberger laboratory classified placental defects across the three primary layers of the mature placenta: the labyrinth, the junctional zone, and the maternally derived decidua. Histopathological scoring was conducted on Haematoxylin and Eosin (H&E) histology sections obtained from the sagittal midline (as detailed in Chapter 6).

The DMDD dataset, supplemented with the placental scores from the Hemberger lab, also contains embryo scores, thus enabling a more thorough exploration of the associations between specific phenotypes. Significant statistical correlations between embryo phenotypes and placental defects were demonstrated, including

heart, brain, and vascular systems abnormalities. Examples include anomalies in forebrain development, heart chamber and septum morphology, subcutaneous oedema, and overall artery and vein topology. These co-associations suggest the existence of co-regulatory or interdependent mechanisms during the development of certain organ systems, such as the placenta and the morphogenesis of the brain, heart, and vascular system.

The research illuminated the fact that placental malformations are more prevalent than previously thought in embryonic-lethal mutations and often co-occur with heart, brain, and vascular network defects. This research underscored the need to include extraembryonic tissues in studies investigating the genetic origins of congenital abnormalities. However, certain limitations of the study were noted, such as the absence of quantitative measurements in the scoring, such as the distorted spongiotrophoblast/labyrinth ratio, which was not assigned a numerical value, thereby subjecting it to human interpretation. A numerical value could yield more precise and reproducible results. Additionally, such a large-scale analysis proved time-consuming due to the manual scoring process.

The challenges of procuring quantitative and replicable measurements can be overcome by employing automated histological analysis techniques. These automated approaches also help decrease the costs of conducting large-scale analyses. The research conducted by the Hemberger lab provides an ideal platform for evaluating the reliability of an automated pipeline.

## 6.2 Chapter Aims

1. Evaluate the efficacy of the phenotyping pipeline presented in Chapter 5 by comparing its output results to those reported by the Hemberger lab on the Deciphering Mouse Developmental Disorders Dataset.

2. Utilise the pipeline to construct a more comprehensive examination of the DMDD dataset.

### 6.2.1 Objectives

1. Develop a method for mouse placental tissue structure analysis.
2. Analyse the DMDD dataset and look for overlap with Hemberger paper results.
3. Look for links between embryonic outcomes and placental phenotypes.

## 6.3 Methods

### 6.3.1 Overview

This section delineates the methodologies employed for tissue analysis, random forest feature analysis, and bioinformatic analysis.

### 6.3.2 Cell Localisation and Classification

The methodology adopted for capturing the positions and classes of all cells in the DMDD dataset is detailed comprehensively in Chapter 5.

### 6.3.3 Tissue Structure Analysis by adapting Concave Hull

Following the capture of cell locations and classes, it is possible to extract further phenotypic insights through automated tissue region analysis. Pathologists often scrutinise cell morphology, tissue structures, variations in texture or colour, or defined abnormalities such as blood clots to provide morphological assessments of tissues via scoring or grading systems. Histological analysis is indispensable for distinguishing and classifying phenotypes to augment our understanding of disease mechanisms and enhance therapeutic approaches.

In the work of the Hemberger group, the principal phenotypes studied encompassed placental size, distorted spongiotrophoblast/labyrinth ratios, cellular density, fibrotic and necrotic areas, labyrinth vasculature, and areas of blood pools. However, due to time constraints, the training data for these phenotypes was not generated to capture blood pools and fibrotic and necrotic areas. Nevertheless, the methodology delineated in Chapter 5 was extended to capture the phenotypes from the paper and beyond.

The method discussed in Chapter 5 accurately captured cell locations and classes, thereby enabling the quantification of cell ratios and counts, which addresses one of the phenotypes (distorted cell ratios) measured by the Hemberger Placental Assessment. Furthermore, the automated approach is more reproducible than manual methods. The subsequent phase involved the capture of other phenotypes, such as placental size, the minor axis of tissue regions, and cellular densities. A point cloud-based approach was leveraged to achieve this, using the cellular locations and classes within specific regions. For instance, to segment the junctional zone, a point cloud was created around glycogen and spongiotrophoblast cells, known to exclusively exist in that region, and then segment the area.

A modified version of the alpha concave hull, a nearest neighbour-based approach, was employed to segment out regions. This approach, which generates non-convex polygons to enclose a cloud of points, was introduced as a concave hull [119]. An alternative approach is the convex hull, but the concave method yields a tighter fit of the point cloud [120].

The concave hull algorithm is an exemplar of shape-based clustering, which is used to generate an optimal concave hull from a given set of points in a given space. This hull is defined as a non-convex shape formed by connecting the points such that no other points from the underlying set are inside the hull. The algorithm works by iteratively removing points from the set, resulting in a more concave hull. The process concludes when the hull contains no more points that can be removed without disrupting its concavity. The concave hull algorithm is widely used in image segmentation, object recognition, and map generation.

The method applied to the placental dataset enables the segmentation of the entire placenta and specific tissue regions, depicted in Figure 6.3. This allows the capture of additional features, including that in table 6.1.

Each of these features provides invaluable insights into the structural and functional changes in the placenta, which can further enhance our understanding of the genetic basis of developmental disorders.

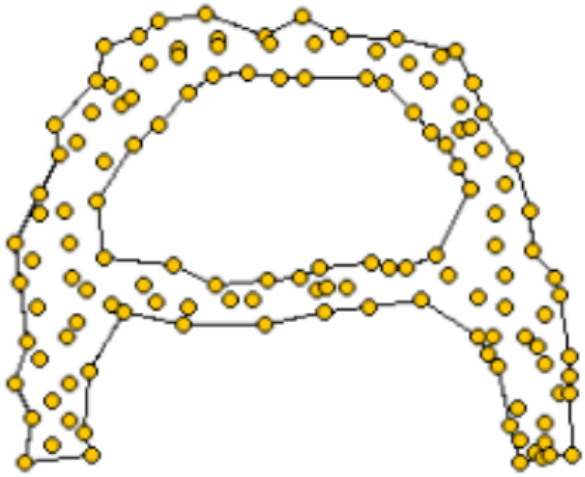


**Figure 6.1:** *This figure provides a visual representation of the concave hull method applied to a set of points forming the shape of the letter 'A'. The concave hull method is a technique used in computational geometry to determine the shape that encloses a given set of points. Unlike the convex hull method, which generates a rubber-band-like shape around the outermost points, the concave hull method generates a shape that more accurately conforms to the configuration of the points, even if they form concave structures. This capability is evident in the figure, where the concave hull accurately captures the shape of the letter 'A'. Figure from Link*

| Metric | Zone |
|---|---|
| Area | Placental |
| Major Axis | Placental |
| Minor Axis | Placental |
| Cell Density | Placental |
| Area | Junctional Zone |
| Major Axis | Junctional Zone |
| Minor Axis | Junctional Zone |
| Cell Density | Junctional Zone |
| Area | Labyrinth Zone |
| Major Axis | Labyrinth Zone |
| Minor Axis | Labyrinth Zone |
| Cell Density | Labyrinth Zone |

**Table 6.1:** Metrics Evaluated in Different Zones of the Placenta

### 6.3.4 Selection of Placental Sections for Analysis

To ensure clarity in the methodology and the integrity of the data analysis, it is important to specify that multiple sections from the same placenta were not used in this study. Instead, a systematic approach was employed to select a single

section from each placenta for detailed examination. This selection was guided by the unique animal ID, which facilitated the identification and inclusion of only one section per animal in the analysis.

The criterion for selecting a particular section was based on the segmented area of the tissue. Among the available sections, the one with the largest area was chosen, under the assumption that a larger area would most likely correspond to a section cut closer to the sagittal midline. This approach was predicated on the hypothesis that sections closer to the midline would provide the most representative and comprehensive view of the placental structure and its cellular constituents, thereby yielding the most reliable data for subsequent analyses.

This methodological decision was used in maintaining the consistency and reliability of the dataset, ensuring that the analysis was not confounded by potential redundancies or biases that could arise from analyzing multiple sections of the same placenta.

### 6.3.5 Statistical Analysis

In order to generate a more comprehensive understanding and interpretation of the results emanating from the phenotype extraction pipeline, we apply bioinformatics, crucial for analysing extensive biological datasets. This approach uses computational tools to enhance our understanding of gene functions.

A key component is hypothesis testing, a statistical method to test assumptions about data. It involves formulating two hypotheses: the null hypothesis (H0), suggesting no effect, and the alternative hypothesis (H1), indicating a possible effect.

For example, in a two-tailed hypothesis test, H0 might state 'there is no difference between variables', while H1 proposes 'there is a difference'. The statistical test's outcome is expressed as a p-value, reflecting the likelihood of observing the data if H0 is true. A p-value below 0.05 typically signals significance, implying the data is unlikely due to chance under H0. However, in large-scale bioinformatics, where multiple tests are common, this threshold can lead to many false positives.

Adjustments like the Bonferroni correction and False Discovery Rate (FDR) help mitigate this. Bonferroni is conservative, adjusting the alpha level based on the number of tests, whereas FDR controls the proportion of false discoveries among rejected hypotheses. Such adjustments are vital in extensive data analyses to maintain statistical reliability.

In this study, we frequently utilise the T-test, applying Bonferroni adjustments for multiple testing. These methods ensure the robustness of our findings, especially given the scale of the data.

### 6.3.6 Linking Placental Phenotypes to Foetal Outcome

The DMDD dataset offers a valuable resource for this purpose, containing detailed embryo annotations and placental histology sections. To gain insights into the influences of various placental phenotypes on embryonic outcomes, researchers need to perform multivariate analysis. This approach, which is an integral part of bioinformatics, involves leveraging multiple variables, such as gene expression levels, to identify the key features that drive placental and embryonic development. Such analysis is essential not only for comprehending the underlying molecular mechanisms but also for identifying potential therapeutic targets for related diseases.

Tree-based models, such as Random Forest and XGBoost, have proven to be highly effective in machine learning analysis, especially when working with tabular data. These models are particularly adept at handling complex variable interactions and non-linear relationships, which are common in biological data. Additionally, their inherent structure allows them to handle mixed data types (numerical and categorical) and missing values, making them suitable for diverse datasets. The advantage of tree-based models over deep learning models lies in their ability to work efficiently with smaller datasets and require less preprocessing, whereas deep learning models need extensive feature engineering and large datasets to capture intricate patterns in tabular data. This makes tree-based models more suitable for scenarios where interpretability and ease of implementation are crucial, such as predicting outcomes from high-dimensional, tabular data in bioinformatics studies.

### 6.3.7 Random Forest

Random Forests, an ensemble learning technique, are widely employed for classification and regression tasks. They are composed of Decision Trees, which are intuitive and commonly used to structure the decision-making process in everyday life. For example, if one needs to determine whether a house is in the United States or the United Kingdom, one could ask questions such as "What currency was the purchased?" and "What is the nearest major city to the house?" This would enable them to come to a conclusion [124].

Decision Trees are a supervised learning algorithm that begins with a root node, from which evaluations are made at branch nodes, culminating in a leaf node where a decision is made. The Gini impurity is calculated to construct the tree, where j represents the number of classes in the labels and P represents the class ratio at the i-th node. The Gini impurity is at its lowest and most optimal when it is 0 and at its highest and least optimal when it is 0.5. The node with the lowest Gini impurity is selected as the root node, followed by the subsequent lowest Gini impurity as the branch, and so on, until a leaf node with a zero Gini impurity is reached, thus giving the decision [125].

While Decision Trees fit a specific data set well, they must generalise better. Thus, Random Forests are used for significantly better classification accuracy. The power of Random Forests comes from assembling many Decision Trees. The most common prediction is then used as the model's predicted value. The specific ensemble method used is bagging. In the bagging method, a random sample of data from the training set is used to train each specific tree independently [126].

Random forests are prone to overfitting when they become too deep, so several hyperparameters can be tuned to prevent this. These include the number of estimators (trees in the forest), the maximum number of features considered when splitting a node, and the maximum depth of the trees. Furthermore, other hyperparameters can be found in the Scikit-Learn documentation. Random forests are relatively easy to use and effectively avoid overfitting; however, their computational complexity can be high depending on the number of trees. One of their most beneficial aspects

is their ability to evaluate the relative importance of each feature in the prediction, which can be done through various methods.

### 6.3.8 XGBoost

Since 2016, the tree-based algorithm, XGBoost (eXtreme Gradient Boosting), has emerged as a dominant force in machine learning. XGBoost implements gradient-boosted decision trees, a supervised learning algorithm that combines weaker decision tree models to produce accurate predictions. XGBoost stands out from other methods, such as random forests, due to its use of boosting, in which each tree adjusts the evaluation criteria based on feedback from the previous tree, and its optimisation through parallelisation, tree pruning, regularisation, sparsity awareness, and weighted quantile sketch [127] . These optimisation methods are described in the XGBoost paper and have enabled XGBoost to be successful in many Kaggle competitions involving small to medium-sized tabular datasets. XGBoost is also a method in which feature analysis can be applied similarly to random forests. Additionally, XGBoost has rated the best method on scikit-learn's 'Make Classification' dataset.

### 6.3.9 Feature Analysis

Measuring feature importance in tree-based approaches can be achieved via several methods. Here I will discuss three of the most popular methods of measuring feature importance, Gini importance, permutation importance, and drop column importance.

Gini importance assigns a score to each feature based on the decrease in impurity across the entire tree that results from splitting on that feature. This score is calculated by multiplying the gain of splitting on the feature by the probability of the feature being used in the split [126].

Permutation importance is a measure of the importance of a feature based on how much the model's performance decreases when the feature is randomly shuffled.

This method can be used to measure both feature and model importance since it measures the impact of a feature on the model's performance.

Drop column importance is a measure of the importance of a feature based on how much the model's performance decreases when the feature is removed from the dataset. This method is useful for understanding how a particular feature affects the model's performance and, thus, its importance in outcome prediction.

## 6.4 Results and Discussion

### 6.4.1 Investigation 1: Computational Tissue Segmentation in Mouse Placenta

**Aim:** The aim was to extend the work in Chapter 5 and develop a comprehensive segmentation approach for whole placental tissue and specific tissue regions.

**Method:** Leveraging the concave hull method, the study attempted to accurately segment key placental structures in the E14.5 placenta. Cell classification, area measurements, cell densities, and major and minor axes were calculated to replicate the measures taken by the Hemberger Lab.

**Result:** Successful segmentation was achieved across all placenta datasets, and additional phenotypes, such as precise placental size, were determined. However, the segmentation success was reduced for the E9.5 placenta, primarily due to less effective cell classification.

**Conclusion:** The concave hull-based methodology proved effective for tissue structure segmentation in the mouse placenta, particularly for E14.5. However, the segmentation of tissue regions in E9.5 could be improved with more effective cell classification models.

**Future Work:** Future research could explore improving cell classification models for E9.5, the segmentation of additional placental regions, and the use of fallback functions to mitigate issues with image quality or gaps that cause poor segmentation.

Histological specimens, which pathologists typically assess to evaluate health or diagnose disease, involve an examination of tissue sections under a microscope. Parameters such as cell and tissue structure, size, shape, and appearance are assessed during this process. Additionally, the presence of any abnormal cell growth,

inflammation or infection, cellular activity, fat or connective tissue, and unusual proteins or molecules can be evaluated. Immunohistochemistry, a method often used to measure the presence of specific proteins or antibodies within the tissue, can be time-consuming and susceptible to human error.

The advent of automated methods for tissue histology analysis may present a solution to enhance the efficiency of pathologists' workflow, thereby addressing the ongoing reproducibility crisis. However, for these automated methods to be successful, they need to emulate the measurements traditionally carried out by pathologists. Within the Hemberger Lab, these measures include the minor axis of the junctional zone, cell densities, and placental size. To achieve this, an adaptation of the concave hull approach was utilised.

The use of the concave hull-based method enabled the segmentation of whole and specific placental tissue regions, facilitating the calculation of area measurements, cell densities, and major and minor axes. Figure 6.3 depicts examples of segmentations of E14.5 placentas, illustrating the application of the concave hull approach. Additionally 6.2 demonstrates that cell are in the regions we expect. A colour legend shows scell classes.

By incorporating all cells in the placenta, a point cloud was generated for segmenting the entire placenta using the concave hull-based approach. The term 'point cloud' refers to a set of data points in space, in this case representing the cells of the placenta. The alpha value parameter was fine-tuned to optimise the performance of the concave hull approach. The generated shapes were then ranked, and the largest continuous shape was selected. This approach aimed to minimise the likelihood of tissue detection outside the central placenta or cell misclassifications in an empty space. Across all placental datasets, segmentation proved highly successful, with any issues arising from poor image quality being addressed during the quality control (QC) procedure, as described in section 4. Following segmentation, additional phenotypes, such as precise placenta size, can be determined. This measurement is later utilised in this chapter to differentiate between multiple histology sections from the same placenta.

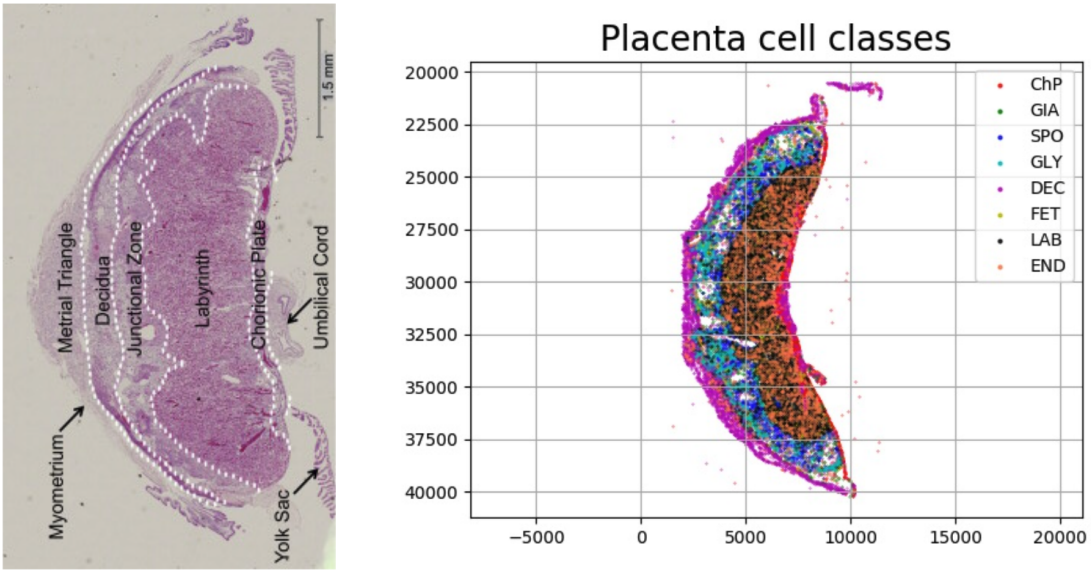



**Figure 6.2:** *This figure combines a histological section of an E14.5 mouse placenta with a detailed scatter plot demonstrating the cellular distribution across different placental regions. The left panel shows a labeled histological section identifying key tissue regions such as the Decidua, Junctional Zone, Labyrinth, and others, offering a comprehensive view of placental structure. The right panel presents a scatter plot where placental cell classes are precisely located within tissue regions where they are biologically expected to be found, such as Trophoblastic Giant cells in the Junctional Zone and Endothelial cells in the Labyrinth. This alignment of cellular distribution with expected histological regions serves as strong evidence of the functionality and accuracy of the cell classification and localization pipeline employed in this study.*

In the Hemberger data, histology sections were taken through the sagittal midline, as discussed in section 4. In some instances, multiple slices were taken. To determine which sections should be used for analysis, the largest section by area was selected, as calculated by whole placenta tissue segmentation. This selection ensured that the most comprehensive histological section was used for the subsequent analysis, thereby providing a representative sample of the placenta.

Following the segmentation of the whole placenta, efforts were concentrated on isolating specific tissue regions. The four regions of the placentas in the dataset include the maternal decidua, the junctional zone, the labyrinth, and the chorionic plate. Although the decidua and chorionic plate regions can often be identified during the harvesting of the placentas, inaccuracies may be present. Consequently, the focus shifted to segmenting the two largest regions—the junctional zone and the labyrinth.

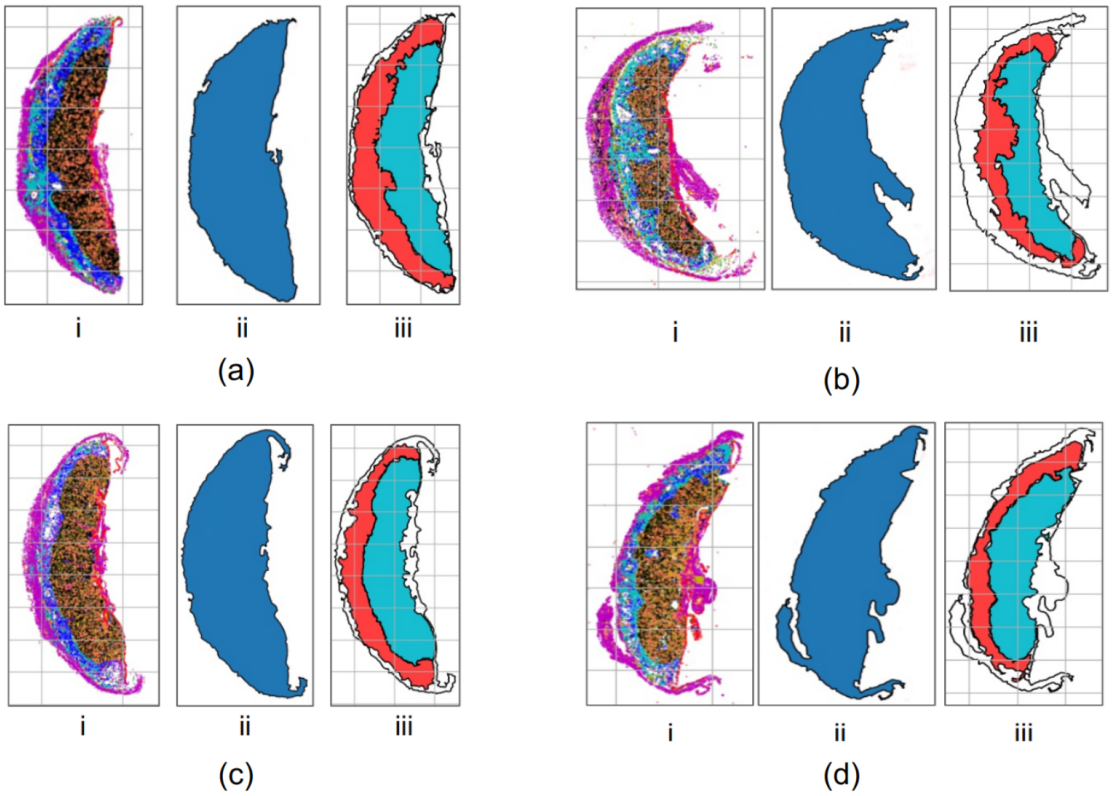



**Figure 6.3:** *This figure demonstrates tissue region segmentation in four E14.5 placental samples (a, b, c, and d). Each sample includes three sections: i) A scatter plot of the cells identified and classified by the deep learning cell classification model and nuclei locator from 5; ii) A comprehensive segmentation of the full placenta, providing an overarching view of the tissue structure; iii) A more focused segmentation that distinctly highlights the junctional zone (depicted in red) and the labyrinth (depicted in blue). These visualisations underscore the successful application of the deep learning model in identifying and segmenting key placental tissue regions, thereby offering a holistic view of the cellular distribution and structure within the E14.5 placenta.*

To segment the junctional zone, spongiotrophoblasts and glycogen cells were utilised. Owing to the lower abundance of these cells in other regions, a point cloud generated by the concave hull method—with the alpha parameter adjusted—was utilised to distinguish the junctional zone. The clouds were assessed and ranked, with the most substantial cloud being selected to ensure precise segmentation of the junctional zone. An analogous approach was adopted to delineate the labyrinth region. In this instance, the point cloud was composed using labyrinth cells, foetal blood cells, and endothelial cells, as these cells are uniquely present in the labyrinth.

This approach proved successful across the vast majority of the E14.5 dataset,

with a few exceptions that had to be addressed during the quality control process. These exceptions are illustrated in figure 6.4. The most common issue was caused by regions such as the junctional zone or the labyrinth being clipped in half, as shown in figure 6.4 (a, c, d). This was due to imperfections in the image resulting in a partial cut across the tissue section or cells in a sparsely populated region. To address this issue, the alpha value could be increased, or the next largest point cloud could be joined if it was above a certain threshold. However, since this issue occurred infrequently, it was simply resolved by removing samples containing the issue in the quality control process. Another occasional issue was regions being over-segmented due to the potential misclassification of cells (figure 6.4 (b)). These may represent actual phenotypic variations, but due to the sheer number of samples,they were removed from analysis due to time constraints.

While segmentation was also carried out on the E9.5 data, the results were less successful. Segmentation of the entire placenta was relatively accurate due to the high density of the point cloud captured from the accurate nuclei detection in both E14.5 and E9.5. However, the segmentation of tissue regions was significantly less effective, with very few cases markedly segmenting the junctional zone and labyrinth. This may be due to the poorer cell classification in E9.5 as discussed in chapter 5, as well as the significantly different tissue structure makeup between E9.5 and E14.5, such as the large maternal decidua in comparison to the rest of the placenta and the relatively small size of the junctional zone. To address these issues, more E9.5 cells could be annotated and better cell classification models trained for E9.5. Furthermore, a reevaluation may be necessary of the tissue regions to be captured in the E9.5 analysis, such as exploring the histological makeup of E9.5.

This study demonstrated that the tissue structure in mouse placenta can be successfully segmented using the concave hull approach. This segmentation allows for the entire placenta and tissue regions, such as the junctional zone and labyrinth, to be effectively captured. However, the method worked well in E14.5 but not in E9.5. To overcome this, better cell classification models could be trained with the E9.5 data. This study could be taken further by segmenting out additional placental

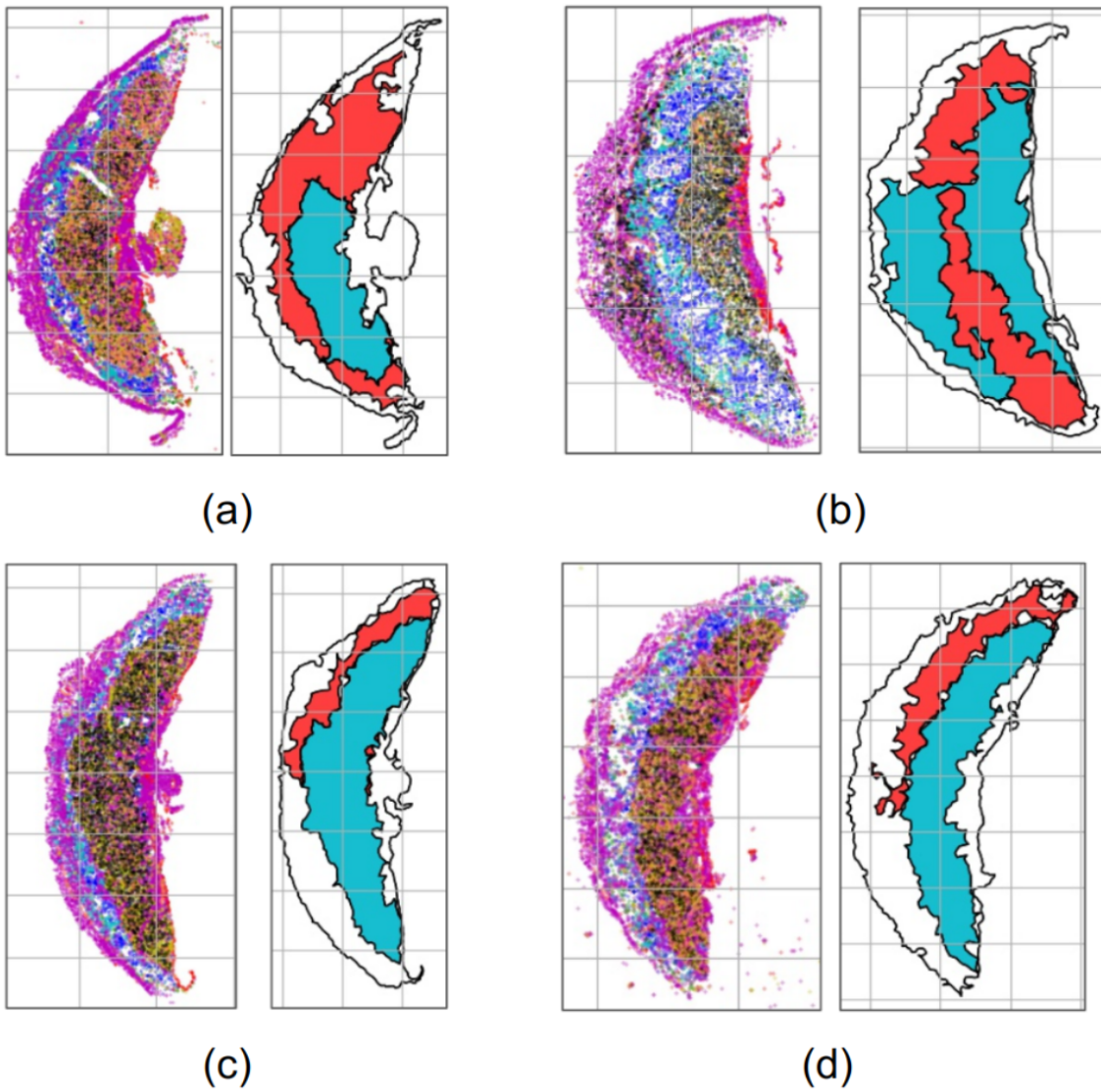



**Figure 6.4:** *This figure presents examples of the rare exceptions encountered during the segmentation of the junctional zone and labyrinth within the E14.5 placental samples. The most frequent issue was observed when the junctional zone or labyrinth regions appeared to be clipped or split, as depicted in samples (a, c, and d). These occurrences were typically due to imperfections in the image or sparsely populated cell regions. The segmentation method also occasionally resulted in the over-segmentation of regions due to potential misclassification of cells, as shown in sample (b). While these anomalies may represent genuine phenotypic variations, they were removed from the final analysis due to their infrequency and the extensive size of the sample set.*

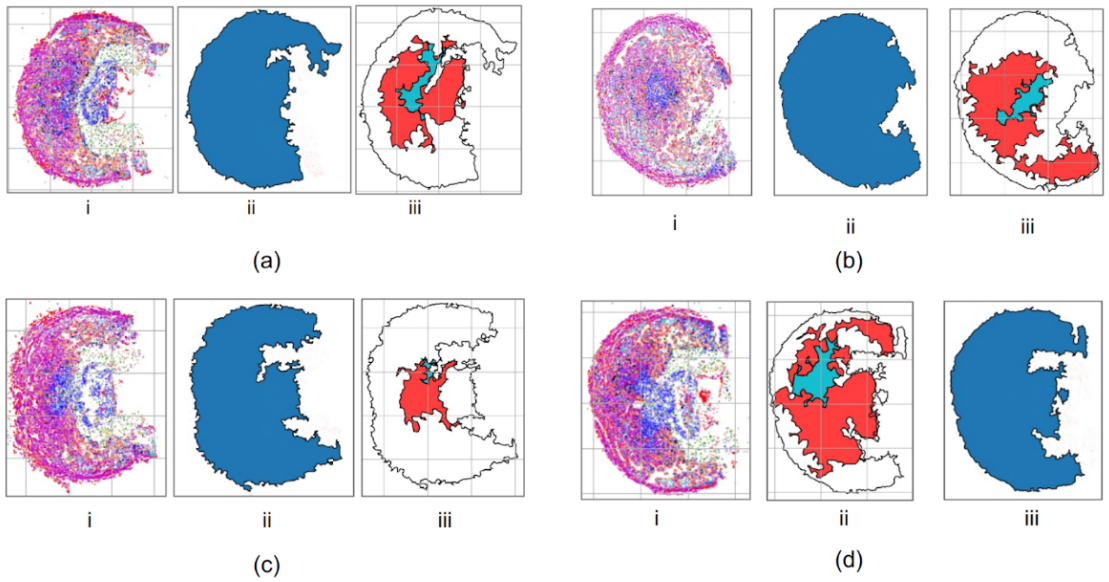



**Figure 6.5:** *The figure presents the tissue region segmentation results for four E9.5 placental samples (a, b, c, and d). Each sample consists of three sections: i) a scatter plot of the cells identified and classified by the deep learning cell classification model and nuclei locator; ii) a comprehensive segmentation of the full placenta; iii) a more focused segmentation, attempting to distinctly highlight the junctional zone and labyrinth. Despite the successful full placenta segmentation due to the high-density point cloud, the segmentation of specific tissue regions, namely the junctional zone and labyrinth, was less successful in the E9.5 samples. This reduced performance could be attributed to the poorer cell classification at E9.5, as well as the considerably different tissue structure makeup between E9.5 and E14.5.*

regions, like the decidua and chorionic plate, and using fallback functions to prevent issues with image quality or gaps causing poor segmentation. This research lays the groundwork for future studies that could further improve the automated analysis of placental tissue, thus contributing significantly to the field of histological analysis.

It is critical to clarify that this analysis did not involve manual segmentations, and therefore, the assessment of segmentation 'success' was strictly qualitative. Success was determined by visual inspection, focusing on the clarity of segment boundaries, the anatomical accuracy of segmented regions against established placental landmarks, and consistency across samples. This qualitative assessment was aimed to ensure that the segmented structures were accurately represented without significant artifacts or misclassifications. The use of manual segmentation, while beneficial for quantitative validation, was not feasible in this context, and as such, segmentation efficacy was judged based on the best available visual evidence.

### 6.4.2 Investigation 2: Mouse Placenta Wild-type Characterisation

**Aim:** The objective was to characterise the wild-type placenta, providing a crucial baseline for genetic studies and enabling a better understanding of the impacts of gene mutations or environmental exposures on placental and embryonic development.

**Method:** The wild-type population average for E14.5 was obtained from 144 randomly selected C57BL/6 samples. The investigation focused on characterising cell ratios, the placental histology section area, and the major and minor axes of the placenta.

**Result:** The results revealed the most abundant cell type to be decidual cells, followed by labyrinth cells, aligning with existing literature. The area of the placenta and its tissue regions matched expectations, with the labyrinth being larger than the junctional zone. However, outliers were identified in the chorionic plate and junctional zone, likely due to cell misclassification and measurement issues.

**Conclusion:** The investigation contributed to a comprehensive characterisation of the wild-type placenta, with findings largely aligning with the current literature. However, there is scope for further refinement in the measurement methodologies and addressing cell misclassification.

**Future Work:** Future research could explore adding more label cells to the training data, implementing advanced augmentation strategies or self-supervised pre-training to overcome cell misclassification, and refining the measurement methodologies for more accurate results.

The objective of this investigation was to characterise the wild-type placenta (Figure 6.6). Achieving a comprehensive understanding of the wild-type placenta provides a crucial baseline for comparison when studying gene mutations or environmental exposure effects on the placenta and embryonic development. The wild-type mouse is a critical model organism in genetics research, providing insights into gene-gene interactions, gene-environment interactions, and epigenetic influences on placental and embryonic development. This investigation focused on characterising cell ratios, the placental histology section area, and the major and minor axes of the placenta.

### Cell Ratios

Cell ratios offer a quantitative measure of the relative abundance of cell types within a tissue, thereby providing insights into the impacts of gene deletions on cell populations [128]. An improved understanding of cell ratios can shed light on the role of specific genes in various cellular processes, such as cell proliferation, differentiation, and migration. In the context of mouse knockout studies, cell ratios are instrumental in assessing the effects of genetic variants on cell populations and cellular processes.

For this investigation, the wild-type population average for E14.5 was obtained from 144 randomly selected C57BL/6 samples. The cell ratios were determined for eight cell types: chorionic plate, giant, spongiotrophoblast, glycogen, decidual, foetal blood, labyrinthine, and endothelial cells. The results revealed the most abundant cell type to be decidual cells, followed by labyrinth cells. This finding aligns with the existing literature, thereby confirming the validity of the results.

However, outliers in the chorionic plate were identified, likely attributable to cell misclassification. A suggested approach to overcome this issue is the addition of more label cells to the training data or implementing advanced augmentation strategies or self-supervised pre-training.

### Placental Histology Section Area

The area of the placenta and its tissue regions can offer critical insights into the organ's health. The results for the wild-type placenta matched expectations, with the labyrinth being larger than the junctional zone, aligning with previously published work [129].

While stereological analysis is recognized as the gold standard for the quantitative morphological assessment of placental structures, its application is often constrained by significant demands on time and resources. In this study, an alternative quantitative approach was utilized to characterize the wild-type placental architecture, primarily due to these limitations. Although this method does not capture the three-dimensional data that stereology offers, it allows for a practical assessment of placental characteristics within the logistical constraints of the project.

Our choice of measurements—cell ratios and placental dimensions—was aimed to approximate the data typically obtained from stereological methods, using the most comparable metrics found in the literature. This decision was informed by the need to balance methodological rigor with feasible execution given the scope of our project. However, it is acknowledged that this approach does not replicate the depth and precision of data provided by stereological techniques.

To substantiate and possibly refine the findings from this study, future research should incorporate stereological analysis. Applying stereological methods would enable direct comparison with established datasets and provide a robust validation of the computational pipeline employed here. Incorporating these techniques would not only affirm the results but also elevate the overall reliability and applicability of the models developed for placental research, bridging the gap between computational estimations and histological realities.

### Placenta Major and Minor Axis

The placenta's major and minor axis measurements also appeared reasonable, with some exceptions in the junctional zone potentially due to measurement issues. Future research could explore taking the average of measurements in all areas to ensure more comparable results with histologists' measurements.

This investigation has contributed to a more comprehensive characterisation of the wild-type placenta. The investigation's findings align with the current literature, suggesting accurate results. However, there is scope for further research to refine the measurement methodologies, particularly for the junctional zone.

It is important to recognize the significance of the mouse strain used in this study. The C57BL/6 strain, a commonly used genetic background in biomedical research, offers a consistent baseline for comparison across genetic studies due to its well-documented genome and typical phenotypic responses. However, mouse strain differences can influence placental and embryonic development, potentially impacting the translatability of findings to other strains or species. Thus, when extending this research or comparing these findings, the specific characteristics

**Table 6.2:** Characterisation of the Wild-Type Placenta at E14.5

| Key | Mean | Std Dev | Median | Min | Max |
|---|---|---|---|---|---|
| nuc_counts | 34273.0 | 3473.0 | 34058.0 | 28176.0 | 44195.0 |
| ChP_count | 3399.0 | 5759.0 | 2475.0 | 829.0 | 39839.0 |
| ChP_ratio | 0.1 | 0.14 | 0.07 | 0.03 | 0.96 |
| GIA_count | 1185.0 | 464.0 | 1149.0 | 72.0 | 3103.0 |
| GIA_ratio | 0.03 | 0.01 | 0.03 | 0.0 | 0.08 |
| SPO_count | 3206.0 | 1005.0 | 3209.0 | 172.0 | 5786.0 |
| SPO_ratio | 0.09 | 0.03 | 0.1 | 0.0 | 0.15 |
| GLY_count | 4014.0 | 1490.0 | 4049.0 | 192.0 | 7578.0 |
| GLY_ratio | 0.12 | 0.04 | 0.12 | 0.0 | 0.23 |
| DEC_count | 8301.0 | 2803.0 | 7971.0 | 605.0 | 23082.0 |
| DEC_ratio | 0.24 | 0.09 | 0.24 | 0.02 | 0.8 |
| FET_count | 4493.0 | 1536.0 | 4268.0 | 197.0 | 9660.0 |
| FET_ratio | 0.13 | 0.04 | 0.13 | 0.0 | 0.28 |
| LAB_count | 6380.0 | 2177.0 | 6567.0 | 25.0 | 11767.0 |
| LAB_ratio | 0.19 | 0.06 | 0.19 | 0.0 | 0.35 |
| END_count | 3299.0 | 1120.0 | 3353.0 | 151.0 | 5536.0 |
| END_ratio | 0.1 | 0.03 | 0.1 | 0.0 | 0.16 |
| area_placenta | 29.02 | 3.18 | 28.87 | 21.64 | 40.26 |
| minor_placenta | 3.14 | 0.41 | 3.11 | 2.36 | 4.43 |
| major_placenta | 7.39 | 0.53 | 7.36 | 6.12 | 9.03 |
| area_JZ | 8.87 | 2.65 | 9.16 | 0.05 | 17.42 |
| minor_JZ | 2.95 | 0.66 | 2.96 | 0.18 | 5.13 |
| major_JZ | 6.36 | 1.26 | 6.56 | 0.26 | 8.36 |
| area_LAB | 11.76 | 2.77 | 11.94 | 0.01 | 18.18 |
| minor_LAB | 1.92 | 0.4 | 1.95 | 0.09 | 2.82 |
| major_LAB | 5.56 | 1.04 | 5.66 | 0.1 | 7.48 |

This table presents a characterisation of the wild-type placenta. The data were derived from 144 randomly selected E14.5 C57BL/6 samples, focusing on cell ratios, placental histology section area, and the placenta's major and minor axes.

of the C57BL/6 strain must be considered. This acknowledgment underpins the importance of strain selection in genetic research and its implications for interpreting and generalizing results in future work.

### 6.4.3 Investigation 3: Sexual Dimorphism in Wild-type

**Aim:** The investigation aimed to examine sexual dimorphism in wild-type placental samples, specifically identifying any significant variations between male and female placentas.

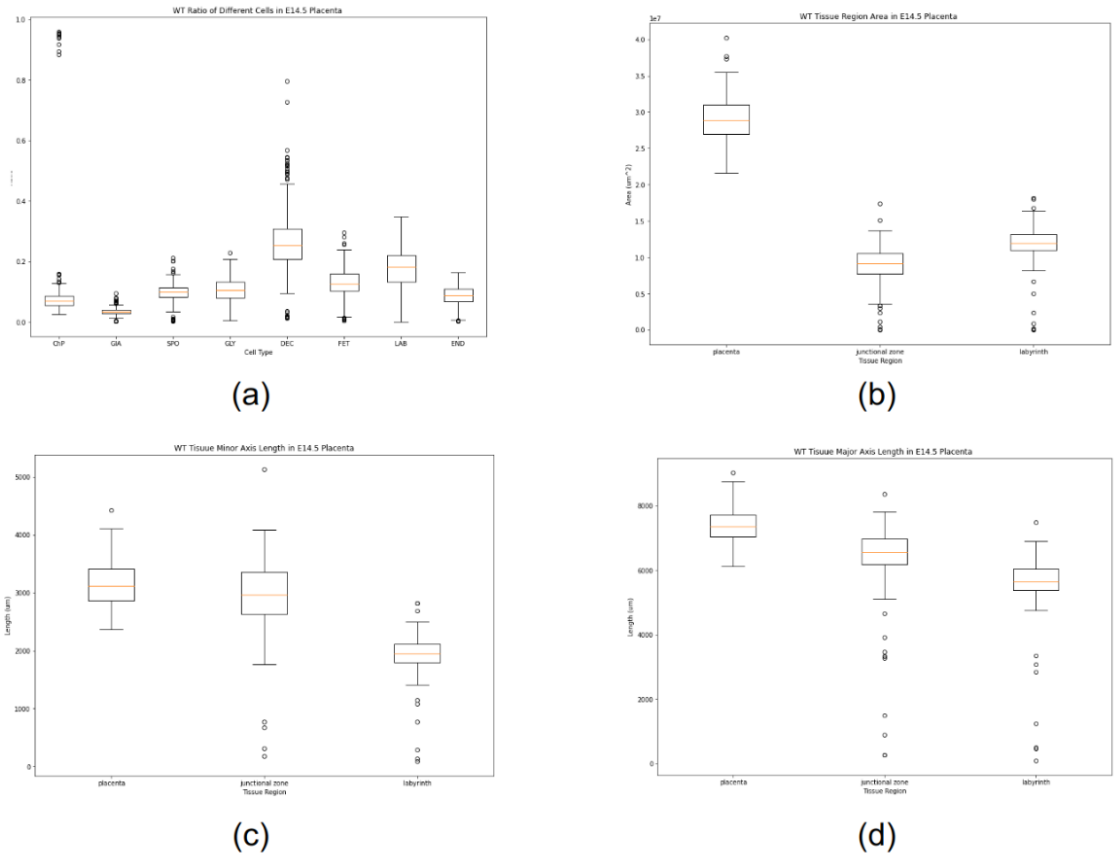



**Figure 6.6:** *This figure presents the characterisation of the wild-type placenta, derived from the analysis of 144 randomly selected E14.5 C57BL/6 samples. Four box plots are displayed: a) Cell ratios of the eight different mouse placental classes, showing the relative abundance of each cell type in the tissue; b) The area of the tissue regions in wild-type placentas, demonstrating the sizes of the placenta, junctional zone, and labyrinth; c) The minor axis measurements of the placenta, junctional zone, and labyrinth, indicating their respective widths; d) The major axis measurements of the placenta, junctional zone, and labyrinth, providing an estimate of their respective lengths. These visualisations offer a comprehensive insight into the wild-type placenta's characteristics, providing a crucial baseline for further comparative studies.*

**Method:** The study compared cell ratios, placental and tissue region areas, and the major and minor axes between E14.5 wild-type male and female placentas from the DMDD dataset. Averages were compared and t-tests were applied to identify statistically significant variations.

**Result:** Statistically significant variation was observed in the placental area, the area of the junctional zone, and the number of glycogen cells, with males generally having larger measurements. In particular, the male placenta appeared larger along its minor axis than its major axis, suggesting a potential nutrient sensitivity and a correlation with the typically larger birth size of male foetuses.

**Conclusion:** The study identified significant differences between male and female wild-type mouse placentas, corroborating existing literature on sexual dimorphism in the species. These findings underscore the importance of considering sex as a variable in placental research, as it can significantly impact the interpretation and generalisability of the results.

**Future Work:** Future research should continue to explore the implications of sexual dimorphism on placental development and function, as well as the potential links to susceptibility to adverse pregnancy outcomes. Further investigation into the role of glycogen cells in late foetal growth could also provide valuable insights.

This investigation aimed to examine sexual dimorphism in wild-type placental samples. Sexual dimorphism, the occurrence of distinct differences between males and females of the same species beyond reproductive organs, is a widespread biological phenomenon [130]. This investigation compared averages and applied t-tests to identify statistically significant variations in male and female placentas.

Males tend to be larger in the mouse species and exhibit darker fur colour than females. Behaviourally, male mice are typically more aggressive and territorial, while females display more passive behaviours and often form social groups.

Until recently, the placenta was considered an asexual organ, leading to many studies overlooking the sex of the embryo. However, contemporary research has highlighted the sex-based differences in placental size, with male placentas generally larger than female placentas. This difference is believed to result in males being larger during gestation but makes them more susceptible to adverse pregnancy outcomes such as preterm birth and low birth weight [131].

Cell ratios, placental and tissue region areas, and the major and minor axes were compared between E14.5 wild-type male and female placentas from the DMDD dataset in this study. Statistically significant variation was observed in the placental area, the area of the junctional zone, and the number of glycogen cells, with males exhibiting larger measurements in all these areas.

The results corroborate existing literature, suggesting that the larger male placenta supports male foetuses' typically larger birth size. Interestingly, the male placenta appeared larger along its minor axis than its major axis. This finding echoes Barker et al.'s (2000) suggestion that growth along the minor axis

may be more nutrient-sensitive, providing a potential explanation for the greater susceptibility of male foetuses to adverse pregnancy outcomes related to growth and maternal environmental disturbances.

Variation between male and female placentas was also notable in the junctional zone along the minor axis. This variation might be attributed to the statistically significant difference in glycogen cell count. Given that glycogen cells play a crucial role in late foetal growth and males typically grow larger, it is plausible that males require larger glycogen stores. This finding also suggests the heightened sensitivity of male foetuses to adverse pregnancy outcomes related to growth and disruptions in the maternal environment, such as food shortages.

This study identified notable differences between male and female wild-type mouse placentas, aligning with the current understanding of sexual dimorphism in the species. These findings highlight the importance of considering sex as a variable in placental research, as it can significantly impact the interpretation and generalisability of the results.

### 6.4.4 Investigation 4: Phenotypic Analysis of E14.5 Gene Knockout Mouse Models

**Aim:** The primary objective of this investigation was to identify gene knockouts that manifest statistically significant phenotypes, thereby validating the pipeline's functionality and contributing to the understanding of developmental disorders.

**Method:** The investigation involved analysing 66 knockouts from 310 individual placental samples using t-tests and Bonferroni corrections for multiple testing to identify statistically significant phenotypes.

**Result:** The study identified 28 knockouts with statistically significant phenotypes. Notably, there was considerable overlap between these findings and those reported in the literature by the Hemberger lab and the Mouse Genome Informatics (MGI) database.

**Conclusion:** The findings suggest that the pipeline accurately detects relevant phenotypes, providing more comprehensive information than traditional methods. This heightened sensitivity allows for the identification of more specific phenotypes and has the potential to improve the understanding of the genetics of placental diseases.

**Table 6.3:** Sexual Dimorphism in Wildtype E14.5 Mouse Placentas

| | Mean M | Std Dev M | Mean F | Std Dev F | P-Val | Corrected P-Val |
|---|---|---|---|---|---|---|
| ChP_count | 4701.79 | 8424.06 | 3161.74 | 4604.66 | 0.27 | 6.99 |
| GIA_count | 1202.63 | 572.58 | 1266.55 | 444.61 | 0.54 | 14.03 |
| SPO_count | 3229.15 | 1136.96 | 2980.94 | 943.64 | 0.24 | 6.32 |
| DEC_count | 8066.58 | 3287.74 | 8101.72 | 2341.82 | 0.95 | 24.75 |
| FET_count | 4264.17 | 1899.18 | 4521.15 | 1237.37 | 0.43 | 11.24 |
| LAB_count | 6119.08 | 2476.83 | 6355.51 | 2043.69 | 0.61 | 15.8 |
| END_count | 3315.15 | 1254.83 | 3494.17 | 930.81 | 0.43 | 11.08 |
| GLY_count | 4359.35 | 1809.49 | 3713.89 | 1314.1 | 0.05 | 1.22 |
| ChP_ratio | 0.13 | 0.2 | 0.09 | 0.12 | 0.33 | 8.64 |
| GIA_ratio | 0.03 | 0.01 | 0.04 | 0.01 | 0.19 | 4.95 |
| SPO_ratio | 0.09 | 0.03 | 0.09 | 0.02 | 0.47 | 12.28 |
| DEC_ratio | 0.23 | 0.11 | 0.24 | 0.07 | 0.62 | 16.07 |
| FET_ratio | 0.12 | 0.05 | 0.13 | 0.04 | 0.12 | 3.09 |
| LAB_ratio | 0.18 | 0.07 | 0.19 | 0.06 | 0.29 | 7.62 |
| END_ratio | 0.09 | 0.03 | 0.1 | 0.03 | 0.12 | 3.04 |
| GLY_ratio | 0.12 | 0.05 | 0.11 | 0.04 | 0.14 | 3.75 |
| area_placenta | 30046471.88 | 3529475.12 | 28326340.96 | 2887783.43 | 0.01 | 0.25 |
| minor_placenta | 3228.79 | 436.28 | 3060 | 396.54 | 0.05 | 1.24 |
| major_placenta | 7496.48 | 559.78 | 7331.83 | 432.22 | 0.11 | 2.79 |
| area_JZ | 9298433.1 | 3210843.46 | 7917646.49 | 2450141.36 | 0.02 | 0.49 |
| minor_JZ | 2977.27 | 845.43 | 2883.66 | 641.68 | 0.54 | 14.03 |
| major_JZ | 6274.56 | 1662.7 | 6219.47 | 1119.99 | 0.85 | 22.07 |
| area_LAB | 11487463.94 | 3678579.91 | 11933380.57 | 2558624.2 | 0.49 | 12.74 |
| minor_LAB | 1885 | 532.36 | 1927.77 | 351.41 | 0.64 | 16.69 |
| major_LAB | 5394.46 | 1448.14 | 5590.09 | 844.62 | 0.42 | 10.92 |

This table demonstrates the sexual dimorphism observed between male and female wildtype E14.5 placentas. The analysis provides insights into variations in cell counts, ratios, and placental dimensions.

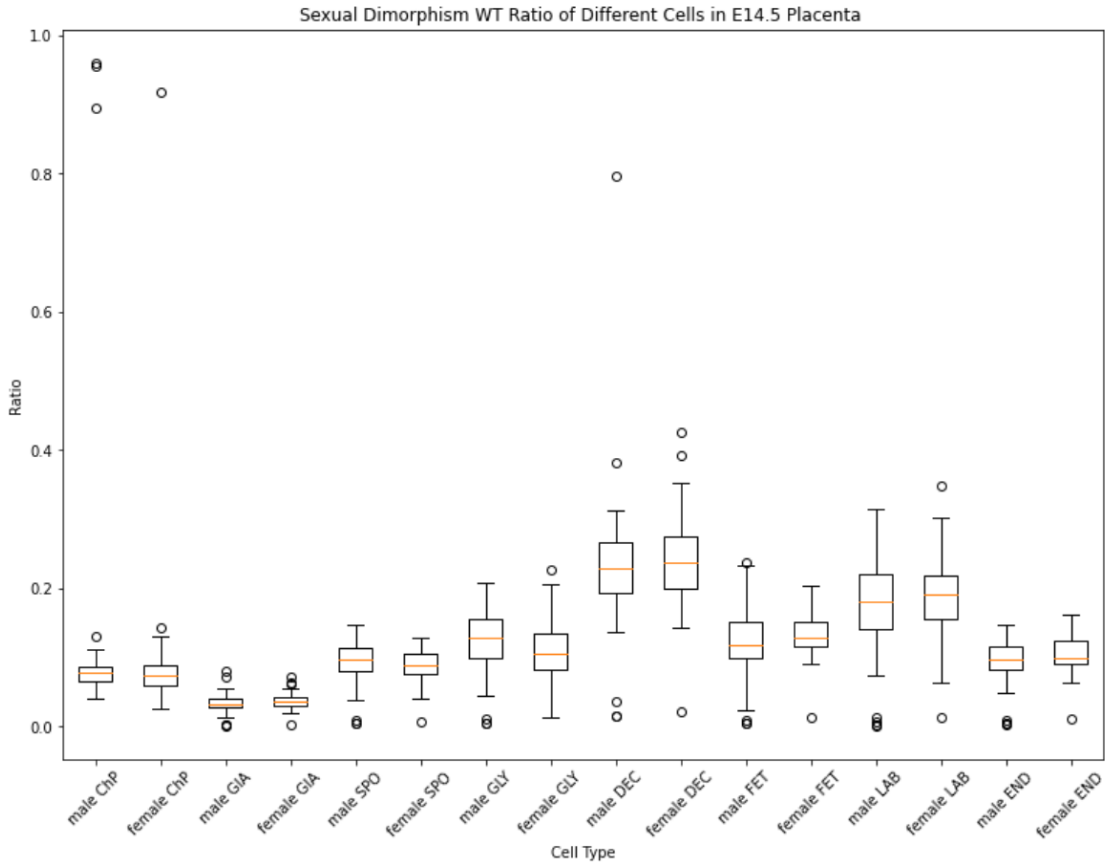



**Figure 6.7:** *This figure provides a visual representation of the sexual dimorphism observed in the cell ratios of the mouse placenta. The box plots depict the mean errors for each cell type in both male and female mice. The plots reveal distinct differences between the sexes, indicating the presence of sexual dimorphism in placental cell ratios. For instance, male placentas tend to have a higher count of glycogen cells, which are crucial for foetal growth. These findings provide valuable insights into the sex-specific characteristics of mouse placentas and highlight potential avenues for further research, such as the impact of placental sex on pregnancy outcomes and susceptibility to environmental disruptions.*

**Future Work:** Future research could focus on evaluating the pipeline's ability to uncover novel genetic targets for placental phenotypes and further refining the methodology for defining statistical significance.

The primary objective of this investigation is to identify gene knockouts that manifest statistically significant phenotypes. These results can validate the pipeline's functionality and contribute to the understanding of developmental disorders.

The investigation involves the application of statistical t-tests and Bonferroni corrections for multiple testing to identify statistically significant phenotypes in genetic knockouts. From the analysis of 66 knockouts derived from 310 individual placental samples, 28 statistically significant phenotypes were identified. Notably,

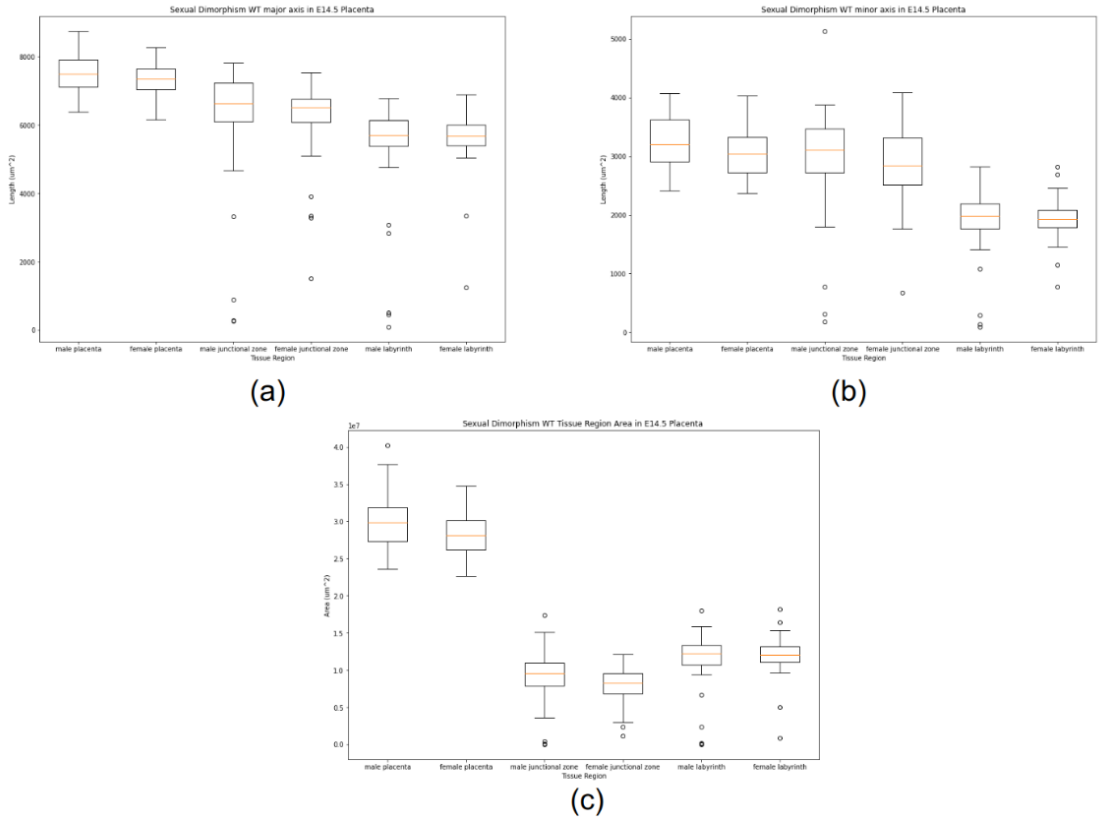



**Figure 6.8:** *This figure presents an analysis of the major and minor axes and area of the placenta, junctional zone, and labyrinth in E14.5 mice. Panel A) displays box plots of the major axis measurements, panel B) shows the minor axis measurements, and panel C) depicts the area measurements. The measurements highlight significant sexual dimorphism within the studied parameters. Notably, male placentas tend to be larger, particularly along the minor axis, and exhibit a higher nuclei count, as depicted in panel B). These observations align with existing literature and suggest potential sensitivity of the minor axis to nutrient availability. Variations in the junctional zone's size, primarily along the minor axis, could be attributed to the statistically significant difference in glycogen cell count, as glycogen cells play a critical role in late foetal growth. These findings provide valuable insights into the sex-specific morphological differences in E14.5 mouse placentas.*

there was considerable overlap between these findings and those reported in the literature by the Hemberger lab and the Mouse Genome Informatics (MGI) database.

Genetic knockout models in mice serve as a formidable tool to decipher the genetics of foetal development. In these models, a target gene is 'knocked out' or deactivated in a mouse embryo. The ensuing developmental changes are studied, offering insights into the gene's role in normal foetal development [132].

This thesis phase focuses on the phenotypic analysis of E14.5 lethal and subviable

| | p-value |
|---|---|
| nuc_counts_p_value | 0.0225257 |
| ChP_count_p_value | 0.268968 |
| GIA_count_p_value | 0.53958 |
| SPO_count_p_value | 0.242926 |
| DEC_count_p_value | 0.951744 |
| FET_count_p_value | 0.432362 |
| LAB_count_p_value | 0.607827 |
| END_count_p_value | 0.426151 |
| GLY_count_p_value | 0.0469858 |
| ChP_ratio_p_value | 0.332267 |
| GIA_ratio_p_value | 0.190308 |
| SPO_ratio_p_value | 0.47215 |
| DEC_ratio_p_value | 0.617998 |
| FET_ratio_p_value | 0.118687 |
| LAB_ratio_p_value | 0.293045 |
| END_ratio_p_value | 0.116766 |
| GLY_ratio_p_value | 0.144334 |
| area_placenta_p_value | 0.00974201 |
| minor_placenta_p_value | 0.0475574 |
| major_placenta_p_value | 0.107336 |
| area_JZ_p_value | 0.0189726 |
| minor_JZ_p_value | 0.539682 |
| major_JZ_p_value | 0.848731 |
| area_LAB_p_value | 0.489923 |
| minor_LAB_p_value | 0.64198 |
| major_LAB_p_value | 0.420081 |

**Figure 6.9:** *This figure presents a table of p-values for various phenotypes, illustrating the statistical significance of observed sexual dimorphism in E14.5 mice. The phenotypes analysed include nuclei counts, glycogen cell count, placental area, placental minor axis, and junctional zone area. The table underscores that all these phenotypes exhibit statistically significant differences between male and female placentas, thus providing strong quantitative evidence of sexual dimorphism. This analysis aligns with existing literature and expands our understanding of sex-specific morphological variations in the mouse placenta.*

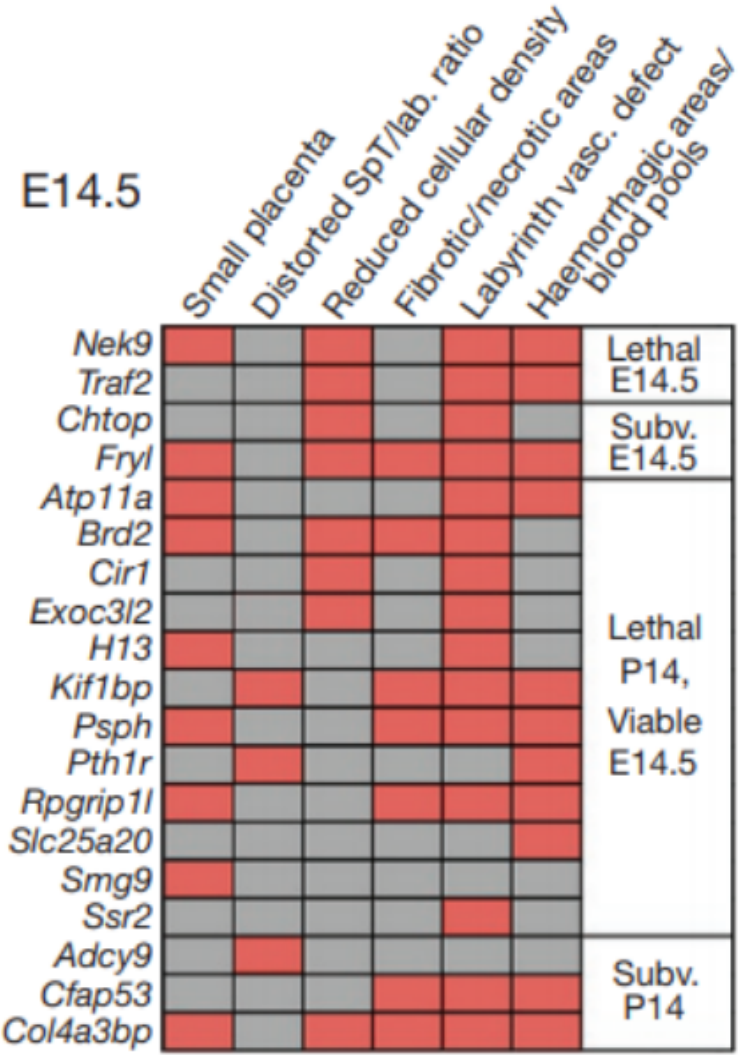



**Figure 6.10:** *This figure, adopted from the study by Perez-Garcia et al., presents the observed E14.5 placental phenotypes in various mutant mouse strains. It comprises examples of histological analyses of E14.5 mutant placentas, highlighting key structural variations and abnormalities. These visualisations provide valuable insights into the significant alterations in placental structure and function associated with specific genetic mutations. They show overlap with the placental phenotyping pipeline from chapter one proving it is an effective tool.*

knockout data, analysing 66 knockouts from 310 individual placental samples. The research methodology involved the application of t-tests to identify variations between the wild-type average and the knockout average, with Bonferroni corrections employed to adjust for multiple testing.

Assessing the overlap between results obtained by this automated pipeline and those manually analysed in the Hemberger lab's paper is crucial. Hemberger et al. (2019) reported phenotypic similarities in E14.5 knockouts, including placenta size, distorted structural ratios, reduced cellular density, and labyrinth vasculature defects.

Compared to the 19 knockout strains with phenotypes identified by the Hemberger lab, this study identified 28 knockouts with statistically significant phenotypes. This discrepancy can be attributed to the more comprehensive pipeline employed here, measuring 32 different placental features for a phenotype, in contrast to the six measurements taken by Hemberger et al. (2019). Moreover, this study employed a more rigorous testing criteria.

Regarding the overlap between the Hemberger lab's results and this study, eleven gene knockouts with comparable phenotypes were identified, indicating a 58% overlap rate. This result suggests the pipeline's effectiveness in measuring phenotypes. The missing phenotypes might be due to the stringent nature of the Bonferroni correction when defining statistical significance. Implementing a less conservative multiple testing correction, such as the False Discovery Rate (FDR), could alleviate this issue. However, given the abundance of results already, retaining the conservative estimates of statistical significance was deemed appropriate.

The Mouse Genome Informatics (MGI) database, an international repository of genetic, genomic, and biological data, was explored to validate the findings further. This research focused on the array of phenotypes for each genetic knockout. Eleven overlapping knockouts were identified in the MGI, in line with the nine found by the Hemberger lab in the DMDD dataset, reinforcing the pipeline's reliability. However, questions arose regarding the pipeline's ability to uncover new phenotypes.

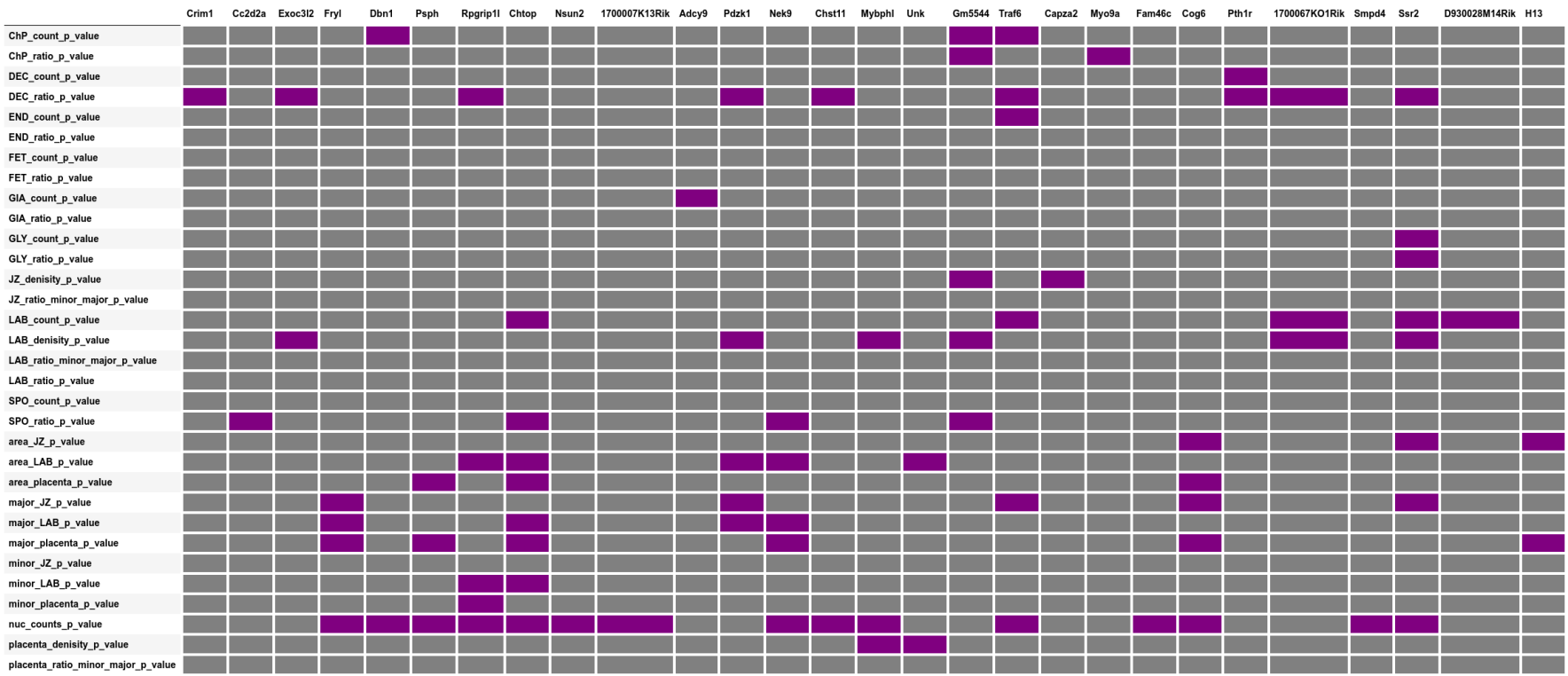



**Figure 6.11:** *This figure displays a heatmap of statistically significant phenotypes for different gene knockouts. The analysis covered 66 knockouts from 310 individual placental samples, revealing 28 statistically significant phenotypes. The heatmap offers a visual representation of these 28 knockouts with phenotypes, revealing considerable overlap with findings reported in the literature. This overlap serves as a validation of the pipeline's functionality and its utility in enhancing our understanding of developmental disorders.The purple boxes represnt statistically significant features after correction for multiple tests. Grey means no significance.*

To evaluate the pipeline's ability to uncover novel genetic targets for placental phenotypes, an analysis was conducted to determine whether placental-related phenotypes were present in the surviving pups or foetuses. The analysis revealed a significant overlap in placental phenotypes between the pups and foetuses, with the exception of smpd4. These findings suggest that the automated analysis pipeline effectively identified known placental phenotypes and potentially uncovered new targets for further research. This is show in **??**

The pipeline's sensitivity was further assessed by analysing the overlap between actual placental phenotypes recorded in the MGI and those detected by the pipeline. The results demonstrated that the pipeline accurately detected phenotypes associated with various issues, such as distorted decidual cell ratios (cr1) and distorted labyrinth cell density (Exoc3l2). Additionally, the pipeline was able to detect phenotypes associated with a small placenta in fry1, rpgrip1l, nek9, and h13, as well as the ratio of spongiotrophobasts in chop.

As discussed earlier, our findings revealed substantial overlap in phenotypes between this work, the Hemberger lab, and the Mouse Genome Informatics (MGI) database. These overlaps, illustrating 'matching or related phenotypes', are clearly depicted in Figure **??**, which presents a comparative analysis demonstrating the consistency across different studies. This visual representation supports the textual discussion and provides direct insight into the similarities and discrepancies in phenotypic expression observed across different datasets.

The findings suggest that the pipeline accurately detects relevant phenotypes, providing more comprehensive information than traditional methods. This heightened sensitivity allows for the identification of more specific phenotypes and has the potential to improve the understanding of the genetics of placental diseases.

The term 'more rigorous testing criteria' in this context refers to the application of Bonferroni corrections to adjust for multiple comparisons in our statistical analyses. This correction method is considered stringent as it reduces the risk of type I errors, which occur when a true null hypothesis is incorrectly rejected. However, while Bonferroni corrections help ensure robustness in findings by controlling the

| Gene KO | Placental Phenotype MGI | Placental Phenotype Hem | Placental related Phenotypes | Specific placental phenotypes |
|---|---|---|---|---|
| cr1 | | | Neurological, Body size | Decidual layer |
| cc2d2a | | | Body size, Cardiac function | |
| exoc3l2 | | | Body size, Cardiac function | Labyrinth Morphology |
| fryl | | | Body size | Labyrinth Morphology, small placenta |
| dbn1 | | | Body size, Cardiac function, Neurological | |
| psph | | | Body size, Cardiac function, Neurological | Labyrinth, small placenta |
| rpgrip1l | | | Body size, Cardiac function, Neurological | Labyrinth, small placenta |
| chtop | | | Body size, Cardiac function, Neurological | Labyrinth, spongiotrophoblast |
| nsun2 | | | Body size, Cardiac function, Neurological | |
| pierce1 | | | Cardiac function | |
| adcy9 | | | Cardiac function | |
| pdzk1 | | | Neurological | |
| nek9 | | | Body SIze | Labyrinth, Trophoblast, Small placenta |
| chst11 | | | Body size, Cardiac function | |
| mybphl | | | Neurological | |
| unk | | | Neurological, Body size | |
| gm5544 | | | Neurological | |
| traf6 | | | Body size, Cardiac function, Neurological | |
| capza2 | | | Cardiac function | |
| myo9a | | | Body size, Cardiac function, Neurological | |
| fam46c | | | Cardiac function, Neurological | |
| cog6 | | | Cardiac function | |
| pth1r | | | Body size, Cardiac function, Neurological | |
| smpd4 | | | No related | |
| ssr2 | | | Body size, Cardiac function, Neurological | |
| d930028m14rik | | | Cardiac function | |
| h13 | | | Body size, Cardiac function, Neurological | Labyrinth, Trophoblast, Small placenta |
| Total | 11 | 9 | | |

**Figure 6.12:** *This figure compares the results of the automated placental phenotyping pipeline with those from the Mouse Genome Informatics (MGI) database and the Hemberger lab. It presents the placental phenotypes of various gene knockouts, along with their corresponding phenotypes in the MGI and Hemberger lab's datasets. The figure also highlights the specific placental phenotypes detected by the pipeline, offering a comprehensive overview of the pipeline's capabilities in identifying known and potential new placental phenotypes.*

family-wise error rate, they can also increase the likelihood of type II errors, where a false null hypothesis is accepted. An alternative approach, such as the False Discovery Rate (FDR) correction, could potentially be less conservative, increasing the power of the study to detect true effects amongst multiple tests. Future research might benefit from comparing these methods to determine the most effective balance between sensitivity and specificity in the context of high-dimensional genetic data.

### 6.4.5 Investigation 5: Predicting Lethality Outcomes from Placental Phenotypes

**Aim:** The aim of this investigation was to ascertain the predictive capability of placental phenotypes in determining lethality outcomes and to identify the most significant features in making these predictions.

**Method:** Random Forest and XGBoost models were employed, utilising Gini importance and drop column importance to determine the significance of each feature. The model was initially trained on three outcome classes (lethal, subviable, and viable) and later simplified to two classes (lethal and viable).

**Result:** An accuracy of 63% was achieved for the three-class model and 70% for the binary model across a ten-fold validation. The labyrinth area emerged as the most significant feature in these predictions.

**Conclusion:** The findings suggest that placental phenotypes can predict lethality outcomes to a reasonable extent, with certain features such as the labyrinth area playing a significant role. However, further refinement of the model and handling of class imbalances are necessary to improve its predictive accuracy.

**Future Work:** Future research could focus on refining the prediction models, exploring oversampling techniques to handle class imbalances, and investigating the potential effects of collinearity among the features.

The primary objective of this investigation is to ascertain the predictive capability of placental phenotypes in determining lethality outcomes and to identify the most significant features in making these predictions.

Random Forest and XGBoost models were employed in this investigation to predict lethality outcomes. Gini importance and drop column importance were utilised to ascertain the significance of each feature in the prediction process.

An accuracy of 63% was achieved across a ten-fold validation in predicting the outcomes of lethal, subviable, and viable embryos. The labyrinth area emerged as the most significant feature in making these predictions.

The findings suggest that placental phenotypes can predict lethality outcomes to a reasonable extent and that certain features are more significant than others in making these predictions. However, further research is required to refine this model and enhance its predictive accuracy.

The Deciphering the Mechanisms of Developmental Disorders (DMDD) programme offers an opportunity to investigate the co-association between specific placental phenotypes and lethality outcomes. In particular, it allows for an exploration of which placental phenotypes contribute the most to adverse lethality

outcomes. This research area could be extended to predict outcomes based on placental phenotypes and specific foetal defects such as cardiac defects.

In this investigation, tree-based classification models were trained, specifically the random forest and XGBoost models. The analysis utilised 283 knockout placental samples across three lethality outcomes at the E14.5 stage. This consisted of 151 viable, 87 lethal, and 45 subviable samples. The outcomes were derived from the DMDD embryo analysis paper [107] and incorporated into the data with placental phenotypes. The data was then cleaned, numericalised, and normalised before being partitioned into training and test sets. The training set consisted of 226 samples, and the test set included 57 samples.

The differences between the random forest and XGBoost models utilised in this study are elaborated upon in methods section of this chapter. Each model's hyperparameters were tuned using a grid search approach [102]. This method generates a grid of hyperparameters and trains different models with varying hyperparameter combinations. The model with the highest accuracy is then selected as the final model. This method is exhaustive, as it tries all possible combinations of hyperparameters. While it can be computationally expensive, it is an effective way to find the best combination of hyperparameters for a given model and can help improve model performance.

Validation was conducted using a ten-fold validation approach. Multi-fold validation is a valuable step in machine learning that reduces the risk of overfitting. It involves dividing the data into multiple subsets and running several iterations of the model on each subset. By testing the model on different data points and comparing the results, the model can be evaluated against data it has not seen before, which helps identify any flaws in the model and prevents it from overfitting. The results of the multiple runs can be used to assess the model's generalisability and accuracy in predicting unseen data. This step is crucial in ensuring the model will produce reliable results when deployed.

In this study, 'accuracy' refers to the percentage of correct predictions made by the model, including correct identifications of both positive and negative

outcomes. Our goal is to continuously improve this metric, striving for the highest possible accuracy to ensure the model's reliability for clinical and research purposes. Improvements may include utilizing advanced machine learning techniques or integrating more comprehensive datasets. However if we acheive greater than baseline we can say we are moving in the right direction.

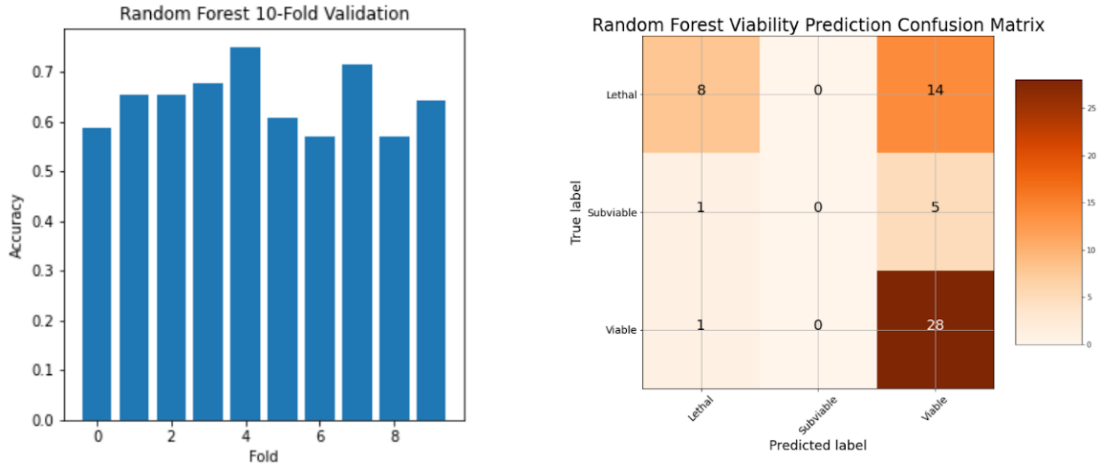
The initial results were obtained using a random forest model, which yielded a 64% accuracy in ten-fold validation. However, as depicted in the confusion matrix in figure 6.13, the results are heavily skewed towards the viable outcome, with zero appearing in the subviable category. This indicates that there is a need for more representation of subviable data. The same imbalance is evident in the XGBoost model, which also reports a 64% ten-fold accuracy heavily skewed towards the viable outcome. To rectify this imbalance, data oversampling methods were applied.

Oversampling is a technique used in machine learning to address class imbalance in datasets. Class imbalance occurs when one class significantly outnumbers others in a dataset, which can lead to poor performance of machine learning models when predicting the minority classes.

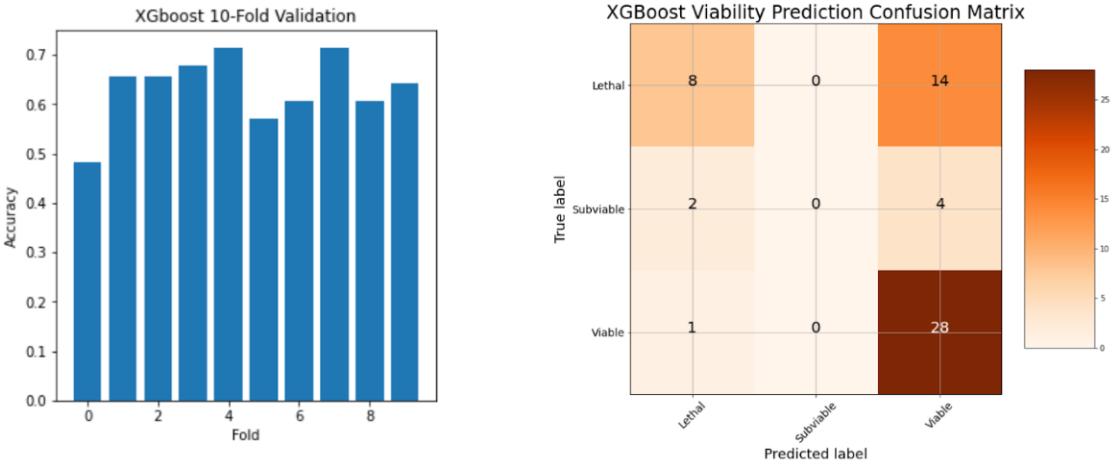
Random oversampling is a straightforward technique that randomly duplicates examples from the minority class, thus increasing the number of examples in the dataset. Although simple, it can lead to overfitting of the data. Synthetic Minority Oversampling Technique (SMOTE) is a more sophisticated technique that creates synthetic examples of the minority class. It does this by randomly selecting a minority class example and its k-nearest neighbours, then generating a new example somewhere between them [133].

There are several other oversampling techniques, such as ADASYN (Adaptive Synthetic Sampling Approach for Imbalanced Learning), which assigns different weights to the examples in the minority class, and Borderline SMOTE, which creates synthetic examples based on the borderline between minority and majority classes. However, in this study, the focus was on random oversampling and SMOTE.

The random oversampling approach generates a balanced dataset, which breaks down into 366 training samples and 87 validation samples. The results from the




**(i)** Random Forest Classification Results

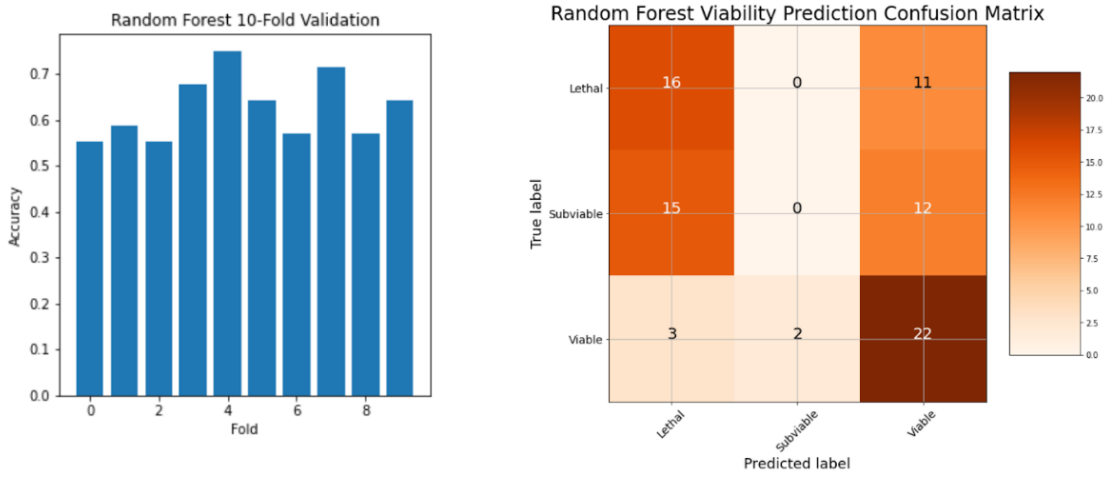



**(ii)** XGBoost Classification Results

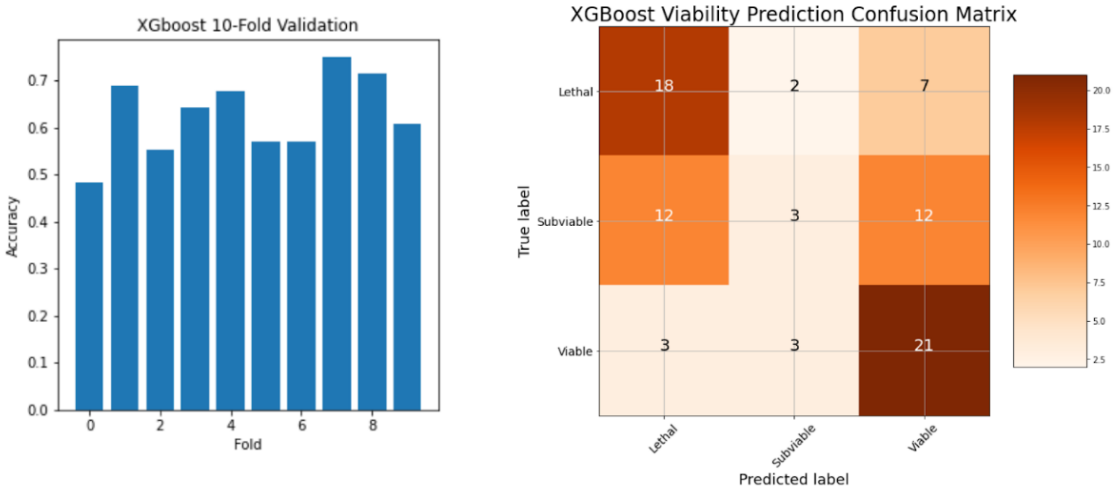
**Figure 6.13:** *These figures present the results of the Random Forest (a) and XGBoost (b) classification models predicting lethality outcomes of lethal, subviable, and viable. The heatmaps represent confusion matrices that visualise the performance of the models across a ten-fold validation. Both models show a prediction bias towards the 'viable' outcome, each with a reported accuracy of 64%.*

random forest method are shown in figure 6.14. The ten-fold validation returned 63% accuracy. However, when we examine the confusion matrix, there are still misclassification issues, with many samples categorised as viable instead of lethal. Moreover, the subviable column remains zero. The same issues occur when the XGBoost method is used. Figure 6.14 again shows a 63% ten-fold validation accuracy, albeit with slightly better results in the subviable column.

Using the SMOTE algorithm, an oversampling approach was applied, and two




**(i)** Random Forest Classification with Oversampling




**(ii)** XGBoost Classification with Oversampling

**Figure 6.14:** *These figures present the results of the Random Forest (a) and XGBoost (b) classification models with random oversampling predicting lethality outcomes of lethal, subviable, and viable. The heatmaps represent confusion matrices that visualise the performance of the models across a ten-fold validation. After applying oversampling, the models show a slight improvement in prediction accuracy but still display misclassifications, particularly in the subviable category.*

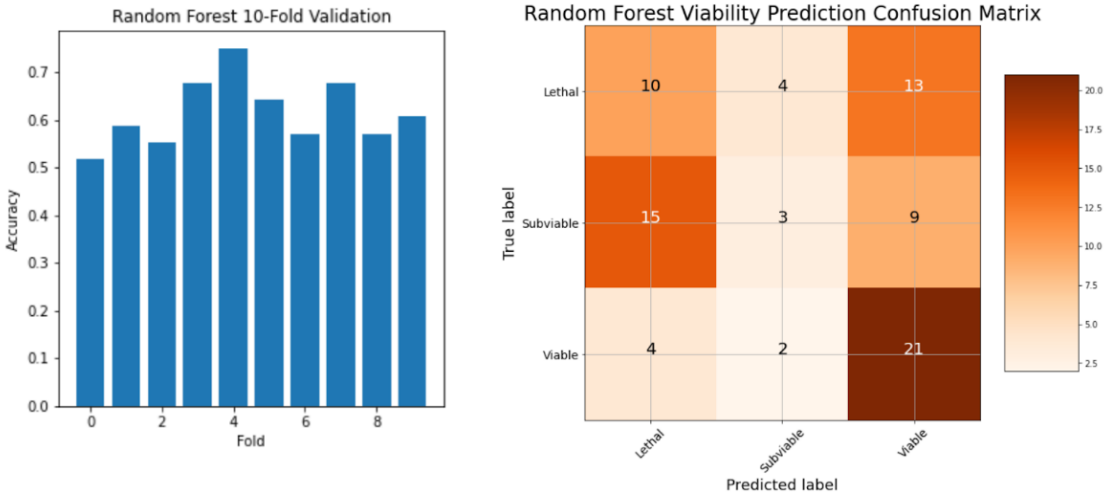
models were trained: a random forest model and an XGBoost model. The random forest model achieved 62% accuracy (Figure 6.15), and the XGBoost model achieved 63% (Figure 6.15). Despite the marginally improved performance, the results still needed improvement, suggesting that the models needed help separating the classes of viable and lethal outcomes. To explore this further, it was hypothesised that results might be improved by reducing the classes to viable and lethal categories.

To simplify the problem, it was converted into a binary classification problem, in which all subviable outcomes were classified as lethal. This was justified because subviable lines are likely to present abnormal placental morphology. Additionally, this process balanced the data points for each class, with 151 viable placentas and 132 lethal placenta samples. An XGBoost model was trained due to its superior performance, resulting in a 10-fold classification accuracy of 70%. The confusion matrix revealed that the results were now skewed towards lethal. In this work, a random forest and oversampling of the datasets were used, with little difference. The results of random oversampling (72%) and SMOTE (73%) in ten-fold validation were similar (Figure 6.16). Thus, the XGBoost model with three classifications from figure 6.14 was chosen for feature analysis.

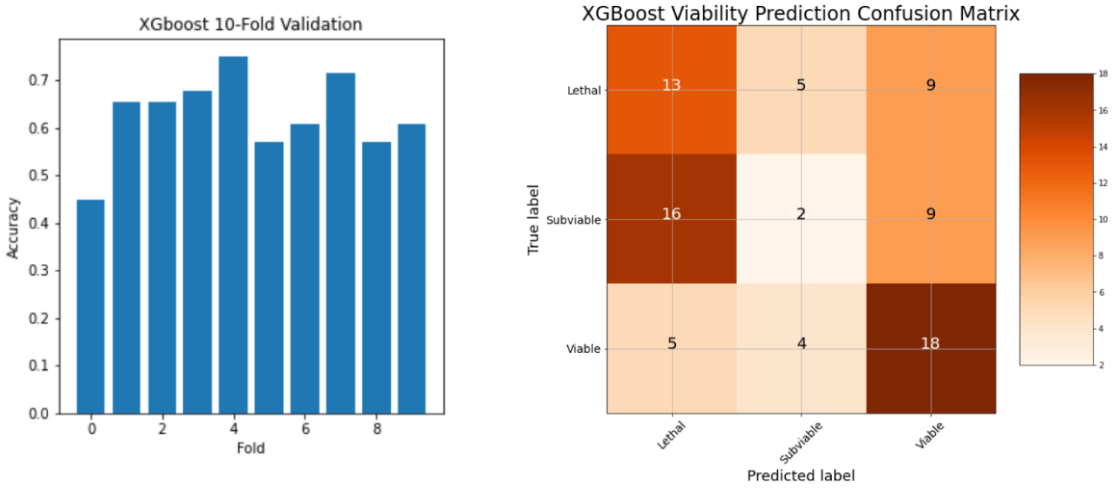
### 6.4.6 Feature Analysis

Understanding the most important features is crucial when working with machine learning models. A better understanding of the underlying data can be achieved by determining the most important features. In the context of the placenta, understanding which features contribute most towards predictions of negative outcomes could provide valuable information in better understanding congenital disease.

There are three methods of measuring feature importance in tree-based models. This study utilised two methods: Gini Importance and Drop Column Importance. Gini importance is a feature importance measure calculated by evaluating the decrease in node impurity from splitting on the feature. It is calculated by subtracting the proportion of samples that would be incorrectly classified if a particular split is made from those that would be incorrectly classified if the split




**(i)** Random Forest Classification with SMOTE




**(ii)** XGBoost Classification with SMOTE

**Figure 6.15:** *These figures present the results of the Random Forest (a) and XGBoost (b) classification models with Synthetic Minority Oversampling Technique (SMOTE) predicting lethality outcomes of lethal, subviable, and viable. The heatmaps represent confusion matrices that visualise the performance of the models across a ten-fold validation. SMOTE, a technique to address class imbalance by generating synthetic examples of the minority class, is applied to improve the performance of the models. Despite the use of SMOTE, the models still display misclassifications, indicating the complexity of the classification task and suggesting the need for further refinements in the models.*

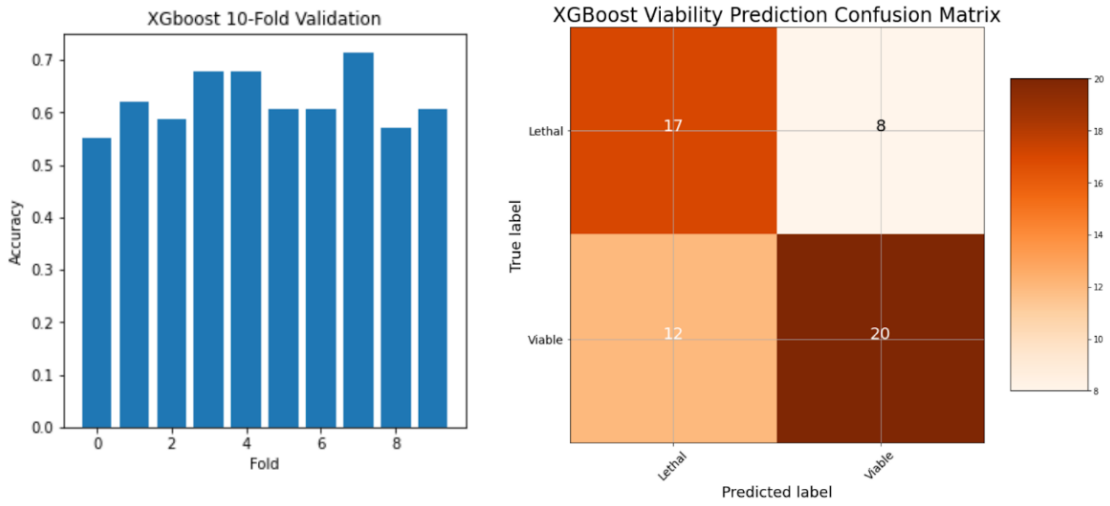



**Figure 6.16:** *This figure presents the results of the XGBoost binary classification model with oversampling predicting lethality outcomes of lethal and viable. The heatmap represents a confusion matrix that visualises the model's performance across a ten-fold validation. After transforming the problem into a binary classification task and applying oversampling techniques, the model shows improved performance with increased prediction accuracies. However, the results also indicate a skew towards the 'lethal' outcome by the model, suggesting the need for further refinements in the model.*

was not made [126]. Drop column importance is a measure of feature importance that is determined by comparing the out-of-sample accuracy of a model trained with all features included to the out-of-sample accuracy of the same model trained with one feature removed. This measure can help determine which features are most important to the model.

Two random number columns were added as a control when calculating the feature importance. Random number column 1 ranged from 1-10, and random number column 2 ranged from 1-1000. These should return zero or very small feature importance as they do not correlate with the result.

The feature importance was calculated on an XGBoost model using the three outcomes: lethal, subviable, and viable. The results for drop column importance are shown in figure 6.17. The first aspect to notice is that the random number columns have negative importance. This means the model performs better without them. This is an excellent sign regarding the feature being shown as important and correlated with the outcome. The most important feature was the labyrinth area. This is an expected result as this is the point of gas exchange. In the Gini

importance feature analysis (Figure 6.18), the labyrinth area again becomes the most important feature in predicting birth outcomes. In this analysis, we again see the random number columns without importance. A key conclusion we can therefore make is labyrinth area is the most critical predictor of birth outcome.

The methodologies for feature analysis do have notable limitations. The apparent disparity in feature importance depicted by different methods raises questions about the accuracy and reliability of these computations, thus inviting scrutiny over the methodological robustness. Discrepancies may be attributed to the need for improved model accuracy and the inherent ambiguities in defining 'viability'. With placentas lacking explicit labelling, the inferred outcomes could introduce unrecorded effects, complicating results interpretation and affecting the calculations of feature importance.

Moreover, the impact of colinearity is significant, arising when high correlations between features distort the coefficients and thus, complicate the understanding of individual feature impacts, potentially leading to model overfitting and compromising result reliability. To mitigate colinearity, correlated features were removed (Figure 6.19), enhancing model interpretability and ensuring insights are reflective of inherent data relationships.

## 6.5 Summary and Conclusion

This chapter has embarked upon an extensive evaluation of the phenotyping pipeline outlined in 5, meticulously comparing its outputs with the findings reported by the Hemberger lab based on the Deciphering the Mouse Developmental Disorders Dataset. Further, it has applied the pipeline to analyse the DMDD dataset, fulfilling the set objectives comprehensively.

The adaptability of point cloud-based methods, such as the concave hull for tissue region segmentation in E14.5 placenta, was demonstrated in the first investigation. While this approach did not yield optimal results with the E9.5 data, this was attributed to misclassifications in cell class carried over from Chapter 5. This issue could be remedied by developing an E9.5-specific model.

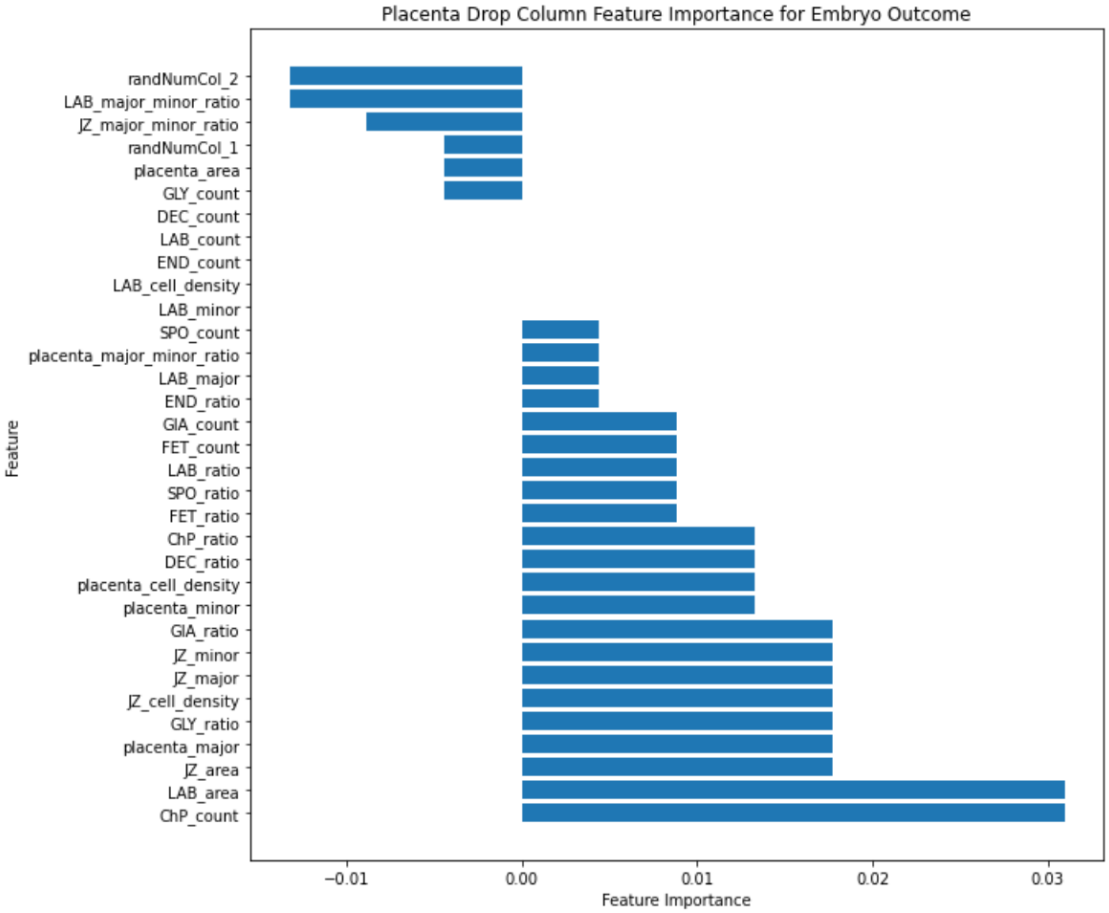



**Figure 6.17:** *This figure presents a bar chart of the drop column importance for each feature in the XGBoost model predicting three lethality outcomes: lethal, subviable, and viable. The y-axis represents the features, and the x-axis represents the drop column importance—a measure of how the model's predictive accuracy decreases when a specific feature is omitted. The higher the value, the more important the feature is for accurate predictions. Notably, the labyrinth area feature has the highest importance. The control features, represented by the two random number columns, confirm their lack of correlation with the outcomes by having negative importance.*

The second investigation characterised the wild-type placenta, revealing overlaps with existing literature and thereby endorsing the accuracy of the pipeline as an automated phenotyping approach. The third investigation applied the pipeline to examine sexual dimorphism in the E14.5 placenta, where it was observed that male placentas were larger than their female counterparts, aligning with existing literature. Interestingly, the study also revealed that male placentas at E14.5 have a higher proportion of glycogen cells compared to females. This finding, although not previously reported to the best of my knowledge, is consistent with the known

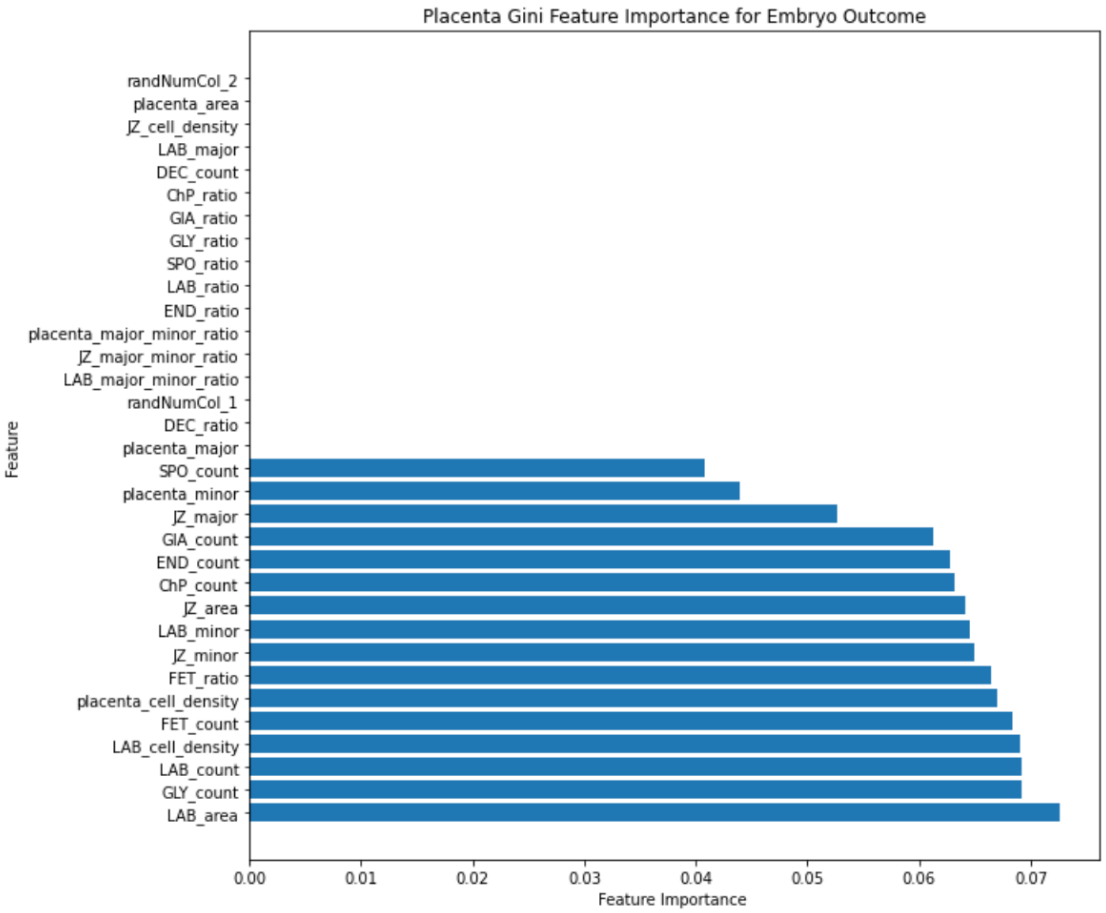



**Figure 6.18:** *This figure illustrates a bar chart of the Gini importance for each feature in the XGBoost model predicting three lethality outcomes: lethal, subviable, and viable. The y-axis represents the features, and the x-axis represents the Gini importance—a measure of the total reduction of the criterion brought by that feature. The higher the Gini importance, the more significant the feature is for the model. The labyrinth area feature emerges as the most critical predictor again. The negligible importance of the two random number columns validates the integrity of the feature importance analysis.*

larger average birth size of males.

Investigation 4 aimed at identifying statistically significant phenotypes in the 66 gene knockouts of the E14.5 dataset. Phenotypes were detected in 28 knockout cases, with overlaps noted with results reported in the Mouse Genome Informatics database. This investigation also uncovered new phenotypes, thereby expanding the existing knowledge base.

Investigation 5 utilised the phenotypes measured in the previous investigations to determine birth outcomes, achieving a predictive accuracy of 63% across three

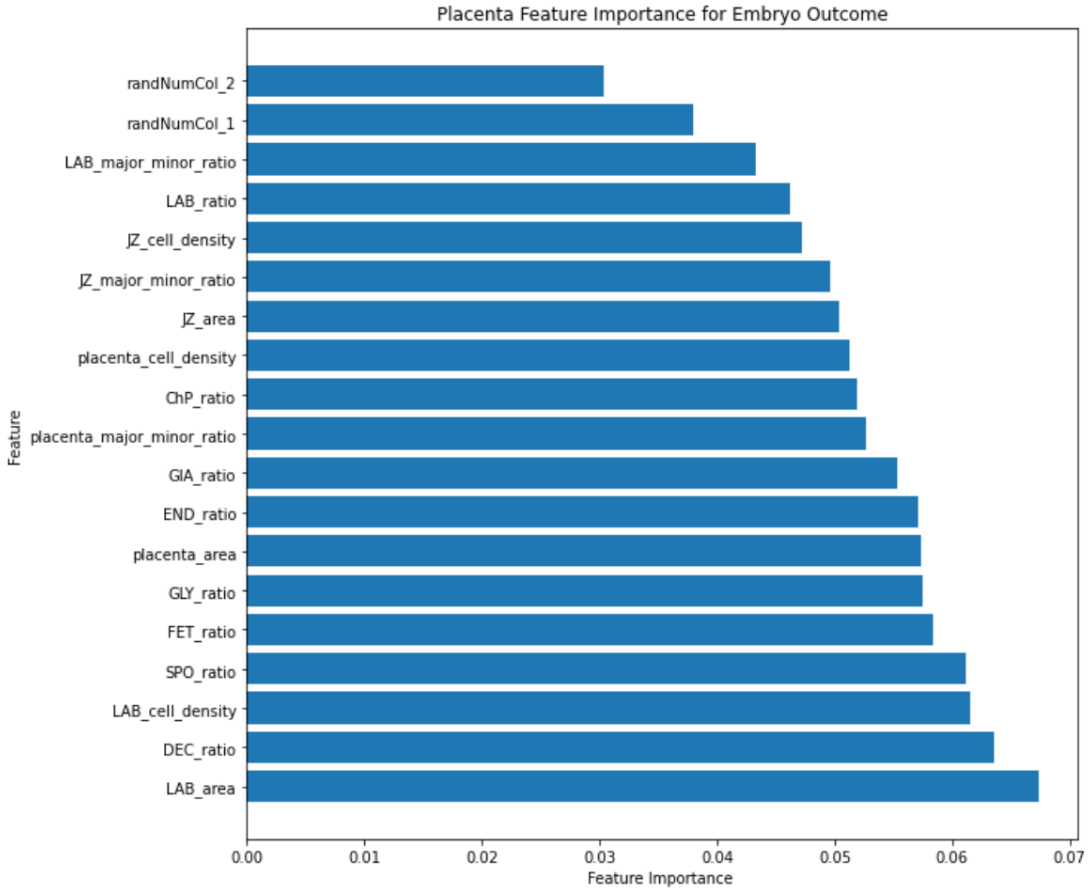



**Figure 6.19:** *This figure displays the feature analysis results after removing features with high colinearity to improve the model's interpretability. The y-axis represents the features, and the x-axis represents their relative importance after the removal of co-linear features. This analysis allows for a more accurate understanding of the impact of each feature on the model's predictions, free from the distortive effects of colinearity.*

outcomes. The labyrinth area was identified as the most critical feature contributing to negative outcomes, which aligns with expectations given its role in the maternal-foetal exchange.

This chapter has replicated existing literature and contributed to its advancement. It has conclusively demonstrated that the automated placental pipeline can effectively detect phenotype in mouse placentas. The pipeline's sensitivity allows for the detection of phenotypes that would otherwise be too time-consuming or impossible to measure accurately by humans. This research has thus unveiled new phenotypes, leading to a better understanding of mouse phenotypes and advancing knowledge of sexual dimorphism in development.

Further research may focus on refining the phenotyping pipeline and expanding its applications, thereby contributing to the continuous advancement of developmental biology.Additionally to further validate and enhance our computational models of placental development, integrating stereology and gene expression data could be invaluable. While computational models provide insights into potential phenotypic changes, stereology offers a robust quantitative method for understanding placental structure, and gene expression analysis provides the molecular context for these changes. Together, these methodologies could greatly improve the accuracy and relevance of our models.

Future research should focus on developing integrative approaches that combine these diverse data types. For instance, applying stereological techniques to quantify placental structures within our dataset could serve as a ground truth for validating our computational predictions. Similarly, incorporating gene expression profiles could help correlate structural changes with molecular pathways, providing a deeper understanding of the underlying biological processes.

Such integrative efforts would require the development of advanced data processing tools and machine learning algorithms capable of handling and interpreting this complex multi-dimensional data. By pursuing these enhancements, we aim to create more robust models that not only predict but also illuminate the mechanisms behind placental and fetal development, ultimately leading to better prenatal management and outcomes.

*"AI is probably the most important thing humanity has ever worked on. I think of it as something more profound than electricity or fire. And we have a lot to come. It will revolutionize many fields in ways we can't imagine."*

— Fei-Fei Li

# 7

# Development of Unsupervised Approach for Detecting Variation in Cell Morphology

## Contents

---

# 7.1 Introduction

## 7.1.1 Exploring Intraclass Variance in Placental Cells

Cellular heterogeneity, also known as intra-class variance within cell populations, is a key characteristic of multicellular organisms [134]. It is observed that cells of the same class or tissue type often exhibit unique gene expression profiles, morphologies, and functional capabilities. This diversity is often the result of genetic, epigenetic, and environmental factors shaping the dynamic cellular landscape within tissues and organs [135].

While cellular heterogeneity plays a critical role in maintaining homeostasis and facilitating tissue adaptation, it can also potentially initiate and drive the progression of diverse diseases. Several mechanisms contribute to the emergence of intra-class variance observed in cells, including Genetic Variability, Epigenetic Modifications, Stochasticity, and Microenvironmental Factors. Genetic variations,

such as single nucleotide polymorphisms (SNPs), copy number variations (CNVs), and chromosomal aberrations, can induce differences in gene expression and protein function, thus promoting cellular heterogeneity. Epigenetic alterations, including DNA methylation, histone modifications, and non-coding RNA molecules, have the potential to modulate gene expression and thus contribute to cellular heterogeneity. These epigenetic modifications can be heritable or induced by environmental factors such as ageing, stress, or exposure to toxins. Stochasticity, or "biological noise," in gene expression and cellular processes engenders phenotypic diversity within cell populations, promoting intra-class variance [136]. Furthermore, cells are highly influenced by their surrounding microenvironment, which includes components of the extracellular matrix, neighbouring cells, and signalling molecules.

The intra-class variance in cells contributes substantially to the onset and progression of various diseases. For instance, intra-class variance in neuronal and glial cells may be implicated in neurodegenerative diseases such as Alzheimer's and Parkinson's. Heterogeneity within populations of neurons and glial cells can influence neurotransmission, synaptic plasticity, and neural circuit function, thereby exacerbating disease progression. Moreover, intra-class variance in immune cells can lead to an imbalanced immune response, promoting the onset of inflammatory and autoimmune diseases.

The principle of inherent heterogeneity extends to placental cells, most notably in the form of trophoblast cells [137]. Trophoblasts, the principal cell type in the placenta, differentiate into various human subpopulations, including cytotrophoblasts, syncytiotrophoblasts, and extravillous trophoblasts. These subpopulations exhibit unique gene expression profiles and functional characteristics integral to placental development and function. In addition, the factors delineated above incite intra-class variance in placental cells [138].

Heterogeneity in placental cells has been linked to a host of pregnancy-associated disorders such as Preeclampsia and Intrauterine Growth Restriction (IUGR), which have been correlated with placental cell heterogeneity, notably in endothelial and

trophoblast populations [138]. Such heterogeneity can lead to inflammation and aberrant signalling, contributing to preterm labour and delivery.

Consequently, the study of placental cell heterogeneity holds significant implications for understanding, diagnosing, and treating pregnancy-associated disorders. Investigating the underlying mechanisms of placental cell heterogeneity can facilitate the identification of novel biomarkers for the early detection of pregnancy complications and pave the way for the development of targeted therapies to modulate placental function [139].

### 7.1.2 Scrutinising Intra-class Variance in Cell Morphology

Investigating intra-class variance in cell morphology poses a multifaceted challenge, primarily attributed to the dynamic nature of cell morphology that evolves concomitantly with shifting cellular functions. This dynamic characteristic is conspicuously evident in the placenta, necessitating morphology assessments at diverse stages, thereby invoking technical limitations [140]. These limitations emerge due to the intricacies in interpreting variations and the voluminous number of cells requiring examination to derive accurate insights. Moreover, the frequency of cell morphology variation can differ, potentially leading human interpreters to overlook crucial aspects. Consequently, an automated approach is imperative to meticulously examine cell morphology across the entire placenta and identify trends in intra-class variance.

Historically, the analysis of cell morphology has been a pivotal facet in cell biology. Advancements in technology have enabled researchers to increasingly leverage automated computational methods such as CellProfiler to scrutinize cell images, measure morphological attributes, and categorize cells based on their morphology [141]. These methods have been instrumental in elucidating the roles of various genes, proteins, and signalling pathways in shaping cell structure and function and have been invaluable in diagnosing and prognosing diseases such as cancer.

The preliminary requirement for analysing cell morphology is cell identification, typically achieved through segmentation, involving threshold-based methods, watershed algorithms, and deep learning techniques, as delineated in previous chapters. Following segmentation, pivotal morphological features like cell size, shape, and texture are extracted and quantified.

Upon extracting morphological features, cells can be profiled and classified based on their phenotypes. Techniques such as principal component analysis (PCA), hierarchical clustering, and t-distributed stochastic neighbour embedding (t-SNE) have been employed to visualise and classify cells in high-dimensional feature spaces [142].

Significant traction has been gained in understanding cell morphology and its underlying mechanisms by machine learning methods, specifically convolutional neural networks (CNNs) [143]. CNNs have been exploited in multiple tasks including cell segmentation, classification, and feature extraction. The reduced reliance on handcrafted features for distinguishing intra-class morphological variance has contributed to the popularity of these methods.

Additionally, autoencoders, unsupervised deep learning algorithms, have been utilized to learn low-dimensional representations of cell morphological features. For instance, variational autoencoders (VAEs) were employed by [144] to learn a low-dimensional representation of cell morphology, enabling segmentation and classification of cells in microscopy images. This approach has aided researchers in identifying subtle morphological variations potentially indicative of cellular function or response to perturbations.

Generative Adversarial Networks (GANs) have been instrumental in generating synthetic cell images, allowing the study of cell morphology under varied conditions [145]. These networks facilitate investigations into how genetic alterations can modulate cell morphology.

In the realm of cell morphology and intra-class variance studies, unsupervised and self-supervised learning methods have been highly advantageous. These methods, not reliant on labelled data, unveil several benefits, including the ability to learn

from extensive unlabeled datasets and identify unknown patterns or subtypes. Unsupervised learning methods facilitate the categorization of cell types and subtypes using various clustering algorithms, such as k-means, hierarchical clustering, and DBSCAN, without prior knowledge of class labels [146]. However, these methods have limitations due to their sensitivity to the quality and relevance of the extracted features impacting the clustering results.

Dimensionality reduction techniques like PCA and t-SNE have been fundamental in visualizing and analyzing high-dimensional morphological feature spaces [142]. They assist in uncovering relationships between cells and discovering unknown subtypes but may only sometimes accurately capture the underlying data structure, and the choice of hyperparameters can significantly affect the results.

Self-supervised learning techniques enable the learning of useful representations of cell images without labeled data, which has been pivotal for studying intra-class variance. While promising, these methods may necessitate large data quantities and can be computationally intensive, presenting challenges in certain scenarios.

Both unsupervised and self-supervised learning methods hold immense potential in advancing the study of cell morphology and intra-class variance. While these methods present numerous advantages, they also come with intrinsic limitations and challenges. The objective of this chapter is to critically assess these methodologies and their feasibility for studying intra-class variance in placental cell morphology to enhance our comprehension of these crucial processes.

## 7.2 Aims

Construct an unsupervised learning framework to explore the intra-class variability among DMDD placental cells and evaluate its applicability to real-world datasets.

### 7.2.1 Objectives

- Investigate using supervised learning models as feature extractors in detecting intra-class variance.

- Investigate the use of unsupervised clustering methods to distinguish between different cell types.
- Explore using advanced self-supervised learning models for feature extraction, dimensionality reduction, and clustering.

## 7.3 Methods

### 7.3.1 Dimensionality Reduction

The importance of dimensionality reduction techniques is foundational in data analysis, machine learning, and data visualisation, mainly when dealing with high-dimensional data [142]. This segment delves into an in-depth exploration of dimensionality reduction strategies, specifically focusing on Principal Component Analysis (PCA), t-Distributed Stochastic Neighbor Embedding (t-SNE), and Uniform Manifold Approximation and Projection (UMAP).

Dimensionality reduction fundamentally targets reducing the number of features or variables in a dataset while retaining the inherent structure and relationships among data points. It is a crucial step in dealing with high-dimensional datasets, often associated with heightened computational complexity and the "curse of dimensionality," which could potentially degrade the efficiency of machine learning algorithms. These challenges can be tackled by mapping data onto a lower-dimensional space, and computational efficiency can be enhanced. Additionally, dimensionality reduction can uncover common patterns in the extracted features, facilitating the identification and separation of classes or subclasses within cells when projected onto two or three-dimensional spaces.

Once features of cells are extracted, dimensionality reduction techniques can be applied to the high-dimensional feature space, thereby identifying cell subclasses. These techniques unravel patterns, clusters, or groupings in the data, revealing distinct cell subpopulations, Such techniques have applications in numerous biological contexts, leading to a richer understanding of cell functions, development, and pathology. This chapter uses three core methods: PCA, t-SNE and UMAP.

### Principal Component Analysis (PCA)

Introduced by Pearson in 1901 and later developed by Hotelling in 1933, PCA is a linear dimensionality reduction method that projects data onto a lower-dimensional subspace by extracting the principal components. These components are orthogonal directions that capture maximum variance in the data [147].

The implementation of PCA usually involves eigenvalue decomposition of the data covariance matrix or singular value decomposition (SVD) of the data matrix. Despite its simplicity and efficiency in capturing the linear structure of the data, PCA exhibits limitations when dealing with nonlinear data, potentially leading to inadequate preservation of local structures in the reduced space [142].

### t-Distributed Stochastic Neighbor Embedding (t-SNE)

Developed by Van Der Maaten and Hinton in 2008, t-SNE is a nonlinear dimensionality reduction method designed to preserve the local structure of high-dimensional data by modelling pairwise similarities between data points in both the original and reduced spaces. In the high-dimensional space, these similarities are modelled using Gaussian probability distributions. In contrast, a heavy-tailed Student's t-distribution in the lower-dimensional space facilitates the separation of dissimilar data points.

The algorithm minimises the Kullback-Leibler (KL) divergence between these two distributions via gradient descent. While t-SNE excels at preserving local structures and has been extensively utilised for visualisation, it has limitations. It is susceptible to hyperparameters, struggles to capture global structures, and encounters scalability issues with large datasets [142].

### Uniform Manifold Approximation and Projection (UMAP)

UMAP is a recent, nonlinear dimensionality reduction technique proposed by McInnes et al. in 2018. Building on the principles of manifold learning and topological data analysis, UMAP strives to preserve both local and global structures in the data by approximating the underlying manifold using a k-nearest neighbour

graph and subsequently optimising a fuzzy simplicial set representation of the graph in the reduced space [148].

UMAP defines the objective function as the cross-entropy between the fuzzy simplicial sets, which is minimised using stochastic gradient descent. By maintaining scalability, reducing sensitivity to hyperparameters, and preserving both local and global structures in the data, UMAP overcomes some of the limitations associated with t-SNE.

### 7.3.2 Comparison and Applications

PCA, t-SNE, and UMAP each bring distinct advantages and drawbacks in dimensionality reduction. PCA is ideal for linear datasets and is computationally efficient but may falter in capturing intricate nonlinear structures. Conversely, t-SNE is excellent at preserving local structures and visualising clusters but needs help with scalability and hyperparameter sensitivity. UMAP offers a balanced approach, preserving local and global structures while demonstrating improved scalability and reduced hyperparameter sensitivity.

These methods have been broadly applied across various fields, including biology (e.g., single-cell RNA sequencing data analysis), computer vision (e.g., feature extraction for image recognition), and natural language processing (e.g., embedding visualisation). The choice of dimensionality reduction technique should be guided by the specific problem at hand, the nature of the dataset, and the desired trade-offs between computational efficiency, scalability, and structure preservation.

### 7.3.3 Clustering Techniques

Clustering techniques are fundamental unsupervised learning approaches. These methodologies unearth underlying patterns and groupings within expansive datasets such as the embeddings of placental cells. The ultimate objective of clustering algorithms is to segregate datasets into distinct clusters or groups [149]. This segregation is achieved by interpreting inherent structures, concurrently minimising intra-cluster distances, and maximising inter-cluster distances. This section

delves into a deeper exploration of four prevalent clustering algorithms: K-means, HDBSCAN, Spectral Clustering, and Gaussian Mixture Models (GMM).

### 7.3.4 K-means Clustering

K-means is a centroid-based clustering algorithm known for its efficiency in partitioning datasets into a specified number of clusters (k), as referenced by [150]. The process begins with the selection of initial centroids, typically using the K-means++ initialization method to reduce sensitivity to initial centroid placement. The algorithm then iteratively adjusts these centroids to minimize the sum of squared distances from each data point to its nearest centroid, concluding when changes in centroid positions become negligible or a set maximum number of iterations is reached.

While k-means is celebrated for its computational efficiency and straightforward implementation, it is fundamentally predicated on the assumption that clusters are spherical and of similar densities when clustering is based on spatial features like coordinates. This assumption facilitates the algorithm's objective to form clusters with minimal variance within them, as each data point is assigned to the nearest centroid. However, it is crucial to note that if k-means is applied using non-spatial features such as size, shape, or diameter of objects, the 'spherical' nature of clusters refers to the similarity in these feature spaces rather than their geometric arrangement in physical space. Therefore, the algorithm's capability to capture the essence of clusters might vary significantly depending on the nature of the features used.

This characteristic implies that while k-means can efficiently partition large datasets when features exhibit spherical distribution patterns in the chosen feature space, its effectiveness may diminish if the underlying distribution deviates from this assumption. The algorithm's requirement for a predefined number of clusters also imposes a limitation, necessitating prior knowledge or additional methods to estimate the optimal number of clusters, such as the silhouette score or the elbow method discussed later.

Given k-means' reliance on the Euclidean distance metric, the choice of features is paramount. When employing k-means, it is essential to consider the scale and type of data. Features must be normalized or standardized if they are on different scales to prevent distortion in distance calculations, which could bias the clustering results. Furthermore, while k-means is generally robust to large datasets with well-separated, spherical clusters in feature space, its performance may be compromised when dealing with complex morphological data where clusters are non-spherical or highly irregular. This makes k-means less suited for datasets where clusters are defined by complex structures or interrelationships between features that do not conform to spherical distributions in the chosen feature space.

### 7.3.5 HDBSCAN

Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) represents a refined version of the DBSCAN clustering algorithm [151]. HDBSCAN functions by transforming the distance matrix of a dataset into a hierarchical structure. The hierarchical methodology of HDBSCAN equips it with the capacity to handle datasets proficiently, showcasing clusters of varying densities and shapes.

HDBSCAN's advantages over K-means include its ability to detect clusters of diverse shapes and sizes, robust noise point differentiation, and automatic determination of the Eps parameter. However, it may be computationally intensive for larger datasets.

### 7.3.6 Spectral Clustering

Spectral clustering employs concepts from spectral graph theory to group similar data points into clusters [152]. Unlike K-means, it can identify non-convex clusters. Spectral clustering creates an affinity matrix that signifies the similarity between data points and then applies eigendecomposition to this matrix to obtain a lower-dimensional representation. Data points are clustered in this reduced space using simpler algorithms such as K-means. Despite its flexibility in handling complex

structures, spectral clustering may be computationally expensive for large datasets and sensitive to the choice of the similarity matrix and the number of clusters.

### 7.3.7 Gaussian Mixture Models (GMM)

Gaussian Mixture Models (GMM) constitute another clustering technique, assuming that the data points are generated from a mixture of a finite number of Gaussian distributions with unknown parameters. GMM uses the Expectation-Maximisation (EM) algorithm to estimate the parameters of these distributions. It offers more flexibility than K-means as it considers the covariance structure of the data, accommodating elliptical and spherical clusters. Also, GMM allows soft clustering, where each data point has a probability of belonging to each clusters. Despite its strengths, GMM can converge to local optima and might struggle with high-dimensional data due to the 'curse of dimensionality'.

### 7.3.8 Critical Comparison of Clustering Techniques in Cell Morphological Variance Identification

The problem of identifying cell morphological variance using embeddings from images of placental cells invites an interesting comparison between the discussed clustering methods: K-means, HDBSCAN, Spectral Clustering, and Gaussian Mixture Models (GMM).

With its ease of implementation and computational efficiency, K-means could be beneficial for initial data exploration and gross morphological delineation. However, its assumption of spherical clusters and equal densities might limit its ability to effectively capture complex morphological variance in placental cells, where shapes might be irregular, and densities vary significantly.

HDBSCAN, on the other hand, has an edge over K-means in handling datasets with diverse cluster shapes and densities. Its noise differentiation ability is especially advantageous in cell morphology analysis, where noise from image processing or cell artefacts is common. Yet, its computational demands could hinder working with many high-dimensional cell embeddings.

Spectral clustering brings a unique advantage in identifying complex, non-convex morphologies in placental cells. It might offer a more nuanced interpretation of cell variance than K-means or HDBSCAN. However, similar to HDBSCAN, Spectral Clustering can be computationally burdensome for larger datasets. Also, its performance hinges significantly on the choice of the similarity matrix and the number of clusters, adding to the complexity of model selection.

Gaussian Mixture Models (GMM) offer a probabilistic perspective to the problem, potentially enabling a softer, more nuanced segregation of cell morphologies. GMM's consideration of the covariance structure of the data could cater to diverse morphological characteristics in placental cells. However, its susceptibility to converge to local optima and potential struggles with high-dimensional data might pose challenges in this application.

Each method has specific strengths and weaknesses that affect its performance in discerning morphological variation in placental cells. The best way to discover which ones perform best is to deploy them on the data; this will be done in the results section.

### 7.3.9 Determination of Optimal k in Clustering: The Silhouette Score and Elbow Method

Determining an optimal number of clusters, or k, is a pivotal aspect of effective clustering. Numerous methods are available to accomplish this, among which the silhouette score and the elbow method are often employed.

### 7.3.10 The Silhouette Score

The silhouette score has earned significant recognition as a metric for assessing the quality of clustering results [153]. It quantifies how comfortably each data point resides within its assigned cluster while concurrently evaluating its distinction from neighbouring clusters. The silhouette score's value ranges between -1 and 1. A higher silhouette score closer to 1 suggests well-separated and cohesive clusters. Conversely, a score approaching zero denotes that data points are precariously near

the decision boundary between clusters. A negative score may imply that data points have been erroneously allocated to clusters, hinting towards overlapping clusters.

The silhouette score can be instrumental in choosing the optimal number of clusters in K-means clustering. This is achieved by computing and comparing the silhouette scores for different values of k. The k corresponding to the highest silhouette score is generally selected as the optimal number of clusters.

### 7.3.11 The Elbow Method

Complementing the silhouette score is the Elbow method, another prevalent technique used to determine the optimal number of clusters in K-means. The Elbow method reduces variance, the sum of the squared distances between each cluster member and its centroid.

In the Elbow method, the total within-cluster sum of squares (WSS) is computed for various k values. As the number of clusters increases, the WSS tends to decrease, with each additional cluster leading to a smaller drop in the WSS. This phenomenon creates an 'elbow-like' curve when the WSS is plotted against the number of clusters.

The 'elbow' of this curve, or the point where adding another cluster does not significantly better the total WSS, offers a reasonable estimate for the optimal number of clusters [154]. While this method is intuitive and easy to implement, identifying the 'elbow' can sometimes be subjective, especially when the curve is relatively smooth or multiple 'elbows' appear.

### 7.3.12 Comparison of the Silhouette Score and Elbow Method in Determining Optimal k

The Silhouette Score and the Elbow Method are well-established techniques for determining a dataset's optimal number of clusters. However, their applications and efficiencies vary depending on the nature of the dataset and the specific clustering problem at hand.

The Silhouette Score measures both cohesion within clusters and separation between clusters. This dual nature makes it more informative, providing insights

into the clustering quality beyond merely evaluating within-cluster variance. This measure is particularly advantageous in cases where inter-cluster separation is as significant as intra-cluster cohesion, such as in the analysis of cell morphological variance. However, this comes at the cost of increased computational complexity, making the silhouette score less efficient for larger datasets. Additionally, the silhouette score might present less reliable results in overlapping or densely packed clusters.

In contrast, the Elbow Method primarily focuses on reducing variance within clusters. It is more computationally efficient and more accessible to implement than the silhouette score, making it an attractive choice for exploratory data analysis or handling large datasets. The simplicity of the Elbow Method extends to its visual interpretability as well, with the 'elbow' in the plot often clearly demarcating the point of diminishing returns for increasing the number of clusters. However, its major limitation lies in its subjectivity; identifying the 'elbow' point may only sometimes be clear-cut, especially when the plot is smooth or exhibits multiple potential 'elbows'. Furthermore, considering only the within-cluster variance may overlook significant aspects of the broader clustering structure, such as inter-cluster separation.

While the Silhouette Score offers a more nuanced and comprehensive measure of clustering quality, it may be computationally intensive for large datasets and less reliable in cases of dense or overlapping clusters. On the other hand, the Elbow Method is simple, efficient, and visually intuitive but can be somewhat subjective and overlook important aspects of cluster separation. Thus, the choice between the two methods will largely depend on the specific requirements of the task, including considerations of computational resources, dataset size, and the balance between intra- and inter-cluster relationships.

### 7.3.13 Performance Metrics for Assessing the Quality of Clustering Methods

The efficacy of a clustering solution can be gauged using different metrics that capture various aspects of the clustering task, such as the similarity of objects within the same cluster and the dissimilarity of objects across different clusters. This section

will discuss six key metrics for assessing the performance of clustering methods: Adjusted Rand Index (ARI), Adjusted Mutual Information (AMI), Homogeneity, Completeness, V-measure, Calinski-Harabasz Index, and Davies-Bouldin Index.

### 7.3.14 Adjusted Rand Index (ARI)

The Adjusted Rand Index (ARI) is a measure that quantifies the similarity between two data clusterings, accounting for chance groupings. ARI modifies the Rand Index, a straightforward measure of the percentage of decisions (i.e., whether a pair of objects should or should not be in the same cluster) that are correct, by introducing an adjustment for chance. It ranges from -1 to 1, with 1 denoting perfect agreement between two clusterings, 0 indicating the result equivalent to random assignments, and negative values suggesting agreement less than what is expected by chance.

### 7.3.15 Adjusted Mutual Information (AMI)

Similar to ARI, Adjusted Mutual Information (AMI) measures the similarity between two clusterings but also corrects the agreement's effect solely due to chance [155]. AMI is based on information theory; it compares the information obtained about one clustering to the other, correcting the score for chance agreement. Like ARI, AMI ranges between -1 and 1, with a similar interpretation.

### 7.3.16 Homogeneity and Completeness

Homogeneity and completeness are two metrics that evaluate whether a clustering result satisfies two fundamental properties: each cluster contains only members of a single class (homogeneity), and all members of a given class are assigned to the same cluster (completeness). Both metrics range from 0 to 1, with 1 representing perfect homogeneity or completeness.

### 7.3.17 V-Measure

V-Measure is a combination of homogeneity and completeness, the harmonic mean of these two metrics [156]. It encapsulates both aspects of good clustering, taking

a balanced view of precision and recall. A V-Measure of 1 indicates perfect homogeneity and completeness.

### 7.3.18 Calinski-Harabasz Index

The Calinski-Harabasz Index (CH Index) quantifies the ratio between cluster dispersion and within-cluster dispersion, thereby measuring cluster separation [157]. A higher value of the CH Index implies better clustering, i.e., higher separation between clusters and tighter clusters.

### 7.3.19 Davies-Bouldin Index

The Davies-Bouldin Index (DB Index) evaluates the average 'similarity' between each cluster and its most similar one, where the similarity is a measure that combines the distances between cluster centres and the size of the clusters [158]. Unlike the CH Index, lower values of the DB Index indicate superior clustering, with 0 being the ideal score. All these metrics offer valuable insights into the performance of clustering solutions from different perspectives, allowing for a comprehensive and robust assessment of the clustering task. The selection of the metric(s) to be used in a given context depends on the specific requirements of the clustering task and the properties of the data at hand.

### 7.3.20 Comparison of Clustering Performance Metrics

The selection of appropriate performance metrics is a critical aspect of clustering analysis. The metrics examined previously - Adjusted Rand Index (ARI), Adjusted Mutual Information (AMI), Homogeneity, Completeness, V-Measure, Calinski-Harabasz Index, and Davies-Bouldin Index - each offers different perspectives on the clustering task. While all are valuable tools for assessing clustering solutions, their applicability and efficacy vary depending on the nature of the dataset and the specific clustering problem.

The Adjusted Rand Index (ARI) and Adjusted Mutual Information (AMI) quantify the agreement between two data clusterings while accounting for chance.

These metrics objectively evaluate clustering solutions, which are particularly useful when ground truth labels are available. However, their reliance on true labels may limit their applicability in real-world scenarios where such labels are often unavailable. Furthermore, these metrics need to consider the geometric structure of the data, which might lead to lower scores for geometrically meaningful clusters but not reflected in the labelling.

Homogeneity, completeness, and the V-measure provide complementary views of the clustering quality, focusing on whether each cluster contains only members of a single class (homogeneity) and whether all members of a given class are assigned to the same cluster (completeness). The V-measure harmonises these two perspectives, offering a balanced assessment of clustering performance. However, these measures, like ARI and AMI, depend on the availability of ground truth labels, limiting their utility in unsupervised scenarios.

The Calinski-Harabasz Index (CH Index) and the Davies-Bouldin Index (DB Index) are two indices that do not require ground truth labels and consider the geometric structure of the data. The CH Index gauges the ratio of between-cluster dispersion to within-cluster dispersion, providing a measure of overall separation between clusters. Conversely, the DB Index assesses the similarity between each cluster and its most similar one, combining the distances between cluster centres and the size of the clusters. While these indices provide valuable insights into the clustering structure, their interpretation may sometimes be complicated, especially in high-dimensional spaces. Furthermore, the CH and DB indices may perform poorly with non-convex clusters or when clusters have different densities.

These clustering performance metrics have distinct strengths and weaknesses. The choice among them hinges on several factors, including the availability of true labels, the importance of capturing geometric structure, the computational resources at disposal, and the specific requirements of the clustering task. While metrics like ARI, AMI, homogeneity, completeness, and the V-measure offer a robust evaluation framework when true labels are available, the CH Index and DB Index are more suited for unsupervised scenarios and capture the geometric

nuances of the data. These methods will all be deployed to access clustering performance in the results section.

### 7.3.21 Self-Supervised Learning in Cell Morphology Differentiation

In the realm of machine learning, Self-Supervised Learning (SSL) is fast gaining recognition as an advantageous approach that liberates the necessity for extensive labelled datasets—assets that are costly, labour-intensive, and time-consuming to amass [68]. In stark contrast to supervised learning, which utilises labelled input-output pairs for training models, SSL thrives on learning rich representations from the data's inherent structure without the requirement of explicit manual labels [159]. This inherent capacity to create and exploit meaningful tasks based on data structure forms the bedrock of SSL. Through these self-generated tasks, models acquire general-purpose features adaptable to various downstream tasks.

The landscape of SSL has witnessed rapid advancements recently, with the advent of contrastive learning methods playing a pivotal role in this progression. Contrastive learning, a subcategory of SSL, operates on the principle of accentuating the agreement between varying perspectives of identical data and concurrently reducing agreement with unrelated data [160]. In computer vision, the frameworks SimCLR [161] and MoCo [162] have revolutionised SSL, substantially improving its performance.

SimCLR is a self-supervised learning framework that deploys contrastive learning to derive valuable visual representations without labelled data. It accomplishes this through data augmentation, base encoding, a projection head, and a contrastive loss function. The model learns to map various augmentations of an image to a shared space, ensuring similar images score highly for similarity and dissimilar ones score low.

Conversely, MoCo, another SSL framework, differentiates itself from SimCLR by implementing a memory bank and a momentum update for the encoder, factors which enhance the stability and efficiency of the learning process.

When differentiating the morphology of cells in histology images, SSL methods can be employed to exploit their ability to extract significant features sans labelled data. A dataset of histology images can be subjected to various data augmentations to generate multiple views for each image, forming the basis for training a contrastive learning model, like SimCLR or MoCo.

Once the SSL model is trained, the extracted features can be used for differentiating cell morphologies. Clustering algorithms such as K-means or hierarchical clustering could be employed to this end. An alternative approach could involve using a small quantity of labelled data to fine-tune the model in a supervised manner, thereby refining its performance in differentiating cell morphology.

The application of SSL in cell morphology differentiation has its challenges. Histology images are inherently complex and often demonstrate significant intraclass variance, making learning discriminative features a challenging task. Overcoming these hurdles may necessitate devising new pretext tasks or data augmentation strategies tailored explicitly for histology data. Additionally, the incorporation of domain-specific knowledge and integration of SSL with other learning paradigms, such as few-shot learning or transfer learning, might serve to enhance the model's performance in distinguishing cell morphologies.

Self-supervised learning offers a novel approach to learning representations without needing explicit labels. With its advancements in computer vision, especially contrastive learning methods, SSL is showing great promise for applications such as cell morphology differentiation in histology images. By employing SSL techniques, we can lessen our dependency on labelled data, thereby enabling the development of more efficient and scalable solutions in various domains, including classifying intraclass variance based on cell morphology. Future research is warranted to tackle the challenges associated with SSL's application to histology images and to investigate the integration of domain-specific knowledge and other learning paradigms to amplify its applicability and performance.

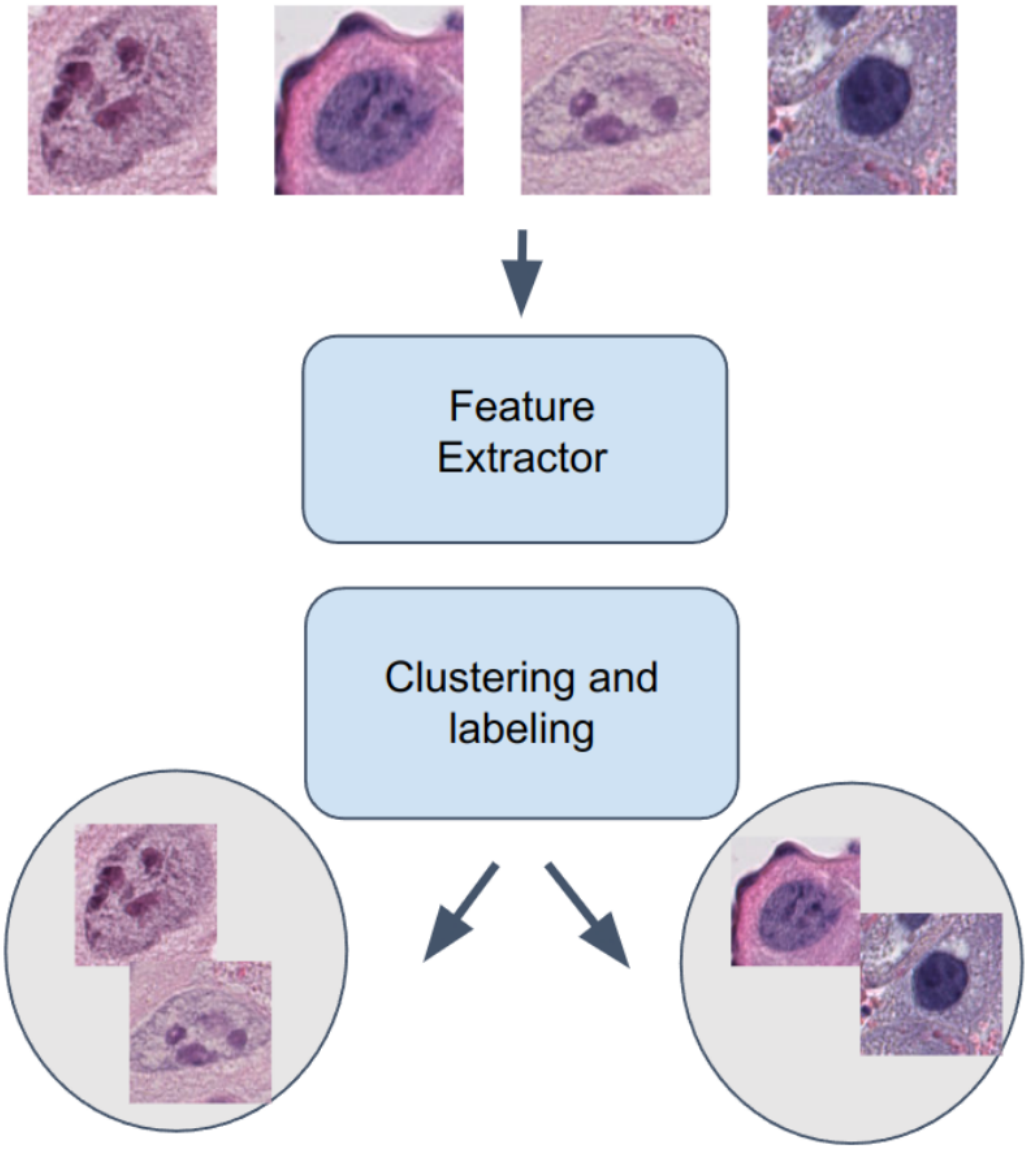



**Figure 7.1:** *This figure illustrates the process flow for the analysis of Trophoblast Giant Cells in mouse placenta. The process begins with the feature extractor, which extracts characteristic attributes from the cells. The extracted features are then passed to a clustering and labelling method, which classifies the cells into distinct groups based on similarities in their features. This methodology aims to discern intraclass variance within the cell classes of the mouse placenta, offering insights into the heterogeneity of Trophoblast Giant Cells and other cell classes.*

## 7.4 Results and Discussion

## 7.5 Investigation 1: Assessment of Dimensionality Reduction Techniques for High-Dimensional Data in the Context of Placental Cell Classification

**Aim:** This investigation critically evaluated three dimensionality reduction techniques, namely Principal Component Analysis (PCA), t-Distributed Stochastic Neighbor Embedding (t-SNE), and Uniform Manifold Approximation and Projection (UMAP), with the aim of discerning interclass and intraclass variance in high-dimensional data.

**Method:** The analysis employed high-dimensional embeddings of placental cells, extracted using the InceptionResNetV2 model previously described in Chapter 5. This model, initially trained for cell classification, provided a rich set of features encompassing various cell characteristics such as texture, morphology, and intensity patterns, encapsulated in a 384-dimensional space.

**Result:** PCA showed limited performance due to the non-linear relationships in the dataset. t-SNE displayed better performance with clear delineation of the clusters, while UMAP yielded the most desirable results with well-separated and compact clusters.

**Conclusion:** UMAP emerged as a suitable technique for handling high-dimensional data in this context, outperforming both PCA and t-SNE in capturing both interclass and intraclass variance. However, the choice of dimensionality reduction technique should be guided by the nature of the dataset and the specific objectives of the study.

**Future Work:** Future research could focus on exploring other dimensionality reduction techniques and comparing their performance on similar datasets. Additionally, research could focus on improving the performance of PCA and t-SNE on high-dimensional and nonlinear datasets.

This analysis delves into the performance of a feature extractor trained for supervised tasks when applied to unsupervised clustering methods. Specifically, I critically evaluate three dimensionality reduction techniques—Principal Component Analysis (PCA), t-distributed Stochastic Neighbor Embedding (t-SNE), and Uniform Manifold Approximation and Projection (UMAP)—with a focus on discerning interclass and intraclass variances in high-dimensional data. Here, 'interclass

variance' signifies observable differences between distinct classes, while 'intraclass variance' encapsulates variations within the same class.

Dimensionality reduction techniques streamline complex data into a more manageable form while striving to preserve intrinsic information. The assessment of these techniques within this study is based on the clarity and separation of visual clusters when the high-dimensional data is represented in two dimensions.

The process involved in each dimensionality reduction technique utilized the high-dimensional features generated by the InceptionResNetV2 model. This model had been optimized in Chapter 5 to classify cells based on their visual characteristics, producing embeddings that reflect deep learned patterns crucial for understanding cell types in placental tissues. The efficacy of PCA, t-SNE, and UMAP was then evaluated based on how well they could reduce these embeddings into two or three dimensions while preserving the most significant features for clustering and visualization purposes. The input to the feature extractor was 100x100pixel images centered arounf the nuclei as trained on in Chapter 5

The assessment of these techniques was mainly qualitative due to the inherent complexities of the dataset. However, by using the actual cell labels from the supervised training as a reference, we were able to verify that cells of the same class generally clustered together in the reduced-dimensional space, indicating that the dimensionality reduction techniques were effectively preserving relevant biological information.

For each dimensionality reduction technique, I observed distinct behaviours and limitations. PCA, traditionally adept at linear correlations, demonstrated a convergence of clusters towards the graph's central axis (Figure 7.2), suggesting a mismatch with the dataset's inherent nonlinearities and complex structure. t-SNE, although offering a better cluster delineation (Figure 7.2), suffered from some cluster dispersion. Despite its ability to handle non-linearity and high dimensionality, it comes with trade-offs such as computational intensity and hyperparameter sensitivity.

UMAP, on the other hand, not only provided clear cluster separation but also presented a more compact cluster formation (Figure 7.2). This compactness is indicative of UMAP's effectiveness in representing the high-density regions of the data, which translates to a more precise modelling of both interclass and intraclass variance. UMAP's delineation of clusters, therefore, is not just about separation but about the meaningfulness of the data representation it provides.

The expected behaviour of a supervised feature extractor in unsupervised clustering is to leverage learnt discriminative features to enhance clustering performance. However, this can lead to biased representations if the feature extractor is overly tailored to the supervised task's classes. This bias was observed in PCA, where the linear discriminative features were insufficient to resolve complex data structures. t-SNE and UMAP, being nonlinear, were more adept at utilising these features for clustering, with UMAP showing a particular strength in maintaining the global data structure, which is critical for unsupervised learning tasks.

In comparing UMAP and t-SNE, while both exhibited seven distinct clusters, the compactness and less dispersed nature of UMAP's clusters provided a clearer distinction between classes. This is particularly beneficial when clusters are used in subsequent analyses, where clear demarcations are crucial. Although this study does not present comparative computational times, literature amd my experience suggests that UMAP is generally more computationally efficient than t-SNE, which is a significant advantage in large-scale data analyses.

These results support the broader academic consensus that considers UMAP an excellent tool for analysing high-dimensional data, especially when it is imperative to maintain both global and local structures. However, it is imperative to stress that the effectiveness of any dimensionality reduction technique is contingent upon the unique characteristics of the dataset at hand. While this investigation underlines the strengths of UMAP, it is not a dismissal of other techniques but a reminder of the importance of selecting a method congruent with the data's intrinsic properties and the specific objectives of the research.

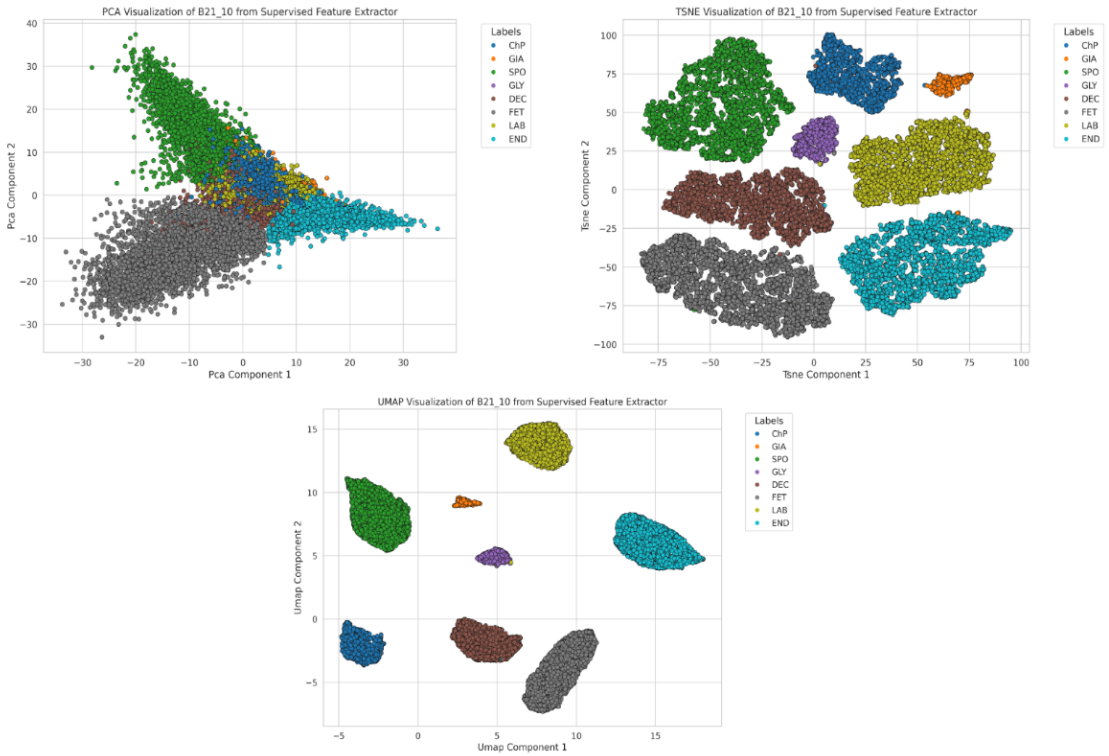



**Figure 7.2:** *This figure displays the results of three distinct dimensionality reduction techniques applied to E14.5 placental cells, each aimed at reducing the original 384-dimensional embedding space to a more interpretable two-dimensional space. Panel (a) shows PCA (Principal Component Analysis), which linearly projects the data into lower dimensions and is seen here with overlapping clusters, reflecting its challenges with complex, nonlinear data structures. Panel (b) illustrates t-SNE (t-Distributed Stochastic Neighbor Embedding), which excels in forming distinct, separated clusters by preserving local relationships, suitable for intricate high-dimensional data. Panel (c) depicts UMAP (Uniform Manifold Approximation and Projection), effectively maintaining both global and local data structures, resulting in well-defined, compact clusters. These visualizations highlight the critical role of choosing the right dimensionality reduction technique tailored to the specific needs and nuances of the dataset.*

## 7.6 Investigation 2: Comparative Evaluation of Unsupervised Clustering Algorithms in the Analysis of Placental Cell Data

**Aim:** This investigation aimed to evaluate four unsupervised clustering methodologies: k-means, spectral clustering, Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), and Gaussian Mixture Models (GMM), with a focus on discovering hidden patterns in the Deciphering Developmental Disorders (DMDD) dataset.

**Method:** The performance of the Elbow Method and Silhouette Score in

estimating the optimal number of clusters was assessed using the k-means clustering algorithm. These methods require an underlying clustering algorithm to function and were applied on top of k-means results to determine a range of potential cluster counts. The Elbow Method evaluates the percentage of variance explained by each cluster number, while the Silhouette Score assesses the average distance between clusters compared to within clusters. This step was crucial in determining the initial guesses for the number of clusters, which were then used to tune other clustering algorithms.

**Result:** The Elbow Method suggested ten clusters as optimal, whereas the Silhouette Score identified five as the primary choice, with eight as a secondary possibility. These results indicated that neither method perfectly aligned with the known ground truth of eight clusters, highlighting potential discrepancies in heuristic evaluations. Following this, all clustering algorithms were standardized and further evaluated with $k = 8$, allowing for a consistent comparison across different methods.

**Conclusion:** Spectral Clustering emerged as the most suitable methodology for this specific dataset due to its ability to identify complex structures and its high performance across all metrics. However, the choice of clustering method should be guided by the characteristics of each dataset and the specific requirements of the analysis.

**Future Work:** Future research could focus on optimizing the parameters for HDBSCAN and GMM, exploring other dimensionality reduction techniques, and investigating how these techniques could be used together for optimal results. Additionally, further evaluations could consider the unique capabilities of HDBSCAN in automatically determining the number of clusters, contrasting it with the preset requirements of other clustering methods.

Unsupervised clustering algorithms are designed to discover hidden groups or patterns within data, devoid of any a priori labels or knowledge. The present study evaluated four principal clustering methodologies: k-means, spectral clustering, Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), and Gaussian Mixture Models (GMM) 7.1.

Most of the evaluated methods, barring HDBSCAN, necessitated a predetermined 'k'-value to function. In the absence of such a value, heuristic techniques, namely the Elbow and the Silhouette scores, were utilised to deduce the optimal cluster count.

The application of these methods to the Deciphering Developmental Disorders (DMDD) dataset, comprising eight clusters, yielded divergent results. The Silhouette

score intimated the existence of five clusters, positing eight as the secondary likelihood. Conversely, the Elbow method indicated a preference for ten clusters. While neither heuristic perfectly mirrored the actual cluster count, they provided reasonable approximations, revealing an innate structure within the dataset.

The analysis found that the Elbow Method's performance could diminish in the context of high-dimensional spaces, such as those with embeddings of length 384. Here, the curse of dimensionality might render distances between data points—and consequently between clusters—less meaningful, potentially elucidating the method's recommendation of ten clusters, in contrast to the true count of eight.

Conversely, the Silhouette Score, which presupposes that clusters are convex and isotropically distributed, may not accurately reflect the complexity of datasets like the placental cell images. This proclivity, along with a tendency to endorse fewer clusters, might account for its suggestion of five as the most probable cluster number.

The Silhouette Score's secondary prediction of eight clusters and its inherent bias towards fewer clusters rendered it a comparatively more credible estimator in this scenario. Nonetheless, given their heuristic nature, these methods offer guidance rather than definitive resolutions. Optimal cluster determination may thus require additional data insights, contextual knowledge, or alternative validation metrics.

A synthesis of these clustering methodologies with domain-specific expertise is likely to facilitate a more accurate cluster count estimation. Subsequent investigations might probe alternative dimensionality reduction techniques and clustering algorithms less affected by the shape and scale of clusters.

To appraise the quality of these clustering methodologies, both visual assessments, via UMAP plots, and quantitative metrics, including the Adjusted Rand Index (ARI), Adjusted Mutual Information (AMI), Homogeneity, Completeness, and V-Measure, were employed. These metrics provided a comprehensive evaluation of the clustering results, reflecting the consistency and unity within clusters and classes. All clustering methods were ultimately evaluated with $k = 8$ , aligning with the 8 recognized cell types in the dataset, to ensure a consistent basis for comparison and validation.

**Table 7.1:** *Comparison of Clustering Methodologies Applied to the Deciphering Mouse Developmental Disorders (DMDD) Dataset.* This table presents a quantitative comparison of four distinct clustering methodologies: K-Means, Spectral Clustering, Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), and Gaussian Mixture Models (GMM). The methods are evaluated across five metrics: Adjusted Rand Index (ARI), Adjusted Mutual Information (AMI), Homogeneity, Completeness, and V-measure. These metrics assess the consistency within each cluster and the ability to group each class into unified clusters. The spectral clustering method outperforms the others across all metrics, indicating its effectiveness in discerning complex structures within the high-dimensional placental cell data.

| Clustering Method | ARI | AMI | Homogeneity | Completeness | V-measure |
|---|---|---|---|---|---|
| K-Means | 0.8 | 0.86 | 0.88 | 0.84 | 0.86 |
| Spectral | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| HDBSCAN | 0.3 | 0.52 | 0.46 | 0.6 | 0.52 |
| GMM | 0.74 | 0.85 | 0.95 | 0.78 | 0.85 |

Spectral clustering was ascertained as the most efficacious method for this specific dataset, outshining the other methodologies across all metrics. The robust scores in ARI, AMI, Homogeneity, Completeness, and V-Measure corroborated spectral clustering's capability in discerning the data's complex structures.

Despite its simplicity and computational efficiency, K-means clustering, with its presupposition of equal-sized and spherical clusters, could not match the performance of spectral clustering. Gaussian Mixture Models (GMM) exhibited notable results, particularly with a high Homogeneity score, suggesting that most clusters predominantly comprised samples from singular classes. However, the Completeness score indicated that certain classes were disseminated across various clusters, detracting from its overall efficacy.

HDBSCAN, despite its advanced density-based clustering algorithm, did not perform optimally in this context. Its sensitivity to varying cluster densities and sizes, coupled with the potential challenges posed by high-dimensional data or noise, may have limited its effectiveness. Unlike k-means, spectral clustering, and GMM, HDBSCAN does not require a predefined number of clusters, allowing it to adaptively determine the cluster count based on data density and connectivity. This characteristic makes HDBSCAN unique in our evaluation, as its performance and optimal cluster count were independently assessed without the initial heuristic

guidance from the Elbow or Silhouette methods. This approach highlights HDBSCAN's ability to identify clusters organically, potentially providing more tailored and accurate clustering results for complex datasets.

While each clustering method has its respective strengths and is applicable in specific contexts, Spectral Clustering proved to be the most appropriate for the placental cell data, courtesy of its capacity to identify complex structures and high metric performance. The choice of clustering methodology should, however, be informed by the dataset's specific characteristics and requirements. Future research endeavours could include the fine-tuning of parameters for HDBSCAN and GMM and investigating how an ensemble of these techniques could be optimally deployed.

## 7.7 Investigation 3: Quantitative Assessment of Intraclass Variance in Placental Cells From a Supervised Feature Extractor

**Aim:** This analysis aimed to examine the intraclass variance within placental cells using clustering techniques, specifically focusing on identifying hidden patterns within predefined cell classes.

**Method:** The investigation utilised a supervised model as a feature extractor, followed by determining the optimal number of clusters using the Silhouette and Elbow methods Figure 7.9. Subsequently, K-means clustering was applied to the features within each specific cell class.

**Result:** Initial visualisation of the intraclass variance using UMAP plots did not display distinct clusters for each cell type. The application of K-Means clustering also did not detect substantial intraclass variance within the data. However, a qualitative investigation by plotting clusters next to each other yielded some notable examples.

**Conclusion:** The techniques employed in this study were unable to detect substantial intraclass variance within the placental cells. This could be due to several factors, such as the high dimensionality of the data or potential overlap of cell classes.

**Future Work:** Future research could focus on exploring other clustering techniques, examining methods to better capture intraclass variance, or investigating the influence of different cell classes' overlap.

This section explores the intraclass variance within placental cells, building upon

the methodologies discussed earlier, which employ a supervised, trained model for feature extraction. The number of clusters was determined using the Silhouette and Elbow methods, with k-means clustering subsequently applied to the cellular data. The preference for k-means, despite spectral clustering's noted performance, stems from its simplicity, effectiveness, and widespread adoption in similar research scenarios. This step represents a shift from the prior method, as it focuses solely on intraclass variance within predefined cell classes.

Utilising the supervised model from Chapter 5—which categorised the cells into one of eight classes with an accuracy of 89%—clustering was confined to the features within each individual class.

Intraclass variance for each cell class was initially visualised using UMAP to plot the 384-dimensional embeddings. Figure 7.4 illustrates the UMAP plots for all eight cell classes in the dataset, with labels denoting the most probable number of clusters as determined by the highest Silhouette scores.

The UMAP plots did not reveal distinct clusters corresponding to each cell type, despite the known presence of intraclass variance, such as the diversity within the labyrinthine layer, which includes syncytiotrophoblasts types 1 and 2. This ambiguity raised questions about the role of high dimensionality and the potential onset of the curse of dimensionality with 384-dimensional embeddings. To probe this further, the size of the embeddings was varied for interclass varience, and the resulting UMAP visualisations were compared.

The embedding lengths tested were 128, 192, 256, 320, and 384 dimensions. Across all sizes, the UMAP plots continued to show consistent clustering within the classes, suggesting that the primary cause of the indistinct intraclass variance was unlikely to be the embedding length. Figure 7.3 provides a visual comparison of the UMAP plots for each embedding size, supporting the conclusion that factors other than the curse of dimensionality are contributing to the observed variance.

Additionally, UMAP hyperparameters could have been poorly tuned, although this is less likely given the performed tuning. Furthermore, it is conceivable that the model needs to extract more low-level features to distinguish intraclass

S

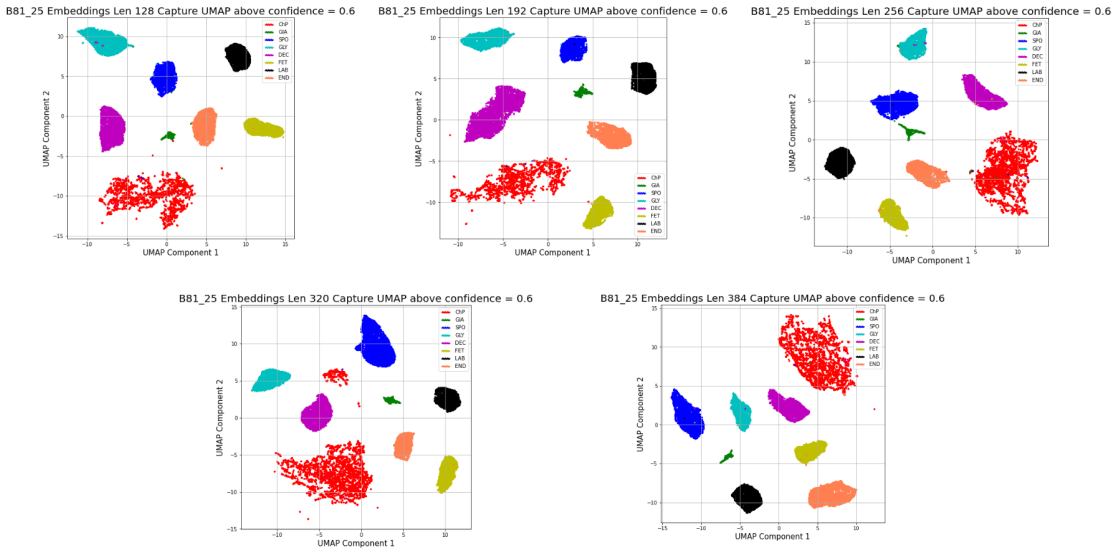



**Figure 7.3:** Comparison of UMAP plots for different embedding sizes for placental cell data. The sizes tested were 128, 192, 256, 320, and 384 dimensions. The plots suggest that the indistinct intraclass variance is not primarily due to the length of the embeddings, as the lack of distinct clustering persists across all sizes tested.

variance. This could be due to class overlap, as with the labyrinth cells grouping syncytiotrophoblasts types 1 and 2, or cell classes with overlapping features. If the variation within each class is less than the variation between classes, overlapping clusters could ensue, obscuring the features and thereby inhibiting the detection of intraclass variance.

To examine whether clustering could illuminate the intraclass variance within cells where UMAP had fallen short, K-means clustering was applied. The number of clusters required for K-means clustering was determined using the Silhouette score and Elbow method, as delineated in the methods section. The analysis of each cell class is presented below with reference to 7.2:

- **Chorionic Plate (ChP) Cells.** For 2, 3, and 4 clusters, the Silhouette method displayed a decreasing Calinski-Harabasz Index, implying that the clusters become less well-defined as their number increases. The Davies-Bouldin Index remained relatively steady, suggesting consistent similarity within the clusters. The Elbow method displayed lower performance with 14,

17, and 19 clusters, suggesting that many clusters may be too granular for this cell class.

- **Decidua (DEC) Cells.** The Silhouette method decreased performance as clusters increased, implying more distinct groupings within this class. The Davies-Bouldin Index from the Elbow method showed a relative consistency, suggesting similar similarity within clusters despite the different numbers of clusters.

- **Endothelial (END) Cells.** Indices varied less with changes in the number of clusters for this class, suggesting less evident distinctions between the clusters. Notably, the Silhouette method displayed a high Calinski-Harabasz Index, indicative of a stronger cluster definition.

- **Foetal (FET) Cells.** The Silhouette method, particularly with 2 clusters, showed a high Calinski-Harabasz Index, denoting a degree of separation between clusters. However, as the number of clusters increased, this value decreased, indicating that an increase in clusters may lead to less defined groups.

- **Giant (GIA) Cells.** This class displayed a relatively consistent Davies-Bouldin Index across different numbers of clusters and methods, implying a similar degree of similarity within clusters regardless of their number. However, the Calinski-Harabasz Index was lower than other classes, indicating less defined clusters.

- **Glycogen (GLY) Cells.**A fairly consistent Davies-Bouldin Index was observed across different cluster numbers and methods, similar to GIA cells. Notably, the highest scores were obtained with two clusters, which decreased with increased clusters.

- **Labyrinth (LAB) Cells.** The Elbow method displayed relatively stable performance across different cluster numbers, whereas the Silhouette method

decreased performance as the number of clusters increased. This suggests that the Elbow method might yield more information.

- **Spongiotrophoblast (SPO) Cells.** Both methods decreased performance as clusters increased, indicating that fewer clusters might result in better-defined groupings within this class. The high Calinski-Harabasz Index with two clusters in the Silhouette method supports this observation.

Overall, the methods fail to detect any substantial intraclass variance within the data, as the Silhouette score consistently predicts two clusters as the best option and the Elbow methods predict a high number of clusters as shown exmplaes of in Figure 7.9. The quantitative analysis of the problem needs to yield clear results. Therefore, the next phase entails a qualitative investigation, achieved by plotting clusters next to each other. Though all cell types and cluster numbers were plotted, only the notable examples are illustrated in the subsequent figure.

**Table 7.2:** *Determination of optimal number of clusters and evaluation of clustering quality for different cell types*

| Class | Method | Num Clusters | Calinski-Harabasz Index | Davies-Bouldin Index |
|---|---|---|---|---|
| ChP | Silhouette | 2 | 627 | 1.55 |
| ChP | Silhouette | 3 | 593 | 1.47 |
| ChP | Silhouette | 4 | 525 | 1.52 |
| ChP | Elbow | 19 | 256 | 1.41 |
| ChP | Elbow | 17 | 268 | 1.43 |
| ChP | Elbow | 14 | 297 | 1.42 |
| DEC | Silhouette | 2 | 1750 | 1.28 |
| DEC | Silhouette | 3 | 1394 | 1.41 |
| DEC | Silhouette | 4 | 1188 | 1.53 |
| DEC | Elbow | 16 | 566 | 1.53 |
| DEC | Elbow | 19 | 516 | 1.49 |
| DEC | Elbow | 12 | 665 | 1.44 |
| END | Silhouette | 2 | 2462 | 1.18 |
| END | Silhouette | 3 | 2041 | 1.29 |
| END | Silhouette | 4 | 1832 | 1.17 |
| END | Elbow | 18 | 848 | 1.44 |
| END | Elbow | 20 | 798 | 1.42 |
| END | Elbow | 14 | 962 | 1.43 |
| FET | Silhouette | 2 | 3223 | 1.09 |

Continued on next page

**Table 7.2 – continued from previous page**

| Class | Method | Num Clusters | Calinski-Harabasz Index | Davies-Bouldin Index |
|---|---|---|---|---|
| FET | Silhouette | 3 | 2384 | 1.39 |
| FET | Silhouette | 4 | 1983 | 1.41 |
| FET | Elbow | 17 | 967 | 1.39 |
| FET | Elbow | 18 | 938 | 1.37 |
| FET | Elbow | 20 | 890 | 1.4 |
| GIA | Silhouette | 2 | 214 | 1.16 |
| GIA | Silhouette | 3 | 195 | 1.14 |
| GIA | Silhouette | 4 | 162 | 1.32 |
| GIA | Elbow | 9 | 115 | 1.29 |
| GIA | Elbow | 18 | 89 | 1.26 |
| GIA | Elbow | 19 | 92 | 1.32 |
| GLY | Silhouette | 2 | 355 | 1.35 |
| GLY | Silhouette | 3 | 320 | 1.26 |
| GLY | Silhouette | 4 | 279 | 1.37 |
| GLY | Elbow | 20 | 125 | 1.37 |
| GLY | Elbow | 19 | 128 | 1.41 |
| GLY | Elbow | 15 | 144 | 1.43 |
| LAB | Silhouette | 2 | 1177 | 1.65 |
| LAB | Silhouette | 3 | 1127 | 1.44 |
| LAB | Silhouette | 4 | 1022 | 1.5 |
| LAB | Elbow | 19 | 481 | 1.45 |
| LAB | Elbow | 16 | 523 | 1.41 |
| LAB | Elbow | 18 | 495 | 1.47 |
| SPO | Silhouette | 2 | 3823 | 0.92 |
| SPO | Silhouette | 3 | 2922 | 1.27 |
| SPO | Silhouette | 4 | 2462 | 1.27 |
| SPO | Elbow | 19 | 1055 | 1.35 |
| SPO | Elbow | 15 | 1190 | 1.34 |
| SPO | Elbow | 11 | 1398 | 1.32 |

## 7.8 Investigation 4: Qualitative Investigation of Intraclass Variance

**Aim:** The investigation aimed to conduct a qualitative analysis of cellular composition within each identified cluster, focusing on the detection of intraclass variance.

**Method:** To explore intraclass variance more effectively, the SimCLR self-supervised learning framework was employed. SimCLR was trained using both labeled and unlabeled data to generate nuanced feature representations. Labeled data comprised training data annotated in

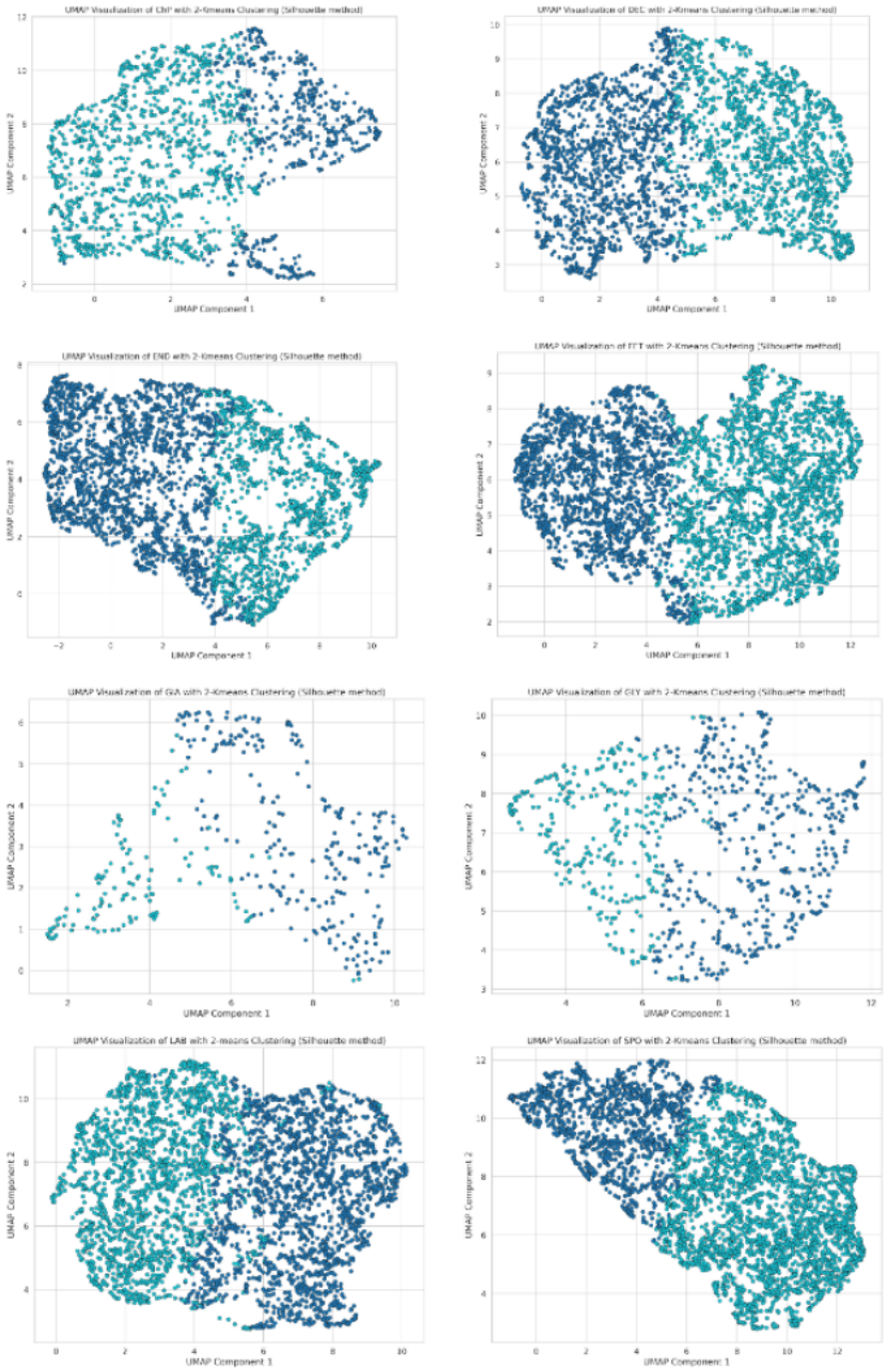


**Figure 7.4:** *This figure presents the results of applying UMAP visualization and K-Means clustering to analyze intraclass variance within eight distinct placental cell classes. Initially, a single model trained on all available data was used to extract features for each cell class separately. These features were then visualized using UMAP with the expectation of distinguishing separate clusters within each class. Subsequently, K-Means clustering was applied to the UMAP outputs with varying numbers of clusters (Ks) to determine the optimal clustering configuration. For each class, the silhouette scores consistently indicated that the highest intra-class differentiation was achieved with K=2. Thus, this figure illustrates these two K-Mean clusters for each class, revealing the challenge of discerning clear clusters and suggesting potential overlap and the need for more nuanced feature extraction methods. The consistent preference for two clusters by the silhouette measure across all classes highlights the subtle but important intraclass variances that necessitate further analytical exploration to fully understand the underlying biological processes of placental cells.*

Chapter 4, while the unlabeled dataset included 3 million samples obtained through a nuclei detection pipeline, where nuclei were detected, and a 100x100 pixel image was extracted around each centered nucleus. Features extracted from SimCLR were then processed using a K-means clustering algorithm across different values of $K$ to evaluate the presence of cluster formations within each cell class.

**Result:** Initial clustering attempts with K-means on top of SimCLR features revealed that the silhouette score consistently suggested the highest silhouette with $K = 2$ for all classes. Consequently, Figures 7.5 and 7.6 illustrate two K-mean clusters per class, underscoring the morphological nuances captured within these clusters. Despite the failure of silohette score to predict cluster seperation on a purely qualitative anysis we can see some seperation, Fig 7.5 shows 4 classes which show difference and 7.6 shows 8. These should not be taken as formal class seperation but instead show the method does clsuter cells with similar charicteristics together and thus has some degree of seperation ability

**Conclusion:** Subtle but distinctive morphological differences suggest a degree of intraclass variance. It indicates that under a refined classification scheme, these differences could be magnified leading to more pronounced intraclass segregation.

**Future Work:** Future strategies include further analysis with varying $K$ values to confirm the consistency of observed patterns and to refine the self-supervised learning framework for enhanced feature extraction capabilities. This could potentially facilitate the development of more granular and nuanced cellular classification systems.

In the pursuit of identifying intraclass variance within cell classes of mouse placental cells, a shift from supervised to self-supervised learning was necessitated. The initial approach using a supervised feature extractor did not yield the expected intraclass variance. It was posited that the supervised model was over-specialised towards identifying variance between different cell classes, rather than discerning the more subtle variance within them. The model's feature extraction capabilities, while adept at high-level classification, appeared insufficient for detecting finer, low-level intraclass differences.

The investigation progressed by employing SimCLR, a self-supervised learning approach, with the potential for a richer feature understanding. Self-supervised learning, by design, generates representations by capturing inherent patterns in the data without explicit class labels. This methodology allows the model to potentially

develop a more nuanced understanding of the low-level features within the cells, which could unveil the sought-after intraclass variance.

The forthcoming qualitative analysis aims to scrutinise the cellular morphology within each cluster identified by the self-supervised learning model. Figures 7.5 and 7.6 will provide illustrative examples of the cellular composition within clusters of Endothelial and Trophoblastic Giant cells, respectively.

Preliminary observations indicated an absence of stark intraclass variance across various cell types. Nonetheless, upon meticulous examination, the Trophoblastic Giant cells displayed subtle morphological variations. For instance, cluster 4 featured a preponderance of overlapping cells, whereas cells in cluster 8 were predominantly larger. Furthermore, cells constituting cluster 6 were distinguished by their elongated shape.

Although these patterns were not universally present, they were indicative of a certain level of intraclass variance, suggesting that cells of the same type could manifest distinct morphological traits when segregated into different clusters. This points to the potential of a more detailed classification approach to accentuate these fine differences, thereby achieving clearer intraclass distinction.

The observations thus far reinforce the notion that self-supervised learning could be advantageous for the qualitative evaluation of cell morphology. By leveraging self-supervised techniques like SimCLR, underlying intraclass variance may be uncovered, enriching our understanding of the intricate nature of cellular structures and their functional roles within the placenta.

Further research should be directed towards perfecting this approach. A refined self-supervised learning model may elucidate novel patterns and facilitate the establishment of an enhanced, more detailed cellular taxonomy. Such advancements could prove instrumental in deepening our comprehension of placental biology and contribute to the broader field of developmental biology.

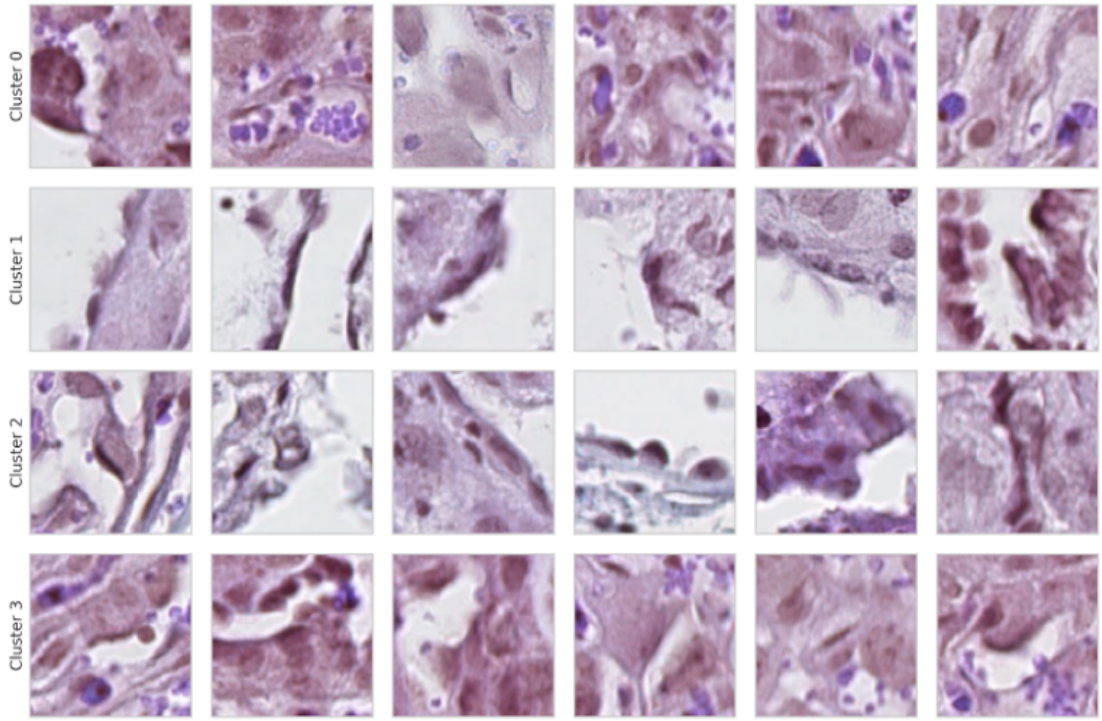



**Figure 7.5:** *This figure presents a representative visualisation of four identified clusters within the Endothelial cell class. Each subplot within the figure corresponds to a single cluster, displaying a grid of cells associated with that specific cluster. Notable morphological differences can be observed across the clusters, suggesting the presence of subtle intraclass variance within the Endothelial cell class.*

## 7.9 Investigation 5: Quantitative Assessment of Intraclass Variance in Placental Cells From a Self-Supervised Feature Extractor

**Aim:** The study aimed to explore the potential of self-supervised learning, specifically the SimCLR framework, in discerning intraclass variance among placental cell types.

**Method:** The SimCLR model was trained on two datasets: a labelled dataset for benchmarking performance against a supervised learning approach, and a larger, unlabeled dataset to test the model's effectiveness in a real-world scenario.

**Result:** The SimCLR model achieved 76% accuracy on the labelled dataset and 58% on the unlabeled dataset. However, neither the Silhouette score nor the Elbow method showed a clear separation of the eight-cell classes.

**Conclusion:** The SimCLR method, in its current implementation, may not be effective in capturing intraclass variance within placental cells. This could

be due to issues with fine-tuning, model architecture, hyperparameters, or the diversity of the training data.

**Future Work:** Future research could involve re-evaluating the model architecture and hyperparameters, training the model with a broader range of data, and exploring different self-supervised learning methods such as SwAV, MoCo, or BYOL.

After conducting an investigation that used a supervised feature extractor, it was discovered that there were limitations in detecting differences between similar types of placental cells. The model's training in distinguishing between broader categories might have affected its ability to identify more subtle differences within the same category. As a result, a hypothesis was formulated that the feature representation lacked a varied range of distinct attributes necessary for detecting such differences. Consequently, the focus of the analysis shifted towards self-supervised learning, with the SimCLR framework being selected to create a more diverse feature landscape potentially.

SimCLR is an approach to learning representations that focuses on the invariances present within data. It achieves this by enforcing consistency between differently augmented views of the same data point. SimCLR uses a self-supervised learning paradigm to provide a more detailed feature set to help distinguish between subtle differences in cell classes.

The usefulness of SimCLR comes from its contrastive learning technique, which efficiently uses unlabelled data to train models in distinguishing between different instances. This method can improve the model's understanding of low-level features and intraclass variance. The architecture was chosen for its robustness in previous studies (Chen et al., 2020) and its potential in domains with complex data structures, such as cellular imaging.

The accuracy of the SimCLR model was gauged using a labelled validation set. The model would assign labels and then compare them against the real labels to get the model's accuracy. The baseline of 8 different classes would be 12.5

SimCLR achieved a 76% accuracy on the labelled dataset. This dataset was the training set I annotated in the DMDD chapter but with the labels removed. The

data was very clean, and the unsupervised model achieved 76% accuracy, although not as good as the supervised model's 89% accuracy, which was still considered a strong result given that the training was performed without explicit labels.

A self-supervised model was trained on a dataset of 3 million placental images with no labels during subsequent experiments. This model was trained over 100 epochs and achieved an accuracy of 58%. This was lower than hoped for for self-supervised models. The confusion matrix for this result is shown in Figure 7.8.

The Silhouette scores indicate that the model has difficulty distinguishing the eight-cell classes. Additionally, the high number of clusters suggested by the Elbow method's inertia graph needed to be supported by the UMAP visualisation, which needed clear cluster separation. This shortfall may be due to various factors, such as the need to fine-tune the model's hyperparameters or the limitations of the model's architecture for this particular task.

The 58% accuracy on the unlabelled dataset falls short of a definitive benchmark set by the supervised trained model of 89%. It reflects the task's challenging nature and the data's complexity. Given the exploratory nature of self-supervised learning for intraclass variance detection, this accuracy is a preliminary benchmark for future work.

Addressing the shortcomings identified will involve meticulous refinement or change of the model's architecture and training regimen. Furthermore, alternative self-supervised learning frameworks might be explored for their efficacy in this context.

## 7.10 Summary and Conclusion

This chapter has undertaken an extensive exploration of machine learning techniques, with a particular focus on dimensionality reduction and clustering methodologies, to examine the interclass and intraclass variance within placental cells. The primary objective of this investigation was to discern meaningful patterns and groupings within high-dimensional data, thereby contributing to a deeper understanding of cell morphology differentiation.

Complexity in placental cellular analysis is highlighted by the variability observed in specific cell types such as giant cells, migratory Glycogen cells, and sinusoidal giant cells, along with the distinct layers of syncytiotrophoblasts (syn1 and syn2). These variations within cell subtypes underline the intricate nature of placental functionality and heterogeneity. Successfully applying unsupervised and self-supervised learning methods could greatly enhance our understanding of these complexities, offering a more nuanced view of cellular behaviors and their implications for placental health. This approach aims to align computational insights with the biological intricacies observed in placental cells, potentially leading to breakthroughs in predictive accuracy and biological relevance.

In the dimensionality reduction realm, Principal Component Analysis (PCA), t-Distributed Stochastic Neighbor Embedding (t-SNE), and Uniform Manifold Approximation and Projection (UMAP) techniques were evaluated. Given their distinct properties and the high-dimensional nature of the dataset, UMAP emerged as the most effective technique. It demonstrated a proficient capability in capturing both global and local structures in the dataset, providing valuable insights into the underlying structure of the high-dimensional data.

For the clustering analysis, methodologies such as k-means, spectral clustering, Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), and Gaussian Mixture Models (GMM) were employed. Among these, spectral clustering stood out, exhibiting the most promising performance by effectively identifying complex structures within the placental cell data.

Despite these promising results, the examination of intraclass variance presented a more challenging task due to the high dimensionality and complexity of the data. Initially, supervised learning methodologies were employed, but they demonstrated limited success in discerning intraclass variance. This led to the exploration of self-supervised learning methods, particularly the SimCLR framework. While the results did not yield definitive success in capturing intraclass variance, they provided valuable insights into the potential and limitations of such techniques.

These findings underscore the importance of method selection in alignment with the nature of the data and the study's objectives. It is clear that future research could benefit from exploring a variety of avenues to enhance the detection of cell morphology differentiation further.

The optimisation of parameters for existing methods is one viable path. This would involve a systematic fine-tuning of the parameters to better align with the specific requirements of the task at hand. This could potentially enhance the performance of the methods and lead to a more accurate detection of interclass and intraclass variance.

Another promising direction for future research is the evaluation of other self-supervised learning methods. Techniques such as SwAV, Momentum Contrast (MoCo), or Bootstrap Your Own Latent (BYOL) could be explored for their ability to discern intraclass variance in placental cells. These methods, like SimCLR, leverage the power of self-supervision and may offer unique strengths that could contribute to a better understanding of the data.

Incorporating domain-specific knowledge into the machine learning methodologies is another strategy that could yield more insightful results. By integrating domain-specific insights into the analysis, the methodologies could be guided towards more meaningful and interpretable results, thereby enhancing the detection of cell morphology differentiation.

In conclusion, this study contributes to the ongoing discourse on leveraging advanced machine learning techniques to enhance our understanding of complex biological data, particularly in the context of placental cell morphology differentiation. Despite the mixed results, this exploration has paved the way for future investigations, highlighting the potential of these advanced techniques and the importance of continual refinement in alignment with the specific demands of the task.

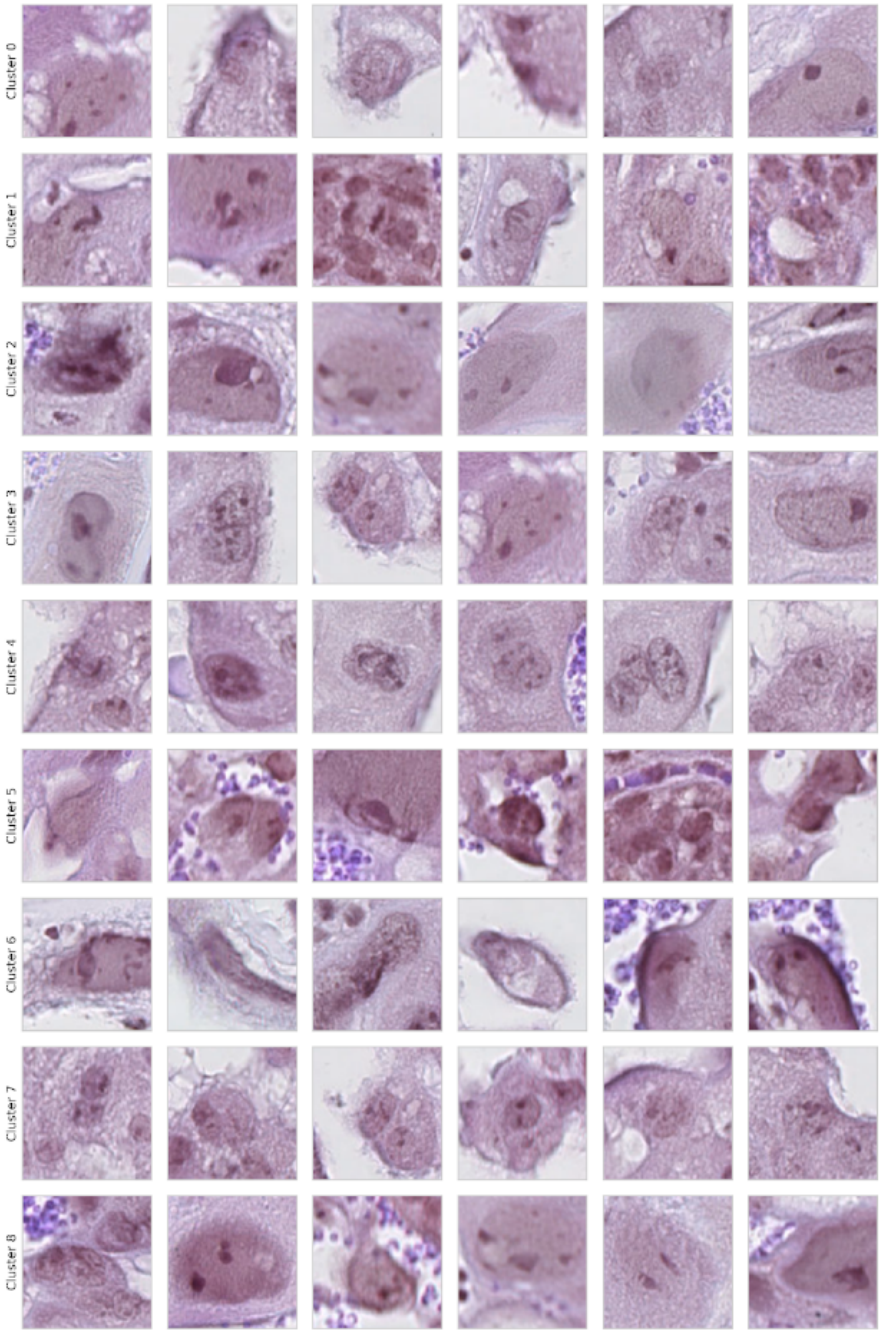



**Figure 7.6:** *This figure depicts a representative visualisation of nine identified clusters within the Trophoblastic Giant cell class. Each subplot corresponds to a single cluster, presenting a grid of cells belonging to that specific cluster. The clusters display distinctive cellular patterns, such as overlapping cells in cluster 4, notably larger cells in cluster 8, and cells with elongated morphology in cluster 6, suggesting a degree of intraclass variance within the Trophoblastic Giant cell class.*

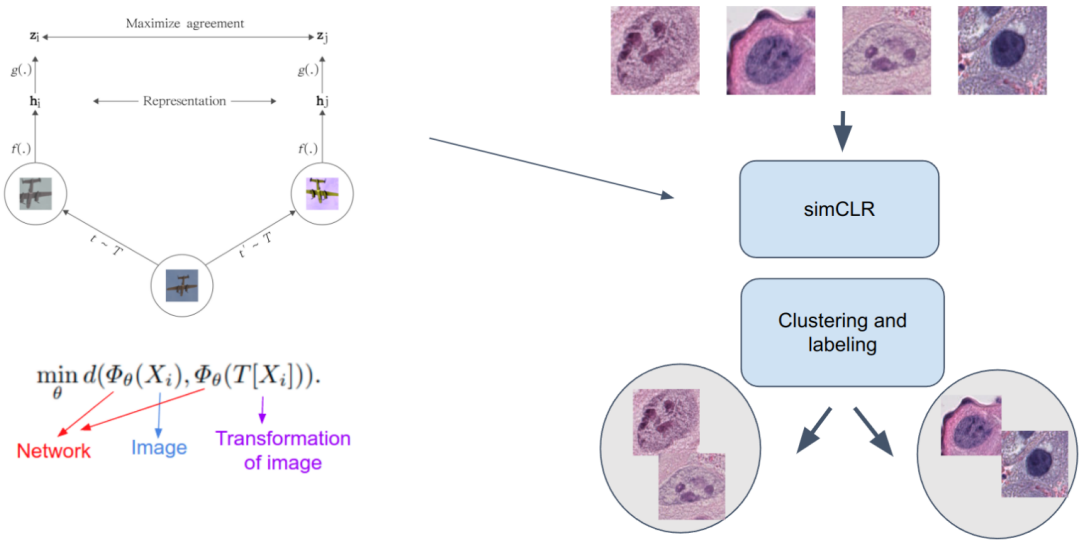



**Figure 7.7:** *This figure illustrates the application of the Simultaneous Contrastive Learning of Representations (SimCLR) framework in the detection of intraclass variance among mouse placental cells. As a self-supervised learning model, SimCLR extracts distinctive low-level features by recognising augmented versions of the same data instance as similar and different instances as dissimilar. The feature extractor's output is subsequently used for clustering and labelling the cells. Despite the model's inability to effectively capture intraclass variance in its current implementation, this representation underscores the potential of self-supervised learning methodologies in enhancing our understanding of cellular heterogeneity.*

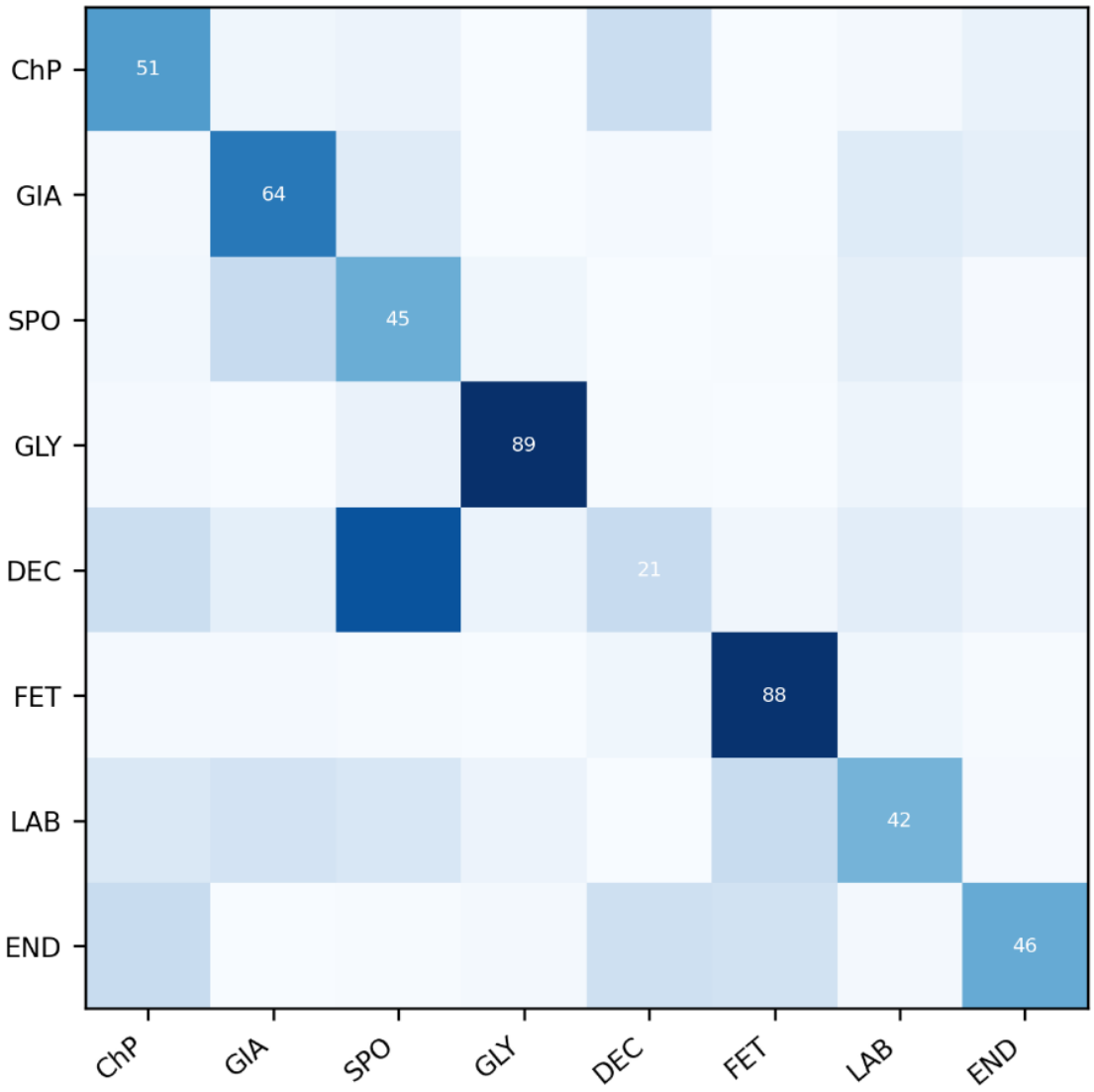



**Figure 7.8:** *This figure presents a confusion matrix of the classification results of the SimCLR model on eight mouse placental cell classes. The SimCLR model was trained on an unlabeled dataset comprising 3 million placental images and validated on a labeled dataset, yielding an accuracy of 58%. Each cell in the matrix represents the model's predictions (columns) against the true labels (rows). The diagonal elements represent instances where the predicted label matches the true label, while off-diagonal elements are instances of misclassifications. Despite the decreased performance compared to the supervised approach, this confusion matrix demonstrates the model's ability to identify certain cell classes, indicating the potential of self-supervised learning methodologies in examining cellular heterogeneity.*

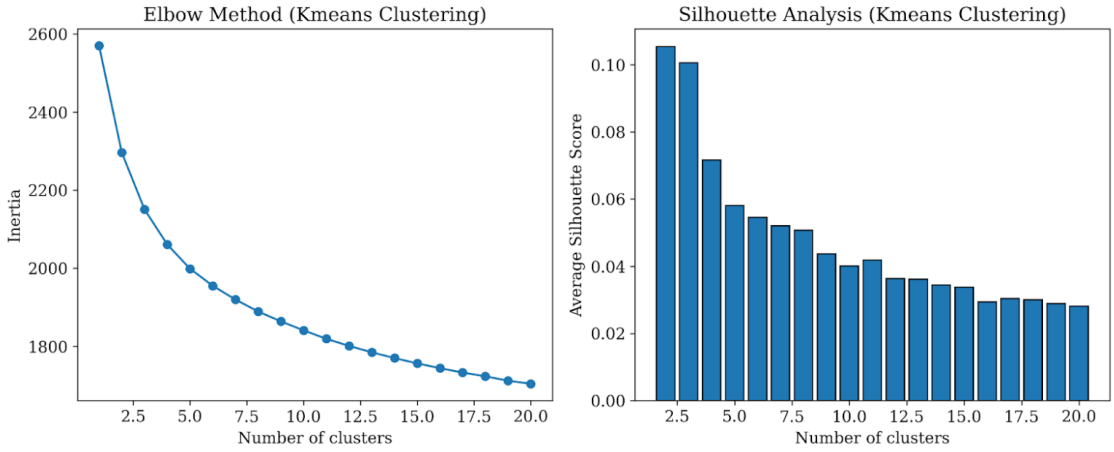



**Figure 7.9:** *This figure showcases the results of the Elbow method and Silhouette scores for determining the optimal number of clusters within each of the eight mouse placental cell classes. The Elbow method plots the total within-cluster sum of squares against the number of clusters, with the optimal number of clusters corresponding to the 'elbow' point in the curve. The Silhouette score measures the quality of clustering, with values closer to 1 indicating better-defined clusters. Despite the SimCLR model's respectable classification performance, both methods failed to clearly distinguish the eight-cell classes. Most notably, the Silhouette score frequently suggested just two clusters, indicating a limited discernment of intraclass variance. This figure underscores the challenges in employing unsupervised learning techniques to elucidate the complex heterogeneity within placental cells.*

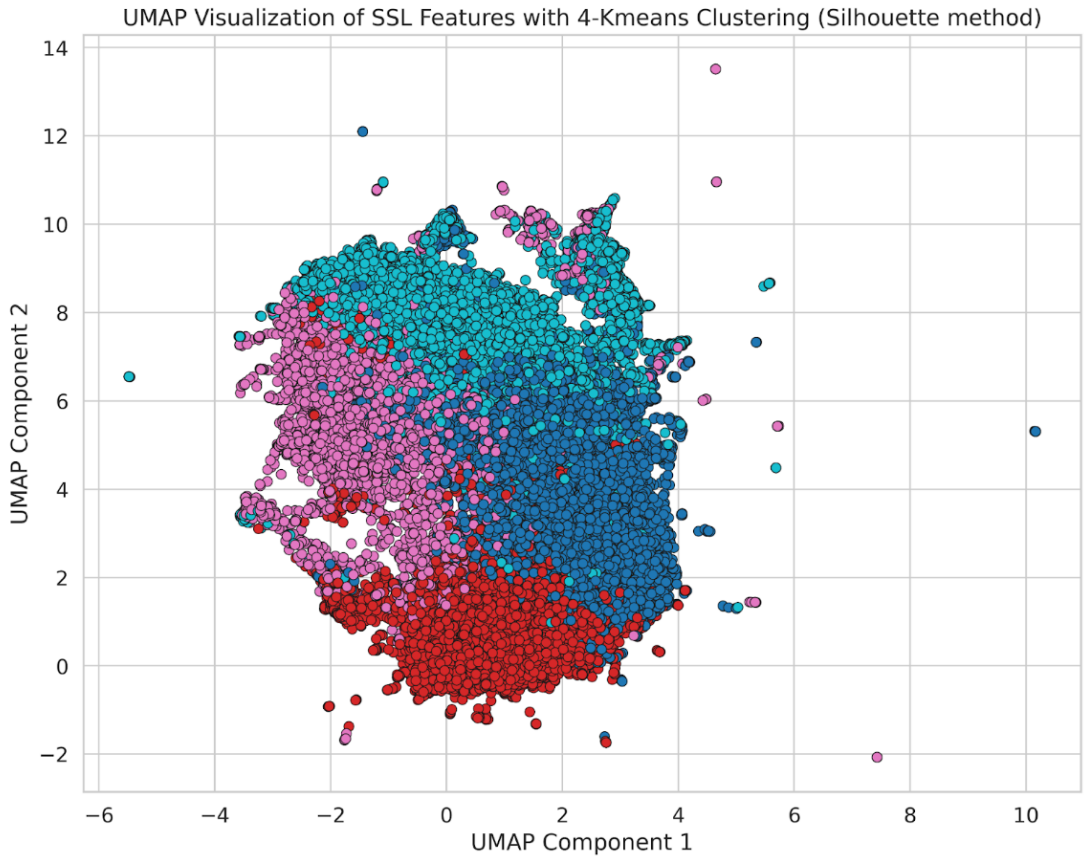



**Figure 7.10:** *This figure presents the Uniform Manifold Approximation and Projection (UMAP) visualisation of the eight mouse placental cell classes with K-means clustering set to 4. Each point represents a single cell, with different colours denoting different cell classes. Despite the application of K-means clustering and the use of the self-supervised SimCLR model for feature extraction, the UMAP visualisation does not show clear, distinct clusters. This observation suggests that the current approach might not effectively capture the intraclass variance within these cell classes. The figure underscores the challenges in discerning the subtle differences within cell classes using unsupervised learning techniques and the need for further investigation and optimisation of these methods.*

*"Medicine is a science of uncertainty and an art of probability, and AI is its newest brush"*

—

# 8 Discussion and Conclusion

## Contents

---

This chapter will review my thesis by discussing and concluding the key research themes. To begin, I will restate the historical context. I will revisit the research questions and describe the methods and investigations employed in each research chapter. Finally, I will conclude the chapter by offering an overall conclusion and discussing the scope for future work.

## 8.1 Context of the Research Recap

The placenta exemplifies biological complexity, which is critical in nutrient transport and protects the maternal and fetal immune systems. Despite its crucial role, the scientific community still needs to explore it. My research has highlighted the placenta's significance, focusing on its multifaceted role and impact on severe conditions such as pre-eclampsia, gestational diabetes, and intrauterine growth restriction. These conditions not only have immediate health implications but may also lead to long-term maternal and fetal complications, underscoring the necessity of placental research in global public health.

Historically, the placenta has been studied primarily through histological investigations. However, these methods have significant limitations. They often yield subjective evaluations, leading to potential inconsistencies. Additionally, traditional methods are time-consuming and can hinder timely clinical interventions crucial for optimal patient outcomes.

In addressing these challenges, it is imperative to incorporate more efficient and objective methodologies. The advent of machine learning and computer vision presents a promising avenue. These advanced computational techniques can detect complex histological patterns often surpassing human observational capabilities. By utilising these state-of-the-art methods, my thesis proposes a shift in histological evaluations, aiming to surmount the constraints of traditional approaches.

## 8.2 The Research Aims of the Thesis

Having recapped the context, the research in my thesis was structured around the following interconnected aims:

1. **Processing and Preparation of the DMDD Placental Dataset**
   The first goal of my thesis involved the detailed processing of the DMDD placental dataset. This step included meticulous data cleansing, annotation, and standardisation, setting the stage for advanced analytical methods.

2. **Development of an Automated Pipeline for Placental Histology**
   My second objective was to create an innovative pipeline for placental histological analysis. This pipeline, leveraging modern machine learning techniques, aimed to outperform traditional methods in both speed and accuracy.

3. **Validation of the DMDD Phenotyping Pipeline**
   The third aim was to evaluate the newly developed phenotyping pipeline. Using the DMDD dataset and benchmarks from the Hemberger Lab, this evaluation focused on assessing the pipeline's effectiveness and identifying opportunities for further improvements.

4. **Investigating Genetic Dynamics in Placental Development**
   Another critical aspect of my research was to enhance understanding of the genetic factors influencing placental development. By conducting comprehensive phenotyping of placental data and correlating it with embryonic outcomes from the HREM annotated fetal dataset, I sought to unravel the complex interactions between developmental genes, placental structures, and fetal outcomes.

5. **Advanced Learning Techniques in Placental Histology**
   The final aim of my thesis was to explore placental cellular morphology through cutting-edge unsupervised and self-supervised learning techniques. This exploration aimed to discover novel insights into placental histology, potentially contributing to the field in significant ways.

In the following sections, I will further explore these objectives, discuss my key findings, and reflect on the wider implications of my research within the scientific community.

## 8.3 Detailed Processing and Preparation of the DMDD Placental Dataset: Discussion and Conclusion

In Chapter 4 of my thesis, I focused on the comprehensive processing of the DMDD placental dataset, a vital resource in mouse developmental biology. This task, central to the first aim of my research, involved meticulous data cleansing, annotation, and standardisation, preparing the dataset for advanced analysis. The significance of this work lies in its potential to facilitate groundbreaking insights in the field of developmental biology, particularly in understanding the complex dynamics of placental development and its broader implications.

Addressing the original research questions, my efforts in data processing culminated in a refined dataset, initially comprising 8713 samples, which was streamlined to 8163 after rigorous post-processing. This refined dataset included 7242 labelled

and 921 unlabelled samples, encompassing observations from 925 unique animals and targeting 103 distinct genes. The developmental stages were categorically divided, with 3262 samples at stage E14.5 and 3980 at stage E9.5. This meticulous categorisation was extended to genotype and gender, offering a comprehensive overview of the dataset's diversity.

However, this process had its limitations. The subjective nature of manual annotation, a necessary step in dataset preparation, introduced potential inconsistencies. This limitation, coupled with time constraints that prevented the annotation of all cell types and subtypes, suggests that future computational analyses may need to be more comprehensive. Furthermore, the inherent complexity of cellular structures and the reliance on visual assessment underscore the need for more robust, perhaps automated, annotation methods in future research.

The potential impact of this refined dataset is significant. When interfaced with computational tools, it could offer transformative insights into developmental biology, reshaping our understanding of genetic dynamics in placental development. Such insights are invaluable, not only for academic advancement but also for informing therapeutic strategies and improving maternal and fetal health outcomes.

In terms of future research, this work highlights the importance of developing more objective and automated methods for data annotation. Integrating computational techniques with human expertise could enhance the accuracy and completeness of datasets like DMDD. This integration could pave the way for more comprehensive studies and, potentially, the discovery of novel insights into placental biology.

## 8.4 Development of an Automated Placental Phenotyping Method: Discussion and Conclusion

In this section of the discussion and conclusion chapter, I revisit the central aim of developing a state-of-the-art phenotyping pipeline for placental histological analysis. This aim was not just a technical endeavour but a response to a critical research question: How can modern machine learning enhance the speed and accuracy of placental histology beyond the capabilities of traditional methodologies? Pursuing

this question led to the creation of a system designed to utilise the power of advanced computational techniques in a field where traditional methods have shown limitations.

The development of this automated pipeline was an intricate process. It began with processing whole-slide histology images. I then employed the RetinaNet model for identifying nuclei, enhancing it with transfer learning and extensive data augmentation techniques. These steps were crucial in ensuring the model could accurately interpret histological variations such as stain depth, a fundamental requirement given the diverse nature of placental tissues.

One of the innovative aspects of the pipeline was the use of overlapping tiles to address challenges at tile edges, a common issue in image processing. This approach ensured data accuracy and showcased the potential of computational methods in overcoming obstacles often encountered in manual analysis. Incorporating the InceptionResNetv2 architecture for cellular classification further highlighted the technical sophistication of the pipeline.

The process culminated in tile integration, providing valuable insights into cellular density and ratios, critical factors in understanding placental morphology. This step was vital as it brought together the various elements of the pipeline, demonstrating the capacity of machine learning to offer a comprehensive analysis of complex biological data.

### 8.4.1 Investigation 1: Replicating the Ferlaino et al. results

In this investigation, I focused on replicating the seminal work of Ferlaino et al. to validate the reliability and consistency of their findings within the scientific community. Utilising Keras and TensorFlow, I mirrored their experimental setup, achieving a mean average precision (mAP) of 65% in nuclei detection and 89% accuracy in classification. This replication effort reaffirms the credibility of Ferlaino et al.'s methods and highlights the robustness of the machine learning frameworks employed. It demonstrates the reproducibility of results in computational biology, a critical factor in establishing trust and a foundational base in scientific research.

### 8.4.2 Investigation 2: Cross-Species Nuclei Detection

This segment of my research involved attempting to capture both human and mouse nuclei detection, a challenging task due to the inherent differences in the biological structures of the two species. Using the supervised learning models initially resulted in a low mAP of 1.85% when applied to mouse data using a model trained on human data. This outcome reveals a significant limitation in the generalisation capabilities of supervised models and underscores the disparity between human cognitive processing and machine learning algorithms.

The model's difficulty adapting to a different species highlights a fundamental challenge in machine learning – the ability of a model trained on one domain to generalise to another. This serves as a reminder of the limitations of current machine-learning techniques in dealing with nuanced biological variations.

This finding is crucial for researchers utilising machine learning in cross-species analyses, emphasising the need for specialised models or advanced transfer learning techniques to better adapt to such variations.

### 8.4.3 Investigation 3: High-Performance Mouse Nuclei Detector

In this investigation, I aimed to enhance the detection of mouse nuclei by redeveloping the approach. This shift significantly improved the outcome, achieving a mean average precision (mAP) of 87% for mouse nuclei detection. This advancement was realised through extensive trials with different model backbones, hyperparameters, and a comprehensively annotated mouse nuclei dataset.

The successful development of a high-performance mouse nuclei detector holds substantial implications for biological research, particularly in developmental biology and disease research. It demonstrates the potential of machine learning to provide accurate and efficient analysis tools in complex biological studies.

### 8.4.4 Investigation 4: Generalisable Whole Slide Nuclei Detector for Mouse Placenta

This investigation aimed to develop a generalisable model for nuclei detection in mouse placenta, capable of consistent performance across diverse tissue samples within the DMDD dataset. Techniques such as bootstrapping and hard negative mining were employed, significantly improving the model's performance.

The model's success demonstrates the effectiveness of these techniques in machine learning. However, it also highlights the challenges in creating a truly generalisable model, as achieving consistent performance across varied samples often requires extensive data and rigorous training.

### 8.4.5 Investigation 5: Removing Multiple Detections of Identical Nuclei

This investigation addressed the challenge posed by multiple detections of identical nuclei due to the tile overlap method in image processing. I employed the K-D tree clustering method to resolve this issue effectively. This approach provided insights into balancing comprehensive coverage and accurate detection in large-scale image analysis.

The success of the K-D tree clustering in consolidating multiple detections illustrates its efficacy in complex image processing tasks. However, its performance depends on the chosen clustering radius, indicating a need for adaptable algorithms tailored to specific data characteristics.

The ability to accurately consolidate multiple detections has practical implications for enhancing the precision of large-scale histopathological analysis. It represents an advancement in computational biology, offering a more efficient and reliable approach to processing complex image data.

Future research could explore dynamic clustering algorithms that adjust the clustering radius based on the variable sizes of nuclei in different samples. This would enhance the versatility of the method and its applicability to a broader range of biological datasets.

### 8.4.6 Investigation 6: Comprehensive Nuclei Detection in the DMDD Dataset

The application of the mouse nuclei detection model, augmented with the overlay technique, demonstrated significant progress in understanding tissue morphology. The model's consistent detection across various placental tissue types and stain depths in the DMDD dataset highlights the potential of generalised computational approaches in biomedical research.

While the model exhibited impressive performance, it is crucial to acknowledge its limitations regarding scalability and potential biases inherent in the training data. These factors may influence the model's applicability to tissue types or staining methods.

The success of this model in detecting nuclei across the DMDD dataset can significantly expedite the process of histological sample analysis, offering a tool for more efficient and precise large-scale investigations in developmental biology and disease research.

Further refinement of the model to include a broader range of tissue samples and regions, and its applicability to datasets beyond DMDD, would be valuable areas for future research. This could be done using semi-supervised learning methods. Additionally, addressing the limitations regarding scalability and bias will be crucial in enhancing the model's effectiveness.

### 8.4.7 Investigation 7: Human Placenta Cell Classification Model

In this investigation, I aimed to improve the cell classification model initially presented by Ferlaino et al. The transition from Keras to PyTorch, coupled with advanced data handling techniques, resulted in discernible improvements in the model's accuracy.

The enhanced performance of the model underscores the dynamic nature of machine learning in biomedical applications. However, it also highlights the

continuous need for model refinement, given the complexities of biological data and the ever-evolving landscape of computational tools.

The improved accuracy of the model has significant implications for histopathological analysis of human placental samples, potentially leading to better diagnostic outcomes and advancing research in prenatal health.

Future efforts could refine the model by exploring more sophisticated machine learning algorithms, expanding the training dataset, and integrating domain-specific knowledge.

### 8.4.8 Investigation 8: Mouse Placenta Cell Classification Model

The final investigation of my thesis aimed to develop a robust model for classifying mouse cells within the DMDD dataset, with a specific goal of accurately differentiating between eight distinct cell types. After thoroughly evaluating various image classification architectures, I selected the InceptionResNetV2 architecture, which proved to be most effective in processing 200x200 pixel images centred around the nucleus.

The decision to use InceptionResNetV2, which has a complex neural network capable of detecting small details, was crucial. However, during the investigation, it became apparent that achieving high accuracy in classification across different developmental stages was challenging. Although the model performed well for E14.5 placenta cells, it was less effective for E9.5 tissue cells, highlighting the inherent variability in biological data and the difficulty it presents in creating machine learning models that can be generalised.

The practical implications of achieving an 87% classification accuracy across eight cell classes are significant, enhancing the reliability and efficiency of automated classification systems in histopathological analysis. This development holds promise for reducing the manual annotation workload for pathologists and researchers, allowing a more focused application of human expertise.

Addressing the performance gap in classifying E9.5 tissue cells is critical for future research. Expanding and annotating the dataset for these earlier developmental stages could improve the model's accuracy and extend its applicability. Moreover, integrating domain-specific knowledge, such as stain normalisation techniques, could further refine the model's learning process.

### 8.4.9 Chapter 5 Overall

In Chapter 5 of my thesis, I developed a state-of-the-art phenotyping pipeline for placental histological analysis. The purpose of this pipeline was to leverage modern machine learning techniques to improve the speed and accuracy of traditional methodologies. This was not only a technical challenge but also a crucial step in achieving the overarching research goal of enhancing the efficiency and precision of histological evaluations.

The creation of this automated pipeline marks a significant departure from traditional patch-based analysis, as it employs whole-slide imaging. This approach allows for a more holistic examination of placental tissues, potentially offering deeper insights into various histological features. However, this shift to a more expansive imaging technique also brought challenges, particularly in managing and processing the substantially larger data volumes inherent in whole-slide images.

After conducting a thorough evaluation of the results, it is evident that the pipeline has made significant improvements in mean average precision (mAP) and classification accuracy. However, it is important to note that these technical achievements come with certain limitations. One of the significant challenges faced was ensuring the robustness of the model across different samples and staining variations, which is a common issue in computational pathology. This limitation emphasises the need to continuously improve and refine the pipeline to ensure its adaptability and reliability in different research settings.

The implications of these advancements in automated placental histology are far-reaching. On a practical level, the pipeline offers the potential to significantly reduce the manual workload involved in histological analysis, thereby enabling

more rapid and reliable assessments of large datasets. Academically, it contributes substantially to biomedical research, successfully integrating theoretical machine-learning concepts with practical histopathological applications.

Looking ahead, there are several promising directions for future research. Firstly, expanding the training datasets is essential. The performance and adaptability of machine learning models depend heavily on the diversity and quality of the data used for training. To this end, efforts should be directed towards increasing the data volume and enriching the dataset with a wider range of tissue samples and developmental stages. Additionally, improving the quality of annotations is crucial; more detailed and accurate annotations will enable the model to learn from a broader spectrum of examples, enhancing its predictive accuracy.

Furthermore, the variation in model performance across different species and tissue types suggests a pressing need for domain adaptation techniques. Multi-task learning could be a viable approach to developing a model capable of generalising across various biological contexts. By training a shared model on multiple related tasks, it might be possible to leverage their commonalities to improve overall performance.

Another promising area of research involves exploring the potential of transformer architectures, renowned for their success in natural language processing, for image analysis tasks. Integrating transformer backbones into the pipeline could significantly boost the model's ability to process and interpret complex tissue structures, paramount in enhancing classification and detection accuracy.

Lastly, variations in staining procedures can adversely affect model performance. Developing advanced image normalisation techniques could mitigate this issue, enabling the model to focus on essential morphological and textural features regardless of staining inconsistencies.

In summary, while the automated pipeline developed in this thesis represents a significant stride in computational histopathology, it also opens up many opportunities for further research and improvement. The lessons learned and the knowledge gained from this research endeavour provide a solid foundation for

future explorations aimed at refining and enhancing the capabilities of machine learning in placental histology.

## 8.5 Automated Analysis of the DMDD Placental Data: Discussion and Conclusion

In Chapter 6 of my thesis, I focused on two critical aspects: validating the DMDD phenotyping pipeline against established metrics and better understanding the genetic dynamics of placental development. This chapter is integral to my thesis, as it ties together the overarching research aim: to enhance the understanding of placental development and its implications on embryonic outcomes.

The first part of this chapter centred on evaluating the effectiveness of the phenotyping pipeline developed in earlier chapters. Using the DMDD dataset, I compared my findings against the benchmarks established by the Hemberger Lab. This comparison was essential to assess the reliability of the pipeline and identify potential areas for improvement.

Upon critical evaluation, while the pipeline showed considerable promise in accuracy and efficiency, certain limitations became apparent. One of the primary challenges was ensuring consistency across different samples. Variabilities in sample preparation and staining sometimes influenced the results, underscoring the need for a more robust approach to accommodate such variations. Additionally, the complexity of the placental tissue structures posed challenges in phenotypic classification, indicating a need for further refinement in the algorithm to enhance its precision.

In the second part of the chapter, I explored the genetic dynamics involved in placental development. To achieve this, I conducted a thorough analysis of the placental data and correlated it with the embryonic outcomes from the annotated fetal dataset using HREM. The goal was to understand better the complex relationship between developmental genes, placental architecture, and fetal outcomes.

This investigation revealed significant correlations between specific placental phenotypes and developmental outcomes, highlighting the placenta's critical role in

embryonic development. However, the complexity of these relationships posed a significant analytical challenge. Advanced machine learning techniques provided some clarity, such as Random Forest and XGBoost. However, the intricacies of genetic interactions in placental development indicate that there is still much to be understood in this field.

### 8.5.1 Investigation 1: Computational Tissue Segmentation in Mouse Placenta

Throughout this investigation, I was able to successfully modify the concave hull approach in order to segment tissue in mouse placenta specimens. This was a significant advancement in automating histological analyses, as the adapted algorithm was able to replicate some of the more intricate assessments that pathologists typically perform. By doing so, the algorithm addressed common issues such as human error and inefficiencies in traditional methods.

The algorithm significantly improved the quantification of cell densities and placental dimensions, which is crucial for accurate histological analysis. However, it was observed that the algorithm was less effective at the E9.5 developmental stage, which highlights the need for further refinement. While the success at the E14.5 stage is commendable, it underscores the importance of developing versatile algorithms that are capable of handling variations across various developmental stages.

This study enhances the application of geometric techniques in computational biology, explicitly validating the concave hull method for biological structure representation. It offers a streamlined workflow for pathologists, particularly for E14.5 mouse placenta, facilitating the transition from segmentation to analysis. The practical application of this methodology signifies a potential shift in how histological analyses are conducted, moving towards more automated and accurate processes.

Future efforts should focus on optimising the algorithm for earlier developmental stages, such as E9.5, and refining cell classification models to extend its applicability. Additionally, exploring the algorithm's adaptability to other tissue types would be a valuable area of research.

### 8.5.2 Investigation 2: Mouse Placenta Wild-type Characterisation

This investigation established a benchmark for the placenta in wild-type mice, providing a quantitative basis for studying the interplay between genetics and placental morphology. Establishing cell ratios and placental dimensions is essential for understanding genetic or environmental influences on development.

While characterising wild-type placenta is instrumental, challenges such as outliers in cell ratio data and measurement variances were encountered. These issues underline the necessity for enhanced accuracy and consistency in data collection and analysis.

The findings emphasize the importance of establishing a wild-type baseline in genetic research, particularly in mouse models, which are widely used. This provides practical insights for gene knockout studies, where deviations from the wild-type baseline can help understand the role of specific genes in placental development. This study is one of the first to characterise wild-type mouse placenta histology computationally.

Future research should use anomaly detection techniques to mitigate outliers and enhance measurement consistency. Moreover, a broader range of environmental and genetic conditions should be considered to understand their effects on placental morphology comprehensively.

### 8.5.3 Investigation 3: Sexual Dimorphism in Wild-type

The study of wild-type placental samples revealed significant differences between male and female specimens, with male placentas being larger and containing more glycogen cells. These findings highlight physical, molecular, and cellular disparities between sexes, suggesting the placenta's role in differential fetal growth rates and vulnerability.

The discovery of sexual dimorphism in the placenta challenges existing paradigms and invites a reevaluation of placental function and its influence on fetal development.

However, limitations such as the need for larger sample sizes and consideration of environmental factors must be addressed in future studies.

Theoretically, this investigation emphasises the placenta's role as a sexually dimorphic organ with potential implications for fetal programming. Practically, understanding these differences could lead to personalised prenatal care interventions.

Further research must elucidate the mechanisms behind sex-based differences in placental development and their impact on post-birth growth and development.

### 8.5.4 Investigation 4: Phenotypic Analysis of E14.5 Gene Knockout Mouse Models

In this investigation, I analysed 66 knockouts in E14.5 mouse models, identifying 28 statistically significant phenotypes. This provided a foundational understanding of the genetic basis of developmental disorders.

While ensuring robustness, the investigation's strict statistical thresholds may have limited the identification of subtler phenotypic changes. This suggests a need for a more balanced statistical approach.

Theoretically, the study contributes insights into the relationship between gene function and placental development. Practically, it confirms the automated pipeline's reliability in detecting phenotypes, which is promising for its application in large-scale studies.

Future studies should consider a balanced approach to statistical correction methods and analyse a broader array of phenotypic features to uncover more nuanced effects of gene knockouts.

### 8.5.5 Investigation 5: Predicting Embryonic Outcomes from Placental Phenotypes

In the fifth investigation, Random Forest and XGBoost models were used to determine the ability of placental phenotypes to predict embryonic outcomes. The study achieved moderate predictive accuracy, with the labyrinth area identified as a significant feature. Although the accuracy of 63% is a decent starting point, there is

room for improvement. The classification outcomes were imbalanced, highlighting the challenges of training models with uneven data distributions.

These results have potential applications in prenatal care, suggesting that specific placental characteristics can indicate fetal growth issues. However, the limitations in predictive accuracy necessitate further refinement of the models.

To improve the predictive performance, it is essential to enhance the dataset size, particularly for underrepresented outcomes. One can also explore more sophisticated algorithms and oversampling techniques. A deeper understanding of embryonic development and placental function can be gained by investigating the physiological implications of the labyrinth area and other tissue areas. Additionally, the DMDD has fetal annotation data of specific organ phenotypes that could be linked to placental phenotypes, assuming a responsible sample size.

### 8.5.6 Chapter 6 Overall

In Chapter 6 of my research, I aimed to achieve two goals: firstly, to validate the DMDD phenotyping pipeline and secondly, to explore the genetic dynamics of placental development. This chapter brings together these two elements, contributing to my overall objective of improving the understanding of placental development and its impact on embryonic outcomes.

To validate the DMDD phenotyping pipeline, I compared my findings with established benchmarks from the Hemberger Lab. This comparison was essential in determining the reliability of the pipeline and identifying areas that required improvement.

While the pipeline demonstrated promise in accuracy and efficiency, it faced challenges maintaining consistency across diverse samples due to sample preparation, staining variabilities, and developmental stage. These challenges highlight a significant opportunity for future research to develop more robust algorithms that can adapt to such variations. Refining the phenotypic classification methods to cater to the complexities of placental tissue structures can further enhance the pipeline's precision. Future studies could focus on algorithmic adaptability, ensuring that such systems can handle a range of developmental stages and tissue types.

During the second phase of my study, I aimed to uncover the complex relationship between developmental genes, placental architecture, and fetal outcomes. Through this exploration, I discovered significant correlations between placental phenotypes and developmental consequences, highlighting the crucial role of the placenta in embryonic development.

However, due to the complexity of genetic interactions, there's still a gap in our understanding, presenting an opportunity for future research. While advanced machine learning techniques like Random Forest and XGBoost provided valuable insights, there's still much to learn about the intricate nature of genetic interactions in placental development. This necessitates further studies, using even more advanced computational methods or multidisciplinary approaches, to fully decipher these relationships.

Theoretically, this research contributes significantly to our understanding of placental biology and embryonic development. Practically, these findings could inform clinical practices and help us better understand developmental biology.

My studies have shown that the predictive accuracy obtained is moderate. This indicates that there is still room for improvement. To enhance the predictive models, future research could focus on expanding the size of the dataset, especially for underrepresented outcomes. Investigating more advanced algorithms and data-handling techniques would also be beneficial. Additionally, a deeper exploration of the physiological implications of placental features could provide novel insights into fetal development and the role of the placenta.

In conclusion, Chapter 6 has significantly contributed to our understanding of placental development, but it also highlights the need for further exploration. The challenges encountered and knowledge gained provide valuable insights into the complexities of placental biology and the potential of computational methods in this field. Therefore, this research adds to the academic discourse and establishes a foundation for future innovations in developmental biology and computational histopathology.

## 8.6 Development of Unsupervised Approach for Detecting Variation in Cell Morphology: Discussion and Conclusion

Chapter 7 of my thesis delved into the complex task of exploring variations in placental cellular morphology through the use of unsupervised and self-supervised learning techniques. This was a crucial step in achieving my ultimate objective of utilising these advanced learning techniques to uncover new insights into placental histology. This chapter connects back to the overarching goal of my research by bridging the gap between sophisticated computational methods and the nuanced world of placental cellular structures.

In my investigation, I utilised unsupervised and self-supervised learning methods to analyse the complexities of placental cell morphology. This approach allowed for a more nuanced understanding of placental histology beyond traditional analysis methods.

My methodology involved the use of dimensionality reduction techniques such as PCA, t-SNE, and UMAP to simplify the complex high-dimensional data and reveal underlying patterns in placental histology. Clustering techniques such as K-means, HDBSCAN, Spectral Clustering, and GMM were used to identify distinct cell groupings which helped in understanding the cellular diversity within the placenta.

To overcome the challenges of limited labelled datasets in traditional supervised learning, I employed self-supervised learning for cell morphology differentiation, utilising contrastive learning within self-supervised learning paradigms.

My research has made a significant contribution to the field of placental histology. By utilising advanced computational methods, I have opened up new avenues for understanding the morphology of placental cells, which can lead to further studies in this field. My findings can serve as a foundation for future research and can potentially pave the way for novel diagnostic and therapeutic approaches to improve placental and fetal health.

I will now review and discuss the investigations in the research chapter and how each contributed to the overall research objective and advanced the field.

### 8.6.1 Investigation 1: Assessment of Dimensionality Reduction Techniques for High-Dimensional Data in the Context of Placental Cell Classification

As part of this study, I evaluated the effectiveness of three different dimensionality reduction techniques, namely PCA, t-SNE, and UMAP, in classifying placental cells. This analysis was crucial in distinguishing between the variations within individual cell classes and across the entire dataset.

PCA's linear approach was found to be less effective, often leading to confusion between clusters due to its inability to capture the complexity of the data. On the other hand, t-SNE, which was proficient in handling complexity, was computationally intensive. UMAP, however, emerged as the most suitable technique. This finding highlights the importance of choosing the right dimensionality reduction technique based on the intrinsic properties of the dataset.

Theoretically, this investigation contributes to the field by validating the effectiveness of non-linear dimensionality reduction techniques in complex biological data analysis.

UMAP's effectiveness suggests its potential applicability in other areas of biological data analysis where similar complexities are present. Future research should focus on optimising these techniques for different biological datasets, possibly exploring hybrid approaches that combine the strengths of various dimensionality reduction methods.

### 8.6.2 Investigation 2: Comparative Evaluation of Unsupervised Clustering Algorithms in the Analysis of Placental Cell Data

The second investigation evaluated the effectiveness of unsupervised clustering algorithms (k-means, spectral clustering, HDBSCAN, and GMM) in identifying underlying groupings within placental cell data. Critical Analysis and Theoretical Implications: Spectral clustering performed better in identifying intricate structures, reflecting its efficacy in handling the high-dimensional data typical in biological studies.

However, heuristic methods like the Elbow and Silhouette scores needed to be more conclusive, highlighting the limitations of current heuristic approaches in complex datasets. Theoretically, this reinforces the need to develop more sophisticated heuristic methods for high-dimensional biological data.

Future research should explore the development of dynamic clustering algorithms that can adapt to various characteristics of biological datasets, enhancing the precision and applicability of clustering methods in computational biology.

### 8.6.3 Investigation 3: Quantitative Assessment of Intraclass Variance in Placental Cells From a Supervised Feature Extractor

This investigation explored the intricacies of intraclass variance among placental cells through the use of a supervised model for feature extraction. The challenge of distinguishing intraclass variance using UMAP and the limitations of K-means clustering in providing clear intra-class distinctions highlight a significant shortcoming in current supervised learning models. These models tend to focus on interclass variance at the expense of more subtle intraclass differences. This suggests a need for more nuanced models to capture these subtle variances, which are crucial in understanding complex biological structures.

These findings indicate the necessity for more advanced models in biological research, especially for tasks requiring fine-grained analysis within classes. Future research should consider employing more sophisticated models that can handle high-dimensional data complexity. Additionally, exploring alternative forms of feature extraction, such as transformers, and reevaluating clustering techniques to suit biological data's complexity better could be valuable.

### 8.6.4 Investigation 4: Quantitative Assessment of Intraclass Variance in Placental Cells From a Self-Supervised Feature Extractor

In order to extract features, I shifted my focus towards self-supervised learning and utilized the SimCLR framework to explore intraclass variance. While SimCLR was

effective in distinguishing between different classes of cells (interclass variance), it struggled to capture the subtler, more nuanced variations within the same class of cells (intraclass variance). This challenge underscores the limitations of the current self-supervised learning approaches when applied to complex biological datasets, particularly in distinguishing finer intraclass details that are critical for fully understanding cellular heterogeneity.

From a practical standpoint, this underscores the potential yet unfulfilled promise of self-supervised learning in biological data analysis. Future research should focus on fine-tuning self-supervised models for biological data, exploring different architectures and training strategies that might be more effective in capturing the nuances of intraclass heterogeneity.

### 8.6.5 Investigation 5: Qualitative Investigation of Intraclass Variance

In the fifth investigation of this chapter, I conducted a qualitative analysis to explore the intraclass variance among placental cells. This investigation aimed to deepen our understanding of cellular morphology in the placenta, focusing on identifying subtle variances within cell clusters.

The results of this investigation suggest that most cell clusters exhibit a general homogeneity, with little evidence of significant intraclass variance that models and visually detected could segment. However, I did observe subtle nuances in the morphology of Trophoblastic Giant cells, such as variations in cell overlap, size, and elongation. These findings suggest the presence of intraclass variance, albeit subtle. The visual representation of these morphological differences played a crucial role in this discovery, highlighting the value of qualitative analysis in cellular analysis.

The subtle nature of the observed intraclass variances makes it challenging to quantify them. However, it suggests that there is a potential for more detailed and nuanced approaches in cellular classification, which could have significant implications for medical interventions, particularly in the field of developmental biology.

Qualitative methods have their limitations as they can introduce subjective biases and are difficult to apply in clinical settings. Therefore, future research should focus on integrating these findings with more quantitative methodologies.

Employing additional self-supervised learning techniques could further dissect the qualitative features identified, potentially revealing deeper intraclass variances. Moreover, incorporating advanced clustering techniques could complement the qualitative assessments, leading to a more robust and objective classification system.

## Chapter 7 Overall

In Chapter 7 of my thesis, I embarked on a comprehensive exploration of placental histology through advanced learning techniques. This investigation was pivotal in fulfilling my final research aim: to utilize unsupervised and self-supervised learning methods to unearth new insights into the complex world of placental cellular morphology.

To achieve my research goal of uncovering subtle patterns within the placental histology data, I utilized dimensionality reduction methods like PCA, t-SNE, and UMAP, along with clustering algorithms such as k-means, spectral clustering, HDBSCAN, and GMM. These techniques were essential in understanding the complexities of dissecting placental cell morphology and offered a more nuanced understanding than traditional methods. The choice of UMAP, in particular, proved to be effective in managing the intricacies of high-dimensional data.

Self-supervised learning, a subset of unsupervised learning techniques, has emerged as a transformative approach in the field of computational biology, particularly for tasks like cellular morphology analysis where labeled data is scarce or expensive to obtain. In my research, I leveraged the power of self-supervised learning to address the challenges posed by the lack of labeled datasets in traditional supervised learning.

One of the key self-supervised techniques I employed was contrastive learning within self-supervised learning paradigms. This method involves training models to differentiate between similar (positive) and dissimilar (negative) pairs of data points.

By doing so, the model learns to recognize intricate patterns and variances within placental cell data, which might be overlooked by traditional analysis methods.

For instance, in Investigation 4, the application of the SimCLR framework, a notable self-supervised learning model, was crucial in exploring intraclass variance among placental cells. While SimCLR effectively distinguished between different classes of cells, illustrating its capability in modeling interclass variation, it faced significant challenges in capturing intraclass variance. The framework's limitations in accurately detecting finer, intraclass details were evident, highlighting the need for more sophisticated approaches in self-supervised learning to capture these subtle differences.

The utilization of self-supervised learning in my research underscores its potential in biological data analysis, especially for tasks requiring detailed analysis within classes. Going forward, it is imperative to focus on refining self-supervised models specifically for biological datasets.

## 8.7 Additional Future Work

While this thesis has advanced histological analysis through computational techniques, the practical application of these innovations in scientific research settings, particularly those involving mouse placenta, is crucial. Future efforts should focus on making the data and computational tools developed in this thesis readily accessible and useful for placenta biologists and developmental biology researchers.

To bridge the gap between computational advances and laboratory application, future work should aim to operationalize the automated phenotyping pipeline for direct use in research labs. This involves developing user-friendly software interfaces that allow researchers to input placental images and receive automated assessments swiftly. These interfaces should provide not only quantitative analysis but also qualitative visual representations of cellular distributions and structural annotations, facilitating deeper insights into placental morphology.

Additionally, the creation of detailed documentation and training materials will ensure that researchers can fully leverage these tools. Workshops or webinars

can be conducted to demonstrate the practical uses of these tools in ongoing research, helping to integrate computational methods into traditional biological studies seamlessly.

Further practical application of this research could involve integrating the machine learning models with image analysis platforms commonly used in biological research. This integration would allow for the seamless processing of placental images, supporting studies on gene expression, developmental anomalies, and phenotype-genotype correlations in mouse models.

Future research should also prioritize the validation and calibration of these computational tools within controlled experimental settings to confirm their accuracy and reproducibility. Collaborative projects with other research labs could provide diverse datasets, enabling the refinement of the models to handle a broader range of placental types and conditions.

Moreover, expanding the dataset to include a wider variety of placental pathologies and developmental stages would enhance the model's applicability and robustness. Developing protocols for incremental learning, where models update and refine their parameters based on new incoming data, would make these tools more adaptable and responsive to the evolving needs of biological research.

Lastly, considering the differences between mouse and human placental structures, future developments could explore the scalability of these models to other species, enhancing their utility in comparative placental biology studies. This would facilitate cross-species insights that could inform both basic biological research and translational studies aimed at understanding placental functions across mammals.

## 8.8 Final Word

In my thesis, I have successfully bridged the interdisciplinary realms of computer science, machine learning, placental biology, and genetics. This endeavour resulted in annotating a comprehensive placental histology dataset for computational analysis, available for future researchers. Furthermore, I developed an end-to-end phenotyping pipeline, distinguished by its ability to localize and classify all

cells within the placenta and automatically annotate tissue regions, setting a new benchmark in the field.

I demonstrated that this phenotyping pipeline has practical applications by replicating and expanding the existing literature, highlighting its significance in enhancing our understanding. Moreover, the pipeline proved effective in predicting fetal outcomes from the captured placental phenotypes, showcasing the potential of an integrated approach to computational and biological sciences.

This thesis has laid the groundwork for future research, including applying advanced techniques like self-supervised learning for placental cell morphology analysis. While these initial efforts have yielded preliminary results, they offer a promising direction for further research.

However, the research journey is far from over. There is immense scope for enhancement, from refining nuclear and cell classification models to capture a more comprehensive array of cells, tissues, and developmental stages to developing models for identifying additional phenotypes such as blood clots, fibrotic, and necrotic tissues.

I envision a user-friendly future for lab researchers. Integrating an intuitive interface into the phenotyping method could revolutionize lab practices. This multimodal interface could analyse images and integrate results with current literature trends through generative AI methods, potentially ushering in a new era of digital pathology.

Moreover, advancing tissue segmentation techniques promise even greater accuracy in phenotyping. Our methods could extend beyond predicting outcomes to forecasting specific fetal organ effects based on placental phenomena, delving deeper into fetal-placental interactions.

Refining cell morphology methods using unsupervised and self-supervised learning involves creating extensive training datasets and innovative architectural and training methodologies across various tissues and species.

In conclusion, this thesis represents a significant milestone at the intersection of computational and biological sciences. Its achievements pave the way for future

research, inviting scholars to expand upon this foundation. I extend my heartfelt thanks to all who have engaged with my work and look forward to the continuing evolution of this fascinating field.

# Appendices

# Bibliography


[1] Michael Ferlaino et al. "Towards Deep Cellular Phenotyping in Placental Histology". In: *arXiv:1804.03270 [cs]* (Apr. 9, 2018). arXiv: 1804.03270. URL: http://arxiv.org/abs/1804.03270 (visited on 02/11/2019).

[2] Graham J. Burton and Abigail L. Fowden. "The placenta: a multifaceted, transient organ". In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 370.1663 (Mar. 5, 2015), p. 20140066. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4305167/ (visited on 08/08/2022).

[3] Foster De Witt. "An Historical Study on Theories of the Placenta to 1900". In: *Journal of the History of Medicine and Allied Sciences* XIV.7 (1959), pp. 360–374. URL: https://academic.oup.com/jhmas/article-lookup/doi/10.1093/jhmas/XIV.7.360 (visited on 12/13/2021).

[4] W. W. Buchanan, W. F. Kean, and D. G. Palmer. "The Contribution of William Hunter (1718-1783) to the Study of Bone and Joint Disease". In: *Clinical Rheumatology* 6.4 (1987), pp. 489–503.

[5] H. W. Mossmann. "Comparative morphogenesis of the fetal membranes and accessory uterine structures". In: *Contributions to Embryology Carnegie Institution Publ no* 479 (Sept. 4, 1937), pp. 26129–246. URL: https://eurekamag.com/research/022/240/022240727.php (visited on 12/13/2021).

[6] P. Kaufmann, T. M. Mayhew, and D. S. Charnock-Jones. "Aspects of Human Fetoplacental Vasculogenesis and Angiogenesis. II. Changes during Normal Pregnancy". In: *Placenta* 25.2 (2004), pp. 114–126.

[7] Graham J. Burton, Abigail L. Fowden, and Kent L. Thornburg. "Placental Origins of Chronic Disease". In: *Physiological Reviews* 96.4 (Oct. 2016), pp. 1509–1565. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5504455/ (visited on 03/21/2020).

[8] Emily Burns. "More Than Clinical Waste? Placenta Rituals Among Australian Home-Birthing Women". In: *The Journal of Perinatal Education* 23.1 (2014), pp. 41–49. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3894590/ (visited on 08/08/2022).

[9] R Michael Roberts, Jonathan A Green, and Laura C Schulz. "The Evolution of the Placenta". In: *Reproduction (Cambridge, England)* 152.5 (Nov. 2016), R179–R189. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5033709/ (visited on 12/14/2021).

[10] Derek E. Wildman et al. “Evolution of the mammalian placenta revealed by phylogenetic analysis”. In: *Proceedings of the National Academy of Sciences of the United States of America* 103.9 (Feb. 28, 2006), pp. 3203–3208. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1413940/ (visited on 08/08/2022).

[11] Stanley J. Robboy and Rana S. Hoda. “Pathology of the Human Placenta. 4th Edition.” In: *International Journal of Gynecological Pathology* 19.4 (Oct. 2000), p. 401. URL: https://journals.lww.com/intjgynpathology/fulltext/2000/10000/pathology_of_the_human_placenta__4th_edition_.20.aspx (visited on 12/14/2021).

[12] null Enders and null Blankenship. “Comparative placental structure”. In: *Advanced Drug Delivery Reviews* 38.1 (June 14, 1999), pp. 3–15.

[13] Anthony M. Carter. “Evolution of Placental Function in Mammals: The Molecular Basis of Gas and Nutrient Transfer, Hormone Secretion, and Immune Responses”. In: *Physiological Reviews* 92.4 (Oct. 1, 2012). Publisher: American Physiological Society, pp. 1543–1576. URL: https://journals.physiology.org/doi/full/10.1152/physrev.00040.2011 (visited on 05/18/2020).

[14] Peter Wooding and Graham Burton. *Comparative Placentation: Structures, Functions and Evolution*. Berlin Heidelberg: Springer-Verlag, 2008. URL: https://www.springer.com/gb/book/9783540787969 (visited on 02/12/2019).

[15] Carolyn M. Salafia, Adrian K. Charles, and Elizabeth M. Maas. “Placenta and fetal growth restriction”. In: *Clinical Obstetrics and Gynecology* 49.2 (June 2006), pp. 236–256.

[16] Margherita Turco and Ashley Moffett. “Development of the human placenta”. In: *Development* 146 (Nov. 15, 2019), dev163428.

[17] Graham J. Burton and Eric Jauniaux. “Development of the Human Placenta and Fetal Heart: Synergic or Independent?” In: *Frontiers in Physiology* 9 (2018). URL: https://www.frontiersin.org/articles/10.3389/fphys.2018.00373 (visited on 08/08/2022).

[18] Emin Maltepe and Susan J. Fisher. “Placenta: The Forgotten Organ”. In: *Annual Review of Cell and Developmental Biology* 31.1 (Nov. 13, 2015), pp. 523–552. URL: https://www.annualreviews.org/doi/10.1146/annurev-cellbio-100814-125620 (visited on 02/11/2019).

[19] Nima Soleymanlou et al. “Molecular Evidence of Placental Hypoxia in Preeclampsia”. In: *The Journal of Clinical Endocrinology and Metabolism* 90.7 (2005), pp. 4299–4308.

[20] Thomas Jansson and Theresa L. Powell. “Role of Placental Nutrient Sensing in Developmental Programming”. In: *Clinical Obstetrics and Gynecology* 56.3 (2013), pp. 591–601.

[21] Gil Mor et al. “Inflammation and Pregnancy: The Role of the Immune System at the Implantation Site”. In: *Annals of the New York Academy of Sciences* 1221 (2011), pp. 80–87.

[22] P. M. Ellery et al. “Evidence for Transcriptional Activity in the Syncytiotrophoblast of the Human Placenta”. In: *Placenta* 30.4 (Apr. 1, 2009), pp. 329–334. URL: https://www.sciencedirect.com/science/article/pii/S0143400409000137 (visited on 12/31/2021).

[23] John D. Aplin. “Developmental Cell Biology of Human Villous Trophoblast: Current Research Problems”. In: *The International Journal of Developmental Biology* 54.2 (2010), pp. 323–329.

[24] Martin Knöfler et al. “Human Placenta and Trophoblast Development: Key Molecular Mechanisms and Model Systems”. In: *Cellular and Molecular Life Sciences* 76.18 (2019), pp. 3479–3496.

[25] Ivo Brosens et al. “The "Great Obstetrical Syndromes" are associated with disorders of deep placentation”. In: *American Journal of Obstetrics and Gynecology* 204.3 (Mar. 2011), pp. 193–201.

[26] Leticia Reyes and Thaddeus G. Golos. “Hofbauer Cells: Their Role in Healthy and Complicated Pregnancy”. In: *Frontiers in Immunology* 9 (2018). URL: https://www.frontiersin.org/articles/10.3389/fimmu.2018.02628 (visited on 09/01/2022).

[27] Yawei Liu et al. “Single-cell RNA-seq reveals the diversity of trophoblast subtypes and patterns of differentiation in the human placenta”. In: *Cell Research* 28.8 (Aug. 2018). Bandiera_abtest: a Cc_license_type: cc_by Cg_type: Nature Research Journals Number: 8 Primary_atype: Research Publisher: Nature Publishing Group Subject_term: Bioinformatics;Developmental biology;Transcriptomics Subject_term_id: bioinformatics;developmental-biology;transcriptomics, pp. 819–832. URL: https://www.nature.com/articles/s41422-018-0066-y (visited on 01/02/2022).

[28] Imran N. Mir, Rachel Leon, and Lina F. Chalak. “Placental origins of neonatal diseases: toward a precision medicine approach”. In: *Pediatric Research* 89.2 (Jan. 2021), pp. 377–383. URL: http://www.nature.com/articles/s41390-020-01293-6 (visited on 01/03/2022).

[29] Susanne Lager and Theresa L. Powell. “Regulation of nutrient transport across the placenta”. In: *Journal of Pregnancy* 2012 (2012), p. 179827.

[30] Ahmet A. Baschat and Kurt Hecher. “Fetal growth restriction due to placental disease”. In: *Seminars in Perinatology* 28.1 (Feb. 2004), pp. 67–80.

[31] Atul Malhotra et al. “Neonatal Morbidities of Fetal Growth Restriction: Pathophysiology and Impact”. In: *Frontiers in Endocrinology* 10 (2019), p. 55.

[32] Johan G. Eriksson et al. “Prenatal Growth and CKD in Older Adults: Longitudinal Findings From the Helsinki Birth Cohort Study, 1924-1944”. In: *American Journal of Kidney Diseases: The Official Journal of the National Kidney Foundation* 71.1 (Jan. 2018), pp. 20–26.

[33] Elizabeth A. Phipps et al. “Pre-eclampsia: pathogenesis, novel diagnostics and therapies”. In: *Nature Reviews Nephrology* 15.5 (May 2019), pp. 275–289. URL: http://www.nature.com/articles/s41581-019-0119-6 (visited on 01/04/2022).

[34] Melina Claussnitzer et al. “A brief history of human disease genetics”. In: *Nature* 577.7789 (Jan. 2020). Bandiera_abtest: a Cg_type: Nature Research Journals Number: 7789 Primary_atype: Reviews Publisher: Nature Publishing Group Subject_term: Functional genomics;Medical genetics Subject_term_id: functional-genomics;medical-genetics, pp. 179–189. URL: https://www.nature.com/articles/s41586-019-1879-7 (visited on 01/04/2022).

[35] Janet Rossant and James C. Cross. “Placental development: Lessons from mouse mutants”. In: *Nature Reviews Genetics* 2.7 (July 2001). Number: 7 Publisher: Nature Publishing Group, pp. 538–548. URL: https://www.nature.com/articles/35080570 (visited on 03/21/2020).

[36] Tsegaselassie Workalemahu et al. “Genetic variations and risk of placental abruption: A genome-wide association study and meta-analysis of genome-wide association studies”. In: *Placenta* 66 (Apr. 1, 2018).

[37] David M. Carty, Christian Delles, and Anna F. Dominiczak. “Preeclampsia and future maternal health”. In: *Journal of Hypertension* 28.7 (July 2010), pp. 1349–1355.

[38] Michael J. Seckl, Neil J. Sebire, and Ross S. Berkowitz. “Gestational trophoblastic disease”. In: *Lancet (London, England)* 376.9742 (Aug. 28, 2010), pp. 717–729.

[39] Raj Rai and Lesley Regan. “Recurrent miscarriage”. In: *Lancet (London, England)* 368.9535 (Aug. 12, 2006), pp. 601–611.

[40] Chong Jai Kim et al. “Acute Chorioamnionitis and Funisitis: Definition, Pathologic Features, and Clinical Significance”. In: *American journal of obstetrics and gynecology* 213.4 (Oct. 2015), S29–S52. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4774647/ (visited on 10/01/2023).

[41] Liesbeth Lewi et al. “The outcome of monochorionic diamniotic twin gestations in the era of invasive fetal therapy: a prospective cohort study”. In: *American Journal of Obstetrics and Gynecology* 199.5 (Nov. 2008), 514.e1–8.

[42] Eric Jauniaux, Sally Collins, and Graham J. Burton. “Placenta accreta spectrum: pathophysiology and evidence-based anatomy for prenatal ultrasound imaging”. In: *American Journal of Obstetrics and Gynecology* 218.1 (Jan. 2018), pp. 75–87.

[43] Mihaela Pavličev et al. “Single-cell transcriptomics of the human placenta: inferring the cell communication network of the maternal-fetal interface”. In: *Genome Research* 27.3 (Mar. 2017), pp. 349–361. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5340963/ (visited on 04/07/2019).

[44] S. Blair Hedges. “The origin and evolution of model organisms”. In: *Nature Reviews Genetics* 3.11 (Nov. 2002). Number: 11 Publisher: Nature Publishing Group, pp. 838–849. URL: https://www.nature.com/articles/nrg929 (visited on 09/01/2022).

[45] Gianni Monaco et al. "A comparison of human and mouse gene co-expression networks reveals conservation and divergence at the tissue, pathway and disease levels". In: *BMC Evolutionary Biology* 15 (Nov. 20, 2015), p. 259. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4654840/ (visited on 09/01/2022).

[46] Jennifer A. Doudna and Emmanuelle Charpentier. "The new frontier of genome engineering with CRISPR-Cas9". In: *Science* 346.6213 (2014), p. 1258096.

[47] Elizabeth C. Bryda. "The Mighty Mouse: The Impact of Rodents on Advances in Biomedical Research". In: *Missouri Medicine* 110.3 (2013), pp. 207–211. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3987984/ (visited on 09/01/2022).

[48] Anne Drapkin Lyerly and Ruth R. Faden. "Mothers Matter: Ethics and Research during Pregnancy". In: *AMA Journal of Ethics* 15.9 (Sept. 1, 2013). Publisher: American Medical Association, pp. 775–778. URL: https://journalofethics.ama-assn.org/article/mothers-matter-ethics-and-research-during-pregnancy/2013-09 (visited on 09/01/2022).

[49] Anthony M. Carter. "Animal models of human pregnancy and placentation: alternatives to the mouse". In: *Reproduction* 160.6 (Dec. 1, 2020). Publisher: Bioscientifica Ltd Section: Reproduction, R129–R143. URL: https://rep.bioscientifica.com/view/journals/rep/160/6/REP-20-0354.xml (visited on 09/01/2022).

[50] Anthony Carter and A Mess. "Evolution of the Placenta in Eutherian Mammals". In: *Placenta* 28 (2007), pp. 259–62.

[51] David G. Simmons, Amanda L. Fortier, and James C. Cross. "Diverse subtypes and developmental origins of trophoblast giant cells in the mouse placenta". In: *Developmental Biology* 304.2 (Apr. 15, 2007), pp. 567–578.

[52] Anshita Rai and James C. Cross. "Development of the hemochorial maternal vascular spaces in the placenta through endothelial and vasculogenic mimicry". In: *Developmental Biology* 387.2 (Mar. 15, 2014), pp. 131–141.

[53] S.J. Tunster et al. "Increased dosage of the imprinted Ascl2 gene restrains two key endocrine lineages of the mouse Placenta". In: *Developmental Biology* 418.1 (Oct. 1, 2016), pp. 55–65. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5040514/ (visited on 09/01/2022).

[54] D. Goldman-Wohl and S. Yagel. "United we stand not dividing: the syncytiotrophoblast and cell senescence". In: *Placenta* 35.6 (June 2014), pp. 341–344.

[55] C. J. P. Jones et al. "A re-appraisal of the morphophenotype and basal lamina coverage of cytotrophoblasts in human term placenta". In: *Placenta* 29.2 (Feb. 2008), pp. 215–219.

[56] Mario R. Capecchi. "Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century". In: *Nature Reviews. Genetics* 6.6 (2005), pp. 507–512.

[57] D. J. Dumont et al. “Vascularization of the mouse embryo: a study of flk-1, tek, tie, and vascular endothelial growth factor expression during development”. In: *Developmental Dynamics: An Official Publication of the American Association of Anatomists* 203.1 (1995), pp. 80–92.

[58] Edward B. Chuong et al. “Endogenous retroviruses function as species-specific enhancer elements in the placenta”. In: *Nature Genetics* 45.3 (Mar. 2013), pp. 325–329.

[59] Richard Szeliski. *Computer Vision: Algorithms and Applications*. Cham: Springer International Publishing, 2022.

[60] *McCulloch & Pitts Publish the First Mathematical Model of a Neural Network : History of Information*. URL: https://www.historyofinformation.com/detail.php?entryid=782 (visited on 08/14/2023).

[61] F. Rosenblatt. “The perceptron: A probabilistic model for information storage and organization in the brain”. In: *Psychological Review* 65.6 (1958). Place: US Publisher: American Psychological Association, pp. 386–408.

[62] D. H. Hubel and T. N. Wiesel. “Receptive fields of single neurones in the cat’s striate cortex”. In: *The Journal of Physiology* 148.3 (Oct. 1959), pp. 574–591. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1363130/ (visited on 01/24/2022).

[63] Henry J. Kelley. “Gradient Theory of Optimal Flight Paths”. In: *ARS Journal* 30.10 (Oct. 1960). Publisher: American Institute of Aeronautics and Astronautics, pp. 947–954. URL: https://arc.aiaa.org/doi/10.2514/8.5282 (visited on 09/06/2022).

[64] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. “Learning representations by back-propagating errors”. In: *Nature* 323.6088 (Oct. 1986), pp. 533–536. URL: http://www.nature.com/articles/323533a0 (visited on 09/06/2022).

[65] Seymour A. Papert. “The Summer Vision Project”. In: (July 1, 1966). Accepted: 2004-10-04T14:40:06Z. URL: https://dspace.mit.edu/handle/1721.1/6125 (visited on 09/06/2022).

[66] Kunihiko Fukushima. “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position”. In: *Biological Cybernetics* 36.4 (Apr. 1, 1980), pp. 193–202. URL: https://doi.org/10.1007/BF00344251 (visited on 09/06/2022).

[67] Yann Lecun et al. “Backpropagation applied to handwritten zip code recognition”. In: *Neural Computation* 1.4 (1989), pp. 541–551.

[68] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. “Deep learning”. In: *Nature* 521.7553 (2015), pp. 436–444.

[69] D.G. Lowe. “Object recognition from local scale-invariant features”. In: *Proceedings of the Seventh IEEE International Conference on Computer Vision*. Proceedings of the Seventh IEEE International Conference on Computer Vision. Kerkyra, Greece: IEEE, 1999, 1150–1157 vol.2. URL: http://ieeexplore.ieee.org/document/790410/ (visited on 09/06/2022).

[70] Mark Everingham et al. "The Pascal Visual Object Classes (VOC) Challenge". In: *International Journal of Computer Vision* 88.2 (June 2010), pp. 303–338. URL: http://link.springer.com/10.1007/s11263-009-0275-4 (visited on 02/19/2019).

[71] Jia Deng et al. "Imagenet: A large-scale hierarchical image database". In: *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.

[72] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems*. Vol. 25. Curran Associates, Inc., 2012. URL: https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html (visited on 09/06/2022).

[73] Karen Simonyan and Andrew Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. Apr. 10, 2015. arXiv: 1409.1556[cs]. URL: http://arxiv.org/abs/1409.1556 (visited on 09/06/2022).

[74] Christian Szegedy et al. "Rethinking the Inception Architecture for Computer Vision". In: *arXiv:1512.00567 [cs]* (Dec. 1, 2015). arXiv: 1512.00567. URL: http://arxiv.org/abs/1512.00567 (visited on 02/13/2019).

[75] Kaiming He et al. "Deep Residual Learning for Image Recognition". In: *arXiv:1512.03385 [cs]* (Dec. 10, 2015). arXiv: 1512.03385. URL: http://arxiv.org/abs/1512.03385 (visited on 02/22/2019).

[76] Christian Szegedy et al. "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning". In: *arXiv:1602.07261 [cs]* (Feb. 23, 2016). arXiv: 1602.07261. URL: http://arxiv.org/abs/1602.07261 (visited on 02/11/2019).

[77] Ross Girshick et al. *Rich feature hierarchies for accurate object detection and semantic segmentation*. Oct. 22, 2014. arXiv: 1311.2524[cs]. URL: http://arxiv.org/abs/1311.2524 (visited on 09/07/2022).

[78] Ross Girshick. *Fast R-CNN*. Sept. 27, 2015. arXiv: 1504.08083[cs]. URL: http://arxiv.org/abs/1504.08083 (visited on 09/07/2022).

[79] Shaoqing Ren et al. *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*. Jan. 6, 2016. arXiv: 1506.01497[cs]. URL: http://arxiv.org/abs/1506.01497 (visited on 09/07/2022).

[80] Joseph Redmon et al. *You Only Look Once: Unified, Real-Time Object Detection*. May 9, 2016. arXiv: 1506.02640[cs]. URL: http://arxiv.org/abs/1506.02640 (visited on 09/07/2022).

[81] Wei Liu et al. "SSD: Single Shot MultiBox Detector". In: *Computer Vision – ECCV 2016*. Ed. by Bastian Leibe et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2016, pp. 21–37.

[82] Tsung-Yi Lin et al. *Focal Loss for Dense Object Detection*. Feb. 7, 2018. arXiv: 1708.02002[cs]. URL: http://arxiv.org/abs/1708.02002 (visited on 08/14/2023).

[83] Dinggang Shen, Guorong Wu, and Heung-Il Suk. "Deep Learning in Medical Image Analysis". In: *Annual Review of Biomedical Engineering* 19.1 (2017), pp. 221–248. URL: https://doi.org/10.1146/annurev-bioeng-071516-044442 (visited on 02/14/2019).

[84] Debashis Ganguly et al. "Medical Imaging: A Review". In: *Security-Enriched Urban Computing and Smart Grid*. Ed. by Tai-hoon Kim, Adrian Stoica, and Ruay-Shiung Chang. Communications in Computer and Information Science. Berlin, Heidelberg: Springer, 2010, pp. 504–516.

[85] Andre Esteva et al. "Deep learning-enabled medical computer vision". In: *npj Digital Medicine* 4.1 (Jan. 8, 2021). Number: 1 Publisher: Nature Publishing Group, pp. 1–9. URL: https://www.nature.com/articles/s41746-020-00376-2 (visited on 09/07/2022).

[86] Neofytos Dimitriou, Ognjen Arandjelović, and Peter D. Caie. "Deep Learning for Whole Slide Image Analysis: An Overview". In: *Frontiers in Medicine* 6 (2019). URL: https://www.frontiersin.org/articles/10.3389/fmed.2019.00264 (visited on 09/07/2022).

[87] Monya Baker. "1,500 scientists lift the lid on reproducibility". In: *Nature* 533.7604 (May 26, 2016), pp. 452–454.

[88] Metin N. Gurcan et al. "Histopathological Image Analysis: A Review". In: *IEEE reviews in biomedical engineering* 2 (2009), pp. 147–171. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2910932/ (visited on 11/09/2022).

[89] Pooya Mobadersany et al. "Predicting cancer outcomes from histology and genomics using convolutional networks". In: *Proceedings of the National Academy of Sciences of the United States of America* 115.13 (2018), E2970–E2979.

[90] Eric J. Topol. "High-performance medicine: the convergence of human and artificial intelligence". In: *Nature Medicine* 25.1 (2019), pp. 44–56.

[91] Kaustav Bera et al. "Artificial intelligence in digital pathology - new tools for diagnosis and precision oncology". In: *Nature Reviews. Clinical Oncology* 16.11 (Nov. 2019), pp. 703–715.

[92] Jakob Nikolas Kather et al. "Multi-class texture analysis in colorectal cancer histology". In: *Scientific Reports* 6 (2016), p. 27988.

[93] Dan C. Cireşan et al. "Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks". In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*. 2013, pp. 411–418.

[94] Varun Gulshan et al. "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs". In: *JAMA* 316.22 (2016), pp. 2402–2410.

[95] Wouter Bulten et al. "Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study". In: *The Lancet. Oncology* 21.2 (Feb. 2020), pp. 233–241.

[96] Zhigang Song et al. "Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning". In: *Nature Communications* 11 (2020), p. 4294.

[97] Shadi Albarqouni et al. "AggNet: Deep Learning From Crowds for Mitosis Detection in Breast Cancer Histology Images". In: *IEEE Transactions on Medical Imaging* 35 (2016), pp. 1313–1321.

[98] Laith Alzubaidi et al. "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions". In: *Journal of Big Data* 8.1 (Mar. 31, 2021), p. 53. URL: https://doi.org/10.1186/s40537-021-00444-8 (visited on 09/08/2022).

[99] Xavier Glorot and Yoshua Bengio. "Understanding the difficulty of training deep feedforward neural networks". In: *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*. 2010, pp. 249–256. URL: https://proceedings.mlr.press/v9/glorot10a.html.

[100] Nitish Srivastava et al. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting". In: *Journal of Machine Learning Research* 15 (2014), pp. 1929–1958.

[101] *Deep Learning*. MIT Press. URL: https://mitpress.mit.edu/9780262035613/deep-learning/ (visited on 08/14/2023).

[102] James Bergstra and Yoshua Bengio. "Random Search for Hyper-Parameter Optimization". In: ().

[103] Allan Bradley et al. "The mammalian gene function resource: the International Knockout Mouse Consortium". In: *Mammalian Genome: Official Journal of the International Mammalian Genome Society* 23.9 (2012), pp. 580–586.

[104] Mary E. Dickinson et al. "High-throughput discovery of novel developmental phenotypes". In: *Nature* 537.7621 (2016), pp. 508–514.

[105] Timothy Mohun et al. "Deciphering the Mechanisms of Developmental Disorders (DMDD): a new programme for phenotyping embryonic lethal mice". In: *Disease Models & Mechanisms* 6.3 (May 2013), pp. 562–566. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3634640/ (visited on 09/21/2022).

[106] Robert Wilson, Christina McGuire, and Timothy Mohun. "Deciphering the mechanisms of developmental disorders: phenotype analysis of embryos from mutant mouse lines". In: *Nucleic Acids Research* 44 (Database issue Jan. 4, 2016), pp. D855–D861. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4702824/ (visited on 02/12/2019).

[107] Vicente Perez-Garcia et al. "Placentation defects are highly prevalent in embryonic lethal mouse mutants". In: *Nature* 555.7697 (Mar. 2018), pp. 463–468. URL: https://www.nature.com/articles/nature26002 (visited on 02/11/2019).

[108] Cande V. Ananth and Anthony M. Vintzileos. "Epidemiology of preterm birth and its clinical subtypes". In: *The Journal of Maternal-Fetal & Neonatal Medicine* 19.12 (2006), pp. 773–782.

[109] Metin N. Gurcan et al. “Histopathological Image Analysis: A Review”. In: *IEEE reviews in biomedical engineering* 2 (2009), pp. 147–171. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2910932/ (visited on 09/07/2022).

[110] Navid Farahani, Anil V Parwani, and Liron Pantanowitz. “Whole slide imaging in pathology: advantages, limitations, and emerging perspectives”. In: *Pathology and Laboratory Medicine International* 7 (2015), pp. 23–33.

[111] Humera Tariq et al. “Otsu’s Segmentation: Review, Visualization and Analysis in Context of Axial Brain MR Slices”. In: *Journal of Theoretical and Applied Information Technology* 3095 (2017), pp. 6042–6055.

[112] Claire McQuin et al. “CellProfiler 3.0: Next-generation Image Processing for Biology”. In: *PLoS Biology* 16.7 (2018), e2005970.

[113] Michael Ferlaino and Craig A Glastonbury. “Towards Deep Cellular Phenotyping in Placental Histology”. In: *ORA - Oxford University Research Archive* (). URL: https://ora.ox.ac.uk/objects/uuid:293bd98d-218b-4d2b-aeb3-caf37dd41518.

[114] Tsung-Yi Lin et al. “Focal Loss for Dense Object Detection”. In: (Aug. 7, 2017). URL: https://arxiv.org/abs/1708.02002v2 (visited on 02/13/2019).

[115] Martin Skrodzki. *The k-d Tree Data Structure and a Proof for Neighborhood Computation in Expected Logarithmic Time*. 2019. URL: http://arxiv.org/abs/1903.04936.

[116] Sinno Jialin Pan and Qiang Yang. “A Survey on Transfer Learning”. In: *IEEE Transactions on Knowledge and Data Engineering* 22.10 (2010), pp. 1345–1359.

[117] C. H. Waddington. “The Epigenotype”. In: *International Journal of Epidemiology* 41.1 (2012), pp. 10–13.

[118] Graham J. Burton and Eric Jauniaux. “Pathophysiology of placental-derived fetal growth restriction”. In: *American Journal of Obstetrics and Gynecology* 218.2 (2018), S745–S761.

[119] Saeed Asaeedi, Farzad Didehvar, and Ali Mohades. “α-Concave hull, a generalization of convex hull”. In: *Theoretical Computer Science* 702 (2017), pp. 48–59.

[120] H. Edelsbrunner, D. Kirkpatrick, and R. Seidel. “On the shape of a set of points in the plane”. In: *IEEE Transactions on Information Theory* 29.4 (1983), pp. 551–559.

[121] Robert Sokal and F. Rohlf. *Biometry : the principles and practice of statistics in biological research*. 2013.

[122] Norman R. Draper and Harry Smith. “The General Regression Situation”. In: *Applied Regression Analysis*. John Wiley & Sons, Ltd, 1998, pp. 135–148.

[123] Richard A. Armstrong. “When to use the Bonferroni correction”. In: *Ophthalmic & Physiological Optics: The Journal of the British College of Ophthalmic Opticians (Optometrists)* 34.5 (2014), pp. 502–508.

[124] Gilles Louppe. *Understanding Random Forests: From Theory to Practice*. 2015. URL: http://arxiv.org/abs/1407.7502.

[125] Ye Yuan, Liji Wu, and Xiangmin Zhang. "Gini-Impurity Index Analysis". In: *IEEE Transactions on Information Forensics and Security* 16 (2021), pp. 3154–3169.

[126] Breiman. "Random Forests". In: *Machine Learning* (2001).

[127] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (Aug. 13, 2016), pp. 785–794. arXiv: 1603.02754. URL: http://arxiv.org/abs/1603.02754 (visited on 11/17/2020).

[128] Charlotte Giesen et al. "Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry". In: *Nature Methods* 11 (2014), pp. 417–422.

[129] P.m. Coan et al. "Origin and characteristics of glycogen cells in the developing murine placenta". In: *Developmental Dynamics* 235 (2006), pp. 3280–3294.

[130] K. H. Cox and E. F. Rissman. "Sex differences in juvenile mouse social behavior are influenced by sex chromosomes and social context". In: *Genes, Brain and Behavior* 10 (2011), pp. 465–472.

[131] Cheryl S. Rosenfeld. "Sex-Specific Placental Responses in Fetal Development". In: *Endocrinology* 156.10 (2015), pp. 3422–3434.

[132] A. G. Smith. "Embryo-derived stem cells: of mice and men". In: *Annual Review of Cell and Developmental Biology* 17 (2001), pp. 435–462.

[133] N. V. Chawla et al. "SMOTE: Synthetic Minority Over-sampling Technique". In: *Journal of Artificial Intelligence Research* 16 (June 1, 2002), pp. 321–357. URL: https://www.jair.org/index.php/jair/article/view/10302 (visited on 01/29/2020).

[134] Andriy Marusyk, Vanessa Almendro, and Kornelia Polyak. "Intra-tumour heterogeneity: a looking glass for cancer?" In: *Nature Reviews. Cancer* 12.5 (Apr. 19, 2012), pp. 323–334.

[135] Gloria Heppner. "Tumor Heterogeneity". In: *Cancer research* 44 (July 1, 1984), pp. 2259–65.

[136] Ronald N. Germain. "The Art of the Probable: System Control in the Adaptive Immune System". In: *Science* 293.5528 (July 13, 2001). Publisher: American Association for the Advancement of Science, pp. 240–245. URL: https://www.science.org/doi/10.1126/science.1062946 (visited on 10/01/2023).

[137] Richard Apps et al. "Human leucocyte antigen (HLA) expression of primary trophoblast cells and placental cell lines, determined using single antigen beads to characterize allotype specificities of anti-HLA antibodies". In: *Immunology* 127.1 (May 2009), pp. 26–39.

[138] R. W. Redline. "Placental pathology: a systematic approach with clinical correlations". In: *Placenta* 29 Suppl A (Mar. 2008), S86–91.

[139] G. J. Burton et al. "Optimising sample collection for placental research". In: *Placenta* 35.1 (Jan. 2014), pp. 9–22.

[140] Gavin E. Murphy et al. “Correlative 3D imaging of whole mammalian cells with light and electron microscopy”. In: *Journal of Structural Biology* 176.3 (Dec. 2011), pp. 268–278.

[141] Anne E. Carpenter et al. “CellProfiler: image analysis software for identifying and quantifying cell phenotypes”. In: *Genome Biology* 7.10 (Oct. 31, 2006), R100. URL: https://doi.org/10.1186/gb-2006-7-10-r100 (visited on 10/01/2023).

[142] Laurens van der Maaten and Geoffrey Hinton. “Visualizing Data using t-SNE”. In: *Journal of Machine Learning Research* 9 (Nov 2008), pp. 2579–2605. URL: http://www.jmlr.org/papers/v9/vandermaaten08a.html (visited on 04/08/2019).

[143] Tanel Pärnamaa and Leopold Parts. “Accurate Classification of Protein Subcellular Localization from High-Throughput Microscopy Images Using Deep Learning”. In: *G3 (Bethesda, Md.)* 7.5 (2017), pp. 1385–1392.

[144] Chawin Ounkomol et al. “Label-free prediction of three-dimensional fluorescence images from transmitted-light microscopy”. In: *Nature Methods* 15.11 (Nov. 2018). Number: 11 Publisher: Nature Publishing Group, pp. 917–920. URL: https://www.nature.com/articles/s41592-018-0111-2 (visited on 10/01/2023).

[145] Eric M. Christiansen et al. “In Silico Labeling: Predicting Fluorescent Labels in Unlabeled Images”. In: *Cell* 173.3 (2018), 792–803.e19.

[146] Rui Xu and Donald C. Wunsch. “Clustering algorithms in biomedical research: a review”. In: *IEEE reviews in biomedical engineering* 3 (2010), pp. 120–154.

[147] H. Hotelling. “Analysis of a complex of statistical variables into principal components”. In: *Journal of Educational Psychology* 24.6 (1933), pp. 417–441.

[148] Leland McInnes, John Healy, and James Melville. “UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction”. In: *arXiv:1802.03426 [cs, stat]* (Sept. 17, 2020). arXiv: 1802.03426. URL: http://arxiv.org/abs/1802.03426 (visited on 10/19/2020).

[149] Anil K. Jain. “Data clustering: 50 years beyond K-means”. In: *Pattern Recognition Letters* 31.8 (2010), pp. 651–666.

[150] *Review on determining the number of Cluster in K-Means Clustering*. n.d. URL: https://www.google.com/search?q=Review+on+determining+the+number+of+Cluster+in+K-Means+Clustering.

[151] Ricardo J. G. B. Campello, Davoud Moulavi, and Joerg Sander. “Density-Based Clustering Based on Hierarchical Density Estimates”. In: *Advances in Knowledge Discovery and Data Mining*. Ed. by Jian Pei et al. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2013, pp. 160–172.

[152] Ulrike von Luxburg. *A Tutorial on Spectral Clustering*. 2007. URL: http://arxiv.org/abs/0711.0189.

[153] Peter J. Rousseeuw. “Silhouettes: A graphical aid to the interpretation and validation of cluster analysis”. In: *Journal of Computational and Applied Mathematics* 20 (Nov. 1, 1987), pp. 53–65. URL: http://www.sciencedirect.com/science/article/pii/0377042787901257 (visited on 10/19/2020).

[154] Trupti Kodinariya and Prashant Makwana. "Review on Determining of Cluster in K-means Clustering". In: *International Journal of Advance Research in Computer Science and Management Studies* 1 (Jan. 1, 2013), pp. 90–95.

[155] Nguyen Xuan Vinh, Julien Epps, and James Bailey. "Information Theoretic Measures for Clusterings Comparison: Variants, Properties, Normalization and Correction for Chance". In: *Journal of Machine Learning Research* 11.95 (2010), pp. 2837–2854. URL: http://jmlr.org/papers/v11/vinh10a.html (visited on 10/01/2023).

[156] Andrew Rosenberg and Julia Hirschberg. "V-Measure: A Conditional Entropy-Based External Cluster Evaluation Measure". In: *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*. EMNLP-CoNLL 2007. Prague, Czech Republic: Association for Computational Linguistics, June 2007, pp. 410–420. URL: https://aclanthology.org/D07-1043 (visited on 10/01/2023).

[157] T. Calinski and J. Harabasz. "A dendrite method for cluster analysis". In: *Communications in Statistics - Theory and Methods* 3.1 (1974), pp. 1–27. URL: http://www.tandfonline.com/doi/abs/10.1080/03610927408827101 (visited on 10/01/2023).

[158] David L. Davies and Donald W. Bouldin. "A Cluster Separation Measure". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-1.2 (Apr. 1979). Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 224–227. URL: https://ieeexplore.ieee.org/document/4766909 (visited on 10/01/2023).

[159] Longlong Jing and Yingli Tian. "Self-Supervised Visual Feature Learning With Deep Neural Networks: A Survey". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 43.11 (Nov. 1, 2021), pp. 4037–4058. URL: https://ieeexplore.ieee.org/document/9086055/ (visited on 10/02/2023).

[160] Aaron van den Oord, Yazhe Li, and Oriol Vinyals. *Representation Learning with Contrastive Predictive Coding*. arXiv.org. July 10, 2018. URL: https://arxiv.org/abs/1807.03748v2 (visited on 10/02/2023).

[161] Ting Chen et al. *A Simple Framework for Contrastive Learning of Visual Representations*. 2020. URL: http://arxiv.org/abs/2002.05709.

[162] Kaiming He et al. *Momentum Contrast for Unsupervised Visual Representation Learning*. 2020. URL: http://arxiv.org/abs/1911.05722.