JGR Machine Learning and Computation

RESEARCH ARTICLE
10.1029/2025JH001124

Key Points:

- We introduce an automated mineral identification and rock classification system tailored specifically for SEM images of lunar mare basalts
- A large, manually annotated SEM data set of Apollo mare basalts was compiled to provide a benchmark for future lunar sample studies
- Despite the high accuracy achieved, manual interpretation remains necessary for detailed mineral phase analysis

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

J.-I. Jung,
jiinjung@stanford.edu

Citation:

Jung, J.-I., Tikoo, S. M., Chung, J., & Nichols, C. I. O. (2026). Automated mineral identification and rock-type classification of lunar mare basalts using SEM images. *Journal of Geophysical Research: Machine Learning and Computation*, *3*, e2025JH001124. https://doi.org/10.1029/2025JH001124

Received 14 DEC 2025
Accepted 21 MAY 2026

Automated Mineral Identification and Rock-Type Classification of Lunar Mare Basalts Using SEM Images

Ji-In Jung[1], Sonia M. Tikoo[1], Jaehong Chung[1], and Claire I. O. Nichols[2]

[1]Department of Geophysics, Stanford University, Stanford, CA, USA, [2]Department of Earth Sciences, University of Oxford, Oxford, UK

Abstract We present an automated system for identifying minerals and classifying rock types in Apollo lunar mare basalts using scanning electron microscopy (SEM) imagery. Mineral segmentation is based on a U-Net architecture, supplemented by two scale-aware models designed to incorporate pixel size information. We find that a single U-Net without explicit scale input achieves performance comparable to the scale-aware models. From a total of 17,248 augmented images (i.e., 1,078 unique base images), 12,800 were used for training, and the final model achieved average pixel-wise accuracies of 0.85, 0.81, and 0.82 on the training, validation, and test sets, respectively. For rock classification, we constructed a separate data set by compiling reported modal mineral abundances from the literature to train both a rule-based classifier and a Gaussian Naive Bayes model. These classifiers were then applied to modal abundances derived from the segmented images, achieving accuracies of >∼ 0.78 in distinguishing between ilmenite, pigeonite, and olivine basalts. Our framework enables rapid, scalable, first-order mineral identification and rock classification of lunar mare basalt petrography. However, it also reveals several limitations, including challenges in identifying minor phases, the need for phase subclassification, and inaccuracies of human bias in training annotations. These limitations underscore the continued importance of expert interpretation for detailed mineralogical and petrological studies. Nevertheless, our system provides a useful baseline and benchmark with applicability to both existing Apollo collections and lunar meteorites, as well as to future returned-sample missions.

Plain Language Summary Scientists who study rocks often spend a great deal of time identifying minerals in thin sections using high-resolution microscopy, such as scanning electron microscopy (SEM). Here, we developed a computer vision–based deep learning model that automatically identifies minerals and classifies rock types in lunar volcanic rock samples using SEM images. Trained on a variety of returned Apollo samples, our model achieved good overall accuracies >80% in mineral identification and >78% in rock type classification. At the same time, our method struggles to reliably identify rare minerals and is influenced by how humans labeled the training images. As a result, expert interpretation is still necessary for detailed mineralogical and petrological analyses. Overall, our system provides a practical and scalable starting point for automated analysis of lunar rocks, enabling much faster mineral identification of existing returned samples, lunar meteorites, and future Moon mission materials.

1. Introduction

The lunar nearside maria preserve over 3 billion years of volcanic history through multiple basalt flow units (Hiesinger et al., 2011; Head, 1976; Q.-L. Li et al., 2021a; Papike et al., 1976). These large-scale, effusive basaltic plains provide valuable insights into the flux and duration of lunar volcanism, thereby enhancing our understanding of the Moon's thermal evolution. Thanks to the Apollo missions and the Chang'e-5 mission, a wealth of lunar samples from these mare regions have been returned and used for petrographic and radiometric analyses (Che et al., 2021; Snape et al., 2019; Stoffler et al., 2006; Taylor et al., 1991). The classification of these rock types and the minerals identified are comprehensively documented in resources such as the *Lunar Sample Compendium* (Meyer, 2005), the *Lunar Sourcebook* (Papike et al., 1991; Taylor et al., 1991), and *the Lunar Sample Information Catalogs* (LSIC Apollo 11, 1977; LSIC Apollo 12, 1970; LSIC Apollo 15, 1971; LSIC Apollo 17, 1973).

Here, we aim to develop an automated system for lunar mare basalt mineral identification and rock classification, combining deep-learning-based semantic segmentation mineral mapping with classical machine-learning classifiers for rock-type classification. Traditionally, mineral identification from rock samples is a complex, time-

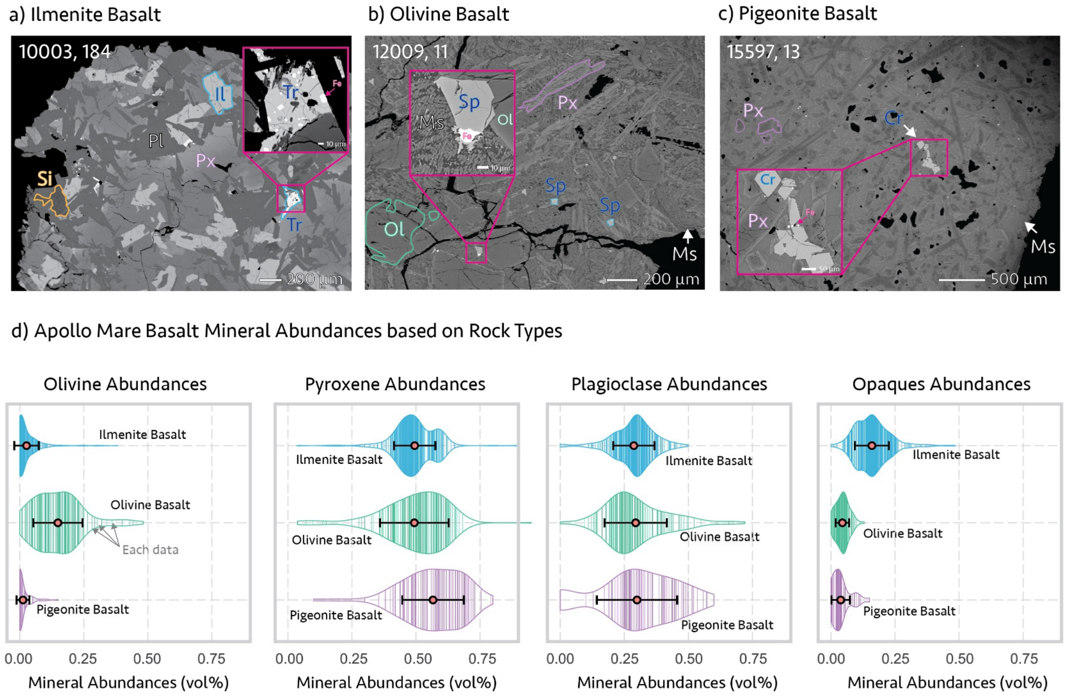

Figure 1. Scanning electron microscopy (backscattered electron) images of Apollo mare basalt varieties: (a) ilmenite basalt (10003,184); (b) olivine basalt (12009,11); (c) pigeonite basalt (15597,13). Abbreviations are as follows: Ol for olivine, Px for pyroxene, Pl for plagioclase, Il for ilmenite, Si for silica minerals, Ms for mesostasis, Sp for spinel (ulvospinel-chromite series), Cr for chromite, Tr for troilite. Panel (d) displays violin plots illustrating the mineral abundances (Olivine, Pyroxene, Plagioclase, Opaques) for each type of Apollo mare basalt, along with the mean and standard deviation. The violin plots are clipped to the data range; each line indicates a single measurement. The Apollo abundances for each rock type depicted in part (d) are sourced from the *Lunar Sample Compendium* (Meyer, 2005) and *the Lunar Sample Information Catalog*s (LSIC Apollo 11, 1977; LSIC Apollo 12, 1970; LSIC Apollo 15, 1971; LSIC Apollo 17, 1973) and are documented in Table S1 of Supporting Information S1.

consuming task requiring a deep understanding of mineralogy, petrology, and proficiency in microscopy. Recently, both shallow and deep learning techniques have been largely applied to the semantic segmentation of various microscopic rock images, including those obtained through optical microscopy (Amaral Pascarelli Ferreira et al., 2024; Baklanova & Shvets, 2014; Izadi et al., 2017; Jiang et al., 2024; Maitre et al., 2019), micro X-ray computer tomography images (Guntoro et al., 2019), and scanning electron microscopy (SEM) images (Bihani et al., 2022; Z. Chen et al., 2020). Initially, shallow machine learning techniques were widely used to identify mineral grains from optical images, using methods such as K-means clustering (Baklanova & Shvets, 2014; Guntoro et al., 2019), simple linear iterative clustering segmentation (Maitre et al., 2019), random forest (Guntoro et al., 2019; Wu et al., 2019), and cascading approaches (Izadi et al., 2017). Recent studies have shown that deep learning architectures—such as DeepLab v3 (Bihani et al., 2022; Saxena et al., 2021) and U-Net variants (Saxena et al., 2021)—are effective for segmenting and analyzing SEM images, particularly for previously unseen samples (C. Li et al., 2021b). More recently, newer architectures, including vision transformers, YOLO-based frameworks, and the Segment Anything Model (SAM), have also been explored for mineral segmentation tasks (Qiu et al., 2025; Wang et al., 2025; Zheng et al., 2024).

However, as previous studies have primarily targeted sediment or sedimentary rock for the identification of pores or grain boundaries (Bihani et al., 2022; Maitre et al., 2019; Saxena et al., 2021; Wu et al., 2019), or an ore sorting algorithm to increase economic benefits (Liu et al., 2023; McCoy & Auret, 2019), these models may face challenges when adapting to various rock types, especially in the absence of data sets specifically trained on fresh igneous rocks. Another critical point is that most studies used a fixed scale across all image data sets, whereas different image scales can reveal unique mineralogical features and textures. For example, a low-magnification image ($\geq\sim 5\ \mu$m/px) of ilmenite mare basalt sample 10003 reveals pyroxene, plagioclase, and ilmenite, which are typical minerals found in lunar mare basalts (Papike et al., 1991; Taylor et al., 1991; Figure 1a). When examined more closely at higher magnification ($\leq\sim 2$ μm/px), distinct textures of opaque minerals, including the eutectic mixture of metallic Fe and troilite (FeS), become visible, which are common features of lunar mare basalts (Cameron, 1970; Skinner, 1970; subset image of Figure 1a). These subtle details, often overlooked in the lunar mineralogy literature, are essential for understanding the composition and properties of these minerals, particularly in the context of lunar paleomagnetism and rock magnetism studies (e.g., Jung et al., 2024, 2025; Weiss & Tikoo, 2014).

In this study, we present an automated system specifically designed to identify constituent minerals and classify lithological types in Apollo lunar mare basalts using SEM backscattered electron (BSE) images. First, SEM images of representative mare basalt samples were collected and manually annotated to delineate individual mineral phases. We then employed U-Net-based deep learning models for mineral segmentation, optimized to handle varying image scales across different data set sizes. Modal mineral abundances derived from the segmentation outputs were subsequently used to classify the lithological type of each sample. Given the relatively well-documented petrological and geochemical characteristics of lunar mare basalts, particularly those returned by the Apollo missions, this approach provides a foundation for building a comprehensive basalt image data set and potentially expediting the classification of existing lunar samples as well as those returned by future missions.

2. Background

2.1. Minerals in Lunar Mare Basalts

Silicate minerals are the primary components of lunar mare basalts, making up to 90% of their composition (Figure 1). The most prevalent minerals include pyroxene, feldspar, olivine, and silica minerals. Pyroxenes $\left((Ca,Fe,Mg)_2Si_2O_6\right)$ occur as orthopyroxene, augite, and pigeonite. Feldspars include plagioclase, which forms a solid-solution series between albite and anorthite $\left((Ca,Na)(Al,Si)_4O_8\right)$ and is typically dominated by anorthitic compositions, as well as K-feldspar ($KAlSi_3O_8$). Olivine $\left((Mg,Fe)_2SiO_4\right)$ commonly exhibits fayalitic compositions, ranging from Fa_{20} to Fa_{70}. Silica minerals (SiO_2) occur as quartz, cristobalite, and tridymite (Meyer, 2005; Papike et al., 1991; Taylor et al., 1991).

In instances of extremely rapid cooling, mineral formation is suppressed, resulting in glass-dominated regions. Late-stage melt compositions (LSMC) regions (e.g., mesostasis or groundmass) are often found, which are mixtures of silicates, glass, and oxides (Figures 1b and 1c). The chemistry of these melts, which are roughly andesitic to dacitic in composition, with low alkali contents and much higher SiO_2 abundances than the bulk compositions of their host rocks, can vary with the extent of fractional crystallization and the development of individual melt pockets (Potts et al., 2016).

Metal oxides are the second most abundant group of minerals in lunar mare basalts, making up to 20% of their total composition. As the TiO_2 concentration in the original magma increases, the abundance of ilmenite, one of the most abundant metal oxides, also increases, reaching up to 15%–20% by volume in many mare basalts from the Apollo 11 and 17 landing sites (see Section 2.2 for the rock classification; Figure 1a). Another group of metal oxides, the spinel group, constitutes up to 10% of the basalts (Figures 1b and 1c). This group shows a broad spectrum of compositions and involves complex solid solutions including ulvöspinel (Fe_2TiO_4), chromite ($FeCr_2O_4$), hercynite ($FeAl_2O_4$), and spinel ($MgAl_2O_4$), which are often displayed as a Johnston compositional prism (Haggerty, 1978; Lucey et al., 2006).

Finally, within the reducing lunar environment, key minerals include native iron, primarily in the forms of kamacite and martensite (in contrast to terrestrial rocks, where iron is more commonly present as oxides such as magnetite or hematite). This native iron typically coexists with troilite (FeS) as part of a eutectic intergrowth (e.g., see the inset in Figure 1a; Cameron, 1970; Skinner, 1970), alongside spinels including chromium-ulvöspinel and titanian-chromite (as shown in the insets of Figures 1b and 1c; Cameron, 1970; Gibb et al., 1970), and within silicates, appearing interstitially, isolated within mesostasis glass, or as inclusions in silicates (Tikoo et al., 2014, 2017). Moreover, rare minerals are present, including zircon and zirconium-containing minerals, as well as phosphate minerals such as apatite and whitlockite (Meyer, 2005; Papike et al., 1991).

2.2. Mare Basalt Rock Types

The Apollo lunar mare basalt samples are primarily categorized as ilmenite basalt, olivine basalt, and pigeonite basalt according to their mineral compositions (Meyer, 2005). Mare regions were typically formed through the extrusion of partial melts from the lunar mantle through dike propagation. Between 3.8 Ga and 3.5 Ga, an early Ti-rich lava (TiO_2 of ~9–13 wt%) flooded large portions of the eastern Apollo 11 and 17 mare basalt regions (and some Apollo 12 regions; Hiesinger et al., 2000, 2011; Taylor et al., 1991). All Apollo 11 and 17 basalts are classified as ilmenite basalts in the *Lunar Sample Compendium* and the *Lunar Sample and Photo Catalog* (https://curator.jsc.nasa.gov/lunar/samplecatalog) due to their high ilmenite content (more than 10%; refer to opaques abundances in (Figure 1d)). These ilmenite basalts have been further classified into multiple schemes, including low-K, high-K, very-low-Ti, and very-high-Ti types, as well as group-based classifications (e.g., Group A, Group B1, and B2), depending on the petrologic criteria (Beaty & Albee, 1980; Papike et al., 1991; Taylor et al., 1991).

Subsequently, less Ti-rich basalts (TiO_2 of <5 wt%), discovered by the Apollo 12 and 15 missions, formed across widespread areas of the Moon between 3.5 Ga and 3.0 Ga in the central nearside regions (Hiesinger et al., 2000, 2011; Taylor et al., 1991). These basalts are classified into olivine basalts (containing 7–20% olivine) and pigeonite basalts (containing 0–2% olivine and pigeonite; refer to olivine and pyroxene abundances in (Figure 1d)). Sometimes, Apollo 15 basalts are classified as olivine-normative and quartz-normative basalts instead (Rhodes & Hubbard, 1973). Additionally, less common basalt types—including plagioclase-rich KREEP basalts or feldspathic basalts, which are interpreted to have formed during pre-mare (>3.9 Ga) magmatic episodes—were recovered at the Apollo 14, 15, and 16 landing sites (Dymek, 1986; Shih, 1977; Taylor et al., 1991).

Crater counting studies indicate that unsampled mare basalt flows resurfaced the western nearside between 3.0 and 1.0 Ga, although no returned samples were available to constrain their ages or compositions through the Apollo missions (Hiesinger et al., 2000, 2011). However, the recent return of Chang'e-5 samples provides the first direct evidence for young (~2.0 Ga) mare basalts, which exhibit low-Ti to high Ti-rich basalt compositions, rather than KREEP-rich or Al-rich basalts (Q.-L. Li et al., 2021a; Q. L. Li et al., 2021b). More recently, Chang'e-6 returned both an ancient (~4.2 Ga) high-Al KREEP basalt and a younger (~2.6 Ga) mare basalt. Despite these advances, large portions of the mare volcanism intervals (e.g., 3.0–2.6 Ga, 2.6–2.0 Ga, and 2.0–1.0 Ga) remain poorly sampled and comparatively understudied, highlighting significant opportunities for future sampling and petrologic classification.

2.3. Backscattered Electron Images

The qualitative examination of minerals in Apollo specimens generally involves both conventional optical microscopy (transmitted and reflected light) and SEM (Brown et al., 1970; Gibb et al., 1970; Meyer, 2005). Optical microscopy allows for thorough mineral identification and petrographic analysis of mineral interrelationships and textures. However, its resolution limit (typically with objectives magnification up to ~40×, corresponding to ~400–1000× total magnification) poses challenges for the qualitative assessment of micro- and nanoscale phases, particularly opaque minerals (Vernon-Parry, 2000). In contrast, electron microscopy utilizes high-energy electron beams (ranging from 2 to 1000 keV), which correspond to much shorter wavelengths (approximately 0.027–0.0009 nm; Ali et al., 2023; Vernon-Parry, 2000). As a result, SEM can achieve very high spatial resolutions. Secondary electron (SE) imaging typically has a resolution of several tens of nanometers, while BSE imaging is fundamentally limited by signal delocalization and therefore exhibits coarser resolutions of several hundreds of nanometers (Ali et al., 2023; Goldstein et al., 2018). The practical quality of electron images also depends strongly on the performance of the detector electronics, including the amplifier stability and signal-to-noise characteristics.

In SEM imaging, BSE mode is widely used in recent lunar petrography and paleomagnetism studies (Bell et al., 2020, 2023; Jung et al., 2024, 2025; McCubbin et al., 2015). BSE imaging, which detects electrons scattered from the sample surface due to interactions with the primary electron beam, provides compositional contrast based on atomic number (Ali et al., 2023). In BSE images, higher atomic number elements appear brighter, while those with lower atomic numbers appear darker, facilitating the distinction of mineral phases in Apollo mare basalt samples. For example, metallic iron (Fe) and troilite (FeS) appear among the brightest due to their high atomic numbers, followed by ilmenite and other spinels, while plagioclase appears relatively dark (Figures 1a–1c). Due to their high resolution and strong compositional contrast, BSE images are particularly wellsuited for machine learning or deep learning applications in mineral identification and rock analysis (Bihani et al., 2022; Z. Chen et al., 2020; Wu et al., 2019).

3. Mineral Identification

3.1. Labeled Mineral Segmentation Data Set

We collected SEM BSE images of various Apollo lunar mare basalt samples, including: ilmenite basalts 10003, 10020, 10069, 10071, and 12022; two olivine basalts, 12008 and 12009; and one pigeonite basalt, 15597. Prior to electron imaging, the samples were coated with approximately 15 nm of carbon using the Leica EM ACE 600 sputter coater in the Stanford Microchemical Analysis Facility (MAF). The SEM images and corresponding electron dispersive spectroscopy (EDS) data were collected for mineral identification using a JEOL JSM-IT500HR field-emission scanning electron microscope. The SEM imaging operating parameters were set at 15–20 keV.

We also incorporated additional SEM BSE images of Apollo mare basalts acquired at comparable operating conditions (15–20 keV). These include olivine and pyroxene basalt samples 15016, 15385, 15536, 15641, 15058, 15065, 15125, and 14595 collected by S. Bell and originally analyzed in Bell et al. (2020, 2023). The corresponding data set is publicly available Bell (2024). Additional images of these basalt types from samples 15555, 15475, and 15499 were shared by J. Day. Images of samples 15016, 15556, and 12017 were obtained from S. Tikoo and were originally used in Tikoo et al. (2012, 2014) and Buz et al. (2015). Finally, images of ilmenite basalts 75035, 75055, and 75075 were collected by C. Nichols and previously used in Nichols et al. (2021). The resulting data set consists of 27 independent rock samples from the Apollo missions (Table 1).

As the number of independent samples is limited, patches derived from the same sample may appear in both the training and validation sets. However, lunar mare basalts of the same lithology often share very similar mineral assemblages and textures across samples, whereas significant heterogeneity may also exist within a single sample, depending on the imaged region and spatial scale (e.g., Figure S1 in Supporting Information S1). As a result, such overlap is unlikely to substantially affect the model's generalization. Following image acquisition, SEM images were cropped into 256 × 256 pixels, which is a commonly adopted input size for modern mineral segmentation (Liu et al., 2023). Multiple SEM images at different magnifications were acquired from each sample, and patches with varying pixel scales ranging from <0.1 μm/px to ∼20 μm/px (Section S1.1 in Supporting Information S1) were extracted for mineral segmentation. We applied data augmentation through 90-degree rotations and four brightness levels. These augmentations reflect typical variations encountered during SEM sample conditions, including differences in sample mounting orientation and contrast adjustments. As our SEM data set was acquired under relatively consistent, high-resolution imaging conditions, we did not apply Gaussian blur or noise augmentation. The resulting data set includes a total of 17,248 images post-augmentation (Table 1).

For mineral classification, we categorized the minerals into 10 groups and created segmented maps accordingly: (a) Void; (b) Metallic Fe (including Fe-Ni-Co alloys); (c) Troilite (FeS); (d) Metal oxides (ilmenite, ulvöspinel, chromite); (e) Pyroxene (pigeonite, augite, orthopyroxene); (f) Plagioclase; (g) Silica minerals (quartz, cristobalite); (h) Olivine; (i) LMSC (Mesostasis or groundmass); (j) Other (including rare phases such as phosphates or zircon that do not fall into the above categories).

3.2. Mineral Segmentation Methods

We employed the U-Net architecture, a convolutional neural network (CNN) specifically designed for image segmentation tasks (Ronneberger et al., 2015; Figure 2), as a backbone to segment minerals from SEM images. The U-Net architecture integrates successive convolutional layers for image contraction, max pooling layers to capture context, symmetric upsampling layers for precise image localization, and skip connections to retain detailed features during upsampling. Within our framework, three distinct segmentation strategies were adopted: one scale-unaware model (Model 1) and two scale-aware models (Models 2 and 3; Figure 3). Using the PyTorch library, we implemented these three models and trained them on five versions of our data set containing 192, 384, 800, 1,600, 3,200, 6,400, and 12,800 samples, all evaluated using the same validation set of 2,048 samples. Each of the training, validation, and test sets maintained a 50:50 ratio of small-scale (pixel size <1.8 μm) and large-scale data (pixel size ≥1.8 μm). The data set was split into training, validation, and test sets before data augmentation to reduce any risk of data leakage.

Table 1
Apollo Thin Section Information Used for the Training and Validation Data Sets

Thin section	Rock types	N	BE (keV)	img ref
10003, 184	Ilmenite Basalt	1904	15, 20	This study
10020, 075	Ilmenite Basalt	992	15	This study
10069, 033	Ilmenite Basalt	976	20	This study
10071, 031	Ilmenite Basalt	1328	15, 20	This study
12022, 114	Ilmenite Basalt	816	20	This study
75035, 084	Ilmenite Basalt	16	–	Nichols et al. (2021)
75055, 046	Ilmenite Basalt	32	–	Nichols et al. (2021)
75075, -	Ilmenite Basalt	32	–	Nichols et al. (2021)
Total		$n = 6096$		
12008, 072	Olivine Basalt	688	20	This study
12009, 011	Olivine Basalt	1072	15, 20	This study
15016, 144	Olivine Basalt	96	15	Tikoo et al. (2014)
15016, 145	Olivine Basalt	896	15	Bell et al. (2020, 2023)
15385, 013	Olivine Basalt	880	15	Bell et al. (2020, 2023)
15536, 005	Olivine Basalt	880	15	Bell et al. (2020, 2023)
15555, -	Olivine Basalt	160	15	James Day*
15556, 132	Olivine Basalt	64	15	Tikoo et al. (2014)
15641, 003	Olivine Basalt	880	15	Bell et al. (2020, 2023)
Total		$n = 5616$		
12017, 024	Pigeonite Basalt	64	15	Buz et al. (2015)
15058, 130	Pigeonite Basalt	832	15	Bell et al. (2020, 2023)
15065, 196	Pigeonite Basalt	880	15	Bell et al. (2020, 2023)
15125, 006	Pigeonite Basalt	896	15	Bell et al. (2020, 2023)
15475, -	Pigeonite Basalt	240	15	James Day*
15495, 092	Pigeonite Basalt	864	15	Bell et al. (2020, 2023)
15499, -	Pigeonite Basalt	160	15	James Day*
15597, 013	Pigeonite Basalt	960	20	This study
15666, 008	Pigeonite Basalt	640	15	Bell et al. (2020, 2023)
Total		$n = 5536$		
Total		$n = 17248$		

Note. * = personal communication; N = number of images; BE = Beam Energy.

Model 1 (Vanilla U-Net, Resolution-Independent): Model 1 employs the conventional U-Net architecture, which is widely used and adapted in microscopic image analysis (Long, 2020; Ronneberger et al., 2015). It operates without the scale parameter, thus working independently of the pixel resolution. We modified our U-Net design to accept input images with dimensions of ($256 \times 256 \times 1$), ensuring it matches our data set.

Model 2 (Scale-Specific U-Net Architecture): This model assigns each image to a distinct U-Net architecture based on its physical scale. As a result, Model 2 incorporates multiple U-Net architectures, each tailored to a specific scale range and trained on a corresponding subset of the data set. In this study, we divided the data into two groups: a small-scale image set (pixel size <1.8 μm) and a large-scale image set (pixel size ≥ 1.8 μm), with each group comprising approximately half of the total data set.

Model 3 (U-Net with Scale as a Learnable Parameter): When image variability with scale becomes considerable, using multiple U-Net architectures (as in Model 2) may become impractical. Model 3 aims to retain a single U-Net architecture while integrating scale as an input parameter. This is achieved by adding a linear

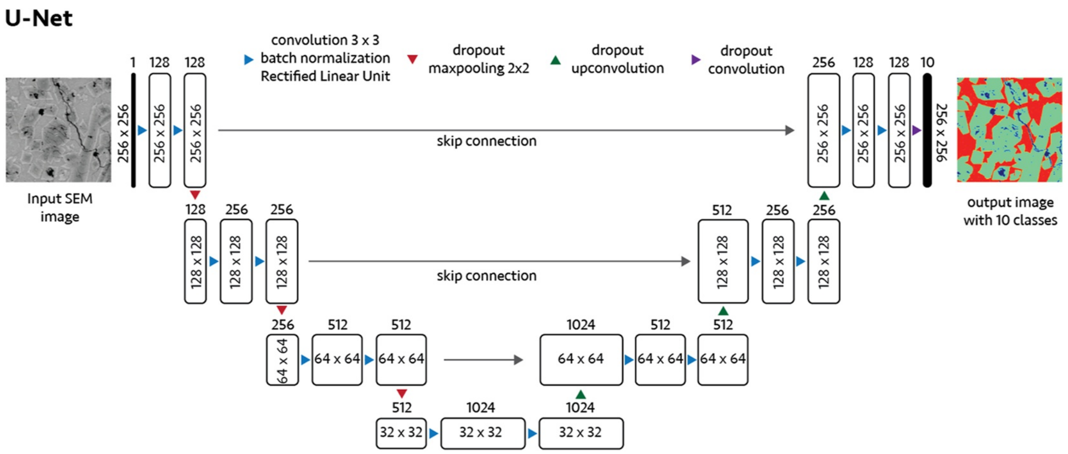

Figure 2. U-Net architecture used for semantic segmentation of SEM images into 10 mineralogical classes. The network takes a 256 × 256 input SEM image and consists of repeated 3 × 3 convolutions with batch normalization and rectified linear unit (ReLU) activation, followed by dropout and max-pooling in the encoder path. The decoder path mirrors the encoder using up-convolutions and convolutional layers with batch normalization, ReLU activation, and dropout to reconstruct the feature maps, producing pixel-wise predictions for 10 mineral classes.

parameter (weight × scale + bias) at the base of the U-Net architecture, enabling the model to adaptively learn and contextualize based on the image scale.

Loss and Evaluation Metrics: To assess model performance, we evaluated three loss configurations: cross-entropy loss (L_{CE}), a combination of cross-entropy and Dice loss (L_{CE+DL}), and a combination of cross-entropy and generalized Dice loss (L_{CE+GDL}). The cross-entropy loss (L_{CE}) is defined as:

$$L_{CE} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} y_{ic}\log\left(p_{ic}\right)$$

where N is the total number of pixels, C is the number of classes (e.g., olivine or plagioclase), y_{ic} is the true label, and p_{ic} is the predicted probability that pixel i belongs to class c. The predicted probabilities are obtained via the Softmax function, which transforms raw model outputs into a probability distribution. This function penalizes incorrect predictions heavily and was employed in the original U-Net segmentation framework (Ronneberger et al., 2015).

This loss function, however, tends to prioritize large, well-represented regions (e.g., pyroxene, plagioclase, and metal oxides) while under-weighting smaller or less abundant classes (e.g., Fe grains). To reduce this bias, we incorporated a standard Dice loss term (L_{DL}) and a generalized Dice loss term (L_{GDL}) (Sudre et al., 2017):

$$L_{CE+Dice} = L_{CE} + L_{Dice}$$

and

$$L_{CE+GDL} = L_{CE} + L_{GDL}.$$

The standard Dice loss (L_{Dice}) measures the overlap between predicted and ground-truth labels across all classes equally:

$$L_{Dice} = 1 - \frac{2\sum_{c=1}^{C}\sum_{i=1}^{N} p_{ic}\, y_{ic}}{\sum_{c=1}^{C}\sum_{i=1}^{N} p_{ic} + \sum_{c=1}^{C}\sum_{i=1}^{N} y_{ic} + \epsilon}$$

The generalized Dice loss (L_{GDL}) extends this formulation by introducing class-specific weights (w_c), thereby penalizing underrepresented classes more strongly (Sudre et al., 2017):

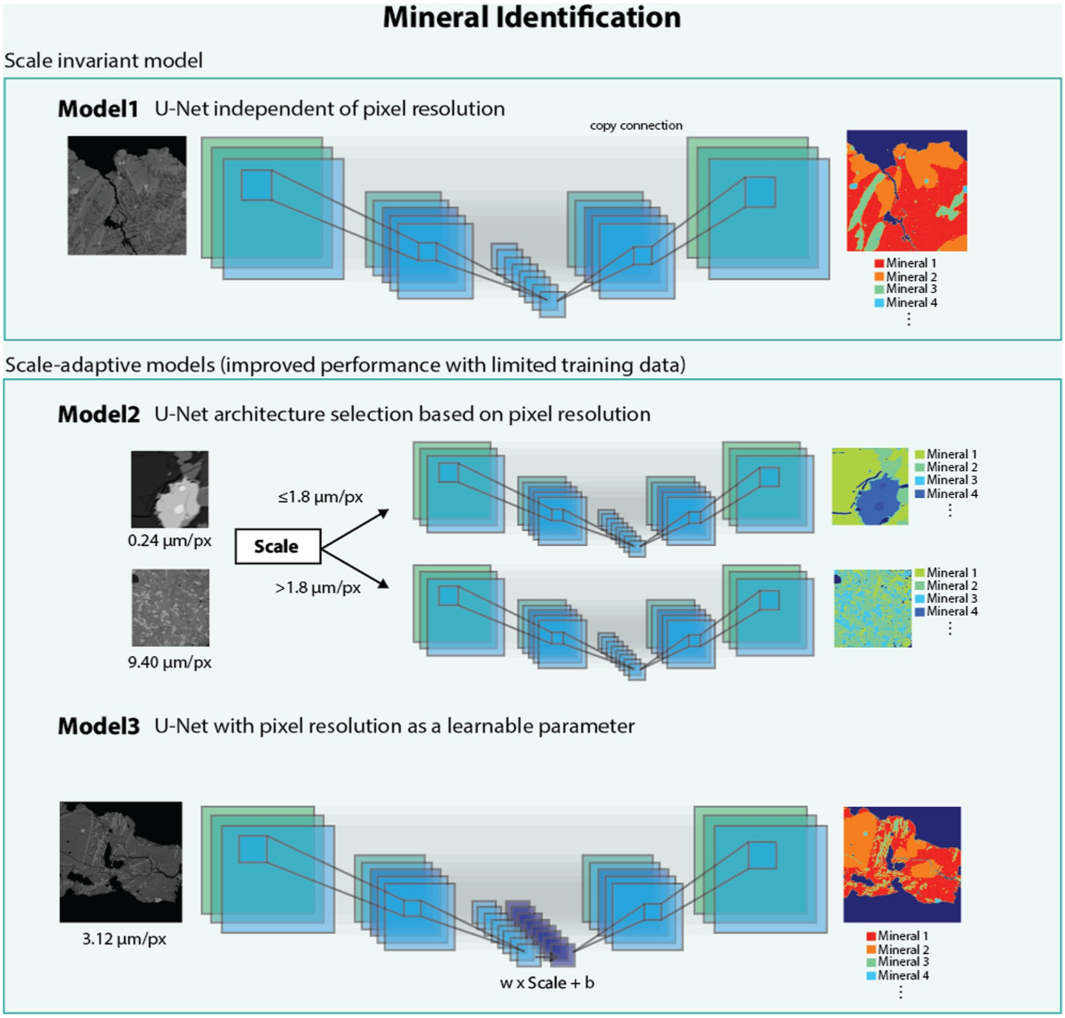

Figure 3. The adapted three U-Net methodologies implemented in this study. Model 1 employs a resolution-independent U-Net applied across all images. Model 2 uses a scale-adaptive approach, selecting different U-Net architectures based on pixel resolution thresholds. Model 3 incorporates pixel resolution as a learnable input parameter (linear modulation as w × scale + b).

$$L_{GDL} = 1 - \frac{2\sum_{c=1}^{C} w_c \sum_{i=1}^{N} p_{ic}\, y_{ic}}{\sum_{c=1}^{C} w_c \left(\sum_{i=1}^{N} p_{ic} + \sum_{i=1}^{N} y_{ic}\right) + \epsilon}$$

where

$$w_c = \frac{1}{\left(\sum_{i=1}^{N} y_{ic}\right)^2}$$

The small constant ϵ ($1e{-}5$) is added to ensure numerical stability and to avoid division by zero.

To assess the performance of the models, we used two metrics: pixel accuracy, defined as the ratio of correctly predicted pixels to the total number of pixels in the image; and intersection over union (IoU), defined as the ratio of the area of overlap to the area of union between predicted and ground-truth segmentations.

Hyperparameters: The initial training of Models 1, 2, and 3 was conducted using fixed hyperparameters: a batch size of 10, a dropout rate of 0.3, and an initial learning rate of 1e-4, with a StepLR scheduler (Goodfellow

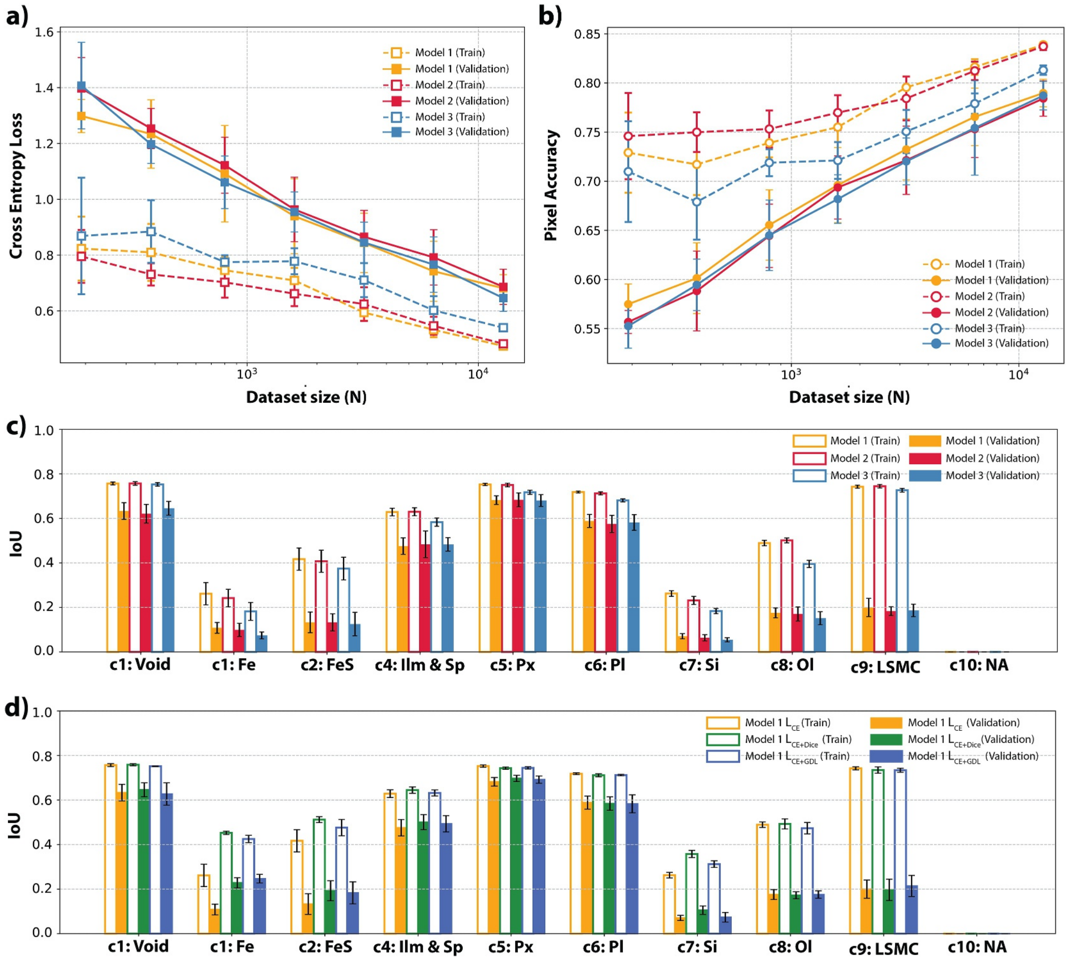

Figure 4. (a) Loss and (b) pixel accuracy as a function of training data set size (200, 400, 800, 1,600, 3,200, 6,400, and 12,800 samples) for Model 1 (orange), Model 2 (red), and Model 3 (blue), trained using cross-entropy loss (L_{CE}). (c) Per-class Intersection over Union (IoU) for the three models trained with L_{CE} using the largest training data set (12,800 samples). (d) Per-class IoU for Model 1 trained with L_{CE} (orange), $L_{CE} + L_{Dice}$ (green), and L_{CE} + generalized Dice loss (L_{GDL}; blue). All models were trained and validated using data sets containing an equal proportion (50:50) of large- and small-scale images, with the validation set fixed at 2,048 samples. For each data set size and model, the best-performing result over 100 training epochs, defined as the epoch with the highest validation accuracy, was selected. Results are shown for five different data set split seeds, and error bars indicate one standard deviation.

et al., 2016) that reduced the learning rate by a factor of 0.1 every 50 epochs. Each model was trained for 100 epochs, and the best-performing model was selected based on validation pixel accuracy. To further improve model performance, we conducted hyperparameter tuning on Model 1, with the full data set with 12800:2048:2400 = training:validation:test split, exploring variations in dropout rates and batch sizes. Dropout rates of 0.3 and 0.5 were tested to mitigate overfitting, while batch sizes of 4, 8, 16, and 32 were evaluated, and training was continued up to 200 epochs to obtain the final model.

3.3. Results

For all models, L_{CE} and pixel accuracy exhibit approximately logarithmic trends with respect to training data set size, indicating overall improved generalization (Figures 4a and 4b). We do not observe a substantial difference in overfitting behavior between smaller and larger training data sets (Section S2 and Figures S4–S13 in Supporting Information S1). Statistical significance was evaluated across different data split seeds using paired *t*-tests and Wilcoxon signed-rank tests with a threshold of $p < 0.05$, and most comparisons did not show statistically significant differences. These results suggest that incorporating pixel-scale information did not substantially enhance model performance, but performance varied across different random data splits, indicating that model performance is driven more by the training data than by the model variant. On the largest training data set (12,800 training samples; 800 unique images), the final pixel-wise accuracies were comparable across models: 84.12% (train)/80.01% (val) for Model 1%, 83.77%/79.04% for Model 2%, and 81.85%/80.00% for Model 3.

All models effectively identified the major mineral phases, including c1: Void (IoU = 0.74–0.75 for training, 0.63–0.66 for validation), c4: Metal oxides (0.60–0.63 for training, 0.44–0.46 for validation), c5: Pyroxene (0.73–0.76 for training, 0.67–0.68 for validation), and c6: Plagioclase (0.69–0.71 for training, 0.61–0.63 for validation; Figure 4c). However, c9: LSMC shows the largest train-validation discrepancy (IoU >0.70 for training; IoU <0.20 for validation), likely because groundmass and mesostasis were not consistently identifiable across all images, indicating the need for more detailed subclassification and representative data (see LSMC examples in (Figures S2c, S2e, S2g, and S2h in Supporting Information S1)). Phases that are less abundant and lack unique visual features also showed lower IoU values. For example, c3: Troilite (IoU >0.40 for training; IoU <0.20 for validation) remains challenging to segment, as it can be confused with slightly darker metal oxides. Likewise, c8: Olivine (IoU >0.40; IoU <0.20) is difficult to segment due to its diverse morphologies, ambiguous boundaries, and similar brightness to pyroxene (see olivine occurrences in (Figure S2d in Supporting Information S1)). Phases that are both less abundant and poorly generalized across samples show low IOU values. Examples include c2: Metallic Fe (IoU <0.30 for training; IoU <0.15 for validation; Figures S2c, S2f, and S2i in Supporting Information S1), c7: Silica minerals (IoU <0.30; IoU <0.15 for validation), and c10: NA (IoU <0.05). When using $L_{CE+Dice}$ or L_{CE+GDL}, Model 1 showed a slight improvement, particularly for minor phases such as metallic Fe (IoU <0.4 for training; IoU <0.2 for validation; see the confusion matrix comparison in (Figure S16 in Supporting Information S1))

The optimal batch size and dropout rate were determined to be 8 and 0.3, respectively (Section S3 in Supporting Information S1). After 200 epochs, the best-performing model (Model 1, trained on a data set of 12,800 samples using L_{CE} with a 12,800:2,048:2,400 training–validation–test split) achieved pixel-wise accuracies of 0.85, 0.81, and 0.82 for the training, validation, and test sets, respectively. Training the same model with $L_{CE+Dice}$ and L_{CE+GDL} yielded comparable performance, with accuracies of 0.85, 0.81, and 0.82, and 0.84, 0.81, and 0.81, respectively. The final model trained with L_{CE} was subsequently used for further rock-type classification and application analyses.

4. Rock Type Classification

In our subsequent rock classification framework, the mineral abundances were used as input features, and the corresponding rock type served as the output label. We employed two primary methods to infer rock type from the mineral maps: (a) a rule-based approach and (b) a Gaussian Naive Bayes (GNB) classifier (Pedregosa et al., 2011). These methods were first tested using a data set compiled from petrology studies (Section 4.1 and Table S1 in Supporting Information S1), and were subsequently applied to the mineral identification maps produced by Model 1 (Table 2). In addition to these two classifiers, we also explored other standard classification algorithms, including support vector machines (SVM), logistic regression, random forest, multilayer perceptron (MLP), and XGBoost (T. Chen & Guestrin, 2016; Pedregosa et al., 2011; Figure 5 and Section S3 in Supporting Information S1). We evaluated these common machine learning models for completeness, but our primary attention focused on these two interpretable classifiers that align with petrologically meaningful thresholds. Finally, we tested rock classification methods on both seen and unseen SEM BSE stitched images sourced from various basalt types (Section S4 in Supporting Information S1).

4.1. Modal Mineral Abundances Data Set

We compiled a data set of modal mineral abundances for Apollo mare basalt samples based on reports in the Lunar Sample Information Catalogs (LSIC Apollo 11, 1977; LSIC Apollo 12, 1970; LSIC Apollo 15, 1971; LSIC Apollo 17, 1973) and studies cited in the Lunar Sample Compendium (Meyer, 2005; Table S1 in Supporting Information S1). Mineralogical modes reported in the *Lunar Sample Information Catalogs* are based on binocular observations of hand samples and petrographic analyses of thin sections. The majority of modal abundances reported in studies cited in the *Lunar Sample Compendium* were derived from several thousand point counts, obtained either through manual optical point counting or automated microprobe-based systems on thin sections. In some cases, modal estimates were instead derived from grain-separation techniques or chemical analyses (see the method column in (Table S1 in Supporting Information S1)). However, binocular observations of hand samples are qualitative estimates made by different petrologists, and therefore naturally contain substantial variation. For thin-section analyses, the measurements were performed by different researchers on different subsamples or thin sections—each containing distinct crystal populations and spatial heterogeneity—so the reported modal mineralogies inherently exhibit user-to-user variability and n–x variation (Meyer, 2005; Neal & Taylor, 1992; Papike et al., 1991). Indeed, in several instances, multiple measurements exist for the same thin section, and we retained all available measurements in our data set to maximize the use of available modal information. As these observations are not strictly independent at the sample level, this may represent a limitation for classifier training and validation in the current data set.

To construct a tabulated data set, we first identified all Apollo samples classified as mare basalts in the Lunar Sample and Photo Catalog (https://curator.jsc.nasa.gov/lunar/samplecatalog) and adopted their reported rock-type classifications. Mare basalts in the catalog are assigned to one of the following categories: ilmenite, olivine, pigeonite, plagioclase, KREEP, feldspathic, or unclassified basalt. We then reduced the data set to include only the mineral abundances of olivine, pyroxene, plagioclase, mesostasis (or groundmass), silica, and opaques (ilmenite, spinels, troilite, metals). The final data set included six normalized mineral abundances, which were used as input features. Our data set comprises 671 measurements from various studies (including repeated measurements of the same samples) spanning all Apollo missions: Apollo 11 (99), Apollo 12 (136), Apollo 14 (14), Apollo 15 (166), and Apollo 17 (256). The mineralogical mode data set includes 397 measurements of ilmenite basalts, 147 of olivine basalts, 106 of pigeonite basalts, 12 of feldspathic basalts, 8 of KREEP basalts, and 1 of an unclassified basalt. The rock type (ilmenite, olivine, or pigeonite basalts) was used as the classification target. An 80:20 train–validation split was used to train the GNB model and other modern classifiers.

Table 2
Rock Classification Performance Results for the Rule-Based (RB) Model and Gaussian Naive Bayes (GNB) Model

Rock type	Precision	Recall	F1-score	Support
RB Accuracy [LSC data set (Test)]: 0.81				
Ilmenite Basalt	0.94	0.80	0.87	397
Olivine Basalt	0.64	0.84	0.73	147
Pigeonite Basalt	0.72	0.80	0.76	106
Macro Average	0.77	0.82	0.78	650
Weighted Average	0.84	0.81	0.82	650
GNB Accuracy [LSC data set (Train)]: 0.88				
Ilmenite Basalt	0.92	0.93	0.93	362
Olivine Basalt	0.85	0.85	0.85	131
Pigeonite Basalt	0.76	0.71	0.73	92
Macro Average	0.84	0.83	0.84	585
Weighted Average	0.88	0.88	0.88	585
GNB Accuracy [LSC data set (Test)]: 0.91				
Ilmenite Basalt	0.89	0.97	0.93	35
Olivine Basalt	1.00	0.94	0.97	16
Pigeonite Basalt	0.83	0.71	0.77	14
Macro Average	0.91	0.87	0.89	65
Weighted Average	0.91	0.91	0.91	65
RB Accuracy [SEM data set (Test)]: 0.80				
Ilmenite Basalt	0.84	0.89	0.86	2128
Olivine Basalt	0.85	0.71	0.77	2800
Pigeonite Basalt	0.76	0.82	0.79	3696
Macro Average	0.81	0.81	0.81	8624
Weighted Average	0.81	0.80	0.80	8624
GNB Accuracy [SEM data set (Test)]: 0.78				
Ilmenite Basalt	0.74	0.87	0.80	2128
Olivine Basalt	0.93	0.61	0.73	2800
Pigeonite Basalt	0.74	0.86	0.80	3696
Macro Average	0.80	0.78	0.78	8624
Weighted Average	0.80	0.78	0.78	8624

4.2. Rock Classification Methods

Rule-Based Classification: We implemented a straightforward rock classification method using predefined rules based on the quantified mineral proportions. Threshold values were established based on Apollo-era data sets (Meyer, 2005; Papike et al., 1991). Samples are classified based on modal mineral abundances as follows: if the olivine abundance is >5 vol%, they are classified as olivine basalts; if the olivine abundance is ≤5 vol% and the opaque mineral abundance is >7 vol%, they are classified as ilmenite basalts; otherwise, samples with ≤5 vol% olivine and ≤7 vol% opaque minerals are classified as pigeonite basalts.

$$\text{Class} = \begin{cases} \text{Olivine Basalt} & \text{if Olivine} > 5\% \\ \text{Ilmenite Basalt} & \text{if Opaques} > 7\% \wedge \text{Olivine} \leq 5\% \\ \text{Pigeonite Basalt} & \text{if Opaques} \leq 7\% \wedge \text{Olivine} \leq 5\% \end{cases}$$

Gaussian Naive Bayes Classification: To further refine rock classification, we applied a Gaussian Naive Bayes classifier. This method assumes that the features $\mathbf{x}_i$ (mineral abundances, where $\mathbf{x}_1$ = olivine, $\mathbf{x}_2$ = pyroxene, $\mathbf{x}_3$ = plagioclase, $\mathbf{x}_4$ = mesostasis, $\mathbf{x}_5$ = silica, $\mathbf{x}_6$ = opaques) are conditionally independent given the rock class label C_k (C_1 = Ilmenite basalt, C_2 = Olivine basalt, C_3 = Pigeonite basalt) and that each feature follows a Gaussian distribution. The probability density function for each feature $\mathbf{x}_i$ given class C_k is defined as:

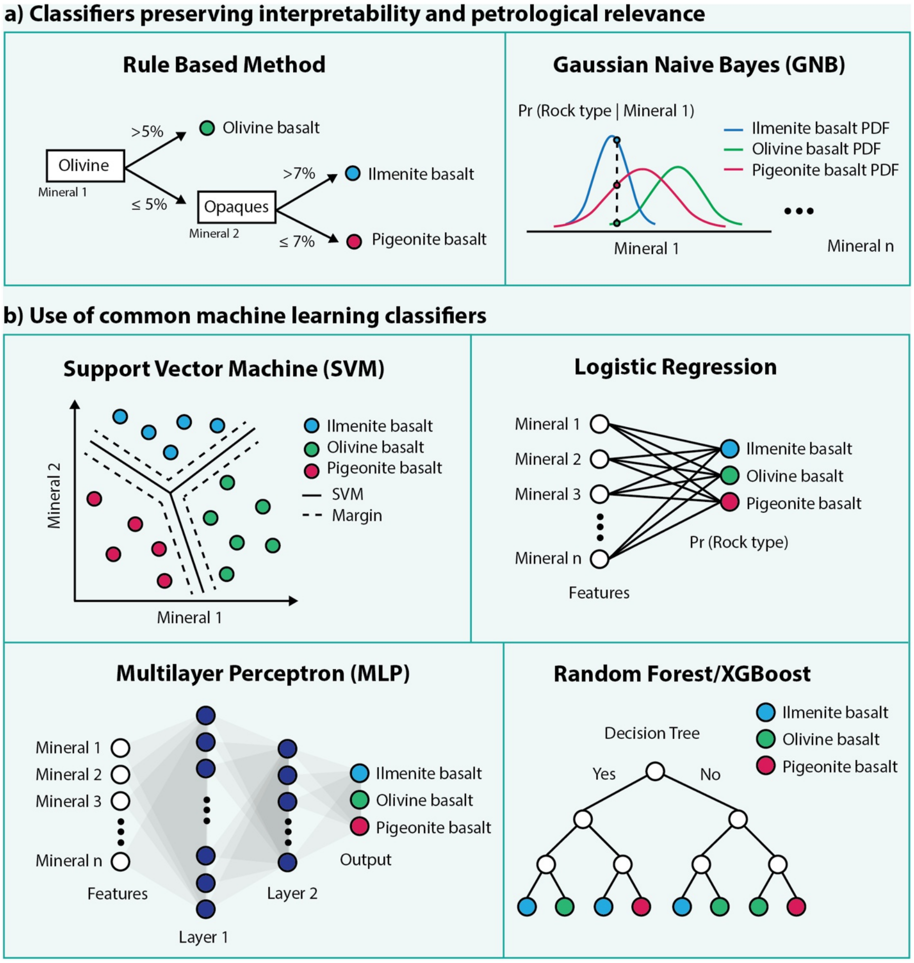

Figure 5. Overview of rock classification framework (a) Two rock classifiers were used to preserve interpretability and petrological relevance: (1) a rule-based method using empirically defined thresholds, and (2) a Gaussian Naive Bayes (GNB) classifier using probability distributions of mineral abundances. (b) Standard machine learning classifiers: support vector machines (SVM), logistic regression, multilayer perceptron (MLP), random forest, and XGBoost.

$$P(\mathbf{x}_i|C_k) = \frac{1}{\sqrt{2\pi\sigma_{k,i}^2}} \exp\left(-\frac{(x_i - \mu_{k,i})^2}{2\sigma_{k,i}^2}\right)$$

where $\mu_{k,i}$ and $\sigma_{k,i}^2$ are the mean and variance of feature i for class C_k. For each class, the likelihood of a data point is given by the product of the probabilities of the individual features:

$$P(\mathbf{x}|C_k) = \prod_{i=1}^{n} P(\mathbf{x}_i|C_k)$$

The Gaussian Naive Bayes classifier calculates the posterior probability $P(C_k|\mathbf{x})$ of a class C_k given a data point $\mathbf{x}$ using Bayes' theorem:

$$P(C_k|\mathbf{x}) = \frac{P(\mathbf{x}|C_k)P(C_k)}{P(\mathbf{x})}$$

Since $P(\mathbf{x})$ is the same for all classes, it can be simplified to:

$$P(C_k|\mathbf{x}) \propto P(\mathbf{x}|C_k)P(C_k)$$

The prior probability $P(C_k)$ is the proportion of class C_k in the training data:

$$P(C_k) = \frac{N_k}{N}$$

where N_k is the number of samples in class C_k and N is the total number of samples. Finally, the predicted class $\hat{C}$ for a new data point $\mathbf{x}$ is determined by:

$$\hat{C} = \underset{C_k}{\operatorname{argmax}} P(C_k|\mathbf{x})$$

4.3. Results

Both methods demonstrated strong performance in classifying the three primary rock types of Apollo mare basalts. For the *Lunar Sample Compendium* data set, the rule-based method and the GNB classifier achieved comparable performance, with accuracies of 81% (training) and 88% (training)/91% (test), respectively (Table 2). Other classifiers achieved accuracies ranging from 89% to 100% (training) and 89%–100% (test; Tables S2–S6 in Supporting Information S1).

We further applied these methods to mineral-segmented maps generated by Model 1. The analysis was limited to large-scale images with pixel sizes >1.8 μm (n = 8,624 images), effectively capturing the full range of mineral abundances to represent the sample. We then quantified the proportions of key minerals listed in the *Lunar Sample Compendium* by pixel counting. Results show that the rule-based method slightly outperformed the GNB classifier, achieving 80% accuracy compared to 78% (Table 2). Classification performance was further evaluated using standard metrics: precision (correctness of predictions), recall (completeness of predictions), F1-score (harmonic mean of precision and recall), support (the number of ground-truth cases for each class), macro-averaged (i.e., per-class average), and weighted-averaged (i.e., sample-proportional average) scores. The rule-based model achieved a more balanced precision and recall across all three rock types, resulting in higher macro- and weighted-average F1-scores. Both models show slightly higher recall but lower precision for ilmenite basalts, suggesting a tendency to overpredict this class. The GNB classifier demonstrated high precision for pigeonite basalts but significantly underperformed in classifying olivine basalts.

To evaluate the validity of our automated system for rock classification, we tested it on both seen and unseen SEM BSE stitched images sourced from various basalt types (Bell et al., 2020, 2023; Jung et al., 2024; Table S7 in Supporting Information S1). Each stitched image was divided into patches with pixel sizes >1.8 μm. For each patch, the abundance of mineral phases was quantified using the Model 1 U-Net architecture, and the rock type was predicted using both a rule-based approach and a GNB classifier. The final rock type for each sample was determined by majority voting across all patches. Although the number of available test stitched images was limited, our system generally agreed with the expected rock types.

5. Discussion

5.1. Sources of Uncertainty in SEM-Based Petrographic Analysis

Our automated system identifies key minerals in lunar mare basalts with >80% accuracy for mineral identification and >78% for rock classification. The remaining 20%–25% inaccuracies arise from the following main sources: First, some mineral grains occur across a wide range of textural and compositional contexts in different rock types, making generalization difficult. For example, ilmenite basalts often contain Fe–metal assemblages associated with troilite, typically forming spherical or rounded Fe grains embedded within FeS or mesostasis (Cameron, 1970; Skinner, 1970). In contrast, olivine and pyroxene basalts more commonly host spherical Fe grains dispersed in the groundmass or angular Fe grains associated with spinels (Jung et al., 2025; Tikoo et al., 2014). This limitation could be mitigated by deploying rock-type-specific, fine-tuned models—such as an ilmenite-basalt model optimized for identifying minerals in ilmenite basalts—following the initial rock-classification step.

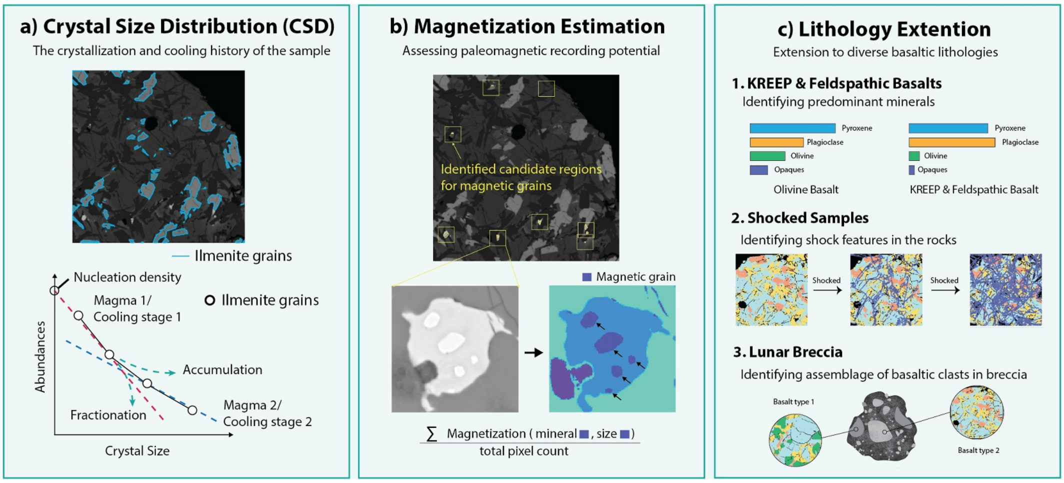

Figure 6. Scientific applications of the resulting segmented mineral maps: (a) Crystal size distribution (CSD) analysis. The relationship between crystal size (estimated from segmentation) and abundance (number density) can provide quantitative constraints on crystal growth processes, crystallization conditions, cooling histories, and magma mixing. (b) Estimation of magnetization potential. Identification of metallic Fe and calculation of its volumetric fraction can enable prediction of the expected thermoremanent magnetization intensity and facilitate the selection of samples suitable for paleomagnetic studies. (c) Extension to diverse basaltic lithologies. This may include classification of KREEP and feldspathic basalts, identification of possible shock features, and characterization of basaltic clasts in breccias.

Second, some phases currently grouped into the same mineral-phase class should not necessarily be treated as the same type. For example, we group mesostasis and groundmass together as LSMC consistent with ambiguities in prior modal compilations from Apollo-era mineralogy and petrology studies (Meyer, 2005). However, in most cases, their occurrences are fundamentally different. In our samples, groundmass in some olivine and pyroxene basalts represents a rapidly quenched vitrophyre background (Figures 1b and 1c), whereas mesostasis typically occurs at the edges of pyroxene or plagioclase and is commonly associated with troilite (Figure S2c in Supporting Information S1). Subclassification of these phases would therefore improve accuracy.

Third, although we tried our best to minimize human bias during manual mineral-phase annotation, some degree of human interpretation or segmentation is unavoidable, particularly along phase boundaries or within extremely small features where mineral contacts are inherently ambiguous. The limitation on human bias also applies to the rock-type modal abundances. Many modal abundance values reported in the literature are derived from manual point counting of optical images (or, in some cases, automated point counting using systems) over a limited area. These measurements are influenced by both counting errors and sampling errors associated with modal analysis (Bayly, 1959; Solomon, 1963). As shown in Table S1 of Supporting Information S1, the specific regions or thin sections selected for analysis can vary substantially between studies, especially given the strong mineralogical heterogeneity of mare basalts. Consequently, it is common in practice to classify the same mare basalt sample into different rock-type categories (e.g., samples 10020, 10045, 10062, 10092, 12006, 70035, 73219, and 74255) are classified as olivine basalts instead of ilmenite basalts (LSIC Apollo 11, 1977; LSIC Apollo 12, 1970; LSIC Apollo 17, 1973); or the differing classifications of the Apollo 12 feldspathic basalt (Neal & Taylor, 1992).

Lastly, it is important to note that modal mineralogy alone does not fully reflect the true petrogenesis of mare basalts and extraterrestrial materials. Bulk chemistry and cooling rate play important roles, as they strongly influence the resulting textures (Bell et al., 2023; Drake & Weill, 1971). While modal classification provides a useful first-order starting point for petrographic analysis, future work may include the incorporation of textural parameters derived from segmentation results.

5.2. Petrological Applications

Crystal Size Distribution: Our automated system may enable a range of downstream applications. One potential application is crystal size distribution (CSD) analysis (Figure 6), a widely used approach in igneous petrology for quantifying the number and size of crystals within a given volume (Bell et al., 2020, 2023; Rannou & Caroff, 2010). CSD analysis provides quantitative information on the rates of nucleation and crystal growth, cooling histories, and possible evidence for magma mixing (Bell et al., 2020; Neal et al., 2015). As such, CSD analyses of lunar rocks have provided key insights into lunar petrology, including estimates of lava cooling rates and whether a basaltic sample represents an endogenous mare basalt or a lunar impact melt (Donohue & Neal, 2015, 2018; Hui et al., 2011; Neal et al., 2015). Since our method distinguishes individual mineral phases, it enables the estimation of approximate CSDs for each mineral phase (see ilmenite CSD examples in (Figure S17 in Supporting Information S1)). However, improved algorithms are required to accurately delineate phase and grain boundaries—especially in the presence of complex textures such as interlocking, interconnection, inclusions, or cross-cutting features—to achieve more reliable mineral segmentation in future CSD studies.

Magnetization Estimation: By identifying the distribution of metallic Fe at larger scales and refining it with higher-resolution imaging (or interactively examining them in real-time during microscopic analysis), we can assess plausible thermoremanent magnetization ranges (M_{tr}) and identify samples that are promising paleomagnetic targets (Figure 6; Liang et al., 2024; Oliveira et al., 2017). If the volume or mass fraction of metallic Fe, the dominant magnetic carrier, can be quantitatively estimated, the bulk saturation magnetization of the rock can be expressed as

$$M_s = cM_s^{Fe},$$

where c is volumetric concentration of magnetic carriers and $M_s^{Fe} = 1.715 \times 10^6 A/m$ is the saturation magnetization of iron (Dunlop & Özdemir, 2007). Using the squareness of hysteresis $s = M_{rs}/M_s$ typical of lunar mare basalts (Oliveira et al., 2017; Strauss et al., 2021), we can then estimate the saturated remanent magnetization M_{rs}. This, in turn, gives an estimation of the expected thermoremanent magnetization intensity M_{tr} under realistic lunar dynamo field strengths ($B = 1\text{–}100\ \mu$ T) using

$$M_{tr} = BM_{rs}/a$$

where $a = 2070\ \mu$ T is an empirically derived calibration factor (Weiss & Tikoo, 2014; Wieczorek et al., 2023). Importantly, this approach also enables us to filter out samples that are unlikely to provide reliable paleointensity records, those whose expected magnetizations fall below the detection limits of standard superconducting magnetometers (sensitivity $\sim 10^{-12}$ Am^2, assuming sample masses $\gtrsim 0.1$ g). For example, samples containing abundant metallic iron (e.g., kamacite or martensite), which typically exhibit $M_{rs} > 10^{-3}$ Am^2/kg and natural remanent magnetization (NRM) values on the order of 10^{-5}–10^{-6} Am^2/kg, can be meaningfully analyzed in paleomagnetic studies (Jung et al., 2024; Nichols et al., 2021; Shea et al., 2012; Strauss et al., 2021; Tikoo et al., 2012). Meanwhile, samples with extremely low abundances of these phases ($M_{rs} < 10^{-5}$ Am^2/kg and NRM < 10^{-7}–10^{-8} Am^2/kg) generally fall below the detection limits of standard instrumentation and thus do not provide meaningful paleointensity constraints. We note, however, that good paleomagnetic recorders (e.g., single-domain or vortex state metallic iron grains) are uncommon in mare basalts, and even when present, they are generally too small to be confidently identified in SEM images (Jung et al., 2025).

Extension to diverse basaltic lithologies: Another consideration is to classify other types of basalts, especially KREEP or feldspathic basalts (Dymek, 1986; Shih, 1977). These basalt types formed during the pre-mare period (>3.9 Ga) and are therefore much less common (only ~10 samples have been officially identified out of more than 350 basalt samples (Meyer, 2005)). The predominant minerals (e.g., pyroxene and plagioclase) in these samples are similar to those found in the common mare basalt types (e.g., high-Ti and low-Ti mare basalts; see (Table S1 in Supporting Information S1)), and therefore the mineral-identification task itself should not differ substantially. However, classifying these rocks based solely on modal abundances may be more challenging due to the small sample sizes and the large variability in plagioclase abundance, which often falls within the modal ranges of common mare basalt types. Note that Chang'e-5 basalts largely fall within high-Ti and low-Ti mare basalts compositionally (Q.-L. Li et al., 2021a), making them relatively straightforward for mineral identification and further rock-type classification. In contrast, the Chang'e-6 basalt sample is a farside KREEP-rich basalt and would therefore behave more similarly to KREEP basalts (Zhang et al., 2025).

For shocked samples, certain shock features for mare basalts—such as mosaicism or heavily fractured domains (<10–30 GPa; Stoffler et al., 2006; Stöffler et al., 2018)- can be identified as minerals containing abundant voids. Newly formed shock products, such as maskelynite (Milton & Carli, 1963; Rubin, 2015), would likely be classified as their precursor phases (e.g., plagioclase) because their bulk composition remains similar and their grayscale contrasts in SEM images are nearly indistinguishable. Introducing shock indicators such as mosaicism or other known deformation features (Stoffler et al., 2006; Stöffler et al., 2018) may help determine whether the rocks are shocked, and could aid in identifying or at least suggesting plausible mineral phases in shocked lunar rocks.

A related challenge emerges for lunar breccias. At present, classification is largely based on observed textures and macroscopic features—such as fragmental breccia, crystalline melt breccia, and regolith breccia—rather than on genetic origins or detailed mineralogical compositions (Taylor et al., 1991). However, both lunar breccia and brecciated meteorites often contain a diverse assemblage of basaltic clasts, and some breccia preserve igneous textures within their matrices (e.g., crystalline melt breccia). Broadly categorizing the types of clasts (on the order of a few millimeters in size) could aid in establishing sample provenance and support more generalized classification schemes.

5.3. Effects of Data Set Size and Training Data Selection on Model Performance

In biomedical imaging, a few hundred augmented images (<100 unique images) are generally sufficient to construct a robust U-Net model (Nemoto et al., 2021; Ronneberger et al., 2015; Shiina et al., 2025). However, due to the diverse pixel scales and textural complexity of rock images, a larger data set with diverse compositions is typically required to achieve comparable robustness. A key strength of this study lies in the construction of a rare and challenging data set comprising high-resolution SEM BSE images and their corresponding segmented mineral maps (>10,000 augmented images, >1,000 unique images).

We find that (a) model performance increases logarithmically with data set size; (b) a standard U-Net without explicit scale input achieves comparable performance to scale-aware models; and (c) performance differences were more strongly influenced by the choice of training data than by the specific model variant. As such, our results suggest that further improvements could be achieved by constructing larger, high-quality data sets, in addition to exploring different model architectures.

6. Conclusion

Our study demonstrates the feasibility of an automated framework for mineral identification and rock classification in lunar mare basalt samples, despite the large diversity in textures, grain sizes, and modal mineral abundances observed across lunar mare basalts. By leveraging high-resolution SEM imagery and a uniquely segmented data set, the framework enables rapid and scalable mineralogical and petrological analyses. However, the remaining inaccuracies stemming from misidentification of minor phases, the need for finer phase subclassification, and residual human bias in training annotations indicate that expert oversight remains essential for mineralogical and petrological interpretations. As the first approach tailored for extraterrestrial samples, this framework holds strong potential for applications in petrological and rock magnetic studies, as well as for automated mineralogical characterization of existing Apollo collections, lunar meteorites, and future returned sample missions.

Conflict of Interest

The authors declare no conflicts of interest relevant to this study.

Availability Statement

The original Apollo SEM images, annotated images used for training and validating the models, the table of Apollo sample modal abundances from the Lunar Sample Compendium, as well as all codes and trained models will be made publicly available on Zenodo, a permanent online data repository supported by CERN (Jung et al., 2026).

Acknowledgments

We thank the editor, Dr. Cedric John, for handling our submission, and three anonymous reviewers for their constructive and insightful comments, which have improved the manuscript. We express our gratitude to R. Ziegler, the NASA Curation Analysis and Planning Team (CAPTEM), and the Lunar Receiving Laboratory at Johnson Space Center for providing and preparing Apollo samples for S. Tikoo. We are grateful to S. Bell, K. Joy, and J. Day for sharing Apollo mare basalt BSE images, assisting with mineral-phase identification, and for many helpful discussions. We also thank Z. Vaci for conducting the stitched SEM imaging on samples 10003 and 10071. We also appreciate D. Burns for his assistance with microscopy measurements. This research was funded by a NASA FINESST award number 80NSSC21K1541, and was directly supported by the NASA Solar System Exploration Research Virtual Institute cooperative agreement number 80NSSC23M0161.

References

Ali, A., Zhang, N., & Santos, R. M. (2023). Mineral characterization using scanning electron microscopy (SEM): A review of the fundamentals, advancements, and research directions. *Applied Sciences*, *13*(23), 12600. https://doi.org/10.3390/app132312600

Amaral Pascarelli Ferreira, B., Soares Augusto, K., César Álvarez Iglesias, J., Dias Pinheiro Caldas, T., Bryan Magalhães Santos, R., & Paciornik, S. (2024). Instance segmentation of quartz in iron ore optical microscopy images by deep learning. *Minerals Engineering*, *211*, 108681. https://doi.org/10.1016/j.mineng.2024.108681

Baklanova, O., & Shvets, O. (2014). Cluster analysis methods for recognition of mineral rocks in the mining industry. In *2014 4th international conference on image processing theory, tools and applications (IPTA)* (pp. 1–5). IEEE. https://doi.org/10.1109/IPTA.2014.7001972

Bayly, M. B. (1959). Modal analysis by point-counter—The choice of sample area. *Journal of the Geological Society of Australia*, *6*(2), 119–129. https://doi.org/10.1080/00167615908728502

Beaty, D. W., & Albee, A. L. (1980). The geology and petrology of the Apollo 11 landing site. In *Proceedings of the lunar and planetary science conference* (Vol. 11, pp. 23–35).

Bell, S. K. (2024). Backscattered electron (BSE) maps of Apollo 15 mare basalt thin sections. *figshare*. Retrieved from https://figshare.com/s/4fc3036c32a9d678ea9f

Bell, S. K., Joy, K. H., Pernet-Fisher, J. F., & Hartley, M. E. (2020). QEMSCAN as a method of semi-automated crystal size distribution analysis: Insights from Apollo 15 mare basalts. *Journal of Petrology*, *61*(4), egaa047. https://doi.org/10.1093/petrology/egaa047

Bell, S. K., Morgan, D., Joy, K., Pernet-Fisher, J., & Hartley, M. (2023). Determining the thermal histories of Apollo 15 mare basalts using diffusion modelling in olivine. *Geochimica et Cosmochimica Acta*, *357*, 77–91. https://doi.org/10.1016/j.gca.2023.08.009

Bihani, A., Daigle, H., Santos, J. E., Landry, C., Prodanović, M., & Milliken, K. (2022). MudrockNet: Semantic segmentation of mudrock SEM images through deep learning. *Computers & Geosciences*, *158*, 104952. https://doi.org/10.1016/j.cageo.2021.104952

Brown, G. M., Emeleus, C. H., Holland, J. G., & Phillips, R. (1970). Mineralogical, chemical and petrological features of Apollo 11 rocks and their relationship to igneous processes. *Geochimica et Cosmochimica Acta - Supplement*, *1*, 195.

Buz, J., Weiss, B. P., Tikoo, S. M., Shuster, D. L., Gattacceca, J., & Grove, T. L. (2015). Magnetism of a very young lunar glass. *Journal of Geophysical Research: Planets*, *120*(10), 1720–1735. https://doi.org/10.1002/2015JE004878

Cameron, E. N. (1970). Opaque minerals in lunar samples. *Science*, *167*(3918), 623–625. https://doi.org/10.1126/science.167.3918.623

Che, X., Nemchin, A., Liu, D., Long, T., Wang, C., Norman, M. D., et al. (2021). Age and composition of young basalts on the moon, measured from samples returned by Chang'e-5. *Science*, *374*(6569), 887–890. https://doi.org/10.1126/science.abl7957

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 785–794).

Chen, Z., Liu, X., Yang, J., Little, E., & Zhou, Y. (2020). Deep learning-based method for SEM image segmentation in mineral characterization, an example from Duvernay Shale samples in Western Canada Sedimentary Basin. *Computers & Geosciences*, *138*, 104450. https://doi.org/10.1016/j.cageo.2020.104450

Donohue, P. H., & Neal, C. R. (2015). Quantitative textural analysis of ilmenite in Apollo 17 high-titanium mare basalts. *Geochimica et Cosmochimica Acta*, *149*, 115–130. https://doi.org/10.1016/j.gca.2014.11.002

Donohue, P. H., & Neal, C. R. (2018). Textural and mineral chemical evidence for the cumulate origin and evolution of high-titanium basalt fragment 71597. *American Mineralogist*, *103*, 284–297. https://doi.org/10.2138/am-2018-6230

Drake, M., & Weill, D. (1971). Petrology of Apollo 11 sample 10071. A differentiated mini-igneous complex. *Earth and Planetary Science Letters*, *13*(1), 61–70. https://doi.org/10.1016/0012-821X(71)90105-1

Dunlop, D., & Özdemir, Ö. (2007). 5.08 - Magnetizations in rocks and minerals. In G. Schubert (Ed.), *Treatise on geophysics* (pp. 277–336). Elsevier. https://doi.org/10.1016/B978-044452748-6.00093-6

Dymek, R. F. (1986). Characterization of the Apollo 15 feldspathic basalt suite. In P. D. Spudis & G. Ryder (Eds.), *Workshop on the geology and petrology of the Apollo 15 landing site* (Vol. 86–03, p. 52). Lunar and Planetary Institute.

Gibb, F., Stumpfl, E., & Zussman, J. (1970). Opaque minerals in an Apollo 12 rock. *Earth and Planetary Science Letters*, *9*(3), 217–224. https://doi.org/10.1016/0012-821X(70)90031-2

Goldstein, J. I., Newbury, D. E., Joy, D. C., Lyman, C. E., Lifshin, E., Sawyer, L., & Michael, J. (2018). *Scanning electron microscopy and x-ray microanalysis* (4th ed.). Springer. https://doi.org/10.1007/978-1-4939-6676-9

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

Guntoro, P. I., Tiu, G., Ghorbani, Y., Lund, C., & Rosenkranz, J. (2019). Application of machine learning techniques in mineral phase segmentation for X-ray microcomputed tomography (μCT) data. *Minerals Engineering*, *142*, 105882. https://doi.org/10.1016/j.mineng.2019.105882

Haggerty, S. E. (1978). Luna 24: Systematics in spinel mineral chemistry in the context of an intrusive petrogenetic grid. In *Mare crisium: The view from Luna 24* (pp. 523–535).

Head, J. W. (1976). Lunar volcanism in space and time. *Reviews of Geophysics*, *14*(2), 265–300. https://doi.org/10.1029/RG014i002p00265

Hiesinger, H., Head, I. J. W., Wolf, U., Jaumann, R., & Neukum, G. (2011). Ages and stratigraphy of lunar mare basalts: A synthesis. In *Recent advances and current research issues in lunar stratigraphy*. Geological Society of America. https://doi.org/10.1130/2011.2477(01

Hiesinger, H., Jaumann, R., Neukum, G., Head, I., & James, W. (2000). Ages of mare basalts on the lunar nearside. *Journal of Geophysical Research*, *105*(E12), 29239–29275. https://doi.org/10.1029/2000JE001244

Hui, H., Oshrin, J. G., & Neal, C. R. (2011). Investigation into the petrogenesis of Apollo 14 high-al basaltic melts through crystal stratigraphy of plagioclase. *Geochimica et Cosmochimica Acta*, *75*(21), 6439–6460. https://doi.org/10.1016/j.gca.2011.08.015

Izadi, H., Sadri, J., & Bayati, M. (2017). An intelligent system for mineral identification in thin sections based on a cascade approach. *Computers & Geosciences*, *99*, 37–49. https://doi.org/10.1016/j.cageo.2016.10.010

Jiang, C., Abdul Halin, A., Yang, B., Abdullah, L. N., Manshor, N., & Perumal, T. (2024). Res-UNet ensemble learning for semantic segmentation of mineral optical microscopy images. *Minerals*, *14*(12), 1281. https://doi.org/10.3390/min14121281

Jung, J. I., Tikoo, S., Chung, J., & Nichols, C. I. O. (2026). Data and software for automated mineral identification and rock-type classification of lunar mare basalts using SEM images [Dataset]. *Zenodo*. https://doi.org/10.5281/zenodo.20698563

Jung, J.-I., Tikoo, S. M., Burns, D., Váci, Z., & Krawczynski, M. J. (2024). Assessing lunar paleointensity variability during the 3.9–3.5 Ga high field epoch. *Earth and Planetary Science Letters*, *638*, 118757. https://doi.org/10.1016/j.epsl.2024.118757

Jung, J.-I., Tikoo, S. M., Váci, Z., Krawczynski, M. J., Solheid, P., Burns, D. H., & Vailionis, A. (2025). Magnetic mineralogy in lunar mare basalts and implications for paleointensity retrieval. *Journal of Geophysical Research: Planets*, *130*(9), e2025JE009030. https://doi.org/10.1029/2025JE009030

Li, C., Wang, D., & Kong, L. (2021b). Application of machine learning techniques in mineral classification for scanning electron microscopy - Energy dispersive x-ray spectroscopy (SEM-EDS) images. *Journal of Petroleum Science and Engineering*, *200*, 108178. https://doi.org/10.1016/j.petrol.2020.108178

Li, Q.-L., Zhou, Q., Liu, Y., Xiao, Z., Lin, Y., Li, J.-H., et al. (2021a). Two-billion-year-old volcanism on the moon from Chang'e-5 basalts. *Nature*, *600*(7887), 54–58. https://doi.org/10.1038/s41586-021-04100-2

Liang, Y., Tikoo, S. M., & Krawczynski, M. J. (2024). Possibility of lunar crustal magmatism producing strong crustal magnetism. *Journal of Geophysical Research: Planets*, *129*(5), e2023JE008179. https://doi.org/10.1029/2023JE008179

Liu, Y., Wang, X., Zhang, Z., & Deng, F. (2023). Deep learning in image segmentation for mineral production: A review. *Computers & Geosciences*, *180*, 105455. https://doi.org/10.1016/j.cageo.2023.105455

Long, F. (2020). Microscopy cell nuclei segmentation with enhanced u-net. *BMC Bioinformatics*, *21*(1), 8. https://doi.org/10.1186/s12859-019-3332-1

LSIC Apollo 11. (1977). *Apollo-11 lunar sample information catalogue (revised)* (Technical Report No. JSC 12522). National Aeronautics and Space Administration, Lyndon B. Johnson Space Center. (PB.SAM.A11.1977.A).

LSIC Apollo 12. (1970). *Apollo 12 lunar-sample information* (Technical Report No. NASA TR R-353). National Aeronautics and Space Administration. (Manned Spacecraft Center, Houston, Texas 77058. PB.SAM.A12.1970.A).

LSIC Apollo 15. (1971). *Lunar sample information catalog: Apollo 15* (Technical Report No. MSC-03209). National Aeronautics and Space Administration (NASA). Manned Spacecraft Center.

LSIC Apollo 17. (1973). *Lunar sample information catalog: Apollo 17* (Technical Report No. MSC 03211). National Aeronautics and Space Administration, Lyndon B. Johnson Space Center.

Lucey, P., Korotev, R. L., Gillis, J. J., Taylor, L. A., Lawrence, D., Campbell, B. A., et al. (2006). Understanding the lunar surface and space-moon interactions. *Reviews in Mineralogy and Geochemistry*, *60*(1), 83–219. https://doi.org/10.2138/rmg.2006.60.2

Maitre, J., Bouchard, K., & Bédard, L. P. (2019). Mineral grains recognition using computer vision and machine learning. *Computers & Geosciences*, *130*, 84–93. https://doi.org/10.1016/j.cageo.2019.05.009

McCoy, J., & Auret, L. (2019). Machine learning applications in minerals processing: A review. *Minerals Engineering*, *132*, 95–109. https://doi.org/10.1016/j.mineng.2018.12.004

McCubbin, F. M., Vander Kaaden, K. E., Tartèse, R., Klima, R. L., Liu, Y., Mortimer, J., et al. (2015). Magmatic volatiles (H, C, N, F, S, Cl) in the lunar mantle, crust, and regolith: Abundances, distributions, processes, and reservoirs†. *American Mineralogist*, *100*(8–9), 1668–1707. https://doi.org/10.2138/am-2015-4934CCBYNCND

Meyer, C. (2005). Lunar sample compendium. Technical Report.

Milton, D. J., & Carli, P. S. D. (1963). Maskelynite: Formation by explosive shock. *Science*, *140*(3567), 670–671. https://doi.org/10.1126/science.140.3567.670

Neal, C. R., Donohue, P., Fagan, A. L., O'Sullivan, K., Oshrin, J., & Roberts, S. (2015). Distinguishing between basalts produced by endogenic volcanism and impact processes: A non-destructive method using quantitative petrography of lunar basaltic samples. *Geochimica et Cosmochimica Acta*, *148*, 62–80. https://doi.org/10.1016/j.gca.2014.08.020

Neal, C. R., & Taylor, L. A. (1992). Petrogenesis of mare basalts: A record of lunar volcanism. *Geochimica et Cosmochimica Acta*, *56*, 2177–2211. https://doi.org/10.1016/0016-7037(92)90585-F

Nemoto, T., Futakami, N., Kunieda, E., Yagi, M., Takeda, A., Akiba, T., et al. (2021). Effects of sample size and data augmentation on U-Net-Based automatic segmentation of various organs. *Radiological Physics and Technology*, *14*(3), 318–327. https://doi.org/10.1007/s12194-021-00630-6

Nichols, C. I. O., Weiss, B. P., Getzin, B. L., Schmitt, H. H., Béguin, A., Rae, A. S. P., & Shah, J. (2021). The palaeoinclination of the ancient lunar magnetic field from an Apollo 17 basalt. *Nature Astronomy*, *5*(12), 1216–1223. https://doi.org/10.1038/s41550-021-01469-y

Oliveira, J. S., Wieczorek, M. A., & Kletetschka, G. (2017). Iron abundances in lunar impact basin melt sheets from orbital magnetic field data. *Journal of Geophysical Research: Planets*, *122*(12), 2429–2444. https://doi.org/10.1002/2017JE005397

Papike, J., Taylor, L., & Simon, S. (1991). Lunar minerals. In G. H. Heiken, D. T. Vaniman, & B. M. French (Eds.), *Lunar sourcebook, a user's guide to the moon* (pp. 121–181).

Papike, J. J., Hodges, F. N., Bence, A. E., Cameron, M., & Rhodes, J. M. (1976). Mare basalts: Crystal chemistry, mineralogy, and petrology. *Reviews of Geophysics*, *14*(4), 475–540. https://doi.org/10.1029/RG014i004p00475

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Potts, N. J., Tartèse, R., Anand, M., van Westrenen, W., Griffiths, A. A., Barrett, T. J., & Franchi, I. A. (2016). Characterization of mesostasis regions in lunar basalts: Understanding late-stage melt evolution and its influence on apatite formation. *Meteoritics & Planetary Science*, *51*(9), 1555–1575. https://doi.org/10.1111/maps.12681

Qiu, Z., Huang, X., Deng, Z., Xu, X., & Qiu, Z. (2025). PS-YOLO-seg: A lightweight instance segmentation method for lithium mineral microscopic images based on improved YOLOV12-seg. *Journal of Imaging*, *11*(7), 230. https://doi.org/10.3390/jimaging11070230

Rannou, E., & Caroff, M. (2010). Crystal size distribution in magmatic rocks: Proposition of a synthetic theoretical model. *Journal of Petrology*, *51*(5), 1087–1098. https://doi.org/10.1093/petrology/egq012

Rhodes, J. M., & Hubbard, N. J. (1973). Chemistry, classification, and petrogenesis of Apollo 15 mare basalts. In *Proceedings of the lunar science conference* (Vol. 4, pp. 1127–1148).

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi (Eds.), *Medical image computing and computer-assisted intervention – MICCAI 2015* (pp. 234–241). Springer International Publishing.

Rubin, A. E. (2015). Maskelynite in asteroidal, lunar and planetary basaltic meteorites: An indicator of shock pressure during impact ejection from their parent bodies. *Icarus*, *257*, 221–229. https://doi.org/10.1016/j.icarus.2015.05.010

Saxena, N., Day-Stirrat, R. J., Hows, A., & Hofmann, R. (2021). Application of deep learning for semantic segmentation of sandstone thin sections. *Computers & Geosciences*, *152*, 104778. https://doi.org/10.1016/j.cageo.2021.104778

Shea, E. K., Weiss, B. P., Cassata, W. S., Shuster, D. L., Tikoo, S. M., Gattacceca, J., et al. (2012). A long-lived lunar core dynamo. *Science*, *335*(6067), 453–456. https://doi.org/10.1126/science.1215359

Shih, C.-Y. (1977). Origins of KREEP basalts. In *Proceedings of the 8th lunar science conference* (Vol. 2, pp. 2375–2401). Pergamon Press.

Shiina, T., Kimura, K., Takemoto, Y., Tanaka, K., & Kato, R. (2025). Importance of dataset design in developing robust u-net models for label-free cell morphology evaluation. *Journal of Bioscience and Bioengineering*, *139*(4), 329–339. https://doi.org/10.1016/j.jbiosc.2025.01.004

Skinner, B. J. (1970). High crystallization temperatures indicated for igneous rocks from tranquillity base. *Science*, *167*(3918), 652–654. https://doi.org/10.1126/science.167.3918.652

Snape, J. F., Nemchin, A. A., Whitehouse, M. J., Merle, R. E., Hopkinson, T., & Anand, M. (2019). The timing of basaltic volcanism at the Apollo landing sites. *Geochimica et Cosmochimica Acta*, *266*, 29–53. https://doi.org/10.1016/j.gca.2019.07.042

Solomon, M. (1963). Counting and sampling errors in modal analysis by point counter. *Journal of Petrology*, *4*(3), 367–382. https://doi.org/10.1093/petrology/4.3.367

Stöffler, D., Hamann, C., & Metzler, K. (2018). Shock metamorphism of planetary silicate rocks and sediments: Proposal for an updated classification system. *Meteoritics & Planetary Science*, *53*(1), 5–49. https://doi.org/10.1111/maps.12912

Stoffler, D., Ryder, G., Ivanov, B. A., Artemieva, N. A., Cintala, M. J., & Grieve, R. A. F. (2006). Cratering history and lunar chronology. *Reviews in Mineralogy and Geochemistry*, *60*(1), 519–596. https://doi.org/10.2138/rmg.2006.60.05

Strauss, B. E., Tikoo, S. M., Gross, J., Setera, J. B., & Turrin, B. (2021). Constraining the decline of the lunar dynamo field at 3.1 Ga through paleomagnetic analyses of Apollo 12 mare basalts. *Journal of Geophysical Research: Planets*, *126*(3), e2020JE006715. https://doi.org/10.1029/2020JE006715

Sudre, C. H., Li, W., Vercauteren, T., Ourselin, S., & Cardoso, M. J. (2017). Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. *arXiv preprint arXiv:1707.03237*. https://doi.org/10.48550/arXiv.1707.03237

Taylor, G. J., Warren, P., Ryder, G., Delano, J., Pieters, C., & Lofgren, G. (1991). Lunar rocks. In G. H. Heiken, D. T. Vaniman, & B. M. French (Eds.), *Lunar sourcebook, a user's guide to the Moon* (pp. 183–284).

Tikoo, S. M., Weiss, B. P., Buz, J., Lima, E. A., Shea, E. K., Melo, G., & Grove, T. L. (2012). Magnetic fidelity of lunar samples and implications for an ancient core dynamo. *Earth and Planetary Science Letters*, *337–338*, 93–103. https://doi.org/10.1016/j.epsl.2012.05.024

Tikoo, S. M., Weiss, B. P., Cassata, W. S., Shuster, D. L., Gattacceca, J., Lima, E. A., et al. (2014). Decline of the lunar core dynamo. *Earth and Planetary Science Letters*, *404*, 89–97. https://doi.org/10.1016/j.epsl.2014.07.010

Tikoo, S. M., Weiss, B. P., Shuster, D. L., Suavet, C., Wang, H., & Grove, T. L. (2017). A two-billion-year history for the lunar dynamo. *Science Advances*, *3*(8), e1700207. https://doi.org/10.1126/sciadv.1700207

Vernon-Parry, K. (2000). Scanning electron microscopy: An introduction. *III-Vs Review*, *13*(4), 40–44. https://doi.org/10.1016/S0961-1290(00)80006-X

Wang, Z., Hou, Z., & Cao, D. (2025). Enhancing sam-based digital rock image segmentation via edge-semantics fusion. *Applied Computing and Geosciences*, *28*, 100292. https://doi.org/10.1016/j.acags.2025.100292

Weiss, B. P., & Tikoo, S. M. (2014). The lunar dynamo. *Science*, *346*(6214), 1246753. https://doi.org/10.1126/science.1246753

Wieczorek, M. A., Weiss, B. P., Breuer, D., Cébron, D., Fuller, M., Garrick-Bethell, I., et al. (2023). Lunar magnetism. *Reviews in Mineralogy and Geochemistry*, *89*(1), 207–241. https://doi.org/10.2138/rmg.2023.89.05

Wu, Y., Misra, S., Sondergeld, C., Curtis, M., & Jernigen, J. (2019). Machine learning for locating organic matter and pores in scanning electron microscopy images of organic-rich shales. *Fuel*, *253*, 662–676. https://doi.org/10.1016/j.fuel.2019.05.017

Zhang, Q. W. L., Yang, M.-H., Li, Q.-L., Liu, Y., Yue, Z.-Y., Zhou, Q., et al. (2025). Lunar farside volcanism 2.8 billion years ago from Chang'e-6 basalts. *Nature*, *643*(8071), 356–360. https://doi.org/10.1038/s41586-024-08382-0

Zheng, D., Hou, L., Hu, X., Hou, M., Dong, K., Hu, S., et al. (2024). Sediment grain segmentation in thin-section images using dual-modal vision transformer. *Computers & Geosciences*, *191*, 105664. https://doi.org/10.1016/j.cageo.2024.105664

References From the Supporting Information

Albee, A. L., Chodos, A. A., & Gancarz, A. J. (1972). Petrology of Apollo 15 sample 15486. In *The Apollo 15 lunar samples* (pp. 20–25). Lunar Science Institute.

Bailey, J. C., Champness, P. E., Dunham, A. C., Esson, J., Fyfe, W. S., & MacKenzie, W. S. (1970). Mineralogy and petrology of Apollo 11 lunar samples. In A. A. Levinson (Ed.), *Proceedings of the Apollo 11 lunar science conference, volume 1: Mineralogy and petrology* (Vol. 1, pp. 169–185). Pergamon Press.

Baldridge, W. S., Beaty, D. W., Hill, S. M. R., & Albee, A. L. (1979). The petrology of the Apollo 12 pigeonite basalt suite. In *Proceedings of the tenth lunar and planetary science conference*. Lunar and Planetary Institute.

Beaty, D. W., & Albee, A. L. (1978). Comparative petrology and possible genetic relations among the Apollo 11 basalts. In *Proceedings of the 9th lunar and planetary science conference, Houston, Texas, March 13–17, 1978* (Vol. 1, pp. 359–463). Pergamon Press.

Beaty, D. W., Hill, S. M. R., Albee, A. L., & Baldridge, W. S. (1979). Apollo 12 feldspathic basalts 12031, 12038 and 12072: Petrology, comparison and interpretations. In *Proceedings of the tenth lunar and planetary science conference* (Vol. 10, pp. 115–139). Pergamon Press.

Boyd, F. R., & Smith, D. (1971). Compositional zoning in pyroxenes from lunar rock 12021, oceanus procellarum. *Journal of Petrology*, *12*(3), 439–464. https://doi.org/10.1093/petrology/12.3.439

Brett, R., Butler, J. P., Meyer, J. C., Reid, A. M., Takeda, H., & Williams, R. (1971). Apollo 12 igneous rocks 12004, 12008, 12009, and 12022: A mineralogical and petrological study. In *Proceedings of the second lunar science conference* (Vol. 2, pp. 301–317). Pergamon Press.

Brown, G. M., Emeleus, C. H., Holland, J. G., Peckett, A., & Phillips, R. (1971). Picrite basalts, ferrobasalts, feldspathic norites, and rhyolites in a strongly fractionated lunar crust. In *Proceedings of the second lunar science conference* (Vol. 2, pp. 583–602). Pergamon Press.

Brown, G. M., Emeleus, C. H., Holland, J. G., Peckett, A., & Phillips, R. (1972). Mineral-chemical variations in Apollo 14 and Apollo 15 basalts and granitic fractions. In *Lunar and planetary science conference proceedings* (Vol. 3, p. 141).

Brown, G. M., Peckett, A., Emeleus, C. H., Phillips, R., & Pinsent, R. H. (1975). Petrology and mineralogy of Apollo 17 mare basalts. In *Proceedings of the sixth lunar science conference* (Vol. 1, pp. 1–13). Pergamon Press.

Bunch, T. E., Keil, K., & Prinz, M. (1972). Mineralogy, petrology and chemistry of lunar rock 12039. *Meteoritics & Planetary Science*, *7*(3), 245–255. https://doi.org/10.1111/j.1945-5100.1972.tb00439.x

Busche, F. D., Conrad, G. H., Keil, K., Prinz, M., Bunch, T. E., Erlichman, J., & Quaide, W. L. (1971). *Electron microprobe analysis of minerals from Apollo 12 lunar samples* (Technical Report No. Special Publication #3). UNM Institute of Meteoritics.

Carlson, I. C., Walton, J., & Wayne, J. A. (1978). *Apollo 14 rock samples*. (Technical Report No. JSC 14240). National Aeronautics and Space Administration, Lyndon B. Johnson Space Center. (Northrop Services, Inc.).

Carter, J. L., & MacGregor, I. D. (1970). Mineralogy, petrology and surface features of some Apollo 11 samples. In A. A. Levinson (Ed.), *Proceedings of the Apollo 11 lunar science conference, volume 1: Mineralogy and petrology* (Vol. 1, p. 247). Pergamon Press.

Christie, J. M., Lally, J. S., Heuer, A. H., Fisher, R. M., Griggs, D. T., & Radcliffe, S. V. (1971). Comparative electron petrography of Apollo 11, Apollo 12, and terrestrial rocks. In *Proceedings of the second lunar science conference* (Vol. 2, pp. 69–89). Pergamon Press.

Dowty, E., Prinz, M., & Keil, K. (1973). Composition, mineralogy, and petrology of 28 mare basalts from Apollo 15 rake samples. In *Proceedings of the fourth lunar science conference* (Vol. 4, pp. 423–444). Pergamon Press.

Dungan, M. A., & Brown, R. W. (1977). The petrology of the Apollo 12 ilmenite basalt suite. In *Proceedings of the eighth lunar science conference*. Lunar and Planetary Institute.

Dymek, R. F., Albee, A. L., & Chodos, A. A. (1975). Comparative mineralogy and petrology of Apollo 17 mare basalts: Samples 70215, 71055, 74255, and 75055. In *Proceedings of the sixth lunar science conference* (Vol. 1, pp. 49–77). Pergamon Press.

French, B. M., Walter, L. S., & Heinrich, K. J. F. (1970). Quantitative mineralogy of an Apollo 11 lunar sample. In A. A. Levinson (Ed.), *Proceedings of the Apollo 11 lunar science conference, volume 1: Mineralogy and petrology* (Vol. 1, p. 433). Pergamon Press.

French, B. M., Walter, L. S., Heinrich, K. F. J., Lowman, P. D., Doan, A. S., & Adler, I. (1972). *Compositions of major and minor minerals in five Apollo 12 crystalline rocks*. (Technical Report No. SP-306). National Aeronautics and Space Administration (NASA).

Gamble, R. P., Coish, R. A., & Taylor, L. A. (1978). The consanguinity of the oldest Apollo 11 mare basalts. In *Proceedings of the 9th lunar and planetary science conference, Houston, Texas, March 13–17, 1978* (Vol. 1, pp. 495–507). Pergamon Press.

Gancarz, A., Albee, A., & Chodos, A. (1971). Petrologic and mineralogic investigation of some crystalline rocks returned by the Apollo 14 mission. *Earth and Planetary Science Letters*, *12*(1), 1–18. https://doi.org/10.1016/0012-821X(71)90050-1

Grove, T. L., & Walker, D. (1977). Cooling histories of Apollo 15 quartz-normative basalts. In *Proceedings of the eighth lunar science conference* (Vol. 3, pp. 1501–1520). Pergamon Press.

Grove, T. L., Walker, D., Longhi, J., Stolper, E., & Hays, J. F. (1973). Petrology of rock 12002 and origin of picritic basalts at Oceanus Procellarum. In *Proceedings of the fourth lunar science conference* (Vol. 4, pp. 995–1011). Pergamon Press.

Haggerty, S. E., Boyd, F. R., Bell, P. M., Finger, L. W., & Bryan, W. B. (1970). Opaque minerals and olivine in lavas and breccias from mare tranquillitatis. In A. A. Levinson (Ed.), *Proceedings of the Apollo 11 lunar science conference, volume 1: Mineralogy and petrology* (Vol. 1, pp. 513–526). Pergamon Press.

Heuer, A. H., Nord, G. L., Radcliffe, S. V., Fischer, R. M., Lally, J. S., Christie, J. M., & Griggs, D. T. (1972). High voltage electron petrographic study of Apollo 15 rocks. In *The Apollo 15 lunar samples* (pp. 98–102). Lunar Science Institute.

Hollister, L. S., & Crawford, M. L. (1977). Melt immiscibility in Apollo 15 KREEP: Origin of Fe-rich mare basalts. In *Proceedings of the eighth lunar science conference* (Vol. 2, pp. 2419–2432). Pergamon Press.

James, O. B., & Jackson, E. D. (1970). Petrology of the Apollo 11 ilmenite basalts. *Journal of Geophysical Research*, *75*(29), 5793–5824. https://doi.org/10.1029/JB075i029p05793

Juan, V. C., Chen, J. C., Huang, C. K., Chen, P. Y., & Wang Lee, C. M. (1972). Petrology and chemistry of some Apollo 15 crystalline rocks. In *The Apollo 15 lunar samples* (pp. 110–115). Lunar Science Institute.

Keil, K., Dowty, E., & Prinz, M. (1974). *Description, classification and inventory of 113 Apollo 17 rake samples from stations 1a, 2, 7 and 8* (Technical Report). National Aeronautics and Space Administration, Lyndon B. Johnson Space Center.

Klein, J. C., Drake, J. C., & Frondel, C. (1971). Mineralogical, petrological, and chemical features of four Apollo 12 lunar microgabbros. In *Proceedings of the second lunar science conference* (Vol. 2, pp. 265–280). Pergamon Press.

Kridelbaugh, S. J., & Weill, D. F. (1973). The mineralogy and petrology of ilmenite basalt 75055 (abs). *EOS Transactions of the American Geophysical Union*, *54*, 597–598.

Kushiro, I., & Nakamura, Y. (1970). Petrology of some lunar crystalline rocks. In A. A. Levinson (Ed.), *Proceedings of the Apollo 11 lunar science conference, volume 1: Mineralogy and petrology* (Vol. 1, p. 607). Pergamon Press.

Li, Q. L., Zhou, Q., Liu, Y., Xiao, Z., Lin, Y., Li, J. H., et al. (2021b). Young volcanism on the moon inferred from Chang'e-5 basalts. *Science*, *374*(6571), 887–890. https://doi.org/10.1126/science.abg8112

Longhi, J., Walker, D., & Hays, J. F. (1972a). Petrography and crystallization history of basalts 14310 and 14072. In *Proceedings of the third lunar science conference* (Vol. 3, pp. 131–144). Pergamon Press.

Longhi, J., Walker, D., Stolper, E. N., Grove, T. L., & Hays, J. F. (1972b). Petrology of mare/rille basalts 15555 and 15065. In *The Apollo 15 lunar samples* (pp. 131–134). Lunar Science Institute.

Longhi, J., Walker, D., Grove, T. L., Stolper, E. M., & Hays, J. F. (1974). The petrology of the Apollo 17 mare basalts. In *Proceedings of the fifth lunar science conference* (Vol. 1, pp. 447–469). Pergamon Press.

Lovering, J. F., & Ware, N. G. (1970). Electron-probe microanalyses of minerals and glasses in Apollo 11 lunar samples. In A. A. Levinson (Ed.), *Proceedings of the Apollo 11 lunar science conference* (Vol. 1, p. 633). Pergamon Press.

Ma, M.-S., Murali, A. V., & Schmitt, R. A. (1976). Chemical constraints for mare basalt genesis. In *Proceedings of the seventh lunar science conference* (Vol. 2, pp. 1673–1695). Pergamon Press.

McGee, P. E., Warner, J. L., & Simonds, C. H. (1977). *Introduction to the Apollo collections. Part 1: Lunar igneous rocks* (Technical Memorandum No. NASA-TM-X-74633). NASA Lyndon B. Johnson Space Center.

McKay, G. A., Wiesmann, H., Nyquist, L. E., Wooden, J. L., & Bansal, B. M. (1978). Petrology, chemistry, and chronology of 14078: Chemical constraints on the origin of KREEP. In *Proceedings of the ninth lunar and planetary science conference* (Vol. 9, pp. 661–687). Pergamon Press.

Meyer, H. O. A., & Boctor, N. Z. (1974b). Opaque mineralogy: Apollo 17, rock 75035. In *Proceedings of the fifth lunar conference (supplement 5, geochimica et cosmochimica acta)* (Vol. 1, pp. 707–716). Lunar and Planetary Institute.

Meyer, C. E., & Wilshire, H. G. (1974a). "dunite" inclusion in lunar basalt 74275. In *Abstracts of the 5th lunar and planetary science conference* (Vol. 5, p. 503).

Neal, C. R., & Taylor, L. A. (1993). *Catalog of Apollo 17 rocks: Volume 3 – Central Valley, Part 2* (Technical Report No. JSC-26088). NASA Johnson Space Center. (Space and Life Sciences Directorate, Solar System Exploration Division, Office of the Curator #87).

Neal, C. R., Taylor, L. A., Hughes, S. S., & Schmitt, R. A. (1990). The significance of fractional crystallization in the petrogenesis of Apollo 17 Type A and B high-Ti basalts. *Geochimica et Cosmochimica Acta*, *54*(6), 1817–1833. https://doi.org/10.1016/0016-7037(90)90410-m

Neal, C. R., Hacker, M. D., Snyder, G. A., Taylor, L. A., Liu, Y.-G., & Schmitt, R. A. (1994). Basalt generation at the Apollo 12 site, Part 1: New data, classification, and re-evaluation. *Meteoritics*, *29*(3), 334–348. https://doi.org/10.1111/j.1945-5100.1994.tb00597.x

Newton, R. C., Anderson, A. T., & Smith, J. V. (1971). Accumulation of olivine in rock 12040 and other basaltic fragments in the light of analysis and syntheses. In *Proceedings of the second lunar science conference* (Vol. 2, pp. 575–582). Pergamon Press.

Nord, G. L. J., Lally, J. S., Heuer, A. H., Christie, J. M., Radcliffe, S. V., & Griggs, D. T. (1973). Petrologic study of igneous and metaigneous rocks from Apollo 15 and 16 using high voltage transmission electron microscopy. In *Proceedings of the fourth lunar science conference* (Vol. 4, pp. 953–970). Pergamon Press.

Patchen, A. P., & Taylor, L. A. (2004). The most reduced rock from the moon – Apollo 14 basalt 14053: Extreme reduction entirely from a reheating event. In *Lunar and planetary science xxxv* (p. 1762). Lunar and Planetary Institute.

PET, P. E. T. (1973). Apollo 17 lunar samples: Chemical and petrographic description. *Science*, *182*(4113), 659–672. https://doi.org/10.1126/science.182.4113.659

Roedder, E., & Weiblen, P. W. (1975). Anomalous low-k silicate melt inclusions in ilmenite from Apollo 17 basalts. In *Proceedings of the sixth lunar science conference* (Vol. 6, pp. 147–164). Pergamon Press.

Ryder, G. (1985). *Catalog of Apollo 15 rocks, Part 3: 15475–15698* (Technical Report Nos. Curatorial Branch Publication 72, JSC 20787). NASA Johnson Space Center.

Schaal, R. B., & Hörz, F. (1977). Shock metamorphism of lunar and terrestrial basalts. In *Proceedings of the 8th lunar science conference* (Vol. 2, pp. 1697–1729). Pergamon Press.

Schnare, D. W., Day, J. M., Norman, M. D., Liu, Y., & Taylor, L. A. (2008). A laser-ablation ICP-MS study of Apollo 15 low-titanium olivine-normative and quartz-normative mare basalts. *Geochimica et Cosmochimica Acta*, *72*(10), 2556–2572. https://doi.org/10.1016/j.gca.2008.02.021

Shervais, J. W., Vetter, S. K., & Linstrom, M. M. (1990). Chemical differences between small subsamples of Apollo 15 olivine-normative basalts. In *Proceedings of the 20th lunar and planetary science conference* (pp. 109–126). Lunar and Planetary Institute.

Steele, I. M., Smith, J. V., & Grossman, L. (1972). Mineralogy and petrology of Apollo 15 rake samples: I. basalts. In *The Apollo 15 lunar samples* (pp. 158–160). Lunar Science Institute.

Taylor, L. A., & Misra, K. C. (1975). Pyroxene-phyric basalt 15075: Petrography and petrogenesis. In *Proceedings of the sixth lunar science conference* (Vol. 6, pp. 165–179). Pergamon Press.

Taylor, L. A., Kullerud, G., & Bryan, W. B. (1971). Opaque mineralogy and textural features of Apollo 12 samples and a comparison with Apollo 11 rocks. In *Proceedings of the second lunar science conference* (Vol. 2, pp. 855–871). Pergamon Press.

Taylor, L. A., McKay, D. S., Graf, J., Patchen, A., Wentworth, S., & Oder, R. (1992). Magnetic beneficiation of high-ti mare basalts: Petrographic analyses. In *Lunar and planetary science conference abstracts* (Vol. 23, p. 1415). Lunar and Planetary Institute.

Warner, R. D., Keil, K., Nehru, C. E., & Taylor, G. J. (1978). *Catalogue of Apollo 17 rake samples from stations 1a, 2, 7, and 8 (No. No. 18)*. Department of Geology and Institute of Meteoritics, University of New Mexico.

Warner, R. D., Keil, K., Prinz, M., Laul, J. C., Murali, A. V., & Schmitt, R. A. (1975). Mineralogy, petrology, and chemistry of mare basalts from Apollo 17 rake samples. In *Proceedings of the sixth lunar science conference* (Vol. 1, pp. 193–220). Pergamon Press.

Warner, R. D., Warren, R. G., Mansker, W. L., Berkley, J. L., & Keil, K. (1976). *Electron microprobe analyses of olivine, pyroxene and plagioclase from Apollo 17 rake sample mare basalts* (Technical Report No. Special Publication 15). UNM Institute of Meteoritics.

Weigand, P. W. (1973). Petrology of a coarse-grained Apollo 17 ilmenite basalt. *EOS Transactions of the American Geophysical Union*, *54*, 621–622.

Weill, D. F., McCallum, I. S., Bottinga, Y., Drake, M. J., & McKay, G. A. (1970). Mineralogy and petrology of some Apollo 11 igneous rocks. In A. A. Levinson (Ed.), *Proceedings of the Apollo 11 lunar science conference, volume 1: Mineralogy and petrology* (Vol. 1, p. 937). Pergamon Press.