

# SPINAL OSTEOPHYTE DETECTION VIA ROBUST PATCH EXTRACTION ON MINIMALLY ANNOTATED X-RAYS

Soumya Snigdha Kundu\*

Yuanhan Mo<sup>†</sup>

Nicharee Srikijkasemwat\*\*

Bartłomiej W. Papież<sup>†</sup>

\* Department of Surgical & Interventional Engineering, King's College London, London, UK

\*\* Department of Engineering Science, University of Oxford, Oxford, UK

<sup>†</sup> Big Data Institute, University of Oxford, Oxford, UK

## ABSTRACT

The development and progression of arthritis is strongly associated with osteophytes, which are small and elusive bone growths. This paper presents one of the first efforts towards automated spinal osteophyte detection in spinal X-rays. A novel automated patch extraction process, called SegPatch, has been proposed based on deep learning-driven vertebrae segmentation and the enlargement of mask contours. A final patch classification accuracy of 84.5% is secured, surpassing a baseline tiling-based patch generation technique by 9.5%. This demonstrates that even with limited annotations, SegPatch can deliver superior performance for detection of tiny structures such as osteophytes. The proposed approach has potential to assist clinicians in expediting the process of manually identifying osteophytes in spinal X-ray.

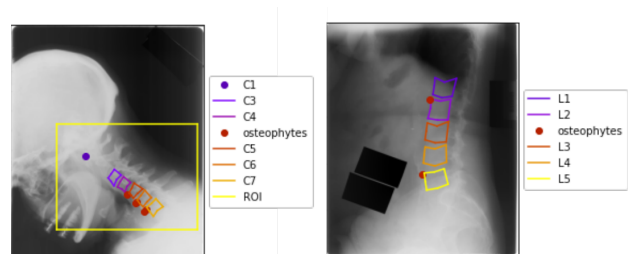
**Index Terms**— Keypoint Identification, Osteophyte Detection, Patch Extraction.

## 1. INTRODUCTION

One of the features observed in osteoarthritis (OA) [1] are osteophytes which are fibrocartilage-capped bony outgrowths. They play a strong role in determining the severity and progression of OA. Osteophytes also cause problems by impinging nerves such as spinal roots [2] and pose significant issues in many other spinal diseases.

The manual identification of osteophytes has been generally done using Magnetic Resonance Imaging (MRI) [3]. In contrast, the use of X-rays has been limited due to the challenges associated with the quality and 2D nature of conventional radiography. Automated detection of spinal osteophytes has also been explored. Wang et al. [4] extracted region proposals and classified them from unwrapped cortical shell maps in Positron Emission Tomography and Computed Tomography scans of the vertebral bodies of the thoracic and lumbar spine. They achieved 85% sensitivity at 2 false positive detections per patient.

\*S.S.K carried out the work as a part of the Big Data Institute Summer Internship Programme at the University of Oxford.

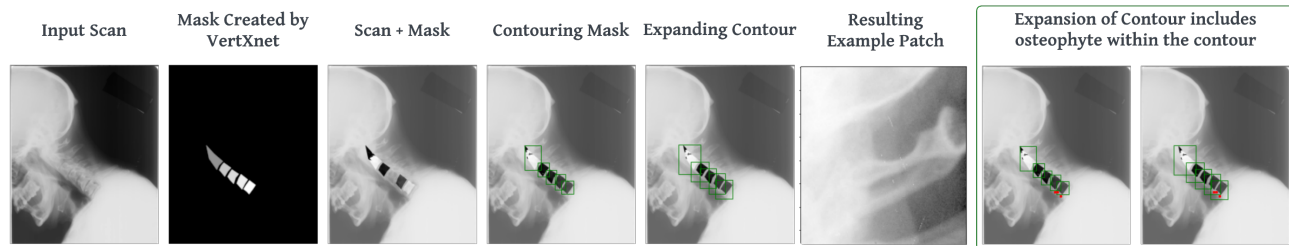


**Fig. 1.** Example of annotations for cervical (left) and lumbar (right) X-ray scan available from NHANESII<sup>2</sup> data set.

Automated approaches to detect osteophytes in other parts of the body has been proposed as well. Ebsim et al., [5] identified hip osteophytes in dual-energy X-ray absorptiometry by segmenting the osteophyte regions using U-Net [6] and manual cropping (to highlight the area of interest). Banerjee et al [7] identified hand osteophytes in X-rays through cellular neural networks (neural networks where communication is allowed only between neighboring units). They automated the ground truth acquisition but they lacked any involvement from an actual clinical representative throughout the acquisition process. The VinDr-SpineXR study [8] provides an extensive dataset with baseline results on multiple spinal diseases, including osteophytes. The poor results (and trend of increasing performance with increasing label size) match with our detection findings in sec 5.3.

The number and size of the osteophytes could be seen as an estimate of the OA progress. Clinical experts' annotations are time-consuming and expensive while being prone to errors (i.e. inter- and intra-reader variability). Especially for tiny structures such as osteophytes. To the best of our knowledge, there has been no automated method for spinal osteophyte detection for spinal X-rays.

Our contributions are as follows - We have developed a method to automatically identify osteophytes from conventional spinal radiography. First, we proposed a robust mechanism SegPatch, to extract patches from the spinal X-ray images through a deep-learning based vertebrae body segmentation approach [9].



**Fig. 2.** A representation of the SegPatch pipeline with a resulting patch. The green box highlights the necessity to expand contours to fully encompass the osteophytes within the boundary. Red dots denote the osteophytes.

Unlike the standard tiling approach, the newly proposed patching generates patches that are associated with potential locations of the osteophytes. Thereafter, we trained multiple neural network models for classification. Our framework was evaluated using the publicly available spine X-ray data set NHANESII. We compared SegPatch with the baseline tiling method. The results show the significant advantage of using SegPatch.

## 2. DATA

The method was developed based on the publicly available data set<sup>1</sup> collected via the Second National Health and Nutrition Examination Survey (NHANES II)<sup>2</sup> conducted by the National Center for Health Statistics during the years 1976-1980. The data encompasses digitized versions of the 17100 X-ray films. 9670 were cervical (of the size 1462 x 1755) and 7430 lumbar (of the size 2048 x 2487). However, only 241 cervical and 232 lumbar X-ray images were annotated for pixel coordinates indicating the presence of osteophytes. Each vertebrae was delineated using six pixel points. A 75:25 split was employed for training and testing data. The exemplar radiographs and the associated annotations have been shown in Fig. 1.

## 3. METHODOLOGIES

Given the challenges associated with the detection of the miniscule osteophytes, the main goal of this research was to create a technique that could ascertain whether a particular image patch contained an osteophyte. As seen in Fig. 1, osteophytes generally tend to occur on the corners of the vertebrae. Based on this observation we developed a pipeline to extract several patches from high-resolution radiographs (Sec. 3.1) and classify those patches for the presence of an osteophyte (Sec. 3.2). Therefore, we set the hypothesis that if a patch is classified as "present", then either of the two corners of the vertebrae will have an osteophyte. If the patch contains only one corner then it is bound to have a single osteophyte.

For an qualitative evaluation, the explainability of the classifier was tested via class activation map generation method SS-CAM [10], to understand and verify the focus of the classifier.

### 3.1. Robust Patch Creation

Two methods of patch creation were investigated in this study.

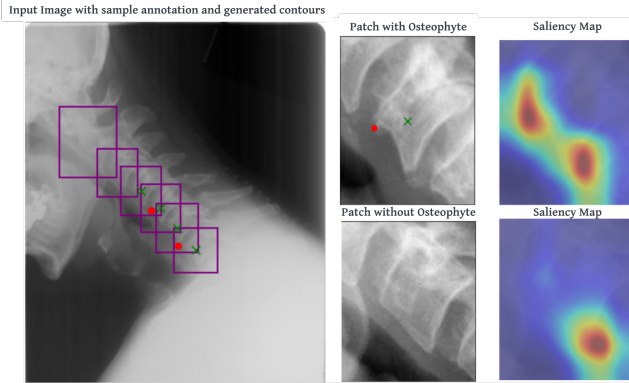
#### 3.1.1. Tiling

The tiling approach was inspired from Ozge et al. [11]. The authors proposed it for a detection task in high-resolution pedestrian and vehicle images. Here, tiling is used as an image pre-processing step to split the image into equal-sized non-overlapping patches and label them based on whether a patch contained an annotation - indicating the presence of an osteophyte. To ensure the corner regions were not split amongst patches, the annotations were considered as boxes. The tiling approach generated a highly imbalanced set of patches. The number of patches containing osteophytes amounted to 1453 (< 5%). Those without equalled to 32160. The boxes used as annotations were also used as the ground truth bounding boxes for the detection task in section 5.3.

#### 3.1.2. SegPatch: Segmentation-driven Patch Extraction

The main challenge in tiling was the significant class imbalance generated by considering the complete scan. Any approach which could focus on the vertebrae region alone, would potentially improve performance as the amount of noise seen by the classifier would reduce. Therefore, a custom patch-generating method called SegPatch was devised to overcome the above limitations.

The SegPatch process begins by utilizing a vertebrae segmentation network VertXNet [9] to localise the vertebrae in the scan. However, while the segmentation accounted for each vertebra in each image, it failed to encapsulate the osteophyte point locations (given as per the dataset). This is most likely because the reader was asked to place the marker along the center of gravity for each vertebra.



**Fig. 3.** Saliency results for a positive and a negative patch. Red dots: Osteophytes. Purple boxes: Contours.

Therefore, a simple contour of the vertebra did not suffice. To account for the inclusion of the osteophyte location into the final resulting patches, another strategy was implemented consisting of expanding the formed contour. The majority of the osteophytes occurred on the left side of the vertebrae hence we first expanded the contour in the  $-X$  direction. To account for the curvature of the spines, we also expanded the contour in the  $+Y$  direction. This allowed for the gaps between the vertebrae to be covered. Different size of expansion was utilized for cervical and lumbar scans to account for the variation in vertebra sizes. SegPatch finally resulted in 1680 present patches and 1517 absent patches, which was more favourably distributed than the tiling method described above. The details and visual depiction are outlined in Fig. 2. A rigid rule and clinical prior based heuristic was used to expand the contour to prevent further processing of the image prior to being fed to the classifier.

### 3.2. Classification of Osteophytes

Multiple state-of-the-art classifiers were tested including ResNet-18/50 and DenseNet-121/169, all pretrained on ImageNet. Given the limited amount of annotated images, we employed data augmentation including random rotations (up to 30 degrees) and random histogram equalization (with a variance of 0.5). The input dimensions for both strategies are the same. The image patches were resized to  $(244 \times 224)$  for SegPatch, and the tiling limit was set to  $(244 \times 224)$ . For both methods, the cross-entropy loss, and the Stochastic Gradient Descent were used. Tiling was fine-tuned with a learning rate of 0.01 and a momentum of 0.7, while SegPatch achieved optimal results with a learning rate of 0.002 and a momentum of 0.09. Additionally, a learning rate scheduler with a step size of 7 and a gamma value of 0.1 was employed. Peak performance was reached at 50 epochs for both approaches.

**Table 1.** Classification results for the proposed patching algorithms.

Model	Accuracy Score			
	SegPatch		Tiling	
	Train	Test	Train	Test
ResNet-18	84.80	80.34	80.10	73.34
ResNet-50	85.34	82.00	79.80	75.40
DenseNet-121	87.05	<b>84.64</b>	77.73	70.05
DenseNet-169	87.20	82.75	76.90	69.70

Scores are averaged over 3 runs. All scores at  $\pm 0.75$ .

## 4. RESULTS

As indicated in Tab. 1, the tiling method produced an average test accuracy of 75.4%, whereas SegPatch surpasses this performance, achieving an average accuracy of 84.5%.

Furthermore, Fig. 3 reveals an interesting observation: in the present patch, the saliency maps generated using SS-CAM are predominantly concentrated over the edges where the osteophyte is present. In the absence of an osteophyte, it primarily focuses on the edge but not on the location of the osteophyte. This observation vividly illustrates that the model is discerning the edges and making determinations regarding the presence or absence of growth by looking specifically at the edges of the vertebrae.

## 5. DISCUSSION

This paper introduces an automated approach for the detection of osteophytes in spinal X-rays. The method employs a specialized automated patch generation technique called SegPatch to generate input data for a patch classifier, which is built upon a fine-tuned DenseNet-121 model. The method secured an overall 84.5% accuracy score and compares favorably with the baseline tiling method.

### 5.1. Solving the class imbalance in Tiling

Using rank-consistent ordinal regression framework (CORN) loss [12] or focal loss [13] to tackle class imbalance did not show any improvement over weighted cross-entropy. Neither did the use of over-sampling. Choosing a fixed number of random patches runs the risk of missing the already diminished number of positive patches. Owing to the significant variability in the images, an algorithmic strategy to eliminate a larger number of negative sample seemed infeasible. Experimentation also involved utilising larger models for the patch classification task but they exhibited over-fitting and were consequently excluded from consideration.



**Fig. 4.** Visualisation of poor performance from off-the-shelf detector - FasterRCNN. Purple circles denote annotation location and squares denote predicted bounding box.

### 5.2. Problems with generating more precise patches

The solution to class imbalance would be to generate more precise patches focused on the edges of the vertebrae. However, generating more precise patches faces quite a few challenges. The primary problem arises with the extreme curvature of the spine in certain scans.

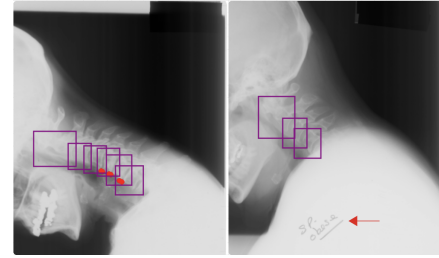
There are also cases with artifacts present on the scan, e.g. text on the scan (see red arrow in Fig. 5). Given the random nature of occurrence of osteophytes, one has to account for each corner of each vertebrae in every scan. There is also a size difference in the Lumbar and Cervical scans. Hence, methods have to be size-invariant like that of SegPatch.

#### 5.2.1. Patch size variability in SegPatch

While tiling produces a uniform patch size, SegPatch generates non-uniformly sized patches which may lead to lower resolution patches for the larger patches generated. Uniform patch size was not feasible as that either resulted in multiple missing osteophytes or a large number (and area) of overlaps between the patches.

### 5.3. Poor Performance from off-the-shelf detectors

Initially, to develop a solution for osteophyte detection, we investigated the state-of-the-art detection method, FasterRCNN [14], and achieved an average mAP and accuracy score of 0.1%, 0.2%, 20% and 30% as the size of the bounding increasing from 10, 18, 36 to 54 pixels, respectively. The results did not improve with different forms of augmentation (including tiling, which is originally devised for detection-based tasks) or utilizing the latest versions of YOLO (NAS, v8). The key positive insight derived from the detection task using FasterRCNN is seen in Fig. 4. The model clearly understands the typical locations where the osteophytes appear.



**Fig. 5.** Example of difficult (but successful) to label scans. (Left): High curvature of spine. (Right): Presence of artifact (text) denoted by red arrow with with no osteophyte labels.

Notably at the corners of the vertebrae. A trend similarly observed in VinDr-SpineXR study [8]. Consequently, it is advisable for future methodologies to place greater emphasis on the identification aspect rather than localization in the context of osteophyte detection.

## 6. CONCLUSION

This paper presents one of the first studies for spinal osteophyte detection. It integrates the use of a specialized patching strategy called SegPatch in conjunction with a fine-tuned DenseNet-121 network. By employing patch classification, the identification process gains enhanced efficiency in detecting osteophytes as osteophytes tend to generally originate in the corner regions of vertebrae. Our method achieves 84.5% accuracy score and a 86.62% specificity score for the patch classifier which generates saliency maps that align well with both positive and negative classifications. A study done by Nevalainen et al., showed a specificity of 75% in detecting osteophytes from MRI [15].

The findings, combined with the analysis of the results in this paper lay the groundwork for a more effective clinical approach to osteophyte detection in X-ray images and other smaller structures. This is all while being robust to multiple orientations of the spine present in the scans, artifacts, minimal annotations and comparatively small training data.

In addition, this work shows that osteophytes can be detected using a single modality instead of multi-modality approaches [4]. This efficiency not only saves valuable time but also has the potential to enhance the accuracy and consistency of diagnoses, ultimately improving patient care and treatment outcomes.

## Notes

1. This research study was conducted retrospectively using human subject data made available in open access<sup>2</sup>. Ethical approval was not required as confirmed by the license attached with the open access data.
2. Dataset information can be found here: <https://data.lhncbc.nlm.nih.gov/public/NHANES/X-rays/README.html>.

## Acknowledgments

S. S. Kundu was supported by the UK Medical Research Council [MR/N013700/1] and the King's College London MRC Doctoral Training Partnership in Biomedical Sciences. Y. Mo is supported by the Oxford BDI-Novartis collaboration for AI in medicine. B. W. Papiez acknowledges Senior Fellowship at Population Health. The computational aspects of this research were supported by the Wellcome Trust Core Award Grant Number 203141 /Z/16/Z and the NIHR Oxford BRC. The authors would also like to thank Mona Furukawa for her edits on the final manuscript.

## 7. REFERENCES

- [1] Peter M Van Der Kraan and Wim B Van Den Berg, "Osteophytes: relevance and biology," *Osteoarthritis and cartilage*, vol. 15, no. 3, pp. 237–244, 2007.
- [2] Thomas Louis Jones and Michael S Hisey, "L5 radiculopathy caused by l5 nerve root entrapment by an l5-s1 anterior osteophyte," *International Journal of Spine Surgery*, vol. 6, pp. 174–177, 2012.
- [3] Ryosuke Tozawa, Yuya Ogawa, Yusuke Minamoto, Taishi Ninomiya, Takahiro Ogura, Shotaro Watanabe, Seiji Kimura, Yuki Shiko, Yohei Kawasaki, Ryuichiro Akagi, et al., "Possible role of mri-detected osteophytes as a predictive biomarker for development of osteoarthritis of the knee: A study using data from the osteoarthritis initiative," *Osteoarthritis and Cartilage Open*, vol. 3, no. 4, pp. 100200, 2021.
- [4] Yinong Wang, Jianhua Yao, Joseph E Burns, Jiamin Liu, and Ronald M Summers, "Detection of degenerative osteophytes of the spine on pet/ct using region-based convolutional neural networks," in *Computational Methods and Clinical Applications for Spine Imaging: 4th International Workshop and Challenge, CSI 2016, Held in Conjunction with MICCAI 2016, Athens, Greece, October 17, 2016, Revised Selected Papers 4*. Springer, 2016, pp. 116–124.
- [5] Raja Ebsim, Benjamin G Faber, Fiona Saunders, Monika Frysz, Jenny Gregory, Nicholas C Harvey, Jonathan H Tobias, Claudia Lindner, and Timothy F Cootes, "Automatic segmentation of hip osteophytes in dxa scans using u-nets," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2022, pp. 3–12.
- [6] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.
- [7] Sreeparna Banerjee, Satyajit Bhunia, and Gerald Schaefer, "Osteophyte detection for hand osteoarthritis identification in x-ray images using cnns," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, 2011, pp. 6196–6199.
- [8] Hieu T Nguyen, Hieu H Pham, Nghia T Nguyen, Ha Q Nguyen, Thang Q Huynh, Minh Dao, and Van Vu, "Vindr-spinexr: A deep learning framework for spinal lesions detection and classification from radiographs," in *Medical Image Computing and Computer Assisted Intervention—MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part V 24*. Springer, 2021, pp. 291–301.
- [9] Yao Chen, Yuanhan Mo, Aimee Readie, Gregory Ligozio, Indrajeet Mandal, Faiz Jabbar, Thibaud Coroller, and Bartłomiej W Papież, "Vertxnet: an ensemble method for vertebral body segmentation and identification from cervical and lumbar spinal x-rays," *Scientific Reports*, vol. 14, no. 1, pp. 3341, 2024.
- [10] Haofan Wang, Rakshit Naidu, Joy Michael, and Soumya Snigdha Kundu, "Ss-cam: Smoothed scorecam for sharper visual feature localization," *arXiv preprint arXiv:2006.14255*, 2020.
- [11] F Ozge Unel, Burak O Ozkalayci, and Cevahir Cigla, "The power of tiling for small object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2019, pp. 0–0.
- [12] Xintong Shi, Wenzhi Cao, and Sebastian Raschka, "Deep neural networks for rank-consistent ordinal regression based on conditional probabilities," *Pattern Analysis and Applications*, vol. 26, no. 3, pp. 941–955, jun 2023.
- [13] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.
- [14] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.
- [15] Mika T Nevalainen, Antti-Pekka Uusimaa, and Simo Saarakkala, "The ultrasound assessment of osteoarthritis: the current status," *Skeletal Radiology*, pp. 1–12, 2023.