Bone mineral density prediction from spine lateral X-ray images

Jinsung Jung

Department of Medicine

The Graduate School, Yonsei University

Bone mineral density prediction from spine lateral X-ray images

Directed by Professor John A. Linton

The Master's Thesis
submitted to the Department of Medicine,
the Graduate School of Yonsei University
in partial fulfillment of the requirements for the degree
of Master of Medical Science

Jinsung Jung

December 2020

This certifies that the Master's Thesis
of Jinsung Jung is approved.

Thesis Supervisor: John A. Linton

Thesis Committee Member#1: Young Han Lee

Thesis Committee Member#2: Hwiyoung Kim

The Graduate School
Yonsei University

December 2020

ACKNOWLEDGEMENTS

First of all, I would like to express my gratitude to Professor John A. Linton for his warmhearted supervision throughout my master's study. I am also sincerely grateful to Professor Young Han Lee and Hwiyoung Kim for their valuable and constructive advice. Their guidance has not only enabled me to excel academically, but also benefited me professionally.

I would like to sincerely thank Siyeong Lee of NAVER LABS for helpful feedback and discussions. I also thank Yeonsu Kwak of Korea Institute of Science and Technology (KIST) for his support and comments.

Finally, I would like to express my appreciation to my family for providing me with invaluable support so that I could embark on this journey.

<TABLE OF CONTENTS>

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Bone mineral density prediction from spine lateral X-ray images

Jinsung Jung

Department of Medicine

The Graduate School, Yonsei University

(Directed by Professor John A. Linton)

Objectives: This study aims to predict bone mineral density with only lumbar X-ray images of the patients.

Materials and methods: Spine lateral X-ray and DXA data were collected from patients who visited Severance Hospital, Seoul, Republic of Korea, between 2005 and 2020 and were aged 50 or over. One thousand X-ray images were randomly extracted and processed. A two-step method was applied for vertebral body detection and BMD prediction using a single X-ray image. For vertebral body detection, both one-stage and two-stage object detection methods were employed. For BMD prediction, the X-ray images were meshed with T-score values and applied to two different models, Basic CNN and ResNet-18.

Results: The AP of vertebral body detection was the highest in RetinaNet-500 with the AP-loss model, with a value of 73.1. Among the classification methods, the accuracy of distinguishing osteoporosis was the highest in both T-score and spine level classification, which were 77.66 and 70.48, respectively.

Conclusion: In this study, each lumbar spine was successfully detected only with spine X-ray images. It was also demonstrated that BMD can be predicted through the lumbar spines from L1 to L4.

Keywords: bone mineral density, osteoporosis, object detection, spine x-ray image, deep learning

Bone mineral density prediction from spine lateral X-ray images

Jinsung Jung
Department of Medicine
The Graduate School, Yonsei University
(Directed by Professor John A. Linton)

I. INTRODUCTION

On a global scale, over 200 million people are suffering from osteoporosis.[1] In the United States, 10 million people have osteoporosis, and over 40 million individuals have osteopenia that is at a pre-stage at risk of osteoporosis. The female to male ratio is 4 to 1. Osteoporosis is a reduction in bones' strength interpreted as an increased risk of fractures. Bone Mineral Density (BMD) refers to a bone density measurement, reflecting the strength of bones. One can be diagnosed with osteoporosis by BMD if the T-score is -2.5 or less at any skeletal site; clinical determinations are usually made in consideration of the lumbar spine and total hip.[2,3]

DEXA or DXA (Dual Energy X-ray Absorptiometry), is a highly accurate X-ray technique that has become the standard for measuring bone density.[4,5] In a clinical situation, a lumbar spine X-ray is routinely performed to evaluate low back pain in various conditions such as trauma, degenerative and neurological symptoms.[2] If a single spine X-ray image could estimate BMD, it could be promising as the standard test for BMD measurement since the cost and radiation exposure are both much lower than it is for DXA.

Various works have been performed to measure bone density through X-rays.[5–7] In 2002, Bouxsein et al. showed strong evidence that digital X-ray radiogrammetry (DXR)-BMD measurement, using a hand X-ray image for

automated BMD analysis, may be used as an alternative to peripheral single-photon X-ray absorptiometry (SXA) and DXA on the hand and, the wrist and for the prediction of fracture risk at the wrist, spine, and hip.[6]

Machine learning-based approaches have recently emerged as a viable option that can be applied to the metadata from the clinical environment. Recently, machine learning techniques have not only been used in image classification and pre-processing[8,9], but have also begun to be used with the aim of predicting BMD through image analysis.[10–14]

In 2019, Nam et al. applied a simple machine learning model to predict the T-score and the osteoporotic vertebrae by measuring the Hounsfield unit of conventional computed tomography (CT).[15] A total of 70 patients were enrolled and then underwent quantitative CT (QCT) and conventional lumbar CT for spine surgery within two months. The baseline T-score was gained from QCT, and it was matched with each vertebra of the conventional CT image. Of the total 198 vertebrae, 40 were the test set, and the classification accuracy precision was 92.5%. This study only applied some of the several CT images' cuts, and the cuts were similar to X-ray images, suggesting that the bone density can be predicted even with an X-ray image.

Unlike CT, the prediction based on X-ray image is a much more affordable option that can be widely used in the clinical environment. In 2020, Lee et al. demonstrated various machine learning algorithms for predicting BMD using simple spine X-ray image features extracted by three deep learning algorithms.[16] Here, BMD values were obtained from DEXA taken on the same day as the spine X-ray. From the feature extraction of the fourth lumbar body center point (150x150 pixel-sized) in 334 lateral spine X-ray images, machine learning was performed to classify those into normal and at-risk (osteoporosis and osteopenia) groups. Among the various models applied, the highest classification accuracy was 71%. However, this study had a small dataset size and the limitation that there was no spine detection process in the X-ray, and that

bone density was deduced with only one level of the spine.

Herein, the two-step method of vertebral body detection and BMD prediction is applied with a single X-ray image. One-stage and two-stage object detection methods are utilized for vertebral body detection. The basic CNN model (denoted as C5-GAP-FC3) and ResNet-18[17] models are used to predict BMD.

II. MATERIALS AND METHODS

1. Overview

The overview of step 1 and step 2 is shown in Figure 1 and Figure 2.

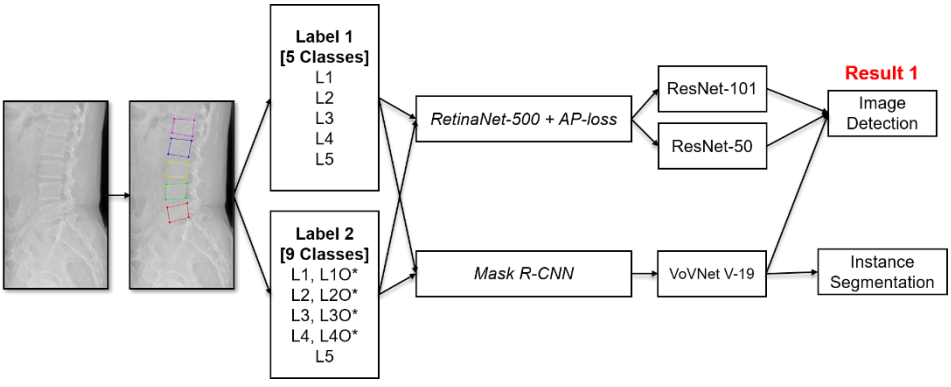

Figure 1. Overview of Step 1: vertebral body detection in the X-ray image (*The asterisk refers to Osteoporosis.)

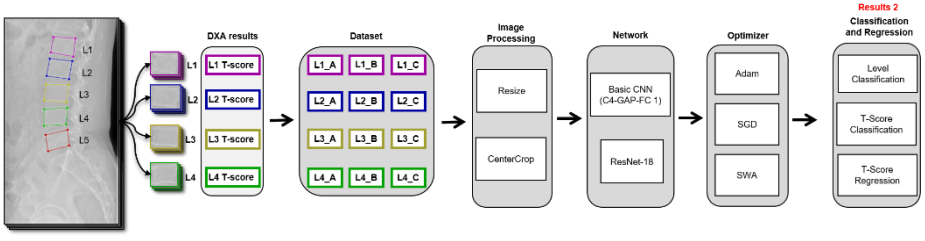

Figure 2. Overview of Step 2: BMD prediction

2. Dataset

A. Data acquisition

DXA and spine lateral X-ray data were collected from patients aged 50 and over who visited Severance Hospital, Seoul, Republic of Korea, between 2005 and 2020. Patients who had done both studies within a month were included to exclude possible changes between the two studies, such as trauma leading to fracture. Out of the total collected data, 1000 X-ray images were randomly

extracted. This study was conducted under the ethical principles of the Declaration of Helsinki.[18] It was approved by the Institutional Review Board (IRB) of Severance Hospital, Seoul, Republic of Korea (IRB approval No. 4-2020-0167).

B. Image annotation

The X-ray images were manually annotated using the LabelMe® program.[19] Annotation data comprised 20 points by annotating four points for each lumbar level. All the X-ray images were manually labeled by the author under the guidance of the other physician and two other non-clinicians who are specialists in computer vision. All the annotated images were blinded, randomly divided, and reviewed to validate the quality of the dataset.

C. Labeling

Two labeling methods were implemented for annotated data. One method labeled the lumbar spine into five parts, from L1 to L5 by the spine level. The other method divided the lumbar spine into a total of nine classes; from L1 to L4, each class has two states according to whether it is an osteoporotic spine (T-score values -2.5 or less) or not. L5 has only a single class, as the osteoporosis in L5 cannot be diagnosed using DXA. The T-score values were recorded by checking the values obtained from the DXA image of the patient.

D. Cropped images of the vertebral body

To predict the BMD, an image dataset was obtained by cropping X-ray images with annotation data. These 4,000 cropped images of vertebral level are matched with the T-scores obtained from DXA, during which 240 images without the T-score were excluded from the experiment (**Figure 3**).

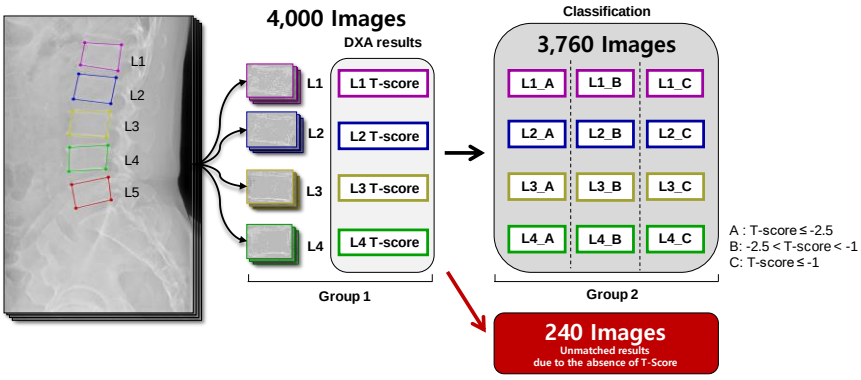

Figure 3. Cropping process of the vertebral body from annotated images (A: Osteoporosis group, B: Osteopenia group, C: Normal group.)

The 3,760 data obtained above were grouped in three ways according to the T-score values by the following criteria: T-score -2.5 or below is the *osteoporosis group*. Those below T-score -2.5 and T-score -1 are the *osteopenia group*. T-score -1 and above were assigned to the *normal group*.

E. Dataset composition

In RetinaNet-500 with the average precision (AP) loss model of step 1, the ratio of the train set to the test set was seven to three. In the Mask R-CNN model, the ratio of train and validation set to test set was seven to three.

In the classification and regression method of step 2, the ratio of the train set to the test set was seven to three.

3. Detection methods

Two different detection methods were employed. RetinaNet-500[20] with AP-loss model (**Figure 4**)[21] was used as a one-stage object detection method, and the Mask R-CNN model (**Figure 5**)[22] was used as a two-stage detection method. The backbone networks ResNet-50 and ResNet-101[17] were applied to RetinaNet-500 with the AP-loss model. VoVNet V-19 FPN3x[23] was used for Mask R-CNN.

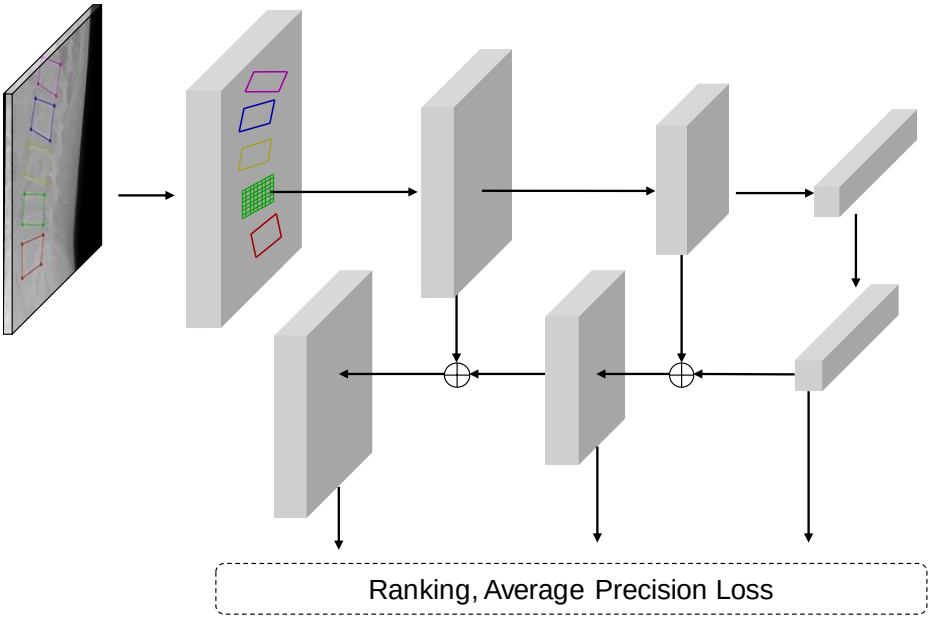

Figure 4. RetinaNet-500 with AP-loss model structure

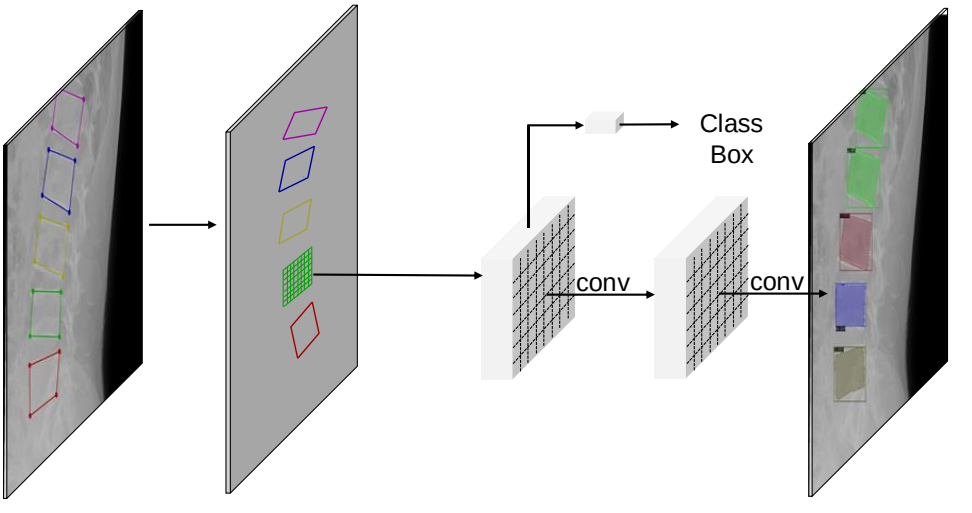

Figure 5. Mask R-CNN model structure

4. Classification and regression methods

To solve the BMD prediction problem, two methods were exploited: classification and regression based on the T-score. There are two main classification methods, one classifying the T-score and the other classifying the spine level of L1 to L5. The T-score classification was also classified three different ways. One was classified into three groups: osteoporosis group, the osteopenia group, and the normal group. Another was divided into the osteoporosis group or non-osteoporosis group. The remaining one is a normal group or abnormal group.

The Basic CNN (C5-GAP-FC3) and ResNet-18 models were used to predict BMD. The Basic CNN consists of five convolutional layers, a global average pooling, and three fully connected (FC) layers. FC layers include three types: T-score classification, Spine level classification, and T-score classification (**Figure 6**). Cross entropy loss was put in classification layers and mean square error (MSE) loss was reflected in the regression layer.

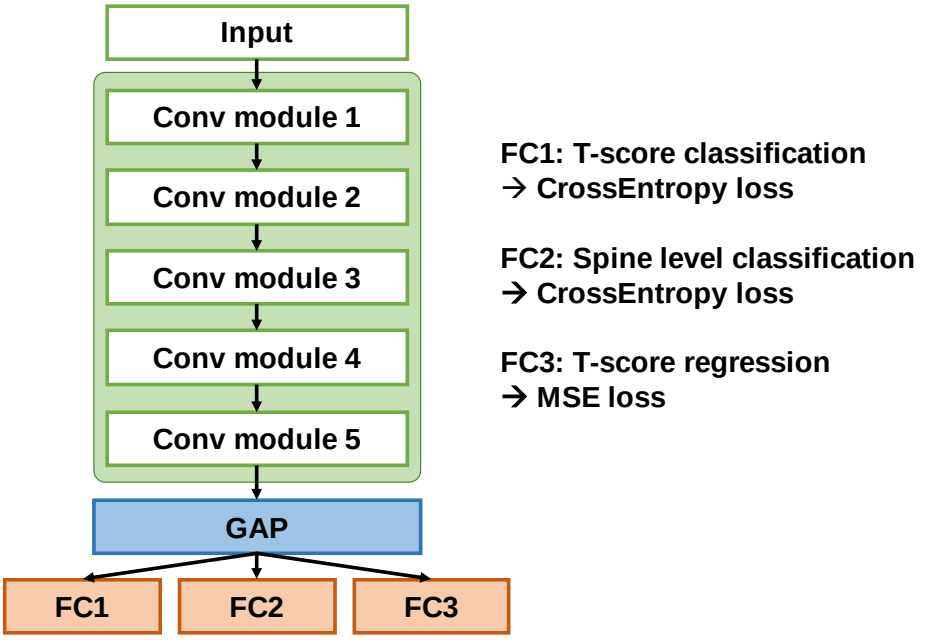

Figure 6. Basic CNN (denoted as C5-GAP-FC3) structure

5. Image processing

Image processing such as resizing and center cropping was attempted. Considering the image size of each vertebral body, 64 x 64 and 128 x 128 were employed, as shown in Figure 7 and Figure 8.

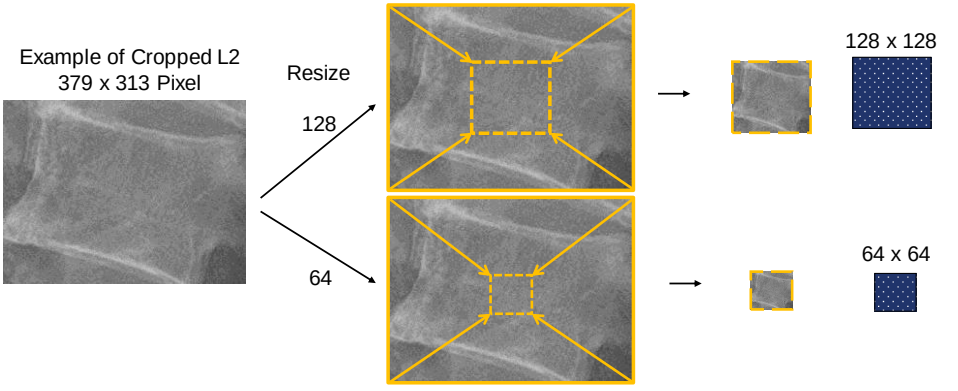

Figure 7. Image processing: Resize 64 and 128.

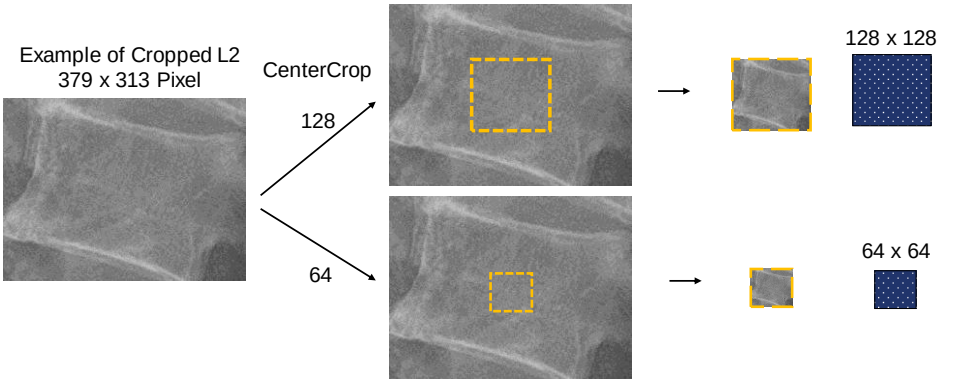

Figure 8. Image processing: CenterCrop 64 and 128.

6. Image Augmentation

Various image augmentation methods were utilized to supplement small datasets, such as random rotation, random horizontal flip, and random vertical flip.[24,25]

7. Optimizer

Adam[26], stochastic gradient descent (SGD)[27], and stochastic weight averaging (SWA)[28] optimizers were attempted to train classification and regression networks.

8. Computing environment

All processes were performed with a personal laptop equipped with Intel® Pentium® CPU G4600 @3.60GHZ, and GeForce GTX 1050 GPU.

III. RESULTS

1. Vertebral body detection

The vertebral detection results for the X-rays are shown in Table 1. In the one stage detection method, using a RetinaNet-500 with AP-loss model with backbone network ResNet-101, the AP of the bounding box was 73.1, which was higher than when ResNet-50 was used. In the two-stage detection method, using a Mask R-CNN with backbone network VoVNet V-19 FPN 3x, the AP of the bounding box was 59.3. Due to the model's characteristics, the instance segmentation was also obtained at the same time, and its AP was 60.83.

Table 1. Evaluation results for the bounding box of 5 classes

Method	Backbone network	AP	AP50	AP75
RetinaNet-500 + AP-loss	ResNet-101	73.1	96.8	92.0
	ResNet-50	72.2	97.1	91.4
Mask R-CNN	VoVNet V-19 FPN 3x	59.3	76.6	74.6

Table 2. Evaluation results for instance segmentation of 5 classes

Method	Backbone network	AP	AP50	AP75
Mask R-CNN	VoVNet V-19 FPN 3x	60.8	74.9	74.9

The results of the nine-class detection for X-rays are shown in Table 3 and Table 4. The AP of the bounding box in the RetinaNet-500 with the AP-loss model was 41.2.

The AP of the bounding box and the instance segmentation in the Mask R-CNN model were 22.0 and 23.0, respectively.

Table 3. Evaluation results for the bounding box of 9 classes

Method	Backbone network	AP	AP50	AP75
RetinaNet-500 + AP-loss	ResNet-101	41.2	56.6	52.4
Mask R-CNN	VoVNet V-19 FPN 3x	22.0	29.2	27.8

Table 4. Evaluation results for instance segmentation of 9 classes

Method	Backbone network	AP	AP50	AP75
Mask R-CNN	VoVNet V-19 FPN 3x	23.0	29.3	27.8

Both X-rays of patients with no previous surgical history (**Figure 9 and Figure 10**) and X-rays of patients who had undergone various surgeries and procedures (**Figure 11 and Figure 12**) were observed in both models. Visualized results for the method of detecting an osteoporotic or non-osteoporotic vertebral body by RetinaNet-500 with the AP-loss model are shown in Figure 13.

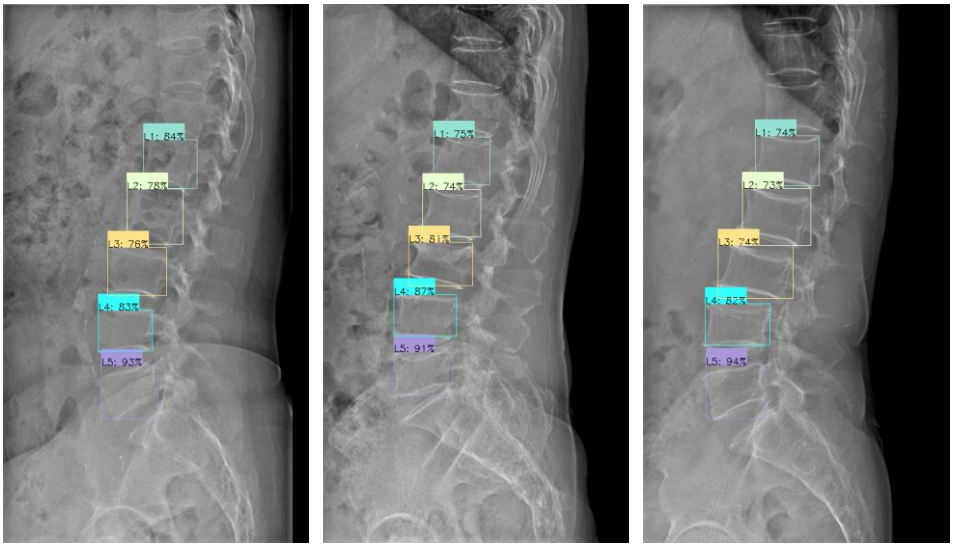

Figure 9. Results of vertebral body detection by RetinaNet-500 with AP-loss model with X-rays of patients with no previous surgical history.

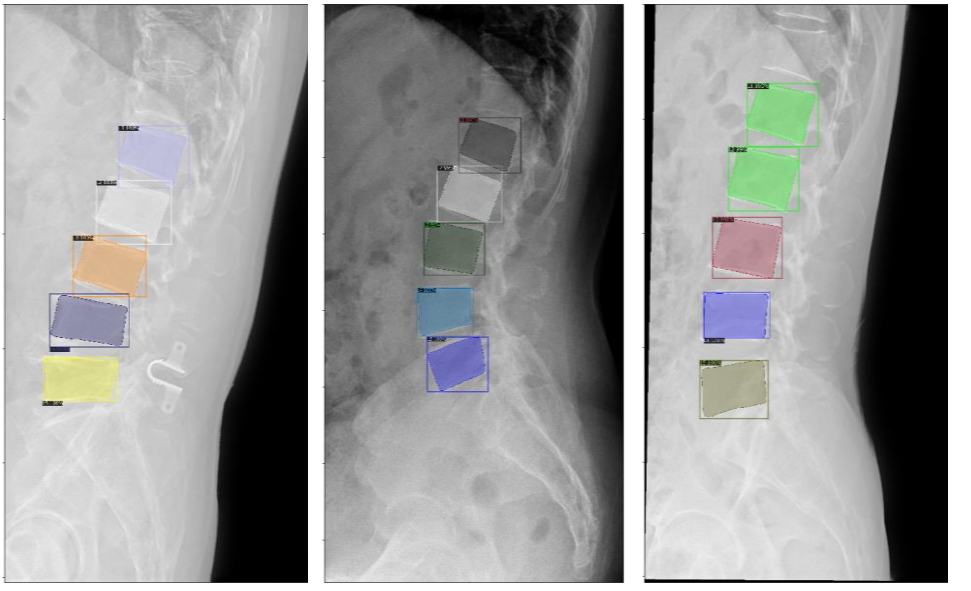

Figure 10. Results of vertebral body detection and instance segmentation by the Mask R-CNN model with X-rays of patients with no previous surgical history.

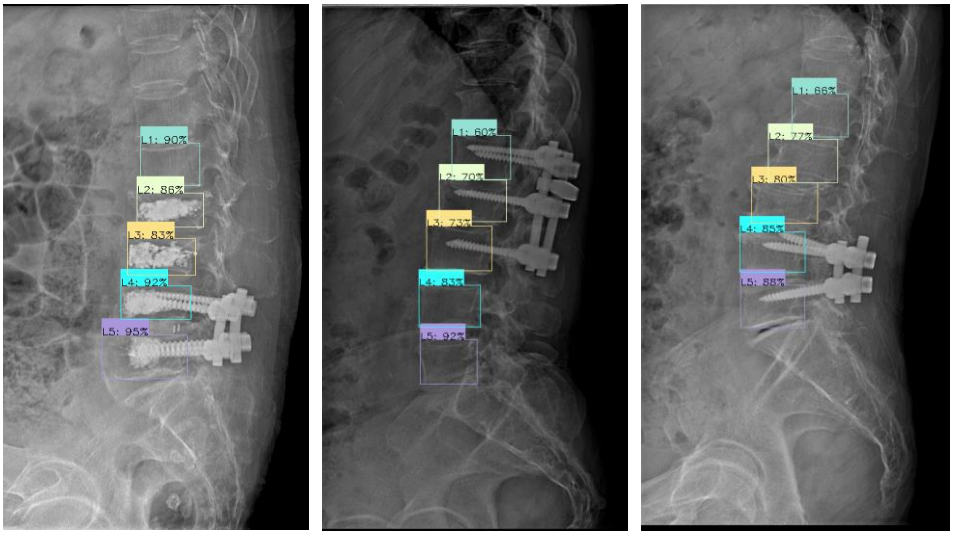

Figure 11. Results of vertebral body detection by RetinaNet-500 with the AP-loss model with X-rays of patients who had undergone various surgeries or procedures.

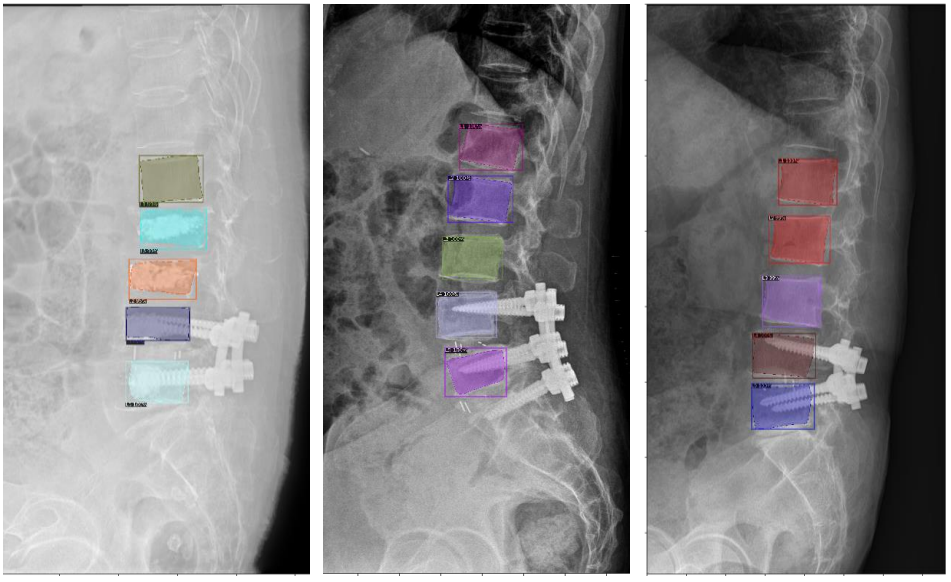

Figure 12. Results of vertebral body detection and instance segmentation by the Mask R-CNN model with X-rays of patients who had undergone various surgeries or procedures.

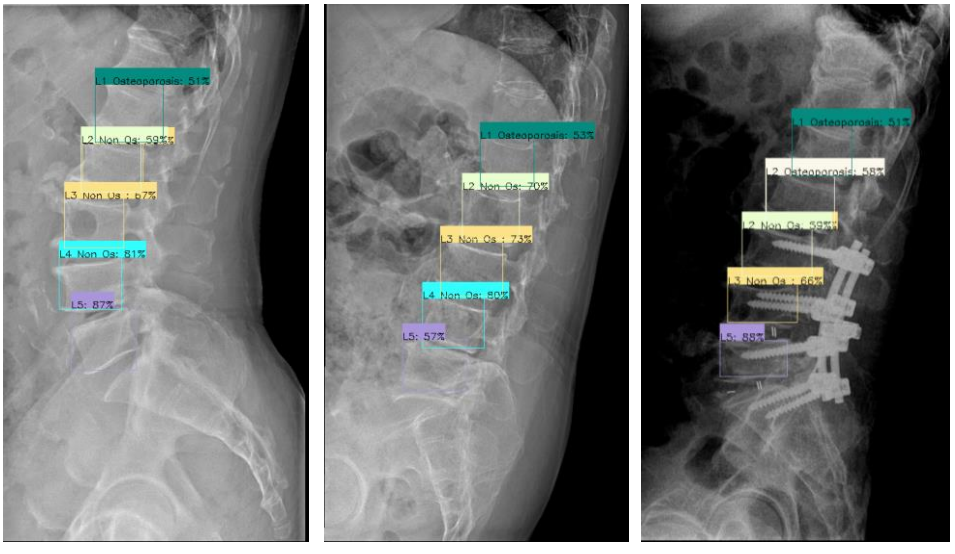

Figure 13. Results of osteoporotic or non-osteoporotic vertebral body detection by RetinaNet-500 with AP-loss model with X-rays.

2. BMD prediction

Table 5 shows the evaluation results of the BMD prediction. Between two models, the Basic CNN outperformed the ResNet-18. Various image processing and augmentation methods as well as optimizers were tested, and the best outcomes of these are demonstrated as follows.

Table 5. Evaluation results of the classification and regression models

Method	**T-score classification accuracy**	**Spine level classification accuracy**	**T-score regression error**
Osteoporosis / Non-osteoporosis	77.66	70.48	6.30
At-risk / Normal	68.53	68.09	5.69
Osteoporosis / Osteopenia / Normal	53.19	70.12	5.64

Among the image processing methods, resizing and center cropping, the resize 64x64 method displayed the best performance. Various augmentation methods were attempted; all augmentation methods somewhat reduced spinal level classification accuracy and negatively affected T-score classification accuracy, resulting in poor performance. Therefore, the performance of the method without the image augmentation was the finest. The performance of the Adam optimizer was the top-rated one among Adam, SGD, and SWA optimizers.

Among the three types of classification methods, the accuracy of distinguishing whether it was osteoporosis or not was the highest in both the T-score and the spine level classification, which were 77.66 and 70.48, respectively. The T-score regression method did not, in any case, produce significant results. In all the experiments, a mean error was found to be higher than 5.

IV. DISCUSSION

The AP of 73.1 in vertebral body detection using ResNet-500 with the AP-loss model refers to a relatively inaccurate detection at some level. Improperly detected images of the lumbar spine are visualized in Figure 14. In this case, the levels can be mismatched, while in some cases some of the levels cannot be detected. This degraded performance seems to result from the comparable shapes among the vertebral bodies and the inherent blurriness of some X-ray images.

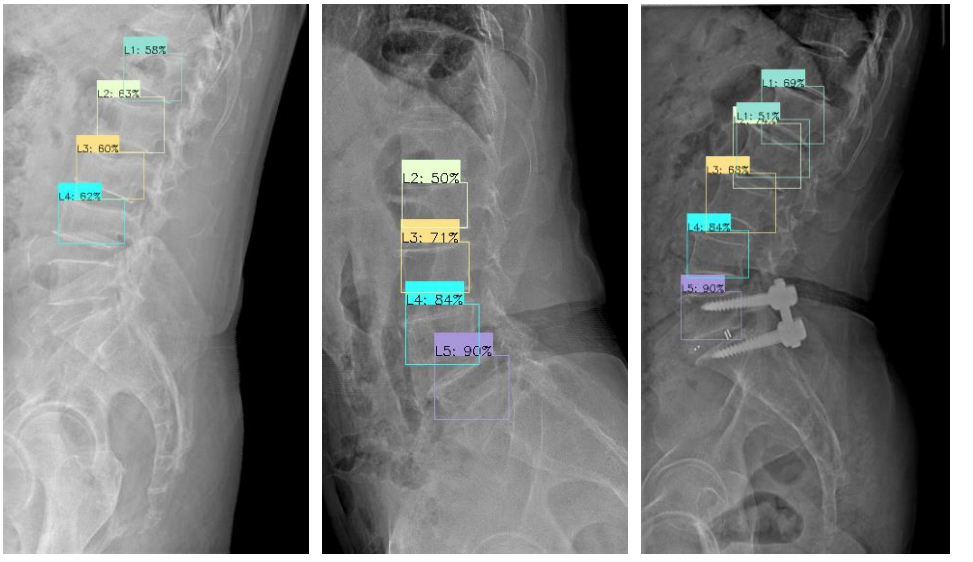

Figure 14. Inaccurate results of vertebral body detection by RetinaNet-500 with AP-loss model with X-rays.

In the radiographic absorptiometry, the X-ray image was standardized, including the aluminum wedge, when the X-ray was taken.[29] On the other hand, the X-ray images used in this study were taken routinely in clinical situations, so no standard element was available. Performance could have been improved by employing standardized images. Also, the T-score is a relative value; it would have produced more significant results if the problems were solved using the absolute value such as bone density.

Lee et al. selected the center crop image processing method.[16] In this study, however, the resizing method performed better than the center crop method. BMD measurement methods, such as DXA, excludes the cortical bone part because it would produce an error.[30] However, this study showed the possibility that the method which includes the cortical bone would improve BMD predicting performance.

Various image augmentation methods were attempted, but the performance was better without the application of those methods. One of the main reasons is that the process of augmentation reduced the performance in spine-level classification. In the case of osteoporosis classification, osteopenia, and normal, the spine level classification without image augmentation performance was 70.12, whereas it was 58.07 for random rotation (-5 to +5 degrees), 66.22 for the random vertical flip, and 64.63 for the random horizontal flip. It should also be mentioned that properly standardized X-ray images could have improved image augmentation performance.

Even with various attempts, the outcome of the T-score regression was not significant. If the problem was to find a BMD rather than a T-score value, a significant result might have been obtained. This study was conducted in two steps; if the lumbar spine detection and BMD prediction accuracy are sufficiently increased in the future, it is expected that the two stages will be processed at once.

The study experimented with a small dataset of 1,000 X-ray images. A larger dataset for model training could help improve performance.

V. CONCLUSION

This study was carried out in two stages: lumbar spine detection and BMD prediction. In the first step, the detection performance was encouraging as the AP of 73.1 was obtained with only the patient's X-ray. Moreover, X-rays of patients with diverse medical histories, regardless of gender, showed a decent performance. This study contributes to the detection of lumbar spines with only spine X-rays. It was showed that a one stage detection model might be more suitable than a two-stage detection model. Although the AP of 43.2 was low when the nine-class classification was applied, the approach itself is of considerable importance and worth being further pursued as it opens up the possibility of BMD prediction only with the detection tool.

For the second step, among the three types of BMD prediction methods, the accuracy of the classification between osteoporosis and non-osteoporosis was the highest in both the T-score and the spine level, 77.66 and 70.48, respectively. To the best of the author's knowledge, this study is the first approach to BMD prediction to use all the lumbar spines of L1 to L4.

REFERENCES

1. Sozen T, Ozisik L, Calik Basaran N. An overview and management of osteoporosis. *Eur J Rheumatol*. 2017;4(1):46-56. doi:10.5152/eurjrheum.2016.048
2. Cosman F, de Beur SJ, LeBoff MS, et al. Clinician's Guide to Prevention and Treatment of Osteoporosis. *Osteoporos Int*. 2014;25(10):2359-2381. doi:10.1007/s00198-014-2794-2
3. Morin SN, Lix LM, Leslie WD. The importance of previous fracture site on osteoporosis diagnosis and incident fractures in women. *J Bone Miner Res*. 2014;29(7):1675-1680. doi:10.1002/jbmr.2204
4. Haidekker MA, Stevens HY, Frangos JA. Computerized methods for X-ray-based small bone densitometry. *Comput Methods Programs Biomed*. 2004;73(1):35-42. doi:10.1016/S0169-2607(02)00164-5
5. Nazia Fathima SM, Tamilselvi R, Beham MP. Role of X-Rays in Assessment of Bone Mineral Density—A Review. *Innov Electron Commun Eng*. 2019;65:51-59. doi:10.1007/978-981-13-3765-9_6
6. Bouxsein ML, Palermo L, Yeung C, Black DM. Digital X-ray radiogrammetry predicts hip, wrist and vertebral fracture risk in elderly women: A prospective analysis from the study of osteoporotic fractures. *Osteoporos Int*. 2002;13(5):358-365. doi:10.1007/s001980200040
7. Choplin RH, Lenchik L, Wuertzer S. A Practical Approach to Interpretation of Dual-Energy X-ray Absorptiometry (DXA) for Assessment of Bone Density. *Curr Radiol Rep*. 2014;2(6). doi:10.1007/s40134-014-0048-x
8. Sohail A, Younas M, Bhatti Y, Li Z, Tunç S, Abid M. Analysis of trabecular bone mechanics using machine learning. *Evol Bioinforma*. 2019;15. doi:10.1177/1176934318825084
9. Mall PK, Singh PK, Yadav D. GLCM based feature extraction and

medical X-RAY image classification using machine learning techniques. *2019 IEEE Conf Inf Commun Technol*. Published online 2019:1-6. doi:10.1109/CICT48419.2019.9066263

10. Kruse C, Eiken P, Vestergaard P. Machine Learning Principles Can Improve Hip Fracture Prediction. *Calcif Tissue Int*. 2017;100(4):348-360. doi:10.1007/s00223-017-0238-7
11. Zhang B, Yu K, Ning Z, et al. Deep learning of lumbar spine X-ray for osteopenia and osteoporosis screening: A multicenter retrospective cohort study. *Bone*. 2020;140(July):115561. doi:10.1016/j.bone.2020.115561
12. Hemke R, Buckless CG, Tsao A, Wang B, Torriani M. Deep learning for automated segmentation of pelvic muscles, fat, and bone from CT studies for body composition assessment. *Skeletal Radiol*. 2020;49(3):387-395. doi:10.1007/s00256-019-03289-8
13. Yasaka K, Akai H, Kunimatsu A, Kiryu S, Abe O. Prediction of bone mineral density from computed tomography: application of deep learning with a convolutional neural network. *Eur Radiol*. 2020;30(6):3549-3557. doi:10.1007/s00330-020-06677-0
14. Rastegar S, Vaziri M, Qasempour Y, et al. Radiomics for classification of bone mineral loss: A machine learning study. *Diagn Interv Imaging*. 2020;101(9):599-610. doi:10.1016/j.diii.2020.01.008
15. Nam KH, Seo I, Kim DH, Lee JI, Choi BK, Han IH. Machine learning model to predict osteoporotic spine with hounsfield units on lumbar computed tomography. *J Korean Neurosurg Soc*. 2019;62(4):442-449. doi:10.3340/jkns.2018.0178
16. Lee S, Choe EK, Kang HY, Yoon JW, Kim HS. The exploration of feature extraction and machine learning for predicting bone density from simple spine X-ray images in a Korean population. *Skeletal Radiol*. 2020;49(4):613-618. doi:10.1007/s00256-019-03342-6

17. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2016;2016-Decem:770-778. doi:10.1109/CVPR.2016.90
18. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *J Am Coll Dent*. 2014;81(3):14-18. doi:10.1093/acprof:oso/9780199241323.003.0025
19. Russell BC, Torralba A, Murphy KP, Freeman WT. LabelMe: A database and web-based tool for image annotation. *Int J Comput Vis*. 2008;77(1-3):157-173. doi:10.1007/s11263-007-0090-8
20. Lin TY, Goyal P, Girshick R, He K, Dollár P. Focal Loss for Dense Object Detection. *IEEE Trans Pattern Anal Mach Intell*. 2020;42(2):318-327. doi:10.1109/TPAMI.2018.2858826
21. Chen K, Li J, Lin W, et al. Towards accurate one-stage object detection with ap-loss. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2019;2019-June:5114-5122. doi:10.1109/CVPR.2019.00526
22. He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. *IEEE Trans Pattern Anal Mach Intell*. 2020;42(2):386-397. doi:10.1109/TPAMI.2018.2844175
23. Lee Y, Hwang JW, Lee S, Bae Y, Park J. An energy and GPU-computation efficient backbone network for real-time object detection. *IEEE Comput Soc Conf Comput Vis Pattern Recognit Work*. 2019;2019-June:752-760. doi:10.1109/CVPRW.2019.00103
24. Frid-Adar M, Diamant I, Klang E, Amitai M, Goldberger J, Greenspan H. GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification. *Neurocomputing*. 2018;321:321-331. doi:10.1016/j.neucom.2018.09.013
25. Shorten C, Khoshgoftaar TM. A survey on Image Data Augmentation for Deep Learning. *J Big Data*. 2019;6(1). doi:10.1186/s40537-019-

0197-0

26. Zhang Z. Improved Adam Optimizer for Deep Neural Networks. *2018 IEEE/ACM 26th Int Symp Qual Serv IWQoS 2018*. Published online 2019:1-2. doi:10.1109/IWQoS.2018.8624183

27. Haras M, Lacatena V, Bah TM, et al. Fabrication of Thin-Film Silicon Membranes with Phononic Crystals for Thermal Conductivity Measurements. *IEEE Electron Device Lett*. 2016;37(10):1358-1361. doi:10.1109/LED.2016.2600590

28. Gupta V, Serrano SA, DeCoste D. Stochastic Weight Averaging in Parallel: Large-Batch Training that Generalizes Well. Published online 2020:1-12. http://arxiv.org/abs/2001.02312

29. Yates AJ, Ross PD, Lydick E, Epstein RS. Radiographic absorptiometry in the diagnosis of osteoporosis. *Am J Med*. 1995;98(2 SUPPL. 1):41S-47S. doi:10.1016/S0002-9343(05)80045-2

30. Lundeen GA, Knecht SL, Vajda EG, Bloebaum RD, Hofmann AA. The contribution of cortical and cancellous bone to dual-energy X-ray absorptiometry measurements in the female proximal femur. *Osteoporos Int*. 2001;12(3):192-198. doi:10.1007/s001980170129

ABSTRACT (IN KOREAN)

척추 측면 엑스레이 이미지에서 골밀도 예측

<지도교수 인요한>

연세대학교 대학원 의학과

정진성

목적: 본 연구는 엑스레이 이미지만을 가지고 환자 척추의 골밀도를 예측하고자 하였다.

재료와 방법: 2005년에서 2020년 사이 세브란스 병원에 방문한 50세 이상의 환자의 척추 측면 엑스레이와 골밀도 검사 데이터들을 이용하였다. 1,000 장의 엑스레이 이미지가 무작위로 추출 및 가공되었다. 첫번째 단계인 요추 검출에서는 한 단계 검출 방법과 두 단계 검출방법 모두를 실험하였다. 두번째 단계인 골밀도 예측에서는 엑스레이 사진과 T-score 값을 결합시킨 후 Basic CNN과 ResNet-18 두 모델에 실험 적용하였다.

결과: 요추 검출의 AP 값은 RetinaNet-500 with AP-loss 모델에서 가장 높았으며 그 값은 73.1 이였다. 다양한 분류 방법들 중 골다공증의 여부를 구별하는 문제에서의 T-score 분류와 척추 레벨 예측의 AP가 가장 높았으며 이는 각각 77.66 그리고 70.48 이였다.

결론: 본 연구는 척추의 엑스레이만을 이용하여 요추를 검출하였다는 데 의의가 있다. 또한, 요추의 L1 에서부터 L4를 포함하여 골밀도를 예측하였다.

핵심되는 말: 골밀도, 골다공증, 사물 검출, 척추 엑스레이, 딥러닝

PUBLICATION LIST

[1] Jung J, Linton JA. Object Detection and Instance Segmentation of Lumbar Vertebral Body from X-ray Images, *Kor J of Fam Med, under review.*