Basic Science

Osteoporotic vertebral compression fracture (OVCF) detection using artificial neural networks model based on the AO spine-DGOU osteoporotic fracture classification system

Wongthawat Liawrungrueang, MD[a], Sung Tan Cho, MD[b], Vit Kotheeranurak, MD[c,d], Khanathip Jitpakdee, MD[e], Pyeoungkee Kim, PhD[f], Peem Sarasombath, MD[g,*]

[a] Department of Orthopaedics, School of Medicine, University of Phayao, Phayao, Thailand
[b] Department of Orthopaedic Surgery, Seoul Seonam Hospital, South Korea
[c] Department of Orthopaedics, Faculty of Medicine, Chulalongkorn University, and King Chulalongkorn Memorial Hospital, Bangkok, Thailand
[d] Center of Excellence in Biomechanics and Innovative Spine Surgery, Chulalongkorn University, Bangkok, Thailand
[e] Department of Orthopedics, Queen Savang Vadhana Memorial Hospital, Sriracha, Chonburi, Thailand
[f] Department of Computer Engineering, Silla University, Busan, South Korea
[g] Department of Orthopaedics, Faculty of Medicine, Chiang Mai University, Chiang Mai, Thailand

ARTICLE INFO

ABSTRACT

Background: Osteoporotic Vertebral Compression Fracture (OVCF) substantially reduces a person's health-related quality of life. Computer Tomography (CT) scan is currently the standard for diagnosis of OVCF. The aim of this paper was to evaluate the OVCF detection potential of artificial neural networks (ANN).
Methods: Models of artificial intelligence based on deep learning hold promise for quickly and automatically identifying and visualizing OVCF. This study investigated the detection, classification, and grading of OVCF using deep artificial neural networks (ANN). Techniques: Annotation techniques were used to segregate the sagittal images of 1,050 OVCF CT pictures with symptomatic low back pain into 934 CT images for a training dataset (89%) and 116 CT images for a test dataset (11%). A radiologist tagged, cleaned, and annotated the training dataset. Disc deterioration was assessed in all lumbar discs using the AO Spine-DGOU Osteoporotic Fracture Classification System. The detection and grading of OVCF were trained using the deep learning ANN model. By putting an automatic model to the test for dataset grading, the outcomes of the ANN model training were confirmed.
Results: The sagittal lumbar CT training dataset included 5,010 OVCF from OF1, 1942 from OF2, 522 from OF3, 336 from OF4, and none from OF5. With overall 96.04% accuracy, the deep ANN model was able to identify and categorize lumbar OVCF.
Conclusions: The ANN model offers a rapid and effective way to classify lumbar OVCF by automatically and consistently evaluating routine CT scans using AO Spine-DGOU osteoporotic fracture classification system.

Introduction

Osteoporosis is characterized by a loss of microarchitecture and a decrease in bone mineral density throughout the body. This increases the likelihood of bone fractures. A chronic progressive bone illness that causes bone density and quality loss, osteoporosis affects 10 million Americans and 50 million worldwide [1,2]. Osteoporotic vertebral com-

pression fractures (OVCF) are the most common osteoporotic fractures and mainly occur at the thoracolumbar junction. Osteoporosis causes a loss in bone density, quality, and strength of the vertebral body, which can occur with or without external stresses [3]. OVCF significantly lowers health-related quality of life. Acute pain, balance and gait changes, muscular exhaustion, an increased risk of falls and fractures, and maybe greater mortality rates are connected with this disorder [4–6]. Anterior-
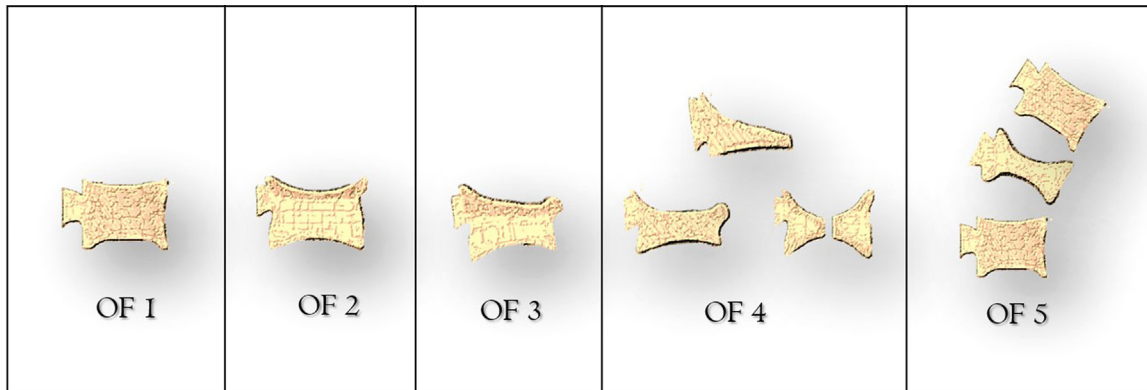
Fig. 1. The AO Spine-DGOU Osteoporotic Fracture Classification System: OF1, OF2, OF3, OF4 and OF5 [8].

posterior and lateral X-ray views are used to initially assess spine injuries, but radiographs have lower sensitivity and specificity compared to CT and MRI scans [7]. The recently common X-ray, CT, and MRI-based classification used by spine surgeons in the interpretation of OVCF is the osteoporotic fracture (OF) classification of the Spine Section of the German Society for Orthopaedics and Trauma (DGOU) for fracture morphology the modify to AO Spine-DGOU Osteoporotic Fracture Classification System [8], which classifies changes in stages from OF1 to OF5 [8,9] (Fig. 1).

The conventional methods, which are X-ray, CT, and MRI take time for radiologists, who are already overloaded with prognosis and patient categorization. A study found that 70% of diagnostic radiography errors are "missed findings". This error rate shows how tough the "detection process" is for humans [10]. In recent years, it has become more common that neural networks and deep learning techniques can be utilized to detect spinal fractures with conventional spinal radiography. Artificial neural networks (ANN), specifically the YOLOv8 model, are diagnostic tools that improve, evaluate, automate, and enhance fracture detection and classification. In addition to improving image quality, patient-centricity, imaging efficiency, and diagnostic precision. AI can also be used to eliminate image artifacts, harmonize images, and shorten the duration of imaging tests [11–14].

The purpose of this study was to assess the capability of ANN for OVCF detection based on DGOU classification. To our best knowledge, this is the first study to present the capability of ANN (YOLOv8 model) for OVCF detection.

Materials and methods

This study was conducted in accordance with the Declaration of Helsinki and with approval from the Ethics Committee and Institutional Review Board (Institutional Review Board (IRB) approval, IRB Number: ORT-2566-09473). Informed consent was not required due to the dataset [15] did not show the identity of the patient.

Study design

This study used 1,050 radiographic views of thoracolumbar and lumbar spine CT images from dataset repositories to create an ANNs model. All the images, including sagittal CT-scan radiographic images contained diagnostic information. The study retrieved the images from an online open-access dataset [15] which did not reveal the patient's identity.

The inclusion criteria were all spinal patients who were symptomatic with back pain with CT scans available from an online open-access OVCF CT image dataset [15]. Exclusion criteria were patients with a spine infection, congenital or other disease, low-quality CT imaging.

A dataset was divided into a training dataset, comprising 89% (934 images), and a test dataset, including 11% (116 images). The radiolo-

gist did a procedure of cleaning, labeling, and annotation on the training dataset. The data images were trained and detected using the deep ANN model, which was used in the deep learning technique. Through the use of an automated testing procedure, the trained ANN model's effectiveness was evaluated. Before determining the model's accuracy, 3 radiologists validated the testing dataset. A group of 3 researchers carried out manual annotation as part of the supervised learning technique. Through a voting procedure, the researchers came to a consensus regarding the best annotation imaging technique. Because there is no training data in the testing dataset, the model hasn't been exposed to the testing image before evaluation.

The dataset was "Osteoporotic vertebral fractures database" from Korez R et al. [15] distributed across the 2 spine regions: 70% males and 30% females in the lumbar spine dataset, and 30% males and 70% females in the thoracolumbar spine dataset. The mean age $\pm$ SD for the subjects is 46.0 ± 13.5 years for the lumbar spine dataset and 23.0 ± 3.7 years for the thoracolumbar spine dataset. Imaging specifications vary slightly between the datasets: the in-plane voxel size ranges from 0.282 to 0.791 mm for the lumbar spine and 0.313 to 0.361 mm for the thoracolumbar spine, while the cross-sectional thickness ranges from 0.725 to 1.530 mm for the lumbar spine and remains consistent at 1 mm for the thoracolumbar spine. Each vertebra in every image within the dataset is accompanied by a reference segmentation binary mask, facilitating precise evaluation and validation of automated frameworks designed for spine and vertebrae detection and segmentation.

Data augmentation

To avoid overfitting due to the short sample size, the Python library augmentor was used to randomly enhance all of the photos. The augmentor software package focuses on offering operations like flipping, rotating, zooming, scaling, cropping, and translating that are commonly used to provide image data for machine learning. This study used a horizontal flip, crop (zoom 0%–20%), rotation (between -15° and 15°), and shear (15° horizontal and 15° vertical) to the standard picture augmentation set. Segmented photos were randomized and scrambled for training. 934 training photos (about 89%) and 116 test images (11%) made up the final data set (Fig. 2).

Deep learning model and ANN algorithme

The ANN model base Deep learning algorithme was employed in the deep learning technique to train, identify, and predicted the OVCF grading. Nevertheless, this model is unable to identify the spinal image segment. By putting an automatic model to the test for dataset grading, the training ANN model was confirmed. For high performance object detection, we employed the YOLOv8 architecture. The main structure of the YOLO network architecture was showed in Fig. 3, which splits
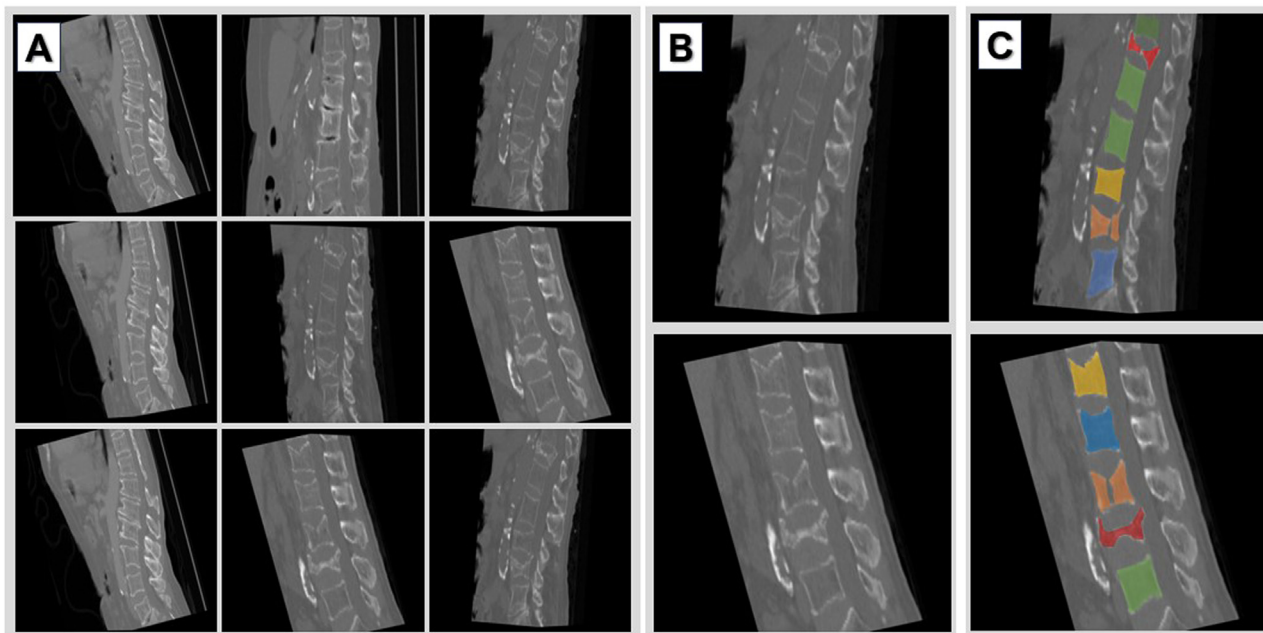
Fig. 2. Dataset preparation: sagittal CT of the thoracolumbar spine with image generation augmented by horizontal flip, cropping (zoom 0%–20%), rotation (between −15° and 15°), shear (±15° horizontal and ±15° vertical) (A), Image before (B) and after (C) annotation by a radiologistes and spin surgeon.
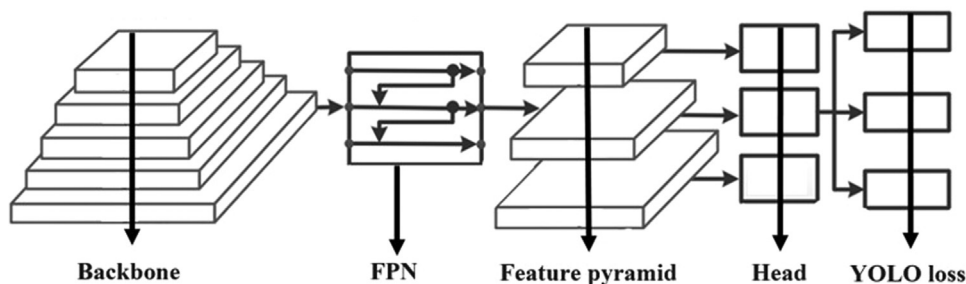

Fig. 3. The main structure of the YOLO network architecture.

an image into a grid system where each grid module detects things inside itself. In addition, real-time object detection based on data streams can be performed on the photos with very little processing overhead. The model was tested across 40 epochs since the accuracy of the model increased up to 100.

Model performance evaluation

Model performance was assessed using accuracy, precision, recall, F1-score, and mean average precision. Accuracy: The proportion of correctly classified instances among the total instances. It's calculated as (TP + TN) / (TP + TN + FP + FN), where TP = True Positives, TN = True Negatives, FP = False Positives, FN = False Negatives. Precision: The accuracy of the positive predictions. It's the ratio of correctly predicted positive observations to the total predicted positive observations. Precision is calculated as TP / (TP + FP). Recall (Sensitivity or True Positive Rate): This metric represents the model's ability to correctly identify all relevant instances. It's the ratio of correctly predicted positive observations to the all observations in the actual class. Recall is calculated as TP / (TP + FN). F1-Score: The harmonic mean of precision and recall. It provides a balance between precision and recall and is particularly useful when there is an uneven class distribution. The F1-score is calculated as 2 x (precision x recall) / (precision + recall). Mean Average Precision (mAP): Often used in information retrieval and object detection tasks, mAP calculates the average precision for each class and then takes the mean across all classes. It measures the quality of the model's ranking. All result were reported under staritic graft reported.

Statistical analysis

All data were collected and analyzed using R version 3.1.0 software (R Foundation for Statistical Computing, Vienna, Austria). A 1-way ANOVA was utilized to assess statistical differences in mean values between the groups. Post hoc testing with Bonferroni correction was performed when significant differences in mean values were observed between groups. A significance level of $p < .05$ was considered statistically significant for all tests.

Ethics

This study was approved by the Ethics Committee and Institutional Review Board of Faculty of Medicine, Chiang Mai University (Ethics Approval Number ORT-2566-09473).

Results

The 7,810 images that make up the sagittal lumbar CT training dataset were taken from an open resource dataset. The data was enhanced, and the models were trained and evaluated using the YOLOv8 model. Based on the OF categorization, the findings showed 5,010 for OF1, 942 for OF2, 522 for OF3, and 336 for OF4. The OF5 could not
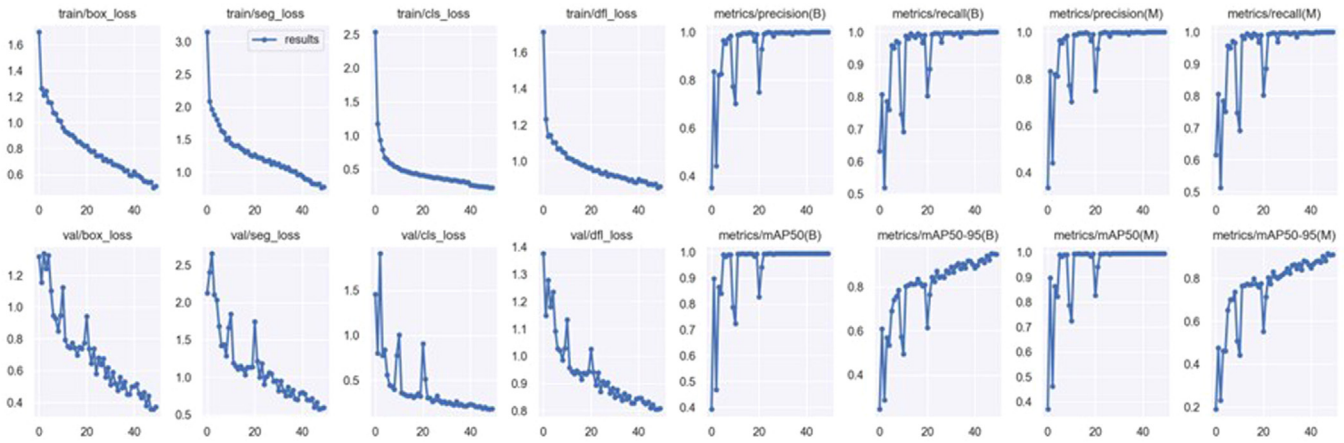
Fig. 4. Performance training analysis plot graphs of the ANN network model.
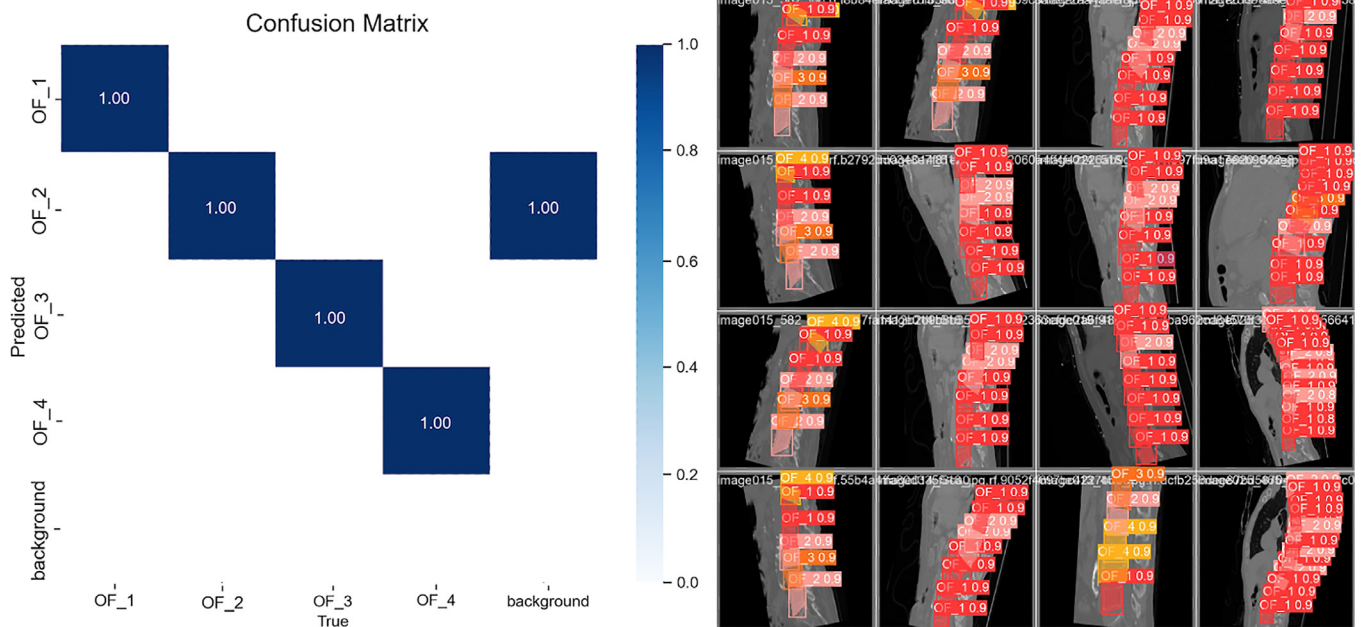

Fig. 5. Confusion matrix representation and accuracy of the training model and A column represents an instance of the actual class, whereas a row represents an instance of the predicted class.

locate this dataset. The accuracy of this model across all severity grades (OF 1–OF 4), a weighted average method is typically employed, taking into account the number of instances in each grade. Here are the accuracies for each grade: OF 1: 95.56%, OF 2: 95.34%, OF 3: 95.68%, and OF 4: 97.59%. The model's overall accuracy across grades OF 1 to OF 4 is 96.04%. This metric provides a comprehensive assessment of the model's performance in identifying osteoporotic vertebral compression fractures across varying severity levels. The bounding box shape data and the OVCF distribution of grades are shown. The training results for the dataset were displayed in Fig. 4. Losses can be classified into 3 categories: box_loss, obj_loss, and cls_loss. The discrepancy between the detected and target disk boxes is represented by a box_loss. The difference in object existence for every grid is called an obj_loss. The misclassification of class between the target and detected items is represented by cls_loss. As the epoch numbers increased, all loss values significantly dropped, indicating the effectiveness of the training. Given that the precision and recall numbers were nearly 1, they were considered good. The mean average accuracy, which is the most commonly utilized statistic, likewise met expectations. The average precision for each class is represented by AP, the average precision for all classes is shown by mAP,

and the average precision for a fixed intersection over union (Boussios, #141) of 0.5 is represented by mAP_0.5. The mAP values between IOU 0.5 and 0.95 intervals, with a step size of 0.05, are represented by the notation mAP_0.5:0.95. The training process was successfully completed for the dataset using the default training hyper-parameters, as evidenced by all of these graphs showing the results.

Every grade of OVCF had accuracy results above 0.98 when the confusion matrix for the ANN was applied to the collected data. It seems likely that the light blue objects in the "background" are those that ought to have been identified but weren't (Fig. 5).

The F1 scores for the confidence values, where each detected OVCF box's identification reliability is used to define confidence. The graph's lower left section was utilized to calculate the AP score. The F1 score displays the relative recognition performance between the precision and recall values and the recall-precision curve (Fig 6). Fig. 7 displays the accuracy vs. confidence values for each OVCF grade. A precision score of more than 0.85 indicates a reasonable degree of confidence.

Results of employing ANN with YOLOv8 for the automatic detection and grading of lumbar degenerative discs. Prediction based on the Osteoporotic Fracture Classification System (AO Spine-DGOU).
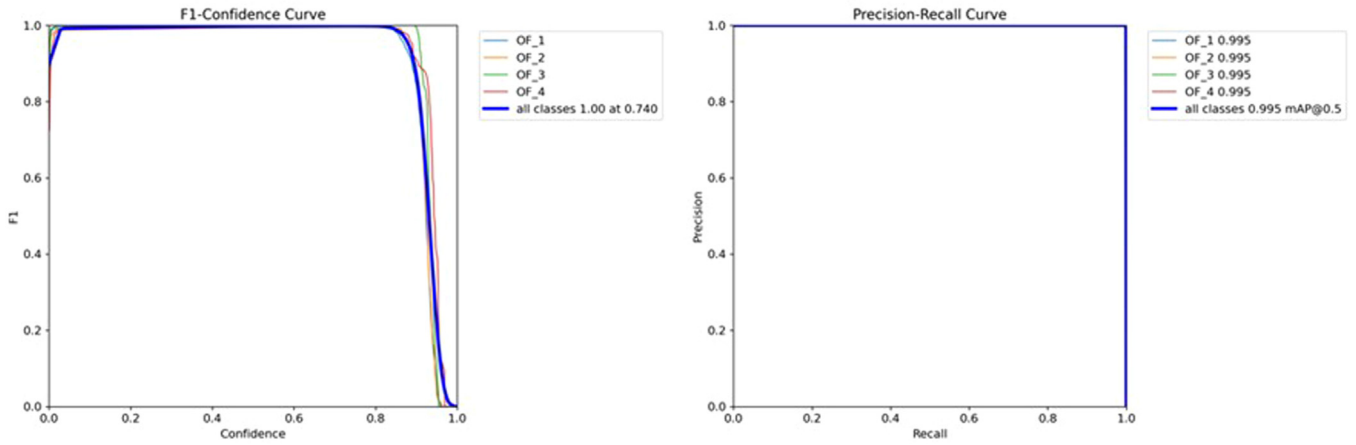
Fig. 6. F1 score graph showing the relationship between F1 and the confidence curve. Precision-Recall graph showing the relationship between recall and precision.
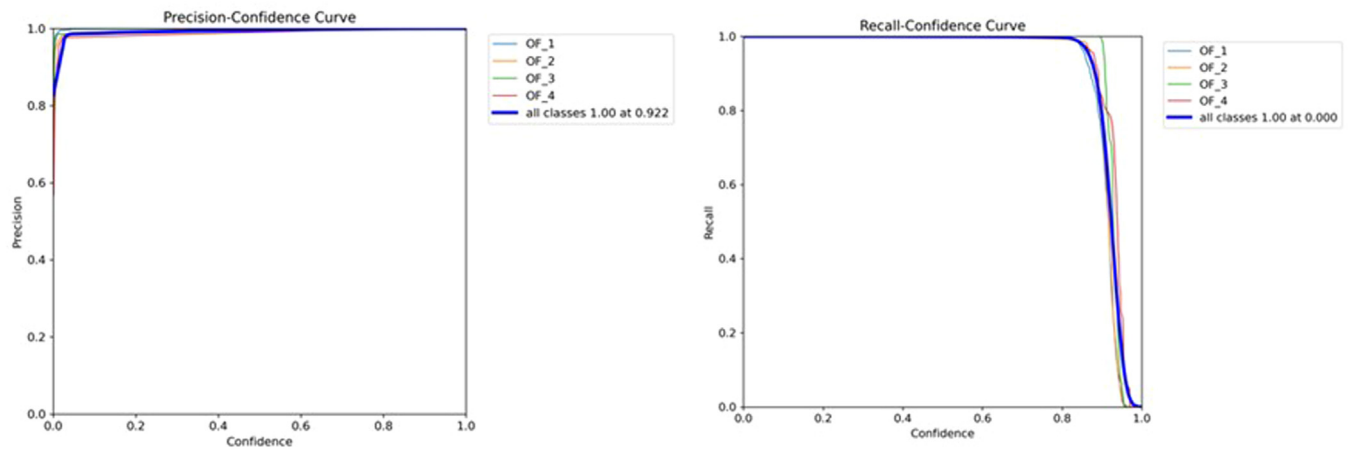

Fig. 7. Precision–confidence graph showing the relationship between precision and confidence. Recall–confidence graph showing the relationship between precision and confidence.
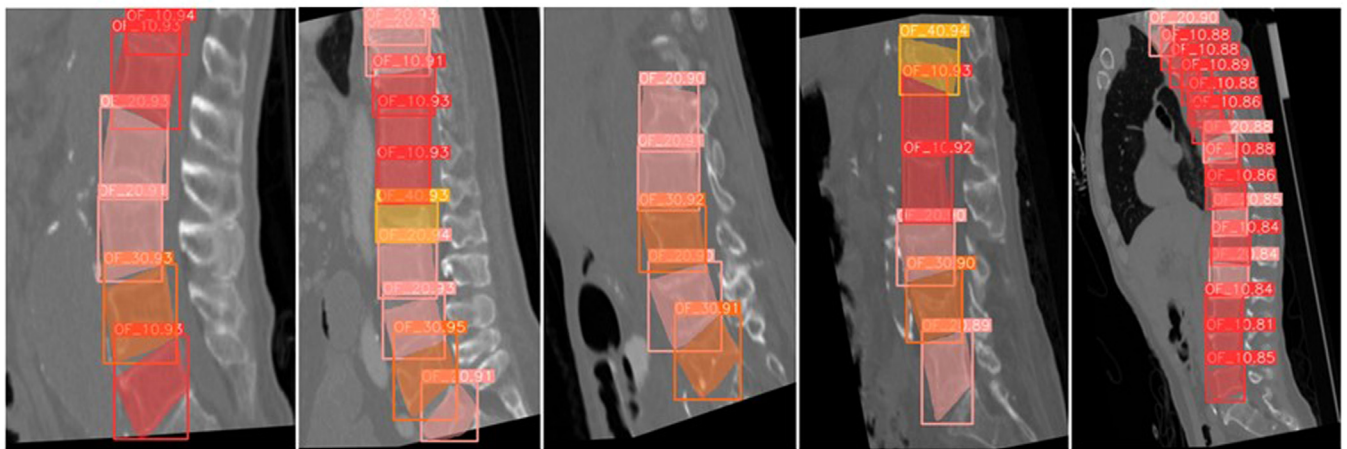
Fig. 8. Results of automatic detection and OVCF with YOLOv8. Prediction based on the AO Spine-DGOU Osteoporotic Fracture Classification System: The sagittal CT scan prediction patients with OF 1(Red), OF2 (Pink), OF3(Orange), and OF4 (Yellow).

The mean average precision and the grade are shown in the picture (Fig. 8).

The authors summary this deep artificial neural network and model performance for detected OVCF (Table 1) and a comparison between grades OF1 to OF5 (Table 2). The accuracy of the models ranged from 95.34% to 97.59% across grades OF1 to OF4, indicating high overall accuracy in detecting OVCF. However, the accuracy for grade OF5 was not applicable (N/A). The p-value of 0.976 suggests that there is no statistically significant difference in accuracy between the different grades. Sensitivity measures the proportion of true positive cases correctly identified by the model. It ranged from 95.85% to 97.35% across grades OF1 to OF4, indicating high sensitivity in detecting OVCF across these

Table 1

Summary deep artificial neural network in this study.

Aspect	Summary
Training dataset composition	- A total of 7,810 sagittal lumbar CT images was used, sourced from an open dataset repository. - The distribution of OVCF grades in the dataset was as follows: OF1 (5,010 images), OF2 (942 images), OF3 (522 images), OF4 (336 images), and OF5 (0 images found).
Model performance	- The deep artificial neural network (ANN) model was trained using the YOLOv8 architecture. - The model's overall accuracy across grades OF 1 to OF 4 is 96.04%. This metric provides a comprehensive assessment of the model's performance in identifying osteoporotic vertebral compression fractures across varying severity levels.
Training process	- Default hyperparameters were used for the training process. Significant improvements in loss values (box_loss, obj_loss, and cls_loss) were observed as the number of training epochs increased.
Evaluation metrics	- Various evaluation metrics including precision, recall, F1-score, mean average precision (mAP), and average precision (AP) were utilized. - Consistently high values of these metrics were observed across different OVCF grades, indicating the model's effectiveness.
Confusion matrix and F1 scores	- The confusion matrix analysis revealed high accuracy results (above 0.98) for each OVCF grade. F1 scores were calculated to assess the relative recognition performance between precision and recall values.
ANN with YOLOv8	- The combination of ANN with the YOLOv8 architecture enabled automatic detection and grading of lumbar degenerative discs based on the AO Spine-DGOU classification system. Mean average precision values were reported to evaluate the model's prediction accuracy.

Table 2

Summary model performance for detected osteoporotic vertebral compression fractures (OVCF) and compare between grading OF1 to OF5.

Parameters	OF 1	OF 2	OF 3	OF 4	OF 5	p-value
Accuracy (%)	95.56	95.34	95.68	97.59	N/A	.976
Sensitivity (%)	96.25	95.85	96.51	97.35	N/A	.665
Specificity (%)	94.91	96.53	94.28	97.86	N/A	.423
Overall Error	8.15	7.86	7.86	5.23	N/A	.568
Cohen's Kappa (κ)	0.855	0.871	0.825	0.864	N/A	.823
Matthews Correlation Coefficient	0.745	0.728	0.736	0.751	N/A	.765
Area Under ROC Curve (AUC)	0.951	0.967	0.953	0.981	N/A	.565

grades. Similar to accuracy, sensitivity for grade OF5 was not applicable (N/A). The p-value of .665 suggests that there is no statistically significant difference in sensitivity between the different grades. Specificity measures the proportion of true negative cases correctly identified by the model. It ranged from 94.28% to 97.86% across grades OF1 to OF4, indicating high specificity in correctly ruling out OVCF. Similar to accuracy and sensitivity, specificity for grade OF5 was not applicable (N/A). The p-value of .423 suggests that there is no statistically significant difference in specificity between the different grades. Overall error represents the overall misclassification rate of the models. It ranged from 7.86% to 8.15% across grades OF1 to OF4, indicating low overall error rates. The lower overall error rates suggest better model performance. Similar to other metrics, overall error for grade OF5 was not applicable (N/A). The p-value of .568 suggests that there is no statistically significant difference in overall error between the different grades. Cohen's Kappa measures the agreement between the model's predictions and the actual classification, corrected for the agreement occurring by chance. Values closer to 1 indicate higher agreement. Cohen's Kappa ranged from 0.825 to 0.871 across grades OF1 to OF4, indicating substantial to almost perfect agreement between the model's predictions and the actual classification. The Matthews Correlation Coefficient (MCC) measures the quality of binary classifications, with values ranging from −1 to 1. Higher values indicate better performance. MCC ranged from 0.728 to 0.751 across grades OF1 to OF4, indicating a moderate to substantial correlation between the predicted and actual classifications. The Area Under ROC Curve represents the model's ability to discriminate between positive and negative cases. Higher values (closer to 1) indicate better discrimination. AUC ranged from 0.951 to 0.981 across grades OF1 to OF4, indicating excellent discrimination ability of the models.

Discussion

The occurrence of traumatic spinal pathologies, demographic risk factors, risks, prevention, therapies, and probable repercussions of spine injuries are needed to distinguish acute OVCF injuries from other pathologies. Radiographic CT imaging is the best diagnostic tool. Grading and classification assist doctors in deciding on nonsurgical or surgical treatment and prognosis. The AO Spine-DGOU Osteoporotic Fracture Classification System is used by radiologists and orthopedists. The

accuracy of the ANN models with overall 96.04% across grades OF1 to OF4, indicating high overall accuracy in detecting and able to identify and categorize lumbar OVCF. However, the accuracy for grade OF5 was not applicable (N/A), possibly due to insufficient data or other factors. The p-value suggests that there is no statistically significant difference in accuracy, sensitivity, and specificity between the different grades.

Artificial intelligence (AI) technology is increasingly being used in medical imaging. Advances in machine learning (ML) and deep learning (DL) have opened up the potential of AI for practical applications. These include improving the accuracy of diagnoses made by radiologists, facilitating operations requiring high cognitive capacity, and many other applications [16]. Recently, Tomita et al. developed a system to detect osteoporotic vertebral fractures on CT exams which can be used for screening and prioritizing possible fracture cases [17]. Niu et al. also succeed in develop another deep-learning system for OVCF [18]. However, this study is the first to use ANN with the YOLO-v8 model to detect OVCF and classify the severity into 4 grades based on the OF grading system with high accuracy. As an optional clinical application, we employed this methodology to identify OVCF in our daily work to support the orthopedists at our facility.

As machine learning researchers and practitioners gain more experience, it will become easier to categorize challenges based on which solution technique makes the most sense. Once enough high-impact software systems grounded in physics, math, computer science, and engineering are routinely employed in clinics, it is anticipated that more of these systems will be approved. But the oversight of the physician is still crucial. Accuracy and model performance will rise when doctorate experience is used to AI performance. By streamlining healthcare processes, minimizing diagnostic errors, and optimizing resource allocation, this could be especially useful for physicians with limited experience diagnosing OVCF. The combination of medical imaging technology and artificial intelligence has resulted in the development of computer-aided detection systems that enhance the performance of lesion detection [19].

The authors utilized CT scans to diagnose OVCF due to the high-resolution images CT provides, which are essential for detecting subtle fractures and assessing spinal anatomy. CT scans are the gold standard in clinical practice for evaluating suspected vertebral fractures, making them an appropriate choice for developing a model aimed at standard

clinical application. The study employed the AO Spine-DGOU Osteoporotic Fracture Classification System, which relies on detailed bone imaging provided by CT scans, to categorize and grade the severity of OVCF accurately [17,20,21].

A limitation of this study is that we used an open assessment dataset of CTs of symptomatic low back pain that was imbalanced and unevenly distributed open evaluation dataset of CT scans with symptomatic low back pain, as seen by the prevalence of grades OF1 and OF2 and the rarity of grades OF3 and OF4 without OF5 [22]. Because they used an open assessment dataset of CTs that was made available, the data is anonymized and does not include information about participant identity, pertinent dates, recruitment periods, eligibility requirements, sources and methods of selection, exposure, follow-up, and data collection. The only person to classify the images using the OF Classification was a radiologist. Differences in how 2 radiologists categorize the same data set may occur. Furthermore, a more extensive dataset spanning multiple centers would have been ideal for further deep learning training, validation, and testing in the future. However, since this study was conducted entirely by computers without the influence of humans and the dataset is publicly available, bias cannot be introduced.

Another limitation is that CT scans cannot definitively differentiate between fractures caused by osteoporosis and those due to other pathological conditions such as tumors. MRI, which offers better soft tissue contrast, would be necessary to detect tumor involvement, and a biopsy might be required for confirmation [23]. The reliance on an open-access dataset without detailed clinical follow-up, such as MRI or biopsy results, means the study assumed the fractures were osteoporotic. This limitation underscores the need for a comprehensive diagnostic workflow involving further imaging and clinical evaluation to accurately diagnose the cause of vertebral fractures.

Conclusion

This work describes a novel approach to automatically detect and classify lumbar OVCF grading, which advances image-guided evaluation and computer-aided diagnosis of spine illnesses. This study shows that OVCF of the spine may be reliably classified and graded in real-time using a deep learning model, offering a rapid and effective procedure. This approach is useful for identifying diseases and anomalies as well as for applications in neurology and orthopedics. It is advised that the model be developed further utilizing a bigger, multicenter data set.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Wright NC, Looker AC, Saag KG, et al. The recent prevalence of osteoporosis and low bone mass in the united states based on bone mineral density at the femoral neck or lumbar spine. J Bone Mineral Res 2014;29:2520–6. doi:10.1002/jbmr.2269.

[2] Johnell O, Kanis JA. An estimate of the worldwide prevalence and disability associated with osteoporotic fractures. Osteoporos Int 2006;17:1726–33. doi:10.1007/s00198-006-0172-4.

[3] Rachner TD, Khosla S, Hofbauer LC. Osteoporosis: now and the future. Lancet 2011;377:1276–87. doi:10.1016/S0140-6736(10)62349-5.

[4] Son HJ, Park S-J, Kim J-K, Park J-S. Mortality risk after the first occurrence of osteoporotic vertebral compression fractures in the general population: a nationwide cohort study. PLoS One 2023;18:e0291561. doi:10.1371/journal.pone.0291561.

[5] Lee DG, Bae JH. Fatty infiltration of the multifidus muscle independently increases osteoporotic vertebral compression fracture risk. BMC Musculoskelet Disord 2023;24:508. doi:10.1186/s12891-023-06640-2.

[6] Gold LS, Suri P, O'Reilly MK, Kallmes DF, Heagerty PJ, Jarvik JG. Mortality among older adults with osteoporotic vertebral fracture. Osteoporos Int 2023;34:1561–75. doi:10.1007/s00198-023-06796-6.

[7] Wáng YXJ, Santiago FR, Deng M, Nogueira-Barbosa MH. Identifying osteoporotic vertebral endplate and cortex fractures. Quantitat Imag Med Surg 2017;7:55591. doi:10.21037/qims.2017.10.05.

[8] AO Spine Classification Systems n.d. https://www.aofoundation.org/spine/clinical-library-and-tools/aospine-classification-systems (Accessed September 24, 2023).

[9] Schnake KJ, Blattert TR, Hahn P, et al. Classification of osteoporotic thoracolumbar spine fractures: recommendations of the spine section of the German Society for Orthopaedics and Trauma (DGOU). Global Spine J 2018;8 46S–49S. doi:10.1177/2192568217717972.

[10] Kim YW, Mansfield LT. Fool me twice: delayed diagnoses in radiology with emphasis on perpetuated errors. AJR Am J Roentgenol 2014;202:465–70. doi:10.2214/AJR.13.11493.

[11] Tang X. The role of artificial intelligence in medical imaging research. BJR Open 2019;2:20190031. doi:10.1259/bjro.20190031.

[12] Hopkins BS, Yamaguchi JT, Garcia R, et al. Using machine learning to predict 30-day readmissions after posterior lumbar fusion: an NSQIP study involving 23,264 patients. J Neurosurg Spine 2019;32:399–406. doi:10.3171/2019.9.SPINE19860.

[13] Murata K, Endo K, Aihara T, et al. Artificial intelligence for the detection of vertebral fractures on plain spinal radiography. Sci Rep 2020;10:20031. doi:10.1038/s41598-020-76866-w.

[14] Zhu G, Jiang B, Tong L, Xie Y, Zaharchuk G, Wintermark M. Applications of deep learning to neuro-imaging techniques. Front Neurol 2019;10:869. doi:10.3389/fneur.2019.00869.

[15] Korez R, Ibragimov B, Likar B, Pernuš F, Vrtovec T. A framework for automated spine and vertebrae interpolation-based detection and model-based segmentation. IEEE Transact Med Imag 2015;34:1649–62. doi:10.1109/TMI.2015.2389334.

[16] Yousefirizi F, Decazes P, Amyar A, Ruan S, Saboury B, Rahmim A. AI-based detection, classification and prediction/prognosis in medical imaging. PET Clin 2022;17:183–212. doi:10.1016/j.cpet.2021.09.010.

[17] Tomita N, Cheung YY, Hassanpour S. Deep neural networks for automatic detection of osteoporotic vertebral fractures on CT scans. Computers in Biology and Medicine 2018;98:8–15. doi:10.1016/j.compbiomed.2018.05.011.

[18] Niu X, Yan W, Li X, et al. A deep-learning system for automatic detection of osteoporotic vertebral compression fractures at thoracolumbar junction using low-dose computed tomography images; 2022. In Review. doi:10.21203/rs.3.rs-2260322/v1.

[19] Obermeyer Z, Emanuel EJ. Predicting the future—big data, machine learning, and clinical medicine. N Engl J Med 2016;375:1216–19. doi:10.1056/NEJMp1606181.

[20] Whitney E, Alastra AJ. Vertebral fracture. StatPearls, Treasure Island (FL): StatPearls Publishing; 2024.

[21] Duarte ML, dos Santos LR, Oliveira ASB, Iared W, Peccin MS. Computed tomography with low-dose radiation versus standard-dose radiation for diagnosing fractures: systematic review and meta-analysis. Sao Paulo Med J 2021;139:388–97. doi:10.1590/1516-3180.2020.0374.R3.1902021.

[22] Shakeri A, Shakeri M, Ojaghzadeh Behrooz M, Behzadmehr R, Ostadi Z, Fouladi DF. Infrarenal aortic diameter, aortoiliac bifurcation level and lumbar disc degenerative changes: a cross-sectional MR study. Eur Spine J 2018;27:1096–104. doi:10.1007/s00586-017-5388-9.

[23] Chung WJ, Chung HW, Shin MJ, et al. MRI to differentiate benign from malignant soft-tissue tumours of the extremities: a simplified systematic imaging approach using depth, size and heterogeneity of signal intensity. Br J Radiol 2012;85:e831–6. doi:10.1259/bjr/27487871.