


# Deep Learning for Discrimination of Early Spinal Tuberculosis from Acute Osteoporotic Vertebral Fracture on CT

Wenjun Liu<sup>1</sup>, Jin Wang<sup>2</sup>, Yiting Lei<sup>1</sup>, Peng Liu<sup>3</sup>, Zhenghan Han<sup>1</sup>, Shichu Wang<sup>1</sup>, Bo Liu<sup>1</sup> 

<sup>1</sup>Department of Orthopedics, First Affiliated Hospital, Chongqing Medical University, Chongqing, People's Republic of China; <sup>2</sup>College of Medical Informatics, Chongqing Medical University, Chongqing, People's Republic of China; <sup>3</sup>Department of Orthopedics, Daping Hospital, Army Medical University, Chongqing, People's Republic of China

Correspondence: Bo Liu, Department of Orthopedics, First Affiliated Hospital, Chongqing Medical University, Chongqing, People's Republic of China, Tel +8613996065698, Email boliu@hospital.cqmu.edu.cn

**Background:** Early differentiation between spinal tuberculosis (STB) and acute osteoporotic vertebral compression fracture (OVCF) is crucial for determining the appropriate clinical management and treatment pathway, thereby significantly impacting patient outcomes.

**Objective:** To evaluate the efficacy of deep learning (DL) models using reconstructed sagittal CT images in the differentiation of early STB from acute OVCF, with the aim of enhancing diagnostic precision, reducing reliance on MRI and biopsies, and minimizing the risks of misdiagnosis.

**Methods:** Data were collected from 373 patients, with 302 patients recruited from a university-affiliated hospital serving as the training and internal validation sets, and an additional 71 patients from another university-affiliated hospital serving as the external validation set. MVITV2, Efficient-Net-B5, ResNet101, and ResNet50 were used as the backbone networks for DL model development, training, and validation. Model evaluation was based on accuracy, precision, sensitivity, F1 score, and area under the curve (AUC). The performance of the DL models was compared with the diagnostic accuracy of two spine surgeons who performed a blinded review.

**Results:** The MVITV2 model outperformed other architectures in the internal validation set, achieving accuracy of 98.98%, precision of 100%, sensitivity of 97.97%, F1 score of 98.98%, and AUC of 0.997. The performance of the DL models notably exceeded that of the spine surgeons, who achieved accuracy rates of 77.38% and 93.56%. The external validation confirmed the models' robustness and generalizability.

**Conclusion:** The DL models significantly improved the differentiation between STB and OVCF, surpassing experienced spine surgeons in diagnostic accuracy. These models offer a promising alternative to traditional imaging and invasive procedures, potentially promoting early and accurate diagnosis, reducing healthcare costs, and improving patient outcomes. The findings underscore the potential of artificial intelligence for revolutionizing spinal disease diagnostics, and have substantial clinical implications.

**Keywords:** deep learning, spinal tuberculosis, osteoporotic vertebral fractures, CT imaging, diagnostic accuracy

## Introduction

Spinal tuberculosis (STB) and acute osteoporotic vertebral compression fracture (OVCF) are common conditions in spinal surgery, being characterized by symptoms such as pain, inability to stand, and even spinal kyphotic deformity and spinal cord compression, potentially leading to neurological deficits and paralysis.<sup>1</sup> With the intensification of population aging, osteoporosis-induced OVCF has emerged as a significant global health concern,<sup>2</sup> with OVCF becoming the third most common fracture worldwide, with an estimated annual incidence of 1.4 million new OVCF cases. Tuberculosis, a severe infectious disease detrimental to health, is caused by the hematogenous spread of *Mycobacterium tuberculosis*.<sup>3</sup> STB is the most common form of extrapulmonary secondary tuberculosis<sup>4</sup> and is a frequent condition in spinal surgery;<sup>5</sup>

worldwide, the annual incidence of STB exceeds 100 000. STB is more prevalent in developing countries,<sup>6</sup> where early detection and treatment can lead to favorable prognoses.<sup>7</sup>

Early-stage STB and vertebral compression fractures often manifest with indistinguishable imaging features and clinical presentations.<sup>8–10</sup> In both cases, patients typically present with primary symptoms of lower back pain accompanied by restricted mobility, and one or multiple vertebrae may exhibit damage or alterations in signal intensity on imaging,<sup>11,12</sup> making early differentiation challenging, especially in elderly patients without a history of trauma.<sup>13</sup> The similarities between the two conditions make their differential diagnosis difficult, hindering early clinical diagnosis. Uncertainty in early diagnosis may lead to inappropriate treatment strategies that exacerbate the patient's condition.<sup>14</sup> With advances in spinal surgery, percutaneous balloon kyphoplasty (PKP) is often the preferred treatment for OVCF;<sup>15,16</sup> however, spinal infectious diseases such as STB are contraindications for PKP, because PKP may lead to the spread of infection and difficulty in eradicating it,<sup>17</sup> and therefore the misdiagnosis of STB as OVCF can have catastrophic consequences.<sup>18</sup> Patients with STB require early and regular anti-tuberculosis drug treatment,<sup>19</sup> regardless of surgical intervention. Given the stark differences in treatment approaches required for STB and OVCF, accurate early diagnosis is crucial for selecting the correct treatment method.

Clinically, the reference standards for diagnosing STB and OVCF are *Mycobacterium* culture<sup>20</sup> and magnetic resonance imaging (MRI), respectively. However, in many countries MRI equipment is not widely available,<sup>21</sup> and even when available, various factors may preclude certain patients from undergoing MRI examinations. These factors include metallic implants or devices, pregnancy (particularly the early stages), claustrophobia, excessive body weight, and inability to remain still.<sup>22</sup> Furthermore, the biopsy techniques used for diagnosing STB are invasive procedures that require skilled spinal surgeons, and many risks are associated with the procedure, including patient discomfort, further spread of the bacteria, and nerve root damage.<sup>23</sup> Additionally, the culture of *Mycobacterium tuberculosis*, a critical step for confirmation, is time-consuming because of the bacterium's slow growth rate.<sup>24</sup> This extended culture time hinders early diagnosis of patients, thereby impeding early treatment.<sup>25,26</sup> Based on the aforementioned reasons, existing diagnostic methods (biopsy, bacterial culture, and MRI examination) fail to meet the widespread demands for differential diagnosis of these two diseases in many countries and populations. Therefore, the development and refinement of a new non-invasive early diagnostic technique is crucial for the treatment and management of these diseases.

Computed tomography (CT) is a routinely utilized diagnostic modality in orthopedics, often being employed to further investigate suspicious features observed on X-ray images.<sup>27</sup> CT has high spatial resolution, enabling differentiation between osteoporotic fractures and pathological fractures on the basis of morphological characteristics such as bone integrity and fracture edge. However, CT is not sensitive to edema or hematoma caused by tuberculosis infection or fresh fractures. Although CT can display certain features of typical STB and OVCF, it cannot differentiate between these two diseases in their early stages, especially when STB coexists with osteoporosis.

Traditional image analysis methods such as threshold-based segmentation, edge detection, and region growing struggle with complex CT images because of their dependence on manual feature extraction, which is inefficient and inflexible. These methods are also prone to image noise, frequently necessitating human intervention, and thereby reducing processing efficiency.<sup>28,29</sup> Furthermore, their effectiveness and adaptability falter with high-dimensional and large-scale CT data, making methods based on complex algorithms challenging to interpret.

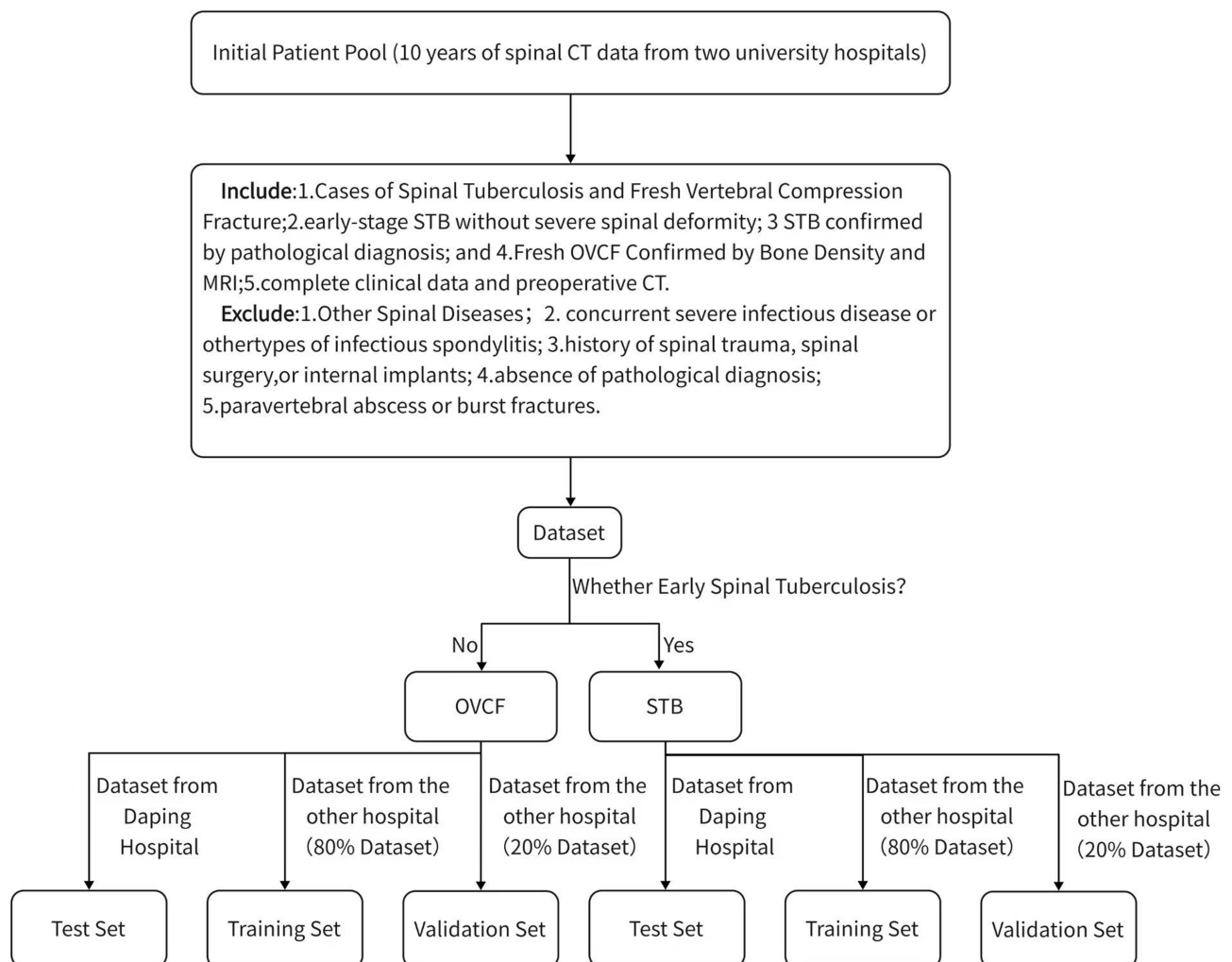
With the development of deep learning and advanced computer vision technologies, artificial intelligence models can automatically learn and extract features, learn complex patterns and features, and capture comprehensive structural information from images, effectively handling tasks primarily focused on image recognition. The application of artificial intelligence in medical image analysis continues to increase, with studies demonstrating that deep learning can serve as an effective technique for fracture diagnosis.<sup>30</sup> Its effectiveness for diagnosing fractures in multiple body parts has been validated,<sup>31</sup> with deep learning having been successfully applied in automatic vertebral segmentation<sup>32</sup> and opportunistic osteoporosis screening.<sup>33</sup> All of this indicates that deep learning models are a promising approach for orthopedic medical image analysis.

Our study aimed to develop deep learning models using sagittal CT images of patients with STB or OVCF, and to evaluate the models' abilities and feasibility for distinguishing between acute OVCF and early-stage STB, thereby guiding early clinical differential diagnosis.

## Methods

### Patient Datasets

Patient images were sourced from the orthopedic CT imaging databases of two affiliated medical university hospitals (Figure 1), with informed consent for the imaging being obtained at admission. Data were collected from individuals diagnosed with acute vertebral compression fracture or STB at these two institutions over a ten-year period. The inclusion criteria were: (1) early-stage STB without severe spinal deformity; (2) STB confirmed by pathological diagnosis; and (3) complete clinical data and preoperative CT. The exclusion criteria included: (1) concurrent severe infectious disease or other types of infectious spondylitis; (2) history of spinal trauma, spinal surgery, or internal implants; (3) absence of pathological diagnosis; and (4) paravertebral abscess or burst fractures. The age, gender, and vertebral lesion location of each patient were obtained from electronic medical records. A total of 373 patients were included in this study, with 302 being from hospital 1 (average age 62.02 years, age range 19–98 years). This group consisted of 149 patients with acute OVCF (average age 72.60 years, age range 36–98 years) and 153 patients with STB (average age 51.61 years, age range 19–96 years). Additionally, 71 patients were recruited from another affiliated hospital (hospital 2; average age 62.03 years, age range 37–94 years), with these consisting of 32 patients with acute OVCF (average age 61.63 years, age range 37–94 years) and 39 patients with STB (average age 62.36 years, age range 48–82 years). The STB cases were confirmed by biopsy results as tuberculosis. All cases of acute OVCF had no known history



**Figure 1** Patient selection and grouping flowchart.

of tuberculosis and were diagnosed with osteoporosis or severe osteoporosis based on bone density results. The CT images were reformatted into sagittal views from the original transverse section images.

This study is reported in line with the STROCSS criteria.<sup>34</sup>

## Deep Learning Models

A senior orthopedic surgeon with 20 years of experience collaborated with a junior orthopedic doctor to analyze the patient images. We specifically selected the most abnormal vertebral segments for deep learning analysis. Regions of interest (ROIs) were manually delineated on the sagittal slices, then a computer program generated rectangles that encompassed the entirety of abnormal areas, which were then input into the deep learning models. For this study, we employed four widely used neural network architectures: ResNet50, ResNet101, Multiscale Vision Transformers (MVITV2), and EfficientNet\_B5. These four models were trained and their effectiveness for differentiating between STB and OVCF was assessed.

ResNet50 is a Residual Network variant with 50 layers, known for its “residual learning” feature for solving deep network training degradation by using skip connections for smoother training. ResNet101 is an extension of ResNet with 101 layers, offering deeper and more complex feature learning than ResNet50 through similar residual learning techniques. MVITV2 is designed for image processing, with the model using vision transformers to analyze images at multiple scales, capturing a wide range of features for complex analysis. EfficientNet\_B5 is known for its efficient scaling of network depth, width, and resolution, and this model from the EfficientNet series achieves high accuracy with fewer parameters, optimizing three key network dimensions.

Our objective was to assess the models’ effectiveness in differentiating between STB and OVCF. The model inputs included the target slice and the four adjacent sagittal slices. Images were resized to a fixed 145×210 matrix and the pixel intensities were normalized to a mean of 0 and standard deviation of 1. The training parameters consisted of 100 training epochs, an AdamW optimizer, a learning rate (ResNet101, 0.00001; MVITV2, 0.000005; EfficientNet\_B5, 0.0001; ResNet50, 0.00005), and a batch size of 32. The patients were randomly split into training and validation sets in a 4:1 ratio. To enhance the dataset, we performed image augmentation through random affine transformations, including translation, rotation, flipping, and scaling, thereby expanding the dataset by 20 times. The diagnostic performance of the models was analyzed and compared using receiver operating characteristics (ROC) curves, area under the curve (AUC), accuracy, precision, sensitivity, F1 score, and confusion matrices. Importantly, our image processing did not involve the inclusion of patient clinical features, such as medical history and laboratory test results.

## Statistical Analysis

Python 3.8 was used for coding and developing the deep learning networks with Pytorch and TensorFlow. The Scikit-learn toolkit was used to implement model comparisons and perform related analyses, such as drawing ROC curves and calculating AUC. A *p* value of less than 0.05 was defined as statistically significant.

## Ethical Approval

This study was approved by the Ethics Committee of the First Affiliated Hospital of Chongqing Medical University (Approval No. K2024-028-01) and the Ethics Committee of the Army Specialty Medical Center of the Chinese People’s Liberation Army (Approval No. [2024] No. 29). All procedures performed in this study were in accordance with the ethical standards of the institutional and national research committees and with the 1964 Declaration of Helsinki and its later amendments.

## Results

### Study Participants

This study included a total of 373 patients (Table 1). Among the 302 patients collected from hospital 1, 153 patients had STB (male/female: 77/76, average age: 51.71 ± 16.52 years, 358 affected vertebrae) and 149 patients had OVCF (male/female: 35/114, average age: 72.24 ± 11.38 years, 175 affected vertebrae). We randomly assigned 80% of the patients

(123 patients with STB and 120 patients with OVCF) to the training set, with the remaining 20% (30 STB patients and 29 OVCF patients) assigned to the internal validation set. From hospital 2, we collected data from 71 patients, including 39 with STB (male/female: 17/22, average age:  $62.36 \pm 10.65$  years, 45 affected vertebrae) and 32 with OVCF (male/female: 16/16, average age:  $61.63 \pm 13.32$  years, 41 affected vertebrae). As shown in Table 1, there were significant differences in age and gender ratios between the STB and OVCF groups ( $p < 0.05$ ). The distributions of lesions across different spinal regions were as follows: in the training set and internal validation set, the STB group consisted of 199 lesions in the thoracic spine, 143 in the lumbar spine, and 16 in the sacral spine, while the OVCF group consisted of 89 lesions in the thoracic spine, 86 in the lumbar spine, and none in the sacral spine. There were no cervical spine lesions in either group. In the external validation set, the STB group consisted of 4 lesions in the cervical spine, 24 in the thoracic spine, and 27 in the lumbar spine, while the OVCF group consisted of 25 lesions in the thoracic spine, 16 in the lumbar spine, and none in the cervical spine. There were no sacral spine lesions in either group.

In the comparative analysis presented in Table 2 and illustrated in Figure 2, the MVITV2 framework demonstrated superior performance to its counterparts, achieving an accuracy rate of 98.98%, precision rate of 100% (indicating that all predictions labeled as fractures were accurately validated, with no false positives), sensitivity of 97.97%, F1 score of 98.98%, and AUC of 0.997. Following closely, the ResNet101 architecture achieved accuracy of 96.95%, precision of 95.86%, sensitivity of 97.89%, F1 score of 96.86%, and AUC of 0.995. The ResNet50 model achieved accuracy of 92.20%, precision of 93.79%, sensitivity of 90.67%, F1 score of 92.20%, and AUC of 0.968. The EfficientNet\_B5 model achieved accuracy of 86.78%, precision of 91.72%, sensitivity of 83.12%, F1 score of 87.21%, and AUC of 0.948.

On the external test set, the MVITV2 model demonstrated excellent performance, with accuracy of 83.77%, precision of 84.84%, sensitivity of 83.09%, F1 score of 83.38%, and AUC of 0.879, as shown in Table 2 and Figure 3. The ResNet101 model achieved accuracy of 72.75%, precision of 72.61%, sensitivity of 72.57%, F1 score of 72.59%, and AUC of 0.751. The ResNet50 model showed slightly poorer performance on the external validation set, with accuracy of 60.00%, precision of 74.05%, sensitivity of 62.58%, F1 score of 55.74%, and AUC of 0.686. The EfficientNet\_B5 model performed well, with accuracy of 76.81%, precision of 77.69%, sensitivity of 77.41%, F1 score of 76.80%, and AUC of 0.844.

In the independent and blind evaluation of the training set and internal test set, the two spine surgeons achieved accuracy rates of 77.38% (junior doctor) and 93.56% (senior surgeon). A confusion matrix (Figure 3) indicates that the neural network models surpassed the level of the junior resident doctor and matched or exceeded the performance of the

**Table 1** Patient Characteristics Comparison Between the STB and OVCF Groups

		ALL	STB	OVCF
Training set and internal validation set				
Age ("±" –standard deviation.)		61.84±17.54	51.71 ±16.52	72.24±11.38
Gender (male/female)		112/190	77/76	35/114
Number of patients in different spinal regions	Thoracic	288	199	89
	Lumbar	229	143	86
	Sacral	16	16	0
External validation set				
Age ("±" –standard deviation.)		62.03 ±11.93	62.36 ± 10.65	61.63 ± 13.32
Gender (male/female)		33/38	17/22	16/16
Number of patients in different spinal regions	Cervical	4	4	0
	Thoracic	49	24	25
	Lumbar	43	27	16

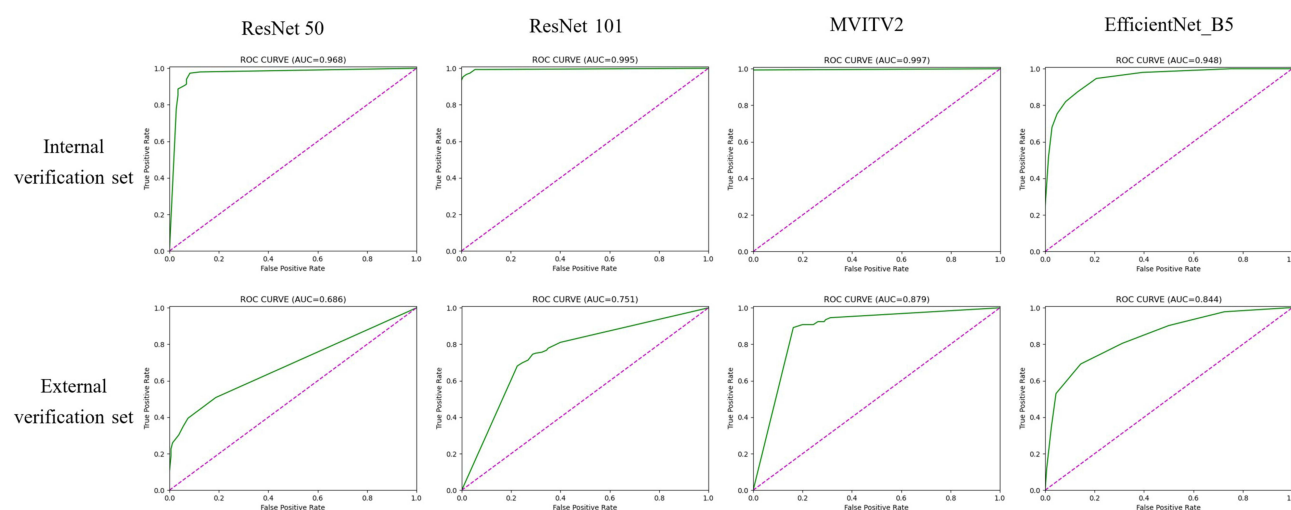
**Table 2** Comparison of Diagnostic Accuracy Across Different Deep Learning Models and Experienced Surgeons

Models	Accuracy	Precision	Sensitivity	F1 Score
Internal validation set				
Resnet50	92.20%	93.79%	90.67%	92.20%
Resnet101	96.95%	95.86%	97.89%	96.86%
EfficientNet-B5	86.78%	91.72%	83.12%	87.21%
MVITV2	98.98%	100%	97.97%	98.98%
External Validation Set				
Resnet50	60.00%	74.05%	62.58%	55.74%
Resnet101	72.75%	72.61%	72.57%	72.59%
EfficientNet-B5	76.81%	77.69%	77.41%	76.80%
MVITV2	83.77%	84.84%	83.09%	83.38%
Surgeons				
Junior doctor	77.38%	76.92%	75.34%	76.12%
Senior doctor	93.56%	94.48%	92.57%	93.52%

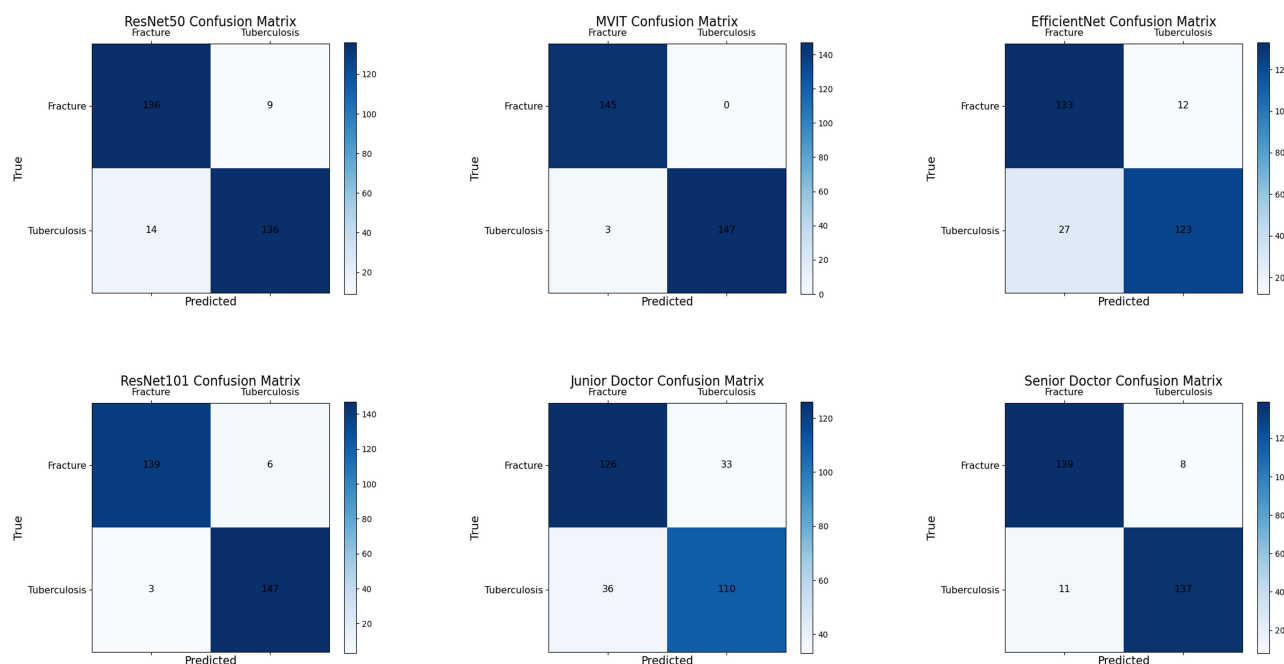
experienced spinal surgeon. Several applications of the models are visually illustrated in Figure 4, which showcases four true positive cases. The CT sagittal images of these four patients were analyzed using the MVITV2 model, and all four patients were not included in the training dataset. The numbers on the images represent the model's probability estimates for the corresponding diseases. The vertebrae with the most evident lesions, denoted by green boxes, indicate the ROIs that the model focused on.

## Discussion

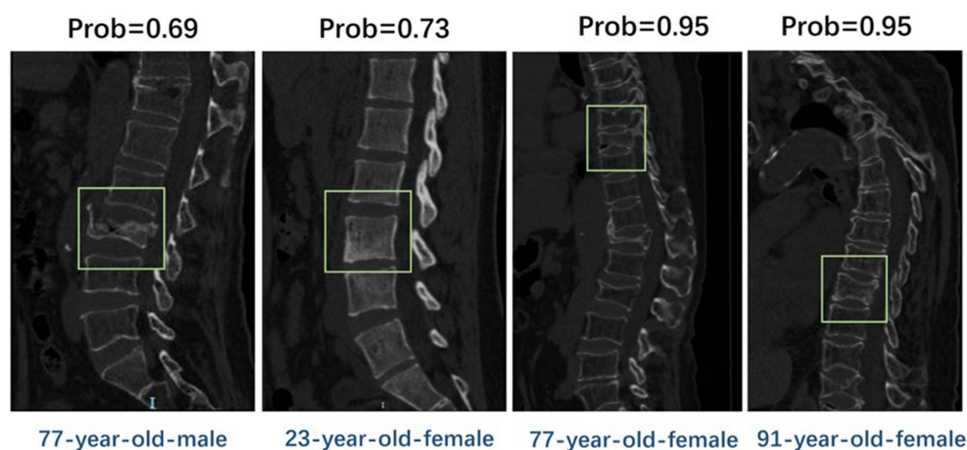
In many countries, STB and OVCF are common, and both require early intervention and customized treatment. In early-stage STB and OVCF, the clinical symptoms may not facilitate early differentiation, with the patient often presenting

**Figure 2** Curves for all models on the validation set, in the following order: ResNet50, ResNet101, MVITV2, EfficientNet\_B5.





**Figure 3** Comparison of results between DL models and surgeons using confusion matrix.



**Figure 4** Visual examples demonstrate four True Positive (TP) cases, all examples of patients not included in the training dataset: The first case involves osteoporosis and vertebral compression changes due to a fall, while the second case, without trauma or osteoporosis, highlighting that spinal tuberculosis can have different appearances on CT scans. The final two cases, both with osteoporosis but no trauma, showcase the model's ability to differentiate osteoporotic compression fractures in CT sagittal images.

with back pain, emphasizing the importance of imaging or laboratory tests in the differential diagnosis.<sup>35,36</sup> However, various limitations may restrict the use of percutaneous biopsy, mycobacterial culture, and MRI examination. In comparison, CT machines are already widely available in most hospitals, and even in primary healthcare centers. Therefore, the application of deep learning models to sagittal CT images allows the models to be more easily applied in clinical practice. Our study aimed to utilize artificial intelligence models trained on reconstructed CT images to differentiate between OVCF and STB on CT images, addressing the current clinical challenges and aiding in early and improved diagnosis and prognostication for patients. Our results showed that all models performed robustly on the internal dataset, with the MVITV2 model achieving 98.98% accuracy and AUC of 0.997. The models also yielded notable outcomes on the external validation dataset, aside from the slightly lower performance of ResNet50. The MVITV2 model achieved 83.77% accuracy and AUC of 0.879. When these values were compared with those of the

spinal surgeons of different seniority, the deep learning model matched the diagnostic abilities of the experienced surgeon, offering a non-invasive tool for junior clinicians to accurately and swiftly differentiate between STB and OVCF, and providing early diagnostic insights for senior surgeons. This demonstrates that the artificial intelligence models (especially the MVITV2 model) developed in this study could effectively differentiate between STB and OVCF on CT images.

Artificial intelligence, particularly deep learning and transfer learning, is increasingly being utilized in medical image analysis for lesion localization, disease diagnosis, tumor classification, and progression prediction.<sup>37,38</sup> Deep learning surpasses traditional machine learning by extracting complex features without predefined parameters.<sup>38</sup> Duan et al<sup>39</sup> used MRI-based deep learning models to distinguish spinal tumors from tuberculosis. Ke Liu et al<sup>40</sup> applied an MRI-based ResNet to identify the origins of spinal metastases, which included lung, kidney, prostate, thyroid, and breast cancer. Yuan Li et al<sup>41</sup> demonstrated that ResNet50 could differentiate benign from malignant vertebral fractures on CT images. Chen et al<sup>42</sup> showed that an attention-based model could separate multiple myeloma from spinal metastases. Kim et al<sup>43</sup> retrained Inception v3 with wrist X-rays for fracture classification. Cheng et al<sup>44</sup> employed Dense-U-Net for segmentation of vertebrae on CT, achieving full spine segmentation. Such neural network models convert medical images into high-dimensional features for tasks like lesion classification and segmentation,<sup>45–48</sup> offering applications in tumor detection, disease differentiation, and pathological analysis.<sup>49–51</sup>

In recent years, the application of artificial intelligence, particularly deep learning, to medical imaging has seen remarkable progress. Previous studies have demonstrated the potential of deep learning models to differentiate between diseases with similar radiologic features. For example, Wang et al<sup>2</sup> successfully employed convolutional neural networks (CNNs) to distinguish between malignant and benign spinal lesions on MRI, while our study extends this approach to the differential diagnosis of STB and OVCF using CT images. Deep learning models offer the advantage of automatically extracting complex features from imaging data, often with better sensitivity and specificity than traditional methods (Zhao et al<sup>52</sup>).

The application of deep learning, convolutional neural networks (CNNs), and Transformer models has significantly impacted medical imaging analysis. These algorithms are designed to autonomously learn from training datasets, enabling them to identify key features.<sup>50,51</sup> In this learning process, they adjust their internal parameters to appropriately weight these features, thereby creating models that can be tested with new datasets. In this study, we trained and evaluated three CNN models (ResNet50, ResNet101, and EfficientNet\_B5) and a representative Transformer model (MVITV2).

The pathogenesis of spinal tuberculosis (STB) involves the direct infection of the vertebrae by *Mycobacterium tuberculosis*, leading to localized granulomatous inflammation and bone destruction. In the early stages, imaging primarily shows inflammatory changes, with minimal abnormal bone density. In contrast, osteoporotic vertebral compression fractures (OVCF) are mainly caused by a decrease in bone density due to osteoporosis. The occurrence of OVCF results from a reduction in the mechanical strength of the vertebrae, leading to compression fractures under minimal external forces. Early imaging may show a combination of hemorrhage, edema, and osteoporotic changes. Based on these mechanisms, our artificial intelligence-based image recognition model can accurately differentiate between these two diseases with high accuracy.

These models collectively represent the forefront of deep learning applications in the field of medical imaging, offering new avenues for accurate and efficient analysis. Our experiments demonstrated that the aforementioned deep learning models are feasible for the differentiation of fresh vertebral compression fractures and STB on CT images.

Although previous studies have highlighted the benefits of incorporating clinical attributes into diagnostic models,<sup>53</sup> our focus was on using CT imaging to differentiate STB from OVCF, especially among elderly women.<sup>54</sup> Given the unreliability of the patients' tuberculosis history, with many cases lacking evidence of pulmonary tuberculosis or disease at other primary sites, and the commonness of osteoporosis in such patients, our research emphasized radiological over clinical data. Our findings suggest that deep learning models, without incorporating extensive medical history data or laboratory results, can match the diagnostic accuracy of experienced spinal surgeons. It is worth noting that the deep learning models developed in this study are still an auxiliary diagnostic tool. When their diagnostic results conflict with medical opinion, the final diagnosis should be comprehensively determined by doctors.



This study has several limitations. First, the relatively small sample size of the dataset may limit the generalization ability of the deep learning network models. Further training and validation with larger sample sizes from more centers are needed. Second, the neural network models developed in this study did not perform lesion localization and automatic segmentation, but were based on manual segmentation and correction by physicians. Third, the retrospective design of the study may introduce case selection bias, potentially affecting the credibility of the results. Our goal for future research is to continuously incorporate new cases, expand the sample size to enhance the efficiency and generalization ability of the model, and introduce vertebral automatic segmentation and localization functions into the model to avoid bias introduced by manual segmentation and improve segmentation efficiency.

In conclusion, this study explored the application of deep learning methods in differentiating spinal tuberculosis (STB) and acute osteoporotic vertebral compression fractures (OVCF) using CT images. Four models (ResNet50, ResNet101, EfficientNet\_B5, and MVITV2) all demonstrated high accuracy, with the MVITV2 model performing exceptionally well on the internal dataset, and showing strong robustness in external validation. However, to further improve the generalizability and diagnostic accuracy of these models, continued training with larger, multi-center datasets is necessary.

From a clinical perspective, the implementation of these deep learning models provides clinicians with a rapid, efficient, and non-invasive diagnostic tool. Particularly in primary care settings where MRI may not be widely available but CT machines are more commonly used, the CT image analysis model developed in this study holds significant value in resource-limited environments. These models can reduce dependence on MRI, lower the need for invasive biopsy procedures and their associated risks, and assist clinicians in identifying lesions more quickly and accurately, thus improving the diagnostic pathway for patients.

Furthermore, the performance of these deep learning models approaches or even surpasses the diagnostic capabilities of some experienced spinal surgeons, providing strong technical support for less experienced clinicians or those in primary care. In challenging cases, these models can also offer supplementary information to senior clinicians, enhancing diagnostic confidence. Particularly for elderly female patients, where medical history may be unreliable and osteoporosis is highly prevalent, this study demonstrates that relying solely on CT imaging data, rather than patient history or laboratory results, can still achieve a high level of diagnostic accuracy. This is crucial for optimizing clinical decision-making processes and reducing diagnostic time.

## Highlights

**Question:** Given the clinical and radiological similarities between early-stage spinal tuberculosis (STB) and osteoporotic vertebral compression fracture (OVCF), how effective are CT image-based deep learning models for differentiating early-stage STB from acute OVCF, and what is their potential impact on early diagnosis?

**Findings:** Using data from 373 patients across multiple centers, a MVITV2 deep learning framework achieved 98.98% accuracy, 100% precision, and an AUC of 0.997, significantly outperforming spine surgeons in differentiating between STB and OVCF.

**Meaning:** The application of deep learning to CT imaging provides a precise non-invasive method for distinguishing early-stage STB from OVCF, offering a critical tool for enhancing diagnostic accuracy and establishing treatment strategies.

## Abbreviations

STB, spinal tuberculosis; OVCF, osteoporotic vertebral compression fracture; CT, Computed Tomography; MVITV2, Multiscale Vision; TransformersResNet, Residual NetworkEfficientNet, Efficient Network.

## Data Sharing Statement

The datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

## Ethics Approval and Consent to Participate

Ethical approval was obtained for this study, and a waiver was granted for the requirement for informed patient consent. This study was conducted in accordance with the World Medical Association Declaration of Helsinki.

## Acknowledgments

The authors would like to thank all et al who assisted in the preparation of this manuscript. We thank Liwen Bianji (Edanz) ([www.liwenbianji.cn](http://www.liwenbianji.cn)) for editing the language of a draft of this manuscript.

## Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

## Funding

There is no funding to report.

## Disclosure

The authors declare that they have no competing interests.

## References

1. Qiao P, Zhao P, Gao Y, Bai Y, Niu G. Differential study of DCE-MRI parameters in spinal metastatic tumors, brucellar spondylitis and spinal tuberculosis. *Chin J Cancer Res Chung-Kuo Yen Cheng Yen Chiu*. 2018;30(4):425–431. doi:10.21147/j.issn.1000-9604.2018.04.05
2. Bell A, Kendler DL, Khan AA, et al. A retrospective observational study of osteoporosis management after a fragility fracture in primary care. *Arch Osteoporos*. 2022;17(1):75. doi:10.1007/s11657-022-01110-z
3. Guo F, Wei J, Song Y, et al. Immunological effects of the PE/PPE family proteins of mycobacterium tuberculosis and related vaccines. *Front Immunol*. 2023;14:1255920. doi:10.3389/fimmu.2023.1255920
4. Khanna K, Sabharwal S. Spinal tuberculosis: a comprehensive review for the modern spine surgeon. *Spine J off J North Am Spine Soc*. 2019;19(11):1858–1870. doi:10.1016/j.spinee.2019.05.002
5. Zhan Y, Kang X, Gao W, et al. Efficacy analysis of one-stage posterior-only surgical treatment for thoracic spinal tuberculosis in the t4-6 segments with minimum 5-year follow-up. *Sci Rep*. 2022;12(1):149. doi:10.1038/s41598-021-04138-2
6. Chen L, Gan Z, Huang S, et al. Blood transfusion risk prediction in spinal tuberculosis surgery: development and assessment of a novel predictive nomogram. *Bmc Musculoskelet Disord*. 2022;23(1):182. doi:10.1186/s12891-022-05132-z
7. Hu X, Zhang H, Li Y, et al. Analysis of the diagnostic efficacy of the QuantiFERON-TB gold in-tube assay for preoperative differential diagnosis of spinal tuberculosis. *Front Cell Infect Microbiol*. 2022;12:983579. doi:10.3389/fcimb.2022.983579
8. Diagnosis and treatment of skipped multifocal spinal tuberculosis lesions – PubMed Available from: <https://pubmed.ncbi.nlm.nih.gov/37186216/>. Accessed December 5, 2023.
9. Ahmadi J, Bajaj A, Destian S, Segall HD, Zee CS. Spinal tuberculosis: atypical observations at MR imaging. *Radiology*. 1993;189(2):489–493. doi:10.1148/radiology.189.2.8210378
10. Pu F, Feng J, Yang L, Zhang L, Xia P. Misdiagnosed and mismanaged atypical spinal tuberculosis: a case series report. *Exp Ther Med*. 2019;18(5):3723–3728. doi:10.3892/etm.2019.8014
11. A predictive model for early clinical diagnosis of spinal tuberculosis based on conventional laboratory indices: a multicenter real-world study – PubMed Available from: <https://pubmed.ncbi.nlm.nih.gov/37033479/>. Accessed December 5, 2023.
12. Li T, Liu T, Jiang Z, Cui X, Sun J. Diagnosing pyogenic, brucella and tuberculous spondylitis using histopathology and MRI: a retrospective study. *Exp Ther Med*. 2016;12(4):2069–2077. doi:10.3892/etm.2016.3602
13. Currie S, Galea-Soler S, Barron D, Chandramohan M, Groves C. MRI characteristics of tuberculous spondylitis. *Clin Radiol*. 2011;66(8):778–787. doi:10.1016/j.crad.2011.02.016
14. Wang B, Gao W, Hao D. Current study of the detection and treatment targets of spinal tuberculosis. *Curr Drug Targets*. 2020;21(4):320–327. doi:10.2174/1389450120666191002151637
15. Zhao Z, Deng L, Hua X, et al. A retrospective study on the efficacy and safety of bone cement in the treatment of endplate fractures. *Front Surg*. 2022;9:999406. doi:10.3389/fsurg.2022.999406
16. Full article: effects of percutaneous kyphoplasty for the treatment of thoracic osteoporotic vertebral compression fractures with or without intravertebral cleft in elderly patients Available from: <https://www.tandfonline.com/doi/full/10.2147/IJGM.S447623>. Accessed February 28, 2024.
17. Liao JC, Lai PL, Chen LH, Niu CC. Surgical outcomes of infectious spondylitis after vertebroplasty, and comparisons between pyogenic and tuberculosis. *BMC Infect Dis*. 2018;18(1). doi:10.1186/s12879-018-3486-x
18. Tuberculous spondylitis after percutaneous vertebroplasty: a case series of 9 cases - PubMed Available from: <https://pubmed.ncbi.nlm.nih.gov/31627871/>. Accessed December 6, 2023.

19. Tang L, Fu CG, Zhou ZY, et al. Clinical features and outcomes of spinal tuberculosis in central China. *Infect Drug Resist.* **2022**;15:6641–6650. doi:10.2147/IDR.S384442
20. Li W, Liu Z, Xiao X, Zhang Z, Wang X. Comparison of anterior transthoracic debridement and fusion with posterior transpedicular debridement and fusion in the treatment of mid-thoracic spinal tuberculosis in adults. *Bmc Musculoskelet Disord.* **2019**;20(1):570. doi:10.1186/s12891-019-2945-x
21. Hossein H, Ali KM, Hosseini M, Sarveazad A, Safari S, Yousefifard M. Value of chest computed tomography scan in diagnosis of COVID-19; a systematic review and meta-analysis. *Clin Transl Imaging.* **2020**;8(6):469–481. doi:10.1007/s40336-020-00387-9
22. Winter L, Seifert F, Zilberti L, Murbach M, Ittermann B. MRI-related heating of implants and devices: a review. *J Magn Reson Imaging JMRI.* **2021**;53(6):1646–1665. doi:10.1002/jmri.27194
23. Wang H, Xu S, bin FK, Dai ZS, Wei GZ, Chen LF. Contrast-enhanced magnetic resonance image segmentation based on improved U-net and inception-ResNet in the diagnosis of spinal metastases. *J Bone Oncol.* **2023**;42:100498. doi:10.1016/j.jbo.2023.100498
24. Huang Y, Ai L, Wang X, Sun Z, Wang F. Review and updates on the diagnosis of tuberculosis. *J Clin Med.* **2022**;11(19):5826. doi:10.3390/jcm11195826
25. Naidoo K, Perumal R, Ngema SL, Shunmugam L, Somboro AM. Rapid diagnosis of drug-resistant tuberculosis—opportunities and challenges. *Pathogens.* **2024**;13(1):27. doi:10.3390/pathogens13010027
26. So C, Ling L, Wong WT, et al. Population study on diagnosis, treatment and outcomes of critically ill patients with tuberculosis (2008–2018). *Thorax.* **2023**;78(7):674–681. doi:10.1136/thorax-2022-218868
27. Guo R, Zou Y, Zhang S, et al. Preclinical validation of a novel deep learning-based metal artifact correction algorithm for orthopedic CT imaging. *J Appl Clin Med Phys.* **2023**;24(11):e14166. doi:10.1002/acm2.14166
28. Javeed MA, Ghaffar MA, Ashraf MA, et al. Lane line detection and object scene segmentation using Otsu thresholding and the fast Hough transform for intelligent vehicles in complex road conditions. *Electronics.* **2023**;12(5):1079. doi:10.3390/electronics12051079
29. Sm G, Sh K. Analysis of medical images using image registration feature-based segmentation techniques. *Int Conference Tech Adv Comput Sci (ICTACS).* **2022**;485–490. doi:10.1109/ICTACS6270.2022.9987895
30. Badgeley MA, Zech JR, Oakden-Rayner L, et al. Deep learning predicts hip fracture using confounding patient and healthcare variables. *Npj Digit Med.* **2019**;2:31. doi:10.1038/s41746-019-0105-1
31. Kubota T, Yamada K, Ito H, Kizu O, Nishimura T. High-resolution imaging of the spine using multidetector-row computed tomography: differentiation between benign and malignant vertebral compression fractures. *J Comput Assist Tomogr.* **2005**;29(5):712–719. doi:10.1097/01.rct.0000175500.41836.24
32. Sekuboyina A, Hussein ME, Bayat A, et al. VerSe: a vertebrae labelling and segmentation benchmark for multi-detector CT images. *Med Image Anal.* **2021**;73:102166. doi:10.1016/j.media.2021.102166
33. Fang Y, Li W, Chen X, et al. Opportunistic osteoporosis screening in multi-detector CT images using deep convolutional neural networks. *Eur Radiol.* **2021**;31(4):1831–1842. doi:10.1007/s00330-020-07312-8
34. Mathew G, Agha R, Albrecht J. STROCSS 2021: strengthening the reporting of cohort, cross-sectional and case-control studies in surgery. *Int j Surg.* **2021**;96:106165. doi:10.1016/j.ijsu.2021.106165
35. Waters R, Held M, Dunn R, Laubscher M, Adikary N, Coussens A. Improving the diagnosis of spinal tuberculosis and our understanding of its pathophysiology and interaction with hiv-1 co-infection. *Orthop Proc.* **2023**;105(SUPP\_15):36. doi:10.1302/1358-992X.2023.15.036
36. Ng GS, Lee LY, Chu EC. Undiagnosed osteoporotic vertebral fractures in an octogenarian during the coronavirus disease pandemic. *Cureus.* **2023**;15(5). doi:10.7759/cureus.38585
37. Mahmood T, Li J, Pei Y, Akhtar F, Rehman MU, Wasti SH. Breast lesions classifications of mammographic images using a deep convolutional neural network-based approach. *PLoS One.* **2022**;17(1):e0263126. doi:10.1371/journal.pone.0263126
38. Torres C, Hammond I. Computed tomography and magnetic resonance imaging in the differentiation of osteoporotic fractures from neoplastic metastatic fractures. *J Clin Densitom off J Int Soc Clin Densitom.* **2016**;19(1):63–69. doi:10.1016/j.jocd.2015.08.008
39. Duan S, Dong W, Hua Y, et al. Accurate differentiation of spinal tuberculosis and spinal metastases using MR-based deep learning algorithms. *Infect Drug Resist.* **2023**;16:4325–4334. doi:10.2147/IDR.S417663
40. Liu K, Qin S, Ning J, et al. Prediction of primary tumor sites in spinal metastases using a ResNet-50 convolutional neural network based on MRI. *Cancers.* **2023**;15(11):2974. doi:10.3390/cancers15112974
41. Li Y, Zhang Y, Zhang E, et al. Differential diagnosis of benign and malignant vertebral fracture on CT using deep learning. *Eur Radiol.* **2021**;31(12):9612–9619. doi:10.1007/s00330-021-08014-5
42. Chen K, Cao J, Zhang X, et al. Differentiation between spinal multiple myeloma and metastases originated from lung using multi-view attention-guided network. *Front Oncol.* **2022**;12:981769. doi:10.3389/fonc.2022.981769
43. Lv M, Zhou Z, Tang Q, et al. Differentiation of usual vertebral compression fractures using CT histogram analysis as quantitative biomarkers: a proof-of-principle study. *Eur J Radiol.* **2020**;131:109264. doi:10.1016/j.ejrad.2020.109264
44. Cheng P, Yang Y, Yu H, He Y. Automatic vertebrae localization and segmentation in CT with a two-stage dense-U-net. *Sci Rep.* **2021**;11(1):22156. doi:10.1038/s41598-021-01296-1
45. Soleimani P, Farezi N. Utilizing deep learning via the 3D U-net neural network for the delineation of brain stroke lesions in MRI image. *Sci Rep.* **2023**;13(1):19808. doi:10.1038/s41598-023-47107-7
46. Xiong X, Smith BJ, Graves SA, Graham MM, Buatti JM, Beichel RR. Head and neck cancer segmentation in FDG PET images: performance comparison of convolutional neural networks and vision transformers. *Tomogr Ann Arbor Mich.* **2023**;9(5):1933–1948. doi:10.3390/tomography9050151
47. Reddy DA, Roy S, Kumar S, Tripathi R. Enhanced U-net segmentation with ensemble convolutional neural network for automated skin disease classification. *Knowl Inf Syst.* **2023**;65(10):4111–4156. doi:10.1007/s10115-023-01865-y
48. A diabetic retinopathy classification method based on improved InceptionV3 neural network model | IEEE conference publication | IEEE xplore. Available from: <https://ieeexplore.ieee.org/document/10326988>. Accessed January 6, 2024.
49. Sailunaz K, Alhajj S, Özyer T, Rokne J, Alhajj R. A survey on brain tumor image analysis. *Med Biol Eng Comput.* **2024**;62(1):1–45. doi:10.1007/s11517-023-02873-4

50. Prakash AM, Aditya S, Diwakar S, Santosh P, Hema N. Tumor detection using deep learning in organs specific to Indian predicament. *IEEE Int Conference Electron, Comput Commun Tech.* 2023;1–6. doi:10.1109/CONECCT57959.2023.10234765
51. Liu M, Qi C, Bao S, et al. Evaluation kidney layer segmentation on whole slide imaging using convolutional neural networks and transformers. *Digit Comput Pathology.* 2023;12933:pp.334–339. doi:10.48550/arXiv.2309.02563
52. Zhao X, Li Y, Wang Q. (2023) exploring new methods in AI for disease diagnosis. *J Health Care Technol.* 2023;12(1):100–110. doi:10.1186/s12938-023-01113-y
53. Ben-Cohen A, Klang E, Diamant I, et al. CT image-based decision support system for categorization of liver metastases into primary cancer sites: initial results. *Acad Radiol.* 2017;24(12):1501–1509. doi:10.1016/j.acra.2017.06.008
54. Ma Y, Lu Q, Wang X, Wang Y, Yuan F, Chen H. Establishment and validation of a nomogram for predicting new fractures after PKP treatment of for osteoporotic vertebral compression fractures in the elderly individuals. *BMC Musculoskelet Disord.* 2023;24(1):728. doi:10.1186/s12891-023-06801-3

## Infection and Drug Resistance

### Publish your work in this journal

Infection and Drug Resistance is an international, peer-reviewed open-access journal that focuses on the optimal treatment of infection (bacterial, fungal and viral) and the development and institution of preventive strategies to minimize the development and spread of resistance. The journal is specifically concerned with the epidemiology of antibiotic resistance and the mechanisms of resistance development and diffusion in both hospitals and the community. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/infection-and-drug-resistance-journal>

**Dovepress**  
Taylor & Francis Group