ORIGINAL RESEARCH

Accurate Differentiation of Spinal Tuberculosis and Spinal Metastases Using MR-Based Deep Learning Algorithms

Shuo Duan [1],*, Weijie Dong[2],*, Yichun Hua[3], Yali Zheng[4], Zengsuonan Ren[5], Guanmei Cao[6], Fangfang Wu[4], Tianhua Rong[1], Baoge Liu[1]

[1]Department of Orthopaedic Surgery, Beijing Tiantan Hospital, Capital Medical University, Beijing, People's Republic of China; [2]Department of Orthopedics, Beijing Chest Hospital, Capital Medical University, Beijing, People's Republic of China; [3]Department of Medical Oncology, Beijing Tiantan Hospital, Capital Medical University, Beijing, People's Republic of China; [4]Department of Respiratory, Critical Care, and Sleep Medicine, Xiang'an Hospital of Xiamen University, School of Medicine, Xiamen University, Xiamen, People's Republic of China; [5]Department of Orthopaedic Surgery, People's Hospital of Hainan Tibetan Autonomous Prefecture, Hainan Tibetan Autonomous Prefecture, Qinghai Province, People's Republic of China; [6]Department of Radiology, Beijing Tiantan Hospital, Capital Medical University, Beijing, People's Republic of China

*These authors contributed equally to this work

Correspondence: Baoge Liu, Department of Orthopaedic Surgery, Beijing Tiantan Hospital, Capital Medical University, No. 119 South 4th Ring West Road, Fengtai District, Beijing, 100070, People's Republic of China, Tel +86 13581521066, Fax +010-59978702, Email baogeliu@hotmail.com

Purpose: To explore the application of deep learning (DL) methods based on T2 sagittal MR images for discriminating between spinal tuberculosis (STB) and spinal metastases (SM).

Patients and Methods: A total of 121 patients with histologically confirmed STB and SM across four institutions were retrospectively analyzed. Data from two institutions were used for developing deep learning models and internal validation, while the remaining institutions' data were used for external testing. Utilizing MVITV2, EfficientNet-B3, ResNet101, and ResNet34 as backbone networks, we developed four distinct DL models and evaluated their diagnostic performance based on metrics such as accuracy (ACC), area under the receiver operating characteristic curve (AUC), F1 score, and confusion matrix. Furthermore, the external test images were blindly evaluated by two spine surgeons with different levels of experience. We also used Gradient-Class Activation Maps to visualize the high-dimensional features of different DL models.

Results: For the internal validation set, MVITV2 outperformed other models with an accuracy of 98.7%, F1 score of 98.6%, and AUC of 0.98. Other models followed in this order: EfficientNet-B3 (ACC: 96.1%, F1 score: 95.9%, AUC: 0.99), ResNet101 (ACC: 85.5%, F1 score: 84.8%, AUC: 0.90), and ResNet34 (ACC: 81.6%, F1 score: 80.7%, AUC: 0.85). For the external test set, MVITV2 again performed excellently with an accuracy of 91.9%, F1 score of 91.5%, and an AUC of 0.95. EfficientNet-B3 came second (ACC: 85.9, F1 score: 91.5%, AUC: 0.91), followed by ResNet101 (ACC:80.8, F1 score: 80.0%, AUC: 0.87) and ResNet34 (ACC: 78.8, F1 score: 77.9%, AUC: 0.86). Additionally, the diagnostic accuracy of the less experienced spine surgeon was 73.7%, while that of the more experienced surgeon was 88.9%.

Conclusion: Deep learning based on T2WI sagittal images can help discriminate between STB and SM, and can achieve a level of diagnostic performance comparable with that produced by experienced spine surgeons.

Keywords: artificial intelligence, deep learning, spinal tuberculosis, spinal metastases, magnetic resonance imaging

Introduction

Spinal tuberculosis (STB) and spinal metastasis (SM) are common spinal disorders characterized by symptoms such as pain, pathological vertebral fractures, spinal kyphosis, and spinal cord compression, potentially leading to neurological deficits or even paralysis.[1–3] STB, also known as Pott's disease, is a benign condition resulting from the hematogenous spread of *Mycobacterium tuberculosis*. The spine is one of the most common sites of extrapulmonary tuberculosis, constituting over 50% of all cases of tuberculosis affecting bones and joints. STB is more prevalent in developing

countries, and early detection and treatment lead to a good prognosis.[4–6] In contrast, SM is a malignant condition that manifests in over 20% of patients diagnosed with cancer. SM requires a diverse range of treatment methodologies, including surgery, radiation therapy, and chemotherapy. However, the prognosis for most individuals with SM is poor.

The similarity in symptomatic presentation between STB and SM often leads to diagnostic ambiguity or delays.[7,8] Furthermore, this overlap could result in the selection of inappropriate treatment strategies, thereby exacerbating the patient's condition. Given the starkly different treatment approaches required for STB and SM, an accurate preoperative diagnosis is essential in informing the selection of the most effective treatment methods. The gold standards for diagnosing STB and SM are *Mycobacterium* culture growth and pathological tissue biopsies, respectively. However, these diagnostic procedures are invasive, potentially leading not only to increased patient discomfort, but also to an elevated risk of further disease dissemination.[9,10] Consequently, prioritizing the development and refinement of non-invasive diagnostic techniques is of paramount importance in the field of spinal disease management.

Spinal magnetic resonance imaging (MRI) allows visualization of the spine, spinal cord, and surrounding soft tissues across multiple sequences, parameters, and planes.[11] This methodology is highly sensitive to changes in tissue signal, making it an excellent tool for locating and characterizing soft tissue lesions. It has become the preferred method for diagnosing spinal diseases.[12] However, early differential diagnosis of STB versus SM can be difficult under some conditions, for example when there is no formation of a paravertebral abscess or involvement of the intervertebral disc, when the lesion is atypical (manifesting as isolated vertebral involvement), or when there is a history of both tuberculosis and tumor.[13–15] Positron Emission Tomography-Computed Tomography (PET/CT) carries good discriminatory power, however its high cost and radiation exposure restrict its application in clinical contexts.[16] In recent years, deep learning (DL) methods have been applied to convert visual features from images into high-dimensional features, which can then be used for disease localization, diagnosis, and discrimination, particularly in the field of medical imaging.[17,18] Prior studies have demonstrated the proficiency of convolutional neural network (CNN) models in accurately classifying lesions in the breast, prostate, kidney, and brain based on MRI scans, achieving high levels of sensitivity and specificity.[19,20] Artificial intelligence (AI)-assisted diagnostic tools could optimize clinical practice by offering an advanced, precise, and patient-friendly diagnostic solution while reducing the need for uncomfortable and risky invasive biopsies. In this study, we evaluated the ability of deep learning methods to discriminate between STB and SM using T2-weighted imaging (T2WI), and compared their performance to that of two spinal surgeons with differing levels of experience in diagnosing these conditions.

Materials and Methods
Study Participants
The study was ethically approved (Ethic number, IRB: #KY2020-073-02) by the Human Investigations Committees at Beijing Tiantan Hospital of the Capital Medical University, and the other three participating sites. All patients signed informed consent for matters regarding participation in the clinical study during hospital admission. We recruited patients diagnosed with spinal metastasis or spinal tuberculosis who were admitted to the four institutions between January 2020 and January 2021. We adopted the following inclusion criteria: (1) age between 18 and 70 years old, (2) STB or SM diagnosis based on pathological findings, (3) presence of complete clinical data and pre-operative MRI. Our exclusion criteria were: (1) concomitant severe infectious diseases or other types of infectious spondylitis, (2) history of spinal trauma or spinal surgery, (3) no pathological diagnosis, and (4) presence of established spinal epidural abscess. The age, gender, and location of the vertebral lesions for each patient were obtained from the electronic medical records.

Image Segmentation and Preprocessing
We obtained complete imaging sequences, including T2WI images. For image preprocessing, we initially employed N4 bias correction and intensity normalization using SimpleITK.[21] To ensure accuracy and reliability, spinal lesions were manually segmented and delineated on T2WI images using Labelme software by a radiologist with 10 years of experience. The segmentations were subsequently reviewed and verified by a spine surgeon with 10 years of experience. In cases of disagreement between the two experts, the segmentations were further evaluated by an additional experienced

surgeon until a consensus was reached. These procedures were repeated until all regions of interest (ROI) for relevant lesions were fully identified. Following this process, we utilized computer software to automatically segment the rectangular images containing the largest lesioned region.

Deep Learning Models

In this research, our aim was to assess the efficacy of deep learning models in distinguishing between STB and SM. To accomplish this, we built models utilizing four widely used neural networks as backbones: Multiscale Vision Transformers V2 (MVITV2), EfficientNet-B3, ResNet101, and ResNet34. We adopted the following training parameters: 200 training cycles, AdamW optimizer, learning rate of 0.0125, and batch size of 32. We leveraged image data from two institutions to train and validate our deep learning models. Patients were randomly assigned to either training or validation sets, with a ratio of 80% to the former and 20% to the latter set. Data from two other institutions were reserved for external testing.

Input images for the model were augmented with random affine transformations, including random translation, rotation, flipping, zooming, and scaling, to expand the dataset by a factor of 20. We used a 5-fold cross-validation technique to assess the diagnostic performance of the models. We analyzed and compared the diagnostic performance of different models using the receiver operating characteristic curve (ROC), the area under the curve (AUC), accuracy (ACC), precision, sensitivity, F1 score, and confusion matrix. We used Gradient-Class Activation Maps (Grad-CAM) to visualize the high-dimensional semantic features of different neural network models.

Manual Evaluation

Image assessment was blind to the clinical features of the associated patients, including medical history, and laboratory tests. Test images were evaluated by two spine surgeons, one with one year of clinical experience (surgeon 1), the other with eight years of experience (surgeon 2). They were tasked with producing a diagnosis of spinal tuberculosis or spinal metastasis. We computed their accuracy and confusion matrix, and compared them with corresponding results from deep learning models.

Statistical Analysis

This study used Python 3.8 language for coding, and Pytorch for developing deep learning networks. The Scikit-learn toolkit was used to implement model comparison and perform related analyses, such as ROC curve plotting and AUC calculation. We defined $p<0.05$ as significant.

Results
Study Participants

We included a total of 121 patients in this study, comprising 54 patients with STB (male/female: 36/18, average age: 44.8 ±14.1 years) and 67 patients with SM (male/female: 36/31, average age: 62.4±12.6 years). Among the 92 patients from Tiantan and Chest hospitals, 42 were diagnosed with STB (28 males, 14 females, average age 46.0±14.1 years, 65 vertebral lesions, 265 lesion images) and 50 with SM (27 males, 23 females, average age 61.8±12.3 years, 75 vertebral lesions, 416 lesion images). We randomly allocated 80% of the patients (34 STB and 40 SM) to the training set, and the remaining 20% (8 STB patients and 10 SM patients) to the internal validation set. Imaging data of 29 patients (12 STB patients with 14 vertebral lesions and 45 lesion images, and 17 SM patients with 19 vertebral lesions and 54 lesion images) from JT and XA hospitals were used for the external testing set. As shown in Table 1 and Table 2, age and gender ratio differed significantly between groups (P<0.05). Patients with lesions in different spinal regions were distributed as follows: 6 cases of cervical vertebrae, 32 cases of thoracic vertebrae, and 29 cases of lumbar vertebrae in the SM group; 4 cases of cervical vertebrae, 24 cases of thoracic vertebrae, and 26 cases of lumbar vertebrae in the STB group. The primary lesions in the SM group were: 37 cases of lung cancer, 8 cases of breast cancer, and 22 cases of other cancers (including liver cancer, prostate cancer, and thyroid cancer).

Table 1 Patient Characteristics Comparison Between the STB and SM Groups

		ALL	STB	SM
Age		54.5 ±15.9	44.8 ±14.1	62.4±12.6
Gender (female/male)		49/72	18/36	31/36
Number of patients in different spinal regions	Cervical	10	4	6
	Thoracic	56	24	32
	Lumbar	55	26	29

**Table 2** Comparison the Clinical Characteristics of Training, Validation and External Test Set

Set	Group	Age	Gender (Female/Male)
Training set All (n=74)		55.4 ±15.3	31/43
	STB (n=34)	46.6 ±14.8	13/21
	SM (n=40)	62.9 ±11.4	18/22
Internal validation set All (n=18)		51.3 ±15.1	6/12
	STB (n=8)	43.5 ±10.9	1/7
	SM (n=10)	57.6 ±15.5	5/5
External test set All (n=29)		54.2 ±17.9	12/17
	STB (n=12)	40.4 ±13.8	4/8
	SM (n=17)	63.9 ±13.7	8/9

Comparison of Diagnostic Accuracy Across Different Deep Learning Models and Experienced Surgeons
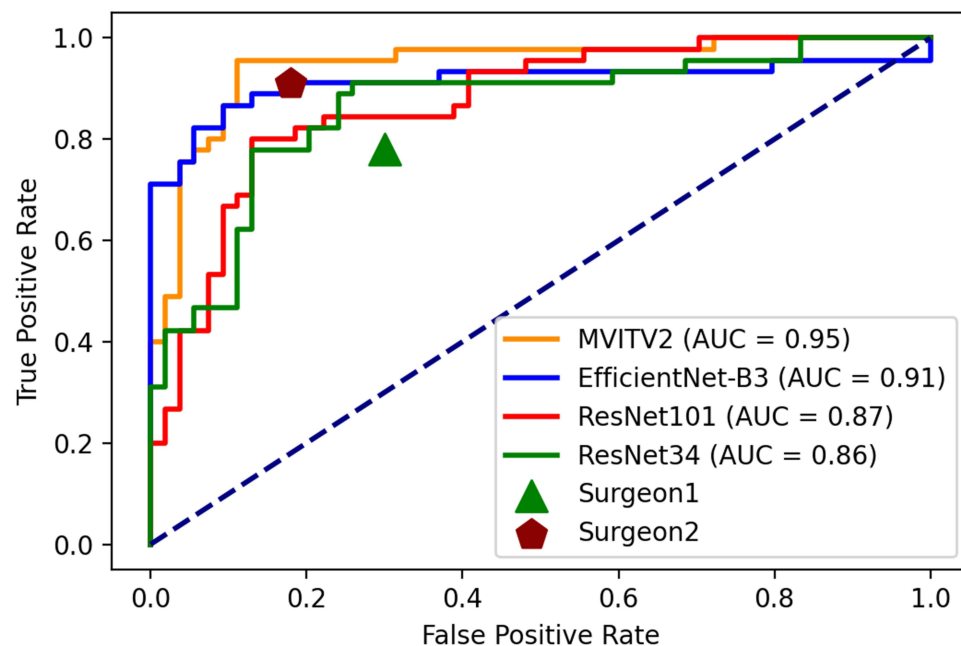
As shown in Table 3 and Figure 1, the MVITV2 model outperformed all others, achieving an accuracy of 98.68%, precision of 98.4%, sensitivity of 98.9%, an F1 score of 98.6%, and an AUC of 0.98. The EfficientNet-B3 model

Table 3 Performances of the DL Models on the Validation and External Test Set

Models	Accuracy	Precision	Sensitivity	F1 Score
Internal validation set				
Resnet34	81.6	80.7	80.7	80.7
Resnet101	85.5	85.0	84.6	84.8
EfficientNet-B3	96.1	95.7	96.2	95.9
MVITV2	98.7	98.4	98.9	98.6
External test set				
Resnet34	78.8	74.0	82.2	77.9
Resnet101	80.8	76.0	84.4	80.0
EfficientNet-B3	85.9	80.4	91.1	91.5
MVITV2	91.9	87.8	95.6	91.5
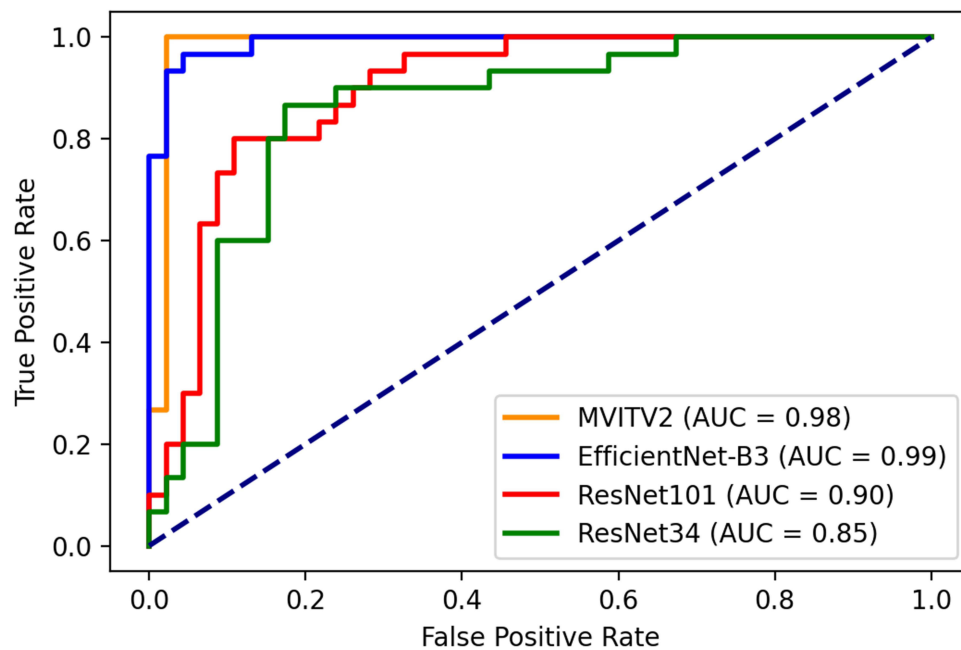
**Figure 1** ROC curves for all models on internal validation set.

Figure 2 ROC curves for all models on external test set and surgeon assessment.

came second with an accuracy of 96.1%, precision of 95.7%, sensitivity of 96.2%, an F1 score of 95.9%, and an AUC of 0.99. It was followed by the ResNet101 model (achieving an accuracy of 85.5%, precision of 85.0%, sensitivity of 84.6%, an F1 score of 84.8%, and an AUC of 0.90) and the ResNet34 model (accuracy of 81.6%, precision of 80.7%, sensitivity of 80.7%, F1 score of 80.7%, and AUC of 0.85).

On the external test set, the MVITV2 model also outperformed all others with an accuracy of 91.9%, precision of 87.8%, sensitivity of 95.6%, F1 score of 91.5%, and an AUC of 0.95, as shown in Table 3 and Figure 2. The EfficientNet-B3 model achieved an accuracy of 85.9%, precision of 80.4%, sensitivity of 91.1%, an F1 score of 91.5%, and an AUC of 0.91. It was followed by the ResNet101 model (achieving an accuracy of 80.8%, precision of 76.0%, sensitivity of 84.4%, F1 score of

80.0%, and AUC of 0.87) and the ResNet34 model (accuracy of 78.8%, precision of 74.0%, sensitivity of 82.2%, F1 score of 77.9%, and AUC of 0.86). Additionally, two spine surgeons independently and blindly diagnosed images from the external test set, achieving ACC values of 73.7% (surgeon 1) and 88.9% (surgeon 2).

The confusion matrix (Figure 3) and ROC curves show that the neural network models were superior to the level of junior resident doctor, and could reach or exceed the level of experienced spine surgeon. We used the Grad-CAM scheme to aid visualization of high-dimensional semantic features associated with neural network models. As shown in Figure 4
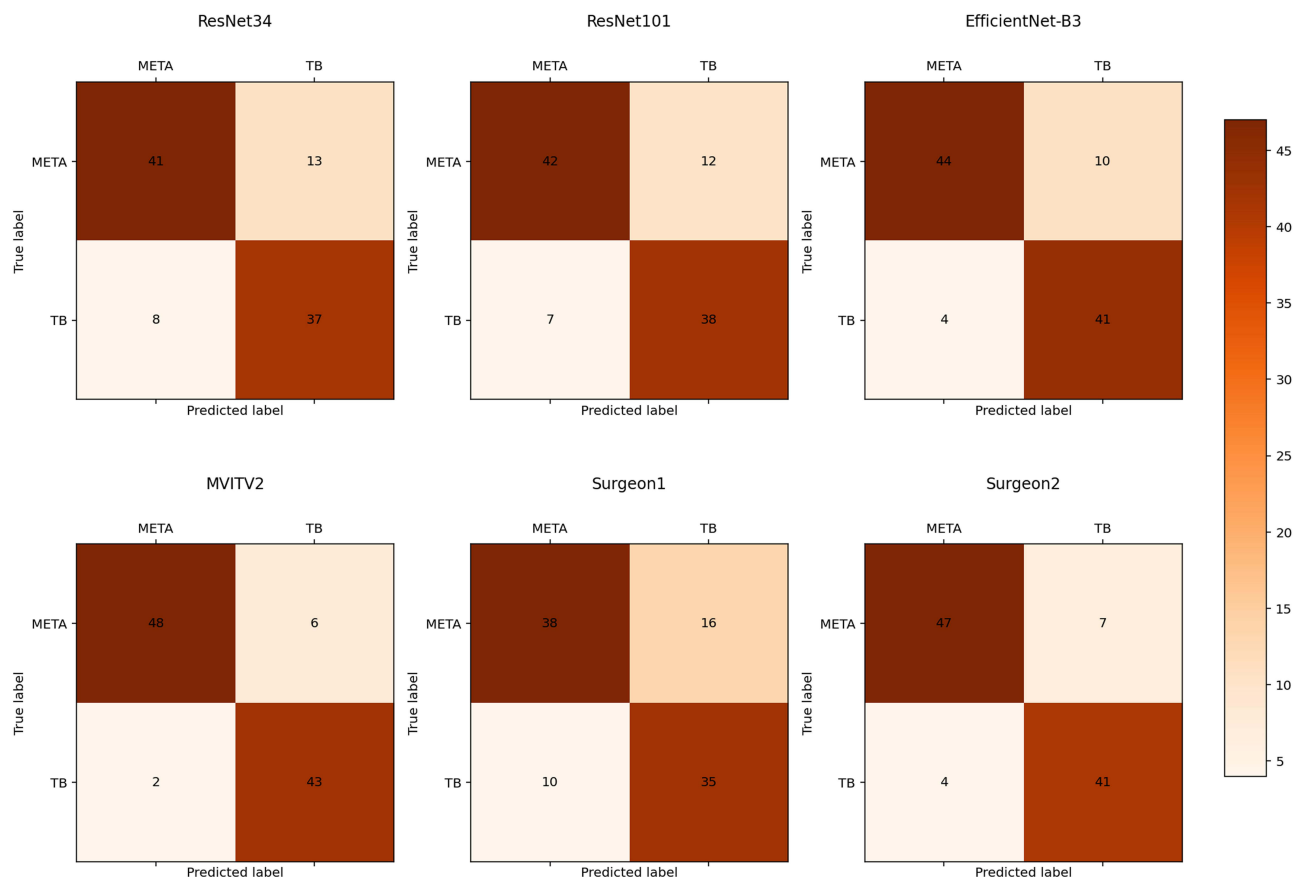
Figure 3 Comparison of results between DL models and surgeons using confusion matrix.
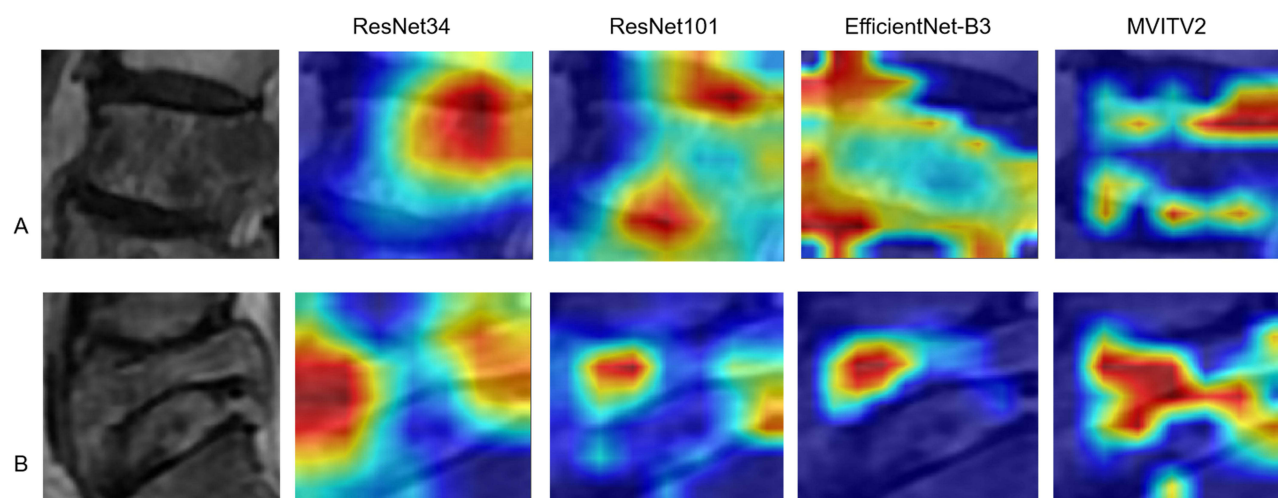
Figure 4 Grad-CAM images of different DL models: (**A**) a SM image, (**B**) a STB image.

higher model accuracy is associated with more precise and focused attention onto the vertebral lesioned area and the intervertebral disc area.

Discussion

Spinal tuberculosis (STB) and spinal metastases (SM) are common diseases in developing countries, with different treatment strategies and outcomes. Early and accurate diagnosis of these conditions is critical for successful intervention. As there are no specific clinical symptoms for STB and SM in the early stages, imaging diagnosis plays an important role in their differentiation. MRI is the preferred advanced imaging technique to assess the suspicious vertebral lesions before proceeding with percutaneous biopsy or other medical intervention. However, for atypical or early lesions, especially those that have not formed abscesses or involved intervertebral discs, vertebral body bone destruction and signal changes represent the main manifestations on MRI, making differential diagnosis difficult.[7,8] In this study, we developed and validated deep learning models for discriminating between STB and SM using information from routine T2WI sagittal-sequence images of the spine. Our results demonstrate that the deep learning model achieved an accuracy of 91.9% and an AUC of 0.95 on the external test set, highlighting the potential clinical utility of this approach in diagnosing spinal pathologies. When compared with two spinal surgeons of different seniority levels, the deep learning models demonstrated diagnostic capabilities that met or exceeded those of experienced spinal surgeons. As a non-invasive technique, deep learning-based algorithm can aid young spinal surgeons or primary care physicians in accurately and rapidly distinguishing between STB and SM.

In recent years, artificial intelligence (AI) methods have been increasingly applied in the medical field, particularly in medical image analysis, with reports of lesion localization, disease diagnosis, tumor classification, and even disease progression prediction models based on deep learning or transfer learning.[17,18,22,23] Liu et al[18] developed a Faster R-CNN model that employs MRI images to differentiate between benign and malignant spinal tumors. Similarly, Marentakis et al[24] demonstrated that deep learning models could predict the histology classification of lung cancer, specifically differentiating between adenocarcinoma and squamous cell carcinoma. Li et al[25] extracted deep learning features from MRI images using ResNet18 and showed that the model can predict the molecular subtypes of gliomas. Furthermore, Chen et al[26] demonstrated that DL models based on attention mechanisms can accurately distinguish between multiple myeloma and spinal metastases. Our recent study showed that deep learning models are capable of identifying the primary tumor source of spinal metastases.[27] Neural network models emulate the structure and function of human neural networks, transforming two-dimensional or three-dimensional medical image data into high-dimensional semantic features. By leveraging these features, neural network models can perform tasks such as lesion classification, localization, and even pixel-level segmentation. These techniques have broad applications in medical image analysis, including tumor detection, organ segmentation, and pathological analysis.[28]

Deep learning models used for image classification diagnosis in medical applications typically belong to supervised learning algorithms.[29] During the training process, image data and associated labels (disease categories) are fed to the model. In this particular study, four popular deep learning network models were selected for training and validation. Among them, ResNet is a CNN that employs residual block structures to address the vanishing gradient problem faced by deep neural networks. ResNet101, with its more intricate structure and increased layer count, has the potential to perform image tasks more effectively than ResNet34.[30] EfficientNet-B3 is a convolutional neural network that scales depth, width, and resolution uniformly using a compound coefficient. This optimization leads to a smaller model size and reduced computation cost while maintaining high accuracy.[31] The MVITV2 model incorporates a multi-head self-attention mechanism to capture richer image features, enabling it to achieve state-of-the-art performance on various visual recognition tasks. The distinct structures of these models result in different high-dimensional features being extracted or selected for the same image sets, which can significantly impact the accuracy of deep learning models. The extracted features are crucial in determining the model's ability to accurately classify and recognize images. In comparison to CNNs, MVITV2 can automatically learn more complex and abstract features from raw image data, making it particularly effective for tasks such as object detection, semantic segmentation, and image classification.

The Grad-CAM method was utilized in this study to visualize the distinct features of each model. We found that models with higher accuracy focused more accurately on the intervertebral disc and vertebral lesion areas. Previous

studies have shown that spinal tuberculosis affects vertebral bodies in more than 98% of cases, resulting in destruction of vertebral bone (47%) or osteolytic changes (34%), often with concomitant involvement of intervertebral discs.[32–34] This is because of the vertebral nutrient artery is a terminal artery, and tuberculosis bacillus can easily infiltrate the vertebral body through the vertebral artery or venous plexus. While Spinal Metastasis (SM) can also manifest as destruction of vertebral bone (osteogenic or osteolytic), it mainly involves the middle and posterior parts of the vertebral bodies and vertebral pedicles. They often appear as flattened bone destruction with increased anteroposterior diameter of the vertebral body, with intervertebral disc involvement being rare. Focused attention of the neural network to key areas of the vertebral body and intervertebral disc is thus critical for successful differentiation.

Previous studies have demonstrated that incorporating clinical features such as gender, age, laboratory indicators, and medical history into the model improves its diagnostic efficiency.[35] However, the inclusion of different types of spinal metastases in various studies, such as the number of breast or prostate cancer patients, can significantly affect the gender ratio and age of the population being studied. Additionally, the age and gender ratio for STB and SM are still controversial.[35–37] Moreover, a history of tumors or tuberculosis is not always reliable: more than 50% of spinal tuberculosis patients present no evidence of pulmonary tuberculosis, and 10–45.6% of SM patients show no clear evidence of primary cancer.[38] Therefore, in our study, we focused more on radiological features and neglected the inclusion of clinical information parameters. It was found that, in the absence of disclosed medical history and laboratory results, the diagnostic accuracy of deep learning models can reach the diagnostic level of experienced spine surgeon.

This study presents the following limitations. First, the generalization ability of our network models maybe limited by the relatively small sample size of our dataset. Further training and validation with larger sample sizes from multiple centers are needed. Second, the neural network models developed in this study did not perform lesion localization and automatic segmentation. In the future, we therefore plan to expand the functional capabilities of our models. Third, this study only included metastatic tumors, leaving out primary spinal tumors, plasmacytoma, and multiple myeloma. In future studies, we will expand our sample size and range to enhance the generalization ability of our models.

Conclusion

Deep learning models can be used to distinguish between spinal tuberculosis and spinal metastasis using T2-weighted sagittal images. The adoption of DL models not only improves the diagnostic accuracy of spinal surgeons, but also enhances work efficiency and reduces medical costs. These findings could potentially facilitate the development of computer-aided diagnostic tools, which may reduce unnecessary referrals from community clinics to specialized centers and limit the need for unnecessary biopsies.

Abbreviations

DL, Deep learning; MVITV2, Multiscale Vision Transformers V2; STB, Spinal tuberculosis; SM, Spinal metastases; MRI, Magnetic resonance imaging; ACC, Accuracy; AUC, Area under the receiver operating characteristic curve; PET/CT, Positron Emission Tomography-Computed Tomography; CNN, Convolutional Neural Network; T2WI, T2-weighted imaging; Grad-CAM, Gradient-Class Activation Maps; ROC, Receiver operating characteristic curve; AI, Artificial intelligence.

Ethics Approval

This study was approved (Ethic number, IRB: #KY2020-073-02) by the Human Investigations Committees at Beijing Tiantan Hospital of the Capital Medical University, and the other three participating sites in accordance with the Declaration of Helsinki. All patients signed informed consent for matters regarding participation in the clinical study during hospital admission.

Consent for Publication

All authors have consented to the publication of the manuscript.

Acknowledgments

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

Disclosure

The authors report no conflicts of interest in this work.

References

1. Jain AK, Kumar J. Tuberculosis of spine: neurological deficit. *Eur Spine J*. 2013;22 Suppl 4(Suppl 4):624–633. doi:10.1007/s00586-012-2335-7
2. Barzilai O, Fisher CG, Bilsky MH. State of the Art Treatment of Spinal Metastatic Disease. *Neurosurgery*. 2018;82(6):757–769. doi:10.1093/neuros/nyx567
3. Shetty A, Kanna RM, Rajasekaran S. TB spine—Current aspects on clinical presentation, diagnosis, and management options. *Semin Spine Surg*. 2016;28(3):150–162. doi:10.1053/j.semss.2015.07.006
4. Dunn RN, Ben Husien M. Spinal tuberculosis: review of current management. *Bone Joint J*. 2018;100-B(4):425–431. doi:10.1302/0301-620x.100b4.Bjj-2017-1040.R1
5. Galgano M, Fridley J, Oyelese A, et al. Surgical management of spinal metastases. *Expert Rev Anticancer Ther*. 2018;18(5):463–472. doi:10.1080/14737140.2018.1453359
6. Zhang HR, Qiao RQ, Yang XG, et al. A multicenter, descriptive epidemiologic survey of the clinical features of spinal metastatic disease in China. *Neurol Res*. 2020;42(9):749–759. doi:10.1080/01616412.2020.1773630
7. Zheng CY, Liu DX, Luo SW, et al. Imaging presentation highly manifested as tuberculosis in a case of spinal metastatic carcinoma. *Orthopedics*. 2011;34(8):e436–438. doi:10.3928/01477447-20110627-32
8. Pu F, Feng J, Yang L, et al. Misdiagnosed and mismanaged atypical spinal tuberculosis: a case series report. *Exp Ther Med*. 2019;18(5):3723–3728. doi:10.3892/etm.2019.8014
9. Filippiadis D, Mazioti A, Kelekis A. Percutaneous, Imaging-Guided Biopsy of Bone Metastases. *Diagnostics*. 2018;8(2). doi:10.3390/diagnostics8020025
10. Piccioli A, Maccauro G, Spinelli MS, et al. Bone metastases of unknown origin: epidemiology and principles of management. *J Orthop Traumatol*. 2015;16(2):81–86. doi:10.1007/s10195-015-0344-0
11. Eweje FR, Bao B, Wu J, et al. Deep Learning for Classification of Bone Lesions on Routine MRI. *EBioMedicine*. 2021;68:103402. doi:10.1016/j.ebiom.2021.103402
12. Patel A, James SL, Davies AM, et al. Spinal imaging update: an introduction to techniques for advanced MRI. *Bone Joint J*. 2015;97-B(12):1683–1692. doi:10.1302/0301-620x.97b12.36164
13. Naim Ur R, El-Bakry A, Jamjoom A, et al. Atypical forms of spinal tuberculosis: case report and review of the literature. *Surg Neurol*. 1999;51(6):602–607. doi:10.1016/s0090-3019(98)00101-3
14. Kumaran SP, Thippeswamy PB, Reddy BN, et al. An Institutional Review of Tuberculosis Spine Mimics on MR Imaging: cases of Mistaken Identity. *Neurol India*. 2019;67(6):1408–1418. doi:10.4103/0028-3886.273630
15. Gao X, Ye XD, Yuan Z, et al. Non-contiguous spinal tuberculosis with a previous presumptive diagnosis of lung cancer spinal metastases. *Eur J Cardiothorac Surg*. 2014;45(5):e178. doi:10.1093/ejcts/ezu047
16. Vorster M, Sathekge MM, Bomanji J. Advances in imaging of tuberculosis: the role of ^{18}F-FDG PET and PET/CT. *Curr Opin Pulm Med*. 2014;20(3):287–293. doi:10.1097/mcp.0000000000000043
17. Biamonte E, Levi R, Carrone F, et al. Artificial intelligence-based radiomics on computed tomography of lumbar spine in subjects with fragility vertebral fractures. *J Endocrinol Invest*. 2022;45(10):2007–2017. doi:10.1007/s40618-022-01837-z
18. Liu H, Jiao M, Yuan Y, et al. Benign and malignant diagnosis of spinal tumors based on deep learning and weighted fusion framework on MRI. *Insights Imaging*. 2022;13(1):87. doi:10.1186/s13244-022-01227-2
19. Song Y, Zhang YD, Yan X, et al. Computer-aided diagnosis of prostate cancer using a deep convolutional neural network from multiparametric MRI. *J Magn Reson Imaging*. 2018;48(6):1570–1577. doi:10.1002/jmri.26047
20. Xi IL, Zhao Y, Wang R, et al. Deep Learning to Distinguish Benign from Malignant Renal Lesions Based on Routine MR Imaging. *Clin Cancer Res*. 2020;26(8):1944–1952. doi:10.1158/1078-0432.Ccr-19-0374
21. Lowekamp BC, Chen DT, Ibáñez L, et al. The Design of SimpleITK. *Front Neuroinform*. 2013;7:45. doi:10.3389/fninf.2013.00045

22. Park T, Yoon MA, Cho YC, et al. Automated segmentation of the fractured vertebrae on CT and its applicability in a radiomics model to predict fracture malignancy. *Sci Rep*. 2022;12(1):6735. doi:10.1038/s41598-022-10807-7

23. Li Z, Wang Y, Yu J, et al. Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma. *Sci Rep*. 2017;7(1):5467. doi:10.1038/s41598-017-05848-2

24. Marentakis P, Karaiskos P, Kouloulias V, et al. Lung cancer histology classification from CT images based on radiomics and deep learning models. *Med Biol Eng Comput*. 2021;59(1):215–226. doi:10.1007/s11517-020-02302-w

25. Li Y, Wei D, Liu X, et al. Molecular subtyping of diffuse gliomas using magnetic resonance imaging: comparison and correlation between radiomics and deep learning. *Eur Radiol*. 2022;32(2):747–758. doi:10.1007/s00330-021-08237-6

26. 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

27. Duan S, Cao G, Hua Y, et al. Identification of origin for spinal metastases from MRI images: comparison between radiomics and deep learning methods. *World Neurosurg*. 2023. doi:10.1016/j.wneu.2023.04.029

28. Chen X, Wang X, Zhang K, et al. Recent advances and clinical applications of deep learning in medical image analysis. *Med Image Anal*. 2022;79:102444. doi:10.1016/j.media.2022.102444

29. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60–88. doi:10.1016/j.media.2017.07.005

30. He F, Liu T, Tao D. Why ResNet Works? Residuals Generalize. *IEEE Trans Neural Netw Learn Syst*. 2020;31(12):5349–5362. doi:10.1109/tnnls.2020.2966319

31. Huang C, Wang W, Zhang X, et al. Tuberculosis Diagnosis using Deep Transferred EfficientNet. *IEEE/ACM Trans Comput Biol Bioinform*. 2022:1–9. doi:10.1109/tcbb.2022.3199572

32. Teo EL, Peh WC. Imaging of tuberculosis of the spine. *Singapore Med J*. 2004;45(9):439–444.

33. Garg RK, Somvanshi DS. Spinal tuberculosis: a review. *J Spinal Cord Med*. 2011;34(5):440–454. doi:10.1179/2045772311y.0000000023

34. Jain AK, Rajasekaran S, Jaggi KR, et al. Tuberculosis of the Spine. *J Bone Joint Surg Am*. 2020;102(7):617–628. doi:10.2106/jbjs.19.00001

35. Du X, She Y, Ou Y, et al. A Scoring System for Outpatient Orthopedist to Preliminarily Distinguish Spinal Metastasis from Spinal Tuberculosis: a Retrospective Analysis of 141 Patients. *Dis Markers*. 2021;2021:6640254. doi:10.1155/2021/6640254

36. Wang P, Liao W, Cao G, et al. Characteristics and Management of Spinal Tuberculosis in Tuberculosis Endemic Area of Guizhou Province: a Retrospective Study of 597 Patients in a Teaching Hospital. *Biomed Res Int*. 2020;2020:1468457. doi:10.1155/2020/1468457

37. Dheda K, Barry CE, Maartens G. Tuberculosis. *Lancet*. 2016;387(10024):1211–1226. doi:10.1016/s0140-6736(15)00151-8

38. Momjian R, George M. Atypical imaging features of tuberculous spondylitis: case report with literature review. *J Radiol Case Rep*. 2014;8(11):1–14. doi:10.3941/jrcr.v8i11.2309