



A deep neural network for MRI spinal inflammation in axial spondyloarthritis

Yingying Lin¹ · Shirley Chiu Wai Chan³ · Ho Yin Chung^{3,4} · Kam Ho Lee² · Peng Cao¹

Received: 31 July 2023 / Revised: 3 November 2023 / Accepted: 9 December 2023
© The Author(s) 2024

Abstract

Objective To develop a deep neural network for the detection of inflammatory spine in short tau inversion recovery (STIR) sequence of magnetic resonance imaging (MRI) on patients with axial spondyloarthritis (axSpA).

Methods A total 330 patients with axSpA were recruited. STIR MRI of the whole spine and clinical data were obtained. Regions of interests (ROIs) were drawn outlining the active inflammatory lesion consisting of bone marrow edema (BME). Spinal inflammation was defined by the presence of an active inflammatory lesion on the STIR sequence. The 'fake-color' images were constructed. Images from 270 and 60 patients were randomly separated into the training/validation and testing sets, respectively. Deep neural network was developed using attention UNet. The neural network performance was compared to the image interpretation by a radiologist blinded to the ground truth.

Results Active inflammatory lesions were identified in 2891 MR images and were absent in 14,590 MR images. The sensitivity and specificity of the derived deep neural network were 0.80 ± 0.03 and 0.88 ± 0.02 , respectively. The Dice coefficient of the true positive lesions was 0.55 ± 0.02 . The area under the curve of the receiver operating characteristic (AUC-ROC) curve of the deep neural network was 0.87 ± 0.02 . The performance of the developed deep neural network was comparable to the interpretation of a radiologist with similar sensitivity and specificity.

Conclusion The developed deep neural network showed similar sensitivity and specificity to a radiologist with four years of experience. The results indicated that the network can provide a reliable and straightforward way of interpreting spinal MRI. The use of this deep neural network has the potential to expand the use of spinal MRI in managing axSpA.

Keywords MRI · Axial spondyloarthritis · Ankylosing spondylitis · Inflammation · Deep learning · Artificial intelligence · Spine

Yingying Lin and Shirley Chiu Wai Chan contributed equally to this study.

✉ Peng Cao
caopeng1@hku.hk

¹ Department of Diagnostic Radiology, The University of Hong Kong, LG3 Sassoon Road No. 5, Pok Fu Lam, Hong Kong

² Department of Radiology, Queen Mary Hospital, Pok Fu Lam, Hong Kong

³ Division of Rheumatology and Clinical Immunology, Department of Medicine, The University of Hong Kong, Pok Fu Lam, Hong Kong

⁴ Chiron Medical, Central, Hong Kong

Introduction

Axial spondyloarthritis predominately affects the sacroiliac joint (SI joint) and spine by causing inflammation [1]. The disease process can individually affect spinal segments, including the cervical, thoracic, and lumbar spine [2]. Active spinal inflammatory lesions are common and occur in over half of the patients with axSpA [3]. Long-term imaging monitoring is crucial for axSpA management [2, 4]. Magnetic resonance imaging (MRI) is essential in diagnosis and disease activity assessment for axSpA [5]. MRI is a noninvasive imaging tool for detecting and monitoring inflammation and disease activity, independent of other biomarkers [5]. As a fat suppression sequence, the short tau inversion recovery (STIR) sequence could depict the signals of active inflammatory lesions consisting of bone marrow edema (BME) obscured by marrow fat signals. Due to the

high sensitivity of detecting the active inflammatory lesion, STIR MRI is commonly used to identify and grade inflammation in axSpA patients [6–8]. According to the available scoring methods, MRI spinal lesions in axial spondyloarthritis [9] and Spondyloarthritis Research Consortium of Canada spine index [6], identifying the spinal inflammation is the first step to score the inflammatory degree of spinal segments. In addition, identifying spinal inflammation has important diagnostic, prognostic, and therapeutic implications [10]. However, the interpretation of MRI is labor intensive, requiring the expertise of specialized personnel, yet variability in interpretation exists even between experienced specialists [11].

Deep learning, a subfield of machine learning, has achieved wide applications in different areas of medical imaging analysis [12]. With the increasing popularity of deep learning for medical imaging analysis, many axSpA studies have applied such a technique. Deep learning in MRI interpretation may be the next crucial step in enabling the widespread application of MRI in managing axSpA, especially in places where expertise is limited. As sacroiliitis on MRI is vital for axSpA, many studies focused on applying deep learning models for sacroiliitis. These studies included various aims like detection of erosion and ankylosis on SI joint CT [13], identification of sacroiliitis [14, 15] or bone marrow edema (BME) of the sacroiliac (SI) joint [16], and detecting the changes of sacroiliitis in MR images on axSpA patients [17]. However, apart from inflammatory structural changes of SI joint, inflammation in the spine could impact physical function [18, 19]. Therefore, early detection of spinal inflammation could assist in the diagnosis of axSpA [2], monitor the disease progress [2], and analyze the correlation of MRI signs with low back pain [20]. Several recent studies explored the feasibility of deep learning on spinal inflammation. These studies focused on images from PET/CT [20], radiographs [21], or the assessment of intervertebral disk (IVD) degeneration in spinal MR images [22]. To our knowledge, no studies tackled the design challenges in identifying inflammation in spinal STIR MRI via deep learning.

Utilizing the attention UNet [23], a U-shaped architecture designed for medical images, and the attention gate (AG) [23] highlighting the regions of interest, we have recently developed a deep neural network for the interpretation of SI MRI [15] by identifying the sacroiliitis. With the increasing importance of spinal MRI interpretation for managing axSpA patients, differentiating the spinal segments with or without inflammation becomes crucial. Therefore, this study aims to develop a deep neural network to identify inflammatory spine on STIR MRI among patients with axSpA.

Methods

The study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (reference number UW 14-085) and local ethics committees.

Deep neural network was developed using STIR MRI of spinal inflammatory lesions from a large prospective cohort designed to investigate clinical applications of MRI in axSpA. Participants with an expert diagnosis of axSpA were consecutively recruited from ten public hospitals in Hong Kong (Queen Mary Hospital, Tung Wah Hospital, Grantham Hospital, Pamela Youde Nethersole Eastern Hospital, Caritas Medical Centre, Tseung Kwan O Hospital, Kwong Wah Hospital, Hong Kong Eye Hospital, Prince of Wales Hospital, and Prince Margaret Hospital) and one rheumatology center in China (University of Hong Kong-Shenzhen Hospital) from April 2014 to April 2021. Participants with pregnancy and inability to undergo MRI scans were excluded. All participants gave written consent before recruitment. Demographic data, including age, sex, ethnicity, smoking, and drinking status, were documented.

MRI acquisition

STIR sequence of the whole spine was obtained using a 3T MR imaging unit (Achieva; Philips Healthcare, Best, the Netherlands). The technical parameters were set as below: repetition times/echo times = 5000/80, fields of view = 150×249 mm², slice thicknesses = 3.5 mm, and acquisition time = 2.48 min. Due to the limited matrix size for each MRI scan, spinal segments were scanned independently, including the cervical, thoracic, and lumbar spine. The spine was covered entirely. Sagittal slices of these individual spinal segments were used to develop the deep neural network. The spinal images from each patient range between 17 and 19 slices per one spinal segment.

Ground-truth MRI interpretation

A rheumatologist and a radiologist with 10- and 6-year experiences in axSpA MRI identified the active inflammatory lesions consisting of bone marrow edema (BME) in the whole spine MRI, according to the Assessment of SpondyloArthritis international Society (ASAS) definition [1]. Spinal inflammation was defined by BME in spinal STIR MRI. The presence and absence of BME in STIR MRI were classified as with and without spinal inflammation, respectively. Different spinal segments, cervical, thoracic, and lumbar spinal segments, were evaluated individually. Discordant interpretations were resolved by consensus between these

two readers. Rheumatologist outlined the active inflammatory lesions, which were set as the ground-truth regions of interest (ROIs) after two readers agreed to the ROIs.

Data preprocessing

A binary labeling system was used to categorize with or without spinal inflammation. A 'fake-color' image comprises of three consecutive slices, which we separate into red–green–blue (RGB) channels. We take the preceding slice in the R-channel, the current slice in G-channel, and the subsequent slice in B-channel. The middle channel (G-channel) of the current slice forms the ground-truth mask of the 'fake-color' images (see Fig. 1). For each MR image, we create a set of 'fake-color' images.

Training, validation, and testing of deep neural network

Participants were classified into two categories based on the presence or absence of active spinal inflammation. A total of 300 participants with active spinal inflammation and 30 without active spinal inflammation were included (Fig. 2). Participants were assigned to (1) the training and validation set consisting of 270 participants with active spinal inflammation and (2) the testing set consisting of 30 participants with active spinal inflammation and 30 participants without active spinal inflammation. Participants were randomly split into training/validation and testing sets. Individual participants only appeared in training and validation, or the testing set.

The training and validation set consisted of a total of 540 spinal segments, which contained a total of 2665 images with inflammation. Additionally, there were 270 spinal segments included, which contained 11,807 images without inflammation from 270 individual participants.

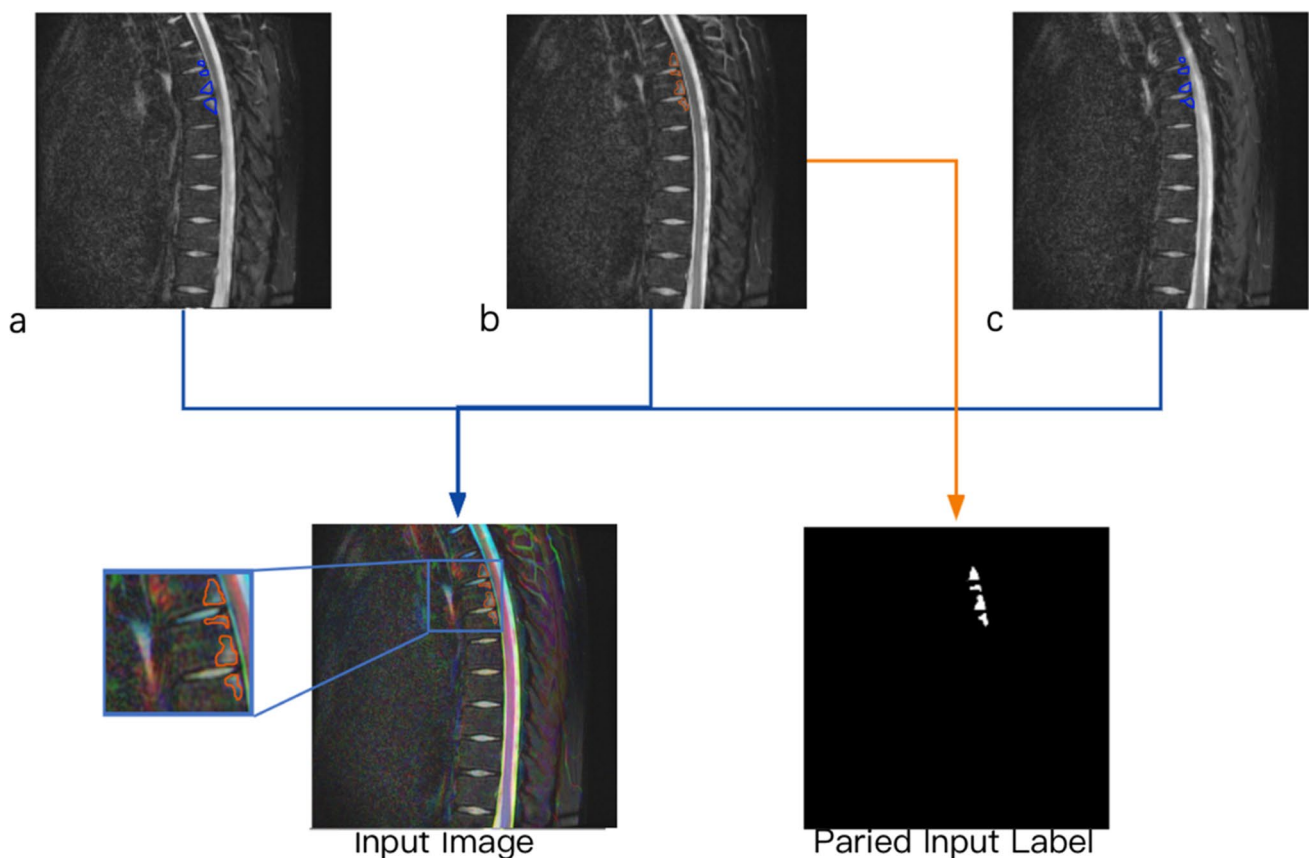


Fig. 1 Process of generating a 'fake-color' input image and paired input label, outlined in orange. The first row includes three consecutive STIR images of cervical segments. The blue arrow highlights the synthesis of the 'fake-color' image by placing consecutive STIR images of the first row into the R-, G-, and B-channels, respectively.

The input label was paired with image b (orange arrow). The blue outline shows the label in the preceding (a) and subsequent (c) slices, which were not the input label. The blue box represents the zoom-in view

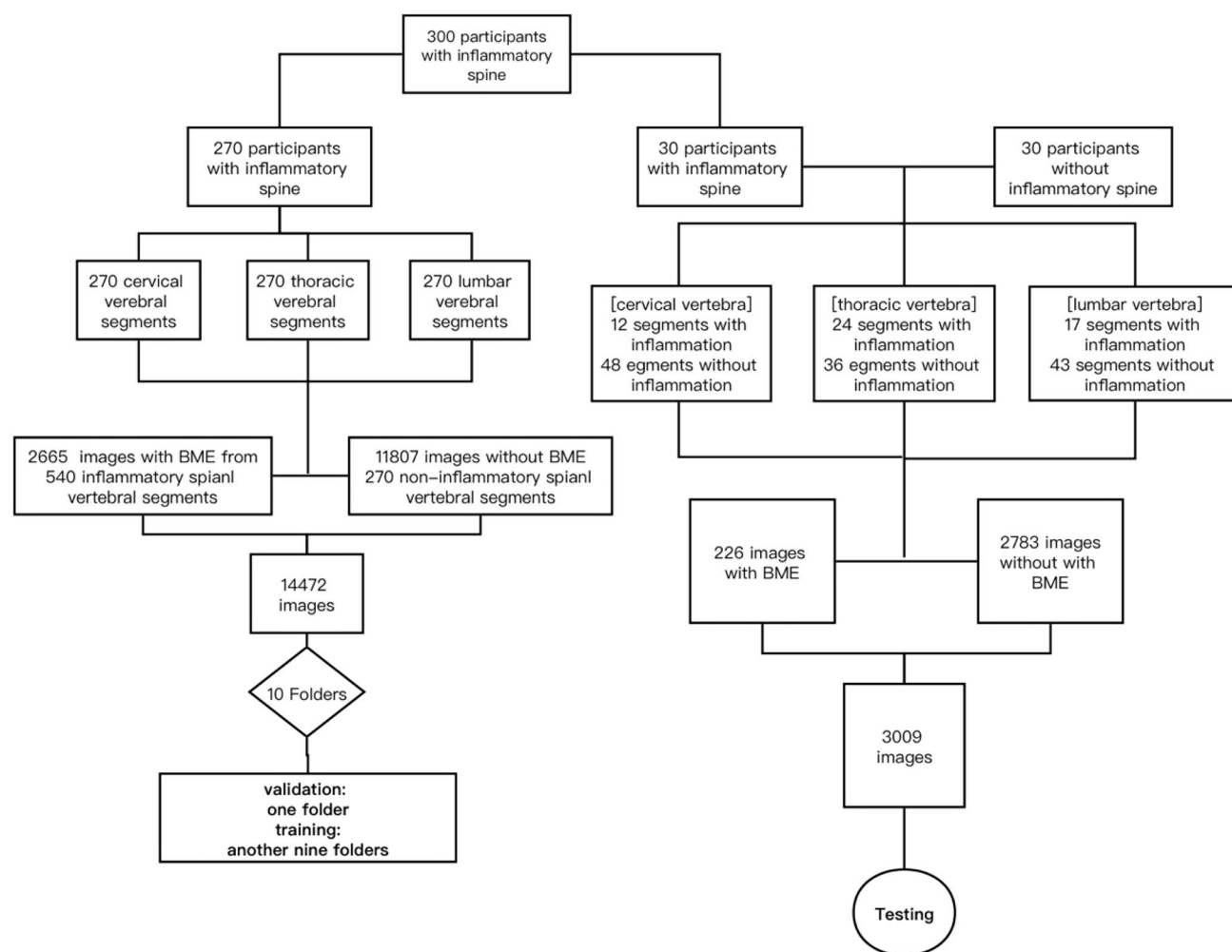


Fig. 2 Data distribution of training, validation, and testing sets for developing the deep neural network

Deep neural network built upon UNet algorithm with AGs was implemented (Fig. 3). The technical details were summarized in our previous publication [7]. A tenfold cross-validation method was used to increase the validity of the deep neural network. Images from the training and validation sets were randomly split into ten folders. Then, the training process was repeated ten times. In each cycle, images from one folder were used for validation, and images from the remaining nine folders were used for training.

The testing set included 53 spinal segments (226 images) with inflammation and 127 spinal segments (2783 images) without inflammation from 60 participants. The testing set was used to infer the final performance of the deep neural network. The performance was evaluated at both the image level and spinal segment level. At the image level, the deep neural network prediction of inflammation in an image was determined as image with inflammation. In contrast, at spinal segment levels, the deep neural network prediction of

inflammation in at least two slices in a spinal segment was defined as spinal segment with inflammation.

Manual labeling

A 4-year experienced radiologist (2 years in musculoskeletal MRI), blinded to the ground-truth masks, identified the BME in the testing set based on ASAS definition of inflammatory spine. Then, the performance of the radiologist was evaluated at image and spinal segment levels using the same standard.

Deep learning neural network

Attention UNet was implemented using TensorFlow-GPU 2.5 and Keras 2.7.0. The input was the ‘fake-color’ image with paired ground-truth BME mask. The output was the predicted BME mask. Only images where the predicted BME overlapped with the ground-truth BME were defined

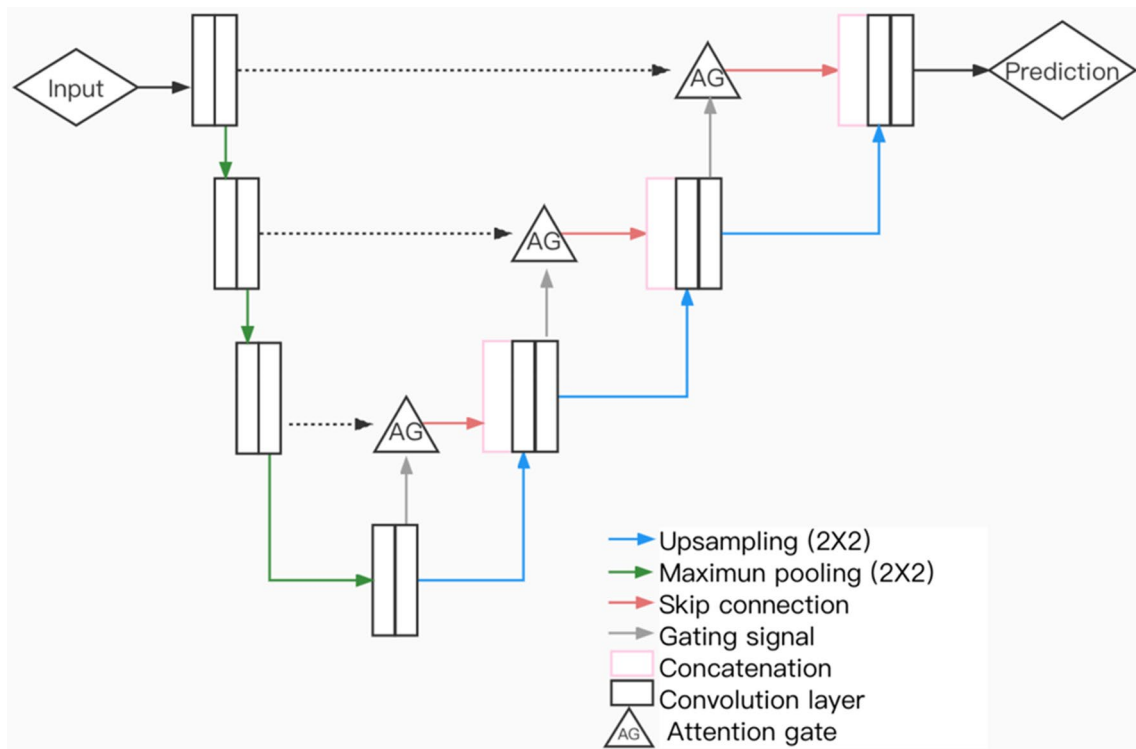


Fig. 3 Architecture of attention UNet

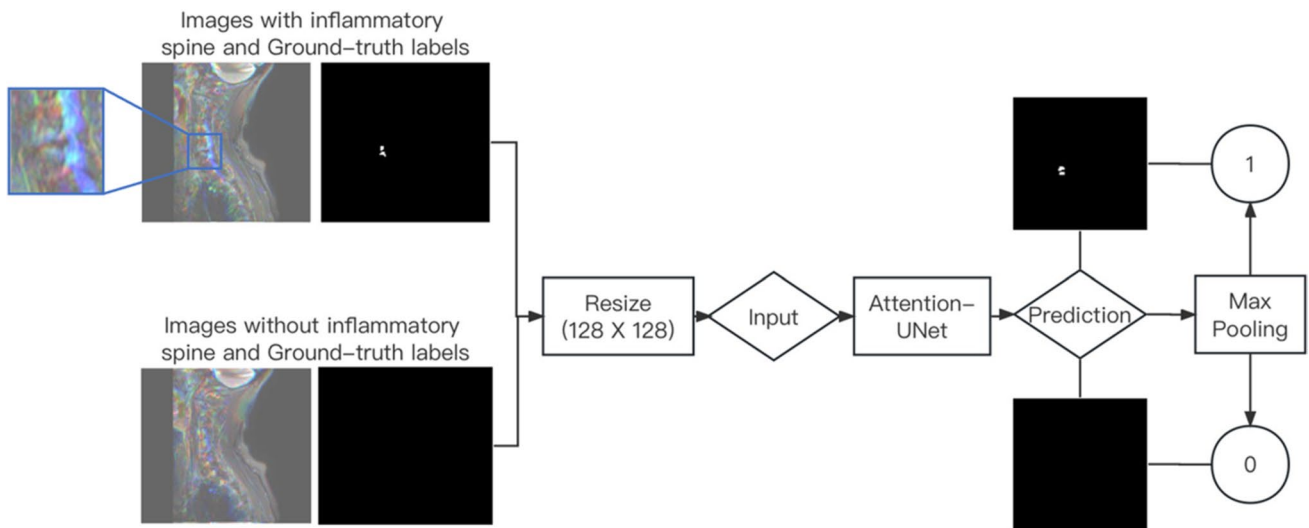


Fig. 4 Flowchart of the process of developing the deep neural network. 'Fake-color' images with (images in first row of the left part) or without (images in second row of the left part) spinal inflammation and their paired labels were resized and then input to the atten-

tion UNet. After training, the developed deep neural network gave the prediction. Finally, max-pooling was applied to define whether the predicted image was image with spinal inflammation (1) or image without spinal inflammation (0). The blue box was the zoom-in view

as images with inflammation (1), while the other images were defined as images without inflammation (0). Please refer to Fig. 4 for a flowchart of the training process.

Statistical analysis

Continuous variables were expressed as mean with standard deviation. The kappa coefficient was used to demonstrate the inter-reader reliability between two readers. The degree of reliability was interpreted as 0.41–0.60, moderate; 0.61–0.80, substantial; and 0.81–1.00, almost perfect.

The performance of the deep neural network was evaluated using the area under the curve (AUC) of receiver operating curve (ROC) according to the probability of the presence of lesion. Sensitivity and specificity were calculated. The spatial accuracy of the automated segmentation of MR images was assessed using the Dice coefficient.

All statistics were performed with IBM SPSS Statistics V27. Listwise deletions were performed for missing values.

Results

A total of 330 patients with axSpA were recruited. Characteristics of patients in the training and validation cohort are summarized in Table 1. Two experienced readers performed MRI interpretation with reasonably high inter-reader reliability and a kappa coefficient of 0.85. Training and validation of the deep neural network for identifying active spinal inflammation were robust according to the result of the tenfold cross-validation. The sensitivity (0.83 ± 0.020) and specificity (0.85 ± 0.026) at the image level during each tenfold cross-validation exhibited minimal fluctuations.

The performance of the deep neural network and that of a radiologist were evaluated in the testing set, as shown in

Table 1 Baseline characteristics of the training and testing cohort

	Training cohort (N=270)	Testing cohort (N=60)		p value
Age	44.6 ± 17.8 (18–86)	45.6 ± 14.0 (19–82)		0.323
Male sex	174/270 (64.4%)	27/60 (45.0%)	0.45 (0.26–0.80)	0.005
Chinese	264/270 (97.8%)	60/60 (100%)	1.23 (1.17–1.30)	0.244
Age of onset	31.6 ± 12.9	33.3 ± 14.4		0.202
Back pain duration	14.0 ± 14.6	12.4 ± 11.3		0.210
Smoker	88/267 (33.0%)	13/59 (22.0%)	0.58 (0.30–1.12)	0.101
Drinker	30/261 (11.5%)	6/56 (10.7%)	0.92 (0.37–2.34)	0.924
HLA-B27	195/252 (77.4%)	34/62 (65.4%)	0.55 (0.29–1.05)	0.068
CRP	0.89 ± 1.90	1.42 ± 3.62		0.140
ESR	31.6 ± 23.8	34.9 ± 27.0		0.198
BASDAI	4.7 ± 2.2	4.7 ± 2.2		0.398
BASFI	3.0 ± 2.4	3.0 ± 2.4		0.278
Sulfasalazine	62/263 (23.6%)	13/59 (22.0%)	0.92 (0.47–1.81)	0.800
Other csDMARDs	32/265 (12.1%)	9/60 (15.0%)	1.29 (0.58–2.86)	0.538

Table 2 Sensitivity and specificity of deep neural network and radiologist in image level and of individual spinal vertebral part

	Sensitivity (true positive/positive)	Specificity (true negative/negative)
	<i>Radiologist</i>	
Level of image	0.82	0.87
Level of individual spinal vertebral part	0.87	0.70
	<i>Deep neural network</i>	
Level of image	0.80 ± 0.03	0.88 ± 0.02
Level of individual spinal vertebral part	0.85 ± 0.04	0.73 ± 0.03
	<i>Deep neural network in individual spinal segment</i>	
Cervical vertebra	0.75 ± 0.02	0.81 ± 0.02
Thoracic vertebra	0.91 ± 0.04	0.62 ± 0.04
Lumbar vertebra	0.82 ± 0.03	0.74 ± 0.02

Table 3 Confusion matrix of active inflammatory lesion prediction by deep neural network per image on the testing set

Ground truth	Prediction	
	With active inflammatory lesions	Without active inflammatory lesions
With active inflammatory lesions	181	45
Without active inflammatory lesions	294	2489

Table 2. The deep neural network demonstrated relatively high sensitivity and specificity at both image and spinal segment levels. The mean sensitivity was 0.80 ± 0.03 at the image level and 0.85 ± 0.02 at the spinal segment level. The mean specificity was 0.88 ± 0.02 at the image level and 0.73 ± 0.03 at the spinal segment level. Confusion matrices of lesion prediction per image are shown in Table 3. The AUC-ROC of the deep neural network was 0.87 ± 0.02 (Fig. 5). The performance of the deep neural network was comparable to a radiologist.

When evaluated based on individual spinal segments (cervical, thoracic, and lumbar), the sensitivity of the deep neural network was highest in the thoracic spine (with sensitivity = 0.90 ± 0.04 and specificity = 0.62 ± 0.04), followed by the lumbar spine (with sensitivity = 0.82 ± 0.03 and specificity = 0.72 ± 0.02) and cervical spine (sensitivity = 0.75 ± 0.02 and specificity = 0.81 ± 0.02). Figure 6 illustrates the different prediction scenarios of the developed deep neural network with reference to the ground

truth. Various lesions were present at the cervical, thoracic, and lumbar spine. The Dice coefficient of the true positive lesions was 0.55 ± 0.02 .

Discussion

Utilizing attention UNet algorithm and 'fake-color' image processing to simulate the interpretation of consecutive images, a deep neural network with good sensitivity and specificity for identifying spinal inflammation in axSpA was firstly developed to the best of our knowledge. The deep neural network performance was comparable to a radiologist with similar sensitivity and specificity at both image level and spinal segment level possessing the potential to assist physicians' interpretation of spinal MRI in axSpA. Furthermore, the satisfied performance of the deep neural network indicated the potential to aid the broader usage of spinal MRI in the management of axSpA. The AUC of the developed deep neural network in this study demonstrated a satisfactory performance compared to other studies [13].

The deep neural network demonstrated higher sensitivity when interpretation was based on spinal segments compared to image level. Similarly, the difference in sensitivity and specificity at the image and spinal segment levels was also observed in image interpretation by the radiologist, who served as the comparator in our study. Determination of inflammation at the spinal segment levels is more clinically relevant as disease activity is usually interpreted based on the overall evaluation of multiple images and lesions.

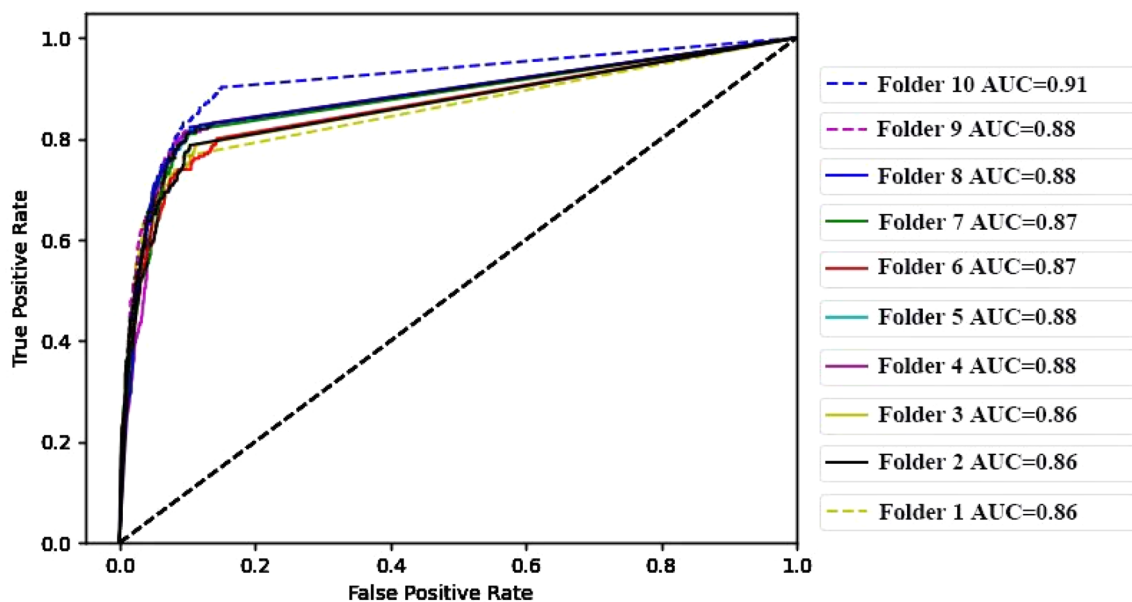
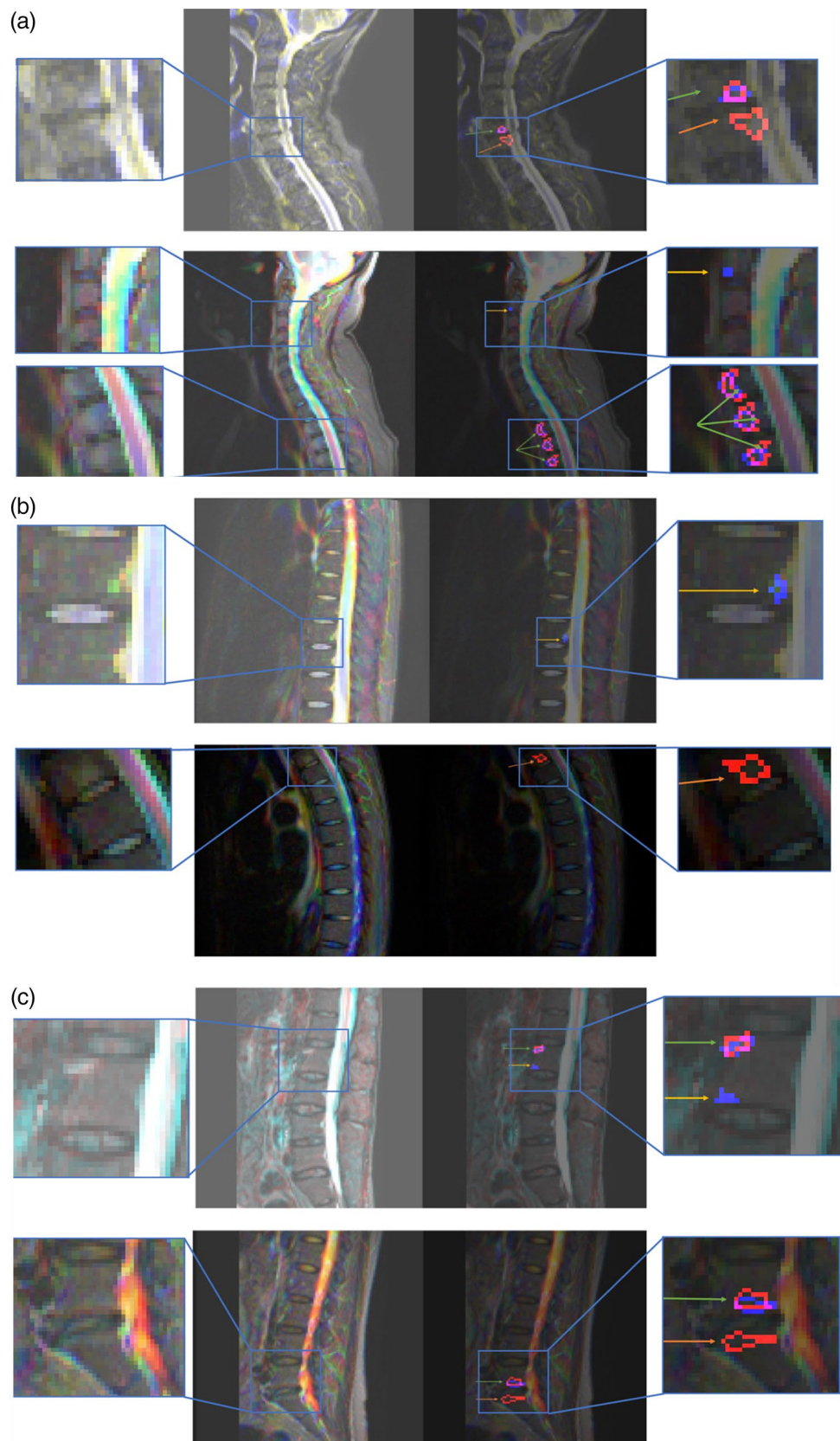
**Fig. 5** ROC curve of the deep neural network

Fig. 6 Examples of the developed deep neural network with image size (128×128). Left was 'fake-color' input image. Right was the ground-truth lesions' outline (red), the predicted lesions' outline (blue), and their overlap outline (rose-red) on 'fake-color' image. The common preprocess of the input, as the normalization, caused the intensity difference between left and right. **a** Two examples of the cervical vertebra. **b** Two examples of a thoracic vertebra. **c** Two examples of lumbar cervical. The green, yellow, and orange arrows pointed out the true positive, false positive, and false negative lesions, respectively. The blue boxes were the zoom-in views



Inflammatory lesions were found at variable frequencies in axSpA depending on the spinal segments and were most common in the thoracic spinal segment due to inherent biomechanics [3]. Therefore, inflammation identified at the thoracic spine tends to be more specific for disease activity and may aid the diagnosis. The deep neural network developed in the current study had the highest sensitivity in identifying thoracic spinal inflammation. Hence, the deep neural network was of clinical relevance and applicability for axSpA.

The 'fake-color' input system by using the information of consecutive images was proved to have better performance as it simulates the real-world MRI interpretation that human reader would compare the consecutive images.[15]. Based on the 'fake-color' image input system providing additional information from adjacent images, our developed deep neural network was comparable to a radiologist. This method may become the next crucial step for the widespread application of MRI in the clinical management of patients with axSpA.

The data imbalance existed as the total number of images with spinal inflammation was far less than the total number of images without spinal inflammation in participants with spinal inflammation. To avoid a severe data imbalance in training, we only included participants without inflammation in the testing set. This helped us evaluate the applicability of the deep neural network on both participants with and without spinal inflammation. The lack of participants without inflammation in training led to a loss in specificity. However, the specificity was similar to the specificity of radiologist.

Our study has several limitations. The ground-truth masks were established by two investigators and contributed to potential bias. This may be overcome by increasing the number of readers. That said, the inter-reader reliability between the two investigators was reasonably high (0.85), which outperforms other studies such as [24] that reported 0.75 inter-reader reliability and 0.8 intra-reader reliability. Furthermore, we expected minimal bias in our study. The relatively low Dice indicated that the deep neural network could not outline the inflammatory lesion precisely. However, this study aimed to identify spinal inflammation rather than the precise outline of the inflammatory lesion. Finally, this study has only proved the satisfactory performance of the deep neural network in identifying the inflammatory spine by evaluating BME in the spine with a binary label, establishing the basis for a SPARCC system rather than output the SPARCC score. It serves as a proof-of-concept study for the potential application of deep neural network in spinal MRI interpretation for axSpA. Future studies are anticipated to develop more advanced deep neural networks or tools for outputting SPARCC score. In addition, external validation from multicenter studies with different MRI modes is necessary in future research. Our team is currently conducting external validation in other cohorts, including patients of different ethnicities and clinical presentations.

Conclusion

A deep neural network was developed to detect spinal inflammation in axSpA. The performance of this deep neural network was comparable to a 4-year experienced radiologist, providing an easy and reliable way to interpret spinal MRI.

Acknowledgements This work is supported by the Hong Kong Society of Rheumatology, Novartis Research, and seed funds from the University of Hong Kong.

Declarations

Conflict of interest All authors have no conflicts of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Sieper J et al (2009) The Assessment of SpondyloArthritis international Society (ASAS) handbook: a guide to assess spondyloarthritis. *Ann Rheum Dis* 68(Suppl 2):ii1–ii7
2. Aiyer S, Udar S, Kharat A, Bhilare P (2022) P Sancheti Utility of selected sequence MRI imaging of the axial skeleton in the diagnosis of axial spondyloarthritis. *J Clin Orthop Trauma* 32:101983–101983
3. Baraliakos X et al (2005) Inflammation in ankylosing spondylitis: a systematic description of the extent and frequency of acute spinal changes using magnetic resonance imaging. *Ann Rheum Dis* 64(5):730–734
4. Braun J et al (2011) 2010 update of the ASAS/EULAR recommendations for the management of ankylosing spondylitis. *Ann Rheum Dis* 70(6):896–904
5. Khmelinskii N, Regel A, Baraliakos X (2018) The role of imaging in diagnosing axial spondyloarthritis. *Front Med* 5:106–106
6. Maksymowych WP et al (2005) Spondyloarthritis research consortium of canada magnetic resonance imaging index for assessment of spinal inflammation in ankylosing spondylitis. *Arthritis Rheum-Arthr* 53(4):502–509
7. Weber U, Kissling RO, Hodler J (2007) Advances in musculoskeletal imaging and their clinical utility in the early diagnosis of spondyloarthritis. *Curr Rheumatol Rep* 9(5):353–360
8. Bray TJP et al (2019) Recommendations for acquisition and interpretation of MRI of the spine and sacroiliac joints in the diagnosis of axial spondyloarthritis in the UK. *Rheumatology (Oxford)* 58(10):1831–1838
9. Braun J et al (2003) Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis, before and after successful therapy with infliximab: evaluation of a new scoring system. *Arthritis Rheum* 48(4):1126–1136

10. van der Heijde D et al (2017) 2016 update of the ASAS-EULAR management recommendations for axial spondyloarthritis. *Ann Rheum Dis* 76(6):978–991
11. Schwartzman M, Maksymowych WP (2019) Is there a role for MRI to establish treatment indications and effectively monitor response in patients with axial spondyloarthritis? *Rheum Dis Clin N Am* 45(3):341–358
12. Liu S et al (2019) Deep learning in medical ultrasound analysis: a review. *Engineering* 5(2):261–275
13. Van Den Berghe T et al (2023) Neural network algorithm for detection of erosions and ankylosis on CT of the sacroiliac joints: multicentre development and validation of diagnostic accuracy. *Eur Radiol* 33:8310–8323
14. Faleiros MC et al (2020) Machine learning techniques for computer-aided classification of active inflammatory sacroiliitis in magnetic resonance imaging. *Adv Rheumatol* 60(1):25–25
15. Lin KYY et al (2022) Deep learning algorithms for magnetic resonance imaging of inflammatory sacroiliitis in axial spondyloarthritis. *Rheumatology (Oxford)* 61:4198–4206. <https://doi.org/10.1093/rheumatology/keac059>
16. Lee KH et al (2021) Method for diagnosing the bone marrow edema of sacroiliac joint in patients with axial spondyloarthritis using magnetic resonance image analysis based on deep learning. *Diagnostics (Basel)* 11(7):1156
17. Bresssem KK et al (2022) Deep learning detects changes indicative of axial spondyloarthritis at MRI of sacroiliac joints. *Radiology* 305(3):655–665
18. Weber U et al (2009) Sensitivity and specificity of spinal inflammatory lesions assessed by whole-body magnetic resonance imaging in patients with ankylosing spondylitis or recent-onset inflammatory back pain. *Arthrit Rheum-Arthr* 61(7):900–908
19. Weber U et al (2010) Assessment of active spinal inflammatory changes in patients with axial spondyloarthritis: validation of whole body MRI against conventional MRI. *Ann Rheum Dis* 69(4):648–653
20. Piri R et al (2022) PET/CT imaging of spinal inflammation and microcalcification in patients with low back pain: a pilot study on the quantification by artificial intelligence-based segmentation. *Clin Physiol Funct Imaging* 42(4):225–232
21. Koo BS et al (2022) A pilot study on deep learning-based grading of corners of vertebral bodies for assessment of radiographic progression in patients with ankylosing spondylitis. *Ther Adv Musculoskelet Dis*. <https://doi.org/10.1177/1759720X221114097>
22. Balzer I et al (2022) A deep learning pipeline for automatized assessment of spinal MRI. *Comput Methods Programs Biomed Update* 2:100081
23. Oktay O, Schlemper J, Le Folgoc L, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B, Glocker B, Rueckert D (2018) Attention U-net: learning where to look for the pancreas. [arXiv:1804.03999](https://arxiv.org/abs/1804.03999)
24. Lorenzin M et al (2020) Spine and sacroiliac joints lesions on magnetic resonance imaging in early axial-spondyloarthritis during 24-months follow-up (Italian Arm of SPACE Study). *Front Immunol* 11:936–936

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.