Automatic MRI Detection and Grading of L-spine Intervertebral Disc Degeneration (IVDD) at 0.05 Tesla via Deep Learning

Junhao Zhang
Department of Electrical and Electronic Engineering
The University of Hong Kong
Hong Kong SAR, China

Yujiao Zhao
Department of Electrical and Electronic Engineering
The University of Hong Kong
Hong Kong SAR, China

Vick Lau
Department of Electrical and Electronic Engineering
The University of Hong Kong
Hong Kong SAR, China

Liubin Wu
Department of Biomedical Engineering
The University of Hong Kong
Hong Kong SAR, China

Xuehong Lin
Department of Biomedical Engineering
The University of Hong Kong
Hong Kong SAR, China

Jiahao Hu
Department of Biomedical Engineering
The University of Hong Kong
Hong Kong SAR, China

Ye Ding
Department of Biomedical Engineering
The University of Hong Kong
Hong Kong SAR, China

Shi Su
Department of Biomedical Engineering
The University of Hong Kong
Hong Kong SAR, China

Kay Chioma Igwe
Department of Biomedical Engineering
Columbia University
New York, USA

Andrew Laine
Department of Radiology and Biomedical Engineering
Columbia University
New York, USA

Chunyi Wen
Department of Biomedical Engineering
The Hong Kong Polytechnic University
Hong Kong SAR, China

Alex T.L. Leong
Department of Electrical and Electronic Engineering
The University of Hong Kong
Hong Kong SAR, China

Ed X. Wu
Department of Electrical and Electronic Engineering
The University of Hong Kong
Hong Kong SAR, China
ewu@eee.hku.hk

***Abstract*—Detection of lumbar spinal degeneration by high-field MRI is clinically valuable but expensive. Such detection via the emerging low-cost ultra-low-field (ULF) MRI remains largely unexplored. Here, we developed one two-stage pipeline for automatic detection of lumbar intervertebral disc degeneration at 0.05 Tesla via deep learning. In the first stage, one well-trained 3D super-resolution model will be used to reconstruct super-resolution images from the corresponding raw ULF MRI images. In the second stage, the super-resolution reconstructed images will be used as model input for one RCNN-based deep learning model to yield potential degeneration annotations. The model for the lumbar spinal degeneration exhibits good performance in terms of degeneration localizing and degeneration level classification. This work demonstrates the first automatic IVDD detection by L-spine MRI at 0.05 Tesla. The results will facilitate low-cost spinal health screening and advance the ULF MRI towards truly accessible and point-of-care clinical applications.**

***Index Terms*—ultra-low-field MRI, intervertebral disc degeneration, super-resolution reconstruction, object detection, deep learning**

I. Introduction

Intervertebral disc degeneration (IVDD) is a major age-related public health concern and a leading cause of low back pain worldwide [1], [2]. In clinical practice, T2W lumbar spine (L–spine) MRI serves as the primary imaging modality for diagnosing and staging IVDD by visualizing structural abnormalities in discs. While ultra–low–field (ULF) MRI scanners that operate below 0.1 Tesla have recently emerged as a promising pathway toward low–cost, portable, and point–of–care imaging, a significant barrier to their application in L–spine imaging is the extremely low SNR due to the drastic field strength reduction, which can severely compromise the image quality for fine spinal structure visualization [3]–[6].

Currently, deep learning has been one powerful tool in the medical image analysis [7]–[9]. In the lesion detection field, there are two mainstreams to realize the lesion detection: a) RCNN–based detection models; b) U–Net based semantic segmentation models. For the RCNN–based models, the models will identify the location using one bounding box at the first

stage and then assign the labels to the identified location at the second stage. For the semantic segmentation models, the models will predict the labels for each pixel or voxel and one mask with specific values corresponding to each pixel/voxel of the input medical images [10], [11].

In this study, we develop a novel framework for automatic IVDD detection from ULF L–spine MRI. Leveraging a newly developed shielding–free 0.05 Tesla whole–body MRI scanner recently developed in our laboratory [4], we integrate a deep learning image formation method to overcome the inherent low SNR limitations. We further develop a new deep learning method for automatic IVDD detection on the deep learning super–resolution (SR) reconstructed ULF L–spine images. Our results have demonstrated preliminary feasibility of automatic MRI IVDD detection and grading at 0.05 Tesla.

II. Methods

A. A. Proposed RCNN–based Disc Degeneration Detections

The framework of our proposed automatic disc degeneration detection is shown in Fig. 1A. Firstly, the 3D T2W k-space data with partial Fourier sampling was acquired at one whole-body built-in MRI scanner and then reconstructed by 3D Fourier Transform (FT). Secondly, the FT reconstruction images are then into the well-trained 3D PF-SR model to yield SR reconstructed images [12]. Finally, the SR reconstructed images will be fed into the well–trained detection model for lesion localization and classification. Specifically, the architecture of the 3D PF–SR model was shown in Fig. 1B. This model integrates multi-scale residual learning and attention mechanisms to effectively extract the features from low-resolution data and then fuse the features to stably reconstruct the high-resolution counterparts [4].

The deep learning pipeline of the T2W L-spine IVDD detection at 0.05T whole-body ULF MRI. A) The 3D T2W k-space data with Partial Fourier sampling is acquired at a whole-body MRI scanner and then reconstructed by 3D Fourier Transform (FT). The reconstructed raw images are fed into the 3D PF super-resolution (PF-SR) model to yield SR reconstructed images. Finally, the SR reconstructed images will be fed into the detection model for prediction. B) The architecture of the 3D PF-SR model. C) The architecture of the RCNN detection model.

The RCNN-based model in Fig. 1A is one region-based convolutional neural network for object detection. It could address the computational inefficiency of running a separate convolutional neural network (CNN) on each region proposal by computing convolutional features once per model input image and sharing convolutional features across all region proposals. The workflow of this detection model is shown in the top part of Fig. 1C. The model input is the SR reconstructed multi-slice data, and the model output is the predicted disc degeneration localization and labels in the central slice of the model input. The model input will firstly go into the ResNet–50 module for deep feature extraction, and then the extracted feature maps will be fed into the region proposal module for roughly proposing the regions corresponding to some potential objects of interest projected onto the feature maps via the window sliding across the feature maps. After the region proposal, the proposed size-varying local feature maps related to objects of interest will be fed into the Region–of–Interest (RoI) alignment module to map all length-varying features from local feature maps to the fixed-length features per region proposal. Finally, the extracted fixed-length features will go into the box regression and classification module for the bounding box coordinate regression and box labeling, and the predicted bounding box is adapted to the original model input size. The architectures of the modules used in the detection model workflow are shown in the bottom part in Fig. 1C. The loss function in the model training is made up of two components: 1) the bounding box coordinate regression smooth L_1 loss L_{reg} calculate the difference between the ground truth bounding box coordinates and the predicted bounding box coordinates; 2) the bounding box classification cross–entropy loss L_{cls} to calculate the difference between the ground truth bounding box labels and the predicted bounding box labels. The total loss function L could be formulated as

$$L = 0.5\, L_{\text{reg}} + 0.5\, L_{\text{cls}} \ . \tag{1}$$

B. Data Preparation and Detection Model Training

Specifically, the 3T multi-slice T2W 2D L–spine training data was from the publicly available RSNA Lumbar Spine Degenerative Classification AI Challenge [13]. The typical spatial resolution is about $0.5\,\text{mm} \times 0.5\,\text{mm} \times 4.0\,\text{mm}$. In the simulation phase, the raw multi-slice T2W L-spine data are down-sampled to about $2.2\,\text{mm} \times 2.2\,\text{mm} \times 8.0\,\text{mm}$ and then added relative Rician noise for synthetic ULF data generation with corresponding annotations; In the meantime, the raw multi-slice T2W L-spine data are down-sampled to about $1.1\,\text{mm} \times 1.1\,\text{mm} \times 4.0\,\text{mm}$ to yield synthetic SR reconstructed data with corresponding annotations. The number of high-quality data is about 1100, and the data could be divided into training/validation/test datasets with the ratio of 7:2:1. Sagittal T2WL-spine datasets were acquired using 3DFSE sequences with TR=1800 ms, TE=172 ms, echo train length (ETL)=31, and NEX=2. Data acquisition adopted 2D k-space partial Fourier (PF) sampling with a fraction of 0.78 along the slice selection direction and a fraction of 0.84 along the phase encoding direction, respectively. Scan time was kept at around 8 mins. After each scan, EMI elimination was first performed using Deep–DSP, and then image reconstruction is finished by 3D Fourier Transform [14].

In the RCNN-based detection model training, the detection model was loaded with pretrained weights from the torchvision library for better initialization in predicting the disc degeneration localizations and corresponding disc degeneration labels [15]. The training process was achieved by minimizing the L loss in Equation (1). The Adam optimizer was carried out for training with $\beta_1 = 0.9$ and $\beta_2 = 0.999$, and the initial learning rate was 0.0001. The training was conducted on 4 NVIDIA A800 80GB PCIe GPUs using the PyTorch package with a batch size of 8 and 300 epochs. The total

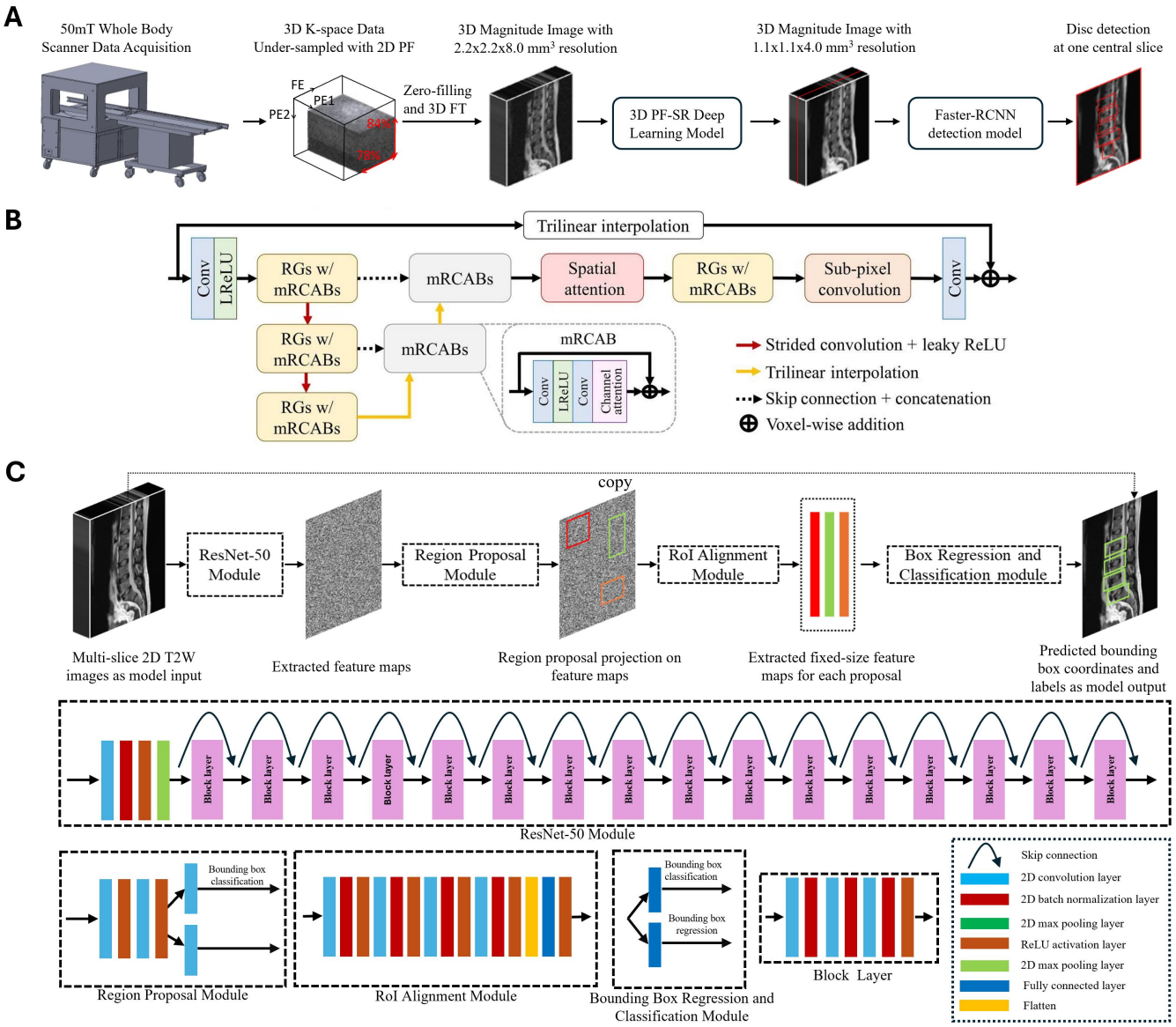

Fig. 1. The deep learning pipeline of the T2W L-spine IVDD detection at 0.05T whole-body ULF MRI. A) The 3D T2W k-space data with Partial Fourier sampling is acquired at a whole-body MRI scanner and then reconstructed by 3D Fourier Transform (FT). The reconstructed raw images are fed into the 3D PF super-resolution (PF-SR) model to yield SR reconstructed images. Finally, the SR reconstructed images will be fed into the detection model for prediction. B) The architecture of the 3D PF-SR model. C) The architecture of the RCNN detection model.

training time was approximately 2.5 hours. For evaluation, here we compared the detection performance using the RCNN-based model on synthetic ULF data and the corresponding synthetic SR reconstructed data qualitatively and quantitatively (sensitivity and specificity) both at the per-patient level and the per-lesion level.

III. Results

Fig. 2 shows the per-lesion disc degeneration detection results on synthetic SR reconstructed data from one healthy subject and one subject with disc degeneration. The model could correctly identify the discs of interest and predict the severity level for each disc for two subjects at the per-lesion level.

Fig. 3 presents the per-lesion detection results of the model on the synthetic ULF data and corresponding synthetic SR reconstructed data from one subject. The model performs better on synthetic SR reconstructed data. Fig. 4 shows the averaged per-lesion and per-patient sensitivity and specificity of the model detection on the synthetic ULF data and synthetic SR reconstructed data, respectively. It is observed that the model using synthetic SR reconstructed data could significantly outperform that on synthetic ULF data in terms of averaged sensitivity and specificity at both the per-patient level and per-lesion level. Fig. 5 shows the per-lesion disc degen-

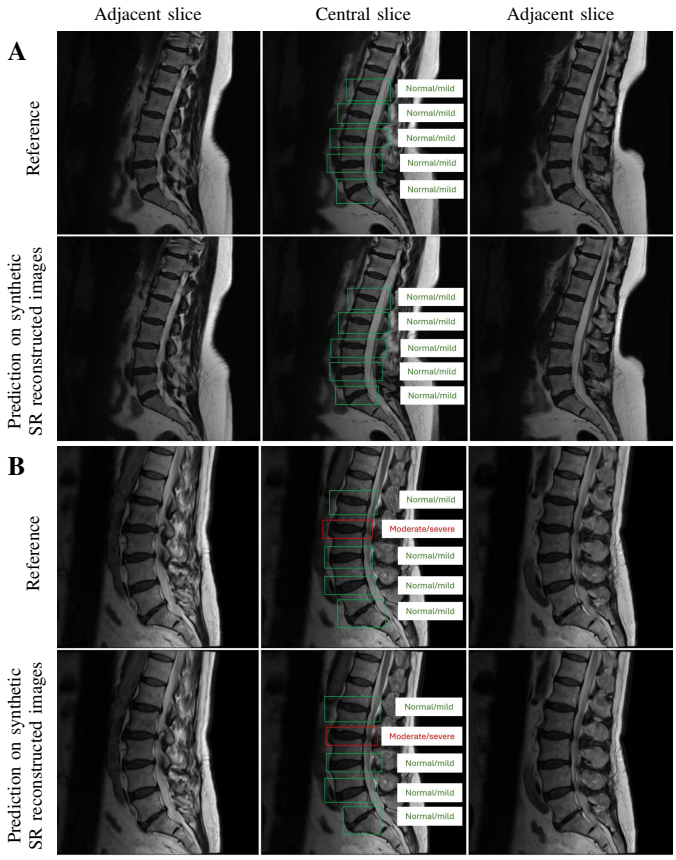

Fig. 2. IVDD detection results of synthetic ULF MRI SR reconstructed data from (A) one healthy subject and (B) one subject with disc degeneration in the per-lesion level. The model utilizes multi-slice information to yield predictions for the central slice. For both two subjects, the model could correctly predict the boxes and box labels for the disc of interest.

eration detection results on experimental SR reconstructed data from one healthy subject and one subject with disc degeneration. The model could consistently perform well in disc degeneration detection.

IV. Discussion

It is shown that the model could have a good performance on both synthetic SR reconstructed images and experimental deep learning reconstructed images in disc degeneration detection at the per-patient and per-lesion level. The deep learning SR reconstruction could help suppress the noise and recover the details, and this brings significant advantages to the disc degeneration detection task. Additionally, multi-slice L-spine images are used for the detection model to explore the spatial continuity between multiple adjacent slices.

This study has several limitations: firstly, here only limited experimental data is used to demonstrate the model detection performance, and there is no 3T reference data and the radiologist's report for validation. In the future, we should recruit more subjects for MRI scanning at both our ULF scanner and the high-field scanner and get references from the radiologists. Secondly, our proposed framework is one two-stage method: first SR reconstruction is applied and then detection on SR reconstructed data. The reconstruction error or imperfections

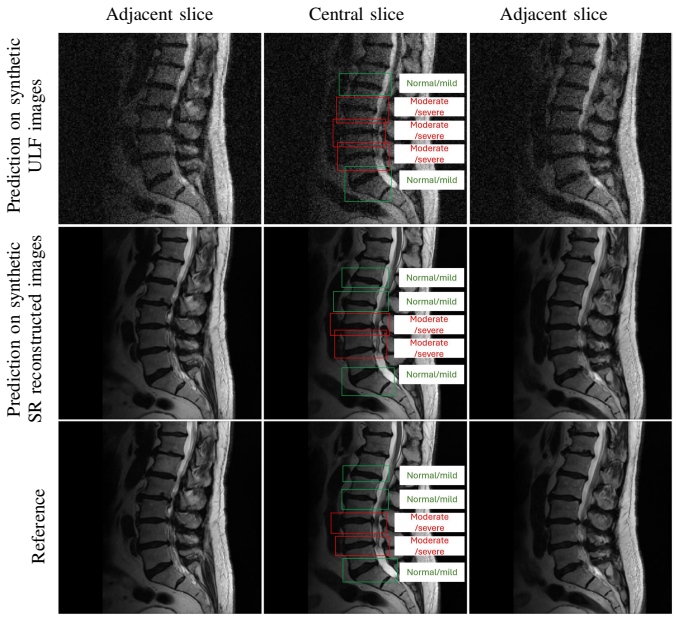

Fig. 3. IVDD detection results of synthetic ULF MRI data and the corresponding synthetic SR reconstructed data from one subject with disc degeneration in the per-lesion level. The model could perform better on synthetic SR reconstructed data. This indicates that the detection model performance could be improved after introducing SR reconstruction.

	Sensitivity	Specificity
Test on synthetic ULF images	Per-lesion level: 81.7%	83.2%
	Per-patient level: 85.1%	100.0%
Test on synthetic SR reconstructed images	Per-lesion level: 93.2%	93.0%
	Per-patient level: 95.0%	100%

Fig. 4. The averaged per-lesion and per-patient sensitivity and specificity of the IVDD detection performance on the synthetic ULF data and synthetic SR reconstructed data. It is observed that the model on synthetic SR reconstructed data could significantly outperform that on synthetic ULF data at both the per-patient and per-lesion level in terms of the averaged sensitivity and specificity.

in the first stage will have a direct effect on the second stage detection. A good SR reconstruction will benefit the whole framework. A one-stage framework may be developed to mitigate the error in the stage transferring. Thirdly, we do not demonstrate the benefits of continuous multi-slice 2D data as the detection model input explicitly by comparing with the model trained on single-slice 2D data. Here, an ablation study to demonstrate the usefulness of information from multiple slices needs to be carried out later. Lastly, currently the training data is limited both for the SR reconstruction model and the detection model, and this will limit the model performance.

V. Conclusion

This study demonstrates, for the first time, that the proposed method enables reliable detection of disc degeneration in ultra low field (ULF) MRI with promising diagnostic performance. The results highlight the feasibility of ULF MRI as a cost effective and accessible imaging modality, capable of providing clinically meaningful information despite inherent hardware limitations. By validating the utility of ULF MRI, this work

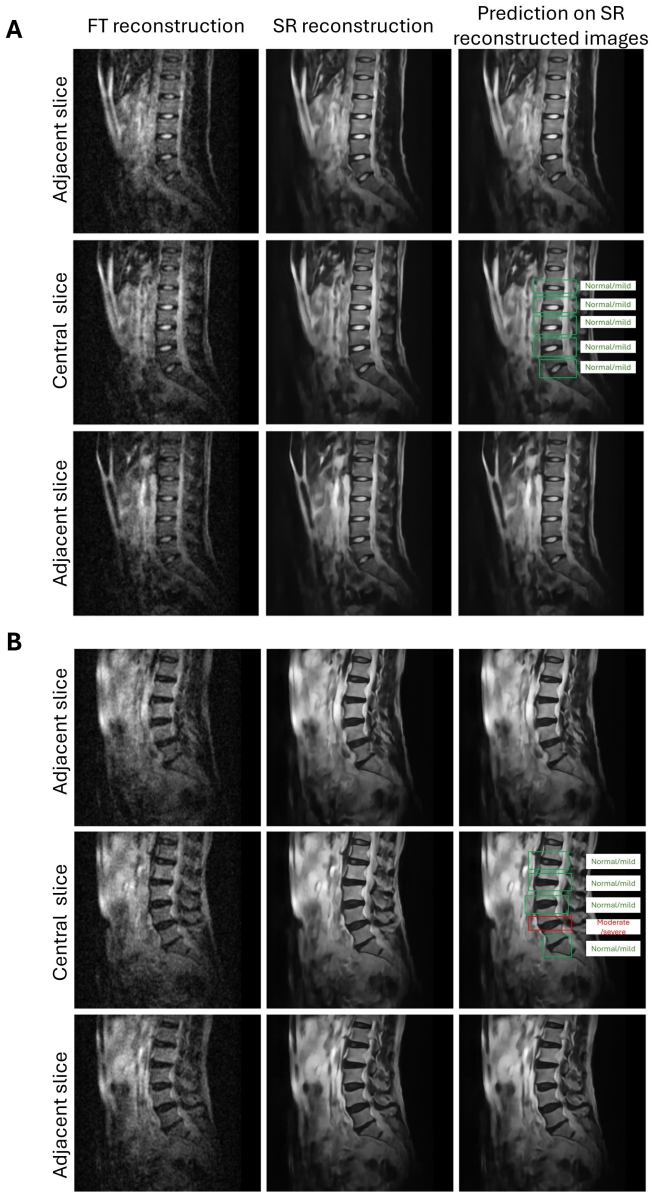

Fig. 5. IVDD detection of experimental 0.05 Tesla ULF MRI data from (A) one healthy subject and (B) one subject with disc degeneration in the per-lesion level. The first column is the raw experiment 3D FT reconstructed ULF images. The second column is the corresponding SR reconstructed images. The third column is the IVDD detection results on the corresponding SR reconstructed images. The detection model could correctly localize discs of interest and classify the labels for each disc for both two subjects on the SR reconstructed images.

expands the scope of low field imaging beyond proof of concept studies and establishes a foundation for future clinical translation. Continued refinement of the method and integration with advanced reconstruction or image enhancement strategies may further improve diagnostic accuracy, paving the way for broader adoption of ULF MRI in routine practice.

ACKNOWLEDGMENTS

This study is supported by Hong Kong Research Grant Council (R7003-19F, HKU17112120, HKU17127121, HKU17127022, HKU17127523, HKU17119224, HKU17130325, HKU17125624, HKU17127723, and HKU17130225), and Lam Woo Foundation.

REFERENCES

[1] Wang, M., Tsang, A., Tam, V., Chan, D., Cao, P. & Wu, E. X. Multiparametric MR investigation of proteoglycan diffusivity, T2 relaxation, and concentration in an ex vivo model of intervertebral disc degeneration. Journal of Magnetic Resonance Imaging 51, 1390–1400 (2020).

[2] Wassenaar, M., van Rijn, R. M., van Tulder, M. W., Verhagen, A. P., van der Windt, D. A., Koes, B. W., de Boer, M. R., Ginai, A. Z. & Ostelo, R. W. Magnetic resonance imaging for diagnosing lumbar spinal pathology in adult patients with low back pain or sciatica: a diagnostic systematic review. European Spine Journal 21, 220—227 (2012).

[3] Liu, Y., Leong, A. T., Zhao, Y., Xiao, L., Mak, H. K., Tsang, A. C. O., Lau, G. K., Leung, G. K. & Wu, E. X. A low–cost and shielding–free ultra–low–field brain MRI scanner. Nature communications 12, 7238 (2021).

[4] Zhao, Y., Ding, Y., Lau, V., Man, C., Su, S., Xiao, L., Leong, A. T. & Wu, E. X. Whole-body magnetic resonance imaging at 0.05 Tesla. Science 384, eadm7168 (2024).

[5] Man, C., Lau, V., Su, S., Zhao, Y., Xiao, L., Ding, Y., Leung, G. K., Leong, A. T. & Wu, E. X. Deep learning enabled fast 3D brain MRI at 0.055 tesla. Science Advances 9, eadi9327 (2023).

[6] Lau, V., Xiao, L., Zhao, Y., Su, S., Ding, Y., Man, C., Wang, X., Tsang, A., Cao, P., Lau, G.K. and Leung, G.K., 2023. Pushing the limits of low-cost ultra-low-field MRI by dual–acquisition deep learning 3D superresolution. Magnetic Resonance in Medicine, 90(2), pp.400-416.

[7] Suganyadevi, S., Seethalakshmi, V., & Balasamy, K. A review on deep learning in medical image analysis. International Journal of Multimedia Information Retrieval 11, 19–38 (2022).

[8] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., Van Der Laak, J. A., Van Ginneken, B. & Sánchez, C. I. A survey on deep learning in medical image analysis. Medical image analysis 42, 60—88 (2017).

[9] Kimberly, W.T., Sorby-Adams, A.J., Webb, A.G., Wu, E.X., Beekman, R., Bowry, R., Schiff, S.J., de Havenon, A., Shen, F.X., Sze, G. and Schaefer, P., 2023. Brain imaging with portable low–field MRI. Nature reviews bioengineering, 1(9), pp.617–630.

[10] Girshick, R. Fast r–cnn in Proceedings of the IEEE international conference on computer vision (2015), 1440—1448.

[11] Jiang, P., Ergu, D., Liu, F., Cai, Y. & Ma, B. A Review of Yolo algorithm developments. Procedia computer science 199, 1066–-1073 (2022).

[12] Margosian, P.M., DeMeester, G. and Liu, H., 1996. Partial Fourier acquisition in MRI. Encyclopedia of nuclear magnetic resonance, 5, pp.3463–3467.

[13] Richards, T. J., Flanders, A. E., Colak, E., Prevedello, L. M., Ball, R. L., Kitamura, F., Mongan, J., Vazirabad, M., Lin, H. M. & Kendell, A. The RSNA Lumbar Degenerative Imaging Spine Classification (LumbarDISC) Dataset. arXiv preprint arXiv:2506.09162 (2025).

[14] Zhao,Y., Xiao,L., Liu, Y., Leong, A. T.& Wu, E.X. Electromagnetic interference elimination via active sensing and deep learning prediction for radiofrequency shielding–free MRI. NMR in Biomedicine 37, e4956 (2024).

[15] Marcel, S. & Rodriguez, Y. Torchvision the machine-vision package of torch in Proceedings of the 18th ACM international conference on Multimedia (2010), 1485—1488.