Deep Learning Approach for Evaluating Lumbar Intervertebral Disc Degeneration: Achieving High Accurate Segmentation for Quantitative Analysis on MRI

Hua-dong Zheng
Shanghai University
Yue-li Sun (yueli_sun@foxmail.com)
Longhua Hospital, Shanghai University of Traditional Chinese Medicine
De-wei Kong
Depart ment of Radiology，Longhua Hospital of Shanghai University of TCM
Meng-chen Yin
Longhua Hospital, Shanghai University of TCM
Jiang Chen
Dongzhimen Hospital of BeijingUniversity of Chinese Medicine
Yong-peng Lin
Dongzhimen Hospital, Beijing University of Chinese Medicine
Xue-feng Ma
Shenzhen Pingle Orthopedics Hospital (Shenzhen Pingshan District Hospital of TCM)
Hong-shen Wang
Guangdong Provincial Hospital of Chinese Medicine
Guangjie Yuan
Shanghai University
Min Yao
Longhua Hospital, Shanghai University of TCM
Xue-jun Cui
Longhua Hospital, Shanghai University of TCM
Yingzhong Tian
Shanghai University
Yongjun Wang
Shanghai University of Traditional Medicine https://orcid.org/0000-0001-9333-2423

Article

Keywords: lumbar disc degeneration, intervertebral disc degeneration, MRI, deep learning and image processing technology

Posted Date: September 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-864336/v1

Version of Record: A version of this preprint was published at Nature Communications on February 11th, 2022. See the published version at https://doi.org/10.1038/s41467-022-28387-5.

Abstract

Purpose: Using deep learning and image processing technology, a standardized automatic segmentation and quantitation network of lumbar disc degeneration based on T2MRI was proposed to help residents accurately evaluate the intervertebral disc (IVD) degeneration.

Materials and Methods: A semantic segmentation network (BianqueNet) consist of self-attention mechanism skip connection module and deep feature extraction module was proposed to achieve high-precision segmentation of IVD related areas. A quantitative method was used to calculate the signal intensity difference (ΔSI) in IVD, average disc height (DH), disc height index (DHI), and disc height-to-diameter ratio (DHR). Quantitative ranges for these IVD parameters in a larger population was established among the 1051 MRI images collected from four hospitals around China.

Results: The average dice coefficients of BianqueNet for vertebral bodies and intervertebral discs segmentation are 97.04% and 94.76%, respectively. This procedure was suitable for different MRI centers and different resolution of lumbar spine T2MRI (ICC=.874~.958). These geographic parameters of IVD degeneration have a significant negative correlation with the modified Pfirrmann Grade, while signal intensity in IVD degeneration had excellent reliability according to the modified Pfirrmann Grade (macroF1=90.63%~92.02%).

Conclusion: we developed a fully automated deep learning-based lumbar spine segmentation network, which demonstrated strong versatility and high reliability to assist residents on IVD degeneration evaluating by means of IVD degeneration quantitation.

Implication for Patient Care: Deep learning–based approaches have the potential to maximize diagnostic performance for detecting disc degeneration and assessing risk of disc herniation while reducing subjectivity, variability, and errors due to distraction and fatigue associated with human interpretation.

Introduction

The intervertebral disc (IVD) plays an important role in distributing loads and absorbing shock in the spine, which is comprised of a gel-like nucleus pulposus (NP), collagenous annulus fibrosis (AF) layers, and ring-like cartilaginous endplates (EP). Identifying IVD structural changes, including IVD deformation, NP dehydration and EP ossification, due to chronic degeneration or acute injury, in patients undergoing MRI of the lumbar spine has many important clinical implications [1]. It has been determined that IVD degeneration is a consequence of aging. Accumulated compressive overload usually lead to functional fatigue fractures in endplates and subsequently IVD herniation[2–4], which may lead to increased inflammation[5], nerve compression[6] and release of pain factors[7]. Lifestyle modifications and surgical interventions are likely to be most effective for treating IVD degeneration or herniation, but it is more important to initiate screening and prevention during the earliest stages of the disease process.

MRI with morphologic cartilage imaging sequences has been shown to have high specificity but only moderate sensitivity for detecting dehydration and deformation within the IVD degeneration[2,4]. Diagnostic performance is highly dependent on the level of reader expertise, and only moderate interobserver agreement between readers has been reported in most studies [2]. Quantitative analysis is efficient and comprehensive in evaluating IVD degeneration by measuring the signal intensity and geometric information. Early research on quantitative measurement of intervertebral discs used general image processing programs for manual measurement[8–10]. However, there is still no universal automatic IVD degeneration analysis tool in this field. The lack of a universal and widely accepted standard definition of IVD degeneration is one of the main reasons.

There has been much recent interest in using deep learning methods in medical imaging [11]. With the wide-spread application of convolutional neural network classifiers in medical images, many studies use the rectangular box surrounding the lumbar IVD as input, and the corresponding degeneration level as the label to train the classifier for learning degenerative features by neural network. However, the input rectangular bounding box of the intervertebral disc needs to be segmented artificially or detected automatically using complex algorithms[1,12–17]. There are also some studies on the quantitative measurement of intervertebral discs based on deep learning, which did not use quantitative data to evaluate intervertebral disc degeneration[18,19].

In this study, a fully automated deep learning-based lumbar spine segmentation network (LSSN) has been developed at our institution by using a deep convolutional neural network (CNN) with the self-attention skip connection, deep feature extraction module and the corresponding loss function. According to IVD degeneration features (water content loss and height decrease) [20], signal intensity difference and geometric parameters of IVD are calculated and validated with the modified Pfirrmann grading system. Finally, baseline ranges of lumbar IVD parameters among different gender and age and lumbar level was established based on a large population around China for quantitative and structured report. The diagram of this study is illustrated in Fig. 1.

Materials And Methods

MRI Data Sets

This study was approved by Institutional Review Board (IRB) in all the participating sites. All retrospective subject data were obtained with a waiver of consent under IRB approval. The data were anonymized before being shared.

Data sets for segmentation training (Data set A & B)

Training and validation of the proposed lumbar spine semantic segmentation method was carried out by performing an institutional review board–approved retrospective analysis of lumbar spine images from

286 subjects who underwent MR imaging in the Longhua Hospital, Shanghai University of TCM between January 1, 2019, to December 31, 2020. Among these, there're 223 subjects using a 1.5-T MRI unit (MAGNETOM Aera XJ, SIEMENS) and 63 subjects using another 1.5-T MRI unit (MAGNETOM Avanto, SIEMENS), which were trained two separate segmentation networks for different resolution of 512*512 (Data set A) and 320*320 (Data set B). Mid-sagittal T2 images of different resolution were exported from Data set A and Data set B respectively, being randomly allocated into each training set or test set (**Fig.1**). All images in the segmentation data set were labeled by LabelMe (version 3.3.6, CSAIL, Massachusetts Institute of Technology) [21]. Based on the structural features mentioned in the modified Pfirrmann grading system, the segmentation area of 14 parts, included 5 vertebral bodies (L1-L5), 5 lumbar IVDs (L1/L2-L5/S1), sacrum (S1), pre-iliac fat area, cerebrospinal fluid area in the spinal canal, and background as **Fig. 2a**.

Data set for quantitative analysis (Data set C)

The proposed LSSN was used to extracted 1051 lumbar spine images as Data set C in four hospitals around China, including Longhua Hospital, Shanghai University of TCM, Guangdong Provincial Hospital of Chinese Medicine, Shenzhen Pingle Orthopedics Hospital, and Dongzhimen Hospital, Beijing University of Chinese Medicine between January 1, 2019, and March 30, 2021. The imaging parameters of all sites are summarized in **Table 1**.

Table 1 Imaging Parameters for the MRI Sequences in the 4 Sites

Site	City	Strength of the Magnet	Company	Model	Coil
Longhua Hospital, Shanghai University of TCM	Shanghai	1.5-Tesla	SIEMENS	MAGNETOM Aera XJ	18-channel Spine Tim 4G coil
Guangdong Provincial Hospital of Chinese Medicine	Guangzhou	3-Tesla	SIEMENS	TIM Systems	32-channel Spine Tim coil
Shenzhen Pingle Orthopedics Hospital	Shenzhen	1.5-Tesla	SIEMENS	MAGNETOM Essenza	8-channel quadrature body coil
Dongzhimen Hospital, Beijing University of Chinese Medicine	Beijing	1.5-Tesla	SIEMENS	MAGNETOM Amira	24-channel quadrature body coil

A research team, composed of a 4-year radiology resident (DW Kong), two 8-year orthopedic resident (J Chen, XF Ma), two 4-year orthopedic resident (YL Sun, YP Lin) and a 2-year orthopedic resident (MC Yin), discussed together for the final segment and Pfirrmann grade for each MR image.

Lumbar Spine Segmentation from MR Images

Convolutional Neural Network (CNN) Training

The critical component of LSSN is an improved deeplabv3+ segmentation network[22] with backbone ResNet-101[23], called BianqueNet. The BianqueNet was built on the basis of deep feature extraction to extract richer semantic information and denser features. An illustration of this semantic segmentation network is shown in **Fig. 2**. The entire network consists of a swin transform skip connection (ST-SC) module and a deep feature extraction (DFE) module. Swin Transform is a hierarchical transform calculated by shifting the window, which has the advantages of high efficiency and low complexity [24]. The skip connections structure designed in this study uses two successive Swin-Transformer blocks, with 1*1 convolutional layers in parallel at the same time, and finally the two output features are spliced. Through the pyramid pooling module, feature information of different depths through pooling operations

of different scales can be obtained. By repeating check with feature map of 4096 channels multi-scale information [25], the network can achieve efficient features extraction with a dense semantic feature map of 256 channels.

32-depth BMP images were exported from raw MRI to train the LSSN as input. In the upsampling phase, a modified upsampling operation with a deconvolution decoder was used to recover more detailed features of the segmentation target. In the feature extraction phase, the feature maps of different resolutions were obtained by down-sampling and output to the ST-SC module, which splices images and extracts features from different resolutions. According to feature pyramid[26], feature maps with low-resolution and high-resolution were integrated to extract more semantic and spatial information, in which a 3*3 double convolutional layer was used for the fused feature map to improve the feature. Finally, a double upsampling operation was performed to obtain a dense prediction image.

Weighted Dice Loss Function

A weighted dice loss function as below was proposed to enhance segmentation performance by estimating difficlties in difference images with typical or atypical structure, which ensured consistent in segmentation:

$$L_{wdice} = \frac{1}{C}\sum_{j=1}^{C} \xi_j \left(1 - \frac{2\sum_{i=1}^{N} p_{1i} g_{1i}}{2\sum_{i=1}^{N} p_{1i} g_{1i} + \sum_{i=1}^{N} p_{0i} g_{1i} + \sum_{i=1}^{N} p_{1i} g_{0i}}\right) \quad (1)$$

This formula was used in the output of the softmax layer, where the is the probability of voxel (target) and is the probability of voxel (non-target). So was for and . represents different segmentation areas, represents the total number of channels, which is taken as 14. represent the weight of different segmentation channels. According to the experimental analysis results, channels weight was set to 0.9, 0.8 and 1 for vertebral body, IVD and the other respectively, which may achieve the best segmentation performance.

For avoiding that the subsequent feature extraction operations are affected, corrosion and expansion operations were used to remove the burrs (**Fig. 2b**).

Lumbar IVD Quantitative Analysis

Parameters Calculation based on Pfirrmann Grading System

Based on previous studies [18, 25–27], some extraction and calculation methods were modified with histogram features of IVD. signal intensity difference (ΔSI) was obtained to quantify the blurring degree of boundary between NP and AF, which indicating water content in IVD. Average disc height (DH), disc

height index (DHI), and disc height-to-diameter ratio (DHR) were obained to quantify structural collapse in IVD degeneration. Specific calculation methods for each parameter are described in the **Supplement File 1.**

Versatility Test for Images with Different Origins

IVD parameters extracted by LSSN in mid-sagittal lumbar MR images with different resolutions were compared with each other. In the data set B, 46 images with resolution of 320*320 were randomly selected to be segmented and quantified by model B. Meanwhile, these images were adjusted to 512*512 for segmentation and quantitation by model A. IVD parameters extracted from these two models were used for versatility test. If IVD parameters from LSSN shows good consistency under different origins of imaging, LSSN will be used into a larger population (Data set C) with different machines and models to establish degenerative ranges of IVD parameters in different people.

Quantitation for IVD Degeneration

Relationship between IVD parameters (including ΔSI, DH, DHI, and HDR) and demographic information (including gender, age, and segment) and correlation between IVD parameters and IVD degeneration (based on the modified Pfirrmann grading system) were analyzed respectively in Data set C. Based on these results, a baseline of IVD degeneration in larger population was established, which may indicate a qualitative IVD degeneration in different population with accordance to the modified Pfirrmann grading system. Details in quantitative protocols were shown in the **Supplement File 2**.

Statistical analysis

The intraclass correlation coefficient (ICC) was used to analyze the consistency between the IVD parameters calculated using the original resolution (320*320) from the data set B and the adjusted resolution (512*512). The macroF1-score and the Kendall correlation coefficient were used to test sensitivity and specificity in IVD degeneration grading performance among deep learning methods and 3 residents in the two data sets with different resolution according to the modified Pfirrmann grading system. An absolute value of r of 0-0.4 was considered as weak correlation, 0.4-0.6 as moderate correlation, and greater than 0.6 as strong correlation.

Spearman rank correlation coefficient between IVD signal intensity and grading score has been calculated via SPSS (version 26, IBM, USA). Multiple regression analysis was performed on IVD quantitative parameters (, DH, DHI, HDR) and baseline information (including gender, age, segment) to describe some characters in larger population via Stata (version 15.1, USA).

Results

Segmentation Performance

The BianqueNet provided good segmentation performance of IVD-related areas. The mean Dice coefficients (mDice) and mean intersection over Union (mIoU) were 94.45% and 89.88% for whole lumbar spine, 96.71% and 93.66% for vertebral body, 94.38% and 89.43% for IVD. Based on deeplabv3+, Adding the three modules of DFE, ST-SC and FPN improved segmentation significantly as shown in **Table 2**. The average training time for BianqueNet was 10 hours in each data set. Segmentation of vertebral bodies and IVDs on the mid-sagittal MRI image for a patient took approximately 1 seconds with the trained network.

Table 2 BianqueNet shows superior segmentation effectiveness demonstrated by the pixel-level Dice and IoU coefficient

Model	Module			Vertebral body		IVD		Lumbar spine	
	DFE	ST-SC	FPN	mDice	mIoU	mDice	mIoU	mDice	mIoU
DeepLabv3+				0.9671	0.9366	0.9438	0.8943	0.9445	0.8988
DeepLabv3++DFE	√			0.9681	0.9384	0.9444	0.8960	0.9455	0.9006
DeepLabv3++DFE+ST-SC	√	√		0.9692	0.9405	0.9458	0.8982	0.9468	0.9028
DeepLabv3++DFE+ST-SC+FPN (BianqueNet)	√	√	√	**0.9703**	**0.9425**	**0.9480**	**0.9019**	**0.9470**	**0.9035**

Versatility test for different resolution

A total of 230 IVDS and 276 vertebral bodies of 46 subjects were segmented after resolution of MRI-exported images had been adjusted from 320*320 to 512*512. The results showed a good consistency in using different parameter calculation algorithms for different resolution of MRI-exported images. Among them, the measurement of intervertebral disc geometric parameters DHI and DWR have extremely high ICC values, which are 0.958 (p=0.000) and 0.956 (p=0.000), respectively, and the ICC value of the *ΔSI* is 0.874 (p=0.000), as shown in **Table 3**.

Table 3 Consistency analysis of intervertebral disc parameters calculated by MRI of different sizes

Measure	Intraclass Correlation[b]	
	ICC[a]	95%CI
ΔSI	.874***	.840.902
DHI	.958***	.943.968
HDR	.956***	.886.978

Two-way mixed effects model where people effects are random and measures effects are fixed. ICC, intraclass correlation coefficient; 95% CI, 95% confidence interval;

a.The estimator is the same, whether the interaction effect is present or not.

b.Type A intraclass correlation coefficients using an absolute agreement definition.

Characteristics of IVD Parameters in a Larger Population

After screening 1508 MRI images in 4 sites around China, a total of 1051 individuals were collected, in which there're 144 excluded for imaging quality and 313 excluded for irregular structures (especially in vertebral bodies). The demographic information (including age and gender) distributed evenly as shown in **Table 4**, which were integrated to conduct correlation analysis with IVD parameters.

Table 4 Included Patient Demographic Information from the Four Sites around China

Site	Number	Age(F/M)					
		20-29	30-39	40-49	50-59	60-69	70-89
Longhua Hospital, Shanghai University of TCM	433	32/21	52/51	49/45	34/35	53/39	12/10
Shenzhen Pingle Orthopedics Hospital	222	16/18	20/20	19/20	18/21	13/23	9/25
Guangdong Provincial Hospital of Chinese Medicine	246	19/24	20/15	23/17	22/17	18/15	22/34
Dongzhimen Hospital, Beijing University of Chinese Medicine	150	7/8	13/18	21/17	13/8	12/11	8/14
Total	1051	74/71	105/104	112/99	87/81	96/88	51/83

Supplement Figure (1-4) and **Table 5** shows comprehensive distribution of IVD parameters in a larger population and multiple regression analysis result among IVD parameters and each demographic information respectively. *ΔSI* in IVDs decreased with age, while DH, DHI and DHR of IVDs increased with age, reaching peak at the age of 50-60 (P<0.01). There're no significant different between male and

female in in IVDs, while DH, DHI and DHR of IVDs were significantly higher in males than those in females (*P*<0.01). In additions, DH, DHI and DHR were significantly higher in lower segmental IVDs (L3-L4, L4-L5 and L5-S1) than upper ones (L1-L2 and L2-L3), and disc height of L4-L5 IVDs was highest(*P*<0.01). In the further analysis of all the IVD height parameters, the influence of segments on the parameters is greater than those of age. For the IVD height, the influence of gender is greater than age. For DHI and HDR, gender and age have similar effects.

Table 5 The results of multiple regression analysis of signal intensity peak difference, DH, DHI, HDR and gender, different ages, and different disc positions

N＝1651	*ΔSI*	DH	DHI	HDR
female	-.0279	-.2541***	-.1121***	.1115***
male	0.000	0.000	0.000	0.000
20-30	0.000	0.000	0.000	0.000
30-40	-.1669***	.0796***	.0557*	.1100***
40-50	-.3802***	.1110***	.0927***	.0980***
50-60	-.4826***	.1612***	.1577***	.0440
60-70	-.6002***	.1427***	.1687***	.0099
70-90	-.5137***	.0328	.0806***	-.0674***
L1-L2	.2800***	-.7181***	-.6708***	-.4932***
L2-L3	.1719***	-.3832***	-.4155***	-.2912***
L3-L4	.0907***	-.1593***	-.1942***	-.1122***
L4-L5	0.000	0.000	0.000	0.000
L5-S1	.1526***	-.0520**	-.0312	.1105***

*** p<0.01 ** p<0.05 * p<0.1

Correlation with IVD Degeneration Grading

Considering structural collapse with IVD degeneration according to the modified Pfirrmann grading system, a regression analysis was conducted to investigate correlation between its certain grading (For analyzing the correlation between degeneration segments and *ΔSI*, the corresponding grading were 1, 2, 3, 4, and (5-8). For analyzing the correlation between the degeneration segments and geometric parameters, the corresponding grading are (1-5), 6, 7, 8)) and IVD parameters in different age, gender and segments. As shown in **Table 6**, IVD parameters showed a good accordance to the modified Pfirrmann grade.

Table 6 Correlations between IVD Parameters and Modified Pfirrmann Grading

lumbar level	*ΔSI*	DH		DHI		HDR	
	-.966***	female	male	female	male	female	male
L1/L2		-.421***	-.296***	-.304***	-.235***	-.473***	-.397***
L2/L3		-.481***	-.417***	-.354***	-.398***	-.575***	-.455***
L3/L4		-.639***	-.470***	-.530***	-.443***	-.626***	-.539***
L4/L5		-.656***	-.696***	-.560***	-.665***	-.709***	-.758***
L5/S1		-.701***	-.687***	-.641***	-.664***	-.744***	-.778***

*** p<0.01 ** p<0.05 * p<0.1

r, Spearman rank correlation coefficients

Regarding water content loss with IVD degeneration, result from a further regression analysis showed a stronger correlation between the modified Pfirrmann grade (1, 2, 3, 4, and (5-8)) and *ΔSI* (R=-0.966, *P*=0.000). Specific ranges of according to the modified Pfirrmann grade (1, 2, 3, 4, and (5-8)) were calculated and listed in **Table 7**.

Table 7 Quantitative ranges of *ΔSI* according to the modified Pfirrmann Grade (1-8)

modified Pfirrmann Grade	1	2	3	4	5-8
Number	154	1130	1622	1315	1034
(mean±SD)	121.97±9.96	95.34±7.20	72.34±7.81	44.63±8.49	20.60±9.28

According to the results of multiple regression analysis, gender and segements have significant correlations with *ΔSI*, while age, gender and segments have significant correlations with geometric parameters. **Fig.3** and **Supplement Table (1-4)** showed comprehensive distribution of IVD parameters in a larger population.

Discussion

Our study described a fully automated deep learning–based lumbar IVD quantitative system utilizing a CNN with the self-attention skip connection, deep feature extraction module and the corresponding loss function for segmenting IVD-related areas to extract geometric and signal parameters. The proposed deep learning approach achieved high accuracy for segmentation and measurement. More specifically, our method showed high consistency with the modified Pfirrmann grading system.

Compared with previously reported conventional image processing methods for lumbar spine MRI, our method is focus on quantitative measurement other than degeneration grade classification. Standard and accurate ranges of in IVD was established to quantify IVD degeneration, which has strong applicability and accuracy for grading IVD degeneration (macroF1: 92.02% and 90.63% in two data sets) as shown in **Table 8**.

Table 8 Accuracy of IVD degeneration grading with ΔSI in IVD

Modified Pfirrmann Grade		1	2	3	4	5-8	macro-average (%)	macroF1(%)
Data set A	Precision (%)	60.76	97.28	99.40	97.89	89.08	88.89	92.02
	Recall (%)	100	90.96	97.84	90.05	98.15	95.40	
Data set B	Precision (%)	/	81.82	93.55	100	85.71	90.27	90.63
	Recall (%)	/	90.00	90.63	83.33	100	90.99	

By means of LSSN, all the IVD parameters will be extracted and quantified from MR images in about half of second, which may describe water content loss and structural collapse in IVD, indicating degeneration process. **Fig.4** shows a potential application for structural MRI report output from the IVD degeneration quantitative analysis.

Pfirrmann grading system, as the most used IVD imaging grading method, was designed based on symptomatic patients with an average age of about 40 years old [15]. Therefore, its reliability for early IVD degeneration or IVD degeneration in the elderly people may be unsatisfied. This study proposed an automatic quantitative method for IVD degeneration assessment in asymptomatic patients of different ages. In addition, multiple quantitative sequences in imaging are generally used together to accurately evaluate IVD degeneration or lesions, which are too time-consuming to popularize MR imaging quantitative analysis in IVD. Our LSSN may meet both patients' affordability and clinical diagnosis needs.

Our study has limitations. First, we only included MR images with relatively regular outline in IVD-related areas, most of which were accurately segmented by LSSN. Second, accuracy of LSSN is dependent on the depth of BMP exported from MR imaging system. Third, the subjects included in this study did not take symptoms (such as low back pain) into account, lacking clinical validation on IVD degeneration.

Finally, as a retrospective study, IVD parameters extracted from Data set C did not exactly represent the real-world setting. Further research is needed to determine the applicability of this LSSN in a prospective multi-institutional study in patients with low back pain.

In conclusion, we developed a fully automated deep learning-based lumbar spine segmentation network, which demonstrated strong versatility and high reliability to assist residents on IVD degeneration grading by means of IVD degeneration quantitation.

Declarations

Acknowledge

This study was supported by the National Natural Science Foundation of China (81930116, 81804115, 81873317, and 81704096).

Author contributions

Guarantor of integrity of entire study, YL Sun, YJ Wang; study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, HD Zheng, MC Yin, M Yao, XJ Cui; clinical studies, YL Sun, DW Kong, MC Yin, J Chen, YP Lin, XF Ma; experimental studies, YZ Tian, HS Wang, GJ Yuan; statistical analysis, HD Zheng, M Yao, XJ Cui; and manuscript editing, HD Zheng, YL Sun, YJ Wang.

Disclosure of Conflict of Interest

All author disclosed no relevant relationships.

References

1. Schnake, K. J., Putzier, M., Haas, N. P. & Kandziora, F. Mechanical concepts for disc regeneration. *Eur. Spine J.* **15**, 354–360 (2006).
2. Myers, E. R. & Wilson, S. E. Biomechanics of osteoporosis and vertebral fracture. *Spine* vol. **22**, 25S–31S (1997).
3. Chu, J. Y., Skrzypiec, D., Pollintine, P. & Adams, M. A. Can compressive stress be measured experimentally within the annulus fibrosus of degenerated intervertebral discs? *Proc. Inst. Mech. Eng. Part H J. Eng. Med.* **222**, 161–170 (2008).

4. Zhao, F. D., Pollintine, P., Hole, B. D., Adams, M. A. & Dolan, P. Vertebral fractures usually affect the cranial endplate because it is thinner and supported by less-dense trabecular bone. *Bone* **44**, 372–379 (2009).
5. Richardson, S. M. *et al.* Degenerate Human Nucleus Pulposus Cells Promote Neurite Outgrowth in Neural Cells. *PLoS One* **7**, (2012).
6. Stefanakis, M. *et al.* Annulus fissures are mechanically and chemically conducive to the ingrowth of nerves and blood vessels. *Spine (Phila. Pa. 1976).* **37**, 1883–1891 (2012).
7. Khan, A. N. *et al.* Inflammatory biomarkers of low back pain and disc degeneration: a review. *Ann. N. Y. Acad. Sci.* **1410**, 68–84 (2017).
8. Christian W.A. Pfirrmann, Alexander Metzdorf, Achim Elfering,Juerg Hodler, N. B. Effect of Aging and Degeneration on Disc Volume and Shape: A Quantitative Study in Asymptomatic Volunteers. *J. Orthop. Res. Sept.* **25**, 1121–1127 (2007).
9. Ma, J. *et al.* Is fractal dimension a reliable imaging biomarker for the quantitative classification of an intervertebral disk? *Eur. Spine J.* **29**, 1175–1180 (2020).
10. Jarman, J. P. *et al.* Intervertebral disc height loss demonstrates the threshold of major pathological changes during degeneration. *Eur. Spine J.* **24**, 1944–1950 (2015).
11. Lecun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
12. Jamaludin, A., Kadir, T. & Zisserman, A. SpineNet: Automated classification and evidence visualization in spinal MRIs. *Med. Image Anal.* **41**, 63–73 (2017).
13. Luoma, K. *et al.* Low back pain in relation to lumbar disc degeneration. *Spine (Phila. Pa. 1976).* **25**, 487–492 (2000).
14. Lootus, M., Kadir, T. & Zisserman, A. Automated radiological grading of spinal MRI. *Lect. Notes Comput. Vis. Biomech.* **20**, 119–130 (2015).
15. Pfirrmann, C. W. A., Metzdorf, A., Zanetti, M., Hodler, J. & Boos, N. Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine (Phila. Pa.* 1976). **26**, 1873–1878 (2001).
16. Castro-Mateos, I., Hua, R., Pozo, J. M., Lazary, A. & Frangi, A. F. Intervertebral disc classification by its degree of degeneration from T2-weighted magnetic resonance images. *Eur. Spine J.* **25**, 2721–2727 (2016).
17. Shinde, J. V., Joshi, Y. V. & Manthalkar, R. R. Multidomain Feature Level Fusion for Classification of Lumbar Intervertebral Disc Using Spine MR Images. *IETE J. Res.* **0**, 1–14 (2020).
18. Huang, J. *et al.* Spine Explorer: a deep learning based fully automated program for efficient and reliable quantifications of the vertebrae and discs on sagittal lumbar spine MR images. *Spine J.* **20**, 590–599 (2020).
19. Pang, S. *et al.* Direct automated quantitative measurement of spine by cascade amplifier regression network with manifold regularization. *Med. Image Anal.* **55**, 103–115 (2019).
20. Griffith, J. F. *et al.* Modified Pfirrmann grading system for lumbar intervertebral disc degeneration. *Spine (Phila. Pa. 1976).* **32**, 708–712 (2007).

21. Marois, B. & Syssau, P. Pratiques des banques françaises en termes d'analyse du risque-pays. *Rev. française Gest.* **32**, 77–94 (2006).
22. Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F. & Adam, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. *Proceedings of the European conference on computer vision (ECCV).* 801–818 (2018).
23. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* 770–778 (2016).
24. Liu, Z. *et al.* Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. *arXiv preprint arXiv:2103.14030* (2021).
25. Zhao, H., Shi, J., Qi, X., Wang, X. & Jia, J. Pyramid scene parsing network. *Proc. – 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017*. 2881–2890 (2017).
26. Li, X. *et al.* Weighted feature pyramid networks for object detection. *Proc. – 2019 IEEE Intl Conf Parallel Distrib. Process. with Appl. Big Data Cloud Comput. Sustain. Comput. Commun. Soc. Comput. Networking, ISPA/BDCloud/SustainCom/SocialCom.* 1500–1504 (2019)
27. Waldenberg, C., Hebelka, H., Brisby, H. & Lagerstrand, K. M. MRI histogram analysis enables objective and continuous classification of intervertebral disc degeneration. *Eur. Spine J.* **27**, 1042–1048 (2018).
28. Abdollah V, Parent EC, Battié MC. Reliability and validity of lumbar disc height quantification methods using magnetic resonance images. *Biomed Tech (Berl).* **64**, 111–117 (2018).
29. Dabbs VM, Dabbs LG. Correlation between disc height narrowing and low-back pain. *Spine (Phila Pa 1976).* **15**, 1366–9 (1990).

Figures

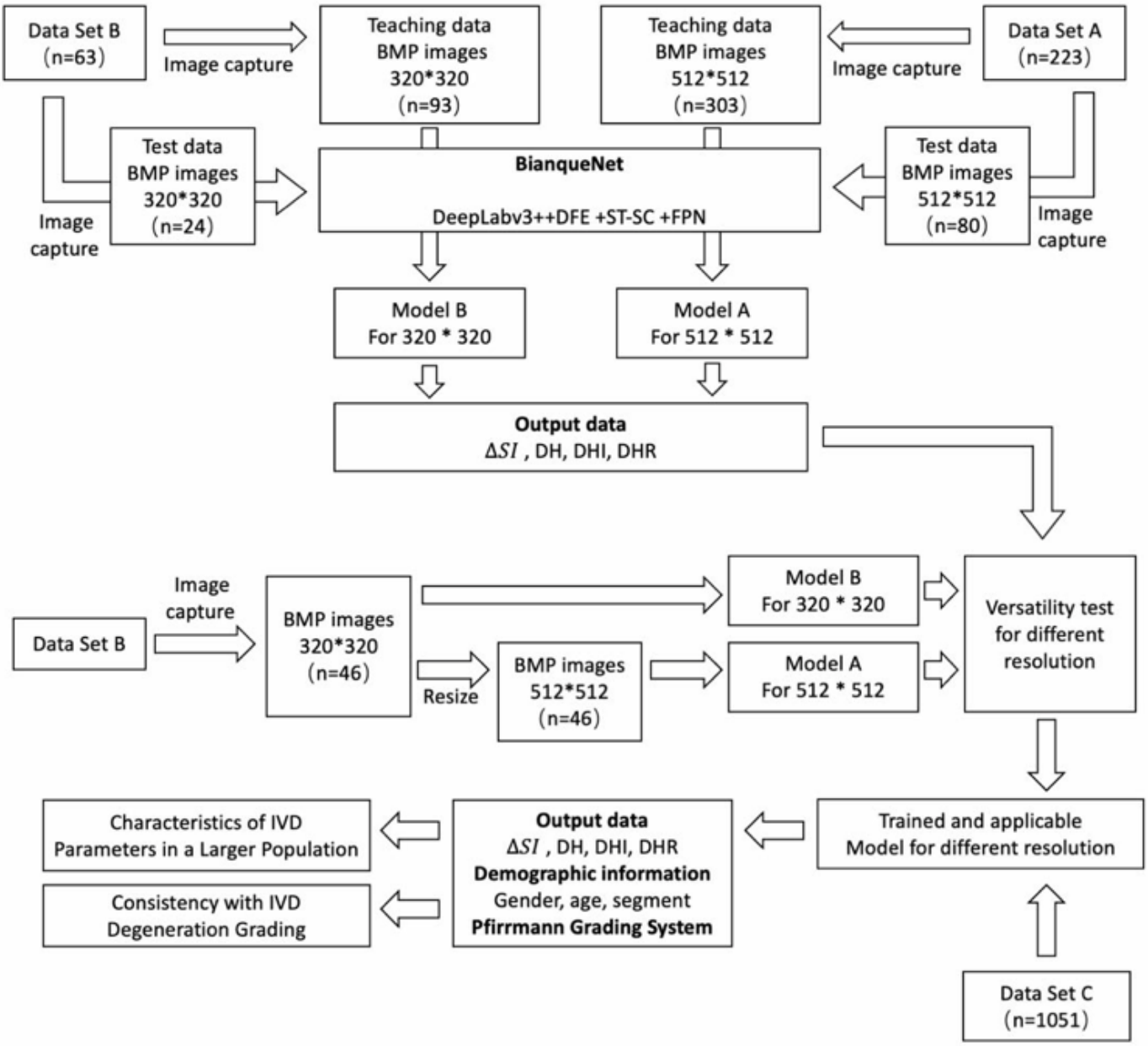

Figure 1

The flowchart of the study process

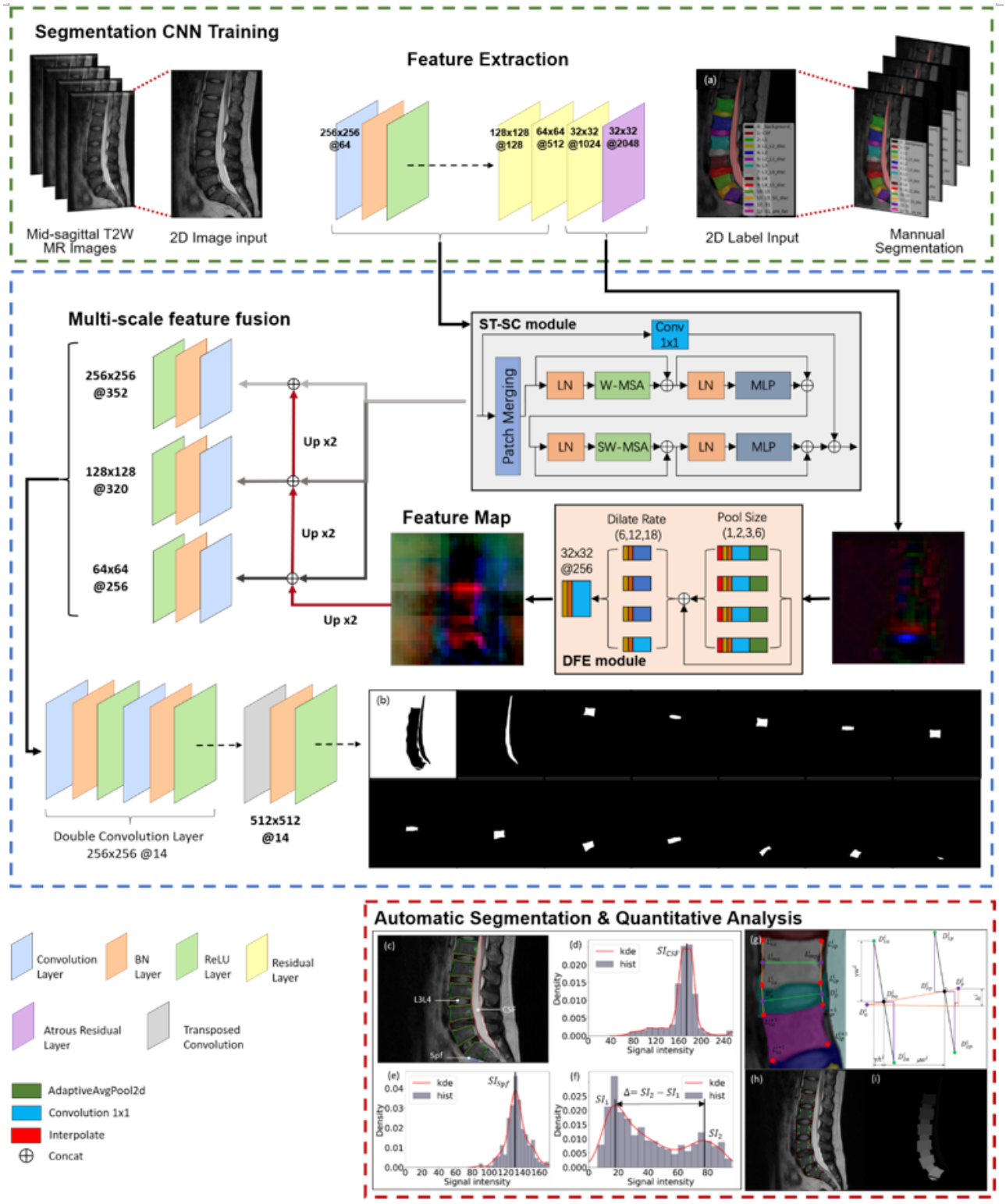

Figure 2

The fully automatic IVD quantitative analysis system based on semantic segmentation network. The proposed method consisted of segmentation CNNs with SW-SC module and DFE module, histogram-based signal intensity quantitation, and area-based fully automated geometric measurement. (a) Segmentation label of lumbar spine related area, (b) Each image channel output by the model corresponds to a segmentation area, (c) The outline of the segmented area is displayed on the original

image. Signal intensity histogram calculation (d) Cerebrospinal fluid area, (e) Presacral fat area, (f) L3L4 intervertebral disc area. (g) Intervertebral disc parameter calculation, (h) Vertebral body corner detection result (red point) and feature point calculation result (green point), (i) 80% area extraction result of the intervertebral disc center.

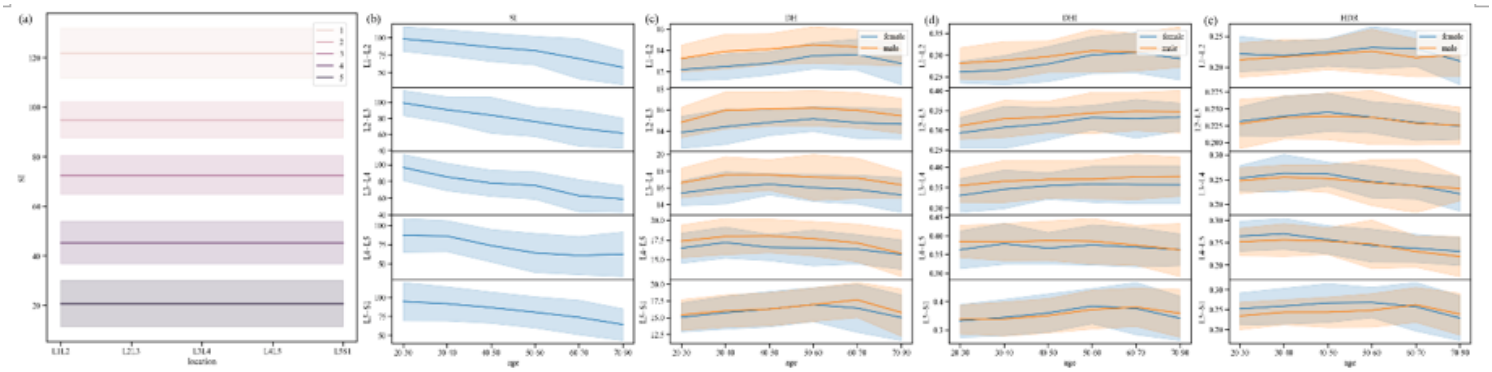

Figure 3

Characteristics of IVD Parameters ((a) The mean and standard deviation (σ) of the ΔSI of each the Modified Pfirrmann Grading System (level 1, 2, 3, 4, 5), ΔSI (b), DH (c), DHI (d) and HDR (e)) in different age, gender, and segments

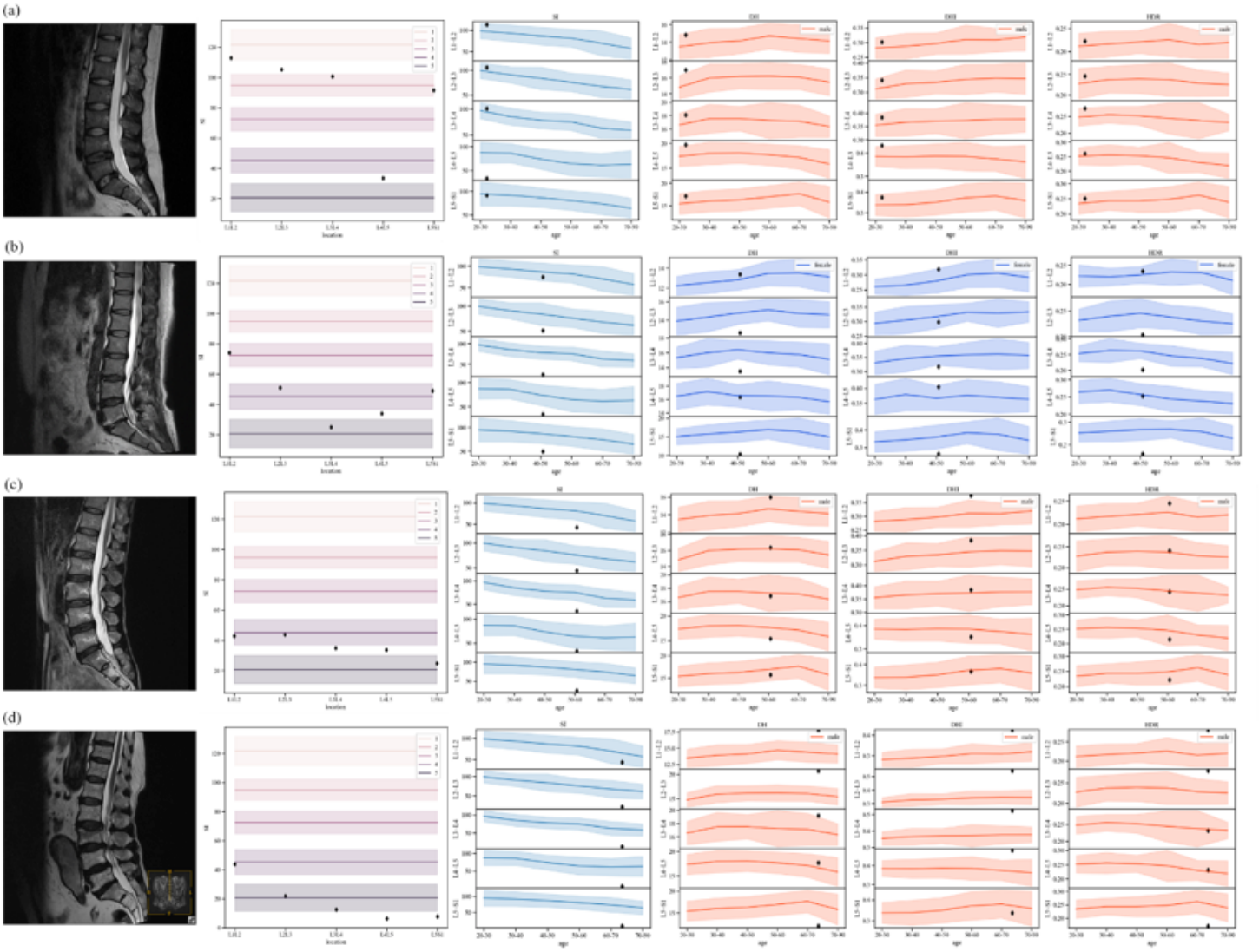

Figure 4

Quantitative analysis results of typical cases. (a) 23-year-old male(Longhua Hospital); (b) 49-year-old female(Dongzhimen Hospital, Beijing University of Chinese Medicine); (c) 63-year-old male(Guangdong Provincial Hospital of Chinese Medicine); (d) 81-year-old male(Shenzhen Pingle Orthopedics Hospital).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementFigures.pdf
- SupplementFile1signalintensityandgeographicmeasurementinIVD.docx
- SupplementFile2IVDquantitativeanalysiswithPfirrmannGrading.docx
- SupplementTable1.docx
- SupplementTable2.docx

- SupplementTable3.docx
- SupplementTable4.docx