Annals of Biomedical Engineering
https://doi.org/10.1007/s10439-025-03840-w

BMES BIOMEDICAL ENGINEERING SOCIETY

ORIGINAL ARTICLE

A CNN Autoencoder for Learning Latent Disc Geometry from Segmented Lumbar Spine MRI

Mattia Perrone[1] · D'Mar M. Moore[1] · Daisuke Ukeba[1] · John T. Martin[1]

Received: 6 March 2025 / Accepted: 3 September 2025
© The Author(s) under exclusive licence to Biomedical Engineering Society 2025

Abstract

Purpose Low back pain is the world's leading cause of disability and pathology of the lumbar intervertebral discs is frequently considered a driver of pain. The geometric characteristics of intervertebral discs offer valuable insights into their mechanical behavior and pathological conditions. In this study, we present a convolutional neural network (CNN) autoencoder to extract latent features from segmented disc MRI. Additionally, we interpret these latent features and demonstrate their utility in identifying disc pathology, providing a complementary perspective to standard geometric measures.

Methods We examined 195 sagittal T1-weighted MRI of the lumbar spine from a publicly available multi-institutional dataset. The proposed pipeline includes five main steps: (1) segmenting MRI, (2) training the CNN autoencoder and extracting latent geometric features, (3) measuring standard geometric features, (4) predicting disc narrowing with latent and/or standard geometric features and (5) determining the relationship between latent and standard geometric features.

Results Our segmentation model achieved an intersection over union (IoU) of 0.82 (95% CI 0.80–0.84) and dice similarity coefficient (DSC) of 0.90 (95% CI 0.89–0.91). The minimum bottleneck size for which the CNN autoencoder converged was 4×1 after 350 epochs (IoU of 0.9984—95% CI 0.9979–0.9989). Combining latent and geometric features improved predictions of disc narrowing compared to use either feature set alone. Latent geometric features encoded for disc shape and angular orientation.

Conclusions This study presents a CNN autoencoder to extract latent features from segmented lumbar disc MRI, enhancing disc narrowing prediction and feature interpretability. Future work will integrate disc voxel intensity to analyze composition.

Keywords Convolutional neural networks (CNN) · Autoencoder · Latent features · MRI segmentation · Features interpretability

Introduction

Low back pain ranks among the most prevalent musculoskeletal conditions worldwide. In 2020, 619 million people were estimated to be impacted by low back pain and projections indicate that this number will surge to 843 million by 2050 [1]. Lumbar disc degeneration is associated with low back pain [2], however the factors that drive disc degeneration and pain in an individual patient are not well understood. Geometric characteristics of intervertebral discs are recognized in the progression of disc degeneration. For example, disc narrowing is a hallmark feature of degeneration identified on clinical imaging [3]. Furthermore, discs with a smaller ratio of disc area to disc height are prone to degeneration [4] and local variations in disc geometry affect stresses at the disc-to-bone interface, which may lead to disc degeneration [5]. Thus, a thorough analysis of disc geometry may reveal underlying drivers of degeneration and low back pain.

Machine learning is often used to identify disc pathologies from imaging data [6–9]. Most algorithms implement convolutional neural networks (CNN) to perform radiological feature classification [10] including disc degeneration [11, 12]. Such models achieve accuracy comparable to that of expert radiologists [10, 13], however the features learned by the model often lack interpretability or are not evaluated and consequently are not explainable to clinicians. Clinicians are more likely to trust and adopt AI tools when they

Associate Editor Dan Ma oversaw the review of this article.

✉ John T. Martin
john_martin@rush.edu

[1] Rush University Medical Center, 1620 W Harrison St., Chicago, IL 60612, USA

Published online: 22 September 2025

comprehend their decision-making process [14], as understanding how a model reaches its conclusions supports informed decision-making [15]. Modeling approaches leveraging tabular patient demographics and extracted disc geometry have also been developed [16–18]. While this ensures interpretability of the predictive features, these models have lower performance than models trained on image data, as the latter exploit a greater amount of information during training [6]. Thus, developing a deep learning model with explainable features may improve the adoption of a model evaluating spine health in clinical practice.

Autoencoders are among the most common models used for computer vision tasks, from denoising [19, 20] to generating new samples [21, 22]. As the autoencoder architecture allows for mapping input data into a lower dimensional space, feature learning is a main application for autoencoders. For example, autoencoders are reliable models for dimensionality reduction and outperform linear [23] and nonlinear [24, 25] dimensionality reduction techniques. In spine research, autoencoders have been applied for automatic diagnosis of degenerative disc diseases. Specifically, extracting disc features in MRI using an autoencoder outperformed other manually extracted geometric features when predicting disc pathology [26]. Similarly, sparse autoencoders and diffusion autoencoders have been employed to enhance interpretability and identify critical features in spinal imaging [27, 28]. While these studies have demonstrated the effectiveness of autoencoders for feature extraction and prediction in spinal imaging, no studies have explored the relationship between these extracted latent features and geometric properties of the disc or the utility in using latent features to identify degenerative disc phenotypes.

This study introduces a CNN autoencoder to extract features from segmented lumbar disc MRI for use in predicting disc narrowing, a common spinal pathology relevant to low back pain [29–31]. We applied our model directly to image masks to isolate the role that spinal geometry plays in driving pathology. We hypothesized that integrating a latent geometric representation with standard geometric determinants (i.e., disc height, anteroposterior width, lateral width and disc inclination angles) will enhance disc narrowing prediction beyond the use of standard geometric measures alone. Furthermore, we provided insight into the latent features by evaluating their relationship to disc geometry.

Methods

Overview

We developed a computational pipeline (Fig. 1) to automatically segment intervertebral discs in lumbar spine MRI and

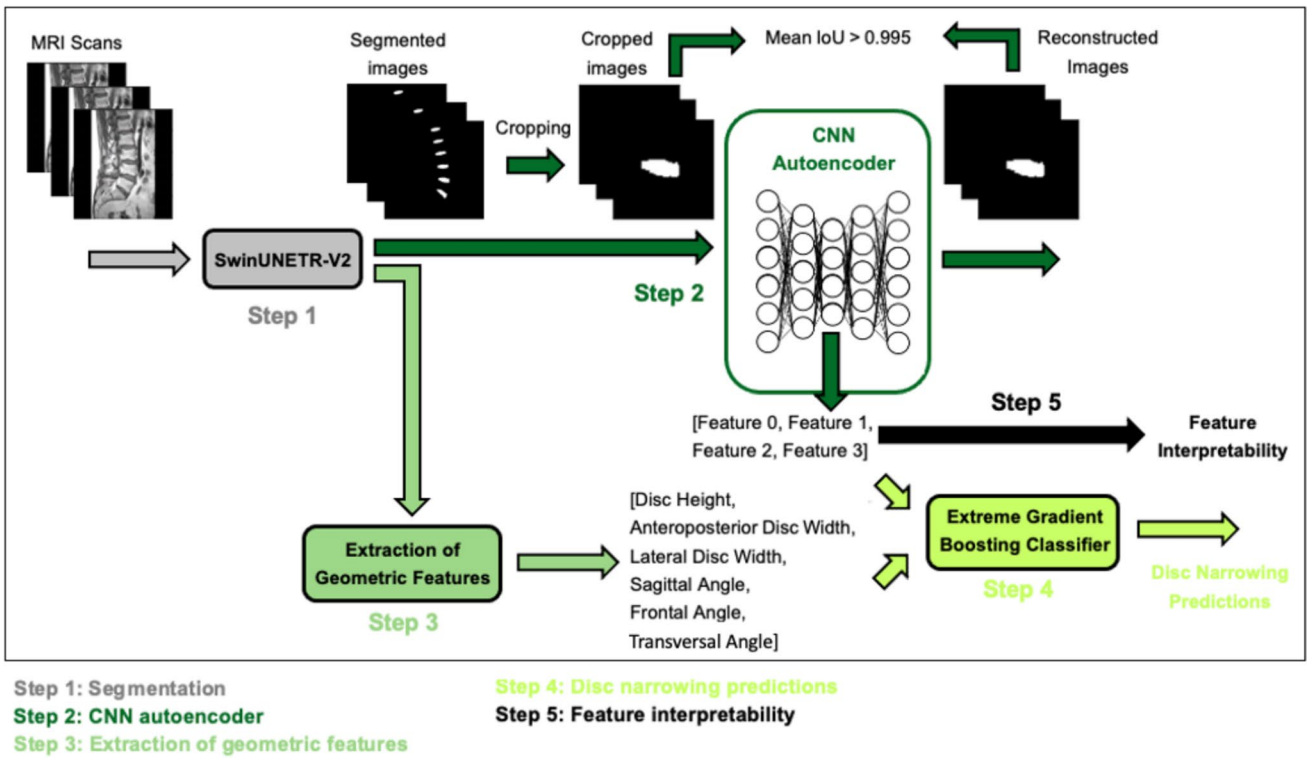

Fig. 1 Pipeline followed in the current study

to extract latent geometric features from each segmented disc. The computational pipeline included five steps: (1) MRI volumes were segmented using a Swin Transformer to generate disc masks; (2) disc masks served as inputs to a CNN autoencoder for unsupervised feature extraction; (3) standard geometric features were calculated from disc masks (4); a gradient boosting classifier was trained on latent and standard disc geometric features to predict disc narrowing; (5) a statistical evaluation was performed to determine the relationship between latent and standard geometric features.

We leveraged a public dataset comprised 195 sagittal T1-weighted lumbar spine MRI scans from patients with chronic low back pain ($N = 75$ Male, 120 female) [32]. A total of 1347 discs were included in our analysis. Imaging data were collected from different institutions (make: Siemens, Philips; field strength: 1.5–3 T; TE: 8 ms-124 ms; TR: 446–5570 ms; flip angle: 80–160°), resulting in a heterogeneous dataset to generalize across diverse imaging settings. For each patient, pixel-wise annotation of lumbar vertebrae, intervertebral discs and spinal cord is included. In addition, disc narrowing was labeled by a practicing spine surgeon (DU) and a medical student (DMM), with disagreements resolved by consensus.

All experiments were run on an NVIDIA RTX A5000 GPU. The whole pipeline required approximately 2 h in total.

Segmentation Model

For MRI automatic segmentation we tested three established architectures: U-Net 3D [33], SegResNet 3D [34] and SwinUNETR-V2 [35]. The best performing model, which was evaluated in terms of intersection over union and dice similarity coefficient, was the SwinUNETR-V2 (Supplementary Table 1). The SwinUNETR-V2 is a shifting window transformer model with convolutional blocks that enhance local feature extraction, achieving most accurate performance on multiple 3D biomedical imaging datasets [35]. We divided the dataset allocating 80% ($n = 156$) for training and validation and 20% for testing ($n = 39$). Preprocessing of MRI scans included normalization of pixel values and image resampling to a resolution of $1.4 \times 1.2 \times 1.5$ mm. The original voxel dimensions are reported in Ref. [35]. The model was trained for 200 epochs with batch size = 4, incorporating an early stopping mechanism when the validation IoU did not improve over 15 consecutive epochs. Batch size and learning rate were tuned on the validation set, resulting in final values of 4 and 0.001, respectively. A custom loss function that includes terms for binary cross-entropy and DSC was implemented for training. The two terms had equal weights in determining the final loss. The model with the highest validation IoU across all training epochs was retained for subsequent analysis.

CNN Autoencoder

For unsupervised feature extraction, a CNN autoencoder was trained using test Masks from the segmentation process. The CNN autoencoder architecture comprises 3D convolutional and deconvolutional layers within the encoder and decoder and batch normalization layers to promote training stability (Fig. 2). Individual discs were isolated from each spine resulting in a training set with 202 disc volumes from 39 spines. All data were used to train the model. These 39 spines correspond to the test set from step 1 (segmentation model) and their Masks were used to train the autoencoder, thus avoiding data leakage. To ensure robust evaluation on this multi-institutional dataset, we designated all patients from one institution as the test cohort of step 1. To ensure the disc is centered in each sample, we computed each segmented disc's center of Mass and cropped a region including the whole Mask around that point. Finally, to ensure consistent input dimensions, we zero-padded cropped disc to $64 \times 64 \times 64$ voxels . Binary cross-entropy was used as the loss function. Model convergence was defined when the IoU between input and output Masks exceeded 0.995,

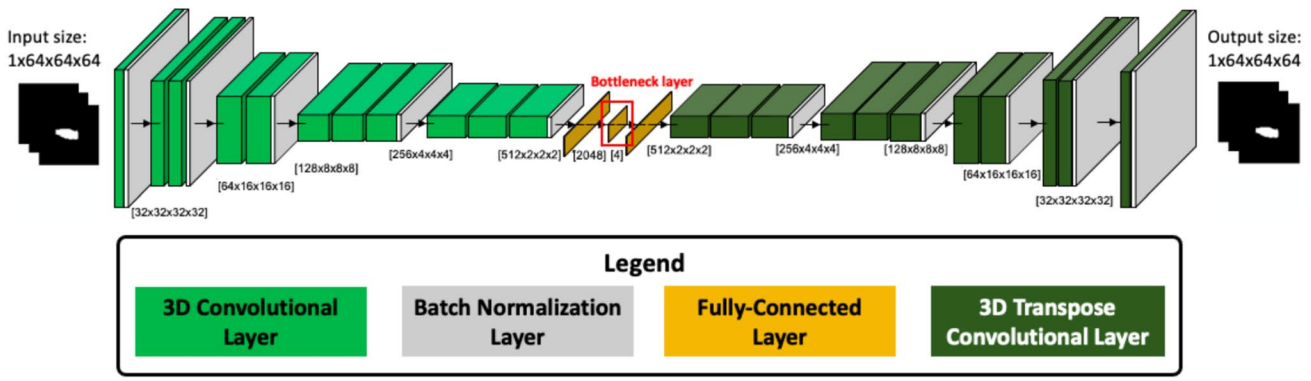

Fig. 2 Autoencoder architecture used for unsupervised feature extraction

indicating that the input mask reconstruction was nearly perfect. The autoencoder bottleneck layer dimension represents a crucial hyperparameter that defines the number of features extracted from the segmented image. The bottleneck layer size was tested with different dimensions (64×1, 8×1, 4×1 and 3×1) and the minimum effective bottleneck size ($n=4$) was selected to reduce multicollinearity among features and enhance their interpretability. The model converged after 350 epochs with batch size = 4. No skip connections were included to guarantee that all the information of the segmented image was mapped into the latent space.

Extraction of Geometric Features

Geometric features were automatically extracted from segmented images including disc height, anteroposterior disc width and lateral disc width, as well as the inclination of the disc relative to the horizontal axis in the sagittal plane and to the vertical axis in the transverse and coronal plane (Fig. 3). All measurements were computed via custom scripts, with no manual annotation. Disc height was calculated as the mean height across all pixels within the segmented disc volume (sagittal plane), anteroposterior disc width as the distance between the two lateral edges of the segmented disc on the middle slice (sagittal plane) and lateral disc width as the distance between the two lateral edges of the segmented disc on the middle slice (transverse plane).

A custom Python script was implemented to compute the angular measurements. This consists of identifying the principal axis of segmented discs using principal component analysis (PCA) on the pixel coordinates and calculating the angle of inclination relative to the horizontal or vertical axis, depending on the anatomical plane.

Gradient Boosting

A gradient boosting classifier (XGBoost [36]) was implemented to predict disc narrowing (max estimator depth = 3, number of estimators = 200, learning rate = 0.1). Since disc narrowing prediction is a binary classification problem, F1-score and area under the receiver operating characteristic curve (AUC ROC) were used as evaluation metrics. F1-score was specifically chosen to account for the data imbalance, as narrowed discs constitute approximately 20% of the dataset. Fivefold cross-validation was performed to ensure the reliability of the results obtained.

Feature Interpretability

To interpret the latent features, we designed an experiment to systematically interrogate how latent features were related to 3D disc geometry. To do so, a synthetic dataset was generated by fixing three of the four latent features at their respective mean and varying the fourth feature from its minimum to Maximum value. For each latent feature, predicted disc

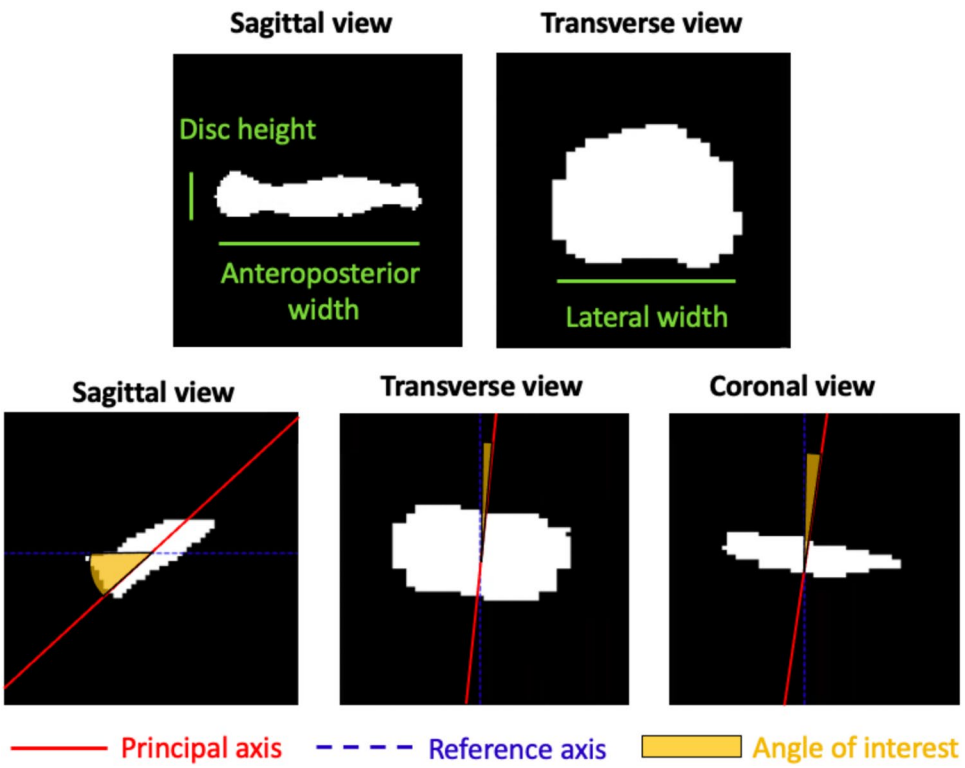

Fig. 3 Visualization of geometric features extracted from lumbar disc masks: disc height, anteroposterior width and lateral width (top row), along with the angle of interest computed between the principal axis (red) and the reference axis (blue) in sagittal, transverse and coronal views (bottom row). All measurements except disc height were performed on the middle slice (dimensions in pixels)

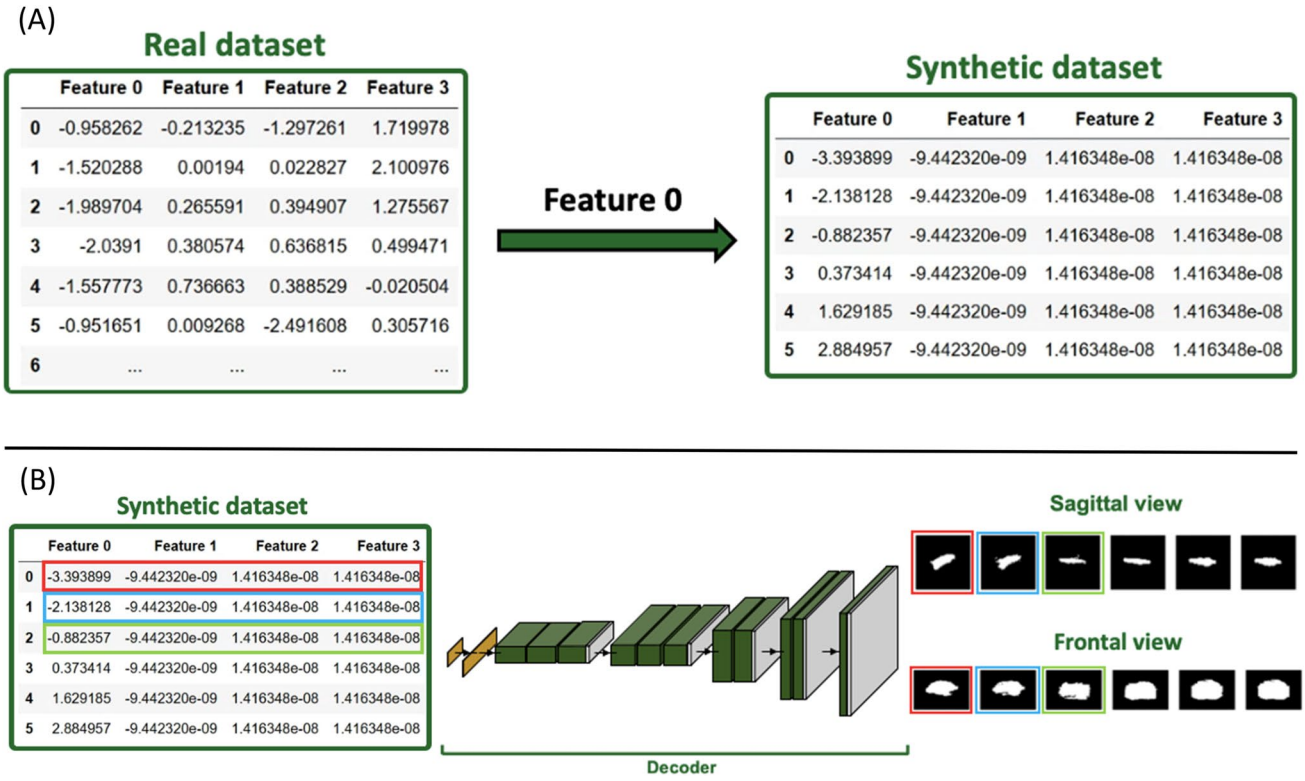

Fig. 4 Overview of the interpretability analysis for latent features: **A** Generation of the synthetic latent feature dataset starting from the real latent feature dataset by varying latent features, with the process illustrated specifically for Feature 0. **B** Decoder architecture used to reconstruct images from the synthetic dataset, with examples of sagittal and frontal views of reconstructed masks. Input features and reconstructed images are marked with matching frame colors, indicating that the former serves as input to the decoder to produce the latter

Masks were generated for geometry measurements and visualization at the minimum, 1/5, 2/5, 3/5 and 4/5, and maximum of the range (Fig. 4).

Results

For segmentation via the Swin Transformer, we obtained an Intersection over Union (IoU) of 0.82 (95% CI 0.80–0.84) and Dice similarity coefficient (DSC) of 0.90 (95% CI 0.89–0.91), which aligns to what reported in the literature for the same dataset [32]. The autoencoder achieved an IoU of 0.9981 (95% CI 0.9976–0.9986) using a bottleneck size of 64×1 in 190 epochs. Reducing the bottleneck to 8×1 improved the IoU to 0.9983 (95% CI 0.9979–0.9987) after 199 epochs, while further reduction to 4×1 resulted in the highest IoU of 0.9984 (95% CI 0.9979–0.9989) in 350 epochs. In contrast, a bottleneck of 3×1 did not converge after 1000 epochs, with the IoU oscillating between 0.6 and 0.8. Based on these results, we selected a bottleneck size of 4×1 to minimize multicollinearity.

Results for predicting disc narrowing are reported in Table 1 (ground truth masks) and Table 2 (predicted masks).

Table 1 Predictions of disc narrowing using latent and geometric features from ground truth masks

Features	F1-score (95% CI)	AUC ROC (95% CI)
Latent features	0.81 (0.77–0.85)	0.75 (0.71–0.79)
Geometric features	079 (0.77–0.81)	0.66 (0.63–0.69)
Geometric and latent features	0.81 (0.79–0.83)	0.76 (0.73–0.78)
Lin et al. [17]	0.78	–
Hung et al. [16]	–	0.75 (0.71–0.79)

Comparisons made to relevant examples in literature

Table 2 Predictions of disc narrowing using latent and geometric features from predicted masks

Features	F1-score (95% CI)	AUC ROC (95% CI)
Latent features	0.77 (0.75–0.78)	0.58 (0.84–0.62)
Geometric features	0.79 (0.77–0.81)	0.66 (0.63–0.69)
Geometric and latent features	0.79 (0.77–0.81)	0.69 (0.63–0.73)

Disc narrowing predictions improved when combining geometric and latent features compared to using either feature set individually. To mitigate the skewness of the dataset, we additionally experimented with class weighted loss function; however, this approach did not lead to improved model performance and was therefore not adopted in the final implementation (Supplementary Table 2)

Finally, we analyzed the correlation between geometric and latent features using correlation coefficients to identify significant relationships between the two (Fig. 5). To do so, we generated a synthetic dataset and visualized how each latent feature influenced disc geometry (Figs. 4 and 5). We then examined the relationship between latent and geometric features by calculating a correlation matrix and found that some latent features are strongly correlated with geometric features—e.g., Feature 2 and anteroposterior width (corr coeff = − 0.80; p-value < 0.001) and Feature 0 and sagittal angles (corr coeff = − 0.75; p-value < 0.001) (Fig. 5). Of note, latent features were not correlated with other latent features (corr coeff < 0.05).

Discussion

Disc geometry plays a key role in disc health and mechanical function [37, 38]. To better understand disc geometry, we developed a CNN autoencoder to extract latent features from segmented Lumbar disc MRI and determine what geometry latent features encode and whether they were related to disc degeneration. We identified that a minimum set of 4 latent features was capable of fully encoding disc geometry, where individual features correlated to disc size and orientation. Furthermore, prediction models using latent features alone

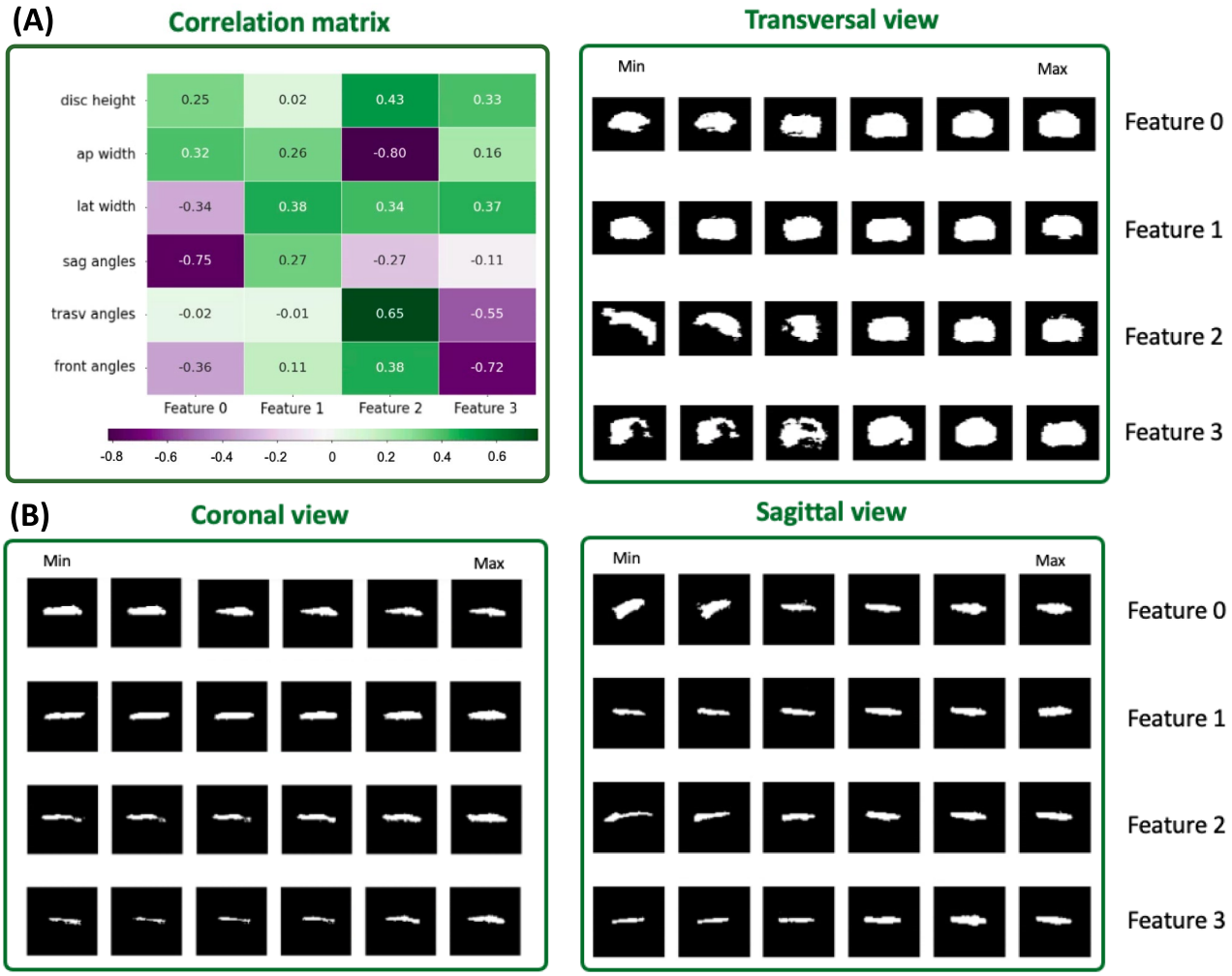

Fig. 5 **A** Correlation analysis of latent features and geometric features. **B** Visualization of reconstructed masks across transversal, coronal and sagittal views for different latent features, ranging from minimum to maximum values. For clarity, only the central slice in each plane is displayed

were as effective at identifying disc narrowing as those using standard disc geometry.

Latent features fully describe disc geometry, are predictive of disc narrowing, and can seamlessly fit into existing computational workflows. Our disc autoencoder generates a set of orthogonal features that spans the latent geometric space. While there are many ways to parameterize disc geometry (anterior height, posterior height, sagittal angle, coronal angle, cross-sectional area, volume, etc.) in establishing a lossless autoencoder, we demonstrate that these latent features are the minimum feature set required to fully recapitulate 3D disc geometry. By establishing correlations between latent features, disc shape (e.g., Feature 2 and anteroposterior width), and orientation (e.g., Feature 0 and sagittal angle), we demonstrate that latent features also have real-world interpretability that may be clinically meaningful. Furthermore, predictive models incorporating latent features are as effective in identifying disc narrowing as those using standard geometry alone, and that combining latent and geometric features modestly boosts AUC-ROC on models trained either both ground truth or predicted masks. Because latent features additional manual labeling and with minimal computational requirements, these descriptors can seamlessly fit into multimodal pipelines that incorporate images, clinical text, or demographic variables. Thus, we propose that latent features may be clinically relevant for spine disease and low back pain, and since this approach is anatomy-independent, it may be useful for other musculoskeletal applications as well. Future work to establish relationships between latent features and clinical symptoms is warranted. Still, some combinations of latent features fail to preserve the anatomical proportions of the discs (e.g., Fig 5, transversal view—row 3, column 1; transversal view—row 4, column 3). We surmise that some unnatural disc configurations exist in the latent feature space, and that development of a new loss function that penalizes unnatural configurations is warranted.

The first step of our computational pipeline is segmentation, a foundational task in medical imaging studies. We utilized a Swin Transformer model [35], which produces results (IoU: 0.82; DSC: 0.90) consistent with those reported by other groups using the same dataset (32). Importantly, the dataset used in this study comprises data from four different institutions, introducing heterogeneity that exists in multi-institutional studies or clinical trials, but likely reduces the segmentation model's performance compared to single-institution studies [39]. Our approach to dimensionality reduction using a CNN autoencoder is a nonlinear extension of the linear approach of Peloquin et al. [40] (principal components). While these two studies use different methods to reduce segmented disc images to latent geometric features (deep learning versus statistical shape modeling; multiple spinal levels versus single level), both studies demonstrate that a parsimonious set of features can parameterize a wide range of disc geometry, revealing an unexpected simplicity in disc shape.

Our approach to predict disc pathology using standard and latent geometric features produced similar results to models that do not use latent geometric features. Lin et al. [17] developed a classification model from disc height and anthropometrics (age, sex and BMI). Hung et al. [16] conducted a similar study, implementing two different models for the same prediction task: one leveraging anthropometrics (age, sex, body height and body weight) and another using both anthropometrics and geometric features (disc height and disc depth). Our model predictions incorporating latent and standard geometric features matched those of Lin et al. [17] and Hung et al. [16]. In both investigations, patient anthropometrics proved relevant for making predictions of disc pathologies. Anthropometric variables may further improve the performance of models including latent and standard geometric features. However, these were absent from the public dataset used in the current study.

This study introduces a CNN autoencoder for extracting latent features from segmented MRI images of lumbar discs. These latent features provide complementary information to standard disc geometric features, as demonstrated by improved metrics in predicting disc narrowing. Furthermore, the proposed approach enhances feature interpretability by elucidating the relationship between latent and geometric features. Limitations of this study include the use of a dataset with relatively few cases of disc narrowing, potentially affecting the generalizability of the findings, which we mitigated by using F1-score to account for class skewness. Moreover, our publicly available dataset did not include any demographics or pain scores. An open source labeled/segmented spine MRI dataset with relevant demographics and clinical factors would advance spine research. Future work will explore integrating latent variables related to disc voxel intensity to incorporate valuable insights from disc composition and look at the relationship between latent features and clinical outcomes, as well as incorporating anatomy-aware constraints to improve biological plausibility of synthetic outputs.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10439-025-03840-w.

Funding This work was supported by the National Institutes of Health (NIH) under Grant R00AR077685.

Declarations

Conflict of interests The authors have no conflicts of interest to declare.

References

1. Ferreira, M. L., K. De Luca, L. M. Haile, J. D. Steinmetz, G. T. Culbreth, M. Cross, et al. Global, regional, and national burden of low back pain, 1990–2020, its attributable risk factors, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021. *Lancet Rheumatol.* 5(6):e316–e329, 2023.
2. De Schepper, E. I. T., J. Damen, J. B. J. Van Meurs, A. Z. Ginai, M. Popham, A. Hofman, et al. The association between lumbar disc degeneration and low back pain: The influence of age, gender, and individual radiographic features. *Spine.* 35(5):531–536, 2010.
3. Pfirrmann, C. W. A., A. Metzdorf, M. Zanetti, J. Hodler, and N. Boos. Magnetic resonance classification of lumbar intervertebral disc degeneration. *Spine.* 26(17):1873–1878, 2001.
4. Natarajan, R. N., and G. B. J. Andersson. The influence of lumbar disc height and cross-sectional area on the mechanical response of the disc to physiologic loading. *Spine.* 24(18):1873–1881, 1999.
5. Fleps I, Newman HR, Elliott DM, Morgan EF. Geometric determinants of the mechanical behavior of image-based finite element models of the intervertebral disc. *J Orthop Res.* 42(6):1343–1355, 2024.
6. Hussain M, Koundal D, Manhas J. Deep learning-based diagnosis of disc degenerative diseases using MRI: A comprehensive review. *Comput Electr Eng.* 105:108524, 2023.
7. Liawrungrueang W, Cholamjiak W, Sarasombath P, Jitpakdee K, Kotheeranurak V. Artificial Intelligence Classification for Detecting and Grading Lumbar Intervertebral Disc Degeneration. *Spine Surg Relat Res.* 8(6):552–559, 2024.
8. Baur, D., R. Bieck, J. Berger, P. Schöfer, T. Stelzner, J. Neumann, et al. Automated three-dimensional imaging and Pfirrmann classification of intervertebral disc using a graphical neural network in sagittal magnetic resonance imaging of the lumbar spine. *J Imaging Inform Med.* 2024. https://doi.org/10.1007/s10278-024-01251-2.
9. Yi, W., J. Zhao, W. Tang, H. Yin, L. Yu, Y. Wang, et al. Deep learning-based high-accuracy detection for lumbar and cervical degenerative disease on T2-weighted MR images. *Eur Spine J.* 32(11):3807–3814, 2023. https://doi.org/10.1007/s00586-023-07641-4.
10. Jamaludin, A., M. Lootus, T. Kadir, A. Zisserman, J. Urban, M. C. Battié, et al. ISSLS PRIZE IN BIOENGINEERING SCIENCE 2017: Automation of reading of radiological features from magnetic resonance images (MRIs) of the lumbar spine without human intervention is comparable with an expert radiologist. *Eur Spine J.* 26(5):1374–1383, 2017.
11. Ma, S., Y. Huang, X. Che, and R. Gu. Faster RCNN-based detection of cervical spinal cord injury and disc degeneration. *J Appl Clin Med Phys.* 21(9):235–243, 2020.
12. Liawrungrueang W, Kim P, Kotheeranurak V, Jitpakdee K, Sarasombath P. Automatic detection, classification, and grading of lumbar intervertebral disc degeneration using an artificial neural network model. *Diagnostics.* 13(4):663, 2023.
13. Compte, R., I. Granville Smith, A. Isaac, N. Danckert, T. McSweeney, P. Liantis, et al. Are current machine learning applications comparable to radiologist classification of degenerate and herniated discs and Modic change? A systematic review and meta-analysis. *Eur Spine J.* 32(11):3764–3787, 2023. https://doi.org/10.1007/s00586-023-07718-0.
14. Amann, J., A. Blasimme, E. Vayena, D. Frey, and V. I. Madai. Explainability for artificial intelligence in healthcare: a multidisciplinary perspective. *BMC Med Inform Decis Mak.* 20(1):1–9, 2020. https://doi.org/10.1186/s12911-020-01332-6.
15. Lu, S. C., C. L. Swisher, C. Chung, D. Jaffray, and C. Sidey-Gibbons. On the importance of interpretable machine learning predictions to inform clinical decision making in oncology. *Front Oncol.* 13:1–13, 2023.
16. Hung, I. Y. J., T. T. F. Shih, B. B. Chen, and Y. L. Guo. Prediction of lumbar disc bulging and protrusion by anthropometric factors and disc morphology. *Int J Environ Res Public Health.* 18(5):1–13, 2021.
17. Lin PC, Chang WS, Hsiao KY, Liu HM, Shia BC, Chen MC, et al. Development of a machine learning algorithm to correlate lumbar disc height on X-rays with disc bulging or herniation. *Diagnostics.* 14(2):134, 2024.
18. Kocaman, H., H. Yıldırım, A. Gökşen, and G. M. Arman. An investigation of machine learning algorithms for prediction of lumbar disc herniation. *Med Biol Eng Comput.* 61(10):2785–2795, 2023. https://doi.org/10.1007/s11517-023-02888-x.
19. Bajaj, K., D. K. Singh, and M. A. Ansari. Autoencoders based deep learner for image denoising. *Procedia Comput Sci.* 2020(171):1535–1541, 2019. https://doi.org/10.1016/j.procs.2020.04.164.
20. Xiang Q, Pang X. Improved Denoising Auto-Encoders for Image Denoising. Proc—2018 11th Int Congr Image Signal Process Biomed Eng Informatics, CISP-BMEI 2018. 2018;1–9.
21. Kingma DP, Welling M. Auto-encoding variational bayes. 2nd Int Conf Learn Represent ICLR 2014—Conf Track Proc. 2014;(Ml):1–14.
22. Yoo, J., H. Lee, and N. Kwak. Unpriortized Autoencoder for Image Generation. *Proc—Int Conf Image Process ICIP.* 2020:763–767, 2020.
23. Wang W, Huang Y, Wang Y, Wang L. Generalized autoencoder: A neural network framework for dimensionality reduction. IEEE Comput Soc Conf Comput Vis Pattern Recognit Work. 2014;496–503.
24. Wang, Y., H. Yao, and S. Zhao. Auto-encoder based dimensionality reduction. *Neurocomputing.* 184:232–242, 2016. https://doi.org/10.1016/j.neucom.2015.08.104.
25. Ardelean, E. R., A. Coporîie, A. M. Ichim, M. Dînşoreanu, and R. C. Mureşan. A study of autoencoders as a feature extraction technique for spike sorting. *PLoS One.* 18(3):1–29, 2023.
26. Oktay, A. B., and Y. S. Akgul. Diagnosis of degenerative intervertebral disc disease with deep networks and SVM BT. *Symp Comput Inf Sci.* 659:253–261, 2016. https://doi.org/10.1007/978-3-319-47217-1_27.
27. Abdulaal A, Fry H, Montaña-Brown N, Ijishakin A, Gao J, Hyland S, et al. An X-ray is worth 15 features: Sparse autoencoders for interpretable radiology report generation. 2024;1–40. http://arxiv.org/abs/2410.03334
28. Atad M, Schinz D, Moeller H, Graf R, Wiestler B, Rueckert D, et al. Counterfactual explanations for medical image classification and regression using diffusion autoencoder. 2024;2(iMIMIC):2103–25. https://arxiv.org/abs/2408.01571v1
29. Raastad, J., M. Reiman, R. Coeytaux, L. Ledbetter, and A. P. Goode. The association between lumbar spine radiographic features and low back pain: A systematic review and meta-analysis. *Semin Arthritis Rheum.* 44(5):571–585, 2015.
30. Goode AP, Carey TS, Jordan JM. Low back pain and lumbar spine osteoarthritis: How are they related? *Curr Rheumatol Rep.* 15(2):305, 2013.
31. Kremer JM. Rheumatological creativity and leflunomide. *J Rheumatol.* 31:203–204, 2004.
32. van der Graaf JW, van Hooff ML, Buckens CFM, Rutten M, van Susante JLC, Kroeze RJ, et al. Lumbar spine segmentation in MR images: a dataset and a public benchmark. 2023; http://arxiv.org/abs/2306.12217
33. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In International conference on medical image

computing and computer-assisted intervention 2016 Oct 2 (pp. 424–432). Cham: Springer International Publishing.
34. MONAI Consortium. SegResNet: A 3D Residual UNet-like Network [Internet]. MONAI Documentation. Available from: https://docs.monai.io/en/stable/networks.html#segresnet. Accessed 25 Jul 2025.
35. He Y, Nath V, Yang D, Tang Y, Myronenko A, Xu D. SwinUNETR-V2: Stronger Swin Transformers with Stagewise Convolutions for 3D Medical Image Segmentation. Vol. 14223 LNCS, Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer Nature Switzerland; 2023, pp 416–426. https://doi.org/10.1007/978-3-031-43901-8_40
36. Chen T, Guestrin C. XGBoost: A scalable tree boosting system. Proc ACM SIGKDD Int Conf Knowl Discov Data Min. 2016;13–17 Aug:785–94.
37. Jeong JG, Kang S, Jung GH, Cho M, Kim H, Kim KT, et al. Biomechanical effect of disc height on the components of the lumbar column at the same axial load: A finite-element study. *J Healthc Eng*. 2022:7069448, 2022.
38. Lim, S., R. D. Huff, J. E. Veres, D. Satish, and G. D. O'Connell. Disc geometry measurement methods affect reported compressive mechanics by up to 65%. *JOR Spine*. 5(3):2–9, 2022.
39. Wang A, Zou C, Yuan S, Fan N, Du P, Wang T, et al. Deep learning assisted segmentation of the lumbar intervertebral disc: A systematic review and meta-analysis. *J Orthop Surg Res.* 19:496, 2024.
40. Peloquin, J. M., J. H. Yoder, N. T. Jacobs, S. M. Moon, A. C. Wright, E. J. Vresilovic, et al. Human L3L4 intervertebral disc mean 3D shape, modes of variation, and their relationship to degeneration. *J Biomech.* 47(10):2452–2459, 2014. https://doi.org/10.1016/j.jbiomech.2014.04.014.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.