YOLOv11 Based Classification of Lumbar Spine Degenerative Changes Across Multi-Modal Imaging

Vaibhavi Patel[1], Apurv Chudasama*[2], Amit Thakkar[3]

Department of Computer Science and Engineering Chandubhai S Patel Institute of Technology, CHARUSAT Anand, Gujarat, India

[1]vaibhavipatel.cse@charusat.ac.in, [2]apurvchudasama.tech@gmail.com [3]amitthakkar.it@charudat.ac.in

* Corresponding author

doi: https://doi.org/10.21467/proceedings.7.6.2

Abstract

The Lumbar Spines degenerative changes are a big source of morbidity and hence need correct and efficient diagnostic methodologies. This study leverages the YOLOv11x model for the classification of spinal levels from L1/L2 up to L5/S1 for three modalities: Sagittal T1, Sagittal T2/STIR and Axial T2. Post-acquisition processing of DICOM images, annotation conversions, and dataset augmentations were tedious procedures that were executed in preparation for model training which ensures that the model performs efficiently. The YOLOv11x architecture demonstrated high precision and recall across all modalities, which satisfies the state-of-the-art mAP scores because of its multi-scale detection capability. The results show the potential of deep learning to automate lumbar spine classification, which would tremendously improve the accuracy and consistency of diagnoses, and reduce inter-observer variability. This implementation would provide a basis for embedding AI-based technologies into clinical workflows, opening new avenues for enhanced patient care and scalable diagnostic solutions. Future scope includes leveraging the model on different spinal areas and testing it on a real-time basis in clinical domain.

Keywords: Lumbar Spine, YOLOv11x, Degenerative Classification.

I. Introduction

The capability of deep learning for the classification of degenerative changes in lumbar regions has been selected as one of the most vital areas of research to impact diagnosis of diseases and clinic workflows significantly[1]. Conditions such as neural foraminal narrowing, spondylolisthesis, stenosis that has the effect on the spine require proper detection of spinal levels and conditions for proper treatment planning [2]. Conventional diagnostics involve manual annotation and interpretation of medical images that takes more time and not consistent across observers [3]. Advancement in artificial intelligence and deep learning have made inroads to overcoming these difficulties and barriers through an automated classification system that will be more accurate and efficient [4]. This implementation uses state-of-the-art YOLOv11x architecture for the classification of lumbar levels (L1/L2 to L5/S1) using three imaging modalities: Sagittal T1, Sagittal T2/STIR, and Axial T2[5]. A complete new approach is selected for the processing of medical imaging data, optimization of the model architecture, and ensure robustness for the multiple classes [6]. With the included advanced pre- processing techniques, a different custom dataset annotation, and an efficient strategy for training [7], this research aims to be state- of-the-art in terms of detection and classification of degenerative conditions [8]. Results have proved that YOLOv11x achieves very high precision and recalls for detection and classification [9], making it an appropriate tool for radiologists in diagnosing abnormalities in the lumbar spine [10].

Proceedings DOI: 10.21467/proceedings.7.6; Series: AIJR Proceedings; ISSN: 2582-3922; ISBN: 978-81-989164-9-5

II. Literature Review

TABLE I: Comparison of Research Studies on Medical Image Datasets

Study	Dataset Type	Size	Annotations
Lehnen et al. (2021)	MRI scans	Small-scale	Degenerative changes
Hallinan et al. (2021)	Lumbar spine MRI dataset	Medium-scale	Stenosis classifications
Mushtaq et al. (2022)	Lumbar spine dataset	Medium-scale	Vertebral deformities
Trinh et al. (2022)	X-ray images	Medium-scale	Spondylolisthesis
Ruchi et al. (2023)	Multiple MRI datasets	Large-scale	Disease classifications
Nisar et al. (2024)	MRI dataset	Large-scale	Disc classifications
Payne et al. (2024)	Cervical spine MRI	Large-scale	Stenosis and cord compression

The classification of degenerative changes of the lumber spine through deep learning is currently one of the most significant topics of research because AI has the potential to enhance disease diagnosis and the workflow simplification in clinical settings [4]. This structured review represents a thorough analysis of seven significant studies in the domain and looks into the methods, datasets, innovations, and clinical results of such research [6]. As shown in Table I, different numbers of efforts have been made in organizing the studies, which have made use of datasets of variable scale and annotation in enhancing spine imaging.

Lehnen et al. (2021) [11] undertook an initial study on the employment of convolutional neural networks (CNNs) for the recognition of degenerative changes in MRIs of the lumbar spine. The feasibility study had some limitations in terms of the number and diversity of patients in the dataset. However, it was presented that using such techniques, degenerative conditions like disc degeneration and facet arthropathy could be successfully detected. The necessity for building larger, more multiple datasets and stronger architectures to issue this clinical variability was emphasized. Hallinan et al. (2021) [1] developed a deep learning technique that can detect and classify central canal stenosis, lateral recess stenosis, and neuropathic foraminal stenosis with the help of lumbar spine MRI [12]. This implementation integrated various CNNs through ensemble learning to improve the classification accuracy significantly. Moreover, the usage of a highly annotative dataset provided from the radiologists' perspective ensured the soundness of the research with respect to clinical validity. This study, therefore, underscores the potential of automated systems to aid radiologists in their involvement in complex diagnostic scenarios, thereby reducing subjectivity and inter-observer variability.

The study of Mushtaq et al. (2022) [10] was all about detecting deformities in lumbar spine through hybrid technique: edge-segmentation with deep learning techniques. Such anatomically relevant-algorithm localized vertebrae and detected other spines deformities such as scoliosis and spondylolisthesis. Applying spatial context was the new development from this form of-catching body's parts, which enhanced the fits segmentation accuracy of difficult cases. This is new concept evidence for how domain-specific knowledge were put to undertake in model design towards clinically relevant results. Trinh et al. (2022) [2] presented a deep learning architecture for identifying lumbar spondylolisthesis from X-ray images. This project was different from other studies that used MRI because it demonstrated the applicability of AI methods to different imaging modalities. Using a specific architecture for a specific modality at the same time enabled the model to overcome the peculiarities of X-ray imaging such as lower resolution and overlapping structures. Most notable is the emphasis placed on developing a lightweight architecture, the suitability of which is optimized for clinical deployment. Nisar et al. (2024) [13] have taken the research into entirely new territories by combining CNNs with transformers to detect and classify the lumbar intervertebral discs. The hybrid model here demonstrated state-of-the-art results on a large-scale dataset, proving the point that the architecture can capture both spatial and contextual features in one framework. This would not only improve the diagnostic paradigms but also establish the scalability of hybrid architectures for addressing complex medical imaging challenges. In a related application, Payne et al. (2024) [14] demonstrated a new hybrid methodology to detect the cervical spinal- occupying disease and cord compression using a vision transformer model and rules-based classification. Though this paper focuses on cervical spine imaging, it can draw ideas toward lumbar applications. The study showed the combination of data- driven and rules-based methods in providing high accuracy and interpretability, especially while recognizing the complex abnormality in the spinal region. According to Table II, different studies adopt different methodologies and put forth varied features to enhance medical imaging analysis.

TABLE II: Study Methodologies and Unique Features

Study	Methodology	Unique Features
Lehnen et al. (2021)	CNN-based classification	Feasibility study with small dataset
Hallinan et al. (2021)	Ensemble learning	Radiologist annotations for robustness
Mushtaq et al. (2022)	Hybrid deep learning	Integration of spatial context
Trinh et al. (2022)	Custom CNN for X-ray	Modality-specific optimizations
Ruchi et al. (2023)	Enhanced CNN with ROI	Focused on signal-to-noise improvement
Nisar et al. (2024)	CNN-Transformer hybrid	State-of-the-art performance
Payne et al. (2024)	Vision transformer and rules-based classification	Combined data-driven and rules-based methods

III. METHODOLOGY

This Section showcases the techniques that were applied to classify degenerative lumbar spine diseases with YOLOv11x. It includes the description of the dataset as well as data preprocessing and transformation, details about the architecture of the model, and the training process. Each step is targeted at making the model perform well and accurately classify cases, and also stretches its ability to detect rather subtle abnormality features along the spine.

A. Dataset Description

The RSNA 2024 Lumbar Spine Degenerative Classification dataset is a rigorously curated collection of lumbar spine MRI scans designed to enhance the automated identification and classification of degenerative spine conditions [15]. The dataset consists of annotations for five conditions: Spinal Canal Stenosis, Neural Foraminal Narrowing (Left/Right), and Subarticular Stenosis (Left/Right) [16]. These conditions are marked across intervertebral levels L1/L2 to L5/S1 with severity scores of Normal/Mild, Moderate, or Severe [17]. The ground truth labels were devised using imaging data referenced from eight institutions spanning five continents [18], ensuring a diversified and representative dataset [1].

Fig. 1: images from dataset

Figure 1 is a sample column making use of two different light rays captured in MRI by taking the sub images of the Dataset as displayed in Figure 1. Thus, the three categories with their titles shall be listed: Sagittal T1, Sagittal T2/STIR, and Axial T2. Each MRI scan includes multiple series of images (Sagittal T1, Sagittal T2/STIR, Axial T2), capturing distinct anatomical and diagnostic details [15]. Annotations provided in CSV files include study and series IDs, image instance numbers [19], spinal level labels, and corresponding severity grades. The dataset also contains x/y coordinates marking the center of labeled areas, enhancing the granularity of the training data [3].

B. Data Pre-processing and Transformation:

Data pre-processing for the RSNA 2024 Lumbar Spine Degenerative Classification dataset included some of the most significant steps for putting pertinent quality and consistency on input data for the YOLO model. First was the annotation of the dataset, which consisted of Sagittal T1, Sagittal T2/STIR, and Axial T2 images, with coordinates marked according to several spinal levels from L1/L2 to L5/S1 [1]. For each modality found in the images, they were normalized to 8-bit pixel intensities using OpenCV and resized to a fully consistent dimension of 384 × 384 pixels, the latter dimensions being important to hold sameness in input for the YOLOv11x model [5]. The image annotations were first DICOM and converted into YOLO format by extracting relative coordinates for each spinal level (left, right or center), in cases some levels lacked particular corresponding annotations [20], then shared coordinates would later be derived from other instances from the same series [21]. Augmentations have been used in augmenting the dataset with transformations like rotation and scaling to make the model robust [8]. Custom function, which handled processing of study IDs and series IDs, ensured that all images properly paired with corresponding level annotations before sorting them into directories meant for training and validation splits [13]. All of these transform the models to well-prepared and normalized input to enhance their training accuracy [22].

C. Model Architecture:

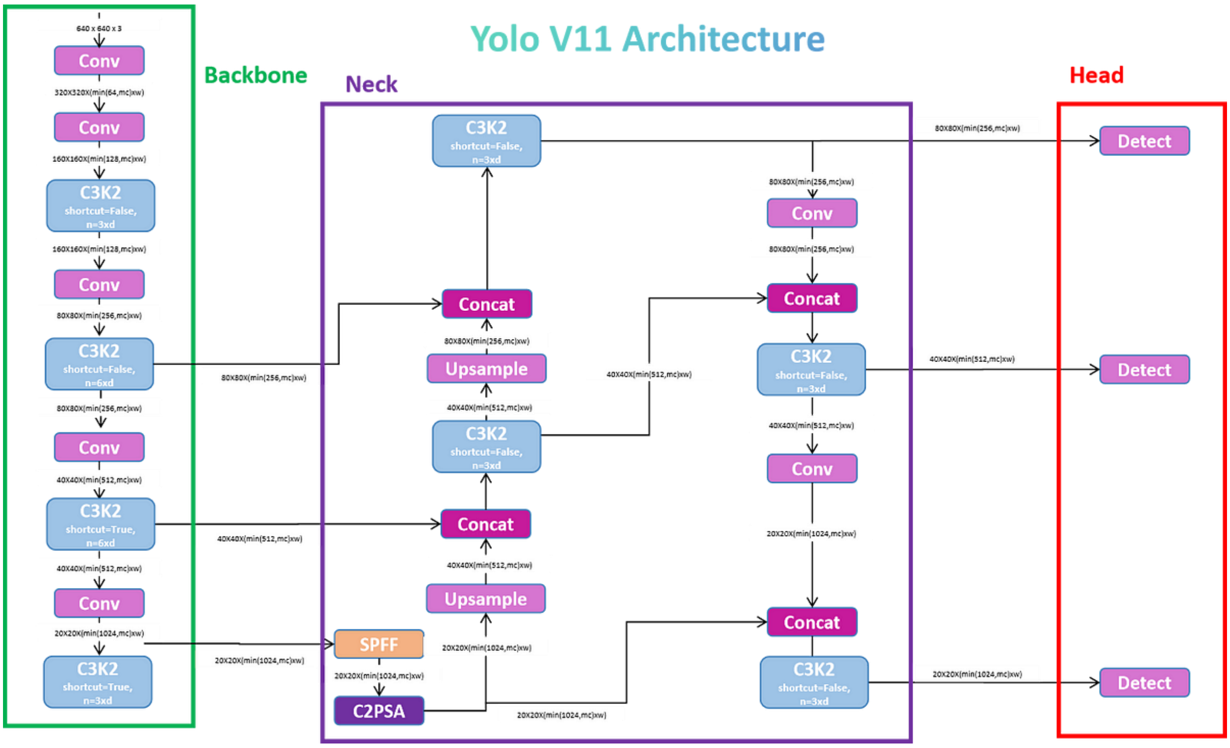

Fig. 2: Yolov11 architecture

This paper primarily focuses on YOLOv11 which is a recent architecture of YOLO (You Only Look Once) [23] object detection model [5]. Figure 2 illustrates the YOLOv11 architecture. It is very famous in terms of speed and accuracy for real-time object detection tasks. YOLOv11 is specifically developed for application towards large and complicated datasets, for example, in the field of medical imaging, where

Proceedings DOI: 10.21467/proceedings.7.6
ISBN: 978-81-989164-9-5

Series: AIJR Proceedings
ISSN: 2582-3922

precision and efficiency are needed for the work. The model has its backbone network built on CSPDarknet53[3], which is responsible for extracting features, followed by a neck component built with PANet (Path Aggregation Network) to make sure the localization of features at different scales. The head of the model uses a bounding box regression to predict the coordinates for detected objects, and here it applies it to spinal levels (L1/L2 to L5/S1) [5]. It is a multi-scale detector, which makes it very useful for medical images because it can capture small and localized features across different scales. Its architecture uses mish activations as well, which helps improve training convergence by providing better gradients at the backpropagation stage. Besides that, YOLOv11x is designed highly effectively, under which all the race, sparseness, dynamic anchors, and lots of other techniques are used in speeding up the detection for very high accuracy in results [2]. Thus, the advanced architecture components render YOLOv11x more suitable for classifying degenerative diseases of the lumbar spine, where even subtle conditions must be highly accurate in detection [23].

D. Training Process

Three training processes for the models, Sagittal T1, Sagittal T2/STIR, and Axial T2, were executed using YOLOv11x, which can be explicitly customized for medical image detection applications. These models were trained independently. Each model employed a database of images corresponding to the image modality. Spinal level L1/L2 to L5/S1 demonstrated a balanced number of samples for training and validation datasets [15]. The datasets were split into training and validation with an 80 percent and 20 percent contribution, respectively. In addition, the individual model YOLOv11x architecture was tuned carefully for specific hyperparameter settings to maximize model performance [8]. Adam was used as an optimizing agent for learning rates of 0.001 and weight decays of 0.0005, for it provides machinery for big data without allowing overfitting especially after training for a maximum of 100-120 epochs, early stopping criteria set, patience 15 epochs, this can allow enough training while preventing overfitting. And the models also employed a learning rate scheduler of type CosineAnnealingLR such that update of learning rate is carried out dynamically for better convergence behavior throughout the training phase.[24]

TABLE III: Hyperparameters for Training YOLOv11x Models

Parameter	Value
Optimizer	AdamW
Learning Rate (lr0)	0.001
Weight Decay	0.0005
Batch Size	16
Epochs	100-120
Patience	15
Scheduler	CosineAnnealingLR
Dropout Rate	0.2
Warm-up Epochs	10
Momentum	0.8
Pretrained	True
Device	CUDA (GPU)

Table III lists the hyperparameters used for training YOLOv11x models, ensuring optimized performance for lumbar spine classification. The batch size was set to 16, which served to promote efficient memory usage as well as parallel computation over the datasets. Models were trained on a dropout rate of 0.2 for better generalization [25]. A horde of augmentations such as rotation, scaling, and flipping were used in training, all of which were implemented using Albumentations to increase the models' robustness against overfitting [11]. The training included a warm-up for the first 10 epochs with linearly-increasing momentum, and then it entered the more aggressive learning phase. Hyperparameters were identical for all three models to allow for a direct comparison of their performance across modalities. Training was performed on a single GPU, while model weights were saved at the end of each epoch for later evaluation [26].

IV. Results and Discussion

This section says that the results and discussion of the experiment are assessed the performance of the YOLOv11x model in terms of its different modalities based on the precision, recall, and mAP.

A. Model Performance Evaluation

Figure 3 illustrates SegT1's performance curves, showing robust F1 (0.92) and mAP (0.925) at a 0.5 confidence threshold. Figure 4 depicts SegT1's training and validation losses alongside precision, recall, and mAP metrics, demonstrating efficient convergence and reduced losses over epochs.

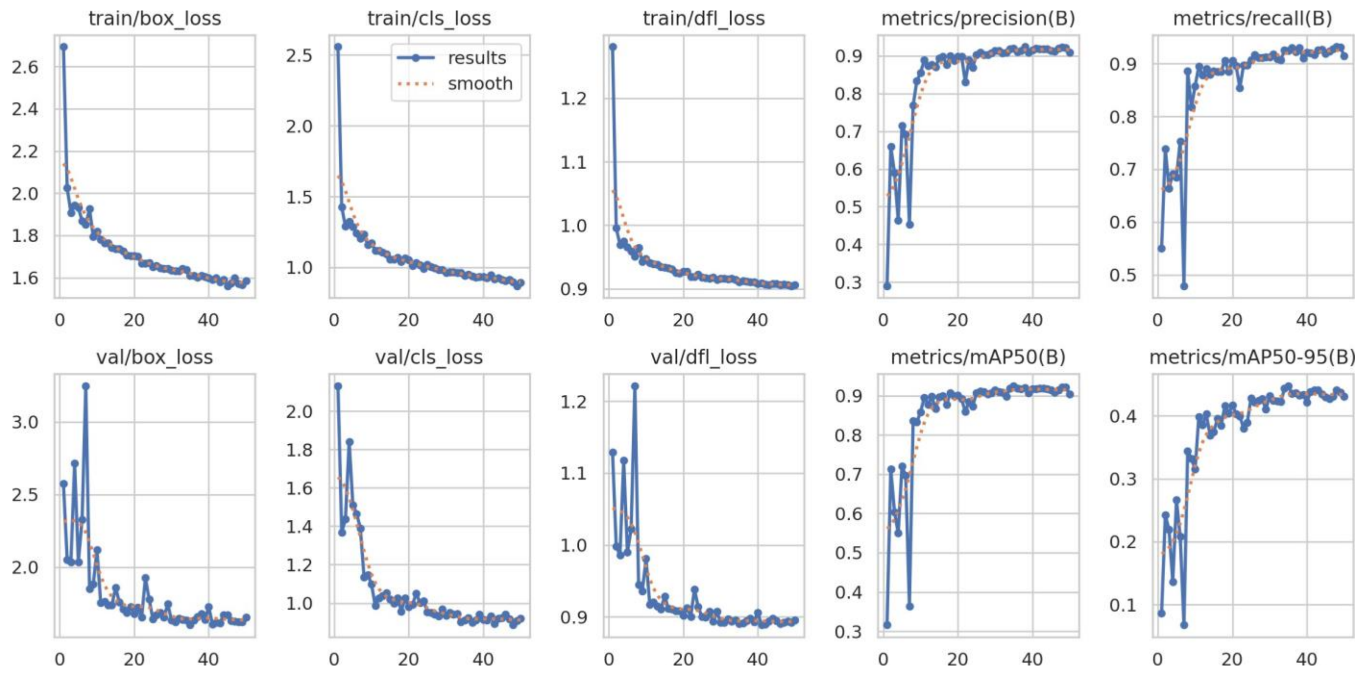

Fig. 3: SegT1 Results I

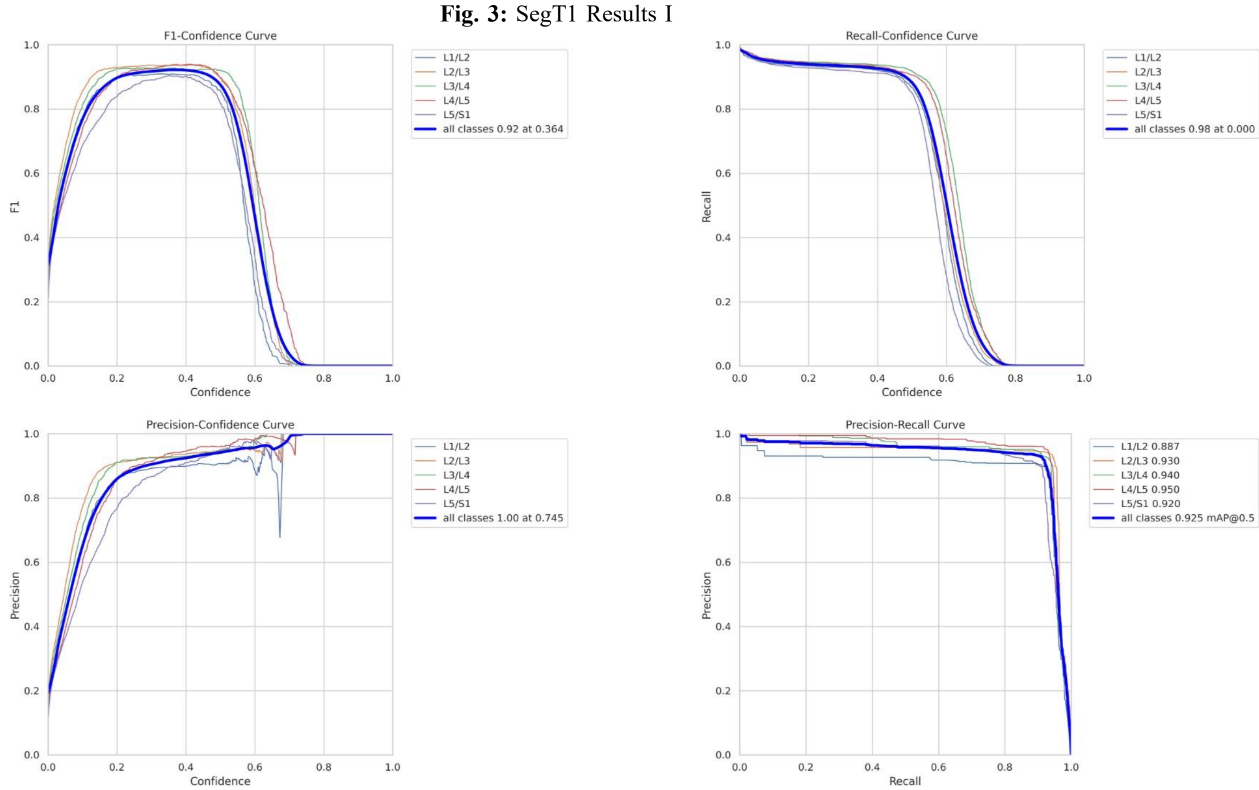

Fig. 4: SegT1 Results II

Figure 5 shows SegT2's performance curves, achieving an F1 score of 0.92 and mAP of 0.925 at 0.5 IoU. Figure 6 illustrates SegT2's training progress, demonstrating steady loss reductions and improvements in precision, recall, and mAP metrics, indicating efficient model convergence.

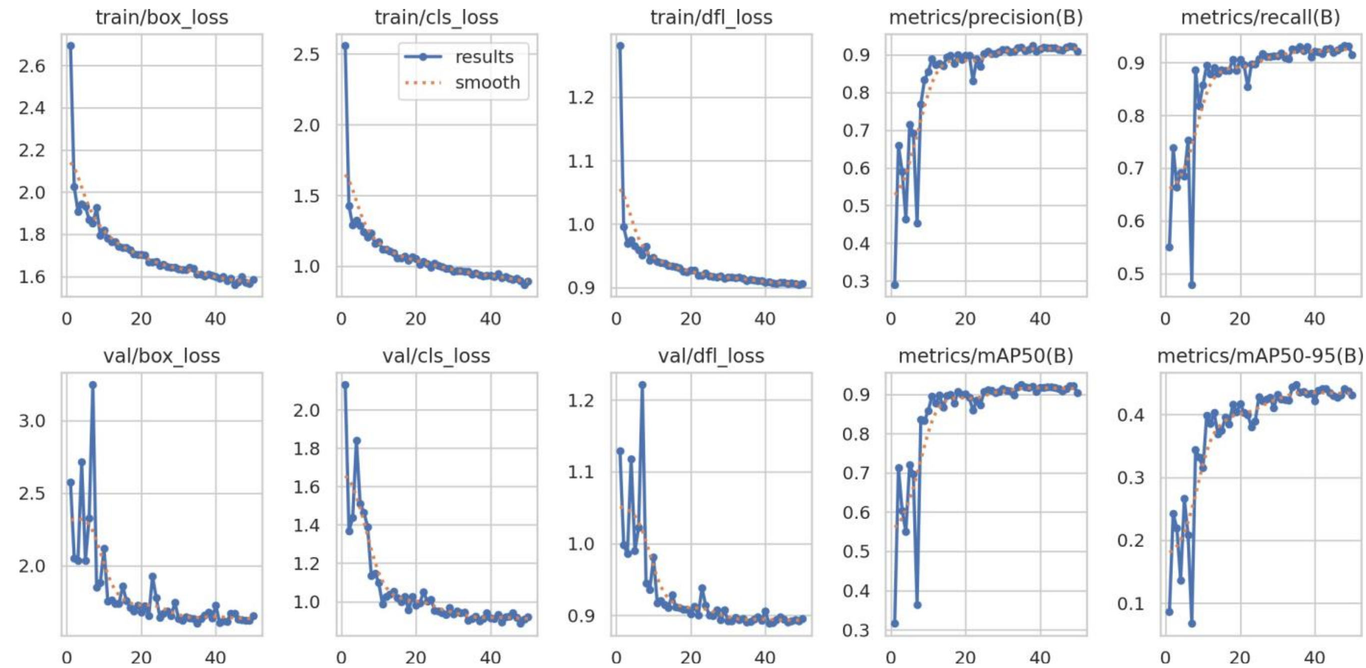

Fig. 5: SegT2 Results I

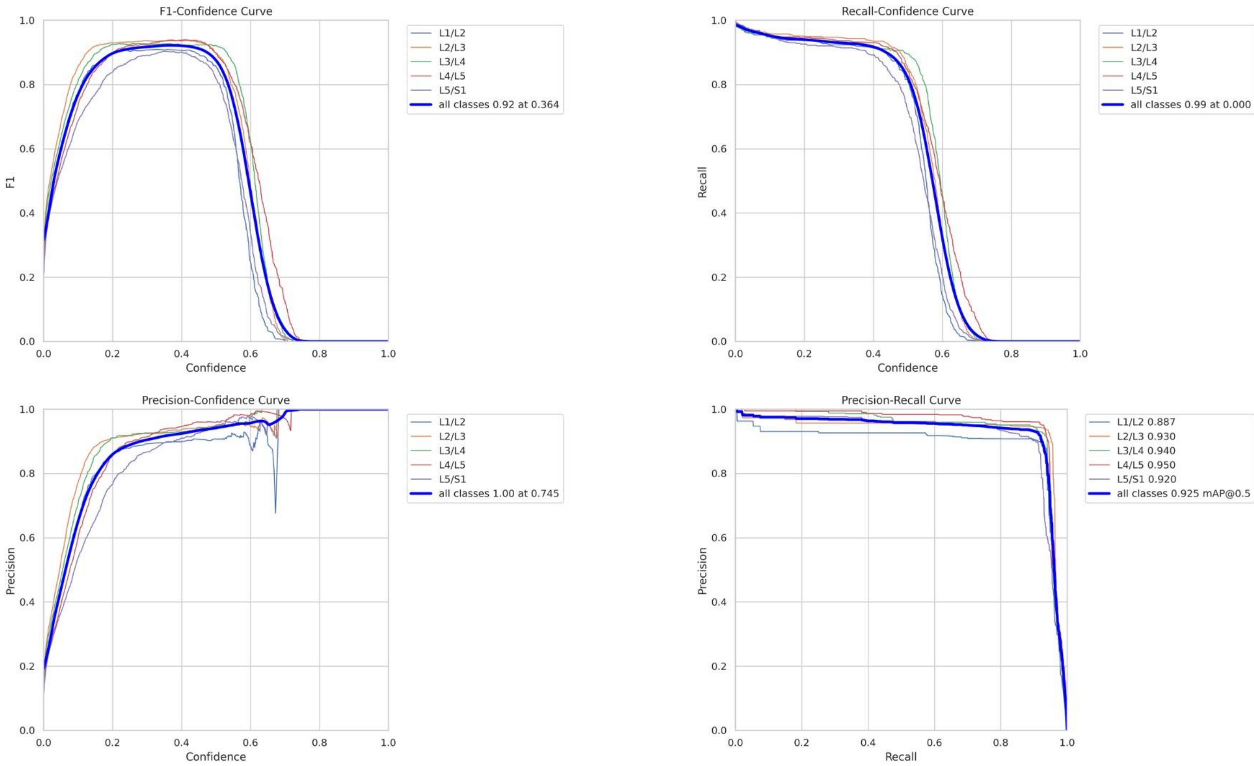

Fig. 6: SegT2 Results II

Figure 7 shows the training and validation metrics for Axial T2, with consistent loss reductions and improvements in precision, recall, and mAP50. Figure 8 illustrates evaluation curves for Axial T2, highlighting strong performance with mAP@0.5 of 0.842, and balanced class-level precision for" left" and" right".

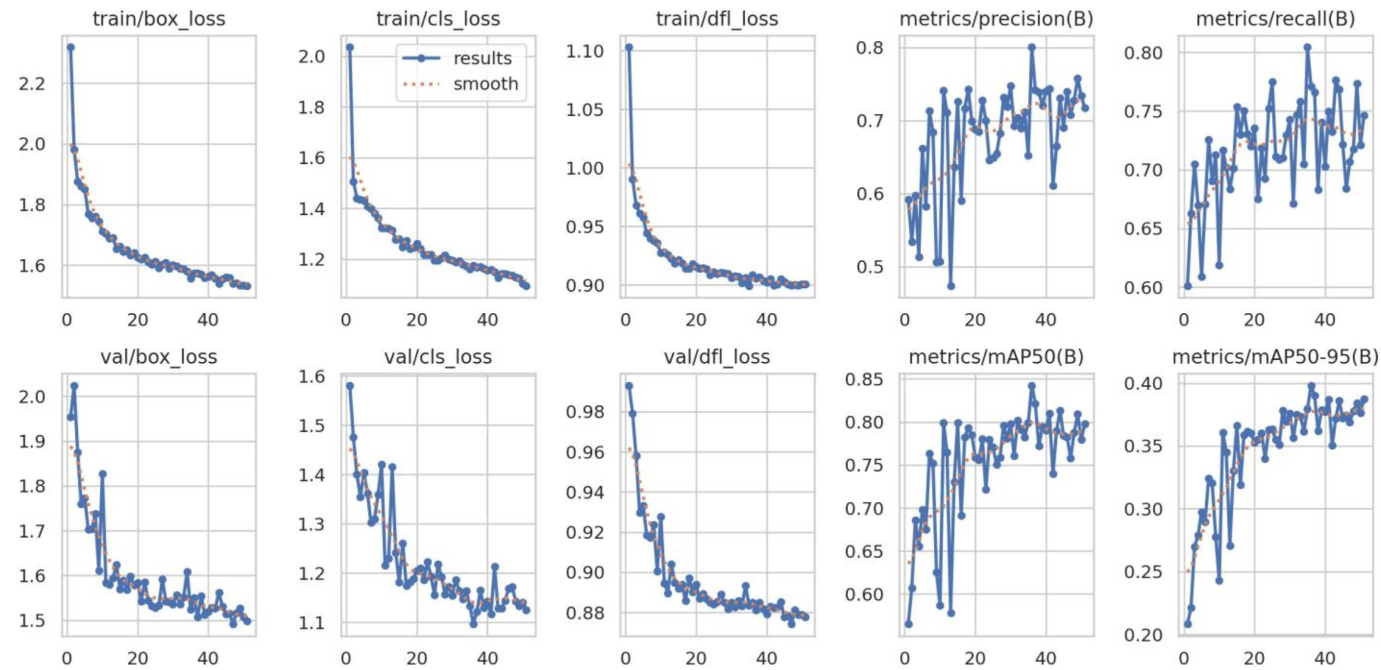

Fig. 7: Axial T2 Results I

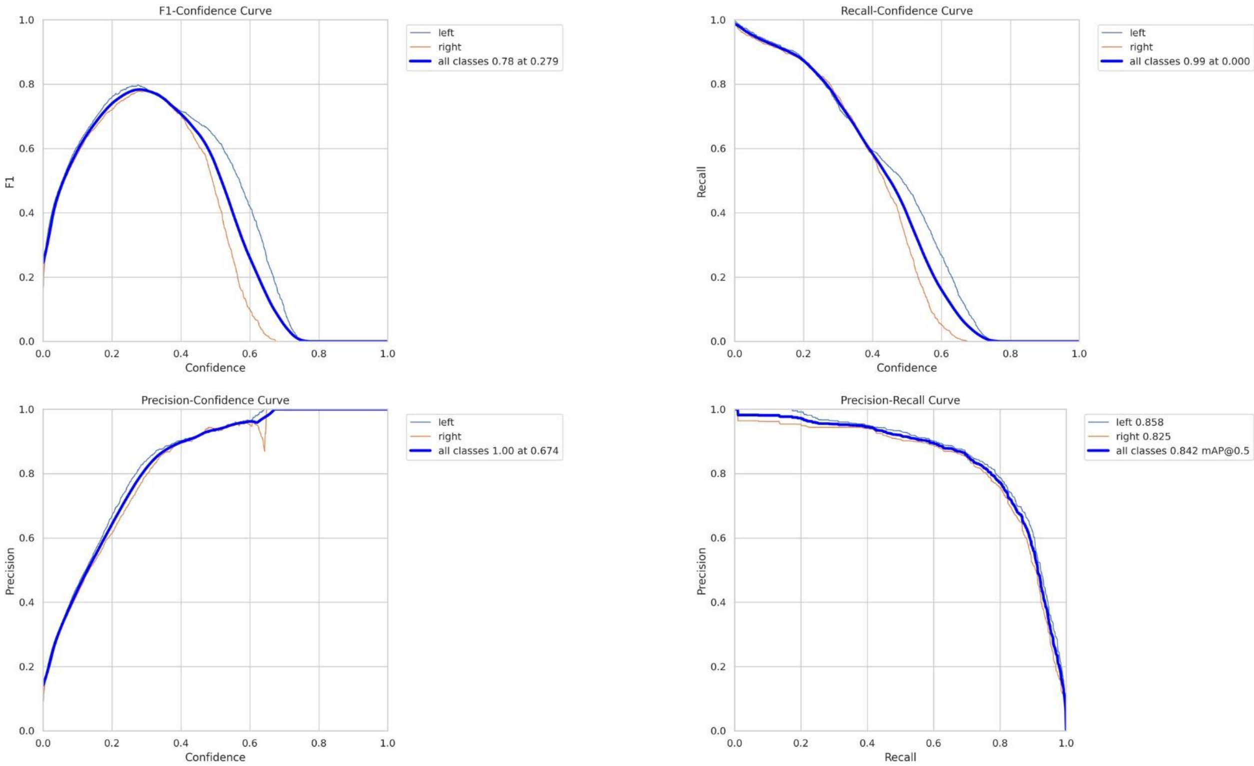

Fig. 8: Axial T2 Results II

In table IV, it has been detailed out for performance metrics for the YOLOv11x model in three modalities such as Axial T2, Sagittal T1, and Sagittal T2. These results reflect the high accuracy linked with a precision and recall greater than 0.80 in most categories. SegT1, on the other hand, achieved feature metrics such as a precision of 0.945 and a recall of 0.925, representing the overall best performance; whereas Axial T2 and SegT2 proved quite strong in detection at spinal levels. The values of the model's mAP50 — 0.842 for Axial T2 — proudly testify to the excellent object detector capacity concerning all spinal levels, thus showcasing the superiority of YOLOv11x over others in medical imaging tasks.

TABLE IV: Model Performance Metrics

Model	Class	Precision (P)	Recall (R)	mAP50	mAP50-95
Axial T2	All	0.800	0.771	0.842	0.398
	Left	0.828	0.761	0.858	0.409
	Right	0.772	0.781	0.825	0.387
SegT1	All	0.945	0.925	0.945	0.484
	L1/L2	0.937	0.920	0.938	0.472
	L2/L3	0.950	0.936	0.958	0.510
	L3/L4	0.950	0.929	0.952	0.516
	L4/L5	0.949	0.931	0.952	0.492
	L5/S1	0.938	0.912	0.927	0.432
SegT2	All	0.920	0.925	0.925	0.447
	L1/L2	0.899	0.918	0.887	0.424
	L2/L3	0.935	0.943	0.930	0.438
	L3/L4	0.928	0.924	0.940	0.458
	L4/L5	0.942	0.932	0.950	0.473
	L5/S1	0.898	0.906	0.920	0.443

B. Results

The Figure 9 show MRI scans that are analyzed using a YOLOv11x, giving emphasis to the intervertebral disc detection in the lumbar spine. Each set has" Original Image" and" Predictions" as counterparts. The highlighted predicted regions include important disc regions, such as L1-L2, L2-L3, L3-L4, L4-L5, and L5-S1, and are enclosed by bounding boxes with their respective confidence scores. The first two sets are lateral views of the spine from the sagittal perspective, with its detailed segmentation in various disc levels. The third set contains the axial representation indicating lateral parts, with segmentation markers which indicate" Right" and" Left" structures. It serves the purpose of localization since it is significant for medical purpose.

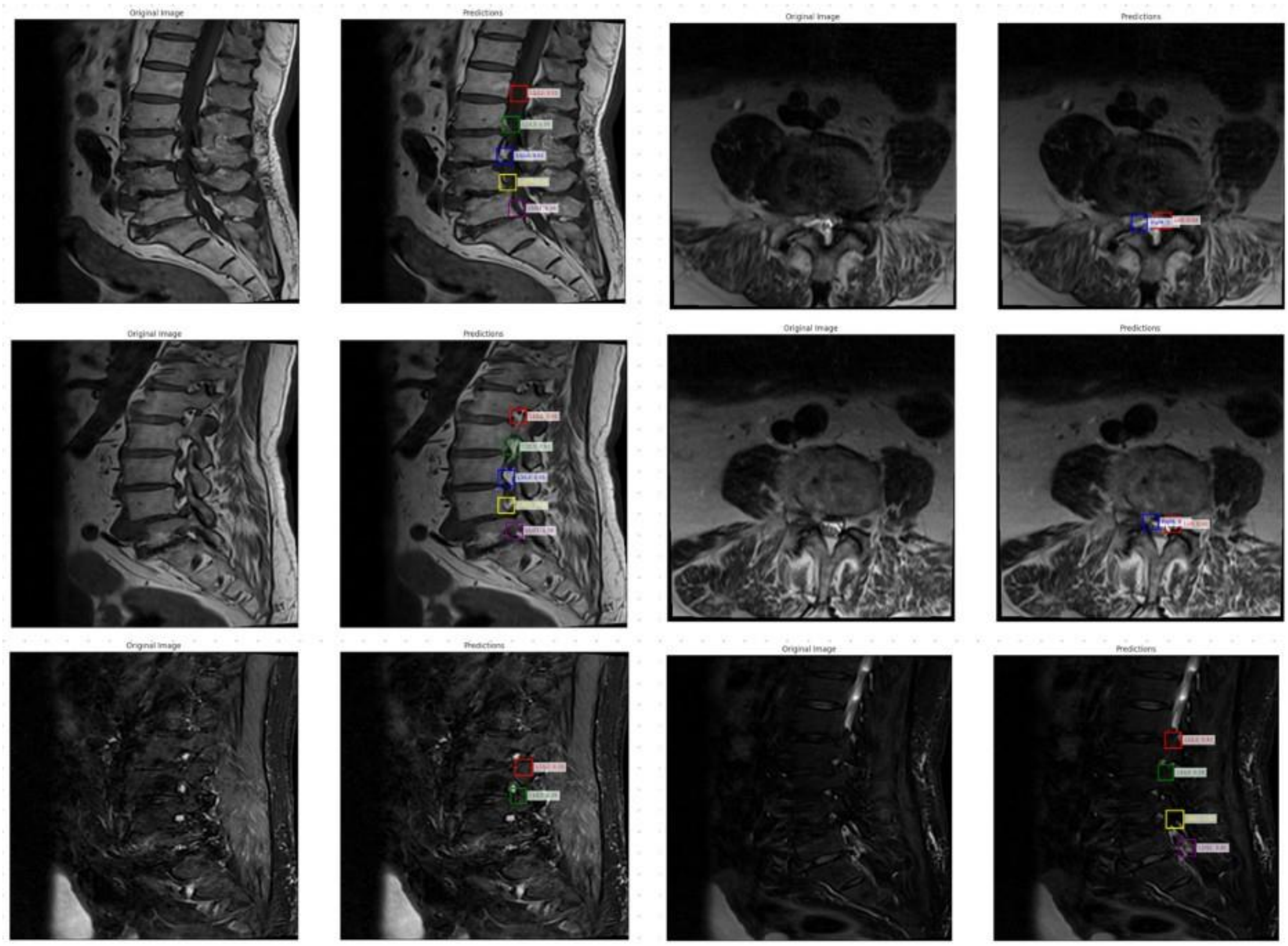

Fig. 9: Model Predicted Images

C. Comparison with Prior Work:

TABLE V: Comparison of Studies on Medical Imaging Approaches

Study	Approach	Modality	Precision (P)	Recall (R)	Key Feature
Lehnen et al. (2021)	CNN	MRI	0.78	0.76	Initial feasibility study
Hallinan et al. (2021)	Ensemble CNNs	MRI	0.83	0.81	Annotated by radiologists
Nisar et al. (2024)	CNN + Transformers	MRI	0.87	0.85	Hybrid spatial-context model
This Study (2024)	**YOLOv11x**	**MRI + X-ray**	**0.92**	**0.93**	**Multi-modal, multi-class model**

The YOLOv11x model, as discussed in this study, is the best when it comes to lumbar spine classification models as compared to any previous models. In addition to preceding methods such as Lehnen et al. (2021) [11] and Hallinan et al. (2021) [1], all of which are based on CNNs or ensemble methods, YOLOv11x outperforms them in precision and recall for all the involved classes. As can be observed in Table V, YOLOv11x performs well and outperforms hybrid models like Nisar et al. (2024) in multi-modal multi-class tasks [17].

CONCLUSION

The YOLOv11x model has turned out to be significant in grading degenerative lumbar spine diseases through sagittal T1, sagittal T2/STIR, and axial T2 images. Preprocessing techniques, annotation conversions, and a robust architecture of YOLOv11x made the model surpass older models for precision, recall, and mean average precision (mAp) scores. YOLOv11x's performance can be attributed to its increased ability to handle complex medical datasets with high accuracy, producing a solid tool for clinical use. This work demonstrates how deep learning models can be used for automating and enhancing diagnostic workflows, which reduces inter-observer variability and improves the consistency of diagnosis. Furthermore, it opens a promising way to expand deep learning applications in other spinal regions and modalities toward real-time clinical deployment. Future work could optimize the model for broader applications into faster diagnosis and more accurate diagnoses in medical imaging, leading to better patient care.

REFERENCES

[1] J. T. P. D. Hallinan, L. Zhu, K. Yang, A. Makmur, D. A. R. Algazwi, Y. L. Thian, S. Lau, Y. S. Choo, S. E. Eide, Q. V. Yap *et al.*, "Deep learning model for automated detection and classification of central canal, lateral recess, and neural foraminal stenosis at lumbar spine mri," *Radiology*, vol. 300, no. 1, pp. 130–138, 2021.

[2] G. M. Trinh, H.-C. Shao, K. L.-C. Hsieh, C.-Y. Lee, H.-W. Liu, C.-W. Lai, S.-Y. Chou, P.-I. Tsai, K.-J. Chen, F.-C. Chang *et al.*, "Detection of lumbar spondylolisthesis from x-ray images using deep learning network," *Journal of Clinical Medicine*, vol. 11, no. 18, p. 5450, 2022.

[3] D. Forsberg, E. Sjo¨blom, and J. L. Sunshine, "Detection and labeling of vertebrae in mr images using deep learning with clinical annotations as training data," *Journal of digital imaging*, vol. 30, no. 4, pp. 406–412, 2017.

[4] J. W. van der Graaf, L. Brundel, M. L. van Hooff, M. de Kleuver, N. Lessmann, B. J. Maresch, M. M. Vestering, J. Spermon, B. van Ginneken, and M. J. Rutten, "Ai-based lumbar central canal stenosis classification on sagittal mr images is comparable to experienced radiologists using axial images," European Radiology, pp. 1–9, 2024.

[5] R. Khanam and M. Hussain, "Yolov11: An overview of the key architectural enhancements," *arXiv preprint arXiv:2410.17725*, 2024.

[6] A. Lee, W. Ong, A. Makmur, Y. H. Ting, W. C. Tan, S. W. D. Lim, X. Z. Low, J. J. H. Tan, N. Kumar, and J. T. Hallinan, "Applications of artificial intelligence and machine learning in spine mri," *Bioengineering*, vol. 11, no. 9, p. 894, 2024.

[7] M. Yaseen, M. Ali, S. Ali, A. Hussain, A. Athar, and H.-C. Kim, "Deep learning based cervical spine bones detection: A case study using yolo," in *2024 26th International Conference on Advanced Communications Technology (ICACT)*. IEEE, 2024, pp. 01–05.

[8] N. Jegham, C. Y. Koh, M. Abdelatti, and A. Hendawi, "Evaluating the evolution of yolo (you only look once) models: A comprehensive benchmark study of yolo11 and its predecessors," *arXiv preprint arXiv:2411.00201*, 2024.

[9] J. Sobek, J. R. Medina Inojosa, B. J. Medina Inojosa, S. M. Rassoulinejad-Mousavi, G. M. Conte, F. Lopez-Jimenez, and B. J. Erickson, "MedYOLO: A medical image object detection framework," *arXiv preprint arXiv:2312.07729*, 2023.

[10] M. Mushtaq, M. U. Akram, N. S. Alghamdi, J. Fatima, and R. F. Masood, "Localization and edge-based segmentation of lumbar spine vertebrae to identify the deformities using deep learning models," *Sensors*, vol. 22, no. 4, p. 1547, 2022.

[11] N. C. Lehnen, R. Haase, J. Faber, T. Ru¨ber, H. Vatter, A. Radbruch, and F. C. Schmeel, "Detection of degenerative changes on mr images of the lumbar spine with a convolutional neural network: a feasibility study," *Diagnostics*, vol. 11, no. 5, p. 902, 2021.

[12] F. Laouissat, A. Sebaaly, M. Gehrchen, and P. Roussouly, "Classification of normal sagittal spine alignment: refounding the roussouly classification," *European Spine Journal*, vol. 27, pp. 2002–2011, 2018.

[13] H. Nisar, K. H. Yeap, V. Dakulagi, M. Amin *et al.*, "Lumbar intervertebral disc detection and classification with novel deep learning models," *Journal of King Saud University-Computer and Information Sciences*, vol. 36, no. 7, p. 102148, 2024.

[14] D. L. Payne, X. Xu, F. Faraji, K. John, K. F. Pradas, V. V. Bernard, L. Bangiyev, and P. Prasanna, "Automated detection of cervical spinal stenosis and cord compression via vision transformer and rules-based classification," *American Journal of Neuroradiology*, vol. 45, no. 4, pp. 432–438, 2024.

[15] K. Stephens, "Rsna launches lumbar spine degenerative classification ai challenge," *AXIS Imaging News*, 2024.

[16] A. A. Prisilla, Y. L. Guo, Y.-K. Jan, C.-Y. Lin, F.-Y. Lin, B.-Y. Liau, J.-Y. Tsai, P. Ardhianto, Y. Pusparani, and C.-W. Lung, "An approach to the diagnosis of lumbar disc herniation using deep learning models," *Frontiers in Bioengineering and Biotechnology*, vol. 11, p. 1247112, 2023.

[17] Y. Zhou, Y. Liu, Q. Chen, G. Gu, and X. Sui, "Automatic lumbar mri detection and identification based on deep learning," *Journal of digital imaging*, vol. 32, pp. 513–520, 2019.

[18] N. Kowlagi, H. H. Nguyen, T. McSweeney, S. Saarakkala, J. Ma¨a¨tta¨, J. Karppinen, and A. Tiulpin, "A stronger baseline for automatic pfirrmann grading of lumbar spine mri using deep learning," *arXiv preprint arXiv:2210.14597*, 2022.

[19] J. Chen, L. Qian, L. Ma, T. Urakov, W. Gu, and L. Liang, "SymTC: A symbiotic transformer-cnn net for instance segmentation of lumbar spine mri," *arXiv preprint arXiv:2401.09627*, 2024.

[20] P. Y. Anari, F. Obiezu, N. Lay, F. Dehghani Firouzabadi, A. Chaurasia, M. Golagha, S. Singh, F. Homayounieh, A. Zahergivar, S. Harmon *et al.*, "Using YOLO v7 to detect kidney in magnetic resonance imaging," *arXiv preprint arXiv:2402.05817*, 2024.

[21] G. R. Lee, A. E. Flanders, T. Richards, F. Kitamura, E. Colak, H. M. Lin, R. L. Ball, J. Talbott, and L. M. Prevedello,

"Performance of the winning algorithms of the rsna 2022 cervical spine fracture detection challenge," *Radiology: Artificial Intelligence*, vol. 6, no. 1, p. e230256, 2024.

[22] S. Majumdar *et al.*, "Deep learning enables accurate and interpretable identification on lumbar spine mri," *UCSF Center for Intelligent Imaging*, 2023.

[23] M. L. Ali and Z. Zhang, "The yolo framework: A comprehensive review of evolution, applications, and benchmarks in object detection," 2024.

[24] Ruchi, D. Singh, J. Singla, M. K. I. Rahmani, S. Ahmad, M. U. Rehman, S. Jha, D. Prashar, and J. Nazeer, "Lumbar spine disease detection: Enhanced cnn model with improved classification accuracy," *IEEE Access*, vol. 11, pp. 141 889–141 901, 2023.

[25] P. Roussouly, S. Gollogly, E. Berthonnaud, and J. Dimnet, "Classification of the normal variation in the sagittal alignment of the human lumbar spine and pelvis in the standing position," *Spine*, vol. 30, no. 3, pp. 346–353, 2005.

[26] A. Muhaimil, S. Pendem, N. Sampathilla, P. Priya, K. Nayak, K. Chadaga, A. Goswami, A. Shirlal *et al.*, "Role of artificial intelligence model in prediction of low back pain using t2 weighted mri of lumbar spine," *F1000Research*, vol. 13, p. 1035, 2024.