

ISSN 1840-4855

e-ISSN 2233-0046

Original scientific article

<http://dx.doi.org/10.70102/afts.2026.1836.674>

MODIFIED ATTENTION U-DENSENET FOR 13-CLASS LUMBAR SPINE MRI SEGMENTATION

Rana Riyadh Saeed^{1*}, Muhammad Amir As'ari², Ahmad Zahran Md Khudzari³,
Ahmed Sabeeh Yousif⁴, Dr. Hadeel Mohamed Farook Alhialy⁵,
Mohamed Muyaser Naif⁶

^{1*}Department of Biomedical Engineering and Health Sciences, Faculty of Electrical Engineering, Universiti Teknologi Malaysia (UTM), Johor, Malaysia; Department of Radiology Techniques, Mosul Medical Technical Institute, Northern Technical University (NTU), Mosul, Iraq. e-mail: ranariyadhsaeed@graduate.utm.my,
orcid: <https://orcid.org/0000-0002-8706-8168>

²Department of Biomedical Engineering and Health Sciences, Faculty of Electrical Engineering, Universiti Teknologi Malaysia (UTM), Johor, Malaysia; Sport Innovation and Technology Centre (SITC), Universiti Teknologi Malaysia, Skudai, Malaysia. e-mail: amir-asari@utm.my, orcid: <https://orcid.org/0000-0003-2122-9536>

³Department of Biomedical Engineering and Health Sciences, Faculty of Electrical Engineering, Universiti Teknologi Malaysia (UTM), Johor, Malaysia; IJN-UTM Cardiovascular Engineering Centre, Institute of Human Centered Engineering, Universiti Teknologi Malaysia, Johor Bahru, Johor, Malaysia. e-mail: zahran@biomedical.utm.my,
orcid: <https://orcid.org/0000-0002-6328-7112>

⁴Department of Information Technology Management, Technical College of Management, Northern Technical University (NTU), Mosul, Iraq. e-mail: ahmedsabeeh123@ntu.edu.iq,
orcid: <https://orcid.org/0009-0006-5740-7030>

⁵Department of Radiology, Mosul Medical College, University of Mosul, Iraq. e-mail: hadeelhiali_radio@uomosul.edu.iq, orcid: <https://orcid.org/0000-0003-2093-5199>

⁶Radiology Unit, Department of Surgery, Nineveh Medical College, University of Nineveh, Iraq. e-mail: mohamed.naif@uoninevah.edu.iq, <https://orcid.org/0000-0003-0134-2013>

Received: May 29, 2026; Revised: July 10, 2026; Accepted: August 20, 2026; Published: September 29, 2026

SUMMARY

Low back pain can be caused by lumbar disc disorders; however, correct segmentation of the structures of the vertebrae and the disc disorders using MRI is crucial for clinical evaluation. In this case, previous deep-learning segmentation approaches have mainly concentrated on the segmentation of anatomical structures and lacked pathology-aware disc segmentation ability. Therefore, a modified attention U-denseNet model for concurrent lumbar anatomical structure segmentation and pathology-aware disc segmentation is proposed in this work using MRI. The approach uses multiscale feature extraction, a channel-spatial attention mechanism, and semantic context gating to increase the performance of the network in discriminating between healthy and abnormal parts of the discs. In this research, a 13-class segmentation task is formulated using 264 sagittal lumbar MRI images. Training, validation and testing were conducted by separating the data based on the subject-wise partitioning strategy to avoid data leakage. Experimentation revealed that the developed model performed with a macro dice of 0.8440, a macro IoU of 0.7985, a boundary F1-score of 0.8693, and a Hausdorff distance of 10.5245 pixels. The performance of the proposed model was better compared to that of the MANet baseline, which yielded a

macro dice of 0.7754 and disc condition segmentation with 0.486 dice. The results obtained reveal that semantic context modelling can improve the pathologic awareness of segmentation and help radiologists to make decisions. The proposed framework provides an effective computer-aided segmentation approach; however, further validation using larger multi-centre datasets and three-dimensional MRI analysis is required.

Key words: *lumbar MRI, medical image segmentation, attention u-DenseNet, disc pathology, deep learning, computer-aided diagnosis.*

INTRODUCTION

The condition of low back pain caused by lumbar disc pathologies is among the most common musculoskeletal ailments globally, where some of the major causes include disc pathologies, such as hernia, bulging and disc height loss as determined by imaging studies [1]. Magnetic Resonance Imaging (MRI) has been adopted as the gold standard modality for the imaging of the lumbar spine owing to its superior ability to visualize soft tissues and offer pathological details of the lumbar spine without exposing the patient to ionizing radiation [2]. In addition to aiding in clinical diagnosis, MRI is also useful in quantitative analysis, including the automatic quantification of structural elements, such as the dural sac area, as well as modelling of the multi-tissue lumbar spine for biomechanics studies [3][4]. However, manual interpretation and annotation of lumbar MRI images require significant expertise and time, creating a need for automated and reliable segmentation methods to assist clinicians.

Deep learning methods for segmentation, especially encoder-decoder models like U-Net, have been shown to perform well in medical image analysis through contextual feature extraction in the encoder and spatial feature recovery in the decoder [5]. Convolutional neural network (CNN)-based approaches have more recently contributed to the improvement of segmentation performance for various medical imaging modalities through better feature extraction and representation learning [6][10]. However, multi-class lumbar spine segmentation has proven to be difficult for several reasons, which include the variation in the size of the anatomical structure, poor contrast of adjacent tissue, narrow disc boundaries, and partial-volume effects. Furthermore, the clinically significant disc diseases occupy only small portions of the MRI image, leading to the overestimation of performance metrics based on overlap for the significant pathological portion.

The above problems can be addressed using recent research that has considered dense connectivity, attention mechanisms, and multi-scale feature extraction approaches. Using dense connectivity will enhance reuse of features and gradient flow because each layer has access to information from all preceding feature maps, and this approach will help in retaining fine anatomical details during feature learning [11]. While foundation models like Segment Anything Model (SAM) and Medical SAM (MedSAM) have exhibited the potential of use for medical image segmentation, specialized models perform better in the tasks of lumbar MRI segmentation, especially in boundary-sensitive evaluation [12][23]. Nonetheless, current lumbar spine segmentation techniques mainly consider the anatomical structures such as vertebrae and discs as individual entities but neglect the intervertebral disc's condition.

The absence of pathology-aware segmentation poses an essential problem, as pathology detection is also essential besides the localization of anatomical structures for clinical decision-making. In addition, down sampling causes loss of information in conventional encoder-decoder architectures, which decreases their capacity for capturing the fine boundary details of small and low-contrast disc areas. Attention-based and multi-scale approaches have been successful at improving feature representation; however, they do not make full use of feature reuse and context integration of network layers [13]. An efficient architecture is needed that would allow the simultaneous segmentation of vertebral structures and disc conditions while improving boundary preservation and pathology detection.

In this study, an attention U-DenseNet model is proposed for 13-class sagittal lumbar spine MRI segmentation [21][22]. While the previous studies on anatomical segmentation have focused solely on the segmentation of the anatomical structures, the proposed algorithm performs anatomical segmentation as well as segmentation of four relevant disc conditions (normal, bulging, herniation, and thinning). The main contributions of this study are summarized as follows:

1. A new 13-class MRI segmentation problem for lumbar spinal regions is proposed, which combines anatomical structures with disc pathology conditions in one unified pixel-level segmentation paradigm with publicly annotated data.
2. An enhanced attention U-DenseNet model is designed through a novel multi-branch and multi-scale attention block that incorporates dense feature extraction, channel-spatial attention, and contextual semantic gating to boost feature learning and preserve boundaries.
3. A complete evaluation approach is discussed based on class-wise segmentation measures, boundary-based evaluation methods, and ablation studies for assessing the role of each building block of the proposed model.

The rest of this paper is structured as follows. Section 2 reviews the existing studies on lumbar spine segmentation and attention-based deep learning techniques. Section 3 introduces the methodology proposed, which covers data set preparation, preprocessing, network design, and training parameters. Section 4 shows the experimental results and comparative analysis. Section 5 provides the discussions, limitations, and clinical significance. Section 6 concludes this paper and presents some future work. Finally, Section 7 conclude this paper.

RELATED WORK

Many current medical image analysis applications rely on convolutional networks, ranging from general reviews of the field to disease classification applications in various imaging modalities [1][2]. Similar deep learning algorithms are employed for solving different problems such as medical image fusion and pathology recognition [3][4]. Variants of U-Net also demonstrate application in various dense prediction tasks, including segmentation of brain tumors from MRI images [5]. Most lumbar and spinal image segmentation tasks involve anatomical structures. Previous works have achieved high image segmentation accuracy of more than 0.93 Dice scores on the spine dataset [6]. Several other studies have looked into segmentation methods using nnU-Net for spine segmentation [7][23]. Studies have looked into hybrid architectures using U-Net and BiSeNet to use the complementary representation of features for spine MRI segmentation, achieving a good level of segmentation accuracy, as well as segmentation techniques based on 2-D U-Net for spine lesion segmentation [8][9][25]. Analogous anatomical pipeline processes have also been discussed for lumbar MRI, which include a combination of ResNet-50 and U-Net, joint vertebrae and discs for disability assessment, a dedicated lumbar segmentation network, and spinal stenosis segmentation models [10][11][12][13]. These studies show that the spine and vertebrae can be segmented with a high degree of accuracy, but they consider the disc as an individual anatomical structure without considering its disease state.

The second path for research involves attention and multi-scale feature fusion in encoder-decoder-based models. There are attention-based U-Net models that have incorporated multi-scale feature fusion along with residual blocks that have been successfully used for lumbar MRI segmentation with excellent Dice results [14]. Further, there are U-Net++ models, which have been modified using modified skip connections and multi-scale feature fusion at the decoder part, which have helped in better segmentation of spinal MRI images [15]. The multi-scale attention-based dual-branch model has been successfully implemented in U-shaped models with multi-scale feature extraction and channel-spatial attention [16]. Another method of increasing the receptive field using fully convolutional models has enhanced multi-scale feature learning, and channel/edge information has been beneficial for the segmentation of spine images [17][18].

Both residual learning and the transformers have been contributing to this trend through the advancement in the feature representation and context modelling for dense prediction. Such approaches are beneficial in enhancing the structural learning but do not fully take advantage of the feature reuse over the entire model depths and have been evaluated in the context of anatomical segmentation only. The foundation models, like SAM and MedSAM, are some recent methods which have been utilized for lumbar spine segmentation. The comparative study has indicated that such models offer the flexibility in zero-shot segmentation but outperform the task-specific supervised networks in lumbar MRI segmentation in terms of overlap and boundary evaluation [19]. Moreover, such models lack disc-condition classification.

The evaluation strategy has an impact on how segmentation is interpreted. The current benchmark for lumbar segmentations uses overlap-based metrics and boundary- and/or distance-based metrics [20]. Large background areas and anatomical structure make it possible to increase overall scores and mask poor performance in small regions related to pathology. Hence, boundary-sensitive evaluation and per-class reporting are necessary to assess the segmentation performance. Previously, disc degeneration classification/grading after segmentation has been considered; however, pixel-level segmentation of multiple disc conditions is not provided. A problem common to all of the previously considered methods is that most methods concentrate on segmentation of anatomical structures and not pathology of discs; moreover, the use of average performance metrics masks issues that arise from the small classes. Table 1 summarizes the previous work based on modality, anatomical structure of interest, the number of classes used, and labelling of disc pathologies. This paper attempts to fill the gap by proposing a segmentation method with 13 classes, including four-disc pathologies.

Table 1. Recent works, showing the modality, anatomical target, number of classes, and whether disc pathology is labelled. The last column is empty for all prior work

Ref.	Model/Study	Modality	Anatomical Target	Classes	Disc Pathology Labelled
[6]	U-Net + 3-D recognition	CT	Vertebrae, Whole Spine	Multi	No
[7]	nnU-Net	CT	Cervical/Full Spine	Multi	No
[8]	U-Net + BiSeNet	MRI	Spine Structures	Multi	No
[9]	2-D U-Net	MRI	Bone Metastases	2	No
[14]	Improved Attention	MRI	Lumbar Discs, Caudal	Multi	No
[15]	U-Net/U-Net++	MRI	Vertebrae, Discs, Vertebral Bodies	Multi	No
[17]	Improved FCN	Natural images	General Scene	Multi	No
[19]	SAM/MedSAM vs. nnU-Net	MRI	Discs, Vertebral Bodies	2	No
This work	Attention U-DenseNet	MRI	Vertebrae, Canal, Disc Conditions	13	Yes

METHODOLOGY

The proposed approach includes four key components, which are the dataset along with 13-class labeling, preprocessing, modified attention U-DenseNet, and finally the training and testing process. Detailed information on each of the components has been provided so that replication and further implementations can be done easily.

Dataset

Dataset Source

Segmentation was conducted by using the publicly available Lumbar Spine MRI Dataset by Sudirman et al., [1]. In this dataset, 264 sagittal images of lumbar MRIs were selected, with a single representative image chosen from each patient. For each individual, the selected image was taken by the radiologist, where a representative slice of the sagittal view which shows the lumbar vertebrae, intervertebral discs, sacral regions, and spinal canal was chosen for characterization. This was done prior to pixel-level characterization and pathology classification and not based on an automatic anatomical marker or distortion of the slice. Annotations were made on a pixel level for 13 classes (See table 2), namely the background class, seven labels for vertebral and sacral segments (namely L1-L5, S1, and S2), the spinal canal (in this study referred to as the spinal cord), and four intervertebral disc conditions, namely the normal disc condition, bulging, herniation, and thinning of the disc. The labelling of intervertebral disc conditions sets apart the proposed annotation schema from the existing anatomy-only datasets. The entire set of 13-class annotations is deposited on Zenodo (DOI: 10.5281/zenodo.20634127). Although this dataset allows for pathology-aware annotations, there is a limitation in the number of cases available.

Ground-Truth Annotation

The ground truth annotations were created by a multidisciplinary team comprising biomedical imaging researchers and diagnostic radiologists. The process of annotating involved experts with expertise in biomedical engineering, medical image analysis, and radiological assessment. The segmentation of the lumbar structures was done using 3D Slicer, whereas the pathological condition of each intervertebral disc was assigned collaboratively with the radiology experts. Only the mid-sagittal lumbar spine MRI PNG images present in the dataset were used in the annotation phase without the use of any other MRI slice, volume study, or clinical report. The review of all cases was done collaboratively, and the ground truth annotation for each disc was achieved by consensus among the experts. The criteria used for labelling the disc conditions are discussed in the next section of the Annotation Protocol. To ensure that there is agreement in the annotation, the radiologists reviewed the labels assigned to the disc pathology after completing the annotation of all 264 images. This review resulted in the discovery of ten images with wrong pathology labels, which were then labelled accordingly.

Annotation Reliability Analysis

The inter-observer agreement among the involved radiologists was conducted to assess the quantitative consistency of the expert annotations. The reliability of the disc condition classification was determined through Cohen's Kappa coefficient (κ), which is a measure of categorical agreement adjusted for the chance of agreement.

The kappa coefficient can be determined through equation (1):

$$\kappa = \frac{P_o - P_e}{1 - P_e} \quad (1)$$

Where P_o represents the observed agreement between experts and P_e represents the expected agreement by chance. A higher κ value indicates stronger agreement between annotators and confirms the reliability of the generated ground-truth labels.

The reliability of annotation analysis helps to ensure the creation of segmentations and labeling with enough consistency among experts, thus increasing the level of trust in the results of model behavior that are associated with its learning ability rather than the inconsistencies in the manual process of annotations.

Annotation Protocol

Each intervertebral disc was given a single unique pathology category label out of the four categories defined as normal, bulging, herniation, and thinning. No multiple pathology category labels were allowed for each intervertebral disc. If an intervertebral disc exhibited features of two different pathologies (thinning and herniation), then the final category label was decided based on the dominant pathology found using a radiologist's expertise. After the category label decision, the whole segmented disc was labelled with the pathology category instead of just the abnormal part of the disc. The borderline cases were discussed with medical experts. In the case of suspicion of a single disc, the case was categorised based on the radiologist's expertise and knowledge. In the case of multiple discs' suspicion, the sample was ignored for categorisation.

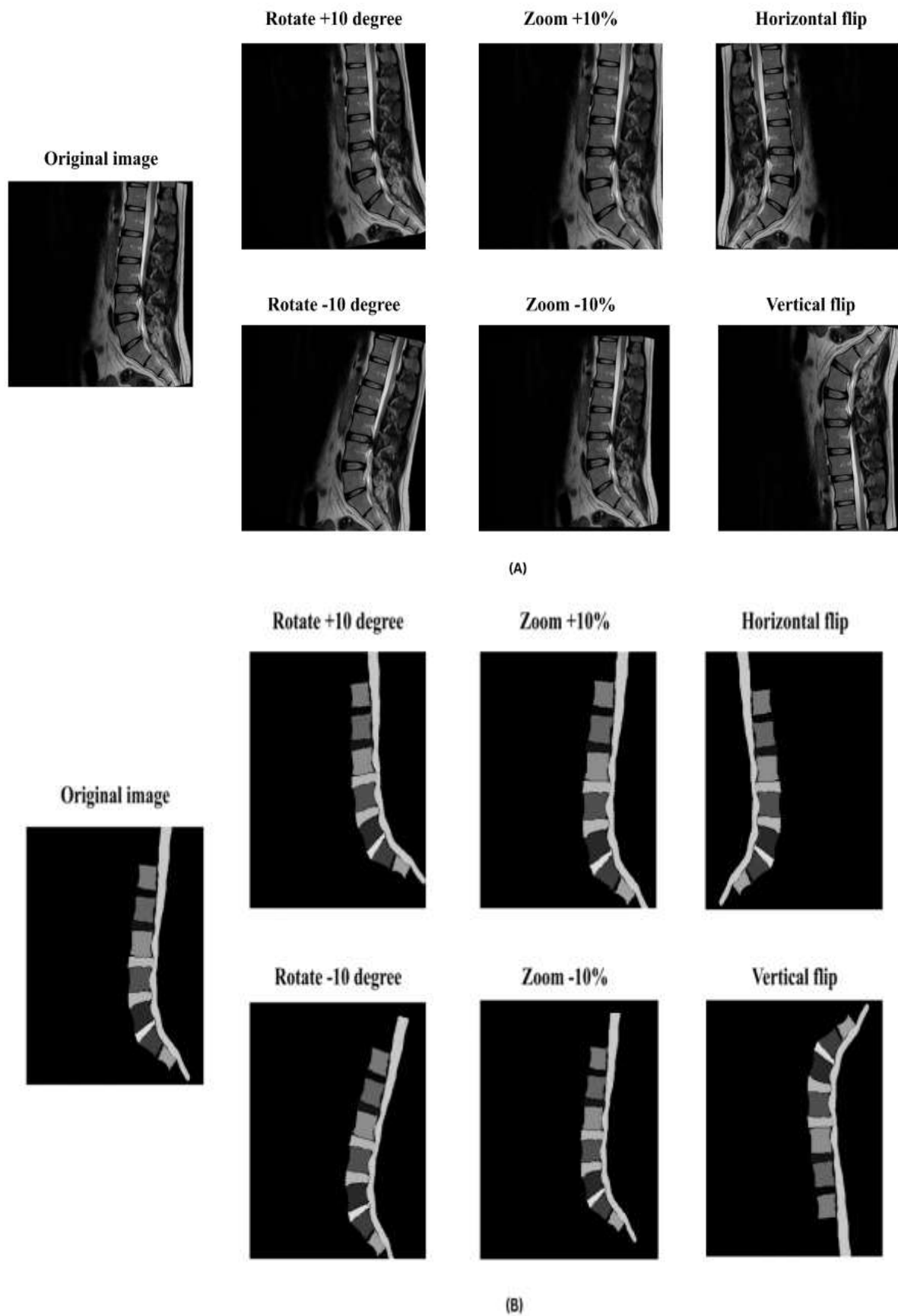


Figure 1. Data augmentation applied to lumbar spine MRI images. (A): Augmented original image. (B): augmented mask

Table 2. Class labels and their corresponding meanings in the dataset

Class ID	Label	Meaning
0	B _G	Background (non-anatomical).
1	C _N	Intervertebral disc region: normal.
2	C _B	Intervertebral disc region: bulging.
3	L5	L5 vertebra.
4	S1	S1 segment.
5	L4	L4 vertebra.
6	L2	L2 vertebra.
7	L1	L1 vertebra.
8	L3	L3 vertebra.
9	S2	S2 segment.
10	C _H	Intervertebral disc region: herniation.
11	S _C	Spinal canal neural structure (labelled as spinal cord in this work).
12	C _T	Intervertebral disc region: thinning (disc height loss / degeneration proxy).

Preprocessing

Data Preparation

Each image in the dataset is related to one mid-sagittal MRI slice of the lumbar spine belonging to one particular patient [1]. The dataset includes 264 MRI images with the original size of 384×384 pixels, having the aspect ratio of 1:1. The pixel spacing in the dataset varied from 0.6771 to 0.8750 mm/pixel, with 0.7292 mm/pixel being the most frequent. During preprocessing, all the images and masks were uniformly downsampled to the size of 256×256 pixels to fit into the input layer of the network. Neither cropping nor padding was performed, thus preserving the initial aspect ratio due to equal scaling of both dimensions. The subjects in the dataset were allocated in such a way that there are 211 training subjects, 26 validation subjects, and 27 testing subjects, making sure that a subject does not belong to more than one subset and avoiding any kind of data leakage. The model was tuned and selected based on the performance achieved by the validation dataset alone. While the test set was completely unseen at the time of evaluating the model performance. But the small sample size of 27 test subjects may not yield statistically significant performance measures; therefore, the future work should include bigger and multi-centered datasets.

Data Augmentation

Augmentation was performed exclusively on the training data set after patient-wise splitting to ensure that there were no augmented images in the validation and test data sets.

Each training image was augmented through random rotation ($\pm 10^\circ$), zoom ($\pm 10\%$), and horizontal flipping, resulting in six new images from each training image and yielding seven samples in total (Table 3).

The augmentation was designed to provide slight changes in patient position and lumbar MRI acquisition in the clinical scenario without affecting anatomical consistency.

The larger rotation or zoom values might result in unrealistic changes in the shape and location of the vertebrae within the image. Thus, the moderate augmentation values were used for diversity training purposes. Figure 1. (A) and (B) illustrates an example of a sample image and its segmentation mask after augmentation.

Table 3. Augmentation techniques used in image preprocessing

Augmentation Type	Parameters	Image Count
Rotation	$\pm 10^\circ$	2
Zoom	$\pm 10\%$ (1.1, 0.9)	2
Flip	Horizontal Flip, Vertical Flip	2

3.2.3 Image Resizing

All images were resized to 256×256 pixels to match the network input requirement. The resizing operation maps the original image dimensions (H, W) into the required network input size equation (2):

$$I_{\text{resized}} = R(I_{\text{original}}, 256, 256) \quad (2)$$

Where $R(\cdot)$ represents the interpolation operator. The MRI images were resized using bilinear interpolation, whereas the corresponding segmentation masks were resized using nearest-neighbour interpolation to preserve discrete class labels.

3.2.4 Normalization

Pixel intensities were normalized into the range $[0,1]$ to improve optimization stability and accelerate convergence. The normalization operation is defined as equation (3):

$$I_{\text{norm}} = \frac{I - I_{\min}}{I_{\max} - I_{\min}} \quad (3)$$

Where $I_{\min}$ and $I_{\max}$ represent the minimum and maximum pixel intensity values, respectively. In the case of MRI 8-bit images employed in this study, the values translate to 0 and 255 respectively. Normalizing minimizes intensity variance and gives a uniform distribution of inputs to train the new Modified Attention U-DenseNet.

Proposed Architecture

Overall Network Architecture

Figure 2 depicts the proposed model architecture, which uses an encoder-decoder (U-shaped) architecture but uses the proposed multi-branch multi-scale attention block instead of a conventional convolutional block. This architecture describes the complete architecture of the Modified Attention U-DenseNet framework, which includes the input MRI image, encoder path, bottleneck representation, decoder reconstruction path, skip connections, feature refinement using attention, and the output 13-class segmentation. The proposed network takes as input a preprocessed sagittal lumbar MRI image of size 256×256 pixels. In the proposed network, the encoder path is composed of four stages to extract hierarchical spatial and semantic features, where the proposed attention mechanism is applied before spatial down sampling in each stage. The multi-branch multi-scale attention block refines the feature representation by using multi-scale dense feature extraction, axial attention, enhanced channel-spatial attention, and class-aware semantic gating. The bottleneck layer extracts high-level contextual information, and the decoder path restores the spatial resolution via up sampling and skip connections. Finally, the output of the model produces probability maps for each of the 13 lumbar spine segmentation classes through a 1×1 convolution and softmax activation layer.

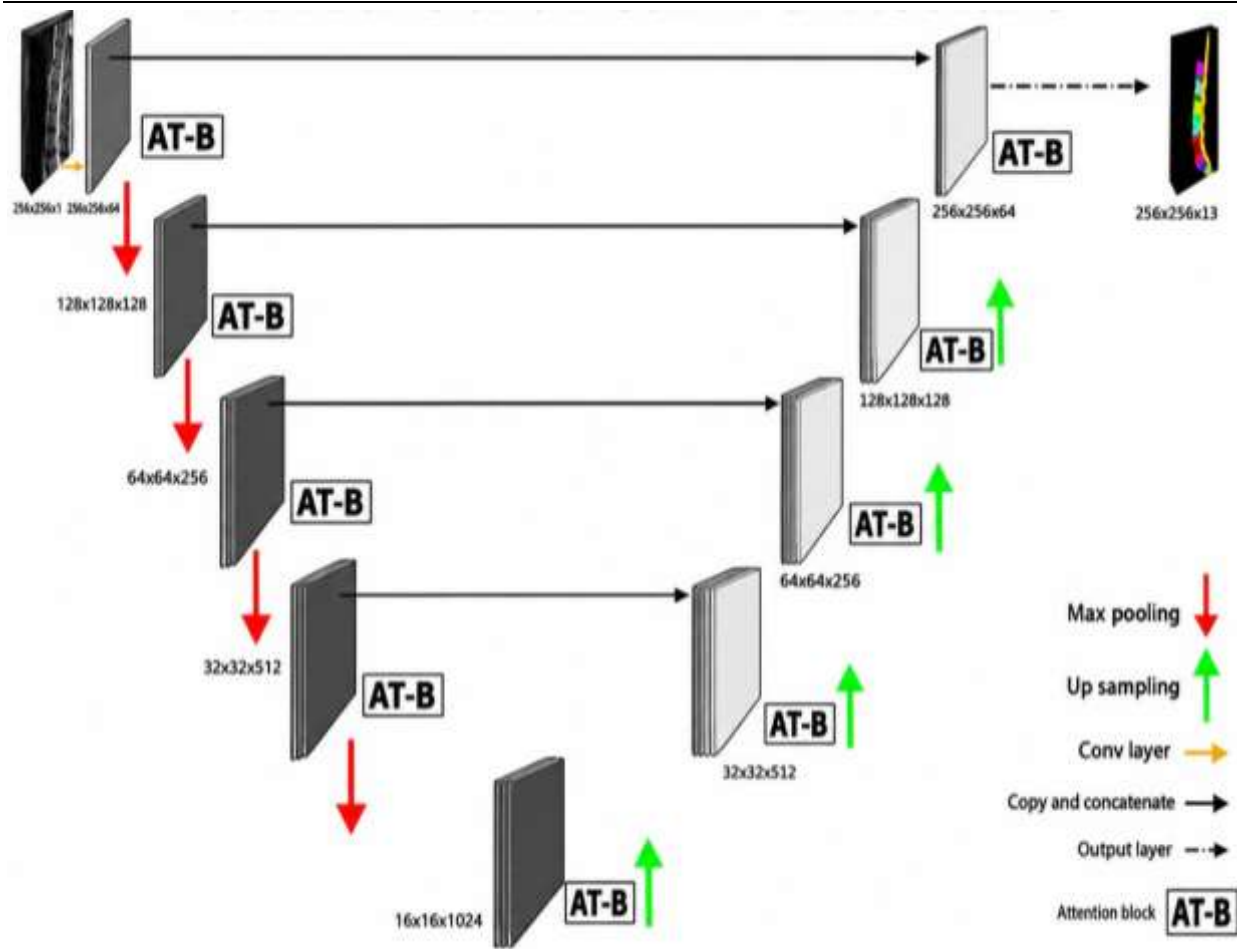


Figure 2. Modified attention U-densenet architecture

The architecture includes four encoder blocks, one bottleneck block, and four symmetrically arranged decoder blocks. The number of feature channels is successively increased from 64, 128, 256, and 512 in the encoder path up to 1024 in the bottleneck layer and decreased again symmetrically from 512, 256, 128, and 64 in the decoder path. Such hierarchy allows for efficient extraction of features necessary for precise segmentation of both low-level boundaries and high-level semantic information of lumbar structures and disc conditions. Skip connections between corresponding encoder and decoder blocks provide for preservation of anatomic details during the process of reconstruction. All convolutional layers use a stride equal to 1 and the same padding. Spatial downscaling is provided only by max pooling. The reconstruction in the decoder part is performed using UpSampling2D(size=(2,2)), nearest-neighbor interpolation. Attention blocks can be applied independently at encoder, bottleneck, and decoder blocks without parameter sharing between various levels. The multi-scale multi-branch attention mechanism is described in detail in Section 3.3. 2. The basic elements of the modified attention U-DenseNet architecture are shown in figure 3 and figure 4.

Encoder Path

The encoder generates the hierarchical feature representations while continuously decreasing the spatial resolution. At every stage, the feature maps are improved via the introduced attention block before being down-sampled, thereby maintaining significant anatomical and boundary information.

The equation for the first encoder layer is given by equation (4):

$$x_0 = \text{ReLU}(\text{BN}(\text{Conv}(X))) \quad (4)$$

Where X represents the input MRI image, $Conv(.)$ denotes convolution operation, $BN(.)$ represents batch normalization, and x_0 is the feature map generated from the first encoder stage.

The feature transformation between encoder stages is formulated as equation (5):

$$x_{i+1} = \text{MaxPool}(A(x_i)) \quad (5)$$

Where x_i represents the input feature map at the i^{th} encoder stage, $A(.)$ denotes the proposed attention block, and x_{i+1} represents the feature map transferred to the next encoder stage.

Bottleneck

This bottleneck connects the encoding path and decoding path by learning high-level semantics in the feature compression step, while retaining context using the attention refinement equation (6).

$$x_b = A(x_n) \quad (6)$$

Where n represents the number of encoder stages, x_n denotes the final encoder feature map, $A(.)$ represents the attention block, and x_b is the bottleneck feature representation.

Decoder Path

Spatial resolution is restored and the segmentation boundaries are fine-tuned through upsampling and skip connections. Decoder features that have been upsampled are combined with encoder features in order to restore the spatial information (Equation 7).

$$x_{dec} = A(\text{Concat}(U, p(x_{dec})x_{enc})) \quad (7)$$

Where $Up(.)$ represents the 2×2 upsampling operation, x_{enc} denotes the encoder feature map, $\text{Concat}(.)$ indicates channel-wise concatenation, and $A(.)$ represents the attention block.

Output Layer

The final feature representation is converted into pixel-wise segmentation probabilities using a 1×1 convolution followed by softmax activation Equation (8):

$$Y = \text{Softmax}(\text{Conv}(x_{final})) \quad (8)$$

Where Y represents the final segmentation probability map for the 13 lumbar spine classes.

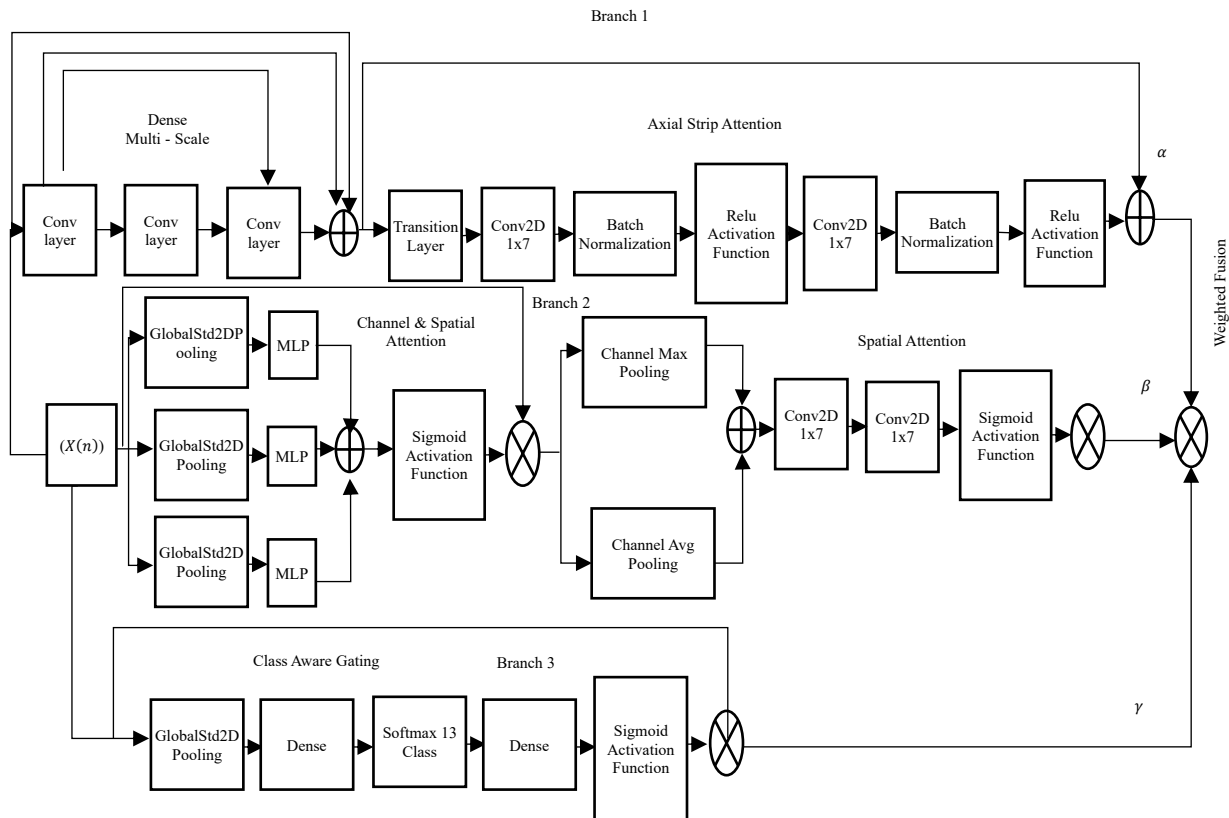


Figure 3. Architecture of the proposed multi-branch multi-scale attention block

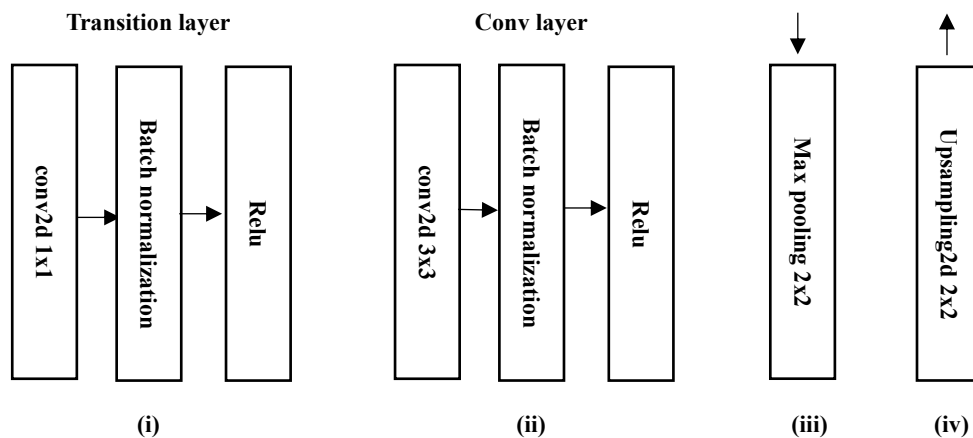


Figure 4. Basic modified attention U-DenseNet MODULE

Training Setup

Training of the suggested model occurred using a consistent optimization setup. ReLU activations were used in all convolutional layers with Glorot Uniform initialization of network weights. The momentum parameter was set to 0.99 for batch normalization layers and $\epsilon=1 \times 10^{-3}$. Dropout or L1/L2 regularizations were not used.

The optimization process was performed using the Adam optimizer with an initial learning rate of 1×10^{-4} . Callbacks ModelCheckpoint, ReduceLROnPlateau, and EarlyStopping were utilized to achieve more stable training. Learning rate reduction by a factor of 0.5 was done after 3 epochs without improvement on the validation dataset, and a minimal learning rate was set to 1×10^{-6} . An EarlyStopping callback with a patience of 5 epochs and restoration of the best model weights was employed.

The training data was shuffled before each epoch, and a batch size of 4 was used owing to computation limitations. Though the maximum training time was limited to 100 epochs, the model converged after epoch 40 according to the stopping criteria. The implementation of the framework has been done with TensorFlow version 2.20.0 and Keras version 3.13.2 run on Google Colab with an NVIDIA T4 GPU (with 16GB memory). The input size used in this experiment was 256×256 pixels with 13 output segmentation classes, four encoder layers, four decoder layers, and 1024 channels at the bottleneck layer. The convolutional layers used 3×3 kernels with the same padding, whereas the decoder reconstruction has been done using the 2×2 UpSampling2D operation. The segmentation map was created by applying 1×1 convolution and Softmax activation on top of that. The random seed value is fixed to 0, and the training time taken is approximately 1 hour and 49 minutes. These settings provide a reproducible and stable training procedure for the proposed Modified Attention U-DenseNet model.

Evaluation Metrics

Performance of the Modified Attention U-DenseNet model has been analyzed based on region-level and boundary-level metrics. The Dice Coefficient and IoU have been calculated per class per test image individually and then the mean has been found out across all the test images. Test images not containing a certain class have been ignored to avoid undefined metric results.

Loss Function

Equation (9) the model was trained on the Sparse Categorical Cross-Entropy loss function which calculates the difference between predicted probability and the truth labels.

$$L = -\frac{1}{N} \sum_{i=1}^N \log(p_{i,y_i}) \quad (9)$$

Where N represents the number of pixels and p_{i,y_i} denotes the predicted probability of the true class.

3.5.2 Accuracy

Accuracy measures the percentage of correctly classified pixels equation (10):

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

Where TP, TN, FP , and FN represent true positives, true negatives, false positives, and false negatives, respectively.

3.5.3 Dice and IoU

The Dice coefficient measures segmentation overlap equation (11):

$$\text{Dice} = \frac{2TP}{2TP + FP + FN} \quad (11)$$

IoU evaluates the intersection between predicted and ground-truth regions equation (12):

$$\text{IoU} = \frac{TP}{TP + FP + FN} \quad (12)$$

3.5.4 Boundary-Based Metrics

The boundary accuracy was determined through the Mean Surface Distance (MSD), Average Surface Distance (ASD), Boundary F1-score (BF1), and Hausdorff Distance (HD). The MSD and ASD measure

average boundary distance while the BF1 determines the boundary precision and recall. The HD computes the maximum boundary difference between the prediction and reference region.

Let ∂P and ∂G represent the sets of boundary points for the predicted region P and the ground-truth region G , respectively. The minimum Euclidean distance from a point $p \in \partial P$ to the boundary ∂G is defined as $d(p, \partial G) = \min_{g \in \partial G} \|p - g\|_2$. The boundary-based metrics are formulated in equations (13–16):

$$\text{MSD} = \frac{1}{|\partial P|} \sum_{p \in \partial P} d(p, \partial G) \quad (13)$$

$$\text{ASD} = \frac{\sum_{p \in \partial P} d(p, \partial G) + \sum_{g \in \partial G} d(g, \partial P)}{|\partial P| + |\partial G|} \quad (14)$$

$$\text{BF1} = \frac{2P_b R_b}{P_b + R_b} \quad (15)$$

$$\text{HD} = \max \left(s, \sup_{p \in \partial P} d(p, \partial G), \sup_{g \in \partial G} d(g, \partial P) \right) \quad (16)$$

Where $|\partial P|$ and $|\partial G|$ denote the total number of boundary pixels in the predicted and ground-truth sets, while P_b and R_b represent boundary precision and boundary recall within a specified distance threshold, respectively.

The combination of measures related to overlapping (Dice coefficient and Intersection over Union) and measures related to boundaries (MSD, ASD, BF1, and HD) will provide a complete assessment of segmentation accuracy. This model can be seen as an intelligent system for decision support in the analysis of the condition of lumbar structures and discs, but not as a replacement for the clinical expertise of radiologists. Final decisions should incorporate radiologist interpretation and patient-specific clinical information.

Algorithm 1: Modified Attention U-DenseNet Segmentation Pipeline

Input: MRI images X , masks Y , learning rate $\eta = 1e-4$, batch size $B = 4$

Output: Trained weights Θ^* , predicted segmentation map Y_{hat}

- 1 Preprocess (X, Y): Resize to 256x256, scale intensities to $[0, 1]$.
- 2 Partition data subject-wise into D_{train} (211), D_{val} (26), D_{test} (27).
- 3 Augment D_{train} via random rotation ($\pm 10^\circ$), zoom ($\pm 10\%$), and flips.
- 4 Initialize Θ using Glorot Uniform; configure Adam optimizer ($\eta = 1e-4$).
- 5 For epoch = 1 to 100 do:
- 6 Shuffle D_{train} and partition into mini-batches (X_b, Y_b).
- 7 Forward Pass: Extract dense features; apply channel-spatial attention
and semantic context gating; compute $Y_{\text{pred}} = \text{Softmax}(1 \times 1 \text{ Conv})$.
- 8 Compute Sparse Categorical Cross-Entropy loss and update Θ .
- 9 Evaluate Val_Loss: Reduce η by 0.5 if stagnant; update Θ^* if improved.

10 If Val_Loss fails to improve for 5 consecutive epochs, Break.

11 Inference: $\hat{Y} = \text{Argmax}(\text{ForwardPass}(X_{\text{test}}, \Theta^*), \text{axis} = -1)$.

12 Return $\hat{Y}$

Algorithm 1 outlines the complete execution flow of the proposed Modified Attention U-DenseNet framework. The pipeline standardizes sagittal lumbar MRI slices and 13-class masks to a resolution of 256×256 pixels with normalized intensities. Data is partitioned subject-wise before data augmentation to prevent leakage. During training, the network learns multi-scale dense features, channel-spatial attention, and semantic context gating, updating parameters using sparse categorical cross-entropy loss alongside adaptive learning rates and early stopping.

RESULTS

The proposed modified attention U-denseNet resulted in a macro-Dice value of 0.844, a macro IoU of 0.799, a macro-HD of 10.52 pixels, and a macro BF1 of 0.869 for 13 segmentation classes in the test set containing 27 patients. As background is known to affect the average performance measures, the class-wise results for disc-condition classes have been taken into consideration for further analysis. Pixel accuracy (0.9991) and validation accuracy (0.9949) have been provided in Table 6 for the sake of completeness, whereas figure 5 (a) and (b) provides stable training and validation curves without any sign of overfitting. The MANet has been used as a baseline and retrained on the proposed 13-class dataset for fair comparison [24]. The proposed system extends the MANet by incorporating an additional branch for the semantic context gating. Table 4 provides three configurations of the MANet baseline, the two-branch attention model, and the proposed multi-branch attention model on the basis of Dice, IoU, HD, and BF1 measures.

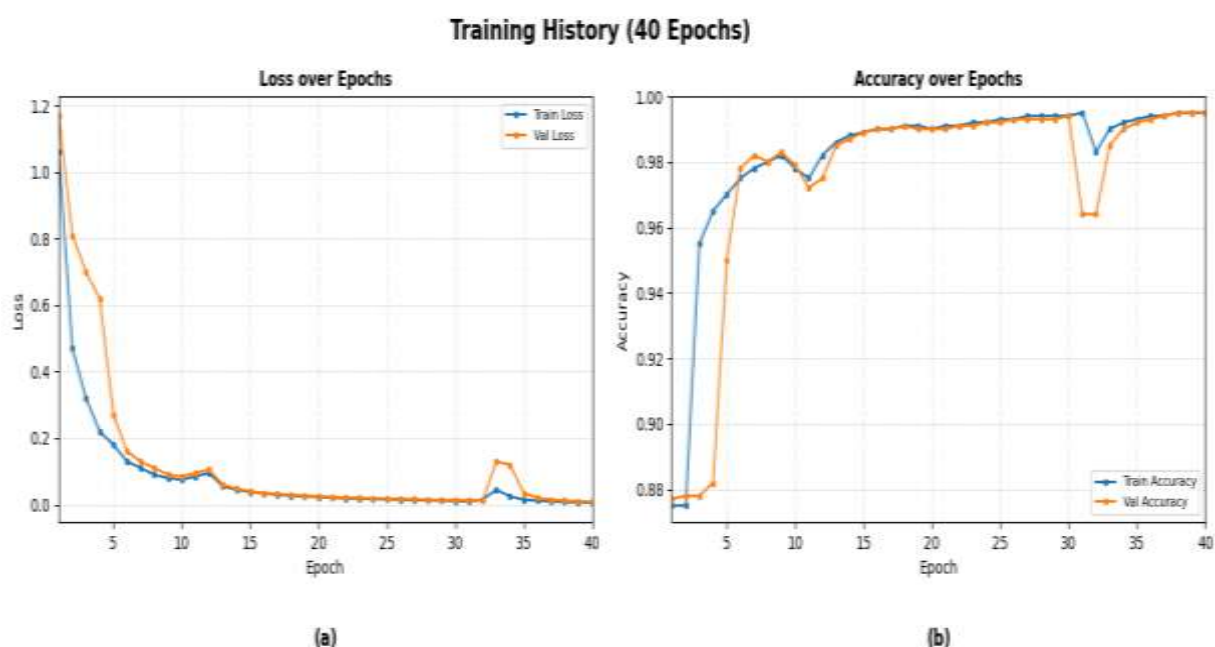


Figure 5. Training and (a) Loss curves and (b) Validation accuracy of the proposed modified attention U-DenseNet model

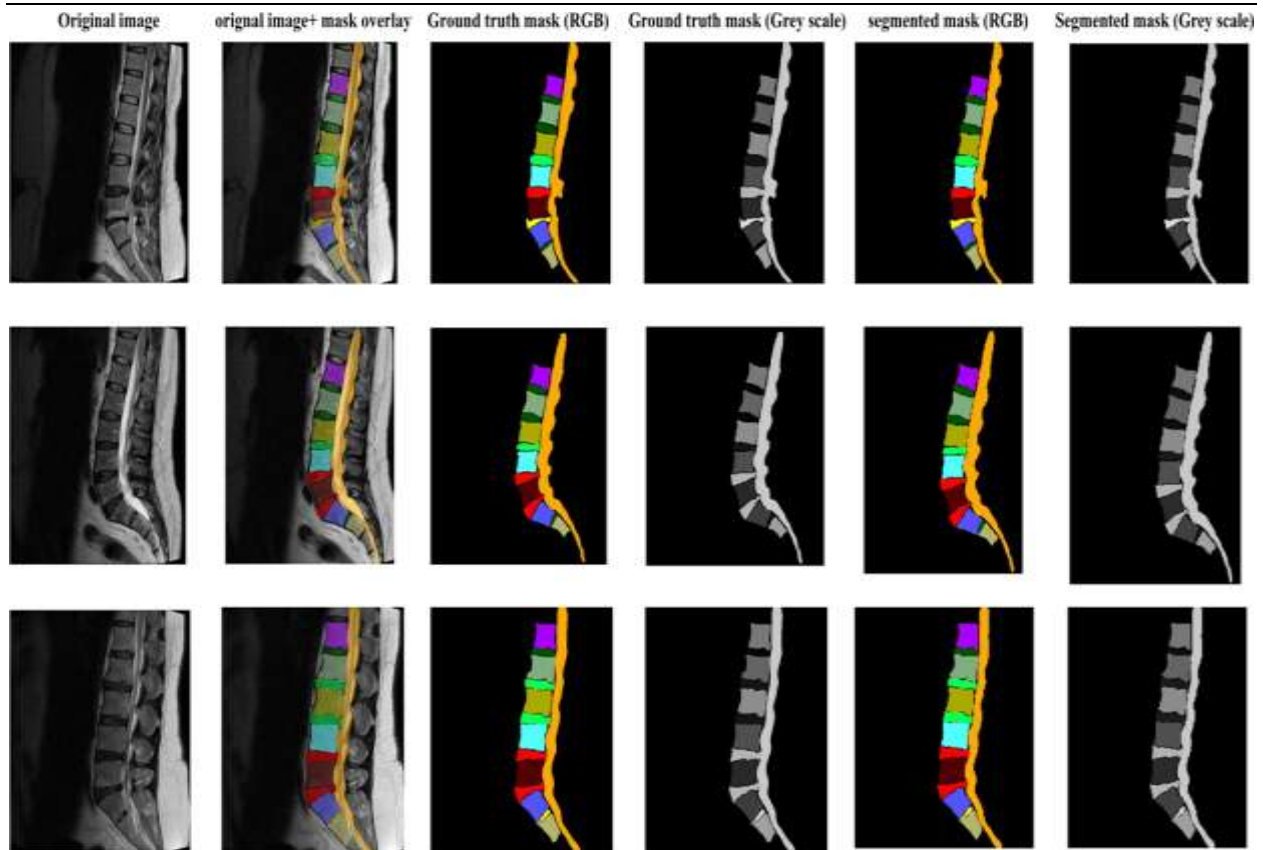


Figure 6. Qualitative segmentation results of the proposed Modified Attention U-DenseNet model on test images

Ablation Analysis and Class-Wise Performance Evaluation

Table 4. Per-class ablation comparison of MANet baseline, two-branch attention model, and proposed modified attention U-densenet using dice, IoU, hausdorff distance (HD), and boundary F1-score (BF1)

Class	Baseline				Two-Branch				Proposed			
	Dice	IoU	HD	BF1	Dice	IoU	HD	BF1	Dice	IoU	HD	BF1
B G	0.9916	0.9833	8.21	0.9875	0.9914	0.9829	7.63	0.9868	0.9971	0.9942	5.95	0.9980
C N (normal)	0.5431	0.3728	79.73	0.6392	0.4623	0.3006	71.52	0.5869	0.7308	0.5758	12.13	0.7781
C B (bulging)	0.4313	0.2749	37.06	0.5061	0.4368	0.2794	42.19	0.4868	0.6591	0.4915	12.05	0.7489
L5	0.8846	0.7931	6.60	0.9330	0.8835	0.7913	8.85	0.9447	0.8994	0.8172	8.35	0.9304
S1	0.9318	0.8723	3.58	0.9885	0.9055	0.8273	3.97	0.9516	0.8994	0.8172	9.28	0.9402
L4	0.8692	0.7687	6.34	0.9093	0.8798	0.7854	7.35	0.9197	0.9495	0.9039	8.35	0.9179
L2	0.8965	0.8124	11.10	0.9490	0.8790	0.7841	15.05	0.9013	0.9494	0.9037	15.71	0.9370
L1	0.9110	0.8365	5.99	0.9700	0.8349	0.7166	12.72	0.8823	0.8876	0.7979	10.12	0.9704
L3	0.8958	0.8113	10.02	0.9266	0.8744	0.7768	8.04	0.9239	0.9382	0.8836	6.98	0.9223
S2	0.8559	0.7481	4.49	0.9606	0.8325	0.7131	4.15	0.9595	0.8879	0.7984	12.62	0.9048
C H (herniation)	0.6442	0.4751	28.18	0.6923	0.7145	0.5558	26.13	0.6906	0.7341	0.5799	13.96	0.7090
S C (canal)	0.8975	0.8141	7.01	0.9769	0.9025	0.8223	7.98	0.9764	0.8878	0.7982	6.49	0.9408
C T (thinning)	0.3273	0.1957	76.51	0.2062	0.2233	0.1257	71.31	0.2098	0.5523	0.3815	14.82	0.6026
Macro	0.7754	0.6332	21.91	0.8189	0.7554	0.6069	22.07	0.8016	0.8440	0.7301	10.52	0.8693

According to table 4, the proposed architecture shows superior results for segmentation over baseline and two-branch versions. The macro-Dice value rises from 0.7754 to 0.8440, and the macro-HD falls from 21.91 pixels to 10.52 pixels. Thus, the boundary localization becomes better in this case. The greatest results are achieved in the disc-condition classes; the Dice values are 0.7308, 0.6591, 0.7341, and 0.5523 for normal, bulging, herniation, and thinning, respectively. This demonstrates that the semantic context gating branch improves discrimination of challenging pathological disc regions.

Group-Level Performance Analysis

As seen in table 5, it is evident that the suggested approach is able to significantly boost the results of disc-condition segmentation. For vertebrae and canal structures, the average Dice value rises from 0.893

to 0.912, while for the disc-condition category, it boosts considerably from 0.486 to 0.669. This confirms that the proposed attention mechanism is effective for small and low-contrast pathological regions.

Table 5. Group-wise mean dice comparison across MANet baseline, two-branch model, and proposed model

Group (Classes)	MANet (Baseline)	Two-Branch	Proposed
Background (1)	0.992	0.991	0.997
Vertebrae + Canal (8)	0.893	0.874	0.912
Disc Conditions (4)	0.486	0.459	0.669
Macro (13)	0.775	0.755	0.844

Auxiliary Metric Analysis

Table 6. Auxiliary performance metrics of the proposed modified attention U-DenseNet model

Metric	Value
Pixel Accuracy (Macro)	0.9991
Validation Accuracy	0.9949

The model under consideration provides an accuracy of 0.9991 at the pixel level and a validation accuracy of 0.9949. These accuracies have been mentioned just for the sake of completeness, as the presence of background class affects the pixel-wise accuracy. Thus, Dice, IoU, HD, and BF1 are the primary metrics table 6.

Segmentation Comparisons for Disc Condition Qualitatively, they are illustrated in figure 7 for the models analysed. The new model achieves better segmentation of disc-condition areas, especially difficult disc boundaries, without compromising the segmentation accuracy of vertebrae and canal regions.

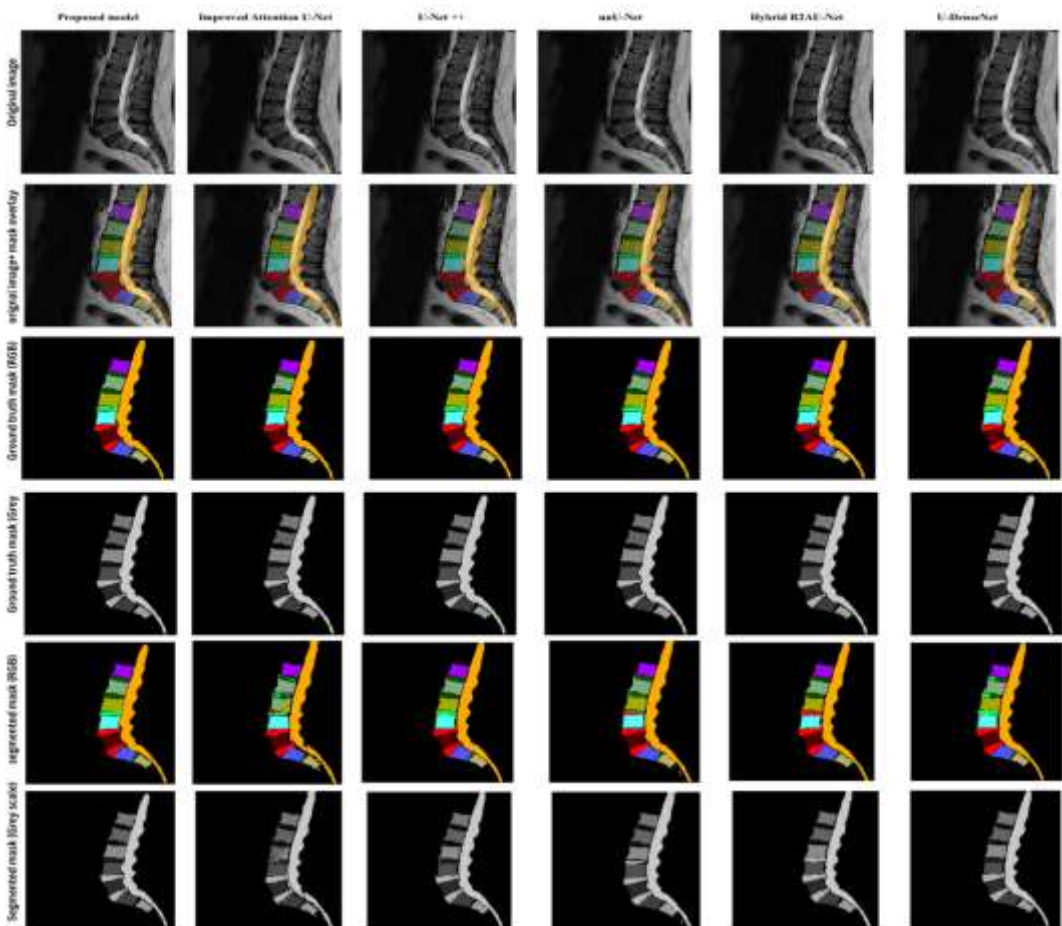


Figure 7. Visual comparison of segmentation results for different models on the 13-class dataset

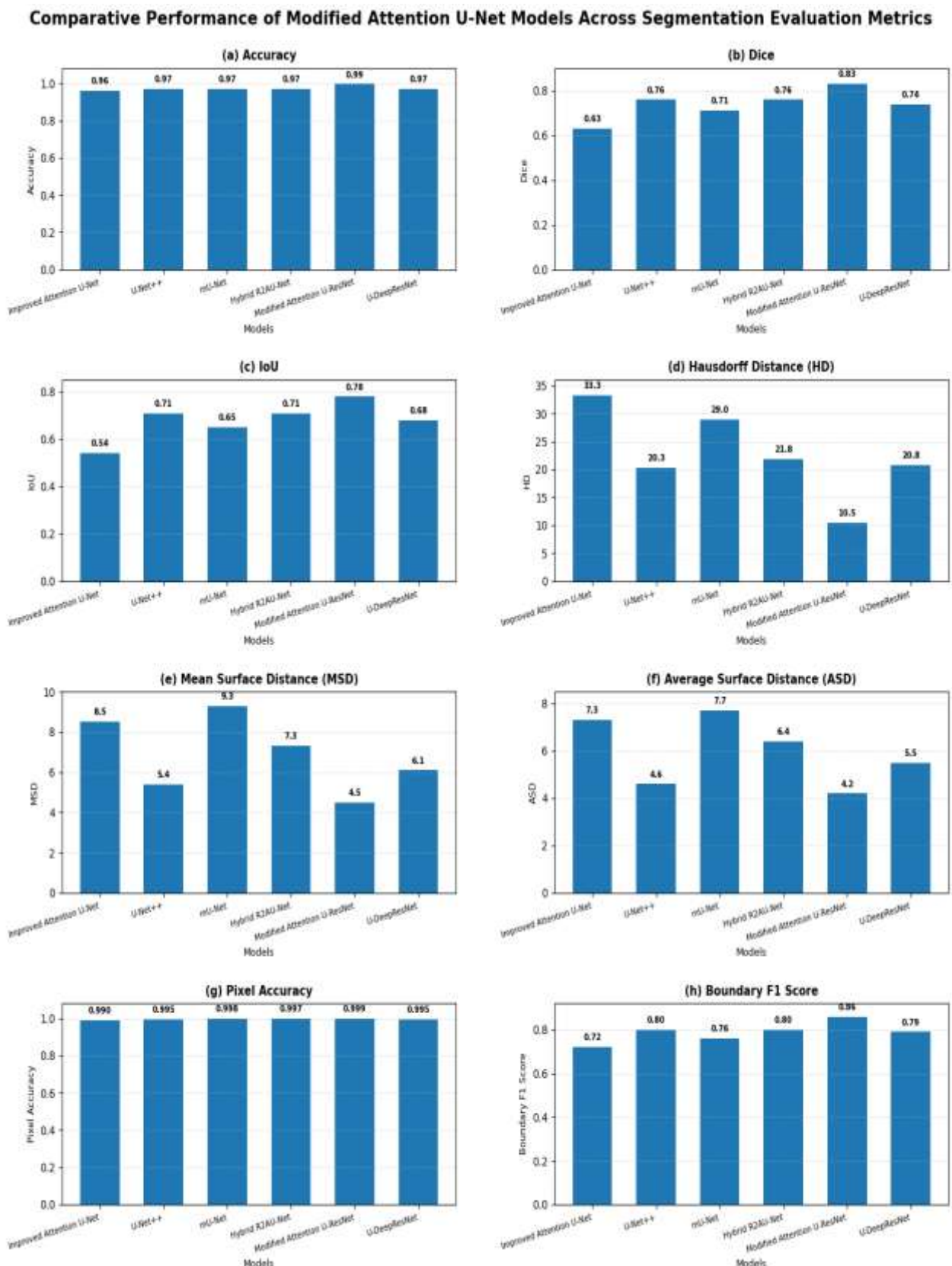


Figure 8. Comparison of segmentation performance metrics for different models on the 13-class dataset accuracy (a), dice score (b), IoU (c), HD (d), MSD (e), ASD (f), pixel accuracy (g), boundary F1 Score (h)

In figure 8, one can see the comparisons between metrics, when the proposed model demonstrates better () Dice, IoU, boundary F1, and distance-wise results. Accuracy is shown in figure 8 (a), Dice score in

figure 8 (b), IoU in figure 8 (c), HD in figure 8 (d), MSD in figure 8 (e), ASD in figure 8 (f), Pixel Accuracy in figure 8 (g), and Boundary F1 Score in figure 8 (h). Together, these plots demonstrate differences in overall segmentation accuracy, region overlap, pixel-level prediction, and boundary delineation performance among the compared models. From the clinical perspective, the found disc-condition Dice values denote that the model will be able to provide only the decision-support function, but not the full automated detection system. Normal condition and herniation have Dice values approximately equal to 0.73, which means that the model will be able to help detect and assess. But thinning (0.552) and bulging (0.659) prove to be more difficult because of their small size and visual proximity to other surrounding objects. So, the suggested model should assist radiologists during automatic segmentation, prioritisation of reviewing areas, and annotation work. Clinical decision-making will depend on the radiologist's analysis alongside patient history and other clinical factors.

Error Analysis

These segmentation errors were mostly encountered with disc-condition classes owing to their small size, low contrast, and similar visual appearance among the different pathologies. Error analysis was done based on the class-wise Dice performance of the proposed approach. The value of segmentation error was defined as $1 - \text{Dice}$, where larger values imply harder segmentations.

From these results, it can be seen that the largest segmentation errors occur with thinned disc due to poor representation and minute differences in disc heights, followed by bulging and herniated classes. Smaller segmentation errors were encountered with normal and herniation classes owing to clear boundaries. This shows that rare pathological classes are still hard for segmentation.

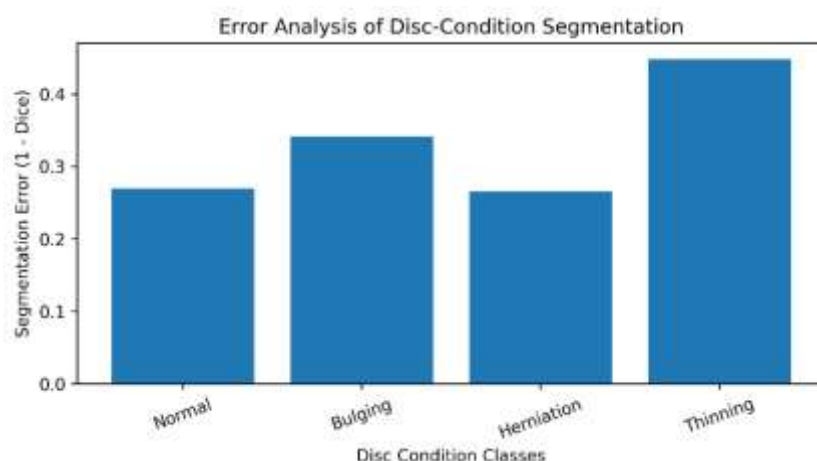


Figure 9: Error analysis of disc-condition segmentation performance

Figure 9 shows that the graph demonstrates the difference in segmentation errors of the classes of discs. This graph helps to explain the results of quantification analysis and also confirms that further research should be done on minority class improvement, optimization by boundaries, and multi-centre validation.

Qualitative and Ablation Discussion

Based on the above findings from the experiments, can deduce that the main enhancement lies in the incorporation of the additional semantic context gating branch. The two-branch model has no significant improvements over the baseline, but the complete model has better disc-condition segmentation accuracy. Most of the errors found are related to the thinning and herniation disc conditions, as they have smaller sizes compared to other disc conditions, which makes them visually similar to neighbouring disc conditions. As shown in figure 6, there are no mistakes in the segmentation of vertebra and canal structures, while disc boundaries have been identified as the source of error. The developed framework serves as a computer-aided diagnosis tool, which will help radiologists make better decisions.

DISCUSSION

The findings from the class-wise evaluation have more significance compared to the macro average because the primary contribution of this work is the segmentation of disc conditions in addition to anatomical structures. This proposed model achieves a better mean Dice on disc condition from 0.486 (MANet) to 0.669 without compromising segmentation of vertebral and canal structures. This shows that proposed attention mechanism has an impact on small regions and pathology-specific structures. From the findings in the ablation experiments, it is clear that the improvement is related to the semantic context gating branch. It was observed that the multi-scale attention branch and the channel-spatial attention branch do not contribute to better segmentation of disc conditions, while the additional global semantic modulation helps to discriminate between visually similar disc classes.

LIMITATIONS AND FUTURE WORK

There are some considerations regarding the presented results that need to be mentioned. First, the used data set is quite small and comes from one source, which might have an impact on the ability of the model to generalize to other scanners and protocols, as well as different types of population. Moreover, the size of the independent test set is only 27 subjects. It may result in possible uncertainty in performance estimation of each class because of the low number of patients with certain disc diseases. Second, each patient provides only one sagittal MRI slice that can affect three-dimensional context and decrease its influence on disc segmentation. Herniation and thinning classes continue to be difficult due to the insufficient number of samples and low visual distinctness.

Future research directions will involve working on imbalance-aware and boundary-aware loss functions, including weighted and focal loss function strategies, for rare class segmentation. Another important research task will be conducting an exhaustive component-wise ablation study to separately analyze the impact of each attention branch. Future work should also include model verification through cross-validation, confidence interval analysis, and statistical significance analysis along with utilizing multi-center datasets. Other future research directions would be working on a lightweight architecture, full 3D analysis based on MRI data, and using transformer models. Ultimately, the goal is to build a reliable CAD system that assists radiologists, while expert clinical judgment remains the final diagnostic authority.

CONCLUSION

The study presented the use of a modified attention U-DenseNet approach for the automatic segmentation of the 13 different classes of lumbar spine MRIs by incorporating multi-scale feature learning, channel and spatial attention, and semantic context gating methods. In contrast to other traditional techniques of segmenting the lumbar spine that consider only the anatomy of the vertebrae, the new framework not only segments the vertebrae regions but also the spinal canal and discs with different clinically important conditions such as normal, bulging, herniation, and thinning. The new approach was tested on 264 sagittal lumbar MRI scans, of which 211 were used for training, 26 for validation, and 27 for testing the model based on the split of subjects. The proposed network has achieved a macro-Dice score of 0.8440, a macro IoU of 0.7985, a macro-Boundary F1-score of 0.8693, and a macro Hausdorff distance of 10.5245 pixels. The maximum improvement was noticed in disc-condition segmentation, wherein the mean Dice coefficient increased from 0.486 to 0.669. In particular, there were improvements in terms of the Dice scores for normal discs (0.5431 to 0.7308), bulging discs (0.4313 to 0.6591), herniation (0.6442 to 0.7341) and thinning (0.3273 to 0.5523). The results of the ablation studies confirmed the contribution of the semantic context gating branch in improving pathology-aware segmentation. Yet, there were certain limitations, like the relatively small size of the single-source dataset, lack of validation outside this data source, one-slice analysis, incomplete component-wise ablation, and lack of statistical significance testing. For future research, it is planned to work on multi-centre validation, analysis based on 3D MRIs, optimisation for imbalanced and boundary-aware learning, full attention-module ablation, lightweight models and combining with transformer-based segmentation approaches. The framework presented in this study may be used as a

computer-aided decision-making system to assist radiologists in segmentation and region detection while leaving diagnosis and treatment to clinicians' expertise.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset generated in this study has been deposited in the Zenodo repository (DOI: 10.5281/zenodo.20634127). The repository record is publicly accessible; however, the dataset files are currently under embargo because they form part of ongoing research and will be made publicly available upon completion of the related study. The dataset was derived from the publicly available Lumbar Spine MRI Dataset by Sudirman et al. (DOI: 10.17632/k57fr854j2.2), distributed under the Creative Commons Attribution 4.0 (CC BY 4.0) license.

Acknowledgments: The authors would like to express their appreciation to Universiti Teknologi Malaysia (UTM) and the Northern Technical University, represented by its president, Dr.Alyaa Al-Attar, for the unlimited support provided to the researcher.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

- [1] Natalia F, Sudirman S, Ruslim D, Al-Kafri A. Lumbar spine MRI annotation with intervertebral disc height and Pfirrmann grade predictions. *PLoS One*. 2024 May 10;19(5): e0302067. <https://doi.org/10.1371/journal.pone.0302067>
- [2] van der Graaf JW, van Hooff ML, Buckens CF, Rutten M, van Susante JL, Kroeze RJ, de Kleuver M, van Ginneken B, Lessmann N. Lumbar spine segmentation in MR images: a dataset and a public benchmark. *Scientific Data*. 2024 Mar 2;11(1):264. <https://doi.org/10.1038/s41597-024-03090-w>
- [3] Ghobrial G, Roth C. Deep learning-based automated segmentation and quantification of the dural sac cross-sectional area in lumbar spine MRI. *Frontiers in Radiology*. 2025 Mar 25; 5:1503625. <https://doi.org/10.3389/fradi.2025.1503625>
- [4] Hess M, Allaire B, Gao KT, Tibrewala R, Inamdar G, Bharadwaj U, Chin C, Padoia V, Bouxsein M, Anderson D, Majumdar S. Deep learning for multi-tissue segmentation and fully automatic personalized biomechanical models from BACPAC clinical lumbar spine MRI. *Pain Medicine*. 2023 Aug 1;24(Supplement_1): S139-48. <https://doi.org/10.1093/pm/pnac142>
- [5] Hu X, Yang H. DRU-net: a novel U-net for biomedical image segmentation. *IET image processing*. 2020 Jan;14(1):192-200. <https://doi.org/10.1049/iet-ipr.2019.0025>
- [6] Litjens G, Kooi T, Bejnordi BE, Setio AA, Ciompi F, Ghafoorian M, Van Der Laak JA, Van Ginneken B, Sánchez CI. A survey on deep learning in medical image analysis. *Medical image analysis*. 2017 Dec 1;42:60-88. <https://doi.org/10.1016/j.media.2017.07.005>
- [7] Ali FH, Jarjees MS. Automated Liver Disease Classification Via Modified ResNet-50 Architecture on CT Images. *NTU Journal of Engineering and Technology*. 2025;5(1):243-51. <https://doi.org/10.56286/qr2kd436>
- [8] Wang K, Zheng M, Wei H, Qi G, Li Y. Multi-modality medical image fusion using convolutional neural network and contrast pyramid. *Sensors*. 2020 Apr 11;20(8):2169. <https://doi.org/10.3390/s20082169>
- [9] Mahmood OA, Yousif AS, Adam A. Brief review for multi-class brain tumor diseases schemes using machine learning techniques. *Al-Kitab Journal for Pure Sciences*. 2024 Jul 7;8(02):94-108. <https://doi.org/10.32441/kjps.08.02.p8>
- [10] Ghazal MT. A robust u-net-based approach for accurate brain tumor segmentation using multimodal MRI data. *NTU Journal of Engineering and Technology*. 2023 Nov 20;2(3). <https://doi.org/10.56286/ntujet.v2i3.692>
- [11] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In 2017 IEEE conference on computer vision and pattern recognition (CVPR) 2017 Jul 21 (pp. 2261-2269). IEEE. <https://doi.org/10.1109/CVPR.2017.243>
- [12] Chang C, Law H, Poon C, Yen S, Lall K, Jamshidi A, Malis V, Hwang D, Bae WC. Segment anything model (SAM) and medical SAM (MedSAM) for lumbar spine MRI. *Sensors*. 2025 Jun 7;25(12):3596. <https://doi.org/10.3390/s25123596>

- [13] Zhang X, Yang Y, Shen YW, Li P, Zhong Y, Zhou J, Zhang KR, Shen CY, Li Y, Zhang MF, Pan LH. SeUneter: Channel attentive U-Net for instance segmentation of the cervical spine MRI medical image. *Frontiers in Physiology*. 2022 Dec 6;13:1081441. <https://doi.org/10.3389/fphys.2022.1081441>
- [14] Saeed MU, Dikaios N, Dastgir A, Ali G, Hamid M, Hajjej F. An automated deep learning approach for spine segmentation and vertebrae recognition using computed tomography images. *Diagnostics*. 2023 Aug 12;13(16):2658. <https://doi.org/10.3390/diagnostics13162658>
- [15] Diotallevi L, Léger P, Beauséjour MH, Mac-Thiong JM, Petit Y. Trained nnU-Net model for semantic segmentation of human adult cervical vertebrae from CT-Scans. In 2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) 2025 Jul 14 (pp. 1-7). IEEE. <https://doi.org/10.1109/EMBC58623.2025.11253480>
- [16] Deng Y, Gu F, Zeng D, Lu J, Liu H, Hou Y, Zhang Q. An effective U-Net and BiSeNet complementary network for spine segmentation. *Biomedical Signal Processing and Control*. 2024 Mar 1;89 :105682. <https://doi.org/10.1016/j.bspc.2023.105682>
- [17] Kim DH, Seo J, Lee JH, Jeon ET, Jeong D, Chae HD, Lee E, Kang JH, Choi YH, Kim HJ, Chai JW. Automated detection and segmentation of bone metastases on spine MRI using U-Net: a multicenter study. *Korean Journal of Radiology*. 2024 Feb 29;25(4):363. <https://doi.org/10.3348/kjr.2023.0671>
- [18] Zeng D, Wang M, Mao F, Wang W, Lei T. Frequency-domain enhanced and feature fusion network for spinal MR image segmentation: D. Zeng et al. *Signal, Image and Video Processing*. 2025 Dec;19(15):1331. <https://doi.org/10.1007/s11760-025-04917-8>
- [19] Alsharif N, Nair R, Aldhyani TH, Farhah NS, Ahmad S, Al-Nefaie AH. Vertebrae and intervertebral discs segmentation using deep learning-based model in disability analysis. *Frontiers in Medicine*. 2026 Feb 25;13 :1723191. <https://doi.org/10.3389/fmed.2026.1723191>
- [20] Lu H, Li M, Yu K, Zhang Y, Yu L. Lumbar spine segmentation method based on deep learning. *Journal of applied clinical medical physics*. 2023 Jun;24(6):e13996. <https://doi.org/10.1002/acm2.13996>
- [21] Altun İ, Altun S, Alkan A. LSS-UNET: Lumbar spinal stenosis semantic segmentation using deep learning. *Multimedia Tools and Applications*. 2023 Nov;82(26):41287-305. <https://doi.org/10.1007/s11042-023-15205-3>
- [22] Wang S, Jiang Z, Yang H, Li X, Yang Z. Automatic segmentation of lumbar spine MRI images based on improved attention U-net. *Computational Intelligence and Neuroscience*. 2022;2022(1):4259471. <https://doi.org/10.1155/2022/4259471>
- [23] Wang Z, Xiao P, Tan H. Spinal magnetic resonance image segmentation based on U-net. *Journal of Radiation Research and Applied Sciences*. 2023 Sep 1;16(3):100627. <https://doi.org/10.1016/j.jrras.2023.100627>
- [24] Li H, Luo H, Huan W, Shi Z, Yan C, Wang L, Mu Y, Liu Y. Automatic lumbar spinal MRI image segmentation with a multi-scale attention network. *Neural Computing and Applications*. 2021 Sep;33(18):11589-602. <https://doi.org/10.1007/s00521-021-05856-4>
- [25] Sadulla S. Foundation Model–Powered Medical Dialogue Agents with Causality-Aware Workflow Reasoning. *Journal of Intelligent Assistive Communication Technologies*. 2026:49-56.