International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

A DEEP LEARNING FRAMEWORK FOR AUTOMATED SEGMENTATION AND CLASSIFICATION OF INTERVERTEBRAL DISC DEGENERATION

Abhishek Dipak Shroff[1]*, Sachin Patel[2], Rajesh Ku. Nagar[3], Jayshri Sonawane[4] Puja Saraf[5]

[1,2,3]*Author's Sage University, Indore 452020, Madhya Pradesh, India*

* Corresponding author's Email: drabhishekshroff@gmail.com

Abstract

Intervertebral disc (IVD) degeneration is a major contributor to chronic low back pain, yet its diagnosis through magnetic resonance imaging (MRI) remains time-consuming and subjective. This study proposes a robust deep learning framework integrating three state-of-the-art architectures U-Net, V-Net, and a novel IVD-ResNet for automated segmentation and classification of IVD degeneration into three categories: normal, mild, and severe. A multi-institutional dataset of 218 patient MRIs encompassing 3,535 discs was used for training and validation. The IVD-ResNet achieved superior classification performance with an overall accuracy of 98.30%, outperforming U-Net (83.31%) and V-Net (76.66%). Evaluation included precision, recall, F1-score, and confusion matrix analyses, with the IVD-ResNet demonstrating strong generalization and reliability across different degeneration levels. Comparative experiments highlight the benefits of hybrid feature integration and the potential of deep learning in clinical workflows. The proposed framework offers an accurate, efficient, and scalable solution for IVD degeneration analysis, with significant implications for early detection, personalized treatment planning, and reducing radiologist workload.

Keywords: Intervertebral Disc Degeneration, Deep Learning, MRI, U-Net, V-Net, ResNet, Medical Image Segmentation, Classification.

1. Introduction

Intervertebral disc (IVD) degeneration is a common musculoskeletal disorder and one of the leading causes of chronic low back pain worldwide [11], [37]. Its progression significantly impairs patient quality of life while increasing healthcare costs [15], [32]. Traditionally, radiologists rely on manual inspection of magnetic resonance imaging (MRI) scans to assess disc condition. However, this process is time-consuming, subjective, and prone to inter-observer variability [33], [36]. Therefore, automated diagnostic systems are increasingly sought after to improve accuracy, consistency, and efficiency [30], [35].

Recent advances in artificial intelligence (AI), particularly deep learning, have revolutionized medical imaging by enabling automated segmentation and classification of

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

complex anatomical structures [2], [6], and [18]. Models such as U-Net, V-Net, and ResNet-based frameworks have demonstrated remarkable performance across segmentation and classification tasks in biomedical imaging [1], [4], [9]. Yet, challenges remain in handling diverse imaging protocols, capturing subtle degenerative features, and balancing classification performance across different severity levels [12], [27], and [29].

This study addresses these challenges by introducing a comprehensive framework that leverages three complementary architectures: U-Net and V-Net for precise segmentation, and a customized ResNet (IVD-ResNet) for robust classification. By integrating multi-center MRI data with advanced pre-processing, augmentation, and hybrid feature learning, this work aims to achieve reliable and clinically applicable automated IVD degeneration diagnosis.

2. Literature Review

Deep learning has significantly advanced medical image analysis, particularly for spinal imaging where precise segmentation and classification of intervertebral discs (IVDs) are critical for accurate diagnosis and treatment planning. This section reviews existing literature in four main areas: segmentation, classification, hybrid and ensemble strategies, and current challenges.

Segmentation of IVDs from MRI is a critical first step for automated analysis. The U-Net architecture and its variants remain the most widely adopted due to their encoder–decoder design with skip connections, which preserve spatial details while enabling high-level feature extraction. U-Net has shown strong performance in biomedical applications even with limited training data, and extensions such as UNet++ and MultiResUNet further refine skip connections to bridge semantic gaps between encoder and decoder layers [1], [18], [24].

For volumetric data, V-Net introduced 3D convolutions, making it particularly suitable for spinal imaging tasks where inter-slice anatomical continuity is important [6]. Studies have demonstrated V-Net's ability to capture fine-grained spatial structures in MRI but also noted its computational demands and susceptibility to overfitting with small datasets. Advanced models such as IVD-Net leveraged multi-modal MRI inputs and dilated inception modules to enhance multi-scale feature learning, outperforming standard U-Net for disc localization and segmentation [4].

Another notable contribution is nnU-Net, which automatically adapts preprocessing, architecture, and training parameters to new datasets, achieving state-of-the-art results across multiple biomedical segmentation challenges without manual tuning [2]. Such self-configuring frameworks mark a shift towards more generalizable solutions. Despite these advances, segmentation accuracy often declines in cases of severe degeneration due to low contrast and structural deformation, highlighting the need for more robust architectures.

While segmentation delineates disc boundaries, classification focuses on determining degeneration severity. Early methods relied on handcrafted features such as texture descriptors, intensity histograms, and statistical measures [32], [33]. Although

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

computationally inexpensive, these approaches were highly sensitive to imaging artifacts and lacked robustness across diverse datasets.

Deep learning significantly improved classification performance by automatically learning hierarchical features. CNN-based pipelines have achieved strong results in multi-class IVD degeneration grading [12], [27]. Attention mechanisms have further enhanced these models by focusing on clinically relevant regions, thereby improving discrimination between subtle degeneration stages. More recent frameworks such as BianqueNet demonstrated high accuracy by linking degeneration patterns with demographic and clinical variables [29]. Similarly, deep networks trained on T1ρ-MRI data successfully detected early biochemical changes in disc tissue, offering new opportunities for pre-symptomatic diagnosis [25].

Although these models report impressive accuracy, most are limited by single-center datasets and lack external validation, raising concerns about their generalizability. Moreover, severe class imbalance where normal discs dominate datasets compared to degenerated discs often reduces classification sensitivity for early or mild degeneration.

To overcome the limitations of purely deep or purely handcrafted approaches, hybrid frameworks have Been proposed. Some studies fused CNN-derived deep features with machine learning classifiers such as Support Vector Machines (SVMs) and Random Forests to improve robustness and interpretability [16]. Others combined handcrafted descriptors (e.g., Gabor features, morphological metrics) with deep features, yielding superior classification accuracy compared to either approach alone [34].

Transfer learning and ensemble strategies have also proven effective in related medical domains. For instance, Efficient Net–ResNet hybrids applied to skin lesion classification achieved superior diagnostic accuracy, demonstrating the benefit of combining complementary architectures [8]. In spinal imaging, similar fusion strategies are emerging, though systematic benchmarking remains limited.

These hybrid approaches highlight the growing recognition that no single model captures all relevant information. Instead, ensembles and multi-feature pipelines may better reflect the complexity of degenerative changes across diverse patient populations and imaging modalities.

Despite substantial progress, several challenges remain unresolved. First, dataset limitations hinder generalization: most published studies rely on relatively small, homogeneous datasets, often from a single MRI scanner or clinical center, which restricts cross-population applicability [30]. Second, class imbalance disproportionately affects models' ability to detect early or mild degeneration, which is clinically important for preventive interventions.

Another challenge lies in robustness and explainability. Deep networks, while highly accurate, often operate as "black boxes," making it difficult for clinicians to trust their

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

predictions without clear interpretability. Furthermore, variability in MRI acquisition protocols across centers introduces heterogeneity in voxel resolution, contrast, and noise, which degrades performance when models are deployed outside their training domain [27], [29].

Most studies treat segmentation and classification as separate tasks. Few frameworks attempt to unify these processes into an integrated pipeline that both delineates anatomical structures and classifies degeneration severity. Such integration would more closely resemble the diagnostic workflow of radiologists and is therefore essential for clinical translation. Prior research has shown the potential of deep learning for automated IVD analysis but also highlighted persistent limitations. Segmentation methods like U-Net, V-Net, and their variants achieve high accuracy under controlled conditions but struggle with severely degenerated discs. Classification methods using CNNs and attention-based architectures achieve impressive performance but suffer from limited generalizability due to dataset imbalance and lack of external validation. Hybrid and ensemble methods offer promising directions but are not yet widely applied in spinal imaging.

This study addresses these gaps by proposing a multi-model framework that integrates segmentation (U-Net, V-Net) and classification (IVD-ResNet) into a unified pipeline. By leveraging multi-center datasets, advanced preprocessing, and hybrid feature learning, the proposed approach aims to deliver a more accurate, reliable, and clinically applicable system for automated IVD degeneration analysis

3. Proposed Methodology

The suggested methodology for classifying intervertebral disc (IVD) degeneration consists of many important stages, as shown in Figure 1. Initially, spinal MRI images are obtained and pre-processed in order to improve picture quality and normalize the input data.

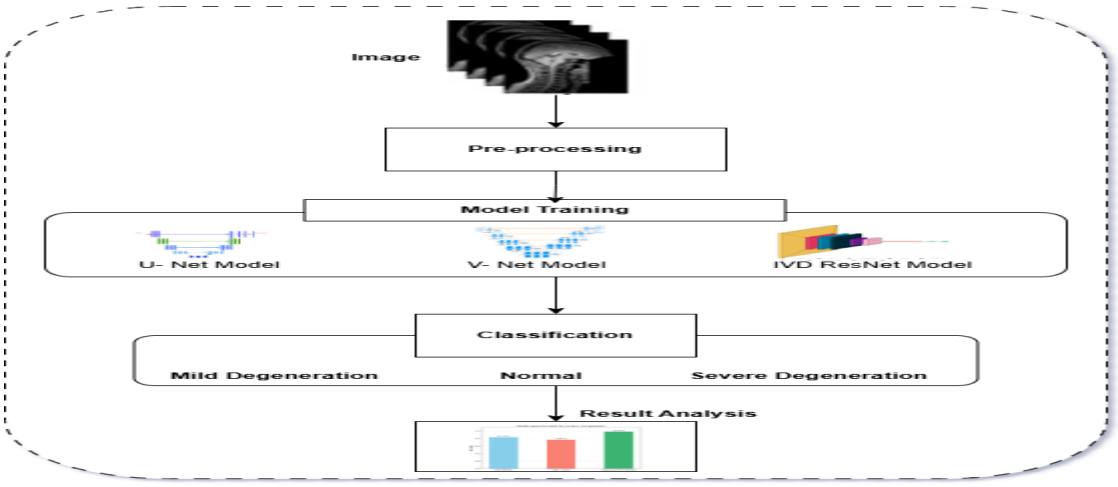

Figure. 1 Proposed Deep Learning Framework for Intervertebral Disc Degeneration Classification

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

This stage guarantees that data is consistent, allowing for effective model training. The pre-processed photos are then used to train three different deep learning models: U-Net, V-Net, and IVD ResNet. The U-Net and V-Net models are generally used for segmentation tasks to identify relevant anatomical structures, whilst the IVD ResNet model is designed for feature extraction and classification. These models are trained on annotated datasets to learn the spatial and textural patterns associated with disc degeneration. The models' outputs are then sent into a classification module, which divides the input photos into three categories: Mild Degeneration, Normal Degeneration, and Severe Degeneration. This multi-class classification gives a detailed examination of disc health state. Finally, a result analysis step evaluates the classification system's performance using relevant statistical and graphical criteria. This study confirms the robustness and reliability of the proposed deep learning pipeline for clinical use in spinal health evaluation.

3.1 Dataset

The proposed research employed MRI image data for the lumbar spine took from the public distribution by Graaf et al 2023. Four Netherlands hospitals collected de-identified MRI readings for research use in retrospective studies. The data collection involves the contributions from three types of healthcare facilities including University Medical Center UMC combined with regional hospitals RH1 and RH2 and orthopedic hospital OH. The time frame for data collection ran from January 2019 through March 2022 specifically for patients with chronic low back pain. The Radboud University Medical Center IRB (IRB 2016–2275) authorized the dataset collection under an IRB 2016–2275 approval which disregarded the need for patient authorization because research occurred on retrospectively analyzed anonymized data. A total of 218 patient MRIs is included which contain 447 MRI series for evaluation of both vertebral bodies and intervertebral discs. discs (IVDs). The research examined 3,535 IVDs from which 707 served testing functions and 2,828 functioned for training and validation purposes. The available MRI sequences consist of T1-weighted and T2-weighted images from the majority of cases while high-resolution T2 SPACE images are less common. Different hospitals possess varied MRI attainment methods that yield voxel resolutions ranging between 0.24×0.24×3.15 mm up to 1.06×1.23×9.63 mm. This multi-faceted data collection from multiple centers enables doctors to observe various spine pathologies within different imaging situations which provides necessary resources for fully automated IVD analysis modeling and system evaluation. The pre-processed dataset enabled scientists to train a CNN model based on ResNet50 architecture for detecting IVD degeneration.

3.2 Classification Framework

IVD-ResNet Model Overview: The proposed IVD-ResNet model operates as a deep convolutional neural network that implements a ResNet-like framework to achieve classification of images. The model works with grayscale images sized at 224×224 before it generates three-class predictions.

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

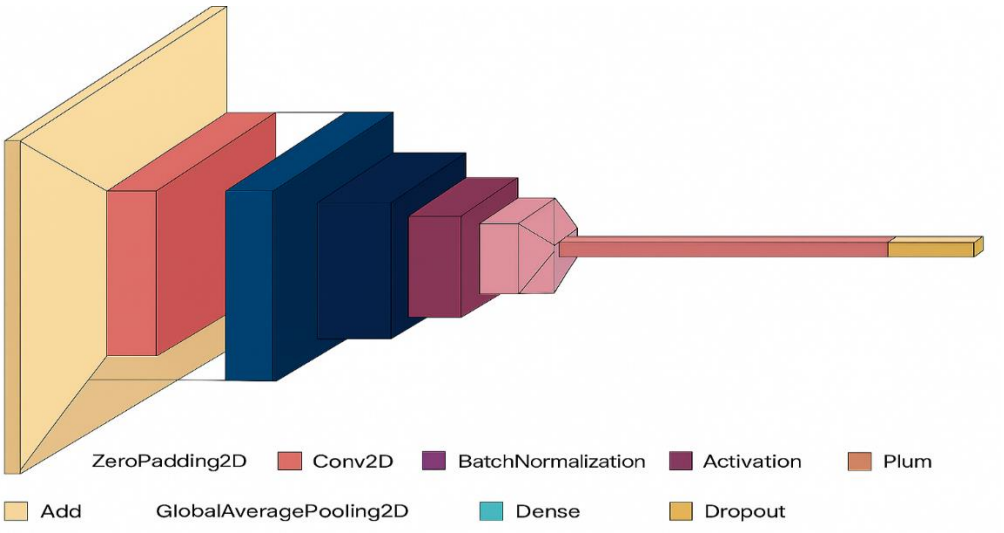

Figure. 2 Layer-wise Architecture of the Model

ResNet50 requires 3-channel RGB input images and therefore the grayscale MRI images went through conversion to meet the input requirements during the procedure. Such incorporation maintains the advantages of transfer learning since it enables 3-channel input processing through unmodified architecture structure and first convolutional layer initialization. The following segment outlines the structural design together with mathematical representation of the model's layers.

3.2.1 Input Layer

Layer Type: InputLayer

Output Shape: $(N, 224,224,1)$

Mathematical Description: here input layer accepts a set of grayscale images, where $N$the batch size, and each image has dimensions 224×224 with a single channel. The arrangement of this input tensor is mathematically represented in Equation (1):

$$X \in \mathbb{R}^{N\times 224\times 224\times 1} \quad (1)$$

3.2.2 Zero Padding Layer

Layer Type: ZeroPadding2D

Output Shape: $(N, 230,230,1)$

Mathematical Description: Zero padding adds a border of zeros around the input image to increase its spatial dimensions. For an input X, the padded output X_{padded}. The mathematical formulation of this padding process is provided in Equation (2):

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$X_{padded}[i,j,k] = \begin{cases} X[i-3, j-3, k] & \text{if } 3 \le i \le 226 \text{ and } 3 \le j \le 226, \\ 0 & \text{otherwise} \end{cases}$$

(2)

3.2.3 Convolutional Layer

Layer Type: Conv2D

Output Shape: $(N, 112,112,64)$

Mathematical Description: The convolutional layer put on a set of learnable filters (kernels) to the input. Each filter performs a cross-correlation operation, mathematically, this operation is defined in Equation (3):

$$Z[i,j,k] = \sum_{m=1}^{f_h} \sum_{n=1}^{f_w} \sum_{c=1}^{1} W[m,n,c,k] \cdot X_{padded}[i \cdot s + m, j \cdot s + n, c] + b[k]$$

(3)

where:

- $W \in \mathbb{R}^{f_h \times f_w \times 1 \times 64}$ is the kernel tensor,
- $b \in \mathbb{R}^{64}$ is the bias vector,
- f_h and f_w are the filter height and width,
- s is the stride (set to 2 in this layer),
- $Z \in \mathbb{R}^{N \times 112 \times 112 \times 64}$ is the output feature map.

3.2.4 Batch Normalization Layer

Layer Type: BatchNormalization

Output Shape: $(N, 112,112,64)$

Mathematical Description: Batch normalization normalizes the activations of the previous layer to improve training stability. For each feature channel k, the output is computed as shown in Equation (4):

$$\hat{Z}[i,j,k] = \frac{Z[i,j,k] - \mu_k}{\sqrt{\sigma_k^2 + \epsilon}} \cdot \gamma_k + \beta_k$$

(4)

where:

- μ_k and σ_k^2 are the mean & variance of the activations for channel k,

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- γ_k and β_k are learnable scale & shift parameters,
- ϵ is a small constant for numerical stability.

3.2.5 Activation Layer

Layer Type: Activation (ReLU)

Output Shape: $(N, 112,112,64)$

Mathematical Description: As defined in Equation (5), the Rectified Linear Unit activation function presents non-linearity by setting all negative activations to zero:

$$\mathrm{A}[i,j,k] = \max\left(0, \hat{\mathrm{Z}}[i,j,k]\right) \tag{5}$$

3.2.6 Max Pooling Layer

Layer Type: MaxPooling2D

Output Shape: $(N, 56,56,64)$

Mathematical Description: Max pooling decreases the spatial dimensions of the feature maps by taking the maximum value in each pooling window as defined in Equation (6):

$$\mathrm{P}[i,j,k] = \max_{m=1}^{p_h} \max_{n=1}^{p_w} \mathrm{A}[i \cdot s + m, j \cdot s + n, k] \tag{6}$$

where p_h and p_w are the pooling window dimensions, and s is the stride.

3.2.7 Residual Blocks

Layer Type: Residual Blocks (e.g., conv2_block1, conv2_block2, etc.)

Output Shape: Varies (e.g., $(N, 56,56,256)$ for conv2_block1)

Mathematical Description: Each residual block comprises of several convolutional layers with batch normalization and ReLU activation. The output of a residual block is computed as defined in Equation (7):

$$\mathrm{Y} = \mathcal{F}(\mathrm{X}) + \mathrm{X} \tag{7}$$

where:

- X input to the block,

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- $\mathcal{F}(\mathrm{X})$ denotes the transformations applied by the convolutional layers,
- Y is the output of the block.

3.2.8 Global Average Pooling Layer

Layer Type: GlobalAveragePooling2D

Output Shape: $(N, 2048)$

Mathematical Description: Global average pooling computes the average value of each feature map as defined in Equation (8):

$$\mathrm{G}[k] = \frac{1}{H \cdot W} \sum_{i=1}^{H} \sum_{j=1}^{W} \mathrm{Y}\,[i, j, k] \tag{8}$$

where H and W are the spatial dimensions of the input feature maps.

3.2.9 Fully Connected Layers

Layer Type: Dense

Output Shape: $(N, 256)$ and $(N, 3)$

Mathematical Description: The fully connected layers accomplish a linear transformation shadowed by an activation function as defined in Equation (9):

$$\mathrm{D}_1 = \mathrm{W}_1 \cdot \mathrm{G} + \mathrm{b}_1, \quad \mathrm{D}_2 = \mathrm{W}_2 \cdot \mathrm{D}_1 + \mathrm{b}_2 \tag{9}$$

Where $\mathrm{W}_1, \mathrm{W}_2$ are weight matrices, and $\mathrm{b}_1, \mathrm{b}_2$ are bias vectors.

3.2.10 Dropout Layer

Layer Type: Dropout

Output Shape: $(N, 256)$

Mathematical Description: As defined in Equation (10), Dropout randomly sets a fraction of the activations to zero during training to prevent overfitting:

$$\mathrm{D}_{\text{dropout}}[i] = \begin{cases} \mathrm{D}_1[i] & \text{with probability } p \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

where p is the dropout rate.

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

3.2.11 Output Layer

Layer Type: Dense (with softmax activation)

Output Shape: $(N, 3)$

Mathematical Description: The final layer usages a softmax activation to produce class possibilities as defined in Equation (11):

$$\mathrm{P}[i] = \frac{\exp(\mathrm{D}_2[i])}{\sum_{j=1}^{3} \exp(\mathrm{D}_2[j])} \qquad (11)$$

The proposed model leverages a ResNet-like architecture with residual blocks to enable the training of a deep network. It processes grayscale images through a series of convolutional, pooling, and fully connected layers, culminating in a 3-class classification output. The use of dropout, global average pooling and batch normalization guarantees vigorous and competent training.

Table 1. Variable / Symbol

Eqn. No.	**Variable / Symbol**	**Description**
(1)	X	Input image tensor
(2)	W,Z,	Convolutional filter weights, Output of convolution operation
(2)	s , b[k]	stride, Bias for k-th channel,
(3)	ϵ , Z,	Small constant to avoid division by zero, Normalized activation,
(3)	μk, σk2	Mean for channel k, Variance for channel k
(3)	βk , γk,	shift parameters & Trainable scale
(4)	A	activated output of ReLU
(5)	P	Output afterward max pooling
(5)	Pw, ph, s	Pooling window width, height, Stride in pooling
(6)	X, Y, F(X)	Input and output to residual block, Residual transformation
(7)	G[k], H,W	Global average value of the k-th feature map, Height and width

Received: August 10, 2025

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

(9)	b1, b2, D1, D2, W1,W2	Bias vector of dense layers, Outputs, weights
(10)	Ddropout, p	Output after applying dropout, Dropout rate
(10)	D2[i], D2[j], P[i]	Logit for class i and j, Probability of class i

Table 2. Hyperparameter Table

Hyperparameter	**Value**
Base Model	ResNet50
Input Shape	(224, 224, 1)
Top Layers	GlobalAveragePooling2D, Dense (256, RELU), Dropout (0.5), Dense (3, softmax)
Number of Classes	3
Image Size	224x224 pixels
Normalization	Pixel values divided by 255.0
Train-Test Split Ratio	80:20
Stratified Split	Yes
Augmentation Strategies, Rotation Strategies	20 degrees
	Width, Height Shift Range, Zoom Range= 0.1
Optimizer	Adam
Learning Rate	1×10^{-4}
Loss Function	Categorical Cross-entropy
Metrics	Accuracy, Precision, Recall, F1-score, Confusion Matrix
Batch Size	16
Epochs	20
Early Stopping	patience=7
Dropout Rate	0.5

3.3 U Net

The U-Net is a convolutional neural network (CNN) designed for semantic segmentation, particularly in biomedical image analysis. It features a symmetric encoder-decoder

Received: August 10, 2025

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

architecture with skip connections to preserve spatial details. The encoder uses convolution and max-pooling to down sample the input, capturing high-level features, while the decoder employs transposed convolutions (or upsampling) to reconstruct the spatial resolution. The skip connections bridge encoder & decoder layers, helping recover fine-grained information lost during downsampling. Unlike V-Net, U-Net primarily processes 2D images (though 3D variants exist) and is widely used for tasks like medical image segmentation (e.g., tumors, cells) due to its precision in localization and efficiency with limited training data.

3.3.1 Input Layer

Layer Type: Input Layer

Output Shape: (N,128,128,1)

Mathematical Description: The input layer takes a set of grayscale images. Each image has dimensions 128×128 with one channel. The input tensor is defined as:

$$X \in \mathbb{R}^{N\times 128\times 128\times 1} \quad (12)$$

3.3.2 Convolutional Layer (Encoder, Bottleneck, Decoder)

Layer Type: Conv2D

Output Shape: Varies by layer, e.g., (N,64,64,32) (N,32,32,64), etc.

Mathematical Description: Applies learnable filters using 2D convolution followed by ReLU activation. For an input tensor X, the output at position (i, j, k):

$$Z[i,j,k] = ReLU(Sum(\, W[m,n,c,k] * X[i+m, j+n, c]) + b[k]) \quad (13)$$

3.3.3 Max Pooling Layer

Layer Type: MaxPooling2D

Output Shape: Reduced spatial size, e.g., (N,64,64,32)→(N,32,32,32)

Mathematical Description: Reduces spatial dimensions via taking the maximum value over a 2×2 Box:

$$P[i,j,k] = \max_{m,n\in[0,1]} Z[i\cdot 2+m, j\cdot 2+n, k] \quad (14)$$

3.3.4 Upsampling & Skip Connection Layer

Layer Type: UpSampling2D + Concatenate

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Output Shape: Doubled spatial size + concatenated channels, e.g., (N,32,32,64+64)

Mathematical Description: Feature maps are upsampled using nearest-neighbor interpolation and concatenated with corresponding encoder feature maps:

$$U[i,j,:] = \text{Concat}\left(\text{UpSample}\left(Z_{\text{prev}}\right)[i,j,:], Z_{\text{skip}}[i,j,:]\right)$$

(15)

3.3.5 Global Average Pooling Layer

Layer Type: GlobalAveragePooling2D

Output Shape: (N, C) where C is the number of channels (e.g., 16)

Mathematical Description: Reduces each feature map to a single average value:

$$G[k] = 1/(H \cdot W) \times \Sigma(i=1 \text{ to } H)\ \Sigma(j=1 \text{ to } W)\ Z[i,j,k]$$

(16)

3.3.6 Dropout Layer

Layer Type: Dropout Output Shape: (N, 64)

Mathematical Description: Randomly sets some activations to zero during training to prevent overfitting:

$$D_{\text{drop}}[i] = \begin{cases} D[i], & \text{with probability } p \\ 0, & \text{otherwise} \end{cases} \quad (17)$$

3.3.7 Output Layer (Softmax Classifier)

Layer Type: Dense with Softmax Activation

Output Shape: (N, 3)

Mathematical Description: Converts the final dense layer output into class probabilities:

$$P[i] = \exp(D\_drop[i]) / (\Sigma(j=1 \text{ to } 3) \exp(D\_drop[j]))$$

(18)

Table 3. Equation Notation List

Symbol	Description

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

N,C	Batch size (number of input images), Number of channels (feature maps)
H,W	Height and width of the feature map/image
X	Input image tensor of shape (N,128,128,1)
W[m,n,c,k]	Weight of the convolutional filter at position (m,n) for input channel c and output channel k
b[k], G[k]	Bias term for the kth output channel, Output of the global average pooling
Z[i,j,k], P[i,j,k]	Output of the convolution operation, Output of the max pooling operation
U[i,j,:]	Output after upsampling and concatenation at position (i,j)
Zprev, ZskipZ	Decoder input feature map to be upsampled, encoder feature map for skip connection
D[i]	Activation from the dense (fully connected) layer
Ddrop[i]	Output after applying dropout to activation D[i]
P, P[i]	Dropout retention probability (e.g., 0.5), Softmax probability for class i
UpSample(·)	Upsampling operation (e.g., nearest-neighbor or bilinear)
Concat(·,·)	Channel-wise concatenation of feature maps
ReLU(·)	Rectified Linear Unit activation function: max(0,x)

Table 4. Hyperparameters

Parameter	**Value**
Input Image Size	128×128×1 (grayscale)
Optimizer	Adam optimizer
Learning Rate	1×10−4
Batch Size	16
Number of Epochs	20
Loss Function	Categorical Cross-entropy
Metrics	Accuracy, Precision, Recall, F1-score, Confusion Matrix
Data Split Ratio	80:20
Augmentation Strategies	- Random rotation (±15°) -

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

	Horizontal/vertical flip - Zoom (±10%) - Contrast adjustment
Dropout Rate	0.5
Normalization	Min-max normalization to [0, 1]
Early Stopping	Patience = 10 epochs

3.4 V Net

The V-Net model represents a convolutional neural network which specializes in volumetric (3D) data analysis particularly for medical applications. The specific V-Net architecture uses 3D convolutions with pooling and upsampling to expand U-Net capabilities for processing three-dimensional spatial relationships along depth and width dimensions. The model variant employs global average pooling followed by a dense layer classifier head for classifying purposes.

3.4.1 Input and Reshape Layer

Layer Type: Input + Reshape

Output Shape: (N, 128, 128, 1, 1)

Mathematical Description: The input is a 2D grayscale image reshaped to introduce a pseudo-depth dimension:

$$X \in \mathbb{R}^{N\times 128\times 128\times 1} \Rightarrow X_{3D} \in \mathbb{R}^{N\times 128\times 128\times 1\times 1} \quad (19)$$

3.4.2 3D Convolution Layer

Layer Type: Conv3D

Output Shape: e.g., (N, 128, 128, 1, 16) after first layer

Mathematical Description: The 3D convolutional layer put on learnable 3D kernels over width, height, and depth:

$$Z[i,j,d,k] = ReLU(\Sigma_m \Sigma_n \Sigma_l \Sigma_c W[m,n,l,c,k] * X[i+m, j+n, d+l, c] + b[k]) \quad (20)$$

Where: $W \in \mathbb{R}^{f_h \times f_w \times f_d \times C_{\text{in}} \times C_{\text{out}}}$

3.4.3 3D Max Pooling Layer

Layer Type: MaxPooling3D

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Output Shape: Reduced spatial size, e.g., (N, 64, 64, 1, 16)

Mathematical Description: Downsamples the feature volume by taking the maximum in a 3D window:

$$P[i,j,d,k] = max\_\{m,n,l \in [0,1]\}Z[i*2+m, j*2+n, d+l, k]$$

(21)

3.4.4 3D Upsampling and Skip Connections

Layer Type: UpSampling3D + Concatenate

Output Shape: Upsampled and merged volume, e.g., (N,128,128,1,32)

Mathematical Description: Feature maps are upsampled in 3D and concatenated with skip-connected encoder features:

$$U[i,j,d,:] = \text{Concat}(\text{UpSample}(Z_{\text{dec}})[i,j,d,:], Z_{\text{enc}}[i,j,d,:])$$

(22)

3.4.5 Global Average Pooling Layer (3D)

Layer Type: GlobalAveragePooling3D

Output Shape: (N, C) e.g., (N,16)

Mathematical Description: Collapses each 3D feature map to a single scalar by averaging all spatial locations:

G[k] = (1 / (H * W * D)) * [Σ_{i=1 to H} Σ_{j=1 to W} Σ_{d=1 to D} Z[i,j,d,k]]

(23)

3.4.6 Fully Connected, Dropout & Output Layers

Layer Type: Dense → Dropout → Dense (Softmax)

Output Shape: (N,64→ (N,64) → (N,3)

Mathematical Description: Final layers project pooled features into class logits and convert them into probabilities:

$$D_1 = \text{ReLU}(W_1 G + b_1)$$

$$D_{\text{drop},i} = \begin{cases} D_{1,i}, & \text{with probability } p \\ 0, & \text{otherwise} \end{cases}$$

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$P_i = \frac{e^{z_i}}{\sum_{j=1}^{3} e^{z_j}} \quad (24)$$

Table 5. Equation Notation Table for V-Net Architecture

Symbol	Description
H, W, D, N	Height, width, and depth of the 3D feature map, Batch size (number of input samples)
Cin, Cout	Number of input channels, Number of output channels
X	Input 2D grayscale image: $X \in \mathbb{R}^{N\times 128\times 128\times 1}$
X_{3D}	Reshaped input with fake depth: $\in \mathbb{R}^{N\times 128\times 128\times 1\times 1}$
W[m,n,l,c,k]	Weight of the 3D convolution filter at position(m,n,l), input channel c, output channel k
Z[i,j,d,k]	Output of 3D convolution at position (i,j,d) and channel k
P[i,j,d,k]	Output of max pooling at spatial location (i,j,d) for channel k
Z_{enc}, Z_{dec}	Encoder feature map used in skip connection, Decoder feature map to be upsampled
U[i,j,d,:]	Concatenated feature map at decoder stage
UpSample(·)	Upsampling operation in 3D
Concat(·,·)	Concatenation along the channel axis
D1, D2[i]	Output of the first dense layer, Logit for class i before softmax
$D_{drop}[i]$	Output after dropout applied to neuron i
P, P[i]	Dropout retention probability (e.g., 0.7), Final softmax probability for class i

Table 6. Hyperparameters

Model Name	Custom 3D V-Net Variant (Functional API)
Input Shape	(128, 128, 1) reshaped to (128, 128, 1, 1)
Total Parameters	59,189
Optimizer Type	Adam

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Learning Rate	1×10−4
Batch Size	16
Number of Epochs	20
Loss Function	Crossentropy
Evaluation Metric(s)	Precision, Recall, F1-score, Accuracy , Confusion Matrix
Data Split Ratio	80:20
Augmentation Strategies	ImageDataGenerator:•rotation_range=20°,width_shift_range=0.1,height_shift_range=0.1,zoom_range=0.1,
Regularization	Dropout (rate=0.3) Dense layer
Classification Head	GAP → Dense(64, relu) → Dropout(0.3) → Dense(3, softmax)

4. Result & Discussion

IVD ResNet:

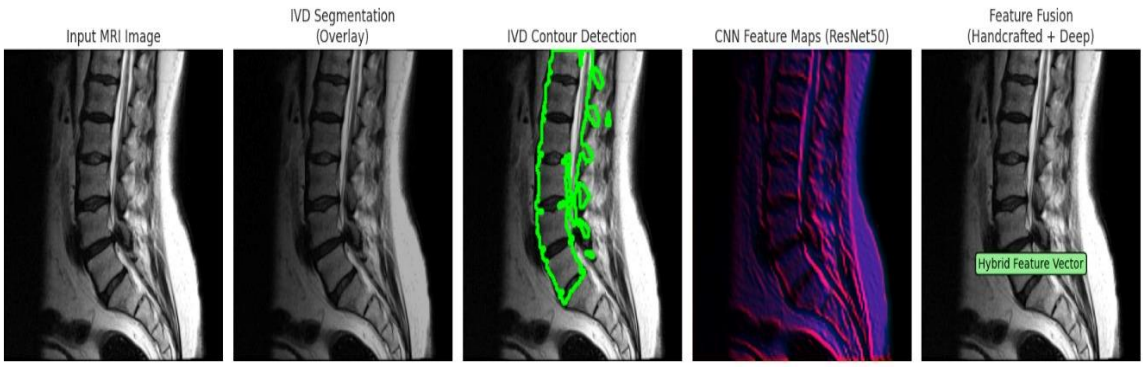

Figure. 3 Overview of the proposed IVD classification pipeline

Figure 3 demonstrates a step-by-step workflow for the automated analysis and classification of intervertebral disc (IVD) conditions using spinal MRI images. It begins with the input of a sagittal MRI scan, which undergoes segmentation to isolate the IVD regions. These segmented areas are overlaid onto the original scan for better visualization. In the next step, contour detection is applied to extract the boundaries of the discs, shown prominently with green outlines. Deep features are then extracted from the segmented MRI using a ResNet50 convolutional neural network, highlighting important anatomical structures through feature maps. These deep features are combined with handcrafted features such as texture, morphological, and statistical descriptors to form a hybrid feature vector. This fusion

Received: August 10, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

ensures a more comprehensive representation of the disc characteristics. The resulting hybrid vector can then be used for downstream tasks like classification, aiding in the diagnosis of spinal conditions with improved accuracy and interpretability.

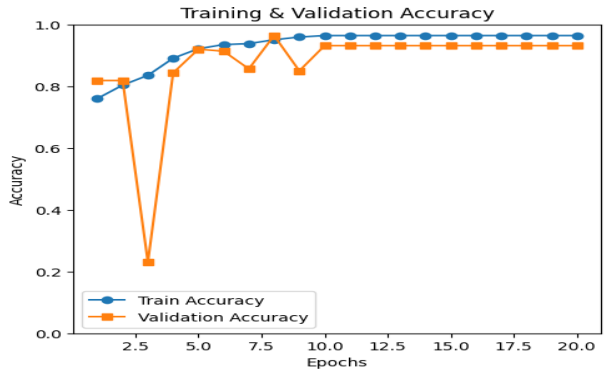

Figure. 4(a). Training and validation accuracy

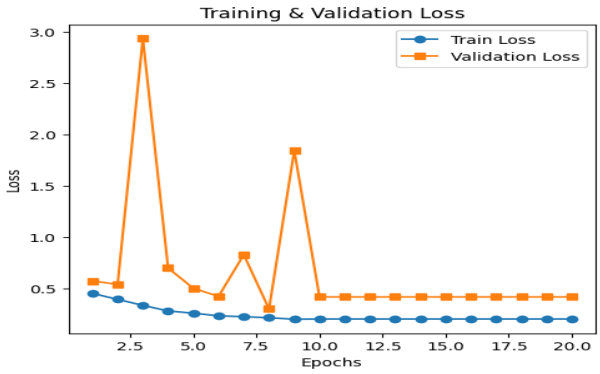

Figure. 4(b). Training and validation loss

The model achieves high accuracy results of 96-97% throughout its five training epochs according to Figure 4. The validation accuracy shows short-term instabilities before reaching stability because it first undergoes an initial period of overfitting which gets corrected. During the first phase of training the loss drops rapidly but validation loss shows major increases until it settles down because of probable early overfitting. The validation loss achieved stability along with training loss convergence during epoch 10 thus indicating better generalization between the datasets. The model shows successful performance and convergence outcomes based on its general trend observation.

Confusion Matrix

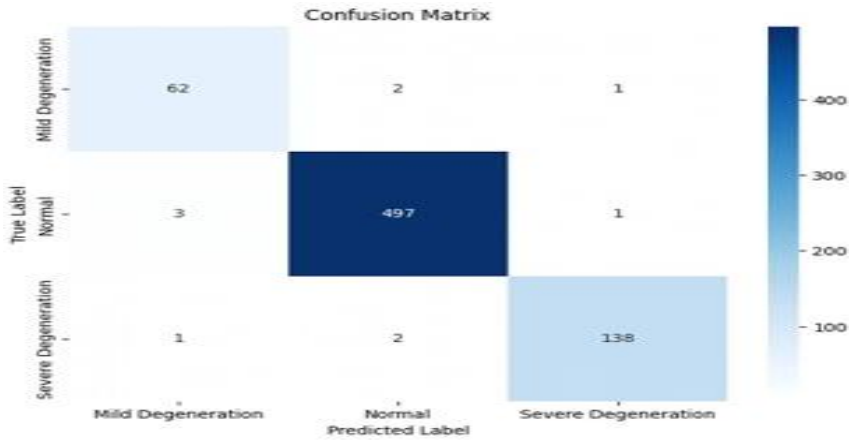

Figure. 5 Confusion matrix for 3-class intervertebral disc degeneration classification using IVD ResNet

The confusion matrix displayed IVD-ResNet model performance regarding intervertebral disc (IVD) degeneration classification through its evaluation of the Mild Degeneration and Normal and Severe Degeneration categories. The model achieves successful classification

Received: August 10, 2025

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

outcomes by identifying 62 Mild Degeneration cases correctly and 497 Normal cases alongside 138 Severe Degeneration cases. Only several cases of Normal tissue received incorrect classification as either Mild or Severe Degeneration and between two classes of degeneration in the entire dataset. The model distinguishes between IVD degeneration categories yet fails to identify 2 Mild cases correctly either as Normal or Severe and misplaces 3 Normal cases under Mild or severe categories while identifying 1 Severe case as Mild and Normal in 2 instances. The model displays high reliability in classifying different degrees of IVD degeneration through its pattern of concentrated values forming a diagonal line and minimal values appearing in off-diagonal positions.

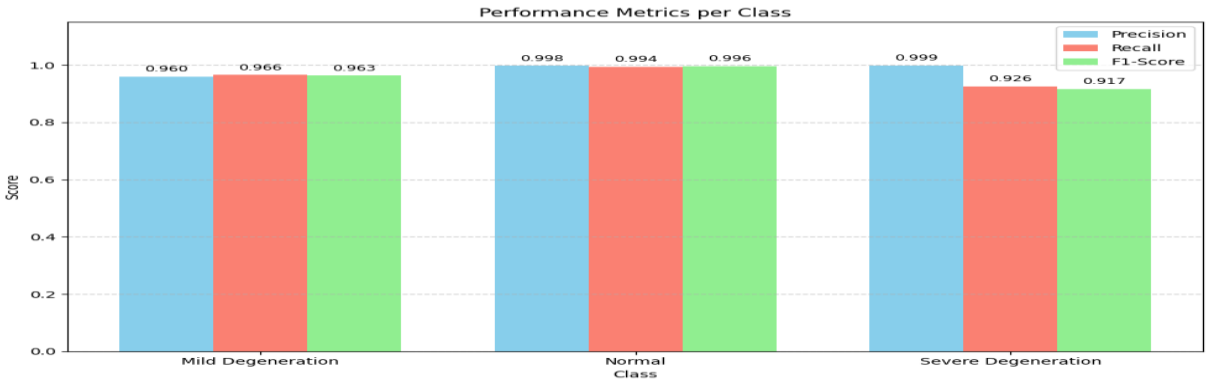

Figure. 6 Class-wise precision, recall, and F1-score of the IVD ResNet model

Figure 6 shows the proposed model's class-wise performance measures. The Normal class had the best results (F1-Score: 0.99, Precision: 0.99, Recall: 0.99), followed by Mild Degeneration (F1-Score: 0.96, Precision: 0.96, Recall: 0.96). The Severe Degeneration class had somewhat poorer results (F1-Score: 0.91, Precision: 0.99, Recall: 0.92), indicating a higher difficulty in correctly diagnosing severe cases. Overall, the results validate the model's robustness and dependability in all categories.

U Net Model:

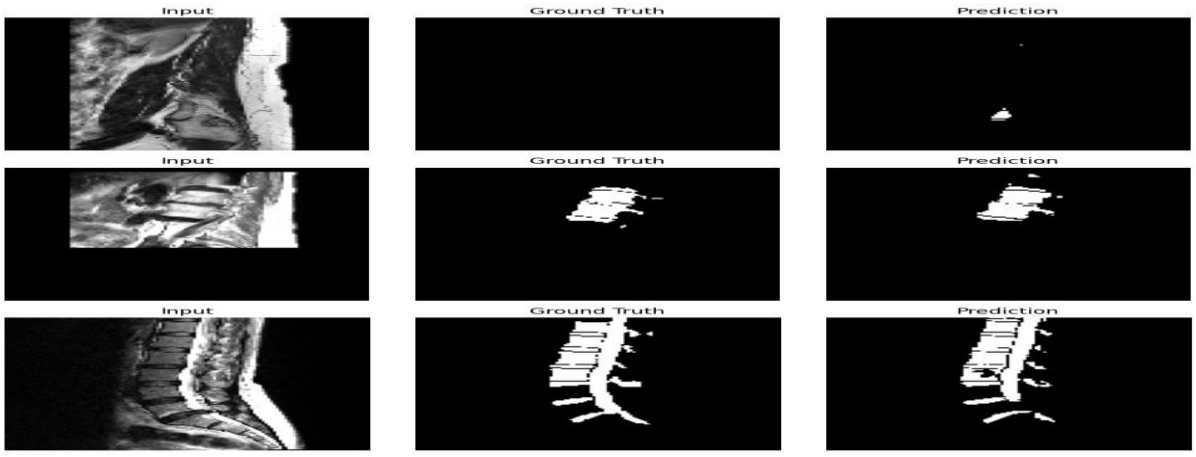

Figure 7. Segmentation results of the proposed U-Net model

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Figure 7 shows Visualization of segmentation results using the U-Net model on lumbar spine MRI slices. Each row illustrates a different case with three columns: the input MRI image, the ground truth mask, and the model prediction. A different clinical case appears in every row of the figure while it shows three visual elements starting with MRI input images followed by ground truth segmentations and predicted results. In the first row the model produces a prediction of disc structure in a region where the ground truth shows no disc structures which could imply potential false positive identification. This case demonstrates moderately complex structures of vertebral discs through successful segmentation with effective matching between projected results and ground truth.

The offered design exhibits excellent performance by matching the complex anatomy of multiple intervertebral discs as well as spinal features with high accuracy between predicted and actual mask outputs.

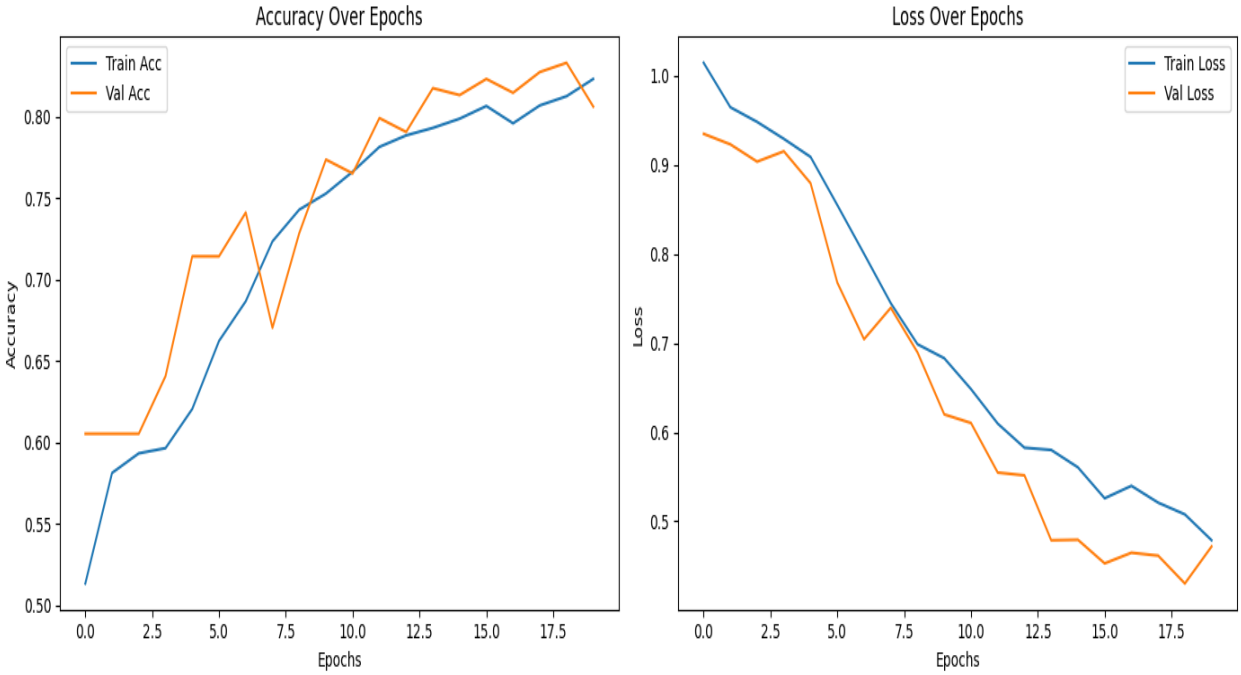

Figure 8. Training and validation accuracy & loss curves of U-Net model

The U-Net model's training and validation accuracy loss data are presented across 20 epochs in the presented plot. Both training and validation phases show consistent accuracy improvement in this plot which achieves validation accuracy greater than 83% indicating good generalization performance. The graphs indicating loss suggest that the model steadily decreased until it reached convergence. The model successfully prevents overfitting because the training and validation trends show strong parallelism throughout the process.

Received: August 10, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

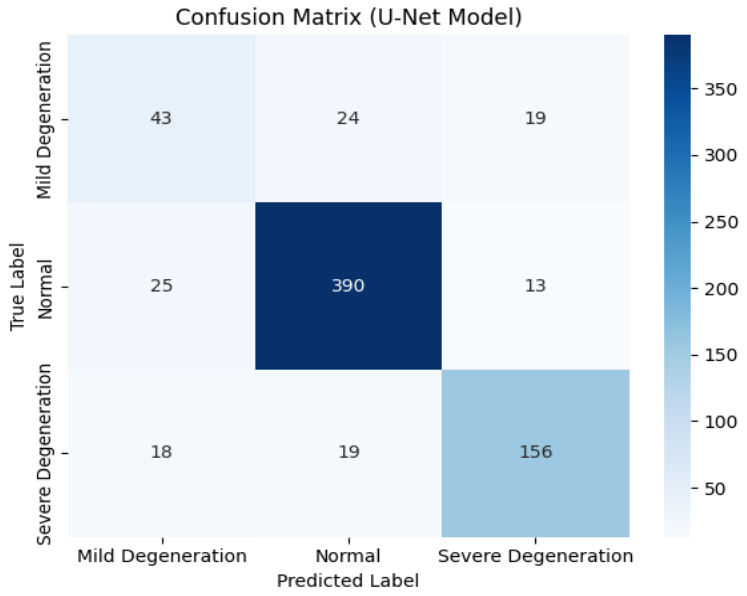

Figure 9. Confusion matrix for 3-class intervertebral disc degeneration classification using U-Net

The U-Net model performance evaluation through confusion matrix displays its ability to identify Mild Degeneration and Normal and Severe Degeneration classes. An analysis of the model indicates proper identification of 390 Normal samples along with 156 Severe Degeneration and 43 Mild Degeneration cases. The majority of mistakes between different classes (Mild vs. Normal) indicate overlapping boundaries of those categories. Through the matrix structure it shows how well the model performs as well as which classes it struggles to identify accurately.

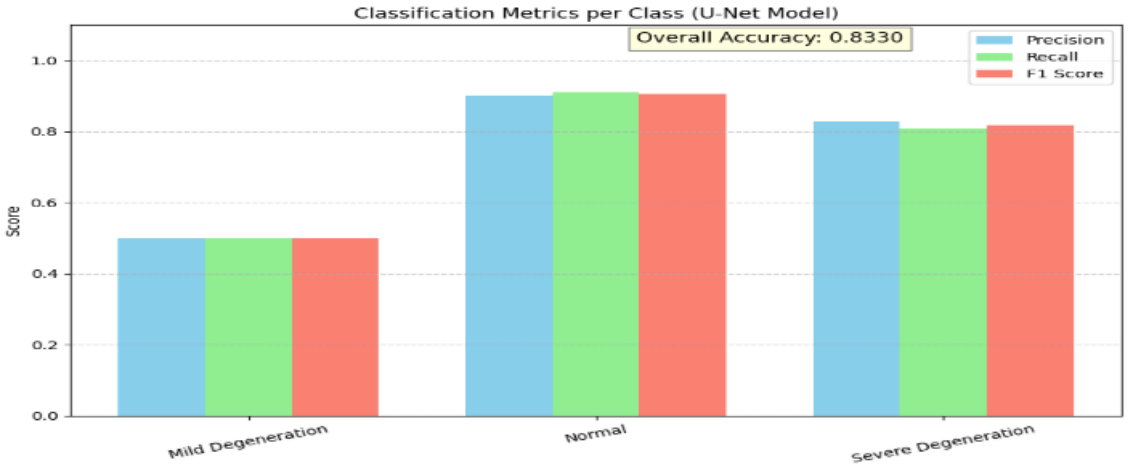

Figure 10. Class-wise precision, recall, and F1-score of the U-Net model

The U-Net model achieved evaluation metrics measurements for F1-score, precision and recall in relation to the three degeneration categories. Precision and recall values approach 0.91 for "Normal" class readings and scores near 0.82 for "Severe Degeneration" class

Received: August 10, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

readings. The "Mild Degeneration" class has lower performance metrics (~0.50) because there is a shortage of available samples and nearby classes have similar visual characteristics. The model reaches 83.3% accuracy in total classifications.

V Net Model:

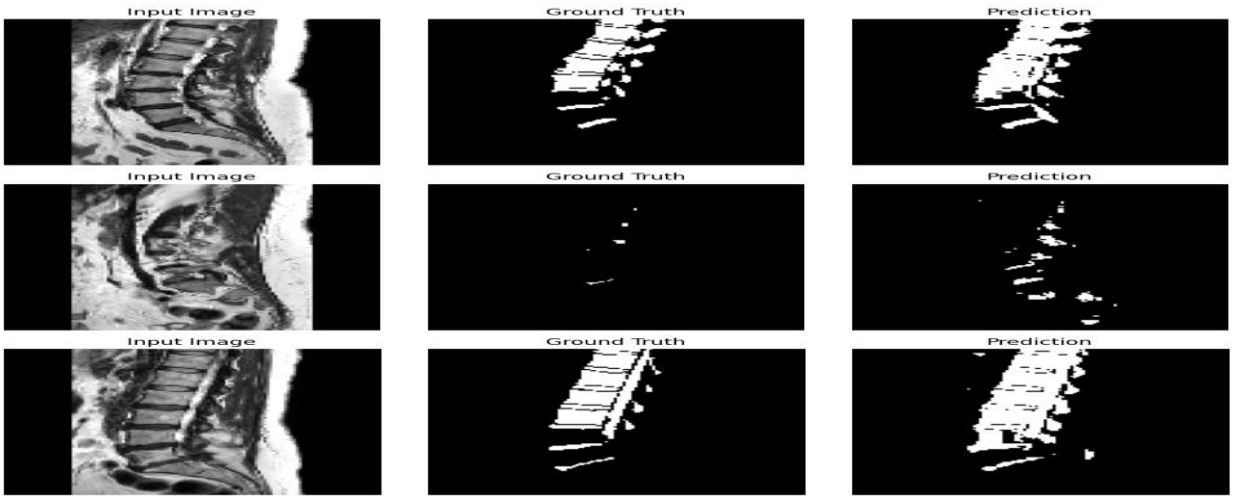

Figure 11. V-Net Predictions vs. Ground Truth

A grid system visually displays input images to ground truth segmentation masks and V-Net predictions and enables quality-based performance evaluation of the model. The predictions match the ground truth accurately for Normal bone structures yet they exhibit differences with ground truth for Mild and Severe Degeneration cases due to segmentation inaccuracies. The visuals from the segmentation process match quantitative values shown in the confusion matrix (Figure X), which demonstrates segmentation difficulties when dealing with degenerated regions. Moreover, such qualitative analysis provides essential information about failure modes to direct subsequent development of improvements which may include multi-scale features or post-processing methods to enhance segmentation boundaries.

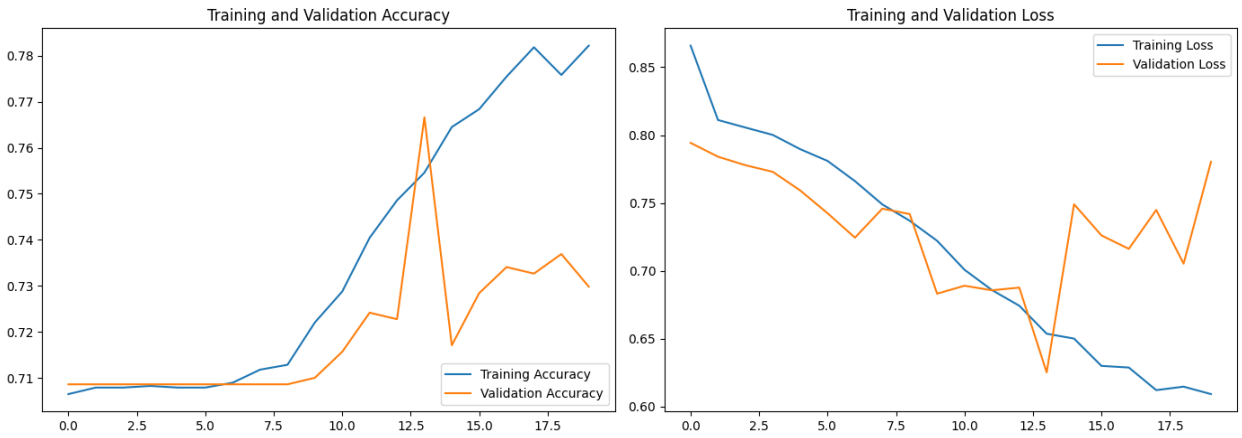

Figure 12. Training and Validation Curves for V-Net Model

Received: August 10, 2025

International Journal of Applied Mathematics

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The V-Net model presents its learning behavior across epochs through training and validation curve representations. The model training reaches a stable point at approximately 0.75 accuracy yet validation accuracy achieves its highest point at 0.85 before showing signs of overfitting behavior. The training loss shows gradual decline yet the validation loss incapacitates or increases after epoch 10. The observed patterns show that the model would benefit from regularizing itself through dropout or early stopping methods to achieve better generalization. Model performance quality depends on continuous evaluation of training and validation metrics because they demonstrate different behavior patterns.

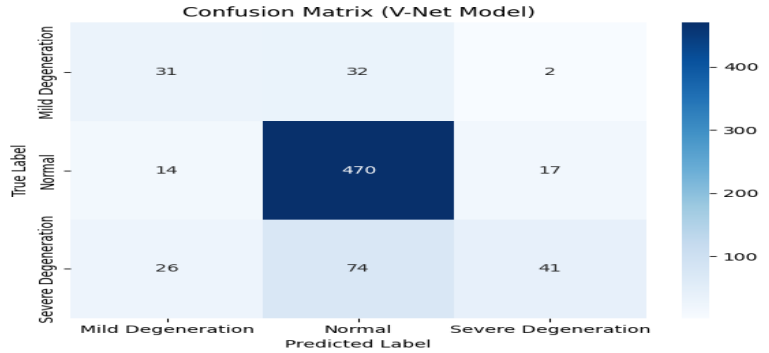

Figure 13. Confusion Matrix for V-Net Model

The V-Net model classification performance excellence can be recorded through a confusion matrix analysis which opposes true label data to predicted label data for Normal and Mild Degeneration and Severe Degeneration tissue types. Stats from the confusion matrix show that Normal cases were accurately classified 31 times while Mild Degeneration and Severe Degeneration had many incorrect predictions (14 and 26 respectively). The matrix reveals a class distribution issue because Severe Degeneration contains only 17 true positive cases while the model misidentifies cases at a high rate. The current model performance indicates it struggles to distinguish different degenerative severities because such separations require advancements both in features extraction methods and supplementary contextual data.

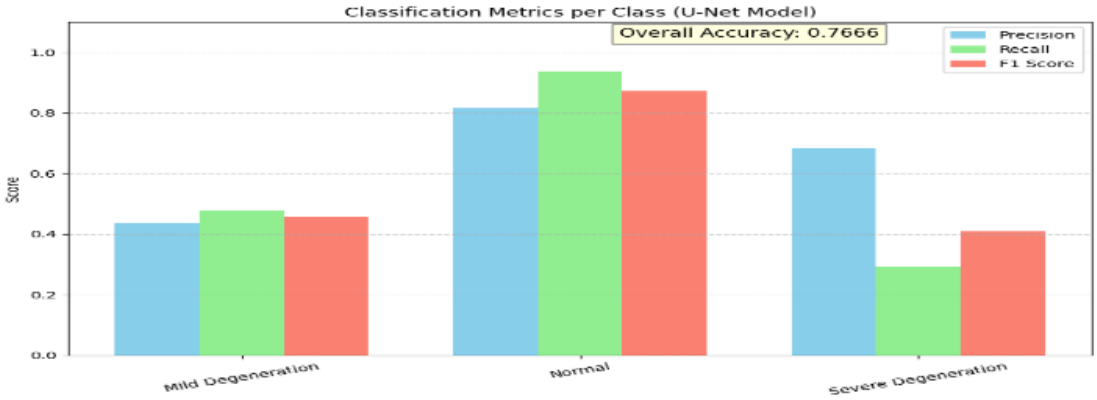

Figure 14. Classification Metrics per Class for V-Net Model

Received: August 10, 2025

International Journal of Applied Mathematics
Volume 38 No. 11s 2025,
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The V-Net model performs testing on three classes named as Normal, Severe Degeneration and Mild Degeneration and while showing its precision, recall and F1 scores through this bar graph representation. The model achieves a total accuracy metric of 0.766 yet demonstrates diverse accomplishment across different categories. Medical segmentation tasks require F1 scores as a balanced assessment metric of model robustness particularly because they effectively measure precision and recall performance. The metrics used for Mild Degeneration classification show lower performance than other classes based on the graph findings. The performance indicates a need to implement specific enhancements like data augmentation techniques alongside class rebalancing methods for better outcomes with underrepresented and complex categories.

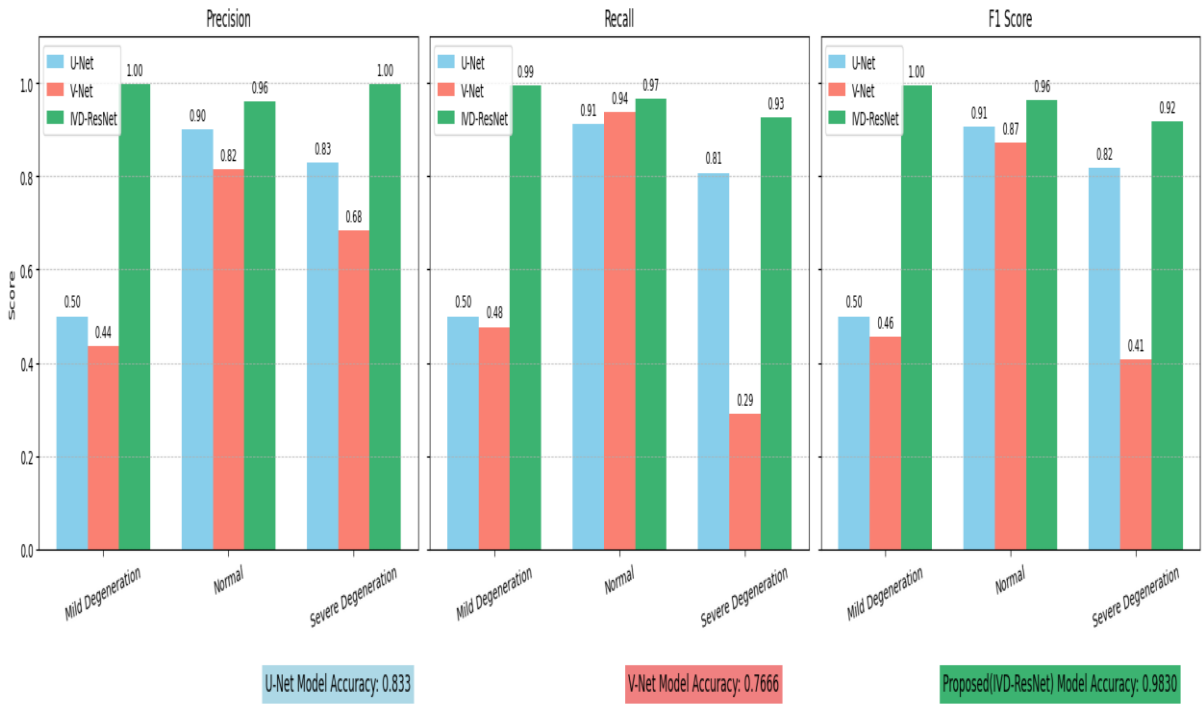

Figure 15. Comparison of U-Net, v-Net and Proposed (IVD-ResNet) Metrics per class

The figure explores the three deep learning models U-Net, V-Net alongside the proposed IVD-ResNet through their precision calculations with recall metrics and F1-score measurements for diagnosing three conditions between mild degeneration, normal and severe degeneration. The IVD-ResNet exhibits complete precision accuracy at scores of 1.00 when classifying both mild and severe degeneration and scores 0.96 for the normal category. The precision statistics from U-Net show effectiveness at 0.50, 0.90, and 0.83 while V-Net performs with 0.44, 0.82, and 0.68 for detecting mild and normal and severe degeneration respectively. The performance of IVD-ResNet regarding recall shows excellence with results of 0.99 for mild degeneration and 0.97 for normal and 0.93 for severe degeneration. The recall measurements of U-Net amount to 0.50 and 0.91 and 0.81 while V-Net recorded 0.48, 0.94 and 0.29 for these categories. The F1-score for IVD-ResNet presents stable results with

Received: August 10, 2025

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

1.00 for mild degeneration and 0.96 for normal degeneration and 0.92 for severe degeneration. The retrieval capability of U-Net surpasses 0.50 while reaching 0.91 and 0.82 while V-Net achieves lower scores of 0.46 and 0.87 and 0.41. The proposed IVD-ResNet achieves a superior overall classification accuracy of 0.9830 while U-Net obtains 0.833 and V-Net reaches 0.7666.

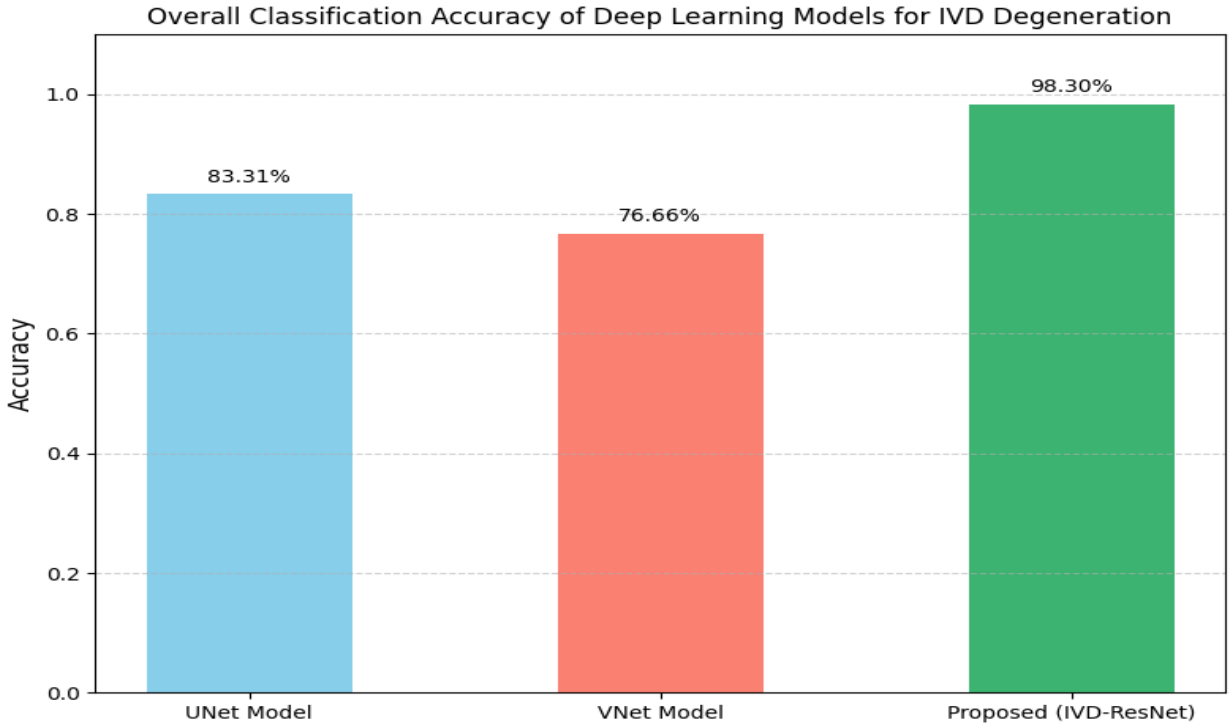

Figure 16. Overall Classification Accuracy Comparison of U-Net, V-Net, and Proposed IVD-ResNet Models for IVD Degeneration

The bar chart shows the total classification accuracy rates where IVD-ResNet surpasses its competitors U-Net and V-Net for identification of intervertebral disc degeneration. The IVD-ResNet model delivers maximum accuracy among all models reaching 98.30% classification precision. The U-Net model achieves 83.31% accuracy but the V-Net model shows 76.66% accuracy in the assessment. The analysis demonstrates that IVD-ResNet provides exceptional accuracy as a current architecture for IVD degeneration stage classification which will support medical applications in diagnosis.

5. Conclusion

A deep learning-built pipeline for the automated categorisation of intervertebral disc degeneration from lumbar spine MRI was effectively created and validated in this study. We presented IVD-ResNet, a brand-new model designed especially for this classification purpose. Using a thorough comparative analysis on a multi-center dataset, we showed that IVD-ResNet outperforms the popular U-Net (83.31%) and V-Net (76.66%) models by achieving a state-of-the-art classification accuracy of 98.30%.

Received: August 10, 2025

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The main finding is that a purpose-built classification model such as IVD-ResNet is far more successful for the complex task of severity grading, even though generic segmentation architectures are strong for localization. The remarkable performance of this model demonstrates its durability and clinical promise, especially its near-perfect precision in differentiating between mild and severe degeneration, a known issue in automated diagnostics.

The main drawbacks of this work are the dataset's class imbalance and the intrinsic subjectivity of the Pfirrmann grading scheme employed for ground truth labels. To improve generalisability, future research will concentrate on verifying this model on bigger, more varied multi-national datasets. In addition, we intend to create a comprehensive system that combines segmentation and classification into a single, smooth process and carry out a prospective clinical study to evaluate its effect on radiologists' diagnostic accuracy and efficiency in practical situations.

In conclusion, automated spinal image analysis has advanced significantly with the IVD-ResNet system. It has great potential as a decision-support tool that can give radiologists rapid, unbiased, and quantitative evaluations of disc degeneration, which would ultimately enhance patient outcomes and diagnostic consistency in spinal medical care.

Conflicts of Interest

The authors declare no conflict of interest.

Author contributions

Abhishek Dipak Shroff is the main author, and he took on the responsibility for planning the research, designing its approach, carrying out the experiments, and writing the manuscript. Sachin Patel, Rajesh Ku. Nagar, Jayshri Sonawane and Puja Saraf helped to draft the paper by supplying direction, key supervision, and clear edits to improve it.

References

[1] Z. Wang, P. Xiao, and H. Tan, “Spinal magnetic resonance image segmentation based on U-net,” *Journal of Radiation Research and Applied Sciences*, vol. 16, no. 3, p. 100627, 2023.

[2] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein, “nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation,” *Nature Methods*, vol. 18, no. 2, pp. 203–211, 2021.

[3] A. Uslucuk and H. Öcal, “Comparative Analysis of Baseline Vnet and Unet Architectures on Pancreas Segmentation,” *Orclever Proceedings of Research and Development*, vol. 3, no. 1, pp. 146–157, 2023.

Received: August 10, 2025

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[4] J. Dolz, C. Desrosiers, and I. Ben Ayed, "IVD-Net: Intervertebral disc localization and segmentation in MRI with a multi-modal UNet," in *International Workshop and Challenge on Computational Methods and Clinical Applications for Spine Imaging*, Cham: Springer International Publishing, pp. 130–143, 2018.

[5] P. M. Adamson, V. Bhattbhatt, S. Principi, S. Beriwal, L. S. Strain, M. Offe, and P. Jordan, "Evaluation of a V-Net autosegmentation algorithm for pediatric CT scans: Performance, generalizability, and application to patient-specific CT dosimetry," *Medical Physics*, vol. 49, no. 4, pp. 2342–2354, 2022.

[6] F. Milletari, N. Navab, and S. A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, pp. 565–571, IEEE, 2016.

[7] G. Zheng, C. Chu, D. L. Belavý, B. Ibragimov, R. Korez, T. Vrtovec, and S. Li, "Evaluation and comparison of 3D intervertebral disc localization and segmentation methods for 3D T2 MR data: A grand challenge," *Medical Image Analysis*, vol. 35, pp. 327–344, 2017.

[8] M. Alruwaili and M. Mohamed, "An Integrated Deep Learning Model with EfficientNet and ResNet for Accurate Multi-Class Skin Disease Classification," *Diagnostics*, vol. 15, no. 5, p. 551, 2025.

[9] H. Liu, S. Lu, and F. Zhao, "MLP-Res-Unet: MLPs and residual blocks-based U-shaped network intervertebral disc segmentation of multi-modal MR spine images," 2024.

[10] X. Dong and G. Zheng, "Automated 3D lumbar intervertebral disc segmentation from MRI data sets," in *Computational Radiology for Orthopaedic Interventions*, pp. 25–40, 2016.

[11] J. Bradley and S. Rajendran, "Developing predictive models for early detection of intervertebral disc degeneration risk," *Healthcare Analytics*, vol. 2, p. 100054, 2022.

[12] H. Nisar, K. H. Yeap, V. Dakulagi, and M. Amin, "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.

[13] A. Shukla, S. Bhardwaj, and M. Singh, "Segmentation for Lumbar Spinal Stenosis Using Convolutional Neural Networks," *Procedia Computer Science*, vol. 218, pp. 2210–2223, 2023.

[14] S. Guinebert, E. Petit, V. Bousson, S. Bodard, N. Amoretti, and B. Kastler, "Automatic semantic segmentation and detection of vertebras and intervertebral discs by neural networks," *Computer Methods and Programs in Biomedicine Update*, vol. 2, p. 100055, 2022.

Received: August 10, 2025

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[15] T. Bassani, A. Cina, F. Galbusera, L. M. Sconfienza, D. Albano, F. Barcellona, and M. Brayda-Bruno, "Automatic classification of the vertebral endplate lesions in magnetic resonance imaging by deep learning model," *Frontiers in Surgery*, vol. 10, p. 1172313, 2023.

[16] F. Natalia, J. C. Young, N. Afriliana, H. Meidia, R. E. Yunus, and S. Sudirman, "Automated selection of mid-height intervertebral disc slice in traverse lumbar spine MRI using a combination of deep learning feature and machine learning classifier," *PLoS One*, vol. 17, no. 1, p. e0261659, 2022.

[17] Z. Chen, W. Wang, X. Chen, F. Dong, G. Cheng, L. He, and S. Zhou, "Deep learning-based quantitative morphological study of anteroposterior digital radiographs of the lumbar spine," *Quantitative Imaging in Medicine and Surgery*, vol. 14, no. 8, pp. 5385, 2023.

[18] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, DLMIA 2018, Granada, Spain, September 20, 2018, Springer International Publishing, pp. 3–11.

[19] C. Wang, Y. Guo, W. Chen, and Z. Yu, "Fully automatic intervertebral disc segmentation using multimodal 3D U-Net," in *2019 IEEE 43rd Annual Computer Software and Applications Conference (COMPSAC)*, vol. 1, pp. 730–739, IEEE, 2019.

[20] I. Ahmed, M. T. Hossain, M. Z. I. Nahid, K. S. Sanjid, M. S. S. Junayed, M. M. Uddin, and M. M. Khan, "Pioneering precision in lumbar spine MRI segmentation with advanced deep learning and data enhancement," *Machine Learning with Applications*, vol. 20, p. 100635, 2025.

[21] Y. K. Cheng, C. L. Lin, Y. C. Huang, G. S. Lin, Z. Y. Lian, and C. H. Chuang, "Accurate intervertebral disc segmentation approach based on deep learning," *Diagnostics*, vol. 14, no. 2, p. 191, 2024.

[22] R. C. Da Costa, S. De Decker, M. J. Lewis, H. Volk, and Canine Spinal Cord Injury Consortium (CANSORT-SCI), "Diagnostic imaging in intervertebral disc disease," *Frontiers in Veterinary Science*, vol. 7, p. 588338, 2020.

[23] M. F. Erkoc, H. Ulutas, and M. E. Sahin, "Intervertebral Cervical Disc Intensity (IVCDI) Detection and Classification on MRI Scans Using Deep Learning Methods," *International Journal of Imaging Systems and Technology*, vol. 34, no. 5, p. e23174, 2024.

[24] Y. K. Cheng, C. L. Lin, Y. C. Huang, J. C. Chen, T. P. Lan, Z. Y. Lian, and C. H. Chuang, "Automatic segmentation of specific intervertebral discs through a two-stage multiresunet model," *Journal of Clinical Medicine*, vol. 10, no. 20, p. 4760, 2021.

Received: August 10, 2025

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[25] Y. Li, M. Hu, J. Chen, Z. Ling, X. Zou, W. Cao, and F. Wei, "Deep learning assisted classification of T1ρ-MR based intervertebral disc degeneration phases," *Journal of Magnetic Resonance Imaging*, vol. 61, no. 3, pp. 1492–1500, 2025.

[26] C. Hou, X. Li, H. Wang, W. Zhang, F. Liu, D. Liu, and Y. Pan, "An MRI image automatic diagnosis model for lumbar disc herniation using semi-supervised learning," *Complex & Intelligent Systems*, vol. 9, no. 5, pp. 5567–5584, 2023.

[27] Z. Soydan, E. Bayramoglu, R. Karasu, I. Sayin, S. Salturk, and H. Uvet, "An automatized deep segmentation and classification model for lumbar disk degeneration and clarification of its impact on clinical decisions," *Glob Spine J.*, 2023.

[28] M. A. Sayed et al., "Automatic detection, classification, and segmentation of sagittal MR images for diagnosing prolapsed lumbar intervertebral disc," *Sci. Rep.*, vol. 15, no. 1, p. 593, 2025.

[29] H. D. Zheng et al., "Deep learning-based high-accuracy quantitation for lumbar intervertebral disc degeneration from MRI," *Nat. Commun.*, vol. 13, no. 1, p. 841, 2022.

[30] A. Wang et al., "Deep learning assisted segmentation of the lumbar intervertebral disc: a systematic review and meta-analysis," *J. Orthop. Surg. Res.*, vol. 19, no. 1, p. 496, 2024.

[31] A. Beulah, T. S. Sharmila, and V. K. Pramod, "Disc bulge diagnostic model in axial lumbar MR images using intervertebral disc descriptor (IdD)," *Multimed. Tools Appl.*, vol. 77, no. 20, pp. 27215–27230, 2018.

[32] C. Waldenberg, H. Hebelka, H. Brisby, and K. M. Lagerstrand, "MRI histogram analysis enables objective and continuous classification of intervertebral disc degeneration," *Eur. Spine J.*, vol. 27, pp. 1042–1048, 2018.

[33] I. Castro-Mateos, R. Hua, J. M. Pozo, A. Lazary, and A. F. Frangi, "Intervertebral disc classification by its degree of degeneration from T2-weighted magnetic resonance images," *Eur. Spine J.*, vol. 25, pp. 2721–2727, 2016.

[34] A. Beulah, T. S. Sharmila, and V. K. Pramod, "Degenerative disc disease diagnosis from lumbar MR images using hybrid features," *Vis. Comput.*, vol. 38, no. 8, pp. 2771–2783, 2022.

[35] Q. Pan et al., "Automatically diagnosing disk bulge and disk herniation with lumbar magnetic resonance images by using deep convolutional neural networks: method development study," *JMIR Med. Inform.*, vol. 9, no. 5, p. e14755, 2021.

[36] B. M. Abuhayi, Y. A. Bezabh, and A. M. Ayalew, "Lumbar disease classification using an involutional neural based VGG Nets (INVGG)," *IEEE Access*, 2024.

Received: August 10, 2025

Volume 38 No. 11s 2025,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[37] W. Liawrungrueang, J. B. Park, W. Cholamjiak, P. Sarasombath, and K. D. Riew, "Artificial intelligence-assisted MRI diagnosis in lumbar degenerative disc disease: a systematic review," *Glob. Spine J.*, 2024.

[38] F. Natalia, S. Sudirman, D. Ruslim, and A. Al-Kafri, "Lumbar spine MRI annotation with intervertebral disc height and Pfirrmann grade predictions," *PLoS ONE*, vol. 19, no. 5, p. e0302067, 2024.

[39] W. V. Graaf, M. L. V. Hooff, and C. F. Buckens, "SPIDER-Lumbar spine segmentation in MR images: a dataset and a public benchmark," *Boston: Zendo*, p. 10, 2023.

Received: August 10, 2025