IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 26 January 2024, accepted 14 February 2024, date of publication 20 February 2024, date of current version 26 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3367774

RESEARCH ARTICLE

Lumbar Disease Classification Using an Involutional Neural Based VGG Nets (INVGG)

BINIYAM MULUGETA ABUHAYI, YOHANNES AGEGNEHU BEZABH, AND ALEKA MELESE AYALEW

Department of Information Technology, University of Gondar, Gondar 196, Ethiopia

Corresponding author: Biniyam Mulugeta Abuhayi (biniyam.mulugeta123@gmail.com)

ABSTRACT Degenerative diseases of the lumbar spine, such as spondylolisthesis, disc degeneration, and lumbar spinal stenosis, are major contributors to global disability. Accurate classification of lumbar diseases is crucial for effective medical diagnosis. This paper introduces an innovative methodology for lumbar disease classification, addressing the limitations of traditional convolutional neural networks (CNNs). We propose a novel approach, InVGG, which combines involutional neural networks with the VGG architecture. Unlike traditional CNNs, InVGG utilizes involution kernels that are location-specific and channel-agnostic, enhancing its adaptability to varied visual patterns in medical images. Our study focuses on a four-class lumbar disease classification problem using sagittal T2 MRI images. The evaluation of InVGG is compared with traditional CNNs (VGG model) and machine learning algorithms, demonstrating superior performance in terms of accuracy, precision, recall, and AUC ROC values. InVGG achieves an impressive 96% accuracy on the testing set and 99% on the training set, showcasing its potential for accurate spinal lumbar disease classification. The reduced parameter count of InVGG compared to CNNs (VGG) makes it more resource-efficient, especially in scenarios with limited computational resources and datasets. The promising results position InVGG as a valuable tool for precise lumbar disease classification, with implications for improving patient care in resource-constrained scenarios.

INDEX TERMS Lumbar disease, involution neural network, convolution neural network, classification, machine learning.

I. INTRODUCTION

Degenerative diseases of the lumbar spine, including conditions such as spondylolisthesis, disc degeneration, and lumbar spinal stenosis, are major contributors to global disability [1]. The worldwide prevalence of degenerative lumbar diseases is substantial, affecting approximately 3.63% of the population, with varying incidences across continents. Notably, low- and middle-income countries bear a fourfold higher burden than high-income countries, emphasizing the need for innovative approaches to enhance diagnostic accuracy and treatment strategies [1]. The lumbar spine's susceptibility to diverse disorders, ranging from degenerative disc diseases to spinal stenosis, underscores the urgency for accurate identification and classification. Lumbar degenerative diseases encompass a spectrum of conditions, such as lumbar disc herniation, lumbar stenosis, and spondylosis, each presenting unique challenges [2], [3]. Surgical interventions, such as oblique lumbar interbody fusion (OLIF) and posterior lumbar interbody fusion (PLIF), offer viable options for lumbar fusion, with OLIF demonstrating advantages in radiographic outcomes and complication rates [4], [5], [6]. Medical image analysis, powered by machine learning techniques, has emerged as a potent tool for pathology classification, particularly in the realm of deep learning. Convolutional neural networks (CNNs), notably effective

The associate editor coordinating the review of this manuscript and approving it for publication was Kumaradevan Punithakumar.

in medical imaging analysis, have demonstrated superior capabilities compared to conventional methods [7], [8]. The integration of deep learning algorithms, specifically CNNs, has revolutionized medical image analysis across various domains, including neurology, cardiology, and pathology [9], [10]. Traditional diagnostic methods, although effective, have limitations in terms of subjectivity and time consumption. The advent of deep learning, especially Involutional Neural Networks (INNs), holds promise for enhancing the efficiency and precision of lumbar disease classification. Involution layers within neural network architectures have exhibited remarkable pattern-capturing abilities in medical imaging datasets, and our objective is to leverage these capabilities to extract nuanced features from lumbar spine images, contributing to the development of a robust classification system. The intersection of medical imaging and cutting-edge machine learning techniques has the potential to revolutionize lumbar diagnostics, offering insights that could significantly impact patient care. This study proposes an improved model incorporating involution neural networks with the VGG architecture, addressing the limitations of traditional convolution networks. The InvVGG net aims to overcome spatial-agnostic and channel-specific properties inherent in convolution, offering an innovative solution to visualize and interpret learned features. Contributions of this work include

- a novel approach to lumbar disease classification using involutional neural networks.
- The introduction of the InvVGG net, and the demonstration of InvNets' ability to achieve accurate classification with fewer parameters than traditional CNNs.
- The paper also evaluates the performance of InvNets on a four-class lumbar disease classification problem and underscores their potential for medical image analysis tasks, particularly in resource-constrained scenarios.

The subsequent sections of the paper are organized as follows: Section II presents literature review, section III details the materials and methods employed, including data collection and analysis procedures. Section IV presents the study findings supported by tables and figures, followed by Section V, which interprets the results, discusses their implications, and addresses study limitations. Section VI summarizes the main findings, reiterates their significance, highlights study limitations, and suggests avenues for future research.

II. LITERATURE REVIEW

This section elaborates on a literature review of the selected papers. A literature review was conducted to identify research gaps and to determine the best possible technique for lumbar spine disease detection and classification. In recent years, deep neural network-based assistant systems have been widely used in the analysis of spinal medical images. The study [11] discussed the use of computer-aided diagnosis (CAD) in the diagnosis and treatment of chronic low back pain (LBP), including lumbar degenerative diseases. It mentions that CAD systems can be based on clinical and physiological data as well as clinical images. After screening and evaluation, 57 articles were included in the review. The paper discusses the application and potential future of CAD systems in managing LBP, highlighting significant heterogeneity across studies in terms of methodology, data source, and outcomes. [12]The paper discussed the need for a computer-assisted diagnostic system for lumbar disc herniation to reduce the burden on radiologists and improve diagnosis efficiency. The model achieved 87.11% accuracy, 87.50% sensitivity, 86.72% specificity, and 0.9487 AUC. The model provides a heatmap to analyze the severity of degenerative changes. The study mentioned limitations in the computer-aided diagnosis of lumbar disc herniation and suggested further study to solve the optimization problem and design a loss function. The Study [13] discussed the challenges in computer-aided diagnosis of lumbar disc disease and proposes a Collaborative Multi-Metadata Fusion classification network (CMMF-Net) to address these challenges. The model achieved a 91.32% accuracy in lumbar spine disease classification, with the Dice coefficient for lumbar disc segmentation reaching 94.39%. The research [14] proposed detection of lumbar spinal diseases by using an enhanced convolutional neural network with MRI data collected from patients. Although it achieved an accuracy of 94%, the study was unable to discuss the type of disease it aimed to detect. The authors [15] discussed the detection of degenerative changes on MRI images of the lumbar spine with a convolutional neural network to assess the feasibility and evaluate the diagnostic performance of a CNN trained on multiple MR features of the lumbar spine. However, the study was unable to achieve better performance, especially for most disease types, with an accuracy range of 51.88%-89.13%. Machine learning techniques, including k-nearest neighbors (KNN), support vector machine (SVM), and MLP, were used for comparison purposes. SVM achieved the highest classification accuracy of 95.23%. The study [16] discussed the use of ANN to detect lumbar disc herniation disease, achieving accuracy scores of 95.23% and 91.90% using SVM classifier and ANN, respectively. [17] proposed lumbar spinal stenosis detection through semantic segmentation and delineation of magnetic resonance images of the axial view of the lumbar spine. The study explored the use of SegNet and achieved an accuracy score of 95 The study [18] proposed Automatic Detection, Classification, and Grading of Lumbar Intervertebral Disc Degeneration Using an Artificial Neural Network Model and YOLOv5 with 800 MR images. The proposed model achieved an accuracy score of 95% in the detection of intervertebral disc degeneration. However, the study only performs binary classification and lacks comparative analysis with other state-of-the-art methods. The study [19] aimed to identify the most significant physical parameters contributing to spinal abnormalities and predict spinal abnormalities based on collected physical spine data using unsupervised machine learning approaches, such as Principal Component Analysis, K-Nearest Neighbors (KNN), and Random Forest (RF). A comparison of the results between the RF classifier and KNN classifier shows that the KNN classifier performs better, with an accuracy score of 85.32%. The study shows that the specificity of the model is much lower than the sensitivity, indicating that the model performs well at correctly identifying positive instances but struggles more with correctly identifying negative instances. This imbalance could be a concern in applications where avoiding false positives is crucial and a balance between sensitivity and specificity is desirable for optimal performance. [20] proposed the development of a deep learning system in which the supplementary CNN model was trained to recorrect the key points located on corners of vertebrae based on the first CNN regression model to measure required characteristics from X-ray images of patients. The study [21] discussed the detection of lumbar Spondylolisthesis from X-ray Images using deep learning. The goal of this study was to develop a computer-aided diagnostic algorithm and evaluate the efficiency of this model in automatically detecting spondylolisthesis using lumbar X-ray images. Even though the proposed model performs better than Unet with a mean intersection over union value of 0.88 in vertebral region segmentation, the accuracy score for the classification task was poor in this case (88%).

Although numerous studies have explored the application of neural networks in lumbar disease detection and classification, our work stands out in several key aspects. Firstly, we propose a novel approach by integrating involutional neural networks with the well-established VGG architecture, addressing the inherent limitations of traditional convolutional networks. This unique combination allows our InVGG model to effectively capture intricate features within medical images, offering a spatially adaptive and channel-agnostic solution. Furthermore, most previous works dealt with detection and binary classification; however, our work focused on multiclass classification. Secondly, our study focuses specifically on a four-class lumbar disease classification problem, providing a detailed analysis of the model's performance in a clinically relevant scenario. A comparison with traditional CNNs (VGG) and machine learning algorithms demonstrates that our InVGG model consistently outperforms these approaches, demonstrating its potential for accurate and efficient lumbar disease classification. Importantly, we have significantly improved the accuracy of the lumbar disease classification. Furthermore, our methodology emphasizes the significance of dataset selection with contributions from multiple sources, ensuring a diverse representation of lumbar pathology. Meticulous image preprocessing techniques such as resizing, augmentation, and denoising contribute to the robustness and generalization capability of our model. In conclusion, our work introduces a unique and effective solution to the lumbar disease classification problem by combining state-of-the-art neural network architectures with careful consideration of dataset characteristics. The superior performance of the InVGG model, reduced parameter count, and adaptability to spatial variations make it a promising advancement in the field with potential implications for improved diagnostic accuracy and patient care.

III. MATERIALS AND METHODS

This study introduces an innovative methodology for lumbar disease classification, integrating Involutional Neural Networks (INN) and VGG architecture. The proposed model follows a systematic process to effectively classify spinal lumbar diseases from sagittal T2 MRI images. Commencing with the collection of the requisite dataset, subsequent stages involve fundamental image processing steps, encompassing resizing, augmentation, and denoising. Following preprocessing, the dataset is partitioned into training and validation sets with an 80:20 ratio. The training set facilitates the iterative adjustment of the model's parameters (weights and biases) during the training process, aiming to minimize the dis difference between predicted and actual values. Conversely, the validation set, serves to assess the model's performance and generalization to unseen data. Its purpose is to impartially evaluate the model's ability to make accurate predictions on previously unseen data. Subsequently, feature extraction is conducted using the proposed InVGG model, and the process advances to the classification phase employing various techniques, including MLP, SVM, DT, and KNN. The trained model is ultimately tested with a separate, unseen dataset. The architecture of the InVGG model is meticulously elucidated, complemented by supporting figures and tables delineating key components and layers. These findings underscore the profound significance of our proposed methodology within the domain of medical image classification, particularly in the context of lumbar disease classification. the overall system architecture is illustrated in Figure 1.

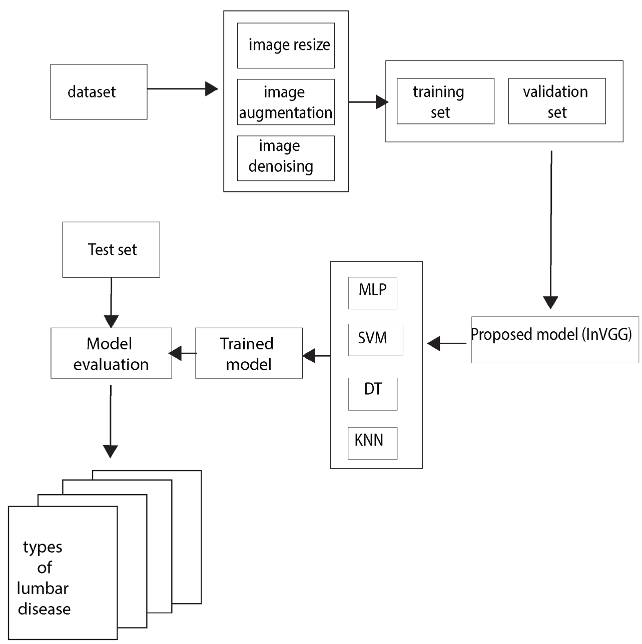


FIGURE 1. Proposed system architecture.

A. DATASET AND IMAGE PREPROCESSING

The dataset was obtained from three sources the first one is an open-source dataset Named Lumbar spine MRI dataset, this dataset contains an anonymized clinical MRI study or set of scans of 515 patients with symptomatic back pain. Each patient data can have one or more MRI studies associated with it. Each study contained slices, individual images taken from sagittal or axial views of the lowest three vertebrae and the lowest three IVDs. The other one is obtained from an open access dataset from the study by Sudirman et al. the final dataset was collected from university of Gondar specialized referral hospital with 107 patients with each patient having multiple MRI scans in sagittal axial view. Sagittal T2 MRI images were obtained from an open-access lumbar spine MRI dataset from Sudirman et al., which contains an anonymized clinical MRI study. We affirm that all procedures performed in this study were in accordance with ethical standards and comply with the 1964 Helsinki Declaration. Patient identification data were not collected. The dataset was partitioned into training, validation, and testing sets utilizing the 'train test split' function from Scikit-Learn, enabling a comprehensive evaluation of the model's performance across distinct data subsets to ensure its robustness. To maintain consistency in input dimensions, all images were resized to a standardized 224 $\times$ 224 pixel format using the target size parameter in the Keras ImageDataGenerator. The RGB color mode was retained for all images to align with the prevalent format in deep learning tasks. For label encoding, categorical labels were transformed into numerical format through one-hot encoding, facilitating the utilization of categorical cross-entropy loss during model training. To bolster the model's generalization capabilities and enhance robustness, data augmentation techniques were implemented during training. These techniques, seamlessly integrated into the data pipeline using the Keras ImageDataGenerator, encompassed random rotations, flips, zoom, and brightness adjustments. The incorporation of data augmentation artificially heightened the diversity of the training data, thereby fostering more effective learning by the model. Basics of the lumbar spine and list of sample images are shown in Figure 2 and Figure 3.

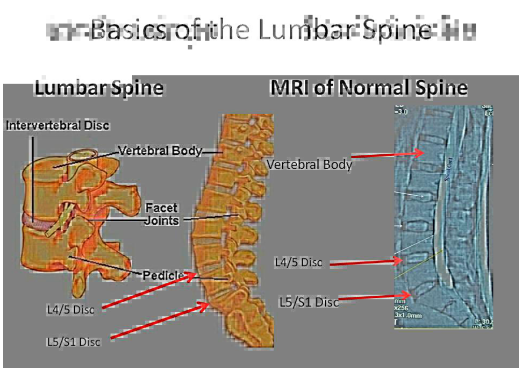


FIGURE 2. Basic structure of lumbar spine.

FIGURE 3. Sagittal T2 MRI images of lumbar diseases.

B. PREPROCESSING

1) IMAGE RESIZING

In order to optimize the efficiency of the proposed model, the extracted images are resized to a standard size of 64x64 RGB. This process of image resizing plays a crucial role in image processing, as it minimizes the image file size, facilitating easier storage, transfer, and manipulation.

2) IMAGE DENOISING

Image denoising is a fundamental pre-processing step in image processing that aims to remove noise from noisy images and recover high-quality clean images. Various algorithms and techniques have been developed for image denoising, including filters such as Median, Mean, Bayesian, Guide, Gaussian, and collaborative filters, as well as artificial neural networks (ANNs), convolutional neural networks (CNNs), and fuzzy algorithms [22]. These techniques aim to preserve detail information while reducing noise and enhancing image quality. Parameters such as Peak Signal-to-Noise Ratio (PSNR), Mean Square Error (MSE), and Structural Similarity Index Measure (SSIM) are commonly used to evaluate the performance of image denoising techniques [23]. Image denoising techniques can also be used for image enhancement, improving the accuracy and quality of images in different forms [24].

Adaptive median filter (AMF) is a non-linear image filtering technique that has been widely used in various image processing applications. AMF has gained popularity due to its ability to effectively remove salt-and-pepper noise while preserving the edges and fine details of the image. One of the main advantages of AMF is its ability to adapt to different levels of noise in the image. Unlike traditional median filter, AMF adjusts the size of the window used for filtering based on the local noise level, making it an effective filter for images with varying levels of noise. For instance, in regions with high noise density, a larger window size is used to ensure that the median value represents the majority of the pixels in the region. In contrast, in regions with low noise density, a smaller window size is used to preserve the image details.

AMF has also been shown to be effective in preserving edges and fine details in the image. Unlike linear filters such as mean and median filters, AMF does not blur the edges and fine details of the image, making it an attractive filter for applications that require edge preservation. In addition, AMF has been compared with other non-linear filters such as the morphological filter, fuzzy filter, and the adaptive filter, and has been shown to outperform these filters in terms of noise reduction and edge preservation.

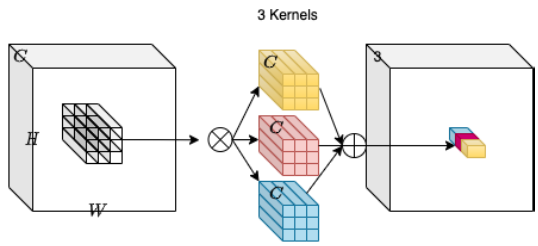


FIGURE 4. Convolution neural network architecture.

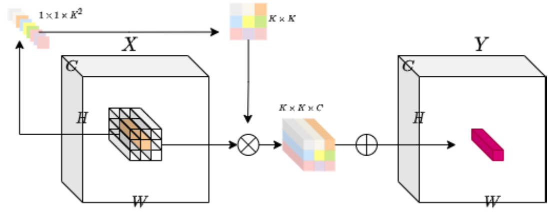


FIGURE 5. Involution neural network architecture.

C. FEATURE EXTRACTION

In the present work, we present an innovative feature extraction technique for lumbar disease classification, utilizing the InVGG network. Our approach involves implementing an involutional neural network, integrating the widely recognized VGG architecture in a novel manner to extract high-level features. The architecture, detailed in Table 1 and illustrated in Figure 6, initiates with an input MRI image. It then progresses through two blocks of double involutional networks with a single pooling layer, followed by three blocks of triple involution neural networks with a single pooling layer. For optimization, we employ the ReLu function. The primary objective is to achieve an operation that is both location-specific and channel-agnostic. Implementing these specific properties poses a challenge, as a fixed number of involution kernels for each spatial position may not effectively process variable-resolution input tensors. To overcome this limitation, we propose a solution by generating each kernel conditioned on specific spatial positions, allowing us to seamlessly process variable-resolution input tensors. The intuitive diagram below (see Figure 5) provides insight into this kernel generation method. In the pivotal feature extraction step of our methodology, we utilize the InVGG network to meticulously capture a diverse range of lumbar spine image features. These features encompass various spatial intricacies, such as the nuanced arrangement of pixels, discernment of spatial patterns, and the identification of contours within the images. The channel-agnostic property of the involutional layer allows us to extract features irrespective of specific color channels. This characteristic is particularly advantageous in medical imaging, where the emphasis often lies on the spatial and structural characteristics of the images rather than color nuances. A key feature extraction process lies in the spatial adaptability introduced by the involutional layer. This functionality enables our model to dynamically focus on specific spatial locations within the lumbar spine images, enhancing its discernment of intricate details and variations that might hold crucial diagnostic information. By leveraging the VGG architecture in conjunction with involutional layers, our feature extraction process extends to the capture of high-level features. These features correspond to complex patterns and structures inherent in lumbar spine images, significantly contributing to the model's ability to discriminate between different classes of lumbar diseases. Furthermore, the InVGG network inherently captures textural and structural features present in medical images. This includes the identification of variations in tissue textures, the spatial arrangement of vertebrae, and any structural abnormalities indicative of lumbar diseases. By incorporating this diverse set of features, our methodology ensures a holistic understanding of the intricate patterns embedded in medical images, ultimately leading to enhanced accuracy in the classification of lumbar diseases. In terms of model size, the InVGG network boasts a total parameter count of 18,952, while the VGG16 architecture, in comparison, has a significantly larger model size, with a total parameter count of 13.8 million. This indicates that the InVGG network is considerably smaller in terms of parameter size compared to the VGG16 model. Despite its smaller size, our methodology demonstrates effective feature extraction by leveraging the unique characteristics of the involutional layer in conjunction with the VGG architecture. This efficient parameter utilization contributes to the model's ability to capture intricate lumbar spine image features, leading to improved accuracy in the classification of lumbar diseases.

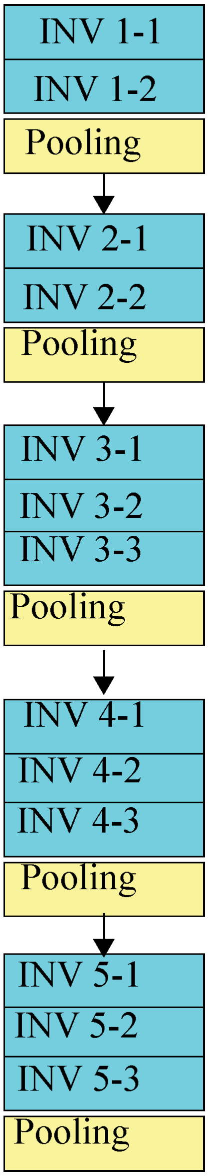


FIGURE 6. Proposed system model.

D. EVALUATION TECHNIQUES

Precise assessment methodologies are crucial for validating the classification efficacy of a system. In this investigation, various evaluation metrics were employed to gauge the precision and efficiency of our proposed model for lumbar disease classification. The initial metric employed was accuracy, which quantifies the proportion of accurately classified images within the test set. Accuracy serves as a widely accepted evaluation metric for classification tasks, offering a prompt and intuitive means of gauging system performance. In addition to accuracy, we conducted a thorough evaluation of precision, recall, and the F1 score by utilizing the confusion matrix. This matrix delineates the predicted and actual class labels, enabling a more intricate analysis of the classification performance. The integration of this combination of evaluation metrics enhances the comprehensive assessment of the proposed lumbar disease classification model. Accuracy: Accuracy is the proportion of correct predictions made by the model out of all predictions made. It measures how well the model classifies the samples. The formula for accuracy is:

$$Accuracy = \frac{TruePositives + TrueNegatives}{TotalSamples}$$

Recall (Sensitivity): Recall is the fraction of positive instances that are correctly identified. It can be represented as:

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

Precision:

Precision is the fraction of positive predictions that are actually correct. It can be represented as:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

where True Positives (TP): The number of instances that were correctly classified as positive by the model. True Negatives (TN): The number of instances that were correctly classified as negative by the model. False Positives (FP): The number of instances that were incorrectly classified as positive by the model. False Negatives (FN): The number of instances that were incorrectly classified as negative by the model. Total Samples: The total number of instances in the dataset.

AUC:

Area Under the Curve (AUC) is a commonly used performance metric machine learning classification problems, which evaluates the overall performance of a classifier. In binary classification, a classifier outputs a predicted probability for each sample to belong to one of two classes, positive or negative. The AUC is the area under the Receiver Operating Characteristic (ROC) curve, which plots the True Positive Rate (TPR) against the False Positive Rate (FPR) for different classification thresholds.

The equation for the ROC curve is:

$$TPR = \frac{TP}{TP + FN}$$
$$FPR = \frac{FP}{FP + TN}$$

where TP, TN, FP, and FN are the number of True Positives, True Negatives, False Positives, and False Negatives, respectively.

The equation for the AUC can be written as:

$$AUC = \int_{-\infty}^{\infty} [\frac{TP}{TP + FN} - \frac{FP}{FP + TN}] ds$$

E. VGG

VGG is a famous model used for various applications such as detecting plant diseases in agriculture [25], detecting lung cancer [26], and classifying breast cancer severities [27]. It is a convolutional neural network (CNN) architecture that has been shown to provide high accuracy in image classification tasks. The VGG-16 model, in particular, has been widely used and has proven to be effective in detecting and diagnosing plant diseases [28], lung cancer, and breast cancer. It uses Rectified Linear Units (ReLU) activation function and is integrated with other layers such as flatten, normalization, dense, and drop-out layers to improve its performance. The VGG-16 model is often used in combination with transfer learning and deep learning techniques to enhance its efficiency and accuracy in classifying images.

F. CLASSIFICATION

To classify lumbar disease, we employed a combination of Multi Layer Perceptron (MLP), traditional machine learning classifiers, and ensemble techniques on top of INVGG. The MLP architecture consisted of two dense layers of 64, and 128 units, followed by a dropout layer of 35%. We also utilized the classical machine learning classifiers including K-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Decision Trees (DT). Finally, we applied the stacking ensemble technique to ensemble the KNN, SVM, and DT classifiers. **MLP:** A dense layer, also known as a fully connected layer, is a type of layer in a neural network architecture. It is a fundamental building block of many neural network models, including Multilayer Perceptrons (MLPs). In a dense layer, each node (neuron) in the layer is connected to every node in the previous layer, forming a dense or fully connected network of neurons. Multilayer perceptron (MLP) is a type of neural network that is commonly used in image classification tasks. MLPs have been compared to convolutional neural networks (CNNs) in terms of their performance and data scale sensitivity for image classification. Studies have shown that MLPs can achieve high accuracies in image classification when optimized with proper activation functions, dropout layers, neurons, and hidden layers [29]. MLPs have been used on datasets such as MNIST, where they achieved accuracies of 97.49%, 97.98%, and 98.24% when trained on different data volumes [29]. MLPs have also been compared to CNNs in terms of their performance and data sensitivity analysis for processing handwriting images. MLPs have shown to be effective in image classification tasks and provide valuable insights for future research in this area [30]. **SVM:** Support Vector Machine (SVM) is a popular machine learning method used for classification problems. It is a linear classifier that aims to maximize the margin between classes by drawing a decision boundary. SVM uses a supervised learning method and is widely used in multi classification problems. The decision boundary separates the classes and is the optimum boundary with maximum margin. The data points closest to the boundary are called support vectors. SVM is known for its solid theoretical foundation and promising performance on different classification problems. However, it can be computationally expensive for large-scale training sets [31], [32]. Support Vector Machines (SVM) are a popular type of machine learning algorithm that can be used for classification and regression tasks. SVM works by finding the best possible hyperplane to separate the classes of data, maximizing the margin between them. This method is effective even when the data is not linearly separable by mapping the data into higher dimensions, where the classes are linearly separable. SVMs have been widely used in various fields, including bioinformatics, chemometrics, and image classification. They have shown to have high accuracy in classification tasks, outperforming other classifiers such as k-Nearest Neighbor (kNN) and Random Forest (RF). SVMs are particularly effective when the number of features is large compared to the number of samples and when dealing with noisy data. Despite the popularity of SVMs, there are still challenges and limitations in its application, including selecting the appropriate kernel function and dealing with imbalanced datasets [33], [34], [35]. We have used support vector machine with a kernel function of Radial Basis Function (RBF). We have applied different numbers as a parameter. For regularization parameter C we used 10 numbers by generating with NumPy log space of 2 as a start number, and 10 as the end. For gamma value we also used 10 numbers generated using NumPy log space of 9 as a start and 3 as the end. To get the best possible combination of parameters, all the given parameters are evaluated using grid search cv. **DT:** A decision tree classifier is a popular data modeling technique used in supervised learning problems. It is a top-down recursive divide-and-conquer algorithm that divides datasets until all instances of a sub dataset belong to the same class value. Decision trees have several advantages, including ease of implementation and understanding, little prior knowledge required, and the ability to handle non linear relationships. They can be used for both classification and regression tasks. Traditional decision tree methods use information gain, gain ratio, and gini values to select the best node. Decision trees can also be scalable to handle big data. Various techniques, such as ensemble classifiers and vertical partitioning, have been proposed to improve decision tree performance [36], [37]. **KNN:** K Nearest Neighbors (KNN) is a simple yet effective machine learning algorithm used for classification and regression analysis. KNN works by finding the k closest training data points to a new test data point and classifying the test data point based on the most common class among its k nearest neighbors. In the context of image classification, KNN can be used as a classifier to predict the class labels of images based on their extracted features. KNN has been shown to be effective in dealing with high-dimensional feature spaces and can handle both linearly and nonlinearly separable data [38]. The KNN algorithm determines the class of an instance by looking at the class labels of the k-nearest neighbors

TABLE 1. Parameter configuration of proposed model.

Layer (type)	Output Shape	Param
input (InputLayer)	(None, 224, 224, 3)	0
inv_1_1 (Involution)	(None, 224, 224, 3)	26
inv_1_2 (Involution)	(None, 224, 224, 3)	26
max_pooling2d(MaxPooling2d)	(None, 112, 112, 3)	0
inv_2_1 (Involution)	(None, 112, 112, 3)	26
inv_2_2 (Involution)	(None, 112, 112, 3)	26
max_pooling2d(MaxPooling2d)	(None, 56, 56, 3)	0
inv_3_1 (Involution)	(None, 56, 56, 3)	60
inv_3_2 (Involution)	(None, 56, 56, 3)	60
inv_3_3 (Involution)	(None, 56, 56, 3)	60
max_pooling2d(MaxPooling2d)	(None, 28, 28, 3)	0
inv_4_1 (Involution)	(None, 28, 28, 3)	60
inv_4_2 (Involution)	(None, 28, 28, 3)	60
inv_4_3 (Involution)	(None, 28, 28, 3)	60
max_pooling2d(MaxPooling2d)	(None, 14, 14, 3)	0
inv_5_1 (Involution)	(None, 14, 14, 3)	60
inv_5_2 (Involution)	(None, 14, 14, 3)	60
inv_5_3 (Involution)	(None, 14, 14, 3)	60
max_pooling2d(MaxPooling2d)	(None, 7, 7, 3)	0
flatten (Flatten)	(None, 147)	0
dense_1 (Dense)	(None, 64)	9472
dense_2 (Dense)	(None, 128)	8320
dense_3 (Dense)	(None, 4)	516
Total params:	—	18,952
Trainable params:	—	18,890
Non-trainable params:	—	62

in the training data. The KNN classifier is known for its ability to handle high-dimensional datasets and its simplicity in implementation. However, the performance of the KNN classifier can be sensitive to the choice of k and may suffer from the curse of dimensionality. Nonetheless, the KNN classifier remains a popular choice due to its simplicity and ease of implementation [30] [39]. In our research, we conducted experiments using various values for 'k'. We explored a range from 1 to 40 to determine the optimal value of 'k'.

IV. RESULTS

A. RESULTS BEFORE AND AFTER DATA AUGMENTATION

Data augmentation offers several advantages in the realm of machine learning. It enables the generation of new datasets from existing ones, thereby augmenting dataset size and enhancing model performance. Techniques such as image rotation, cropping, flipping, and the introduction of noise to audio signals contribute to the efficacy of data augmentation, particularly beneficial when dealing with limited or unbalanced original datasets [31] [40]. This approach proves instrumental in mitigating overfitting by expanding the pool of training data incorporated into the model. Furthermore, data augmentation strategies find application in computer vision and natural language processing models to address issues of data scarcity and lack of diversity. In summary, data augmentation contributes to the overall improvement of robustness, accuracy, and generalization in machine learning models. In our study, we conducted experiments on our proposed model prior to applying data augmentation to assess its impact on involutional neural networks. This investigation is motivated by the well-established findings in convolutional neural networks, where studies have demonstrated that the application of data augmentation, such as flipping, zooming, rotation, and others, influences not only the generalization ability of models but also enhances the dataset size, especially when dealing with a limited number of datasets [41]. The obtained results are presented in Table 2 and Table 3. As seen in the table, the value of DT has slightly decreased after applying data augmentation. This decrease is attributed to the relationship with the dataset size, which depends on the dimensionality of the feature space. In high-dimensional spaces, Decision Trees may struggle to find optimal splits. Another reason is that hyperparameters of the Decision Tree, such as maximum depth, minimum samples per leaf, and others, can significantly impact its performance. Proper tuning of these hyperparameters is essential, especially when dealing with larger datasets.

TABLE 2. Results of proposed study before data augmentation.

classifier	Accuracy (%)	Precision (%)	Recall (%)
MLP	89.23	88.98	88.64
SVM	84.62	83.61	82.78
KNN	79.72	78.38	79.02
DT	81.36	81.12	79.33

TABLE 3. Results of proposed study after data augmentation.

classifier	Accuracy (%)	Precision (%)	Recall (%)
MLP	93.73	93.230	92.85
SVM	84.91	83.75	82.59
KNN	83.64	82.40	83.11
DT	80.09	79.24	78.93

B. RESULTS AFTER APPLYING NOISE REMOVAL TECHNIQUE

we conducted experiments on our proposed model using different image denoising techniques. The model exhibited a training accuracy of 99.06% and a testing accuracy of 95.11%. In this investigation, we utilized three distinct image filtering techniques Gaussian filter, Median filter, and Adaptive median filter to preprocess sagittal T2 MRI images of lumbar diseases. The primary objective was to assess the efficacy of these techniques in minimizing image noise and enhancing the accuracy of classification outcomes. The Gaussian filter, known for its linear smoothing capabilities, is commonly employed in image preprocessing to effectively eliminate Gaussian noise. The Median filter, a non-linear approach, substitutes each pixel with the median value of neighboring pixels, thereby reducing noise while preserving edges. The Adaptive median filter, a variant of the Median filter, dynamically adjusts the filter window size based on image content, rendering it more proficient in noise reduction, especially in regions with varying pixel intensities (see Table 4).

C. RESULTS OF HYBRID MODEL

After identifying the optimal parameters for SVM, DT, and KNN classifiers using the grid search CV algorithm,

TABLE 4. Results after applying noise removal technique.

classifier	Accuracy (%)
Gaussian Filter	94.03
Median Filter	93.53
Adaptive Filter	95.11

TABLE 5. Results of hybrid models.

classifier	Accuracy (%)	Precision (%)	Recall (%)
SVM	89.71	88.68	87.06
KNN	85.98	86.65	85.22
DT	86.30	85.37	87.01

TABLE 6. Results of proposed study.

classifier	Accuracy (%)	Precision (%)	Recall (%)
MLP	96.71	95.89	96.11
SVM	89.71	88.68	87.06
KNN	85.98	86.65	85.22
DT	86.30	85.37	87.01

we proceeded with the experiment. For the SVM classifier, we utilized 'C': 46415.888336127726 and 'gamma': 0.1. For the DT classifier, the chosen parameters were 'max leaf nodes': 54 and 'min samples split': 6. Finally, for the KNN classifier, 'n_neighbors': 1 was employed. The obtained results are presented in Table 5, and furthermore, the outcomes are illustrated using AUC-ROC curves as shown in Figure 7. Upon examining the results, it is evident that the SVM classifier exhibited superior performance with an accuracy of 89.71%. Subsequently, the DT classifier achieved an accuracy score of 86.30%, and the KNN classifier demonstrated a comparable performance to the decision tree classifier, also yielding an accuracy score of 85.98% on testing set.

D. RESULTS OF PROPOSED MODEL

Following the selection of hyperparameters and the extraction of features from Sagittal T2 MRI images, we utilized them to train our proposed model employing an MLP classifier and softmax function. Subsequently, through 80 epochs of training, we observed training, validation, and testing accuracies of 99.67%, 98.89%, and 96.71% respectively. These results are visually represented in the subsequent figure. Further insight into the outcomes is provided through the application of a confusion matrix, depicted in the accompanying figure. Notably, from the 686 images utilized for testing, only 8 images from two classes were misclassified, as indicated by the confusion matrix. The detailed results are presented in Table 6 and Figure 8.

E. RESULTS OF VGG16 MODEL

we have made a comparative study to identify if the proposed model was better than its corresponding VGG architecture with the original parameter values using Convolutional neural networks. the results are displayed in Figure 9. As we can see from the training accuracy curve even though

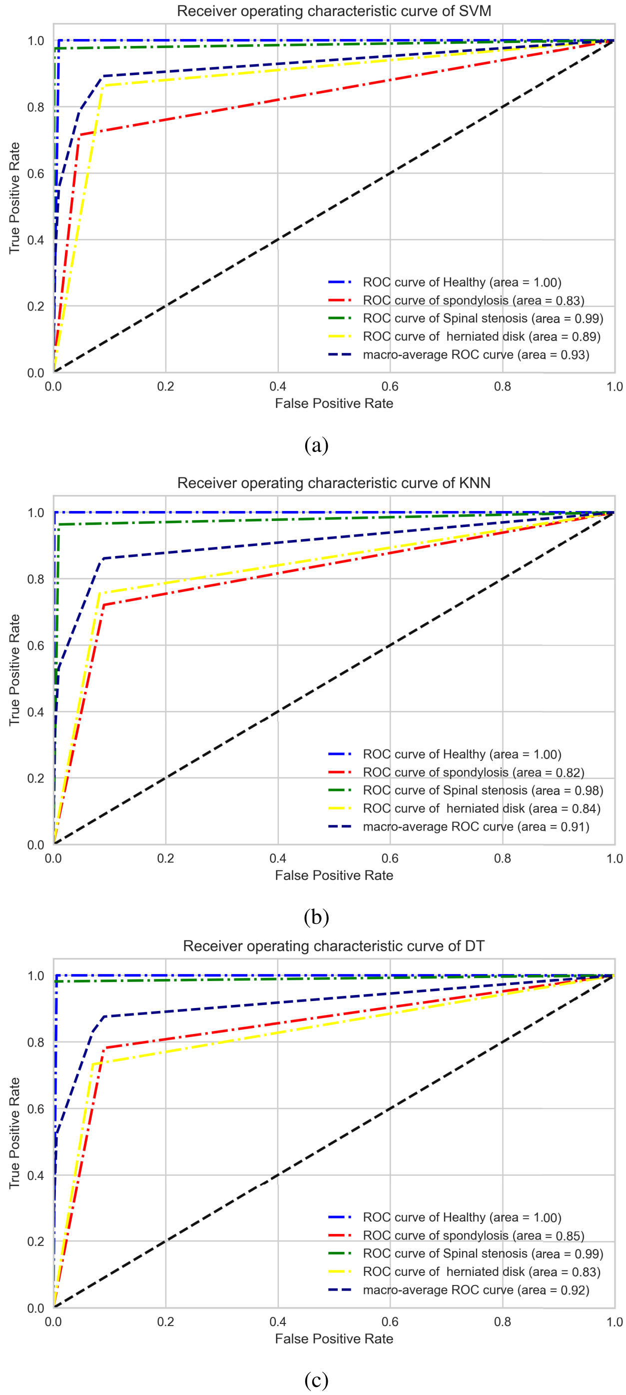
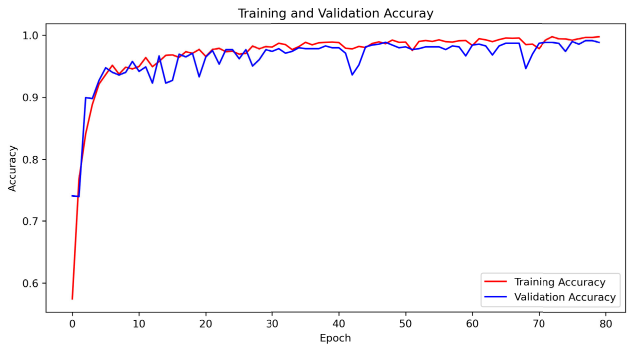


FIGURE 7. ROC curves for three different classification algorithms (a) SVM, (b) KNN, and (c) DT. The average area under the curve (AUC) values for each algorithm are 0.93, 0.91, and 0.92, respectively.

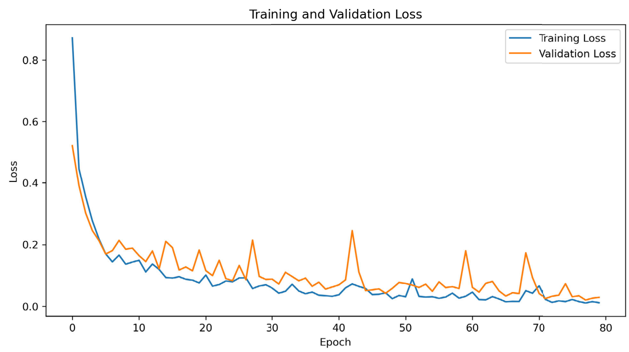
the model gained good accuracy on the training data it suffered from overfitting problem. whereas the proposed model outperformed the original VGG 16 model with a smaller number of parameters.

V. DISCUSSION

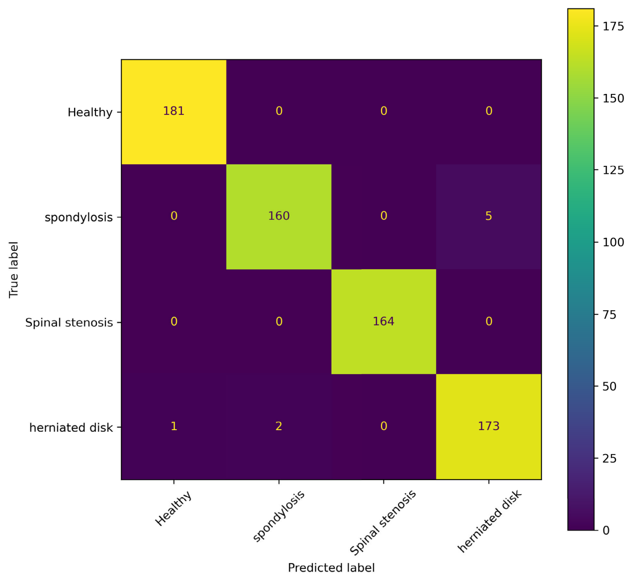
This paper introduces an innovative approach to lumbar disease classification through the integration of Involutional



(a)



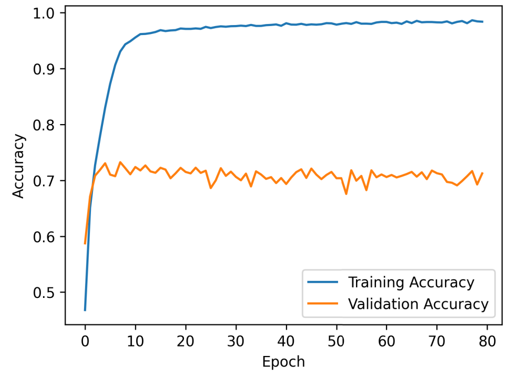
(b)



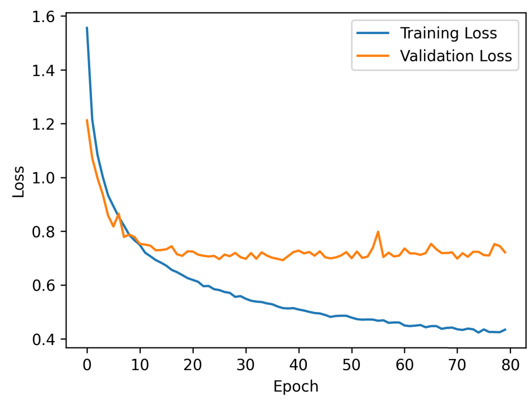
(c)

FIGURE 8. **Learning curve of proposed model (a) training and validation accuracy, (b) training and validation loss, and (c) confusion matrix.**

Neural Networks (InvNets) with the VGG architecture. Our research addresses a four-class lumbar disease classification problem, with the objective of distinguishing between various lumbar disease types based on MRI data. A comparative analysis of InvNets and traditional Machine Learning classifiers is presented, employing diverse evaluation parameters. Our results indicate that InVGG surpasses conventional machine learning classifiers in terms of accuracy, precision, recall, and



(a)



(b)

FIGURE 9. **Learning curve of VGG16 model with convolutional neural network (a) training and validation accuracy, and (b) training and validation loss.**

other values. This enhanced performance is attributed to the distinctive features of InvNets. In contrast to conventional CNNs that utilize spatial-agnostic and channel-specific convolution kernels, InvNets employ location-specific and channel-agnostic involution kernels. This design enables the network to adapt to varied visual patterns across spatial locations, augmenting its capacity to capture intricate features in medical images. The evaluation outcomes of InVGG demonstrate an impressive accuracy rate of 96%. This noteworthy accuracy, coupled with a significantly reduced parameter count, underscores the efficacy of InvNets for tasks in medical image analysis. Particularly in resource-constrained environments, InvNets emerge as a promising solution for accurate lumbar disease classification. Comparative analyses with alternative machine learning methods underscore the superiority of our proposed model. Decision Trees, KNN, and SVM methods exhibit lower accuracy rates and potential overfitting concerns. In contrast, our hierarchical approach, amalgamating VGG architecture with InvNets, achieves a balanced trade-off between accuracy and computational efficiency. Nevertheless, it is crucial to acknowledge certain limitations in our study. The model's performance was assessed using standard metrics and benchmarks, warranting further validation on larger and more diverse datasets from various sources. Additionally, our model exclusively considered T2 sagittal MRI images, prompting exploration of its effectiveness across different modalities. Despite these considerations, the proposed model's robust performance and reduced computational requirements position it as a promising candidate for practical applications in the medical domain. Subsequent research endeavors could delve into refining the model architecture and exploring its potential in diverse medical imaging tasks.

VI. CONCLUSION

The present study introduces a pioneering methodology utilizing involutional neural networks with VGG architecture, denoted as InVGG. InVGG mitigates the computational load associated with CNNs by employing an involution kernel, which is location-specific and channel-agnostic. This spatial adaptability enhances the model's capability to capture intricate features present in medical images, particularly in the context of a four-class lumbar disease classification problem using MRI data. Comparative analysis with traditional machine learning algorithms demonstrates the remarkable efficiency of InVGG, requiring significantly fewer parameters while achieving competitive accuracy rates. Performance evaluations encompass a range of standard metrics, including accuracy, precision, recall, and ROC values. The results consistently show that InVGG outperforms conventional machine learning algorithms and other pretrained models in the classification task, achieving an impressive 96% accuracy rate. The paper further conducts a detailed analysis of the model's performance through metrics such as confusion matrices and ROC curves. These evaluations provide comprehensive insights into the model's ability to accurately identify different lumbar disease classes, as well as its strengths and limitations in doing so. Overall, InVGG demonstrates substantial potential for precise lumbar disease classification.

Grid search cv is a method used to determine the optimal model parameters in machine learning algorithms. It helps in finding the best combination of hyperparameters for a given model by exhaustively searching through a specified parameter grid. This technique is used to improve the performance and accuracy of machine learning models. Grid search cv involves evaluating the model's performance on different combinations of hyperparameters and selecting the one that gives the best results.

DECLARATIONS

- Conflict of interest
  On behalf of all authors, the corresponding author states that there is no conflict of interest.
- Consent for publication
  The authors declare that they are in agreement with this submission and for the paper to be published if accepted.
- Funding
  The authors have no funding to report.
- Data Availability
  The data are available in the following link: https://data.mendeley.com/datasets/k57fr854j2/2
- Ethics Approval
  This retrospective study uses previously collected data for analysis in accordance with ethical standards and privacy regulations.

REFERENCES

[1] V. M. Ravindra, S. S. Senglaub, A. Rattani, M. C. Dewan, R. Härtl, E. Bisson, K. B. Park, and M. G. Shrime, "Degenerative lumbar spine disease: Estimating global incidence and worldwide volume," *Global Spine J.*, vol. 8, no. 8, pp. 784–794, Dec. 2018.

[2] R. Battisti, C. E. A. S. D. Oliveira, and R. D. R. Pratali, "Impact of the grade of degenerative lumbar disease on the occurrence of spinal deformity," *Coluna/Columna*, vol. 20, no. 1, pp. 30–33, Mar. 2021.

[3] X.-Y. Li, Y. Wang, K. Yang, C.-X. Liu, W.-G. Zhu, C. Kong, and S.-B. Lu, "Four types of global spine sagittal alignment and compensation mechanism in adult patients with lumbar degenerative disease," *J. Orthopaedic Sci.*, vol. 28, no. 1, pp. 98–104, Jan. 2023.

[4] B. R. Subach, T. C. Schuler, M. R. McLaughlin, P. J. Slosar, C. H. Comey, and N. M. Thomas, "Discogenic lumbar disease," *Adv. Orthopedics*, vol. 2012, p. 1, Jun. 2012.

[5] M. Lacroix, C. Nguyen, R. Burns, A. Laporte, F. Rannou, and A. Feydy, "Degenerative lumbar spine disease: Imaging and biomechanics," in *Seminars in Musculoskeletal Radiology*, vol. 26. New York, NY, USA: Thieme Medical Publishers, 2022, pp. 424–438.

[6] G. H. Kang, D. Son, J. S. Lee, S. H. Lee, S. H. Bae, S. W. Lee, and G. S. Song, "Surgical treatment for degenerative lumbar disease with neurologic deficits: Comparison between oblique lumbar interbody fusion and posterior lumbar interbody fusion," *Korean J. Neurotrauma*, vol. 18, no. 2, p. 277, 2022.

[7] J. Zhang, F. Zhao, F.-L. Wang, Y.-F. Yang, C. Zhang, Y. Cao, Y.-L. Wang, X.-J. Shi, Y. Wan, M. Zhang, M.-Q. Liu, C.-G. Zuo, and H.-Q. Wang, "Identification of lumbar disc disease hallmarks: A large cross-sectional study," *SpringerPlus*, vol. 5, no. 1, pp. 1–8, Dec. 2016.

[8] M. Hassouna, M. Al-Antary, M. Saleh, and N. B. Al Barghuthi, "Applications of deep learning in medical imaging: A brief review," in *Proc. Adv. Sci. Eng. Technol. Int. Conf. (ASET)*, Feb. 2023, pp. 1–4.

[9] K. Suzuki, "Overview of deep learning in medical imaging," *Radiolog. Phys. Technol.*, vol. 10, no. 3, pp. 257–273, Sep. 2017.

[10] S. Maitra, R. Ghosh, and K. Ghosh, "Applications of deep learning in medical imaging," in *Handbook of Deep Learning Applications*. New York, NY, USA: Springer, 2019, pp. 111–127.

[11] F. D'Antoni, F. Russo, L. Ambrosio, L. Bacco, L. Vollero, G. Vadalà, M. Merone, R. Papalia, and V. Denaro, "Artificial intelligence and computer aided diagnosis in chronic low back pain: A systematic review," *Int. J. Environ. Res. Public Health*, vol. 19, no. 10, p. 5971, May 2022.

[12] 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 Intell. Syst.*, vol. 9, no. 5, pp. 5567–5584, Oct. 2023.

[13] S. Lu, J. Liu, X. Wang, and Y. Zhou, "Collaborative multi-metadata fusion to improve the classification of lumbar disc herniation," *IEEE Trans. Med. Imag.*, vol. 42, no. 12, pp. 3590–3601, Jul. 2023.

[14] Ruchi, D. Singh, J. Singla, M. K. I. Rahmani, S. Ahmad, M. Ur Rehman, S. Jha, D. Prashar, and J. Nazeer, "Lumbar spine disease detection: Enhanced CNN model with improved classification accuracy," *IEEE Access*, vol. 11, pp. 141889–141901, 2023.

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

[16] E. Ebrahimzadeh, F. Fayaz, F. Ahmadi, and M. Nikravan, "A machine learning-based method in order to diagnose lumbar disc herniation disease by MR image processing," *MedLife Open Access*, vol. 1, no. 1, p. 1, 2018.

[17] A. S. Al-Kafri, S. Sudirman, A. Hussain, D. Al-Jumeily, F. Natalia, H. Meidia, N. Afriliana, W. Al-Rashdan, M. Bashtawi, and M. Al-Jumaily, "Boundary delineation of MRI images for lumbar spinal stenosis detection through semantic segmentation using deep neural networks," *IEEE Access*, vol. 7, pp. 43487–43501, 2019.

[18] W. Liawrungrueang, P. Kim, V. Kotheeranurak, K. Jitpakdee, and P. Sarasombath, "Automatic detection, classification, and grading of lumbar intervertebral disc degeneration using an artificial neural network model," *Diagnostics*, vol. 13, no. 4, p. 663, Feb. 2023.

[19] A. A. Abdullah, A. Yaakob, and Z. Ibrahim, "Prediction of spinal abnormalities using machine learning techniques," in *Proc. Int. Conf. Comput. Approach Smart Syst. Design Appl. (ICASSDA)*, Aug. 2018, pp. 1–6.

[20] T. P. Nguyen, D.-S. Chae, S.-J. Park, K.-Y. Kang, and J. Yoon, "Deep learning system for meyerding classification and segmental motion measurement in diagnosis of lumbar spondylolisthesis," *Biomed. Signal Process. Control*, vol. 65, Mar. 2021, Art. no. 102371.

[21] G. M. Trinh, H.-C. Shao, K. L.-C. Hsieh, C.-Y. Lee, H.-W. Liu, C.-W. Lai, S.-Y. Chou, P.-I. Tsai, K.-J. Chen, F.-C. Chang, M.-H. Wu, and T.-J. Huang, "Detection of lumbar spondylolisthesis from X-ray images using deep learning network," *J. Clin. Med.*, vol. 11, no. 18, p. 5450, Sep. 2022.

[22] D. B. Kuttan, S. Kaur, B. Goyal, and A. Dogra, "Image denoising: Pre-processing for enhanced subsequent CAD analysis," in *Proc. 2nd Int. Conf. Smart Electron. Commun. (ICOSEC)*, Oct. 2021, pp. 1406–1411.

[23] S. E. Kuzhali and D. S. Suresh, "Automated image denoising model: Contribution towards optimized internal and external basis," *Int. J. Image Graph.*, vol. 21, no. 4, Oct. 2021, Art. no. 2150051.

[24] R. S. Jebur, C. S. Der, and D. A. Hammood, "A review and taxonomy of image denoising techniques," in *Proc. 6th Int. Conf. Interact. Digit. Media (ICIDM)*, Dec. 2020, pp. 1–6.

[25] S. Balaji, B. N. Shivacharan, A. Nissar, and S. V. Bhaskar, "Plant infirmity detection using vgg -16 convolutional neural network," in *Proc. Int. Conf. Comput. Intell. Sustain. Eng. Solutions (CISES)*, Apr. 2023, pp. 485–491.

[26] P. H. S. Prasad, N. M. V. S. Daswanth, C. V. S. P. Kumar, N. Yeeramally, V. M. Mohan, and T. Satish, "Detection of lung cancer using VGG-16," in *Proc. 7th Int. Conf. Comput. Methodologies Commun. (ICCMC)*, Feb. 2023, pp. 860–865.

[27] N. H. Hider, A. S. B. M. Khairuddin, and E. B. Hanafi, "VGG classification model for lung cancer diagnosis," in *AIJR Proceedings*. Malaysia: AIJR, 2022, pp. 57–62.

[28] C. S. Sannasi and H. Rajaguru, "Deep learning-based breast cancer detection using VGG-NiN architecture," *AIP Conf. Proc.*, vol. 2725, Apr. 2023, Art. no. 020001.

[29] Y. Mansour, K. Lin, and R. Heckel, "Image-to-image MLP-mixer for image reconstruction," 2022, *arXiv:2202.02018*.

[30] C. Xu, "Applying MLP and CNN on handwriting images for image classification task," in *Proc. 5th Int. Conf. Adv. Electron. Mater., Comput. Softw. Eng. (AEMCSE)*, Apr. 2022, pp. 830–835.

[31] A. Besrour and R. Ksantini, "Incremental subclass support vector machine," *Int. J. Artif. Intell. Tools*, vol. 28, no. 7, Nov. 2019, Art. no. 1950020.

[32] J. Zou, C. Yuan, X. Zhang, G. Zou, and A. T. K. Wan, "Model averaging for support vector machine by cross-validation," 2021, *arXiv:2112.14578*.

[33] R. G. Brereton and G. R. Lloyd, "Support vector machines for classification and regression," *Analyst*, vol. 135, no. 2, pp. 230–267, 2010.

[34] M. A. Chandra and S. S. Bedi, "Survey on SVM and their application in image classification," *Int. J. Inf. Technol.*, vol. 13, no. 5, pp. 1–11, Oct. 2021.

[35] W. S. Noble, "What is a support vector machine?" *Nature Biotechnol.*, vol. 24, no. 12, pp. 1565–1567, Dec. 2006.

[36] Z. Saurav, M. M. Mitu, N. S. Ritu, M. A. Hasan, S. Arefin, and D. M. Farid, "A new method for learning decision tree classifier," in *Proc. Int. Conf. Electr., Comput. Commun. Eng. (ECCE)*, Feb. 2023, pp. 1–6.

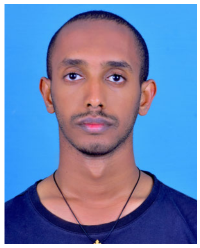
[37] M. S. G. N. Nithish, S. Babu, and V. Yogesh, "Detection of DDOS attack using decision tree classifier in SDN environment," *J. Ubiquitous Comput. Commun. Technol.*, vol. 5, no. 2, pp. 193–202, Jun. 2023.

[38] M. A. Zardari, L. T. Jung, and N. Zakaria, "K-NN classifier for data confidentiality in cloud computing," in *Proc. Int. Conf. Comput. Inf. Sci. (ICCOINS)*, Jun. 2014, pp. 1–6.

[39] J. Deng, A. C. Berg, K. Li, and L. Fei-Fei, "What does classifying more than 10,000 image categories tell us?" in *Proc. 11th Eur. Conf. Comput. Vis.*, Crete, Greece. Cham, Switzerland: Springer, 2010, pp. 71–84.

[40] K. Nanthini, D. Sivabalaselvamani, K. Chitra, P. Gokul, S. KavinKumar, and S. Kishore, "A survey on data augmentation techniques," in *Proc. 7th Int. Conf. Comput. Methodologies Commun. (ICCMC)*, Feb. 2023, pp. 913–920.

[41] B. M. Abuhayi and A. A. Mossa, "Coffee disease classification using convolutional neural network based on feature concatenation," *Informat. Med. Unlocked*, vol. 39, Feb. 2023, Art. no. 101245.



BINIYAM MULUGETA ABUHAYI received the B.S. degree in electrical and computer engineering and the M.Sc. degree in information technology from the University of Gondar, Gondar, Ethiopia, in 2016 and 2022, respectively. He is currently with the University of Gondar. He is the author of two books and many articles. His research interests include machine learning, natural language processing, image processing, and computer vision. His research has been published in several reputable international conferences and major international journals. He serves as a reviewer for several reputable international journals.

YOHANNES AGEGNEHU BEZABH received the B.S. degree in information technology from Jijiga University and the M.Sc. degree in information technology from the University of Gondar, Gondar, Ethiopia. He is currently with the University of Gondar. He is the author of multiple articles. His research interests include computer vision, image processing, machine learning, deep learning, and data science.

ALEKA MELESE AYALEW received the B.Sc. degree in information technology from Adigrat University, Tigray, Ethiopia, and the M.Sc. degree in information technology from the University of Gondar, Gondar, Ethiopia. His research interests include computer vision, image processing, machine learning, deep learning, and data science. His research has been published in several reputable international conferences and major international journals. With a strong academic background, he serves as a reviewer for several reputable international journals.