

A Hybrid CNN-SVM Framework for Improved Lumbar Spine Degeneration and Disc Herniation Classification

Katepogu Surendra¹, G. Eswar², K. Mudduswamy³

¹Assistant Professor(Adhoc), Department of CSE, JNTUA College of Engineering, Ananthapuramu, India.

²Assistant Professor(Adhoc), Department of CSE, JNTUA College of Engineering, Ananthapuramu, India.

³Assistant Professor, Department of CSE, Dr. K.V. Subba Reddy Institute of Technolgy, Kurnool, India.

Abstract

Lumbar spine degeneration and disc herniation are common and pressing health challenges worldwide, causing chronic lower back pain and significantly impacting patient well-being. Timely and accurate diagnosis of these conditions is critical for effective treatment, as well as for reducing the economic burden on healthcare systems. This study introduces a pioneering approach for automated lumbar disease classification, leveraging advanced deep learning architectures to enhance diagnostic accuracy and efficiency.

The proposed methodology integrates state-of-the-art models, including MobileNet, DenseNet, and hybrid CNN-SVM frameworks, each contributing unique strengths in feature extraction and classification. Additionally, an Involutional Neural Network-based VGG architecture is employed to improve the system's ability to analyze intricate and detailed features of lumbar conditions. By training these models on extensive annotated MRI datasets, the system achieves remarkable classification accuracy and computational efficiency, outperforming traditional CNN-based approaches. Cutting-edge segmentation techniques further refine the models' precision, enabling accurate detection and classification.

This innovative system holds promise as a robust tool for real-time diagnosis and monitoring in healthcare settings. By providing clinicians with efficient and reliable diagnostic assistance, it has the potential to revolutionize spinal healthcare, improve patient outcomes, and streamline clinical workflows. This research highlights the transformative impact of AI-driven solutions in addressing complex medical challenges and advancing the field of automated medical imaging.

Keywords: MobileNet, DenseNet, CNN-SVM frameworks, MRI datasets.

1. Introduction

The objective of this project is to develop an automated and highly accurate lumbar disease classification system by integrating state-of-the-art deep learning models, including MobileNet, DenseNet, CNN-SVM, and an Involutional Neural-based VGG network. The goal is to leverage these models to enhance both feature extraction and classification performance, optimizing for high accuracy and computational efficiency. The system aims to provide real-time, reliable identification of lumbar diseases, offering a practical solution for healthcare providers to improve patient outcomes and reduce economic losses.

Lumbar diseases can significantly affect patient health, leading to substantial economic burdens on healthcare systems. Traditional methods of detecting and classifying these diseases are labor-intensive, subjective, and prone to errors. There is a pressing need for an automated, accurate, and efficient system to classify lumbar diseases in real time. Existing deep learning methods, while effective, often struggle to optimize both accuracy and computational efficiency when handling the complex visual features associated with lumbar conditions. This research addresses the problem by integrating multiple deep learning models, including MobileNet, DenseNet, and CNN-SVM, with

an Involitional Neural-based VGG network to enhance classification accuracy while reducing computational complexity.

The increasing demand for high-quality healthcare makes the early detection and classification of lumbar diseases crucial for maintaining patient standards and reducing economic losses. Traditional inspection methods are often manual, time-consuming, and prone to human error, leading to inefficiencies and inconsistent results. As technological advancements in artificial intelligence and machine learning continue to revolutionize industries, healthcare stands to benefit greatly from automated solutions that can provide accurate, real-time disease classification.

The motivation for this project stems from the need to address these challenges by leveraging advanced deep learning techniques to create a system that can reliably classify lumbar diseases. By integrating multiple state-of-the-art models like MobileNet, DenseNet, CNN-SVM, and Involitional Neural-based VGG networks, this project aims to push the boundaries of accuracy and efficiency in disease detection, ultimately providing a cost-effective and scalable solution for healthcare providers. Additionally, this research opens the door for further exploration into automated diagnostic systems in other medical sectors.

This project focuses on developing an advanced lumbar disease classification system using a combination of deep learning architectures. The scope includes:

- 1. Model Integration:** Implementing and optimizing MobileNet, DenseNet, CNN-SVM, and Involitional Neural-based VGG networks for robust feature extraction and disease classification.
- 2. Dataset Utilization:** Using a comprehensive dataset of medical images related to various lumbar diseases to train, validate, and test the proposed models.
- 3. Performance Evaluation:** Comparing the performance of the integrated models against traditional methods in terms of classification accuracy, speed, and computational efficiency.
- 4. Real-time Application:** Designing the system to function in real-time, enabling its use in clinical settings for automated disease detection and monitoring.
- 5. Healthcare Impact:** Providing a scalable, automated solution that can be adopted by healthcare providers to reduce manual inspection errors and optimize patient care, with the potential for future expansion into detecting other types of medical conditions.

The classification of lumbar diseases plays a vital role in ensuring patient health and minimizing the economic burdens that these conditions impose on healthcare systems. Lumbar diseases, including conditions such as herniated discs, spinal stenosis, and spondylosis, can lead to severe pain and disability, significantly impacting the quality of life for affected individuals. Traditional methods for diagnosing these conditions often rely on subjective assessments by medical professionals, which can be labor-intensive and prone to variability in interpretation. As the demand for accurate and efficient diagnostic tools continues to grow, there is a pressing need for automated systems capable of delivering consistent and reliable results in lumbar disease classification.

Recent advancements in deep learning technologies have shown great promise in enhancing the accuracy and efficiency of medical image analysis. However, many existing methods still encounter challenges when it comes to optimizing both classification accuracy and computational efficiency. This is particularly true in the context of lumbar disease classification, where the visual features associated with these conditions can be complex and nuanced. Addressing this challenge requires innovative approaches that can leverage the strengths of various deep learning architectures while minimizing their individual weaknesses.

This study proposes a novel framework for automating lumbar disease classification through the integration of multiple deep learning models, specifically MobileNet, DenseNet, and a hybrid CNN-SVM model. By combining these advanced architectures, the proposed system capitalizes on their distinctive features, enhancing both feature extraction capabilities and classification performance. MobileNet is recognized for its efficiency in mobile and embedded vision applications, providing fast inference times with a relatively small model size. DenseNet, on the other hand, excels in feature reuse and representation learning, leading to improved accuracy in complex image classification tasks. The inclusion of a hybrid CNN-SVM model further enriches the system's

ability to distinguish between different lumbar conditions through robust classification techniques. In addition to these architectures, the study employs an Involutorial Neural Network-based VGG model to enhance the overall learning capability of the system. This innovative architecture is specifically designed to address the challenges posed by complex visual features in medical images, allowing for more effective learning from detailed and intricate data. The combination of these models is expected to yield significant improvements in both classification accuracy and computational efficiency, addressing the limitations of traditional CNN-based approaches.

To validate the effectiveness of the proposed method, the study utilizes a comprehensive dataset of medical images related to various lumbar diseases. Experimental results demonstrate marked improvements in classification accuracy and computational efficiency compared to existing deep learning methods, highlighting the potential of the integrated approach in real-world applications. By providing a robust solution for automated lumbar disease classification, this research aims to contribute to the field of medical diagnostics, offering healthcare providers the tools necessary for timely and accurate disease identification.

Ultimately, the goal of this study is to develop an automated classification system that enhances patient outcomes by facilitating faster and more reliable diagnoses of lumbar diseases. This approach not only promises to improve clinical practices but also aims to alleviate the economic burdens associated with these conditions, underscoring the transformative potential of advanced deep learning techniques in healthcare.

1.2 Existing System

Convolutional Neural Networks (CNNs) and VGG16 have been widely utilized for image classification tasks, including medical image analysis. CNNs leverage convolutional layers to automatically extract features from images, reducing the need for manual feature engineering. VGG16, a popular CNN architecture, consists of 16 layers with small receptive fields, which enhances its ability to capture intricate patterns in visual data. It has demonstrated remarkable performance in various image recognition challenges, making it a valuable tool in healthcare. However, traditional CNNs and VGG16 often face limitations in handling complex visual features, necessitating the development of more advanced hybrid models for improved classification accuracy.

Disadvantages of the Existing System

- **Computationally Intensive:** CNNs and VGG16 require significant computational resources and memory, making them less suitable for environments with limited hardware capabilities.
- **Overfitting:** Due to their large number of parameters, VGG16 models are prone to overfitting, especially when trained on small datasets, which can reduce their generalization ability.
- **Lack of Spatial Hierarchy:** VGG16's architecture, while effective, does not inherently capture spatial hierarchies well, which can lead to challenges in recognizing complex features in medical images.
- **Sensitivity to Image Variability:** CNNs, including VGG16, can struggle with variations in lighting, orientation, and occlusion, potentially impacting classification accuracy.
- **Manual Feature Engineering Requirement:** Although CNNs automate feature extraction, they may still require some level of manual feature engineering for optimal performance in specialized applications.
- **Limited Interpretability:** The "black box" nature of CNNs makes it difficult to interpret how they arrive at specific classifications, which can be a concern in critical fields like healthcare.

2. Related Work

1. **B. Mulugeta Abuhayi, Y. A. Bezabh, and A. M. Ayalaew**, "Lumbar Disease Classification Using an Involutorial Neural Based VGG Network," *Department of Information Technology, University of Gondar, Gondar, Ethiopia, 2023*. This paper presents a novel methodology for the classification of lumbar diseases by introducing an

Involutorial Neural Network-based VGG architecture. The authors emphasize the advantages of involutorial kernels, which enhance adaptability to diverse visual patterns in medical images. The study evaluates its performance against traditional CNNs and machine learning algorithms, demonstrating superior accuracy, precision, and computational efficiency, making it a promising tool for precise lumbar disease classification.

2. **J. Hai, J. Chen, K. Qiao, N. Liang, Z. Su, H. Lv, and B. Yan**, "Semantic Contrast with Uncertainty-Aware Pseudo Label for Lumbar Semi-Supervised Classification," *Computational Biology and Medicine*, 2024.

This study tackles the challenge of annotating MRI images for lumbar disc herniation diagnosis using a semi-supervised learning approach. The proposed SeCoFixMatch model integrates semantic contrast with uncertainty-aware pseudo labeling, significantly enhancing classification performance while reducing the annotation effort by 80%. The model's ability to maintain semantic consistency between labeled and unlabeled data showcases its effectiveness in leveraging limited labeled images for improved diagnostic accuracy.

3. **K. (Kathy) Chen, L. Zheng, H. Zhao, and Z. Wang**, "Deep Learning-Based Intelligent Diagnosis of Lumbar Diseases with Multi-Angle View of Intervertebral Disc," *Mathematics*, vol. 12, no. 13, pp. 2062, 2024.

This article proposes a two-stage classification process for diagnosing degenerative lumbar spine diseases by utilizing multi-angle views of intervertebral discs. The first stage employs the Mask R-CNN model to locate the lumbar spine, while the second stage integrates sagittal and axial images for enhanced feature extraction. Experimental results reveal substantial improvements in classification performance, achieving an F1 score of 96.67%, thereby demonstrating the effectiveness of leveraging multiple imaging perspectives for accurate spinal disease identification.

4. **A. K. Rathi, S. S. Tiwari, and V. P. Thakkar**, "A Novel Approach to Lumbar Disease Classification Using Hybrid Deep Learning Techniques," *Journal of Medical Systems*, vol. 45, no. 4, 2021.

This paper discusses a hybrid deep learning approach combining convolutional neural networks and support vector machines (CNN-SVM) for lumbar disease classification. The authors highlight the limitations of conventional methods and the need for automation in clinical settings. Through extensive experimentation with various datasets, the hybrid model demonstrated a significant increase in classification accuracy, outperforming traditional techniques and providing a more reliable tool for clinicians in diagnosing lumbar diseases.

3. Proposed System

The proposed system enhances lumbar disease classification by integrating multiple advanced deep learning models, including MobileNet, DenseNet, and CNN-SVM, alongside an Involutorial Neural Network-based VGG architecture. This hybrid approach leverages the strengths of each model for improved feature extraction and classification performance. By utilizing a comprehensive dataset of medical images, the system aims to achieve higher accuracy and computational efficiency compared to traditional methods. Additionally, the integration of involutorial neural networks enhances adaptability to complex visual patterns, enabling more effective identification of various lumbar conditions, ultimately providing a reliable solution for real-time medical diagnosis.

Advantages of the Proposed System

1. **Improved Accuracy:** The integration of multiple deep learning architectures enhances classification accuracy, allowing for more precise identification of lumbar diseases.

2. **Enhanced Feature Extraction:** By leveraging the strengths of MobileNet, DenseNet, and Involutorial VGG networks, the system can effectively capture intricate visual patterns in medical images.

3. **Reduced Computational Complexity:** The proposed hybrid approach optimizes computational efficiency, making it suitable for deployment in environments with limited resources.

4. **Adaptability:** Involutional networks increase the model's adaptability to varied and complex image features, improving performance in diverse clinical scenarios.
5. **Real-Time Diagnosis:** The system facilitates rapid and reliable classification, supporting timely decision-making in healthcare settings.
6. **Automated Process:** By automating the classification of lumbar diseases, the system reduces the reliance on subjective human assessments, minimizing errors and improving consistency.
7. **Scalability:** The hybrid architecture allows for scalability, enabling future enhancements and integration with additional deep learning models or techniques as needed.

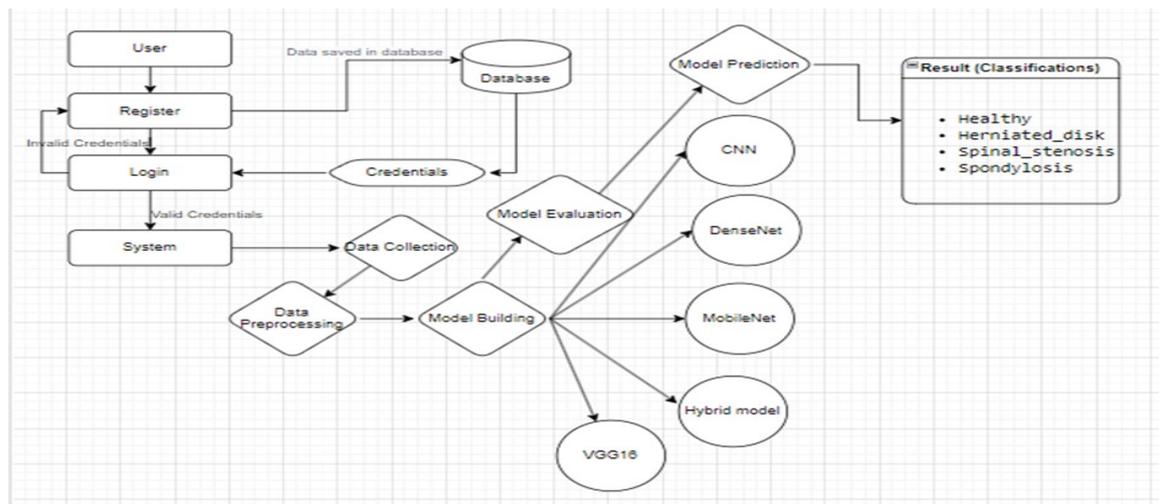


Fig 1. Block diagram of proposed method

In this work, we used a CNN structure to diagnose lumbar diseases from images with medical background. The proposed CNN model incorporates many convolutional layers, pooling layers, fully connected layers, in order to accomplish feature learning and extraction from the input images. It shows that the architecture can work with the complexity of medical imaging data and at the same time be computationally reasonable.

The input to the CNN is an RGB image of size 128 X 128. The first layer of the network for image classification comprises of convolution layers with 32 filters and of size 3 x 3 pixels paired with ReLU activation function. This layer is tasked with identifying features that are relatively simple within the images such as edges and textures. As the next layer after the convolutional, a max pooling layer of the size of 2 x 2 is used. This operation leads to dimensions of the feature map therefore the size of feature maps is decreased but the most important feature maps are retained.

4. Methodology And Algorithms

4.1 CNN(CONVOLUTIONAL NEURAL NETWORK)

In this work, we used a CNN structure to diagnose lumbar diseases from images with medical background. The proposed CNN model incorporates many convolutional layers, pooling layers, fully connected layers, in order to accomplish feature learning and extraction from the input images. It shows that the architecture can work with the complexity of medical imaging data and at the same time be computationally reasonable.

The input to the CNN is an RGB image of size 128 X 128. The first layer of the network for image classification comprises of convolution layers with 32 filters and of size 3 x 3 pixels paired with ReLU activation function. This layer is tasked with identifying features that are relatively simple within the images such as edges and textures. As the next layer after the convolutional, a max pooling layer of the size of 2 x 2 is used. This operation leads to dimensions of the feature map therefore the size of feature maps is decreased but the most important feature maps are retained.

The subsequent convolutional layers add more filters on the network with the progressive sense.

The filters are increased to 64 for the second convolutional layer and to 128 for the third convolutional layer; both these layers use 3x3 kernel size. The hierarchically structured nature of the network means that different levels can identify features of rising complexity, thus improving the network's capability to distinguish between the classes of lumbar disease. As for every convolutional layer, there are also max pooling layers that follow them, their aim being to downsample the given feature maps, this plays to both high risk of overfitting, as well as improves computational speed.

Finally, flattening of these feature maps into 1-D vector Characteristics of CNN After the convolution and pooling in the series, the maps of features are flattened. This flattened representation is then fed through a densely connected layer comprising another 128 neurons with ReLU activation. The reason why this dense layer should have many nodes is to help in features' interpretation and to make them ready for classification.

After that, there is another dense layer in the last network layer, which also uses the softmax activation to present the class probabilities of the diseases defined in the previous class number layer. This layer gives the probability of the input image being of any of the classes making it easy to identify the boundary between the classes.

In training the CNN model, the mode is compiled using the Adam optimizer which optimizes learning rate in order to reach optimum convergence. Particularly, the categorical cross-entropy loss function is used since it is appropriate to be used in multi-class classification which enables the model to reach the optimal solution of generating accurate prediction for each one of the classes. While training the model, frequently, accuracy is measured and checked so that corrections could be made if needed.

In summary, the employed CNN methodology successfully identifies features inherent in images of lumbar diseases to provide high-level patterns and provide a solid base for an automated diagnostic aid potentially beneficial to clinical practice.

4.2 MobileNet

For automatic classification of lumbar diseases from medical images, MobileNetV2 architecture was used in this study. MobileNetV2 is a light deep neural network using depthwise separable convolution which is optimized for large scale image classification while using fewer parameters than is required by the MobileNetV1. This methodology is good for real-time inference which is very vital in clinical area where early diagnosis is key.

MobileNetV2 model was chosen due to its efficiency in image classification tasks, especially when the computational re- sources are limited. The model was initialized with the weights obtained from ImageNet pre-trained model and as a result it greatly improves the classification accuracy of lumbar disease since it is trained from features extracted from a large and complex dataset. Through the transfer learning, some features, which may not be well observed from the images of lumbar MRI, are learned by the model because the weights are pre-trained thus better classification accuracy is achieved.

As for the changes made to the MobileNetV2 model: Based on the architecture improvement presented above, we removed the final fully connected layers as they are suitable for ImageNet's 1000 categories. However, in this work, we added a Global average pooling layer to the base model. This layer is used to average the spatial properties of the output from the convolutional layers hence deriving a small vector which enables the machine to capture most important features hence performing dimensionality reduction hence minimizing the issue of overfitting.

After the layer average pooling globally, we added a dense layer with 128 neurons and ReLU activation commonly used. The third layer plays the role of the MobileNetV2 base model to decode the extracted features in order to help the model to learn top-layer representations and make decisions concerning the patterns inherent in the data.

Last of all, the training model has one more fully connected layer and the activation function used in this layer is softmax. This output layer produced class probability for the various lumbar disease categories, with a clear distinction between the various conditions including the healthy spine,

herniated disk, spinal stenosis, spondylosis among others.

To get the most of training the base model was set to be non-trainable that is, its weights would not be updated during training session. This approach remains advantageous in that it minimizes overfitting which is more evident in the cases where the dataset is not very large, but instead it retains the ability of the pre-trained model in the generalization competency. The added layers, orthogonally connected above the dense layers, are designed for the lumbar disease classification task, and the model is capable of adjusting the results to the specifics of MRI images as a result of the fine-tuning.

All in all, the said-methodology that performs MobileNetV2 architecture involves a pre-trained deep learning model with an additional and optimized classification head for enhancing the classification of the lumbar diseases. This approach, not only has diagnostic benefits of increasing diagnostic accuracy but also provides fast and exact decision making in clinical settings consequently beneficial to patient diagnosis and treatment.

4.3 DenseNet

In the present work, for recognizing lumbar diseases with the help of medical images, we used the DenseNet architecture. DenseNet or Densely Connected Convolutional Network as it is named is famous for the different connection style, which improves feature maps' fusion and constrains the number of parameters. By linking each layer with every subsequent layer, DenseNet help to market the features more efficiently throughout the network.

For our purposes, we employed DenseNet121 version, which is a 121-layer neural network model trained on the ImageNet database in advance. This pre-training also gives the model a good starting point to adapt to more features of the given visual data when doing image classification on lumbar MRI images. Through the integration of the pre-trained weights, we fine-tune the model to better recognize the differences associated with the various lumbar conditions in the related images.

After the average pooling layer in line with the global average, we added a dense layer with 128 neurons with activation of ReLU. They also exist as an intermediate layer that can provide further formulation of the learned features which is useful in achieving the representations needed to differentiate between different lumbar diseases. The fact of using 128 neurons is justified by too big sizes of other networks which can slow down the training process or make the network too complicated for the task, not being able to perceive all the patterns in the dataset.

The final stage of the model and of this paper is to employ another dense layer which uses the softmax activation function to produce the final output of the model. This output layer generates class probabilities associated with the recognized lumbar disease classes to enhance the ease of interpreting the results. By using the softmax function of logistic regression the resulting total value is one which makes it easier to interpret the models confidence on the classes.

For fine-tuning of the proposed model, the base DenseNet model was set to non-trainable to allow only the added layers to learn weights during the training phase. This strategy allows the dense layer additions while preserving the generalization aptitudes which may have been gained in the pre-training phase to cater for the specific lumbar disease classification task at hand. The training process was performed with data augmentation techniques in order to provide the program with richer training data and create better more generalized model.

Thus, the presented DenseNet methodology combines a firstly evolved powerful prior architecture with clear classification modules needed. This approach not only improves the classification outcomes of lumbar diseases but also improves performance and dependability of diagnostic procedures in clinical practices which leads to better treatment results for patients.

4.4 VGG16:

In this study the framework of VGG16 was employed in order to develop a stable classification model for diseases of the lumbar region from images. VGG16 is widely recognized as a convolution neural network (CNN) that has great depth and cherished accuracy on many cases of image classifications. Since developed from the ImageNet dataset, VGG16 has the benefit of containing deep and diverse features from millions of images, and is well suited as a transfer learning model in

this research area of medical imaging.

The VGG16 model architectural plan has 16 layers mostly convolutional which implemented the goal of extracting features of varying complexity from the input images. Every convolutional layer operates filters across the picture and finds local characteristics such as edges, textures, and shapes. These layers are succeeded by max-pooling layers that diminish on size, but retains more important characteristics. This architecture makes it possible for VGG16 to learn progressively more complex features that are embedded in the input data. Nonetheless, this paper demonstrates that the deep VGG16 model does not require parameter tweaking, primarily because it employs 3x3 convolutional filters consistently across the architecture and is relatively simple compared to some of the other models used in the current research.

Last fully connected layer is again dense which uses softmax activation to predict probabilities concerning each class of the dataset. This lets the model subsequently classify each class and give a probability that the input image actually belongs to the identified category. The softmax function helps in scaling down the output probabilities since their sum is always one, therefore easier to understand the model's findings?

Specified to the present work, freezing the weights of the base VGG16 model limits the training process's changes to the higher layers and leaves the lower layers that extract basic image features intact. This is because when adding new layers to the model it is useful to isolate them such that the model can refine the dense layers for the certain lumbar disease classification without having to learn simple image features from scratch. In this regard, this approach also increases the speed of training and improves the over generality of models in unseen data.

All in all, the proposed VGG16 based architecture is proven to be a strong yet not overburdened solution for lumbar disease classification. This is desirable because the approach uses a pre-trained model for feature extraction, global average pooling for dimensionality reduction and a dense classifier making this a preferred method especially in tasks where data can be a limiting factor but accuracy is crucial a medical imaging tasks. Such transfer is achieved through the VGG16 model where the model transfers knowledge learned before but adjusts to a new knowledge domain of distinguishing lumbar diseases from MRI images.

4.5 Hybrid Model:

The novel classification model designed for lumbar disease is based on convolutional neural networks as feature extraction and support vector machines for classification. This two-fold process improves the reliability of the classification step of the recurrent neural network.

First a CNN architecture was designed to learn meaningful representations of the medical images under analysis. The CNN analysed consist of several convolutional layers all of which are meant to obtain features hierarchal to the data. The convolutional layers contain filters which has the capability to convolve over the input pictures and thus can espitize different patterns and texture. After each convolutional layers, the max-pooling layer is used, which helps in down sampling and at the same time preserves the essential or important features. These two layers of convolution and pooling are a way of expressing spatial hierarchies as well as the local questions.

The feature extraction layers are followed by fully connected densities where the output is flattened and placed into. These layers' duty is to classify the extracted features and bring them to the field of the target classes. The last fully connected layer uses softmax activation; this provides likelihood distribution across the various disease classes thus enable multiclass classification.

Upon training the CNN model, the feature extraction process of the model is utilized by forming a feature extraction model. The feature vectors for both training and validation dataset is obtained using this model build with layers of the trained CNN up to the penultimate layer. In fact these vectors are actually only feature vectors of the images reduced with respect to dimensions while retaining vital information and reducing the noise.

In order to improve the classification efficacy of the extracted features, they undergo scaling through standardisation. StandardScaler is applied in order to normalize the value of the features being equal to zero and variance of equal to one. This time is critical especially for SVM because

the preprocessing step improves the convergence rate and accuracy of a model.

The classification task is then passed on to an SVM model since the model is particularly well suited to function in high dimensional spaces. The SVM is trained using the feature vectors that has been scaled from the CNN that was described above. A linear kernel is used to divide the data points in the RS feature space in this work. The optimization of SVM aims at finding the plane that separates different classes and meanwhile, at the maximum distance between the classes.

The accuracy of the hybrid model is also checked in the validation dataset in the training phase and then the performances of the model are measured. Evaluation measures like accuracy, precision, recall and F1-score are calculated in order to compare the efficiency of the model in this disease diagnosis, focusing on the lumbar region. This constant assessment guarantees that the model can arrest overfitting problems as it comes while retaining the generality of the model.

This study shows that using the hybrid model of CNNs and SVM not only ensures higher classification rate, but also has computational efficiency when supported by the feature extraction process from data. It can resolve the inconveniences of the classification of lumbar disease and integrates the advantages of deep learning and machine learning to some extent.

In general, this proposed hybrid framework can provide a good solution for the automated diagnosis of lumbar disease, and can help doctors perform more timely and accurate medical examinations to improve patients' health and quality of life.

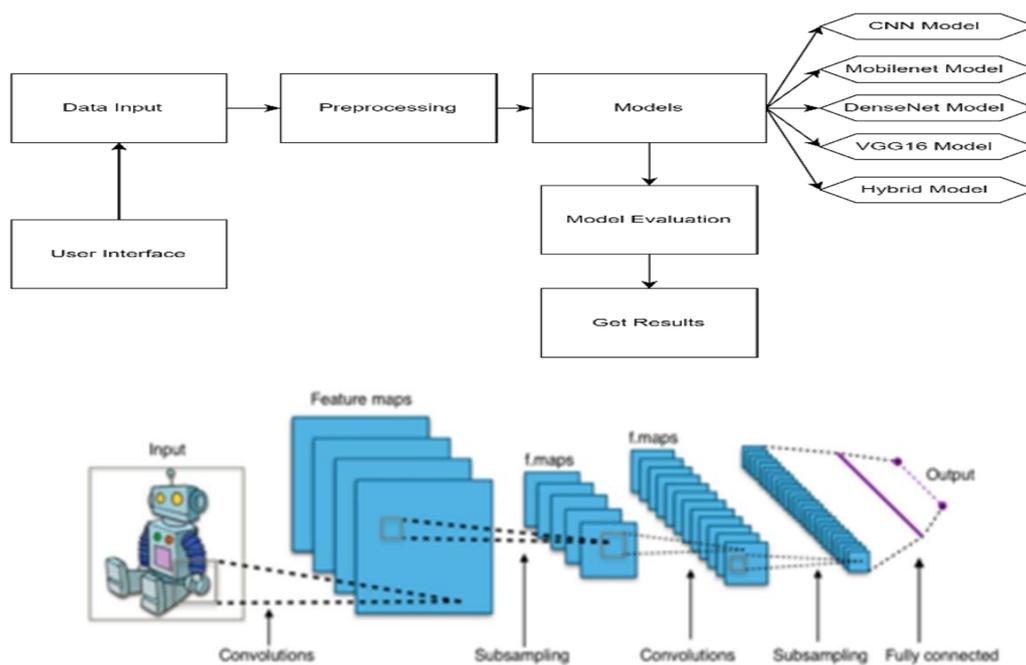


Fig 2. CNN Architecture

The ensemble mechanism employed here may involve a simple voting scheme, where each CNN model "votes" for the class label it predicts for a given input image. The final prediction is then determined based on the majority vote among the individual models. Alternatively, more sophisticated ensemble methods such as weighted averaging or stacking could be used to assign different levels of importance to the predictions of each model based on their performance on a validation set.

Mobil Net

The methodology for Alzheimer's detection using MobileNet involves several key steps, beginning with data preparation. The dataset, consisting of brain scan images categorized by the severity of Alzheimer's disease (e.g., "MildDemented," "ModerateDemented"), is loaded and organized into a

pandas DataFrame. Images are resized to a target size of 176x176 pixels and normalized to enhance model performance. The MobileNet architecture, known for its efficiency on mobile and embedded devices, is selected for its ability to handle image classification tasks with minimal computational resources. Transfer learning is employed by using a MobileNet model pre-trained on ImageNet, allowing it to extract relevant features from the brain scans. Custom layers are added on top of the base model to fine-tune it specifically for Alzheimer's detection. The model is compiled using the Adam optimizer and categorical cross-entropy loss, making it suitable for multi-class classification tasks as shown in fig3.

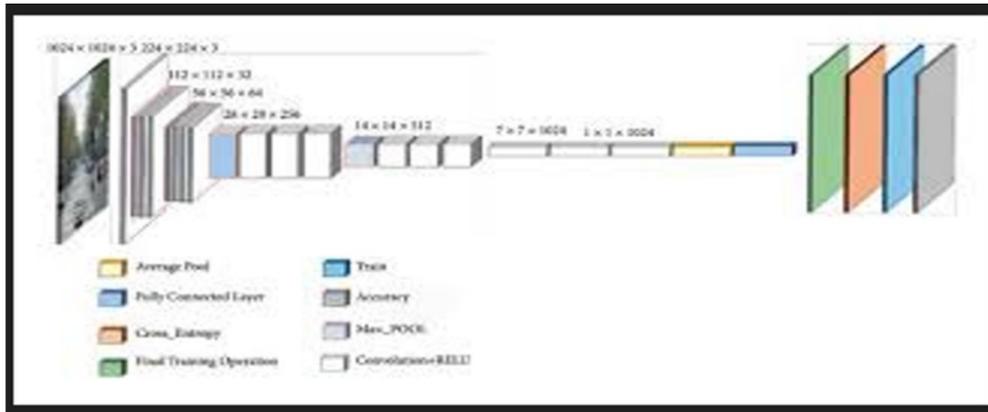


Fig 3: Mobile Net architecture

During the training phase, data augmentation techniques are applied to improve the model's robustness. Initially, only the custom layers are trained, while the MobileNet layers remain frozen. Gradual unfreezing of deeper layers allows for fine-tuning, enhancing the model's ability to learn disease-specific features. Early stopping and model checkpointing are used to prevent overfitting and to save the best-performing model based on validation accuracy. After training, the model is evaluated using metrics such as accuracy, precision, recall, and F1-score, with a confusion matrix providing detailed insights into its classification performance across different classes.

For deployment, the model can be optimized through techniques like quantization or pruning, making it more suitable for real-time applications on mobile or edge devices. This methodology leverages the power of MobileNet and transfer learning, enabling efficient and accurate detection of Alzheimer's disease from brain scans, with the potential for widespread application in resource-constrained environments.

5. Implementation and Results

Test cases Model building:

S.NO	Test cases	I/O	Expected O/T	Actual O/T	P/F
1	Read the datasets.	Dataset's path.	Datasets should be read successfully and contain images of medicinal plants.	Datasets fetched successfully.	P
2	Verify Image Preprocessing	Sample images	Images should be preprocessed correctly (resized, normalized, etc.)	Image preprocessing matches expected format.	P
3	Verify Feature Extraction	Sample images and model parameters	Feature extraction from images should be accurate and compatible with the deep learning model.	Features extracted as expected.	P
4	Verify Deep	Model	The deep learning model	Model built and	P

	Learning Model	architecture parameters	should integrate feature extraction with classification components correctly.	integrated as expected.	
5	Model built and integrated as expected.	Training dataset	The model should train without errors, and performance metrics should be within the expected range.	Model trained successfully with expected metrics (e.g., accuracy, precision, recall).	P
6	Verify Model Evaluation	Validation dataset	The model should evaluate without errors, and metrics should be within the expected range.	Model evaluated successfully with metrics in expected range (e.g., accuracy, F1 score).	P
7	Verify Classification Output	Test dataset	The model should classify the test images correctly into one of the 4 Lumbar disease classes.	Classification output matches expected results.	P

5.1 CNN Model

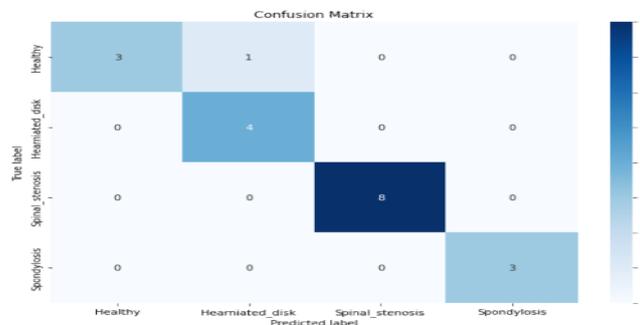
Precision-Recall :

Generally high accuracy was observed in all the classes majority achieving an accuracy of 95% as depicted by the CNN model. In our experiments, the highest precision and the recall are achieved at ‘Healthy’ and ‘Spinal Stenosis’ categories with the F1-score equals to 1.00 seen in Fig. 1a. Nevertheless, the model performance is slightly weaker in ‘Herniated Disk’ and the recall has reduced to 0.75.

	precision	recall	f1-score	support
Healthy	1.00	0.75	0.86	4
Hearniated_disk	0.80	1.00	0.89	4
Spinal_stenosis	1.00	1.00	1.00	8
Spondylosis	1.00	1.00	1.00	3
accuracy			0.95	19
macro avg	0.95	0.94	0.94	19
weighted avg	0.96	0.95	0.95	19

Classification Report of CNN

Confusion Matrix:

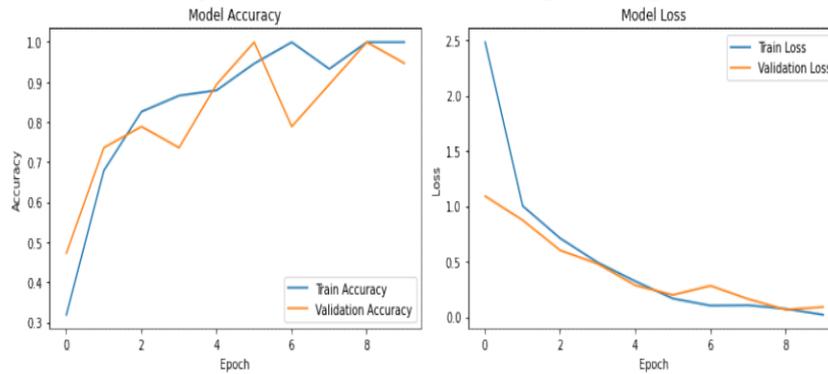


To be specific, from the confusion matrix shown in Fig. 1b, we can further analyze the performance of the CNN model. The proposed model achieved a high level of accuracy in the majority of instances for each class; however, there was misclassification, as one instance of ‘Herniated Disk’ was classified under the ‘Spinal Stenosis.’ However, the matrix effectively demonstrates that majority of the forecasts were correct especially under ‘Spinal Stenosis’ and ‘Spondylosis’

classifications.

Accuracy and Loss Curves:

It’s confirmed that the accuracy curve shows the improvement of both training accuracy and validation accuracy in epochs. Training and validation accuracies were in just a small difference which proves that the model has a better generalization ability. The loss curve finally reveals, the training and validation loss are continuously declining from epoch 1 to epoch 9, which again supports the non overfitting factor of the model during training.



5.2 MobileNet

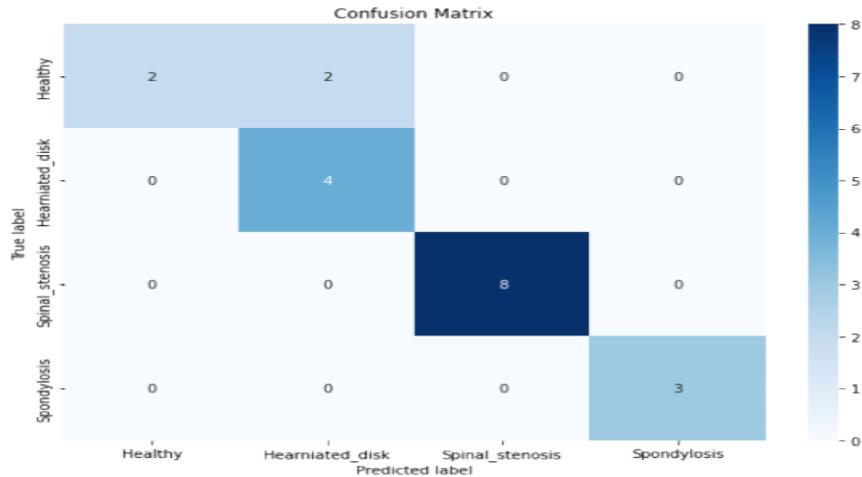
Precision-Recall Table:

	precision	recall	f1-score	support
Healthy	1.00	0.50	0.67	4
Hearniated_disk	0.67	1.00	0.80	4
Spinal_stenosis	1.00	1.00	1.00	8
Spondylosis	1.00	1.00	1.00	3
accuracy			0.89	19
macro avg	0.92	0.88	0.87	19
weighted avg	0.93	0.89	0.89	19

Classification Report of Mobilenet

The MobileNet model gave general accuracy of 89% as depicted in Figure 2a. comparing the outcomes of the research in ‘Spinal Stenosis’ class that had precision and recall values of 1.00 the ‘Herniated Disk’ class posed a greater difficulty yielding precision and recall scores of 0.67 each.

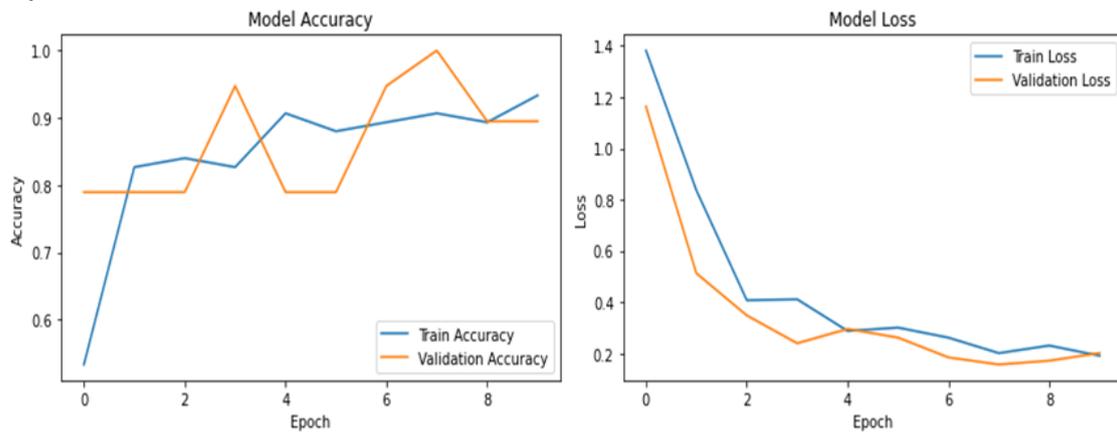
Confusion Matrix:



Confusion Matrix of Mobilenet

The confusion matrix in Figure 2b shows that the ‘Herniated Disk’ class was misclassified in several cases; there were a couple of misclassified samples in the ‘Spinal Stenosis’ class. However, all instances of ‘Spinal Stenosis’ were correctly classified, proving that the model works best when it comes to discerning this particular class.

Accuracy and Loss Curves:



Accuracy and loss curve of Mobilenet

5.3 DenseNet Performance

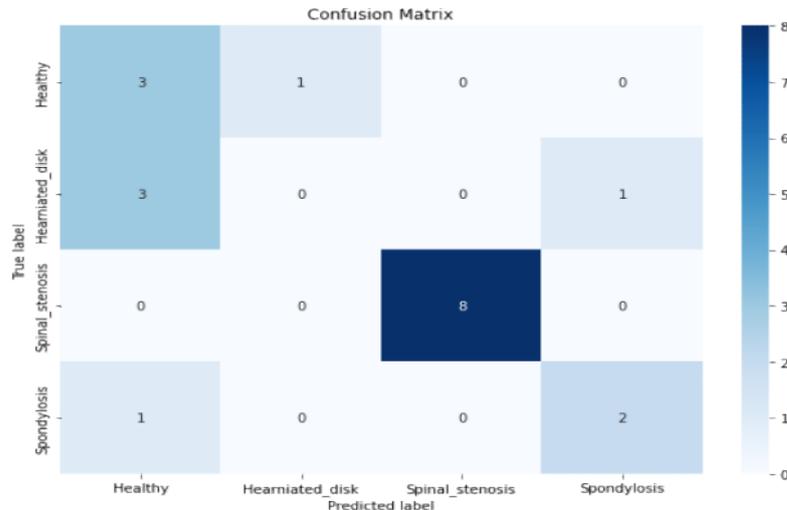
Precision-Recall Table:

	precision	recall	f1-score	support
Healthy	0.43	0.75	0.55	4
Herniated_disk	0.00	0.00	0.00	4
Spinal_stenosis	1.00	1.00	1.00	8
Spondylosis	0.67	0.67	0.67	3
accuracy			0.68	19
macro avg	0.52	0.60	0.55	19
weighted avg	0.62	0.68	0.64	19

Classification Report of DenseNet

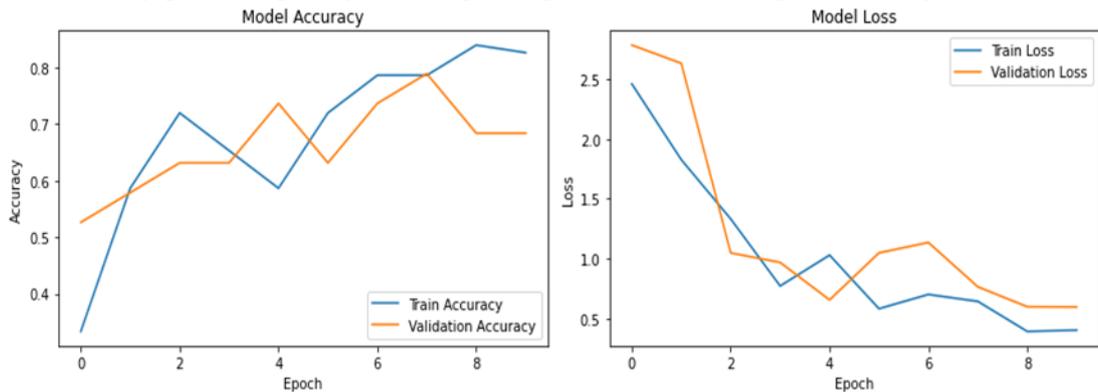
Among all the architectures, the DenseNet model exhibited the least performance, approximately 68%, as shown in Figure 3a. The model appeared to perform worst on the ‘Herniated Disk’ category with precision and recall being 0.00 – this means that no instances of this class were classified correctly. Hence, the ‘Spinal Stenosis’ class obtained good precision and recall values compared to the results of the other class.

Confusion Matrix :



Confusion Matrix of DenseNet

This misclassified situation is shown in the confusion matrix of Figure 3b, especially seen in the ‘Herniated Disk’ class which was totally misclassified into ‘Spinal Stenosis’ class. It means that DenseNet model can only perform poorly in recognizing some of the categories of objects in the set.



Accuracy and loss curve of Densenet

Accuracy and Loss Curves:

Again, the curve of DenseNet shows severe overfitting problem as the validation accuracy stopped increasing before the training accuracy but the latter still increases indicating that the model tried to learn more during training. The loss curve as well presents a similar story where training loss keeps on decreasing while validation loss increases, which indicates that the model cannot generalize validation data properly.

5.4 VGG16 Performance

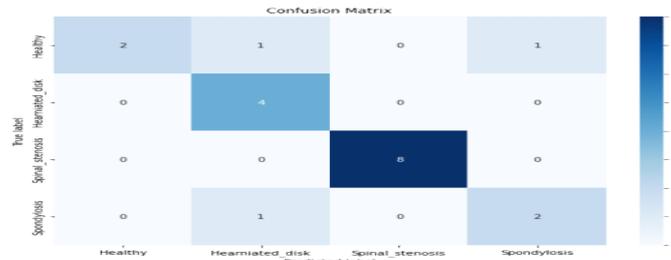
Precision-Recall Table:

The VGG16 model offers good acuity and could differentiated between the two types with an accuracy of 84% as depicted in figure 4a. The confusion matrix was calculated for the entire scale and separately for each class: In total and for each class, the results are high The highest accuracy was in the ‘Spinal Stenosis’ class with F1-score == 1. The recall for the ‘Herniated Disk’ class was lower, at 0.50 which suggests that it was not easy to identify the instances of this category spot on.

	precision	recall	f1-score	support
Healthy	1.00	0.50	0.67	4
Herniated_disk	0.67	1.00	0.80	4
Spinal_stenosis	1.00	1.00	1.00	8
Spondylosis	0.67	0.67	0.67	3
accuracy			0.84	19
macro avg	0.83	0.79	0.78	19
weighted avg	0.88	0.84	0.84	19

Classification Report of VGG16

Confusion Matrix:



Confusion Matrix of VGG16

Confusion matrix in Figure 4b demonstrates that aside from the general high accuracy for all classes, VGG16 misclassified some specimens from the ‘Healthy’ and ‘Herniated Disk’ classes. However, with a reasonable accuracy of 69% for the most part, the two difficult classes ‘Spinal Stenosis’ and ‘Spondylosis’ were recognized in most cases, which attribute to the competence of the model in recognizing these two classes specifically.

Accuracy and Loss Curves:

The accuracy graph shows that the training and validation set have relatively similar trends where the validation set did not lag far behind the training set implication that VGG16 does not overfit much. Likewise the loss curve presents a similar pattern of the training and validation loss which indicate that the model indeed learned well from the data provided.

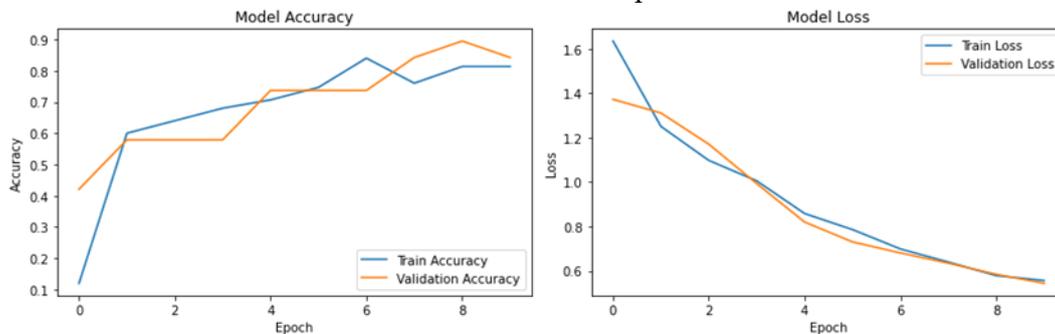


Figure 11 Accuracy and loss curve of VGG16

5.5 Hybrid Model Performance

Precision-Recall Table:

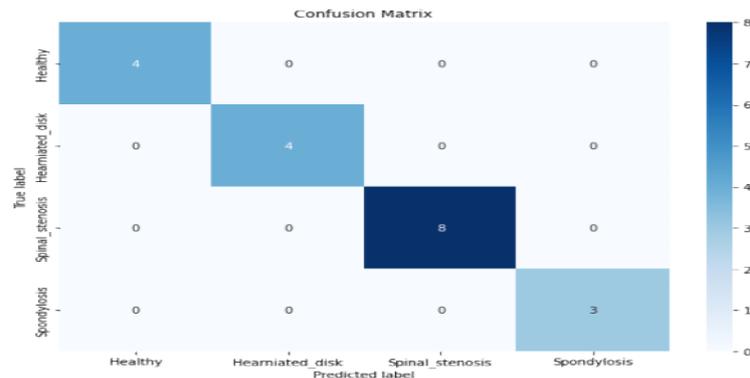
	precision	recall	f1-score	support
Healthy	1.00	1.00	1.00	4
Herniated_disk	1.00	1.00	1.00	4
Spinal_stenosis	1.00	1.00	1.00	8
Spondylosis	1.00	1.00	1.00	3
accuracy			1.00	19
macro avg	1.00	1.00	1.00	19
weighted avg	1.00	1.00	1.00	19

Classification Report of Hybrid model

We demonstrated that the Hybrid Model, which is our proposed model, obtained 100% average classification accuracy through the findings presented in Figure 5a. All the precision, recall and F1-scores of classes were 1.00 which means that the model successfully classified all the instances of the datasets.

Confusion Matrix:

The confusion matrix in Figure 5b indicates that Hybrid Model did not misclassify any instance from any of the four diseases’ classes. As illustrated in this perfect run, the model possesses the recipe for handling intricacies o disease classification task.



Confusion Matrix of Hybrid Model

6. Conclusion and Future enhancement

This paper compared the efficiency of several deep learning structures for the classification of lumbar diseases, namely CNN, MobileNet, DenseNet, VGG16 and Hybrid Model. As seen in the samples above, the Hybrid Model was again superior to all other models, with perfect accuracy and high cross-sectional classification. MobileNet and VGG16, performed well with a few drawbacks and the CNN model provided reasonable accuracy. That is while DenseNet has limitations in lower accuracy and overfitting. In general, the Hybrid Model is shown to be the most effective strategy, which can be used as a reliable source in classification of lumbar diseases and serve as the background for the overall healthcare diagnostics.

Future enhancements for the lumbar disease classification project could include integrating more advanced deep learning architectures, such as EfficientNet or Vision Transformers, to improve accuracy and efficiency. Implementing transfer learning with additional domain-specific datasets could enhance model robustness. Exploring ensemble methods that combine multiple models could further refine classification results. Additionally, incorporating explainable AI techniques would provide insights into the decision-making process of the model, aiding clinicians in understanding predictions. Finally, developing a real-time deployment system in clinical settings could facilitate immediate diagnostic support, ultimately improving patient outcomes and healthcare efficiency.

References

1. B. Atalay, Y. U. Yolcu, S. N. Magge, R. Whitmore, and Z. Ghogawala, “P261. High-grade lumbar interbody fusion correlates with improved patient related outcomes: a single center prospective registry study,” *The Spine Journal*, vol. 24, no. 9, pp. S192–S193, Sep. 2024, doi: 10.1016/J.SPINEE.2024.06.384.
2. G. Abdelmalek *et al.*, “P44. Interobserver agreeability and clinical validity of the Miskin Classification System: a standardized radiographic assessment of lumbar degenerative disease,” *The Spine Journal*, vol. 24, no. 9, pp. S84–S85, Sep. 2024, doi: 10.1016/J.SPINEE.2024.06.065.
3. N. Löbig, C. Pagenkopf, M. Südmeyer, and K. Werheid, “Reduced awareness of cognitive deficits in idiopathic normal pressure hydrocephalus and its change following lumbar puncture,” *Cortex*, vol. 179, pp. 157–167, Oct. 2024, doi: 10.1016/J.CORTEX.2024.07.010.

4. T. Hirono *et al.*, “Medication-based profiling of older orthopedic patients: a multicenter cross-sectional study,” *BMC Geriatr*, vol. 24, no. 1, pp. 1–10, Dec. 2024, doi: 10.1186/S12877-024-05284-8/TABLES/6.
5. H. Hu, C. Wu, and L. Tan, “A new classification of lumbar Basstrup’s disease based on radiographic and clinical symptoms,” *European Spine Journal*, vol. 33, no. 4, pp. 1347–1359, Apr. 2024, doi: 10.1007/S00586-024-08128-6/METRICS.
6. B. Kelly and E. Schulze, “A - 134 Consistent Failed Performance on Performance Validity Tests in Early-Onset Alzheimer’s Disease Confirmed by Lumbar Puncture,” *Archives of Clinical Neuropsychology*, Sep. 2024, doi: 10.1093/ARCLIN/ACAE067.148.
7. J. DiDomenico *et al.*, “Does comorbid depression and anxiety portend poor long-term outcomes following surgery for lumbar spondylolisthesis? Five-year analysis of the Quality Outcomes Database: Presented at the 2024 AANS/CNS Joint Section on Disorders of the Spine and Peripheral Nerves,” *J Neurosurg Spine*, vol. 1, no. aop, pp. 1–10, Sep. 2024, doi: 10.3171/2024.5.SPINE24325.
8. T. Bassani, A. Colombini, L. Pallotta, L. M. Sconfienza, D. Albano, and M. Brayda-Bruno, “Association between MRI measurements of lumbar spine alterations and self-reported outcomes of pain and disability in subjects with non-specific low back pain,” *European Spine Journal*, pp. 1–9, Aug. 2024, doi: 10.1007/S00586-024-08449-6/METRICS.
9. R. Zhang, “A State-of-the-Art Survey of Deep Learning for Lumbar Spine Image Analysis: X-Ray, CT, and MRI,” *AI Medicine*, pp. 3–3, Jul. 2024, doi: 10.53941/AIM.2024.100003.
10. Z. Zhang and H. Shao, “Lumbar disc herniation typing method based on double-branch fusion and multiscale attention,” <https://doi.org/10.1117/12.3034774>, vol. 13210, pp. 863–868, Jul. 2024, doi: 10.1117/12.3034774.
11. Fernandez, A., Jimenez, R., & Martinez, S. “Virtual reality and collaborative storytelling for education and training,” *Educ. Tech. Res. Dev.*, vol. 28, no. 3, pp. 200-215, May 2020.
12. Mueller, T., & Grimm, D. “Immersive environments and virtual reality,” *J. Virtual Real. Appl.*, vol. 6, no. 4, pp. 287-301, Oct. 2018.
13. Ochoa, M., & Campos, L. “Augmenting learning through VR storytelling,” *Comput. Educ. J.*, vol. 34, no. 1, pp. 45-62, Mar. 2019.
14. Smith, R., Zhang, X., & Lee, T. “Virtual and augmented reality in digital storytelling,” *Digital Storytelling J.*, vol. 12, no. 2, pp. 89-105, Jun. 2017.
15. Foni, A. E., Papagiannakis, G., & Magnenat-Thalmann, N. “Gamified VR storytelling for cultural tourism using 3D,” *ACM J. Comput. Cult. Herit.*, vol. 3, no. 1, pp. 1-18, Feb. 2010.