Lumbar Spine Degeneration

1st Nina Burden
Department of Mathematics
Data Science
Hoboken, USA
nburden@stevens.edu

2nd Swapnil Gautam
Department of Electrical and Computer Engineering
Applied Artificial Intelligence
Hoboken, USA
sgautam3@stevens.edu

3rd Harjot Kaur
Department of Mathematics
Data Science
Hoboken, USA
hkaur5@stevens.edu

***Abstract*—This project explored using two machine learning algorithms, Convolutional Neural Networks (CNNs) and Support Vector Machines (SVMs), to classify the severity of spinal stenosis. Given that millions of individuals suffer from spinal degeneration as year, the opportunity to automate lumbar spine degeneration classification opens the doors to more accurate diagnoses and more targeted treatment. The RSNA 2024 Lumbar Spine Degenerative Classification dataset, comprised of 35.34 GB of DICOM images and CSV files totaling 147,320 files, provided the basis for training and evaluating the SVM and CNN models used to classify lumbar spine degeneration. Both models were well-suited for this image-based classification task due to their unique strengths. SVMs excel at identifying optimal boundaries between data classes, especially when dealing with high-dimensional data like the flattened image arrays used in this project. Using a linear kernel, the SVM model achieved a remarkable 94.39 percent accuracy in differentiating between "Normal" and "Severe" cases. This high performance highlights the potential of SVMs as a valuable tool for assisting radiologists in diagnosing spinal stenosis. CNNs, in contrast, are specifically designed for analyzing image data and can automatically learn and extract relevant features from the images without the need for manual feature engineering. This is achieved through a hierarchical approach, where initial layers detect basic edges and textures, while deeper layers identify increasingly complex patterns. CNNs are also adept at handling variations in image size, orientation, and lighting, making them well-suited for medical images, which often exhibit such variations.**

In this project four models were developed and evaluated: a binary CNN and a binary SVM to classify "Normal" and "Severe" cases, and a multiclass CNN and SVM to classify images into "Normal," "Moderate," and "Severe" categories. The binary CNN achieved a validation accuracy of 85.21 percent, while the binary SVM outperformed it, achieving 94.39 percent accuracy on the test set. The multiclass CNN achieved a validation accuracy of 70.49 percent while the SVM achieved a 65.20 percent, indicating that distinguishing between three classes is more challenging than binary classification. Although the binary SVM demonstrated the highest accuracy in this project, future work could explore combining the strengths of both CNNs and SVMs for potentially even better performance. One approach could involve using a CNN to extract features from the images and then feeding those features to an SVM for classification, leveraging the advantages of both algorithms.

I. Introduction

In this project, we tackle the widespread issue of lower back pain that affects millions globally. The main cause is the deterioration of discs in the lumbar spine. Traditional methods for diagnosis and assessment can be very time consuming and may vary with interpretation among practitioners, leading to potential inaccuracies.

Our methodology key steps consist of data preparation, feature extraction, spine level classification, degeneration detection, severity assessment, model evaluation, and final testing. In data preparation, we preprocess the lumbar spine dataset to ensure the images are uniform and to improve on the model's performance. In feature extraction, the model will learn to recognize patterns to focus on key areas related to degeneration to help simplify the image and to make accurate predictions. In spine level classification, the model identifies and labels sections of the lumbar spine to accurately detect degeneration in the identified area. In degeneration detection, the model will identify signs of degeneration while identifying where it is located. In severity assessment, the model will grade the severity of the degeneration from mild to severe which would understand the impact of the condition. In model evaluation, we would test the accuracy of the model and how well it performs with its prediction against the actual result to show the model's reliability. In final testing, we would experiment the model on unseen data to see how well it would perform in practical situations.

By training our model on the dataset, we are able to accurately classify the images while scoring the degeneration levels. This would boost the accuracy in diagnosis, ensuring consistent evaluations among practitioners while diagnosing degeneration earlier to improve in patient care.

II. Related Work

There have been several studies conducted using deep learning techniques to diagnose degenerative diseases in the spine. Convolutional neural networks have proven to be advantageous in this region, particularly with detecting features within MRI images. How these tools have been leveraged varies from study to study. In "Deep learning-based diagnosis of disc degenerative diseases (DDD) using MRI: A comprehensive review", the authors conducted a systematic review to assess DDD diagnosis through deep learning methodologies using MRI images (Hussain). It was found that convolutional neural networks (CNN) in addition to a few deep learning methods considerably outperformed traditional machine learning techniques, obtaining over 95 percent accuracy (Hussain). It was found that pooling layers in the model reduces dimensionality and prevents overfitting, this also cuts the cost of computations and allows for a better generalization when new data is applied. There is also a constant learning mechanism with CNN, which opens the door to optimize every layer of the neural network, thus improving the overall model efficiency. Lastly, one can apply a Rectified Linear Unit, a widely used activation function that gives the model non-linear properties.

This is widely useful for identifying minute and complex patterns such as those presented in medical data.

In another study, "An Automated Deep Learning Approach for Spine Segmentation and Vertebrae Recognition Using Computed Tomography Images" the goal was similar to this current project: develop an automated, deep learning method for correct spine segmentation and vertebrae recognition for diagnosis and surgical planning (Saeed). This process is extremely useful as the spine's anatomy is naturally complex, in addition to being encased by muscle and surrounding tissue. Two different deep learning techniques were proposed: CHASPPRAU-Net for spine segmentation and 3D MRU-Net for vertebrae recognition. The CHASPPRAU-Net Model was useful for extracting deep features from the CT images, while 3D MRU-Net employed an architecture that created 3D feature maps out of axial, sagittal, and coronal CT images which helped significantly with vertebrae recognition. Both methodologies achieved over a 90 percent dice score, a test that measures the overlap between predicted segmentation and actual labels (Saeed). With these results, on a broader scale, these tools can automate and optimize clinical diagnostic workflows by providing highly accurate 3D spinal models.

Facets of machine learning are being employed in all aspects of healthcare to extract features from large medical datasets, evaluate different combinations of therapies and treatments, and streamline routine tasks. As previously explored, there are several different approaches and methodologies that can be used. Support Vector Machines can be used on labeled image datasets where the features have been extracted. Convolutional neural networks and scale-invariant feature transforms are two ways for obtaining these features. In biomedical image classification tasks it is difficult to deal with high noise, pixel intensity variation and blurriness. One study proposed using a two-stage method to deal with these inconsistencies. In a paper titled "Biomedical image classification based on a cascade of an SVM with a reject option and subspace analysis" images were classified based on confidence scores, those with low confidence scores got reclassified in the second stage. Stage 2 utilized an analysis technique called "eigenfeature regularization and extraction" to focus on important features while ignoring unnecessary details that may have overwhelmed the process in Stage 1. "The proposed method significantly outperformed several competing state-of-the-art methods in terms of classification accuracy…The proposed cascade structure via an SVM with a reject option could be extended to other domains of image classification when two or more complementary features or classifiers could be used to jointly solve the problem" (Dongyun). In terms of streamlining SVM applications, leveraging computing power capabilities and power is another way to enhance the model, this time in terms of runtime. Algorithms using the MadReduce model from the Hadoop platform open the door for parallel processing and classification. With this approach there has been a significant reduction in runtime while achieving over 95 percent accuracy in some image classification processes. In one study's case, an 80,000 image SVM model had a runtime 1/5th of that of a single-node architecture algorithm (Cao). Advancements beyond the algorithms that look to engage all facets of both data science and computer engineering to move image classification in healthcare forward, faster.

III. Our Solution

A. Description of Dataset

The RSNA 2024 Lumbar Spine Degenerative Classification dataset is composed of 35.34 GB of Digital Image and Communications in Medicine image format and Comma Separated Values types, amounting to 147,320 files. This includes four files that contain the study condition description, sample images, descriptive labels for the training and test images, and bounding box coordinates for the training data. Each image contains a study id number to uniquely identify its characteristics in the corresponding tables. Each image was classified 4 ways: orientation/imaging process, location, disease and level of degeneration. The images are all collected via Magnetic Resonance Imaging (MRI) and labeled Sagittal T1, Sagittal T2/STIR and Axial T2. T1 and T2 refer to the imaging type where T1 tends to highlight anatomy and marks fluids in dark intensities while T2 focuses more on pathology where fluids are marked in brighter intensities. Next, the image was identified by its location in the intervertebral disc levels: L1/L2, L2/L3, L3/L4, L4/L5, and L5/S1. There were five different lumbar spine degenerative conditions that were observed: Left Neural Foraminal Narrowing, Right Neural Foraminal Narrowing, Left Subarticular Stenosis, Right Subarticular Stenosis, and Spinal Canal Stenosis. Lastly each sample was provided a severity score: Normal/Mild, Moderate, or Severe. Due to the vastness of this data, both in size and depth of characterization, it was decided to focus on one of each of the orientation/imaging process, location, disease classifications while building a model that will undertake the scoring of the level of degeneration.

Visualizing the spread of the data was the best way to determine which types of characterizations to isolate for the study. In Figure 1 below, the images are characterized by type of disease, then divided by its diseased location (right, left or neither) within each vertebrae. For example there is a histogram titled "Foraminal Distribution" where there are ten bars, one for each intervertebral disc level on the left and right. Then the count of each disc level is broken into three categories: Normal/Mild, Moderate and Severe. It was decided that in order to have a uniform dataset where the model is consistently looking at one boundary within the images, the grouping of data with a Subarticular Stenosis assessment was the best pick. This is due to the fact that the degeneration will occur within the same region.

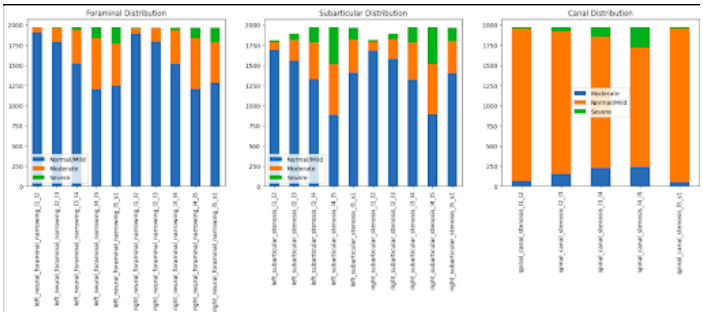


Fig. 1. Spread of RSNA 2024 Lumbar Spine Degenerative Classification

Next, it was decided that Axial T2 would be a plane view used to visualize the images. This is due to the main indicator

of Subarticular Stenosis being a clear narrowing of the spinal canal, which is best observed through the axial segmentation, no matter the level, as seen in Figure 2. Using the descriptive label file and the corresponding study id numbers, each training image that was diagnosed with Subarticular could be extracted from the main dataset to form the curated dataset that will be used for this study. This resulted in 1,285 dcm files being collected and stored in a folder for future use. The same process was done with the testing images so when the model is completely trained, it will only be tested using images that fall within the same training parameters i.e. Subarticular and Axial T2 slices. An example of this sample image can be viewed in Figure 3. Additionally, a new descriptive label file was generated for only the Subarticular and Axial T2 image slices as this will be easier to navigate and encode throughout the analysis process. The first step was downloading the data.

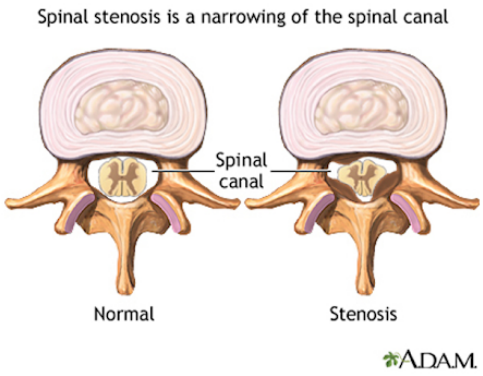


Fig. 2. Axial View of the Spinal Canal with Normal and Stenosis Conditions

This was achieved by utilizing the Kaggle API to download the "RSNA 2024 Lumbar Spine Degenerative Classification" dataset. These files included training images in a DCM format, each with a respective study and series ID, a CSV file that contained the series description used to identify Axial T2 images, and another CSV that contained key coordinates to draw boundary boxes around the targeted region. After downloading the data, it was unzipped and saved to the local system.

The next step was to explore and visualize the Kaggle dataset components to construct an ideal dataset for the study. Using a DICOM file reader, images were visualized. The metadata for the dataset, which was stored in a CSV file, was also loaded, and the contents were explored, noting the column names and distribution of labels. This was important as the locations of Axial T2 data were needed in order to isolate them and create a new dataset. To extract the Axial T2 series from the dataset, a specific filtering process was employed using the train series descriptions CSV file. This file contained information about each series in the dataset, including the series description. To isolate the Axial T2 series, the data was filtered to retain only rows where the 'SeriesDescription' column matched 'Axial T2'. This resulted in a new DataFrame, axial t2 df, containing only the entries corresponding to the Axial T2 series. This filtered DataFrame was then saved to a new CSV file, train series descriptions Axial T2.csv. This process ensured that only the Axial T2 series were extracted for further analysis and model training.

Since the study focused only on subarticular stenosis cases, the Axial T2 data needed to be merged with the DataFrame containing the case data. This process began by loading two CSV files: train Axial T2 filter.csv, which contained information about subarticular stenosis cases, and train series descriptions Axial T2.csv, which contained information about the Axial T2 series. These two DataFrames were merged based on the 'study id' column using a left merge, resulting in a new DataFrame called merged df. This left merge used the first input DataFrame as the base, discarding any non-matches in the second DataFrame. In this merged DataFrame, the 'series id' column was reordered to be placed immediately after the 'study id' column. The merged df was then saved as a new CSV file, Axial T2 data.csv, which combined the relevant information from both input files. Next, another CSV file, train label Left Subarticular L4 L5.csv, was loaded, which contained labels for left subarticular stenosis at the L4 L5 level. This file was merged with Axial T2 data.csv, using the 'series id' as the common column, to create a new DataFrame. This merging step ensured that each image in the train labels DataFrame was associated with its corresponding subarticular stenosis label. The 'severity' column in the train labels DataFrame was then filled with the 'left subarticular stenosis l4 l5' values from the merged data. Finally, the updated train labels DataFrame was saved as train label Axial T2 Data.csv, which contained a comprehensive dataset with labels for Axial T2 images related to subarticular stenosis at the L4 L5 level. To ensure data quality, any rows with empty 'severity' values were removed from train label Axial T2 Data.csv, and the cleaned dataset was saved as Axial T2 labels left l4 5.csv. This final CSV file provided a labeled dataset ready for further preprocessing and analysis.

The final stage of data preprocessing involved organizing the extracted regions of interest (ROIs) into a structured dataset for model training. This process utilized the subfolders ("Normal," "Moderate," and "Severe") created in the local system to categorize the images according to their severity levels. The Axial T2 labels left l4 5.csv file was then used to categorize the images into folders based on their severity. This involved iterating through each row of the DataFrame, extracting the study id, series id, instance number, and severity label. Before being placed into the appropriate "Normal," "Moderate," or "Severe" folder, each image's ROI was resized to 112 x 112 pixels. This resizing was done to standardize the input size for the machine learning models. The code accomplished this resizing using the cv2 resize function from the OpenCV library. The specific interpolation method used for resizing was cv2 INTER AREA, which was generally preferred for shrinking images, as it helped to preserve image details while reducing artifacts. The resizing process was embedded within a loop that iterated through each row of the Axial T2 labels left l4 5.csv file. After the ROI was cropped from the original image, the cv2 resize function was applied to resize the ROI to the desired 112 x 112 pixel dimensions. Once resized, the ROI, now in PNG format, was saved into the appropriate folder based on its severity label.

This organization created a dataset structured as a main folder containing three subfolders representing the three sever-

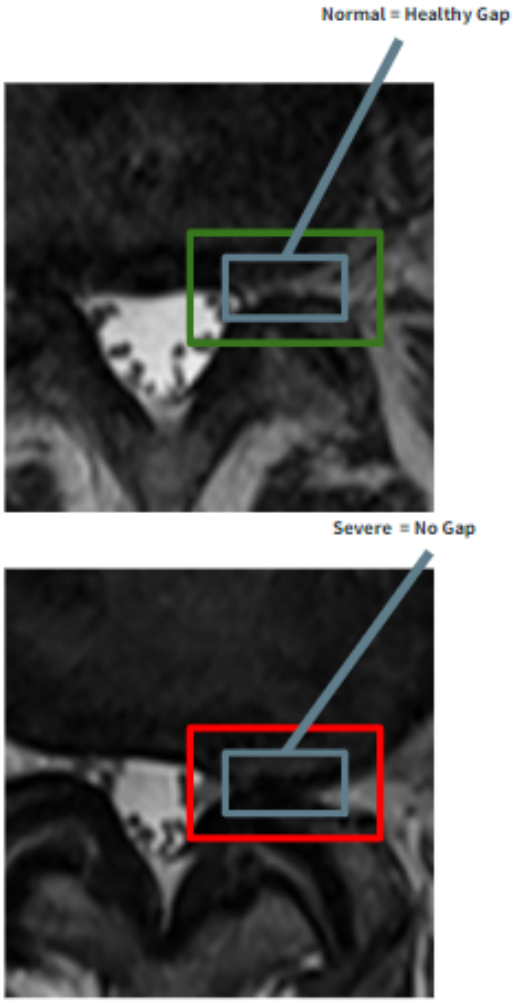


Fig. 3. Images of the Normal and the Severe Cases

ity levels. The final output of this preprocessing stage was a well-structured dataset ready for input into a machine learning model that could be split for training and evaluation.

B. Machine Learning Algorithms

Two machine learning algorithms we used in our project were the Convolutional Neural Network (CNN) and Support Vector Machines (SVM) to classify and assess spine degeneration. The CNN was designed to analyze images by breaking them into small parts to recognize patterns, helping the model identify specific details in the data. CNNs were appropriate for our problem because they identified the specific structure of the spine while detecting signs of degeneration. CNNs classified the degeneration with high accuracy by analyzing image pixels.The structure that CNN uses is convolutional layers, max pooling layers, flatten layer and dense layers. The key parameters were the image dimensions = 112*112, batch size = 32, optimizer = Adam, loss function = Binary Cross-Entropy and metrics= accuracy. The Support Vector Machines (SVM) algorithm was used for classification tasks, effectively separating data into distinct groups. SVMs were suitable for our problem as they classified spine degeneration with precision across categories such as normal, mild, moderate, and severe levels. The SVM uses linear kernel and C parameter of 1.

C. Implementation Details

The CNN was built using the Keras library in TensorFlow. After preprocessing the data, we created a training dataset enhanced through data augmentation techniques such as random rotation (rotation range=20), width and height shifts (width shift range=0.2, height shift range=0.2), and horizontal flips (horizontal flip=True). These transformations increased data variety, helping the model generalize better. Input images were resized to 112×112 (img height, img width=112, 112), rescaled to a range of 0-1 (rescale=1./255), and split into training and validation sets with a validation split of 0.2. The CNN architecture included convolutional layers, max-pooling layers, a flatten layer, and a dense layer. The ReLU (Rectified Linear Unit) activation function was used in the convolutional and dense layers to introduce non-linearity into the model. This non-linearity allowed the network to learn complex patterns and relationships in the data, which were essential for accurate classification. ReLU was selected because it is efficient and mitigates the gradient problem, which can occur with other activation functions like sigmoid. For the output layer, a sigmoid activation function was applied. The sigmoid function was crucial for binary classification, as it squashes the output into a range between 0 and 1, effectively representing the probability of each class. This probabilistic output made it easier to interpret the results and determine the likelihood of spine degeneration being present in an image. Convolutional layers detected patterns in images, max pooling layers reduced feature map sizes to focus on key features and improve efficiency, the flatten layer converted features for compatibility with the dense layer, and the dense layer combined features for classification. The model was optimized using Adam (optimizer='adam'), and binary cross-entropy (loss='binary crossentropy') was used as the loss function to calculate prediction accuracy. Training ran for 50 epochs (epochs=50) with early stopping (patience=5, restore best weights=True) to prevent overfitting. Performance during training and validation was visualized using accuracy and loss curves, allowing for an assessment of the model's generalization ability. Testing and validation were performed using the data split into training and validation sets with a ratio of 80:20 (validation split=0.2). Additionally, a train-test split was performed using scikit-learn, where the test size was set to 0.2, and the random state was fixed at 42 for reproducibility. This ensured that the model's performance was evaluated on unseen data to verify its ability to generalize.

Accuracy traces how well is the model predicting during the training and evaluation. The model is trained for a set number of epochs and early stopping is executed to stop the training if the validation loss doesn't improve to help avoid overfitting. We visualize the model's performance by plotting the accuracy and loss curves for training and validation. When training is concluded, the model is used to predict new images by loading and preprocessing the new images through the model to acquire a prediction. Support Vector Machines algorithm is designed for classification tasks. They work by identifying a specific feature that would separate data in distinct groups. SVMs are appropriate for our problem because it will help classify the spine degeneration with precision

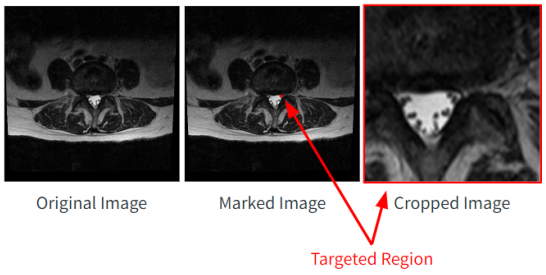


Fig. 4. Images showing differences between the original and the results after preprocessing.

between categories such as normal, mild, moderate and severe levels. The algorithm will ensure that any important differences between categories are recognized which would ensure the accurate classification. In our case, the model is built using the scikit-learn library. After loading the images and the labels, the data is divided into training and testing sets using the method train test split. We initialize the model with a linear kernel and a regularization parameter. After that, the algorithm is conditioned on the training data and it produces the labels for the test data. To assess the performance of the SVM model we acquire the accuracy and an evaluation report.

Binary Classfication includes two classes: Normal vs. Severe. For CNN, the accuracy computes to 85.21 percent while for SVM the accuracy computes to 94.39 percent. The Multi-Class Classfication includes three classes: Normal, Modederate and Severe. For the accuracy, CNN computes a value of 85.21 percent while for SVM it computes a value of 94.39 percent.

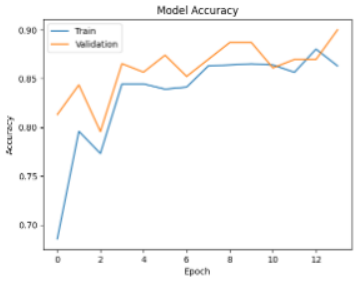
D. Comparison

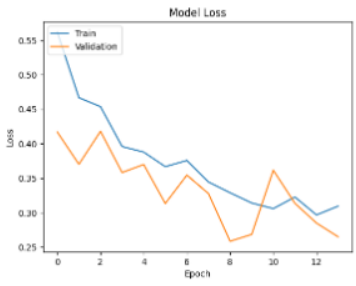
The first two models, a binary CNN and a binary SVM, were trained to differentiate between ”Normal” and ”Severe” cases. The binary CNN achieved a validation accuracy of 90% after 14 epochs of training. Its performance improved consistently throughout the training process, as evidenced by increasing accuracy and decreasing loss values with each epoch. The binary SVM demonstrated even higher accuracy, achieving 94.4% on the test set. It exhibited strong performance metrics, with a precision of 0.97, recall of 0.95, and F1-score of 0.96 for the ”Normal” class, and a precision of 0.89, recall of 0.93, and F1-score of 0.91 for the ”Severe” class.

The third model, a multiclass CNN, was trained to classify images into three categories: ”Normal,” ”Moderate,” and ”Severe.” This model achieved a validation accuracy of 70.88% after 16 epochs. The results indicated that distinguishing between three classes was more challenging than binary classification. Among the three models, the binary SVM demonstrated the best performance, achieving the highest accuracy of 94.4% on the test set. Its high precision, recall, and F1-scores underscored its effectiveness in distinguishing between ”Normal” and ”Severe” cases of spinal stenosis based on the provided image data.

For comparison, an existing solution on Kaggle, titled *“84% Accuracy in Spine Classification Using CNN”* (Ceng 49), employed a CNN algorithm and achieved an accuracy of approximately 75% for multi-class classification. In contrast,

Convolutional Neural Network





Support Vector Machine

	precision	recall	f-score	support
0	0.97	0.95	0.96	159
1	0.89	0.93	0.91	73
accuracy			0.9439	232
macro avg.	0.93	0.94	0.94	232
weighted avg.	0.94	0.94	0.94	232

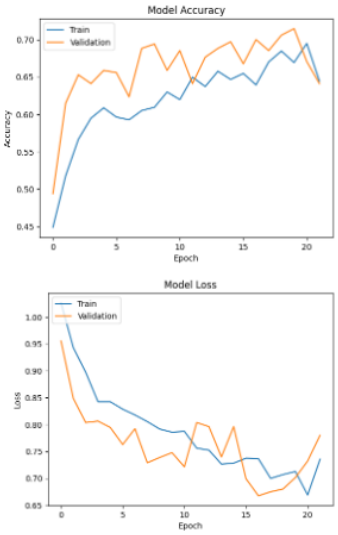
Fig. 5. Results provided for CNN and SVM algorithms using Binary Classfication

our binary classification models achieved significantly higher accuracy, with the binary SVM reaching 94%. In preprocessing and data augmentation, we utilized full images and applied extensive transformations to enhance the model’s ability to learn. The existing solution, in contrast, focused on small parts of the images, potentially missing important details. Furthermore, our approach addressed class imbalance effectively, ensuring robust performance across all categories, while the existing solution lacked this feature, leading to poorer performance on underrepresented groups. These differences highlighted that our models outperformed the existing solution in both accuracy and overall robustness.

IV. Future Directions

The sources showcase a project focused on classifying spinal stenosis severity using SVM and CNN models, highlighting their potential for future applications in healthcare. One promising application is in Computer-Aided Diagnosis (CAD) systems. The high accuracy of the binary SVM, reaching 94.4 percent in differentiating between ”Normal” and ”Severe” cases, makes it a valuable tool for assisting radiologists in diagnosing spinal stenosis. Integrated into CAD systems, these models could pre-screen images, identify potential cases requiring attention, and offer a second opinion, ultimately improving diagnostic accuracy and efficiency. Another key application lies in severity assessment and treatment planning.

Convolutional Neural Network



Support Vector Machine

	precision	recall	f-score	support
0	0.77	0.77	0.77	158
1	0.48	0.52	0.50	108
2	0.68	0.59	0.63	76
accuracy			0.6520	342
macro avg.	0.64	0.63	0.63	342
weighted avg.	0.66	0.65	0.65	342

Fig. 6. Results provided for CNN and SVM algorithms using Multi-Class Classfication

Accurate classification, enabled by these models, is crucial for guiding treatment decisions. Clinicians can use the model's output to determine whether conservative management or surgical intervention is necessary.

There are several areas for future development to improve the accuracy of the models. First, more layers and convolutions can be added to the model, which will allow it to detect more features. Next, it can be expanded to work with the entire RSNA dataset, incorporating different image orientations, locations, and diseases. Overall performance can be improved by using advanced data augmentation, leveraging pre-trained CNN models, and combining the strengths of CNNs with SVMs. Automating the ROI detection and cropping process using models like RCNN or YOLOv8 is another specific area for enhancement. Finally, there is an opportunity to incorporate clinical data and refine the methods for multi-class classification to enhance accuracy.

The potential for personalized medicine is also noteworthy. With further development, these models could be customized for individual patients, factoring in their age, medical history, and other imaging findings to create more tailored treatment plans. This personalized approach holds the promise of improved patient outcomes. Moreover, this project contributes to the broader field of spinal stenosis research. The dataset used to train the models, along with the trained models themselves, can be utilized to investigate new imaging features, develop novel classification algorithms, and validate existing research findings. Additionally, the models could be implemented in telemedicine platforms, allowing for remote diagnosis and severity assessment of spinal stenosis. This application could significantly improve access to specialized healthcare, particularly in areas with limited resources.

However, several considerations should be addressed for successful future applications. Ensuring the generalizability of the models is crucial. Since they were trained on a specific dataset focused on left subarticular stenosis at the L4-L5 level, validation on larger and more diverse datasets, encompassing various spinal levels and stenosis types, is essential. Another important factor is the explainability and interpretability of the models' predictions. While they demonstrate promising performance, understanding how they arrive at their conclusions is vital for clinical acceptance. Finally, seamless integration of these models into existing clinical workflows is key for practical implementation. This would require user-friendly interfaces and efficient data management systems to facilitate adoption by healthcare professionals.

V. Conclusion

In conclusion, our report addressed the classification of lumbar spine degeneration using CNN and SVM. The SVM performed well in binary classification, achieving an accuracy of 94 percent. However, in multi-class classification, both CNN and SVM encountered challenges, with accuracies ranging from 65 percent to 70 percent. For multi-class tasks, CNNs were better suited due to their ability to automatically learn image features, while SVMs demonstrated superior performance in binary classification tasks.

In the future, we can use the entire RSNA dataset plus the CNN model would best because it performing well in detecting and differentiating features. Additionally, incorporating clinical data could further improve the accuracy of predictions. Future efforts could focus on refining methods to handle multi-class classification challenges and boost overall accuracy.

REFERENCES

Cao, J., Wang, M., Li, Y., Zhang, Q. (2019). Improved support vector machine classification algorithm based on adaptive feature weight updating in the Hadoop cluster environment. PloS one, 14(4), e0215136. https://doi.org/10.1371/journal.pone.0215136

Dongyun, L., Sun, L., Toh, K.-A., Bo Zhang, J., Lin, Z. (2018, March 12). Biomedical Image Classification based on a cascade of an SVM with a reject option and subspace analysis. Computers in Biology and Medicine. https://www.sciencedirect.com/science/article/pii/S0010482518300568

Hussain, M., Koundal, D., Manhas, J. (2023, January). Deep learning-based diagnosis of disc degenerative diseases using MRI: A comprehensive review. Computers and Electrical Engineering. https://www.sciencedirect.com/science/article/pii/S004579062200739X

Saeed, M. U., Dikaios, N., Dastgir, A., Ali, G., Hamid, M., Hajjej, F. (2023, August 12). An automated deep learning approach for spine segmentation and vertebrae recognition using computed tomography images. MDPI. https://www.mdpi.com/2075-4418/13/16/2658

Ceng49. "84% Accuracy in Spine Classification Using CNN." Kaggle. www.kaggle.com/code/ceng49/84-accuracy-in-spine-classification-using-cnn.