A Novel Approach for a Three-Way Classification of Lumbar Spine Degeneration using Multimodal Architecture

Ahmed Abdullah[1,*], Burhan Ahmed[1,*], Ibtehaj Ali[1,*], and Tahmooras Khan[2]

[1]FAST-NU, School of Computing
[2]FAST-NU, School of Computing
[*]These authors contributed equally to this work.

September 24, 2024

Abstract

This paper discusses how we can use MRI scans to identify problems with the lower back, specifically conditions that cause pain and discomfort. We are devising a computer algorithm based on machine learning and multimodal input that automatically classify various back issues, such as narrowing of the spinal canal and nerve passageways. We have a labeled dataset of MRI images to highlight the affected areas and how severe the problems are. Our goal is to train a computer model using these images so that it can accurately detect and classify these back conditions, helping doctors make better and faster diagnoses.

1 Introduction

Degenerative spine diseases, particularly in the lumbar region, are a leading cause of disability and pain worldwide, with over 266 million cases reported each year [12]. These conditions, such as spinal canal stenosis and foraminal narrowing, are commonly diagnosed using MRI imaging. Manual diagnosis is time-consuming and requires expertise. To address this, we aim to automate the classification process using machine learning techniques, particularly convolutional neural networks (CNNs), which have shown success in image-based medical diagnoses [1] [2] [3].

In this work, we develop a model to automatically detect and classify various forms of lumbar spine degeneration based on labeled MRI images. This can potentially assist radiologists in improving diagnosis accuracy and efficiency.

2 Problem Statement

The challenge addressed in this research is the automatic classification of lumbar spine degeneration conditions from MRI scans. Existing diagnostic techniques depend on manual labeling and expert evaluation, which are both time-consuming and subjective. The goal is to build a deep learning-based classification model that can accurately identify and classify different degeneration types, such as spinal canal stenosis and foraminal narrowing, across multiple patients and spine levels.

3 Dataset Description

The dataset consists of MRI scans from multiple patients, with labels corresponding to various forms of lumbar spine degeneration. The labeled conditions include:

- Spinal Canal Stenosis (red)
- Right Neural Foraminal Narrowing (pink)
- Left Neural Foraminal Narrowing (pink)
- Left Subarticular Stenosis (yellow)
- Right Subarticular Stenosis (yellow)

Each MRI image is associated with specific labels marking the region of degeneration and its severity. The images are stored in DICOM format, and the metadata provides details such as spine levels (L1-L5) and degeneration components.

The dataset was initially normalized to rescale the MRI images for consistency during the training phase. However, to improve feature extraction and preserve important image details, we have applied *denormalization* techniques during preprocessing. This helps restore spatial and intensity information critical for detecting subtle degeneration patterns in MRI scans.

Furthermore, the dataset was originally organized into separate dataframes, including base data, label data, and description data. These dataframes have been *merged* to create a unified structure. This merging process was carried out to ensure we could find a common dataframe that is available in both the training and testing phases. By merging the dataframes, we streamline the data input pipeline, making the same structured information accessible to the model during both phases.

3.1 Data Available During Training

- **Base Data:** Contains descriptions of the degeneration severity for each spine level (L1-L5) and condition (e.g., stenosis or foraminal narrowing).

Available only during training to provide insights into the different patterns of degeneration across patients.

- **Label Data:** Consists of precise coordinate-based annotations that indicate the location of degeneration, such as stenosis or foraminal narrowing, for specific spine levels. These annotations guide the model in learning to localize the degeneration points.

3.2 Data Available During Both Training and Testing

- **Description Data:** Metadata such as *Study ID*, *Series ID*, and MRI scan parameters (e.g., angle and resolution). This information is used in both training and testing phases, ensuring that the model has consistent metadata input during evaluation. By merging the dataframes, we were able to ensure that this common structure (i.e., the description data) is available for both phases.

3.3 Data Available During Testing

- **Unlabeled MRI Images:** During testing, only the MRI images and description data are available for the model to predict degeneration. The base data and label data used in training are not provided. The model must classify the degeneration type, location, and severity based solely on the images and associated metadata.

4 Methodology

Our approach involves developing a convolutional neural network (CNN) to classify spine degeneration based on the labeled MRI dataset. The architecture of our model includes several convolutional layers, followed by max-pooling and fully connected layers.

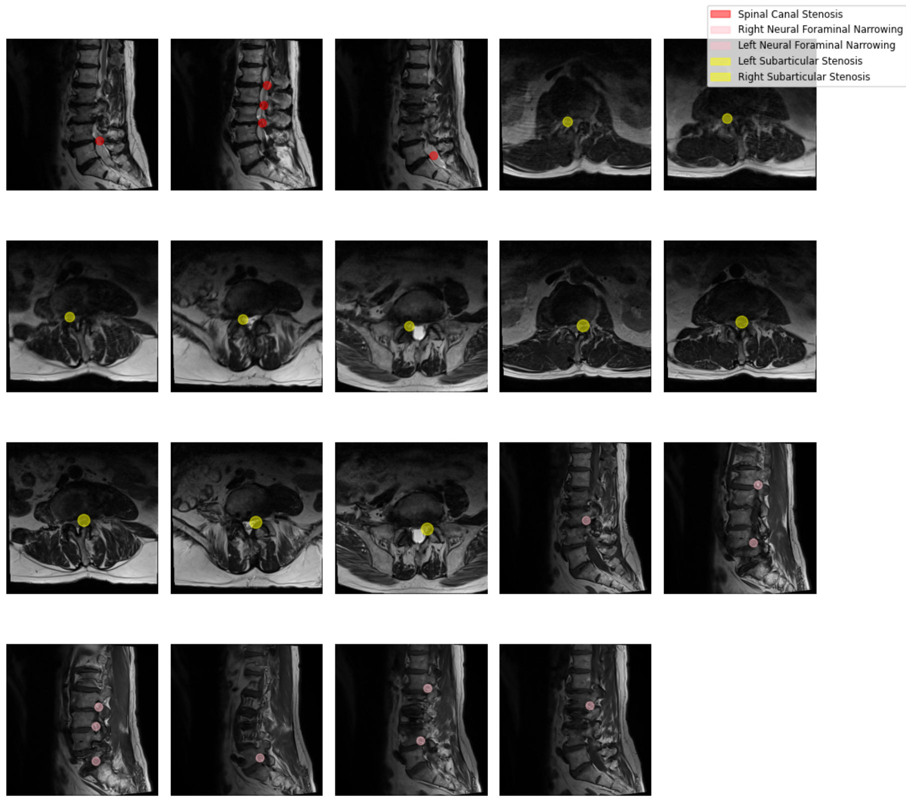

Figure 1: MRI scans of the lumbar spine with annotations for various degeneration types: Spinal Canal Stenosis (red), Neural Foraminal Narrowing (pink), Subarticular Stenosis (yellow)

The key steps in our methodology include:

1. Preprocessing the MRI scans, including normalization and augmentation techniques.
2. Annotating regions of interest (ROI) using the provided labels (see Figure 1).
3. Training a CNN on the labeled data with a softmax classifier to distinguish between different degeneration types.
4. Evaluating the model's performance using metrics such as accuracy, precision, and recall.

We experiment with different architectures, including:

- **ResNet50** [7]
- **EfficientNet** [5]
- **3D U-Net** [8]
- **CLIP** [6]
- **BERT** [11]

These models are fine-tuned using transfer learning and trained on the annotated dataset.

5 Conclusion

In this work, we have presented an automated method for classifying lumbar spine degeneration conditions using CNNs. By leveraging annotated MRI scans, we aim to assist radiologists in diagnosing spine degeneration more effectively. Future work could involve refining the model's performance and expanding the dataset to include more diverse spine conditions.

References

[1] Saeed, M. S., Nawaz, S., Zaheer, M. Z., Khan, M. H., Nandakumar, K., Yousaf, M. H., Sajjad, H., De Schepper, T., and Schedl, M., "Modality Invariant Multimodal Learning to Handle Missing Modalities: A Single-Branch Approach," *arXiv preprint arXiv:2408.07445*, 2024.

[2] Saeed, M. S., Nawaz, S., Khan, M. H., Zaheer, M. Z., Nandakumar, K., Yousaf, M. H., and Mahmood, A., "Single-branch Network for Multimodal Training," *arXiv preprint arXiv:2303.06129*, 2023.

[3] Osama, S., Zafar, K., and Sadiq, M. U., "Predicting Clinical Outcome in Acute Ischemic Stroke Using Parallel Multi-Parametric Feature Embedded Siamese Network," *MDPI Diagnostics*, vol. 10, no. 11, 2020.

[4] Tamoor, M., Naseer, A., Khan, A., and Zafar, K., "Skin Lesion Segmentation Using an Ensemble of Different Image Processing Methods," *MDPI Diagnostics*, vol. 13, no. 16, 2023.

[5] Khojasteh, H. A., Ansari, E., Razzaghi, P., and Karimi, A., "Deep Multimodal Image-Text Embeddings for Automatic Cross-Media Retrieval," *arXiv preprint arXiv:2002.10016*, 2020.

[6] Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., and Sutskever, I., "Learning Transferable Visual Models From Natural Language Supervision," *arXiv preprint arXiv:2103.00020*, 2021.

[7] He, K., Zhang, X., Ren, S., and Sun, J., "Deep Residual Learning for Image Recognition," *arXiv preprint arXiv:1512.03385*, 2015.

[8] Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., and Ronneberger, O., "3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation," *arXiv preprint arXiv:1606.06650*, 2016.

[9] Liaqat, M. I., Nawaz, S., Zaheer, M. Z., Saeed, M. S., Sajjad, H., De Schepper, T., Nandakumar, K., Khan, M. H., and Schedl, M., "Chameleon: Images Are What You Need For Multimodal Learning Robust To Missing Modalities," *arXiv preprint arXiv:2407.16243*, 2024.

[10] T. Richards et al., "RSNA 2024 Lumbar Spine Degenerative Classification," *Kaggle*, 2024. Available: `https://kaggle.com/competitions/rsna-2024-lumbar-spine-degenerative-classification`.

[11] Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805.*

[12] Ravindra, V. M., Senglaub, S. S., Rattani, A., Dewan, M. C., Härtl, R., Bisson, E., Park, K. B., & Shrime, M. G. (2018). Degenerative lumbar spine disease: Estimating global incidence and worldwide volume. *Global Spine Journal*, 8(8), 784-794. doi: 10.1177/2192568218770769. PMID: 30560029; PMCID: PMC6293435.