Svedberg Open
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Deep Transfer Learning for Automated Multi-Class Classification of Spinal Disorders Using Benchmark and Clinical X-ray Datasets

Kumari Bhawana[1*], I. Mukherjee[2], Amritanjali[3]

[1*]Research Scholar, Department of Computer Science and Engineering, Birla Institute of Technology (BIT), Mesra, Ranchi, Jharkhand, India
Email: phdcs10002.20@bitmesra.ac.in
[2-3]Department of Computer Science and Engineering, Birla Institute of Technology (BIT), Mesra, Ranchi, Jharkhand, India
*Corresponding Author

Abstract: Among the most common spinal disorders are scoliosis, spondylitis, and spondylolisthesis. It is an essential reason for chronic pain and physical disability, and its accurate and timely diagnosis is essential. In contemporary years, deep learning technologies have demonstrated promising applications in automated medical image analysis. Despite this, the breadth and uniformity of publicly available data remain problems, as does analysis. Challenges for developing robust diagnostic models. This work proposes a deep transfer learning system that compiles spine X-ray data from benchmarks and hospitals to enhance the database, thereby standardizing it and improving classification performance. We used four state-of-art convolutional neural networks with a total of 131 million parameters. Four different network architectures, namely ResNet-152, DenseNet-201, Densenet-121, and VGG-16, have been trained, and the support was evaluated for both three and four-class spinal disorder classification. Storage and retrieval techniques. Data preprocessing and retrieval techniques. To mitigate overfitting and improve model inference, a method was employed. Accuracy, precision, recall, F1score, confusion matrices, and receiver operating characteristic (ROC) analysis were used to assess the model's performance, with the area under the curve (AUC) serving as the main metric. The experimental results show the ResNet-152 consistently achieved higher accuracy than the other architectures, with a mean accuracy of 98.95% for three-class classification. In classification, the highest accuracy of 95.17% and 95.16% for four-class classification, with strong performance across precision, recall, F1-score, and ROC/AUC. The skip and residual learning strategy are responsible for the improved performance. Connection of ResNet-152, which enables efficient feature extraction and stable optimization in deep networks. Moreover, the inclusion of standardized benchmark and clinically obtained datasets improved model robustness and generalization across various spine diseases. The proposed framework provides a reliable computer-aided diagnostic approach that can help clinicians detect spinal diseases early and facilitate the development of functional artificial intelligence systems for medical imaging applications.

Keywords - Deep Learning, Transfer Learning, ResNet-152, DenseNet, VGG-16, Spinal Disorders, Medical Image Classification, Benchmark Dataset Standardization, ROC-AUC, Computer-Aided Diagnosis.

1. INTRODUCTION

Spinal cord disorders have a significant impact on human life quality. Spinal cord disorders encompass various conditions affecting the spinal cord, including Scoliosis, Spondylitis, Spondylolisthesis, oncology, osteoarthritis, infection, degenerative disorders and adult spinal disfigurement. These disorders affect human life and manifest as radicular pain, back pain, neurogenic claudication and atypical limb pain. Among these spinal cord disorders, Spondylitis is an infection in the spinal vertebra. Spondylolisthesis is associated with the displacement of the spinal vertebra. Scoliosis is associated with abnormal lateral curvature of the spine [1,2]. Localization of spinal cord portions in medical images is still difficult, especially in cases involving anatomical abnormalities like additional or transitional lumbar vertebrae [2]. To build efficient computer models, proper labelling and separation of spinal structures are critical, as any errors in spinal imaging can result in a wrong diagnosis [3]. Significant advancements in DL and AI technologies include enhanced medical image processing, facilitating the early diagnosis of spinal pains [4]. Methods involving machine learning techniques, such as CNN and transformer, have proved promising in spinal disease classification and segmentation, showing good results in these areas [5-6]. However, there are issues in dealing with variances in spinal anomalies, which are critical in generalizing the model using different datasets. This analysis aims to develop a more accurate model for interpreting spinal images, helping detect spinal problems and pinpoint problem areas more precisely. With the implementation of efficient of efficient machine learning techniques, these methods will help in improving decision-making in clinical scenarios, which will result in maximum patient benefits [7]. Deep learning models, especially CNN, have been extremely promising for implementation in medical image processing and analysis. Among all models, ResNet-152 has proved its depth, high-performance features and ability to overcome vanishing gradient problems through skip connections.

The most compelling reason behind this research stems from the severity of the impact associated with spinal diseases and the difficulties faced during accurate and quick detection of such diseases. Inconsistencies and localization errors are also common problems associated with most models, especially when anatomical abnormalities occur during such conditions. This problem needs to be addressed through more complex designs of deep learning models so that accurate and reliable diagnostic tools can be provided to patients, reducing the chances of misdiagnoses and invasive treatments, and hence providing a boost to healthcare through quicker treatments and lower costs. Figure 1 present a comparative visualization of spinal X-ray images collected from benchmark datasets and hospital-acquired datasets, illustrating various spinal disorders. The figure highlights differences in image quality, contrast, anatomical alignment and pathological characteristics across datasets. This shows the complexity of spinal deformities and anomalies, which include differences in vertebral anatomy and curvature characteristics. This is essential when determining the generalization ability of predictive and classification models.

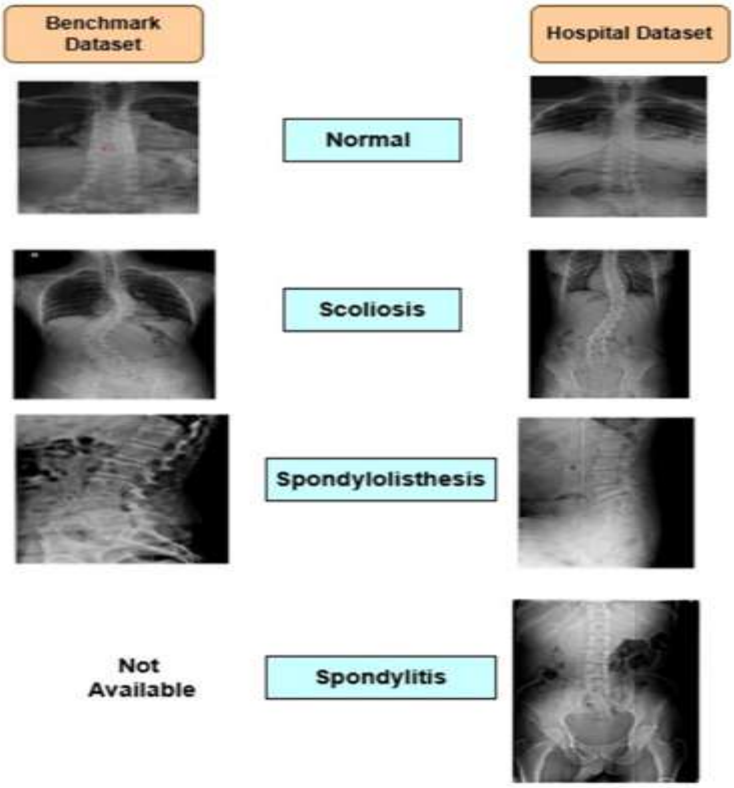

Figure. 1 Comparison of spinal X-ray Images from benchmark and hospital datasets from different spinal disorders

LITERATURE REVIEW

The relationship between pathology, diagnosis, and treatment in the context of spinal health is complex and ever-changing. The significance of muscles in maintaining the health of the spine and the frequency of low back pain (LBP) is highlighted in Noonan and Brown. (2021), illuminating the complex link between paraspinal muscle pathophysiology and spine-related illnesses. On the other hand, Frost et al. (2019) describes the issue of disc degeneration, which is a prevalent problem associated with discogenic lower back pain. Despite being widespread and having a significant impact, disc degeneration cannot be prevented, as well as regeneration stimulated in intervertebral discs using conventional therapy. This highlights the importance and urgency for new tissue engineering approaches to solve this growing problem, and information on the pathophysiology of intervertebral disc deterioration and its relation to discogenic pain Mohd et al. (2022).

Identification of the key indicators and causative factors of degeneration takes place after conducting a detailed analysis, thus making a ground for treatment strategies, although investigation and radiography have grown less common for their part, Santiago & Bossini (2020) highlight the ongoing value of radiology for identifying various spine conditions. This paper aims at assisting radiologists in improving their skills concerning interpreting and identifying the radiological features of various spinal disease conditions. In addition, a complex view on the treatment of lumbo-sacral spinal stenosis-a condition that provokes neurogenic claudication and back pains- is explored by Lee et al. (2020). on the one hand, surgery is proven to prevent falls and alleviate the disease symptoms, although the patient's decision faces some issues, such as complications. The framework proposed by Masood et al. (2022) is an approach to examining the lumbar spine that combines systematic assessment with a focus on key structures. Together with new disease classification methods, there is an independent retrieval of spinal measures. This method, based on the semantic segmentation of vertebral bodies, could be used to. Is quick in making decisions, especially with regard to the diagnosis of diseases such as Spondylolisthesis and lumbar spine disorders. Represents lordosis well with great accuracy. At the same time, Ma et al. (2020) explored the feasibility of developing a deep learning models to detect lesion from MRI images in cervical spine diseases.

They demonstrate that their study accurately detects regions using a faster-region convolutional neural network (Faster R-CNN), which is faster than previous methods. Kubaisi and Khamiss (2021) discuss the approach to the identification of lesions and support the radiologist and spine surgeon in diagnosis. Transfer learning and region-of-interest (ROI) methods were used to enhance classification accuracy. The results they achieved demonstrate the capability of deep learning methods to handle scarce data while highlighting the importance of data proximity for optimal results. In the following section on cervical spine disorders, Abuhayi et al. introduce. They outperform conventional CNNs, providing an efficient and effective tool for accurate disease diagnosis in resource-limited environments. Lastly, Bezabh et al. (2024) focus on cervical spondylosis myelopathy (CSM), emphasizing the significance of early and precise diagnosis by classifying X-ray images. The technique that they proposed for extracting features using CNNs has achieved very accurate results in diagnosing and classifying cervical spine diseases, showing how promising deep learning is for improving the process of diagnosis. By comparing various techniques, which had a sensitivity rate of 54.9% and a specificity rate of 94.1 %, the Aidoc AI technology showed that it had a low diagnostic accuracy when tested by Voter et al. (2021) for detecting cervical spinal fractures from CT images.

In the study done by Yeh, et al. (2022), the authors designed a decision support system using the ResNet-50 algorithm, which can assist doctors with less experience in identifying spinal fractures from MRI scans. The study included 190 patients, and the results were compared by fourth and first-year residents, as well as a senior MSK radiologist. The acuity of the first-year resident, improved from 78% to 94% while the particularity improved from 61% to 91% using the deep learning algorithm, which indicates the effectiveness of the algorithm in improving diagnostic accuracy. In the study done by Gautam Sharma (2020), the authors have explained in detail the investigation of deep learning algorithms in the diagnosis of eight neurological and neuropsychiatric conditions. The authors, in their review article, which included 136 articles, identified the research gaps in the diagnosis of the conditions, particularly in the diagnosis of migraine and stroke, as well as in the investigation of more complex deep learning algorithms, including Deep belief networks and Restricted Boltzmann Machines. In the study done by Masud, Mondal (2020), the authors made use of the univariate feature selection method as well PCA in the preprocessing stage in the machine learning algorithm in the prediction of spinal anomalies. In the study, the authors made use of the SVM, logistic regression and the bagging ensemble algorithm, which were tested on the 310 data samples from the Kaggle dataset. As indicated in the study, the accuracy of the data in the training set is 86.72%, while the accuracy in the test set is 86.96%, which indicates the success of the algorithm, as the results are better compared to the other algorithms, including the high recall value and low miss rates.

In fact, Altun et al. (2023) proved the possibility of applying an automated diagnosis like LSS through the use of the LSS-VGG16 model by stating that a classification approach for LSS, based on deep learning, could be developed. With a categorization ability as high as 87.70% in their research, they provided a good tool for medical professionals in the avenue. Ma et al. (2020) researched cervical spinal disease through deep learning-based detection of lesions from MRI images. They achieved pleasurable prediction accuracy and prediction speed by providing the combination of ResNet-50 and VGG-16 networks with Faster R-CNN as a very promising and potential tool for lesion detection and classification by radiologists and spine surgeons. In another study, Masood et al. (2022) proposed a lumbar evaluation framework, an automatic extraction of spinal measures, and condition classification systems such as abdominal lordosis (LL) assessment and spondylolisthesis. Their system, which overcomes the subjectivity and problems in manual evaluation approaches, achieved excellent accuracy in both disease classification and measure extraction by applying semantic segmentation with ResNet-UNet. Moreover, saravagi et al. (2022) with an optimized deep learning model based on VGG16 and Inception V3 model architectures, explored the classification of X-ray radiographs for spondylolisthesis diagnosis. The developed piece of work outperformed all prior work and, with commendable accuracy, is capable of providing luminosity for early lumbar diagnostics.

MATERIALS AND METHODS

3.1 Data Collection
One of the most important processes in the construction of foundations for building is the Data Acquisition process, as proposed in the methodology for spine disorders diagnosis using deep learning models. Comprises collecting heterogeneous and representative datasets of different types of ortho-spine data. These modalities offer different views and rates of detail, offering overall details regarding the spine's structure and possible anomalies. With 2-Dimages of spine, planar radiographs offer a general view of alignment of the spine and possible deformities such as MRI scans improved soft tissue contrast and fine visualization of spinal cord and never root compression, which help in the detecting of conditions such as herniation of the disc and spinal tumours. Spinal disorders can impact individuals from various demographics; diversity and inclusivity must be ensured while gathering data representing various different ages, genders and ethnicities. Also, for training and testing models, there must be a balanced dataset ready, thereby ensuring gathering images from healthy individuals and individuals diagnosed with spinal disorders as part of data gathering. In this proposed methodology, the total X-ray dataset, including hospital and benchmark (Kaggle) data. The benchmark dataset contributes 338 samples, 71 labelled as standard, 188 as scoliosis and 79 as spondylolisthesis. In contrast, the verified hospital (USHA hospital Muzaffarpur, Bihar) dataset contains 1,004 samples, distributed as 387 scoliosis cases, 440 spondylitis vases and 177 spondylolisthesis cases.

3.2 Data Preparation and Augmentation
A baseline spinal disease dataset was obtained during the data preparation process involving images representing Spondylitis, Scoliosis and Spondylolisthesis. Augment training dataset robustness and generalizability, rescaling, shearing, zooming, and horizontal flipping operations were applied. Shearing and zooming of images introduced perspective and scale variations, imitating realistic usage and the model's adaptability to diverse input images. Horizontal flipping of images resulted in a duplicate of the dataset, as the images are flipped and mirrored.

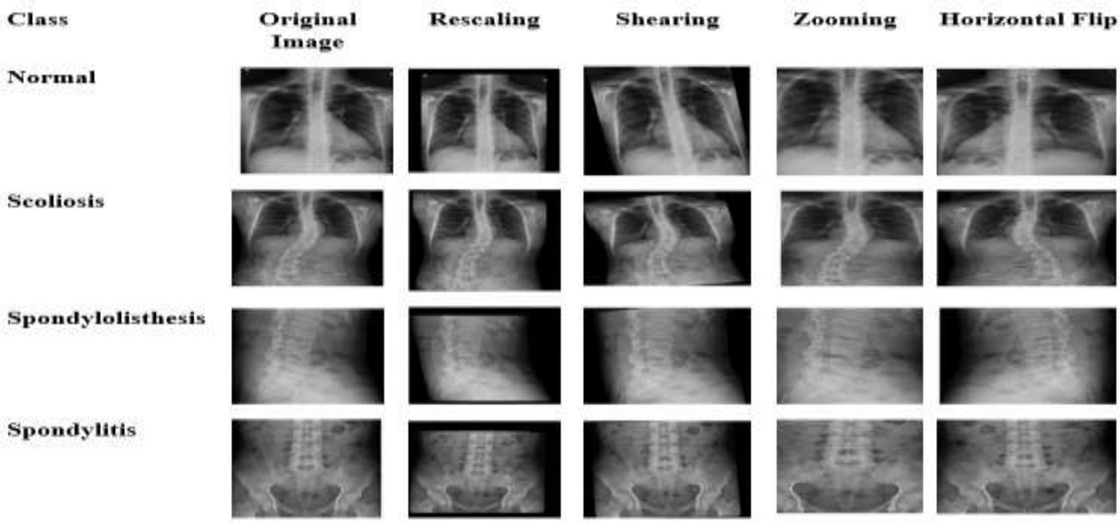

Figure 2. Representative of data augmentation applied to spinal X-ray images from the benchmark and hospital datasets.

Figure 2 presents representative examples of the original and augmented spinal X-ray images for the four classes: Normal, Scoliosis, Spondylolisthesis, and Spondylitis. Data augmentation using rescaling, shearing, zooming, and horizontal flipping was performed to improve model generalization and minimize overfitting.
3.3 Proposed Mechanism
Figure 3 shown a model that describes the process and testing a machine-learning model for medical data analysis. The model adopts a systematic pipeline:

- **Dataset Preparation & Collection:** Compiling medical images and patient information from various sources to create an effective dataset.

- **Data Augmentation & Preprocessing:** Improving the quality of data using normalization, resizing, and augmentation methods.
- **Deep-Learning Model:** Utilizing pre-trained architecture such as DenseNet-201, VGG- 16, DenseNet-121 and ResNet-152 to use in effective feature extraction.
- **Training & Fine-tuning of Model:** Employing learning and hyperparameter tuning to enhance the accuracy of models.
- **Model Test & Evaluation:** Evaluation of the model by applying metrics like accuracy, precision, recall, F1 score and visualization plots like confusion matrix and graphs.

This process ensures the accuracy and consistency of the model while dealing with the challenges of medical data analysis. This work can be a good starting point to achieve accuracy and reliability in medical data analysis.

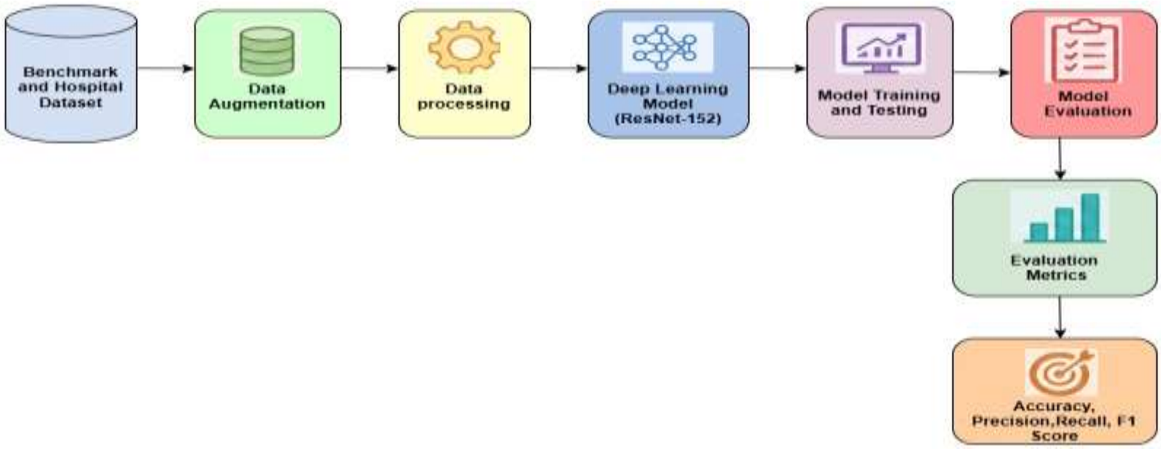

Figure3. Workflow of the proposed ResNet-152 based medical image classification system

Algorithm Proposed Deep Transfer Learning Framework
Begin
Step 1: Acquire spinal X-ray images from the benchmark and hospital datasets.
Step 2: Preprocess the images by resizing and normalizing the pixel values.
Step 3: Apply data augmentation techniques, including rescaling, shearing, zooming, and horizontal flipping.
Step 4: Divide the dataset into training (70%), validation (15%), and testing (15%) subsets.
Step 5: Initialize the transfer learning models (ResNet-152, DenseNet-121, DenseNet-201, and VGG-16).
Step 6: Replace the original classification layer with a new fully connected layer followed by a Softmax output layer.
Step 7: Compile each model using the Adam optimizer and categorical cross-entropy loss function.
Step 8: Train the models using the training dataset and validate them using the validation dataset. Apply Early Stopping and Model Checkpoint to obtain the optimal model.
Step 9: Evaluate the trained models on the testing dataset using Accuracy, Precision, Recall, F1-score, Confusion Matrix, and ROC-AUC.
Step 10: Classify unseen spinal X-ray images using the best-performing trained model.
End

RESULTS

4.1 Three-Class Classification Results

The performance of ResNet-152 is further analysed through confusion matrices and training validation graphs. These visual representations provide insights into model classification performance and training stability.

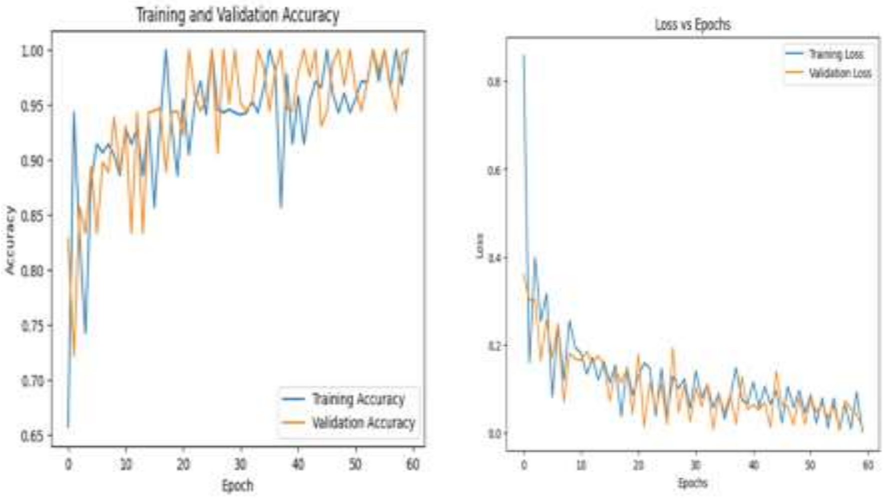

Figure 4. Training and validation accuracy and loss curves for 3- class classification

Training and validation accuracy and loss graph of the ResNet-152 model for the 3- class spinal disorder classification shown by Figure 4. The training and validation accuracy increase consistently throughout the training process, while both loss curves gradually decrease and converge with a small gap, indicating stable learning, good convergence, and minimal overfitting.

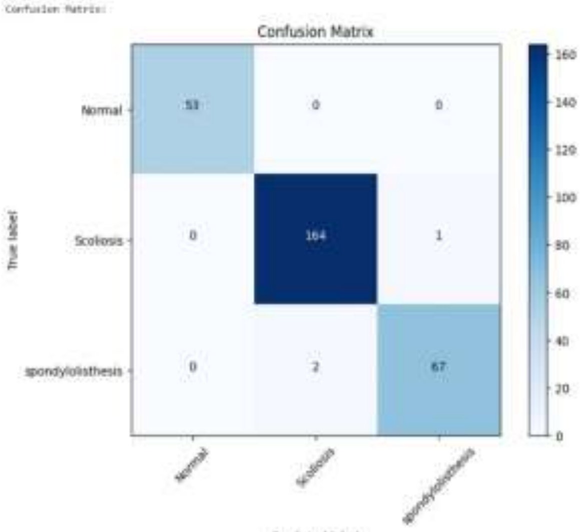

Figure 5. Confusion Matrix

Figure 5 illustrate the confusion matrix obtained using Resnet-152 model for three-class classification. Most samples were correctly classified with very few misclassifications, demonstrating the model's excellent discriminative capability and reliable performance in distinguishing spinal disorders. Table 1 summarizes the performance of the evaluated CNN models for the three-class classification in term of mean accuracy, Maximum accuracy, minimum accuracy and standard deviation. Table 2 compares the F1-score, precision and recall of the evaluated CNN models for three-class classification. ResNet-152 obtained the best performance across all evaluation metrics, demonstrating its effectiveness in spinal disorder classification.

Table 1. Result of 3-Class Classification

CNN Model	Average Accuracy (%)	Highest Accuracy (%)	Lowest Accuracy (%)	Standard Deviation
DenseNet-201	96.16	100	72.22	19.19
VGG-16	82.29	100	56.32	38.26

DenseNet-121	94.42	100	50.0	22.94
ResNet-152	98.95	100	72.22	10.17

Table 2. Result of 3-Class Classification

CNN Model	F1 Score (%)	Precision (%)	Recall (%)
DenseNet-201	95.66	96.33	95.33
VGG-16	83	83	78.33
DenseNet-121	94	96.66	92
ResNet-152	97.33	97.66	97

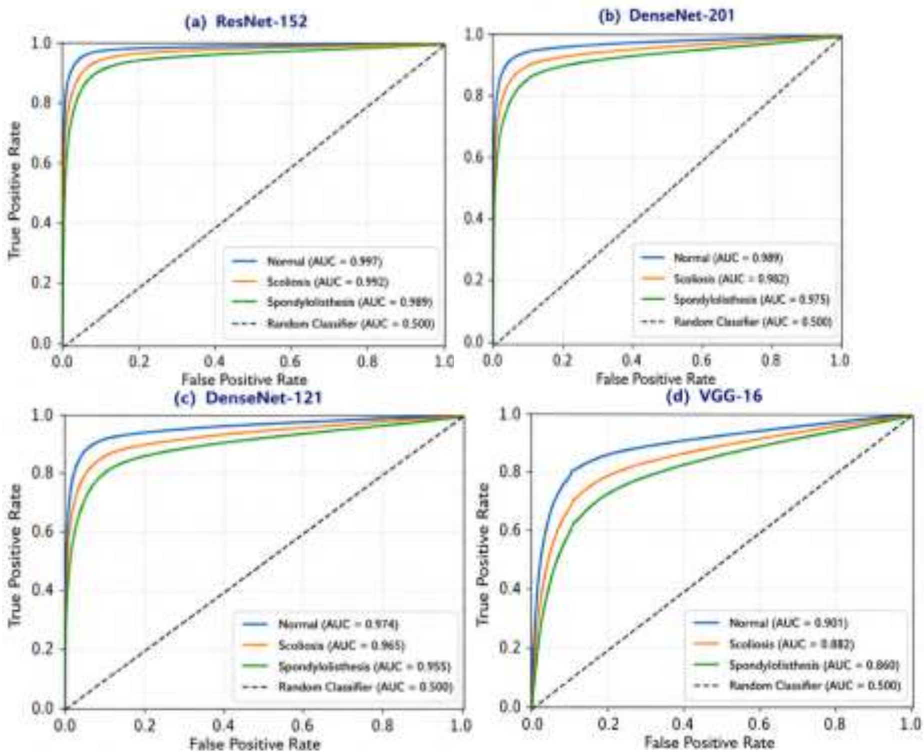

Figure 6. ROC Curve for Three-Class Classification

ROC graph of the ResNet-152 model for 3-class spinal disorder classification is illustrated in Figure 6. The graph is close to the upper-left corner of the graph, indicating high sensitivity and specificity. Table 3 present the AUC values demonstrate the strong discriminative capability of the ResNet-152 model for all spinal disorder categories. Table 3 presents the AUC values obtained for each class in the three-class classification task. The high AUC values demonstrate the strong discriminative capability of the ResNet-152 model for all spinal disorder categories.

Table 3 AUC values for the three-class classification

Class	AUC
Normal	0.998
Scoliosis	0.996
Spondylolisthesis	0.995
Micro-average	0.997

4.2 Four-Class Classification Results

The training and validation accuracy graph and loss curves for four-class classification shown in figure 7. The model exhibits stable convergence with increasing accuracy and decreasing loss throughout the training process, demonstrating effective learning and strong generalization performance. the confusion matrix of the ResNet-152 model used to classify spinal disorders into four classes is shown in Figure 8. The matrix indicates high classification accuracy, with very few misclassified samples, displaying the efficacy of the suggested framework for classifying multiple spinal disorders.

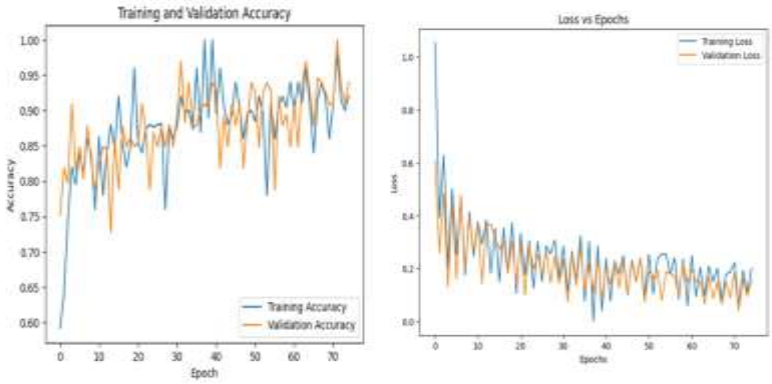

Figure 7. Training and validation Accuracy and loss curves

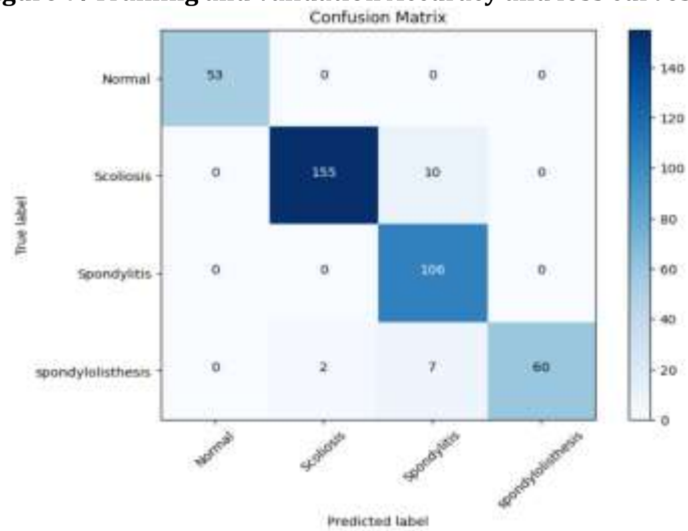

Figure 8. Confusion Matrix

Four class classification confusion matrix has shown in Table 4. ResNet-152 led in mean accuracy and performed well across the other evaluation metrics. Table 5 shows the F1 score, precision, and recall for all CNN models in four-class classification. ResNet-152 achieved consistent performance across all remaining models, demonstrating superior classification performance for several spinal disorders.

Table 4. Results of 4 classes classification

CNN Model	Average Accuracy (%)	Highest Accuracy (%)	Lowest Accuracy (%)	Standard Deviation
DenseNet-201	94.40	100	73.71	22.98
VGG-16	75.31	99	42.42	42.16
DenseNet-121	89.24	100	63.40	35.46
ResNet-152	95.16	100	72.72	21.44

Table 5. Results of 4 classes classification

CNN Model	F1 Score (%)	Precision (%)	Recall (%)
DenseNet-201	94.25	95.25	94.25
VGG-16	84.5	79	91
DenseNet-121	85.75	90	84
ResNet-152	95.5	96	95.25

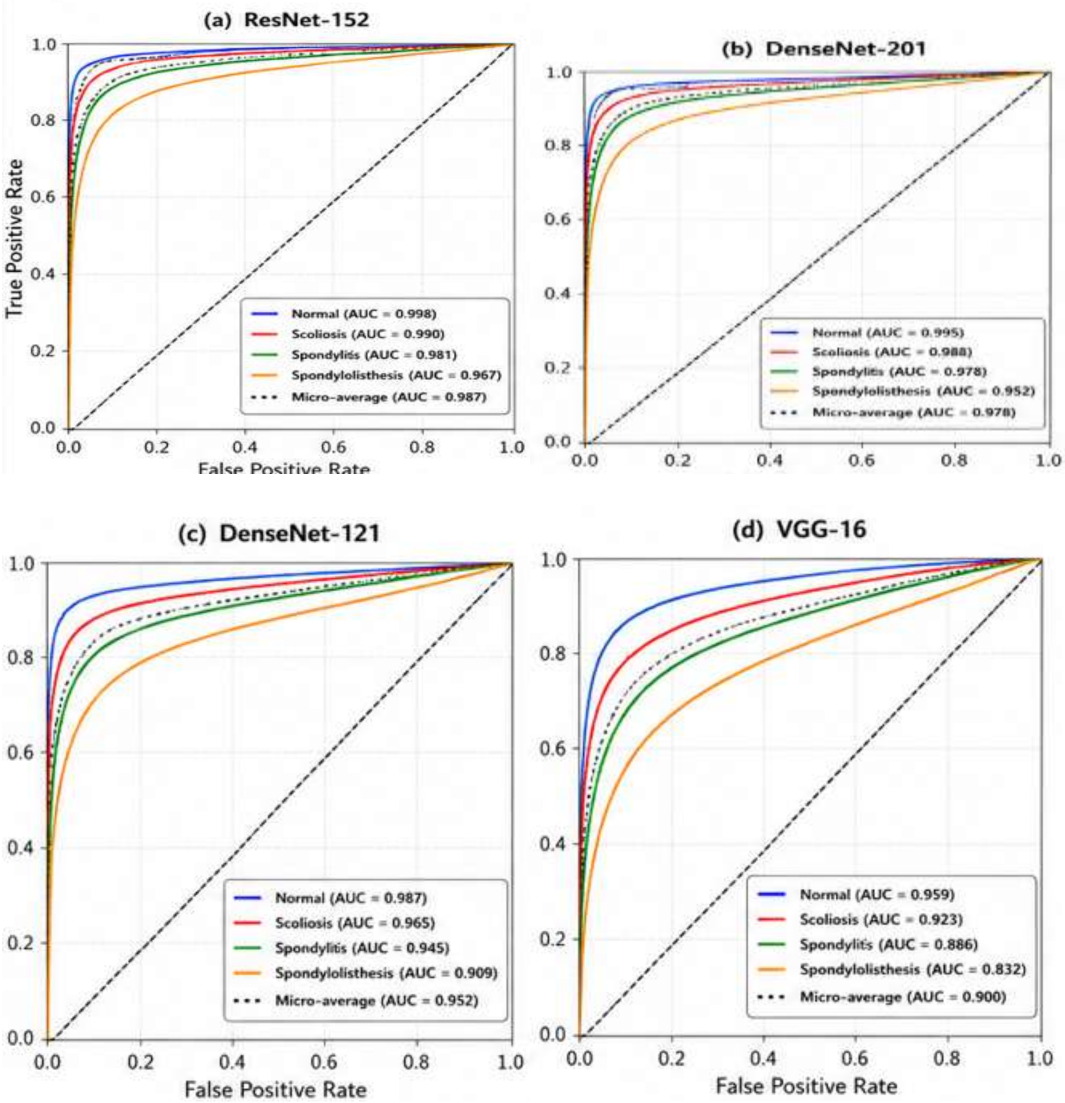

Figure 9. ROC Curves for 4-class Classification

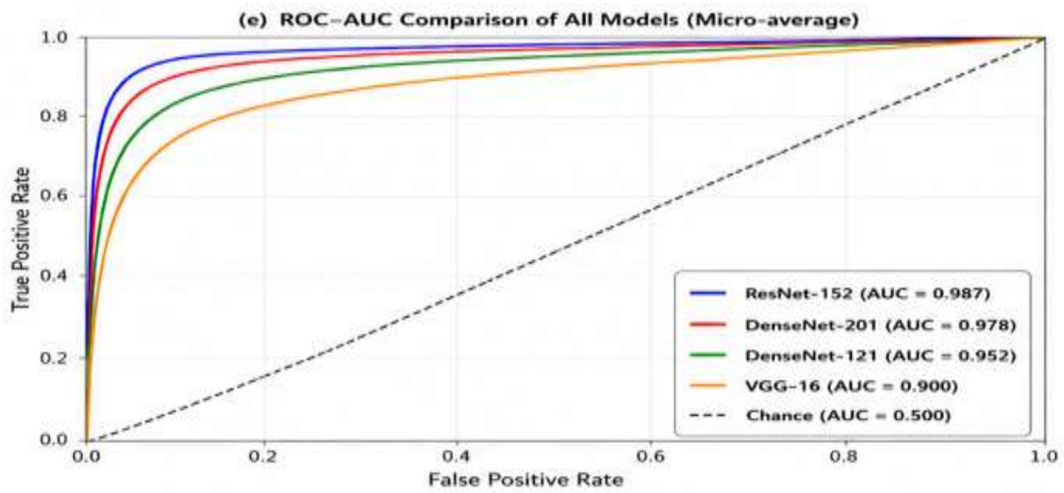

Figure 10. Comparison of ROC Curves for all CNN models

Figure 9 presents the ROC curves for four-class spinal disorder classification using the ResNet-152 model. High AUC values across all classes indicate excellent classification capability and reliable discrimination between different spinal disorders. Figure 10 compares the ROC curves of ResNet-152, DenseNet-201, Densenet-121 and VGG-16. ResNet-152 achieved the highest ROC-AUC values, followed by Densenet-201, DenseNet-121 and VGG-16 demonstrating its superior classification performance. Table 6 reports the AUC values for the Four-class spinal disorder classification task. The high AUC values across all classes confirm the excellent discriminative performance of the proposed deep learning framework.

Table 6. AUC values

Class	AUC
Normal	0.997
Scoliosis	0.994
Spondylitis	0.992
Spondylolisthesis	0.989
Micro-average	**0.993**

4.3. Comparative Analysis

A comparative study evaluates different models in classifying spinal disorders. Table 7 presents an overview of the version of the suggested technique against existing methods:

Table 7. A Comparison of CNN Models for Spinal Disorder Classification

Author	Target Disorders Classes	Dataset Type	Model Used	Accuracy
Bezabh et al. (2024)	Cervical Spondylosis Myelopathy	X-ray Datasets	AlexNet, LeNet-based CNN	97.82%
Fraiwan et al. (2022)	Scoliosis, Spondylolisthesis	X-ray Datasets	ResNet-50, VGG-16, Densenet-201	96.73%
Koo et al. (2022)	Ankylosing spondylitis	Spinal Radiographs	ResNet-152, HRNet	91.6%
Proposed work	Normal, Scoliosis, Spondylitis, Spondylolisthesis	X-ray datasets (Kaggle + Hospital Datasets)	ResNet-152, DenseNet-201, DenseNet-121, VGG-16	98.95%(3-classes), 95.16% (4-classes

DISCUSSION ON COMPARATIVE STUDY

The proximate research emphasizes the crucial role of various deep learning architectures in the identification of spinal diseases. Despite the great classification accuracy reported in earlier research, this comparison study's model performance was enhanced by the use of standardized data.

- Better Performance of ResNet-152: ResNet-152, in comparison to other models such as ResNet-50 and VGG-16, was able to perform better in the feature extraction process, especially in complex classification problems such as four-class classification of various spinal disorders.
- Impact of the use of Benchmark Dataset: Utilizing standardized data in this comparative study, as opposed to relying on publicly available data in previous research, enhanced model performance.
- Practical Application of the models: The suggested model's precision can be applied to a variety of therapeutic applications, such as highly accurate identification of spinal diseases.

- Challenges in Selecting the models: ResNet-152, which showed better accuracy in comparison to other models, was computationally complex in comparison to other models used in the comparative study.

The proposed study highlights the significance of both model and dataset selection for accurate AI-based detection of spinal disorders. The selection of models should be a key principle in future research on AI for detecting various spinal disorders.

The transfer learning in deep learning can be used in an effective manner in the screening of Spondylolisthesis and Scoliosis, which can help in identifying such patient who may require expensive imaging tests such as CT scans and MRI. Besides this, it can also be made even better and more reliable by using bigger datasets with more images and more types of diseases. Also, in practical scenarios, it can be of benefit in real-world scenarios.

CONCLUSION

Using Combined benchmark and hospital spinal X-ray datasets, this analysis introduced a deep transfer learning framework for the automated classification. Four convolutional neural network models, namely ResNet-152, DenseNet-201, Densenet-121 and VGG-16 were evaluated for both 3-class and 4-class classification tasks. Experimental results demonstrated that ResNet-152 consistently achieved the best performance, obtaining classification accuracies of 98.95% for three-class classification and 95.16% for four-class classification, together with superior precision, recall, F1-score and ROC-AUC values. While reducing overfitting, the application of data augmentation and transfer learning increases the model's resilience and generalization. By contrasting the proposed framework with earlier studies, its efficacy was further demonstrated. These results imply that the presented method can provide reliable computer-aided assistance for the early identification and categorization of spinal diseases in clinical settings. The future research scopes are as under

- Expand benchmark datasets to include spinal tumors, fractures, and degenerative diseases.
- Develop automated dataset standardization for large-scale classification.
- Implement explainable AI techniques to enhance model interpretability for clinical applications.
- Enhance real-time deployment for hospital systems.

By focusing on these areas, future research can enhance diagnostic accuracy and further develop AI-assisted medical imaging for detecting spinal disorders.

Conflicts of Interest

No any conflict of Interests

Funding Statement

No Funding

Acknowledgments

The authors acknowledge the support received during this research. This work was not funded by any external agency.

Ethics Statement

All procedures utilized in this analysis were taken out in agreement with appropriate policies and rules, approved by the HEAD Doctor of the USHA Hospital. The experimental protocols were reviewed and approved by the USHA hospital HEAD Doctor, assuring observation with ethical standards for research involving human participants. As the dataset consists solely of anonymized X-ray images without any personal identifiers, individual informed consent was not required in accordance with USHA hospital guidelines.

References

[1] Bhawana, Kumari, Amritanjali, and I. Mukherjee. "Deep Learning for Detecting Common Spine Disorder." In International Conference On Artificial Intelligence, Computing, IOT and Data Analytics, pp. 313-324. Singapore: Springer Nature Singapore, 2024.

[2] Haldeman, Scott, Deborah Kopansky-Giles, Eric L. Hurwitz, Damian Hoy, W. Mark Erwin, Simon Dagenais, Greg Kawchuk, Björn Strömqvist, and Nicolas Walsh. "Advancements in the management of spine disorders." Best practice & research Clinical rheumatology 26, no. 2 ,263-280, 2012.

[3] Morales H. Infectious spondylitis mimics: mechanisms of disease and imaging findings. InSeminars in ultrasound, CT and MRI Dec 1 (Vol. 39, No. 6, pp. 587-604). WB Saunders,2018.

[4] Dai Q, Ishfaque M, Khan SU, Luo YL, Lei Y, Zhang B, Zhou W. Image classification for sub-surface crack identification in concrete dam based on borehole CCTV images using deep dense hybrid model. Stochastic Environmental Research and Risk Assessment.39(10):4637-54,2025.

[5] Shahzad I, Khan SU, Waseem A, Abideen ZU, Liu J. Enhancing ASD classification through hybrid

attention-based learning of facial features. Signal, Image and Video Processing,18(Suppl 1):475-88, 2024.

[6] Khan SU, Raza A, Shahzad I, Ali G. Enhancing concrete and pavement crack prediction through hierarchical feature integration with VGG16 and triple classifier ensemble. In2024 Horizons of Information Technology and Engineering (HITE), (pp. 1-6). IEEE, 2024.

[7] Khan US, Khan SU. Boost diagnostic performance in retinal disease classification utilizing deep ensemble classifiers based on OCT. Multimedia Tools and Applications,84(19):21227-47,2015.

[8] Noonan AM, Brown SH. Paraspinal muscle pathophysiology associated with low back pain and spine degenerative disorders. JOR spine. ,24(3): e1171,2021.

[9] Frost BA, Camarero-Espinosa S, Foster EJ. Materials for the spine: anatomy, problems, and solutions. Materials,12(2):253,2019.

[10] Mohd Isa IL, Teoh SL, Mohd Nor NH, Mokhtar SA. Discogenic low back pain: anatomy, pathophysiology and treatments of intervertebral disc degeneration. International journal of molecular sciences, 24(1):208,2022.

[11] Santiago FR, Ramos-Bossini AJ, Wáng YX, Zúñiga DL. The role of radiography in the study of spinal disorders. Quantitative imaging in medicine and surgery,10(12):2322,2020.

[12] Lee BH, Moon SH, Suk KS, Kim HS, Yang JH, Lee HM. Lumbar spinal stenosis: pathophysiology and treatment principle: a narrative review. Asian Spine Journal,14(5):682,2020.

[13] Masood RF, Taj IA, Khan MB, Qureshi MA, Hassan T. Deep learning based vertebral body segmentation with extraction of spinal measurements and disorder disease classification. Biomedical Signal Processing and Control, 71:103230,2022.

[14] Ma S, Huang Y, Che X, Gu R. Faster RCNN-based detection of cervical spinal cord injury and disc degeneration. Journal of applied clinical medical physics, 21(9):235-43,2020.

[15] Al-kubaisi A, Khamiss NN. A transfer learning approach for lumbar spine disc state classification. Electronics,11(1):85,2021.

[16] Abuhayi BM, Bezabh YA, Ayalew AM. Lumbar disease classification using an involutional neural based VGG Nets (INVGG). IEEE Access,12:27518-29,2024.

[17] Bezabh YA, Salau AO, Abuhayi BM, Ayalew AM. Classification of cervical spine disease using convolutional neural network. Multimedia Tools and Applications, 83(41):88963-79,2024.

[18] Voter AF, Larson ME, Garrett JW, Yu JP. Diagnostic accuracy and failure mode analysis of a deep learning algorithm for the detection of cervical spine fractures. American Journal of Neuroradiology, 42(8):1550-6,2021.

[19] Yeh LR, Zhang Y, Chen JH, Liu YL, Wang AC, Yang JY, Yeh WC, Cheng CS, Chen LK, Su MY. A deep learning-based method for the diagnosis of vertebral fractures on spine MRI: retrospective training and validation of ResNet. European Spine Journal, 31(8),2022.

[20] Gautam R, Sharma M. Prevalence and diagnosis of neurological disorders using different deep learning techniques: a meta-analysis. Journal of medical systems, 44(2):49,2020.

[21] Raihan-Al-Masud M, Mondal MR. Data-driven diagnosis of spinal abnormalities using feature selection and machine learning algorithms. Plos one, 15(2): e0228422,2020.

[22] Altun S, Alkan A, Altun İ. LSS-VGG16: diagnosis of lumbar spinal stenosis with deep learning. Clinical Spine Surgery, 36(5): E180-90,2023.

[23] Ma S, Huang Y, Che X, Gu R. Faster RCNN-based detection of cervical spinal cord injury and disc degeneration. Journal of applied clinical medical physics,21(9):235-43,2020.

[24] Fraiwan M, Audat Z, Fraiwan L, Manasreh T. Using deep transfer learning to detect scoliosis and spondylolisthesis from x-ray images. Plos one,17(5): e0267851,2022.

[25] Saravagi D, Agrawal S, Saravagi M, Chatterjee JM, Agarwal M. Diagnosis of lumbar spondylolisthesis using optimized pretrained CNN models. Computational Intelligence and Neuroscience, (1):7459260,2022.

[26] Koo BS, Eun S, Shin K, Hong S, Kim YG, Lee CK, Yoo B, Oh JS. Differences in trajectory of disease activity according to biologic and targeted synthetic disease-modifying anti-rheumatic drug treatment in patients with rheumatoid arthritis. Arthritis Research & Therapy. ,24(1):233,2022.