Received: 25 July 2025 | Revised: 28 September 2025 | Accepted: 15 November 2025 | Published online: 31 December 2025

RESEARCH ARTICLE

Journal of Computational and Cognitive Engineering
2025, Vol. 00(00) 1-13
DOI: 10.47852/bonviewJCCE52026933

BON VIEW PUBLISHING

Lightweight CNN-Enabled Framework for Automated Detection of Scoliosis and Spondylosis Using Spine X-rays

Gobalakrishnan Natesan[1,*], Anbarasan Murugesan[1], and Ramshankar Nagarajan[1]

[1] *Department of Computer Science and Engineering, Saveetha Engineering College, India*

Abstract: The spine serves as a fundamental structure that supports bodily functions, movements, protection of vital neural pathways, and overall well-being. Any abnormalities or injuries to the spine can significantly impact a person's quality of life. Early detection of these conditions is crucial for effective treatment and management. Hence, this study focuses on the detection and classification of spine abnormalities using advanced deep learning algorithms. In recent years, deep learning methods, particularly convolutional neural networks (CNNs), have demonstrated promising performance in medical image processing applications. This study investigates the effectiveness of advanced CNN architectures, specifically AlexNet and ResNet, in detecting spine abnormalities and distinguishing scoliosis, spondylosis, and normal spine. Leveraging a dataset comprising 3341 diverse spine X-rays images, this study aims to not only compare the effectiveness of AlexNet and ResNet but also determine the most accurate model for deployment in clinical settings. The results demonstrate the comparable performance of AlexNet and ResNet in detecting spine abnormalities. Insights gained from this comparative analysis can inform healthcare practitioners and researchers on the optimal choice of deep learning architecture for spine abnormality detection, ultimately contributing to improved diagnostic accuracy and patient outcomes.

Keywords: spine abnormalities, convolutional neural networks, AlexNet, ResNet

1. Introduction

Spine abnormalities are conditions widely ranging from congenital malformations to degenerative alterations that can impact the structure, alignment, and function of the spinal area. Examples include spinal stenosis, disc herniation, kyphosis, scoliosis, and vertebral fractures. These frequently present as pain, limited movement, and neurological deficits, which substantially influence an individual's quality of life [1–3]. Conventional diagnostic techniques that depend on subjective interpretation and manual evaluation are time-consuming and prone to errors.

This study presents a novel approach that improves and automates the detection of spine abnormalities using deep learning techniques particularly focusing on spondylosis (involves small fractures between vertebrae), scoliosis (defined by sideways bending), and normal spine shape. Advanced convolutional neural network (CNN) architectures like AlexNet and ResNet are used to accurately identify and categorize abnormalities in medical X-rays that have been annotated by skilled radiologists [4–6].

The goal of this research is to equip medical practitioners with improved diagnostic capabilities in the field of spine health by creating a dependable and user-friendly tool [7, 8], thereby enabling early intervention and improved patient outcomes through prompt treatments and individualized treatment plans [9, 10]. The major motivation of this work is to resolve the evolving challenges in the detection of spine abnormality, in which the existing diagnostic methods heavily rely on manual interpretation that is highly time-consuming and prone to variability among radiologists. The major motivation of the proposed work is provided as below:

1) **Multiclass clinical-based detection:** Different from the conventional models that effectively concentrate on single abnormality, the proposed method was designed to effectively and accurately classify normal spine, spondylosis, and scoliosis, to achieve support to reach a more comprehensive decision.
2) **Comparative evaluation of CNN architectures:** By comparing the AlexNet and ResNet on similar datasets, the more effective model (i.e., AlexNet with 97% accuracy) would be determined, and researchers and clinicians would gain actionable insights regarding selection of architectures for spine abnormality detection.
3) **Optimized deployment from lightweight architecture:** Although ResNet struggled with overfitting and instability in validation, the AlexNet achieved higher accuracy with lesser computational requirements due to its shallower yet efficient regularization.

The remainder of this research work is organized as follows: In Section 2, we discuss various spine abnormality identification methods carried out using machine learning (ML) and deep learning techniques in image processing applications. In Section 3, we describe the overall system design of the spine abnormality in image processing and deep learning techniques. In Section 4, we present the performance evaluation and analysis of the results obtained using ALexNet had been compared with other state-of-the-art CNN model. In Section 5, our conclusion summarizes the contributions of the study along with suggestions for future research.

2. Literature Review

Wang et al. [11] presented SpineHRNet+, a novel hybrid model that combines rule-based and artificial intelligence methods. The model uses deep learning models HRNet and UNet to automate the detection

***Corresponding author:** Gobalakrishnan Natesan, Department of Computer Science and Engineering, Saveetha Engineering College, India. Email: gobalakrishnann@saveetha.ac.in

of endplate landmarks and the segmentation of the spine region by employing biplanar radiographs, an EOS machine, and manually annotated landmarks as references. The model's dependability is verified by assessment using Bland–Altman plots, confusion matrices, and linear regression. SpineHRNet+ enables continuous auto-analysis of spine alignment, which may improve clinical practice and make large-scale studies easier. Because of its interpretability, users can make adjustments to improve clinical insights. Biplanar radiographs of 1,542 patients with scoliosis are included in the study of Wang et al. [11]. This technique for spine alignment analysis [11] has clinical applicability and scalability, which bodes well for improvements in patient care and future research.

Germann et al. [12] described a deep learning-based automated method for segmenting the lumbar vertebrae. It uses a three-dimensional X-Unet for segmentation and a UNet-based lumbar spine localization network for placement. Comparison with conventional approaches shows improved performance on lumbar spine CT images from the VerSe 2020 public dataset and an internal hospital dataset (containing 500 and 156 samples, respectively). The two-stage approach shows potential for use in the diagnosis of spinal anomalies and support of surgical treatment by efficiently positioning and segmenting lumbar vertebrae. Notably, it achieves accurate lumbar vertebrae segmentation better than traditional UNet, IFCN, and nn-Unet approaches.

In a paper by Natalia et al. [13], a major problem among tumor-related bone metastases is the vital need for precise detection of spinal metastases, which is addressed in the present study where we present a novel method for predicting and diagnosing spinal metastases using MRI data and deep learning, more especially a multilayer CNN. The model obtains an amazing accuracy rate of 96.45% by analyzing a large dataset consisting of 941 patients who had spinal metastases. It uses manifold learning techniques and SoftMax classifiers to identify spinal metastases with high precision and to reveal fracture risk indicators on X-rays. The present study underscores the capacity of sophisticated computational techniques to revolutionize the identification and handling of spinal metastases, presenting auspicious prospects for augmenting patient results and therapeutic approaches.

Merali et al. [14] employed a deep convolutional neural network (DCNN) to conduct vertebral body measurements and detect insufficiency fractures on lumbar spine MRI scans. The dataset included lumbar spine MRI data from multiple institutions as well as a retrospective examination of 1000 vertebral bodies in 200 patients. Of these patients, 61 suffered fractures resulting in vertebral body insufficiency. The results showed that the DCNN performed exceptionally well in terms of both vertebral body measurements and the diagnosis of insufficiency fractures in a variety of lumbar spine MRI scans. Especially, it shows resilience to changes in radiological facility, a kind of MRI scanner, and intensity of magnetic field. For vertebral body measurements, the DCNN achieved a sensitivity of 0.941, specificity of 0.969, and accuracy of 0.962 in diagnosing insufficiency. It also demonstrated good inter-reader agreement with radiologists.

Biercher et al. [15] focused on automating the selection of transverse lumbar spine MRI scans for detecting anomalies in the intervertebral disc (IVD), notably in the diagnosis of lumbar spinal stenosis (LSS). They investigated a variety of ML techniques, including DCNN, dimensionality reduction, feature selection, and various ML algorithms, utilizing a dataset of MRI scans, and concluded that using the support vector machine technique with a small Gaussian kernel on full-length features taken from the DenseNet201 model produces the best results. It achieved per-class classification performance of approximately 0.88, with precision (0.95–0.99) and recall (0.93–1.0). This study by Biercher et al. [15] significantly contributes to the literature with their thorough investigation of ML techniques for automatic image selection, providing valuable insights into LSS diagnosis and suggestions regarding selection of suitable methodologies.

Merali et al. [16] trained a CNN to detect various spinal cord diseases in thoracolumbar MRI scans of dogs and evaluated its performance. The CNN was trained to recognize intervertebral disc extrusion (IVDE), intervertebral disc protrusion (IVDP), fibro-cartilaginous embolism (FCE), acute non-compressive nucleus pulposus extrusion (ANNPE), syringomyelia, and neoplasia using MRI scans from 500 dogs (2,693 for training and 7,695 for testing). The results demonstrated the CNN's ability to recognize and identify various diseases, with particular success in detecting IVDPs and IVDEs. Notably, on sagittal T1-weighted images, the CNN detected IVDPs with a sensitivity of 100% and a specificity of 95.1%. It also performed well in detecting IVDEs, FCEs, and ANNPEs, with different sensitivity and specificity [17, 18]. This study by Merali et al. [16] demonstrated CNNs' potential for accurately and specifically detecting complex spinal cord disorders in canine MR imaging.

Current work has explored highly advanced architectures beyond conventional CNN. For instance, the UIU-Net, a variant of U-Net that enables uncertainty-based mechanism and enhanced encoder-decoder pathways, has achieved improved performance compared with traditional CNN architectures in medical anomaly detection tasks especially in segmentation enabled spine analysis [19, 20]. Such models are computationally intensive and require larger annotation datasets. In contrast, the proposed model enables lighter CNN architecture which trade-off computational efficiency and accuracy for real time diagnostic support.

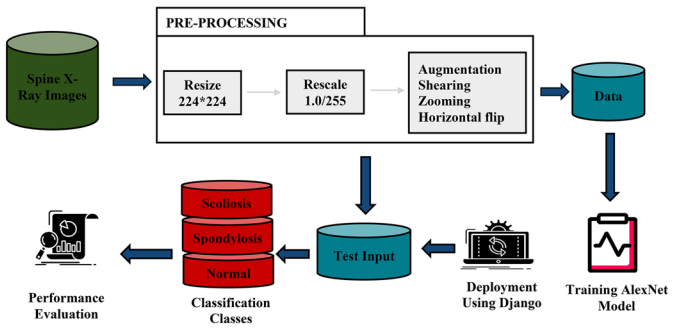
3. System Design

The system to detect and classify spine abnormalities using advanced CNN techniques uses a comprehensive approach encompassing data preprocessing, model architecture design, training, and evaluation. Initially, a diverse dataset of X-ray images of the spine comprise both normal and abnormal spine conditions, and these are collected and preprocessed. Preprocessing techniques such as rescaling, shear transformations, zoom transformations, flipping images horizontally are performed to enhance model generalization. The dataset is then divided into subsets for testing, validation, and training. The model architecture is developed using advanced CNN techniques such convolutional layers with different filter sizes, batch normalization, dropout layers, advanced activation functions, and pooling layers for spatial down sampling (Figure 1). To improve performance, deeper learning models are also being investigated. The model is optimized during training by using suitable optimizers [21, 22]. To evaluate the effectiveness and accuracy of the categorization, a variety of performance metrics are computed and applied.

The model's effectiveness and flexibility to changing patient demographics and diagnostic standards are guaranteed by regular retraining and ongoing monitoring. An efficient system for spine abnormality detection and classification employing advanced CNN techniques is obtained through this methodical approach, which may have a big impact in the medical field.

Figure 1

Proposed system architecture of spine abnormality



3.1. Data setting

In this study, we leveraged the dataset sourced from Fraiwan et al. [23], which comprised 3341 vertebrae X-ray images depicting a diverse representation of spine conditions including spondylosis, scoliosis, and normal spine, facilitating a comprehensive model training. As shown in Figure 2, the system can learn to classify X-ray images accurately, distinguishing between different spine conditions and detecting them with high precision. This combination of hardware and software ensures robust, efficient, and consistent development environments suitable for the complex task of medical image analysis and spine abnormality detection.

3.2. Preprocessing images

The model starts by importing images from the specified dataset directory using the OpenCV library. The images are read from their storage location and formatted to function with the training and inference deep learning framework. X-ray images usually have different resolutions and sizes. Therefore, the images are downsized to a specified resolution of 224 × 224 pixels, which is a typical input size for many deep learning models including AlexNet and ResNet, to ensure consistency and compliance with the model's input size. X-ray images are usually saved in the project as digital pictures with pixel values between 0 and 255. For images with 8-bit depth, each pixel is represented by an integer value between 0 (black) and 255 (white), indicating the intensity or brightness of the pixel. In this model, rescaling is applied using the formula: rescaled pixel value = original pixel value/255, where 1.0/255 is used to perform element-wise division, effectively scaling the pixel values to the range [0, 1]. Rescaling is part of the pipeline along with other data preparation steps like resizing and data augmentation. Every pixel in the input image is rescaled before being sent to the deep learning model for training or inference. The X-ray images are subjected to data augmentation in order to produce more training samples and enhance the model's capacity to generalize to previously unknown data.

The following data augmentation techniques are employed in this project:

a. Shearing: Shear transformations introduce geometric distortions to the image by displacing a portion of the image in either the horizontal or vertical direction. The shear range parameter establishes the maximum intensity of the shear transformation that is applied to the images. Here, the shearing range was set to 0.2, allowing for a maximum shear angle of 20°.
b. Zooming: Randomly zooming into or out of the image allows simulation of variations in scale. The zoom range was set to 0.2, allowing for a maximum zoom factor of 20%.
c. Horizontal flipping: This step randomly flips images horizontally. It efficiently doubles the training data set without requiring additional picture collections by introducing left-right symmetry variances. Horizontal flipping was applied with a probability of 50%, resulting in a mirror image of the original with equal likelihood.

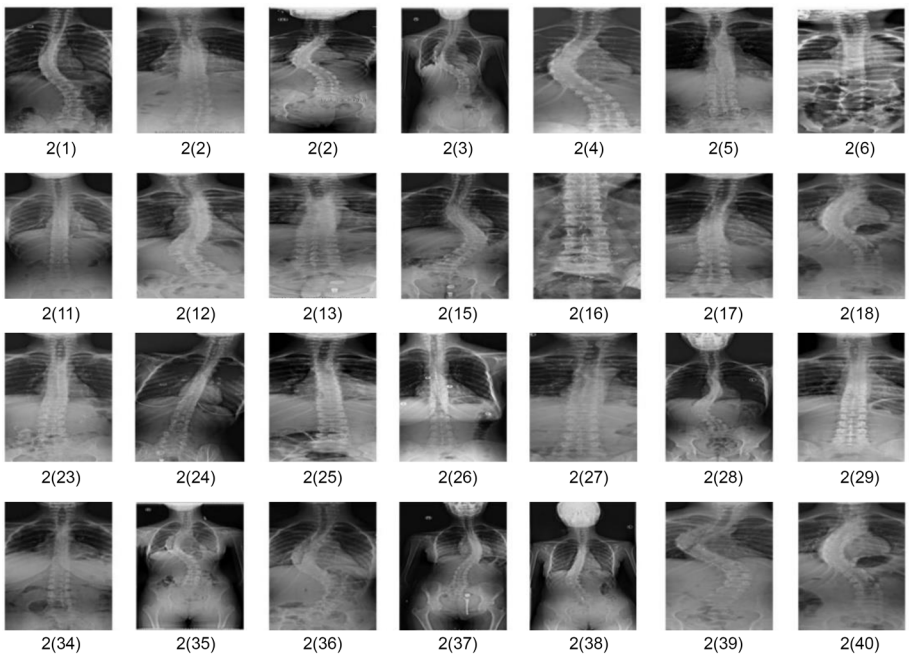
3.3. Model building and training

Before proceeding with the detailed exploration of model architectures, it is imperative to acknowledge the initial phase of the project where both AlexNet and ResNet architectures were meticulously trained and evaluated for spine abnormality detection. Subsequently, the model exhibiting the highest accuracy was judiciously selected for deployment in the spine abnormality detection system. This methodical approach ensured a thorough comparison between AlexNet and ResNet architectures, ultimately guaranteeing the deployment of the most accurate and reliable model for real-world implementation.

3.3.1. AlexNet model

AlexNet's input layer creates a placeholder for input data and defines its intended format. The input shape is (224, 224, 3), suggesting images having height and width dimensions of 224 pixels and three RGB color channels. Convolutional layers are critical components of CNNs as they extract significant information from input images. These

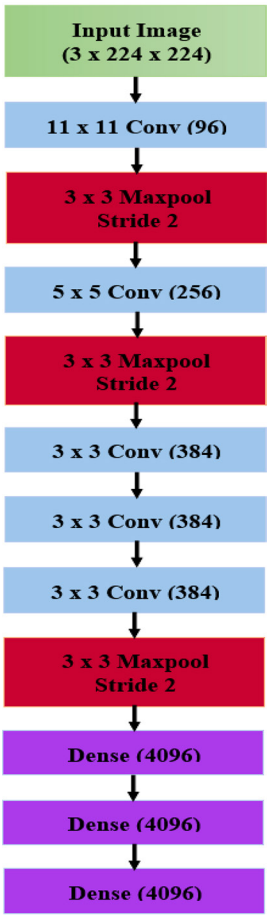
Figure 2
Vertebrae X-ray images



layers conduct convolution operations on the input image using a set of learnable filters or kernels to identify characteristics like edges, textures, and patterns. Here, the AlexNet model consisted of 5 convolutional layers and other layers:

a. First convolutional layer: The input images were subjected to 96 filters, each measuring 11 × 11 pixels, using the first convolutional layer of AlexNet. These filters captured large-scale patterns and structures, like edges and color gradients, by swiping across the input with a stride of 4. The output feature maps maintained the spatial dimensions of the input. Rectified linear unit (ReLU) activation was used to add nonlinearity after the convolution procedure, enabling the network to describe intricate relationships in the data. The activations were subsequently normalized by batch normalization, which lowered the internal covariate shift and improved training stability and efficiency.
b. Second convolutional layer: enhancing the first feature extraction, 256 filters of 5 × 5 pixels were used in the second convolutional layer. The previous layer's feature maps can be used to extract more precise and localized features, thanks to these smaller filters. Nonlinearity was again introduced using ReLU activation, which made it easier to identify complex patterns in the data.
c. Third, fourth, and fifth convolutional layers: These next convolutional layers were made up of 384 filters, each measuring 3 × 3 pixels. The network may capture increasingly abstract and high-level representations of the input data by stacking many layers with lower filter sizes. In order to encode more intricate and discriminative information pertinent to the classification task, these layers concentrated on honing the features that were extracted by the layers that came before them. While batch normalization was applied after these layers, ReLU activation was applied to create nonlinearity.
d. Normalization layers (batch normalization): By normalizing the output of the convolutional layers, batch normalization improves stability and speeds up training. Normalizing each batch's activations aids in lowering the internal covariate shift. Each convolutional layer was followed by batch normalization.
e. Max pooling layers: To downsample the feature maps and improve their resilience to spatial translations, AlexNet's Max Pooling layers were arranged after the first, second, and fifth convolutional layers. The most prominent characteristics were retained, while the spatial dimensions of the feature maps were effectively cut in half with each pooling operation, which has a size of 3 × 3 and a stride of 2. Max Pooling assisted in capturing the most salient aspects while eliminating extraneous information by keeping only the maximal activation within each pooling region.
f. Flattening layer: The flattening layer acts as a link between the convolutional and fully connected layers. It collapses spatial information while maintaining feature representation by converting the multidimensional convolutional layer output into a one-dimensional tensor. Spatial information with dense connections may be seamlessly integrated, thanks to the flattening layer, which converts the convolutional feature maps into a linear array. As a result, it makes it easier for learnt features to spread to next layers, guaranteeing efficient feature extraction and classification in challenging tasks like image recognition.
g. Fully connected (dense) layers: The central hub of the completely connected layers of AlexNet was made up of two dense layers, each with 4096 neurons. These layers added nonlinearity to the network by applying ReLU activation to every neuron. Furthermore, by permitting only positive values to flow through, the ReLU activation function mitigated the vanishing gradient issue and speeded up the network's calculation. The AlexNet architecture culminates in a soft-max output layer, which generates probability distributions over the three classes, enabling the model to make predictions about the presence of spine abnormalities in the input X-ray images. The architecture is compiled using the Adam optimizer, categorical cross-entropy loss function, and metrics such as accuracy and precision to evaluate model performance during training and testing phases. Figure 3 shows the architecture of ALexNet.

Figure 3
AlexNet architecture



3.3.2. ResNet model

In this spine abnormality detection project, the ResNet architecture utilized began with an input layer tailored to receive RGB images of dimensions 224 × 224 pixels.

a. Convolutional layers: The initial convolutional layer utilized a 7 × 7 kernel size with 64 filters and a stride of 2 to process input images. Each filter extracted different features, such as textures and edges, contributing to the model's ability to capture intricate patterns in the input data.
b. Max pooling layers: After convolutional processing, max pooling layers significantly reduced the spatial dimensions of feature maps, facilitating more effective processing and better generalization. In this downsampling process, the maximum value was set within each 3 × 3 region of the feature maps, in order to capture essential information while reducing computational complexity.
c. Residual blocks: The ResNet architecture incorporated residual blocks to address the vanishing gradient problem commonly encountered in deep networks. Each residual block consisted of convolutional layers followed by batch normalization and ReLU activation functions, enhancing stability and convergence during training. Skip connections within residual blocks ensured the preservation of original data from earlier layers, mitigating degradation issues in deep networks.

d. Global average pooling layer: Following residual blocks, global average pooling condensed feature maps into fixed-size vectors, efficiently summarizing distinct features while preserving spatial information. This process enhanced generalization and reduced the risk of overfitting by ensuring the model's output remains consistent across varying input sizes.
e. Fully connected layer: The flattened layer converted 3D feature maps into a 1D vector, preparing them for further processing. Subsequently, a fully connected layer processed the flattened vector, mapping retrieved features to class probabilities.
f. Activation function: The final output of the fully connected layer was transformed using an activation function, typically soft-max, which converted the output values into probabilities and normalized them across all classes, facilitating interpretation by providing a likelihood distribution over potential class.
g. Optimizer and loss function: The ResNet model was trained using the Adam optimizer, a popular optimization algorithm known for its efficiency and effectiveness in minimizing the model's loss function. During training, the difference between expected and actual class probabilities was measured using the categorical cross-entropy loss function.

3.3.3. Training

Both AlexNet and ResNet architectures were trained using a batch size of 32 images per batch, over multiple epochs. The training process iterated over the dataset for a predefined number of epochs to update model parameters. The models were compiled with the Adam optimizer, categorical cross-entropy loss function, and evaluation metrics such as accuracy and precision. Early stopping criteria, based on validation loss or performance plateauing, were implemented to halt training and prevent model over fitting. Hyper-parameter tuning was performed to optimize model performance, including adjustments to learning rate, dropout rate, and other model parameters.

3.4. Model evaluation

Following training, the efficiency of the trained models (ResNet and AlexNet) in identifying spine anomalies was assessed. During preprocessing, the dataset was divided into training and validation sets. While the validation set was kept reserved for evaluation, the training set was utilized to train the models. This guarantees that the models are evaluated on untested data, resulting in a more precise assessment of their generalization capacity. To measure the models' effectiveness, accuracy, precision, recall (sensitivity), and F1-score were calculated.

3.4.1. Accuracy

Accuracy refers to the proportion of correctly classified images out of the total number of images in the validation set. The formula used to evaluate accuracy is represented in Equation (1):

$$Accuracy = \frac{Number\ of\ Correct\ Prediction}{Total\ Number\ of\ Prediction} \quad (1)$$

3.4.2. Precision

Precision refers to the ratio of true positive predictions to the total number of positive predictions. It measures the model's ability to correctly identify positive cases (e.g., spine abnormalities) without misclassifying negative cases (e.g., normal spine) using Equation (2), where TP represents true positives and FP represents false positives.

$$Precision = \frac{TP}{(TP+FP)} \quad (2)$$

3.4.3. Recall (sensitivity)

Recall is calculated using Equation (3), where TP represents true positives and FN represents false negatives. It is the ratio of true positive predictions to the total number of actual positive cases in the validation set. Recall measures the model's ability to correctly detect all positive cases.

$$Recall = \frac{TP}{(TP+FN)} \quad (3)$$

3.4.4. F1-Score

This metric measures the harmonic mean of precision and recall using Equation (4), providing a balanced measure of the model's performance.

$$F1\ Score = 2 \times \left[Precision\left(\frac{Recall}{Precision+Recall}\right)\right] \quad (4)$$

3.4.5. Model selection and deployment

The selection of the optimal model for deployment is a critical decision influenced by various factors, including accuracy, performance metrics, and computational efficiency. After thorough evaluation, it was determined that AlexNet demonstrated superior accuracy and precision in classifying spine abnormalities, compared to ResNet, making it the optimal choice for deployment within the web application framework. Detailed analysis and comparison of the performance metrics between AlexNet and ResNet is presented in Table 1.

AlexNet Algorithm
Input : x = dataset
Output : Classification accuracy, sensitivity, specificity, precision and F1 score
1. X is divided into a training set and a testing set.
2. Preprocessing: resize the training set
3. Augment: Rand Rotation [-5,-5] , Rand X Reflection 1, Rand Y Reflection 1, Rand X Shear [-0.05, 0.05], Rand Y Shear [-0.05,0.05], Rand X Shear [0.5, 1], Rand Y Shear [0.5,1], Rand X Translation [-5,-5], Rand Y Translation [-5,-5]
4. Initialize: net = AlexNet, S = sparsity, N= the number of Iterations, T= threshold, A=array of scores
5. For i= 1 to N
6. Calculate: A (Dlupdate)
7. Sort (A)
8. Calculate: The threshold by Threshold=Iteration Scheme (x) ×A
9. Prune (net,t)
10. Set: number of pruned parameters.
11. End
12. Return Classification accuracy, sensitivity, specificity, precision and F1-score

Table 1
AlexNet and ResNet performance metrics

Evaluation parameters	AlexNet	ResNet
Overall accuracy	0.9687	0.94792
Validation loss	0.2704	3.3679
Precision	0.9344	0.9158
Recall (sensitivity)	0.8566	0.7867
Validation accuracy	0.9531	0.3906

The trained AlexNet model was saved in the .h5 format and integrated into a web application using the Django framework for deployment. The building of a web application allowed users to upload a spinal X-ray image for categorization. When an image was uploaded using the web interface, it was preprocessed (e.g., resized and normalized) to meet the AlexNet model's input specifications. The preprocessed image was then run through the deployed AlexNet model to make an inference. The web application showed the user the matching outcome based on the model's prediction, which assigned a probability distribution over classes Normal, Scoliosis, or Spondylosis [24–26]. This method made it easier to understand medical photos and gives patients and healthcare providers quick information about spinal disorders right from the website, all without the need for specialist knowledge or software. Users can interact with the web application through a user-friendly interface, facilitating the diagnosis of spine abnormalities using deep learning technology.

4. Implementation and Results

4.1. Environmental setup

This step created the environment needed to make it easier to develop and train neural network models where Jupyter notebook was used for training and Django framework was used to deploy the model. Keras, TensorFlow, and other necessary packages and libraries were loaded.

4.2. Data preparation

To create batches of photos and their labels for feeding into the neural network model during training and assessment, the code first set up "ImageDataGenerator" objects for both training and testing datasets. It then preprocessed and augmented the images for training.

4.3. Setting up model checkpoint

To save the best-performing model during training, both snippets set up a model checkpointing mechanism. When a new best model was saved, the "ModelCheckpoint" callback published a message, monitored training accuracy, and only saved the model when the accuracy was better than the previous best. The file path (model_path) was where the saved model was kept.

4.4. Fitting the model

The AlexNet and ResNet models were trained on their respective training datasets (train_data) using the fit technique in both snippets. The training procedure processed batches of data defined by steps_per_epoch at each epoch, iterating over the number of epochs indicated by epochs. The model's performance was tracked throughout training using the validation dataset (test_data), as indicated by validation_data. Additionally, the validation procedure ran in batches specified by validation_steps. Performance analysis and post-training visualization were made possible by the training history (history), which recorded the training and validation metrics (e.g., loss and accuracy) at the end of each epoch.

4.5. Model accuracy analysis

AlexNet vs. ResNet training trends examining these charts provided insights on how effectively the model was learning from the training data and whether changes to the training procedure or model architecture were necessary.

Figure 4 shows the accuracy of AlexNet, with a steady increase and a maximum accuracy of 96.88%. Figure 5 shows the accuracy of ResNet with slight fluctuations and gradual increase obtaining a maximum of 94.79%.

Figure 4
Accuracy of AlexNet model

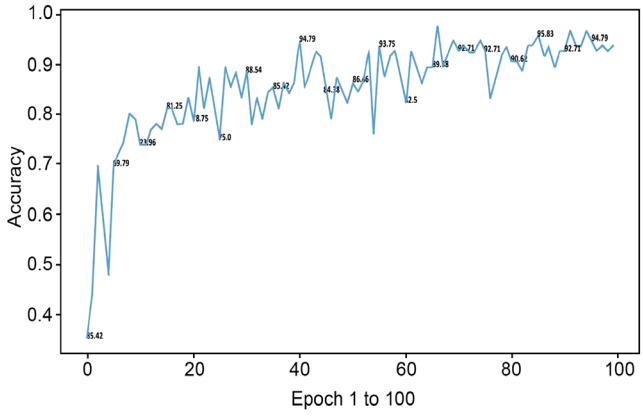
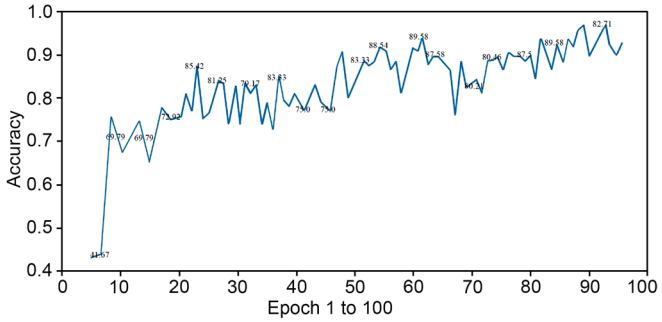


Figure 5
Accuracy of ResNet model



4.6. Visualizing training loss over epochs

Figure 6 illustrates how the loss varied throughout the training process by examining these loss charts. AlexNet showed a stable declining loss, indicating that the model was learning. Figure 7 represents the ResNet model with a plateau or rising loss, which could be an indication of overfitting or other problems that call for changes to the model or training procedure. A declining loss shows that the model was learning.

Figure 6
Loss of AlexNet model

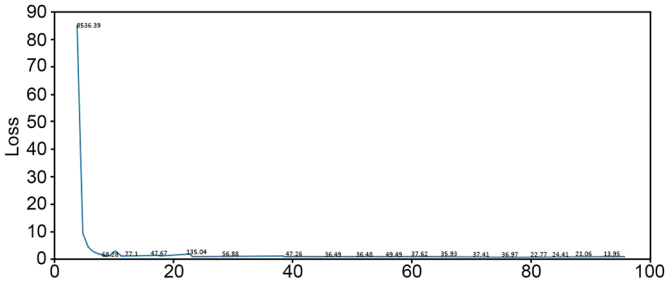
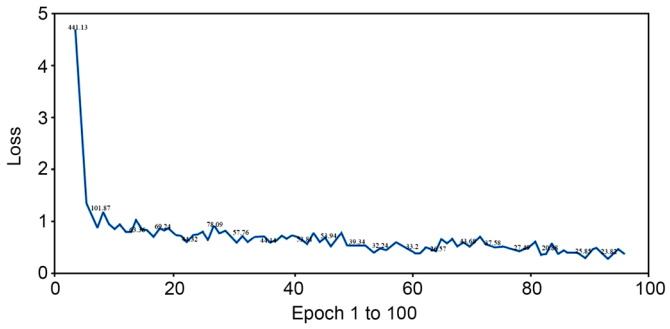


Figure 7
Loss of ResNet model



4.7. Testing and classification of spine abnormality

Deploying AlexNet with Django facilitated model testing using dataset pictures, enabling real-world application and evaluation of its clinical usefulness (Figure 8). This integration provided vital insights into performance using unseen data. Django's flexibility ensured a smooth interaction with the user interface, enhancing accessibility and usability of the spinal anomaly detection system. This holistic approach underscored the significance of selecting appropriate neural network architectures and integrating deep learning models into practical applications to yield meaningful results in healthcare and beyond.

Step 1: Upload an X-ray image from test dataset in the user interface shown in Figure 9.

Step 2: Click submit to classify the spine abnormality of uploaded image. Spine abnormality is classified into three categories: normal spine, scoliosis spine, and spondylosis spine. The sample images are shown in Figure 10 [23]. A normal healthy spine is straight when viewed from the front, with gentle curves when viewed from the side (Figure 10(a)). Scoliosis is defined by an unnatural curving of the spine to the

Figure 8
Implementation of Django framework

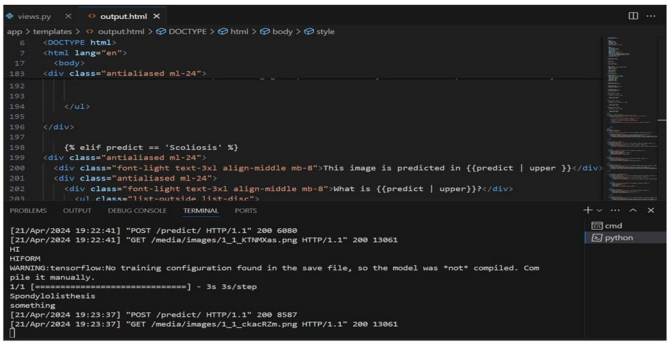


Figure 9
Image uploading interface

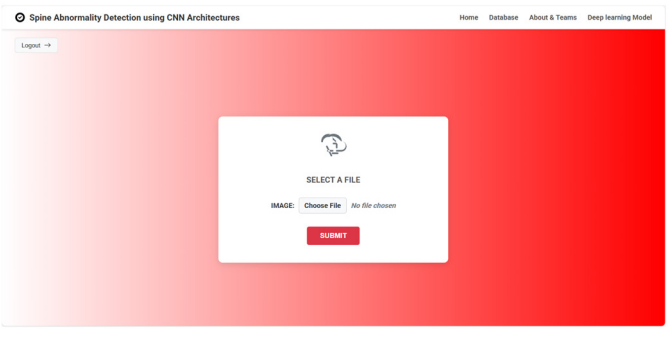
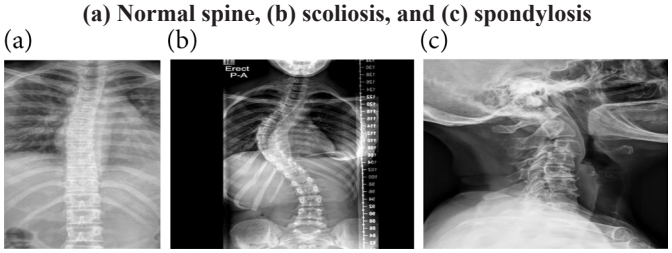


Figure 10
(a) Normal spine, (b) scoliosis, and (c) spondylosis



side. The spine may not line vertically but instead curve sideways to form an S or C shape. Figure 10(b) shows an image of a scoliosis spine. Although this curvature can arise at any age and range from moderate to severe, it is typically diagnosed in adolescence. Spondylosis refers to degenerative changes in the spine, including the formation of bone spurs, disc degeneration, and thickening of ligaments (Figure 10(c)). It can cause pain, stiffness, and decreased mobility in affected areas.

4.8. Performance metrics analysis of AlexNet and ResNet

The performance metrics for AlexNet and ResNet are presented in Table 1, encompassing the various evaluation measures. For instance, the image preprocessing (normalization and rescaling) enabled uniform distribution of input whereas data augmentation (horizontal flipping, zooming, and shearing) predominantly enhanced the generalization and diminished overfitting. Similarly, the inclusion of dropout regularization in the AlexNet showed its superior stability than ResNet with a higher validation loss ad overfitting.

4.9. Training and validation metrics of ResNet and AlexNet

The model's capacity to generalize to new, previously unseen data was measured by validation loss, which compared its predictions to the actual values in the validation set. Similar to training loss, validation loss was frequently calculated using metrics such as mean squared error or cross-entropy loss. Figure 11 demonstrates that both training and validation accuracy have reached roughly 97%. Figure 12 depicts the loss and accuracy metrics throughout epochs, revealing that by the 20th epoch, the model's loss for both training and validation is almost zero. Furthermore, training and validation accuracy rapidly improved, reaching approximately 97% by the 75th epoch.

Figure 11
Performance metrics of AlexNet

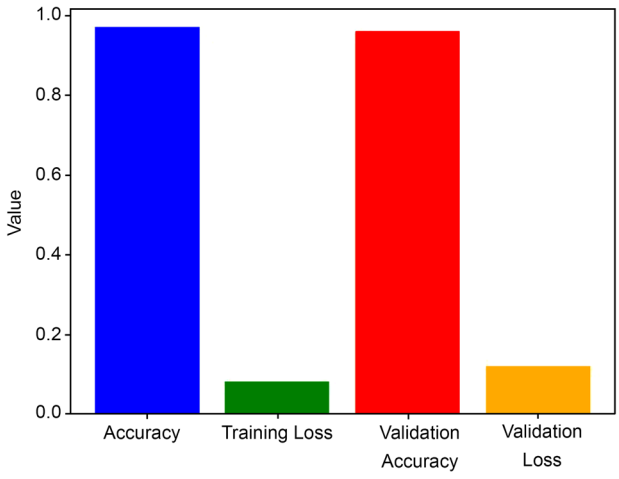


Figure 12
AlexNet loss and accuracy graph

Figure 13 is a graphical depiction of the precision–recall (PR) curve of the trade-off between precision and recall for various threshold levels. When there is an imbalance in the classrooms, it is helpful. Precision provides an answer to the following query: What percentage of the model's positive predictions were accurate? While recall answers, how many of the actual positive events were accurately predicted by the model? Here, the precision and recall curve values of the model are constantly around 95–100%.

Figure 13
Precision–recall curve

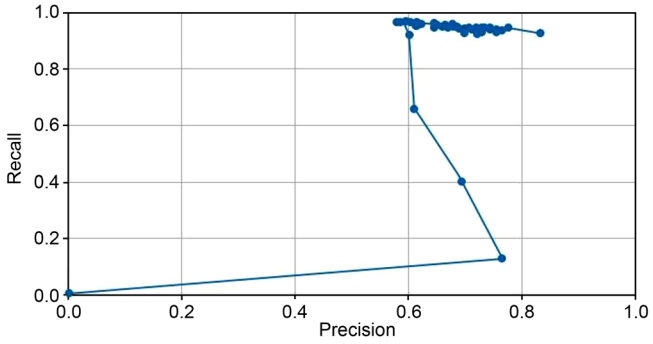
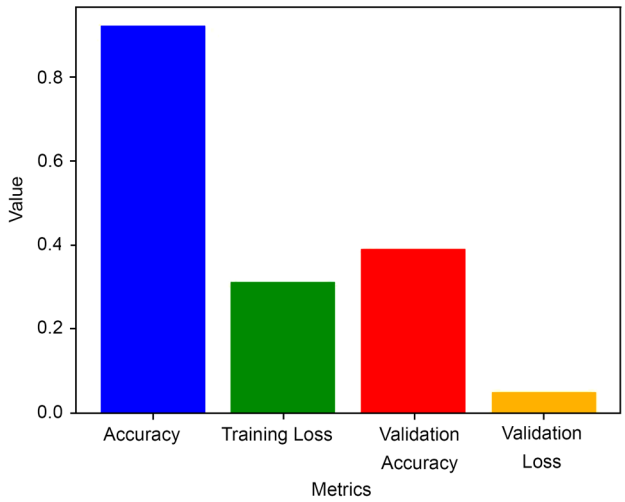


Figure 14
Performance metrics of ResNet



The ResNet model showed a significant difference in performance between the training and validation phases, indicating possible overfitting. In Figure 14, despite reaching a high training accuracy of roughly 94%, the validation accuracy was significantly lower at around 40%. This mismatch was visibly represented in the graph, with the validation loss reaching 3.3679. Such results indicated that, while the model learned efficiently from training data, it struggled to generalize well to unseen data, emphasizing the importance of regularization approaches or model changes to alleviate overfitting and enhance validation accuracy.

The AlexNet and ResNet models were evaluated using a range of performance parameters, including accuracy, validation loss, precision, recall, and validation accuracy, which are all represented in a comparative graph in Figure 15. AlexNet outperformed ResNet with an accuracy of 96.87%, which is close to 97%, and a validation accuracy of 95.31%. It yielded a less significant validation loss of 0.27%, demonstrating robust generalization capabilities. AlexNet also

Figure 15
Comparison of evaluation parameters of AlexNet and ResNet

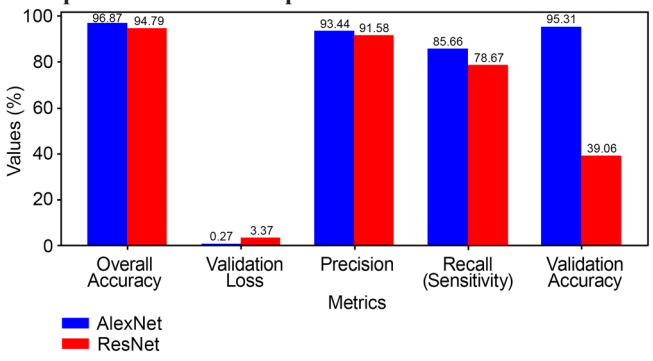


Table 2
Computational complexity and runtime comparison

Metric	AlexNet	ResNet-50
Approx parameters	~60 million	~25.6 million
Model size	~240 MB	~98 MB
FLOPs	~724 MFLOPs	~3.8 GFLOPs
Training time	~3.5 min	~ 7.8 min
Inference time	~ 3.2 ms	~6.5 ms

had remarkable precision and recall scores of 93.44% and 85.66%, respectively, suggesting its efficacy in correctly detecting important instances.

In contrast, ResNet had a slightly lower accuracy of 94.79%, with a significant difference between training and validation accuracy of 94% and 39.06%, respectively. ResNet also had a significantly larger validation loss (3.37%), indicating probable overfitting and low generalization capacity. The precision and recall scores for ResNet were 91.58% and 78.67%, respectively, showing its potential to reliably categorize instances but with lesser recall compared to AlexNet.

4.10. Computational complexity and running time analysis

In order to further examine the proposed work practicality, this research analyze the computational complexity and running time of both the ResNet and AlexNet models. Table 2 shows the comparison in terms of average inference time, FLOPs, model size, and parameter count.

5. Conclusion

To identify spine abnormalities, two common CNN architectures, ResNet and AlexNet, were used. The dataset included 3341 X-ray images divided into three categories: normal spine, scoliosis, and spondylitis. ResNet, which is recognized for its deeper architecture and skip connections, obtained 94% accuracy, whereas AlexNet, a pioneering CNN with a shallower design, achieved 96%. Despite ResNet's resistance to over-fitting and greater performance on a variety of tasks, including picture classification, AlexNet achieved slightly higher accuracy in this particular situation. The decision to choose AlexNet over ResNet in the comparative model indicates that, despite its simpler architecture, AlexNet was better suited to the spine abnormality classification assignment for this dataset. AlexNet's improved performance in this case can be ascribed to a variety of variables. Its shorter architecture may have been better suited to the dataset's size and task complexity, needing fewer parameters and computational resources than ResNet. Furthermore, AlexNet's dropout regularization may have effectively reduced overfitting, adding to its improved accuracy. Furthermore, the unique qualities of the spine abnormality photos, such as the existence of distinguishing features that AlexNet could effectively capture, may have favored its performance over ResNet's. Furthermore, conducting clinical trials or working with medical specialists to test the model's predictions and examine their influence on patient care will increase its credibility and utility in clinical settings. Furthermore, investigating strategies for model interpretability and explain ability would increase clinicians' faith in the system and allow its integration into medical workflows. Finally, continued development and validation of the model are required to ensure its usefulness and safety in identifying spine problems and improving patient outcomes.

Future Directions

Various avenues remains open future work. Those includes expanding datasets with multicenter and multimodal images, which are compared against SOTA models includes vision transformers, UIU-Net along with complete ablation and interpretability studies that combines system validation and multimodal clinical data in real world hospital workflows.

Ethical Statement

The X-ray images were obtained from a publicly available and fully de-identified dataset (Fraiwan et al. [23]; Mendeley Data, DOI: 10.17632/xkt857dsxk.1). Therefore, ethical approval and informed consent were not required.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in Mendeley Data at http://doi.org/10.17632/xkt857dsxk.1, reference number [23].

Author Contribution Statement

Gobalakrishnan Natesan: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, Project administration. **Anbarasan Murugesan:** Software, Formal analysis, Investigation, Resources, Data curation, Writing – review & editing, Visualization, Supervision. **Ramshankar Nagarajan:** Software, Formal analysis, Investigation, Resources, Data curation, Writing – review & editing, Visualization.

References

[1] Meng, N., Cheung, J. P. Y., Wong, K.-Y. K., Dokos, S., Li, S., Choy, R. W., ..., & Zhang, T. (2022). An artificial intelligence powered platform for auto-analyses of spine alignment irrespective of image quality with prospective validation. *eClinical Medicine*, *43*, 101252. https://doi.org/10.1016/j.eclinm.2021.101252

[2] Velmurugan, N., Rajeswari, R., Naganjaneyulu, S., & Anupama, A. (2025). Rat swarm political optimizer based deep learning approach for lung lobe segmentation and lung cancer detection using CT images. *Biomedical Signal Processing and Control*, *105*, 107612. https://doi.org/10.1016/j.bspc.2025.107612

[3] Kaplan, K. M., Spivak, J. M., & Bendo, J. A. (2005). Embryology of the spine and associated congenital abnormalities. *The Spine Journal*, *5*(5), 564–576. https://doi.org/10.1016/j.spinee.2004.10.044

[4] Lu, H., Li, M., Yu, K., Zhang, Y., & Yu, L. (2023). Lumbar spine segmentation method based on deep learning. *Journal of Applied Clinical Medical Physics*, *24*(6), e13996. https://doi.org/10.1002/acm2.13996

[5] Singh, A., Nagabhooshanam, N., Kumar, R., Verma, R., Mohanasundaram, S., Manjith, R., ..., & Rajaram, A. (2025). Deep learning based coronary artery disease detection and segmentation using ultrasound imaging with adaptive gated SCNN models. *Biomedical Signal Processing and Control*, *105*, 107637. https://doi.org/10.1016/j.bspc.2025.107637

[6] Viancy, V., & Gobalakrishnan, N. (2025). Intelligent food quality monitoring: A hybrid dilated convolutional network and BiLSTM-MSVM model for robust beef quality assessment. *Food Control*, *176*, 111396. https://doi.org/10.1016/j.foodcont.2025.111396

[7] Peng, Y., Wang, Y., Hu, F., He, M., Mao, Z., Huang, X., & Ding, J. (2024). Predictive modeling of flexible EHD pumps using Kolmogorov–Arnold Networks. *Biomimetic Intelligence and Robotics*, *4*(4), 100184. https://doi.org/10.1016/j.birob.2024.100184

[8] Li, C., Zhang, B., Hong, D., Jia, X., Plaza, A., & Chanussot, J. (2024). Learning disentangled priors for hyperspectral anomaly detection: A coupling model-driven and data-driven paradigm. *IEEE Transactions on Neural Networks and Learning Systems*, *36*(4), 6883-6896. https://doi.org/10.1109/TNNLS.2024.3401589

[9] McKay, S. D., Al-Omari, A., Tomlinson, L. A., & Dormans, J. P. (2012). Review of cervical spine anomalies in genetic syndromes. *Spine*, *37*(5), E269–E277. https://doi.org/10.1097/BRS.0b013e31823b3ded

[10] Bradford, D. S., Heithoff, K. B., & Cohen, M. (1991). Intraspinal abnormalities and congenital spine deformities: A radiographic and MRI study. *Journal of Pediatric Orthopaedics*, *11*(1), 36–41. https://doi.org/10.1097/01241398-199101000-00009

[11] Wang, D., Sun, Y., Tang, X., Liu, C., & Liu, R. (2023). Deep learning-based magnetic resonance imaging of the spine in the diagnosis and physiological evaluation of spinal metastases. *Journal of Bone Oncology*, *40*, 100483. https://doi.org/10.1016/j.jbo.2023.100483

[12] Germann, C., Meyer, A. N., Staib, M., Sutter, R., & Fritz, B. (2023). Performance of a deep convolutional neural network for MRI-based vertebral body measurements and insufficiency fracture detection. *European Radiology*, *33*(5), 3188–3199. https://doi.org/10.1007/s00330-022-09354-6

[13] Natalia, F., Young, J. C., Afriliana, N., Meidia, H., Yunus, R. E., & Sudirman, S. (2022). Automated selection of mid-height intervertebral disc slice in traverse lumbar spine MRI using a combination of deep learning feature and machine learning classifier. *PLoS One*, *17*(1), e0261659. https://doi.org/10.1371/journal.pone.0261659

[14] Merali, Z., Wang, J. Z., Badhiwala, J. H., Witiw, C. D., Wilson, J. R., & Fehlings, M. G. (2021). A deep learning model for detection of cervical spinal cord compression in MRI scans. *Scientific Reports*, *11*(1), 10473. https://doi.org/10.1038/s41598-021-89848-3

[15] Biercher, A., Meller, S., Wendt, J., Caspari, N., Schmidt-Mosig, J., de Decker, S., & Volk, H. A. (2021). Using Deep Learning to detect spinal cord diseases on thoracolumbar magnetic resonance images of dogs. *Frontiers in Veterinary Science*, *8*, 721167. https://doi.org/10.3389/fvets.2021.721167

[16] Merali, Z. A., Colak, E., & Wilson, J. R. (2021). Applications of machine learning to imaging of spinal disorders: Current status and future directions. *Global Spine Journal*, *11*(1_suppl), 23S–29S. https://doi.org/10.1177/2192568220961353

[17] Almansour, H., Herrmann, J., Gassenmaier, S., Afat, S., Jacoby, J., Koerzdoerfer, G., Nickel, D., ..., & Othman, A. E. (2022). Deep learning reconstruction for accelerated spine MRI: Prospective analysis of interchangeability. *Radiology*, *306*(3), e212922. https://doi.org/10.1148/radiol.212922

[18] Tavana, P., Akraminia, M., Koochari, A., & Bagherifard, A. (2023). Classification of spinal curvature types using radiography images: Deep learning versus classical methods. *Artificial Intelligence Review*, *56*(11), 13259–13291. https://doi.org/10.1007/s10462-023-10480-w

[19] Kim, H. J., Yang, J. H., Chang, D.-G., Lenke, L. G., Suh, S. W., Nam, Y., ..., & Suk, S.-I. (2022). Adult spinal deformity: A comprehensive review of current advances and future directions. *Asian Spine Journal*, *16*(5), 776–788. https://doi.org/10.31616/asj.2022.0376

[20] Tavana, P., Akraminia, M., Koochari, A., & Bagherifard, A. (2023). An efficient ensemble method for detecting spinal curvature type using deep transfer learning and soft voting classifier. *Expert Systems with Applications*, *213*, 119290. https://doi.org/10.1016/j.eswa.2022.119290

[21] Mbarki, W., Bouchouicha, M., Frizzi, S., Tshibasu, F., Farhat, L. B., & Sayadi, M. (2020). Lumbar spine discs classification based on deep convolutional neural networks using axial view MRI. *Interdisciplinary Neurosurgery*, *22*, 100837. https://doi.org/10.1016/j.inat.2020.100837

[22] Cousins, J. P., & Haughton, V. M. (2009). Magnetic resonance imaging of the spine. *Journal of the American Academy of Orthopaedic Surgeons*, *17*(1), 22–30. https://doi.org/10.5435/00124635-200901000-00004

[23] Fraiwan, M., Audat, Z., & Manasreh, T. (2022). A dataset of scoliosis, spondylolisthesis, and normal vertebrae X-ray images. *Mendeley Data*, *1*, 2022. http://doi.org/10.17632/xkt857dsxk.1

[24] Kim, M., Yun, J., Cho, Y., Shin, K., Jang, R., Bae, H., & Kim, N. (2019). Deep learning in medical imaging. *Neurospine*, *16*(4), 657–668. https://doi.org/10.14245/ns.1938396.198

[25] Nanda, R. H., Hua, C.-H., Flampouri, S., Eaton, B., Kaste, S., Patni, T., ..., & Esiashvili, N. (2024). Risks of spinal abnormalities and growth impairment after radiation to the spine in childhood cancer survivors: A PENTEC comprehensive review. *International Journal of Radiation Oncology*Biology*Physics*, *119*(2), 507–521. https://doi.org/10.1016/j.ijrobp.2023.10.039

[26] Takai, K., Endo, T., & Komori, T. (2025). Calcified Hofmann's ligaments as the cause of spinal cerebrospinal fluid leaks associated with spinal ventral dural tears. *Journal of Neurosurgery: Spine*, *42*(1), 43–48. https://doi.org/10.3171/2024.7.SPINE24480

How to Cite: Natesan, G., Murugesan, A., & Nagarajan, R. (2025). Lightweight CNN-Enabled Framework for Automated Detection of Scoliosis and Spondylosis Using Spine X-rays. *Journal of Computational and Cognitive Engineering*. https://doi.org/10.47852/bonviewJCCE52026933