IJASST
International Journal of
Applied Sciences and Smart Technologies

Volume 06
Issue 02

p-ISSN 2655-8564, e-ISSN 2685-943, pp. 321-336, 2024

Automated Detection of Spine Deformities: Advancing Orthopedic Care with Convolutional Neural Networks

Deepesh Pratap[1], Saran Sinha[1], A. Charan Kumari[1*], K. Srinivas[1]

[1]*Department of Electrical Engineering, Faculty of Engineering, Dayalbagh Educational Institute, Dayalbagh, Agra, India*
[*]*Corresponding Author: charankumari@dei.ac.in*

(Received 24-07-2024; Revised 04-08-2024; Accepted 06-08-2024)

Abstract

This paper proposes Spine-CNN, a deep learning model for the detection of spinal deformities that can assist orthopedic doctors as a reliable tool for diagnosis. This technology promises to dramatically simplify the diagnostic process, freeing valuable time, and resources for healthcare professionals. To achieve this objective, a dataset of spine deformity X-ray images was curated from the PhysioNet database. The Spine-CNN was specially designed for detecting the spine deformity by incorporating features to leverage its ability to extract intricate features from radiographic images and by fine tuning the hyperparameters to properly train the model. Model performance was evaluated using standard metrics. Results from the Spine-CNN demonstrated promising performance in detecting spinal deformities. The model achieved an accuracy of 74%, with precision, recall, and F1-score values of 77%, 70%, and 73% respectively. Specifically, this research work introduces a Spine-CNN that underscore the potential of deep learning techniques to revolutionize diagnostic practices in orthopedic medicine, leading to improved treatment outcomes and patient care.

Keywords: Computer-aided detection, Convolutional neural network, Image classification, Spine Deformation, X-ray imaging

1 Introduction

Malformations of the spine, for instance, scoliosis or kyphosis, are physical impairments that affect the patient's health situation and his or her quality of life. Spinal deformities inevitably lead to discomfort, if left untreated, to life-threatening complications. Just counting injuries to the spinal cord, between 250,000 and 500,000 people are affected per year worldwide. The lethality of injuries to the spine is two to five times higher than it is without injuries, in low and middle-income countries even higher. The orthopedist must recognize these deformities early and make a precise diagnosis to initiate effective medical treatment.

The evaluation of spine deformity in conventional form leans mainly on straightforward measurements and radiographic analysis. Long-term methods present few obstacles in the direction of precision and performance. Straight measurements are partly subject to careless fallacy, and radiographic inspection requests particular resources and skills. Additionally, they tend to be more time-intensive than these approaches potentially slow up correct diagnosis and effective administration.

Over the last few years, deep learning has made major strides, particularly with Convolutional Neural Networks (CNNs). This technology has shown great promise in taking up medical image analysis. CNNs are artificial neural networks that are particularly well suited to pulling out characteristics from complex visual data. This talent makes them good candidates to look for structural irregularities in medical images, like those used in diagnosing spinal deformities.

The aim of this research is to improve the accuracy and swiftness of spine deformity detection by leveraging the capabilities of CNNs. Employing a dataset made up of medical imaging, the Spine-CNN trained in this research offers the possibility of establishing a resilient system. Although this is accurate, various other crucial determinants are necessary when deciding on the course of treatment. A patient's global health, current condition, and the extent to which the curvature has altered their breathing are all significant facets. Additionally, it is important to recommend and develop additional methods that aid in diagnosing scoliosis abnormalities with a high level of accuracy.

Several works have represented strategies that rely on various geometric models centered on depicting spinal curvature. These models lead to a detailed and exact depiction of spinal shape and flexure, benefitting diagnosis and treatment of spinal deformity. By fusing these geometric models with the power of CNNs, this work enabled a holistic and convincing framework for spine deformity detection and treatment.

The primary contributions of this research paper are as follows:

1. Implementation of Spine-CNN, a deep learning model to effectively categorize spinal deformities from image datasets, presenting a novel approach to addressing a critical medical challenge
2. A meticulously curated dataset is used that encompasses diverse images representing various spinal disorders, facilitating robust model training to tackle real-world complexities and enhance performance in handling clinical scenarios.
3. Enhancement of model performance through rigorous hyperparameter optimization, ensuring optimal training and fine-tuning of Spine-CNN for accurate spine deformity classification.
4. A comprehensive evaluation of the model's effectiveness by employing sensitivity, specificity, and precision calculations, providing insights into its ability to precisely classify spinal irregularities.
5. To leverage the capabilities of CNNs to speed up accurately diagnosing spine deformities, with a vision to develop a computerized diagnostic system equipped with extensive medical image data, thereby empowering CNNs to achieve advanced diagnostic capabilities.

The following is the arrangement of the upcoming sections of the paper: A concise synopsis of the literature study is given in Section 2. The employed methodology is explained in Section 3. Section 4 highlights the obtained results together with its analysis. The conclusions from the study's findings and the direction of future research are summarized in Section 5.

2 Literature Review

This section presents a brief overview of spine deformity detection strategies in a concise way, covering both conventional image processing methods and the most recent developments in deep learning. The objective is to present a comprehensive analysis of the current literature, highlighting the advantages and disadvantages of different strategies to provide insightful background information for the selected methodology.

Zhang L. et al. [1] performed human X-ray image spine model positioning based on an R-CNN. They concentrated on the placement, detection, and segmentation of human radiograph spine models using Mask R-CNN, as well as object detection and image segmentation. They achieved good accuracy by concentrating on radiographic spine model location while considering application instructions. Lin. H. [2] implemented a multilayer feed-forward, back-propagation (MLFF/BP) artificial neural network (ANN) to identify the classification patterns of the scoliosis spinal deformity using X-ray images. He achieved fairly good accuracy. Lee et al. [3] created a CNN model to diagnose CSM using only one lateral cervical spine radiograph, with an acceptable diagnostic accuracy. However, the study was limited by the small number of subjects with MR images. Kim et al. [4] proposed a technique for analyzing a moire image of a human back in a 2-D way to automate the primary screening of spinal deformity detection based on neural network. Saravi et al. [5] concluded that AI-based decision-making tools in spine surgery utilize multimodal data to predict outcomes and detect disease patterns, requiring collaboration between healthcare providers and industries to implement good machine learning principles. Leveraging techniques like Federated Learning and continuous updates with new data are essential for safe integration into clinical practice. Mezghani N. et al. [6] studied Computer applications employing fuzzy clustering, support vector classifiers, artificial neural networks (ANN), and surface topography algorithms aid in managing Adolescent Idiopathic Scoliosis (AIS) by regrouping similar spine geometries, predicting Cobb angles accurately, and enhancing classification reliability. Salehi et al. [7] applied the deep convolutional neural networks (CNNs) in computer-aided diagnosis of three types of disc herniation disease based on lumbar axial MR Images. Pinheiro et al. [8] proposed and validated a novel computerized methodology for detecting elliptical patterns from X-ray images to evaluate the extent of the underlying scoliotic deformity. Hieu T. Nguyen et al. [9] aimed at developing and evaluating a deep learning-based framework, named VinDr-SpineXR, for the classification and localization of abnormalities from spine X-rays. B. L. Qasthari et al. [10] used a pre-trained model VGG19 to categorize histological images of lung and colon cancer into five labels to aid medical professionals' categorization job.

The literature review revealed several potential designs for identifying spinal deformities, with an emphasis on the CNN model. While some research has produced promising results in the recognition and categorization of spinal deformities, other studies encountered challenges with achieving greater accuracy. Pretrained models and transfer learning techniques have shown potential for improving performance across multiple tasks.

3 Methodology

This section presents the description of the dataset, Spine-CNN architecture, the model parameters, and performance evaluation metrics.

Dataset

This research work uses PhysioNet dataset contributed by Pham *et al.* [11] supported with comma-separated values (CSV) labels for the spinal X-ray DICOM images with a total of 10466 images. It has 13 types of abnormalities and contains basic demographic information. The dataset was divided into training, testing and validation sets of images, with 8389 images in the training dataset and 2077 in validation and testing.

All collected images are organized into two separate folders based on their respective classes: "Deform" and "Normal". This structuring facilitates the efficient handling and processing of data during model development. The dataset is split into three subsets: training, testing, and validation sets - 80% of the data is allocated for training and 20% of the data is reserved for testing and validation.

Model Architecture of Spine-CNN

The model summary of the Spine-CNN architecture is shown in Table 1.

Table 1. Architecture of Spine-CNN

Layer (type)	**Output Shape**	**Param**
conv2d (Conv2D)	(None, 222, 222, 16)	160
batch_normalization (BatchNormalization)	(None, 222, 222, 16)	64
activation (Activation)	(None, 222, 222, 16)	0
max_pooling2d (MaxPooling2d)	(None, 111, 111, 16)	0
conv2d_1 (Conv2D)	(None, 109, 109, 32)	4640

batch_normalization_1 (BatchNormalization)	(None, 109, 109, 32)	128
activation_1 (Activation)	(None, 109, 109, 32)	0
max_pooling2d_1 (MaxPooling2D)	(None, 54, 54, 32)	0
conv2d_2 (Conv2D)	(None, 52, 52, 32)	9248
batch_normalization_2 (BatchNormalization)	(None, 52, 52, 32)	128
activation_2 (Activation)	(None, 52, 52, 32)	0
dropout (Dropout)	(None, 52, 52, 32)	0
max_pooling2d_2 (MaxPooling2D)	(None, 26, 26, 32)	0
conv2d_3 (Conv2D)	(None, 24, 24, 64)	18496
batch_normalization_3 (BatchNormalization)	(None, 24, 24, 64)	256
activation_3 (Activation)	(None, 24, 24, 64)	0
dropout_1 (Dropout)	(None, 24, 24, 64)	0
max_pooling2d_3 (MaxPooling2D)	(None, 12, 12, 64)	0
flatten (Flatten)	(None, 9216)	0
dense (Dense)	(None, 256)	2359552
batch_normalization_4 (BatchNormalization)	(None, 256)	1024
activation_4 (Activation)	(None, 256)	0
dropout_2 (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 1)	257
Total params: 2393953 (9.13 MB)		
Trainable params: 2393153 (9.13 MB)		
Non-trainable params: 800 (3.12 KB)		

It consists of the following layers.

A. Convolutional Layers (Conv2D):

Four convolutional layers are employed to extract hierarchical features from input images. These layers use 3x3 filters with varying depths (16, 32, 32, 64) to capture different levels of abstraction.

B. Batch Normalization Layers:

The training process is stabilized and accelerated by interleaving convolutional layers with batch normalization layers. They normalize activations, improving gradient flow and mitigating the vanishing gradient problem.

C. Activation Layers (Activation):

ReLU activation functions introduce non-linearity, enabling the model to learn complex patterns in the data. They follow each convolutional layer to introduce non-linear transformations.

D. Max Pooling Layers (MaxPooling2D):

Max pooling layers downsample feature maps, reducing computational complexity and prevent overfitting. 2x2 pooling windows are used to retain the most salient features.

E. Dropout Layer:

A total of three dropout layers are utilized. The Convolution layer consists of two dropouts at a probability of 0.2, while the dense layer is equipped with a single dropout at a probability of 0.3.

F. Flatten Layer (Flatten):

The flatten layer comes after the last max pooling layer. It transforms the 2D feature maps into a 1D vector, so that they can be fed into the fully connected layers.

G. Fully Connected Layers:

The model contains two dense layers for classification, with the initial dense layer possessing 256 neurons that can capture high-level features and the last layer consisting of only a single neuron aimed at tasks focused on binary classification, thereby outputting classes. The model is made up of 2,393,953 parameters.

The process flow of the spine deformation detection of Spine-CNN is presented in Fig. 1, where process starts with the input X-ray image, which is preprocessed by resizing to 224x224 pixels and rescaling pixel values for normalization. This preprocessed image is passed through the Spine-CNN Model explained above which uses a sigmoid activation function to output a probability score, which is thresholded to classify the image as either "Deform" (indicating a spinal deformity) or "Normal" (indicating no deformity).

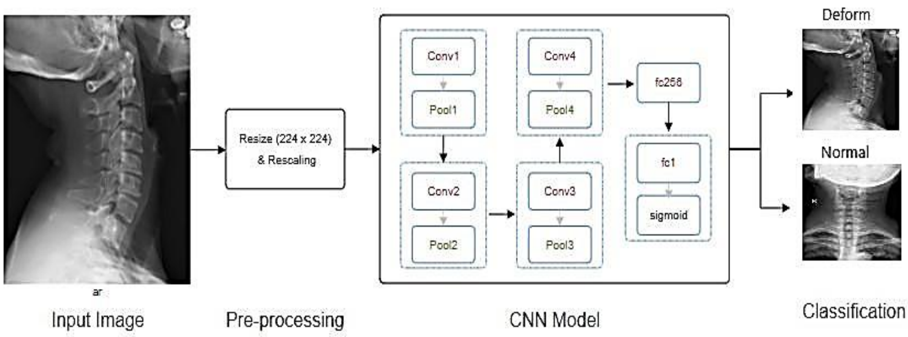

Figure 1. Flow diagram for Spine deformity detection

3.3 Model parameters

The following parameters are used in the Spine-CNN model:

- Epochs: 50
- Batch Size: 64
- Loss Function: Binary Cross-Entropy
- Optimizer: Adam
- Learning rate = 0.001

Learning rate scheduler: ReduceLROnPlateau (monitor='val_loss')

Evaluation metrics

While evaluating a Convolutional Neural Network (CNN) model for image classification, several metrics are used to assess its performance. A brief description of these metrics is presented below:

1. Accuracy:

Accuracy is defined as the proportion of accurately identified instances to all instances as shown in equation 1.

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

2. Precision:

The precision metric, as described in equation 2, quantifies the percentage of true positive predictions among all the positive predictions generated by the model. It shows that the model can prevent false positives.

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \tag{2}$$

3. Recall (Sensitivity):

The percentage of true positive predictions among all actual positive data instances is known as recall. It illustrates how the model may prevent false negatives by capturing all positive cases as depicted in equation 3.

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \tag{3}$$

4. F1-Score:

The F1-score, as described in equation 4, is the harmonic mean between recall and precision. It offers an equitable assessment of a model's effectiveness, particularly in cases where the dataset exhibits class imbalances.

$$F1 - Score = 2\ \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

5. Confusion Matrix:

A tabular overview of the model's predictions compared to the actual class labels is given by a confusion matrix. It allows for a more in-depth analysis of the model's performance, showing the counts of true negatives, true positives, false negatives, and false positives.

6. Receiver Operating Characteristic Curve (ROC) and Area Under the Curve (AUC):

ROC curves illustrate how threshold values alter the trade-off between the true positive rate (sensitivity) and the false positive rate (1-specificity). The model's overall performance is represented by a single scalar value, or AUC, which summarizes the ROC curve.

4 Results

This section reports the results obtained by Spine-CNN. Table 2 presents the results of the model after 50 epochs of training.

Accuracy: The accuracy shows overall accuracy of the model's predictions. The model achieves an accuracy of 74%, suggesting that it correctly identifies spinal deformities approximately three-fourths of the time. Although accuracy is a valuable metric, it should be read in conjunction with other metrics to provide a complete picture of the model's performance.

Precision: The percentage of true positive predictions among all positive predictions the model made is referred to as precision. The model's precision is 77%, meaning that 77% of the time, it accurately predicts a spine deformity. In medical applications, a low percentage of false positives is indicative of high precision and helps prevent misdiagnosis.

Recall: The percentage of true positive predictions among all actual positive cases in the dataset is determined by recall, which can also be referred to as sensitivity. Here, the model achieves a recall of 70%, implying that it identifies 70% of all actual spine

Table 2. Performance of Spine-CNN

Metric	Value
Accuracy	74%
Precision	77%
Recall	70%
F1-Score	73%
ROC AUC	81%

deformities. A high recall indicates that the model effectively captures most instances of spinal deformities, reducing the chances of false negatives.

F1-Score: The harmonic mean of precision and recall results in the F1-score, which provides a balance between the two measures. It is especially helpful in cases where there is an imbalance in classes. The F1-score in this instance is 73%, suggesting that recall and precision are fairly balanced.

ROC AUC: The Receiver Operating Characteristic Area Under Curve (ROC AUC) evaluates the model's ability to differentiate between classes that are negative and positive at different thresholds. The model's 81% AUC indicates that it can distinguish between positive and negative occurrences with reasonable accuracy.

Fig. 2 shows the obtained ROC curve, where the x-axis represents the False Positive Rate (FPR), while the y-axis represents the True Positive Rate (TPR). The curve plots TPR against FPR at various threshold settings, illustrating the trade-off between sensitivity (true positive rate) and specificity (false positive rate).

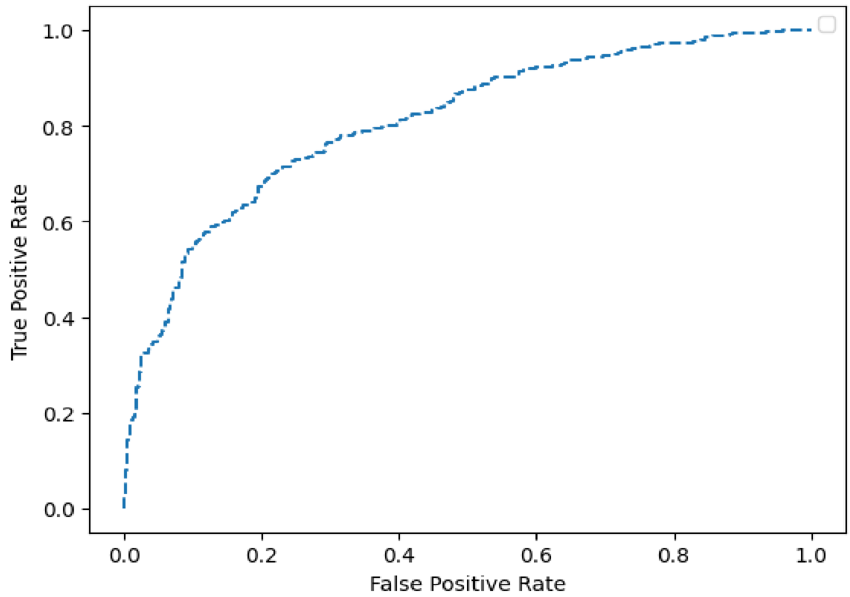

Figure 2. ROC Curve

Overall, the results demonstrate promising performance of Spine-CNN for spinal deformity detection. The precision and recall values indicate a balanced performance in correctly identifying deformities while minimizing false positives and false negatives. Althorough analysis of the model's predictions is shown with a confusion matrix in Fig. 3

True Positives (TP): A total of 374 instances of spine deformities were accurately predicted by the model. (Positive class)

False Positives (FP): 114 instances were incorrectly classified by the model as spine abnormalities when they weren't. (Negative class)

True Negatives (TN): For 400 instances, the model accurately predicted that there were no spine deformities. (Negative class)

False Negatives (FN): 159 instances were wrongly predicted by the model to not have spinal abnormalities when in fact they had. (Positive class).

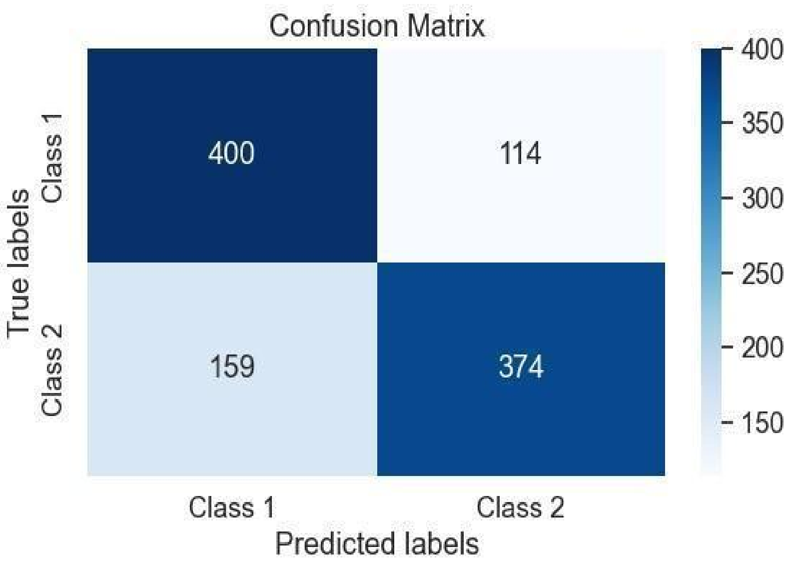

Figure 3. Confusion Matrix obtained by Spine-CNN

5 Conclusions and Future work

To conclude, this research introduces a method for deep learning utilizing Convolutional Neural Networks (CNNs) to detect spinal abnormalities using Spine-CNN. This study designed and implemented a CNN model to detect and differentiate kinds of spine deformities substantially. This work succeeded in presenting that through intensive experimentation and fine-tuning a CNN of substantial size which could achieve an accuracy of 74%. In terms of recall, precision, and F1 score, the Spine-CNN achieved impressive results, with 70%, 77%, and 73%, respectively. The study emphasizes the potential of deep learning technology, specifically Convolutional Neural Networks (CNNs), to change orthopedic medicine. With deep learning, the CNN model fuels a powerful engine for automatically detecting intricate details from radiographic imaging, which could make the diagnosis process smarter, thereby improving subsequent care and perhaps saving time and resources for all clinicians.

Future research endeavors aimed at enhancing the diagnostic capabilities of Spine-CNN could investigate the potential integration of 3D imaging data obtained from modalities such as MRI or CT scans, offering a more comprehensive understanding of spinal structures. Furthermore, the incorporation of multi-modal data, such as integrating X-ray images with a patient's medical history or other clinical information, holds significant potential for a thorough evaluation, thereby improving diagnostic outcomes.

To ensure Spine-CNN's predictions are reliable in clinical settings, it can be integrated with explainable AI (XAI) techniques that allow clinicians to interpret the model's decisions. By providing insights into which features of the X-ray images the model focuses on, XAI will enable doctors to understand the rationale behind each prediction. This transparency will certainly gain clinician confidence, as it allows for cross-verification with medical knowledge and enhances the model's utility in real-world clinical settings to ensure its practical applicability.

Acknowledgements

The authors are extremely grateful to the Revered Prof. P.S. Satsangi, Chairman, Advisory Committee on Education, Dayalbagh for continued guidance and support.

References

[1]. L. Zhang, J. Zhang, and S. Gao, "Region-based convolutional neural network-based spine model positioning of X-ray images," BioMed Research International, vol. 2022, Art. no. 7512445, 2022. [Online]. Available: https://doi.org/10.1155/2022/7512445.

[2]. Lin, "Identification of spinal deformity classification with total curvature analysis and artificial neural network," IEEE Transactions on Biomedical Engineering, vol. 55, no. 1, pp. 376-382, 2008. [Online]. Available: https://doi.org/10.1109/TBME.2007.894831.

[3]. G. W. Lee, H. Shin, and M. C. Chang, "Deep learning algorithm to evaluate cervical spondylotic myelopathy using lateral cervical spine radiograph," BMC Neurology, vol. 22, no. 1, Art. no. 147, 2022. [Online]. Available: https://doi.org/10.1186/s12883-022-02670-w.

[4]. H. Kim, S. Ishikawa, M. Khalid, Y. Otsuka, H. Shimizu, Y. Nakada, T. Shinomiya, and M. A. Viergever, "Automatic spinal deformity detection based on neural network," in Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003, R. E. Ellis and T. M. Peters, Eds., Lecture Notes in Computer Science, vol. 2878. Berlin, Heidelberg: Springer, 2003, pp. 1014-1016. [Online]. Available: https://doi.org/10.1007/978-3-540-39899-8_98.

[5]. B. Saravi, F. Hassel, S. Ülkümen, A. Zink, V. Shavlokhova, S. Couillard-Despres, M. Boeker, P. Obid, and G. M. Lang, "Artificial intelligence-driven prediction modeling and decision making in spine surgery using hybrid machine learning models," Journal of Personalized Medicine, vol. 12, no. 4, Art. no. 509, 2022. [Online]. Available: https://doi.org/10.3390/jpm12040509.

[6]. N. Mezghani, P. Phan, H. Labelle, C. Aubin, and J. Guise, "Computer-aided Lenke classification of scoliotic spines," International Journal of Computer and

Information Engineering, vol. 3, no. 5, pp. 1351-1354, 2009. World Academy of Science, Engineering and Technology, Open Science Index 29.

[7]. E. Salehi, S. Khanbare, H. Yousefi, H. Sharpasand, and O. S. Sheyjani, "Deep convolutional neural networks for automated diagnosis of disc herniation on axial MRI," in Proc. 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), Istanbul, Turkey, 2019, pp. 1-6. doi: 10.1109/EBBT.2019.8741895.

[8]. A. P. Pinheiro, J. C. Coelho, A. C. P. Veiga, and T. Vrtovec, "A computerized method for evaluating scoliotic deformities using elliptical pattern recognition in X-ray spine images," Computer Methods and Programs in Biomedicine, vol. 161, pp. 85–92, 2018. doi: 10.1016/j.cmpb.2018.04.015.

[9]. H. T. Nguyen, H. H. Pham, N. T. Nguyen, H. Q. Nguyen, T. Q. Huynh, M. Dao, and V. Vu, "VinDr-SpineXR: A deep learning framework for spinal lesions detection and classification from radiographs," in Proc. International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2021.

[10]. B. L. Qasthari, E. Susanti, and M. Sholeh, "Classification of lung and colon cancer histopathological images using convolutional neural network (CNN) method on pre-trained models," International Journal of Applied Sciences and Smart Technologies, vol. 5, no. 1, pp. 133-142. [Online]. Available: https://doi.org/10.24071/ijasst.v5i1.6325.

[11]. H. H. Pham, H. Nguyen Trung, and H. Q. Nguyen, "VinDr-SpineXR: A large annotated medical image dataset for spinal lesions detection and classification from radiographs (version 1.0.0)," PhysioNet, 2021. doi: 10.13026/q45h-5h59.

This page intentionally left blank