A DEEP LEARNING-BASED APPROACH TO DETECTION OF SPINAL DEFORMITIES

Tuğba Özmen
Departmartment of Biomedical Engineering, Institute of Natural Sciences, Sakarya University;
https://orcid.org/0009-0001-1161-4505
tugba.ozmen@ogr.sakarya.edu.tr
Corresponding author

Hüseyin Eski,
Assistant Professor
Department of Computer Engineering, Sakarya University
https://orcid.org/0000-0002-6006-3228
heski@sakarya.edu.tr

ABSTRACT
As spinal deformities reach advanced levels, they reduce the patient's quality of life and make diagnosis and treatment difficult. Therefore, early diagnosis is very important in order to stop the progression of the deformity and start the treatment process on time. The diagnostic procedure for deformities is based on manual measurements made by specialist doctors on medical images. Manual measurements are error-prone, time-consuming and they rely on subjective opinion. In this study, a deep learning-based automatic diagnosis method is presented to eliminate these disadvantages in the diagnosis of spinal deformities.
In the light of literature research, it has been observed that deep learning-based convolutional neural networks produce high-accuracy results in image processing and segmentation processes. In this study, the X-ray image classification performances of ResNet and GoogleNet architectures, which are convolutional neural network architectures reported to have high accuracy and performance, will be compared in terms of accuracy, sensitivity, precision, F1 score, specificity, training and testing times.
As a result of the experiments conducted within the scope of this study, it was seen that ResNet showed superior performance than GoogleNet in terms of accuracy, precision, F1 score, specificity and sensitivity, and when a comparison was made in terms of training and testing times, it was seen that GoogleNet processed faster than ResNet.
Keywords: Scoliosis, Spondylolisthesis, CNN, GoogleNet, ResNet

1. Introduction

The spine is a multi-jointed bony structure that extends between the head and pelvis, supports the trunk during daily activity and standing, contains the most important component of the central nervous system, and whose primary function is to protect the spinal cord from external factors [1]. The spine also maintains mechanical balance in order to adapt to gravitational stress, has a straight form in the coronal plane and an S-form in the sagittal plane. For the sagittal plane, the normal boundaries of the spine are determined by the lumbar lordosis and thoracic kyphosis angles [2]. Abnormal changes that occur when the spinal alignment goes beyond these defined normal limits are called spinal deformities. While the spine, which should have a straight form in the coronal plan, angles more than 10 degrees to the right, left, or both right and left, is called scoliosis, while in the sagittal plane, the lumbar lordosis and thoracic kyphosis angles are larger than they should be, called hyperkyphosis and hyperlordosis. In addition, spondylolisthesis, another deformity that occurs in the sagittal plane, is defined as the sliding of adjacent vertebrae over each other, creating the appearance of a staircase [3], [4], [5], [6].

Although spinal deformities are generally idiopathic, they can also occur as a result of various traumas or as a result of aging. When deformities reach advanced levels, they reduce the patient's quality of life and make diagnosis and treatment difficult. Therefore, early diagnosis is very important to stop the progression of the deformity and start the treatment process on time [2]. The diagnostic procedure for deformities is based on manual measurements made by specialist doctors on medical images. Manual measurements are error-prone, time-consuming and they rely on subjective opinion. In this study, a deep learning-based automatic diagnosis method is presented to eliminate these disadvantages in diagnosing spinal deformities. The aim of shortening the diagnosis and treatment processes with automatic deformity diagnosis method is to reduce errors, optimize hospital costs and develop an objective diagnostic method.

In studies conducted for the purpose of diagnosis from deep learning-based medical images, convolutional neural networks are used as a popular and reliable method and provide successful performance results. Fraiwan M. and his colleagues compared the classification performances of the convolutional neural network architectures they used in their deep transfer learning-based studies for the automatic diagnosis of scoliosis and spondylolisthesis. They reported that they achieved the highest accuracy values with the DensNet-201 architecture [7]. In studies on deep learning-based spondylolisthesis diagnosis, Trinh G. M. and his colleagues compared the performance of convolutional neural network architectures LumbarNet and UNet and reported that LumbarNet had higher accuracy [8], while Zhang J. and his colleagues used RCNN and RetinaNet architectures and reported that RCNN had higher performance [9]. In their study on the diagnosis of spondylolisthesis, Varçın F. and his colleagues compared the performances of AlexNet and GoogleNet architectures and reported that GoogleNet had higher classification accuracy [10]. In studies on automatic scoliosis angle measurement, Alharbi R. and his colleagues used ResNet, one of the convolutional neural network architectures, and reported that they achieved high accuracy [11]. Tu Y. and his colleagues used DU-Net deep architecture for automatic Cobb angle measurement and compared the results with the measurements of expert doctors. They reported that the automatic system they designed gave similar results to specialist doctor measurements [12]. In the light of literature research, it has been observed that deep learning-based convolutional neural networks produce high-accuracy results in image processing and segmentation processes. In this study, the X-ray image classification performances of ResNet and GoogleNet architectures, which are convolutional neural network architectures reported to have high accuracy and performance, will be compared.

2. Material and Method

2.1 Deep Learning

Deep learning is a representation-learning technique that has representations at many levels, obtained by combining non-linear and simple layers that transform each representation into a representation at a higher, more abstract level for operations such as classification and feature extraction. Deep learning architectures, which have a structure based on learning many features of the data simultaneously, highlight the important and distinctive parts and suppress the unimportant parts. Inputs from the lower layer determine the features in the upper layer and thus a hierarchical learning representation is formed [13]. In deep learning, which is based on learning from the representation of data, when the image comes to the first layer as input as a series of pixel values, the presence or absence of features in certain positions and directions in the image is represented, while various edges and patterns are detected in the second layer. The patterns detected here are brought together in larger combinations in the third layer, resembling parts of familiar objects, and subsequent layers perceive the object as combinations of these parts. The learning algorithms required for feature extraction are not designed by engineers; deep learning does this by itself learning from data with a general-purpose learning procedure [13]. Deep learning, which is increasingly gaining ground in the medical and health sector, has functions of analyzing medical images, making predictions about the course of diseases in cancer research and playing a guiding role in diagnosis and treatment processes[13].

2.2 Convolutional Neural Network (CNN)

Many deep learning architectures have been developed according to usage areas and requirements, different data properties and different layer types. Convolutional neural networks (CNN), the most commonly used deep learning-based technique in the field of image processing, are inspired by the animal vision system and work on the principle of filtering. CNNs, which have multi-layer feed-forward features, make classification by revealing the defining features of the image with the filters they use [14]. CNNs, which provide successful results in different fields such as audio processing, face recognition, biomedical fields and natural language processing in addition to image processing, include many convolution layers, fully connected layers, activation layers, pooling layers, classification layers and additional layers (Figure 2.1.).

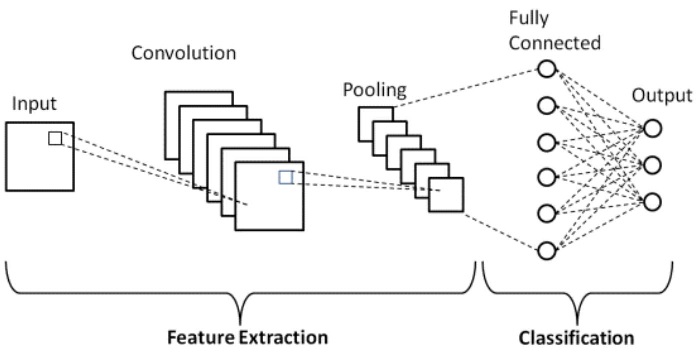

Figure 2.1. CNN architecture [15]

The input layer is the first layer of CNN where data is received in raw form. The size selection of the data received here is very important because the data size affects the memory status, processing and network performance [16].

Convolutional layers are the layers that form the basis of CNNs, where convolution operations are applied to the input images through filters. In the convolution process, a filter is passed over the image step by step, allowing patterns, edges, and features in the image to be detected. After the convolution processes, activation maps are created showing the regions where the features of each filter are discovered [13], [16].

The rectified linear unit layer (ReLU) works as the rectifier unit and is located after the convolution layer. This layer, which accelerates the learning of the network, is also called the activation layer and transforms the network, which entered a linear structure after convolution operations, into a non-linear form again [16].

Pooling layers reduce the size of the input to be sent to the next convolution layer while ensuring that the features learned by the network remain constant even in the face of small changes. Certain filters can be applied to the data at this layer, and the presence of this layer is optional in CNN architectures [16].

Fully connected layers are connected to all areas of the previous layer within the architecture, and the features learned in the convolution layer are integrated by this layer and regression or classification operations are performed [16]. The results produced after the classification process are sent to the output layer and the number of outputs from the output layer depends on the number of objects to be classified.

Transfer learning means that a model previously trained for a different task is later reused for a secondary task. Thanks to transfer learning, instead of training a model from scratch, new models are created with new data sets using the information in the source database [7], [17].

2.3 CNN Architectures Used in the Study

GoogleNet consists of 22 layers and is the 2014 winner of the ImageNet competition. The feature that distinguishes GoogleNet, designed by Szegedy and his friends, from others is the Inception module (Figure 2.2.), which helps keep costs low while increasing the depth and breadth of the network. GoogleNet architecture, parallel config.ures many filters of different sizes and thanks to its parallel modules it optimized the memory cost and prevented the system from memorizing [18].

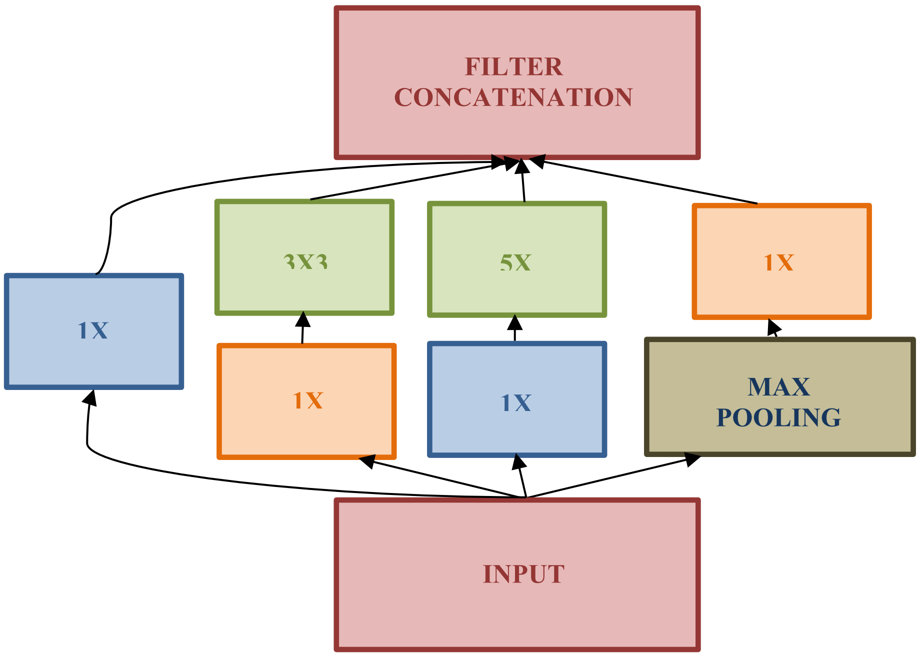

Figure 2.2. Inception module

ResNet consists of 152 layers and is the 2015 winner of the ImageNet competition. The feature that distinguishes ResNet, designed by Zhang and his colleagues, from other architectures is the Residual blocks (Figure 2.3.). Increasing the number of layers in the network causes the efficiency of the network to reach saturation, but then causes a rapid decline. Zhang and his colleagues, who prevented this decrease thanks to the shortcuts they added with the residual block, demonstrated superior performance with an error rate of 3.57% [19].

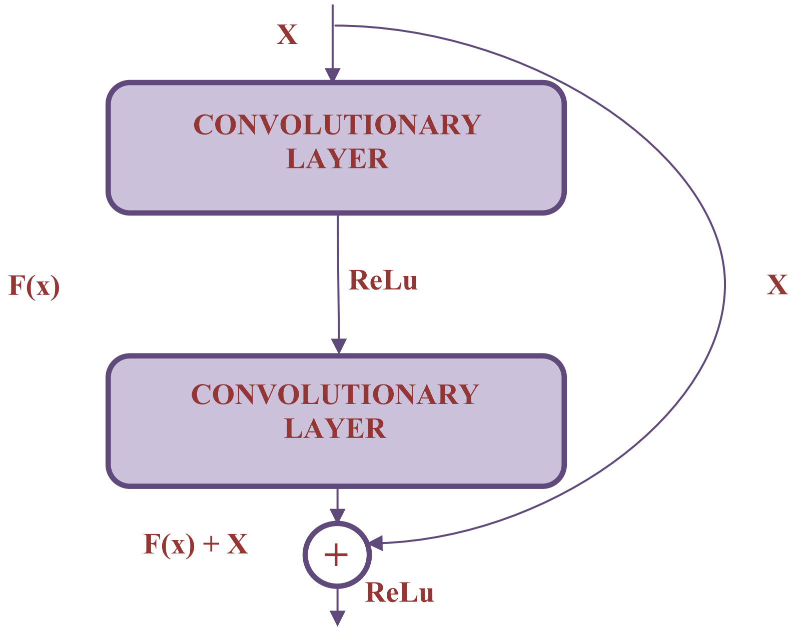

Figure 2.3. Residual block

2.4 Data Set

In this study, a dataset obtained from open sources via Mendeley Data was used, and X-ray images were collected at King Abdullah University Hospital in Jordan [7]. X-ray images of 338 people, 240 women and 98 men, were divided into three groups: 79 with spondylolisthesis, 188 with scoliosis and 71 healthy, and were used in the experimental stages of this study.

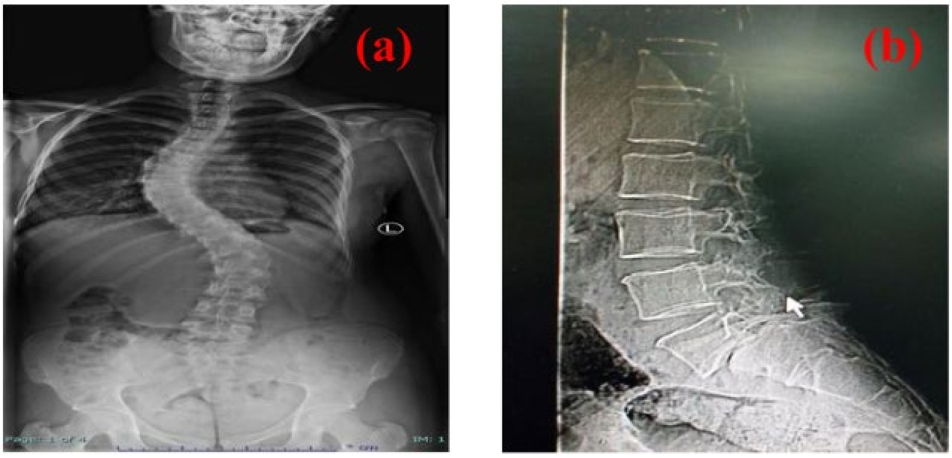

Figure 2.4. X-ray images in dataset **(a)** scoliosis, **(b)** spondylolisthesis

2.5 Experimental Studies

In this study, in which we designed an automatic spinal deformity diagnosis method with deep transfer learning techniques, ResNet and GoogleNet architectures were compared in terms of their performance in classifying X-ray images.

While training and testing the transfer learning-based CNN models, a computer with 8 GB RAM, NVIDIA ® GEFORCE 940MX, Intel ® Core ™ i5-7200U CPU was used and Python programming language was chosen.

At the beginning of the training, the images in the data set were loaded with their own class labels (scol, spond, normal). The data set is divided into 80% training and 20% testing to be used in the training and testing stages. Each architecture was created in its own structural form and initialized to pre-trained intervals in the ImageNet database. After adding the necessary layers for training and validation, the models were trained and performance evaluation metrics were recorded. The training process was graphed and classification reports were created.

3. Results

In this study for the automatic diagnosis of spinal deformities, the classification performances of GoogleNet and ResNet architectures trained with the deep transfer learning approach were compared in terms of accuracy, sensitivity, sensitivity, specificity, F1 score and the result graphs were drawn. When the classification performances of the architectures are compared in terms of accuracy, the ResNet architecture stands out with an accuracy of 0.92 and surpasses GoogleNet in terms of sensitivity with 0.91. According to Figure 3.1, ResNet surpassed GoogleNet in terms of specificity, F1 score and precision values and showed a superior classification performance. Based on these results, it is possible to say that ResNet is superior in terms of classification metrics.

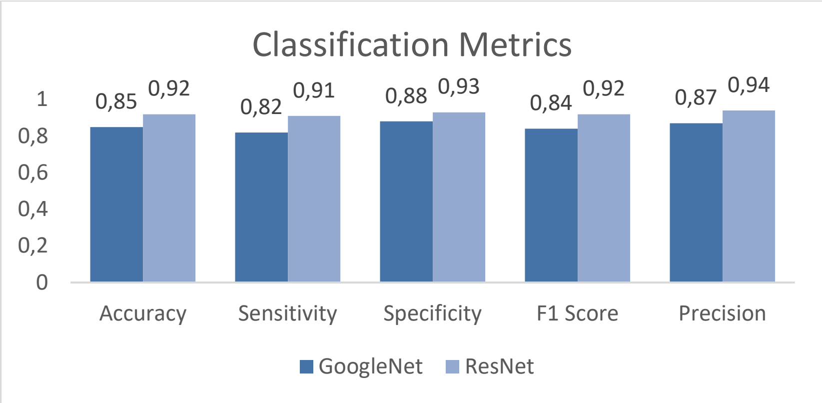

Figure 3.1. Classification Metrics for CNN Architectures

When compared in terms of training and testing times, it was seen that ResNet was behind GoogleNet's 30-second training time with 45 seconds, and GoogleNet gave faster results than ResNet with 0.5 seconds to a 0.6-second test time. According to these results, it is possible to say that GoogleNet is more successful in terms of training and testing times.

4. Conclusion

Artificial intelligence-based techniques are increasing their usage areas day by day and providing successful results. Artificial intelligence, which is expanding its use and application areas in the healthcare sector, has started to take an active role in diagnosis and treatment processes and offers the advantages of optimizing these processes.

In this study, when the classification performances of GoogleNet and ResNet, which are popular convolutional neural networks architectures, were compared in terms of accuracy, sensitivity, precision, F1 score and specificity values, it was concluded that ResNet showed superior performance than GoogleNet. On the other hand, when a comparison was made in terms of training and testing times, it was seen that GoogleNet processed faster. In light of these results, the performance of deep learning-based CNN architectures in classifying X-ray images was found to be successful and it became possible to say that they can be used in automatic diagnosis systems for spinal deformities.

In order for deep learning-based automatic diagnosis systems to operate with higher performance, the scope of this study can be expanded by using larger data sets, images of different diseases and different CNN architectures. Architectures supported by different image types and different parameters can have higher processing capabilities by increasing the number of cycles. In line with the optimized data, higher performance CNN architectures can be designed and higher accuracy automatic diagnosis systems can be developed.

References

[1] R. Izzo, G. Guarnieri, G. Guglielmi, and M. Muto, “Biomechanics of the spine. Part I: Spinal stability,” *European Journal of Radiology*, vol. 82, no. 1. pp. 118–126, Jan. 2013. doi: 10.1016/j.ejrad.2012.07.024.

[2] J. Iorio *et al.*, “The Effect of Aging on Cervical Parameters in a Normative North American Population,” *Global Spine J*, vol. 8, no. 7, pp. 709–715, Oct. 2018, doi: 10.1177/2192568218765400.

[3] C. R. , Good, J. D. Auerbach, P. T. O’Leary, and T. C. Schuler, “Adult spine deformity,” *Current Reviews in Musculoskeletal Medicine*, vol. 4, pp. 159–167, 2011.

[4] P. Guigui and E. Ferrero, “Surgical treatment of degenerative spondylolisthesis,” *Orthopaedics and Traumatology: Surgery and Research*, vol. 103, no. 1. Elsevier Masson SAS, pp. S11–S20, Feb. 01, 2017. doi: 10.1016/j.otsr.2016.06.022.

[5] S. , Haleem and C. , Nnadi, “Scoliosis: a review,” 2018.
[6] E. Moustafa and B. S. Tur, “Spondylolisthesis: Etiology, Diagnosis, Clinical Features and Treatment,” *Journal of Ankara University Faculty of Medicine*, vol. 71, no. 3, pp. 118–126, Dec. 2018, doi: 10.4274/atfm.36844.
[7] M. Fraiwan, Z. Audat, L. Fraiwan, and T. Manasreh, “Using deep transfer learning to detect scoliosis and spondylolisthesis from x-ray images,” *PLoS One*, vol. 17, no. 5 May, May 2022, doi: 10.1371/journal.pone.0267851.
[8] G. M. Trinh *et al.*, “Detection of Lumbar Spondylolisthesis from X-ray Images Using Deep Learning Network,” *J Clin Med*, vol. 11, no. 18, Sep. 2022, doi: 10.3390/jcm11185450.
[9] J. Zhang *et al.*, “Deep learning system assisted detection and localization of lumbar spondylolisthesis,” *Front Bioeng Biotechnol*, vol. 11, 2023, doi: 10.3389/fbioe.2023.1194009.
[10] F. Varçın, Hasan Erbay, Eyüp Çetin, İhsan Çetin, and Turgut Kültür, “Diagnosis of Lumbar Spondylolisthesis via Convolutional Neural Networks Ihsan Çetin,” 2019.
[11] Roa H. Alharbi, Meshal B. Alshaye, Maryam M. Alkanhal, and Najla M. Alharbi, “Deep Learning Based Algorithm For Automatic Scoliosis Angle Measurement,” *National Center for Communication Systems and Command and Control Technology*, 2020.
[12] Y. Tu, N. Wang, F. Tong, and H. Chen, “Automatic measurement algorithm of scoliosis Cobb angle based on deep learning,” in *Journal of Physics: Conference Series*, Institute of Physics Publishing, May 2019. doi: 10.1088/1742-6596/1187/4/042100.
[13] Y. Lecun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553. Nature Publishing Group, pp. 436–444, May 27, 2015. doi: 10.1038/nature14539.
[14] F. Doğan and İ. Türkoğlu, “Derin Öğrenme Modelleri ve Uygulama Alanlarına İlişkin Bir Derleme,” *DÜMF Mühendislik Dergisi*, vol. 10, no. 2, pp. 409–445, Jun. 2019, doi: 10.24012/dumf.411130.
[15] Pranshu Sharma, “Basic Introduction to Convolutional Neural Network in Deep Learning,” Analytics Vidhya. Accessed: May 17, 2024. [Online]. Available: https://www.analyticsvidhya.com/blog/2022/03/basic-introduction-to-convolutional-neural-network-in-deep-learning/
[16] Ö. İnik *et al.*, “Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri,” *GAZİOSMANPAŞA BİLİMSEL ARAŞTIRMA DERGİSİ*, vol. 6, no. 3, pp. 85–104, 2017, [Online]. Available: http://dergipark.gov.tr/gbad
[17] P. Zhao, S. C. H. Hoi, J. Wang, and B. Li, “Online transfer learning,” *Artif Intell*, vol. 216, pp. 76–102, 2014, doi: 10.1016/j.artint.2014.06.003.
[18] C. Szegedy *et al.*, “Going Deeper with Convolutions,” Sep. 2014, [Online]. Available: http://arxiv.org/abs/1409.4842
[19] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Dec. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.