Comparison of Deep Learning Models and Optimization Algorithms in the Detection of Scoliosis and Spondylolisthesis from X-Ray Images

Harun Güneş[1*] , Cengiz Hark[2] , Abdullah Erhan Akkaya[2]

[1] Hakkari University, Vocational School of Health Services, Medical Services and Techniques, Hakkari, Türkiye, harungunes@hakkari.edu.tr
[2] Inonu University, Faculty of Engineering, Department of Computer Engineering, Malatya, Türkiye, cengiz.hark@inonu.edu.tr, erhan.akkaya@inonu.edu.tr
*Corresponding Author

ARTICLE INFO

ABSTRACT

The spine is composed of pieces of bone called vertebrae that lie between the skull and the tailbone. Various medical conditions can affect the spine. In this study, two types of degenerative diseases, scoliosis, and spondylolisthesis, were studied. Deep AI architectures have recently enabled further disease diagnosis innovation using medical images. Various traditional and deep learning studies use medical images for disease diagnosis in the literature. This study aims to classify spine X-ray images according to three possible conditions (Normal, Scoliosis, and Spondylolisthesis) and to exploit the potential of these X-ray images to detect possible diseases occurring in the spine. The performance of deep learning models and optimization algorithms used in this process was evaluated. The study uses a data set created and/or analyzed during an existing study. This data set consists of images that belong to three different classes: scoliosis, spondylolisthesis, or x-ray images of normal (i.e. healthy) individuals. A total of 338 spine X-ray images, 188 scoliosis images, 79 spondylolisthesis images, and 71 normal images. Six different deep-learning architectures have been used in the study. These architectures are Alexnet, GoogLeNet, ResNet-18, ResNet-50, ResNet-101, and EfficientNet-bo. While working on these deep architectures, each model has been evaluated using different optimization algorithms. These optimization algorithms are RmsProp, SGDM, and Adam. According to the classification processes, the deep learning model with the highest accuracy value was Alexnet, and the optimization algorithm used with it, Sgdm (99.01%), and the training time lasted 38 seconds. According to the classification processes, the deep learning model with the fastest completion time (30 seconds) was Alexnet and the optimization algorithm used with it was RmsProp. An accuracy rate of 98.02% has been obtained in the training of this model.

1. Introduction

The spine is composed of pieces of bone called vertebrae that lie between the skull and the tailbone. Intervertebral discs and facet joints are arranged symmetrically between the vertebrae. While these joints increase the spine's mobility, they also increase the strength and flexibility of the spine. The tissue in front of each vertebra is called the intervertebral disc, and the tissue behind it is called the facet joint [1]. Among these structures, soft tissue extending from the skull to the tailbone surrounds and protects the spinal cord, which is an important part of the central nervous system. Like the brain, the spinal wire is wrapped through membranes known as the meninges. The cerebrospinal fluid is between this membranous structure representing the cerebral cortex's and spinal cord's continuity. The spine of the human skeleton consists of 7 cervical vertebrae, 12 thoracic or dorsal vertebrae, 5 lumbar vertebrae, pseudolumbar vertebrae, and 5 sacral vertebrae. The spine is also supported by

connective and muscle tissue, known medically as ligaments [2].

Various medical conditions can affect the spine. In this study, two different spinal diseases, scoliosis, and spondylolisthesis, were examined. Scoliosis is the most common three-dimensional spinal deformity among spinal deformities and causes severe postural dysfunction in advanced stages. The diagnosis of scoliosis, a degenerative disease, can be made by an expert using magnetic resonance imaging (MRI) or an X-ray of the spine. The rate of scoliosis varies from country to country in the world and it is considered to be one of the most important spinal diseases with a prevalence of 0.47-5.2% [3].

A Cobb angle value greater than 10° on spine images indicates scoliosis [4]. According to estimates and studies, it is stated that 6-9 million people in the United States have some degree of scoliosis [5]. Spondylolisthesis is caused by a damaged vertebral carrier or anterior displacement of the underlying vertebra. It is generally classified in various degrees (low-grade, high-grade, etc.) according to the degree of slip [6]. Spondylosis disease can be seen in 6% of adults. It can lead to weakness and numbness in the legs of individuals with this disease [6]. Spondylolisthesis is the change of the upper segment of the spine relative to the lower segment. It is a deformation in which it moves forward [7]. Spondylolisthesis can be seen as a focal abnormality in the sagittal plane, in which the spine slides from the adjacent anterior or posterior plane [8].

Vertebral deformities are often accompanied by cosmetic concerns or concerns about the progression of the deformity. The most common are; adults experiencing symptoms, pain, and limitation in daily life. Therefore, it is very important to document the type, severity, location, and change of patient pain over time [9]. In addition, it is necessary to clarify which factors lead to increases and decreases in axial low back pain as well as leg pain and their effects [10].

In recent years, they have contributed to the literature in the field of deep learning in disease diagnosis using medical images. Fraiwan et al. (2022) examined the utilization of deep transfer learning techniques for diagnosing scoliosis and spondylolisthesis from X-ray images. Their research focuses on the application of 14 deep transfer learning algorithms to automate the diagnosis of scoliosis and spondylolisthesis. Deep learning networks are employed to perform classification and diagnostic operations on the X-ray image data.

The authors' study investigates the effectiveness of deep transfer learning techniques in identifying scoliosis and spondylolisthesis from X-ray images. The most favorable outcome was achieved by DensNet-201, which exhibited a success rate of 99.01%. The article's findings provide information regarding the diagnostic efficacy and accuracy of these procedures[11]. Rao Farhat Masood et al. (2022), in their study, focus on deep learning-based spinal body segmentation, extraction of spine measurements, and disease classification. In this article, the segmentation of spinal objects in spine images using deep learning methods is discussed. In addition, the usability of deep learning-based approaches for spine measurements and disease classification is also examined. This study investigates how automated segmentation and classification methods can be potentially helpful tool in the diagnosis of spinal diseases. Using the angular deviation metric for spondylolisthesis classification, they achieved 89% accuracy, while calculating the area within the closed lumbar curve region, they achieved 93% accuracy in determining adequacy/inadequacy in the LL assessment [12].

Mahsa Tajdari et al. (2021) focus on image-based modeling for Idiopathic Scoliosis in adolescents. The article focuses on understanding and predicting Idiopathic Scoliosis using image-based modeling methods, focusing on mechanistic machine learning analysis and prediction. This study explores the usability of deep learning and machine learning techniques to understand the mechanistic features and progression of Idiopathic Scoliosis [13].

The aim of this study is to classify spine X-ray images according to three possible conditions (Normal, Scoliosis, and Spondylolisthesis) and to exploit the potential of these X-ray images to

detect possible diseases occurring in the spine. Performance evaluation and comparison of 6 different deep learning architectures, which can be classified into three categories, on the same dataset were made using different optimization algorithms. The neural networks used in this study were used to investigate the optimal estimation method. Other parts of the study continue as follows. Chapter 2 discusses materials and methods. This section explains data sets, deep learning architectures, and optimization algorithms used in experimental studies. Chapter 3 presents the experimental studies and results.

2. Material and Methods

In this study; data set created and/or analyzed during an existing study was used [14]. This dataset was created from three different classes of images: chest X-ray images of scoliosis, spondylolisthesis, or normal (i.e. healthy) people. There are 338 x-ray images in total, with 188 scoliosis images, 79 spondylolisthesis images, and 71 normal images. In this study, 6 different deep-learning architectures were studied. These architectures; Alexnet, GoogLeNet, ResNet-18, ResNet-50, ResNet-101 and EfficientNet-bo. While working on these deep learning architectures, each model was evaluated using different optimization algorithms. These optimization algorithms are RmsProp, SGDM, and Adam.

2.1.Dataset

In this study, a data set created and/or analyzed during an existing study was used [13]. The dataset included 338 subjects (240 females, 98 males) aged 9 months to 79 years with a mean ± SD of 24.9 ± 18.58 years. There were 71 radiographically normal subjects (40 females, 31 males) aged 9 months to 56 years, with a mean ± SD of 19.41 ± 11.19. 79 people (49 females, 30 males) aged between 15 and 79 were diagnosed with spondylolisthesis, with a mean ± SD of 53.59 ± 14.02. The number of people aged 5-35 years diagnosed with scoliosis was 188 (151 females, 37 males), and the mean ± SD was 14.73 ± 3.36.

This dataset is composed of images belonging to three different classes: scoliosis, spondylolisthesis, or vertebral X-ray images of normal (i.e. healthy) people. There are 338 X-ray images in total, with 188 scoliosis images, 79 spondylolisthesis images, and 71 normal images. Figure 1 shows image examples of scoliosis, spondylolisthesis, and normal (i.e. healthy) individuals from this dataset [14].
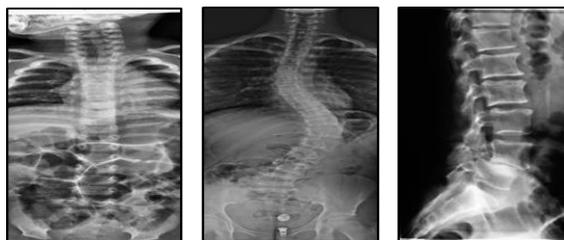
Figure 1. Sample normal, scoliosis, and spondylolisthesis x-ray images in the dataset

2.1. Convolutional Neural Network (CNN)

Deep learning, a hot topic of late, is a machine learning approach that has emerged with the deepening of multilayer feedforward neural networks. Due to the limited number of hardware products, the number of layers in traditional neural networks is limited by learned parameters, and the relationships between layers are computationally intensive. High-end computer generation makes it possible to train deep architectures using multilevel neural networks [15].

It is a high-performing convolutional neural network method in many areas such as deep learning, image processing, machine learning, speech recognition, and object tracking. A CNN (Convolutional Neural Network) is a type of multilayer neural network. One of the main advantages of CNN algorithms is feature extraction, which minimizes preprocessing steps. Therefore, no preliminary search is required to find features in the image [16].

Currently, many deep learning architectures have been developed and used in various research. Some of these architectures are LeNet, AlexNet, ZFNet, GoogLeNet, VGGNet, LSTM, RNN, SequeezeNet, ResNet, and EfficientNet. The mentioned architectures are just a few of the deep learning architectures. In this study, AlexNet, GoogLeNet, ResNet-18, ResNet-50, ResNet-

101, and EfficientNet-b0 architectures were used.

2.2.1. Alexnet

It is a deep neural network architecture developed by Krizhevsky, Sutskever, and Hinton. It won the 2012 ImageNet competition, making deep learning a global voice. There are sequential convolution and link layers. This architecture increased the performance of computer-aided object identification from 10.8% to 83.6% [17].

Alexnet; is a neural network with 60 million parameters and 650 000 neurons. Most consist of 5 layers of convolution followed by a max pooling layer and 3 fully connected layers.

The ImageNet dataset contains 1000 different image classes. Therefore, the output layer consists of 1000 units. The resulting model will be a deeper and larger model, but the architecture is very similar to LeNet. In the AlexNet diagram in Figure 2, we can see that he splits the problem into two parts. Half are running on GPU1 and half on GPU2 [18]. This way, it keeps the communication load low, resulting in good overall performance. The data processing of the two channels intersects only at the third feature extraction layer. ReLU as the activation function, a dilution method is applied to prevent overfitting [18].
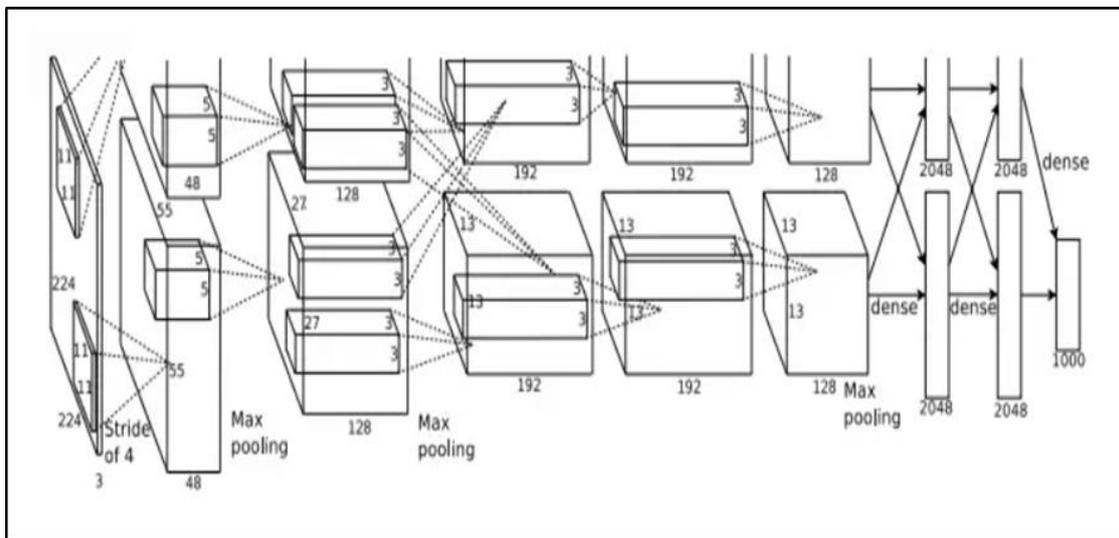
Figure 2. Alexnet

2.2.2. Googlenet

GoogLeNet (Szegedy, Liu et al. 2015) is a complex architecture due to its output modules. In 2014, GoogLeNet beat ImageNet's competition with 22 layers and 5.7% error rate [19]. This architecture is usually one of the first CNN architectures to move away from stacking convolutional and pooling layers in a sequential structure. Storage and power consumption also play an important role in this new model. Stacking all the layers and adding many filters increases computation and storage costs and increases the potential for memorization. GoogLeNet overcomes this by using modules connected in parallel. The GoogLeNet network architecture is shown in figure 3 [19].
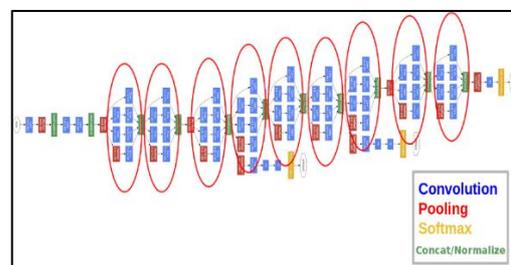
Figure 3. GoogLeNet

2.2.3. ResNet

In 2015, ResNet ranked first and the error rate detected by GoogLeNet decreased from 6.67% the previous year to 3.57%. In this deep network model, a different approach is applied by adding a new structure called a residual block.

The structure of the residual layer is shown in Figure 4 [19]. As you can see, the block output is equal to (F(x) + x), where x is the block input. Here F(x) represents the weight layer output for x input data.
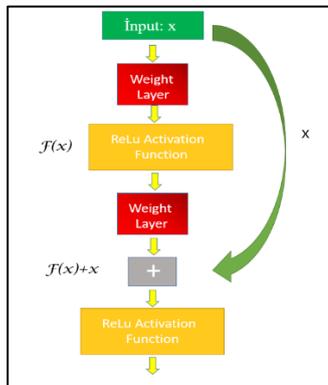
Figure 4. Residual block structure in ResNet architecture

a. ResNet-18

A residual block structure is introduced to solve the vanishing gradient problem. The disappearing gradient problem causes the error derivative to shrink and disappear during training. For this reason, the updating of the weights is interrupted and the training is terminated. Theoretically, the training error should decrease as the number of layers in the network increases. However, in practice, gradient flux decreases and training error increases as layers are added to the network. With the deepening network, thanks to residual blocks in ResNet, a way to reduce the training error has been found [19].

The ResNet model has three different versions, ResNet-18, ResNet-50, and ResNet101, depending on the number of deep layers involved.

b. ResNet-50

Resnet 50 is obtained by replacing each 2-layer block in a 34-layer network with a 3-layer bottleneck block. For each remainder function F, a three-layer stack is used instead of two. The dimensions of these three layers are 1x1, 3x3, and 1x1. Here the 1×1 layer causes the size to decrease and then increase (return). On the other

hand, 3×3 layers still have the bottleneck of small input/output sizes. This structure is shown in figure 5 [19].
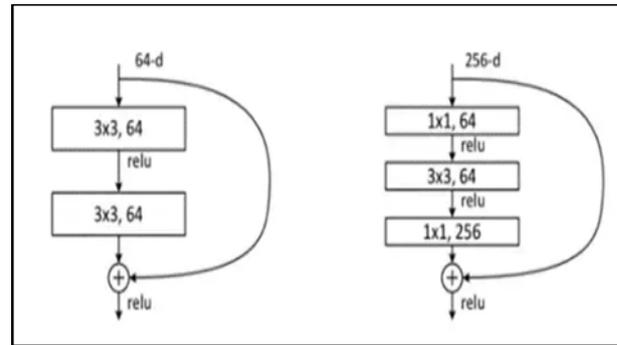
Figure 5. Left: A building block for ResNet-34 Right: A "bottleneck" building block for ResNet-50/101/152

c. ResNet-101-152

Other 3-layer blocks are used to create 101 and 152-layer ResNets. A 50/101/152 layer ResNet is much more accurate than a 34-layer ResNet. The problem of distortion is avoided thanks to the residual mesh, and the increased depth greatly improves the accuracy. Figure 6 shows the classification errors of some imagenet models [19].
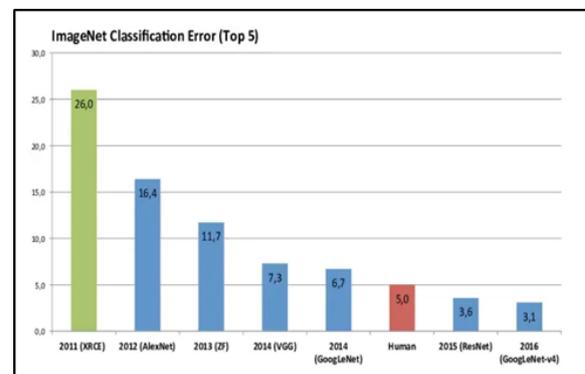
Figure 6. Imagenet classification error (Top 5)

As a result, ResNet uses shortcut links to skip some layers, thus improving model performance in deeper networks and overcoming optimization/degradation issues in deeper networks.

The building blocks in the Resnet model are shown in brackets in Figure 7 with their stacked block numbers [19].

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3\times3,\ 64 \\ 3\times3,\ 64 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3,\ 64 \\ 3\times3,\ 64 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1,\ 64 \\ 3\times3,\ 64 \\ 1\times1,\ 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1,\ 64 \\ 3\times3,\ 64 \\ 1\times1,\ 256 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1,\ 64 \\ 3\times3,\ 64 \\ 1\times1,\ 256 \end{bmatrix}\times3$
conv3_x	28×28	$\begin{bmatrix} 3\times3,\ 128 \\ 3\times3,\ 128 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3,\ 128 \\ 3\times3,\ 128 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1,\ 128 \\ 3\times3,\ 128 \\ 1\times1,\ 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1,\ 128 \\ 3\times3,\ 128 \\ 1\times1,\ 512 \end{bmatrix}\times4$	$\begin{bmatrix} 1\times1,\ 128 \\ 3\times3,\ 128 \\ 1\times1,\ 512 \end{bmatrix}\times8$
conv4_x	14×14	$\begin{bmatrix} 3\times3,\ 256 \\ 3\times3,\ 256 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3,\ 256 \\ 3\times3,\ 256 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1,\ 256 \\ 3\times3,\ 256 \\ 1\times1,\ 1024 \end{bmatrix}\times6$	$\begin{bmatrix} 1\times1,\ 256 \\ 3\times3,\ 256 \\ 1\times1,\ 1024 \end{bmatrix}\times23$	$\begin{bmatrix} 1\times1,\ 256 \\ 3\times3,\ 256 \\ 1\times1,\ 1024 \end{bmatrix}\times36$
conv5_x	7×7	$\begin{bmatrix} 3\times3,\ 512 \\ 3\times3,\ 512 \end{bmatrix}\times2$	$\begin{bmatrix} 3\times3,\ 512 \\ 3\times3,\ 512 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1,\ 512 \\ 3\times3,\ 512 \\ 1\times1,\ 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1,\ 512 \\ 3\times3,\ 512 \\ 1\times1,\ 2048 \end{bmatrix}\times3$	$\begin{bmatrix} 1\times1,\ 512 \\ 3\times3,\ 512 \\ 1\times1,\ 2048 \end{bmatrix}\times3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^{9}	3.6×10^{9}	3.8×10^{9}	7.6×10^{9}	11.3×10^{9}

Figure 7. Building blocks are shown in brackets with their stacked block numbers. Down-sampling is performed by conv3, conv4 1, and conv5 in steps of 2

2.2.4. EfficientNet-b0

EfficientNet is a CNN architecture that achieves high accuracy with fewer data compared to other convolutional neural networks. The EfficientNet architecture is based on the principle of scaling several dimensions of the network simultaneously, such as depth, width, and image resolution, using fixed complex factors [20].

In fact, this composite scaling idea also works with existing MobileNet and ResNet architectures. It should be optimized for accuracy and efficiency measured in floating point operations per second (FLOPS). This advanced architecture uses MBConv (Mobile Inverted Bottleneck Convolution). Unlike other cutting-edge models, EfficientNet achieves more efficient results by scaling depth, width, and resolution equally while downsizing the model.

The EfficientNet family consists of eight models, from EfficientNet-B0 to EfficientNet-B7. As the model number increases, the number of parameters and mesh depth increases. Therefore, higher performance is achieved. A block diagram of the basic network structure of the EfficientNet B0 model is shown in figure 8 [21].

Figure 8. EfficientNet B0 basic network structure block representation

2.2. Optimization algorithms

There are six commonly used optimization methods to minimize machine learning error rates. These methods are; SGD, Momentum, Adagrad, RMSProp, Adadelta and Adam methods. In Table 1 [22], the characteristics of the gradient descent methods are compared. The evolution map of gradient descent is shown in Figure 9 [22].

Table 1. Comparison of Gradient Descent Methods

Algorithm	Year	Learning Coefficient	Gradient
SGD	1951	√	√
Momentum	1964		√
Adam	2014	√	√
AdaGrad	2011	√	
RMSProp	2012	√	
Adadelta	2012	√	

Figure 9. Evolutionary map of gradient descent methods

figure 10 b the SGD with momentum reaches a solution with fewer oscillations [23].

(a)   (b)
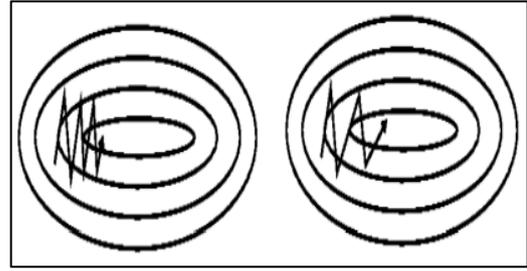
Figure 10. a.) SGD without momentum b.) SGD with momentum

2.3.1.SGDM

Many studies in the literature use SGD as stochastic gradient descent. SGD randomly updates weights using some but not all gradients. It is necessary to update the current weights (wt) by multiplying the current gradient ($\partial L / \partial wt$) by the learning factor (a) [22].

$$w_{t+1} = w_t - a \frac{\partial L}{\partial w_t} \qquad (1)$$

There are many variations when looking for the best spot in SGD. The Momentum method is recommended to reduce these vibrations and increase the speed of reaching your goals [22]. This method uses pulsed gradients instead of existing gradients. In fact, the name Monemtum is somewhat disappointing and the method can be described as 'controlled beats' [23].

$$w_{t+1} = w_t - a v_t \qquad (2)$$

$$v_t = \beta v_{t-1} + (1-\beta) \frac{\partial L}{\partial w_t} \qquad (3)$$

Here the initial value of Vt is 0. β ranges from 0 to 1, and the commonly used value of 0.9 is used to set how much of the historical gradient is included in the process. In figure 10 we see that the momentumless SGD oscillates strongly on the way to the solution. On the other hand, in

$$v_t = \beta_1 v_{t-1} + (1-\beta_1) \frac{\partial L}{\partial w_t} \qquad (9)$$

$$S_t = \beta_2 S_{t-1} + (1-\beta_2) \left[\frac{\partial L}{\partial w_t} \right]^2 \qquad (10)$$

2.3.2.RMSprop

It is proposed to solve the continuous learning coefficient problem like Adagrad. The difference is that the Adagrad method does not square the slope, it squares the slope by momentum [24].

$$w_{t+1} = w_t - \frac{a}{\sqrt{S_t + \varepsilon}} \cdot \frac{\partial L}{\partial w_t} \qquad (4)$$

$$S_t = \beta S_{t-1} + (1-\beta) \left[\frac{\partial L}{\partial w_t} \right]^2 \qquad (5)$$

Here, S is initially prioritized as 0, a=0.001, β=0.9, ϵ=10^{-6} [23].

2.3.3.Adam

Gradient descent is proposed by combining the advantages of RMSprop and impulse methods. V is used for the pulse method and S is used for rmsprop [24].

$$w_{t-1} = w_t - \frac{a}{\sqrt{\hat{S} + \varepsilon}} \cdot \hat{v}_t \qquad (6)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_1^t} \qquad (7)$$

$$\hat{S}_t = \frac{S_t}{1 - \beta_2^t} \qquad (8)$$

Here, firstly S and V are initially 0; a=0.001; β1=0.9; β2=0.999; ϵ is preferred as 10^{-8} [24].

2.3.Method used

The aim of this study is to take X-Ray images of the spine; scoliosis, spondylolisthesis, and normal (i.e. healthy) to classify according to three possible conditions and to detect the disease. While doing this process, 6 different deep-learning architectures were used. These architectures are: AlexNet, GoogleNet, ResNet-18, ResNet50, ResNet100, and EfficientNet-b0. During the training process, 3 different optimization algorithms were used in each model. These optimization algorithms are Sgdm, Adam, and RmsProp. After performing the classification process, the success rate and training time of the models and algorithms were compared. Performance evaluation and comparison of 6 different deep learning techniques, which can be classified into three categories, on the same dataset were made using different optimization algorithms.

The best estimation method was investigated with the neural networks used in this study (Figure 11).
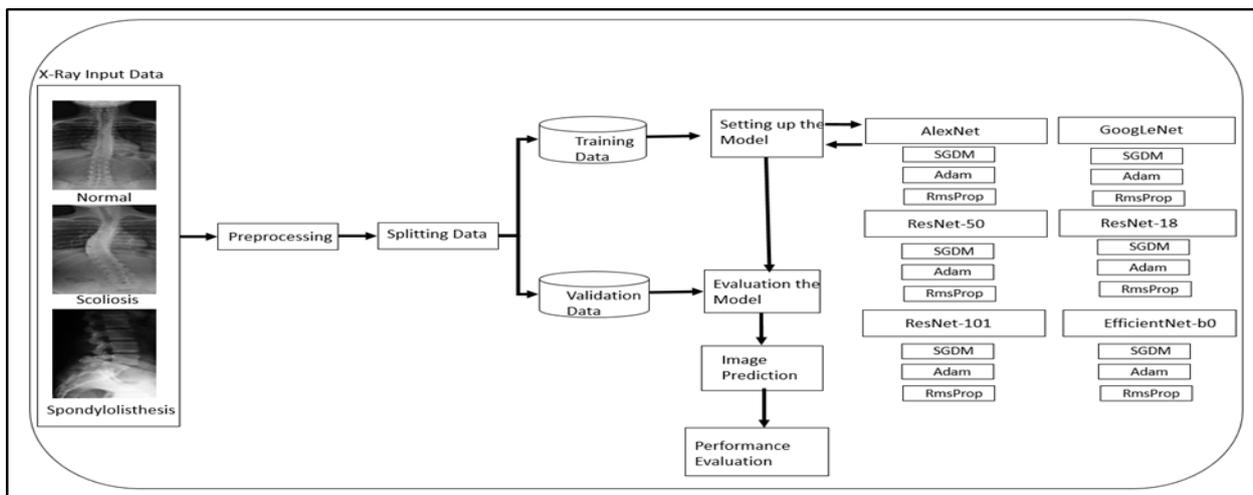
Figure 11. Architecture used to compare different models in x-ray image classification with deep learning technique

3. Results and Discussion

In this study, the classification process and disease detection were tried to be done by using deep-learning neural networks from X-Ray vertebra images. It is trained with previously trained deep neural networks and 80% of the X-Ray dataset is taken randomly. The system was then validated with 20% of the data taken randomly. While this process is being done, the classification process in neural network models is done separately with optimization algorithms (Sgdm, Adam, RmsProp). The application was prepared in a Matlab environment [25]. The work done; It is built on Intel Core-i7 6800K 3.4GHz processor, GPU Nvidia GeForce RTX 3060 Ti graphics card, 16 GB RAM, and 64-bit Windows 10 hardware. After the classification process was performed, the best accuracy rate was obtained with the Alexnet model and Sgdm optimization algorithm with a rate of 99.01%. When comparing the training times, the model that completed the training time the fastest was the Alexnet and RmsProp optimization algorithm with 30 seconds. In Table 2, the sample numbers and rates in the classification process for all models are presented.

Table 2. X-Ray dataset sample numbers, training and validation rate

Dataset	% Rate	Number of samples
Total	%100	338
Train	%80	270
Validation	%20	68

In table 3, the results obtained as a result of the classification process and the training periods are presented. According to the classification processes, the deep learning model that gave the highest accuracy value was Alexnet and the optimization algorithm used with it, Sgdm (99.01%), and the training time lasted 38 seconds. Performance graphics of this process are presented in figure 12 and figure 13.

Table 3. Deep-learning models used, Optimization algorithms, accuracy values , and training time

Deep Learning Model	Optimization Algorithm	Accuracy (%)	Model Training Time (Second)
Alexnet	**Sgdm**	**99.01**	38
Alexnet	Adam	96.04	40
Alexnet	**RmsProp**	98.02	**30**
GoogLeNet	Sgdm	94.06	67
GoogLeNet	Adam	98.02	94
GoogLeNet	RmsProp	97.03	76
ResNet-18	Sgdm	97.06	67
ResNet-18	Adam	92.65	42
ResNet-18	RmsProp	97.06	40
ResNet-50	Sgdm	97.03	97
ResNet-50	Adam	97.06	96
ResNet-50	RmsProp	98.53	131
ResNet-101	Sgdm	98.02	221
ResNet-101	Adam	98.02	286
ResNet-101	RmsProp	98.02	210
EfficientNet-b0	Sgdm	98.02	607
EfficientNet-b0	Adam	99.00	647
EfficientNet-b0	RmsProp	94.06	625
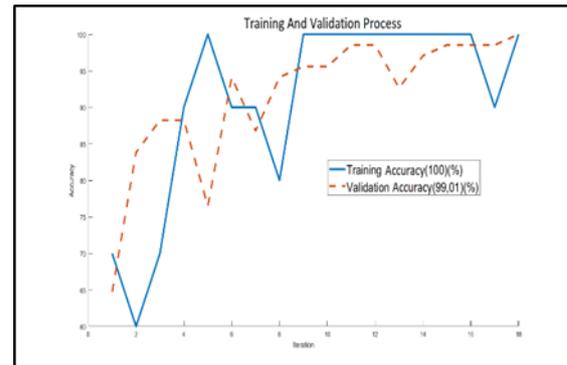
Figure 12. Alexnet and sgdm performance graph (Accuracy)
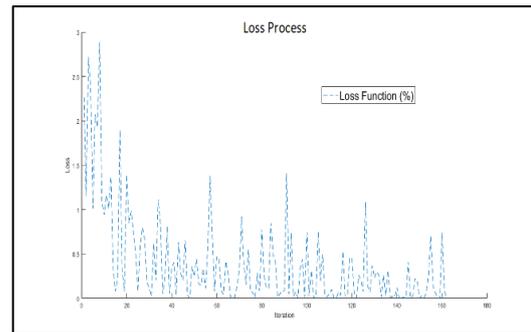
Figure 13. Alexnet and sgdm performance graph (Loss)

The complexity matrices of the training and validation data obtained after the completion of the training of the model are given in figure 14 and figure 15.
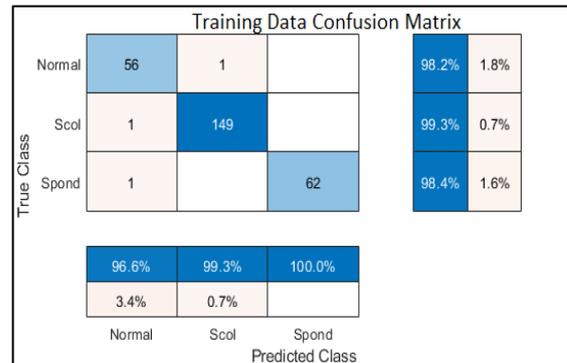
Figure 14. Alexnet and Sgdm; Training Data Confusion Matrix

446
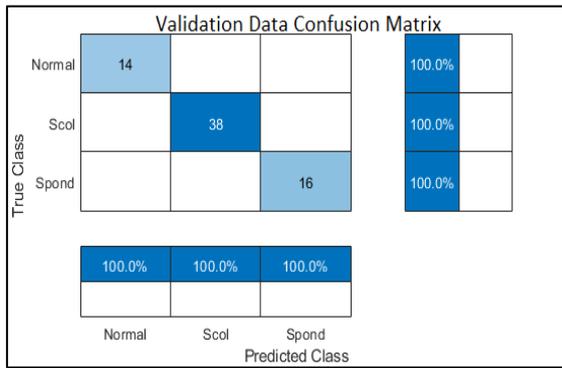
Figure 15. Alexnet and Sgdm; Validation Data Confusion Matrix

After the training of the model is completed, the ratios of determining the classes to which the X-Ray images randomly sent to the classifier belong are given in Figure 16.

As shown in Figure 16a, an X-ray image sent randomly to the classifier is evaluated and it is determined that the image belongs to the normal class at the rate of 95% by the classifier. By following the same path, it was determined that the randomly selected image in figure 16b belongs to the scoliosis class at the rate of 100%, and the image in figure 16c is 100% of the spondylolisthesis class.
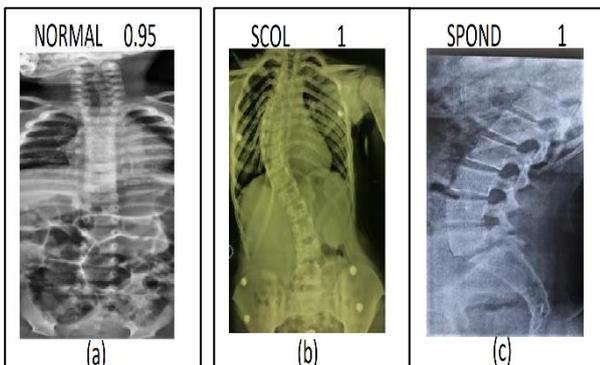
Figure 16. Alexnet and sgdm; x-ray image detection

According to the classification processes, the deep learning model with the fastest completion of the training period was Alexnet, and the optimization algorithm used with it, RmsProp. An accuracy rate of 98.02% was obtained in the training of this model. Performance graphics of this process are presented in figure 17 and figure 18.

Figure 17. Alexnet and RmsProp performance graph (Accuracy)
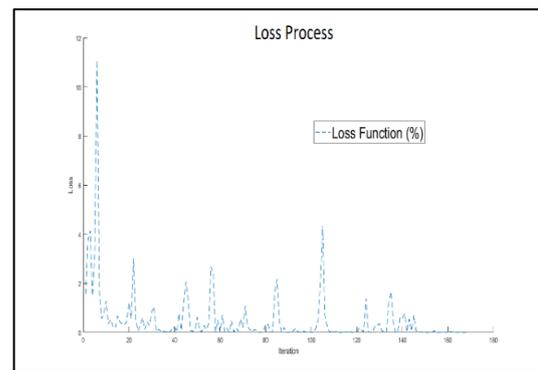
Figure 18. Alexnet and rmsprop performance graph (Loss)

The complexity matrices of the training and validation data obtained after the completion of the training of the model are given in figure 19 and figure 20.
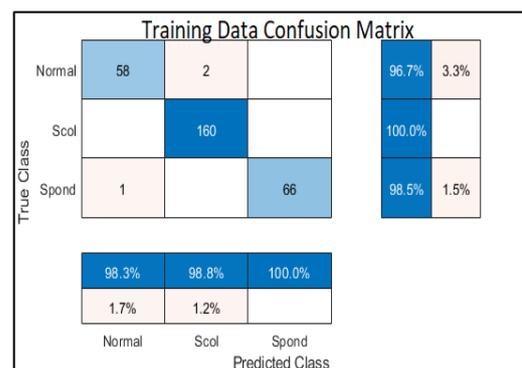
Figure 19. Alexnet and rmsrrop; training data confusion matrix
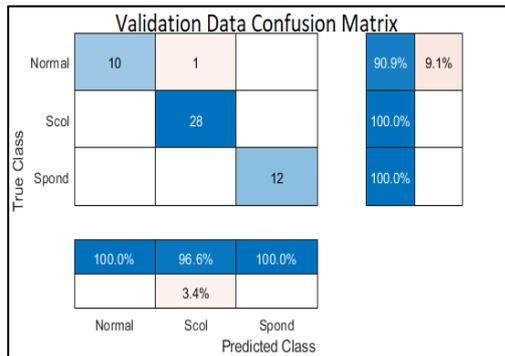
Figure 20. Alexnet and rmsprop; validation data confusion matrix
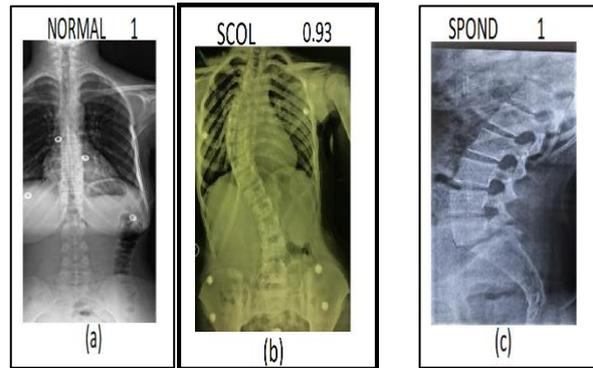
Figure 21. Alexnet and rmsprop; x-ray image detection

After the training of the model is completed, the ratios of determining the classes to which the X-Ray images randomly sent to the classifier belong are given in Figure 21.

As shown in Figure 21a, an X-ray image sent randomly to the classifier is evaluated and it is determined that the image belongs to the normal class at 100% by the classifier. Following the same path, it was determined that the randomly selected image in figure 21b belongs to the scoliosis class at a rate of 93% and the image in figure 21c belongs to the spondylolisthesis class at a rate of 100%.

Table 4 illustrates a performance comparison between relevant studies. Although previous studies [1, 26, 29] achieved notable levels of accuracy, their methodologies necessitate extensive and error-prone measurements of biomechanical parameters, which may not be essential for the specific detection of disease cases. To the best of our knowledge, no other investigation has employed deep learning techniques for the classification of scoliosis and spondylolisthesis using normal X-ray images. Kolombo et al. [27] focused on discriminating scoliosis from a healthy condition and attained a maximum accuracy of approximately 85%. Likewise, Wang et al. [28] did not achieve satisfactory accuracy in detecting scoliosis progression, while Yang et al. obtained an average accuracy of around 80% in differentiating scoliosis severity based on Cobb angles ($< 10°$, $10°$-$19°$, $20°$-$44°$, or $\geq 45°$).

Table 4. Comparison with related studies in the literature.

Study	Classification Problem	Dataset	Accuracy (%)
Alafeef et al. [1]	3-class classification	422 subjects	99.5
Fraiwan et al. [3]	3-class pair-wise classification	331 subjects	96.34-99.33
Yang et al. [4]	4-class for scoliosis severity	3640 back images	80
Reshi et al. [26]	3-class classification	310 records	99.5
Colombo et al. [27]	Healthy vs scoliosis	272 scoliosis and 20 healthy	85
Wang et al. [28]	Progressing vs non-progressive scoliosis	490 subjects	76
Unal et al. [29]	Pairwise	310 records	96
This Work	3-class pair-wise classification	331 subjects	92.65-99.01

However, despite the absence of directly comparable literature, our current study demonstrates superior accuracy with a reduced number of input processing and measurements. Unal et al. [29] investigated a pairwise Fuzzy C-Means based feature weighting method to improve the classification of spinal diseases. This study, which stands out to overcome the limitations of traditional methods, was able to achieve more accurate classification results while demonstrating the effective use of artificial intelligence and data mining in disease diagnosis. This article focuses on the comparison of deep learning models and optimization algorithms in detecting spinal diseases such as scoliosis and spondylolisthesis from X-ray images, and it contributes to the literature in various ways. Here are the potential contributions of this article to the literature:

1. Effectiveness of deep learning models: The article investigates the use of six different deep learning architectures in detecting spinal diseases such as scoliosis and spondylolisthesis. This is an important contribution that evaluates the effectiveness of deep learning models in accurately diagnosing these diseases.

2. Comparison of optimization algorithms: The article examines the impact of different optimization algorithms on the performance of deep learning models. This can provide guidance to researchers on which optimization algorithms yield better results in this type of disease detection.

3. Dataset creation: The article discusses the creation and utilization of a dataset containing spinal images of scoliosis, spondylolisthesis, and normal individuals. This provides a foundation for similar studies and enables researchers to use this dataset in their own work.

4. Performance evaluation: The article presents a comprehensive analysis evaluating the performance of different deep learning models and optimization algorithms. This can assist researchers in comparing their performance when conducting similar studies.

These contributions of the article emphasize the importance of using deep learning and optimization techniques in the diagnosis of spinal diseases such as scoliosis and spondylolisthesis. This study inspires progress in the relevant field of literature and encourages further research.

Article Information Form

Funding
The author(s) received no financial support for the research, authorship, or publication of this study.

Authors' Contribution
All authors contributed equally to the writing of this article. All authors have read and approved the final manuscript.

The Declaration of Conflict of Interest/ Common Interest
No conflicts of interest or common interest have been stated by the authors.

The Declaration of Ethics Committee Approval
This study did not require permission from the ethics committee or any special permission.

The Declaration of Research and Publication Ethics
The authors of the article declare to respect SAUJS scientific, ethical, and citation rules in all processes of the paper and that they did not falsify the collected data. Furthermore, they state that the Sakarya University Journal of Science and its editorial board are not responsible for any ethical violations that may be encountered and that this study has not been evaluated in any environment. Scholarly publications other than the Sakarya University Journal of Science.

References

[1] M. Alafeef, M. Fraiwan, H. Alkhalaf, Z. Audat, "Shannon entropy and fuzzy C-means weighting for AI-based diagnosis of vertebral column diseases" Journal of Ambient Intelligence and Humanized

Computing, Springer, 2020, vol. 11, pp. 2557-2566.

[2] M. R. Konieczny, H. Senyurt, R. Krauspe, "Epidemiology of adolescent idiopathic scoliosis," Journal of children's orthopaedics, London, England, vol 7, no. 1, pp. 3-9, 2013.

[3] M. Alafeef, Z. Audat, L. Fraiwan, T. Manasreh, "Using deep transfer learning to detect scoliosis and spondylolisthesis from X-ray images," Plos one, vol. 17, no. 5, pp. e0267851, 2022.

[4] J. Yang, K. Zhang, H. Fan, Z. Huang, Y. Xiang, J. Yang, H. Lin, "Development and validation of deep learning algorithms for scoliosis screening using back images,"Communications biology, vol. 2, no. 1, pp. 390, 2019.

[5] American Association of Neurological Surgeons, "Scoliosis", Jan. 22, 2023. [Online]. Available:https://www.aans.org/Patients/ Neurosurgical-Conditions- andTreatments/Scoliosis

[6] S. The American Academy of Orthopaedic Surgeons, "Spondylolysis and Spondylolisthesis", Jan. 22, 2023. [Online]. Available:https://orthoinfo.aaos.org/en/dis eases--conditions/spondylolysis-and- spondylolisthesis

[7] M. A. Deveci, A. Şenköylü, "Gelişimsel spondilolisteziste bel ağrısı: tanı ve tedavi yaklaşımı, " TOTBİD Dergisi:14, pp. 282-289.

[8] J. S. Smith, C. I. Shaffrey, K. M. Fu, J. K. Scheer, S. Bess, V. Lafage, C.P. Ames, "Clinical and radiographic evaluation of the adult spinal deformity patient," in Neurosurgery Clinics, Elsevier, vol. 24, no. 2, pp. 143-156, 2013.

[9] S. Bess, O. Boachie-Adjei, D. Burton, M. Cunningham, C. Shaffrey, A. Shelokov, International Spine Study Group, "Pain

and disability determine treatment modality for older patients with adult scoliosis, while deformity guides treatment for younger patients," Spine, vol. 34, no. 20, pp. 2186-2190, 2009.

[10] J. S. Smith, C. I. Shaffrey, S. D. Glassman, S. H. Berven, F. J. Schwab, C. L. Hamill, Spinal Deformity Study Group, "Risk-benefit assessment of surgery for adult scoliosis: an analysis based on patient age," Spine, vol. 36, no.10, pp. 817-824, 2019.

[11] M. Fraiwan, Z. Audat, L. Fraiwan, T. Manasreh, "Using deep transfer learning to detect scoliosis and spondylolisthesis from X-ray images," Plos one, vol.17, no. 5, pp. e0267851, 2022.

[12] R. F. Masood, I. A Taj, M. B. Khan, M. A. Qureshi, T. Hassan, "Deep learning based vertebral body segmentation with extraction of spinal measurements and disorder disease classification," in Biomedical Signal Processing and Control, Elsevier, vol. 71, pp. 103230, 2022.

[13] M. Tajdari, A. Pawar, H. Li, F. Tajdari, A. Maqsood, E. Cleary, W. K. Liu, "Image-based modeling for adolescent idiopathic scoliosis: mechanistic machine learning analysis and prediction," Computer methods in applied mechanics and engineering, Elsevier, vol. 374, pp. 113590, 2021.

[14] M. Fraiwan, Z. Audat, L. Fraiwan, T. Manasreh, "Using deep transfer learning to detect scoliosis and spondylolisthesis from X-ray images," Plos one, vol. 17, no. 5, pp. e0267851, 2022.

[15] I. Özkan, E. Ülker, "Derin öğrenme ve görüntü analizinde kullanılan derin öğrenme modelleri," Gaziosmanpaşa Bilimsel Araştırma Dergisi, vol. 6, no. 3, pp. 85-104, 2017.

[16] M. Turkoglu, D. Hanbay, A. Sengur, "Multi-model LSTM-based convolutional

neural networks for detection of apple diseases and pests," in Journal of Ambient Intelligence and Humanized Computing, Springer, pp. 1-11, 2019.

[17] A. Krizhevsky, I. Sutskever, G. E. Hinton, "Imagenet classification with deep convolutional neural networks," Communications of the ACM, vol. 60, no. 6, pp. 84-90, 2017.

[18] S. Gökalp, İ. Aydın, "Farklı Derin Sinir Ağı Modellerinin Duygu Tanımadaki Performansların Karşılaştırılması," Muş Alparslan Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, vol. 2, no.1, pp. 35-43, 2021.

[19] K. He, X. Zhang, S. Ren, j. Sun, "Deep residual learning for image recognition" In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016.

[20] B. Anadolu, "Dijital hikâye anlatıcılığı bağlamında yapay zekânın sinemaya etkisi: Sunspring ve It's No Game filmlerinin analizi," Erciyes İletişim Dergisi, vol.1, pp. 39-56, 2019.

[21] M. Tan, Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," In International conference on machine learning, PMLR, pp. 6105-6114, 2019.

[22] E. Seyyarer, F. Ayata, T. Uçkan, A. Krci, "Derin öğrenmede kullanilan optimizasyon algoritmalarnin uygulanmasi ve kiyaslanmasi," Computer Science, vol. 5, no. 2, pp. 90-98, 2020.

[23] M. F. Akca "Gradient Descent Nedir?" Jan.22,2023[Online]. Available:https://medium.com/deep-learning-turkiye/gradient-descent-nedir-3ec6afcb9900.

[24] S. Ruder "An overview of gradient descent optimization algorithms" Jan. 22,2023 [Online].

Available:https://ruder.io/optimizing-gradient-descent/

[25] Matlab [Online] Available:https:www.mathworks.com/products/matlab.html

[26] A. A. Reshi, I. Ashraf, F. Rustam, H. F. Shahzad, A. Mehmood, G. S. Choi, "Diagnosis of vertebral column pathologies using concatenated resampling with machine learning algorithms," Peer Journal Computer Science, vol 7, pp. e547, 2021.

[27] T. Colombo, M. Mangone, F. Agostini, A. Bernetti, M. Paoloni, V. Santilli, L. Palagi, "Supervised and unsupervised learning to classify scoliosis and healthy subjects based on non-invasive rasterstereography analysis," Plos one, vol.16, no. 12, pp. e0261511, 2021.

[28] H. Wang, T. Zhang, K. M. C. Cheung, G. K. H. Shea, "Application of deep learning upon spinal radiographs to predict progression in adolescent idiopathic scoliosis at first clinic visit," EClinicalMedicine, vol 42, pp. 101220, 2021.

[29] Y. Unal, K. Polat, H. E. Kocer, "Pairwise FCM based feature weighting for improved classification of vertebral column disorders,"Computers in biology and medicine,vol. 46, pp. 61-70, 2014.