SAĞLIK BİLİMLERİNDE YAPAY ZEKA DERGİSİ
Journal of Artificial Intelligence in Health Sciences

Cilt | Volume: 4 • **Sayı | Issue:** 2 - Ağustos | August 2024

Optimized AI-Assisted Diagnosis of Spinal Anomalies Using Convolutional Neural Networks by Enhancing Feature Extraction in Small Datasets

Küçük Veri Setlerinde Özellik Çıkarmayı Geliştirerek Evrişimli Sinir Ağları ile Omurga Anomalilerinin Yapay Zeka Destekli Teşhisinin Optimizasyonu

Ozan Durgut*,[1] | Gökhan Bora Esmer[1]

[1] Marmara University, Faculty of Engineering, Istanbul, Türkiye.

ORCID ID: OD: 0009-0003-1068-9462 **GBE:** 0000-0003-2405-0777

Makale Bilgisi | Article Information
Makale Türü | Article Type: Araştırma Makalesi | Research Article
Doi:
Geliş Tarihi | Received: 27.05.2024
Kabul Tarihi | Accepted: 08.07.2024
Yayım Tarihi | Published: 30.08.2024

Sorumlu Yazar | Correspondence Author
Ozan Durgut
ozandurgut.2001@hotmail.com
Address for Correspondence: Marmara University, Faculty of Engineering, Istanbul, Türkiye.

Atıf | Cite As
Durgut, O., & Esmer, G. B. Optimized ai-assisted diagnosis of spinal anomalies using convolutional neural networks by enhancing feature extraction in small datasets. Journal of Artificial Intelligence in Health Sciences. 2024;4(2): 1-10. DOI:

Çıkar çatışması
Yazarlar herhangi bir çıkar çatışması veya finansal destek bildirmemişlerdir.

Finansman
Bu çalışma herhangi bir kurum ya da kuruluş tarafından finanse edilmemektedir.

Etik Onay
Çalışmada kullanılan tüm veriler, açık kaynaklı platformlardan elde edilmiştir. Dolayısıyla, bu kaynakların kullanımıyla ilgili etik onay sürecine ihtiyaç duyulmamıştır.

Yazar Katkıları
Motivasyon / Konsept: OD, GBE.
Çalışma Tasarımı: OD, GBE.
Kontrol / Gözetim: OD, GBE.
Veri Toplanması ve / veya İşlemesi: OD, GBE.
Analiz ve / veya Yorum: OD, GBE.
Literatür incelemesi: OD, GBE.
Makalenin Yazılması: OD, GBE.
Eleştirel İnceleme: OD, GBE.

Optimized AI-Assisted Diagnosis of Spinal Anomalies Using Convolutional Neural Networks by Enhancing Feature Extraction in Small Datasets

ABSTRACT

Purpose: Major spinal health anomalies, particularly idiopathic scoliosis and spondylolisthesis are primarily caused by abnormal vertebral displacements. Early diagnosis is critical for effective treatment and management. However, diagnosis of these conditions requires the analysis of X-ray images by expert physicians, and when the number of patients increases, the amount of required time to have a diagnosis may take longer duration. Also, the concentration of the physician may be lost. As a result of this, physician may have an erroneous decision related to diagnosis. As a solution to the problem, we suggest a method based on artificial intelligence that helps physician to come up with the correct diagnosis.

Materials and Methods: To address the issue of insufficient datasets, we use a customized convolutional neural network model and the Leaky ReLU activation function. This approach helps us extract features better while reducing computational complexity.

Results: In our experiments, we achieve success rates of 98.51% in accuracy, 98.63% in precision, 98.53% in recall, and 98.51% in the F1 score. When we compare these results to another study using the same dataset, we see increases of 2.25% in accuracy, 1.04% in precision, 2.67% in recall, and 4.11% in the F1 score. To avoid misleading results from small or imbalanced datasets, we use a balanced version of the dataset for comparison. When we compare the model trained on the imbalanced dataset with the version trained on the balanced dataset, we find a minimal performance decrease of only 0.787% in the F1 score and an average decrease of 0.721% in the other metrics. This shows that the model performs well regardless of potential issues from dataset imbalance. We also test the model with challenging data and obtain successful metrics.

Conclusion: We achieve the objectives of increasing the success rate by reducing computational complexity and improving feature extraction for small datasets. Furthermore, experiments with challenging datasets show that our method remains generalizable and usable even on small datasets.

Keywords: Spine health, Scoliosis, Artificial intelligence, Convolutional neural network, Leaky ReLU

Küçük Veri Setlerinde Özellik Çıkarmayı Geliştirerek Evrişimli Sinir Ağları ile Omurga Anomalilerinin Yapay Zeka Destekli Teşhisinin Optimizasyonu

ÖZET

Amaç: Omurga sağlığıyla ilgili başlıca anomaliler, özellikle idiopatik skolyoz ve spondilolistezis, esasen anormal vertebral kaymalar nedeniyle ortaya çıkar. Erken teşhis, etkili tedavi ve yönetim için kritiktir. Ancak, bu durumların teşhisi, X-ray görüntülerinin uzman hekimler tarafından analiz edilmesini gerektirir ve hasta sayısı arttığında teşhis süresi daha uzun olabilir. Ayrıca, hekimin dikkati dağılabilir. Bunun sonucunda, hekim teşhisle ilgili yanlış bir karar verebilir. Soruna çözüm olarak, hekime doğru teşhis koymasına yardımcı olacak yapay zeka temelli bir yöntem öneririz.

Gereç ve Yöntem: Veri seti yetersizliği sorununu ele almak için, özelleştirilmiş bir konvolüsyonel sinir ağı modeli ve Leaky ReLU aktivasyon fonksiyonunu kullanırız. Bu yaklaşım, daha iyi özellik çıkarımı yapmamıza ve hesaplama karmaşıklığını azaltmamıza yardımcı olur.

Bulgular: Deneylerimizde, doğrulukta %98.51, kesinlikte %98.63, duyarlılık %98.53 ve F1 skorunda %98.51 başarı oranları elde ediyoruz. Bu sonuçları aynı veri setini kullanan başka bir çalışma ile karşılaştırdığımızda, sırasıyla doğrulukta %2.25, kesinlikte %1.04, geri çağırmada %2.67 ve F1 skorunda %4.11 artışlar görüyoruz. Küçük veya dengesiz veri setlerinden kaynaklanabilecek yanıltıcı sonuçları önlemek için, karşılaştırma için dengelenmiş bir veri seti kullanırız. Dengesiz veri seti ile eğitilen modeli, dengelenmiş veri seti ile eğitilen versiyonu ile karşılaştırdığımızda, F1 skorunda yalnızca %0.787'lik minimal bir performans düşüşü ve diğer metriklerde ortalama %0.721'lik bir düşüş buluruz. Bu, modelin veri setinin dengesizliklerinden kaynaklanabilecek potansiyel sorunlara rağmen iyi performans gösterdiğini gösterir. Ayrıca, modeli zorlu verilerle test ediyoruz ve başarılı metrikler elde ediyoruz.

Sonuç: Hesaplama karmaşıklığını azaltarak ve küçük veri setleri için özellik çıkarımını artırarak başarı oranını artırma hedeflerini yakalarız. Ayrıca, zorlu veri setleri ile yapılan deneyler, yöntemimizin küçük veri setlerinde bile genellenebilir ve kullanılabilir olduğunu gösterir.

Anahtar Kelimeler: Omurga sağlığı, Skolyoz, Yapay zekâ, Evrişimli sinir ağı, Sızdıran ReLU

Introduction

The spine is a crucial component of the skeletal system, supporting the body and housing the spinal cord, which facilitates communication between the brain and the rest of the body [1]-[3]. Maintaining spinal health is essential for overall well-being. Anomalies in the spine, whether hereditary, congenital, or acquired later in life, can significantly impair an individual's quality of life. Among these, scoliosis and spondylolisthesis are the most prevalent conditions [4][5].

Scoliosis, characterized by a lateral deviation of the spine from its vertical axis, often manifests without a known cause, termed idiopathic scoliosis [6]. This condition can lead to postural abnormalities and neurological issues due to spinal pressure [7][8]. Conversely, spondylolisthesis involves the abnormal forward displacement of one vertebra over another, potentially resulting in pain and restricted mobility [9]. Early diagnosis of these conditions is critical to prevent disease progression and the development of severe deformities [10]– [12].

Traditionally, the diagnosis and assessment of these spinal conditions rely on the analysis of X-ray images from various angles, including frontal, lateral, and sagittal views [13][14]. However, the manual examination of these images is time-consuming and requires a significant number of skilled physicians, a demand that exceeds current supply in many healthcare settings. Consequently, there is an increasing interest in leveraging artificial intelligence (AI) and machine learning techniques to expedite the diagnostic process, improve decision-making, and meet growing patient needs [15]– [19].

Numerous AI-based applications have been developed for their early detection and appropriate treatment. For instance, Shrestha et al. [15] employed linear regression and support vector machine methods for scoliosis detection, achieving success rates of 68% and 85%, respectively. Bernstein et al. [16] focused on accurately measuring spinal curvature angles, a critical factor in tracking disease progression, using segmentation methods. Fatima et al. [17] reported a 94.69% success rate in scoliosis detection using the YOLOv5 algorithm. Such comparative analyses highlight the current gaps in the literature and areas for improvement. Chen et al. [18] provided a comprehensive review of machine learning applications in scoliosis clinical trials, underscoring the need for AI models in healthcare to adhere to stringent medical norms and deliver robust results.

However, the scarcity of large medical datasets poses a significant challenge to the effectiveness of these AI models. Shaikhina et al. [19] addressed the small dataset problem in medical AI by exploring various methodological approaches. Unlike large datasets, small datasets often fail to capture sufficient features for effective model training. This study aims to design an optimized model specifically for small datasets, enhancing the accuracy and reliability of scoliosis and spondylolisthesis detection.

Material and Method

In this section, we explain the model structure and leaky rectified linear unit (leaky ReLU). Then we summarize the datasets used in experiments. Afterward, we describe the testing environment.

Model structure

This study employs a customized convolutional neural network (CNN), a method frequently used in image classification tasks. Despite the similarity in appearance among X-ray images of varying spinal slopes, we designed the model to discern subtle differences effectively. It comprises multiple layers, including convolution, normalization, leaky ReLU activation, and sampling, as illustrated in Figure 1. The primary objectives are to achieve rapid results and optimize feature processing by determining appropriate weights. We choose the leaky ReLU activation function to enhance the proposed method's efficiency. leaky ReLU extends the traditional ReLU activation function's capabilities [20].

Leaky ReLU

In designing an artificial neural network model, selecting, and optimizing components is crucial. The activation function, a key component, manipulates neuron outputs to aid in feature extraction. Common activation functions, including Sigmoid, Tanh, ReLU, and SoftMax, present challenges when applied to small datasets [21],[22]. Sigmoid and Tanh can lead to the vanishing gradient problem [23], while ReLU may cause the 'dying ReLU' issue [24][25]. SoftMax can also result in overfitting. Given the need for high accuracy and precision in medical applications, expanding datasets through synthetic methods is often impractical. To mitigate these challenges, the study employs leaky ReLU, a variant of ReLU. Unlike ReLU, leaky ReLU assigns small negative values to negative inputs, effectively preventing the 'dying ReLU' problem, which occurs when neurons consistently produce negative outputs during training, rendering them inactive for feature extraction [20]. This characteristic makes leaky ReLU particularly effective for small datasets.

Dataset

Selecting an appropriate dataset is essential for ensuring reliable and generalizable results in healthcare studies. Fraiwan et al. [26] conducted image classification for scoliosis and spondylolisthesis detection using deep transfer learning methods, achieving a 96.34% accuracy with the DenseNet-101 model. The dataset, which includes 338 X-ray images, comprises 71 healthy, 79 spondylolisthesis, and 188 scoliosis cases from 240 female and 98 male patients. We resized the images to 224x224 pixels without preprocessing before training. Then, we divided them into training, validation, and test sets containing 213, 40, and 85 images, respectively. Despite its utility, the dataset's imbalanced structure may lead to misleading findings. To address this, we balanced the dataset by reducing the number of scoliosis images while preserving the dataset's original characteristics. We achieved this by calculating correlation and variance values within the class through feature extraction, resulting in a subset that retained the original dataset's integrity.

Challenging dataset

The generalizability of a model is as important as its accuracy. To evaluate the proposed model's adaptability, we prepared a challenging dataset comprising 255 new X-ray images from two different sources. As illustrated in Figure 3 [27][28], these images were obtained using various X-ray devices and projected onto different film types, providing diverse and robust test data. This approach ensures that the model's performance is assessed under varied conditions, which is crucial for determining its applicability in real-world scenarios.

Experimental environment

Model generation, training, and testing are conducted using the PyTorch library, an open-source platform developed for artificial intelligence research [29]. The proposed convolutional neural network model comprises 150,528 input neurons, three output neurons, and 24 layers, as shown in Figure 1. After parameter optimization, the model is trained and tested in a cloud-based programming environment, Google Colab, utilizing a Tesla T4 graphics processor [30].

Results

The obtained results are analyzed under two categories. The first category includes the results which used the data provided by Fraiwan et al. [26] and the second category includes the results obtained with challenging dataset.

Dataset

The model is evaluated using the dataset provided by Fraiwan et al. [26], which contains 338 X-ray images. The primary objective is to enhance the model's effectiveness for smaller datasets. The training process involved 1,000 iterations, with the employed parameters summarized in Table 1. Performance metrics, including accuracy, precision, recall, and F1 score, are detailed in Table 2. Additionally, inferred images from the test dataset are presented in Figures 4.a through 4.d. To ensure a more deterministic comparison, the model is also trained using a balanced version of the dataset. The metrics obtained from these trainings, alongside comparisons between the balanced and unbalanced datasets, are shown in Table 3.

Table 1. Model Training Parameters

Model	Learning Rate (At start)	Optimizer	Batch Size	Epochs
The proposed model	1e-3	Adam	4	1000

The training parameters of the proposed CNN based model for the dataset provided by Frawian et al. [26].

Table 2. Model Performance (Original Dataset)

Model	F1 Score	Precision	Recall	Accuracy
The proposed model	98.51%	98.63%	98.53%	98.51%

The performance of the proposed CNN based model when the dataset provided by Fraiwan et al. [26] is used.

Table 3. Balanced vs. Imbalanced Dataset Performance

Dataset	F1 Score	Precision	Recall	Accuracy
Balanced	97.74 %	97.96%	97.83%	97.83%
Imbalanced	98.51%	98.63%	98.53%	98.51%
Difference	**0.787%**	**0.684%**	**0.716%**	**0.695%**

Comparison of the performance metrics for the proposed model using balanced and imbalanced(original) version of the dataset respectively [26].

Challenging data

To assess the model's generalizability, we tested it on an unseen dataset comprising 255 new X-ray images, which extended the original dataset. These new images, derived from different X-ray devices and film types, provided diverse characteristics, offering a robust test environment. The model's performance on this challenging dataset is summarized in Table 5, where the corresponding metrics indicate its adaptability and effectiveness across varied conditions.

Table 4. Performance on Challenging Dataset

Model	F1 Score	Precision	Recall	Accuracy
The proposed model	86.57%	89.55%	83.37%	84.11%

The performance of the proposed model when new dataset with 255 different x-ray images is used in testing [27], [28].

Table 5. Performance Comparison with Related Study.

Model	F1 Score	Precision	Recall	Accuracy
The proposed model	98.51%	98.63%	98.53%	98.51%
Fraiwan et al. - DenseNet-201	95.97%	97.61%	94.62%	96.34%
Improvement	**2.67%**	**1.04%**	**4.11%**	**2.25%**

Comparison of the performance results between the proposed model and the model presented by Fraiwan et al [26].

Discussion

The proposed model demonstrated improved accuracy and precision when compared to the results obtained by Fraiwan et al. [26]. This improvement is crucial, particularly in medical applications where lower tolerance levels are essential due to the direct impact on patient outcomes. Table 5 presents a performance comparison between our model and the DenseNet-201 model used by Fraiwan et al. [26]. Since the paper includes models with a large number of parameters, it is also possible to infer the idea of using smaller scale models for small dataset optimization, which is the goal of this study.

The proposed model outperformed DenseNet-201, largely due to the use of the leaky ReLU activation function. Most of the methods tested in this paper use the ReLU activation function. This is due to the higher computational complexity of methods like leaky ReLU compared to standard ReLU. Models requires vast number of parameters are usually intended to be run on large datasets (ImageNet, COCO, CIFAR-100, etc.) and the computational complexity is needed to be kept at an acceptable level [31], [32]. Nevertheless, when the objective is to work with small datasets, ability of leaky ReLU to solve problems such as vanishing gradient compensates its drawbacks such as computational complexity. As a result of this we can achieve higher precision and accuracy for small datasets. Meanwhile, it provides feasible computational complexity by dealing with fewer parameters. The accuracy, precision, sensitivity and F1 score metrics improved by 2.25%, 1.04%, 4.11% and 2.67%, respectively when leaky ReLU is utilized. In calculating improvement, the DenseNet-201 results of Frawian et al.'s [26] study are considered. In addition to this increase, it should be emphasized that DenseNet-201 has over 20 million parameters, while the convolutional neural network model proposed in this paper works with only around 400 thousand parameters. Using models with a high number of parameters when working with small datasets leads to various training problems, especially overfitting, and creates an unnecessary level of computational complexity.

We also conducted a comparative analysis between the model trained on the balanced dataset and the one trained on the unbalanced dataset. This comparison aimed to determine whether the proposed method provides a genuine improvement, particularly when addressing small, unbalanced datasets that might otherwise produce misleading results. As shown in Table 4, the F1 score—a metric that balances precision and recall—decreased by only 0.787%, while all metrics saw an average decrease of 0.721%. This minimal reduction demonstrates that the proposed model maintains high performance and reliability, even when potential biases in the dataset are accounted for.

When tested with the challenging dataset, the model achieved an acceptable level of success across all metrics. This evaluation is conducted to determine the model's robustness and to ensure that no overfitting occurred. Despite the specific optimizations and improvements made for small datasets, the architecture and methods employed in this study demonstrate potential success in other studies facing similar data constraints. These results underscore the model's generalizability and its ability to perform effectively across varied conditions.

Conclusion

The spine is one of the most important structures for maintaining a healthy body. The common problems that endanger the health of the spine are scoliosis and spondylolisthesis, anomalies for which early treatment is very important. Their detection by physicians takes significant amount of time and resources. For this reason, a method has been proposed to detect anomalies based on artificial intelligence. However, the difficulty in finding large medical datasets has led to the creation of more optimized models for small datasets. Hence, a model based on convolutional neural networks and the leaky ReLU activation function has developed. The purpose of using leaky ReLU is to overcome the "Dead ReLU" problem of the standard ReLU function and to better extract features for small datasets. Conducted experiments, and the calculated performance results outperformed the ones obtained in the studies of Fraiwan et al. [26] by 2.52% on average. Besides, the comparison of our proposed model trained on both balanced and unbalanced datasets demonstrates its true effectiveness, showing a minimal F1 score decrease of only 0.787%. This indicates that our model maintains high performance and reliability, even when accounting for potential biases in the dataset. The proposed model achieves successful results. Tests with challenging dataset also show that the model can be successful and generalizable for different small datasets. This study not only advances the field of AI-assisted medical diagnosis but also provides a viable solution to the challenges posed by small datasets in healthcare. Future research should continue to refine and validate these approaches, ensuring their readiness for widespread clinical adoption.

References

1. Rohit Aiyer. Chapter 1 - an overview on the anatomy of the spine. In Alaa Abd-Elsayed, editor, *Decompressive Techniques (First Edition)*, Atlas of Interventional Pain Management Series, pages 1–12. Elsevier, New Delhi, first edition, 2024.
2. Adrese Michael Kandahari, Varun Puvanesarajah, Francis H. Shen, Jon Raso, and Hamid Hassanzadeh. 1 - anatomy of the spine. In Dino Samartzis, Jaro I. Karppinen, and Frances M.K. Williams, editors, *Spine Phenotypes*, pages 1–34. Academic Press, 2022.
3. John H. Bland and Dallas R. Boushey. Anatomy and physiology of the cervical spine. *Seminars in Arthritis and Rheumatism*, 20(1):1–20, 1990.
4. Max Aebi. The adult scoliosis. *European spine journal*, 14:925–948, 2005.
5. Norman Capener. Spondylolisthesis. *The British Journal of Surgery,* volume 19, pages 374-386, 1932.
6. Jack C Cheng, René M Castelein, Winnie C Chu, Aina J Danielsson, Matthew B Dobbs, Theodoros B Grivas, Christina A Gurnett, Keith D Luk, Alain Moreau, Peter O Newton, et al. Adolescent idiopathic scoliosis. *Nature reviews disease primers*, 1(1):1–21, 2015.
7. Alex MacLennan. Scoliosis. *The British Medical Journal*, pages 864–866, 1922.
8. Caroline J Goldberg, David P Moore, Esmond E Fogarty, and Frank E Dowling. Scoliosis: a review. *Pediatric surgery international*, 24:129–144, 2008.
9. JA Fitzgerald and PH Newman. Degenerative spondylolisthesis. *The Journal of Bone & Joint Surgery British Volume*, 58(2):184–192, 1976.
10. Joseph A Janicki and Benjamin Alman. Scoliosis: Review of diagnosis and treatment. *Paediatrics & child health*, 12(9):771–776, 2007.
11. Joseph S Lombardi, Leon L Wiltse, James Reynolds, Eric H Widell, and CURTIS SPENCER III. Treatment of degenerative spondylolisthesis.

Spine, 10(9):821–827, 1985.

12. JOHN R FISK, JOHN H MOE, and ROBERT B WINTER. Scoliosis,
13. spondylolysis, and spondylolisthesis: their relationship as reviewed in 539 patients. *Spine*, 3(3):234–245, 1978.
14. Yeliz Basar, Deniz Alis, Deniz Esin Tekcan Sanli, Tugana Akbas, and Ercan Karaarslan. Whole-body mri for preventive health screening: Management strategies and clinical implications. *European Journal of Radiology*, 137:109584, 2021.
15. Tomasz Kotwicki. Evaluation of scoliosis today: examination, x-rays and beyond. *Disability and rehabilitation*, 30(10):742–751, 2008.
16. Pratik Shrestha, Aachal Singh, Riya Garg, Ishika Sarraf, TR Mahesh, and G Sindhu Madhuri. Early stage detection of scoliosis using machine learning algorithms. In *2021 International Conference on Forensics, Analytics, Big Data, Security (FABS)*, volume 1, pages 1–4. IEEE, 2021.
17. Peter Bernstein, Johannes Metzler, Marlene Weinzierl, Carl Seifert, Wadim Kisel, and Markus Wacker. Radiographic scoliosis angle esti- mation: spline-based measurement reveals superior reliability compared to traditional cobb method. *European spine journal*, 30:676–685, 2021.
18. Joddat Fatima, Mashood Mohsan, Amina Jameel, Muhammad Usman Akram, and Adeel Muzaffar Syed. Vertebrae localization and spine segmentation on radiographic images for feature-based curvature clas- sification for scoliosis. *Concurrency and Computation: Practice and Experience*, 34(26):e7300, 2022.
19. Kai Chen, Xiao Zhai, Kaiqiang Sun, Haojue Wang, Changwei Yang, and Ming Li. A narrative review of machine learning as promising revolution in clinical practice of scoliosis. *Annals of Translational Medicine*, 9(1), 2021.
20. Torgyn Shaikhina and Natalia A Khovanova. Handling limited datasets with neural networks in medical applications: A small-data approach. *Artificial intelligence in medicine*, 75:51–63, 2017.
21. Andrew L Maas, Awni Y Hannun, Andrew Y Ng, et al. Rectifier nonlinearities improve neural network acoustic models. In *Proc. icml*, volume 30, page 3. Atlanta, GA, 2013.
22. Andrinandrasana David Rasamoelina, Fouzia Adjailia, and Peter Sincák. A review of activation function for artificial neural network. In *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, pages 281–286. IEEE, 2020.
23. Bin Ding, Huimin Qian, and Jun Zhou. Activation functions and their characteristics in deep neural networks. In *2018 Chinese control and decision conference (CCDC)*, pages 1836–1841. IEEE, 2018.
24. Sepp Hochreiter. The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 6(02):107–116, 1998.
25. Lu Lu, Yeonjong Shin, Yanhui Su, and George Em Karniadakis. Dying relu and initialization: Theory and numerical examples. *arXiv preprint arXiv:1903.06733*, 2019.
26. Scott C Douglas and Jiutian Yu. Why relu units sometimes die: Analysis of single-unit error backpropagation in neural networks. In *2018 52nd Asilomar Conference on Signals, Systems, and Computers*, pages 864– 868. IEEE, 2018.
27. Mohammad Fraiwan, Ziad Audat, Luay Fraiwan, and Tarek Manasreh. Using deep transfer learning to detect scoliosis and spondylolisthesis from x-ray images. *Plos one*, 17(5):e0267851, 2022.
28. Hieu T Nguyen, Hieu H Pham, Nghia T Nguyen, Ha Q Nguyen, Thang Q Huynh, Minh Dao, and Van Vu. Vindr-spinexr: A deep learning framework for spinal lesions detection and classification from radiographs. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France,*

September 27–October 1, 2021, Proceedings, Part V 24, pages 291–301. Springer, 2021.

29. Scoliosis test dataset, accurate automated spinal curvature estimation. Online, 2019.
30. Eli Stevens, Luca Antiga, and Thomas Viehmann. *Deep learning with PyTorch*. Manning Publications, 2020.
31. Ekaba Bisong and Ekaba Bisong. Google colaboratory. *Building machine learning and deep learning models on google cloud platform: a comprehensive guide for beginners*, pages 59–64, 2019.
32. Alexandra L'heureux, Katarina Grolinger, Hany F Elyamany, and Mi- riam AM Capretz. Machine learning with big data: Challenges and approaches. *Ieee Access*, 5:7776–7797, 2017.
33. Lina Zhou, Shimei Pan, Jianwu Wang, and Athanasios V Vasilakos. Machine learning on big data: Opportunities and challenges. *Neuro-computing*, 237:350–361, 2017.

Supplementary Material

Figure 1. Model structure of the proposed CNN based model.

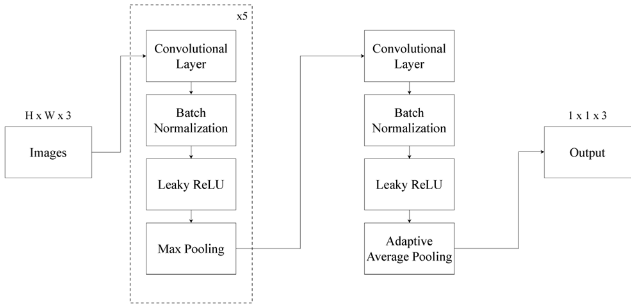
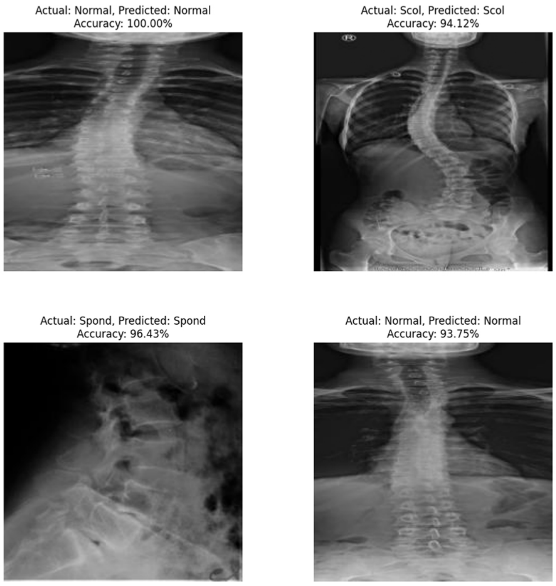


Figure 2. X-ray image example of a scoliosis from the dataset provided by Fraiwan et al. [26].

Figure 3. X-ray image example of a scoliosis from the challenging data [27], [28].

Figure 4. Inferenced images from test dataset provided by Fraiwan et al. [26].



4.a) Prediction of normal vertebrae.
4.b) Prediction of scoliosis.
4.c) Prediction of spondylolisthesis
4.d) Prediction of normal vertebrae.