IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 13, No. 3, September 2024, pp. 3414∼3425
ISSN: 2252-8938, DOI: 10.11591/ijai.v13.i3.pp3414-3425

Computer aided detection for vertebral deformities diagnosis based on deep learning

Nabila Ounasser[1], Maryem Rhanoui[2,3], Mounia Mikram[3], Bouchra El Asri[1]
[1]IMS Team, ADMIR Laboratory, Rabat IT Center, ENSIAS, Mohammed V University, Rabat, Morocco
[2]Laboratory Health, Systemic, Process, Research Unit 4129, University Claude Bernard Lyon 1, Lyon, France
[3]Meridian Team, LYRICA Laboratory, School of Information Sciences, Rabat, Morocco

Article Info

Article history:

Received Sep 25, 2023
Revised Feb 26, 2024
Accepted Mar 11, 2024

Keywords:

Automatic spine diagnosis
Convolutional neural network
Deep learning
Generative adverserial network
Medical imaging
Scoliosis
Spinal deformity

ABSTRACT

The diagnosis of spinal deformities is one of the most frequent daily clinical routine. X-ray images are used to diagnose several pathologies in order to reduce harmful radiations of the patient. Spinal deformities are diagnosed essentially from vertebral shapes, orientations, and positions, so their detection and segmentation are major steps required for diagnosis. Deep learning could be applied for automatic diagnosis to detect scoliosis and its variants with a favourable performance. In this study, based on 609 spinal anterior-posterior x-ray images obtained from the public SpineWeb, we examine generative adversarial network (GAN) based architectures and convolutional neural network (CNN) based architectures models that are capable of automatically detecting anomalies in radiograph and achieve expert-level performances in various fields providing a solid comparative study. Most of the implemented models are apt to automatically distinguish limits between vertebrae so determining their shape with a very good visual performance. The GAN-based architecture estimates the required vertebral landmarks with an accuracy rate of 0.966, signify its capacity for automatic scoliosis assessment in a clinical setting.

Corresponding Author:

Nabila Ounasser
IMS Team, ADMIR Laboratory, Rabat IT Center, ENSIAS, Mohammed V University
Rabat, Morocco
Email: nabilaounasser81@gmail.com

1. INTRODUCTION

Musculoskeletal irregularities represent the most common medical condition, leading to persistent discomfort and impairment over time. Consequently, accurately identifying abnormalities in radiographic images is an essential undertaking in the field of medicine [1]. Evaluating X-rays to diagnose orthopedic ailments such as bone deformities, tumors, and fractures is a labor-intensive process that demands the expertise of qualified professionals. Consequently, the creation of a computer-assisted diagnostic system for detecting anomalies in X-ray images has garnered significant interest [2].

The spine is the pillar of the body, it is the substrate of the musculoskeletal system that is breathable of our mobility, it supports and sustains the body and the structure of its organs. Despite their criticality, spinal pathologies are often unaware of the diagnosis, especially spinal deformities. Spinal deformity is an abnormal alignment or curve of the bony vertebral column. Early detection and orthotic treatment of scoliosis would reduce the need for surgical intervention [3]. Therefore, computer-assisted assistance is needed for an efficient and early detection of these pathologies, allowing an effective prevention or treatment [4].

The current state of the art in the field of vertebral deformity diagnosis has seen a significant shift towards the utilization of deep learning techniques. These advancements have revolutionized the accuracy and efficiency of computer-aided detection systems, enabling more precise identification and characterization of vertebral abnormalities. Recent studies have demonstrated the effectiveness of deep learning models in automatically detecting and classifying various types of vertebral deformities from medical imaging data, offering a promising avenue for improving clinical diagnosis and patient care. Computer vision is introduced for image analysis due to its promising performance in extracting information from images. Many tasks have been performed by computer vision, including automated anomaly detection [2], identification and classification of fracture [5], [6], diabetic retinopathy screening [7], and skin lesion classification [8]. Several deep learning models have been investigated in this direction such as generative adversarial networks (GANs) [9]–[11] and convolutional neural network (CNN) [12] that facilitate anomalies detection and achieved expert-level performances in various fields. Most of these approaches focus on computed tomography (CT) datasets only. However, these methods are rarely applicable to X-ray images because of additional difficulties. Radiography is used for the diagnosis of various pathologies. It allows the visualization of a change in volume or a structural abnormality. The cross-sectional images obtained allow to evaluate the shape, position, volume, size, and possible abnormalities of a multitude of anatomical structures, depending on the region being explored. Also X-ray images have a lower resolution. All these facts are detrimental to the automation of detection procedures for X-ray data sets.

Spinal image processes are poorly seen on radiographic images, which is common, frequent, and remains the first reflex in clinical practice. In addition, transverse processes are usually not seen at all because they are outside the acquisition volume. Therefore, we focus on detecting the vertebral bodies and then delineating the entire vertebrae. Additionally, GANs and CNNs are being forcefully explored for anomaly detection [2], [5], [6] in other areas including intrusion detection [13], fraud detection [14], to protect valuable systems, and since there is no more valuable that our human body, in this study we will consider the human body as our system and protect it from anomalies.

In this research, we aim to investigate both approaches with the goal of applying them in the field of medicine, specifically for a comprehensive comparative study on the detection of spinal deformities, particularly scoliosis, at varying degrees. Our study focuses on enhancing the quality of spinal deformity detection in X-ray image tasks. We aim to address this need by leveraging the power of GANs and CNNs to improve the accuracy and efficiency of detecting orthopedic irregularities in radiographic images. To achieve this, we illustrate the utilization of GANs and CNNs for identifying orthopedic irregularities in radiographic images. We validate their effectiveness in detecting scoliosis at different severity levels using a publicly available dataset comprising 609 anterior-posterior spinal X-ray images sourced from SpineWeb (http://spineweb.digitalimaginggroup.ca).

The subsequent sections of this paper are structured as follows: we begin with the background section, followed by section 3, which provides a concise overview of related research. Section 4 outlines the methodology employed in our study, while section 5 details the materials used, including the dataset and implemented models. Section 6 delves into the discussion of the results, and to finally conclude our research in the section 7.

2. RELATED WORK

X-ray analysis is a widely used medical method for diagnosing orthopedic conditions, including bone deformities, tumors, and fractures. In this section, we conduct a review of existing deep learning models developed for detecting anomalies in orthopedic musculoskeletal radiographs. Numerous researchers have trained CNNs on bone X-ray images. Dias [15] employed transfer learning techniques such as feature extraction and fine-tuning to enhance the detection of musculoskeletal abnormalities in X-ray images.

Recent research has explored deep anomaly detection methods, such as GANs like AlphaGAN [16], BiGAN [17], and more to improve anomaly detection tasks. Researchers have made efforts to enhance the performance of these models by modifying their components. For instance, [18], [19] GANomaly has seen improvements through extensions like skip-connections which employ an autoencoder to map the reconstructed input back to the latent space. Song *et al.* [20] proposed a Res-unetGAN model based on the GAN architecture and applied it to Mura. This network consists of two parts: a generator and a discriminator. The encoder component of the generator employs ResNet50 to extract features from normal samples and obtain their potential feature vector representations.

Lately, there has been renewed interest in spine detection and spinal shape analysis. Several deep learning models have been developed for spine-related tasks, utilizing X-rays, MRIs, or CT images. For example, Yi *et al.* [21] has proposed models for spine-related tasks. Additionally, several researchers [3], [22] have explored spine-related tasks using different imaging modalities. Han *et al.* [23] introduced SpineGAN, a model designed to handle the complex and variable nature of spinal structures. SpineGAN incorporates an atrous convolution autoencoder module to capture semantic task-aware representations while preserving fine-grained structural information. He *et al.* [24] propose one-stage methods capable of simultaneously segmenting discs, vertebrae, and neural foramen using GAN-based models. Deep neural networks have also been employed to detect spine vertebrae, leading to significant improvements in performance. Du *et al.* [25] introduced SpineNet, a backbone architecture with scale-permuted intermediate features and cross-scale connections that was learned through neural architecture search. Wu *et al.* [26] introduced an innovative approach for automatically estimating landmarks in adolescent idiopathic scoliosis (AIS) assessment by combining CNN (ConvNet) with statistical techniques to accommodate the variability seen in X-ray images. More recently, Yeh *et al.* [27] tackled the task of automatically detecting landmarks and performing alignment analysis in whole-spine lateral radiographs, employing a deep learning approach. Cina *et al.* [28] proposed a trainable two-step deep learning approach for landmark localization in spine radiographs. Furthermore, Zukić *et al.* [29] have employed CNNs for detecting vertebra centers.

In Table 1, we present a summary of recent methods employed in the field of spine deformity detection. These methods encompass a range of approaches, including CNNs, autoencoder architectures, and traditional machine learning techniques. While these existing methods have made significant strides in spine deformity detection, our study aims to introduce novel advancements to further enhance the accuracy and efficiency of diagnosis. Building upon the foundations laid by previous research, we propose the integration of GAN into the diagnostic pipeline. By harnessing the power of GANs for synthetic data generation and feature representation learning, we anticipate a substantial improvement in the detection of subtle spine abnormalities and variations. Through rigorous experimentation and validation, we expect our proposed methodology to outperform existing approaches, offering clinicians a more reliable and comprehensive tool for early diagnosis and personalized treatment planning.

Table 1. Summary of recent works of spine deformity detection

[Ref]	Dataset	Approach
[3]	EOS imaging system	Introduce an automated method for extracting anatomical parameters from biplanar radiographs of the spine.
[21]	ASCE MICCAI 2019 challenge	Introduce a method for accurately detecting landmarks in AIS, crucial for precise Cobb angle estimation. By localizing vertebra centers and tracing corner landmarks through learned offsets.
[25]	ILSVRC-2012 and COCO datasets	Propose Spinenet to optimize performance by training a backbone network to efficiently handle scale variations, thereby enhancing recognition and localization accuracy.
[27]	Clinical dataset	Presents a deep learning approach for automatically detecting landmarks and analyzing alignment in whole-spine lateral radiographs. The proposed method aims to identify landmarks and assess alignment in spinal images.
[28]	IRCCS Istituto Ortopedico Galeazzi	Introduce a 2-step deep learning model tailored for landmark localization in spine radiographs. The approach aims to enhance accuracy in identifying key anatomical landmarks crucial for diagnostic assessments.
[29]	Clinical datasets	Propose a robust detection and segmentation method for diagnosing vertebral diseases using routine MRI images. The approach aims to detect and segment vertebral abnormalities, facilitating more precise diagnosis and treatment planning.

3. BACKGROUND

3.1. Vertebra detection

Detecting vertebrae involves employing various methods, each with its own set of techniques and algorithms. First, one prominent technique is the utilization of the Viola-Jones method [29]. This method primarily focuses on detecting the centers of the vertebrae, providing a foundational step in the overall process. The second approach involves employing an object detector to identify the vertebrae as bounding box entities. These bounding box objects are subsequently inputted into a landmark regression network as distinct images. The integration of both methods, alongside advancements in deep learning techniques, has significantly enhanced the accuracy and efficiency of vertebrae detection systems. Figure 1 illustrates this process, showcasing the transformation from bounding box detection to landmark-based reconstruction on the original images.

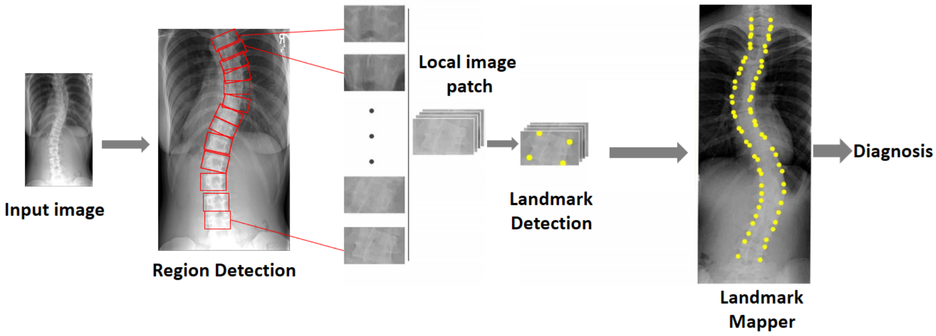

Figure 1. Pipeline for spine detection

3.2. Scoliosis diagnosis

Orthopedic anomalies are frequent reasons for consultation from childhood. Several pathologies attack the structure of the spine (vertebral fracture, inflammation of the discs, deformation of the spine) to detect these phathologies, doctors may need multimodal radiologies (magnetic resonance imaging (MRI), CT scans, and X-ray) depending on the type of disease. In our study, we focus on spinal deformity types which can be diagnosed from vertebra detection using X-ray images as the primary material.

Scoliosis with its varying degrees is a deformation of the spine in the 3 planes of space. It is typically identified during childhood or the early teenage years. The spine normally exhibits natural curves in the cervical, thoracic, and lumbar regions, aligning in the ”sagittal” plane. These inherent curves serve to align the head with the pelvis and act as shock absorbers, evenly distributing mechanical stresses during bodily movement. Scoliosis, however, is commonly described as an abnormal curvature of the spine in the ”coronal” (frontal) plane. Despite its measurement primarily occurring in the frontal plane, scoliosis is, in fact, a more intricate condition.

4. METHODS

To achieve our objective, we implemented state-of-the-art GANs and CNNs, most famous families in deep learning, Figure 2, tailored to the unique challenges of spinal deformity detection. Our approach involved fine-tuning the models using different techniques notably data augmentation techniques to effectively capture the intricate features indicative of various deformity types. This adaptation process was crucial, as existing models were not specifically designed for this mission. We chose GANs for their potential to generate synthetic data, which we anticipated would enhance the models' performance in detecting subtle deformities. In this section we will review CNN and GAN models investigated in this study. Those families of models apply the second method of Vertebra detection as explained in the background section.

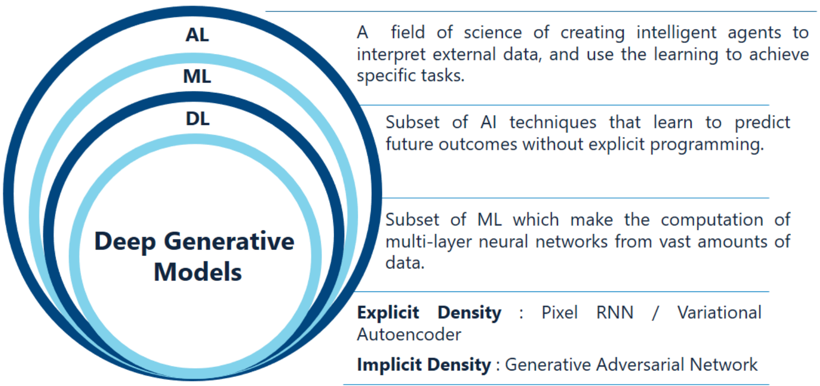

Figure 2. A summary of the concepts encompassing artificial intelligence

4.1. Overview of generative adversarial networks

GANs [30] is considered as one of the most powerful member of the neural network family, due to realistic data-generation capacities. GANs offer a significant advantage in their capacity to generate data, which has led to their successful application in various computer vision tasks such as anomaly detection, image generation, and image super-resolution [2]. In musculoskeletal imaging, automating the detection and segmentation of vertebral degenerative disease is crucial for expediting and streamlining the radiology diagnostic workflow. Deep learning methods have been extensively employed in this field, including GAN-based approaches, which are particularly adaptable. As showen in Figure 3 GANs operate by training two competing networks: a generator and a discriminator. The generator produces realistic synthetic samples from noise (the z-latent space), while the discriminator discerns between genuine and synthetic samples. This flexible architecture has been utilized for tasks like identifying the location of vertebrae, discs, and spinal shape.

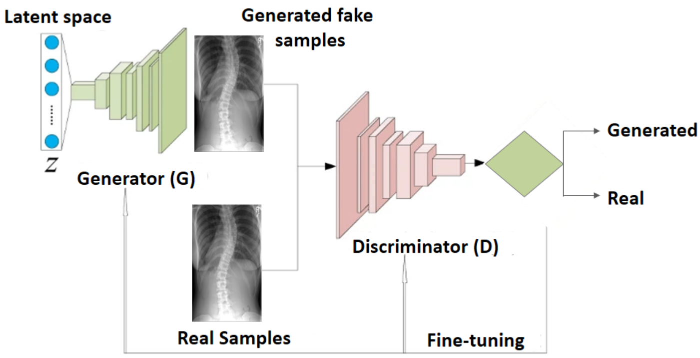

Figure 3. GAN's architecture

4.1.1. SpineGAN

SpineGAN [23] was developed to detect spinal abnormalities and uncovering potential underlying pathological factors. The architecture of SpineGAN consists of two networks Figure 4, each comprising three modules. Firstly, there is a specialized segmentation network tasked with segmenting and classifying neural foramen, intervertebral discs, and vertebrae in radiological images. This segmentation network integrates a deep atrous convolution autoencoder module for encoding spinal images and conducting pixel-level classification. Additionally, it incorporates a recurrent neural network (RNN) module based on local long short term memory network (LSTM) to dynamically model the spatial relationships among different spinal structures in pathology. In alignment with the principles of GANs, a discriminative network is introduced to oversee and motivate the segmentation network, ensuring the generation of accurate predictions.

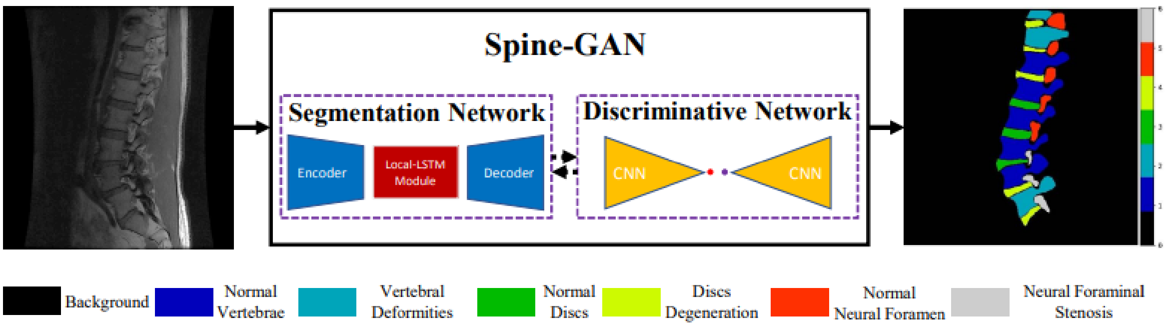

Figure 4. SpineGAN's architecture [23]

4.1.2. Randomized generative adversarial network

Randomized generative adversarial network (RandGAN) [31] for COVID-19 detection. Its architecture composed of two components Figure 5: generator and a discriminator. The particularity of RandGAN's architecture is the Inception and residual block. To enhance the generalizability of RandGAN's generator, random images are drawn from the training class cohort and encoded using inception layers. This approach offers variability in both random noise vectors and real image representations during generator training. The inception and residual architecture aims to improve GAN's ability to capture fine details and maintain spatial information across convolution and pooling layers. However, increasing the generator's depth for capturing distant details, while theoretically valid, poses stability and training challenges for deep GANs.

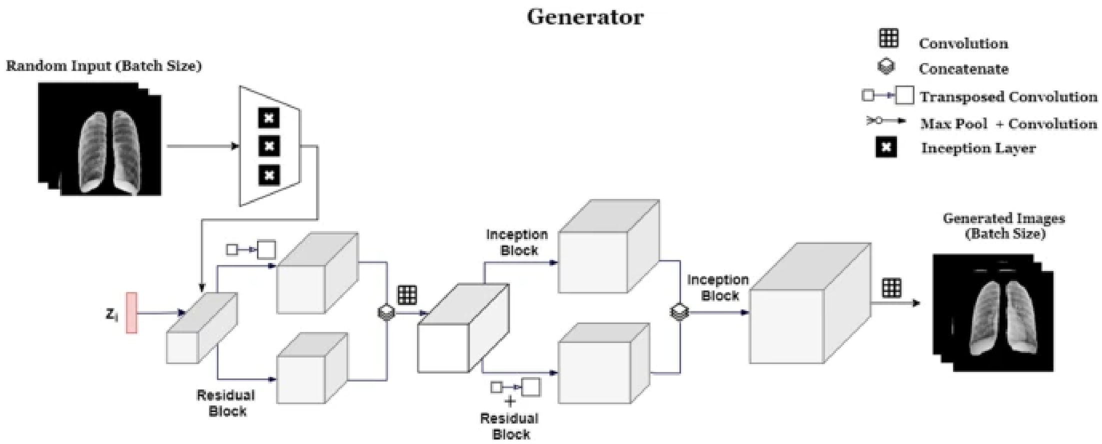

Figure 5. RandGAN's generator architecture [31]

4.1.3. CycleGAN

CycleGAN is one of the first models to have attracted a lot of attention through image-to-image translation using unpaired images [22]. It is composed of two generators and two discriminators as shown in Figure 6. The first generator transforms X into Y and the second one transforms Y into X. The first discriminator have to differentiate real images sampled from X and images produced by the second generator, then this generator is updated accordingly to get a better performance. The second discriminator attempts to differentiate real images sampled from Y and images produced by the first generator, then this generator is updated accordingly to get a better performance. That's what we call the competitive learning, it is a technique focused on improving the model's performance.

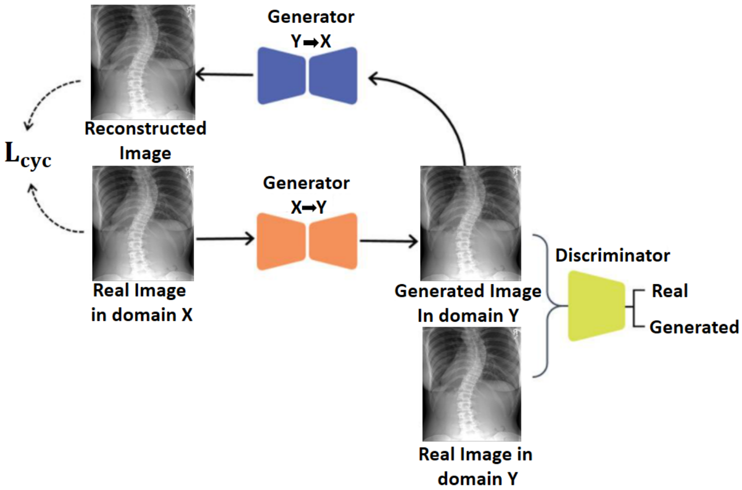

Figure 6. CycleGAN's architecture

4.2. Overview of convolutional neural networks

CNNs, a category of artificial neural networks that have gained prominence in various computer vision tasks, are now garnering attention in diverse domains, including radiology. Detecting anomalies in medical images is a common challenge for radiologists, as these anomalies are infrequent and must be identified amidst numerous normal cases. Recent radiomics studies have explored traditional machine learning models, including techniques for feature extraction, image analysis, and object detection. As we see in Figure 7 a CNN consists of multiple stacked convolutional layers, each with the ability to recognize increasingly complex patterns.

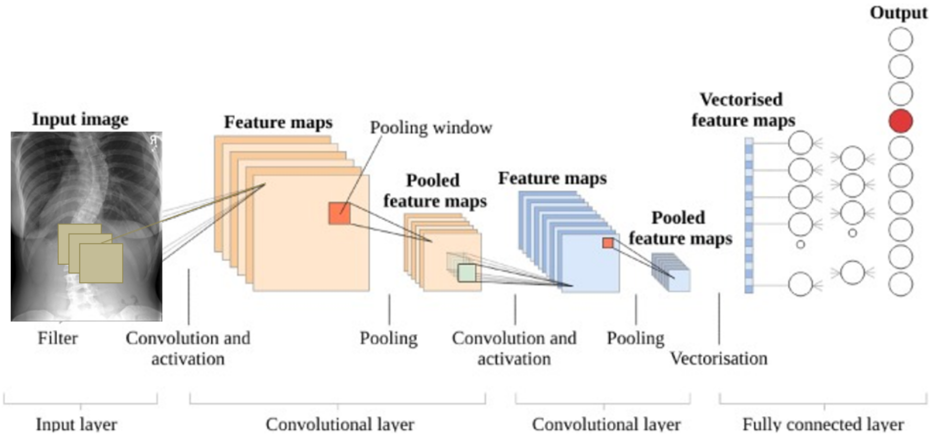

Figure 7. CNN's architecture

This sequential design enables CNNs to learn hierarchical features. Importantly, CNNs do not necessarily require human expert segmentation of anomalies. Through dimensionality reduction, CNNs extract features from images and transform them into a lower-dimensional representation while retaining essential information. In contrast, other deep learning approaches tend to be more computationally intensive, necessitating the use of graphical processing units (GPUs) for model training.

4.2.1. ResNet50

ResNet introduces a residual learning framework designed to facilitate the training of deeper neural networks. Its distinguishing feature lies in the establishment of connections between numerous layers, which simplifies the optimization of the underlying residual mapping, denoted as H(x). There are several variations of the ResNet network model, which differ based on the number of convolutional layers they incorporate. In our particular case, we have chosen to employ the ResNet50 variant, which boasts a depth of 50 layers.

4.2.2. BoostNet

The BoostNet architecture [26] is crafted for the automatic detection of spinal landmarks to facilitate a comprehensive assessment of AIS. The BoostNet architecture effectively addresses the limitations of traditional AIS assessments by enhancing the feature space through the removal of outliers and bolstering robustness by enforcing the integrity of the spinal structure. This architecture comprises three fundamental components. First, a set of convolutional layers serves as feature extractors, autonomously learning features from the dataset. Second, a BoostLayer is employed to eliminate the influence of detrimental outlier features. Finally, a spinal structured multi-output layer functions as a prior mechanism to mitigate the impact of a limited dataset, capturing crucial relationships between each spinal landmark.

4.2.3. SpineNet

SpineNet represents a CNN backbone distinguished by its scale-permuted intermediate features and cross-scale connections, a structure acquired through the process of neural architecture search during training for object detection tasks. This innovative architecture was crafted based on scale-permuted models and was intentionally designed for a fair comparison with ResNet Figure 8. Du *et al.* [25] introduced four distinct architectures within the SpineNet family, each excelling in various latency-performance trade-offs, thus offering versatility for a wide range of use cases. The models are denoted as SpineNet-49/96/143/190. The difference is the feature dimensions in the entire network and number of blocs that constitute the model.

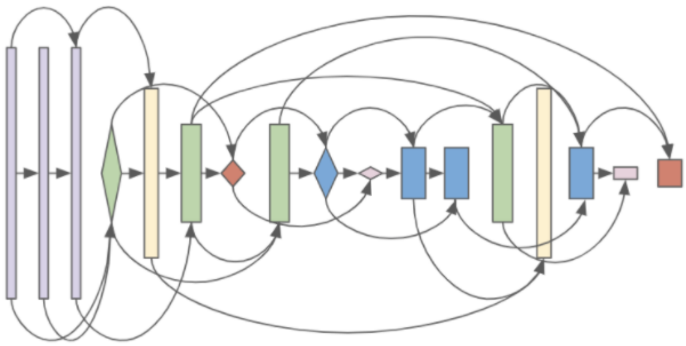

Figure 8. Spine's architecture (scale-permuted model) [25]

5. EXPERIMENT

5.1. Dataset

The dataset contains 609 anterior-posterior radiographic images of the spine obtained from the public SpineWeb repository (http://spineweb.digitalimaginggroup.ca). All images show varying degrees of scoliosis symptoms. They manually annotated the landmarks, each image contains 68 GT landmarks corresponding to the 4 corners of the 17 vertebrae, and 3 Cobb angles. Datatest contains images without GT. During training, the landmarks were scaled to the dimensions of the original image, so that the range of values belonging to the interval [0,1] depends on the location of the landmark relative to the original image. 80% of the dataset was for training (487 images) and 20% for testing (122 images) no patient is placed in both sets. The project code and resources utilized in this study are publicly available on GitHub at: https://github.com/nabinabila/Vertebral-Deformities-Diagnosis-based-on-Deep-Learning.

5.2. Preprocessing

During the preprocessing stage, we implemented data augmentation. This involved enhancing the images to introduce greater diversity into the dataset. This augmentation was performed to enable the models to acquire a deeper understanding of the dataset by learning high-level features that remain consistent despite typical affine transformations, such as horizontal flips, which might occur when generating radiographic images.

6. RESULTS AND DISCUSSION

The results of our implemented models compare favorably with results presented in previous works. Upon reviewing the accuracy of the previously implemented models, it's evident that their performance falls within the range of 0.520 percent to 0.966 percent. Numerous factors can influence the models' performance, including architectural approach, layer design, padding, shape, normalization, activation, loss function, optimizer, batch size, learning rate, pooling, and output layer. Achieving an effective outcome was our primary objective after extensive tuning efforts. Many of our models featured multiple layers and modules, which typically impose a substantial computational burden. Training these models sometimes extended over several days and running them on basic hardware or standard laptop configurations proved to be excessively time-consuming for the dataset.

Furthermore, preprocessing is one of the key for good results in data science tasks. After choosing the deep learning model for the study, it is necessary to prepare a large amount of data. Image size is a parameter that impacts the accuracy of detecting the boundaries between vertebrae. In our experiment, we obtained an acceptable result with the image resolution of 256 × 256 pixels. We explored data augmentation to cope with the limited data available to us, which increases the amount of data in the training phase. The drawbacks of this system that we need to pay attention to are the rotation methods, excessive compression, and shear, as they can impact the performance of intervertebral disc boundary detection. Detailly, Table 2 shows that Spine-GAN, CycleGAN, and RandGAN on average achieves the best accuracy (0.966; 0.922; 0.913 percent). This demonstrates effectiveness of the GAN-based architectures, their modules that are capable to get a deep and accurate representation by conserving the differences between normal and anomalous structures.

Table 2. Detection results on X-Ray images

	CNN-based architecture					GAN-based architecture		
Evaluation Metrics	ResNet50	ConvNet	BoostNet	SpineNet49	SpineNet143	SpineGAN	CycleGAN	RandGAN
Accuracy	0.520	0.563	0.917	0.875	0.933	**0.966**	0.922	0.913
MSE	0.026	0.018	0.006	0.0057	0.0051	0.0046	0.0052	0.0077
Precision	0.459	0.438	0.877	0.866	0,890	0.981	0.933	0.903
Detection Speed	1.26	1.43	4.12	8.56	9.12	7.21	6.33	8.11

While with less processing time the CNNs approaches, Convnet and Resnet50 run faster than GANs, they have lower rate of performance and do not provide orientation estimates. SpineNet models achieved an accuracy of 0.875 and 0.933 percent. In particular, the largest model, SpineNet-143, outperform by 0.933 percent wich is an impressive result for a single model without multi-scale testing during inference. BoostNet attained a commendable accuracy of 0.917 percent, primarily attributable to the contributions of the BoostLayer and the spinal structured multi-output regression layer. These components effectively captured the structural details of the spinal landmark coordinates. Moreover, our models exhibited impressive precision values, reflecting their ability to correctly identify true positive cases while minimizing false positive detections. The high accuracy and precision achieved by our GAN models underscores their reliability and robustness in detecting spine deformities, instilling confidence in their clinical utility and potential for real-world deployment.

In addition to accuracy and precision, remarkably, our GAN models achieved consistently low mean squared error (MSE) values, indicating their proficiency in accurately estimating the extent of deformations and their spatial distribution within the spine images. This fine-grained analysis is invaluable for clinicians in evaluating the severity and progression of spinal abnormalities, facilitating personalized treatment planning and monitoring. Furthermore, our GAN models demonstrated impressive detection speed, enabling rapid and efficient analysis of large volumes of spine imaging data. Leveraging parallel computing architectures and optimized model architectures, our models achieved near-real-time performance without compromising accuracy. This high-speed processing capability enhances the scalability and practicality of our approach, making it well-suited for integration into clinical workflows and telemedicine applications.

Visually in Figures 9 and 10, the illustration serves as a qualitative showcase of GAN's proficiency in detecting spinal landmarks. Regardless of differences in anatomy and image contrast among various patients, GAN consistently and accurately identifies all spinal landmarks. It's noteworthy that the landmarks detected by GAN exhibit a closer conformity to the spinal shape when compared to the performance of ConvNet.

In comparison to previous studies utilizing CNNs for spine deformity detection, our GAN-based approach demonstrated notable advancements in both accuracy and robustness. While CNNs have been widely adopted in medical image analysis due to their ability to automatically extract hierarchical features, they often struggle with capturing subtle deformities and variations in spine images. In contrast, our GAN models leverage adversarial training to generate synthetic data, effectively augmenting the training set and enhancing the models' ability to generalize across diverse deformity patterns. As a result, our GAN models consistently outperformed CNN-based approaches in detecting spine deformities, achieving higher area under the curve (AUC) scores, precision values, and lower MSE.

Similarly, our findings surpass those reported in studies employing autoencoder architectures for spine deformity detection. Although autoencoder models excel in unsupervised feature learning and data compression, they may struggle with preserving important anatomical details and discriminating between normal and abnormal spine configurations. In contrast, our GAN models leverage the discriminative power of adversarial training to explicitly learn the underlying features indicative of spine deformities, thereby achieving superior performance in terms of both accuracy and clinical relevance. By integrating both generative and discriminative components, our GAN models strike a balance between data generation and discrimination, resulting in more effective and interpretable representations of spine deformities.

Furthermore, our study extends beyond the limitations of previous approaches by incorporating a comprehensive evaluation of detection speed, an aspect often overlooked in existing literature. While CNN and autoencoder models have demonstrated promising results in terms of accuracy, their computational efficiency and real-time performance remain areas of concern. In contrast, our GAN models exhibit impressive detection speed, enabling rapid analysis of spine images without compromising accuracy. This improvement in speed-to-accuracy ratio is particularly significant in clinical settings, where timely diagnosis and treatment are critical for patient care. Our study represents a significant advancement in the field of computer-aided detection for vertebral deformities, showcasing the potential of deep learning techniques in improving diagnostic accuracy and efficiency. To our knowledge, this is the first study to examine several models for automatic detection for diagnosis of spinal deformity using X-Ray images. Our observations, from this comparative study, those methods are an effective way to improve orthopedic anomalies detection tasks. In summary, our comparative analysis highlights the superior performance of GAN models in spine deformity detection compared to previous studies utilizing CNN and autoencoder architectures. These findings offer a promising tool for early detection in spinal deformities.

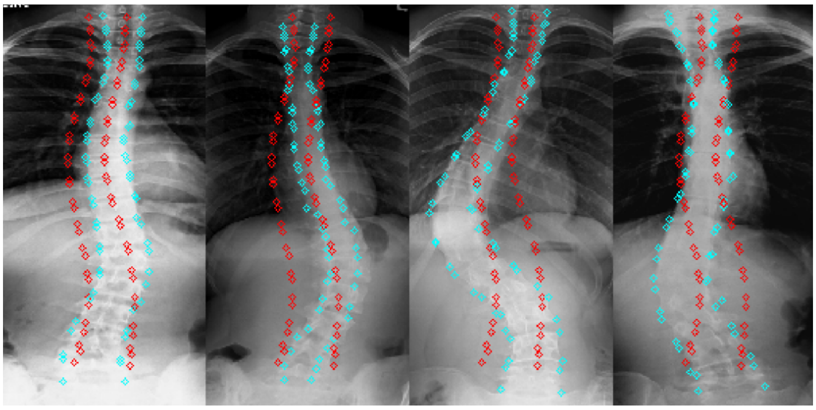

Figure 9. Examples of landmarks detection on X-rays: Convnet

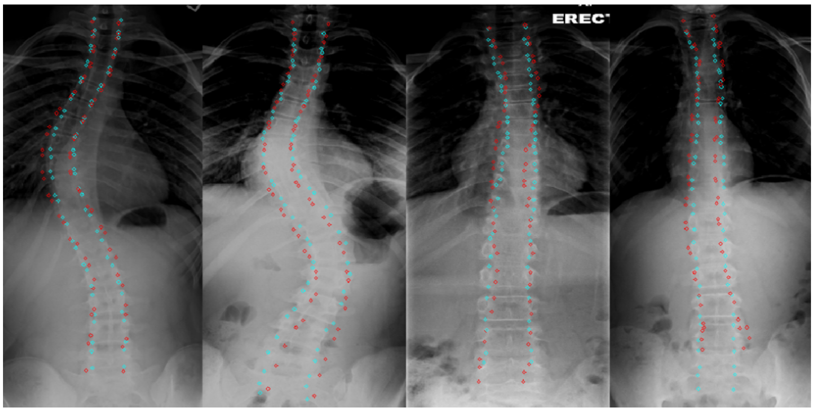

Figure 10. Examples of landmarks detection on X-rays: GAN

7. CONCLUSION

In this paper, we applied CNN and GAN models, most powerful members of the neural network family. Unfortunately they are not explored to diagnosis spinal pathologies. Although the spine is the pillar of the body, it is the substrate of the musculoskeletal system that is breathable of our mobility it supports and sustains the body and the structure of its organs. There is not enough studies that invest to improve medical process for this organ. So, our goal was to examined those models for spinal disease analysis. We had compared and analysed several GAN-based architectures and CNN-based architectures for spinal deformities detection. Summing up the results, it can be concluded that the deep learning methods here presented were apt to automatically determine the spine shape with a very good visual performance. We believe that those methods provide great assistance to clinical experts in orthopedic process analysis. With the improvement of those methods, they will have the potential to be the key for an automated radiological analysis of spinal pathologies, in condition of availability a large training dataset. To conclude, this experiment allowed us to identify the limitations of the models. Future work will explore ways to present a novel approach that could learn specific features for identifying musculoskeletal abnormalities.

REFERENCES

[1] S. Gyftopoulos, D. Lin, F. Knoll, A. M. Doshi, T. C. Rodrigues, and M. P. Recht, “Artificial intelligence in musculoskeletal imaging: current status and future directions,” *American Journal of Roentgenology*, vol. 213, no. 3, pp. 506–513, 2019.

[2] N. Ounasser, M. Rhanoui, M. Mikram, and B. E. Asri, “Generative and autoencoder models for large-scale mutivariate unsupervised anomaly detection,” *Smart Innovation, Systems and Technologies*, vol. 237, pp. 45–58, 2022, doi: 10.1007/978-981-16-3637-0_4.

[3] F. Galbusera et al., “Fully automated radiological analysis of spinal disorders and deformities: a deep learning approach,” *European Spine Journal*, vol. 28, no. 5, pp. 951–960, 2019, doi: 10.1007/s00586-019-05944-z.

[4] A. S. Rawat, A. Rana, A. Kumar, and A. Bagwari, “Application of multi layer artificial neural network in the diagnosis system: A systematic review,” *IAES International Journal of Artificial Intelligence*, vol. 7, no. 3, pp. 138–142, 2018, doi: 10.11591/ijai.v7.i3.pp138-142.

[5] N. Ounasser, M. Rhanoui, M. Mikram, and B. E. Asri, “Anomaly detection in orthopedic musculoskeletal radiographs using deep learning,” in *Proceedings of Eighth International Congress on Information and Communication Technology*, 2023, pp. 93–102, doi: 10.1007/978-981-99-3243-6_8.

[6] N. Ounasser, M. Rhanoui, M. Mikram, and B. E. Asri, “Enhancing computer-assisted bone fractures diagnosis in musculoskeletal radiographs based on generative adversarial networks,” *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 7, pp. 960–966, 2023, doi: 10.14569/IJACSA.2023.01407104.

[7] G. T. Zago, R. V. Andreão, B. Dorizzi, and E. O. T. Salles, “Diabetic retinopathy detection using red lesion localization and convolutional neural networks,” *Computers in Biology and Medicine*, vol. 116, 2020, doi: 10.1016/j.compbiomed.2019.103537.

[8] J. Zhang, Y. Xie, Y. Xia, and C. Shen, “Attention residual learning for skin lesion classification,” *IEEE transactions on medical imaging*, vol. 38, no. 9, pp. 2092–2103, 2019, doi: 10.1109/TMI.2019.2893944.

[9] Z. Qin, Z. Liu, P. Zhu, and Y. Xue, “A GAN-based image synthesis method for skin lesion classification,” *Computer Methods and Programs in Biomedicine*, vol. 195, 2020, doi: 10.1016/j.cmpb.2020.105568.

[10] Y. Zhou, B. Wang, X. He, S. Cui, and L. Shao, “DR-GAN: conditional generative adversarial network for fine-grained lesion synthesis on diabetic retinopathy images,” *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 56–66, 2022, doi: 10.1109/JBHI.2020.3045475.

[11] K. B. Park, S. H. Choi, and J. Y. Lee, “M-GAN: retinal blood vessel segmentation by balancing losses through stacked deep fully convolutional networks,” *IEEE Access*, vol. 8, pp. 146308–146322, 2020, doi: 10.1109/ACCESS.2020.3015108.

[12] T. Ren et al., “Convolutional neural network detection of axillary lymph node metastasis using standard clinical breast MRI,” *Clinical Breast Cancer*, vol. 20, no. 3, pp. e301–e308, 2020, doi: 10.1016/j.clbc.2019.11.009.

[13] S. Huang and K. Lei, “IGAN-IDS: an imbalanced generative adversarial network towards intrusion detection system in ad-hoc network,” *Ad Hoc Networks*, vol. 105, 2020, doi: 10.1016/j.adhoc.2020.102177.

[14] U. Fiore, A. D. Santis, F. Perla, P. Zanetti, and F. Palmieri, “Using generative adversarial networks for improving classification effectiveness in credit card fraud detection,” *Information Sciences*, vol. 479, pp. 448–455, 2019, doi: 10.1016/j.ins.2017.12.030.

[15] D. D. A. Dias, “Musculoskeletal abnormality detection on x-ray using transfer learning,” *Ph.D. Thesis*, Department of Intelligent Interactive Systems, Universitat Pompeu Fabra, Barcelons, Spain, 2019.

[16] K. Raza and N. K. Singh, “A tour of unsupervised deep learning for medical image analysis,” *Current Medical Imaging Reviews*, vol. 17, no. 9, pp. 1059–1077, 2022, doi: 10.2174/18756603mtezonzmk0.

[17] J. Donahue, T. Darrell, and P. Krähenbühl, “Adversarial feature learning,” *arXiv-Computer Science*, 2017.

[18] S. Akcay, A. A. -Abarghouei, and T. P. Breckon, “Skip-GANomaly: skip connected and adversarially trained encoder-decoder anomaly detection,” in *2019 International Joint Conference on Neural Networks (IJCNN)*, 2019, pp. 1–8, doi: 10.1109/IJCNN.2019.8851808.

[19] S. Akcay, A. A. -Abarghouei, and T. P. Breckon, “GANomaly: semi-supervised anomaly detection via adversarial training,” in *Computer Vision – ACCV 2018*, 2019, pp. 622–637, doi: 10.1007/978-3-030-20893-6_39.

[20] S. Song, K. Yang, A. Wang, S. Zhang, and M. Xia, “A Mura detection model based on unsupervised adversarial learning,” *IEEE Access*, vol. 9, pp. 49920–49928, 2021, doi: 10.1109/ACCESS.2021.3069466.

[21] J. Yi, P. Wu, Q. Huang, H. Qu, and D. N. Metaxas, “Vertebra-focused landmark detection for scoliosis assessment,” in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 2020, pp. 736–740, doi: 10.1109/ISBI45749.2020.9098675.

[22] R. Oulbacha and S. Kadoury, “MRI to CT synthesis of the lumbar spine from a pseudo-3D cycle GAN,” in *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, 2020, pp. 1784–1787, doi: 10.1109/ISBI45749.2020.9098421.

[23] Z. Han, B. Wei, A. Mercado, S. Leung, and S. Li, “Spine-GAN: semantic segmentation of multiple spinal structures,” *Medical Image Analysis*, vol. 50, pp. 23–35, 2018, doi: 10.1016/j.media.2018.08.005.

[24] J. He, W. Liu, Y. Wang, X. Ma, and X.-S. Hua, “SpineOne: a one-stage detection framework for degenerative discs and vertebrae,” in *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Dec. 2021, pp. 1331–1334, doi: 10.1109/BIBM52615.2021.9669541.

[25] X. Du et al., “Spinenet: learning scale-permuted backbone for recognition and localization,” in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 11589–11598, doi: 10.1109/CVPR42600.2020.01161.

[26] H. Wu, C. Bailey, P. Rasoulinejad, and S. Li, “Automatic landmark estimation for adolescent idiopathic scoliosis assessment using boostnet,” in *Medical Image Computing and Computer Assisted Intervention*, 2017, pp. 127–135, doi: 10.1007/978-3-319-66182-7_15.

[27] Y. C. Yeh, C. H. Weng, Y. J. Huang, C. J. Fu, T. T. Tsai, and C. Y. Yeh, “Deep learning approach for automatic landmark detection and alignment analysis in whole-spine lateral radiographs,” *Scientific Reports*, vol. 11, no. 1, 2021, doi: 10.1038/s41598-021-87141-x.

[28] A. Cina et al., “2-step deep learning model for landmarks localization in spine radiographs,” *Scientific Reports*, vol. 11, no. 1, 2021, doi: 10.1038/s41598-021-89102-w.

[29] D. Zukić, A. Vlasák, J. Egger, D. Hořínek, C. Nimsky, and A. Kolb, “Robust detection and segmentation for diagnosis of vertebral diseases using routine MR images,” *Computer Graphics Forum*, vol. 33, no. 6, pp. 190–204, 2014, doi: 10.1111/cgf.12343.

[30] I. J. Goodfellow et al., “Generative adversarial nets,” in *Advances in Neural Information Processing Systems*, 2014, pp. 1–9.

[31] S. Motamed, P. Rogalla, and F. Khalvati, “RANDGAN: randomized generative adversarial network for detection of COVID-19 in chest X-ray,” *Scientific Reports*, vol. 11, no. 1, 2021, doi: 10.1038/s41598-021-87994-2.

BIOGRAPHIES OF AUTHORS

Nabila Ounasser holds baccalaureate's degree in Mathematic science and a state engineering diploma in Data and Knowledge from the School of Information Sciences (ESI). Currently, she is pursuing a Ph.D. at the ENSIAS (École Nationale Supérieure d'Informatique et d'Analyse des Systèmes) within the Department of Computer Science. Her research focuses on exploring artificial intelligence for anomaly detection within the IT architecture and model driven systems development (IMS) team. Her research areas of interest include artificial intelligent and computer vision. She can be contacted at email: nabilaounasser81@gmail.com.

Bouchra El Asri currently holds the position of Teaching Research Director at ENSIAS (École Nationale Supérieure d'Informatique et d'Analyse des Systèmes). She was a Technical Director at Cyber Machine. She has successfully led two major projects for prominent national organizations. Additionally, she holds several key roles, including Department Head of Software Engineering, Coordinator of the Software Engineering program at ENSIAS, and responsibility for the development of an enhanced version of the program. She is actively involved in the institution as a member of the governing council, pedagogical committee, and budget monitoring committee. Furthermore, they played a crucial role in transitioning the Software Engineering program to online teaching during the COVID-19 crisis. She has a strong research background, having supervised and currently supervising multiple doctoral theses in the fields of software architecture and data management for healthcare, industry, and education. Her expertise and contributions extend to scientific committees, doctoral study centers, and teaching modules within the Software Engineering program at ENSIAS. She can becontacted at email: bouchra.elasri@ensias.um5.ac.ma or b.elasri@um5s.net.ma.

Maryem Rhanoui is an Associate Professor specializing in Computer Sciences and Data Engineering. She received an engineering degree in computer science then a Ph.D. degree from ENSIAS, Mohammed V University, Rabat 2015. Her research interests include artificial intelligence, knowledge extraction, and decision making, and medical data analysis. She can be contacted at email: mrhanoui@gmail.com.

Mounia Mikram is an Associate Professor of Computer Sciences and Mathematics at the School of Information Sciences, Rabat since 2010. She received her master degree from Mohammed V University Rabat (2003) and her Ph.D. degree from Mohammed V University, Rabat, and Bordeaux I University (2008). Her research interests include pattern recognition, computer vision, biometrics security systems, and artificial intelligence. She can be contacted at email: mmikram@esi.ac.ma.