Enhancing Computer-assisted Bone Fractures Diagnosis in Musculoskeletal Radiographs Based on Generative Adversarial Networks

Nabila Ounasser[1], Maryem Rhanoui[2], Mounia Mikram[3], Bouchra El Asri[4]
IMS Team, ADMIR Laboratory, ENSIAS, Mohammed V University in Rabat[1,4]
Meridian Team, LYRICA Laboratory, School of Information Sciences, Rabat, Morocco[2,3]

Abstract—**Computer-Assisted Bone Fractures Diagnosis in musculoskeletal radiographs plays a crucial role in aiding medical professionals in accurate and timely fracture detection. In this work, we explore a Generative Adversarial Network based approach for this task, which is a powerful deep learning model capable of generating realistic images and detecting anomalies. Our proposed approach leverages the potential of GANs to generate synthetic radiographs with fractures and identify anomalous patterns, thereby enhancing fracture diagnosis. Through extensive experimentation and evaluation on musculoskeletal radiograph datasets (MURA), we demonstrate the effectiveness of GAN-based models in improving fracture detection performance by adopting several evaluation metrics notably accuracy, precision, F1-score and detection speed. These findings highlight the potential of integrating GANs into computer-assisted diagnosis, contributing to the advancement of fracture diagnosis methodologies in orthopedics. It is important to note that GANs operate by training a generator network to produce synthetic images and a discriminator network to distinguish between real and generated images. This adversarial process fosters the generation of realistic radiographs with fractures, enabling accurate and automated detection. Our findings contribute to the advancement of fracture diagnosis methodologies and pave the way for more efficient and precise diagnostic tools in the field of orthopedics.**

Keywords*—*Deep learning; generative adversarial network; diagnosis; orthopedics; fracture detection; x-ray image

I. Introduction

Accurate and timely diagnosis of bone fractures is crucial in musculoskeletal radiology for effective patient care and treatment planning[1], [2], [3]. Conventional methods heavily rely on the expertise of radiologists [4], which can be subjective and time-consuming. However, recent advancements in artificial intelligence and deep learning techniques have paved the way for computer-assisted diagnosis systems that can aid radiologists in detecting and classifying fractures with improved accuracy [5].

In this article, we explore the application of generative adversarial networks (GANs) for computer-assisted bone fracture diagnosis in musculoskeletal radiographs. GANs are a class of deep learning models comprising two neural networks: a generator and a discriminator[6]. Its architecture give a promising future for anomaly detection in several fields[7], [8], [9]. The integration of GANs in fracture diagnosis presents several advantages. Firstly, it offers the potential to reduce radiologists' workload by automating the initial screening process. This allows radiologists to allocate more time and attention to complex cases, ultimately improving patient care. Secondly, GANs have shown promise in enhancing diagnostic accuracy by providing a reliable second opinion. By learning from a large dataset of medical images, GANs can capture subtle fracture features and aid radiologists in making more informed decisions. Moreover, GAN-based systems have the potential to contribute to the standardization of fracture diagnosis. By learning from diverse cases and incorporating a wide range of fracture patterns, these systems can help minimize inter-observer variability and increase diagnostic consistency across different healthcare settings. Furthermore, the continuous learning capabilities of GANs enable the system to adapt to new data and improve its diagnostic performance over time. Despite these promising advancements, several challenges must be addressed. Ensuring the robustness and generalizability of GAN-based fracture diagnosis systems across various patient populations, imaging modalities, and fracture types is essential. Furthermore, ethical considerations, data privacy, and regulatory frameworks need to be carefully considered to ensure the responsible and safe implementation of these technologies in clinical practice.

In this work, we will investigate how the integration of GANs in computer-assisted bone fracture diagnosis can hold a great potential in revolutionizing musculoskeletal radiology. By combining the expertise of radiologists with the power of deep learning, these systems can enhance diagnostic accuracy, streamline workflow, and contribute to standardized fracture diagnosis. As ongoing research and development in this field continue to unfold, it is expected that GAN-based computer-assisted diagnosis systems will play a pivotal role in improving fracture detection and patient care in the near future. To this end, we will explore MURA dataset [10], the largest public radiographic image datasets, to see what GANs may be drawn on orthopedic anomaly detection on X-ray images.

The contribution through this work can be summarized as follows:

- Efficient and accurate anomaly detection techniques based on deep learning to detect bone fractures.
- Applying several GANs models in one work to have a rich comparative study.
- Review the examined deep learning models, approaches and architectures.
- Identification of the most suitable data pre-processing techniques for our study, especially for the dataset em-

ployed. This analysis provides valuable insights into the optimal preprocessing steps required to improve the overall effectiveness of the models.

- Optimize the performance of the examined models to overcome the performance of recent and relevant works.
- Evaluates the effectiveness of the examined models using different evaluation metrics to have a comparison and analysis study based on their results.

The rest of the paper is designed as follows: In Section I we provide the essential background to understand the rest of the paper. Section II lays out a brief summary of the related work in the same field. Section III presents our paper's methodology. Section IV presents the experiment, materials used, dataset description and evaluation metrics for a promising comparative analysis study. Section V presents models' results and discusses the outcomes based on several evaluation metrics. Finally, the conclusion and the findings of this work are in Section VI.

II. Related Work

Fractures are highly accurate indicators of orthopedic pathology in most hospitals. However, analyzing medical images to identify bone fractures can be time-consuming and requires the expertise of qualified professionals. To address these challenges, scientists have been investigating ways to reduce diagnosis time and improve decision precision, aiming to assist doctors in their diagnostic processes [11], [12], [5], [13]. Several studies have demonstrated the potential of AI/DL in supporting medical professionals and decision-makers by developing automated tools that enhance the accuracy of physician interpretation [14], [5], [15], [16], [17], [18], [19], [20] and facilitate the creation of effective and cost-efficient treatment plans [15]. While many studies have focused on accurately detecting musculoskeletal abnormalities using CNN models, our research specifically explores the application of generative adversarial networks (GANs) [21] as a novel approach as shown in Table I. In this article [22], we propose the Res-UnetGAN network, an unsupervised anomaly detection approach based on GANs. The architecture combines the ResNet50 and UNet models within the generative network to calculate the normal distribution of samples. The discriminator employs a deep separable convolution-based convolutional neural network model to facilitate the adversarial training process. Anomaly identification is achieved by evaluating the reconstruction error score, which measures the quality of reconstruction and detects the presence of defects. Extensive testing on the Mura dataset demonstrates that our proposed method outperforms several other models in terms of defect detection accuracy.

In addition, [23] Davletshina et al. highlighted the value of unsupervised techniques trained on radiographic images without anomalies. Their approach aims to improve diagnostic accuracy and reduce the possibility of overlooking critical areas. By leveraging cutting-edge unsupervised learning techniques, they successfully identify anomalies and demonstrate the justifiability of the results.

Through our research, we aim to contribute to the advancement of fracture diagnosis by integrating the capabilities of GANs in framework will be used by radiographs. Hopefully this approach will have the potential to enhance the efficiency and accuracy of fracture diagnosis, enabling timely interventions and improved patient care.

III. Methods

In this section we will present our proposed framework and we will review GAN models investigated in this study.

A. Proposed Framework

In the proposed framework for fracture diagnosis as shown in Fig. 1, the process begins with the acquisition of the images using medical imaging devices; then will be integrated into computers. These medical images are then passed through a deep learning system that performs preprocessing tasks such as image enhancement, augmentation, rotation, and normalization. The preprocessed images are then fed into the fracture detection algorithm, which utilizes GAN models and CAM technique to identify and visualize fractures within the medical image. Once the detection process is complete, the system evaluates the results using various evaluation metrics such as accuracy, precision, F1-score, and detection speed. These metrics provide quantitative measures of the system's performance in correctly identifying fractures. Based on these metrics, a final decision is generated, indicating the presence or absence of fractures in the input image. This decision is crucial in assisting radiologists in making accurate diagnoses and treatment decisions. The framework aims to enhance the efficiency and time to diagnosis fracture detection, ultimately improving patient care in the field of musculoskeletal radiology.

B. Overview of GANs

Generative Adversarial Network (GAN) is a powerful deep learning framework that has gained considerable attention in recent years [6]. As in Fig. 2 GANs are composed of two main components: a generator and a discriminator. The concept behind GANs is to train the generator and discriminator in a competitive manner. Initially, the generator produces random samples, and the discriminator tries to correctly identify them as fake. As training progresses, both the generator and discriminator learn and improve their capabilities. The generator aims to generate samples that are increasingly difficult for the discriminator to differentiate from real data, while the discriminator continuously adapts to distinguish the real and generated samples accurately.

One of the significant advantages of GANs is their ability to generate new and realistic data that captures the underlying distribution of the training data. GANs have been widely used for various applications, including image synthesis, text generation, anomaly detection, medical diagnosis, and data augmentation [5], [8], [7], [24], [26], [27]. The generated samples can be used to enhance training datasets, generate novel and diverse content, or assist in data analysis tasks. Despite their remarkable capabilities, GANs come with their own challenges. Training GANs can be complex and prone to instability, often requiring careful hyperparameter tuning and architectural considerations. Issues like mode collapse, where the generator fails to explore the entire distribution of

TABLE I. SUMMARY OF RECENT WORKS OF FRACTURE DETECTION IN MUSCULOSKELETAL RADIOGRAPHS

[Ref](Year)	Dataset	Approach	Results
[24](2021)	MURA dataset	A proposed unsupervised anomaly detection method is the Res-UNetGAN Network. This method incorporates a GAN that merges ResNet50 and UNet architectures to create an autoencoder framework. This structure enables the system to learn distinctive characteristics from the input data.	Res-UnetGAN: 0.92 GANomaly: 0.81 Skip-GANomaly: 0.90 CVAE-GAN-Based: 0.86 EGBAD: 0.80
[23](2020)	MURA dataset	Comparative study between GAN and AE models on anomaly detection.	CAE: 0.57 VAE: 0.48 DC-GAN: 0.53 BiGAN: 0.54 AlphaGAN: 0.60
[25](2020)	MURA dataset	Computer Based Diagnosis (CBDs) model based on DenseNet201 and Inception V3 models, they were used to classify the given dataset as abnormal or normal.	DenseNet201: 87.15 InceptionV3: 86.11 Ensemble: 88.54

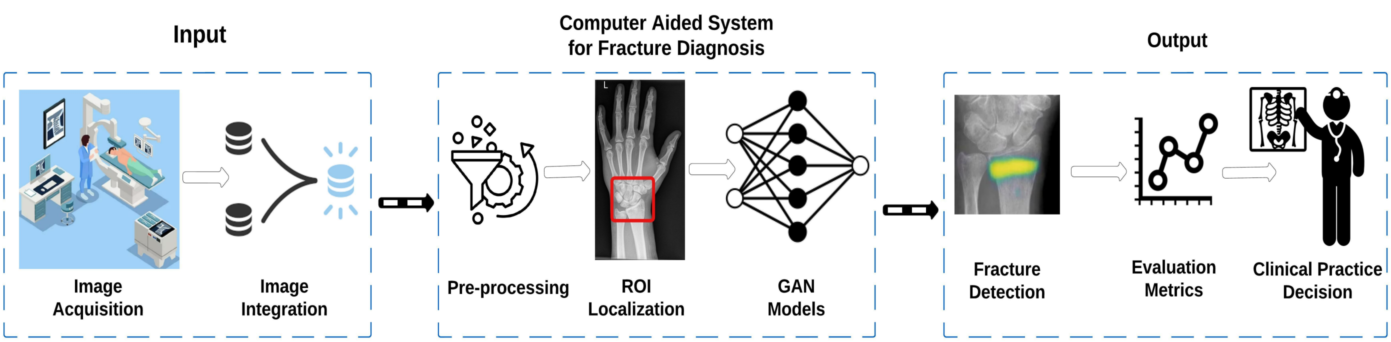

Fig. 1. Our framework for fracture diagnosis.

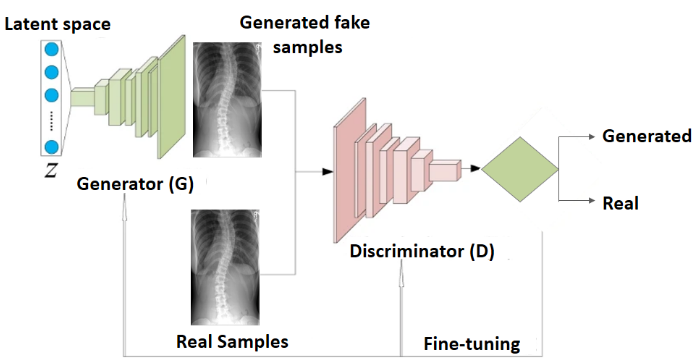

Fig. 2. Architecture of a generative adversarial network.

real data, can also arise. However, ongoing research aims to address these challenges and further improve the performance and stability of GANs.

To sum up, the ability of GANs to generate realistic and novel data has opened up new possibilities in various domains, including computer vision and data analysis. With continued advancements and research, GANs hold great promise for generating high-quality synthetic data and pushing the boundaries of generative modeling even further.

In the following Table II, we describe the different GAN models implemented in this work.

IV. EXPERIMENT

A. Experimentation Setup

The following subsection provide an overview of the dataset utilized in this study, including a description of its characteristics. The training settings and evaluation metrics employed.

1) Dataset description: The MURA dataset is a widely recognized and utilized dataset in the field of musculoskeletal imaging[10]. It comprises a large collection of radiographic images across different anatomical regions, including upper extremities, lower extremities, and the torso. The dataset focuses on various musculoskeletal conditions, especially fractures. The MURA dataset serves as a valuable resource for developing and evaluating algorithms and models in the domain of musculoskeletal radiography. Researchers and practitioners leverage this dataset to advance the field and improve diagnostic accuracy, automated diagnosis systems, and computer-assisted fracture detection techniques. The dataset is provide by the Stanford Program for Artificial Intelligence in Medicine: https://stanfordmlgroup.github.io/competitions/mura/.

2) Evaluation metrics: **Accuracy:** Accuracy is a widely used evaluation metric that measures the overall correctness of a fracture detection model [30]. It calculates the percentage

TABLE II. DESCRIPTION OF IMPLEMENTED MODELS

Model	Type	Description
[24] GANomaly	GAN-based anomaly detection model	Ganomaly combines GAN architecture with anomaly detection techniques to identify fractures in radiographs. It learns to generate normal images and detects anomalies based on reconstruction error.
[24] SkipGANomaly	GAN-based anomaly detection model	SkipGanomaly extends the GAN architecture by incorporating skip connections to improve the quality of reconstructed images.
[26] AnoGAN	GAN-based anomaly detection model	AnoGAN combines GAN architecture with unsupervised learning to detect anomalies in images.
[27] MadGAN	GAN-based anomaly detection model	MadGAN is a GAN architecture that utilizes multiple discriminators to enhance the detection of anomalies.
[28] AttentionGAN	GAN model image-to-image translation	AttentionGAN is a type of GAN that incorporates an attention mechanism to improve the quality of generated images. It selectively focuses on important regions, capturing fine details and producing realistic outputs.
[5] DCGAN	GAN model with deep convolutional layers	DCGAN is a foundational GAN model that employs deep convolutional layers for image generation. It can be utilized to generate synthetic radiographs with fractures for training or augmenting the dataset.
[5] CycleGAN	GAN model for image-to-image translation	CycleGAN is primarily used for domain adaptation and image translation tasks. Although not directly designed for fracture detection, it can potentially be employed to translate normal radiographs to fractured ones, facilitating the identification of fractures based on the translated images.
[29] SAGAN	GAN model with self-attention mechanism	SAGAN incorporates self-attention mechanisms to improve the quality and coherence of generated images.

of correctly identified fractures out of all the samples in the dataset.

$$Accuracy = \frac{TruePositives + TrueNegatives}{TotalSamples}$$

Precision: Precision is a metric that focuses on the positive predictions made by the fracture detection model[30]. It measures the proportion of correctly identified fractures out of all the predicted fractures. Precision helps assess the model's ability to minimize false positives, indicating how reliable the model is when it identifies a sample as a fracture.

$$Precision = \frac{TruePositives}{TruePositives + FalsePositives}$$

F1-Score: The F1-Score is a combined metric that takes into account both precision and recall, providing a balanced measure of the fracture detection model's performance[30]. It considers the trade-off between false positives and false negatives. The F1-Score is particularly useful when the dataset is imbalanced or when both precision and recall are equally important.

$$F1_{Score} = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Detection Speed: Detection speed is an essential evaluation aspect that measures the efficiency of a fracture detection model in processing and analyzing musculoskeletal radiographs[30]. It quantifies the time taken by the model to detect fractures in a given dataset or per image. Faster detection speeds are desirable, particularly in clinical settings where time is of the essence.

By considering these evaluation metrics - accuracy, precision, F1-Score, and detection speed - we can thoroughly assess the performance and efficiency of fracture detection models. These metrics provide valuable insights into the model's ability to correctly identify fractures, minimize false positives, achieve a balance between precision and recall, and process radiographic images efficiently.

B. Preprocessing

The MURA dataset, presents a challenge due to its diverse range of images of bone abnormalities with different formats and sizes. To address this issue and make the data more uniform, we employed image pre-processing techniques to enhance the image quality. Initially, we applied binary thresholding to identify the Region of Interest (RoI) within the image and extract its contours. This process enabled us to isolate the relevant region for classification and crop it accordingly.

Data augmentation played a vital role in our research, helping to expand the dataset and improve the learning algorithm. Several augmentation approaches were employed, including horizontal image flipping, random rotation within a range of 30 degrees, scaling within the range of 95-130 percent, and randomly adjusting brightness within the range of 80-120 percent.

Prior to initiating the learning process, radiographs were normalized to have the same mean and standard deviation as the images in the ImageNet training set. This normalization step ensured consistency and facilitated subsequent stages of the project.

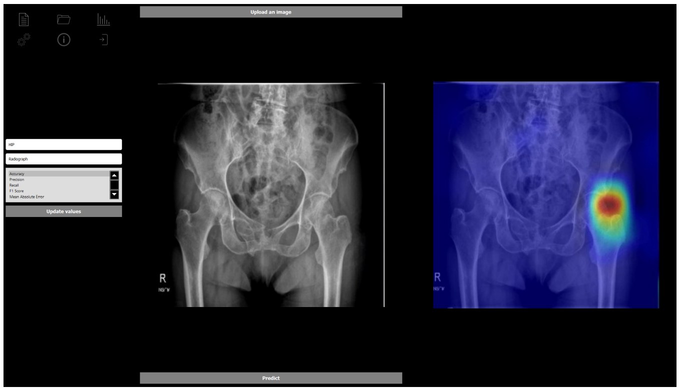

Fig. 3. Fracture diagnosis results: Hip.

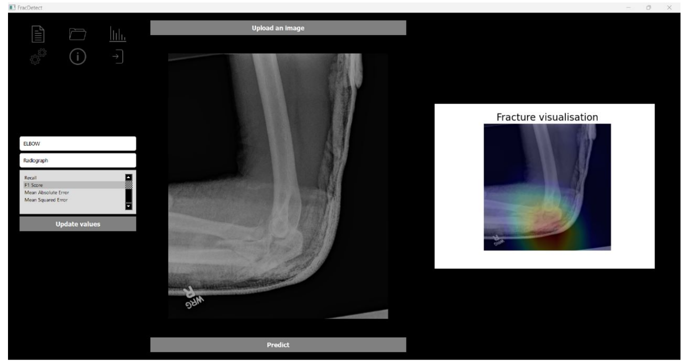

Fig. 4. Fracture diagnosis results: Elbow.

Following data augmentation and normalization contributed to enhancing the quality, consistency, and effectiveness of the dataset for our research purposes.

V. Results and Discussion

In this section, we discuss the performance of GAN models for bone fracture detection in radiographs, focusing on key evaluation metrics such as accuracy, precision, F1-score, and detection speed. These metrics provide valuable insights into the effectiveness and efficiency of the models in identifying fractures and aiding in clinical decision-making. Various factors influence the performance of the models, including architectural approach, layer design, padding, shape, normalization, activation, loss function, optimizer, batch size, learning rate, pooling, and output layer. Consequently, achieving an effective result required multiple tuning iterations. Many of our models comprised computationally expensive layers and modules, necessitating long training durations that were often impractical on basic hardware or laptop configurations. Preprocessing played a crucial role in obtaining good results in our deep learning tasks. After selecting the appropriate GAN models for the study based on a benchmark between the different GAN model, extensive data preparation was necessary. The image size proved to be a significant parameter impacting the accuracy of detecting fractures. To this end, several treatments were involved on our dataset as mentioned in preprocessing section. Also, to overcome the limitations of the available data, we employed data augmentation techniques, which augmented the amount of data during the training phase. However, we had to be cautious regarding the rotation methods, excessive compression, and shear, as they could negatively affect the performance of bone fracture diagnosis. By considering these factors and conducting thorough experimentation, we aimed to optimize our models and enhance the accuracy of our results. It is evident that preprocessing, data augmentation, data normalization and careful parameter selection are vital considerations when striving for accurate and reliable outcomes in deep learning tasks related to bone fracture detection. Overall, as shown in Fig. 3 and 4 after selecting the input image we configure our framework by choosing body part treated, image modality and evaluation metrics that the user want to display. We see the visualized results by Grad-CAM and obtained from our implemented GAN models for bone fracture detection in radiographs were promising. The models demonstrated a high level of accuracy, achieving performance within the range of 0.7% to 0.954%. These results indicate that the GAN-based approach holds considerable potential for improving fracture diagnosis in the field of orthopedics.

Comparing our GAN models with previous works, it is evident that they compare favorably in terms of performance. The accuracy achieved by our models aligns with some results reported in works with the same task with different techniques and surpasses others results reported in relevant literature. This suggests that the utilization of GANs for bone fracture diagnosis can yield significant improvements in diagnostic accuracy and support medical professionals in their decision-making processes.

Detailly, Table III shows that MadGAN, CycleGAN, SAGAN and SkipGANomaly on average achieves the best accuracy (0.954%; 0.922%; 0.9%;0.901%). These accuracy scores indicate that the models were successful in correctly classifying fractures in radiographs, contributing to improved diagnostic capabilities. While with less processing time MadGAN and SAGAN have speed detection higher than other models. So we can consider MagGAN and SAGAN are the most powerful models. This is due to their architectures. MadGAN has the incorporation of multiple discriminators that can enhance the detection and identification of fractures in radiographs as anomalous patterns. SAGAN use self-attention mechanisms can improve the coherence and quality of generated images. Although not specifically tailored for fracture detection, its ability to produce visually consistent radiographs was the main cause to have an accurate fracture identification. Then we had GANomaly, AttentionGAN and AnoGAN achieved an accuracy of 0.861% and 0.842%. Those are reputable results even those models were not explicitly designed for fracture detection, but it could give a better performance in anomaly detection in other fields and with other parameters notably type of data, size of dataset, experiment setting, etc. For DCGAN, its performance in fracture detection be less compared to the models explicitly designed for anomaly detection or fracture identification. Because its architecture is not explicitly designed for fracture detection also due to the high number of convolutional layers that take more time to give the predictive result.

One notable advantage of GAN models is their ability to generate synthetic data, which can be beneficial in addressing

TABLE III. Comparison of Models based on Performance Metrics

Models	Accuracy	Precision	F1-Score	Detection Speed
Ganomaly	0.861	0.873	2.863	2.987
SkipGanomaly	0.901	0.903	0.907	5.837
AttentionGAN	0.842	0.877	0.8418	2.291
MadGAN	0.954	0.958	0.953	1,418
SAGAN	0.9	0.905	0.907	0,2
AnoGAN	0.838	0.848	0.842	3.738
DCGAN	0.7	0.724	0.705	15.233
CycleGAN	0.932	0.917	0.928	8.412

the issue of limited labeled datasets. By leveraging the generative capabilities of GANs, it becomes possible to augment the available data and improve the robustness and generalization of the models. This can be particularly valuable in medical imaging, where acquiring large annotated datasets can be challenging. Despite the promising results, it is important to acknowledge the limitations and challenges associated with GAN models for fracture detection. One key consideration is the computational complexity and resource requirements of training and deploying GAN models. The training process of GANs can be computationally intensive and time-consuming, necessitating powerful hardware and substantial computational resources. This can pose practical limitations, especially in clinical settings where quick and efficient diagnosis is crucial.

To sum up, our study demonstrates the potential of GAN models in detecting bone fractures in radiographs. The achieved accuracy and performance indicate that GANs can serve as valuable tools for assisting medical professionals in fracture diagnosis. However, challenges related to computational complexity, data availability, and interpretability need to be addressed for broader adoption and real-world application. Future research should focus on optimizing GAN architectures, addressing dataset limitations to further enhance the performance and practicality of GAN models in fracture detection.

VI. Conclusion

In conclusion, this article has presented an exploration of Computer-assisted Bone Fractures Diagnosis in musculoskeletal radiographs using Generative Adversarial Networks (GANs). The use of GANs in medical image analysis has shown great potential in aiding clinicians and radiologists in the accurate and efficient detection of bone fractures. By leveraging the power of GANs, we have demonstrated the ability to generate realistic radiographs with fractures, detect anomalies, and improve the overall diagnostic process. Through our research, we have observed promising results in utilizing GAN-based models such as MadGAN, CycleGAN, SkipGanomaly, and SAGAN for fracture detection. These models have demonstrated their effectiveness in generating high-quality images, identifying anomalies, and translating normal radiographs to fractured ones. The performance of these models, although influenced by various factors such as dataset size, training configuration, and preprocessing techniques, has shown significant potential in enhancing fracture diagnosis accuracy and reducing the reliance on manual interpretation. However, it is important to acknowledge the challenges that lie ahead. Further research and development are required to address limitations such as data heterogeneity, model generalization, and interpretability. Additionally, the ethical implications, including patient privacy and the need for human oversight in the diagnostic process, should be carefully considered and addressed. In conclusion, the application of Generative Adversarial Networks for Computer-Assisted Bone Fractures Diagnosis in musculoskeletal radiographs holds great promise. By harnessing the power of GANs, we can improve the accuracy, efficiency, and overall quality of fracture detection, ultimately benefiting both clinicians and patients. Continued advancements in this field have the potential to revolutionize musculoskeletal radiography and pave the way for more effective and precise diagnostic tools in the future.

Acknowledgment

I would like to express my sincere gratitude to my supervisors and colleagues whose invaluable contributions and support have greatly enriched this article. Their guidance and insights have been instrumental in shaping its content and direction.

References

[1] P. H. Kalmet, S. Sanduleanu, S. Primakov, G. Wu, A. Jochems, T. Refaee, A. Ibrahim, L. v. Hulst, P. Lambin, and M. Poeze, "Deep learning in fracture detection: a narrative review," *Acta orthopaedica*, vol. 91, no. 2, pp. 215–220, 2020.

[2] A. D. Woolf and B. Pfleger, "Burden of major musculoskeletal conditions," *Bulletin of the world health organization*, vol. 81, pp. 646–656, 2003.

[3] 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," *AJR. American journal of roentgenology*, vol. 213, no. 3, p. 506, 2019.

[4] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights into imaging*, vol. 9, no. 4, pp. 611–629, 2018.

[5] Y. Shin, J. Yang, and Y. H. Lee, "Deep generative adversarial networks: applications in musculoskeletal imaging," *Radiology: Artificial Intelligence*, vol. 3, no. 3, p. e200157, 2021.

[6] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.

[7] N. Ounasser, M. Rhanoui, M. Mikram, and B. E. Asri, "Generative and autoencoder models for large-scale mutivariate unsupervised anomaly detection," in *Networking, Intelligent Systems and Security: Proceedings of NISS 2021*. Springer, 2021, pp. 45–58.

[8] ——, "Anomaly detection in orthopedic musculoskeletal radiographs using deep learning," in *International Conference on Computing and Communication Networks 2021 (ICCCN 2021)*. Springer, 2022.

[9] L. Yao, X. Guan, X. Song, Y. Tan, C. Wang, C. Jin, M. Chen, H. Wang, and M. Zhang, "Rib fracture detection system based on deep learning," *Scientific Reports*, vol. 11, no. 1, p. 23513, 2021.

[10] P. Rajpurkar, J. Irvin, A. Bagul, D. Ding, T. Duan, H. Mehta, B. Yang, K. Zhu, D. Laird, R. L. Ball *et al.*, "Mura: Large dataset for abnormality detection in musculoskeletal radiographs," *arXiv preprint arXiv:1712.06957*, 2017.

[11] G. Moon, S. Kim, W. Kim, Y. Kim, Y. Jeong, and H.-S. Choi, "Computer aided facial bone fracture diagnosis (ca-fbfd) system based on object detection model," *IEEE Access*, vol. 10, pp. 79 061–79 070, 2022.

[12] L. Sathish Kumar, A. Prabu, V. Pandimurugan, S. Rajasoundaran, P. P. Malla, and S. Routray, "A comparative experimental analysis and deep evaluation practices on human bone fracture detection using x-ray images," *Concurrency and Computation: Practice and Experience*, vol. 34, no. 26, p. e7307, 2022.

[13] M. Wu, Z. Chai, G. Qian, H. Lin, Q. Wang, L. Wang, and H. Chen, "Development and evaluation of a deep learning algorithm for rib segmentation and fracture detection from multicenter chest ct images," *Radiology: Artificial Intelligence*, vol. 3, no. 5, p. e200248, 2021.

[14] L. Jin, J. Yang, K. Kuang, B. Ni, Y. Gao, Y. Sun, P. Gao, W. Ma, M. Tan, H. Kang *et al.*, "Deep-learning-assisted detection and segmentation of rib fractures from ct scans: Development and validation of fracnet," *EBioMedicine*, vol. 62, p. 103106, 2020.

[15] G. Mehr, "Automating abnormality detection in musculoskeletal radiographs through deep learning," *arXiv preprint arXiv:2010.12030*, 2020.

[16] L.-W. Cheng, H.-H. Chou, K.-Y. Huang, C.-C. Hsieh, P.-L. Chu, and S.-Y. Hsieh, "Automated diagnosis of vertebral fractures using radiographs and machine learning," in *International Conference on Intelligent Computing*. Springer, 2022, pp. 726–738.

[17] S. Mutasa, S. Varada, A. Goel, T. T. Wong, and M. J. Rasiej, "Advanced deep learning techniques applied to automated femoral neck fracture detection and classification," *Journal of Digital Imaging*, vol. 33, pp. 1209–1217, 2020.

[18] B. Zhang, C. Jia, R. Wu, B. Lv, B. Li, F. Li, G. Du, Z. Sun, and X. Li, "Improving rib fracture detection accuracy and reading efficiency with deep learning-based detection software: a clinical evaluation," *The British Journal of Radiology*, vol. 94, no. 1118, p. 20200870, 2021.

[19] B. Guan, G. Zhang, J. Yao, X. Wang, and M. Wang, "Arm fracture detection in x-rays based on improved deep convolutional neural network," *Computers & Electrical Engineering*, vol. 81, p. 106530, 2020.

[20] D. Kim and T. MacKinnon, "Artificial intelligence in fracture detection: transfer learning from deep convolutional neural networks," *Clinical radiology*, vol. 73, no. 5, pp. 439–445, 2018.

[21] A. Spahr, B. Bozorgtabar, and J.-P. Thiran, "Self-taught semi-supervised anomaly detection on upper limb x-rays," in *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2021, pp. 1632–1636.

[22] 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. 49 920–49 928, 2021.

[23] D. Davletshina, V. Melnychuk, V. Tran, H. Singla, M. Berrendorf, E. Faerman, M. Fromm, and M. Schubert, "Unsupervised anomaly detection for x-ray images," *arXiv preprint arXiv:2001.10883*, 2020.

[24] S. Song, A. Yang, Kechengand Wang, S. Zhang, and M. Xia, "A mura detection model based on unsupervised adversarial learning," *IEEE Transactions on Knowledge and Data Engineering*, 2021.

[25] N. K. Namit Chawla, "Automating abnormality detection in musculoskeletal radiographs through deep learning," *RIA - Revue d'Intelligence Artificielle - IIETA*, 2020.

[26] T. Schlegl, P. Seeböck, S. M. Waldstein, G. Langs, and U. Schmidt-Erfurth, "f-anogan: Fast unsupervised anomaly detection with generative adversarial networks," *Medical image analysis*, vol. 54, pp. 30–44, 2019.

[27] D. Li, D. Chen, B. Jin, L. Shi, J. Goh, and S.-K. Ng, "Mad-gan: Multivariate anomaly detection for time series data with generative adversarial networks," in *Artificial Neural Networks and Machine Learning–ICANN 2019: Text and Time Series: 28th International Conference on Artificial Neural Networks, Munich, Germany, September 17–19, 2019, Proceedings, Part IV*. Springer, 2019, pp. 703–716.

[28] H. Tang, H. Liu, D. Xu, P. H. Torr, and N. Sebe, "Attentiongan: Unpaired image-to-image translation using attention-guided generative adversarial networks," *IEEE transactions on neural networks and learning systems*, 2021.

[29] H. Zhang, I. Goodfellow, D. Metaxas, and A. Odena, "Self-attention generative adversarial networks," in *International conference on machine learning*. PMLR, 2019, pp. 7354–7363.

[30] D. Joshi and T. P. Singh, "A survey of fracture detection techniques in bone x-ray images," *Artificial Intelligence Review*, vol. 53, no. 6, pp. 4475–4517, 2020.