Machine Learning Approach for Detect Anomalies in Medical Images

Zinga Banda Firmino René

[1] Universidade Eötvös Loránd, Egyetem tér 1-3, 1053, Budapest - Hungary
[2] (06 1) 411 6500

Abstract. Anomaly detection in medical imaging is crucial for recognizing anomalies that depart from predicted patterns, which aids in early diagnosis and therapy. The detection of anomalies in medical images is widely used in several specialties, allowing for faster and more accurate diagnoses. CNN and GANs are used in pneumology to identify pneumonia, Covid-19, and pulmonary nodules on X-rays and CT scans. In cardiology, images such as electrocardiograms and cardiac resonances are analyzed to identify coronary diseases, cardiac insufficiency, and arrhythmia. This study presents a new approach to evaluate the performance of supervised anomaly detection techniques specifically for chest X-ray medical images, leveraging validation metrics widely used in clustering algorithms associated with generative model GANs. The experiments were conducted on datasets such as the NIH14 Chest X-ray and Chest X-ray (Pneumonia) images, which include extensive collections of chest X-rays used to detect medical anomalies such as pneumonia and other abnormalities. By generating high-quality synthetic images, the results demonstrate the effectiveness and reliability of the proposed metrics in evaluating the performance of anomaly detection, highlighting their potential to improve the quality and consistency of automated diagnostic systems in medical images.

Keywords: Medical imaging · Abnormalities detection · Machine learning · Supervised learning · Algorithm · Evaluation · Generative Models

1 Introduction

X-ray radiography allows for non-destructive inspection of the internal structure of an object. Industrial radiography-based inspection is an anomaly detection task, which aims at distinguishing between defective and non-defective samples. Identifying possible defective objects solely from the observation of the object radiographs can be a hard task and is often defect-dependent. [20]

Anomaly detection has been employed in the following medical areas:

Brain Imaging: Tumors, lesions, cerebrovascular disorders, and metastases can be identified using anomaly detection in MRI images. Furthermore, fMRI and rs-fMRI approaches are used to identify visual artifacts and cognitive alterations.

Mammograms: Mammogram images are used to diagnose breast cancer and other abnormalities, such as compressions or implants. Ophthalmologists employ fundus photography and optical coherence tomography (OCT) to detect retinal lesions and disorders like glaucoma and macular degeneration.

Histopathology: The examination of histological pictures enables the detection of malignancies in biological tissues.

Breast Ultrasound: The discovery of irregularities during a breast ultrasound aids in distinguishing between normal, benign, and cancerous breast tissue.

Chest X-rays: are used to diagnose pneumonia, lung disorders (including COVID-19), and pleural effusions. [33]. Anomaly detection is also utilized. One example is patient monitoring, where electrocardiography (ECG) signals or other body sensors are used to detect critical, possibly life-threatening situations [16].

In this project, Chest X-ray (CXR) was chosen as the primary dataset due to its relevance in the diagnosis of lung diseases such as pneumonia, COVID-19, pleural effusions, and other respiratory anomalies. Chest X-rays are widely used in medical applications due to their accessibility, low cost, and ability to provide detailed information about the internal structure of the chest. Furthermore, the use of CXR allows for a non-invasive approach to patient inspection and monitoring, making it a valuable resource for anomaly detection tasks in clinical settings.

The selection of CXR also aligns with the project's goal of employing anomaly detection techniques to identify unusual patterns that may indicate the presence of disease. Compared to other imaging modalities such as magnetic resonance imaging (MRI) or computed tomography (CT), chest X-rays offer simplicity and scalability, making them widely adopted in hospitals and clinics worldwide. Thus, its inclusion in this project maximizes the practical applicability of the results, while exploring the potential of advanced machine learning models to overcome the challenges related to identifying anomalous patterns in medical images.

2 RELATED WORK

2.1 Overview of Anomalies

The main idea of supervised anomaly detection algorithms is to detect data instances in a dataset, which deviate from the norm. However, there are a variety of cases in practice where this basic assumption is ambiguous. 5 illustrates some of these cases using a simple two dimensional dataset. Two anomalies can be easily identified by eye: x1 and x2 are very different from the dense areas with respect to their attributes and are therefore called global anomalies. When you are looking at the dataset globally, x3 can be seen as a normal record since it is not too far away from the cluster c2. However, when we focus only on cluster c2 and compare it with x3 while neglecting all the other instances, it can be seen as an anomaly. Therefore, x3 is called a local anomaly, since it is only anomalous when compared with its close-by neighborhood. It depends on the application, whether local anomalies are of interest or not. Another interesting question is whether the instances of the cluster c3 should be seen as three anomalies or as a (small) regular cluster. These phenomena is called micro cluster and anomaly detection algorithms should assign scores to its members larger than the normal instances, but smaller values than the obvious anomalies. This simple example already illustrates that anomalies are not always obvious and a score is much more useful than a binary label assignment. [6]

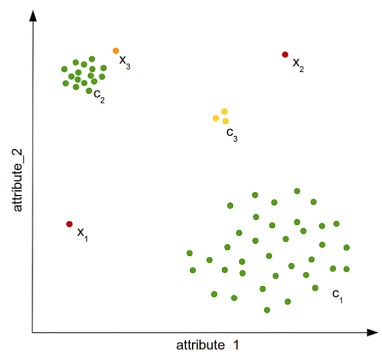

Fig. 1. Anomalies in 2D environment

2.2 Overview of Medical Imaging Modalities

Medical Image Modalities: There are many ways to acquire medical photographs, each with their own benefits and drawbacks. X-rays are used in CT scans to create precise images of the inside organs and tissues of the body. X-rays are suitable for regular diagnosis of a variety of disorders such as fractures, tumors, and lung ailments since they are rapid and non-invasive. [34] A powerful magnetic field and radio waves are used in MRI to provide precise images of the body's soft tissues. MRIs are more thorough than CT scans and are frequently used to identify tumors, joint issues, and cancers [24]. Using high-frequency sound waves, ultrasound can provide images of the internal organs of the body. Monitoring fetal development, identifying gallbladder and liver conditions, and assisting with biopsies can all be done with the safe, non-invasive use of ultrasound. The best modality to choose will depend on the clinical question at hand as well as the patient's health. Each modality has its advantages and disadvantages. [27]

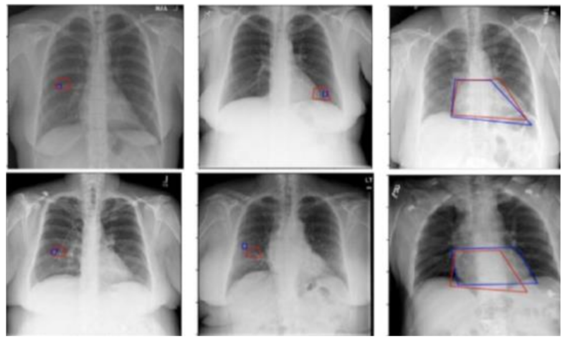

Fig. 2. Medical image of Lungs

Machine Learning Algorithms for Medical Image Analysis: By automating the process of finding patterns and anomalies in medical images, machine learning algorithms play a crucial role in image analysis. Supervised, unsupervised, and deep learning algorithms are three broad categories of machine learning algorithms. Support vector machines and random forests are two examples of supervised learning techniques that employ labelled training data to create a model that can classify fresh images into established categories [7]. Unsupervised learning methods, including clustering and principal component analysis, find patterns and structures in the data instead of using labelled training data. By enabling automated feature extraction and delivering cutting-edge performance on a variety of medical image analysis tasks, deep learning techniques, such as convolutional neural networks and recurrent neural networks, have revolutionized the field of medical image analysis. [27]

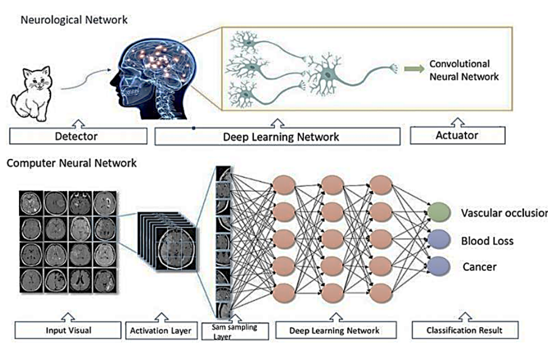

Fig. 3. Medical image of Lungs

Supervised Learning Algorithms: Algorithms for supervised learning are frequently employed in classification and segmentation tasks in medical image analysis. For these algorithms to develop a model that can identify or segment fresh images, training data must be labelled. Support vector machines (SVM), random forests, and artificial neural networks are examples of popular supervised learning techniques used in medical image processing. While random forests are frequently utilized for multi-class classification problems like tissue categorization, SVMs are frequently employed for binary classification tasks like tumor detection [15] . Convolutional and recurrent neural networks, two types of artificial neural networks, have attained cutting-edge performance on a variety of medical image processing tasks such image classification, segmentation, and registration.

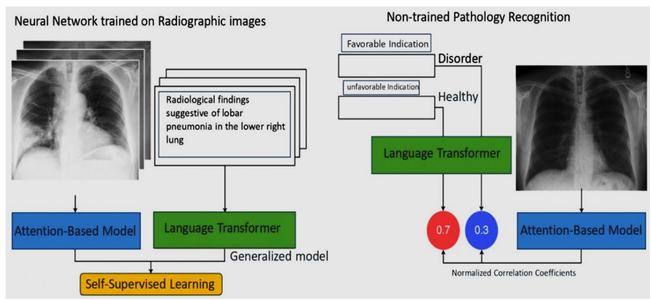

Fig. 4. Medical image of Lungs

Deep Learning Algorithms: By enabling automated feature extraction and attaining cutting edge performance on a variety of medical image analysis tasks, deep learning algorithms have completely changed the field of medical image analysis. In comparison to conventional machine learning algorithms, deep learning algorithms, such as CNNs and RNNs, perform better and can automatically learn hierarchical features from raw medical pictures [32]. In medical image analysis, CNNs are frequently employed for tasks like image classification, segmentation, and registration [32]. RNNs are employed in jobs involving sequence data, such as time-series data and medical reports, in medical image analysis.

Applications of Machine Learning Algorithms in Medical Image Analysis: Image segmentation, classification, registration, and reconstruction are just a few of the medical image analysis tasks for which machine learning techniques have been applied. Early-stage cancer detection, the identification of brain tumors, the detection of lung nodules, the diagnosis of bone fractures, and the monitoring of fetal development have all been accomplished using machine learning algorithms [22]. The quantification of illness development, the prediction of therapeutic outcomes, and the evaluation of therapeutic response have all been accomplished using machine learning algorithms.

Challenges in Machine Learning Algorithms in Medical Image Analysis: Despite the enormous advancements made in machine learning algorithms for medical image interpretation, there are still a number of issues that need to be resolved. Lack of extensive annotated datasets, which restricts the development and assessment of machine learning algorithms, is one of the key issues. The interpretability of deep learning algorithms is another issue that prevents their use in therapeutic contexts [29]. To ensure that machine learning algorithms work well across a range of patient populations and imaging modalities, robustness and generalizability must also be increased.

By enabling automated picture analysis and increasing diagnosis accuracy, machine learning algorithms have shown considerable promise in the field of medical image analysis. Due to the accessibility of big medical image datasets and the advancements in deep learning algorithms, the use of machine learning techniques in medical image analysis has risen quickly in recent years [23]. Despite the tremendous advancements gained, a number of issues still need to be resolved in order to guarantee the secure and efficient application of machine learning algorithms in clinical settings.

Types of Medical Images: A wide variety of medical disorders can be diagnosed and treated with the use of medical imaging. A number of modalities are used in medical imaging to take pictures of the body's internal organs, tissues, and structures. These pictures can be used to detect anomalies and make medical diagnoses [5]. Medical imaging can take many various forms, each with specific features and advantages, such as CT scans, X-rays, and MRIs.

CT Scans: X-rays are used in Computed Tomography (CT) scans to produce finely detailed pictures of the inside organs. The detection of bone fractures and other injuries, as well as the detection of anomalies in soft tissues and organs including the liver, lungs, and brain, are among the many uses for CT scans. CT scans can also be used to keep track of how well diseases like cancer and heart disease are being treated. The fact that CT scans are non-invasive and deliver high-quality images with a very quick scanning period is one of their advantages. However, ionizing radiation, which can be dangerous in large levels, is exposed to patients during CT scans [30]. As a result, it's crucial to carefully consider the advantages and disadvantages of CT scans for each patient.

X-rays: X-rays employ electromagnetic radiation to produce images of the inside organs and tissues of the body. The most frequent applications of X-rays are the detection of lung and chest abnormalities, as well as the detection of bone fractures and other traumas. X-rays are frequently used in dentistry to detect tooth decay and other problems with the mouth's health. X-rays are rapid and non-invasive, just like CT scans. Ionizing radiation, which can be dangerous at high doses, is further exposed to patients during X-rays [35]. As a result, it's crucial to utilize X-rays sparingly and to carefully weigh the advantages and disadvantages for each patient.

MRI's: A powerful magnetic field and radio waves are used in magnetic resonance imaging (MRI) to provide precise pictures of the inside organs of the body. The brain, liver, and heart are examples of soft tissues and organs where MRIs are very helpful in detecting problems. MRIs can also be used to find injuries, tumors, and other types of illnesses. The fact that MRIs don't subject patients to ionizing radiation makes them a safer option for individuals who may need several scans. MRIs cost more money and take longer to complete than other kinds of medical pictures, though. Additionally, due to safety concerns, people with certain medical disorders, such as pacemakers, may not be allowed to have MRIs. In conclusion, medical imaging is essential for the detection and management of a variety of medical problems. Three of the most popular medical picture kinds are CT scans, X-rays, and MRIs, each with special characteristics and advantages [26]. Healthcare professionals should carefully weigh the advantages and disadvantages of each form of medical imaging before selecting the best modality for each patient.

2.3 Limitation in Medical Image

Medical imaging is indispensable in modern health care, guiding diagnostics, surgeries, treatment assessments, and disease monitoring. The growing volume of images poses challenges [2] for radiologists and physicians to maintain workflow efficiency without technological support. There are significant challenges to train accurate and reliable Machine Learning or Deep Learning diagnostic models. Key issues include the scarcity of extensive and diverse datasets [3], stringent data privacy regulations, and inherent dataset imbalances. These imbalances can lead to biased models that struggle with rare conditions, and even minor errors can have negative implications.
Traditional data augmentation techniques, such as random rotations, flipping, cropping, and noise injection, have been widely employed to expand training datasets. While useful, these methods merely manipulate existing samples and fail to introduce the kind of fundamental variability needed for robust model training [13]. In contrast, generative models [4], GANs and DDPM, have revolutionized image synthesis by creating entirely new data points. These models offer promising solutions to the challenges of imbalanced datasets, particularly in the field of medical imaging, where the availability of labeled data is limited. Generative models typically require large and diverse datasets [1]. This creates a paradox: if such large labeled datasets were available, effective models could be trained directly. Therefore, generative models are only viable if they can work effectively with scarce datasets. The present paper addresses this challenge by proposing a comprehensive framework for generating synthetic medical images from both small and imbalanced datasets using two generative models: PGGANs [11] and DDPM [9].
GANs are prominent implicit density models that consist of two competing neural networks, a generator, that creates synthetic images from a latent space, and a discriminator which evaluates resemblance of generated images to real images, engaging in a zero sum game. Generally, it is hard to train GANs due to training instability [31]. In medical imaging, GANs mainly have been used to enhance classification and segmentation deep learning models [12]. The work in [3] uses GANs on a small CT scan dataset to generate eye fundus images which confirm to the given masks. [17] also used mask to generate lung images, and only those synthetic images that fulfilled informativeness criteria calculated

by Bayesian neural networks were used to improve the classifier model. In [4], GANs are employed to synthesize high quality focal liver lesions of multiple conditions to enhance a CNN-classifier. Moreover, GANs have been successful at synthesizing prostate lesions [14].

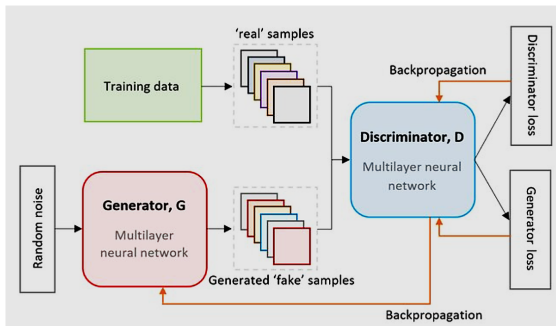

Fig. 5. GANs architecture

3 Literature review

The paper *"Deep Learning for Pneumonia Detection in Chest X-ray Images: A Comprehensive Survey"* is a study of the analysis and use of deep learning (DL) techniques to detect pneumonia in chest radiographs (CXR), focusing on work between 2012 and 2023. The datasets include sets such as NIH Chest X-ray Dataset and RSNA Pneumonia Dataset. Convolutional neural networks (CNNs), transfer learning, and hybrid models were used. The evaluation metrics were accuracy, sensitivity, specificity, F1-score, and AUC-ROC. Transformer-based models, such as Vision Transformers (ViTs), have shown significant promise but suffer from limitations, including class imbalance, lack of explainability, and vulnerability to adversarial attacks. [10].

The study *"Deep Learning Applications to Detect Pneumonia on Chest X-ray: A Systematic Study"* explores different deep learning architectures to detect pneumonia on CXRs, using databases such as ChestX-ray14 and CheXpert. Methods such as AlexNet, ResNet, DenseNet, and VGG were evaluated. Metrics include accuracy (up to 98%), sensitivity, specificity, and F1-score. Data augmentation techniques were used extensively to deal with small datasets. The paper highlights that CNNs are widely applied due to their efficiency in feature extraction and robustness in classifications. Key gaps include the need for larger datasets and better integration of data augmentation methods for higher accuracy. [25]

In the scientific article *"Diagnostic Performance of a Deep Learning Model Deployed at a National COVID-19 Screening Facility for Detection of Pneumo-*

nia on Frontal Chest Radiographs (2022)", the study evaluated a deep learning model for detecting pneumonia on chest radiographs (CXR) in a national COVID-19 screening center. Using the public ChestX-ray14 database and an in-house set of 4,277 images for fine-tuning, the DenseNet121 model was refined with transfer learning, data augmentation techniques, and an ensemble of seven models to improve performance in a clinical setting. The model achieved an AUC of 0.95, specificity of 97%, and sensitivity of 79%, reducing the time to generate medical reports by 22%. Although effective, the study highlighted limitations such as class imbalance, lack of data from multiple institutions, and the absence of integration with additional clinical information, indicating the need for future research to improve the generalizability and robustness of the model. [28]

"Automated Pneumonia Detection Using Deep Features in Chest X-ray Images (2023)" This paper proposes a model for automatic detection of pneumonia in chest radiographs using the Kermany dataset with 5,856 images, applying preprocessing with Histogram Equalization (HE) and CLAHE. The approach consists of a two-stream CNN architecture for feature extraction, followed by classifiers such as KNN, SVM and LDA, with KNN achieving the best accuracy of 97.86%. The study used metrics such as precision, sensitivity and F1-score to evaluate the performance. Despite the promising results, challenges include the scarcity of images in some classes and the dependence on a single dataset, indicating the need for future studies to improve generalization and address pneumonia subtypes. [18]

Pneumonia Detection from Chest X-rays Using the CheXNet Deep Learning Algorithm (2024): The study used the CheXNet chest radiograph dataset, consisting of 5,863 images labeled as “Normal” and “Pneumonia,” and divided into training, validation, and testing. The CheXNet model, based on the DenseNet121 architecture pre-trained on ImageNet, was implemented and fine-tuned to detect pneumonia in X-ray images. Preprocessing included rescaling and normalization, while training was conducted with Adam optimization and the binary cross-entropy loss function. Evaluation metrics include accuracy, sensitivity, F1-score, and AUC-ROC, with the model achieving an accuracy of 91% on the test set (624 images), accuracy of 0.90, sensitivity of 0.96, and F1-score of 0.93 for pneumonia cases. Despite the promising performance, limitations include the scarcity of diverse data and the need for larger datasets to improve generalizability and reduce bias. The study recommends future investigations to integrate broader data and explore advanced architectures to improve diagnostic reliability in clinical settings. [21]

"Proposed Convolutional Neural Network for Early Pneumonia Diagnosis (2023)" The study proposed a 22-layer CNN to detect pneumonia in chest radiographs (CXR), using a Kaggle dataset composed of 5,863 images, divided into "Normal" and "Pneumonia" categories. To address class imbalance, data augmentation techniques such as rotation and mirroring were applied. The model was trained and evaluated with metrics such as precision, recall, F1-score, and accuracy, presenting a final accuracy of 92.63%. During training, the loss analysis showed a significant reduction in error across epochs, while the confusion matrix

indicated robust performance in classifying both classes. Despite the promising results, the study highlights the limitation of using a single dataset and recommends exploring more diverse datasets to improve generalization and applicability in different clinical contexts. [19]

"Efficient Pneumonia Detection in Chest X-ray Images Using Deep Transfer Learning (2020)" This study used the Guangzhou Women and Children's Medical Center dataset, which consists of 5,836 chest X-rays, including 3,873 pneumonia cases and 1,283 normal images in training. To mitigate class imbalance, data augmentation techniques were applied to the normal image set. The methodology proposed a weighted classifier that combines the results of five pre-trained models: ResNet18, DenseNet121, InceptionV3, Xception, and MobileNetV2, fine-tuned with transfer learning. Evaluation metrics include accuracy, precision, recall, F1-score, and AUC-ROC. The combined model achieved an accuracy of 98.43% and an AUC of 99.76%, outperforming the individual models. Despite the promising results, the study highlighted limitations, including the scarcity of diverse data and the difficulty of generalization to other populations. [8]

The reviewed articles highlight significant limitations related to the scarcity and imbalance of data categories in chest radiograph (CXR) pneumonia detection studies. Although different approaches such as transfer learning, data augmentation, and weighted classifiers have been employed, the insufficiency of images for certain categories, especially normal cases, has been consistently identified as a critical challenge. CNN-based models such as DenseNet121, ResNet18, and MobileNetV2 have demonstrated robust results, but their effectiveness is limited by the diversity and volume of data available, compromising generalization to new clinical scenarios.

The reliance on restricted datasets such as CheXNet and Guangzhou Women and Children's Medical Center highlights the need for innovative strategies to overcome data limitations. Data augmentation with conventional techniques such as rotations and mirroring partially helps but does not fully address the lack of variability in the datasets. Furthermore, the performance of the models is hampered by the scarcity of labeled data for subcategories such as viral and bacterial pneumonia.

Our research project proposes a promising approach using generative adversarial models (GANs) to overcome data scarcity. GANs have the potential to generate realistic synthetic data that can augment datasets and balance underrepresented categories, providing greater diversity and aiding in model generalization. This solution could fill an important gap in current research by providing a more solid foundation for training and validating deep learning models in medical diagnostics.

4 Methodology

The flow of the research is divided into three main stages that cover all the steps of the workflow as shown in Figure 1 8. From data preparation to model evaluation.

In Stage 1 (Data Preparation) , the datasets are collected and enriched with synthetic images generated by generative models such as GANs and DDPM. Next, the data undergoes a preprocessing step for normalization, data augmentation, and validation of the quality of synthetic images, ensuring they are representative of real anomalies. Finally, the data is divided into training, validation, and test sets, establishing a solid foundation for model training. This step is crucial for dealing with possible class imbalances and ensuring efficient training.

In Stage 2 (Model Development) , the supervised machine learning model DenseNet121 is trained using both real and synthetic data, evaluating the impact of artificially generated images on the model's performance. In Stage 3 (Evaluation and Optimization), the models are evaluated using metrics such as precision, recall, F1-score, and AUC. Finally, the model results are analyzed to validate their robustness and generalization, highlighting the importance of using synthetic data in improving anomaly detection tasks.

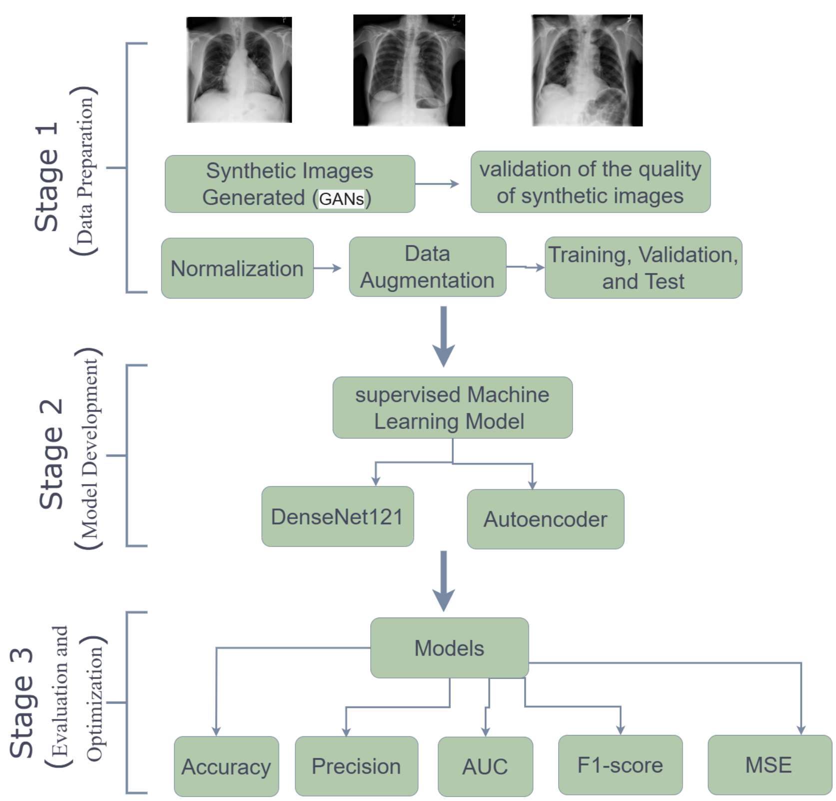

Fig. 6. Methology

4.1 Dataset

The NIH Chest X-ray14 dataset contains approximately 112,000 chest X-ray images, annotated with 14 different conditions, including pneumonia, nodules, and pleural effusions. It is widely used in medical projects due to its diversity and the richness of the conditions represented.

The Chest X-Ray Images (Pneumonia) dataset, in turn, focuses on pneumonia classification and includes images divided into three categories: normal, viral pneumonia, and bacterial pneumonia. It was essential in the training and validation of supervised models, such as DenseNet121, to evaluate the effectiveness of the generated synthetic images.

In the project, this dataset were used to train generative model GANs, which created synthetic images representative of real anomalies, providing a solid foundation for analyses and improvements in anomaly detection.

5 Experiments and Results

In this study, a Generative Adversarial Network (GAN) was implemented for supervised anomaly detection in medical imaging datasets. The GAN's generator synthesized realistic medical images, while its discriminator distinguished between real and synthetic images. Anomalies were identified based on the reconstruction errors of the generated images compared to the original inputs. The model's performance was assessed using precision, recall, F1-score, and accuracy measures, providing a comprehensive understanding of its classification capabilities. Additionally, execution time, hardware resources, and training refinement processes were considered critical factors the get the bellow result.

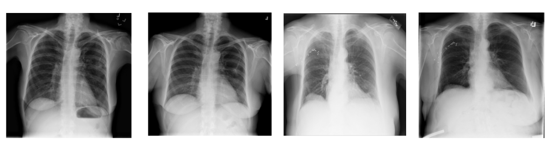

Fig. 7. Real images of NIH Chest X-ray14 dataset

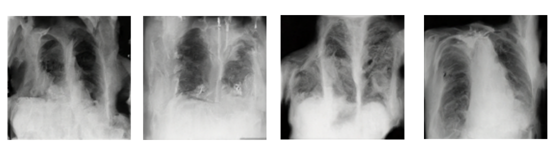

Fig. 8. First synthetic image results using the GAN

5.1 Discussion of Results

Based on the results shown in the examined images, the experiment will be modified further to produce even higher-quality outcomes. The adopted approach will focus on three primary fronts:

1. Improvements to the GAN Model and Hyperparameters Hyperparameters such as the learning rate, batch size, and architecture of both the generator and the discriminator will be refined to improve picture reconstruction and anomaly detection. Regularization techniques and mechanisms such as gradient penalty will also be used to stabilize adversarial learning during training.

2. Data augmentation and diversification To improve the model's robustness, more synthetic photos generated using advanced data augmentation techniques such as geometric modifications, contrast changes, and realistic medical noise simulations will be added. This diversification is intended to increase the model's generalizability and eliminate bias in the outcomes.

With these approaches, it is expected not only to improve the precision, recall, and F1-Score metrics but also to consolidate the use of the methodology as a practical and efficient solution for anomaly detection in clinical settings. The next cycle of experimentation will rigorously monitor the gains in performance and computational efficiency, with the aim of ensuring greater applicability in real-world environments.

6 Conclusion

Significance of Anomaly Detection in Medical Imaging: This work emphasizes the crucial requirement of advanced methods, such as Generative Adversarial Networks, for enabling enhanced anomaly detection across diverse medical imaging modalities, particularly chest X-rays, hence enabling enhanced diagnostic accuracy and patient outcomes. The expected achievements, despite challenges like paucity of data and data imbalance, the generation of synthetic images through GANs offers a novel solution to dataset augmentation and thus ensures diagnostic model robustness and usability. As Future Directions, continued model and hyperparameter development, as well as the creation of new data augmentation methodologies, is vital for the continued growth of the applicability and validity of anomaly detection systems in clinical practice.

References

1. Ajay Bandi, Pydi Venkata Satya Ramesh Adapa, and Yudu Eswar Vinay Pratap Kumar Kuchi. The power of generative ai: A review of requirements, models, input–output formats, evaluation metrics, and challenges. *Future Internet*, 15(8):260, 2023.
2. Heang-Ping Chan, Ravi K Samala, Lubomir M Hadjiiski, and Chuan Zhou. Deep learning in medical image analysis. *Deep learning in medical image analysis: challenges and applications*, pages 3–21, 2020.
3. Pedro Costa, Adrian Galdran, Maria Inês Meyer, Michael David Abramoff, Meindert Niemeijer, Ana Maria Mendonça, and Aurélio Campilho. Towards adversarial retinal image synthesis. *arXiv preprint arXiv:1701.08974*, 2017.
4. Maayan Frid-Adar, Idit Diamant, Eyal Klang, Michal Amitai, Jacob Goldberger, and Hayit Greenspan. Gan-based synthetic medical image augmentation for increased cnn performance in liver lesion classification. *Neurocomputing*, 321:321–331, 2018.
5. Tomoyuki Fujioka, Kazunori Kubota, Mio Mori, Yuka Kikuchi, Leona Katsuta, Mizuki Kimura, Emi Yamaga, Mio Adachi, Goshi Oda, Tsuyoshi Nakagawa, et al. Efficient anomaly detection with generative adversarial network for breast ultrasound imaging. *Diagnostics*, 10(7):456, 2020.
6. Markus Goldstein and Seiichi Uchida. A comparative evaluation of unsupervised anomaly detection algorithms for multivariate data. *PloS one*, 11(4):e0152173, 2016.
7. Dong Gong, Lingqiao Liu, Vuong Le, Budhaditya Saha, Moussa Reda Mansour, Svetha Venkatesh, and Anton van den Hengel. Memorizing normality to detect anomaly: Memory-augmented deep autoencoder for unsupervised anomaly detection. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 1705–1714, 2019.
8. Mohammad Farukh Hashmi, Satyarth Katiyar, Avinash G Keskar, Neeraj Dhanraj Bokde, and Zong Woo Geem. Efficient pneumonia detection in chest xray images using deep transfer learning. *Diagnostics*, 10(6):417, 2020.
9. Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.
10. Zhenxing Huang, Xinfeng Liu, Rongpin Wang, Mudan Zhang, Xianchun Zeng, Jun Liu, Yongfeng Yang, Xin Liu, Hairong Zheng, Dong Liang, et al. Fanet: fast assessment network for the novel coronavirus (covid-19) pneumonia based on 3d ct imaging and clinical symptoms. *Applied Intelligence*, 51:2838–2849, 2021.
11. Tero Karras. Progressive growing of gans for improved quality, stability, and variation. *arXiv preprint arXiv:1710.10196*, 2017.
12. Salome Kazeminia, Christoph Baur, Arjan Kuijper, Bram van Ginneken, Nassir Navab, Shadi Albarqouni, and Anirban Mukhopadhyay. Gans for medical image analysis. *Artificial intelligence in medicine*, 109:101938, 2020.
13. Aghiles Kebaili, Jérôme Lapuyade-Lahorgue, and Su Ruan. Deep learning approaches for data augmentation in medical imaging: a review. *Journal of Imaging*, 9(4):81, 2023.
14. Andy Kitchen and Jarrel Seah. Deep generative adversarial neural networks for realistic prostate lesion mri synthesis. *arXiv preprint arXiv:1708.00129*, 2017.
15. Ke Li, Chuyang Ye, Zhen Yang, Aaron Carass, Sarah H Ying, and Jerry L Prince. Quality assurance using outlier detection on an automatic segmentation method for the cerebellar peduncles. In *Medical Imaging 2016: Image Processing*, volume 9784, pages 398–404. SPIE, 2016.

16. Jessica Lin, Eamonn Keogh, Ada Fu, and Helga Van Herle. Approximations to magic: Finding unusual medical time series. In *18th IEEE Symposium on Computer-Based Medical Systems (CBMS'05)*, pages 329–334. IEEE, 2005.
17. Dwarikanath Mahapatra, Behzad Bozorgtabar, Jean-Philippe Thiran, and Mauricio Reyes. Efficient active learning for image classification and segmentation using a sample selection and conditional generative adversarial network. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 580–588. Springer, 2018.
18. Taoufik Ouleddroun, Ayoub Ellahyani, and Mohamed El Ansari. Automated pneumonia detection using deep features in chest x-ray images. In *2023 3rd International Conference on Intelligent Communication and Computational Techniques (ICCT)*, pages 1–6. IEEE, 2023.
19. Rudresh Pillai, Rupesh Gupta, Neha Sharma, and Rajesh Kumar Bansal. Proposed convolutional neural network for early pneumonia diagnosis. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–6. IEEE, 2023.
20. Alice Presenti, Zhihua Liang, Luis F Alves Pereira, Jan Sijbers, and Jan De Beenhouwer. Automatic anomaly detection from x-ray images based on autoencoders. *Nondestructive Testing and Evaluation*, 37(5):552–565, 2022.
21. P Rajpurkar. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *ArXiv abs/1711*, 5225, 2017.
22. Justus Randolph. A guide to writing the dissertation literature review. *Practical assessment, research, and evaluation*, 14(1):13, 2019.
23. Daisuke Sato, Shouhei Hanaoka, Yukihiro Nomura, Tomomi Takenaga, Soichiro Miki, Takeharu Yoshikawa, Naoto Hayashi, and Osamu Abe. A primitive study on unsupervised anomaly detection with an autoencoder in emergency head ct volumes. In *Medical Imaging 2018: Computer-Aided Diagnosis*, volume 10575, pages 388–393. SPIE, 2018.
24. Thomas Schlegl, Philipp Seeböck, Sebastian M Waldstein, Ursula Schmidt-Erfurth, and Georg Langs. Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In *International conference on information processing in medical imaging*, pages 146–157. Springer, 2017.
25. Neenu Sebastian and B Ankayarkanni. Deep learning applications to detect pneumonia on chest x-ray: A systematic study. In *2022 International Conference on Advanced Computing Technologies and Applications (ICACTA)*, pages 1–5. IEEE, 2022.
26. Philipp Seeböck, José Ignacio Orlando, Thomas Schlegl, Sebastian M Waldstein, Hrvoje Bogunović, Sophie Klimscha, Georg Langs, and Ursula Schmidt-Erfurth. Exploiting epistemic uncertainty of anatomy segmentation for anomaly detection in retinal oct. *IEEE transactions on medical imaging*, 39(1):87–98, 2019.
27. Moazzam Siddiq. Ml-based medical image analysis for anomaly detection in ct scans, x-rays, and mris. *Devotion: Journal of Research and Community Service*, 2(1):53–64, 2020.
28. Jordan ZT Sim, Yong-Han Ting, Yuan Tang, Yangqin Feng, Xiaofeng Lei, Xiaohong Wang, Wen-Xiang Chen, Su Huang, Sum-Thai Wong, Zhongkang Lu, et al. Diagnostic performance of a deep learning model deployed at a national covid-19 screening facility for detection of pneumonia on frontal chest radiographs. In *Healthcare*, volume 10, page 175. MDPI, 2022.
29. Dejan Štepec and Danijel Skočaj. Image synthesis as a pretext for unsupervised histopathological diagnosis. In *Simulation and Synthesis in Medical Imaging: 5th*

International Workshop, SASHIMI 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4, 2020, Proceedings 5, pages 174–183. Springer, 2020.
30. Liyakathunisa Syed, Saima Jabeen, and S Manimala. Telemammography: a novel approach for early detection of breast cancer through wavelets based image processing and machine learning techniques. *Advances in soft computing and machine learning in image processing*, pages 149–183, 2018.
31. Shichang Tang. Lessons learned from the training of gans on artificial datasets. *IEEE Access*, 8:165044–165055, 2020.
32. David MJ Tax and Robert PW Duin. Uniform object generation for optimizing one-class classifiers. *Journal of machine learning research*, 2(Dec):155–173, 2001.
33. Maximilian E Tschuchnig and Michael Gadermayr. Anomaly detection in medical imaging-a mini review. In *International Data Science Conference*, pages 33–38. Springer, 2021.
34. Hristina Uzunova, Sandra Schultz, Heinz Handels, and Jan Ehrhardt. Unsupervised pathology detection in medical images using conditional variational autoencoders. *International journal of computer assisted radiology and surgery*, 14:451–461, 2019.
35. Shashanka Venkataramanan, Kuan-Chuan Peng, Rajat Vikram Singh, and Abhijit Mahalanobis. Attention guided anomaly localization in images. In *European Conference on Computer Vision*, pages 485–503. Springer, 2020.