PEOPLE'S DEMOCRATIC REPUBLIC OF ALGERIA
Ministry of Higher Education and Scientific Research
Mohamed Khider University – BISKRA
Faculty of Exact Sciences, Natural and Life Sciences
Computer Science department

Thesis

Presented to obtain the academic master's degree in

Computer Science

Course : **Artificial Intelligence (AI)**

Semi Supervised Learning For Medical Image

By:

Mamen Lahcene

Defended on 11/06/2023 before the jury composed of :

Youkana Imen	MCB	President
Ammari Asma	MCB	Supervisor
Belaala Abir	MCB	Examiner

Academic year 2023-2024

Acknowledgments

First and foremost, I praise and thank God, the Almighty, for his showers of blessings that enabled me to successfully complete my studies.

I would like to express my deepest gratitude to my research supervisors, Dr. Asma Ammari and Prof. Kahloul Laid, for providing me with the invaluable opportunity to pursue this research. Their guidance, encouragement, and unwavering support have been instrumental in the completion of this work.

I am profoundly grateful to Prof. Okba Kazar, whose fatherly advice and mentorship have been a cornerstone of my academic journey. His insights and wisdom have greatly contributed to my growth and success.

My heartfelt thanks go to my research friend, Zakaria Abdellah Sellam, and my brother, Abdelkarim Mamen, a PhD student, for inspiring me with the idea for this thesis and convincing me to undertake this significant endeavor. Their belief in the importance of this work has been a driving force behind my efforts.

I would also like to extend my sincere appreciation to my friends, Kazar Nazim and Mohamed Taki Eddine Abdesselam, for their invaluable assistance with coding challenges. Their support and technical expertise were crucial during the most challenging moments of this project. Additionally, I want to thank my friend Nadjib Mebarki, who is from the medical field. He provided invaluable advice on where to start to understand the medical field and always helped clarify things when I was confused.

I extend my gratitude to Ilyes Ghamri, a nurse who was the first person I met at the hospital. Without him, I would not have had the chance to communicate with the medical staff. I am also grateful to Dr. Tarhlissia Radhouane, who invested his time and effort to help me realize this project.

Lastly, but most importantly, I would like to express my deepest love and gratitude to my family. Your unwavering support, patience, and encouragement have been my anchor throughout this journey. Your love has been the foundation upon which I have built my aspirations, and for that, I am eternally thankful.

Abstract

Since the inception of medical imaging technologies, significant advancements have continually reshaped the landscape of healthcare diagnostics and treatment planning. In recent years, the integration of Artificial Intelligence (AI), particularly Deep Learning (DL), has propelled medical imaging to new heights. Despite these advancements, the accurate classification of medical images remains challenging, particularly in the presence of limited labeled data and diverse pathologies. This thesis addresses these challenges by proposing a Semi-Supervised Learning (SSL) approach to enhance the classification accuracy of medical images, with a focus on respiratory diseases such as pneumonia.

In this thesis, a comprehensive approach is implemented utilizing Denoising Diffusion Probabilistic Models (DDPM) and ConvNeXt architectures to overcome the limitations of existing methods. The DDPM is employed for generating high-fidelity synthetic medical images, which are then used to augment the training dataset. Concurrently, the ConvNeXt architecture is utilized for robust pseudo-labeling and classification, leveraging both labeled and unlabeled data to improve model performance.

The proposed method is tested on a comprehensive dataset of chest X-radiation (X-ray) images, demonstrating significant improvements in the classification accuracy of pneumonia, even in scenarios with limited labeled data. The approach not only addresses the class imbalance issues but also enhances the model's ability to generalize across different patient demographics and imaging conditions.

Furthermore, the practical applications of this method are explored through case studies involving the classification of pneumonia in various patient demographics and imaging conditions. The results indicate that the proposed approach achieves high precision in identifying and classifying pneumonia, outperforming traditional methods and other state-of-the-art models.

The effectiveness of the proposed SSL framework highlights its potential for broader applications in medical imaging, including the diagnosis of other respiratory conditions and the integration into real-time clinical workflows. This study sets the stage for future research aimed at tackling more complex medical imaging challenges, thereby enhancing patient care and diagnostic accuracy.

Keywords: SSL, Medical Imaging, DL, Pneumonia Classification, Diffusion Denoising Probabilistic Models, ConvNeXt, AI in Healthcare, Respiratory Diseases.

Résumé

Depuis l'apparition des technologies d'imagerie médicale, des avancées significatives ont continuellement remodelé le paysage du diagnostic médical et de la planification des traitements. Ces dernières années, l'intégration de l'intelligence artificielle (IA), en particulier l'apprentissage profond (DL), a propulsé l'imagerie médicale vers de nouveaux sommets. Malgré ces avancées, la classification précise des images médicales demeure un défi, notamment en présence de données étiquetées limitées et de pathologies diversifiées. Cette thèse aborde ces défis en proposant une approche d'apprentissage semi-supervisé pour améliorer la précision de la classification des images médicales, en se concentrant sur les maladies respiratoires telles que la pneumonie.

Dans cette thèse, une approche détaillée a été mise en œuvre en utilisant les modèles probabilistiques de débruitage par diffusion (DDPM) et les architectures ConvNeXt pour surmonter les limitations des méthodes existantes. Le DDPM est employé pour générer des images médicales synthétiques de haute fidélité, qui sont ensuite utilisées pour augmenter le jeu de données d'entraînement. Parallèlement, l'architecture ConvNeXt est utilisée pour un pseudo-étiquetage et une classification robustes, tirant parti à la fois des données étiquetées et non étiquetées pour améliorer les performances du modèle.

La méthode proposée est testée sur un ensemble de données complet d'images radiographiques thoraciques, démontrant des améliorations significatives de la précision de la classification de la pneumonie, même dans des scénarios avec des données étiquetées limitées. L'approche non seulement aborde les problèmes de déséquilibre des classes, mais améliore également la capacité du modèle à généraliser à travers différentes démographies de patients et conditions d'imagerie.

En outre, les applications pratiques de cette méthode sont explorées à travers des études de cas impliquant la classification de la pneumonie dans diverses démographies de patients et conditions d'imagerie. Les résultats indiquent que l'approche proposée atteint une haute précision dans l'identification et la classification de la pneumonie, surpassant les méthodes traditionnelles et d'autres modèles à la pointe de la technologie. L'efficacité du framework d'apprentissage semi-supervisé proposé met en évidence son potentiel pour des applications plus larges en imagerie médicale, y compris le diagnostic d'autres conditions respiratoires et l'intégration dans les flux de travail cliniques en temps réel. Cette étude ouvre la voie à de futures recherches visant à relever des défis plus complexes en imagerie médicale, améliorant ainsi les soins aux patients et la précision des diagnostics.

Mots-clés : Apprentissage Semi-Supervisé, Imagerie Médicale, Apprentissage Profond, Classification de la Pneumonie, Modèles Probabilistiques de Débruitage par Diffusion, ConvNeXt, IA en Santé, Maladies Respiratoires

ملخص

منذ نشأة تقنيات التصوير الطبي، أعادت التطورات الكبيرة باستمرار تشكيل مشهد تشخيص الرعاية الصحية وتخطيط العلاج. في السنوات الأخيرة، أدى دمج الذكاء الاصطناعي، وخاصةً التعلّم العميق، إلى دفع التصوير الطبي إلى آفاق جديدة. على الرغم من هذه التطورات، لا يزال التصنيف الدقيق للصور الطبية يمثل تحديًا، لا سيما في ظل وجود بيانات موسومة محدودة وأمراض متنوعة. تعالج هذه الأطروحة هذه التحديات من خلال اقتراح نهج تعلّم شبه خاضع للإشراف لتعزيز دقة تصنيف الصور الطبية، مع التركيز على أمراض الجهاز التنفسي مثل الالتهاب الرئوي.

في هذه الأطروحة، يتم تنفيذ نهج شامل باستخدام نماذج احتمالية الانتشار الضوضائي المقلل للانتشار وبنى كونفنكست للتغلب على قيود الطرق الحالية. يتم استخدام النماذج الاحتمالية للانتشار الضوضائي لتوليد صور طبية اصطناعية عالية الدقة، والتي يتم استعمالها بعد ذلك لزيادة مجموعة بيانات التدريب. وفي الوقت نفسه، يتم استخدام بنية كونفنكست لوضع علامات زائفة قوية وتصنيفها، والاستفادة من البيانات المصنفة وغير المصنفة لتحسين أداء النموذج.

يتم اختبار الطريقة المقترحة على مجموعة بيانات شاملة من صور الأشعة السينية للصدر، مما يدل على تحسينات كبيرة في دقة تصنيف الالتهاب الرئوي، حتى في السيناريوهات ذات البيانات المحدودة ذات التسميات المحدودة. لا يعالج النهج مشاكل اختلال التوازن الفئوي فحسب، بل يعزز أيضًا قدرة النموذج على التعميم عبر مختلف الخصائص الديموغرافية للمرضى وظروف التصوير.
وعلاوة على ذلك، يتم استكشاف التطبيقات العملية لهذه الطريقة من خلال دراسات حالة تتضمن تصنيف الالتهاب الرئوي في مختلف الحالات الديموغرافية للمرضى وظروف التصوير. تشير النتائج إلى أن المنهج المقترح يحقق دقة عالية في تحديد وتصنيف الالتهاب الرئوي ويتفوق على الطرق التقليدية وغيرها من النماذج الحديثة.

تسلط فعالية إطار التعلّم شبه الخاضع للإشراف المقترح الضوء على إمكاناته في تطبيقات أوسع في التصوير الطبي، بما في ذلك تشخيص حالات الجهاز التنفسي الأخرى والاندماج في تدفقات العمل السريرية في الوقت الحقيقي. تمهد هذه الدراسة الطريق لأبحاث مستقبلية تهدف إلى معالجة تحديات التصوير الطبي الأكثر تعقيداً، وبالتالي تعزيز رعاية المرضى ودقة التشخيص.

الكلمات الدالة: التعلم شبه الخاضع للإشراف، التصوير الطبي، التعلم العميق.

Contents

List of Figures

List of Tables

List of Acronyms

General Introduction

Medical imaging has been a cornerstone of modern healthcare since the advent of X-ray technology in the late 19Th century [1]. Over the past two decades, technological advancements, particularly in the fields of Machine Learning (ML) and AI, have revolutionized how medical images are analyzed and interpreted. Among these advancements, SSL has emerged as a promising approach, leveraging both labeled and unlabeled data to enhance diagnostic efficiency [2]. This is particularly crucial in the realm of respiratory diseases, where accurate and timely diagnosis can significantly impact patient outcomes [3].

In recent years, the integration of AI in medical imaging has seen significant progress, especially in the detection and classification of pneumonia from chest X-ray [4]. Traditional methods, while effective, often fall short due to the high variability in medical images and the need for extensive labeled datasets. SSL addresses these challenges by utilizing the vast amounts of unlabeled data available, thereby improving model performance even with limited labeled data [2].

This thesis implements a multi-method approach employing both DDPM and ConvNeXt architectures to enhance pneumonia classification. The proposed system is designed to overcome the limitations of existing methods by improving data augmentation, pseudo-labeling, and classification processes. The current work discusses the application of these models to real-world medical imaging datasets, focusing on their ability to generalize and accurately classify pneumonia in diverse patient populations.

The remaining of this thesis is organized into three main chapters. Accordingly, the first chapter provides an in-depth exploration of the medical background relevant to respiratory diseases, the importance of medical imaging, and the significant advancements in imaging technology. It also highlights the role of AI in modern healthcare and the specific challenges faced in medical imaging.

The second chapter delves into the theoretical foundations of AI and ML, with a particular emphasis on SSL. It reviews various learning paradigms and presents recent advancements in SSL applied to medical image analysis. Additionally, this chapter introduces the technological choices made for this research, such as the use of DDPM and ConvNeXt architectures.

In the third chapter, the designed approach, implementation, and evaluation are detailed towards accurate pneumonia classification. This includes a comprehensive overview of the data flow, processing steps, and specific challenges encountered during the development process. Also, the obtained experimental results are presented and discussed, analyzing the performance of the proposed models in terms of accuracy, precision, recall, and other relevant metrics. The iterative refinement of the models and the incorporation of expert feedback were crucial in achieving high-quality results.

Eventually, a general conclusion is provided to summarize the key objectives and contributions of this research and to discuss the broader implications of the achieved results for the field of medical imaging. Besides, it outlines potential future directions for further improving SSL techniques and synthetic data generation in medical imaging.

Chapter 1: Medical Background

Chapter 1

Medical Background

1.1 Introduction

Medical imaging plays a pivotal role in modern healthcare, serving as a cornerstone for diagnosis, treatment planning, and monitoring. Understanding the anatomy and physiology of the respiratory system, alongside the advances in imaging technology, provides a foundation for addressing respiratory diseases. This chapter explores the significance of medical imaging, the structure and function of the respiratory system, common respiratory diseases, diagnostic techniques, and the emerging role of AI in medical imaging.

1.2 Medical Imaging

Medical imaging is a backbone of modern healthcare, providing non-invasive methods to visualize the interior of the body [5]. These techniques are crucial for diagnosing diseases, planning treatments, and monitoring the progress of various medical conditions. The evolution of medical imaging technology has significantly enhanced the ability to detect and treat illnesses, leading to improved healthcare outcomes [6].

1.2.1 Importance of Medical Imaging

Medical imaging plays a crucial role in modern healthcare by enabling the visualization of internal body structures and functions without the need for invasive procedures [6]. This capability is essential for several reasons:

- **Diagnosis:** Medical imaging allows for the accurate and early detection of diseases, such as cancer, cardiovascular conditions, and neurological disorders. Techniques like X-ray, MRI, and CT scans provide detailed images that help clinicians identify abnormalities and make precise diagnoses [5].

- **Treatment Planning:** Imaging is vital in planning surgical procedures and other treatments. For instance, MRI and CT scans can guide surgeons in navigating complex anatomical structures, ensuring precision and reducing the risk of complications[7].

- **Monitoring:** Medical imaging is also used to monitor the effectiveness of treatments and track the progression of diseases. For example, repeated imaging studies can assess tumor shrinkage in response to chemotherapy or radiation therapy, allowing for adjustments in treatment plans as needed [5].

In conclusion, medical imaging is indispensable in modern healthcare for its role in diagnosing conditions, guiding treatment strategies, and monitoring patient progress, ultimately leading to better clinical practices.

1.2.2 Advances in Medical Imaging Technology

The field of medical imaging has seen remarkable technological advancements over the past few decades [8]. These advancements have improved image quality, reduced scan times, and expanded the range of detectable conditions. Key developments in imaging modalities, illustrated in Figure 1.1, include:

- **X-ray:** The development of digital X-ray has enhanced image resolution and reduced radiation exposure. Digital radiography allows for faster image acquisition and easier storage and sharing of images [1].

- **MRI:** Advances in MRI technology have led to higher-resolution images and faster scanning techniques. Functional Magnetic Resonance Imaging (fMRI) now enables the visualization of brain activity in real-time, offering insights into neurological functions and disorders [9].

- **CT:** Modern CT scanners provide high-resolution images with reduced radiation doses. Innovations such as dual-energy CT and iterative reconstruction algorithms have improved image quality and diagnostic accuracy [10].

- **PET:** PET scans, often combined with CT, offer detailed images of metabolic activity in the body. This technology is particularly valuable in oncology for detecting cancer and assessing treatment response [11].

- **Ultrasound:** Advances in ultrasound technology, including 3D and 4D imaging, have expanded its applications beyond obstetrics to areas such as cardiology, musculoskeletal imaging, and abdominal imaging [12].

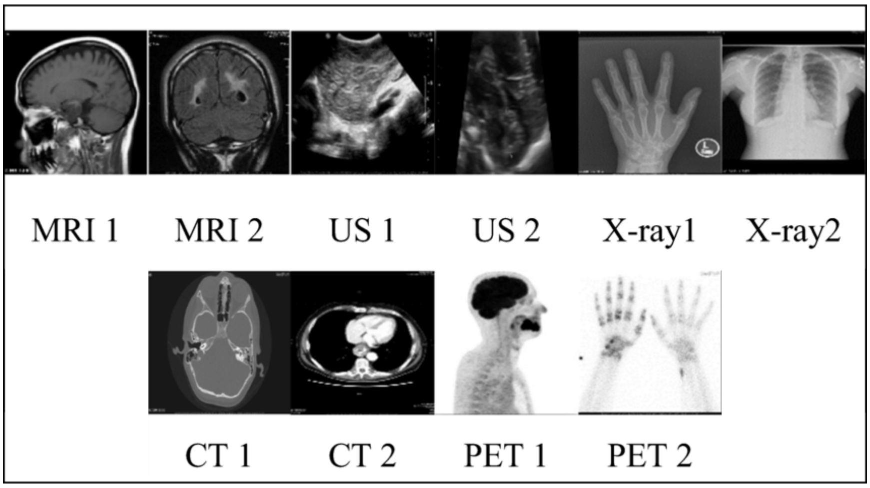


Figure 1.1: Medical Imaging Modalities: MRI, Ultrasound, X-ray, CT, PET [11]

Emerging imaging technologies are poised to further revolutionize healthcare. Techniques such as photoacoustic imaging, which combines optical and ultrasound methods, and advanced molecular imaging, which allows for the visualization of specific biological processes at the molecular level, hold great promise [13]. These innovations are expected to improve early disease detection, enable personalized treatment plans, and enhance patient care.

In conclusion, the continuous advancements in medical imaging technology have significantly enhanced the capabilities of healthcare providers to diagnose, treat, and monitor diseases, leading to better patient outcomes and more efficient healthcare delivery.

1.3 Anatomy and Physiology of the Respiratory System

Understanding the anatomy and physiology of the respiratory system is crucial for comprehending how medical imaging can aid in diagnosing and treating respiratory conditions [14]. This section delves into the structure and function of the lungs, the role of alveoli in gas exchange, and the mechanics of respiration.

1.3.1 Structure and Function of the Lungs

The lungs are essential organs in the respiratory system, responsible for the exchange of gases between the air and the blood [14]. As illustrated in Figure 1.2, key anatomical components include:

- **Lobes:** The right lung consists of three lobes (superior, middle, and inferior), while the left lung has two lobes (superior and inferior). Each lobe is further divided into segments, facilitating efficient gas exchange [15].
- **Pleura:** The lungs are encased in a double-layered membrane called the pleura. The visceral pleura covers the lungs, and the parietal pleura lines the chest cavity. The pleural cavity between these layers contains fluid that reduces friction during breathing [16].
- **Diaphragm:** The diaphragm is a dome-shaped muscle at the base of the lungs, playing a critical role in breathing by contracting and relaxing to change the volume of the thoracic cavity [17].

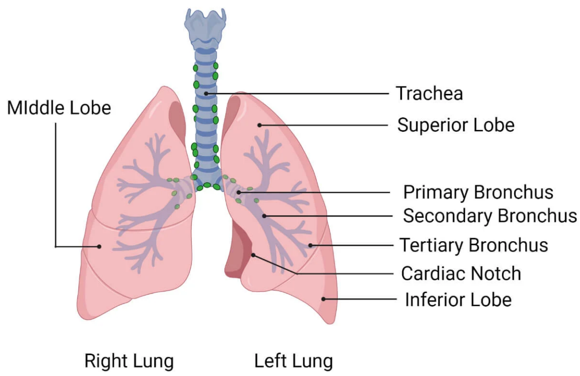


Figure 1.2: Structure of the Lungs [14]

1.3.2 The Role of Alveoli in Gas Exchange

The alveoli are the primary sites for gas exchange in the lungs [18], as shown in Figure 1.3. These structures play a critical role in respiratory function due to several key features:

- **Structure and Function:** Alveoli are tiny, balloon-like structures surrounded by a network of capillaries. The walls of the alveoli are extremely thin, allowing for efficient diffusion of gases [18].

- **Alveolar Surface Area:** The large surface area of the alveoli, combined with their extensive capillary network, maximizes the exchange of oxygen and carbon dioxide. Oxygen diffuses from the alveoli into the blood, while carbon dioxide diffuses from the blood into the alveoli to be exhaled [18].

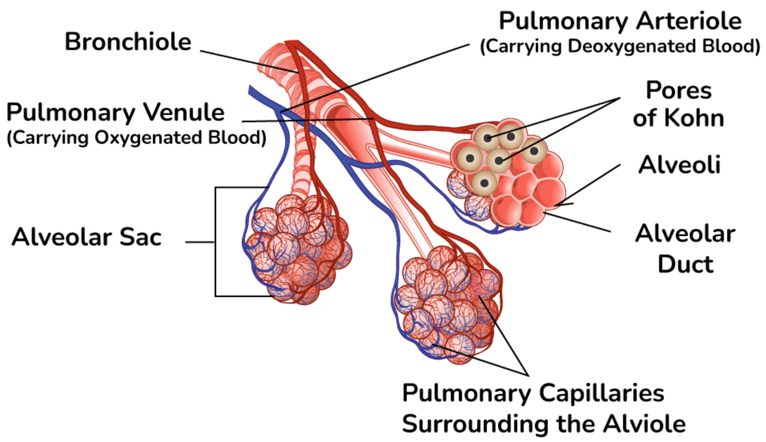


Figure 1.3: Structure of the Alveoli [18]

1.3.3 Respiratory Mechanics

Breathing involves the coordinated effort of several muscles and the movement of air into and out of the lungs. Key components include (as illustrated in Figure 1.4):

- **Diaphragm and Intercostal Muscles:** The diaphragm contracts and moves downward during inhalation, increasing the thoracic cavity's volume and allowing air to enter the lungs. The intercostal muscles, located between the ribs, also contract to expand the chest cavity. During exhalation, the diaphragm relaxes and moves upward, while the intercostal muscles relax, decreasing the thoracic cavity's volume and expelling air from the lungs [17].

- **Mechanisms of Inhalation and Exhalation:** Inhalation is an active process driven by the contraction of the diaphragm and intercostal muscles, creating a negative pressure within the thoracic cavity that draws air into the lungs. Exhalation is typically a passive process resulting from the relaxation of these muscles, although it can become active during forceful breathing [19].

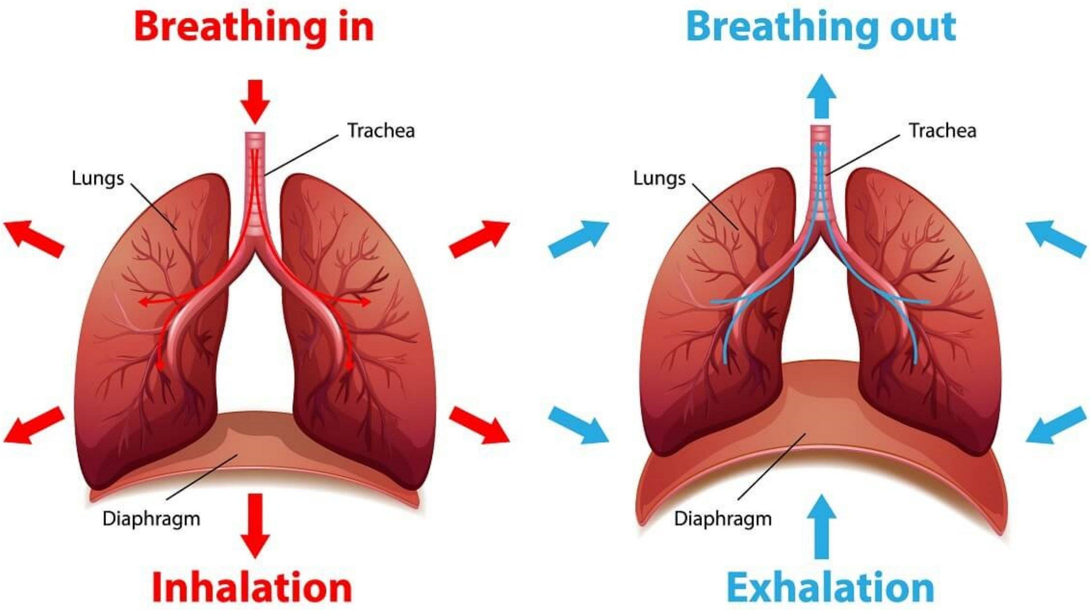


Figure 1.4: Mechanism of Breathing [17]

Understanding the anatomy and physiology of the respiratory system provides a foundational knowledge necessary for interpreting medical imaging and addressing respiratory diseases effectively.

1.4 Respiratory Diseases and Their Impact

Respiratory diseases significantly affect the function of the respiratory system, leading to various health complications and reduced quality of life [20]. This section provides an overview of common respiratory diseases, the impact of pneumonia on alveolar function, and other notable respiratory infections and conditions.

1.4.1 Overview of Common Respiratory Diseases

Understanding the causes, symptoms, and impacts of common respiratory diseases is essential for diagnosing and treating these conditions effectively.

A. Asthma:

- **Causes:** Asthma is a chronic inflammatory disease of the airways that can be triggered by allergens, environmental factors, exercise, and stress [20].

- **Symptoms:** Common symptoms include wheezing, shortness of breath, chest tightness, and coughing, particularly at night or early morning [20].

- **Impact on the Respiratory System:** Asthma leads to airway hyperresponsiveness, inflammation, and obstruction, which can cause recurrent episodes of breathing difficulties and reduce overall lung function [20].

B. Chronic Obstructive Pulmonary Disease (COPD):

- **Pathophysiology:** COPD is characterized by persistent respiratory symptoms and airflow limitation due to airway and/or alveolar abnormalities, usually caused by significant exposure to noxious particles or gases, such as smoking [21].

- **Clinical Presentation:** Patients with COPD often experience chronic cough, sputum production, and progressive dyspnea. The disease is also associated with frequent respiratory infections and exacerbations that can lead to hospitalization [21].

C. Pneumonia:

- **Causes:** Pneumonia is an infection that inflames the air sacs in one or both lungs. It can be caused by bacteria, viruses, or fungi [22].

- **Types:** There are various types of pneumonia, including community-acquired pneumonia, hospital-acquired pneumonia, and aspiration pneumonia [22].

- **Epidemiology:** Pneumonia is a significant cause of morbidity and mortality worldwide, particularly among the elderly, infants, and individuals with weakened immune systems [22].

1.4.2 Impact of Pneumonia on Alveolar Function

Pneumonia can severely affect the alveoli, which are crucial for gas exchange, by causing inflammation and filling them with fluid and pus (see Figure 1.5).

· **Pathogenesis of Pneumonia:** The infection causes the alveoli to fill with pus and fluid, leading to impaired gas exchange. The inflammation can spread throughout the lungs, causing widespread respiratory distress [23].

· **Clinical Symptoms and Complications:** Symptoms of pneumonia include fever, chills, cough with phlegm, chest pain, and difficulty breathing. Severe cases can lead to complications such as pleural effusion, lung abscesses, and Acute Respiratory Distress Syndrome (ARDS), which can be life-threatening [23].

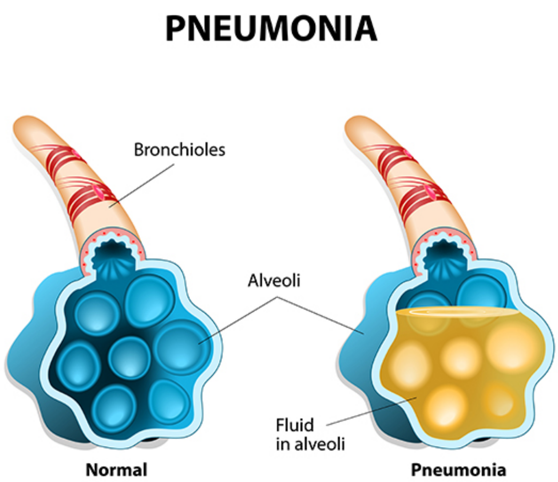


Figure 1.5: Pneumonia Effect on Alveolar Function [23]

1.4.3 Other Respiratory Infections and Conditions

Other respiratory infections and conditions also pose significant health challenges and require careful diagnosis and management. These diseases, as illustrated in Figure 1.6, underscore the broad spectrum of respiratory conditions affecting global health.

A.Tuberculosis (TB):

· **Impact on Lungs:** TB is a bacterial infection caused by *Mycobacterium tuberculosis*. It primarily affects the lungs, causing chronic cough, hemoptysis, fever, night sweats, and weight loss [24].

· **Diagnostic Challenges:** Diagnosing TB can be challenging due to its slow progression and the need for specialized tests, such as sputum culture, chest X-ray, and molecular assays [24].

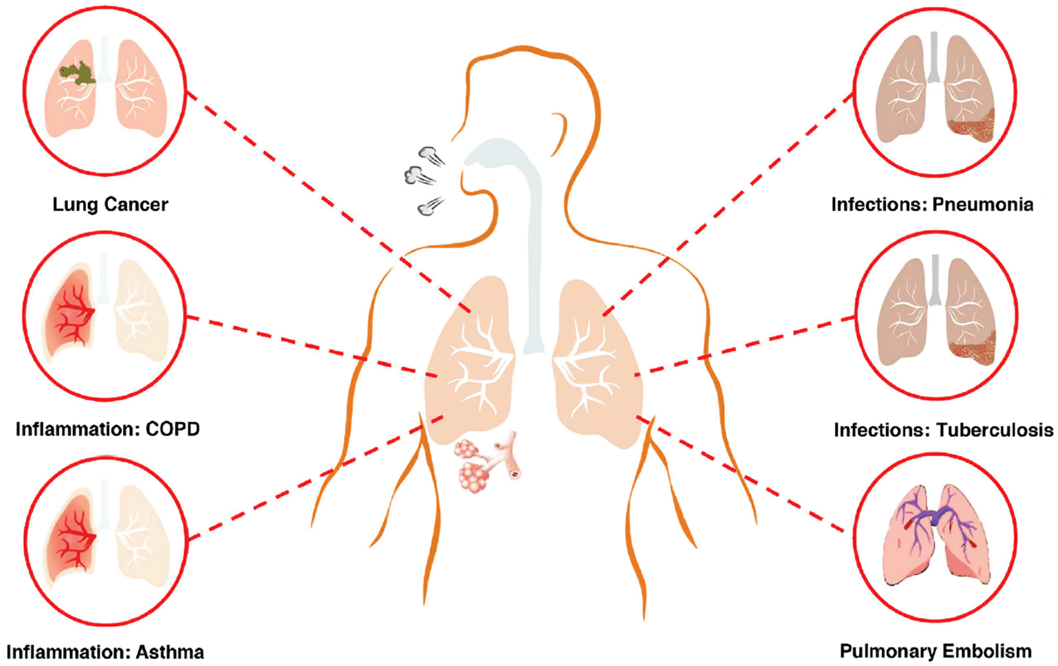


Figure 1.6: Common Respiratory Diseases [25]

B. Lung Cancer:

- **Types:** Lung cancer is broadly categorized into Non-Small Cell Lung Cancer (NSCLC) and Small Cell Lung Cancer (SCLC). NSCLC is more common and includes subtypes such as adenocarcinoma, squamous cell carcinoma, and large cell carcinoma [25].

- **Stages:** The staging of lung cancer, from I to IV, indicates the extent of disease spread and helps guide treatment decisions [25].

- **Significance of Early Detection:** Early detection of lung cancer significantly improves the chances of successful treatment and survival. Imaging techniques such as CT scans play a critical role in early diagnosis [25].

C. Pulmonary Fibrosis:

- **Causes:** Pulmonary fibrosis is a condition characterized by the thickening and scarring of lung tissue, which can result from various factors, including chronic inflammatory processes, environmental exposures, and certain medications [26].

- **Effects on Lung Function:** The fibrosis leads to a progressive decline in lung function, causing symptoms such as chronic dry cough, shortness of breath, fatigue, and diminished exercise tolerance [26].

Understanding these respiratory diseases and their impact on lung function is crucial for developing effective diagnostic and therapeutic strategies.

1.5 Diagnostic Techniques in Respiratory Medicine

Effective diagnosis of respiratory diseases relies on a combination of clinical evaluation, laboratory tests, and advanced imaging techniques [27]. This section discusses the various diagnostic methods used in respiratory medicine, with a particular focus on diagnosing pneumonia and the presentation and interpretation of X-ray images.

1.5.1 Diagnosis of Pneumonia

Pneumonia is diagnosed using a variety of methods to assess the presence and severity of the infection.

A. Diagnostic Methods:

- **Blood Tests:** Blood tests help identify the presence of an infection and the body's response to it. Elevated white blood cell counts and other markers of inflammation can indicate pneumonia [28].

- **Pulse Oximetry:** This non-invasive test measures the oxygen saturation in the blood. Lower than normal levels can indicate impaired lung function due to pneumonia [29].

- **Sputum Analysis:** Examining sputum (mucus coughed up from the lungs) under a microscope can identify the specific pathogen causing the infection, helping to guide antibiotic therapy [30].

- **Bronchoscopy:** This procedure involves inserting a thin, flexible tube with a camera into the airways to directly visualize the lungs and collect samples for further analysis [31].

B. The Role of Chest X-ray and CT Scans:

- **Chest X-ray:** X-ray are often the first imaging test performed when pneumonia is suspected. They can reveal areas of opacity (infiltrates) in the lungs, indicative of infection [32]. Figure 1.7 shows a comparison between a normal chest X-ray and a chest X-ray indicating pneumonia.

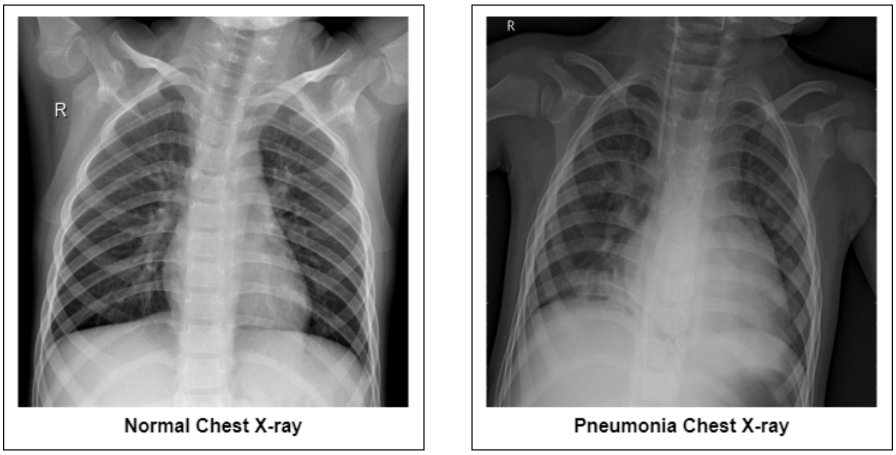


Figure 1.7: Comparison of Normal Chest X-ray and Pneumonia Chest X-ray [32]

- **CT Scans:** CT scans provide more detailed images than X-ray, allowing for a more precise assessment of the extent and severity of the infection, as well as identifying complications such as abscesses or pleural effusion [32].

1.5.2 Presentation of X-ray Images

X-ray are a fundamental tool in diagnosing and managing respiratory diseases [32]. This section explores the historical background, functionality, and specific application of X-ray in pulmonary radiography.

A. Historical Background of X-ray Discovery and Evolution

- **Wilhelm Conrad Roentgen's Discovery:** In 1895, Wilhelm Conrad Roentgen discovered X-ray, a form of electromagnetic radiation, while experimenting with cathode rays. This groundbreaking discovery revolutionized medical diagnostics by enabling the visualization of internal structures without invasive procedures [33]. The first X-ray photograph taken by Roentgen, as shown in Figure 1.8, depicts the hand of Roentgen's wife, with her wedding ring clearly visible on her finger bones.

- **Impact on Medicine:** The introduction of X-ray rapidly transformed clinical practice, allowing for the early detection of fractures, infections, and tumors. Over the years, advancements in X-ray technology have continued to enhance its diagnostic capabilities [33].

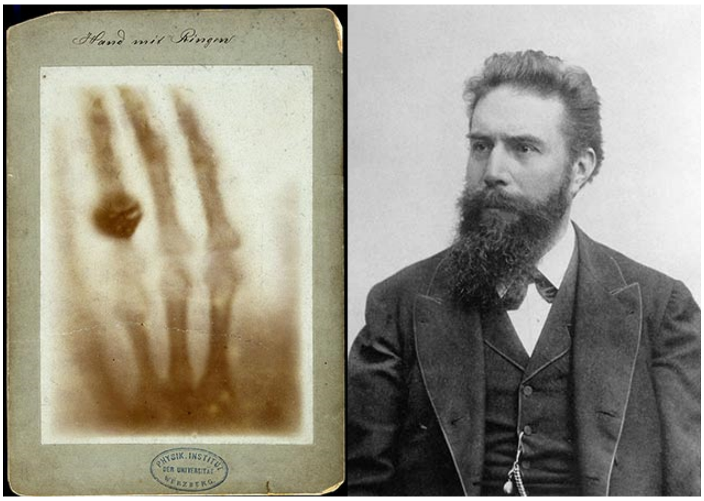


Figure 1.8: First X-ray Photography by Wilhelm Conrad Roentgen [33]

B. Functioning of X-ray Machines

- **Principles of X-ray Production:** X-ray are produced when high-energy electrons collide with a metal target, typically tungsten, inside the X-ray tube. This collision generates X-ray, which are directed towards the body part being examined [33].

- **Image Formation:** As X-ray pass through the body, they are absorbed to varying degrees by different tissues. Dense tissues, such as bone, absorb more X-ray and appear white on the resulting image, while softer tissues absorb fewer X-ray and appear darker [33].

- **Interpretation of X-ray Images:** Interpreting X-ray images involves analyzing the contrast between different tissue densities. Radiologists look for abnormalities such as infiltrates, masses, or fluid accumulation to diagnose conditions like pneumonia, fractures, or tumors [33].

C. Pulmonary Radiography

- **Process and Techniques:** Pulmonary radiography involves taking X-ray images specifically of the lungs and thoracic cavity. The patient is usually positioned in various ways (e.g., standing, lying down) to obtain multiple views, ensuring a comprehensive assessment [34].

- **Diagnostic Value:** Pulmonary X-ray are invaluable for diagnosing lung diseases. They can reveal patterns of infection, such as the consolidation seen in pneumonia, detect fluid in the pleural space, identify tumors, and monitor the progress of chronic conditions like COPD and pulmonary fibrosis [34]. Figure 1.9 shows the normal anatomy of a chest X-ray, labeling important anatomical structures such as the clavicle, ribs, heart, lungs, and diaphragm. Understanding these structures is crucial for accurately interpreting pulmonary radiographs.

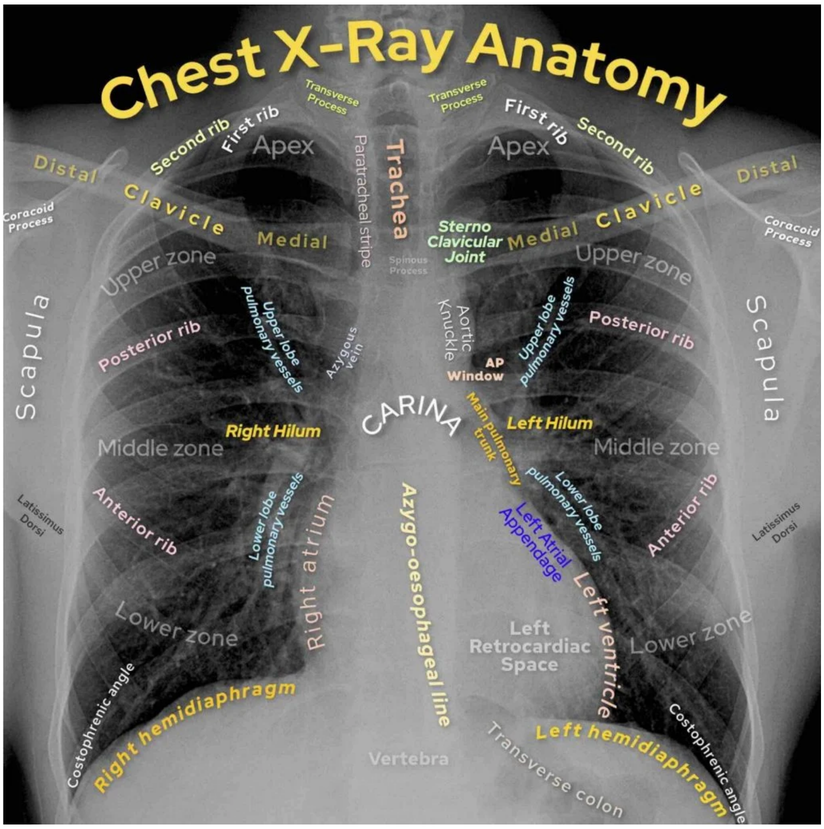


Figure 1.9: Normal Chest X-ray Anatomy [34]

Eventually, the use of various diagnostic techniques, including blood tests, pulse oximetry, sputum analysis, bronchoscopy, and advanced imaging methods like X-ray and CT scans, is essential for accurately diagnosing and managing respiratory diseases. Understanding the historical evolution and technical aspects of X-ray imaging further underscores its importance in modern respiratory medicine.

1.6 Advanced Imaging Techniques

Advanced imaging techniques have revolutionized the field of respiratory medicine, providing detailed insights into the anatomy and pathology of the respiratory system. This section explores the principles, advantages, and clinical applications of CT scans, MRI, and PET.

1.6.1 Computed Tomography (CT) Scans

CT scans are a vital tool in modern diagnostic imaging, offering detailed cross-sectional images of the body.

A. Principles of CT Imaging: CT imaging involves rotating X-ray beams around the patient to capture multiple images from different angles. These images are then processed by a computer to create detailed cross-sectional views of the body [35]. The process of CT imaging is illustrated in Figure 1.10, which shows the key components involved:

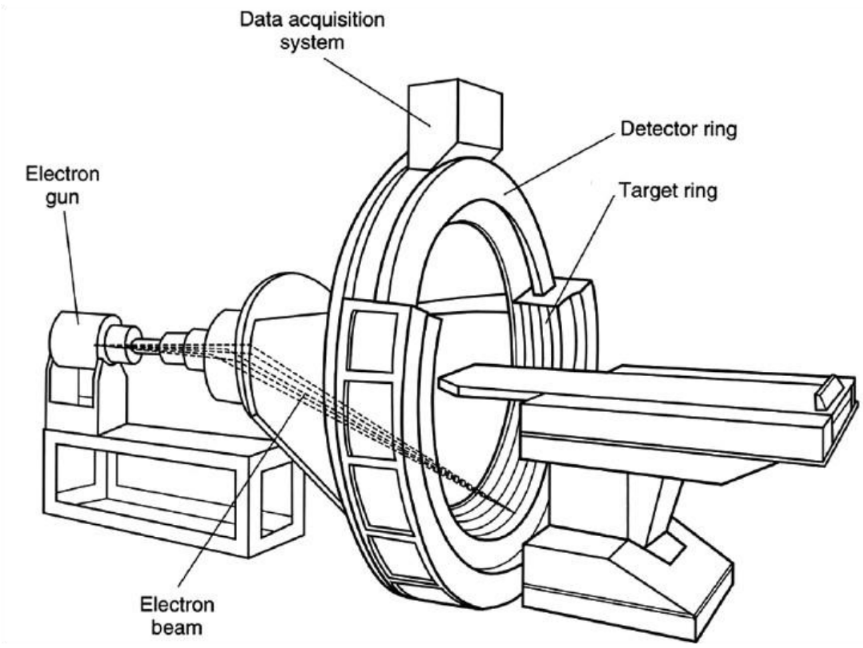


Figure 1.10: Schematic of a CT Machine [35]

- **Electron Gun:** Generates and focuses a beam of electrons.

- **Electron Beam:** Directed towards a target to produce X-rays.
- **Data Acquisition System:** Collects data from the detector ring, including the intensity of the X-rays after passing through the patient's body.
- **Detector Ring:** Contains multiple detectors capturing X-rays after passing through the body, allowing the creation of cross-sectional images.
- **Target Ring:** The area where the electron beam is focused to generate X-rays.

B. Advantages over Traditional X-ray:

- CT scans provide greater detail and clarity than traditional X-ray. They can distinguish between different types of tissues with high precision, allowing for the identification of subtle abnormalities that might not be visible on standard X-ray [35].

C. Clinical Applications of CT Scans in Respiratory Medicine:

- CT scans are widely used to diagnose and monitor respiratory conditions such as pneumonia, lung cancer, pulmonary embolism, and interstitial lung diseases. They provide comprehensive images that help in assessing the extent and severity of these conditions [35].
- They are also used in guiding biopsies and other interventional procedures, ensuring accuracy and safety [35].

1.6.2 Magnetic Resonance Imaging (MRI)

MRI is a powerful imaging modality particularly useful for soft tissue imaging.

A. How MRI Works:

- MRI uses strong magnetic fields and radio waves to generate detailed images of the body's internal structures. When placed in a magnetic field, hydrogen atoms in the body align with the field. Radiofrequency pulses disrupt this alignment, and as the atoms return to their original state, they emit signals that are captured and converted into images by the MRI scanner [36].

B. Benefits in Soft Tissue Imaging:

- MRI provides superior contrast resolution compared to CT, making it especially valuable for imaging soft tissues, such as the brain, muscles, and organs [36].

C. Uses of MRI in Diagnosing Respiratory Conditions:

- MRI is particularly useful in evaluating chest wall and pleural diseases, vascular abnormalities, and congenital lung conditions. It can provide detailed images of soft tissue structures within the thoracic cavity that are not as clearly seen with CT [36].

- Functional MRI (fMRI) can assess lung function and ventilation patterns, offering insights into conditions like asthma and COPD [36].

The schematic of an MRI machine, shown in Figure 1.11, illustrates the concentric arrangement of coils, including the magnet, gradient coils, radio frequency coils, and bore.

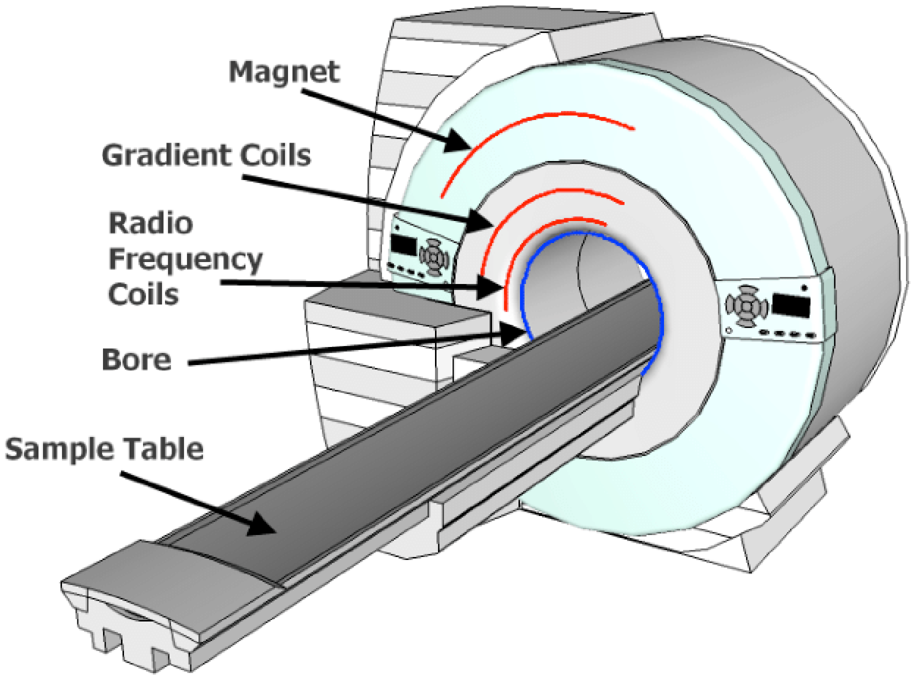


Figure 1.11: Schematic of an MRI Machine [36]

1.6.3 Positron Emission Tomography (PET)

PET scans are essential for detecting metabolic activity within the body.

A. Functionality of PET Scans:

· PET imaging involves the use of radioactive tracers that are absorbed by active cells. The scanner detects the gamma rays emitted by the tracer and creates images that reflect metabolic activity. This allows for the identification of areas with increased or decreased metabolic function [37].

· The included schematic in Figure 1.12 illustrates the process of a PET scan. The radiotracer is injected into the antecubital vein, travels through the bloodstream, and accumulates in regions with high metabolic activity. The PET scanner detects the emitted gamma rays and constructs detailed images of the brain and other organs.

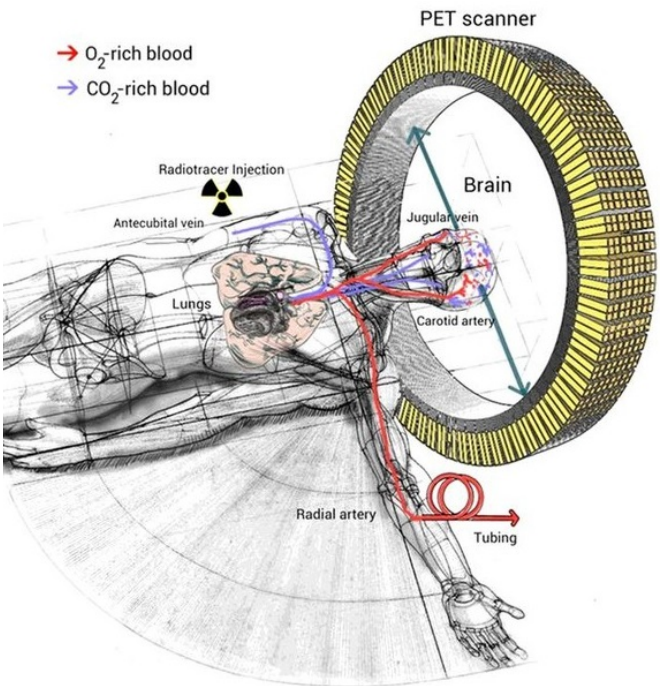


Figure 1.12: Schematic of an PET Machine [37]

B. Role in Detecting Metabolic Activity:

· PET scans can identify cancerous cells, which typically have higher metabolic rates than normal cells, as well as areas of inflammation and infection [37].

· As shown in the schematic 1.12, the radiotracer travels through the carotid artery to the brain, highlighting regions of metabolic activity. The scanner captures these areas, helping to diagnose and monitor various conditions.

C. Applications of PET in Lung Cancer and Inflammatory Diseases:

- **Lung Cancer:** PET scans are extensively used in the diagnosis, staging, and monitoring of lung cancer. They help detect primary tumors, assess metastasis, and evaluate the response to treatment [37].

- **Inflammatory Diseases:** PET imaging can identify areas of inflammation, making it useful in diagnosing and monitoring conditions such as sarcoidosis, tuberculosis, and other inflammatory lung diseases [37].

In summary, advanced imaging techniques like CT, MRI, and PET scans have significantly enhanced the diagnostic capabilities in respiratory medicine. These technologies provide detailed and precise images, facilitating accurate diagnosis, effective treatment planning, and thorough monitoring of various respiratory conditions.

1.7 Challenges in Medical Imaging-based Diagnosis

Despite significant advancements in medical imaging technologies, several challenges remain in accurately diagnosing and managing respiratory conditions. This section discusses the limitations of current diagnostic methods, the growing demand for improved diagnostic tools, and the potential solutions offered by SSL approaches.

1.7.1 Limitations of Current Diagnostic Methods

Current diagnostic methods in medical imaging, though advanced, still face several limitations that can impact the accuracy and reliability of diagnoses.

A. Inaccuracy and Limitations of Traditional Diagnostic Techniques:

- Traditional imaging techniques like X-ray and basic ultrasound can sometimes lack the resolution and specificity needed to detect early or subtle abnormalities. This can lead to missed diagnoses or incorrect assessments of disease severity [38].

- Techniques such as CT and MRI, while more advanced, can still have limitations related to patient movement, the presence of artifacts, and variations in interpretation[38].

B. Challenges in Interpreting Medical Images and Variability in Diagnoses:

- Interpreting medical images requires significant expertise and experience. Variability in interpretation can occur due to differences in radiologists' training, experience, and subjective judgment. This variability can lead to inconsistent diagnoses and treatment plans [38].

- Complex cases often require consensus from multiple specialists, further highlighting the potential for variability and the need for standardized interpretation protocols[38].

C. The Need for Large Labeled Datasets for Training Supervised Models:

- Supervised ML models require extensive labeled datasets to train accurately. Obtaining these large datasets can be challenging, as it involves meticulous labeling by experts, which is time-consuming and resource-intensive [38].

- The lack of sufficient labeled data can limit the effectiveness of ML models, leading to less accurate predictions and classifications [38].

1.7.2 The Growing Demand for Improved Diagnostic Tools

The increasing complexity and volume of medical data necessitate the development of more advanced and efficient diagnostic tools to keep pace with the demands of modern healthcare[38].

A. Increasing Complexity and Volume of Medical Data:

- The sheer volume of medical imaging data generated daily is overwhelming, making it challenging for healthcare providers to manage and analyze this data efficiently. The complexity of data also increases with advancements in imaging modalities that produce higher resolution and multi-dimensional images [38].

- Efficient data management and analysis tools are required to handle this influx of information, ensuring timely and accurate diagnoses [38].

B. Need for Faster, More Accurate Diagnostic Methods:

- As medical conditions become more complex and patient numbers increase, there is a pressing need for diagnostic methods that can deliver quick and accurate results. Delays in diagnosis can lead to worsening of conditions and reduced patient outcomes [38].
- Innovations in imaging technology and analytical methods are essential to meet these needs, providing rapid and reliable diagnostic information [38].

C. Addressing the Lack of Labeled Data with SSL Approaches:

- SSL, which combines a small amount of labeled data with a large amount of unlabeled data, offers a promising solution to the challenge of insufficient labeled datasets. This approach can enhance the learning process and improve the accuracy of diagnostic models [38].
- SSL techniques can leverage the vast amounts of unlabeled medical imaging data available, reducing the dependency on exhaustive manual labeling and accelerating the development of robust diagnostic algorithms [38].

D. Handling Variability in Data Quality and Imaging Conditions:

- Medical images can vary significantly in quality due to differences in imaging equipment, patient positioning, and technical settings. These variations can complicate the diagnostic process and reduce the consistency of image interpretation [38].
- Advanced image processing algorithms and ML models can help standardize and enhance image quality, making it easier to achieve consistent and accurate diagnoses across different settings and conditions [38].

In conclusion, while current diagnostic methods in medical imaging are highly advanced, they still face significant challenges that can impact their effectiveness. Addressing these limitations through improved diagnostic tools, leveraging SSL approaches, and standardizing image quality and interpretation are critical steps toward enhancing the accuracy and reliability of medical imaging in respiratory medicine.

1.8 The Role of AI in Medical Imaging

AI is revolutionizing the field of medical imaging by enhancing diagnostic accuracy, reducing interpretation times, and managing large datasets. This section provides an overview of AI in medicine and discusses how AI addresses current challenges [38][39].

1.8.1 Introduction to AI in Medicine

AI involves the development of algorithms that enable computers to learn from and make predictions based on data. In the context of medical imaging, AI can significantly improve diagnostic processes and outcomes [38][39].

- Potential of AI in Transforming Medical Diagnostics:

- AI algorithms can analyze vast amounts of medical imaging data quickly and accurately, identifying patterns and anomalies that may be missed by human observers [38][39].

- These technologies can assist radiologists by providing decision support, reducing diagnostic errors, and improving overall efficiency in clinical workflows [38][39].

1.8.2 Addressing Current Challenges with AI

AI technologies offer solutions to many of the limitations faced by traditional imaging techniques. AI can significantly improve accuracy, reduce diagnostic times, and efficiently handle large datasets, addressing several key challenges in medical imaging.

A. Improved Accuracy:

- AI algorithms can analyze images with a level of detail and consistency that surpasses human capabilities. This reduces diagnostic errors and improves patient outcomes by providing precise measurements and identifying subtle changes in medical images [38][39].

- AI-powered tools can aid in the early detection and monitoring of diseases, ensuring timely medical interventions [38][39].

B. Reduced Diagnostic Times:

- Automated image analysis with AI can significantly speed up the diagnostic process. By providing faster results, AI enables timely medical interventions and improves patient care [38][39].

- This efficiency allows radiologists to focus on more complex cases, enhancing overall workflow efficiency [38][39].

C. Handling Large Datasets:

- AI systems are well-suited for managing and analyzing large volumes of imaging data. They can extract valuable insights that would be difficult for human experts to discern, integrating information from multiple sources to provide comprehensive diagnostic insights [38][39].

- This capability ensures that the increasing complexity and volume of medical data are managed effectively, supporting accurate and timely diagnoses [38][39].

Accordingly, the integration of AI in medical imaging holds great promise for enhancing diagnostic accuracy, efficiency, and the management of large datasets. These technologies are poised to transform the landscape of medical diagnostics, improving outcomes for patients worldwide.

1.9 Conclusion

This chapter highlights the indispensable role of medical imaging in modern healthcare, with a particular focus on its applications in diagnosing and managing respiratory diseases. The discussion starts with the importance of medical imaging and the significant advancements in imaging technologies such as X-ray, MRI, CT, and PET. Each technology's specific applications in respiratory medicine were examined, showcasing how these tools contribute to better diagnosis, treatment planning, and disease monitoring. Besides, an in-depth look at the anatomy and physiology of the respiratory system provided a foundation for understanding how imaging techniques are utilized.

Significant challenges in medical imaging are also highlighted, including the limitations of current diagnostic methods, such as inaccuracy, variability in image interpretation, and the need for large labeled datasets to tackle AI-based methods. The growing demand for improved diagnostic tools is discussed, highlighting the need for faster and more accurate diagnostic methods. The chapter is concluded by emphasizing the emerging role of AI in enhancing medical imaging. It underscored how AI can overcome traditional imaging limitations by improving accuracy, reducing diagnostic times, and efficiently managing large datasets.

In the following chapter, the challenge of labeled datasets availability will be covered by reviewing and synthesizing a full state-of-the-art about the current semi-supervised approaches investigated to handle the datasets-related issues.

Chapter 2 : Theoretical and Technical Background

Chapter 2

Theoretical and Technical Background

2.1 Introduction

The integration of AI in medical imaging has ushered in a new era of precision diagnostics and personalized treatment plans. AI's capacity to process and analyze large datasets efficiently has been particularly transformative in the field of medical imaging, where rapid and accurate image interpretation is critical. This chapter delves into the theoretical foundations of AI in medical imaging, with a particular focus on SSL methods. We will explore ML and DL techniques, various learning paradigms, and specific advancements in SSL applied to medical image analysis. Additionally, the chapter introduces cutting-edge technological choices like diffusion models and ConvNeXt architectures, which have shown significant promise in enhancing the accuracy and efficiency of medical image classification and generation.

2.2 Overview of AI in Medical Imaging

Advancements in artificial intelligence (AI) have revolutionized various fields, including medical imaging. By leveraging AI, healthcare professionals can enhance the accuracy, efficiency, and scope of diagnostic imaging, ultimately improving patient outcomes.

2.2.1 Definition and Scope

AI in medical imaging refers to the use of advanced computational techniques to interpret and analyze medical images with minimal human intervention. AI encompasses various methods, including ML and DL, to process large datasets, identify patterns, and make predictions. The scope of AI in medical imaging includes enhancing image quality, automating image analysis, improving diagnostic accuracy, and aiding in treatment planning. AI technologies are designed to perform tasks traditionally carried out by radiologists and other medical professionals, thus augmenting their capabilities and enabling more efficient workflows [40][41]. As illustrated in Figure 2.1, AI includes a broad range of technologies, with ML and DL being crucial components that enable sophisticated image processing and analysis.

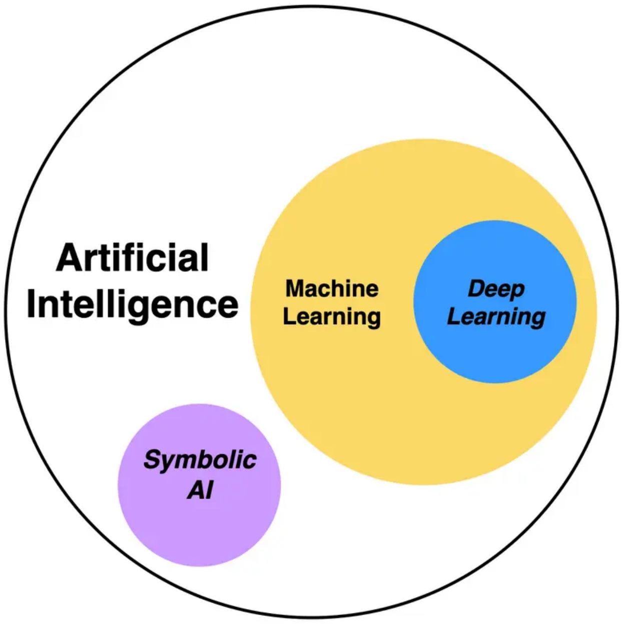


Figure 2.1: Relationship between Artificial Intelligence, Machine Learning, and Deep Learning [42]

2.2.2 Impact on Radiology

AI has had a profound impact on the field of radiology, transforming how images are processed and interpreted. Key benefits include:

- **Enhanced Image Analysis:** AI algorithms can detect and quantify abnormalities, such as tumors or fractures, with high precision, often surpassing human capabilities in speed and accuracy [43].
- **Workflow Efficiency:** Automated image processing reduces the time radiologists spend on routine tasks, allowing them to focus on complex cases and patient interactions [43].

- **Early Detection and Diagnosis:** AI systems can identify subtle changes in imaging data, facilitating early diagnosis of conditions like cancer, thereby improving patient outcomes [43].
- **Personalized Treatment Plans:** By analyzing vast amounts of imaging data, AI can assist in developing tailored treatment strategies based on individual patient profiles [43].

2.2.3 Machine Learning in Medical Imaging

ML has become a cornerstone in the evolution of medical imaging, offering innovative solutions that enhance diagnostic accuracy, efficiency, and patient outcomes. By leveraging the capabilities of ML, healthcare professionals can process and analyze large volumes of medical images, leading to significant advancements in various medical disciplines [44].

A. General Overview

ML, a subset of AI, focuses on developing algorithms that enable computers to learn from and make predictions or decisions based on data. In the context of medical imaging, ML involves training models on vast amounts of medical image data to identify patterns, detect anomalies, and support diagnostic and prognostic decisions.

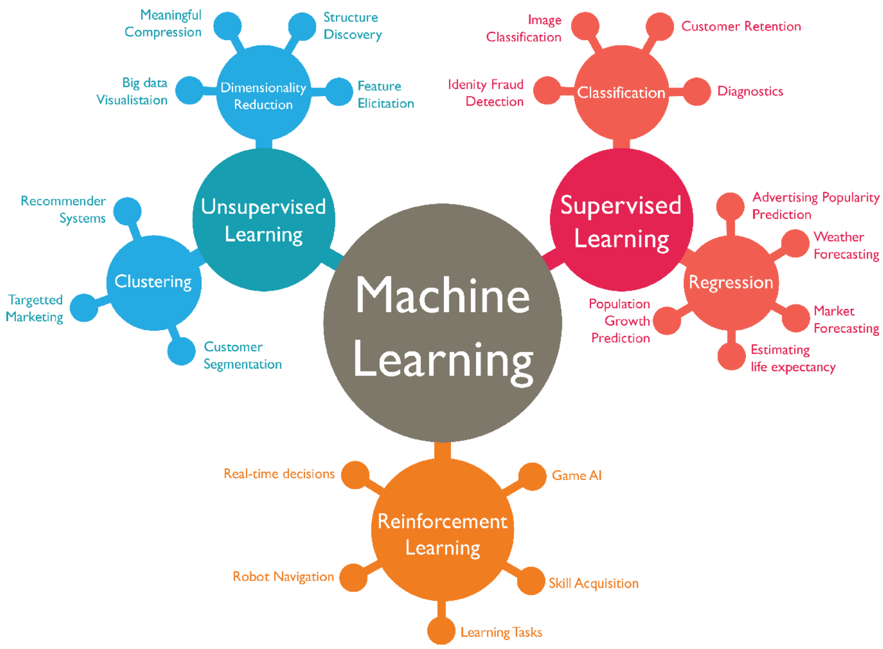


Figure 2.2: Types of Machine Learning and Their Applications [44]

The primary types of ML techniques used in medical imaging include supervised learning, unsupervised learning, and reinforcement learning (as illustrated in Figure 2.2). ML's capacity to process and analyze large datasets swiftly and accurately makes it a powerful tool in advancing medical imaging technologies [44].

B. Applications and Impact

The applications of ML in medical imaging are diverse and significantly impactful across various medical disciplines. Some key applications and their impacts include:

- **Image Classification:** ML algorithms can classify medical images into different categories, such as distinguishing between benign and malignant tumors in radiology. This capability enhances the accuracy of diagnoses and helps in the early detection of diseases [45].
- **Segmentation:** ML techniques, especially DL models like Convolutional Neural Networks (CNNs), are adept at segmenting medical images to outline structures such as organs, tissues, or pathological regions. Accurate segmentation is crucial for treatment planning, surgical interventions, and monitoring disease progression [45].
- **Anomaly Detection:** ML models can be trained to recognize normal patterns in medical images and identify deviations that may indicate abnormalities. This is particularly useful in screening programs, where early detection of conditions like breast cancer in mammograms can lead to better patient outcomes [45].
- **Image Reconstruction:** ML enhances image reconstruction processes in modalities like MRI and CT scans, leading to higher quality images with reduced noise and artifacts. This improvement in image quality aids radiologists in making more precise evaluations and diagnoses [45].

The integration of ML into medical imaging not only enhances the accuracy and efficiency of diagnostic processes but also contributes to the development of personalized medicine and improved patient outcomes. As ML technologies continue to evolve, their applications in medical imaging are expected to expand, further revolutionizing the field.

2.2.4 Deep Learning in Medical Imaging

DL has emerged as a transformative force in the realm of medical imaging, building upon the foundations laid by ML to deliver unprecedented capabilities in image analysis and interpretation. The application of DL in medical imaging harnesses the power of neural networks to process and analyze vast amounts of complex data, driving significant advancements in diagnostic accuracy and efficiency [46].

A. General Overview

DL is a specialized branch of ML that employs neural networks with multiple layers (deep neural networks) to analyze and interpret complex data. In medical imaging, DL models, particularly CNNs, have become the cornerstone for tasks such as image classification, segmentation, and detection due to their ability to automatically learn and extract features from raw data. Unlike traditional ML methods that rely on manual feature extraction, DL models can process large amounts of unstructured data, capturing intricate patterns and relationships within medical images. This capability has made DL a powerful tool in enhancing the precision and efficiency of medical image analysis [46][47]. Figure 2.3 visually compares the machine learning and deep learning processes, highlighting how deep learning integrates feature extraction and classification into a single, streamlined process.

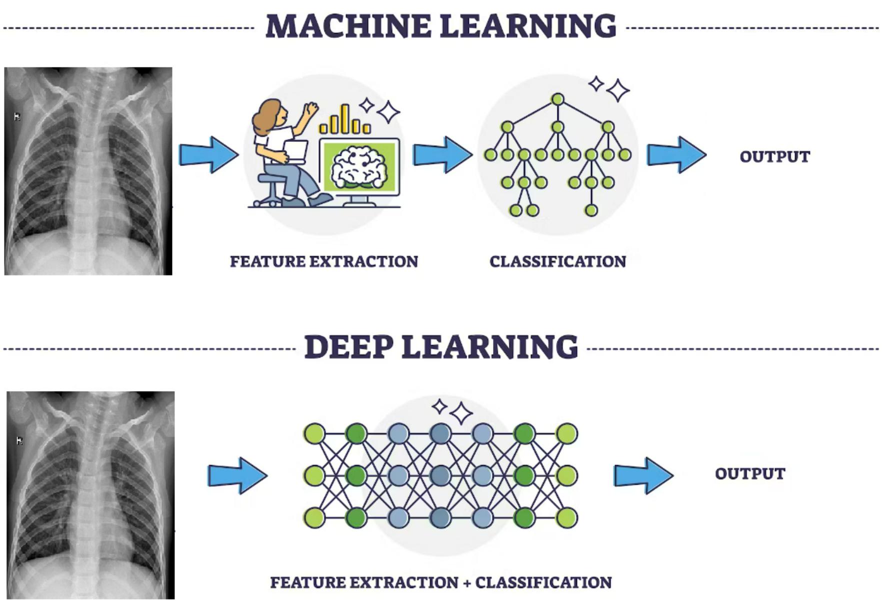


Figure 2.3: Comparison of Machine Learning and Deep Learning Processes [47]

B. Applications and Impact

The application of DL in medical imaging has led to numerous groundbreaking achievements, revolutionizing various aspects of the field. Some of the key achievements include:

- **High-Accuracy Diagnostics:** DL models have achieved remarkable accuracy in diagnosing diseases from medical images, often surpassing human experts. For instance, DL algorithms can detect diabetic retinopathy from retinal images, classify skin lesions from dermoscopic images, and identify pulmonary nodules from CT scans with high precision [47].

- **Automated Image Segmentation:** DL techniques, such as U-Net and its variants, have significantly advanced the accuracy and efficiency of image segmentation. These models can delineate anatomical structures, tumors, and other regions of interest in medical images with high fidelity, which is critical for treatment planning and monitoring disease progression [47].

- **Enhanced Image Generation and Reconstruction:** Generative models like Generative Adversarial Networks (GANs) have been used to generate high-quality synthetic medical images, augmenting training datasets for other DL models. Additionally, DL has improved image reconstruction in modalities such as MRI and CT, reducing noise and artifacts to produce clearer images [47].

- **Real-Time Image Analysis:** The speed and efficiency of DL models allow for real-time analysis of medical images. This capability is particularly valuable in settings like emergency rooms or during surgical procedures, where immediate image interpretation is essential [47].

- **Transfer Learning and Pretrained Models:** The use of transfer learning, where models pretrained on large datasets like ImageNet are fine-tuned on medical imaging datasets, has significantly reduced the need for extensive labeled medical data. This approach has accelerated the deployment of DL solutions in clinical practice [47].

The advancements in DL have not only improved diagnostic accuracy and efficiency but also contributed to personalized medicine by providing insights into patient-specific conditions and responses to treatment. As DL technologies continue to evolve, their impact on medical imaging is expected to grow, further transforming the field and enhancing patient care.

2.3 AI Learning Strategies

In the field of AI include a variety of approaches and techniques used to enable computers to learn from data and make decisions. These techniques can be broadly categorized into different types based on the nature of the learning process and the type of feedback provided to the model. Figure 2.4 illustrates an overview of these categories and techniques.

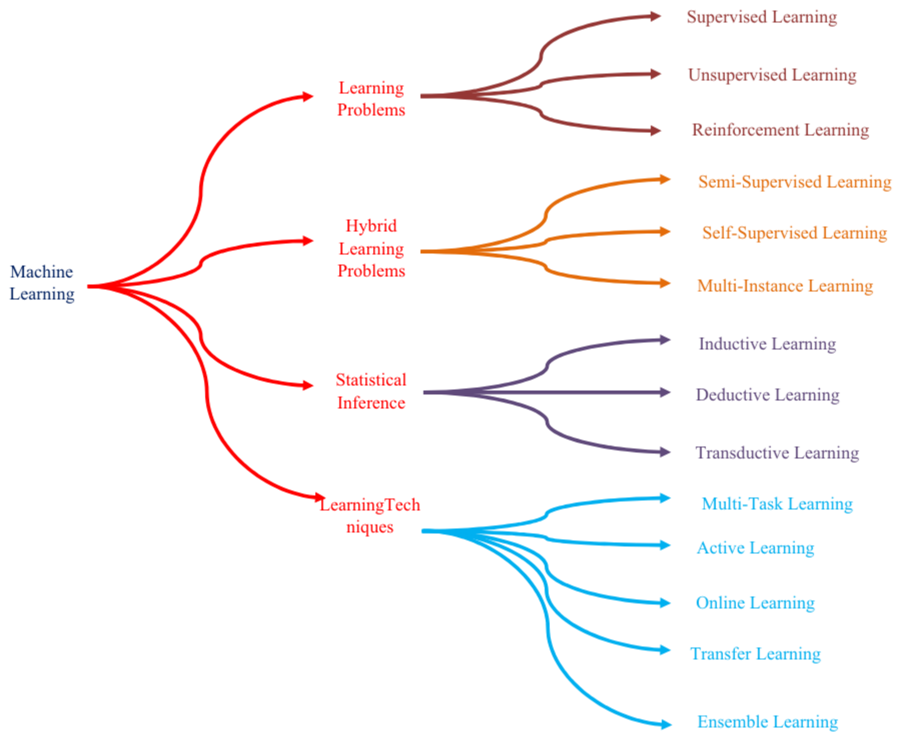


Figure 2.4: Overview of Machine Learning Categories and Techniques [48]

2.3.1 Supervised Learning

Supervised learning is a type of ML where the model is trained on a labeled dataset, meaning that each training example is paired with an output label. The goal is for the model to learn the mapping from inputs to outputs so it can predict the labels for new, unseen data [49][50]. Common applications include classification tasks (e.g., diagnosing diseases from medical images) and regression tasks (e.g., predicting patient outcomes based on clinical data) [49].

2.3.2 Unsupervised Learning

Unsupervised learning involves training a model on data without labeled responses. The model tries to learn the underlying structure of the data through techniques like clustering (grouping similar data points together) and dimensionality reduction (reducing the number of variables under consideration) [51]. In medical imaging, unsupervised learning can be used to identify patterns and anomalies in large datasets without prior knowledge of what the patterns represent.

2.3.3 Reinforcement Learning

Reinforcement learning involves scenarios where an individual must adapt to a specific environment using feedback. Similar to supervised learning, reinforcement learning uses feedback that might be delayed. Additionally, due to the inherently noisy nature of the model, it only provides limited responses for learning. This makes it difficult for both the entity and the model to establish causal connections [52].

2.3.4 Hybrid Learning Problems

Hybrid learning problems combine elements of supervised and unsupervised learning to leverage the strengths of both approaches.

A. Semi-Supervised Learning

SSL uses a small amount of labeled data along with a large amount of unlabeled data during training. This method is particularly useful in medical imaging, where obtaining labeled data is expensive and time-consuming. SSL can improve model performance by utilizing the vast amounts of available unlabeled data [53].

B. Self-Supervised Learning

Self-supervised learning is a type of unsupervised learning where the data itself provides the supervision. The model is trained on a pretext task where the labels are derived from the input data, such as predicting the rotation angle of an image. This technique allows models to learn useful representations from unlabeled data, which can then be fine-tuned for specific tasks [53].

C. Multi-Instance Learning

Multi-instance learning involves training a model on labeled groups (bags) of instances, rather than on individual instances. Each bag is labeled, but individual instances within the bag are not. This approach is useful in scenarios where the relationship between instances and labels is complex, such as in identifying pathological regions within medical images where the presence of disease might only be indicated by a subset of instances [54].

2.3.5 Statistical Inference in Learning

Statistical inference methods are used to draw conclusions from data subject to random variation.

Inductive Learning

Inductive learning involves using specific instances to learn general principles or models. In medical imaging, this means learning general diagnostic rules from a set of labeled images, which can then be applied to new, unseen images [55].

A. Deductive Inference

Deductive inference uses general principles to make predictions about specific instances. This approach is less common in ML but can be used to apply known medical knowledge to interpret new imaging data [48].

B. Transductive Learning

Transductive learning aims to make predictions directly on the specific test instances without developing a general model. This method is useful in medical imaging when the test data is available at training time, allowing the model to tailor its predictions to the specific examples it will encounter [56].

2.3.6 Learning Techniques

Various advanced learning techniques enhance the effectiveness and efficiency of AI models in medical imaging.

A. Multi-Task Learning

Multi-task learning involves training a single model on multiple related tasks simultaneously. This approach can improve generalization by leveraging commonalities among tasks. For example, a model could be trained to identify different types of diseases from medical images, benefiting from shared features across tasks [57].

B. Active Learning

Active learning is a technique where the model can query a human annotator to label new data points it finds most informative. This approach is valuable in medical imaging to maximize the efficiency of the labeling process by focusing on the most uncertain or ambiguous cases [58].

C. Online Learning

Online learning involves continuously updating the model as new data becomes available, rather than retraining it from scratch. This is particularly useful in medical settings where new imaging data is constantly being generated, allowing the model to stay up-to-date with the latest information [59].

D. Transfer Learning

Transfer learning involves taking a pretrained model on a large dataset and fine-tuning it for a specific task with a smaller dataset. This technique is widely used in medical imaging to leverage models trained on general image datasets and adapt them to medical-specific tasks, significantly reducing the need for large labeled datasets [60].

E. Ensemble Learning

Ensemble learning combines the predictions of multiple models to improve overall performance. Techniques such as bagging, boosting, and stacking are used to create a robust model that reduces variance and bias, leading to more accurate and reliable predictions in medical imaging tasks [61].

To tackle the earlier mentioned challenges related to labeled dataset availability. the focus of the upcoming sections surrounds particularly Semi-supervised methods.

2.4 Semi-Supervised approaches for Medical Imaging

SSL is emerging as a pivotal approach in the field of medical imaging, addressing the critical challenge of limited labeled data. By harnessing both labeled and unlabeled datasets, SSL techniques enable the development of more accurate and efficient models for medical image analysis, offering significant potential to enhance diagnostic processes and patient outcomes[62].

2.4.1 Overview and Importance

SSL occupies a middle ground between supervised and unsupervised learning, utilizing both labeled and unlabeled data for training. This approach is particularly significant in medical imaging, where acquiring labeled data can be challenging and expensive due to the need for expert annotations. SSL methods enable the development of robust models by leveraging the abundant unlabeled data, thereby improving the performance of medical image analysis tasks such as classification, segmentation, and anomaly detection [62].

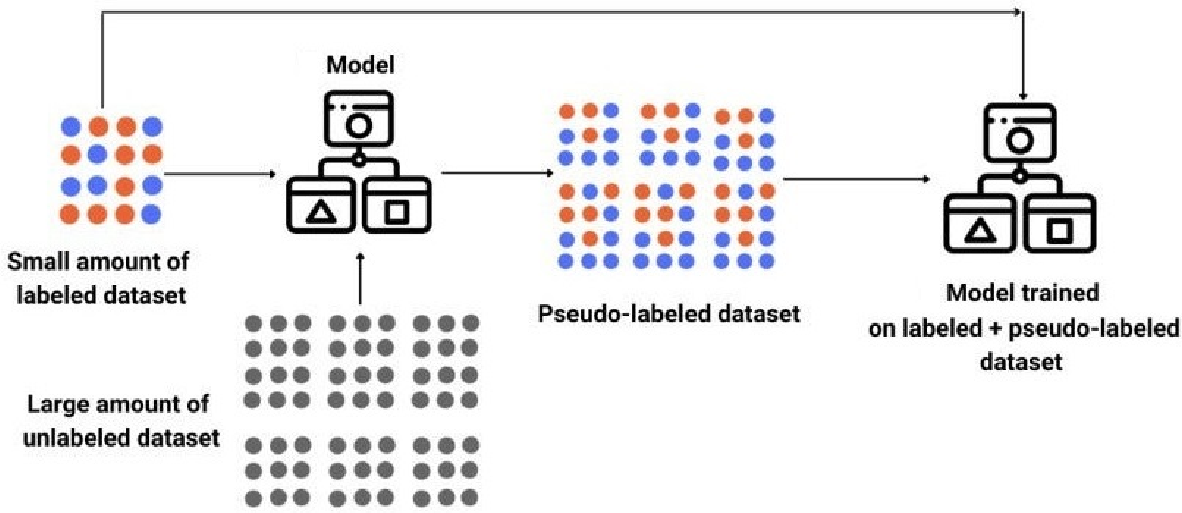


Figure 2.5: Semi-Supervised Learning Process[63]

As illustrated in Figure 2.5, the semi-supervised learning process begins with a small labeled dataset and a large amount of unlabeled data. Initially, a model is trained on the small labeled dataset. The trained model is then used to predict labels for the unlabeled data, resulting in a pseudo-labeled dataset. Finally, the model is retrained using both the original labeled data and the pseudo-labeled data, enhancing its performance [63].

2.4.2 Challenges related to Labeled Data Acquisition

The acquisition of labeled medical images faces several challenges:

- **Expertise Requirement:** Labeling medical images requires specialized knowledge from radiologists or medical professionals, making the process time-consuming and costly [63].
- **Data Privacy and Security:** Patient data is sensitive and often subject to strict privacy regulations, limiting the availability of publicly accessible labeled datasets[63].
- **Class Imbalance:** Certain medical conditions are rare, resulting in an imbalance where the number of normal images far exceeds those depicting pathological cases, complicating the training process [63].

2.4.3 Benefits of SSL for Medical Image Classification

SSL offers several benefits for medical image classification:

- **Enhanced Model Performance:** By utilizing both labeled and unlabeled data, SSL methods can improve the accuracy and robustness of models, especially in scenarios with limited labeled data [63].
- **Cost Efficiency:** Reducing the reliance on extensive labeled datasets lowers the costs associated with data annotation and allows for faster model development[63].
- **Better Generalization:** SSL models can learn more generalized representations from the combined data, leading to better performance on unseen cases and diverse patient populations [63].

2.5 Review of SSL-based methods

Various SSL methodologies have been developed to address the unique challenges of medical image classification.

2.5.1 Consistency-Based Methods

Consistency-based methods encourage the model to produce similar outputs when presented with slightly perturbed versions of the same input. This approach leverages data augmentation to improve the robustness of the model [64].

A. Embarrassingly Simple Consistency Regularization (SCR) Method

This method introduces a straightforward consistency regularization strategy to enhance semi-supervised medical image segmentation. It utilizes unlabeled data effectively, boosting the performance of DL models with limited labeled data. The authors claim that this method outperforms many leading SSL techniques while requiring fewer annotations [65].

B. Medical Image Classification with Relation-driven Self-ensembling Model

This approach employs a relation-driven self-ensembling model to improve classification accuracy by understanding the relationships between different classes. The model creates pseudo-labels for unlabeled data, outperforming traditional methods in both single and multi-label classification scenarios [66].

2.5.2 Graph-Based Methods

Graph-based SSL methods represent data as graphs, with nodes indicating samples and edges representing relationships. These methods propagate labels from labeled to unlabeled samples based on graph structure insights [67].

A. The GraphNet Zoo

The GraphNet Zoo integrates various graph-based SSL techniques into a unified framework, demonstrating effective medical image classification with minimal labeled data. This approach matches the performance of fully supervised methods, highlighting its potential in resource-constrained settings [68].

B. Graph-BAS3Net

Graph-BAS3Net uses bilateral graph convolution for boundary-aware semi-supervised segmentation. This method enhances segmentation accuracy by generating pseudo-labels for unlabeled data, showing superior performance in both labeled and unlabeled conditions [69].

2.5.3 Adversarial Methods

Adversarial SSL methods use adversarial training to improve model robustness. Two popular approaches are Virtual Adversarial Training (VAT) and GANs [70][71].

A. Semi-Supervised Generative Adversarial Network and Pseudo-Labelling

This method combines GANs with pseudo-labeling for medical image classification. GANs generate realistic images, and pseudo-labels classify unlabeled data effectively, enhancing classification accuracy in scenarios with limited labeled data [72].

B. Semi-Supervised Attention-Guided Generative Adversarial Network

Semi-Supervised Attention-Guided Generative Adversarial Network (SAG-GAN) employs semi-supervised attention-guided GANs for data augmentation in medical imaging. This technique generates synthetic images to improve the performance of models like ResNet18, particularly on small datasets of tumor MRI images [73].

A summary of the above SSL-based approaches for medical image analysis is provided in Table 2.1. Each reviewed work includes details about its category of semi-supervised methodology, the datasets employed, and the key results. This table offers a quick reference to compare various approaches and their performance, showcasing the advancements and effectiveness of SSL-based approaches in different medical imaging contexts. By examining these studies, we can understand the current trends and identify areas for further research and improvement in medical image analysis.

Table 2.1: Summary of the reviewed SSL-Based Methods

Categories	Ref	Objective	Methods	Dataset	Results
Consistency Based Methods	[65]	Cardiac Images Segmentation	Simple Consistency Regularization	ACDC2017, MMWHS datasets	**ACDC2017:** DSC = 89.8% HD = 4.47 **MMWHS:** DSC = 79.83% HD = 3.04
	[66]	Thorax disease classification	Relation-Driven SSL	ChestX-ray, ISIC 2018 datasets	**ChestX-ray:** AUC=79.23% **ISIC 2018:** AUC= 93.58% Sens= 71.47% Spec= 92.72% Acc= 92.54% F1= 60.68%
Graph-Based Methods	[68]	Medical Image Classification	GraphNet Zoo + SSL	ChestX-ray, Malaria Cells, CBIS-DDSM, datasets	**ChestX-ray:** AUC= 81.5% (20% labels) **Malaria Cells:** Acc= 92.7% (20% labels) **CBIS-DDSM:** AUC= 73.5% (20% labels)
	[69]	Lungs Segmentation	Graph-BAS3Net	LiTS, COVID-19 datasets	**LiTS:** DSC= 95.58% (100% labeled) DSC= 93.19% (10% labeled) **COVID-19:** DSC= 82.09% (100% labeled) DSC= 74.22% (10% labeled)
Adversarial Methods	[73]	Data Augmentation	SAG-GAN	BraTS, BraTSS, UNS, datasets	**BraTS:** Acc= 95.03% AUC= 96.97% **BraTSS:** Acc= 87.59% AUC= 95.30% **UNS:** Acc= 76.50% AUC= 82.97%
	[72]	Medical Image Classification	GAN with pseudo-labeling	ChestX-ray, BreakHis, datasets	**ChestX-ray:** Acc= 93.15% **BreakHis:** Acc= 96.87%

2.5.4 Other Methods

In the realm of less common ML techniques, various innovative methods are briefly reviewed, including pseudo labeling, which trains supervised models on labeled data before applying them to unlabeled data to enhance training sets with high-confidence pseudo-labels. However, this approach can be biased towards dominant classes. To counteract these limitations, it's often combined with other methods like consistency-based or adversarial techniques. Additionally, Also we have Federated Learning, which allows for decentralized training across multiple institutions without sharing private data. Other methods discussed include Learning by Association, which seeks to maximize correct associations between data embeddings, and Contrastive Learning, which focuses on distinguishing between data augmentations to improve generalization across varied datasets. MixMatch, another notable method, integrates multiple loss functions to refine model training using augmented data, balancing consistency, entropy minimization and generalization [74].

2.6 Case Study: Pneumonia Detection

Pneumonia detection has become a significant focus in medical imaging research due to its potential to improve diagnostic accuracy and patient outcomes. By leveraging advanced imaging techniques and AI, researchers aim to enhance the detection and classification of pneumonia, particularly in early stages[23].

2.6.1 Overview of Medical Imaging for Pneumonia Diagnosis

Pneumonia is a severe respiratory condition characterized by inflammation of the air sacs in one or both lungs, which can fill with fluid or pus. Diagnosing pneumonia typically involves analyzing chest X-ray images to identify signs of infection such as lung opacities. Early and accurate detection is crucial for effective treatment and management [23]. Traditional diagnostic methods rely heavily on radiologists' expertise, but the integration of AI, particularly through DL, has significantly enhanced the accuracy and speed of pneumonia detection and classification.

2.6.2 Recent Work and Advancements

Recent advancements in AI have led to the development of various DL models and techniques for pneumonia detection and classification, leveraging large datasets of chest X-ray images. These advancements highlight the potential of AI to surpass traditional diagnostic methods in both accuracy and efficiency.

A. Pneumonia Detection: An Efficient Approach Using Deep Learning

This study developed an AI network to identify pneumonia from chest X-ray images using an ensemble of two DL models, ResNet-34 based U-Net (a type of convolutional neural network) (U-Net) and EfficientNet-B4 based U-Net. The ensemble approach effectively handled class imbalance and improved diagnostic precision. Trained on a Kaggle dataset, the model achieved an accuracy of 90%, precision of 87%, recall of 99%, and an F1-score of 92%. This robust performance across various metrics indicates the model's reliability in clinical settings [75].

Limitation: Challenges in handling class imbalance effectively, despite the ensemble model's high accuracy.

B. Pneumonia Classification using Deep Learning in Healthcare

This study aimed to classify pneumonia from chest X-ray images using a DL model, specifically a convolutional neural network (CNN) built from scratch. The CNN was developed with layers for convolving images to extract features, applying activation functions, and classifying the images. Using a Kaggle dataset and various data augmentation techniques, the model achieved an accuracy of 94.36%, underscoring its potential to improve diagnostic accuracy over traditional methods [76].

Limitation: Potential overfitting issues due to building models from scratch without leveraging pre-trained networks.

C. Detection of Pneumonia from Chest X-ray Images Utilizing MobileNet Model

This study explored the use of MobileNet and other pre-trained models to differentiate between normal and severe cases of pneumonia. The researchers applied eight pre-trained models on datasets, including the Kaggle dataset and the ChestX-ray14 dataset. The MobileNet model achieved an accuracy of 93.75%, demonstrating a balance between computational efficiency and accuracy, suitable for real-time clinical applications. However, the reliance on pre-trained models may limit customization for specific diagnostic needs [77].

Limitation: Reliance on pre-trained models may limit customization for specific diagnostic needs.

D. GANs 'N Lungs: Improving Pneumonia Prediction

This research addressed the challenge of class imbalance in medical imaging datasets using CycleGAN for data augmentation. The study enhanced the training of a binary classifier based on the DenseNet-121 architecture by oversampling the minority class. The model achieved Receiver Operating Characteristic Area Under the Curve (ROC AUC) scores increasing from 0.9745 to 0.9929 and Precision-Recall Area Under the Curve (PR AUC) scores from 0.9580 to 0.9865. The use of GANs for data augmentation proved effective in balancing the dataset, thereby improving model performance. However, GAN-generated data may introduce artifacts, requiring careful evaluation to ensure clinical reliability [78].

Limitation: Potential introduction of artifacts in GAN-generated data, requiring careful evaluation.

E. A Deep Learning Based Approach towards the Automatic Diagnosis of Pneumonia from Chest Radiographs

This study proposed an automatic detection system for pneumonia from chest radiography images using a deep Siamese-based neural network. The focus was on distinguishing between viral and bacterial pneumonia infections by analyzing the spread of white substance in chest X-ray images. Utilizing a Kaggle dataset, the Deep Siamese Network achieved an accuracy of 89.67%. This approach of comparing symmetric structures in chest X-ray to quantify infection spread demonstrated the model's innovative potential in differentiating types of pneumonia [79].

Limitation: The focus on symmetry may not capture all relevant features for distinguishing pneumonia types.

F. Pneumonia Detection on Chest X-ray Images Using Ensemble of Deep Convolutional Neural Networks

This study enhanced the accuracy of pneumonia detection using an ensemble learning method combining pre-trained CNN models. The approach employed a combination of DenseNet169, MobileNetV2, and Vision Transformers (ViT) models pre-trained on the ImageNet database, fine-tuned on chest X-ray datasets, and integrated using ensemble learning. The model achieved an accuracy of 93.91% and an F1-score of 93.88%. This demonstrates the effectiveness of ensemble methods in leveraging the strengths of multiple models to improve diagnostic accuracy [80].

Limitation: Complexity in integrating and fine-tuning multiple pre-trained models may pose practical challenges.

2.6.3 Summary of Recent Studies

Following Table 2.2 provides a concise summary of recent studies on pneumonia detection using deep learning. It highlights the authors, publication years, datasets used, and key results, offering a quick reference for comparing different approaches and their effectiveness.

Table 2.2: Summary of deep learning approaches for pneumonia detection from chest X-ray images.

Paper	Authors	Methods	Data Sets	Results
[75]	Ayush et al.	ResNet-34 + U-Net	Pneumonia dataset	Accuracy: 0.90
[76]	Garima et al.	CNN	Pneumonia dataset	Accuracy: 0.9436
[77]	Mana et al.	MobileNet	Pneumonia dataset	Accuracy: 93.75%
[78]	Tatiana et al.	CycleGAN	CheXnext dataset	ROC AUC: 0.9929
[79]	Anuja et al.	Siamese-based CNN	Pneumonia dataset	Accuracy: 89.67%
[80]	Alhassan et al.	DCNN + ViT	Pneumonia dataset	Accuracy: 93.91%

Addressing these limitations, this work aims to overcome these issues by combining the strengths of DDPM for data generation and ConvNeXt for robust pseudo-labeling and classification. This approach mitigates class imbalance, reduces overfitting, allows for greater customization, and ensures high reliability without introducing significant artifacts, thereby enhancing the overall performance and applicability of pneumonia detection models.

2.7 Technical Background

In the rapidly evolving field of AI, selecting the appropriate technological approaches is crucial for achieving optimal results in various applications. This section delves into the rationale behind choosing specific models and architectures for image generation and classification tasks, focusing on the advancements that have significantly impacted the field.

2.7.1 Diffusion Models for Image Generation

Why Diffusion Models for Image Generation?

DDPM represent a groundbreaking shift in the landscape of image synthesis, providing a compelling alternative to the well-established GANs. The seminal work, 'Diffusion Models Beat GANs on Image Synthesis,' documents the superior performance of DDPM, particularly in their ability to generate high-fidelity images. Unlike GANs, which can suffer from instability and the notorious issue of mode collapse—where the model overlooks significant variations within the input data—DDPM excel in capturing the rich diversity inherent in complex datasets. They achieve this through a gradual denoising process, as illustrated in Figure 2.6, which methodically reconstructs the data distribution in a controlled manner, leading to remarkably consistent and diverse outputs. Moreover, DDPM provide a clearer and more interpretable training dynamic, as each step of the diffusion process is a simple Gaussian transformation, making the model's behavior during training easier to monitor and adjust. This inherent stability, combined with their capacity for high-quality generation without requiring adversarial training dynamics, positions DDPM as a powerful tool for image synthesis, pushing the boundaries of what is possible in creating realistic and varied synthetic imagery [81][82].

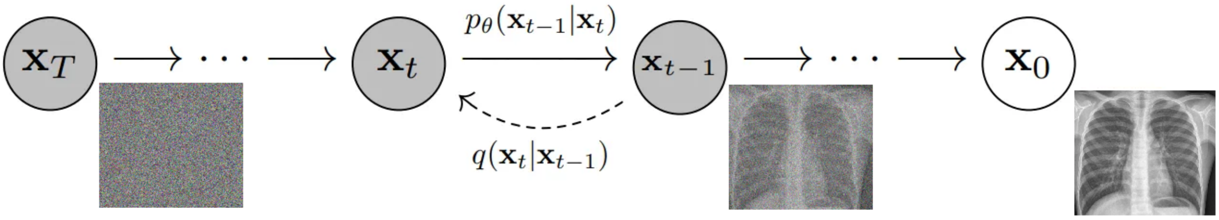


Figure 2.6: Detailed Illustration of the Denoising Diffusion Probabilistic Model (DDPM) Process [81]

Understanding Diffusion Models

Diffusion models are generative models that learn to create data samples resembling the original dataset by reversing a diffusion process. This process incrementally corrupts the input data by gradually adding Gaussian noise over a series of steps. These models are formulated as parameterized Markov chains and are trained using the principles of variational inference [83].

Training Methodology The training of diffusion models is based on minimizing the variational lower bound of the negative log-likelihood of the data. Specifically, the model learns to denoise by estimating the parameters of the reverse Markov chain [83]. This process involves a series of Gaussian transitions, as illustrated in Figure 2.7, defined as shown in Equation 2.1.

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)) \tag{2.1}$$

where x_t represents the data at timestep t, μ_θ and Σ_θ are functions modeled by the neural network that predict the mean and variance of the Gaussian distribution at each step.

U-Net Architecture in DDPM A U-Net architecture is employed to parameterize μ_θ and Σ_θ. This architecture effectively captures both local features and broader context, which are crucial for reconstructing high-quality images from heavily noised data (see Equation 2.2).

$$\mu_\theta(x_t, t), \Sigma_\theta(x_t, t) = \text{U-Net}(x_t, t) \tag{2.2}$$

The Forward Process The forward process of a diffusion model is a predefined or learned noise-adding procedure, expressed as a Markov chain (see Equation 2.3).

$$q(x_t|x_{t-1}) = \mathcal{N}(x_t; \sqrt{1-\beta_t}x_{t-1}, \beta_t I) \tag{2.3}$$

where β_t are the variance schedules determining the amount of noise added at each step. This process gradually transitions the data distribution towards a standard Gaussian distribution.

The Reverse Process The reverse process is where the model learns to reconstruct the original data by iteratively denoising (see Equation 2.4).

$$p_\theta(x_{t-1}|x_t) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, t), \Sigma_\theta(x_t, t)) \tag{2.4}$$

During training, the model optimizes the parameters θ to minimize the difference between the noisy data x_t and the predictions from the reverse process, effectively learning to denoise the data [83].

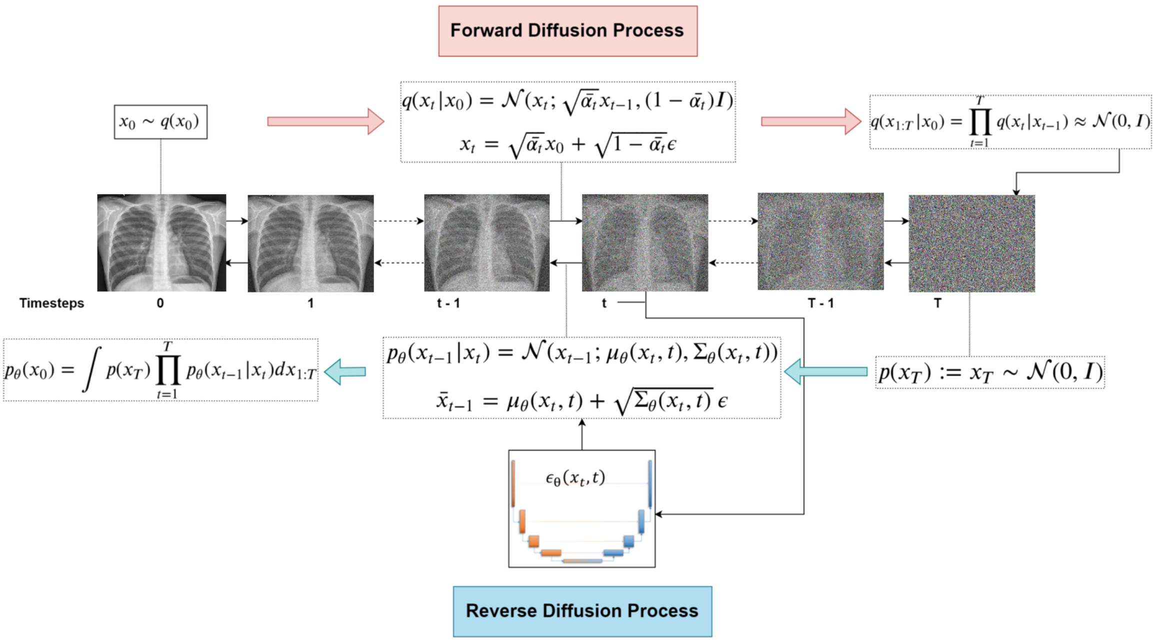


Figure 2.7: Denoising Diffusion Probabilistic Model (DDPM) Training Process [83]

Application in Medical Imaging

The profound impact of Diffusion Probabilistic Models (DPM) extends significantly into the realm of medical imaging, as demonstrated by the research findings in 'Diffusion Probabilistic

Models Beat GANs on Medical 2D Images.' This study, which has been approved by the local ethical committee (Ethical Committee (EK) 22-319), underscores the capability of DPMs to outperform GANs in generating more accurate and detailed 2D medical images. Importantly, all experiments were conducted in strict accordance with the Declaration of Helsinki and the International Ethical Guidelines for Biomedical Research Involving Human Subjects by the Council for International Organizations of Medical Sciences (CIOMS). The precision and reliability of DPMs are especially valuable in medical settings, where the clarity and detail of images directly influence diagnostic accuracy and treatment decisions. Unlike GANs, which can sometimes produce artifacts or obscure critical information, DPMs systematically reconstruct images through a noise reduction process that preserves essential features and details. This attribute, coupled with their stability during training, makes DPMs exceptionally reliable and reproducible, which is crucial in medical applications where consistency and ethical compliance are paramount [82].

2.7.2 ConvNext for Pseudo-Labeling and Classification

Understanding ConvNext?

ConvNeXt is a state-of-the-art CNNs architecture that was developed as a modern reinterpretation of the traditional CNNs to meet the demands of contemporary visual recognition tasks. Inspired by the advances in ViT, ConvNeXt integrates Transformer-like elements into the CNNs framework to enhance learning efficacy and adaptability. This architecture revisits the design choices of standard CNNs, optimizing layer configurations, scaling strategies, and normalization techniques, as illustrated in Figure 2.8. The adjustments made in ConvNeXt aim to capitalize on the strengths of both CNNs and Transformers, resulting in improved performance on a variety of computer vision benchmarks. By refining the building blocks of CNNs, ConvNeXt sets a new standard for convolutional networks, making it highly effective for tasks involving complex image recognition and processing [84].

Training ConvNeXt Models ConvNeXt models are trained using a series of strategic architectural transformations, each contributing to incremental performance improvements[84]. The learning process is encapsulated by optimizing a loss function that typically involves a

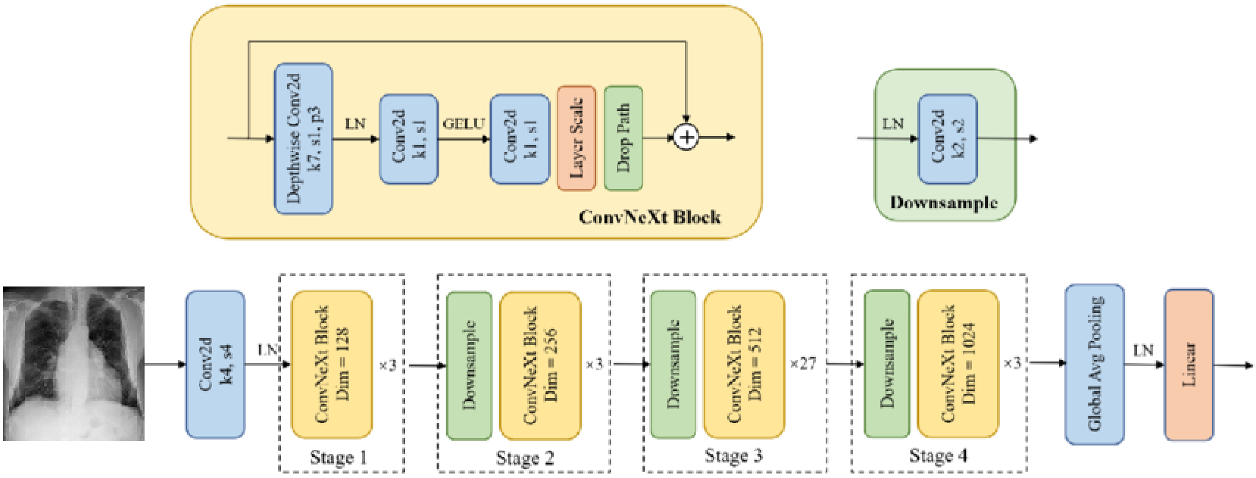


Figure 2.8: Architecture of the ConvNeXt Network [84]

softmax output over the model's predictions to ensure class probabilities(see Equation 2.5).

$$L = -\sum_i y_i \log(p_i) \tag{2.5}$$

where y_i are the target labels and p_i are the predicted probabilities by the model.

Network Design Evolution Starting from a baseline ResNet, the ConvNeXt model undergoes various modifications, enhancing its macro and micro design elements (as illustrated in Figure 2.9):

- **Stem Cell Design:** Changing from a complex to a simple "patchify" layer, which uses a 4x4 convolution with stride 4, processing input images similarly to Transformers but retaining the essence of ConvNets.
- **Macro Design Adjustments:** Adapting the compute ratio across different network stages, changing the block distribution from (3, 4, 6, 3) to (3, 3, 9, 3) to better allocate processing power.
- **Depthwise Convolution:** Integrating larger kernel sizes (5x5, 7x7, 9x9, 11x11) and depthwise separable convolutions to decrease parameter count and computation overhead, while increasing network width to maintain high performance.
- **Inverted Bottlenecks:** Implementing an inverted bottleneck design where the hidden dimension of the Multilayer Perceptron (MLP) block is four times wider than the input dimension.

- **Normalization and Activation:** Replacing BatchNorm with LayerNorm and ReLU with Gaussian Error Linear Unit (GELU) for better performance and training stability.

Forward and Backward Passes In training ConvNeXt models, both the forward and backward propagation phases are crucial. During the forward pass, input data is transformed through layers (see Equation 2.5).

$$x_{out} = f(W * x_{in} + b) \tag{2.6}$$

where W and b represent weights and biases of the convolutional layers, and f denotes an activation function such as GELU. During the backward pass, gradients are computed to update these parameters, minimizing the loss function [84].

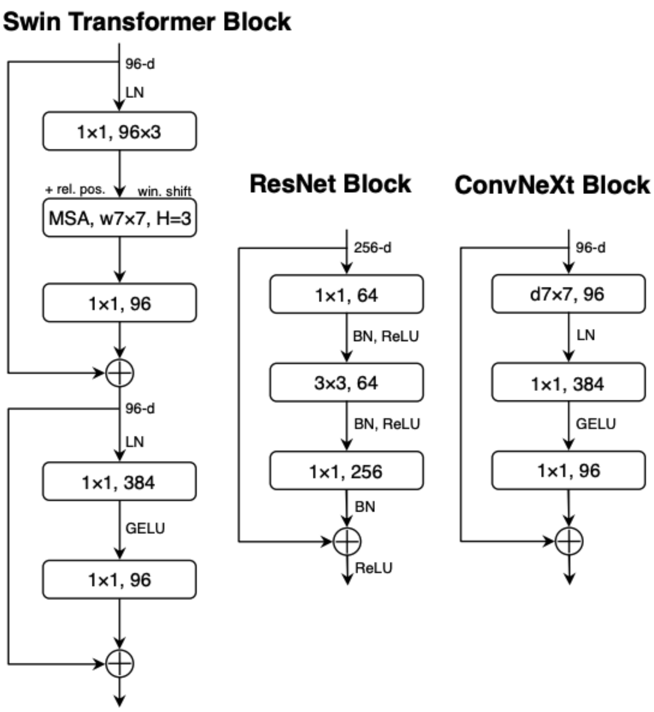


Figure 2.9: Comparison of Swin Transformer, ResNet, and ConvNeXt Blocks [84]

Why ConvNext for Pseudo-Labeling and Classification?

ConvNeXt is particularly well-suited for pseudo-labeling and classification tasks, especially when faced with the challenge of using both labeled and unlabeled data effectively. This architecture leverages the strengths of DL, minimizing the limitations often encountered with traditional models. ConvNeXt's advanced convolutional encoder is designed to excel in complex classification scenarios, providing robust feature extraction capabilities that are essential for high accuracy in semi-supervised settings. The use of ConvNeXt enables enhanced handling of sparse and varied data, a critical feature in environments where data labeling is exhaustive and inconsistent. Its capacity to adapt to SSL conditions and optimize the use of unlabeled data sets ConvNeXt apart, allowing for superior classification performance even when labeled datasets are limited. Extensive testing on multiple public datasets has demonstrated that ConvNeXt consistently surpasses other state-of-the-art DL models, showcasing its effectiveness in tasks demanding high precision and adaptability without the reliance on extensive labeled data [85].

Application in Medical Imaging

The ConvNeXt architecture marks a significant advancement in the field of medical imaging, effectively addressing the challenge of limited annotated datasets. This architecture integrates elements from Transformer blocks to modernize traditional convolutional networks. Designed for environments with scarce data, ConvNeXt employs residual blocks that adeptly manage upscaling and downscaling, maintaining semantic detail at varying scales. Additionally, it features a novel approach to incrementally increase kernel sizes, thereby overcoming performance limitations typical in medical imaging due to dataset constraints. The architecture also uses compound scaling across depth, width, and kernel size to enhance performance in various imaging tasks. With its high precision and reliability, ConvNeXt sets a new benchmark in medical imaging technology [86].

2.7.3 Approaches for Image Synthesis and Classification

Following Table 2.3 provides a concise summary of recent studies on image synthesis and classification using deep learning. It details the authors, publication years, datasets utilized, and key results achieved in each study. This summary offers a quick comparison of various methodologies and their effectiveness, highlighting the advancements in generating and classifying images through deep learning techniques. By reviewing these studies, we can gain insights into the current state of research and identify potential directions for future work in image synthesis and classification.

Table 2.3: Summary of deep learning approaches for image synthesis and classification.

Paper	Authors	Methods	Data Sets	Results
[81]	Prafulla et al.	Diffusion models	ImageNet Dataset	Achieved a record-low FID score
[82]	Gustav et al.	DDPM	CheXpert Dataset	Medfusion achieved significantly lower FID scores
[84]	Zhuang et al.	modernized ConvNet	ImageNet-1K Dataset	ConvNeXt models achieved up to 87.8%
[85]	Farouq et al.	ConvNeXt + SSL	4-Weeds Dataset	the model outperformed other state-of-the-art models

2.8 Conclusion

This chapter provides a detailed examination of AI's role in medical imaging, focusing on ML, DL, and various learning paradigms. It has highlighted the significance of SSL in medical image analysis, particularly in scenarios with limited labeled data. The introduction of advanced technological choices, such as diffusion models and ConvNeXt architectures, underscores the potential for enhanced accuracy and efficiency in medical image classification and generation. This foundational understanding sets the stage for further exploration and application of these techniques in addressing specific medical imaging challenges. Based on the current state of the art, the following chapter will be devoted to experiment with a personalized SSL-based approach toward accurate pneumonia classification using X-ray images.

Chapter 3 : Methodology and Implementation

Chapter 3

Methodology and Implementation

3.1 Introduction

The increasing prevalence of pneumonia, a critical respiratory condition, necessitates advancements in diagnostic techniques to enhance early detection and treatment outcomes. This chapter describes the innovative architecture developed for the classification of pneumonia using SSL, combining data generation through DDPM and classification via ConvNext. We will explore the detailed design of the architecture, the implementation process, and the evaluation of its performance, highlighting how these components synergistically contribute to addressing the challenges associated with limited labeled data in medical imaging. The overview of the system architecture is illustrated in Figure 3.1, which shows the process from feeding the chest X-ray image to the ConvNext classifier to obtaining the classification result.

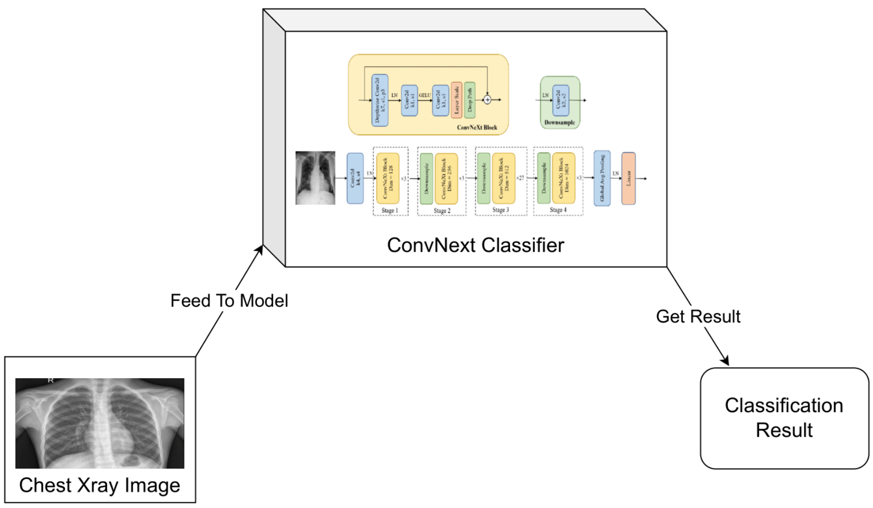


Figure 3.1: System Architecture Overview

3.2 Dataset Description

The dataset used in this research was sourced from Kaggle and consists of a total of 5856 images categorized into two classes.

3.2.1 Data Splitting and Usage

The dataset was split into training, validation, and test sets with the following proportions:

- **Test Set:** 15% (labeled)
- **Validation Set:** 15% (labeled)
- **Training Set:** 70%
    - 10% labeled
    - 60% unlabeled

Additionally, 1800 images were generated for use in training.

3.2.2 Dataset Configuration for DDPM Models

For training DDPM the entire 70% training set was used. The preprocessing steps applied to the images included resizing them to 256x256 pixels, applying random horizontal flips, converting the images to tensors, and normalizing them with a mean of 0.5 and a standard deviation of 0.5.

3.2.3 Dataset Configuration for ConvNext Models

For the ConvNext models, the data was split similarly:

- **Test Set:** 15% (labeled)
- **Validation Set:** 15% (labeled)
- **Training Set:** 70%
    - 10% labeled

- 60% unlabeled

The transformation configuration for the ConvNext models involved resizing the images to 224x224 pixels, converting them to tensors, and normalizing them using mean and standard deviation values commonly used in image classification tasks.

3.2.4 Models and Configurations

Three ConvNext models were used in this research, each with different configurations:

- **Model 1:** Trained without using generated images.
- **Model 2:** Trained using generated images.
- **Model 3:** Pre-trained and fine-tuned with generated images.

The dataset played a crucial role in training these models, with specific attention to semi-supervised learning methods for the ConvNext models. The semi-supervised learning approach involved using a combination of labeled and unlabeled data to improve the classification performance.

3.3 Detailed System Architecture

Our system architecture is designed to maximize the utility of both labeled and unlabeled data in the training process. Figure 3.2 illustrates a high-level diagram of our system, consisting of three main components: data generation using DDPM, classification using ConvNext, and integration of SSL techniques for enhancing training with pseudo-labeled data.

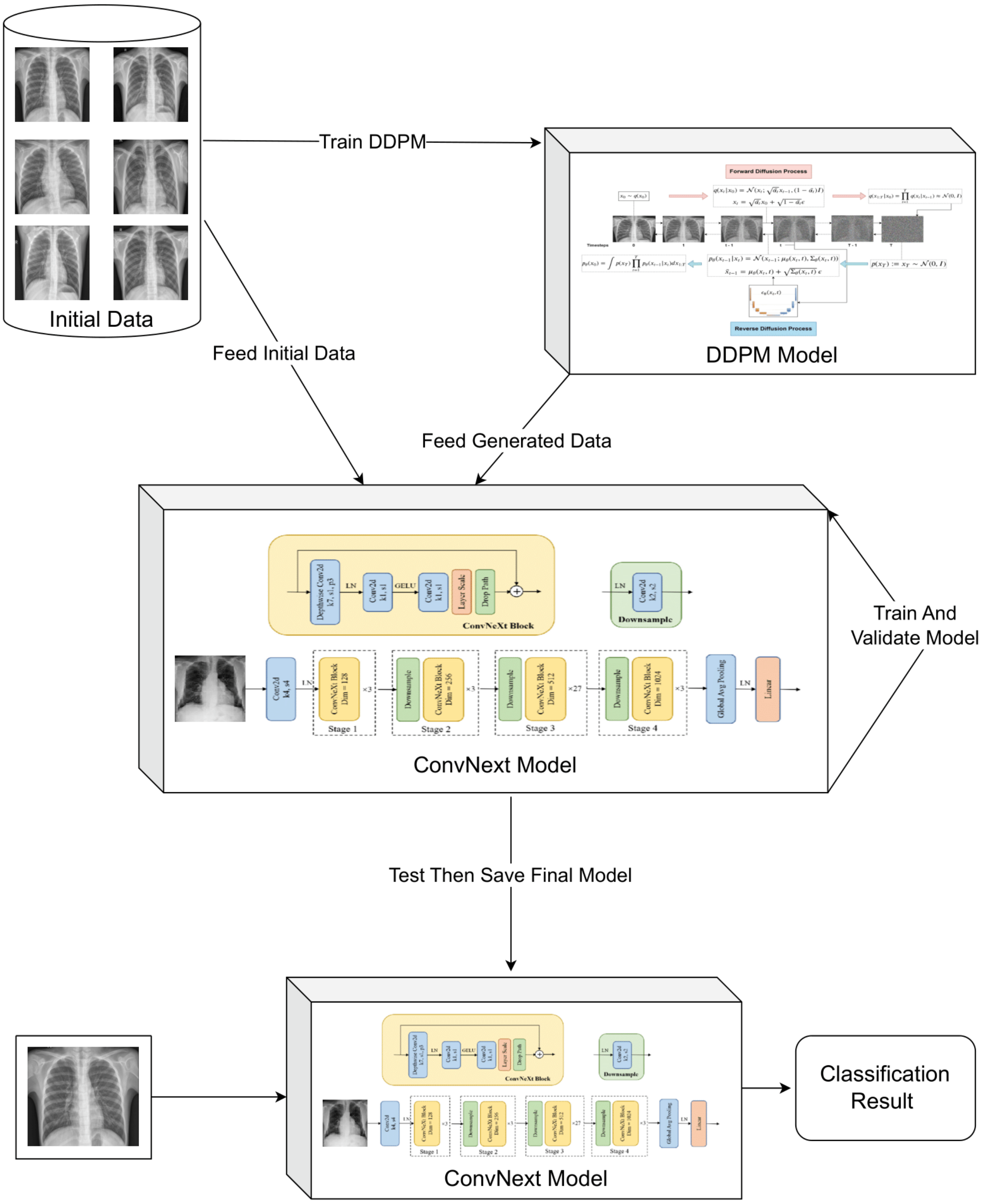


Figure 3.2: Detailed System Architecture

3.4 Detailed Design and Implementation

This section delves into the detailed design and implementation of the main two primary components: DDPM for image generation and the ConvNext model for image classification. The integration of these components forms the core of the proposed approach to enhancing medical image analysis through semi-supervised learning.

3.4.1 DDPM Generation

DDPM are chosen for their ability to generate realistic and diverse images from noise by modeling the data distribution of existing medical images. This capability is crucial for augmenting dataset, where labeled examples are scarce but critical for training robust models. As illustrated in Figure 3.3, the DDPM model follows a detailed training process to achieve this generation capability.

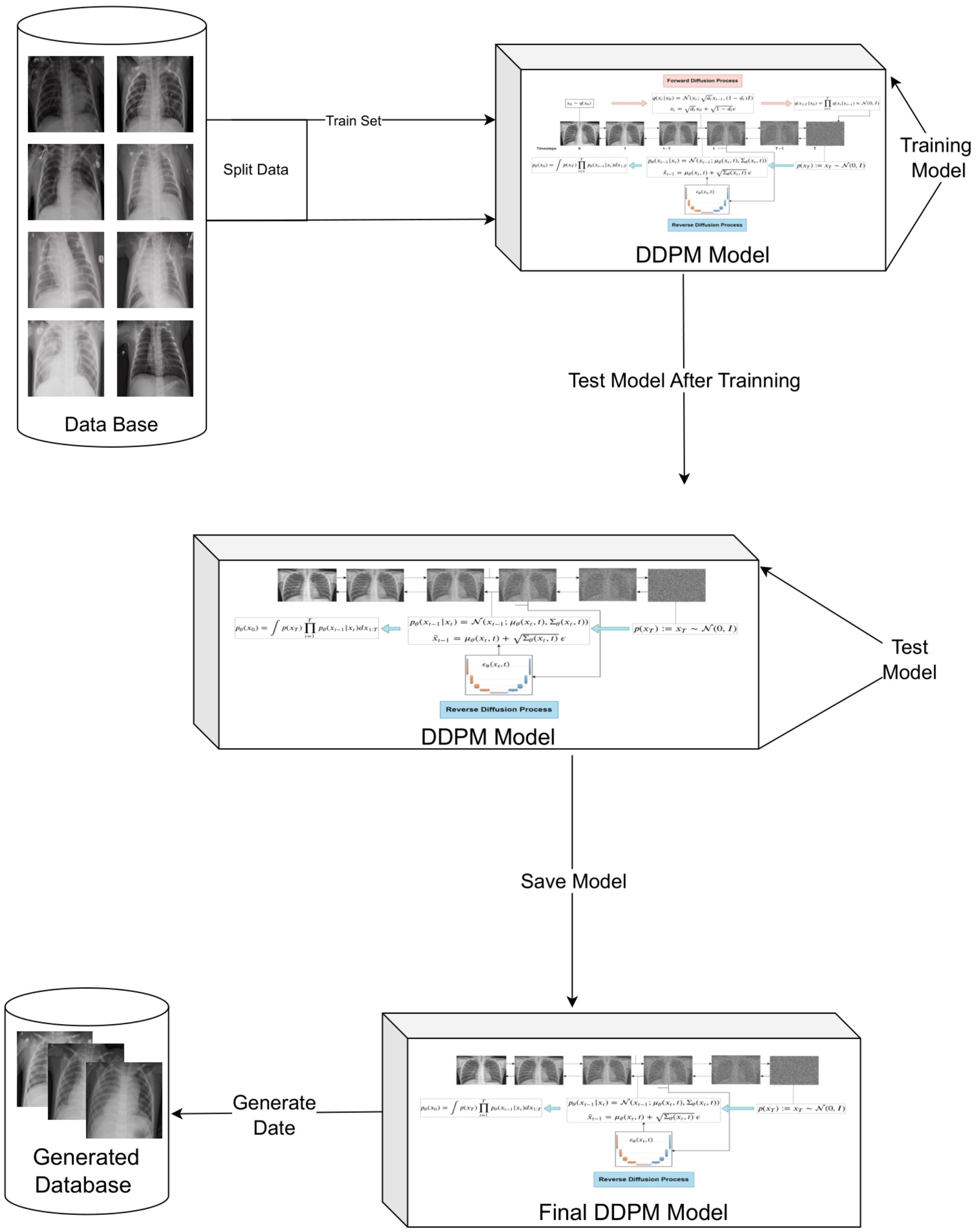


Figure 3.3: Detailed Design of DDPM Model Training

A. Development Path:

- **Initial Model (128x128):** The initial DDPM model was developed with a resolution of 128x128 pixels. However, the generated images did not meet our quality expectations.

- **Improved Model (256x256 with One Attention Layer):** The resolution is increased to 256x256 pixels and added one attention layer in the UNet architecture. Despite these improvements, the generated images still lacked the desired quality.

- **Final Model (256x256 with Two Attention Layers):** After consulting with medical experts, who demanded higher quality images, a second attention layer is incorporated to the model. This adjustment significantly enhanced the image quality, and the generated images were finally approved by the doctors for use as synthetic chest X-ray, as shown in Figures : 3.4 and 3.5.

B. Model Configuration:

- Batch size: 8

- Learning rate: 1e-4

- Epochs: 40

- Noise schedule: DDPMScheduler

C. Training Process:

The DDPM training involves processing a dataset of 4977 chest X-ray (4099 for training and 878 for validation). The model learns to reverse the process of adding Gaussian noise to clean images, gradually generating clean images from noisy inputs. This process enables the model to create high-quality medical images from noise by the end of training.

Confirmation of Generated Chest X-ray Images

Student Information

- **Thesis Title:** Semi Supervised Learning for Medical Images
- **Student's Name:** Lahcene Mamen
- **Supervisor's :** Pr.Laid Kahloul, Dr.Asma Ammari
- **University:** University of Mohamed Khider Biskra
- **Date:** 24/05/2024

Introduction: This document is to inform about the use of Diffusion Probabilistic Models (DDPM) in generating synthetic chest X-ray images as part of my thesis work. The goal of this work is to augment the existing dataset to enhance the training process of a ConvNext classifier for improved diagnostic performance.
Method and Results: We used DDPM to generate synthetic chest X-ray images. These images were created to increase the diversity of our training dataset. The augmented data was then used to train a ConvNext classifier, resulting in improved model accuracy.
Sample Results: Below are some results from the DDPM-generated images:

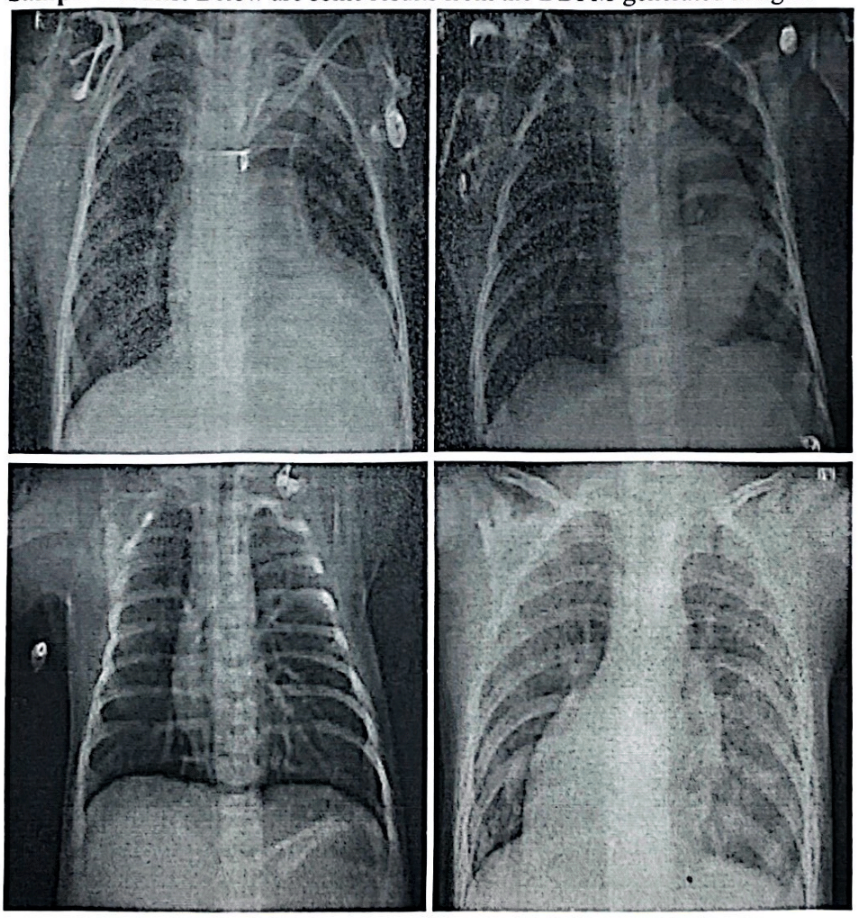
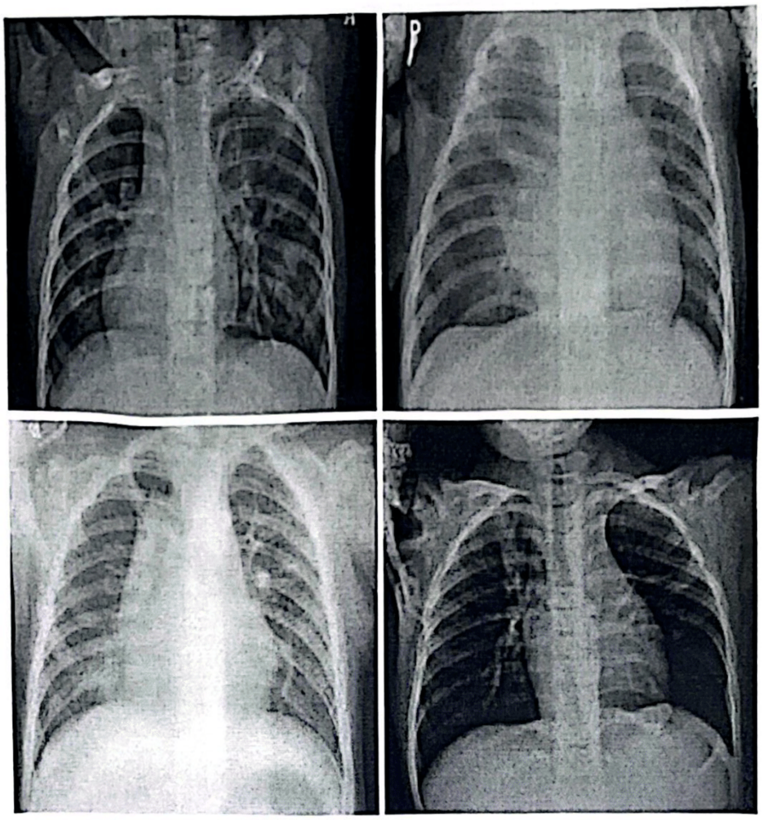


Figure 3.4: Medical Approval: Part 1



Doctor's Acknowledgement The undersigned doctors, specialists in their respective fields, have reviewed the generated images. By signing below, they confirm that the generated images appear realistic and can be used as part of the augmented dataset in the thesis.

Doctor's Information

Avis d'un chirurgien Thoracique :
Des image de Radiographie thoracique simule des image
Réales dans 80%..

DOCTEUR
DOURAKH KAISS
Chirurgien Thoracique

Avis de Medecine générale :
· Des images (Rx) facile et clair à interpréter
et aide de préciser le diagnostic

Dr: Tachlissia Radhouane
Docteur en Médecine

Figure 3.5: Medical Approval: Part 2

3.4.2 ConvNext for Classification

ConvNext, a modern CNNs architecture, is utilized for its efficiency and effectiveness in image classification tasks. The model is trained on labeled data to establish a baseline ability to identify and classify pneumonia from chest X-ray.

3.4.3 Integration of Semi-Supervised Learning

To leverage unlabeled data, the generated data is integrated with the unlabeled pool during the training loop, as illustrated in Figure 3.6. This allows pseudo-labeling to occur dynamically as part of the training process. The model's predictions on the unlabeled data are treated as pseudo-labels and used to continuously refine and improve the model's performance, effectively expanding the labeled dataset without additional expert annotation.

A. Model Setup:

- Batch size: 16
- Learning rate: 5e-5
- Epochs: 100

B. Development Path:

- **Initial Approach (7% Labeled, 63% Unlabeled):** A first step correspond to training the ConvNext model using 7% of the dataset as labeled data (409 images) and 63% as unlabeled data (3690 images). The validation dataset consisted of 15% of the images (878 images), and the test dataset also consisted of 15% of the images (879 images).
- **Enhanced Approach (7% Labeled, 63% Unlabeled with DDPM Images):** The DDPM-generated images (1800 images) is then included into the unlabeled dataset. This integration improved the model's performance significantly. The total unlabeled dataset size became 5490 images. The validation dataset consisted of 15% of the images (878 images), and the test dataset also consisted of 15% of the images (879 images).

- **Best Approach (Pretrained Model with 7% Labeled, 63% Unlabeled with DDPM Images):** Finally, a pretrained ConvNext-tiny-224 model is used, adopting the same configuration as before, with 7% labeled data (350 images) and 63% unlabeled data (5490 images), including our DDPM-generated images (1800 images). This approach yielded the best results. The validation dataset consisted of 15% of the images (879 images), and the test dataset also consisted of 15% of the images (879 images).

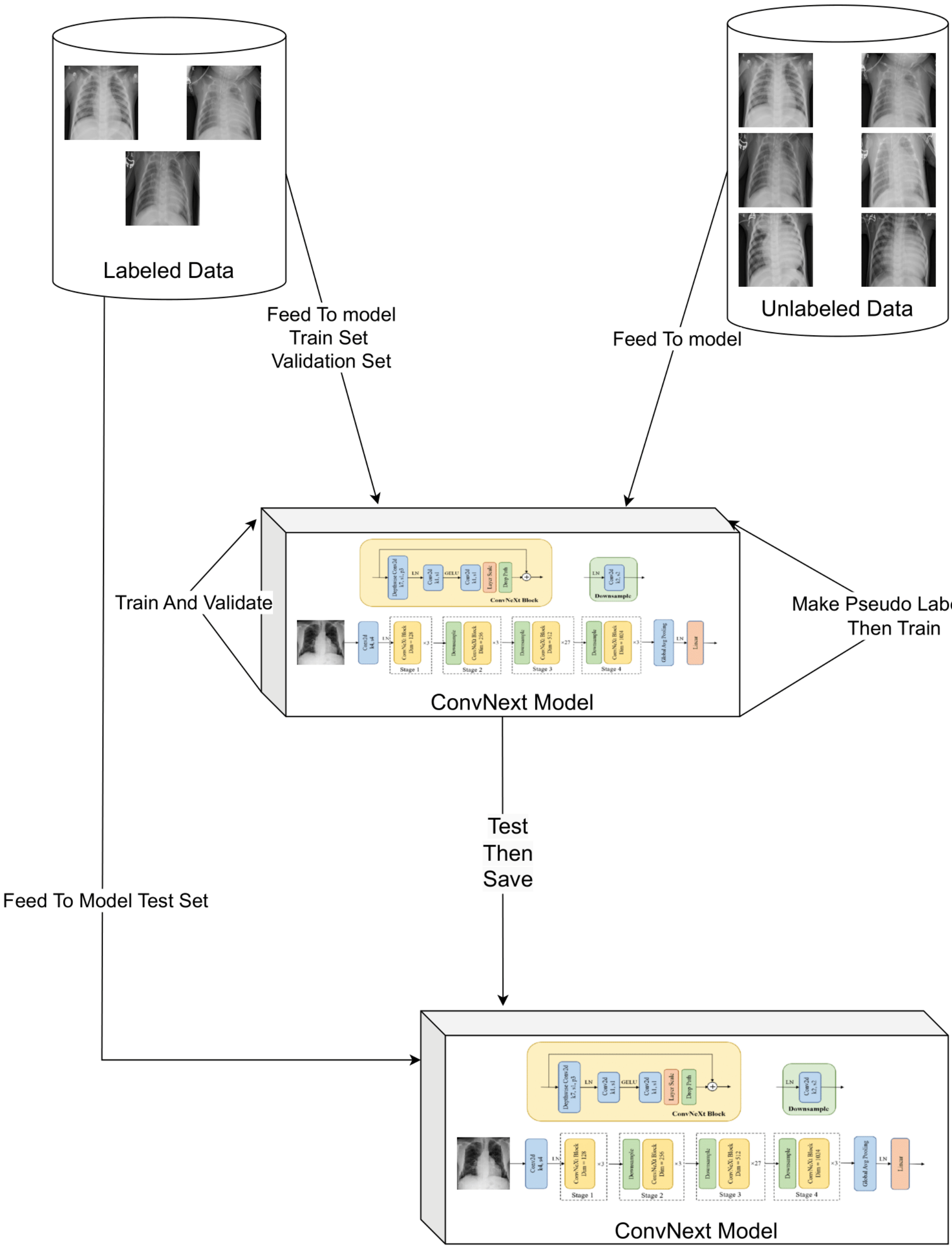


Figure 3.6: Detailed Design of ConvNext Model Training

3.5 Data Flow and Processing

Data flow through the current approach is meticulously designed to ensure seamless integration and processing. Initially, all available data is split into training, validation, and test datasets. The training set is further divided into labeled and unlabeled pools. Post DDPM training, the generated data is combined with the unlabeled pool, and the enhanced dataset re-enters the training cycle, allowing the ConvNext model to be exposed to varied data scenarios, as illustrated in Figure 3.7, improving its diagnostic accuracy.

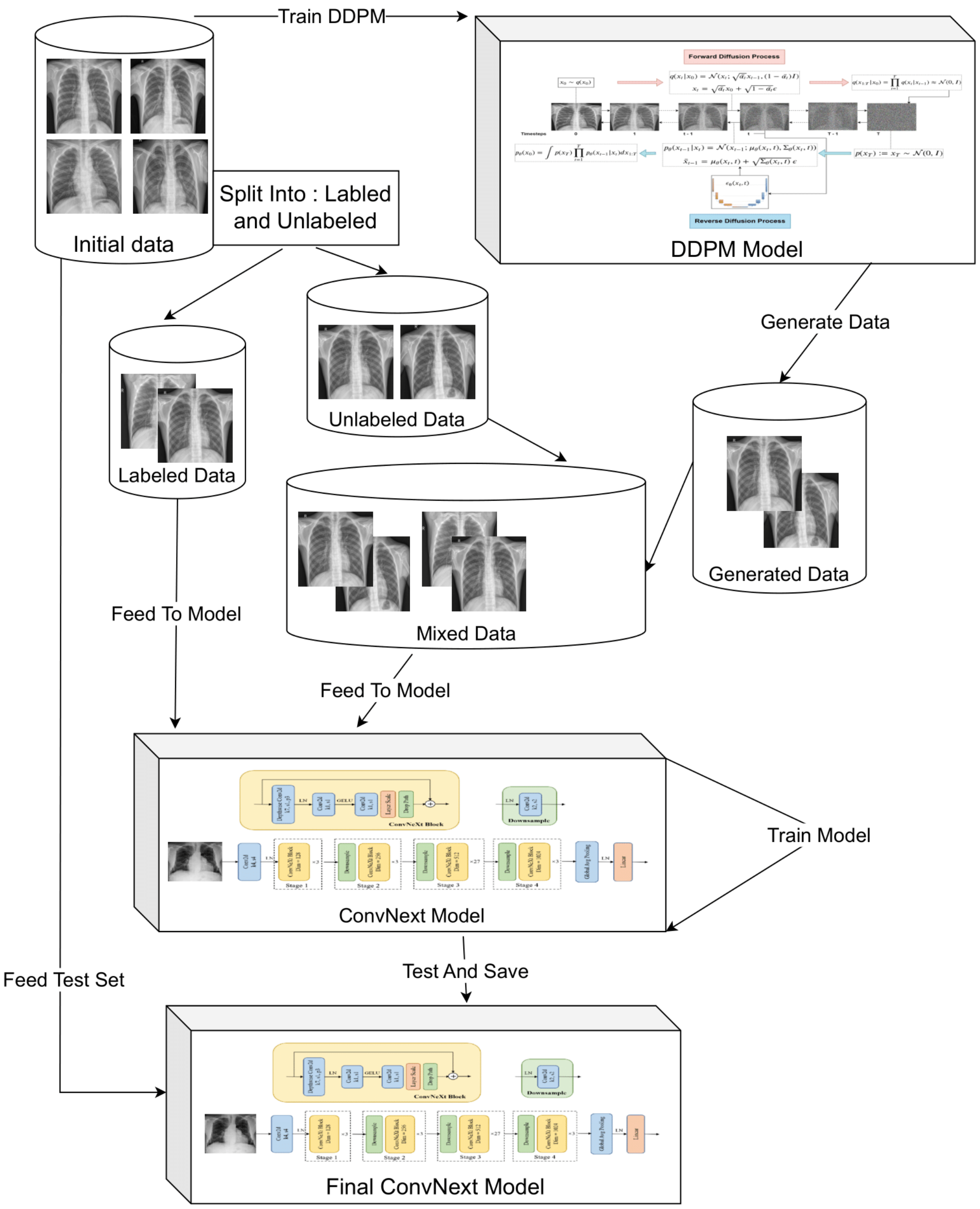


Figure 3.7: Data Flow Design

3.6 Design Challenges and Considerations

Several challenges were encountered in the design phase, including the stability of DDPM during training and ensuring the quality of generated images. Furthermore, integrating pseudo-labeled data without introducing significant noise into the training process was critical. These challenges were addressed through iterative testing and refinement of model parameters and thresholds for pseudo-labeling.

3.7 Implementation and Evaluation

The implementation and system testing phase focused on validating the entire pipeline, ensuring seamless interaction between components, and optimizing the performance of the system. This phase was crucial in identifying and resolving any integration issues that could impact the stability and efficiency of the training and inference processes.

3.7.1 Hardware and Software Requirements

The Training of this system divided into:

A. Hardware For DDPM:

- Platform: Google Colab Pro+
- Graphics Processing Unit (GPU): NVIDIA A100 with 40 GB memory
- Random Access Memory (RAM): 64 GB
- Storage: 2TB Solid State Drive (SSD)

B. Hardware For ConvNext:

- GPU: NVIDIA RTX 4070
- Central Processing Unit (CPU): Intel Core i7 13th Gen
- RAM: 64 GB
- Storage: 2TB SSD

3.7.2 Configuration Settings

- Operating System: Linux-based distribution
- Programming Language: Python 3.8
- DL Framework: PyTorch 1.7
- Key Libraries:
  - **torch**: The core library of PyTorch, providing tensors and dynamic neural networks in Python with strong GPU acceleration.
  - **torchvision**: Provides datasets, model architectures, and image transformations for computer vision tasks. It's widely used for preprocessing images and loading popular vision datasets.
  - **numpy**: A fundamental library for numerical computations in Python. It supports large, multi-dimensional arrays and matrices, along with a collection of mathematical functions to operate on these arrays.
  - **scikit-learn**: A ML library that provides simple and efficient tools for data mining and data analysis. It's built on NumPy, SciPy, and matplotlib and is used for implementing standard ML algorithms.
  - **diffusers**: A library for diffusion probabilistic models, useful for generating high-quality images from noise. It includes tools for training and using diffusion models in various applications, including image generation.
  - **transformers**: Developed by Hugging Face, this library provides state-of-the-art general-purpose architectures for natural language understanding and generation. It includes models like BERT, GPT, and many others.
  - **datasets**: Also developed by Hugging Face, this library offers a simple way to share and load datasets. It provides access to a wide variety of datasets and makes it easy to prepare data for ML tasks.
  - **accelerate**: Another Hugging Face library, designed to streamline and optimize the training of DL models across different hardware configurations,

such as CPU, GPU, and TPUs. It allows for easier scaling and distributed training.

3.7.3 Integration Testing

Comprehensive testing was conducted to ensure data flowed correctly through all approach steps, from data generation to classification. Integration points were rigorously tested to handle discrepancies in data formats or quality.

3.8 Results and Analysis

In this section, we delve into the results and analysis of the investigated study, focusing on the performance metrics and insights gained from the integration of DDPM and SSL techniques. The aim is to evaluate the effectiveness of these advanced methodologies in enhancing the accuracy and generalization of the ConvNext model for pneumonia classification. Detailed results, including training and validation metrics, are presented to highlight the significant improvements achieved through our approach.

3.8.1 Presentation of Results

The system achieved an accuracy of 97% along with a precison equal to 98.6% on both validation set and testing set, marking a significant improvement over traditional methods with using just 7% labeled data from whole dataset. Therefore, integrating DDPM-generated data and utilizing SSL significantly improved model generalization on unseen data. In the remaining Sub-sections more details about the achieved results are discussed.

3.8.2 Training and Validation Metrics

A. DDPM:

- Training Loss: The average loss during the training phase of DDPM demonstrates rapid convergence, with a steep initial decline and stabilization at low levels, indicating effective training, as illustrated in Figure 3.8.

· Validation Loss: The average loss during the validation phase closely follows the training loss, suggesting minimal overfitting and good generalization of the DDPM model, as illustrated in Figure 3.8

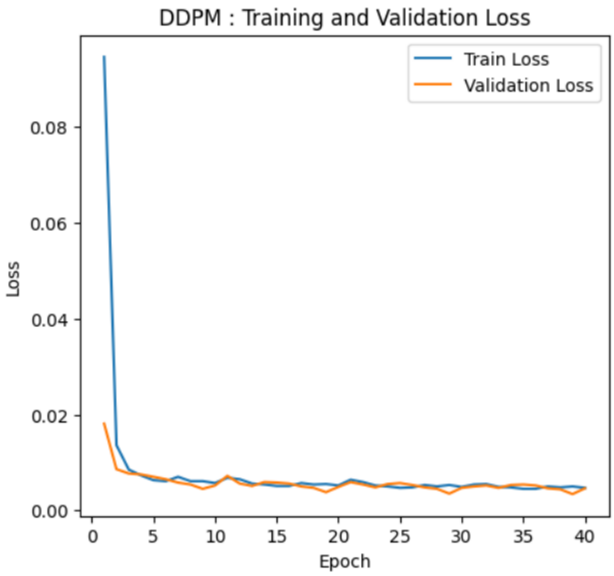


Figure 3.8: Training and Validation Loss for DDPM

B. ConvNext:

· Training Loss: The average loss during the training phase of ConvNext, showing a decline over time indicative of learning.

· Validation Loss: The average loss during the validation phase for ConvNext, which tends to be higher than when combined with DDPM, indicating potential model limitations when standalone (refer to the right graph in Figure 3.9).

· Training Accuracy: The accuracy on the training dataset for ConvNext increases over epochs, reflecting the model's ability to learn from the training data.

· Validation Accuracy: The accuracy on the validation dataset for ConvNext, which is lower compared to when enhanced with DDPM, suggesting improved performance with the integration of DDPM (see the left graph in Figure 3.9).

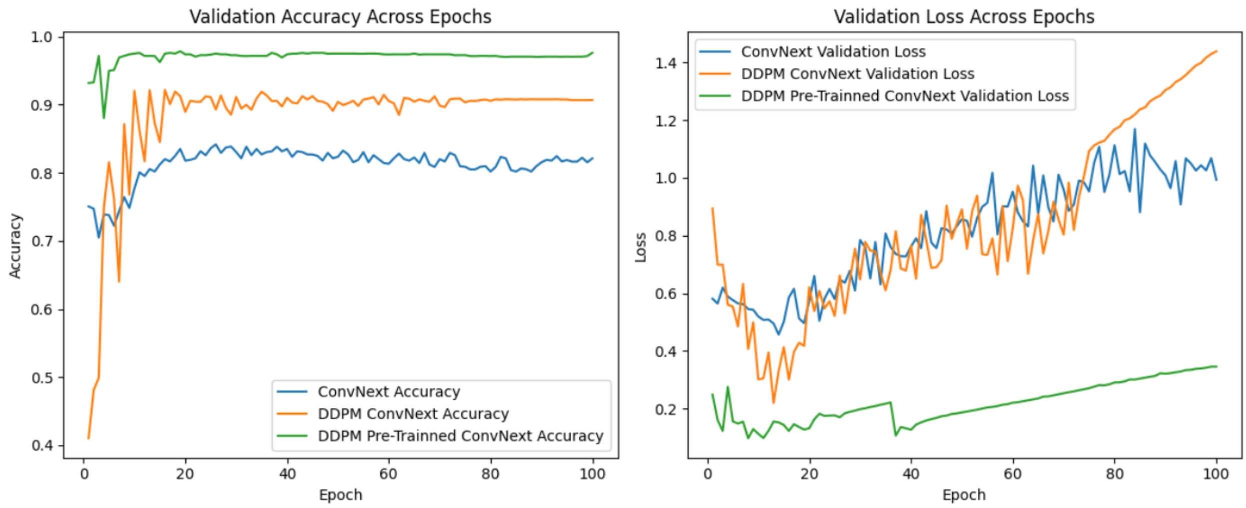


Figure 3.9: Validation Loss and Accuracy for Different Methods

3.8.3 Empirical Result

To assess the performance of the current study, we utilized several key metrics, each quantifying different aspects of model performance. Below are the used metrics and their respective mathematical formulations:

Metrics For DDPM:

· Frechet Inception Distance (FID): Measures the similarity between two datasets of images.

· Learned Perceptual Image Patch Similarity (LPIPS): Measures perceptual similarity between images.

· Peak Signal-to-Noise Ratio (PSNR): Measures the ratio between the maximum possible power of a signal and the power of corrupting noise (see Equation 3.1).

$$\text{PSNR} = 20 \cdot \log_{10} \left(\frac{\text{MAX}_I}{\sqrt{\text{MSE}}} \right) \tag{3.1}$$

where MAX_I is the maximum possible pixel value of the image and MSE is the mean squared error.

· Structural Similarity Index (SSIM): As shown in Equation 3.2 measures the similarity between two images.

$$\text{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{3.2}$$

Result Of Generated Image Using DDPM:

The performance scores for the generated chest X-ray images using the DDPM are presented in the Table 3.1. It summarizes the performance of generated images at different resolutions and configurations, including the use of attention mechanisms.

Table 3.1: Performance Scores for Generated Images

Image Type	FID	LPIPS	PSNR	SSIM
DDPM 128x128	111.85	0.4211	9.304	0.4326
DDPM 256x256 With 1 Attention	50.376	0.3434	10.405	0.4252
DDPM 256x256 With 2 Attention	**21.579**	**0.3178**	**12.869**	**0.4365**

As illustrated in Table 3.1, the performance metrics include the Fréchet Inception Distance (FID), Learned Perceptual Image Patch Similarity (LPIPS), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index Measure (SSIM). The results indicate that higher resolution images (256x256) with attention mechanisms significantly improve the quality of generated images, as reflected by lower FID and LPIPS scores, and higher PSNR and SSIM values. Notably, the DDPM model with two attention mechanisms achieves the best performance across all metrics.

Metrics For ConvNext:

· Accuracy: The proportion of total predictions that were correct (see Equation 3.3).

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3.3}$$

where TP is True Positives, TN is True Negatives, FP is False Positives, and FN is False Negatives.

- Precision: The ratio of correctly predicted positive observations to the total predicted positives (see Equation 3.4).

$$\text{Precision} = \frac{TP}{TP + FP} \tag{3.4}$$

- Recall (Sensitivity): The ratio of correctly predicted positive observations to all observations in the actual class (see Equation 3.5).

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3.5}$$

- F1-Score: The weighted average of Precision and Recall, useful in cases of uneven class distribution (see Equation 3.6).

$$\text{F1-Score} = 2 \times \left(\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \right) \tag{3.6}$$

- Specificity: The ratio of correctly predicted negative observations to all actual negatives (see Equation 3.7).

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{3.7}$$

- False Positive Rate (FPR): The ratio of incorrectly predicted positive observations to all actual negatives(see Equation 3.8).

$$\text{FPR} = \frac{FP}{FP + TN} \tag{3.8}$$

Result Of Classification Using ConvNext: The ConvNext model, trained with both real and pseudo-labeled images generated by DDPM, showed significant improvement in classification metrics compared to the baseline model trained only with real labeled data.

Confusion Matrices: The performance improvement is clearly illustrated in the confusion matrices shown in Figure 3.10. The ConvNext model alone demonstrates a balanced performance, but when enhanced with DDPM-generated pseudo-labeled data, the ConvNext + DDPM model exhibits a marked reduction in false positives and false negatives. Specifically, the true positives increased from 620 to 659, and the true negatives from 131 to 154. Furthermore, the Pre-Trained ConvNext + DDPM model shows the best performance among the three configurations, with the highest true positive (669) and true negative (186) counts, and the lowest false positive (8) and false negative (16) counts. These results underscore the efficacy of incorporating DDPM-generated data and pre-training in improving classification accuracy.

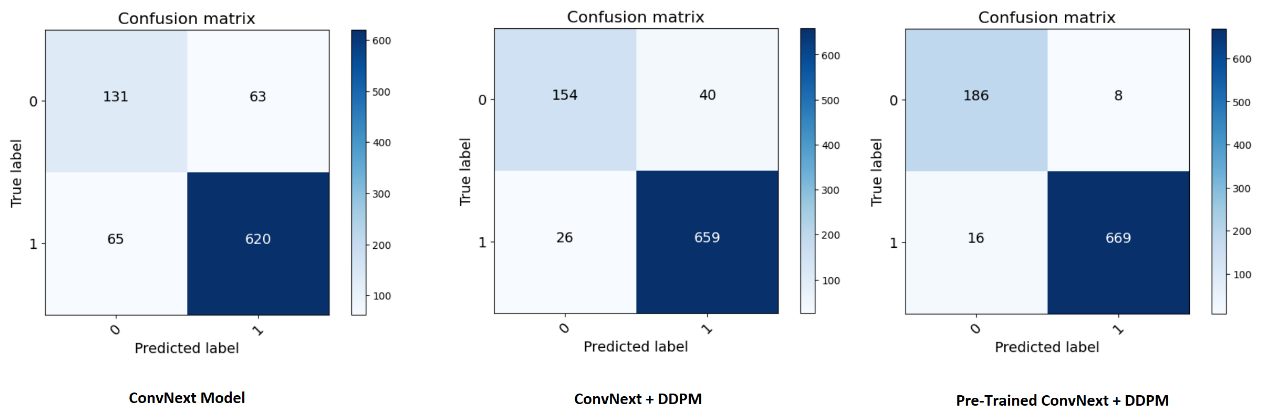


Figure 3.10: Confusion matrices for the three model configurations: ConvNext Model, ConvNext + DDPM, and Pre-Trained ConvNext + DDPM.

Performance Metrics

- **Accuracy:** Improved from 85.4% in the baseline to 92% and up to 97% with Pre-Trained ConvNext + DDPM.
- **Precision:** Increased from 90.8% to 88%, and further to 98.6% with Pre-Trained ConvNext + DDPM.
- **Recall:** Rose from 90.5% to 90%, and further to 97.6% with Pre-Trained ConvNext + DDPM.
- **F1-Score:** Enhanced from 90.6% to 89%, and up to 98.1% with Pre-Trained ConvNext + DDPM.
- **Specificity:** Reached 90.5%, consistent with Pre-Trained ConvNext + DDPM results.
- **False Positive Rate (FPR):** Reduced from 32.5% in the baseline to 20.6%, and further to 4.02% with Pre-Trained ConvNext + DDPM.

3.9 Performance Comparison

In this section, the effectiveness of the proposed approach is evaluated and discussed based on various metrics. Accordingly, Table 3.2 presents the performance comparison of different model configurations, specifically ConvNext, ConvNext+DDPM, and Pre-Trained ConvNext+DDPM (PT-ConvNext+DDPM). The evaluation metrics include Accuracy, Precision, Recall, F1 Score, Specificity, and False Positive Rate (FPR). The results demonstrate a significant improvement in performance when using the DDPM enhancement, with the PT-ConvNext+DDPM configuration achieving the highest scores across all metrics.

Table 3.2: Performance Comparison of Model Configurations

Metric	ConvNext	ConvNext+DDPM	PT-ConvNext+DDPM
Accuracy	0.8544	0.9249	0.9716
Precision	0.9078	0.9428	0.9861
Recall	0.9051	0.9620	0.9756
F1 Score	0.9064	0.9523	0.9808
Specificity	0.9051	0.9620	0.9756
FPR	0.3247	0.2062	0.0402

3.10 Comparison with Existing Research

The current investigated approach extends existing research in several key ways. The study "Pneumonia Detection: An Efficient Approach Using Deep Learning" by Ayush Pant et al. (2020) achieved an accuracy of 90% using an ensemble of ResNet-34 based U-Net and EfficientNet-B4 based U-Net. While effective, this method faced challenges in handling class imbalance. In contrast, our use of DDPM for data augmentation addresses class imbalance more effectively, leading to a higher accuracy.

Similarly, Garima Verma et al. (2020) reported an accuracy of 94.36% using a CNN built from scratch. While impressive, this approach may suffer from potential overfitting due to the lack of pre-trained models. Our method, by leveraging ConvNext and SSL, achieves better generalization with less risk of overfitting.

Mana Saleh Al Reshan et al. (2023) utilized MobileNet, achieving an accuracy of 93.75%, balancing computational efficiency and accuracy. However, their reliance on pre-trained models limits customization. Our approach, combining DDPM and ConvNext, provides flexibility and customization while maintaining high accuracy.

Tatiana Malygina et al. (2019) employed GANs for data augmentation, improving ROC AUC scores from 0.9745 to 0.9929. While effective, GAN-generated data can introduce artifacts. Our use of DDPM mitigates this risk by generating high-quality images approved by medical experts.

Finally, the ensemble learning approach by Alhassan Mabrouk et al. (2023) achieved an accuracy of 93.91% using a combination of DenseNet169, MobileNetV2, and Vision Transformer models. While effective, the complexity of integrating multiple models poses practical challenges. The proposed approach experimented in this study simplifies the architecture while achieving higher accuracy.

Table 3.3 provides a comparative summary of these deep learning approaches for pneumonia detection from chest X-ray images, comparing previous works with the current proposed method.

Table 3.3: Summary of deep learning approaches for pneumonia detection from chest X-ray images.

Paper	**Year**	**Results**
[75]	2020	Accuracy: 90.00%
[76]	2020	Accuracy: 94.36%
[77]	2023	Accuracy: 93.75%
[79]	2020	Accuracy: 89.67%
[80]	2023	Accuracy: 93.91%
Proposed approach	2024	Accuracy: 97.16%

3.11 Conclusion

This chapter presents an architecture for pneumonia classification using SSL. The approach combines DDPM for data generation and ConvNext for classification, effectively using both labeled and unlabeled data. Iterative DDPM development improved image quality, and SSL with pseudo-labels enhanced performance. Challenges like DDPM stability and integrating pseudo-labeled data are also addressed through testing and refinement. The proposed method achieved 97% accuracy, and 98.6% precision on validation and testing sets. The discussed results has proven to overcome the existing research.

Conclusion and Perspectives

This thesis presents an innovative architecture for pneumonia classification using semi-supervised SSL techniques, combining DDPM for data generation and ConvNext for classification. The architecture effectively utilized both labeled and unlabeled data, achieving a 98.6% as a precision on validation and testing sets. Key contributions from this study include the iterative development of the DDPM, involving increased resolution and added attention layers, which successfully produced high-quality synthetic chest X-ray images [83][72]. These images were validated by medical experts, ensuring their utility in augmenting the training dataset. The ConvNext model, when integrated with SSL, demonstrated significant improvements in pneumonia classification accuracy compared to traditional methods. The model's performance was notably enhanced by the inclusion of DDPM-generated images. Furthermore, the incorporation of pseudo-labeled data into the training process allowed the model to leverage a large amount of unlabeled data, significantly improving its generalization and robustness [53].

The corresponding results have several important implications for the field of medical imaging. The significant improvement in classification accuracy suggests that SSL, combined with synthetic data generation [73], can effectively address the challenge of limited labeled data in medical imaging [53]. This can lead to more accurate and early diagnosis of pneumonia, ultimately improving patient outcomes [3]. Additionally, the ability to generate high-quality synthetic images reduces the dependency on extensive labeled datasets, which are often expensive and time-consuming to produce [63]. This approach can democratize access to advanced diagnostic tools, particularly in resource-constrained settings. While this study focused on pneumonia classification, the underlying principles of combining DDPM and ConvNext with SSL can be applied to other medical imaging tasks, such as tumor detection, organ segmentation, and more.

Despite the promising results, this study has several limitations. The synthetic images generated by the DDPM, although validated by medical experts, may not capture the full diversity of real-world medical images. This limitation could affect the generalizability of the model. Furthermore, the training and generation processes for both DDPM and ConvNext require significant computational resources, which may limit the practical deployment of the system in low-resource settings. Moreover, the methods and models were specifically optimized for pneumonia classification. Adapting the architecture to other medical conditions may require additional modifications and validations.

Building on the results of this study, several avenues for future research can be explored. Investigating the applicability of the combined DDPM and ConvNext architecture to other medical imaging tasks, such as cancer detection, cardiovascular imaging, and neurological imaging, could extend the impact of this work. Developing more efficient algorithms and architectures to reduce the computational complexity and resource requirements of the system would make it more accessible for broader clinical use. Evaluating the effectiveness of other SSL techniques, such as self-training, co-training, and multi-view learning, in improving model performance could further enhance the methodology. Conducting longitudinal studies to assess the long-term impact and reliability of using synthetic data in clinical settings is crucial for ensuring that these methods remain robust over time and across diverse patient populations. Collaborating with healthcare institutions to implement and test the proposed architecture in real-world clinical environments, gathering feedback, and data to further refine and optimize the system, is also an important step forward.

In conclusion, this research has demonstrated the potential of combining DDPM and ConvNext within a SSL framework to significantly enhance pneumonia classification accuracy. The successful integration of synthetic data generation and SSL addresses key challenges in medical imaging, particularly the scarcity of labeled data. The obtained results underscore the importance of leveraging advanced AI techniques to improve diagnostic accuracy and patient outcomes. Future research should continue to explore and refine these methods, ensuring their broader applicability and deployment in diverse clinical settings. The contributions of this study mark a significant step forward in the field of medical imaging, paving the way for more accessible and accurate diagnostic tools.

Bibliography

[1] Xiangyu Ou, Xue Chen, Xianning Xu, Lili Xie, Xiaofeng Chen, Zhongzhu Hong, Hua Bai, Xiaowang Liu, Qiushui Chen, Lin Li, and Huanghao Yang. Recent development in x-ray imaging technology: Future and challenges. *Research*, 2021:Article ID 9892152, 2021. https://doi.org/10.34133/2021/9892152.

[2] Rushi Jiao, Yichi Zhang, Le Ding, Bingsen Xue, Jicong Zhang, Rong Cai, and Cheng Jin. Learning with limited annotations: A survey on deep semi-supervised learning for medical image segmentation. *Computers in Biology and Medicine*, 169, 2024. https://doi.org/10.1016/j.compbiomed.2023.107840.

[3] Gillian E Doe, Marie T Williams, Stacey Chantrell, Michael C Steiner, Natalie Armstrong, Ann Hutchinson, and Rachael A Evans. Diagnostic delays for breathlessness in primary care: a qualitative study to investigate current care and inform future pathways. *British Journal of General Practice*, 73(731):e468–e477, 2023. https://doi.org/10.3399/BJGP.2022.0475.

[4] Qiuyu An, Wei Chen, and Wei Shao. A deep convolutional neural network for pneumonia detection in x-ray images with attention ensemble. *Diagnostics*, 14:390, 2024. https://doi.org/10.3390/diagnostics14040390.

[5] Shah Hussain, Iqra Mubeen, Niamat Ullah, Syed Shahab Ud Din Shah, Bakhtawar Abduljalil Khan, Muhammad Zahoor, Riaz Ullah, Farhat Ali Khan, and Mujeeb A. Sultan. Modern diagnostic imaging technique applications and risk factors in the medical field: A review. *BioMed Research International*, 2022. https://doi.org/10.1155/2022/5164970.

[6] Michael Friebe and Alfredo Illanes. Advancements in medical imaging and image-guided procedures: A potential—or rather likely—paradigm shift in diagnosis and therapy: Un-

derstand disruption and take advantage of it! *Applied Sciences*, 13(16):9218, 2023. https://doi.org/10.3390/app13169218.

[7] Amir M. Owrangi, Peter B. Greer, and Carri K. Glide-Hurst. Mri-only treatment planning: benefits and challenges. *Phys Med Biol*, 63(5):05TR01, 2019. https://doi.org/10.1088/1361-6560/aaaca4.

[8] John Rego and KM Tan. Advances in imaging—the changing environment for the imaging specialist. *The Permanente Journal*, 10(1):26–28, 2006. https://doi.org/10.7812/tpp/05-118.

[9] Bradford C. Dickerson. Advances in functional magnetic resonance imaging: Technology and clinical applications. *Neurotherapeutics*, 4(3):360–370, 2007. https://pubmed.ncbi.nlm.nih.gov/17599702/.

[10] Cynthia H. McCollough, Shuai Leng, Lifeng Yu, and Joel G. Fletcher. Dual- and multi-energy ct: Principles, technical approaches, and clinical applications. *Radiology*, 276(3):637–653, 2015. https://doi.org/10.1148/radiol.2015142631.

[11] Fahim-Ul-Hassan and Gary J Cook. Pet/ct in clinical oncology. *Clinical Medicine*, 12(4):368–372, 2012. https://doi.org/10.7861/clinmedicine.12-4-368.

[12] Aladin Carovac, Fahrudin Smajlovic, and Dzelaludin Junuzovic. Application of ultrasound in medicine. *Acta Informatica Medica*, 19(3):168–171, 2011. https://pubmed.ncbi.nlm.nih.gov/23408755/.

[13] Haeni Lee, Seongyi Han, Hyunjun Kye, Tae-Kyoung Kim, Wonseok Choi, and Jeesu Kim. A review on the roles of photoacoustic imaging for conventional and novel clinical diagnostic applications. *Photonics*, 10(8):904, 2023. https://doi.org/10.3390/photonics10080904.

[14] Apeksh Patwa and Amit Shah. Anatomy and physiology of respiratory system relevant to anaesthesia. *Indian Journal of Anaesthesia*, 59(9):533–541, 2015. https://doi.org/10.4103/0019-5049.165849.

[15] Zewei Zhang, Jialiang Ren, Xiuli Tao, Wei Tang, Shijun Zhao, Lina Zhou, Yao Huang, Jianwei Wang, and Ning Wu. Automatic segmentation of pulmonary lobes on low-dose

computed tomography using deep learning. *Annals of Translational Medicine*, 9(4):291, 2021. https://doi.org/10.21037/atm-20-5060.

[16] Charalampos Charalampidis, Andrianna Youroukou, George Lazaridis, Sofia Baka, Ioannis Mpoukovinas, Vasilis Karavasilis, Ioannis Kioumis, Georgia Pitsiou, Antonis Papaiwannou, Anastasia Karavergou, Kosmas Tsakiridis, Nikolaos Katsikogiannis, Eirini Sarika, Konstantinos Kapanidis, Leonidas Sakkas, Ipokratis Korantzis, Sofia Lampaki, Konstantinos Zarogoulidis, and Paul Zarogoulidis. Pleura space anatomy. *Journal of Thoracic Disease*, 7(S1):S27–S32, 2015. https://doi.org/10.3978/j.issn.2072-1439.2015.01.48.

[17] Matthew J. Fogarty, Carlos B. Mantilla, and Gary C. Sieck. Breathing: Motor control of diaphragm muscle. *Physiology*, 33(2):113–126, 2018. https://doi.org/10.1152/physiol.00002.2018.

[18] Lars Knudsen and Matthias Ochs. The micromechanics of lung alveoli: Structure and function of surfactant and tissue components. *Histochemistry and Cell Biology*, 150(6):661–676, 2018. https://doi.org/10.1007/s00418-018-1747-9.

[19] Joachim D. Pleil, M. Ariel Geer Wallace, Michael D. Davis, and Christopher M. Matty. The physics of human breathing: flow, timing, volume, and pressure parameters for normal, on-demand, and ventilator respiration. *Journal of Breath Research*, 15(4), 2022. https://doi.org/10.1088/1752-7163/ac2589.

[20] Nicola Ullmann, Virginia Mirra, Antonio Di Marco, Martino Pavone, Federica Porcaro, Valentina Negro, Alessandro Onofri, and Renato Cutrera. Asthma: Differential diagnosis and comorbidities. *Frontiers in Pediatrics*, 6:276, 2018. https://doi.org/10.3389/fped.2018.00276.

[21] Stephen I. Rennard, James K. Stoller, and Helen Hollingsworth. Chronic obstructive pulmonary disease: Definition, clinical manifestations, diagnosis, and staging. *UpToDate*, 2015. https://doi.org/10.3978/j.issn.2072-1439.2015.01.48.

[22] Grant Mackenzie. The definition and classification of pneumonia. *Pneumonia*, 8:14, 2016. https://doi.org/10.1186/s41479-016-0012-z.

[23] Kyung-Yil Lee. Pneumonia, acute respiratory distress syndrome, and early immune-modulator therapy. *International Journal of Molecular Sciences*, 18(2):388, 2017. https://doi.org/10.3390/ijms18020388.

[24] Yon Ju Ryu. Diagnosis of pulmonary tuberculosis: Recent advances and diagnostic algorithms. *Tuberculosis and Respiratory Diseases*, 78(2):64–71, 2015. https://doi.org/10.4046/trd.2015.78.2.64.

[25] Caichen Li, Huiting Wang, Yu Jiang, Wenhai Fu, Xiwen Liu, Ran Zhong, Bo Cheng, Feng Zhu, Yang Xiang, Jianxing He, and Wenhua Liang. Advances in lung cancer screening and early detection. *Cancer Biology & Medicine*, 19(5):591–604, 2022. https://doi.org/10.20892/j.issn.2095-3941.2021.0690.

[26] Jürgen Behr, Maurits Demedts, Roland Buhl, Ulrich Costabel, Richard P. N. Dekhuijzen, Henk M. Jansen, William MacNee, Michiel Thomeer, Benoit Wallaert, Francois Laurent, Andrew G. Nicholson, Eric K. Verbeken, Johny Verschakelen, CDR Flower, Stefano Petruzzelli, Paul De Vuyst, JMM van den Bosch, Eulogio Rodriguez-Becerra, Ida Lankhorst, Marco Sardina, Gabrielle Boissard, and the Ifigenia study group. Lung function in idiopathic pulmonary fibrosis - extended analyses of the ifigenia trial. *Respiratory Research*, 10:101, 2009. https://doi.org/10.1186/1465-9921-10-101.

[27] Olivia R. Ferry, Yao C. Huang, Philip J. Masel, Michael Hamilton, Kwun M. Fong, Rayleen V. Bowman, Scott C. McKenzie, and Ian A. Yang. Diagnostic approach to chronic dyspnoea in adults. *Journal of Thoracic Disease*, 11(S17):S2117–S2128, 2019. https://doi.org/10.21037/jtd.2019.10.53.

[28] Jose Curbelo, Sergio Luquero Bueno, Jose Maria Galvan-Roman, Mara Ortega-Gomez, Olga Rajas, Guillermo Fernandez-Jimenez, Lorena Vega-Piris, Francisco Rodrıguez-Salvanes, Belen Arnalich, Ana Dıaz, Ramon Costa, Hortensia de la Fuente, Angel Lancho, Carmen Suarez, Julio Ancochea, and Javier Aspa. Inflammation biomarkers in blood as mortality predictors in community-acquired pneumonia admitted patients: Importance of comparison with neutrophil count percentage or neutrophil-lymphocyte ratio. *PLOS ONE*, 12(3):e0173947, 2017. https://doi.org/10.1371/journal.pone.0173947.

[29] Chloé Fischer, José Knüsli, Loïc Lhopitallier, Estelle Tenisch, Marie-Garance Meuwly, Pauline Douek, Jean-Yves Meuwly, Valérie D'Acremont, Andreas Kronenberg, Isabella Locatelli, Yolanda Mueller, Nicolas Senn, and Noémie Boillat-Blanco. Pulse oximetry as an aid to rule out pneumonia among patients with a lower respiratory tract infection in primary care. *Antibiotics*, 12(3):496, 2023. https://doi.org/10.3390/antibiotics12030496.

[30] Liying Sun, Chi Zhang, Shuhua An, Xiangpeng Chen, Yamei Li, Leshan Xiu, Baoping Xu, Zhengde Xie, and Junping Peng. Comprehensive description of pathogens and antibiotic treatment guidance in children with community-acquired pneumonia using combined mass spectrometry methods. *Frontiers in Cellular and Infection Microbiology*, 11:695134, 2021. https://doi.org/10.3389/fcimb.2021.695134.

[31] Yu Sato, Kengo Murata, Miake Yamamoto, Tsukasa Ishiwata, Miyako Kitazono-Saitoh, Akihiko Wada, and Mikio Takamori. Risk factors for post-bronchoscopy pneumonia: a case–control study. *Scientific Reports*, 10(19983), 2020. https://doi.org/10.1038/s41598-020-76998-z.

[32] Dima Ibrahim, Abdul Rahman Bizri, Mohammad Ali El Amine, and Zeina Halabi. Chest computed tomography and chest x-ray in the diagnosis of community-acquired pneumonia: a retrospective observational study. *Journal of International Medical Research*, 49(8), 2021. https://doi.org/10.1177/03000605211039791.

[33] Rade R. Babic, Gordana Stankovic Babic, Strahinja R. Babic, and Nevena R. Babic. 120 years since the discovery of x-rays. *Med Pregl*, 69(9-10):323–330, 2016. https://doi.org/10.2298/MPNS1610323B.

[34] Yuzo Yamasaki, Takeshi Kamitani, Koji Sagiyama, Takuya Hino, Megumi Kisanuki, Kosuke Tabata, Takuro Isoda, Yoshiyuki Kitamura, Kohtaro Abe, Kazuya Hosokawa, Daisuke Toyomura, Shohei Moriyama, Masateru Kawakubo, Hidetake Yabuuchi, and Kousei Ishigami. Dynamic chest radiography for pulmonary vascular diseases: clinical applications and correlation with other imaging modalities. *Japanese Journal of Radiology*, 42:126–144, 2024. https://doi.org/10.1007/s11604-023-01483-2.

[35] Catherine Beigelman-Aubry and Sabine Schmidt. Pulmonary infections: Imaging with ct. *Multidetector-Row CT of the Thorax, Medical Radiology*, 2016. https://doi.org/10.1007/978-3-319-30355-0$_8$.

[36] Mateusz C. Florkow, Koen Willemsen, Vasco V. Mascarenhas, Edwin H.G. Oei, Marijn van Stralen, and Peter R. Seevinck. Magnetic resonance imaging versus computed tomography for three-dimensional bone imaging of musculoskeletal pathologies: A review. *Journal of Magnetic Resonance Imaging*, 56:11–34, 2022. https://doi.org/10.1002/jmri.28067.

[37] Daniel Lee Aizhi Zhu and Hyunsuk Shim. Metabolic pet imaging in cancer detection and therapy response. *Seminars in Oncology*, 38(1):55–69, 2011. https://doi.org/10.1053/j.seminoncol.2010.11.012.

[38] Andreas S. Panayides, Amir Amini, Nenad D. Filipovic, Ashish Sharma, Sotirios A. Tsaftaris, Alistair Young, David Foran, Nhan Do, Spyretta Golemati, Tahsin Kurc, Kun Huang, Konstantina S. Nikita, Ben P. Veasey, Michalis Zervakis, Joel H. Saltz, and Constantinos S. Pattichis. Ai in medical imaging informatics: Current challenges and future directions. *IEEE Journal of Biomedical and Health Informatics*, 24(7):1837–1857, 2020. https://doi.org/10.1109/JBHI.2020.2991043.

[39] Luís Pinto-Coelho. How artificial intelligence is shaping medical imaging technology: A survey of innovations and applications. *Bioengineering*, 10(12):1435, 2023. https://doi.org/10.3390/bioengineering10121435.

[40] Filippo Pesapane, Marina Codari, and Francesco Sardanelli. Artificial intelligence in medical imaging: threat or opportunity? radiologists again at the forefront of innovation in medicine. *European Radiology Experimental*, 2:35, 2018. https://doi.org/10.1186/s41747-018-0061-6.

[41] Mingyu Kim, Jihye Yun, Yongwon Cho, Keewon Shin, Ryoungwoo Jang, Hyun jin Bae, and Namkug Kim. Deep learning in medical imaging. *Neurospine*, 16(4):657–668, 2019. https://doi.org/10.14245/ns.1938396.198.

[42] Wenli Yang, Yuchen Wei, Hanyu Wei, Yanyu Chen, Guan Huang, Xiang Li, Renjie Li, Naimeng Yao, Xinyi Wang, Xiaotong Gu, Muhammad Bilal Amin, and Byeong Kang.

Survey on explainable ai: From approaches, limitations and applications aspects. *Human-Centric Intelligent Systems*, 3, 2023. https://doi.org/10.1007/s44230-023-00038-y.

[43] European Society of Radiology (ESR). Impact of artificial intelligence on radiology: a euroaim survey among members of the european society of radiology. *Insights into Imaging*, 10(1):105, 2019. https://doi.org/10.1186/s13244-019-0798-3.

[44] James A. Nichols, Hsien W. Herbert Chan, and Matthew A. B. Baker. Machine learning: applications of artificial intelligence to imaging and diagnosis. *Biophysical Reviews*, 11:111–118, 2019. https://doi.org/10.1007/s12551-018-0449-9.

[45] Meghavi Rana and Megha Bhushan. Machine learning and deep learning approach for medical image analysis: diagnosis to detection. *Multimedia Tools and Applications*, 82:26731–26769, 2023. https://doi.org/10.1007/s11042-022-14305-w.

[46] Lei Cai, Jingyang Gao, and Di Zhao. A review of the application of deep learning in medical image classification and segmentation. *Annals of Translational Medicine*, 8(11):713, 2020. https://doi.org/10.21037/atm.2020.02.44.

[47] Samir S. Yadav and Shivajirao M. Jadhav. Deep convolutional neural network based medical image classification for disease diagnosis. *Journal of Big Data*, 6:113, 2019. https://doi.org/10.1186/s40537-019-0276-2.

[48] Fei Yu, Hongbo Zhang, Prayag Tiwari, and Benyou Wang. Natural language reasoning, a survey. *arXiv*, 1(1):1–36, 2023. https://doi.org/10.48550/arXiv.2303.14725.

[49] Taufik Rahmat, Azlan Ismail, and Sharifah Aliman. Chest x-ray image classification using faster r-cnn. *Malaysian Journal of Computing*, 4(1):225–236, 2019. https://doi.org/10.24191/mjoc.v4i1.6095.

[50] Govardhan Jain, Deepti Mittal, Daksh Thakur, and Madhup K. Mittal. A deep learning approach to detect covid-19 coronavirus with x-ray images. *Biocybernetics and Biomedical Engineering*, 40:1391–1405, 2020. https://doi.org/10.1016/j.bbe.2020.08.008.

[51] Prabira Kumar Sethy, Santi Kumari Behera, Pradyumna Kumar Ratha, and Preesat Biswas. Detection of coronavirus disease (covid-19) based on deep features and support

vector machine. *International Journal of Mathematical, Engineering and Management Sciences*, 5(4):643–651, 2020. https://doi.org/10.33889/IJMEMS.2020.5.4.052.

[52] Yaron Anavi, Ilya Kogan, Elad Gelbart, Ofer Geva, and Hayit Greenspan. A comparative study for chest radiograph image retrieval using binary texture and deep learning classification. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 136–140, 2015. https://doi.org/10.1109/EMBC.2015.7319008.

[53] S. Suganyadevi, V. Seethalakshmi, and K. Balasamy. A review on deep learning in medical image analysis. *International Journal of Multimedia Information Retrieval*, 11:19–38, 2022. https://doi.org/10.1007/s13735-021-00218-1.

[54] Soumya Ray, Stephen Scott, and Hendrik Blockeel. Multi-instance learning. In Claude Sammut and Geoffrey I. Webb, editors, *Encyclopedia of Machine Learning*, pages 701–710. Springer, Boston, MA, 2011. https://doi.org/10.1007/978-0-387-30164-8$_5$69.

[55] Emanuel Kitzelmann. Inductive programming: A survey of program synthesis techniques. In Rinus Plasmeijer Ute Schmid, Emanuel Kitzelmann, editor, *Approaches and Applications of Inductive Programming*, volume 5812 of *Lecture Notes in Computer Science*, pages 50–73. Springer, Berlin, Heidelberg, 2010. https://doi.org/10.4230/DagRep.9.5.58.

[56] Sumit Gulwani. Inductive programming: A survey of program synthesis techniques. In Zhenjiang Hu, editor, *Programming Languages and Systems*, volume 6014 of *Lecture Notes in Computer Science*, pages 1–7. Springer, Berlin, Heidelberg, 2010. https://doi.org/10.1007/978-3-642-11931-6$_3$.

[57] Andrew Janowczyk and Anant Madabhushi. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *Journal of Pathology Informatics*, 7:29, 2016. https://doi.org/10.4103/2153-3539.186902.

[58] Davide Cacciarelli and Murat Kulahci. Active learning for data streams: a survey. *Machine Learning*, 113:185–239, 2024. https://doi.org/10.1007/s10994-023-06454-2.

[59] Steven C. H. Hoi, Doyen Sahoo, Jing Lu, and Peilin Zhao. Online learning: A comprehensive survey. *arXiv*, 1:1–100, 2018. https://doi.org/10.1016/j.neucom.2021.04.112.

[60] Karl Weiss, Taghi M. Khoshgoftaar, and DingDing Wang. A survey of transfer learning. *Journal of Big Data*, 3:9, 2016. https://doi.org/10.1186/s40537-016-0043-6.

[61] S. Suganyadevi, D. Shamia, K. Balasamy, SK Hafizul Islam, and Debabrata Samanta. An iot-based diet monitoring healthcare system for women. In Kuan-Ching Li, Ching-Hsien Hsu, Dong Zheng, Nian Xie, Tomoya Enokido, Hugo Filipe Martins, and Pijush Samui, editors, *Smart Healthcare System Design: Security and Privacy Aspects*, pages 150–171. Wiley, 2021. https://doi.org/10.1002/9781119792253.ch8.

[62] Zhengxiang Shi, Francesco Tonolini, Nikolaos Aletras, Emine Yilmaz, Gabriella Kazai, and Yunlong Jiao. Rethinking semi-supervised learning with language models. *arXiv*, 1:1–22, 2023. https://doi.org/10.18653/v1/2023.findings-acl.347.

[63] Hao Wu, Jun Sun, and Qi You. Semi-supervised learning for medical image classification based on anti-curriculum learning. *Mathematics*, 11:1306, 2023. https://doi.org/10.3390/math11061306.

[64] Olivier Chapelle, Bernhard Schölkopf, and Alexander Zien, editors. *Semi-Supervised Learning*. MIT Press, 2006. https://doi.org/10.1109/TNN.2009.2015974.

[65] Hritam Basak, Rajarshi Bhattacharya, Rukhshanda Hussain, and Agniv Chatterjee. An embarrassingly simple consistency regularization method for semi-supervised medical image segmentation. *arXiv*, 1:1–10, 2022. https://doi.org/10.48550/arXiv.2202.00677.

[66] Quande Liu, Lequan Yu, Luyang Luo, Qi Dou, and Pheng Ann Heng. Semi-supervised medical image classification with relation-driven self-ensembling model. *IEEE Transactions on Medical Imaging*, pages 1–13, 2020. https://doi.org/10.1109/TMI.2020.2995518.

[67] Thi Kieu Khanh Ho, Ali Karami, and Narges Armanfard. Graph anomaly detection in time series: A survey. *arXiv*, 1:1–22, 2023. https://doi.org/10.48550/arXiv.2302.00058.

[68] Marianne de Vriendt, Philip Sellars, and Angelica I. Aviles-Rivero. The graphnet zoo: An all-in-one graph based deep semi-supervised framework for medical image classification. *arXiv*, 1:1–9, 2020. https://doi.org/10.1007/978-3-030-60365-6$_1$8.

[69] Huimin Huang, Lanfen Lin, Yue Zhang, Yingying Xu, Jing Zheng, XiongWei Mao, Xiaohan Qian, Zhiyi Peng, Jianying Zhou, and Yen-Wei Chen. Graph-bas3net: Boundary-aware semi-supervised segmentation network with bilateral graph convolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 136–140, 2021. https://doi.org/10.1109/ICCV48922.2021.00729.

[70] 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. https://doi.org/10.1016/j.artmed.2020.101938.

[71] Zeyuan Allen-Zhu and Yuanzhi Li. Feature purification: How adversarial training performs robust deep learning. *arXiv*, 1:1–34, 2020. https://doi.org/10.1109/FOCS52979.2021.00098.

[72] Kun Liu, Xiaolin Ning, and Sidong Liu. Medical image classification based on semi-supervised generative adversarial network and pseudo-labelling. *Sensors*, 22(24):9967, 2022. https://doi.org/10.3390/s22249967.

[73] Chang Qi, Junyang Chen, Guizhi Xu, Zhenghua Xu, Thomas Lukasiewicz, and Yang Liu. Sag-gan: Semi-supervised attention-guided gans for data augmentation on medical images. *arXiv*, 1:1–12, 2020. https://doi.org/10.48550/arXiv.2011.07534.

[74] Zahra Solatidehkordi and Imran Zualkernan. Survey on recent trends in medical image classification using semi-supervised learning. *Applied Sciences*, 12(23):12094, 2022. https://doi.org/10.3390/app122312094.

[75] Ayush Pant, Akshat Jain, Kiran C Nayak, Daksh Gandhi, and Dr. B. G. Prasad. Pneumonia detection: An efficient approach using deep learning. In *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*, pages 1–7, 2020. https://doi.org/10.1109/ICCCNT49239.2020.9225543.

[76] Garima Verma and Shiva Prakash. Pneumonia classification using deep learning in healthcare. *International Journal of Innovative Technology and Exploring Engineering*, 9(4):643–651, 2020. https://doi.org/10.35940/ijitee.D1599.029420.

[77] Mana Saleh Al Reshan, Kanwarpartap Singh Gill, Vatsala Anand, Sheifali Gupta, Hani Alshahrani, Adel Sulaiman, and Asadullah Shaikh. Detection of pneumonia from chest x-ray images utilizing mobilenet model. *Healthcare*, 11(11):1561, 2023. https://doi.org/10.3390/healthcare11111561.

[78] Tatiana Malygina, Elena Ericheva, and Ivan Drokin. Gans 'n lungs: improving pneumonia prediction. *Medical Imaging with Deep Learning 2019 MIDL 2019 – Extended Abstract Track*, 2019. https://doi.org/10.13140/RG.2.2.21979.00800.

[79] Anuja Kumar Acharya and Rajalakshmi Satapathy. A deep learning based approach towards the automatic diagnosis of pneumonia from chest radio-graphs. *Biomedical and Pharmacology Journal*, 13(1):449–455, 2020. https://doi.org/10.13005/bpj/1905.

[80] Alhassan Mabrouk, Rebeca P. Díaz Redondo, Abdelghani Dahou, Mohamed Abd Elaziz, and Mohammed Kayed. Pneumonia detection on chest x-ray images using ensemble of deep convolutional neural networks. *Applied Sciences*, 12(13):6448, 2022. https://doi.org/10.3390/app12136448.

[81] Prafulla Dhariwal and Alex Nichol. Diffusion models beat gans on image synthesis. *arXiv*, 1:1–16, 2021. https://doi.org/10.48550/arXiv.2105.05233.

[82] Gustav Müller-Franzes, Jan Moritz Niehues, Firas Khader, Soroosh Tayebi Arasteh, Christoph Haarburger, Christiane Kuhl, Tianci Wang, Tianyu Han, Sven Nebelung, Jakob Nikolas Kather, and Daniel Truhn. Diffusion probabilistic models beat gans on medical images. *arXiv*, 1:1–13, 2022. https://doi.org/10.48550/arXiv.2212.07501.

[83] Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *arXiv*, 1:1–12, 2020. https://doi.org/10.48550/arXiv.2006.11239.

[84] Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, and Saining Xie. A convnet for the 2020s. *arXiv*, 1:1–17, 2022. https://doi.org/10.48550/arXiv.2201.03545.

[85] Farouq Benchallal, Adel Hafiane, Nicolas Ragot, and Raphaël Canals. Convnext based semi-supervised approach with consistency regularization for weeds classification. *Expert Systems with Applications*, 239:122222, 2024. https://doi.org/10.1016/j.eswa.2023.122222.

[86] Saikat Roy, Gregor Koehler, Constantin Ulrich, Michael Baumgartner, Jens Petersen, Fabian Isensee, Paul F. Jaeger, and Klaus Maier-Hein. Mednext: Transformer-driven scaling of convnets for medical image segmentation. *arXiv*, 1:1–22, 2023. https://doi.org/10.48550/arXiv.2303.09975.