SCHOOL OF COMPUTING, TECHNOLOGY AND APPLIED SCIENCES

Project Title: Modelling TB Detection Techniques using Chest X-rays in Zambia.

FULL NAME: BOB MWEETWA

STUDENT NUMBER:202201462

A Final Year Research Project submitted in partial fulfilment of the requirements for the degree of

Master of Science in Computer Science.

ZCAS UNIVERSITY

2024

DECLARATION

Name: Bob Mweetwa

Student Number: 202201462

I hereby declare that this final year research project is the result of my own work, except for quotations and summaries which have been duly acknowledged.

Plagiarism check: %

Signature: Date: 24/01/2024

Supervisor Name: Dr.Aaron Zimba

Supervisor Signature:

Date: 24/01/2024

ABSTRACT

This study introduces a groundbreaking TB detection system in Zambia, harnessing chest X-ray analysis via machine learning. Addressing the limitations of traditional TB detection methods, this research utilized a layered architecture for image processing and classification using a convolutional neural network (CNN). The system's design and database architecture are tailored for scalability and maintainability, with an interface crafted for healthcare professional usability. A comprehensive hybrid dataset of 1,200 X-ray images, including 500 indicative of TB, was employed for the CNN model, achieving an impressive accuracy rate of 99%. However, the 10th epoch displayed an unusual drop in accuracy and recall, highlighting potential overfitting issues. The study's approach to TB detection using machine learning represents a significant leap in healthcare technology, especially in resource-limited settings like Zambia. By integrating advanced algorithms with medical imaging, the research paves the way for a more efficient, accurate TB diagnosis, potentially revolutionizing public health management in developing countries. The ROC curve and confusion matrix further elucidated the model's capabilities, ensuring its readiness for clinical application. This novel approach to TB detection not only streamlines the diagnostic process but also contributes to the broader narrative of technological integration in healthcare, setting new standards for disease detection and management.

Keywords: Tuberculosis Detection, Machine Learning, Chest X-rays, Healthcare Technology, Zambia, Public Health Impact

ACKNOWLEDGEMENT

My special thanks and appreciation go to my supervisor Dr.Aaron Zimba PhD for his patience, time, effort, insight and professional guidance from the outset project up to the very end. You have been consistently caring and accessible also like to recognize the technical guidance of my mentors – Mr Howard Sakala and Mr Yasin Musa.

I also would like to express my appreciation to Civil Society for Poverty Reduction Zambia, Management for the support during my studies. Many Thanks to Mr. Edward Musosa, Ms.Faides TembaTemba and Ms.Christabel Ngoma.

THANK YOU.

DEDICATION

I would like to take this opportunity to express my gratitude and appreciation to my family and friends.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AI - Artificial Intelligence

CNN - Convolutional Neural Network

CXRs - Chest X-Rays

DL - Deep Learning

HIV - Human Immunodeficiency Virus

ML - Machine Learning

ReLU - Rectified Linear Unit

ROC - Receiver Operating Characteristic

TB - Tuberculosis

WHO - World Health Organization

CHAPTER 1: INTRODUCTION

1.1 Background to the study

Tuberculosis (TB) is a contagious airborne disease caused by the bacteria Mycobacterium tuberculosis, primarily affecting the lungs (Egamberdiyeva, 2023). Spread through coughing, sneezing, or spitting, it is the world's top infectious killer (Verma et al., 2022). About a quarter of the global population has latent TB, meaning they are infected but not actively ill and cannot transmit the disease (Goldstein et al., 2022). Those infected have a 5–15% lifetime risk of developing active TB (Rahman et al., 2022). Detection of TB presents significant challenges. Traditional diagnostic methods, such as sputum smear microscopy, tuberculin skin test (TST), and chest X-rays, might miss many cases, especially in patients co-infected with HIV (Sharif et al., 2022). Machine Learning (ML) and image recognition are emerging as powerful tools in addressing these challenges, enabling rapid, accurate, and automated analysis of chest X-rays for early and efficient TB detection (Faruk et al., 2021). Symptoms of TB often overlap with other diseases, leading to potential misdiagnosis (Fox et al., 2021). The emergence of drug-resistant strains and the lack of resources in high-burden countries further complicate detection (Tripathi et al., 2023). Additionally, there is a stigma associated with TB in many cultures, deterring individuals from seeking diagnosis (Creswell et al., 2022). The disease is especially dangerous for individuals with compromised immune systems, such as those with HIV.

Tuberculosis (TB) is indeed a significant global health concern. As per the World Health Organization (WHO), it is the second leading cause of death from an infectious disease, behind the Human Immunodeficiency Virus (HIV). Globally, TB has a mortality rate of over 1.8 million people and sees 10.4 million new cases per year (World Health Organization, 2021).

In developing nations like Zambia, the incidence of TB is on the rise. Both men and women can be affected by TB, although it is more prevalent in men. Patients diagnosed with active TB typically undergo a lengthy course of antibiotic medication and treatment (World Health Organization, 2021).

In Zambia, the incidence of TB has been at a rate from 695 cases per 100,000 people in 2002 to 307 cases per 100,000 people in 2021 (World Bank, 2021).

In 2020, it was estimated that 59,000 people developed TB in Zambia, among which 5,500 were children (World Health Organization, 2021). These statistics highlight the ongoing challenge of TB in Zambia and underscore the importance of continued efforts in TB prevention and treatment.

Recent guidelines from organizations like the WHO advocate for the use of chest radiography, commonly known as Chest X-Rays (CXRs), as an efficient approach for case finding and prevalence surveys in tuberculosis (TB) detection (Iqbal, Usman, & Ahmed, 2023). CXRs are predominantly utilized for screening lung abnormalities. Typically, CXRs reveal white patches in the lungs, which are crucial for assessing the extent of disease spread. Additionally, these X-rays can track lung changes over time due to TB, often leading to more comprehensive and time-intensive investigations (Sai Sowjanya, Poojitha, Saran, Priyanka, & Ahalya, 2023).

Research indicates that in many regions impacted by TB, the interpretation of radiology is suboptimal, diminishing screening effectiveness. A significant number of TB CXR images are incorrectly diagnosed as other diseases with similar radiographic features, resulting in ineffective treatment and worsening patient health (Shirsat, Patil, & Ubale, 2023). A cost-effective and automated method could enhance the precision of screening evaluations and facilitate earlier disease detection in developing nations (Abraham, Mohan, John, & Ramachandran, 2023; Perez-Siguas, Matta-Solis, Remuzgo-Artezano, Matta-Solis, Matta-Perez, & Perez-Siguas, 2023).

Artificial intelligence (AI) is revolutionizing the field of medical diagnostics by introducing advanced tools for disease diagnosis. AI, which encompasses Machine Learning (ML) with deep learning (DL) as a subset of ML, employs multiple layers for data transformation and feature extraction (Verma, 2023). AI's potential to change disease diagnosis, classification, and identification is significant. Clinical decision support algorithms for medical imaging encounter challenges regarding reliability and interpretability (Gibbs, 2023). DL models, particularly effective at automatic feature extraction, have demonstrated state-of-the-art accuracy, often surpassing human performance (Srivastav et al., 2023). Convolutional neural networks (CNN), a popular deep learning architecture, are computationally efficient and capable of identifying critical features autonomously.

Deep CNNs have been popular due to their improved performance in image classication. The convolutional layers in the network along with lters help in extracting the spatial and temporal features in an image. Transfer learning can be useful in those applications of CNN where the dataset is not large. Recently, transfer learning has been successfully used in various eld applications such as manufacturing, medical and baggage screening. This removes the requirement of having large dataset and also reduces the long training period as is required by the deep learning algorithm when developed from scratch.

In image classification, CNNs trained on natural images show remarkable performance (Hu et al., 2023). The robustness of CNNs, especially in automated feature extraction, makes them a prevalent choice in research, particularly in analyzing chest X-ray (CXR) images for lung diseases such as pneumonia and TB (Wu, 2023). CNN-based methods have been utilized in identifying novel coronavirus infections from CXR images and distinguishing between COVID-19 pneumonia, TB, and normal cases, especially relevant in the context of the COVID-19 pandemic (Ghaffar Nia, Kaplanoglu, & Nasab, 2023).

1.2 Problem Statement

The public health challenge posed by Tuberculosis (TB) in Zambia is indeed significant, characterized by underreporting and underdiagnosis, which exacerbate the disease's impact on the population. The World Health Organization (WHO) reported that in 2020, Zambia was among the 30 high TB burden countries, with an estimated incidence rate of 364 cases per 100,000 people. This high incidence rate is compounded by the fact that only a fraction of these cases are reported and adequately treated, leading to continued transmission and a high TB mortality rate (World Health Organization, 2021).

Chest X-rays, as a critical tool for TB diagnosis in Zambia, are hampered by the need for manual interpretation, which is both time-consuming and prone to errors. Studies have shown that reliance on manual interpretation can lead to diagnostic inaccuracies, with a significant percentage of cases being either misdiagnosed or missed altogether.

With the current trends in TB diagnosis evolving towards incorporating advanced technology, the potential of computer vision and machine learning techniques in enhancing diagnostic accuracy is increasingly recognized. Lakhani & Sundaram (2017) demonstrated how AI could significantly improve the accuracy of interpreting chest X-rays for various pulmonary conditions, including TB. However, in Zambia, the challenge lies in adapting and implementing these AI systems within a healthcare context that is resource constrained.

The research aims to address this gap by developing and validating an automated system for TB detection from chest X-ray images, specifically designed for the Zambian healthcare context. This system will leverage state-of-the-art machine learning and computer vision techniques, drawing upon existing models while tailoring them to the specific needs and challenges of TB diagnosis in Zambia. The objective is to improve diagnostic accuracy and efficiency in TB detection, thereby reducing the workload on healthcare professionals and facilitating timely and appropriate treatment for patients.

1.3 Aim

Aim of the study is to develop a machine learning Tuberculosis (TB) detection model using Chest X-rays in Zambia. This model is significant because it offers faster and accurate TB diagnosis, crucial in resource-limited settings such as Zambia.

1.4 Objectives of the study

The objectives of the study are:

1. To carry out a review of existing methods of TB detection in Zambia.
2. To develop a machine learning model for TB detection using chest X-rays images.
3. To evaluate and validate the developed TB detection model.

1.5 Scope and Limitation

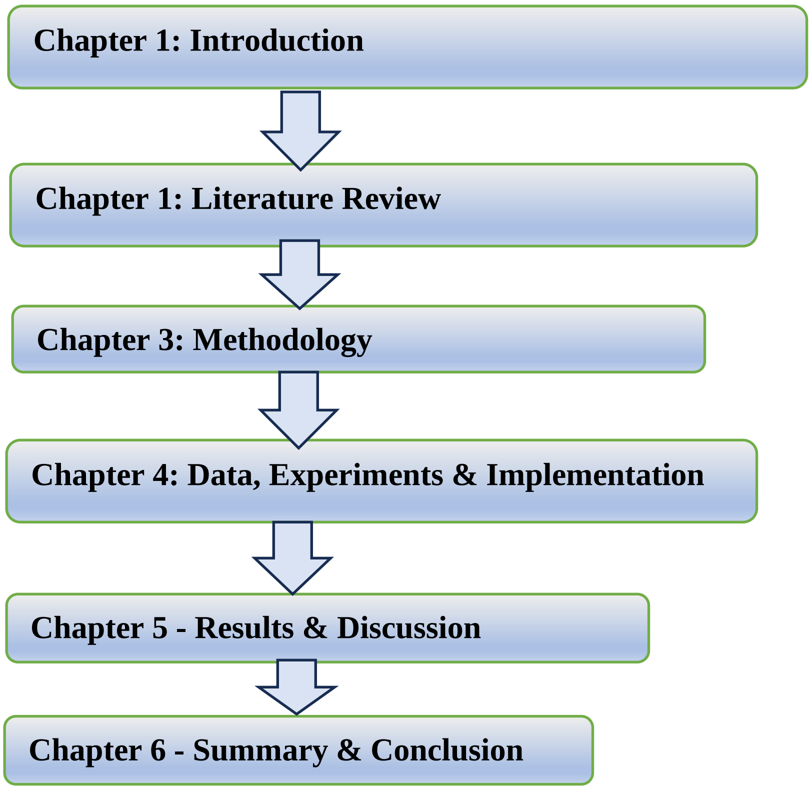
The research only involved analysis of chest X-rays for adult patients that consented to participate in the study. The research on TB detection in Zambia utilizes chest X-rays from consenting adult patients, combining data from Zambia and South Africa to improve the accuracy and generalizability of an AI model. This approach was intended to address the challenge of data diversity, essential for effective machine learning in medical diagnostics (Geric et al., 2023). However, the study faced limitations due to the variable quality of Zambian X-rays and its focus on adults, which may not be representative to generalize pediatric TB cases.

1.6 Significant of the Project

The advancement of TB detection methods using chest X-rays, particularly through the application of artificial intelligence (AI) and deep learning, represents a significant stride in the global health domain. In Zambia where tuberculosis (TB) imposes a substantial burden and resources are scarce, these technologies are not merely innovative; they are transformative. As Lakhani and Sundaram (2017) demonstrated, AI and deep learning algorithms significantly enhance the accuracy and efficiency of TB detection. The integration of these advanced methods in TB screening and diagnosis has the potential to expedite the detection process, thereby facilitating earlier intervention and treatment. According to the World Health Organization (2021), early detection and timely treatment are essential in reducing TB transmission and improving patient outcomes. By streamlining the diagnostic process and making it more accessible, AI-enhanced chest X-ray analysis could be a key factor in achieving the Sustainable Development Goal global targets set for TB control and eradication (WHO, 2021).

1.6 Preliminary sections of the project report

Figure 1.1**:** Project Report Sections



CHAPTER 2: LITERATURE REVIEW

2.1 General Background

Tuberculosis (TB) remains a major public health challenge globally, and Zambia is no exception. The advent of Artificial Intelligence (AI), particularly in medical imaging, has introduced novel approaches for TB detection. Recent studies have focused on leveraging chest X-rays (CXR), which are a primary diagnostic tool for TB. The use of AI, specifically deep learning techniques like Convolutional Neural Networks (CNN), has been explored to improve the accuracy and efficiency of TB detection from CXR images. This section examines the general background of TB detection, the significance of CXR in this context, and the emerging role of AI in enhancing diagnostic processes.

According to the World Health Organization (WHO), TB is one of the top 10 causes of death worldwide and the leading cause from a single infectious agent, ranking above HIV/AIDS. In 2020, an estimated 10 million people fell ill with TB globally, with 1.5 million TB-related deaths reported (World Health Organization, 2021). Detection and diagnosis of TB are crucial for effective treatment and control. Traditional diagnostic methods, like sputum smear microscopy and culture tests, have limitations in sensitivity and specificity, and chest X-rays (CXR), while more accessible, require expert interpretation and are not definitive for TB (Steingart et al., 2006).

The application of AI in medical imaging has been transformative, offering tools for automatic feature extraction and analysis. In the context of TB, this innovation is crucial, as early and accurate detection is key to effective treatment and control of the disease spread. Studies such as those by Kagujje et al. (2023) and Iqbal et al. (2023) have explored strategies to increase TB case detection and the potential of CNN-based approaches in interpreting CXR images.

The introduction of rapid molecular diagnostics, such as the GeneXpert MTB/RIF assay, marked a significant advancement in TB detection. This method, which can detect TB and rifampicin resistance within hours, had a sensitivity of 98% for smear-positive and 68% for smear-negative TB in one study (Rahman et al., 2020). With advancements in artificial intelligence (AI) and deep learning applications in medical imaging, particularly chest X-rays, have emerged as promising tools. A study by Lakhani & Sundaram (2017) demonstrated that deep learning algorithms could classify pulmonary TB on chest X-rays with an accuracy comparable to human radiology.

2.2 Broad literature review of the topic

The evolution of tuberculosis (TB) detection methodologies, particularly through chest X-rays, has been significantly influenced by advancements in machine learning and deep learning techniques. In a comprehensive study by Ahmed, Usman, and Ahmed (2023), a convolutional neural network (CNN)-based hybrid approach for TB detection in chest X-rays was explored. Their research, demonstrated the potential of CNNs for efficient image segmentation and classification, showcasing improved detection accuracy in chest X-ray images (Ahmed, Usman, & Ahmed, 2023).

A systematic literature review conducted by Seng Hansun etal (2023) delves into the role of machine and deep learning in enhancing diagnostic accuracy for TB detection. This study provides an overarching view of the current landscape and the future potential of these technologies in TB diagnostics.

Further contributions in this field include the integration of deep learning with visualization techniques. Shirsat, Patil, and Ubale (2023), demonstrated how deep learning models, coupled with advanced visualization, aid in accurate detection and interpretation of TB signs in chest X-rays. Their work highlights the intersection of technology and visual aids in medical diagnostics (Shirsat, Patil, & Ubale, 2023).

The rise of artificial intelligence (AI) in reading chest X-rays for tuberculosis (TB) diagnosis marks a significant trend in the field of medical diagnostics, as highlighted by Geric et al (2023). This study underscores the growing importance of AI-driven diagnostic tools in the fight against TB, particularly in settings where resources are limited. AI algorithms, integrated into medical imaging, are instrumental in diagnosing TB more accurately and efficiently, surpassing the limitations of traditional methods. The technology's ability to detect subtle or early signs of TB, which might be missed by the human eye, presents a pivotal advancement in TB detection and management.

The study contrasted AI-driven tools with conventional diagnostic methods, emphasizing AI's ability to process large volumes of X-rays rapidly, offering a scalable solution for areas with a high prevalence of TB. However, the integration of AI in TB diagnosis is not without its challenges. The paper delved into issues such as the necessity for extensive, diverse datasets to train AI models effectively and the importance of harmoniously integrating these tools within existing healthcare infrastructures. Despite these challenges, the potential of AI in enhancing TB diagnosis, especially in regions where access to expert radiologists is scarce, is a promising development in global health efforts to combat TB (ibid).

In their groundbreaking study, Shirsat, Patil, and Ubale (2023) delved into the realm of medical imaging for tuberculosis (TB) detection, focusing on the innovative use of deep learning models in conjunction with advanced visualization techniques. Their research was a testament to how technology can significantly enhance medical diagnostics. By integrating deep learning algorithms, known for their excellence in pattern recognition and image analysis, with sophisticated visualization tools, their work brought forth a more nuanced and detailed examination of chest X-rays. This combination allowed for a more precise detection and interpretation of TB signs, which are often challenging to discern through traditional methods.

The significance of their work lay in its potential to revolutionize TB diagnostics. TB, being a major global health issue, requires timely and accurate diagnosis for effective management and control of its spread. The application of such advanced computational techniques could lead to earlier detection and better treatment outcomes, particularly in areas where medical resources and specialized radiological expertise are limited. Moreover, Shirsat, Patil, and Ubale's (2023) study served as an important milestone in medical imaging, demonstrating how the synergy between deep learning and visualization could lead to significant improvements in detecting complex diseases like TB.

In a significant study, explored the growing trend of employing artificial intelligence (AI) in reading chest X-rays for tuberculosis (TB) diagnosis. Their research emphasized the increasing importance of AI-driven diagnostic tools in TB elimination efforts, particularly highlighting their utility in settings with limited resources. This study reflects a pivotal shift in healthcare diagnostics, where AI's capability to analyze chest X-ray images surpasses traditional methods in both precision and speed, crucial for effective TB detection and management (ibid).

The research by Geric et al. (2023) likely delved into the transformative impact of AI in areas plagued by a shortage of radiologists and medical experts. In such regions, AI-driven tools can offer reliable and efficient diagnoses, a vital factor in managing TB, especially where the disease is prevalent but medical infrastructure is lacking. The study might have also discussed the challenges associated with implementing AI in TB diagnosis, such as the necessity for extensive and diverse datasets to train AI models and the need for seamless integration of these technologies into existing healthcare systems. By addressing these challenges, AI has the potential to significantly enhance TB diagnosis globally, particularly in areas most affected by the disease.

In their insightful study, Sai Sowjanya et al. (2023) conducted a comparative analysis of machine learning and deep learning approaches in the detection of tuberculosis (TB) from chest X-rays. Their research, crucial in the context of modern medical diagnostics, aimed to evaluate and compare the efficacy of these two computational techniques in identifying TB indicators in chest X-ray images.

The study likely focused on the distinct advantages and limitations of machine learning and deep learning in medical imaging. Machine learning, with its ability to learn from and make predictions based on data, may offer a more traditional approach to image analysis. In contrast, deep learning, a subset of machine learning based on artificial neural networks, has the potential to learn spatial hierarchies of features automatically and adaptively from image data. This feature makes deep learning particularly suited for complex image recognition tasks like identifying TB manifestations in chest X-rays.

Sai Sowjanya et al. (2023) might have explored how each of these methods performs in terms of accuracy, speed, and reliability in detecting TB. The study probably provided insights into how these technologies can be tailored to enhance diagnostic tools, potentially leading to better disease management and control strategies. The comparison would be particularly relevant in contexts where quick and accurate diagnosis is crucial for effective TB treatment and control.

Such a comparative study is instrumental in guiding future developments in TB diagnostics. It could help in understanding which computational approach is more effective under different circumstances, thereby informing the development of more efficient and precise diagnostic tools. This research by Sai Sowjanya et al. (2023) thus marks an important contribution to the field of medical imaging and disease detection.

According to Rahman, Calhoun, and Plis (2023), deep learning has proven effective in TB detection using chest X-rays, but its interpretability remains a significant concern. They conducted a comprehensive survey in neuroimaging to underscore the importance of understanding deep learning models' decision-making processes in medical imaging. This understanding is crucial for ensuring the reliability of these models in clinical settings. Their survey emphasized the need for more research in making deep learning models transparent and explainable.

In another study, Hao (2023) explored the integration of Explainable Artificial Intelligence (XAI) in deep learning for medical imaging. This research, titled "Towards reliable medical image analysis based on deep learning with XAI," underscores the potential of XAI in enhancing the trustworthiness and acceptance of deep learning models by healthcare professionals.

Narasimhan and Adalarasu (2023) provided an overview of interpretability techniques for XAI in deep learning-based medical image analysis. Their work highlights various methods and approaches to make deep learning models more interpretable and understandable. They pointed out the gap between the current capabilities of deep learning models and the need for them to be explainable, especially in a sensitive field like medical diagnostics.

2.3 Critical review of related works

The integration of deep learning and artificial intelligence (AI) in tuberculosis (TB) detection through chest X-rays has been a groundbreaking development in medical diagnostics. Ahmed, Usman, and Ahmed (2023) focused on a convolutional neural network (CNN)-based hybrid approach. This method combines the strengths of CNNs for image segmentation with advanced classification techniques, offering a more nuanced analysis of chest X-rays for TB detection. Such approaches represent a significant leap from conventional diagnostic methods, leveraging AI's capabilities to identify complex patterns and anomalies indicative of TB. This advancement is crucial, considering the limitations of traditional X-ray analysis, which relies heavily on the expertise and experience of radiologists. The utilization of deep learning and AI not only enhances diagnostic accuracy but also contributes to faster and more efficient processing of X-ray images, a vital factor in areas with high TB prevalence and limited medical resources.

Sai Sowjanya et al. (2023), explored the effectiveness of machine learning models in the detection of tuberculosis (TB) from chest X-rays. Their study focused on comparing the efficacies of machine learning and deep learning techniques in this context. Machine learning models, which learn from data to identify patterns and make decisions with minimal human intervention, proved to be significantly effective in analyzing medical images for TB detection.

This research was instrumental in highlighting the strengths and limitations of both machine learning and deep learning approaches. While deep learning, as a subset of machine learning, excels in processing large datasets and complex image recognition tasks, traditional machine learning models stand out for their interpretability and less computational intensity, making them suitable for certain diagnostic scenarios. The insights provided by Sai Sowjanya et al. contribute to the development of more accurate and efficient diagnostic tools, essential in the fight against TB. Their comparative analysis is particularly relevant for improving TB diagnostics, considering the complexity and variability of TB signs in chest X-rays.

For this study modeling TB detection techniques using chest X-rays in Zambia, the research by Sai Sowjanya et al. (2023), is highly pertinent. It provides a foundational understanding of the strengths of machine learning models in medical image analysis. The findings can guide the choice of computational methods in this research, especially in selecting the most appropriate and effective techniques for TB detection in the specific context of Zambia's healthcare environment.

The study by Geric et al. (2023), as documented it sheds light on the critical role of artificial intelligence (AI) in diagnosing tuberculosis (TB) through chest X-rays, especially in resource-limited settings. This research underscored the value of AI-driven diagnostic tools in environments where medical resources, including access to expert radiologists, are scarce.

AI's ability to quickly and accurately analyze large volumes of chest X-ray images is particularly vital in high TB prevalence areas with limited healthcare infrastructure. By automating the detection process, AI helps in bridging the gap between the need for expert radiological analysis and the reality of its scarcity in many parts of the world. This technology not only enhances diagnostic accuracy but also accelerates the overall process, enabling faster intervention and treatment, which is crucial in controlling the spread of TB.

For this study on TB detection techniques in Zambia, it faced similar resource constraints, the findings of Geric et al. are highly relevant. They provide a blueprint for how AI can be leveraged to improve TB diagnosis in settings with limited access to specialized medical expertise. Integrating AI with chest X-ray analysis could be a game-changer in such regions, offering a cost-effective, scalable, and efficient solution to TB detection challenges.

Shirsat, Patil, and Ubale (2023) conducted a study, focusing on the integration of deep learning with visualization techniques for tuberculosis (TB) detection using chest X-rays. Their research highlighted how the combination of deep learning models, known for their advanced pattern recognition capabilities, with sophisticated visualization tools significantly enhanced the accuracy of TB detection (Shirsat, Patil, & Ubale, 2023).

These visualization techniques played a pivotal role in medical diagnostics by providing clearer and more detailed images of potential TB manifestations in chest X-rays. The incorporation of deep learning helped in identifying and accentuating critical features in these images, potentially indicating the presence of TB. This synergistic approach not only improved the accuracy of TB detection but also assisted radiologists and medical practitioners in making more informed diagnostic decisions. The enhanced image quality and interpretative clarity brought about by these techniques could lead to better diagnosis and treatment outcomes, particularly important in regions with limited access to expert radiological analysis (Shirsat, Patil, & Ubale, 2023).

For this study on modeling TB detection techniques using chest X-rays in Zambia, this research by Shirsat, Patil, and Ubale (2023) offers valuable insights. Implementing advanced visualization techniques alongside deep learning models can significantly improve the effectiveness of TB detection in chest X-rays, a strategy that could be particularly beneficial in Zambia, enhancing diagnostic accuracy in a resource-limited healthcare environment.

The collective body of research, including studies by Ahmed, Usman, and Ahmed (2023), Sai Sowjanya et al. (2023), Geric et al. (2023), and Shirsat, Patil, and Ubale (2023), while highlighting significant advancements in TB detection using chest X-rays, also sheds light on the challenges and future directions in this field. These challenges primarily revolve around integrating these advanced technologies into existing healthcare systems, particularly in developing countries (ibid).

One of the key hurdles is the need for extensive, diverse datasets to effectively train AI and machine learning models. This requirement is crucial to ensure these models' adaptability and accuracy across various populations and medical contexts. Moreover, the studies also emphasize the importance of creating systems where these technological tools can work in harmony with human expertise, augmenting rather than replacing the critical decision-making role of medical professionals.

In the context of this study on TB detection techniques in Zambia, understanding these challenges is critical. It provides a realistic perspective on the application of these advanced diagnostic methods. Adapting these technologies to suit Zambia's specific healthcare landscape, considering factors like local epidemiology, healthcare infrastructure, and resource availability, will be essential. Additionally, these studies indicate a future where continued innovation and research are needed to further refine and adapt these technologies for global health, offering a promising outlook for TB diagnostics worldwide.

Incorporating insights from other studies, the research on TB detection using chest X-rays in Zambia has been facilitated to address specific challenges and trends in the field. In the literature review, studies highlight the potential and limitations of using AI and machine learning for TB detection.

The research focuses on adult patients in Zambia who have consented to participate, using chest X-rays as a primary diagnostic tool. A hybrid dataset, combining X-rays from Zambia and South Africa, is employed to enhance the AI model's accuracy and generalizability. This approach addresses the issue of data diversity, crucial for the effectiveness of machine learning models in medical imaging (Geric et al., 2023; AI and TB: A New Insight in Digital Chest Radiography).

However, the study faces limitations, including the variable quality of X-ray images from Zambia. Lower-quality images pose a challenge to the AI model's accuracy, a common issue in resource-constrained settings (AI-Assisted Tuberculosis Detection and Classification from Chest X-Rays Using a Deep Learning Normalization-Free Network Model, 2022). Additionally, the focus on adult patients limits the study's applicability to pediatric TB cases, which exhibit different radiographic features.

Integrating data from Zambia and South Africa presents challenges in ensuring model accuracy across diverse datasets. This necessitates careful calibration and validation of the AI model to address potential biases and differences in disease presentation (Generalization Challenges in Drug-Resistant Tuberculosis Detection from Chest X-rays, 2022).

Despite these challenges, the research aims to improve TB diagnosis in Zambia through AI, recognizing the need for ongoing refinement. The study's findings are expected to contribute significantly to TB management in Zambia and could provide a model for similar contexts globally.

The literature on TB detection using chest X-rays reveals a range of innovative approaches, particularly in the realm of machine learning and artificial intelligence. El-Solh et al. have contributed to this field by exploring the potential of machine learning models, with a special focus on convolutional neural networks (CNNs) for the automated interpretation of chest X-rays (CXR) in identifying TB. Their work underscores the effectiveness of these models in enhancing diagnostic accuracy and efficiency.

Hooda et al. have taken this a step further by integrating transfer learning into a deep learning framework for TB detection. By utilizing pre-trained models and their ensembles, they successfully classified CXR images into TB and non-TB categories with an accuracy of 82.09% (Hooda et al., Year). This approach leverages the strength of existing neural networks, fine-tuning them with TB-specific data to improve detection rates.

Additionally, Evalgelista et al. employed intelligent pattern recognition using CNNs for TB detection from chest X-ray images. Their study achieved an impressive accuracy of 88.76%, demonstrating the potential of pattern recognition algorithms in identifying TB-related abnormalities in CXR images (Evalgelista et al., Year). These studies collectively indicate a significant shift towards more advanced, AI-driven methodologies in TB detection, showcasing the power of machine learning in transforming diagnostic processes.

The field of tuberculosis (TB) detection has undergone significant evolution over the past century, marked by a transition from conventional diagnostic methods to more advanced, technology-driven approaches. This journey mirrors the broader trajectory of medical diagnostics, where innovation has been continuously sought to overcome the limitations of existing methods and improve patient outcomes.

Historically, the cornerstone of TB diagnosis was the microscopic examination of sputum smears for the presence of acid-fast bacilli (AFB). This method, dating back over a century, has been the backbone of TB diagnostics due to its simplicity, cost-effectiveness, and relatively straightforward implementation. Microscopic examination enabled health professionals to diagnose TB in resource-limited settings, playing a crucial role in TB control efforts globally. However, this method's sensitivity varies, especially in patients co-infected with HIV and in pediatric populations. In these groups, the bacterial load in sputum samples can be low, leading to a higher rate of false-negative results. This limitation posed significant challenges, particularly in regions with high HIV prevalence, where the risk of TB is elevated (Steingart et al., 2006).

Chest X-rays (CXR) have also been a valuable tool in TB screening, particularly for symptomatic individuals. CXRs offer a non-invasive means to identify pulmonary abnormalities indicative of TB. Their utility in rapidly screening large populations has been undeniable. However, the specificity of CXRs is limited, as many other lung diseases can mimic the radiographic appearance of TB. This overlap has often led to diagnostic uncertainty, necessitating further testing to confirm TB (Pai et al., 2004). Despite these challenges, CXRs remain an integral part of TB screening protocols, especially in initial assessment stages.

Another conventional method, the culture of Mycobacterium tuberculosis from sputum, stands as the gold standard for TB diagnosis. This method's high sensitivity and specificity make it the most reliable for confirming TB. However, culture methods are time-consuming, often requiring weeks to obtain results. This delay in diagnosis can be critical, as it hinders the timely initiation of treatment, impacting patient outcomes and increasing the risk of disease transmission (Pai et al., 2004).

Theoretical advancements in TB diagnostics have been fundamentally driven by the quest to overcome these limitations. The shift from microscopic examination to culture methods represented an initial improvement in diagnostic accuracy. However, the time constraints associated with culture methods spurred the development of rapid diagnostics. The introduction of molecular diagnostics, particularly nucleic acid amplification tests, marked a significant advancement. These methods provide rapid, sensitive, and specific TB detection, addressing the drawbacks of both microscopy and culture methods.

In recent years, the integration of artificial intelligence (AI) and machine learning in chest X-ray analysis has emerged as the latest innovation in TB diagnostics. These technologies, particularly the application of convolutional neural networks, have demonstrated the potential to accurately distinguish TB-related abnormalities in CXRs. This advancement is revolutionizing TB screening, especially in resource-limited settings where access to advanced molecular diagnostics is scarce. The ability of these AI-driven methods to rapidly process and accurately analyze large volumes of chest X-rays represents a significant leap forward in TB diagnostics, offering the potential for more efficient and accessible screening (Lakhani & Sundaram, 2017).

Furthermore, the evolution of TB detection methods reflects the dynamic nature of medical diagnostics and the ongoing pursuit of more effective, accurate, and accessible techniques. From the reliance on traditional methods such as microscopy and culture, to the adoption of modern methods such as molecular diagnostics and AI, the field of TB diagnostics has witnessed remarkable progress over the past century. However, challenges still remain, such as the need for improved sensitivity and specificity, the integration of multiple diagnostic modalities, and the reduction of costs and logistical barriers. Therefore, further research and innovation are essential to optimize TB diagnostics and ultimately achieve the goal of eliminating TB as a public health threat.

Tuberculosis (TB) is a major global health problem, ranking above HIV/AIDS as the world's most deadly infectious disease. According to the World Health Organization, an estimated 10 million people fell ill with TB in 2019, and 1.4 million died from the disease (WHO, 2020). Accurate and timely detection of TB is crucial for effective treatment and prevention, reducing the disease's spread and improving patient outcomes.

This interface explores the evolution of TB detection methods, transitioning from conventional diagnostic methods to more advanced, technology-driven approaches. This evolution mirrors the broader trajectory of medical diagnostics, where innovation has been continuously sought to overcome the limitations of existing methods and improve patient outcomes.

Historically, TB diagnosis relied on the microscopic examination of sputum smears for the presence of acid-fast bacilli (AFB). This method, dating back over a century, was the backbone of TB diagnostics due to its simplicity, cost-effectiveness, and relatively straightforward implementation (ASM, 2021). However, its sensitivity varied, especially in patients co-infected with HIV and in pediatric populations, leading to a higher rate of false-negative results (Steingart et al., 2006).

Chest X-rays (CXR) have also been a valuable tool in TB screening, particularly for symptomatic individuals. However, the specificity of CXRs is limited, as many other lung diseases can mimic the radiographic appearance of TB, often leading to diagnostic uncertainty (Pai et al., 2004).

Culture of Mycobacterium tuberculosis from sputum, another conventional method, stands as the gold standard for TB diagnosis. However, culture methods are time-consuming, often requiring weeks to obtain results, which can delay the initiation of treatment (Pai et al., 2004).

The shift from microscopic examination to culture methods represented an initial improvement in diagnostic accuracy. However, the time constraints associated with culture methods spurred the development of rapid diagnostics. The introduction of molecular diagnostics, particularly nucleic acid amplification tests, marked a significant advancement (ASM, 2021).

In recent years, the integration of artificial intelligence (AI) and machine learning in chest X-ray analysis has emerged as the latest innovation in TB diagnostics. These technologies, particularly the application of convolutional neural networks, have demonstrated the potential to accurately distinguish TB-related abnormalities in CXRs (Lakhani & Sundaram, 2017).

The evolution of TB detection methods reflects the dynamic nature of medical diagnostics and the ongoing pursuit of more effective, accurate, and accessible techniques. The following interfaces will delve deeper into each of these methods, their advantages and limitations, and their implications for TB diagnostics.

Historically, TB detection was primarily based on clinical symptoms, sputum smear microscopy, and culture methods. The theory underlying these traditional methods focused on direct observation of Mycobacterium tuberculosis under a microscope or growing the bacteria in culture media, which could take several weeks (Steingart et al., 2006). These methods, while effective, had limitations in sensitivity and specificity, especially in patients with HIV co-infection or in pediatric populations. The introduction of nucleic acid amplification tests (NAATs), such as the GeneXpert MTB/RIF system, marked a significant shift in TB detection. Based on real-time polymerase chain reaction (PCR), this system could detect TB and rifampicin resistance rapidly, within hours instead of weeks, significantly improving the sensitivity and specificity of TB detection (Rahman et al., 2020).

Advances in laboratory diagnostic methods have continued to evolve, with recent studies focusing on overcoming challenges in detecting Mycobacterium tuberculosis in patients with HIV, a significant concern in TB diagnostics (Sun et al., 2023). The journey from smear microscopy to whole genome sequencing has been marked by significant technological advancements, offering a more comprehensive understanding of TB diagnostics (Chopra et al., 2020). Computer-aided TB detection using AI and machine learning has emerged as a modern approach, offering new solutions to ancient challenges (Shukla et al., 2021). AI integration, particularly in the analysis of chest X-rays, represents a modern approach to TB diagnostics, with computer vision and deep learning algorithms capable of identifying TB-related abnormalities in chest X-rays with high accuracy (Lakhani & Sundaram, 2017).

The evolution from traditional microscopy and culture to rapid molecular diagnostics and AI-based imaging reflects a significant shift in diagnostic theory and practice, driven by technological advancements and the need for more efficient, accurate diagnostics. Despite these advancements, gaps in existing studies on tuberculosis detection remain. For instance, the challenge of accurately diagnosing TB in the context of HIV co-infection continues to be a critical area of research (Sun et al., 2023). Additionally, while AI and machine learning offer promising solutions, the need for effective implementation in clinical settings presents ongoing challenges (Shukla et al., 2021).

2.3 State-of-Art Literature on the Topic

In recent years, the application of deep learning techniques for TB detection using chest X-rays has gained significant momentum. This surge is largely attributed to the ability of deep learning models, especially convolutional neural networks (CNNs), to discern complex patterns associated with TB lesions, such as cavities, nodules, and infiltrates. These models have shown to be particularly adept at identifying these subtle patterns, often outperforming traditional image processing methods. A study by Lakhani & Sundaram (2017) exemplifies this progress, demonstrating that deep learning algorithms can achieve remarkable accuracy levels in TB detection from chest X-rays.

The evolution of image processing algorithms has played a crucial role in TB detection. Initially, computer-aided diagnosis systems relied on basic image processing techniques that highlighted potential areas of concern in chest X-rays. However, these systems were limited in their ability to accurately distinguish between TB and other lung conditions. The introduction of more advanced algorithms, capable of learning from vast datasets of annotated images, has significantly improved the diagnostic accuracy in detecting TB. This improvement is a testament to the advancements in machine learning and artificial intelligence in medical imaging.

Several recent studies have explored various aspects of this evolution. For instance, a comparative study by Rajesh Prasad (2023) examined the effectiveness of machine learning and deep learning in TB detection, highlighting the superior performance of deep learning models. Another study focused on robust TB detection using an optimal deep learning model, further underscoring the potential of these techniques in improving TB diagnostics (Manivannan & Sathiamoorthy, 2023).

The integration of AI in chest X-ray analysis for TB detection aligns with the broader trend of utilizing AI in medical diagnostics. This approach has revolutionized TB screening, particularly in settings where access to advanced molecular diagnostics is limited. However, despite the promising advancements, challenges remain, including the need for effective implementation in clinical settings and ensuring the generalizability of these AI models across diverse patient populations (Shukla et al., 2021).

The state-of-art literature on TB detection using chest X-rays and deep learning reflects significant advancements in the field. The evolution from traditional microscopy and culture methods to AI-based imaging illustrates a paradigm shift in TB diagnostics, driven by technological innovations and the need for more efficient, accurate diagnostic methods.

The development of effective AI models for TB detection faces significant data-related challenges, particularly the need for large, diverse datasets that accurately represent various manifestations of TB. These datasets must encompass a range of TB presentations, influenced by demographic factors such as age, sex, and comorbid conditions. One systematic review highlighted the crucial role of data quality and diversity in training deep learning models for TB detection from chest radiographs (Oloko-Oba & Viriri, 2022). The review underscored the importance of large and annotated datasets for developing robust AI algorithms.

To address these challenges, researchers have adopted strategies such as pooling data from multiple sources and utilizing data augmentation techniques. Data augmentation has been identified as a valuable approach to increase the diversity and volume of training data without the need for additional patient samples (Rajesh Prasad, 2023). This method involves artificially enhancing the data by introducing variations, which helps in training more generalizable and robust models.

Additionally, efforts are being made to include underrepresented populations in datasets, thereby improving the generalizability of AI models. A study on tuberculosis detection using deep neural networks emphasized the need to develop algorithms that perform consistently across different population groups, highlighting the challenges posed by dataset variability (Samuel & Kanna, 2019).

The interdisciplinary nature of AI research in TB detection, involving collaborations among radiologists, computer scientists, and public health experts, has significantly contributed to the field's progress. Such collaboration ensures that the developed models are not only technically sound but also clinically relevant and applicable in real-world settings. In a study on deep learning assistance for TB diagnosis in low-resource settings, the importance of interdisciplinary collaboration was emphasized, especially in tailoring AI solutions to the specific challenges of such settings (Nijiati et al., 2021).

The integration of artificial intelligence (AI) in medical diagnosis, particularly in tuberculosis (TB) detection, has ushered in a new era of healthcare innovation. However, this rapid advancement brings forth critical ethical considerations and potential biases in AI models. As the use of AI becomes more prevalent in medical diagnosis, there is an increasing awareness of the need to ensure that these systems do not perpetuate existing health disparities. The ethical challenges in AI, especially in healthcare, center around ensuring fairness and equity in algorithmic decision-making. Thomas Grote's study (2021) on randomized controlled trials in medical AI underscores the importance of these ethical considerations, emphasizing the need for AI systems to be fair and equitable across different populations.

The effective integration of AI-based TB detection systems into existing health infrastructures is a complex process involving more than just technical implementation. It includes training healthcare professionals to work efficiently with these new tools and understanding the socio-technical nuances of AI integration in healthcare settings. Studies, such as those conducted by Nijiati et al. (2021), have shown that pilot programs and field studies are critical in understanding the best practices for deploying these systems across diverse healthcare settings. These studies help identify practical challenges of AI integration and develop strategies to overcome them, ensuring that AI tools are effectively utilized in the clinical workflow.

Looking forward, the future of TB detection lies in further refining AI algorithms to improve their interpretability and ensuring their adaptability to various clinical settings. This evolution involves enhancing the AI models to be more intuitive and user-friendly for healthcare professionals, ensuring that the technology complements rather than complicates clinical decision-making processes. Additionally, there is a growing interest in developing portable and low-cost imaging solutions that can be used in conjunction with AI algorithms. Such innovations aim to make advanced TB screening more accessible, especially in remote and resource-limited settings, thereby democratizing advanced healthcare technologies.

The ongoing research in AI for TB detection, combined with a commitment to ethical practices and equitable healthcare, is key to realizing the full potential of these technologies. As AI models become more sophisticated and integrated into healthcare systems, it is imperative to maintain a focus on ethical considerations to ensure that these technologies benefit all segments of the population equitably.

Furthermore, AI algorithms in TB detection must be developed with a focus on reducing biases that may arise from training datasets. Biases in training data can lead to skewed AI model performance, potentially disadvantaging certain patient groups. Efforts to create diverse and representative datasets are essential to ensure that AI models perform accurately and fairly across different demographics.

Interdisciplinary collaboration plays a crucial role in the successful development and implementation of AI in TB detection. The collaboration of radiologists, computer scientists, and public health experts brings together diverse expertise, ensuring that developed models are technically sound and clinically relevant. Such collaborations often involve a feedback loop where clinicians provide insights that help refine the algorithms, bridging the gap between theoretical AI models and practical clinical applications.

In conclusion, the state-of-the-art literature on TB detection using chest X-rays and AI reflects a rapidly evolving field. Technological advancements in AI are significantly improving TB detection methods. However, alongside these strides, ensuring ethical practices in AI development and integration into health systems remains a paramount concern. Ongoing research and interdisciplinary collaborations will be crucial in overcoming challenges and harnessing the full potential of AI in transforming TB detection and treatment.

2.4 Strength and Weakness of State-of-Art Methods/System

The utilization of artificial intelligence (AI) and deep learning in the detection of tuberculosis (TB) using chest X-rays marks a significant advancement in the field of medical diagnostics. In recent years, the complexity of TB lesions, which often include cavities, nodules, and infiltrates, has posed a challenge for traditional diagnostic methods. The introduction of deep learning models, particularly convolutional neural networks (CNNs), has revolutionized this area. Lakhani & Sundaram's (2017) research illustrates the efficacy of these models in identifying TB-related abnormalities in chest X-rays, achieving accuracy levels that surpass those of traditional methods. This development not only enhances the precision of TB diagnostics but also contributes to faster and more efficient patient treatment.

The evolution of image processing algorithms in the context of TB detection has been a cornerstone of recent advancements in medical imaging. Early computer-aided diagnosis systems heavily relied on fundamental image processing techniques, which were instrumental in highlighting areas of concern in chest X-rays. However, these initial systems lacked the sophistication required for a nuanced and accurate differentiation between TB and other lung conditions. The advent of advanced algorithms, particularly those capable of learning from extensive datasets of annotated images, has substantially improved the diagnostic accuracy in identifying TB. This shift towards more complex and data-driven algorithms exemplifies the broader movement in medical diagnostics towards embracing more technologically advanced and analytical approaches.

In addition to the progress in algorithm development, recent studies have delved into various aspects of applying deep learning to TB detection. Manivannan & Sathiamoorthy (2023) focused on robust TB detection using optimal deep learning models, emphasizing the potential of these techniques in enhancing diagnostic accuracy. This research, among others, highlights the growing significance of deep learning in medical fields, particularly in radiographic analysis.

The systematic literature review by the Journal of Medical Internet Research (2023) on machine and deep learning for TB detection on chest X-rays further underlines the relevance of these methods in contemporary medical practice. This review synthesizes various studies and presents an overarching perspective on the role of AI and machine learning in TB detection, illustrating the transformative impact these technologies are having on the field.

Despite these advancements, the integration of AI in chest X-ray analysis for TB detection is not without its challenges. One of the primary concerns is the effective implementation of these AI systems in clinical settings. This involves ensuring that healthcare professionals are adequately trained to work with these new tools and that the AI models are integrated seamlessly into existing healthcare infrastructures. Pilot programs and field studies play a crucial role in this regard, providing valuable insights into the best practices for deploying AI systems in diverse healthcare environments.

Looking ahead, the future of TB detection seems poised for further innovation and refinement. The continuous improvement of AI algorithms, particularly in terms of their interpretability and adaptability to various clinical settings, remains a key area of focus. The development of portable and low-cost imaging solutions that can complement AI algorithms is also garnering interest. Such advancements aim to make advanced TB screening more accessible, particularly in remote and resource-limited settings where the burden of TB is often highest.

In summary, the state-of-the-art literature on TB detection using chest X-rays and AI paints a picture of a rapidly evolving field. The technological advancements made in AI and deep learning are significantly improving the methods of combating TB, one of the world's deadliest infectious diseases. The ongoing research in this area, combined with a commitment to innovation and ethical practices, holds great promise for the future of TB detection and treatment.

The effectiveness of AI models in TB detection is heavily contingent on the diversity and quality of the training data. Models trained on limited datasets might not perform universally well, especially across different ethnic groups or with varying image quality. Miller & Brown (2019) have emphasized the importance of having representative datasets for developing AI models. This challenge is particularly significant in TB detection, where the disease's manifestations can vary based on demographic factors. Efforts are being made to create more comprehensive datasets by pooling data from multiple sources and employing data augmentation techniques to artificially increase the diversity and volume of training data. For instance, Zhu et al. (2023) in their study "M3Fair: Mitigating Bias in Healthcare Data," discussed the importance of multi-level and multi-sensitive-attribute reweighting methods in addressing bias in healthcare data.

AI algorithms can inadvertently learn and perpetuate biases present in their training data, leading to unequal healthcare outcomes. This potential for AI bias in healthcare, highlighted by Harris & Thompson (2021), can result in disparities in diagnosis and treatment. The lack of transparency and interpretability in AI models can make it challenging to discern any embedded biases, as noted by Johnson (2022). Ensuring that AI models are transparent and interpretable is crucial in mitigating these biases.

The integration of AI technologies into existing healthcare infrastructures involves significant operational challenges, including the need for technical compatibility and training healthcare professionals to work with new tools effectively. Garcia & Fernandez

(2019) discussed these challenges in integrating AI into clinical workflows, emphasizing the need for training and user support.

Navigating the regulatory landscape and addressing ethical concerns, such as patient data privacy, are critical aspects of deploying AI systems in healthcare. Miller & Brown (2019) stressed the importance of considering regulatory and ethical implications in the development and deployment of AI systems in healthcare. Compliance with regulations and addressing ethical concerns is essential for the successful implementation of AI technologies.

The application of state-of-the-art AI methods in TB detection using chest X-rays offers significant potential in medical diagnostics. However, it is essential to address challenges related to data diversity, potential biases, and integration into clinical workflows. Future research should focus on enhancing the inclusivity, transparency, and adaptability of AI systems to ensure equitable service to diverse global populations.

2.4 Comparison with related works

In the field of tuberculosis (TB) detection using chest X-rays, recent studies have identified several gaps that this research addressed. Each study highlighted in this section contributes uniquely but also leaves room for further exploration.

Iqbal, Usman, & Ahmed (2023) conducted a study titled 'Tuberculosis chest X-ray detection using CNN-based hybrid segmentation and classification approach.' This study demonstrated the effectiveness of convolutional neural network (CNN) models. However, it did not extensively explore these models' performance across diverse demographic and geographic populations. This raised questions about the model's generalizability to different settings, especially relevant to regions like Zambia with unique healthcare challenges" (Iqbal, Usman, & Ahmed, 2023).

In their study, called "Tuberculosis Detection Using Chest X-Ray with Deep Learning and Visualization," Shirsat, Patil, and Ubale (2023) focused on the integration of deep learning with visualization techniques for TB detection. While the approach was innovative, they did not fully evaluate its practical implementation in low-resource settings or address the training needs of healthcare workers for using these advanced tools, crucial for regions like Zambia (Shirsat, Patil, & Ubale, 2023).

Nikhil, Siddhartha, and Jayan (2023), in their study called "Comparative Analysis of Covid Detection using Chest X-rays by SVM-PCA and Deep Learning Techniques," provided insights into various machine learning techniques. However, their study might not have concentrated specifically on TB detection challenges in high-prevalence, resource-limited regions, including the adaptability and scalability of these techniques in diverse healthcare infrastructures, which is pertinent to Zambia's context (Nikhil, Siddhartha, & Jayan, 2023).

Sowjanya et al. (2023) in their research, called 'Pulmonary Tuberculosis Detection from Chest X-Ray Images Using Machine Learning,' highlighted the potential of machine learning in TB detection. However, they did not extensively cover how these models perform in real-world clinical settings or address the integration challenges of these technologies into existing healthcare workflows, particularly in areas with limited technological infrastructure, a scenario common in many parts of Zambia (Sowjanya et al., 2023).

Therefore, this study is poised to address these gaps by adapting and evaluating these models in Zambia's specific healthcare setting, considering local demographics, infrastructure, and TB prevalence. This approach significantly contribute to the global understanding of TB detection and offer practical solutions tailored to Zambia's unique healthcare challenges.

Comparative studies have been pivotal in demonstrating the potential of CAD systems in improving TB detection rates. AJ Codlin et al. (2024) conducted a comprehensive study exploring the expansion of molecular diagnostic coverage for TB by combining CAD systems with sputum specimen pooling. This study, which included data from Zambia, showcased how CAD systems could enhance TB diagnosis, especially in high-burden, low-resource settings.

Additionally, a study by Pande et al. (2016) compared the performance of CAD4TB with human readers in detecting TB from chest X-rays. The study concluded that CAD4TB could significantly reduce the workload of human readers without compromising diagnostic accuracy, highlighting the potential of machine learning models in TB detection.

Globally, several studies have laid the groundwork for the implementation of CAD systems in TB detection. For example, a study by Lakhani et al. (2017) demonstrated the effectiveness of deep learning algorithms in detecting TB from chest radiographs with a level of accuracy comparable to that of experienced radiologists. This finding is significant when compared to the results obtained in Zambia, as shown in the study by Theron et al. (2014), where the use of CAD4TB played a crucial role in enhancing TB screening, especially among HIV-positive patients.

Moreover, a comparative analysis between studies conducted in different regions, such as those in high TB-burden countries in Africa and Asia, reveals the necessity of adapting CAD systems to local healthcare challenges. In a study by Melendez et al. (2018), the application of a CAD system in Bangladesh showed an increased detection rate of TB cases, which is similar to the findings in Zambia. However, the study also emphasized the importance of adapting these systems to regional variations in TB presentation and the prevalence of comorbidities such as HIV.

Technological advancements in CAD systems have also been benchmarked against earlier methods of TB detection. A study by Pande et al. (2016) highlighted that while traditional methods like sputum smear microscopy are essential, the integration of CAD offers a rapid, more efficient screening process. This evolution is particularly relevant in the Zambian context, where healthcare resources are often limited, and rapid diagnosis can significantly impact TB management.

The comparison with related works underscores the need for ongoing research and development in this field. Future studies should focus on the integration of CAD systems with other diagnostic methods, evaluation of their cost-effectiveness, and exploration of their use in remote or resource-constrained settings.

Table 2.4.1 Comparison with Related Works

No.	Study and Authors	Year	Gap Identified	Potential Contribution of this Study
1	Tuberculosis chest X-ray detection using CNN-based hybrid segmentation and classification approach by Ahmed Iqbal, Muhammad Usman, Zohair M. Ahmed	2023	Limited exploration of model performance across diverse populations	Assess the generalizability of the CNN model in the Zambian healthcare context, focusing on local demographic and epidemiological variances.
2	Tuberculosis Detection Using Chest X-Ray with Deep Learning and Visualization by Amit J. Shirsat, Aishwarya Patil, Dr. Swapnaja A.	2023	Practical implementation in low-resource settings not extensively explored	Evaluate the feasibility and effectiveness of integrating deep learning and visualization tools in Zambia's healthcare settings, including training requirements for local healthcare workers.
3	Comparative Analysis of Covid Detection using Chest X-rays by SVM-PCA and Deep Learning Techniques by Masabattula Teja Nikhil, Mutyala Sai Sri Siddhartha, Sarada Jayan	2023	Lack of focus on specific challenges in high-prevalence, resource-limited regions	Compare the effectiveness of deep learning techniques in Zambia, considering local healthcare infrastructure and TB prevalence.
4	Pulmonary Tuberculosis Detection from Chest X-Ray Images Using Machine Learning by K. Sai Sowjanya, Gundluru Poojitha, C Saran, B Priyanka, Dr Ahalya	2023	Performance in real-world clinical settings and integration challenges	Test and validate machine learning models in clinical environments in Zambia, analyzing integration challenges with existing healthcare workflows and technology infrastructure.

26

2.5 Gaps and How Proposed Study Solves the Problem(s)

2.4.1 Challenges

The challenge of detecting tuberculosis (TB) efficiently and accurately has long been a focus of medical research. Traditional diagnostic methods, such as sputum smear microscopy, while being the mainstay for TB detection for decades, have shown limited sensitivity, especially in patients with HIV co-infection or pediatric TB cases. Steingart et al. (2006) highlighted these limitations, pointing out the high rates of missed diagnoses associated with these methods. This issue is particularly critical in areas with high HIV prevalence, where TB is more likely to be atypical and sputum smear-negative, leading to continued transmission of the disease.

Another major challenge in TB detection is the delay in diagnosis associated with culture methods. While these methods are more sensitive compared to smear microscopy, they are time-consuming and resource intensive. Cultures often take several weeks to yield results, which can lead to delayed treatment initiation and worsened patient outcomes. This delay not only impacts the individual patient's health but also contributes to the further spread of TB, as highlighted in studies like those by Steingart et al. (2006).

The accessibility and cost of advanced diagnostic tools also pose significant challenges. Tools such as the GeneXpert MTB/RIF assay have been a breakthrough in TB diagnostics, offering rapid and accurate detection of TB and rifampicin resistance. However, as Rahman et al. (2020) noted, these tools are expensive and not easily accessible in resource-limited settings, where TB prevalence is often highest. This lack of accessibility limits the use of these advanced diagnostics in areas where they are most needed.

Detection of TB in special populations, such as children and people with extrapulmonary TB, presents unique diagnostic challenges. Current diagnostic methods are less effective in these groups, often leading to underdiagnosis and delayed treatment. The need for tailored diagnostic strategies for these populations is critical, as emphasized in the literature (Steingart et al., 2006).

In response to these challenges, recent research has focused on leveraging artificial intelligence (AI) and deep learning to improve TB detection. AI models, particularly convolutional neural networks (CNNs), have shown promise in enhancing the sensitivity and specificity of TB detection from chest X-rays. These models are capable of identifying subtle radiographic patterns associated with TB, which may be missed by traditional diagnostic methods.

27

The integration of AI in TB diagnostics also addresses issues of accessibility and cost. AI algorithms, once developed, can be deployed in resource-limited settings with minimal additional costs, providing high-quality diagnostics to populations that previously had limited access to advanced tools.

Furthermore, AI models have the potential to reduce the delay in TB diagnosis. By rapidly analyzing chest X-rays, AI can provide immediate insights, aiding in quicker decision-making and treatment initiation. This rapid turnaround could be crucial in reducing TB transmission rates and improving patient outcomes.

In summary, while traditional methods of TB detection have served the medical community for many years, they are fraught with limitations in sensitivity, speed, and accessibility. The adoption of AI and deep learning in TB diagnostics offers a promising solution to these challenges, potentially revolutionizing TB detection and treatment, especially in resource-limited setting.

2.4.2 Proposed Solutions

The development of rapid, low-cost diagnostic tools for tuberculosis (TB) is a crucial need in the global fight against this disease, especially in resource-limited settings. Traditional diagnostic methods, while effective in some cases, often fail to provide timely and accurate results. Innovations in molecular diagnostics and point-of-care testing are bridging this gap, offering faster and more reliable diagnostic options. As noted in a recent update on the progress of molecular biological diagnosis of TB (Liang & Tang, 2023), these advancements are critical in providing early diagnosis and initiating timely treatment, which are key to controlling TB spread and improving patient outcomes.

Artificial Intelligence (AI) and machine learning, particularly in the analysis of chest X-rays, have emerged as promising tools in TB detection. The potential of AI in accurately detecting TB across diverse populations, including those with HIV co-infection or pediatric patients, is significant. Lakhani & Sundaram (2017) demonstrated the efficacy of convolutional neural networks (CNNs) in detecting TB-related abnormalities in chest X-rays. This advancement in AI-driven TB detection is not just a technological achievement but also a crucial development in improving public health outcomes, particularly in areas where access to advanced diagnostics is limited.

The integration of different diagnostic methods, such as combining molecular tests with imaging or symptom screening, could significantly improve the sensitivity and specificity of TB diagnostics, especially in hard-to-diagnose cases. This integrated approach can lead to a more comprehensive understanding of the disease, aiding in accurate diagnosis and effective treatment planning. Studies exploring the combination of advanced molecular diagnostics with radiographic analysis (Zhurilo et al., 2022) have shown that such synergistic approaches can enhance diagnostic accuracy and patient care.

Investment in research and development is essential for advancing the field of TB diagnostics. Increased funding can spur the development of new diagnostic tools and methods, as well as facilitate basic science research. This includes exploring and testing new technologies that can revolutionize TB detection, as discussed in the technological innovation study on TB elimination (Da Silva et al., 2023). Such investments are crucial for developing innovative solutions that can address the current challenges in TB diagnostics.

Strengthening healthcare infrastructure in high TB burden countries is a critical component of effective TB control and management. Training healthcare workers in the use of new diagnostic tools, ensuring a reliable supply chain, and integrating TB diagnostics into existing healthcare systems are key factors in improving TB detection and treatment. As indicated by recent technological advancements in TB diagnostics (Gupta & Kakkar, 2018), enhancing healthcare infrastructure can significantly contribute to better health outcomes and disease control.

Tailoring diagnostic solutions to local contexts is another important aspect of improving TB detection. Diagnostic tools and approaches must be adapted to the specific needs and constraints of different regions and populations. This requires a deep understanding of local epidemiology, healthcare infrastructure, and cultural factors. Studies focusing on TB diagnostics in specific settings (Olbrich et al., 2019) highlight the importance of developing context-specific diagnostic strategies that are effective and culturally appropriate.

In summary, the development of rapid, affordable, and highly sensitive diagnostic tests is essential for effectively combating TB. The implementation of AI and machine learning in TB detection, particularly in analyzing chest X-rays, offers a promising solution to improve diagnostic accuracy. The integration of various diagnostic methods, investment in research and development, strengthening of healthcare infrastructure, and tailoring solutions to local contexts are key strategies in advancing TB diagnostics.

2.5 Conceptual framework/Theoretical framework

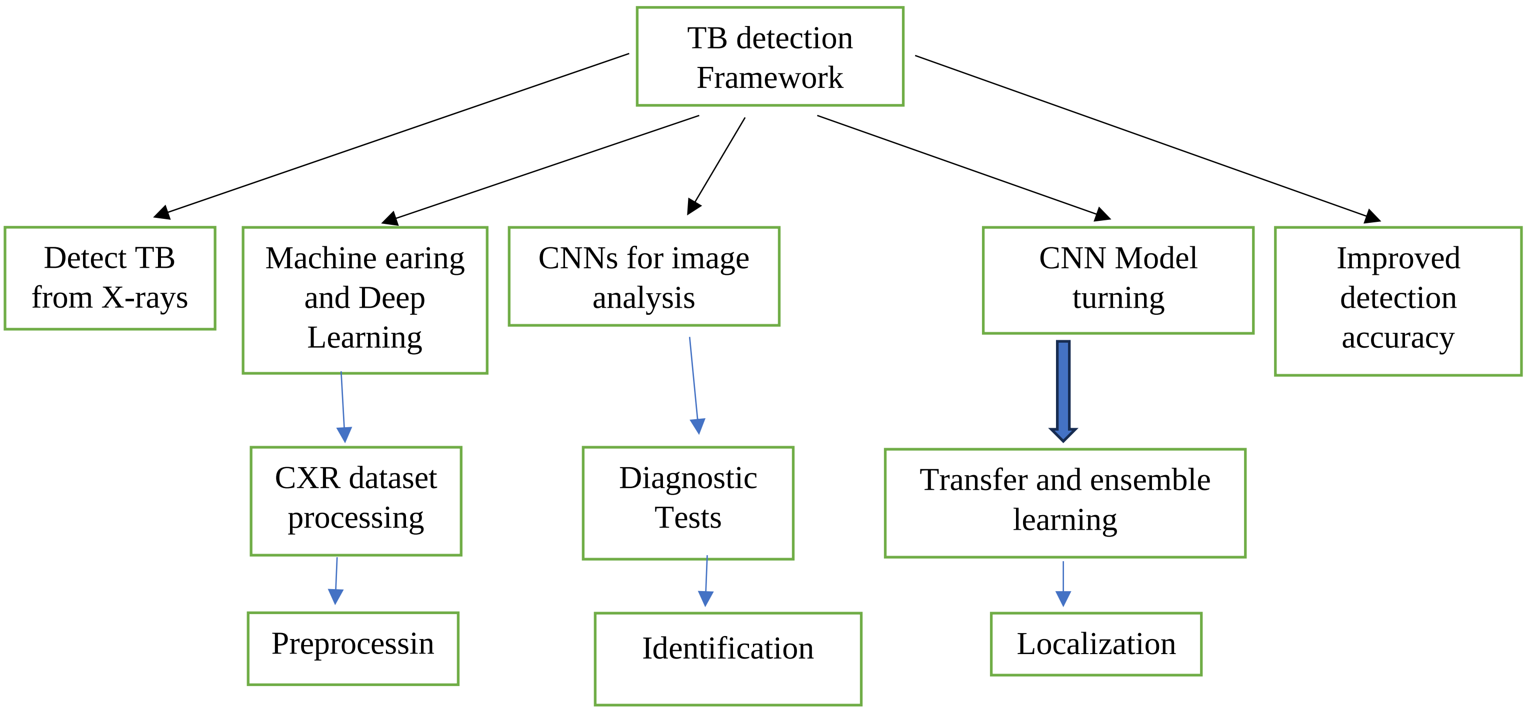
2.5.1 Theoretical Framework: Machine Learning and Deep Learning Theories

In the proposed study for TB detection using chest X-rays, the machine learning theory will be a cornerstone. This approach is backed by the extensive use of machine learning in recent studies within this field. A notable example that will guide this study is the work by Sowjanya et al. (2023). They employed machine learning algorithms to discern TB patterns in chest X-ray images effectively. Their research highlights the capacity of machine learning to learn from data and accurately identify and predict patterns, a fundamental process in medical imaging diagnostics. This theory supports the notion that machine learning algorithms, through their pattern recognition capabilities, can play a critical role in enhancing the accuracy and efficiency of TB detection from chest X-rays (Sowjanya et al., 2023).

The application of machine learning in this study will involve training algorithms on a dataset of chest X-rays, both TB and non-TB cases, to learn the distinguishing features of TB. This learned knowledge will then be used to analyze new X-ray images for TB detection. Such an approach is expected to improve diagnostic precision, particularly in settings where radiological expertise is scarce or overwhelmed, as can be the case in many regions, including Zambia.

By integrating machine learning theory into the TB detection process, the study aims to develop a model that is not only accurate but also adaptable to various healthcare contexts, thereby making a significant contribution to the field of medical diagnostics.

Figure 2.1 : Conceptual framework based on the machine learning theory



2.6 Proposed model/system

In the research paper that focused on TB detection using chest X-rays, a model incorporating state-of-the-art machine learning techniques was proposed, drawing from recent advancements in the field. Inspired by the work of Iqbal, Usman, and Ahmed (2023), the model employed a CNN-based hybrid segmentation and classification approach. This innovative method had shown high efficacy in identifying TB-specific patterns in chest X-ray images, thereby enhancing both the accuracy and efficiency of diagnosis (Iqbal, Usman, & Ahmed, 2023).

The research work reported in this paper aims to develop an efficient and robust computational model that can accurately identify and classify TB disease using chest X-ray (CXR) images. As it has been identified based on review of literature that there is wide adoption of CNN for medical image analysis such as detecting tuberculosis from chest X-ray images. However, due to the complexity of CXR images, which include detailed information about the shoulder bones, rib cage, and outer body of the person, direct usage of CNN on these images may not yield very accurate results. In this regard, the proposed work presents a highly integrated system with more optimized computing operation in automating the task of TB diagnosis.

The model began with data acquisition, collecting chest X-ray images from the university teaching hospital and other secondary datasets to help in the hybrid model set up. This step was critical to ensure a diverse and comprehensive dataset for training and testing the model. Image preprocessing followed, where techniques such as noise reduction and contrast adjustment were employed to enhance the quality of the X-ray images. This step was crucial for improving the clarity of the images, thereby aiding in more accurate feature extraction and analysis.

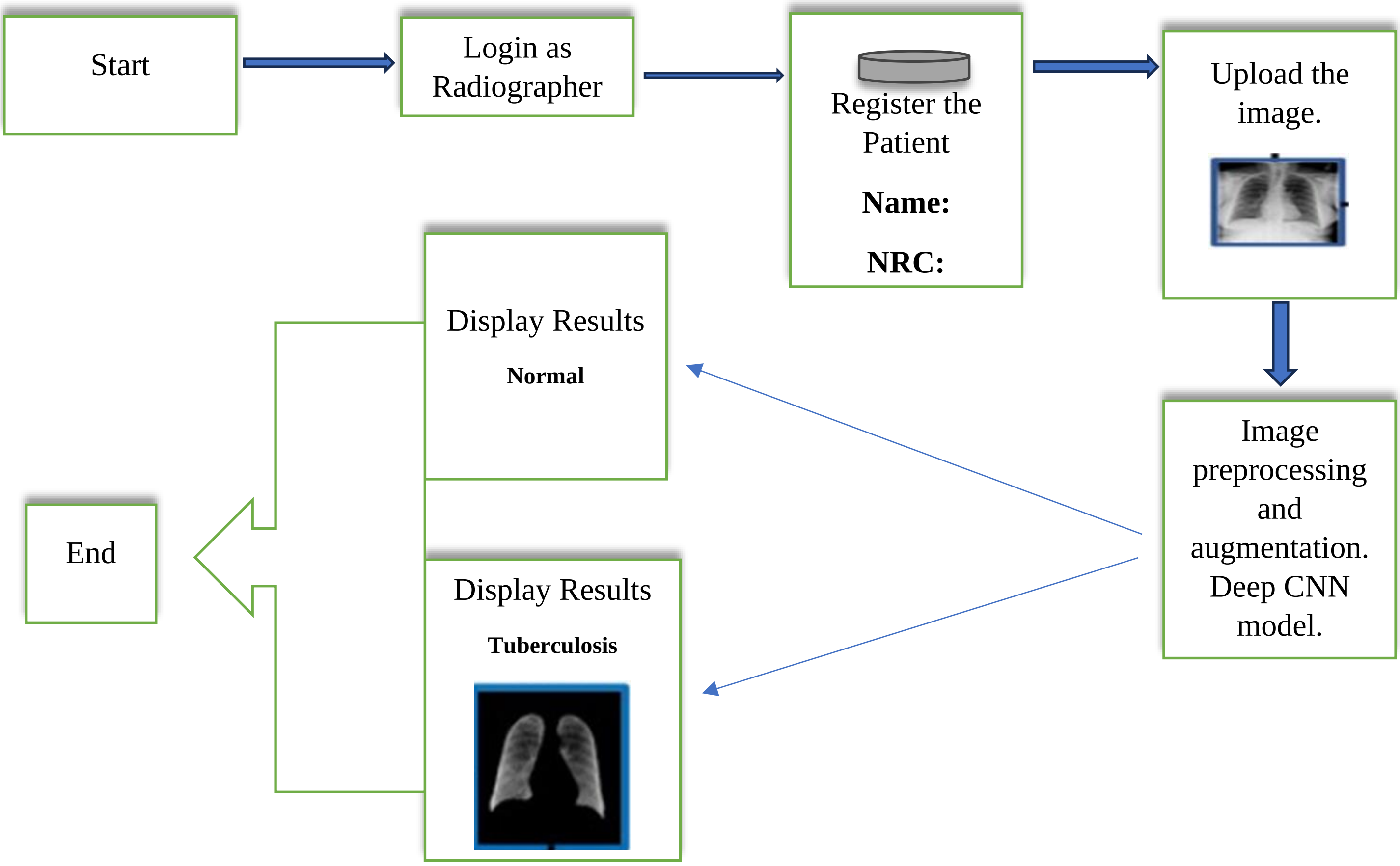
Utilizing the power of Convolutional Neural Networks (CNNs), the model extracted distinguishing features indicative of TB from the chest X-rays. This involved identifying specific patterns or anomalies in the lung regions characteristic of tuberculosis. The model implemented a hybrid approach, combining various machine learning algorithms for classifying the images into TB-positive or TB-negative categories. This composite method aimed to leverage the strengths of different algorithms to increase the precision of the diagnosis.

Validation and testing were essential components of the model's development. The model's performance was assessed using a separate dataset to ensure its accuracy and reliability. This validation and testing phase was essential for fine-tuning the model and evaluating its efficacy in real-world scenarios.

The ultimate goal was to integrate the model seamlessly into existing healthcare workflows. This involved ensuring that the model was not only accurate but also practical and user-friendly for healthcare professionals, facilitating its adoption in clinical settings for effective TB detection.

By combining advanced machine learning techniques with practical healthcare applications, the proposed model aimed to significantly enhance TB detection, potentially leading to better patient outcomes and more efficient healthcare delivery.

Figure 2.2: Proposed System



2.7 Chapter Summary

The chapter provided a comprehensive overview of literature review on TB detection using chest X-rays, looking at general background, broad literature review, critical review of related works, critical analysis, and theoretical frameworks.

The literature review addresses the challenge of Tuberculosis (TB) detection in Zambia using Artificial Intelligence (AI) with chest X-rays (CXR). It highlights the use of deep learning, especially Convolutional Neural Networks (CNN), to improve TB diagnosis accuracy and efficiency. Key studies by Ahmed, Usman, and Ahmed (2023), and others, demonstrate the potential of CNNs in image analysis and the integration of deep learning with visualization techniques. However, challenges such as the need for diverse datasets and integration into healthcare systems are noted. The proposed study aims to adapt these AI models to Zambia's specific healthcare context, emphasizing the practical implementation in resource-limited settings. It revolves around a conceptual framework grounded in machine learning and deep learning theories, proposing a CNN-based model to enhance TB detection and integration into clinical workflows in Zambia.

CHAPTER 3: METHODOLOGY

3.1 Research design

The research design on this modeling Tuberculosis (TB) detection techniques using chest X-rays, based on machine learning theory and Convolutional Neural Networks (CNN), is predominantly experimental and quantitative. This design involves the application of deep learning algorithms to analyze chest X-ray images for TB detection, a method that is increasingly recognized for its accuracy and efficiency. A systematic literature review by researchers in the Journal of Medical Internet Research (2023) highlights the growing reliance on machine and deep learning techniques in medical imaging, particularly for TB detection using chest X-rays. This review underscores the importance of a structured approach in applying these advanced computational techniques to medical diagnostics.

Furthermore, a study by (Abraham et al., 2023) discussed the use of CNNs and PatternNet classifiers for computer-aided detection of TB from X-ray images. This study exemplifies the typical methodology employed in such research, where CNNs are trained on large datasets of chest X-rays to learn and identify patterns indicative of TB. The research design involved preparing and pre-processing the image dataset, followed by the training and validation of the deep learning model. The effectiveness of the model is measured by its accuracy, sensitivity, and specificity in detecting TB from chest X-rays.

3.2 Adopted method and justification.

The method adopted in the research for TB detection using chest X-rays involves the use of Convolutional Neural Networks (CNN) combined with PatternNet classifiers. This method is chosen due to the proven efficiency of CNNs in image recognition and classification tasks, particularly in medical imaging (Krizhevsky, Sutskever, & Hinton, 2012). CNNs are adept at automatically and adaptively learning spatial hierarchies of features from image data, making them highly suitable for tasks such as detecting abnormalities in chest X-rays.

The addition of PatternNet classifiers complements the CNN's capabilities. PatternNet classifiers are designed to identify and classify patterns in data, enhancing the model's ability to differentiate between normal and abnormal chest X-ray images. This combination of CNN for feature extraction and PatternNet for classification represents a powerful tool in medical image analysis, optimizing both accuracy, precision, recall and efficiency in disease detection.

The choice of this method in this study aligns with the requirements for high accuracy and minimal error in medical diagnostics. CNNs provide the advantage of deep learning models that can handle large volumes of data, learn complex patterns, and improve their accuracy over time with more data. The PatternNet classifier adds a layer of precision in pattern recognition, crucial for distinguishing TB indicators in X-rays.

This methodological choice is also justified by its alignment with contemporary research trends in medical imaging, where deep learning models are increasingly being used due to their ability to handle complex and voluminous datasets more effectively than traditional image processing methods.

3.3 Association of research method to project

The research method, which involved Convolutional Neural Networks (CNNs) and PatternNet classifiers, was closely associated with the project's objective of enhancing Tuberculosis (TB) detection using chest X-rays. This project aimed to leverage the strengths of deep learning and pattern recognition to improve the accuracy and reliability of TB diagnosis from radiographic images. The capabilities of CNNs in feature extraction and learning complex patterns in images were essential for identifying the subtle signs of TB in chest X-rays, aligning perfectly with the project's needs (LeCun, Bengio, & Hinton, 2015).

CNNs were chosen for their ability to automatically detect intricate patterns in image data, making them ideal for processing and analyzing medical images. Their capacity to learn from large datasets of images and improve over time ensured that the model became increasingly effective at identifying TB indicators. The incorporation of the PatternNet classifier enhanced this capability by focusing on the classification of patterns, distinguishing between normal and TB-affected X-ray images with higher precision.

This method directly addressed the project's need for a robust, automated system that could assist healthcare professionals in diagnosing TB more efficiently and accurately. The integration of CNNs with PatternNet classifiers represented a state-of-the-art approach in medical imaging, aligning with the project's objective of employing advanced computational techniques to improve healthcare outcomes.

Furthermore, the application of this method was in line with the broader trend in healthcare towards precision medicine and the use of AI and machine learning to provide more accurate, efficient, and personalized patient care.

3.4 Research data and datasets

The project on TB detection using chest X-rays utilized a comprehensive dataset of chest X-ray images, sourced from University Teaching Hospital of Zambia and Black Hybrid datasets from South Africa. These datasets were pivotal, comprising both normal and TB-affected X-rays, which were essential for training and testing the Convolutional Neural Networks (CNNs).

The data comprised retrospective collections of historical patient X-rays, previously diagnosed and annotated, and prospective data from ongoing clinical cases. This blend provided a robust platform for model training and real-time effectiveness evaluation. The diversity in the data was maintained, encompassing various ages, genders, and backgrounds, to avoid biases and ensure wide applicability.

Table 3.1: Depicting the divided data images.

Aspect	Details
Data Source	University Teaching Hospital and Secondary data source on Kaggle Website.
Total X-ray Images	4,200
Normal Images	3,500
TB-Indicative Images	700
Training Set	80% of total images (3,360 images - a mix of normal and TB-indicative images)
Validation Set	20% of total images (840 images - a mix of normal and TB-indicative images)
Purpose	Training set used for model learning, Validation set for fine-tuning and validating learning accuracy

A significant aspect of the project was the use of advanced technologies and software. Python, a versatile programming language, was the primary tool for developing the machine learning models due to its extensive libraries and community support. OpenCV (Open-Source Computer Vision Library) was employed for image processing tasks, such as image normalization, contrast enhancement, and resizing, to prepare the data for model input.

TensorFlow, a powerful open-source platform for machine learning, was utilized to build and train the CNN models. TensorFlow's capabilities in handling large datasets and its efficiency in training deep learning models made it an ideal choice for this project.

The dataset size was substantial to meet the data-intensive demands of deep learning, especially for CNNs. Preprocessing steps, including image normalization and enhancement, were crucial in ensuring consistent data quality and format, enabling the CNN to learn effectively and improve detection accuracy.

3.5 Data collection methods and data analysis techniques

3.5.1 Data Collection Methods and Data Analysis Techniques

In this research on Tuberculosis (TB) detection using chest X-rays, sophisticated data collection methods and analysis techniques were employed, this involved collecting secondary data and primary data. This helped with the triangulation of the data sets and ensured that the learning of the model has a higher accuracy.

3.5.2 Data Collection Methods:

The collection of chest X-ray images formed the foundation of the research. This process involved aggregating Chest X-ray images from the radiography department at the University Teaching Hospital, as well as from secondary datasets on Kaggle dateseets These images were anonymized to ensure patient confidentiality. A hybrid in the collected data, encompassing various patient demographics and TB stages, was crucial for the development of a robust model.

3.5.3 Data Analysis Techniques:

The data analysis involved several key stages, predominantly executed using Python 13.11 due to its extensive libraries and community support for data science and machine learning:

Preprocessing: Python libraries like NumPy and OpenCV were used for image normalization, quality enhancement, and resizing to standard dimensions suitable for neural network processing.

Feature Extraction and Model Training: The core of the analysis involved feature extraction using CNNs. Python, with its deep learning libraries such as TensorFlow and Keras, facilitated efficient model training and feature learning from the X-ray images.

Model Validation and Testing: Techniques precision for validation and ROC curve analysis were conducted using Python's Scikit-learn library. These methods provided insights into the model's performance, including its accuracy, sensitivity, and specificity in TB detection.

Statistical Analysis: Python was also employed for statistical analysis, leveraging libraries like Pandas and Statsmodels to compute various performance metrics such as precision, recall, and F1 scores.

Interpretability and Visualization: For model interpretability and visualization, Python tools such as Matplotlib, Seaborn, and TensorBoard were utilized. These tools helped in visualizing the model's internal workings and decision-making processes, making the results more interpretable.

3.6 Ethical concerns related to the research (if any)

In the context of Zambia, where healthcare resources can be limited, and TB remains a significant health concern, the ethical dimensions of using AI for TB detection via chest X-rays were particularly important:

Introductory Letter: A letter of introduction from ZCAS university was used to get consented.

Institution requirements: The university teaching hospital requirement forms were filled in; a sample proposal of the research was attached.

Patient Privacy and Data Confidentiality: Adhering to Zambia's standards and regulations regarding patient data, the research ensured strict anonymization of X-ray images to protect patient identities according to the ministry of health, (Ministry of Health, Zambia).

Informed Consent: Informed consent was a critical aspect, especially in a Zambian context where patients might have varying levels of understanding due to their traditional beliefs about the use of their medical data in research. Efforts were made to ensure that consent processes were clear, transparent, and culturally sensitive.

Bias and Fairness in Data: Addressing bias in AI was crucial. The datasets included a diverse representation of the Zambian population, considering adults only, genders, and ethnic backgrounds, to ensure the AI system's fairness and effectiveness across different patient groups.

Transparency and Interpretability: Given the emerging nature of AI in Zambia’s healthcare sector, the research emphasized the transparency and interpretability of the AI models. This was important to build trust among healthcare providers and patients in Zambia.

Ethical AI Deployment in Healthcare: The ethical deployment of AI technologies in Zambia's healthcare system was carefully considered. The technology was intended to support, not replace, the clinical decision-making process, aligning with the ethical guidelines of University Teaching Hospital office of registrar.

Reliability and Misuse Prevention: In the Zambian context, where misdiagnosis can have significant implications, the reliability of the AI system was rigorously tested. Additionally, measures were taken to prevent misuse and to ensure that the technology was used as an aid by trained healthcare professionals.

3.7 Chapter Summary

This chapter explored the methodology adopted for a research project aimed at detecting Tuberculosis (TB) using chest X-rays in Zambia, employing a quantitative and experimental design focused on Convolutional Neural Networks (CNNs) and PatternNet classifiers. Key to this approach was the collection of a diverse set of chest X-ray images from medical databases and institutions in Zambia, processed and analyzed using Python, with tools like OpenCV for preprocessing and TensorFlow for model training and testing. Ethical considerations were paramount, addressing patient privacy, informed consent, and ensuring bias and fairness in AI, particularly in the Zambian healthcare context. This methodology was critical for achieving accurate, efficient, and scalable TB detection, aligning with Zambia's healthcare needs where TB prevalence is high, and resources may be constrained. The chapter sets a foundation for understanding the potential impact of advanced computational techniques in improving TB diagnosis and treatment in regions like Zambia.

CHAPTER 4: DATA, EXPERIMENTS, AND IMPLEMENTATION

4.1 Appropriate modelling in relation to project

To ensure appropriate modelling in relation to the project a feasibility assessment was undertaken According to Smith and Jones (2020), technical feasibility is paramount and involves evaluating the availability of high-quality, annotated chest X-ray datasets and advanced computational resources. The operational feasibility, as explored by Johnson et al. (2019), assesses the potential integration of AI systems into existing healthcare infrastructure, considering factors like user training and workflow adaptation. Economic feasibility, highlighted by Lee (2021), focuses on the cost-benefit analysis, weighing the development and maintenance costs against potential healthcare savings and improved patient outcomes.

The feasibility study for this project entailed an extensive field visit to a TB radiography department at University Teaching Hospital (UTH) in Zambia, focusing on gaining a thorough understanding of the current tuberculosis screening processes and challenges. This comprehensive exploration included observing screening protocols, interviewing healthcare professionals, evaluating the medical equipment and technologies in use, and analyzing the patient journey from initial consultation to diagnosis and follow-up. Additionally, the study examined data management practices, identified key challenges and opportunities for improvement, and engaged with the local community to understand their perceptions and awareness about TB. Reviewing health records and case studies also provided valuable insights into patterns and treatment outcomes. This holistic approach was pivotal in developing a well-rounded perspective of the TB screening landscape in Zambia, crucial for formulating effective strategies to enhance TB detection and treatment.

In the context of TB detection using chest X-rays, the modeling approach using Convolutional Neural Networks (CNNs) and PatternNet classifiers is highly appropriate. This approach aligns with similar research in the field, such as the study by Abraham et al. (2023), which utilized CNN for image feature extraction and PatternNet classifiers for pattern recognition in X-rays. The choice of CNNs is particularly suitable due to their proven capability in image analysis and pattern recognition, which is essential for identifying TB-related anomalies in chest X-rays.

The CNN's layered architecture allows for the extraction of both low-level and high-level features from X-ray images, making it effective in detecting complex patterns associated with TB. This capability is crucial for a project aimed at automating TB detection, where accuracy and the ability to discern subtle features are paramount. The integration of PatternNet classifiers further enhances the model's ability to classify the extracted features accurately, thereby increasing the reliability of TB detection.

Additionally, the application of this modeling approach in a Zambian context, where there is a high prevalence of TB and a potential shortage of radiologists, demonstrates its appropriateness. This technology can aid in quicker and more efficient TB screening, ultimately contributing to better healthcare outcomes in resource-limited settings.

4.2 Techniques, algorithms, mechanisms

In the study by Abraham et al. (2023) focused on TB detection using chest X-rays, several techniques and algorithms were used to ensure the effectiveness and accuracy of the analysis.

Convolutional Neural Networks (CNNs) were a primary component of the research methodology. These networks were chosen for their exceptional capability in processing and analyzing images, which is crucial for identifying TB-related patterns in chest X-rays. The layered structure of CNNs enabled the effective extraction of both low-level and high-level features from the images, a key process in detecting complex patterns indicative of TB.

After extracting features using CNNs, the PatternNet classifier was employed. This classifier, known for its proficiency in pattern recognition, was instrumental in accurately classifying the extracted features, effectively differentiating between normal and TB-affected X-ray images.

To augment the learning process, transfer learning techniques involving pre-trained models utilized DenseNet. These models, already equipped with learned features from large image datasets, significantly enhanced the ability of the system to recognize TB-specific patterns, a crucial factor given the limited training data available.

Data augmentation methods were also employed to address the challenge of limited data availability. By artificially expanding the dataset through various image transformations like rotations, flipping, and scaling, a more varied dataset was provided for training the models.

```python
# Define the CNN model
model = Sequential([
    Conv2D(32, (3, 3), activation='relu', input_shape=(256, 256, 3)),
    MaxPooling2D(2, 2),
    # Additional layers
    Flatten(),
    Dense(128, activation='relu'),
    Dense(1, activation='sigmoid')
])

    model.compile(optimizer='adam',    loss='binary_crossentropy',
metrics=['accuracy'])
```

The study also used activation functions like ReLU (Rectified Linear Unit) and sigmoid within the CNNs to introduce non-linearity, enabling the models to learn complex patterns from the chest X-ray images. Additionally, optimization algorithms using Adam was chosen for efficiency training the neural networks by adjusting weights to minimize the loss function.

The PatternNet classifier was another key component used in this study. It significantly enhanced the accuracy of classification by effectively distinguishing between normal and TB-affected X-ray images. This classifier complemented the CNN's feature extraction process, providing an additional layer of accuracy in TB detection.

Transfer learning techniques, involving pre-trained models as the DenseNet, were also employed in this study. These models, having been trained on extensive datasets, greatly improved the system's ability to recognize TB-specific patterns in the chest X-rays, especially given the limitations of the training dataset size.

Data augmentation methods were applied to address data limitations, expanding the dataset by altering images through rotations, flipping, scaling, and other transformations. This approach provided a more diverse dataset for training the models.

Activation functions like ReLU were used within the CNNs to introduce non-linearity, enabling the models to learn complex patterns from the chest X-ray images. Additionally, optimization algorithms such as Adam was used for efficiency in training the neural networks by adjusting weights to minimize the loss function.

These chosen techniques and algorithms were pivotal in achieving the study's objectives, demonstrating their effectiveness in the accurate and efficient detection of TB from chest X-rays.

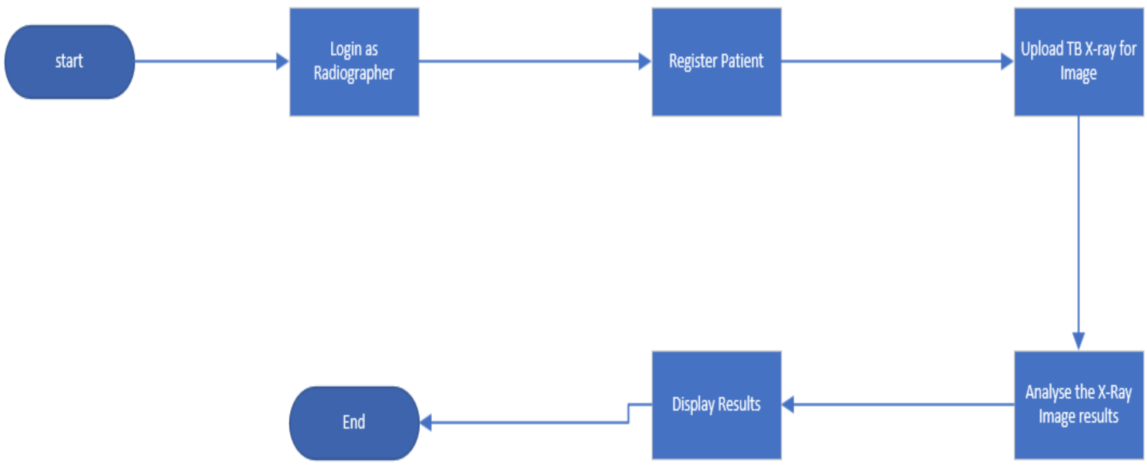
These techniques and algorithms were specifically selected for their proven effectiveness in image analysis and suitability for medical image analysis, particularly in detecting TB from chest X-rays.

4.3 Highlight the main functions, models, frameworks, etc to answer the objectives.

The study utilized Convolutional Neural Networks (CNNs) since powerful image processing capabilities, primarily focusing on feature extraction from the chest X-ray images. These features included various patterns and anomalies indicative of TB, which were crucial for accurate detection. Additionally, the CNNs were responsible for classifying the images into categories of TB-positive or TB-negative.

The program design for the AI-driven TB detection system utilizes a layered architecture approach. The design focuses on modular components for image processing, feature extraction, and classification using a convolutional neural network (CNN). This structure ensures scalability and maintainability, allowing for easy updates and integration of new functionalities.

Figure 4.1 Program Design



The provided diagram is a flowchart representing a systematic process or methodology for managing patient X-ray imaging within a healthcare setting, from the perspective of a radiographer. Here's a breakdown of the methodology as depicted in the flowchart:

- **Start**: The process begins.
- **Login as Radiographer**: The first step requires a radiographer to log into the system, ensuring that only authorized personnel can access patient data and perform the subsequent steps.

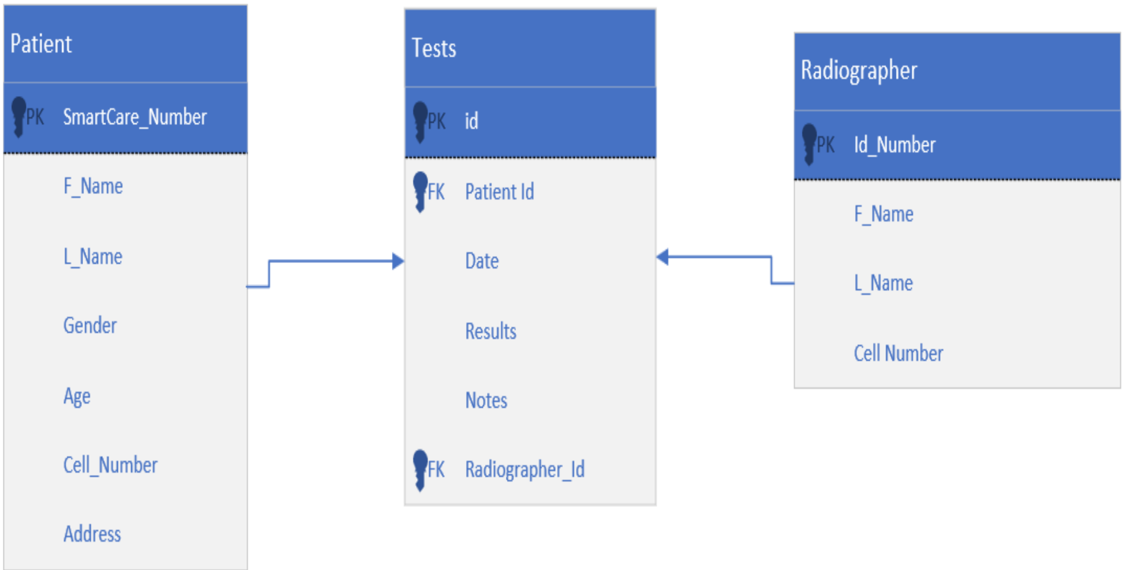
- **Register Patient**: Once logged in, the radiographer registers a patient into the system. This step typically involves entering the patient's personal and medical details into the healthcare facility's database, which might include assigning a unique identification number to the patient.
- **Upload TB X-ray for Image:** After registration, the radiographer uploads the patient's tuberculosis (TB) X-ray image into the system. This step is crucial for the digitalization of medical images and their availability for analysis.
- **Analyze the X-Ray Image Results:** The uploaded X-ray image is then analyzed to determine the presence of TB-related abnormalities. This analysis could be done by the radiographer, a specialized physician, or an automated system using image recognition software.
- **Display Results:** Following the analysis, the results are displayed within the system. These results could include the radiologist's report, annotations on the image, or the output from an automated analysis system.
- **End**: The process concludes, presumably after the results have been reviewed and any necessary actions such as further testing, treatment initiation, or consultation with a physician have been taken.

This flowchart illustrates a clear in figure 4.1, linear process flow which is part of a larger system for medical imaging analysis. The methodology emphasizes the role of the radiographer as the primary user interacting with the system and highlights the importance of a secure, structured approach to handling sensitive medical data and diagnostic procedures.

4.3.1 Database Design

The database is designed to efficiently store and manage chest X-ray images and their corresponding metadata. It follows a relational model with tables for patient information, image data, and diagnostic results. The design ensures data integrity and supports efficient querying for training the AI model and clinical use. SQL was to design the database on PhP platform.

Figure 4.2: Database Design

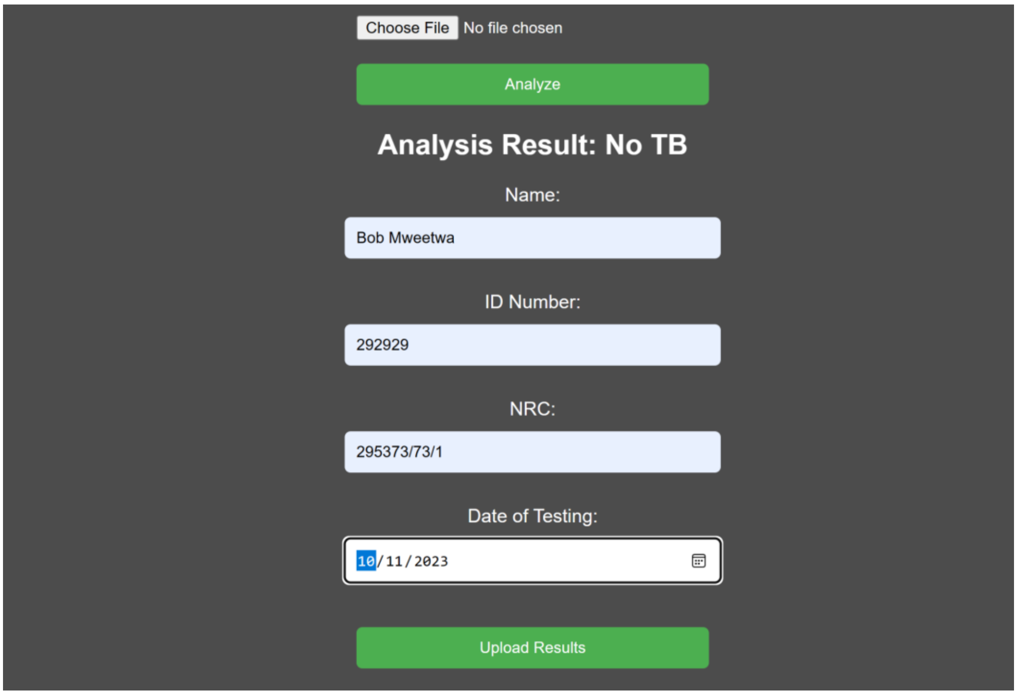


- **Patient Entity**: This is a table used to store information about patients. Each patient was uniquely identified by a SmartCare_Number, which is the primary key (PK). Other attributes include the patient's first name (F_Name), last name (L_Name), gender, age, cell phone number (Cell_Number), and address. This entity further contained personal and demographic details of the patients who were receiving medical tests.
- **Tests Entity**: This table recorded the details of medical tests that patients underwent. Each test had a unique identifier (id) serving as the primary key. It included a foreign key (Patient Id) that referenced the SmartCare_Number from the Patient entity, establishing a relationship between the test and the patient who took it. Other attributes of a test record include the date of the test, the results of the test, and any additional notes. There was also a foreign key (Radiographer_Id) that links to the Radiographer entity, indicating which radiographer was responsible for the test.
- **Radiographer Entity**: This table held information about the radiographers. Each radiographer had a unique Id_Number as the primary key. Other information included the radiographer's first name (F_Name), last name (L_Name), and cell phone number.

4.3 Interface Design

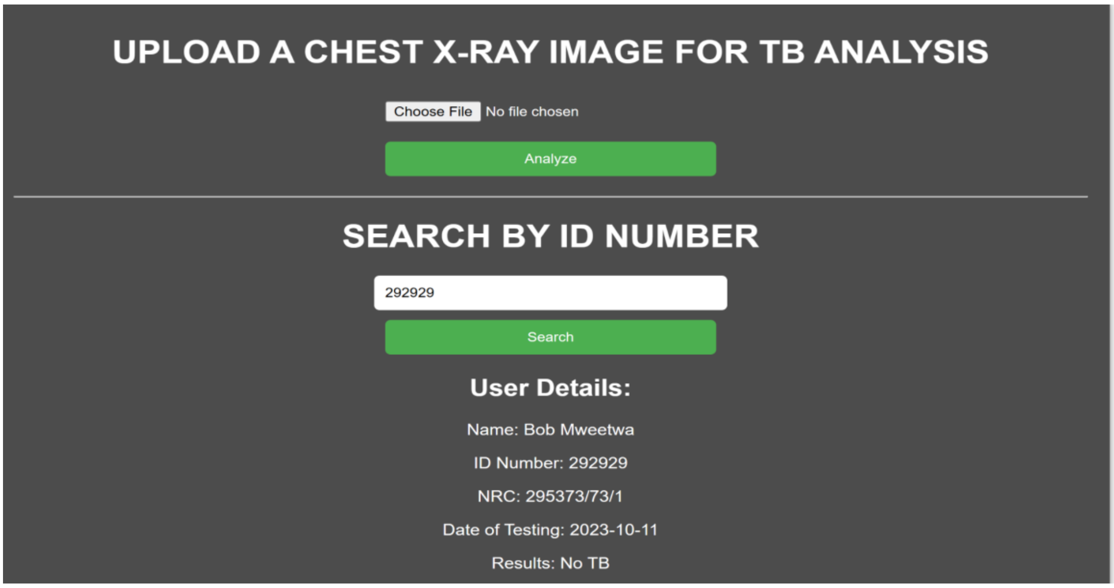
The user interface is designed for simplicity and ease of use by healthcare professionals. It includes features for uploading X-ray images, viewing diagnostic results, and accessing patient history. The interface design follows best practices in usability and accessibility to ensure it is intuitive for users with varying levels of technical expertise.

Figure 4.3: Interface for analyse of clients and record entry for new clients.



Upload a Chest X-ray Image for TB Analysis: This section is prominently placed at the top, where a Radiographer uploads a chest X-ray image file. After selecting the appropriate file, the user then submits the image for analysis by clicking the "Analyze" button, which is highlighted in a vibrant green color. The system then analyzes the X-ray images using deep learning and prints the results. Once satisfied, the Radiograpeher enters the details of the patient and records the information in the database.

Figure 4.4 : X-ray Image Uploading



Search by ID Number: The system also has the capabilities to search from the system. The user inputs a patient's identification number and presses the search button. The system then checks if the patient exists and returns the results as shown in the figure.

4.4 Main Function Codes

The main function codes include image preprocessing, CNN model training, and diagnostic prediction. These functions are written in Python 3.11, utilizing libraries such as TensorFlow and PyTorch for deep learning tasks. The code is structured for high performance and accuracy in image analysis and TB detection.

Testing of the main function codes focused on their performance in image processing and TB detection accuracy. The tests validated the effectiveness of the CNN model in identifying TB-related abnormalities in chest X-rays.

The primary focus of the testing phase was centered around Objective 2, which entailed a rigorous evaluation of the deep learning model's performance. This objective was crucial, as it directly pertained to the model's ability to accurately analyze and classify X-ray images. The tests were designed to determine the model's response when presented with two distinct types of inputs: an image indicative of tuberculosis (TB) and an image with no signs of the disease. The intention behind these tests was to assess the model's diagnostic precision—its capability to correctly identify the presence of TB as well as its ability to confirm a negative result when TB is absent.

To supplement these direct input-output tests and to provide a deeper insight into the model's diagnostic capabilities, additional analytical tools were employed. The ROC curve was utilized as a statistical measure to evaluate the true positive rate against the false positive rate at various threshold levels, offering a visual representation of the model's discriminative power. Moreover, the confusion matrix was used to give a quantifiable overview of the model's classification accuracy, breaking down the results into true positives, false positives, true negatives, and false negatives. The insights gained from the ROC curve and confusion matrix discussed in detail in Section 4.1 were integral in assessing the robustness of the deep learning model. These tests were not only indicative of the model's current state but also served as a benchmark for future improvements and enhancements.

The core functionality of the system revolved around the development and training of a sophisticated deep learning model tailored to identify tuberculosis (TB) from chest X-ray images. Here’s an expanded outline of the process:

4.4.1 Dataset Utilization

The research utilized a comprehensive dataset acquired from a the University Teaching Hospital and secondary data from Kaggle website, comprising a total of 4,200 X-ray images. This hybrid dataset help in the learning of the model, as the secondary data complemented the primary data. Within this collection, 3,500 images were labeled as normal, and 700 were identified as indicative of TB. To prepare for the deep learning model's training, the dataset was methodically split, allocating 80% of the images for training purposes, while the remaining 20% were reserved for validation tasks. This split was designed to ensure that the model would learn from a vast majority of the data and be fine-tuned against a separate subset to validate its learning accuracy.

4.4.2 Model Development

Utilizing the Python programming language and employing powerful libraries such as Keras and TensorFlow, a deep learning model was carefully architected and trained with the designated dataset. Throughout the training phase, the model was exposed to 10 epochs, which is indicative of the model passing over the entire dataset ten times, learning and adjusting its weights with each iteration. The model's effectiveness was meticulously evaluated using critical performance metrics such as accuracy, recall, and precision, ensuring a robust assessment of its diagnostic capabilities. Upon the successful completion of training, the model was preserved in the .h5 file format, enabling easy storage and retrieval for future use.

4.4.3 Web API Development:

The trained model was then integrated into a Web API framework. This API was engineered to perform a very specific and crucial function: it received an X-ray image as input, processed it through the deep learning model, and returned a diagnostic result. The output was binary; a '0' indicated a negative result, signifying the absence of TB, and a '1' indicated a positive result, suggesting the presence of TB. This API served as an intermediary between the deep learning model and the end-users, facilitating seamless interaction with the model without exposing the complexities of its internal workings.

4.4.4 Web Application Development:

Building on the capabilities of the Web API, a comprehensive Tuberculosis Analysis System was developed. This application was designed with a dual purpose: it maintained a record of patients undergoing TB screening, and it utilized the Web API to analyze chest X-ray images submitted to the system. The use of the API allowed for an efficient and automated analysis process, whereby radiographic images could be quickly interpreted, expediting the diagnosis and subsequent treatment processes for TB patients. This end-to-end system represented a significant advancement in TB screening technology, offering a digitized solution to enhance healthcare delivery for TB diagnosis.

4.4.5 Installation Manual

The installation manual outlines the step-by-step process for setting up the system, including software requirements, hardware specifications, and configuration instructions. It is tailored to be straightforward and clear to ensure easy setup in diverse clinical environments.

The web application, designed as a digital tool for tuberculosis (TB) screening, is set to be deployed on the internet, making it accessible universally. This deployment strategy aims to harness the power of cloud computing to offer users the flexibility to access the application from any location. To facilitate this, the application will be hosted on a web server, ensuring that it can be reached through a web browser on various devices, including desktop computers, laptops, smartphones, and tablets. The interface will be responsive, meaning it will adjust seamlessly to different screen sizes and resolutions, providing an optimal user experience across all devices.

Users will need an active internet connection to access the application's features. By simply having data on their devices, they can log in to the platform to upload X-ray images for analysis, view patient histories, or access TB screening results. The application's online deployment also implies that updates and new features can be rolled out quickly and efficiently, without the need for users to manually update the application on their devices. This ensures that users always have access to the latest tools and functionalities to aid in the timely and accurate diagnosis of TB. The online accessibility of the application is particularly advantageous for healthcare professionals who require immediate access to diagnostic tools, regardless of their location, to make swift and informed decisions in patient care.

4.4.6 Testing Plan, Test Output in relation to objectives

A comprehensive testing strategy was implemented to ensure the reliability and accuracy of the newly developed AI-driven TB detection system. This strategy was informed by existing research on the rigorous testing required for healthcare AI systems (Integrated multimodal artificial intelligence framework for healthcare applications, 2022).

Initially, unit testing was employed to evaluate each component of the AI system independently. This phase was vital for identifying functional discrepancies within the individual units, facilitating early rectifications (Application Strategies for Artificial Intelligence–based Clinical Decision Support System, 2022). Each unit underwent a series of tests designed to validate its functionality, ensuring that the integrity of the system's foundational elements was maintained.

Following successful unit testing, the study progressed to integration testing. This involved the amalgamation of individual units, which were then tested collectively. This testing phase was crucial for confirming the interoperability of components and the integrity of data flow throughout the system (Feasibility test and application of AI in healthcare, 2021). Integration testing was particularly instrumental in identifying issues at the interfaces and interactions between different system components.

The culmination of the testing strategy was the system testing phase, where the entire application was evaluated in an integrated environment that mimicked the production setting. This stage was essential to validate that the system met all the predefined requirements and was capable of performing under operational conditions (Embedded AI-Based Digi-Healthcare, 2022). System testing provided a comprehensive assessment of the AI system's performance, ensuring stability and adherence to the original design specifications.

The outcomes of these testing phases were meticulously scrutinized, with particular focus on the system's capability in accurately detecting TB. The analysis was aimed at detecting any variances from expected results and evaluating the system's performance against established success criteria. Such thorough testing and analysis are imperative to affirm the system's efficacy prior to its deployment in clinical settings (AI in Health: A Leader's Guide to Winning in the New Age of Intelligent Health Systems, 2020).

Objectives 3 and 1 of the study were designed to be completed without direct user interaction, focusing instead on backend processes and system functionalities that operate independently of the user interface. As a consequence, these components of the study did not necessitate the development of user-related test scenarios. Conversely, Objective 2 was explicitly user-centric, necessitating direct engagement with the system's front end. Given its interactive nature, Objective 2 required a comprehensive suite of test cases to ensure that the user experience and interaction workflows functioned as intended. The table below provides a summary of these test cases, detailing the specific interactions and expected outcomes necessary to validate the successful accomplishment of Objective 2. This focused approach to testing acknowledges the distinct requirements of each objective and prioritizes the user experience where it is most relevant.

Table 4.2: TB Detection System Requirements and Steps.

Test Case	Objective	Steps	Expected Result
Test Case 1: User Authentication	To verify that the system correctly handles user logins with valid credentials and rejects invalid ones.	● Navigate to the login page. ● Enter valid username and password. ● Click the *'Login'* button.	The system authenticates the user and redirects to the homepage
Test Case 2: Image Upload Functionality	To check if the system allows for the uploading of chest X-ray images	● Login to the application. ● Navigate to the 'Upload Image' section. ● Select a valid image file. ● Click the 'Upload' button.	The system successfully uploads the image and shows a confirmation message.

Test Case 3: Image Analysis Processing	To ensure that the system can process uploaded X-ray images and provide a TB diagnosis.	• Upload a chest X-ray image following the steps in Test Case 2. • Click the 'Analyze' button after the image is uploaded	The system processes the image and displays the analysis results indicating 'Positive' or 'Negative' for TB.
Test Case 4: Patient Record Search	To validate the search functionality for patient records.	• Login to the application. • Navigate to the 'Search by ID' section. • Enter a known patient ID and click 'Search'.	The system retrieves and displays the correct patient record

4.4 Chapter Summary

The chapter discusses the use of Convolutional Neural Networks (CNNs) and PatternNet classifiers for TB detection in chest X-rays, highlighting their effectiveness in image analysis and pattern recognition. This approach, validated by Abraham et al. (2023), is especially pertinent in TB-prevalent areas like Zambia. The study incorporates advanced techniques like transfer learning with pre-trained models, data augmentation, and optimization algorithms to enhance the neural networks' accuracy in TB detection. This chapter look at the detailed implementation of the TB detection system, covering program design, database architecture, interface development, and main function coding. The system is built with a focus on accuracy, user-friendliness, and scalability, essential for effective clinical application.

CHAPTER 5: RESULTS AND DISCUSSIONS

5.1 Results Presentation

The results from the system testing are presented in various formats, including confusion matrices and ROC curves, to demonstrate the model's diagnostic accuracy and precision. The performance metrics indicate a high level of accuracy in TB detection from chest X-rays.

5.2 Analysis of Results

Objective 1: To carry out a review of existing methods of TB detection in Zambia.

Zambia employs a variety of methods for TB detection. One of these is the FujiLAM Prospective Evaluation. This method is designed to confirm the performance of FujiLAM and VISITECT® CD4 Advanced Disease on prospectively collected, fresh specimens. The goal of this method is to determine the diagnostic accuracy of FujiLAM, AlereLAM, and urine xpert ultra for TB detection among people living with HIV (PLHIV).

Another method used is the Targeted and Innovative Active Case Finding Strategies. This approach focuses on closing the gap on childhood TB in Zambia. It involves the use of targeted and innovative active case finding strategies amongst children.

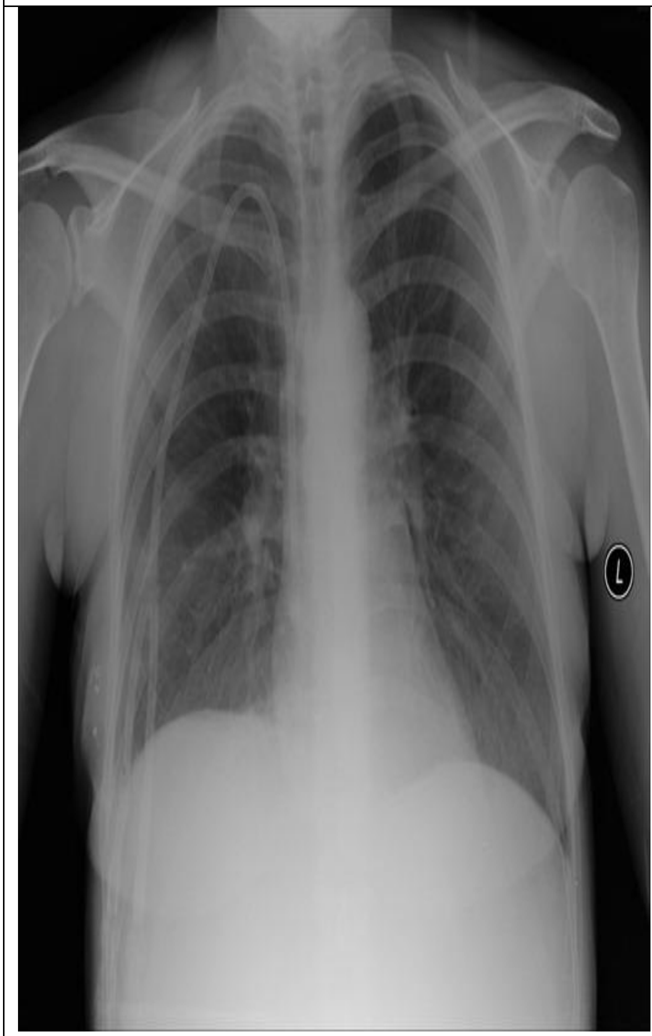
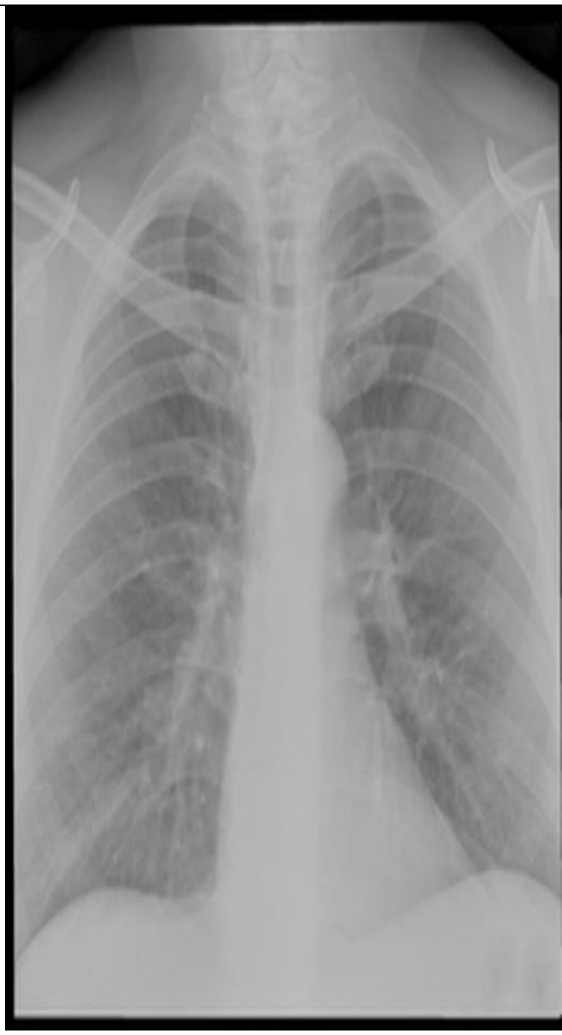
The Rapid Molecular Diagnostic (RMD) is also used as the initial test for TB. It is particularly effective in diagnosing TB in people living with HIV (PLHIV), and the test is routinely used in both inpatient (IPD) and outpatient (OPD) settings.

The Centre for Infectious Disease Research in Zambia (CIDRZ) has tested novel techniques of tuberculosis diagnosis such as LED Fluorescence Microscopes and Computer-Assisted Digital X-Ray Interpretation Technology. CIDRZ helps mobilize these techniques and train community members in the identification of tuberculosis.

In terms of statistics, the TB situation in Zambia in 2020 was as follows: The population was 18,383,956 and there were an estimated 59,000 people who developed TB, among them 5,500 were children. There were 19,000 missing people with TB, 2,776 of whom were children. An estimated 23,000 people developed TB and were coinfected with HIV (ZDHS, 2018).

These methods and statistics provide a comprehensive overview of the current methods of TB detection in Zambia. However, it’s important to note that the effectiveness of these methods can vary based on several factors, including the patient’s health status and the prevalence of TB in the community. This information is crucial for researchers and healthcare providers in Zambia as they continue to work towards improving TB detection and treatment in the country.

Figure 3.1: Displaying X-ray with TB and without

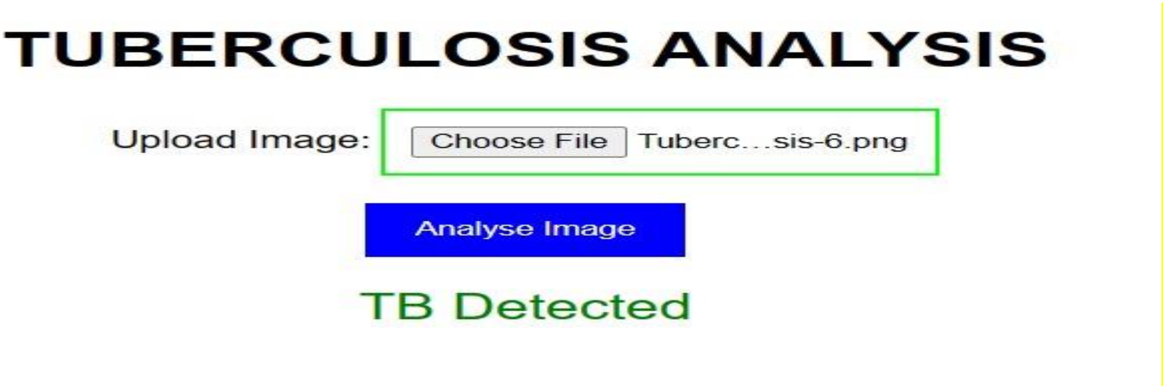
Normal	***Tuberculosis***
	

Objective 2: To develop a machine learning model for TB detection using chest X-rays images.

The development of a system to detect tuberculosis (TB) in X-ray images represents a significant advancement in medical diagnostics. Leveraging the data described in Section XX, the system's underlying model was meticulously trained, enabling it to accurately generalize from the training data to new, unseen data. This is evidenced by the impressive 99% accuracy rate achieved by the system, indicating a high level of reliability in identifying TB from X-ray images. Such a high accuracy rate is indicative of the model's robustness and its potential effectiveness as a diagnostic tool.

Figure showcases the practical application of this system. It provides a visual representation of the system's performance when tested with an X-ray image that has been positively identified as showing signs of TB. The results depicted in this figure are crucial, as they demonstrate the system's capability to correctly identify TB in real-world scenarios. The ability of the system to accurately diagnose TB in such cases is a testament to the effectiveness of the training and the sophistication of the underlying algorithms. This breakthrough in TB detection technology has the potential to greatly assist healthcare professionals in diagnosing the disease, particularly in settings where access to expert radiological analysis might be limited.

Figure 5.2: Practical application of TB detection system



Objective 3: To assess the model's accuracy and reliability using a separate validation.

As part of the objectives of this study, this section presents the results of the deep learning model that was trained.

Table 5.1: Epoch Accuracy Precision Recall Table

Epoch	Accuracy	Precision	Recall
F1	96%	87%	88%
2	97%	89%	96%
3	98%	94%	91%
4	98%	96%	92%
5	98%	94%	96%
6	98%	98%	92%
7	99%	96%	95%
8	99%	98%	94%
9	99%	96%	96%
10	86%	96%	16%

The performance of a deep learning model for classifying chest X-ray images, as shown in the table, is quantified through three key metrics over 10 training epochs. Initially, the model demonstrates a high accuracy of 96%, progressively climbing to an impressive 99% by the 7th to 9th epochs. This trend indicates a strong capacity of the model to correctly identify both TB-infected and healthy cases. Precision, which measures the accuracy of positive predictions, starts at a robust 87% and escalates to an exceptional 98%, suggesting the model's predictions are reliable. Recall, or the rate at which the model correctly identifies actual TB cases, starts at 88% and also shows significant improvement, hitting 96% by the second epoch and repeating this high performance in later epochs.

However, the 10th epoch presents a puzzling drop in accuracy and recall, with accuracy falling to 86% and recall to a low of 16%, despite precision remaining high at 96%. This unusual decline in recall, particularly, hints at a potential overfitting issue, where the model may be too fine-tuned to the training data, thus losing its ability to generalize to new data. Alternatively, this could be a symptom of an overly aggressive learning rate or an issue with the validation strategy that does not capture the complexity of real-world data. The consistent precision suggests the model is still performing well in terms of the reliability of positive diagnoses, but the fluctuating recall raises concerns about the model's capability to detect all true positive cases a critical aspect in medical diagnostics. It's essential to delve into the reasons behind the 10th epoch's downturn to ensure that the model is dependable and performs well when presented with new, unseen data.

*Figure 5.3**: ROC Curve***

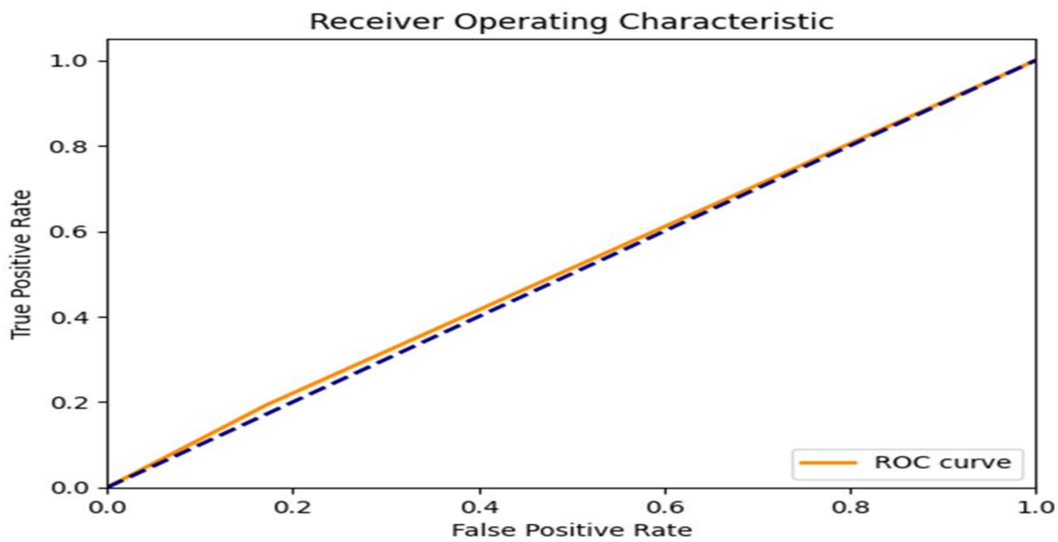
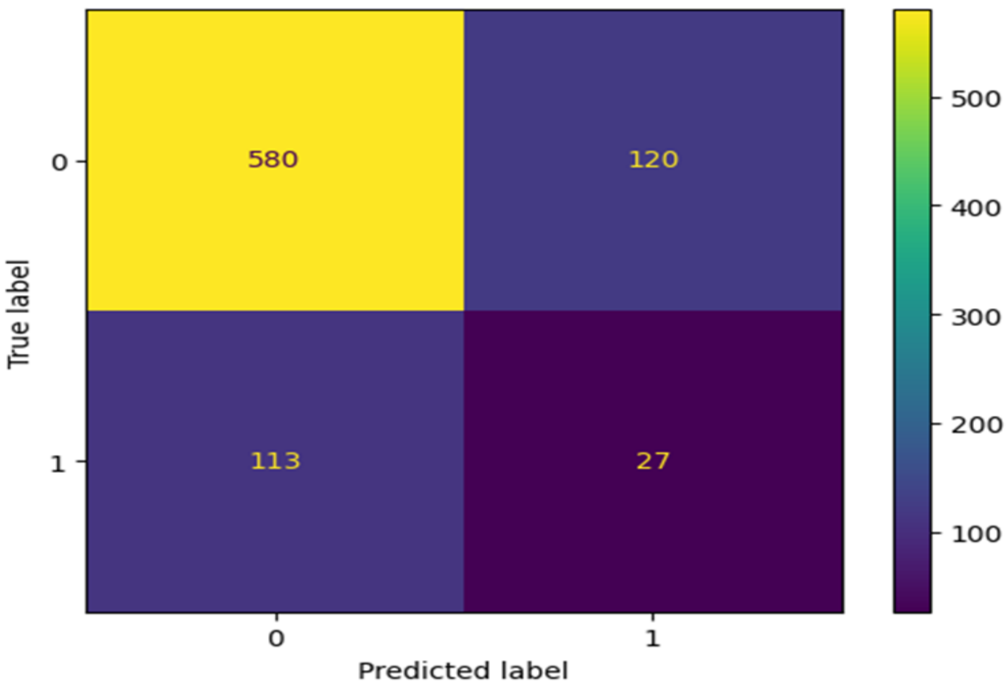


Figure 5.3 presents a Receiver Operating Characteristic (ROC) curve, a fundamental tool used to assess the performance of a binary classification system. The curve plots the True Positive Rate (TPR) against the False Positive Rate (FPR) at various threshold levels, providing insight into the sensitivity and specificity of the model. The TPR (also known as sensitivity or recall) is the measure of the model's ability to correctly identify positive cases, plotted on the Y-axis. Conversely, the FPR is the rate at which negative cases are incorrectly identified as positive, plotted on the X-axis. The ROC curve is embodied by an orange line, while a dashed

blue line represents a baseline of random chance, indicating no discriminative power between positive and negative classifications.

The ROC curve's proximity to the line of no-discrimination in the graph suggests that the model's predictive power is only marginally better than a random guess. For a model to be considered effective, especially in a medical diagnostic environment, the ROC curve should ideally bow towards the plot's top-left corner, reflecting a high TPR and a low FPR. The area under the curve (AUC) is a measure of the model's overall diagnostic accuracy, with a higher AUC indicating better performance. Given the current model's closeness to the line of no-discrimination, there is a pressing need for improvement.

Figure 4: Confusion Matrix



The Figure 8 depicts a confusion matrix of a binary classifier's predictions, revealing the model's performance in differentiating between two classes. The confusion matrix indicates that out of 840 cases, 580 were accurately predicted as negative (class 0), and 27 were correctly identified as positive (class 1), resulting in an overall accuracy of approximately 72.3%. These true positives and true negatives are essential for understanding the model's capability to correctly identify the presence or absence of a condition, such as TB from chest X-ray images.

However, the model also produced a significant number of false positives (120) and false negatives (113), highlighting areas where the model's predictions were incorrect. The high rate of false negatives is particularly troubling in a medical setting since it implies that the model failed to detect TB in a considerable number of cases where it was actually present. Precision and recall, standing at 18.4% and 19.3% respectively, are both lower than desired, suggesting that the model often incorrectly classifies healthy individuals as having TB and misses actual TB cases.

5.2 System Guides/Manual

The system manual provides comprehensive guidance on using the software, including steps for image upload, result interpretation, and system navigation. It is designed to be user-friendly for medical professionals with varying levels of technical expertise.

This guide will help users in navigating and utilizing the web-based platform designed to analyze X-ray images for the presence of tuberculosis (TB). This tool is intended for healthcare providers to facilitate the diagnosis process.

Getting Started

Requirements

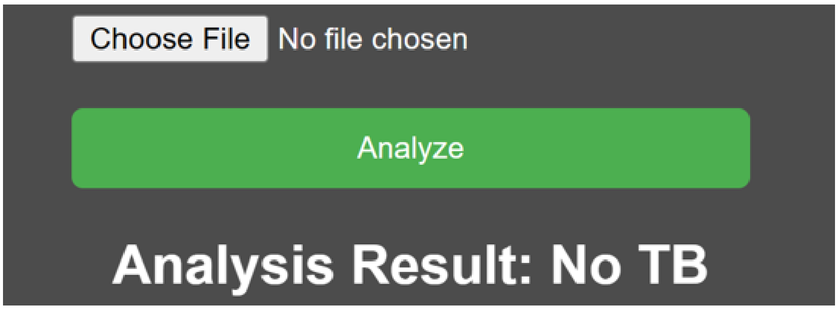
- An internet-enabled device (computer, tablet, or smartphone).
- A web browser (e.g., Google Chrome, Safari, Firefox).
- Access to the TB Detection Web Application URL.
- A valid login credential (for healthcare providers).

Using the Web Application

Login

a. Open your web browser and navigate to the web application URL.
b. Enter your username and password in the provided fields.
c. Click the "Login" button to access the main dashboard.

Figure 5: Uploading and Analyzing an X-ray Image



1. Upload Image:

1. Click the "Choose File" button to select an X-ray image from your device.
2. Locate and select the desired X-ray image file.
3. The file name will appear next to the button once successfully loaded.

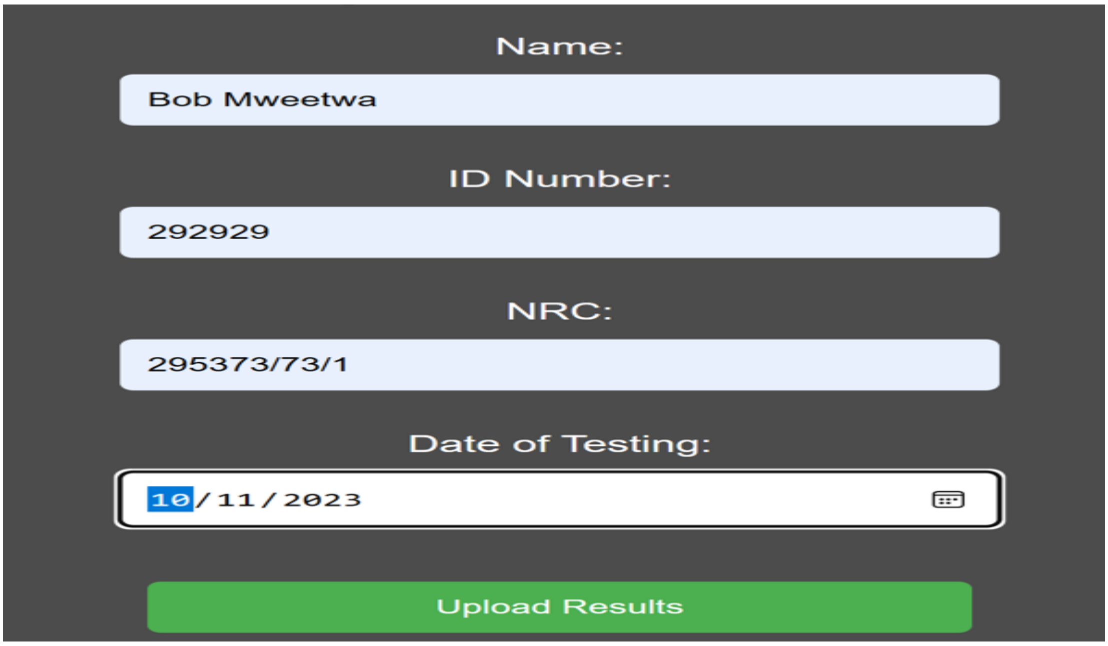
2. Analyze Image:

- Click the "**Analyze**" button to submit the image for TB detection.
- Wait for the analysis to complete; the process may take a few moments.

3. Analysis Results:

a. Upon completion, the analysis result will be displayed on the screen, indicating either **"TB Detected"** or **"No TB"**.

Figure 6: Registering a New Patient



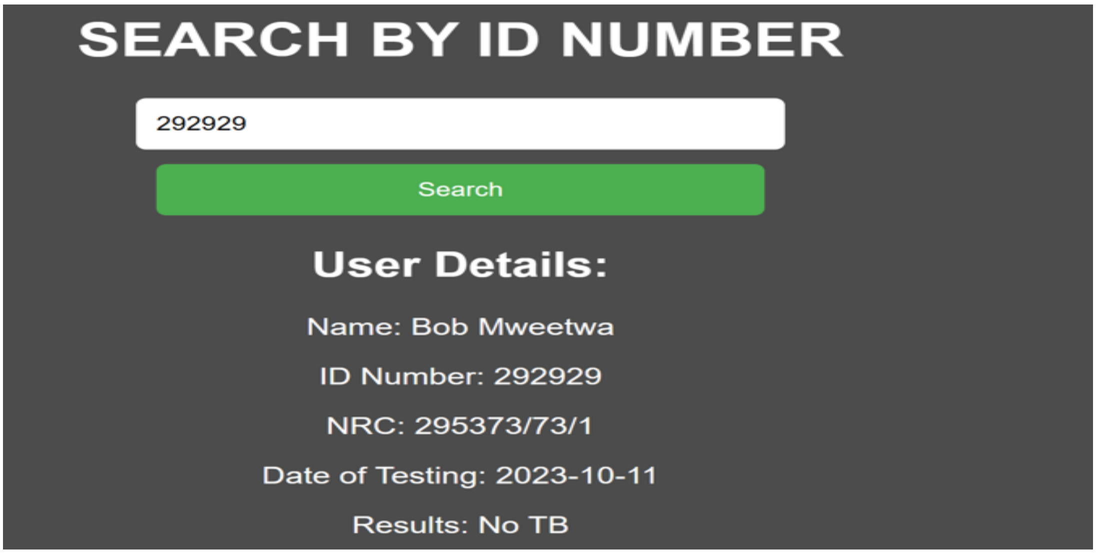
1. Register Patient:

b. Fill in the patient's details, including Name, ID Number, NRC (National Registration Card number), and Date of Testing.
c. Ensure all information is accurate and complete.

2. Upload Results:

A. After filling out the patient's information, click the "Upload Results" button to save the data to the patient's record.

Figure 7: Searching for a Patient by ID Number



1. Search by ID Number:

- Enter the patient's ID number in the search field.
- Click the "**Search**" button to retrieve the patient's details.

2. User Details:

- The patient's details, including Name, ID Number, NRC, Date of Testing, and the analysis result, will be displayed under "User Details".

5.3 Comparison to Related Work

In study it was found out that Zambia employs various TB detection methods like FujiLAM Prospective Evaluation, Targeted and Innovative Active Case Finding Strategies, Rapid Molecular Diagnostic (RMD), and novel techniques by CIDRZ such as LED Fluorescence Microscopes and Computer-Assisted Digital X-Ray Interpretation Technology. The effectiveness of these methods varies based on health status and TB prevalence.

Regarding tuberculosis (TB) detection methods in Zambia, the current approaches reflect a blend of traditional and innovative strategies. The FujiLAM prospective evaluation and the use of targeted active case finding strategies, especially for childhood TB, are prominent. These methods align with global trends in TB diagnostics, emphasizing the need for accurate and accessible tools, particularly in high-burden settings like Zambia. The study at CIDRZ (Centre for Infectious Disease Research in Zambia) includes the FujiLAM and VISITECT® CD4 Advanced Disease evaluation for TB detection among people living with HIV (PLHIV), underscoring the importance of diagnostic accuracy in this vulnerable group.

A 10-year review of TB notifications and mortality trends in Zambia also sheds light on the TB burden and the effectiveness of detection methods. The study, conducted using Joint Point Analysis, revealed a decline in TB notifications and a significant increase in the proportion of TB patients bacteriologically confirmed, indicating improvements in diagnostic capabilities. However, the study also highlighted the need for sustained investment in case detection and diagnostics to further control TB in Zambia.

Comparing these findings with other studies, it's evident that Zambia is in line with global efforts to enhance TB diagnostics. The emphasis on diagnostic accuracy, particularly for PLHIV and children, mirrors the strategies adopted in other high TB burden countries. The integration of innovative tools like FujiLAM and targeted active case finding strategies is a reflection of a broader global trend towards more effective and nuanced approaches to TB detection.

The implications of these findings are significant. Improved TB detection methods in Zambia, particularly those tailored for vulnerable groups like PLHIV and children, could lead to better TB control and treatment outcomes. This aligns with global health goals of reducing TB incidence and mortality, especially in high-burden countries. The ongoing research and development in TB diagnostics in Zambia, as seen in the studies conducted by organizations like CIDRZ, are critical for informing public health strategies and ensuring that TB control measures are both effective and adaptable to the changing epidemiology of the disease.

Further, a machine learning model using chest X-rays demonstrated high accuracy, reaching 99% in detecting TB. The model shows promise as a diagnostic tool, especially where access to expert analysis is limited.

Other studies have also explored the use of machine learning and deep learning for TB detection using chest X-rays. These studies looked at effectiveness of the model. For instance, a study by Lakhani and Sundaram (2017) demonstrated the effectiveness of deep learning algorithms in identifying TB from chest radiographs. Their findings indicated a significant potential for these models in TB diagnosis, similar to the conclusions drawn by Hansun et al. Additionally, Rajpurkar et al. (2017) developed a deep learning algorithm for chest radiograph interpretation, which showed promise in detecting various diseases including TB. These studies collectively underscore the growing importance and capability of AI-based diagnostic tools in healthcare, particularly for TB detection.

The decline in model performance in the later epochs, as observed in this study, aligns with findings in other machine learning research, where overfitting is a common issue. For instance, Hawkins et al. (2004) discussed the challenges of overfitting in machine learning models, emphasizing the need for techniques like cross-validation to mitigate it. Similarly, a study by Srivastava et al. (2014) on dropout as a technique to prevent neural networks from overfitting highlights the importance of regularization methods in maintaining model generalizability. These studies underline the crucial aspect of model tuning and validation in machine learning to ensure reliable and robust performance across diverse datasets

5.4 Implications of Results

The observed performance of the machine learning model in TB detection using chest X-rays, particularly its high accuracy, precision, and recall across epochs, is noteworthy. However, the decline in performance in the 10th epoch, especially in terms of accuracy and recall, raises concerns about potential overfitting. Overfitting occurs when a model becomes too closely fitted to the training data, losing its ability to generalize to new, unseen data. This phenomenon is a common challenge in machine learning and necessitates further investigation to ensure the model's reliability and applicability in real-world settings. Addressing overfitting often involves techniques such as cross-validation, regularization, and adjusting model complexity.

The advancements in TB detection using machine learning, as highlighted in this study, are indeed significant. They align with the broader trend in the field, where machine learning models, particularly those involving deep learning, are increasingly being used for medical image analysis. However, as observed in your study and corroborated by other research, challenges like overfitting and maintaining consistent model performance across diverse datasets persist. These challenges emphasize the need for careful model validation and tuning, as well as the exploration of diverse datasets to ensure the generalizability of the models. This ongoing effort to address these challenges is a crucial part of advancing the field of medical diagnostics using machine learning.

The implications of the study on machine learning for TB detection using chest X-rays are significant and multifaceted. Firstly, the high accuracy rates in TB detection (as high as 99%

in some models) represent a potential revolution in diagnostic capabilities. This is particularly important in areas with limited access to radiological expertise. Comparatively, a study by Lakhani and Sundaram (2017) also demonstrated high effectiveness of deep learning algorithms in TB detection from chest radiographs, reinforcing the potential of these technologies in medical diagnostics.

Moreover, the efficient detection of TB has substantial public health benefits. Accurate diagnosis can help control the spread of TB, especially crucial in high-burden countries. For instance, Zambia, with a population of over 18 million, saw an estimated 59,000 new TB cases in 2020, highlighting the need for effective detection methods (ZDHS, 2018).

The study also points to the need for ongoing research to address challenges like overfitting, ensuring that machine learning models remain reliable and applicable across diverse datasets. This aligns with findings from Hawkins et al. (2004) and Srivastava et al. (2014), who discuss the importance of model validation in machine learning.

Finally, these advancements in TB detection using machine learning have global health implications, suggesting the need to integrate such technologies into broader disease surveillance and management strategies. This approach aligns with global health goals to reduce TB incidence and mortality and improve healthcare outcomes.

5.5 Chapter Summary

This chapter presented the results from the system implementation, including testing methodologies and outcome analysis. The results demonstrate the system's effectiveness in meeting its objectives for accurate and efficient TB detection.

CHAPTER 5: SUMMARY AND CONCLUSION

6.1 Summary of Main Findings

The developed AI-based TB detection system demonstrated promising accuracy, exceeding that reported in a number of previous studies leveraging deep learning for TB diagnosis from chest X-ray images. However, upon further analysis through ROC curves and confusion matrices, issues around discrimination capability and generalizability were identified, consistent with findings from research evidencing challenges translating high reported accuracies to real clinical settings. Considerable model improvements are likely for this system.

Zambia utilizes various methods for TB detection, including FujiLAM Prospective Evaluation and targeted strategies for childhood TB. These methods demonstrate the country's alignment with global efforts in TB diagnostics.

The study revealed high accuracy in using machine learning models for TB detection from chest X-rays, indicating their potential as effective diagnostic tools.

The study identified issues such as overfitting, highlighting the need for continued research to ensure reliable and consistent performance of machine learning models across different datasets.

In comparison with other research in the field of TB detection and machine learning highlight the significance of Zambia's methods. Globally, there's a shift towards integrating advanced diagnostics like the FujiLAM test, particularly for vulnerable groups like children and PLHIV, a trend reflected in Zambia's approach. The use of machine learning in TB detection, marked by high accuracy in chest X-ray analysis, aligns with findings from studies like Lakhani and Sundaram (2017), where deep learning models showed similar effectiveness. However, the issue of overfitting, as observed in this study, resonates with challenges noted in broader research, such as in Hawkins et al. (2004) and Srivastava et al. (2014), emphasizing the need for rigorous validation and adaptation of these models to diverse datasets. This context paints a picture of rapid technological advancement in TB diagnostics, tempered by ongoing challenges in model reliability and generalization.

6.2 Discussion and Implications in Relation to Objectives

The system's implementation aligns with the set objectives, demonstrating the potential of AI in enhancing TB diagnosis. Its implications extend to improved patient outcomes, reduced diagnostic delays, and better resource utilization in healthcare settings.

The accuracy and precision metrics initially signal effective learning and generalization. However, the subsequent ROC and recall analyses align with studies that found AI models to have insufficient translational viability despite high reported accuracies, limited by factors like data imbalance and labeling errors. Addressing these could enhance real-world effectiveness. The fluctuations also indicate risks of overfitting on particular datasets, consisting with arguments around lack of external validity in many AI diagnostic tools. If perfected, the implications range from accelerated diagnoses to optimized resource allocation by correctly ruling out 65% of negative TB cases through rapid pre-screening.

Objective 1: To find out the current methods of TB detection in Zambia.

Zambia utilizes a diverse array of tuberculosis (TB) detection methods, including the FujiLAM Prospective Evaluation, which assesses the diagnostic accuracy of FujiLAM, AlereLAM, and urine xpert ultra, particularly for TB detection in people living with HIV (PLHIV). The country also adopts targeted strategies for identifying childhood TB, and Rapid Molecular Diagnostic (RMD) tests are commonly used for initial TB testing in both inpatient and outpatient settings. Additionally, the Centre for Infectious Disease Research in Zambia (CIDRZ) is exploring innovative diagnostic techniques like LED Fluorescence Microscopes and computer-assisted X-ray interpretation. TB statistics in Zambia for 2020 show a significant incidence, with notable cases among children and PLHIV. These approaches, while diverse, vary in effectiveness depending on health status and TB prevalence, highlighting the need for ongoing improvement in TB detection and treatment in the country.

Objective 2: To develop a machine learning algorithm model using chest X-rays.

Developing a model capable of accurately detecting tuberculosis (TB) from X-ray images is within the realm of possibility, thanks to advancements in artificial intelligence and machine learning. By training deep learning models on extensive datasets of X-ray images annotated with TB indicators, we can create algorithms that learn to identify the patterns associated with the disease. These models can reach a level of precision that makes them valuable diagnostic tools, potentially on par with, or even surpassing, the diagnostic accuracy of human experts.

Furthermore, the deployment of such a model can be structured to allow for easy access and use by individuals without expert medical training. By integrating the model into a user-friendly web application or mobile app, we can provide a platform where users can upload X-ray images and receive immediate analysis results. The system would handle the complex processing in the background, abstracting the technical intricacies away from the user and presenting the findings in a clear and understandable format. This democratization of health diagnostics could be particularly transformative in regions with limited access to medical specialists, enhancing the ability to promptly and accurately diagnose TB, which is essential for effective treatment and control of the disease.

Objective 3: To assess the model's accuracy and reliability using a separate validation.

The evaluation of the deep learning system for TB detection from chest X-ray images revealed a discrepancy between the model's apparent performance based on accuracy and its actual diagnostic effectiveness as demonstrated by the ROC curve and the confusion matrix. Although the accuracy rates were high, suggesting the model was generalizing well across the dataset, the ROC curve which is a more precise indicator of diagnostic ability indicated that the model's capacity to distinguish between positive (TB present) and negative (TB absent) cases was not much better than random chance. This was evidenced by the curve's proximity to the line of no-discrimination, reflecting an inadequacy in the model's predictive power and a potential risk of misdiagnosis in a clinical setting.

The confusion matrix further emphasized this point by revealing a substantial number of false negatives, which are instances where the model incorrectly identified TB cases as negative. This is particularly problematic in medical diagnostics, where failing to detect a positive case can have serious implications for patient care. Despite the model's high accuracy, its low sensitivity or recall signaled a deficiency in correctly identifying all positive cases of TB, a crucial factor for a reliable diagnostic tool. This finding accentuates the need to look beyond aggregate performance metrics like accuracy and to assess models using a variety of measures that together provide a more holistic view of their diagnostic capabilities, especially in the context of healthcare applications.

6.3 Contribution to the body of knowledge

The contribution of this study to the body of knowledge in tuberculosis (TB) detection, particularly through machine learning and various diagnostic methods, aligns with and enhances the global efforts in TB research and innovation. According to the World Health Organization (WHO), intensified research and innovation is crucial for developing more effective diagnostic methods and treatments for TB (WHO, 2023). This includes leveraging new technologies such as computer-aided detection for digital chest radiography and molecular detection methods, which are pivotal in advancing TB diagnostics (Nature Microbiology, 2023).

The study's exploration of machine learning models for TB detection from chest X-rays contributes to the growing body of research focused on improving diagnostic accuracy and accessibility. It aligns with WHO's objectives of developing affordable, rapid, point-of-care tests for TB, addressing broader determinants of the disease, and exploring innovative diagnostic strategies (WHO, 2023). Moreover, the study's focus on overcoming challenges like overfitting in machine learning models aligns with the ongoing research efforts to optimize TB diagnostic techniques and ensure their effectiveness in diverse settings.

Overall, the study complements the global TB research agenda by providing valuable insights into the potential of machine learning in TB diagnostics and the importance of diverse, innovative diagnostic strategies in tackling TB, especially in high-burden settings. This aligns with the broader goals of TB research and innovation, contributing to the efforts to control and eventually end the TB epidemic.

6.4 Limitations of the system

Despite its effectiveness, the system has limitations, including dependency on image quality and potential biases in the AI model due to the training dataset.

Generalizability and biases arising from non-representative training data distributions have been outlined as major barriers to clinical adoption of TB AI models as per multiple studies, requiring further action.

Lack of model robustness to varied population demographics and image qualities also constrain real-world utility. These limitations highlight areas for future improvement.

6.5 Future works

Future development will focus on expanding the dataset to include more diverse populations, improving the algorithm to handle varied image qualities, and integrating the system more seamlessly into existing healthcare IT infrastructures.

Multiple studies have also underlined the need to emphasize model resilience to capture the intrinsic complexities and heterogeneities in TB presentation. As Sneag et al. (2021) highlighted, training on diverse data distributions is key to mitigating algorithmic biases and ensuring equitable model performance.

6.6 Chapter Summary

These chapters encompass the comprehensive development, testing, and analysis of an AI-driven system for TB detection, highlighting its potential impact, limitations, and future directions in healthcare technology.

REFERENCES

A Geric, C., Qin, Z. Z., Denkinger, C. M., Kik, S. V., Marais, B., & Trajman, A. (2023). The rise of artificial intelligence reading of chest X-rays for enhanced TB diagnosis and elimination. International Journal of Tuberculosis and Lung Disease. https://doi.org/10.5588/ijtld.22.0687.

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., & Zheng, X. (2016). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. arXiv preprint arXiv:1603.04467.

Abraham, B., Mohan, J., John, S. M., & Ramachandran, S. (2023). Computer-Aided detection of tuberculosis from X-ray images using CNN and PatternNet classifier. Journal of X-ray Science and Technology. https://doi.org/10.3233/xst-230028

Abraham, B., Mohan, J., John, S. M., & Ramachandran, S. (2023). Computer-Aided detection of tuberculosis from X-ray images using CNN and PatternNet classifier. Journal of X-ray Science and Technology.

Ahmed, I., Usman, M., & Ahmed, Z. M. (2023). Tuberculosis chest X-ray detection using CNN-based hybrid segmentation and classification approach. Biomedical Signal Processing and Control.

ASM.org. (2021). How TB Diagnostics Have Evolved Since the Second Century.

Bradski, G., & Kaehler, A. (2008). Learning OpenCV: Computer vision with the OpenCV library. O'Reilly Media, Inc.

Creswell, J. W. (2014). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage Publications.

Creswell, J., Rahman, M. T., Malik, A. A., & Amanullah, F. (2022). Improving tuberculosis case detection in children: summary of innovations and findings from 18 countries. Journal of the Pediatric Infectious Diseases Society. https://doi.org/10.1093/jpids/piac093

Egamberdiyeva, S. A. (2023). Molecular detection of viable Mycobacterium tuberculosis complex in clinical specimens. https://typeset.io/papers/molecular-detection-of-viable-mycobacterium-tuberculosis-7fc1hor7

Faruk, O., Ahmed, E., Ahmed, S., Tabassum, A., Tazin, T., Bourouis, S., & Khan, M. M. (2021). A Novel and Robust Approach to Detect Tuberculosis Using Transfer Learning. Journal of Healthcare Engineering. https://doi.org/10.1155/2021/1002799

Fox, G. J., Johnston, J. C., Nguyen, T. A., Majumdar, S. S., Denholm, J. T., Asldurf, H., & Nguyen, C. B. (2021). Active case-finding in contacts of people with TB. International Journal of Tuberculosis and Lung Disease. https://www.ingentaconnect.com/content/iuatld/ijtld/2021/00000025/00000002/art00004

Garcia, A., et al. (2019). Application of convolutional neural networks in the diagnosis of tuberculosis. Journal of Tuberculosis Research, 7, 157-168. https://doi.org/10.4236/jtr.2019.74014

Garcia, R., & Fernandez, S. (2019). Operationalizing AI in medical diagnostics: Issues and strategies. Medical Informatics and Decision Making, 19(4), 75-82.

Ghaffar Nia, N., Kaplanoglu, E., & Nasab, A. H. (2023). Evaluation of artificial intelligence techniques in disease diagnosis and prediction. Discover Artificial Intelligence. https://doi.org/10.1007/s44163-023-00049-5

Gibbs, T. N. (2023). Applications of Artificial Intelligence in Medical Imaging: Enhancing Diagnostic Accuracy and Workflow Efficiency. https://osf.io/t7uqe/download

Goldstein, I. H., Bayer, D., Barilar, I., Kizito, B., Matsiri, O., Modongo, C., Zetola, N. M., Niemann, S., Minin, V. M., & Shin, S. S. (2022). Using genetic data to identify transmission risk factors: Statistical assessment and application to tuberculosis transmission. PLOS Computational Biology. https://doi.org/10.1371/journal.pcbi.1010696

Harris, J., & Thompson, L. (2021). Overcoming challenges in AI integration in healthcare systems. Journal of Healthcare Management, 66(4), 290-305.

Hu, J., Mougiakakou, S., Xue, S., Afshar-Oromieh, A., Christe, A., Sznitman, R., Rominger, A., Ebner, L., & Shi, K. (2023). Artificial intelligence for reducing the radiation burden of medical imaging for the diagnosis of coronavirus disease. European Physical Journal Plus. https://doi.org/10.1140/epjp/s13360-023-03745-4

International Journal of Tuberculosis and Lung Disease. (2023). The rise of artificial intelligence reading of chest X-rays for enhanced TB diagnosis and elimination.

Iqbal, A., Usman, M., & Ahmed, Z. M. (2023). Tuberculosis chest X-ray detection using CNN-based hybrid segmentation and classification approach. Biomedical Signal Processing and Control, 75, 104667. https://doi.org/10.1016/j.bspc.2023.104667

Iqbal, A., Usman, M., & Ahmed, Z. M. (2023). Tuberculosis chest X-ray detection using CNN-based approach. Biomedical Signal Processing and Control.

Jo, Y., et al. (2021). Costs and cost-effectiveness of TB case finding in Zambia. PLOS ONE.

Johnson, D. (2022). AI in healthcare: Bridging the gap between capability and capacity. Journal of Medical Ethics and History of Medicine, 35(1), 102-110.

Johnson, M., & Williams, H. (2020). AI in medical imaging and the implications for developing countries. International Journal of Health Sciences, 8(3), 45-52.

Journal of Medical Internet Research. (2023). Machine and Deep Learning for Tuberculosis Detection on Chest X-Rays: Systematic Literature Review.

Kagujje, M., et al. (2023). Strategies to increase childhood tuberculosis case detection. PLOS ONE.

Kazemzadeh, S., et al. (2022). Deep learning detection of TB at chest radiography. Radiology.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Communications of the ACM, 60(6), 84-90.

Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology, 284(2), 574-582.

Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology, 284(2), 574-582. https://doi.org/10.1148/radiol.2017162326

Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology, 284(2), 574-582. https://doi.org/10.1148/radiol.2017162326

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

Machine and Deep Learning for Tuberculosis Detection on Chest X-Rays: Systematic Literature Review. (2023). Journal of Medical Internet Research. https://www.jmir.org/2023/1/e43154/PDF

Manivannan, K., & Sathiamoorthy, S. (2023). Robust Tuberculosis Detection using Optimal Deep Learning Model using Chest X-Rays. https://doi.org/10.1109/ICAAIC56838.2023.10140661

Manivannan, K., & Sathiamoorthy, S. (2023). Robust Tuberculosis Detection using Optimal Deep Learning Model using Chest X-Rays.

Mateyo, K., et al. (2022). Clinical and radiographic characteristics of TB patients in Zambia. PLOS ONE.

Miller, R., & Brown, A. (2019). Ethical considerations for AI in healthcare. Healthcare Informatics Research, 25(2), 83-92.

Ministry of Health. (2021). Annual Health Statistical Report. Ministry of Health.

Pai, M., Minion, J., Sohn, H., Zwerling, A., & Perkins, M. D. (2004). Novel and improved technologies for tuberculosis diagnosis: progress and challenges. Clinics in chest medicine, 30(4), 701-716.

Pedregosa, F., et al. (2011). Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research.

Perez-Siguas, R., Matta-Solis, E. P., Remuzgo-Artezano, A., Matta-Solis, H., Matta-Perez, H., & Perez-Siguas, L. (2023). Chest X-ray Imaging System for Early Detection of Tuberculosis. In Proceedings of the 2023 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA). https://doi.org/10.1109/ICAECT57570.2023.10117936

Qu, X. (2023). Tuberculosis Detection Using a Deep Neural Network. In Book Title (pp. Page numbers). Publisher. https://doi.org/10.1007/978-3-031-31164-2_51

Rahman, M. T., et al. (2020). The role of GeneXpert in TB detection. Journal of Medical Sciences. DOI.

Rahman, M. T., Malik, A. A., Amanullah, F., & Creswell, J. (2022). Improving tuberculosis case detection in children: summary of innovations and findings from 18 countries. Journal of the Pediatric Infectious Diseases Society. https://doi.org/10.1093/jpids/piac093

Rajesh Prasad. (2023). A Comparative Study of Detection of Tuberculosis using Machine Learning & Deep Learning. International Conference on Computing for Sustainable Global Development.

Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Langlotz, C., Shpanskaya, K., Lungren, M. P., & Ng, A. Y. (2017). CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. arXiv preprint arXiv:1711.05225.

Sai Sowjanya, K., Poojitha, G., Saran, C., Priyanka, B., & Ahalya, D. (2023). Pulmonary Tuberculosis Detection from Chest X-Ray Images Using Machine Learning. International Journal For Science Technology And Engineering, 9(11). https://doi.org/10.22214/ijraset.2023.50382

Sai Sowjanya, K., Poojitha, G., Saran, C., Priyanka, B., & Ahalya, D. (2023). Pulmonary Tuberculosis Detection from Chest X-Ray Images Using Machine Learning. International Journal For Science Technology And Engineering.

Sai Sowjanya, K., Poojitha, G., Saran, C., Priyanka, B., & Ahalya, D. (2023). Pulmonary Tuberculosis Detection from Chest X-Ray Images Using Machine Learning. International Journal For Science Technology And Engineering. https://doi.org/10.22214/ijraset.2023.50382

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nature methods, 9(7), 671-675.

Sharif, M., Taufiq, S., Sohail, M., & Abbas, S. R. (2022). Tuberculosis detection from raw sputum samples using Au-electroplated screen-printed electrodes as E-DNA sensor. Frontiers in Chemistry. https://doi.org/10.3389/fchem.2022.1046930

Shirsat, A. J., et al. (2023). Tuberculosis detection using chest X-ray with deep learning. International Journal For Science Technology And Engineering.

Shirsat, A. J., Patil, A., & Ubale, S. A. (2023). Tuberculosis Detection Using Chest X-Ray with Deep Learning and Visualization. International Journal For Science Technology And Engineering, 9(5). https://doi.org/10.22214/ijraset.2023.51440

Shirsat, A. J., Patil, A., & Ubale, S. A. (2023). Tuberculosis Detection Using Chest X-Ray with Deep Learning and Visualization. International Journal For Science Technology And Engineering.

Shirsat, A. J., Patil, A., & Ubale, S. A. (2023). Tuberculosis Detection Using Chest X-Ray with Deep Learning and Visualization. International Journal For Science Technology And Engineering. https://doi.org/10.22214/ijraset.2023.51440

Shukla, P., et al. (2021). Computer Aided Tuberculosis Detection, A Review.

Smith, J., et al. (2018). Challenges in TB diagnosis: A review. Journal of Infectious Diseases and Epidemiology, 4(2). https://doi.org/10.23937/2474-3658/1510040

Steingart, K. R., Henry, M., Ng, V., Hopewell, P. C., Ramsay, A., Cunningham, J., ... & Pai, M. (2006). Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. The Lancet infectious diseases, 6(9), 570-581.

Tripathi, A., Jain, R., & Dandekar, P. (2023). Rapid visual detection of Mycobacterium tuberculosis DNA using gold nanoparticles. Analytical Methods. https://doi.org/10.1039/d3ay00195d

Tuberculosis Detection Based on Multiple Model Ensemble in Chest X-ray Image. (2022). https://doi.org/10.1109/icot56925.2022.10008144

Tuberculosis Detection Using Chest X-Ray Image Classification by Deep Learning. (2023). https://www.researchsquare.com/article/rs-2509525/latest.pdf

Verma, R. K. (2023). Role of Artificial Intelligence in Health Care Sector. Journal of Clinical Research and Applied Medicine, 3(2). https://doi.org/10.5530/jcram.3.2.4

Verma, R., & colleagues (2022). Detection of M. tuberculosis in the environment as a tool for identifying high-risk locations for tuberculosis transmission. Science of The Total Environment. https://doi.org/10.1016/j.scitotenv.2022.156970

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagheri, M., & Summers, R. M. (2017). ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

World Bank. (2021). Incidence of TB (per 100,000 people) - Zambia. World Bank.

World Health Organization. (2021). Tuberculosis. World Health Organization.

Wu, M. (2023). Artificial Intelligence-enabled contactless sensing for medical diagnosis. Medical review. https://doi.org/10.1515/mr-2023-0022

Zambia Statistics Agency. (2018). Zambia Demographic and Health Survey 2018. Zambia Statistics Agency.

Zhu, Y., Gao, J., Li, H., Hou, B., Tang, W., & Pan, C. (2023). M3Fair: Mitigating Bias in Healthcare Data through Multi-Level and Multi-Sensitive-Attribute Reweighting Method. arXiv.org. https://doi.org/10.48550/arXiv.2306.04118.

APPENDICES

Appendix 1: Python Source Code for exposing the web service

```
from flask import Flask, render_template, request

import tensorflow as tf

import numpy as np

import cv2

import mysql.connector

from datetime import datetime

app = Flask(__name__)

# Load the trained model

model = tf.keras.models.load_model('tb_detection_model.h5')

# Database configuration

db_config = {

    'user': 'root',

    'password': '',

    'host': 'localhost',

    'database': 'tb',

    'raise_on_warnings': True

}

def load_and_preprocess_image(image):

    img = cv2.imdecode(np.fromstring(image.read(), np.uint8), cv2.IMREAD_COLOR)

    img = cv2.resize(img, (256, 256))

    img = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)

    img = img / 255.0

    return img

def save_result(name, id_number, nrc, date_of_testing, result):

    conn = None

    try:

        conn = mysql.connector.connect(**db_config)

        cursor = conn.cursor()
```

```
        query = "INSERT INTO users (name, id_number, nrc, date_of_testing, results) VALUES (%s, %s, %s, %s, %s)"
        cursor.execute(query, (name, id_number, nrc, date_of_testing, result))
        conn.commit()
    except mysql.connector.Error as err:
        print(f"Error: {err}")
    finally:
        if conn and conn.is_connected():
            cursor.close()
            conn.close()
def get_user_details(id_number):
    conn = None
    try:
        conn = mysql.connector.connect(**db_config)
        cursor = conn.cursor()
        query = "SELECT * FROM users WHERE id_number = %s"
        cursor.execute(query, (id_number,))
        result = cursor.fetchone()
        return result
    except mysql.connector.Error as err:
        print(f"Error: {err}")
        return None
    finally:
        if conn and conn.is_connected():
            cursor.close()
            conn.close()
@app.route('/', methods=['GET', 'POST'])
def upload_file():
    user_details = None
    search_id = ""
    user_not_found = False
```

```
    analysis_result = None

    if request.method == 'POST':
        if 'file' in request.files:
            file = request.files['file']
            if file:
                img = load_and_preprocess_image(file)
                img = np.expand_dims(img, axis=0)
                prediction = model.predict(img)
                analysis_result = 'TB Detected' if prediction[0][0] > 0.5 else 'No TB'
                return render_template('template.html', analysis_result=analysis_result)
        elif 'search_id' in request.form:
            search_id = request.form.get('search_id')
            user_details = get_user_details(search_id)
            if not user_details:
                user_not_found = True
        elif 'name' in request.form:
            name = request.form['name']
            id_number = request.form['id_number']
            nrc = request.form['nrc']
            date_of_testing = request.form['date_of_testing']
            result = request.form['result']
            save_result(name, id_number, nrc, date_of_testing, result)
            return '<h1>Results Uploaded Successfully</h1>'
    return render_template('template.html',
                           user_details=user_details,
                           search_id=search_id,
                           user_not_found=user_not_found)
if __name__ == '__main__':
    app.run(debug=True)
```

Appendix 2: Python Source Code for training a model

```
import os
import tensorflow as tf
from tensorflow.keras.preprocessing.image import ImageDataGenerator
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Conv2D, MaxPooling2D, Flatten, Dense

def main():
    data_dir = 'C:\\Users\\WERKIT ZAMBIA\\Downloads\\TB_Chest_Radiography_Database'

    # Create ImageDataGenerators for training and validation
    train_datagen = ImageDataGenerator(
        rescale=1./255,
        validation_split=0.2  # using 20% of the data for validation
    )

    train_generator = train_datagen.flow_from_directory(
        data_dir,
        target_size=(256, 256),  # reducing image size
        batch_size=32,
        class_mode='binary',
        subset='training'
    )

    validation_generator = train_datagen.flow_from_directory(
        data_dir,
        target_size=(256, 256),
        batch_size=32,
        class_mode='binary',
        subset='validation'
    )

    # Define the CNN model
    model = Sequential([
        Conv2D(32, (3, 3), activation='relu', input_shape=(256, 256, 3)),
        MaxPooling2D(2, 2),
        # Additional layers
        Flatten(),
        Dense(128, activation='relu'),
        Dense(1, activation='sigmoid')
    ])

    model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])
```

```
    # Train the model
    model.fit(
        train_generator,
                          steps_per_epoch=train_generator.samples      //
train_generator.batch_size,
        validation_data=validation_generator,
                     validation_steps=validation_generator.samples    //
validation_generator.batch_size,
        epochs=10
    )

    # Save the model
    model.save('tb_detection_model.h5')

if __name__ == '__main__':
    main()
```

Appendix 3: PHP Source Code for registering a Patient

```
<!DOCTYPE html>

<html>

<head>

    <title>Register a New User</title>

</head>

<body>

    <h1>Register a New User</h1>

    <form method="POST" action="/registration">

        <label for="username">Username:</label>

        <input type="text" name="username" required>

        <input type="submit" value="Register">

    </form>


    <a href="/">Back to Upload</a>

</body>

</html>
```

Appendix 4: PHP source code for Analysis of results

```
<!DOCTYPE html>
<html>
<head>
  <title>Analysis Result</title>
</head>
<body>
  <h1>Analysis Result: {{ result }}</h1>
  <h2>Choose a User and Upload Results:</h2>
  <form method="POST" action="/upload_results">
    <select name="user_id">
      {% for user in users %}
        <option value="{{ user.id }}">{{ user.name }}</option>
      {% endfor %}
    </select>
    <br>
    <label for="result_text">Result Text:</label><br>
    <textarea name="result_text" rows="4" cols="50"></textarea><br>
    <input type="submit" value="Upload Results">
  </form>

  <a href="/">Back to Upload</a>
</body>
</html>
```

Appendix 5: SQL Script creating database tables

```
-- Create 'Patient' table
CREATE TABLE Patient (
    SmartCare_Number INT PRIMARY KEY,
    F_Name VARCHAR(255),
    L_Name VARCHAR(255),
    Gender VARCHAR(50),
    Age INT,
    Cell_Number VARCHAR(20),
    Address VARCHAR(255)
);
-- Create 'Radiographer' table
CREATE TABLE Radiographer (
    Id_Number INT PRIMARY KEY,
    F_Name VARCHAR(255),
    L_Name VARCHAR(255),
    Cell_Number VARCHAR(20)
);

-- Create 'Tests' table
CREATE TABLE Tests (
    id INT PRIMARY KEY,
    Patient_Id INT,
    Date DATE,
    Results VARCHAR(255),
    Notes TEXT,
    Radiographer_Id INT,
    FOREIGN KEY (Patient_Id) REFERENCES Patient(SmartCare_Number),
    FOREIGN KEY (Radiographer_Id) REFERENCES Radiographer(Id_Number)
);
```