RESEARCH PAPER

Computer-aided Detection of Tuberculosis from Microbiological and Radiographic Images

Abdullahi Umar Ibrahim[1†], Ayse Gunnay Kibarer[1], Fadi Al-Turjman[2]

[1]Department of Biomedical Engineering, Near East University, Nicosia, Mersin 10, Turkey

[2]Department of Artificial Intelligence, Research Centre for AI and IoT, Near East University, Nicosia, Mersin 10, Turkey

Keywords: Tuberculosis; Deep Learning; Pretrained AlexNet; Chest X-ray; Microscopic slide

Citation: Ibrahim, A.U., Kibarer, A.G., Al-Turjman, F.: Computer-aided Detection of Tuberculosis from Microbiological and Radiographic Images. Data Intelligence 5 (2023). doi: 10.1162/dint_a_00198

Received: September 5, 2022; Revised: October 22, 2022; Accepted: January 4, 2023

ABSTRACT

Tuberculosis caused by *Mycobacterium tuberculosis* have been a major challenge for medical and healthcare sectors in many underdeveloped countries with limited diagnosis tools. Tuberculosis can be detected from microscopic slides and chest X-ray but as a result of the high cases of tuberculosis, this method can be tedious for both Microbiologists and Radiologists and can lead to miss-diagnosis. These challenges can be solved by employing Computer-Aided Detection (CAD)via AI-driven models which learn features based on convolution and result in an output with high accuracy. In this paper, we described automated discrimination of X-ray and microscope slide images into tuberculosis and non-tuberculosis cases using pretrained AlexNet Models. The study employed Chest X-ray dataset made available on Kaggle repository and microscopic slide images from both Near East University Hospital and Kaggle repository. For classification of tuberculosis using microscopic slide images, the model achieved 90.56% accuracy, 97.78% sensitivity and 83.33% specificity for 70: 30 splits. For classification of tuberculosis using X-ray images, the model achieved 93.89% accuracy, 96.67% sensitivity and 91.11% specificity for 70:30 splits. Our result is in line with the notion that CNN models can be used for classifying medical images with higher accuracy and precision.

† Corresponding author: Abdullahi Umar Ibrahim (E-mail: Abdullahi.umaribrahim@neu.edu.tr; ORCID: 0000-0003-3850-9921).

1. INTRODUCTION

According to WHO as of 2020, there were over 10 million people estimated to be infected with Tuberculosis (TB) globally. However, it is reported that more than 1.4 million people died from the infection which include more than 200 thousand people suffering with HIV. Globally, TB is among the top 20 causes of death with HIV, cancer, pneumonia, cardiovascular disease etc. TB is rate higher than HIV/AIDS in terms of leading cause of deaths from a single pathogen [1, 2].

The major challenges regarding TB are that children and adolescent are mostly overlooked by healthcare providers which makes it difficult of diagnose and treat. Secondly, there have been thousand cases of multi-drug resistance TB which is difficult to treat using the current drugs and thus, pose serious health burden and concern. The number is estimated to have increase by 10% as of 2018 with total number of 187 thousand people infected with the resistant TB. However, in 2019, it was estimated that over 200 thousand people are infected with multi-drug resistance TB or rifampicin resistant TB [2].

Early screening and detection of the bacteria is critical for preventing widespread and increasing survival rate of patients. Medical laboratory technicians, pathologists, radiologists and microbiologists utilize several diagnostic assays or techniques for accurate detection of the bacteria. Some of the approaches include tissue biopsy assay, true NAT assay, tuberculin test, smear sputum culture, Xpert ultra and Xpert mycobacterium TB assay (which is highly recommended by WHO) chest X-ray (CXR) images and compute tomography (CT) scans [3].

The advent of Artificial intelligence (AI) has transformed computer science and other technological fields, medicine, agriculture, natural science etc. The technology has been utilized for data analysis, image classification, prediction etc. In medicine, AI has been utilized for prediction of disease, identification of drugs, image classification of disease such as cancer, pneumonia, COVID-19, skin lesions, diabetes retinotopy etc. [4].

Due to low sensitivity of existing assays and interpretation of result obtained from this laboratory procedures which is very tedious for medical experts, scientists developed the use of AI-driven models for image-based screening of TB from microscopic slide and radiographic images [5]. Thus, the integration of this technique for automated detection offer promise of lessening the human burden due to the lack of adequate testing materials especially in developed countries and remote areas.

1.1 Computer Aided Diagnosis

Computed Aided Diagnosis (CAD) also referred to Computer Aided Detection (CAD) is a computer-based application that assists medical experts in decision-making [6]. In healthcare system, a medical practitioner uses medical images to evaluate information such as abnormality from images for proper diagnosis. Interpretation of medical images is very critical in the medical field due to the fact that any miss-diagnosis can be detrimental [7]. Different fields of medicine deal with specific types of images such as microscopic slide images by microbiologists, pathological and histological stain slide by pathologists and oncologists, CT scans or Chest Xray (CXR) by radiologists as well as ultrasound and endoscopy [8].

CAD technology revolves around the use of multiple concepts such as medical image processing, computer vision and AI. The primary function of CAD system is detection of abnormality in medical images such as providing quantified image metrics to compute probabilities of different diagnoses and identification of potential Regions of Interests (ROIs) [9]. These techniques have been applied for detection of different grades of tumors (such as colon, prostate, breast and lung cancer) from pathological stain images, detection of *Mycobacterium tuberculosis* from both microscopic slide images and radiographic images, detection of pneumonia from CT scans and diabetic retinopathy. Other application of CAD systems includes diagnosis of Alzheimer's disease, pathological brain detection, coronary artery disease, bone metastases etc. [7].

1.1.1 Artificial Intelligence (AI)

The notions of AI have been trending throughout the last 6 decades. The definition of the concept varies among scholars. However, AI is termed as any technique that enables computers to mimic human behaviour. The term AI was coined in the 1950s and subsequently Machine Learning (ML) as a subfield of AI was introduced. ML is coined in the 1980s which is categorized in supervised ML (SML), unsupervised ML and Reinforcement ML [10].

1.1.2 Machine Learning (ML)

ML is coined in 1959 by Arthur Samuel as "*a field of AI that gives computers the ability to learn without being explicitly programmed*". ML algorithms are grouped into 3 major categories, Supervised ML, Unsupervised ML and Reinforcements ML. Supervised ML algorithms are the most common and widely used ML techniques adopted by medical practitioners in which data are labelled and the network learn features to recognise patterns in data for prediction or classification. The most popular SML techniques include Neural Networks (NNs), Support Vector Machine (SVM), Random Forest (RF), decision Tree (DT), etc. [10, 11].

In Unsupervised ML, models learn from unlabelled data. These models learn to forecast result from the data according to the patterns learned. The most widely used unsupervised ML models include Rule Mining (RM) and Clustering algorithms. Meanwhile, Reinforcement ML is termed as "when a computer program learns from Experience (E) with respect to some Task (T) based on the Performance (P) measures, if P at T is measured by P improve with experience [10, 11].

1.1.3 Deep learning and Artificial Neural Networks (ANNs)

Deep learning (DL) is a subfield of ML which is inspired by how human brain's function due to connections or synapses of nerve cells or neurons. Model learn as a result of data connection between neurons in the network. A simple neural network is termed as perceptron which take input as data set and produced an output as classification or prediction outcome. DL neural networks are made of multiple perceptron's with an input layer (IL), and many hidden layers (HL) before output layer (OL) [12, 13]. Since the emergence of DL in 2010, scientists have designed different models using convolutional neural networks that can classify and analyse medical images such as cancer, tuberculosis, radiological images for diagnosis of diseases [14].

Convolutional Neural Network (CNN) is a class of ANN with multi-layer perceptron which are fully connected network in which each neuron from one layer is connected to all neurons in the next layer. CNNs are termed as networks that utilize series of mathematical operations knows as "Convolution". There are various neural networks architectures developed. Some of the architectures have performed better than others in terms of regression, classification and denoising images. The current best models include AlexNet with 8 layers, VGGNet with 19 and 16 layers, Inception module also known as GoogleNet with 22 layers and 9 modules and Residual or ResNet with 152 layers. In order to train a NNs, a backpropagation algorithm is used to adjust the weight according to the data pattern and optimize the error between predicted output and actual output [15, 16].

1.1.4 AlexNet

AlexNet is the first CNN developed that outperform other models in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competition in 2012. AlexNet is developed by Alex Krizhevsky. The model consists of overall 8 layers, in which 5 are convolutional layers and 3 fully connected layers. The first two convolutional layers are made of 3 operations which include convolution, pooling and normalization. AlexNet use Rectified Linear Unit (ReLu) as an activation function unlike Tanh and sigmoid functions that are used in traditional ML. ReLU converts negative numbers to zeros and help models learn non-linear functions [17, 18].

Max pooling is the most common pooling methods which main function is to down sample or to reduce image size by pooling most important feature or by pooling out the number with highest pixel value. The next 2 layers are mainly convolution layers without pooling and normalization and the final convolution layer consists of only convolution and pooling without normalization. The first 2 fully connected layers are dropout layers which main function is to reduce overfitting through reduction of number of neurons. The final FCL is basically for classification as shown in figure 1 [19, 20].

1.1.5 Transfer Learning (TL)

TL is defined as ML approach where a model trained on a specific task is re-purposed on other related task or a means to extract knowledge from a source setting and apply it to a different target setting. TL can also be described as a process where what models learned from a specific task or setting is harnessed to improve better outcome in another task or setting. Some of the advantages of TL or pretrained models over the ones developed from scratch include low computation, better performance, small learning rates, time saving, etc. [21].

1.2 Tuberculosis (TB)

TB is an airborne infection caused by a bacterium known as *Mycobacterium tuberculosis* (MBTB) which are slender, rod shape microbes with length ranging from 1–10μm and strict aerobes (need oxygen to survive). They possess a waxy cell wall as a result of formation of "Mycolic Acid" making them "Acid

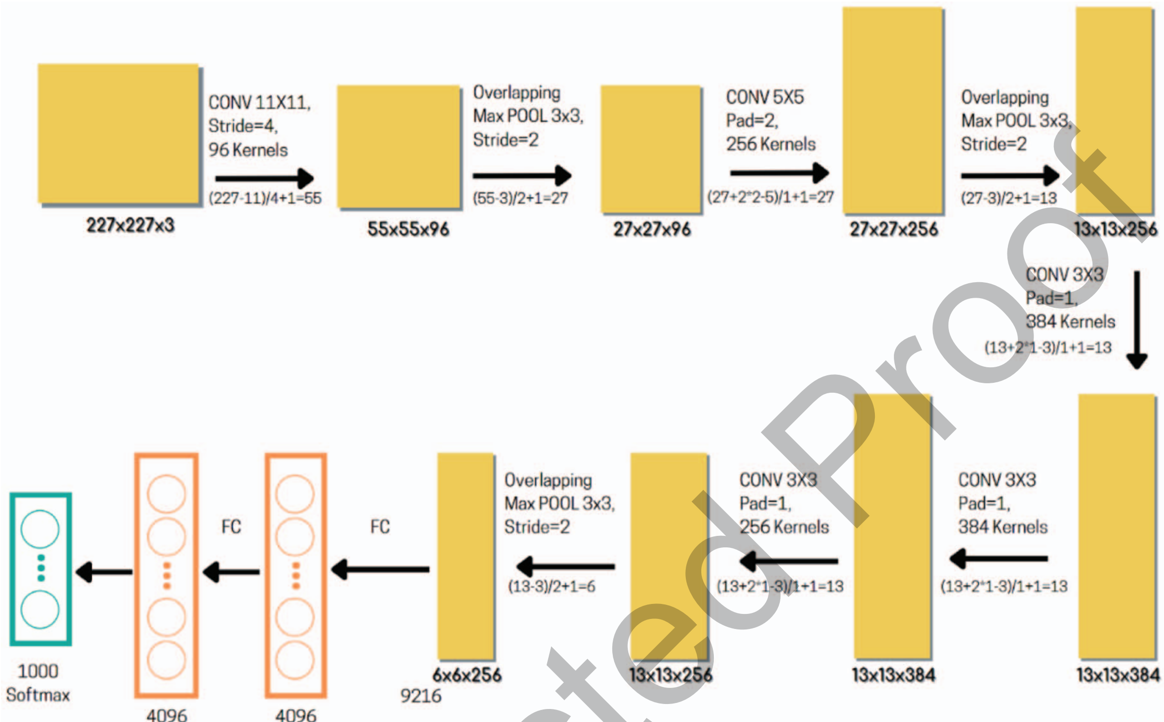

Figure 1. AlexNet architecture.

fast" which signifies that they can retain on to a dye or stain in spite of being exposed to alcohol, thus, making the bacilli look red in color when Ziehl-Nelson stain is applied. However, due to the nature of their waxy cell wall, they tend to repel weak disinfectants and can survive on a dry surface for a long period of time [22]. According to WHO 2019 report, deaths as a result of TB-related disease decreased from 1.6 million in 2017 to 1.5 million in 2018. It was estimated that 10 million people fell ill as a result of TB in 2018, with the majority of patients coming from India, Pakistan and China. The symptoms of tuberculosis include haemolysis (coughing up blood), fever, weight loss, night sweats etc. [1].

1.2.1 Diagnosis

There are many approaches adopted by pathologist for the detection of TB; some of the techniques include Tuberculin Skin Test (TST), microscopy, chest X-ray, Purified Protein Derivative (PPD), GeneXpert, culture test and Interferon γ-release assay. However, among these techniques, microscopic sputum smear evaluation using microscope remains the most common approach globally, especially in underdeveloped and developing countries due to its affordability, simplicity, speed and maintenance when compared to other techniques [3].

1.3 Challenges

Majority of microbiological diagnostic depends on chemical and analytical assays and interpretation by professionals in the field. Thus, due to the tedious nature of the assays, lack of adequate and reliable testing kit and reagents, or expiration of the existing chemicals especially in underdeveloped countries, this result in miss-diagnosis, irreproducibility, false-positive results and inaccuracy. Despite the fact that the use of microscope for identification of microbial pathogens is very crucial for screening and diagnosis, this method is hindered by several challenges which can lead to miss-diagnosis and false-positive cases [23].

Some of the challenges of microscopic examination of *Mycobacterium tuberculosis* include the overlapping of the bacteria on top of each other, the size of the bacteria which is very small (i.e., 1–10 μm), low background contrast, irregular appearance, heterogenous shape, faint boundaries. Moreover, screening large number of microscopic slides images can be very hectic and tedious for pathologists. In order to resolve these challenges and limitations, there is need to develop more rigorous, robust, accurate, fast, reliable technique that can be utilized for screening and diagnosis of TB [3, 23].

The aim of this work is to applied DL-TL for the detection of TB. However, there are several studies that addressed this issue. One of the distinctions of our study with existing literature is that we trained and validated AlexNet model using microscopic slide images and X-ray while existing studies only focus on one type of medical image. Secondly, majority of studies evaluated model based on conventional performance metrics, however, in this study, we conducted 10k cross validation in order to provides average performance of the model rather than using single evaluation.

1.4 Contribution

- The use of pretrained AlexNet for classification of TB from microscopic slide and X-ray images.
- The use of pretrained AlexNet for classification of TB from microscopic slide images.
- Performance evaluation based on Accuracy, sensitivity and specificity.
- Evaluation of model using 10k-folds cross validation
- Realistic comparison between pretrained AlexNet with the current state-of-the-art.

2. RELATED WORK

2.1 Microscopic Slide

Automated detection of *Mycobacterium tuberculosis* has aided in accurate diagnosis of the disease. The integration of computer-based technology into healthcare has transformed the sector and contribute to increase in efficiency. Several studies in the literature have demonstrated the benefit of integrating CAD systems. Smith et al. [24] utilized a high-quality microscope designed to collect high resolution stain slide images. In order to make the bacilli visible, the researchers amplified the number of colonies and stained using dye to acquired 25000 images. The study conducted different types of image augmentation which are used to train the CNN model in order to discriminate between various types of bacteria (rod, chain and round-shapes). The model achieved an overall accuracy of 95%. The classification of *Mycobacterium tuberculosis* and normal cases using ANN is provided by Khan et al. [25]. The study utilized over 12 thousand images which are partitioned into 70% for training of the model and 30% for validation. The image training was carried out using feedforward backpropagation model and the model achieved 94% testing accuracy. A CNN Model built from scratch by Xiong et al. [5] name TB-AI was used for detection of *Mycobacterium tuberculosis* bacillus. The model was trained using 45 total samples with 15 as negative cases and 30 as positive cases which are tissue samples that were treated using acid-fast stain. The result has shown TB-AI achieved 83.65% specificity and 97.94% sensitivity.

Panicker et al. [26] utilized CNN approach to detect *Mycobacterium tuberculosis* bacillus from microscopic sputum spear images. The dataset was obtained from a public domain with 120 images which were cropped to 900 patches for both positive and negative samples for segmentation method. The model accomplished sensitivity of 97.13% and specificity of 78.4%. A study based on the use of SVM for the detection of *Mycobacterium tuberculosis* is provided by Costa Filho et al. [27]. The study employed 120 smear microscopic slide image from 12 cases. Prior to training, the images undergo conventional smear microscopy and segmentation. The study reported an error rate of 3.38% and sensitivity of 96.8%. The use of microscopic slide images for the classification of TB and non-Tb cases is proposed by Ibrahim et al. [23]. The study acquired microscopic slide images from Near East University (NEU) Hospital which contain 530 images. Data augmentation based on rotation and cropping were also conducted to increase the training set to 2444 images. The images were trained and tested using pretrained AlexNet model which resulted into 98.73% accuracy, 98.59% sensitivity and 98.84% specificity. Machine vs Human experiments were conducted where DL model outperformed human pathologists.

A deeper model was employed by El-Melegy et al. [28] for detection of Ziehl Nelson-stained sputum smear of TB and healthy images. The research utilized 500 images which are divided into 80% and 20% for training and validation respectively and train using Faster Region-based convolutional neural network plus CNN (F-R-CNN+CNN) and Region-based convolutional neural network F-R-CNN. F-R-CNN+CNN achieved 85.1% sensitivity and 98.4% accuracy while F-R-CNN achieved 82.6% sensitivity and 98.3% accuracy. Due to the prevalence of TB in Uganda, Muyama et al. [29] utilized 3 TL models based on ResNet (inception V3), GoogleNet and VGGNet for computer assisted-detection of TB from Ziehl-Nelson sputum smear slide images. The study made used of dataset obtained from an online database and the ones captured using cell phone's camera in the university microbiology laboratory. The 2 datasets are combined together and partitioned into 80% for training and 20% for validation. The models were trained according to (I) find-tunning (II) with augmentation and (III) without augmentation. However, among all the pre-trained models, ResNet achieved the highest accuracy score of 86.7%. The summary for detection of tuberculosis from microscopic slide using AI-driven tools is presented in Table 1.

Table 1. Detection of tuberculosis from microscopic slide using AI-driven tools.

Reference	Neural Network	Sample type	Dataset	Results
[25]	CNN	Microscopic Stained image	12,636	94% accuracy
[27]	SVM & CNN	Microscopic stained images	120	96.80% accuracy
[24]	CNN	Microscopic stained image	25,000 images	95% accuracy
[28]	CNN	Microscopic stained image	500 images	F-R-CNN achieved 98.3% accuracy and 82.6% sensitivity F-R-CNN+CNN achieved 98.4% accuracy and 85.1% sensitivity
[3]	CNN	Microscopic stained image	Cases: 45	97% sensitive and 83.65% specific
[29]	ResNet, GoogleNet and VGGNet	Ziehl-Nelson-stained smear sputum slides	Not specified	86.7% accuracy
[23]	Pretrained AlexNet	Ziehl-Nelson-stained smear sputum slides	2464 images	98.73% accuracy, 98.59% sensitivity and 98.84% specificity

2.2 Chest X-ray

Klassen et al. [30] employed CNN to discriminate between normal and pulmonary TB using radiographs. The study utilized 1007 posterior chest radiographs which were partition into training, validation and testing respectively. The images were trained using 2 models which include GoogleNet and AlexNet. The two models ensembled together to achieved AUC of 0.99 with a sensitivity of 97.3% and specificity of 94.7%. The same approach was adopted by Yahiaoui et al. [31] to classify TB from healthy samples using SVM. The model was trained using 15 total CXR images (acquired from 50 patients suffering from TB and 100 healthy samples). However, the model achieved 96.7% accuracy.

The use of TL based on VGGNet and SVM as the model classifier is reported by Ahsan et al. [32]. The study utilized dataset obtained from Shenzhen hospital, China and from Montgomery County Tuberculosis Control Program (MCTCP) in the form of CXR images. The images were trained based on (I) with augmentation and (II) without augmentation. The model achieved 81.25% validation accuracy with augmentation and 80% validation accuracy without augmentation. Chang et al. [33] proposed a 2-stage classification of TB based on TL on CNN. The study made used of 1727 cases of TB culture acquired from Tao-Yuan general hospital, Taiwan. The dataset was trained using VGGNet, YOLO and CNN designed from scratch. By targeting the result of the model on non-negative class, the proposed system achieved 98% recall and 99% precision.

To classify normal and abnormal X-ray images of individuals suffering from TB, Abbas and Abdelsamea [34] utilized a pretrained AlexNet model. The model was trained based on 138 total CXR images (58 normal and 80 abnormal X-ray images). To increase number of training set, the X-ray images undergo data augmentation. The model hyperparameters were turned according to deep-tuning, shallow-tuning and fine-tuning techniques. The study revealed that hyperparametric fine-tuning of pretrained AlexNet outperform other tuning techniques with 0.998 AUC score, 99.7% sensitivity and 99.9% specificity. The application of hybrid model for automated detection of TB is proposed by Sahool et al. [35]. The hybrid model comprises of MobileNet with 88 layers and feature selector in the form of Artificial Ecosystem-based optimization (AEO) algorithm. The model was trained on X-ray dataset acquired from Shenzhen hospital, China with 662 totals frontal CXR images (of which 336 are positive and 336 negative). The model achieved 90.2% best classification accuracy, 93.85% sensitivity and 86.76% specificity.

The study conducted by Ibrahim et al. [36] applied 2 classifiers for the classification of TB and non-TB images. 2 Pretrained AlexNet models (AlexNet+Softmax and AlexNet+SVM) are trained and tested using 3871 TB images and 3500 non-TB images. The evaluation of the model performance yielded 98.19% accuracy using AlexNet+Softmax and 98.38% using AlexNet+SVM. The study conducted by Nafisah and Muhammad [37] applied DL models for the classification of CXR images into TB and non-TB cases. Several DL models which include (EfficientNetB3, MobileNet, ResNeXt-50, Inception-ResNet-V2 and Xception) were trained and tested using 692 TB and 406 non-TB segmented images. The evaluation of the model performance has shown that EfficientNetB3 achieved the highest result with 99.1% accuracy, 98.7% average accuracy and 99.9% ROC. The summary for detection of tuberculosis from Chest X-ray using AI-driven tools is presented in table 2.

Table 2. Detection of tuberculosis from chest X-ray using AI-driven tools.

Reference	Neural Network	Sample type	Dataset	Results
[32]	Automated recognition and pattern analysis	Flougraphic chest mages	Negative = 238 Positive = 70	sensitivity 75.0–87.2%, specificity 53.5–60.0%,
[31]	SVM	Chest X-ray images	150 cases 50 positive and 100 negatives	96.68% accuracy
[33]	Pretrained AlexNet model	Chest X-ray images	138 (58 negative and 80 positive cases)	0.998 AUC score, 99.7% sensitivity and 99.9% specificity
[34]	Hybrid model (MobileNet and Artificial Ecosystem-based optimization algorithm	Frontal chest X-ray	662 (336 positive and 336 negative)	90.2% best classification accuracy, 93.85% sensitivity and 86.76% specificity.
[36]	AlexNet+Softmax and AlexNet+SVM	Chest X-ray images	3871 TB images and 3500 non-TB images.	98.19% accuracy using AlexNet+Softmax and 98.38% using AlexNet+SVM.
[37]	EfficientNetB3, MobileNet, ResNeXt-50, Inception-ResNet-V2 and Xception)	Chest X-ray images	692 TB and 406 non-TB segmented images	EfficientNetB3 with 99.1% accuracy, 98.7% average accuracy and 99.9% ROC

The field of CAD is changing the landscape of disease diagnosis through improving diagnostic efficiency, reducing the use of toxic chemicals and workload. Several researches in the existing literature have demonstrated how AI-driven models aid in diagnosis of several diseases ranging from cancer (breast, melanoma, colorectal, prostate etc.), TB, pneumonia (COVID-19 and non-COVID-19). Screening of TB is crucial for early diagnosis and timely treatment especially in underdeveloped countries with prevalence of the disease. Majority of the studies in the literature either focus solely on the application of AI-driven models on microscopic slide images or X-ray images. However, this study includes both medical images and provide comparative analysis of the result obtained and their distinctions. Consequently, majority of exiting studies evaluated models based on conventional metrics (which include accuracy, sensitivity, specificity, AUC, ROC etc.) of training and testing sets. While in this study we include the use of cross validation in order to provides average performance of the model rather than using single evaluation.

3. METHODOLOGY

This chapter discuss about the overall methodology which include data collection, image pre-processing, data labelling, data split, cross validation pretrained model adjustment, data training and mode of evaluation.

3.1 Overview

The summary of the overall methodology is illustrated in Figure 2. Dataset were obtained from 2 different sources (1) X-ray images from Kaggle repository and (2) microscopic slide from Near East University Hospital and Kaggle repository. In order to fit the images into AlexNet model, data processing step was conducted by reducing the image the required pixel size and fed into the pretrained AlexNet model. The model was subsequently evaluated on the basis of sensitivity, specificity and Accuracy.

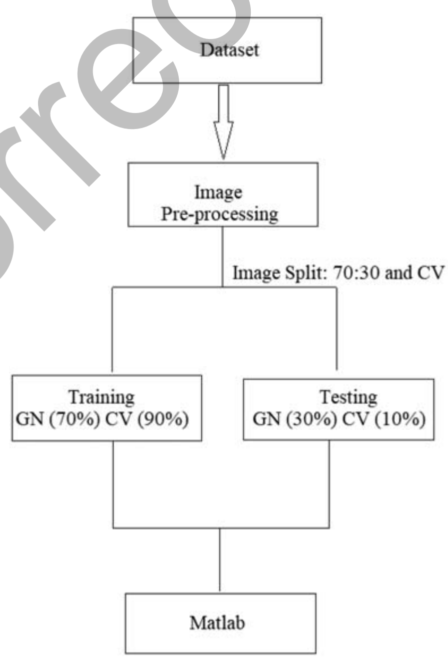

Figure 2. Flow Chart (GN: General Dataset and CV: Cross Validation).

3.2 Data Collection

3.2.1 Microscopic Slide Images

Microscopic slide images have been used extensively in clinical diagnosis of pathogenic disease such as TB, Pneumonia, Anthrax, Cholera, Meningitis, Gonorrhoea etc. Microscopic slides imaging is the most popular conventional approach for the detection of TB. The slides use in this study were prepared through smearing of sputum collected from suspected patients. The basic protocol includes smearing the sputum on clean slides follow by flaming using low flame in order to fix the air-dried smear. Nexus is the spray of Auramine O stain for 10–15 minutes. The slide is then wash using alcohol and subsequently rinsing using deionized water. The final step is based on addition of potassium permanganate to counter stain the slide which result into a clear contrast background which can be view using electronic microscope. The images obtained from the microscope can be save in a computer for further analysis. 2 sample of microscopic slides employed in this study are shown in Figure 3. The overall microscopic slide images are 600 (300 positive and 300 negative).

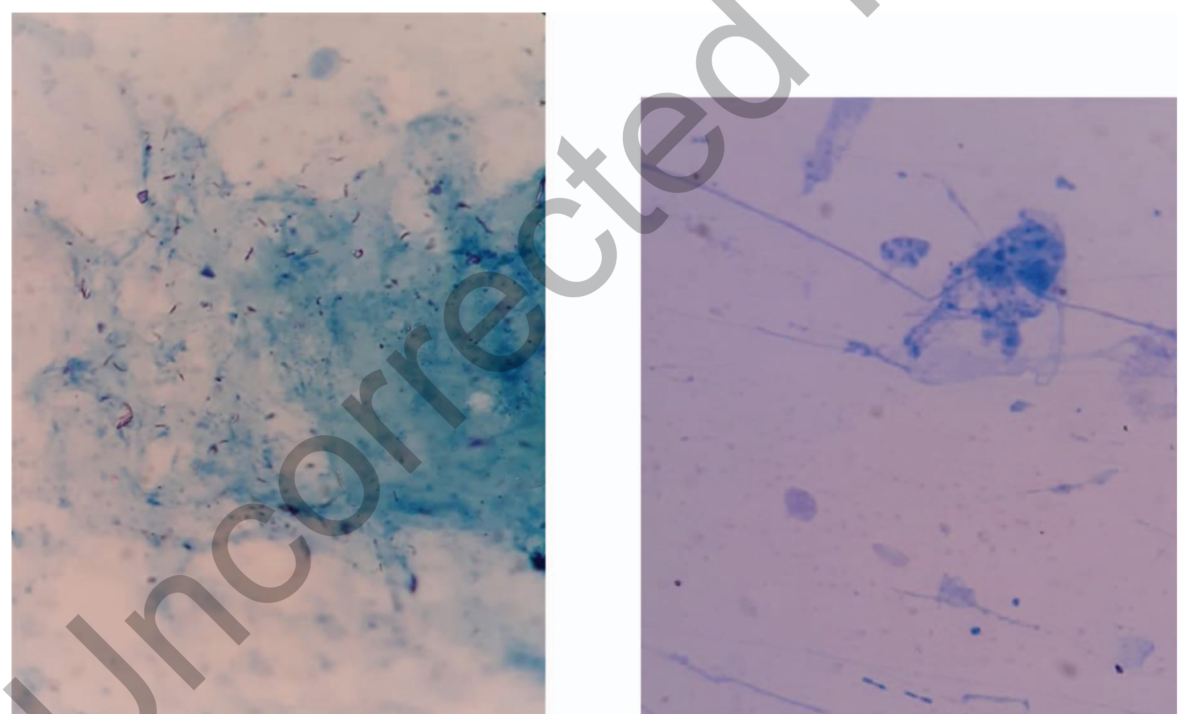

Figure 3. Microscopic Slide Images. Right: Positive (tuberculosis) microscopic slide image. The purple and red thick bacilli depict mycobacterium tuberculosis. Left: Negative.

3.2.2 Chest X-ray

The use of radiographic images provides medical expert with a clear view of patient's organs which aid in diagnosis and clinical decision. Chest X-ray images are used to diagnose disease associated with the lungs such as pneumonia (bacterial, viral which include COVID-19) effusion and TB as shown in Figure 4. The X-ray images used in this study are acquired from publicly available website known as Kaggle. The total X-ray images used in this study is 600 (300 positive and 300 negative or normal images).

Figure 4. Chest X-ray Images. Right: Positive. Left: Negative.

3.3 Data Processing

The microscopic slides images employed for this research were first evaluated and labelled by clinical pathologists into positive and negative cases. The X-ray images used are already labelled in the website. Both the microscopic slide images and chest X-ray images are too large ranging from 1MB to 6MB per each single image. In order to fit the images into AlexNet model, the images are reduced using an online free website known as Birme (https://www.birme.net/). The images are resized into 227x227x3. 227 denoted the pixel size for both vertical and horizontal while 3 denoted channels which include Red, Green and Blue (RGB).

3.4 Data Split

After resizing, the images are split into 70:30 for training and testing respectively. Based on the 600 images obtained, 420 (210 positive and 210 negative are used for training (I.e., 70%) and 180 (90 positive and 90 negative) for testing (I.e., 30%) as shown in Table 3.

Table 3. Data split of Microscopic Slides Images and Chest X-ray Images.

Datasets	Training (70%)		Testing (30%)		Total
	Positive	Negative	Positive	Negative	
Microscopic Slide	210	210	90	90	600
Chest X-ray	210	210	90	90	600

3.5 Cross Validation (CV)

Evaluation of model performance is crucial for determining overall accuracy and generalization. Evaluation of model performance based on conventional ML approach involve training computer models using training set and evaluating their performance using testing set (i.e., holdout set). However, this method is not reliable and accurate for finalizing model performances. In order to resolve this issue, scientists proposed the use of CV. This technique is use for assessing generalization of independent data using statistical analysis [38, 39]. Moreover, this technique is useful for evaluating ML models based on training models with subsets of the input dataset (i.e., 90%) and evaluating the performance using complementary subset (i.e., 10%). CV differs with conventional ML performance approach as it allow the use of each subset as part of the training set and testing set (i.e., reshuffling of subset where each data point get an equal chance to be included in both training and testing set. Cross validation is conducted using k folds, where k can be 3, 5 or 10. Some of the advantages of CV is that it can help scientists detect overfitting, provides average performance of the model rather than using single evaluation (i.e., 70:30%) [39]. Therefore, CV is crucial for more accurate evaluation of model performances.

CV is regarded as one of the most important approaches adopted in DL for selection of parameters and evaluation of model performance. CV enable the use of the entire dataset as training and testing respectively. Each fold is use as training and testing. In this study, 10k folds are used, 9 folds for training and 1-fold for testing for both microscopic slides and X-ray images. Each dataset contains 600 images (microscopic slides images and Chest X-ray images). Thus, the images are split into 9 folds (270) and 1-fold (30) for both positive and negative cases.

3.6 Model Training

For both microscopic slide images and chest X-ray images, 70% split was used for training in Matlab install with DL, image processing toolbox and AlexNet driver. The program is run on a PC with 8GB random access memory, intel R core-i7-3537U, GPU and window 64-bit. The remaining 30% was reserved for testing. The subsequent training of the model for general dataset and cross validation is adjusted to 20 epoch and 0.0001 learning rate. The code use for training and testing is included in the supplementary file.

3.7 Evaluation Metrics

To evaluate the performance of the learned models, some certain parameters are utilized; accuracy, Precision also known as sensitivity and specificity or recall. Accuracy is defined as the ratio of properly classified images over total sum of images; it is also described as the sum of precision and recall. For evaluating the accuracy and loss of the learned model the resulting formulas are employed:

$$Loss = -\frac{1}{n}\sum_{i=1}^{n}\log PC \tag{1}$$

$$Accuracy = -\frac{C}{N} \tag{2}$$

Where N is the overall number of images during training and testing, n is the number of images and PC is the probability of the correctly classified images.

Confusion matrix is the common approach used for evaluation of model performance based on True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN). TPs is the number of samples that are correctly identified by the model as positive cases or number of cases who actually have TB according to each model. TNs is the number of samples that are correctly identified by the model as negative cases or number of cases who are actually healthy (normal) and classified as negative according to each model. FPs are the number of samples that are incorrectly classified as negative by the model or number of cases that are actually negative (normal or healthy) but classified as TB according to each model. FNs are the number of samples that are incorrectly classified as positive by the model or number of cases that are actually positive (TB) but classified as normal or healthy according to each model as shown in Table 4.

True Positive rate (Sensitivity) is the portion of positive cases or samples which are precisely classified as positive sample (i.e., it describes the ration of positive cases that are correctly identified as positives).

$$\text{Sensitivity}\frac{\text{TPs}}{\text{TPs}+\text{FNs}} \tag{3}$$

False positive rate (FPR) also known as Specificity is the portion of positive cases or samples which are wrongly classified as positive samples (i.e., it describes the ratio of negative samples that are incorrectly classified as positives).

$$\text{Specificity}\frac{\text{TNs}}{\text{TNs}+\text{FPs}} \tag{4}$$

Table 4. Confusion matrix.

Predictions	Actual Positive	Actual Negatives
Positive Predictions	TP	FP
Negative Predictions	FN	TN

4. RESULT AND DISCUSSION

This section focusses on the results obtained as a result of training and testing of the pretrained AlexNet model using both microscopic slide images and chest X-ray images. The chapter also present the result obtained from cross validation of 2 dataset and comparison with state of the art.

4.1 Classification of Tuberculosis using Microscopic Slide Images

4.1.1 General Dataset: Microscopic Slide Images

Pretrained AlexNet was used to classify positive and negative cases of TB from microscopic slide images. The dataset was split into 70:30, 80:20 and 90:10 for training. In the case of 70:30 split, training of the model resulted in 580 iterations and 20 epochs as shown in Figure 5. The remaining 30% are used for testing. The model achieved 96.83%% training accuracy, 98.41% training sensitivity, 95.24% training specificity, 90. 56% testing accuracy, 97.78% testing sensitivity and 83.33% testing specificity as shown in Table 6 and Figure 5. In the case of 80:20 split, training of the model resulted in 660 iterations and 20 epochs. The remaining 20% are used for testing. The model achieved 96.53% training accuracy, 93.03% training sensitivity, 100% training specificity, 100% testing accuracy, 100% testing sensitivity and 100%

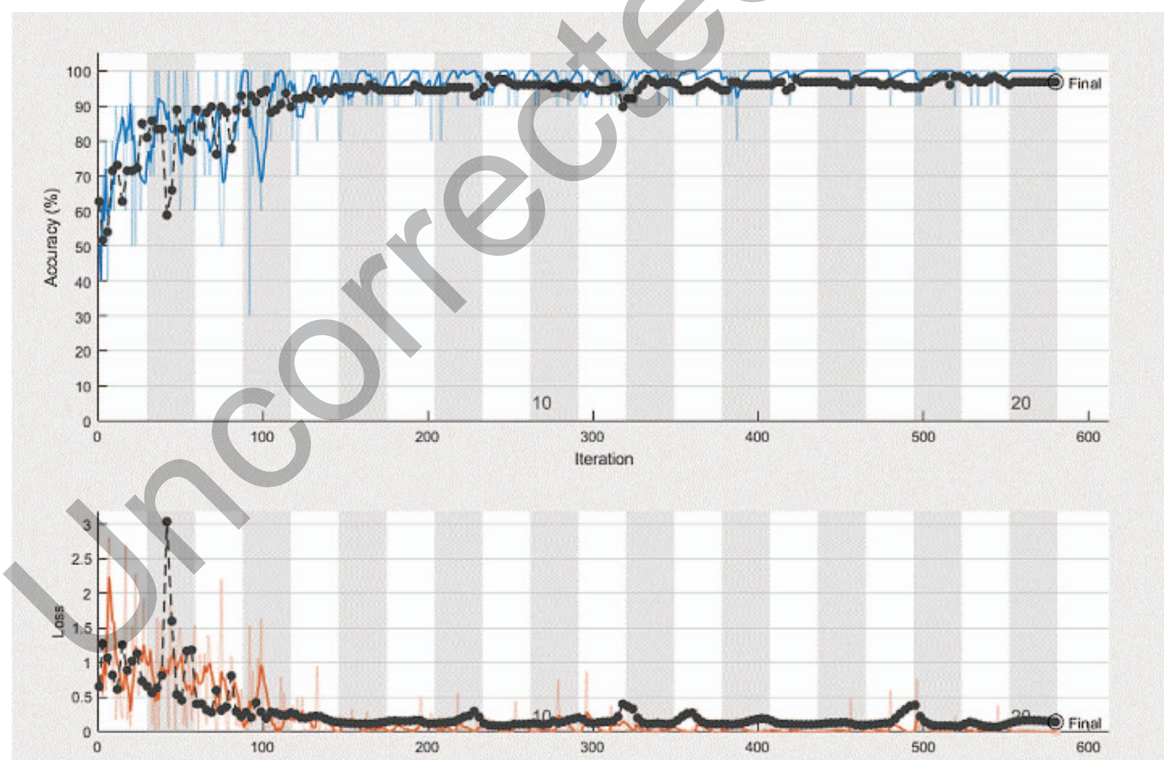

Figure 5. Accuracy and loss function for Training AlexNet Model using microscopic slide images.

testing specificity as shown in Table 6. In the case of 90:10 split, training of the model resulted in 740 iterations and 20 epochs. The remaining 10% are used for testing. The model achieved 97.53% training accuracy, 95.06% training sensitivity, 100% training specificity, 96.67% testing accuracy, 93.33% testing sensitivity and 100% testing specificity as shown in Table 5 and Figure 6. The results presented in table 5 has shown that increasing the number of dataset lead to increase in training accuracy. However, our results are in line with the study carried out by Prashanth et al. [40] based on data splits from 50%–90%. Moreover, 80:20 split achieved higher testing accuracy which shows that model is "fit" compare to 70:30 (with less training set) and 90:10 which is relatively "overfit" due to testing on small number of datasets.

Table 5. General Dataset Microscopic slide (70:30, 80:20 and 90:10).

Parameters	70:30		80:20		90:10	
	Training	Testing	Training	Testing	Training	Testing
Accuracy (%)	96.83	90.56	96.53	100	97.53	96.67
Sensitivity (%)	98.41	97.78	93.06	100	95.06	93.33
Specificity (%)	95.24	83.33	100	100	100	100

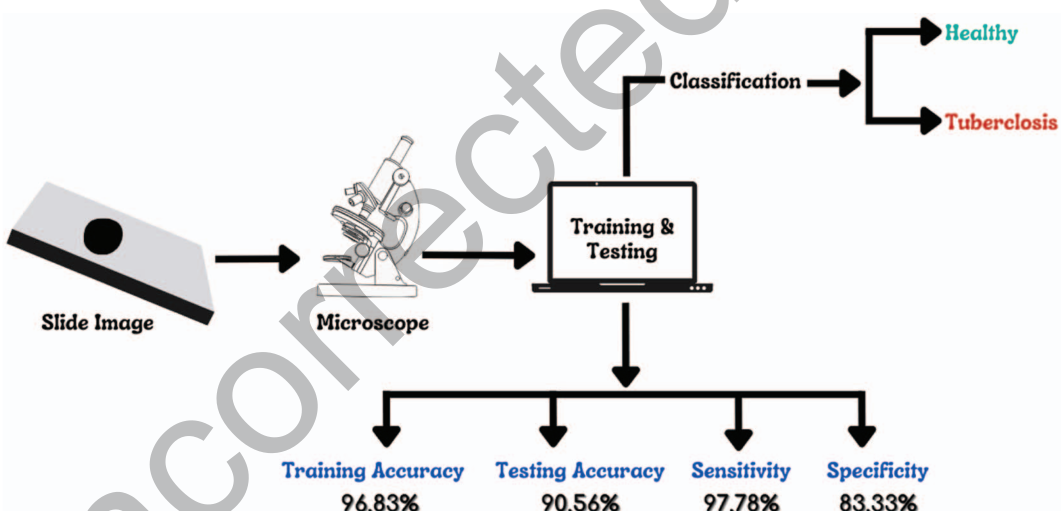

Figure 6. General Microscopic Slide Images Result.

4.1.2 Cross Validation for Microscopic Slide Images

CV was conducted in order to evaluate the general performance of the model by employing every K-fold as training and testing respectively. The model achieved average training accuracy of 88.70%, average training sensitivity of 99.50%, average training specificity of 97.78%. In terms of testing, the model achieved average testing accuracy of 94.66%, average testing sensitivity of 98.33% and average testing specificity of 91.00% as shown in Table 6.

Table 6. Cross Validation for Microscopic Slide Images.

K-fold	Training A*	Training SV*	Training SP*	Testing A*	Testing SV*	Testing SP*
K1	99.38	98.77	100	100	100	100
K2	98.77	98.77	98.77	90.00	100	80.00
K3	98.77	98.77	98.77	76.67	96.67	56.67
K4	99.38	100	98.77	93.33	96.67	90.00
K5	99.38	100	98.77	98.33	100	96.67
K6	98.15	100	96.30	95.00	90.00	100
K7	98.77	100	97.53	100	100	100
K8	95.68	98.77	92.59	93.33	100	86.67
K9	99.38	100	98.77	100	100	100
K10	98.77	100	97.53	100	100	100
Average	88.70	99.50	97.78	94.66	98.33	91.00

A*=Accuracy SV*=Sensitivity SP*=Specificity

4.2 Classification of Tuberculosis Using Chest X-ray Images

4.2.1 General Dataset: Chest X-ray Images

Pretrained AlexNet was used to classify positive and negative cases of TB from MS images. The dataset was split into 70:30, 80:20 and 90:10 for training and testing. In the case of 70:30 split, training of the model resulted in 580 iterations and 20 epochs as shown in figure 7. The remaining 30% are used for testing. The model achieved 99.21% training accuracy, 100% training sensitivity, 98.41% training specificity, 93.39% testing accuracy, 96.67% testing sensitivity and 91.11% testing specificity as shown in Table 8 and Figure 6. In the case of 80:20 split, training of the model resulted in 660 iterations and 20 epochs. The remaining 20% are used for testing. The model achieved 96.53% training accuracy, 97.22% training sensitivity, 95.83% training specificity, 94.17% testing accuracy, 98.33% testing sensitivity and 90% testing specificity as shown in Table 8. In the case of 90:10 split, training of the model resulted in 7540 iterations and 20 epochs. The remaining 10% are used for testing. The model achieved 99.38% training accuracy, 100% training sensitivity, 98.77% training specificity, 98.33% testing accuracy, 100% testing sensitivity and 96.67% testing specificity as shown in Table 7 and Figure 8.

Table 7. General Dataset Chest X-ray (70:30; 80:20 and 90:10).

Parameters	70:30		80:20		90:10	
	Training	Testing	Training	Testing	Training	Testing
Accuracy (%)	99.21	93.89	96.53	94.17	99.38	98.33
Sensitivity (%)	100	96.67	97.22	98.33	100	100
Specificity (%)	98.41	91.11	95.83	90.00	98.77	96.67

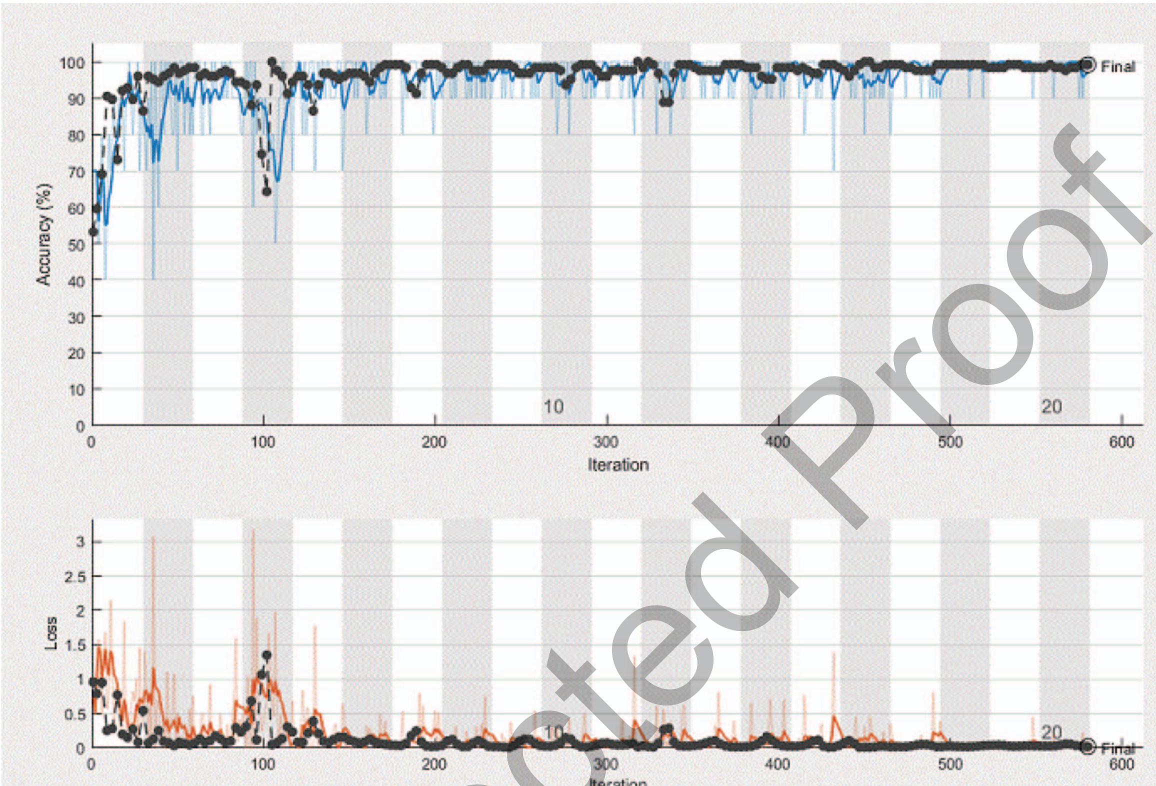

Figure 7. Accuracy and loss function for Training AlexNet Model using chest X-ray images.

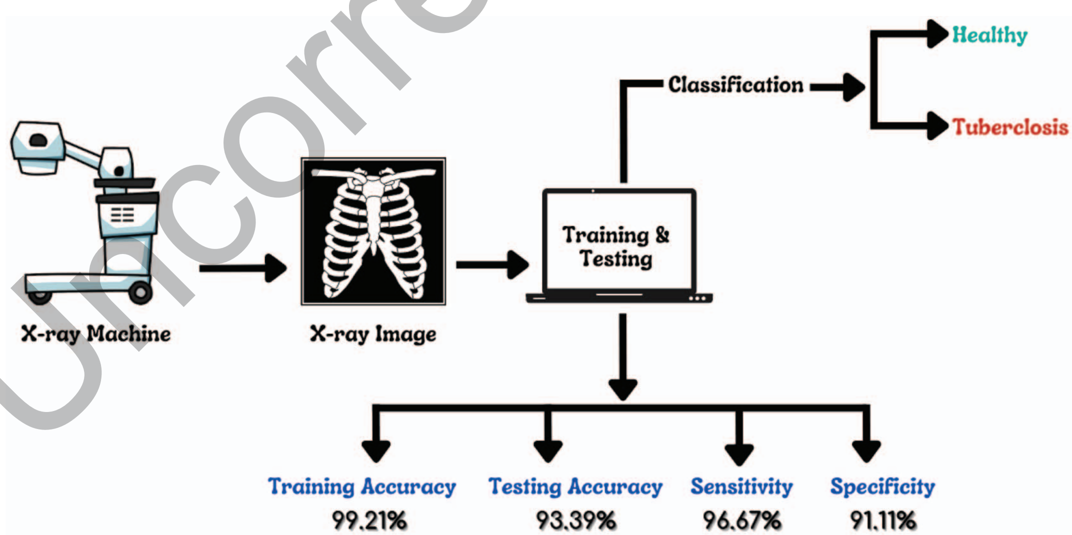

Figure 8. General Chest X-ray Images Result.

4.2.2 Cross Validation for Chest X-ray Images

CV was conducted in order to evaluate the general performance of the model by employing every K-fold as training and testing respectively. The model achieved average training accuracy of 97.22%, average training sensitivity of 99.87%, average training specificity of 95.58%. In terms of testing, the model achieved average testing accuracy of 94.00%, average testing sensitivity of 98.33% and average testing specificity of 98.33% as shown in Table 8.

Table 8. Cross Validation for Chest Xray Images.

K-fold	Training A*	Training SV*	Training SP*	Testing A*	Testing SV*	Testing SP*
K1	96.30	100	92.59	100	100	100
K2	99.38	98.77	100	98.33	96.67	100
K3	95.06	100	90.12	95.00	100	90.00
K4	96.91	100	93.83	91.67	100	83.33
K5	97.53	100	95.06	95.00	100	90.00
K6	93.83	100	87.65	85.00	100	70.00
K7	98.15	100	96.30	96.67	100	93.33
K8	96.30	100	92.59	91.67	100	83.33
K9	98.77	100	97.53	90.00	96.67	83.33
K10	100	100	100	95.00	90.00	100
Average	97.22	99.87	95.58	94.00	98.33	98.33

A*=Accuracy SV*=Sensitivity SP*=Specificity

4.3 Comparison with State of the Art

4.3.1 Comparison with State of the Art for Microscopic Slide Images

70:30 split is used for comparison as a result of support from many articles in the literature. The result obtained for general dataset yield 90.56% testing accuracy, 97.78% sensitivity and 83.33% specificity and average cross validation resulted in 94.66% average testing accuracy, 98.33% average testing sensitivity and 91.00% average testing specificity. Compare to study conducted by Khan et al. [25] with 94% accuracy. Despite the fact that the study trained and tested the model using over 12,000 microscopic slide images, our model which is trained and tested using 600 images achieved similar accuracy based on average CV and less based on 70:30 data split. According to literature, using SVM as a classifier increase model performance. Thus, the result achieved by Costa Filho et al. [27] (96.80%) outperform our model (both 70:30 spilt and CV). This can be attributed to the use of SVM instead of SoftMax use by our model. Smith et al. [24] utilized 25 thousand dataset and achieved 95% accuracy compared with our model which utilized 600 dataset and achieved 94.66% average CV and 90.56% accuracy. The study conducted by El-Melegy et al. [28] achieved higher accuracy (I.e., 98.4%) but achieved lower sensitivity ((I.e., 85.1%) compare to our model (97.78% for 70:30 split and 98.33% average sensitivity). The study conducted by Ibrahim et al. [23] achieved higher result (98.73% accuracy, 98.59% sensitivity and 98.84% specificity). This can be attributed to training pretrained AlexNet with over 2000 images compare to our study with 600 images. The result obtained by Xiong et al. [5] (97% sensitivity and 83.65% specificity) perform lower than our model in terms of sensitivity and specificity. Our model outperforms the result obtained by Muyama et al. [29] with 86.7% accuracy as shown in table 9.

Table 9. Comparison of AlexNet Model with state of the art for classification of microscopic slide.

Reference	Neural Network	Number of Dataset	Results		
			Accuracy	Sensitivity	Specificity
[25]	CNN	12,636	94%	-	-
[27]	SVM & CNN	120	96.80%	-	-
[24]	CNN	25,000 images	95%	-	-
[28]	CNN	500 images	98.4%	85.1%	
[5]	CNN	Cases: 45	-	97%	83.65%
[29]	ResNet	Not specified	86.7%	-	-
[23]	Pretrained AlexNet	2464	98.73%	98.59%	98.84%
70:30	Pretrained AlexNet	600	90.56%	97.78%	83.33%
CV	Pretrained AlexNet	600	94.66%	98.33%	91.00%

4.3.2 Comparison with State of the Art for Chest X-ray Images

The result obtained for general dataset yielded 93.89% testing accuracy, 96.67% testing sensitivity and 91.11% testing specificity and average cross validation result in 94.00% average testing accuracy, 98.33% average testing sensitivity and 98.33% average testing specificity. By comparing our result with the study conducted by Klassen et al. [30], our model achieved higher sensitivity and specificity as presented compare to the 87.2% sensitivity and 60% specificity achieved by Klassen and colleagues. The study conducted by Yahiaoui et al. [31] with 96.68% accuracy and the study conducted by Chang et al. [33] with 99.7% sensitivity and 99.9% specificity achieved better performance than our model. The most realistic comparison is based on the study conducted by Abbas and Abdelsamea [34] who utilized 662 X-ray images which is closer to our dataset of 600 X-ray images. However, despite the difference of 62 images our model performs better compare to the use of hybrid model proposed the 2 authors which 90.2% accuracy, 93.85% sensitivity and 86.76% specificity

The study conducted by Ibrahim et al. [36] using AlexNet+SVM achieved higher result (98.38%, 98.71% 98.04% accuracy, sensitivity and specificity respectively) compare to our study, this can be attributed with the use of large number of datasets (i.e., 7371 images). Moreover, the study conducted by Nafisah and Muhammad [37] achieved higher accuracy and average accuracy (99.1% and 98.7%) compare to our model. This result can be attributed to the use of EfficientNetB3 which contains more layers than AlexNet model as shown in Table 10.

Table 10. Comparison of AlexNet Model with state of the art for classification of Chest X-ray Images.

Reference	Model	Dataset	Results		
			Accuracy/AUC	Sensitivity	Specificity
[30]	Automated recognition and pattern analysis	308	-	87.2%,	60.0%
[31]	SVM	150	96.68%	-	-
[33]	Pretrained AlexNet	138	0.998 AUC	99.7%	99.9%
[34]	Hybrid model	662	90.2%	93.85%	86.76%
[36]	AlexNet+SVM	7371	98.38%	98.71	98.04
[37]	EfficientNetB3	1098	99.1%	-	-
70:30	**Pretrained AlexNet**	**600**	**93.89%**	**96.67%**	**91.11%**
CV	**Pretrained AlexNet**	**600**	**94.00%**	**98.33%**	**98.33%**

5. CONCLUSION

TB is one of the most common diseases that affect the lungs. The disease is very contagious and can spread from infected patients to healthy ones. Thus, diagnosis of TB is very crucial for early treatment, prevention and control. Medical expert employs different approaches for the detection of the disease, some of which includes tuberculin test, blood test, microscopic slide sputum smear test, genetic testing and radiographic imaging. The 2 common diagnosis of TB include microscopic slide sputum test and chest X-ray imaging. Despite the wide approaches employ by clinicians for the detection TB, it is still limited or hindered by so many challenges which includes lack of point of point care, low sensitivity, low background images, small size of bacilli, the use of toxic chemical reagents, time consuming, tediousness, the need for trained clinicians.

In order to address some of these challenges, we proposed the use of pretrained AlexNet model which is a DL model that recognize pattern integrated with classifier for classification of disease into positive and negative classes. The study employed 2 set of datasets which consist of microscopic slide images and chest X-ray images obtained from Near East University and Kaggle website respectively. The dataset undergoes image processing, labelling, splitting and training using Matlab software. The performance evaluation of the model based on training and testing are assess based on accuracy, sensitivity and sensitivity.

For classification of TB using MS images, t the model achieved 90.56% testing accuracy, 97.78% sensitivity and 83.33% specificity in the case of 70: 30 splits. In the case of 80:20 split, the model achieved 100% testing accuracy, 100% testing sensitivity and 100% testing specificity. In the case of 90:10 split, the model achieved 96.67% testing accuracy, 93.33% testing sensitivity and 100% testing specificity. In the case of CV, the model achieved 94.66% average testing accuracy, 98.33% average testing sensitivity and 91.00% average testing specificity.

For classification of TB using X-ray images, the model achieved 93.89% testing accuracy, 96.67% testing sensitivity and 91.11% testing specificity in the case of 70: 30 splits. In the case of 80:20 split, the model achieved 96.53% training accuracy, 97.22% training sensitivity, 95.83% training specificity, 94.17% testing accuracy, 98.33% testing sensitivity and 90% testing specificity. In the case of 90:10 split, the model achieved 99.38% training accuracy, 100% training sensitivity, 98.77% training specificity, 98.33% testing accuracy, 100% testing sensitivity and 96.67% testing specificity. In the case of CV, the model achieved 94.00% average testing accuracy, 98.33% average testing sensitivity and 98.33% average testing specificity. The success of these models on both microscopic slide images and chest X-ray images is in line with the notion that DL models can be useful for classification and discrimination of diseases into different groups with high accuracy, precision, sensitivity and specificity.

Some of the limitations of this study include the lack of sufficient datasets, different stained images, the use of frontal view posterior X-ray images. Thus, with sufficient amount of dataset with different variations (I.e., lateral and frontal view) we can expand the work using AlexNet (SoftMax or Support Vector Machine classifier), DenseNet, inception V3, InceptionResNet V2, MobileNet V2, NashNetmobile, VGG-16, VGG-19 and Xception. Moreover, the study can be expanded using ensemble models or hybrid models and the use of different classifiers and other techniques such as segmentation and data augmentation.

AUTHOR CONTRIBUTION STATEMENT

Abdullahi Umar Ibrahim (Abdullahi.umaribrahim@neu.edu.tr; ORCID: 0000-0003-3850-9921), Curation of Dataset, Experimental set up and model training, result and discussion; Ayse Gunnay Kibarer (aysegunay.kibarer@neu.edu.tr), Overview literature/related work and proofreading; Fadi Al-Turjman (fadi.alturjman@neu.edu.tr; 0000-0001-5418-873X), Overview Introduction and supervised the whole work.

REFERENCES

[1] World Health Organization Report on Tuberculosis (2020). Available at https://apps.who.int/iris/bitstream/handle/10665/329368/9789241565714-eng.pdf?ua=1. Access on 11 September, 2021

[2] Narasimhan, P., Wood, J., MacIntyre, C.R., Mathai, D.: Risk factors for tuberculosis. Pulmonary Medicine (2013). DOI: https://doi.org/10.1155/2013/828939

[3] Priya, E., Srinivasan, S.: Separation of overlapping bacilli in microscopic digital TB images. Biocybernetics and Biomedical Engineering 35(2), 87–99 (2015). DOI: https://doi.org/10.1016/j.bbe.2014.08.002

[4] Briganti, G., Le Moine, O.: Artificial intelligence in medicine: today and tomorrow. Frontiers in Medicine 7, 27 (2020). DOI: https://doi.org/10.3389/fmed.2020.0002

[5] Xiong, Y., Ba, X., Hou, A., Zhang, K., Chen, L., Li, T.: Automatic detection of mycobacterium tuberculosis using artificial intelligence. Journal of Thoracic Disease 10(3), 1936 (2018). DOI: https://doi.org/10.21037/jtd.2018.01.91

[6] Doi, K.: Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Computerized Medical Imaging and Graphics 31(4–5), 198–211 (2007). DOI: https://doi.org/10.1016/j.compmedimag.2007.02.002

[7] Halalli, B., Makandar, A.: Computer aided diagnosis-medical image analysis techniques. Breast Imaging 85 (2018)

[8] Chen, C.M., Chou, Y.H., Tagawa, N., Do, Y.: Computer-aided detection and diagnosis in medical imaging. Computational and Mathematical Methods in Medicine 2013. DOI: https://doi.org/10.1155/2013/790608
[9] Cicerone, M.T., Camp Jr, C.H.: Potential roles for spectroscopic coherent Raman imaging for histopathology and biomedicine. Neurophotonics and Biomedical Spectroscopy 547–570 (2019). Elsevier. DOI: https://doi.org/10.1016/B978-0-323-48067-3.00021-4
[10] Alloghani, M., Al-Jumeily, D., Mustafina, J., et al.: A systematic review on supervised and unsupervised machine learning algorithms for data science. Supervised and Unsupervised Learning for Data Science 3–21 (2020). DOI: https://doi.org/10.1007/978-3-030-22475-2_1
[11] Paiva, J.S., Cardoso, J., Pereira, T.: Supervised learning methods for pathological arterial pulse wave differentiation: a SVM and neural networks approach. International Journal of Medical Informatics 109, 30–38 (2018). DOI: https://doi.org/10.1016/j.ijmedinf.2017.10.011
[12] Helwan, A., Abiyev, R.: Shape and texture features for the identification of breast cancer. In Proceedings of the world congress on engineering and computer science, Vol. 2, pp. 19–21 (October 2016)
[13] Abiyev, R.H., Ma'aitaH, M.K.S.: Deep convolutional neural networks for chest diseases detection. Journal of Healthcare Engineering (2018). DOI: https://doi.org/10.1155/2018/4168538
[14] Mnih, A., Hinton, G.E.: A scalable hierarchical distributed language model. In: Proceedings of the 21st International Conference on Neural Information Processing Systems, pp. 1081–1088 (2008)
[15] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
[16] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770–778 (2016)
[17] Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems 25 (2012)
[18] Aloysius, N., Geetha, M.: A review on deep convolutional neural networks. In: 2017 international conference on communication and signal processing (ICCSP), pp. 0588–0592 (2017). IEEE. DOI: https://doi.org/10.1109/ICCSP.2017.8286426
[19] Liang, J., Liu, R.: Stacked denoising autoencoder and dropout together to prevent overfitting in deep neural network. In: 2015 8th international congress on image and signal processing (CISP), pp. 697–701 (2015). IEEE. DOI: https://doi.org/10.1109/CISP.2015.7407967
[20] Han, X., Zhong, Y., Cao, L., Zhang, L.: Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification. Remote Sensing 9(8), 848 (2017). DOI: https://doi.org/10.3390/rs9080848
[21] Raghu, M., Zhang, C., Kleinberg, J., et al.: Transfusion: Understanding transfer learning for medical imaging. Advances in Neural Information Processing Systems 32 (2019)
[22] Tsai, K.S., Chang, H.L., Chien, S.T., Chen, K.L., Chen, K.H., Mai, M.H., et al.: Childhood tuberculosis: epidemiology, diagnosis, treatment, and vaccination. Pediatrics & Neonatology 54(5), 295–302 (2013). DOI: https://doi.org/10.1016/j.pedneo.2013.01.019
[23] Ibrahim, A.U., Guler, E., Guvenir, M., et al: Automated detection of Mycobacterium tuberculosis using transfer learning. The Journal of Infection in Developing Countries 15(05), 678–686 (2021). DOI: https://doi.org/10.3855/jidc.13532
[24] Smith, K.P., Kang, A.D., Kirby, J.E.: Automated interpretation of blood culture gram stains by use of a deep convolutional neural network. Journal of Clinical Microbiology 56(3), e01521–17 (2018). DOI: DOI: https://doi.org/10.1128/JCM.01521-17
[25] Khan, M.T., Kaushik, A.C., Ji, L., Malik, S.I., Ali, S., Wei, D.Q.: Artificial neural networks for prediction of tuberculosis disease. Frontiers in Microbiology 10, 395 (2019). DOI: https://doi.org/10.3389/fmicb.2019.00395

[26] Panicker, R.O., Kalmady, K.S., Rajan, J., Sabu, M.K.: Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods. Biocybernetics and Biomedical Engineering 38(3), 691–699 (2018). DOI: https://doi.org/10.1016/j.bbe.2018.05.007

[27] Costa Filho, C.F.F., Levy, P.C., Xavier, C.D.M., Fujimoto, L.B.M., Costa, M.G.F.: Automatic identification of tuberculosis mycobacterium. Research on Biomedical Engineering 31, 33–43 (2015). DOI: https://doi.org/10.1590/2446-4740.0524

[28] El-Melegy, M., Mohamed, D., ElMelegy, T.: Automatic detection of tuberculosis bacilli from microscopic sputum smear images using faster r-cnn, transfer learning and augmentation. In Iberian Conference on Pattern Recognition and Image Analysis, pp. 270–278 (July 2019). Springer, Cham.

[29] Muyama, L., Nakatumba-Nabende, J., Mudali, D.: Automated detection of tuberculosis from sputum smear microscopic images using transfer learning techniques. In International Conference on Intelligent Systems Design and Applications, pp. 59–68 (December 2019). Springer, Cham. DOI: https://doi.org/10.1007/978-3-030-49342-4_6

[30] Klassen, V.I., Safin, A.A., Maltsev, A.V., Andrianov, N.G., Morozov, S.P., Vladzymyrskyy, A.V.: AI-based screening of pulmonary tuberculosis: diagnostic accuracy. Journal of eHealth Technology and Application 16(1), 28–32 (2018)

[31] Yahiaoui, A., Er, O., Yumuşak, N.: A new method of automatic recognition for tuberculosis disease diagnosis using support vector machines Biomedical Research 28(9), 4208–4212 (2017)

[32] Ahsan, M., Gomes, R., Denton, A.: Application of a convolutional neural network using transfer learning for tuberculosis detection. In: 2019 IEEE International Conference on Electro Information Technology (EIT), IEEE, pp. 427–433 (2019)

[33] Chang, R.I., Chiu, Y.H., Lin, J.W.: Two-stage classification of tuberculosis culture diagnosis using convolutional neural network with transfer learning. The Journal of Supercomputing 76(11), 8641–8656 (2020). DOI: https://doi.org/10.1007/s11227-020-03152-x

[34] Abbas, A., Abdelsamea, M.M.: Learning transformations for automated classification of manifestation of tuberculosis using convolutional neural network. In: 2018 13th International Conference on Computer Engineering and Systems (ICCES), IEEE, pp. 122–126 (2018)

[35] Sahlol, A.T., Abd Elaziz, M., Tariq Jamal, A., et al.: A novel method for detection of tuberculosis in chest radiographs using artificial ecosystem-based optimisation of deep neural network features. Symmetry 12(7), 1146 (2020). DOI: https://doi.org/10.3390/sym12071146

[36] Ibrahim, A.U., Al-Turjman, F., Ozsoz, M., Serte, S.: Computer aided detection of tuberculosis using two classifiers. Biomedical Engineering/Biomedizinische Technik 67(6), 513–524 (2022). DOI: https://doi.org/10.1515/bmt-2021-0310

[37] Nafisah, S.I., Muhammad, G.: Tuberculosis detection in chest radiograph using convolutional neural network architecture and explainable artificial intelligence. Neural Computing and Applications 1–21 (2022). DOI: https://doi.org/10.1007/s00521-022-07258-6

[38] Ghojogh, B., Crowley, M.: The theory behind overfitting, cross validation, regularization, bagging, and boosting: tutorial. arXiv preprint arXiv:1905.12787 (2019)

[39] King, R.D., Orhobor, O.I., Taylor, C.C.: Cross-validation is safe to use. Nature Machine Intelligence 3(4), 276–276 (2021). DOI: https://doi.org/10.1038/s42256-021-00332-z

[40] Prashanth, D.S., Mehta, R.V.K., Sharma, N.: Classification of Handwritten Devanagari Number–An analysis of Pattern Recognition Tool using Neural Network and CNN. Procedia Computer Science 167, 2445–2457 (2020). DOI: https://doi.org/10.1016/j.procs.2020.03.297

AUTHOR BIOGRAPHY

Abdullahi Umar Ibrahim was born in Zaria, Kaduna state on the 25th of October 1989. He studied at Nigerian Institute of Science and leather technology Zaria to obtain both OND and HND in Laboratory Science technology. He served at General Hospital Mina Niger state for 1 year. He studied MSc in Bioengineering at Cyprus International university and worked as Student Assistant at the university's information center. He completed his Doctoral degree in Biomedical Engineering and also worked as a research Assistant in the same department. He worked at Kaduna State University as a laboratory technology. Currently, he is an Assistant professor and also the vice chairman in Biomedical Engineering. His research interest is related to CRISPR as genetic engineering tool and application of Artificial Intelligence in Medicine.

Prof. Dr. **Ayse Günay Kibarer** received her BSc from Middle East Technical University (METU), Department of Chemistry in 1979. She completed her MSc in 1982. She was awarded with Fulbright Scholarship in 1983 and was invited by the late Prof. William J. Bailey to the University of Maryland. College Park, US to participate in his working group for her PhD studies. She returned to METU in 1984 and completed her PhD at METU in 1989. She stayed in the same department and was appointed as Assist. Prof. in the newly founded major branch of Biotechnology at Hacettepe University, Department of Biology. After receiving her associate professorship, she was invited by the chair of the Department of Chemistry at the same university in 1997 and continued her studies. In 2008, she was invited as a visiting professor from the University of Akron, OH, US. Recently, she has been appointed as a full-time professor by Near East University (2018) and has started working in the Department of Biomedical Engineering. Her main research interests are cold plasma and gold nanoparticles for cancer treatment, immobilization of microorganisms to produce enzymes, enzyme immobilization, architectural synthesis of microcarriers via raft polymerization technique, preparation of nanocomposites for antitumor activity and design of nanocomposite adhesives in solid melt working within a wide range of temperature for space applications.

Prof. Dr. **Fadi Al-Turjman** received his Ph.D. in computer science from Queen's University, Canada, in 2011. He is a full professor and a research center director at Near East University, Nicosia, Cyprus. Prof. Al-Turjman is a leading authority in the areas of smart/intelligent IoT systems, wireless, and mobile networks' architectures, protocols, deployments, and performance evaluation in Artificial Intelligence of Things (AIoT). His publication history spans over 350 SCI/E publications, in addition to numerous keynotes and plenary talks at flagship venues. He has authored and edited more than 40 books about cognition, security, and wireless sensor networks' deployments in smart IoT environments, which have been published by well-reputed publishers such as Taylor and Francis, Elsevier, IET, and Springer. He has received several recognitions and best papers' awards at top international conferences. He also received the prestigious Best Research Paper Award from Elsevier Computer Communications Journal for the period 2015–2018, in addition to the Top Researcher Award for 2018 at Antalya Bilim University, Turkey. Prof. Al-Turjman has led a number of international symposia and workshops in flagship communication society conferences. Currently, he serves as book series editor and the lead guest/associate editor for several top tier journals, including the IEEE Communications Surveys and Tutorials (IF 23.9) and the Elsevier Sustainable Cities and Society (IF 5.7), in addition to organizing international conferences and symposiums on the most up to date research topics in AI and IoT.