International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 8 Aug. 2022, pp: 317-323 www.ijaem.net ISSN: 2395-5252

Tuberculosis Detection Using Deep Learning

Rohan Rana[1,*], Abhishek Gaba[2],
[1]M.Tech Research Scholar, ECE Department, RIEIT, Railmajra, SBS Nagar, India
[2]Assistant Professor, ECE Department, RIEIT, Railmajra SBS Nagar, India

Submitted: 01-08-2022 Revised: 07-08-2022 Accepted: 10-08-2022

ABSTRACT : As a rising and developing nation, India demands concerted effort to combat tuberculosis (TB), a global infectious disease. This pandemic is threatened by resource-poor conditions such as socioeconomic inequity, inadequate infrastructure, and a lack of local specialists in the healthcare sector. Significant researchers are focusing on automated diagnosis via the application of statistical approaches to medical pictures, eliminating the need for individual image analysis and greatly lowering total expenses.
Deep Learning is becoming more popular for a variety of applications, including but not limited to autonomous driving, manufacturing, and medical imaging. Chest X-rays are the most cost-effective and widely utilized imaging technology, and they are employed as a dataset in research. Several deep learning architectures were investigated, with VGGnet achieving the greatest performance with 91.14 percent validation accuracy and 97 percent sensitivity. This concept has the added virtue of being able to function in low-resource settings, which may serve as the initial step toward enabling automated access for diagnostic purposes in underdeveloped parts of India and other developing nations.
Keywords : CNN(Convolution Neural Network), TB(Tuberculosis).

I. INTRODUCTION

1.1 Preamble: Mycobacterium tuberculosis is the causal agent of Tuberculosis, an infectious illness (TB). Tuberculosis and HIV/AIDS are the two most common infectious illnesses in the world, both of which may be fatal. Instead of indiscriminate mass screening for TB, the World Health Organization (WHO) has recommended quick molecular diagnosis and broader use of chest x-rays for high-risk populations. Tuberculosis screening methods include posteroanterior chest x-ray screening, which is regarded to be a significant element of the procedure. Furthermore, automated imaging detections serve a critical role in identifying individuals who may have pulmonary tuberculosis.

Deep learning makes advantage of Deep learning use neural networks to learn usable representations of characteristics directly from data. Image analysis has been shown to be very useful to deep convolutional neural networks, resulting in all of the winning submissions since 2012[1].

Deep learning-based Convolutional Neural Network (CNN) algorithms that can automatically learn from data pictures are successful on medical imaging [2]. CNN models can perform automated TB screening since they are more accurate than other approaches [3]. CNNs are made up of convolutional, pooling, and fully linked layers. It runs an image through the network layers and returns a final class, which is learned from the picture's characteristics. Filters of varying resolutions are applied to a picture, and convolution with the previous image is performed before the result is sent to the next layer [4]. Layers effectively learn basic and complicated characteristics. The final layer's features are convolved by the max pooling layer. The last layers are totally linked and soft max. With deep learning, feature extraction is a straightforward job, and learnt features aid in the training of a classifier[5]. There are numerous methods for training a deep learning network, including training from scratch, transfer learning, and semantic segmentation[6]. The transfer learning method is mostly used. Tuning the training settings is also critical for getting better outcomes[7]. Data augmentation improves performance with smaller datasets.

In 2012, Alex Krizhevsky employed CNNs to achieve the top 5 error rate of 15.4 percent[8]. Furthermore, in 2013-14, Mathew Zeiler'sZFnet[9] and Zisseraman'sVGGnet[10] were constructed with the goal of achieving lower mistake rates in the ILSVRC competition, and they were successful. Googlenet provided for the fact that as the depth and breadth of the network increased, improved error rates were found as a consequence, and it also gained first place in the relevant 2014 competition[11]. As a consequence of the creation of the Resnet model, which was derived from residual connections, the classification error rate decreased from around 25% in 2011 to 3.6 percent in 2015[12]. Researchers discovered the efficacy of employing four extracted CNNS characteristics as a highly competent base for categorization[13].

1.2 Training Techniques:

1.2.1 Starting from scratch:

Initialization: Gaussian random variables and biases that are reduced to zero are utilized to initialize the network parameters. This is known as Layer-sequential unit-variance (LSUV) initialization. Pixel intensities are calculated by dividing the standard deviation by the mean. The fundamental behind the operation of LSUV is the preservation of unit variance of input. Pretrained designs keep the standard deviation and mean.

A loss function indicates how effective our present classifier is. It determines the badness of every given classification, i.e., how poor our classifier is at categorizing the sample photos. The loss across the dataset is the average of the losses over all samples in the dataset. A loss function may be given the weight, biases, and instances from the training set as inputs. For example, the loss may be the number of photos successfully categorized. However, analgorithm comparable to the Stochastic Gradient Descent (SGD) is one of the most efficient methods for determining weight and biases in relation to the number of parameters.

1.2.2 Transfer Learning Elements Extraction:

The network learns the different characteristics from passing instances via a sequence of convolution and pooling operations. This is the stage at which, for example, if you showed the network an image of a cat, it would identify its tail, eyes, four legs, and ears. Alterations to the examples, such as rotation, are also feasible. The linked characteristics are then combined by averaging or stacking.

Adjustment: It entails training a pretrained neural network on our own dataset. In general, the final fully connected layer and the softmax layer are reset such that the size of the last layer equals the number of categories or classes in our dataset. The learning rate of the previous layers is reduced so that the network can adapt to the new dataset's characteristics. The more dissimilar the necessary dataset is from the original dataset, the more the parameters/layers must be reset.

1.2.3 Normalization Approaches:

Any data may be memorized by deep and massive neural networks. During training, their accuracy on the trainset often improves while it worsens on the testset. This is referred to as overfitting. Approaches are used to circumvent this regularization.

Regularization L2: The conventional weight decay strategy, which adds a term to the cost function to penalize the parameters in each dimension, prevents the network from precisely mimicking the training data and so helps generalize to new examples:

$$Err(x, y) = Loss(x, y) + \Sigma\, \theta i\, 2$$

θ: a vector that contains all of the network parameters.

Data augmentation: To avoid the CNN model from overfitting, the training set size must be increased, which is accomplished by data augmentation. This strategy is most often employed when the dataset is tiny. As a result, rather of adding new photos to the collection, tiny changes such as translations, flips, and rotations are done to the current dataset. They enhancements trick our neural network into thinking these are separate photos. Furthermore, augmentation facilitates the neural network's detection of the same kind of picture captured from varied viewpoints, settings, and noise levels. As a result, the network becomes more resilient.

Dropout: This is a new technology with a high success rate. Dropout is an effective and straightforward regularization approach. During training, randomly chosen neurons in each layer are dropped out and ignored, i.e., their activations are set to zero. This causes the network to become less sensitive to the individual weights of neurons, resulting in less overfitting and greater dataset generalization.

II. PRESENT WORK

CNN model with appropriate accuracy in a short computing time is necessary. Denser networks offer higher classification accuracy but may be more computationally demanding. For most medical diagnoses, Alexnet, Resnet, and Googlenet are utilized. The performance of other Pretrained models is yet to be determined. Deep learning model training often takes many hours. Model hyperparameters must be specified ahead of time to reduce computational complexity.The available

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 8 Aug. 2022, pp: 317-323 www.ijaem.net ISSN: 2395-5252

models are identified, and the best model for training the dataset is selected.

III. PROPOSED APPROACH

The initial step was to determine which CNN model performed best on our dataset. For this data, preprocessing was performed, and our dataset was divided into two discrete datasets, one for training and the other for testing, in a 70:30 ratio. This implies that 70% of the original dataset is selected at random to function as the training dataset, which will be used to train the model, and the other 30% is utilized to further evaluate the model's dataset accuracy. The next stage is model selection and hyperparameter configuration. A variety of CNN models may be employed to classify our data. Resnet50, googlenet, and vggnet were employed because they produced better results in a short period of time. Then, utilizing the hit-and-trial process, its many hyperparameters are tuned to provide the greatest outcomes for our purpose. The third stage is to apply supervised learning to the CNN model, which modifies the weight of the CNN model to effectively categorize our dataset.

There are 3500 TB images, and 3500 normal images in Qatar University, Doha, Qatar, and the University of Dhaka, Bangladesh, in conjunction with Malaysian partners and medical professionals from Hamad Medical Corporation and Bangladesh, have produced a database of chest X-ray. The X-rays are in PNG format and have a resolution of about 512 x 512 pixels

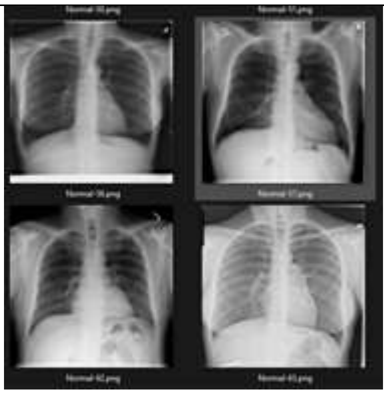
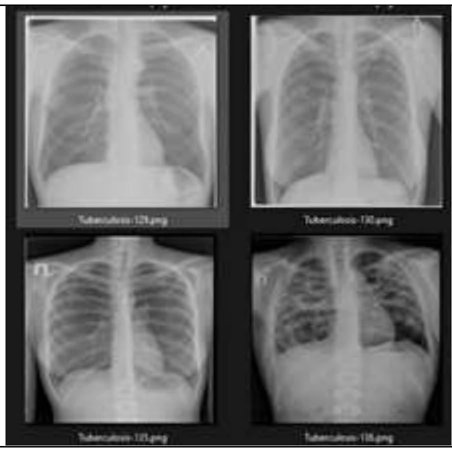
Normal	Effected
	

Fig 1 Sample images from the dataset

Convolutional neural nets are the primary neural networks used in image processing (CNNs). LeCun et alwork .'s identifying handwritten postal zip codes in the United States was one of the early uses of CNNs. Many additional applications followed, but it wasn't until the results of the 2012 ImageNet Challenge were announced that CNNs were extensively employed in image recognition tasks. Several CNN architectures were investigated, including LeNet, VGG, AlexNet, ResNet, and GoogleNet.A Comparison of Different CNN Architectures.

CNN Architecture	Error rate	Year	Developer	Parameter number
Alexnet	16 %	2012	Alex krizhevsky. IyaSutskever	60 Million
Vgg-16	7.3%	2014	Simonyan,Zisserman	138 Million
LENet	-	-	Yann LeCun et al	60 thousand
GoogLeNet	6.67%	2014	Google	4 Million
ResNet	3.6%	2015	Kaiming He	-

Table 1: A Comparison of Different CNN Architectures

By training the network via transfer learning, the performance of VGG16net on the CXR dataset is examined. Feature extraction and segmentation are not necessary in CNN-based models since the layers determine which characteristics are relevant and which are not. Table 3 lists the parameters utilized to train the VGG16. The parameters are therefore selected to provide the best outcomes.

Parameter name	Value
Batch size	8
Epoch	20
Optimizer	Adam
Loss	Categorical crossentropy
Metrics	Accuracy

Table 2: Hyperparameters Configuration

The first stage in training is to input data into the augment image datastore. It is recommended to load the pretrained network before extracting the layer graph from the trained network. Replace the last layers and freeze the first layers as well. The network is then trained using supplemented training data, and classification accuracy is measured with validation data. Our technique achieved 95.4 percent training accuracy and 99.34 percent validation accuracy, which are similar to the accuracies reported by other researchers on the same dataset.

Authors	Total No of samples	Accuracy
Mostofa Ahsan, et al.	388	81.25%
Sonaal Kant, et al.	400	83.78%
Xudong Liu, et al.	635	87%
Ravi Ghorakavi, et al.	800	81.33%
Lopes, et al.	800	82.8%
Rahul Hooda, et al.	800	82.09%
Marlon F. Alcantara, et al.	4701	89.6%
TawsifurRahman	3500	98.6%
Propose method	3500	99.34%

Table 3: Comparison of proposed research work

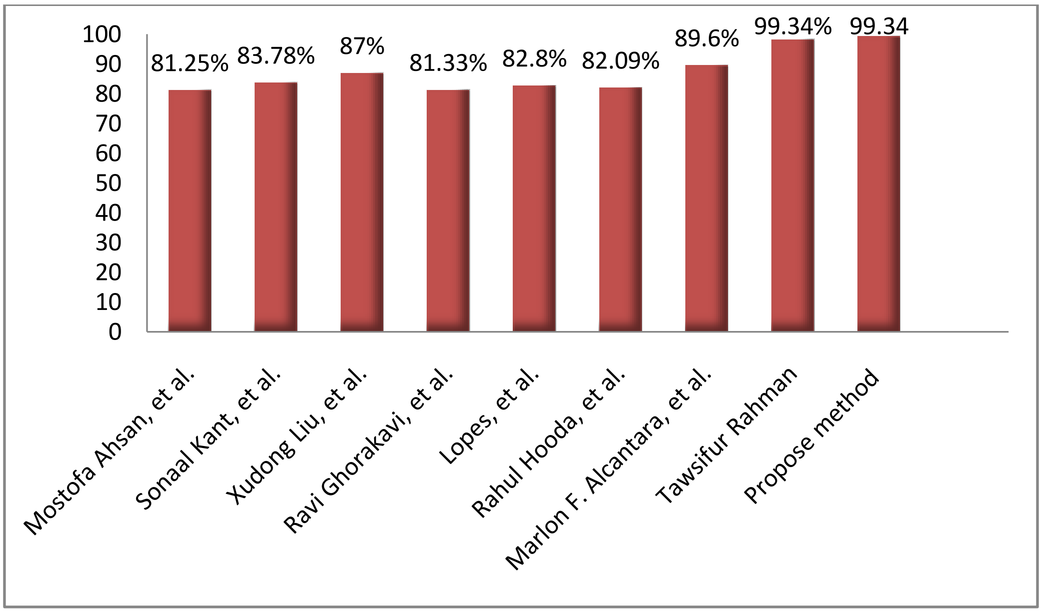


Fig 2 Comparison with others

The following parameters were used to evaluate the model's performance:

Verification Accuracy: A measure of information used to evaluate categorization models. Accuracy is defined as the percentage of accurate predictions among all queries provided to the model. A model's accuracy is often expressed as a percentage. It is a

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 8 Aug. 2022, pp: 317-323 www.ijaem.net ISSN: 2395-5252

measure of how close the model's guess is to the real label.

Receptivity: The test's sensitivity, also known as its recall, is defined as the fraction of patients who get a positive result who are actually positive for tuberculosis.

Particularity: The test's specificity is defined as the percentage of patients with a negative result who are actually negative or normal without any illness.

Precision: Precision is defined as the percentage of TB positive patients recognized out of all cases positive examples.

F1 Score: The F1 score, which is the harmonic mean of accuracy and memory, assesses the precision and recall.The CNN system's categorization abilities into normal and pathological instances.

Filters: The many validation formulae used for accuracy, precision , sensitivity and specificity. Real Effected (RE) cases are those that are accurately expected to have TB manifestations, i.e., the test result is positive and the case is actually positive. Not Effected (NE) is the number of instances properly anticipated as normal or healthy, i.e., the test result is negative and the case is actually negative. Wrong Effected (WE) refers to the number of healthy cases wrongly associated with TB manifestation, i.e., a positive test result that is really negative. The number of TB manifesting cases that are mistakenly forecasted as healthy, i.e. the test result is negative but is really positive, is shown by Wrong Negative (WN).

$$\text{Accuracy}=\frac{R_E+N_E}{R_E+N_E+W_E+W_N} \quad (1)$$

$$\text{Precision}=\frac{R_E}{R_E+W_E} \quad (2)$$

$$\text{Recall}=\frac{R_E}{W_N+R_E} \quad (3)$$

$$\text{Specificity}=\frac{N_E}{N_P+W_E} \quad (4)$$

IV. RESULTS AND DISCUSSION

The CXR dataset is used to train deep learning models such as GoogleNet, Resnet, and VGGnet. The characteristics supplied with the CXR image datasets were age, gender, and if the patient had TB. The total number of data sets 3500 TB images, and 3500 normal images in Qatar University, Doha, Qatar, and the University of Dhaka, Bangladesh, in conjunction with Malaysian partners and medical professionals from Hamad Medical Corporation and Bangladesh, have produced a database of chest X-ray. The X-rays are in PNG format and have a resolution of about 512 x 512 pixels.

The first trials were carried out by simply applying the resend, googlenet, vggnet, squeezenet, and alexnet models on the dataset, yielding accuracy of 79.84 %, 75.6 %, 72 %, 64.58 %, and 71%, respectively. Vggnet, Resnet, and Googlenet had the highest accuracies on the CXR dataset and were therefore chosen for further testing in the first studies. The dataset was divided into many folds, and each fold was subjected to analysis and iteration.

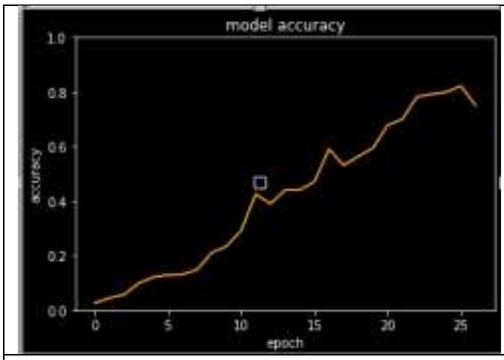


Fig 3 Googlenet model training

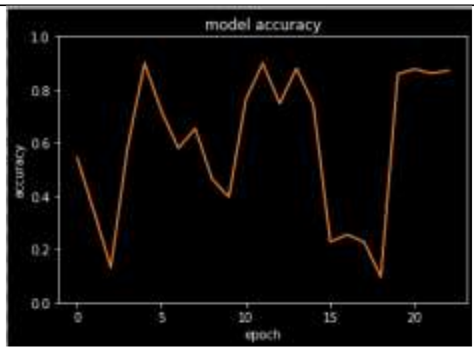


Fig 4Resnet model training

International Journal of Advances in Engineering and Management (IJAEM)
Volume 4, Issue 8 Aug. 2022, pp: 317-323 www.ijaem.net ISSN: 2395-5252

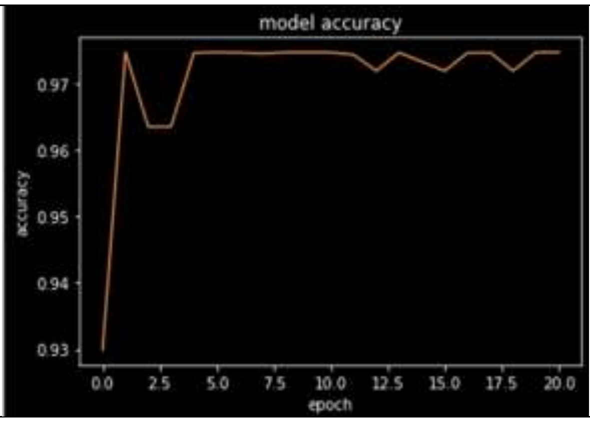


Fig 5VGG16 model training

Substantiation of model	Accuracy	Specificity	F1 score
Vggnet	99.34 %	94.5%	95.8
Resnet50	87.36%	88.03%	88.8
Googlenet	88.61%	94.08%	88.2

Table 4:Comparison of proposed method using different CNN models

With a validation accuracy of 99.34% percent, sensitivity of 98.5 percent, and f-score of 95.8, vgg16net produces the greatest results. Though vggnet produced superior overall results on the dataset, it had a slower computational speed (epoch per minute) than the models resnet and googlenet.

V. CONCLUSION AND PROSPECTS

This study provided a deep learning-based system for TB illness medical diagnosis. Convolutional neural networks are used for a variety of approaches. Based on calculation criteria, the suggested study demonstrated that VGGnet is superior than ResNet and Googlenet for the purpose of categorizing CXR pictures to diagnose TB in patients, with a high accuracy of 99.34 percent. VGGnet, ResNet, and Googlenet have F1-scores of 95.8, 88.8, and 88.2, respectively. CNNs can categorize the outcomes with high accuracy by learning various characteristics from the raw data.

REFERENCES

[1]. U. K. Lopes and J. F. Valiati, "Pre-trained convolutional neural networks as feature extractors for TB diagnosis," Computers in Biology and Medicine, doi: 10.1016/j.compbiomed.2017.08.001, 2017.

[2]. S. Hwang, H.-E. Kim, J. Jeong, and H.-J. Kim, "A new technique for TB screening based on deep convolutional neural networks," in Medical Imaging 2016: Computer-Aided Diagnosis, p. 97852W, Mar. 2016, doi: 10.1117/12.2216198.

[3]. P. Lakhani and B. Sundaram, "Deep learning at chest radiography: Automated categorization of pulmonary TB using convolutional neural networks," Radiology, doi: 10.1148/radiol.2017162326.

[4]. Y. Cao et al., "Improving Tuberculosis Diagnostics Using Deep Learning and Mobile Health Technologies in Resource-Poor and Marginalized Communities," in IEEE First International Conference on Connected Health: Applications, Systems, and Engineering Technologies (CHASE), Washington, DC, June 2016, pp. 274–281, doi: 10.1109/CHASE.2016.18.

[5]. M. Ahsan, R. Gomes, and A. Denton, "Application of a Convolutional Neural Network Using Transfer Learning for Tuberculosis Detection," in 2019 IEEE International Conference on Electro Information Technology (EIT), May 2019, pp. 427–433, doi: 10.1109/EIT.2019.8833768.

[6]. R. S. Ghorakavi, "TBNet: Pulmonary Tuberculosis Diagnosis System Using Deep Neural Networks," arXiv.

[7]. M. F. Alcantara et al., "Improving TB diagnosis in resource-poor areas in Peru using deep learning and mobile health technology," Smart Health, 2017, doi: 10.1016/j.smhl.2017.04.003.

[8]. A. Krizhevsky et al., "ImageNet Classification with Deep Convolutional Neural Networks Alex," 32nd International

Conference on Machine Learning (ICML-14), 2012.
[9]. R. Fergus and M. D. Zeiler, "Visualizing and Understanding Convolutional Networks," Computer Vision – ECCV 2014, vol. 8689, 2014. Springer International Publishing, Cham, 2014, pp. 818–833. D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds.
[10]. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv:1409.1556 [cs], Accessed: Dec. 04, 2020. [Online]. Available: \shttp://arxiv.org/abs/1409.1556.
[11]. C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, pp. 1-9, doi: 10.1109/CVPR.2015.7298594.
[12]. K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770-778, doi: 10.1109/CVPR.2016.90.
[13]. A. S. Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "CNN Features Off-the-Shelf: An Astounding Baseline for Recognition," 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp. 512-519, doi: 10.1109/CVPRW.2014.131.