# Shallow Cnn with Lstm Layer for Tuberculosis Detection in Microscopic Image

**Anson Simon, Vinayakumar R, Sowmya V, K P Soman**

***Abstract: Tuberculosis or TB, a disease mainly affecting lungs is infected by bacterium mycobacterium tuberculosis and diagnosed by careful examination of microscopic images taken from sputum specimen. Diagnosis of disease using microscopy and computer vision methods are applied for many previous practical problems. Recently, deep learning is playing major role in computer vision applications producing remarkable performance. But, computational complexity always remains as an obstacle in the application of deep learning in many aspects. So in this paper, a shallow CNN with LSTM layer is used for detecting the tubercle bacillus, mycobacterium tuberculosis, from the microscopic images of the specimen collected from the patients. The specified model is producing better performance than state of the art model and also have reduced number of learnable parameters, which requires comparatively less computation than the existing model.***

**Keywords:** ***CNN; deep learning; LSTM layer; tubercle bacillus; tuberculosis***

## I. INTRODUCTION

The remarkable growth in computer vision let us to use vast amount of medical images data for the diagnosis, treatment and detailed report of the diseases. By extracting the features like shape, texture, colour, and prior knowledge from the provided medical images data, modern computer vision technologies like machine learning, image segmentation, pattern classification etc. enables an efficient way of disease diagnosis than laboratory experts. Since images are provided in a 2D format containing pixel values, it is easier to make use of basic image processing techniques, which will be a boost for efficient disease diagnosis and treatment using computer vision methods. But, it is always a big challenge in managing computer vision in medical imaging due to the complexity in the medical images.

**Revised Manuscript Received on March 25, 2019.**

**Anson Simon**, Center for Computational Engineering and Networking (CEN), Amrita School of Engineering, Coimbatore
**Vinayakumar R**, Center for Computational Engineering and Networking (CEN), Amrita School of Engineering, Coimbatore
**Sowmya V**, Center for Computational Engineering and Networking (CEN), Amrita School of Engineering, Coimbatore
**K P Soman**, Center for Computational Engineering and Networking (CEN), Amrita School of Engineering, Coimbatore

Microscopy technology of using microscopes to view areas of interests, which cannot be seen by naked eye is mostly used as source for medical imaging data. Conventional optical or light microscopy involves transmitting visible light through and reflected back from the specimen through lens to view the sample magnified. The resultant image can be directly seen by the eye or can be digitally captured. In low-resource disease prone areas, microscopy is commonly used for diagnostic tasks, because of being both simple and versatile. New technologies like, diagnosis based on molecular biology and flow cytometry cannot be afforded in such places. Even if microscopes are available, due to the shortage of skilled technicians, disease diagnosis often depend upon disease symptoms and clinical signs alone, which is prone to error and results in death, drug resistance and loss of money for buying unnecessary drugs. Therefore, there should be another efficient alternative for quality diagnosis that is not available today in many places [1].Evolution of deep learning methods results in huge advancement in the field of computer vision to the extent that automated object recognition with accuracy outperforming human capability. Many previous problems depend upon hand engineered image features for a particular task, which are used by machine learning algorithms. Rather than relying on this, there should be a common approach in medical imaging, which is able to learn useful image representations automatically with adequate layers in the model. Therefore, this paper particularly focuses on the application of deep learning to the microscopy based point of care diagnosis of tuberculosis. The images are captured by economical device consists of smart phone attached microscope, which is intended to use in a resource constrained environment. Also, the proposed model aims to reduce the computational complexity by reducing the number of learnable parameters in the model. The following sections of this paper include the following discussions: Section 2 provides an overview on the previous works related to the task. Background study about disease TB is detailed in section 3. Methodology used to attain solution for the detection of TB is detailed in section 4. Section 5 shows the results obtained. Analysis of the results and conclusion is provided in section 6.

## II. RELATED WORKS

Success in the field of image digitizing and processing, joined with the accessibility of reagents for genome-scale activation, have empowered deliberate investigations of biological phenotypes.

Protein localization results in genesiological or hereditary perturbation casts a dominant role for protein in biological reaction [3]. High quality microscopes empowers examination of proteome wide transformations in diverse conditions, which creates data required to know the transitions in biological systems. In recent past, many trials have been attempted to create computational techniques for precise and appreciable investigation of proteome flow in yeast and different cells. Altogether 60 binary support vector machines (SVM) as a group is used here to classify images of single yeast cells from Open Reading frame - Green Fluorescent Protein (ORF-GFP) screens into 15 different sub-cellular localizations [4]. Each SVM classifier was trained on a training set consists of more than 70,000 cell images, which are manually annotated. This ensemble classifier (ensLOC) outperforms manual assessment by performing with 70% precision and recall. This approach also surpassed the earlier automated methods [5] [6], which depend upon SVMs for the classification of ORF-GFP collection. In the case of applying ensLOC classifier to other new microscopy datasets needs re-engineering and additional training processes. This is due to the diversity in the features used to train the classifier for each individual dataset. Mostly, image segmentation is needed in order to crop single cells from images and many features like pixel intensity and significant patterns are figured out for each cell. Also, if the feature space is high dimensional, choosing suitable relevant features for the classification or dimensionality reduction techniques will be utilized before the training process of a classifying model. These techniques like segmentation and feature reduction are not the same for all the datasets, which necessitate the researchers to retrain and tune classifiers for each new dataset. Deep learning technology can run over these obstructions related with changeable feature sets for different datasets, by learning only relevant features and performing classification precisely from pixel level information [7]. CNN in particular have surpassed human performance in object recognition tasks and now, it is being widely used in the biological imaging field. In recent past, progression of deep learning technology has been utilized for the classification in imaging flow cytometry [10], protein localization in yeast cells [8] [9] and also in the classification of aberrant mor phology in MFC7 breast cancer cells [11]. Also, a CNN model is implemented for the detection of patches for the diagnosis of certain diseases [1]. Here, in this proposed model, a combination of CNN with LSTM layer is used as classification model to detect and classify the patches present in the microscopic images in order to diagnose the disease TB and to provide better treatment. This model overcomes the drawbacks associated with prevailing machine learning classifiers and deep learning models with respect to performance, computation and transferability across various datasets.

## III. BACKGROUND

Tuberculosis is a potentially infectious disease worldwide which can lead into mortality by mainly affecting lungs. Mycobacterium tuberculosis, bacteria that causes tuberculosis are spread from one person to another through air when a person with TB coughs, sneezes, spits. Speedy shielding against TB is possible, but in rural low resource-high burden TB areas, current expensive diagnostic methods consisting of specialized equipment is not proper fitting. Mostly, diagnosis of TB is done by careful analysis and illustration of the existence of mycobacteria in the clinical specimens by sputum tests and smear microscopy. Acid-fast stains such as the fluorescent auraminerhodamine stain or the Ziehl-Neelsen (ZN) stain are favored for mycobacteria. One of the above mentioned stains helps to make a composite mixture between the mycolic acids of the mycobacterial cell wall and the pigment (carbon fuchsin), which secure the mycobacteria resisting to decolorization. After the preparation of the smear, it should be carefully examined by microscopes with high power to find the presence of tubercle bacilli, which may occur singly, V shaped or as clusters of bacilli. The proposed model detects the presence of tubercle bacilli in the microscopic images better than humans, which make the diagnosis and treatment of TB effectual.

## IV. METHODOLOGY

Process of training the proposed deep learning model from annotated images [2] to learn the useful complex pattern of images and testing this model in order to spot the microbes are described in this section.

### A. Train/Test dataset generation

Laboratory experts called pathologists annotate the regions with bounding boxes in microscopic images of sputum samples where the tuberculosis bacilli which can be seen by naked eye using an annotation software. This annotated dataset consists of 1265 sputum images with bounding boxes around bacilli. Each image in the dataset was down-scaled and split up into overlying patches according to scaling constant and patch size, which will be fixed by the type of pathogen to be detected. Each individual case should need to be inspected visually in order to check whether the patches are large enough to include all microbes but without needless details which creates extra processing load and also should be amplified enough to clearly view the objects by eye. Annotations are done by centering the bounding boxes around positive patches (bacilli). Negative patches (other objects such as staining artifacts, impurities except bacilli) are taken from random locations of images without intersecting any bounding boxes. Excessively large number of negative patches are present in the images in comparison with positive patches makes the model biased. Therefore, in order to solve the problem two ideas were implemented. First, randomly selected negative patches were castaway such that it became 100 times the number of positive patches. Next method is to design new positive patches by augmenting the patches i.e. by twisting and flipping, creating 7 additional positive samples for each genuine sample.

### B. Model architecture

Using a network with fully connected layers is not commonly used to classify images, because of the fact that such a network architecture does not consider spatial information of the images. It gives the same importance to the input pixels which are close together and far apart. This problem introduces the CNN such that, the information about the spatial structure is captured from the training data. CNN uses special architecture, which gets adapted efficiently for classifying the inputs. Design of CNN includes input layer, multiple hidden layers and an output layer. Basically, hidden layers consist of convolutional layer, pooling layer and fully connected layer. CNNs use shared weights in convolutional layer, which means same filter is used for each local receptive field in the input tensor, which reduces the number of learnable parameters and improves performance. In the proposed model, we used an additional layer called LSTM (Long Short-Term Memory units) layer in network.

### C. Convolutional layer

The image is passed to this convolutional layer, where number of filters (kernels) are set as one of the hyper parameters. Kernels defined are usually random while training from scratch and also each filters are one of a kind. A feature map is generated by panning the filters across the entire image. The feature map generated by convolution details enhanced representations of key features in the image like edges and unique patterns. Subsequent application of other weight filters to image successively generates a set of feature maps representing the same image but, different features.

### D. Pooling layer

Pooling is a technique, which is used for the dimension reduction and it is applied after one or more convolutional layers. This can play a major role in building a CNN as any extra layers worsens the computing time. Maxpooling is done to reduce the dimension of the image.

### E. Fully connected layer

In fully connected layer or dense layer, every neuron in the layer connects to every other neuron in the previous layer. Convolutional layers are actually producing meaningful, low dimensional, high level features in the data. Using a dense layer is the usual way of learning non-linear combinations of that features.

### F. LSTM layer

LSTM layer is a Recurrent Neural Network (RNN) layer that often used for time series and sequential data in a model. But it is slightly different from RNN. RNN modifies the complete existing information by executing a function. This reveals that RNN has no forethought on ‘necessary’ and ‘not so necessary’ information. On the other hand, LSTMs produce limited changes to the information by addition and multiplication. In LSTM, flow of information is done through cell-states which can selectively remember or forget information. The proposed model architecture is a shallow CNN with one convolution and pooling layer appended with one LSTM layer shown in Fig. 1.

**Table. 1:** Performance comparison results between state of the art model and the proposed model.

| Model | State of the Art : CNN [1] | **Proposed Model : CNN-LSTM** |
|---|---|---|
| No. of parameters | 77646 | **8698** |
| ROC : AUC | 0.99 | **0.99** |
| Precision-Recall :AP | 0.93 | **0.99** |
| Accuracy (%) | 95.21 | **96.29** |
| Precision | 0.95 | **0.96** |
| Recall | 0.95 | **0.96** |
| F1-Score | 0.95 | **0.96** |

### G. Detection in test dataset

After finishing the model training procedure, the developed model is competent enough to label the presence or absence of pathogen in small image patches. Above described method is used to classify patches in a limited part of an image, which consist of objects of interest. It is required to divide the image into patches in order to detect pathogens in an entire image or field of view. Then, these patches are evaluated one by one using the trained model and select those patches which have leading activation scores. Still, this unassisted technique alone influences the model to classify many overlying patches for each original patch present in the image used for testing process. This problem arises especially when small stride length is applied during the creation of duplicate patches resulting in overlapping. A technique called non-maximum suppression is used to work out this problem, which helps to sort out one activation per patch within the image. This method first catches the overlaps among the chosen patches in the test image. From the selected patches, sorting procedure should be carried out in order to choose the one with highest probability score and overthrowing the others.

## V.RESULTS

Trained network model created is then tested to the corresponding test set of tuberculosis images containing tubercle bacilli. By careful examination of detection results in Fig. 2, it is observed that the middle row which have high scored negative labelled test patch (false detection) contain bacilli and thus spotted annotation errors. This highlights the performance of the model over human diagnosis. Receiver Operating Characteristics (ROC) and Precision-Recall (P-R) curves are plotted in Fig. 3.

A solid and legitimate area under ROC curve (AUC) value gives the probability that the model will produce a higher score to an arbitrarily picked positive case than to an arbitrarily picked negative case. Average Precision (AP) outlines precision-recall curve as the weighted mean of precision produced at each point, with the hike in recall from the previous point utilized as the weight. TPR-FPR curve is also plotted. LSTMs are arguably most effective neural network, which is consistent in applying for images also. In comparison with convolution and pooling layer which only has local context window, LSTM layer establishes the fact that each feature activation in the output is a stimulation at the peculiar point in the whole input 2D tensor. LSTM layer in the network passes information of the entire input matrix through lateral connections, while CNN alone exploits only the local information. This helps to extract well-packed feature representation of the input with the help of lateral connections discarding unnecessary repeated features at distinct locations of the input. This allow the model to solve slight shifts of the features across numerous successive patches. Main limitation of the usage of LSTM layer is that, it blocks the parallel processing because of the sequential lateral connections. This limitation applies only to model parallelism and any technique of data parallelism is possible for the model. But, this initiative to use LSTM layer added with shallow CNN resulted in reduced number of parameters with overall less computations, although this should be further researched. In comparison to the state of the model [1], metrics of the proposed model are tabulated in Table 1. Comparison of the annotations done by both pathologists (white bounding boxes) and proposed model (red bounding boxes) is shown in Fig.5. Model annotations are more for the detection of bacilli than human annotations, shows the model can suggest more regions where likely to contain bacilli, which cannot be observed by the naked eye of pathologists.

## VI. CONCLUSION AND FUTURE WORK

Microscopic image analysis is a huge research area, which will be helpful for the development in pathology detection, better diagnosis and treatment of diseases. Earlier, deep CNNs have been used for the detection of intestinal parasites, tuberculosis and malaria on images from microscopes. In this paper, the proposed model shallow CNNLSTM architecture have shown performance improvement comparing to earlier state of the art models, which uses deep learning. In comparison with machine learning models with hand engineered features, the proposed model learns relevant representations of the data directly from pixel level information. This achievement can lead point of care (POC) diagnostics into next level, which is especially significant in the developing world, where the smart phones and microscopes are easily accessible than laboratory experts. Even with the presence of laboratory staff, the proposed model can be used as a decision backing tool for detecting possible pathogens in an image, which helps to take the final and accurate decision by the technician. This approach can be a helping hand for laboratory staff to attain flexibility in the diagnosis by concentrating more on the model detected regions of the image, where more chance lies for the occurrence of pathogens. This helps to provide better diagnosis and treatment for the disease.

In the future analysis, flexibility of the model should be checked on different microscopic datasets. This will help in creating a generalized model for diverse microscopic image datasets which helps in detecting patches, microbes or proteins etc. which assists the medical field in providing better diagnosis and treatment for people.

## REFERENCES


1. J. A. Quinn, R. Nakasi, P. K. Mugagga, P. Byanyima, W. Lubega and A. Andama, “Deep convolutional neural networks for microscopy based point of care diagnostics”, Machine Learning for Healthcare Conference, (2016), pp:271-281.
2. M. Bates and A. Zumla, “Rapid infectious diseases diagnostics using smartphones”, Annals of translational medicine, Vol.3, No.15, (2015).
3. M. Breker and M. Schuldiner, “The emergence of proteome-wide technologies: systematic analysis of proteins comes of age”, Nature reviews Molecular cell biology, Vol.15, No.7, (2014), pp:453–464.
4. B. T. Grys, D. S. Lo, N. Sahin, O. Z. Kraus, Q. Morris, C. Boone and B. J. Andrews, “Machine learning and computer vision approaches for phenotypic profiling”, J Cell Biol, Vol.216, No.1, (2017), pp: 65–71.
5. S. C. Chen, T. Zhao, G. J. Gordon and R. F. Murphy, “Automated image analysis of protein localization in budding yeast”, Bioinformatics, Vol.23, No.13, (2007), pp:i66–i71.
6. S. Huh, D. Lee, and R. F. Murphy, “Efficient framework for automated classification of subcellular patterns in budding yeast,”, Cytometry Part A: The Journal of the International Society for Advancement of Cytometry, Vol.75, No.11, (2009), pp: 934–940.
7. Y. LeCun, Y. Bengio and G. Hinton, “Deep learning”, Nature, Vol.521, No.7553, (2015), pp:436–444.
8. O. Z. Kraus, J. L. Ba and B. J. Frey, “Classifying and segmenting microscopy images with deep multiple instance learning”, Bioinformatics, Vol.32, No.12, (2016), pp:i52–i59.
9. T. P¨arnamaa and L. Parts, “Accurate classification of protein subcellular localization from high-throughput microscopy images using deep learning”, G3: Genes, Genomes, Genetics, Vol.7, No.5, (2017), pp:1385–1392.
10. P. Eulenberg, N. Koehler, T. Blasi, A. Filby, A. Carpenter, P. Rees, F. Theis and F. Wolf, “Deep learning for imaging flow cytometry:cell cycle analysis of jurkat cells”, bioRxiv,, (2016), pp:081364.
11. O. D¨urr and B. Sick, “Single-cell phenotype classification using deep convolutional neural networks”, Journal of biomolecular screening, Vol.21, No.9, (2016), pp:998– 1003.
12. F. Visin, K. Kastner, K. Cho, M. Matteucci, A. Courville, and Y. Bengio. “Renet: A recurrent neural network based alternative to convolutional networks”, arXiv preprint arXiv:1505.00393, (2015).
13. D. C. Cires¸an, A. Giusti, L. M. Gambardella, and J. Schmidhuber, “Mitosis detection in breast cancer histology images with deep neural networks”, International Conference on Medical Image Computing and Computer-assisted Intervention, Springer, (2013), pp:411–418.
14. R. Nayak, V. P. Shenoy, and R. R. Galigekere, “A new algorithm for automatic assessment of the degree of tb-infection using images of ZN-stained sputum smear”, International Conference on Systems in Medicine and Biology (ICSMB), IEEE, (2010), pp:294–299.
15. J. A. Quinn, A. Andama, I. Munabi, and F. N. Kiwanuka, “Automated blood smear analysis for mobile malaria diagnosis”, Mobile Point-of-Care Monitors and Diagnostic Device Design, Vol.31, (2014), pp:15.
16. V. Sujadevi, K. P. Soman, and R. Vinayakumar, “Real-time detection of atrial fibrillation from short time single lead ecg traces using recurrent neural networks”, The International Symposium on Intelligent Systems Technologies and Applications, Springer, (2017), pp:212–221.
17. Neethu Mohan, K. P. Soman, R.Vinayakumar, “Deep power: deep learning architectures for power quality disturbances classification”, 2017 International Conference on Technological Advancements in Power and Energy (TAP Energy, IEEE, (2017), pp:1-6.
18. Rajesh, M., and J. M. Gnanasekar. "Path Observation Based Physical Routing Protocol for Wireless Ad Hoc Networks." Wireless Personal Communications 97.1 (2017): 1267-1289.
19. Rajesh, M., and J. M. Gnanasekar. "Sector Routing Protocol (SRP) in Ad-hoc Networks." Control Network and Complex Systems 5.7 (2015): 1-4.

20. Rajesh, M. "A Review on Excellence Analysis of Relationship Spur Advance in Wireless Ad Hoc Networks." International Journal of Pure and Applied Mathematics 118.9 (2018): 407-412.
21. Rajesh, M., et al. "SENSITIVE DATA SECURITY IN CLOUD COMPUTING AID OF DIFFERENT ENCRYPTION TECHNIQUES." Journal of Advanced Research in Dynamical and Control Systems 18.
22. Rajesh, M. "A signature based information security system for vitality proficient information accumulation in wireless sensor systems." International Journal of Pure and Applied Mathematics 118.9 (2018): 367-387.
23. Rajesh, M., K. Balasubramaniaswamy, and S. Aravindh. "MEBCK from Web using NLP Techniques." Computer Engineering and Intelligent Systems 6.8: 24-26.