Abstract

Tuberculosis is a deadly infectious disease prevalent around the world. Due to the lack of proper technology in place, the early detection of this disease is unattainable. Also, the available methods to detect Tuberculosis is not up-to a commendable standards due to their dependency on unnecessary features, this make such technology obsolete for a reliable health-care technology. In this paper, I propose a deep-learning based system which diagnoses tuberculosis based on the important features in Chest X-rays along with original chest X-rays. Employing our system will accelerate the process of tuberculosis diagnosis by overcoming the need to perform the time-consuming sputum-based testing method (Diagnostic Microbiology). In contrast to the previous methods [4, 15], our work utilizes the state-of-the-art ResNet [2] with proper data augmentation using traditional robust features like Haar [23, 22] and LBP [16, 17]. I observed that such a procedure enhances the rate of tuberculosis detection to a highly satisfactory level. Our work uses the publicly available pulmonary chest X-ray dataset to train our network [3]. Nevertheless, the publicly available dataset is very small and is inadequate to achieve the best accuracy. To overcome this issue I have devised an intuitive feature based data augmentation pipeline. Our approach shall help the deep neural network [14, 2, 6] to focus its training on tuberculosis affected regions making it more robust and accurate, when compared to other conventional methods that use procedures like mirroring and rotation. By using our simple yet powerful techniques, I observed a 10% boost in performance accuracy.

1. Introduction

Tuberculosis (TB) is a deadly communicable disease. This disease can spread via air, water or any other forms of day-to-day activities. As a result, its one of the widely spread illness all around the world, making it one of the
deadly disease. To prevent the spread of this disease many countries have taken serious initiative by employing high health care standards and technology. However, the standards and technology employed to diagnose this disease usually need human interventions, as a result, they are costly, time-consuming and highly prone to errors. Such prolonged testing methods have a high chance of faulty outcome, which can result in mis-diagnosis of the patients. In medicine, it’s well known that Tuberculosis is curable when caught at an early stage, but due to such an obsolete procedure in place, it difficult for doctors to diagnose TB on time, resulting in an incurable state of tuberculosis.

Diagnosing TB involves the “chest radiography” and “sputum smear” testing. The “chest radiography” is used to check the abnormalities in the lung regions indicating the signs of tuberculosis. However, other abnormalities caused due to cough, chest pains, etc. observed in the chest X-rays do not unequivocally indicate that a person has tuberculosis. As a result, doctors perform an extra sputum test to confirm the inference from the chest X-ray. In this test sputum excreted by the infection is observed for the bacteria responsible for tuberculosis called “acid-fast bacilli” bacteria under a microscope. This method is very effective in diagnosing a person with TB but preparing the appropriate medium for this test takes at least 24 hours to make it viable for diagnosis, which is time consuming. Additionally, all these tests require diligent human involvement to perform the test.

As humans dependent testing methods are subject to errors, we need a high precision autonomous system to perform this task. As alluded to above, the testing method can be performed using sputum or X-ray. Although sputum-based tuberculosis system can generate very accurate results, it takes at least a day to complete the TB diagnosis. Therefore, we need an alternative method which is accurate, fast and consumes less time to address this issue. Leveraging on the success of modern deep neural networks on image detection and recognition [2, 6, 14], I utilized X-ray data to supply a fast and reliable system for TB detection.

It is well-studied in deep learning that deep neural networks requires a large number of training data samples to perform well [6]. However in medical imaging, collecting a large dataset is not easy as it requires individual consent. Consequently, previous approaches use Support Vector Machine (SVM)[1] based algorithms on publicly available small dataset to perform the tuberculosis detection, hence, not scalable. In contrast, in this work I propose a new data augmentation technique which helps deep neural network to work more efficiently on limited TB X-ray images. Our data augmentation technique conveniently utilizes traditional fast Haar features [22], LBP feature [16], and cropping of region of interest to improve the performance of TB detection.

In this paper, I have focused on developing a deep neural-network based algorithm for diagnosing tuberculosis using chest X-rays. For this purpose, I have used the pulmonary chest X-rays of 800 different people taken from the dataset created by National Library of Medicine, Maryland, the USA. the dataset is composed of the data collected from Shenzhen hospital and Montgomery hospital. The dataset collected from Shenzhen hospital has 662 chest X-ray images, and the Montgomery hospital has 138 chest X-ray images. This dataset is available as a free source for the educational and research purpose. I used both of these datasets to train a deep neural network (ResNet). Our work makes the following contributions:

(a) I present a method that can be used to diagnose a person for tuberculosis based on the pulmonary chest X-rays.

(b) In our method, I used the traditional features like Haar, LBP and cropped region of interest to perform data augmentation, which improves the performance of the deep neural network by making it more focused on tuning itself with the important features in the chest X-rays.

(c) I have also studied the effect of increase in testing accuracy on the network when the incomplete feature data is added in data augmentation step.

2. Related Work and Motivation

There are many computer-vision and machine-learning based tuberculosis detection algorithms [15, 13, 19, 21]. Nevertheless, there are two prominent ways to solve this problem. One way is to use the microscopic images of sputum to perform the feature detection and use machine learning based approaches [1] to detect the presence of tuberculosis. The other way is to use chest radiograph image features and use machine learning approaches to identify the abnormalities caused in the lungs due to tuberculosis [13].

Initial research focused on designing tuberculosis detection system uses microscopic images of the sputum [20, 21, 19]. In such methods, the image processing algorithms were utilized to focus on identifying, isolating and detecting the bacteria causing tuberculosis called Acid-fast bacilli (AFB) from its surroundings. Diagnosing a person with this method sometimes may not be very accurate as the bacteria can be misleading, resulting in false positive. Moreover, developing the culture to visualize the bacteria in the microscope is time-consuming. Despite such limitations, the sputum-based approach is used in practice for tuberculosis detection due to its par accuracy.

The second kind of methods use chest X-rays images to observe the abnormalities in the lung regions. These methods apply several image processing techniques to identify and isolate the abnormal regions in the lungs. As these methods for TB detection is based on X-rays image pro-
cessing algorithm which supplies similar response for other diseases too, hence, it’s not reliable and accurate. Therefore, we need a method which can keep track of minute differences that can be used to differentiate tuberculosis from other diseases. As a result, relying solely on image-processing based approaches to address this problem is not recommended. This motivates to design an algorithm that learns the salient features from the X-ray images which can reliably supply information about tuberculosis.

In this paper, I have applied a supervised approach to detect tuberculosis. Our algorithm is trained based on the doctor’s diagnostic results on the chest X-ray images. Since the publicly available datasets are very small, support vector machine algorithm’s are often used to solve this problem. However, I show that by cleverly exploiting the same small scale dataset, I can make the large-scale data-driven deep neural network work as good as SVM. This makes our algorithm applicable to large-scale tuberculosis detection systems and flexible for further usage as more and more dataset may get publicly available in the future.

Recently numerous attempt has been made to design a tuberculosis diagnosing system. Out of these approaches, a couple of recent research motivated us to pursue this work to achieve faster and accurate solution to this problem. Firstly, Melendez et al. [15] proposed a feature based diagnostic system which utilizes computer-aided detection (CAD) tool to identify and isolate the infected tuberculosis regions in the chest X-rays. Secondly, the method proposed by Lakhani et al. [13] train deep neural network both on the publicly available dataset, and data collected from the hospitals Belarus TB Public Health Program and Thomas Jefferson University Hospital, which are not publicly available for research purpose. This has created an added advantage for this work to generate the best accuracy of 99%.

(a) Melendez et al. approach [15]: This method proposed a framework that uses the combination of Computer-aided detection (CAD) score automatically computed from a chest radiograph reading (CXR) and 12 clinical x-ray features to achieve high accuracy for TB screening. This method was devised to overcome the drawbacks of radiological interpretation systems in TB-endemic countries. This method uses the dataset collected at Gugulethu TB Clinic (Cape Town, South Africa) claims to achieves an effective accuracy of 84%.

(b) Lakhani et al. approach [13]: This method uses a deep convolutional neural network to diagnose a person based on the chest radiographs. They employ DeepConvNet [14], AlexNet [6], and GoogLeNet [18] to train these network on their dataset. The dataset that they used is generated by collecting the data from four different hospitals to create a large pool of chest X-ray images. Before training, they have applied conventional data augmentation methods like 90, 180, and 270 rotation and Contrast Limited Adaptive Histogram Equalization to further increase the volume of the available dataset. Using this augmented dataset and pre-trained GoogLeNet, they were able to achieve the classification accuracy of 98% using GoogLeNet over 120 epochs.

In this paper, I argue that the accuracy claimed by Lakhani et al. [13] may not generalize to real-world TB diagnosis system due to over-reliability on self-made dataset. Also, CAD detection tools are expensive and require additional hardware interfacing. In contrast to the these methods, I devised a new method which significantly reduces the training time by increasing the training loss convergence rate using salient features like Haar and LBP. It’s well-known that training a convolutional neural network like GoogLeNet and AlexNet for 120 epochs results in over-fitting of the data. To make the system more robust and fast, I suggest an add-on technique which involves Haar features, LBP feature extraction methods with additional focus on region of interest via cropping lungs regions, in the data augmentation process. Such procedure increases the focus of training a neural network on the salient tuberculosis features. Our approach of data augmentation based on salient features improves both the performance and response time of TB diagnosis system, even with a small number of data samples in hand.

3. Deep Learning based classification model

Inspired from the recent work [15, 13], I developed a method which is designed to improve the accuracy of tuberculosis detection system. Our system uses deep neural network with an efficient way of data augmentation. To generate an accuracy that can be applied in the real-world application with the deep neural networks, we need at least 20,000 x-ray images, but publicly available dataset is composed of just 800 images. So, I need a data augmentation method to make our system more reliable and accurate. Our approach of data augmentation significantly improves the performance of the neural network. I first perform few experiments to decide which network can predict tuberculosis with reasonable precision on original dataset. For this, I use two networks, first network is composed of two convolution layers and two fully connected layers and the other is pre-trained ResNet18. The accuracy generated by ResNet18 when trained on original dataset (65.77181%) which is significantly higher when compared to the first network (52.34899%). As a result, I used our base network as the pre-trained ResNet18 to perform experiments.

In our approach to increase the accuracy of ResNet18, I have performed the data augmentation process using feature like Haar [22], LBP [16] and cropped region of interest. These process generates the data of salient features which
can be added to the pre-existing dataset of 800 images. It is well-known that large number of dataset can make the performance of deep neural network better than traditional machine learning algorithms like gradient descent, SVM etc. Using such notions, I augmented the salient feature with original images to improve the accuracy of tuberculosis diagnosing system. This operation of adding salient features with original images directs the network to focus on important information of the chest X-ray images resulting in significant and fast convergence of the network. Before going into details of the data augmentation process, let us investigate how feature extraction methods work.

(a) Haar Feature Extraction: This method is widely used in extracting features from a digital image to perform object recognition operation. Each object in nature has features which makes it stand out in the image. One way to extract these features is by comparing the intensity values of them with its neighbors. This operation is done pixel wise by considering each pixel to be an independent feature. In this algorithm, I perform the averaging operation on the pixel intensities surrounding the pixel of interest and use it to compare with the required average pattern of the average pixel intensity values of the reference lung image. The reference average pixel intensity values of an image are nothing but the predefined Haar features of lungs. The output of this algorithm is discussed in results section.

(b) LBP (Local Binary Patterns) Feature extraction: In this method, I perform the similar operation as Haar feature extraction method, but instead of averaging the neighboring pixel intensity values, I compare the intensity value of the interested pixel with its neighboring pixel intensities and assign a bit pattern to each neighboring pixel via comparison results. Later this bit pattern in the window of interest is compared with the reference bit pattern of the object of interest (lungs) to identify lungs. The reference bit pattern of the object of interest is nothing but predefined LBP features of lungs. The output of this algorithm is discussed in results section.

Both the methods were very effective in isolation the lung region from the chest X-rays. The accuracy of the Haar feature extraction method is 71.125% where 569 lung regions out of 800 are correctly segmented. Whereas in the LBP feature extraction method the accuracy of salient feature extraction is 79.125% where 633 lung regions out of 800 are correctly isolated. Later I have cropped the region of interest from the original dataset to increase the robustness of our data augmented method. The incorrectly identified lung regions generated in LBP feature extraction were grouped separately and are called noisy data further in this paper. Finally, to test the new approach to data augment, I have created three cases based approach. The description of these cases is as follows.

- **Case 1**: Data augmentation with one feature extraction Algorithm. (No noisy data added).
- **Case 2**: Data augmentation with two feature extraction Algorithm and cropping region of interest (Noisy data generated from LBP feature extraction is added).
- **Case 3**: Data augmentation with two feature extraction Algorithm and cropping region of interest (No noisy data added).

In case 1 I have added features generated using Haar feature extraction along with the original dataset to check if
this approach increases the accuracy. As the above test is successful, I have tested the case 2 and 3, where I performed the same operation as case 1 but instead of just one feature extraction method added data from all the three salient feature extraction methods. The difference in case 2 and 3 is that I have added the noisy data generated by the LBP feature extraction method is added in case 2 to test if the noisy data increases the performance of the neural network. Our idea was if the data generated by feature extraction method has at least a few features related to the region of interest, considering this data in training process should increase the accuracy of the network. This theory was proven to be accurate by generating the highest accuracy in the case 2 over the other cases. The detailed discussion of these results are done in the section (4).

I used widely employed cross entropy loss-function to train both Conv-Net and ResNet 18. Let \( \hat{y}_n \) and \( y_n \) be the predicted and supplied probability of having tuberculosis respectively. I defined our cross-entropy as

\[
L = -\frac{1}{N} \sum_{i=1}^{N} [y_n \log(\hat{y}_n) + (1 - y_n) \log(1 - \hat{y}_n)]
\]

Here, ‘\( N_a \)’ denoted the total number of samples after data augmentation. The weights of the loss are updated using ADAM optimization [5] i.e.

\[
\begin{align*}
m_{w}^{t+1} & \leftarrow \beta_1 m_w^t + (1 - \beta_1) \nabla_w L^t \\
v_{w}^{t+1} & \leftarrow \beta_2 v_w^t + (1 - \beta_2) (\nabla_w L^t)^2 \\
\hat{m}_w & := \frac{m_{w}^{t+1}}{1 - \beta_1^{t+1}} \\
\hat{v}_w & := \frac{v_{w}^{t+1}}{1 - \beta_2^{t+1}} \\
w^{t+1} & \leftarrow w^t - \frac{\eta \hat{m}_w}{\sqrt{\hat{v}_w} + \epsilon}
\end{align*}
\]

where \( m, v \) are the first and second moment vector respectively. \( \beta_1 \) and \( \beta_2 \) are the forgetting factors for gradients and second moments of gradient. \( \epsilon \) is a small scalar value to avoid division by zero. The overall procedure of our method is presented in Algorithm-1.

### 4. Experiments and Results

To simulate this model, I have used computing system with 16GB RAM, 6GB GTX1060 GPU with CUDA 9.2 interfacing. Our algorithm is simulated on python3.6. Our implementation can be also be realized on commodity desktop machine. Our experimentation is performed in two steps: In the first step I observed the effects of training Deep-ConvNet and ResNet18 when trained on original dataset. The first network i.e. Deep-ConvNet, is composed of two convolution layers and two fully connected layers which are tuned to generate the classification output during the training process see Fig.(2). The detailed specifications of the Deep-ConvNet are listed below in order of execution.

- \( 5 \times 5 \) Convolutional Layer with 32 filters, stride 1 and padding 2.
- ReLU Activation Layer
- Batch Normalization Layer
- \( 2 \times 2 \) Max Pooling Layer with a stride of 2
- \( 3 \times 3 \) Convolutional Layer with 64 filters, stride 1 and padding 1.
- ReLU Activation Layer
- Batch Normalization Layer
- \( 2 \times 2 \) Max Pooling Layer with a stride of 2
- Fully-connected layer with 1024 output units
- ReLU Activation Layer
- Fully-connected layer with 2 output unit

Later I have tuned the pre-trained ResNet18 on original data to compare its performance with Deep-ConvNet. The detailed architecture of ResNet 18 can be observed in the Fig.(3). As the network parameters in Deep-ConvNet are less when compared with the ResNet18 this Increase will create more room for the network to train, resulting in better accuracy when compared with Deep-ConvNet. Sometimes parameterized network might result in over fitting, with this experiment I decided the best network that can be utilized

---

**Algorithm 1 : TBNet**

**Input:** Augmented Dataset obtained using Haar, LBP features and cropped region of interest.

**Output:** Diagnosis result of tuberculosis.

1. Salient feature Extraction.
   - Haar Features extraction.
   - LBP Features extraction.
   - Cropping region of interest.

2. Augmentation with Haar, LBP features and Cropped region of interest over original dataset.

3. Splitting Augmented dataset : test, train and validation set splitting based on the configuration shown in Table (1) and Table (2)

4. Tuning pre-trained ResNet18 using the Cross Entropy Loss Eq.(1) and Adam optimizer Eq.(2).

5. Diagnosis result.

---

#### Algorithm-1

```python
Algorithm 1: TBNet

**Input:** Augmented Dataset obtained using Haar, LBP features and cropped region of interest.

**Output:** Diagnosis result of tuberculosis.

1. Salient feature Extraction.
   - Haar Features extraction.
   - LBP Features extraction.
   - Cropping region of interest.

2. Augmentation with Haar, LBP features and Cropped region of interest over original dataset.

3. Splitting Augmented dataset : test, train and validation set splitting based on the configuration shown in Table (1) and Table (2)

4. Tuning pre-trained ResNet18 using the Cross Entropy Loss Eq.(1) and Adam optimizer Eq.(2).

5. Diagnosis result.
```
The data extracted using LBP feature extraction is more accurate than the Haar feature extraction. The LBP feature extraction has the accuracy of \( 79.125\% \) in identifying the features whereas the Haar feature extraction has the accuracy of \( 71.125\% \). The output from the feature extractors is shown in the figure (5). While performing the feature extraction both the processes generate noisy data as in the figure (6) which is due to the similarities in the both the left and right lung regions. For example from the first couple of images it is evident that the similarities of left lung with right lung resulted in one lung being repeated twice. Sometimes, the lungs were also reversed in position as in the second image couple.

After extraction of the salient features, the pooling of data is done in three cases to clearly observe the effects of data augmentation using feature extraction methods. In the first case, I add the data generated using Haar feature extraction to the original data. In the second case I have added the data generated from both Haar and LBP feature images with original dataset along with the manually cropped salient feature images. I also added noisy image data generated during LBP feature extraction method to test if partially extracted features can improve the accuracy of the testing accuracy of the network. Finally in the Third case, I have added the data generated from both Haar and LBP feature extractions with original dataset along with the manually cropped salient feature extraction. These data cannot be ignored as they have few features related to TB. So, these data are pooled together and are called noisy data.

The results of the three cases are discussed in the Table (2). By comparing the testing accuracy of case 1 and case 3, the data augmentation performed with multiple feature
Figure 4: (a) Training loss Convergence of Network1 (Deep-ConvNet Fig. (2)) and Network2 (ResNet18 Fig. (3)) (b). Rate of Convergence of training loss on all three cases discussed in §3 along with network 2 trained on original dataset (c). Case 2 §3 training loss graph for 30 epochs.

| Database       | Dataset size (#images) | Training set | Validation set | Testing set | Validation Accuracy (10th epoch) | Testing Accuracy |
|----------------|------------------------|--------------|----------------|-------------|----------------------------------|-----------------|
| Original data  | 800                    | 650          | 100            | 150         | 70.0%                            | 65.77181%       |
| Case 1         | 1,379                  | 1000         | 100            | 379         | 64.0%                            | 72.48322%       |
| Case 2         | 2,969                  | 2557         | 150            | 412         | 81.333%                          | 75.42579%       |
| Case 3         | 2,812                  | 2400         | 150            | 412         | 78.667%                          | 74.69586%       |

Table 2: Case1: Performance using original image and Harr features. Case3: Performance using Harr, LBP without noisy features. Case2: Performance using Harr, LBP with noisy features. By observing the statistics of Case 2, it can observed that by using salient features like Harr, LBP along with noisy features can help boost the performance of the deep neural network.

Figure 5: From left to right Raw data, Haar data, LBP data and manually cropped data.

Figure 6: From left to right, Original image, Bad Haar feature extraction result (noisy data), Original image, bad LBP feature extraction result (noisy data)

that, when I compare the case 2 and case 3, the case 2 with noise data addition has generated best accuracy when compared with case 3, hence proving our idea, discussed in the section (3). Moreover, the rate of convergence is case 2 is much smoother and faster than the other cases, indicating the increase in performance over other approaches in data augmentation, as shown in the Fig.(4(b)).

5. Limitations

From the above results, I observed that the accuracy of the network is much better in case 2 with multiple salient feature based data augmentation with original dataset. But if the network is trained for more epochs (30 epochs) the op-
6. Conclusion and Future Work

I witnessed that by proper data augmentation techniques, we can make the state-of-the-art deep neural network work very efficiently on a small sample of available chest X-ray image dataset. In this paper, I have also demonstrated that salient features like Haar and LBP can be quite effective in improving the performance of the current state-of-the-art algorithms. The proposed approach in addressing this problem by fusing feature extraction methods and deep learning model has shown a substantial increase in the performance accuracy of the tuberculosis diagnosis system in very few epochs. Hence, such a system can be put to work on a real-time platform shall contribute in early diagnosis of TB. Lastly, I believe this algorithm can help save time, effort and money for health care organizations invested in treating TB. Furthermore, as my future work, I hope to employ 3D chest X-ray data to obtain structure from motion of lungs and achieve efficient TB abnormalities detection. The recent work done by Kumar et al. [10, 9, 11, 8, 7, 12] look very promising and can help improve the TB detection using 3D chest X-ray data.

References

[1] C. Cortes and V. Vapnik. Support-vector networks. Machine learning, 20(3):273–297, 1995. 2
[2] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016. 1, 2, 6
[3] S. Jaeger, S. Candemir, S. Antani, Y.-X. J. Wáng, P.-X. Lu, and G. Thoma. Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. Quantitative imaging in medicine and surgery, 4(6):475, 2014. 1
[4] S. Kant and M. M. Srivastava. Towards automated tuberculosis detection using deep learning. arXiv preprint arXiv:1801.07080, 2018. 1
[5] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014. 5
[6] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pages 1097–1105, 2012. 1, 2, 3
[7] S. Kumar. Jumping manifolds: Geometry aware dense non-rigid structure from motion. arXiv preprint arXiv:1902.01077, 2019. 8
[8] S. Kumar, A. Cherian, Y. Dai, and H. Li. Scalable dense non-rigid structure-from-motion: A grassmanian perspective. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2018. 8
[9] S. Kumar, Y. Dai, and H.Li. Spatio-temporal union of subspaces for multi-body non-rigid structure-from-motion. Pattern Recognition, 71:428–443, May 2017. 8
[10] S. Kumar, Y. Dai, and H. Li. Multi-body non-rigid structure-from-motion. In 3D Vision (3DV), 2016 Fourth International Conference on, pages 148–156. IEEE, 2016. 8
[11] S. Kumar, Y. Dai, and H. Li. Monocular dense 3D reconstruction of a complex dynamic scene from two perspective frames. In IEEE International Conference on Computer Vision (ICCV), pages 4649–4657, Oct 2017. 8
[12] S. Kumar, R. S. Ghorakavi, Y. Dai, and H. Li. A motion free approach to dense depth estimation in complex dynamic scene. arXiv preprint arXiv:1902.03791, 2019. 8
[13] P. Lakhani and B. Sundaram. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology, 284(2):574–582, 2017. 2, 3
[14] Y. LeCun, Y. Bengio, and G. Hinton. Deep learning. nature, 521(7553):436, 2015. 1, 2, 3
[15] J. Melendez, C. I. Sánchez, R. H. Philipson, P. Maduskar, R. Dawson, G. Theron, K. Dheda, and B. Van Ginneken. An automated tuberculosis screening strategy combining x-ray-based computer-aided detection and clinical information. Scientific reports, 6:25265, 2016. 1, 2, 3
[16] T. Ojala, M. Pietikäinen, and D. Harwood. Performance evaluation of texture measures with classification based on kull-back discrimination of distributions. In Pattern Recognition, 1994. Vol. 1-Conference A: Computer Vision & Image Processing., Proceedings of the 12th IAPR International Conference on, volume 1, pages 582–585. IEEE, 1994. 1, 2, 3
[17] T. Ojala, M. Pietikäinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. Pattern recognition, 29(1):51–59, 1996. 1
[18] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1–9, 2015. 3
[19] K. Veropoulos, C. Campbell, and G. Learmonth. Image processing and neural computing used in the diagnosis of tuberculosis. *IET*, 1998.

[20] K. Veropoulos, C. Campbell, G. Learmonth, B. Knight, and J. Simpson. The automated identification of tubercle bacilli using image processing and neural computing techniques. In *ICANN 98*, pages 797–802. Springer, 1998.

[21] K. Veropoulos, G. Learmonth, C. Campbell, B. Knight, and J. Simpson. Automated identification of tubercle bacilli in sputum: a preliminary investigation. *Analytical and quantitative cytology and histology*, 21(4):277–82, 1999.

[22] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, volume 1, pages I–I. IEEE, 2001.

[23] P. Viola, M. J. Jones, and D. Snow. Detecting pedestrians using patterns of motion and appearance. *International Journal of Computer Vision*, 63(2):153–161, 2005.