

Computer-Aided Detection of Tuberculosis using Convolutional Neural Networks

Agnus James

Department of Information Technology College of Engineering Perumon Perinad P O, Kollam, Kerala

Adarsh S

Department of Information Technology College of Engineering Perumon Perinad P O, Kollam, Kerala

Harisree P

Department of Information Technology College of Engineering Perumon Perinad P O, Kollam, Kerala

Vishnu S

Department of Information Technology College of Engineering Perumon Perinad P O, Kollam, Kerala

Abstract - Early TB screening and diagnosis are critical in the control and treatment of tuberculosis infections. For the identification of several types of TB lesions in chest radiographs, an integrated computer-aided approach based on deep learning is presented. Chest radiographs are evaluated in clinical practice by competent physicians for the identification of tuberculosis. This is, however, a time-consuming and subjective procedure. X-ray pictures of TB are frequently misclassified as other illnesses with similar radiologic patterns, resulting in patients receiving incorrect treatment and deteriorating their health. In the available literature, transfer learning utilising Alexnet, Google-Net, Resnet, and other methods has been used to diagnose TB. However, transfer learning using pretrained CNNs such as Nasnet Mobile, Nasnet Large, Darknet-19, and other pre-trained CNNs has yet to be thoroughly studied. The proposed research seeks to create a completely automated computer-assisted TB detection system.

Keywords - CNN; Tuberculosis; Segmentation; Computer-Aided detection, Transfer Learning; Pretrained networks

1. INTRODUCTION

Tuberculosis (TB) is a bacterial-caused chronic lung disease that is among the world's top 10 leading causes of mortality. Early tuberculosis screening and diagnosis are important for tuberculosis infection control and therapy. Computer-aided detection techniques might help radiologists diagnose TB from X-ray images. The objective of the proposed research is to develop computer-assisted methods for identifying TB from X-ray images. In practical practice, skilled physicians analyze chest radiographs for TB detection. This is, however, a subjective and time-consuming procedure. Subjective differences in disease diagnosis based on radiographs are inescapable. Importantly, TB CXR images are sometimes misclassified as other illnesses with similar radiologic patterns, resulting in patients receiving inappropriate treatment and deteriorating their health. In low-resource countries (LRCs), radiologists are also in short supply, particularly in rural regions. In this context, computer-aided diagnostic (CAD) equipment that analyze chest X-ray pictures can be quite useful in mass TB screening. Large-scale labelled datasets and deep convolutional neural networks (CNNs) are now readily available, leading in huge picture recognition success. CNNs make it possible to learn data-driven, highly representative, hierarchical picture features from adequate training data, but gathering data sets as thoroughly annotated as ImageNet in the medical imaging industry is still a challenge. Computed tomography (CT) is now the most frequently utilized technique of TB detection. Chest X-rays (CXRs) are utilized to confirm TB diagnosis in the majority of early cases due to the radiation dosage, cost, availability, and capability to disclose unexpected pathologic abnormalities among TB detection methods. For decades, scientists have been developing a computer-aided detection (CAD) system for the preliminary diagnosis

of tuberculosis-related diseases using medical imaging. In order to provide meaningful quantitative insight in the early stages, CAD relies on rule-based algorithms to identify and extract valuable pathogenic aspects from images; however, such methods are time-consuming, relying primarily on the artificial extraction of patterns containing usable information. Because the manifestation of many illnesses typically covers such a tiny portion of the overall picture, the difficulty of the feature identification method quickly increases. Furthermore, issues like poor data transferability and in-consistent performance with freshly generated data have hindered the CAD system from forming a well-founded conclusion with high accuracy using cumulative medical imaging data and changing disease mutations. Computer-aided detection techniques might help radiologists diagnose TB from X-ray images. The objective of the proposed research is to develop computer-assisted methods for identifying TB from X-ray images.

II. LITERATURE REVIEW AND RELATED WORKS

Deep CNNs have gained popularity as a result of their increased image classification performance. The network's convolutional layers, together with filters aid in retrieving spatial and temporal characteristics from a Image. In CNN applications where the dataset is not huge, transfer learning can be effective. Transfer learning has recently been used successfully in a variety of field applications such as manufacturing, medical, and baggage screening [2]–[4]. This eliminates the need for a big dataset and shortens the training period required by the deep learning algorithm when created from scratch [5][6]. For TB identification, nine prominent pre-trained deep learning CNNs were utilized, including ResNet18, ResNet50, ResNet101 [7], DenseNet201 [8], ChexNet [9], SqueezeNet [10], InceptionV3 [11], VGG19 [12], and MobileNetV2 [13]. Except for ChexNet, all of these networks were first trained on the ImageNet database. Residual Network (or ResNet) was created to tackle the vanishing gradient and degradation problem [7]. ResNet has various versions dependent on the number of layers in the residual network: ResNet18, ResNet50, ResNet101, and ResNet152. ResNet was effectively utilized for transfer learning in biomedical picture categorization. During training, deep neural network layers typically learn low or high level features, whereas ResNet learns residuals instead of features [14]. Because it does not train duplicate feature maps, the Dense Convolutional Network (DenseNet) requires fewer parameters than a standard CNN. DenseNet's layers are relatively thin, resulting in a smaller number of new feature-maps. DenseNet is available in four distinct versions: DenseNet121, DenseNet169, DenseNet201, and DenseNet264. DenseNet provides direct access to the original input picture as well as gradients from the loss function to each layer. As a result, the computational cost of DenseNet has been considerably decreased, making it a superior alternative for picture categoriza-

tion. ChexNet Pretrained is a customized version of DenseNet121 that has been trained specifically on a large number of chest X-ray images[9]. In comparison to the other networks, SqueezeNet and MobileNetv2 are extremely small networks. ResNet was effectively utilized for transfer learning in biomedical picture categorization. During training, deep neural network layers typically learn low or high level features, whereas ResNet learns residuals rather than features. Because it does not train duplicate feature maps, the Dense Convolutional Network (DenseNet) requires fewer parameters than a standard CNN. DenseNet's layers are relatively thin, resulting in a smaller number of new feature-maps. DenseNet is available in four distinct versions: DenseNet121, DenseNet169, DenseNet201, and DenseNet264. DenseNet provides direct access to the original input picture as well as gradients from the loss function to each layer. As a result, the computational cost of DenseNet has been considerably decreased, making it a superior alternative for picture categorization. ChexNet Pretrained model is a modified DenseNet121 network that has been trained on a large number of chest X-ray images. In comparison to the other networks, SqueezeNet and MobileNetv2 are extremely small networks. A fire module, which comprises of the Squeeze Layer and Expand Layer, is the foundation of the Squeeze Net network. Only 1*1 filters are used in the Squeeze layer, which feeds into an Expand layer with a combination of 1*1 and 3*3 convolution filters. VGG focuses on one of the most essential aspects of CNNs: depth. The VGG network's convolutional layers have a relatively narrow receptive field. A linear transformation of the input is performed using 11 convolution filters, which are followed by a rectified linear unit (ReLU) layer. The convolution stride is set at 1 pixel in order to retain spatial resolution following convolution. VGG comes in two varieties: VGG16 and VGG19. Except for the first layer, which is a complete convolution, the MobileNet structure is constructed on depth-wise separable convolutions. With the exception of the last fully connected layer, which has no nonlinearity and feeds into a Softmax layer for classification, all layers are followed by batch normalization and ReLU nonlinearity. Before the completely linked layer, a last average pooling decreases the spatial resolution to 1. MobileNet contains 28 layers when depth-wise and pointwise convolutions are counted separately. Convolutional Neural Networks employ Inception modules to enable for more efficient computation and deeper networks by reducing dimensionality using stacked 11 convolutions. The modules were created to address concerns such as computational cost and overfitting, among others.

A. Dataset

For Tuberculosis detection, the proposed approach uses two publicly available datasets: the Shenzhen dataset and the Montgomery dataset (TB) Jaeger S, Candemir S, Antani S, Wang YXJ, Lu PX, Thoma G (2014) Two public chest x-ray data sets for computer-aided screening of pulmonary diseases. Quantitative imaging in medicine and surgery 4(6):475 [1]. The Shenzhen dataset includes 326 normal and 336 TB X-ray pictures in all. Shenzhen Hospital in China collected it. The chest X-rays were obtained from outpatient clinics and were taken on a daily basis for a month, primarily in 2012, with the use of a specialist medical diagnostic system. It comprises 662 frontal chest X-ray pictures, 326 of which depict benign cases and 336 of which depict TB (malignant) patients. The naming pattern for all picture files is the same: CHNCXRX.png, where X can be "0" stands for Non-TB (benign) or "1" Tuberculosis (malignant) X-ray. For each X-ray, a clinical report is accessible in a file with the same format, which includes the patient's age, gender, and any abnormalities observed in the lungs. The Montgomery dataset contains 138 frontal chest x-ray pictures, 80 of which are CRs of healthy lungs and 58 of which are tuberculosis-infected lungs. The health department of Montgomery County in Maryland, USA, gathered all of the pictures in this collection. The resolution of the radiographs is either 4,020x4,892 pixels or 4892x4020 pixels. Additional pictures with manually produced lung segmentation masks for Eve are included in this collection.

III. METHODOLOGY

A. Data Preprocessing

The dimensions of the X-ray pictures in the two databases are different. The pictures must be scaled to fit the input size that pre-trained CNNs can handle. Different CNNs have different input sizes. For example, Alexnet's input size is 227 x 227 x 3 while Nasnet mobile's is 224 x 224 x 3. Rotation, zooming, and other picture enhancement techniques are presented. Pretrained networks are used to segment tuberculosis-infected areas. The most effective CNN for medical picture segmentation is the U-Net[14] architecture presented by Olaf Ronneberger, Phillip Fischer, and Thomas Brox in 2015. We suggest utilizing U-Net CNN to separate TB-infected regions from X-ray images (fig 1). A shrinking path and an expanding path make up the Unet CNN. The contracting route consists of two 33 convolutions (unpadded convolutions) applied repeatedly, each followed by a ReLU and a 22 max pooling operation with stride 2 for down sampling. An up sample of the feature map is followed by a 22 convolution ("up-convolution") that halves the number of feature channels, a concatenation with the similarly cropped feature map from the contracting route, and two 33 con-

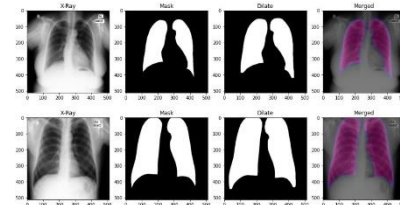


Fig. 1. segmented lung images.

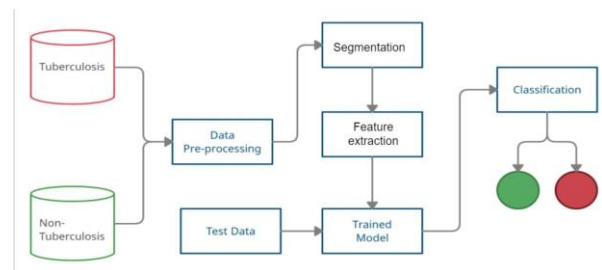


Fig. 2. Data flow diagram

volutions, each followed by a ReLU, in the expanding path. The network employs a total of 23 convolutional layers. A contracting path and an extensive path make up the U-Net architecture. So that pictures may be divided continually, the pixels in the border region are proportionately added around the image.

B. Feature Extraction and Classification

Convolutional Neural Networks (CNN) have recently become the most popular feature extraction approach. A vast quantity of labelled data is required to train a CNN from scratch. However, a substantial number of pictures are not available for medical imaging applications such as TB diagnosis. This is due to ethical issues at hospitals and diagnostic centers over the distribution of patient data. Pre-trained networks, which are CNNs trained using a large number of natural pictures, such as the ImageNet database, can be utilized in these situations. The key idea behind transfer learning with extremely deep neural networks is to retrain a CNN model that was previously trained on the ImageNet dataset (on our dataset) (about 1.2 million images). Because the dataset contains a wide range of items (1000 distinct categories), the model may learn a variety of different sorts of features, which can subsequently be applied for additional classification tasks. Any layer activation in pre-trained CNN can be utilized as a feature. There are numerous pre-trained networks, such as NasnetMobile and Nasnet Large, that have yet to be used in the detection of tuberculosis. The Softmax layer of pre-trained CNN performs classification. We suggest using an underutilized pre-trained CNN for tuberculosis detection and adding extra layers to it to create a network that can identify tuberculosis. The proposed method uses VGG-16 pre-trained CNN for feature extraction and classification.

IV. RESULT

Early TB diagnosis and treatment are critical for tuberculosis infection control and treatment. As a result, a computer-aided detection system must be implemented as a rapid alternative diagnosis option to prevent Tuberculosis from spreading among

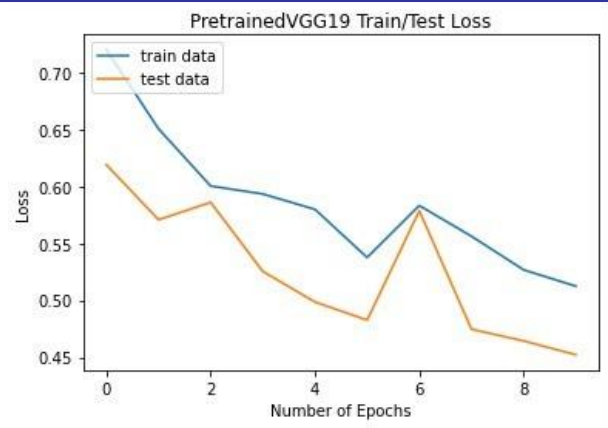


Fig. 4. Train/test model loss

individuals. Computer-aided automated diagnostic tools might become more trustworthy if the accuracy of TB diagnosis from chest radiographs could be improved with a robust and adaptable technique. Radiologists may be able to diagnose tuberculosis from X-ray pictures using computer-aided detection techniques. The detection of tuberculosis was achieved with 80 percentage accuracy using our suggested method. The fig 3 depicts the training/test accuracy using the VGG19 that has been pre-trained. The loss in training and test data using the pretrained CNN is shown in Fig 4. The findings from confusion matrix is shown in Table 1. A confusion matrix is a table that shows how well a classification model (or "classifier") performs on a set of test data for which the real values are known. This contains the sensitivity, accuracy, F1 score, and other parameter values. A receiver operating characteristic curve (ROC curve) is a graph that shows how well a classification model performs across all categorization levels. Two parameters are shown on this curve: Rate of True Positives and Rate of False Positives. ROC curve of the method is illustrated in Figure 6. The proposed method achieved an Area under the ROC Curve (AUC) of 0.88.

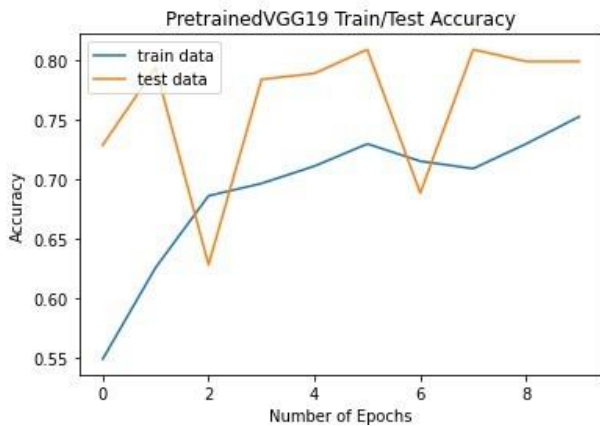


Fig. 3. Train/test model accuracy

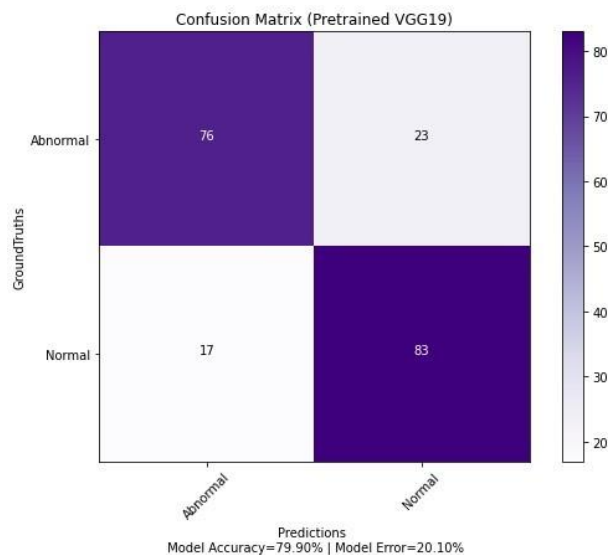


Fig. 5. confusion matrix

Measure	values
Sensitivity	0.8172
Specificity	0.783
Precision	0.7677
False Positive Rate	0.2170
False Discovery Rate	0.2323
False Negative Rate	0.1828
Accuracy	0.7990
F1 Score	0.7917
Matthews Correlation Coefficient	0.5989

Table 1. summarization

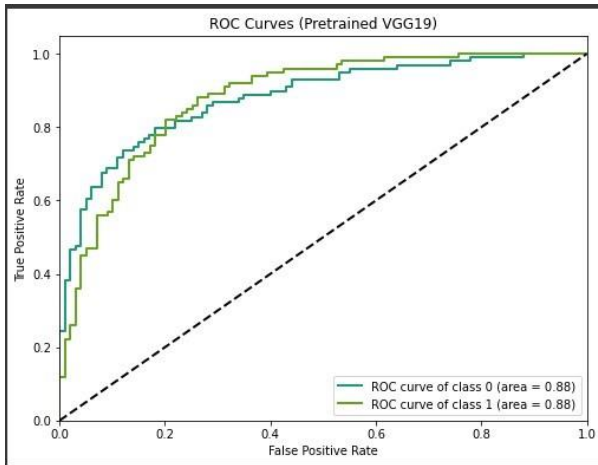


Fig. 6. ROC curve

V. CONCLUSION

This paper proposes a deep learning-based computer-aided method for detecting several types of TB lesions on radiographs. Transfer learning utilizing Alexnet, GoogleNet, Resnet, and other methods has previously been done in the literature. In this study, segmentation is accomplished through the use of CNNs such as UNET and feature extraction and classification using VGG-16 pre-trained CNN. To extract additional data from restricted datasets, data augmentation techniques are utilized. In the future, pretrained CNN from Keras such as MobileNet, Efficientnet B0, and Darknet-53 will be utilized. Tuberculosis may be detected from other X-ray pictures with the aid of CNN.

REFERENCES

- [1] Jaeger S, Candemir S, Antani S, Wang YXJ, Lu PX, Thoma G (2014) Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. *Quantitative imaging in medicine and surgery* 4(6):475.
- [2] S. Christodoulidis, M. Anthimopoulos, L. Ebner, A. Christe, and S. Mougiakakou, "Multisource transfer learning with convolutional neural networks for lung pattern analysis," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 1, pp. 76–84, Jan. 2017.
- [3] H. Yang, S. Mei, K. Song, B. Tao, and Z. Yin, "Transfer-learning-based online Mura defect classification," *IEEE Trans. Semicond. Manuf.*, vol. 31, no. 1, pp. 116–123, Feb. 2018.
- [4] S. Akçay, M. E. Kundegorski, M. Devereux, and T. P. Breckon, "Transfer learning using convolutional neural networks for object classification within X-ray baggage security imagery," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2016, pp. 1057–1061.
- [5] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang, "Convolutional neural networks for medical image analysis: Full training or fine tuning?" *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1299–1312, May 2016.
- [6] S. Jialin Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.
- [7] ResNet, AlexNet, VGGNet, Inception: Understanding Various Architectures of Convolutional Networks. Accessed: Jul. 5, 2020. [Online]. Available: <https://cv-tricks.com/cnn/understand-resnet-alexnet-vgg-inception/>
- [8] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 4700–4708.
- [9] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng, "CheXNet: Radiologist-level pneumonia detection on chest X-Rays with deep learning," 2017, arXiv:1711.05225. [Online]. Available: <http://arxiv.org/abs/1711.05225>
- [10] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and 0.5 MB model size," 2016, arXiv:1602.07360. [Online]. Available: <http://arxiv.org/abs/1602.07360>
- [11] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.
- [12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, arXiv:1409.1556. [Online]. Available: <http://arxiv.org/abs/1409.1556>
- [13] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 4510–4520.
- [14] Y. LeCun, K. Kavukcuoglu, and C. Farabet, "Convolutional networks and applications in vision," in *Proc. IEEE Int. Symp. Circuits Syst.*, Jun. 2010, pp. 253–256.
- [15] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2015, pp. 234–241.