Digital Image Processing application to predict the TB using CXR images: An case study for optimization of medical images

Pawan Kumar Jatwa[a*], Suraj Yadav[b]
a,b Computer Science Department, Jagannath University, Jaipur (Rajasthan), India

Abstract: TB is a lethal malignant infectious illness in most parts of the world. Image Processing analysis and categorization of chest X-rays (CXRs) into TB and non-TB can be a valid alternative to predict the TB disease. The work proposes an automated TB detection technique employing sophisticated IP (image processing) models. A large section of a CXR picture is black, preventing diagnosis and misleading IP models. Therefore, in the proposed approach, we use advanced segmentation networks to extract the region of interest from multi-modal CXRs. IP models get segmented images to create the model development. Researcher employ explainable machine learning to depict TB-infected lungs for the subjective judgment. In present study various techniques are compare with various techniques to find the best prediction method. ImageJ software was used for the present study. Images uses for the present investigation was collected from public domain like kaggle.

Keywords: Image Processing, Tuberculosis, Segmentation, classification, ImageJ.

I. INTRODUCTION

Patients diagnosed with tuberculosis may appear with cavitation, fibrosis, or nodular infiltrates, or they may have a combination of these three findings. Patients suffering from persistent TB frequently develop cavities in their lungs. It is necessary to employ thoracic x-ray imaging in order to locate this hollow since these pictures include a significant amount of medical information. On the other hand, the benefits of such unfiltered pictures are fairly limited, and it might be rather challenging to locate the lesion. The amount of time required for the examinations is substantial, and the findings are frequently unreliable. Tuberculosis can display a wide variety of abnormalities as seen on x-rays of the chest [**01-09**].

Picture segmentation refers to the process of recognizing and then dividing into discrete portions the various components of an image. Before proceeding to the next stage of processing, geometric analysis, it is common practice to partition the data being processed. It is generally accepted that segmentation, one of the several operations that can be performed on recorded pictures using a computer, is the most significant activity. Despite decades of work, segmentation continues to be a difficult problem in image processing and computer vision [**10**],[**11**].

Because the degree of subdivision that may be applied to an image is inversely proportional to the intricacy of the issues being addressed, the process of image segmentation would cease as soon as the delineation stage was accomplished. Before the image's attributes can be rebuilt, occluded objects must be isolated from one another and then extracted from the image's backdrop [**12**], [**13**].

The approach of feature analysis requires a correct description of the shape or appearance of something, as well as the enhancement of this representation so that it may be utilized for further classification. In several image processing applications, appearance, form, spectral, textural, and contextual characteristics are among the fundamental feature categories that have been digitized. These feature classes are determined by human interpretations of image data. The portrayal of the appearance of medical images is the most discerning of these several types of characteristics, which simplifies the categorizing process [**10**]. The literature review of the base paper was reviewed for the present study. Researchers present an automated technique for diagnosing tuberculosis based on deep learning (DL) models. A CXR picture is mostly black, which provides little diagnostic information and may lead to confusion with DL models. We apply complex segmentation networks to multimodal CXRs to determine the region of interest. DL models are then fed segmented images. We employ explainable AI to depict lung regions affected by tuberculosis for subjective evaluation. Using three publicly available CXR datasets, we evaluate the classification accuracy of several CNN models during our tests. EfficientNetB3 has an accuracy of 99.1%, a receiver operating characteristic of 99.9%, and an average accuracy of 98.7%. This is CNN's most accurate report. The findings of the experiment indicate that segmented

Volume No.-III , Issue No.–II, November, 2022, ISSN: 2582-6263

lung CXR images perform better than raw images [**14**].

Table 1 Tuberculosis (TB) Image Processing Model development using Dataset in some recent works

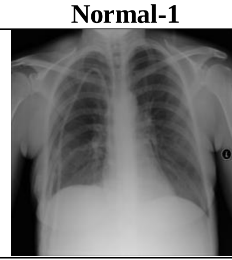
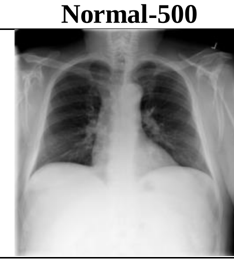
Author Name	**TB Dataset**	**ML/DL Models used by Researchers**	**Remarks**
Nafisah et al., 2022 (Base Paper) [14]	Montgomery, Shenzhen, Belarus dataset, Kaggle datasource	ANN Classifier, CNN,	In this paper researcher found CNN is better model for image processing and classification
Rahman et al., 2020 [15]	Montgomery Shenzhen Belarus Health NIAID TB dataset	Segmented and augmented lung CXRs	Accuracy level in present study was more then 95% which is good prediction of the images for medical industry
Heo et al., 2019 [16]	Montgomery Shenzhen	CNN methods and image processing	Accuracy level in present study was more then 95% which is good prediction of the images for medical industry
Nguyen et al.,, 2019 [17]	Montgomery Shenzhen	CNN methods and image processing	AUC level is more than 99% which is greatest value for image processing
Hwang et al., 2019 [18]	Seoul National Hospital	Deep learning with image processing	AUC level is more than 97% which is greatest value for image processing
Stirenko et al., 2018 [19]	Montgomery -Shenzhen	CNN methods and image processing	Accuracy levels is in range of 85% from dataset
Antani et al., 2018 [20]	From India and Kenya	CNN methods and image processing	Accuracy level is in range of 95% with ADNI data source
Present Study	**Kaggle data set [21]**	**Image Segmentation and Classification using Decision Tree**	**Accuracy level is more than 95%**

II. CXR DATASET

The data collection on CXR imaging that is made freely available by the National Library of Medicine (NLM) is what has been detailed up above. Researchers have been granted free access to this data in the expectation that it may enable them to treat lung-related diseases such as TB more effectively and at an earlier stage. A total of 700 patients, comprised of both tuberculosis (TB) sufferers (700 photographs) and normal individuals, were selected for inclusion in the data collection that is housed inside the NLM Data Set (which is housed within the National Library of Medicine) (3500 images). Users are welcome to go through the picture data and clinical data that are included in this data repository. This dataset contains both CXR images and clinical data from people who have TB as well as normal participants, and the software has access to both types of data. The data that comprise the NLM dataset are broken out in excruciating detail in Table 2, which may be seen below.

Some important datasets available for TB disease image processing work. All these data source are used for the present study to get the final data source which is available in kaggle open source platform [**21**].

Table 2 Normal people CXR images from data set

Normal-1	**Normal-500**
	
Normal-1000	**Normal-1500**

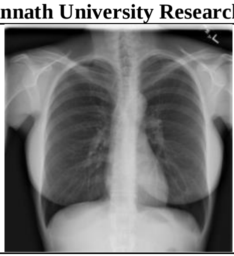
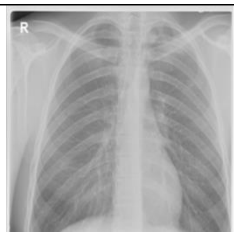
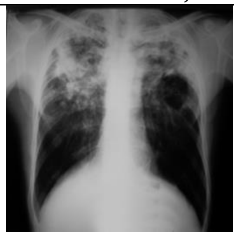
	
Normal-2000	**Normal-2500**
	
Normal-3000	**Normal-3500**
	

Table 3 TB Patients CXR images from data set

Tuberculosis-1	**Tuberculosis-100**
	
Tuberculosis-200	**Tuberculosis-300**
	
Tuberculosis-400	**Tuberculosis-500**
	
Tuberculosis-600	**Tuberculosis-700**

III. RESEARCH METHODOLOGY

For the purpose of this investigation, a dataset was employed that had lung CT pictures acquired from many individuals. These images included both aberrant and normal lung tissue. The patients included in this study are both male and female, and their ages range from 15 to 78 years. Images are taken from them. The final research methodology adopted for the present study was shown in figure 1. in this figure all important segment of the image processing tools are presented. The same model was used for both type of the CXR images provided by the data set, researcher adopted.

IV. RESULT AND DISCUSSION

In this section, the each segment required for the image processing is presented for all possible available tools in Image J software. Both normal and TB patients are used for these sample investigations.

Image Filtration

Image preprocessing refers to the processes that are carried out on images in order to get them ready for use in the model training and inference phases of the data analysis process. Adjustments may be made to the image's orientation, size, and color, among other things, but this is not restricted to just those things. The term "image augmentation" refers to the process of applying various alterations to medical images in order to produce distinct versions of material that is comparable for the purpose of exposing the model to a greater variety of training cases. These distinct versions of material can then be used for the purpose of exposing the model to a wider variety of training cases. For instance, sporadically altering an input image's rotation, brightness, or size requires a model that can evaluate what a specific image topic seems to look like in a variety of settings. This is because the image is being interpreted differently each time.

Jagannath University
Volume No.-III , Issue No.–II, November, 2022, ISSN: 2582-6263

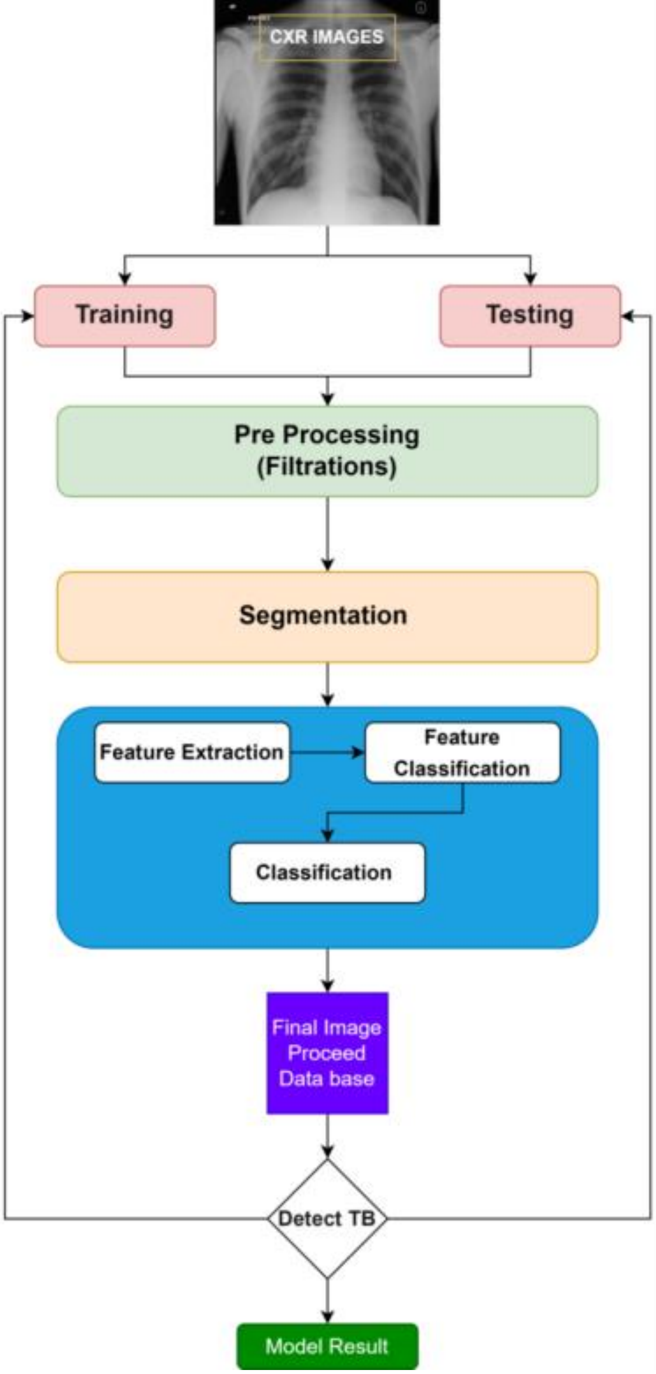


Figure 1 Research methodology required for the present Image Processing research work

Original CXR Images	
Normal-148	TB-65
	
Convolve Filters	

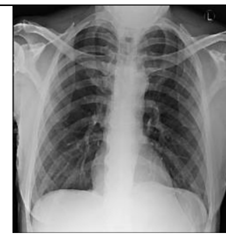
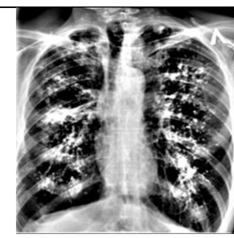
	
Kuwahara Filters	
	
Normalize Local Contract	
	

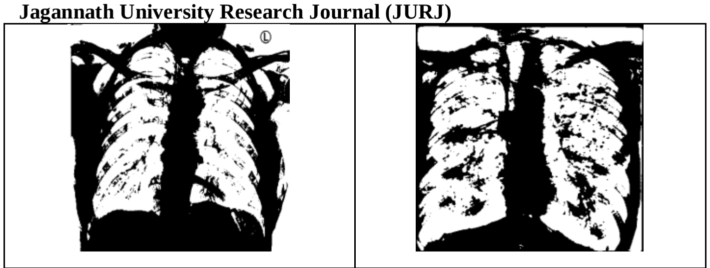
Figure 2 Pre Processing of the CXR images using Filter tools

Image Segmentation

As an alternative to evaluating the full picture, the detector might take as its input a portion of the picture that has been selected by applying a segmentation approach. This helps to save time.

Original CXR Images	
Normal-148	TB-65
	
Threshold Based Segmentation	

Volume No.-III , Issue No.–II, November, 2022, ISSN: 2582-6263



Region-Based Segmentation (Left Side Only)

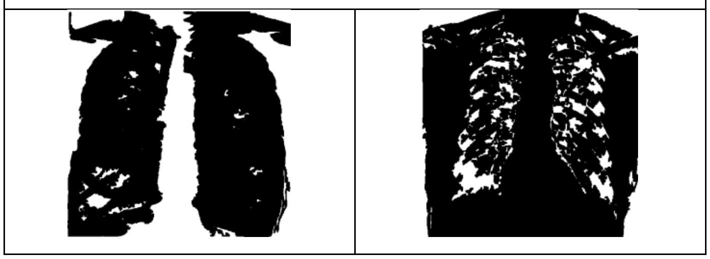


Figure 3 Segmentation of the CXR images

Feature Extraction

The practice of automatically extracting features from signals or photographs without the need for any input from a person by making use of specialized algorithms or deep neural networks is referred to as automated feature extraction. This approach has the potential to be highly efficient if there is a need to make a rapid transition from the raw data to the construction of machine learning algorithms. Wavelet scattering is a technique that may be used to demonstrate the process of automated feature extraction. The simple example of the feature extraction are shown in figure 4.

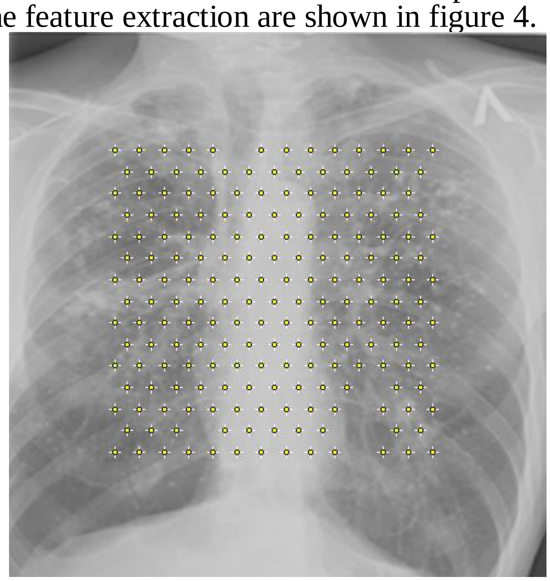


Figure 4 Feature extraction on CXR image (TB-65) using imajeJ software

In figure 5, the feature extraction is combined with classification methods, and the results, which were rather encouraging, are shown.

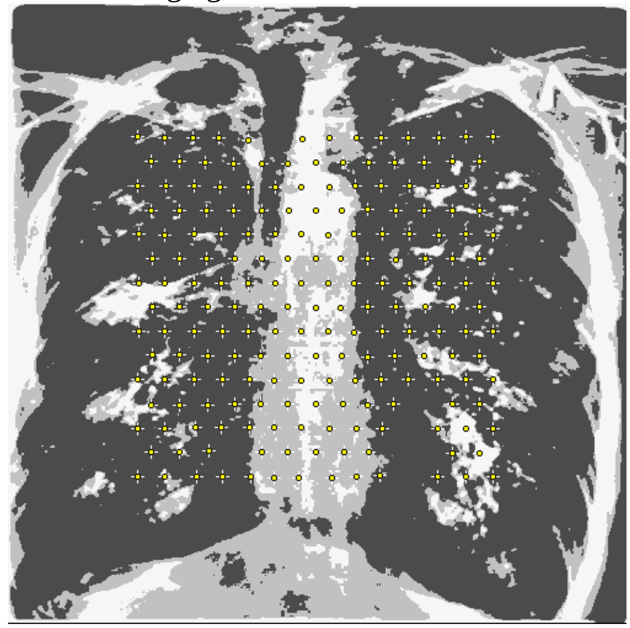


Figure 5 Feature extraction and classification on CXR image (TB-65) using imajeJ software

V. CONCLUSION

In proposed research study, image processing tools are used to find the results for the prediction of the TB in any patient which have TB in moderate stage. In this work classification of the results are performed using ANN (Artificial neural network) method. Remains results are gathered by using image processing techniques, which hare already discussed in previous chapters. The final conclusion of the study have some important conclusions which are following:

It is possible to draw the conclusion, based on the findings obtained from the CXR images of pulmonary tuberculosis, that it is possible to improve the performance of the classifiers by performing a preprocessing stage that enhances contrast and luminance before the segmentation stage. This would come before the stage in which the segmentation would take place.

As can be observed from the results that were obtained, the recommended preprocessing approaches are both well-known and simple to put into practice. They do not need a lengthy amount of preprocessing time or a significant amount of computational work, but they do improve.

REFERENCES

[1]. Lakhani P, Sundaram B (2017) Deep learning at chest radiography: automated classification of Pulmonary Tuberculosis byusing convolutional neural networks. Radiology 284(2):574–582

[2]. Chandra TB et al (2020) Automatic detection of Tuberculosisrelated abnormalities in chest Xray images using hierarchicalfeature extraction scheme. Expert Syst Appl 158(15):113514

[3]. Liu et al., (2017), ''TX-CNN: Detecting tuberculosis in chestX-ray images using convolutional neural network. 2017 IEEEinternational conference on image processing (ICIP), Beijing,pp. 2314–2318

[4]. Hunter, R. L. (2011). Pathology of post primary tuberculosis of the lung: an illustrated critical review. Tuberculosis, 91(6), 497-509.

[5]. Li, T.-T., Yan, X., Zhou, T., & Yu, Z.-X. (2018). A Challenge: Pulmonary Sclerosing Haemangioma. Chinese Medical Journal, 131(19), 2390–2391. Doi: 10.4103/0366-6999.239689.

[6]. Handa, A., Dhooria, S., Sehgal, I. S., & Agarwal, R. (2018). Primary cavitary sarcoidosis: A case report, systematic review, and proposal of new diagnostic criteria. Lung India: Official Organ of Indian Chest Society, 35(1), 41–46.

[7]. Petrella, F., Casiraghi, M., Prisciandaro, E., Gherzi, L., & Spaggiari, L. (2019). Incidental diagnosis of pulmonary mycobacteriosis among patients scheduled for lung cancer surgery: results from a series of 3224 consecutive operations. Heliyon, 5(3), e01395. Doi: 10.1016/j.heliyon.2019.e01395.

[8]. Hellbach, K., Baehr, A., De Marco, F., Willer, K., Gromann, L. B., Herzen, J., ... & Rummeny, E. J. (2018). Depiction of pneumothoraces in a large animal model using x-ray dark-field radiography. Scientific reports, 8(1), 1-8. Doi: 10.1038/s41598-018-20985-y.

[9]. Khlifa, I., Vabre, A., Hočevar, M., Fezzaa, K., Fuzier, S., Roussette, O., &Coutier-Delgosha, O. (2017). Fast X-ray imaging of cavitating flows. Experiments in Fluids, 58(157), 3-22. Doi: 10.1007/s00348-017-2426-7.

[10]. Udupa, J.K., LeBlanc, V.R., Zhuge, Y., Imielinska, C., Schmidt, H., Currie, L.M, Hirsch B.E. and Woodburn, J. "A framework for evaluating image segmentation algorithms", Comput. Med. Imaging Graph., Vol.30, No.2, pp.75-87, 2006.

[11]. Du, X. and Dua, S. "Segmentation of fluorescence microscopy cell images using unsupervised mining", Open Med. Inform. J., Vol.4, pp.41-49, 2010.

[12]. Wahlby, C., Lindblad, J., Vondrus, M., Bengtsson, E. and Bjorkesten, L. "Algorithms for cytoplasm segmentation of fluorescence labelled cells", Anal. Cell. Pathol., Vol.24, No.2-3, pp.101-111, 2002.

[13]. Khare, M. and Srivastava, R.K. "Level set method for segmentation of medical images without re-initialization", J. Med. Imaging Health Inf., Vol.2, No.2, pp.158-167, 2012.

[14]. Nafisah, S.I., Muhammad, G. Tuberculosis detection in chest radiograph using convolutional neural network architecture and explainable artificial intelligence. Neural Comput & Applic (2022). https://doi.org/10.1007/s00521-022-07258-6

[15]. Rahman T et al (2020) Reliable Tuberculosis detection using chest X-Ray with deep learning, segmentation and visualization. IEEE Access 8:191586–191601

[16]. Heo SJ, Kim Y, Yun S, Lim SS, Kim J, Nam CM, Park EC, Jung I, Yoon JH (2019) Deep learning algorithms with demographic information Help to detect Tuberculosis in chest radiographs inannual workers' health examination data. Int J Environ Res Public Health 16:250

[17]. Nguyen et al., (2019) Deep learning models for Tuberculosisdetection from chest X-ray images. 26th International Conference on Telecommunications (ICT), Hanoi, Vietnam, pp. 381–385

[18]. Hwang EJ et al (2019) Development and validation of a deep learning–based automatic detection algorithm for active pulmonary tuberculosis on chest radiographs. Clin Infect Dis 69(5):739–747

[19]. Stirenko et al., Chest X-Ray analysis of Tuberculosis by deep learning with segmentation and augmentation. In: 2018 IEEE 38th international conference on electronics and nanotechnology (ELNANO), Kiev, 2018, pp. 422–428

[20]. Liz H, Sa´nchez-Montan˜e´s M, Tagarro A, Domı´nguez-Rodrı´guez S, Dagan R, Camacho D, Ensembles of Convolutional Neural Network models for pediatric pneumonia diagnosis,'' arXiv:2010.02007 [eess.IV].

[21]. Tawsifur Rahman, Amith Khandakar, Muhammad A. Kadir, Khandaker R. Islam, Khandaker F. Islam, Zaid B. Mahbub, Mohamed Arselene Ayari, Muhammad E. H. Chowdhury. (2020) "Reliable Tuberculosis Detection using Chest X-ray with Deep Learning, Segmentation and Visualization". IEEE Access, Vol. 8, pp 191586 - 191601. DOI. 10.1109/ACCESS.2020.3031384.

AUTHOR'S BIOGRAPHIES

Pawan Kumar Jatwa has engineering graduate from Computer Scienece engineering from RTU, Kota. He is doing MTECH in computer science from Jagannath University, Jaipur (Rajasthan). his area of interest are machine learning and Application of AI in medical science.

Suraj Yadav has MTECH in Computer Science Engineering from RTU, Kota. Currently he is working with Jagan Nath University, Jaipur, as Assistant Professor in Computer Science Department. His area of interest is machine learning, AI, Optimization methods etc.