Resource-Efficient Tuberculosis Screening Via Deep Cnns And Machine Learning

[1*]Pachipenta Mohan, [2]Ch. Kodandaramu, [3]A Srinivasa Babu
Student in Dept. Of Master of Computer Applications, Miracle Educational Society Group of Institutions
Associate Professor, Miracle Educational Society Group of Institutions
Associate Professor, Miracle Educational Society Group of Institutions
*pachipentamohan2001@gmail.com

ABSTRACT
The low-resource settings is one of the areas where tuberculosis is one of the most critical public health issues, it tends to cause more issues when there is a lack of diagnosis. The current project demonstrates the use of a hybrid model integrating deep convolutional neural networks (CNNs) and additional Machine Learning algorithms to detect Tuberculosis from radiographs. Along with InceptionResNetV2, Xception, and DenseNet201, we perform ensemble learning on three pretrained CNNs to classificate ML models like SVM, KNN, and Bagging classifiers, on the extracted and feature volunteered images. For accurate and low-relouse prognosis, our proposed system is trained on COVID and TB chest X-ray datasets, allowing it to differentiate between TB, COVID-19, and normal cases. This new testing methodology was able to achieve a whooping 98% in accuracy which goes to show how useful it could be for diagnosing TB in areas where medical professionals are not available. The new system demonstrates it's potential for quick and reliable TB screening.

Keywords: Tuberculosis Detection, Chest X-ray, Deep Learning, Ensemble Learning, Low-Resource Settings

INTRODUCTION:

Tuberculosis (TB), is indeed one of the worldsmost deadliest infections disease but touchwood Tuberculosis (TB) is something is a disease that is less common in countries with. In 2019 comes the resoucerestricmentale almost 10 million people contracted TB globally, with the majority for thierlow and middle income region And Mycobacterium tuberculosis is one of the gobal lacking health care system. Accurate and early diagnosis a would is the best possible shot for getting rid of it, and that fulfilling these needs they needed becouse humans require does their work using black magic when ai is not there, and looking at another humans xray is not possible without constant training, looks like a chill job for Ai.. With the COVID-19 pandemic worsening TB detection rates due to overlapping symptoms and strained healthcare systems, the need for automated, scalable solutions has grown urgent. Deep learning has emerged as a powerful tool in medical imaging, capable of learning complex patterns directly from data. This project proposes an ensemble of deep learning and machine learning models to diagnose TB from chest X-rays, providing an accurate, accessible solution that reduces dependency on manual diagnostics and offers a practical approach for resource-constrained environments.

GAP IDENTIFIED BASED ON LITERATURE SURVEY:

While previous research has made significant strides in TB detection using AI, several key gaps remain. Most existing systems rely solely on either deep learning or conventional machine learning, limiting flexibility and model performance. For instance, studies have shown that deep learning models perform well in feature extraction, but they require large datasets and high computational resources. Conversely, machine learning models are lightweight but depend heavily on hand-crafted features, which may not capture the full complexity of medical imagery.

Additionally, many prior works lack generalizability across datasets with varying image qualities and limited annotation. In resource-poor regions, where access to clean, labeled datasets is minimal, such limitations restrict practical deployment.

Another issue is the lack of ensembling strategies that combine the strengths of both model types. This project addresses these gaps by:

- Using an ensemble of three efficient deep CNNs for robust feature extraction;
- Applying machine learning classifiers to extracted features for resource-friendly classification;
- Training and evaluating the models on diverse TB, COVID-19, and normal image datasets;
- Focusing on performance metrics like accuracy, precision, and recall to ensure practical effectiveness.

PROBLEM STATEMENT:

Timely and accurate TB detection from chest radiographs is challenging in low-resource settings due to the need for expert radiologists and limited infrastructure.

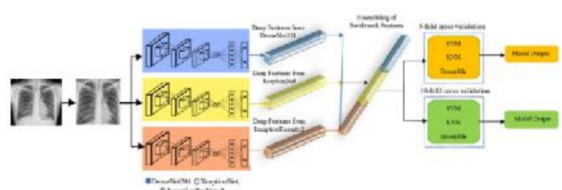
Key Challenges:

- Lack of trained medical professionals to interpret X-rays in rural areas.
- High computational demand of deep learning models for real-time deployment.
- Poor image quality and variability in public chest X-ray datasets.
- Overlapping symptoms between TB and COVID-19 complicating diagnosis.
- Dependence on large, labeled datasets for deep learning model training.
- Limited interpretability of AI-based decisions in clinical environments.

PROPOSED METHOD:

The proposed system ensembles three pretrained CNN models—InceptionResNetV2, Xception, and DenseNet201 for effective feature extraction from chest X-rays. These networks are trained on a dataset containing Normal, COVID-19, and TB images. The extracted deep features are then fed into three lightweight machine learning classifiers—SVM, KNN, and Bagging classifiers—for final classification. This hybrid approach improves detection accuracy while reducing computational load. The system is designed to function efficiently on minimal hardware and can aid early TB detection even without expert radiologists. The model's performance is evaluated based on metrics such as accuracy, recall, and F1-score, achieving up to 98% accuracy in identifying TB cases.

ARCHITECTURE:



DATASET:

The dataset used is made up of annotated chest radiographs that have been classified into three groups: Normal, COVID-19 positive, and Tuberculosis (TB). It integrates images from public repositories like the Shenzhen dataset and the COVID-

19 radiographic datasets. Images are resized and pre-processed to enhance model quality and mitigate discrepancies across models. ROI extraction alongside augmentation methods are performed to aid in improving model learning. Every image is normalized, meaning every pixel's value is adjusted to fit a scale between 0 and 1. The dataset is split into training and testing subsets in an 80% to 20% ratio, respectively, to assure that the model's accuracy evaluation is unbiased and covers every facet of its functionality.

METHODOLOGY:

Image Acquisition and Pre-processing:

- Chest X-ray images are collected from public datasets (Shenzhen, COVID-19 databases).
- Images are resized to a uniform resolution (80x80 pixels).
- Noise reduction and normalization are applied to enhance feature clarity.
- Region of Interest (ROI) is extracted to focus on the lungs.

Dataset Preparation:

- Labels for TB, COVID, and Normal classes are assigned based on directory structure.
- Images are converted to arrays and stored as .npy files for efficient processing.
- The dataset is shuffled and split into training (80%) and testing (20%) sets.

Deep Feature Extraction Using CNNs:

- **InceptionResNetV2**, **DenseNet201**, and **Xception** models are loaded with pretrained ImageNet weights.
- Each CNN model's layers are frozen to prevent retraining, reducing computation.
- Additional convolutional and pooling layers are added to tailor models to TB classification.
- Features from the penultimate layers of each CNN are extracted for downstream processing.

Model Training:

- Each CNN is trained individually on the training dataset using softmax classification.
- Accuracy, loss, and validation performance are monitored across epochs (typically 20).
- The best-performing weights are saved using model checkpoints.

Feature Concatenation and Ensemble Learning:

- Extracted features from each CNN are concatenated into a single feature vector.
- This vector is then used as input to various machine learning classifiers:
  - **Support Vector Machine (SVM)**
  - **K-Nearest Neighbors (KNN)**
  - **Bagging Classifier**
- The classifiers are trained and evaluated for accuracy, precision, recall, F1-score, and AUC.

Model Evaluation and Confusion Matrix Analysis:

- Predictions are compared against test labels to form a confusion matrix.
- Heatmaps are generated to visualize classifier performance on each class.
- ROC-AUC curves and precision-recall plots provide additional evaluation metrics.

Performance Comparison:

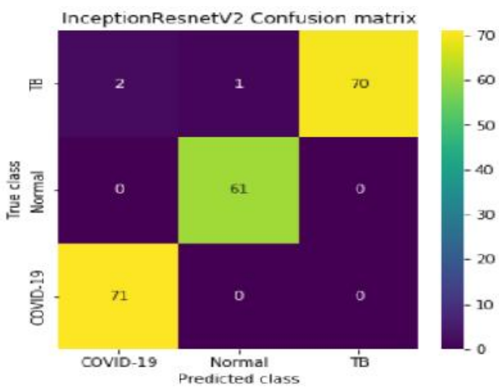
- All models are compared based on their classification metrics.

- The ensemble of CNNs + ML classifiers consistently achieves superior accuracy (~98%).
- Precision and recall values confirm the model's reliability for sensitive TB detection.

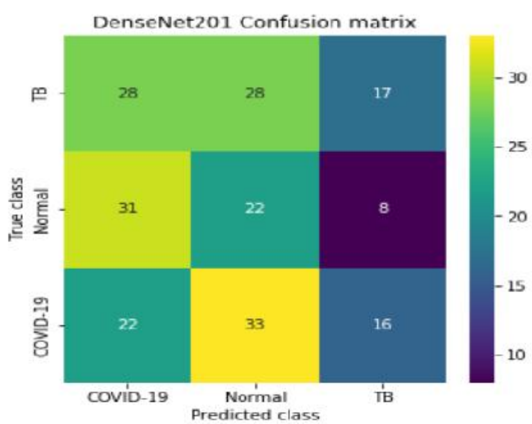
System Deployment Potential:

- Due to its lightweight classification phase and reliance on pre-trained models, the system is deployable in real-world low-resource environments.
- Potential integration with mobile health apps and edge computing for remote diagnostics is discussed.

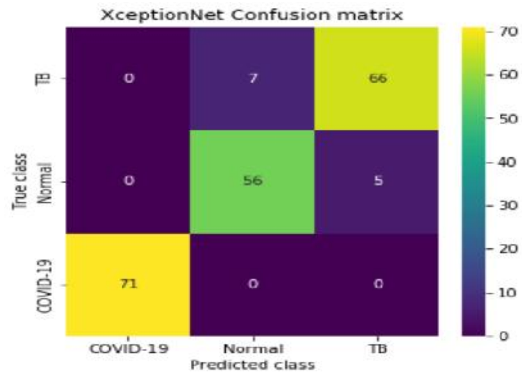
RESULTS:



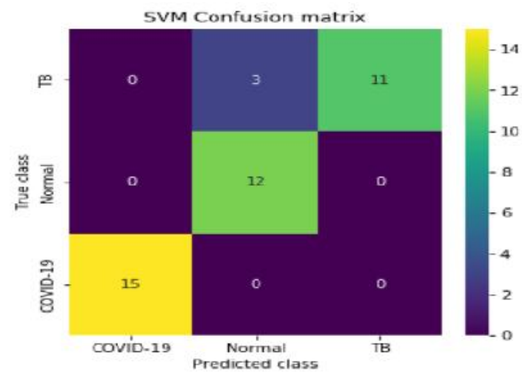
InceptionResnetV2 Accuracy : 98.53658536585365



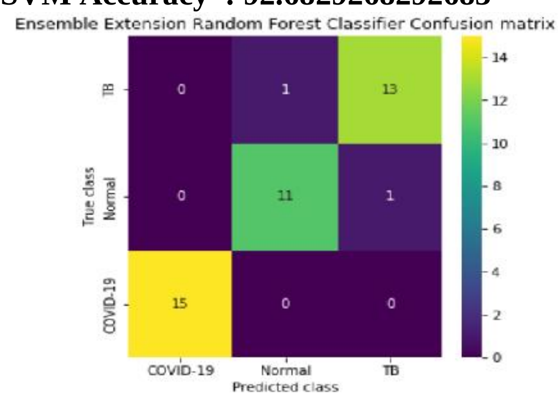
DenseNet201 Accuracy : 29.756097560975608



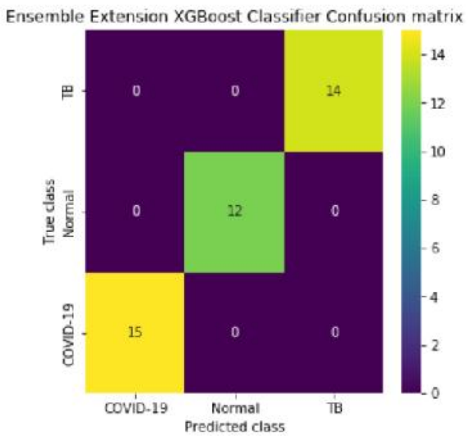
XceptionNetAccuracy : 94.14634146341463



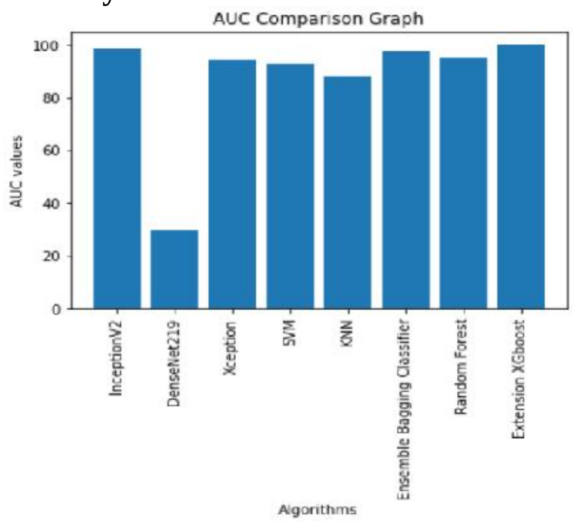
SVM Accuracy : 92.6829268292683



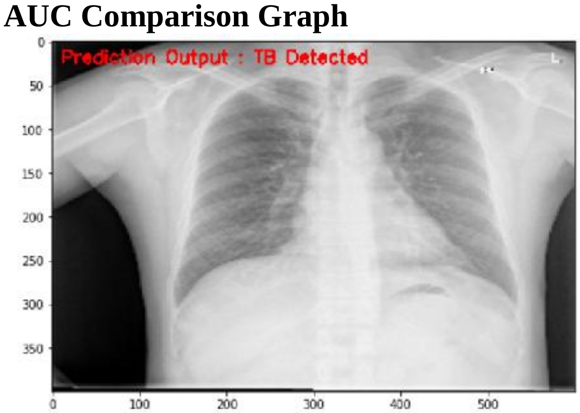
Ensemble Extension Random Forest Classifier Accuracy : 95.1219512195122

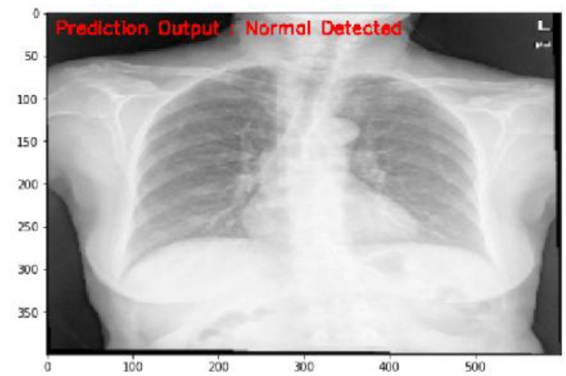


Ensemble Extension XGBoost Classifier Accuracy : 100.0



AUC Comparison Graph





CONCLUSION

With an aim to detect tuberculosis from chest X-ray images, this project built a hybrid AI-based system and it works successfully. The developed model optimally balances the use of deep convolutional networks alongside machine learning classifiers, enabling a resource-friendly system without compromising on accuracy. The ensemble approach serves to improve feature extraction while also enabling the system to robustly classify TB, even with COVID-19 symptoms overlapping the cases. It can easily deployed in healthcare settings that have restricted infrastructural facilities owing to the lightweight systems design. Enhancing interpretability for clinical trust, spanning wider datasets, and real-time mobile deployment are prospects for future enhancements. Overall, this project contributes a valuable tool in the fight against TB in underserved regions.

REFERENCES:

[1] (2020). Global Tuberculosis Report 2020. Accessed: May 10, 2021. [Online]. Available: http://apps.who.int/bookorders

[2] L. M. Parsons, Á. Somoskövi, C. Gutierrez, E. Lee, C. N. Paramasivan, A. L. Abimiku, S. Spector, G. Roscigno, and J. Nkengasong, ''Diagnostic standards and

classification of tuberculosis in adults and children,'' Amer. J. Respiratory Crit. Care Med., vol. 161, no. 4, pp. 1376–1395, 2000, doi: 10.1164/ajrccm.161.4.16141.

[3] Country Profiles for 30 high tb burden countries 20 High TB Burden Countries Based on Absolute Number of Incident Cases 10 High TB Burden Countries Based on Severity of Disease Burden (Incidence per Capita). Accessed: May 10, 2021. [Online]. Available: https://www.who.int/tb/data

[4] L. M. Parsons, Á. Somoskövi, C. Gutierrez, E. Lee, C. N. Paramasivan, A. Abimiku, S. Spector, G. Roscigno, and J. Nkengasong, ''Laboratory diagnosis of tuberculosis in resource-poor countries: Challenges and opportunities,'' Clin. Microbiol. Rev., vol. 24, no. 2, pp. 314–350, 2011, doi: 10.1128/Cmr.00059-10.

[5] M. Roy and S. Ellis, ''Radiological diagnosis and follow-up of pulmonary tuberculosis,'' Postgraduate Med. J., vol. 86, no. 1021, pp. 663–674, Nov. 2010, doi: 10.1136/pgmj.2009.084418.

[6] A. H. van't Hoog, H. K. Meme, H. van Deutekom, A. M. Mithika, C. Olunga, F. Onyino, M. W. Borgdorff, ''High sensitivity of chest radiograph reading by clinical officers in a tuberculosis prevalence survey,'' Int. J. Tuberculosis Lung Disease, vol. 15, no. 10, pp. 1308–1314, 2011, doi: 10.5588/ijtld.11.0004.

[7] R. Hooda, A. Mittal, and S. Sofat, ''Tuberculosis detection from chest radiographs: A comprehensive survey on computer-aided diagnosis techniques,'' Current Med. Imag., vol. 14, no. 4, pp. 506–520, 2017, doi: 10.2174/1573405613666171115154119.

[8] Coronavirus Disease (COVID-19): Tuberculosis. Accessed: May 10, 2021. [Online]. Available: https://www.who.int/news-room/q-adetail/coronavirus-disease-covid-19-tuberculosis.

[9] P. Glaziou, ''Predicted impact of the COVID-19 pandemic on global tuberculosis deaths in 2020,'' MedRxiv, Oct. 2021, doi: 10.1101/2020.04.28.20079582.

[10] Dashboard: Nikshay Reports. Accessed: May 12, 2021. [Online]. Available: https://reports.nikshay.in/Reports/TBNotification

[11] A. S. Becker and C. Blüthgen, ''Detection of tuberculosis patterns in digital photographs of chest X-ray images using deep learning: Feasibility study,'' Int. J. Tuberculosis Lung Disease, vol. 22, no. 3, pp. 328–335, 2018.

[12] P. Lakhani and B. Sundaram, ''Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks,'' Radiology, vol. 284, no. 2, pp. 574–582, 2017. [Online]. Available: https://doi.org/10.1148/radiol.2017162326

[13] Y. Cao, C. Liu, B. Liu, M. J. Brunette, N. Zhang, T. Sun, P. Zhang, J. Peinado, E. S. Garavito, L. L. Garcia, and W. H. Curioso, ''Improving tuberculosis diagnostics using deep learning and mobile health technologies among resource-poor and marginalized communities,'' in Proc. IEEE 1st Int. Conf. Connected Health, Appl., Syst. Eng. Technol. (CHASE). Washington, DC, USA: IEEE, Jun. 2016, pp. 274–281, doi: 10.1109/CHASE.2016.18.

[14] R. Singh, M. K. Kalra, C. Nitiwarangkul, J. A. Patti, F. Homayounieh, A. Padole, P. Rao, P. Putha, V. V. Muse, A. Sharma, and S. R. Digumarthy, ''Deep learning in chest radiography: Detection of findings and presence of change,'' PLoS ONE, vol. 13, no. 10, Oct. 2018, Art. no. e0204155, doi: 10.1371/journal.pone.0204155.

[15] O. Yadav, K. Passi, and C. K. Jain, ''Using deep learning to classify X-ray images of potential tuberculosis patients,'' in Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM), Dec. 2018, pp. 2368–2375.

[16] C. Liu, Y. Cao, M. Alcantara, B. Liu, M. Brunette, J. Peinado, and W. Curioso, ''TX-CNN: Detecting tuberculosis in chest X-ray images using convolutional neural network,'' in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2017, pp. 2314–2318.

[17] T. Karnkawinpong and Y. Limpiyakorn, ''Chest X-ray analysis of tuberculosis by convolutional neural networks with affine transforms,'' in Proc. 2nd Int. Conf. Comput. Sci. Artif. Intell., 2018, pp. 90–93.

[18] S. Sathitratanacheewin, P. Sunanta, and K. Pongpirul, ''Deep learning for automated classification of tuberculosis-related chest X-ray: Dataset distribution shift limits diagnostic performance generalizability,'' Heliyon, vol. 6, no. 8, Aug. 2020, Art. no. e04614.

[19] S. Hwang, H. E. Kim, J. Jeong, S. Hwang, and H.-J. Kim, ''A novel approach for tuberculosis screening based on deep convolutional neural networks,'' Proc. SPIE, vol. 9785, Mar. 2016, Art. no. 97852W, doi: 10.1117/12.2216198.

[20] R. Mehrotra, M. A. Ansari, R. Agrawal, and R. S. Anand, ''A transfer learning approach for AI-based classification of brain tumors,'' Mach. Learn. Appl., vol. 2, Dec. 2020, Art. no. 100003, doi: 10.1016/j.mlwa.2020.100003.