A Comparison of 2D and 3D CNN for Lung CT Image Tuberculosis Severity Assessment

David Olayemi Alebiosu[1,*], Farrukh Hassan[1], Adeola Folayan[2], and Mathew Yit Hang Yeow[1]

[1]Department of Computing and Information Systems, Sunway University, Bandar Sunway, Malaysia
[2]Jeffrey Cheah School of Medicine and Health Sciences, Monash University, Malaysia
Email: davida@sunway.edu.my (D.A.); farrukh@sunway.edu.my (F.H.H.); adeola.folayan@monash.edu (A.F.); matthewyyh@sunway.edu.my (M.Y.H.Y.)
*Corresponding author

Manuscript received July 29, 2024; revised September 9, 2024; accepted December 3, 2024; published February 14, 2025

***Abstract*—Tuberculosis is an infectious disease that usually affects the lungs. However, early diagnosis of tuberculosis increases the chance of cure. Practical analysis of lung Computed Tomography (CT) images from tuberculosis patients is one of the primary methods used to determine the severity of the disease. Handcrafted CT image analysis techniques such as grey level concurrence matrix, Fourier transform, etc, used for medical image pre-processing techniques have been ineffective due to their limitations in extracting discriminating features from the images. The application of deep learning, a branch of machine learning, is gaining increased acceptance in medical image analysis. The challenges such as high cost, human error, and slow speed encountered during manual labelling are gradually eliminated in various scales with deep learning techniques. This study explores two deep-learning approaches to classify TB severity in Lung CT Images. Two-Dimensional (2D) and three-Dimensional (3D) convolutional neural networks (CNNs) were used separately to classify the ImageCLEF 2021 lung CT dataset into 'High' and 'Low' severity categories. The proposed 3D-CNN in this study outperformed the 2D CNNs; it produced an overall average accuracy and Area Under the ROC curve (AUC) of 0.9929 and 0.9982 respectively.**

***Keywords*—tuberculosis, lung CT, classification**

I. INTRODUCTION

The 2022 World Health Organisation (WHO) global tuberculosis (TB) report stated an estimated TB incidence of 10.6 million cases in 2021, making TB a global health challenge [1]. Also, with the increase in the reports of Multi-Drug-Resistant TB (MDR-TB), the havoc of TB on humanity is increasingly becoming an unbearable Challenge. The treatment of TB in the last 50 years has centred on the use of a short-course chemotherapy regimen. This technique has achieved little result in treating TB [2].

The goal of the WHO remains focused on achieving a timely and effective diagnosis of TB, but the current technologies have limited support to actualize the WHO goal.

The current technologies are costly and complicated and display poor sensitivity. Hence, there is an urgent need to develop modern diagnostic techniques for treating TB to meet the WHO goal. In recent years, deep learning [3] techniques have been applied in medical image processing for various tasks ranging from disease detection to segmentation and image classification. Convolutional Neural Networks (CNN) [4] have achieved outstanding results in medical-related applications. Deep learning techniques have been successfully applied with the availability of ImageNet pre-trained networks on hundreds of thousands of images for medical image analysis [5]. Researchers have employed various deep-learning approaches to classify medical CT images for TB severity assessment [6, 7]. Some Studies have decomposed 3D volumes of interest into 2D views and augmented the images before training CNN architectures on the image. Other studies have combined 2D CNN to analyze 3D medical images [8].

This study proposes a unique 14-layer 3D CNN model and compares its performance with six popularly used pre-trained CNN models to classify TB-infected lung CT images. The images were classified into 'High' and 'Low' severity categories. The six pre-trained neural networks employed include VGGNet-16 [9], VGGNet-19 [10], DenseNet 121 [11], GoogleNet [12], ResNet-50 [13] and Inception V4 [14]. Both the proposed 3D CNN and the pre-trained neural networks were separately used as feature extraction models used to extract deep discriminative image features from the dataset. The extracted image feature vectors were later used to classify the testing images into two severity categories through the Softmax classification layer.

We evaluated both the 2D pre-trained CNN and our proposed 3D CNN model on the ImageCLEF 2021TB lung CT image dataset [15]. The evaluation results from the classification model were analyzed and compared with some recently employed state-of-the-art 3D CNN models used on similar datasets.

II. LITERATURE REVIEW

For over a decade, several studies have employed 2D and 3D CNNs for medical image analysis. The architecture of CNN models was used, initializing the training with random weights. Some of the studies that used pre-trained 2D CNN for medical image disease analysis, such as detection, segmentation and classification, include [16–19], and [20]. Others include the automatic classification of pulmonary nodules in computed tomography [21], federated learning approach with pre-trained deep learning models for COVID-19 detection [22], early prediction of lung cancer using pre-trained neural networks [23], TB diagnosis using CNN [24], and automatic analysis of active and non-active TB using a pre-trained neural network [25].

Other studies employed pre-trained CNN as a transfer learning mechanism on a new 3D CNN. Some of these studies include improving TB severity assessment in CT Images [7], lung CT automatic identification in CT images [24], and detection of tuberculosis diseases using pre-trained CNN [26].

However, despite the outstanding results reported in many of these works, the use of 2D pre-trained neural networks has been limited. This is because lung CT images are 3D in structure, and analyzing them with 2D CNN leads to the potential loss of spatial contextual information. In most cases, the third dimension of 3D images is often ignored by the 2D CNN during feature extractions [27]. Table 1 shows the summary of some of the recently employed CNN models for medical image analysis.

This study alleviates the limitation of 2D pre-trained neural networks in lung 3D CT scan images for TB severity assessment by exploring the potential of a 3D CNN model. Our proposed 3D CNN model provides a more robust technique for analyzing lung CT scan diseases such as TB.

Table 1. Some state-of-the-art-3D-CNN Models used for different medical image modalities

Author	Modality	ACC
[16]	X-ray	0.9464
[7]	CT	0.8119
[18]	CT	0.8671
[20]	CT	0.9649
[21]	CT	0.9534
[28]	X-ray	0.090
[24]	CT/X-ray	0.9447
[17]	X-ray	0.9720
[29]	CT	98.25

III. MATERIALS

A. ImageCLEF 2021 Lung CT Dataset

The ImageCLEF 2021 CT datasets [30] comprise 1,338 TB patients. The training images are 917, while the testing images are 421. Each CT image is captured from one unique patient and corresponds to only one TB type. As obtained in the ImageCLEF2019 TB challenge, additional meta-information containing the CT report was provided. The entire dataset was used to evaluate the proposed models in this study. The testing dataset was classified into two severity categories: the 'HIGH' and 'LOW' categories. From five discrete values of 1 to 5, the values 1, 2, and 3 belong to 'HIGH', and 4 and 5 belong to the 'LOW' severity category. Fig. 1 shows examples of lung TB CT slices in ImageCLEF 2021 dataset.

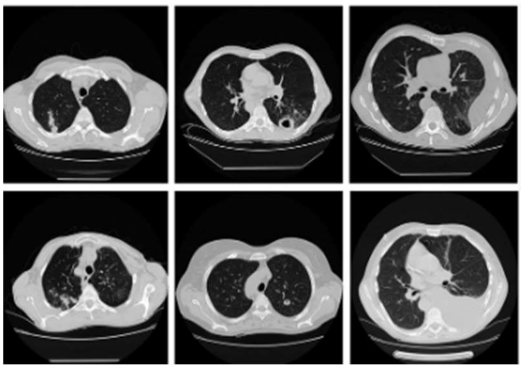


Fig. 1. Examples of TB lung CT slices in the ImageCLEF 2021 dataset.

B. Data Augmentation

Four different random transformations were employed to augment the training datasets. The random transformation includes rotation (90 to 90), shear (−30 to 30), translation (−15 to 15), and cropping. After the augmentation, the training dataset increased to four times the original size. In total, the training dataset increased to 3,668 images. Fig. 2 shows a TB-infected lung CT scan before and after augmentation.

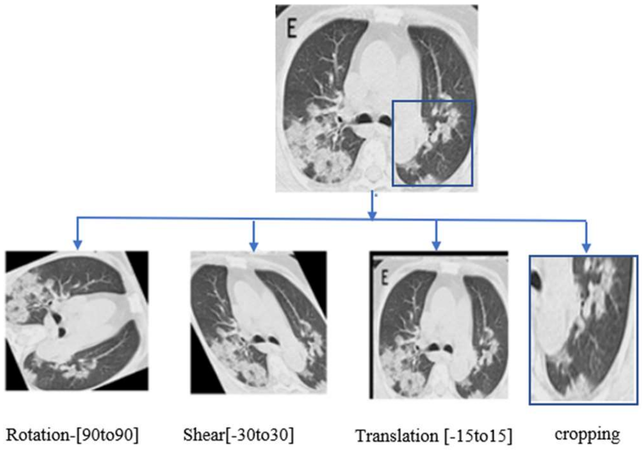


Fig. 2. A TB-infected lung CT scan sample undergoing four different random transformations.

IV. METHODS

The methodology for this study consists of the 2D pre-trained CNNs and our proposed 3D CNN. First, each component of the framework, which is the pre-trained 2D CNNs and the proposed 3D CNN model, is explained. Second, the components of our proposed 3D CNN are described in detail. All six pre-trained neural networks consist of 1000 classes, trained on 1.28 million images and 100,000 testing images evaluated on 50,000 validation images. Each of the CNNs was used separately to extract the image features. The extracted features were later passed into the Softmax classification layer to be classified into High and low-severity classes. Fig. 3 depicts the general architecture of the 2D CNN Model.

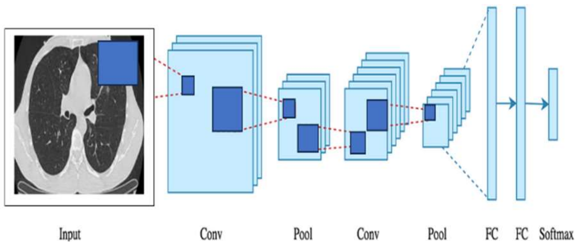


Fig. 3. The general architecture of the 2D CNN model. It shows a lung CT scan as input to the model.

A. VGGNet-16

VGGNet-16 [9] was first mentioned by two researchers at the University of Oxford and achieved 97% (top 5 test) accuracy in the ImageNet challenge. ImageNet consists of over 14 million images distributed across 1000 classes. Compared with AlexNet [31], which uses an 11×11 filter size, VGGNet-16 uses a smaller filter of 3×3 size. The smaller filter size helps reduce the network parameters, which improves the efficiency of the model

B. VGGNet-19

VGGNet-19 [10] is a deeper architecture compared with VGGNet-16 and has more weight. It comprised 19 deep trainable layers for convolution. The 19 trainable layers are fully connected with max pooling and dropout layers. It has fully connected nodes of size 574 M, whereas VGGNet-16 has fewer fully connected nodes of size 533 M.

C. DenseNet 101

DenseNet 121 [11] has been employed by various researchers for image classification and has achieved state-of-the-art results on different datasets. DenseNet family is known for its ability to reuse features from deep CNN architectures. DenseNet uses its feature maps as input to the next one while treating the feature maps of all preceding layers as independent input for each layer.

D. GoogLeNet

The GoogLeNet architecture [12] has a deeper architecture compared with AlexNet [31]; It comprises 1×1 convolution and global average pooling. The global average pooling is used at the end of the network, and it takes a feature map of a dimension of 7×7 and averages it to 1×1. This helps to decrease the number of parameters. GoogleNet has an overall architecture of 22 layers and two auxiliary classifier layers.

E. ResNet-50

ResNet-50 was developed to allow for the training of very deep networks consisting of hundreds of layers. Resnet-50 consists of 50 layers divided into five blocks. Each of the 50 layers consists of residual blocks that allow information preservation from previous layers. The residual blocks help the networks learn better input data representations. Aside from the convolutional layer, residual and fully connected layers are present in Res-Net-50. They also consist of skip connections constructed by adding a previous layer's output to the later layer's output.

The six pre-trained 2D CNNs were used first to extract the deep image features. The extracted feature vector is passed into the classification layers to classify the testing dataset into two categories. Fig. 4 shows the architecture of the pre-trained 2D CNNs classification process.

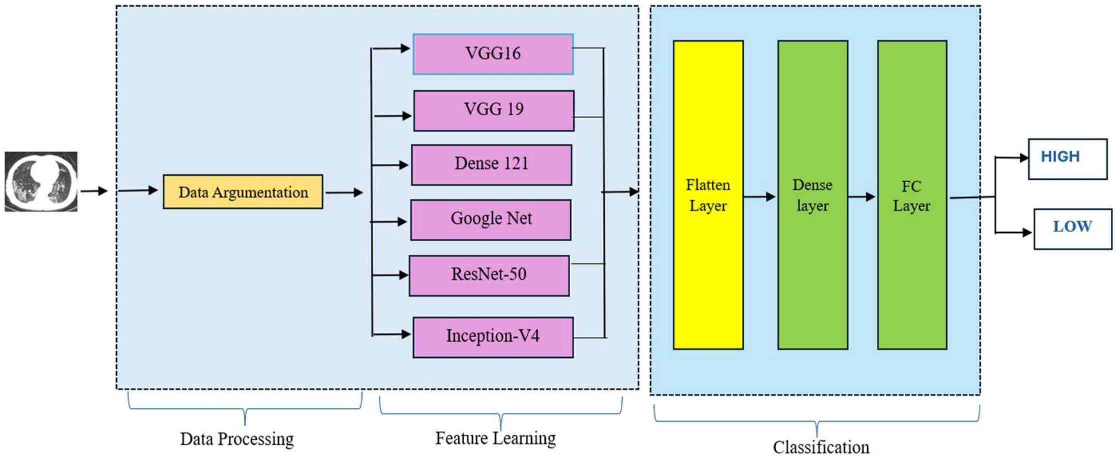


Fig. 4. The architecture of the pre-trained 2D CNNs, showing the input lung CT, the feature learning and the classification layers.

F. Inception V4

Inception V4 [14] is a deep convolutional network, the fourth iteration of the Inception family of networks. The Inception V4 efficiently combines various features from the previous Inception architectures. Unlike the previous Inception models, Inception-v4 is without residual connections. It has a more uniform, simplified architecture when compared with Inception V3. Fig. 5 shows the overall architecture of Inception V4.

G. Our Proposed 3D CNN

The proposed 3D CNN for this study is a 14-layer 3D CNN, which comprises three 3D convolutional (CONV) layers. The first layer consists of 64 filters, the second layer consists of 128 filters and the third layer with 256 filters. It has a kernel size of 3×3×3. Each CONV layer is followed by a max-pooling (MAXPOOL) layer, with a stride of 2 and ReLU activation. The model ends with a batch normalization (BN) layer. Our 3D CNN model consists of three CONV-MAXPOOL-BN modules. The output from the feature extraction block is passed into a 512-neuron fully connected layer. Two dropout layers were used, having an effective dropout rate of 60%.

The SoftMax activation function [32] with two neurons served as the last activation function, used to classify the dataset into two classes, which are the 'High' and 'Low' severities. To avoid too many parameters and overfitting [33] of our 3D CNN model, we keep the network relatively simple, with only 8,397,992 learnable parameters. The same classification process applied to the pre-trained 2D CNN networks was used for our 3D CNN model. Fig. 5 shows the architecture of our proposed 3D CNN model.

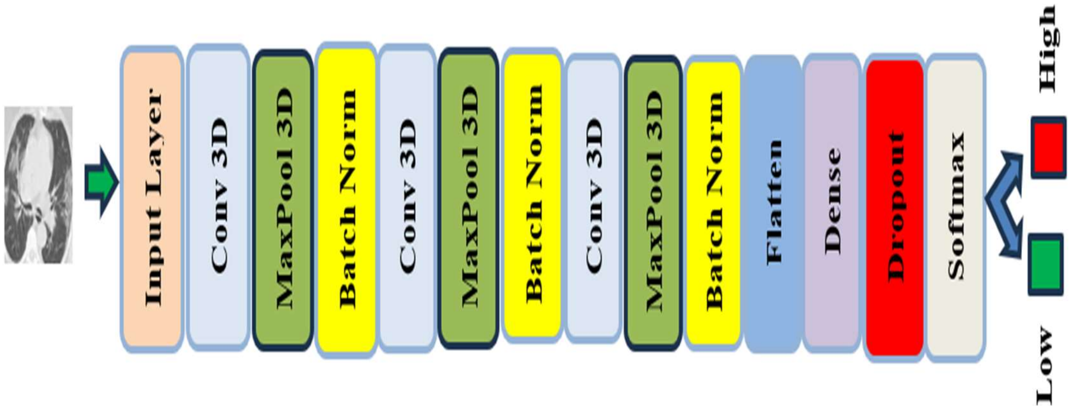


Fig. 5. The architecture of our proposed 14-layer 3D CNN. It consists of 3 convolutional layers, MaxPool layers, dense, and dropout layers. It classifies the input lung slice into 'High' or 'Low' severity categories.

The SoftMax layer is an integral part of a neural network. The SoftMax activation function provides non-linearity to the neural network. There are different types of activation functions, but the SoftMax activation function is the most used in CNN. The SoftMax activation function calculates the relative probabilities to determine the final probability value. The SoftMax activation function was used to determine the two categories of the testing dataset. Unlike Sigmoid, which is used for multi-class classification, the Softmax classification layer provides a range of probabilities in a vector. Each vector represents the probability for one of the classes.

V. Experimental Results

This section provides detailed evaluation results of the testing dataset on the pre-trained 2D CNNs and the proposed 3D CNN. It explains the evaluation metrics, hyperparameters, cross-validation, and ablation study of previous state-of-the-art methods.

A. *Evaluation Metrics*

Popular classification evaluation metrics, which include ACC and The Area Under the ROC Curve (AUC), were used to evaluate the 2D pre-trained neural networks and the 3D CNN on the dataset. AUC and ACC metrics were selected for the evaluation as demanded by the ImageCLEF challenge rule. The priority lies on the AUC. All the works submitted for the ImageCLEF TB SVR challenge used the same metrics and were evaluated as a binary classification task.

B. *Hyperparameter*

The values most appropriate for the convergence of models are chosen for all the models. We adopt the hit-and-trial method to explore all the possibilities of determining the best parameter. Also, the hit and trial method was used to determine the kernel size and the number of filters. We experimented with various commonly used kernel sizes, such as 3×3 and 5×5, and observed that a kernel size of 3 is most appropriate for the final experiment. An early stopping method that terminates the training when the validation loss increases was employed to choose the number of training epochs.

Since the weight of a network adjustment with the gradient loss is determined by the learning rate, it is essential to choose the learning rate carefully. The learning rate for the experiment was examined between the range of 0.0001 and 0.01 while keeping all other hyperparameters the same. Relatively small batch sizes were used to converge the model [34] rapidly. Large batch size, though, speeds up the training process; it consumes more memory space [35]. The experiment used a 20% dropout rate to avoid overfitting. Randomly chosen neurons in the training phase were ignored by the dropout technique. However, the dropout technique temporarily disconnects the neurons that were ignored during the forward past to prevent any changes to their weights during the backward pass. TB severity assessment is a binary classification task. Therefore, the binary cross-entropy function was used as the objective function. Softmax classifier assigns a probability distribution to each category. The Softmax classifier outputs values 0 and 1, representing the number of categories ('High' and 'Low'). Table 2 shows the summary of Hyperparameters used in this study.

Table 2. Hyperparameters of our proposed 3D CNN

Hyper-parameter	3D-CNN
Learning rate	0.0001
Batch size	20
Optimizer	Adam
No. of epochs	60
Activation function	ReLU/ Softmax
Dropout	0.2
Loss function	Categorical cross-entropy

C. *Cross Validation*

A five-fold cross-validation was employed to determine the final ACC and AUC of the 2D pre-trained CNN models and our proposed 3D-CNN on the dataset. The training and testing datasets were summed together and divided into five groups. While one group served as the testing dataset per iteration, the other four groups served as the training dataset. The process was repeated until five iterations were completed, and each group served as a testing set. The average value was chosen for each of the metrics. Table 2 shows the ACC cross-validation comparison result, while Table 4 shows the AUC cross-validation results across the five folds. Table 5 summarizes the average ACC and AUC performance of the 2D and 3D CNN models. It shows the processing time per test and the testing time per epoch.

From the results of the cross-validation in Tables 3 and 4, it can be observed that our proposed 3D CNN outperformed the pre-trained 2D CNN models. The proposed 3D CNN generated an overall ACC of 0.9929 and an AUC of 0.9982. The second best-performed model, which is the Inception V4 model, produced an average ACC of 0.9857 and an average AUC of 0.9837. Our proposed 3D CNN improved on the Inception V4 model with a difference of 0.0072 in ACC and 0.0145 in AUC values. Fig. 6 shows the AUC curve of our proposed 3D CNN model.

Table 3. Five-fold cross-validation of ACC on the 2D CNNS and the proposed 3D CNN model

Group	VGGNet-16	VGGNet-19	DesnseNet 121	Google-Net	ResNet-50	Inception V4	Proposed 3D CNN
1	0.9772	0.9767	0.9737	0.9799	0.9835	0.9848	0.9906
2	0.9732	0.9755	0.9732	0.9733	0.9822	0.9858	0.9924
3	0.9712	0.9796	0.9894	0.9716	0.9853	0.9845	0.9978
4	0.9792	0.9793	0.9891	0.9806	0.9817	0.9867	0.9919
5	0.9808	0.9795	0.9812	0.9622	0.9878	0.9867	0.9918
Avg.	0.9763	0.9781	0.9813	0.9735	0.9840	0.9857	**0.9929**

Table 4. Five-fold cross-validation of AUC on the 2D CNNs and the proposed 3D CNN model

Group	VGGNet-16	VGGNet-19	DesnseNet 121	Google-Net	ResNet-50	Inception V4	Proposed 3D CNN
1	0.9779	0.9767	0.9837	0.9771	0.9735	0.9848	0.9988
2	0.9788	0.9755	0.9823	0.9783	0.9742	0.9858	0.9957
3	0.9728	0.9796	0.9897	0.9699	0.9865	0.9845	0.9979
4	0.9821	0.9793	0.9897	0.9721	0.9899	0.9867	0.9992
5	0.9831	0.9795	0.9897	0.9835	0.9887	0.9867	0.9996
Avg.	0.9789	0.9781	0.9870	0.9762	0.9826	0.9837	**0.9982**

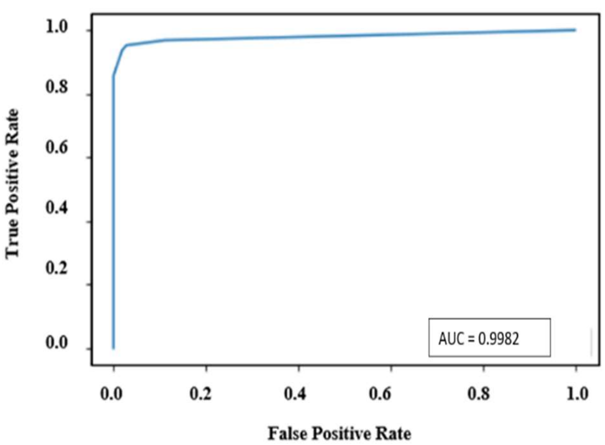


Fig. 6. The AUC curve of our proposed 3D CNN model.

Table 5. Summary of the evaluation performance among the 2D CNNs and the proposed 3D CNN

Model	ACC (%)	AUC (%)	Processing time/test (sc)	Testing time/epoch (sc)
VGGNet-16	0.9763	0.9789	2.20	59.23
VGGNet-19	0.9781	0.9790	2.22	59.11
DenseNet 121	0.9813	0.9870	1.97	58.90
GoogLeNet	0.9735	0.9762	2.32	59.22
ResNet-50	0.9840	0.9826	1.76	57.21
InceptionV4	0.9857	0.9837	1.34	53.54
Proposed 3D CNN Model	**0.9929**	**0.9982**	**0.95**	**49.22**

D. Evaluation Performance Comparison of the Proposed 3D CNN with Some State-of-the-art CNN Models on Similar Datasets

The experimental result of our proposed 3D CNN was further compared with some recently used state-of-the-art CNN methods on lung CT and X-ray images for tuberculosis severity assessment. Table 6 shows a detailed comparison of the state-of-the-art techniques.

Table 6. Comparison of our proposed 3D CNN model with recently used state-of-the-art CNN techniques.

Author	Modality	ACC
[36]	CT	0.6750
[7]	CT	0.8119
[17]	X-ray	0.9720
[18]	CT	0.8671
[20]	CT	0.9649
[21],	CT	0.9534
[28]	X-ray	0.090
[24]	CT/X-ray	0.9447
[37]	CT	0.9690
[29]	CT	0.9105
Our 3D CNN	**CT**	**0.9929**

Overall, the proposed 3D CNN model generated an ACC of and an AUC of 0.9929 and 0.9982. When compared with the ACC and AUC values of the pre-trained 2D CNN, it can be observed that the proposed 3D CNN produced a higher value difference of 0.0072 ACC and 0.0145 AUC. This is a significant difference considering the various limitations, such as inter-class similarity associated with the medical image domain. In addition, the proposed 3D CNN has a lower processing time per test and testing time per epoch of 0.95 and 49.22 respectively. This is due to our diligence in selecting the best hyperparameters for the final experiment". The ablation study in Table 6 shows that our proposed 3D CNN has the highest ACC and AUC values compared with the recently used similar techniques.

E. Analysis of the Misclassified Testing Images Using the Confusion Matrix

The misclassified images on our proposed 3D CNN and the best performer on the pre-trained 2D CNNs were further analysed using the confusion matrix [38]. The confusion matrix shows the number of correctly classified images and the number of misclassified images. It calculates the True Positive (TP), the True Negative (TN), the False Positive (FP) and the False Negative (FN). Table 7 shows the formula for

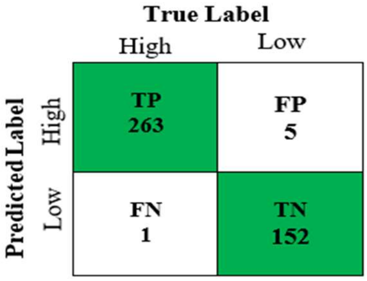
ACC and the error rate of the confusion matrix.

Table 7. The ACC and Error rate formula using the confusion matrix

Accuracy (Recognition rate)	Error Rate (Misclassification)
(TP+TN)/(TP+FP+TN+FN)	(FP+FN)/ (TP+FP+TN+FN)

F. Evaluation of the Accuracy and the Error Rate

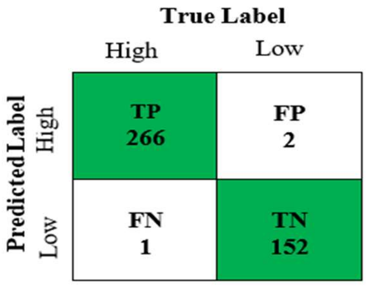
Fig. 7 shows the confusion matrix of the Inception V4 and our proposed 3D CNN on the ImageCLEF 2021 testing dataset. It can be observed that the Inception V4 model has a total of six misclassified images, while our proposed 3D CNN has only three misclassified testing images. The Inception V4 generated an error rate of 0.0142, while our proposed 3D CNN produced a lesser error rate of 0.0071.



(A)

$$ACC = \frac{263+152}{421} = 0.9857$$

$$Error\ rate = \frac{5+1}{421} = 0.0142$$



(B)

$$ACC = \frac{266+1}{421} = 0.9929$$

$$Error\ rate = \frac{2+1}{421} = 0.0071$$

Fig. 7. The confusion matrix of the Inception V4 model (A) and our proposed 3D CNN model (B). It shows the values for ACC and the error rate.

E. Implementation Details

The proposed model's framework comprises Window10 Pro, GPU: NVIDIA Tesla K40c, memory: 12.0 GB, Image processing: Python 3.8, OpenCV 3.10, and NumPy 1.6.

VI. Discussion and Conclusion

In this study, we have compared two different approaches to TB severity assessment in lung CT images. Firstly, we explored the potential of six different 2D pre-trained neural networks, and secondly, we proposed a new 3D CNN. The models were evaluated on the ImageCLEF 2021 TB dataset using the ACC and AUC classification metrics. From the experimental evaluation results of the pre-trained CNN, it can be observed that the Inception V4 model has the highest ACC and AUC values of 0.9857 and 0.9837 respectively. Our proposed 3D CNN produced higher ACC and AUC values compared with the Inception V4 network. It generated ACC and AUC values of 0.9929 and 0.9982 respectively. In addition, our proposed 3D CNN has the lowest processing time per test of 0.95 and testing time per epoch of 49.22.

Our proposed 3D CNN model, compared with the 2D pre-trained CNNs, prevented the loss of contextual information associated with 2D CNNs in analyzing 3D images such as lung CT images. Our proposed 3D CNN is a robust 14-layer CNN model with relatively low parameters. The low parameters and the dropout employment help prevent overfitting of the model. It enhances the model performance in efficiently classifying the testing dataset into the 'High' and 'Low' TB severity categories.

The analysis of the misclassified testing images using the confusion matrix showed that our proposed 3D CNN has the lowest error rate and fewer misclassified images. In addition, the five-fold cross-validation method employed in this study improves our proposed 3D CNN robustness.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Author DO Conceptualization, methodology, software, and first draft. Author FH, software and validation, Author AF and MY, reviewing and editing, visualization; all authors had approved the final version.

References

[1] M. A. Abdou, "Literature review: Efficient deep neural networks techniques for medical image analysis," *Neural Computing and Applications*, vol. 34, no. 8, pp. 5791–5812, 2022.
[2] A. F. Ismoilovich, "Tuberculosis Diagnostics with Modern Solutions (Literature Review)," *Central Asian Journal of Medical and Natural Science*, vol. 3, no. 3, pp. 377–383, 2022.
[3] J. D. Kelleher, *Deep Learning*; Massachusetts: MIT press, 2019.
[4] N. Ketkar and J. Moolayil, "*Convolutional Neural Networks*," in *Deep Learning with Python: Learn Best Practices of Deep Learning Models with Pytorch*, CA: Apress Berkeley, 2021, pp. 197–242,
[5] D. O. Alebiosu and F. P. Muhammad, "Medical image classification: a comparison of deep pre-trained neural networks," in *Proc. 2019 IEEE Student Conference on Research and Development (SCOReD)*, 2019, pp. 306–310.
[6] M. Singh, G. V. Pujar, and S. A. Kumar, "Evolution of machine learning in tuberculosis diagnosis: a review of deep learning-based medical applications," *Electronics*, vol. 11, no. 17, p. 2634, 2022.
[7] D. O. Alebiosu, A. Dharmaratne, and C. H. Lim, "Improving tuberculosis severity assessment in computed tomography images using novel DAvoU-Net segmentation and deep learning framework," *Expert Systems with Applications*, vol. 213, 119287, 2023.
[8] H. Jiang, Z. Diao, and T. Shi, "A review of deep learning-based multiple-lesion recognition from medical images: classification, detection and segmentation," *Computers in Biology and Medicine*, vol. 157, 106726, 2023.
[9] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv Print, arXiv:1409.1556, 2014. doi: 10.48550/arXiv.1409.1556
[10] V. Sudha and T. Ganeshbabu, "A convolutional neural network classifier VGG-19 architecture for lesion detection and grading in diabetic retinopathy based on deep learning," *Computers, Materials & Continua*, vol. 66, no. 1, 2021.
[11] G. Huang, "Densely connected convolutional networks," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2017.
[12] C. Szegedy, "Going deeper with convolutions," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2015.

[13] B. Koonce and B. Koonce, "ResNet 50," in *Convolutional Neural Networks with Swift for Tensorflow: Image Recognition and Dataset Categorization*, CA: Apress Berkeley, 2021, pp. 63–72.

[14] C. Szegedy, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Proc. AAAI Conference on Artificial Intelligence*, 2017.

[15] S. Kozlovski, "Overview of ImageCLEFtuberculosis 2021: CT-based tuberculosis type classification," in *Proc. CLEF 2021 Conference and Labs of the Evaluation Forum*, 21–24 September 2021.

[16] Y. X. Tang, "Automated abnormality classification of chest radiographs using deep convolutional neural networks," *NPJ Digital Medicine*, vol. 3, no. 1, p. 70, 2020.

[17] S. H. Karaddi and L. D. Sharma, "Automated multi-class classification of lung diseases from CXR-images using pre-trained convolutional neural networks," *Expert Systems with Applications*, vol. 211, 118650, 2023.

[18] C. C. Chiet, "A Lung Cancer Detection with Pre-Trained CNN Models," *Journal of Informatics and Web Engineering*, vol. 3, no. 1, pp. 41–54, 2024.

[19] O. Adebisi, "Automatic classification of lung nodules on computed tomography images using a pre-trained convolutional neural network," *International Journal of Engineering Science Invention (IJESI)*, vol. 9, no. 1, pp. 63–66, 2020.

[20] P. Ardimento, "Deep neural networks ensemble for lung nodule detection on chest CT scans," in *Proc. 2021 International Joint Conference on Neural Networks (IJCNN)*, 2021

[21] T. Lima, "Automatic classification of pulmonary nodules in computed tomography images using pre-trained networks and bag of features," *Multimedia Tools and Applications*, vol. 82, no. 27, pp. 42977–42993, 2023.

[22] L. M. Florescu,"Federated learning approach with pre-trained deep learning models for COVID-19 detection from unsegmented CT images," *Life*, vol. 12, no. 7, p. 958. 2022

[23] K. Ramana, "Early prediction of lung cancers using deep saliency capsule and pre-trained deep learning frameworks," *Frontiers in Oncology*, vol. 12, 886739, 2022.

[24] D. Han, "Diagnosis of active pulmonary tuberculosis and community acquired pneumonia using convolution neural network based on transfer learning," *Academic Radiology*, vol. 29, no. 10, p. 1486–1492, 2022

[25] M. Nijiati, "Deep learning based CT images automatic analysis model for active/non-active pulmonary tuberculosis differential diagnosis," *Frontiers in Molecular Biosciences*, vol. 9, 1086047, 2022.

[26] Ç. Bakır and M. Babalık, "Detection of tuberculosis disease with convolutional neural networks," in *Proc. International Conference on Advanced Technologies*, 2023.

[27] S. Ghaderizadeh, "Hyperspectral image classification using a hybrid 3D-2D convolutional neural networks," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 7570–7588, 2021.

[28] M. A. Elashmawy, "Detection of tuberculosis based on hybridized pre-processing deep learning method," *International Journal of Advanced Computer Science and Applications*, 2023.

[29] C. Yan, "A fully automatic artificial intelligence–based CT image analysis system for accurate detection, diagnosis, and quantitative severity evaluation of pulmonary tuberculosis," *European Radiology*, pp. 1–12, 2022.

[30] S. Kozlovski, "Overview of ImageCLEFtuberculosis 2021-CT-based Tuberculosis Type Classification," in *Proc. CLEF (Working Notes)*, 2021.

[31] A. Krizhevsky, I. Sutskever, and G.E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, 2012

[32] J. Wu, "On the effectiveness of sampled softmax loss for item recommendation," *ACM Transactions on Information Systems*, vol. 42, no. 4, pp. 1–26, 2024.

[33] D. García-Martín, M. Larocca, and M. Cerezo, "Effects of noise on the overparametrization of quantum neural networks," *Physical Review Research*, vol. 6, no. 1, 013295**,** 2024.

[34] H. Zulkifli, "Understanding Learning Rates and How It Improves Performance in Deep Learning," *Towards Data Science*, vol. 21, no. 23, 2018.

[35] R. A. Hamad, "Efficient activity recognition in smart homes using delayed fuzzy temporal windows on binary sensors," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 2, pp. 387–395, 2019.

[36] H. Zunair, "Uniformizing techniques to process CT scans with 3D CNNs for tuberculosis prediction," *Predictive Intelligence in Medicine: Third International Workshop, PRIME 2020*, Held in Conjunction with *MICCAI 2020*, Lima, Peru: Springer, Oct. 2020.

[37] R. Aswiga and A. Shanthi, "A multilevel transfer learning technique and LSTM framework for generating medical captions for limited CT and DBT images," *Journal of Digital Imaging*, vol. 35, no. 3, pp. 564–580, 2022.

[38] F. Amin and M. Mahmoud, "Confusion matrix in binary classification problems: A step-by-step tutorial," *Journal of Engineering Research*, vol. 6, no. 5, 2022.