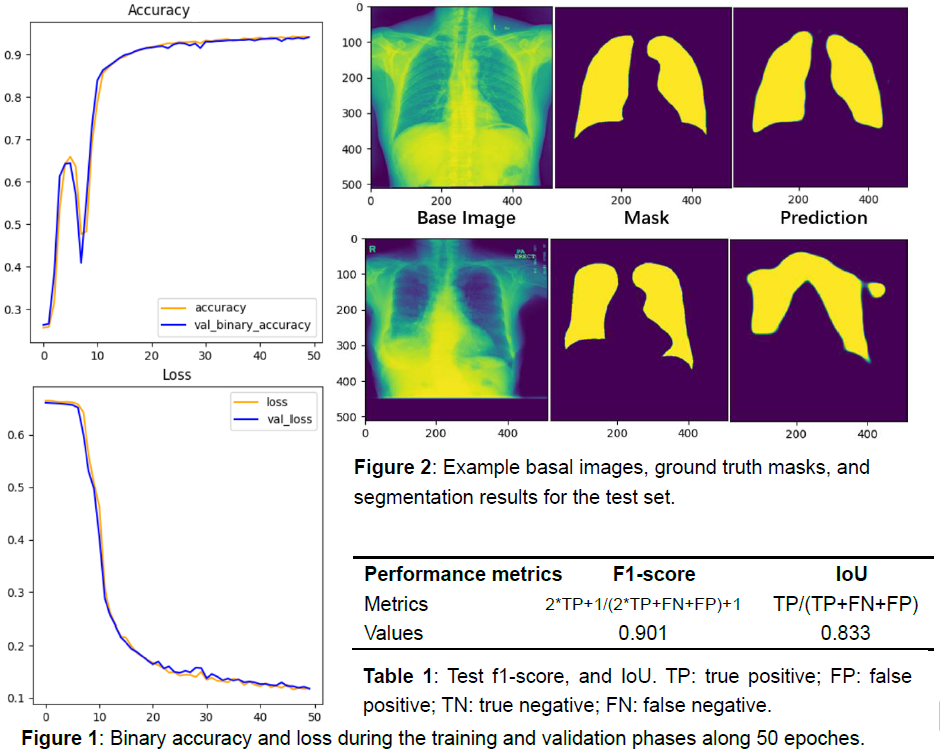
**A Handy Software for Chest X-ray Mycobacterium Tuberculosis (M.TB) Diagnosis**

The project aims to develop a software for effective diagnosis of M. TB infection by chest X-ray.

**Lung Segmentation.**

A modified U-net model based on previous work by Dr. Pandey was used for lung segmentation. Our U-net comprises 28 layers falling into a contracting and an expanding path. Four repeats of two 3\*3 convolution layers (activation function: *ReLU*, zero padding) and a 2\*2 max-pooling layer (stride: 2) comprises the contracting path, while four repeats of one 2\*2 and two 3\*3 up-convolutional layers (*ReLU*) concatenated with the feature map from the corresponding down-sampling layers, made up the expanding path. The Kaggle dataset was divided into training, validation, and testing sets at an 8:1:1 ratio. The assessment of model performance in the training and validation procedure was achieved with three metrics: binary accuracy, loss, and dice coefficient. Accuracy and loss curves were plotted against epochs to visualize the training procedure, as shown in Figure 1. Notably, binary accuracy in the training and validation phases reached approximately 96%. In contrast, the loss(proportionof miss-segmentedregions to total ground truth and prediction) was reduced to below 0.1 in both stages within 50 epochs, indicating a high accuracy without apparent overfitting.



We used dice coefficient (f1-score) and intersection over union (IoU) to evaluate our model using testingdata.Modelevaluationatthetesting phase revealed a dice coefficient of 0.901 and an IoU of 0.833, indicating overall reliable segmentation performance, despite a potential risk for miss-segmentations. Indeed, as shown in Figure 2, the segmented pictures retain major characteristics of the ground truth masks (Figure 2a). However, occasionally exterior regions are miss-classified as the lung. For example, tracheal areas in Figure 2b were miss-segmented, while part of bottom lung areas was neglected, which may confound M. TB diagnosis on segmented images. Several reasons may underlie the occasional low-quality segmentations. From one perspective, the power of our model may be insufficient, and adjustment of the learning rate, along with increased computing power of our device, may improvenetworkperformance.Fromanother,imagequalitymayimpactsegmentation. For example, Figure 2bseemsoverexposed,withexcessivelungtransparencyandparenchymalbrightness.Carefulimage preprocessing, such as brightness equalization, may improve segmentation and M. TB diagnosis in the future.

**Mycobacterium Tuberculosis (M.TB) Infection Classification.**

***Model construction.***

For classification, we used the TB\_Chest\_Radiography\_Database containing 3500 healthy and 700 M.TB chest X-ray images, which we divided into training, validation, and testing set at 7:2:1 ratio. Athree-layerconvolutional neural network (CNN) was constructedand implemented with *TensorFlow* (1.13.1) and *Keras* (2.3.1) in *Python3.7*, which is then used for classifying unsegmented lung images andsegmentedlungimages,plusevaluatingtheimpactofsegmentationon classification performance. Each convolutional layer used 4\*4 kernels with the activation function *ReLU* and was followed by a 4\*4 max-pooling layer. The convolutional layers were followed by a fully-connected layer consisting of1024neurons, adropout layerwith adropout rateof 0.5, andanoutputlayerusinga sigmoid function. RMSprop was selected as the optimizer and binary cross-entropy as the loss function. Finally, All images were converted to RGB format, and M.TB images underwent quadrupling data augmentation to balance normal and disease images in the training set before model training.

***Hyperparameter tuning.***

Table2. Comparative performance of CNN model with different hyperparameters

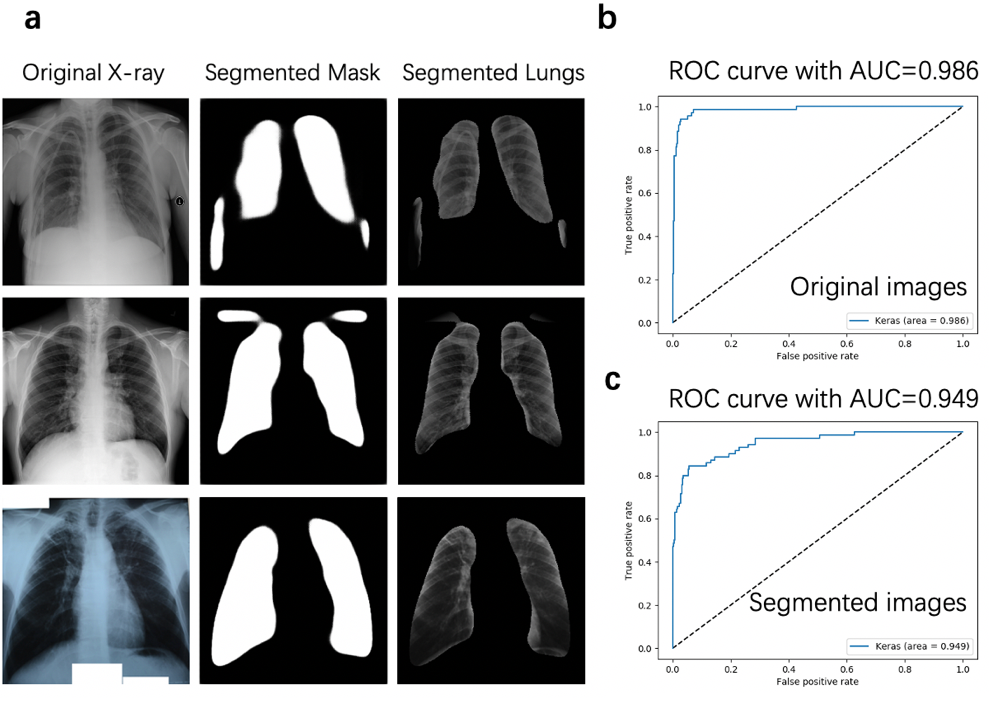
|  |  |  |  |
| --- | --- | --- | --- |
| Batch Size  Optimizer | 8 | 16 | 32 |
| RMSprop | 0.95 | 0.98 | 0.98 |
| Adam | 0.86 | 0.98 | 0.96 |
| SGD | 0.96 | 0.97 | **0.96** |

**图片包含 图示

描述已自动生成**Several models were trained and compared for hyperparameter tuning. Table 2 shows the comparative performance of models with different batch sizes and optimizers. The model with a batch size of 32 and the optimizer SGD was selected as the final model (Figure 3b). Other models either have insufficient classification capacity or show apparent overfitting (Figure 3a).

**Figure3. Model evaluation. (a)** The accuracy curve of nine models with batch size 8, 16, 32 and optimizer RMSprop, Adam and SGD. **(b)** The accuracy and loss trends of the model with best perform (batch size=32, SGD).

***Model evaluation.***

We evaluate our model with several methods. The ‘model.evaluate()’ function was called to print the accuracy of the test set after the whole training process of 10 epochs. Then, ROC curves withAUCvalues wereplotted toevaluate the true positive and false-positive rates of our model. The CNN models were trained with original and segmented X-ray images (Figure 4a). The models achieved an accuracy of 96% and an AUC value of 0.986 with original images (Figure 4b), along with an accuracy of 91%, and an AUC value of 0.949 with segmented images (Figure 4c), indicating reliable classification performance. Hence, we successfully developed a software for M.TB diagnosis.

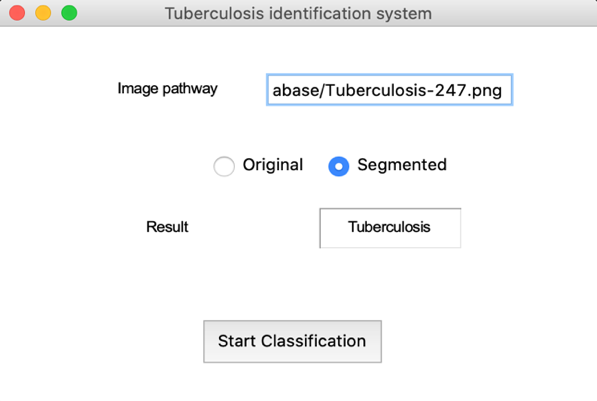
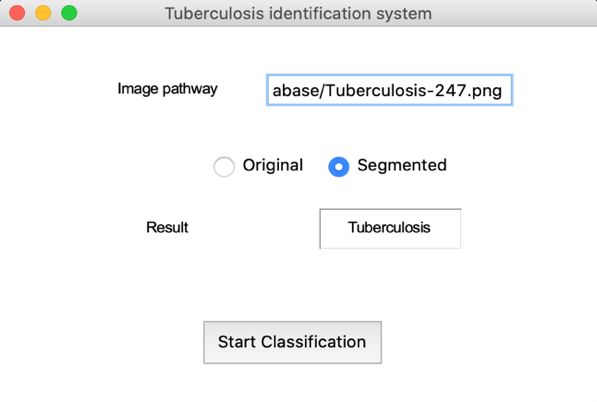
**Figure4. original and segmented images and ROC curves.** **(a)** The mask images were generated by our segmentation model. The lung images were segmented by performing the ostu function with original images and corresponding masks. **(b)** The ROC curve of CNN model with original images. **(c)** The ROC curve of CNN model with segmented images.

**Graphic User Interface.**

Finally, to improve customer experience, a graphical user interface (GUI) with both original image classification and segmented image classification functions was provided (Figure 5). For example, when the image path to a segmented chest X-ray image from an M.TB infected patient was loaded into the GUI, the classification result “ Tuberculosis” was presented in the result box. The process from image pathway to final classification result is firstly model loading, which is the body of machinelearning.Then,theinputimagewillbeextractedfromtheimagepathwayandthen processed through machine analysis. Lastly, the predicted classes and probabilities will be returned to GUI.

**Codes available at:** <https://github.com/CJJ8848/BIA4_miniproject_group8.git>

**Segmentation dataset**: <https://www.kaggle.com/nikhilpandey360/chest-xray-masks-and-labels> **Classification dataset:** <https://www.kaggle.com/tawsifurrahman/tuberculosis-tb-chest-xray-dataset>



**Figure5. Graphical user interface (GUI).** User simply type in image pathway and then get the classification result. User can choose to input either original or segmented X-ray images by clicking the radio button.