Your Name: A. S.

Name of your Device: PND

Algorithm Description

1. General Information

Intended Use Statement: This algorithm detects pneumonia according to the patient's x-ray taken from the chest.

Indications for Use: This algorithm is designed for male and female patients less than 70 years old. Patients with cardiomegaly and infiltration may falsely be diagnosed.

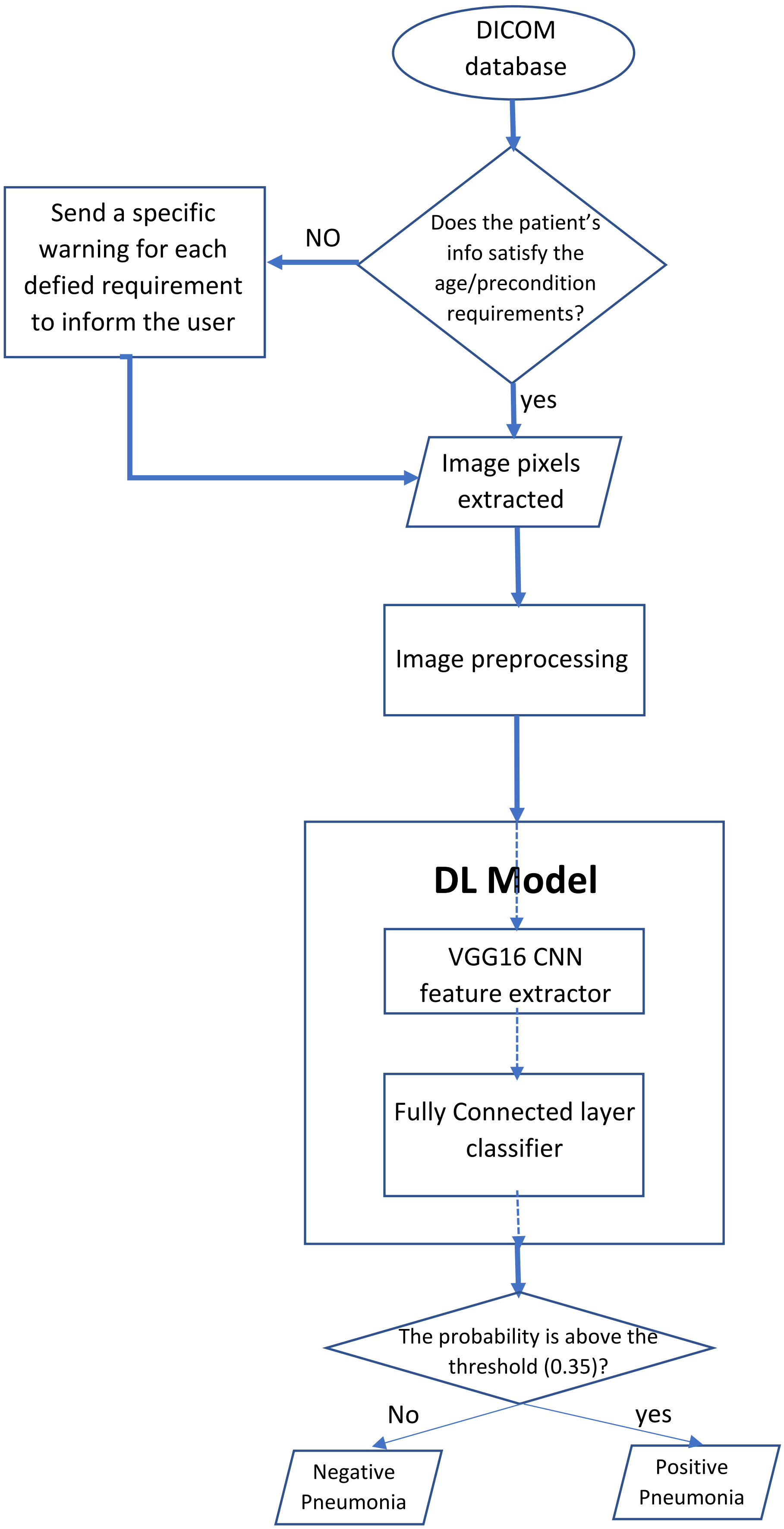
Device Limitations: patients with implant such as pacemaker would interfere with algorithm image processing stage which could lead to poor analysis. Furthermore, patents with lung disease history such as infiltration and effusion may not be suitable candidate for pneumonia diagnosis using the proposed algorithm.

Clinical Impact of Performance: This algorithm is able to expedite and considerably assist pneumonia diagnosis process. Since pneumonia is a serious health condition, false negative result can lead to severe consequences for the patients, while, positive readings can always be forwarded to secondary confirming test. Hence, this algorithm provide sufficient sensitivity to rule out pneumonia by lowering false positive rates. Therefore, i.e. if a patient is tested negative with this algorithm, there is a high chance and reliability that the patient does not have pneumonia.

2. Algorithm Design and Function

DICOM Checking Steps: Patient's demographic including medical history is studied to ensure their age is less than 70 years old and does not have prior cardiovascular issues to require any implants. Furthermore, they are verified if they have history of infiltration and/or effusion. All prerequisites confirmed, the image info (pixels) will be extracted to be fed into the developed model.

check_dicom() function: check_dicom will read the important data form DICOM file and compare it with the algorithm requirements such as age, scanned body part, and patient position. If any of the aforementioned requirements denied, instead of rejecting prediction, the patient's image will be passed onto the model to predict the presence of pneumonia, however, a warning message will also be included in the final report indicating the flaw. To compare the result, Study Description from DICOM file is also printed.


DICOM database
Does the patient's info satisfy the age/precondition requirements?
NO
Send a specific warning for each defied requirement to inform the user
yes
Image pixels extracted
Image preprocessing
DL Model
VGG16 CNN feature extractor
Fully Connected layer classifier
The probability is above the threshold (0.35)?
No
yes
Negative Pneumonia
Positive Pneumonia

Preprocessing Steps: The image is initially cropped to only focus on the lung section of the xray image followed by normalization and standardization with average of [0.485, 0.456, 0.406], and standard deviation of [0.229, 0.224, 0.225].

CNN Architecture: pretrained VGG16 model is utilized for the CNN section of the algorithm. Following is the summary of the CNN architecture:

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #   
=================================================================
input_1 (InputLayer)         (None, 224, 224, 3)       0         
_________________________________________________________________
block1_conv1 (Conv2D)        (None, 224, 224, 64)      1792      
_________________________________________________________________
block1_conv2 (Conv2D)        (None, 224, 224, 64)      36928     
_________________________________________________________________
block1_pool (MaxPooling2D)   (None, 112, 112, 64)      0         
_________________________________________________________________
block2_conv1 (Conv2D)        (None, 112, 112, 128)     73856     
_________________________________________________________________
block2_conv2 (Conv2D)        (None, 112, 112, 128)     147584    
_________________________________________________________________
block2_pool (MaxPooling2D)   (None, 56, 56, 128)       0         
_________________________________________________________________
block3_conv1 (Conv2D)        (None, 56, 56, 256)       295168    
_________________________________________________________________
block3_conv2 (Conv2D)        (None, 56, 56, 256)       590080    
_________________________________________________________________
block3_conv3 (Conv2D)        (None, 56, 56, 256)       590080    
_________________________________________________________________
block3_pool (MaxPooling2D)   (None, 28, 28, 256)       0         
_________________________________________________________________
block4_conv1 (Conv2D)        (None, 28, 28, 512)       1180160   
_________________________________________________________________
block4_conv2 (Conv2D)        (None, 28, 28, 512)       2359808   
_________________________________________________________________
block4_conv3 (Conv2D)        (None, 28, 28, 512)       2359808   
_________________________________________________________________
block4_pool (MaxPooling2D)   (None, 14, 14, 512)       0         
_________________________________________________________________
block5_conv1 (Conv2D)        (None, 14, 14, 512)       2359808   
_________________________________________________________________
block5_conv2 (Conv2D)        (None, 14, 14, 512)       2359808   
_________________________________________________________________
block5_conv3 (Conv2D)        (None, 14, 14, 512)       2359808   
_________________________________________________________________
block5_pool (MaxPooling2D)   (None, 7, 7, 512)         0         
_________________________________________________________________
flatten (Flatten)            (None, 25088)             0         
=================================================================
Total params: 14,714,688
Trainable params: 2,359,808
Non-trainable params: 12,354,880
_________________________________________________________________
```

Figure 1- VGG16 model summary

Fully connected classifier: previous CNN layer extracts the important features of the image and pass them onto the fully connected layer where the their influence is adjusted according to achieve the correct prediction. In order to improve model generalization, dropout layers are also added in between the fully connected layers. The last layer is a sigmoid function computing the probability of pneumonia.

3. Algorithm Training

Augmentation: images are randomly flipped horizontally. Moreover, their height and width are randomly shifted by 1 pixel. Random rotation range is also set to 10 degree.

Batch size: 32

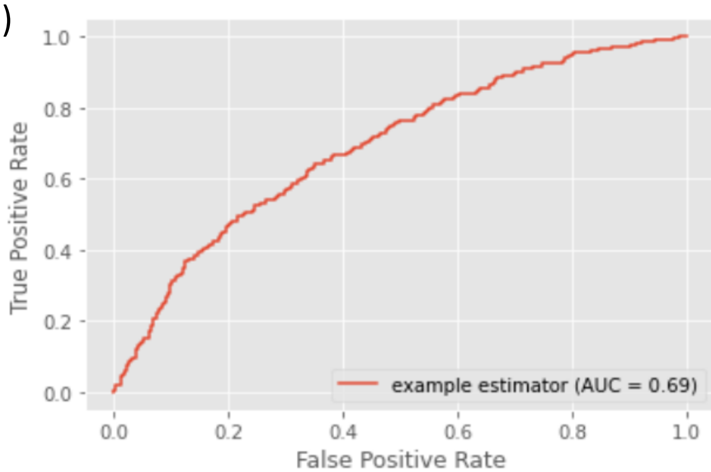
Optimizer learning rate: Adam optimizer is selected with initial learning rate of 0.001 and beta_1 of 0.3. the learning rate is scheduled to decrease by 10 fold when there is no improvement in validation set accuracy for 5 successive epoch.

Layers of pre-existing architecture that were frozen: according to Figure 1, sixteen VGG layers have been frozen (upto block5_conv2).

Layers of pre-existing architecture that were fine-tuned: layer block5_conv3

Layers added to pre-existing architecture: Three fully connected layers with a dropout layer in between followed by a final fully connected layer using sigmoid function for binary classification.

a)



b)

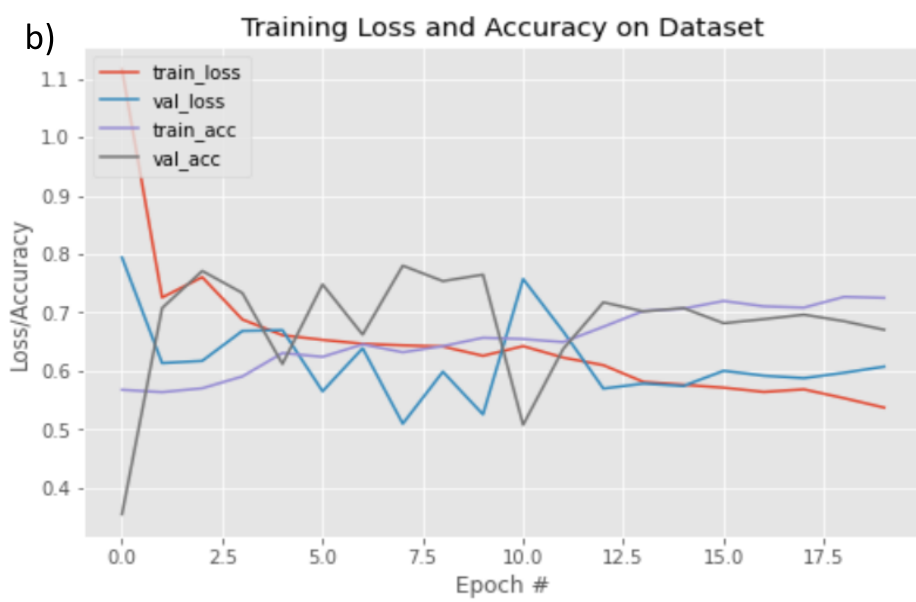


Figure 2- a) ROC curve, b) training metrics changes versus number of epochs

Final Threshold and Explanation: The threshold is found 0.35 based on maximum F1 value. This has resulted in sensitivity of 0.64, precision of 0.31, and F1 score of 0.42

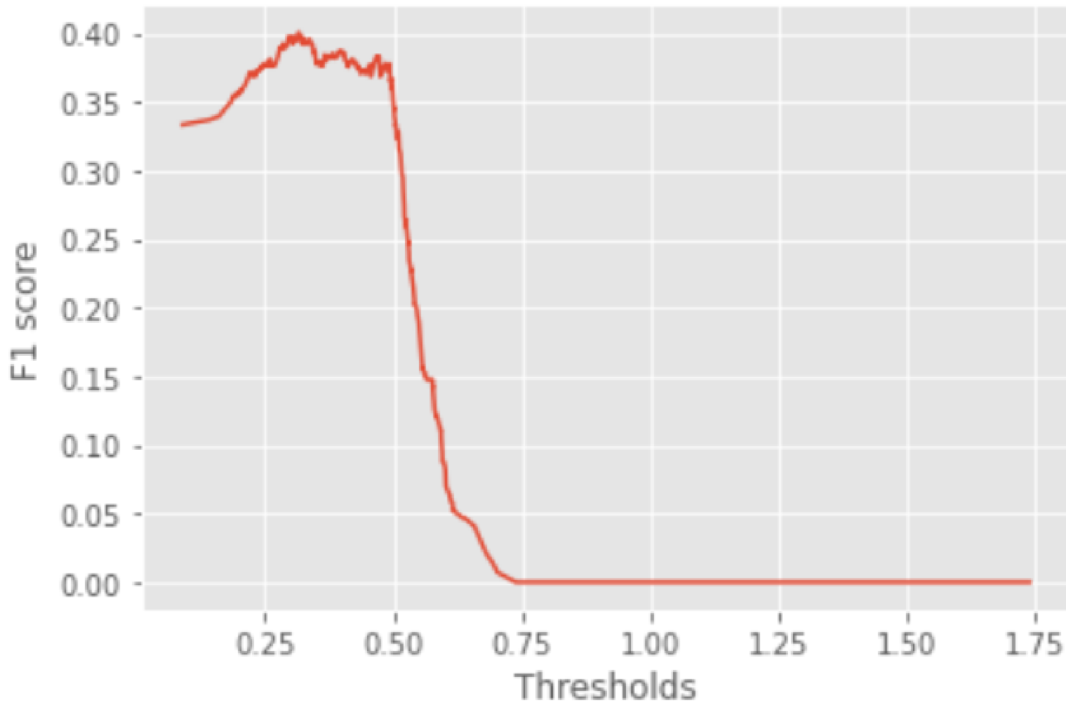


Figure 3- F1 score according to the thershold

4. Databases

Training Dataset: This dataset is composed of 2288 patients half of which has confirmed pneumonia. The age ranges from 10 to 90 with more focus on 20 to 70 years old.

Validation Dataset: This dataset is composed of 1430 patients 1/5 of which has been diagnosed with pneumonia. This ratio is proportional to the prevalence rate of pneumonia. This data set follows similar age distribution as training dataset.

Following gaphs provide info ragrading age distribution in training and validation dataset. Figure 5 shows age distribution of male and female in both dataset which indicate fairly similar trend.

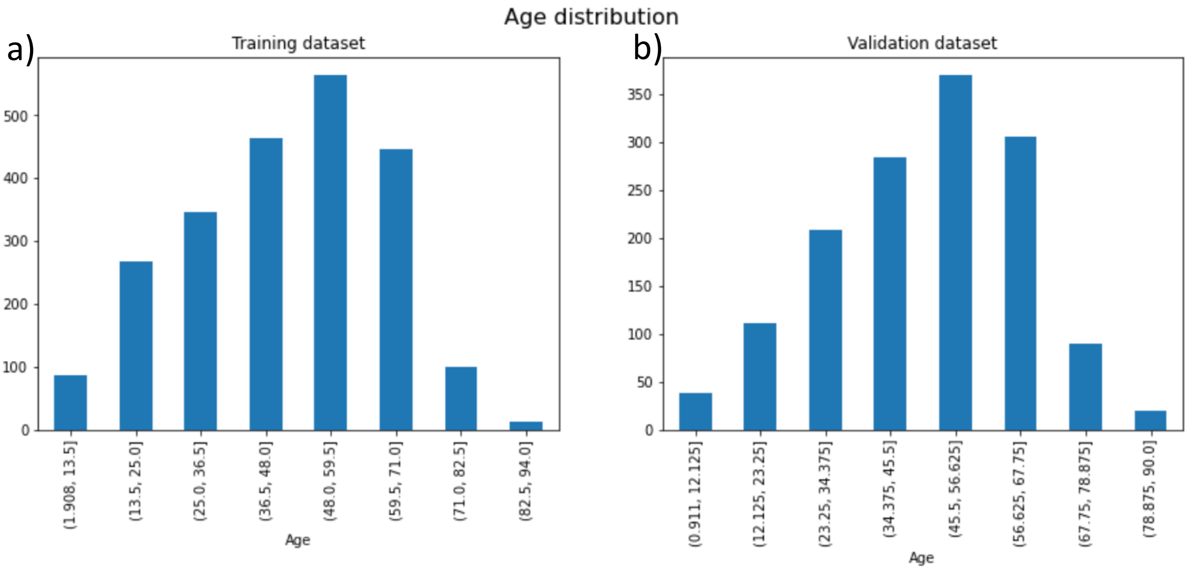


Figure 4- a) Age distribution in training dataset, b) Age distribution in validation dataset

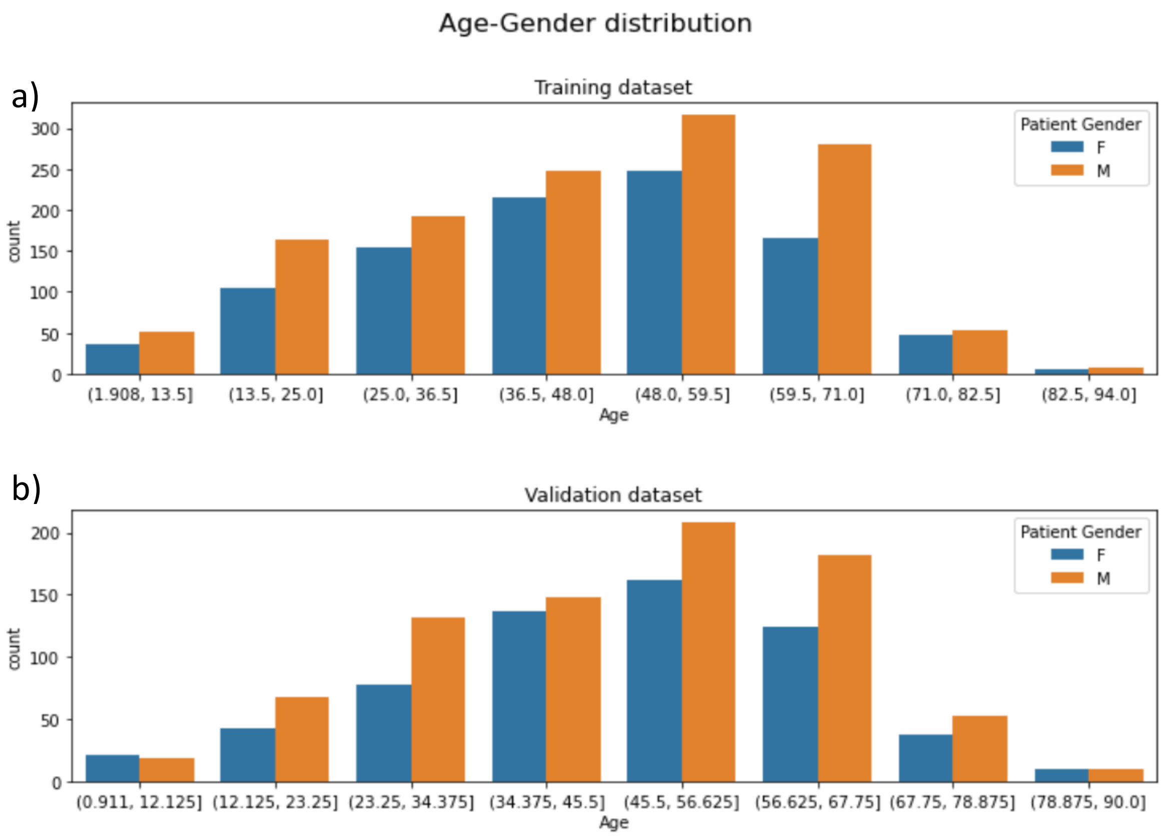


Figure 5- a) Age-Gender distribution in training dataset, b) Age-Gender distribution in validation dataset

5. Ground Truth

This NIH Chest X-ray Dataset is comprised of 112,120 X-ray images with disease labels from 30,805 unique patients. labeling process in this Open Access paper: "ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases." (*Wang et al.*)

6. FDA Validation Plan

Patient Population Description for FDA Validation Dataset: collaborating with Toronto General Hospital (TGH), pneumonia diagnosed patients with age ranging from 20 to 70 are selected with balanced gender distribution. The selected patients would not have cardiovascular issues, implant, and history of infiltration and effusion.

Ground Truth Acquisition Methodology: chest xray is a gold standard for pneumonia diagnosis. The dataset is confirmed with the department of radiology at TGH.

Algorithm Performance Standard: In order to achieve a balanced performance on sensitivity and precision, F1 score has been selected as the metric standard. Although, the goal here is to achieve high F1 score, one could also aim for high sensitivity or precision, according to their priority. The following paper is considered as the base standard to evaluate our proposed method performance.

Wang, Xiaosong, et al. "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.