FDA Submission

Your Name: ABDELRAHMAN USAMA

Name of your Device: PneuNET

Algorithm Description

1. General Information

Intended Use Statement:

For assisting the radiologist in detecting pneumonia in x-ray chest images.

Indications for Use:

Indicated for use in chest x-ray studies, where the patient's age is in the range between 20 and 80 for both genders (male and female), considering the screening position to be either AP or PA, in which it can assist radiologists in detecting pneumonia by giving an initial classification of the case, and the beginning of the workflow.

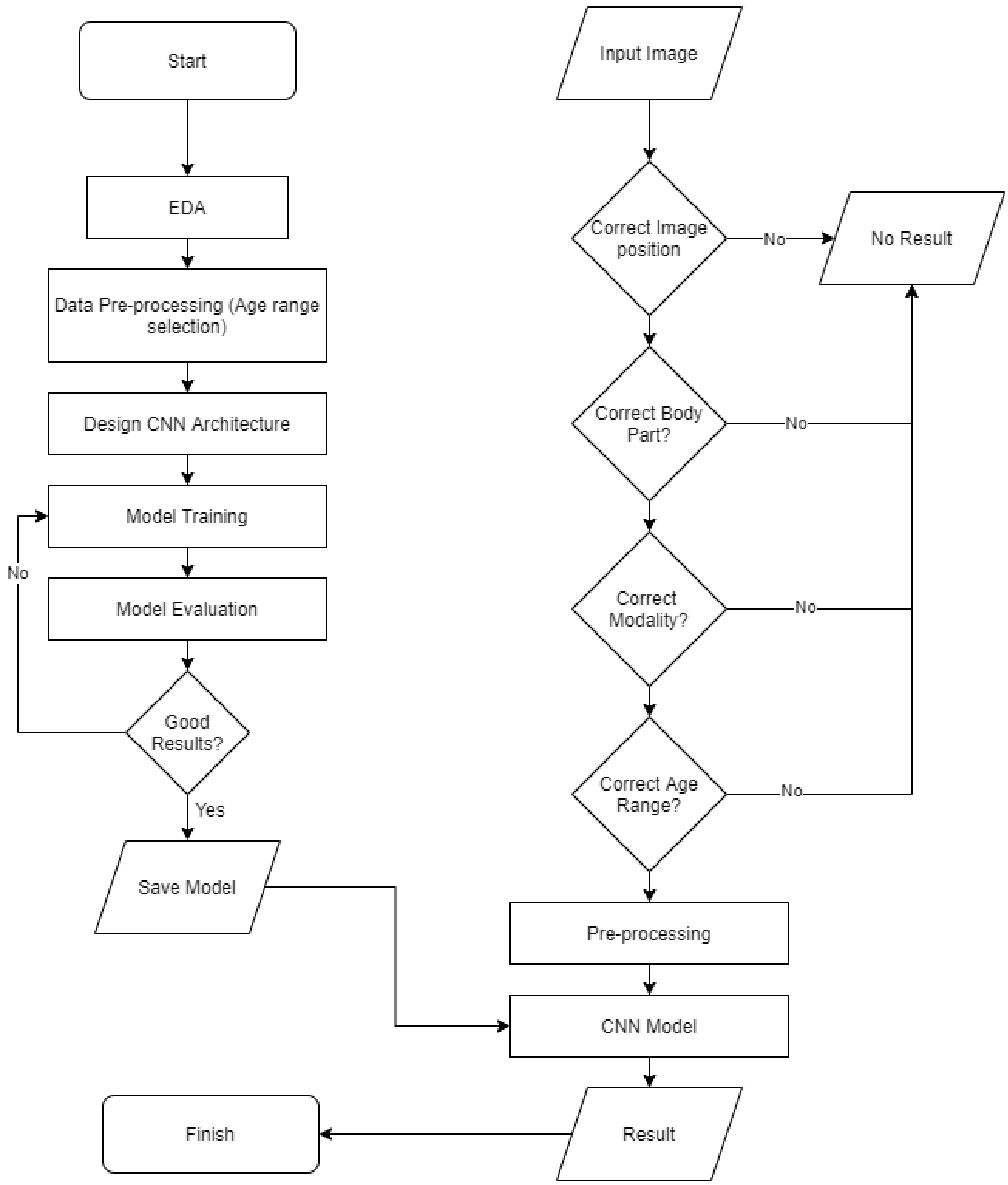
Device Limitations:

As a result of the EDA process, the algorithm might fail in classifying Pneumonia with the presence of the following comorbid: - Nodule - Fibrosis

Clinical Impact of Performance:
Considering this algorithm to assist and supportradiologists by giving an initial opinion about the case, it is quite important toconsider false positive and false negative cases as they can influence the radiologists.

False positive results can result in incorrect treatment which for some cases can have negative impact on the patient. On the other hand, false negatives can result in sending the patient home while he is in the need of the treatment. False negative cases are more critical and dangerous as patient with pneumonia must receive treatment and not sent home.

2. Algorithm Design and Function



DICOM Checking Steps:

1. Check the patient position is either 'PA' or 'AP'
2. Check if the examined body part is 'CHEST' or not.
3. Check if the modality is 'DX' or not.

Preprocessing Steps:

1. For training and validation, patient age below 20 and above 80 are excluded.
2. Images are resized to fit the pre-trained model architecture input size.
3. Images intensities are normalized to be in the range between 0 and 1.

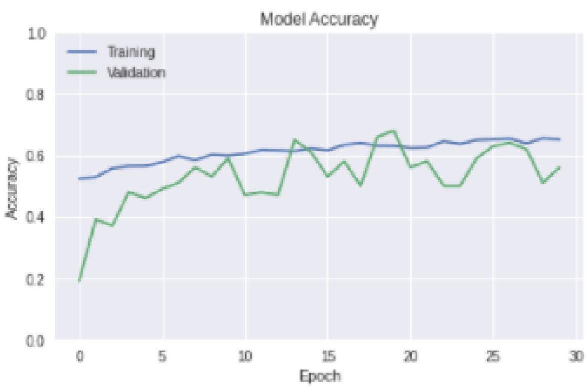
CNN Architecture:

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
model_1 (Model)              (None, 7, 7, 512)         14714688
_________________________________________________________________
flatten_1 (Flatten)          (None, 25088)             0
_________________________________________________________________
dropout_1 (Dropout)          (None, 25088)             0
_________________________________________________________________
dense_1 (Dense)              (None, 1024)              25691136
_________________________________________________________________
dropout_2 (Dropout)          (None, 1024)              0
_________________________________________________________________
dense_2 (Dense)              (None, 512)               524800
_________________________________________________________________
dropout_3 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_3 (Dense)              (None, 256)               131328
_________________________________________________________________
dense_4 (Dense)              (None, 1)                 257
=================================================================
Total params: 41,062,209
Trainable params: 28,707,329
Non-trainable params: 12,354,880
```

3. Algorithm Training

Parameters:

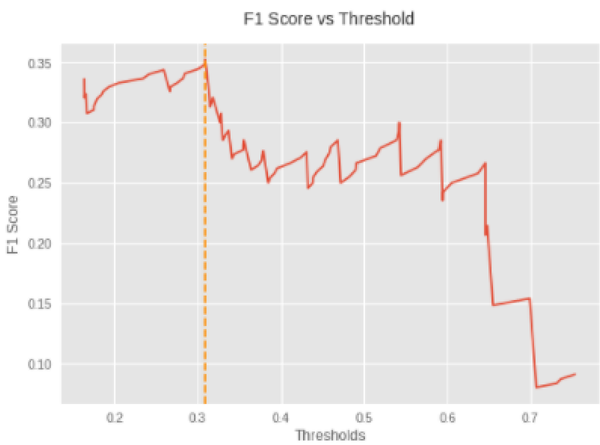
- Types of augmentation used during training
  - Rescale (1/255)
  - Horizontal Flip
  - Height Shift Range = 0.1
  - Width Shift Range = 0.1
  - Rotation Range = 20
  - Shear Range = 0.1
  - Zoom Range = 0.1
- Batch size
  - Train = 20
  - Validation = 100
- Optimizer learning rate = 1e-4
- Layers of pre-existing architecture that were frozen
  - input_1
  - block1_conv1
  - block1_conv2
  - block1_pool
  - block2_conv1
  - block2_conv2
  - block2_pool
  - block3_conv1
  - block3_conv2
  - block3_conv3
  - block3_pool
  - block4_conv1
  - block4_conv2
  - block4_conv3
  - block4_pool
  - block5_conv1
  - block5_conv2
- Layers of pre-existing architecture that were fine-tuned
  - block5_conv3
  - block5_pool
- Layers added to pre-existing architecture
  - Dropout(0.5)
  - Dense(1024, activation='relu')
  - Dropout(0.5)
  - (Dense(512, activation='relu')
  - Dropout(0.5)
  - (Dense(256, activation='relu')
  - Dense(1, activation='sigmoid')



The precision-recall curve is as follows:

Final Threshold and Explanation:

Threshold value of 0.31 had been chosen as it showed the maximum accuracy and F1-score the algorithm can achieve.



4. Databases

The dataset curated by the NIH specifically to address the problem of a lack of large x-ray datasets with ground truth labels to be used in the creation of disease detection algorithms. It contains a total of 112,120 X-ray images with disease labels from 30,805 unique patients.

Both training and validation dataset were for population age ranging between 20 and 80 years old, both genders, and pictures were taken from the same modality.

Description of Training Dataset:

The training dataset contained 80% of all thepositive pneumonia cases. Most of the training dataset were approximated to be menalmost 1200 patient data. As for the age, the range of 50-60 was more prominent.

Description of Validation Dataset:

The validation dataset had 20% positive pneumoniacases and 80% non-pneumonia. No finding case was more prominent and had a similar agerange as training dataset. Men cases were more than female, almost 700 and 600 accordingly.

5. Ground Truth

As driving clinical diagnosis from chest x-rays can be challenging, and diagnosis of pneumonia could be difficult as it can overlap with other diagnosis, it is quite important to consider a silver standard for this case to validate the model trained. As for the available dataset, the ground truth was obtained using Natural Language Processing (NLP) to mine the associated radiological reports. This approach had an accuracy estimated to be > 90%.

6. FDA Validation Plan

Patient Population Description for FDA Validation Dataset:

I would want to collect avalidation set that was made up of chest x-rays studies for both men and women betweenthe ages of 20 and 80. The dataset may also contain pneumonia with other comorbid such as atelectasis, cardiomegaly, consolidation, edema, effusion, emphysema, fibrosis, hernia, infiltration, mass, nodule, pleural thickening, and pneumothorax. The images acquisition tool must be digital x-ray where images are taken can be taken fromtwo different positions, both 'AP' and 'PA'. Considering the imaging modality to collect the data, Digital X-ray for digital radiography should be used. The dataset should also contain at least 1.2% of the cases as true cases (positive pneumonia cases) among all other types of diseases or comorbid.

Ground Truth Acquisition Methodology:

Considering the ideal case, a silver-standard approach would be preferred where 4 radiologists scan the images and use voting systemto reduce possible errors. As if we were to consider the NLP approach, the usage of NLPwould be used to label the dataset and then compare the results with the algorithm classification output.

Algorithm Performance Standard:

Considering the radiologist average performance of F1-score equal to 0.387 as cited in Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta,H., Duan, T., ... & Ng, A. Y. (2017). Chexnet: Radiologist-level pneumonia detectionon chest x-rays with deep learning. arXiv preprint arXiv:1711.05225, our algorithm wasable acheive 0.4927 which indicates that the algorithm has good performance for the scope of study.