FDA SUBMISSION

Masinde

2022-05-19

FDA Submission

Masinde Mtesigwa Masinde

Name of your Device: MassCancello

Algorithm Description

1. General Information

Intended Use Statement: This algorithm is intended for use on Caucasian, Hispanic, and African American women from the ages of 35-61 who have been administered a screening mammography study on a Hologic mammography machine and have never before demonstrated an abnormal mammography study.

Indications for Use: MassCancello is an image processing software that provides qualitative and quantitative analysis of the chest from x-ray images to support clinicians in the evaluation and assessment of pneumonia disease. The MassCancello software provides the following functionality:

Device Limitations: The results above indicate that the presence of infiltrations in a chest x-ray is a limitation of this algorithm, and that the algorithm performs very poorly on the accurate detection of pneumonia in the presence of infiltration. The presence of nodules and pneumothorax have a slight impact on the algorithm's sensitivity and may reduce the ability to detect pneumonia, while the presence of effusion has a slight impact on specificity and may increase the number of false positive pneumonia classifications.

Clinical Impact of Performance:

2. Algorithm Design and Function

« Insert Algorithm Flowchart »

DICOM Checking Steps:

Preprocessing Steps:

CNN Architecture: The CNN algorithm which is used is VGG16

3. Algorithm Training

Parameters: * Types of augmentation used during training * Batch size * Optimizer learning rate * Layers of pre-existing architecture that were frozen * Layers of pre-existing architecture that were fine-tuned * Layers added to pre-existing architecture

« Insert algorithm training performance visualization »

« Insert P-R curve »

Final Threshold and Explanation: CNN models output a probability ranging from 0-1 that indicates how likely the image belongs to a class. We will need a cut-off value called threshold to assist in making the

decision if the probability is high enough to belong to one class. Recall and precision vary when a different threshold is chosen.

4. Databases

(For the below, include visualizations as they are useful and relevant)

Description of Training Dataset:

Description of Validation Dataset:

5. Ground Truth

1. Asking several radiologists to label images and take the majority vote
2. Using a biopsy sample
3. NLP-extracting labels from reports

Gold standard The gold standard for a particular type of data refers to the method that detects disease with the highest sensitivity and accuracy. Any new method that is developed can be compared to this to determine its performance. The gold standard is different for different diseases.

Ground truth Often times, the gold standard is unattainable for an algorithm developer. So, you still need to establish the ground truth to compare your algorithm.

Ground truths can be created in many different ways. Typical sources of ground truth are

Biopsy-based labeling. Limitations: difficult and expensive to obtain. NLP extraction. Limitations: may not be accurate. Expert (radiologist) labeling. Limitations: expensive and requires a lot of time to come up with labeling protocols. Labeling by another state-of-the-art algorithm. Limitations: may not be accurate. Silver standard The silver standard involves hiring several radiologists to each make their own diagnosis of an image. The final diagnosis is then determined by a voting system across all of the radiologists' labels for each image. Note, sometimes radiologists' experience levels are taken into account and votes are weighted by years of experience.

The silver standard approach of using several radiologists would be more optimal for Algorithm A because I gave you the hint that it's really hard for radiologists to agree on breast density labels.

For Algorithm B, a single radiologist's labels would probably suffice, because I gave you the hint that they're really good at labeling 'normal' v. 'abnormal.'

6. FDA Validation Plan

Patient Population Description for FDA Validation Dataset:

Ground Truth Acquisition Methodology:

Algorithm Performance Standard: