Context learning for pneumonia detection on Chest Radiographs

Stanley Liang, Zhiyun Xue, Sivarama Rajaraman, Niccolo Marini, Sameer Antani

IRD, NLM, NIH

The context learning algorithm for pneumonia detection is an ensemble model that provides pixel level detections of lung opacities from front chest x-ray images (CXR). Our goal is to minimize the impact of irrelevant pixels such as scapulars, clavicles, and jaw bone with high intense pixel values, which usually caused overfitting. It improves the model performance by the proper learning context. The algorithm consists of four neural networks: a YOLOv11n classifier, a YOLOv9s detector for lung regions, a YOLOv9s detector for opacities, and a MobileSAM for further extract the lung tissues.

The classifier was trained by 1,000 MIDRC CXR images in two classes: 500 COVID-19 positive images with mRALE score from 18 to 24 as positive, and another 500 images with zero mRALE score as negative. The YOLOv9s lung regions was trained by 800 annotated images to detect three regions: right lung, left lung, and lungs and heart region. The YOLOv9s lung opacities detector is a weak detector trained with 200 annotated images with lung opacities on both lung regions. The MobileSAM is a pretrained lightweight SAM model to extract the lung-relevant pixels based on the bounding boxes detected by the YOLO models.

During the inference, the input DICOM image is first converted to PNG format, and the YOLO region detector will respectively detect, crop, and generate the right lung, the left lung, and the lungs and heart region (close-up image) images. The close-up image will be sent to the classifier for image level classification. If it is classified as negative with a confidence score less than 0.9, it is considered negative. A likelihood score of (1-confidence, as likelihood for positive) and an all-zero pixel level classification map will be output. If it is classified as positive or as negative with a confidence score greater than 0.9, it is considered positive. Next, we use the YOLO opacity detector to detect the pneumonia opacities respectively from the right lung and the left lung images, followed by the MobileSAM to further extract and merge the lung relevant pixels to form the masks for the predicted pneumonia opacity regions on the whole image. An Eigen GradCAM feature map is generated based on the outputs of the classifier limited to the lungs and hear region of the input image. The pixel values are rescaled to (0, 255) and binned to four values: 0 if the raw output is least than 8, 85 if the output is between 8 and 64, 170 if the output is between 64 and 179, and 255 if the output is greater than 255. Finally, the output will be a likelihood score for positive and a pixel level prediction map with the positive pixels calculated by the Eigen GradCAM and trimmed by YOLO and MobileSAM.

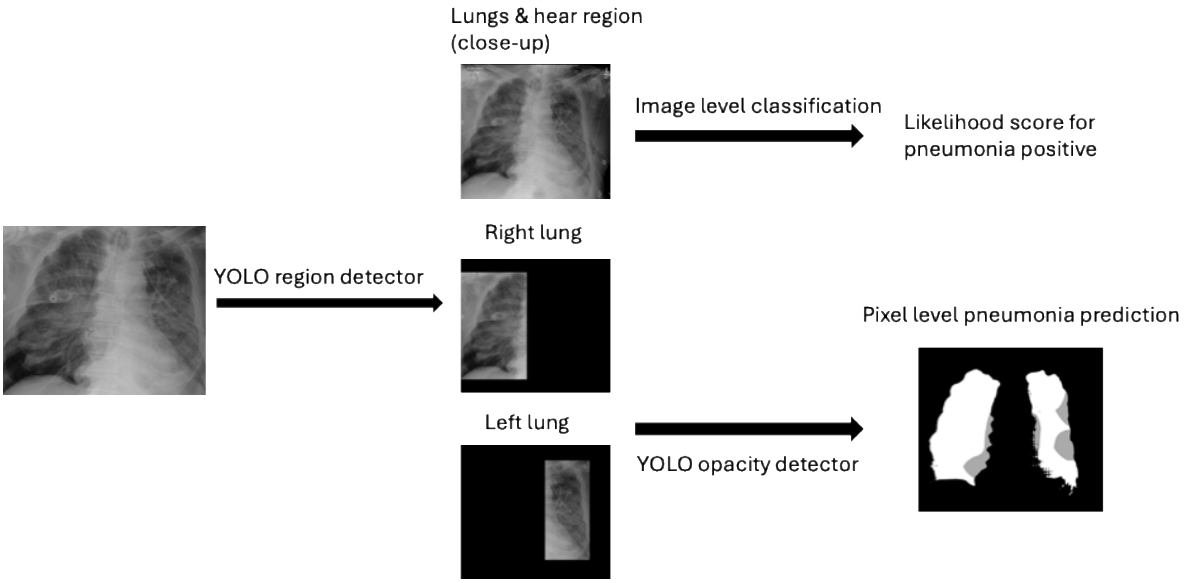

Fig 1. Prediction flowchart