Algorithm Description

Version: xai_submission_f1

Models

The algorithm uses a combination of models, including PSPNet, YOLOv5, and a custom segmentation model named PneumoniaSegNet, to conduct multi-stage image analysis and generate mask predictions for pneumonia.

- **PSPNet**: Pretrained model from torchxrayvision. Used for segmenting key anatomical regions such as the lungs, heart, and spine, creating a region of interest (ROI) where pneumonia is likely to be present.
- **YOLOv5**: YOLOv5 models (RSNA, QATA, VinBrain) are used for pneumonia detection in different regions.
- **PneumoniaSegNet**: A custom segmentation model based on the nnUNet backbone, trained to obtain pneumonia segmentation in the lung region.

A combination of results from the detection and segmentation models are ensembled to obtain the final result, which is described step-by-step in detail below.

Workflow

1. Load

DICOM images are read and normalized using min-max normalization. Then, black edges are removed, and the remaining image is resized to 512 × 512.

2. Process

The workflow starts by reading DICOM images from the input directory.

1. **Mask of ROI**: PSPNet is employed to segment key anatomical regions such as the lungs, heart, and spine. The output masks undergo horizontal and vertical dilation to refine their boundaries, followed by combining lung and heart masks to create a region of interest, while excluding the spine.

2. **Prediction (Detection)**: Three YOLOv5 models, each trained for pneumonia detection on different datasets (RSNA, QATA, VinBrain), are used to detect pneumonia. Each model processes the 512 × 512 black-edge-removed image, and bounding boxes are converted to masks by accumulating 0.5 for each detected bounding box.
3. **Prediction (Segmentation)**: Due to early-stage setup, the segmentation model only works for the cropped lung region, which is obtained using the lung masks generated by PSPNet. The lung region is then resized to 512 × 512. PneumoniaSegNet, based on the nnUNet backbone, is used for refined pneumonia segmentation in the lung region. The lung segmentation mask is used to generate a bounding box, which is filled with the ratio between the area of segmentation as positive vs. area of the bounding box to cover more positive regions.

4. Ensemble and Post-Processing

The predictions from all models are combined to form a probability map (pmap) = (0.3 * bbox + 0.3 * segmentation + 0.4 * (bbox * segmentation)). The probability map is further refined using dilation and multiplication, where pmap values greater than 0.8 are set to 1, and other values greater than 0.8 are divided by 0.8. These refinements are tuned based on the training data. Finally, the scores are calculated as the summation of the probability map, representing the probability of pneumonia. The scores are not normalized, as they are used to calculate the AUC, where only the order of values matters.

Finally, the predictions are saved as PNG files, and the scores for each image are saved in a CSV file for further analysis. The script efficiently combines segmentation, object detection, and pneumonia classification to provide detailed insights from medical images.