**Introduction**

Pneumonia is one of the leading causes of death in the U.S. Timely, accurate diagnosis is a critical factor in determining patient outcomes. Currently, pneumonia diagnosis is made by highly trained clinical experts interpreting chest radiographs (CXR) and laboratory exams. Areas of increased opacity are usually clear indicators for the presence of pneumonia, because the fluid, characteristic of pneumonia, preferentially attenuates the x-ray beam and therefore appears more opaque than the surrounding area. Our proposed model takes in chest radiographs as input and outputs bounding boxes localized to opaque regions indicating presence of pneumonia. An autonomous method for accurately identifying cases of radiological evidence for pneumonia would speed diagnosis time and hopefully reduce the number of deaths caused by pneumonia.

**Model Architecture**

The main compared were YOLOv3 and RetinaNet. RetinaNet uses ResNet51 backbone. YOLOv3 uses DarkNet53 as backbone. Both models are one-stage detectors where one optimizes speed and the other is lauded for its high accuracy/focal loss framework. Much of our investigations centered around recovering similar accuracy using YOLOv3 as the main component for bounding box prediction.

**Discussion**

- The NH weight initialization improved performance of the CheXnet F1 score and convergence speed.
- The category No OpaqueNet normal produced the most misclassification errors.
- Chairing YOLO and CheXNet showed decreased performance the most likely due to reduced training size and noisy training space (mix of positive and negative).
- Chairing YOLO and CheXNet increased train time considerably.
- We were unable to effectively compare class activation maps of the various models to each other and will be left for future work.
- RetinaNet + Yolo ensemble: The state of the art retinanet model returned a much higher AP IOU score.

**Dataset & Features**

The chest radiographs and the corresponding bounding boxes are provided by the Radiological Society of North America (RSNA) via the Pneumonia Detection Kaggle challenge. The dataset contains approximately 33,000 unique patient IDs labeled as 31% with opacity, 43% no lung opacity (normal), and 24% other (not normal, no opacity). The images are stored in dicom format at 1024x1024 solution. The images have been converted to .jpg and scaled down (various sizes) for further analysis. We have split the data into an 80/10/10 train/dev/test split.

**Comparison of Loss**

CheXnet Classification Loss

\[ L(C) = \sum_{i=1}^{n} \log(1+e^{-y_i}) \]

Improved classification performance comes from the deep layer architectures rather than loss which is just log loss on the 13 class predictions.

YOLO Loss vs. Focal Loss

\[ L_{YOLO} = \sum_{i=1}^{n} \sum_{j=1}^{k} (I_{cell_{ij}} - \hat{I}_{cell_{ij}})^2 \]

Both YOLO and RetinaNet are one-stage detectors in which classification and object localization occur in the same model. The difference in performance is largely due to calculation of loss.

In YOLOv3 the loss is calculated as the cross entropy loss of the class probabilities and the confidence as well as the error in the overlap of the predicted bounding boxes. In contrast focal loss now weights negative and positive images differently so that a model isn’t incentivized to maximize reward classification of ‘easy’ negatives.

**Workflow Optimization**

- Trained models on only positive images to reduce train time
- Scaled image resolution to reduce train time
- Concatenated positive/negative images to improve performance

**Future Work**

Although CheXnet + YOLO boosted accuracy, DenseNet121 is expensive to train. Future work could be:

- YOLO with cheaper classifiers that maintains accuracy
- All model class activation map comparison
- Local pooling
- Crop image based on box distribution
- Focus on No Lung Opacity / Not Normal prediction - predicting 3rd class, concatenating train images, etc.
- Dropout regularization to sync classification and bounding box prediction

There is value in a cheap algorithm like YOLO that could be run in near real-time with low resolution data for areas with limited access to computational resources.

**References**

5. DarkNet - See paper for paper details.