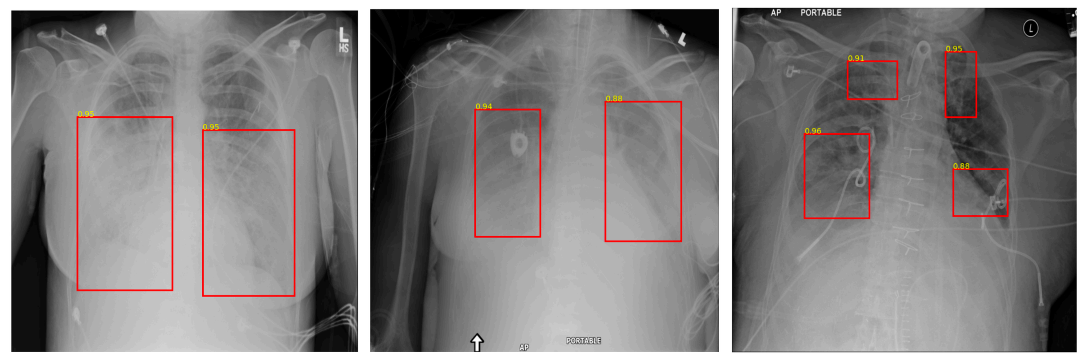
Pneumonia Detection Through Fine-Tuning FRCNN and FCOS Models

Colin Sullivan, Surya Sanapala
Stanford CS230 Final Project, Department of Computer Science

Overview

Problem:

- Computer vision model that can detect pneumonia in chest radiographs and draw a bounding box around the location of infection.



Approach:

- Fine tuned Faster Region-Based Convolutional Neural Network (FRCNN) with Custom Anchors and Generalized Intersection over Union (GIoU) Loss.
- Fine tuned Fully Convolutional One-Stage (FCOS) Object Detection with Weighted Loss.
- Implemented Data Augmentations.

Findings:

- Best model was FCOS + Weighted Loss

Overall Findings

	Model	mAP at 50%	mAP at 50%:95%	Recall_100
1	Baseline	76.47%	55.74%	60.65%
2	FRCNN + Custom Anchors + GIoU	**80.52%**	69.72%	74.21%
3	FCOS + Weighted Loss	80.21%	**71.25%**	**78.24%**
4	FCOS + Weighted Loss + Data Augmentation	58.05%	25.0%	51.56%

Future Work

- Multi-GPU Training
- Fine-tune the DEtection and TRansformer (DERT) model, proposed by Meta AI.
- DERT uses uses a set-based global loss function that forces unique predictions through a transformer encoder-decoder architecture and bipartite matching.

References

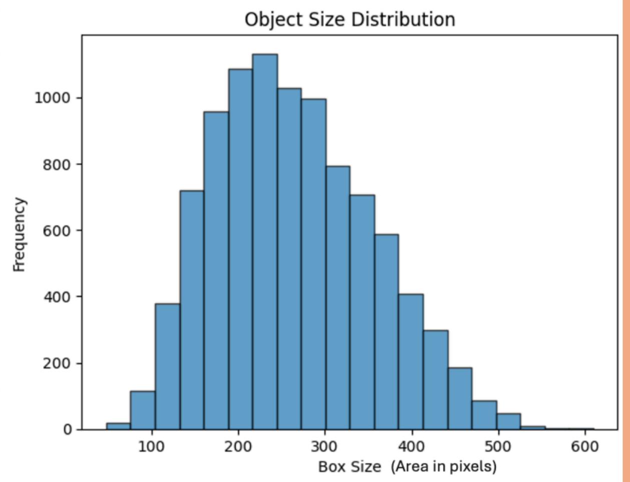
[1] Alapat DJ, Menon MV, Ashok S. A Review on Detection of Pneumonia in Chest X-ray Images Using Neural Networks. J Biomed Phys Eng. 2022 Dec 1;12(6):551-558. doi: 10.31661/jbpe.v0i0.2202-1461. PMID: 36569568; PMCID: PMC9759647.
[2] Anouk Stein, MD, Carol Wu, Chris Carr, George Shih, Jamie Dulkowski, kalpathy, Leon Chen, Luciano Prevedello, Marc Kohli, MD, Mark McDonald, Peter, Phil Culliton, Safwan Halabi MD, Tian Xia. (2018). RSNA Pneumonia Detection Challenge. Kaggle. https://kaggle.com/competitions/rsna-pneumonia-detection-challenge
[3] Ren, Shaoqing. "Faster r-cnn: Towards real-time object detection with region proposal networks." arXiv preprint arXiv:1506.01497 (2015).
[4] Tian, Zhi, et al. "Fully convolutional one-stage 3d object detection on lidar range images." Advances in Neural Information Processing Systems 35 (2022): 34899-34911.
[5] Wu, Linghua et al. "Pneumonia detection based on RSNA dataset and anchor-free deep learning detector." Scientific reports vol. 14,1 1929. 22 Jan. 2024, doi:10.1038/s41598-024-52156-7 [17] Chunhua Shen et al. "Fully Convolutional One-Stage Object Detection". https://arxiv.org/pdf/1904.01355

Dataset

RSNA Pneumonia Detection Challenge dataset [2]

- 26,685 chest radiographs
- For each label, if pneumonia is present, one or more sets of bounding box coordinates indicate location of the infection
- We randomly sample 6000 positive and 6000 negative examples from the dataset, then shuffle and divide this into 80% train, 10% validation, and 10% test

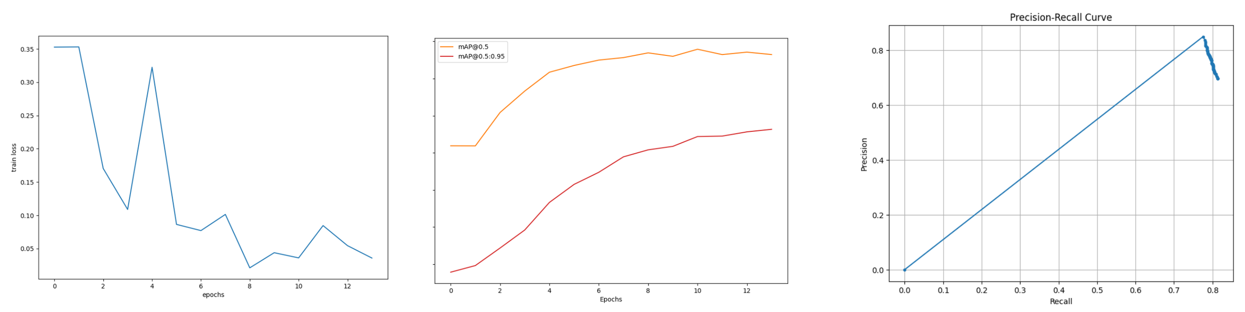
Description	Values
Positive Pneumonia Radiographs	6,012
Negative Pneumonia Radiographs	20,672
Total Positive Labels	9,555
Total Negative Labels	20,672
Bounding box distribution (small / medium / large)	0 / 86 / 30,141



FRCNN

Two-stage Object detection architecture

- Depends on predefined anchors of different sizes and aspect ratios generated by first stage RPN
- Between 2 stages it computes loss_objectness, loss_rpn_box_reg, loss_classifier,l oss_box_reg respectively.
- With RSNA dataset, gives a moderate performance but it is over confident in predicting high scores to the incorrect examples(shown in PR Curve) at higher thresholds

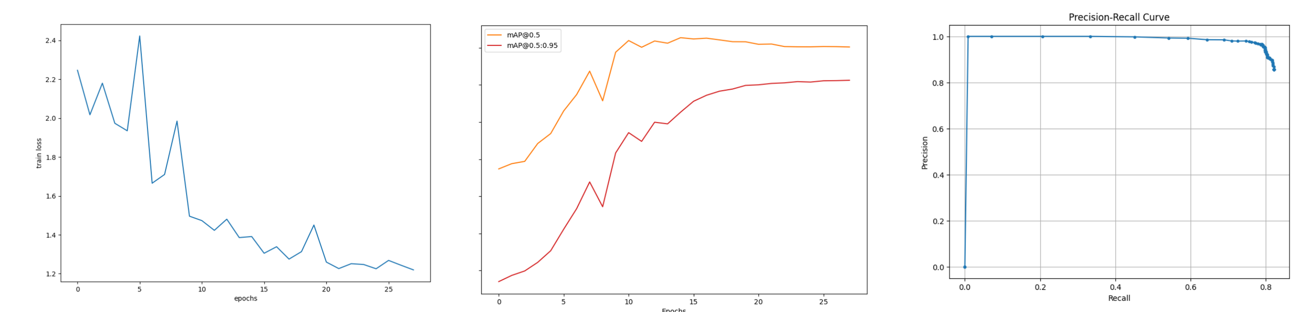


mAP@50	mAP@50:95	mAR_100	mAR_10	mAR_1
76.47%	55.74%	60.65%	60.65%	30.54%

FCOS with weighted loss

One-stage Object detection architecture

- Anchor free architecture and easy to train
- Uses Focal loss for classification and GIoU for regression
- Computes losses loss_cls, loss_reg ,loss_centerness.
- Fine tune weights for all 3 and find the weights that gives optimal performance
  - *losses = loss_cls + lambda_reg * loss_reg + lambda_cen * loss_centerness*
- Further enhancements can be done by fine turning values of alpha and gamma of Focal loss function



mAP@50	mAP@50:95	mAR_100	mAR_10	mAR_1
80.21%	71.25%	78.24%	78.24%	42.87%

Dataset → Evaluation → FRCNN → FRCNN w/ Custom Anchors and GIoU → FCOS with Weighted Loss → Data Augmenta-tions

Evaluation

- Calculate mAP at given Intersection over Union (IoU) thresholds.
- Calculate recall at 100 hits per image.
- Experimented with two evaluation implementations:
  1. Implemented custom functions where model predictions are filtered by a score threshold of 0.83, then IoU is calculated, and finally non-max suppression filters outputs with a threshold of 0.5
  2. Utilize Python library TorchMetrics

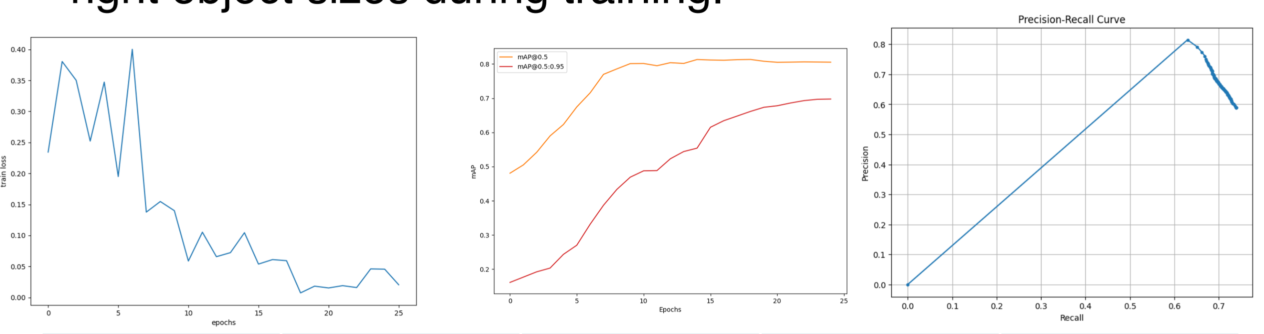
$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative}$$

$$Intersection\ over\ Union\ (IoU) = \frac{Area\ of\ Overlap}{Area\ of\ Union}$$

FRCNN With Custom Anchor Generator and GIoU reg loss

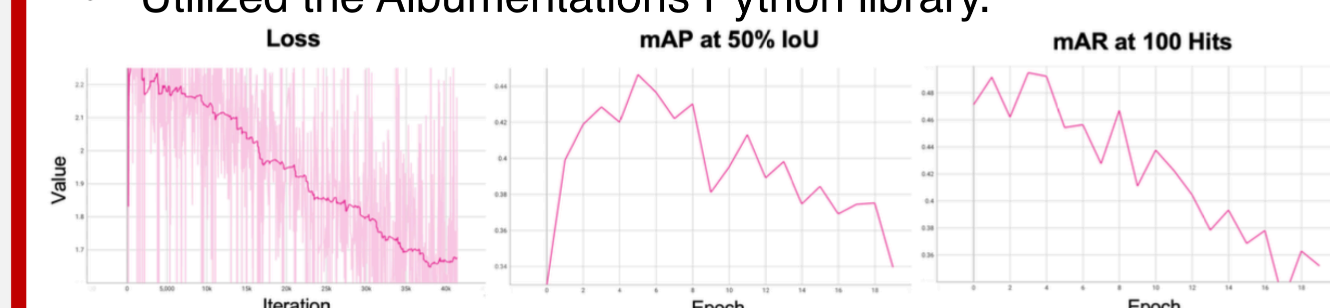
- Smooth L1 loss used for box regression does not give meaningful gradients when there is no overlap between predicted box and ground truth box.
- Replace Smooth L1 loss with GIoU loss which enhances learning process by providing meaningful gradients.
- Depending on the dataset distribution, fine tune the anchor sizes and aspect ratios so that RPN can focus on right object sizes during training.

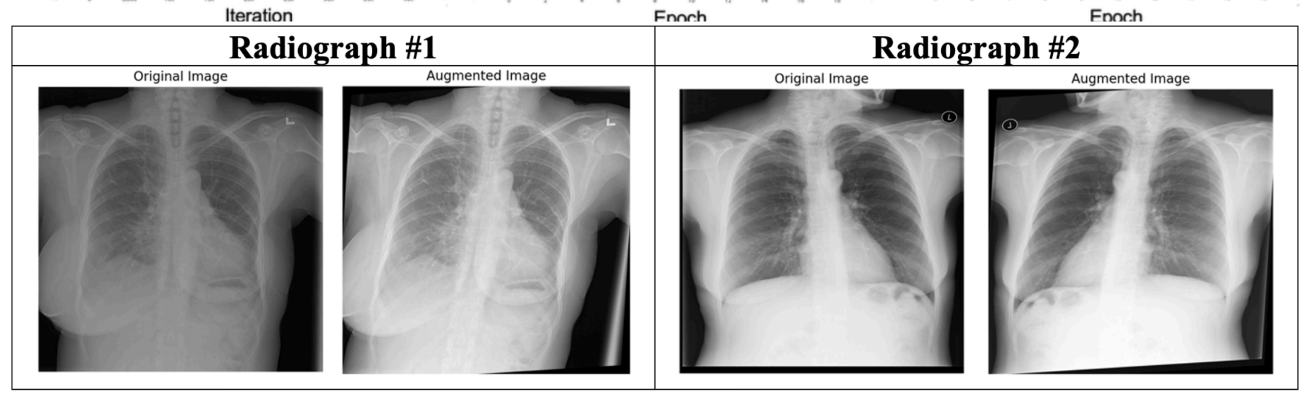


mAP@50	mAP@50:95	mAR_100	mAR_10	mAR_1
80.52%	69.72%	67.22%	67.22%	33.78%

Data Augmentation

- Five data augmentation techniques including horizontal/ vertical flips, Gaussian Blur, rotation, shear, and brightness/color adjustments to 10% of the training data.
- Utilized the Albumentations library.





mAP@50	mAP@50:95	mAR_100
58.05%	25.0%	51.56%