RSNA Pneumonia Detection Challenge

張〇〇、郭〇〇、邱泓儒

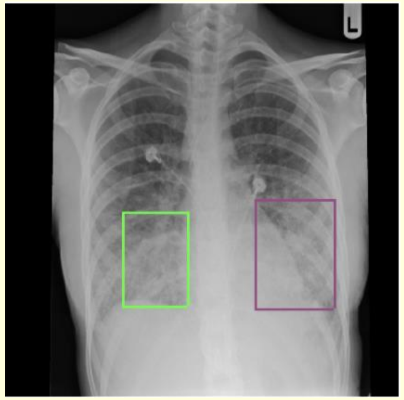
Outline

- Introduction
- Related work
- Proposed approach
- Experimental results
- Conclusions

Outline

- Introduction
- Related work
- Proposed approach
- Experimental results
- Conclusions

What’s our task?

F() = 

Input: an CXR(Chest X-Ray)
Output: bounding boxes of pneumonia regions

Structure Learning!

What’s the value of it?

Analyzing each radiographs ➔ time consuming

How a deep neural network help? ➔ **filtering** ➔ doctors spend time on other tasks

What’s a CXR?

Metal → whitest
Bone & Calcium(鈣) → white
Tissues & liquid → gray → Lung opacities
Air → black

Pneumonia on CXR?

Healthy lungs: full of air → black (except ribs are white and the heart is gray)

Pneumonia: liquid, bacteria, immune system cells → more gray(whiter)

Target: Find the gray/white regions supposed to be black!

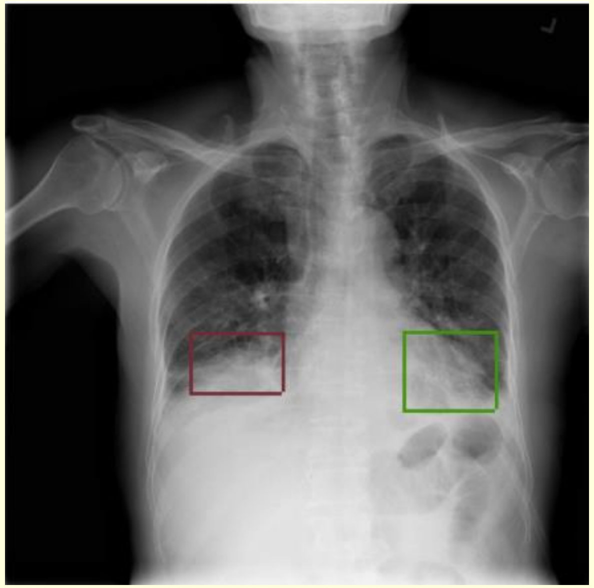
Any other issues?

- Opacities out of lungs?
  → segment the lung parts

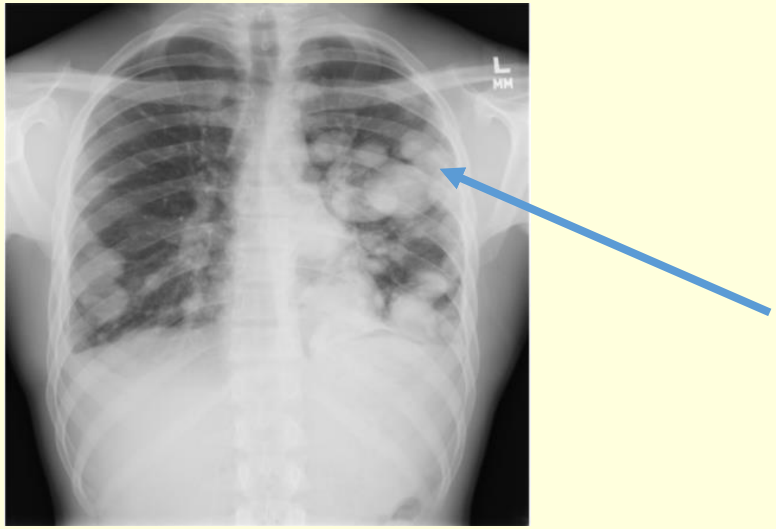
- All opacities in the lungs are pneumonia?
  → No, can be tumors
  → Distinguish?

Pneumonia vs Tumors

Pneumonia:
blurry, hazy, no clear shape

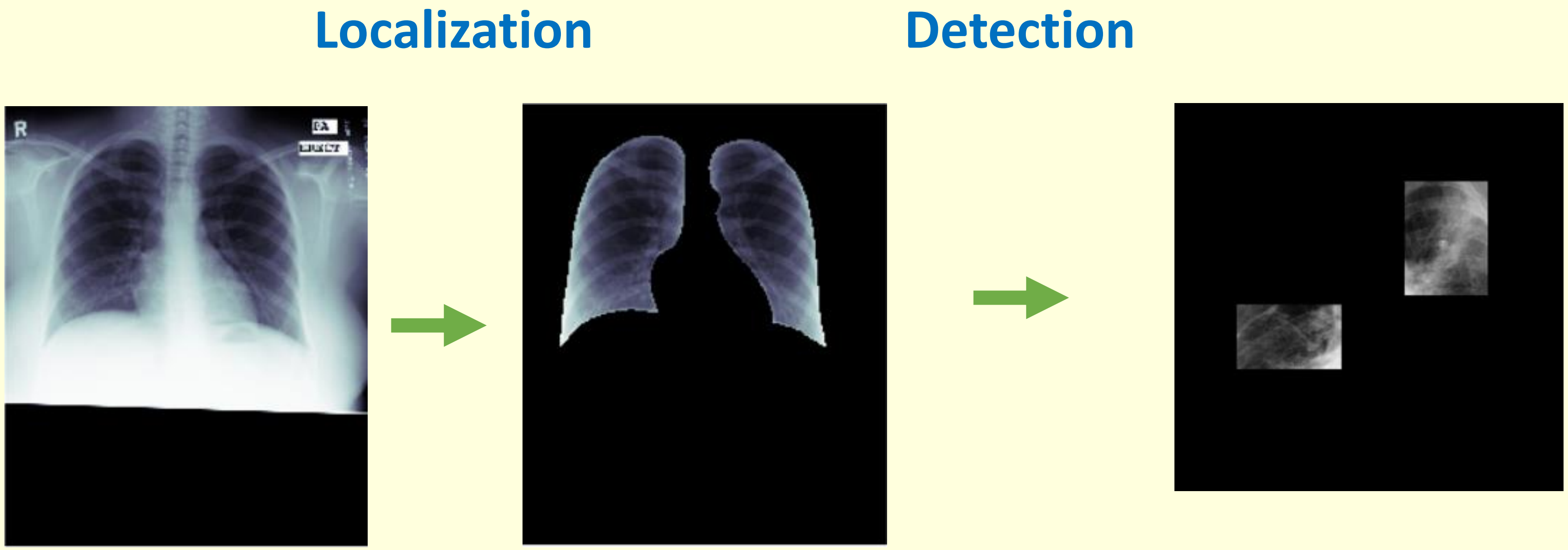


Tumor:
clear round shaped opacities



Design a proper approach?

2-step approach: **localization** & **detection**



Outline

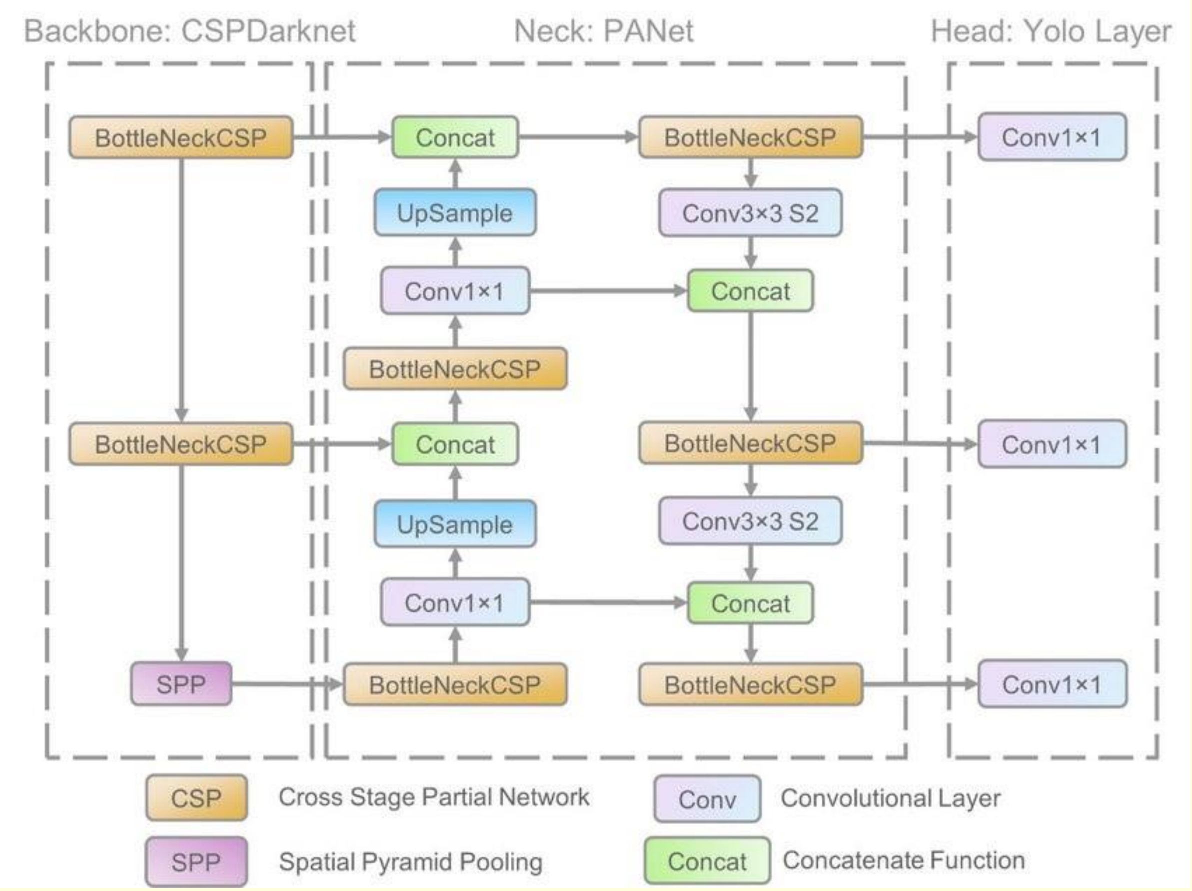
- Introduction
- Related work
- Proposed approach
- Experimental results
- Conclusions

Object detection

- Two stage detector
  - R-CNN, fast R-CNN, faster R-CNN
  - Region Proposals → Object Recognition
  - more accurate but require a higher computational time
- One stage detector
  - YOLO, SSD, RetinaNet
  - predict the bounding boxes and class probabilities directly
  - much faster but provide less accurate results

YOLOv5

- Backbone: feature extraction
- Neck: feature fusion (dealing with feature pyramids)
- Head (one-stage): object detector
- Activation function:
  - Layers in the middle → Leaky ReLU
  - Layers in the end → Sigmoid
- Number of layers: 191

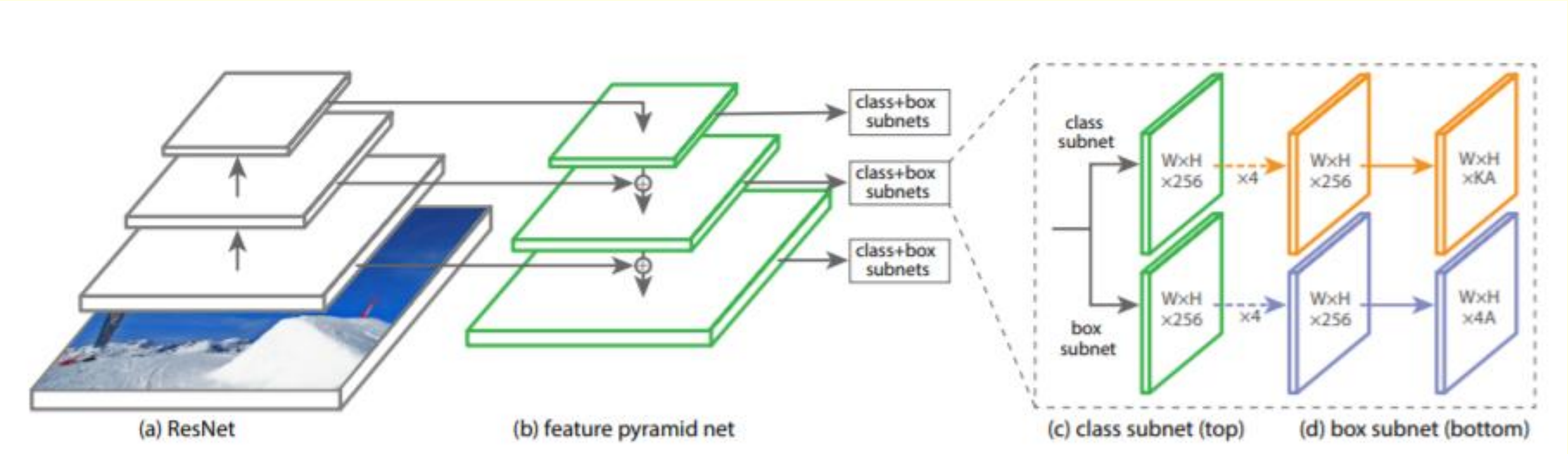


Outline

- Introduction
- Related work
- Proposed approach
- Experimental results
- Conclusions

RetinaNet - model structure

- Residual Network(ResNet)
- Feature Pyramid Network(FPN)
- Class Subnet and Box Subnet



RetinaNet - Loss function

- One stage detectors have worse accuracy than two-stage detectors. The authors think that the extreme class imbalance is the central cause
- They propose a new loss function：Focal Loss

Class imbalance problem

- One stage detectors generate 10000 - 100000 candidate boxes per image but only a few boxes contain objects.
- This causes two problems:
  - training is inefficient
  - may lead to degenerate models.

Focal Loss

$$\mathrm{CE}(p, y) = \mathrm{CE}(p_t) = -\log(p_t)$$

$$\mathrm{FL}(p_t) = -(1 - p_t)^\gamma \log(p_t).$$

$$\mathrm{FL}(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t).$$

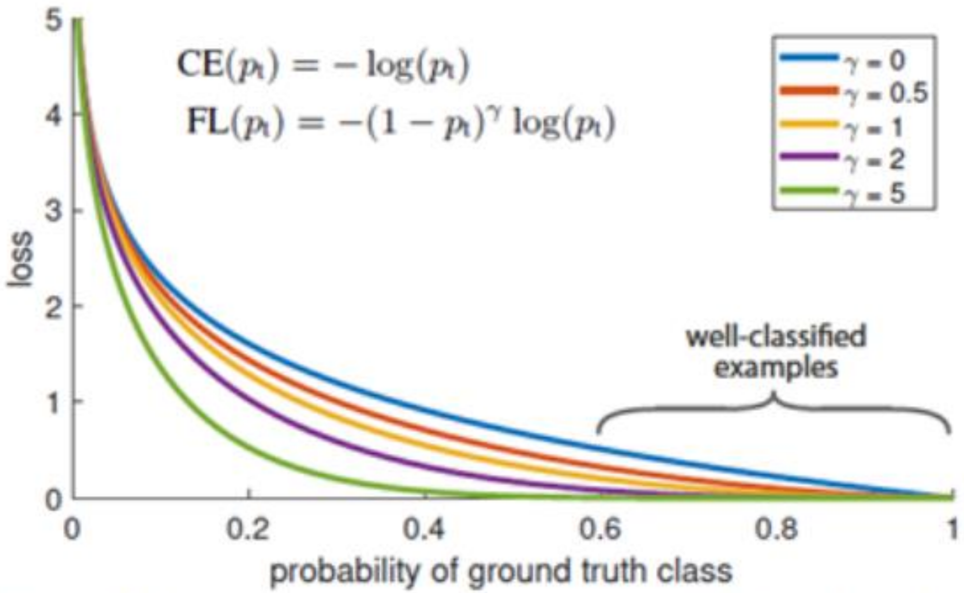


Figure 1. We propose a novel loss we term the *Focal Loss* that adds a factor $(1 - p_t)^\gamma$ to the standard cross entropy criterion. Setting $\gamma > 0$ reduces the relative loss for well-classified examples ($p_t > .5$), putting more focus on hard, misclassified examples. As our experiments will demonstrate, the proposed focal loss enables training highly accurate dense object detectors in the presence of vast numbers of easy background examples.

Pre-process

- Resize image to 224*224
- divide the 25684 training images into a large training set (24399 images, 95%) and a small validation set (1285 images, 5%)

Training

- resnet-50 and resnet-101 backbones
- SGD with learning rate = 0.01, momentum =0.9, decay = 0.0001
- 75 epochs with 2500 steps per epoch and batch size of 8
- Augmentation, such as：
  - Rotation
  - Translation
  - Scaling
  - horizontal flipping

Ensemble two models

- bounding box score threshold：
  - resnet-50：0.04
  - resnet-101：0.05
- overlapping bounding boxes in each network：NMS=0
- overlapping bounding boxes from both networks：take weighted averages
- bounding boxes that did not overlap between two networks：use a separate higher threshold value(=0.15)
- shrinking all final bounding boxes by a factor of 0.83

Outline

- Introduction
- Related work
- Proposed approach
- Experimental results
- Conclusions

YOLOv5

model	# of epochs	Bbox_th	Private score	Public score
YOLOv5s	10	1e-1	0.12432	0.02160
		5e-2	0.11700	0.03395
		15e-2	0.11157	0.00793
	20	1e-1	0.15505	0.08730
		5e-2	0.12931	0.03240
		15e-2	0.13562	0.08730
	30	1e-1	0.13893	0.08641
		5e-2	0.12057	0.06134
		15e-2	0.13076	0.02380
	40	1e-1	0.13858	0.04761

Highly possible that overfitting occurs

Different bounding box score thresholds

Model	BBox_th	Private score	Public score
ResNet50	0.05	0.15871	0.05709
	0.1	0.17121	0.07341
	0.15	0.16844	0.07341
	0.2	0.15372	0.07341
ResNet101	0.05	0.16476	0.04166
	0.1	0.16257	0.04861
	0.15	0.14555	0.05555
	0.2	0.13139	0.02380

Different NMS thresholds

Model	NMS_th	Private score	Public score
Ensemble	0	0.17532	0.05753
	0.1	0.17548	0.05753
	0.2	0.17585	0.05753
	0.3	0.17133	0.05753
	0.4	0.16598	0.05753

Our best model

Model	BBox_th	NMS_th	Private score	Public score
Resnet50	0.04	-	0.14698	0.04282
ResNet101	0.05	-	0.16476	0.04166
Ensemble	-	0	0.17916	0.05034

Outline

- Introduction
- Related work
- Proposed approach
- Experimental results
- Conclusions

Conclusions

- We try YOLOv5, RetinaNet with Resnet50 and Resnet101
- Our best model is the ensemble model of two RetinaNet with private score 0.179
- Choosing proper bounding boxes and NMS thresholds is important
- Ensemble model may be better when single model is not good enough

GitHub link of our code

- https://github.com/Lucas-Kuo/VR_DL_Final
- yahan1011/VRDL_FP (github.com)

Reference

- https://medium.com/@kimujan/retinanet-%E8%AE%80%E5%BE%8C%E6%95%B4%E7%90%86-af821dcd4558
- https://arxiv.org/pdf/1708.02002.pdf
- https://github.com/pmcheng/rsna-pneumonia
- https://www.researchgate.net/figure/The-network-architecture-of-Yolov5-It-consists-of-three-parts-1-Backbone-CSPDarknet_fig1_349299852
- https://jonathan-hui.medium.com/understanding-feature-pyramid-networks-for-object-detection-fpn-45b227b9106c
- https://medium.com/image-processing-and-ml-note/yolo-v5-explained-and-demystified-22f041ad316e
- https://medium.com/swlh/focal-loss-what-why-and-how-df6735f26616
- https://www.kaggle.com/zahaviguy/what-are-lung-opacities/notebook