Using Deep Learning for Breast Cancer Diagnosis

LYU1706

- Li Qi (1155062147) , Li Wei(1155062148)
- Supervisor: Prof. Michael R. Lyu
- CUHK CSE

01. Introduction

Introduction

01 Motivation

02 Background

03 Objective

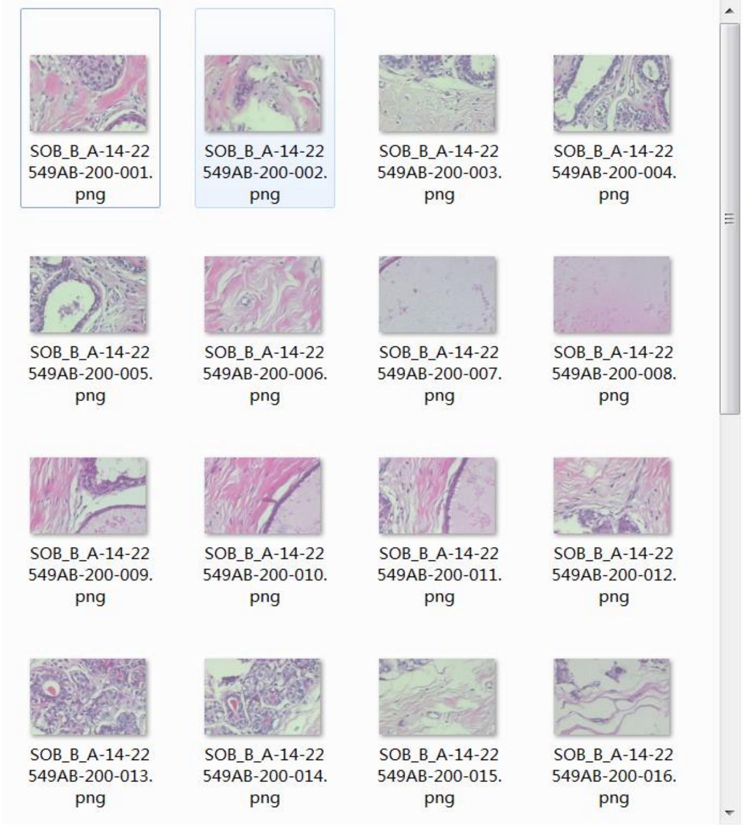
Introduction

01 Motivation

02 Background

03 Objective

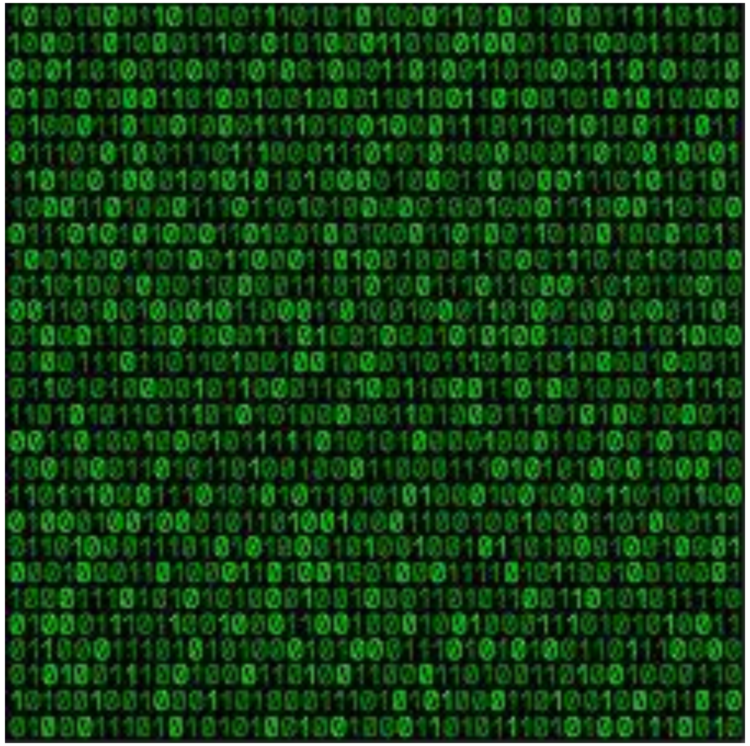
Introduction: Motivation



Breast cancer diagnosis

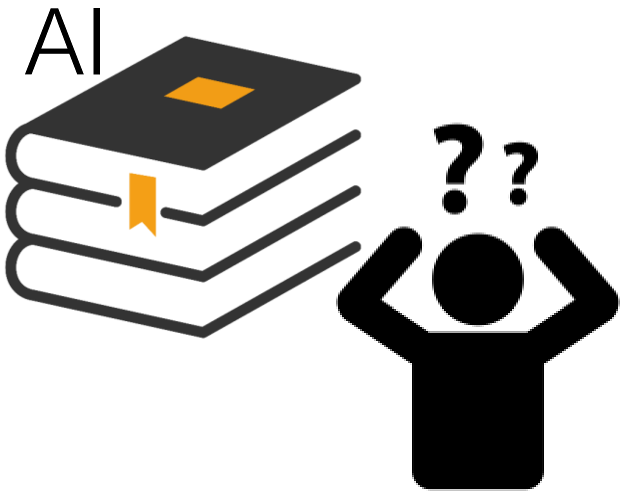
- **10+** gigapixels per patient
- agreement in diagnosis < **48%**

Introduction: Motivation



Current automatic diagnosis

- Statistics
- Jargons
- Codes



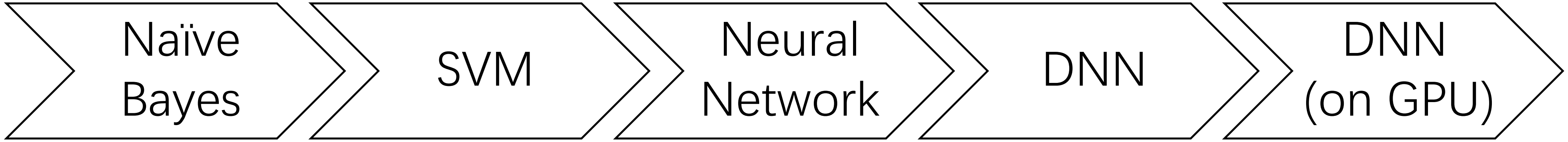
Introduction

01 Motivation

02 Background

03 Objective

Introduction: Background



Development of Classification

Introduction: Background

Feature Matching

Viola-Jones

RCNN

Fast RCNN

Faster RCN

Mask RCNN

Development of Object Detection

Introduction: Background

- Deep Learning for Medical Images
  - structures detection
  - segmentation
  - labeling and captioning
  - computer aided detection or diagnosis

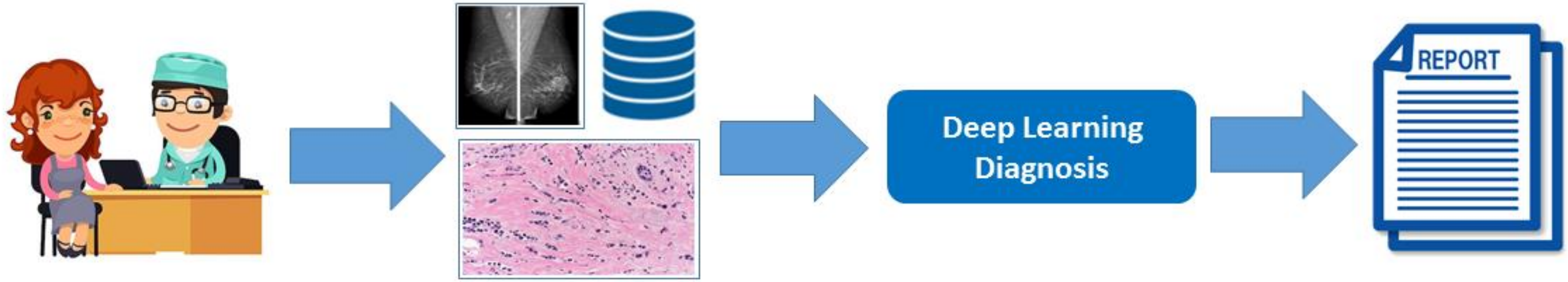
Introduction

01 Motivation

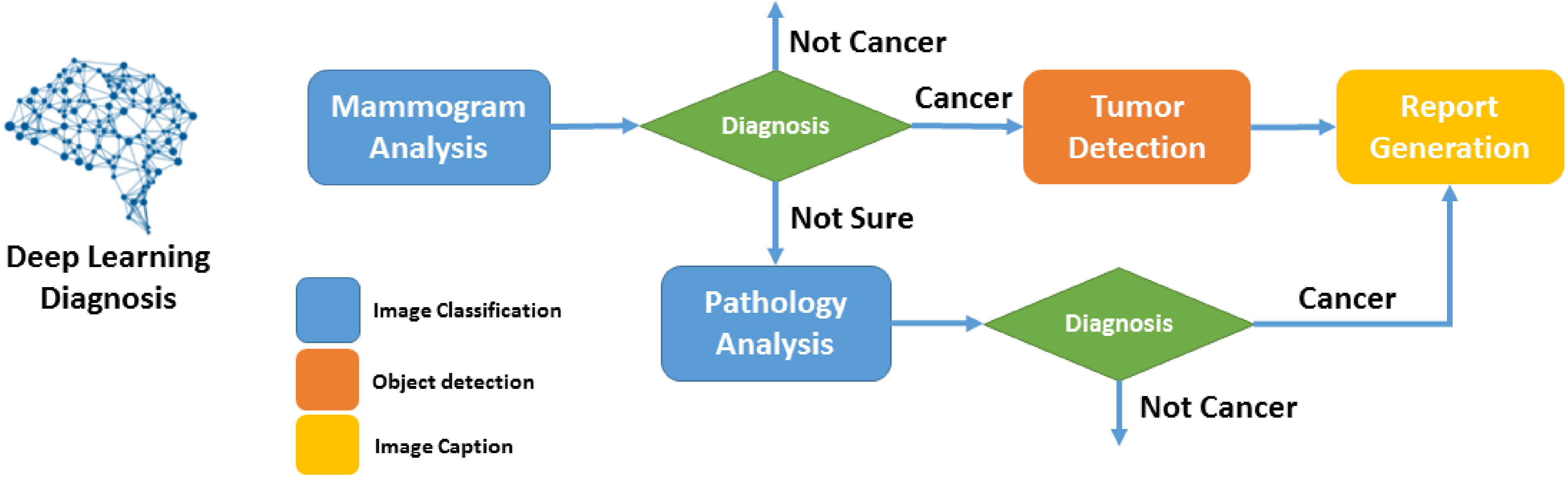
02 Background

03 Objective

Introduction: Objective



Introduction: Objective



02. Term One Review

Term One Review

01 Overview

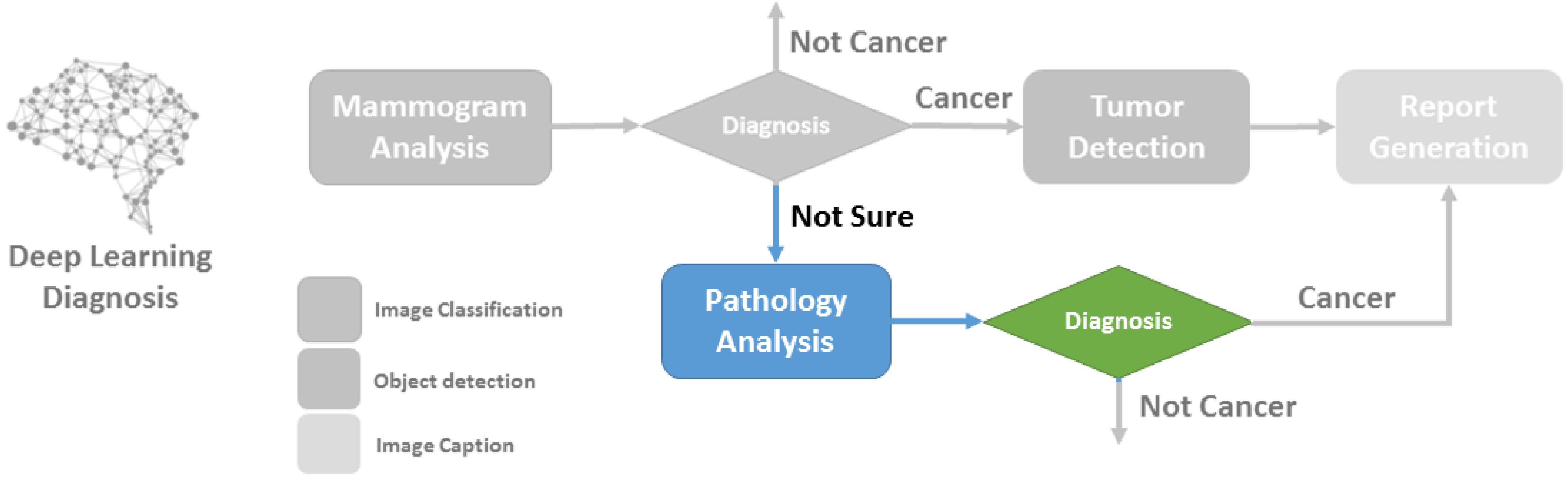
02 Dataset

03 Model Architecture

04 Result

Term One Review

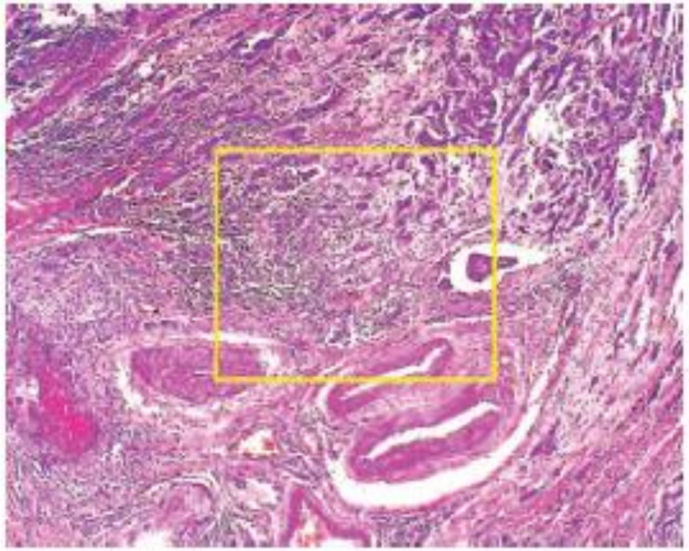
Term One Review: Overview

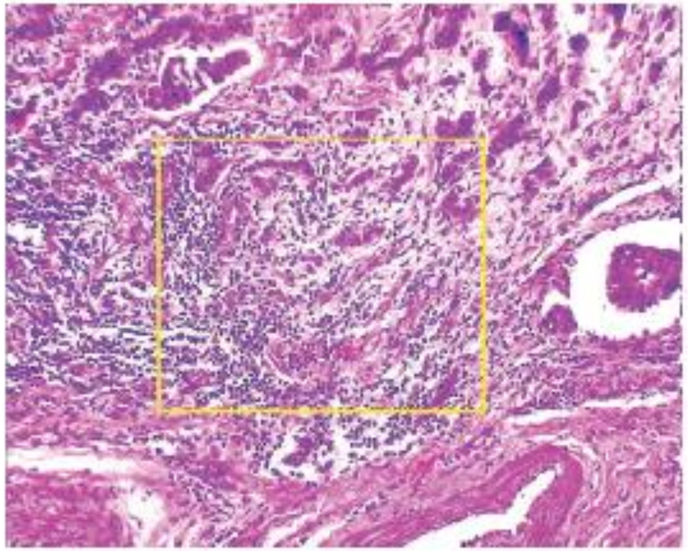


Term One Review

Term One Review: Dataset

Breast Cancer Histopathological Image Classification (BreakHis)





different magnifying factors (40x, 100x, 200x, and 400x)

Term One Review

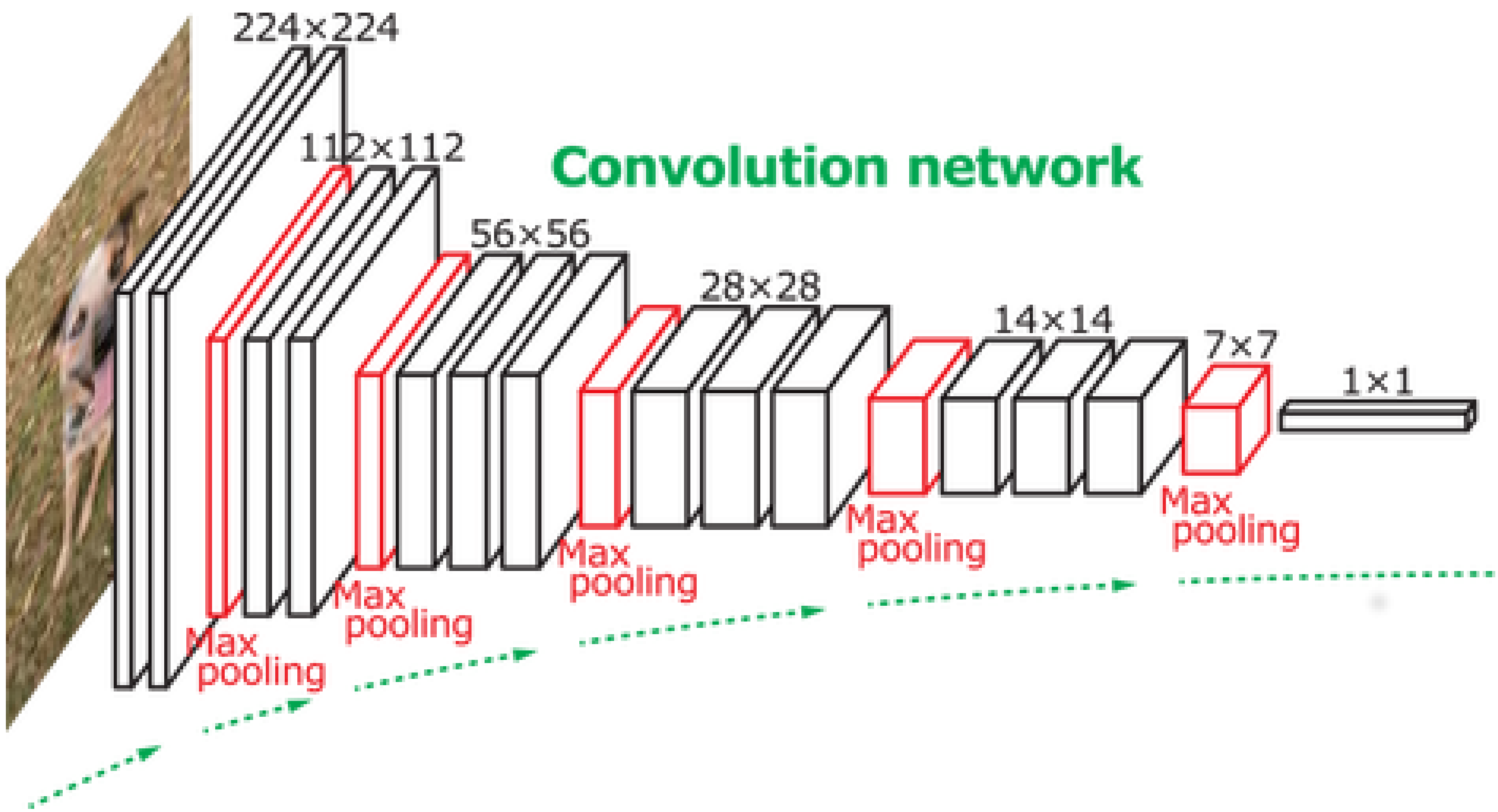
01 Overview

02 Dataset

03 Model Architecture

04 Result

Term One Review: Model Architecture



Term One Review: Model Architecture

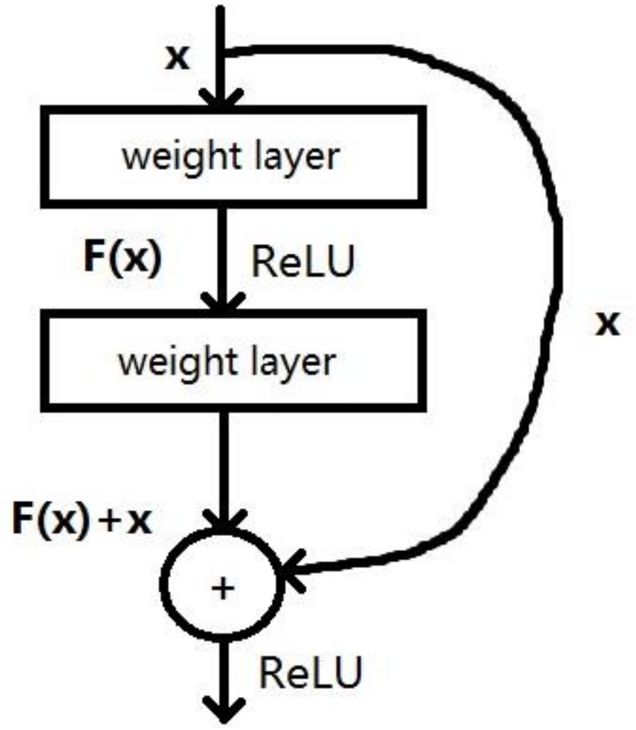
Residual Blocks: fix degradation problem

$$\mathrm{H}(\mathrm{x}) - \mathrm{x} \rightarrow \mathrm{F}(\mathrm{x})$$

$$\mathrm{H}(\mathrm{x}) = \mathrm{F}(\mathrm{x}) + \mathrm{x}$$

Term One Review: Model Architecture

Residual Blocks: fix degradation problem



ImageNet Large Scale Visual Recognition Challenge 2015 winner

Term One Review

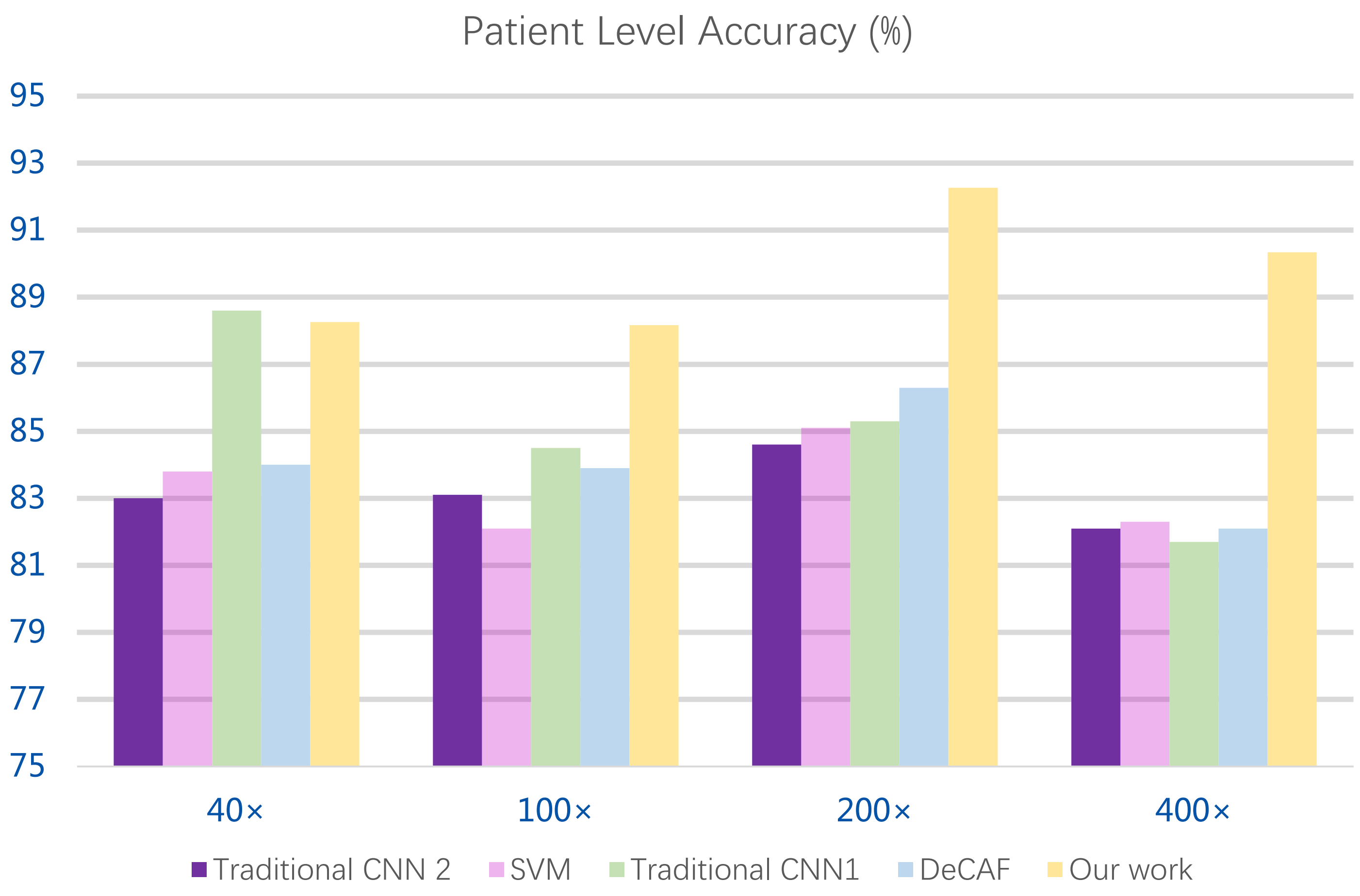
01 Overview

02 Dataset

03 Model Architecture

04 Result

Term One Review: Result



- Our work is **better** than other research using same dataset in **almost all of cases**
- The difference can be as large as **5%** in most cases.
- low magnification factors, such as 40× and 100×, has a **fewer information** and features for model to catch and learn

03. Literature Review

Related Work

01 Deep Multi-instance Networks with Sparse Label

02 Mass Segmentation via Cascaded Random Forests

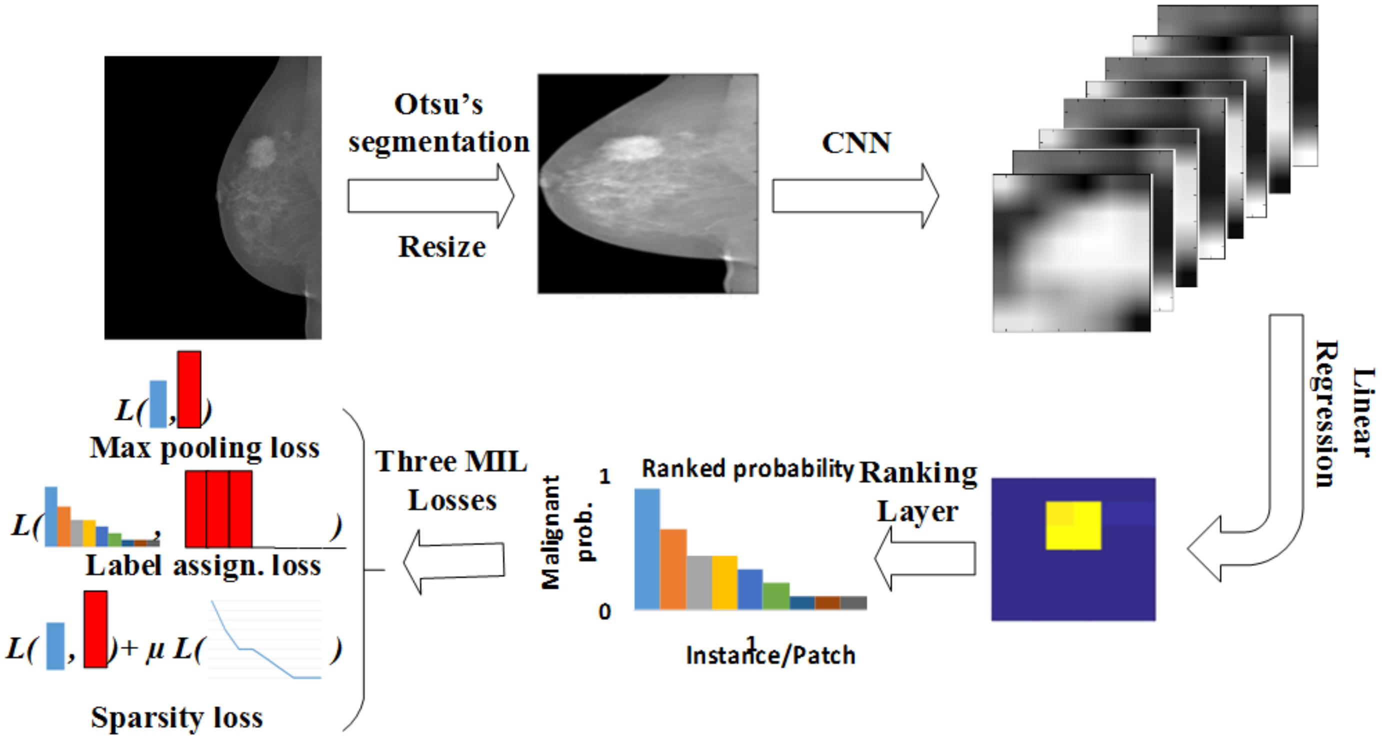
Related Work

01 Deep Multi-instance Networks with Sparse Label

02 Mass Segmentation via Cascaded Random Forests

Related Work: Deep Multi-instance Networks

- End-to-end network
- Multi-instance learning
  - Max pooling based loss
  - Label assignment based loss
  - Sparse loss
- Whole mammogram as input



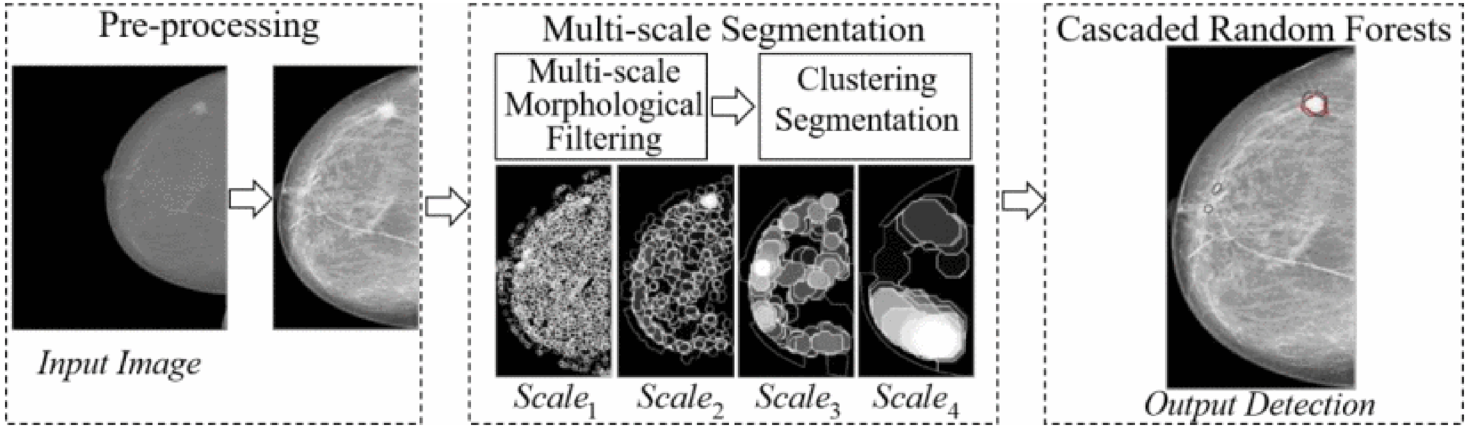
W. Zhu et al.

Related Work

01 Deep Multi-instance Networks with Sparse Label

02 Mass Segmentation via Cascaded Random Forests

Related Work: Cascaded Random Forests



- Filters at several scales
- Self-adjusting #layers
- Narrowing down false-positives

H. Min et al.

04. Method

Method

01 Dataset

02 Preprocess

03 Model Architecture

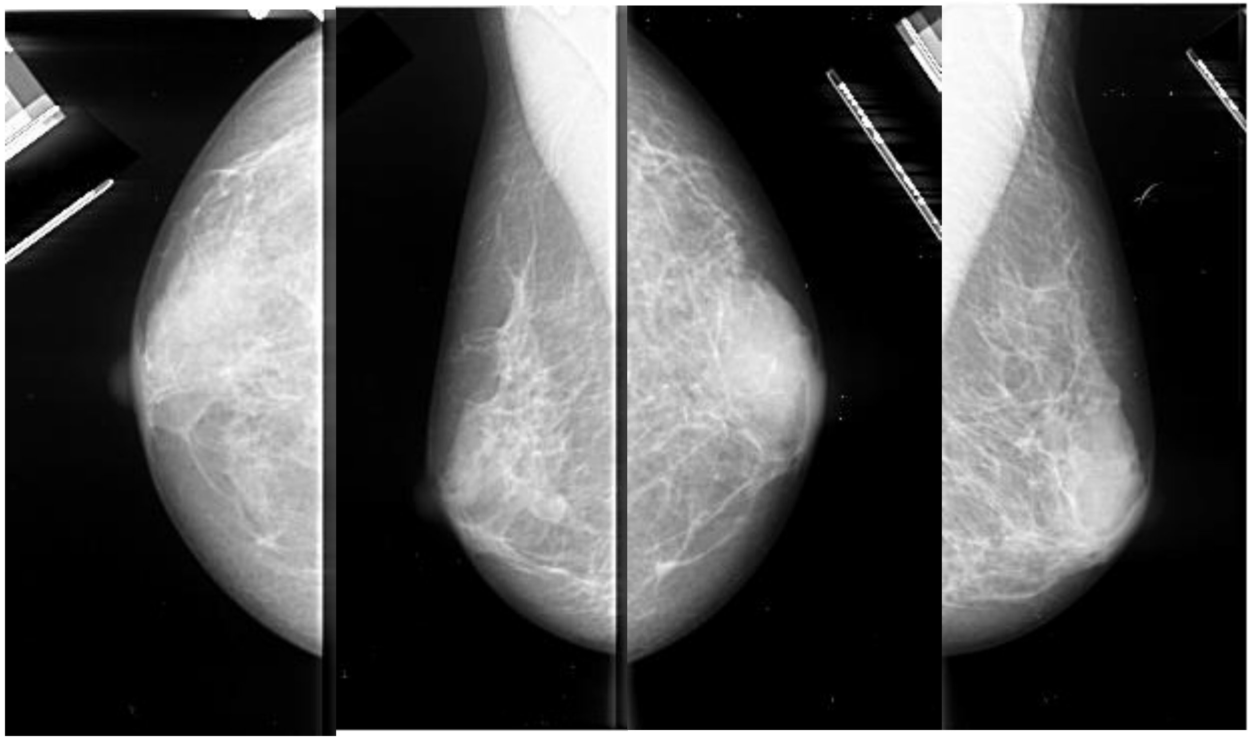
04 Loss Function

05 Evaluation

Method

Method: Dataset

Digital Database for Screening Mammography (DDSM)



two views of both side (left CC+MLO, right CC+MLO)

Method: Dataset

Digital Database for Screening Mammography (DDSM)

2479 cases	benign	870
	malignant	914
	normal	695

Method: Dataset

Digital Database for Screening Mammography (DDSM)

rich meta information

Time of study: 5 3 1991

Patient age: 63

Scanner resolution: 42

Keyword description: 2

...

Method

01 Dataset

02 **Preprocess**

03 Model Architecture

04 Loss Function

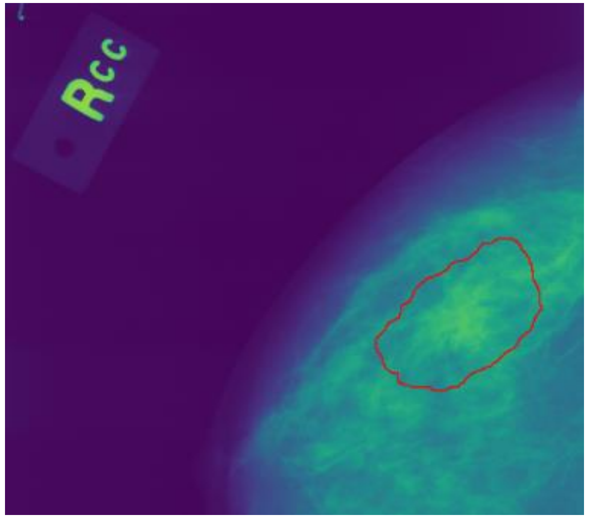
05 Evaluation

Method: Preprocess

01 LJPEG and Chain Code

Idea: decompress the data

	X→		
Y ↓	7	0	1
	6	·	2
	5	4	3

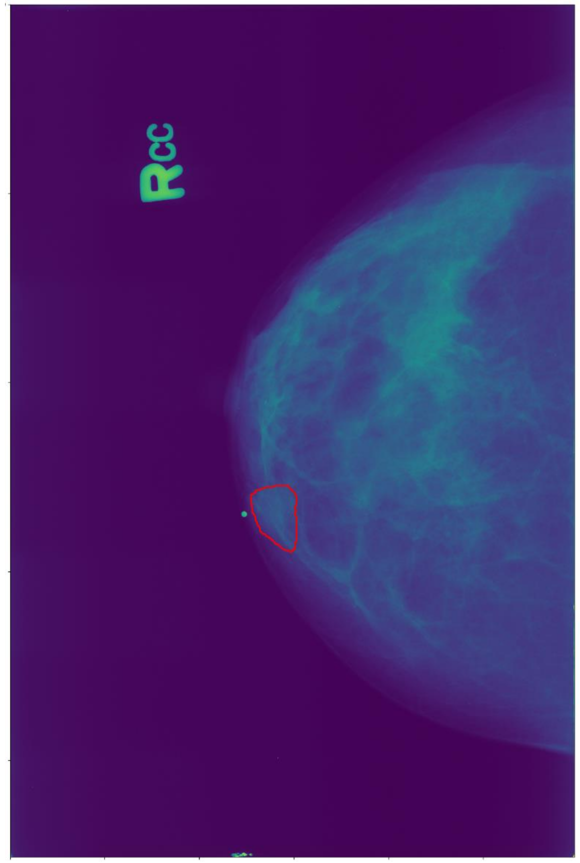


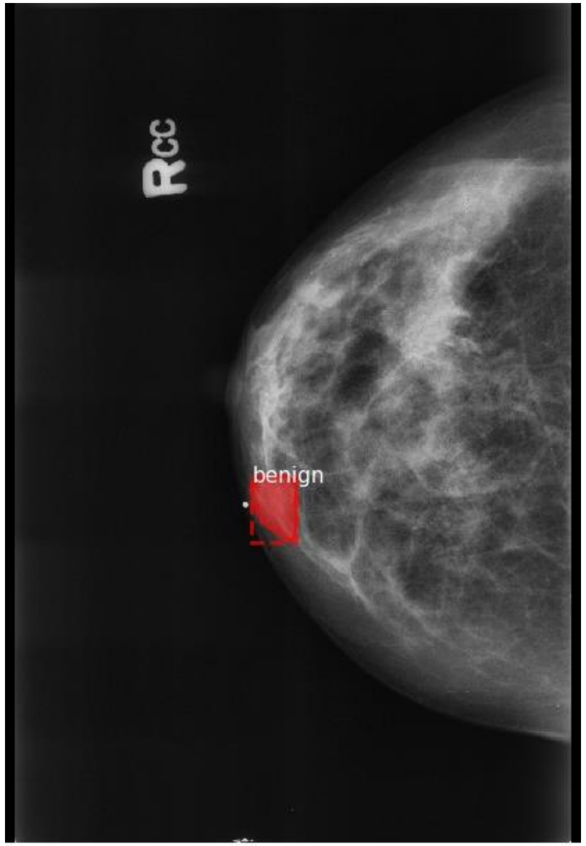
Method: Preprocess

01 LJPEG and Chain Code

02 Contrast Limited AHE

Idea: make image clearer





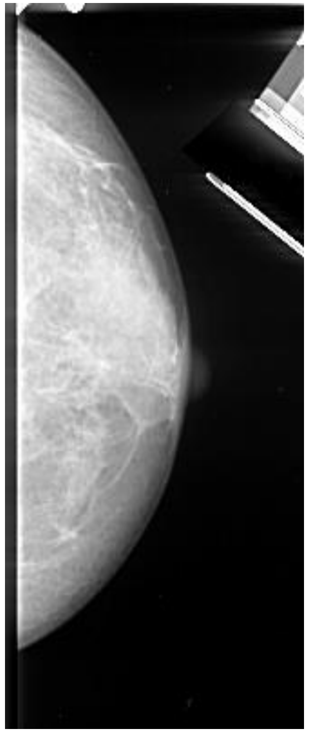
Method: Preprocess

01 LJPEG and Chain Code

02 Contrast Limited AHE

03 Image Augmentation

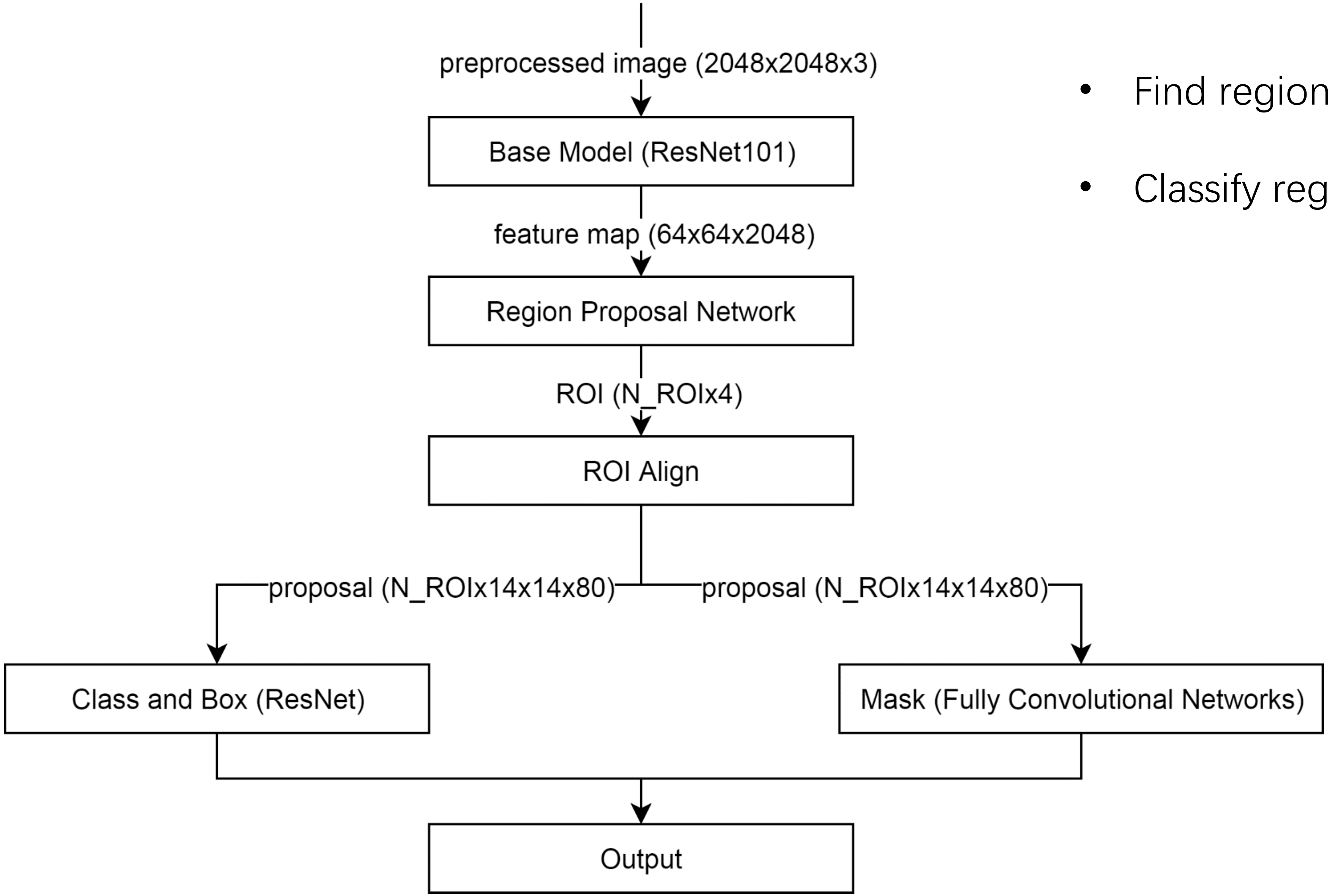
Idea: make dataset larger



Method

01	Dataset
02	Preprocess
03	Model Architecture
04	Loss Function
05	Evaluation

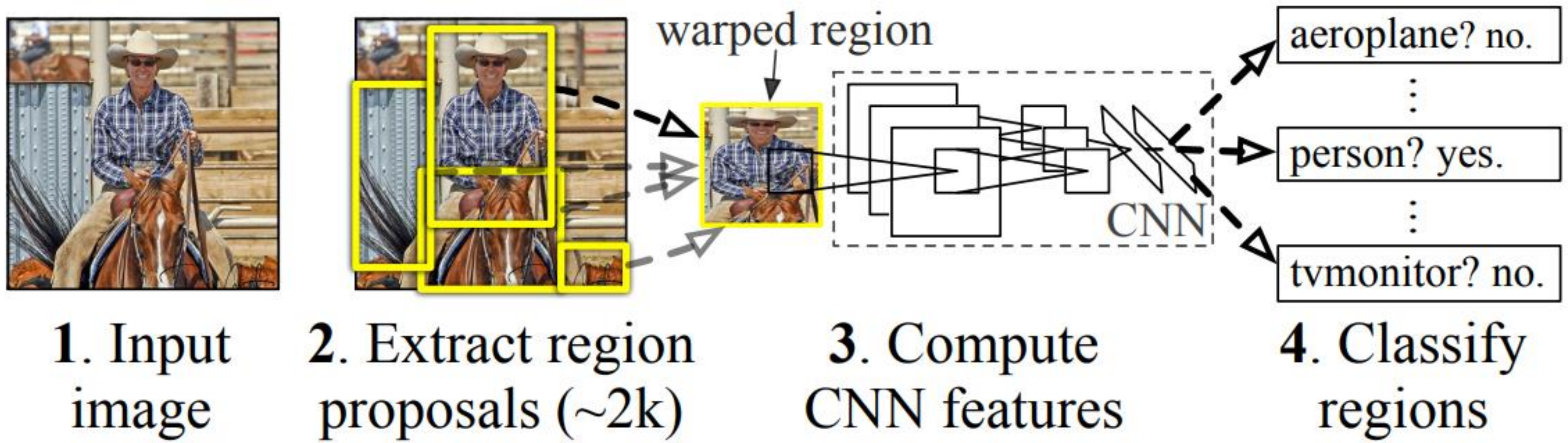
Method: Model Architecture



- Find region proposals
- Classify region proposals

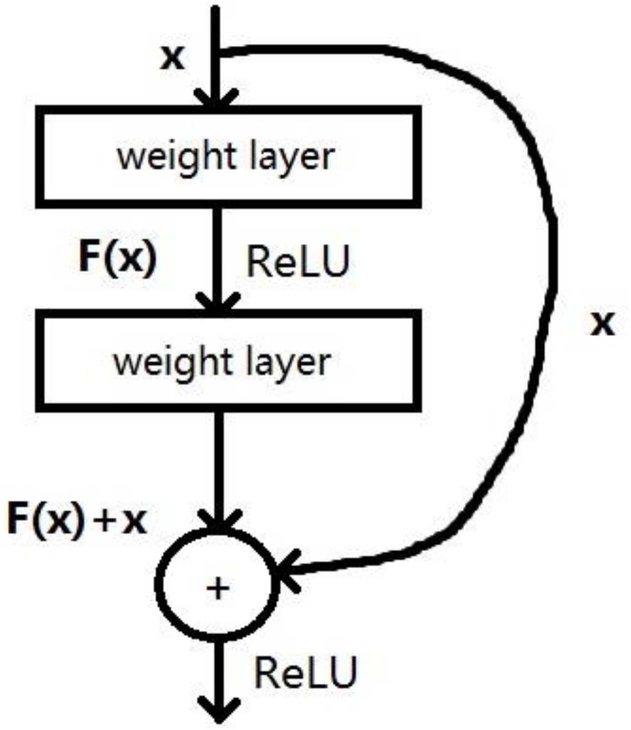
Method: Model Architecture

- Find region proposals
- Classify region proposals



Method: Model Architecture

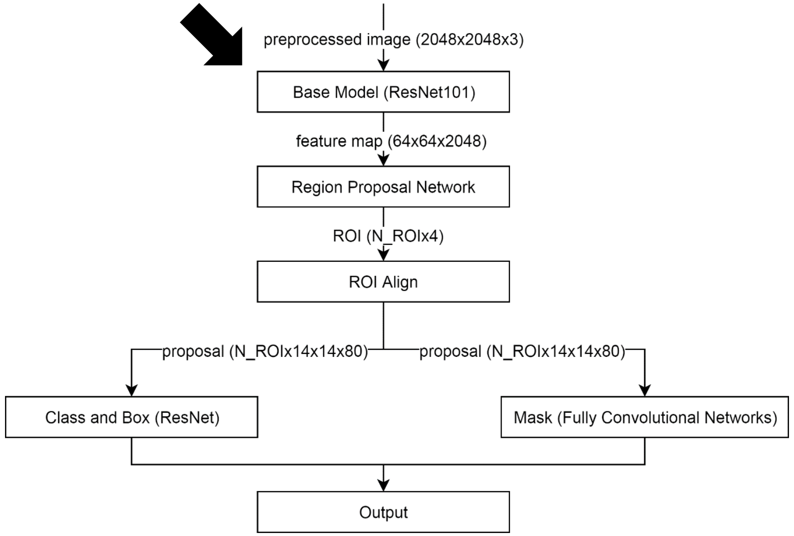
Residual Network: fix degradation problem



ImageNet Large Scale Visual Recognition Challenge 2015 winner

Method: Model Architecture

01 Base Model



ResNet 101 → feature map

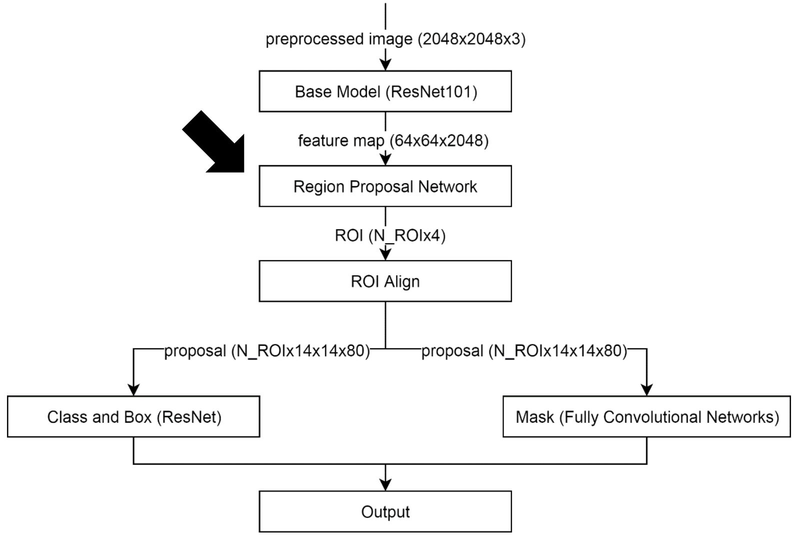
$\left[\begin{array}{c}1\times1, 64\\ 3\times3, 64\\ 1\times1, 256\end{array}\right]\times3$
$\left[\begin{array}{c}1\times1, 128\\ 3\times3, 128\\ 1\times1, 512\end{array}\right]\times4$
$\left[\begin{array}{c}1\times1, 256\\ 3\times3, 256\\ 1\times1, 1024\end{array}\right]\times23$
$\left[\begin{array}{c}1\times1, 512\\ 3\times3, 512\\ 1\times1, 2048\end{array}\right]\times3$

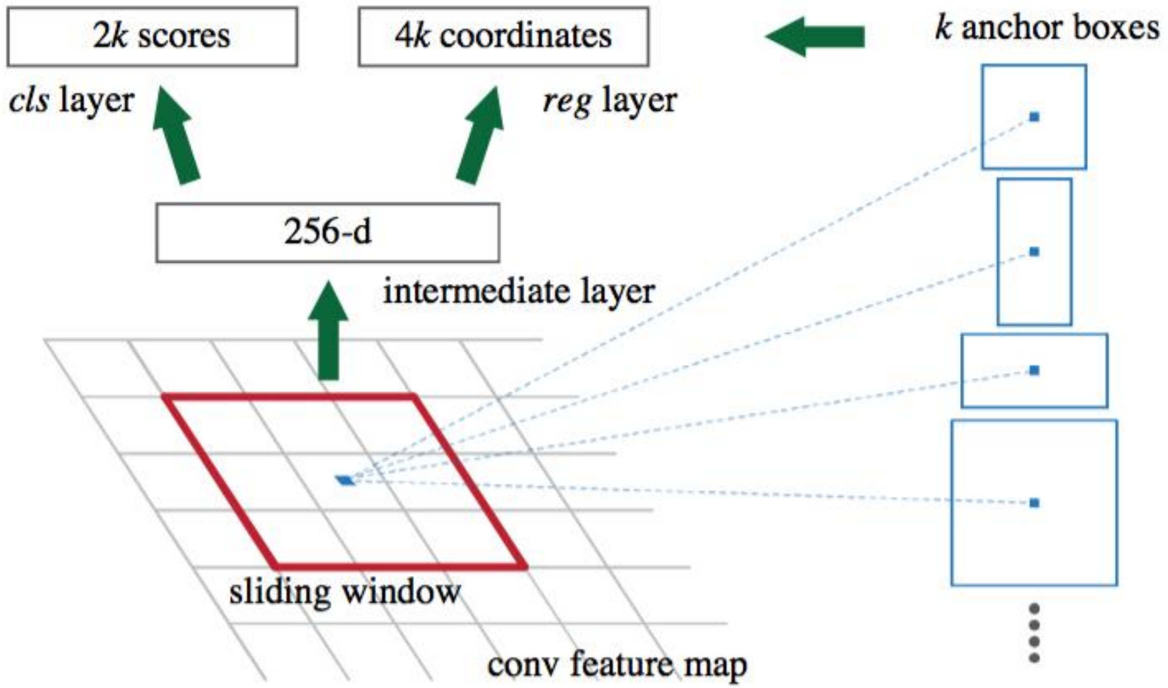
Method: Model Architecture

01 Base Model

02 Region Proposal Network

feature map → RPN →region of interest





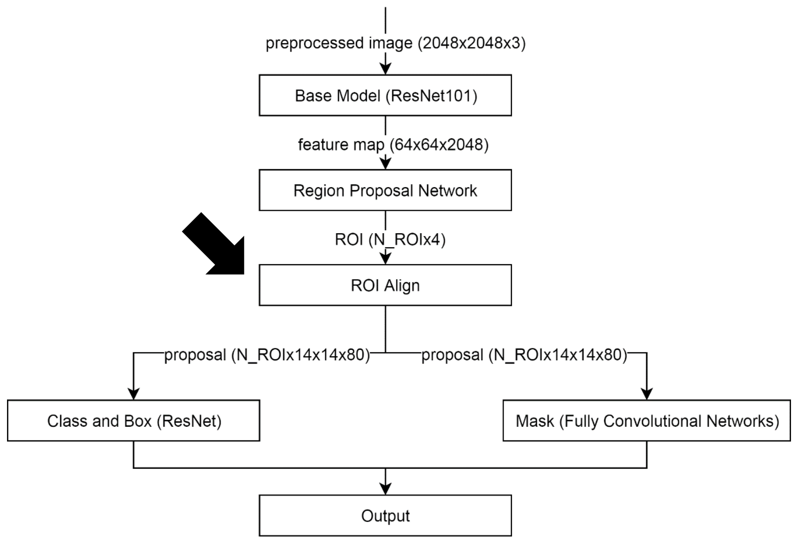
Method: Model Architecture

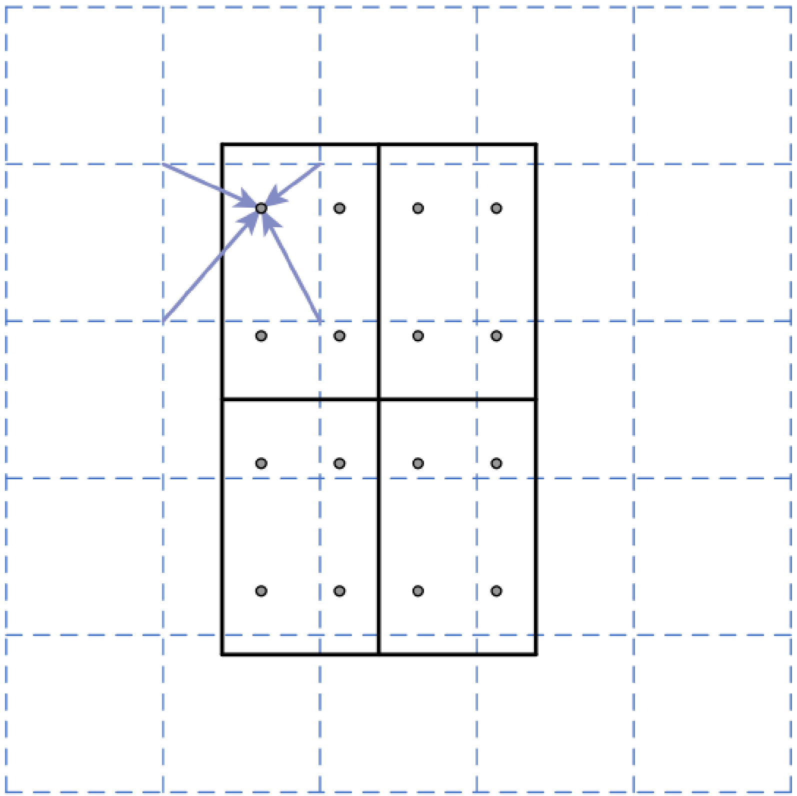
01 Base Model

02 Region Proposal Network

03 ROI Align

region of interest → ROI Align →region proposal





Method: Model Architecture

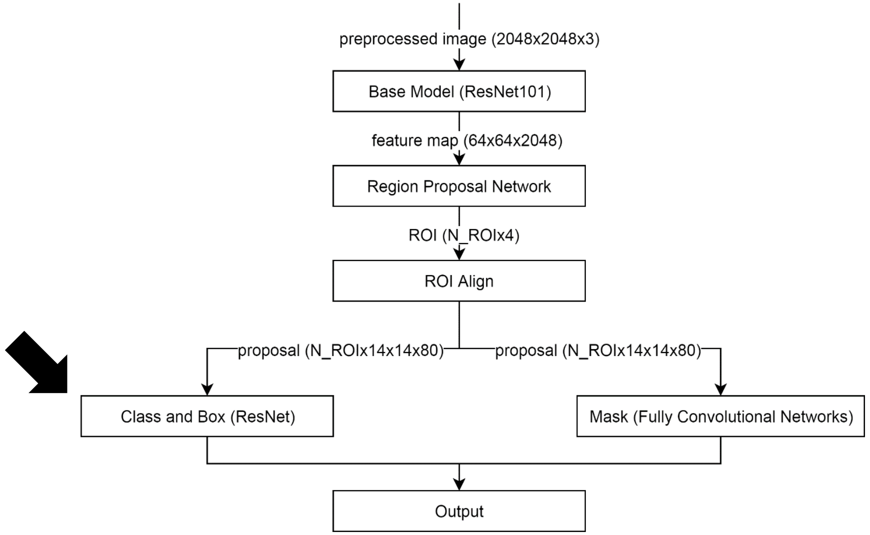
01 Base Model

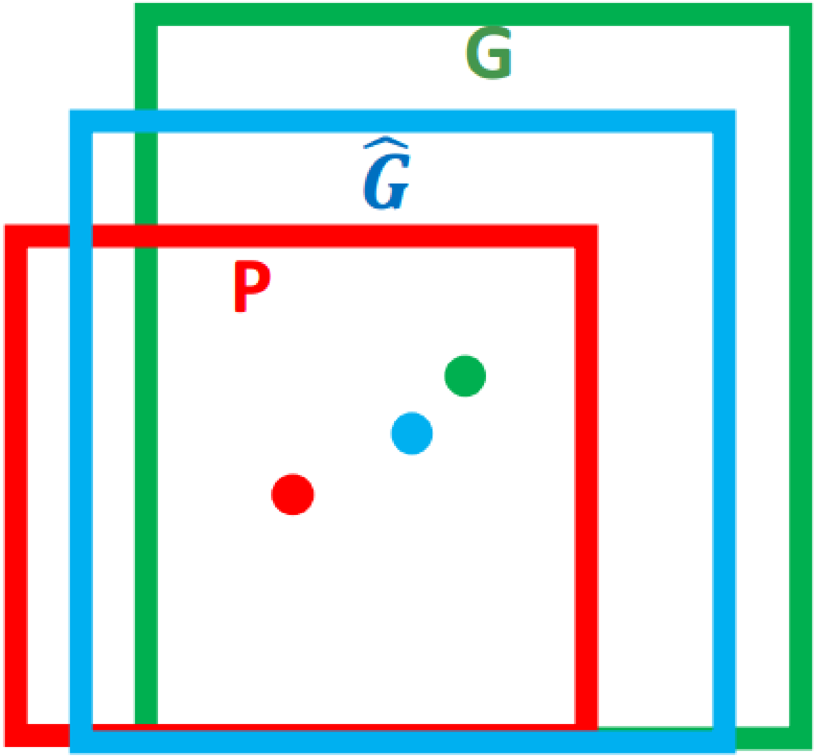
02 Region Proposal Network

03 ROI Align

04 Class and Box Generation

region proposal → ResNet→ class + box





Method: Model Architecture

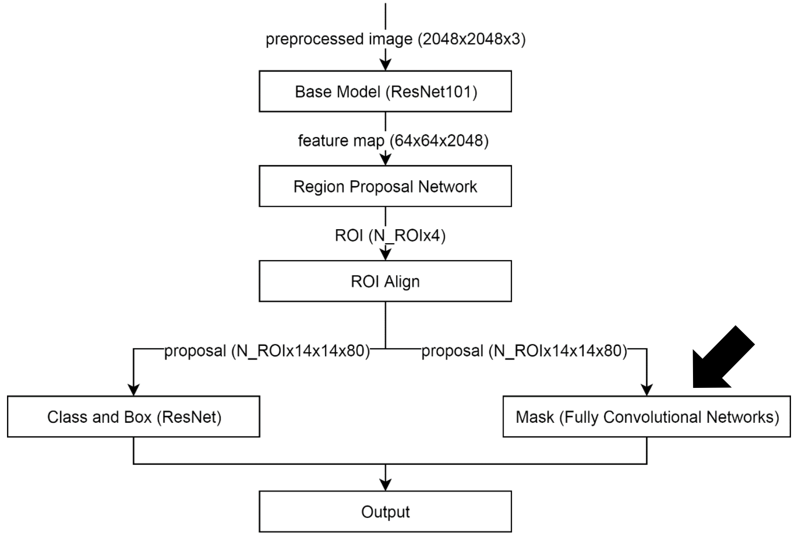
01 Base Model

02 Region Proposal Network

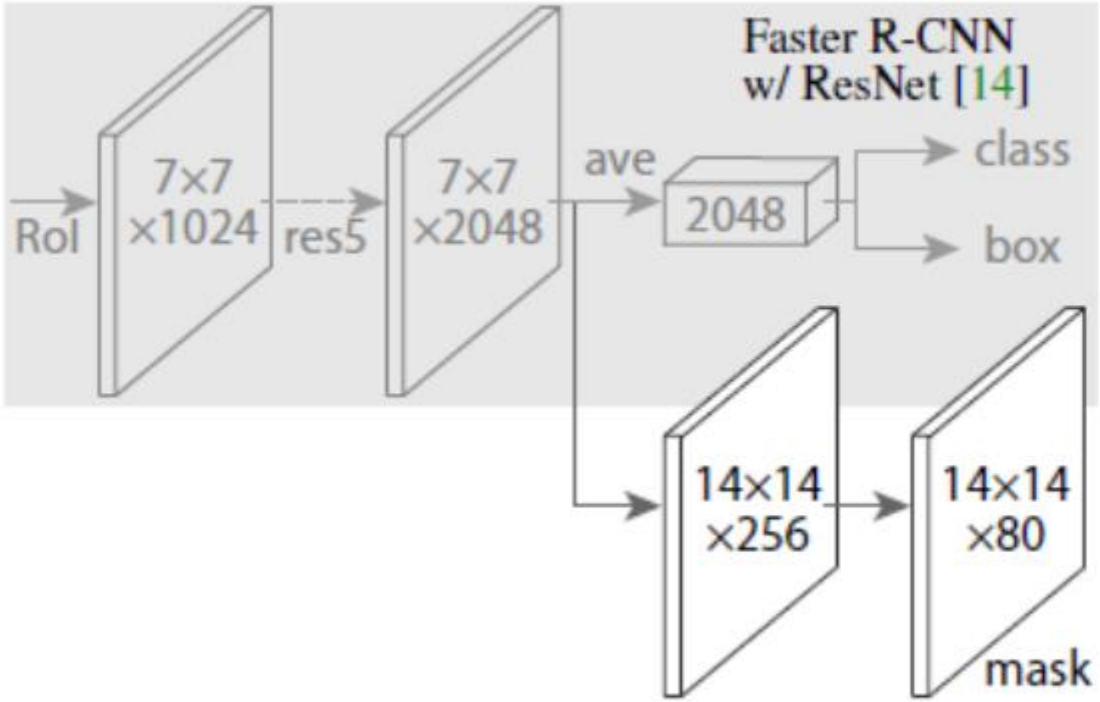
03 ROI Align

04 Class and Box Generation

05 **Mask Generation**



region proposal → mask



Method

Method: Loss Function

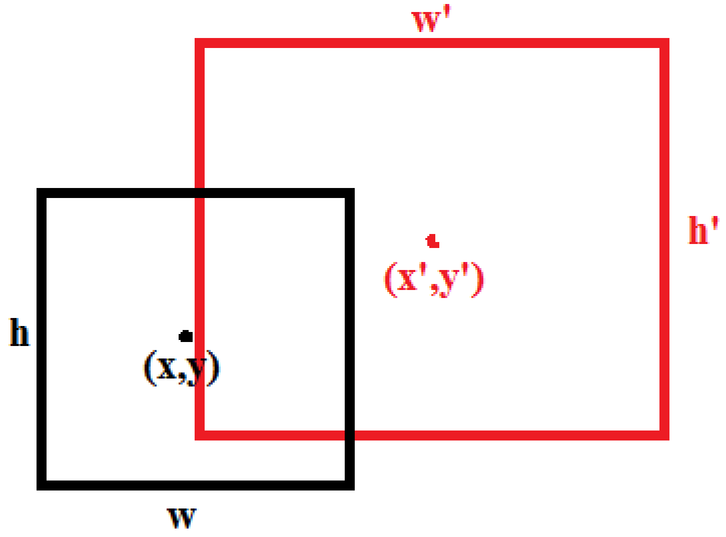
01 Original Loss Function

$$t_x = \frac{x - x'}{w'}$$

$$t_y = \frac{y - y'}{h'}$$

$$t_w = \log \frac{w}{w'}$$

$$t_h = \log \frac{h}{h'}$$



Black: Predicted box
Red: True box

Method: Loss Function

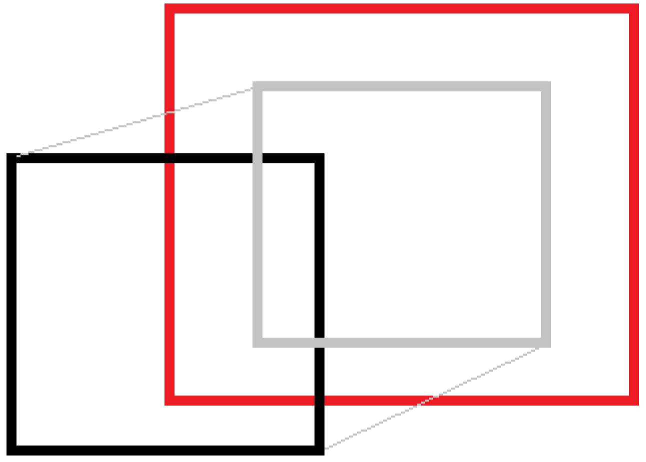
01 Original Loss Function

$$t_x = \frac{x - x'}{w'}$$

$$t_y = \frac{y - y'}{h'}$$

$$t_w = \log \frac{w}{w'}$$

$$t_h = \log \frac{h}{h'}$$

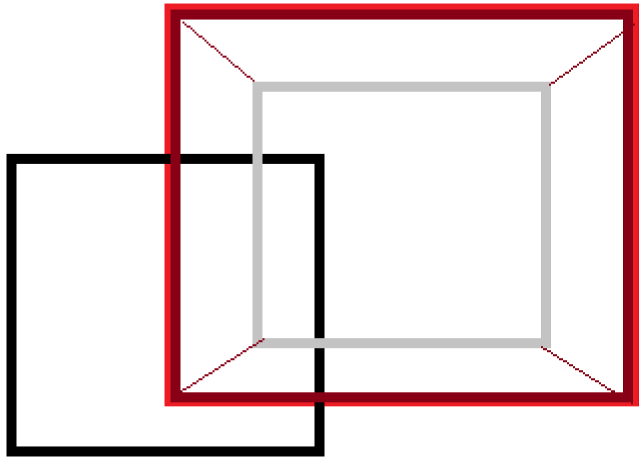


Black: Predicted box
Red: True box
Gray: Shifted box

Method: Loss Function

01 Original Loss Function

$$t_x = \frac{x - x'}{w'}$$
$$t_y = \frac{y - y'}{h'}$$
$$t_w = \log \frac{w}{w'}$$
$$t_h = \log \frac{h}{h'}$$



Black: Predicted box
Red: True box
Dark Red: Scaled box

Motivation: make IOU 100%

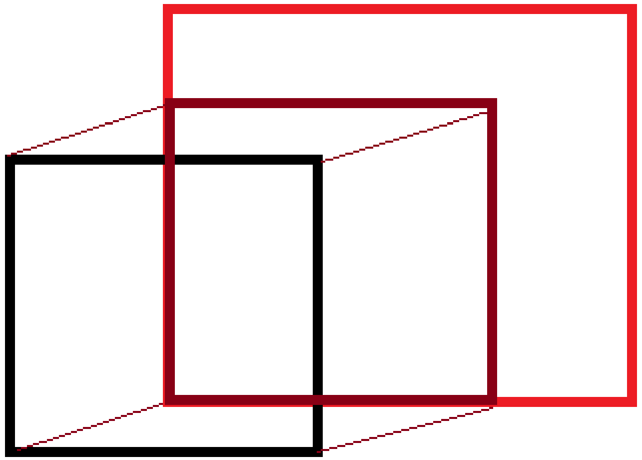
$$IoU = \frac{DetectionResult \cap GroundTruth}{DetectionResult \cup GroundTruth}$$

Method: Loss Function

01 Original Loss Function

02 Our New Loss Function

$$t_x = \begin{cases} 0, & \text{if } \operatorname{int}(x_0' > x_0) + \operatorname{int}(x_3' > x_3) = 1 \\ \dfrac{\max(x_0' - x_0, x_3' - x_3)}{w}, & \text{otherwise} \end{cases}$$



Black: Predicted box
Red: True box
Dark Red: Shifted box

Motivation: make OR 100%

$$OR = \frac{DetectionResult \cap GroundTruth}{DetectionResult}$$

Method

01 Dataset

02 Preprocess

03 Model Architecture

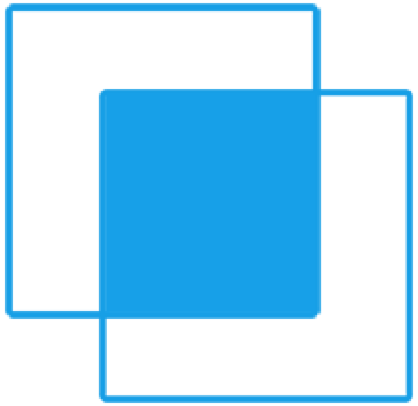
04 Loss Function

05 Evaluation

Method: Evaluation

01 Intersection over Union

$$IoU = \frac{DetectionResult \cap GroundTruth}{DetectionResult \cup GroundTruth} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$



Idea: exact match is best

Method: Evaluation

01 Intersection of Union

02 Overlapping Ratio

$$OR = \frac{DetectionResult \cap GroundTruth}{DetectionResult} = \frac{\text{Area of Overlap}}{\text{Area of Detection}}$$

Idea: one tumor cell spoils the whole sample

Method: Evaluation

01 Intersection of Union

02 Overlapping Ratio

03 Mean Average Precision

$$mAP = \frac{1}{|classese|} \sum_{e \in classes} \frac{\#TP(x)}{\#TP(c) + \#FP(c)}$$

Idea: precision of all test data. The probability of successful prediction for each predicted mask. Higher is better

Method: Evaluation

01 Intersection of Union

02 Overlapping Ratio

03 Mean Average Precision

04 False Positive Per Image

$$FPPI = \text{average number of false positive samples}$$

Idea: false positive of all test data. The number of wrong predicted masks per image. Lower is better

Method: Evaluation

01 Intersection of Union

02 Overlapping Ratio

03 Mean Average Precision

04 False Positive Per Image

05 **Mean Sensitive**

$$Sensitive = \frac{\#(\text{successfully predicted truth boxes})}{\#(\text{all truth boxes})}$$

Idea: true positive of all test data. The probability of successful prediction for each existing mammogram mass. Higher is better

05. Results

Results

01 Experiment Results

02 Analysis and Discussion

03 Limitations

Results

01 Experiment Results

02 Analysis and Discussion

03 Limitations

Experiment Results

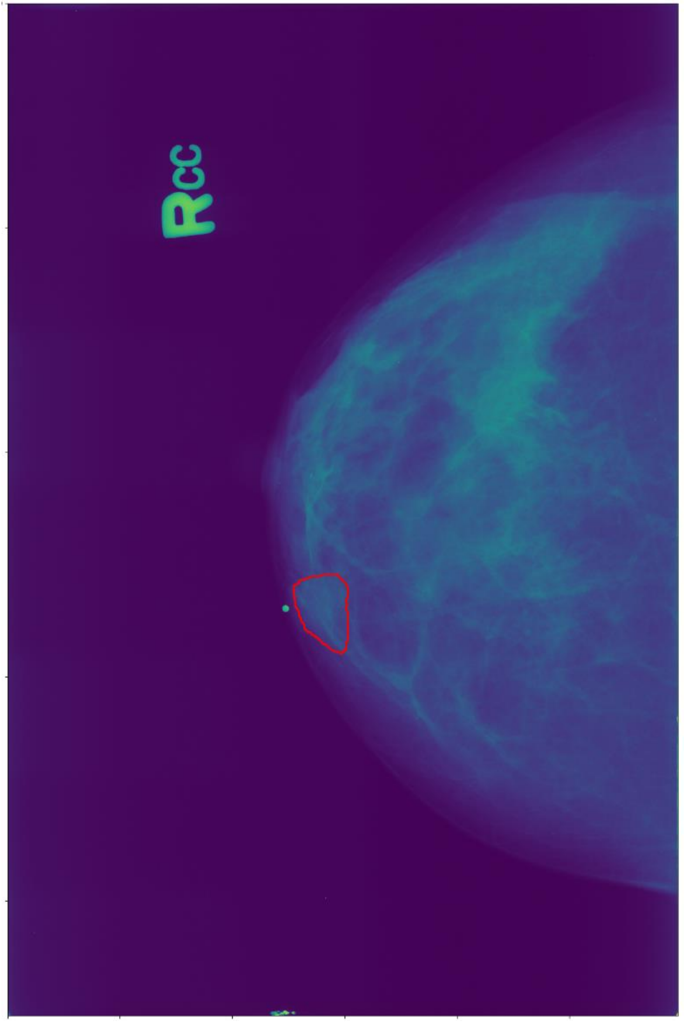
1. Preprocess
2. Positive Decision
3. Loss Function

Experiment Results

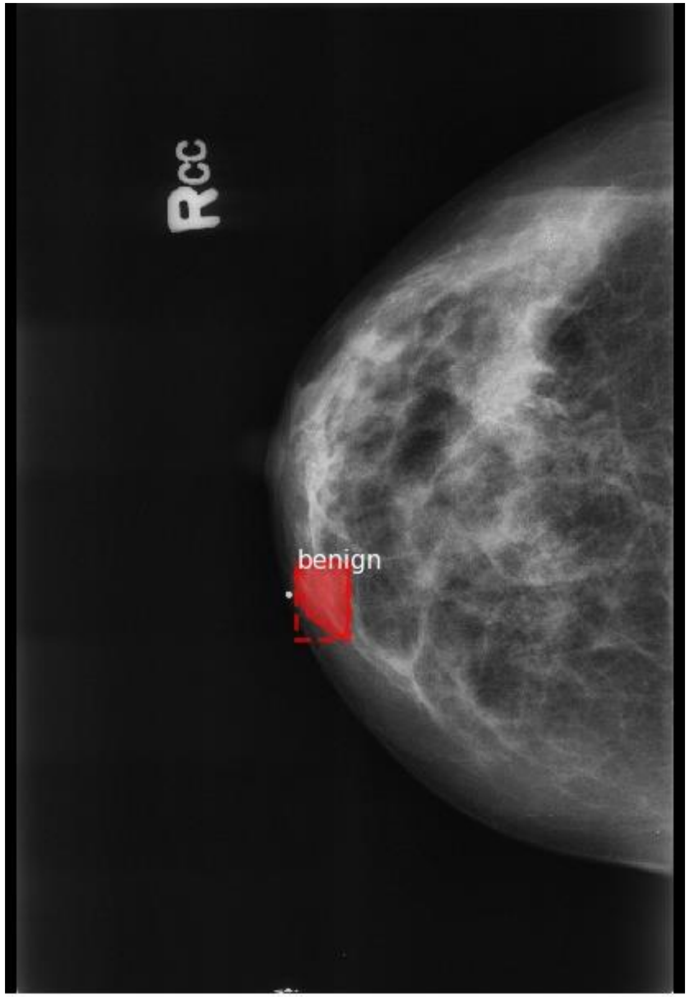
1. Preprocess
2. Positive Decision
3. Loss Function

Experiment Results: Preprocess

01 Original



02 Contrast Limited AHE



Experiment Results

1 Preprocess

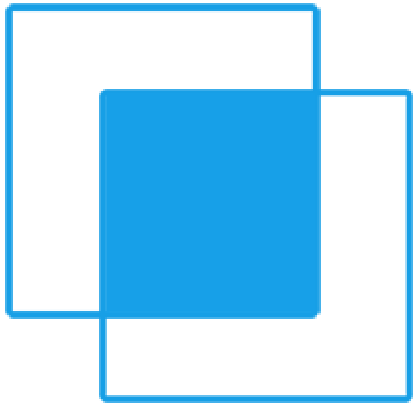
2 Positive Decision

3 Loss Function

Experiment Results: Positive Decision

01 Intersection over Union

$$IoU = \frac{DetectionResult \cap GroundTruth}{DetectionResult \cup GroundTruth} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$



Experiment Results: Positive Decision

01 Intersection of Union

02 Overlapping Ratios

$$OR = \frac{DetectionResult \cap GroundTruth}{DetectionResult} = \frac{\text{Area of Overlap}}{\text{Area of Detection}}$$

Experiment Results

1 Preprocess

2 Positive Decision

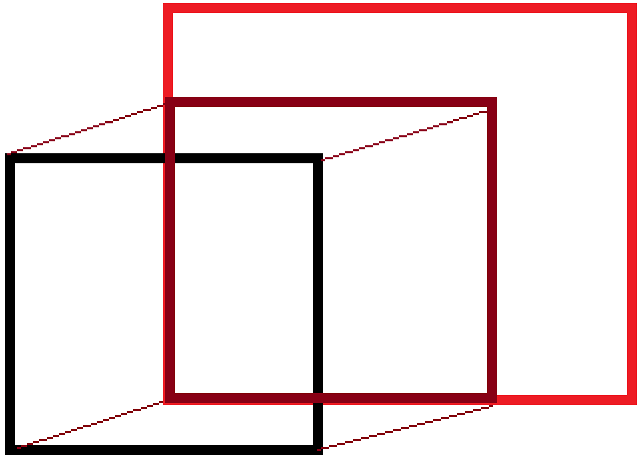
3 Loss Function

Method: Loss Function

01 Original Loss Function

02 Our New Loss Function

$$t_x = \begin{cases} 0, & \text{if } \mathrm{int}(x_0' > x_0) + \mathrm{int}(x_3' > x_3) = 1 \\ \dfrac{\max(x_0' - x_0, x_3' - x_3)}{w}, & \text{otherwise} \end{cases}$$



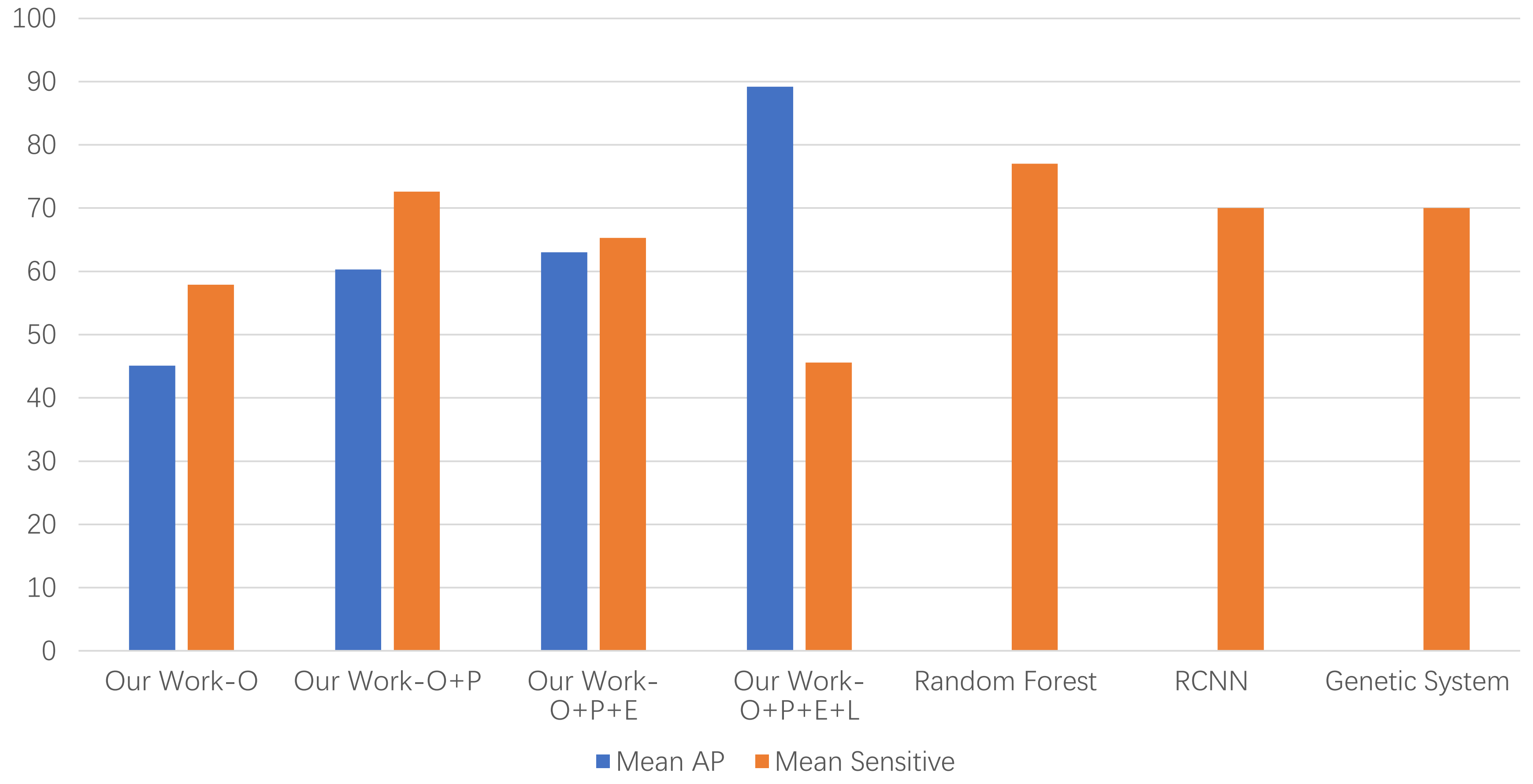
Black: Predicted box
Red: True box
Dark Red: Shifted box

Motivation: make OR 100%

Experiment Results

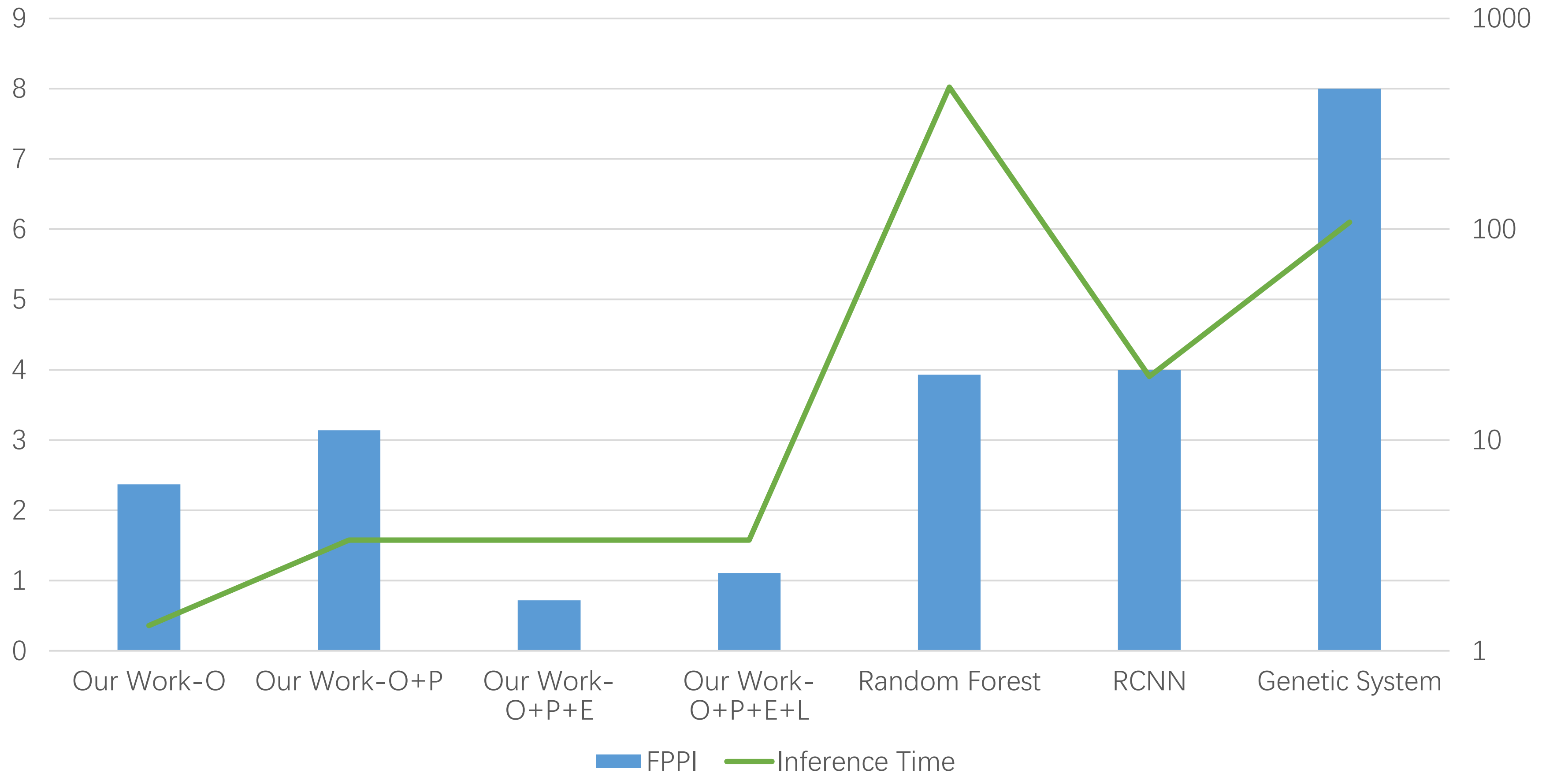
1. Preprocess
2. Positive Decision
3. Loss Function

Experiment Results



- O: original method
- P: new preprocess method
- E: new positive decision
- L: new loss function
- Our work using new preprocess method gets a comparable mean sensitive (73%) with previous work
- Our work using new loss function gets a impressive mean AP but the mean sensitive is not satisfactory.

Experiment Results



- All of our results behave much better than other works on inference time.
- Also, our work outperforms other works with the lowest FPPI

Results

01 Experiment Results

02 Analysis and Discussion

03 Limitations

Analysis and Discussion

1. Confidence Threshold
2. Training Strategy

Analysis and Discussion

1

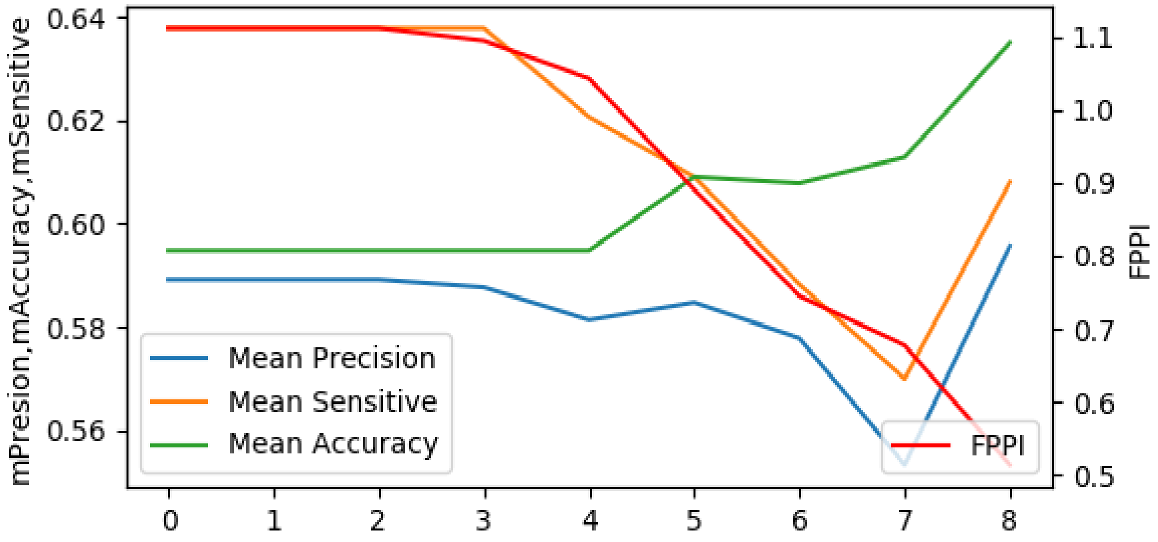
Confidence Threshold

2

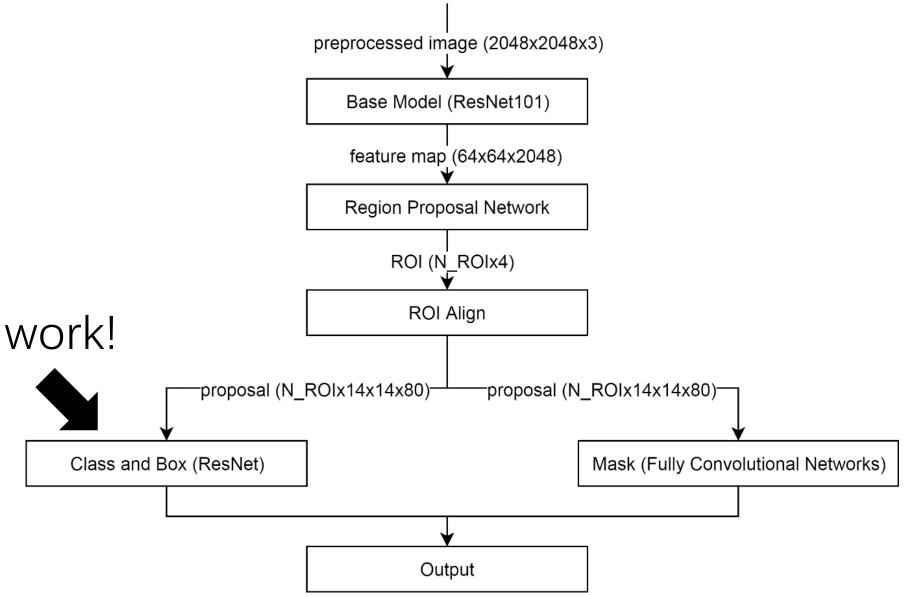
Training Strategy

Analysis and Discussion: Confidence Threshold

- X: confidence threshold * 10
- When the confidence threshold becomes larger, **the mean sensitive does not increase** as we expect!
- We conjecture that the reason is classification and bounding box regression part doesn’ t work!



Don’ t work!



Analysis and Discussion

1 Confidence Threshold

2 Training Strategy

Analysis and Discussion: Training Strategy

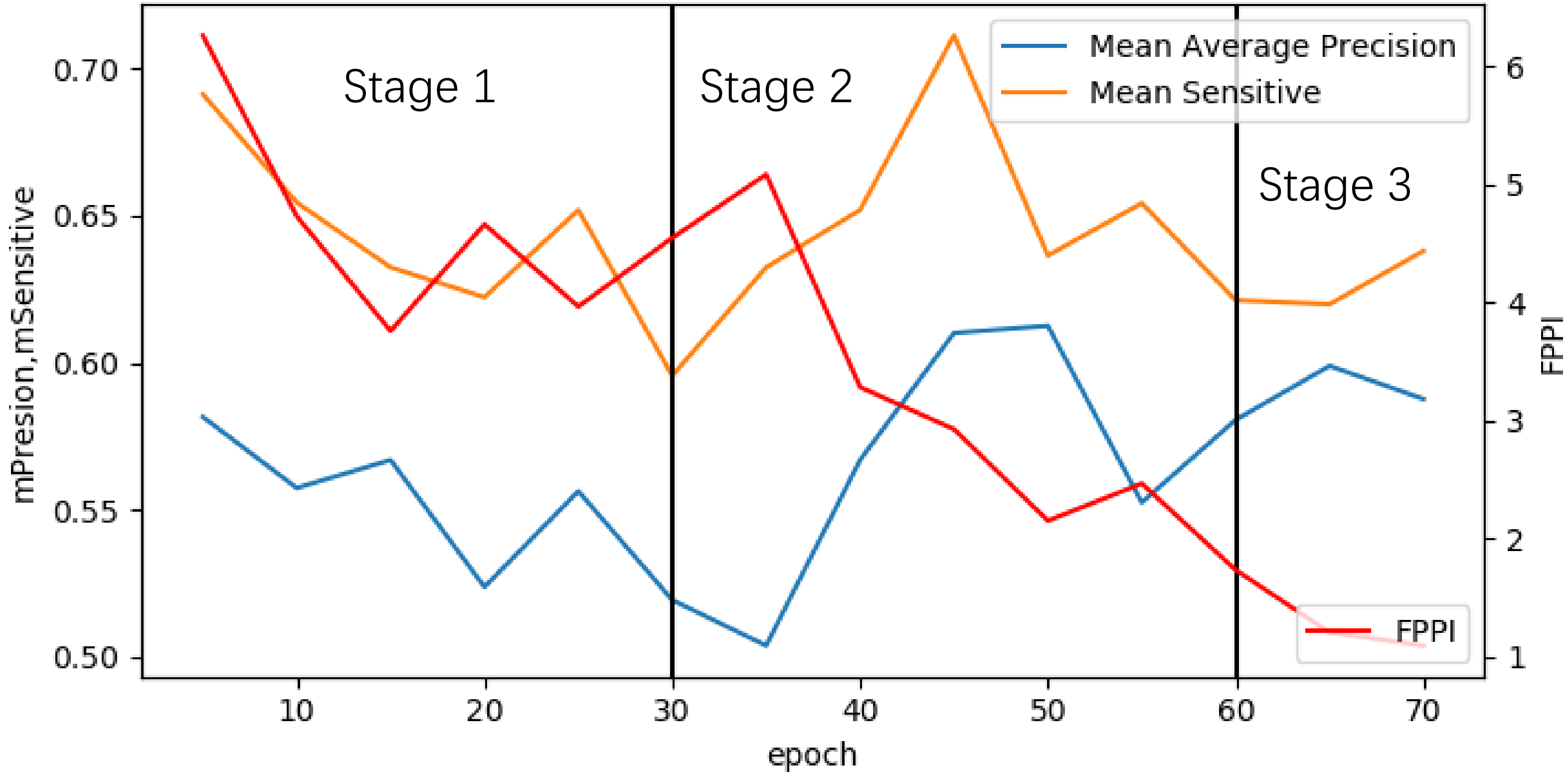
Stage 1: train different layers one by one

Stage 2: train base model

Stage 2: fine tune all layers

Mean sensitive benefits a lot from the training of base model

For each stage, 15 epochs are enough to avoid overfitting



Results

01 Experiment Results

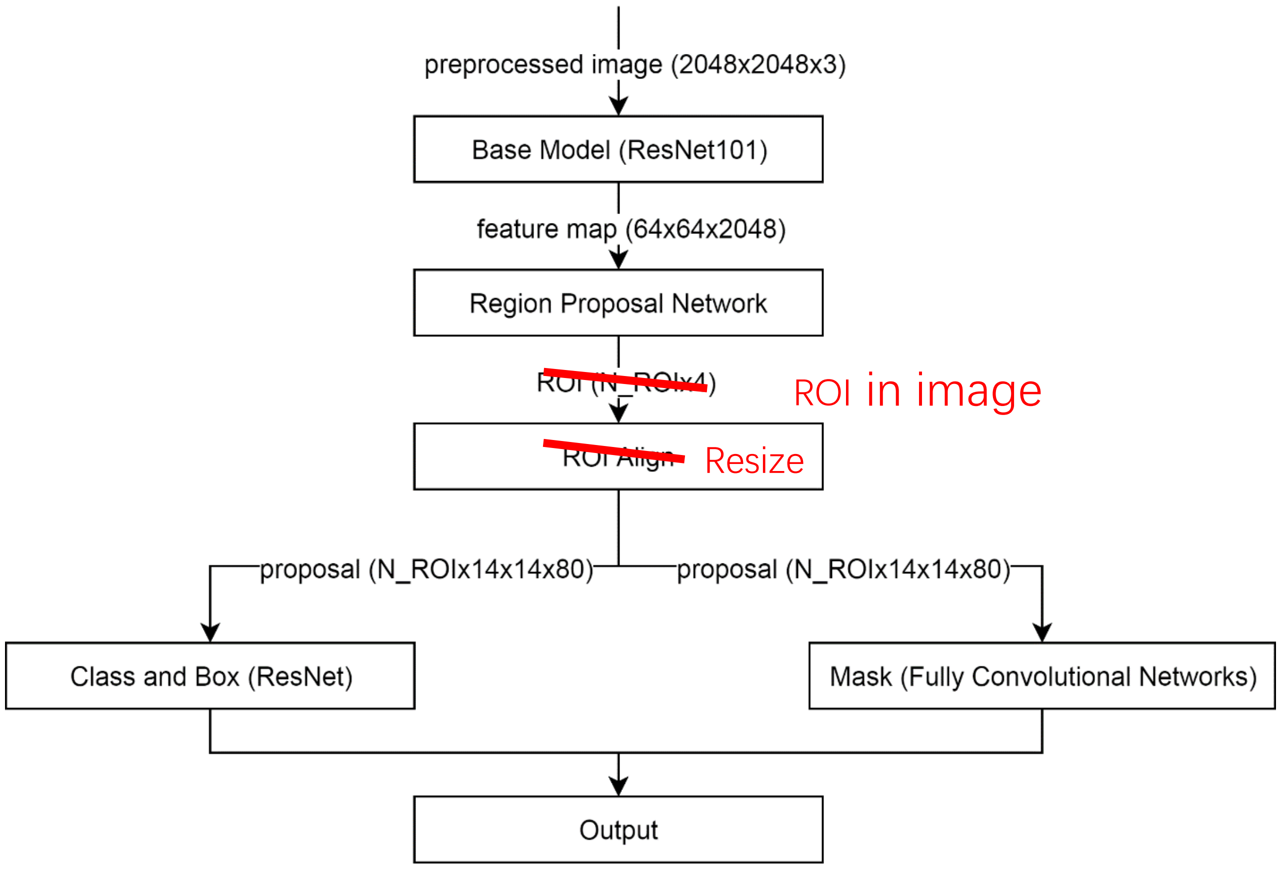
02 Analysis and Discussion

03 Limitations

Limitations: Small Feature Maps

01 Small Feature Maps

- The corresponding feature maps of mask is too small so that it has no enough representation information
- One direct idea is that using another network which adopts masks in image as input, instead of the feature maps to get the final classification and regression score.



Limitations: Too Few Training Data

01 Small Feature Maps

02 Few Training Data

- the key point of a successful is not the **power of model**, but the **power of dataset**
- Although we outperform other work using same dataset, but the results are still not impressive using private dataset stored in hospital and university.

06. User Interface

User Interface

- 01 Web Portal
- 02 Human Readable Report

User Interface

01 Web Portal

02 Human Readable Report

User Interface: Web Portal

01 Authentication

http://127.0.0.1:4899/index.html

User Interface: Web Portal

01 Authentication

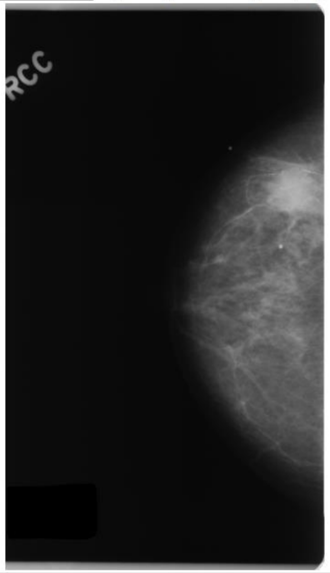
02 Submission

Breast Cancer Diagnosis

Username qli5

Password ••••••

PNG file 选择文件 C_0195_1.RIGHT_CC.png



Submit

It may take up to 120s to process an image. Please wait...

1. Upload Image
2. Process
3. Generate Report

Report:

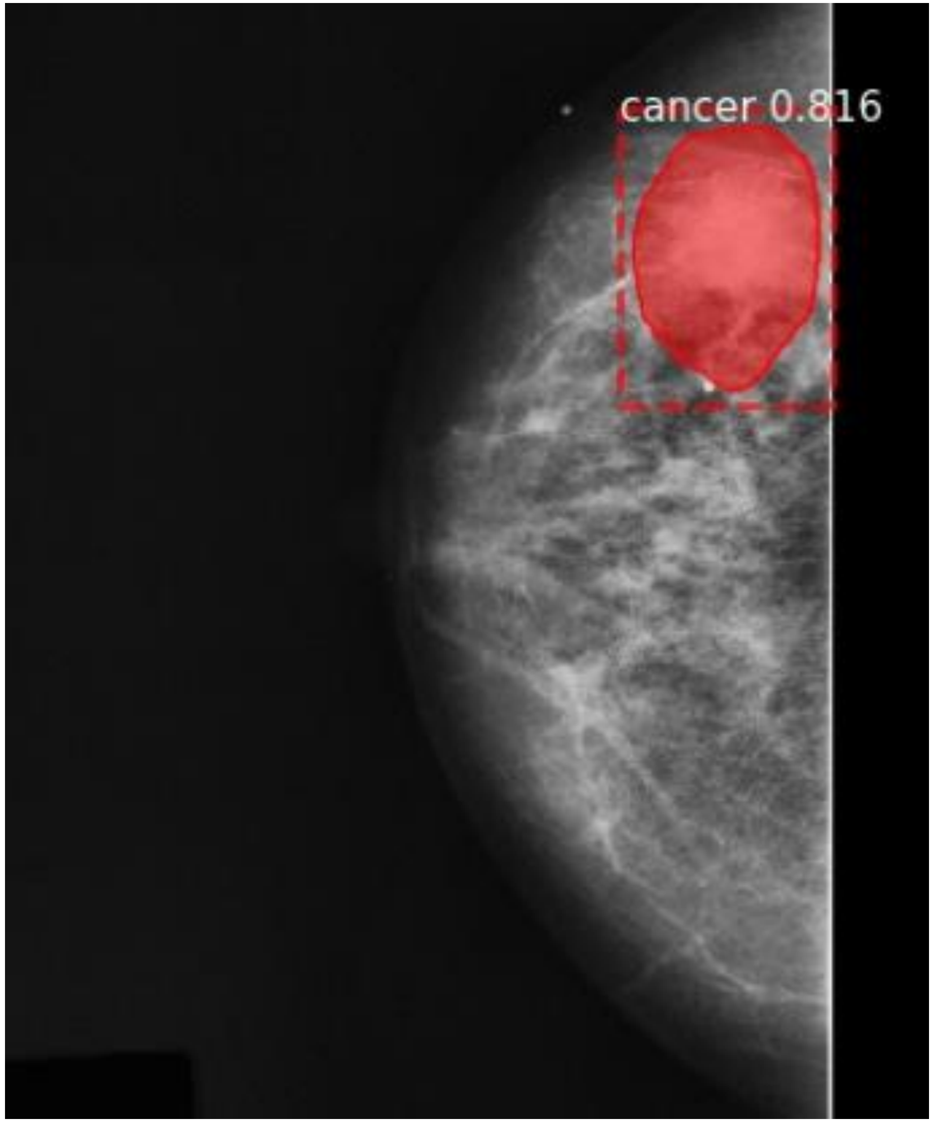
http://127.0.0.1:4899/index.html

User Interface

01 Web Portal

02 Human Readable Report

User Interface: Human Readable Report



- Bounding box
- Region mask
- Short description
- Confidence level
- Different color for different classes

07. Conclusion

Conclusion

01 Project Review

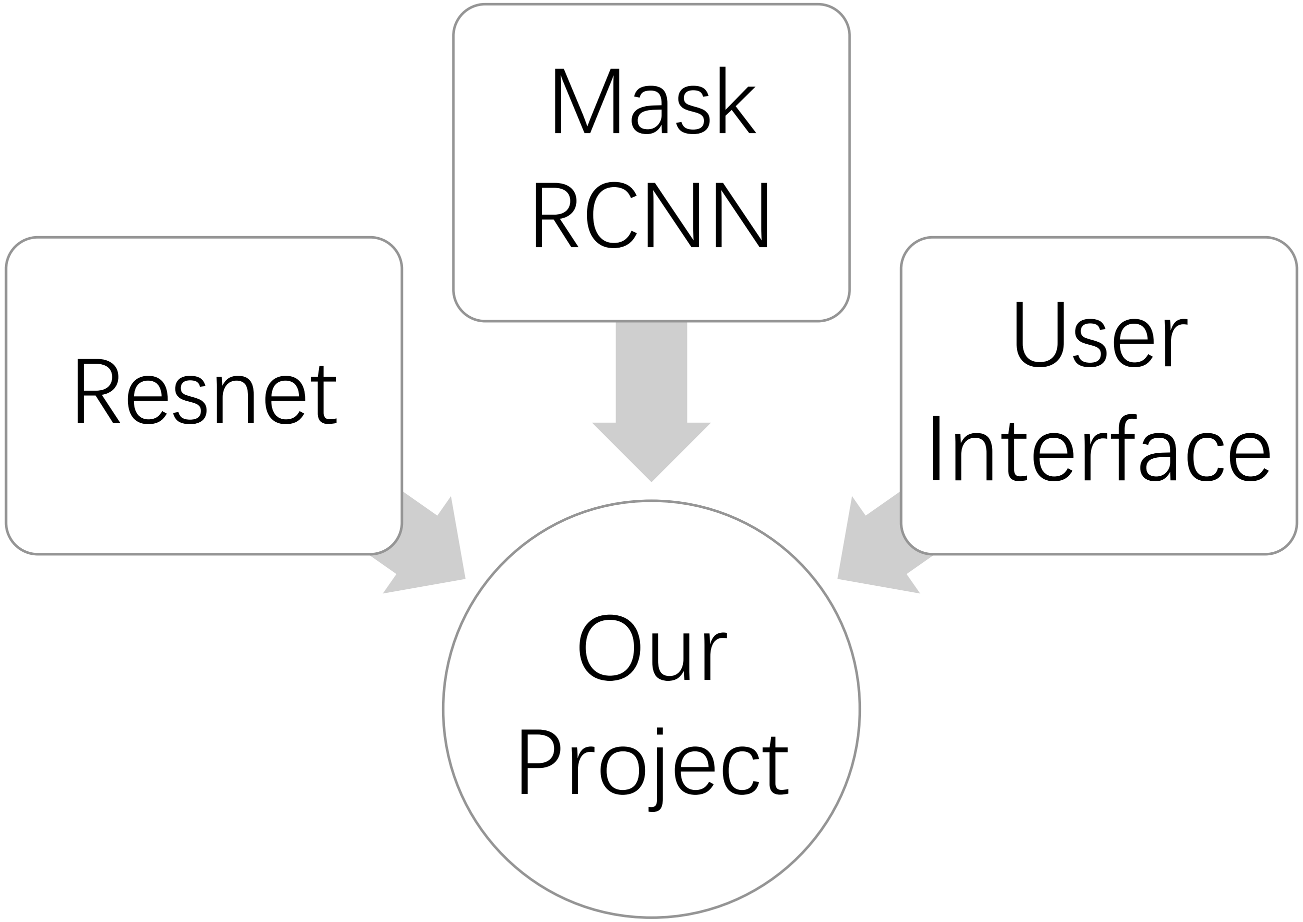
02 Future Work

Conclusion

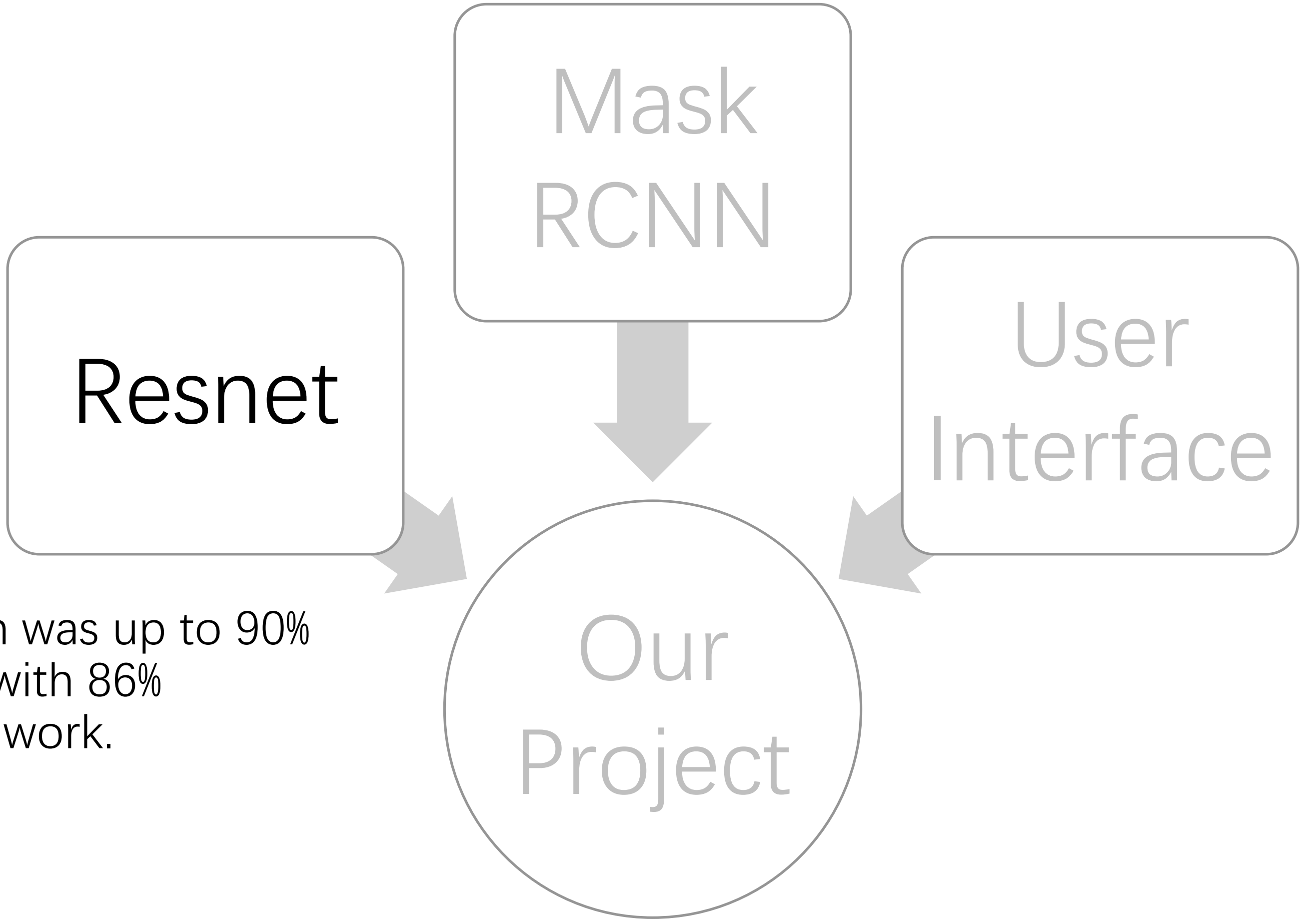
01 Project Review

02 Future Work

Conclusion: Project Review

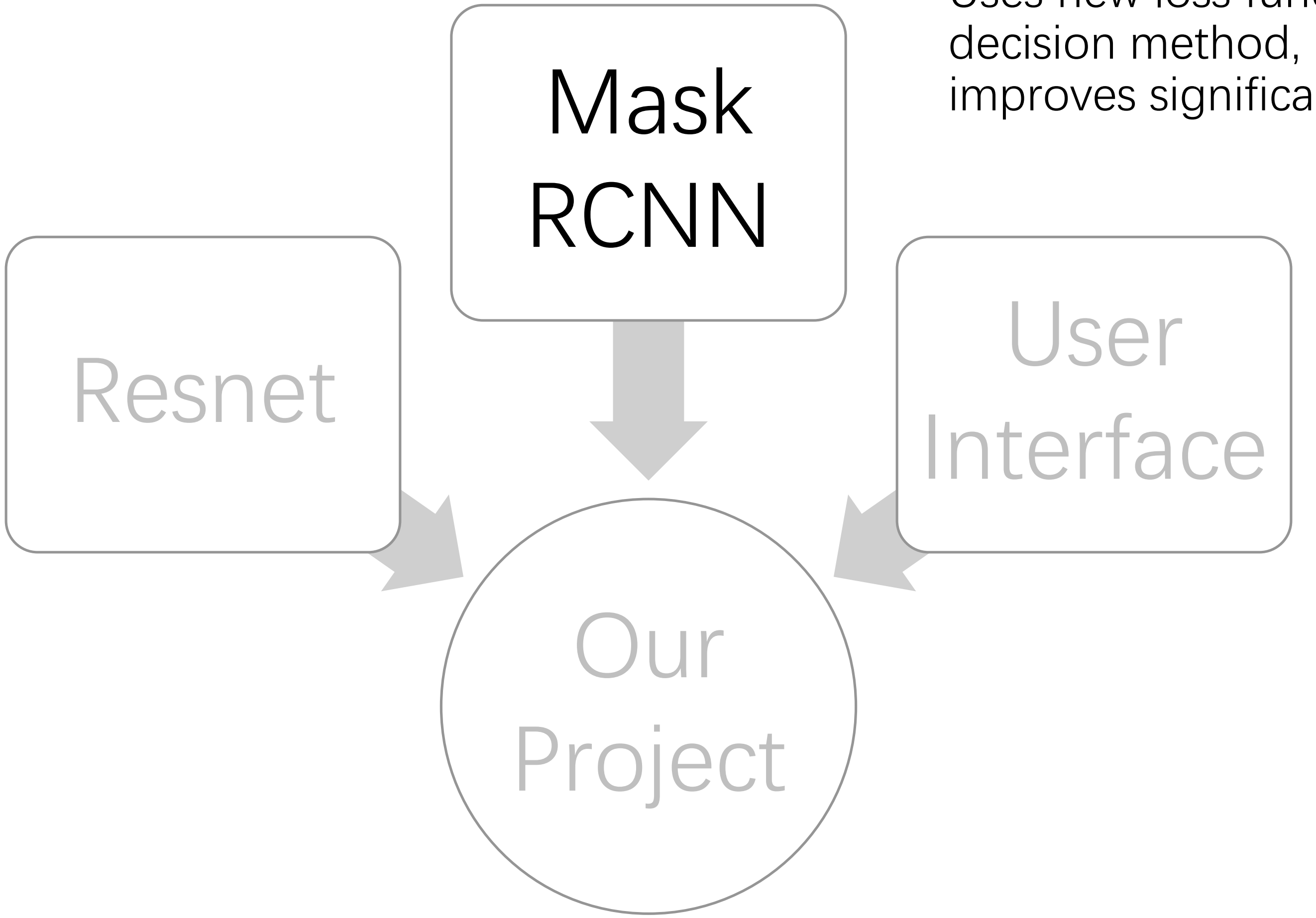


Conclusion: Project Review



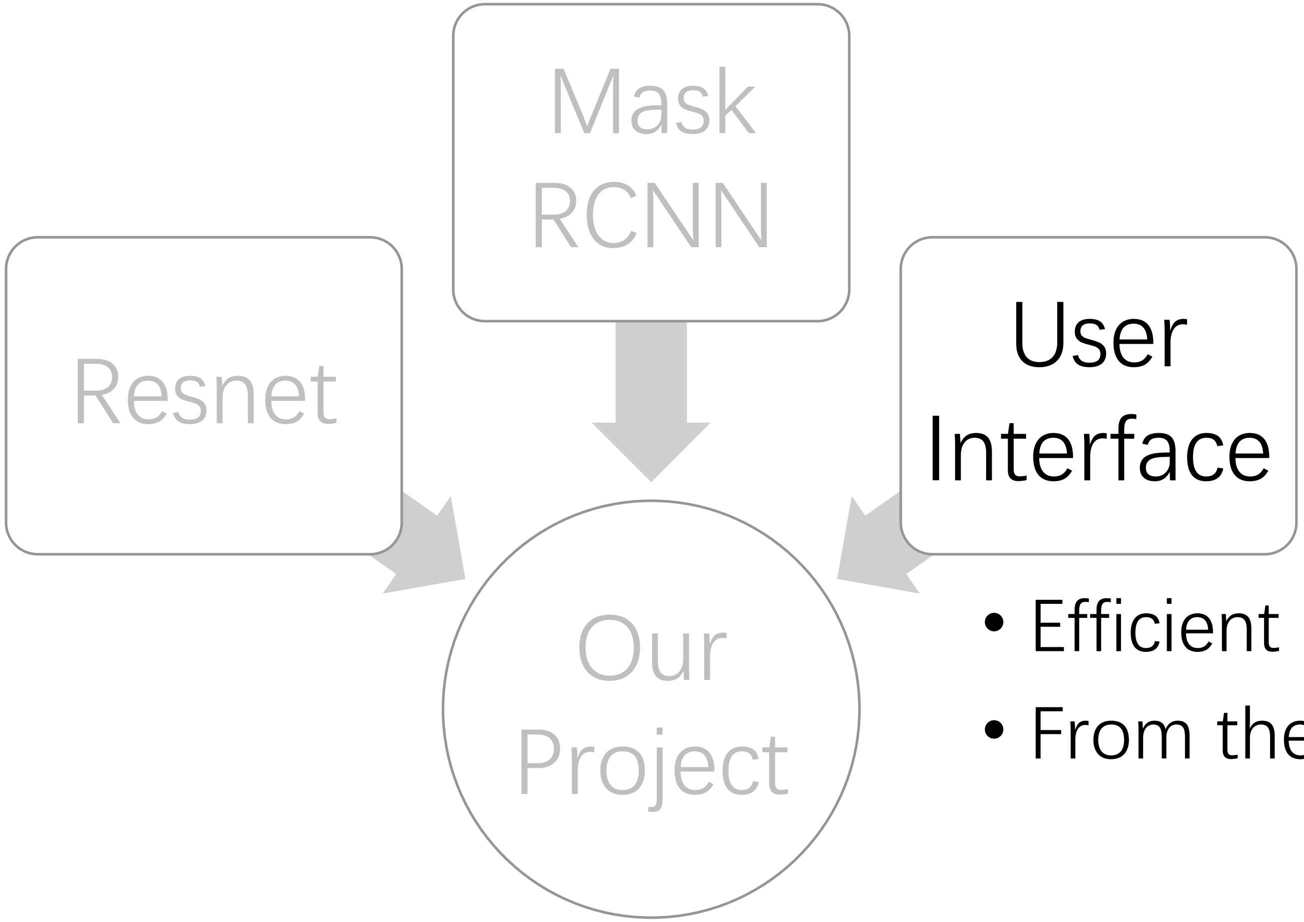
- High accuracy which was up to 90% average compared with 86% average in previous work.

Conclusion: Project Review



- Achieves a comparable result with a fewer FPPI
- Uses new loss function and positive sample decision method, making precision improves significantly from 60% level to 90%

Conclusion: Project Review



- Efficient and user-friendly
- From theory to production

Conclusion

01 Project Review

02 Future Work

Conclusion: Future Work

Metainfo

Pre-trained Model

New Network

More Data

Thank you