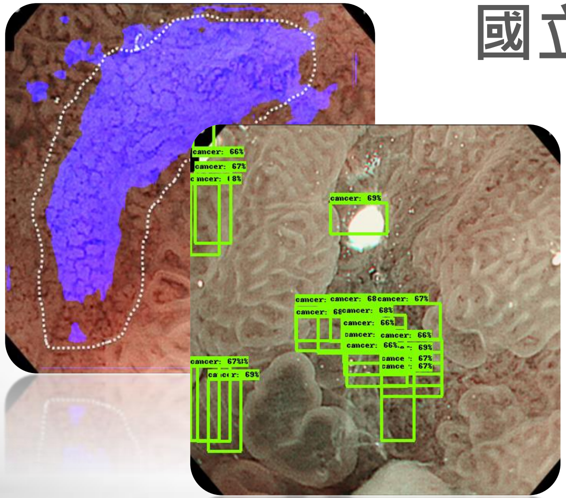
# 基於輕量型密集連接網路之擴大窄頻影像圖像胃癌區域檢測

# Gastric Cancer Region Detection for Magnified Narrow Band Imaging images Based on Lightweight Densely Connected Convolutional Networks



國立屏東科技大學 資訊管理系

報告人：簡志宇

指導教授： 許志仲 助理教授

# Outline

- Introduction
- Related Work
- Motivation
- Limitations of the Study
- Method
- Experiment Result
- Conclusion

# Introduction

- The demand for endoscopy of gastric cancer is increasing, but not all doctors can determine the condition of a patient's symptoms immediately.
- Computer-aided detection (CADe) can effectively help doctors in areas that require special attention during the detection

  ※ *Reducing errors caused by fatigue caused by repeated detections by doctors.*

# Introduction

- High-performance detection network, conditions must include:
  - <u>The Sensitivity of the lesion area must be high</u>
  - <u>The False-negative must be low</u>
  - <u>The detection time must be short</u>
- It is difficult to achieve these performances only by using the published network for deep learning training.

# Introduction

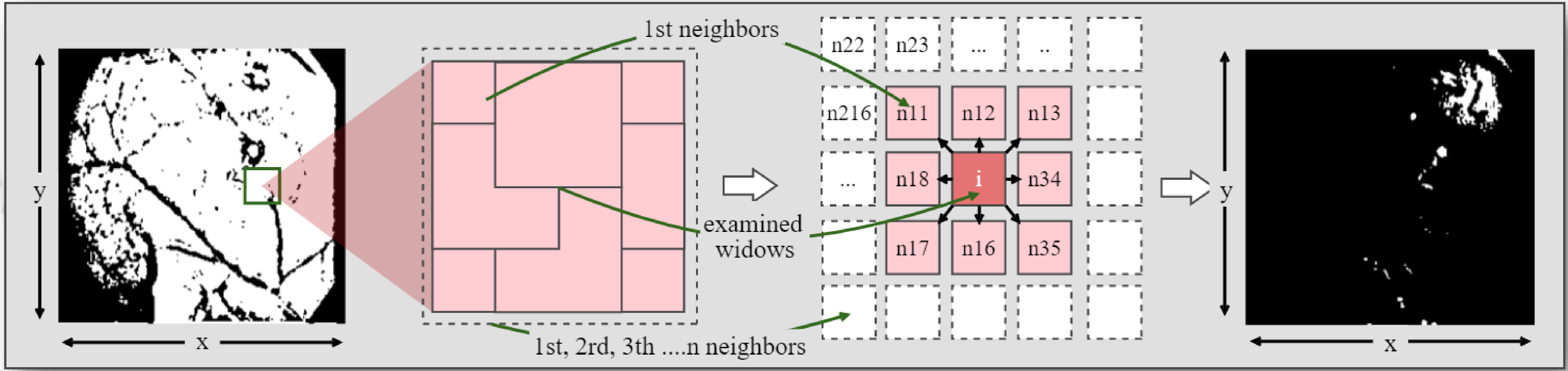
Semantic Segmentation network

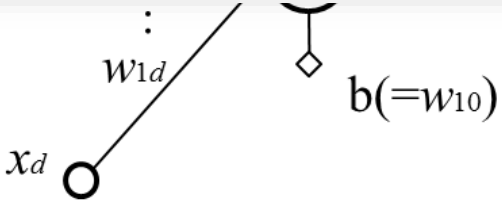
※not fast enough !

Object Detection network ＋ Semantic Segmentation network ＝ Combined Two Network

※ more faster?
※ same performance or inefficient?

# Related Work – Related technologies I

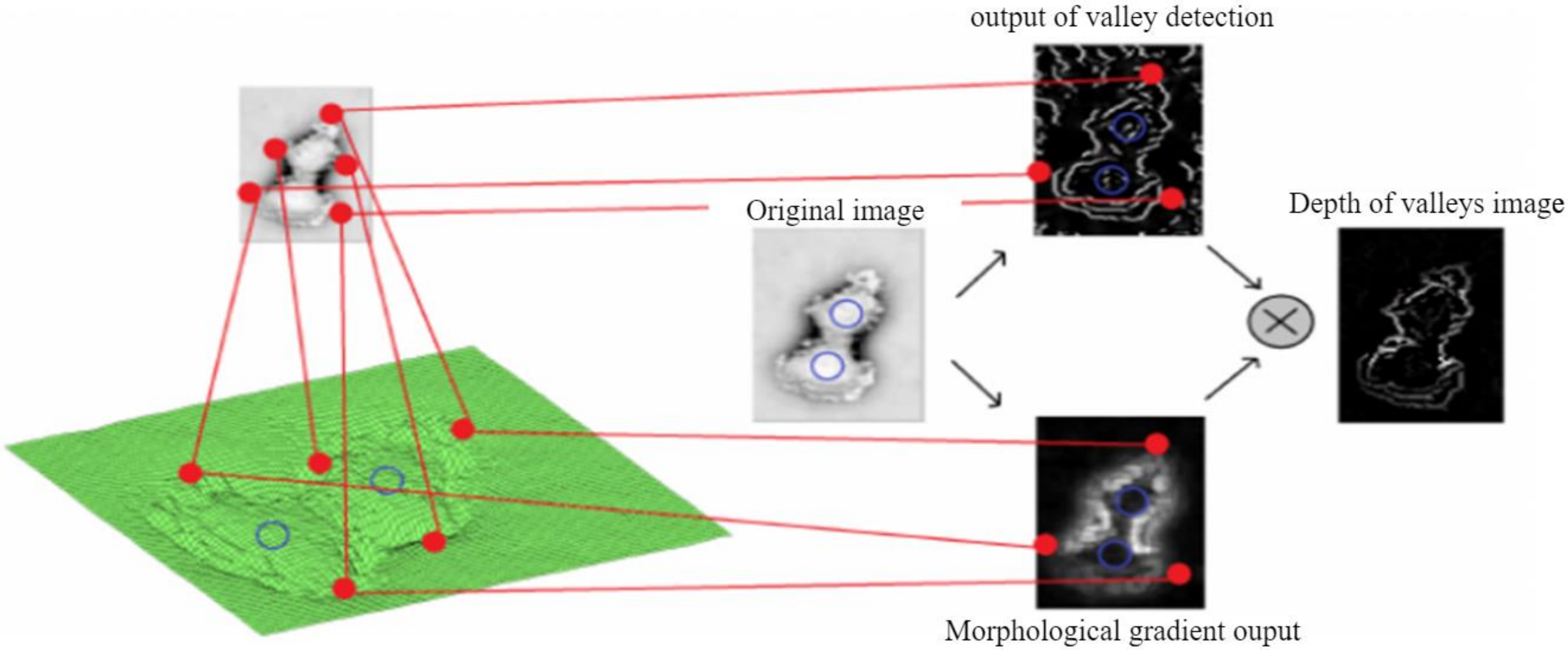




Maroulis, Dimitrios E., et al., "CoLD: a versatile detection system for colorectal lesions in endoscopy video-frames," Computer Methods and Programs in Biomedicine , Vol. 70, No.2, pp.151-166, 2003.

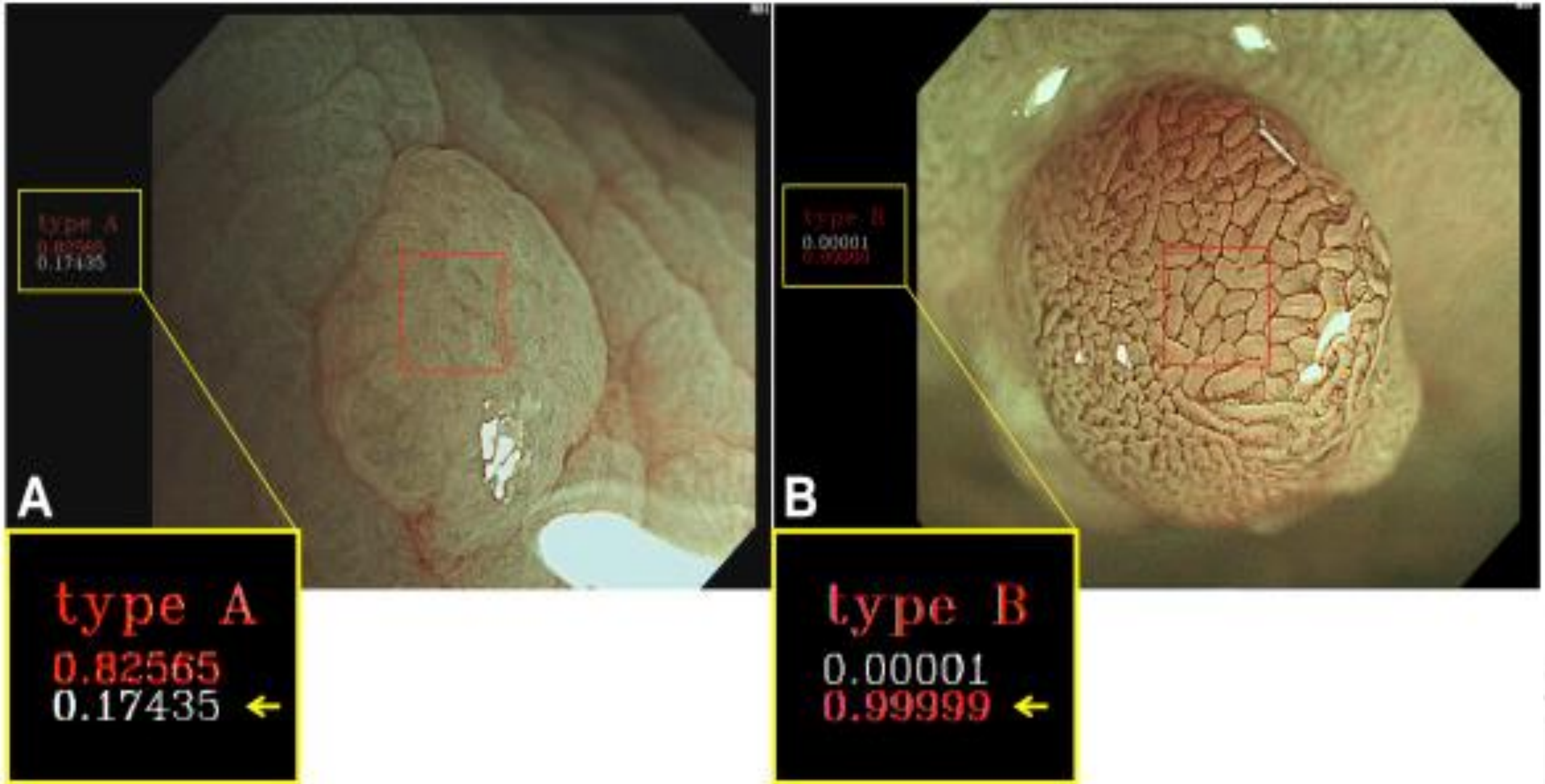
# Related Work – Related technologies II





Bernal, Jorge, et al., "WM-DOVA maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians," Computerized Medical Imaging and Graphics , Vol. 43, pp. 99-111, 2015

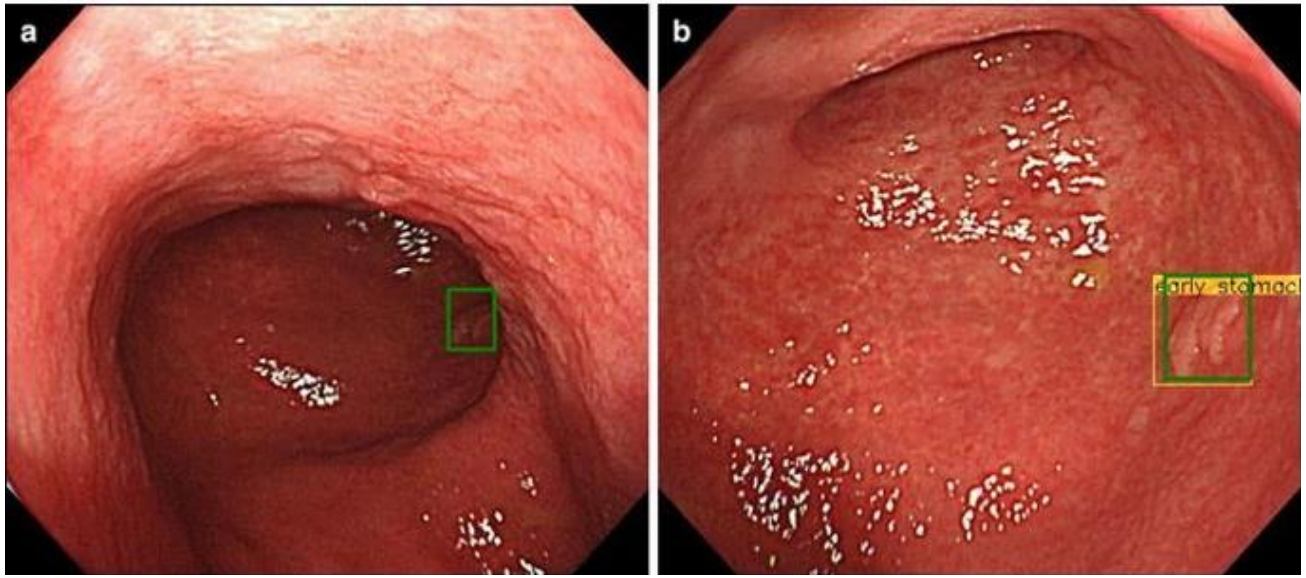
# Related Work – Related technologies III




imaging magnifying colonoscopy," Gastrointestinal endoscopy , Vol. 83, No.3, pp. 643-649,2016.

# Related Work – Related technologies IV

- Using Deep Learning for detecting gastric cancer
  - Using SSD (Single Shot Multibox Detector) net build detect network




Green box = ground truth

Yellow box = predicted result

Hirasawa, Toshiaki, et al., "Application of artificial intelligence using a convolutional neural network for detecting gastric cancer in endoscopic images." Gastric Cancer , Vol. 21, No. 4 ,pp. 653-660, 2018.

# Related Work – Related technologies IV

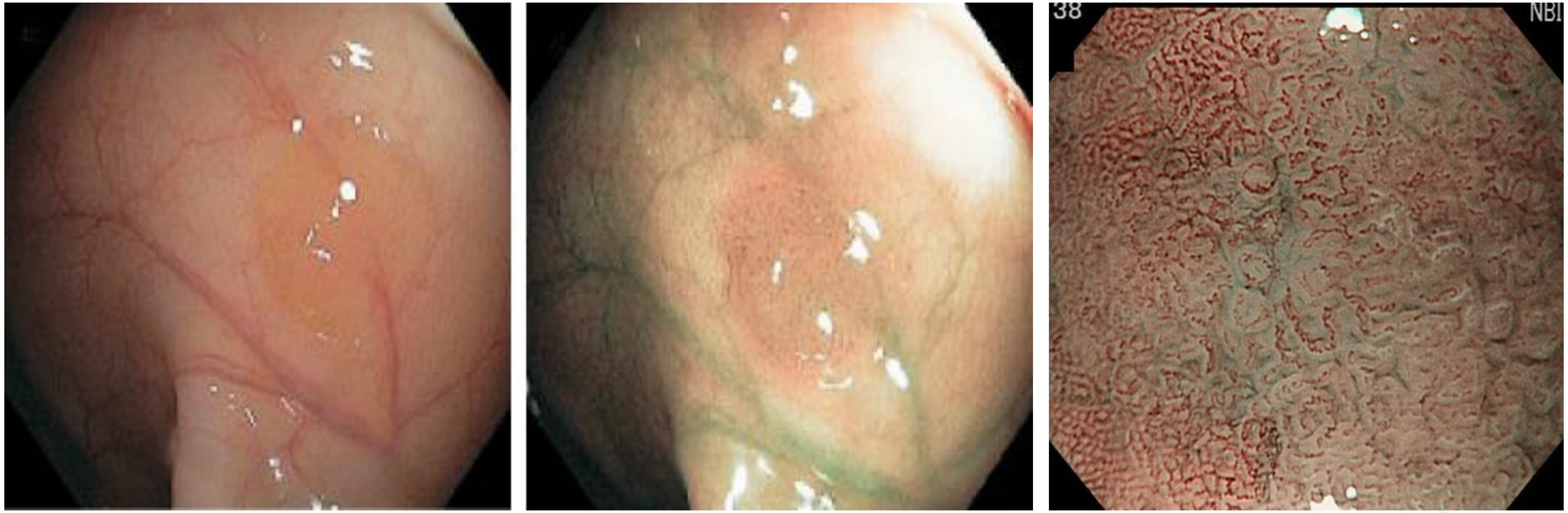
- Current CAD testing equipment can help doctors quickly find diseased areas.
- Most device are very expensive.
- Difficult to detection early diseased cells.

---

- It must needs a lot of data to train a good deep learning network.
- Need a doctor to help make a good dataset.

# Related Work – M-NBI

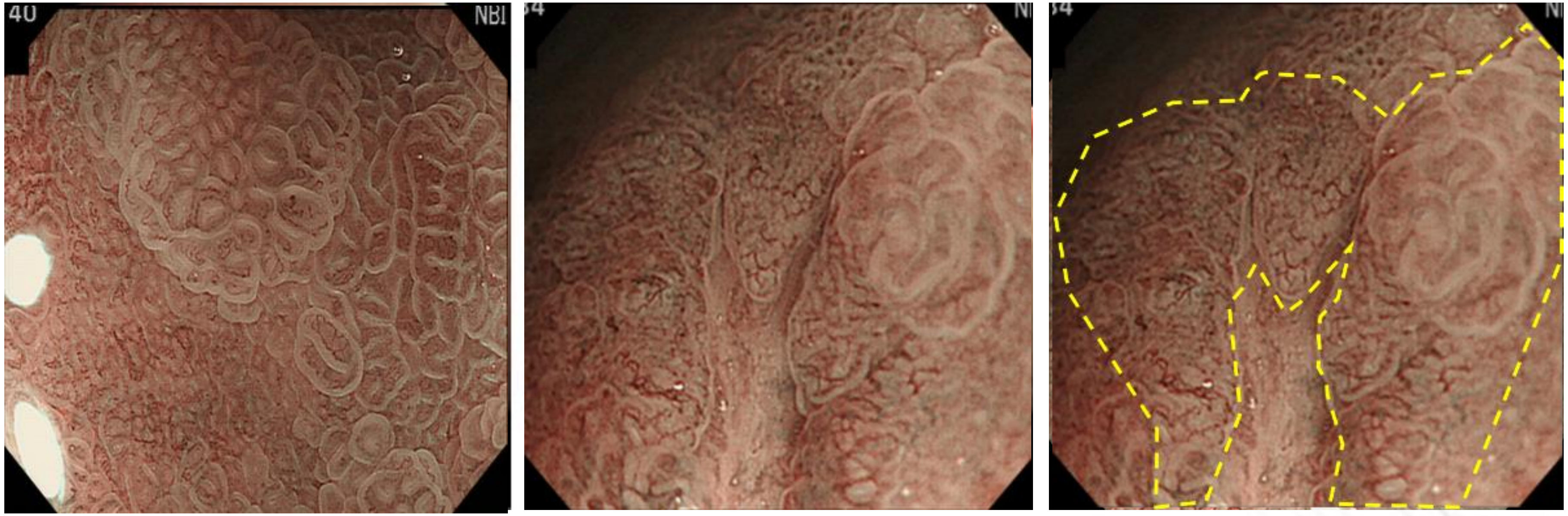
- M-NBI




| Endoscopy | NBI | M-NBI |
| --- | --- | --- |

# Related Work – M-NBI

- M-NBI




| | | |
|---|---|---|
| No-cancer | Cancer | Segmentation Map (Ground Truth) |

# Related Work – Data Augmentation

- Data Augmentation


=4 + 1 =5

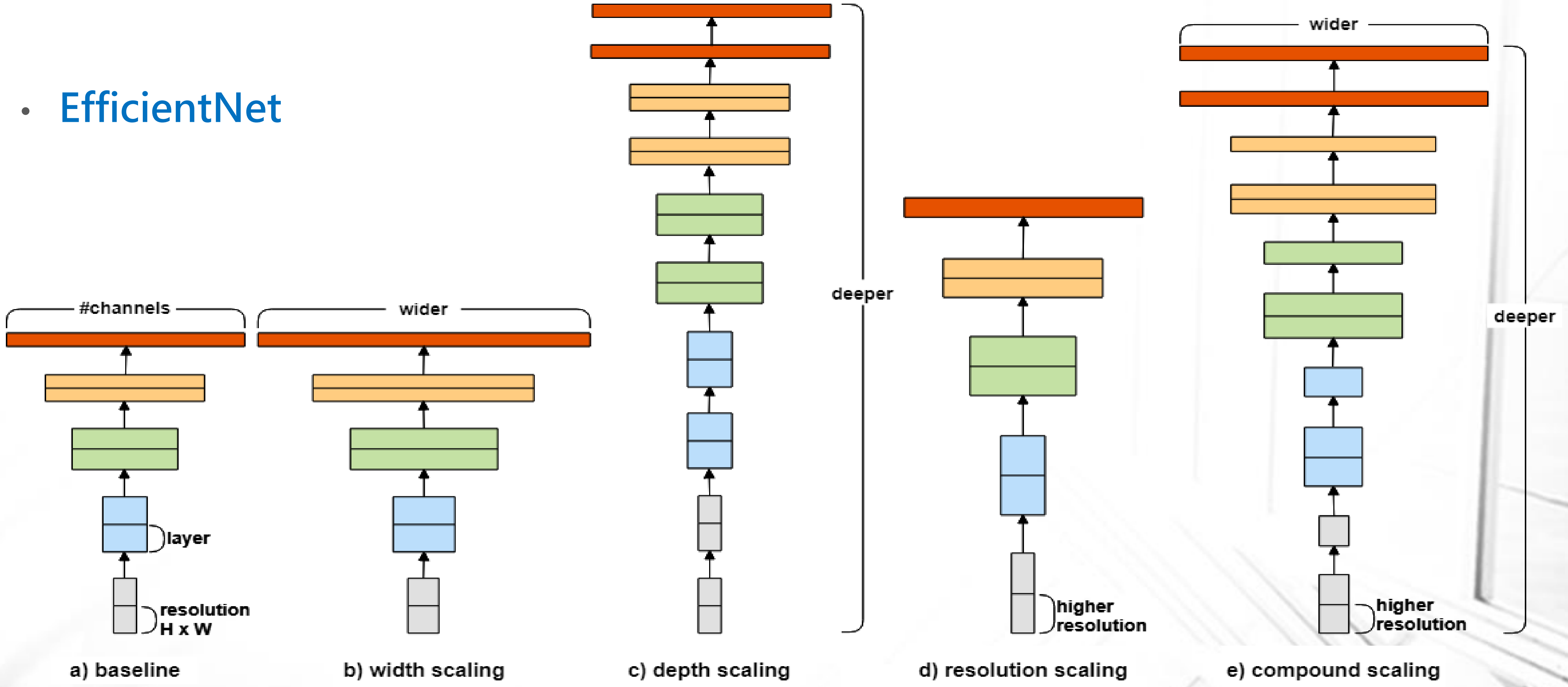
=4 + 1 =5

=4 + 4 + 1 =9


Zhong, Zhun, et al., "Random erasing data augmentation," arXiv preprint arXiv:1708.04896 , 2017.

# Related Work – Backbone Network

- EfficientNet

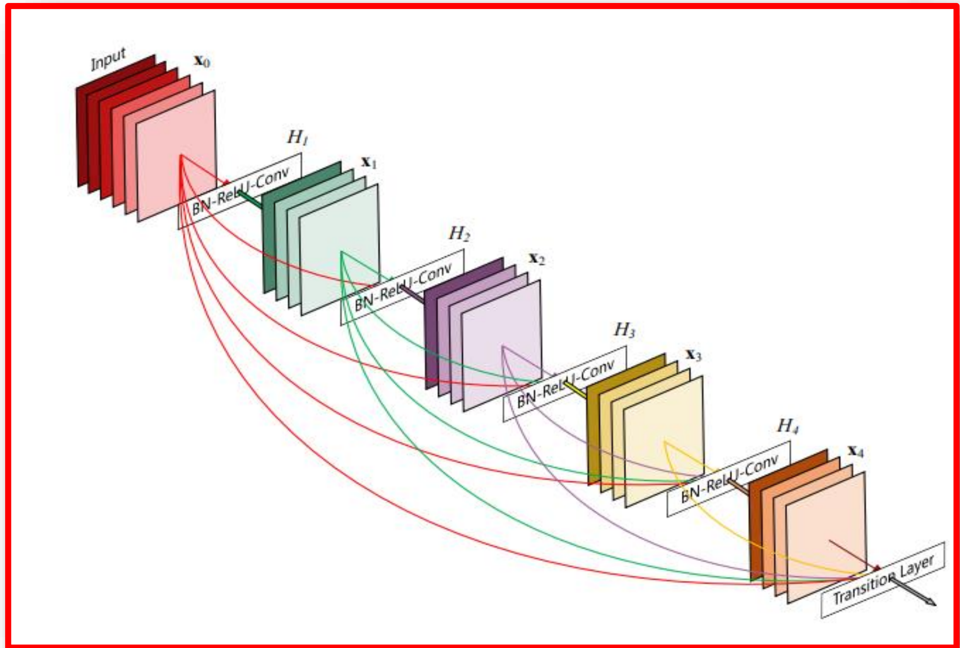




Tan, Mingxing, and Quoc V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," arXiv preprint arXiv:1905.11946, 2019.

# Related Work – Backbone Network

- DenseNet
  - Solve the overfitting problem
  - Greatly reduce the number of parameters
  - The features of each layer can be reused





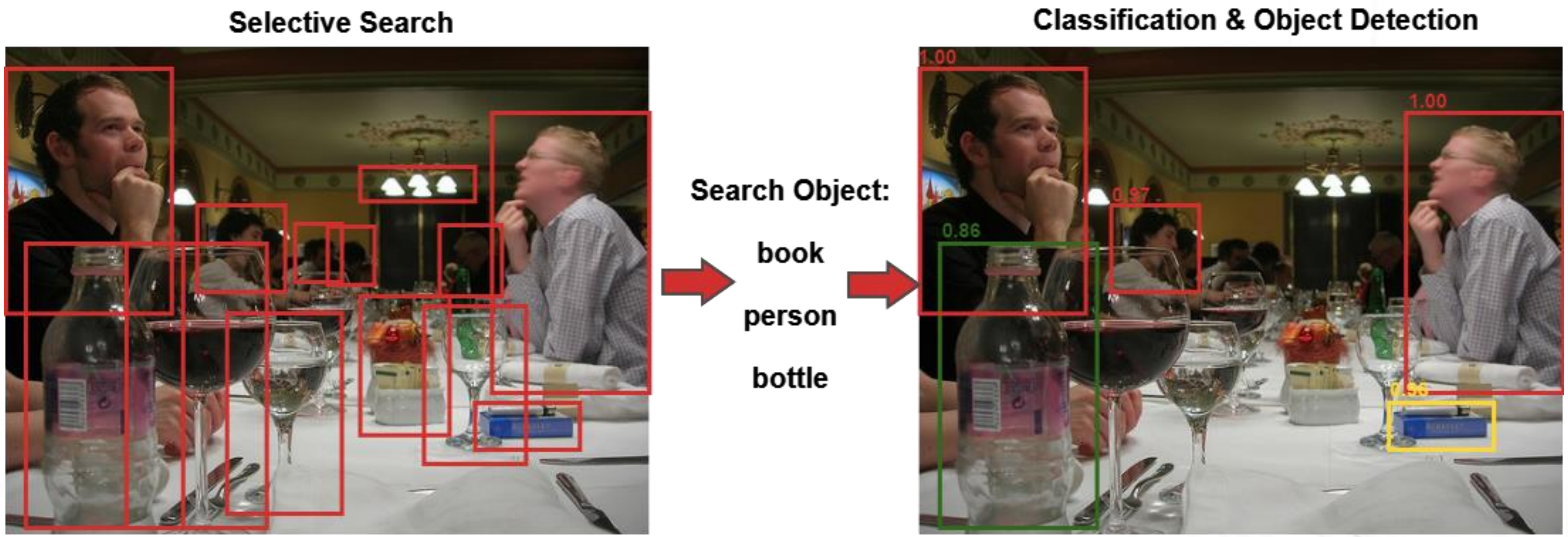
Huang, Gao, et al., "Densely connected convolutional networks," Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700-4708, 2017.

# Related Work – Object Detection

- Object Detection Network

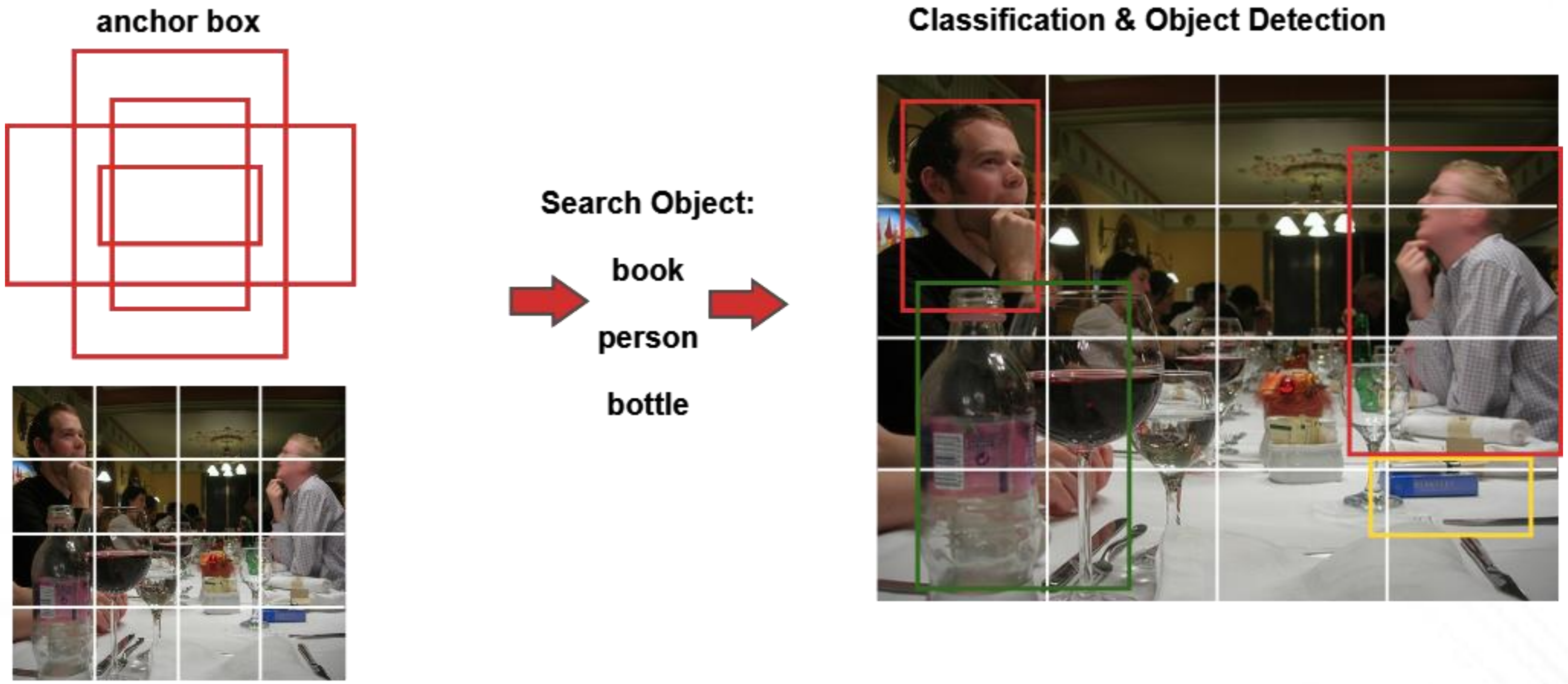
# Related Work – Object Detection

- Object Detection Network – Two Stage



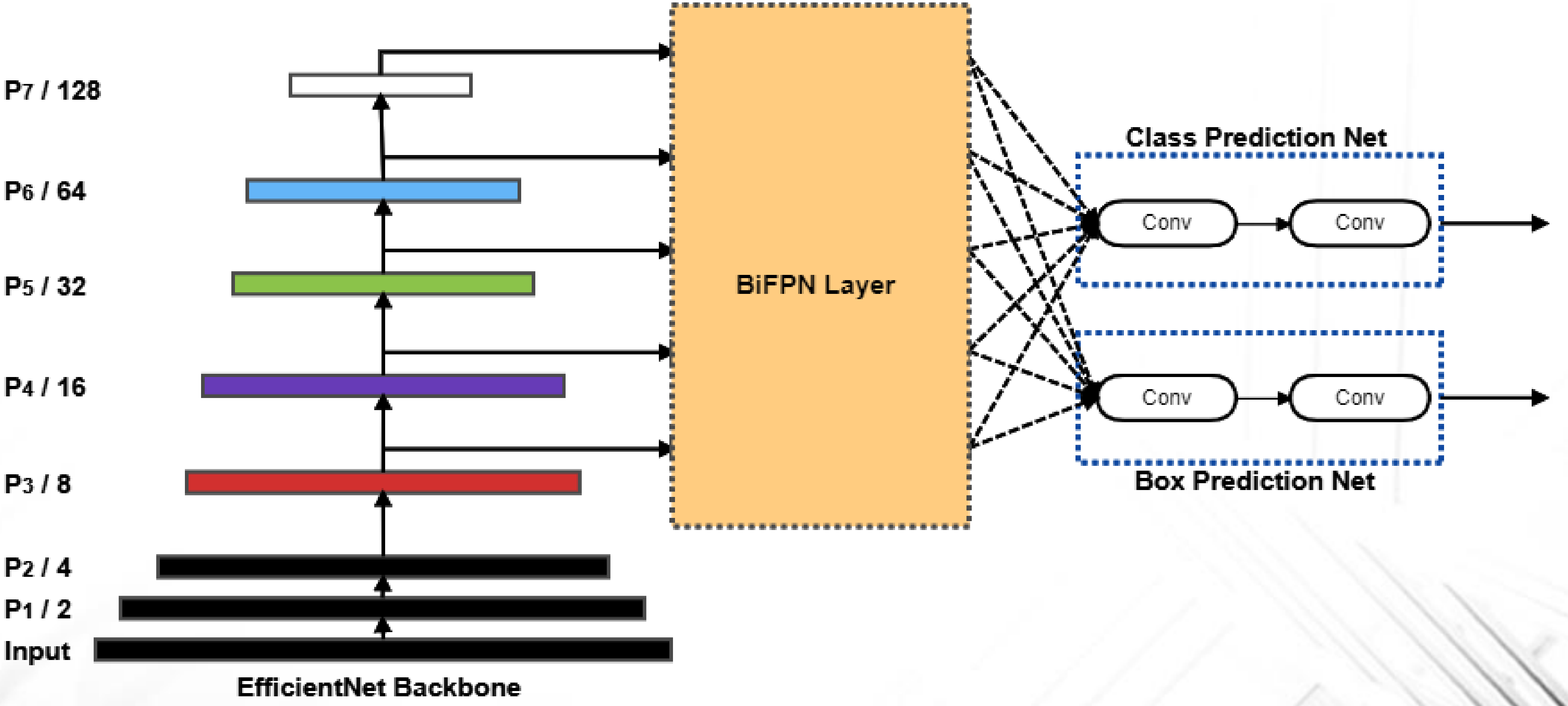
# Related Work – Object Detection

- Object Detection Network – One Stage



# Related Work – Object Detection

- Object Detection Network – EfficientDet

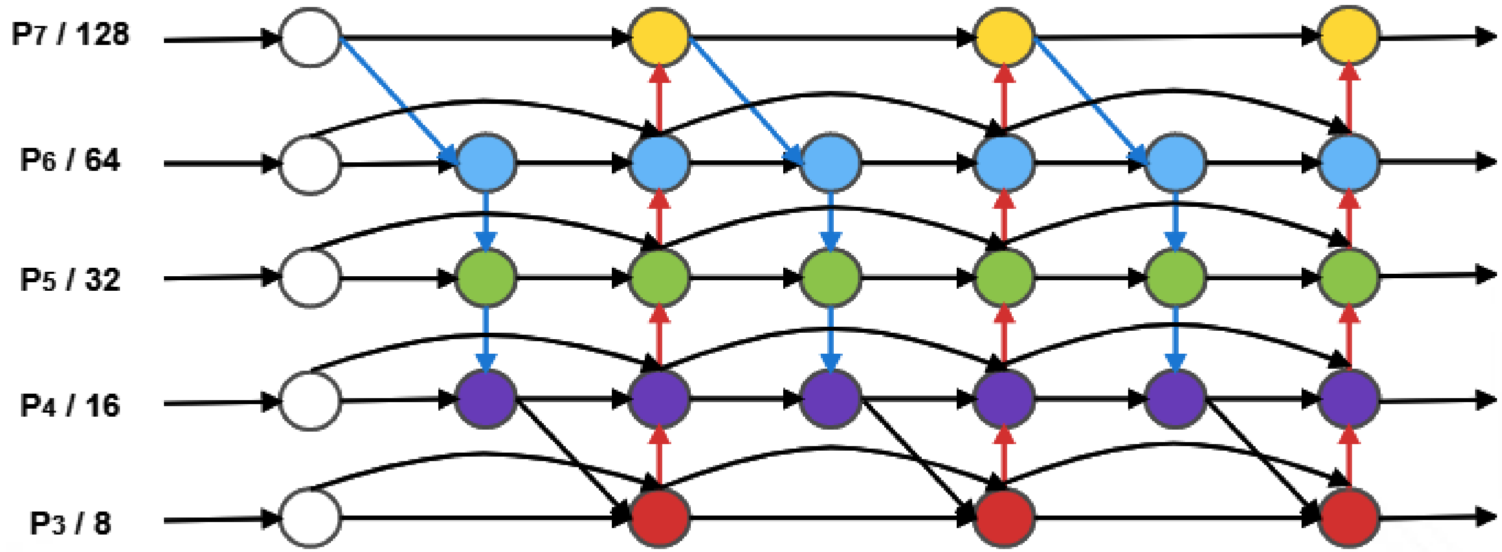





Tan, M., R. Pang, and Q. V. Le, "Efficientdet: Scalable and efficient object detection. arXiv," arXiv preprint arXiv:1911.09070, 2019.

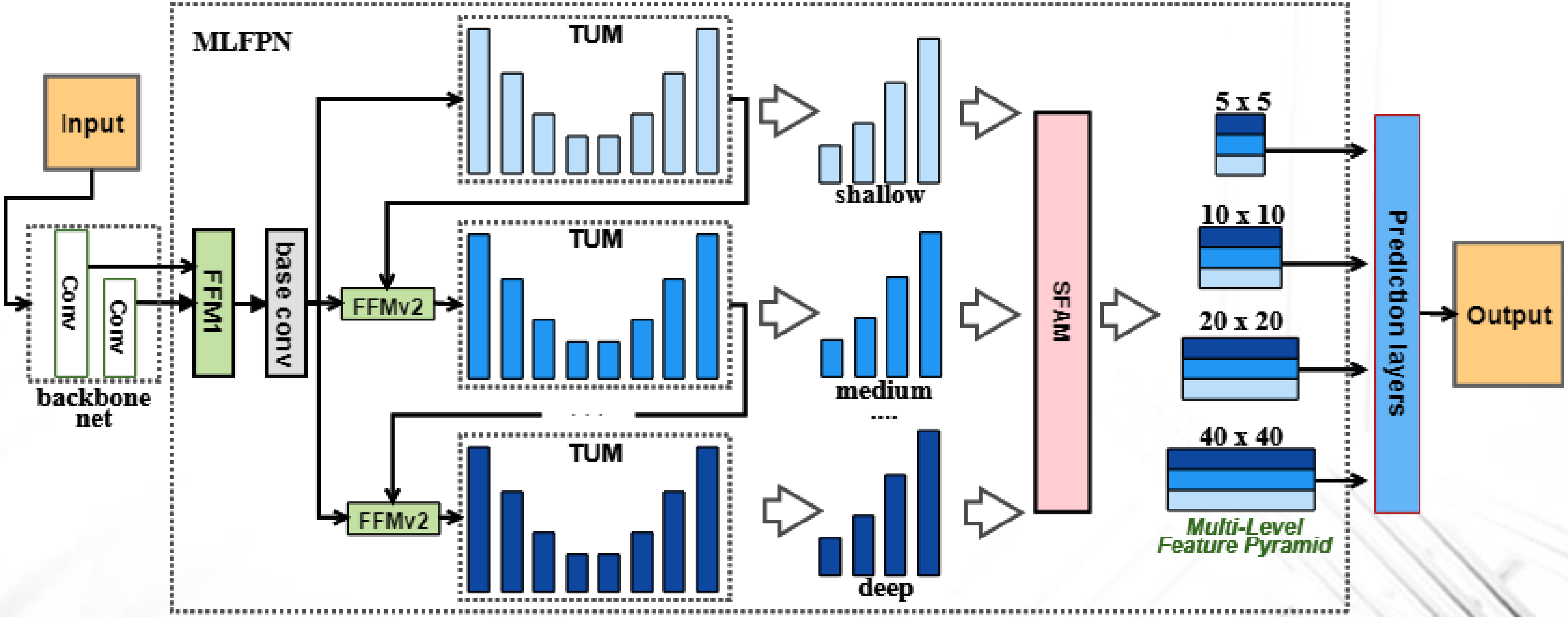
# Related Work – Object Detection

- Object Detection Network – EfficientDet



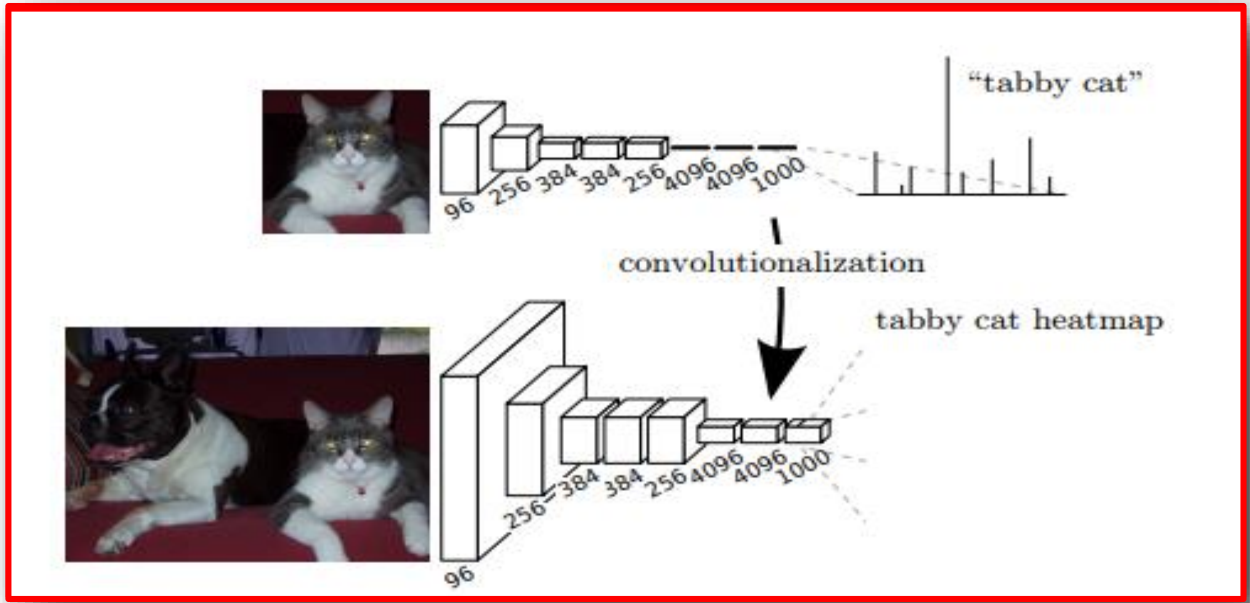
# Related Work – Object Detection

- Object Detection Network – M2Det



# Related Work – Semantic Segmentation

- Fully Convolutional Network




Flattening

00100011000010000100

Spatial Info

| | | | | |
|---|---|---|---|---|
| 0 | 0 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 |
| 0 | 0 | 1 | 0 | 0 |


Long, Jonathan, Evan Shelhamer, and Trevor Darrell, "Fully convolutional networks for semantic segmentation," Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3431-3440, 2015.

# Motivation

| Model | backbone | fps | accuracy |
|---|---|---|---|
| Object Detection Network (COCO dataset，$1280 \times 1080$) | | | |
| EfficientDet-D1 | EfficientNet | $27 \pm 5$ | 45% (mAP) |
| YOLOv3 | Darknet53 | $40 \pm 5$ | 31% (mAP) |
| Semantic Segmentation Network (COCO dataset，$1280 \times 1080$) | | | |
| R-FCN | VGG16 | $5.9 \pm 5$ | 29% (mIoU) |
| DeepLab v3+ | Xception | $8 \pm 5$ | 85% (mIoU) |

# Motivation

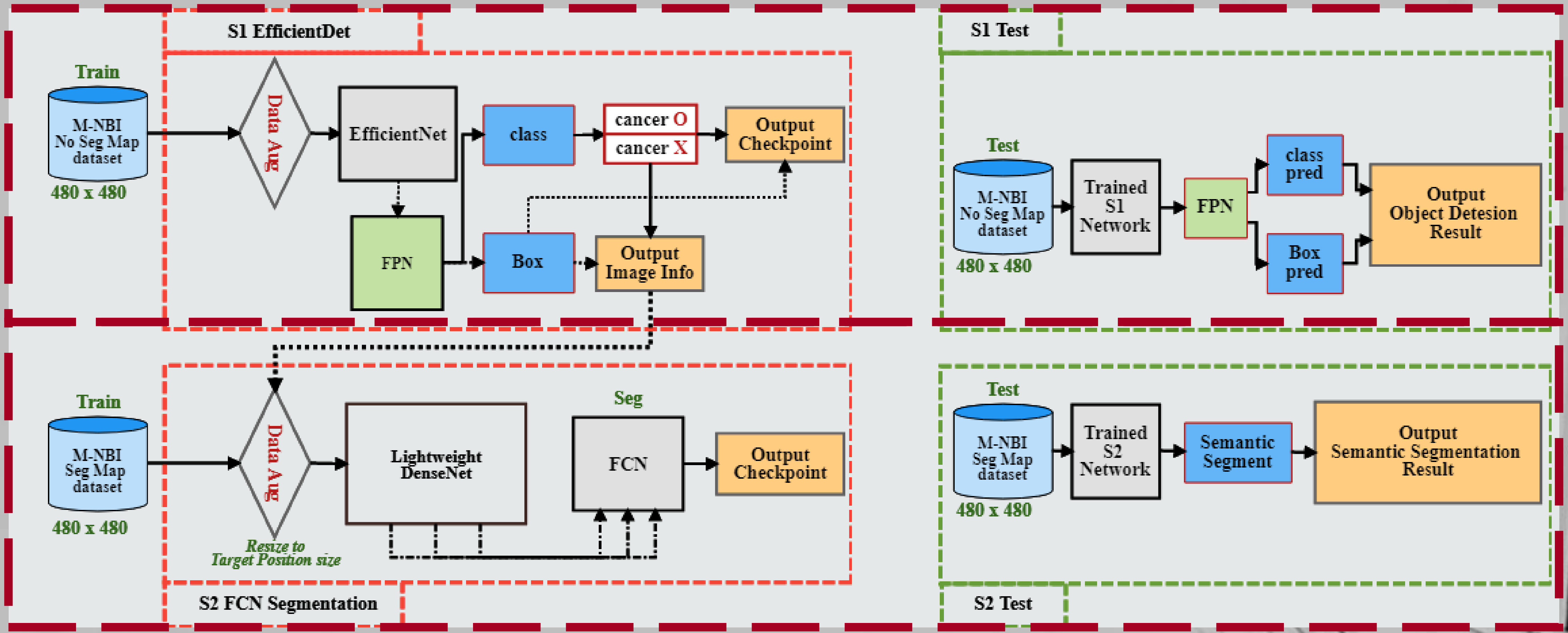
- Need to improve the current method of Data Augmentation.

  ※Generate more datasets.

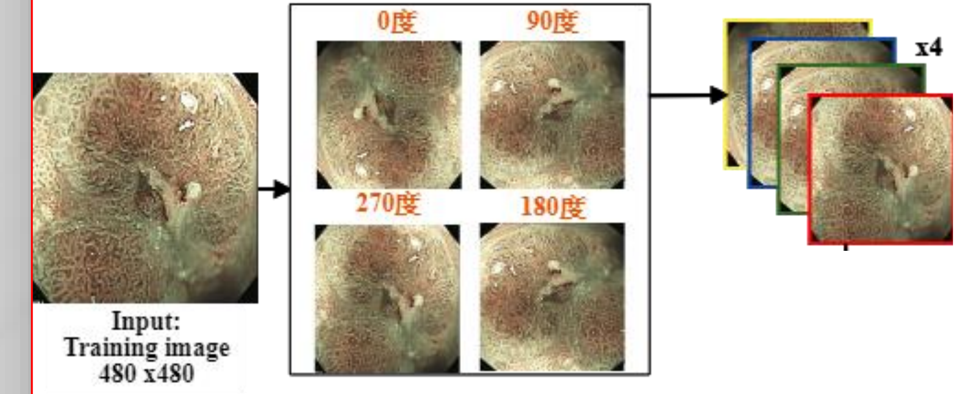
# Method

- Propose Method – S1 and S2
  - S1  Stage : Use EfficientDet for Object Detection and Classification.
  - S2 Stage : Use the adjusted FCN Net and Dense Net for Semantic Segmentation.
- Combined Two Network (S1+S2)
- Data Augmentation
  - Use special pre-process method to increase the number of samples before starting training.

# Method – End-to-End Training Framework



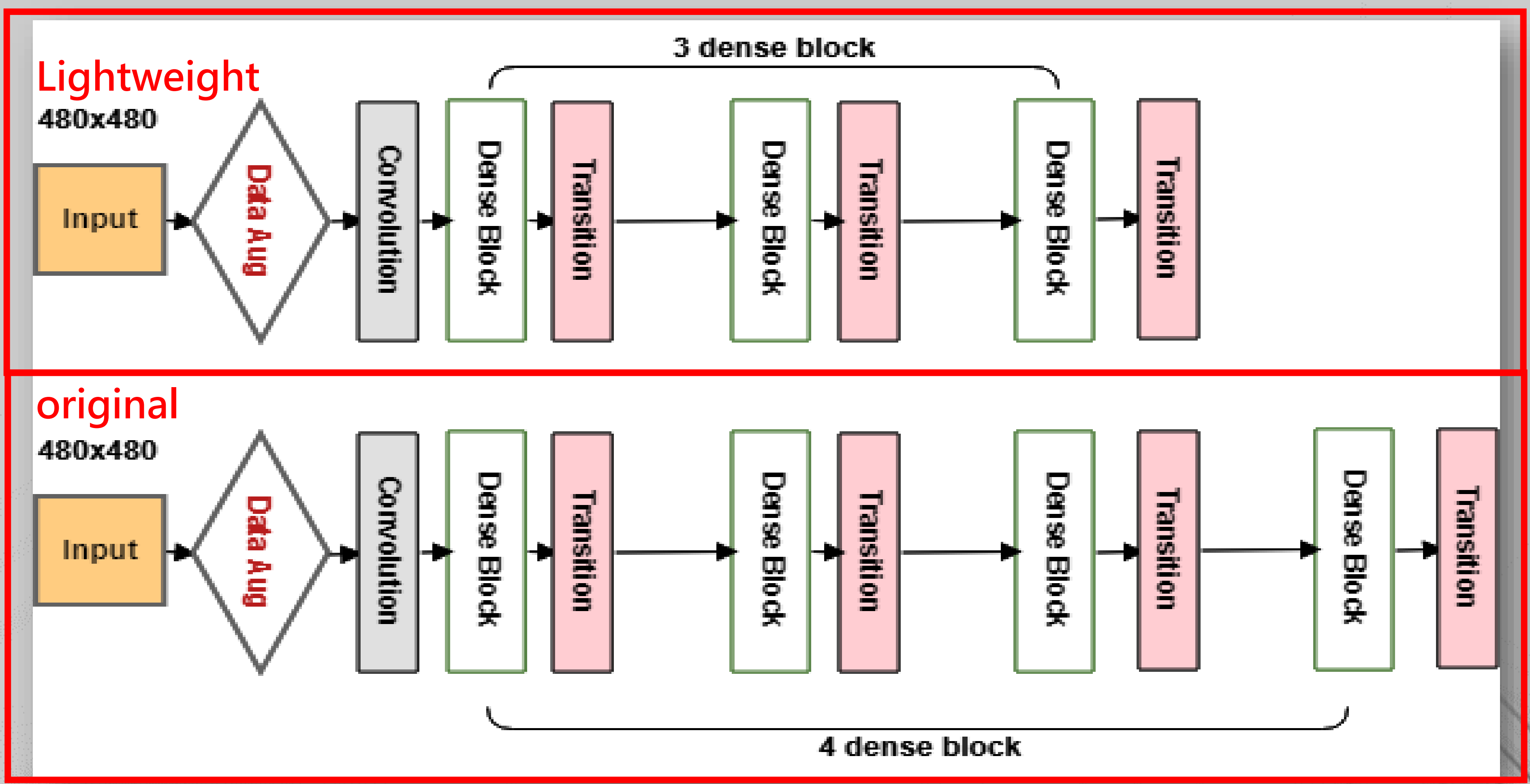
# Method – Special Data Augmentation



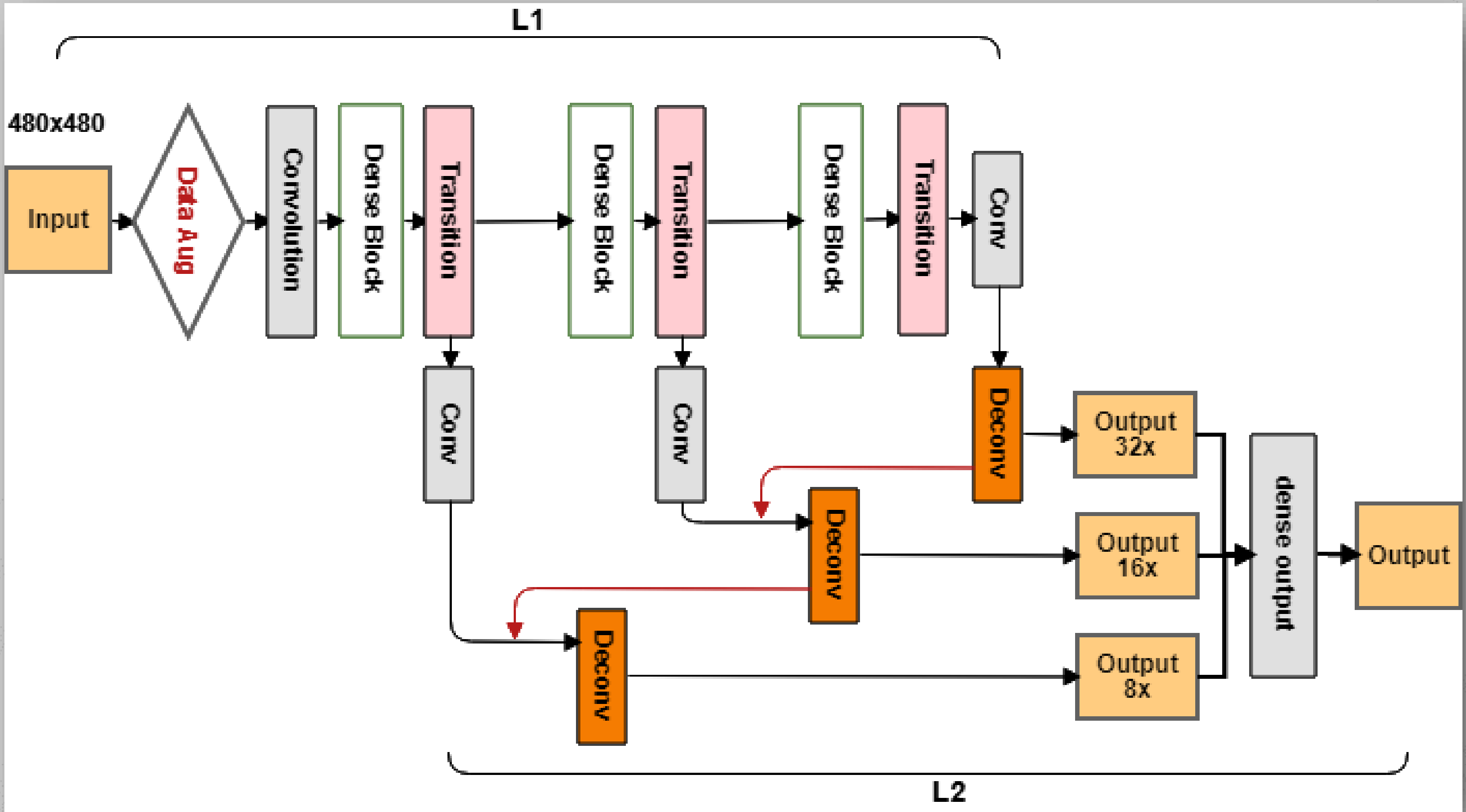
# Method – EfficientDet

| Model (Backbone) | Top-1 Acc | Top-5 Acc | Params | FLOPS | Detection Network |
|---|---|---|---|---|---|
| EfficientNet-B0 | 76.3% | 93.2% | 5.3M | 0.39 B | EfficientDet-D0 |
| EfficientNet-B1 | 78.8% | 94.4% | 7.8 M | 0.70 B | EfficientDet-D1 |
| EfficientNet-B2 | 79.8% | 94.9% | 9.2 M | 1.0 B | EfficientDet-D2 |
| EfficientNet-B3 | 81.1% | 95.5% | 12 M | 1.8 B | EfficientDet-D3 |
| EfficientNet-B4 | 82.6% | 96.3% | 19 M | 4.2 B | EfficientDet-D4 |
| EfficientNet-B5 | 83.3% | 96.7% | 30 M | 9.9 B | EfficientDet-D5 |
| EfficientNet-B6 | 84.0% | 96.9% | 43 M | 19 B | EfficientDet-D6 |
| EfficientNet-B7 | 84.4% | 97.1% | 66 M | 37 B | EfficientDet-D7 |

# Method – Lightweight DenseNet



# Method – Lightweight FCN



# Method – Evaluation Metrics

## Fast Normalized Fusion

$$O = \sum_i \frac{w_i}{\epsilon + \sum_j w_j} \cdot I_i \quad , \quad w_i = \max(0, x)$$

## Cross Entropy

$$H = \sum_{i=1}^{n} -y_{C,i} \log_2(p_{C,i})$$

# Method – Evaluation Metrics

## Confusion Matrix

| | | Actual Class | |
|---|---|---|---|
| Predicted Class | | True | False |
| | Positive | True Positive (TP) | False Negative (FN) |
| | Negative | False Positive (FP) | True Negative (TN) |

| | |
|---|---|
| **True Positive** | 是癌症圖像且模型判斷是癌症圖像 |
| **False Negative** | 是癌症圖像但模型沒有判斷出來是癌症圖像 |
| **False Positive** | 不是癌症圖像但模型判斷是癌症圖像 |
| **True Negative** | 不是癌症圖像且模型沒有判斷是癌症圖像 |

# Method – Evaluation Metrics

| | |
|---|---|
| $Specificity = \frac{TN(nocancer)}{FP + TN} \times 100\%$ | $Sensitivity = \frac{TP(cancer)}{TP + FN} \times 100\%$ |
| $F1 - score = 2 \times \frac{Specificity \times Sensitivity}{Specificity + Sensitivity} \times 100\%$ | $\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + TN}$ |
| $\text{mAP} = \frac{\sum_{th=\{0.5,0.55,\ldots,0.75\}} AP(IoU_{th} = th)}{x}$ , $\text{AP} = \frac{1}{n}\sum_{r} P_{interp}(r)$ | |

# Experiment Result

| First stage network parameter setting | | | | | | |
|---|---|---|---|---|---|---|
| **Parameter** | Input size | Optimizer | Learning Rate | momentum | Activation Function | iteration |
| **Value** | 480 | Adam | $10^{-4}$ | 0.997 | Relu | 1000 |

| Second stage network parameter setting | | | | | | |
|---|---|---|---|---|---|---|
| **Parameter** | Input size | Optimizer | Learning Rate | momentum | Activation Function | iteration |
| **Value** | 480 | Adam | $5 \times 10^{-5}$ | 0.9 | Relu | 1000 |

# Experiment Result - Outline

- Experiment Result (Multiple box and Single box)
- Experiment Result I (S1 - Object Detection Network Result)
- Experiment Result II (S2 - Semantic Segmentation Network Result)
- Experiment Result III (S1+S2 - Combined Network Result)

# Experiment Result

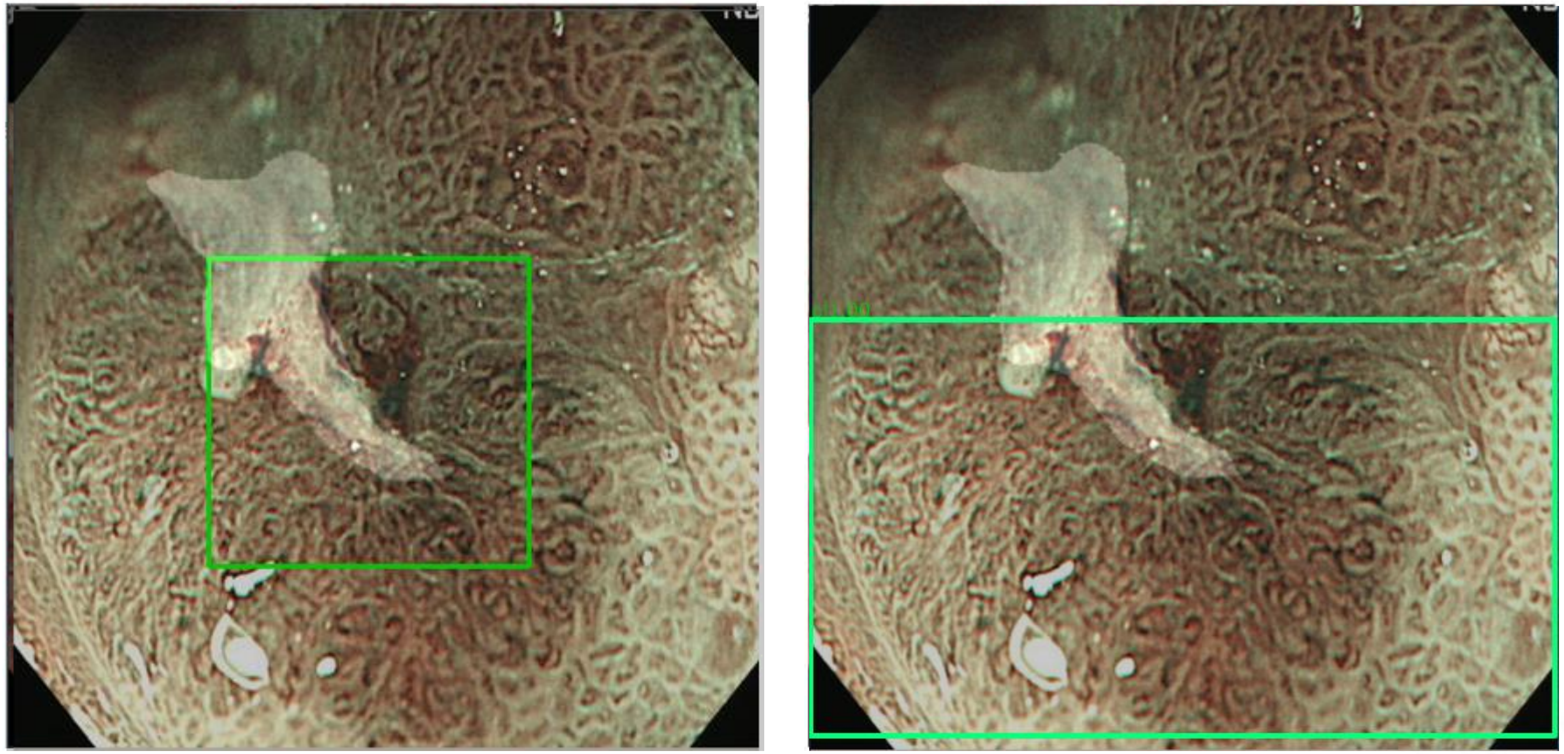


Fig 1. Multiple box function and Single box function.

# Experiment Result I


| AP (IoU) | FPS |
|---|---|
| 36.4% | 17±2 |
| 74.6% | 17±2 |

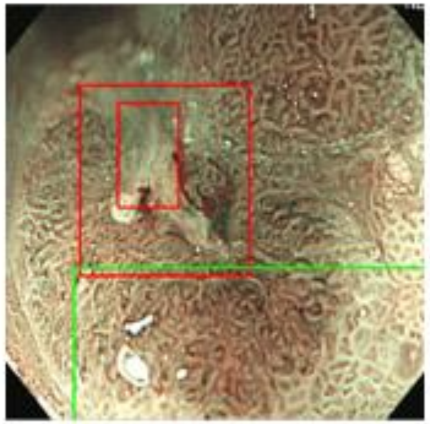
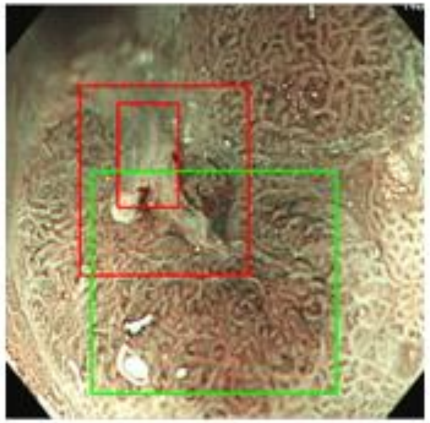
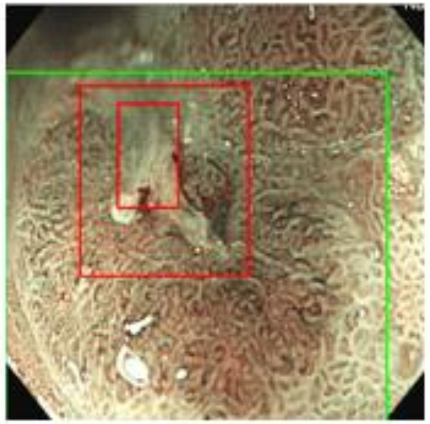
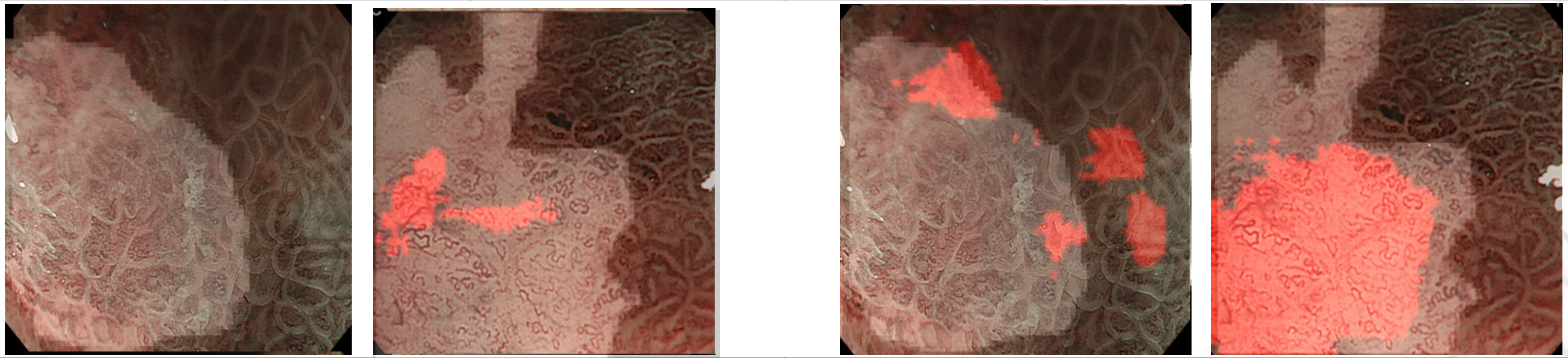
| Single Detection | Average Detection | Max Detection |
|---|---|---|
| Detection Box= 5 | | |
|  |  |  |

Fig 2. Each detection box result.

# Experiment Result II

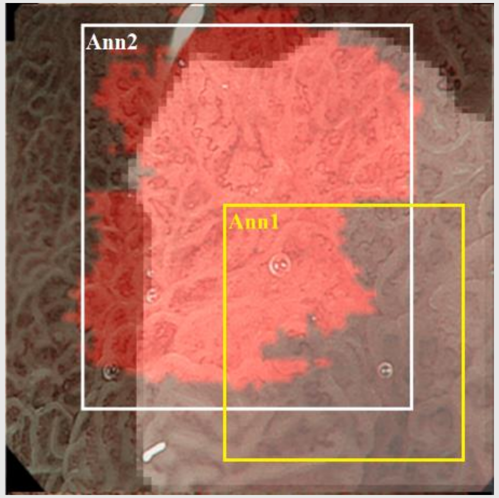
| Epoch=1200, Input Size=$480 \times 480$, Test data = 20 張 | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Model** | My Data Augmentation | Sensitivity | F1-socre | Accuracy | TP | FN | FPS |
| **Lightweight DenseNet** | X | 42.8% | 59.1% | 60% | 6 | 8 | 60 |
| **Lightweight DenseNet** | V | 64.2% | 72.5% | 70% | 9 | 5 | 60 |
| **Epoch=900, Input Size=$480 \times 480$, Test data = 20 張** | | | | | | | |
| **Lightweight DenseNet** | X | 40% | 57% | 55% | 6 | 9 | 60 |
| **Lightweight DenseNet** | V | 28.5% | 42.5% | 45% | 4 | 10 | 60 |



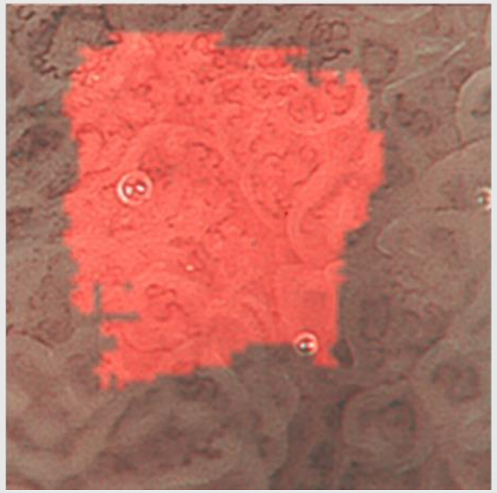
| None Data Augmentation | Data Augmentation |
|---|---|

# Experiment Result III

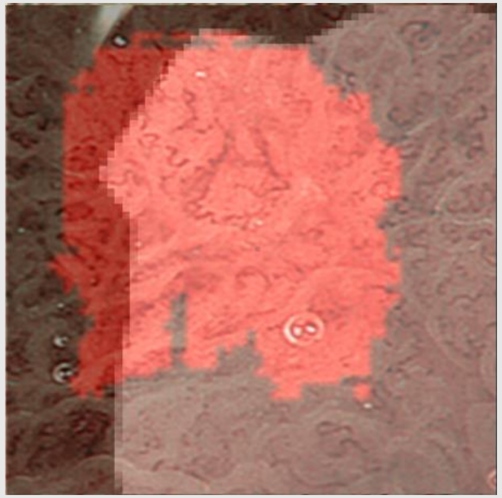
| Model | Sensitivity | F1-score | TP | FN | Accuracy |
|---|---|---|---|---|---|
| Semantic Segmentation Network | 64.2% | 72.5% | 9 | 5 | 70% |
| Two Stage Detection Network-Ann1 | 83.3% | 85.3% | 10 | 2 | 87.5% |
| Two Stage Detection Network-Ann2 | 84.6% | 91.6% | 11 | 2 | 89.7% |




Semantic Segmentation Network



Two Stage Detection Network-Ann1



Two Stage Detection Network-Ann2

# Conclusion

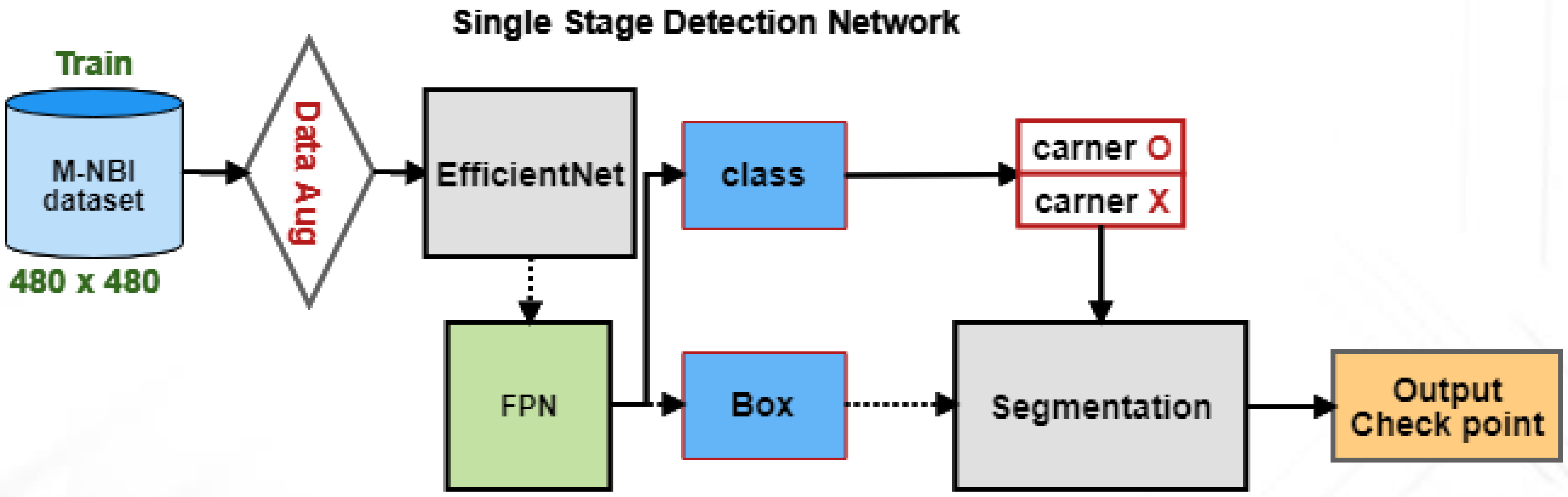
- Build a two-stage high-performance cancerous detection network based on EfficientDet and lightweight DenseNet.
- Using 「Special Data Augmentation」 method can make Small-Datasets comparable to Large-Datasets.

---

- EfficientDet || AP=90% ; FPS=70
- Lightweight DenseNet FCN || Accuracy=70% ; FPS=60
- Combined Network || Accuracy=89% ; FPS=50

# Conclusion

- The accuracy of the combined network is less than 90%



# The End

Thank you for your listening !!!