Medical Vision Seminar

Jinyue Cai
2021.7.7

Outline:

- DARCNN: Domain Adaptive Region-based Convolutional Neural Network for Unsupervised Instance Segmentation in Biomedical Images (CVPR2021)

- Towards Unbiased COVID-19 Lesion Localisation and Segmentation via Weakly Supervised Learning (ISBI2021)

(CVPR2021)

DARCNN: Domain Adaptive Region-based Convolutional Neural Network for Unsupervised Instance Segmentation in Biomedical Images

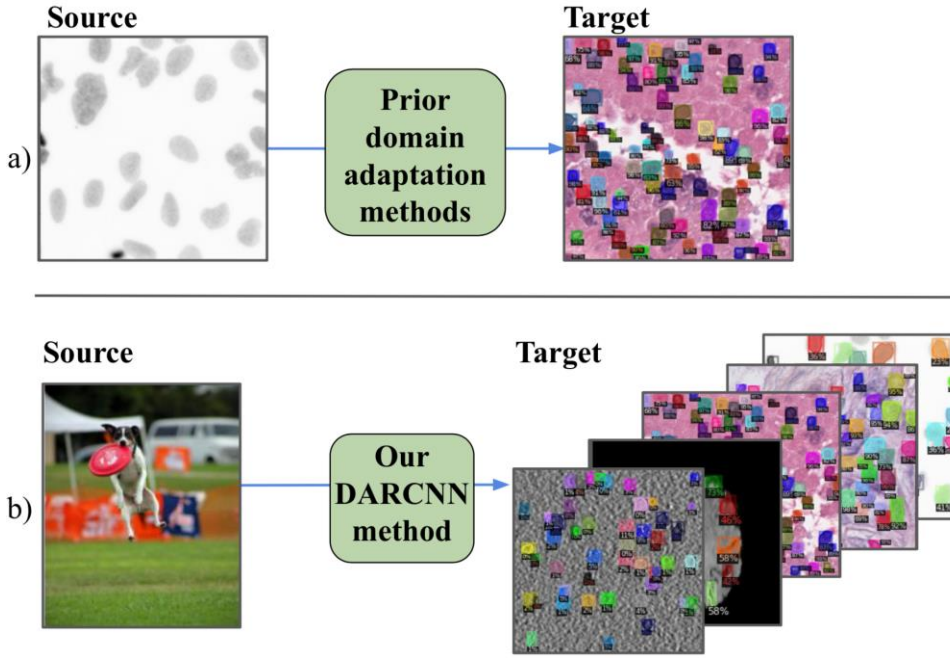
Joy Hsu, Wah Chiu, Serena Yeung

Stanford University

Introduction

Problem:

1. Labelled domain specific datasets are often expensive to obtain
2. It is not always feasible to find similar labelled biomedical datasets

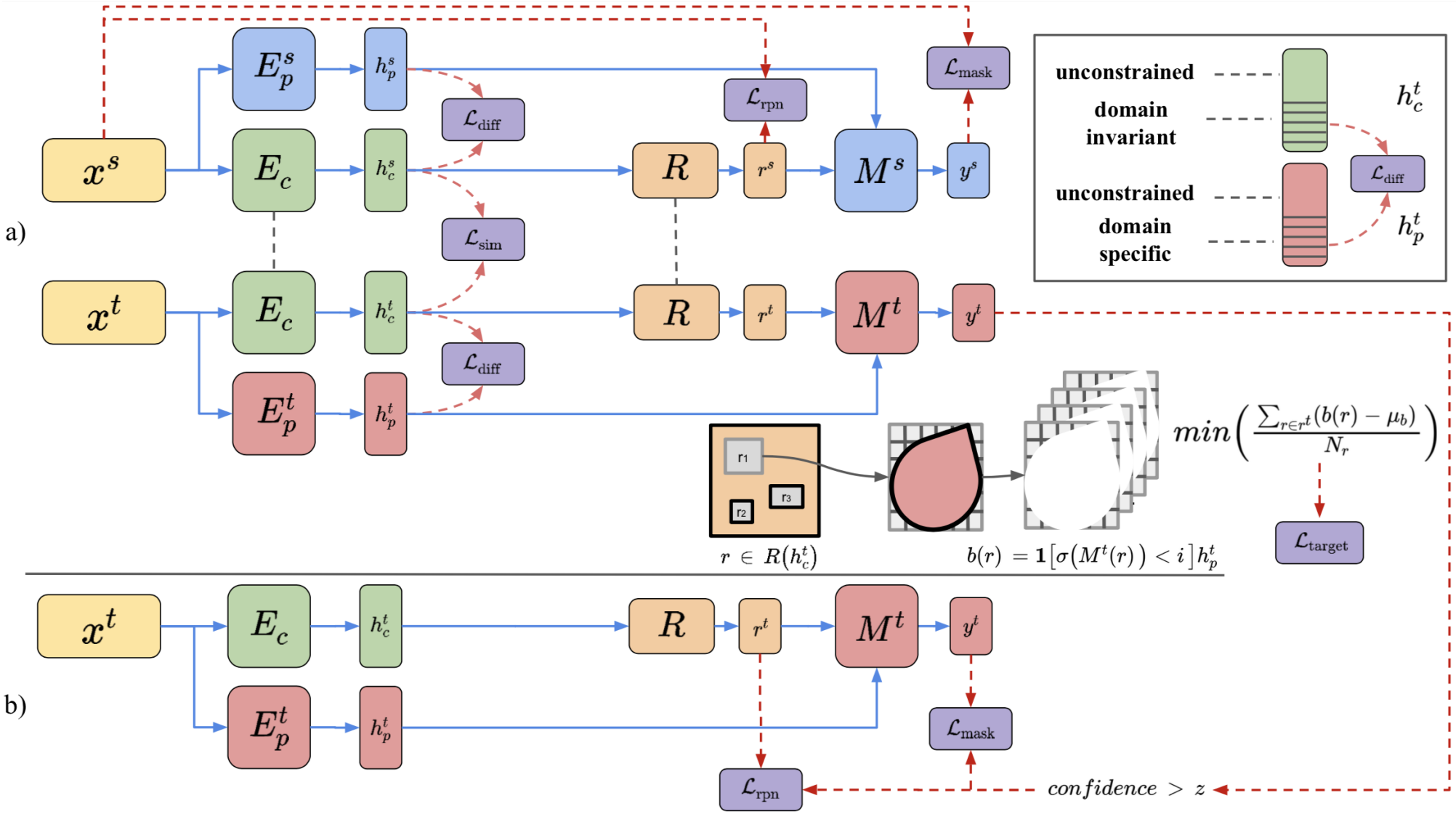


Contribution

1. Unsupervised instance segmentation adapting from COCO to biomedical images
2. Introduce a domain separation module to learn domain invariant and domain specific features
3. Propose a self-supervised background consistency loss
4. Utilize pseudo-labelling with data augmentation

Methods

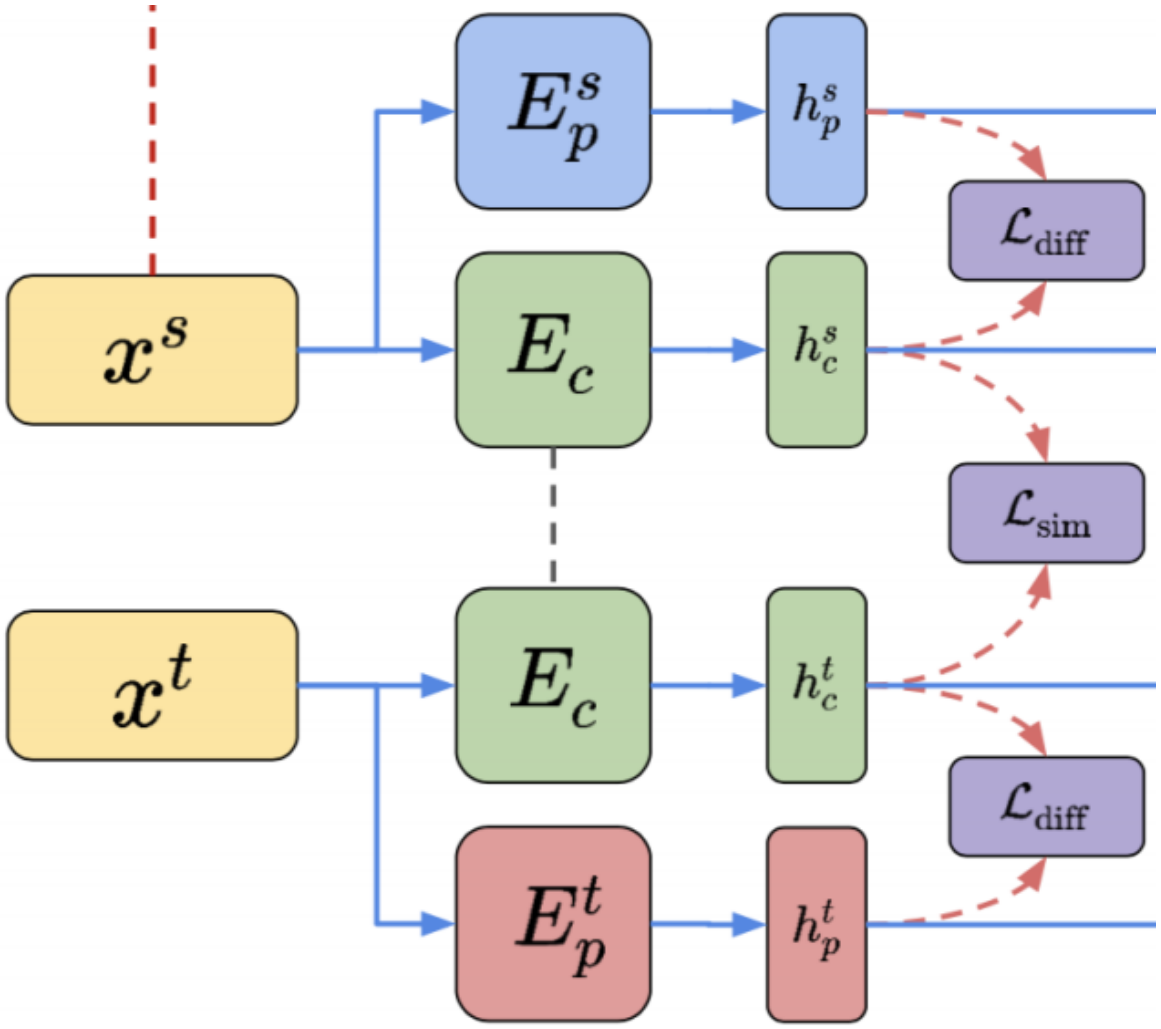
- Overview



$$L_{\text{DARCNN}} = \alpha L_{\text{sim}} + \beta L_{\text{diff}} + \gamma L_{\text{target}} + L_{\text{source}}$$

Methods

- Domain Separation Module

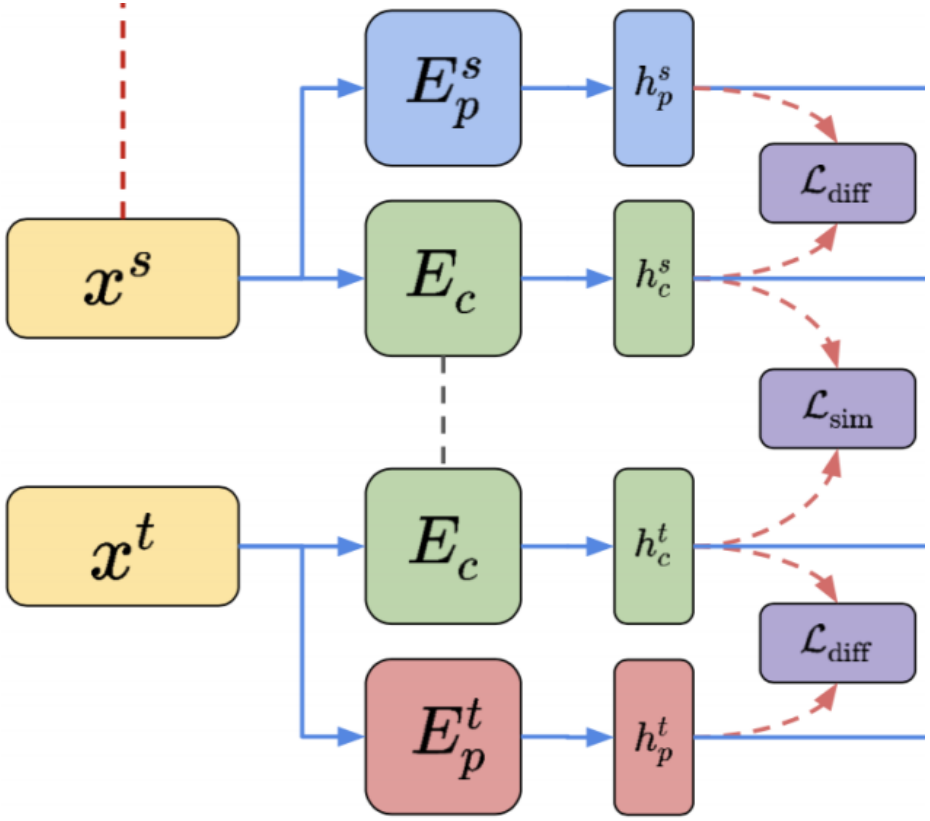


domain invariant features encode objectness of the source and target domain in a joint representational subspace

domain specific features capture discriminability of each domain as well as contain additional unconstrained embedding space

Methods

- Domain Invariant Features

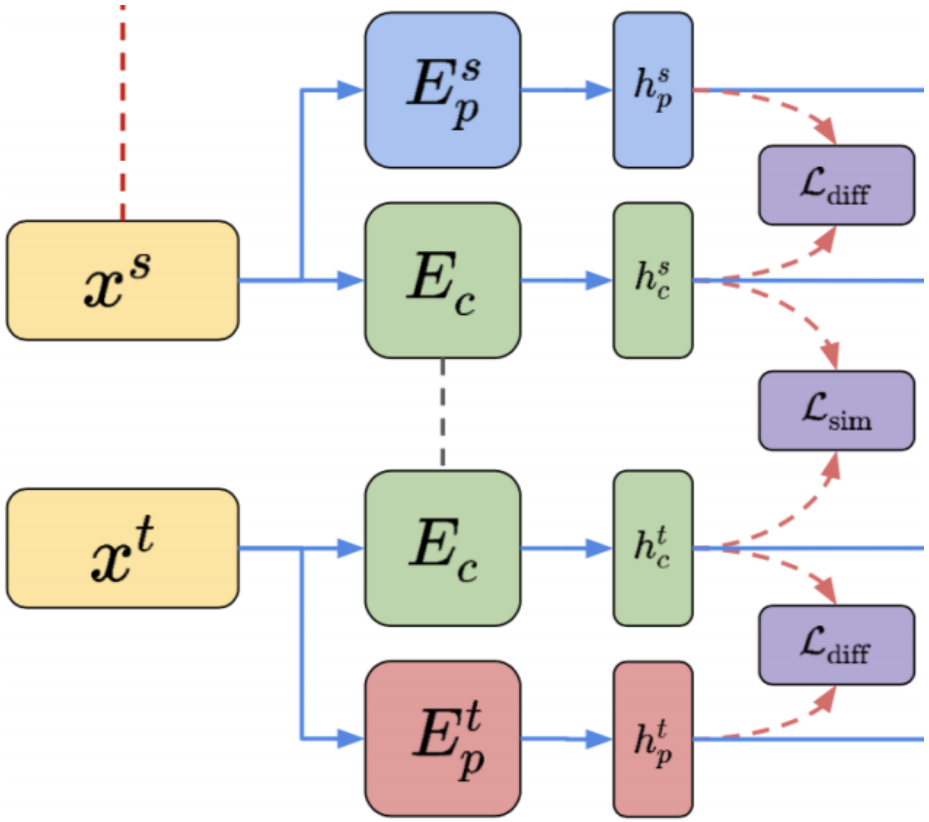


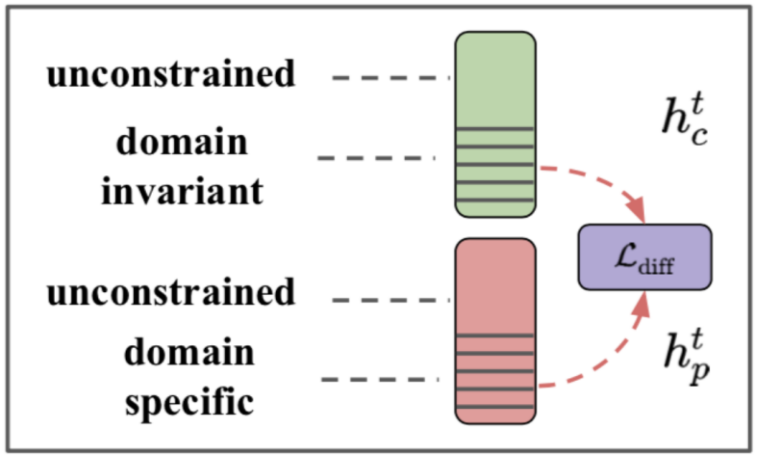
$$L_{\text{sim}} = \frac{1}{(N^{\text{s}})^2} \sum_{i,j=0}^{N^{\text{s}}} \kappa(h_c^{s,i}, h_c^{s,j}) \qquad (2)$$

$$- \frac{2}{N^{\text{s}} N^{\text{t}}} \sum_{i,j=0}^{N^{\text{s}} N^{\text{t}}} \kappa(h_c^{s,i}, h_c^{t,j}) + \frac{1}{(N^{\text{t}})^2} \sum_{i,j=0}^{N^{\text{t}}} \kappa(h_c^{t,i}, h_c^{t,j})$$

Methods

- Domain Specific Features





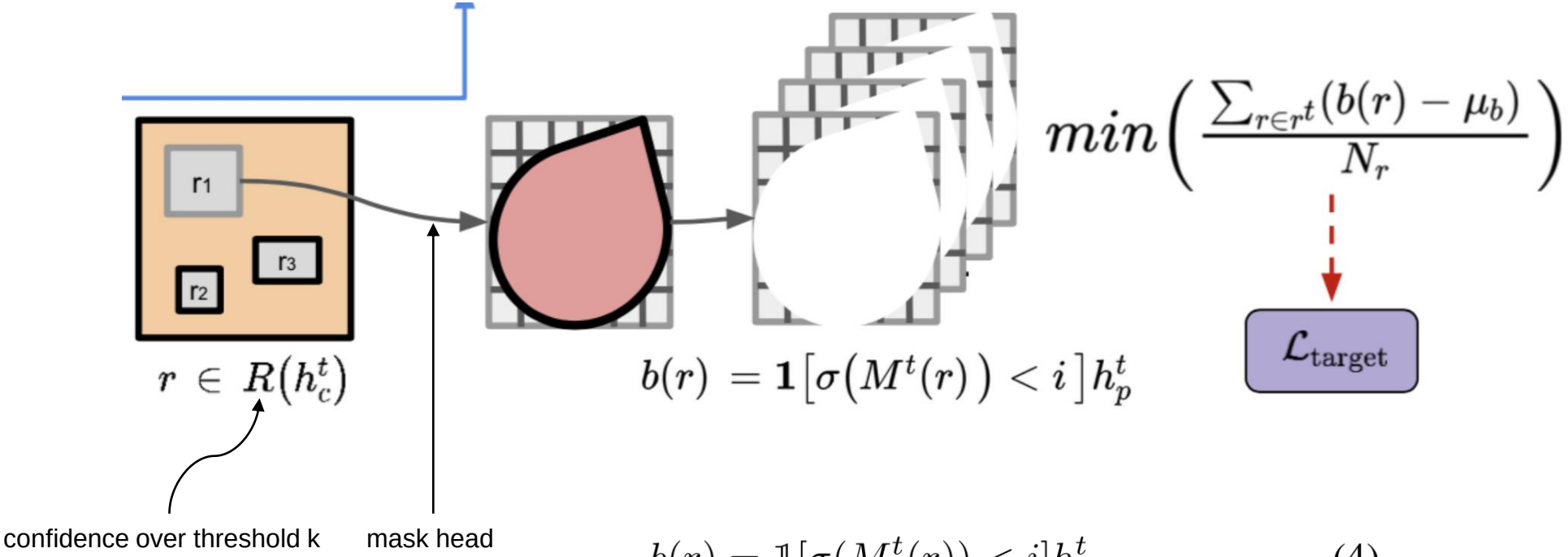
$$L_{\text{diff}} = ||H_c^{s\top} H_p^s||_F^2 + ||H_c^{t\top} H_p^t||_F^2$$

Methods

- Self-supervised Background Consistency Loss

Assumption: the target images contain homogeneous backgrounds

To minimize the differences between background representations of each predicted instance in a image



$$b(r) = \mathbb{1}[\sigma(M^t(r)) < i]h_p^t \tag{4}$$

$$\mu_b = \frac{1}{N_p} \sum_{p \in R(h_c^t)} b(p) \tag{5}$$

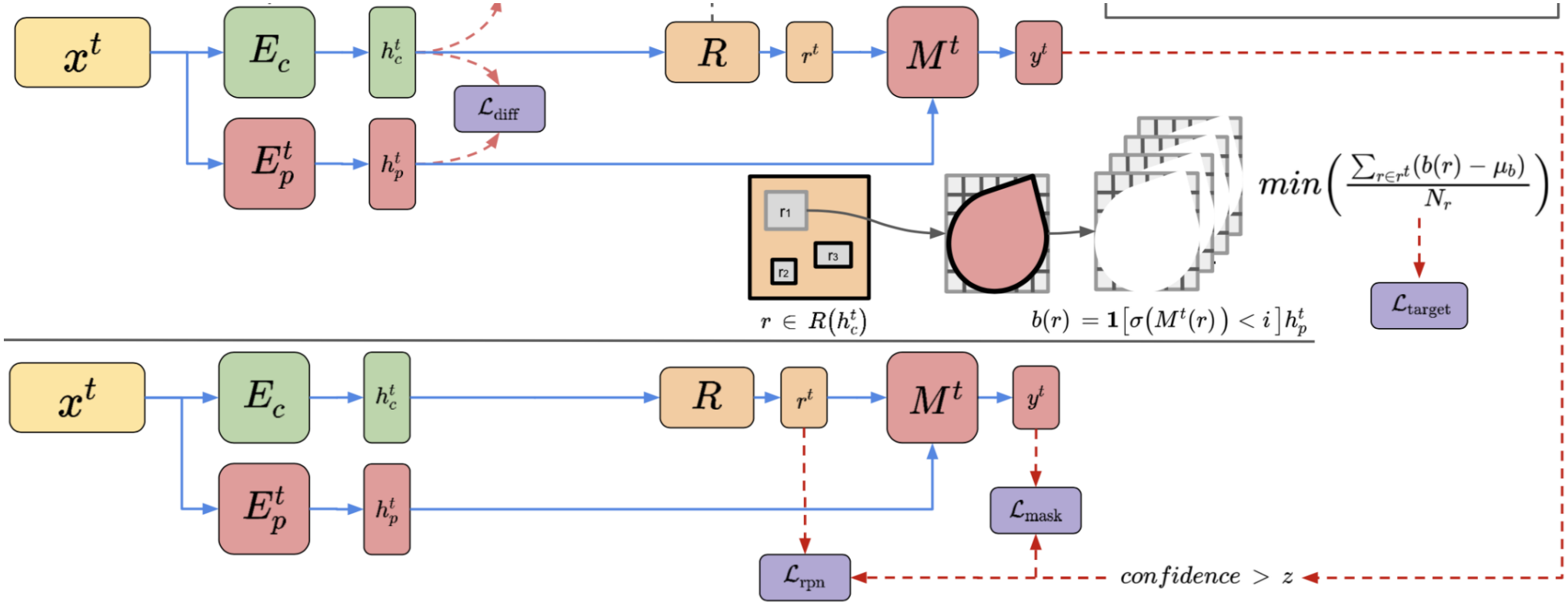
$$L_{\text{target}} = \frac{1}{N_r} \sum_{r \in R(h_c^t)} | b(r) - \mu_b | \tag{6}$$

Methods

- Augmented Pseudo-Labelling

Pseudo-labels: retrieve high confidence (threshold z) predictions from our first stage DARCNN

Data augmentations: lighting, contrast, and blur



Experiments

- Adaptation from Microscopy or COCO to Histopathology

Source Dataset: BBBC (a fluorescence microscopy dataset) or COCO

Target Dataset: Kumar or TNCB (histopathology datasets)

Method	AJI	Pixel-F1	Object-F1
Chen et al. [7]	0.4407	0.6405	0.6289
DDMRL [17]	0.4642	0.7000	0.6872
SIFA [6]	0.4662	0.6994	0.6698
CyCADA [13]	0.4721	0.7048	0.6866
Hou et al. [14]	0.4775	0.7029	0.6779
Liu et al. [21]	0.5672	0.7593	0.7478
Ours from BBBC	0.5120	0.7175	0.6436
Ours from COCO	0.4906	0.6998	0.6396

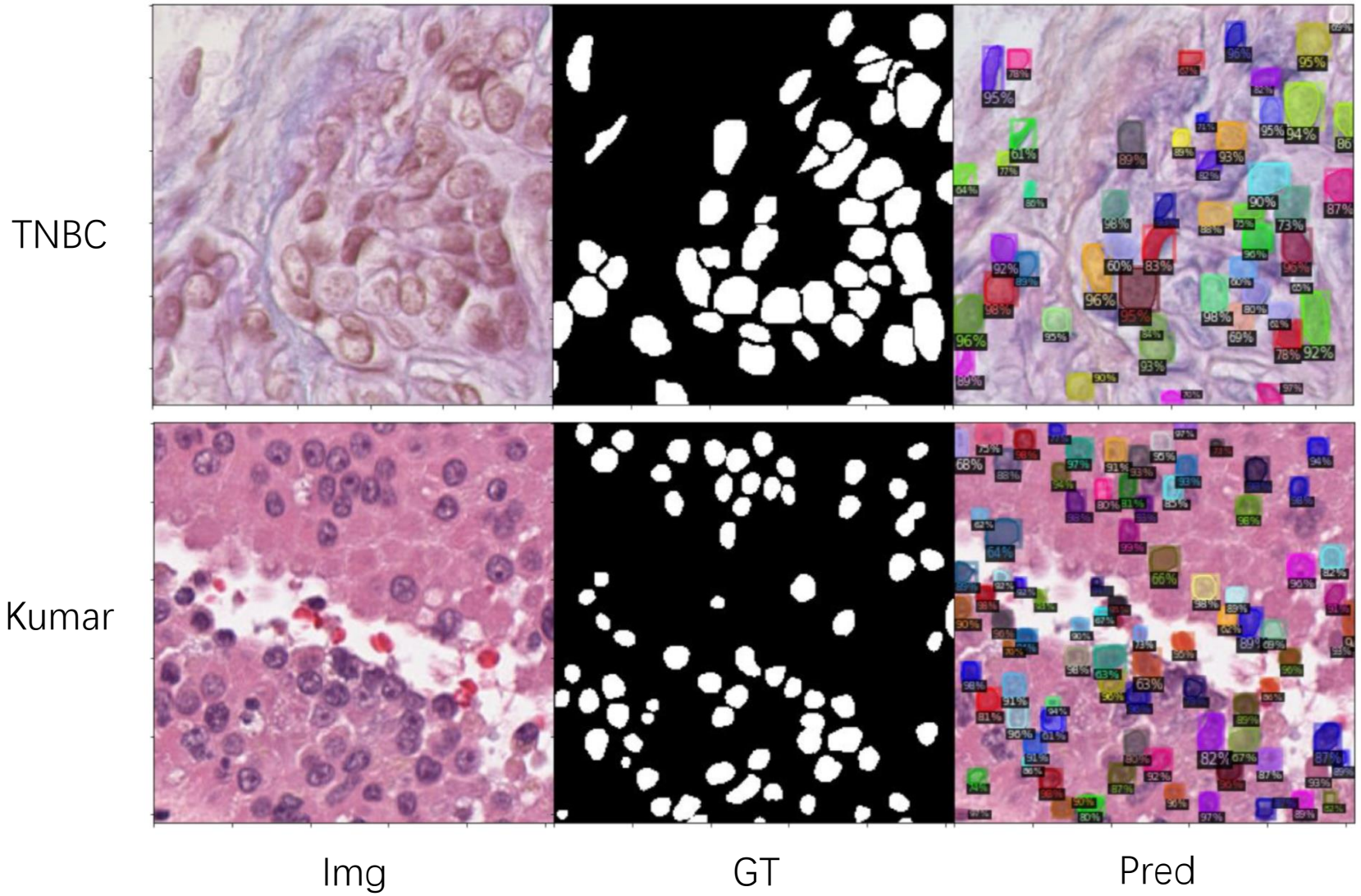
From BBBC or COCO to TNBC

Method	AJI	Pixel-F1	Object-F1
Chen et al. [7]	0.3756	0.6337	0.5737
SIFA [6]	0.3924	0.6880	0.6008
CyCADA [13]	0.4447	0.7220	0.6567
DDMRL [17]	0.4860	0.7109	0.6833
Hou et al. [14]	0.4980	0.7500	0.6890
Liu et al. [21]	0.5610	0.7882	0.7483
Ours from BBBC	0.4461	0.6619	0.5410
Ours from COCO	0.4421	0.6549	0.5104

From BBBC or COCO to Kumar

Experiments

- Qualitative Results (from COCO to TNBC and Kumar)



Ablation Studies

- Qualitative Results (from COCO to TNBC and Kumar)

Method	AJI	Pixel-F1	Object-F1
Mask R-CNN			
w/ COCO pre-trained	0.0060	0.2769	0.0181
w/ synthesized images	0.3332	0.5782	0.6061
First stage DARCNN			
Domain sim. only	0.3687	0.6023	0.6099
Bg. consistency only	0.3808	0.6120	0.5470
Full 1st stage DARCNN	0.4071	0.6353	0.5986
Second stage DARCNN			
Pseudo-label w/o aug	0.4463	0.6781	0.6339
Full 2nd stage DARCNN	0.4906	0.6998	0.6396

Table 3. Ablation study adapting from COCO as source to TNBC.

Method	AJI	Pixel-F1	Object-F1
Mask R-CNN			
w/ COCO pre-trained	0.0315	0.3144	0.0818
First stage DARCNN			
Domain sim. only	0.1414	0.4905	0.4295
Bg. consistency only	0.3250	0.7128	0.5720
Full 1st stage DARCNN	0.3371	0.6409	0.5904
Second stage DARCNN			
Pseudo-label w/o aug	0.4349	0.6914	0.7151
Full 2nd stage DARCNN	0.4725	0.6586	0.6733

Table 4. Ablation study adapting from COCO as source to BBBC.

(ISBI2021)

Towards Unbiased COVID-19 Lesion Localisation and Segmentation via Weakly Supervised Learning

Yang Yang, Jiancong Chen, Ruixuan Wang, Ting Ma, Lingwei Wang,

Jie Chen, Wei-Shi Zheng, Tong Zhang

Artificial Intelligence Research Center, Peng Cheng Laboratory, Shenzhen, China

School of Computer Science and Engineering, Sun Yat-sen University, China

School of Electronic and Information Engineering, Harbin Institute of Technology, Shenzhen, China

Shenzhen People's Hospital, Shenzhen Institute of Respiratory Diseases, Shenzhen, China

School of Electronic and Computer Engineering, Peking University Shenzhen Graduate School, China

Introduction

Problem:

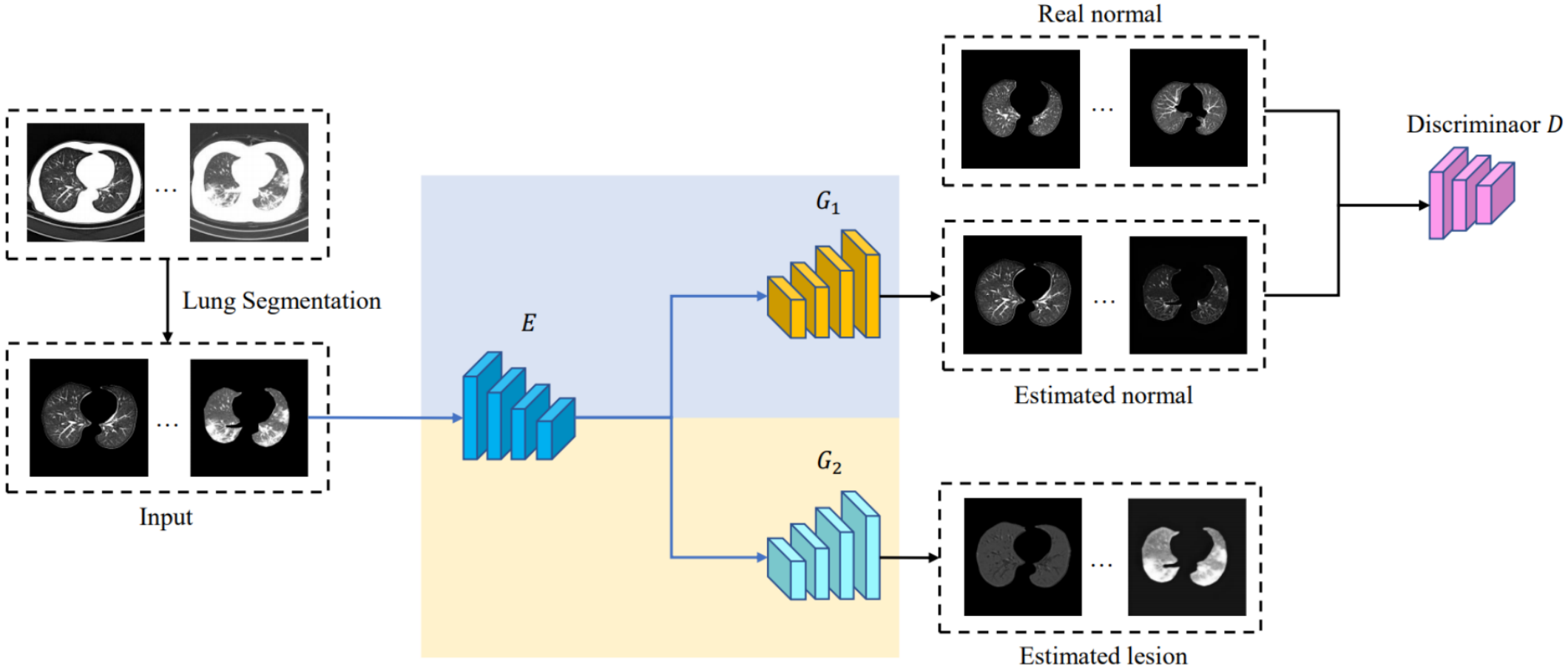
1. Pixel-wise annotation of image dataset is expensive and time-consuming

Contribution

1. Propose a framework for localization and segmentation of COVID-19 pneumonia lesions only with image-level label supervision
2. Decompose biomedical image to normal version and lesion version

Methods

- Overview

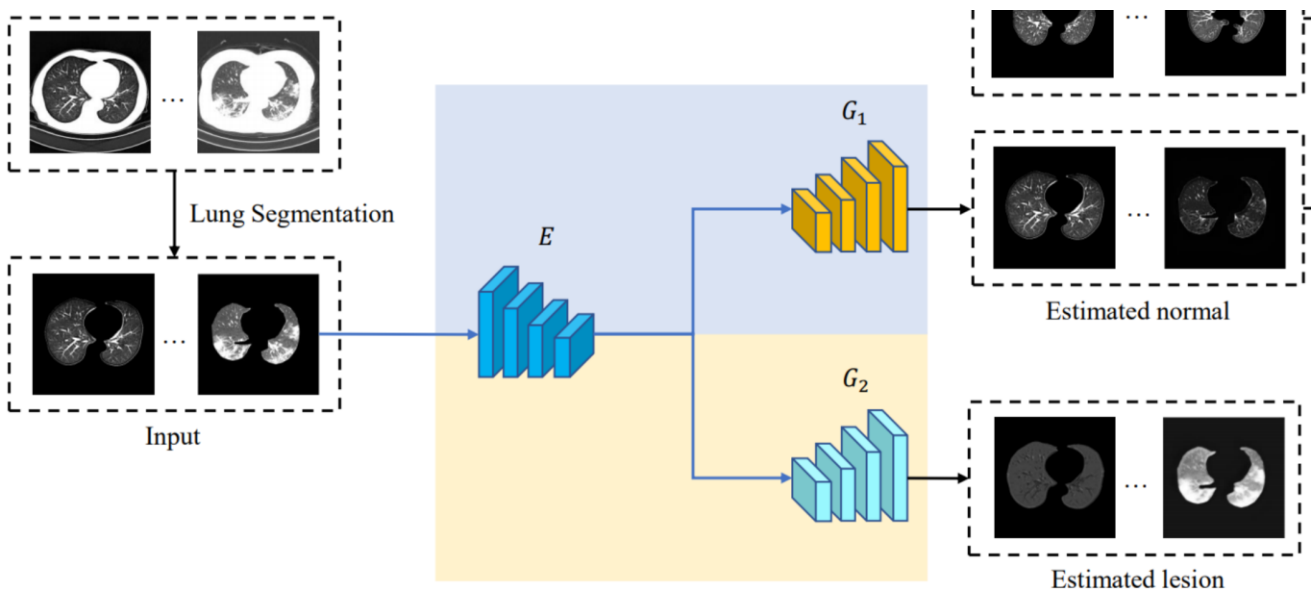


$$L_g = \alpha_1 L_a + \alpha_2 L_r + \alpha_3 L_n \qquad (5)$$

$$L_c = \frac{1}{N}\sum_{i=1}^{N} D(G_1(E(\mathbf{x}_i))) - \frac{1}{N_1}\sum_{j=1}^{N_1} D(\mathbf{x}_j) + \lambda \cdot GP \qquad (3)$$

Methods

- Reconstruction Loss

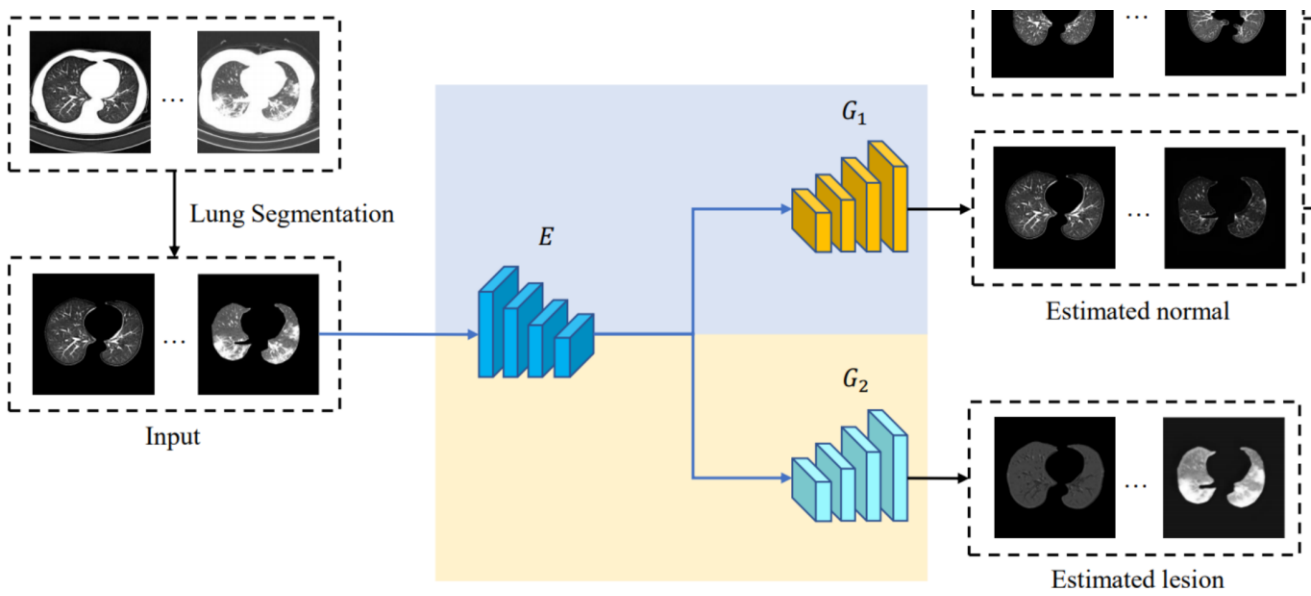


If the decomposition process works well, the recombination of the two decomposed components should be close to the original input

$$L_r = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{x}_i - G_1(E(\mathbf{x}_i)) - G_2(E(\mathbf{x}_i))\| \qquad (1)$$

Methods

- Normal Fidelity Loss

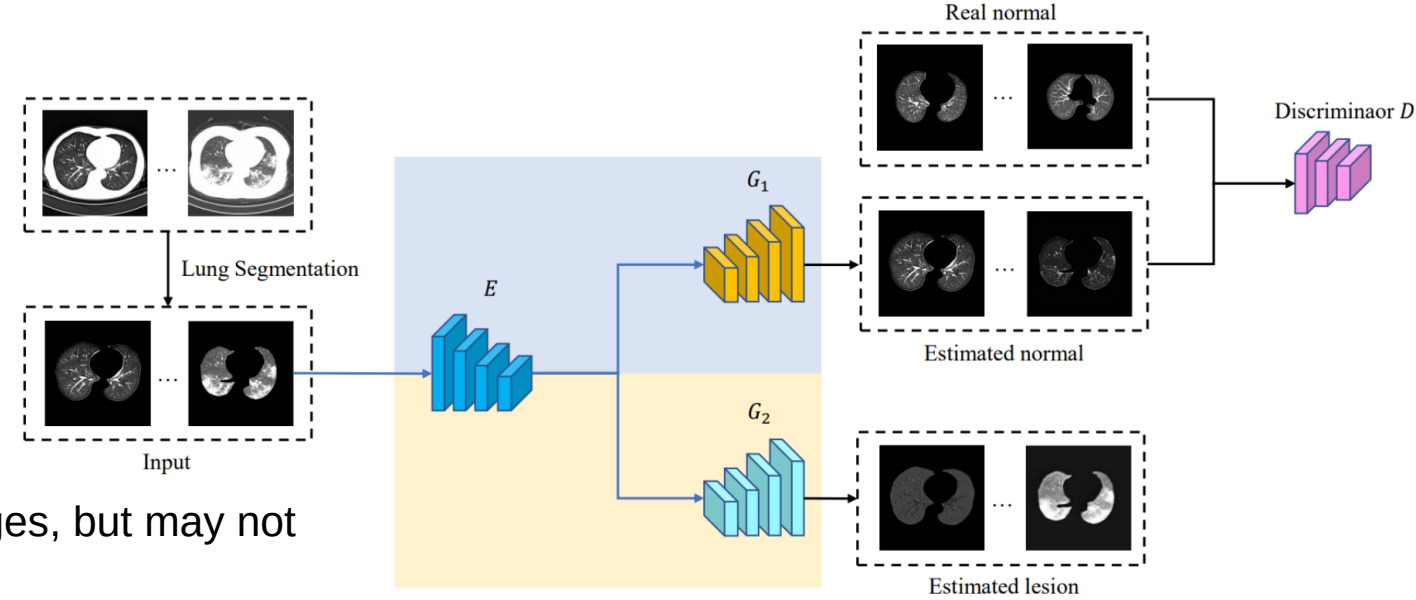


Since normal images contain no lesion, if the decomposition works well, the normal version $G_1(E(xj))$ itself should be close to the original input for any normal input x_j

$$L_n = \frac{1}{N_1} \sum_{j=1}^{N_1} \|\mathbf{x}_j - G_1(E(\mathbf{x}_j))\| \qquad (2)$$

Methods

- Critic Loss



L_r and L_n may help the model well reconstruct normal images, but may not enough to correctly estimate the lesion information.

An extreme case is that to minimize L_r, G_1 would always output the original input, no matter whether the input contains lesion or not.

$$L_r = \frac{1}{N}\sum_{i=1}^{N} \|\mathbf{x}_i - G_1(E(\mathbf{x}_i)) - G_2(E(\mathbf{x}_i))\| \qquad (1)$$

To well separate lesion in lesioned images, the framework uses a discriminator to judge whether the decomposed normal versions are real normal images or not.

$$L_c = \frac{1}{N}\sum_{i=1}^{N} D(G_1(E(\mathbf{x}_i))) - \frac{1}{N_1}\sum_{j=1}^{N_1} D(\mathbf{x}_j) + \lambda \cdot GP \quad (3)$$

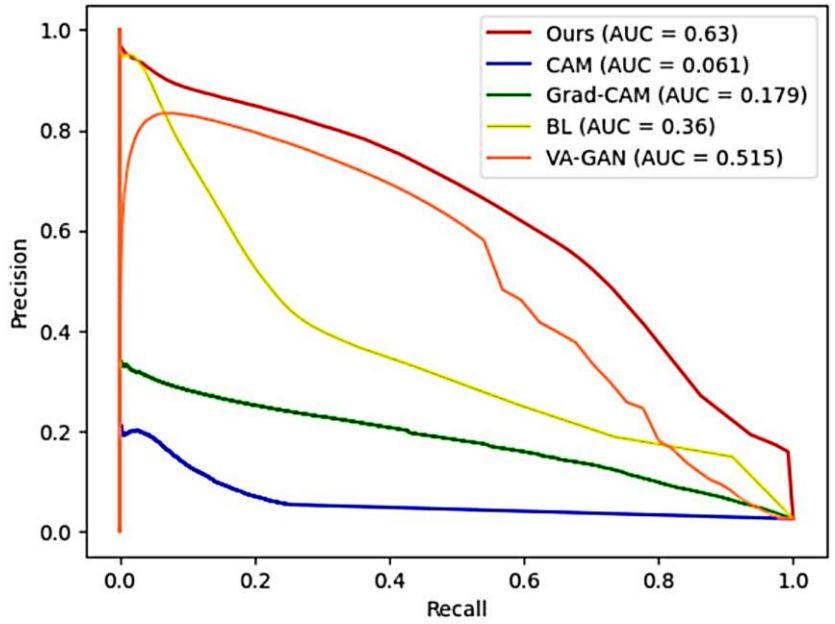
$$L_a = -\frac{1}{N}\sum_{i=1}^{N} D(G_1(E(\mathbf{x}_i))) \qquad (4)$$

Experiments

- Dataset

  Training set: 2007 normal lung CT slices and 870 lesioned slices randomly sampled from COVID-cell
  Test set 1: 128 lesioned slices randomly sampled from COVID-cell
  Test set 2: 493 lesioned slices from COVID-19 Image Data Collection

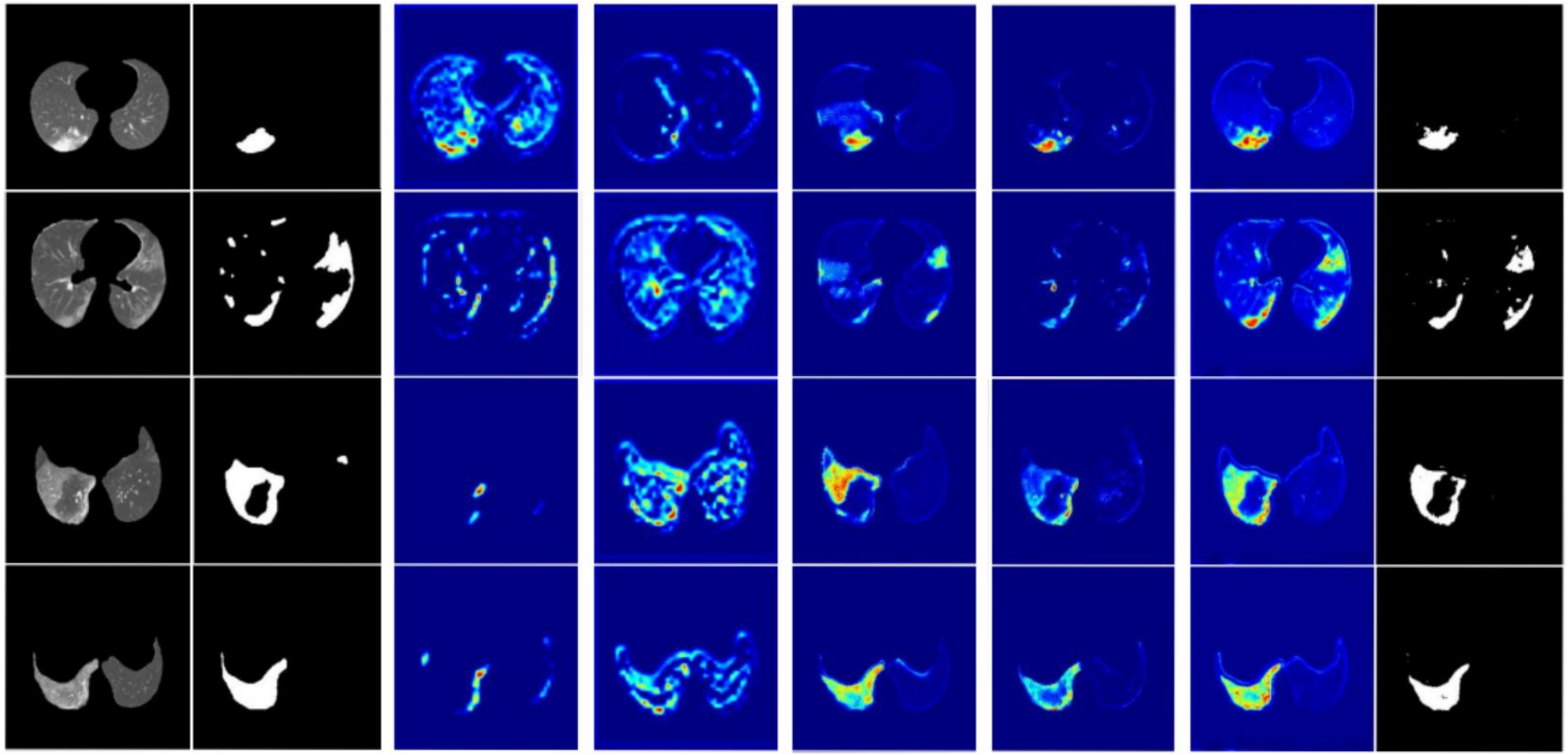
- Quantitative Evaluation



PR curve for each method

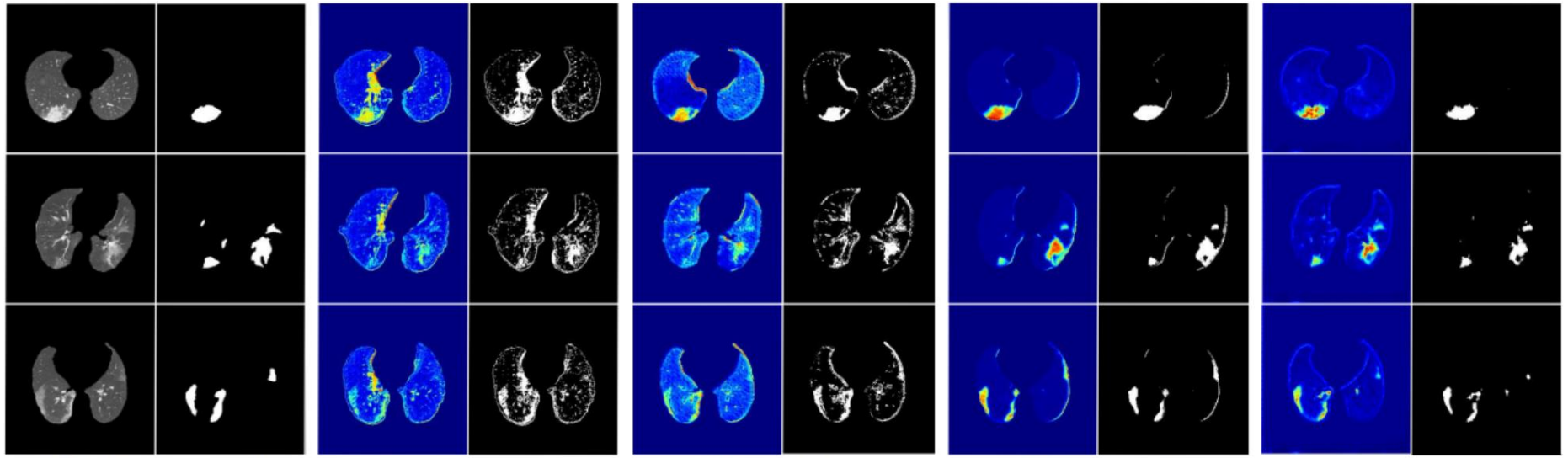
Experiments

- Qualitative Evaluation



Ablation Study

- Qualitative Evaluation



- Quantitative Evaluation