北京大學

Cell Edge Detection

PKU

Xinran Dong
2301210603

Content

- Introduction
- Pre-processing
- Architecture
- Training
- Post-processing
- Result
- Conclusion

Introduction

- **U**ltra**-**high **R**esolution Scanning Electron Microscope (SEM) **I**mages **S**egmentation **C**hallenge is a competition hosted by BAAI and PKU, which provides many high-resolution SEM mouse retinal neurons with per-pixel level annotation for boundary segmentation.

- There are two tracks in the competition: the **Simple** one and the **Complex** one:

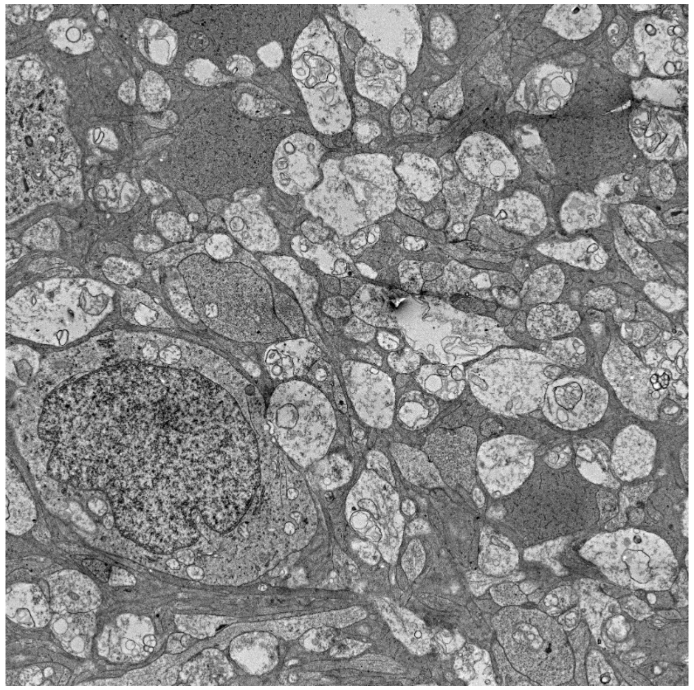
  - The Simple track has fewer cell counts, smaller image size, lower resolution and fewer pixels on the cell membrane.

  - The Complex track has more cell counts, larger image size, higher resolution and more pixels on the cell membrane.

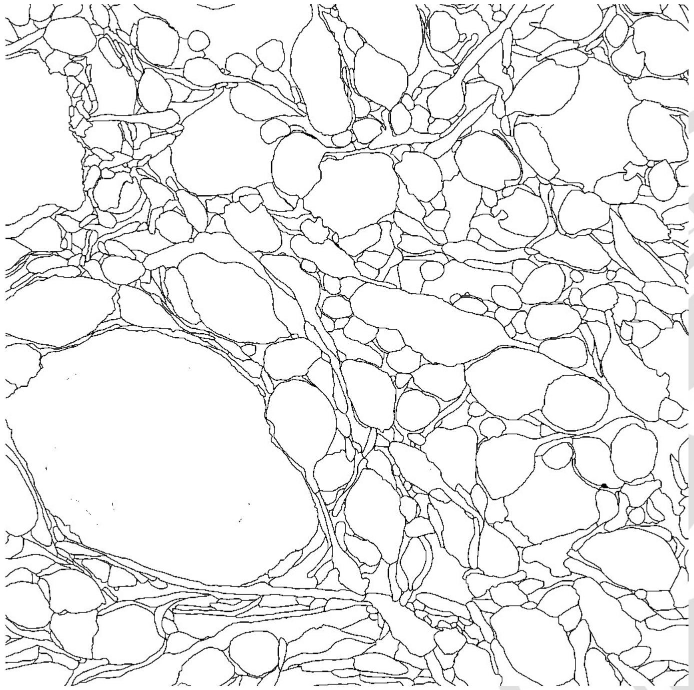
Introduction

- Due to the limitation of computing resources, we choose the Simple track, which is actually more difficult because of the extremely small amount of training data.

- There are **30** training images, **9** validation images (10 originally but containing a mismatched image and label pair) and **30** test images, the resolution of each image and label is **1024×1024**.



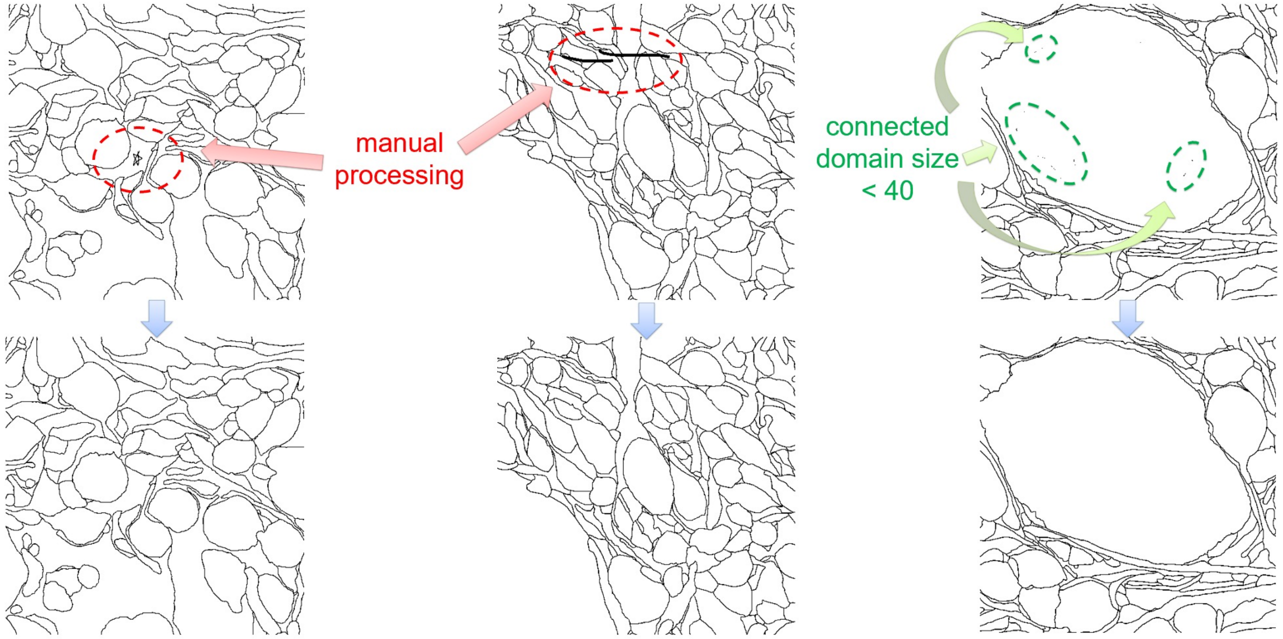
image



label

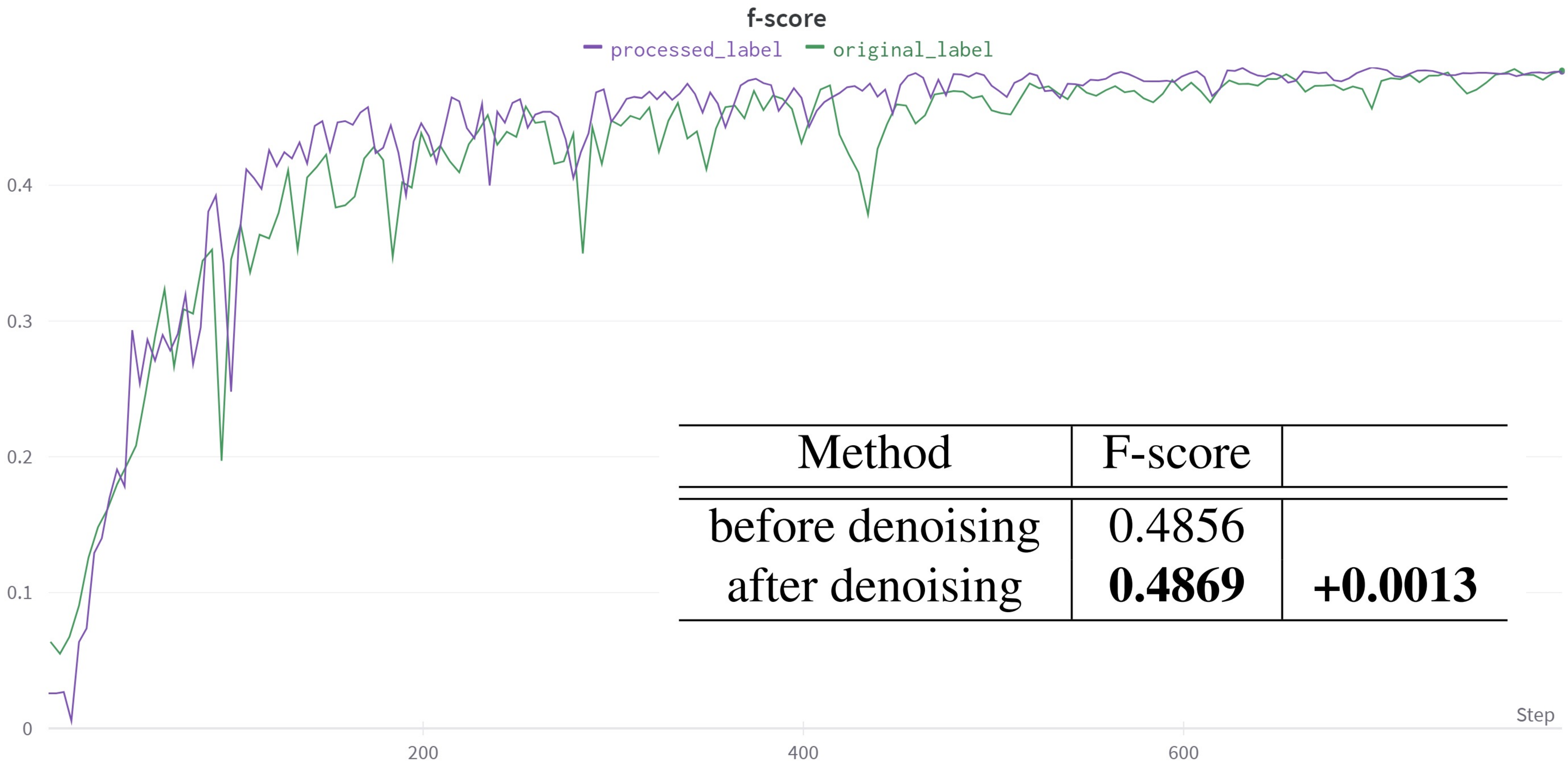
Pre-processing: denoising

- Like other datasets, the raw data in the U-RISC dataset contains a lot of noise, including large areas of mislabeling and small noise points.
- Since data quality determines the upper limit of model performance, we do the denoising before training.
  - For errors with large area and difficult to be processed by morphological method, we use manual processing to remove them.
  - For the noise points with small local area, we directly remove the area whose connected domain size < **40**.



Pre-processing: denoising

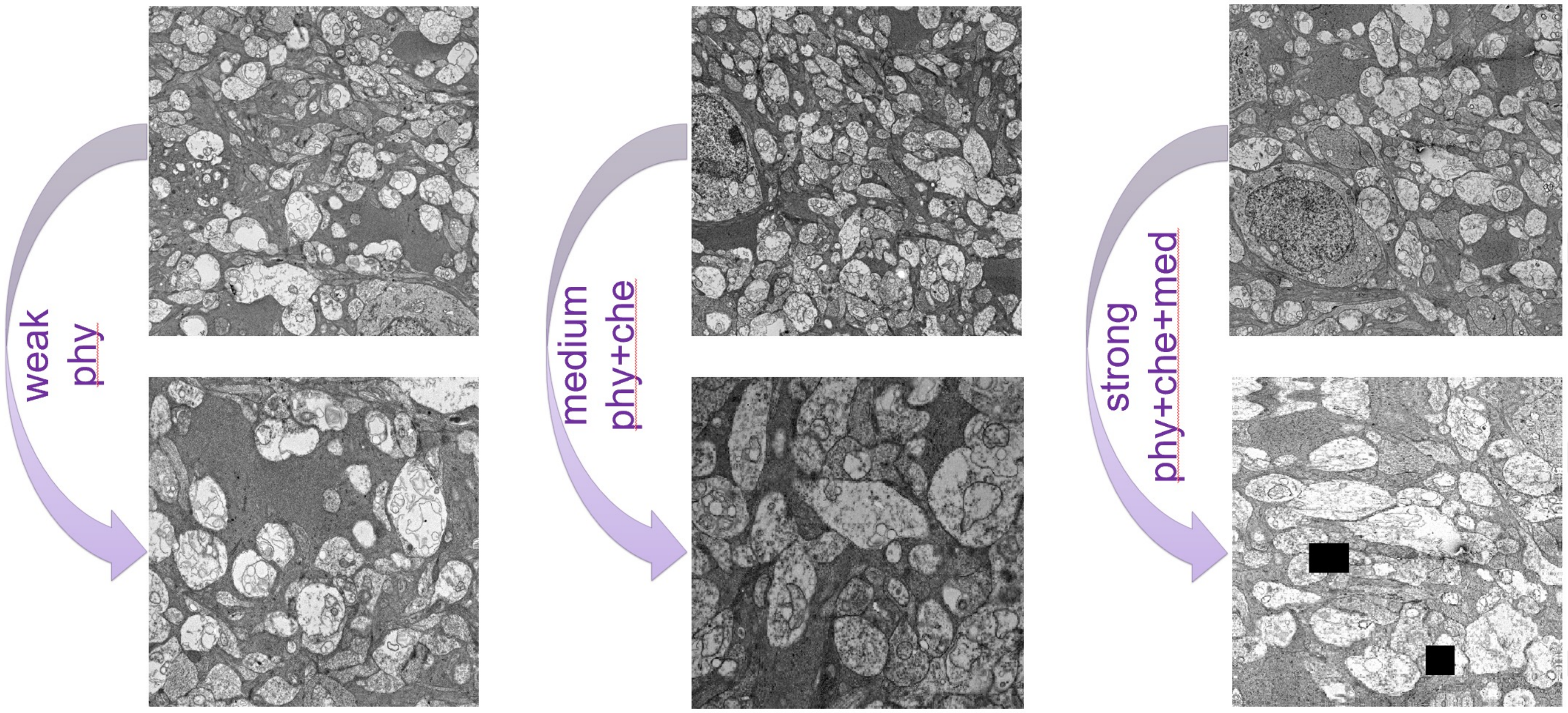
- Comparison of the results before/after denoising



Method	F-score	
before denoising	0.4856	
after denoising	**0.4869**	**+0.0013**

Pre-processing: augmentation

- Due to the lack of training data in simple track, we adopt a series of physical and chemical transformations for strong data augmentation in order to avoid overfitting.
  - Physical: RandomResizedCrop, HorizontalFlip, VerticalFlip, RandomRotate90
  - Chemical: RandomBrightnessContrast, MultiplicativeNoise
  - Medical: ElasticTransform, GridDistortion, OpticalDistortion, CoarseDropout



The augmentation transforms are from https://github.com/albumentations-team/albumentations

Pre-processing: augmentation

- Comparison of different data augmentations

Method	F-score	
WAUG	0.4703	
MAUG	0.4724	**+0.0021**
SAUG	**0.4785**	**+0.0061**

Architecture

- We use a pretrained ResNet-50 as backbone for all network structures.
- We also use model ensemble to improve robustness which is very important in the medical field.

Training

- Loss

  Focal Loss (BCELoss with weight): $\alpha = 0.7,\ \gamma = 2$

  F–score Loss (1 – F–score, derivable version, without binarization)

  Near Edge Loss (proposed in UNet)

  TotalLoss = 0.5×FocalLoss + 0.5×F–scoreLoss + 0.4×NearEdgeLoss

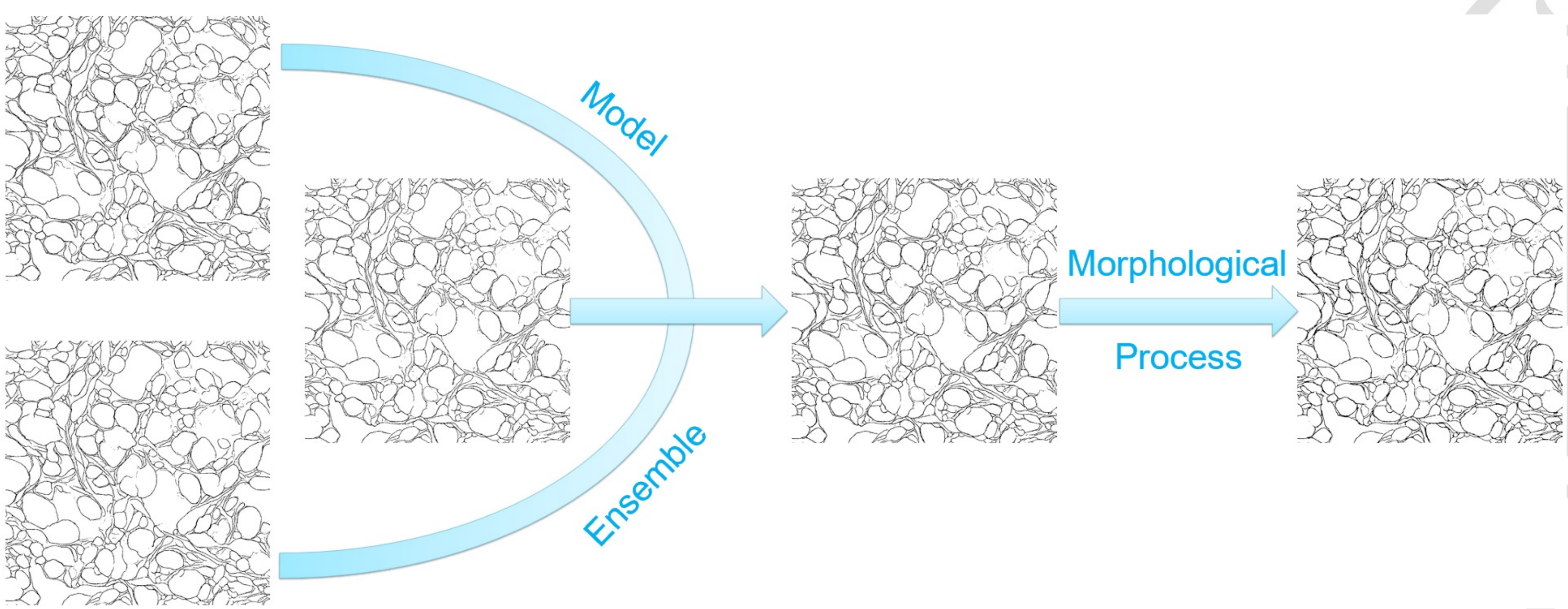
Training

- hyperparameters

Hyperparameters	Simple	Complex
learning rate	backbone: 0.001, others: 0.01	
optimizer	Adam	
batch size	8	4
crop size	960	1280
training epoch	500	100
γ of Focal Loss	2	
α of Focal Loss	0.7	
λ	0.8	
pre-binarization thres	122	
post-binarization thres	127.5	

Post-processing: morphological processing

- During the training process, we found that the **precision** of model's prediction was always lower than the **recall**.
- According to the calculation method of F-score, a better balance between the two could improve the model's performance.
  - $F - score = \dfrac{2 \times precision \times recall}{precision + recall}$
- Therefore, after the model ensemble, morphological processing methods such as **opening** and **closing** operation are adopted to improve precision at the expense of slightly sacrificing recall, thus achieving a higher Fscore.



Result

- We have also tried a variety of advanced network architectures for image segmentation and edge detection, including DDS, CASENet and DFF. The results of Simple-Track validation dataset are as follows:

Architecture	Backbone	F-score
DDS	ResNet-50	0.49058
DDS	ResNet-152	0.48914
CASENet	ResNet-50	0.49196
CASENet	ResNet-152	0.48699
DFF	ResNet-50	**0.51182**

Result

In Simple–Track, we choose DFF as our network and pretrained ResNet–50 as backbone. After 500 epochs of training, our method has achieved an F–score of 0.55667 on the test dataset. Our method performs better than the fourth place in the original U–RISC Competition.

Ranking	Research Institution	F-score
	Human 1st	0.96915 ± 0.014
	Human 2nd	0.99334 ± 0.008
1	UCAS	0.56932 ± 0.053
2	NJU	0.56213 ± 0.055
3	HDU	0.56136 ± 0.049
4	UCAS	0.55170 ± 0.046
5	THU	0.55103 ± 0.047
6	SCU	0.54847 ± 0.053
-	Ours	0.55667

Conclusion

- Contribution
  - We are the first to use a combination of neural networks and morphological processing to do EM cytomembrane segmentation.
  - We use strong data augmentation to avoid overfitting in the small dataset, model ensemble to improve robustness which is very important in the medical field and morphological processing to improve the precision at the expense of slightly sacrificing the recall.
  - We outperform previous state-of-the-art methods by a significant margin on U-RISC dataset (more than 0.02 F-score on val set).

Conclusion

- Improvement
  - Add more morphological processing in the preprocessing stage to help the neural network extract image features better.
  - Adopt two-stage or cascade network structure to better classify details.

Thanks for listening!