Master's Thesis Presentation for the 140th Dissertation Evaluation

A study on Effective Tiny Object Semantic Segmentation using Skip Connection and Cross-Channel Features

스킵 커넥션과 교차 채널 특징을 활용한 작은 객체 시맨틱 세그멘테이션에 관한 연구

Taekyung Song

songteagyong@gmail.com

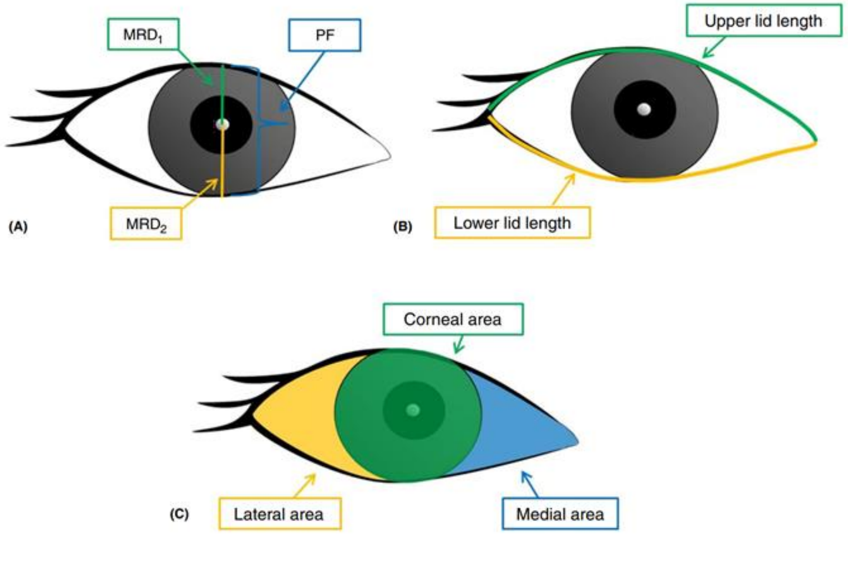
18th Oct, 2023

Contents

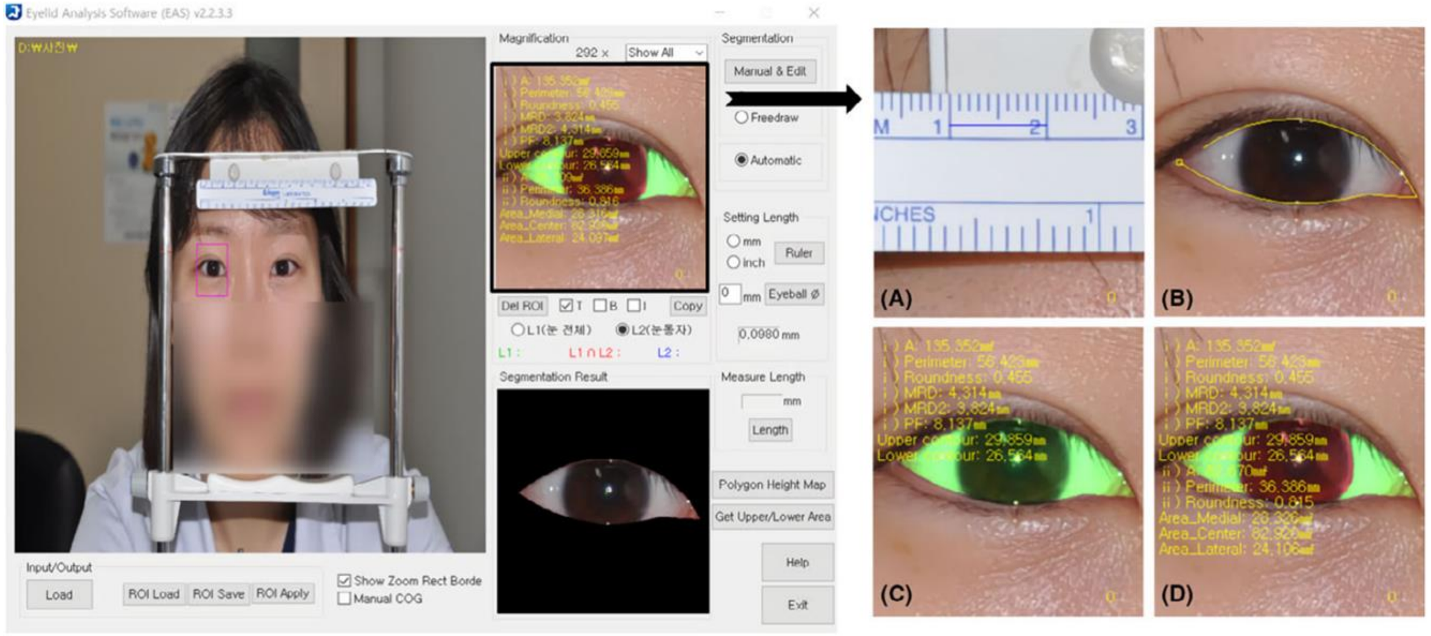
Introduction

Thyroid Ophthalmopathy Patient’s Eyelid Length

- In thyroid ophthalmopathy, eyelid length is measured to monitor disease progression.
- There is no automatic measurement system, so measurements must be taken manually.
- Manual measurements can be inaccurate, as the length can vary depending on the person measuring it.



Lenth measurement element

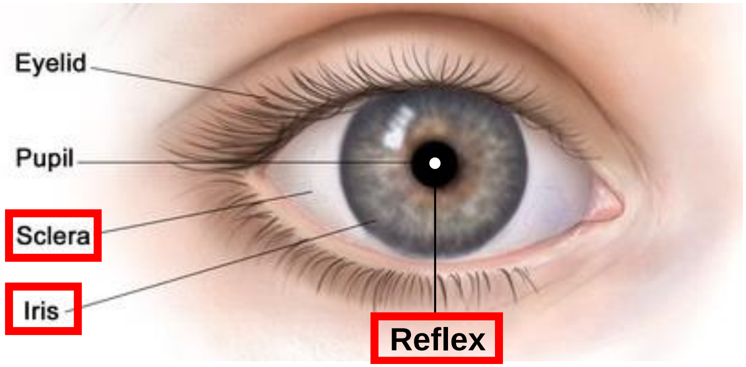


Eyelid analysis software

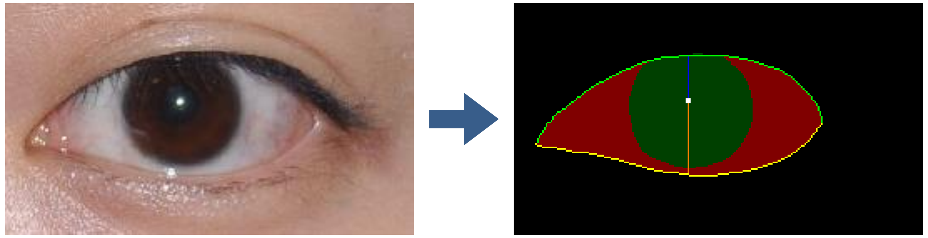
Introduction

Eyelid Length Measurement

- Reflex is the central point of length measurement.
- Because the size of the reflex is very tiny, segmenting a reflex is quite a challenging problem.



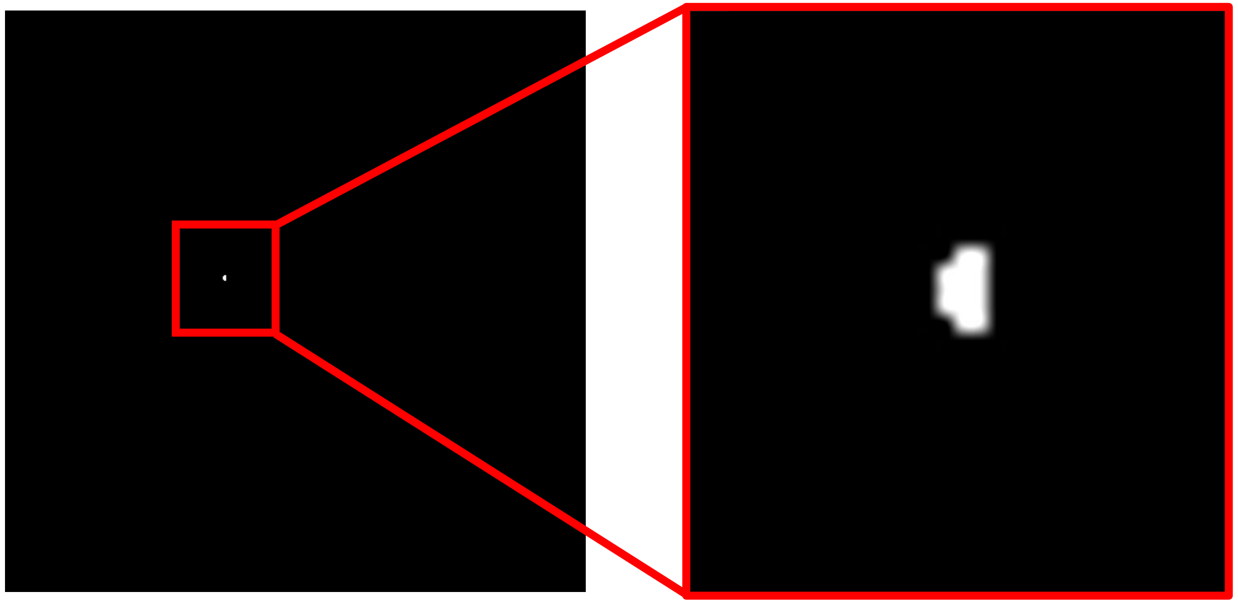
Target of segmentation



Part	MRD1	MRD2	Upper lid	Lower lid
Length	3.18	7.41	40.62	37.18

(mm)

The result of eyelid length measurement



(512, 512, 1) mask

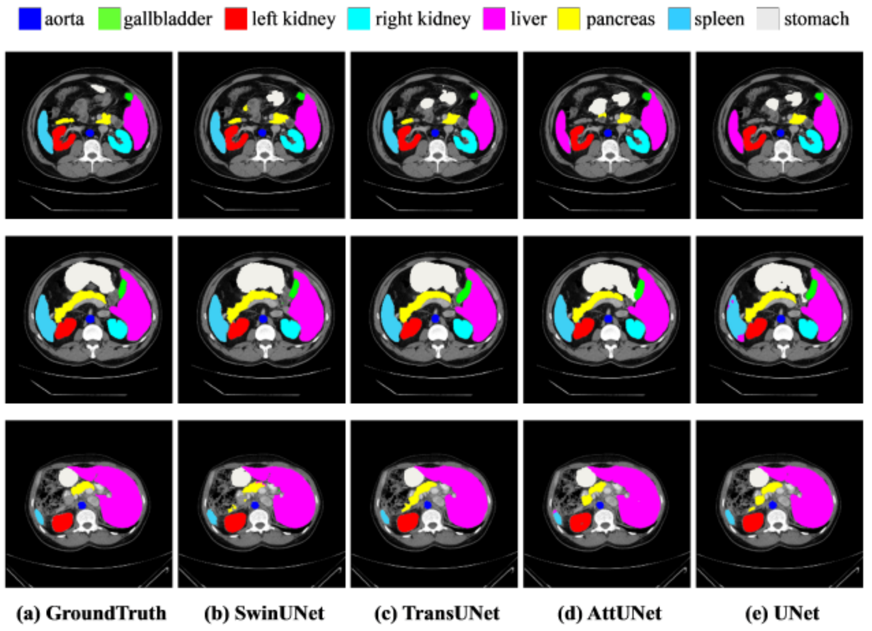
x20 zoom

Reflex mask

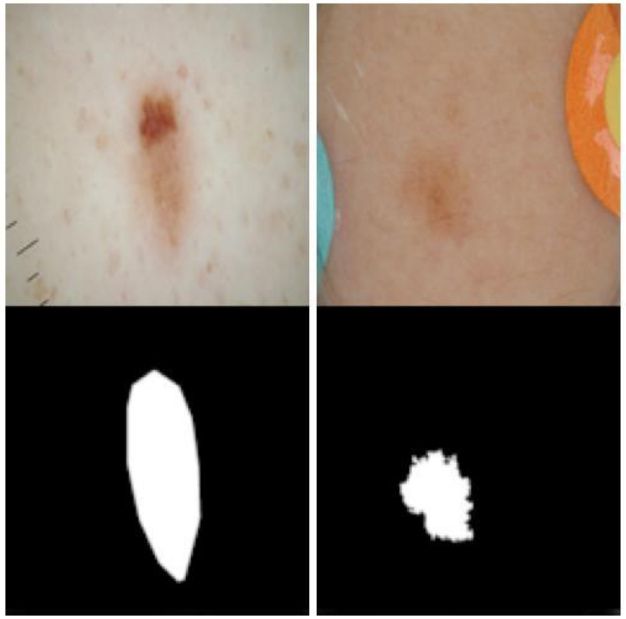
Related Work

Medical Image Segmentation

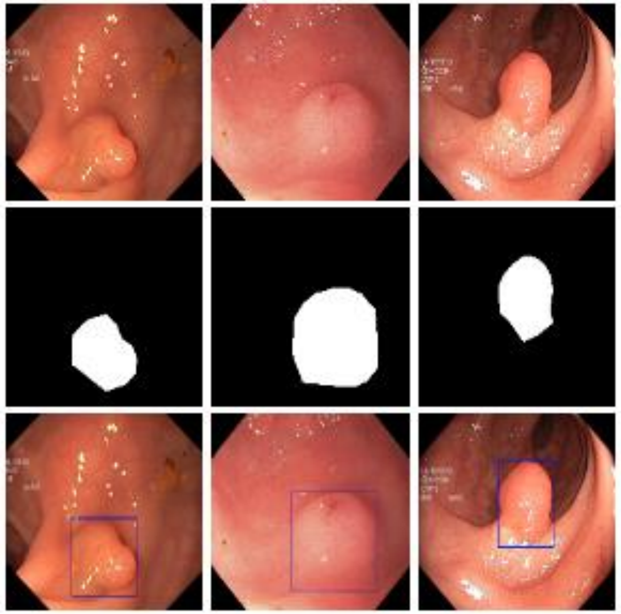
- Segmentation in the medical image is a requirement for medical application such as brain segmentation, Computed Tomography (CT) image segmentation, lesion segmentation, and polyp segmentation.
- It is very common in medical image that the anatomy of interest only occupies a very small portion.
- Even though the anatomy of interest are small, it is very important to detect them when they are small.



Synapse multi-organ CT



Skin cancer data

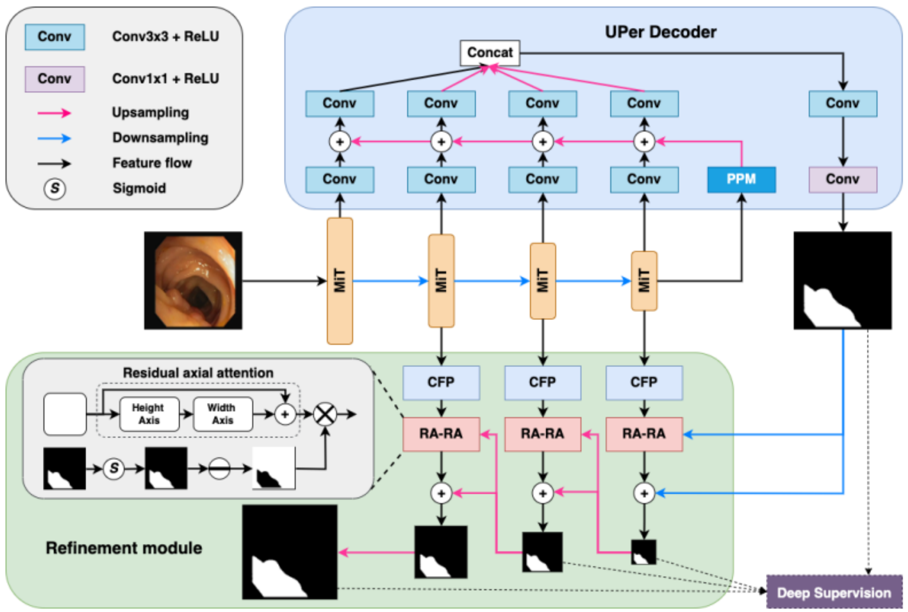


Kvasir-SEG

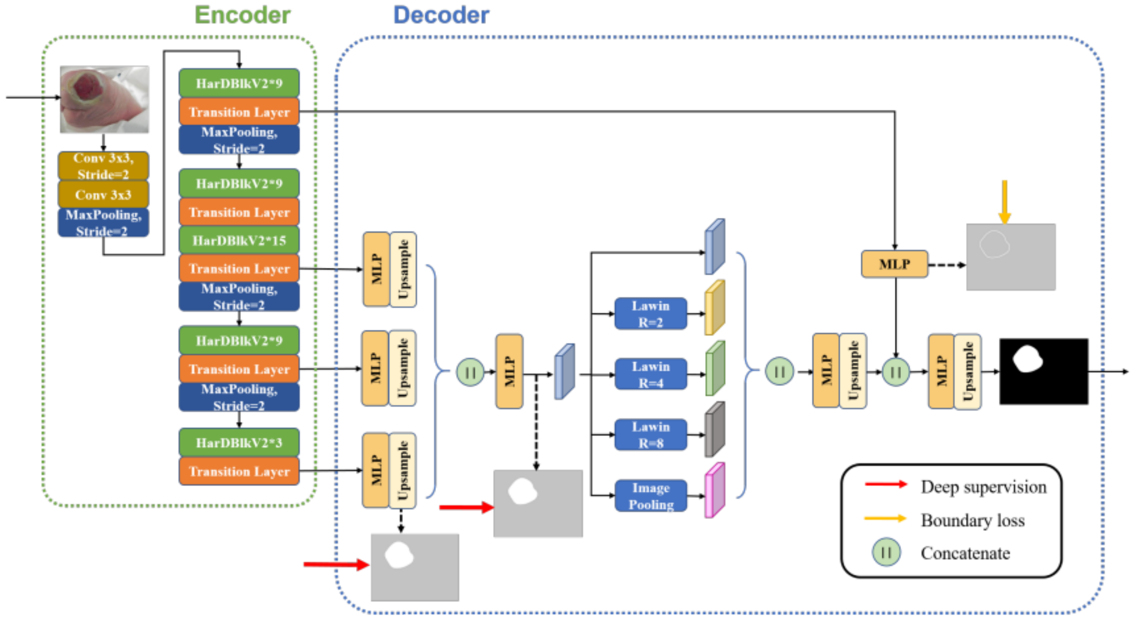
Related Work

Medical Image Segmentation

- Recently, many researchers have applied convolutional neural networks to segmenting medical image and achieved good results.
- ColonFormer uses the pyramid pooling and the refinement module to refine the boundary of the object.
- HarDNet-DFUS uses HarDBlock in the encoder and replaces the decoder with Dense Aggregation.



ColonFormer

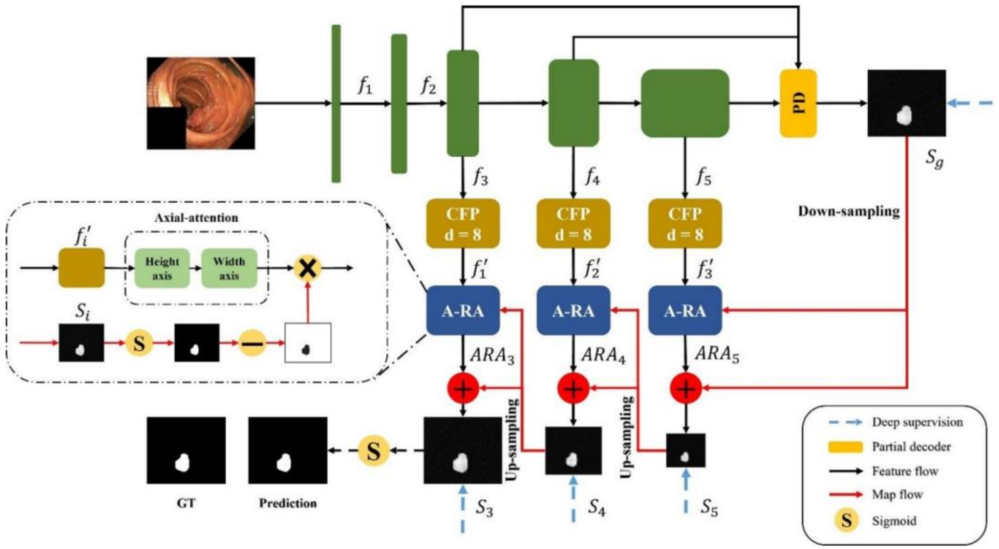


HarDNet-DFUS

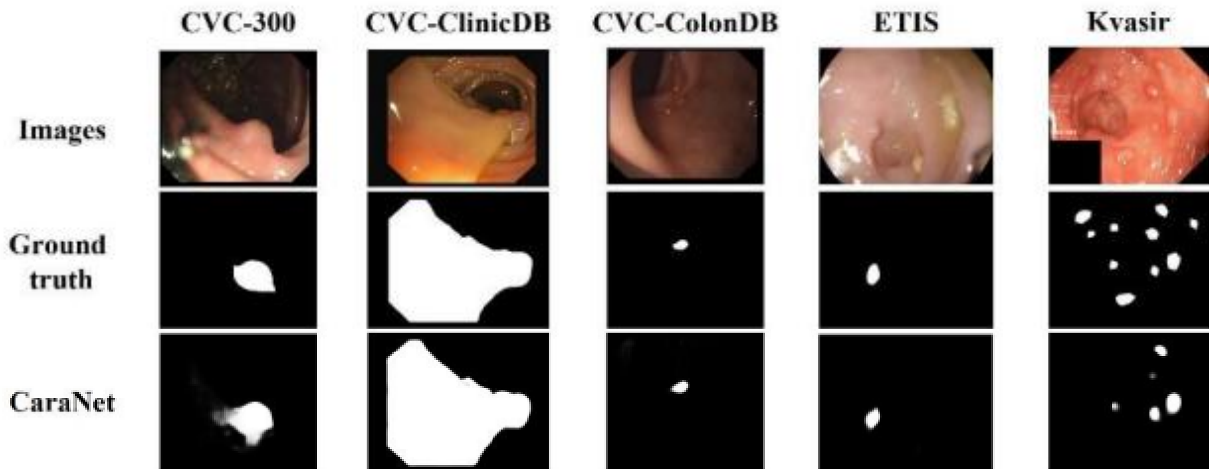
Related Work

Segmentation of Small Medical Objects

- CaraNet proposed segmentation network for small medical objects.
- CaraNet applies axial reserve attention and channel-wise feature pyramid module to dig feature information of small medical object.
- They mainly focus on the small areas whose size ratios are smaller than 5%.



CaraNet

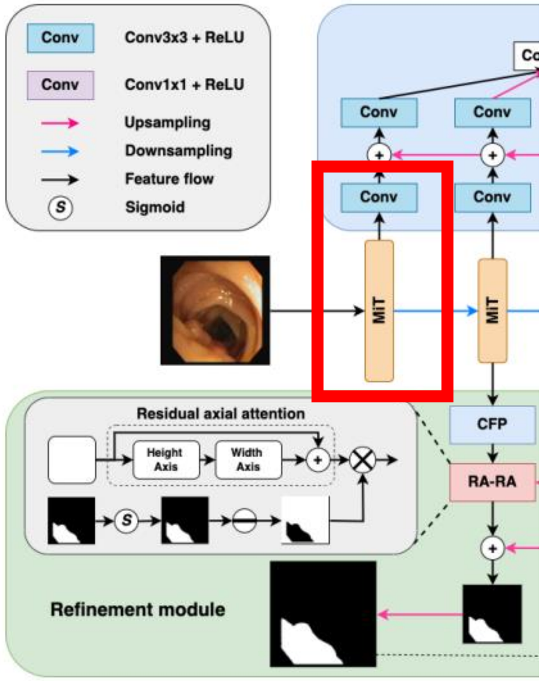


CaraNet Polyp segmentation results

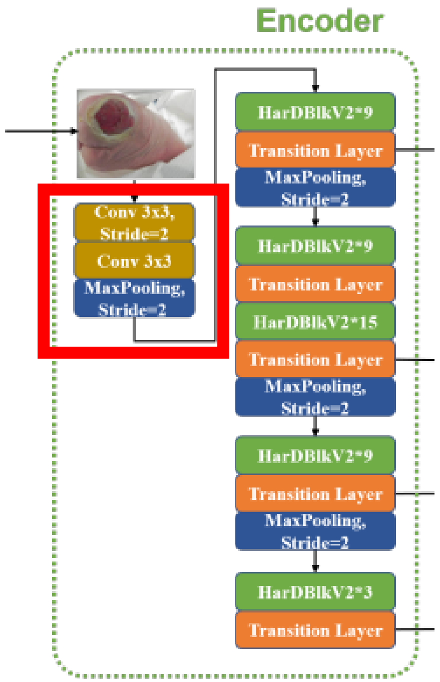
Related Work

Common Drawback of Conventional Methods

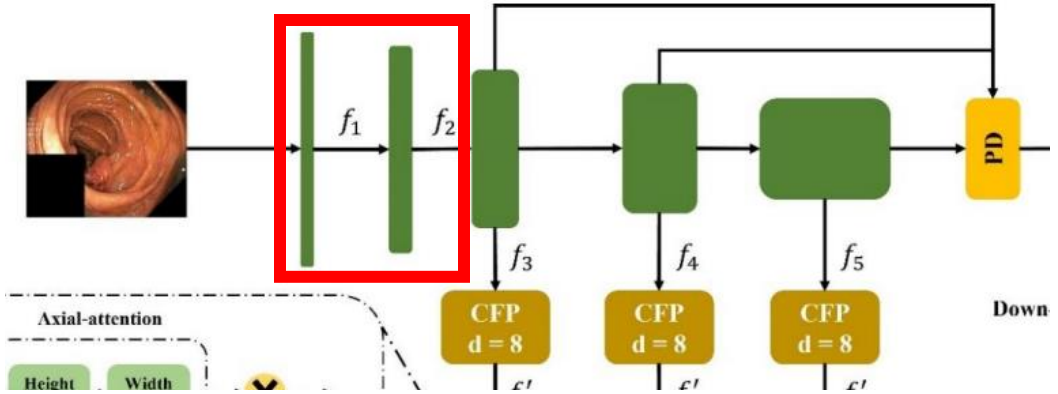
- Due to the nature of medical imaging, it is important to make good use of shallow features.
- A way to take advantage of this is to use skip connections.
- The previous neural networks do not use the features of first down-sampling.



ColonFormer’s encoder

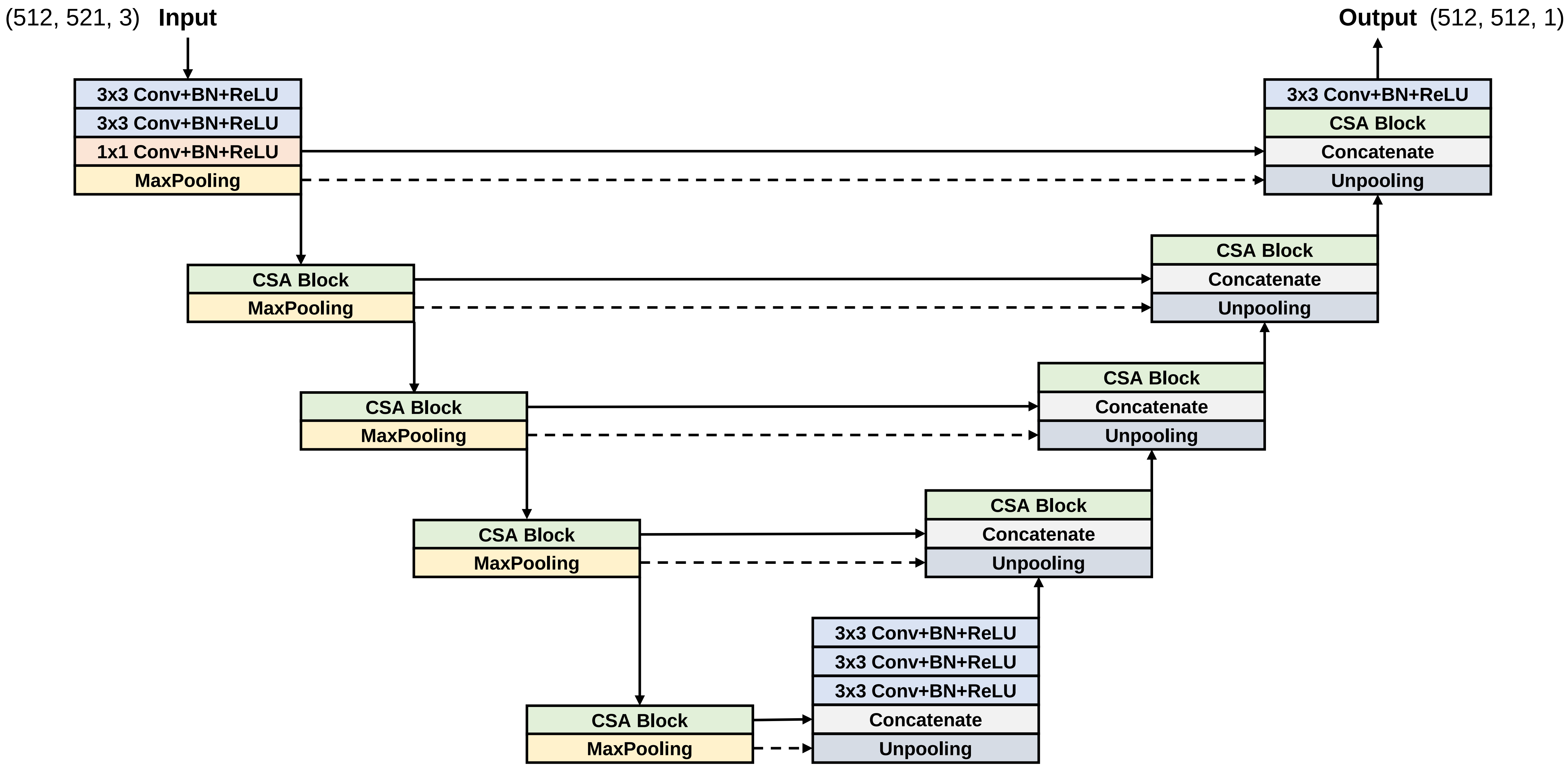


HarDNet-DFUS’s encoder



CaraNet’s encoder

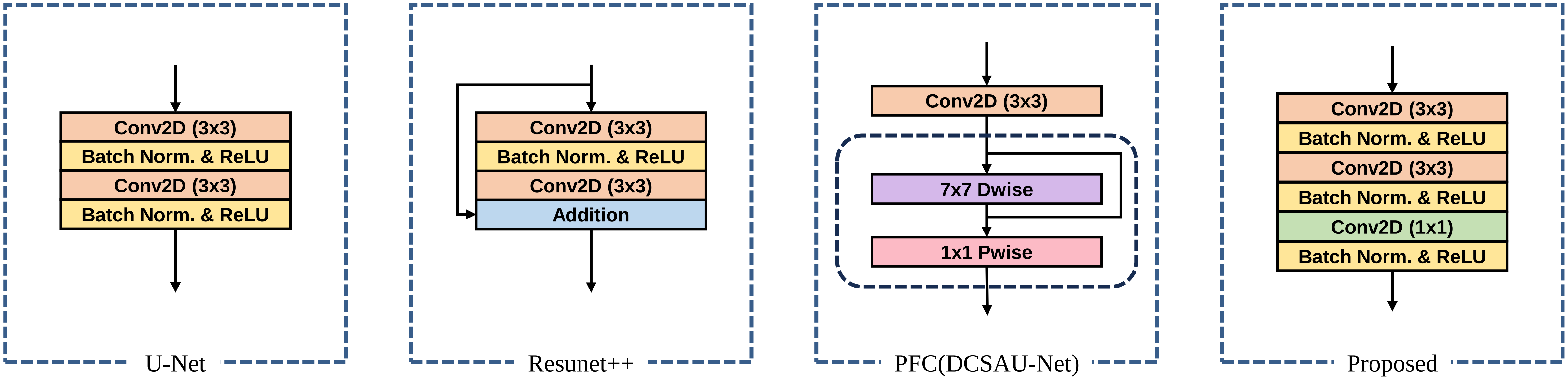
Proposed Method



Proposed Method

Importance of first down-sampling in medical images

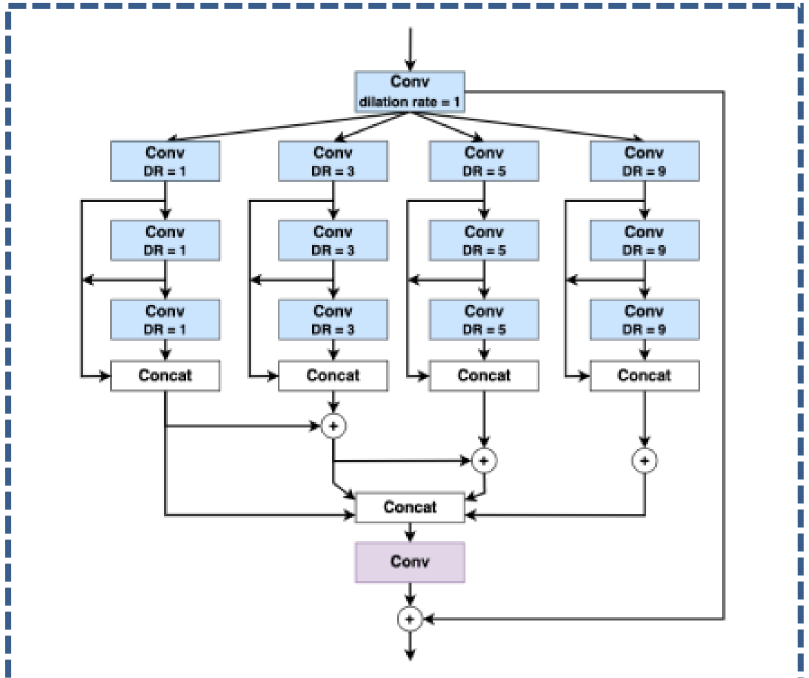
- Medical images have the characteristic that the boundary with the surroundings is not clear.
- In most medical images, the first down-sampling block is used to extract low-level semantic information.
- To extract low-level information, we add a kernel-size 1x1 convolution to the first down-sampling block.



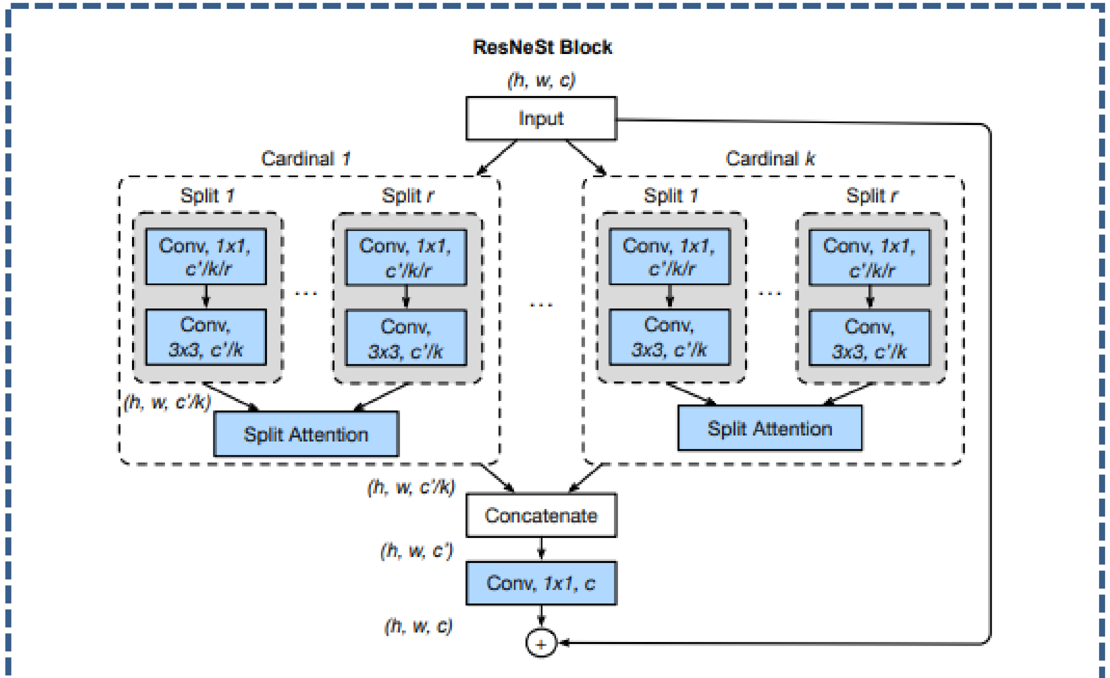
Proposed Method

Cross-Channel Feature

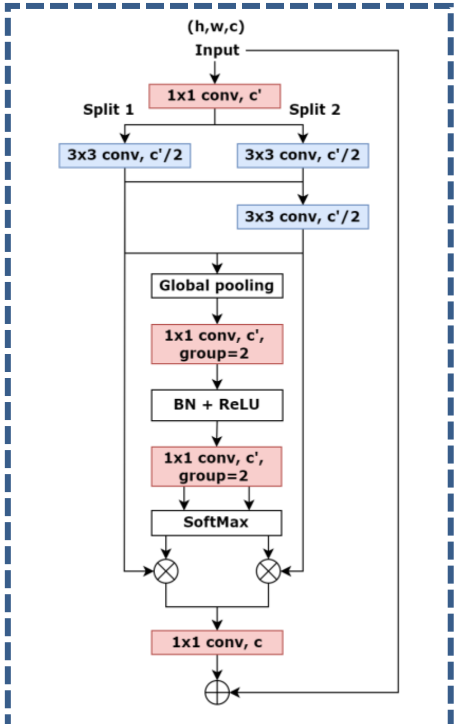
- The representation of multi-scale features is beneficial for the model to perceive data features.
- Recently various models learning the representation via cross-channel features have been proposed.
- The proposed network was designed using Compact Split Attention (CSA) block exploiting cross-channel features to focus more on small objects.



CFP block



ResNeSt block

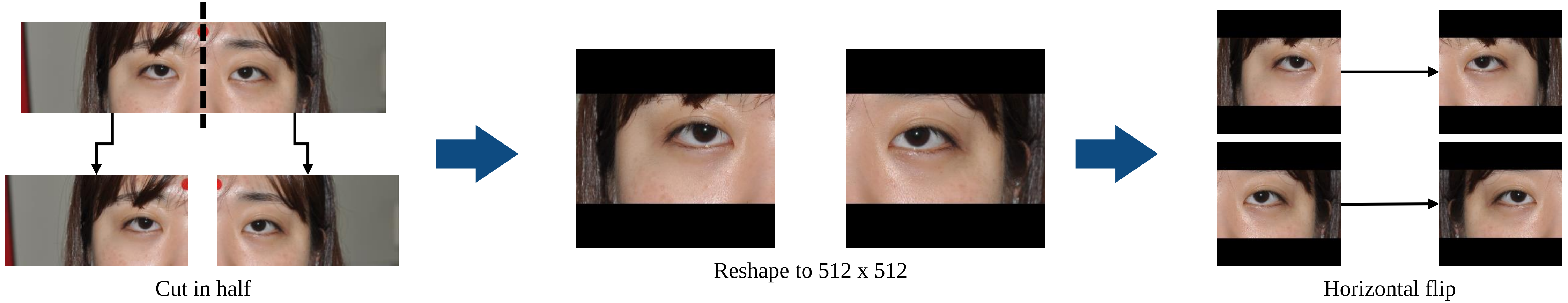


CSA block

Experimental Settings

Dataset

- 1004 images acquired from 251 patients are used to train the network.
- The image is cut in half and used for learning without considering the left and right sides.
- For augmentation, horizontal flip was applied.



Data preprocessing process

Cross Validation

- Number of Iterations: 30
- Data Split Ratio (Train: Validation: Test) = 0.6 : 0.2: 0.2

Experimental Settings

Hyperparameter Settings

- Batch size: 8
- Epochs : 50
- Loss function: Focal loss
- Optimizer : Adam
- Learning Rate: 1e-4

Evaluation Measures

- *Dice Coefficient*

$$Dice = \frac{2 * |A \cap B|}{|A + B|} = \frac{2 * TP}{(TP + FP) + (TP + FN)}$$

- *IoU*

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} = \frac{A \cap B}{A \cup B}$$

Overall Experimental Results

Experimental Results

- The proposed model show better performance than the other comparison models.

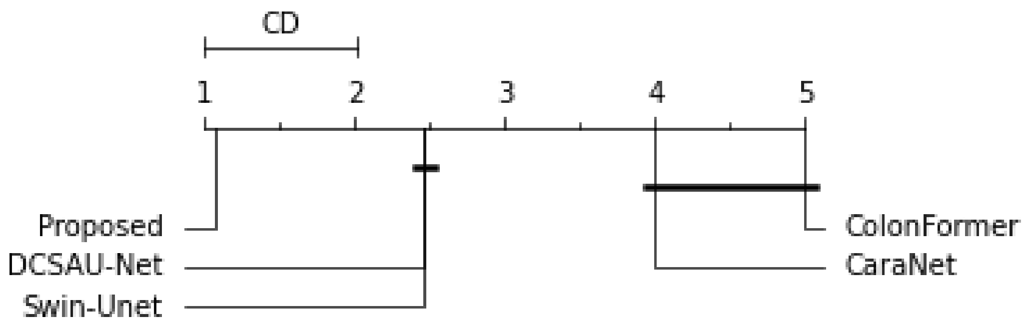
Model	Dice	IoU	Rank
CaraNet	0.654 ± 0.010	0.497 ± 0.012	4.067
DCSAU-Net	0.825 ± 0.007	0.715 ± 0.023	2.267
Swin-Unet	0.818 ± 0.007	0.714 ± 0.009	2.733
ColonFormer	0.614 ± 0.015	0.467 ± 0.016	4.933
Proposed	**0.880 ± 0.021**	**0.800 ± 0.032**	1.033

Overall Experimental Results

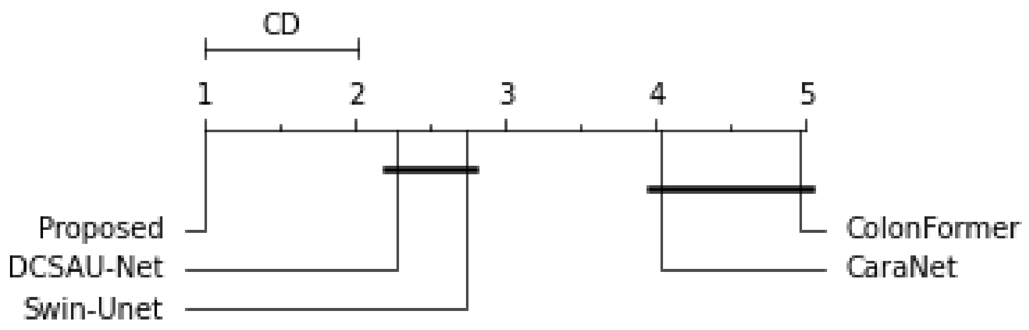
Friedman Test

Evaluation Measure	Friedman statistic	Critical Values (α = 0.05)
IoU	111.68	2.276
Dice	114.533	2.276

Bonferroni-Dunn Test



IoU



Dice

In-Depth Analysis

Experimental Results

- Unlike other segmentation tasks, in tiny object segmentation, even a single pixel can have a huge impact.
- CaraNet and ColonFormer, which do not utilize the first down-sampling information, performed poorly.
- When up-sampling, using the information from the 1x1 kernel resulted in more accurate segmentation.

Ground Truth | proposed | DCSAU-Net | Swin-Unet | CaraNet | ColonFormer

Prediction mask results for each model (x20 zoom)

Conclusion and Contribution

Conclusion

- A network is proposed, which can segment tiny object in medical images.
- The proposed model outperforms the existing models.
- When up-sampling the last layer, it worked to use the first down-sampling information via skip-connection.
- Using cross-channel features was effective in learning information from different receptive fields.

Contribution

- In Domestic Competition
  - B-cell epitope (linear peptide) prediction competition, AI Graduate School Challenge, 2022, 3rd Prize
  - Medical Technology for Patients with Thyroid-associated Ophthalmopathy, AI Technology-based Employment and Start-up Workshop, 2023, 3rd Prize
- In International Conferences
  - Effective Eyelid Length Measurement Using Semantic Segmentation, ***ICEIC***, 2023
- In Domestic Conferences
  - Comparison of Deep Learning based Segmentation Models for Measuring Eyelid Contour of TAO Patients, ***IEIE***, 2022

Thank you

Reference

[1] Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., & Wang, M. (2022, October). Swin-unet: Unet-like pure transformer for medical image segmentation. In *European conference on computer vision* (pp. 205-218). Cham: Springer Nature Switzerland.

[2] Lou, A., Guan, S., & Loew, M. (2023). Caranet: Context axial reverse attention network for segmentation of small medical objects. *Journal of Medical Imaging*, *10*(1), 014005-014005.

[3] Xu, Q., Ma, Z., Na, H. E., & Duan, W. (2023). DCSAU-Net: A deeper and more compact split-attention U-Net for medical image segmentation. *Computers in Biology and Medicine*, *154*, 106626.

[4] Duc, N. T., Oanh, N. T., Thuy, N. T., Triet, T. M., & Dinh, V. S. (2022). Colonformer: An efficient transformer based method for colon polyp segmentation. *IEEE Access*, *10*, 80575-80586.

[5] Dumitru, R. G., Peteleaza, D., & Craciun, C. (2023). Using DUCK-Net for polyp image segmentation. *Scientific Reports*, *13*(1), 9803.

[6] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18* (pp. 234-241). Springer International Publishing.

[7] Jha, D., Smedsrud, P. H., Riegler, M. A., Johansen, D., De Lange, T., Halvorsen, P., & Johansen, H. D. (2019, December). Resunet++: An advanced architecture for medical image segmentation. In *2019 IEEE international symposium on multimedia (ISM)* (pp. 225-2255). IEEE.

[8] Zhang, H., Wu, C., Zhang, Z., Zhu, Y., Lin, H., Zhang, Z., ... & Smola, A. (2022). Resnest: Split-attention networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 2736-2746).

[9] Lou, A., & Loew, M. (2021, September). Cfpnet: Channel-wise feature pyramid for real-time semantic segmentation. In *2021 IEEE International Conference on Image Processing (ICIP)* (pp. 1894-1898). IEEE.

[10] Liao, T. Y., Yang, C. H., Lo, Y. W., Lai, K. Y., Shen, P. H., & Lin, Y. L. (2022). HarDNet-DFUS: Enhancing Backbone and Decoder of HarDNet-MSEG for Diabetic Foot Ulcer Image Segmentation. In *Diabetic Foot Ulcers Grand Challenge* (pp. 21-30). Cham: Springer International Publishing.

Appendix

Additional Experiments

- Comparison with Split Attention (SA) block

Method	Dice	IoU
Proposed + SA	0.802 ± 0.012	0.695 ± 0.014

- Kernel size adjustment in first down-sampling

Method	Dice	IoU
Kernel 3x3 3x3	0.838 ± 0.013	0.74 ± 0.017
Kernel 3x3 1x1	0.809 ± 0.013	0.705 ± 0.017
Kernel 7x7 3x3 1x1	0.799 ± 0.014	0.692 ± 0.015