A NOVEL CONVOLUTIONAL NEURAL NETWORK BASED ON ADAPTIVE MULTI-SCALE AGGREGATION AND BOUNDARY-AWARE FOR LATERAL VENTRICLE SEGMENTATION ON MR IMAGES

Fei Ye[1,⋆], *Zhiqiang Wang*[2,3,⋆], *Sheng Zhu*[2], *Xuanya Li*[4,†], *and Kai Hu*[1,3,†],

[1]Key Laboratory of Intelligent Computing and Information Processing of Ministry of Education, Xiangtan University, Xiangtan 411105, China
[2]Department of Radiology, Affiliated Hospital of Xiangnan University, Chenzhou 423000, China
[3]Key Laboratory of Medical Imaging and Artificial Intelligence of Hunan Province, Xiangnan University, Chenzhou 423000, China
[4]Baidu Inc., Beijing 100085, China
[⋆]Co-first authors, contributed equally to this work
[†]Corresponding authors, email: lixuanya@baidu.com, kaihu@xtu.edu.cn

ABSTRACT

In this paper, we propose a novel convolutional neural network based on adaptive multi-scale feature aggregation and boundary-aware for lateral ventricle segmentation (MB-Net), which mainly includes three parts, i.e., an adaptive multi-scale feature aggregation module (AMSFM), an embedded boundary refinement module (EBRM), and a local feature extraction module (LFM). Specifically, the AMSFM is used to extract multi-scale features through the different receptive fields to effectively solve the problem of distinct target regions on magnetic resonance (MR) images. The EBRM is intended to extract boundary information to effectively solve blurred boundary problems. The LFM can make the extraction of local information based on spatial and channel attention mechanisms to solve the problem of irregular shapes. Finally, extensive experiments are conducted from different perspectives to evaluate the performance of the proposed MB-Net. Furthermore, we also verify the robustness of the model on other public datasets, i.e., COVID-SemiSeg and CHASE_DB1. The results show that our MB-Net can achieve competitive results when compared with state-of-the-art methods.

***Index Terms*—** Lateral ventricle segmentation, Adaptive multi-scale feature aggregation, Boundary-aware, Convolutional neural network

1. INTRODUCTION

The lateral ventricle, located deep in the brain, is an essential part of the brain. A number of studies have shown that the measurement of the lateral ventricle region on magnetic resonance (MR) images plays a significant role in the diagnosis of neurodegenerative diseases [1]. Therefore, the clinical physician needs to evaluate the volume of the lateral ventricles in the diagnosis, treatment, and prognosis of diseases. However, rapid manual segmentation by clinicians still presents many challenges due to its subjective and time-consuming nature. Thus, it is necessary to propose an efficient and accurate computer-aided method for automatically segmenting the region of the lateral ventricle on MR images.

In the past decade, the research on lateral ventricle segmentation is mainly realized by traditional methods, such as fast multi-atlas likelihood fusion [2], region growing [3], etc. These methods are susceptible to noise, and the selection of model hyperparameters is random, which leads to poor generalization performance of the model.

In recent years, deep learning has made great breakthroughs on natural or medical image processing including lateral ventricle segmentation. For example, Shao *et al.* [4] proposed a deep convolutional neural network (CNN) to segment the lateral ventricle region. In 2019, Wu *et al.* [5] designed a 3D encoding and 2D decoding convolutional neural network to segment the lateral ventricle. However, the research on lateral ventricle segmentation based on deep learning is still in the preliminary stage, and there is still much room for improvement in segmentation accuracy due to the following reasons. (1) MR images of lateral ventricles are characterized by blurred boundary, low contrast, and differentiated target regions, which makes it difficult to segment the lateral ventricles. (2) Currently, existing deep learning algorithms do not carry out specific design for the difficulties of lateral ventricle segmentation. (3) There are no publicly available datasets for lateral ventricle segmentation.

In this paper, we first construct a polycentric and diverse lateral ventricle MR images segmentation (LV-Seg) dataset. In particular, the LV-Seg dataset fully integrates many difficulties under some complicated clinical manifestations from the perspective of doctors. Therefore, it is more consistent

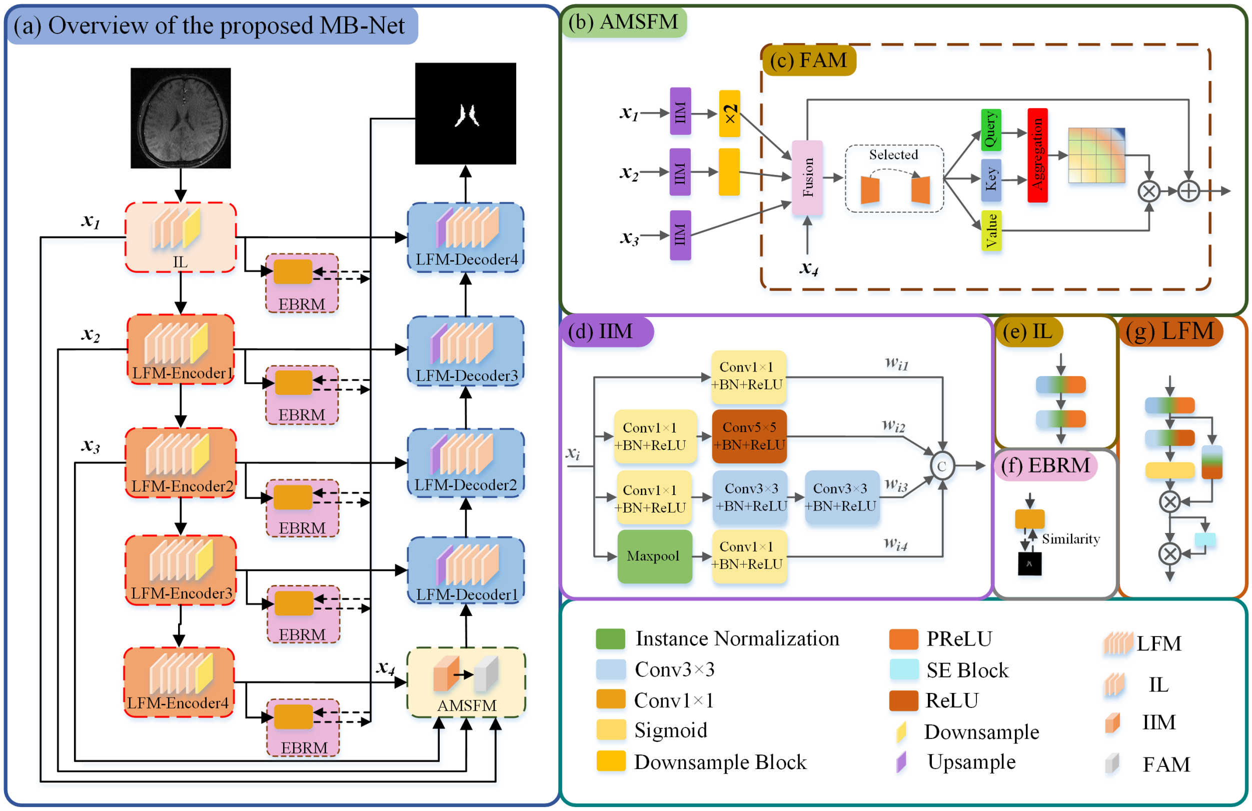

Fig. 1. The network framework of MB-Net. (a) Overview of the proposed MB-Net. (b) Adaptive multi-scale feature aggregation module. (c) Feature aggregation mudule. (d) Improved inception module. (e) Init layer. (e) Embedded boundary refinement module. (g) Local feature extraction module.

with clinical practice and more conducive to assisting doctors in diagnosing diseases. Secondly, we propose an effective and accurate CNN based on adaptive multi-scale feature aggregation and boundary-aware (MB-Net) for lateral ventricle segmentation. Our MB-Net mainly includes three parts, i.e., an adaptive multi-scale feature aggregation module (AMSFM), an embedded boundary refinement module (EBRM), and a local semantic feature extraction module (LFM). Especially, the AMSFM can extract different scale different information based on the fact that different receptive fields can observe various information. Furthermore, the EBRM can guide our model to enhance boundary learning by extracting the boundary feature information in the process of our model learning. Besides, the LFM can make the extraction of local information more amiable based on spatial and channel attention mechanisms. Finally, we conduct a series of experiments to evaluate the validity of these modules and the overall segmentation framework. Moreover, additional experiments are also conducted to verify the robustness of MB-Net on two public datasets, i.e., COVID-SemiSeg and CHASE_DB1. The results show that our MB-Net outperforms the state-of-the-art methods and has good performance for multiple medical image segmentation tasks.

2. METHODS

The overview of the proposed MB-Net is shown in Fig. 1 (a). For a single image, we first input it into an init layer (IL) to change the number of channels of the image and then input the image into four encoders with LFM for local feature information extraction and dimension reduction. Then, a AMSFM is applied to locate potential multi-scale objects in the first three stages of the encoder. Furthermore, five EBRMs are applied to guide the model to learn boundary information. Finally, four decoders with LFM are used to progressively detect and remove false positive and false negative interference, so that lateral ventricle objects could be accurately identified.

2.1. Local Feature Extraction Module

To enhance local information extraction and reduce information loss caused by downsampling, we propose a local feature extraction module (see Fig. 1(g)). When input features are processed by a convolution block, a series of feature maps are generated. These feature maps then go through two branches, one for further extraction of feature information and the other for generating probability maps about pixels. The results of the two branches are then multiplied to produce a feature map with increased importance of pixels. Next, a squeeze-and-excitation (SE) block is employed to further enhance the importance of the different feature maps. Finally, local feature information is obtained. In particular, the module well considers the spatial location information of the target region of the lateral ventricle in the feature map and the characteristic information of different channels, which make our model more amiable to the extraction of local information.

2.2. Adaptive multi-scale feature aggregation module

There are some challenges for segmenting the region of the lateral ventricle on MR images due to its distinct size in different slices. It is worth noting that a multi-scale module is the most popular solution to this problem. Therefore, we propose a new adaptive multi-scale feature aggregation module, which can comprehensively extract feature information of different scales based on different receptive fields to help the model better learn useful features.

Specifically, the AMSFM can be divided into two parts, i.e., improved inception module (IIM) and feature aggregation module (FAM), which are shown in Fig. 1 (b), (c), and (d), respectively. Among them, the IIM is derived from part of InceptionNet-V3 [6]. Different from it, IIM takes into account the importance of features extracted from different receptive fields to help the model learn useful feature information. Similarly, different from existing feature aggregation methods, we design a feature filtering mechanism, which considers the information redundancy that may occur in the process of feature aggregation. Besides, FAM can fully consider the similarity between different feature maps and realize the deep fusion of information at different scales.

2.3. Embedded boundary refinement module

The boundary of the lateral ventricle is blurred in MR images because plenty of choroids are densely distributed in or around the lateral ventricle. Therefore, the EBRM is designed to address the problem (see Fig. 1 (f)). Specifically, we first extract the boundary feature of every encoder stage by a convolutional layer with a 1×1 kernel. Then, we calculate the similarity to the label in the training process. Next, we use this similarity to guide the model to learn the boundary information. The similarity can be calculated by

$$Similarity = \frac{1}{n}\sum_{i=1}^{n}(y_i - \bar{y}_i)^2 \tag{1}$$

where y_i is the ground truth and $\bar{y}_i$ is the predicted value. n is the number of samples.

Compared with the traditional boundary extraction methods (including the Canny operator and gradient algorithm), the EBRM can extract the boundary information in the process of model learning, and then guide the model to realize the learning of boundary information. Notably, we remove the EBRM to avoid the involvement of label information during testing.

3. EXPERIMENTS AND RESULTS

To verify the performance of MB-Net, we conduct several experiments from different perspectives. (1) The ablation experiment of MB-Net is used to verify the validity of each module (i.e., AMSFM, EBRM, and LFM) on the LV-Seg dataset. Besides, we compare MB-Net with classical segmentation methods published in top journals or conferences in recent years to verify the accuracy of our model. (2) We also verify the robustness performance of MB-Net on other public datasets, i.e., COVID-SemiSeg and CHASE_DB1.

3.1. LV-Seg Dataset

We collected 173 cases from two hospitals, i.e., the Affiliated Hospital of Xiangnan University and the First People's Hospital of Chenzhou. We obtained 9314 slides from 109 cases, which ultimately constituted a lateral ventricle segmentation dataset with different ages, genders, and clinical manifestations. Especially, ground truths are annotated by experienced clinicians. We randomly divide the LV-Seg into a training set and a testing set in 8 to 2 ratios. The training set has 86 cases with 7310 slices and the test has 23 cases with 2004 slices.

3.2. Implementation Details

During the training process, we optimize our MB-Net using Adam optimizer with a learning rate of 0.0001. The maximum epoch is set as 35. Data augmentation methods (i.e., rotation, saturation, and contrast) are adopted to alleviate individual differences on our LV-Seg dataset. Besides, we use average dice coefficient (Dice), Jaccard (Jac), precision (Pre), and recall (Rec) to evaluate the segmentation results. Our MB-Net is trained using the PyTorch on an NVIDIA GEFORCE 2080Ti GPU.

3.3. Ablation experiment

To verify the effectiveness of the three modules, i.e., AMSFM, EBRM, and LFM, some ablation experiments are conducted to evaluate the influence of each module and its various combinations. It is worth noting that when verifying the necessity of LFM, we replace it with the IL. The numerical results are shown in Table 1.

Table 1. Quantitative comparison with different combination of our proposed modules.

AMSFM	EBRM	LFM	Dice	Jac	Pre	Rec
√			0.8976	0.8153	0.8965	0.9001
	√		0.8986	0.8169	0.9031	0.8953
		√	0.9038	0.8254	0.9052	0.9031
√	√		0.8986	0.8169	0.8911	0.9072
√		√	0.9038	0.8254	**0.9099**	0.8985
	√	√	0.9037	0.8252	0.8982	**0.9100**
√	√	√	**0.9042**	**0.8261**	0.9068	**0.9025**

As seen from Table 1, different modules contribute to the model performance for segmenting the lateral ventricle in different degrees. Besides, MB-Net achieves the best performance when AMSFM, EBRM, and LFM are all used, which demonstrates that our model has some potential for the lateral ventricle segmentation breaking through the problems of blur boundary, distinct size, and irregular shape of the lateral ventricle region on MRI.

3.4. Comparison with the state-of-the-art methods

To verify the performance of the model for lateral ventricle segmentation, we compare MB-Net with some state-of-the-art segmentation models published in top journals or conferences in recent three years. For a fair comparison, we adjusted each model to ensure that all models obtained the best segmentation performance. The numerical results and visual

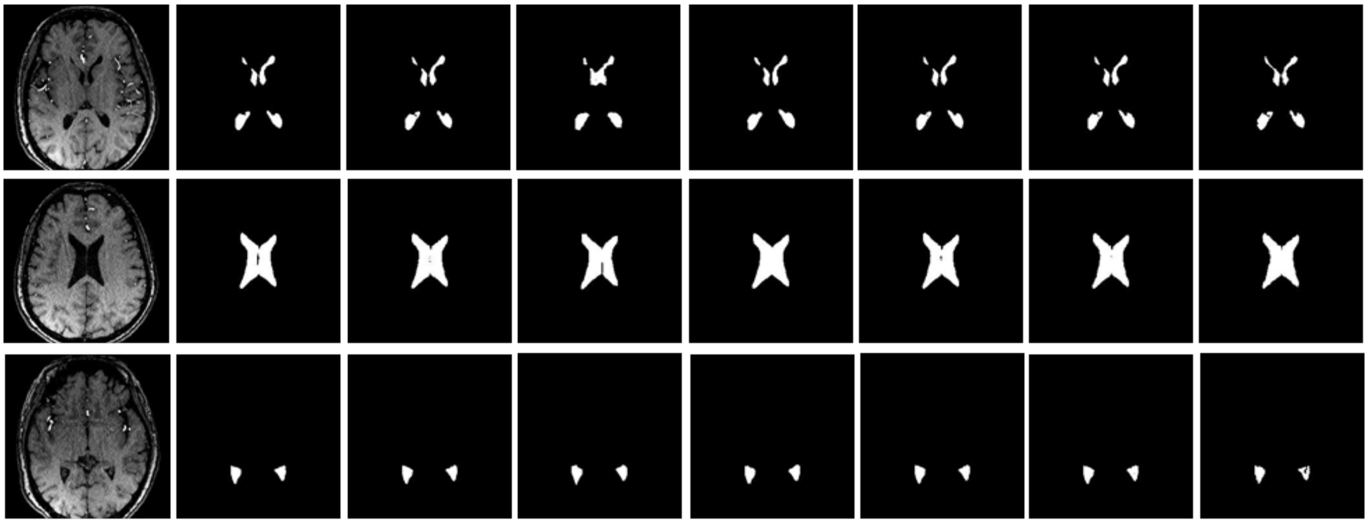

Fig. 2. Visual segmentation effects using different methods. (a) Original image. Segmentation results obtained by (b) GSCNN, (c) CA-Net, (d) PraNet, (e) PFNet, (f) TransUNet, and (g) MB-Net. (h) Ground truth.

effects are shown in Table 2 and Fig. 2, respectively. From the results, we can observe that our model outperforms the state-of-the-art methods.

Table 2. Comparisons with different segmentation methods.

Method	Year	Dice	Jac	Pre	Rec
GSCNN [7]	2019	0.8962	0.8129	0.8934	0.9001
CA-Net [8]	2020	0.9005	0.8199	0.9021	0.8997
PraNet [9]	2020	0.8407	0.7274	0.7883	0.9029
PFNet [10]	2021	0.8752	0.7796	0.8257	**0.9323**
TransUNet [11]	2021	0.8986	0.8168	0.9006	0.8973
MB-Net (Ours)	-	**0.9042**	**0.8261**	**0.9068**	0.9025

3.5. Validation of robustness performance

To verify the robustness of the model, we further conduct two experiments on COVID-SemiSeg [12] and CHASE_DB1 [13], respectively. We also compare the results with existing segmentation methods on the two datasets. The results are shown in Tables 3 and 4.

Table 3. Comparisons with the state-of-art methods on the COVID-SemiSeg dataset.

Method	Year	Dice	Sen	Spe
Unet++ [14]	2018	0.581	0.672	0.902
Gated-Unet [15]	2019	0.623	0.658	0.926
Dense-Unet [16]	2020	0.515	0.594	0.840
Inf-Net [12]	2020	0.682	0.692	**0.953**
MB-Net (Ours)	-	**0.762**	**0.789**	0.952

As can be seen from Table 3, compared with existing methods, our model has significantly improved on the COVID-SemiSeg, because Dice and Sen increased by 8% and 9.7%, respectively. Besides, Table 4 also demonstrates that MB-Net obtains a better performance on the CHASE_DB1. It implies that our model has great advantages in the segmentation of the target region, and also verifies that our model has an excellent robustness performance for medical image segmentation.

Table 4. Comparisons with the state-of-art methods on the CHASE_DB1 dataset. Evaluation metrics are consistent with [13], including Area Under the Curve (AUC), Sensitivity (Sen), Specificity (Spe), and Accuracy (Acc).

Method	Year	Auc	Sen	Spe	Acc
Yan *et al.* [17]	2019	0.9776	0.7641	0.9806	0.9607
Wang *et al.* [18]	2020	0.9871	0.8239	0.9813	0.9670
Li *et al.* [19]	2021	0.9810	0.7818	0.9719	0.9635
Wang *et al.* [13]	2021	0.9872	**0.8435**	0.9782	0.9630
MB-Net (Ours)	-	**0.9887**	0.7886	**0.9886**	**0.9674**

4. CONCLUSION

In this study, we first established a multi-center lateral ventricle segmentation dataset (LV-Seg) based on clinical practice. Then, an efficient and accurate convolutional neural network based on adaptive multi-scale feature aggregation and boundary-aware was proposed for lateral ventricle segmentation. Next, extensive experiments were conducted on LV-Seg to verify the performance of the model. Besides, to verify the robustness of the model, we also conducted some experiments on two public datasets, i.e., COVID-SemiSeg and CHASE_DB1. The results show that our MB-Net is superior to other methods in both LV-Seg and two public datasets, indicating that our model has great potential in medical image segmentation. In the future, we will replicate our model on the open source deep learning platform paddle.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grants 61802328, 61972333 and 61771415, the Research Foundation of Education Department of Hunan Province of China under Grants 21B0172 and 19B561, and the Baidu Program.

5. REFERENCES

[1] Gan K, "Automated segmentation of the lateral ventricle in mr images of human brain," in *2015 IEEE International Conference on Digital Signal Processing (DSP)*. IEEE, 2015, pp. 139–142.

[2] X Tang, K Kutten, C Ceritoglu, S Mori, and M I Miller, "Simultaneous skull-stripping and lateral ventricle segmentation via fast multi-atlas likelihood fusion," in *Medical Imaging 2015: Image Processing*. International Society for Optics and Photonics, 2015, vol. 9413, p. 94130W.

[3] HG Schnack, Hulshoff Pol HE, WF Baaré, MA Viergever, and RS Kahn, "Automatic segmentation of the ventricular system from mr images of the human brain," *Neuroimage*, vol. 14, no. 1, pp. 95–104, 2001.

[4] Muhan Shao, Shuo Han, Aaron Carass, Xiang Li, Ari M Blitz, Jerry L Prince, and Lotta M Ellingsen, "Shortcomings of ventricle segmentation using deep convolutional networks," in *Understanding and Interpreting Machine Learning in Medical Image Computing Applications*, pp. 79–86. Springer, 2018.

[5] Jiong Wu, Yue Zhang, and Xiaoying Tang, "Simultaneous tissue classification and lateral ventricle segmentation via a 2d u-net driven by a 3d fully convolutional neural network," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2019, pp. 5928–5931.

[6] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.

[7] Towaki Takikawa, David Acuna, Varun Jampani, and Sanja Fidler, "Gated-scnn: Gated shape cnns for semantic segmentation," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE, pp. 5228–5237.

[8] Ran Gu, Guotai Wang, Tao Song, Rui Huang, Michael Aertsen, Jan Deprest, Sébastien Ourselin, Tom Vercauteren, and Shaoting Zhang, "Ca-net: Comprehensive attention convolutional neural networks for explainable medical image segmentation," *IEEE transactions on medical imaging*, vol. 40, no. 2, pp. 699–711, 2020.

[9] Deng-Ping Fan, Ge-Peng Ji, Tao Zhou, Geng Chen, Huazhu Fu, Jianbing Shen, and Ling Shao, "Pranet: Parallel reverse attention network for polyp segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2020, pp. 263–273.

[10] Haiyang Mei, Ge-Peng Ji, Ziqi Wei, Xin Yang, Xiaopeng Wei, and Deng-Ping Fan, "Camouflaged object segmentation with distraction mining," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 8772–8781.

[11] Jieneng Chen, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, Yan Wang, Le Lu, Alan L Yuille, and Yuyin Zhou, "Transunet: Transformers make strong encoders for medical image segmentation," *arXiv preprint arXiv:2102.04306*, 2021.

[12] Deng-Ping Fan, Tao Zhou, Ge-Peng Ji, Yi Zhou, Geng Chen, Huazhu Fu, Jianbing Shen, and Ling Shao, "Inf-net: Automatic covid-19 lung infection segmentation from ct images," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2626–2637, 2020.

[13] Chang Wang, Zongya Zhao, and Yi Yu, "Fine retinal vessel segmentation by combining nest u-net and patch-learning," *Soft Computing*, vol. 25, no. 7, pp. 5519–5532, 2021.

[14] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang, "Unet++: A nested u-net architecture for medical image segmentation," *arXiv e-prints*, pp. arXiv–1807, 2018.

[15] Jo Schlemper, Ozan Oktay, Michiel Schaap, Mattias Heinrich, Bernhard Kainz, Ben Glocker, and Daniel Rueckert, "Attention gated networks: Learning to leverage salient regions in medical images," *Medical image analysis*, vol. 53, pp. 197–207, 2019.

[16] S Cai, Y Tian, H Lui, H Zeng, Y Wu, and G Chen, "Dense-unet: a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network.," *Quantitative Imaging in Medicine and Surgery*, vol. 10, no. 6, pp. 1275–1285, 2020.

[17] Zengqiang Yan, Xin Yang, and Kwang Ting Cheng, "A three-stage deep learning model for accurate retinal vessel segmentation," *IEEE Journal of Biomedical and Health Informatics*, 2019.

[18] D Wang, A Haytham, J Pottenburgh, O Saeedi, and Y Tao, "Hard attention net for automatic retinal vessel segmentation," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 12, pp. 3384–3396, 2020.

[19] Xiang Li, Yuchen Jiang, Minglei Li, and Shen Yin, "Lightweight attention convolutional neural network for retinal vessel image segmentation," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 1958–1967, 2021.