

## Letter

## Dendritic Deep Learning for Medical Segmentation

Zhipeng Liu<sup>1b</sup>, Zhiming Zhang<sup>1b</sup>, Zhenyu Lei<sup>1b</sup>,  
Masaaki Omura<sup>1b</sup>, Rong-Long Wang<sup>1b</sup>, and  
Shangce Gao<sup>1b</sup>, *Senior Member, IEEE*

Dear Editor,

This letter presents a novel segmentation approach that leverages dendritic neurons to tackle the challenges of medical imaging segmentation. In this study, we enhance the segmentation accuracy based on a SegNet variant including an encoder-decoder structure, an upsampling index, and a deep supervision method. Furthermore, we introduce a dendritic neuron-based convolutional block to enable nonlinear feature mapping, thereby further improving the effectiveness of our approach. The proposed method is evaluated on medical imaging segmentation datasets, and the experimental results demonstrate that it is superior to state-of-the-art methods in terms of performance.

**Introduction:** The significance of medical image segmentation in computer-assisted diagnosis and treatment planning cannot be overstated. It plays a crucial role in accurately identifying and outlining regions of interest, which aids in further analysis and medical decision-making. Nevertheless, the complexity and diversity of image content present challenges in the process of medical image segmentation [1]. Particularly, precisely identifying regions of interest while minimizing errors and artifacts that may occur during image capture and processing can be particularly difficult. For example, the presence of tumors with unique textures compared to the surrounding tissue can impede their accurate identification and segmentation. Furthermore, it is worth noting that the existing simple feature mapping techniques in computer vision exhibit limitations in effectively capturing the intricate details and fine-grained structures present in medical images. This deficiency considerably impedes the accurate representation and analysis of the images complex nuances and subtle characteristics, consequently compromising the quality of segmentation outcomes.

In recent years, the utilization of deep learning techniques for medical image segmentation has gained substantial interest, driven by the availability of an increasing number of medical image datasets and remarkable advancements in methodologies. Fully convolutional networks (FCN) simplify image segmentation by converting it into a pixel-level classification task, but may suffer from challenges in capturing overall context and preserving fine details [2]. Another widely used FCN architecture is U-Net, which incorporates skip connections to fuse high-level and low-level features, however, downsampling operations utilized in the network may result in information loss [3]. Therefore, through the use of pooling indices derived from the corresponding encoder's max pooling step, SegNet [4] can implement non-linear upsampling in the decoder, thus ensuring the effective

preservation of spatial information in the original image and the accurate restoration of fine image details. Although it has exhibited noteworthy efficacy in diverse semantic segmentation tasks, it may suffer from information loss when dealing with non-uniform sampled images.

To address the limitations of current medical image segmentation approaches, various researchers have introduced novel mechanisms to enhance segmentation accuracy. Abraham and Khan [5] proposed a deeply supervised attention U-Net for tumor segmentation in BUS images. The model was enhanced with a multi-scaled input image pyramid to improve the quality of intermediate feature representations. Chen *et al.* [6] added six side-out deep supervision modules, enabling the network to learn to predict accurate segmentation masks at multiple scales. We have introduced a multi-scale approach for constructing the network, demonstrating the effectiveness of the training strategy involving deep supervision for addressing these challenges.

Larkum [7] emphasized the importance of dendritic structures in neural understanding, revealing limitations in the conventional neuron description that neglect dendrites' computational properties. This omission has impeded progress in comprehending higher-level neural processes. Therefore, the utilization of dendritic neurons as a substitute for McCulloch-Pitts (MCP) neurons is both biologically interpretable and necessary. In alignment with this notion, the dendritic neuron model (DNM) has emerged as a prominent approach for classification and prediction tasks, leveraging its inherent nonlinear information processing capability [8]. Gao *et al.* [9] extend the DNM from a real-valued domain to a complex-valued domain and conduct an extensive evaluation of the resulting model, while we further extend the utilization of dendritic neurons to image segmentation.

Expanding on prior research, we propose a novel method called dendritic deep supervision neural network (DDNet), which combines biologically interpretable dendritic neurons with deep supervision during the training process. The use of dendritic neurons is expected to capture intricate details and fine-grained structures through nonlinear feature mapping, while deep supervision facilitates the learning of hierarchical representations and improves training efficiency and accuracy, thereby conferring supplementary advantages for medical image segmentation. The contributions of this study are summarized as follows:

- 1) Novel network architecture: We propose a novel network architecture that integrates dendritic neurons and a deep supervision mechanism into the SegNet framework, specifically designed to enhance medical image segmentation.
- 2) Improved feature representation: Our dendritic neuron-based feature extractor captures precise and informative features from medical images while enabling access to nonlinear feature mappings, further enhancing the performance of our approach.
- 3) Enhanced training effectiveness: Our approach incorporates a deep supervision mechanism and a customized loss function, optimizing the training process. These enhancements improve efficiency and effectiveness by providing additional guidance for feature acquisition and facilitating the capture of fine-grained details and hierarchical representations.
- 4) Superior performance: The effectiveness of our proposed method is demonstrated through its outperformance of existing networks in experiments conducted on several datasets.

**Proposed method:** As illustrated in Fig. 1, the comprehensive framework comprises two fundamental components: deep supervision SegNet (DSegNet) and DNM modules. The DSegNet module employs deep supervision by utilizing the feature results from both deep and shallow layers of the SegNet variant. The final feature maps generated by the last layer of DSegNet are further processed by the DNM module to perform nonlinear feature mapping, resulting in the desired segmentation outcomes. These structures work in tandem to achieve improved medical image segmentation performance.

Corresponding author: Shangce Gao.

Citation: Z. Liu, Z. Zhang, Z. Lei, M. Omura, R.-L. Wang, and S. Gao, "Dendritic deep learning for medical segmentation," *IEEE/CAA J. Autom. Sinica*, vol. 11, no. 3, pp. 803–805, Mar. 2024.

Z. Liu, Z. Zhang, Z. Lei, M. Omura, and S. Gao are with the Faculty of Engineering, University of Toyama, Toyama-shi 930-8555, Japan (e-mail: m22c8006@ems.u-toyama.ac.jp; d2272007@ems.u-toyama.ac.jp; leizg@eng.u-toyama.ac.jp; momura@eng.u-toyama.ac.jp; gaosc@eng.u-toyama.ac.jp).

R.-L. Wang is with the Faculty of Engineering, University of Fukui, Fukui-shi 910-8507, Japan (e-mail: wang@u-fukui.ac.jp).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JAS.2023.123813

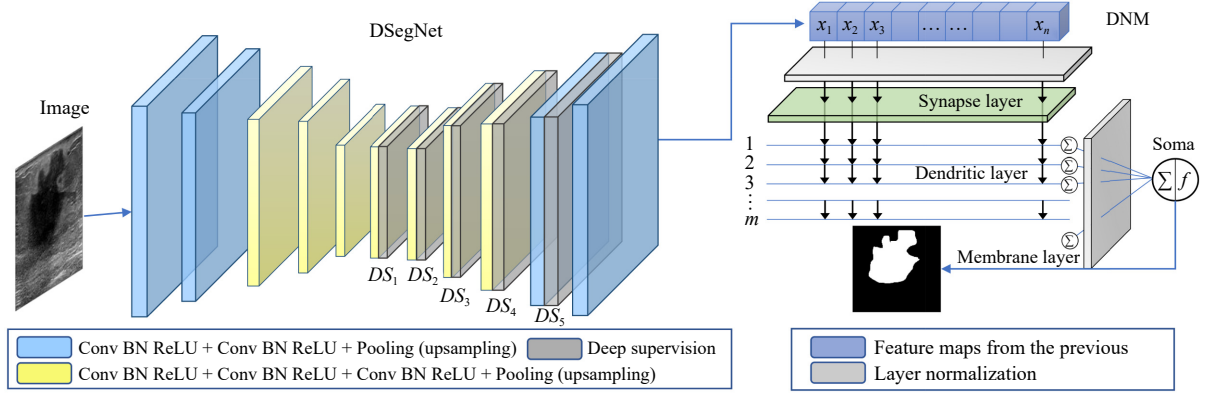


Fig. 1. The network structure of DDNet consists of DSegNet and DNM.

**DSegNet:** Within the proposed framework, the DSegNet module effectively captures multi-scale information for more accurate segmentation outcomes. To improve the SegNet architecture, we introduce an additional layer for deep supervision and feature enrichment. By incorporating multiple deep supervision signals ( $DS_1$ ,  $DS_2$ ,  $DS_3$ ,  $DS_4$ , and  $DS_5$ ) in Fig. 1 from intermediate layers and upsampled feature maps, we refine segmentation results at various scales, effectively utilizing both local and global contextual information. This deep supervision mechanism enhances the model's capacity to handle images with irregular sampling patterns and accurately restore intricate image details. Furthermore, SegNet's pooling indices in the decoder path preserve spatial information during upsampling. This feature ensures the faithful restoration of fine image details and robust handling of unevenly sampled images.

**DNM:** To further improve the effectiveness of the network, we introduce the DNM module, which leverages the unique computational properties of dendritic neurons. In Fig. 1, the DsegNet output is processed by the DNM module through the synapse layer, dendritic layer, membrane layer, and soma layer. According to the final feature maps generated by the last layer of SegNet, the DNM module performs nonlinear feature mapping, enhancing the model's ability to capture intricate details and fine-grained structures in medical images. This nonlinear mapping facilitates the extraction of more discriminative and informative features, contributing to improved segmentation accuracy. Fig. 2 visually illustrates the distinct segmentation results obtained by different methods, where sub-figures located in the top three rows represent distinct segmentation for DatasetB, while the rest of the figure is the illustration of Polyp. The green and red curves represent the contours of the masks, while the white pixels indicate the predicted areas. This visual representation highlights the accurate prediction and delineation capabilities of our method for the segmented regions.

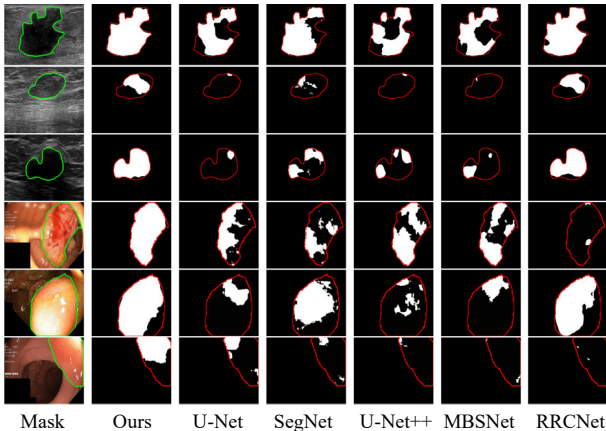


Fig. 2. Segmentation results of different methods.

In contrast to prior methods, such as the utilization of proposed neurons high-order coverage function neural network (HCFNN) instead of fully connected layers [10], our method focuses on optimizing convolutions and incorporates additional preprocessing steps to enhance the utilization of the dendritic layer. Specifically, to optimize the utilization of the dendritic layer, we perform several dimension transformations on the input feature. Next, we apply normalization to the input feature, ensuring that the data falls within a suitable range for optimal processing. Following, we replicate the input feature vectors along the channel dimension to match the number of dendritic branches to achieve the desired effect of feature reuse. These modifications enable the proposed neurons to serve as replacements for neuron feature maps, expanding their role beyond simple feature extraction in fully connected layers. In addition, the learnable parameters, including  $k$ ,  $w_{ij}$ , and  $q_{ij}$ , are randomly initialized within the range of (0, 1).

$$Z_{ij} = \max(0, k \times (w_{ij} \times x - q_{ij})) \quad (1)$$

where  $Z_{ij}$  represents the output of the  $j$ -th dendritic branch for the  $i$ -th element of the input  $x$ . The operation involves element-wise multiplication between the weights  $w_{ij}$  and the input  $x$ , followed by subtracting the threshold  $q_{ij}$ . The rectified linear unit (ReLU) activation function is then applied, which clips negative values to zero, introducing non-linearity and facilitating the preservation of informative signals while suppressing noise and irrelevant information. The amplification of the resulting signal is controlled by the parameter  $k$ . Equation (1) captures the essential information integration and processing within the dendritic layer.

Then the dendritic layer receives signals from the previous synaptic layer and performs a summation operation for each dendritic branch. In this context, the  $j$ -th dendritic branch aggregates  $N$  input signals  $Z_{ij}$ , resulting in the computation:  $V_j = \sum_{i=1}^N Z_{ij}$ .

Subsequently, the membrane layer accumulates the signals from all dendritic branches through another summation operation. This layer combines the outputs of  $M$  dendritic branches to generate a collective representation denoted as  $Y = \sum_{j=1}^M V_j$ .  $Y$  represents the overall input integration from the dendritic layer.

Finally, the soma layer processes the output of the membrane layer using a sigmoid activation function. This function determines the firing behavior of the neuron, and the final output prediction  $P$  is given by

$$P = \frac{1}{1 + e^{-k_s(Y - q_s)}} \quad (2)$$

where  $k_s$  and  $q_s$  are additional learnable parameters, which are both initially randomly set within the range of (0, 1). During the training phase, we employ the Adam optimizer to optimize the learnable parameters.

**Loss function:** A combined loss function is proposed in our framework, comprising a binary cross-entropy (BCE) loss from the final DNM module and a depth-supervised focal loss for upsampling layers

$$\begin{aligned}
\mathcal{B} &= -\frac{1}{n} \sum_{i=1}^n [T_i \log P_i + (1 - T_i) \log(1 - P_i)] \\
\mathcal{F} &= -\frac{1}{n} \sum_{i=1}^n [\alpha(1 - e^{-\mathcal{B}})^{\gamma} \mathcal{B}] \\
L &= \mathcal{B} + \lambda \sum_{j=1}^5 \mathcal{F}_j
\end{aligned} \tag{3}$$

where  $\mathcal{B}$  measures dissimilarity between predicted and ground truth labels, aiding in representation acquisition.  $\mathcal{F}$  incorporates intermediate supervision signals, leveraging parameters  $\alpha$  and  $\gamma$  to handle class imbalance and prioritize challenging samples ( $\alpha = 0.8$ ,  $\gamma = 2$ ). In the loss function,  $n$  denotes the pixel count,  $T_i$  and  $P_i$  represent true labels and prediction results, respectively, and  $\lambda$  denotes the deep supervision loss coefficient.

**Experiment:** This section presents experimental evaluations on three distinct datasets using four-fold cross-validation. Additionally, we performed ablation experiments and parameter analysis to enhance the evaluation of our proposed approach. The datasets employed in our experiments encompass the DatasetB (primarily designed for tumor segmentation tasks), the STU dataset (comprising a smaller tumor dataset comprising a mere 42 images), and an extensively utilized dataset for Polyp segmentation tasks.

As shown in Table 1, we compare DDNet with classical and state-of-the-art networks, among which RRCNet utilizes a deep supervision technique and employs a refinement residual convolutional network for breast tumor segmentation, MBSnet is a multi-branch medical image segmentation network that captures local and global information [11]. To ensure fairness, all comparative experiments are conducted with the optimal configuration of parameters and loss functions using the same experimental setup.

Table 1. Comparison of Results

Methods	DatasetB		STU		Polyp	
	IoU	Dice	IoU	Dice	IoU	Dice
U-Net	69.5 ± 7.2	81.5 ± 5.4	75.9 ± 3.7	86.1 ± 2.4	68.8 ± 1.2	80.8 ± 1.0
SegNet	67.3 ± 4.3	79.8 ± 3.4	72.0 ± 5.0	83.6 ± 3.4	66.0 ± 2.9	78.7 ± 2.4
U-Net++	71.4 ± 3.7	82.8 ± 2.8	74.5 ± 5.6	85.1 ± 3.7	67.4 ± 0.6	79.9 ± 0.6
MBSNet	69.2 ± 3.7	80.1 ± 3.7	75.7 ± 5.2	86.0 ± 3.5	67.5 ± 2.1	80.0 ± 1.6
RRCnet	71.81 ± 2.7	80.4 ± 2.3	76.5 ± 4.7	86.5 ± 3.0	69.1 ± 1.7	81.1 ± 0.0
Ours	<b>72.6 ± 1.3</b>	<b>83.5 ± 0.7</b>	<b>77.3 ± 6.9</b>	<b>87.1 ± 3.8</b>	<b>70.6 ± 1.4</b>	<b>81.9 ± 1.3</b>

Moreover, the results of the parameter discussion are presented in Tables 2 and 3. The optimal result is achieved when setting  $\lambda$  to 0.1 and  $M$  to 10 in the framework. It is worth noting that we conduct separate parameter discussions on both DatasetB and the STU dataset, and the effectiveness of the chosen parameters is confirmed in each case. Additionally, we configure the epoch to 100, the batch size to 8, and the learning rate to  $4E-5$ . Table 4 displays the outcomes of the ablation experiments performed to assess the efficacy of our proposed methodology. The objective of these experiments is to evaluate the influence of incorporating deep supervision (DS) and deep supervision with a dendritic neuron model (DS+DNM) on the performance of the SegNet baseline. Intersection over union (IoU), Dice coefficient, F1 score, and recall are adopted as evaluation metrics to quantify the enhancements achieved by integrating these components.

Table 2. Discussion on the Parameters  $M$  and  $\lambda$  in DatasetB

$\lambda$	IoU	Dice	$M$	IoU	Dice
0.01	71.3 ± 3.6	80.0 ± 4.4	5	69.7 ± 6.0	81.4 ± 4.5
0.1	<b>72.6 ± 1.3</b>	<b>83.5 ± 0.7</b>	10	<b>72.6 ± 1.3</b>	<b>83.5 ± 0.7</b>
0.3	70.6 ± 2.1	79.5 ± 2.5	20	70.6 ± 5.0	81.8 ± 4.1

**Conclusion:** This investigation explores the issue of segmentation in medical images. In order to obtain better segmentation outcomes,

Table 3. Discussion on the Parameters  $M$  and  $\lambda$  in STU

$\lambda$	IoU	Dice	$M$	IoU	Dice
0.01	77.1 ± 7.8	86.8 ± 5.3	5	74.3 ± 9.2	84.5 ± 6.8
0.1	<b>77.3 ± 6.9</b>	<b>87.1 ± 3.8</b>	10	<b>77.3 ± 6.9</b>	<b>87.1 ± 3.8</b>
0.3	76.7 ± 7.2	86.6 ± 4.5	20	76.3 ± 3.2	86.5 ± 2.0

Table 4. Ablation Experiments on DatasetB

	IoU	Dice	F1	Recall
SegNet	67.3 ± 4.3	79.8 ± 3.4	82.8 ± 3.4	75.9 ± 3.2
SegNet+DS	70.4 ± 1.7	81.8 ± 1.2	84.7 ± 1.5	78.8 ± 2.9
SegNet+DNM	69.8 ± 6.4	81.6 ± 4.5	84.6 ± 2.7	77.5 ± 8.6
SegNet+DS+DNM	<b>72.6 ± 1.3</b>	<b>83.5 ± 0.7</b>	<b>85.9 ± 1.4</b>	<b>80.4 ± 1.2</b>

we propose a novel network that integrates biologically interpretable dendritic neurons and employs deep supervision during training to enhance the model's efficacy. To evaluate the effectiveness of our proposed methodology, we conducted comparative trials on datasets STU, Polyp, and DatasetB. The experiments demonstrate the superiority of our proposed approach.

**Acknowledgments:** This work was partially supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (JP22H03643), Japan Science and Technology Agency (JST) Support for Pioneering Research Initiated by the Next Generation (SPRING) (JPMJSP2145) and JST through the Establishment of University Fellowships Towards the Creation of Science Technology Innovation (JPMJFS2115).

## References

- [1] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3523–3542, 2022.
- [2] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.
- [3] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Medical Image Computing and Computer-Assisted Intervention*, 2015, pp. 234–241.
- [4] V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.
- [5] N. Abraham and N. M. Khan, "A novel focal tversky loss function with improved attention U-Net for lesion segmentation," in *Proc. IEEE 16th Int. Symp. Biomedical Imaging*, 2019, pp. 683–687.
- [6] G. Chen, Y. Dai, and J. Zhang, "RRCNet: Refinement residual convolutional network for breast ultrasound images segmentation," *Engineering Applications of Artificial Intelligence*, vol. 117, p. 105601, 2023.
- [7] M. E. Larkum, "Are dendrites conceptually useful?" *Neuroscience*, vol. 489, pp. 4–14, 2022.
- [8] S. Gao, M. Zhou, Y. Wang, J. Cheng, H. Yachi, and J. Wang, "Dendritic neuron model with effective learning algorithms for classification, approximation, and prediction," *IEEE Trans. Neural Networks and Learning Systems*, vol. 30, no. 2, pp. 601–614, 2019.
- [9] S. Gao, M. Zhou, Z. Wang, D. Sugiyama, J. Cheng, J. Wang, and Y. Todo, "Fully complex-valued dendritic neuron model," *IEEE Trans. Neural Networks and Learning Systems*, vol. 34, no. 4, pp. 2105–2118, 2023.
- [10] X. Ning, W. Tian, Z. Yu, W. Li, X. Bai, and Y. Wang, "HCFNN: Higher order coverage function neural network for image classification," *Pattern Recognition*, vol. 131, p. 108873, 2022.
- [11] S. Jin, S. Yu, J. Peng, H. Wang, and Y. Zhao, "A novel medical image segmentation approach by using multi-branch segmentation network based on local and global information synchronous learning," *Scientific Reports*, vol. 13, no. 1, p. 6762, 2023.