Blood Vessel Segmentation from Low-Contrast and Wide-Field Optical Microscopic Images of Cranial Window by Attention-Gate-Based Network

Yunheng Wu[1] Masahiro Oda[1] Yuichiro Hayashi[1] Takanori Takebe[2,3,4]
Shogo Nagata[5] Cheng Wang[1] Kensaku Mori[1,6,*]
[1]Nagoya University, Japan [2]Tokyo Medical and Dental University (TMDU), Japan
[3]Cincinnati Children's Hospital Medical Center, USA [4]University of Cincinnati, USA
[5]Keio University, Japan [6]National Institute of Informatics, Japan
{yunhengwu, moda, yhayashi}@mori.m.is.nagoya-u.ac.jp, Takanori.Takebe@cchmc.org,
shogo.nagata@keio.jp, kensaku@is.nagoya-u.ac.jp

Abstract

The stereomicroscope, which is an optical microscope, is used to observe the organoids cultured in cranial windows. A cranial window is a light accessible observation window made on the brain of mice through craniotomy. Organoids research is often conducted on cranial windows. Hence, the observation of blood vessels in them is important for organoid research, like organoid vascularization. Therefore, achieving a simple, low-cost method that extracts blood vessel structures would significantly help researchers observe the blood vessels in cranial windows from microscopic images. However, wide-field optical microscopic images taken by stereomicroscope suffer from low contrast and dura mater occlusion, complicating the observation of the blood vessels in such images. To address such problems and assist researchers who are observing vascular structures, we propose a method that segments blood vessels in cranial windows from low-contrast and wide-field microscopic images. Our method is based on the Attention U-Net framework and clDice, which considers the connectivity of blood vessels. In addition, for low-contrast and partial occlusion problems, we used contrast enhancement and dehazing as preprocessing steps. Our method achieved a Dice score of 75.56%, a clDice score of 79.95%, and the Accuracy of 91.41% on our microscopic image dataset, suggesting that our method can extract blood vessels from low-contrast and wide-field microscopic images better than other methods.

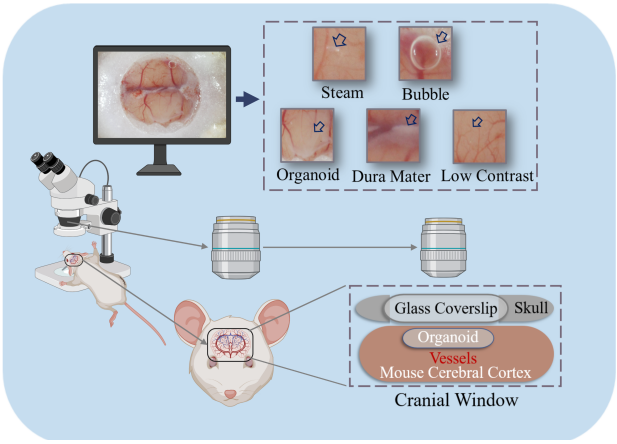

Figure 1. Optical stereomicroscope observes a cranial window in organoid experiment. There are blood vessels are on the mouse cerebral cortex, where organoid experiment is conducted. A glass coverslip protects cranial window to ensure a stable environment for long-term imaging. Biomedical researchers observe blood vessels and organoids in a cranial window through a stereomicroscope across a glass coverslip. Steam, bubbles, organoids and dura mater exist in images, all of which impact blood-vessel observation.

1. Introduction

Cranial window (CW) technology refers to making an optically accessible window by craniotomy on mice for long-term imaging and experimental observations of the mouse cerebral cortex. The structure and the state of blood vessels in CWs are essential for many biological experiments implemented in the mouse cerebral cortex [6]. Organoids, also known as the Mini Organ, are complex three-dimensional tissue with similar architectures and functionalities to organs [19]. Organoids are widely used in many kinds of biomedical research, including regenerative

*Corresponding author

medicine [3], drug development [22], and disease modeling [20]. In 2014, Takebe *et al*. [25] generated a vascularized and functional human liver in a cranial window. Since the structure of blood vessels in cranial windows is essential to observe the anastomosis between vascularized organoids and blood vessels in the mouse cerebral cortex, we focus on blood vessels observation in the optical microscopic images for organoid cultivation experiments in CWs.

However, as shown in Fig. 1, the microscopic images of cranial windows taken by stereomicroscope have many problems that complicate the observation of blood vessels. First, the blood vessels in the mouse cerebral cortex are covered with the cerebrospinal fluid, resulting in low contrast between the vessels and the background. Second, the cranial window needs to be covered with a glass coverslip to prevent contamination from external matter and to enable a stable fixation in favor of long-term observation [28]. An optical microscope is used to observe and photograph cranial windows through a glass coverslip that reduces optical transmittance. In addition, during a craniotomy in mice, entirely removing the thick, opaque dura mater is difficult, and it also easily regenerates, which hinders the optical transparency of longitudinal imaging [5]. Finally, organoids are cultivated in a living mouse that obviously needs to eat and move around. Consequently, the glass coverslip's outside becomes contaminated with feed or dust, which degrades the observation and analysis of microscopic images.

To the best of our knowledge, there are currently three kinds of methods for improving the light accessibility and observability of cranial windows: using some material with good light transmittance to replace glass coverslips [5], improving the manufacturing process to get a clear cranial window [7], and improving the imaging quality [9]. These methods can make optical microscopes with which it is easier to observe such brain tissues as the blood vessels in the CWs. However, widely applying these methods is difficult because of the limitations of complicated operations or expensive special materials. As far as we know, few studies on have improved the observability of the optical microscopic images of cranial windows by methods based on computer vision.

Based on the progress of artificial intelligence and computer vision technology, we anticipate a method that can assist researchers who are observing cranial windows by processing microscopic images through artificial intelligence rather than complex operations or expensive materials. In 2021, Wu *et al*. [32] used the U-Net framework to segment blood vessels in a cranial window. However, error segmentation occurred in some blood vessels' edges occluded by dura mater and organoids. In addition, for some small vessels, the connectivity of the segmented vascular structures is inferior. Therefore, to solve these problems, we use microscopic images taken in multiple periods of organoid culture to make a data set named the Microscopic Image of Cranial Window (MICW) for vascular structure segmentation. In addition, we propose a deep-learning-based method that segments the vascular structure from low-contrast and wide-field microscopic images.

The following are our primary contributions:

- From the perspective of computer vision and deep learning, we propose a simple, low-cost method to segment the blood vessels in the microscopic images of cranial windows to help researchers observe the blood vessel structure of organoid cultures without complex operations or expensive materials.
- We propose a method for blood vessel segmentation in the microscopic images of cranial windows. It uses inverted gray images to perform an advanced attention operation on the feature maps from skip connections before they are input into the attention gate, further increasing the Attention U-Net focus on the blood vessel regions.
- We introduce the image dehazing method as a preprocessing step to better segment the boundary between dura mater and blood vessels and reduce the low-contrast problem caused by the glass coverslip.

This paper is structured as follows. In Sec. 2, we introduce related works on segmentation. In Sec. 3, we describe our blood vessel segmentation method from microscopic images in detail. We evaluate our method for the segmentation of blood vessels in cranial windows in Sec. 4. Conclusions and future work are described in Sec. 5.

2. Related Work

Blood vessel segmentation in medical images. Blood vessel segmentation in fundus images resembles our work and, provides many references for our work. The retinal blood vessel structure in fundus images has clinical importance for doctors when they diagnose diseases such as diabetic retinopathy. Thus, to assist doctors who are observing the vascular structure, Wang *et al*. [30] extracted retinal blood vessels from fundus images by combining the Convolutional Neural Network (CNN) and Random Forest (RF). In 2015, Fully Convolutional Networks (FCN) [11] and U-Net [18] were proposed and applied in many image segmentation tasks and achieved favourable performance. Subsequently, Wang *et al*. [31] applied the U-Net to retinal blood vessel segmentation and outperformed the previous methods.

The advantages of the attention mechanism in image processing also been shown in previous research. Anderson *et al*. [1] proposed a combined bottom-up and top-down visual attention mechanism to propose image regions and applied this approach to image captioning. Oktay *et al*. [15]

proposed a novel self-attention gating module to segment the pancreas in 3D abdominal CT images and improve the prediction performance. Subsequently, Li *et al*. [10] proposed a connection sensitive Attention U-Net and a connection sensitive loss to segment the retinal blood vessels and improve the connectivity of segmented results. Shit *et al*. [23] proposed the clDice loss function based on topology information to improve the connectivity of results for tubular-structure segmentation. The corresponding evaluation metric clDice score is used to appraise the segmentation results of tubular structures. The clDice loss function achieved exceptional performance in various datasets for tubular-structures segmentation.

Blood vessel segmentation in microscopic images. Some previous research achieved blood vessel segmentation from microscopic images to analyze blood vessels in biological experiments. In 2012, Soetikno *et al*. [24] extracted the vascular structure from the optical resolution photoacoustic microscopic images of mice brains to analyze the ischemic strokes of mice. Virtudazo *et al*. [29] segmented blood vessels from microcirculation images to measure the blood flow velocity. Deep-learning-based segmentation methods have recently achieved exceptional performance in many image segmentation tasks. Prentašić *et al*. [17] used a deep learning network to extract the foveal microvasculature in optical coherence tomography angiography (OCTA) images. Hu *et al*. [8] used CNN to segment cerebral vessels in light-sheet microscopic images. Todorov *et al*. [26] used CNN and the transfer learning approach for whole mouse brain segmentation from confocal microscopic images, which have higher resolution and higher contrast than optical microscopic images. However, previous research focued on segmenting blood vessels from high-resolution and high-contrast images, like photoacoustic and confocal microscopic images. Blood vessels are easily distinguished from their background. To our knowledge, research is at an early stage on blood vessel segmentation in optical microscopic images with low resolution and low contrast. Wu *et al*. [33] used a line detector to extract blood vessels from low-contrast stereomicroscope images. Nercessiany *et al*. [14] extracted cortical blood vessels from craniotomy images taken by an operating microscope during craniotomy to simplify how doctors can observe how the brain might have shifted in both preoperative and intraoperative situations.

3. Method

3.1. Overview

In this section, we specifically introduce our method, including preprocessing, network architecture, and loss function. Fig. 2 shows how our method segments blood vessels from microscopic images.

3.2. Preprocessing Microscope Images of CW

We use a series of preprocessing steps to solve the current problems in the microscopic images of CWs to improve vessel segmentation, as shown in Fig. 2 (b).

Image dehazing. Due to the occlusion of the glass coverslip and the dura mater, the light transmittance is reduced, which affects the segmentation of blood vessels. Similarly, a haze affects light transmittance during imaging, degrading the images. In 2010, He *et al*. [4] proposed a dark-channel-based method to remove the haze from a single image. Therefore, in our work, we use the dehazing method based on the dark channel to preprocess the microscopic images of cranial windows to reduce the influence of the occlusion of the glass coverslip and the dura mater on blood vessel segmentation.

Image denoising and Gamma correction. After dehazing, images become dark, and the noise in them is amplified, which is not conducive to subsequent segmentation. Accordingly, we apply the Gamma Correction to improve the image brightness and used the Bilateral filter to remove the noise in the images. The Bilateral filter [27] can maintain the edge of small blood vessels during denoising.

Green channel extraction. Inspired by many blood vessel segmentation methods for fundus images [34], we separated RGB channels and visualized the microscopic images of each one. Compared with RGB images, we found that the extracted green channel images have higher contrast, and their blood vessel structure is observed more clearly. For this reason, we extracted the green channel images in preprocessing. In Sec. 4, we conducted an ablation study and discussed the influence of both green channel and RGB images on the blood vessel segmentation results on our dataset.

Contrast enhancement. The color of the blood vessels resembles the surrounding tissues and cerebrospinal fluid is also present in the mouse cerebral cortex. A stereomicroscope also takes images across a glass coverslip. These problems increase the difficulty of distinguishing between the blood vessels and the surrounding tissues: that is, the low-contrast problem. Therefore, we use the Contrast Limited Adaptive Histogram Equalization (CLAHE) [36] to enhance the contrast of the images to simplify distinguishing the vascular structure in optical microscopic images.

3.3. Attention U-Net with Advance Vessel Attention

The Attention U-Net. Many deep-learning-based segmentation models like the FCN framework [11] have been proposed and achieved excellent performance for diverse segmentation tasks. Various non-vascular regions in the wide-field microscopic images of the cranial window will cause rough segmentation results. In this paper, to tackle the problem and make the model focus more on blood vessel regions, we follow the Attention U-Net framework proposed by Oktay *et al*. [15] who achieved exceptional segmenta-

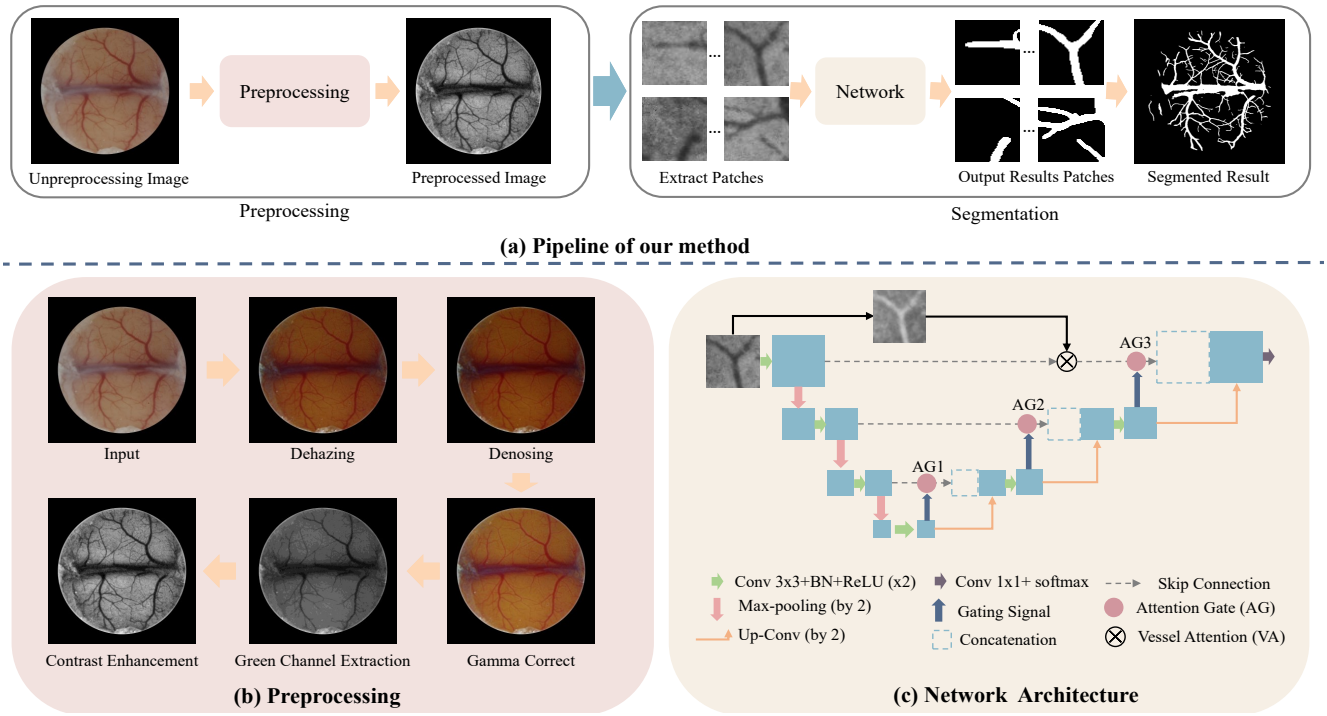

Figure 2. Details of our method: (a) pipeline of our method; (b) our specific preprocessing method; (c) architecture of the Attention U-Net with vessel attention for blood vessel segmentation.

tion results for pancreas segmentation in 3D abdominal CT images. The Attention Gate (AG), which reduces irrelevant features, is the most critical method for the Attention U-Net. The feature maps from the skip connection and the previous upsampled layer are the Attention Gate's input.

Advance Vessel Attention. Inspired by [35] and [21], we introduce gray-inversion input images as an advanced attention coefficient to emphasize the vessel regions, which are used in feature maps from the skip connection in a 4-layer Attention U-Net. Specifically, the Attention U-Net's input images are gray images whose pixel values have been normalized between 0 and 1. The pixel values of the blood vessels are generally lower than the other pixels in gray images. Therefore, after inverting the input images, the pixel values of the blood vessel region are higher, and those of the non-vessel region are lower. The gray-inversion images can be used as a coefficient (α), which we call THE Vessel Attention Coefficient (VAC). Before inputting the feature maps from the skip connection into the AG, we perform an advanced attention computation to the feature maps. The advance attention coefficient can be described as follows:

$$\hat{x}_i^n = \alpha_i^n x_i^n, \tag{1}$$

where x_i^n represents the value of the i-th pixel in the feature map from the n-th skip connection. α_i^n is the Vessel Attention Coefficient for i-th pixel in n-th feature map from the connection skip. $\hat{x}_i^n$ is the new value of the i-th pixel in the feature map from the n-th skip connection. After the advanced vessel attention operation, additive attention [15] calculates the gating coefficient, which can be described as follow:

$$\beta_i^n = \sigma(q_{att}^n(\hat{x}_i^n, g_i; \theta_{att})), \tag{2}$$

where $\sigma(\cdot)$ is sigmoid activation function. The attention gate is characterized by set of parameters θ_{att}. q_{att}^n represents a series computation on $\hat{x}_i^n$ and g_i by parameters θ_{att}. The gating signal g_i comes from a coarser scale [15]. In our work, we just used an advance vessel attention (α_i^3) in the last feature maps from the connection skip as shown in Fig. 2 (c). This is because gray-inversion images need to be downsampled to be used as a vessel attention map in any other connection skip, but the gray inversion image loses many vessel feature after multiple downsampling.

3.4. Loss Function Considering Vessel Topology

Diverse loss functions exist for image segmentation tasks. For example, the Dice coefficient [13] is commonly used as a loss function that mainly calculates the training by measuring the overlap of the prediction and the ground truth. However, for blood vessel segmentation, connectivity is a factor that has to be considered. For this reason, in 2021 Shit *et al.* [23] proposed a loss function that considered vessel topology and connectivity called the centerline

Dice (clDice), which is calculated on the intersection of the segmentation results and their morphological skeletons. In addition, clDice can evaluate the vascular structure and vessel connectivity of results. In our work, we use clDice to train the network. The clDice function can be expressed as follows:

$$T_p(S_P, V_L) = \frac{|S_P \cap V_L|}{|S_P|}, \tag{3}$$

$$T_s(S_L, V_P) = \frac{|S_L \cap V_P|}{|S_L|}, \tag{4}$$

$$clDice(V_P, V_L) = 2\frac{T_p(S_P, V_L)T_s(S_L, V_P)}{T_p(S_P, V_L) + T_s(S_L, V_P)}, \tag{5}$$

where V_P and V_L are the predicted segmentation mask and the ground truth mask, respectively. S_P and S_L are the skeletons which are extracted from V_P and V_L, respectively. $Tp(S_P, V_L)$ refers to the topology precision, and $Ts(S_L, V_P)$ is the corresponding sensitivity. As the Dice definition, clDice is defined as the harmonic mean of two measures. The final clDice function, which combines the Dice function and clDice function, is expressed as:

$$\mathcal{L} = (1 - \lambda)(1 - Dice) + \lambda(1 - clDice), \tag{6}$$

where λ is scale factor.

4. Experiments and Results

4.1. Dataset and Metrics

Dataset. Raw images of 5184×3456 pixels were taken by a experienced doctor using a camera on the left eyepiece of a stereomicroscope. These microscopic images were taken at different periods in organoid cultivation on different mice. A stereomicroscope is a kind of optical microscope that can directly observe objects without slicing. As shown in Fig. 3, we made our microscopic image dataset (MICW) for blood vessel segmentation in cranial windows follows: (1) First, we cut the images with an aspect ratio of 1:1 and resized them to 3000×3000 pixels to remove such irrelevant regions as bone cement. (2) Second, we masked the remaining regions that could not be cropped and kept the circular cranial window areas to reduce the subsequent computation cost. (3) Finally, we made the FOV masks and used MIPAV software [12] to make Ground Truth (GT) for each masked images. Therefore, our dataset included 20 microscopic images of cranial windows of 3000×3000 pixels for training and testing of blood vessel segmentation.

Metrics. For evaluation, we used Dice Score, Accuracy, Sensitivity, and Specificity, which are commonly used to appraise such tubular-structure segmentation tasks as retinal blood vessel and brain vessel segmentation.

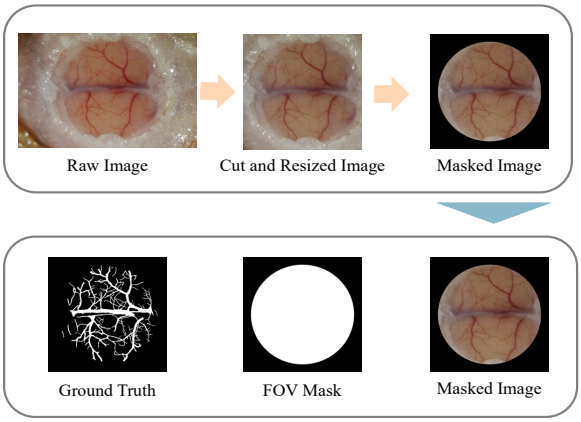

Figure 3. Process of making our dataset: There are corresponding the GT and the FOV masks for each masked image.

To evaluate the connectivity of the extracted blood vessels, we also used the clDice score *et al*. [23], which is based on morphology and corresponds with a soft clDice. The clDice score's formula is shown in Eq. 5, which calculates the intersection of the segmentation masks and their morphological skeletons.

4.2. Implementation Detail

Our experiments were implemented by the Pytorch [16] deep learning library using NVIDIA GeForce GTX 1080Ti. The loss functions were used for optimization with the Adam, and the learning rate was 0.0005.

In our experiment, each microscopic image used for network training was 3000×3000 pixels. Due to GPU limitations, we downsampled each image to 1024×1024 pixels and randomly cut the images into 2560 patches with 64×64 pixels and input the patches into the networks for training. For testing, the test image was also cut into patches of 64×64 pixels and input into the networks. After outputting the prediction patches, we spliced the output prediction patches into the whole prediction, and the pixel values with a blood vessel probability of more than 0.5 were regarded as a blood vessel region to generate the final prediction result. Because we used a small dataset, we flipped the patch in the up-down and left-right directions with a probability of 0.5 and rotated the patches at three angles (90°, 180°, 270°) to augment the data. Additionally, we conducted a five-fold cross-validation experiment. We used 16 images for training and four images for testing in each fold experiment. The final quantitative results of our experiment were obtained by calculating the mean and standard deviation of the five-fold cross-validation results.

4.3. Evaluation of Results

To prove the effectiveness of our method, we compared it with some previous methods. Fig. 4 shows the qualitative

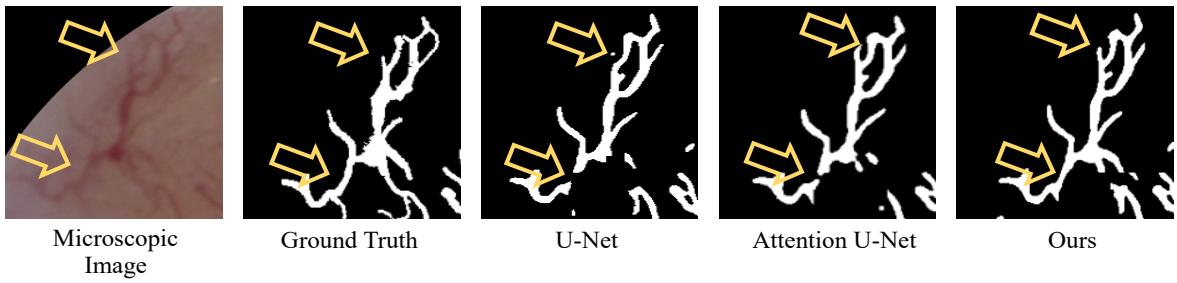

Figure 4. Qualitative results of the U-Net [18], the Attention U-Net [15], and ours. Note that the three networks have identical 4-layers encoder-decoder structure trained with the same loss function clDice($\lambda = 0.1$).

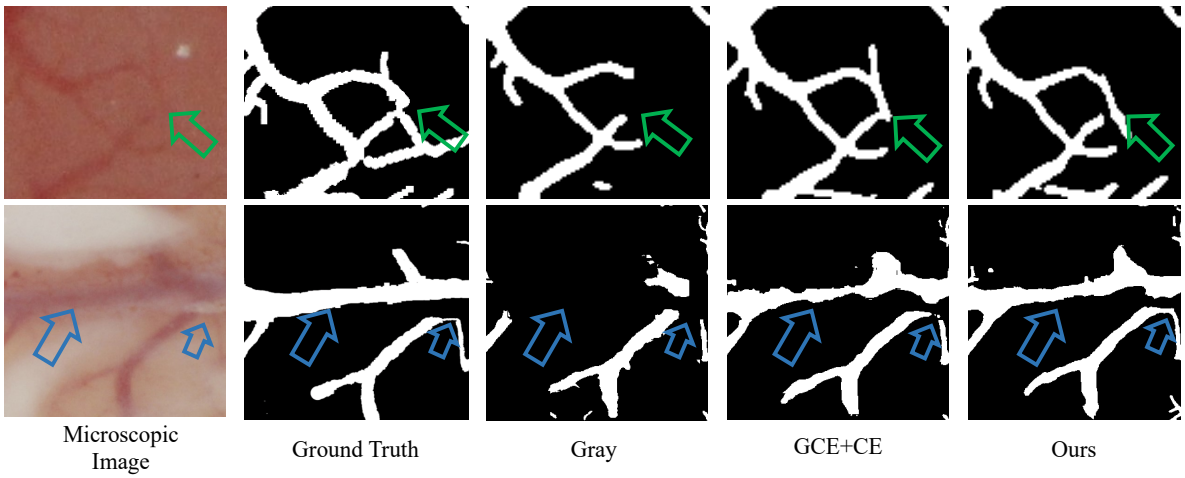

Figure 5. Visualization results for same Attention U-Net with vessel attention and loss function clDice($\lambda = 0.1$) with different preprocessed input images, including gray images, images after preprocessing by green channel extraction and contrast enhancement (GCE+CE), and images after preprocessing by dehazing, gamma correction, denoising, green channel extraction and contrast enhancement. We use green arrows in first row and blue arrows in second row to show segmentation results of small blood vessels and partially occluded blood vessels of each network, respectively.

result of our method on our microscopic image dataset compared to the U-Net and Attention U-Net results. Note that we used a 4-layer encode-decode structure and the clDice loss function for all the networks to compare them. As shown in Fig. 4, we found that the Attention U-Net with attention gate outperformed the U-Net. After further adding vascular information, the Attention U-Net achieved better segmentation results. The quantitative results in Tab. 1 show that after adding the attention gate and vessel attention, the Dice score reached 75.56%, and the clDice score reached 79.95%.

Figure 5 shows the performance of different preprocessing methods when we used the same Attention U-Net framework with vessel attention. From the first row in Fig. 5, we found that compared with the gray images (Gray), the images after preprocessing by green channel extraction and contrast enhancement (GCE+CE) as input images had better results for small blood vessels segmentation in some regions with low contrast (green arrows). From the second row in Fig. 5, we learned that for the partially occluded blood vessels regions, compared with the input images without dehazing (Gray and GCE+CE), the extraction results of the input images with dehazing (Ours) extracted the blood vessels more accurately and better maintained the connectivity of arteries (blue arrows). According to the results shown in Tab. 2, our preprocessing scheme improved the Dice score by 2.02% and the clDice score by 1.54%, compared with a scheme that used gray images without preprocessing.

4.4. Ablation Study

Attention map and feature map analysis. As shown in Fig. 6, we visualized the attention coefficient (β_3) of AG3 in Fig. 2 (c), obtained from test images concerning the training epoch. The blue arrows in Figs. 6 (a) and (b) show that with the increase of training epochs, the model pays more attention to vessel pixels and reduce its concentration on non-vascular pixels. In addition, we also visualized the

Methods	Accuracy ↑	Sensitivity ↑	Specificity ↑	Dice ↑	clDice ↑
U-Net [18]	91.33 (0.61)	**74.48 (2.28)**	95.09 (1.21)	75.30 (2.88)	**79.77 (2.94)**
Attention U-Net [15]	**91.40 (0.70)**	74.23 (2.58)	**95.23 (1.29)**	**75.38 (3.20)**	79.74 (2.87)
Ours	**91.41 (0.88)**	**74.76 (1.74)**	**95.13 (1.40)**	**75.56 (3.21)**	**79.95 (2.88)**

Table 1. Results of Attention U-Net with vessel attention (Ours) and other encode-decode networks (the U-Net and the Attention U-Net) on our dataset. We reported mean (and standard deviation) of five-fold cross-validation experiment for five metrics. **The best** and **the second-best** results are marked in bold red and black, respectively.

Methods	Accuracy ↑	Sensitivity ↑	Specificity ↑	Dice ↑	clDice ↑
Gray	90.69 (0.56)	72.96 (6.55)	**94.76 (1.61)**	73.54 (2.20)	78.41 (2.47)
GCE+CE	**91.16 (0.64)**	**75.67 (5.03)**	94.73 (1.59)	**75.29 (1.48)**	**80.30 (1.61)**
Ours	**91.41 (0.88)**	**74.76 (1.74)**	**95.13 (1.40)**	**75.56 (3.21)**	**79.95 (2.88)**

Table 2. Quantitative results using different preprocessing on our dataset. The Attention U-net with vessel attention is trained using gray images without preprocessing (Gray), gray images with green channel extraction and contrast enhancement (GCE+CE), and the images with dehazing, gamma correction, denoising, green channel extraction and contrast enhancement (Ours). We reported mean (and standard deviation) of five-fold cross-validation experiment for five metrics. **The best** and **the second-best** results are marked in bold red and black, respectively.

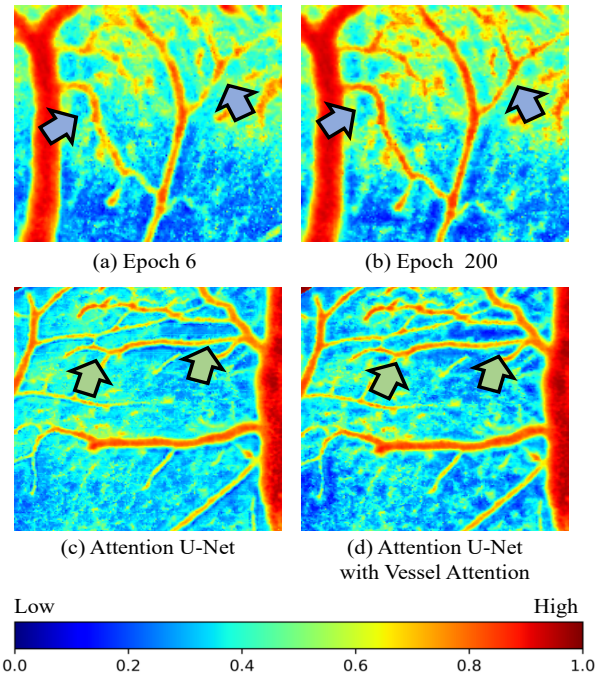

Figure 6. We visualized the attention coefficient β^3 for *6-th* and *200-th* training epochs in (a) and (b), respectively. In (c) and (d), we visualized the attention coefficient β^3 for the Attention U-Net and Attention U-Net with vessel attention in identical *200-th* training epochs. Equivalent colormap is shown at bottom, which indicates that different colors in attention maps represent different attention coefficient values. The more biased to red, the values are higher; and the closer to blue, the values are lower.

attention coefficient to compare Attention U-Net and the Attention U-Net with the vessel attention as shown in Figs. 6 (c) and (d). The green arrows show that the Attention U-Net with the vessel attention not only focuses more on the blood vessel pixels on our dataset but also further reduce the attention to the non-vascular pixels compared with the Attention U-Net.

Color images and green channel images. Finally, as shown in Tab. 3, the RGB images as the segmentation network input have more unsatisfactory performance on our dataset. The reason is that many pixels are not blood vessels, although their color closely resemble that of the blood vessels in the RGB images of the cranial window in our dataset, which leads to segmentation results that are worse than one-channel green images. Moreover, Zhang *et al*. [34] proved gray images can increase the contrast between the blood vessels and the background, conducive to blood vessel segmentation. Therefore, we extracted the green channel image instead of the RGB image in preprocessing.

Loss function analysis. In our work, we used the clDice as a loss function to train the network. To prove that the clDice performs blood vessel segmentation better on our microscopic images dataset, we compared the clDice with the Cross-Entropy (CE) [2] and the Dice Loss Function [13]. As shown in Tab. 4, we found that when using the clDice [23] as the loss function, we obtained 75.56% Dice and 79.95% clDice score, respectively, which are more than using the CE and the Dice loss function. Moreover, compared with the CE, Dice and clDice (λ=0.5) loss function, clDice (λ=0.1) achieved the best connectivity for the segmented blood vessels shown in Fig. 7.

Input Images	Accuracy ↑	Sensitivity ↑	Specificity ↑	Dice ↑	clDice ↑
RGB Images	90.01 (1.49)	71.54 (4.74)	**94.18 (2.97)**	71.81 (4.24)	77.21 (3.73)
Green Channel Images	**90.31 (0.33)**	**73.58 (7.38)**	94.16 (2.20)	**72.84 (2.72)**	**77.81 (3.22)**

Table 3. Quantitative results of three-channel RGB image and one-channel green image as network input. Note that three experiments used identical network structure (Attention U-Net without vessel attention) and loss function (clDice λ=0.1) to evaluate impact of different input images on the results. We reported mean (and standard deviation) of five-fold cross-validation experiment for five metrics. **The best** results are marked in bold red.

Loss Function	Accuracy ↑	Sensitivity ↑	Specificity ↑	Dice ↑	clDice ↑
Cross-Entropy (CE) [2]	**91.47 (0.70)**	72.46 (3.09)	**95.74 (1.10)**	**75.13 (2.74)**	**78.37 (2.12)**
Dice [13]	**91.53 (0.58)**	71.44 (2.32)	**96.02 (0.99)**	75.01 (2.21)	78.17 (2.29)
clDice [23] (λ=0.5)	86.46 (3.67)	**76.79 (6.28)**	88.84 (5.85)	67.29 (5.52)	74.85 (2.80)
clDice [23] (λ=0.1)	91.41 (0.88)	**74.76 (1.74)**	95.13 (1.40)	**75.56 (3.21)**	**79.95 (2.88)**

Table 4. Quantitative results of the Attention U-Net with vessel attention trained with different loss functions. We reported mean (and standard deviation) for five metrics. **The best** and **the second-best** results are marked in bold red and black, respectively.

5. Conclusions

In this paper, we proposed a method using Attention U-Net with vessel attention to segment blood vessels in cranial windows and created a microscopic image dataset for blood vessel segmentation. By segmenting the blood vessels from the wide-field and low-contrast microscopic images, we achieved a simple and low-cost method, which is a novel work for long-term cranial window observation. The Attention U-Net with vessel attention made the Attention U-Net concentrate more on the blood vessel regions, reduced the influence of other non-vascular regions like organoids, and used the clDice to train the model to improve the connectivity of segmented vessels. In addition, for the low-contrast problem, we used the CLAHE method as one preproccessing step to enhance the contrast of the microscopic images. We introduced dehazing to preprocess the images to improve the segmented results of the blood vessels occluded by organoids or dura mater. The experimental results showed that our proposed method achieved excellent performance in blood vessel segmentation from low-contrast and wide-field microscopic images of cranial windows.

However, because a cranial window's radius is only about a few millimeters, some small blood vessel edges in the wide-field microscopic images taken by stereomicroscopes are difficult to observe even by the human eye. Thus, it is not easy to make ground truth. In the future, we will combine high-resolution image information to assist the segmentation task and further improve the segmentation accuracy of small blood vessels and apply it to extract small blood vessels in vascularized organoids. Also, consider the uncertainty of the vascular region for blood vessel segmentation. In addition, we also plan to register the extracted 2D vascular structures with the 3D mouse's brain vascular temple to help the biological researchers better observe the changes of blood vessels when the experiment is implemented in cranial windows.

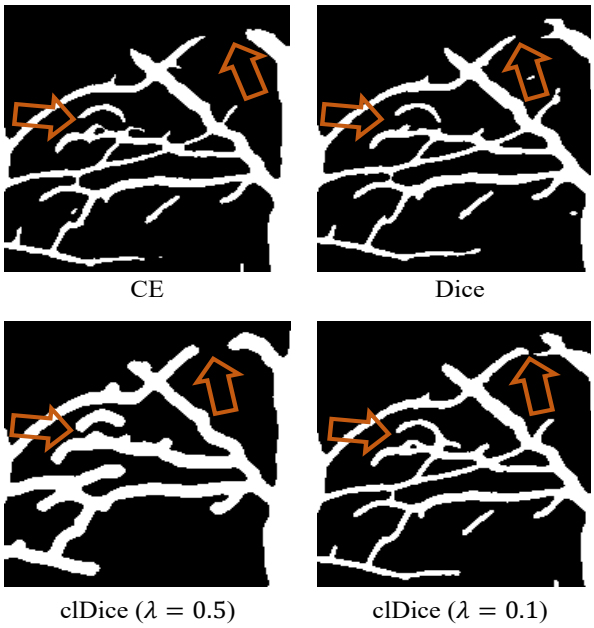

Figure 7. We visualized segmentation results when using identical network trained with different loss functions.

Acknowledgement: Parts of this research were supported by the JST Moonshot R&D Grant Number JPMJMS2033, the JST CREST Grant Number JPMJCR20D5, and the fellowship of Nagoya University CIBoG program from MEXT WISE program.

References

[1] Peter Anderson, Xiaodong He, Chris Buehler, Damien Teney, Mark Johnson, Stephen Gould, and Lei Zhang. Bottom-up and top-down attention for image captioning and visual question answering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 6077–6086, 2018. 2

[2] Pieter-Tjerk De Boer, Dirk P Kroese, Shie Mannor, and Reuven Y Rubinstein. A tutorial on the cross-entropy method. *Annals of operations research*, 134(1):19–67, 2005. 7, 8

[3] Thomas Geuens, Clemens A van Blitterswijk, and Vanessa LS LaPointe. Overcoming kidney organoid challenges for regenerative medicine. *NPJ Regenerative medicine*, 5(1):1–6, 2020. 2

[4] Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353, 2010. 3

[5] Chaejeong Heo, Hyejin Park, Yong-Tae Kim, Eunha Baeg, Yong Ho Kim, Seong-Gi Kim, and Minah Suh. A soft, transparent, freely accessible cranial window for chronic imaging and electrophysiology. *Scientific reports*, 6(1):1–11, 2016. 2

[6] Anthony Holtmaat, Tobias Bonhoeffer, David K Chow, Jyoti Chuckowree, Vincenzo De Paola, Sonja B Hofer, Mark Hübener, Tara Keck, Graham Knott, Wei-Chung A Lee, et al. Long-term, high-resolution imaging in the mouse neocortex through a chronic cranial window. *Nature protocols*, 4(8):1128–1144, 2009. 1

[7] Anthony Holtmaat, Tobias Bonhoeffer, David K Chow, Jyoti Chuckowree, Vincenzo De Paola, Sonja B Hofer, Mark Hübener, Tara Keck, Graham Knott, Wei-Chung A Lee, et al. Long-term, high-resolution imaging in the mouse neocortex through a chronic cranial window. *Nature protocols*, 4(8):1128–1144, 2009. 2

[8] Chaoen Hu, Hui Hui, Shuo Wang, Di Dong, Xia Liu, Xin Yang, and Jie Tian. Cerebral vessels segmentation for light-sheet microscopy image using convolutional neural networks. In *Medical Imaging 2017: Biomedical Applications in Molecular, Structural, and Functional Imaging*, volume 10137, pages 138–143. SPIE, 2017. 3

[9] Juergen C Jung, Amit D Mehta, Emre Aksay, Raymond Stepnoski, and Mark J Schnitzer. In vivo mammalian brain imaging using one-and two-photon fluorescence microendoscopy. *Journal of neurophysiology*, 92(5):3121–3133, 2004. 2

[10] Ruirui Li, Mingming Li, Jiacheng Li, and Yating Zhou. Connection sensitive attention u-net for accurate retinal vessel segmentation. *arXiv preprint arXiv:1903.05558*, 2019. 3

[11] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015. 2, 3

[12] Matthew J McAuliffe, Francois M Lalonde, Delia McGarry, William Gandler, Karl Csaky, and Benes L Trus. Medical image processing, analysis and visualization in clinical research. In *Proceedings 14th IEEE Symposium on Computer-Based Medical Systems. CBMS 2001*, pages 381–386. IEEE, 2001. 5

[13] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)*, pages 565–571. IEEE, 2016. 4, 7, 8

[14] Michael Nercessiany, Nazim Haouchine, Parikshit Juvekar, Sarah Frisken, and Alexandra Golby. Deep cortical vessel segmentation driven by data augmentation with neural image analogy. In *2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI)*, pages 721–724. IEEE, 2021. 3

[15] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y Hammerla, Bernhard Kainz, et al. Attention U-Net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*, 2018. 2, 3, 4, 6, 7

[16] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019. 5

[17] Pavle Prentašić, Morgan Heisler, Zaid Mammo, Sieun Lee, Andrew Merkur, Eduardo Navajas, Mirza Faisal Beg, Marinko Šarunic, and Sven Lončarić. Segmentation of the foveal microvasculature using deep learning networks. *Journal of biomedical optics*, 21(7):075008, 2016. 3

[18] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, pages 234–241. Springer, 2015. 2, 6, 7

[19] Giuliana Rossi, Andrea Manfrin, and Matthias P Lutolf. Progress and potential in organoid research. *Nature Reviews Genetics*, 19(11):671–687, 2018. 1

[20] Kokoro Sakabe, Takanori Takebe, and Akihiro Asai. Organoid medicine in hepatology. *Clinical Liver Disease*, 15(1):3, 2020. 2

[21] Chen Shen, Holger R Roth, Hayashi Yuichiro, Masahiro Oda, Tadaaki Miyamoto, Gen Sato, and Kensaku Mori. Attention-guided pancreatic duct segmentation from abdominal CT volumes. In *Clinical Image-Based Procedures, Distributed and Collaborative Learning, Artificial Intelligence for Combating COVID-19 and Secure and Privacy-Preserving Machine Learning*, pages 46–55. Springer, 2021. 4

[22] Tadahiro Shinozawa, Masaki Kimura, Yuqi Cai, Norikazu Saiki, Yosuke Yoneyama, Rie Ouchi, Hiroyuki Koike, Mari Maezawa, Ran-Ran Zhang, Andrew Dunn, et al. High-fidelity drug-induced liver injury screen using human pluripotent stem cell–derived organoids. *Gastroenterology*, 160(3):831–846, 2021. 2

[23] Suprosanna Shit, Johannes C. Paetzold, Anjany Sekuboyina, Ivan Ezhov, Alexander Unger, Andrey Zhylka, Josien P.W. Pluim, Ulrich Bauer, and Bjoern H. Menze. clDice-a novel topology-preserving loss function for tubular structure segmentation. In *Proceedings of the IEEE/CVF Conference*

on Computer Vision and Pattern Recognition, pages 16560–16569, 2021. 3, 4, 5, 7, 8

[24] Brian Soetikno, Song Hu, Ernie Gonzales, Qiaonan Zhong, Konstantin Maslov, Jin-Moo Lee, and Lihong V Wang. Vessel segmentation analysis of ischemic stroke images acquired with photoacoustic microscopy. In *Photons Plus Ultrasound: Imaging and Sensing 2012*, volume 8223, pages 765–769. SPIE, 2012. 3

[25] Takanori Takebe, Ran-Ran Zhang, Hiroyuki Koike, Masaki Kimura, Emi Yoshizawa, Masahiro Enomura, Naoto Koike, Keisuke Sekine, and Hideki Taniguchi. Generation of a vascularized and functional human liver from an ipsc-derived organ bud transplant. *Nature protocols*, 9(2):396–409, 2014. 2

[26] Mihail Ivilinov Todorov, Johannes Christian Paetzold, Oliver Schoppe, Giles Tetteh, Suprosanna Shit, Velizar Efremov, Katalin Todorov-Völgyi, Marco Düring, Martin Dichgans, Marie Piraud, et al. Machine learning analysis of whole mouse brain vasculature. *Nature methods*, 17(4):442–449, 2020. 3

[27] Carlo Tomasi and Roberto Manduchi. Bilateral filtering for gray and color images. In *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*, pages 839–846. IEEE, 1998. 3

[28] Yohei Tsukada, Fumitaka Muramatsu, Yumiko Hayashi, Chiaki Inagaki, Hang Su, Tomohiro Iba, Hiroyasu Kidoya, and Nobuyuki Takakura. An in vivo model allowing continuous observation of human vascular formation in the same animal over time. *Scientific reports*, 11(1):1–9, 2021. 2

[29] Bea Lyn M Virtudazo, Jimmy Hasugian, Wen-Chen Lin, Mei-Fen Chen, and Kang-Ping Lin. Blood vessel segmentation from microcirculation images. In *International Conference on Biomedical and Health Informatics*, pages 133–137. Springer, 2017. 3

[30] Shuangling Wang, Yilong Yin, Guibao Cao, Benzheng Wei, Yuanjie Zheng, and Gongping Yang. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. *Neurocomputing*, 149:708–717, 2015. 2

[31] Xiancheng Wang, Wei Li, Bingyi Miao, He Jing, Zhangwei Jiang, Wen Xu, Zhenyan Ji, Gu Hong, and Shen Zhaomeng. Retina blood vessel segmentation using a U-Net based Convolutional Nneural Network. In *Procedia Computer Science: International Conference on Data Science (ICDS 2018)*, pages 8–9, 2018. 2

[32] Yunheng Wu, Masahiro Oda, Yuichiro Hayashi, Takanori Takebe, and Kensaku Mori. Investigation on blood vessels segmentation from images of mouse's cranial window based on deep learning. *IEICE Tech. Rep. MI2021-84*, 121(347):174–179, 2021. 2

[33] Yunheng Wu, Masahiro Oda, Yuichiro Hayashi, Takanori Takebe, and Kensaku Mori. Vascular structure segmentation in stereomicroscope image. In *The 40th the Japanese Society of Medical Imaging Technology Annual Meeting 2021*, pages 229–234, 2021. 3

[34] Jingdan Zhang, Yingjie Cui, Wuhan Jiang, and Le Wang. Blood vessel segmentation of retinal images based on neural network. In *International Conference on Image and Graphics*, pages 11–17. Springer, 2015. 3, 7

[35] Chao Zhou, Xiaogang Zhang, and Hua Chen. A new robust method for blood vessel segmentation in retinal fundus images based on weighted line detector and hidden Markov model. *Computer methods and programs in biomedicine*, 187:105231, 2020. 4

[36] Karel Zuiderveld. Contrast limited adaptive histogram equalization. *Graphics gems*, pages 474–485, 1994. 3