Image Anal Stereol 2024;43:295-304 doi: 105566/ias.3055
Original Research Paper

AUTOMATIC ASSESSMENT OF OCT SKIN THICKNESS IN MICE BASED ON TRANSFER LEARNING AND ATTENTION MECHANISMS

CHANGKE WANG[1,2], QIONG MA[1], YU WEI[1,3], NAN YU[1,2], YUQING WANG[1], QINGYU CAI[2,4], HAIYANG SUN[2,4], AND HONGXIANG KANG✉ [1]

[1]Beijing Institute of Radiation Medicine, 27 Taiping Road, Beijing, 100850, China; [2]College of Information Engineering, Henan University of Science and Technology, 263 Kaiyuan Avenue, Luoyang, 471023, China; [3]College of Life Sciences, Hebei University, 180 East Wusi Road, Baoding, 071000, China; [4]Hunan SANY Industrial Vocational Technical College, Hanli Industrial Park, Changsha, 410129, China
e-mail: shuji_wck@163.com; maqiong666@163.com; 15133201723@163.com; 1378313191@qq.com; 1334051981@qq.com; 13875982232@163.com; sunhy@nudt.edu.cn; khx007@163.com
(Received October 10, 2023; revised September 18, 2024; accepted September 18, 2024)

ABSTRACT

Optical coherence tomography (OCT) is characterized by high resolution and noninvasiveness; thus, it has been widely used to analyze skin tissues in recent years. Previous studies have evaluated skin OCT images using traditional algorithms with low accuracy for complex tissue structure images. Although a few studies have used deep learning methods to assess tissue structure in OCT images, they lack quantitative assessment of deeper skin tissue thickness and are limited to the epidermal layer. Thus, in the present study, we proposed an automated segmentation and quantitative evaluation method. The skin OCT images were first pre-processed, and the attention mechanism was added to U-Net based on transfer learning to segment the images and quantify the thickness of mouse skin tissue structure. The results showed that U-Net combined with the coordinate attention (CA) mechanism had better segmentation performance with 93.94% mean intersection of union (MIoU) value and 96.99% Dice similarity coefficient; the segmentation errors were 0.6 μm, 2.2 μm, 3.8 μm, and 6.0 μm for the epidermis layer, subcutaneous fat layer, muscle fascia layer, and the overall skin tissue structure of mice, respectively. The overall skin tissue thickness of the four mice were 235 ± 20 μm, 264 ± 42 μm, 275 ± 40 μm, and 774 ± 91 μm, respectively. The present study provides a rapid and accurate method for the automated measurement of skin tissue thickness.

Keywords: Deep learning; Image segmentation; Optical coherence tomography; U-Net.

INTRODUCTION

Optical coherence tomography (OCT) (Huang *et al*., 1991) is a noninvasive, high-resolution biomicroscopic imaging technique that acquires high-resolution (micron level) three-dimensional (3D) cross-sectional images inside biological tissues (Tadrous, 2000). Thus, it has wide applications in biomedical ophthalmology (Pan and Chen, 2023), oncology (Sacha *et al*., 2020), and dermatology (Kim *et al*., 2012). In the past two decades, clinical attempts have been made to evaluate normal skin tissue structures, including the epidermis, dermis, hair follicles, and sweat glands, using OCT (Pathania *et al*., 2022).

Investigators in dermatology have used traditional noninvasive optical imaging methods such as decision trees and edge detection algorithms to segment and classify lesion regions in skin OCT images (Bonne *et al*., 2015; Cobb *et al*., 2006; Gao *et al*., 2016). However, the conventional methods used for such analyses are not accurate in evaluating the OCT images of structurally complex skin and cannot assess deeper areas of imaging depth. Additionally, the traditional methods are time-consuming and cannot meet the needs of real-time evaluation in clinical settings.

Therefore, to address the above limitations, Gao et al. (Gao *et al*., 2022) used U-Net to segment the three layers of the mouse skin tissue structure in OCT images and quantify the volume of laser-damaged areas. Keep et al. (Keep *et al*., 2019) added dense linked blocks to U-Net for segmenting the skin tissue structures in mouse skin OCT. Strikingly, the deep learning method segmented more accurately with higher Dice similarity coefficients than the previously proposed random forest RF+GC-based algorithm. Ji *et al.* (Ji *et al*., 2022) used multiple deep-learning models to segment the epidermal and scabbing areas in OCT mouse skin images, based on which epidermal and crustal region thicknesses were

calculated during healing in rodent skin injury models. Nonetheless, these studies were limited to the structural segmentation of skin tissue or thickness measurements of the epidermal layer and did not involve quantitative analysis of deeper skin layers. Also, the studies used supervised learning methods that did not facilitate obtaining a large number of high-quality labeled biological images, which might have affected the accuracy of subsequent qualitative and quantitative analyses. Transfer learning (Zhuang *et al*., 2021) can achieve feature migration and parameter sharing, can improve the generalization ability of the model, and has been proven an effective model in several studies (Xu *et al*., 2022; Abubakar *et al*., 2020). Specifically, it can obtain improved training results from small sample training data and, hence, has been widely used in image processing in medicine and biology (Shahoveisi *et al*., 2023). Attentional mechanisms (Guo *et al*., 2022) mimic the human visual and cognitive systems and are widely used in natural language processing and computer vision (Li *et al*., 2020). The introduction of an attention mechanism facilitates automatic learning of neural networks and selective focus on vital information in the input, thereby improving the performance and generalization ability of the model. It also performs excellently in extracting information about lesions from pathological images (Gong *et al*., 2023; Chen *et al*., 2023).

The present study used U-Net and U-Net with an added attention mechanism to segment the epidermis layer (EDL), subcutaneous fat layer (SFL), and muscle fascia layer (FML) in the skin tissue structure of mice based on transfer learning and to quantify the thickness of the three tissue structure layers with a segmentation quantization depth of up to 800 μm. This approach will have potential applications for noninvasive skin testing and automated qualitative and quantitative assessment of the tissue structures.

MATERIALS AND METHODS

In this study, the image was first preprocessed, based on which network models were used for training, and the images were segmented using the obtained model. Finally, the thickness of the three-layer mouse skin structure was calculated and analyzed quantitatively using the thickness calculation algorithm. The process is illustrated in Fig 1.

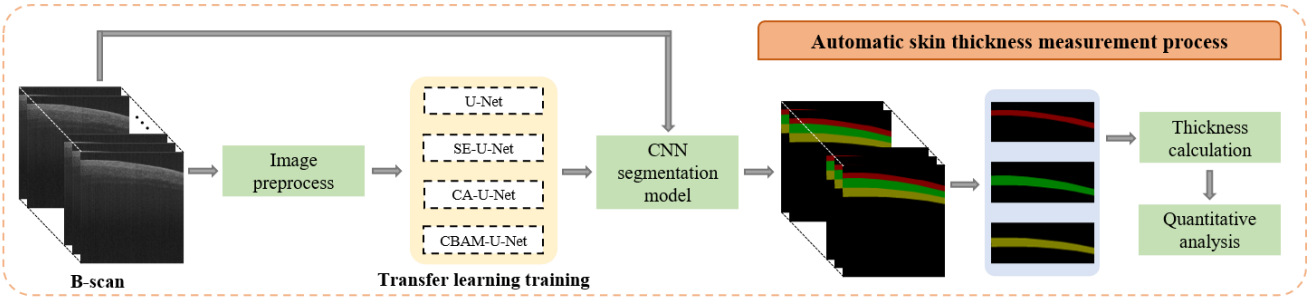

Fig. 1. *Flow of three-layer structure thickness assessment of mouse skin tissue*

EXPERIMENTAL ANIMALS AND SYSTEM CONFIGURATION

A total of four specific pathogen free grade Kumming mice (two males and two females; named samples 1–4), weighing 20–25 g, were purchased from Beijing Keyu Animal Breeding Center (Beijing, China). The animals were routinely housed in the Experimental Animal Center of Military Medical Research Institute (Beijing, China) for 3 days before image acquisition and were found to be free of abnormalities before the experiments. For pre-image acquisition by OCT (OptoMedic Company, Model No.LVM-1000; Guangzhou, China), mice were anesthetized intraperitoneally with 1% pentobarbital sodium and shaved (dorsal skin was depilated using depilatory creams, and the skin was cleaned with saline three times to remove maximal depilatory agent). All experiments were approved by the Animal Ethics Committee of The Beijing Institute of Radiation Medicine (Beijing, China) and conducted according to the guidelines of the IACUC-DWZX-2019-502.

In this study, a swept-source OCT imaging system with a central wavelength of 1,310 nm, a main frequency scan of 100 kHz, and spectral broadband of 91.5 nm was used to scan and image the skin area on the back of the mice; a longitudinal resolution of 22 μm and a lateral resolution of 12 μm was applied to generate 200 images per scan cycle.

DATASETS AND PREPROCESSING

A total of 50 OCT images were acquired from each mouse at different locations of the dorsal skin, and each 3D image block contained 200 B-scan images 460 ×

500 in size. 5/200 B-scan images were randomly selected from each OCT image block for manual annotation of three specialized physicians with two days of training. They identified the boundaries between the different skin layers by comparing the OCT B-scan images with the hematoxylin-eosin (HE)-stained images, followed by annotation. The labeling software was Labelme (MIT Computer Science and Artificial Intelligence Laboratory, version number 5.1.1). The image dataset comprised 1,000 B-scan images (each 460 × 500) that were divided into training, validation, and test sets in the ratio of 7:2:1. Each image was resized before training to 512 × 512.

To highlight the boundary of each layer in the image and obtain an enhanced segmentation effect, we combined non-local mean filtering and bilateral filtering for noise reduction on B-scan images. The radius of the non-local mean filter neighborhood window was 2, the radius of the search window was 5, the smoothing parameter of the Gaussian function was 10, the spatial distance parameter of the bilateral filtering was 5, the range of the color difference was, and the sigma value of the coordinate space was 150. In order to prevent overfitting, the image was data-augmented based on noise reduction in the original B-scan by flipping the image horizontally with a random rotation, a probability of 50%, and an angle rotation of ±15°.

DEEP LEARNING ALGORITHM

U-Net was selected as the base network for optimal performance in semantic segmentation. In order to improve the training efficiency and model segmentation accuracy, an attention mechanism was added to the U-Net network for training. Unlike Gao et al., we used migration training to improve the training efficiency, and two new attention mechanisms, CBAM and CA, were added to improve the segmentation accuracy of the network model.

Network structure

Furthermore, to improve the segmentation accuracy of the model, we added attention modules in the feature map connection part of the U-Net decoder section. The network model is depicted in Fig 2.

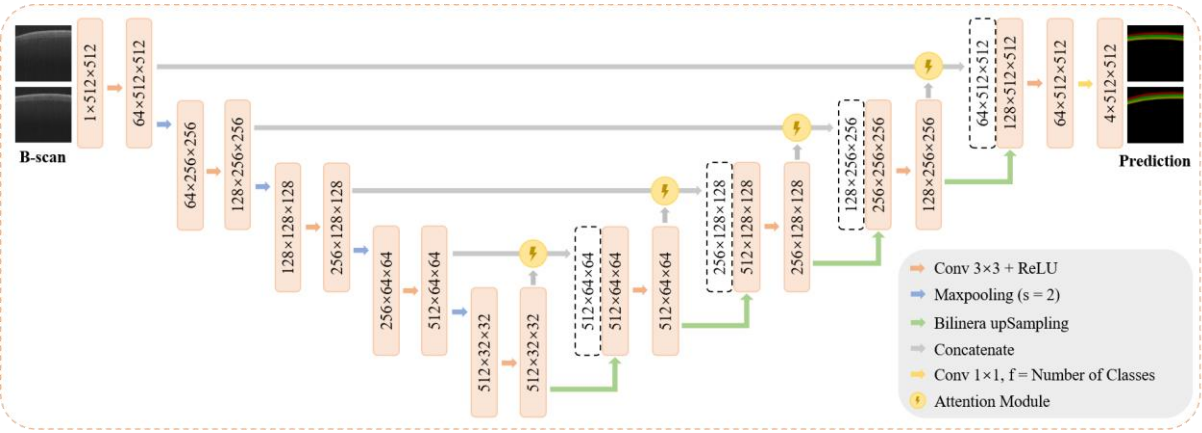

Fig. 2. *U-Net with the attention module.*

The selected attention mechanisms were squeeze-and-excitation attention (SE) (Geoffrey *et al.*, 2015), coordinate attention (CA) (Hou *et al.*, 2021), and convolutional block attention module (CBAM) (Woo *et al.*, 2018). The attention mechanism structure is shown in Fig 3.

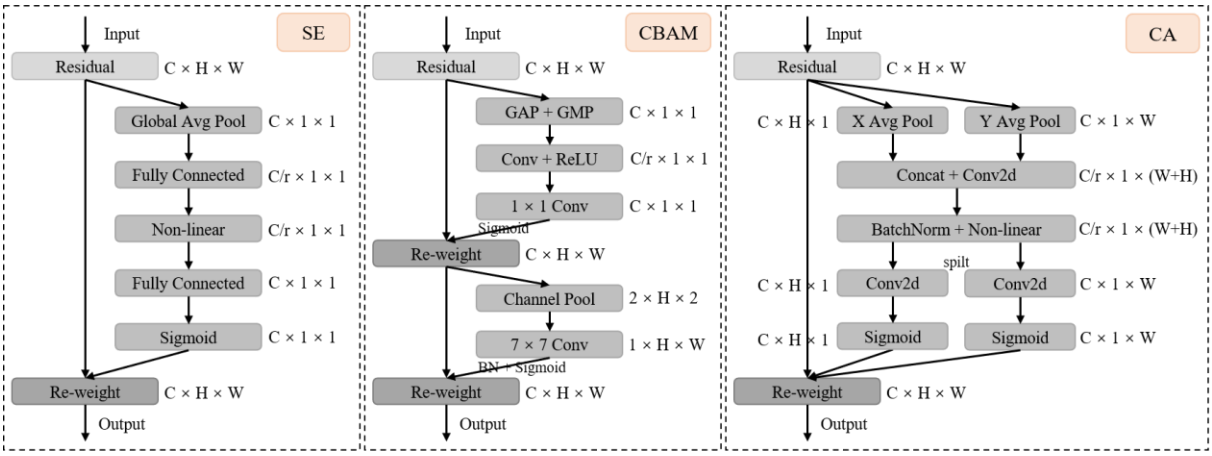

Fig. 3. *SE, CBAM, CA attention mechanism structure.*

Evaluation metrics

Three different evaluation metrics were used to assess the accuracy of model segmentation quantitatively: the mean intersection of union (MIoU), mean pixel accuracy (MPA), and precision.

MIoU is a standard measure of semantic partitioning to calculate the difference between the intersection and the union of the two sets of predicted and formal labels within each class, followed by the average of all classes. MIoU can be calculated as in Eq. (1):

$$MIoU = \frac{1}{k+1}\sum_{i=0}^{k}\frac{TP}{FN+FP+TP} \tag{1}$$

where k is the number of images to be classified in the image, in this paper the value of k is 3, TP represents true-positives, indicating the number of positive samples that the model correctly predicts as positive; FN represents false-negatives, indicating the number of positive samples that the model incorrectly predicts as negative, i.e., the number of positive samples that the model fails to identify correctly; FP represents false-positives, indicating the number of negative samples that the model incorrectly predicts as positive, i.e., the number of negative samples that the model incorrectly identifies as positive. MPA is used to calculate the proportion of correctly classified pixels within each class. Then, the average of all classes is expressed as in Eq. (2):

$$MPA = \frac{1}{k+1}\sum_{i=0}^{k}\frac{TP}{TP+FP} \tag{2}$$

Precision indicates the proportion of samples with positive predicted outcomes that are actually positive and can be expressed as in Eq. (3):

$$precision = \frac{TP}{TP+FP} \tag{3}$$

Training

In this study, the official pretraining model ResNet50 provided by PyTorch was used in the backbone network part of the U-Net model. The network models were trained using Adam optimizer and the cross-entropy loss function. The exponential decay rates of β1 and β2 of Adam were 0.9 and 0.999, respectively, and the initial learning rate was 0.0001. The learning rate was updated using the cosine learning rate decline strategy; the minimum learning rate was 0.01 times the initial learning rate. Each network model was trained for 200 epochs, with a batch size of 8 sets.

Herein, the training and testing of the network were conducted using Python 3.8, PyTorch 1.11, and CUDA 12.0.89 on an Intel Xeon w-2255 CPU, Nvidia RTX A5000.

SKIN LAYER THICKNESS MEASUREMENT METHOD

Next, we calculated the area of different skin layers according to Eq. (4). The percentage of pixel points corresponding to that layer in the whole image was measured, and then the average thickness of different layers in the B-scan image was determined.

$$L = \frac{S_{B-scan}\cdot\sum_{i=0}^{n} I(l_i)}{N_{B-Scan}\cdot W} \tag{4}$$

where $S_{B\text{-}scan}$ is the area of the B-scan image, the size is 0.5 cm²; the I(li) function is used to determine the class of pixels in the segmented image, and the total number of pixels of that class is obtained by accumulation; $N_{B\text{-}scan}$ is the number of pixels of the B-scan image with size 230,000; W is the width of the image to be calculated.

RESULTS

IMAGE PREPROCESSING RESULTS

Scattering noise in OCT images reduces the resolution and contrast of the image, masks the image details, and blurs the image (Fig 4(a)). The effect of speckle noise in the images on the training effect of the model was reduced using non-local mean filtering and bilateral filtering. After non-local mean filtering, the visual effect of the image was less grainy, and the noise was suppressed, as shown in Fig 4(c). In contrast, Fig 4(d) shows

the results of secondary denoising using bilateral filtering based on non-local mean filtering. The scattered noise in the filtered image was suppressed, which clarified the image and, thus, the layers. Fig 4(b) shows the labeled image: red indicates EDL, green indicates SFL, and yellow indicates FML.

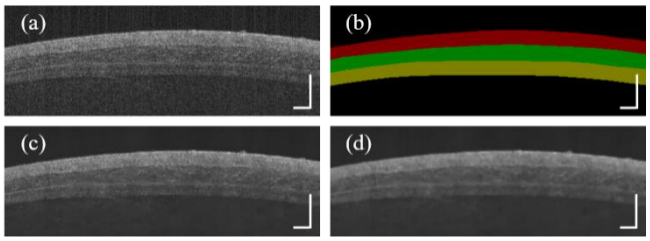

Fig. 4. *B-scan image denoising effect. a. Original image; b. Artificially labeled image: red indicates EDL, green indicates SFL, and yellow indicates FML; c. Non-local mean filter; d. Non-local mean filter and bilateral filter. Scale bar is 500 μm.*

EFFECTIVENESS OF TRANSFER LEARNING

To verify the effectiveness of transfer learning, we compared the validation loss of the four network models with and without transfer learning (Fig 5). The results showed that in the early stage of training, the model with migration training adapts better to the semantic segmentation task after adjusting the parameters and converges faster than the model without migration training. In addition, the model with migration training is more stable throughout the first and middle stages of training, with little fluctuation in validation loss than that without migration training. Finally, all four models showed a decrease in validation loss after using transfer learning, thus confirming that transfer learning can accelerate the convergence speed of the model and improve its performance.

QUALITATIVE ANALYSIS OF NETWORK MODEL SEGMENTATION PERFORMANCE

The trained models were used to segment the images in the test set, and the segmentation results of the four different network models were visualized and compared with the manually labeled images for qualitative analysis. Subsequently, three representative B-scan images in the test set were compared to assess the segmentation effects (Fig 6).

In terms of the overall structure, the segmentation results of the four deep learning models showed high visual consistency with the manual segmentation results. Although all the models accurately segmented the three layers of the mouse skin tissue structure, the segmentation results of different models did not differ markedly.

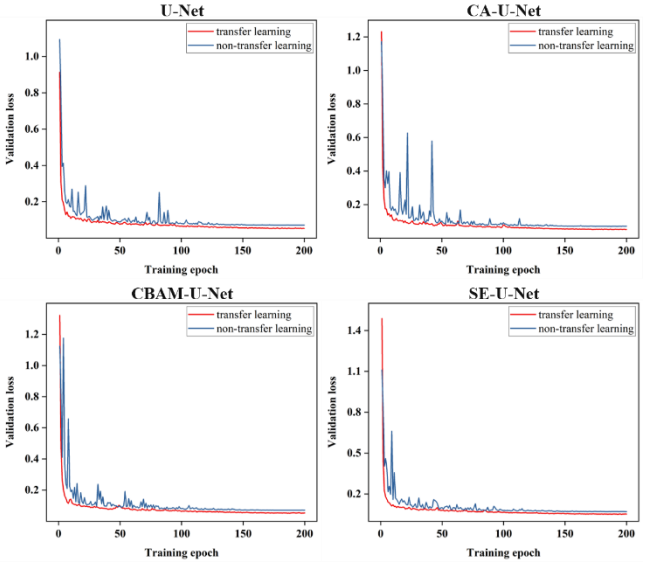

Fig. 5. *Validation loss curves for whether the four models use transfer learning.*

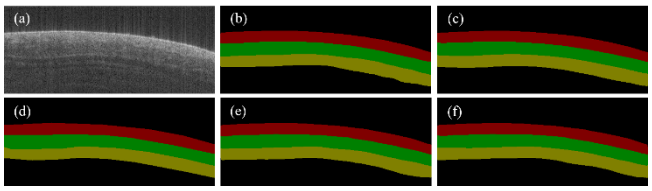

Fig. 6. *Visual comparison of the segmentation results of the four models. a-f are the B-Scan grey-scale map, manual labelling results, U-Net, CA-U-Net, CBAM-U-Net and SE-U-Net segmentation results respectively.*

Based on Fig 6, we extracted the contours of the segmentation results to further compare the segmentation differences between the different models. As shown in Fig. 7, the segmentation results of U-Net on edges with the addition of the attention mechanism have a better overlap between the segmentation results and the real values, with higher accuracy.CA-U-Net is smoother in visual effect, and the segmentation results on SFL and FML are also better than those of CBAM-U-Net and SE-U-Net. Another aspect, segmenting the EDL and SFL was better than segmenting the FML. This might be because the FML is deeper than the EDL and FML and is more affected by the attenuation of light signals and blurred boundary information, resulting in greater errors in segmentation results.

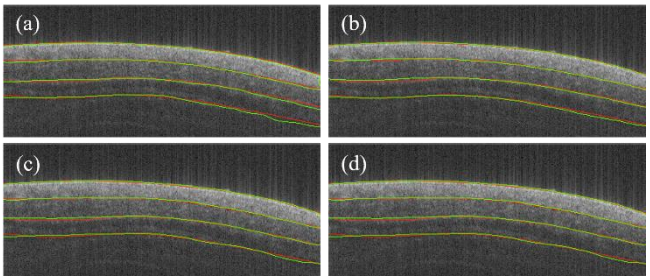

Fig. 7. *Visual comparison of the segmentation results of the four models. The red and green lines are the three-layer organisational structure outlines of manual annotation and deep learning model segmentation, respectively. a-d are the segmentation results of U-Net, CA-U-Net, CBAM-U-Net and SE-U-Net, respectively.*

QUANTITATIVE ANALYSIS OF NETWORK MODEL SEGMENTATION PERFORMANCE

Next, to compare the network performance quantitatively, the prediction results of the network models on the test set were evaluated using manually labeled images; the evaluation metrics were MIoU, MPA, and precision.

As shown in Table 1, the three U-Nets incorporating the attention mechanism showed improved segmentation performance on different skin tissue structures compared to the original U-Net, proving the effectiveness of the attention mechanism. Taken together, CA-U-Net had the best performance with MIoU values of 93.94%, 91.29%, and 89.74% on EDL, SFL, and FML, respectively, and precision values of 96.86% and 94.21% on EDL and FML, respectively. However, the MPA value on EDL and precision value on SFL were lower at 96.99% and 95.46%, respectively, for SE-U-Net. According to EDL, SFL, and FML evaluation indices, the deeper the tissue structure, the lower the evaluation index. The phenomenon might be attributed to the lighter attenuation effect in OCT imaging; the deeper the imaging depth, the stronger the light attenuation, and the worse the imaging quality, resulting in poor segmentation results.

Table 1. *Average evaluation metrics of network models on the test set: MIoU, MPA and Precision(Bold highlights the best results)*

Parameters	Models	EDL	SFL	FML
MIoU/%	U-Net	92.77	90.04	88.67
	SE-U-Net	93.57	90.97	89.34
	CA-U-Net	**93.94**	**91.29**	**89.74**
	CBAM-U-Net	93.73	91.08	89.29
MPA/%	U-Net	95.49	94.44	94.11
	SE-U-Net	**96.99**	95.01	94.72
	CA-U-Net	96.90	**95.58**	**95.27**
	CBAM-U-Net	96.33	95.19	94.89
Precision/%	U-Net	95.87	94.17	93.00
	SE-U-Net	96.38	**95.46**	93.95
	CA-U-Net	**96.86**	95.31	**94.21**
	CBAM-U-Net	96.51	95.39	94.05

In order to analyze and observe the improvement of U-Net on different evaluation metrics intuitively after the addition of the attention mechanism, box-and-line plots were used to analyze the results of the four models on the test set, as shown in Fig 8. The results showed that the U-Net with added attention mechanism improved and stabilized the prediction results. However, the CA-U-Net with the CA attention mechanism added displayed the best prediction results.

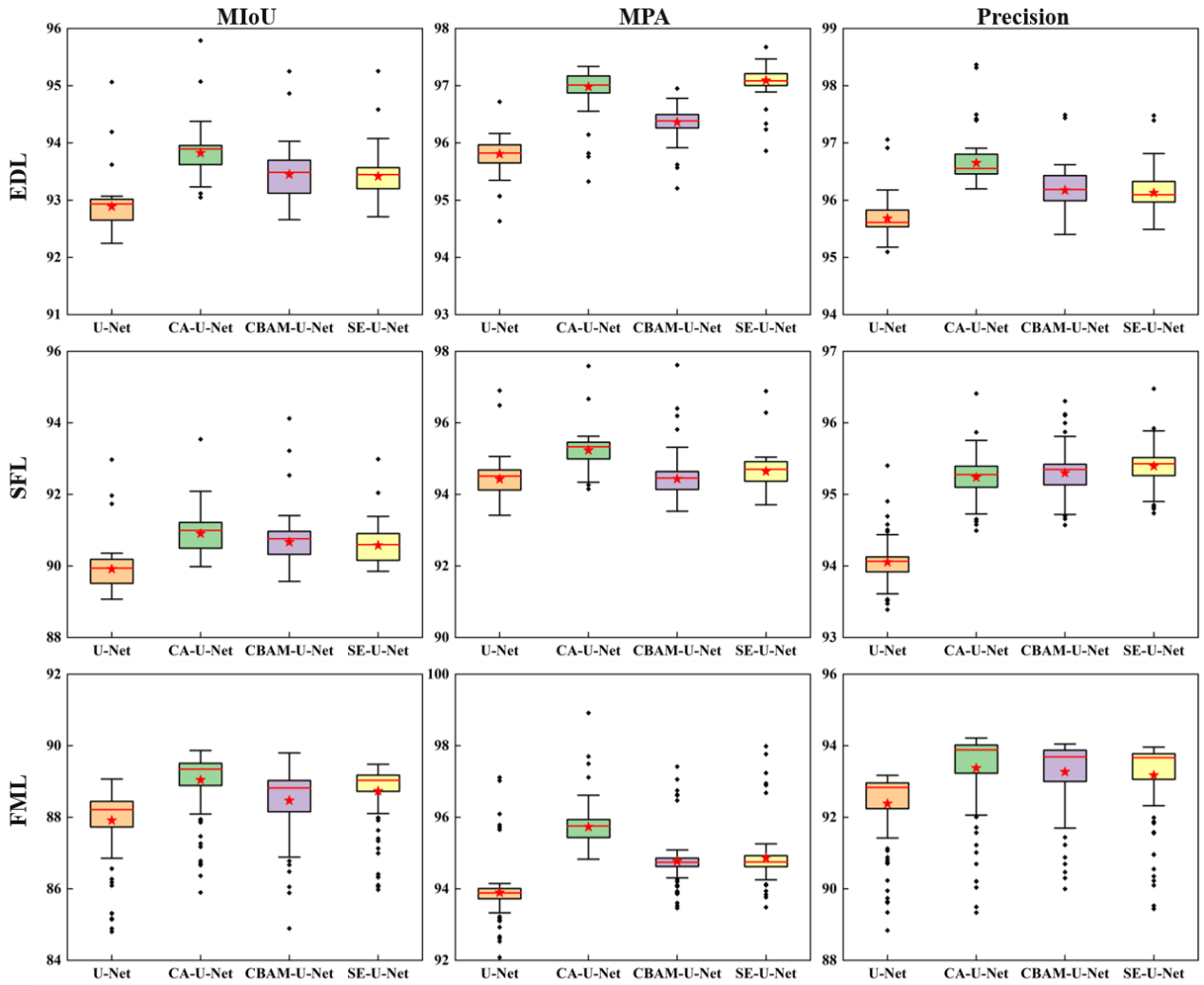

Fig. 8. *Box line plots of quantitative assessment results for the four models. SE, CA, and CBAM for U-Net with the addition of the three attention modules.*

QUANTITATIVE ASSESSMENT OF SKIN LAYER THICKNESS IN MICE

To verify the scientific validity and effectiveness of the deep learning method, we quantified different skin layers of four mice and compared them with manually labeled images using the CA-U-Net deep learning model. The three tissue layers were quantitatively evaluated using the thickness calculation method proposed in the Methods section: skin layer thickness measurement.

Fig 9(a) shows the thickness and average thickness of the EDL of the four mice. Compared to the labeled image, the average thickness segmentation error of CA-U-Net in the four mice was 0.6 μm, proving its highest segmentation accuracy for EDL. Nonetheless, the imaging quality decreased with increasing imaging depth (Fig 9(b),(c)). The prediction errors of CA-U-Net for SFL and FML increased gradually, and the average thickness prediction errors of the four mice were 2.2 μm and 3.8 μm, respectively. Fig 9(d) shows the overall thickness of the mouse skin tissue structure, and the average thickness segmentation error of CA-U-Net for four mice was 6.0 μm. The comparison of the quantified skin thickness results of the four mice revealed that the segmentation errors of the EDL were within 1 μm and did not differ significantly. Also, the prediction errors of Samples 1, 3, and 4 were not significantly different from the mean prediction errors for SFL, MFL, and overall skin tissue thickness in mice. On the other hand, the prediction errors of Sample 2 were 2.6 μm, 3.4 μm, and 5.7 μm for SFL, FML, and overall skin tissue thickness in mice, respectively. These errors were larger than the prediction errors of other mice and the average prediction errors because of the lack of curved skin tissue structure images in Sample 2. The mean thicknesses of the EDL, SFL, FML, and skin tissue as a whole in the four mice were 235 ± 20 μm, 264 ± 42 μm, 275±40 μm, and 774±91 μm, respectively.

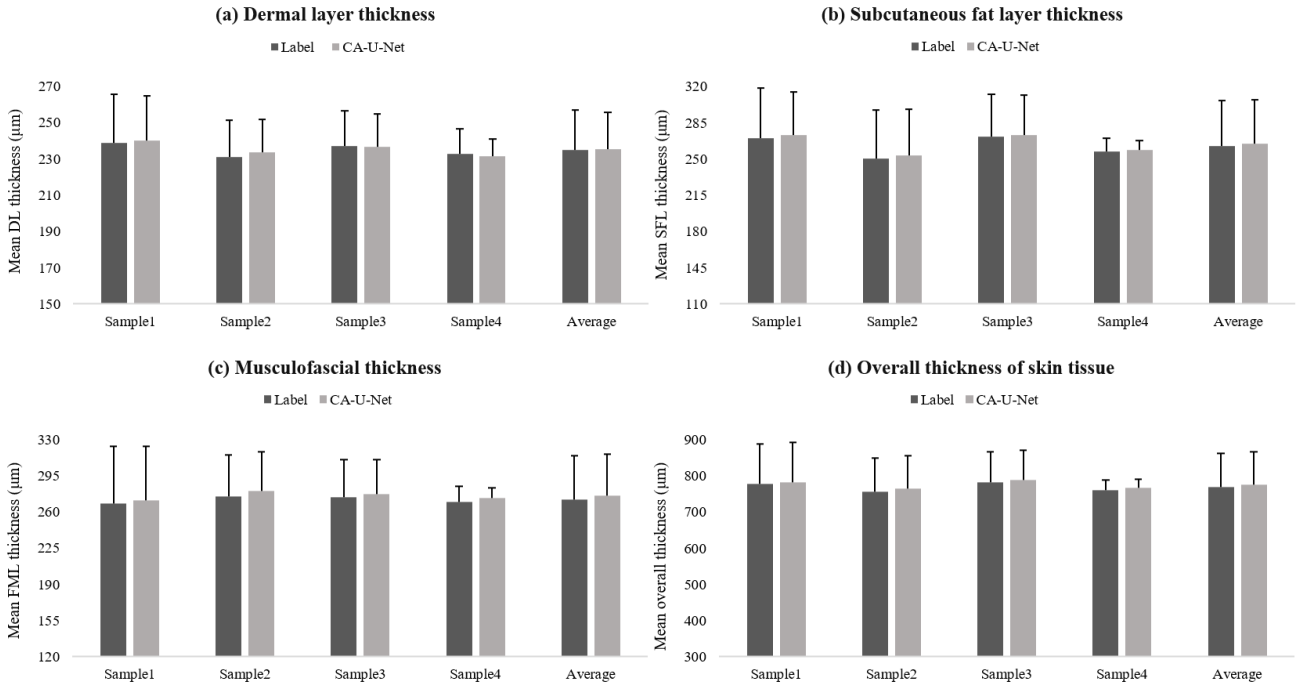

Fig. 9. *Structure and overall thickness of the three layers of mouse skin. Average thickness of labeled images with CA-U-Net predicted the average skin thickness in four mice; (a) is the average thickness of EDL, (b) is the average thickness of SFL, (c) is the average thickness of FML, and (d) is the average thickness of overall skin tissue structure. Samples 1–4 represent four mice. Average indicates the average thickness of skin tissue in four mice. Error bars indicate standard deviation; t-test was used to verify the differences between CA-U-Net prediction results and labeled images in each group, and the p-values were > 0.05. No significant difference was detected between CA-U-Net prediction results and labeled images in each group.*

DISCUSSION

In recent years, deep learning methods have been applied gradually in the skin OCT. In this study, we used U-Net and U-Net with attention mechanism, based on transfer learning, for accurate segmentation of the three-layer structure of mouse skin tissue. Herein, a method for automatic segmentation and quantitative assessment of the thickness of the three-layer structure of mouse skin tissue is proposed. The quantification of the skin tissue structure thickness deemed the CA-U-Net model segmentation most effective as the prediction error of the mouse skin three-layer tissue structure thickness was within 7 μm.

In the present study, we first combined non-local mean filtering with bilateral filtering for noise reduction in B-scan images, distinguished the layer boundaries of the mouse skin tissue, and improved the segmentation accuracy of the network model to obtain accurate results for subsequent quantification of skin tissue thickness. To prevent overfitting, data augmentation was performed using flip and random rotation, followed by training using U-Net and U-Net with different attention mechanisms (SE, CA, and CBAM). The quantitative results indicated that the U-Net incorporating all three attention mechanisms performed better than the base U-Net, improving all performance metrics (Table 1, Fig 8). Among these, the CA-U-Net with CA attention mechanism achieved the best segmentation performance with MIoU and precision of 93.94% and 96.86%, respectively, better than the other models. This finding could be attributed to the fact that the CA attention mechanism not only acquires the inter-channel information but also considers the direction-related position information, which can simultaneously consider the inter-channel relationship and long-distance position information, thereby enhancing the feature extraction. In a previous study, Gao et al. estimated the Dice values of the improved model for the three layers of skin segmentation as 0.93, 0.84, and 0.86, respectively, which are lower than our evaluation metrics, indicating higher segmentation accuracy in this study. The prediction effects of the four network models were visualized and compared with respect to the images with different thicknesses and degrees of curvature of the skin tissue structures. All four

models accurately segmented the three-layer structure of the mouse skin tissue, proving the scientific effectiveness of the deep learning method. Then, based on the segmentation effect of the three structural layers, we deduced that due to the relatively shallow imaging depth of the EDL, the imaging quality was higher, the segmentation accuracy was higher than that of the SFL and the FML, and the overlap with the labeled image was also higher. The upper and lower boundaries of the EDL were visually smoother and closest to the label image, and with increasing imaging depth, the segmentation accuracy of SFL and FML was lower than that of EDL.

Furthermore, B-scan image data from the skin of four mice were input into the CA-U-Net model with the best segmentation performance for prediction, and the thickness of the predicted images was calculated using the thickness calculation algorithm. Compared to the labeled image thickness results, CA-U-Net segmented the EDL with high accuracy, small prediction error, and an average prediction error of 0.6 μm in all four mice. However, due to the deeper imaging depth of the SFL and the FML, the boundary information was more blurred than that of the EDL, resulting in a larger prediction error in this region, with an average prediction error of 2.2 μm *vs*. 3.8 μm on the four mice and an error of 6.0 μm in the overall skin tissue structure thickness of the mice. The mean thicknesses of the EDL, SFL, DFL, and overall skin tissue of the four mice were 235 ± 20 μm, 264 ± 42 μm, 275 ± 40 μm, and 774 ± 91 μm, respectively. T-test analysis of the CA-U-Net prediction results compared to the labeled images revealed that the *p*-values of the three-layer structure and the overall structure of each sample were > 0.05, i.e., no significant difference, proving the scientific validity and effectiveness of the deep learning method in this task. Taken together, the quantitative assessment results of skin tissue thickness in mice indicated that the deep learning method has high accuracy and small error. Hence, it can be applied to measure skin thickness and matrix quality in esthetic medicine and skin injury treatment.

Nevertheless, the present study has some limitations. First, the low resolution of the OCT device led to scattered noise in the images, blurring the boundary information in the mouse skin tissue structure, affecting the subsequent network segmentation performance, and causing errors in the assessment of the skin layer thickness. Second, although CA-U-Net shows better performance than the other three network models, the segmentation effect can be improved further. Therefore, future studies will focus on improving the model, increasing its segmentation accuracy, and extending the current results to segmenting and assessing damaged skin tissue structures in mice to enhance the application prospects of the method.

REFERENCES

Abubakar A, Ugail H, Bukar AM (2020). Assessment of human skin burns: A deep transfer learning approach. J Med Biol Eng 40: 321-33.

Boone M, Marneffe A, Suppa M, Miyamoto M, Alarcon I, Wellenhof RH, Malvehy J, Pellacani G, Marmol VD (2015). High-definition optical coherence tomography algorithm for the discrimination of actinic keratosis from normal skin and from squamous cell carcinoma. J Eur Acad Dermatol 29: 1606-15.

Chen X, Cai Q, Ma N, Li HS (2023). ChroSegNet: An attention-based model for chromosome segmentation with enhanced processing. Appl Sci 13: 2308.

Cobb MJ, Chen Y, Underwood RA, Usui ML, Olerud J, Li X (2006). Marmol Noninvasive assessment of cutaneous wound healing using ultrahigh-resolution optical coherence tomography. J Biomed Opt 11: 064002.

Gao TX, Liu S, Gao E, Wang AC, Tang XY, Fan YW (2022). Automatic segmentation of laser-induced injury OCT images based on a deep neural network model. Int J Mol Sci 23: 11079-91.

Gao W, Zakharov VP, Myakinin O, Bratchenko IA, Artemyev DN, Kornilin DV (2016). Medical images classification for skin cancer using quantitative image features with optical coherence tomography. J Innov Opt Heal Sci 9: 1650003.

Geoffrey H, Vinyals O, Dean J (2015). Distilling the Knowledge in a Neural Network. arXiv:1503.02531.

Gong P, Cheng L, Zhang Z, Meng A, Li E, Chen J, Zhang L (2023). Multi-omics integration method based on attention deep learning network for biomedical data classification. Comput Meth Prog Bio 231: 107377.

Guo MH, Xu TX, Liu J, Liu ZN, Jiang PT, Mu TJ, Zhang SH, Martin R, Chen M, Hu SM (2022). Attention mechanisms in computer vision: A survey. Comput Vis Media 8: 331-68.

Hou QB, Zhou D, Feng J. Coordinate Attention for Efficient Mobile Network Design (2021). in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) 13708-17.

Huang D, Swanson EA, Lin CP, Stinson WG, Chang W, Hee GR, Flotte T, Gregory K, Puliafito CA (1991). Optical coherence tomography. Science 254: 1178-81.

Ji Y, Yang S, Zhou K, Rocliffe H, Pellicoro A, Cash JL, Wang R, Li C, Huang Z (2022). Deep-learning approach for automated thickness measurement of epithelial tissue and scab using optical coherence tomography. J Biomed Opt 27: 015002.

Kepp T, Droigk C, Casper M, Evers M, Hüttmann G, Salma N, Manstein D, Heinrich MP, Handels H (2019). Segmentation of mouse skin layers in optical coherence tomography image data using deep convolutional neural networks. Biomed Opt Express 10: 3484-96.

Kim KH, Pierce MC, Maguluri GN, Park BH, Yoon SJ, Lydon M, Sheridan R, Boer JF (2012). In vivo imaging of human burn injuries with polarization-sensitive optical coherence tomography. J Biomed Opt 17: 066012.

Li X, Jiang Y, Li M, Yin S (2020). Lightweight attention convolutional neural network for retinal vessel image segmentation. IEEE Trans Industr Inform 17: 1958-67.

Pan L, Chen X (2023). Retinal OCT image registration: Methods and applications. IEEE Rev Biomed Eng 16: 307-18.

Pathania YS, Apalla Z, Salerni G, Patil A, Grabbe S, Goldust M (2022). Non-invasive diagnostic techniques in pigmentary skin disorders and skin cancer. J Cosmet Dermatol 21: 444-50.

Sacha NS, Arnaud M, Celia M, Stephanie B, Sandra L, Jean PC (2020). Swept-source and spectral domain OCT imaging of conjunctival tumors. Ophthalmology 128: 947-50.

Shahoveisi F, Gorji HT, Shahabi S, Hosseinirad S, Markell S, Markell S, Vasefi F (2023). Application of image processing and transfer learning for the detection of rust disease. Sci Rep 13: 5133.

Tadrous PT (2000). Methods for imaging the structure and function of living tissues and cells: optical coherence tomography. J Pathology 191: 115-9.

Woo SH, Park JC, Lee JY, Kweon IS (2018). CBAM: Convolutional Block Attention Module. arXiv: 1807.06521

Xu Z, Kui CZ, Jing G, Wei H, Peng L, Zhang JN (2022). A two-stage deep transfer learning model and its application for medical image processing in traditional chinese medicine. Knowl Based Syst 239: 108060.

Zhuang FZ, Qi ZY, Duan KY, Xi DB, Zhu Y, Zhu HS. Xiong H, He Q (2021). A comprehensive survey on transfer learning. P IEEE 109: 43-76.