2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Generative Adversarial Training with Dual-Attention for Vascular Segmentation and Topological Analysis

Xueying Wang[1], Xiaoya Liu[2], Li Lin[1], Qiongyu Guo[2,*], and Xiaoying Tang[1,*]

***Abstract*— The vascular topology is of vital importance in building a chemotherapy model for the liver cancer in rats. And segmentation of vessels in the liver is an indispensable part of vessels' topological analysis. In this paper, we proposed and validated a novel pipeline for segmenting liver vessels and extracting their skeletons for topological analysis. We employed a dual-attention based U-Net trained in a generative adversarial network (GAN) fashion to obtain precise segmentations of vessels. For subsequent topological analysis, the vessels' skeletons are extracted and classified according to their lengths and bifurcation orders. Based on 40 samples with carefully-annotated ground truth labels, our experiments revealed consistent superiority in terms of both segmentation accuracy and topology correctness, demonstrating the robustness of the proposed pipeline.**

***Clinical relevance*— This work provides a useful and practical tool for analyzing the topology of anticancer drug-diffused vessels.**

I. INTRODUCTION

Liver cancer has become a major killer in recent years. As a predominant pathological type of liver malignant tumors, hepatocellular carcinoma (HCC) accounts for $85\% - 90\%$ of primary liver cancers. Transcatheter arterial chemoembolization (TACE) is the main therapeutic option for HCC [1]. To effectively evaluate the effectiveness of anticancer drugs in the TACE treatment, it is important to build a drug diffusion model in vitro. Gao *et al*. [2] tried to build an *ex vivo* liver chemotherapy model by hyalinizing a decellularized rat liver and injecting anticancer drugs into liver vessels. They then took photographs of the liver at regular intervals, manually annotated the extent and skeleton of the drug-diffused vessels, and measured the length, diameter and area of the vessel branches. The entire process is time-consuming since it involves a lot of manual work, and thus it is necessary to design an automatic pipeline for analyzing the drug diffusion process.

Accurate vascular segmentation is the first and most important step in the entire pipeline, which significantly affects the topology of vessels. Recently, with the development of deep learning, deep convolutional neural networks (CNNs) have been the main research tool for a wide variety of biomedical image analysis problems [3] [4]. To enhance the performance of CNNs, various backbones have been proposed. For example, VGGNet [5] replaces large convolution kernels with small stacking blocks of the same shape to reduce the training parameters. To make the network deeper, ResNet [6] stacks residual blocks along with skip connection. Apart from these backbones, a variety of network structures have also emerged, such as U-Net [7], LinkNet [8], FPN [9] and so on.

The aforementioned methods mainly rely on pixel-wise loss functions, and thus may generate discrete prediction results with wrong topology. Generative adversarial networks (GANs) attend to make the predicted output indistinguishable from ground truth by alternatively training a discriminator and a generator [10]. In the process of adversarial training, the discriminator can automatically learn the continuous topology of manual annotation as a supplement loss for the pixel-wise loss. In addition to generative adversarial training, since the drug concentration within vessels gradually decreases, it may result in blurry edges and endings. That motivates the utility of an attention mechanism to focus on important features and suppress unnecessary ones [11].

In such context, we propose a pipeline for liver vessels' segmentation and topological analysis in the drug-diffusion process of TACE, by taking GAN as a general framework and incorporating a dual-attention module into U-Net as GAN's generator. After segmentation, we iteratively delete redundant points to obtain the vascular skeleton, the leaf nodes of which are then classified according to branch length and bifurcation order to characterize the vessels' topology. Fig. 1 shows the flowchart of our overall pipeline. Experiments show that the proposed pipeline can generate precise segmentation results with correct topology.

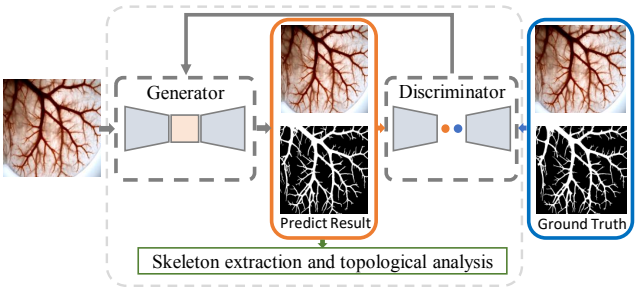

Fig. 1. Flowchart of the overall pipeline.

*This study was supported by the National Natural Science Foundation of China (62071210), the Shenzhen Basic Research Program (JCYJ20190809120205578), the National Key R&D Program of China (2017YFC0112404), and the High-level University Fund (G02236002). The authors declare that they have no competing financial interests.

[1]Department of Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China
[2]Department of Biomedical Engineering, Southern University of Science and Technology, Shenzhen, China
*Correspondence at tangxy@sustech.edu.cn and guoqy@sustech.edu.cn

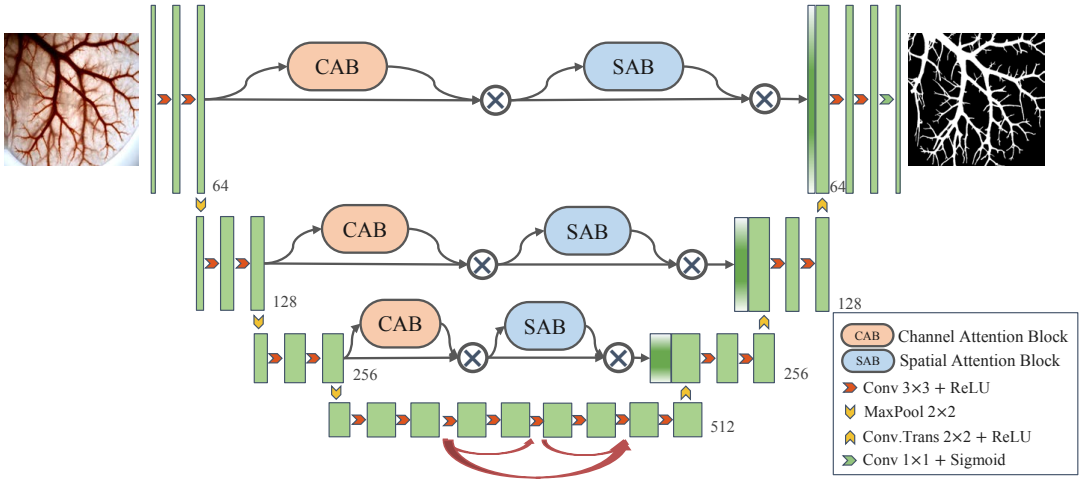

Fig. 2. Architecture of the generator in our proposed DAU-GAN.

II. METHOD

A. Image Preprocessing

The sample size of our dataset is 40, and thus data augmentation is needed. A typically-employed data augmentation approach is to randomly crop images into patches, but experimental results turn out that this strategy works poorly for our task. After exploration, we adopt random rotation and flipping as well as random brightness and contrast adjustment for data augmentation, resulting in 5460 images. Afterwards, all images are normalized to be of zero mean and unit variance.

B. Network Architecture

The network structure is a U-Net based GAN with a dual-attention module incorporated, abbreviated as DAU-GAN. It contains a generator (G) and a discriminator (D). G is trained to produce a probability mask of vessels that cannot be easily distinguished from "real" masks (the manually-annotated ground truth), while D is trained to perform as well as possible at detecting the generator's "faked" outputs.

Architecture of the generator With its encoding-decoding structure and skip connection, U-Net has greatly succeeded in medical image segmentation tasks. As Fig. 2 shows, we choose U-Net as the backbone of our generator. To have the network more properly focus on the target object, we replace the simple skip connection in U-Net with two attention-based feature refining blocks, namely the Channel Attention Block (CAB) and the Spatial Attention Block (SAB). They are arranged sequentially and generate feature maps $\mathbf{M_c}$ and $\mathbf{M_s}$ respectively. Fig. 3 shows details of those two blocks.

To exploit the inter-channel relationship of features, the CAB squeezes the input feature map's spatial dimension with max-pooling and average-pooling operations. Then the two pooled descriptors share multi-layer perceptron (MLP) with one hidden layer. After that, we apply element-wise addition to merge the output feature vectors and conduct the rectified linear unit (ReLU) activation.

To aggregate channel information, the SAB conducts max-pooling and average-pooling along the channel axis and concatenate them to generate a feature descriptor. Then a convolution layer is used to produce a spatial attention map, following the ReLU activation.

Assume the input feature map is $\mathbf{F}$, the overall flow can be described as

$$\begin{aligned} \mathbf{F}' &= \mathbf{M_c}(\mathbf{F}) \otimes \mathbf{F}, \\ \mathbf{F}'' &= \mathbf{M_s}(\mathbf{F}') \otimes \mathbf{F}', \end{aligned} \tag{1}$$

where $\otimes$ indicates element-wise product.

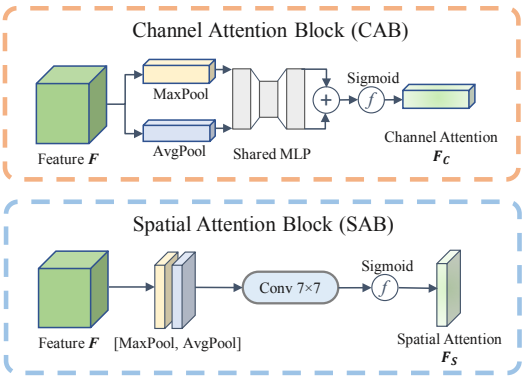

Fig. 3. The two attention blocks in our proposed DAU-GAN.

Architecture of the discriminator There are five convolution blocks in our discriminator. Each block consists of two 3×3 convolution layers followed by group normalization and ReLU. There is a 2×2 max-pooling layer between every two blocks. Finally, a fully connected layer and a sigmoid layer are employed to determine the classification.

Objective function The loss function of GAN can be formulated as

$$\begin{aligned} \mathscr{L}_{GAN} =& E_{x,y\sim p_{\text{data}}(x,y)}[\log D(x,y)] \\ &+ E_{x\sim p_{\text{data}}(x)}[\log(1-D(x,G(x))], \end{aligned} \tag{2}$$

where G tries to minimize $\mathscr{L}_{GAN}$ against an adversarial D that tries to maximize $\mathscr{L}_{GAN}$. The x, y respectively represent an

original vascular image and the corresponding ground truth. D maps a pair of input data into binary categories $\{0,1\}$, where 0 and 1 respectively denote the paired sample is either model-generated or ground truth. G is designed to not only fool D but also make the output be close to the ground truth. Thus we also introduce the binary cross-entropy loss $\mathscr{L}_{SEG}$ between the ground truth and the automated prediction. The overall objective function is $\mathscr{L}_{total} = \mathscr{L}_{GAN} + \lambda \mathscr{L}_{SEG}$. λ is used to balance the two parts of the objective function. The minimax objective function is $\arg\min_G \max_D \mathscr{L}_{total}$.

C. Topological Analysis

Skeleton extraction We extract the vascular skeleton to describe the drug diffusion status at the specific moment of photographing. Skeleton extraction is based on a thinning algorithm that contains two sub-iterative processes [12]. One aims at deleting redundant points around the north-west corner and south-east boundary. The other one aims at deleting redundant points around the south-east corner and the north-west boundary. In the end, only connected skeleton of a single-pixel width is preserved in a binary skeleton matrix, where 1 (0) represents skeleton (background).

Leaf nodes classification Classifying the vascular branches into T classes according to the skeleton's length and bifurcation order can more accurately quantify the extent of drug diffusion. Fig. 4 shows one representative example of our topological analysis. The nonzero values in the skeleton matrix from the previous step are divided into two parts: key nodes (**KN**) and node-to-node edges. The **KN** can be further divided into bifurcation nodes (**BN**) and leaf nodes (**LN**), as shown in Fig. 4. All **KN** are incrementally numbered. Then we build an undirected graph (*Graph*) with interconnected **KN** and their corresponding edges. The number of pixels on the edge represents the branch's length. In the iterative process, we automatically choose the **KN** located at the vessel's root as root nodes (**RN**), determine the lengths between **RN** and **LN** with Dijkstra algorithm [13]. The skeleton branch with the maximum length is defined as the FIRST class skeleton. Each **BN** in the FIRST class skeleton extends out a sub-skeleton, and we specify that **BN** as a new **RN** of the sub-skeleton and identify the skeleton branch of the maximum length as the SECOND class skeleton. The rest are done in the same manner until T classes are calculated, and we end up with the vascular branches' pathways and lengths of each class.

III. EXPERIMENTS

A. Datasets and Implementation Details

Our drug-diffused rat liver vascular dataset contains 40 images in total, each of which is 2056×2464 in size[1]. The vessels were annotated by three experts. To ensure consistency of manually annotating standards, one expert marked the drug-diffused vessels, and the other two reviewed and modified if necessary.

[1]This dataset will be published later. All experimental procedures were approved by the Institutional Animal Care and Use Committee at the Sourthern University of Science and Technology.

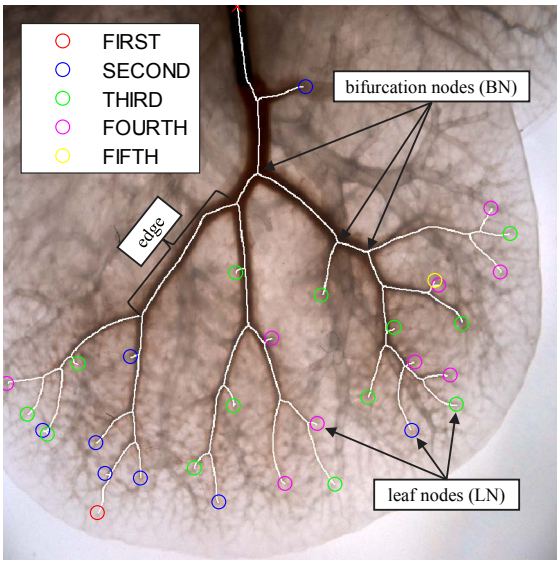

Fig. 4. Representative example of vascular skeleton for topological analysis. The white lines represent the vascular skeleton, and the colored circles represent different vascular classes.

We employ a 4-fold cross-validation strategy. In the training process, we use Adam optimizer. The learning rate is initialized as 1e-4. The coefficient λ is set to be 10.

B. Qualitative Evaluation

To prove the superiority of our proposed pipeline, we test and compare the performance of four different networks: FPN, U-Net, V-GAN [14], and the proposed DAU-GAN. For FPN and U-Net, we use VGG16 as their backbone. V-GAN without the attention module is used to prove the validity of DAU-GAN. Fig. 5 shows qualitative comparisons of different methods. The *Graph* is built based on interconnecting nodes, and thus we only keep connected vessels of the maximum size before skeleton extraction to avoid disconnected nodes. To make a more intuitive comparison of the topological structure, we thicken the skeleton in red and overlay it on the predicted vascular segmentation. It can be clearly seen that our proposed method produces results that are closest to the ground truth, not only in segmentation but also in topology.

C. Quantitative Evaluation

Table I shows the quantitative comparisons of the aforementioned four segmentation networks. The Dice Similarity Coefficient (DSC) focuses on pixel-level similarity between a segmentation result and its ground truth. The Average Symmetric Surface Distance (ASSD) and Maximum Symmetric Surface Distance (MSSD) describe two types of surface boundary distances between a segmentation result and the ground truth. The Topology Similarity (TS) is defined as $|v1 - v2|/(0.5 \times (v1 + v2))$, where $v1$, $v2$ respectively represent the total number of nodes belonging to a specific class in the ground truth and the prediction, and we take an average of the TS results of all class nodes.

Although the DSC of U-Net is much better than that of LinkNet and FPN, the ASSD and MSSD of U-Net are high,

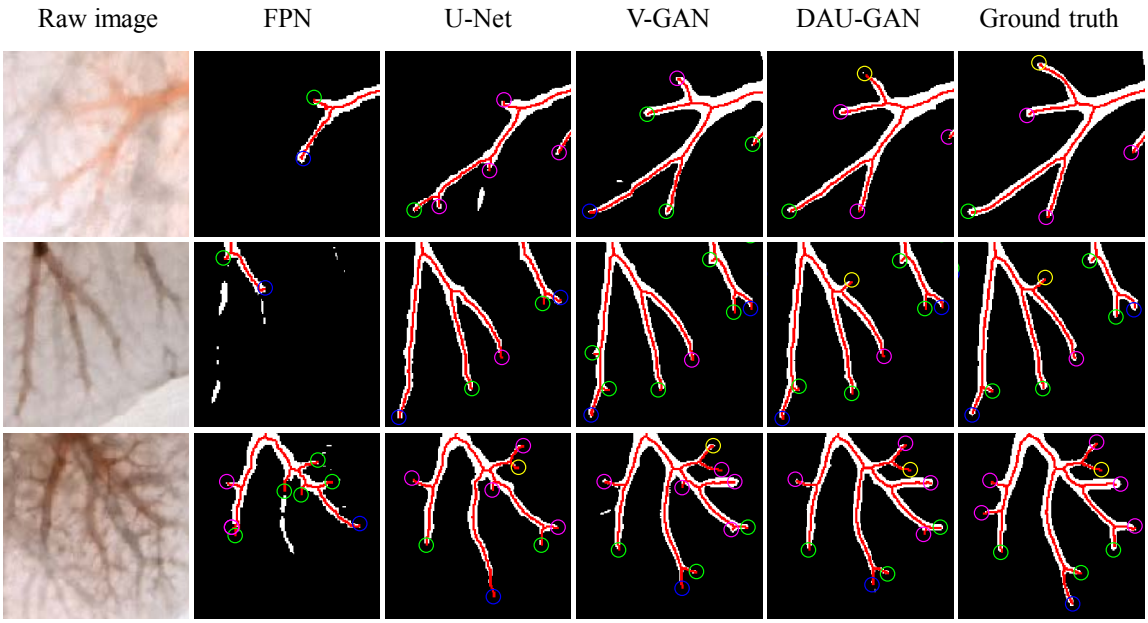

Fig. 5. Qualitative evaluations of the topology of results from different methods.

indicating that U-Net is not very effective in segmenting vascular endings. By applying U-Net in GAN as V-GAN, all four indicators have improved. More importantly, our proposed DAU-GAN, after incorporating both channel and spatial attention blocks, achieved the best overall performance.

TABLE I

QUANTITATIVE COMPARISON RESULTS.

Method	DSC[%]↑	MSSD ↓	ASSD↓	TS ↓
LinkNet	82.80	34.13	6.23	3.68
FPN	82.93	40.78	6.40	2.61
U-Net	87.39	33.14	6.20	2.02
V-GAN	90.74	14.18	1.93	1.80
DAU-GAN (proposed)	**91.22**	**13.12**	**1.87**	**1.33**

* The ↑ and ↓ respectively denote the higher (the lower) the better.

IV. CONCLUSION

In this paper, we proposed DAU-GAN, a novel pipeline for segmenting drug-diffused vessels and analyzing their topological structure. To the best of our knowledge, this study is the first one of such kind, due to the innovative nature of the dataset used in this work. The neural network was designed based on the GAN framework with U-Net dual-attention blocks. Then we extracted the segmentation results' skeletons and classified them according to lengths and bifurcation orders. Experiments showed the proposed DAU-GAN generated precise segmentation results, in terms of not only overlap accuracy but also topology correctness.

REFERENCES

[1] J.-M. Idée and B. Guiu, "Use of lipiodol as a drug-delivery system for transcatheter arterial chemoembolization of hepatocellular carcinoma: a review," *Critical reviews in oncology/hematology*, vol. 88, no. 3, pp. 530–549, 2013.

[2] Y. Gao, Z. Li, Y. Hong, T. Li, X. Hu, L. Sun, Z. Chen, Z. Chen, Z. Luo, X. Wang, J. Kong, G. Li, H.-L. Wang, H. L. Leo, H. Yu, L. Xi, and Q. Guo, "Decellularized liver as a translucent ex vivo model for vascular embolization evaluation," *Biomaterials*, vol. 240, p. 119855, 2020.

[3] Y. Zhang, J. Wu, Y. Liu, Y. Chen, E. X. Wu, and X. Tang, "A 3d+ 2d cnn approach incorporating boundary loss for stroke lesion segmentation," in *International Workshop on Machine Learning in Medical Imaging*. Springer, 2020, pp. 101–110.

[4] J. Lyu, P. Cheng, and X. Tang, "Fundus image based retinal vessel segmentation utilizing a fast and accurate fully convolutional network," in *International Workshop on Ophthalmic Medical Image Analysis*. Springer, 2019, pp. 112–120.

[5] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *Computer Science*, 2014.

[6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[7] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[8] A. Chaurasia and E. Culurciello, "Linknet: Exploiting encoder representations for efficient semantic segmentation," in *2017 IEEE Visual Communications and Image Processing (VCIP)*. IEEE, 2017, pp. 1–4.

[9] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2117–2125.

[10] L. Lin, J. Wu, P. Cheng, K. Wang, and X. Tang, "Blu-gan: Bi-directional convlstm u-net with generative adversarial training for retinal vessel segmentation," in *BenchCouncil Federated Intelligent Computing and Block Chain Conferences*, 2020.

[11] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "Cbam: Convolutional block attention module," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 3–19.

[12] S. Wang, Y. Zhang, and X. Tang, "Fully automatic diameter measurement method based on retinal vessel segmentation," in *Proceedings of the Third International Symposium on Image Computing and Digital Medicine*, 2019, pp. 218–221.

[13] E. W. Dijkstra *et al.*, "A note on two problems in connexion with graphs," *Numerische mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

[14] J. Son, S. J. Park, and K.-H. Jung, "Retinal vessel segmentation in fundoscopic images with generative adversarial networks," *arXiv preprint arXiv:1706.09318*, 2017.