Simulation Modelling Practice and Theory 121 (2022) 102653

Contents lists available at ScienceDirect

Simulation Modelling Practice and Theory

journal homepage: www.elsevier.com/locate/simpat

Graph-based convolution feature aggregation for retinal vessel segmentation ☆

Cao Shi [a], Canhui Xu [a,*], Jianfei He [a], Yinong Chen [b], Yuanzhi Cheng [a], Qi Yang [a], Haitao Qiu [a]

[a] School of Information Science and Technology, Qingdao University of Science and Technology, Qingdao, China
[b] School of Computing and Augmented Intelligence, Arizona State University, Tempe, AZ 85287-8809, USA

ARTICLE INFO

Keywords:
Retinal vessel segmentation
Graph neural network
IoT

ABSTRACT

Retinal vessel segmentation is a crucial task in computer assistant diagnosis of eye diseases. Instead of relying heavily on crafted features, high dimensional deep learning convolution features has provided better representation. However, the neglecting of inherent relation among multiple features is problematic. To explore non-local contextual dependencies, we proposed a graph-based convolution feature aggregation network (GCFAN) for segmenting retinal vessel and enhancing image non-vessel region simultaneously relying on graph to propagate and aggregate message of cross-level features. Specifically, the model consists of three modules: multi-level feature extraction module (MFEM), graph-based high level convolution feature aggregation module (GHFAM), graph-based low level convolution feature aggregation module (GLFAM). The multi-level feature representations are extracted from retinal image in MFEM. GHFAM utilizes more semantic information to reconstruct retinal image without vessel, which facilitates diagnosis. GLFAM utilizes more boundary information to segment vessel. Competitive experimental results on four retinal image datasets validate the efficacy of the proposed model, which achieves segmentation and reconstruct retinal image without vessel, indicating its potential clinical application. Finally, an IoT framework which integrates our algorithm is built to analyze image from various fundus cameras in different places and display results on PC and mobile phone simultaneously, which will facilitate doctor diagnose.

1. Introduction

Eye is the vital organ processing visual information, which contributes significantly to human perception. Its structure, function, and physiology has been focused on by researchers for decades [1]. As the senor responding to light, retina transmits impulses from two eyes to the brain through optic nerves. Its complex structure with at least nine layers and vital role in vision let human perceive color. Hence, research on retina has been continued [2–4].

Various retinal diseases caused by dysfunctions in different retinal layers. For example, drusen at the outside layer of retina, called retinal pigment epithelium, can cause age-related macular degeneration (AMD); other diseases like retinal and choroidal tumors affect several layers of retina, etc. These diseases can be diagnosed by ophthalmologists analyzing fundus photographs.

☆ This work was supported in part by the National Natural Science Foundation of China under Grant No. 61806107 and 61702135, Shandong Key Laboratory of Wisdom Mine Information Technology, and the Opening Project of State Key Laboratory of Digital Publishing Technology.

* Corresponding author.
E-mail address: ccxu09@yeah.net (C. Xu).

https://doi.org/10.1016/j.simpat.2022.102653

Available online 13 September 2022

Recently, the rapid development of graphics processing units (GPUs) from Nvidia [5] brings many advances that leads to deep learning effective enough to be used for medical image processing. Deep learning offers a complete "end-to-end" method, which employs a single deep neural network to fulfill multi-tasks such as input, feature extraction[6], detection[7,8], segmentation[9], classification, etc. Deep learning has obtained remarkable achievement in computer vision, so that the medical image processing community has taken great interest in exploring usages of deep learning in medical image tasks. Dhungel et al. [10] combined conditional random field (CRF), structured support machine (SSVM), deep belief network and convolutional neural network for breast mass segmentation in mammograms. Different from breast attaching outside to the chest wall, heart is in the chest. It also attracted researchers to analysis of cardiac image based on deep learning [11,12]. Besides organs such as breast and heart which locate in relatively fixed parts of body, blood vessels are tubular organs all around body for transporting blood. This kind of organs have aroused also researchers' interest. Zheng et al. [13] combined Haar wavelet features with features in deep learning to detect carotid artery bifurcation. Ghesu et al. [14] proposed a deep learning based shape model to estimate boundary of the aortic valve in ultrasound. And Wolterink et al. [15] exploited paired convolutional neural networks to automatically score coronary artery calcium in cardiac CT angiography.

As above mentioned organs, eye's function cannot be also ignored. It is hard to imagine how to perceive the world accurately without eye, because visual information accounts for 70% of information come into and processed in brain [16]. Naturally, algorithms on computer-aided diagnosis (CAD) for eye also exploited deep learning. As narrated before, many eye diseases involve the retina, and fundus image can be separated into vessels and non-vessel areas. Therefore, many algorithms were proposed to segment vessels from fundus image, like hybrids [17–19] of traditional approach and deep learning, or pure end-to-end architectures [20–23]. In these algorithms, most employed deep neural networks (DNNs) as trainable feature extractors, which were combined with features from non-end-to-end approaches. Convolutional neural network (CNN) was applied into retinal vessel segmentation in [17], however, for higher accuracy in [17], features from different layers of CNN were fed into random forests (RF) as inputs. Fully convolutional network (FCN) was employed in [20], which described a three-layer neural network predicting a label for each pixel. Inspired by FCN, a holistically nested edge detection (HED) [18] method was proposed, in which features from hidden layers of neural network are connected directly with the output layer, to generate vessel probability map. And CRF was utilized on the probability map to improve performance of segmentation. In order to gain speed up of algorithm, Wu et al. [19] combined pixel-wise classification with vessel tracking, in which vessels were traced using Monte Carlo random sampling. Most of the above methods adopt encoder–decoder architecture, in which the encoder module encodes information and decoder module decodes information, without carefully considering the inherent relationship with multi-layer during encode process. We explored the inherent relationship of features in graph convolutional network (GCN) [24] fed with retina images, because GCN could represent well intrinsic relationship between pixels or blocks (super pixels) in vessels and non-vessel areas. The feature is transformed into graph node form, and then the relationship between feature nodes at different levels was explored through the ability of graph transmit information. Thus, in our model, we use graph to propagate representation information, which can explore dependency and relationship between multi-layers.

In this paper, we proposed a graph-based convolution feature aggregation network(GCFAN) for retinal vessel segmentation. Our main contributions are summarized as follows:

- Propose a feature aggregation strategy, in which low level feature containing enriched details utilized for segmenting vessel, and high level feature for the reconstruction of non-vessel areas.
- Present graph-based low level convolution feature aggregation module, which makes well use of boundary information to segment vessel.
- Develop graph-based high level convolution feature aggregation module, which employs semantic information to reconstruct retinal image without vessel for facilitating diagnosis.

2. Related works

Many diseases are concerned with vessels in fundus image, such as macular telangiectasia, polypoidal choroidal vasculopathy, cotton-wool spots, artery obstruction, etc. As shown in Fig. 1, (b) and (c) illustrate macular telangiectasia. Capillaries dilate in (c) so that it looks fine pink, and the blood vessels dilating consecutively leads to rupture in (b). Idiopathic macular hole appears in (d), and it looks dark red like the outflow of blood from vessels. Hypertension results in retinal vessel occlusion so that blood leakage occurs in (e). Central retinal vein obstruction also shows blood leakage in (f), and it is just more serious. Therefore, segmentation of retinal vessels provides explicit reference for diagnosis. Wang et al. [17] integrated CNN and RF as a pipeline, in which RF took features generated by CNN as input for RF, in order to segment blood vessels in fundus image. For making well use of internal information of data, the work [20] modeled retinal vessel segmentation as a task to construct a function realizing cross-modality data transformation. Different from considering inherent characteristics of data, ensemble learning skills [25] was adapted in [21] for robust detection of vessels in fundus image. To further investigate fundus image of a particular disease, a novel microaneurysm (MA) detector was proposed in [26], which fulfilled early screening of diabetic retinopathy because MA was usually an indicator. Wu et al. [19] proposed deep vessel tracking, to combine CNN and traditional detectors for more efficient segmentation of retinal vessel.

Recently, graph convolutional network (GCN) [24] was adapted in fundus image analysis [22,23], since GCN is good at representing spatial information. The idea of directly modeling graph data instead of using one dimension vector was proposed in [27]. Moreover, GCN was used in [28] to represent non-Euclidean graph data. Considering computational complexity [29] of

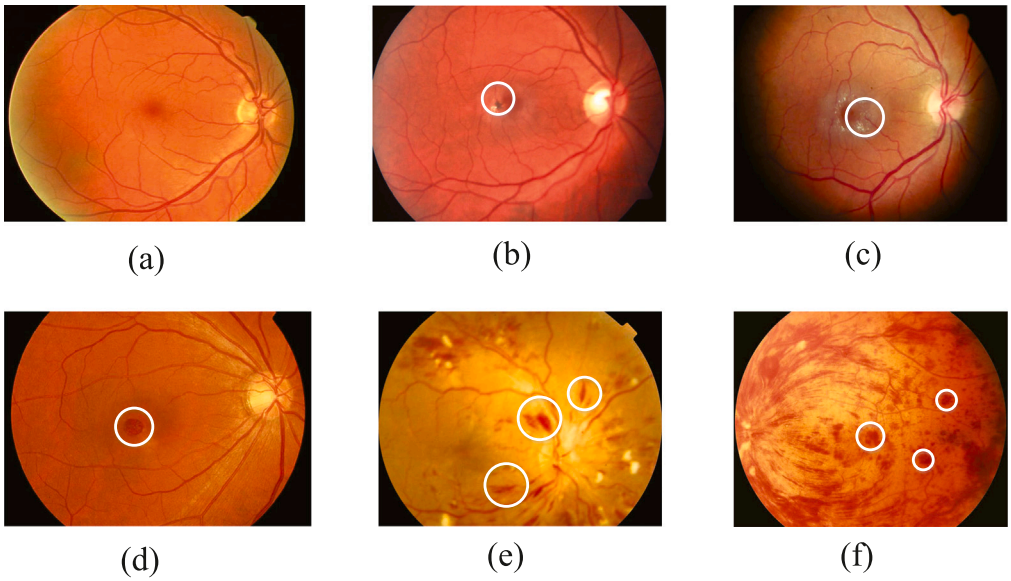

Fig. 1. Several common diseases of the retina. (a) Normal retinal fundus images. (b) and (c) Macular telangiectasia: The blood vessel dilates and leaks under macula, generating swelling of macula and vision loss. (d) Idiopathic macular hole: Small holes appear in the macula, which is responsible for reading and observing detail. (e) Hypertensive Retinopathy: Flame-shaped hemorrhages and yellow hard Exudates from leaking vessels occur on the surface of the retina. (f) Central Retinal Vein Obstruction: Clinical manifestations are macular edema or abnormal blood vessels, which can make vision blurred and bring other eye problems.

GCN, how to simplify computation was focused on by research works, like [24,30]. GCN is naturally utilized in medical image analysis, since many tissues or organs in medical image have structures similar to graph structures. For example, blood vessels in fundus images have a tree topology. A novel retinal vessel segmentation algorithm [23] based on GCN was proposed, with two parallel pipelines: a dynamic channel graph convolutional network and edge enhancement blocks. Region and boundary aggregation method for retinal vessel segmentation based on GCN in [22] made well use of topological associations in fundus image to achieve better segmentation performance.

On the other hand, with the invention of hardware prerequisite GPU [5], CAD based on deep learning has been applied in clinical medicine. The DeepMind [31] was used for diagnoses of eye diseases in Moorfields Eye Hospital [32,33]. A preoperative aesthetic assessment based on CNN was proposed in [34] to investigate the impact on face and age before cosmetic surgery. Furthermore, as the Internet of things (IoT) has been wildly used in various fields, it also has aroused great interest in medical image processing. In [35], 3D medical image processing pipeline were distributed in an IoT environment, in which X-ray scanning data were transmitted through ZigBee component and devices were terminals connected in the IoT. An optimal feature selection model for medical image analysis in IoT was proposed in [36]. Wang et al. [37] proposed a CT segmentation method based on deep learning in Internet-of-Medical-Things (IoMT) domain. Dhevi et al. [38] used Cognitive Multiple-input Multiple-output (MIMO) Multicarrier Code-division-multiple-access (MC-CDMA) to transmit images on monitoring patient in IoT.

3. Graph-based convolution feature aggregation network

3.1. Network architecture

The architecture of the proposed graph-based convolution feature aggregation network is shown in Fig. 2, including three parts: multi-level feature extraction module(MFEM) in the left of Fig. 2, graph-based high level convolution feature aggregation module(GHFAM) in the top, graph-based low level convolution feature aggregation module(GLFAM) in the bottom. MFEM adopts ResNet [39] as the backbone network for visual feature extraction. It aims to extract multi-level features as inputs for subsequent modules GHFAM and GLFAM. There is a dual flow including GHFAM and GLFAM in Fig. 2. The top one GHFAM makes well use of semantic features to gain non-vessel fundus images, whereas, The GLFAM at the bottom focuses on detailed features on vessels to obtain the vessel segmentation.

The Alg. 1 gives an overview of the processing flow of our approach. A retinal image is fed to the proposed network, and features at various scales are extracted by MFEM (the left part of Fig. 2, Alg.1 line 1). Graph nodes in both GHFAM and LFAM are initialized (Alg.1 line 2). And then extracted features with different resolutions are put into GHFAM and GLFAM. In Alg. 1, lines 3 to 9 correspond to the top branch GHFAM in Fig. 2. Here, extracted features are used to reconstruct retina images without vessels. On the other hand, the vessel segmentation is fulfilled through the bottom branch GLFAM in Fig. 2 corresponding to Alg. 1 lines 10 to 16.

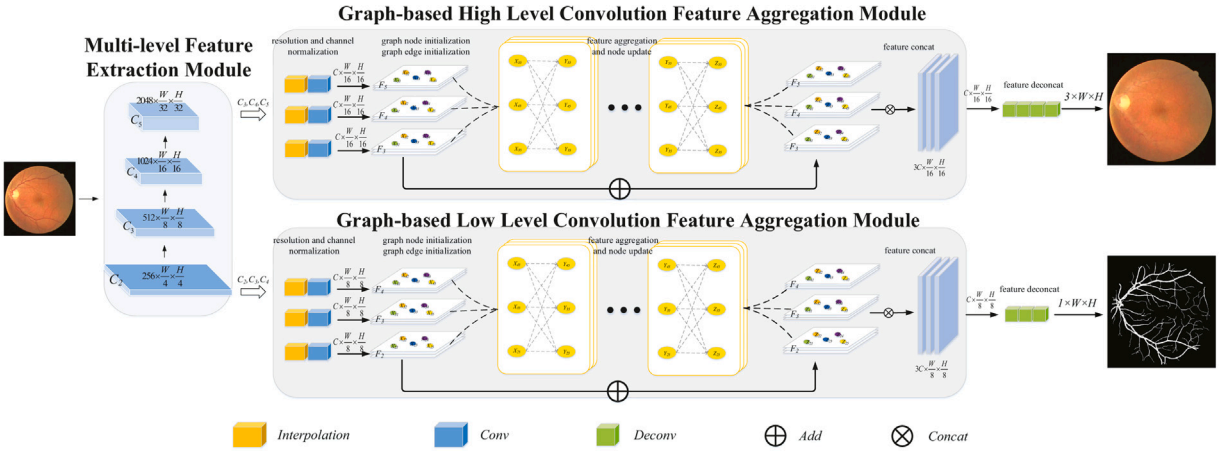

Fig. 2. The architecture of our proposed graph-based convolution feature aggregation network (GCFAN). With backbone feature maps, we construct a group of node graphs across different layers. Then, attribute adjacency matrix could be obtained from node graphs to capture relationships and dependencies among nodes. After multiple iterations, node information can be updated relying on its relationships with surrounding nodes for predicting results.

Algorithm 1: graph-based convolution feature aggregation network

```
• Given a retinal image, using a backbone network to extract feature maps;
• Initialize graph node base feature maps;
if In retinal construct branch then
    for t = 1 to T do
        calculate edge between graph nodes;
        aggregate feature and update nodes' information;
    end
    obtain image without vessel;
end
if In retinal vessel segmentation branch then
    for t = 1 to T do
        calculate edge between graph nodes;
        aggregate feature and update nodes' information;
    end
    obtain vessel segmentation result;
end
```

3.1.1. Multi-level feature extraction module

In MFEM, ResNet [39] outputs a feature group $\{C_i\}$. The interpolation operation is utilized to resize resolutions of cross-level features from bottom to top (C_2 to C_5), with reduction rate 0.5, and the numbers of channel of C_2, C_3, C_4, C_5 are 256, 512, 1024 and 2056 respectively. Parameters W and H of MFEM in Fig. 2 are width and height of input retina image. C_i can be calculated iteratively using:

$$C_{i+1} = Conv(Interp(C_i)) \quad i = \{2, 3, 4\} \tag{1}$$

where $Interp$ denotes interpolation, $Conv$ is 3×3 convolution. As shown in Fig. 2, the fundamental feature representation C_i is fed into graph convolution network GHFAM and GLFAM to explore context and dependency information across feature representations.

3.1.2. Graph-based high level convolution feature aggregation module

From outputs of MFEM, GHFAM takes three as inputs C_3, C_4, C_5, of which each is interpolated and convolved to get the same number of channels 64 and the same resolution $W/16 \times H/16$. And then these features are used for graph node initialization and graph edge initialization. After graph initialization, feature will be aggregated and node will be updated. Finally, features are concatenated and deconvolved to reconstruct non-vessel fundus images.

The GHFAM flow can be summarized as following steps:

<i> resolution and channel normalization;

<ii> graph node initialization;

<iii> graph edge initialization;

<iv> feature aggregation and node update;
<v> feature concat and deconvolution.

3.1.3. Graph-based low level convolution feature aggregation module

Similar to GHFAM, GLFAM takes the lower layers C_2, C_3, C_4 as inputs. Each of them is interpolated and convolved to get the same number of channels 64 and the same resolution $W/8\times H/8$. Actually, except selection of feature layers and interpolation parameters, GLFAM has the same implementation steps <i>~<v> in GHFAM.

GLFAM uses features C_2, C_3, C_4 from lower three layers of MFEM with higher resolution $W/8\times H/8$ than GHFAM's $W/16\times H/16$, because higher resolution means more details in feature maps. GLFAM is designed for segment vessels, hence it needs more details than GHFAM. Whereas, lower resolution features C_3, C_4, C_5 gains better semantic information for GHFAM to reconstruct non-vessel fundus images, which have less high frequency information than vessels.

3.2. Detailed operations of the proposed graph-based network

Given a graph $G = (V, E)$, graph nodes $v_i \in V$ are relevant with initial features, and graph edges $e_{ij} \in E$ denotes representative relation between pair-wise nodes (v_i, v_j). As shown in Fig. 2, GHFAM's inputs C_3, C_4, C_5 are interpolated and convolved to F_3, F_4, F_5. Every F_i has the same size $64\times W/16\times H/16$. Take a pixel X_{31} from F_3, and apparently its size is 64×1. Similarly, take X_{41}, X_{51} from F_4, F_5, respectively. And then X_{31}, X_{41}, X_{51} comprise a nodes set V. In the GHFAM flow, features F_3, F_4, F_5 are aggregated iteratively. In the aggregation process, graph nodes X_{31}, X_{41}, X_{51} are updated iteratively. After the definition of graph, the implementation steps <ii>, <iii> and <iv> of GHFAM and GLFAM are defined as follows:

3.2.1. Graph node initialization

In graph G, every node's information can be updated from neighborhood nodes to explore relationship and dependency between nodes. Directly employing pixels of input images as graph nodes will impose a substantial computational burden. Thus, graph nodes $V = \{v_1, v_2, \ldots, v_n\}$ are built on pixels of the extracted feature map. For example, in Fig. 2, as mentioned before, X_{31}, X_{41}, X_{51} from F_3, F_4, F_5 in GHFAM (or GLFAM) comprise a set V. There are two reasons for building Graph using pixels from different feature maps: First, the features from different levels have different correlated information, which can be utilized to explore the dependencies of nodes among different level features. Second, Graph can propagate message between nodes, and all nodes can be updated during an iteration with other nodes.

3.2.2. Graph edge initialization

In graph neural network(GNN), graph edges represent relations of nodes, especially, bigger weight indicates tighter relationship. Inspired by [40], the edge e_{ij} from v_i to v_j is denoted as:

$$e_{ij} = Conv(Concat(v_i - v_j, v_j)) \tag{2}$$

where $Concat$ is a channel-wise concatenation, $Conv$ is a convolution with kernel size 1×1. Similarly, the edge e_{ji} from v_j to v_i is defined as:

$$e_{ji} = Conv(Concat(v_j - v_i, v_i)) \tag{3}$$

Using (2) and (3) edges between nodes has been built, which describe relation.

3.2.3. Feature aggregation and node update

To propagate and exchange information between graph nodes, features of current node are aggregated from neighborhood nodes. The processing can be represented as follows:

$$m_i^T = \sum_j^{n-1} ReLU(e_{ji}^{T-1}) \otimes v_j^{T-1} \tag{4}$$

where $\otimes$ is an element-wise multiplication. $ReLU$ is a non-linear active function, in order to convert edge to link weight. T is the number of iterations. Considering continuity of vessel structure, relationships between nodes can be built with their neighborhoods. Affections of discontinuity of segmentation can be relieved with Eq. (4). Because ResNet can mitigated gradient disappearance and gradient explosion so as to make training of deep networks possible, we use a residual connection to update the node embedding. Every node can be updated according to its own information and surrounding nodes, after T times information propagation and aggregation:

$$v_i^T = m_i^T + v_i^{T-1} \tag{5}$$

Using Eq. (5), output nodes (the left-hand side of the equation) is generated through channel-wise concatenation, following by 1×1 convolution.

Table 1
The data distributions of our four public dataset.

Dataset	Number	Train-test	Origin-resolution	Reshape-size
DRIVE	20	15–5	584 × 565	256 × 256
CHASE_DB1	28	24–4	999 × 960	256 × 256
HRF	45	40–5	3504 × 2336	256 × 256
IOSTAR	24	20–4	1024 × 1024	256 × 256
Total	117	99–18	–	256 × 256

3.2.4. Loss function

In the retinal vessel segmentation (the graph-based low level convolution feature aggregation module in Fig. 2), we adopt Dice loss and Cross Entropy (CE) loss. Dice loss originates from Sørensen–Dice coefficient, which calculate similarity between two binary inputs, and can automatically assign weights to samples of different classes to establish the right balance between foreground and background pixels, so that the sample foreground and background imbalance problem is alleviated. The Dice loss can be defined as follows:

$$DL(y, \hat{p}) = 1 - \frac{2y\hat{p} + \epsilon}{y + \hat{p} + \epsilon} \tag{6}$$

where y represents label, $\hat{p}$ denotes prediction, ϵ is a small value which is normally assigned 1.

CE loss is a popular loss function in classification and segmentation, and it can be defined as follows:

$$CE(y, \hat{p}) = -ylog(\hat{p}) - (1 - y)log(1 - (\hat{p})) \tag{7}$$

where y represents label, $\hat{p}$ denotes prediction.

In the retinal image reconstruction branch (the graph-based high level convolution feature aggregation module in Fig. 2), we adopt Mean Squared Error (MSE) loss, which can be defined as follows:

$$ML(y, \hat{p}) = \frac{1}{n} \times \sum_{i}^{n} (y_i - \hat{p}_i)^2 \tag{8}$$

where n is the number of pixels, y represents label, $\hat{p}$ denotes prediction.

4. Experiments

4.1. Dataset

To comprehensively evaluate the performance of our proposed network GFAM, a series of experiments is conducted on four public retinal image datasets, including DRIVE [41], CHASE_DB1 [42], HRF [43], IOSTAR [44]. The detail information of the four datasets is summarized in Table 1.

The Digital Retinal Images for Vessel Extraction(DRIVE) dataset consists of 40 color fundus images, with resolution 584 × 565 pixels. In the experiments, all image are resized to 512 × 512 pixels, and the ratio of training to testing data is 1 : 1.

CHASE_DB1 contains 28 color fundus images of 14 children's left and right eyes, with the size of 999 × 960 pixels, which are resized to 512 × 512 pixels for training.

HRF consists of 45 color fundus images, with resolution 3304 × 2336 pixels. These images can be further divided into 15 groups. Every group contains 3 images, one healthy, one image of patient with diabetic retinopathy and one glaucoma image.

IOSTAR vessel segmentation dataset includes 30 color fundus images with resolution of 1024 × 1024 pixels. All images resize to 512 × 512 pixels.

In total, 117 dolor fundus images with label are used in our designed experiments, in which 99 are regarded as training set and the remaining 18 as testing set.

4.2. Implementation details

Our network GCFAN is implemented using PyTorch 1.8.0, and monai 1.7.0. We train it on one RTX 3080Ti GPU with the Stochastic Gradient Descent (SGD) optimizer and a momentum of 0.9. Initial learning rate is set to 0.008, and damping to last time's 0.8 every 10,000 iterations. The number of iteration is 50,000, and the batch size is set to 2. Monai's DiceCELoss is adapted as loss function in the training. In order to train the GHFAM branch of our method, non-vessel retinal images are manually generated from the four retinal image datasets.

Table 2
Comparison of different model experimental results.

	Accuracy	Precision	Recall	Dice
U-Net	0.9642	0.7839	0.7437	0.7605
SegNet	0.9644	0.7825	0.7664	0.7720
Attention U-Net	0.9625	0.7629	**0.7813**	0.7657
ours	**0.9679**	**0.8189**	0.7521	**0.7825**

4.3. Evaluation metrics

Evaluation metrics such as Accuracy, Precision, Recall, Dice are utilized to evaluate the performance of the proposed model and other models, they are defined by the following formula:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

$$Precision = \frac{TP}{TP + FP} \tag{10}$$

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

$$Dice = \frac{2TP}{2TP + FP + FN} \tag{12}$$

where TP, TN, FP, FN denotes true positive, true negative, false positive, and false negative pixels, respectively, in the segmentation result. Specially, accuracy presents the overall performance of vessel and non-vessel. Precision measures the percentage of true vessel region in predict vessel region. Recall denotes the rate of predict positive subject as negative subject. Dice calculates the overlap between prediction and ground truth, which can measure segmentation result in imbalance label condition, such as retina vessel segmentation.

4.4. Results

To demonstrate the performance of the proposed network, we compare it with existing segmentation methods, such as U-Net [45], SegNet [46] and Attention U-Net [47], on four datasets under the same experimental configurations. The overall experimental results are shown in Table 2, which lists Accuracy, Precision, Recall, Dice. In addition, the ultimate visual comparison is shown in Fig. 3 and segmentation result can be presented on different terminal devices in IoT, and its architecture is shown in Fig. 4.

Table 2 compares the performance of four methods with the same metrics, and the last row shows our method. As for $Accuracy$, our method obtains the best value 0.9679, because our method segments vessels and non-vessel areas simultaneously. Compared with other methods in Table 2, the GHFAM branch in the proposed method is able to segment non-vessel areas well. TP represents how much pixels of vessels a method labels out of all the pixels of vessels, whereas, TN measures how much pixels of non-vessel areas a method labels out of all the pixels of non-vessel areas. The GHFAM branch in our method let TN increase, so that $Accuracy$ increases. Because of the same reason, our method gains the best precision 0.8189 and the best dice 0.7825. Hence, it is safe to conclude that our method shows excellent performance to segment vessels.

Visual comparison of several methods implemented under the same experimental configurations are displayed in Fig. 3. The first column lists four retinal images (a)–(d). The second column includes vessel images corresponding to the first column. The red circles mark areas being zoomed in to show more detailed vessels in the third column. The third and the fourth show vessel segmentation results from U-Net, and the sixth and the seventh are ours. The yellow circle and the green circle locate different areas, which are used to compare the segmentation performance between U-Net and our method. From Fig. 3, experimental results show the efficiency of information aggregation from multi-feature using graph. Compared with U-Net in the fifth column in row (a), the green circle in the seven sub-figure in row (a) indicates our methods can generate finer results because of aggregating message from low-level features which contain more detail information. In the rows (a), (b), by comparing red circles between our method and U-Net, it is apparent that our predictions are more continuous because of it considers the relationship and dependency with surrounding nodes in the graph, and improve the precision of results. Actually, blood vessels are inherently continuous. As shown in the column (c), due to exploring non-local contextual dependencies, our results minimize unnecessary prediction and contain less noise.

4.5. Internet of Things (IoT) platform

The Internet of Things (IoT) aims to build connection between people and every device by internet. Therefore, all devices can collect, share data, and work better together. In our model, as shown in Fig. 4, fundus camera instruments in different palaces can post their image to the Data Cloud which stores data. Then, the server can read image data and analysis image by our retina vessel segmentation algorithm, the results will be uploaded to Cloud Data again, for terminal display. Through this process, doctors and

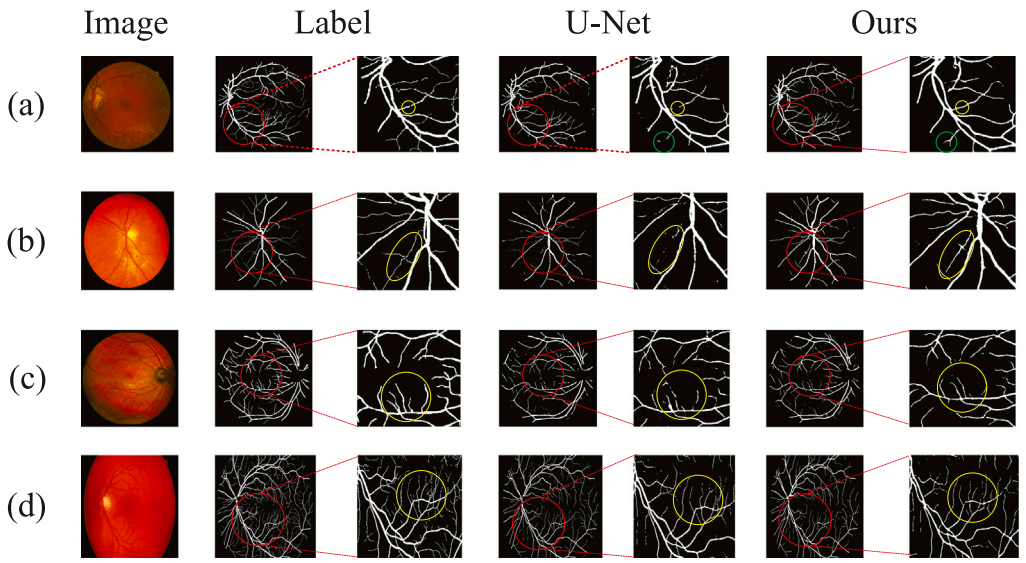

Fig. 3. Visual comparison of several methods implemented under the same experimental conditions.

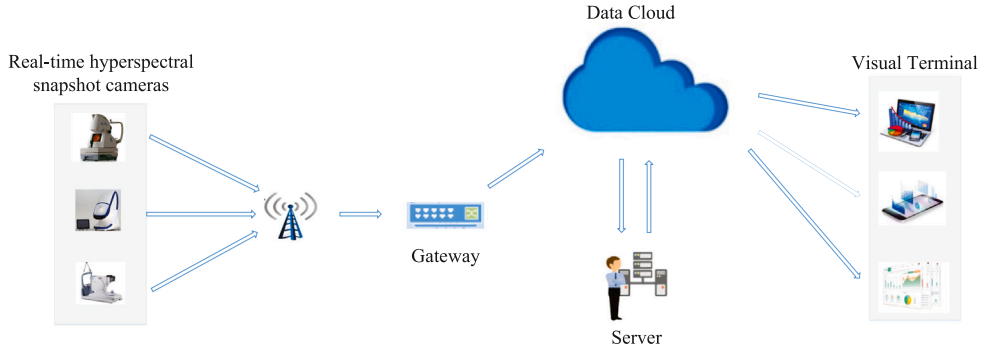

Fig. 4. The architecture of IoT platform. The images collected by the fundus camera are transmitted to Cloud Data through the internet, then the server downloads images from Could Data and processes them with our model. After calculation, the processed data are re-uploaded to Cloud Data. Finally, the enhanced images are displayed in various terminals to help doctors diagnose and propose treatments easily and efficiently.

patients can expediently and easily gain roughly information about retinal current state, to take appropriate action with the help of AI and IoT. Final results can be displayed in PC and Mobile Phone, as shown in Fig. 5.

IoT healthcare facilities are beneficial for retina image collection which can be used to study retinal diseases for doctors. Accessing remote data through the use of IoT, hospitals and medical institutions can reduce costs by sharing retinal image data. Of course, IoT has some drawbacks, especially, the privacy. It is difficult for IoT to grant the security of data which are acquired by IoT but used by different users and devices. Meanwhile, if device integration and IoT implementation are not standardized, the potential capability of IoT will be limited. Hence, it is important to take a tradeoff between advantages and disadvantages of IoT to maximize its usage.

5. Conclusion

In this paper, we propose graph-based convolution feature aggregation network (GCFAN). It consists of three modules: multi-level feature extraction module, graph-based high level convolution feature aggregation module and graph-based low level convolution feature aggregation module. Relying on the propagating and aggregating information ability of graph, our model explores non-local contextual relationship and dependency between paired-wise nodes and make predictions are more continuous. In addition, due to aggregation across different level features (low level features contain enriched detail information and high level features contain more contextual information), the attenuation of vessel edge during the sampling and convention process is effectively alleviated. A series of experiments are conducted on four public retinal vessel datasets, and the results demonstrate the efficacy of our proposed method. Finally, we integrate our network into a IoT framework, aiming to facilitate doctor diagnosis. In this way, all image data from various fundus camera in different places can be analyzed online by Artificial intelligence (AI) and the results can be displayed on PC and mobile phones simultaneously.

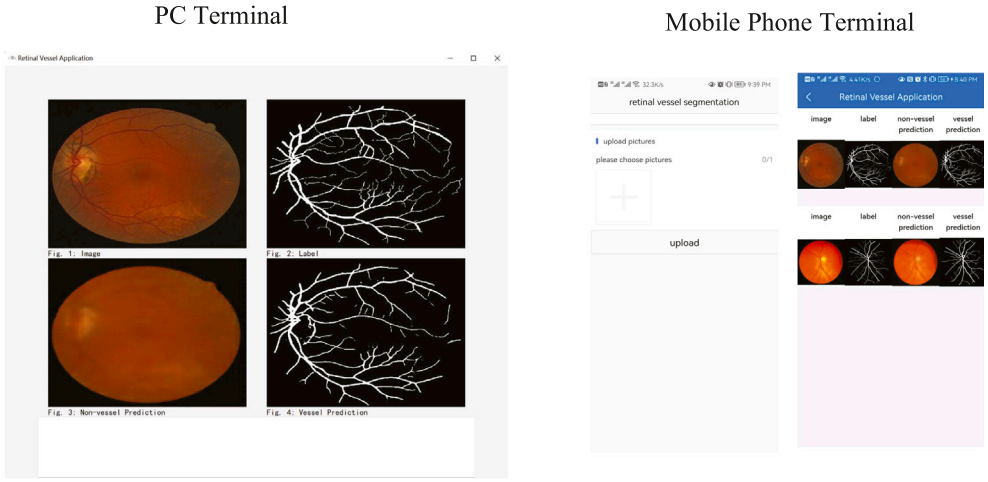

Fig. 5. Visualization on PC and Mobile Phone Devices.

Data availability

Data will be made available on request.

References

[1] N.R. Galloway, W.M.K. Amoaku, Common eye diseases and their management, Springer, 1999.
[2] A.A.E.F. Elsharif, S.S. Abu-Naser, Retina diseases diagnosis using deep learning, Int. J. Acad. Eng. Res. (IJAER) 6 (2) (2022).
[3] F. Abbas, S. Becker, B.W. Jones, L.S. Mure, S. Panda, A. Hanneken, F. Vinberg, Revival of light signalling in the postmortem mouse and human retina, Nature (2022) 1–7.
[4] A.U. Kale, A. Serrano, X. Liu, B. Balasubramaniam, P.A. Keane, D.J. Moore, V. Llorenç, A.K. Denniston, Measuring inflammation in the vitreous and retina: A narrative review, Ocular Immunol. Inflamm. (2022) 1–10.
[5] Nvidia, 2022, Website https://www.nvidia.com/.
[6] C. Shi, C. Xu, H. Bi, Y. Cheng, Y. Li, H. Zhang, Lateral feature enhancement network for page object detection, IEEE Trans. Instrum. Meas. 71 (2022) 1–10, http://dx.doi.org/10.1109/TIM.2022.3201546.
[7] C. Xu, C. Shi, H. Bi, C. Liu, Y. Yuan, H. Guo, Y. Chen, A page object detection method based on mask R-CNN, IEEE Access 9 (2021) 143448–143457, http://dx.doi.org/10.1109/ACCESS.2021.3121152.
[8] H. Bi, C. Xu, C. Shi, G. Liu, Y. Li, H. Zhang, J. Qu, SRRV: a novel document object detector based on spatial-related relation and vision, IEEE Trans. Multimedia (2022) 1–1, http://dx.doi.org/10.1109/TMM.2022.3165717.
[9] C.-h. Xu, C. Shi, Y.-n. Chen, End-to-end dilated convolution network for document image semantic segmentation, J. Cent. South Univ. 28 (6) (2021) 1765–1774.
[10] N. Dhungel, G. Carneiro, A.P. Bradley, Deep learning and structured prediction for the segmentation of mass in mammograms, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 605–612.
[11] M.R. Avendi, A. Kheradvar, H. Jafarkhani, A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI, Med. Image Anal. 30 (2016) 108–119.
[12] T. Anh Ngo, G. Carneiro, Fully automated non-rigid segmentation with distance regularized level set evolution initialized and constrained by deep-structured inference, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 3118–3125.
[13] Y. Zheng, D. Liu, B. Georgescu, H. Nguyen, D. Comaniciu, 3D deep learning for efficient and robust landmark detection in volumetric data, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 565–572.
[14] F.C. Ghesu, E. Krubasik, B. Georgescu, V. Singh, Y. Zheng, J. Hornegger, D. Comaniciu, Marginal space deep learning: efficient architecture for volumetric image parsing, IEEE Trans. Med. Imaging 35 (5) (2016) 1217–1228.
[15] J.M. Wolterink, T. Leiner, B.D. de Vos, R.W. van Hamersvelt, M.A. Viergever, I. Išgum, Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks, Med. Image Anal. 34 (2016) 123–136.
[16] D. Williams, Autism–an Inside-Out Approach: An Innovative Look At the Mechanics of'Autism'and Its Developmental'Cousins', Jessica Kingsley Publishers, 1996.
[17] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, G. Yang, Hierarchical retinal blood vessel segmentation based on feature and ensemble learning, Neurocomputing 149 (2015) 708–717.
[18] H. Fu, Y. Xu, D.W.K. Wong, J. Liu, Retinal vessel segmentation via deep learning network and fully-connected conditional random fields, in: 2016 IEEE 13th International Symposium on Biomedical Imaging, ISBI, IEEE, 2016, pp. 698–701.
[19] A. Wu, Z. Xu, M. Gao, M. Buty, D.J. Mollura, Deep vessel tracking: A generalized probabilistic approach via deep learning, in: 2016 IEEE 13th International Symposium on Biomedical Imaging, ISBI, IEEE, 2016, pp. 1363–1367.
[20] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, T. Wang, A cross-modality learning approach for vessel segmentation in retinal images, IEEE Trans. Med. Imaging 35 (1) (2015) 109–118.
[21] D. Maji, A. Santara, P. Mitra, D. Sheet, Ensemble of deep convolutional neural networks for learning to detect retinal vessels in fundus images, 2016, arXiv preprint arXiv:1603.04833.
[22] Y. Meng, H. Zhang, Y. Zhao, X. Yang, Y. Qiao, I.J.C. MacCormick, X. Huang, Y. Zheng, Graph-based region and boundary aggregation for biomedical image segmentation, IEEE Trans. Med. Imaging 41 (3) (2022) 690–701, http://dx.doi.org/10.1109/TMI.2021.3123567.

[23] Y. Li, Y. Zhang, W. Cui, B. Lei, X. Kuang, T. Zhang, Dual encoder-based dynamic-channel graph convolutional network with edge enhancement for retinal vessel segmentation, IEEE Trans. Med. Imaging (2022) 1, http://dx.doi.org/10.1109/TMI.2022.3151666.
[24] T.N. Kipf, M. Welling, Semi-supervised classification with graph convolutional networks, 2016, arXiv preprint arXiv:1609.02907.
[25] T.G. Dietterich, Ensemble methods in machine learning, in: International Workshop on Multiple Classifier Systems, Springer, 2000, pp. 1–15.
[26] M. Haloi, Improved microaneurysm detection using deep neural networks, 2015, arXiv preprint arXiv:1505.04424.
[27] M. Gori, G. Monfardini, F. Scarselli, A new model for learning in graph domains, in: Proceedings. 2005 IEEE International Joint Conference on Neural Networks, Vol. 2, (2005) 2005, pp. 729–734.
[28] S. Wan, C. Gong, P. Zhong, B. Du, L. Zhang, J. Yang, Multiscale dynamic graph convolutional network for hyperspectral image classification, IEEE Trans. Geosci. Remote Sens. 58 (5) (2020) 3162–3177, http://dx.doi.org/10.1109/TGRS.2019.2949180.
[29] F. Scarselli, M. Gori, A.C. Tsoi, M. Hagenbuchner, G. Monfardini, Computational capabilities of graph neural networks, IEEE Trans. Neural Netw. 20 (1) (2009) 81–102, http://dx.doi.org/10.1109/TNN.2008.2005141.
[30] F. Wu, T. Zhang, A.H. de Souza, C. Fifty, T. Yu, K.Q. Weinberger, Simplifying graph convolutional networks (supplementary material), 2019.
[31] Deepmind, 2022, Website https://www.deepmind.com/.
[32] Moorfields researchers show how AI could help detect eye disease in diabetes patients, 2022, Website https://www.moorfields.nhs.uk/news/moorfields-researchers-show-how-ai-could-help-detect-eye-disease-diabetes-patients.
[33] M. Burgess, Now deepmind's ai can spot eye disease just as well as your doctor, Wired August 13 (2018).
[34] R. Patcas, D.A. Bernini, A. Volokitin, E. Agustsson, R. Rothe, R. Timofte, Applying artificial intelligence to assess the impact of orthognathic treatment on facial attractiveness and estimated age, Int. J. Oral Maxillofac. Surg. 48 (1) (2019) 77–83.
[35] H. Pan, X. Yang, Application of Internet of Things technology in 3D medical image model, IEEE Access 7 (2019) 5508–5518, http://dx.doi.org/10.1109/ACCESS.2018.2886223.
[36] R.J.S. Raj, S.J. Shobana, I.V. Pustokhina, D.A. Pustokhin, D. Gupta, K. Shankar, Optimal feature selection-based medical image classification using deep learning model in internet of medical things, IEEE Access 8 (2020) 58006–58017, http://dx.doi.org/10.1109/ACCESS.2020.2981337.
[37] E.K. Wang, C.-M. Chen, M.M. Hassan, A. Almogren, A deep learning based medical image segmentation technique in internet-of-medical-things domain, Future Gener. Comput. Syst. 108 (2020) 135–144, http://dx.doi.org/10.1016/j.future.2020.02.054, URL https://www.sciencedirect.com/science/article/pii/S0167739X19323313.
[38] B. Lakshmi Dhevi, K. Vishvaksenan, K. Senthamil Selvan, A. Rajalakshmi, Patient monitoring system using cognitive internet of things, J. Med. Syst. 42 (11) (2018) 1–9.
[39] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, 2016, pp. 770–778.
[40] S. Woo, J. Park, J.-Y. Lee, I.S. Kweon, Cbam: Convolutional block attention module, in: Proceedings of the European Conference on Computer Vision, ECCV, 2018, pp. 3–19.
[41] J. Staal, M.D. Abràmoff, M. Niemeijer, M.A. Viergever, B. van Ginneken, Ridge-based vessel segmentation in color images of the retina, IEEE Trans. Med. Imaging 23 (2004) 501–509.
[42] M.M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, S.A. Barman, An ensemble classification-based approach applied to retinal blood vessel segmentation, IEEE Trans. Biomed. Eng. 59 (2012) 2538–2548.
[43] J.I. Orlando, E. Prokofyeva, M.B. Blaschko, A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images, IEEE Trans. Biomed. Eng. 64 (2017) 16–27.
[44] J. Zhang, B. Dashtbozorg, E.J. Bekkers, J.P.W. Pluim, R. Duits, B.M. ter Haar Romeny, Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores, IEEE Trans. Med. Imaging 35 (2016) 2631–2644.
[45] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, 2015, ArXiv, arXiv:1505.04597.
[46] V. Badrinarayanan, A. Kendall, R. Cipolla, SegNet: A deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (2017) 2481–2495.
[47] O. Oktay, J. Schlemper, L.L. Folgoc, M.J. Lee, M.P. Heinrich, K. Misawa, K. Mori, S.G. McDonagh, N.Y. Hammerla, B. Kainz, B. Glocker, D. Rueckert, Attention U-net: Learning where to look for the pancreas, 2018, ArXiv, arXiv:1804.03999.