

[文章编号] 1003-4684(2022)04-0024-05

基于多层特征融合的视网膜血管分割模型

刘延龙, 吴 聪

(湖北工业大学计算机学院, 湖北 武汉 430068)

[摘 要] 针对现有的视网膜血管分割算法在下采样的学习层中不能有效捕捉更多特征,且低层特征没有得到充分利用从而提高视网膜血管分割的准确性,提出了一种新颖的多层特征融合网络。首先设计了一个具有连续记忆力机制的模块来增加网络的深度,以捕获更多的特征。在此基础上提出了一个多层特征融合模块将低层特征和高层特征进行融合。在 DRIVE 和 CHASE_DB1 数据集上进行训练和测试,结果表明:与其它流行的深度学习方法相比,该方法在两个数据集上获得了更好的结果,视网膜血管分割结果更加准确。

[关键词] 医学图像处理; 视网膜血管分割; 连续记忆力机制; 多层特征融合

[中图分类号] TP391 **[文献标识码]** A

眼睛是人类接收和获取外界信息的重要途径之一,占接受信息总和的 80%。因此,眼睛的健康对人们日常的生活和工作学习有着至关重要的影响。医学临床中,眼底视网膜检查是眼部检查的重要步骤,对眼底视网膜图像的判读和分析,有助于医生对相关疾病进行检查和诊断。而检查最重要的则是提取出眼底视网膜血管的形态和结构。在视网膜血管图像中,视网膜血管结构和形态的改变是眼底病变最常见的形式。视网膜血管的血液循环受到其周围的局部组织影响,往往是许多其他心脑血管疾病发生的原因。例如,糖尿病视网膜病变可以通过视网膜血管结构的改变来诊断。它会导致失明,这意味着早期发现是至关重要的。高血压是另一种视网膜疾病,高血压性视网膜病变。高血压患者血管曲度增加或血管狭窄。因此,视网膜血管分割的准确性可作为相关疾病判断的重要依据。但是人工分割视网膜血管不仅费时费力,还需要丰富的经验和熟练的技能。因此,开发高精度、省时省力的计算机系统辅助视网膜血管检测是目前广泛需要的。

目前,深度学习方法在各个领域都取得了巨大的成功,受到了许多研究者的关注,也成为了医学图像处理的首选方法。深度学习是一种机器学习方法。卷积神经网络(CNN)在图像分类上取得了很大的成功,但它并没有解决分割任务的问题。为了解决这个问题,Long 等^[1]提出了一种全卷积神经网络(FCN)。这允许从任意大小的输入图像生成分割图像。同年,Ronneberger 等^[2]改进了 FCN,提出

了一种医学图像分割模型 U-Net。U-Net 可以利用小样本进行训练,得到较好的分割结果。借鉴这些先进的思想,文献[3]将深度学习方法应用于医学图像分析。但这些方法都有一定的局限性,随着网络深度的增加,容易出现梯度消失。同时,这些方法在小血管检测上也有局限性,很容易遗漏一些细小结构。为了解决梯度随网络深度增加而消失的问题,K. He 等^[4]提出了残差网络(Residual Network, ResNet),它引入了残差块,使训练这些深度网络变得容易,缺点是它有很多参数,需要大量的计算开销。针对此问题,G. Huang 等^[5]提出了密集连接,命名为密集连接卷积网络(Dense Connected Convolutional Networks, DenseNet)。它引入了密集块,以减少参数,并提供更好的性能。Foivos I 等^[6]提出了一个 Resunet,其骨干是 U-Net。文献[7]提出了 Resunet ++,进一步改进了 U-Net。文献[8]提出了 DenseUNet。该算法在 U-Net 中采用密集连接的方法对肝脏和肿瘤进行分割。Jegou 等^[9]提出了全卷积密集网络,他们在 FCN 中引入了密集块,并在实验中取得了良好的效果。

近年来,许多基于深度学习的方法被应用于视网膜血管分割领域。并取得了不错的视网膜血管分割结果。Song Guo 等^[10]采用短连接来缩小输出层之间的语义差距,命名为多尺度深度监督网络(multi-scale deeply supervised network, BTS-DSN)。Juntang Zhuang 等^[11]提出了一种具有多个编码器和解码器的网络。它采用了跳跃连接,使网

[收稿日期] 2021-10-04

[第一作者] 刘延龙(1995-),男,陕西延安人,湖北工业大学硕士研究生,研究方向为医学图像处理

络具有更多路径的信息传输。M. Z. Alom 等^[12]提出了一种 R2U-Net,他们将 RNN 和 ResNet 的结构集成到了编码器-解码器结构中。Changlu Guo 等^[13]提出了一种深度密集残差网络结构(a deep dense residual network structure,DRNet)。该方法将残差连接和密集连接的思想相结合应用于视网膜图像的血管分割。文献[14]提出了利用 atrous 卷积(attention guided U-Net with atrous convolution,AA-UNet)的注意力引导 U-Net,重复使用特征进行血管分割。在文献[15]中,提出了一种多路径递归 U 网深度学习体系结构。该结构结合了递归神经网络和卷积神经网络,对原有的 U-Net 和递归单元进行了相应的改进。Xiaoyu Guo 等^[16]提出了在密集 U-Net 中使用 Inception 模块进行视网膜血管分割。Zhun Fan 等^[17]提出了一种采用不同于传统卷积的卷积方式来获取更多信息的方法。它可以更好地捕捉不同形状和大小的视网膜血管。Aashis Khanal 等^[18]在网络中使用动态卷积学习更多的特征用于视网膜血管分割,从而提高对细小血管的检测能力。

综上所述,这些方法都是端到端的,没有对低层特征进行有效的利用。本文提出多层特征融合方法,充分利用了低层视觉特征,是视网膜血管分割的一种新思路。

1 基于多层特征融合的视网膜血管分割模型

本文提出了一种视网膜血管分割的多层次特征融合模型(图 1)。

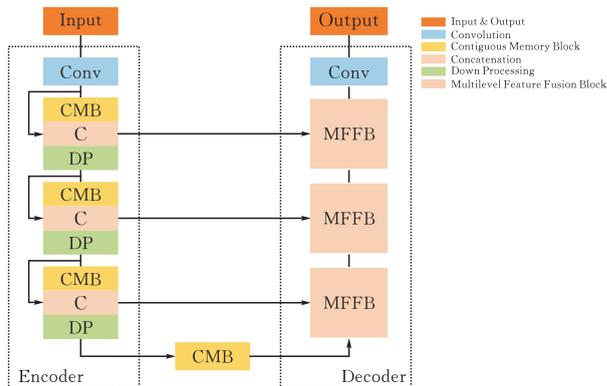


图 1 MFFNet 的网络结构

2.1 多层特征融合模型(MFFNet, Multilevel Feature Fusion Net)

本文提出的多层特征融合模型包含两个部分:包含连续记忆力模块(CMB, Contiguous Memory Block)的编码器和包含多层特征融合模块(MFFB, Multilevel Feature Fusion Block)的解码器。定义

I_{raw} 作为输入和 I_{seg} 作为最后的输出。在编码器中,首先使用卷积层来从输入的原始图像捕获特征 F_0 。

$$F_0 = O_{conv}(I_{raw}) \quad (1)$$

式(1)中, O_{conv} 是一个卷积运算。 F_0 用作第一个连续记忆力模块(CMB)的输入。MFFNet 中采用了四个连续的记忆力模块。定义 F_d 是 d_{th} 连续记忆力模块的输出。

$$F_d = O_{cmb,d}(F_{d-1}) \quad (2)$$

式(2)中 $O_{cmb,d}$ 是 d_{th} 连续记忆力模块的运算,是一个包含卷积运算和矫正线性单元(ReLU)的复合函数。其中, F_d 充分利用了连续记忆力模块中的每个卷积层,因此编码器实现了一种连续的记忆和存储机制。关于连续记忆力模块的更多细节将在 2.2 介绍。利用连续记忆力模块捕获特征后,得到一个充分利用所有卷积层信息的特征。将其用于解码器中,设计一个用于上采样的多层特征融合块。最后,经过一个卷积层,MFFNet 的输出可以由下式得到:

$$I_{seg} = O_{MFF}(I_{raw}) \quad (3)$$

式(3)中 O_{MFF} 是整个 MFFNet 的操作。

这里的 DP 包含 BN, Relu, 1×1 卷积和一个大小 2×2 的最大池化。

2.2 编码器中的连续记忆力模块(CMB, Contiguous Memory Block)

在下采样过程中,本文采用了下面的特征传递方式,来学习更多关于低层特征的信息。图 2 展示了连续记忆力模块中特征传递的细节,它是通过将前一个 CMB 的信息传输到当前 CMB 的每一层,从而形成了一种连续记忆机制。

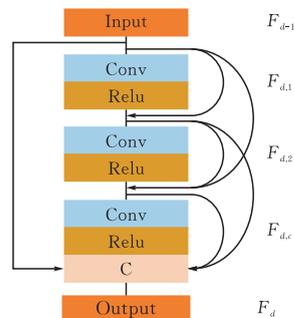


图 2 连续记忆力模块结构

F_{d-1} 和 F_d 作为在 d_{th} CMB 的输入和输出,它们有相同的特征图 G_0 。这意味着在 CMB 中保持相同的特征图。 d_{th} CMB 的 c_{th} 卷积层的输出可通过以下途径获取:

$$F_{d,c} = \sigma[W_{d,c}(F_{d-1}, F_{d,1}, F_{d,2}, \dots, F_{d,c-1})]$$

其中 σ 代表 ReLU 函数, $W_{d,c}$ 是 C_{th} 卷积层的权重。让 $F_{d,c}$ 有 G_0 特征图, G 代表增长率。 $[F_{d-1}, F_{d,1}, F_{d,2}, \dots, F_{d,c-1}]$ 是 $(d-1)_{th}$ CMB 中的特征, CMB 中的卷积层 $1, 2, \dots, c-1$ 的输出可以由 $G_0 + (c-$

1) $\times G$ 得到。输出可以直接利用于所有后续层中的每一层,传递需要保存的特征。可以从下式得到:

$$F_{d,c} = C_d(F_{d,1}, F_{d,2}, \dots, F_{d,c})$$

其中 C_d 表示在 d_{th} CMB 中的连接操作,充分利用了现有 CMB 的低层次特征。最终的 d_{th} CMB 输出可通过下面方法得到:

$$F_d = R(F_{d-1}, F_{d,c})$$

上式中 R 为 ResNet 的跳跃连接。最后,在下采样过程中的最后一个 CMB 的输出特征(F),将用于上采样过程中的 MFFB,计算方法如下:

$$F = \sigma[W(F_d)] = \sigma[W[R(F_{d-1}, F_{d,c})]] \quad (4)$$

2.3 解码器中的多层特征融合模块(The Multilevel Feature Fusion Block, MFFB)

图 3 是解码器中的多层特征融合模块,该模块输入有两层特征。将下采样过程中学习到的视觉特征称之为低层特征,上采样学习到的语义特征称为高层特征。多层特征融合块的左侧是一个低层语义特征处理的过程,CMB 对训练好的低层特征进行调整。这个连续记忆力模块有助于消除无效的特征并增强有用的低层特征。然后跟随一个 3×3 的卷积学习一个新的权重特征融合。该操作可以通过以下方式获取:

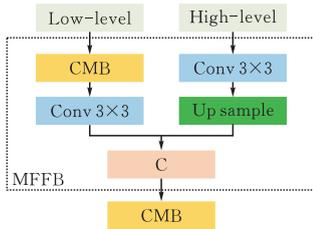


图 3 多层特征融合模块

$$F_{low-level} = W(\omega_0 F_{d-1}, \omega_1 F_{d-1}, \dots, \omega_c F_{d,c})$$

$F_{d,c}$ 是 d_{th} CMB 下采样特征学习过程得到的每个特征图。 W 代表卷积运算,也是特征的学习权值。

右边是高层特征处理,通过卷积操作调节输入的高级特征的权重,然后通过上采样过程返回到与低层特征相同的通道数和维数。然后,对两个层次特征进行连接和加权。最后,在经过 CMB 处理后,它将用于下一个 MFFB。MFFB 的输出可以通过下式得到:

$$F_{high-level} = W(aF_{new-high-level}, bF_{low-level})$$

$F_{new-high-level}$ 是当前多层特征融合块的输出, $F_{low-level}$ 是下采样过程学习的特征, a 和 b 代表 $F_{new-high-level}$ 和 $F_{low-level}$ 的权重。 W 是像原始 U-Net 一样的连接操作。

在多层特征融合块中采用连续记忆块的原因是为了产生一种新的高级特征,然后对其进行后处理以提高结果的准确性。可以调节框架中每一层的权

重,允许两个层次的特征更有效地融合。但是,如果在 CMB 中直接连接输入和输出,将导致特征的数量增加,意味着更多的计算。为了解决这个问题,CMB 的输入并不直接与其输出连接。在 CMB 的每一层中先采用 1×1 卷积,再采用 3×3 卷积,以抑制特征图数量的增加。一方面,经过 CMB 学习后,低阶特征的维数总是相同的。另一方面,通过上采样使高层特征的通道数与低层特征通道数保持一致。down processing (DP) 包括 BN、Relu、 1×1 conv 和大小为 2×2 的 max-pooling。上采样是由一个步长为 2 的转置卷积组成的。MFFNet 的体系结构见图 4。

| | | Input | |
|-------|----------------------------|------------------------------|----------|
| Conv1 | | Conv 7×7 | stride=2 |
| CMB1 | | CMB (N=4) | |
| DP1 | | Conv 1×1 | stride=1 |
| | | Maxpool 2×2 | stride=2 |
| CMB2 | | CMB (N=6) | |
| DP2 | | Conv 1×1 | stride=1 |
| | | Maxpool 2×2 | stride=2 |
| CMB3 | | CMB (N=8) | |
| DP3 | | Conv 1×1 | stride=1 |
| | | Maxpool 2×2 | stride=2 |
| CMB4 | | CMB (N=10) | |
| MFFB1 | CMB (N=10) | Conv 3×3 | stride=1 |
| | Conv 3×3 stride=1 | Transposed Conv 3×3 | stride=2 |
| MFFB2 | CMB (N=8) | CMB (N=8) | |
| | Conv 3×3 stride=1 | Conv 3×3 | stride=1 |
| | | Transposed Conv 3×3 | stride=2 |
| MFFB3 | CMB (N=6) | CMB (N=6) | |
| | Conv 3×3 stride=1 | Conv 3×3 | stride=1 |
| | | Transposed Conv 3×3 | stride=2 |
| Conv2 | | CMB (N=4) | |
| | | Transposed Conv 3×3 | stride=2 |
| | | Output | |

图 4 MFFNet 的网络结构

3 实验结果和分析

在不同的视网膜血管数据集上实验,以检查 MFFNet 的性能。

3.1 数据准备及预处理

DRIVE: 该数据集是从荷兰糖尿病视网膜病变 (DR) 的一个筛查项目中收集的。像素为 565×584 , 包含 40 张图像。为了进行训练,分成 20 张图像进行训练,20 张图像进行测试。

CHASE_DB1: 本数据集来自英国儿童听力与健康协会的一个研究项目。像素为 999×960 , 包含 28 张图像。分成 20 张图像进行训练,其余 8 张图像进行测试。两个数据集的具体信息见表 1。

表 1 DRIVE 和 CHASE_DB1

| 数据集 | Drive | CHASE_DB1 |
|---------|------------------|--------------------|
| 总数 | 40 | 28 |
| 训练集/测试集 | 20/20 | 20/8 |
| 像素 | 584×565 | 999×960 |
| 调整后大小 | 592×592 | 1008×1008 |

增强方法: 1) 随机旋转; 2) 添加高斯噪声; 3) 水平、垂直和对角线翻转

3.2 评价指标

本文实验中使用的评估方法性能的指标包括准确度(AC)、灵敏度(SE)和特异性(SP)。不同指标的计算如表 2 所示。

表 2 本文使用的评价标准

| 评价标准 | 方法描述 |
|-----------------|--|
| Accuracy(AC) | $AC = \frac{TP + TN}{TP + TN + FP + FN}$ |
| Sensitivity(SE) | $SE = \frac{TP}{TP + FN}$ |
| Specificity(SP) | $SP = \frac{TN}{TN + FP}$ |

其中 TP 为真阳性,表示有血管的区域分类正确。反之,如果该区域分类错误,则命名为 false negative (FN)。非血管区分类正确称之为 True negative(TN)。如果背景区域分类错误,则称为假阳性(FP)。

3.3 实验结果与分析

本文所有的实验都是在带有 GPU Titan X 的 windows PC 上进行的。其中,第一个卷积层是一个步长为 2 的 7×7 卷积,其他所有层的设置如图 4 所示。本文实验使用的网络层数如图 5 所示。在这项工作中,设定的增长率(k)是 16。

| | | |
|--------------|-------------------|-------------|
| Conv1 | Conv 7×7 | $\times 1$ |
| CMB1 (N=4) | Conv 1×1 | $\times 4$ |
| | Conv 3×3 | $\times 4$ |
| DP1 | Conv 1×1 | $\times 1$ |
| CMB2 (N=6) | Conv 1×1 | $\times 6$ |
| | Conv 3×3 | $\times 6$ |
| DP2 | Conv 1×1 | $\times 1$ |
| CMB3 (N=8) | Conv 1×1 | $\times 8$ |
| | Conv 3×3 | $\times 8$ |
| DP3 | Conv 1×1 | $\times 1$ |
| CMB4 (N=10) | Conv 1×1 | $\times 10$ |
| | Conv 3×3 | $\times 10$ |
| MFFB1 | Conv 1×1 | $\times 18$ |
| | Conv 3×3 | $\times 21$ |
| MFFB2 | Conv 1×1 | $\times 14$ |
| | Conv 3×3 | $\times 17$ |
| MFFB3 | Conv 1×1 | $\times 10$ |
| | Conv 3×3 | $\times 13$ |
| Conv2 | Conv 3×3 | $\times 1$ |
| Total layers | | 144 |

图 5 采用深度为 4 的 MFFNet 进行视网膜血管分割

在实验结果可视化对比(图 6)中,展示了 U-Net、U-Net+MFFB(在 U-Net 的上采样过程中使用 MFFB)和 MFFNet 在 DRIVE 和 CHASE_DB1 上的分割结果。可以看出,MFFNet 的分割性能更好,检测出了一些容易遗漏的细小结构。表 3 显示了 U-Net、残差 U-Net、LadderNet、U-Net+MFFB 和 MFFNet 在 DRIVE 和 CHASE_DB1 上的分割结果。结果表明:1)U-Net+MFFB 和 MFFNet 的性能与其它模型对比都有所提升,证明了引入多层特征融合的策略是有效的。2)MFFNet 的 AC 和 AUC 分别比 U-Net 在 DRIVE 和 CHASE_DB1 上的 AC 和 AUC 高 1.59%/1.7%和 0.97%/1.2%,证

明包含了 CMB 和 MFFB 的 MFFNet 性能更好。在医学图像分割任务需要的高精度来讲,改进效果显著。

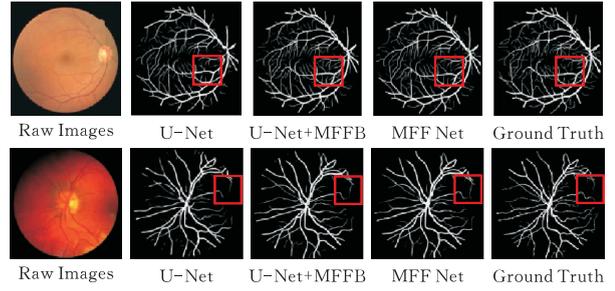


图 6 在 DRIVE 和 CHASE_DB1 上的分割结果

表 3 在 DRIVE 和 CHASE_DB1 上的测试结果

| | Methods | AC | SE | SP | AUC |
|-----------|----------------|---------------|---------------|---------------|---------------|
| DRIVE | U-Net | 0.9531 | 0.7537 | 0.9820 | 0.9755 |
| | Residual U-Net | 0.9553 | 0.7726 | 0.9820 | 0.9779 |
| | LadderNet | 0.9561 | 0.7856 | 0.9810 | 0.9793 |
| | UNet+MFFB | 0.9675 | 0.7902 | 0.9844 | 0.9821 |
| | MFFNet | 0.9690 | 0.8148 | 0.9838 | 0.9852 |
| CHASE_DB1 | U-Net | 0.9578 | 0.8288 | 0.9701 | 0.9772 |
| | Residual U-Net | 0.9553 | 0.7726 | 0.9820 | 0.9779 |
| | LadderNet | 0.9656 | 0.7978 | 0.9818 | 0.9839 |
| | UNet+MFFB | 0.9734 | 0.8328 | 0.9833 | 0.9869 |
| | MFFNet | 0.9748 | 0.8484 | 0.9832 | 0.9892 |

4 结论

本文提出一种多层次特征融合网络(MFF-Net),该网络能够增强特征学习和传输,用于视网膜血管分割。本文设计的连续记忆力模块意味着 CMB 的输出可以直接利用下一个 CMB 中的每一层,CMB 中的每一卷次层可以利用后面的每一层传递有用的特征。然后,提出一种多层特征融合块来融合两层特征。它不同于大多数直接连接低层特征和高层特征端到端方法,而是高效的提取低层视觉特征并将其与高层语义特征相融合从而提升分割性能。通过对所提出的 MFFNet 在 DRIVE 和 CHASE_DB1 数据集上进行测试,可以发现与其他流行的方法相比,该方法在该分割任务上表现更好。

[参 考 文 献]

- [1] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4):640-651.
- [2] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[C]// International Conference on Medical image

- computing and computer-assisted intervention. Springer, Cham, 2015, 234-241.
- [3] FANG T B, SU R, XIE L P, et al. Retinal vessel landmark detection using deep learning and hessian matrix [C] // 2015 8th International Congress on Image and Signal Processing (CISP). IEEE, 2015, 387-392.
- [4] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] // Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, 770-778.
- [5] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks [C] // Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, 4700-4708.
- [6] FOIVOS. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 162:94-114.
- [7] JHA D, SMEDSRUD P H, RIEGLER M A, et al. ResUNet++: an advanced architecture for medical image segmentation [C] // 2019 IEEE International Symposium on Multimedia (ISM). IEEE, 2019, 225-2255.
- [8] LI X, CHEN H, QI X, et al. H-DenseUNet: hybrid densely connected unet for liver and tumor segmentation from CT volumes [J]. IEEE Transactions on Medical Imaging, 2018, 37(12): 2663-2674.
- [9] S JÉGOU, DROZDZAL M, DVAZQUEZ, et al. The one hundred layers tiramisu: fully convolutional densenets for semantic segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, 11-19.
- [10] GUO S, WANG K, KANG H, et al. BTS-DSN: deeply supervised neural network with short connections for retinal vessel segmentation [J]. International Journal of Medical Informatics, 2019, 126:105-113.
- [11] ZHUANG J. LadderNet: Multi-path networks based on U-Net for medical image segmentation [J/OL]. arXiv preprint arXiv:1810.07810, 2018. [2021-10-01]. <https://arxiv.org/abs/1810.07810>.
- [12] ALOM M Z, HASAN M, YAKOPCIC C, et al. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation [J/OL]. [2021-10-01]. <https://arxiv.org/abs/1802.06955>.
- [13] GUO C, SZEMENYEI M, YI Y, et al. Dense residual network for retinal vessel segmentation [C] // ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2020, 1374-1378.
- [14] LV Y, MA H, LI J, et al. Attention guided U-Net with atrous convolution for accurate retinal vessels segmentation [J]. IEEE Access, 2020, 8: 32 826-32 839.
- [15] JIANG Y, WANG F, GAO J, et al. Multi-path recurrent U-Net segmentation of retinal fundus image [J]. Applied Sciences, 2020, 10(11):3777.
- [16] GUO X, CHEN C, LU Y, et al. Retinal vessel segmentation combined with generative adversarial networks and dense U-Net [J]. IEEE Access, 2020, 8:194 551-194 560.
- [17] FAN Z, MO J, QIU B, et al. Accurate retinal vessel segmentation via octave convolution neural network [J/OL]. [2021-10-01]. <https://arxiv.org/abs/1906.12193>.
- [18] KHANAL A, ESTRADA R. Dynamic deep networks for retinal vessel segmentation [J/OL]. Frontiers in Computer Science, 2020, 2:35. [2021-10-01]. <https://www.frontiersin.org/articles/10.3389/fcomp.2020.00035/full>.

A Multilevel Feature Fusion Model for Retinal Vessels Segmentation

LIU Yanlong, WU Cong

(School of Computer Science, Hubei Univ. of Tech., Wuhan 430068, China)

Abstract: Existing retinal vessel segmentation methods cannot capture more features effectively in the learning layers, and the low-level features are not fully utilized to improve the accuracy of retinal vessel segmentation. We propose a novel multilevel feature fusion network. First, we designed block with a continuous memory mechanism to increase the depth of the network to capture more features. On this basis, we proposed a multilevel feature fusion block to fuse the low-level features and the high-level features. In DRIVE and CHASE_DB1, the results show that our method achieves better performance than other popular deep learning methods.

Keywords: medical image processing; retinal vessel segmentation; continuous memory mechanism; multilevel feature fusion

[责任编辑: 张岩芳]