

# 基于条件深度卷积生成对抗网络的视网膜血管分割

蒋芸<sup>1</sup> 谭宁<sup>1</sup>

**摘要** 视网膜血管的分割帮助医生对眼底疾病进行诊断有着重要的意义. 但现有方法对视网膜血管的分割存在着各种问题, 例如对血管分割不足, 抗噪声干扰能力弱, 对病灶敏感等. 针对现有血管分割方法的缺陷, 本文提出使用条件深度卷积生成对抗网络的方法对视网膜血管进行分割. 我们主要对生成器的网络结构进行了改进, 在卷积层引入残差模块进行差值学习使得网络结构对输出的改变变得敏感, 从而更好地对生成器的权重进行调整. 为了降低参数数目和计算, 在使用大卷积核之前使用小卷积核对输入特征图的通道数进行减半处理. 通过使用 U 型网络的思想将卷积层的输出与反卷积层的输出进行连接从而避免低级信息共享. 通过在 DRIVE 和 STARE 数据集上对本文的方法进行了验证, 其分割准确率分别为 96.08%、97.71%, 灵敏度分别达到了 82.74%、85.34%,  $F$  度量分别达到了 82.08% 和 85.02%, 灵敏度比 R2U-Net 的灵敏度分别高了 4.82%、2.4%.

**关键词** 生成对抗网络, 残差网络, 视网膜血管分割, 条件模型, 卷积神经网络

**引用格式** 蒋芸, 谭宁. 基于条件深度卷积生成对抗网络的视网膜血管分割. 自动化学报, 2021, 47(1): 136–147

**DOI** 10.16383/j.aas.c180285

## Retinal Vessel Segmentation Based on Conditional Deep Convolutional Generative Adversarial Networks

JIANG Yun<sup>1</sup> TAN Ning<sup>1</sup>

**Abstract** The segmentation of retinal vessels is of significance for doctors to diagnose the fundus diseases. However, existing methods have various problems in the segmentation of the retinal vessels, such as insufficient segmentation of retinal vessels, weak anti-noise interference ability, and sensitivity to lesions, etc. Aiming to the shortcomings of existed methods, this paper proposes the use of conditional deep convolutional generative adversarial networks to segment the retinal vessels. We mainly improve the network structure of the generator. The introduction of the residual module at the convolutional layer for residual learning makes the network structure sensitive to changes in the output, as to better adjust the weight of the generator. In order to reduce the number of parameters and calculations, using a small convolution kernel to halve the number of channels in the input signature before using a large convolution kernel. By used the idea of a U-net to connect the output of the convolutional layer with the output of the deconvolution layer to avoid low-level information sharing. By verifying the method on the DRIVE and STARE datasets, the segmentation accuracy rate is 96.08% and 97.71%, the sensitivity reaches 82.74% and 85.34%, respectively, and the  $F$ -measure reaches 82.08% and 85.02%, respectively. The sensitivity is 4.82% and 2.4% higher than that of R2U-Net.

**Key words** Generative adversarial network (GAN), residual networks, retinal vessel segmentation, conditional models, convolutional neural networks (CNNs)

**Citation** Jiang Yun, Tan Ning. Retinal vessel segmentation based on conditional deep convolutional generative adversarial networks. *Acta Automatica Sinica*, 2021, 47(1): 136–147

收稿日期 2018-05-07 录用日期 2018-08-02

Manuscript received May 7, 2018; accepted August 2, 2018

国家自然科学基金 (61962054, 61163036), 2016 年甘肃省科技计划资助自然科学基金项目 (1606RJZA047), 2012 年度甘肃省高校基本科研业务费专项资金项目, 甘肃省高校研究生导师项目 (1201-16), 西北师范大学第三期知识与创新工程科研骨干项目 (nwnu-kjxcg-03-67) 资助

Supported by National Natural Science Foundation of China (61962054, 61163036), 2016 Gansu Provincial Science and Technology Plan Funded Natural Science Fund Project (1606RJZA047), 2012 Gansu Provincial University Basic Research Business Expenses Special Fund Project, Gansu Provincial University Graduate Tutor Project (1201-16), The third phase of the Northwest Normal University Knowledge and Innovation Engineering Research Backbone Project (nwnu-kjxcg-03-67)

本文责任编辑 黄庆明

血管作为血液循环流动的管道, 遍布人体的各个部位, 尤其是眼底中含有大量的动脉血管, 是人身体内唯独采用无损伤手段就能够直接观察到的较深层次的微血管系统, 为医生提供了有关眼睛状况和一般系统状态的丰富信息. 眼科医生可以检测到高血压和糖尿病引起的全身血管负荷增加以及视网膜血管疾病如视网膜静脉阻塞和视网膜动脉阻塞等血管结构异常的早期征兆, 血管和血管系统引起的疾病

Recommended by Associate Editor HUANG Qing-Ming

1. 西北师范大学计算机科学与工程学院 兰州 730000  
1. College of Computer Science and Engineering, Northwest Normal University, Lanzhou 730000

会造成患者失明. 随着技术的发展, 探索一种自动化方法分割视网膜血管已经被广泛研究. 从而辅助医师诊断、分析患者病情, 对眼底疾病做出尽早的预防和诊治, 可有效避免病变所引发的视觉损失. 目前在国内外的血管分割算法有匹配滤波器法<sup>[1]</sup>、多阈值的血管检测<sup>[2]</sup>、以形态学为基础技术血管分割<sup>[3]</sup>、区域生长法、使用神经网络的血管分割算法<sup>[4-5]</sup>、多尺度层分解和局部自适应阈值血管分割方法<sup>[6]</sup>、基于活动轮廓模型的血管分割<sup>[7]</sup> 以及基于模糊聚类的分割方法<sup>[8]</sup> 等. 文献 [9] 提出了一种基于多尺度 2D Gabor 小波变换和形态学重构的血管分割方法. 采用不同尺度的 2D Gabor 小波对视网膜图像进行变换, 并分别应用形态学重构和区域生长法对变换后的图像进行分割, 最后, 对以上两种方法分割的视网膜血管和背景像素点重新标记识别, 得到视网膜血管最终分割结果, 但对于血管与背景对比度低的图片, 分割假阳性率高. 文献 [10] 提出一种融合区域能量拟合信息和形状先验的水平集血管分割方法, 通过形态学算子去除血管中心亮线, 并与原图像和掩模分别进行减法和点乘运算增强视网膜血管图像, 然后分析 Hessian 矩阵的特征值在血管、背景和病灶上不同的几何性质, 利用 Hessian 矩阵特征值重构血管响应函数, 最大化不同结构的差异, 从而获得视网膜血管初步图像, 但因为高斯卷积算子的影响, 导致复杂的血管交叉处分割不足. 文献 [11] 提出通过三阶段对血管进行分割, 首先, 对眼底图像的绿色平面进行预处理, 并针对血管区域中提取二值图像, 然后, 使用基于像素邻域一阶和二阶梯度图像提取的 8 个特征集合, 使用高斯混合模型分类器对两个二值图像中的所有其余像素进行分类, 血管的主要部分与分类的血管像素结合, 但存在微血管易断的问题. 文献 [12] 使用将分割任务重塑为从视网膜图像到血管图的跨模态数据转换的分割监督方法, 提出了一种具有较强诱导能力的广泛深度神经网络模型, 并给出了一种有效的训练策略. 文献 [13] 一种将异构情境感知功能与区分性学习框架相结合的方式对眼底血管图片进行分割. 文献 [14] 使用监督分割

技术, 该技术使用在全局对比度归一化, 零相位白化以及使用几何变换和校正进行预处理的样本上训练深度神经网络, 对噪声有较强的适应能力, 但依旧无法克服微血管易断的问题, 分割的准确率也有待提高. 综上所述, 虽然国内外研究人员已提出很多血管分割方法, 但大部分分割的结果精度不高, 性能也有很大的改善空间. 针对现有方法对视网膜血管分割不足, 准确率不高, 对噪声、病灶敏感等问题, 于是使用生成式思想分割出高精度、高准确率的视网膜血管图像. 本文主要工作包括:

1) 对于之前所提出的方法中存在的局限性进行了详细的分析后, 针对之前方法对视网膜血管分割效果不理想的问题, 本文使用生成式框架对眼底视网膜血管进行分割, 通过生成器和判别器两者之间进行对抗训练, 相互优化各自的网络结构, 从而实现了一种自动分割视网膜血管的方法.

2) 提出了一个新的网络结构如图 5, 利用残差网络的思想, 使得网络结构对输出的变化和权重的改变变得敏感, 从而更好地对权重进行了调整, 提高分割的效果, 也缓解了梯度减少, 解决了梯度消失的问题. 在使用  $3 \times 3$  的卷积之前使用  $1 \times 1$  的卷积作为瓶颈层对输入的特征图进行降维, 然后用  $3 \times 3$  的卷积核为维度进行还原, 降低了网络的复杂度, 保持了精度又减少了计算量.

3) 由于编码器-解码器这种网络结构要求所有的信息流通过所有的网络层, 在输入和输出之间共享大量本可直接穿过网络层的低级信息. 为了避免出现这种情况, 本文使用了 U 型网络结构的思想, 对于  $n$  层的网络, 将第  $i$  层的输出和第  $n-i$  层的输出进行拼接, 作为第  $n-i+1$  层的输入.

## 1 生成对抗

近些年来, 无监督学习已经成为了研究的热点, 变分自编码器<sup>[15]</sup>、生成对抗网络 (Generative adversarial networks, GAN)<sup>[16]</sup> 等无监督模型受到越来越多的关注. 在人工智能高速发展的时代, GAN 的提出不仅满足了相关领域的研究和应用和需求,

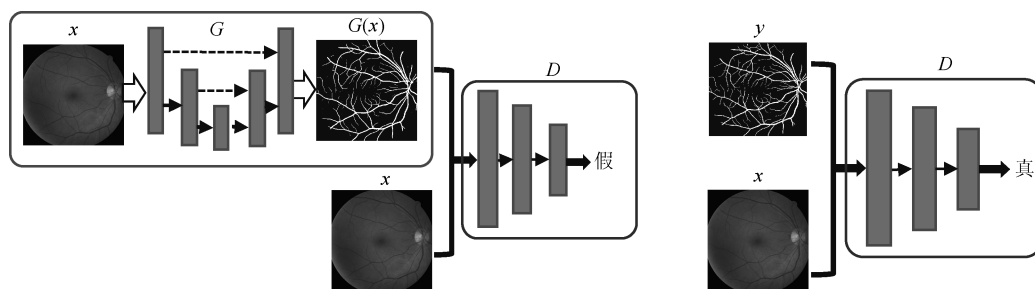


图 1 视网膜血管图像分割模型

Fig. 1 Retinal vessels image segmentation model

也带来了新的发展动力<sup>[17]</sup>. 特别是在图像和视觉领域中对 GAN 的研究和应用最为广泛, 已经可以通过随机数字生成人脸、从低分辨率图像生成高分辨率图像等. 此外, GAN 已经开始逐渐应用到医学图像处理中: 模拟超声探头的空间位置上有条件地对解剖学精确的图像进行采样<sup>[18]</sup>、检测恶性前列腺癌<sup>[19]</sup> 等问题的研究中.

### 1.1 生成对抗网络原理

生成对抗网络由生成器 ( $G$ ) 和判别器 ( $D$ ) 两部分组成, 生成器主要负责通过随机噪声生成服从真实数据分布的样本, 判别器负责通过识别输入的数据样本来自生成器还是真实数据, 两者相互迭代优化提升各自的性能, 最终使得判别器无法判别输入的数据来源时, 则认为生成器已经学习到了真实数据的分布. 其损失函数为:

$$\max_D \min_G (\theta_G, \theta_D) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\lg D(x)] + \mathbb{E}_{x \sim p_z(z)} [\lg(1 - D(G(z)))] \quad (1)$$

### 1.2 条件生成对抗网络

条件生成对抗网络 (Conditional generative adversarial networks, CGAN) 是在 GAN 的基础上加入了辅助信息  $y$ , 通过额外的辅助信息  $y$  控制生成器对数据的生成.  $y$  可以是任何类型的辅助信息, 比如类标签. 其模型结构如图 2 所示:

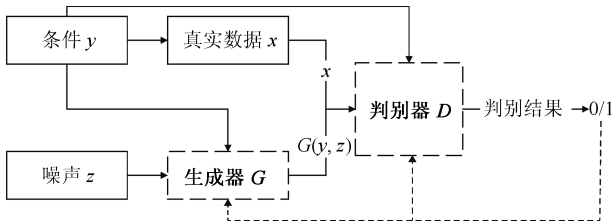


图 2 条件生成对抗网络模型

Fig. 2 Condition generation adversarial networks model

损失函数表示为:

$$\max_D \min_G \mathcal{L}(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\lg D(x|y)] + \mathbb{E}_{z \sim p_z(z)} [\lg(1 - D(G(z|y)))] \quad (2)$$

## 2 基于条件深度卷积生成对抗网络的视网膜血管分割

本文通过结合深度卷积生成对抗网络 (Deep convolutional generative adversarial networks, DCGAN)<sup>[20]</sup> 和 U 形卷积神经网络<sup>[21]</sup> 的特点, 在生成器使用 U 型卷积神经网络模型的思想, 利用 U 型卷积网络的边缘检测能力对视网膜图像中的血管进行分割, 将分割出的视网膜血管图像和原

图像一同输入判别器中进行判断, 最后直到判别器无法正确区分输入的视网膜血管图像的来源 (专家分割或者生成器分割).

生成对抗神经网络的学习优化过程是寻找到生成器和判别器之间的一个纳什均衡, 生成对抗网络分割眼底视网膜问题的目标函数如下所示:

$$\mathcal{L}_{cGAN}(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x), y} [\lg D(x, y)] + \mathbb{E}_{x \sim p_{\text{data}}(x)} [\lg(1 - D(x, G(x)))] \quad (3)$$

其中,  $x$  为输入的视网膜图像,  $y$  为专家分割的视网膜血管图像,  $D(x, y)$  表示视网膜血管图像来源于专家分割的概率,  $D(x, G(x))$  表示视网膜血管图像来自生成器分割的概率. 生成器 ( $G$ ) 尝试最小化目标函数, 判别器 ( $D$ ) 尝试最大化目标函数, GAN 通过对下面的函数进行优化:

$$G^* = \min_G \max_D \mathcal{L}_{cGAN}(G, D) \quad (4)$$

通过实验发现, 将现有的 GAN 目标函数与传统的损失 (例如  $L1$  距离函数) 相结合时, 生成出的视网膜血管图像更加趋近于专家分割的视网膜血管图像, 判别器的作用保持不变, 但生成器的任务不仅要欺骗判别器, 而且要最小化生成的视网膜血管图像与人工分割出的视网膜血管图像直接的  $L1$  距离.

$$\mathcal{L}_{L1}(G) = \mathbb{E}_{x \sim p_{\text{data}}(x), y} [\|y - G(x, y)\|] \quad (5)$$

通过将 GAN 的目标函数与  $L1$  距离函数进行结合, 最终的目标函数如下:

$$G^* = \arg \min_G \max_D \mathcal{L}_{cGAN}(G, D) + \lambda \mathcal{L}_{L1}(G) \quad (6)$$

$\lambda$  用于平衡两个目标函数, 推荐  $\lambda$  的值为 100.

## 3 模型结构

分割视网膜血管的模型如图 1 所示, 通过输入一张视网膜图像  $x$  到生成器 ( $G$ ) 中, 生成器通过学习到从眼底视网膜图像  $x$  到眼底血管图像  $y$  之间的映射关系,  $G: x \rightarrow y$  最终输出一张分割后的视网膜血管图像. 判别器 ( $D$ ) 通过学习输入图像对  $\{x, y\}$  与  $\{x, G(x)\}$  之间的分布差异从而正确判别输入图像对来源的二分类器  $\{0, 1\}$ , 判别器输出 1 表示输入的视网膜血管图像来源于人工分割, 输出 0 表示视网膜血管图像来源于生成器.

### 3.1 残差思想

在一定程度上, 随着网络结构深度的增加, 会出现梯度消失和梯度爆炸的问题使模型出现退化的情况<sup>[22]</sup>, 为了避免这情况, 通过使用跳跃连接 (Skip-connection), 将某一层的特征图反馈给下一层, 甚

至更深层, 其过程通过如下公式所示:

$$\mathcal{H}(x) = x^l + \sum_{k=l}^L \mathcal{F}(x^k) \quad (7)$$

$$\mathcal{F}(x^k) = f(C_0^k, C_1^k, \dots, C_c^k) \quad (8)$$

$$C_m^k = \sum_{n \in \mathcal{X}_m} x_i^{k-1} * k_{nm}^l + b_m^l \quad (9)$$

其中,  $\mathcal{H}(x)$  代表下一个残差模块的输入,  $L$  代表跳跃的层数,  $\mathcal{F}(x^k)$  代表对  $C^k$  进行激活后的输出,  $C_m^k$  代表第  $k$  层的特征图的第  $m$  个通道,  $f(\cdot)$  称为激活函数,  $\mathcal{X}_m$  表示用于计算  $C_i^k$  的输入特征图子集,  $k_{nm}^l$  代表卷积核矩阵,  $*$  代表卷积。

图 3(a) 是跳跃两层 ( $L = 2$ ) 卷积层的残差模块, 其函数表达式为:  $H(x) = x + F(x)$ , 图 3(b) 是标准的卷积层, 其函数表达式为:  $H(x) = F(x)$ 。

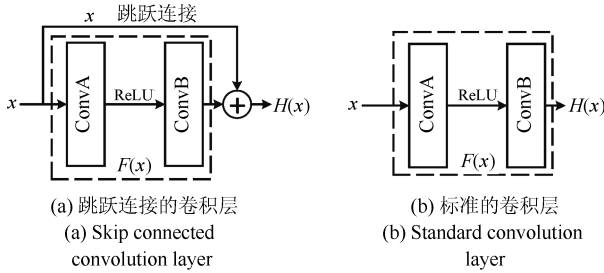


图 3 卷积层

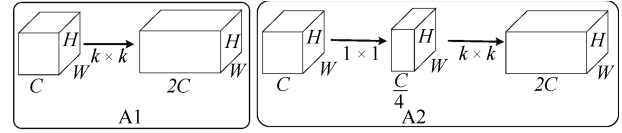
Fig. 3 Convolution layer

残差模块与标准的卷积模块相比较具有如下优势: 1) 在前向传播中, 假如图 3(a) 中卷积层中的权重没有学习到任何信息, 相当于进行了恒等变换, 如果卷积层学到了一些有用信息, 那么可能比学习恒等函数表现更好. 2) 在反向传播中, 引入残差模块后对输出的变化更敏感, 可以比标准的卷积层更细微的调整权重。

### 3.2 瓶颈层

随着网络层数的增加, 使用标准的卷积层将要训练大量的参数, 模型的参数越多, 运行所需要的计算资源和存储资源越多, 且会增加训练的时间, 为了在减少参数的情况下不影响模型的效果, 因此通过将标准的卷积层和  $1 \times 1$  卷积相结合如图 4(b) 所示, 通过中间产生一个瓶颈层, 达到减少参数和计算量, 加快网络模型的收敛。

模块 A1 将一个  $H \times W \times C$  大小的特征图作为输入, 直接输出一个大小为  $H \times W \times 2C$  的特征图. 模块 A2 通过在中间加入一个卷积核为  $1 \times 1$  的卷积层作为瓶颈层, 先缩小特征图的通道数, 然后再输出大小为  $H \times W \times 2C$  的特征图.  $H$  和  $W$  为输入和输出的特征图的尺寸.  $C$  代表输入和输入特征图的通道数。



(a) 标准的卷积神经网络模块 (b) 改进后的卷积层  
(a) Standard convolution layer (b) Improved convolution layer

图 4 卷积层

Fig. 4 Convolution layer

模块 A1 的计算量为:

$$C_S = k \times k \times C \times 2C \times H \times W \quad (10)$$

参数的数量为:

$$G_S = k \times k \times C \times 2C \quad (11)$$

加入  $1 \times 1$  卷积核后的模块 A2 的计算量为:

$$C_A = 1 \times 1 \times C \times \frac{C}{4} \times H \times W + k \times k \times \frac{C}{4} \times 2C \times H \times W \quad (12)$$

其中,  $k$  为卷积核的尺寸, 参数的数量为:

$$G_A = 1 \times 1 \times C \times \frac{C}{4} + k \times k \times \frac{C}{4} \times 2C \quad (13)$$

使用结合  $1 \times 1$  卷积核后的模块 A2 替换标准网络结构中的模块 A1, 计算量比值以及参数数量的比值均为:

$$R = \frac{C_A}{C_S} = \frac{G_A}{G_S} = \frac{1 + 2k^2}{8k^2} \quad (14)$$

当使用  $3 \times 3$  大小的卷积核时 ( $k = 3$ ), 模块 A2 计算量和参数大概是模块 A1 的 5/18。

### 3.3 生成器网络结构

生成器中使用  $x \in \mathbf{R}^{w \times h \times c}$  作为输入图像,  $w = n = 512, c = 3$ , 网络结构 (27 层) 如图 5 所示, 之前许多分割方法的网络结构使用编码器-解码器 (Encoder-decoder)<sup>[23]</sup>, 这种网络结构通过不断地向下采样, 直到达到一个瓶颈层, 将提取的信息变为一个一维向量, 然后在这一点上进行反卷积, 逐渐向上采样, 最后还原成图像. 这样的网络结构要求所有的信息流通过所有的网络层, 包括瓶颈层. 在许多图像分割问题中, 输入和输出之间共享大量本可直接穿过网络层的低级信息. 为了使生成器能够避免出现这种情况, 本文使用了 U 型网络结构的思想, 对于  $n$  层的网络结构, 将第  $i$  层网络的输出和第  $n-i$  层的输出进行连接, 作为第  $n-i+1$  层节点的输

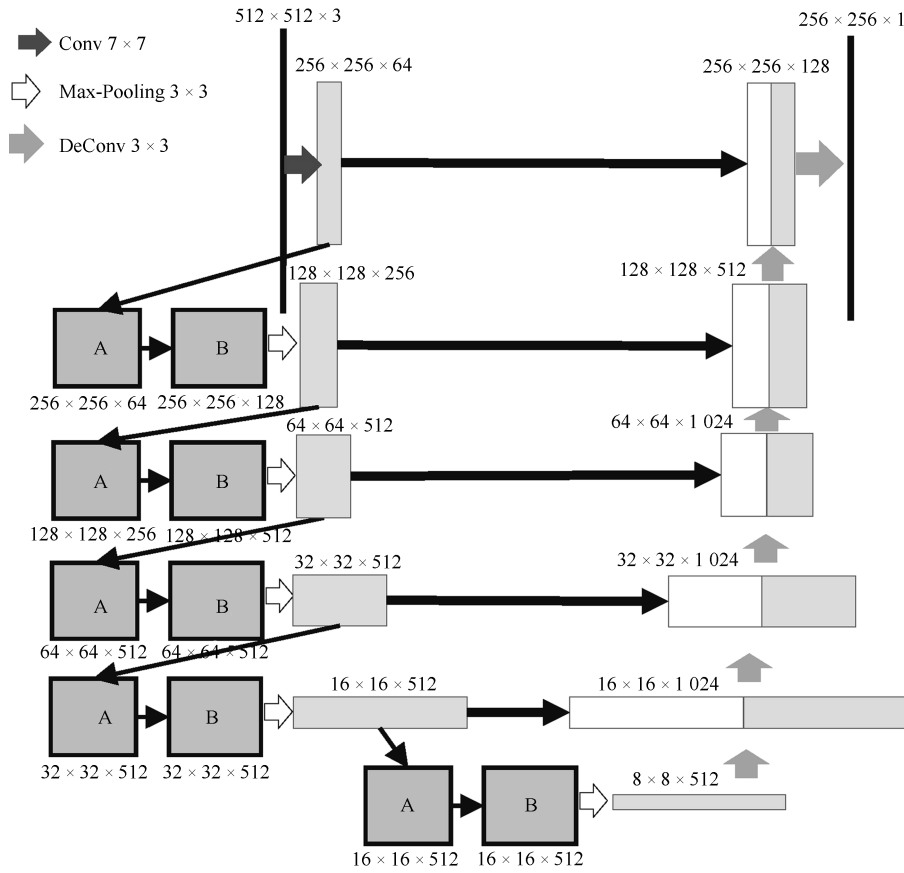


图5 生成器网络结构

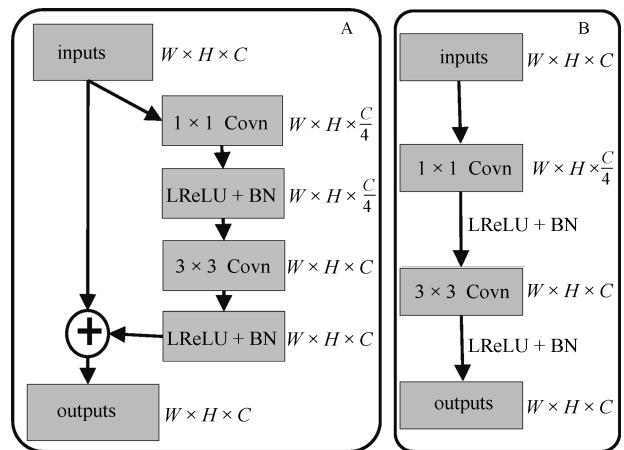
Fig. 5 Generator network structure

入. 在编码层使用 Leaky-ReLU 作为非线性激活函数, 每层都使用批量归一化 (Batch normalization, BN)<sup>[24]</sup>. 通过归一化当前层输入 ( $\mu = 0, \sigma = 1$ ), 有利于加速整个网络的收敛, 提高层与层之间的独立性. 解码层使用 ReLU 作为非线性函数, 在最后一层使用 Tanh 激活函数生成图像. 为了提高模型分割的准确率, 我们在生成器中加入残差网络结构, 结构如图 6 (a) 所示, 由于残差可以缓解梯度减少, 解决梯度消失的问题, 提高对网络结构对权重改变的敏感度, 使得生成器能够充分学习到视网膜血管图像的分布, 从而提高分割的效果. 为了降低模型的复杂度, 减少计算量和训练的参数数量, 在每次使用  $3 \times 3$  的卷积核之前加入  $1 \times 1$  卷积核作为瓶颈层对输入的特征层进行降维. 将通道数目降为原来的一半, 然后通过  $3 \times 3$  卷积核恢复到原来的通道数, 具体结构如图 6 (b) 所示.

3.4 判别器网络结构

判别器的网络结构如图 7 所示, 总共 7 层, 使用  $\{x_i, y_i\}_{i=1}^N$  或者  $\{x_i, G(x_i)\}_{i=1}^N (x_i \in X, y_i \in Y)$  作为输入,  $x \in \mathbf{R}^{512 \times 512 \times 3}, y \in \mathbf{R}^{512 \times 512 \times 1}$ . 对于输入为  $\{x_i, y_i\}$  时, 判别器正确的输出为 1, 输入

$\{x_i, G(x_i)\}$  判别器正确的输出为 0. 在判别器中的每一个卷积层中使用 Leaky-ReLU 作为非线性激活函数, 使用带步幅 (Strided convolutions) 的卷积代替池化层 (Pooling) 的效果.



(a) 残差卷积单元 (a) Forward convolutional units (b) Residual convolutional unit

图6 卷积的不同变体

Fig. 6 Different variants of convolutional

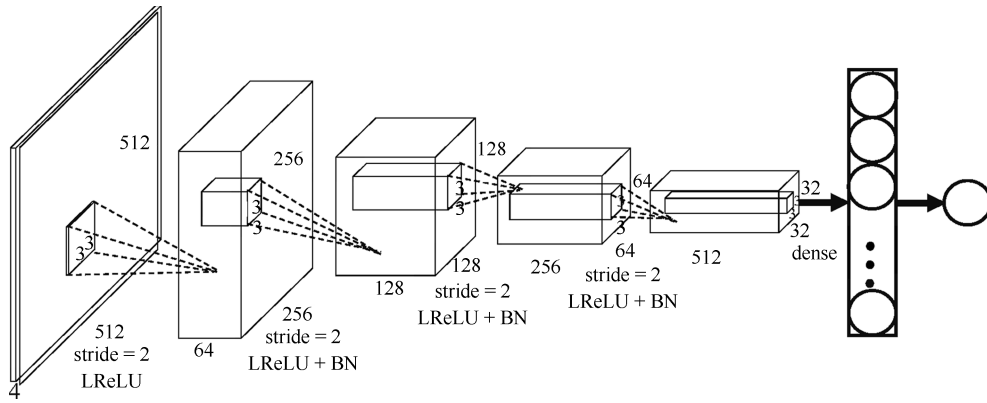


图 7 判别器网络结构

Fig. 7 Discriminator network structure

### 4 训练

本论文工作基于深度卷积对抗神经网络 (DCGAN) 和条件生成对抗神经网络 (CGAN)<sup>[25]</sup> 实现, 使用 DCGAN 中推荐的训练参数进行训练, 训练时使用 Adam 优化方式进行梯度下降 ( $\beta_1 = 0.5$ ,  $\beta_2 = 0.999$ ,  $\varepsilon = 10^{-8}$ ), 学习率为  $lr = 0.0002$ ,  $mini-batch = 1$ , 卷积提取特征时使用 Leaky-ReLU 作为激活函数, Leak 斜率  $\alpha$  设置为 0.2, 反卷积还原图像时, 使用 ReLU 激活函数. 训练时总共迭代 800 个周期, 输出的分割图像的大小为  $512 \times 512 \times 1$ .

生成器的训练过程如图 8 所示, 主要通过两个途径调整权重: 1) 通过比较生成器分割出的视网膜图像  $G(x)$  与专家分割出的血管图像  $y$  之间的差值, 通过差值调整权重, 使  $\|y - G(x)\|$  的值最小化; 2) 将视网膜图像与生成器分割出的血管图像  $G(x)$  输入到判别器中, 根据判别器输出的结果  $D(G(x), y)$  与 1 之间的差值对生成器的权重进行调整. 通过上面两种方式对生成器权值的调整, 使生成器分割出的视网膜血管图像更加接近专家分割后的血管图像.

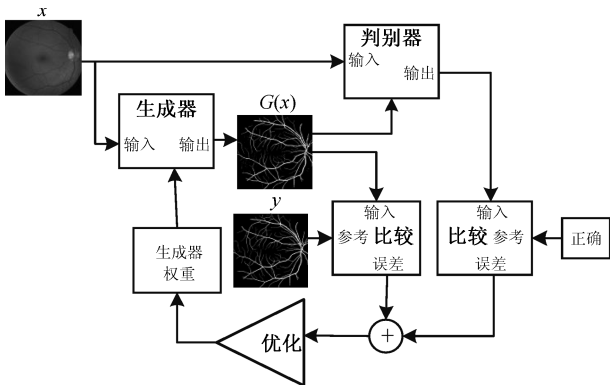


图 8 生成器训练的过程

Fig. 8 Generative training process

判别器的训练过程如图 9 所示, 也是通过两个

途径调整权重: 1) 输入视网膜图像和专家分割的视网膜血管图像, 通过比较判别器输出的值  $D(x, y)$  与标准值 1 直接差值调整权重, 使判别器输出的值更加接近 1; 2) 输入视网膜图像和生成器分割出的血管图像  $G(x)$ , 比较判别器输出的结果  $D(G(x), y)$  与 0 之间的差值, 通过差值调整权重. 通过这两种方式对权值的调整, 使得判别器能更准确地区分视网膜血管图像是通过专家分割的还是来自生成器分割.

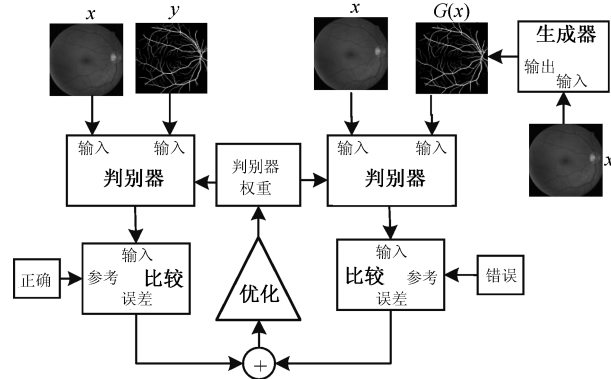


图 9 判别器训练的过程

Fig. 9 Discriminator training process

### 5 实验结果与分析

我们在两个公开的眼底视网膜数据集 (STARE<sup>[9]</sup> 和 DRIVE<sup>[26]</sup>) 上对我们的方法进行了测试, 在 DRIVE 数据集中, 总共包含 40 张视网膜图像和对应的视网膜血管图像, 用前 20 张图像做训练样本集, 后 20 张图像做测试样本集, 我们在第一类视网膜血管标签图像上进行训练和测试. 在 STARE 数据集中我们用前 10 张图像做训练样本集, 后 10 张图像当测试集. 由于数据集中只有少量的样本集对网络结构进行训练, 但神经网络对参数的训练需要大量的训练样本, 因此我们通过对图像的旋转、水平翻转、垂直翻转、平移变换、图像

灰度值的变化等方法对训练数据集进行扩充,这对于提高分割的准确率、防止过拟合和网络的鲁棒性至关重要。

### 5.1 性能评价指标

为了评估分割的效果,我们通过对灵敏性 (*Sensitivity*), 特效性 (*Specificity*), 准确率 (*Accuracy*), *F* 度量 (*F-measure*) 等评估指标进行了比较。

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \quad (15)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (16)$$

$$Specificity = \frac{TN}{TN + FP} \quad (17)$$

$$Precision = \frac{TP}{TP + FP} \quad (18)$$

$$Recall = \frac{TP}{TP + FN} \quad (19)$$

$$F\text{-measure} = 2 \times Precision \times \frac{Recall}{Precision + Recall} \quad (20)$$

其中, *TP* 为被正确分割的血管像素的数目, *TN* 为正确分割的背景像素数目, *FP* 为错误分割为血管像素的背景像素, *FN* 为被错误标记为背景像素的血管像素。

### 5.2 模型改进前后结果的对比

为了验证引用的结构对改进模型的有效性,我们对本文所使用的网络结构在引用前和引用后对视网膜血管分割的效果进行了对比,主要对比了如表 1

中的 4 种结构: 1) 不做任何改进的情况下,单独使用 U 型网络结构. 2) 生成对抗网络结构中的生成器使用 U 型网络结构. 3) 在 U 型网络结构中加入残差模块. 4) 生成对抗网络中的生成器中使用结合了残差模块的 U 型网络结构. 通过对这 4 种方式进行实验,结果表明第 4 种结构的性能最好. 因此可以看出,残差模块对权重改变的敏感性能对梯度进行更好的调整,以及生成对抗网络模型中判别器与生成器的相互对抗,相互学习,使得生成器的网络结构能够对视网膜血管进行更准确的分割。

为了证明第 3.2 节中使用瓶颈层减少参数后对模型准确率的影响,我们通过对分别使用图 4(a) 的结构和使用图 4(b) 中的结构后模型对眼底视网膜图像分割后的结果进行对比. 通过实验可以证明 (表 2), 使用  $1 \times 1$  卷积核的瓶颈层在减少参数和计算量的情况下基本不会影响模型对视网膜血管图像分割的准确率。

### 5.3 不同分割算法结果的对比

我们分别在 STARE、DRIVE 数据集上进行了大量的血管分割实验来证明方法的可行性,图 10 和图 11 分别展示了在 DRIVE 数据集上和 STARE 数据集上文献 [14]、文献 [27] 与本文方法对其中两张视网膜图像的分割结果的对比,在图 10 和图 11 中, (a) 为原始视网膜图像, (b) 为专家分割的视网膜血管图像, (c) 和 (d) 分别为文献 [14] 和文献 [27] 的分割结果,其中图 11 (a) 中第一张和第二张分别为健康人,青光眼视网膜图像,尽管图 11 (a) 的第二张图像中有血管有病变,存在病灶,但本文的方法可以较好的克服这些因素,分割出一张准确的血管图像,从图 10 和图 11 中的 (c) 和 (d) 中可以看出文献 [14] 和文献 [27] 分割的视网膜血管图像中包含了大量的噪声,易受病变视网膜图像中病灶的影响,对细小的

表 1 模型改进前后分割的结果

Table 1 The segmentation results before and after model improvement

数据集	方法	<i>F</i> 度量	准确率
DRIVE/STARE	U-net	0.8142/0.8373	0.9531/0.9690
	GAN+U-net	0.8150/0.8398	0.9583/0.9710
	U-net+Residual	0.8149/0.8388	0.9553/0.9700
	GAN+U-net+Residual (本文结构)	<b>0.8208/0.8506</b>	<b>0.9608/0.9771</b>

表 2 使用瓶颈层前后分割的结果

Table 2 The result of segmentation before and after using the bottleneck layer

数据集	方法	参数	计算量 ( <i>GFLOPS</i> )	<i>F</i> 度量	准确率
DRIVE/STARE	No Bottleneck	19.8 M	183.8	0.8210/0.8504	0.9612/0.9772
	Bottleneck	5.2 M	48.5	0.8208/0.8502	0.9608/0.9771

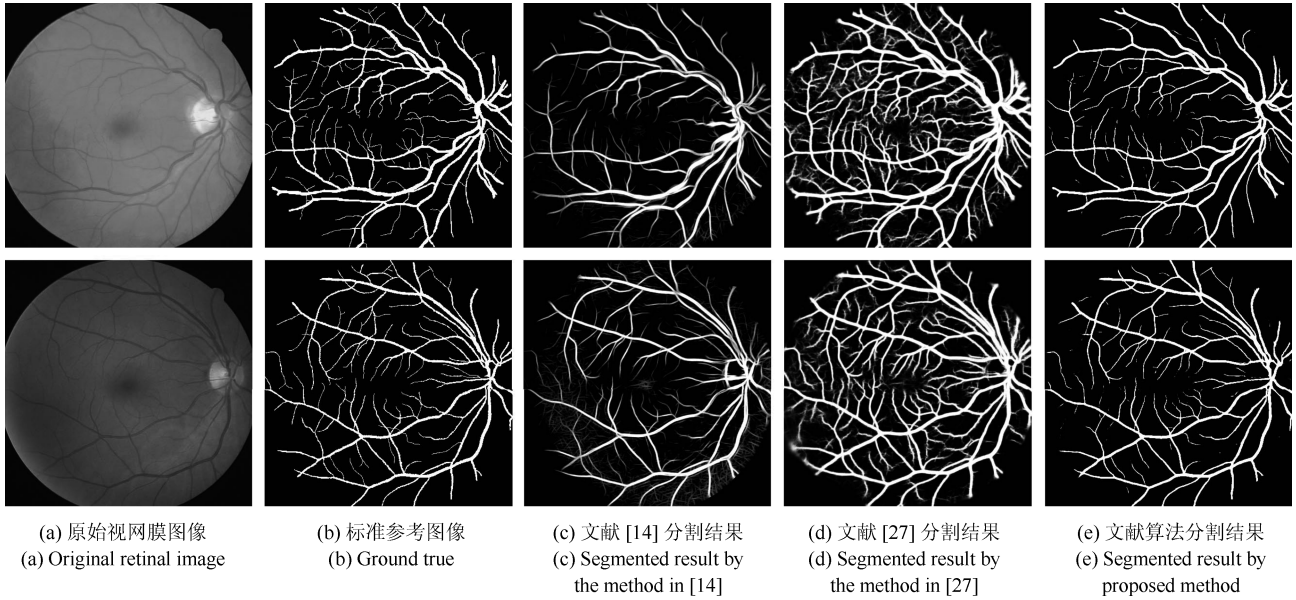


图 10 DRIVE 数据库视网膜血管分割结果比较

Fig. 10 Comparisons of segmentation results on DRIVE database

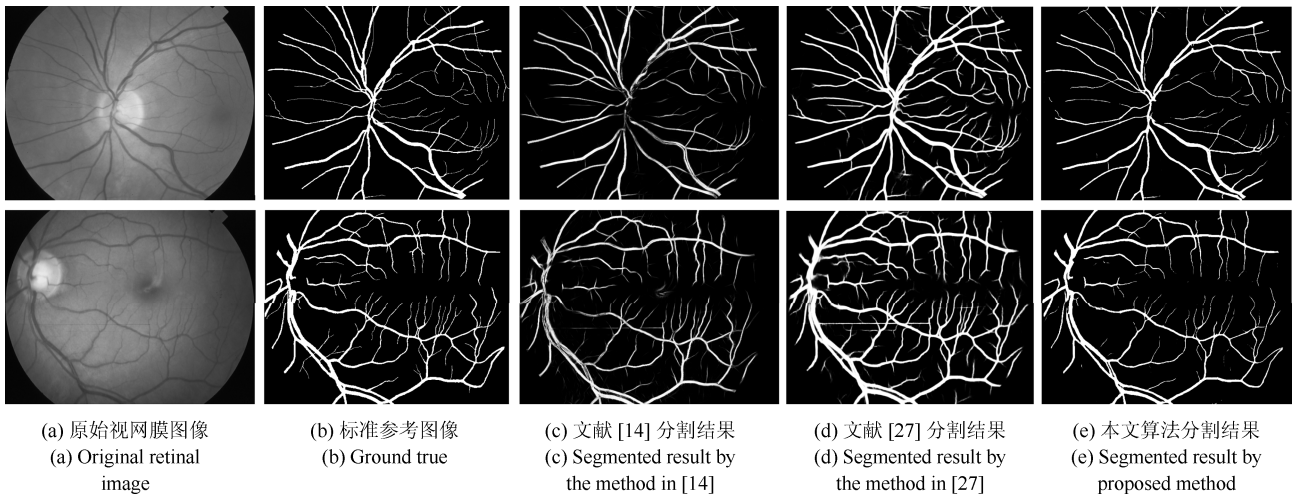


图 11 STARE 数据库视网膜血管分割结果比较

Fig. 11 Comparisons of segmentation results on STARE database

血管分割比较模糊, 且 (d) 中分割出的血管尺寸比标准图像中血管尺寸偏大. 这些细小的血管对于分析视网膜疾病有着重要的意义. 为了对细小的血管进行准确的分割, 本文使用了生成对抗网络, 通过判别器不断地对生成器进行优化, 两者相互对抗学习, 提高模型的分割效果. 并且生成器中使用的 U 型网络结构克服了无法对细小的血管进行分割准确分割的缺陷. 与以上两种方法相比可知, 本文方法所分割出的眼底血管图像中的噪声水平更低, 对细小的血管分割更加全面, 且比较清晰, 有更好的鲁棒性和精度.

为了更加突出本文方法分割视网膜血管的优势,

我们对 STARE 数据集中编号为 im0240 的视网膜图像使用不同方法分割后的结果进行局部放大, 如图 12 所示. 其中图 12 (a)、(c)、(e)、(g)、(i) 分别为文献 [14]、文献 [27]、文献 [28]、本文方法和专家分割的结果, 图 12 (b)、(d)、(f)、(h)、(j) 为与之对应的对右下角部分进行放大后的图像, 图中的白色圆圈表示不准确的分割, 字母 A 和字母 B 分别表示邻近的血管和细小的血管. 比较图 12 (a)、(c)、(e)、(g) 可看出, 本文的方法分割出的血管更加的清晰, 血管的尺寸也比较合适, 包含的噪声更少. 从图 12 (a) 中可以看到文献 [14] 分割出的血管比较虚幻, 且不能完整的分割, 含有大量的噪声. 图



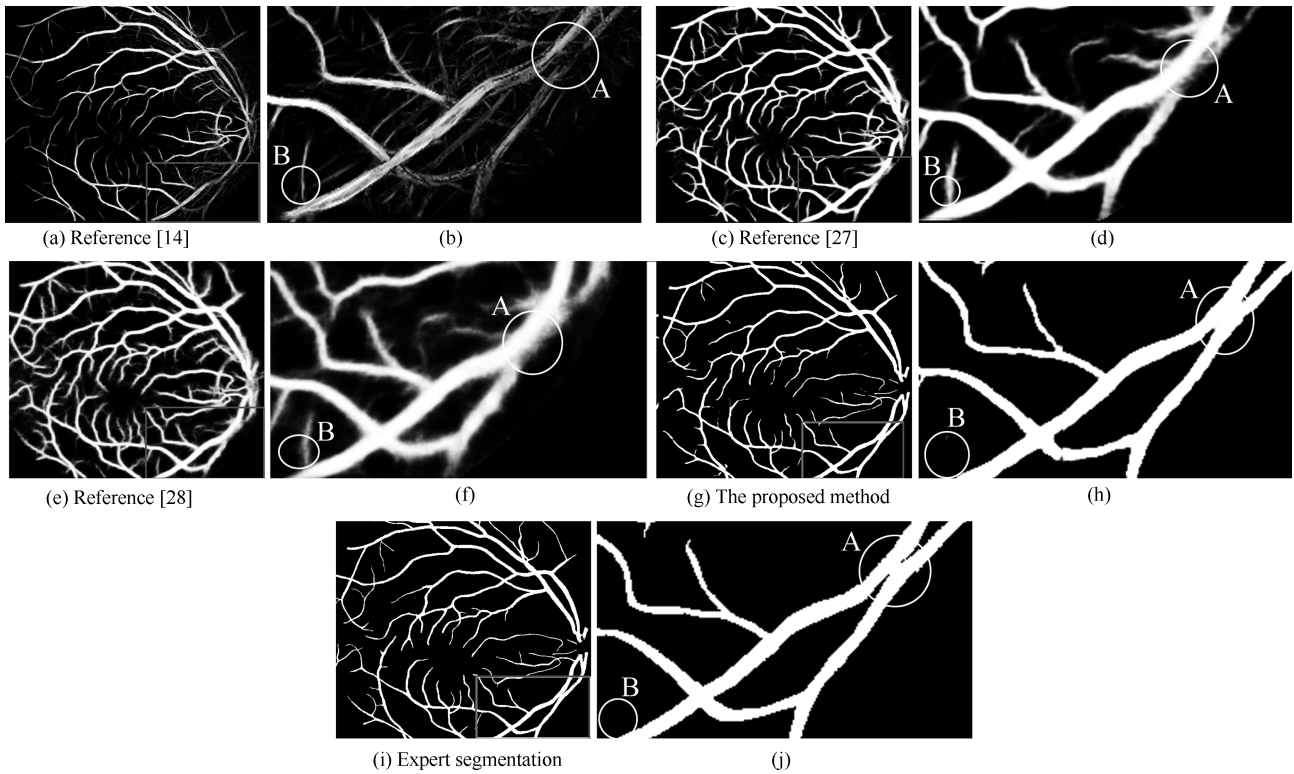


图 12 不同算法的视网膜血管分割局部放大图

Fig. 12 Different methods of partial retinal vessel segmentation

12(b) 为文献 [27] 分割的结果, 噪声比前面两种方法要少, 分割的结果相对要好, 但分割效果还是不理想. 从图 12(f) 中可以看到文献 [28] 的分割的血管比实际的血管要大, 血管周围出现了大量的噪声使血管比较模糊. 在图 12(b)、(d)、(f) 中可以看到其他的方法在 A 处对血管分割不全, 无法进行准确的分割, 在 B 处都出现错误的分割. 为了减少这种错误的分割我们加入残差模块, 利用其对权重改变的敏感性使得生成器能够更加充分地学习到眼底视网膜图像与眼底血管图像之间的映射关系, 提取更多用于分割血管的特征, 从而进行更加准确的分割, 减少误分割和漏分割, 从图 12(h) 可以看到, 本文分割出的视网膜血管图像中基本上没有噪声、更加清晰, 血管的大小也接近标准大小, 含有的错误分割少, 更加接近标准的视网膜血管图像.

#### 5.4 不同分割算法结果的评估

为了进一步证明本文算法对视网膜血管分割的有效性, 在 STARE、DRIVE 数据集上将本文的方法分别与文献 [5]、文献 [6]、文献 [12]、文献 [13]、文献 [28–33] 中的方法在敏感性、特效性、准确率、 $F$  度量等几个指标对视网膜血管分割的性能进行了比较. 其中灵敏度表示正确分割出的血管像素占真实血管像素的百分比, 特异性为正确分割的背景像素占真实背景像素的百分比, 准确率为正确分割血管

和背景像素占整个图像的百分比,  $F$  度量表示精度和召回率直接的调和均值. 表 3 和表 4 分别给出了不同方法在 DRIVE 和 STARE 数据集上视网膜分割的准确率. 从表 3 和表 4 中可知, 本文算法在不同的数据集上灵敏度、准确率、 $F$  度量等都比文献 [5]、文献 [6]、文献 [12]、文献 [13]、文献 [29]、文献 [30] 要高. 虽然文献 [27]、文献 [28]、文献 [33] 的灵敏度比本文方法要高, 但是分割出的血管尺寸比实际的要大, 分割出的血管包含比较多的噪声, 且本文算法有最高的  $F$  度量, 灵敏性和特效性都保持相对稳定, 分割图像中含有的噪声比较少, 但对细小血管分割效果不是很好, 有时会出现断裂的情况. 对于 DRIVE 数据集, 本文方法对视网膜血管分割的  $F$  度量达到 82.08%, 比文献 [33] 高了 0.37%, 灵敏度比文献 [33] 高了 4.82%. 在 STARE 数据集上, 我们的方法在  $F$  度量、灵敏度上比文献 [33] 分别高了 0.27%、2.4%. 因此由表 3 和表 4 中血管分割的评估指标展示可知, 在 DRIVE 和 STARE 数据集上, 本文的方法对血管、背景的认识和提取不同特征能力均优于其他有监督血管分割方法.

图 13 给出了本文方法和其他方法的  $F$  度量性能评价曲线, 在 DRIVE 数据集上分别与文献 [14]、文献 [27]、文献 [28]、文献 [30] 进行了比较, 在 STARE 数据集上分别与文献 [27]、文献 [28]、文献 [30] 进行了比较. 从图 13 可以看出, 在 DRIVE 数

表 3 DRIVE 数据库视网膜血管分割结果

Table 3 Segmentation performance of retinal vessel on the DRIVE database

数据集	方法	年份	$F$ 度量	灵敏度	特效性	准确率
DRIVE	Chen <sup>[13]</sup>	2014	—	0.7252	0.9798	0.9474
	N <sup>4</sup> -Fields <sup>[30]</sup>	2014	0.7970	0.8437	0.9743	0.9626
	Azzopardi <sup>[5]</sup>	2015	—	0.7655	0.9704	0.9442
	Roychowdhury <sup>[12]</sup>	2016	—	0.7250	0.9830	0.9520
	Liskowsk <sup>[14]</sup>	2016	—	0.7763	0.9768	0.9495
	Qiaoliang Li <sup>[12]</sup>	2016	—	0.7569	0.9816	0.9527
	DRIU <sup>[27]</sup>	2016	0.6701	0.9696	0.9115	0.9165
	HED <sup>[28]</sup>	2017	0.6400	0.9563	0.9007	0.9054
	U-Net <sup>[33]</sup>	2018	0.8142	0.7537	0.9820	0.9531
	Residual U-Net <sup>[33]</sup>	2018	0.8149	0.7726	0.9820	0.9553
	Recurrent U-Net <sup>[33]</sup>	2018	0.8155	0.7751	0.9816	0.9556
	R2U-Net <sup>[33]</sup>	2018	0.8171	0.7792	0.9813	0.9556
本文方法	2018	<b>0.8208</b>	<b>0.8274</b>	<b>0.9775</b>	<b>0.9608</b>	

表 4 STARE 数据库视网膜血管分割结果

Table 4 Segmentation performance of retinal vessel on the STARE database

数据集	方法	年份	$F$ 度量	灵敏度	特效性	准确率
STARE	Marin <sup>[31]</sup>	2011	—	0.6940	0.9770	0.9520
	Fraz <sup>[32]</sup>	2012	—	0.7548	0.9763	0.9534
	Liskowsk <sup>[14]</sup>	2016	—	0.7867	0.9754	0.9566
	Roychowdhury <sup>[12]</sup>	2016	—	0.7720	0.9730	0.9510
	Qiaoliang Li <sup>[12]</sup>	2016	—	0.7726	0.9844	0.9628
	DRIU <sup>[27]</sup>	2016	0.7385	0.6066	0.9956	0.9499
	HED <sup>[28]</sup>	2017	0.6990	0.5555	<b>0.9955</b>	0.9378
	U-Net <sup>[33]</sup>	2018	0.8373	0.8270	0.9842	0.9690
	Residual U-Net <sup>[33]</sup>	2018	0.8388	0.8203	0.9856	0.9700
	Recurrent U-Net <sup>[33]</sup>	2018	0.8396	0.8108	0.9871	0.9706
	R2U-Net <sup>[33]</sup>	2018	0.8475	0.8298	0.9862	0.9712
	本文方法	2018	<b>0.8502</b>	<b>0.8538</b>	0.9878	<b>0.9771</b>

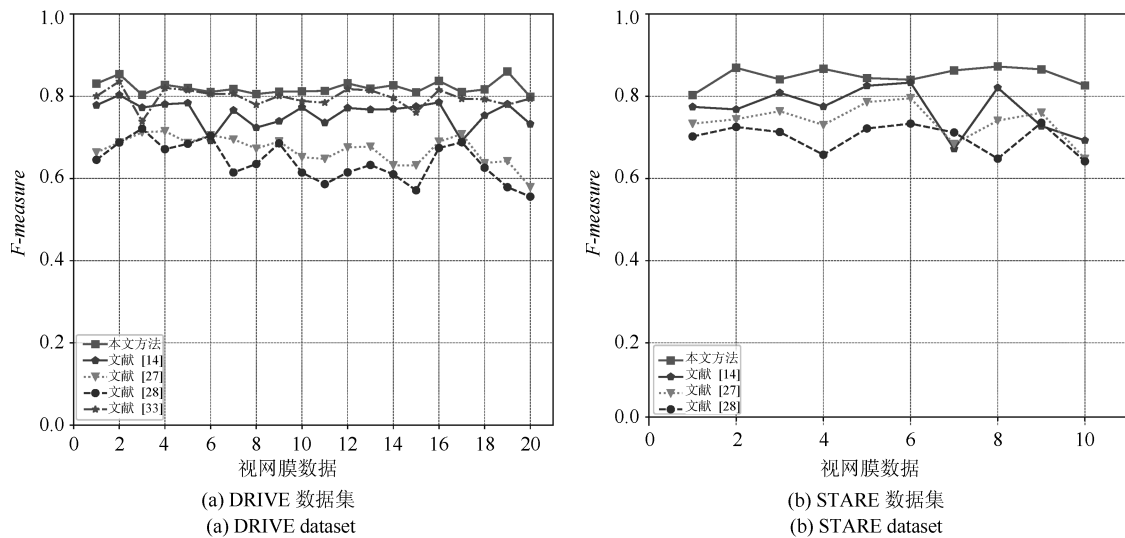


图 13 不同算法的  $F$  度量性能评价曲线

Fig. 13 Different methods of  $F$ -measure performance evaluation curve

据集和 STARE 数据集上, 本文方法比文献 [14]、文献 [27]、文献 [28]、文献 [30] 中的方法效果都要好, 且波动比较小, 模型的泛化性比较强, 对特征的提取和识别能力比其他方法更好, 能对各种眼底图像的血管进行很好的分割, 无论在健康还是病变的视网膜图像上, 都能进行保持相对稳定的分割效果, 能够很好地克服病灶的影响。

## 6 结论

视网膜血管的正确分割对帮助医生进行眼底疾病的诊断具有重大的实际意义. 本文使用条件深度卷积生成对抗网络对眼底视网膜进行分割, 在生成器中使用卷积对图像的特征进行了若干次提取后通过反卷积生成对应的视网膜的血管图像, 我们在生成器中加入残差网络模块, 由于残差网络对特征值的改变非常敏感性, 使得提取的特征更加准确. 为了降低模型的规模, 在每次进行  $3 \times 3$  的卷积之前通过  $1 \times 1$  的卷积进行降维, 在保证分割的准确性的情况下减少了参数的数量和计算量. 我们分别 DRIVE 和 STARE 数据集上对本文提出的方法的可行性进行了验证, 在 DRIVE 和 STARE 数据集上的准确率分别达到了 96.08% 和 97.71%,  $F$  度量分别达到了 82.08% 和 85.02%, 通过对分割出的视网膜血管图像进行分析和比较, 本文提出的方法与其他方法相比更具有优势.

## References

- Zhang B, Zhang L, Zhang L, Karray F. Retinal vessel extraction by matched filter with first-order derivative of Gaussian. *Computers in Biology and Medicine*, 2010, **40**(4): 438–445
- Jiang X Y, Mojon D. Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(1): 131–137
- Zana F, Klein J C. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation. *IEEE Transactions on Image Processing*, 2001, **10**(7): 1010–1019
- Mehrotra A, Tripathi S, Singh K K, Khandelwal P. Blood vessel extraction for retinal images using morphological operator and KCN clustering. In: Proceedings of the 2014 IEEE International Advance Computing Conference. Gurgaon, India: IEEE, 2014. 1142–1146
- Azzopardi G, Strisciuglio N, Vento M, Petkov N. Trainable COSFIRE filters for vessel delineation with application to retinal images. *Medical Image Analysis*, 2015, **19**(1): 46–57
- Wang Y F, Ji G R, Lin P, Trucco E. Retinal vessel segmentation using multiwavelet kernels and multiscale hierarchical decomposition. *Pattern Recognition*, 2013, **46**(8): 2117–2133
- Guo Z L, Lin P, Ji G R, Wang Y F. Retinal vessel segmentation using a finite element based binary level set method. *Inverse Problems and Imaging*, 2014, **8**(2): 459–473
- Tolias Y A, Panas S M. A fuzzy vessel tracking algorithm for retinal images based on fuzzy clustering. *IEEE Transactions on Medical Imaging*, 1998, **17**(2): 263–273
- Wang Xiao-Hong, Zhao Yu-Qian, Liao Miao, Zou Bei-Ji. Automatic segmentation for retinal vessel based on multi-scale 2D Gabor wavelet. *Acta Automatica Sinica*, 2015, **41**(5): 970–980  
(王晓红, 赵于前, 廖苗, 邹北骥. 基于多尺度 2D Gabor 小波的视网膜血管自动分割. 自动化学报, 2015, **41**(5): 970–980)
- Liang Li-Ming, Huang Chao-Lin, Shi Fei, Wu Jian, Jiang Hong-Jiu, Chen Xin-Jian. Retinal vessel segmentation using level set combined with shape priori. *Chinese Journal of Computers*, 2018, **41**(7): 1678–1692  
(梁礼明, 黄朝林, 石霁, 吴健, 江弘九, 陈新建. 融合形状先验的水平集眼底图像血管分割. 计算机学报, 2018, **41**(7): 1678–1692)
- Roychowdhury S, Koozekanani D D, Parhi K K. Blood vessel segmentation of fundus images by major vessel extraction and subimage classification. *IEEE Journal of Biomedical and Health Informatics*, 2015, **19**(3): 1118–1128
- Li Q L, Feng B W, Xie L P, Liang P, Zhang H S, Wang T F. A cross-modality learning approach for vessel segmentation in retinal images. *IEEE Transactions on Medical Imaging*, 2016, **35**(1): 109–118
- Cheng E, Du L, Wu Y, Zhu Y J, Megalooikonomou V, Ling H B. Discriminative vessel segmentation in retinal images by fusing context-aware hybrid features. *Machine Vision and Applications*, 2014, **25**(7): 1779–1792
- Liskowski P, Krawiec K. Segmenting retinal blood vessels with deep neural networks. *IEEE Transactions on Medical Imaging*, 2016, **35**(11): 2369–2380
- Kingma D P, Welling M. Auto-encoding variational Bayes. arXiv: 1312.6114, 2014
- Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative adversarial nets. In: Proceedings of the 27th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press, 2014. 2672–2680
- Wang Kun-Feng, Gou Chao, Duan Yan-Jie, Lin Yi-Lun, Zheng Xin-Hu, Wang Fei-Yue. Generative adversarial networks: the state of the art and beyond. *Acta Automatica Sinica*, 2017, **43**(3): 321–332  
(王坤峰, 苟超, 段艳杰, 林懿伦, 郑心湖, 王飞跃. 生成式对抗网络 GAN 的研究进展与展望. 自动化学报, 2017, **43**(3): 321–332)
- Hu Y P, Gibson E, Lee L L, Xie W D, Barratt D C, Vercauteren T, et al. Freehand ultrasound image simulation with spatially-conditioned generative adversarial networks. In: Proceedings of Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment. Québec, Canada: Springer, 2017. 105–115
- Kohl S, Bonekamp D, Schlemmer H P, Yaqubi K, Hohenfellner M, Hadaschik B, et al. Adversarial networks for the detection of aggressive prostate cancer. arXiv: 1702.08014, 2017.

- 20 Radford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv: 1511.06434, 2015.
- 21 Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. arXiv preprint arXiv: 1505.04597v1, 2015.
- 22 He K M, Sun J. Convolutional neural networks at constrained time cost. In: Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition. Boston, Massachusetts: IEEE, 2015. 5353–5360
- 23 Pathak D, Krähenbühl P, Donahue J, Darrell T, Efros A A. Context encoders: feature learning by inpainting. In: Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, Nevada: IEEE, 2016. 2536–2544
- 24 Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv: 1502.03167, 2015.
- 25 Mirza M, Osindero S. Conditional generative adversarial nets. arXiv: 1411.1784, 2014.
- 26 Staal J, Abrámoff M D, Niemeijer M, Viergever M A, Van Ginneken B. Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 2004, **23**(4): 501–509
- 27 Maninis K K, Pont-Tuset J, Arbeláez P, Van Gool L. Deep retinal image understanding. arXiv preprint arXiv: 1609.01103, 2016.
- 28 Xie S, Tu Z. Holistically-nested edge detection. *International Journal of Computer Vision*, 2017, **125**(1–3): 3–18
- 29 Soares J V B, Leandro J J G, Cesar R M, Jelinek H F, Cree M J. Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification. *IEEE Transactions on Medical Imaging*, 2006, **25**(9): 1214–1222
- 30 Ganin Y, Lempitsky V.  $N^4$ -fields: neural network nearest neighbor fields for image transforms. arXiv preprint arXiv: 1406.6558, 2014.
- 31 Marin D, Aquino A, Gegundez-Arias M E, Bravo J M. A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. *IEEE Transactions on Medical Imaging*, 2011, **30**(1): 146–158
- 32 Fraz M M, Remagnino P, Hoppe A, Uyyanonvara B, Rudnicka A R, Owen C G, et al. An ensemble classification-based approach applied to retinal blood vessel segmentation. *IEEE Transactions on Biomedical Engineering*, 2012, **59**(9): 2538–2548
- 33 Alom M Z, Hasan M, Yakopcic C, Taha T M, Asari V K. Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation. arXiv preprint arXiv: 1802.06955, 2018.



蒋芸 工学博士. 西北师范大学计算机科学与工程学院教授. 主要研究方向为数据挖掘, 粗糙集理论及应用.  
E-mail: jiangyun@nwnu.edu.cn  
(JIANG Yun Ph. D., professor at the College of Computer Science and Engineering, Northwest Normal University. Her research interest covers data mining, rough set theory and application.)



谭宁 西北师范大学计算机科学与工程学院硕士研究生. 主要研究方向为深度学习, 医学图像处理. 本文通信作者.  
E-mail: tanning2315@126.com  
(TAN Ning Master student at the College of Computer Science and Engineering, Northwest Normal University. His research interest covers deep learning, medical image processing. Corresponding author of this paper.)