

DOI: 10.13382/j.jemi.B2103943

基于 GAN 的少样本视网膜血管分割研究*

张思杰 方 翔 魏 赋

(重庆大学微电子与通信工程学院 重庆 400044)

摘 要:视网膜血管分割是自动筛查糖尿病视网膜病变的重要步骤,当前大部分深度学习方法都使用大样本进行网络训练,但医学领域带标签样本难以获取,且存在健康人样本与患者样本不平衡问题。提出了一种基于生成对抗网络的少样本视网膜血管分割方法,生成器部分对图像做反色等预处理后,通过旋转增扩充了数据集,网络部分使用 U-Net 结构,判别器部分使用卷积神经网络。在实验阶段,在 DRIVE 数据集和 HRF 数据集上进行训练测试,训练时只使用训练集的 6 个样本,测试时使用全部测试集样本,最终在两个数据集下的 ROC 曲线下面积分别达到了 0.97 和 0.95,准确率达到了 0.95 和 0.94。与少样本情况下的 U-Net 相比,分割性能提升很大,表明本方法针对少样本视网膜血管分割任务确实有效。

关键词: 视网膜血管分割;生成对抗网络;少样本学习;类不平衡

中图分类号: TP391;TH786 **文献标识码:** A **国家标准学科分类代码:** 520.2060;320.37

Research on retinal vascular segmentation based on GAN using few samples

Zhang Sijie Fang Xiang Wei Fu

(College of Microelectronics and Communication Engineering, Chongqing University, Chongqing 400044, China)

Abstract: Retinal vessel segmentation is an important step for automatic screening of diabetic retinopathy. Currently, most deep learning methods use a large number of labeled samples for network training, but it is difficult to obtain labeled samples in the medical field, and healthy samples and patient samples are imbalanced. In this paper, we have proposed a method of retinal vessel segmentation using few samples based on generating adversarial network. In the generator part, after preprocessing the image by inverting color and other methods, the dataset is expanded by rotation. The U-Net structure is used in the network part and the discriminator uses CNN network. In the experimental stage, the training test was applied to DRIVE dataset and HRF dataset. Only 6 samples of the training set were used in the training step, and all the test samples were used in the test step. Finally, the area under the ROC curve of the two datasets reached 0.97 and 0.95, and the accuracy rate reached 0.95 and 0.94. Compared with U-Net in condition of few samples, segmentation performance is improved greatly. It shows that this method is effective for the task of low sample segmentation in retina vessel.

Keywords: retina vessel segmentation; generation of adversarial network; few sample learning; class-imbalance

0 引 言

眼睛的深层血管只能以非侵入性的方式通过视网膜观察,所以视网膜图像是很重要的健康指标。而视网膜血管是在视网膜中检测到眼底图像进一步分割处理下得到的,其结构和特征变化能够反映了各类心血管疾病的程度,如白内障、糖尿病视网膜病变、高血压等^[1]。

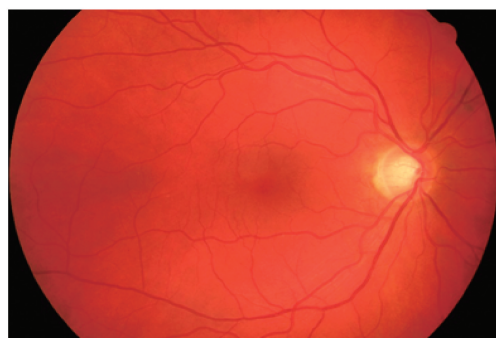
视网膜图像分析中的计算机辅助疾病诊断可以为大规模的筛查工作提供强有力的支持。随着深度学习的发展,越来越多成熟的深度学习方法进入医学图像处理领域,为视网膜图像处理提供了新的解决方法^[2]。但传统深度网络一般需求大量带标签样本,否则随之而来的是网络过拟合等问题,但在医学领域这类带标签样本难以获取,需要专业的医生诊断标注才能得到,耗时长开销大,且存在专家个体差异带来的不同标准^[3]。与此同时,

多数数据集中健康者的样本和患者的样本数量不平衡,从而对网络性能产生影响。而利用生成对抗网络(generative adversarial networks, GAN)的少样本血管分割技术则针对此问题提供了一种新的解决思路^[4]。

近年来关于眼底血管分割的深度学习方法层出不穷,2015 年, Ronneberger 等^[5]针对细胞分割提出的网络框架 U-Net, 因其形似 U 型而得名。U-Net 在 FCN^[6]的网络架构上修改并扩大了网络框架,使之成为一种编码器——解码器的端到端训练结构,能够使用较少(30 张)的训练图像得到较好的分割性能,由于其中的跳级连接能够保证特征更加精细,所以很多医学图像分割任务都使用了 U-Net 这样的结构。2019 年, Jin 等^[7]提出一种利用了视网膜血管的局部特征的基于 CNN 的视网膜血管分割可变形 UNet (Deformable U-Net, DUNet) 模型。DUNet 使用了上采样算子使得输出分辨率提高,并保留上下文信息,同时整合高层次和低层次的特征,方便具体定位。图像的各种预处理步骤来增强血管的细节、提高图像质量,并且通过调节接收域的大小来提取不同尺度和形状的视网膜血管,最终分割准确率达到 96%。尽管 U-Net 及其各类变种已经可以在几十张图像的训练集上训练出较为理想的模型,但在医学图像领域里,训练网络时需要面对带标签样本的一个潜在问题:样本不平衡^[8]。

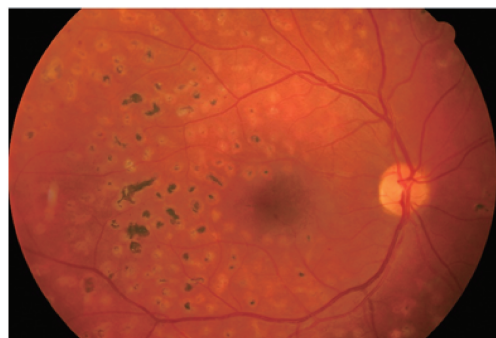
由于患病者的样本总比健康者的样本稀缺,同理重症样本比轻症样本稀缺,用不平衡的样本做分类训练可能会导致训练结果向比重大的那部分偏移^[9]。而对于血管分割任务,其本质是一个二分类问题,糖尿病视网膜病变重症样本和轻症、健康样本之间的血管差异主要在血管直径、是否存在淤血凝结、是否存在血管瘤等等,如图 1 所示,血管的差异度很大,不平衡的样本训练下泛化能力很差。

目前大多数基于深度学习的视网膜血管分割方法都是在 DRIVE 数据集 STARE 数据集和 CHASE_DB1 数据集上训练测试,DRIVE 数据集总计 40 张图像,仅 7 张为轻度早期糖尿病视网膜病变图像,其余 33 张为健康者图像,病变图像中 1 张为脉络膜病变,2 张为蝶形黄斑病,4 张为糖尿病视网膜病变;STARE 数据集总计 20 张图像,其中 10 张为病变患者图像,另 10 张为健康者图像;CHASE_DB1 数据集共 28 张图像,均为 14 岁健康儿童双眼视网膜图像。DRIVE 数据集和 CHASE_DB1 数据集中的视网膜图像多为健康者,健康者的视网膜图像血管边界分明,背景区分度高,无新生血管,分割难度低。在 DRIVE 测试集上轻度糖尿病的样本较少,测试时由样本不平衡而产生的影响较小,所以不平衡的问题没有体现出来。然而在实际临床诊断中参加检测的人群多是患者居多,由不平衡的训练集训练出来的网络对重度患者的进行血管分割的话,分割结果势必不够准确。



(a) 健康人眼底图片

(a) The fundus images of healthy people



(b) 重症糖尿病视网膜病变患者的眼底图片

(b) The fundus images of severe diabetic retinopathy patients

图 1 健康人和糖尿病视网膜病变患者的眼底图片

Fig. 1 The fundus images of healthy people and diabetic retinopathy patients

本文针对这一问题,提出将各个程度的病变图像等量纳入,并控制训练集的总量,使得网络利用少量样本进行训练,减少了对带标签样本的需求,同时也保证了样本平衡。

鉴于在传统深度网络中直接进行少样本训练会带来过拟合^[10],现阶段针对少样本学习主要有 3 种解决思路:1)通过扩充数据集,在少样本的情况下尽量增加参加训练的样本数,如 U-Net 随机切割大量图像块参与训练、使用在图像合成领域颇有成效的生成对抗网络生成成对的图像和标签^[11]等方法;2)通过损失函数增加正则项,在训练过程中增加了限制项来提高最终测试的准确度^[12];3)元学习,即通过预先设置的先验知识对少量带标签的样本任务进行学习推理,通过掌握目标核心替代海量样本战术。

由于视网膜血管分割这类样本的标签标注难度大,通过重复采样扩充样本尚可一试,但若用生成对抗网络生成新的视网膜图像,新生图像的血管标注工作无法完成,等同于无标签样本,对少样本学习没有帮助。所以本文结合了思路 1)和 2),在预处理阶段通过旋转扩充数据集,在网络结构部分借鉴生成对抗网络在图像风格迁移^[13]和条件生成对抗网络(conditional generative

adversarial nets, CGAN)^[14] 的思想,利用标签对图像分割进行约束,利用对抗系统对生成器训练的损失函数添加限制项。生成器输出分割后的图像,判别器分辨生成器的输出与人工标注图,尽量将网络生成图定为假,将人工标注的标准图定为真,并将判别结果反馈于生成器,使其提高分割性能。

1 基于 GAN 的少样本视网膜血管分割

1.1 少样本血管分割 GAN 介绍

原始的 GAN 主要用于生成与真实图像类似的假图像,其中生成器 G 利用初始噪声生成图像,判别器 D 则通过比对假图像和真实图像来确定究竟谁真谁假。GAN 生成器 G 的训练过程,本质就是以初始噪声的分布作为输入 z ,经过神经网络形成一个复杂的分布 $f(z)$,尽可能得让网络生成的分布 $f(z)$ 接近真实图像的分布。而判别器输入真假图像,尽力区分两者,判别器的好坏会影响生成器的性能^[15]。

本文提出的使用少量样本训练的 GAN 网络 (few samples generative adversarial networks, FS-GAN)。训练所用数据集仅有 6 张图像。在训练之前需要对图像预处理借此扩大数据集并提高分割性能^[16]。主要操作有图像反色、按规定角度旋转^[17],各个步骤处理后的结果如图 2 所示,预处理操作可以提高后续血管分割的性能,特别是旋转步骤既可以扩充数据集,同时也能给网络提供旋转不变性,这对于待测目标结构比较简单的医学图像分割非常重要,可以保证无论图像中的需要识别的对象的方向、角度如何变化都可以识别出来。

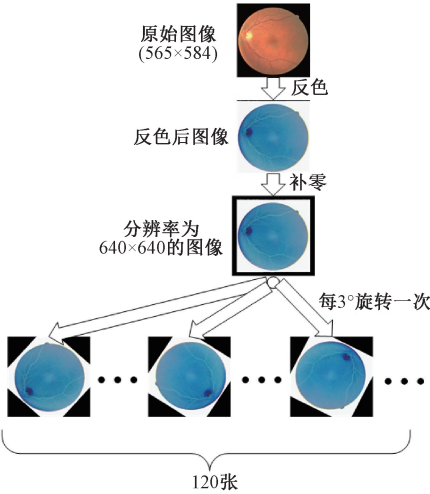


图 2 对原始图像进行的预处理操作步骤及结果
Fig. 2 The preprocessing operation steps for original image and result of each operation

FS-GAN 的整体网络结构如图 3 所示,网络依旧沿用原始 GAN 的单生成器、单判别器的结构。其中生成器网络需要进行图像分割,其输入为预处理后的眼底血管图,输出为眼底血管分割图。FS-GAN 的生成器使用了编码、解码形式的 U-Net 结构,如图 4 所示。编码段的每一维度的特征图都通过跳层连接融合到解码段的对应特征图上,通道叠加丰富了特征图的尺度信息^[18],低层的高分辨率特征图提供精细的细节,高层特征图由于多次卷积和池化 (pooling)、反卷积 (deconvolution) 得到较为模糊的轮廓信息,两者融合以保证图像分割时不仅考虑了细节特征同时也纳入上下文信息,提高了网络细分血管的精度。

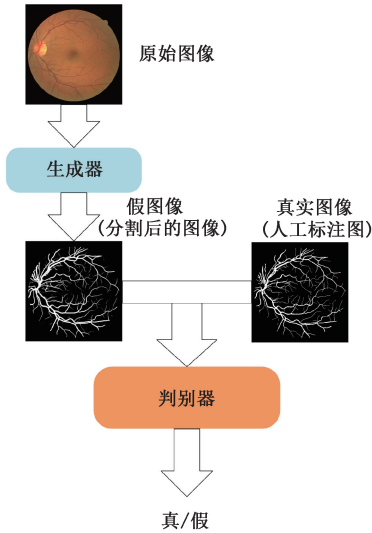


图 3 FS-GAN 网络结构
Fig. 3 Network structure of FS-GAN

判别器则使用了 VGG16^[19] 的变种,网络结构如图 5 所示,有 11 层卷积层、4 次最大池化、1 次全局池化和 1 层全连接层 (dense),网络任务仍是尽力区分真实图像 (人工分割的血管图) 和假图像 (生成器提供的分割血管图)。判别器的网络在 VGG16 结构上进行了修改,移除了最后的三层全连接网络,换成了全局平均池化,将每一个通道的全部像素进行均值计算,得到 1 024 个数据构成 $1 \times 1 \times 1\,024$ 的特征图,用全局池化代替全连接层一方面整合了全局信息,提高了网络的鲁棒性,另一方面也去除了全连接层的大量参数。全局池化后接全连接层,通过非线性映射将所有特征图整合为一个数值,以此判断输入图像是真实图像还是假图像。

1.2 代价函数

原始 GAN 的目标函数即为寻求生成器与判别器的纳什均衡点,其判别器需要最小化的损失函数如下:

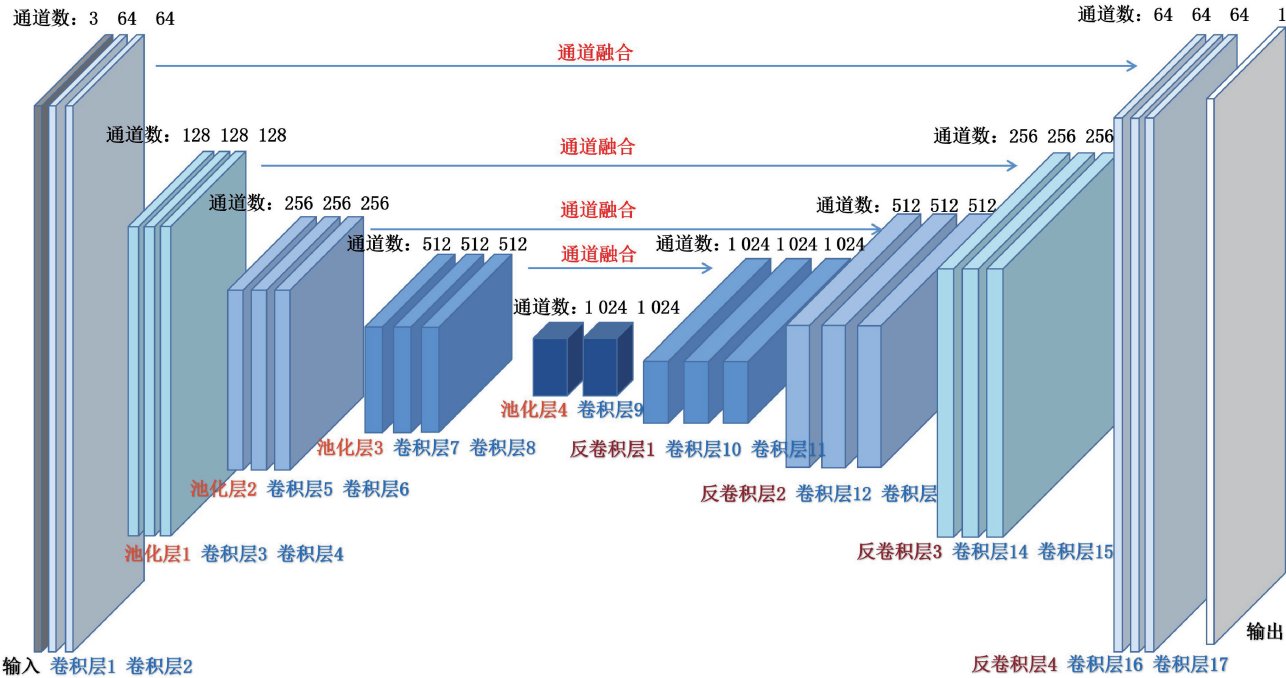


图 4 生成器网络模型
Fig. 4 Network model of generator

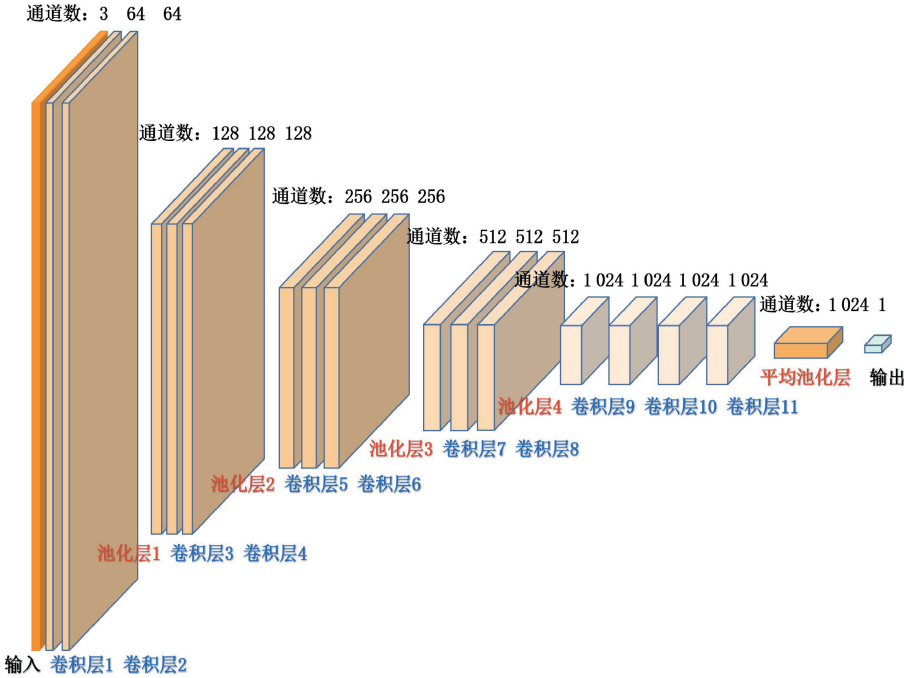


图 5 判别器网络模型
Fig. 5 Network model of discriminator

$$V_D(G,D) = \{-E_{x \sim p_d(x)}[\log(D(x))]\} - E_{z \sim p_z(z)}[\log(1 - D(G(z)))] \quad (1)$$

而生成器的损失函数如下:

$$V_G(D) = E_{z \sim p_z(z)}[\log(-D(G(z)))] \quad (2)$$

式中: x 为真实数据; z 为生成器所需的随机噪声信号

(如高斯分布), $G(z)$ 为生成器生成的假数据。

而对于解决分割任务而提出的 FS-GAN 而言, 判别器的输入数据有输入的图像 x 与真实图像 y 、输入的图像与假图像 $G(x)$ 两对, 实际上这部分生成对抗系统的损失函数使用的依旧是原始 GAN 使用的交叉熵损失函

数^[20],所以生成对抗系统的目标函数比较类似,判别器端的损失函数如下:

$$L_D(G, D) = -E_{x, y \sim p_d(x, y)} [\log(D(x, y))] - E_{x \sim p_d(x)} [\log(1 - D(x, G(x)))] \quad (3)$$

由于生成器和判别器需要异步训练,所以使用的是交替优化方式,在 FS-GAN 的训练过程中,第一步是固定判别器 D 的参数,对生成器 G 进行优化得到最优的参数,尽量骗过判别器;第二步是固定优化后的生成器 G 的参数,对判别器 D 进行优化得到最优的参数,让判别器的判别能力尽可能好。这一部分是在原始 GAN 的基础上修改而成,但由于 FS-GAN 的生成器本质上是做分割任务,所以在式(3)之外需要额外添加生成器中血管分割的损失函数。

既然 FS-GAN 的生成器主要任务是血管分割,作为一类二分类任务,可以使用 0~1 交叉熵作为损失函数,因此 FS-GAN 的生成器端的损失函数如下:

$$L_G(G) = E_{x, y \sim p_d(x, y)} [\log(D(x, y))] + E_{x \sim p_d(x)} [\log(1 - D(x, G(x)))] + \theta E_{x, y \sim p_d(x, y)} [-y \cdot \log G(x) - (1 - y) \log(1 - G(x))] \quad (4)$$

在 FS-GAN 训练中,首先将生成器分割任务的损失函数(4)最小化作为生成器的优化目标,然后固定 G 的参数,训练判别器。这里将判别器的损失函数(3)与生成器的损失函数(4)的加权和视为 FS-GAN 整体需要优化的对象,加权系数 θ 在网络训练时属于常量,可通过改变数值多次训练取最佳数值。FS-GAN 网络的目标函数为:

$$\min_c \max_b L(G, D) = \min_c \max_b L_D(G, D) + \min_c L_G(G) \quad (5)$$

2 实验结果与分析

本实验使用一台 CPU 为 6 路 E5-2678 v3、内存 64 GB、显卡为 2080Ti 的计算机,所有的算法均在 Keras 框架下实现,判别器和生成对抗系统的优化器都采用初始学习率为 0.000 2、 β_1 系数为 0.5 的 Adam,批处理大小为 1,预设的训练轮次为 10。

2.1 数据集介绍

本文在数字视网膜图像血管提取数据集(digital retinal images for vessel extraction, DRIVE)与高分辨率眼底图像数据集(High-Resolution Fundus, HRF)上分别进行训练、测试。DRIVE 数据集^[21]是使用最广泛的视网膜血管分割数据集,其中共有 40 张分辨率为 565×584 的图像,其中 7 张是轻度眼底病变患者的眼底图像,其余均为健康人眼底图像。训练集是在原始 DRIVE 数据集中选取了 6 张图像,3 张为健康人眼底图像,3 张为患者眼底

图像,测试集沿用原始 DRIVE 的测试集,共 20 张图像。

HRF 数据集^[22]是针对精细眼底血管分割而制作的数据集,其中含有 15 例健康者的图像,15 例糖尿病性视网膜病变患者的图像和 15 例青光眼患者的图像,图像分辨率为 3 504×2 336。由于 HRF 数据集中没有划分训练、测试数据,所以本方法在 HRF 数据集上进行处理,建立了一个混合 HRF 数据集。训练集有 6 张图像,其中健康者眼底图像,轻度患者眼底图像(糖尿病性视网膜病变患者)和重度患者眼底图像(青光眼患者)各有 2 张。测试集则使用 HRF 数据集中剩下的 39 张图像。受网络模型大小限制,输入图像尺寸需要控制在 640×640 以内,HRF 数据集的图像在训练测试前已经做了压缩处理,处理后尺寸为 640×426。

2.2 DRIVE 数据集上各网络实验结果

本次实验网络在 DRIVE 数据集上预设训练迭代次数为 10,加权系数 θ 的多次实验测试指标如图 6 所示,通过对比取值的结果后,发现其他条件不变的情况下,在 θ 取值在 25 附近时网络性能最佳,多次实验的测试指标如图 6 所示,网络推理后取一幅代表图,其网络不同轮次下推理的输出结果如图 7 所示。由于 DRIVE 数据集在前几次迭代后,血管的主干部分便已经分割出来,但是细小血管和背景难以区分,分割图存在大量噪点。第 7 次迭代之后,细小血管已经较为清晰地分割出来,但仍旧存在背景区域的部分错误分割。第 10 次迭代后,背景部分基本剔除干净,血管分割效果与人工标注从视觉效果来看区别不大,但也存在少量被忽略的微小血管。不同迭代次数下测试指标如图 8 所示,结合图 7,第 1~8 轮迭代网络性能一直上升,而第 10 轮迭代后网络性能开始波动,网络整体损失并没有一直下降而是频繁波动,同时也考虑到时间成本,所以本实验将网络迭代次数选定为 10 次。

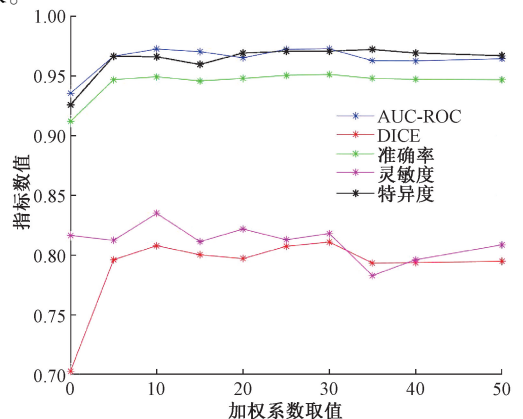


图 6 DRIVE 数据集上不同加权系数 θ 的 FS-GAN 网络输出性能指标折线

Fig. 6 Broken line diagram of FS-GAN output performance in different weighted coefficient θ on DRIVE dataset

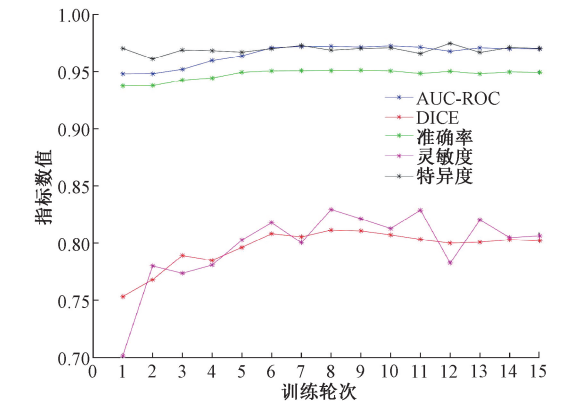


图 7 DRIVE 数据集下不同训练轮次的 FS-GAN 网络输出性能指标折线图

Fig. 7 Broken line diagram of FS-GAN output performance in different training rounds on DRIVE dataset

表 1 比较了现有视网膜血管分割方法中性能较好的 DUNet 和 IterNet^[23]与本文提出的 FS-GAN 和原始的 U-Net 在 DRIVE 数据集上的性能,评价指标包括曲线下面积 (area under curve- receiver operating characteristic curve, AUC-ROC)、F1 值、准确率、训练轮次和训练加推理时间。可以看出 FS-GAN 在少样本的训练集限制下,对视网膜血管的分割性能却能非常接近其他使用完整数据集的血管分割方法。同时通过和使用 6 样本和 20 样本训练的 U-Net 性能对比可以看出,在使用少样本进行训练后,U-Net 的网络性能小幅下降,AUC-ROC 降低了 4%,F1 值降低了 15%,但是本文提出的 FS-GAN 通过对 U-Net 添加对抗系统,在少样本情况下利用判别器监督网络输出,提高血管分割性能。同时将 FS-GAN 和 DUNet、IterNet 在少样本的情况下进行比较,可以发现本方法在 AUC-ROC、F1 值和全局准确率上均位列第一,虽然和其他算

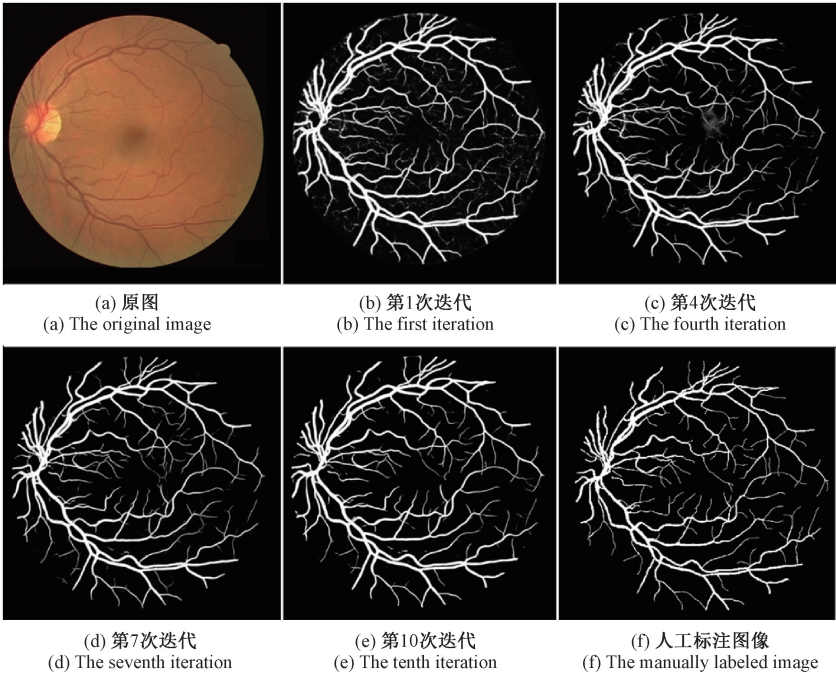


图 8 FS-GAN 网络训练 DRIVE 数据集的输出结果与人工标注图对比

Fig. 8 Comparison of segmentation result from FS-GAN and manual labeled image on DRIVE dataset

表 1 本文提出的方法与其他方法在 DRIVE 数据集上分割结果的比较

Table 1 Comparison of proposed method segmentation result with other methods on DRIVE dataset

方法	AUC-ROC	F1 值	准确率	训练轮次	训练+推理时间/min
FS-GAN(6 样本)	0.976 2	0.807 4	0.953 0	10	60+5
U-Net(6 样本)	0.934 8	0.665 9	0.948 3	10	20+5
IterNet(6 样本)	0.975 1	0.790 0	0.952 3	200	300+30
DUNet(6 样本)	0.972 2	0.799 9	0.952 0	100	200+20
U-Net(20 样本)	0.975 5	0.814 2	0.932 9	10	60+5
IterNet(20 样本)	0.981 6	0.820 5	0.957 3	200	480+30
DUNet(20 样本)	0.980 2	0.823 7	0.969 7	100	400+20

法差距不大,但考虑到网络结构中生成器部分使用的是性能并不卓越、但训练效率较高的 U-Net,在使用本方法后确实能大幅提高少样本训练下的血管分割精度。

从训练时间上来看,本文提出的 FS-GAN 由于在 U-Net 的基础上添加了生成对抗结构,所以训练时间比 U-Net 耗时更久,需要 1 h 左右才能完成 10 轮次训练,达到较好的分割效果。而 IterNet 和 DUNet 由于网络复杂度高并且为达到较好效果所需的训练轮次较多,所以两者的训练时间较长。并且 IterNet 和 DUNet 的推理阶段需要将待测试图片切割成小块再进行血管分割,导致两者的推理时间较长。然而,FS-GAN 中并未使用原始 U-Net 中的切割图像块进行训练的步骤,而是将图像旋转多次后输入网络,生成器网络的输入尺寸为原始图像尺寸,这使得 FS-GAN 的推理过程非常简单,只需将待分割的图像输入生成器网络即可得到分割后的图像,所以推理花费的时间较少。

2.3 HRF 数据集上各网络实验结果

本次实验网络在混合 HRF 数据集上训练迭代次数通过比较后依旧确定为 10,加权系数 θ 的多次实验测试指标如图 9 所示,通过对比取值的结果后,发现其他条件不变的情况下,在 θ 取值在 35 附近时网络性能最佳。迭代后取一幅代表图,其通过网络推理输出的结果如图 10

所示,由于 HRF 数据集比起 DRIVE 数据集,人工标注图中拥有更多的细小血管,细节标注的精度极高,对血管分割任务来说是一个有挑战性的数据集。在 10 次迭代训练之后,网络分割出了主干血管和其分支血管和接近全部的细小血管,FS-GAN 在 HRF 数据集上的分割性能指标如表 2 所示,FS-GAN 面对这些专家标注都存在挑战性的细小血管达到了 94% 的分割准确度,ROC 的曲线下面积也达到了 0.95。

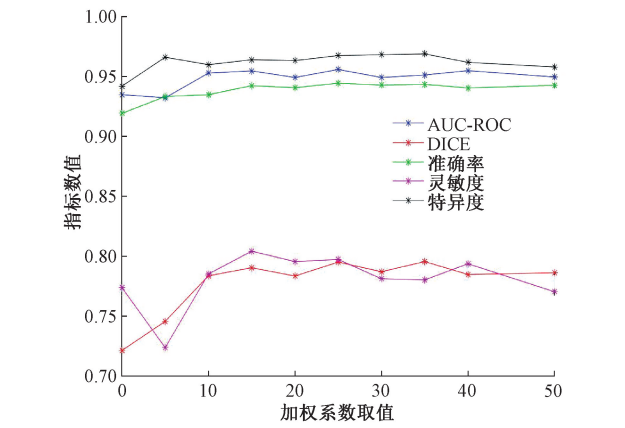


图 9 HRF 数据集下不同加权系数 θ 的 FS-GAN 网络输出性能指标折线图

Fig. 9 Broken line diagram of FS-GAN output performance in different weighted coefficient θ on HRF dataset

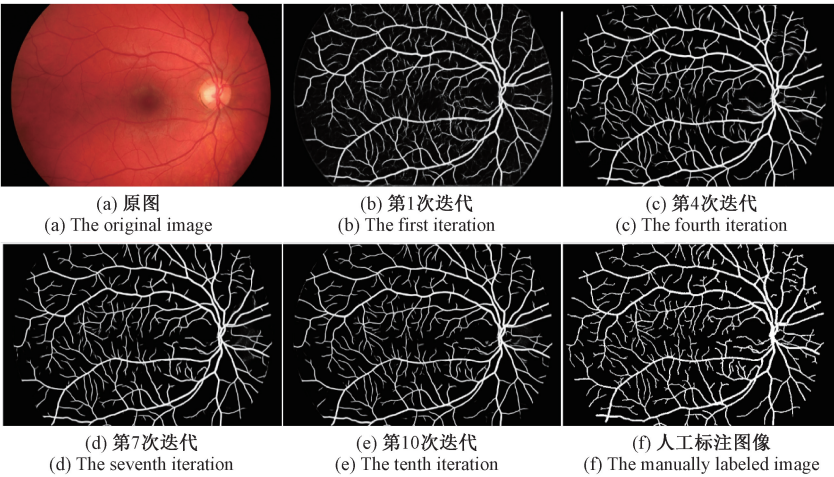


图 10 FS-GAN 网络训练 HRF 数据集的输出结果与人工标注图对比

Fig. 10 Comparison of segmentation result from FS-GAN network and manual labeled image on HRF dataset

表 2 本文提出的方法与其他方法在 HRF 数据集上分割结果比较

方法	AUC-ROC	特异度	灵敏度	F1 值	准确率	训练+推理时间/min
FS-GAN(6 样本)	0.951 2	0.968 8	0.780 4	0.802 1	0.943 4	40+5
IterNet(6 样本)	0.950 3	0.959 0	0.785 2	0.788 9	0.930 1	300+60
DUNet(6 样本)	0.950 2	0.969 7	0.695 5	0.796 1	0.940 8	200+30

从表 2 的分割指标来看,FS-GAN 在 HRF 数据集上的大部分性能指标均处于 DUNet、IterNet 平均水平之上,有着最高的准确率、F1 值和训练时间,特异度与灵敏度均处于较高水平。这表明 FS-GAN 在面对复杂血管时,错分概率不大,大部分血管能够检测出来,漏掉了部分细小血管,且极少出现将背景划分为成血管的情况。而 IterNet 灵敏度最高,特异度最低,其血管分割时正确划分血管的概率高,但不可避免地会将部分背景划为血管,导致分割出的血管超出实际,会多出很多虚假血管。

DUNet 特异度最高,灵敏度最低,其对血管的分割较为谨慎,很少将背景划为血管,但会导致部分血管被划入背景之列。3 种网络的分割比较如图 11 所示。明显可以看出,IterNet 分割的血管图中血管边缘模糊不清,DUNet 分割的血管图中有接近半数的细小血管未能分割出来。从指标上来说,FS-GAN 在 3 个网络中处于较高水平,且从观感来看,FS-GAN 的分割图的视觉效果也比其他两种方法要好,鉴于临床时眼科医生需要根据血管图对病人进行诊断,视觉效果好坏也是一个重要评价标准。

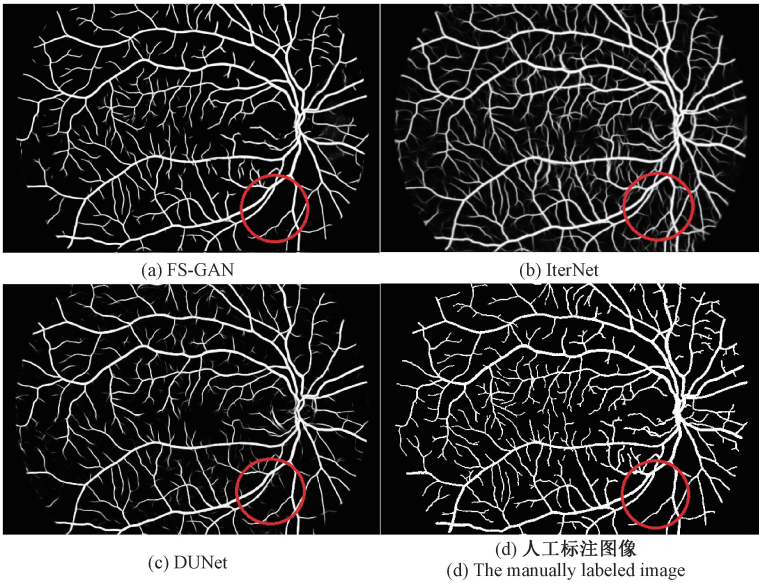


图 11 FS-GAN、IterNet、DUNet 和人工标注的分割图对比

Fig. 11 Comparison of the segmentation images from FS-Gan, IterNet, DUNet and manual labeled

表 3 为 FS-GAN 对 HRF 数据集中不同人群的分割结果统计,可以看出 FS-GAN 对三者的血管分割能力略有差别,对健康人的血管分割效果极其优秀,在 HRF 数据集中平均准确率为 0.943 4,对健康人的准确率直接增加了 1 个百分点达到了 0.953 2,而同样的,对青光眼病变患者和糖尿病性视网膜病变患者的血管分割效果各有一定程度的衰减,尤其是对糖尿病性视网膜病变患者的分割指标大幅下降。原因是糖尿病性视网膜病变患者的视网膜图像中存在大量出血块,网络需要学习如何去去除掉这部分。尽管 FS-GAN 通过缩减训练集做到了样本集平衡,但实现这一目标牺牲了大样本带来的先验知识,导致网络对区分复杂血管和复杂背景的学习不够充分。

2.4 测试样本不平衡对网络性能的影响

本次实验针对样本不平衡问题进行对比实验,重组 HRF 数据集中的不同来源的数据,构建了两种数据集,数据集 1 是在 2.1 节建立的混合数据集上进行修改,保持训练集不变,测试集去掉了 3 张健康者的图像;数据集 2 是在数据集 1 的基础上保持测试集不变,训练集则删去糖尿病性视网膜病变患者 1 张图像和青光眼患者的 2 张图像,加入 3 张健康者的图像,这样训练集就有 5 张健康者的图像和 1 张糖尿病性视网膜病变患者图像。两次训练后测试集的推理结果如表 3 所示。数据集 1 和 2 的推理分割结果代表图如图 12 所示。

表 3 本文提出的方法在 HRF 数据集上对不同人群的分割结果统计

Table 3 Statistics of proposed method segmentation result of different groups on HRF dataset

人群分类	DICE	特异度	灵敏度	F1 值	准确率
健康人	0.833 5	0.974 0	0.841 6	0.848 4	0.953 2
青光眼病变患者	0.789 5	0.969 4	0.752 8	0.795 1	0.942 7
糖尿病性视网膜病变患者	0.744 9	0.963 1	0.746 2	0.762 4	0.934 3

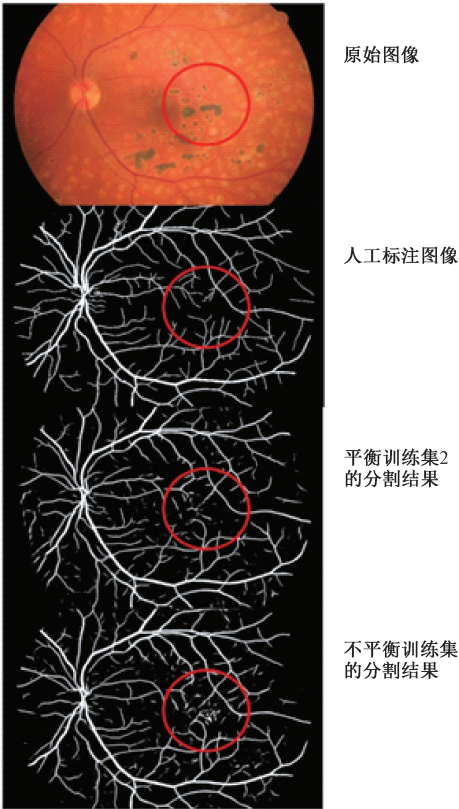


图 12 本文提出的方法在数据集 1 和 2 上分割结果比较
Fig. 12 Comparison of proposed method segmentation result on dataset a and dataset b

表 4 本文提出的方法在数据集 1 和 2 上分割结果比较
Table 4 Comparison of proposed method segmentation result on dataset 1 and dataset 2

方法	数据集 1(不平衡)	数据集 2	变化度(数据集 2-1)
AUC-ROC	0.945 1	0.951 1	+0.006 0
特异度	0.965 2	0.969 3	+0.004 1
灵敏度	0.765 2	0.770 4	+0.005 2
DICE	0.770 5	0.783 4	+0.012 9
准确率	0.938 0	0.942 4	+0.004 4

从表 4 可以看出在数据集 2 上得到的推理结果中,AUC-ROC、F1 值和全局准确率均比数据集 1 的结果低 0.5%~1%,从数据来看似乎差别不大,但在图 9 中圈出的区域中,不平衡的训练集迭代多次后仍然无法对淤血区域分割出清晰的边界,说明本文提出的 FS-GAN 方法在使用不平衡的训练集训练的最终分割效果比使用平衡标签的训练集得到的结果更差,所以通过构建一个少样本训练架构来保证训练集平衡是一项必要的任务。

2.5 针对生成对抗结构的消融实验

本次消融实验通过去除 FS-GAN 中的生成对抗结构,观察去除前后的网络分割性能变化定性分析生成

对抗结构对本网络的影响。数据集沿用 2.1 节建立的混合 HRF 数据集。两次训练后测试集的推理结果如表 5 所示。消融后数据集的部分推理分割结果如图 13 所示。

表 5 本文提出的 FS-GAN 和消融了生成对抗结构的 FS-GAN 在 HRF 数据集上分割结果比较
Table 5 Comparison of proposed method segmentation result with ablating GAN structure network on HRF dataset

方法	FS-GAN	消融后	变化度
AUC-ROC	0.951 2	0.948 1	-0.003 1
特异度	0.969 3	0.970 8	+0.001 5
灵敏度	0.780 4	0.742 0	-0.038 4
DICE	0.785 1	0.769 6	-0.015 5
准确率	0.943 4	0.939 9	-0.003 5

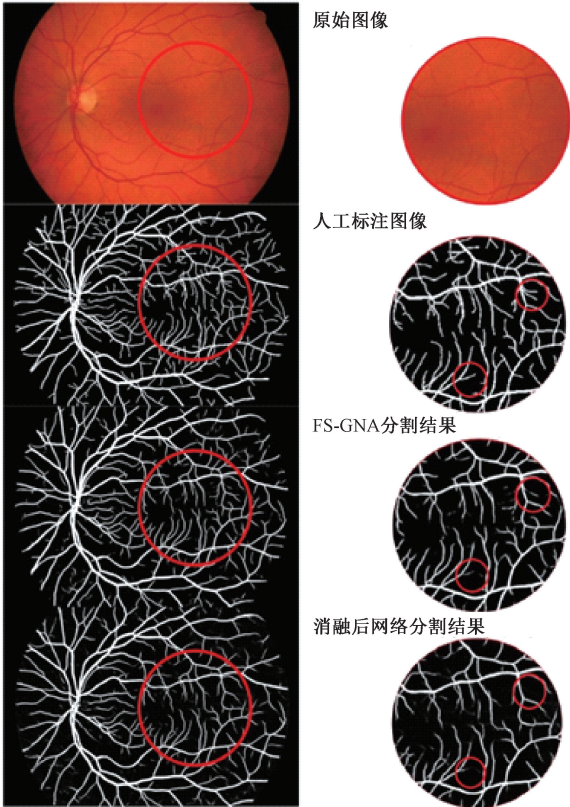


图 13 本文提出的方法与消融后的网络在 HRF 数据集上分割结果比较
Fig. 13 Comparison of proposed method segmentation result with ablating GAN structure network on HRF dataset

表 5 的性能指标在消融对抗结构后,网络仅靠旋转扩充训练集和 U-Net 网络对血管进行分割,性能指标中除了特异度之外均有所下降,其中灵敏度下降幅度很大,灵敏度越低,则意味着对细小血管的分割精度低。单从

指标数据来看,消融前后差异很小,但通过图 13 中圈出的区域对比,可以看出缺少了对抗机制对生成器(U-Net 网络)的约束,消融后网络对主干血管分割的效果基本不变,但对于复杂的细小血管、血管密集部分,分割精度有所降低。说明本文提出的对 U-Net 加装生成对抗系统的架构能够提高分割网络对细小血管以及复杂区域的分割精度。

3 结 论

本文针对医学图像处理领域样本稀缺和样本不平衡等问题,在 U-Net 网络的基础上构建了对抗学习模型,借此提出一种基于生成对抗网络的少样本视网膜血管分割方法,并进行研究实验。通过将实验结果与现有算法比较,本文的方法在使用少样本训练集的情况下在 DRIVE 数据集上 ROC 的曲线下面积达到了 0.97,准确率达到 0.95。各项指标均接近其他使用完整训练集的血管分割算法比如 IterNet 和 DU-Net,并且本文提出的方法在少样本下与其他方法的比较中均处于较好水平,且明显优于少样本训练下的 U-Net 网络,证明了本文提出的方法在少样本视网膜血管分割上的可行性。但在 HRF 数据集上进行的实验表明本方法对细小血管的分割存在一些小问题,对细小血管的分割准确度还不够高,因此在 HRF 数据集上的分割精度比 DRIVE 数据集要低。后续研究方向是在 U-Net 的跳层连接中增添自学习的注意力机制^[24],将细小血管和主干血管的特征有选择性地传递到高层网络中,结合生成对抗机制,进一步提高细小血管的分割精度。

参考文献

- [1] FAN Z, LU J, WEI C, et al. A hierarchical image matting model for blood vessel segmentation in fundus images[J]. IEEE Transactions on Image Processing, 2019, 28(5):2367-2377.
- [2] LITJENS, GEERT, KOOI, et al. A survey on deep learning in medical image analysis[J]. Medical Image Analysis, 2017, 42(9):60-88.
- [3] 徐莹莹, 沈红斌. 基于模式识别的生物学医学图像处理研究现状[J]. 电子与信息学报, 2020, 42(1): 201-213.
XU Y Y, SHEN H B. Review of research on biomedical image processing based on pattern recognition[J]. Journal of Electronics and Information Technology, 2020, 42(1):201-213.
- [4] 罗佳, 黄晋英. 生成式对抗网络研究综述[J]. 仪器仪表学报, 2019, 40(3):74-84.
LUO J, HUANG J Y. Generative adversarial network;

- An overview [J]. Chinese Journal of Scientific Instrument, 2019, 40(3):74-84.
- [5] 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, MICCAI, 2015: 234-241.
- [6] 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.
- [7] JIN Q, MENG Z, PHAM T D, et al. DU-Net: A deformable network for retinal vessel segmentation[J]. Knowledge-based Systems, 2019, 178(15):149-162.
- [8] ZHOU Z H, LIU X Y. Training cost-sensitive neural networks with methods addressing the class imbalance problem[J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 18(1):63-77.
- [9] TAN A C, GILBERT D, DEVILLE Y. Multi-class protein fold classification using a new ensemble machine learning approach[J]. Genome Informatics, 2003, 14: 206-217.
- [10] HINTON G E, SRIVASTAVA N, KRIZHEVSKY A, et al. Improving neural networks by preventing co-adaptation of feature detectors [EB/OL]. Computer Research Repository, 2012, <https://arxiv.org/abs/1207.0580>.
- [11] ZHU J, PARK T, ISOLA P, et al. Unpaired image-to-image translation using cycle-consistent adversarial networks[C]. 2017 IEEE International Conference on Computer Vision (ICCV), 2017:2242-2251.
- [12] WANG Y, YAO Q, KWOK J T, et al. Generalizing from a few examples: A survey on few-shot learning[J]. ACM Computing Surveys, 2020, 53(3):63:1-63:34.
- [13] ZHANG Z, LI M, YU J. D2PGGAN: Two discriminators used in progressive growing of GANS[C]. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2019:3177-3181.
- [14] MIRZA M, OSINDERO S. Conditional generative adversarial nets [J]. Computer Science, 2014: 2672-2680.
- [15] GOODFELLOW J I, POUGET-ABADIE J, et al. Generative adversarial nets [J]. Advances in Neural Information Processing Systems, 2014:2672-2680.
- [16] 于宁波, 刘嘉男, 高丽, 等. 基于深度学习的膝关节 MR 图像自动分割方法[J]. 仪器仪表学报, 2020, 41(6):140-149.
YU N B, LIU J N, GAO L, et al. Auto-segmentation

- method based on deep learning for the knee joint in MR images [J]. Chinese Journal of Scientific Instrument, 2020, 41(6):140-149.
- [17] GUIZARD N, COUPÉ P, FONOV V S, et al. Rotation-invariant multi-contrast non-local means for MS lesion segmentation [J]. NeuroImage: Clinical, 2015, 8: 376-389.
- [18] ZHOU Z, SIDDIQUEE MMR, TAJBAKHS N, et al. UNet++: A nested u-net architecture for medical image segmentation [C]. Medical Image Computing and Computer-Assisted Intervention Workshop, 2018:3-11.
- [19] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [C]. International Conference on Learning Representations, 2015:1-14.
- [20] RUBINSTEIN R. The cross-entropy method for combinatorial and continuous optimization [J]. Methodology & Computing in Applied Probability, 1999 (1):127-190.
- [21] STAAL J, ABRAMOFF M D, NIEMEIJER M, et al. Ridge-based vessel segmentation in color images of the retina [J]. IEEE Transactions on Medical Imaging, 2004, 23(4):501-509.
- [22] BUDAI A, BOCK R, MAIER A, et al. Robust vessel segmentation in fundus images [J]. International Journal of Biomedical Imaging, 2013, 2013(6):154860.
- [23] LI L, VERMA M, NAKASHIMA Y, et al. IterNet: Retinal image segmentation utilizing structural redundancy in vessel networks [C]. 2020 IEEE Winter Conference on Applications of Computer Vision, 2020: 3645-3654.
- [24] LEE K, OGNIBENE D, CHANG H J, et al. STARE: Spatio-temporal attention relocation for multiple structured activities detection [J]. IEEE Transactions on Image Processing, 2015, 24(12):5916-5927.

作者简介



张思杰(通信作者), 分别在 1990 于重庆大学获得学士学位, 2002 年和 2005 年于重庆大学获得硕士学位和博士学位, 现为重庆大学副教授, 主要研究方向为生物医学信号处理、生物微机械电子系统等。

E-mail: zhangsj@cqu.edu.cn

Zhang Sijie (Corresponding author) received B. Sc. degree from Chongqing University in 1990, M. Sc. and Ph. D. degree both from Chongqing University in 2002 and 2005, respectively. Now he is an associate professor at Chongqing University. His main research interests include biomedical signal processing, BioMEMS, etc.



方翔, 2018 年于重庆大学获得学士学位, 现为重庆大学硕士研究生, 主要研究方向为医学图像分割。

E-mail: fx@cqu.edu.cn

Fang Xiang received his B. Sc. degree from Chongqing University in 2017. He is currently a M. Sc. candidate at Chongqing University. His main research interest includes medical image segmentation.