

文章编号: 2096-1618(2018)05-0525-06

基于 Dense Connected 深度卷积神经网络的 自动视网膜血管分割方法

唐明轩, 李孝杰, 周激流

(成都信息工程大学计算机学院, 四川 成都 610225)

摘要:深度卷积神经网络(DCNN)在自然图像分类和分割问题中具有优越的性能。眼底视网膜血管作为可无创直接观察到的血管,对其结构的分析是眼科病变诊断的重要依据之一。如毛细血管增生等变化为糖尿病等眼科疾病的诊断提供了重要的指导意义。因此,如何正确高效地分割视网膜血管成为一种临床需求。在不使用任何前后期处理的条件下,提出一种基于 Densely Connect 深度卷积神经网络的自动视网膜血管分割方法。方法通过使用稠密连接(Densely connect),批规范化(Batch Normalization)等技术构建一种新型的深度卷积神经网络,并结合带孔卷积(Dilated convolution)增加网络分割精度,在更少人为处理的情况下提高视网膜血管的分割性能。在对比实验中,提出网络的平均精确度,敏感度,特异性达到0.9617,0.7325,0.9839,像素级的 AUC 指标达到0.978,优于对比的机器学习方法和对比深度卷积神经网络。验证了所提方法在视网膜血管分割中的有效性。

关键词:深度学习;卷积神经网络;图像分割;视网膜血管

中图分类号:TP181

文献标志码:A

doi:10.16836/j.cnki.jcuit.2018.05.007

0 引言

眼底视网膜血管作为可无创直接观察到的血管,其结构信息是眼科病变诊断的重要依据之一。视网膜血管分割在很多疾病的诊断中都有重要作用。除了视网膜本身的病变会对视网膜眼底血管产生影响,一些其他的疾病(如糖尿病等)也会对视网膜血管产生影响。糖尿病患者因血液成分的变化会产生视网膜血管增生,出血,产生微动脉瘤等症状。如果不能及早发现并治疗,很可能导致视力受损甚至失明。传统的分割算法虽然可以一定程度上完成分割,但通常存在分割效果欠佳的情况。同时,传统的分割方法(如血管追踪,匹配滤波等)往往需要大量的前期处理(如手工特征选择,参数选择等)在增加人工成本的情况下减弱了分割方法的泛化能力。深度学习作为机器学习的分支,通过深度卷积神经网络以及前向反向传播算法,可以自动的提取血管特征,完成血管的分割。该类自动分割方法对眼科疾病的诊断具有重要意义。

1 神经网络

1.1 卷积神经网络

人工神经网络是一种模拟人类神经系统进行信息

处理的网络,在很多场景显示了其高效性。深层神经网络是一种多层结构的网络。其首先通过建立多层的网络结构,将输入特征映射至输出空间。网络结构中的每一层可以认为是一个特征映射函数,如:

$$y=f(x;o) \quad (1)$$

其中 x 表示输入, y 表示输出, o 表示需要训练的参数。通过逐层相连接,形成一种复合的链式结构:

$$F(x)=f_n(Af_3(f_2(f_1(x)))) \quad (2)$$

将这个结构从内到外的每个映射依次称为第1层,第2层,……,第 n 层。 n 为网络的深度,其第1层又称为输入层,最后1层称为输出层。因中间层的特征并没有明确的定义,而是由学习算法来自动确定每一层的输出结果,所以被称为隐藏层。其中每个映射产生的结果称为特征图(featuremap),特征图上每个点是对应其感受野的特征映射。

在图像处理中,早期使用神经网络进行图像处理的网络,如 LeNet 被应用在手写体数字识别,达到了可被当时邮政系统接受的识别效率。方法的核心思想是通过卷积运算自动提取图像特征,已是目前深度神经网络在图像处理的基本方法。卷积神经网络具有稀疏交互,权值共享等变表示等特点^[1],可对图像的特征数量进行极大的压缩,同时增强网络对于平移变换的适应性。而在 imageNet2012 上大放异彩的 AlexNet^[2]。其主要的贡献在于增加 dropout 以增强网络的泛化性,使用 ReLU 激活函数替换 Sigmoid 激活函数加速网络

训练,以及应用 GPU 加速网络训练。由于卷积神经网络的优越性能,促使大量研究者进入深度学习的研究中,同时催生了 Caffe, Tensorflow^[3], mxnet 等一系列优秀的深度学习框架。

1.2 全连接神经网络

最初的深度卷积神经网络仅被使用于图片分类中,由前部分的卷积层提取特征,并通过后续的全连接层进行分类。设计能有效分离特征提取和分类方法。可以在不修改特征提取部分的情况下,仅修改全连接的输出就能适配网络至其他数量,种类的应用中。但它仅能处理图片的分类问题。最简单的将其应用至分割问题的办法是对图像每个像素点提取 patch,逐个对 patch 的中心点进行预测,通过拼接所有点分类结果来将分割问题转换成分类问题。这类处理方式需要提取大量 patch 用于制作训练集,而在预测结果时,需要针对图像上每个点生成对应的 patch 进行分类,并拼接所有的分类结果生成分割结果。这些繁琐的操作使得网络在训练和实际应用中效率极其低下。

2015 年,Jonathan Long 提出使用全卷积网络 (fully

convolution networks, FCN)^[3] 进行图像分割,其核心思想是在网络的后半段并不使用全连接层输出类别,而是使用反卷积层将图像放大到原尺寸的方法,实现端到端 (end-to-end) 的分割目标。为实现更具细节的分割效果,前层的特征图被保存并融合到对应的还原层中。这使得图像分割网络在预测的精度上获得了提升,同时大幅降低了图像分割在训练和测试时的运行时间。在此之后延伸出了大量的类似网络,如 SegNet^[4], U-Net^[5] 等。

U-Net 是 Olaf Ronneberger 等提出的深度卷积神经网络,在 ISBI cell tracking challenge 2015 中取得了最好的成绩,并且广泛应用于医学图像处理中。U-Net 拥有类似 FCN 的网络结构,输入图像首先经过两个卷积层和 ReLU 激活层,然后进行 max-pooling 下采样,之后是重复的卷积层,激活层,然后再次下采样。每次下采样后的卷积层输出的通道数都是上一层的两倍。在经过 3 次下采样后,网络开始还原特征。这时 U-Net 使用上采样还原特征图尺寸,并结合前层的低级特征还原细节特征。在经过 3 次上采样后,特征图大小还原为原始大小,并使用卷积输出分割结果。一个使用两次下采样的 U-Net 网络结构如图 1 所示。

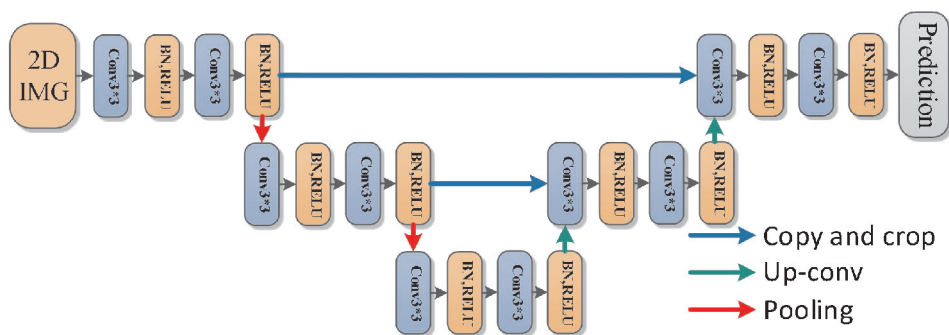


图1 两次下采样的 U-Net 网络结构图

1.3 更深层的神经网络

在传统的神经网络中,随着网络的加深,网络很容易出现梯度消失或者梯度爆炸的问题。但更深的神经网络结构能够处理更非线性的抽象问题,并且通常可以获得更优的性能。为增加网络的深度以获得更好的抽象能力,致力于克服梯度问题的网络结构分别被提了出来。在 AlexNet 中作为替代 sigmoid 成为激活函数的线性整流函数 (rectified linear unit, ReLU) 目前已经是主流网络的默认激活函数。已有研究证明,使用 ReLU 类激活函数可以使得网络深度加深时更容易收敛。ReLU 激活函数公式如下:

$$G(z) = \max(0, z) \quad (3)$$

Batch Normalization^[6] 作为一种可学习的归一化层,可以在不引入太多计算的情况下解决梯度消失和

梯度爆炸等问题。该类处理的核心思想是通过逐层归一化,减小反向传播时较远层因梯度传递造成的影响。应当注意在训练时使用每个批次的数据进行归一化,但在测试时使用均值进行归一化。具体归一化过程如下式所示:

$$\mu_B = \frac{1}{m} \sum_{i=1}^m x_i \quad (4)$$

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^m (x_i - \mu_B)^2 \quad (5)$$

公式(4)和公式(5)分别求取特征图的均值和方差,其中 x_i 表示本层第 i 个 batch, m 表示 batch 的个数。

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \quad (6)$$

$$y_i = \gamma \hat{x}_i + \beta \quad (7)$$

公式(6)和公式(7)则对特征进行重新映射,使得特征

值为被归一化到期望为 0,方差为 1 的范围内,同时学习参数和参数来还原原始的特征。公式(6)中的 ε 为一个极小的正值,用来防止产生除 0 的情况。公式(7)中的 γ 和 β 是可学习的参数,用来还原原始的特征。批归一化操作在加深网络深度的同时使得卷积层可以舍弃掉 bias 参数。

Kaiming he^[7]提出了 residual 模型。该模型的基本思想是在网络结构加深层次时,即使某些层无法学习到有效特征,仅将输入特征直接输出也不应降低网络的性能。显然网络无法学习到这种直接传递的方法,因此在加深网络深度的过程中,网络的性能反而会下降。通过将前层特征和卷积层输出求和,使得特征学习从学习映射函数 $F(x)=f(x)$ 变成 $F(x)=f(x)+x$ 进行残差的学习,让浅层的特征更容易传递到深层。Gao Huang 等^[8]提出了 Densely Connect 方法,其网络结构称为 DenseNet。该方法通过直接连接前层特征至后层,可以直接将特征传递给后续的网络层。这种直接传递特征的方法可以有效的重用特征并使得梯度的

传播更加良好(见图 2 中 DenseBlock)。由图 2 可知,DenseNet 使用一种分块的结构进行网络结构的组织,每个块中是一系列的批归一化,ReLU 和卷积层。前层的输出通过直接连接传递到每个后层作为输入。在每个 block 块最后是一个可选的 BottleNeck 层,用来压缩 featuremap 的通道(channel)数量。

通过搭建额外的特征传递通路,现代深度神经网络已经可以达到相当高的深度。如 ResNet 使用传递残差的方式,将网络的深度扩大到了千层。此后,简单增加网络深度的方法已经无法取得更好的效果。

文中提出了一种新型卷积神经网络结构,大幅改进了 DenseNet 的网络结构。通过增加块间跳跃连接,相比于普通编码-解码网络,提高了特征复用,加强前层细节的传递。同时引入带孔卷积,通过带孔卷积模拟 pooling 操作,增强毛细血管分割效果。层间的 batch normalization 可以降低梯度消失等风险,并加速网络的收敛。最后使用加权交叉熵作为损失函数,优化分割结果。

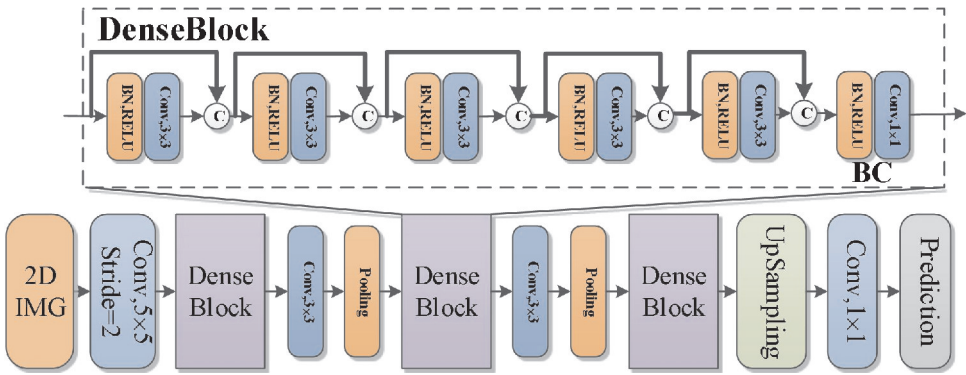


图 2 DenseNet 网络结构

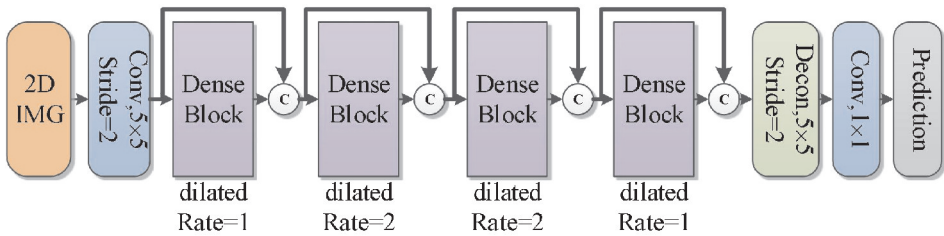


图 3 DNP 网络结构

2 DenseNoPooling 网络

采用稠密连接,批规范化等技术构建了一种新型的深度卷积神经网络,DenseNoPooling (DNP)。该结构使用 DenseNet 作为网络框架,舍弃了 pooling 进行下采样的操作,采用带孔卷积(dilated convolution)^[9]获得近似于 pooling 的感受野,以此来获得更精确的分割效果。在此之上,网络使用了块间稠密连接来传递浅

层的细节特征。块间稠密连接在块间通过直接连接的方式传递特征,这种稠密连接是一种冗余的跳跃连接,可以优化浅层特征的传递并提高分割细节。

2.1 DenseNoPooling 网络结构

DNP 网络如图 3 所示,该网络结构使用与 DenseNet 类似的成块结构进行网络的堆叠。这种简单的构造方法使得网络便于调整以适应不同的需求。仅通过增减 block 数量以及 block 内卷积层的数量,就可

以调整网络的深度,容量,感受野,使得网络可以满足不同复杂程度的应用。

提出网络在块内使用了与 DenseNet 相同的稠密连接,这种短距离的特征传递可以通过特征重用提高网络性能。相较于传统神经网络,这种结构可以减少参数数量,并使得网络同时兼顾浅层特征和深层特征。

同时,在 DenseNet 的基础上,剔除了层间的卷积和池化层的部分。正如前面所介绍的,传统神经网络中,特别是分类网络中,最大池化(maxpooling)作为一种增加感受野的下采样方法,被广泛使用。但在分割网络中,需要输入与输出有相同的尺寸。常见分割网络通常使用反卷积等上采样的方法进行特征恢复,但是反卷积或者插值会造成一定的精度损失。考虑到血管分割是一个对毛细血管敏感的分割任务,提出方法剔除了 pooling 层,通过使用带孔卷积的方法来弥补未使用 pooling 层所损失的感受野。其中,带孔卷积是一种特殊的卷积方式,相比于传统的卷积,带孔卷积的卷积拥有更大的感受野。通过使用有间隔的卷积核,带孔卷积可以在使用相同参数数量的情况下获得更大的感受野。卷积核中的间隔也被称为孔径(dilated rate)。图4(a)为一个普通的卷积核,图中每个红点是中心点进行卷积运算的计算点,绿色区域则为中心点的感受野。在这种情况下,网络每个点拥有一个 3×3 的感受野。如果在图4(a)的基础上使用一个孔径为2的卷积,则结果如图4(b)所示。该方法中卷积核在与 feature map 进行计算时会增加一个间隔,这种不连续的卷积会使网络结构拥有更大的感受野。在图4(c)中,通过使用孔径为4的卷积进行计算,结果中每个点的感受野最终扩大到了 15×15 ,而3个连续的普通卷积只能获得 7×7 的感受野。

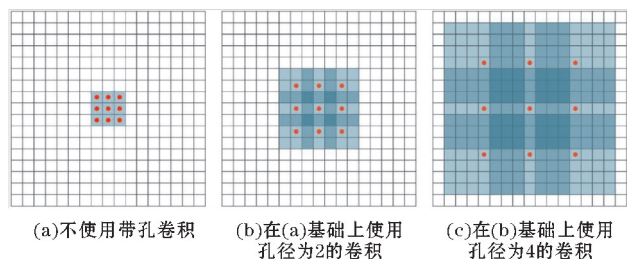


图4 带孔卷积

2.2 损失函数

在二分类分割任务中,一般使用交叉熵作为损失函数,具体结构如下式所示:

$$L = tar \times (-\log(S(l))) + (1 - tar) \times (-\log(1 - S(l))) \quad (6)$$

其中, tar 为标签,其值1为血管点,值为0表示非血管点, $S(\cdot)$ 为 sigmoid 函数, l 为网络输出。公式(6)

根据网络输出的预测概率结合标签进行损失计算。可知,当网络预测正确且得出预测的概率较高时,损失函数有更低的值。在分割任务中,总的损失函数使用每个分割点取平均的方法获得。但是视网膜血管分割是一类正负样本极为不均的二分类任务,其中大量像素点都为背景像素(非血管点),仅有少数点为正样本点(血管点)。如图5所示,当使用交叉熵进行训练时,可能导致网络将所有点预测为负样本点。因此,使用加权交叉熵作为损失函数。

$$L = tar \times (-\log(S(l))) \times w + (1 - tar) \times (-\log(1 - S(l))) \quad (7)$$

其中,权重 w 作为超参数,通常赋值为负样本与正样本之比。额外的权重可以帮助网络获得更有效的收敛,对正样本错误预测会产生更大的惩罚。这种惩罚能有效防止网络因正负样本不均而将所有结果预测为负样本。同时,使用加权交叉熵可以有效的提高网络的检出率,这在医学分类上有一定的意义,更高的检出率能有效的减少漏诊率。

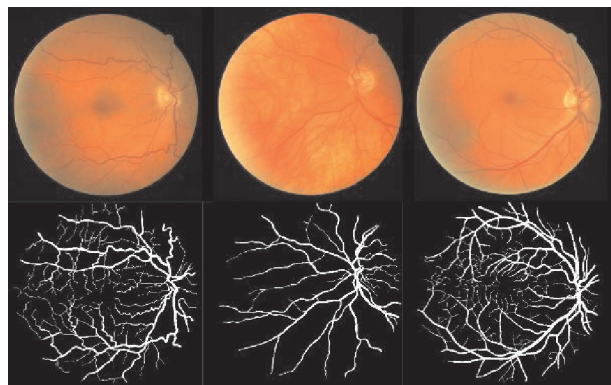


图5 DRIVE 数据集

3 实验与分析

3.1 实验设置

实验硬件环境为 Intel Xeon E5 $\times 2$, 256G 内存, Nvidia Tesla m40, 软件环境为 ubuntu16.04, cuda8.0, tensorflow1.4。为了验证网络性能,选取了性能较好的 U-Net 作为对比网络。所有网络均使用 adam 作为网络的优化器算法,初始学习率设置为 0.01,并在每 10 次循环(epoch)后减少至原学习率的 90%。在经过 200 次循环后网络均达到收敛,停止训练。

3.2 数据集

DRIVE^[10]是一个广泛采用的视网膜血管数据库,这个数据库来自于荷兰的一个糖尿病视网膜病变的筛查项目,筛查的人群包括 400 位年龄在 25 ~ 90 岁的糖尿病患者。研究人员从中随机选择了 40 例图片作为整个数据集数据,其中 33 名病人无糖尿病视网膜病变,7

名显示出轻微的视网膜病变。数据的格式统一为 565×584 的 JPEG 图片。其中 20 例病人被选取为训练集,并提供了对应的金标准分割结果作为训练标签。另外 20 例病人作为测试集并提供了两个参考的分割结果。图 5 是数据集中的 3 份样例,其中上一行为彩色眼底视网膜照片,下一行为照片对应的医生标注血管。

在所有网络的训练中,都使用了同样的数据增强方法。对所有输入图片和标签同时随机缩放0.9~1.2倍大小,并随机裁剪至 512×512 尺寸。

3.3 实验结果

实验采用精确度,敏感度,特异性,以及 ROC (receiver operating characteristic) 曲线, AUC (area under curve) 作为评测指标。精确度,敏感度和特异性的计算公式如下表 1 所示。

表 1 评估指标及计算方法	
性能评价指标	计算公式
特异性	$\frac{TN}{TN+FP}$
敏感度	$\frac{TP}{TP+FN}$
精确度	$\frac{TP+TN}{TP+TN+FP+FN}$

其中 TP 为正确分类的血管点, TN 为正确分类的非血管点, FP 为误分成血管点, FN 为误分为非血管点的。ROC 曲线指标用真阳率作为纵轴,假阳率作为横轴,通过从 1 到 0 减小分类阈值,在网络给出的概率分别进行真阳率和假阳率的计算。最后绘制成 ROC 曲线图。AUC 指标是在 ROC 基础上,计算曲线与横轴中间区域

的面积所得的指标,值越高表示分类结果越好。

表 2 评估指标及计算方法				
方法名称	精确度	敏感度	特异性	AUC
DNP	0.9617	0.7325	0.9839	0.97824
U-Net	0.9577	0.719	0.980	0.95257
文献 ^[11]	0.9606	0.7447	0.9838	0.9632

由表 2 可知,提出的网络在精确度,敏感度,特异性, AUC 4 项指标中都取得了超过对比网络的结果。图 6 展示了提出网络 and 对比网络的 ROC 曲线。

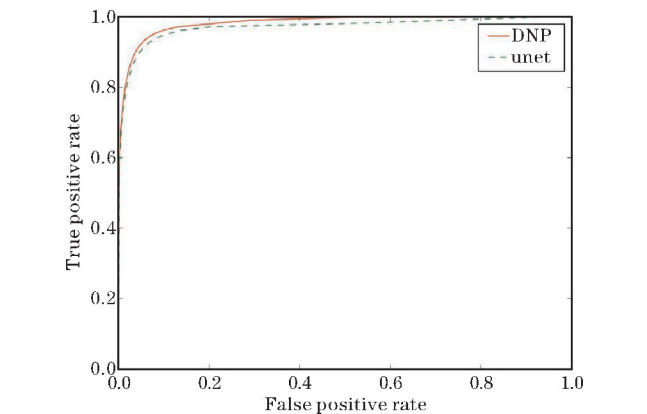


图 6 ROC 曲线

可视化血管分割结果如图 7 所示。可以明显的看出,提出的网络可以获得更加接近于医生标注的分割结果。对比提出网络 and 对比网络的分割结果可知,方法在主血管和血管上的毛细分支都获得了较好的分割。相比于对比网络,文中网络可产生更少的噪音分割。而在 U-Net 的分割结果中,有大量细小且不连续的分割结果,而在提出网络中,这种噪音相对较少。

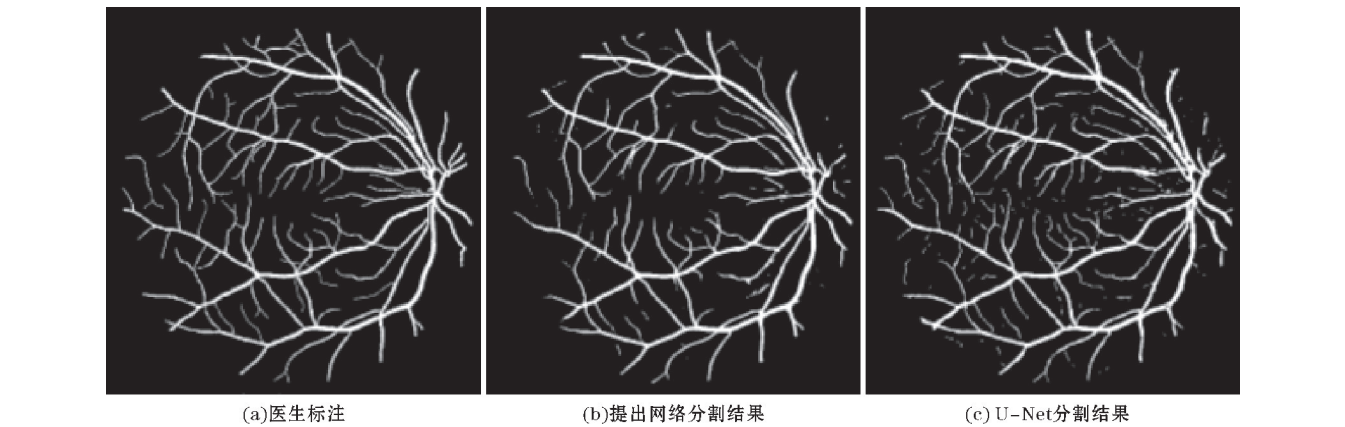


图 7 提出网络 and 对比网络的分割结果

4 结束语

针对视网膜血管分布广泛,毛细血管较细,分割较难的情况,提出了一种新型的卷积神经网络结构,通过

使用带孔卷积,在保证卷积神经网络感受野的同时提高了分割的精度。同时引入块间的稠密连接,加强了梯度传递。经过对比实验,验证了提出网络的分割性能。

致谢:感谢成都信息工程大学科研基金 (KYTZ201608) 对本文的资助

参考文献:

- [1] Lan Goodfellow, Yoshua Bengio, Aaron Courville, 等. 深度学习 [M]. 北京: 人民邮电出版社, 2017: 203–307.
- [2] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C]. International Conference on Neural Information Processing Systems. Red Hook, NY: Curran Associates, Inc. 2012: 1097–1105.
- [3] Abadi M, Agarwal A, Barham P, et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems [OL]. <https://arxiv.org/abs/1603.04467v1>, 2016.
- [4] Shelhamer E, Long J, Darrell T. Fully Convolutional Networks for Semantic Segmentation [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014, 39(4): 640–651.
- [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. Europe: Springer, Cham, 2015: 234–241.
- [6] Ioffe S, and Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [C]. International Conference on International Conference on Machine Learning. Lille, France, 2015: 448–456.
- [7] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition [C]. IEEE Conference on Computer Vision and Pattern Recognition. United States: IEEE, 2016: 770–778.
- [8] Huang G, Liu Z, Maaten L V D, et al. Densely Connected Convolutional Networks [C]. IEEE Conference on Computer Vision and Pattern Recognition. United States: IEEE, 2017: 2261–2269.
- [9] Yu F, Koltun V. Multi-Scale Context Aggregation by Dilated Convolutions [OL]. <https://arxiv.org/abs/1511.07122>, 2015.
- [10] Staal J, Abràmoff M D, Niemeijer M, et al. Ridge-based vessel segmentation in color images of the retina [J]. IEEE Transactions on Medical Imaging. United States: IEEE, 2004, 23(4): 501–509.
- [11] 朱承璋, 崔锦恺, 邹北骥, 等. 基于多特征融合和随机森林的视网膜血管分割 [J]. 计算机辅助设计与图形学学报, 2017, 29(4): 584–592.

Automatic Retinal Vascular Segmentation Method based on Densely Connected Convolution Neural Network

TANG Ming-xuan, LI Xiao-jie, ZHOU Ji-liu

(College of Computer Science and Technology, Chengdu University of Information Technology, Chengdu 610225, china)

Abstract: Deep convolutional neural network (DCNN) has shown its superior performance in image classification and segmentation problems. It has been extensively studied and also promotes the medical image segmentation development. Fundus retinal blood vessels are non-invasive directly observed blood vessel, which can provide one of the most important evidences for the diagnosis of ophthalmic diseases. For example, capillary proliferation provides important guidance for the diagnosis of ocular diseases such as diabetes. Therefore, a correct and efficient method of retinal blood vessel segmentation becomes a clinical requirement. In this paper, we propose a method of densely connected-based convolution neural network for retinal blood vessel segmentation. The proposed innovative network employs densely connect for reusing features and enhancing feature delivery. This method uses batch normalization enabling the network to converge to better results with less time. Combined with the dilated convolution, the network can get more accurate segmentation results. By this method, we can gain a better performance without using pre-processing and post-processing. Through the comparison experiments with traditional machine learning methods and other DCNN segmentation methods, the proposed method can achieve better average accuracy and sensitivity. The specificity reached 0.9617, 0.7325, 0.9839, and a pixel wide AUC reached 0.978. Experimental results demonstrate the effectiveness and efficiency of our method.

Keywords: deep learning; convolutional neural network; image segmentation; retinal blood vessels