

硕士学位论文





学位论文原创性声明

本论文是我个人在导师指导下进行的工作研究及取得的研究成果。论文中 除了特别加以标注和致谢的地方外,不包含其他人或其它机构已经发表或撰写 过的研究成果。对本文的研究做出贡献的个人和集体,均已在论文中以明确方 式标明。本人完全意识到本声明的法律责任由本人承担。

作者签名: ______ 日期: _2021 年_6月_1日

学位论文使用授权声明

本人授权汕头大学保存本学位论文的电子和纸质文档,允许论文被查阅和 借阅;学校可将本学位论文的全部或部分内容编入有关数据库进行检索,可以 采用影印、缩印或其它复制手段保存和汇编论文;学校可以向国家有关部门或 机构送交论文并授权其保存、借阅或上网公布本学位论文的全部或部分内容。 对于保密的论文,按照保密的有关规定和程序处理。

作者签名:6) 译	导师签名:
日期:年6月日	日期: <u>2*21</u> 年 <u>6</u> 月 <u>1</u> 日



硕士学位论文

- 论文中文题目: 基于自校准卷积的视网膜血管分割
- 论文英文题名: Self-Calibrated Convolution U-Net for Retinal Vessel Segmentation
- 指导教师: 范衡
- 申 请 人 : 熊宇
- 论文答辩委员会成员
- 主席: 陈耀文 教授 (汕头大学生物医学工程系)
- 委员: 庄哲民 教授 (汕头大学电子系)
 - 熊智 教授 (汕头大学计算机系)
 - 邢阳辉 副教授 (汕头大学生物医学工程系)
 - 杨金耀 教授级高工 (汕头市超声仪器研究所有限公司)

摘要

眼底图像中的视网膜血管作为人体唯一能够直接观察到的深层微血管,通过分析 血管的长度、直径、分支模式、弯曲度和角度信息是医生诊断许多严重疾病的重要方 法,比如患有心血管疾病、糖尿病、眼科疾病的患者,其视网膜血管结构通常会发生 改变且血管周围存在病变区域。眼底图像具有获取方便快捷的特点,而视网膜血管的 精确分割是诊断和筛查这些疾病的基本步骤。人工分割眼底血管耗时费力,分割效果 依赖于专科医生的主观感受,因此越来越多的自动分割技术被提出来解决这一问题, 但是许多方法在分割准确率上还不够。因此有必要设计一个更准确高效的算法来实现 视网膜血管的自动分割。

在本文中,本人提出了一种自校准卷积 U-Net(SCCU-Net)来实现视网膜血管的精确分割。在 SCCU-Net 中,采用自校准卷积模块代替传统的卷积运算,自校准卷积不像传统卷积一样均匀地对原始空间中的输入特征执行所有卷积,而是首先通过下采样将输入特征转换为低维特征并卷积,然后通过卷积变换后的低维特征来校准另一部分中的卷积变换。得益于这种异构卷积,可以有效地扩大每个空间位置的感受野,可以获得更丰富的输出信息。此外,提出了一种改进的空间注意模块,该模块在卷积运算前通过获取空间注意权值来获得更完整的空间信息。

本文提出的算法在 DRIVE, STARE 和 CHASE DB1 三个公开可用的数据集上进行了评估。通过对比实验结果表明,SCCU-Net 相对于 U-Net 在分割结果上有较大提升。与当前先进的方法相比,所提出的方法也能取得更好或相当的视网膜血管分割结果。

关键词:视网膜血管分割;自校准卷积;空间注意力机制

I

Abstract

Retinal vessels in fundus images are the only deep microvessels that can be directly observed in the human body, it is an important method for doctors to diagnose many serious diseases by analyzing the length, diameter, branching mode, curvature and angle information of vessels, for example, in patients with cardiovascular disease, diabetes, and eye disease, the structure of the retinal vessels is often changed and there are disaffected areas around the vessels. Fundus images are easy to obtain and accurate segmentation of retinal vessels is the basic step for diagnosis and screening of these diseases. Artificial segmentation of fundus vessels is time-consuming and labors, and the segmentation effect depends on the subjective perception of the specialist. Therefore, more and more automatic segmentation technologies have been proposed to solve this problem, however, many methods are not enough in segmentation accuracy. Therefore, it is necessary to design a more accurate and efficient algorithm to realize the automatic segmentation of retinal vessels.

In this paper, I propose a self-calibrated convolutional U-Net (SCCU-Net) to achieve accurate segmentation of retinal vessels. Convolution in SCCU-Net, using the calibration module instead of the traditional convolution operation, unlike traditional convolution, which uniformly performs all convolution on the input features in the original space, self-calibration convolution first transforms the input features into low-dimensional features and convolves them through subsampling, and then calibrates the convolution transform in the other part through the low-dimensional features after convolution transformation. Thanks to this heterogeneous convolution, the receptive field of each spatial location can be effectively expanded and richer output information can be obtained. In addition, an improved spatial attention module is proposed, which can obtain more complete spatial information by obtaining spatial attention weights before convolution operation.

The proposed algorithm was evaluated on three publicly available datasets, DRIVE, STARE and CHASE DB1. The comparative experimental results show that SCCU-Net has a great improvement in segmentation results compared with U-NET. Compared with the current state of the art methods, the proposed method can also achieve better or comparable results in the segmentation of retinal vessels.

Keywords: Retinal Vessel Segmentation, Self-Calibrated Convolutions, Spatial Attention Module

	ヨ
日	求

摘要I
AbstractII
目 录IV
第1章 绪论1
1.1 研究背景及意义1
1.2 相关研究概况
1.2.1 基于无监督的学习方法
1.2.2 基于有监督的学习方法4
1.3 主要研究内容7
1.4 论文章节安排
第2章 相关工作9
2.1 卷积神经网络9
2.1.1 卷积神经网络基本组件9
2.1.2 经典的卷积神经网络结构11
2.2 全卷积神经网络15
2.3 基于编码器解码器的神经网络16
2.3.1 SegNet
2.3.2 U-Net
2.4 注意力机制
2.4.1 注意力机制的分类18
2.4.2 常见的注意力模型18
第3章 基于自校准卷积的视网膜血管分割算法22
3.1 整体网络结构设计22
3.2 自校准卷积算法
3.2.1 自校准卷积的结构24
3.2.2 自校准卷积操作
3.3 空间注意力机制的设计25
3.4 损失函数
第4章 实验结果与分析
4.1 数据集概述

4.2 评价指标	
4.3 实验设计	
4.3.1 实验环境	31
4.3.2 实验参数设置	31
4.3.3 实验设计	32
4.4 结果与分析	
4.4.1 对比实验结果与分析	
4.4.2 SCCU-Net 详细分割结果	
4.4.3 与先进视网膜血管分割方法比较	
第5章 结论及展望	42
5.1 结论	
5.2 课题展望	
参考文献	44
附录 A DRIVE 数据库测试集详细实验结果	50
附录 B STARE 数据库测试集详细实验结果	51
附录 C CHASE DB1 数据库测试集详细实验结果	52
攻读学位期间主要研究成果	53
致谢	54

第1章 绪论

1.1 研究背景及意义

随着经济社会的快速发展,伴随着人类生活水平的不断提高,越来越多人的工作 方式逐步由体力劳动转变为脑力劳动,许多工作者由于长时间面对电脑手机等电子设 备,眼部得不到休息,这部分人群中患眼科疾病的数量呈不断上升的趋势。青光眼、 糖尿病视网膜病变等疾病为比较常见的眼科疾病,这些眼科疾病严重的可导致失明。 据相关报道,糖尿病视网膜病变是老年人失明的主要诱因,至2019年,中国糖尿病 患者的人数已达1.22亿,高居全球首位,预计到2028年这一人数将升至1.49亿。因 此定期对眼部进行筛查是很有必要的,眼睛的每个部分都对拥有清晰的视力很重要。 视网膜作为人类眼睛最基本的组成部分,提取其视网膜血管的形态学特征,如长度、 直径、分支模式、弯曲度和角度信息,是医生诊断许多眼科疾病的基础。例如,狭窄 且反射亮度高的小动脉血管能诊断高血压,而分叉角度和弯曲可以帮助检测糖尿病视 网膜病变和心血管疾病。因此视网膜眼底图像的分析离不开视网膜血管特征的提取, 而视网膜血管分割是提取视网膜血管特征的主要方法。为了定量分析疾病,视网膜血 管分割已成为检测分析眼部疾病相关工作的最为重要的一步,对人类疾病的诊断具有 指导意义。

在临床实践中,视网膜血管通常由眼科医生选取眼底图像手动标注。这种手动标 注不仅占用了医生大量的时间,而且容易出错。再者由于不同眼科医生的经验不同, 主观感受不同,即使标注同一幅眼底图像,也会出现标注的结果不统一的情况。并且 当数据量增加时,人工对视网膜血管进行标注将耗费大量的人力物力,而分割精度参 差不齐。因此,为了减少人工分割的工作量,提高分割的精度和速度,近几十年来, 大量的研究工作致力于视网膜眼底血管的自动分割方法,研究这些自动分割视网膜血 管的算法具有重大的临床实践意义。

与此同时,目前存在的视网膜血管分割算法依然难以满足辅助计算机医疗诊断的 要求,实现视网膜血管的高精度分割依旧是一个充满挑战的研究课题。在视网膜血管 分割中存在诸多不利因素影响着分割效果。主要因素包括:

(1) 血管形态多样,粗细不一。微细血管的直径仅仅只有1个像素点(图1-1a), 而粗血管的直径能达十几个甚至几十个像素点(图1-1b),如果算法专注粗血管, 那么微细血管会出现被漏检的情况。而算法过于关注微细血管,会大幅降低粗血 管的分割精度,出现断层或模糊的情况。 (2) 成像缺陷。视网膜图像的获取是将眼球的三维结构映射到二维平面上,这个 过程很容易产生对比度低的视网膜血管图像(图 1-1 c),这些眼底图像的血管区 域像素与背景像素对比度较低,图像本身就模糊不清,即使是人类肉眼观察也很 难分辨出来血管部分,这对于算法而言是一个重大的考验。

(3) 病灶图像及眼底周边结构的影响。现有许多算法能够在健康的视网膜图像中 取得优异的分割结果,但是在临床诊断中,由于患者自身具有影响正常血管结构 的疾病,因此在临床中采集的眼底图像与正常的眼底图像具有较大差别。例如: 由于病变导致的血管出血(图1-1 d)、渗出物、棉絮状斑(Cotton Wool Spots) 都会影响视网膜血管的分割。此外,其他组织结构的背景图像也会出现在视网膜 眼底图像中,如视盘(Optic Disc)(图1-1 e)、黄斑(Macular)等结构,这些结 构容易对算法产生干扰,导致算法将这些结构误判为血管。



图 1-1 形态各异的视网膜血管。从(a)到(e)分别为细血管、粗血管、对比度低血管、病灶区 域血管、视盘附近血管。

基于以上因素,目前辅助计算机医疗诊断的算法还不完善,本文提出基于自校准 卷积的 U-Net 网络模型可以一定程度上解决以上问题,不仅对正常的眼底图像有良好 的检测分割能力,对存在病灶区域的视网膜图像的检测分割也有较好鲁棒性。

1.2 相关研究概况

在医学临床实践中,视网膜血管分割任务作为最重要的一环具有重大意义,因此 越来越多的研究者致力于开发出高效精准的视网膜血管分割方法,在现有方法中,用 于视网膜血管分割任务中主要有两类机器学习方法:基于无监督的学习方法和基于有 监督的学习方法。前者发现一个模型能够发现隐藏在数据中有效的特征信息及其概率 分布,而无需任何先验知识或相关反馈等指导性信息,而后者从一组已经标记的特征 中学习数据模型,基于无监督学习和有监督学习的视网膜血管分割方法分别在 1.2.1 和 1.2.2 节中介绍。由于无监督学习不需要金标准图像,它对于公开的数据集在当前 任务中不可用的情况以及探索性数据分析是有用的。另一方面,监督学习需要将金标 准图像用于训练学习模型,训练的计算成本取决于所采用的监督学习方法。然而,在 测试阶段,计算成本通常可以忽略不计。

1.2.1 基于无监督的学习方法

1.2.1.1 基于匹配滤波的方法

匹配滤波(Matched Filter)是一种经典的模板匹配技术,由 Chaudhuri 等人^[1]首 先提出。匹配滤波方法假定血管局部呈线性,根据血管强度轮廓进行建模并作为二维 滤波器,利用局部图像区域和滤波核之间的相关性,可以增强血管与背景的对比度, 从而得到更优的血管结构。

Ricci 等人^[2]提出将线检测器作为移动滤波器用于视网膜血管的分割。线检测器从 不同方向通过目标像素沿固定长度的线计算平均强度。此外,Nguyen 等人^[3]认为, 通过改变基本线检测器的长度,可以将线检测器与多尺度技术相结合。最终的分割是 通过不同尺度的直线响应的线性组合来实现的。带线检测器的匹配滤波器以原始血管 图像作为输入,但有时会在含有微细血管密集和背景噪声严重的眼底图像上失去作用。 因此,Ding 等人^[4],提出了一种用于三维血管分割的集成滤波器。该方法结合了多尺 度线检测器和 Hessian 矩阵滤波器的优点,在微细血管分割中表现良好。

基于 Hessian 矩阵和线检测器方法有一个共同的局限性。他们隐含的前提是图像局部只有单一管状结构,这在复杂血管的应用中^[3,5]是不利的。Qian 等人^[6]提出了一种不依赖单一管状结构图像的分割方法。它捕获了极性邻域强度分布在所有类型血管点中的共同特性。该方法提高了基于球面极坐标邻域强度剖面的血管分叉性能。然而,它对接近背景结构的结构噪声很敏感。因此,Rivest-Henault等人^[7]定义了一种基于结构球(SBall)分析的三维曲线结构检测滤波器。结构球是通过计算各个方向的二阶数值差来构造的。由于它不是检测特定的血管轮廓的检测滤波器,它可以用来表示交叉的结构。

除了在空间域中使用单一特征外,匹配滤波器还可以采用多种设计方法。Zhang 等人^[8]采用零均值高斯滤波器和高斯函数的一阶导数来检测血管。Lajevardi 等人^[9]使 用一系列频域匹配滤波器提取血管,其中血管模板定义为从视网膜血管网络派生的形 式空间图。此外,Kovacs 等人^[10]提出了一种基于广义 Gabor 函数的匹配滤波模板。 他们利用提取的中心线和从训练数据库中获得的先验知识重建血管的轮廓。

1.2.1.2 基于血管追踪的方法

血管追踪和跟踪方法通常从边缘或中心线选择的初始种子点开始跟踪血管,然后 根据血管路径进行跟踪直到结束,这样就确定出了与起始点之间的血管结构。该方法 使用局部信息,通常在具有检测难度的分岔和交叉点提供血管宽度、连接性信息。

Vazquez 等人^[11]利用 Snack 模型、梯度方向和最小路径提取血管轮廓特征,通过 k-means 聚类将血管段分类为动脉和静脉。Bekkers 等人^[12]采用不同尺度的线和多重 方向检测技术,对比较难以检测的分叉区域进行血管边缘和中心线跟踪。Khan 等人^[13]

3

利用 Otsu 的阈值和张量着色生成二元血管图,从而更好的分割微细血管。Wang 等人 ^[14]提出了一种多尺度形态学和种子点跟踪方法来增强并分割视网膜血管的方法。该方 法首先通过顶帽和底帽变换二者的综合应用获取每个尺度的高对比度血管图像,然后 通过基于多阈值的血管跟踪技术对血管图像进行分割,在跟踪阶段使用血管像素的比 例自适应获得阈值,并且可以在该过程中自动计算停止条件。

1.2.1.3 基于形态学的方法

传统的数学形态学方法利用结构元素(圆形结构、正方形结构、线性结构等)进行 医学图像分析,依赖于腐蚀、膨胀、开闭运算等基本操作^[15]。在提取图像边缘轮廓时, 这些数学形态学操作通常被用作预处理,通过腐蚀膨胀操作可以获取图像边缘轮廓, 而开闭运算操作可以用来填补提取边缘轮廓过程中产生的空洞^[16]。后来,更复杂的形 态学方法被开发出来,包括分水岭算法、灰度级击中或击不中变换算法、连接过滤器 算法等,这些方法也被应用到三维血管图像分割^[17,18]。

Fan 等人^[19]针对眼底图像中不确定的区域提出了一种用于血管分割的分层抠图 模型,有效利用了不确定区域中的像素,将不确定区域中的像素通过该模型归类为血 管像素或者为背景像素,最终取得了显著的效果。Fraz 等人^[20]首先通过高斯滤波器提 取血管中心线,然后通过形态学操作顶帽变换对视网膜血管图像进行增强,最后得到 分割出的血管网络图像。Babin 等人^[21,22]通过引入更复杂的环形结构元素扩展了形态 轮廓,并进一步定义了新的基于方向的线形结构元素。因此,线状结构元素将多尺度 方法的思想融入到多方向形态学方法中,提高了血管分割的计算效率。此外,Dufour 等^[23]基于校正的方向场对三维脑血管图像进行空间变化的形态学闭合操作。在 Hessian 矩阵阈值度量的指导下,采用细化处理和膨胀操作,对由 Hessian 矩阵计算的 初始方向场进行了细化。Sigurosson 等人^[24]将定向数学形态学与模糊分类相结合。利 用基于血管对比及其线性连通性的先进形态定向滤波器进行特征提取后,根据这些特 征的融合进行像素分类。

得益于结构元素的精美设计,复杂数学形态方法相对于基本数学形态方法的明显 优势在于其自适应的方向和多尺度识别能力。然而,上述复杂的形态学方法在进行方 向估计时,往往无法对微细血管进行判断。因此,利用复杂数学形态学方法对多尺度 血管的准确操作还需要进一步的研究。

1.2.2 基于有监督的学习方法

基于有监督技术的方法需要一些先验知识来判断像素是血管像素还是非血管像 素,这些方法通过人工处理的金标准视网膜图像训练出一个判别模型。真实准确的视 网膜图像数据是这些方法的前提条件。有监督的分割方法中的数据是已经提前分类好 的数据,其训练样本中通常包含有标签信息,所以一般来说,它们的性能要优于无监 督的分割方法。

1.2.2.1 基于人工神经网络的方法

人工神经网络(ANN)已被大量研究用于分类。Nekovei和 Sun^[25]提出了一种基于反向传播神经网络的方法,用于 x 射线血管造影对视网膜血管的识别。将人工标注血管图像的金标准图像作为神经网络权值的训练集。为了找到这些权重,应用了一个自定义的增量规则,该方法只对血管像素进行标记,忽略眼底图像中血管网络的提取。

Rahebi 和 Hardalaç^[26]提出了一种通过图像灰度共生矩阵(Gray Level Co-occurrence Matrix, GLCM)提取特征进行视网膜血管分割的方法。首先选取视网 膜彩色图像的最佳波段进行预处理, 然后应用新的局部处理函数对视网膜图像中的亮度进行修改, 该过程可以得到明亮均匀的图像, 并采用多层感知器进行像素分类。

Li 等人^[27]提出了一种基于跨模态图像变换的系统。通过使用基于卷积神经网络 (CNN)的映射函数,将相同尺寸的图像作为第一模态,并将其对应的血管映射作为第 二模态。利用绿色通道视网膜图像的小块,制作相应血管标签的训练数据。

Liskowski 等人^[28]提出了一种基于卷积神经网络(CNN)的方法,对40万张样本 图像进行了训练,用于血管分割。在预处理阶段,对图像进行全局对比度归一化

(Global Contrast Normalization)、几何变换、零相位白化(Zero Phase Whitening)和 Gamma 校正。该系统通过结构化预测对多个像素同时进行分类。

1.2.2.2 基于 SVM 的方法

Ricci 和 Perfetti^[2]提出了一种利用两个正交的线检测器提取特征的方法,并将其应用于建立支持向量机(Support Vector Machine, SVM)分割血管的特征向量,该方法通过线强度的局部微分计算来构建强线性特征描述子,该方法的优点是运用简单的特征提取就能完成视网膜血管分割的有监督学习,且只需少量的样本数据。

Osareh 和 Shadgar^[29]使用多尺度 Gabor 滤波器检测视网膜图像中的候选血管,并使用 PCA 进行特征提取,最后采用 SVM 和高斯混合模型(GMM)对血管进行分割。

Xu和 Luo^[30]提出了一种基于多种图像处理技术的支持向量机分组的视网膜血管 分类方法。利用这些技术对绿色道背景进行归一化处理,分割血管,分离视盘边缘。 然后利用不同尺度的小波对原始图像进行特征提取。为了识别微细血管,使用了线检 测测器,对视网膜二值图像中的每个像素建立了12维特征向量。为了得到血管网络, 支持向量机从每个片段中分化出细小的血管段。最后,由 Hessian 矩阵得到血管路径 和特征向量。

1.2.2.3 基于统计学习的方法

Niemeijer 等人^[31]提出了一种提取像素绿色通道分量特征的 KNN 分类器。为了计算血管像素的概率,每个特征被标准化为零均值和单位方差,采用阈值分割和概率映

射的方法对血管进行分割。

Soares 等人^[32]提出了一种基于"二维 Gabor 小波"的视网膜血管分割技术,该技术 通过计算每个像素在不同尺度下二维 Gabor 小波变换的像素强度和响应,形成一个特 征向量,用于视网膜血管的 GMM 分类器分割。

Xiao 等人^[33]使用带有空间约束的贝叶斯模型,提出了一种新的能量函数来提取 血管,并采用自定义水平集方法来降低能量函数。该模型假设每个像素的后验概率取 决于其邻近像素的后验概率。该方法由于利用了丰富的空间信息,在检测低对比度血 管结构上有良好的表现。

Roychowdhury^[34]提出了一种基于高斯混合模型(GMM)分类器的新方法。为了 得到血管区域,在图像绿色通道上进行高通滤波后提取二值图像,并通过形态学重建 增强图像对血管区域提取新的二值图像。在此基础上,利用 GMM 分类器提取两幅图 像中的相同区域进行血管分割。

Azzopardi^[35]在现有 COSFIRE 滤波器的基础上,提出一种移位滤波器响应组合 (B-COSFIRE),能准确检测不同方向的主干血管和血管末端,尤其对微细血管有很好 的检测效果。

1.2.2.4 基于集成分类器的方法

Fraz 等人^[36]提出了由增强决策树(Boosted Decision Trees)和袋装决策树(Bagged Decision Trees)组成的集成分类器方法。该方法采用了包含血管特征图、Gabor 滤波响应、线强度测度、形态变换等9维特征向量。袋装集成算法计算了分类的准确性,这些特征是在不含测试数据的训练过程中获得的。

Wang 等人^[37]提出基于卷积神经网络(CNN)和随机森林(Random Forest, RF) 的有监督血管分割方法。该方法利用传统分类器的优点,在特征提取阶段利用卷积神 经网络提取血管的多层次特征,然后送入随机森林进一步训练并得到最终的血管分割 结果。该算法有效避免完全利用人工特征参数设计量大的弊端,且能较好地保证血管 特征的旋转不变性和尺度不变性,总体鲁棒性较高。

Fan 等人^[38]提出了一种基于整体通道特征和随机森林的数字视网膜图像血管分割算法。第一阶段是对血管的候选像素进行预处理,然后基于整体通道对每个候选像素提取一系列简单特征。此外,利用集成决策树构建随机森林对候选像素进行分类作为血管或非血管。最后,采用后处理的方法对被分类血管中的像素间隙进行填充。

Akram 等人^[39]提出了一种基于集成分类器的糖尿病视网膜病变检测方法。该方法 分为三个阶段:第一阶段包括视网膜图像预处理、血管分割和视盘定位,第二阶段包 括利用滤波器组提取候选病变区域,第三阶段包含集成分类器(GMM 和 m-Mediods) 的特征集制定,以实现精确分类。 Memari 等人^[40]提出了一种匹配滤波技术和 AdaBoost 分类器相结合的方法,该方 法采用形态学运算和对比度受限的自适应直方图均衡化(CLAHE)对视网膜图像进行 预处理。血管图像经过 B-COSFIRE 和 Frangi 匹配过滤器增强后,在 AdaBoost 分类 器中使用各种基于统计特征的向量从视网膜图像中提取血管网络。

1.2.2.5 基于深度学习的方法

得益于卷积神经网络(CNN)自动学习图像特征的能力,其不需要人工预先提取特征作为先验知识,通过对深度神经网络结构和参数的优化,实现网络模型对图像特征的自动提取与识别。目前,基于深度学习的方法被广泛应用于医学图像分割、计算机视觉等领域。

Ronneberger 等人^[41]提出了一种基于全卷积网络(FCN)^[42]的语义分割网络U-Net, U-Net 采用编码器和解码器的结构,并在编码器和解码器之间采用跳跃连接,在解码 器中引入编码器的信息可以更好地恢复特征图的细节。

Li 等人^[27]将血管分割视为从视网膜图像到血管特征图的跨模态数据转换问题, 视网膜图像与血管映射之间的关系需要一种宽而深、拟合能力强的神经网络来建模。 这种方法减少了主观因素的影响,因为它不需要预处理,但代价是增加了计算量。

Alom 等人^[43]使用基于 U-Net 的递归卷积神经网络和基于 U-Net 模型的递归残差 卷积神经网络对视网膜血管进行分割,显著改善了分割的效果。

Wang 等人^[44]提出了双编码 U-Net(DEU-Net),显著提高了网络端到端和像素到像 素分割视网膜血管的能力。

1.3 主要研究内容

本文将实现高精度的视网膜血管分割任务作为核心问题,运用深度学习中的图像 处理技术,结合卷积神经网络强大的图像特征表达能力,提出基于融合自校准卷积与 空间注意力模块的视网膜血管分割算法,实现高精度的视网膜血管分割。具体研究内 容如下:

(1) 研究基于编码器-解码器架构的全卷积神经网络,实现端到端的视网膜血管 图像语义分割。

(2) 研究自校准卷积与改进的空间注意力模块在视网膜血管分割任务中的有效性。

(3) 研究融合自校准卷积与改进的空间注意力模块的视网膜血管分割算法,实现 视网膜血管的高精度分割。

1.4 论文章节安排

本文共包含五个章节,各章节具体内容如下:

第一章:绪论。本章节简述了视网膜血管分割任务的背景意义以及国内外的研究 现状,并确立论文的研究内容与论文章节安排。

第二章:相关工作。本章节主要介绍卷积神经网络、全卷积神经网络、基于编码 器解码器的神经网络以及注意力机制等理论基础知识。

第三章:基于自校准卷积的视网膜血管分割算法。本章节主要阐述本论文提出的 融合自校准卷积与改进的空间注意力模块的视网膜血管分割算法,以及对各个子模块 的介绍。最后采用交叉熵损失函数实现对视网膜血管的高精度分割。

第四章:实验结果与分析。本章节主要阐述实验所用的数据库、评价指标,以及 证明所提算法的有效性的对比实验设计,对实验结果进行具体分析,最后与其他先进 方法作对比。

第五章:总结与展望。本章节主要对全文的研究工作进行总结,阐述本文主要贡献,并对未来工作进行展望。

第2章 相关工作

2.1 卷积神经网络

20世纪60年代,生物学家们在猫的视觉皮层细胞实验中发现猫的视觉皮层神经 元对特定区域图像比较敏感,这一特定区域被称作感受野。后来感受野的概念被引入 到卷积神经网络(Convolution Neural Network, CNN),在一些简单的图像识别实验 中发现物体经过轻微位移或者形变后,卷积神经网络仍然能很好的识别图像中的物体。

近些年来,CNN 被广泛的应用在图像处理相关任务中。在计算机视觉领域,图像识别(Image Recognition)、图像分类(Image Classification)、图像语义分割(Image Semantic Segmentation)等子领域得益于CNN的成功,取得了非常不错的技术成果,这些技术成果也推动着于人类生活智能化的进程。此外,随着CNN研究的深入,在自然语言处理(Natural Language Processing)领域中,虽然CNN的应用要少于循环卷积网络(Recurrent Neural Network, RNN),但是将CNN与其他算法结合起来在语音处理中取得了不错的效果,例如将卷积神经网络与长短记忆模型(Long Short Term Memory model, LSTM)相结合可以很好地对输入句子进行补全。

2.1.1 卷积神经网络基本组件

卷积神经网络基本组件主要有卷积层、池化层以及全连接层,针对不同的任务, 卷积神经网络具有不同的结构,随着卷积神经网络的迅速发展,全连接层在许多任务 中被全卷积层代替。

2.1.1.1 卷积层

卷积层的作用是提取一个图像局部区域的特征,低层卷积层提取图像的浅层特征, 例如线条、端点、颜色、轮廓等。深层卷积层可以提取更高层的信息,结合提取到的 前层特征,可以得到丰富的特征语义信息。卷积核为提取图像特征的提取器,通常卷 积核由具有三维结构的矩阵组成,其尺寸为高度H×宽度W×深度D。特征映射

(Feature Map)为图像经过卷积处理后提取到的特征,为了得到更为丰富的特征映射,通常在同个卷积层使用不同的卷积核得到不同的特征映射。假设一个卷积层的结构如下:

(1) 输入特征映射: $X \in R^{H \times W \times D}$ 是具有三维结构的张量,组成输入特征 X 的切片 矩阵 $X^{d} \in R^{H \times W}$ 为其中的一个输入特征映射,其中 $1 \le d \le D$;

(2) 输出特征映射: $Y \in R^{H' \times W' \times C}$ 为具有三维结构的张量,由D个 $H' \times W' \times C$ 的特征映射相加得到,其中 $Y^d \in R^{H' \times W' \times C}$, $1 \le d \le D$ 。

(3) 卷积核: $K \in R^{h \times w \times C \times D}$ 为四维张量, C 为卷积核的通道数, D 为输入特征的 深度,卷积核中每个切片矩阵 $K^{C,d} \in R^{h \times w}$ 为一个两维卷积核,其中1 $\leq d \leq D$ 。



图 2-1 卷积层中从输入特征映射 X 到输出特征映射 y 的计算示例

卷积层的计算过程如图 2–1 所示。首先通过卷积核 *K*^{C1},*K*^{C2},…,*K*^{CD} 分别对输入 特征映射 *X*¹,*X*²,…,*X*^D 进行卷积操作,然后将卷积处理得到的子特征映射依次相加, 得到结果后加上偏置量 b,最后经过非线性激活函数 ReLU^[45]后得到输出特征映射 *Y*。

$$Y = f\left(K \otimes X + b\right) = f\left(\sum_{d=1}^{D} K^{C,d} \otimes X^{d} + b\right)$$
(2-1)

公式 2–1 中 $K \in R^{h \times w \times C}$ 为三维卷积核,h = w分别表示卷积核的长和宽,C表示卷积核的通道数,f(*)表示非线性激活函数,在实际使用中一般用 ReLU^[45]函数。 2.1.1.2 池化层

池化层是卷积神经网络中常用的网络层,通常跟随在卷积层后面,对卷积层的输出特征图进行处理。池化操作通常使得输入的特征图在空间维度上变为原来的一半, 避免过拟合的同时减少网络参数数量,带来更小的计算资源消耗。常见的池化方法有 最大池化(Max Pooling)和平均池化(Average Pooling),其中,最大池化利用统计 局部感受野内的最大值作为池化层输出,平均池化采用计算感受野中所有值的平均值 作为特征图经过池化层后得到的特征图。图 2-2 为最大池化操作的示例。可以看出, 在深度保持不变时,空间上特征映射的神经元数量减少为原来的四分之一,而且在局 部感受野内特征发生轻微位移不会影响池化的结果,使得网络对特征微小的位移有平 移不变性。



图 2-2 最大池化示例

在实际应用中,池化层通常是将每个特征映射划分为尺寸大小为2×2的不重叠区 域,通常使用最大池化进行下采样,有效降低网络参数数量。此外,设置过大的采样 区域会大幅减少特征映射中神经元的数量,在经过一个池化层后就造成严重的信息丢 失,因此通常在池化层中使用大小为2×2的窗口。

2.1.2 经典的卷积神经网络结构

2.1.2.1 LeNet-5

LeNet-5^[46]是由科学家 Yann LeCun 于 1994 年提出的第一个神经网络模型,四年后,Yann LeCun 在 LeNet-5 神经网络模型上采用反向传播算法实现监督训练。LeNet-5 的网络结构如图 2–3 所示。



图 2-3 LeNet-5 网络结构^[46]

由图可知,不计输入层,LeNet-5 共有7层,其每一层的结构信息如下:

(1) C1 层: C1 表示卷积层,经过 6 组 5×5 的卷积核进行卷积处理后,输出 6 组 特征图,其空间尺寸为 28×28。

(2) S2 层: S2 表示池化层,使用滑动窗口对特征图进行平均池化操作,其窗口 尺寸为2×2,输出6组空间维度减半的特征图,其空间尺寸为14×14。

(3) C3 层: C3 表示卷积层,经过 16 组 5×5 的卷积核进行卷积处理后,输出 16 组特征图,其空间尺寸为10×10。

(4) S4 层: S4 表示池化层,使用滑动窗口对特征图进行平均池化操作,其窗口 尺寸为2×2,输出 16 组空间维度减半的特征图,其空间尺寸为5×5。

(5) C5 层: C5 表示卷积层,使用120×16=1920个5×5的卷积核,得到120 组大小为1×1的特征图。

(6) F6 层: F6 表示全连接层,该层有 84 个神经元,与(5)中输出的特征图进行全连接。

(7) 输出层:输出层具有 10 个神经元,与(6)中 120 个神经元全连接,得到 10 个输出结果代表 10 个分类的类别。

2.1.2.2 AlexNet

LeNet-5 提出时,受限于当时的数据不足与硬件水平,LeNet-5 并没有引起太大的关注。2012年,Alex 等人^[47]提出首个现代深度卷积网络模型 AlexNet(图 2-4),当时以远低于第二名模型分类错误率的成绩获得 ImageNet 图像分类竞赛的冠军。在AlexNet 网络模型上采取的训练神经网络的技术方法也被沿用至今,通过使用两台GPU 进行并行计算,大幅提升网络模型的训练速度,此外采用 ReLU 激活函数将非线性和稀疏性加入到网络中,并使用 Dropout 方法以特定概率丢弃网络中的神经元,防止过拟合,最后通过数据增广来增大数据集,从而提高模型准确率。



图 2-4 AlexNet 网络结构^[47]

由图 2-4 可知, AlexNet 包括 5 个卷积层、3 个全连接层和 1 个 softmax 层。AlexNet 每一层结构信息如下:

(1) 卷积层 C1: 使用两组11×11×3×48的卷积核,以步长 4,不填充边缘的方法 得到两个55×55×48的特征图。

(2) 池化层 S1:使用大小为3×3、步长为2的最大池化操作,得到两个27×27×48 的特征图。

(3) 卷积层 C2: 使用两组 5×5×48×128 的卷积核,以步长 1,边缘填充为 2 的方 法得到两个 27×27×128 的特征图。

(4) 池化层 S2:使用大小为3×3、步长为2的最大池化操作,得到两个13×13×128的特征图。

(5) 卷积层 C3: 使用一组3×3×256×384的卷积核,以步长1,边缘填充为1的 方法得到一个13×13×384的特征图。

(6) 卷积层 C4: 使用两组 3×3×192×192 的卷积核,以步长 1,边缘填充为 1 的 方法得到两个13×13×192 的特征图。

(7) 卷积层 C5: 使用两组 3×3×192×128 的卷积核,以步长 1,边缘填充为 1 的 方法得到一个13×13×128 的特征图。

(8) 池化层 S3: 使用大小为 3×3、步长为 2 的最大池化操作,得到两个 6×6×128 的特征图。

(9) 三个全连接层: F6、F7、F8 的神经元个数分别是 4096、4096、1000, 最终 得到 1000 个分类结果。

2.1.2.3 Inception

在卷积网络中,卷积核大小的选择对神经网络模型的训练有重要影响。在 Inception 网络中,通过采用大小为1×1、3×3、5×5的卷积核拓宽网络结构,称为 Inception 模块,采用不同大小的卷积核使得网络具有不同尺度大小的感受野,经过卷 积处理后将特征映射相加能够表达更丰富的输入数据的内容。图 2–5 给出了 Inception v1^[48]模块,采用4组平行且大小不同的特征描述子来提取特征,分别为1×1、3×3、 5×5的卷积以及3×3的最大池化。Inception 模块在进行3×3、5×5的卷积之前、3×3 的最大池化之后,进行一次1×1的卷积来减少特征映射的深度,以此达到减少参数数 量的目的,提高计算效率。





图 2-5 Inception v1 模块结构^[48]

Inception 网络经过几年来的发展,衍生出了多个提升的 Inception 网络版本,其 中影响力较大的有 Inception v2^[49]网络模型和 Inception v3 网络模型^[49,50]。Inception v2 网模型络在 Inception v1 的基础上主要有两点改进:(1) Inception v1 中的卷积核大小 为 5×5 的卷积被两层卷积核大小为 3×3 的卷积代替,这种采用多层连续小卷积核替 换大卷积核的方法,其主要目的是在保持卷积核对图像感受野不变的情况下还能减少 计算量,缩减网络模型参数数量;(2)使用连续的卷积核大小为 n×1 和 1×n 的卷积 层来替换卷积核大小 n×n 的卷积层。另外,Inception v3^[50]在 Inception v2 网络的基础 上引入了标签平滑以及批量标准化(Batch Normalization, BN)等优化方法进行训练。 最后,Szegedy 等人还提出了结合残差模块的深度网络模型 Inception v4^[51],该网络模 型以更深的网络结构进一步取得更优的模型分类精度。

2.1.2.4 残差网络

近十年来,在有关深度神经网络模型的实践中得知,增加神经网络模型复杂度的两种方式是通过增加网络的深度或者网络的宽度,而研究者们发现增加网络的深度相对于增加宽度更容易提升模型的性能。然而,随着神经网络深度的增加,给网络的训练会带来更大的挑战。其主要原因是在网络训练过程中,误差信号经过多层反向传播后,回传的梯度过小或者过大从而产生梯度消失或者梯度爆炸。He等人^[52]提出使用权重初始化的方法以及批标准化方法可以有效解决这个问题。但是新的问题出现了,当深度神经网络收敛时,如果想要继续降低训练数据的训练误差,根据增加网络深度有利于提升网络模型性能这一特点,应当增加网络深度,但是训练误差并没有降低反而升高,而这不能被解释为过拟合,后来被称之网络退化,这个问题一度困扰着研究者们设计更深层卷积神经网络及其应用。ResNet^[52]的提出使得上述问题得到解决。

ResNet 主要由残差模块构成,如图 2-6(a)所示,在残差模块中,假设目标函数为H(x),我们可以将目标函数拆为恒等函数x与残差函数F(x)=H(x)-x。这两个表达式具有相同的效果,但优化的难度不同,比如学习恒等映射函数H(x)=x,在残差

函数中只需要学习 F(x) = 0,因为每层网络参数的初始化值一般偏向为 0,相比于更 新该层的参数来学习 H(x) = x,学习 F(x) = 0 能更快的收敛,实际训练中使用残差函 数更容易学习。



图 2-6 残差模块[52]

图 2-6(b)为上面提到的常规残差模块,由两个3×3卷积堆叠而成,但是随着 网络的深度越来越深,这种残差模块在模型训练过程中表现出一定的局限性。图 2-6 (c)为瓶颈残差模块(Bottleneck Residual Module),这种残差模块被广泛应用在深 度网络模型中,该残差模块依次由1×1,3×3和1×1的卷积层构成,为了提高网络模 型的计算效率,采用1×1卷积对特征图进行降通道,使得计算复杂的3×3的卷积在相 对较低维度的输入特征图上进行,最后通过1×1卷积升通道数至目标维数。在这里, 首先用1×1卷积把输入特征降到 64 维,然后在 64 维上使用3×3的卷积,最后通过1×1 卷积恢复特征维度到 256 维。

2.2 全卷积神经网络

全卷积神经网络(Fully Convolutional Neural Network, FCN)由 Long 等人^[42]提出, 由于 FCN 可以对图像进行像素级的分类,从而作为一种框架被广泛用于语义分割 (Semantic Segmentation)问题。通过使用卷积层来替换神经网络中所有的全连接层, 其目的是使得 FCN 能够获取任意尺寸大小的输入图像,从而不需要采用图像补丁的 方式把数据输入网络,这样减少了图像补丁之间由于重叠造成的冗余计算,提高网络 分割效率。另外,FCN 通过反卷积对特征图进行上采样,使特征图恢复到输入图像 相同的分辨率,并且通过反卷积计算后的图像保证了特征所在区域的权重。FCN 框 架如图 2-7 所示。



图 2-7 全卷积网络框架[42]

2.3 基于编码器解码器的神经网络

2.3.1 SegNet

目前常用的分割网络大多都是基于编码器解码器架构的神经网络,2015年,剑桥大学团队提出了 SegNet^[53]分割网络模型(图 2-8),该网络模型主要应用在自动驾驶领域,用于解决图像语义分割问题。



图 2-8 SegNet 网络结构

SegNet 和 FCN 思路十分相似,只是在编码器与解码器中采用了不同的技术。在 编码器部分 SegNet 采用的是与 VGG16^[54]的前 13 层卷积具有相同拓扑结构的卷积网 络,去除了最后三层全连接层。每个编码器层都对应着一个解码器层,在经过最后一 个解码器后,采用 Softmax 分类器对输出的特征图进行分类,得到图像每个像素的分 类概率。对比 FCN 可以发现, SegNet 在上采样时用下采样保存的最大池化位置索引 信息,直接将数据放回对应位置,后面再接卷积层训练学习。这个上采样过程不需要 训练学习,只带来一些存储空间上的消耗,这个操作在分割中可以保留图像中更多的 高频信息,但是对于低分辨率的特征图,这个操作会丢失索引位置邻近的像素信息。

2.3.2 U-Net

虽然 FCN 解决了网络输入特征尺寸固定的约束,但是在图像分割任务中还存在 一个问题:经过下采样后特征位置信息丢失。我们知道,在经过池化层后,输入的特 征图尺寸会一层层的减小,在缩小的特征图像进行特征提取能够有更大的感受野,但 同时带来的是特征空间位置信息的缺失,由于特征图尺寸较小,虽然能够表达更高层 级的特征,但是难以表达该特征信息在输入图像的原位置信息,这对于医学图像分割 任务而言是非常不利的。针对这个问题,在 FCN 的基础上,Ronneberger 等人^[41]提出 了基于编码器解码器架构(Encoder-Decoder Architecture)的卷积神经网络 U-Net,如 图 2–9 所示。



图 2-9 U-Net 网络框架

由图可知,U-Net 网络主要分为三部分:编码器、解码器、跳跃连接。在编码器 部分通过卷积与池化提取特征,在解码器阶段通过上采样逐步恢复特征图尺寸为输入 图像相同的尺寸,前提是在卷积层采用 padding 为 same 的设置。这里和 SegNet 不同 的是没有采用池化索引做特征映射,而是直接通过跳跃连接(Skip Connection)将特 征位置信息更丰富的各编码器输出特征图作为各解码器输入,并通过 Concate 方式融 合浅层特征信息与深层特征信息,在编码器池化中丢失的相关特征信息可以通过这个 操作补充到解码器中,从而可以帮助神经网络恢复特征的位置信息,达到更好的分割 效果。

2.4 注意力机制

注意力机制(Attention Mechanisms)^[55]的研究目前在深度学习中是一个重要的 方向。受到人类视觉系统的启发,注意力(Attention)这一概念被提出来。当人类用 视觉系统观察事物时,最先注意到的往往不是这个事物的整体,而是倾向于根据人类 自身需要选择性的去关注被观察事物的某些重要部分,比如当我们看一个人的时候, 第一个关注的地方往往是这个人的脸,因为分辨人类最明显的特征就是人脸,然后再 把其他不同区域的信息组合起来,例如身高、胖瘦、着装打扮等,从而形成一个对被 观察人类的整体印象。同理,在计算机领域中,注意力机制往往通过对输入的每个部 分设置不同的权重,使模型着重关注重要的特征信息,从而提高模型判别的准确率, 同时模型不会产生更大的计算和存储开销,这也是注意力机制应用如此广泛的原因。

2.4.1 注意力机制的分类

我们知道,注意力机制可以通过对输入的每个部分设置不同的权重来表示我们对 各个部分的关注度,例如在一幅输入的图像中,对存在重要特征信息的部分设置比较 大的权重,对不需要关注的部分赋予较小的权重。按照可微性来划分,注意力机制可 分为硬注意力(Hard Attention)与软注意力(Soft Attention)。硬注意力就是二分类问 题,其实就是哪些区域需要被关注,而哪些区域不需要关注,最经典的应用就是图像 裁剪(Image Crop),只保留关注的部分,而不需要关注的部分都被裁减掉。软注意 力与硬注意力不同的是,软注意力是表示每个区域被关注的程度高低问题,一个非常 重要的地方是软注意力是可微的,注意力可微分就说明可以利用神经网络的反向传播 计算梯度,从而来学习注意力权重。

2.4.2 常见的注意力模型

2.4.2.1 Attention

Attention 早期的模型主要有 SENet^[56]和 CBAM^[57],二者通常都是在空间或通道 上计算注意力分布。

由 Hu 等人^[56]提出的 SENet (图 2–10),获得了 ImageNet 2017 图像分类赛冠军。 SENet 是早期注意力机制,其本质就是学习通道特征间的关系,产生通道间注意力分 布,进而提高模型表现。这个注意力机制分成三个部分:挤压(Squeeze)、激励 (Excitation)、变换(Scale)。挤压就是采用全局平均池化对输入特征图进行处理,得到 输出的通道数和输入的特征通道数相同的特征向量,该操作将输入特征的尺寸从 *h×w×c*变为1×1×*c*;激励就是采用包含两个全连接层的 bottleneck 结构,其中第一个 FC 层先压缩通道数,第二个的 FC 层恢复原始的维度,最后采用 *Sigmoid* 激活函数得 到 0-1 之间的注意力权重。变换的的操作就是将激励输出的注意力权值通过逐通道相 乘到原始输入的特征上,从而在通道维度上的对输入特征进行重新标记。



图 2-10 Squeeze-and-Excitation 模块[56]

Woo 等人^[57]提出了 CBAM(Convolutional Block Attention Module)(图 2–11)用 于提升 CNN 的表现。CBAM 采用通道注意力与空间注意力串联的方式来使网络模型 知道分别在通道和空间上关注什么,以及在哪里关注。在通道注意力上同时采用最大 池化与平均池化得到两个不同描述的特征图,然后将一个隐藏层的多层感知器(MLP) 应用于两个通道,最后通过 Sigmoid 激活函数获得通道注意力权重图。空间注意力首 先在通道上对输入特征进行降维,并利用最大池化与平均池化获得两个 h×w×1的特 征描述,然后将这两个特征图经过 Concate 操作拼接到一起,经过一个7×7的卷积操 作和 Sigmoid 激活函数后,可以得到空间注意力权重图。



图 2-11 Convolutional Block Attention Module^[57]

2.4.2.1 Self Attention



图 2-12 Self Attention 基本结构[58]

Self Attention^[58]的基本结构(图 2–12)有三个分支:Q(Query)、K(Key)和V(Value)。计算时通常分为三步:

- (1) 将Q和每个K进行相似度计算得到权重;
- (2) 一般是使用一个 softmax 函数对这些权重进行归一化;
- (3) 将权重和相应的键值 V 进行加权求和得到最后的注意力权重图。

Wang 等人^[59]提出了 Non-Local Neural Network (图 2–13),该模型利用 Self Attention 的思想,使图像中每一点的输出受所有像素点的影响,而传统的 CNN 一个 像素点对应的输出只受其邻域的影响。模型结构简单,效果提升显著,且可以方便的 嵌入到现有网络中,对于在空间域的图像识别准确率有提升。



图 2-13 Non-local 模块^[59]

Fu 等人^[60]提出了 DANet (Dual Attention Network)(图 2–14)用于场景分割,该 模型和 CBAM^[57]相似,只是在通道域和空间域结合 Self Attention 的思想建立全局上 下文关系,改善分割效果明显。



21

第3章 基于自校准卷积的视网膜血管分割算法

3.1 整体网络结构设计

本文针对视网膜血管分割任务,在 U-Net 网络的基础上提出了一种端到端的视网 膜血管分割网络 SCCU-Net,所提出的网络可以在不经过任何图像预处理的情况下实 现对输入的源数据自动提取特征,网络框架如图 3-1 所示,该网络主要由编码部分、 空间注意力模块、解码部分构成。

在编码部分里,我们使用自校准卷积替代原来的卷积层,并且每个自校准卷积后 都连接了一个 BN 和 ReLU,相较于传统的卷积层,引用自校准卷积的网络能够提取 更加丰富的特征信息。空间注意力模块主要用来获取更加完整的空间特征信息,提升 网络模型的分割表现。解码部分主要通过一个转置卷积来进行上采样,然后与经过跳 连操作(Skip Connection)过来的特征进行特征融合,最后,和编码部分里的自校准卷 积一样,我们使用自校准卷积替代原来的卷积层,以获得更加丰富的解码特征信息。 另外需要指出的是,由于眼底血管数据的数量少,即使做了一些数据增广,训练一个 比较深的网络也是非常困难的,因此我们减少的下采样的次数,由4次变为3次,这 样能有效减少网络的深度,也可以减少网络模型的参数量,加快网络训练。



3.2 自校准卷积算法

卷积操作是提取神经网络中特征信息的常见操作,但是由于统一的卷积核尺寸 (例如使用3×3卷积核),这使得在一个卷积操作中提取不同尺度的特征信息变得困 难。因此在文献^[61]中,Liu等人提出了一种自校准卷积的操作,所提出的自校准卷积 将卷积滤波器分割成多个部分,然而和分组卷积^[62]不同的是,滤波器的每个部分都不 是相同的,而是负责一个特殊的功能。

3.2.1 自校准卷积的结构

图 3-2 显示的是自校准卷积的网络结构图。由于输入的通道数和输出的通道数相同,为了方便描述,输入特征与输出特征的通道数都记为 C。首先定义 4 个卷积操作 {*Conv1;Conv2;Conv3;Conv4*},其中每个卷积操作的卷积滤波器为(*C*/2,*C*/2, k_h , k_w), k_h 和 k_w 分别表示卷积核的高度和宽度,在本文中,我们设定 $k_h = k_w = 3$,这样就得到了四 个不同功能的卷积滤波器。



图 3-2 自校准卷积模块

给定了卷积滤波器的四个部分,每个部分都具有不同的功能。然后我们将输入特征 F 均匀地分成两个部分 { F_1 , F_2 },然后每个部分分别送入到一个特殊的路径中,用于 关注不同类型的上下文特征信息。在第一个路径中,利用 { $Conv_1$; $Conv_2$; $Conv_3$ }对 F_1 执行自校准操作,得到 F_1' 。在第二个路径中,我们执行一个简单的卷积运算: $F_2' = Conv4(F_2) = f_4(F_2)$,其作用是保持原始的空间上下文特征信息。在得到两个中间输出 { F_1',F_2' }后,经过 Concate 操作将 { F_1',F_2' }连接在一起作为输出 F',输出 F'可由公式 3–1 表示。在 3.2.2 节中我们详细描述如何在第一个路径中执行自校准操作。

$$F' = Concate \{F'_1; F'_2\}$$
(3-1)

3.2.2 自校准卷积操作

为了高效有效地为每个空间位置收集信息丰富的上下文信息,从而在两个不同的 尺度空间进行卷积特征提取:一个是原始的尺度空间,其中特征图与输入特征具有相 同的尺寸,另一个是经过下采样后得到的尺度空间。由于经过下采样后特征图对应的 卷积滤波器具有较大的感受野,因此可以作为参考来指导原始特征空间中的特征提取 过程。自校准卷积操作步骤如下:

- (1) 对输入的原始特征图 $F_1 \in R^{H \times W \times C/2}$ 进行一次下采样,特征图尺寸从 $H \times W \times C/2$ 变为 $H/2 \times W/2 \times C/2$,操作定义为Down(*)。
- (2) 对下采样后的特征图进行卷积,该操作记为 f₁(*)。
- (3) 对(2)中卷积后的特征图进行一次上采样操作,特征图尺寸从*H*/2×*W*/2×*C*/2 恢复为*H*×*W*×*C*/2,该操作记为*Up*(*)。
- (4) 将输入的原特征图 $F_1 \in R^{H \times W \times C/2}$ 与(3)中得到的特征图相加,再经过激活函数 Sigmoid,得到尺度为 $H \times W \times C/2$ 的特征图,该操作定义为 $\sigma(*)$ 。
- (5) 对原始特征图 $F_1 \in \mathbb{R}^{H \times W \times C/2}$ 进行卷积操作,该操作记为 $f_2(*)$ 。
- (6) 将(4)中得到的特征图与(5)中得到的特征图相乘,该操作记为(*)⊗(*)
- (7) 对(6)中得到的特征图进行卷积操作,得到输出 $F'_1 \in R^{H \times W \times C/2}$,该操作记为 $f_3(*)$ 。 综合以上内容,自校准卷积操作的输出 $F'_1 \in R^{H \times W \times C/2}$ 可表示为:

$$F_{1}' = f_{3} \left[\sigma \left(Up \left(f_{1} \left(Down \left(F_{1} \right) \right) \right) \oplus F_{1} \right) \otimes f_{2} \left(F_{1} \right) \right]$$
(3-2)

3.3 空间注意力机制的设计

受到 CBAM^[57]利用通道注意力模块和空间注意力模块用于分类和检测的启发。 我们利用空间注意模块生成空间注意图,在空间注意模块的基础上,设计了改进的空 间注意模块。图 3-3显示了原始空间注意模块(a)和我们改进后的空间注意模块(b)。 我们知道,经过卷积运算后,部分特征图的空间信息将丢失。为了获得更完整的空间 信息,我们在改进的空间注意模块中在执行卷积操作前生成空间注意力图。这样可以 利用更完整的空间特征图信息来提高分割性能。



图 3-3 空间注意力模块

图 3 中 b 图为改进的空间注意力模块,给定输入特征图 $F_s \in R^{H \times W \times C}$,其主要操作步骤如下:

(1) 对输入原始特征分别进行最大池化操作和平均池化操作,得到特征图 $F_{MP} \in R^{H \times W \times 1}$ 和 $F_{AP} \in R^{H \times W \times 1}$ 。

(2) 对(1)中得到的特征图执行 Concate 操作融合特征图,得到尺寸为H×W×2的特征图。然后将得到的特征图用尺寸为7×7卷积滤波器进行卷积操作,该操作记为 f^{7×7}(*)。

(3) 将(2)中得到的特征图经过激活函数 Sigmoid ,得到尺度为 $H \times W \times 1$ 的空间注意力图 M_s ,该操作记为 $\sigma(*)$ 。

(4) 对输入原始特征进行自校准卷积操作,该操作记为SCC(*)。

(5) 最后将(3)中得到的空间注意力权重与经过两次自校准卷积的特征图相乘,得到输出特征图 $F'_{s} \in R^{H \times W \times 2C}$ 。

综合以上内容,空间注意力模块的输出F's可表示为:

$$F'_{S} = SCC(SCC(F_{S})) \otimes \sigma(f^{7 \times 7}(F_{MP}; F_{AP}))$$
(3-3)

3.4 损失函数

我们的任务是找出眼底图像的血管和非血管部分,因此视网膜血管分割的任务可 以看作是一个二分类问题。交叉熵损失函数被广泛应用于深度学习网络中,用于处理 图像的二分类问题,计算预测某一类或非某一类的概率。在下面的介绍中, t_i表示金标准图像第i个像素对应的标签,用1表示血管像素,0表示背景像素。网络的输出为图像 y_i,表示第i个像素被分类为1的概率。损失函数定义如公式 3-4 所示:

$$Loss = -\sum_{i} \left[t_{i} \log y_{i} + (1 - t_{i}) \log (1 - y_{i}) \right]$$
(3-4)

第4章 实验结果与分析

4.1 数据集概述

为了更公平的对比本文所提出的视网膜血管分割算法的效果,本文选取了三个公 开发表的眼底图像数据集来验证我们所提算法的有效性,他们分别是 DRIVE 数据集、 STARE 数据集以及 CHASE DB1 数据集,这些数据集中均包含有专科医生手动标注 的血管分割金标准。关于 DRIVE 数据集、STARE 数据集以及 CHASE DB1 数据集的 具体信息如下表 4-1 所示。

数据集	DRIVE	STARE	CHASE DB1
图像总数	40	20	20
训练集数/测试集数	20/20	19/1	20/8
训练集数/验证集数	234/26	162/9	247/13
原图像大小	565×584	700×605	999×960
调整图像后大小	592×592	704×704	1008×1008
数据增强方法	(1)(2)(3)(4)(5)	(2)(3)(4)	(1)(2)(3)(4)(5)

表 4-1 本文选用的彩色视网膜图像公共数据库介绍

其中 DRIVE 数据集由 40 张彩色眼底照片组成,这些照片是从一个糖尿病视网膜 病变筛查项目中获得的,其中 20 个样本用于训练,其余 20 个样本用于测试。每幅眼 底图像的大小为 565×584 像素,并配有由专科医生标注的视网膜血管分割金标准。 STARE 数据集由 20 张视网膜图像组成,每张图像的大小为 700×605 像素,这份数据 集的一半图像采样自患病患者的眼底图像,另一半是健康者的眼底图像,需要指出的 是,由于 STARE 数据集样本较少,采用留一法(Leave One Out)的训练测试方法, 即每次将 19 张眼底图像用于训练,余下一张图像用于测试,如此单独重复 20 次实验。 CHASE DB1 数据集包含 28 张视网膜图像,每张图像大小为 999×960 像素。本数据 集中的图像采集自 14 名儿童的左眼和右眼。数据集分为两组,前 20 个样本用于训练, 其余 8 个样本用于测试。

在将眼底图像送入到分割网络之前,我们需要调整输入图像到网络合适的大小。 对于 DRIVE 数据集,我们将其大小调整为 592×592;对于 STARE 数据集,我们将其 大小调整为 704×704;对于 CHASE 数据集,我们将它们的大小调整为 1008×1008。 这三个数据集的训练集都采用在原图像边界到新图像边界的四个边距内填充 0 的方 法。虽然在训练时对图像大小做了调整,但在评估期间将尺寸裁剪为原始图像尺寸。

汕头大学硕士学位论文

另外为了增强数据,我们对眼底图像数据采用了以下五种数据增强方法:(1)随机旋转;(2)添加高斯噪声;(3)颜色增强(利用图像亮度,饱和度,对比度变化来增加数据量);(4)水平、垂直翻转;(5)对角线翻转。其中需要补充的是对于 STARE 数据 集的数据增强没有用到方法(1)(5),这是由于 STARE 数据集眼底图像的长宽差距 过大,如果进行随机旋转或对角线翻转操作,数据将会严重失真,这反而会影响模型 的准确率。图 4-1 展示了一组经过数据增强的眼底图像的样例。



图 4-1 眼底图像数据增强样例: (a)视网膜眼底图像, (b) 经过随机旋转后的增强图像, (c) 经过添加高斯噪声后的增强图像, (d) 经过随机颜色增强后的眼底图像, (e) 经过水平翻转的 眼底图像, (f) 经过垂直翻转的眼底图像, (g) 经过对角线翻转的眼底图像

4.2 评价指标

考虑到视网膜血管分割的任务就是对眼底图像的像素点进行二分类,正类为血管 像素点,负类为非血管像素点。且大多数视网膜血管分割的文献中采用的也是二分类 的评价方法。因此本文使用几个针对二分类问题的常见评价指标来对视网膜血管分割 网络模型的表现进行评估。例如,准确率(Accuracy, ACC)、敏感度(Sensitivity, SE)、 特异性(Specificity, SP)、F1分数(F1-score)、模型评价曲线下面积(Area Under Curve, AUC)。

将视网膜血管分割算法得到的分割结果与金标准相比较可得到下表 4-2,表示的 是用于二分类问题的混淆矩阵 (Confusion Matrix)。

汕头大学硕士学位论文

表 4-2 二分类问题的混淆矩阵

	金标准正类	金标准负类
算法预测正类	真阳性(True Positive, TP)	假阳性(False Positive, FP)
算法预测负类	假阴性(False Negative, FN)	真阴性 (True Negative, TN)

根据得到的混淆矩阵,可以计算不同评价指标。准确率(Accuracy, ACC)是指 被正确分割的像素与总像素的比例,如公式 4-1 所示:

$$ACC = \frac{TP + TN}{TP + FN + TN + FP}$$
(4-1)

敏感度(Sensitivity, SE)表示分割正确的正类像素点与所有被分类成正类的像素点的比例,如公式 4-2 所示:

$$SE = \frac{TP}{TP + FN} \tag{4-2}$$

特异性(Specificity, SP)表示分割正确的负类像素点与所有被分类成负类的像素点的比例,如公式 4-3 所示:

$$SP = \frac{TN}{TN + FP} \tag{4-3}$$

F1 分数(F1-score)表示精确率(Precision,公式 4-4)与召回率(Recall,公式 4-5)的调和平均数,如公式 4-6 所示:

$$Precision = \frac{TP}{TP + FP}$$
(4-4)

$$Recall = \frac{TP}{TP + FN}$$
(4-5)

$$Fl = 2*\frac{Precision*Recall}{Precision+Recall} = \frac{2*TP}{2*TP+FP+FN}$$
(4-6)

我们可以采用不同的颜色来标注分类的不同结果,从而可以将得到的分割结果可 视化。具体的,用绿色像素点表示正确分割的像素点(金标准为正类,算法预测正类), 即 TP 真阳性。红色像素点表示将背景像素点误判为血管像素点的像素点(金标准为 负类,算法预测为正类),即假阳性 FP。蓝色像素点为漏判的血管像素点(金标准为 正类,算法预测为负类),即假阴性 FN。黑色像素点表示分类正确的背景像素点(金 标准为负类,算法预测为负类),即真阴性 TN。下图 4-2 展示的是分割结果可视化样 例,图中可以看出该算法存在局部区域将背景像素点预测为血管像素点的部分,同时 蓝色像素点也表明该算法丢失了一部分细小血管像素点。 汕头大学硕士学位论文



图 4-2 视网膜血管分割结果可视化样例: (a)视网膜眼底图像, (b)金标准图像, (c)算法 预测概率图像, (d)算法分割结果可视化图像

4.3 实验设计

4.3.1 实验环境

表 4-3 实验硬件平台及软件环境

类别	环境
CPU	Intel Xeon E5-2680 v4 @2.4GHz, 56 核心
GPU	NVIDIA TITAN Xp(12G) × 4
操作系统	Linux Ubuntu 18.04
编程语言	Python 3.6
框架	Keras 2.3.1、NumPy 1.15.4 等

本文实验所需的硬件平台及软件环境如表 4-3 所示,需要指出的是,实验的训练 是在 4 块 NVIDIA TITAN Xp GPU 搭建成的 GPU 集群上完成的,每块 GPU 的显存为 12G,使用 GPU 集群可以加速神经网络模型的训练,从而大大减少网络模型的训练 时间。

4.3.2 实验参数设置

对于训练参数,本文采用随机梯度下降算法(SGD)^[63]来最小化损失函数,并将

学习率设置为 0.001, 模型训练的最大迭代数设置为 40, 批量训练数设置为 4。

本实验网络主要参数设置如下表 4-4 所示,其中表达式 H×W×C 表示输入或输出 特征图的尺寸,H 为特征图的高度,W 为特征图的宽度,C 为特征图的通道数。s 表 示步长。

模块名称	输入尺寸(H×W×C) DRIVE/STARE/CHAS E DB1	输出尺寸(H×W×C) DRIVE/STARE/CHAS E DB1	网络结构
Encoder_ layer1	592×592×3 /704×704×3 /1008×1008×3	296×296×16 /352×352×16 /504×504×16	$\begin{bmatrix} SCC_block\\ BN+ReLU \end{bmatrix} \times 2$ Max - Pooling, s = 2
Encoder_ layer2	296×296×16 /352×352×16 /504×504×16	148×148×32 /176×176×32 /252×252×32	$\begin{bmatrix} SCC_block \\ BN+ReLU \end{bmatrix} \times 2$ Max - Pooling, s = 2
Encoder_ layer3	148×148×32 /176×176×32 /252×252×32	74×74×64 /88×88×64 /126×126×64	$\begin{bmatrix} SCC_block \\ BN+ReLU \end{bmatrix} \times 2$ Max - Pooling, s = 2
ISAM	74×74×64 /88×88×64 /126×126×64	74×74×128 /88×88×128 /126×126×128	ISAM
Decoder_ layer1	74×74×128 /88×88×128 /126×126×128	148×148×64 /176×176×64 /252×252×64	Transposed_Conv, s = 2 $\begin{bmatrix} SCC_block\\ BN + ReLu \end{bmatrix} \times 2$
Decoder_ layer2	148×148×64 /176×176×64 /252×252×64	296×296×32 /352×352×32 /504×504×32	Transposed_Conv, s = 2 $\begin{bmatrix} SCC_block\\ BN + ReLu \end{bmatrix} \times 2$
Decoder_ layer3	296×296×32 /352×352×32 /504×504×32	592×592×16 /704×704×16 /1008×1008×16	Transposed_Conv, s = 2 $\begin{bmatrix} SCC_block\\ BN + ReLu \end{bmatrix} \times 2$
Conv_lay er1	592×592×16 /704×704×16 /1008×1008×16	592×592×1 /704×704×1 /1008×1008×1	$Conv 3 \times 3, s = 1$ Sigmoid

表 4-4 网络主要参数设置

4.3.3 实验设计

为了证明本文所采用的各个方法的有效性,本文设计了如下四个对比实验,并分别在 DRIVE 数据集和 CHASE DB1 数据集上进行实验,对于 STARE 数据集,由于采用的是留一法进行训练测试,所需时间代价非常大,因此在 STARE 数据集上只验证

本文所提方法的有效性,即只完成实验五。下表 4-5 为相应对比实验详细信息。

- (1) 实验一(U-Net): 采用 U-Net 网络作为实验基本网络框架;
- (2) 实验二(U-Net+SCC): 在 U-Net 网络的基础上引入自校准卷积模块 (Self-Calibrated Convolution),使用自校准卷积替换 U-Net 网络的原卷积操作;
- (3) 实验三(U-Net+SAM): 在 U-Net 网络的基础上引入空间注意力模块(Spatial Attention Module);
- (4) 实验四(U-Net+ISAM): 在 U-Net 网络的基础上引入改进的空间注意力模块 (Improved Spatial Attention Module);
- (5) 实验五(U-Net+SCC+SAM): 在 U-Net 网络的基础上引入自校准卷积模块 (Self-Calibrated Convolution)和改进的空间注意力模块(Improved Spatial Attention Module)。

实验/模块	U-Net	SSC	SAM	ISAM	模型大小 (MB)
实验一					6.38
实验二	\checkmark	\checkmark			8.68
实验三	\checkmark		\checkmark		6.39
实验四	\checkmark			\checkmark	6.39
实验五				\checkmark	8.69

表 4-5 对比实验详细信息

4.4 结果与分析

本文在这一小节展示实验结果以及对相应实验结果进行分析。主要包括两部分内 容:一是通过五个实验的比较来验证所采用方法的有效性,并证明所添加的模块能够 提升视网膜血管分割的各项评价指标;二是展示本文所提方法更加详细的分割结果; 三是通过我们所提出的视网膜血管分割算法与先进的视网膜血管分割算法对比,比较 结果说明本文所提的算法在视网膜血管分割任务上具有良好的表现。

4.4.1 对比实验结果与分析

从第三章所介绍的方法可以得知: 自校准卷积操作可以利用下采样获取更大尺度 感受野,从而输出更加丰富的语义信息; 改进的空间注意力机制可以利用平均池化与 最大池化来生成空间注意力权值图,并且该空间注意力权值图是在卷积操作之前完成 的,从而使空间注意力权值图中保留了更加完整的特征图信息,避免了由于提前卷积 带来的部分特征丢失。因此,本文设计了五个对比实验来证明所采用模块的有效性。 通过实验二与实验一比较,可以验证自校准卷积在视网膜血管分割任务中的有效性; 通过实验四与实验三比较,可以证明改进的空间注意力机制可以提升视网膜血管分割 的表现;通过实验五与前四个实验比较,可以证明本文所提出方法在视网膜血管分割 中能够取得良好的分割结果。

数据集	实验	ACC	SE	SP	F 1	AUC
	实验一	0.9646	0.7895	0.9814	0.7963	0.9791
	实验二	0.9665	0.8067	0.9818	0.8084	0.9810
DRIVE	实验三	0.9659	0.7795	0.9838	0.8001	0.9780
	实验四	0.9663	0.8066	0.9820	0.8089	0.9815
	实验五	0.9680	0.8036	0.9840	0.8138	0.9840
	实验一	0.9733	0.7817	0.9862	0.7870	0.9810
	实验二	0.9744	0.8158	0.9851	0.8006	0.9872
CHASE DB1	实验三	0.9740	0.7933	0.9861	0.7934	0.9859
DDI	实验四	0.9751	0.8052	0.9865	0.8029	0.9843
	实验五	0.9756	0.8118	0.9867	0.8068	0.9888

表 4-6 在 DRIVE 数据集和 CHASE DB1 上的对比实验结果

从表 4-6 展示的对比实验在 DRIVE 数据集和 CHASE DB1 数据集上的实验结果 我们可以发现,在两个数据集上,采用添加 SCC 或者 ISAM 后的网络模型所有的评 价指标都比 U-Net 网络模型所得的分割结果高。特别是对于综合指标调和平均数 F1 分数与模型评估曲线下面积 AUC, 全部超过了 U-Net 所得到的分割结果, 即采用自 校准卷积模块和改进的空间注意力模块对视网膜血管分割结果的提升是有效的。仔细 观察整个表 4--6, 可以发现实验五所取得的视网膜血管分割结果除了少数指标略低于 前面的实验结果,而其他指标近乎能取得最优值。在 DRIVE 数据集中,实验二在特 异性不输实验一的前提下,灵敏度 SE 与 F1 分数分别提高了 1.72%与 1.21%,并且其 他两个评价指标也是高于实验一的。实验三与实验一相比,指标灵敏度 SE 与模型评 估曲线下面积 AUC 略低于实验一,其他三个指标高于实验一。除特异性 SP,实验四 其他所有的评价指标都优于实验三,另外,实验四整体结果都优于实验一。实验五所 得的灵敏度 SE 比实验二略低,但在准确率 ACC、SP、F1 分数以及 AUC 上是所有实 验中是最高的。在 CHASE DB1 数据集上的实验结果可以得出:对比实验一,实验二 在特异性上略输于实验一,但是其他四个评价指标均高于实验一取得的实验结果,特 别是对于灵敏度 SE, 在实验一的基础上提升了 3.41%; 对比实验三, 实验四取得的 实验结果几乎所有评价指标都较优,此外,实验四对于所有评价指标均优于实验一; 实验五取得的视网膜血管分割结果在指标 SE 上略逊于实验二,但在其他评价指标上 取得了所有实验的最优值。在 DRIVE 数据集与 CHASE DB1 数据集上,实验二在灵

34

敏度 SE 上都取得了最好的表现,这同时也表明加入自校准卷积模块对于准确分割血 管具有良好的表现。

综上信息可以知道将自校准卷积模块与改进的空间注意力模块相结合的方法在 视网膜血管分割任务中是有效的。为了更直观的体现本文所提方法的优势,我们对分 割的结果图进行了可视化,在 DRIVE 数据集与 CHASE DB1 数据集上的分割结果分 别如图 4-3、4-4 所示。



图 4-3 在 DRIVE 数据集的对比实验视网膜血管分割结果可视化: (a)视网膜眼底图像, (b) 金标准图像, (c)实验一算法分割结果可视化图像, (d)实验二算法分割结果可视化图像, (e) 实验三算法分割结果可视化图像, (f)实验四算法分割结果可视化图像, (g)实验五(本文方 法)算法分割结果可视化图像

图 4-3 为 DRIVE 数据集测试示例的可视化,包括眼底图像、对应的金标准图像、 实验一(U-Net)、实验二(U-Net+SCC)、实验三(U-Net+SAM)、实验四(U-Net+ISAM) 和实验五(SCCU-Net)的视网膜血管分割结果。与 U-Net 相比,U-Net+SCC 实现了 更精确的血管分割,证明了自校准卷积模块的有效性。与 U-Net 相比,U-Net+SAM 产生更少的噪声,但是存在细小血管的漏检。与 U-Net+SAM 相比,U-Net+ISAM 能 保留更多的血管信息,表明了改进空间注意力的有效性。与前面四组实验得到的结果 相比,本文提出的 SCCU-Net (U-Net+SCC+ISAM)能够产生更准确的连续血管,表 明了本文方法的有效性。

汕头大学硕士学位论文



图 4-4 在 CHASE DB1 数据集的对比实验视网膜血管分割结果可视化: (a) 视网膜眼底图像, (b) 金标准图像, (c) 实验一算法分割结果可视化图像, (d) 实验二算法分割结果可视化图 像, (e) 实验三算法分割结果可视化图像, (f) 实验四(本文方法) 算法分割结果可视化图像, (g) 实验五(本文方法) 算法分割结果可视化图像

图 4-4 为 CHASE DB1 数据集测试示例的可视化,包括眼底图像、对应的金标准 图像、实验一(U-Net)、实验二(U-Net+SCC)、实验三(U-Net+SAM)、实验四 (U-Net+ISAM)和实验五(SCCU-Net)的视网膜血管分割结果。从可视化结果图像 中可以看出,U-Net 生成的分割结果图像里存在较多区域的血管丢失,而加入自校准 卷积模块或改进的空间注意力模块后的分割结果图中,细小血管能较好的被检测出来, 但同时也存在小部分区域的血管丢失。这表明虽然我们提出的方法对血管有良好的检 测分割能力,但是由于原眼底图像部分区域血管像素与背景像素的低对比度,导致算 法不能很好的检测出血管像素。

4.4.2 SCCU-Net 详细分割结果

(e)



图 4-5 本文所提方法在 DRIVE 数据集上最好(第一行)与最差(第二行)的视网膜血管分割结果, (a)和(e)为彩色眼底图像, (b)和(f)为金标准图像, (c)和(g)为算法分割概率 图, (d)和(h)为分割结果可视化图

(f)

	ACC	SE	SP	F1	AUC
最好(19)	0.9730	0.9191	0.9779	0.8497	0.9912
最坏(06)	0.9640	0.7280	0.9895	0.7976	0.9776
平均值	0.9680	0.8036	0.9840	0.8138	0.9840
标准差	0.0028	0.0547	0.0041	0.0172	0.0042

表 4-7 SCCU-Net 在 DRIVE 数据集上的实验结果

(g)

(h)

汕头大学硕士学位论文



图 4-6 本文所提方法在 STARE 数据集上最好(第一行)与最差(第二行)的视网膜血管分割结果,(a)和(e)为彩色眼底图像,(b)和(f)为金标准图像,(c)和(g)为算法分割概率 图,(d)和(h)为分割结果可视化图

	ACC	SE	SP	F1	AUC
最好(07)	0.9799	0.8866	0.9877	0.8722	0.9946
最坏(19)	0.9673	0.7071	0.9859	0.7428	0.9709
平均值	0.9728	0.8186	0.9856	0.8179	0.9870
标准差	0.0066	0.0718	0.0044	0.0424	0.0063

表 4-8 SCCU-Net 在 STARE 数据集上的实验结果

汕头大学硕士学位论文



图 4-7 本文所提方法在 CHASE DB1 数据集上最好(第一行)与最差(第二行)的视网膜血管分割结果, (a)和(e)为彩色眼底图像, (b)和(f)为金标准图像, (c)和(g)为算法分割 概率图, (d)和(h)为分割结果可视化图

	ACC	SE	SP	F1	AUC
最好(02)	0.9802	0.8636	0.9867	0.8228	0.9928
最坏(06)	0.9716	0.7334	0.9877	0.7658	0.9839
平均值	0.9756	0.8117	0.9867	0.8068	0.9888
标准差	0.0035	0.0404	0.0022	0.0220	0.0036

表 4-9 SCCU-Net 在 CHASE DB1 数据集上的实验结果

本文提出的方法 SCCU-Net 分别在 DRIVE, STARE 和 CHASE DB1 数据库上测 试的最好和最差情况如图 4-5, 4-6, 4-7 所示。所有数据集上情况最好的案例只有很 少的细小血管丢失。另一方面,在最坏的情况下显示了光照不均匀对 SCCU-Net 的影 响。无论是在最好的还是最坏的情况下,该方法都能够区分非血管结构和血管结构。 为了展示更详细的结果,对 DRIVE, STARE 和 CHASE DB1 测试集的每一张图像进 行了性能评估,详细结果见附录 A, 附录 B 和附录 C。

4.4.3 与先进视网膜血管分割方法比较

表 4-7 SCCU-Net 与先进视网膜血管分割方法在 DRIVE 上的比较结果

Methods	Year	ACC	SE	SP	F1	AUC
Ricci et al. ^[2]	2007	0.9563	-	-	-	0.9558
Azzopardi et al. ^[35]	2015	0.9442	0.7655	0.9704	-	0.9614
Roychowdhury et al. ^[34]	2016	0.9520	0.7250	0.9830	-	0.9620

Methods	Year	ACC	SE	SP	F1	AUC	
Liskowsk and Krawiec. ^[28]	2016	0.9495	0.7763	0.9768	-	0.9720	
Li et al. ^[27]	2016	0.9527	0.7569	0.9816	-	0.9738	
Fan et al. ^[38]	2016	0.9614	0.7191	0.9849	-	-	
MS-NFN ^[64]	2018	0.9567	0.7844	0.9819	-	0.9807	
R2U-Net ^[43]	2019	0.9556	0.7792	0.9813	0.8171	0.9784	
DEU-Net ^[44]	2019	0.9567	0.7940	0.9816	0.8270	0.9772	
SCCU-Net(Proposed)	2021	0.9680	0.8036	0.9840	0.8138	0.9840	

汕头大学硕士学位论文

表 4-7 SCCU-Net 与先进视网膜血管分割方法在 DRIVE 上的比较结果(续)

表 4-8 SCCU-Net 与先进视网膜血管分割方法在 STARE 上的比较结果

Methods	Year	ACC	SE	SP	F1	AUC
Ricci et al. ^[2]	2007	0.9584	-	-	-	0.9602
Azzopardi et al. ^[35]	2015	0.9497	0.7716	0.9701	-	0.9563
Roychowdhury et al. ^[34]	2016	0.9510	0.7720	0.9730	-	0.9690
Liskowsk and Krawiec. ^[28]	2016	0.9729	0.8554	0.9862	-	0.9928
Li et al. ^[27]	2016	0.9628	0.7726	0.9844	-	0.9879
Fan et al. ^[38]	2016	0.9588	0.6996	0.9787	-	-
R2U-Net ^[43]	2019	0.9712	0.8292	0.9862	0.8475	0.9914
SCCU-Net(Proposed)	2021	0.9728	0.8186	0.9856	0.8179	0.9870

表 4-9 SCCU-Net 与先进视网膜血管分割方法在 CHASE DB1 上的比较结果

Methods	Year	ACC	SE	SP	F1	AUC
Azzopardi et al. ^[35]	2015	0.9387	0.7585	0.9587	-	0.9487
Roychowdhury et al. ^[34]	2016	0.9530	0.7201	0.9824	-	0.9532
Li et al. ^[27]	2016	0.9581	0.7507	0.9793	-	0.9793
MS-NFN ^[64]	2018	0.9637	0.7538	0.9847	-	0.9825
R2U-Net ^[43]	2019	0.9634	0.7756	0.9820	0.7928	0.9815
DEU-Net ^[44]	2019	0.9661	0.8074	0.9821	0.8037	0.9812
SCCU-Net(Proposed)	2021	0.9756	0.8117	0.9867	0.8068	0.9888

表 4-7,表 4-8 和表 4-9 分别列举了了本文提出的方法和其他最先进方法在 DRIVE 数据集,STARE 数据集和 CHASE DB1 上的性能比较。结果表明,除 STARE 数据集外,本文提出的方法在其他所有数据集测试中取得了比最先进结果更好或者相 当的性能。在 DRIVE 数据集上,SCCU-Net 在除 F1 外的所有评价指标上均取得了较 好的结果。与 DEU-Net 相比, SCCU-Net 的准确率 ACC 比 DEU-Net 高 1.13%, 达到 了 96.80%, AUC 提高了 0.68%, 在灵敏度 SE 与特异性 SP 上分别提高了 0.96%和 0.24%, 且均优于其他方法。在 STARE 数据集上,本文算法对比最先进算法稍有逊色,但除 灵敏度 SE 与 F1 分数,其他性能指标均取得与最先进评价指标相当的结果。在 CHASE DB1 数据集上,所提出的 SCCU-Net 在所有指标上都达到了最佳性能。以上结果表明, SCCU-Net 的分割性能优于以往的视网膜血管分割方法,表明所提出的 SCCU-Net 是 一种有效的视网膜血管分割方法。

第5章 结论及展望

5.1 结论

眼底视网膜血管的结构分析在针对青光眼、白内障等眼科疾病、糖尿病、高血压 和动脉硬化等心血管类疾病的诊断发挥着越来越重要的作用,而对视网膜血管精确的 分割是诊断此类疾病的首要步骤,也是最为重要的一步。在传统的临床诊断中,往往 是由专科医生人工的在眼底图像中挑选出血管的部分,这个方式耗时耗力,并且分割 的准确率也不高,所以借助计算机对视网膜血管进行自动化分割对眼科类疾病或心血 管类等疾病的诊断具有重要的意义。而深度学习技术在视网膜血管分割的任务中具有 较好的表现,许多深度学习的方法依赖大量的训练数据,为了获取足够的训练数据, 往往采用对眼底图像进行随机切片的方法来增加训练数据,但是这种方法割裂了切片 图像之间的上下文信息联系。此外,由于眼底图像中血管与背景对比度低、视网膜病 变、血管粗细差别较大等影响,自动分割高精度的视网膜血管也是一项极具挑战的任 务。

针对上述视网膜血管分割任务中存在的问题,本文基于端到端的深度学习方法,提出了自校准卷积 U-Net(SCCU-Net)算法,在视网膜血管分割任务中实现了较好的结果。本文的贡献点主要包括:

(1) 针对单一传统卷积操作无法获得不同尺度的感受野,在不显著增加参数量与 计算量的情况下,采用自校准卷积模块能够产生更大的感受野,相比于标准卷积, 该模块产生的特征图更具有区分度。我们将自校准卷积模块(Self-Calibrated Convolution)结合到 U-Net 网络中,并提出一种新的端到端的深度神经网络 SCCU-Net。

(2) 我们在 CBAM 的基础上设计了一种改进的空间注意力模块,该模块通过在 卷积操作前计算空间注意力权值图,可以保留更加完整的空间特征信息。

(3) 本文通过丰富的对比实验验证了所采用方法在视网膜血管分割任务中的有效性。性能评价指标结果说明了我们提出方法的优势,可以实现更好的或可与最 先进的视网膜血管分割方法相媲美的结果。

5.2 课题展望

深度学习技术在医学领域的发展受到越来越多研究者的关注。本文基于深度学习 提出的自校准卷积 U-Net(SCCU-Net)在视网膜血管分割中取得了良好的表现,但 是仍存在一些问题,对未来的视网膜血管分割任务而言是一项艰巨的挑战。 (1) 出于对患者隐私性的保护,在全世界范围公开的眼底图像数据非常稀少,并 且由于眼底图像信息不统一、数据获取困难等原因研究者往往很难去训练出比较 好的视网膜血管分割模型。在未来,或许可以利用生成对抗网络(GAN)去生成 数据集,从而扩大训练所需的数据。

(2) 本文提出的算法虽然可以检测分割出大部分血管像素,但是针对某些血管与 背景像素对比度低的区域、血管较细的像素点,我们提出的算法表现的还不够好, 存在一定范围的血管漏检。

(3) 在今后的工作中,本文提出的算法模型可以尝试加入模型剪枝技术,减少网络模型参数,以便于移植到移动设备中,开发出集数据提取、数据处理、结果显示一体的移动端视网膜血管检测设备,这将会产生深远的医疗工业应用价值。 针对这些问题,还需要研究者们结合大量理论和先验知识继续努力的探索。

参考文献

[1] Chaudhuri S, Chatterjee S, Katz N, et al. Detection of blood vessels in retinal images using two-dimensional matched filters[J]. IEEE Transactions on Medical Imaging, 1989, 8 (3): 263-269.

[2] Ricci E, Perfetti R. Retinal blood vessel segmentation using line operators and support vector classification[J]. IEEE Transactions on Medical Imaging, 2007, 26 (10): 1357-1365.

[3] Nguyen U T V, Bhuiyan A, Park L A F, et al. An effective retinal blood vessel segmentation method using multi-scale line detection[J]. Pattern Recognition, Pergamon, 2013, 46 (3): 703-715.

[4] Ding Y C, Ward W O C, Wästerlid T, et al. Three-dimen sional vessel segmentation using a novel combinatory filter framework[J]. Physics in Medicine and Biology, 2014, 59 (22): 7013-7029.

[5] Frangi A F, Niessen W J, Vincken K L, et al. Multiscale vessel enhancement filtering[C]. Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention 1998. Cambridge. USA. 1998. 130-137.

[6] Qian X N, Brennan M P, Dione D P, et al. A non-parametric vessel detection method for complex vascular structures[J]. Medical Image Analysis, 2009, 13 (1): 49-61.

[7] Rivest-Hénault D, Cheriet M. 3-D curvilinear structure detection filter via structure-ball analysis[J]. IEEE Transactions on Image Processing, 2013, 22 (7): 2849-2863.

[8] Zhang B, Zhang L, Zhang L, et al. Retinal vessel extraction by matched filter with first-order derivative of Gaussian[J]. Computers in Biology and Medicine, 2010, 40 (4): 438-445.

[9] Lajevardi S M, Arakala A, Davis S A, et al. Retina verification system based on biometric graph matching[J]. IEEE Transactions on Image Processing, 2013, 22 (9): 3625-3635.

[10] Kovács G, Hajdu A. A self-calibrating approach for the segmentation of retinal vessels by template matching and contour reconstruction[J]. Medical Image Analysis, 2016: 24-46.

[11] Vazquez S G, Cancela B, Barreira N, et al. Improving retinal artery and vein classification by means of a minimal path approach[J]. Machine Vision & Applications, 2013, 24 (5): 919-930.

[12] Bekkers E, Duits R, Berendschot T, et al. A multi-orientation analysis approach to retinal vessel tracking[J]. Journal of Mathematical Imaging & Vision, 2014, 49 (3): 583-610.

[13] Khan K B, Khaliq A A, Jalil A, et al. A robust technique based on VLM and Frangi filter for retinal vessel extraction and denoising[J]. Plos One, 2018, 13 (2): e0192203.

[14] Wang W, Zhang J, Wu W, et al. An automatic approach for retinal vessel segmentation by multi-scale morphology and seed point tracking[J]. Journal of Medical Imaging & Health Informatics, 2018, 8 (2): 262-274.

[15] Najman L, Talbot H. Introduction to mathematical morphology[J]. Computer Vision Graphics and Image Processing, 1986, 35 (3): 283-305.

[16] Zana F, Klein J C. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation[J]. IEEE Transactions on Image Processing, 2001, 10 (7): 1010-1019.

[17] Passat N, Ronse C, Baruthio J, et al. Watershed and multimodal data for brain vessel segmentation[J]. Image and Vision Computing, 2007, 25 (4): 512-521.

[18] Bouraoui B, Ronse C, Baruthio J, et al. 3D segmentation of coronary arteries based on advanced mathematical morphology techniques[J]. Computerized Medical Imaging and Graphics the Official Journal of the Computerized Medical Imaging Society, 2010, 34 (5): 377-387.

[19] Fan Z, Lu J W, Wei C M, 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.

[20] Fraz M M, Barman S A, Remagnino P, et al. An approach to localize the retinal blood vessels using bit planes and centerline detection[J]. Computer Methods and Programs in Biomedicine, 2012, 108 (2): 600-616.

[21] Babin D, Pizurica A, Bellens R, et al. Generalized pixel profiling and comparative segmentation with application to arteriovenous malformation segmentation[J]. Medical Image Analysis, 2012, 16 (5): 991-1002.

[22] Babin D, Pizurica A, Vylder J D, et al. Brain blood vessel segmentation using line-shaped profiles[J]. Physics in Medicine and Biology, 2013, 58 (22): 8041-61.

[23] Dufour A, Tankyevych O, Naegel B, et al. Filtering and segmentation of 3D angiographic data: Advances based on mathematical morphology[J]. Medical image analysis, 2013, 17 (2): 147-164.

[24] Sigurosson E M, Valero S, Benediktsson J A, et al. Automatic retinal vessel extraction based on directional mathematical morphology and fuzzy classification[J]. Pattern Recognition Letters, 2014, 47 (1): 164-171.

[25] Nekovei R, Sun Y. Back-propagation network and its configuration for blood vessel detection in angiograms[J]. IEEE Transactions on Neural Networks, 2002, 6 (1): 64-72.

[26] Rahebi J, Hardalaç, F. Retinal blood vessel segmentation with neural network by using gray-level co-occurrence matrix-based features[J]. Journal of Medical Systems, 2014, 38 (8): 1-12.

[27] Li Q L, Feng B W, Xie L P, et al. A cross-modality learning approach for vessel segmentation in retinal images[J]. IEEE Transactions on Medical Imaging, 2015, 35 (1) : 109-118.

[28] Liskowski P, Krawiec K. Segmenting retinal blood vessels with deep neural networks[J]. IEEE Transactions on Medical Imaging, 2016, 35 (11): 2369-2380.

[29] Osareh A, Shadgar B. Automatic blood vessel segmentation in color images of retina[J]. Iranian Journal of Science and Technology Transaction B: Engineering, 2009, 33 (B2): 191-206.

[30] Xu L L, Luo S Q. A novel method for blood vessel detection from retinal images[J]. BioMedical Engineering OnLine, 2010, 9 (1): 14.

[31] NIEMEIJER M, STAAL J J, GINNEKEN B V, et al. Comparative study of retinal vessel segmentation methods on a new publicly available database[J]. Proceedings of SPIE - The International Society for Optical Engineering, 2004, 5370: 648-656.

[32] Soares J V B, Leandro J J G, Cesar R M, et al. Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification[J]. IEEE Transactions on Medical Imaging, 2006, 25 (9): 1214-1222.

[33] Xiao Z Y, Adel M, Bourennane S. Bayesian method with spatial constraint for retinal vessel segmentation[J]. Computational and Mathematical Methods in Medicine, 2013: 1-9.

[34] Roychowdhury S, Koozekanani D D, Parhi K K. Blood vessel segmentation of fundus images by major vessel extraction and subimage classification[J]. IEEE Journal of Biomedical and Health Informatics, 2015, 19 (3): 1118-1128.

[35] Azzopardi G, Strisciuglio N, Vento M, et al. Trainable COSFIRE filters for vessel delineation with application to retinal images[J]. Medical Image Analysis, 2015, 19 (1): 46-57.

[36] Fraz M M, Remagnino P, Hoppe A, et al. An ensemble classification-based approach applied to retinal blood vessel segmentation[J]. IEEE Transactions on Biomedical Engineering, 2012, 59 (9): 2538-2548.

[37] Wang S L, Yin Y L, Cao G B, et al. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning[J]. Neurocomputing, 2015, 149 (pt.b): 708-717.

[38] Fan Z, Rong Y B, Lu J W, et al. Automated blood vessel segmentation in fundus image based on integral channel features and random forests[C]. World Congress on Intelligent Control and Automation 2016. Guilin. China. 2016. 2063-2068.

[39] Akram M U, Khalid S, Tariq A, et al. Detection and classification of retinal lesions for grading of diabetic retinopathy[J]. Computers in Biology and Medicine, 2014, 45: 161-171.

[40] Memari N, Ramli A R, Saripan M I B, et al. Supervised retinal vessel segmentation from color fundus images based on matched filtering and AdaBoost classifier[J]. Plos One, 2017, 12 (12): e0188939.

[41] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation[C]. Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention 2015. Munich. Germany. 2015: 234-241.

[42] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2015. Boston. USA. 2015. 3431-3440.

[43] 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]. arXiv preprint arXiv:1802.06955, 2018.

[44] Wang B, Qiu S, He H G. Dual encoding U-Net for retinal vessel segmentation[C]. Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention 2019. Shenzhen. China. 2019: 84-92.

[45] Glorot X, Bordes A, Bengio Y. Deep sparse rectifier neural networks[C].Proceedings of the 14th International Conference on Artificial Intelligence and Statistics 2011. Ft. Lauderdale. USA. 2011. 315-323.

[46] Maas A L, Hannun A Y, Ng A Y. Rectifier nonlinearities improve neural network acoustic models[C]. Proceedings of the International Conference on Machine Learning 2013. Atlanta. USA. 30 (1): 3.

[47] He K M, Zhang X Y, Ren S Q, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification[C]. Proceedings of the IEEE International Conference on Computer Vision 2015. Santiago. Chile. 2015. 1026-1034.

[48] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C].Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).IEEE, 2014: 1-9.

[49] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2016. Las Vegas. USA. 2016. 2818-2826.

[50] Ioffe S, Szegedy C. Batch Normalization: Accelerating deep network training by reducing internal covariate shift[J]. 2015.

[51] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning[J]. 2016.

[52] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2016. Las Vegas. USA. 2016. 770-778.

[53] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39 (12): 2481-2495.

[54] Simonyan K, Zisserman A. Very Deep convolutional networks for large-scale image recognition[J]. Computer Science, 2014.

[55] Anderson J R. Cognitive psychology and its implications[M]. WH Freeman/Times Books/Henry Holt & Co, 1990.

[56] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 42 (8): 2011-2023.

[57] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C].Proceedings of the European Conference on Computer Vision 2018. Munich. Germany.2018. 3-19.

[58] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C].Conference and Workshop on Neural Information Processing Systems 2017. Los Angle.USA. 2017. 6000-6010.

48

[59] Wang X L, Girshick R, Gupta A, et al. Non-local neural networks[C].Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2018.Salt Lake City. USA. 2018. 7794-7803.

[60] Fu J, Liu J, Tian H J, et al. Dual attention network for scene segmentation[C].Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2019.Los Angle. USA. 2019. 3146-3154.

[61] Liu J J, Hou Q, Cheng M M, et al. Improving convolutional networks with self-calibrated convolutions[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2020.Virtual. 2020. 10093-10102.

[62] Ioannou Y, Robertson D, Cipolla R, et al. Deep roots: Improving CNN efficiency with hierarchical filter groups[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2017. Honolulu. Hawaii. 2017. 5977-5986.

[63] Bottou L. Large-scale machine learning with stochastic gradient descent[J]. Proceedings of COMPSTAT, 2010: 177-186.

[64] Wu Y C, Xia Y, Song Y, et al. Multiscale network followed network model for retinal vessel segmentation[C]. Proceedings of the International Conference on Medical Image Computing & Computer-assisted Intervention 2018. Strasbourg. France. 2018. 119-126.

	ACC	SE	SP	F1	AUC
1	0.9670	0.8687	0.9766	0.8245	0.9886
2	0.9683	0.8475	0.9821	0.8456	0.9877
3	0.9611	0.7288	0.9868	0.7888	0.9786
4	0.9695	0.8027	0.9864	0.8289	0.9804
5	0.9660	0.7409	0.9893	0.8033	0.9784
6	0.9640	0.7280	0.9895	0.7976	0.9776
7	0.9681	0.7581	0.9892	0.8129	0.9805
8	0.9660	0.7248	0.9887	0.7857	0.9787
9	0.9691	0.7260	0.9905	0.7920	0.9803
10	0.9688	0.8038	0.9836	0.8091	0.9827
11	0.9648	0.8092	0.9801	0.8044	0.9807
12	0.9691	0.8226	0.9829	0.8212	0.9871
13	0.9660	0.7941	0.9846	0.8203	0.9819
14	0.9703	0.8429	0.9815	0.8211	0.9879
15	0.9720	0.8359	0.9824	0.8101	0.9870
16	0.9702	0.8212	0.9850	0.8325	0.9876
17	0.9681	0.7777	0.9856	0.8045	0.9857
18	0.9690	0.8651	0.9780	0.8158	0.9890
19	0.9730	0.9191	0.9779	0.8497	0.9912
20	0.9703	0.8544	0.9795	0.8087	0.9894
平均值	0.9680	0.8036	0.9840	0.8138	0.9840
标准差	0.0028	0.0547	0.0041	0.0172	0.0042

附录 A DRIVE 数据库测试集详细实验结果

	ACC	SE	SP	F1	AUC
1	0.9695	0.8080	0.9835	0.8086	0.9868
2	0.9766	0.7715	0.9912	0.8144	0.9855
3	0.9772	0.8426	0.9858	0.8158	0.9879
4	0.9660	0.6990	0.9874	0.7532	0.9835
5	0.9629	0.7404	0.9850	0.7831	0.9827
6	0.9738	0.8811	0.9802	0.8240	0.9883
7	0.9799	0.8866	0.9877	0.8722	0.9946
8	0.9774	0.8924	0.9857	0.8755	0.9940
9	0.9766	0.8909	0.9852	0.8735	0.9935
10	0.9721	0.8615	0.9826	0.8421	0.9895
11	0.9762	0.8731	0.9852	0.8548	0.9931
12	0.9773	0.9046	0.9831	0.8560	0.9942
13	0.9802	0.8792	0.9889	0.8749	0.9939
14	0.9702	0.8643	0.9795	0.8235	0.9889
15	0.9574	0.6517	0.9921	0.7573	0.9743
16	0.9721	0.7765	0.9913	0.8367	0.9855
17	0.9786	0.8210	0.9870	0.7952	0.9844
18	0.9826	0.7980	0.9909	0.7977	0.9878
19	0.9673	0.7071	0.9859	0.7428	0.9709
20	0.9631	0.8232	0.9736	0.7567	0.9816
平均值	0.9728	0.8186	0.9856	0.8179	0.9870
标准差	0.0066	0.0718	0.0044	0.0424	0.0063

附录 B STARE 数据库测试集详细实验结果

	100	2 0	G D	11	AUG
	ACC	SE	SP	Fl	AUC
1	0.9806	0.8554	0.9876	0.8245	0.9927
2	0.9802	0.8636	0.9867	0.8228	0.9928
3	0.9703	0.8117	0.9826	0.7972	0.9850
4	0.9733	0.8372	0.9841	0.8217	0.9893
5	0.9757	0.8102	0.9865	0.8038	0.9895
6	0.9716	0.7334	0.9877	0.7658	0.9839
7	0.9778	0.8118	0.9901	0.8344	0.9926
8	0.9751	0.7711	0.9878	0.7840	0.9843
平均值	0.9756	0.8118	0.9867	0.8068	0.9888
标准差	0.0035	0.0404	0.0022	0.0220	0.0036

附录 C CHASE DB1 数据库测试集详细实验结果

攻读学位期间主要研究成果

- 一、发表的学术论文
- (一) 已见刊或录用论文
- [1] 范衠, 熊宇, 罗晨林 等. 电厂集控室智能监盘机器人软件系统设计与研发[J]. 广 东电力, doi:10.3969/j.issn.1007-290X.2020.S1.037

二、专利

[1] 一种手眼触屏机械臂, 专利号: ZL201921899215.6

三、奖项

[1] 第十五届中国研究生电子设计竞赛华南赛区二等奖

致谢

三年的研究生学习生活转瞬即逝,刚入学时的迷茫仿佛还在眼前,入学后,老师同学、 师兄师姐都给予过我非常大的帮助,借此简短的篇幅,对在我研究生期间所有帮助过我的 人表示最诚挚的谢意。

首先我要感谢我的导师范衠教授。范老师学识渊博,在学术上严谨务实的态度指引我 懂得如何去做科研,从踏入实验室开始,范老师就以严谨的态度时时刻刻教导我们做事情。 在学习工作中,范老师毫不吝啬自己的资源,给我们提供了充足的资源平台,帮助我们做 出更加优秀的成果。与此同时,范老师更多的是教会我为人处世的道理,言传身教。在此 向我的导师范衠教授致以崇高的敬意,衷心感谢范教授的知遇之恩。

其次,我要感谢实验室的小伙伴们,无论是学习中还是生活中都给我带来了很大的帮助。感谢邱本章、安康师兄,他们积极的工作态度,扎实的理论知识深深指引着我前进。 感谢同窗谢梓歆、胡星晨、罗晨林等同学,感谢他们在学习生活中三年的陪伴,我们互相 勉励,共同进步,我将记住我们一起奋斗的这段岁月。

最后我想感谢我的家人,他们是我读研究生背后最大的支持,感谢他们在背后默默地 付出,帮助我度过一个个难关。

作者: 熊宇

2021年4月8日