



汕頭大學

SHANTOU UNIVERSITY

硕士学位论文

题目 基于分频卷积神经网络的视网膜血管分割

英文题目 Octave Convolution Network for Retinal
Vessel Segmentation

姓名 莫嘉杰 学号 11609023

所在学院 工学院 导师姓名 范衡

专业 信息与通信工程

入学日期 2016年9月 答辩日期 2019年6月

汕头大学学位论文原创性声明和学位论文使用授权声明

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： 莫嘉杰

日期： 2019 年 5 月 13 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： 莫嘉杰

日期： 2019 年 5 月 13 日

导师签名： 张

日期： 2019 年 5 月 13 日

摘要

视网膜血管结构被广泛地用于眼科疾病，糖尿病，心血管类疾病的诊断，筛查和临床研究。然而，由于视网膜病变、血管中心反光现象、血管背景对比度低、血管分支和拓扑结构复杂等影响，视网膜血管分割是一项具有挑战性的任务。

为了克服上述挑战，本文提出了基于分频卷积神经网络（Octave UNet）的视网膜血管分割方法，实现了高精度快速的视网膜眼底图像血管自动分割。为了提高网络特征对血管分割的判别能力，基于分频卷积的特征编码器被用于高低分频的卷积特征的提取。此外，为了让模型具备学习如何解码这些高低分频的卷积特征的能力，本文提出了一种分频转置卷积操作，并在其基础上提出了 Octave UNet 网络架构。

本文所提出的方法在四个公开数据库：DRIVE, STARE, CHASE_DB1 和 HRF 数据库上进行了算法评估。实验结果表明，所提出的方法已获得与现有先进方法相比更好或相当的视网膜血管分割结果。此外，本文所提出的方法在对各类视网膜病变和阈值变化具有一定的鲁棒性。

关键词：视网膜眼底图像，血管分割，分频转置卷积，分频卷积神经网络

Octave Convolution Network for Retinal Vessel Segmentation

Jiajie Mo (Information and Communication Engineering)

Directed by Zhun Fan

ABSTRACT

Retinal vascular structure is widely used in diagnosis, screening and clinical research of ophthalmic diseases, diabetes, and cardiovascular diseases. However, segmenting retinal vessel from fundus images is a challenging task because of the presence of pathologies and vessel central reflection, low contrast of blood vessel and background, complex vascular topology and branching pattern.

To overcome these challenges, an encoder-decoder based convolution neural network named Octave UNet was proposed in this paper to achieve accurate and fast retinal vessel segmentation. To improve the discriminative capacity of Octave UNet, features with multiply spatial frequencies are learned by adopting octave convolution inside the encoder blocks. Moreover, in order to provide the capability for learning how to decoder these multi-frequencies features, a novel operation named octave transposed convolution was proposed and upon which the Octave UNet architecture is build.

The proposed method was evaluated on four publicly available databases, DRIVE, STARE, CHASE_DB1 and HRF databases. Extensive experimental results demonstrated that the proposed method has achieved better or comparable results to state-of-the-art methods. Furthermore, the proposed method has shown robustness against a variety of pathologies and varying threshold.

KEY WORDS: Retinal Fundus Images, Vessel Segmentation, Octave UNet

目录

摘要.....	I
ABSTRACT.....	II
目录.....	III
第一章 绪论.....	1
1.1 研究背景与意义.....	1
1.1.1 视网膜血管分割的研究意义.....	1
1.1.2 视网膜血管分割算法的研究意义.....	2
1.1.3 视网膜血管分割的难点和挑战.....	2
1.2 相关研究概况.....	3
1.2.1 传统数字图像预处理方法.....	4
1.2.2 基于无监督机器学习的方法.....	4
1.2.3 基于有监督机器学习与深度学习的方法.....	5
1.3 主要研究内容.....	8
1.4 论文章节安排.....	8
第二章 基础技术研究.....	9
2.1 公共数据库与算法评价指标.....	9
2.1.1 彩色眼底图像公共数据库.....	9
2.1.2 视网膜血管分割算法评价指标.....	9
2.2 深度神经网络概述.....	12
2.2.1 人工神经网络.....	13
2.2.2 卷积神经网络.....	17
2.2.3 全卷积神经网络.....	20
2.2.4 基于编码器解码器架构的深度卷积网络.....	22
第三章 基于分频卷积网络的视网膜血管分割算法.....	25
3.1 分频卷积神经网络.....	25
3.1.1 图像金字塔特征与低频图像空域中的冗余.....	25
3.1.2 高低分频卷积特征.....	26
3.1.3 分频卷积操作.....	27
3.2 基于编码器解码器架构的分频卷积网络.....	28
3.2.1 分频转置卷积层.....	29
3.2.2 Octave UNet 架构设计.....	31
3.2.3 网络参数优化.....	38
第四章 实验结果与分析.....	39
4.1 实验设计.....	39
4.1.1 训练集和测试集设置.....	39
4.1.2 实验环境及参数设置.....	39
4.1.3 实验数据库及评价指标说明.....	39
4.2 实验结果与分析.....	40
4.2.1 各数据库上的实验结果.....	40

4.2.2	与先进方法的结果对比.....	44
4.2.3	阈值敏感度分析.....	47
4.2.4	案例分析.....	49
第五章	结论及展望.....	53
5.1	结论.....	53
5.2	研究课题展望.....	54
参考文献	55
附录 A	DRIVE 数据库测试集详细实验结果.....	59
附录 B	STARE 数据库详细实验结果.....	61
附录 C	CHASE_DB1 数据库详细实验结果.....	63
附录 D	HRF 数据库详细实验结果.....	65
附录 E	模型训练过程可视化.....	67
致谢	69
攻读硕士学位期间的主要工作成果	71

第一章 绪论

1.1 研究背景与意义

视网膜眼底图像分析被广泛用于青光眼和白内障等眼科疾病，糖尿病，高血压和动脉硬化等心血管类疾病的诊断，筛查和临床研究^[1-5]。其中，视网膜血管结构作为可以通过非入侵方式观察到的人体血管结构，对视网膜血管的分割是视网膜眼底图像分析中最重要的一个步骤，也是过去二十年间大量医学图像处理工作的研究热点。

在临床实践中，视网膜血管分割一般由眼科医师标注完成，是一项需要熟练技巧，乏味而耗时的工作。为了减少眼科医师的工作量，提高疾病诊断效率，能由计算机自动化完成视网膜血管分割的智能算法和计算机辅助诊断系统具有重大的研究意义和临床价值。

1.1.1 视网膜血管分割的研究意义

1.1.1.1 视网膜血管结构在临床实践中的作用

近年来医学研究表明，糖尿病视网膜病变(Diabetic Retinopathy)，青光眼(Glaucoma)和年龄相关性黄斑变性(Age-related Macular Degeneration)是导致老龄人群失明的三大原因^[6-8]。视网膜疾病已经成为全世界老龄人口最重要的公共健康问题之一。

医学研究表明，糖尿病视网膜病变和视网膜血管结构变化有着密切关系^[7,9]。糖尿病患者因为其血液成分的变化导致视网膜血管出血和视网膜毛细血管异常增生。甚至可能形成增殖性糖尿病视网膜病变(Proliferative Diabetic Retinopathy)，其主要病征是为了补偿受损血管而产生的视网膜血管异常增生。如果不能在早期发现糖尿病视网膜病变，患者可能因此而视力受损，甚至失明^[7,9]。而视网膜血管结构病变作为一种可以通过非入侵方式观察到早期糖尿病病征，在糖尿病的早期诊断，筛查和治疗中具有重要的意义。

此外，视网膜血管分割还能帮助其他视网膜病灶的准确诊断^[10]。虽然视网膜血管分割并不能直接参与如微动脉瘤(Microaneurysms)，硬性渗出(Hard Exudates)等病灶的定位和诊断，但是因为这些病灶通常出现在容易出血的高弯曲度的血管结构和较薄的血管结构附近，对视网膜血管的精确分割能有效地帮助微动脉瘤和硬性渗出等视网膜眼底病灶的定位和诊断。

1.1.1.2 视网膜血管结构在大规模人口调查和医学研究的作用

视网膜血管结构是一种可以通过非入侵方式观察到的人体较深层的血管结构，其图像的获取较为方便，被广泛地用于各种大规模人口调查和医学研究。文献中已有大量与视网膜血管有关的大规模人口调查和医学研究^[11-16]。例如，相关研究^[11,14]中报导了视网膜血管结构变化与高血压有着密切关系，研究^[15,16]报导了中风和视网膜血管结构变化有着统计相关性。

1.1.1.3 视网膜血管结构在生物特征识别中的作用

生物特征识别技术的相关研究表明，每个人双眼的视网膜血管分布都是因人而异的，具有唯一性，而且视网膜血管结构和指纹、人脸等生物特征不同，难以通过外科手术等方式进行修改，具有较强的稳定性^[17,18]。

而另一方面，视网膜血管结构作为人体一种可以通过非入侵方式观察到的血管结构，视网膜血管图像的获取也较为方便快捷。通过视网膜血管分割算法提取得到的视网膜血管结构非常适合用于基于生物特征的个体识别。

视网膜血管分割算法是基于视网膜血管结构的个人认证系统（Personal Authentication System）中的一个核心组成部分，具有重要的研究意义和商业应用价值。

1.1.2 视网膜血管分割算法的研究意义

虽然视网膜血管结构的研究有着重要的临床实践、医学研究和商业应用价值，但传统的视网膜血管分割工作一般由眼科医师等具备相关专业知识和技能的观察者手工标注完成。这是一项需要熟练技巧，乏味而耗时的的工作。再者，不同观察者的经验判断，主观感受都不一样，即便是对同一张视网膜眼底图像进行血管分割标注，不同观察者给出的血管分割结果也各不相同。因此传统的手工视网膜血管分割不仅耗费大量人力，而且没有办法保证血管分割的精度和效率。

为了减少工作量，提高疾病诊断效率，能由计算机执行，自动化完成视网膜血管分割的智能算法和计算机辅助诊断系统（Computer Aided Diagnostic, CAD）具有重大的研究意义和临床实践价值。

1.1.3 视网膜血管分割的难点和挑战

1.1.3.1 视网膜眼底图像的复杂性

视网膜眼底图像的获取过程，是将三维眼球的视网膜结构投影到二维成像平面的过程。在这个过程中，非血管组织结构和视网膜血管结构重叠成像，降低了视网膜血管结构的清晰度，特别是宽度较小的血管分支和眼底图像背景的对比度通常较低，使得

精确分割细小的视网膜血管分支成为一个难题^[2,3,5]。

此外，在每个人的眼球深度不一，不同眼底相机成像效果参差不齐等等因素的影响下，获取到的视网膜眼底图像可能会存在如模糊（Blur），光照不均（Uneven Illumination），眼底图像的数据分布与成像设备高度相关（Inter-camera Variability）等难题^[2,5]。

1.1.3.2 复杂的视网膜血管拓扑结构

视网膜血管结构主要由动脉（Arteries），静脉（Veins）和其他在视网膜眼底图像中可见的血管分支组成。不同视网膜眼底图像中的血管结构形态学差异巨大，根据眼底图像分辨率和血管粗细的不同，血管图像的像素宽度可以从1个像素到20多个像素不等^[2]。动态范围变化大的血管尺度使得视网膜血管分割算法容易出现两种极端，或为了兼顾分割小尺度的血管分支而过度分割大尺度的血管结构，或为了精确分割血管主干而没法识别小尺度血管。

而在年幼个体的视网膜眼底图像中，高弯曲度的血管结构更为常见^[2]。高弯曲度的血管分割精度困难较大。此外，因为年幼个体的血氧含量较高，血管中心反光（Central Vessel Reflex）现象也更加显著^[2,19]。该现象成因是血管中心反光区域的图像像素灰度值较亮，和普通血管像素灰度值的分布范围差异较大。血管中心反光使得带有中心反光的血管结构容易被误判为两条平行的血管，对这类血管的精确分割也是一种挑战。

1.1.3.3 来自其他图像背景的干扰

除了图像质量和复杂血管结构带来的干扰以外，正常的视网膜眼底图像中还包含着如视盘（Optic Disc）、黄斑（Macular）等其他组织结构的背景图像。此外，异常的视网膜图像中，还可能存在着如渗出（Exudates）、血管出血（Hemorrhages）、棉絮状斑（Cotton Wool Spots）和微动脉瘤（Microaneurysms）等病灶。眼底图像的复杂图像背景信息给视网膜血管分割带来了一定的难度。

1.2 相关研究概况

视网膜血管分割任务一方面具有重大的研究意义和临床价值，另一方面还存在着许多待解决的困难和挑战，因此而吸引着医学图像处理、图像分割算法和计算智能领域研究者的关注，大量的相关科研文献被发表，一直是在近二十年间医学图像处理领域的研究热点。

在现有文献中，视网膜血管分割任务的解决方法种类繁多，研究思路和针对的具体问题各不相同。主要可以分为三大类：传统数字图像预处理方法、基于无监督机器学习的方法和基于有监督机器学习的方法。

1.2.1 传统数字图像预处理方法

视网膜眼底图像因为各种成像条件限制，图像质量常会受如噪声、光照不均衡和血管和背景对比度低等因素影响。其图像预处理常常在实际血管分割之前进行，能有效改善算法的分割精度。

1.2.1.1 血管中心反光去除

血管中心反光现象表现为血氧含量较高血管的图像中，血管中心比血管壁更亮。这容易使血管分割算法将一条带有中心反光的血管误判为两条平行的小尺度血管。

去除血管中心反光的经典方法是属于形态学操作（Morphological Operations）的开操作（Opening Operation），在需要选择适当的尺度来定义结构元素体（Structural Element）后，用元素体中心遍历眼底图像，根据预先定义好的连通域关系，通过先腐蚀（Erosion）再膨胀（Dilation），通过灰度值，从而去除中心反光区域。

1.2.1.2 眼底图像光照不均衡矫正

因为眼底图像采集过程中需要主动补光，采集到的眼底图像受到光照不均的影响时，不同区域图像的灰度值变化较大，文献中通常采用直方图均衡（Histogram Equalization）来矫正灰度值不均衡。

Foracchia et al.^[20]提出了一种用于视网膜眼底图像对比度归一化（Contrast Normalization）的方法。该方法首先估算眼底图像的背景亮度和对比度，然后用估算所得统计量来对全图归一化。

1.2.1.3 血管图像增强

如顶帽变换（Top Hat Transformation）等形态学操作常被用于血管图像增强^[21]。这些方法的基本思路是通过形态学处理的开操作算子（Opening Operator）对眼底图像的局部背景进行估计，然后从原图中减去估计所得的背景，进而增强血管像素和背景的对对比度。

1.2.2 基于无监督机器学习的方法

1.2.2.1 匹配滤波

匹配滤波（Matched Filters）方法利用了二维滤波器对视网膜血管边缘和眼底图像背景交界处的像素灰度值变化进行建模，并用建模所得的二维滤波器模板对眼底图像进行卷积操作，计算滤波器模板与图像局部区域的相似程度，从而增强血管和背景的对对比度，进而方便后续的血管分割操作。

针对视网膜眼底图像中背景噪音强，小尺度的血管和背景对比度低，难以分辨的难题，Chaudhuri et al.^[22]首次提出了基于匹配滤波的血管图像增强方法。该方法利用二维高斯滤波器对血管边缘和背景相交处的像素灰度值变化进行建模，并用这些二维高斯滤波器对眼底图像进行检测，最后利用自适应阈值处理实现视网膜血管的分割。

此后，基于特征值的滤波器(Eigenvalue-based Filters)^[23]、多尺度线性滤波器(Multi-Scales Linear Operators)^[24]、小波(Wavelet)^[25]等不同滤波方式的像素灰度值匹配滤波方法被文献报导用于视网膜血管分割。Ricci et al.^[26]提出了一种用于视网膜血管分割的线性特征描述子(Line Detector)，利用目标像素和不同方向基线的平均灰度响应，能高效地处理血管中心反光，但同时也可能将临近的两条平行血管误判为带中心反光的一条血管。Nguyen et al.^[27]在该方法的基础上，引入了多尺度的基线检测器(Multi-Scale Base Line Detector)，该方法能够正确识别带有中心反光的血管，且利用了中心反光区域像素数量比其周围血管壁像素数量少的特征，减少了融合平行血管的误判现象。

1.2.2.2 血管跟踪方法

血管跟踪(Vessel Tracking)方法利用血管树状结构的连续性，通过定义种子点(Seed Point)作为初始生长点，通过定义局部生长过程的约束方式，指导跟踪血管中心线和血管边缘。种子点的定义可以是半自动地手动标定，或者通过其他方法自动获得。而不同血管跟踪方法地差异主要在于其跟踪约束的不同。Cohen et al.^[28]提出了一种利用主动轮廓模型(Active Contour Model)的方法，该方法通过最小化能量约束函数，寻找两个种子点之间最短路径实现视网膜血管跟踪。

1.2.2.3 形态学方法

形态学操作是一类强大的数字图像处理工具，常用于提取复杂图像结构，构造有用的特征表达，描述如边界(Boundary)、骨架(Skeleton)和凸包(Convex Hull)等图像区域形状。形态学操作具有速度快、抗噪声能力强的优点，但是在处理高弯曲度地血管时，效果较差。Fraz et al.^[29]提出了一种形态学方法用于分割视网膜血管中心线和视网膜血管结构，该方法首先利用高斯滤波器地一阶导数识别血管中心线，然后通过多方向的形态学顶帽变换来实现血管的分割。

1.2.3 基于有监督机器学习与深度学习的方法

有监督机器学习方法的基本思路是将视网膜血管分割任务建模为一个像素级别的二元分类问题。通过利用带有眼科专家标定的金标准(Ground Truth, or Gold Stander)的数据样本对，训练一种判别模型，根据局部图像特征，对一个视网膜眼底图像像素是否属于血管像素做出判断。

1.2.3.1 基于特征工程和传统机器学习的方法

早期的有监督机器学习方法在视网膜血管分割任务上的应用，通常需要利用特征工程（Feature Engineering）构建有效的血管特征描述子，经过特征提取后，再利用各类分类器对眼底图像像素进行二元分类。

Ricci et al.^[26]提出了一种利用两个正交的线性特征描述子提取血管特征，并用支持向量机（Support Vector Machine, SVM）作为分类器，实现了视网膜血管分割模型的有监督训练。

Fraz et al.^[30,31]通过结合包括梯度矢量场方向分析（Orientation Analysis of Gradient Vector Field）、形态学变换（Morphological Transformation）、线性强度度量（Line Strength Measure）和 Gabor 滤波器响应等多种特征提取技术，利用如决策树（Decision Tree）等弱分类器进行集成学习（Ensemble Learning），实现了眼底图像的像素级别的血管分割。

Fan et al.^[32]通过多种形态学操作构建大量不同的特征通道，对每个待分类像素的领域图像补丁（Patch），基于结构相似性准则进行特征通道的筛选，构建像素级别的特征描述方式，最后以集成决策树构建随机森林（Random Forest）进行分类训练，从而对图像补丁特征和血管标定的关系进行建模，实现像素级别的视网膜血管分割。

Fan et al.^[33]提出了一种将贪婪逐层训练和整体微调相结合的消噪自动编码器（Denoising Auto-Encoder），实现了对图像补丁中心像素是否属于血管像素的二元分类。

Orlando et al.^[34]首次引入条件随机场（Conditional Random Field, CRF）用于视网膜血管分割，该方法利用一个结构化输出的支持向量机来训练全连接条件随机场。通过全连接条件随机场给两个待分类的眼底图像像素加入空间约束：相邻越近的两个像素被分类为同一类别的可能性越大。这种空间约束来提升血管分割算法的精度。

此外，在视网膜血管分割任务中应用人工神经网络时，最初也是用作有监督学习的分类器，对像素领域补丁的特征矢量进行二元分类。Marin et al.^[35]利用像素领域提取灰度值特征和矩不变性特征（Moment Invariant based Feature）来描述视网膜眼底图像像素，并通过有监督学习训练一个含三隐藏层的全连接前馈神经网络（Fully Connected Feedforward Neural Network）^①，利用 Sigmoid 函数激活的输出层对输入图像补丁中心像素属于血管的概率进行建模，最后利用阈值处理二值化该概率值得到每个像素的分类结果。

1.2.3.2 基于深度神经网络和深度学习的方法

基于传统的有监督机器学习算法实现视网膜血管分割的方法，通常涉及复杂的特

^① 全连接前馈神经网络也被称为多层感知机（Multi-Layer Perceptron, MLP）

征设计、提取和筛选，将原始数据转换成合适的特征向量，才能更有效被机器学习算法处理。然而，这些特征工程需要大量的领域相关专家知识，且人工设计的特征描述泛化能力有限。

近年来，通过对神经网络的结构和参数优化，取代传统特征工程实现特征设计和提取的深度学习方法，在自然语言处理（Natural Language Processing, NLP）、机器视觉（Computer Vision, CV）等领域的应用研究中取得了突破性进展^[36,37]。特别是在数字图像分割领域，卷积神经网络（Convolutional Neural Network, CNN）已经被普遍用于图像特征提取和图像分割。

在视网膜血管分割任务中，大量的研究工作应用了卷积神经网络。Wang et al.^[38]利用卷积神经网络提取多层次特征（Hierarchical Feature），随后通过随机森林算法对这些特征进行分类，实现视网膜血管分割。

除了利用卷积神经网络作为特征提取器外，Liskowski et al.^[39]将特征提取模型和分类器融合在一个卷积神经网络的训练过程中，实现了端对端（End-to-End）的视网膜血管分割。这种方法一方面能够加快卷积神经网络的训练，但另一方面，因为卷积神经网络中的全连接层可训练参数数量巨大，需要大量的不同的训练样本以减轻网络过拟合（Overfitting）。Liskowski et al.^[39]还提出了利用多种图像预处理方法对原始视网膜眼底图像训练数据库进行数据增广（Data Augmentation），并报导了数据增广能提升卷积神经网络对视网膜血管的分割效果。Maninis et al.^[40]提出了一种利用卷积神经网络对眼底图像中的视网膜血管和视盘同时分割的同一框架。

Fu et al.^[41]提出了结合卷积神经网络和条件随机场的视网膜血管分割方法，优化了分割效果。Luo et al.^[41]也提出了类似的方法，通过条件随机场加强了卷积神经网络对图像背景噪声和光照不均带来的灰度值骤变的鲁棒性。

与上述的卷积神经网络不同，不包含全连接层（Fully Connected Layer）和全局池化层（Global Average Pooling Layer）的全卷积神经网络（Fully Convolutional Network）更加适合图像分割任务。Dasgupta^[42]提出了一种利用全卷积神经网络的视网膜血管分割方法，和基于带有全连接层的卷积神经网络相比，该方法耗时更短，并且具有更加精确的分割效果。Mo et al.^[43]提出了一种基于深度监督（Deep Supervision）训练的全卷积神经网络，实现了高精度的视网膜血管分割。基于图像补丁中心点分类的建模方式，Alom et al.^[44]提出了通过结合残差连接和循环卷积来改进 UNet^[45]的网络架构，改善了视网膜血管分割效果。类似地，Xiao et al.^[45]提出了利用残差模块代替 UNet 中的子模块，构建了 Res-UNet 用于视网膜血管分割。

此外，近年来深度学习的研究热点，对抗生成网络（Generative Adversarial Network, GAN）也被用于视网膜血管分割。Lahiri et al.^[44]提出了一种基于对抗学习的视网膜血管分割方法。其中生成网络负责处理输入的眼底图像并生成候选的血管分割图，而另一

个判别网络则负责在输入的视网膜眼底图像和其血管分割图的样本对中，鉴别出哪些样本对是由金标准组成的样本对，哪些样本对是由生成网络输出的血管分割图组成的样本对。最终，对抗训练完成的生成网络会真正被用于视网膜血管分割任务。

1.3 主要研究内容

本文以视网膜眼底图像中的血管分割任务为核心问题，以深度学习及神经网络为主要技术手段，利用其强大的图像特征表达能力，构建基于分频卷积神经网络的视网膜血管分割算法，实现快速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，实现端对端的视网膜血管分割。针对视网膜血管分割问题的类别不均衡（Class Imbalance）特性，研究适用于全卷积神经网络的端对端训练方式。
- (2) 研究基于分频卷积神经网络的架构设计，设计并提出适用于编码器解码器架构的分频全卷积神经网络，研究实现基于分频全卷积神经网络的视网膜血管分割算法，实现多频率、多尺度特征学习，优化视网膜血管分割精度和速度。

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。本章节主要介绍视网膜血管分割的研究背景和意义，调研国内外最新的研究进展，并确立论文的研究内容和各章节安排。

第二章：基础技术研究。本章节主要介绍视网膜血管分割的公共数据库和算法评价指标，此外，还会普及深度学习及人工神经网络相关的基础技术和理论知识。

第三章：基于分频卷积神经网络的视网膜血管分割算法。在分频卷积操作的基础上，提出一种基于分频转置卷积的基础操作，用于分频全卷积神经网络中的分频卷积特征的解码。并利用分频卷积层和分频转置卷积层作为基础模块，提出一种基于编码器解码器架构的分频全卷积神经网络（Octave UNet）。针对视网膜血管分割任务中的类别不均衡问题，以加权交叉熵函数作为损耗函数训练分频卷积神经网络，实现端对端的高精度视网膜血管分割。

第四章：实验结果与分析。分别在 DRIVE、STARE、CHASE_DB1 和 HRF 数据库上对本文提出的方法进行测试，分析实验结果，并与其他最先进方法作对比。

第五章：总结与展望。本章节对全文的研究内容进行总结概括，提炼本文主要贡献，并对本课题的研究前景和未来工作进行展望。

第二章 基础技术研究

2.1 公共数据库与算法评价指标

2.1.1 彩色眼底图像公共数据库

为了能更直观、更公平地对比不同视网膜血管分割算法的效果，大量的研究工作都报告了所提出算法在一些彩色眼底图像公共数据库上获得的结果。这些公共数据库大多由医学机构组织眼科医师或相关专家收集整理，制作并公开发表。本文选择了DRIVE^[45]数据库，STARE^[46]数据库，CHASE_DB1^[47]数据库和HRF^[48]数据库共四个常用的彩色眼底图像公共数据库，这些数据库均包含了彩色眼底图像和对应的由专家手工标定的血管分割图金标准（Ground Truth）。

表 2.1 本文选用的彩色视网膜图像公共数据库介绍

	发表年份	图像分辨率	数据库规模
DRIVE	2004	565 * 584	训练集 20 张 测试集 20 张
STARE	2000	700 * 605	20 张
CHASE_DB1	2011	999 * 960	28 张
HRF	2011	3304 * 2336	正常 15 张 青光眼 15 张 糖尿病 15 张

2.1.2 视网膜血管分割算法评价指标

视网膜血管分割算法的评价通常需要将其输出血管分割图与专家标定的金标准进行比较衡量。因视网膜血管分割任务常被描述成像素级别的二元分类问题，即对眼底图像中每个像素点属于血管像素还是背景像素进行二元分类，其算法的衡量指标也和二元分类问题的衡量指标对应。

2.1.2.1 二元分类问题的混淆矩阵

针对视网膜血管分割任务，眼底图像中的像素点可以根据专家标定的金标准分为两类：正类像素点，即血管像素点，以及负类像素点，即背景像素点。

将待评价的视网膜血管分割算法的分割结果与对应的金标准相比较，可以得到如表 2.2 中的二元分类问题的混淆矩阵（Confusion Matrix）。

表 2.2 二元分类问题的混淆矩阵

	人工标定为正类	人工标定为负类
算法判别为正类	真阳性 (True Positive, TP)	假阳性 (False Positive, FP)
算法判别为负类	假阴性 (False Negative, FN)	真阴性 (True Negative, TN)

2.1.2.2 基于混淆矩阵的分割结果可视化

将视网膜血管分割算法的分割结果与对应的金标准相比较, 统计得到混淆矩阵后, 将分割结果中的不同类别用不同颜色像素标注, 可以得到如图 2.1 中的可视化效果。

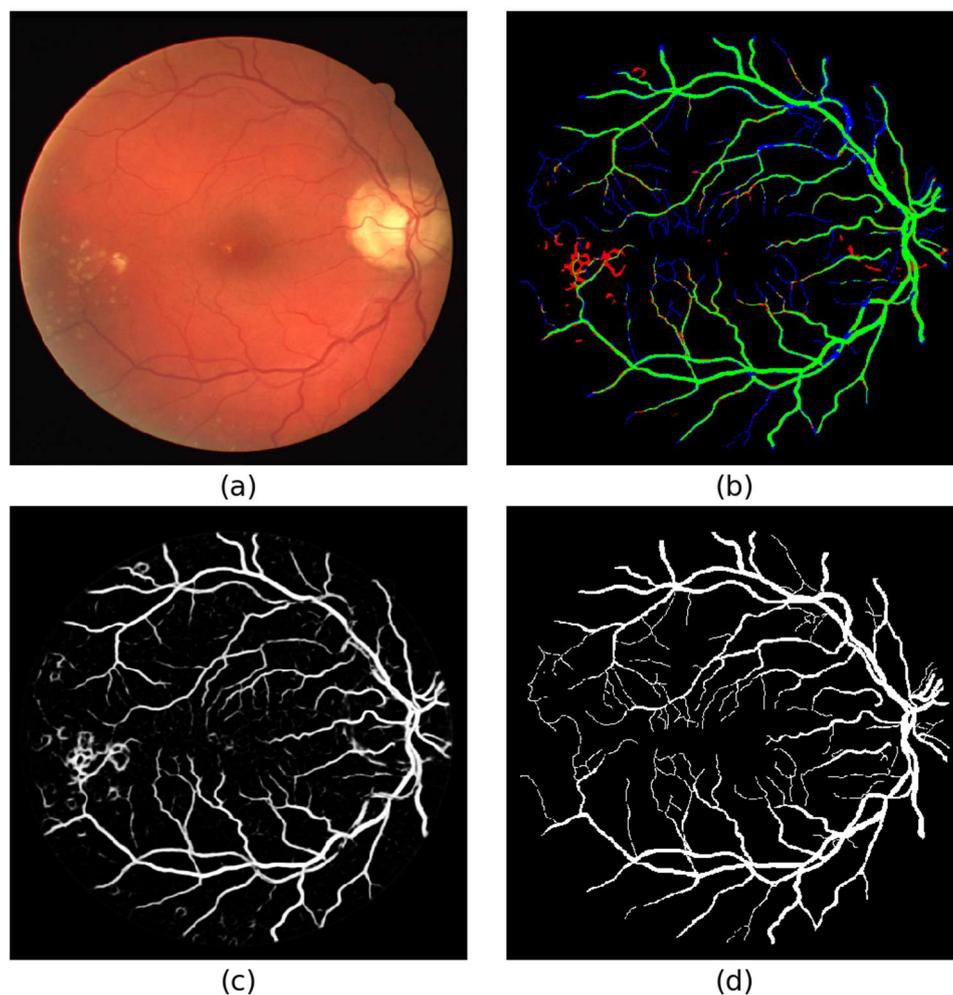


图 2.1 视网膜血管分割结果可视化示例: (a)视网膜眼底图像, (b)算法分割结果分析图, (c)算法输出概率图, (d)人工标注金标准。

在分割结果可视化图中, 绿色像素点为正确分割的视网膜血管像素点, 即真阳性。红色像素点为被误判为视网膜血管像素的背景像素点, 即假阳性。蓝色像素点为漏判的视网膜血管像素点, 即假阴性。最后, 黑色像素点为真阴性。从上图可以看出, 该算法将病灶周围的背景误判为血管像素点, 且漏判了大量细小的分支血管。

2.1.2.3 基于混淆矩阵的数值指标

在统计得到混淆矩阵后，可以根据 TP 、 FP 、 FN 和 TN 计算一系列的数值指标。下面针对视网膜血管任务中常用的评价指标做简单介绍：

(1) 正确率 (Accuracy, ACC)

正确率是分类问题中最常用的评价指标，反映了算法判别正确结果占所有样本的比例。

$$ACC = \frac{TP + TN}{TP + FP + FP + FN} \quad (2-1)$$

仅靠正确率往往不足够用于视网膜血管分割算法的对比。在眼底图像中，血管像素点的数量占总像素点数量的6% – 10%左右。即便一个对所有像素判断为负类点的简单分类规则的正确率也有90% – 94%左右。因此，对于类别不均衡的分类问题评价，除正确率外，往往还需要考虑其他指标。

(2) 敏感度 (Sensitivity, SE)

在机器学习领域，敏感度也称召回率 (Recall)，反映了在所有的正类样本中，算法能正确找回的正样本比例。

$$SE = \frac{TP}{TP + FN} \quad (2-2)$$

上面的例子中，对所有像素判断为负类点的简单分类规则的召回率始终为零。

(3) 特异性 (Specificity, SP)

类似的，特异性反映了在所有的负类样本中，算法能正确分类的负样本比例。在视网膜血管分割任务中，特异性受到假阳性判别数量的影响较大，模型判别的假血管点越多，特异性指标越低。

$$SP = \frac{TN}{TN + FP} \quad (2-3)$$

(4) F1 分数 (F1 Score)

F1 分数是敏感度和特异性的调和平均值，同时反映了敏感度和特异性指标高低。是类别不均衡的分类问题中的常用指标之一。

$$F1 = \frac{2 * TP}{2 * TP + FP + FN} \quad (2-4)$$

2.1.2.4 对阈值不敏感的指标

多数视网膜血管分割算法的分割都是对二值化阈值敏感的^①。这些算法先是输出一个像素级别的概率分割图，用于表达每个像素属于血管的置信度大小，然后通过阈值二值化得到最终的血管分割二值图。只有置信度高于阈值的像素点才会被判别为正类，

^① 如无特别说明，因为事先对模型输出的置信度概率图的分布没有任何先验知识，故本文中对二值化阈值敏感的指标计算，均在取阈值 $T = 0.5$ 的情况下计算所得。

而阈值设置得越高，越多像素点会被算法判别为正类，通常敏感度就会越高，但也越容易出现FP而使得特异性下降。对同一个分割算法，衡量阈值选取不同，包括正确率、敏感度、特异性和 F1 分数在内的基于混淆矩阵计算的数值指标也会不同。

当要综合考虑算法在不同阈值的表现时，就要引入对阈值不敏感的指标。比如图 2.2 中的 AUPRC (Area Under Precision-Recall Curve) 和 AUROC (Area Under Receiver Operating Characteristic curve, ROC curve)。这两种指标的计算都设计遍历预定的阈值梯度作图，然后计算所得曲线下的面积大小。

AUPRC 通过计算不同阈值下的 Precision 和 Recall，作如图 2.2(b)中的 Precision-Recall 曲线，并计算 Precision-Recall 曲线下的面积大小。其中：

$$Precision = \frac{TP}{TP + FP} \quad (2-5)$$

AUROC 则通过计算不同阈值下的 FPR (False Positive Rate) 和 TPR (True Positive Rate, Sensitivity, Recall)，作如图 2.2(a)中的 ROC 曲线，即 FPR-TPR 曲线，并计算 ROC 曲线下的面积大小。其中：

$$FPR = \frac{FP}{TP + FN} \quad (2-6)$$

本文仅选用了在视网膜血管分割领域其他文献中更常用的 AUROC 指标。

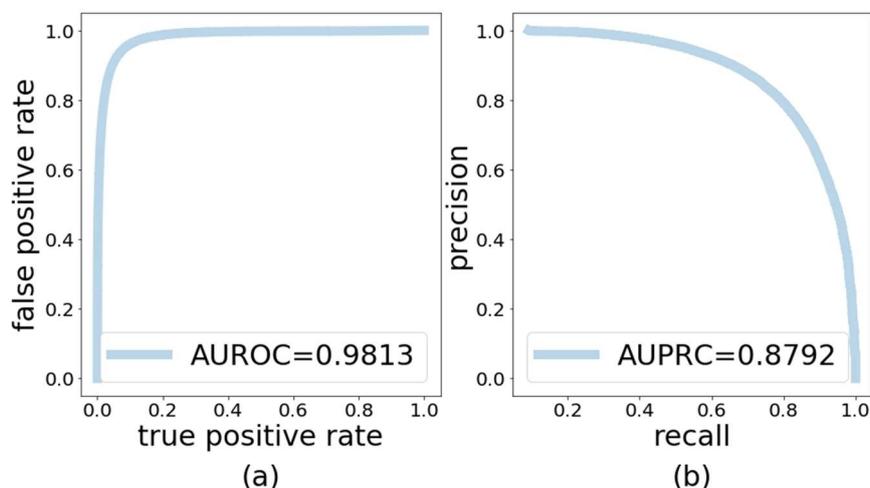


图 2.2 (a)为 ROC 曲线样例，(b)为 Precision-Recall 曲线样例

2.2 深度神经网络概述

近年来，以深度神经网络为核心的深度学习技术被广泛运用在自然语言处理 (Natural Language Processing)、计算机视觉 (Computer Vision)、医学图像分析 (Medical Image Analysis) 等各个方面。

在这些应用领域，大多数任务都依赖于将数据映射到合适的特征空间，比如，检测图像的饱和度 (Saturation) 的任务在映射到 HSV (Hue- Saturation-Value) 格式的图像

上进行, 会比在 RGB (Red-Green-Blue) 格式的图像上进行更合适。但是对于目标检测、图像分割、医学图像分析等复杂任务, 通过人手设计适合于当前复杂任务的特征描述方式非常困难, 耗时费力且效果一般。

神经网络正是用层级特征学习 (Hierarchical Feature Learning) 的方式解决了这一主要难题。其核心思路是利用低抽象层级的特征描述子来表达高抽象层级的特征描述子, 比如, 利用检测边缘的特征来表达检测拐角点的特征, 再利用检测拐角点的特征来表达检测三角形的特征等等。通过这样重复的多层组合, 代替传统特征工程, 实现适合于复杂任务的特征学习。神经网络以强大的特征学习能力著称, 具有巨大的潜能和重要的研究价值。

2.2.1 神经网络

最早的人工神经网络模型, 是一种对受生物神经科学启发的数学计算模型^[49], 人工神经网络一方面是对生物脑逆向工程的尝试, 另一方面也希望可以帮助理解生物脑的工作原理。虽然人工神经网络或受到了生物脑神经科学的启发和影响, 但随着机器学习学科和深度学习技术的发展, 人工神经网络技术的发展并不以为生物脑神经科学为刚性指导, 而更多的受如信息论、数字信号处理等工程学科的影响。

从本质上看, 现代神经网络^①是一种数学计算图模型 (Computational Graph Model), 它将神经元模型 (Neuron Model) 作为基本操作单元 (Operation Unit), 通过并联基本操作单元组成神经网络层 (Neural Network Layer)、以多层级的方式连接大量基本操作单元进行计算, 通过优化算法对各基本单元的可调参数进行优化, 实现复杂特征的学习和复杂非线性映射的拟合。

计算图模型的观点对神经网络的理解、分析和工程实现都非常有帮助。如同电子电路设计中, 将三极管等元器件抽象封装成门电路, 再将门电路抽象封装为芯片, 从而简化电路设计过程一样, 计算图模型将子计算节点抽象封装的过程也可以帮助理解, 设计和搭建神经网络模型。另一方面, 基于自动微分 (Automatic Differentiation) 的梯度计算图也方便了神经网络的优化算法的实现。本文将尝试结合计算图模型介绍神经网络基础理论。

2.2.1.1 基本组成单元

神经网络的基本组成单元是神经元模型, 主要包括: 对生物神经元脉冲放电时间建模的脉冲神经元模型 (Spiking Neuron Model)^[50]和模拟累积电位超过阈值后激活的 MP 神经元模型 (McCulloch-Pitts Neuron Model)^[49], 其中, 后者对现代神经网络

^① 如无特别说明, 本文中神经网络, 神经元等相关术语均指代人工神经网络, 人工神经元模型等, 而非其对应的生物学概念。

的各类基本单元有着重大而深远的影响。现代神经网络中，最常见的神经元模型是整流线性单元（Rectified Linear Unit, ReLU）。与 MP 神经元模型中用阶跃函数作为激活函数（Activation Function）不同，整流线性单元中的激活函数为：

$$g(x) = \max(0, x) \tag{2-7}$$

整流线性单元主要包含两组可训练参数：权重（Weights） W_j 和偏置（Bias） b_j ，以及 $\max(*)$ 非线性变换。整流线性单元首先利用权重和偏置对输入特征向量作仿射变换，然后通过非线性函数激活输出。具体而言，对于处于神经网络第 (l) 层的第 j 个整流线性单元，当输入特征向量 \mathbf{x} ，该整流线性单元的输出 \hat{y}_j 的计算公式为：

$$\hat{y}_j = \max\left(0, \sum_i w_{i,j}^{(l)} x_i + b_j\right) \tag{2-8}$$

整流线性单元的计算图模型如图 2.3 所示。

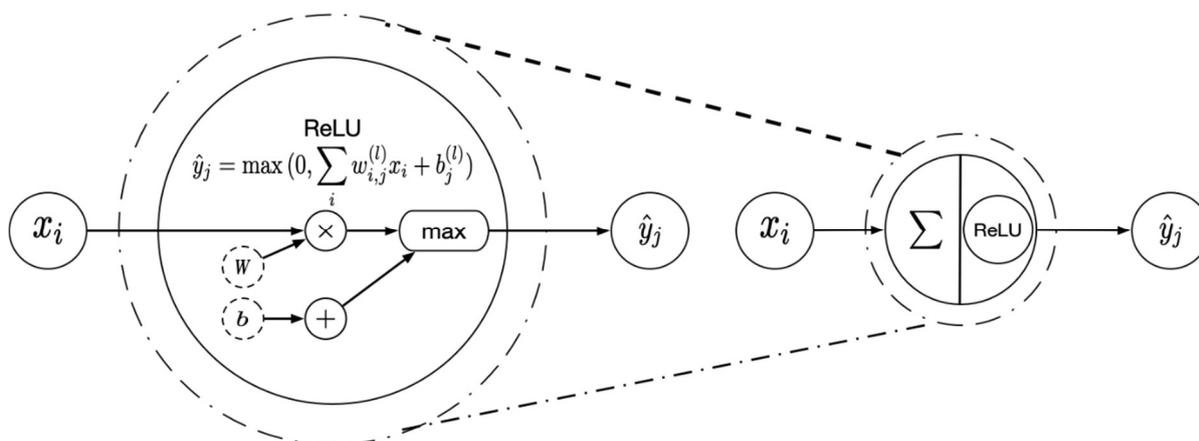


图 2.3 整流线性单元计算图模型

和大多数神经元模型一样，整流线性单元是一种非线性单元。在激活（Activated）状态时（即 $\sum_i w_{i,j}^{(l)} x_i + b_j > 0$ 时），它一阶导数保持较大的值且二阶导数为 0。这使得整流线性单元的梯度方向极其有用，且由于不用考虑二阶梯度的影响，使得其容易被仅利用一阶导数的优化算法训练。

但在整流线性单元进入饱和（Saturated）状态后（即 $\sum_i w_{i,j}^{(l)} x_i + b_j < 0$ 时），其输出为 0，各阶导数也为 0，基于梯度的优化算法没办法使饱和态整流线性单元重新激活，所以整流线性单元没有办法通过基于梯度的优化算法学习那些能使其进入饱和状态的样本。

针对整流线性神经元的饱和问题，Mass at el^[51]提出了渗漏整流线性单元（Leaky Rectified Linear Unit, Leaky ReLU），其核心思路是设计激活方程使得单元进入饱和状态后，其一阶导数是一个非零值。相似的，对于处于神经网络第 (l) 层的第 j 个渗漏整流线性单元，当输入特征向量 \mathbf{x} ，该渗漏整流线性单元的输出 \hat{y}_j 的计算公式为：

$$\hat{y}_j = \max\left(0, \sum_i w_{i,j}^{(l)} x_i + b_j^{(l)}\right) + \alpha_j \min\left(0, \sum_i w_{i,j}^{(l)} x_i + b_j^{(l)}\right) \quad (2-9)$$

渗漏整流线性单元将 α_j 设为一个较小的固定值，通过非零梯度的调整，渗漏整流线性单元有离开饱和状态重新激活。He et al^[52]将 α_j 作为可训练参数，在渗漏整流线性单元的基础上提出了参数化整流线性单元（Parametric Rectified Linear Unit, PReLU）。

除了整流线性单元及其系列扩展以外，针对二元分类问题，神经网络输出层单元常用Sigmoid单元，它们和整流线性单元的区别在于非线性函数的不同。Sigmoid单元以Sigmoid函数作为非线性激活函数。相似的，对于处于神经网络第 (l) 层的第 j 个Sigmoid单元，当输入特征向量 \mathbf{x} ，该单元的输出 \hat{y}_j 的计算公式为：

$$\hat{y}_j = \text{Sigmoid}\left(\sum_i w_{i,j}^{(l)} x_i + b_j\right) = \frac{1}{1 + e^{-(\sum_i w_{i,j}^{(l)} x_i + b_j)}} \quad (2-10)$$

Sigmoid单元的计算图模型如下图所示：

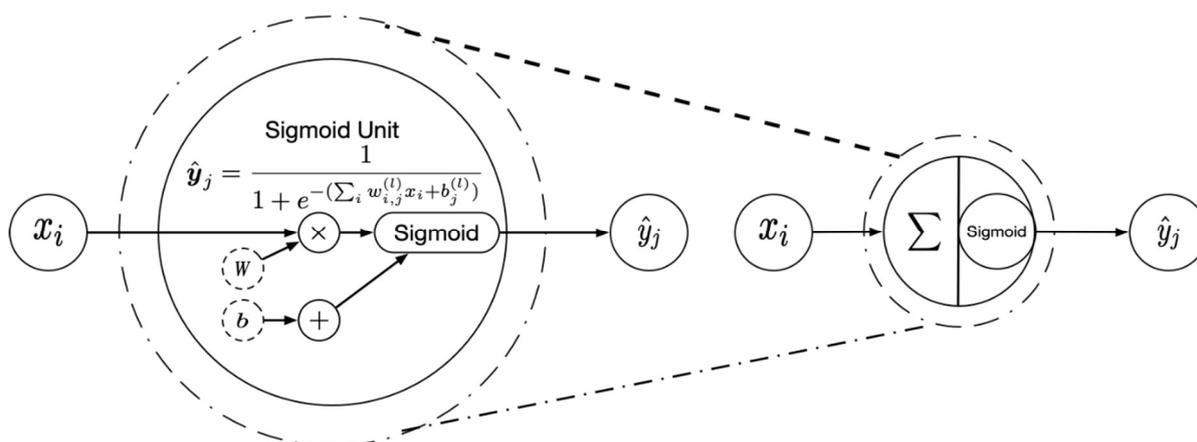


图 2.4 Sigmoid单元计算图模型

此外，针对不同的任务，神经网络输出层还会考虑使用不同的神经元模型，如针对多类别分类问题的Softmax神经元模型，针对回归问题的线性神经元模型等。

2.2.1.2 全连接前馈神经网络

生物脑神经科学实验发现：通过重新连接雪貂大脑中不同功能的区域，雪貂大脑能够利用负责听觉处理的区域学会处理视觉信号，这也暗示着生物大脑能够使用单一的算法解决不同任务^[53]。

这一发现启发了神经网络的架构（Architecture）设计：深度神经网络中大多包含大量相同或相似的基础操作单元。通过逐层地连接这些基本操作单元，逐步地利用简单特征表达构建复杂特征表达。

全连接前馈神经网络（Fully Connected Feedforward Neural Network），或称多层感

知机 (Multi-Layer Perceptron, MLP) 是经典的神经网络架构之一, 它的计算图模型是一个有向无环图。其网络架构设计思路主要为:

- (1) 利用单个神经元模型学习简单特征描述子。
- (2) 通过并联多个初始化参数不同的神经元模型组成网络层, 学习多种不同的特征描述方式。其中, 同层并联神经元的数量被称为神经网络的宽度。
- (3) 通过逐层连接各非线性的网络层^①, 利用简单特征描述子逐层构建复杂特征描述子, 通过端对端梯度优化参数实现特征学习。其中, 网络层的数量被称为神经网络的深度。
- (4) 神经元所在网络层越接近输入层, 其特征描述子的抽象层级也就越低。比如, 较底层隐藏层可能检测不同角度的边缘特征, 较高层的隐藏层检测边缘组成的拐角点特征等等。

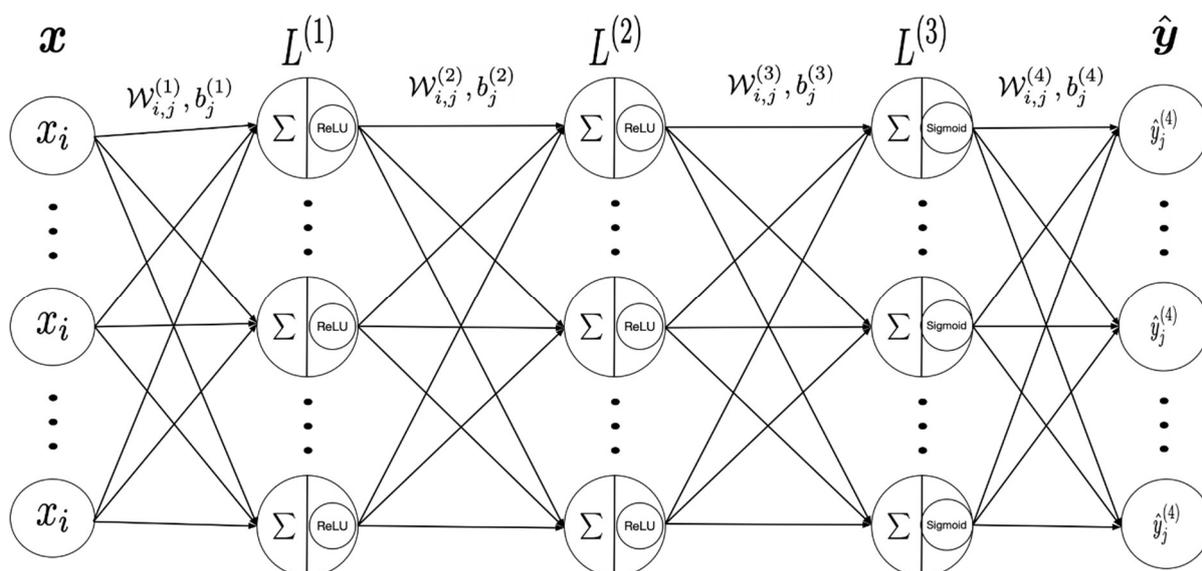


图 2.5 全连接前馈神经网络样例的计算图模型

在上图的样例中, 该全连接前馈神经网络包含 1 层输入层、3 层隐藏层和 1 层输出层。记不包含可训练参数的输入层为第 0 层, 但输入样本 x 后, 第 l 层第 j 个神经元输出的计算公式为:

$$\hat{y}_j = f^{(l)}(\mathbf{x}; W_j^{(l)}, b_j^{(l)}) = f^{(l)}\left(\sum_i w_{i,j}^{(l)} x_i + b_j^{(l)}\right) \quad (2-11)$$

其中 $f^{(l)}$ 为第 l 层的非线性激活函数, $w_{i,j}^{(l)}$ 为第 l 层权重, $b_j^{(l)}$ 为第 l 层偏置。则整个全连接前馈神经网络输出的计算公式为:

$$\hat{\mathbf{y}} = f^{(4)}(f^{(3)}(f^{(2)}(f^{(1)}(\mathbf{x}; W^{(1)}, b^{(1)}); W^{(2)}, b^{(2)}); W^{(3)}, b^{(3)}); W^{(4)}, b^{(4)}) \quad (2-12)$$

^① 值得注意的是, 简单层叠线性网络层的神经网络将退化成线性回归模型。

形如上式的计算图模型被称为是前馈（Feedforward）的，输入 \mathbf{x} 逐层流经各网络层的子计算图，最终达到输出 \mathbf{y} 的过程中，不包含任何反馈（Feedback）。

全连接前馈网络的计算图呈链式架构，方便了在对网络参数优化时，微分链式法则的运用，通过逐层求解目标方程对可训练参数的梯度，利用梯度下降法（Gradient Descent）可以实现端对端的网络参数优化。

但另一方面，参数过多的全连接的网络架构也会给前馈神经网络带来一些问题：

- (1) 需要训练的参数过多，深度前馈网络模型训练时间较长。
- (2) 网络模型表达能力过强，若没有进行合适的正则化（Regularization），容易发生拟合现象。
- (3) 参数过多使得运算复杂度高，运算速度较慢。
- (4) 可扩展性差，不能端对端地处理较大的数据样本，比如在图像处理应用中，全连接架构的前馈神经网络没有办法直接处理高分辨率的图片，只能先将高分辨率的图片裁剪成图像补丁，分别通过前馈神经网络处理获得输出后，在进行整张输出图的合成。这种做法不但计算存在冗余且速度慢，还可能引入人造地图像虚影。
- (5) 连接架构学习分布式特征存在冗余，通常深度神经网络对复杂的全局特征的学习是通过简单的局部特征逐层构建组合而成的，但是全连接前馈网络的各层神经元都能获取上一层所有神经元的输出的全局信息，这也意味着全连接的网络架构可能存在大量冗余连接权重。

2.2.2 卷积神经网络

在实际应用中的数据通常会具有一些空间结构特性，比如音频和文本数据可以看作是具有一维网格结构的数据，图像数据可以看作是二维网格结构的数据，视频数据则可以看作具有三维网格结构的数据。卷积神经网络（Convolution Neural Network, CNN）是一类特化于处理具有网格结构数据的神经网络架构。卷积神经网络一方面是生物闹神经学研究成功应用于人工神经网络架构设计的关键例子，另一方面是最早的表现良好且被用于解决重要商业问题的深度学习模型之一。卷积神经网络不但在深度学习技术发展历史中发挥了重要意义，而且依旧是当今深度学习技术的关键研究热点之一。

2.2.2.1 卷积操作

与前馈神经网络的全连接方式不同，卷积神经网络的卷积层（Convolution Layer）采用了类似卷积（Convolution）这一数学运算的运算规则，避免了因全连接方式下参数过多带来的种种问题。

以特化于处理图像数据的二维卷积神经网络为例子，卷积层包含了一组称为卷积核（Convolution Kernel）的权重参数和可选的偏置参数。和传统图像处理中的特征描述子相似，这组卷积核的学习到的参数分布不同决定了该卷积操作所能提取的特征性质不同。所以，通常一层卷积层会包含多个不同的卷积核，分别对应输出特征图（Feature Map）的不同通道，从而提取同抽象层级的不同特征。

再者，因为卷积核的参数是通过学习获得的，卷积核翻转（Flip）这一数学卷积操作的步骤在卷积神经网络中显得不再重要^①，所以大多数神经网络的实现都会将这一步骤省略，不再实现严格的数学卷积，而是实现互相关函数（Cross-Correlation）的运算。

在激活函数为 f ，卷积核步长为1，无边缘补值，空洞系数（Dilation）为1的二维卷积层^②中，记对应某一输出通道的卷积核为 W ，偏置为 b ，当输入二维网格数据 X 后，该通道输出的特征图计算公式为：

$$\hat{y}_{(i,j)} = f((W * X)_{(i,j)} + b) = f\left(\sum_m \sum_n W_{\left(m+\frac{k-1}{2}, n+\frac{k-1}{2}\right)}^T X_{(i+m, j+n)} + b\right) \quad (2-13)$$

该卷积操作样例的计算细节如下图所示：

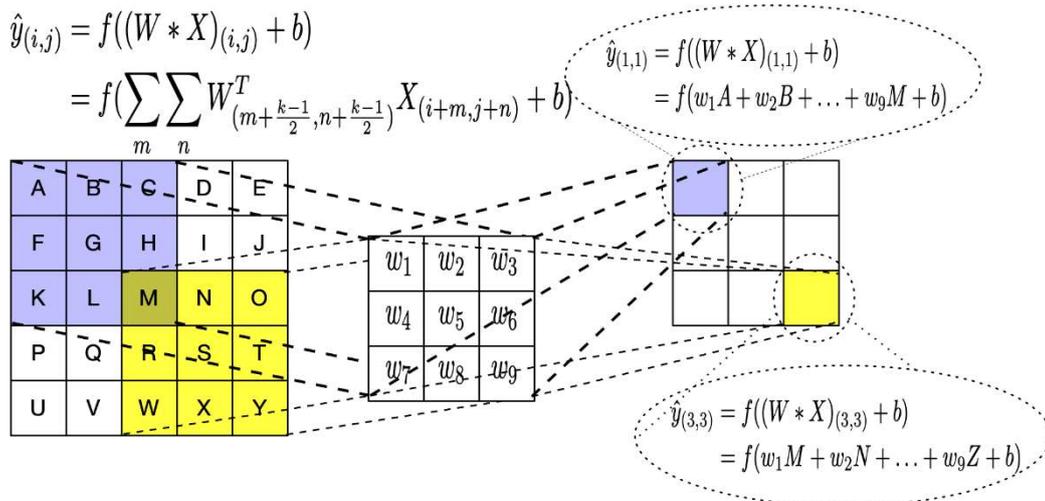


图 2.6 单通道卷积操作样例的计算细节，图中仅标注第一和最后一步的运算细节

通常一个卷积层会包含多个卷积核 $W = W_1, W_2, \dots, W_k$ 和对应的偏置 $\beta = b_1, b_2, \dots, b_k$ ，不同组卷积核和偏置将提取不同的特征，生成该组对应的特征图通道，最后构成该卷积层的 k 个通道的输出特征图。

与全连接神经网络层相比，特化于处理网格数据的卷积层使用了稀疏连接（Sparse Connectivity）的设计。输出特征图网格中的神经元与输入特征图网格中的神经元之间

^① 卷积核翻转使得卷积操作中，随着输入索引增大，卷积核索引变小。卷积核反转这一步骤保证了卷积操作的可交换性，该性质常用于涉及卷积操作的数学证明推导。

^② 但本文将按照机器学习领域文献的惯例，把实现互相关运算的网络层称为卷积层。

的连接是稀疏的，每个输出特征图网格中的神经元仅能接受其对应局部感受野（Receptive Field）中的输入，而局部感受野的位置是由神经元所在输出特征图中的位置，卷积核大小、稀疏系数和步长，以及边缘补值等因素共同决定的。稀疏连接与局部感受野的引入，一方面即满足了神经元提取局部特征的输入需求，另一方面也减少了卷积神经网络的可训练参数量，避免了因参数过多带来的种种问题。

此外，在传统数字图像处理中空域滤波（Spatial Filtering）中，因为输入数据具有空间结构，而我们并没有办法事先知道一种局部特征会出现在输入数据的哪个空间位置中，所以通常需要将该局部特征描述子对输入数据的所有领域进行遍历，从而提取不同空间位置的特征响应。相似地，卷积层也会利用同一个卷积核，对每一个局部感受野进行互相关运算，得到输出特征图中不同空间位置的特征响应值。这意味着位于相同特征图上的不同神经元输出的运算用到了相同的一组卷积核，这种设计被称为参数共享（Parameter Sharing）。参数共享的设计满足了在输入的不同空间位置提取特征的需求，也进一步减少了卷积神经网络的可训练参数量，加强了网络的可扩展性。

通过稀疏连接和参数共享的设计，卷积神经网络层具备了对特征平移的等变性（Equivariance），这意味着某张图片的卷积特征响应会随着该图的平移变换而平移，这种性质使得卷积特征图不但能对输入中包含哪些特征进行编码，还能通过特征图不同空间位置的特征响应对特征在输入的哪个空间位置进行编码，因此也更适合处理如图像等具有空间结构的数据。

2.2.2.2 池化操作

在卷积神经网络中，除卷积操作外，另一种常见的网络层是池化层（Pooling Layer）。池化层通常用于处理卷积层的输出特征图，池化层通常利用某一位置邻域内各值的统计量取代该位置的输出。其中，最大池化层（Max Pooling Layer）和平均池化（Average Pooling Layer）是最常见的两种池化层实例，它们分别用局部感受野区域中的最大值和统计值代替局部感受野中心的输出值。

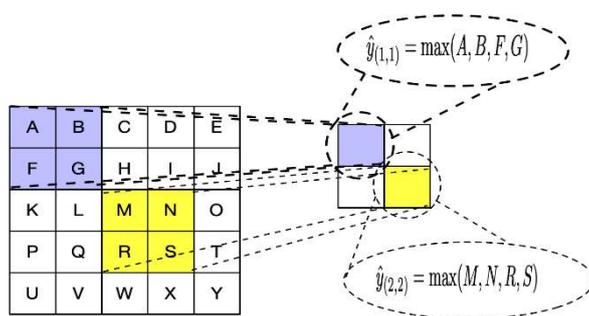


图 2.7 单通道最大池化操作样例，其中局部感受野为 2×2 网格，步长为 2

因为局部感受野内的输入值进行平移后，如最大值和平均值等统计量通常会保持不变，这也使得池化层具有对特征平移的不变性（Invariant）。池化层一方面通过减少特征图的空间尺度从而减少了后续网络的运算量，但另一方面也因此而丢失了部分位置信息。

2.2.2.3 基础卷积神经网络架构

早期的卷积神经网络多被用于图像分类应用中，比如 LeNet^[54]和 AlexNet^[55]，它们都是较为浅层的基础卷积神经网络架构，且都在抽象层级低（越靠近输入）的卷积层使用了尺度较大的卷积核，而在抽象层级高的卷积层使用了尺度较小的卷积核。LeNet 和 AlexNet 的架构较为基础，对现代卷积神经网络的发展影响深远，其中 LeNet 是最早解决重要商业应用的卷积神经网络之一，而 AlexNet 提出了包括用整流线性单元作为激活函数、Dropout、数据增广和 GPU 加速训练等贡献。

此后，各类更深层的卷积神经网络被提出。比如 VGGNet^[56]提出采用固定尺度的小卷积核来实现深度卷积网络。卷积神经网络除了不断往更深层的方向发展以外，更复杂的基本网络层也不断被提出。GoogLeNet^[57]提出了名为 Inception Block 的网络层作为深度卷积神经网络的基本构造单元，Inception Block 中包含了不同大小卷积核的卷积操作和池化操作，从而实现了在同一层级提取不同尺度的特征。ResNet^[58]提出了利用 Residual Block 进行残差学习，成功训练了更深层的卷积神经网络网络。而 DenseNet^[59]提出了 Dense Block 加强了各抽象层级特征图的利，在 Residual Block 基础上作了进一步扩展。

2.2.3 全卷积神经网络

由卷积层和池化层反复，并以全连接层作为输出层的深度卷积网络通常是针对图像分类（Image Classification）类任务设计的架构。将这些包含全连接层的神经网络架构用于图像分割任务时，一种常用的做法是首先提取图像补丁，然后通过这类网络对

图像补丁的中心点类别进行分类。这是因为全连接层的权重张量的维度与网络输入相关，所以一个网络的输入数据必须保持相同维度。这种做法效率低下，计算存在冗余，且容易引入人造虚影，所以以全连接层作为输出层的神经网络架构并不适合用于图像分割任务。

针对图像语义分割（Semantic Segmentation）和医学图像分割任务，Long et al.^[60]提出了全卷积神经网络（Fully Convolutional Neural Network, FCN）。全卷积神经网络提出以卷积层代替卷积神经网络中所有的全连接层，以摆脱固定网络输入大小的约束，使利用同一个全卷积神经网络对任意输入大小进行处理成为可能。因为全卷积神经网络不需要提取并处理重叠的图像补丁，计算冗余也得以去除，加快了图像分割的速度。此后，几乎所有用于图像分割的卷积神经网络都摒弃了全连接层。

此外，全卷积神经网络提出了利用转置卷积层（Transposed Convolution Layer）对卷积特征图进行解码，从而产生与输出大小等大的分割图。与简单地利用双线性插值（Bilinear Interpolation）将特征图进行空间尺度的进行扩展做法不同，转置卷积层是一种可训练的上采样方法。

2.2.3.1 转置卷积操作

与简单地利用双线性插值等上采样规则不同，转置卷积层是一种可训练的上采样运算。与卷积类似操作，转置卷积层也包含了一组可训练的卷积核和偏置参数。并且都采用局部感受野和稀疏连接。

在激活函数为 f ，卷积核步长为1，无边缘补值，空洞系数（Dilation）为1的二维转置卷积层中，记对应某一输出通道的卷积核为 W ，偏置为 b ，当输入二维网格数据 X 后，该通道输出的特征图计算公式为：

$$\hat{y}_{(i,j)} = f \left(\sum_m \sum_n X_{\left(m+\frac{k-1}{2}, n+\frac{k-1}{2}\right)}^T W_{(i+m, j+n)} + b \right) \quad (2-14)$$

该转置卷积操作样例的计算细节如下图所示：

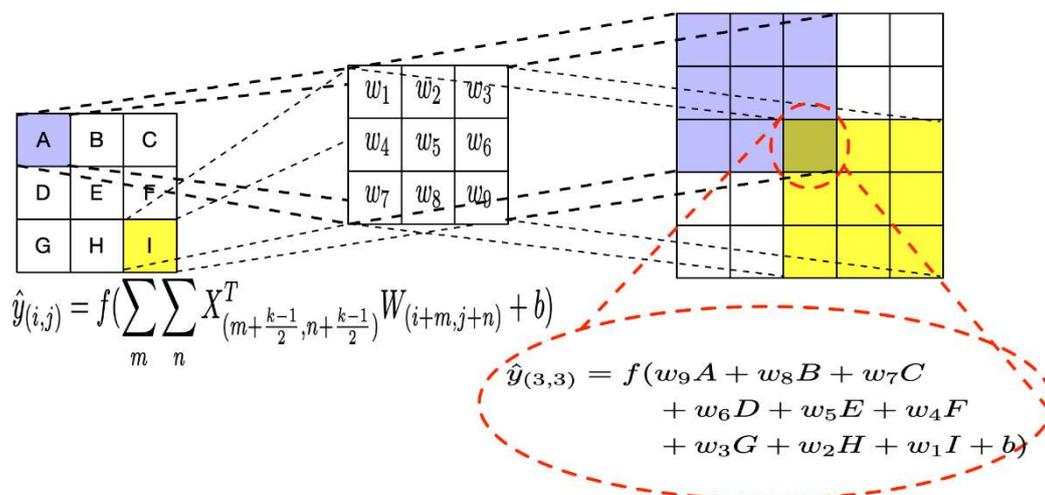


图 2.8 单通道转置卷积操作样例的运算样例

和卷积操作类似，通常一个转置卷积层也会包含多个卷积核 $W = W_1, W_2, \dots, W_k$ 和对应的偏置 $\beta = b_1, b_2, \dots, b_k$ ，不同组卷积核和偏置将解码得到不同的特征，生成该组对应的特征图通道，最后构成该转置卷积层的 k 个通道的输出特征图。

2.2.4 基于编码器解码器架构的深度卷积网络

在全卷积神经网络架构解决了网络输入固定的约束后，卷积神经网络用于图像分割任务还存在一个问题：特征图位置信息的丢失。通常卷积层和池化层输出的特征图尺度会被渐渐缩小，这虽然给神经元带来更大的局部感受野，但同时也会使得特征的位置信息被渐渐丢失：处于网络越深位置的特征图抽象层级更高，更容易表达复杂的特征，但是因为特征图的空间尺度越小，也更容易表达该特征在输入的哪个位置。显然，这种特性对于需要进行像素级别精准分类的自然图像语义分割和医学图像分割任务而言是非常不利的。

针对特征图空间信息丢失的问题，Ronneberger et al.^[45]提出了名为 UNet 的基于编码器解码器架构（Encoder-Decoder Architecture）的卷积神经网络。UNet 主要包含两大部分：特征编码器和特征解码器。

特征编码器是一个基础神经网络架构的子计算图，它通过层叠卷积层和池化层作为特征描述子，逐层编码复杂特征。在这个过程中，网络层越深，其输出特征图的特征抽象层级越高，通常丢失的特征位置信息也越多，具体样例可参考图 3.7 中本文提出的方法所提取的各层特征响应图。

只利用抽象层级高而位置信息匮乏的特征来生成分割图往往是效果不佳的，而 UNet 则提出了利用转置卷积层、卷积层和池化层组成多层解码器架构。和特征编码过

程类似的，UNet 逐层地对不同抽象层级地特征进行解码，具体样例可参考图 3.8 中本文提出的方法所提取的各层特征响应图。在解码过程中，通过跳连（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，以帮助神经网络恢复特征的位置信息，优化分割效果。

以 UNet 为代表的基于编码器解码器架构的深度学习神经网络已经成为医学图像处理任务中最常用的技术手段和基准模型。比如，Alom et al^[44]提出了通过结合残差连接和循环卷积来改进 UNet 的网络架构，改善了基于图像补丁中心点分类的视网膜血管分割效果。类似地，Xiao et al.^[45]提出了利用残差模块代替 UNet 中的子模块，构建了 Res-UNet 用于视网膜血管分割。

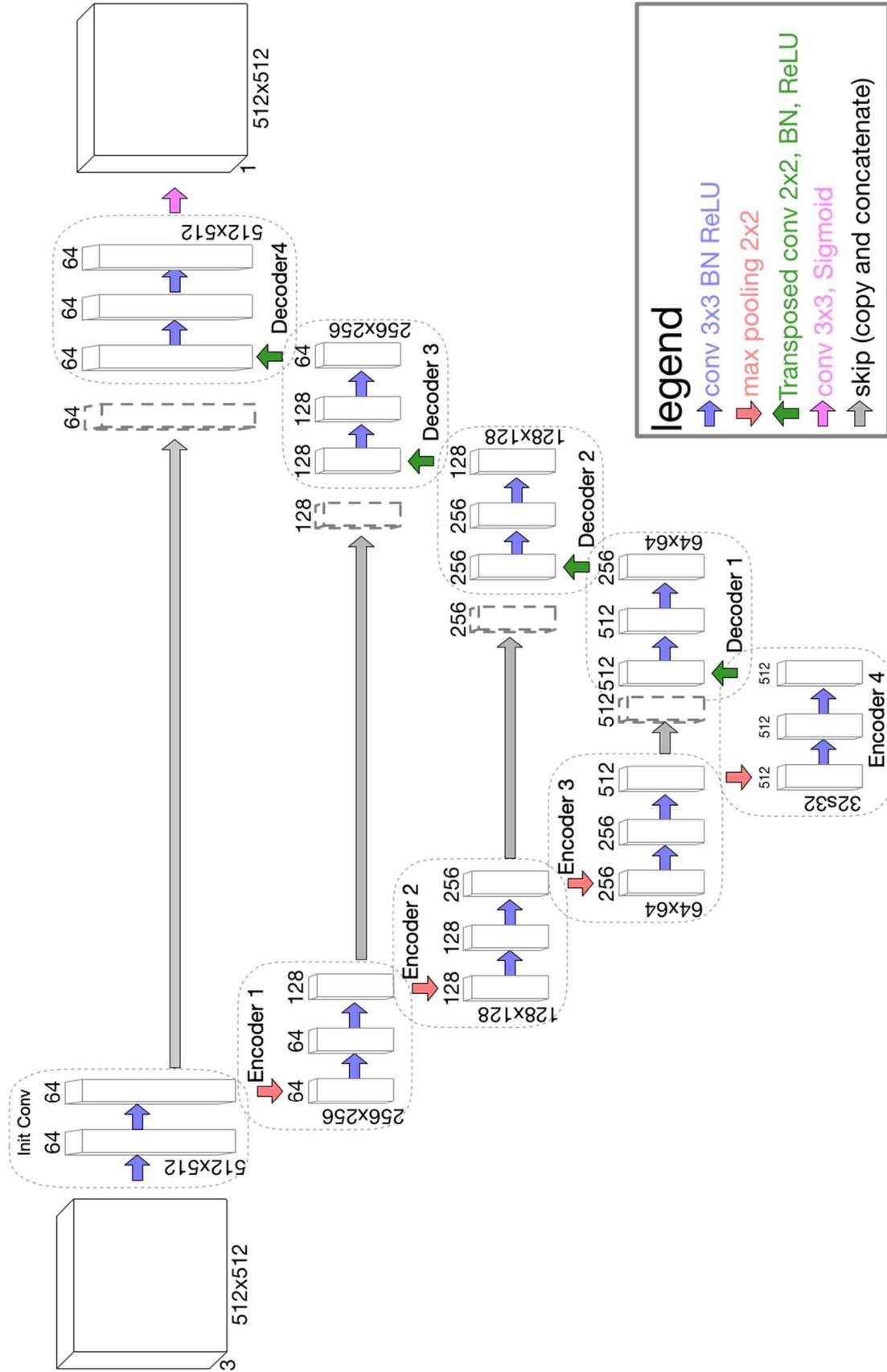


图 2.9 UNet 网络架构样例

第三章 基于分频卷积网络的视网膜血管分割算法

本文针对视网膜血管分割任务中的血管结构形态学特征差异巨大，血管粗细变化动态范围大的任务难点，结合通过高低分频的卷积神经网络特征分别提取细微血管轮廓和视网膜血管主干结构的思路，在分频卷积神经网络的基础上进行扩展，提出了基于分频转置卷积操作的解码器，搭建了名为 Octave UNet 的基于编码器解码器架构的分频卷积神经网络，实现了 Octave UNet 的端对端训练和高精度的视网膜血管自动分割，获得了与最前沿方法可比的分割效果。

3.1 分频卷积神经网络

3.1.1 图像金字塔特征与低频图像空域中的冗余

自然图像可以被分解为高频分量图和低频分量图。其中高频分量图主要包含图像局部的边缘和细节纹理信息，在图像空域 (Spatial Dimension) 中像素值的变化较剧烈。而低频分量图则包含了图像整体的结构信息，在图像空域中像素值变化较为平缓。

低频分量图在空域常常是存在冗余的，适当压缩较模糊的低频分量图像后丢失的信息较少。在如低通图像金字塔表征 (Lowpass Image Pyramid Representation) 和 SIFT 特征描述子 (Scale-Invariant Feature Transform)^[62] 中高斯金字塔等尺度空间 (Scale Space) 的构建过程中，金字塔中每层 Octave^① 都包含了一张或一系列通过高斯滤波器模糊和下采样 (Down Sampling) 获得的相同尺度的不同图像。而相邻两层 Octave 中的图像尺度相差一半，其空域频率也相差一半。

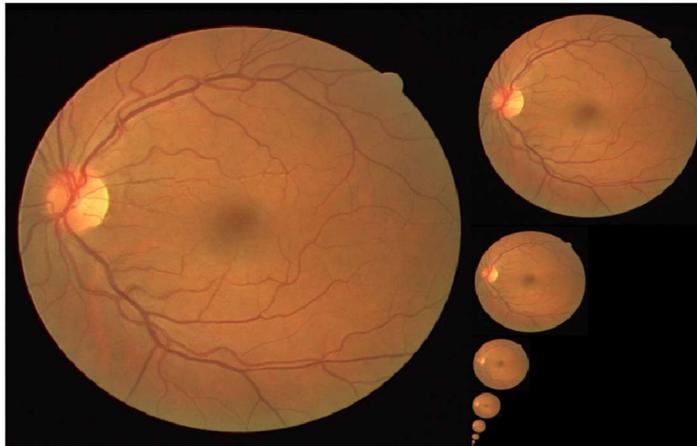


图 3.1 眼底图像金字塔表征样例，其中每个 Octave 只包含一张低频分量图

^① Octave 原指相邻两个音乐符号间的频率间隔。相邻两个音符中，低频音符的频率是高频音符的一半。而在机器学习领域，SIFT 特征描述子^[62]用 Octave 指代尺度空间中的不同频率阶层。

3.1.2 高低频卷积特征

从图 3.2 中可以看出视网膜血管图像的高频分量主要描述血管主干等整体结构信息，而低频分量则描述边缘和细节信息。如果能通过卷积神经网络逐层地同时提取高频和低频卷积特征图，将有可能更好地提取血管主干和细微血管细节地特征。

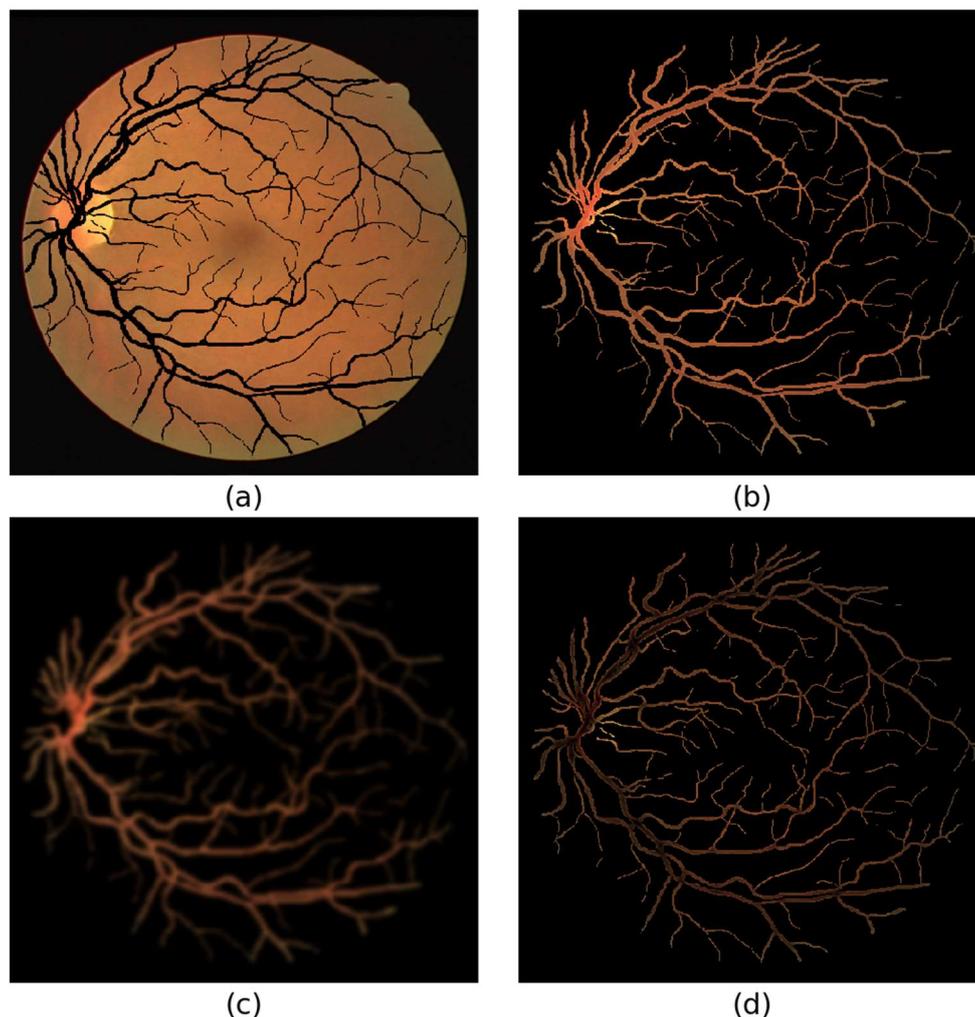


图 3.2 利用给人工标注分离的眼底背景图像(a)和血管图像(b)，
(c)为血管图像(b)的低频分量，(d)为血管图像(b)的高频分量

Chen et al.^[63]提出了卷积神经网络中的特征图也可分解为高频分量和低频分量的设想，并将普通卷积特征图的不同通道划分为：

- (1) 描述图像整体结构的低频通道 $X^L \in \mathbb{R}^{\alpha c \times \frac{h}{2} \times \frac{w}{2}}$ 。
- (2) 描述局部细节特征的高频通道 $X^H \in \mathbb{R}^{(1-\alpha)c \times h \times w}$ 。

其中，高低频共有 c 个通道， $\alpha \in [0, 1]$ 是用于控制低频和高频通道数比例的参数。低频通道特征图的空间分辨率 $0.5h \times 0.5w$ 是高频通道分辨率 $h \times w$ 的一半。

3.1.3 分频卷积操作

从输入图像中获取第一层高低分频卷积特征并不困难：例如利用不同步长的普通分组卷积（Grouped Convolution）、或者利用相同步长的普通卷积分别对原输入图像和下采样后的输入图像进行处理。

但是上述的卷积操作均不能对高低分频的特征图进行处理，在普通卷积层输出的卷积特征图中，不同通道的特征图均由相同的卷积核大小、相同的卷积步长、相同的稀疏系数和相同的边缘补值方式的互相关运算后获得，所以不同通道的特征图都具有相同的空域分辨率。将卷积层输出的特征图张量记为 $X \in \mathbb{R}^{c \times h \times w}$ ，那么其中 c 个特征图通道的空域分辨率均为 $h \times w$ 。为了设计一种能够高效处理高低分频卷积特征的基础运算，Chen et al.^[63]提出了分频卷积（Octave Convolution, OctConv）^①。

记输入分频特征图为 $X = \{X^L, X^H\}$ ，输出分频特征图为 $\hat{Y} = \{\hat{Y}^L, \hat{Y}^H\}$ ，分频卷积主要包含两大部分：

(1) 频率内的卷积处理：

$$\hat{Y}^{L \rightarrow L} = \text{conv}(X^L; W^{L \rightarrow L}) \quad (3-1)$$

$$\hat{Y}^{H \rightarrow H} = \text{conv}(X^H; W^{H \rightarrow H}) \quad (3-2)$$

其中 $\text{conv}(\ast)$ 为普通卷积运算， $W^{L \rightarrow L}$ 和 $W^{H \rightarrow H}$ 分别为描述低频映射到低频的卷积核和描述高频映射到高频的卷积核。

(2) 频率间的信息交互：

$$\hat{Y}^{L \rightarrow H} = \text{upsample}(\text{conv}(X^L; W^{L \rightarrow H}), 2) \quad (3-3)$$

$$\hat{Y}^{H \rightarrow L} = \text{conv}(\text{pool}(X^H, 2); W^{H \rightarrow L}) \quad (3-4)$$

其中 $\text{upsample}(\ast, 2)$ 为双线性插值的双倍上采样，上采样保证了从低频映射到高频的映射特征具有和低频特征一样的空域尺度。 $\text{pool}(\ast, 2)$ 为平均池化的双倍下采样，类似的，下采样保证了从高频映射到低频的映射特征具有和低频特征一样的空域尺度。从而保证了频率间信息交互。

分频卷积最终输出的分频卷积特征的高低分量均由其对应的频率内卷积特征和频率间交互特征组成，记分频卷积层的非线性激活函数为 $f(\ast)$ ，那么高频特征图和低频特征图中 (i, j) 处的特征响应 $\hat{Y}_{i,j}^H$ 和 $\hat{Y}_{i,j}^L$ 的计算公式分别为：

$$\begin{aligned} \hat{Y}_{i,j}^H &= \hat{Y}^{H \rightarrow H} + \hat{Y}^{L \rightarrow H} \\ &= f\left(\sum_m \sum_n W_{\left(m+\frac{k-1}{2}, n+\frac{k-1}{2}\right)}^{H \rightarrow H} X_{(i+m, j+n)}^H + b\right) \\ &\quad + f\left(\sum_m \sum_n W_{\left(m+\frac{k-1}{2}, n+\frac{k-1}{2}\right)}^{L \rightarrow H} X_{(2*i+0.5+m, 2*j+0.5+n)}^L + b\right) \end{aligned} \quad (3-5)$$

^① 本文选择将 Octave Convolution 意译为分频卷积。

$$\begin{aligned}
 \hat{Y}_{i,j}^L &= \hat{Y}^{L \rightarrow L} + \hat{Y}^{H \rightarrow L} \\
 &= f \left(\sum_m \sum_n W_{\left(m+\frac{k-1}{2}, n+\frac{k-1}{2}\right)}^{L \rightarrow L} X_{(i+m, j+n)}^L + b \right) \\
 &\quad + f \left(\sum_m \sum_n W_{\left(m+\frac{k-1}{2}, n+\frac{k-1}{2}\right)}^{H \rightarrow L} X_{\left(\lfloor \frac{i}{2} \rfloor + m, \lfloor \frac{j}{2} \rfloor + n\right)}^H + b \right)
 \end{aligned} \tag{3-6}$$

其中 $\lfloor * \rfloor$ 为取整运算，卷积核大小 k 通常设计为奇数，满足权重矩阵下标为整约束的 m 和 n 的取值决定了以 $X_{(i,j)}$ 为中心像素点的局部感受野在输入 X 上的范围。

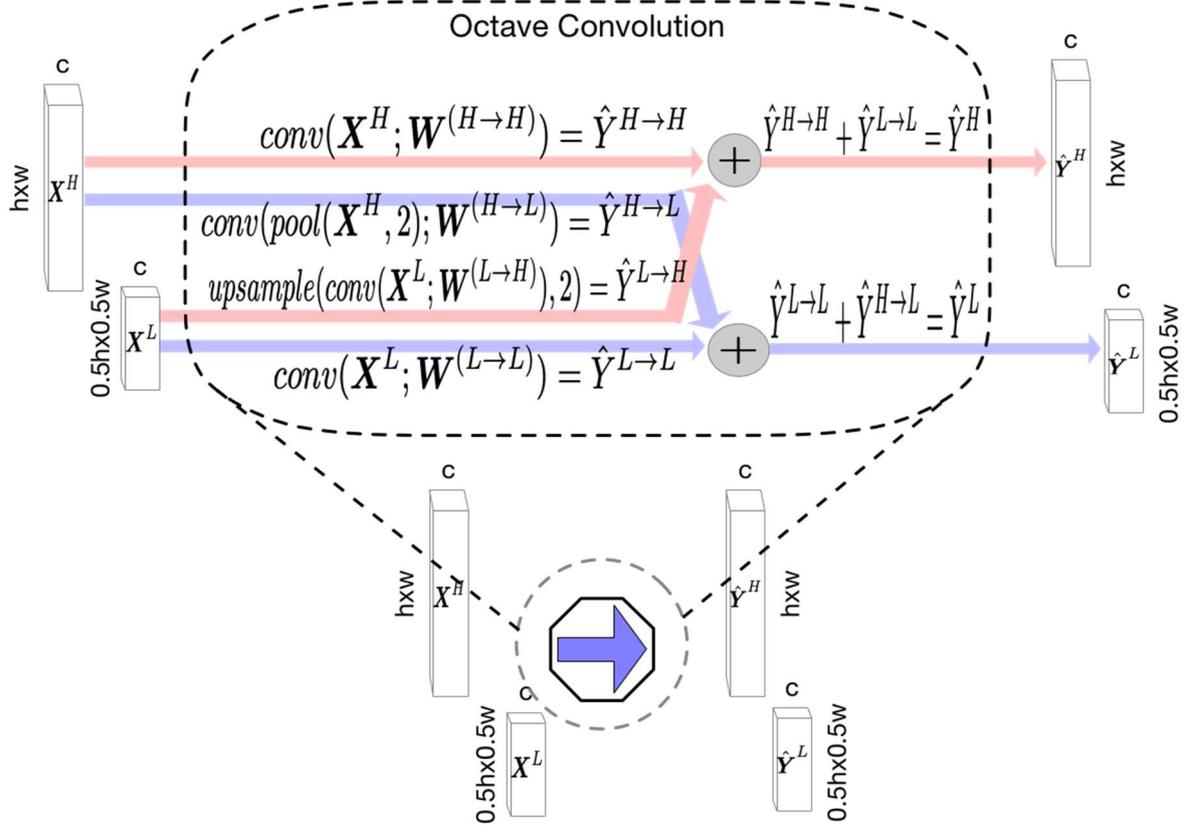


图 3.3 分频卷积计算图

3.2 基于编码器解码器架构的分频卷积网络

分频卷积运算满足了特征编码器中提取高低分频卷积特征的需求。但是，针对包括视网膜血管分割任务在内的图像语义分割和医学图像分割等任务而言，通常还需要一种解码器架构用于对编码器提取到的各层特征进行解码，恢复特征图空域尺度从而实现像素级别的分割效果。

3.2.1 分频转置卷积层

针对分频卷积编码器提取到的分频卷积特征图，本文提出了一种分频转置卷积运算（Octave Transposed Convolution）用于对分频卷积特征图进行高效解码。

与双线性插值等方法不同，转置卷积运算（Transposed Convolution, TranspConv）是一种参数可训练的上采样方式。但是与普通卷积操作类似，普通的转置卷积操作也不能对高低频的特征图进行处理。将转置卷积输出的特征图张量记为 $X \in \mathbb{R}^{c \times h \times w}$ ，那么其中 c 个特征图通道的空域分辨率均为 $h \times w$ 。为了设计一种能够高效解码高低频卷积特征的基础运算，一种简单的设计是频率对应的转置卷积（Frequency-wise Transposed Convolution），即利用多个普通的转置卷积运算分别对不同频率的特征通道进行处理，将不同频率通道分别解码，映射到对应的特征通道。具体而言，记输入分频特征为 $X = \{X^L, X^H\}$ ，输出分频特征为 $\hat{Y} = \{\hat{Y}^L, \hat{Y}^H\}$ ，则有：

$$\hat{Y}^L = \text{transp_conv}(X^L, 2; W^{L \rightarrow L}) \quad (3-7)$$

$$\hat{Y}^H = \text{transp_conv}(X^H, 2; W^{H \rightarrow H}) \quad (3-8)$$

其中 $\text{transp_conv}(*, 2; W)$ 是以 W 为卷积核，步长为2的普通转置卷积运算，该运算会将特征图的空间尺度扩展为原来的2倍。 $W^{L \rightarrow L}$ 为负责将低频解码到低频的卷积核，而 $W^{H \rightarrow H}$ 则为负责将高频解码到高频的卷积核。

这种简单设计能满足输入高低频卷积特征图，并解码输出空间尺度更大的高低频卷积特征图的需求。但是因为缺乏了频率间的信息交互，无法实现利用频率间信息协作解码。比如血管主干结构的低频特征信息或许可以帮助找到细微血管的局部细节，而细微血管分支信息则可以避免血管中心反光等干扰，帮助血管主干结构的正确分割。

本文在分频卷积运算基础上扩展，并提出了一种分频转置卷积（Octave Transposed Convolution, OctTranspConv）。记输入分频特征为 $X = \{X^L, X^H\}$ ，输出分频特征为 $\hat{Y} = \{\hat{Y}^L, \hat{Y}^H\}$ ，分频转置卷积主要包含两大部分：

(1) 频率内的解码处理：

$$\hat{Y}^{L \rightarrow L} = \text{transp_conv}(X^L, 2; W^{L \rightarrow L}) \quad (3-9)$$

$$\hat{Y}^{H \rightarrow H} = \text{transp_conv}(X^H, 2; W^{H \rightarrow H}) \quad (3-10)$$

与频率对应的转置卷积类似，其中 $\text{transp_conv}(*, 2; W)$ 是以 W 为卷积核，步长为2的普通转置卷积运算，该运算会将特征图的空间尺度扩展为原来的2倍。 $W^{L \rightarrow L}$ 为负责将低频解码到低频的卷积核，而 $W^{H \rightarrow H}$ 则为负责将高频解码到高频的卷积核。

(2) 频率间的信息交互：

$$\hat{Y}^{L \rightarrow H} = \text{upsample}(\text{transp_conv}(X^L, 2; W^{L \rightarrow H}), 2) \quad (3-11)$$

$$\hat{Y}^{H \rightarrow L} = \text{transp_conv}(\text{pool}(X^H, 2), 2; W^{H \rightarrow L}) \quad (3-12)$$

其中 $\text{upsample}(*, 2)$ 为双线性插值的双倍上采样，上采样保证了从低频映射到高频的映射特征具有和低频特征一样的空域尺度。 $\text{pool}(*, 2)$ 为平均池化的双倍下采样，类

似的，下采样保证了从高频映射到低频的映射特征具有和低频特征一样的空域尺度。从而保证了频率间信息交互。

分频卷积最终输出的分频卷积特征的高低分量均由其对应的频率内卷积特征和频率间交互特征组成，记分频卷积层的非线性激活函数为 $f(\cdot)$ ，那么高频特征图和低频特征图中 (i, j) 处的特征响应 $\hat{Y}_{i,j}^H$ 和 $\hat{Y}_{i,j}^L$ 的计算公式分别为：

$$\begin{aligned} \hat{Y}_{i,j}^H &= \hat{Y}^{H \rightarrow H} + \hat{Y}^{L \rightarrow H} \\ &= f \left(\sum_m \sum_n X_{(m+\frac{k-1}{2}, n+\frac{k-1}{2})}^H W_{(i+m, j+n)}^{H \rightarrow H} + b \right) \\ &\quad + f \left(\sum_m \sum_n X_{(m+\frac{k-1}{2}, n+\frac{k-1}{2})}^L W_{(2*i+0.5+m, 2*j+0.5+n)}^{L \rightarrow H} + b \right) \end{aligned} \quad (3-13)$$

$$\begin{aligned} \hat{Y}_{i,j}^L &= \hat{Y}^{L \rightarrow L} + \hat{Y}^{H \rightarrow L} \\ &= f \left(\sum_m \sum_n X_{(m+\frac{k-1}{2}, n+\frac{k-1}{2})}^L W_{(i+m, j+n)}^{L \rightarrow L} + b \right) \\ &\quad + f \left(\sum_m \sum_n X_{(m+\frac{k-1}{2}, n+\frac{k-1}{2})}^H W_{(\lfloor \frac{i}{2} \rfloor + m, \lfloor \frac{j}{2} \rfloor + n)}^{H \rightarrow L} + b \right) \end{aligned} \quad (3-14)$$

其中 $\lfloor * \rfloor$ 为取整运算，卷积核大小 k 通常设计为奇数，满足权重矩阵和输入张量的下标为整数且在有效范围内的 m 和 n 取值决定了以 $\hat{Y}_{(i,j)}$ 为中心像素点的邻域在输出张量 \hat{Y} 上的范围。

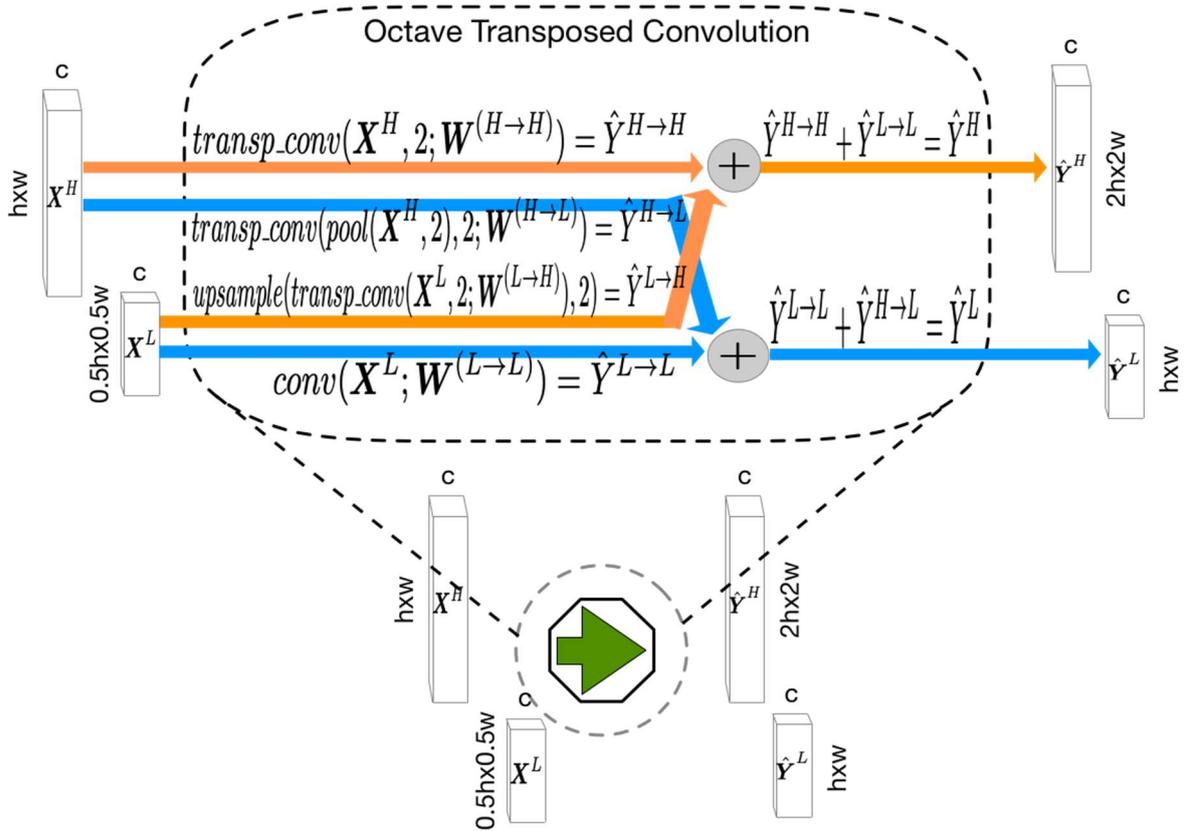


图 3.4 分频转置卷积计算图

3.2.2 Octave UNet 架构设计

Octave UNet 网络架构主要包括两个主要组成部分：由多个特征编码器（Encoder Block）组成的编码通路（Encoding Path）和由多个特征解码器（Decoder Block）组成的解码通路（Decoding Path）。

3.2.2.1 编码器架构

用于对图像进行特征编码的编码器是一个主要由池化层和分频卷积层组成的子计算图。记编码器输入的分频特征为 $X = \{X^L, X^H\}$ ，其中 $X^L \in \mathbb{R}^{ac \times h \times w}$ 为低频特征分量， $X^H \in \mathbb{R}^{(1-\alpha)c \times 2h \times 2w}$ 为高频特征分量，分频特征共有 c 个通道， $\alpha \in [0, 1]$ 是用于控制低频和高频通道数比例参数。

编码器首先利用频率对应的最大值池化层（Frequency-wise Max Pooling Layer）对高低频特征分别进行最大值池化，然后输入两层分频卷积层进行特征编码，编码器输出记为 $Y = \{Y^L, Y^H\}$ ， $Y^L \in \mathbb{R}^{ac \times 0.5h \times 0.5w}$ ， $Y^H \in \mathbb{R}^{(1-\alpha)c \times h \times w}$ ，其计算逻辑为：

$$Y = \text{OctConv}(\text{OctConv}(\text{max_pool}(X, 2))) \quad (3-15)$$

其中 $OctConv(*)$ 为分频卷积操作，具体计算逻辑参考章节 3.1.3。 $max_pool(*,2)$ 由两个分别处理高低频的最大池化层组成，具体计算逻辑参考章节 2.2.2.2。在 $max_pool(*,2)$ 将输入的高低频特征图分别压缩一半空域尺度后，整个编码器内的各组分频特征图的空域尺度保持不变。此外，值得一提的是，最后一层分频卷积需要输出激活值在范围 $[0,1]$ 内的代表像素属于正类的概率置信度图，故选用 $Sigmoid(*)$ 函数作为非线性激活函数。其余的 $OctConv(*)$ 中的非线性激活函数选择为均为 $ReLU(*)$ 函数，且在应用激活函数对未激活值（Pre-activation）进行非线性映射之前，应用了频率对应的批归一化操作（Frequency-wise Batch Normalization, Frequency-wise BN）。

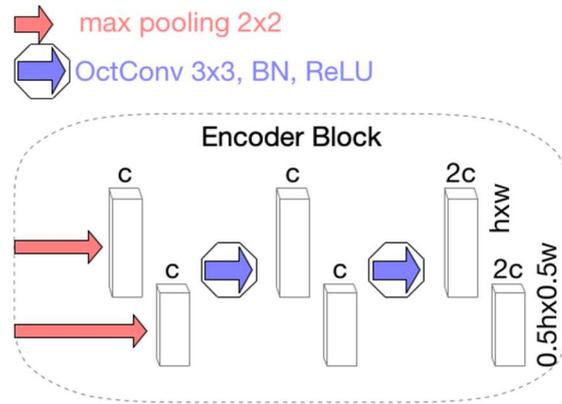


图 3.5 分频卷积编码器架构

3.2.2.2 解码器架构

用于对特征图像进行解码的解码器是一个主要由池化层和分频卷积层组成的子计算图。解码器的输入包括上一层网络层输出的分频卷积特征图和从上一抽象层级的编码器跳跃式连接而来的分频特征图。

记上一层网络层输出的分频特征为 $X = \{X^L, X^H\}$ ，其中 $X^L \in \mathbb{R}^{ac \times 0.5h \times 0.5w}$ 为低频特征分量， $X^H \in \mathbb{R}^{(1-\alpha)c \times h \times w}$ 为高频特征分量。编码器首先会通过一层分频转置卷积层 $OctTranspConv(*,2)$ 对 X 进行解码，并恢复其空域尺度为原来的一倍，得到 $\hat{X} = \{\hat{X}^L, \hat{X}^H\}$ ，其中 $\hat{X}^L \in \mathbb{R}^{ac \times h \times w}$ ， $\hat{X}^H \in \mathbb{R}^{(1-\alpha)c \times 2h \times 2w}$ 。此外，从上一抽象层级的编码器跳跃式连接而来的分频特征图 $K = \{K^L, K^H\}$ 中，低频特征分量 $K^L \in \mathbb{R}^{ac \times h \times w}$ 和高频特征分量 $K^H \in \mathbb{R}^{(1-\alpha)c \times 2h \times 2w}$ ，必须具有与 \hat{X} 匹配的空域尺度。将 K 与 \hat{X} 沿着各频段通道维度拼接（Concatenate）后，将其作为后两层分频卷积层的输入，经过特征变换后得到解码器的输出 $Y = \{Y^L, Y^H\}$ ， $Y^L \in \mathbb{R}^{ac \times h \times w}$ ， $Y^H \in \mathbb{R}^{(1-\alpha)c \times 2h \times 2w}$ ，其计算逻辑为：

$$Y = OctConv(OctConv(Concat(OctTranspConv(X, 2), K))) \quad (3-16)$$

其中 $OctTranspConv(*)$ 为分频转置卷积运算，参考章节 3.2.1。 $Concat(*)$ 为沿着通道维度的张量拼接操作。这一步骤是为了将特征位置信息更丰富的上一抽象层级的编码器输出的分屏特征图作为解码器特征变换的额外输入，从而补充因为编码过程不

断压缩特征图空域尺度而不断丢失的特征位置信息。而解码器的特征变换过程与编码器类似，均为两层 $OctConv(*)$ ，参考章节 3.1.3。

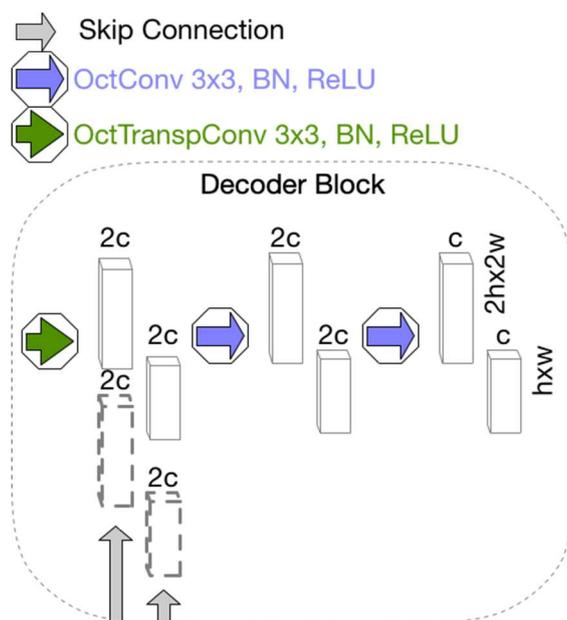


图 3.6 分频卷积解码器架构

3.2.2.3 Octave UNet 网络架构

本文提出了一种名为 Octave UNet 的基于编码器和解码器架构的分频卷积神经网络。在 UNet^[61]网络架构的启发下，结合章节 3.2.2.1 中定义的基于分频卷积的编码器架构和章节 3.2.2.2 中定义的基于分频转置卷积的解码器架构，搭建了主体由编码通路和解码通路构成的卷积神经网络。

如图 3.9 所示，Octave UNet 主要包含以下组成：

- (1) 初始分频卷积模块：该模块由输入层和 2 层分频卷积层组成。输入层为整个网络读入图像数据，并产生一个通道优先(Channel First)的 $3 \times 512 \times 512$ 维度的张量，值得一提的是，Octave UNet 中并没有用全连接和全局池化等参数规模对输入尺度敏感的网络层，所以 Octave UNet 可以处理任意合适大小的图像^①，但是为了加速训练过程^②，选用了固定 512×512 像素的输入图像尺寸。而在 2 层分频卷积层中，卷积核大小均设置为 3×3 ，步长均为 1，边缘补值方式为等大补零 (Same Padding)，即根据卷积核大小和步长，在边缘补 1 像素的常数灰度值 0，空洞系数为 1，即没有使用空洞卷积。此外，每层卷积核个数为 64，即输出分频特征图中高低频通道总数为 64，低高频比率为 α 。所以初始分频卷积模块输出的分频卷积特征图包含 $64(1 - \alpha) \times$

^① 输入图像应该足够大并能所有网络层中产生特征响应图，即输入图像最小尺寸与最深网络层的局部感受野大小一致。

^② 工程实现相关，CUDA 的 Benchmark 模式会针对固定的数据规模配置进行优化，从而加速训练过程。

512 × 512 维度的高频特征张量和 $64\alpha \times 256 \times 256$ 维度的低频特征张量。

- (2) 4 组编码器模块：由多组编码器模块组成的编码通路采用了深度学习逐层抽象的思路，逐层地构建更复杂特征描述。其中每组编码器模块的计算逻辑如章节 3.2.2.1 中所述，主要包含将输入高低频特征张量的空间尺度分别压缩一半的频率对应的最大池化层和 2 组分频卷积层。在整个编码器模块内部，特征张量的维度保持一致，而通道数则由最后一层分频卷积运算翻倍。这是因为随着抽象层级的增加，复杂特征的种类会比简单特征种类多，故用于学习不同特征的卷积核数量及其产生的不同通道数量也应增加。比如在第 1 组编码器中，输入为上一层的输出的 $64(1 - \alpha) \times 512 \times 512$ 维度的高频特征张量和 $64\alpha \times 256 \times 256$ 维度的低频特征张量，经过池化层后，高低分频分量分别被下采样至 $64(1 - \alpha) \times 256 \times 256$ 维度和 $64\alpha \times 128 \times 128$ 维度，经过 2 层分频卷积运算的特征变换后，整个编码器的输出的分频卷积特征图包含 $128(1 - \alpha) \times 256 \times 256$ 维度的高频特征张量和 $128\alpha \times 128 \times 128$ 维度的低频特征分量。提取到的特征响应图参考图 3.7。
- (3) 4 组跳跃式连接：跳跃式连接没有任何特征变换过程，仅仅是将低一抽象层级的分频特征图沿各通道拼接到高一抽象层级的解码器的第一层转置卷积层的输出特征图中。比如，初始分频卷积模块的输出的 $64(1 - \alpha) \times 512 \times 512$ 维度的高频特征张量和 $64\alpha \times 256 \times 256$ 维度的低频特征张量将会作为输入传给解码器 4，而编码器 1 的输出特征图将输入给解码器 3 等等。
- (4) 4 组解码器模块：由多组解码器模块组成的解码通路也采用了深度学习逐层抽象的思路，逐层地解码分频特征图，实现最后像素级分割图的重构。其中每组解码器模块的计算逻辑如章节 3.2.2.2 中所述，主要包含将输入高低频特征张量的空间尺度分别恢复到原来一倍的分频转置卷积层和 2 组分频卷积层。与编码器类似，在整个解码器模块内部，特征张量的维度保持一致。但通道数则由最后一层分频卷积运算减半。比如在第 1 组编码器中，输入为编码器 4 的输出的 $512(1 - \alpha) \times 32 \times 32$ 维度的高频特征张量和 $512\alpha \times 16 \times 16$ 维度的低频特征张量，经过分频转置卷积后，高低分频分量分别被变换至 $512(1 - \alpha) \times 64 \times 64$ 维度和 $512\alpha \times 32 \times 32$ 维度，经过 2 层分频卷积运算的特征变换后，整个编码器的输出的分频卷积特征图包含 $256(1 - \alpha) \times 64 \times 64$ 维度的高频特征张量和 $256\alpha \times 32 \times 32$ 维度的低频特征分量。解码分频特征所得的响应图参考图 3.8。
- (5) 分频卷积输出层：通常神经网络最后的输出层需要根据任务设计。Octave UNet 最后的分频卷积输出层是一个较特殊的分频卷积层，其运算逻辑与章节 3.1.3 中类似，但不包含与低频特征图的子计算图，仅仅包含与输出高频

特征分量相关的运算。此外该分频卷积层不包含 Batch Normalization, 激活函数选用了针对像素级别二分类问题的 Sigmoid(*) 函数。

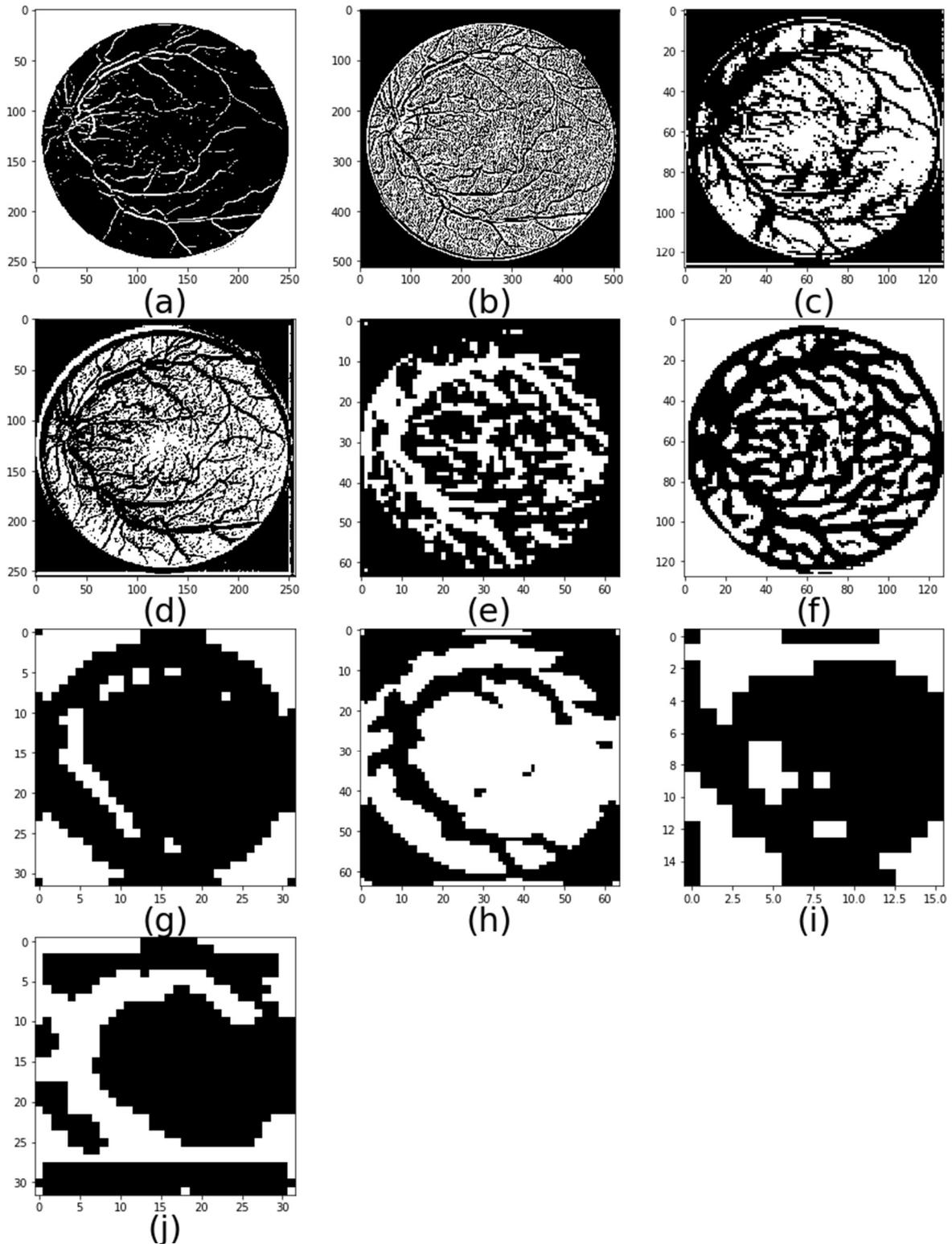


图 3.7 Octave UNet 编码过程提取到的某通道的特征响应图, (a)和(b)对应初始化分频卷积的低频和高频特征响应图, (c)和(d)对应 Encoder1 的低频和高频特征响应图, (e)和(f)对应 Encoder2 的低频和高频特征响应图, (g)和(h)对应 Encoder3 的低频和高频特征响应图, (i)和(j)对应 Encoder4 的低频和高频特征响应图

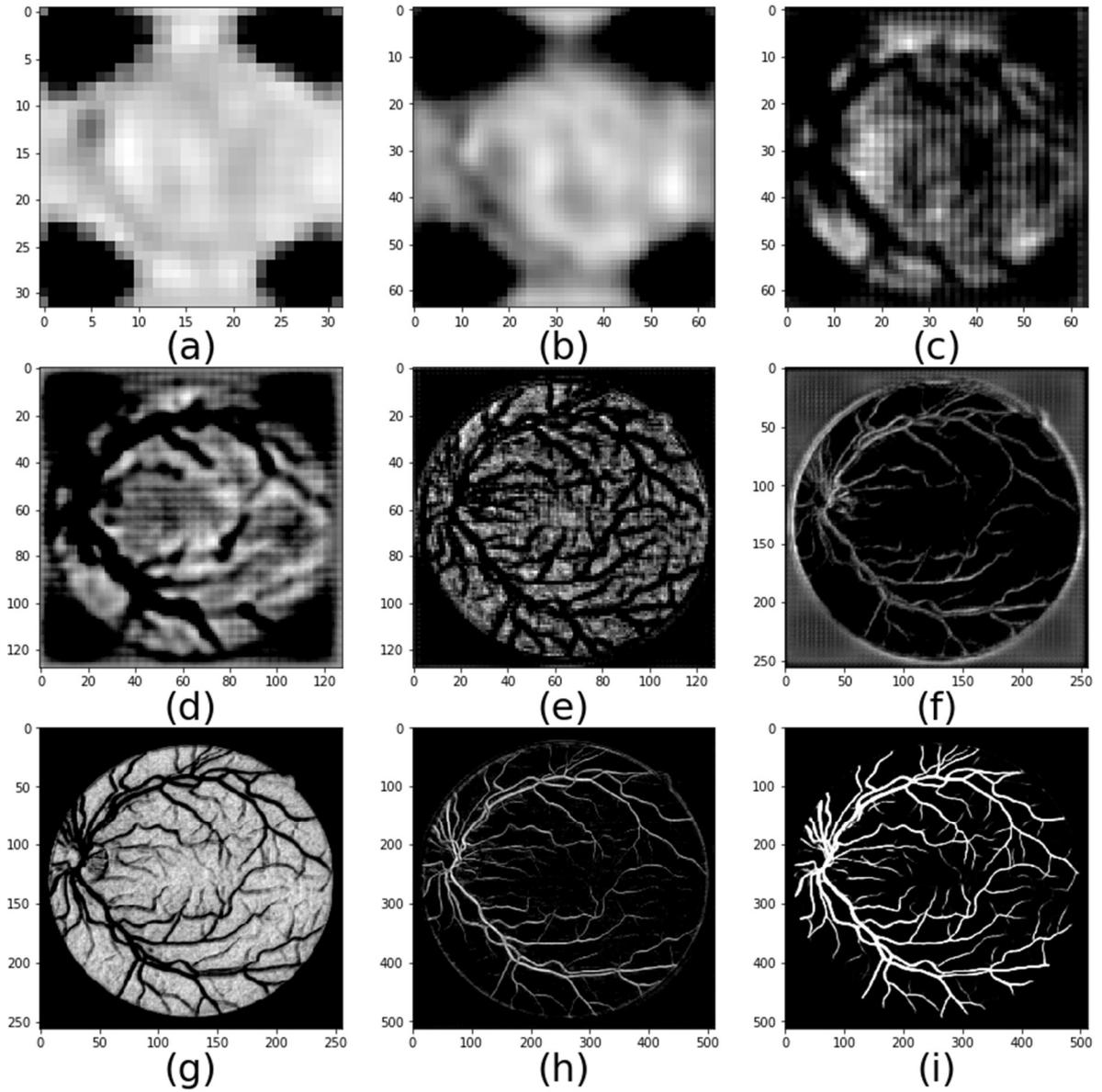


图 3.8 Octave UNet 解码过程提取到的某通道的特征响应图，(a)和(b)对应 Decoder1 的低频和高频特征响应图，(c)和(d)对应 Encoder2 的低频和高频特征响应图，(e)和(f)对应 Encoder3 的低频和高频特征响应图，(g)和(h)对应 Encoder4 的低频和高频特征响应图，(i)对应最终输出

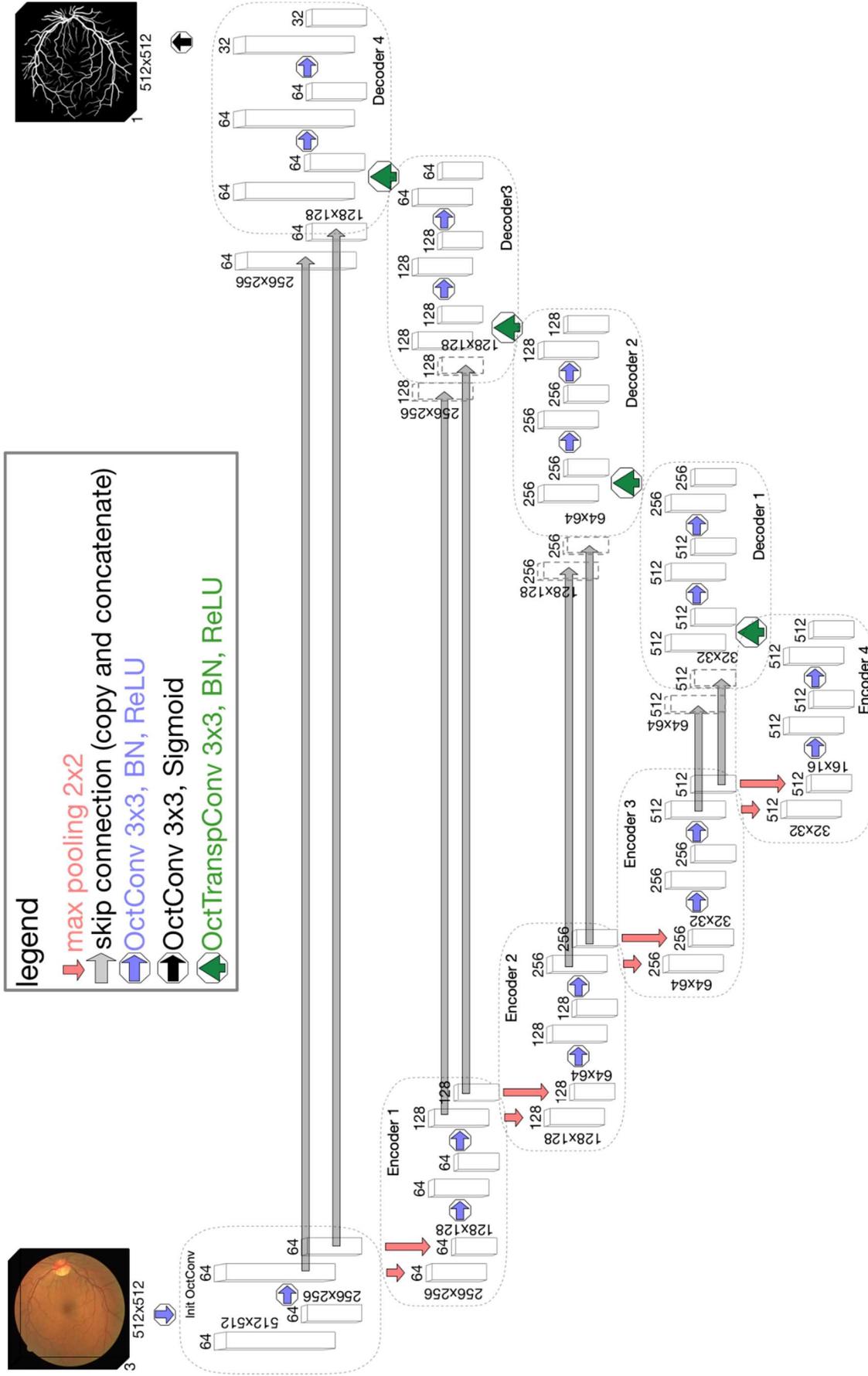


图 3.9 Octave UNet 网络架构

3.2.3 网络参数优化

通常深度神经网络模型在搭建好网络架构后，还需要针对任务设计损耗方程，并利用随机梯度下降等优化算法对网络中的可训练参数进行优化。

3.2.3.1 类别不均衡问题

针对视网膜血管分割任务而言，眼底图像中的血管像素点的数量占眼底图像像素点总量的6% – 10%左右，这意味着正负类别像素的样本比例严重不均衡。这种类别不均衡问题（Class Imbalance Problem）无论是对模型训练，还是评价指标设计而言都是一种挑战^[64]。

针对类别不均衡的视网膜血管分割任务，一种基于图像补丁中心点分类的传统做法是：以图像补丁中心点的标定类别作为整个图像补丁的类别，并以此来作为对图像补丁重采样（Re-sampling）的依据，通过过采样正类图像补丁或者欠采样负类图像补丁来缓解类别不均衡问题。除了受到基于补丁的方法的缺点影响外，这种做法对图像补丁的定义是非常不严谨的：可能存在图像补丁主体包含大量血管像素点，但因补丁中心像素类别非正类而被定义为背景图像补丁的样本。这些样本在训练模型时，不但不能提供准确的指导信息，反而会影响训练。

针对类别不均衡问题，本文选择从模型训练的损耗方程入手，采用了加权二元交叉熵函数（Class Weighted Binary Cross Entropy）作为网络参数的损耗方程。

$$l(\hat{y}_i, y_i) = -w_+ y \log_2 \hat{y}_i + (1 - y) \log_2 (1 - \hat{y}_i) \quad (3-17)$$

$$l(\hat{Y}, Y) = \sum_i l(\hat{y}_i, y_i) \quad (3-18)$$

其中 \hat{y}_i 为 i 样本的模型预测， y_i 为专家标定的金标准， w_+ 则为负责加大误判正类样本时的损耗值的权重，故通常有 $w_+ > 1$ ，而针对视网膜血管分割， w_+ 的值通过统计训练集正负类像素点数量获得。

3.2.3.2 优化算法选择

在优化算法选择上，本文的采用了 Adam（Adaptative Momentum）算法^[65]作为网络参数寻优的优化器。Adam 优化器是一种结合了动量和 RMSProp 算法^[36]的随机梯度下降算法。Adam 优化器对不同超参数选择具有良好的鲁棒性，通常被用作训练深度神经网络的基准算法。

此外，本文还采用了一种自适应的学习率调节策略（Learning Rate Schedule）。当模型评价指标持续一定训练代数（Epoch）没有得到改善后，训练算法会以一定倍率减少学习率，使模型训练继续收敛。

第四章 实验结果与分析

4.1 实验设计

4.1.1 训练集和测试集设置

为了与其他科研文献作对比，在 DRIVE 数据库上的实验均使用了数据库原有的训练集和测试集划分方式。而在 STARE、CHASE_DB1 和 HRF 这三个数据库上的实验结果均通过 Leave One Out 策略划分训练集和测试样本，进行多次独立实验获得。以 STARE 数据库为例，STARE 数据库包含 20 个数据样本对，在每个独立实验中，都会把其中 1 个数据样本对作为测试样本评价算法指标，其余的 19 个样本对作为训练集，用于网络训练。通过共 20 次独立实验，获得 20 张 STARE 数据库中视网膜眼底图像样本的分割效果。本文中所有数据库中的样本编号均在数据原文件名降序排序所得。

4.1.2 实验环境及参数设置

本文实验以 GPU 集群 Linux 服务器作为实验硬件平台，其主要计算资源为 4 块 12GB 显存，计算能力为 3.6 的 NVIDIA-GTX-Titan-XP 型号的 GPU。除了训练时可以采用多 GPU 数据并行 (Data Parrellel) 的方式加速训练外，如无特别说明，本文提出的 Octave UNet 模型均在单 GPU 模式下训练、测试及部署。

实验参数方面，模型训练的最大代数为 500，并以此为模型训练终止条件。CPU 和 GPU 随机数生成器的随机种子均设置并固定为 42。模型优化器 Adam 的初始学习率设置为 0.01，其余超参设置为默认参数。自适应学习率策略为：监测指标持续 10 代没有得到优化，则下降当前学习率至 95%，最小学习率下限为 0.00001。值得一提的是，本文方法没有采用任何图像预处理、图像后处理和图像增广技术。

架构如图 3.9 所示的 Octave UNet 模型在处理一张 512×512 分辨率的 3 通道 RGB 彩色眼底图像时，平均耗时约 0.4 秒。方法代码实现基于 PyTorch^① 开源框架，模型训练过程的可视化工具则选用了 TensorBoard^②，附录 E 展示了一个模型过程可视化样例。

4.1.3 实验数据库及评价指标说明

本文选用的视网膜眼底图像数据库包括 DRIVE 数据库、STARE 数据库、CHASE_DB1 数据库和 HRF 数据库。数据库的介绍详见章节 2.1.1。

本文采用的针对视网膜血管分割算法的评价指标包括：正确率 (ACC)、敏感度 (SE)、

^① <https://pytorch.org/>

^② <https://github.com/tensorflow/tensorboard>

特异性 (SP)、F1 分数 (F1) 和 ROC 曲线下面积 (AUROC)。各评价指标的定义详见章节 2.1.2。如无特别说明，本文中报告的所有对阈值敏感的评价指标，包括正确率，敏感度，特异性和 F1 分数，均是在取阈值 $T = 0.5$ 的情况下计算所得。而对阈值步敏感的评价指标，如 AUROC 则是在遍历阈值梯度 $T = 0.01, 0.02, 0.03, \dots, 0.98, 0.99$ 的情况下计算所得。

4.2 实验结果与分析

4.2.1 各数据库上的实验结果

表 4.3 Octave UNet 在 DRIVE 数据库测试集上的实验结果

	ACC	SE	SP	F1	AUROC
19 (最佳)	0.970478	0.897508	0.977105	0.835049	0.990026
03 (最差)	0.956635	0.667089	0.988573	0.753481	0.967355
均值	0.966067	0.795736	0.982697	0.803288	0.981833
标准差	0.003321	0.058350	0.004398	0.018434	0.005816

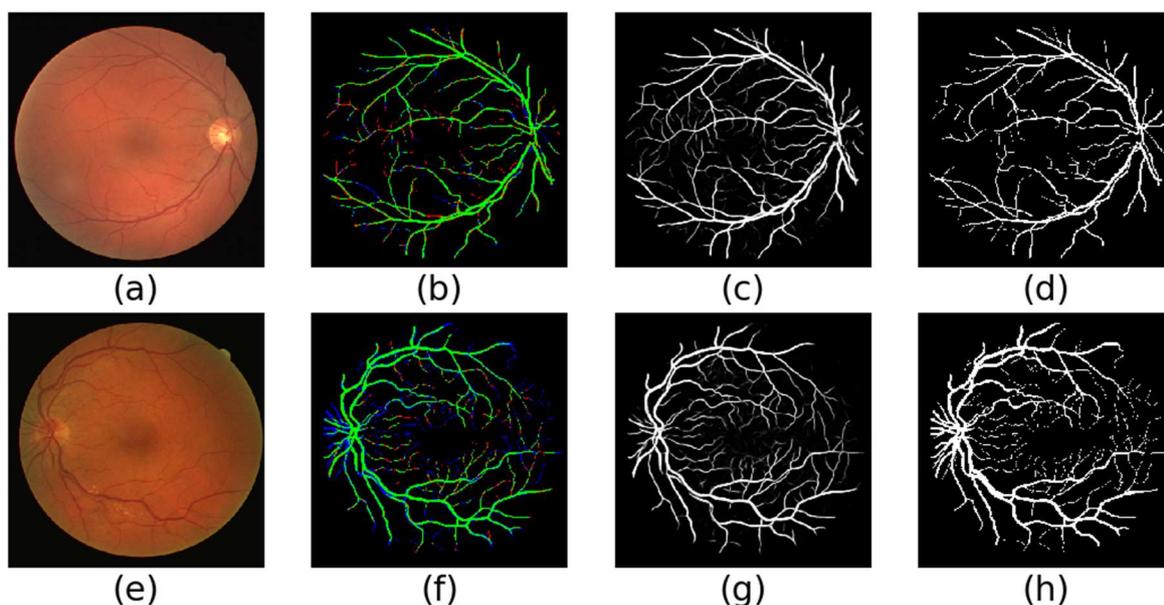


图 4.1 本文方法在 DRIVE 测试样本上最佳（第一行）和最差（第二行）的实验结果，(a)和(e)为彩色眼底图像，(b)和(f)为结果分析图，(c)和(g)为模型输出的血管概率置信度图，(d)和(h)为专家标注血管结构

表 4.4 Octave UNet 在 STARE 数据库上的实验结果

	ACC	SE	SP	F1	AUROC
12 (Best)	0.979469	0.869667	0.988604	0.866782	0.995187
20 (Worse)	0.968586	0.646472	0.991692	0.733676	0.975118
均值	0.974190	0.816438	0.986994	0.824953	0.989191
标准差	0.005509	0.064650	0.003361	0.037742	0.005520

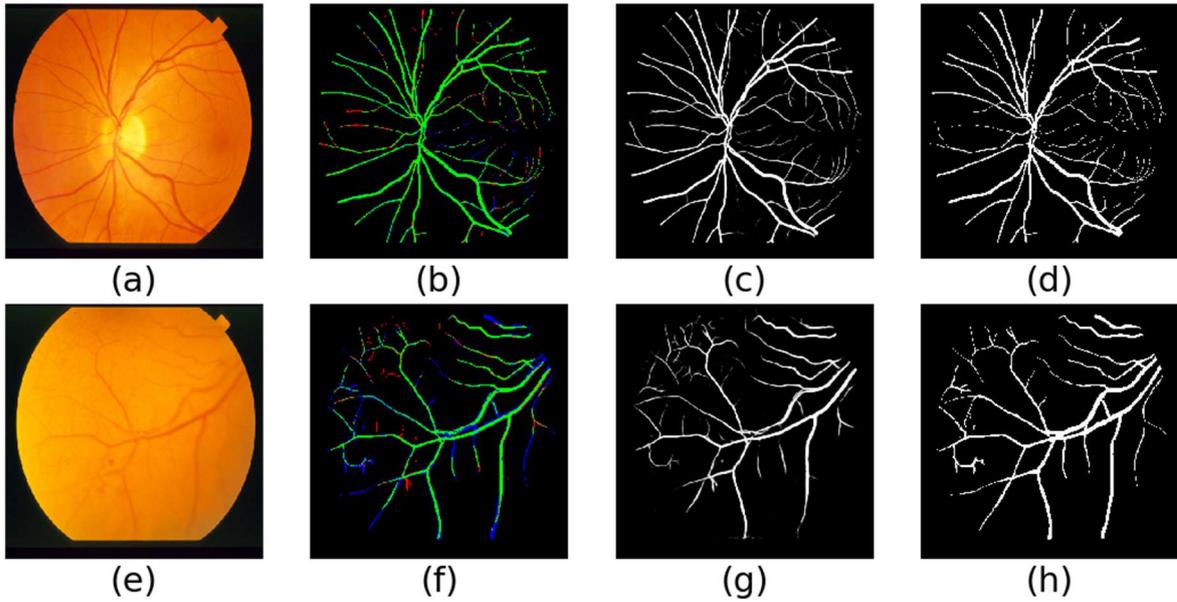


图 4.2 本文方法在 STARE 测试样本上最佳（第一行）和最差（第二行）的实验结果，
 (a)和(e)为彩色眼底图像，(b)和(f)为结果分析图，
 (c)和(g)为模型输出的血管概率置信度图，(d)和(h)为专家标注血管结构

表 4.5 Octave UNet 在 CHASE_DB1 数据库上的实验结果

	ACC	SE	SP	F1	AUROC
7 (best)	0.977844	0.881950	0.986020	0.862148	0.993532
26 (worse)	0.973591	0.741978	0.989166	0.779768	0.982529
均值	0.976251	0.824447	0.987432	0.826300	0.989097
标准差	0.002940	0.033417	0.003061	0.023189	0.003093

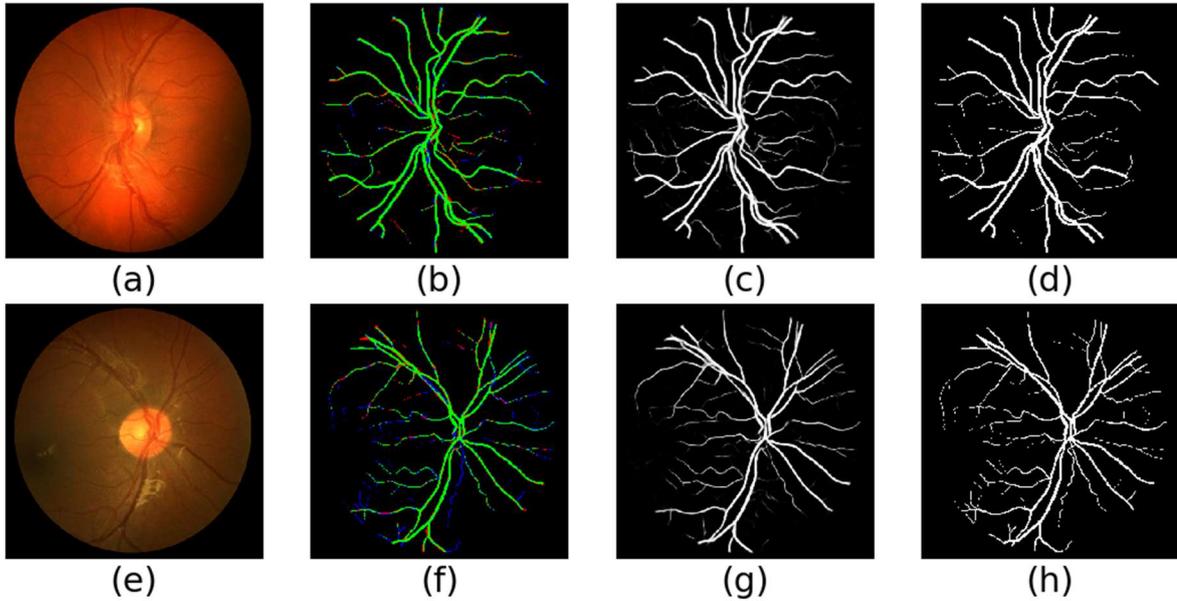


图 4.3 本文方法在 CHASE_DB1 测试样本上最佳（第一行）和最差（第二行）的实验结果，
 (a)和(e)为彩色眼底图像，(b)和(f)为结果分析图，
 (c)和(g)为模型输出的血管概率置信度图，(d)和(h)为专家标注血管结构

表 4.6 Octave UNet 在 HRF 数据库上的实验结果

	ACC	SE	SP	F1	AUROC
16 (worse)	0.959465	0.712489	0.979556	0.725625	0.962333
45 (best)	0.978550	0.793366	0.994361	0.853350	0.994412
均值	0.971383	0.802003	0.985281	0.807909	0.985131
标准差	0.004702	0.039377	0.004157	0.041109	0.006799

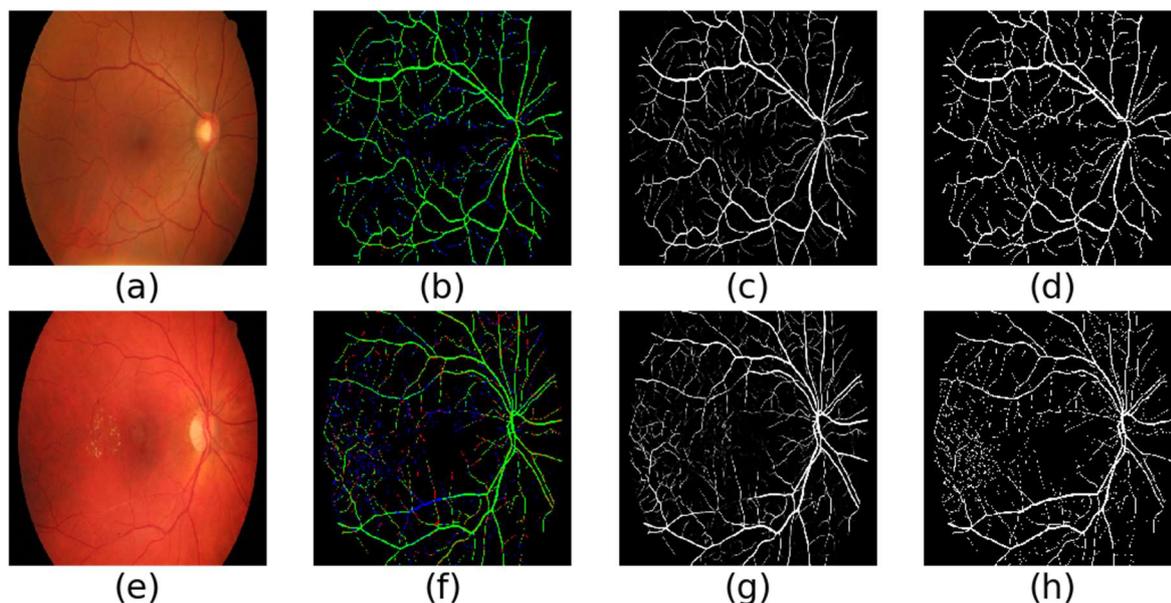


图 4.4 本文方法在 HRF 测试样本上最佳（第一行）和最差（第二行）的实验结果，
 (a)和(e)为彩色眼底图像，(b)和(f)为结果分析图，
 (c)和(g)为模型输出的血管概率置信度图，(d)和(h)为专家标注血管结构

上述数值指标和可视化结果^①说明：在 DRIVE 数据库上，本文方法能较好地分割视网膜血管结构，但在结果分析图中仍然存在少量的假阳性（红色像素）和假阴性（蓝色像素）点。在 STARE、CHASE_DB1 和 HRF 数据库中，本文算法也能较好的分割视网膜血管结构，但仍存在一些漏判边缘模糊的血管结构的现象。

^① 结果分析图定义详见章节 2.1.2.2。

4.2.2 与先进方法的结果对比

表 4.7 本文方法在 DRIVE 数据库测试集上的实验结果

Methods	Year	ACC	SE	SP	F1	AUROC
Unsupervised Methods						
Zena et al. ^[66]	2001	0.9377	0.6971	0.9769	-	0.8984
Mendonca et al. ^[67]	2006	0.9452	0.7344	0.9764	-	-
Al-Diri et al. ^[68]	2009	0.9258	0.7282	0.9551	-	-
Miri et al. ^[69]	2010	0.9458	0.7352	0.9795	-	-
You et al. ^[70]	2011	0.9434	0.7410	0.9751	-	-
Fraz et al. ^[29]	2012	0.9430	0.7152	0.9768	-	-
Fathi et al. ^[71]	2013	-	0.7768	0.9759	0.7669	-
Sreejini et al. ^[72]	2015	0.9633	0.7132	0.9866	-	-
Roychowdhury et al. ^[73]	2015	0.9494	0.7395	0.9782	-	-
Fan et al. ^[32]	2019	0.9600	0.7360	0.9810	-	-
Supervised Methods						
Staal et al. ^[45]	2004	0.9441	0.7194	0.9773	-	0.9520
Ricci et al. ^[26]	2007	0.9563	-	-	-	0.9558
Marin et al. ^[35]	2011	0.9452	0.7067	0.9801	-	0.9588
Fraz et al. ^[31]	2012	0.9480	0.7460	0.9807	-	0.9747
Cheng et al. ^[74]	2014	0.9472	0.7252	0.9778	-	0.9648
Orlando et al. ^[75]	2014	-	0.7850	0.9670	0.7810	-
Vega et al. ^[76]	2015	0.9412	0.7444	0.9612	0.6884	-
Fan et al. ^[32]	2016	0.9614	0.7191	0.9849	-	-
Fan et al. ^[33]	2016	0.9612	0.7814	0.9788	-	-
Liskowski et al. ^[39]	2016	0.9535	0.7811	0.9807	-	0.9790
Li et al. ^[77]	2016	0.9527	0.7569	0.9816	-	0.9738
Orlando et al. ^[34]	2016	-	0.7897	0.9684	0.7857	-
Mo et al. ^[43]	2017	0.9521	0.7779	0.9780	-	0.9782
Xiao et al. ^[45]	2018	0.9655	0.7715	-	-	-
Alom et al. ^[44]	2018	0.9556	0.7792	0.9813	0.8171	0.9784
Proposed Method	2019	0.9661	0.7957	0.9827	0.8033	0.9818

表 4.8 本文方法在 STARE 数据库测试集上的实验结果

Methods	Year	ACC	SE	SP	F1	AUROC
Unsupervised Methods						
Hoover et al. ^[78]	1998	0.9264	0.6747	0.9565	-	-
Mendonca et al. ^[67]	2006	0.9440	0.6996	0.9730	-	-
Al-Diri et al. ^[68]	2009	-	0.7521	0.9681	-	-
You et al. ^[70]	2011	0.9497	0.7260	0.9756	-	-
Fraz et al. ^[29]	2012	0.9442	0.7311	0.9680	-	-
Fathi et al. ^[71]	2013	-	0.8061	0.9717	0.7509	-
Roychowdhury et al. ^[73]	2015	0.9560	0.7317	0.9842	-	-
Fan et al. ^[32]	2019	0.9570	0.7910	0.9700	-	-
Supervised Methods						
Staal et al. ^[45]	2004	0.9516	-	-	-	0.9614
Ricci et al. ^[26]	2007	0.9584	-	-	-	0.9602
Marin et al. ^[35]	2011	0.9526	0.6944	0.9819	-	0.9769
Fraz et al. ^[31]	2012	0.9534	0.7548	0.9763	-	0.9768
Vega et al. ^[76]	2015	0.9483	0.7019	0.9671	0.6614	-
Fan et al. ^[32]	2016	0.9588	0.6996	0.9787	-	-
Fan et al. ^[33]	2016	0.9654	0.7834	0.9799	-	-
Liskowski et al. ^[39]	2016	0.9729	0.8554	0.9862	-	0.9928
Li et al. ^[77]	2016	0.9628	0.7726	0.9844	-	0.9879
Orlando et al. ^[34]	2016	-	0.7680	0.9738	0.7644	-
Mo et al. ^[43]	2017	0.9674	0.8147	0.9844	-	0.9885
Xiao et al. ^[79]	2018	0.9693	0.7469	-	-	-
Alom et al. ^[79]	2018	0.9712	0.8292	0.9862	0.8475	0.9914
Proposed Method	2019	0.9741	0.8164	0.9870	0.8250	0.9892

表 4.9 本文方法在 CHASE_DB1 数据库测试集上的实验结果

Methods	Year	ACC	SE	SP	F1	AUROC
Unsupervised Methods						
Fraz et al. ^[30]	2014	-	0.7259	0.9770	0.7488	-
Roychowdhury et al. ^[73]	2015	0.9467	0.7615	0.9575	-	-
Fan et al. ^[32]	2019	0.9510	0.6570	0.9730	-	-
Supervised Methods						
Fraz et al. ^[31]	2012	0.9469	0.7224	0.9711	-	0.9712
Fan et al. ^[33]	2016	0.9573	0.7656	0.9704	-	-
Liskowski et al. ^[39]	2016	0.9628	0.7816	0.9836	-	0.9823
Li et al. ^[77]	2016	0.9527	0.7569	0.9816	-	0.9738
Orlando et al. ^[34]	2016	-	0.7277	0.9712	0.7332	-
Mo et al. ^[43]	2017	0.9581	0.7661	0.9793	-	0.9812
Alom et al. ^[79]	2018	0.9634	0.7756	0.9820	0.7928	0.9815
Proposed Method	2019	0.9714	0.8020	0.9853	0.8079	0.9851

表 4.10 本文方法在 HRF 数据库测试集上的实验结果

Methods	Year	ACC	SE	SP	F1	AUROC
Unsupervised Methods						
Roychowdhury et al. ^[73]	2015	0.9467	0.7615	0.9575	-	-
Supervised Methods						
Odstreilik et al. ^[48]	2013	-	0.7794	0.9584	0.7158	-
Orlando et al. ^[34]	2016	-	0.7874	0.9584	0.7158	-
Proposed Method	2019	0.9763	0.8244	0.9874	0.8079	0.9891

本文方法在 DRIVE、CHASE_DB1 和 HRF 数据库上均取得了最先进的结果。而在 STARE 数据库上则取得了最先进的正确率和特异性，以及与其他先进方法可比的召回率、F1 分数和 AUROC^①。

^① 值得一提的是，本文方法在表 4.7，表 4.8，表 4.9 和表 4.10 中汇报的结果均采用了 $T = 0.5$ 的二值化阈值，若采用 $T = 0.3$ 的阈值，则在 STARE 数据库上，本文算法指标有：正确率 0.9768，召回率 0.9112，特异性 0.9852 和 F1 分数 0.8576，与其他算法相比最为先进。

4.2.3 阈值敏感度分析

多数视网膜血管分割算法的模型输出是一张与输入图像尺度大小一致的概率图，图上像素点的概率值代表该像素点属于正类血管像素的置信度大小。为了做出最终的分割判定，通常需要用阈值处理将概率图二值化。

本文在视网膜血管分割算法衡量时采用了四种对阈值敏感的评价指标：正确率（ACC）、召回率（SE）、特异性（SP）和 F1 分数（F1）。

图 4.5、图 4.8、图 4.7 和图 4.6 中的阈值敏感度曲线具有相同的趋势变化。随着二值化阈值大小的增加，像素点需要的置信度越高才会被判定为正类，这意味着正类判定更加严格，召回率会随之下降，而特异性会随之上升。而本文方法在各数据库上正确率、特异性和 F1 分数的敏感度曲线在 $[0.3, 0.7]$ 阈值范围内变化平稳，说明了本文方法具有一定的鲁棒性。此外，通过减少本文采用的0.5的固定阈值大小，可以在牺牲少量特异性过的同时，有效提高召回率。这意味着即便放宽血管正类的判定，本文方法对误判背景像素为血管像素的频率也较低。

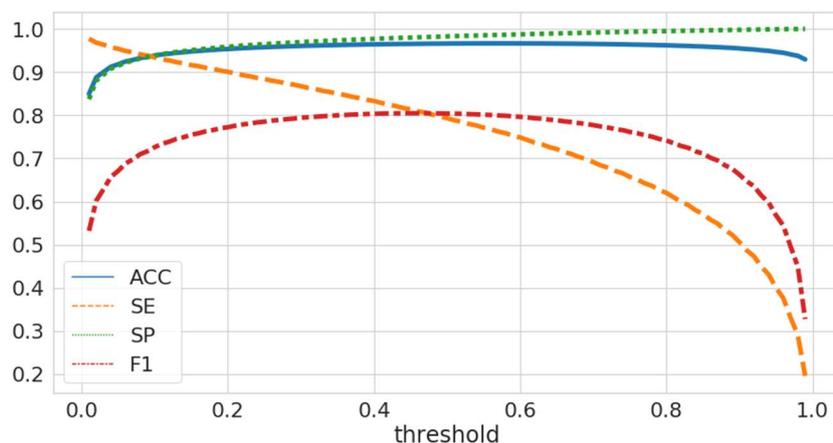


图 4.5 本文方法在 DRIVE 数据库上的敏感度分析

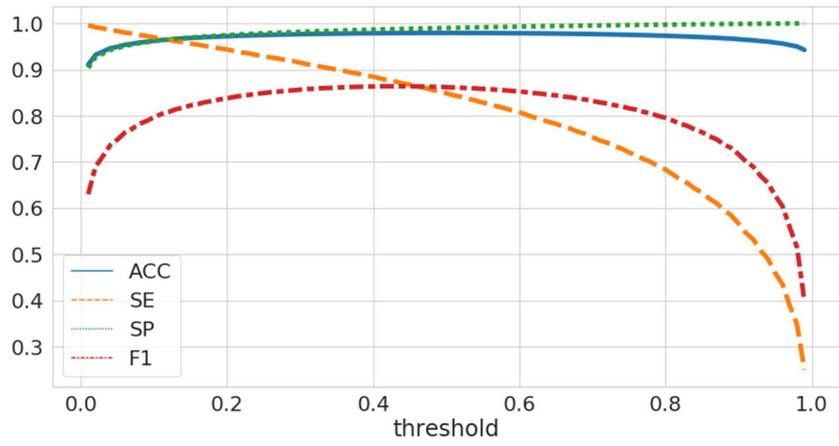


图 4.8 本文方法在 STARE 数据库上的敏感度分析

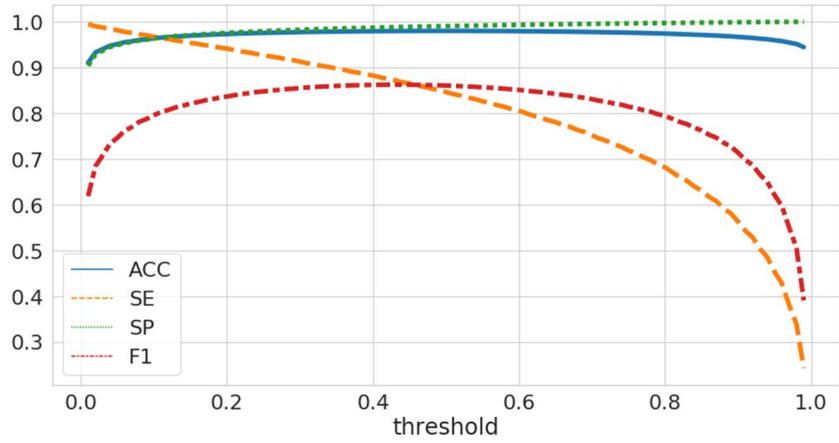


图 4.7 本文方法在 CHASE_DB1 数据库上的敏感度分析

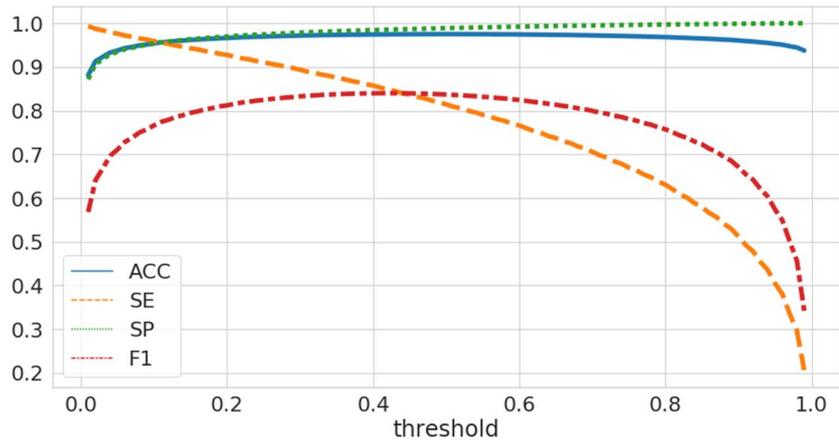


图 4.6 本文方法在 HRF 数据库上的敏感度分析

4.2.4 案例分析

4.2.4.1 低对比度血管结构的精确分割

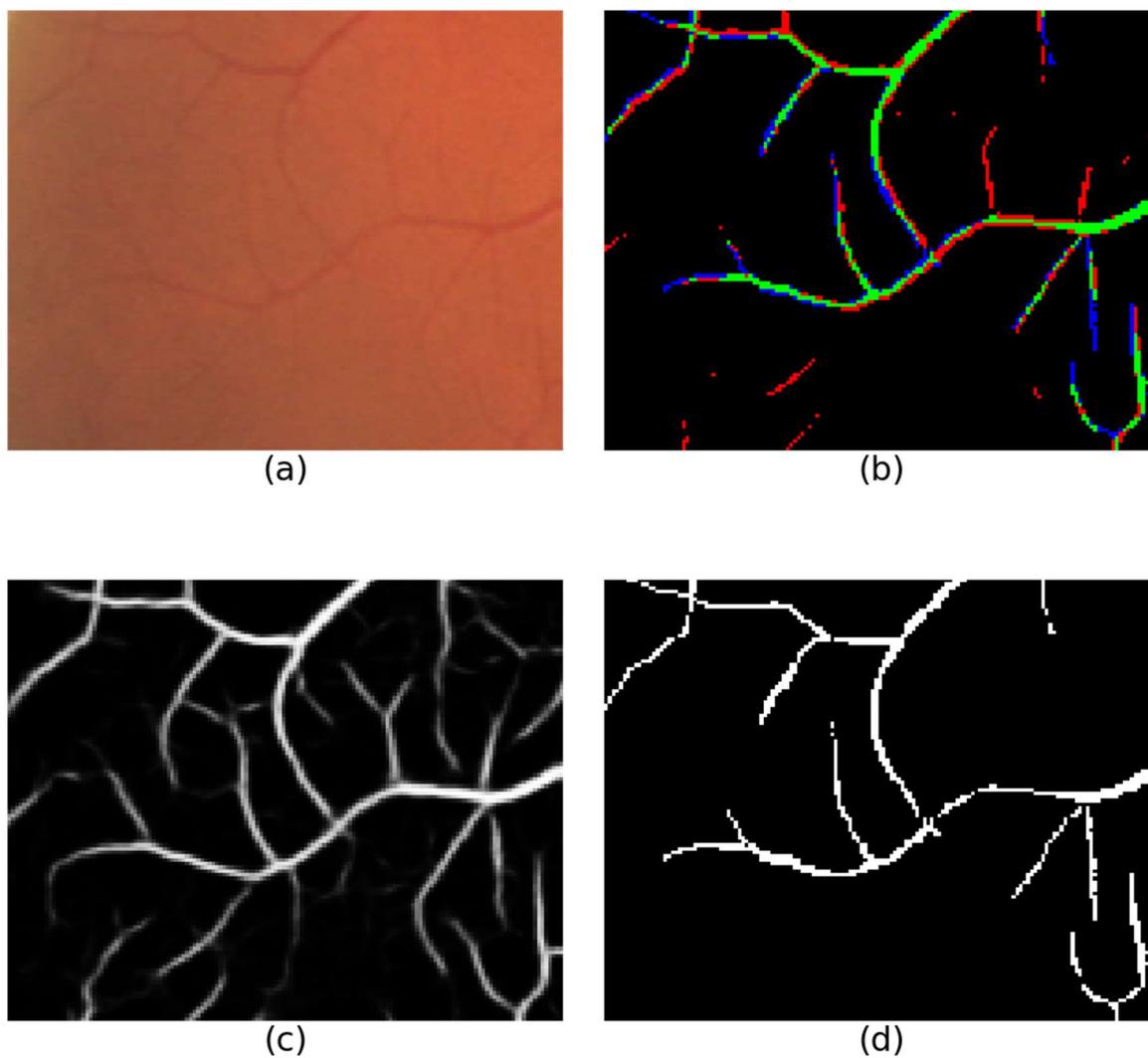


图 4.9 本文算法能有效对低对比度的血管结构精确分割，如(b)中的红色像素和(c)中的细小分叉结构很可能是专家标定漏判的低对比度血管结构

4.2.4.2 抗眼底图像病灶干扰

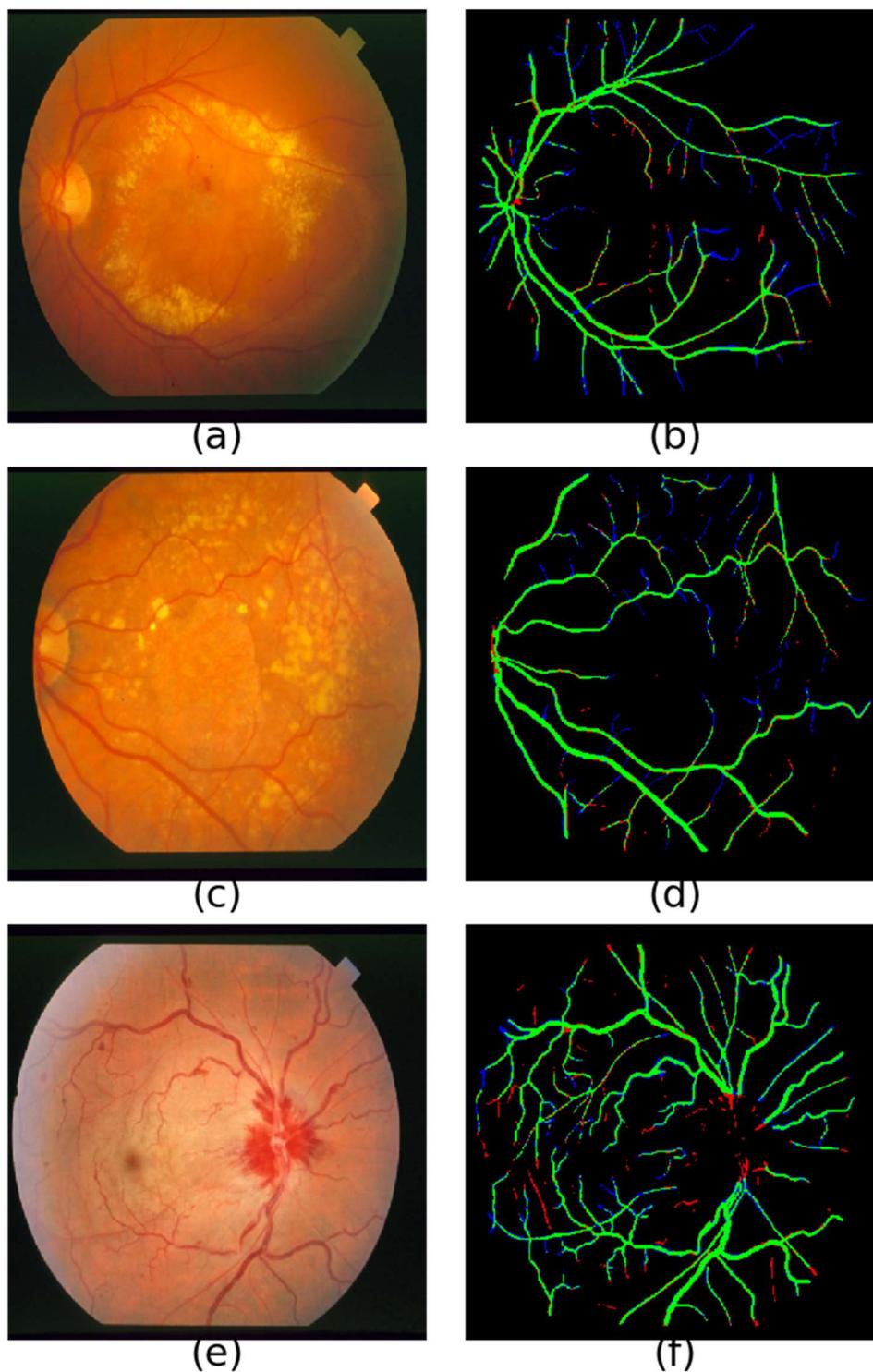


图 4.10 本文方法对复杂的眼底图像病灶具有一定的鲁棒性，第一行是对带有硬性渗出病灶的样本的血管分割结果，第二行是对带有糖尿病病变的样本的血管分割结果，第三行是对带有出血病灶的样本的血管分割结果

4.2.4.3 抗血管中心反光干扰

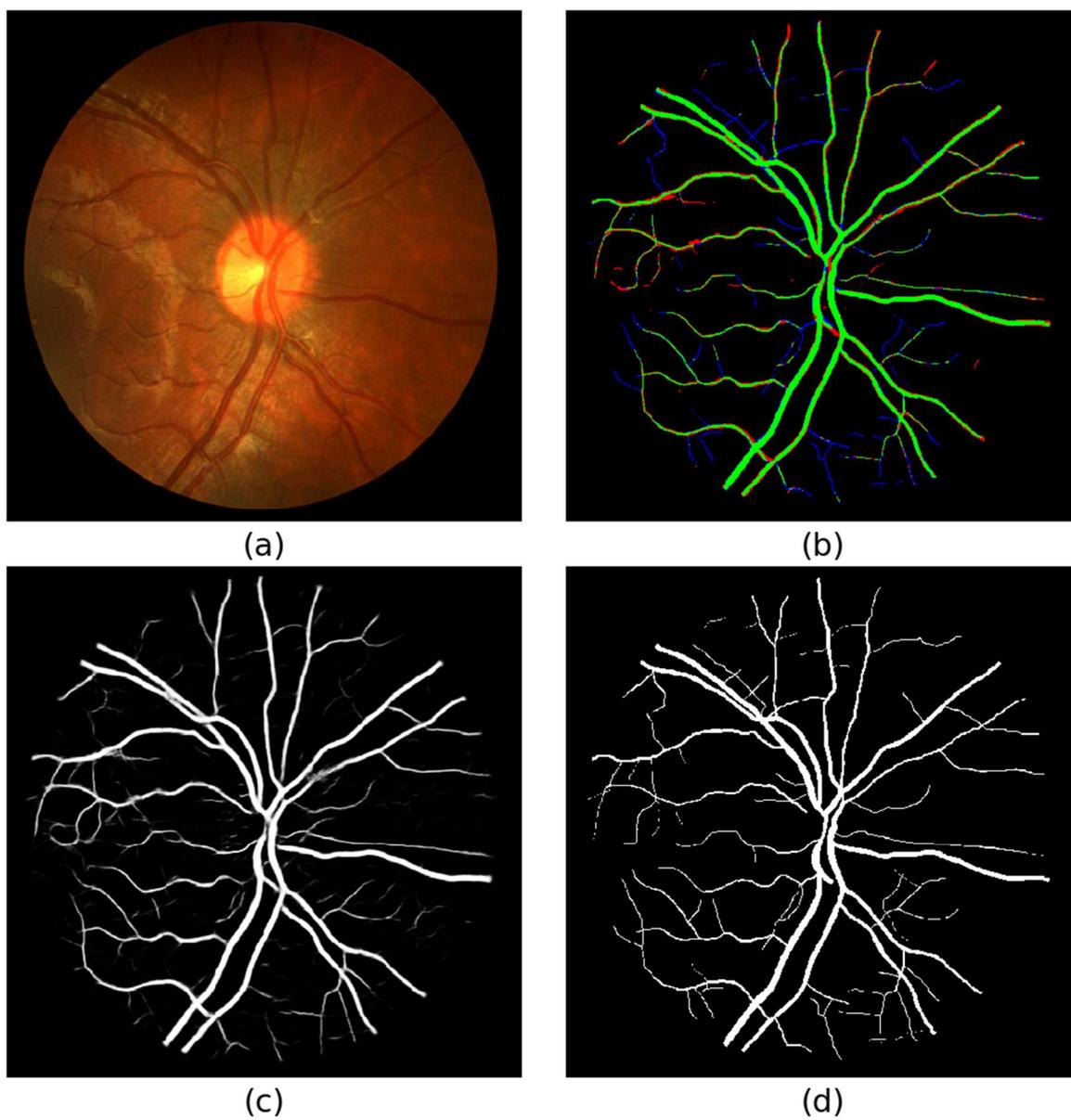


图 4.11 本文方法对血管中心反光现象的干扰具有一定鲁棒性

4.2.4.4 抗光照不平衡干扰

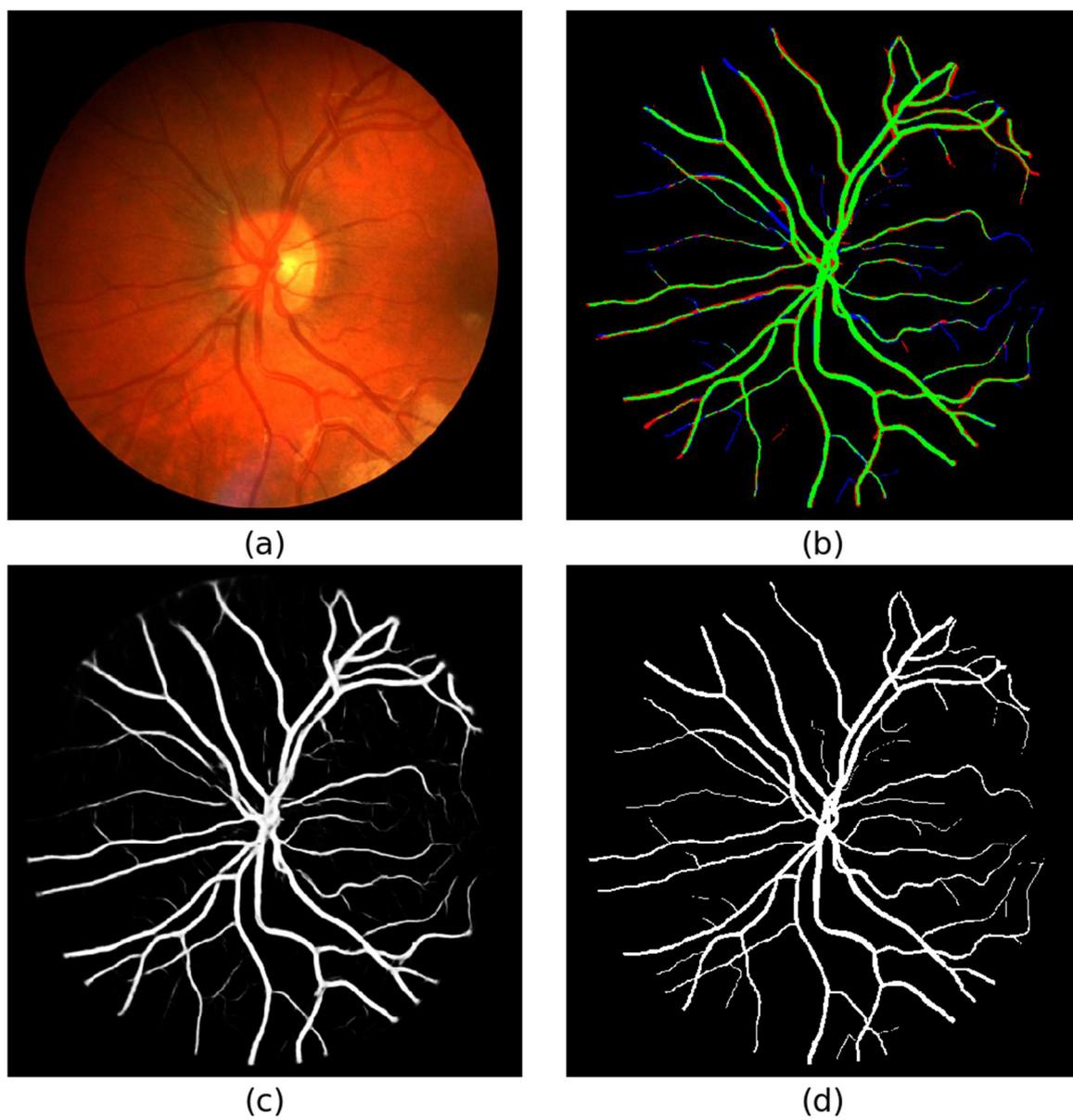


图 4.12 本文方法对光照不平衡现象的干扰具有一定鲁棒性

第五章 结论及展望

5.1 结论

视网膜眼底图像分析被广泛用于如青光眼和白内障等眼科疾病、糖尿病、高血压和动脉硬化等心血管类疾病的诊断，筛查和临床研究。对视网膜血管的精确分割是视网膜眼底图像分析中最重要的一个步骤，视网膜血管不但能反映如糖尿病视网膜病变等疾病的病情，还能帮助如微动脉瘤和硬性渗出等视网膜眼底病灶的定位和诊断。此外，视网膜血管结构作为人体一种可以通过非入侵方式观察到的深层血管结构，视网膜血管图像的获取较为方便快捷，可以用作生物特征识别的依据。

在临床实践中，视网膜血管分割一般由眼科医师或专家标注完成，是一项需要熟练技巧，乏味而耗时的工作。基于智能算法的视网膜血管分割具有重大的研究意义和临床及商业价值。但是，由于视网膜病变、血管中心反光现象、血管背景对比度低、血管分支和拓扑结构复杂等影响，视网膜血管分割算法的设计是一项具有挑战性地任务。

为了克服上述挑战，本文提出了基于编码器解码器架构分频卷积神经网络（Octave UNet）的视网膜血管分割方法，实现了快速而高精度的眼底图像中的视网膜血管自动分割。本文的主要贡献主要包括：

- (1) 从深度学习角度出发，以深度全卷积神经网络为基础，通过神经网络逐层地学习复杂血管特征描述子的构造，提高模型对血管像素地辨识能力。在算法设计过程中避免了传统的人工设计特征描述和对视网膜眼底图像的复杂而启发式强的预处理，尽可能地利用彩色眼底图像的原始数据特征，验证了 Octave UNet 的有效性。
- (2) 以计算图为工具，逐层抽象，采用分频卷积操作搭建了全卷积特征编码器，在同一抽象层级提取高低分频的特征响应，从而加强 Octave UNet 对血管特征的描述能力和辨识能力。
- (3) 为了让网络具备学习如何解码这些高低分频特征的能力，本文提出了一种新的分频转置卷积操作，该操作通过利用多组可训练的转置卷积核学习同频特征变换和异频信息交互，从而恢复特征响应图的空域尺度，完成高低分频特征响应的解码。
- (4) 在分频编码器和分频解码器架构的基础上，本文还提出了一种名为 Octave UNet 的分频全卷积神经网络，该网络架构通过层叠分屏编码器逐层学习分频特征，利用跳跃式连接将抽象层级低而位置信息丰富的分频特征列输入高一抽象层级的分频解码器，并以此来加强抽象层级高而位置信息匮乏的分频特征响应的解码，从而实现高精度的血管分割。

- (5) 摒弃了基于图像补丁中心点分类的问题建模方式，而采用更适合医学图像分割的更精确的像素级别分类的问题建模方式。通过采用损耗方程设计的方式缓解了视网膜血管分割任务类别不均衡问题的影响，实现了 Ocatve UNet 的端对端训练。

Ocatve UNet 在四个公开数据库：DRIVE，STARE，CHASE_DB1 和 HRF 数据库上进行了模型评估。其实验结果表明，本文所提出的方法已经获得了与先进视网膜血管分割方法相比更好或相当的结果。而案例分析则体现了针对视网膜血管分割任务，Ocatve UNet 在对各类视网膜病变和阈值变化具有一定的鲁棒性。

5.2 研究课题展望

在视网膜血管分割这一研究领域，目前公开数据库数量并不多，带有人工标定金标准的数据库则更少，这对于数据驱动的深度学习方法而言是一种挑战，但同时基于无监督学习 (Unsupervised Learning) 或半监督学习 (Semi-Supervised Learning) 的方法对数据标定地依赖较有监督学习方法弱，这类在视网膜血管分割任务上的应用具有重大的潜力和重要的研究意义。此外，不同数据库中的数据样本分布差异较大，在某一数据库上训练的模型直接用于另一个数据库样本的测试时往往效果欠佳，这对于实际临床应用而言是非常不利的，增量学习 (Incremental Learning) 和在线学习 (Online Learning) 技术对于部署临床视网膜血管分割系统或许能提供重要指导。

传统机器学习算法和数字图像处理方法为视网膜血管分割任务积累了大量的基础理论和先验知识，如何利用这些先验知识，改善深度学习模型在视网膜血管分割任务上的效果是一个重要的研究热点。

参考文献

- [1] ABRAMOFF M D, GARVIN M K, SONKA M. Retinal imaging and image analysis[J]. *IEEE Reviews in Biomedical Engineering*, 2010, 3: 169–208.
- [2] L SRINIDHI C, APARNA P, RAJAN J. Recent Advancements in Retinal Vessel Segmentation[J]. *Journal of Medical Systems*, 2017, 41(4): 70.
- [3] MOCCIA S, DE MOMI E, EL HADJI S, et al. Blood vessel segmentation algorithms — Review of methods, datasets and evaluation metrics[J]. *Computer Methods and Programs in Biomedicine*, 2018, 158: 71–91.
- [4] VOSTATEK P, CLARIDGE E, UUSITALO H, et al. Performance comparison of publicly available retinal blood vessel segmentation methods[J]. *Computerized Medical Imaging and Graphics*, Pergamon, 2017, 55: 2–12.
- [5] FRAZ M M, REMAGNINO P, HOPPE A, et al. Blood vessel segmentation methodologies in retinal images - A survey[J]. *Computer Methods and Programs in Biomedicine*, 2012, 108(1): 407–433.
- [6] WONG W L, SU X, LI X, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: A systematic review and meta-analysis[J]. *The Lancet Global Health*, Elsevier, 2014, 2(2): 106–116.
- [7] YAU J W Y, ROGERS S L, KAWASAKI R, et al. Global prevalence and major risk factors of diabetic retinopathy[J]. *Diabetes Care*, American Diabetes Association, 2012, 35(3): 556–564.
- [8] THAM Y C, LI X, WONG T Y, et al. Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis[J]. *Ophthalmology*, Elsevier, 2014, 121(11): 2081–2090.
- [9] WELIKALA R A, DEHMESHKI J, HOPPE A, et al. Automated detection of proliferative diabetic retinopathy using a modified line operator and dual classification[J]. *Computer Methods and Programs in Biomedicine*, 2014, 114(3): 247–261.
- [10] YOUSSEF D, SOLOUMA N H. Accurate detection of blood vessels improves the detection of exudates in color fundus images[J]. *Computer Methods and Programs in Biomedicine*, Elsevier, 2012, 108(3): 1052–1061.
- [11] WONG T Y, KLEIN R, KLEIN B E K, et al. Retinal microvascular abnormalities and their relationship with hypertension, cardiovascular disease, and mortality[J]. *Survey of Ophthalmology*, Elsevier, 2001, 46(1): 59–80.
- [12] WONG T Y, KLEIN R, COUPER D J, et al. Retinal microvascular abnormalities and incident stroke: The Atherosclerosis Risk in Communities Study[J]. *Lancet*, Elsevier, 2001, 358(9288): 1134–1140.
- [13] SMITH W, WANG J J, WONG T Y, et al. Retinal arteriolar narrowing is associated with 5-year incident severe hypertension: The Blue Mountains Eye Study[J]. *Hypertension*, Lippincott Williams & Wilkins, 2004, 44(4): 442–447.
- [14] WONG T, MITCHELL P. The eye in hypertension[J]. *Lancet*, Elsevier, 2007, 369(9559): 425–435.
- [15] DOUBAL F N, HOKKE P E, WARDLAW J M. Retinal microvascular abnormalities and stroke: A systematic review[J]. *Journal of Neurology, Neurosurgery and Psychiatry*, BMJ Publishing Group Ltd, 2009, 80(2): 158–165.
- [16] HUGHES A D, FALASCETTI E, WITT N, et al. Association of retinopathy and retinal microvascular abnormalities with stroke and cerebrovascular disease[J]. *Stroke*, 2016, 47(11): 2862–2864.
- [17] KÖSE C, IKIBAŞŞ C. A personal identification system using retinal vasculature in retinal fundus images[J]. *Expert Systems with Applications*, 2011, 38(11): 13670–13681.
- [18] MARIÑO C, PENEDO M G, PENAS M, et al. Personal authentication using digital retinal images[J]. *Pattern Analysis and Applications*, 2006, 9(1): 21–33.
- [19] NARASIMHA-IYER H, MAHADEVAN V, BEACH J M, et al. Improved detection of the central reflex in retinal vessels using a generalized dual-Gaussian model and robust hypothesis testing[J]. *IEEE Transactions on Information Technology in Biomedicine*, 2008, 12(3): 406–410.
- [20] FORACCHIA M, GRISAN E, RUGGERI A. Luminosity and contrast normalization in retinal images[J]. *Medical Image Analysis*, Elsevier, 2005, 9(3): 179–190.
- [21] SIGURSSON 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*, North-Holland, 2014, 47: 164–171.
- [22] 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.
- [23] FRANGI A F, NIESSEN W J, VINCKEN K L, et al. Multiscale vessel enhancement filtering[G]//Springer, Berlin, Heidelberg, 2006: 130–137.
- [24] PALOMERA-PÉREZ M A, MARTINEZ-PEREZ M E, BENÍTEZ-PÉREZ H, et al. Parallel multiscale feature extraction and region growing: Application in retinal blood vessel detection[J]. *IEEE Transactions on Information Technology in Biomedicine*, 2010, 14(2): 500–506.
- [25] BANKHEAD P, SCHOLFIELD C N, MCGEOWN J G, et al. Fast retinal vessel detection and measurement using wavelets and edge location refinement[J]. SERRANO-GOTARREDONA T. *PLoS ONE*, Public Library of Science, 2012, 7(3): e32435.
- [26] 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.
- [27] 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.
- [28] COHEN L D, KIMMEL R. Global minimum for active contour models: A minimal path approach[J]. *International Journal of Computer Vision*, Kluwer Academic Publishers, 1997, 24(1): 57–78.
- [29] 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*, Elsevier, 2012, 108(2): 600–616.
- [30] FRAZ M M, RUDNICKA A R, OWEN C G, et al. Delineation of blood vessels in pediatric retinal images using decision trees-based ensemble classification[J]. *International Journal of Computer Assisted Radiology and Surgery*, Springer Berlin Heidelberg, 2014, 9(5): 795–811.
- [31] 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.
- [32] FAN Z, RONG Y, LU J, et al. Automated blood vessel segmentation in fundus image based on integral channel features and random forests[C]//Proceedings of the World Congress on Intelligent Control and Automation. Institute of Electrical and Electronics Engineers Inc., 2016: 2063–2068.
- [33] FAN Z, MO J. Automated blood vessel segmentation based on de-noising auto-encoder and neural network[C]//Proceedings of the International Conference on Machine Learning and Cybernetics. IEEE, 2017, 2: 849–856.
- [34] ORLANDO J I J I J I, PROKOFYEVA E, BLASCHKO M B. A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images[J]. *IEEE transactions on Biomedical Engineering*, IEEE, 2016, 64(1): 16–27.
- [35] MARÍN D, AQUINO A, GEGÚNDEZ-ARIAS M E, et al. A New Supervised Method for Blood Vessel Segmentation in Retinal Images by Using Gray-Level and Moment Invariants-Based Features[J]. *IEEE Transactions on Medical Imaging*, 2011, 30(1): 146–158.
- [36] IAN G, YOSHUA B, AARON C. *Deep Learning*[M]. MIT Press, 2016.
- [37] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. *Nature*, Nature Publishing Group, 2015, 521(7553): 436–444.
- [38] WANG S, YIN Y, CAO G, et al. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning[J]. *Neurocomputing*, Elsevier, 2015, 149: 708–717.
- [39] LISKOWSKI PAWEŁAND KRAWIEC K, LISKOWSKI P, KRAWIEC K, et al. Segmenting retinal blood vessels with deep neural networks[J]. *IEEE transactions on medical imaging*, IEEE, 2016, 35(11): 2369–2380.
- [40] MANINIS K-K, PONT-TUSET J, ARBELÁEZ P, et al. Deep Retinal Image Understanding[C]//OURSELIN S, JOSKOWICZ L, SABUNCU M R, et al. *Medical Image Computing and Computer-Assisted Intervention -- MICCAI 2016*. Cham: Springer International Publishing, 2016: 140–148.
- [41] FU H, XU Y, LIN S, et al. DeepVessel: Retinal Vessel Segmentation via Deep Learning and Conditional

- Random Field[C]//OURSELIN S, JOSKOWICZ L, SABUNCU M R, et al. International conference on medical image computing and computer-assisted intervention. Cham: Springer International Publishing, 2016: 132–139.
- [42] DASGUPTA A, SINGH S. A fully convolutional neural network based structured prediction approach towards the retinal vessel segmentation[C]//Proceedings - International Symposium on Biomedical Imaging. IEEE, 2017: 248–251.
- [43] MO J, ZHANG L. Multi-level deep supervised networks for retinal vessel segmentation[J]. International Journal of Computer Assisted Radiology and Surgery, 2017, 12(12): 2181–2193.
- [44] 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.
- [45] 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. 2015: 234–241.
- [46] XIAO X, LIAN S, LUO Z, et al. Weighted Res-UNet for High-Quality Retina Vessel Segmentation[J]. Proceedings - 9th International Conference on Information Technology in Medicine and Education, ITME 2018, 2018: 327–331.
- [47] LAHIRI A, AYUSH K, BISWAS P K, et al. Generative Adversarial Learning for Reducing Manual Annotation in Semantic Segmentation on Large Scale Microscopy Images: Automated Vessel Segmentation in Retinal Fundus Image as Test Case[C]//IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops. 2017, 2017-July: 794–800.
- [48] 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, 2004, 23(4): 501–509.
- [49] HOOVER A. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response[J]. IEEE Transactions on Medical Imaging, 2000, 19(3): 203–210.
- [50] OWEN C G, RUDNICKA A R, MULLEN R, et al. Measuring retinal vessel tortuosity in 10-year-old children: validation of the computer-assisted image analysis of the retina (CAIAR) program[J]. Investigative ophthalmology & visual science, The Association for Research in Vision and Ophthalmology, 2009, 50(5): 2004–2010.
- [51] KOLAR R, KUBENA T, CERNOSEK P, et al. Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database[J]. IET Image Processing, 2013, 7(4): 373–383.
- [52] MCCULLOCH W S, PITTS W. A logical calculus of the ideas immanent in nervous activity[J]. The Bulletin of Mathematical Biophysics, Kluwer Academic Publishers, 1943, 5(4): 115–133.
- [53] IZHIKEVICH E M. Simple model of spiking neurons[J]. IEEE Transactions on Neural Networks, 2003, 14(6): 1569–1572.
- [54] MAAS A L, HANNUN A Y, NG A Y. Rectifier nonlinearities improve neural network acoustic models[C]//Proceedings of Machine Learning Research. 2013, 30(1): 3.
- [55] HE K, ZHANG X, REN S, et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification[C]//Proceedings of the IEEE International Conference on Computer Vision. 2015: 1026–1034.
- [56] VON MELCHNER L, PALLAS S L, SUR M. Visual behaviour mediated by retinal projections directed to the auditory pathway[J]. Nature, Nature Publishing Group, 2000, 404(6780): 871–876.
- [57] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278–2323.
- [58] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet Classification with Deep Convolutional Neural Networks[G]//PEREIRA F, BURGESS C J C, BOTTOU L, et al. Advances in Neural Information Processing Systems 25. Curran Associates, Inc., 2012: 1097–1105.
- [59] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition[J]. arXiv preprint arXiv:1409.1556, 2014.
- [60] SZEGEDY C, LIU W, JIA Y, et al. Going Deeper with Convolutions (GoogleLeNet)[C]//Journal of Chemical Technology and Biotechnology. 2016, 91(8): 2322–2330.
- [61] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the

- IEEE conference on computer vision and pattern recognition. 2016: 770–778.
- [62] HUANG G, LIU Z, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 4700–4708.
- [63] LONG J, SHEHMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 3431–3440.
- [64] G L D. Distinctive Image Features from Scale-Invariant Keypoints[J]. *International Journal of Computer Vision*, Springer, 2004, 60(2): 91.
- [65] CHEN Y, FANG H, XU B, et al. Drop an Octave: Reducing Spatial Redundancy in Convolutional Neural Networks with Octave Convolution[J]. *arXiv preprint arXiv:1904.05049*, 2019, 1.
- [66] JAPKOWICZ N, STEPHEN S. The class imbalance problem: A systematic study[J]. *Intelligent data analysis*, IOS Press, 2002, 6(5): 429–449.
- [67] KINGMA D P, BA J. Adam: A method for stochastic optimization[J]. *arXiv preprint arXiv:1412.6980*, 2014.
- [68] ZANA F, KLEIN J-C. Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation[J]. *IEEE transactions on image processing*, IEEE, 2001, 10(7): 1010–1019.
- [69] MENDONCA A M, CAMPILHO A. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction[J]. *IEEE transactions on medical imaging*, IEEE, 2006, 25(9): 1200–1213.
- [70] AL-DIRI B, HUNTER A, STEEL D. An active contour model for segmenting and measuring retinal vessels[J]. *IEEE Transactions on Medical imaging*, IEEE, 2009, 28(9): 1488–1497.
- [71] MIRI M S, MAHLOOJIFAR A. Retinal image analysis using curvelet transform and multistructure elements morphology by reconstruction[J]. *IEEE Transactions on Biomedical Engineering*, IEEE, 2010, 58(5): 1183–1192.
- [72] YOU X, PENG Q, YUAN Y, et al. Segmentation of retinal blood vessels using the radial projection and semi-supervised approach[J]. *Pattern Recognition*, Elsevier, 2011, 44(10–11): 2314–2324.
- [73] FATHI A, NAGHSH-NILCHI A R. Automatic wavelet-based retinal blood vessels segmentation and vessel diameter estimation[J]. *Biomedical Signal Processing and Control*, Elsevier, 2013, 8(1): 71–80.
- [74] SREEJINI K S, GOVINDAN V K. Improved multiscale matched filter for retina vessel segmentation using PSO algorithm[J]. *Egyptian Informatics Journal*, Elsevier, 2015, 16(3): 253–260.
- [75] ROYCHOWDHURY S, KOOZEKANANI D D, PARHI K K. Iterative vessel segmentation of fundus images[J]. *IEEE Transactions on Biomedical Engineering*, IEEE, 2015, 62(7): 1738–1749.
- [76] CHENG E, DU L, WU Y, et al. Discriminative vessel segmentation in retinal images by fusing context-aware hybrid features[J]. *Machine vision and applications*, Springer, 2014, 25(7): 1779–1792.
- [77] ORLANDO J I, BLASCHKO M. Learning fully-connected CRFs for blood vessel segmentation in retinal images[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. 2014: 634–641.
- [78] VEGAR, SANCHEZ-ANTE G, FALCON-MORALES L E, et al. Retinal vessel extraction using Lattice Neural Networks with dendritic processing[J]. *Computers in Biology and Medicine*, Pergamon, 2015, 58: 20–30.
- [79] LI Q, FENG B, XIE L, et al. A cross-modality learning approach for vessel segmentation in retinal images[J]. *IEEE transactions on medical imaging*, IEEE, 2015, 35(1): 109–118.
- [80] HOOVER A, KOUZNETSOVA V, GOLDBAUM M. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response.[C]//Proceedings of the AMIA Symposium. 1998: 931.

附录A DRIVE 数据库测试集详细实验结果

	ACC	SE	SP	F1	AUROC
1	0.964211	0.870812	0.973338	0.812465	0.987231
2	0.966236	0.838802	0.980796	0.835925	0.986380
3	0.956635	0.667089	0.988573	0.753481	0.967355
4	0.967209	0.778888	0.986209	0.813212	0.976252
5	0.965137	0.768123	0.985587	0.805590	0.978641
6	0.961231	0.717085	0.987563	0.782670	0.975725
7	0.965481	0.730608	0.989311	0.795886	0.976916
8	0.965797	0.744105	0.986716	0.789550	0.977921
9	0.967915	0.746177	0.987476	0.790380	0.980735
10	0.966591	0.805994	0.980954	0.798426	0.980912
11	0.962730	0.796675	0.979003	0.792340	0.977612
12	0.967564	0.813521	0.982088	0.812101	0.985407
13	0.962830	0.766860	0.984155	0.801959	0.977047
14	0.968288	0.846013	0.979128	0.812901	0.986130
15	0.970142	0.847692	0.979548	0.801993	0.986303
16	0.968498	0.823045	0.982933	0.825102	0.987359
17	0.966751	0.762386	0.985718	0.795696	0.983030
18	0.968014	0.855755	0.977632	0.808522	0.987777
19	0.970478	0.897508	0.977105	0.835049	0.990026
均值	0.966067	0.795736	0.982697	0.803288	0.981833
标准差	0.003321	0.058350	0.004398	0.018434	0.005816

附录B STARE 数据库详细实验结果

	ACC	SE	SP	F1	AUROC
1	0.969513	0.822765	0.982244	0.811634	0.987500
2	0.974136	0.755678	0.989702	0.795364	0.977733
3	0.980423	0.815160	0.990932	0.832747	0.991535
4	0.965931	0.699121	0.987329	0.752912	0.984591
5	0.966999	0.775611	0.986101	0.810089	0.988431
6	0.972141	0.845725	0.981624	0.809037	0.990188
7	0.977615	0.897780	0.984566	0.865302	0.994581
8	0.976822	0.893063	0.983608	0.852410	0.994221
9	0.978405	0.872147	0.987464	0.863863	0.993864
10	0.971073	0.867965	0.980078	0.828178	0.990797
11	0.974800	0.810308	0.987396	0.820606	0.990165
12	0.979469	0.869667	0.988604	0.866782	0.995187
13	0.976246	0.853009	0.988303	0.864874	0.994013
14	0.977299	0.869041	0.988113	0.874274	0.994151
15	0.973431	0.844569	0.985654	0.846339	0.991441
16	0.962673	0.758500	0.985803	0.805278	0.982547
17	0.973240	0.836860	0.986544	0.847536	0.989206
18	0.982624	0.800060	0.992342	0.823141	0.987573
19	0.982372	0.795261	0.990776	0.795014	0.990971
20	0.968586	0.646472	0.991692	0.733676	0.975118
均值	0.974190	0.816438	0.986994	0.824953	0.989191
标准差	0.005509	0.064650	0.003361	0.037742	0.005520

附录C CHASE_DB1 数据库详细实验结果

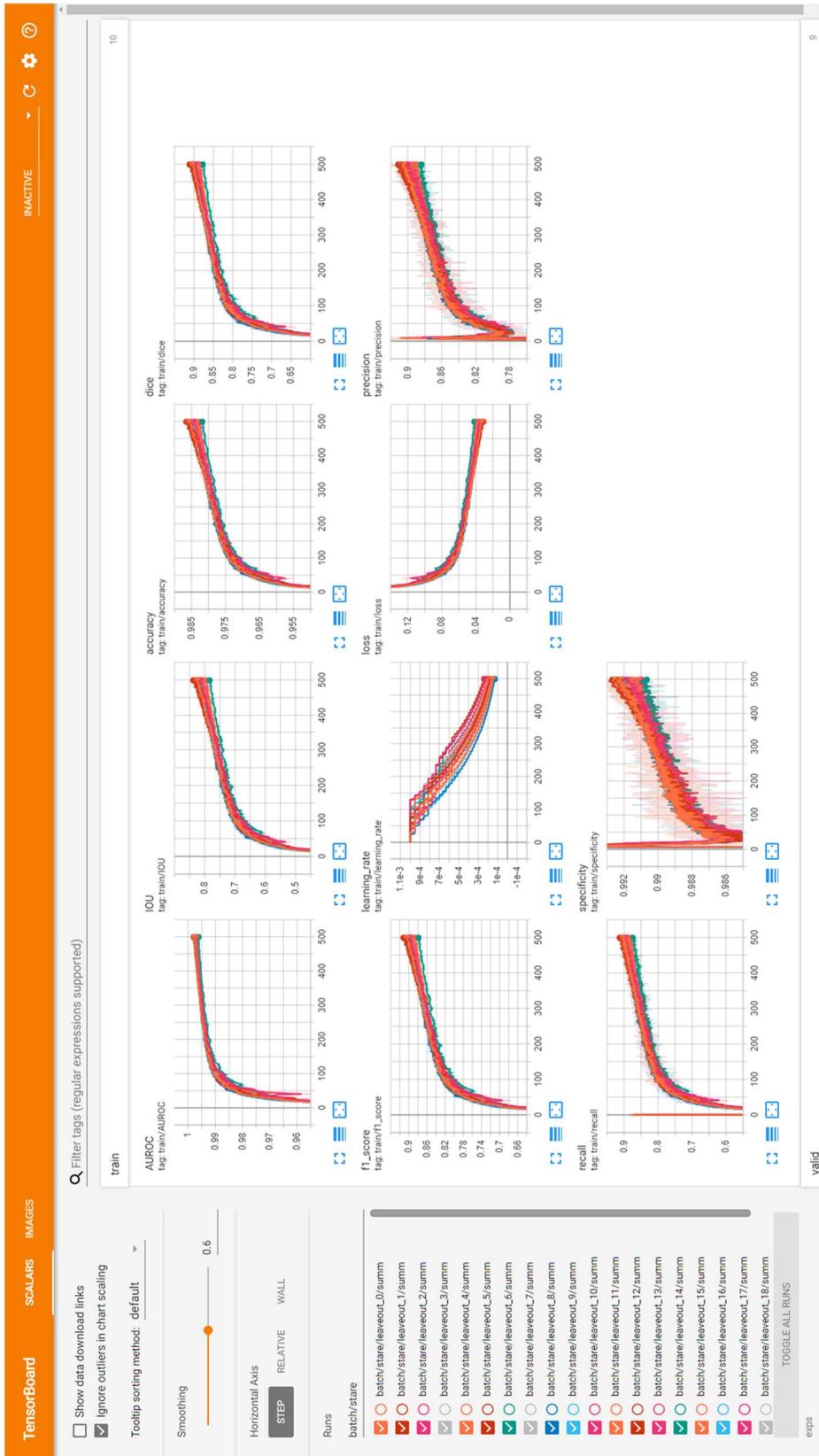
	ACC	SE	SP	F1	AUROC
1	0.980991	0.887499	0.987976	0.866504	0.993482
2	0.978065	0.804069	0.992363	0.847722	0.991885
3	0.972481	0.835641	0.984542	0.831062	0.988077
4	0.973911	0.797107	0.988882	0.826698	0.987601
5	0.975838	0.852825	0.986299	0.846931	0.990941
6	0.976227	0.855022	0.986131	0.844573	0.992191
7	0.977844	0.881950	0.986020	0.862148	0.993532
8	0.975609	0.808109	0.989518	0.835545	0.986158
9	0.976601	0.851029	0.987421	0.852300	0.992112
10	0.975883	0.839259	0.988553	0.855213	0.990487
11	0.974834	0.791499	0.989859	0.826527	0.983554
12	0.976692	0.832916	0.988288	0.842127	0.989494
13	0.975288	0.831828	0.987445	0.840239	0.989525
14	0.975719	0.827826	0.987918	0.838595	0.990102
15	0.976494	0.813101	0.987797	0.817390	0.989995
16	0.975510	0.835115	0.985034	0.812478	0.990752
17	0.980724	0.822549	0.989152	0.811925	0.988704
18	0.979919	0.737659	0.993043	0.790596	0.984280
19	0.976810	0.804298	0.988344	0.812983	0.989335
20	0.971542	0.822500	0.981380	0.781616	0.983749
21	0.980835	0.838030	0.988860	0.823099	0.989002
22	0.981674	0.844564	0.989398	0.830952	0.992964
23	0.968811	0.854384	0.977649	0.797091	0.986031
24	0.973072	0.840921	0.983618	0.821936	0.990606
25	0.976471	0.810965	0.987318	0.809158	0.989735
26	0.973591	0.741978	0.989166	0.779768	0.982529
27	0.977493	0.819042	0.989228	0.833850	0.991944
28	0.976097	0.802826	0.986879	0.797374	0.985955
均值	0.976251	0.824447	0.987432	0.826300	0.989097
标准差	0.002940	0.033417	0.003061	0.023189	0.003093

附录D HRF 数据库详细实验结果

	ACC	SE	SP	F1	AUROC
1	0.968990	0.771142	0.979617	0.717124	0.981423
2	0.967381	0.787676	0.980244	0.763373	0.983022
3	0.968029	0.838188	0.982759	0.842335	0.983163
4	0.972141	0.774110	0.985503	0.778408	0.985915
5	0.970875	0.817253	0.983203	0.806547	0.987327
6	0.975998	0.876713	0.986886	0.878340	0.993692
7	0.974327	0.806704	0.985108	0.791563	0.986294
8	0.975170	0.790826	0.986409	0.785429	0.987067
9	0.967648	0.814583	0.985705	0.841616	0.986713
10	0.976868	0.790767	0.988050	0.794886	0.984711
11	0.973957	0.784258	0.987006	0.794930	0.986138
12	0.975483	0.875283	0.985927	0.870796	0.993817
13	0.973972	0.826882	0.983692	0.797507	0.987961
14	0.976574	0.828319	0.986932	0.822005	0.990246
15	0.974113	0.783921	0.992981	0.845344	0.990785
16	0.959465	0.712489	0.979556	0.725625	0.962333
17	0.974648	0.817623	0.986161	0.815029	0.989675
18	0.975021	0.850168	0.989099	0.873381	0.992467
19	0.970310	0.799606	0.984662	0.806849	0.987076
20	0.976219	0.807371	0.987918	0.814806	0.987060
21	0.977818	0.871016	0.988228	0.874604	0.993741
22	0.968670	0.794542	0.982858	0.792596	0.983487
23	0.975227	0.830617	0.985758	0.819881	0.989088
24	0.969120	0.770962	0.991227	0.833652	0.983069
25	0.966862	0.714189	0.985185	0.744538	0.979774
26	0.974220	0.789979	0.987225	0.801620	0.984322
27	0.977108	0.786800	0.993162	0.842464	0.988797
28	0.958759	0.808554	0.973161	0.774305	0.974294
29	0.974262	0.773646	0.988314	0.797369	0.987936

	ACC	SE	SP	F1	AUROC
30	0.972557	0.760498	0.992516	0.826626	0.981132
31	0.968636	0.816114	0.982750	0.815086	0.987151
32	0.966759	0.783567	0.981936	0.782942	0.973842
33	0.976902	0.869313	0.987987	0.875481	0.993411
34	0.964390	0.724002	0.981922	0.734326	0.967373
35	0.968594	0.804345	0.983334	0.808370	0.982521
36	0.968460	0.806115	0.987440	0.842532	0.986274
37	0.972549	0.782005	0.986602	0.796470	0.986878
38	0.966793	0.775867	0.980744	0.760884	0.973728
39	0.976028	0.862525	0.986854	0.862374	0.992751
40	0.965855	0.767663	0.982199	0.774039	0.981175
41	0.966827	0.760336	0.982576	0.764617	0.979110
42	0.973652	0.865176	0.984011	0.851299	0.991553
43	0.968395	0.826948	0.978344	0.774711	0.985493
44	0.968052	0.798110	0.981526	0.785888	0.982682
45	0.978550	0.793366	0.994361	0.853350	0.994412
均值	0.971383	0.802003	0.985281	0.807909	0.985131
标准差	0.004702	0.039377	0.004157	0.041109	0.006799

附录E 模型训练过程可视化



TensorBoard

SCALARS

IMAGES

INACTIVE

⚙️

🔄

🔍 Filter tags (regular expressions supported)

valid

valid/analytic
step: 490
Mon May 06 2019 16:51:15 GMT+0800 (China Standard Time)

valid/analytic
step: 490
Mon May 06 2019 18:39:53 GMT+0800 (China Standard Time)

valid/images
step: 490
Mon May 06 2019 16:51:15 GMT+0800 (China Standard Time)

valid/images
step: 490
Mon May 06 2019 18:39:52 GMT+0800 (China Standard Time)

Show actual image size

Brightness adjustment RESET

Contrast adjustment RESET

Runs

batch/stare

- batch/stare/leaveout_0/summm
- batch/stare/leaveout_1/summm
- batch/stare/leaveout_2/summm
- batch/stare/leaveout_3/summm
- batch/stare/leaveout_4/summm
- batch/stare/leaveout_5/summm
- batch/stare/leaveout_6/summm
- batch/stare/leaveout_7/summm
- batch/stare/leaveout_8/summm
- batch/stare/leaveout_9/summm
- batch/stare/leaveout_10/summm
- batch/stare/leaveout_11/summm
- batch/stare/leaveout_12/summm
- batch/stare/leaveout_13/summm
- batch/stare/leaveout_14/summm
- batch/stare/leaveout_15/summm
- batch/stare/leaveout_16/summm
- batch/stare/leaveout_17/summm
- batch/stare/leaveout_18/summm
- batch/stare/leaveout_19/summm

TOGGLE ALL RUNS

exps

致谢

时光如白驹过隙，转眼之间三年的研究生生活即将结束。值此论文付梓之际，借此简短篇幅，向在我整个研究生学习、生活期间所有关心和帮助我的人们致以我最真诚的谢意！

感谢我的导师范衡教授，在我整个研究生生活期间给予的悉心指导和热情帮助。导师渊博的知识，严谨务实的治学态度，求真务实的工作作风，宽容大度的长者作风使我受益匪浅；导师平易近人，坦荡的胸襟，谦虚的为人更是我学习的楷模。在此谨向导师致以崇高的敬意和衷心的感谢！

感谢李文姬，容毅标，谢红辉，游煜根，卢杰威，伍宇明，邱本章等实验室小组成员。他们对学术研究的热枕，深深的感染了我。在每周的学术讨论会上，他们让我开阔了视野，极大地调动了我的学习积极性。

感谢同班的各位同学，在汕头大学的三年里，我们一起学习，一起生活，互相陪伴，互相扶持，共同进步，我将永记我们一起奋斗的这段美好时光。

感谢我的家人，你们的全力支持和殷切期望，让我有勇气面对困境，让我在学术路上没有后顾之忧。在此，对你们致以最诚挚的敬意和衷心的祝福！

攻读硕士学位期间的主要工作成果

- [1] FAN Z, MO J. Automated blood vessel segmentation based on de-noising auto-encoder and neural network[C]//Proceedings of the International Conference on Machine Learning and Cybernetics. IEEE, 2017, 2: 849–856.
- [2] FAN Z, RONG Y, LU J, et al. Automated blood vessel segmentation in fundus image based on integral channel features and random forests[C]//Proceedings of the World Congress on Intelligent Control and Automation. Institute of Electrical and Electronics Engineers Inc., 2016, 2016-September: 2063–2068.
- [3] MO J, FAN Z, LI W, et al. Multi-factorial evolutionary algorithm based on M2M decomposition[C]//SHI Y, TAN K C, ZHANG M, et al. Asia-Pacific Conference on Simulated Evolution and Learning. Cham: Springer International Publishing, 2017: 134–144.
- [4] FAN Z, LI W, CAI X, et al. An improved epsilon constraint-handling method in MOEA/D for CMOPs with large infeasible regions[J]. Soft Computing, Springer Verlag, 2017: 1–20.
- [5] Li Wenji, Fang Yi, Mo Jiajie and Fan Zhun, Evolutionary Multi-task Optimization (Multi-Objective Track) Competition Runner-Up in IEEE 2017 Congress on Evolutionary Computation (CEC2017)
- [6] 范衡, 卢杰威, 谢红辉, 朱贵杰, 莫嘉杰, 李文姬. 一种基于深度卷积神经网络的烟草植株识别方法[P]. 中国发明专利, 申请号: 201710404295.2
- [7] 范衡, 伍宇明, 卢杰威, 朱贵杰, 莫嘉杰, 黄龙涛. 一种基于深度学习的蝴蝶自动分类方法[P]. 中国发明专利, 申请号: 201811070920.5