

摘 要

眼底图像超分辨率重建和血管分割关键技术的研究

眼底图像在医学诊断中扮演着重要作用,可用于多种眼部疾病和全身疾病的诊断和筛查。在眼底图像中包含多种组织结构如微血管瘤、黄斑和血管等,其中血管形态特征是医生诊断的重要依据,诊断重点包括血管的走向、长度、宽度、曲折度和分支角度。在临床中,医生通常需要手工标注眼底图像中的血管,然而这一做法存在一定程度的局限性。首先,由于成像效果的限制以及医生专业水平的不同,手工标注存在主观性和不确定性,尤其在微细血管的标注上可能存在漏误。其次,手工标注的过程需要耗费大量时间,造成医生诊断效率低下,同时可能导致急重症患者的病情延误。目前,计算机视觉领域的超分辨率重建和图像分割技术被认为是解决成像效果和手工标注问题的关键。

为了解决眼底图像成像效果不足和数据集数据量少、分辨率较低的问题,本文首先提出了一种基于 ZSSR 的图像超分辨率重建方法。该方法的意义是对分割网络使用的数据集进行超分辨率重建,在解决成像图像模糊、感兴趣区域不显著等问题的同时,使经过预处理和数据增强后的图像块能以更大尺寸和更高的清晰度送入分割网络,提升分割精度。在模型中引入本文提出的残差信息蒸馏模块,在不增加网络深度的前提下提取不同层次的特征并保留浅层特征,使模型更加轻量化。通过特征融合方式组合本文提出的多重大核注意力块,使各深度特征被融合提取,实现特征的精炼与增强。采用 ZSSR 提出的内部图片训练方式对网络进行训练,在实现无监督学习的同时,进一步提升了针对数据量较少的眼底图像超分辨率任务的性能表现。通过在眼底图像数据集上进行实验,证明了该方法较其他经典方法在眼底图像上具有更好的超分辨率重建效果。

在认真研究眼底图像医学特征和现有语义分割方法的基础上,本文提出了一种基于 U-Net 的眼底图像血管分割网络 DSENet。在该网络中,本文将数据集进行灰度化和 CLAHE 预处理,增强图像的对比度使血管细节部分较背景更加突出。提出了大核空间金字塔(LKSP)模块,通过不同大小感受野的卷积核获取图像的多尺度信息。引入了多视角注意力(MVA)块,以捕捉多视角交叉维度的相互作用,增强编码器的特征同时辅助解码器生成更准确的

掩模图。在多个眼底图像数据集上采用消融实验方式证明了本文提出模块的有效性，同时综合比较 DSENet 与经典网络进行血管分割任务时的定量与定性表现，证明本文提出网络具有更高的分割精度。最后，采用第 3 章提出的超分辨率方法对数据集进行进一步处理，发现较不使用超分辨率方法时分割精度明显提升，证明了本文提出的超分辨率重建方法与血管分割方法相结合的优势。

关键词：

眼底图像，血管分割，超分辨率重建，深度学习，注意力机制

Abstract

The Research on the Key Technology of Super-resolution Reconstruction and Vessel Segmentation for Fundus Image

Fundus images play an important role in medical diagnosis and can be used to diagnose and screen for a variety of ocular and systemic diseases. Fundus images contain a variety of tissue structures such as microvascular tumors, macula and blood vessels, among which the morphological characteristics of blood vessels are an important basis for diagnosis. In clinical practice, physicians usually need to manually label blood vessels in fundus images, however, this practice has some degree of limitations. First, due to the limitation of imaging effect and the professional level of doctors, manual labeling is subjective and uncertain, especially in the labeling of microscopic vessels, which may be missed. Secondly, the manual annotation process takes a lot of time, resulting in inefficient diagnosis and possible delays for patients with acute illnesses. Currently, super-resolution reconstruction and image segmentation techniques in the field of computer vision are considered to be the key to solve the problems of imaging effects and manual annotation.

In order to solve the problems of insufficient imaging effect of fundus images and low data volume and resolution of datasets, this paper firstly proposes a ZSSR-based image super-resolution reconstruction method. The significance of this method is to perform super-resolution reconstruction on the dataset used by the segmentation network, and to solve the problems of blurred imaging images and insignificant regions of interest while enabling the image blocks after pre-processing and data enhancement to be fed into the segmentation network with larger size and higher resolution to improve the segmentation accuracy. The residual information distillation module proposed in this paper is introduced in the model to extract features at different levels and retain shallow features without increasing the network depth, making the model more lightweight. The multi-significant kernel attention blocks

proposed in this paper are combined by feature fusion so that each depth feature is fused and extracted to achieve feature refinement and enhancement. The network is trained using the internal image training method proposed by ZSSR, which further improves the performance performance for the super-resolution task of fundus images with less data volume while achieving unsupervised learning. By conducting experiments on fundus image datasets, it is demonstrated that the method has better super-resolution reconstruction results on fundus images than other classical methods.

Based on the careful study of medical features of fundus images and existing semantic segmentation methods, this paper proposes a U-Net-based vascular segmentation network DSENet for fundus images, in which the dataset is grayed out and CLAHE preprocessed to enhance the contrast of the image to make the vascular detail part more prominent than the background. The Large Kernel Space Pyramid (LKSP) module is proposed to obtain multi-scale information of images by convolutional kernels with different sizes of sensory fields. Multi-view attention (MVA) block is introduced to capture the interactions of multi-view cross dimensions to enhance the features of the encoder while assisting the decoder to generate more accurate mask maps. The effectiveness of the proposed module is demonstrated using ablation experiments on several fundus image datasets, and the quantitative and qualitative performance of the DSENet is also compared with the classical network for the vascular segmentation task in a comprehensive manner, demonstrating that the proposed network has higher segmentation accuracy. Finally, the data set was further processed using the super-resolution method proposed in Chapter 3, and it was found that the segmentation accuracy was significantly improved compared with that without the super-resolution method, demonstrating the advantages of combining the super-resolution reconstruction method proposed in this paper with the vessel segmentation method.

Key words:

Fundus image, Blood vessel segmentation, Super-resolution reconstruction, Deep learning, Attention mechanism

目 录	
第 1 章 绪 论	1
1.1 研究背景与研究意义	1
1.2 国内外研究现状	2
1.2.1 图像超分辨率重建	2
1.2.2 眼底图像血管分割	4
1.3 本文主要工作	6
1.4 论文结构	6
第 2 章 相关理论介绍	8
2.1 计算机视觉与深度学习	8
2.1.1 卷积神经网络基础	8
2.1.2 注意力机制	10
2.1.3 图像超分辨率重建网络	12
2.1.4 医学图像分割网络	14
2.2 眼底图像相关数据集	16
2.3 本章小结	17
第 3 章 基于 ZSSR 的图像超分辨率方法	18
3.1 引言	18
3.2 方法	19
3.2.1 模型的整体结构	19

3.2.2 残差信息蒸馏块	20
3.2.3 多重注意力特征融合提取模块	21
3.2.4 内部图片训练方式	23
3.2.5 模型损失函数	24
3.3 实验	24
3.3.1 数据集处理	24
3.3.2 评价指标	25
3.3.3 训练细节	26
3.3.4 消融实验	27
3.3.5 实验结果	31
3.4 本章小结	36
第4章 基于 U-Net 的眼底图像血管分割网络	37
4.1 引言	37
4.2 方法	37
4.2.1 模型的整体结构	37
4.2.2 大核空间金字塔模块	38
4.2.3 多视角注意力	40
4.2.4 模型损失函数	41
4.2.5 数据集预处理	42
4.3 实验	43
4.3.1 评价指标	43

4.3.2 训练细节	44
4.3.3 图像灰度化与 CLAHE 处理效果实验	45
4.3.4 消融实验	46
4.3.5 实验结果	48
4.3.6 结合超分辨率方法后实验结果	51
4.4 本章小结	53
第 5 章 总结与展望	54
5.1 全文工作总结	54
5.2 未来工作展望	54
参考文献	56
作者简介及科研成果	62
致 谢	63

第1章 绪论

1.1 研究背景与研究意义

眼底图像在医学诊断中扮演着重要作用，它可用于对眼部疾病如青光眼、白内障、黄斑病变等进行诊断，也可以用于其他疾病如肾炎、心脑血管疾病、糖尿病等的筛查。眼底图像中包含微血管瘤、黄斑和血管等结构，其中血管形态特征的病变是医生诊断的重要依据，主要检查点是血管的走向、长度、宽度、曲折度和分支角度。在临床中，医生通常需要手工标注眼底图像中的血管，然而这一做法有着一些局限性。首先，由于医生个人水平的不同以及成像效果的限制，手工标注存在主观性和不确定性，尤其在微细血管的标注上可能存在漏误。其次，手工标注的过程需要耗费大量时间，造成医生诊断效率低下，同时可能导致急重症患者的病情延误。目前，计算机视觉领域的超分辨率重建和图像分割技术被认为是解决成像效果和手工标注问题的关键。

图像超分辨率重建指将一幅低分辨率图像恢复成对应的高分辨率图像，其目的是克服或补偿由于图像采集系统或采集环境本身的限制导致的成像图像模糊、质量低下、感兴趣区域不显著等问题。现有的超分辨率重建技术更多研究的是自然图像，而自然图像与眼底图像存在一定程度的差异。自然图像中噪声分布绝大多数情况下可以认为是均匀的，可以近似为高斯噪声，因为直射和散射造成光场分布可以被认为是均匀的，但在眼底图像中由于单一光源、探测手段和人体的影响，噪声可以近似为泊松噪声，因此在自然图像降噪问题上表现良好的方法并不简单的适用于眼底图像。此外，眼底图像某些细微结构的变化可能代表着某些病变，且这些变化常常被隐藏在相似度极高的背景中，因此在自然图像超分辨率重建中表现良好的方法在眼底图像上未必适用。且由于超分辨率重建本质上是一种图像数据插补技术，对丢失数据的过度补充可能造成“无中生有”的问题，这一问题发生在在极度重视细节的眼底图像上将造成严重的后果。

医学图像分割的主要任务是从医学图像中分割出兴趣区域，如特定器官部位或兴趣目标（如血管、肿瘤）等。目前，基于深度学习的分割方法展现出较传统方法更优的性能，被广泛应用于医学图像的分割任务上，并表现出了较好的性能，但应用在眼底图像的血管分割任务时仍存在一些不足之处。首先，带有专家标注血管标签的眼底图像公开数据集较

少,且数据集内数据量较少、分辨率较低,即使进行分块、旋转、翻转等数据增强操作后数据量仍有所不足,不利于复杂网络的训练。其次,较传统的医学图像分割网络如 U-Net^[1]及其改进型等虽然已被应用至眼底图像,然而其性能表现仍存在一定程度的提升空间,而较新的网络如 ResUNet++^[2]等在眼底图像上表现如何尚不明朗。最后,应用于眼底图像血管分割任务的网络为了解决数据集不足的问题,通常进行分块等数据增强操作,但受限于常用开源数据集的分辨率因素,这通常会使得送入网络的图像块分辨率很小,而经验表明送入网络的图像块大小一定程度上会影响网络的性能,在眼底图像血管分割问题上主要表现在微细血管的分割效果上。综上所述,针对于眼底图像的超分辨率重建及血管分割方法在网络模型结构和最终的性能表现上仍需进行突破,具有重要的研究意义。

1.2 国内外研究现状

1.2.1 图像超分辨率重建

1. 传统超分辨率重建方法

在过去的研究中,超分辨率重建领域有许多出色的工作,如基于插值的超分辨率重建方法、基于稀疏字典的超分辨率重建方法和最大后验概率估计方法。基于插值的这类方法最为传统,主要采用基函数和插值和对图像高频信息进行计算和填充,典型方法包括最近邻插值法、双线性插值法和双三次插值法^[3]。插值方法相较如今流行的深度学习方法重构效果较差,但其也拥有重构时间短、不易产生过多无关细节的特点,因此对于眼底图像的超分辨率重建仍有一定的研究意义。

基于稀疏字典的方法本质上是一种浅层学习方法,主要学习 LR (Low Resolution) 图像与 HR (High Resolution) 图像之间的映射关系与规律,如 J. T. Y. Yang^[4]等提出的经典方法,通过计算图像低维空间的投影线性关系完成超分辨率重建。然而,这类方法由于需要大量图像以构建 LR、HR 图像的稀疏字典,虽然重建效果相对较好,但计算复杂度较高、运算效果差。

最大后验概率估计方法主要基于贝叶斯理论,是一种统计学方法,其基本思想是使用观测公式对模型进行观测,进而采用观测得到的 LR 退化图像来估计出 HR 清晰低噪图像,使得 HR 图像的后验概率最大。这类估计方法具有良好的灵活性和鲁棒性,是应用较多的

一类传统超分辨率图像重建方法。

2. 基于深度学习的超分辨率重建方法

目前,将针对自然图像的深度学习算法迁移至医学图像的研究上并加以改进是一种比较常见的做法。在以下论文发表之前,已有学者将在自然图像上表现良好的 SRCNN^[5]、ESPCN^[6]、SRGAN^[7]、ESRGAN^[8]等经典模型加以改进后应用到医学图像领域,取得了较传统算法更加优越的表现。

Gao^[9]等提出了一种更适用于医学图像无纹理、具有独特重复结构特点的网络,它相比于基准的 SRCNN 网络将单个的卷积层替换为两个连续卷积层,采用重叠池化层对数据降维,采用跳跃连接的结构提高网络的收敛性能,根据 VDSR^[10]的经验加深网络层数,在眼底图像数据集上表现出了良好效果。FMISR^[11]由 Zhang 等于 2018 年提出,主要针对 CT 图像的超分辨率重建问题,基于 SRCNN 的三层网络架构进行改进,使用亚像素卷积解决了 SRCNN 插值预处理和计算量大的问题,同时用堆叠的小卷积核代替大卷积核有效提高了重建速度。MDRN^[12]由 Ren 等于 2019 年提出,该网络将去除 BN 层的残差块应用到了超分辨率模型中,在包括眼底图像的多种医学影像上取得了良好效果。ISRCNN^[13]由 Yan Lv 等于 2021 年提出,该网络在 SRCNN 的基础上在特征映射部分添加了 DRR(Dense and Rezero Residual)模块,该模块可以通过局部密集连接充分利用 DRR 块中所有层的信息,之后利用局部残差学习融合保留的自适应特征。

除了提高网络的重建效果,如何使网络更加轻量化也是目前的研究热点,许多论文针对这一问题进行了大量工作。IDN^[14]由 Hui Z 等于 2018 年提出,主要创新点是提出了信息蒸馏结构对浅层特征进行分离,使得一部分特征被进一步处理以提取深层特征,另一部分被蒸馏出来并送入更深层,这种方式可以更有效的提取局部的短路径与长路径特征。

IMDN^[15]在 IDN 研究的基础上提出了两大创新点,一是提出了信息多蒸馏结构在不加深网络层数的前提下更有效提取深度特征,二是通过实验发现网络的深度是影响网络运行速度的主要因素。MAFFSRN^[16]由特征融合组 (Feature Fusion Group, FFG) 组成,该网络取得了 2020 年自然图像超分辨率重建的 SOTA 成绩,在性能表现优秀的同时较为轻量化,易于在样本量较少的数据集上进行训练。LESRCNN^[17]由 Tian C 等于 2021 年提出,通过一种异构的结构将图像的高频和低频特征信息有效结合提取,防止了深层网络易丢失低频特征的问题,并采用一组自适应的子像素卷积计算实现盲超分辨率重建。

1.2.2 眼底图像血管分割

眼底图像血管分割方法可大致分为基于传统图像处理方法、基于机器学习的方法和基于深度学习的方法。基于传统图像处理的方法无需训练、实时性高，缺点是对血管分割的泛化性差、分割精度不佳。基于机器学习的方法针对特定数据集分割效果不错，缺点是依赖专家个人能力对数据集针对性选择特征学习模板和判别器，在其他数据集上的泛用性差。基于深度学习的方法不针对某一特定数据集，拥有良好的泛化性、鲁棒性和分割精度，因此逐渐在眼底图像血管分割领域占据了主导地位。

1. 基于传统图像处理的眼底图像血管分割方法

用于眼底图像血管分割的传统图像处理方法主要包括匹配滤波器法、形态学处理法和血管跟踪法等。Chaudhuri^[18]等采用图像的光学和空间特征提取算子，结合匹配滤波器算法完成血管的分段检测分割。Zhang^[19]等提出一种基于3D旋转帧的李氏群函数滤波器，借助小波变换和垂直于线性结构的多尺度二阶高斯导数将结构分解为对应平面并增强血管特征。Hoove^[20]等将局部与全局特征协同，实现眼底图像中血管的分割。Zana^[21]等采用数学形态和曲率评估的方式，提高了血管检测的准确率和鲁棒性。Chutatape^[22]等设计了一种跟踪过程中的分支检测方法，使眼底图像中血管能更加完整的检测出。

2. 基于机器学习的眼底图像血管分割方法

Yambal^[23]等对传统模糊C均值聚类方法进行改进，在其目标函数中加入了空间特征，以无监督学习的方式完成分割任务。You^[24]等采用径向投影和半监督方法降低了血管分割任务的误判率，分割效果良好。Orland^[25]等对条件随机场模型进行改进，并使用结构化输出的支持向量机模型进行自主学习和参数优化，有效提升了分割性能。Han^[26]等结合灰度和形态特征，采用多尺度局部统计方式对血管特征进行提取。Soras^[27]等使用不同尺度的二维Gabor小波和像素值建立特征向量空间，并通过贝叶斯或高斯混合模型分类器划分背景和前景。Lupasc^[28]将像素点表示为代表滤波器输出和数值估计等测量值集合的41维特征向量，并训练不同的弱分类器最终整合成强分类器。

3. 基于深度学习的眼底图像血管分割方法

随着深度学习方法在人脸识别、超分辨率、目标检测等领域表现出更强的提取和表达特征的能力，以卷积神经网络(CNN)为主的各种浅层、小型神经网络在医学图像分割领域逐渐占据了主导地位。在此之后，从自然语言处理领域迁移至计算机视觉领域的

Transformer 表现出超越传统 CNN 网络的优越性能, 以 TransUNet^[29]为代表开始进军医学图像分割领域。

FCN^[30]将传统 CNN 的全连接层替换为了卷积层, 通过跳跃连接改善其粗糙度, 并通过转置卷积上采样使得网络可以输出与输入图像相同尺寸的分割图, 开创了在语义分割领域应用深度学习的先河。SegNet^[31]在 FCN 的基础上进行改进, 编码器使用 VGG-16 网络并采用最大池化索引对解码器输入进行非线性上采样, 采用 Softmax 分类器对解码器输出进行像素级分类, 可用于解决医学图像语义分割问题。

Olaf Ronneberger 等于 2015 年提出了 U-Net^[1]网络, 采用类似 FCN 的编码器-解码器架构, 在同层编码器-解码器的跳跃连接处采用了拼接(concat)操作, 解码器部分采用上卷积, 该网络参数量少不易过拟合、可用于多模态图像, 高度适用于医学图像数据集, 后续有许多研究者基于其进行改进, 其改进型几乎占据了医学图像分割领域的半壁江山。

Ozan Oktay 等^[32]在 U-Net 编码器与解码器跳跃连接的位置添加了注意力门(Attention Gate, AG), 用于将同一层级的编码器信息进行注意力加权后再输入至同层解码器中。Zongwei Zhou 等^[33]将简单的跳跃连接修改为稠密连接以实现多层次深度的编码器特征融合, 并通过在测试时剪枝的方式验证网络模型的最优深度, 使出现在多个尺度的病变能够更准确的分割。Luo^[34]等修改了 U-Net 网络的跳跃连接部分, 在编码器-解码器跳跃连接的基础上新增了编码器、解码器内部的跳跃连接, 并采用全局随机场将全局信息集成。Jin^[35]等在 U-Net 基础上添加了可变形卷积^[36], 以根据血管的大小和形状对其自适应捕获, 使最终分割的血管细节丰富。Wu^[37]等将 U-Net 与生成对抗网络的思想进行融合, 并添加密集连接和 Attention Gate 使分割性能大幅提升。

除了上述基于全卷积或 U-Net 的方法, 基于多标签处理或多尺度连接以及其他类型的方法同样在眼底图像的血管分割领域占有一席之地。Ma^[38]等采用空间激活机制设计网络的输出模块, 并在其中集成了一些常用图像预处理方法, 使网络能够对眼底图像中动脉和静脉血管自动分割。Fu^[39]等将多尺度和多层卷积网络共同使用, 并利用条件随机场和模拟像素的远程交互将分割任务转化为边界检测任务。Guo^[40]等提出了一种多尺度深度学习网络, 在侧面输出层添加短连接将底层语义信息传递给浅层的同时降低噪声。Yan^[41]等针对眼底图像中粗细血管数量级差异, 将粗细血管在两个架构的网络中分别训练并优化了训练损失。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/118027124072006046>