



计算机工程与科学  
Computer Engineering & Science  
ISSN 1007-130X, CN 43-1258/TP

## 《计算机工程与科学》网络首发论文

题目: 多尺度深度监督的高度近视萎缩病变分割算法  
作者: 曾增峰, 环宇翔, 邹卓, 郑立荣  
收稿日期: 2020-03-26  
网络首发日期: 2020-10-21  
引用格式: 曾增峰, 环宇翔, 邹卓, 郑立荣. 多尺度深度监督的高度近视萎缩病变分割算法[J/OL]. 计算机工程与科学.  
<https://kns.cnki.net/kcms/detail/43.1258.TP.20201020.0952.002.html>



**网络首发:** 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

**出版确认:** 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

# 多尺度深度监督的高度近视萎缩病变分割算法

曾增峰, 环宇翔, 邹卓, 郑立荣

(复旦大学信息科学与工程学院微纳系统中心, 上海 200000)

**摘要:**为提升眼底图像的高度近视萎缩病变分割精度, 针对不同个体的眼底图像质量良莠不齐及因萎缩病变与相邻组织之间边界较为模糊等引起分割困难的问题, 提出具有多尺度深度监督思想的高度近视萎缩病变分割方法。首先开发优化算法使得眼底图像组织结构清晰、风格统一, 降低复杂特征的区分难度。由于 V-Net 只能够得到较低的分割精度, 因此, 通过融合高层与低层的特征形成多尺度特征学习的 MS-V-Net, 能够提取不同尺度图像中语义信息。更为重要的是, 最终对 MS-V-Net 每个多尺度模块的深度监督形成紧密监督的 MSS-V-Net, 与原始 V-Net 分割方法相比, 提高了网络对重要语义信息的判别性及泛化能力。实验结果表明, 本文方法的 Dice 盒图呈现异常值变少, 中位数变大, 盒子长度变短, 上下间隔变小, 盒外的两条线变短的趋势, 说明有效提升了高度近视萎缩病变图像的分割精度。

**关键词:**深度学习; 多尺度深度监督; 高度近视; 图像分割

**中图分类号:** TP391

**文献标志码:** A

## Segmentation algorithm for high myopia atrophy lesions based on multi-scale depth supervision

ZENG Zeng-feng, HUAN Yu-xiang, ZOU Zhuo, ZHENG Li-rong

(Micro nano System Center, School of Information Science and Technology,  
Fudan University, Shanghai 200000, China)

**Abstract:** In order to improve the segmentation accuracy of atrophic lesions of high myopia in fundus images, aiming at the problems of poor quality of the fundus images of different individuals and the difficulty of segmentation due to the blurred border between atrophic lesions and adjacent tissues, high myopia with multi-scale depth supervision is proposed Atrophy lesion segmentation method. Firstly, an optimization algorithm is developed to make the fundus image organization structure clear and uniform in style, reducing the difficulty of distinguishing complex features. Because V-Net can only obtain lower segmentation accuracy, the MS-V-Net that combines multi-level and low-level features to form multi-scale feature learning can extract semantic information in images of different scales. More importantly, the deep supervision of each multi-scale module of MS-V-Net eventually forms a closely supervised MSS-V-Net. Compared with the original V-Net segmentation method, it improves the discrimination of important semantic information by the network and generalization ability. The experimental results show that the Box-plot of Dice of the method in this paper presents a trend of fewer outliers, larger median, shorter box length, smaller upper and lower interval, and shorter two lines outside the box, effectively improving the segmentation of atrophic lesion images of high myopia.

**Key words:** deep learning; multi-scale depth supervision; high myopia; image segmentation

## 1 引言

由于眼底相机的成像能够清晰地显示眼底组织信息,并在眼底图像上较好地显示出病变,而广泛应用于眼科医学成像领域。从眼底图像中精确地分割出高度近视萎缩病变,是现代科学技术辅助诊断高度近视及手术规划的关键步骤。然而,可靠而精确地分割高度近视萎缩病变在眼底图像分割的研究中还存在以下问题:

(1) 不同个体的眼底图像质量良莠不齐,萎缩病变形状、大小有很大差异;

(2) 萎缩病变与相邻组织之间边界较为模糊,难以区分。

存在该问题图像如图 1 所示。针对上述问题,如何提升算法的分割精度是医学图像科研人员所面临的极具挑战性任务<sup>[1]</sup>。为此,本文将开发一种基于灰度变换的眼底图像优化算法。

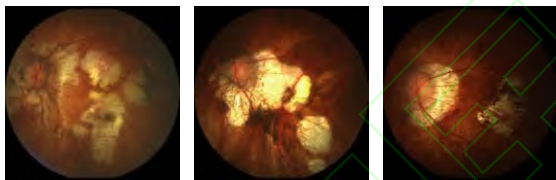


Figure1<sup>[2]</sup> Fundus image display  
图 1<sup>[2]</sup> 眼底图像展示

目前,深度学习已经被广泛应用到解决计算机视觉和医学图像语义分割等领域的众多问题。Shelhamer 等<sup>[3]</sup>提出了一个全卷积神经网络(Fully Convolutional Network, FCN)的图像语义分割方法,该方法用 $1 \times 1$ 的卷积层代替全连接层,使得可以网络接受任意尺寸的输入图像,并通过反卷积使得经过多次卷积缩小的特征图能够恢复到原始图像尺寸大小,同时更好地保留了图像空间信息。Ronneberger 等<sup>[4]</sup>、Drozdz 等<sup>[5]</sup>和 Abdulkadir 等<sup>[6]</sup>对 U-Net 分割方法做了详细的描述和论证,并列举了各自方法的优缺点和原理差异,通过上述三个文献可以得出,现有 U-Net 结构所应用的方法都是使用 2D 卷积层和 2D 转置卷积层建立下采样层和上采样层组合的对称结构,可在下采样层和上采样层之间添加跳跃连接操作来提升分割精度,并能够在卷积层中添加了 2D 注释断层图做 3D 医学图像疾病分割。Chen 等<sup>[7]</sup>用膨胀卷积核代替经典的卷积核,解决了下采样降低特征图大小和网络参数较多的问题。上述的分割算法,假如图像中目标特征与背景差异明显情况下可以取得不错的分割效果,但当特征差异不明显时,分割

精度则有待提高。另外,深度学习算法一般要对图像进行预处理,预处理的好坏会直接影响算法的分割精度。

受深度学习在多学科交叉领域取得丰硕研究成果的启发,研究人员将其应用到了医学图像处理与分析中。在实践的过程中,针对医学图像中的不同任务,分别获得了较好的结果。大多数的分割方法往往需要一定的知识和经验积累,采用一些复杂的预处理操作和辅助其他方法,如 K-means<sup>[8]</sup>、支持向量机<sup>[9]</sup>等,但 these 方法不能充分利用图像中丰富的语义信息。为了更好地辅助医生对高度近视的诊断与治疗,提升深度学习和医疗技术结合的应用前景。本文构建了一种具有多尺度深度监督的高度近视萎缩病变分割算法,可以较好地解决上述问题,完成高度近视萎缩病变眼底图像的精准分割,如图 2 所示。



(a)原始图像 (b)标签 (c)图像标签可视化  
Figure2<sup>[2]</sup> Data set label example  
图 2<sup>[2]</sup> 数据集标签示例

## 2 相关方法介绍

### 2.1 V-Net

Fausto Milletari 等<sup>[10]</sup>提出了一种只需要少量数据就可训练的网络模型,并提出 Dice 作为损失函数,该网络模型结构如图 3 所示。研究表明,V-Net 只需少量的医学图像数据集就能够得到了不错的分割精度,并且网络的训练速度很快,为此广泛应用于医学图像分割领域。在医学图像语义分割的应用中,V-Net 能够做到端对端训练、像素对像素的分类,即便没有关于数据集的预处理,也可以取得较好的分割效果,可见其具有很好的医学图像特征空间的表征能力。V-Net 网络结构以及特征学习过程如图 3 所示,且详细结构及参数可参考其文献。

V-Net 的网络结构由下采样路径和上采样路径组成,并且整个网络都使用全 0 补齐方式的卷积使得图像的输入和输出尺寸一致。

V-Net 左边的下采样路径是十分经典的卷积神经网络,它是由卷积块结构组成的,每个卷积块结构中均有多个卷积,类似 Resnet<sup>[11]</sup>网络中的残差结构来解决深度学习梯度消失的问



题,卷积层中卷积核大小均为 $5 \times 5$ ,采用 PRelu 作为激活函数,采用卷积核的步数来控制特征图尺寸的大小。每一次卷积块操作后特征图的通道数量都翻倍,但是特征图尺寸减半。

V-Net 右边的上采样路径的每一个卷积块之前都是反卷积,除了第一个反卷积不改变特征图的通道数量,别的反卷积都使得特征图的通道数量减半并使特征图尺寸加倍。反卷积过后,将反卷积的结果与下采样路径中对应大小的特征图拼接起来,组成跳跃拼接的结构。对拼接后的特征图进行多次 $5 \times 5$ 的卷积,其类似下采样路径中的卷积块。上采样路径中的最后 1 层的卷积核大小为 $1 \times 1$ ,将特征图的 32 个通道转化为分类个数,例如本文图像分割分为前景和背景,所以分类个数为 2<sup>[12]</sup>。

V-Net 结构与 U-Net 结构十分类似,看起来呈 V 形状,故叫 V-Net。通过对比 V-Net、U-Net 的网络结构,可以得出两者主要差别为 V-Net 在上采样和下采样的每个阶段都采用了 ResNet 的短路连接方式,其相当于在 U-Net 中引入了 ResBlock,可以一定程度上解决深度网络退化问题,缓解梯度弥散问题,即残差连接使得信息前后向传播更加顺畅。同时,现有研究表明 V-Net 性能更好,并且 V-Net 具有上述结构与优点,使其十分适合本文的眼底图像分割任务,为此,作者决定改进原有 V-Net 的网络结构来优化分割任务。

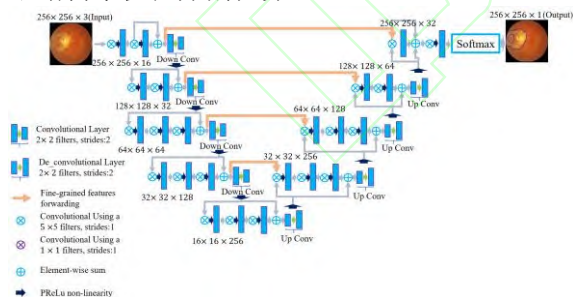


Figure3<sup>[2]</sup> V-net

图 3<sup>[2]</sup> V-Net

## 2.2 特征金字塔网络

Girshick R 等<sup>[13]</sup>提出一种特征金字塔网络 (Feature Pyramid Networks, FPN), 该 FPN 网络的大概结构如图 4 所示。

FPN 左边的下采样路径是典型的深度卷积神经网络,通过卷积操作使得特征图大小保持不变或者减半。网络右边的上采样路径的每一步都采用反卷积,使得特征图尺寸保持不变或者加倍。反卷积过后,将反卷积结果与上采样中对应大小一致的特征图拼接起来。

由于特征金字塔网络能够从不同尺度图像提取语义信息,充分利用低层特征和高层特征。为此作者决定将其运用到 V-Net 中组成具有多尺度融合功能的 Multi-Scale V-Net(MS-V-Net),即在 V-Net 的网络结构上改变一些连接,在增加少量参数计算量情况下来提高网络的分割精度。

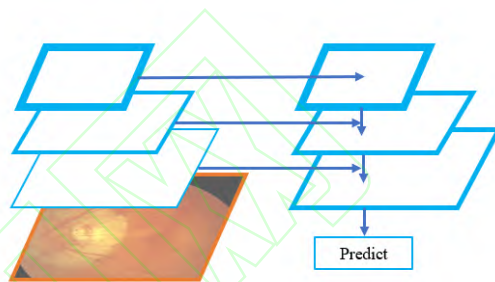


Figure4<sup>[2]</sup> Feature Pyramid Networks

图 4<sup>[2]</sup> 特征金字塔网络

## 2.3 深度监督网络

Lee 等<sup>[14]</sup>提出一种深度监督网络 (Deeply-supervised nets, DSN), 该 DSN 网络的大概结构如图 5 所示。

DSN 左边的不同卷积块引入了分类目标函数,对多个连续卷积块学习到的信息特征提供直接监督,不像传统的深度卷积网络主要依赖最后分类层计算预测误差再进行反向传播来指导前面层的学习,该方法每个卷积块之间形成紧密监督的形式。

由于深度监督网络能够对前面层的学习起到监督的作用,从而保证整个网络的学习效果。为此,作者决定将其运用到在 MS-V-Net 中组成了具有多尺度深度监督功能的 Multi-Scale Supervised V-Net(MSS-V-Net),即在 MS-V-Net 的网络结构上添加多个目标函数来提高网络的分割精度。

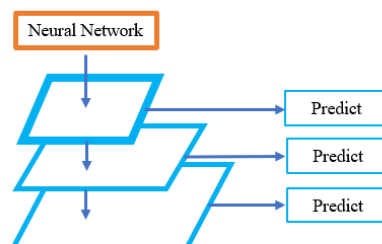


Figure5 Deeply-Supervised Nets

图 5 深度监督网络

### 3 眼底图像的优化算法

由于不同个体的眼底图像质量差异较大及因萎缩病变与相邻组织之间边界较为模糊等问题。研究现有眼底图像优化算法<sup>[15-17]</sup>的原理,大部分对眼底图像的局部采用对比度增强、色彩校正和直方图均衡化等方法。这些方法尽管具有一定的优化作用,但结果仍不令人满意。另外,可以使用美图秀秀、GIMP 和 PhotoShop 等图像处理软件,人工处理达到图像优化的效果,但是该方法需要耗费巨大人力,并且效率十分低下。况且不同人员对像素敏感性差异较大,修图效果存在差别,无法保证图像效果一致,这对于深度学习是一个巨大障碍。为解决上述问题,本文构建眼底图像的优化算法旨在使眼底图像组织结构清晰、质量风格统一。

#### 3.1 算法原理

为了更好的描述算法原理,本文需要定义中央区域和非中央区域两个概念:

**定义 1** 图像灰度直方图中去掉灰度值最大、最小那部分所对应像素坐标点的区域为中央区域(Central area, Ca)。

**定义 2** 图像灰度直方图中非中央区域那部分坐标点的区域为非中央区域(Non-Central area, N-Ca)。

根据定义 1 和定义 2 获得眼底图像的优化算法步骤如下:

#### 算法 眼底图像的优化算法

**输入:** 待优化图像。

**输出:** 优化后图像。

#### 步骤 1 调整图像局部区域的灰度;

在输入待优化图像后,首先需要分离眼底图像的 RGB 颜色通道,并获取它们各自的灰度直方图。其次,令眼底图像的眼底区域作为感兴趣区域,并保证只在该区域进行操作。最后,在每个颜色通道的灰度直方图中提取感兴趣区域,并对其中央区域的灰度值做灰度变换。灰度变换公式如下所示。

$$A_t = \frac{(A_o - Q_o) \times C_t}{C_o} + Q_t \times CE_c \quad (1)$$

其中,  $A_o$  和  $A_t$  分别为变换前后对应坐标点的灰度值,  $Q_o$  和  $Q_t$  分别为变换前后图像的灰度极小值,  $C_o$  和  $C_t$  为变换前后图像的对比度,  $CE_c$  为不同通道的灰度变换权重系数。

#### 步骤 2 平衡图像全部区域的亮度。

经过调整眼底图像中央区域的灰度后,由于非中央区域的灰度值没有经过任何处理,导致中央区域和非中央区域的灰度值会存在较大差异,为此,首先将非中央区域的灰度最大最小值替换为变换后中央区域的灰度最大最小值。其次,对整幅图像的所有像素点的灰度值叠加上平衡值( $O_a$ ),该值为实验经验值。最后,将优化后的三个颜色通道的灰度图像融合成彩色的眼底图像。

#### 3.2 实验结果及分析

本文的实验数据由两部分构成,即公开数据集 Grand Challenge of PALM<sup>[2]</sup>以及 Grand Challenge of ODIR2019<sup>[18]</sup>。本节实验环境为: Windows10 64 位操作系统, CPU i5-9300H, 软件编程环境为 Anaconda Spyder -Python3.6。

在眼底图像的优化算法实验过程中,其中几个关键步骤所包含的重要参数设置为, B、G、R 三个通道对应的  $O_a$ 、 $CE_c$ 、 $Q_t$ 、 $C_o$  值分别为 38、0.9、0、64; 90、1.0、0、200; 180、0.9、112、240。

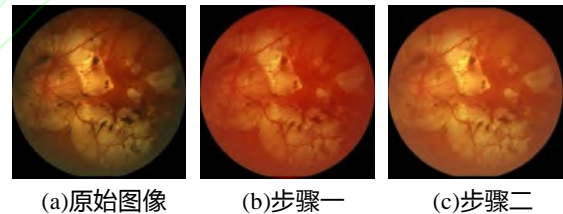


Figure 6<sup>[2]</sup> Image optimization display

图 6<sup>[2]</sup> 图像优化展示

从图 6 中可以看出,图 6(b)为在经过算法步骤 1 后的图像,可以看出该图像相对于原始图像已经起到一定的优化效果,但是由于还没对非中央区域的灰度最大最小值进行替换,图像呈现较为深红并略显失真。图 6(c)为经过算法步骤 2 后的图像,眼底图像已经变得组织结构清晰、质量风格统一。两个步骤处理后的图像相对于图 6(a)原始图像都呈现出较好的图像质量。

图 7 展示了病理性高度近视的三张眼底图像优化前后对比。可以看出,优化算法明显提高了图像质量的一致性,使得眼底图像较为清晰显示多种生理结构,具体表现为眼底边缘轮廓更为清晰,主要是清除了眼底的黑色边缘,更清晰显示眼底中的血管、视神经盘等组织结构,并且让萎缩病变与相邻组织之间边界的区别较为明显。

由此可以得出, 该算法为优化眼底图像提供了一种较为新颖的方法, 具有提高眼底图像的可读性和统一眼底图像色调及清晰显示眼底组织的作用。

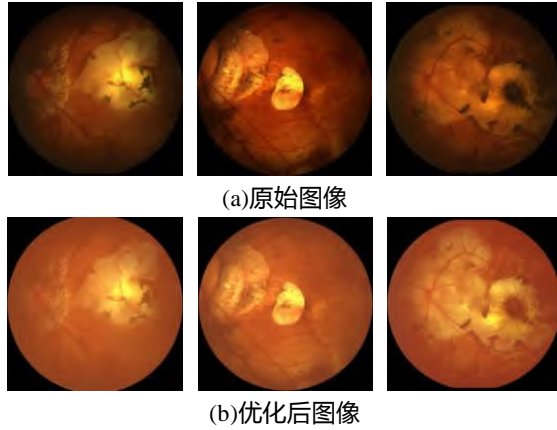


Figure 7<sup>[2]</sup> Before and after image optimization

图 7<sup>[2]</sup> 图像优化前后对比

## 4 基于 V-Net 的 MSS-V-Net

在传统方法需要较为复杂的图像预处理, 如需要预先用形态学提取高度近视病变特征等操作, 这些预处理导致网络分割效果偏差较大。为此, 本文研究多尺度深度监督条件下的高度近视病变分割算法, 相对于传统的分割方法具有较大优势, 深度学习方法只需提供分割标签就能够自动学习并取得较好分割结果。

### 4.1 MS-V-Net

如图 8 所示, 相对于 V-Net 网络结构, 可以看出 MS-V-Net 网络在 V-Net 网络上采样路径中, 额外增加了一条上采样路径, 即 MS-V-Net 分为两条上采样路径。第一条上采样路径为原始 V-Net 网络的上采样路径; 第二条上采样路径为多尺度上采样路径, 其是由第一条上采样路径所引申出来的, 在第二条上采样路径的反卷积之前都需要对第一条上采样路径的特征图进行  $1 \times 1$  的卷积, 再将反卷积结果与第一条上采样路径中对应大小一致的特征图拼接起来, 并且为了消除拼接操作可能导致的混叠效应, 需要对拼接后的特征图再卷积的操作, 最终经过 Softmax 函数进行像素分类, 从而可以分割出萎缩病变区域。

在测试阶段, 同 V-Net 网络一样采用最后一层的分割结果作为预测结果。

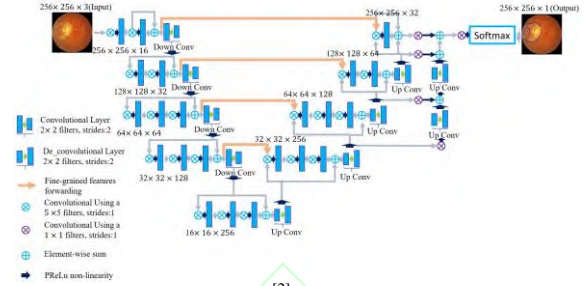


Figure 8<sup>[2]</sup> MS-V-Net

图 8<sup>[2]</sup> MS-V-Net

### 4.2 MSS-V-Net

由 2.3 节的深度监督思想可知, 卷积神经网络前面层学习效果越好越有利于后面层提取信息特征, 即前面层提取特征越丰富, 则越能够保证后面层具有稳健的学习能力。为此, 如图 9 所示, 在 MS-V-Net 网络结构的基础上, 分别在第二条路径的每个阶段引入 Softmax 目标函数, 即每个阶段都将进行语义分割, 三个 Softmax 目标函数分别对应三个 Dice 损失函数, 实现多个损失函数共同调整模型的作用, 能够提升模型的泛化能力和鲁棒性。

具体来说, MSS-V-Net 利用多尺度特征融合后每个阶段信息引入分类目标函数做不同尺度的语义分割, 每个阶段形成紧密监督的形式, 训练过程中实现了能够在不同多尺度层独立预测的功能。

在训练的过程中, 会存在三个 Dice 损失函数, 重要性权重分别设置为 0.2、0.2 和 0.6, 但是只监控最后一层分类输出的 Dice 数值, 并只保证该数值达到最大才存储权重模型。

在测试阶段, 由于实验经验表明最后层的结果特征提取能力最强, 前面层只是起到辅助提取特征的作用, 并监督最后一层达到较好的分割结果。因此, 并不使用前面几个尺度层的测试结果, 而是采用最后层的分割结果作为 MSS-V-Net 的预测结果。

对于 MSS-V-Net 的结构, 引入深层监督目标函数主要有以下两个方面的作用:

(1) 深度监督可以起到正则项的作用, 提升网络对重要信息特征的判别性。

(2) 深度监督的作用可以理解为一程度上解决了深度学习中的梯度消失和梯度爆炸等问题, 提高深度学习网络对重要信息特征的泛化性能力。



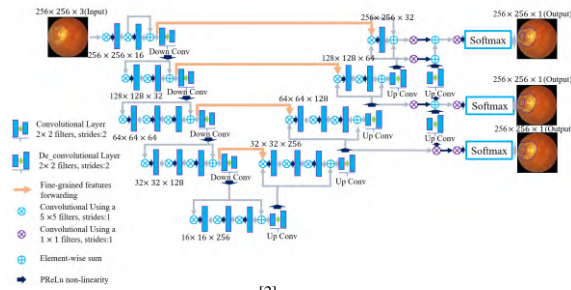
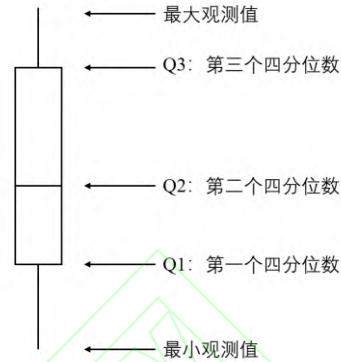
Figure 9<sup>[2]</sup> MSS-V-Net图 9<sup>[2]</sup> MSS-V-Net

Figure 10 Box-plot

图 10 盒图

## 5 实验及评估标准

### 5.1 实验平台

本文实验环境为: Windows10 64 位操作系统, 使用 NVIDIA GeForce GTX 1060Ti 显卡, 核心频率为 2.30GHz, 显存为 6G, 软件编程环境为 Python3.6, CUDA 显卡 GPU 运算平台和 cuDNN 神经网络 GPU 加速库, 基于 Tensorflow 后端的 Keras 深度学习框架。

### 5.2 评价指标

定义 1 将训练样本中的每幅图像序号记为  $n$ ,  $n = 1, 2, 3, \dots, m$ 。其中,  $n$  为训练数据集和验证集的图像数量。

定义 2 第  $m$  幅图像的萎缩病变像素的真实数目为  $G(m)$ 。

定义 3 第  $m$  幅图像的萎缩病变像素预测数目为  $P(m)$ 。

定义 4 由于文献<sup>[10]</sup>已经验证, Dice 损失函数能够更好处理前景和背景像素的数目之间存在强烈的不平衡的情况, 所以采用 Dice 作为评价指标, 其公式定义如下所示:

$$Dice(P, G) = \frac{2 \times N(\Delta_P \cap \Delta_G) + Smooth}{N(\Delta_P) + N(\Delta_G) + Smooth} \quad (2)$$

其中,  $P$  和  $G$  分别表示预测结果和金标准标签;  $\Delta_P$  和  $\Delta_G$  表示  $P$  和  $G$  包含的闭合区域;  $N(\cdot)$  表示区域中的像素总数;  $Smooth = 1$ , 具有防止除 0 的作用;  $Dice(\cdot)$  的数值越大则表示算法的分割精度越高。

定义 5 盒图一种用作显示一组数据分散情况资料的统计图, 能够直观显示数据离散程度, 如图 10 所示。

其中, Q1 和 Q3 之间的距离表示的是被数据中间一半所覆盖的范围, 该距离称为四分位数极差(IQR), 定义为:

$$IQR = Q3 - Q1 \quad (3)$$

在分析数据的过程中, 借助盒图可以有效地识别数据的特性<sup>[19]</sup>:

(1) 盒图的下端点是第一个四分位数上, 上端点是第三个四分位数上, 使得盒的长度是第三个四分位数和第一个四分位数的极差 IQR。

(2) 盒内的线标记了第二个四分位数, 即中位数。

(3) 盒外的两条线(亦称作胡须)延伸到最小观测值和最大观测值。

本文用盒图作为评价指标衡量测试数据集的 Dice 分布情况, 在处理 Dice 观测值时, 盒图能够绘制出可能的离群点。仅当最大观测值和最小观测值超过四分位数不到  $1.5 \times IQR$  时, 值得个别地绘出可能的离群点。否则, 胡须则在出现在四分位数的  $1.5 \times IQR$  之内的最极端的观测值处终止<sup>[20]</sup>。

### 5.3 实验过程

实验数据采用随机挑选的方法将数据集划分为训练集 260 张、验证集 30 张和测试集 35 张。在调整超参数时, 参考相关文献及实践经验, 从训练集随机抽取的一份小规模数据集对超参数进行大致的网络探索, 以节省模型参数调整的时间。本文的参数设置为: Adam 优化算法, 0.0001 基本学习率, 4Mini-batch, 0.00001 权重衰减因子, 90 次最大迭代次数。

尽管深度学习分割网络降低了数据预处理难度, 只需批量输入图像数据, 就能够自动学习萎缩病变特征, 达到病变分割的目的。但是由于训练深度学习网络需要较多图像数据, 并为了避免过拟合, 本文采用数据增广方法来

扩增高度近视数据集,实践了以下三种数据扩增方法:

(1) 翻转。本文采用水平翻转的数据扩增方法。

(2) 图像亮度的调整,将随机生成亮度值叠加到原图像的像素值上。

(3) 旋转和尺度变换。在一定角度之间随机地选择一个角度进行旋转和对图像坐标做尺度变换。

为了更好地说明 MSS-V-Net 模型的图像语义分割效果,三个模型分别在原始数据集优化数据集上进行 6 组对比实验。

在训练过程中,每个 epoch 都用验证集对模型进行评估,并只保留 Dice 数值最大的模型作为最终模型。

#### 5.4 实验结果及分析

在本文的实验过程中,训练集和验证集 Dice 损失函数,在每个模型训练的初始阶段存在较大的差异,但是最终模型都达到收敛状态且收敛时 Dice 值相差不大,无法判断模型好坏。为此,本实验主要用测试集判定模型的泛化能力。

在测试集上,分别提取 V-Net、MS-V-Net 和 MSS-V-Net 在原始图像和优化图像的轮廓并计算 Dice 的盒图作为评价指标。

图 11(a、i)为两张原始图像(对应优化后图像没有展示),分别对应图 11(e、m)标签,图 11(b、c、d)和图(j、k、l)分别对应为两组 V-Net、MS-V-Net 和 MSS-V-Net 网络在原始图像的分割结果,图 11(f、g、h)和图(n、o、p)分别对应为两组 V-Net、MS-V-Net 和 MSS-V-Net 网络在优化图像的分割结果。从图 11 可以看出, MSS-V-Net 和 MS-V-Net 两个方法的分割结果相较于 V-Net 的分割结果,更好地拟合高度近视萎缩病变的轮廓,存在较少的误分割连通轮廓,这些都初步验证了本文方法在高度近视萎缩病变分割上的可行性和有效性。

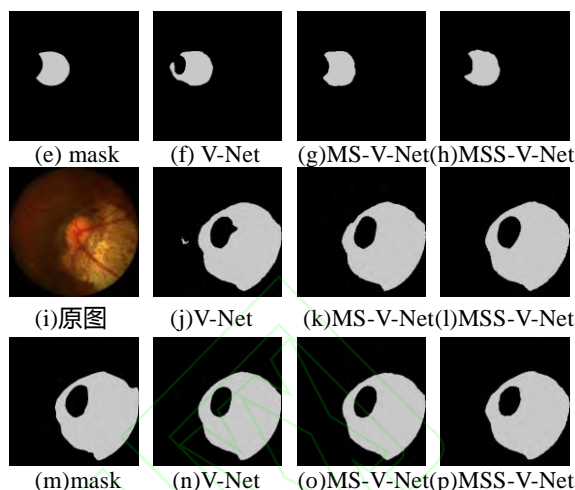
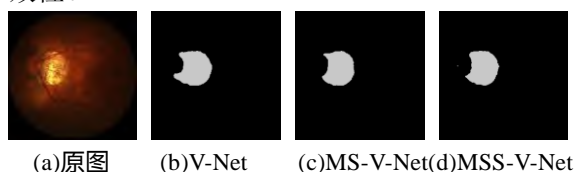


Figure 11<sup>[2]</sup> The segmentation results of the data set

图 11<sup>[2]</sup> 数据集的分割结果

为了进一步验证本文方法,从统计学评价指标来分析模型的分割精度。在图 12 中可知,只有 MSS-V-Net 和 MS-V-Net 两个新颖网络在优化数据的情况下不存在离群点, V-Net 在原始、优化数据集上分别有 1、3 个离群点, MSS-V-Net、MS-V-Net 在原始数据集上都有 2 个离群点。图 12 中 V-Net、MS-V-Net 和 MSS-V-Net 的 Dice 盒图都呈现中位数变大,盒子长度变短,上下间隔变小,盒外的两条线长度变短的趋势。可知,总体上 MSS-V-Net 和 MS-V-Net 两个模型的鲁棒性均高于 V-Net,并且 MSS-V-Net 也比 MS-V-Net 具有更好的泛化能力和分割精度。

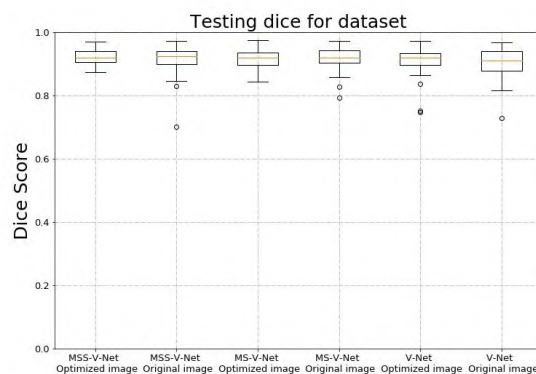


Figure 12 Box-plot of Dice

图 12 Dice 盒图

值得一提的是,分割结果可能存在一些零散分割区域,这存在两种情况。第一种情况,这些区域并不是病变区域;第二情况,这些区域依附在主体的分割病变区域上,大概率这离散区域本身就是病变区域。为此,两种情况的零散分割区域能够相互抵消,导致了即使经过



消除实验, Dice 也是相差无几, 消除实验的作用在该任务中显得微乎其微, 最终本文只分析原始分割结果数据作为结论。

可知 MS-V-Net 网络能够更细粒度地利用上采样过程中的底层特征和高层特征的语义信息, 通过融合这些信息特征以提升网络分割精度。MSS-V-Net 通过在不同融合层引入多个分类目标函数, 实现对连续融合层学习到的信息特征提供直接监督, 不像传统的深度卷积网络主要依赖最后层计算预测误差再进行反向传播来指导前面层的学习。

## 6 结束语

在计算机辅助诊断高度近视的诊断与治疗的系统中, 图像分割系统能够预测到高度近视病变严重程度, 供医生进一步确认患者病情, 这些工作不仅增加了诊断效率和精确度, 而且提供了病变筛查的可行措施。本文采用了具有多尺度深度监督思想的模型来分割高度近视萎缩病变眼底图像区域。该模型不需要人工提取病变信息特征, 只需利用分割网络强大的信息表征能力进行自动学习, 即可直接获得分割结果。针对图像分割问题, 该模型与原始的 V-Net 方法相比, 充分地利用了眼底图像丰富的语义信息, 更为有效地适应病理高度近视的复杂特征, 且具有强大的自主学习能力、泛化能力、鲁棒性等优点。值得一提的是, 本文开发的图像优化算法无需复杂的处理工作, 就能够获得结构清晰、风格统一的眼底图像, 并使模型可以获得相对较好的分割精度。

更为重要的是, 作者使用本文的多尺度深度监督的深度思想参加了 MICCAI 2019 收录文章的生物医学图像分析的 Grand Challenge-KiTS19 Challenge<sup>[21]</sup>竞赛, 在分割 CT 图像的肾脏与肿瘤的任务中, 取得了第七名<sup>[22]</sup>的比赛成绩<sup>[23]</sup>, 虽然不是针对眼底图像的相关疾病分割的竞赛, 但是也一定程度上验证了本文方法的可行性和有效性。

## 参考文献:

- [1] GUO Shuxu, MA Shuzhi, et al. Fully convolutional neural network for liver segmentation in CT image [J]. Computer Engineering and Applications, 2017, 053(018):126-131.
- [2] Yanwu Xu, Hrvo Bogunovic, et al. Pathologic Myopia Challenge[EB/OL]. <https://palm.grand-challenge.org/2019>.
- [3] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, 2015: 3431-3440.
- [4] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation[J]. Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, 2015:234-241.
- [5] Drozdal M, Vorontsov E, Chartrand G, et al. The Importance of Skip Connections in Biomedical Image Segmentation[C]. International Workshop on Large-Scale Annotation of Biomedical Data and Expert Label Synthesis International Workshop on Deep Learning in Medical Image Analysis. Springer International Publishing, 2016: 179-187.
- [6] Çiçek, Özgün, Abdulkadir A, Lienkamp S S, et al. 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation[J]. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016, 2016: 424-432.
- [7] Chen LC, Papandreou G, Kokkinos I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2018, 40(4):834-848.
- [8] J.A. Hartigan, M.A. Wong. A K-means clustering algorithm[J]. Appl Stat, 2013, 28(1):100-108.
- [9] Luo Man, Huang Jing, Yang Feng. Multimodal 3D Convolutional Neural Networks Features for Brain Tumor Segmentation[J]. Science Technology and Engineering, 2014(31):84-89.
- [10] Milletari F, Navab N, Ahmadi S A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation[J]. 2016 Fourth International Conference on 3D Vision, Stanford, CA, 2016:565-571.
- [11] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[J]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 770-778.
- [12] Wang Wei-min, Fu Shou-fu, et al. An insect image segmentation and counting method based on convolutional neural network[J]. Computer Engineering and Applications, 2019:110-116.
- [13] Lin T Y, Dollár P, Girshick R, et al. Feature Pyramid Networks for Object Detection[J]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016: 2117-2125.

- [14] Lee C Y,Xie S,Gallagher P, et al. Deeply-supervised nets. Artificial Intelligence and Statistics.2015:562-570.
- [15] Chen J, Tie J, Tang Z, et al. Retinal vessel enhancement and extraction based on directional field[J].Journal of X-Ray Science and Technology,2008, 16:189-201.
- [16] Sandler M, Howard A, Zhu M, et al. MobileNetV2: Inverted Residuals and Linear Bottlenecks[J]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) ,2018: 4510-4520.
- [17] Hubbard L D, Danis R P, Neider M W, et al. Brightness, Contrast, and Color Balance of Digital versus Film Retinal Images in the Age-Related Eye Disease-Study2[J].Investigative Ophthalmology &V-isual Science,2008,49(8):3269-3282.
- [18] Peking University, et al.Peking University International Competition on Ocular Disease Intelligent Recognition[EB/OL].  
<https://odir2019.grand-challenge.org/>,2019.
- [19] Jiawei Han, Micheline Kamber.Data mining concepts and technologies [M].Beijing: Machinery Industry Press,2012:0-350.
- [20] Zhang Yu. Structural analysis method based on experimental data mining and cellular automata [D].Ha'er'bin:Harbin Institute of Technology , 2010.
- [21] Christopher W, Niranjan S,et al.KiTS 19-Challenge[EB/OL].<https://kits19.grand-challenge.org/>, 2019.
- [22] Christopher W, Niranjan S, et al.MICCAI-KiTS2019Results[EB/OL]:  
<http://results.kits-challenge.org/miccai2019/>,2019.
- [23] Wenshuai Z, Zengfeng Z. Multi Scale Supervised 3D U-Net for Kidney and Tumor Segmentation [J].Cornell University,2019:0-7.
- 2017, 053(018): 126-131.
- [9] 罗蔓,黄靖,杨丰.基于多模态 3D-CNNs 特征提取的 MRI 脑肿瘤分割方法[J].科学技术与工程, 2014(31): 84-89.
- [12] GB/T 7714 王卫民, 符首夫, 顾榕蓉,等. 基于卷积神经网络的虫情图像分割和计数方法[J].计算机工程与科学, 2020, 042(001):110-116.
- [19] 范明,孟小峰.数据挖掘概念与技术[M].北京:机械工业出版社,2012: 0-350.
- [20] 张瑀.基于实验数据挖掘与细胞自动机的结构分析方法[D].哈尔滨工业大学, 2010.

### 作者简介:



曾增峰 (1994-), 男, 福建漳州人, 硕士, 研究方向为计算机视觉。E-mail: 17210720170@fudan.edu.cn  
Zeng Zeng-feng, born in 1994, MS candidate, his research interest includes computer vision.



环宇翔 (1991-), 男, 江苏南通人, 博士, 研究方向为智能芯片与系统。E-mail: yxhuan@fudan.edu.cn  
HUAN Yu-xiang, born in 1991, Ph.D., as assistant researcher, his research interest includes smart chips and systems.



邹卓 (1982-), 男, 四川成都人, 博士, 教授, 研究方向为物联网应用的智能系统研究。E-mail: zhuo@fudan.edu.cn  
Zou Zhuo, born in 1982, Ph.D., as professor, his research interest includes smart electronic systems for IoT applications.



郑立荣 (1969-), 男, 浙江绍兴人, 博士, 教授, 研究方向为无线片上系统等。E-mail: lrzheng@fudan.edu.cn  
ZHENG Li-rong, born in 1969, Ph.D., as professor, his research interest includes wireless system on chip et al.

### 附中文参考文献:

- [1] 郭树旭,马树志,李晶, et al.基于全卷积神经网络的肝脏 CT 影像分割研究[J].计算机工程与应用,