

RCAR-UNet: 基于粗糙通道注意力机制的视网膜血管分割网络

孙颖 丁卫平 黄嘉爽 鞠恒荣 李铭 耿宇

(南通大学信息科学技术学院 江苏南通 226019)

(17805056265@163.com)

RCAR-UNet: Retinal Vessels Segmentation Network Based on Rough Channel Attention Mechanism

Sun Ying, Ding Weiping, Huang Jiashuang, Ju Hengrong, Li Ming, and Geng Yu

(School of Information Science and Technology, Nantong University, Nantong, Jiangsu 226019)

Abstract The health of retinal vessels in fundus images is of great significance for the early diagnosis of various ophthalmic diseases and diabetic cardiovascular diseases, etc. However, the retinal blood vessels are delicate, distributed irregularly and the boundary is ambiguous. Therefore, it is difficult to accurately segment them. Based on the characteristics of retinal blood vessels, we propose a U-shaped network—rough channel attention residual U-Net (RCAR-UNet), which combines rough neurons and channel attention mechanism. Firstly, the network introduces the concept of upper and lower approximation in rough set theory to design rough neurons. Secondly, the rough channel attention module is constructed based on rough neurons, and the module uses global max pooling and global average pooling in U-Net skip connections to construct upper and lower approximation neurons, and performs weighted summation between neurons to reasonably rough the established channel dependencies, which not only contain global information but also have local characteristics, and can effectively achieve accurate rescaling of the extracted retinal vessels features. Then adding residual connections to transfer features directly from the lower to the higher layers, to help solve the network performance degradation problem and effectively extract richer retinal vascular features. Finally, in order to verify the effectiveness of the proposed RCAR-UNet model, comparison experiments are performed on three public fundus image datasets with traditional network models such as U-Net, Attention U-Net, etc. The results show that the RCAR-UNet model has high superiority in the accuracy, sensitivity and similarity of blood vessel segmentation.

Key words rough set; rough neuron; attention mechanism; fundus retinal blood vessels; image segmentation; residual connection

收稿日期: 2021-07-06; 修回日期: 2022-06-23

基金项目: 国家自然科学基金面上项目(61976120); 国家自然科学基金青年科学基金项目(62006128, 62102199); 江苏省自然科学基金项目(BK20191445); 江苏省双创博士计划项目; 江苏省研究生科研与实践创新计划项目(SJCX20_1150); 江苏省高等学校自然科学研究重大项目(21KJA510004); 江苏省高等学校自然科学基金面上项目(20KJB520009); 南通市科技局基础科学研究项目(JC2020141, JC2021122); 江苏“青蓝工程”项目

This work was supported by the General Program of the National Natural Science Foundation of China (61976120), the National Natural Science Foundation of China for Young Scientists (62006128, 62102199), the Natural Science Foundation of Jiangsu Province (BK20191445), the Double-Creation Doctoral Program of Jiangsu Province, the Postgraduate Research & Practice Innovation Program of Jiangsu Province (SJCX20_1150), the Major Program of the Natural Science Research of Jiangsu Province Higher Education Institutions (21KJA510004), the General Program of the Natural Science Foundation of Jiangsu Province Higher Education Institutions (20KJB520009), the Basic Science Research Program of Nantong Science and Technology Bureau (JC2020141, JC2021122), and the Qing Lan Project of Jiangsu Province.

通信作者: 丁卫平 (dwp9988@163.com)

摘要 眼底图像中视网膜血管的健康状况对早期诊断各种眼科疾病及糖尿病心脑血管疾病等具有重要意义,然而视网膜血管结构细微、边界模糊且分布不规则,对其进行准确分割存在较大的难度.针对视网膜血管的这些特征,提出一种粗糙通道注意力残差 U型网——粗糙通道注意力残差 U型网络 (RCAR-UNet).该网络首先引入粗糙集理论中上下近似概念设计粗糙神经元;接着基于粗糙神经元构建粗糙通道注意力模块,该模块在 U-Net 跳跃连接中采用全局最大池化和全局平均池化构造上下近似神经元,并进行神经元间的加权求和,对所建立的通道依赖关系进行合理的粗糙化,该依赖关系不仅包含全局信息,同时具有局部特性,可有效实现对所提取视网膜血管特征的准确重标定;然后添加残差连接,将特征直接从低层传递给高层,有助于解决网络性能退化问题,并有效提取更加丰富的视网膜血管特征;最后为了验证所提视网膜分割网络的有效性,在 3 个眼底视网膜公开图像数据集上与 U-Net, Attention U-Net 等传统网络模型进行对比实验,实验结果表明,所提视网膜分割网络在血管分割准确率、敏感度和相似度等方面具有较高的优越性.

关键词 粗糙集;粗糙神经元;注意力机制;眼底视网膜血管;图像分割;残差连接

中图法分类号 TP391

眼健康作为国民健康中的重要组成部分,涉及到公共卫生和社会领域的民生福祉,引起了国家健康委的高度重视^[1].视网膜血管中蕴含丰富的形态特征,如血管直径、旁支角度、尺寸和弯曲度等^[2].各种眼科疾病以及心脑血管疾病都会导致视网膜血管出现形态结构变化、出血等不同程度的病变,从而导致视力受损^[3],所以在临幊上医生广泛使用眼底图像来分析视网膜血管的形态变化并辅助诊断各种眼科及心脑血管等疾病具有重要意义^[4].然而眼底图像中视网膜血管分布密集而无规律,存在大量易与背景混淆、对比度较低的细小血管,血管边界模糊不清,同时易受采集设备和光照以及病变组织的影响^[5].这些问题使得临幊上手动分割视网膜血管不仅工作量巨大而且对医疗人员的经验和技能要求颇高.此外,不同专家对同一张图像的血管提取也存在主观上的差异,手动分割已不能满足临幊的需要.

随着计算机技术的不断发展,实现眼底视网膜血管的智能分割并对眼科疾病进行辅助诊断和决策,成为国内外学者关注的研究热点.深度学习凭借其在识别应用中超高的预测准确率,在图像处理领域获得了极大关注^[6-9].与传统方法相比,深度学习模型能够以端到端的方式自动提取特征.全卷积神经网络^[10-11](fully convolutional network, FCN)是首个应用于图像语义分割任务的神经网络,在此基础上形成目前分割任务中最流行的编解码结构^[12-13].而 U-Net 模型不仅具有编解码结构,同时在 U型对称层之间添加跳跃连接,实现低层特征和高层特征的拼接和融合,在图像分割领域有显著的优势,在医学图像分割领域获得较好的效果.何慧等人^[14]利用改进预测编码器 U-Net 模型实现 PET 肿瘤的自动分割,实现了

更准确、快速、稳定的肿瘤分割,分割结果可以达到金标准的 88.5%;Rundo 等人^[15]为了解决前列腺区域分割任务,将 Squeeze-and-Excitation 块合并到 U-Net 中,提出了一种新的卷积神经网络,称为 USE-Net;Jin 等人^[16]将可变形卷积集成到 U-Net 模型中,提出 DUNet 网络模型,根据血管的大小和形状自适应地调节感受野来捕获各种形状和大小的视网膜血管等.

Attention 机制是模仿人类注意力而提出的一种解决问题的方法,是一种能让模型对与任务相关的重要信息重点关注的技术,能够作用于任何序列模型,其应用领域包括文本、图片等. Basiri 等人^[17]将 Attention 机制和长短期记忆模型相融合,提出一种基于注意力的双向 CNN-RNN(convolutional neural network-recurrent neural network)深度模型用于情感分析; Haut 等人^[18]将 Attention 机制和残差网络模型相结合用于分析遥感高光谱图像,得到更为准确的分类准确率.

进一步地,将符合人类视觉机制的 Attention 机制与能够实现低层特征和高层特征融合的 U-Net 模型相结合,实现对显著性区域的关注,在医学图像分割中广泛应用. Guo 等人^[19]开发了一个 3D 深度注意力 U-Net,从冠状动脉计算机断层扫描血管造影中分割左心室心肌轮廓; Cui 等人^[20]将注意力机制和 U-Net 模型相结合,在短轴磁共振成像图像中进行心脏分割.基于注意力机制的 U-Net 模型在眼底视网膜血管分割中也取得不错效果^[21-24], Guo 等人^[21]提出了一种空间注意力 U-Net 的轻量级网络实现对视网膜血管的精确分割; Tang 等人^[22]提出多尺度的通道注意力模块和空间信息定位模块来提高血管末端的分割准确率.

上述模型与方法将 Attention 机制融合 U-Net 网络用于眼底视网膜血管分割中仍然存在一定的不足, 视网膜血管分布密集、杂乱且形状不规则, 血管粗细不一, 存在大量细小血管, 血管边界不清晰, 且易受采集设备和光照以及病变组织等噪声的影响, 上述模型无法解决血管边界的不确定性和细小血管的分割. 为弥补该不足之处, 本文引入能有效分析不精确、不一致、不完整等各种不完备信息的粗糙集理论中上下近似概念设计粗糙神经元, 对特征通道依赖关系进行合理粗糙化.

本文针对视网膜血管分布密集而杂乱, 存在大量对比度低的细小血管, 血管边界模糊, 且易受采集设备和光照以及病变区域等噪声影响的特点, 提出一种粗糙通道注意力残差 U 型网络 (RCAR-UNet), 有效提高对细微血管的分割精度. 该网络以 U-Net 模型为主干, 首先引入粗糙集上下近似概念设计粗糙神经元. 然后利用粗糙神经元对每一层下采样视网膜血管特征图建立通道之间的依赖关系, 构建粗糙通道注意力机制, 将全局最大池化和全局平均池化分别作为通道重要性权重的上下近似神经元, 将全局平均池化作为通道重要性权重的下近似神经元, 对每个通道重要性权重设置上下限, 并对上下限赋予自适应权重系数, 进行神经元间加权求和, 得到更加合理的 Attention 系数; 对下采样的视网膜血管特征图进行特征的重标定, 并与 U-Net 模型对称层的上采样视网膜血管特征图进行高低层特征之间的拼接融合. 最后为解决经典 U-Net 网络的退化问题, 在该模型中添加残差连接, 实现将低层视网膜特征直接传递给高层, 可有效提取更加准确的视网膜血管特征.

1 相关工作

1.1 粗糙集理论

粗糙集于 1982 年由波兰数学家 Pawlak^[25] 提出, 是一种可以定量分析处理不精确、不一致、不完整信息与知识的数学工具. 粗糙集理论基于不可分辨关系^[26] 对数据进行划分, 利用上下近似集对目标进行描述, 形成正域、负域和边界域 3 个互不重叠的区域^[27].

假设决策信息系统表示为 $S = (U, AT, f, V)$, 其中, $U = \{x_1, x_2, \dots, x_n\}$ 表示非空有限对象集合, n 表示系统中对象的个数; AT 表示信息系统的属性集合, 它由条件属性集合 C 和决策属性集合 D 共同组成; $V = \bigcup_{a \in AT} V_a$ 表示所有属性的值域; $f: U \times AT \rightarrow V$ 表示

信息函数, $\forall a \in AT, x \in U, f(x, a) \in V_a$.

粗糙集可以通过上、下近似集对目标概念 X 进行逼近. 粗糙集在属性集合 $R \subseteq A$ 对论域 U 的任意对象子集 X 中的对象进行上、下近似集的划分, 那么 X 基于属性集合 R 的上近似集 $\bar{R}(X)$ 和下近似集 $\underline{R}(X)$ 定义为

$$\bar{R}(X) = \{x \in U | [x]_R \cap X \neq \emptyset\}, \quad (1)$$

$$\underline{R}(X) = \{x \in U | [x]_R \subseteq X\}. \quad (2)$$

目标概念 X 基于属性集合 R 的正域 $POS_R(X)$ 、负域 $NEG_R(X)$ 和边界域 $BND_R(X)$ 定义为

$$POS_R(X) = \{x \in U | [x]_R \subseteq X\}, \quad (3)$$

$$NEG_R(X) = \{x \in U | [x]_R \cap X \neq \emptyset\}, \quad (4)$$

$$BND_R(X) = \bar{R}(X) - \underline{R}(X). \quad (5)$$

正域 $POS_R(X)$ 表示论域 U 中基于不可分辨关系 R 一定属于目标概念 X 的对象集合; 负域 $NEG_R(X)$ 表示论域 U 中基于不可分辨关系 R 一定不属于目标概念 X 的对象集合; 边界域 $BND_R(X)$ 表示论域 U 中基于不可分辨关系 R 可能属于目标概念 X , 也可能不属于目标概念 X 的对象集合. 边界域 $BND_R(X)$ 描述了 X 的粗糙度, 若 $BND_R(X) = \emptyset$, 说明其是精确的集合; 若 $BND_R(X) \neq \emptyset$, 说明其是粗糙集.

1.2 Attention U-Net 模型

Attention U-Net 模型以 U-Net 编解码的 U 形网络结构为主干, 其中 U-Net 网络结构包括:

1) 收缩路径(编码器)包括卷积层、激活层和池化层. 卷积层具有局部感知、参数共享等特性, 用于图像局部特征的自动提取; 激活层的输出都是对上一层输入的线性映射, 常用激活函数 ReLu 和激活函数 Sigmoid; 池化层则是对所提取的特征进行数据的降维和压缩操作, 常用平均池化和最大池化.

2) 扩展路径(解码器)包括反卷积层和卷积层, 经过下采样特征图尺寸会越来越小, 所以上采样主要用于恢复特征图的细节和尺寸.

3) 在上采样和下采样的同层结构中添加跳跃连接, 将高层图像特征信息与低层图像特征信息相联合, 实现上下特征信息的融合.

U-Net 模型最主要的思想是在收缩路径上捕获全局特征, 在扩展路径上实现精确定位. 为解决扩展路径的上采样过程中重建的空间信息不精确问题, 使用跳跃连接将下采样的空间信息与上采样的空间信息相结合, 然而跳跃连接也带来了许多冗余的底层特征, 造成大量的无效计算, 浪费大量的计算资源. 对此, Oktay 等人^[28] 将注意力门作用于 U-Net 模型的

跳跃连接中, 提出 Attention U-Net 模型, 该模型能够有效地抑制不相关区域中的激活, 从而减少冗余特征的数量, 并且不会引入大量的参数和计算量, 得到更加符合人类视觉机制的网络模型, 其结构如图 1 所示。同一张图像的不同区域关注度不同, 对此, 通过注意力门系数控制关注区域, 得到更加具有语义

的特征图。注意力门具体结构如图 2 所示。将收缩路径得到的特征图 $g \in \mathbb{R}^{H_g \times W_g \times C}$ 与扩展路径中同层通道数相同的特征图 $p \in \mathbb{R}^{H_p \times W_p \times C}$ 相加, 经过一系列的线性变换, 得到注意力系数 $A \in \mathbb{R}^{H_p \times W_p}$, 将其与特征图 p 相乘, 实现对特征图 p 的不同区域特征值的重新标定, 得到新的特征图 p' 。

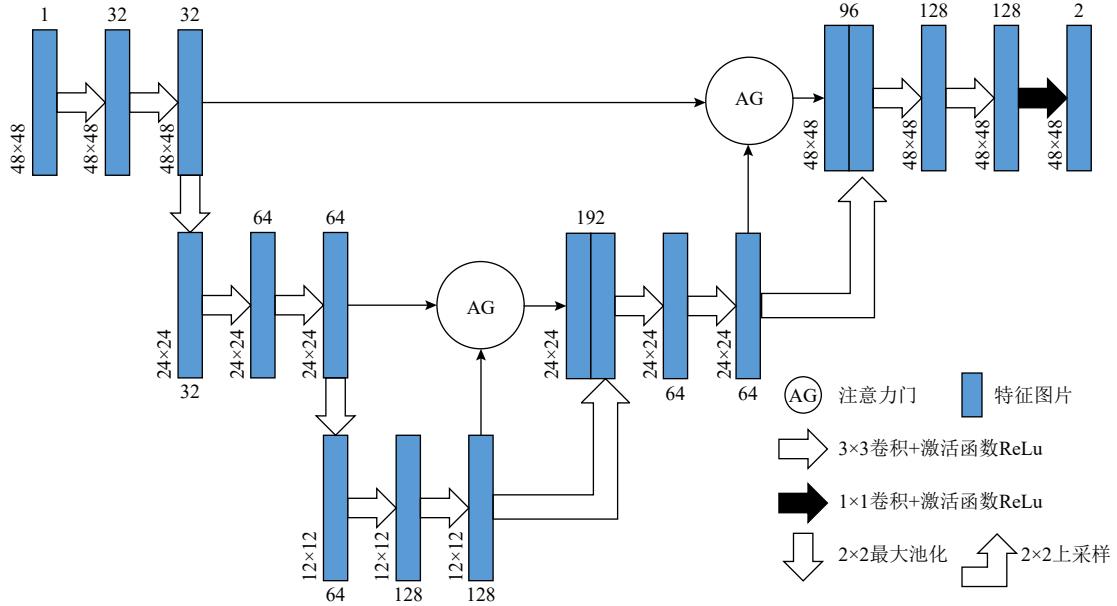


Fig. 1 Attention U-Net network architecture

图 1 Attention U-Net 网络结构

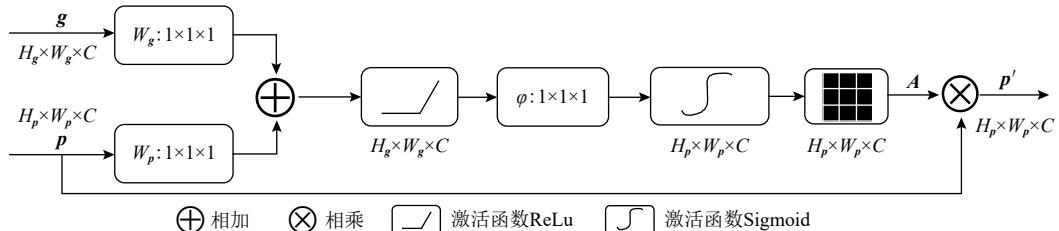


Fig. 2 Attention gate architecture

图 2 注意力门结构

2 粗糙神经网络

图像信息复杂多样, 其内容蕴含着诸如随机性、模糊性等不确定性信息, 例如眼底视网膜血管分布杂乱, 形状不规则并且边缘模糊不清, 且易受采集设备和光照以及病变影响, 这些眼底视网膜血管图像中蕴含的不确定信息使得深度神经网络效果不佳。1996 年, Lingras^[29] 使用上下界的一般概念引入了粗糙神经网络。深度神经网络在学习复杂特征方面的成功以及粗糙神经网络处理不确定性的能力, 促进了将粗糙神经网络与深度架构相结合^[30-31]。

2.1 粗糙神经元

上下限的概念已在人工智能的各种应用中使用, 特别是粗糙集理论中上下近似集证明了上下限在规则生成中的有效性, 粗糙集理论利用上下近似集对目标概念进行逼近, 本文引入粗糙集的上下近似集的思想, 构建上下近似神经元对注意力模块所得的注意力系数进行合理粗糙化, 粗糙神经元^[32-35]结构如图 3 所示。

上近似神经元的参数 $\theta_U = \{W_U, b_U, \alpha\}$, 下近似神经元的参数 $\theta_L = \{W_L, b_L, \beta\}$, 其中 W_U, b_U 表示上近似神经元的权重和偏差, W_L, b_L 表示下近似神经元的权重和偏差。图 3 中 O_U, O_L 表示上下近似神经元的输出, 与

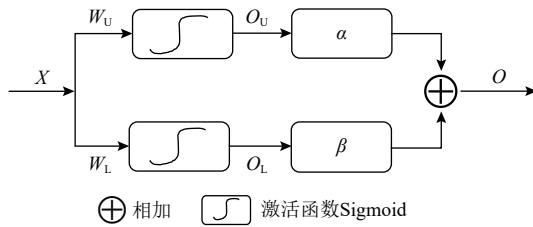


Fig. 3 Rough neuron architecture

图 3 粗糙神经元结构

常规神经元的单个输出值不同的是,粗糙神经元的输出是一对上下限,其计算公式为:

$$O_U = f(W_U \times X + b_U), \quad (6)$$

$$O_L = f(W_L \times X + b_L). \quad (7)$$

最终给定上下边界神经元一定的权重 α 和 β ,加权求和得到最终粗糙神经元的输出 O ,表示为

$$O = \alpha \times O_U + \beta \times O_L. \quad (8)$$

粗糙神经元是对输入的数值映射为一个粗糙的范围,最后的输出取决于上下神经元的输出,对于不同的输出赋予一定的权重值,以区分不同神经元对最终输出的贡献.该方法能够对输入的确定数值粗糙化,对确定性的数值进行重标定,得到更合理、准

确的输出值.

2.2 通道注意力机制

特征图的通道数取决于卷积操作中卷积核的个数,以往认为得到的多通道特征图中每个通道所蕴含的信息重要程度相同,即不对特征通道的重要性加以区分.然而同一张图像的不同区域关注度不同,同样,每个特征通道的关注度也有所差异.例如1张眼底视网膜血管图像有2个通道,分割目标是血管,那么会更加关注与分割任务相关的血管通道,也就是需要赋予每个特征通道一个表征通道蕴含信息重要度的权重,以此对不同特征通道信息加以区分.将表征每个特征通道信息重要性的权重与原特征图对应的每个通道相乘得到的对应值就是真实的特征图值.

为了建立特征通道之间的关系,引用一个新的维度信息来表征特征通道重要度,进行特征通道之间的融合,即通过训练网络参数的方式计算每个特征通道信息的重要性权重,依照所得到的特征通道重要性权重提升与当前分割任务相关的特征并抑制对当前分割任务无关的特征.通道注意力机制^[36-38]主要包括激励部分,其结构如图4所示.

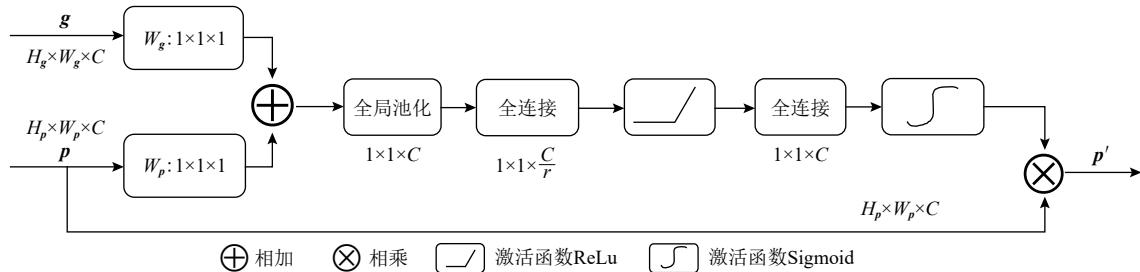


Fig. 4 Channel attention mechanism

图 4 通道注意力机制

首先是压缩操作,对特征通道维度进行特征图的压缩,通过全局池化,将维度为 $H_p \times W_p \times C$ 的特征图的每个通道的2维数据转化为一个实数 \mathbb{R} ;然后是激励操作,包括2层全连接:第1个全连接层是把 C 个通道压缩成 C/r 个通道来减少模型所需的参数和计算时间,第2个全连接层将特征图恢复为 C 个通道,使模型更加非线性,更好地适应通道之间的复杂关系,激活函数 ReLu 降低了梯度消失的概率,激活函数 Sigmoid 使 C 中每个通道的特征权值在 0~1 之间;最后是特征重标定,重新计算特征通道值,通过权值乘法运算完成特征通道内对原始各个特征通道重要程度的重新标定,并作为下一级的输入数据.

2.3 基于粗糙通道注意力机制 U-Net 模型

基于粗糙神经元和通道注意力机制,提出粗糙

通道注意力机制模块,并将其嵌入到 U-Net 模型的跳跃连接中. RCAR-UNet 模型结构分别如图 5 和表 1 所示.

在 U-Net 模型的特征编码部分,构建一个简单的特征提取模块包括:残差连接的 2 个 3×3 的卷积层和 1 个 2×2 的最大池化层.加入残差连接是因为残差连接能够实现将低层特征直接传播到高层,在一定程度上解决了网络退化的问题,以此提高模型性能.

在特征解码部分,使用粗糙通道注意力机制来融合特征图的上下语义信息,以便生成更具代表性的特征图.粗糙通道注意力机制融合了粗糙神经元处理不确定信息的能力和通道注意力机制能够区分不同特征通道重要性的优点.通道注意力机制可以实现对特征图不同特征通道重要度的区分,一般选用

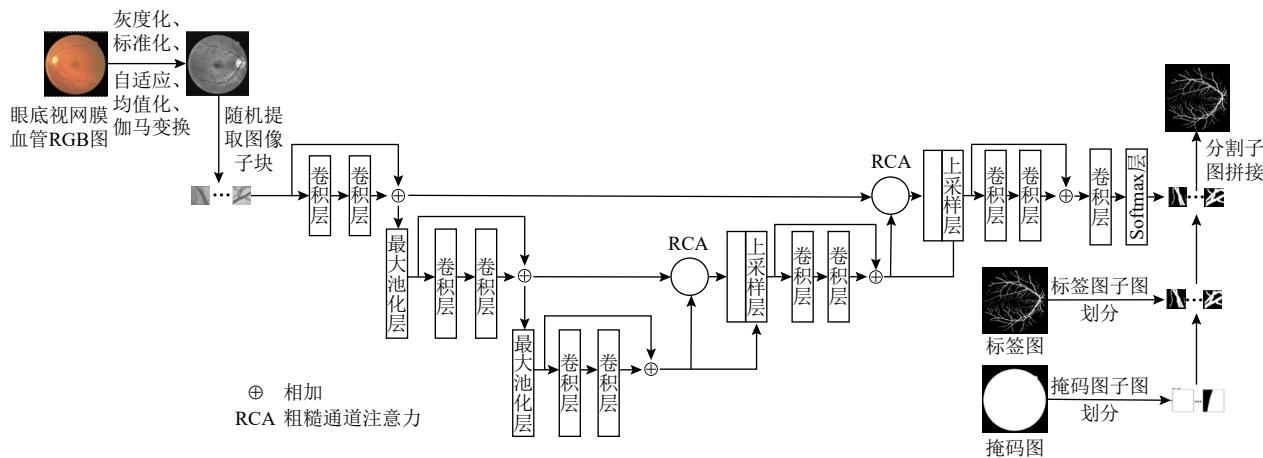


Fig. 5 Architecture of RCAR-UNet

图 5 RCAR-UNet 模型结构

Table 1 U-Net Model Architecture Based on Rough Channel Attention

表 1 基于粗糙通道注意力 U-Net 模型结构

编号	模块	层	图像尺寸
1	输入模块	输入层	$48 \times 48 \times 1$
		卷积层 1 (3×3卷积核, 激活函数 ReLu)	$48 \times 48 \times 32$
		卷积层 2 (3×3卷积核, 激活函数 ReLu)	$48 \times 48 \times 32$
2	编码模块 1	残差连接	$48 \times 48 \times 32$
		2×2最大池化层, 步长为 2	$24 \times 24 \times 32$
		卷积层 3 (3×3卷积核, 激活函数 ReLu)	$24 \times 24 \times 64$
		卷积层 4 (3×3卷积核, 激活函数 ReLu)	$24 \times 24 \times 64$
3	编码模块 2	残差连接	$24 \times 24 \times 64$
		2×2最大池化层, 步长为 2	$12 \times 12 \times 64$
		卷积层 5 (3×3卷积核, 激活函数 ReLu)	$12 \times 12 \times 128$
4	编码模块 3	卷积层 6 (3×3卷积核, 激活函数 ReLu)	$12 \times 12 \times 128$
		残差连接	$12 \times 12 \times 128$
		2×2上采样	$24 \times 24 \times 128$
		拼接层 1 上采样 1 ($24 \times 24 \times 128$) 粗糙通道注意力 1 ($24 \times 24 \times 64$)	$24 \times 24 \times 192$
5	解码模块 1	卷积层 7 (3×3卷积核, 激活函数 ReLu)	$24 \times 24 \times 64$
		卷积层 8 (3×3卷积核, 激活函数 ReLu)	$24 \times 24 \times 64$
		残差连接	$24 \times 24 \times 64$
		2×2上采样	$48 \times 48 \times 64$
		拼接层 2 上采样 2 ($48 \times 48 \times 64$) 粗糙通道注意力 2 ($48 \times 48 \times 32$)	$48 \times 48 \times 96$
6	解码模块 2	卷积层 9 (3×3卷积核, 激活函数 ReLu)	$48 \times 48 \times 32$
		卷积层 10 (3×3卷积核, 激活函数 ReLu)	$48 \times 48 \times 32$
		残差连接	$48 \times 48 \times 32$
7		卷积层 11 (1×1卷积核)	$48 \times 48 \times 2$
8		Softmax 激活层	$48 \times 48 \times 2$
9	输出模块	输出层	$48 \times 48 \times 2$

全局平均池化对特征图进行压缩, 建立特征通道之间的关系, 其值在一定程度上具有全局感受野. 然而视网膜血管粗细不一, 存在大量对比度较低的细微

血管, 其结构细长, 只有一个或者几个像素的宽度, 局部图像细节信息也不容忽视, 因此加以利用全局最大池化对特征图进行压缩, 使得其值在一定程度

上具有局部感受野. 对此, 本文在构建粗糙神经元时, 使用全局最大池化和全局平均池化作为上下近似神经元对特征通道重要性进行上下限描述, 使得 Attention

系数值兼具全局信息和局部细节信息. 在此基础上进行特征重标定, 得到新的视网膜血管特征图. 设计粗糙通道注意力机制结构如图 6 所示.

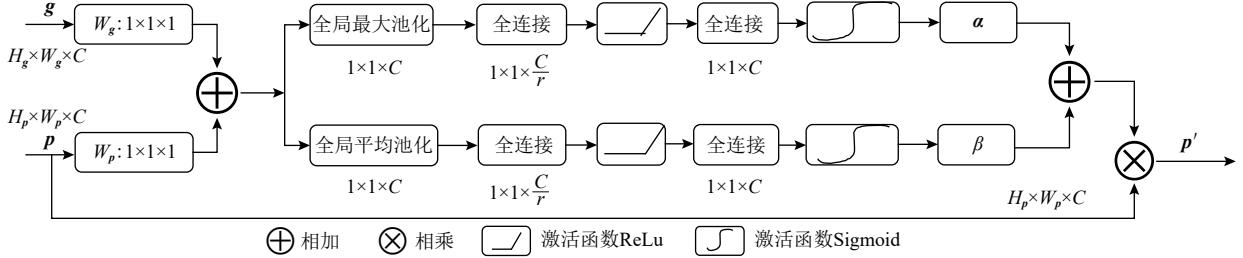


Fig. 6 Rough channel attention mechanism

图 6 粗糙通道注意力机制

如图 6 所示的具体步骤为: 首先定义具有 C 个通道的低层特征图 $\mathbf{L}_{\text{feature}} = (\mathbf{L}_{\text{feature}_1}, \mathbf{L}_{\text{feature}_2}, \dots, \mathbf{L}_{\text{feature}_c})$, 其中 $\mathbf{L}_{\text{feature}_c} \in \mathbb{R}^{H \times W \times C}$. $\mathbf{L}_{\text{feature}_k} \in \mathbb{R}^{H \times W}$ 表示低层特征图的第 k 个特征通道, H 、 W 分别表示低层特征图的高度和宽度; 定义具有 C 个特征通道的高层特征图 $\mathbf{H}_{\text{feature}} = (\mathbf{H}_{\text{feature}_1}, \mathbf{H}_{\text{feature}_2}, \dots, \mathbf{H}_{\text{feature}_c})$, 其中 $\mathbf{H}_{\text{feature}} \in \mathbb{R}^{H' \times W' \times C}$. $\mathbf{H}_{\text{feature}_k} \in \mathbb{R}^{H' \times W'}$ 表示高层特征图的第 k 个特征通道, H' 、 W' 分别表示高层特征图的高度和宽度.

对高低特征图进行相加操作, 得到融合特征图. 表示为

$$\begin{aligned} \mathbf{New}_{\text{feature}} &= \mathbf{L}_{\text{feature}} + \mathbf{H}_{\text{feature}}, \\ \mathbf{New}_{\text{feature}} &= (\mathbf{New}_{\text{feature}_1}, \mathbf{New}_{\text{feature}_2}, \dots, \mathbf{New}_{\text{feature}_c}). \end{aligned} \quad (9)$$

利用全局最大池化层建立通道间的依赖关系, 并保留全局信息, 表示为

$$\mathbf{F}_{\text{feature}_c}^{\text{High}} = \max(\mathbf{New}_{\text{feature}_c}(i, j)). \quad (10)$$

其中 $0 < i \leq H$, $0 < j \leq W$, $0 < c \leq C$, 全局最大池化即计算每个特征通道中像素的最大值, 经过全局最大池化层, 得到一个 $1 \times 1 \times C$ 的张量.

利用全局平均池化层建立通道之间的依赖关系, 并保留局部信息, 表示为

$$\mathbf{F}_{\text{feature}_c}^{\text{Low}} = \frac{1}{H' \cdot W'} \sum_{i=1}^{H'} \sum_{j=1}^{W'} \mathbf{New}_{\text{feature}_c}(i, j), \quad (11)$$

其中 $0 < i \leq H'$, $0 < j \leq W'$, $0 < c \leq C$, 全局平均池化即计算每个特征通道中像素和的平均值, 经过全局平均池化层, 同样得到一个 $1 \times 1 \times C$ 的张量.

分别对全局平均池化和全局最大池化的 $2 \times 1 \times C$ 的张量进行激励操作: 第 1 个全连接层是把 C 个通道压缩成 C/r 个通道来减少模型所需的参数和计算时间; 第 2 个全连接层将特征图恢复为 C 个通道, 使模型更加非线性, 更好地适应通道之间的复杂关

系, r 是指压缩的比例, 本文设置 $r = 16$. 激活函数 ReLu 降低了梯度消失的概率, 激活函数 Sigmoid 使 C 中每个通道的特征权值在 0~1 之间, 得到 Attention 系数的上下限值, 分别表示为

$$\mathbf{F}_{\text{feature}_c}^{\text{High}} = \text{Sigmoid}(\mathbf{W}_2 \cdot \text{ReLu}(\mathbf{W}_1 \cdot \mathbf{F}_{\text{feature}_c}^{\text{High}})), \quad (12)$$

$$\mathbf{F}_{\text{feature}_c}^{\text{Low}} = \text{Sigmoid}(\mathbf{W}_2 \cdot \text{ReLu}(\mathbf{W}_1 \cdot \mathbf{F}_{\text{feature}_c}^{\text{Low}})). \quad (13)$$

其中 $\mathbf{W}_1 \in \mathbb{R}^{\frac{C}{r} \times C}$, $\mathbf{W}_2 \in \mathbb{R}^{\frac{C}{r} \times C}$ 分别表示 2 个全连接层的权值矩阵.

为了使得通道之间的依赖关系既包含全局信息, 又包含局部细节信息, 对所得到的通道重要性值的上下限进行加权操作, 得到新的通道之间的依赖关系:

$$\mathbf{F}_{\text{feature}_c} = \alpha \cdot \mathbf{F}_{\text{feature}_c}^{\text{High}} + \beta \cdot \mathbf{F}_{\text{feature}_c}^{\text{Low}}, \quad (14)$$

其中 α, β 分别表示上下神经元的权重信息.

利用输出的通道之间的依赖关系对特征图进行重新标定, 得到新的特征图表示:

$$\mathbf{F}'_{\text{feature}_c} = \mathbf{F}_{\text{feature}_c} \cdot \mathbf{L}_{\text{feature}_c}. \quad (15)$$

将 $\mathbf{F}'_{\text{feature}_c}$ 与对称层具有相同维度的上采样特征图进行拼接, 作为网络下一层的输入.

在 U-Net 模型的特征解码部分, 首先构建一个简单的恢复特征模块包括: 一个 2×2 上采样层和残差连接的 2 个 3×3 的卷积层; 然后因为该任务包含血管类和非血管类 2 类, 所以在恢复特征尺寸后使用 2 个 1×1 的卷积核进行卷积操作得到一个通道数为 2 的特征图, 一个通道表示血管类, 另一个通道表示非血管类; 最后使用 Softmax 层输出每个像素属于血管类的概率.

眼底图像分割的过程如算法 1 所示.

算法 1. RCA-UNet 模型的视网膜血管分割算法.

输入: 彩色眼底视网膜血管图像 $\text{input} \in \mathbb{R}^{H \times W \times 3}$ 、掩码图 $\text{mask} \in \mathbb{R}^{H \times W \times 1}$;

输出: 视网膜血管分割图 $output \in \mathbb{R}^{H \times W \times 2}$.

1) 对视网膜血管原图 $input \in \mathbb{R}^{H \times W \times 3}$ 进行灰度化、自适应均衡化等预处理, 得到视网膜血管预处理图 $input' \in \mathbb{R}^{H \times W \times 1}$;

2) 以步长 $patch_stride$ 从视网膜血管预处理图 $input' \in \mathbb{R}^{H \times W \times 1}$ 中进行视网膜血管图像子块 $patch_image \in \mathbb{R}^{H' \times W' \times 1}$ 的划分, 保证视网膜血管图像子块覆盖视网膜血管图;

3) 将视网膜血管图像子块 $patch_image \in \mathbb{R}^{H' \times W' \times 1}$ 输入到训练好的 RCAR-UNet 模型中, 得到视网膜血管子图 $patch_output \in \mathbb{R}^{H' \times W' \times 2}$, 其具体步骤为:

① 利用卷积操作和残差连接提取眼底视网膜血管特征.

② 利用最大池化对眼底视网膜血管特征进行降维.

③ 利用粗糙通道注意力机制求得较为合理的 Attention 系数对收缩路径提取的视网膜血管特征进行准确重标定.

④ 利用上采样操作恢复视网膜血管特征图的尺寸, 直到网络输入的视网膜血管图尺寸保持一致为止.

⑤ 采用 2 个 1×1 的卷积核进行操作得到通道为 2 的视网膜血管特征图;

⑥ 利用 Softmax 层得到视网膜血管特征图中的每个像素点属于血管类别和非血管类别的概率;

⑦ 设置阈值 $\theta=0.5$, 如果视网膜血管特征图中的像素点属于血管类的概率值大于 θ , 则属于血管类; 反之, 属于非血管类.

4) 将视网膜血管图像子块 $patch_output \in \mathbb{R}^{H' \times W' \times 2}$ 按照提取顺序进行图像拼接, 得到完整的视网膜血管分割图 $output' \in \mathbb{R}^{H \times W \times 2}$;

5) 对视网膜血管掩码 $mask \in \mathbb{R}^{H \times W \times 1}$ 进行通道的修改, 使其通道数与视网膜血管网络分割图 $output' \in \mathbb{R}^{H \times W \times 2}$ 保持一致, 得到视网膜血管掩码图 $mask' \in \mathbb{R}^{H \times W \times 2}$;

6) 利用视网膜血管掩码从视网膜血管网络输出图中抠出眼球部分, 得到最终视网膜血管分割图 $output \in \mathbb{R}^{H \times W \times 2}$.

RCA-UNet 模型的时间复杂度可以表示为

$$T \sim O\left(\sum_{i=1}^N F_i^2 \times K_i^2 \times C_{i-1} \times C_i\right). \quad (16)$$

时间复杂度与每一层的输出特征图尺寸 F_i 、每一个卷积核的尺寸 K_i 、上一层输出特征图的通道数 C_{i-1} 以及当前层输出特征图的通道数 C_i 有关, 与此同时每一层特征图的输出尺寸 F 与输入尺寸 X 、卷积

核的步长 $stride$ 、填充的数量 $padding$ 以及卷积核的尺寸 K 等相关, 可表示为

$$F = (X - K + 2 \times padding) / stride + 1. \quad (17)$$

3 实验

3.1 数据集和预处理

3.1.1 数据集

DRIVE 数据集^[39-41] 发布于 2004 年, 包含 40 张格式为 tif、尺寸为 565×584 的彩色眼底图像, 每张图像包含 2 位专家手工标注的金标准图, 并且自带视网膜血管的掩膜图.

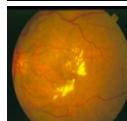
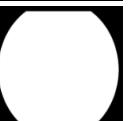
Stare 数据集^[39-41] 是 1975 年 Michael Goldbaum 发起的项目, 用来进行视网膜血管分割的彩色眼底图数据库, 包括 20 张格式为 ppm、尺寸为 605×700 的彩色眼底图像, 同样每张图像对应 2 位专家手工标注的金标准图, 相对应的掩膜需要自己通过代码进行掩膜的设置.

CHASE DB1 数据集^[39-41] 包括从 14 名学童的双眼中拍摄的 28 张格式为 jpg、尺寸为 999×960 的视网膜图像. 每张图像具有 2 位专家的手工分割标签, 相对应的掩膜需要自己通过代码进行掩膜的设置. 一般情况下将前 20 张图像用于训练, 其余 8 张图像用于测试.

3 个眼底视网膜血管图像数据集信息如表 2 所示.

Table 2 Information of Retinal Vessels Image Datasets

表 2 眼底视网膜血管图像数据集信息

数据集	原图	金标准图	掩码图
DRIVE			
Stare			
CHASE DB1			

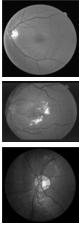
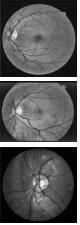
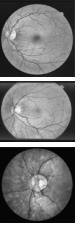
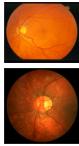
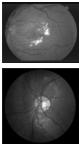
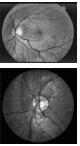
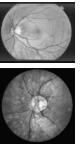
3.1.2 预处理

由于眼底图像照明不均匀以及血管与背景之间的对比度较低等因素, 为了捕获细小血管的更多特征并提高血管分割的准确性, 需要将输入网络的眼底图像进行预处理. 首先对 RGB 图像进行通道分离,

发现 G 通道的血管与背景之间的对比度最高, 选取 RGB 图像的 G 通道完成图像的灰度变换, 对视网膜血管灰度图像进行归一化; 然后采用对比度受限的自适应直方图均衡化, 在不放大眼底视网膜图像噪声的情况下增强视网膜血管与背景之间的对比度, 以使眼底图像中血管的结构和特征更容易受到关注; 最后使用 Gamma 变换进行图像增强, 对过白或者过暗的图像区域进行校正。**表 3** 分别表示原图和预处理之后的图像。

Table 3 Information of Original Images and Preprocessed Images

表 3 原图与预处理图信息

数据集	原图	G 通道	均衡化	伽马变换
DRIVE				
Stare				
CHASE DB1				

3.1.3 数据集扩充

由于深度卷积神经网络的复杂性, 训练一个深度卷积神经网络进行图像分割, 通常需要大量标签图像, 然而, 只有几十张视网膜血管图像具有像素级的标签, 因此, 设计用于视网膜血管分割的深度学习模型容易出现过拟合现象。对此, 本文采用对原图进行分块的方式进行数据扩充, 将训练图像以及相应的掩码图划分为大小为 48×48 的图像子块, 从中随机选取一定数量的图像子块数, **图 7** 中展示了整合部分的图像子块以及相对应的掩码子块图。

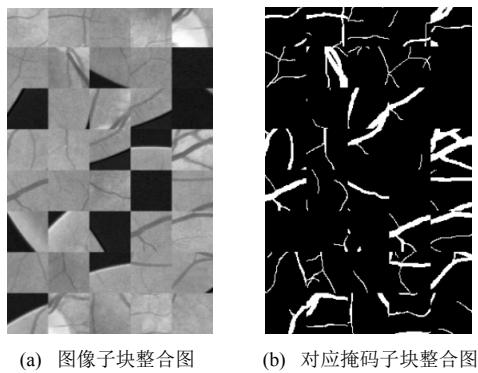


Fig. 7 Sub-image and corresponding mask sub-image

图 7 图像子块以及相对应的掩码子块图

3.2 实验参数设置

本研究使用基于 Windows10 系统的工作站, 运

行在 Intel® Core™ i7-10750H CPU @2.60GHz 上, 拥有 16GB 内存和 NVidia GeForce RTX 2060 6.0GB GPU, 使用 TensorFlow 和 Keras 等框架构建文中所用的网络模型。在模型的训练过程中, 采用交叉熵损失函数作为训练的损失函数, 批大小 $batch_size$ 设置为 32, 模型迭代的次数 N_epochs 设置为 50, 模型初始学习率设为 0.01, 选择 SGD 随机梯度下降法作为优化器对模型的参数进行更新。

3.3 评价指标

视网膜血管分割任务的实质是像素级的分类, 判断像素点是血管类还是非血管类。血管是需要检测分割的目标, 称作正类; 非血管类的部位, 称作负类。分割算法的结果与真实值比较可以得到混淆矩阵中的真阳性 N_{TP} 、假阳性 N_{FP} 、假阴性 N_{TN} 、真阴性 N_{FN} , 如**表 4** 所示。其中 N_{TP} 是将血管类正确分类为血管类的像素点数, N_{FP} 是将非血管类错分为血管类的像素点数, N_{TN} 是将非血管类分类为非血管类的像素点数, N_{FN} 是将血管类错分为非血管类的像素点数。

Table 4 Confusion Matrix

表 4 混淆矩阵

真实值	预测为正类 (血管)	预测为负类 (背景)
标签为正类 (血管)	N_{TP}	N_{FN}
标签为负类 (背景)	N_{FP}	N_{TN}

为了评价视网膜血管分割算法的好坏, 选用准确率 Acc 、灵敏度 Sen 、特异性 Spe 、精确率 Pre 等评价指标, 其中 Acc 表示将血管类和背景类分类正确的概率, Sen 表示将血管类分类正确的概率, Spe 表示将背景类分类正确的概率, Pre 表示预测为血管类的样本中真正为血管类所占的比例。为了进一步地评估分割模型的性能, 利用 Jaccard 相似度描述金标准图 *truth* 与分割图 *result* 之间的相似性和差异性, Jaccard 值越大说明相似度越高。各评价指标公式表示为:

$$Acc = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}}, \quad (18)$$

$$Sen = \frac{N_{TP}}{N_{TP} + N_{FN}}, \quad (19)$$

$$Spe = \frac{N_{TN}}{N_{TN} + N_{FP}}, \quad (20)$$

$$Pre = \frac{N_{TP}}{N_{TP} + N_{FP}}, \quad (21)$$

$$J(truth, result) = \frac{|truth \cap result|}{|truth \cup result|}, \quad (22)$$

$$F_1 = 2 \times \frac{Pre \times Sen}{Pre + Sen}. \quad (23)$$

3.4 实验结果分析

本文所提模型是以 U-Net 网络为基础网络, 并针对视网膜血管的独特特性, 将 U-Net 网路和本文所提的粗糙通道注意力机制相融合。首先, 为了验证基础网络 U-Net 的有效性, 选取目前较常使用的分割网络 FCN 和 Seg-Net^[42] 在眼底图像视网膜血管 DRIVE, Stare, CHASE DB1 数据集上进行对比实验, 从准确率 *Acc*、灵敏度 *Sen*、特异性 *Spe*、精确率 *Pre*、Jaccard 相似度 *J* 等评价指标对实验结果进行比较。

FCN, Seg-Net, U-Net 模型在视网膜血管的 3 个数据集上的对比结果分别如表 5~7 所示, 从表 5~7 中可以发现, U-Net 模型具有更好的表现, 总体表现优于其他 2 个模型。更具体地说, U-Net 具有更高的分割准确率, 对血管像素具有更好的识别能力, 所得分割图和金标准图具有更高的相似度, 对背景像素的识别能力也是具有竞争力的, 综合考虑分割精度和敏感度, 取得更高的 *F₁* 值。

Table 5 Comparison Results on DRIVE Dataset

表 5 DRIVE 数据集对比结果

方法	<i>Acc</i>	<i>Sen</i>	<i>Spe</i>	<i>Pre</i>	<i>J</i>	<i>F₁</i>
FCN	0.900 1	0.258 9	0.994 6	0.875 4	0.249 7	0.399 6
Seg-Net	0.944 0	0.695 0	0.977 4	0.822 2	0.619 3	0.764 9
U-Net	0.950 9	0.703 1	0.987 1	0.888 2	0.645 9	0.784 9

注: 黑体数字表示最优值。

Table 6 Comparison Results on Stare Dataset

表 6 Stare 数据集对比结果

方法	<i>Acc</i>	<i>Sen</i>	<i>Spe</i>	<i>Pre</i>	<i>J</i>	<i>F₁</i>
FCN	0.920 4	0.268 1	0.997 8	0.937 9	0.263 5	0.417 1
Seg-Net	0.928 9	0.360 8	0.996 3	0.920 8	0.349 9	0.518 5
U-Net	0.952 6	0.612 4	0.992 9	0.911 8	0.578 2	0.732 7

注: 黑体数字表示最优值。

Table 7 Comparison Results on CHASE DB1 Dataset

表 7 CHASE DB1 数据集对比结果

方法	<i>Acc</i>	<i>Sen</i>	<i>Spe</i>	<i>Pre</i>	<i>J</i>	<i>F₁</i>
FCN	0.934 0	0.475 6	0.978 4	0.680 9	0.388 9	0.560 0
Seg-Net	0.937 7	0.543 9	0.975 8	0.685 6	0.435 3	0.606 6
U-Net	0.951 7	0.636 5	0.982 2	0.766 7	0.537 9	0.699 5

注: 黑体数字表示最优值。

上述结果表明 Seg-Net 和 U-Net 这类基于编解码

结构的网络在医学图像分割上具有更好的竞争力, 而增加了跳跃连接结构的 U-Net 在上采样过程中将对称层的特征图进行通道上的拼接, 实现低层特征和高级特征的信息融合使得网络能够提取和保留更多视网膜血管局部细节信息, 从而提高了图像分割精度。基于此, 选用 U-Net 作为基础网络具有一定的有效性。

为了进一步验证本文所提粗糙通道注意力机制的有效性, 选取 U-Net, Attention U-Net, RCA-UNet 在 3 个眼底视网膜血管数据集中进行对比, 从准确率 *Acc*、灵敏度 *Sen*、特异性 *Spe*、精确率 *Pre*、Jaccard 相似度 *J* 等评价指标对实验结果进行比较, 实验结果分别如表 8~10 所示。实验结果表明: 在 3 个数据集中 RCA-UNet 模型都具有相对较好的性能。具体而言, RCA-UNet 模型具有较高的灵敏度 *Sen*, 对血管类具有更好的识别能力, 在 3 个数据集上相对于 U-Net 模型提高了 2.97%, 2.9%, 10.33%, 相对于 Attention U-Net 模型提高了 1.35%, 1.23%, 0.54%。在考虑模型性能的时候, 单纯地追求精度 *Pre* 或者灵敏度 *Sen* 的提升并没有太大作用, 在实际分割任务中, 需要结合正负样本比进行综合评价, 对此从精度 *Pre* 与灵敏度 *Sen* 的调和平均值 *F₁* 可以看出 RCA-UNet 模型具有较好的性能, 提高 0.19%~1.43%, 0.88%~2.03%, 1.24%~4.71%。另外, RCAR-UNet 所得的分割图与金标准的 Jaccard 相似度更高, 相较于其他模型相似度提高了 0.26%~1.96%, 1.12%~2.57% 和 1.57%~5.78%。上述实验结果清楚地表明利用全局最大池化和全局平均池化构建上下近似神经元的粗糙通道注意力机制的有效性。分析其原因在于考虑了特征通道之间的依赖关系, 借助粗糙集的上下近似原理, 利用全局最大池化和全局平均池化分别构建上下近似神经元, 并赋予一定的自适应权重, 得到较为合理的 Attention 系数, 并对特征图进行相应的重标定操作, 使得在上采样的过程中, 得到更加细致的特征信息。

Table 8 Comparison Results on DRIVE Dataset

表 8 DRIVE 数据集对比结果

方法	<i>Acc</i>	<i>Sen</i>	<i>Spe</i>	<i>Pre</i>	<i>J</i>	<i>F₁</i>
U-Net	0.950 9	0.703 1	0.987 1	0.888 2	0.645 9	0.784 9
Attention U-Net	0.952 7	0.729 7	0.985 3	0.878 6	0.662 9	0.797 3
RCA-UNet	0.953 1	0.732 8	0.985 2	0.878 8	0.665 5	0.799 2
RCAR-UNet	0.953 7	0.748 7	0.983 6	0.869 6	0.673 2	0.804 7

注: 黑体数字表示最优值。

为了在说明粗糙通道注意力机制的有效性的同

时, 进一步验证残差连接的有效性, 本文将 RCA-UNet 和 RCAR-UNet 在 3 个数据集上进行对比实验, 实验结果分别如表 8~10 所示。实验结果表明, 本文所提模型 RCAR-UNet 能够得到更好的视网膜血管分割效果。从 3 个表中可以发现, 在 3 个视网膜血管数据集中各个评价指标都有所提升。从表 8 中可以很直观地看出在 DRIVE 数据集上, RCAR-UNet 相对于 RCA-UNet 模型, Acc 提高了 0.06%, Sen 提高了 1.59%, 相似度 J 提高了 0.77%, F_1 提高了 0.55%; 从表 9 中可以看出, RCAR-UNet 在 Acc , Sen , J , F_1 指标上都有明显的提升, 分别提升了 0.41%, 5.65%, 4.22%, 3.2%; 从表 10 中可以看出, RCAR-UNet 在 CHASE DB1 数据集上相对于 RCA-UNet 模型的提升不是很大, 但是都有所改善, Acc 提高了 0.1%, Sen 提高了 0.33%, Pre 提高了 2.11%, Jaccard 相似度提高 0.26%, F_1 指标提高了 0.04%。上述实验结果充分说明模型中添加残差连接实现特征映射的有效性, 将低层特征直接传递给高层特征的短跳跃连接方式, 不仅丰富网络特征提取, 并且有助于训练模型时梯度的反向传播, 可有效解决网络退化问题。

Table 9 Comparison Results on Stare Dataset

表 9 Stare 数据集对比结果

方法	Acc	Sen	Spe	Pre	J	F_1
U-Net	0.9526	0.6124	0.9929	0.9118	0.5782	0.7327
Attention U-Net	0.9541	0.6291	0.9927	0.9109	0.5927	0.7442
RCA-UNet	0.9553	0.6414	0.9926	0.9119	0.6039	0.7530
RCAR-UNet	0.9594	0.6979	0.9905	0.9069	0.6461	0.7850

注: 黑体数字表示最优值。

Table 10 Comparison Results on CHASE DB1 Dataset

表 10 CHASE DB1 数据集对比结果

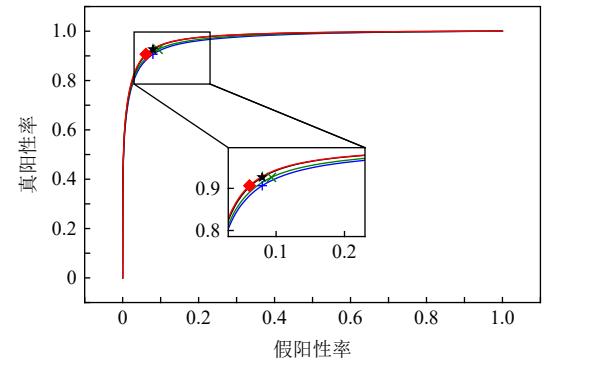
方法	Acc	Sen	Spe	Pre	J	F_1
U-Net	0.9517	0.6365	0.9822	0.7667	0.5379	0.6995
Attention U-Net	0.9530	0.7344	0.9742	0.7340	0.5800	0.7342
RCA-UNet	0.9556	0.7398	0.9765	0.7536	0.5957	0.7466
RCAR-UNet	0.9566	0.7475	0.9798	0.7747	0.5983	0.7470

注: 黑体数字表示最优值。

最后, 本文还采用 ROC 曲线和 PR 曲线进行模型的评价。ROC 曲线是对假阳性率和真阳性率的安全考虑, 横坐标表示假阳性率, 纵坐标表示真阳性率; PR 曲线是对召回率和精确率的综合考虑, 其横坐标表示召回率 $Recall$ (和灵敏度 Sen 一样), 纵坐标表示精确率 Pre 。曲线与坐标系围成的面积表示 AUC 面积,

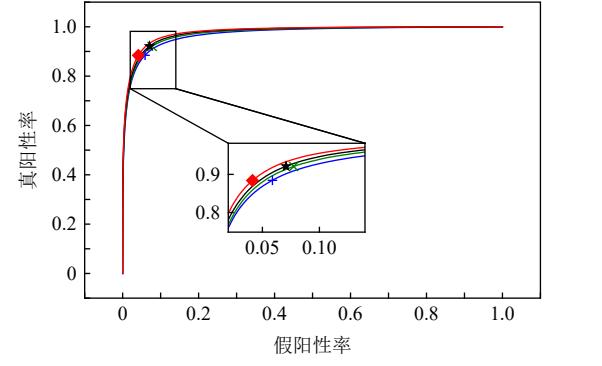
AUC 值越大, 表明模型具有更好的性能。4 个网络模型在不同数据集上的 ROC 曲线和 PR 曲线对比实验结果分别如图 8 和图 9 所示。

图 8 展示了不同模型在三个数据集上的 ROC 曲线对比实验结果, 包括 DRIVE、Stare 和 CHASE DB1。图中展示了 U-Net、Attention U-Net、RCA-UNet 和 RCAR-UNet 的 AUC 值。



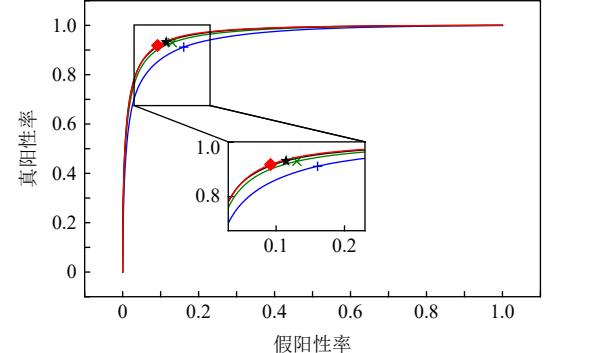
(a) DRIVE 数据集 ROC 曲线对比实验

图 8 展示了不同模型在三个数据集上的 ROC 曲线对比实验结果, 包括 DRIVE、Stare 和 CHASE DB1。图中展示了 U-Net、Attention U-Net、RCA-UNet 和 RCAR-UNet 的 AUC 值。



(b) Stare 数据集 ROC 曲线对比实验

图 8 展示了不同模型在三个数据集上的 ROC 曲线对比实验结果, 包括 DRIVE、Stare 和 CHASE DB1。图中展示了 U-Net、Attention U-Net、RCA-UNet 和 RCAR-UNet 的 AUC 值。



(c) CHASE DB1 数据集 ROC 曲线对比实验

Fig. 8 Comparison of ROC curves of different models on different datasets

图 8 不同模型在不同数据集上的 ROC 曲线对比

从图 8 中可以看出, 本文所提模型 RCAR-UNet 在 3 个数据集上的 ROC 曲线的 AUC 值为 0.9759, 0.9801, 0.9695, 相对于经典 U-Net 模型, 提高了 0.73%, 0.93%,

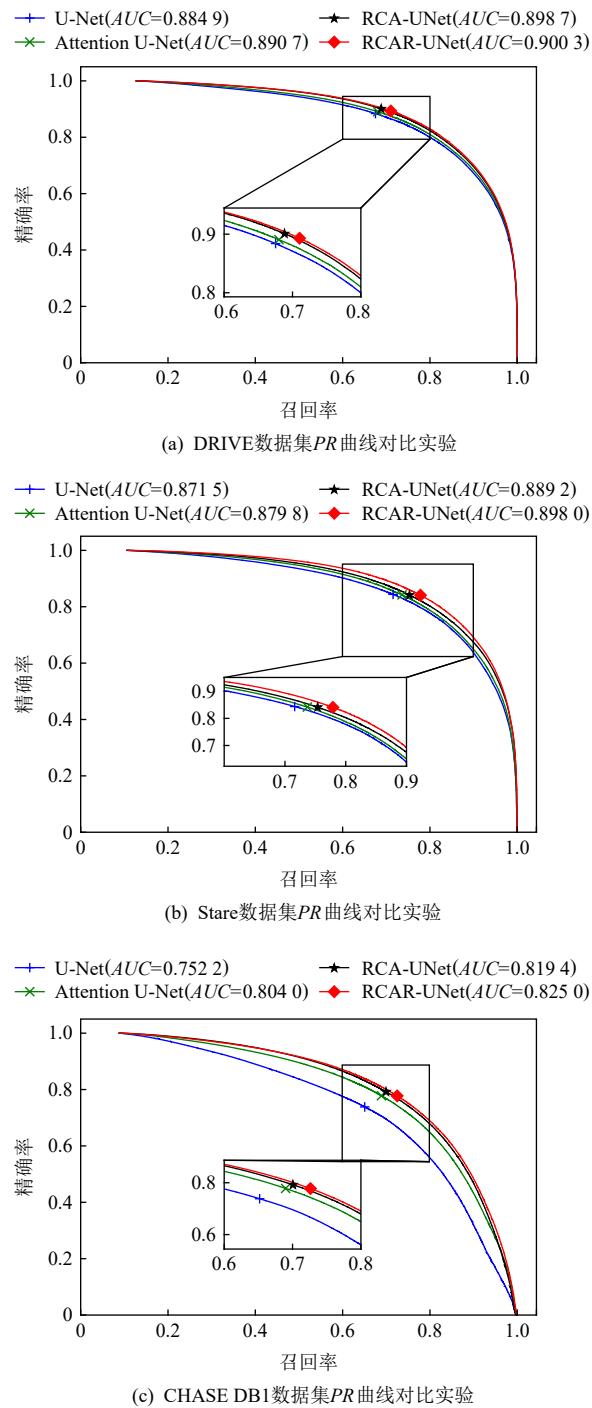


Fig. 9 Comparison of PR curves of different models on different datasets

图 9 不同模型在不同数据集上的 PR 曲线对比

2.05%. ROC 曲线的纵坐标为真阳性率, 表示实际正样本中被预测成正样本的占比, 也就是实际正样本的预测正确率, 所以是越大越好; 而其横坐标为假阳性率, 表示实际负样本中被预测成正样本的占比, 也就是实际正样本的预测错误率, 所以是越小越好. 最为理想的点是(0,1), 综上所述, ROC 曲线的 AUC 值越大, 性能越好, 所以 RCAR-UNet 在 3 个眼底视网膜

血管数据集上具有较好的性能. PR 曲线描述的是精准率与召回率的关系, 在分割过程中, 期望精确率和召回率值均相对较高, 达到较好的平衡点, 同样利用曲线与坐标系围成的 AUC 面积对模型进行评估, 图 9 结果表明 RCAR-UNet 具有较好的性能. RCAR-UNet 在 3 个数据集上的 AUC 值为 0.9003, 0.8980, 0.8250, 相对于其他模型都有所提高.

综上所述, RCAR-UNet 相对于经典 U-Net, Attention U-Net 等模型能够得到更好的眼底视网膜血管分割精度, 分析其主要原因在于: RCAR-UNet 在编码-解码的 U-Net 模型主干上, 利用粗糙集中上下近似理论构建粗糙上下近似神经元, 对眼底视网膜血管特征通道重要性进行区分, 对视网膜血管特征进行准确重标定, 并融合残差连接, 进行特征的映射.

为了进一步直观地说明 RCAR-UNet 对于视网膜血管分割具有较好的分割效果, 将之与 U-Net, Attention U-Net 等模型的分割图进行比较, 各模型分割效果图如图 10 所示. 从视网膜血管分割效果图可以直观看出, 本文所提 RCAR-UNet 模型的分割结果和专家分割标准图基本一致, 特别是在矩形框区域内细微血管的分割上有更好的效果.

4 总 结

针对视网膜血管结构复杂、血管较细、易受光照影响等, 本文考虑了特征通道之间的关系, 引入一种新的通道注意力机制来增加网络的鉴别能力. 具体而言: 首先在编码-解码结构的 U-Net 模型上利用粗糙集的上下近似概念构建粗糙注意力模块, 对注意力系数设置上下限, 利用全局最大池化构造上神经元作为注意力系数的上限, 利用全局平均池化构造下神经元作为注意力系数的下限, 通过赋予上下限一定的权重系数并求和得到较具语义的注意力系数, 该注意力系数不仅包含全局信息也具有了一定的局部信息; 然后引入残差连接, 实现将低层特征直接传播到高层, 在一定程度上解决了网络退化的问题, 以此提高模型性能; 最后在 3 个眼底视网膜血管数据集进行网络模型性能验证, 实验结果表明本文所提网络 RCAR-UNet 模型能够分割出视网膜血管末端的细小分支, 得到较好的视网膜血管分割精度. 我们会进一步研究将粗糙集理论融合到深度神经网络模型的卷积和池化层中, 处理图像特征中不确定性、不精确性信息.

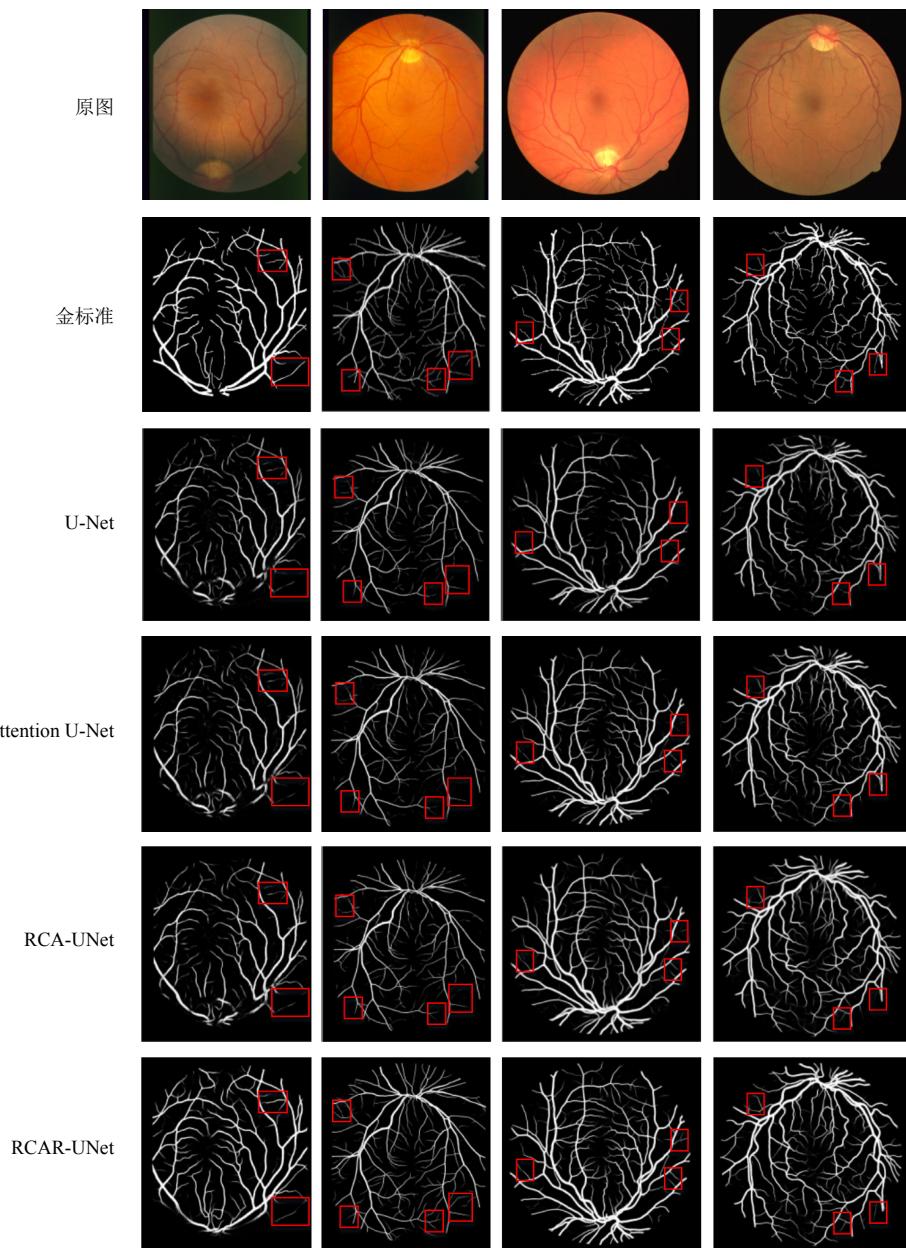


Fig. 10 Segmentation effect diagram of retinal blood vessels of each model

图 10 各模型视网膜血管分割效果图

作者贡献声明: 孙颖提出了算法的核心思想,设计了实验方案,完成实验并撰写论文初稿;丁卫平提出了整个算法的框架并对整个算法思想进行完善,修改了论文;黄嘉爽、鞠恒荣完善了算法的思路,指导了论文写作并修改论文;李铭、耿宇协助实验数据的处理并修改论文.

参 考 文 献

- [1] Liang Liming, Huang Chaolin, Shi Fei, et al. Vessels segmentation of fundus images with level set fusion with shape priors[J]. *Chinese Journal of Computers*, 2018, 41(7): 1678–1692 (in Chinese)

[梁礼明, 黄朝林, 石霏, 等. 融合形状先验的水平集眼底图像血管分割[J]. *计算机学报*, 2018, 41(7): 1678–1692)

- [2] Wu Yicheng, Xia Yong, Song Yang, et al. NFN+: A novel network followed network for retinal vessel segmentation[J]. *Neural Networks*, 2020, 126: 153–162
- [3] Li Xiang, Jiang Yuchen, Li Minglei, et al. Lightweight attention convolutional neural network for retinal vessel image segmentation[J]. *IEEE Transactions on Industrial Informatics*, 2021, 17(3): 1958–1967
- [4] Cheung C Y, Xu Dejiang, Cheng Chingyun, et al. A deep-learning system for the assessment of cardiovascular disease risk via the measurement of retinal-vessel calibre[J]. *Nature Biomedical Engineering*, 2021, 5(6): 498–508
- [5] Abràmoff M D, Garvin M K, Sonka M. Retinal imaging and image

- analysis[J]. *IEEE Reviews in Biomedical Engineering*, 2010, 3: 169–208
- [6] Chen Jia, Chen Yasong, Li Weihao, et al. Application and prospect of deep learning in video object segmentation[J]. *Chinese Journal of Computers*, 2021, 44(3): 609–631 (in Chinese)
(陈加, 陈亚松, 李伟浩, 等. 深度学习在视频对象分割中的应用与展望[J]. *计算机学报*, 2021, 44(3): 609–631)
- [7] Sun Liang, Shao Wei, Wang Mingliang, et al. High-order feature learning for multi-atlas based label fusion: Application to brain segmentation with MRI[J]. *IEEE Transactions on Image Processing*, 2019, 29: 2702–2713
- [8] Sun Liang, Shao Wei, Zhang Daoqiang, et al. Anatomical attention guided deep networks for ROI segmentation of brain MR images[J]. *IEEE Transactions on Medical Imaging*, 2019, 39(6): 2000–2012
- [9] Hesamian M H, Jia Wenjing, He Xiangjian, et al. Deep learning techniques for medical image segmentation: Achievements and challenges[J]. *Journal of Digital Imaging*, 2019, 32(4): 582–596
- [10] Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 39(4): 640–651
- [11] Sun Jindong, Peng Yanjun, Guo Yanfei, et al. Segmentation of the multimodal brain tumor image used the multi-pathway architecture method based on 3D FCN[J]. *Neurocomputing*, 2021, 423: 34–45
- [12] Ding Henghui, Jiang Xudong, Shuai Bing, et al. Semantic segmentation with context encoding and multi-path decoding[J]. *IEEE Transactions on Image Processing*, 2020, 29: 3520–3533
- [13] Wu Congzhong, Sun Jun, Wang Jing, et al. Encoding-decoding network with pyramid self-attention module for retinal vessel segmentation[J]. *International Journal of Automation and Computing*, 2021, 18(6): 973–980
- [14] He Hui, Chen Sheng. Improved pretrained encoder U-Net model for automatic segmentation of PET tumors[J]. *Chinese Journal of Image and Graphics*, 2020, 25(1): 171–179 (in Chinese)
(何慧, 陈胜. 改进预训练编码器U-Net模型的PET肿瘤自动分割[J]. *中国图象图形学报*, 2020, 25(1): 171–179)
- [15] Rundo L, Han C, Nagano Y, et al. USE-Net: Incorporating squeeze-and-excitation blocks into U-Net for prostate zonal segmentation of multi-institutional MRI datasets[J]. *Neurocomputing*, 2019, 365: 31–43
- [16] Jin Qiangguo, Meng Zhaopeng, Pham T D, et al. DUNet: A deformable network for retinal vessel segmentation[J]. *Knowledge-Based Systems*, 2019, 178: 149–162
- [17] Basiri M E, Nemati S, Abdar M, et al. ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis[J]. *Future Generation Computer Systems*, 2021, 115: 279–294
- [18] Haut J M, Paoletti M E, Plaza J, et al. Visual attention-driven hyperspectral image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(10): 8065–8080
- [19] Guo Bangjun, He Xiuxiu, Lei Yang, et al. Automated left ventricular myocardium segmentation using 3D deeply supervised attention U-Net for coronary computed tomography angiography: CT myocardium segmentation[J]. *Medical Physics*, 2020, 47(4): 1775–1785
- [20] Cui Hengfei, Yuwen Chang, Jiang Lei, et al. Multiscale attention guided U-Net architecture for cardiac segmentation in short-axis MRI images[J]. *Computer Methods and Programs in Biomedicine*, 2021, 206: 106142
- [21] Guo Changlu, Szemenyei M, Yi Yugen, et al. SA-UNet: Spatial attention U-Net for retinal vessel segmentation[C] //Proc of the 25th Int Conf on Pattern Recognition (ICPR2020). Piscataway, NJ: IEEE, 2021: 1236–1242
- [22] Tang Xianlun, Zhong Bing, Peng Jiangping, et al. Multi-scale channel importance sorting and spatial attention mechanism for retinal vessels segmentation[J]. *Applied Soft Computing*, 2020, 93: 106353
- [23] Gu Ran, Wang Guotian, Song Tao, et al. CA-Net: Comprehensive attention convolutional neural networks for explainable medical image segmentation[J]. *IEEE Transactions on Medical Imaging*, 2020, 40(2): 699–711
- [24] Mou Lei, Zhao Yitian, Chen Li, et al. CS-Net: Channel and spatial attention network for curvilinear structure segmentation[C] //Proc of the Int Conf on Medical Image Computing and Computer-Assisted Intervention. Berlin: Springer, 2019: 721–730
- [25] Pawlak Z. Rough sets[J]. *International Journal of Computer and Information Sciences*, 1982, 11(5): 341–356
- [26] Liu Wen, Mi Jusheng, Sun Yan. Axiomatic characterization of new hesitant fuzzy rough approximation operators[J]. *Journal of Computer Research and Development*, 2021, 58(9): 2062–2070 (in Chinese)
(刘文, 米据生, 孙妍. 一种新的犹豫模糊粗糙近似算子的公理刻画[J]. *计算机研究与发展*, 2021, 58(9): 2062–2070)
- [27] Miao Duoqian, Xu Feifei, Yao Yiyu, et al. A set-theoretic description of granular computing[J]. *Chinese Journal of Computers*, 2012, 35(2): 2351–2363 (in Chinese)
(苗夺谦, 徐菲菲, 姚一豫, 等. 粒计算的集合论描述[J]. *计算机学报*, 2012, 35(2): 2351–2363)
- [28] Oktay O, Schlemper J, Folgoc L L, et al. Attention U-Net: Learning whereto look for the pancreas[J]. arXiv preprint, arXiv: 1804.03999, 2018
- [29] Lingras P. Rough neural networks[C] //Proc of the 6th Int Conf on Information Processing and Management of Uncertainty in Knowledge Based Systems. Piscataway, NJ: IEEE, 1996: 1445 – 1450
- [30] Jelonek J, Krawiec K, Slowiński R. Rough set reduction of attributes, and their domains for neural networks[J]. *Computational Intelligence*, 1995, 11(2): 339–347
- [31] Khodayar M, Kaynak O, Khodayar M E. Rough deep neural architecture for short-term wind speed forecasting[J]. *IEEE Transactions on Industrial Informatics*, 2017, 13(6): 2770–2779
- [32] Cao Bin, Zhao Jianwei, Lv Zhihan, et al. Multiobjective evolution of fuzzy rough neural network via distributed parallelism for stock prediction[J]. *IEEE Transactions on Fuzzy Systems*, 2020, 28(5): 939–952
- [33] Liao Hongmei, Ding Shifei, Wang Miaomiao, et al. An overview on rough neural networks[J]. *Neural Computing and Applications*, 2016, 27(7): 1805–1816
- [34] Jahangir H, Tayarani H, Baghali S, et al. A novel electricity price forecasting approach based on dimension reduction strategy and rough artificial neural networks[J]. *IEEE Transactions on Industrial Informatics*, 2019, 16(4): 2369–2381
- [35] Sabzalian M H, Mohammadzadeh A, Lin Shuyi, et al. A robust control

- of a class of induction motors using rough type-2 fuzzy neural networks[J]. *Soft Computing*, 2020, 24(13): 9809–9819
- [36] Ning Shangming, Teng Fei, Li Tianrui. Entity relation extraction from electronic medical records based on multi-channel self-attention mechanism[J]. *Chinese Journal of Computers*, 2020, 43(5): 916–929 (in Chinese)
(宁尚明, 滕飞, 李天瑞. 基于多通道自注意力机制的电子病历实体关系抽取[J]. *计算机学报*, 2020, 43(5): 916–929)
- [37] Chen Yong, Liu Xi, Liu Huanlin. Occluded pedestrian detection method based on feature channel and spatial joint attention mechanism[J]. *Journal of Electronics and Information*, 2020, 42(6): 1486–1493 (in Chinese)
(陈勇, 刘曦, 刘换淋. 基于特征通道和空间联合注意机制的遮挡行人检测方法[J]. *电子与信息学报*, 2020, 42(6): 1486–1493)
- [38] Ma Jiayi, Zhang Hao, Yi Peng, et al. SCSCN: A separated channel-spatial convolution net with attention for single-view reconstruction[J]. *IEEE Transactions on Industrial Electronics*, 2019, 67(10): 8649–8658
- [39] Song Jie, Xiao Liang, Lian Zhichao, et al. Review and prospect of digital pathology image segmentation based on deep learning[J]. *Journal of Software*, 2021, 32(5): 1427–1460 (in Chinese)
(宋杰, 肖亮, 莲智超, 等. 基于深度学习的数字病理图像分割综述与展望[J]. *软件学报*, 2021, 32(5): 1427–1460)
- [40] Liang Liming, Liu Bowen, Yang Hailong, et al. Supervised retinal vessel extraction based on multi-feature fusion[J]. *Chinese Journal of Computers*, 2018, 41(11): 2566–2580 (in Chinese)
(梁礼明, 刘博文, 杨海龙, 等. 基于多特征融合的有监督视网膜血管提取[J]. *计算机学报*, 2018, 41(11): 2566–2580)
- [41] Liskowski P, Krawiec K. Segmenting retinal blood vessels with deep neural networks[J]. *IEEE Transactions on Medical Imaging*, 2016, 35(11): 2369–2380
- [42] Badrinarayanan V, Kendall A, Cipolla R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481–2495



Sun Ying, born in 1997. Master candidate. Her main research interests include granular computing, rough sets, and deep learning.
孙 颖, 1997 年生. 硕士研究生. 主要研究方向为粒计算、粗糙集和深度学习.



Ding Weiping, born in 1979. PhD, professor, PhD supervisor. Senior member of CCF. His main research interests include data mining, machine learning, granular computing, evolutionary computing, and big data analytics.

丁卫平, 1979 年生. 博士, 教授, 博士生导师, CCF 高级会员. 主要研究方向为数据挖掘、机器学习、粒计算、演化计算和大数据分析.



Huang Jiashuang, born in 1988. PhD, lecturer. His main research interests include brain network analysis and deep learning. (hjshdym@163.com)
黄嘉爽, 1988 年生. 博士, 讲师. 主要研究方向为脑网络分析和深度学习.



Ju Hengrong, born in 1989. PhD, associate professor. His main research interests include granular computing, rough sets, machine learning, and knowledge discovery. (juhengrong@ntu.edu.cn)

鞠恒荣, 1989 年生. 博士, 副教授. 主要研究方向为粒计算、粗糙集、机器学习和知识发现.



Li Ming, born in 1996. Master candidate. His main research interests include data mining, granular computing, and big data analytics. (liming_2014@163.com)

李 铭, 1996 年生. 硕士研究生. 主要研究方向为数据挖掘、粒度计算和大数据分析.



Geng Yu, born in 1998. Master candidate. His main research interests include granular computing, machine learning, and deep learning. (tian19981999@163.com)

耿 宇, 1998 年生. 硕士研究生. 主要研究方向为粒计算、机器学习和深度学习.