

文章编号:1673-9469(2025)06-0105-08

DOI:10.3969/j.issn.1673-9469.2025.06.013

基于对比与重构的无监督多粒度眼底图像分割

倪健,王峥,赵伟康,王子鹏,韩宇轩,王毅飞

(河北工程大学 信息与电气工程学院,河北 邯郸 056038)

摘要: 针对眼底图像分割任务对大规模标注数据依赖强、特征表达不足以及分割结果单一的问题,提出了一种无监督多粒度视网膜图像分割方法。构建了新型的全卷积编码器-解码器结构,用以充分捕捉图像的局部细节与全局语义特征,实现多层次特征的高效重构。在此基础上,设计了一种综合损失函数,将像素级补丁对比损失、表征级对比学习损失与全局重构损失进行联合优化,通过多尺度特征约束强化模型的表征能力,使特征空间更适配分割任务的结构分布。在得到的表征空间中引入扩散凝聚算法,以聚合多尺度语义信息,提升分割边界的精确性与整体结构的连贯性,并生成具有层次性和多粒度特征的分割结果。在公开的视网膜眼底数据集上对提出的方法进行验证,结果显示,该方法在无监督条件下的 Dice 系数较当前主流无监督分割模型平均提升 3.7%,在分割细节保真度与结构一致性方面均表现出显著优势,该方法能够有效实现眼底图像的高精度、多粒度分割。

关键词: 眼底图像分割;无监督;对比学习;全局重构;扩散凝聚

中图分类号: TP399

文献标识码: A

Unsupervised Multi-granularity Segmentation of Fundus Images Based on Comparison and Reconstruction

NI Jian, WANG Zheng, ZHAO Weikang, WANG Zipeng, HAN Yuxuan, WANG Yifei

(School of Information and Electrical Engineering, Hebei University of Engineering, Handan, Hebei 056038, China)

Abstract: To address the challenges of strong dependence on large-scale labeled data, insufficient feature representation, and single-granularity segmentation results in retinal fundus image segmentation, an unsupervised multi-granularity segmentation method for retinal images is proposed. A novel fully convolutional encoder-decoder architecture was designed to effectively capture local details and global semantic features of images, achieving efficient reconstruction of multi-level representations. On this basis, a comprehensive loss function was constructed by integrating pixel-level patch contrastive loss, representation-level contrastive learning loss, and global reconstruction loss. This joint optimization constrained the model across multiple feature scales, enhancing the representation capability and aligning the feature space with the structural distribution of the segmentation task. Subsequently, a diffusion-condensation algorithm was applied within the representation space to aggregate multi-scale semantic information, improving boundary precision and structural coherence, and generating segmentation results with hierarchical and multi-granular characteristics. Experiments conducted on publicly available retinal fundus datasets demonstrated that the proposed method achieved a 3.7% improvement in Dice coefficient compared with state-of-the-art unsupervised segmentation approaches, showing superior performance in both detail fidelity and structural consistency. The results indicated that this method enabled accurate and multi-granularity segmentation of retinal fundus images.

收稿日期:2024-06-28

基金项目:河北省自然科学基金资助项目(F2023402011)

第一作者:倪健(1975—),女,河北邯郸人,硕士,教授,从事计算机视觉方面研究。

Key words: fundus image segmentation; unsupervised learning; contrastive learning; global reconstruction; diffusion condensation

随着中国人口老龄化程度不断加深,老年人口持续增长,给医疗公共资源带来了前所未有的压力。年龄相关性黄斑变性(Age-related Macular Degeneration, AMD)是一种常见的老年疾病,其晚期形态——地图状萎缩(Geographic Atrophy, GA)可导致不可逆的视力损伤,该疾病已引起广泛关注^[1]。在此背景下,深度学习在医学图像分割领域的应用为AMD的诊断提供了新视角,基于神经网络模型,可以有效识别和定位视网膜眼底图像中的关键区域,如不同器官组织和细胞类型,从而为临床诊断提供更精准的依据。

尽管深度学习在性能和效率方面取得显著提升,但仍面临诸多挑战,如对大量标注数据的依赖、医学数据隐私保护问题、跨设备泛化能力差以及分割结果粒度单一等^[2-5]。特别是在处理GA等复杂医学图像时,由于血管结构复杂且萎缩程度多样,分割难度显著上升。为解决这一问题,引入不同粒度的分割结果图显得尤为重要,这些图像能够有效区分萎缩程度,从而为病因的精准诊断提供关键依据。

近年来,对比学习作为一种新的学习范式,通过学习图像的特征表示来解决标注数据稀缺的问题。然而现有的大多数对比学习方法^[6-11]侧重于全局图像表示,忽略了局部细节信息,并依赖大规模有标签数据进行预训练,导致图像分割效果不理想。局部对比学习无监督多粒度分割方法虽然关注局部信息并生成多粒度结果图,但忽视了全局特征^[12],而全局特征对于保持图像结构一致性和提升分割精度至关重要,尤其在多粒度分割任务中。

针对现有方法的局限性,本文提出了一种基于对比与重构的无监督多粒度眼底图像分割方法。该方法设计了一种新型的全卷积编码解码网络,通过融合局部对比与全局重构损失,构建了适配分割任务的特征空间,在特征空间中引入扩散凝聚算法。本文所建方法显著提升了分割精度,实现了多样化的多粒度分割。

1 方法

1.1 算法总体框架

本研究设计的模型旨在通过重设计网络结

构,引入新学习策略,更准确地捕捉图像的局部细节和全局信息,从而提高分割的准确性。模型整体架构分为两部分,如图1所示。第一部分:输入与特征提取。本文采用结合局部补丁对比损失、像素补丁对比损失和全局重构损失等策略,利用视网膜眼底图像数据集对编码器网络进行训练,确保模型有效学习图像的结构和纹理信息。第二部分:分割与后处理。在学习到的特征中,通过扩散凝聚方法生成多粒度的视网膜分割。

1.2 全卷积主干网络

全卷积编码器-解码器主干网络的核心主要包括3个方面:(1)创新性地利用卷积操作代替传统网络中显式的图像补丁裁剪,实现局部特征的高效聚合;(2)采用独特的8层卷积设计,基于镜像对称的卷积-转置卷积布局,适应任意尺寸图像输入;(3)摒弃传统卷积网络中的池化操作,有效避免池化过程中产生的特征流失,为对比学习与重构网络提供了充分特征信息支撑,助力构建优质的图像分割特征空间。

1.2.1 全卷积编码器

编码器的作用是将“像素为中心的补丁” \mathcal{P}_{ij} 从原始的低维图像空间映射到扩展的高维特征向量空间 R^d ,从而实现局部补丁的对比学习。为了在像素中心定义隐式的补丁,本文基于卷积的局部特性,通过单层卷积在每个像素的 $k \times k$ 邻域中整合信息,这相当于进行了 $k \times k$ 的补丁裁剪。鉴于实际感受野通常小于理论值,本文引入多层卷积,卷积层数越多,信息整合效果越明显。最终,位于 (i, j) 的嵌入 z_{ij} 汇聚了远超单一像素的广泛区域信息,其范围由感受野界定,从而可作为“以像素 (i, j) 为中心的补丁”的嵌入向量。该策略高效整合了周围信息,无需直接裁剪图像即可实现局部补丁特征提取。

本文的具体操作如下:编码器采用4层 5×5 卷积核的卷积层,以整合每个像素周围的广泛上下文信息。随着网络深度的增加,卷积核的数量逐步增加。最后一个卷积层配备了 d 个卷积核,实现每个输入像素向 d 维嵌入空间的映射。在所有实验中, d 设定为128。因此,宽度 W 、高度 H 、通道数 c 的图像被编码器从实数空间 $R^{H \times W \times c}$ 转换

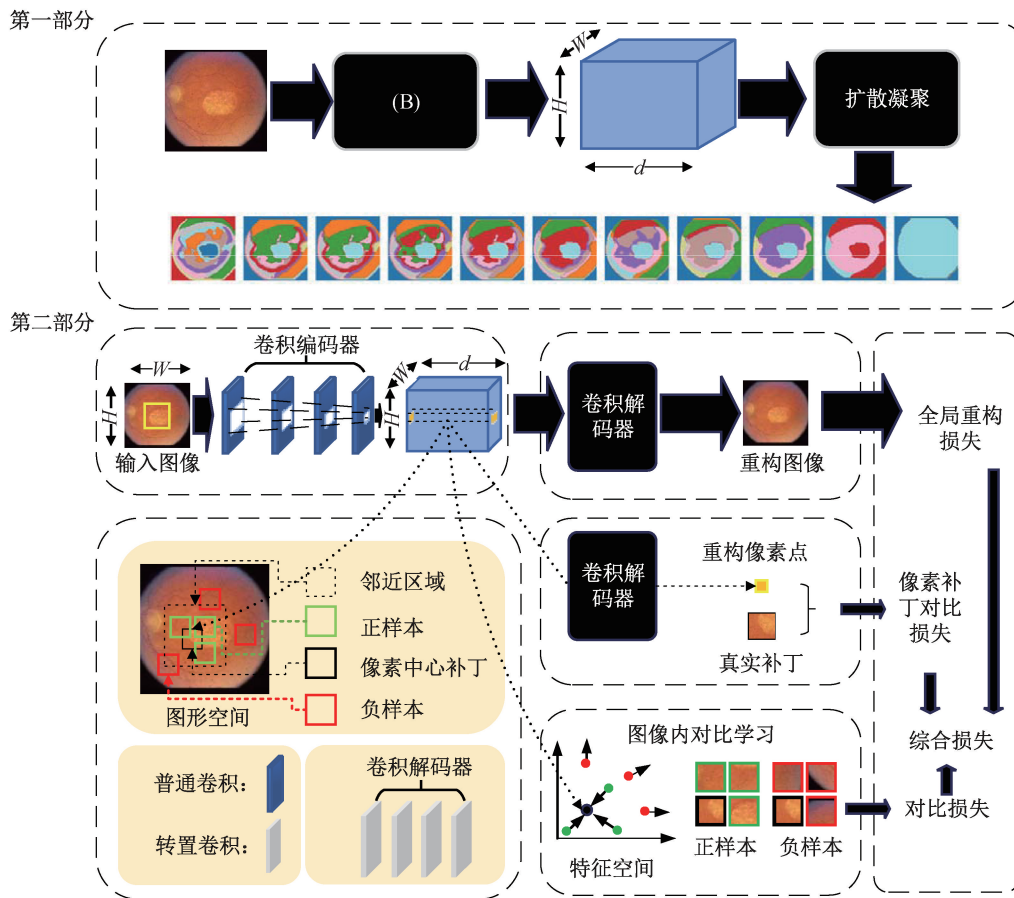


图 1 网络框架图

Fig. 1 Network framework diagram

为 $R^{H \times W \times d}$ 。虽然 4 层卷积神经网络(具有 5×5 卷积核和单位步长)的理论感受野为 17×17 ,但实际有效感受野通常小于理论值。为确保充分保留局部补丁信息,本文设置的补丁宽度 \mathcal{P} 略小于理论感受野,从而得到像素补丁 $\mathcal{P}_{ij}(\mathcal{P}_{ij} \in R^{5 \times 5 \times c})$ 。经过 4 层卷积处理,得到的特征图中每个特征向量代表了以特定像素为中心的补丁特征向量 $z_{ij} \in R^{H \times W \times d}$,具体细节如表 1 所示。

表 1 卷积编码器各层参数设置

Tab. 1 Parameter settings for each layer of the convolutional encoder

结构名称	卷积核	步长	输出尺寸($H \times W \times d$)
卷积层 1	(5, 5)	1	$128 \times 128 \times 16$
卷积层 2	(5, 5)	1	$128 \times 128 \times 32$
卷积层 3	(5, 5)	1	$128 \times 128 \times 64$
卷积层 4	(5, 5)	1	$128 \times 128 \times 128$

1.2.2 卷积解码器

解码器采用 4 个转置卷积层,与编码器镜像对称,确保空间特征向量在传递中有效恢复图像的

全局和局部信息,精准再现图像特征细节。编码器和解码器协同工作构成一个高效的全卷积网络框架,支持多种尺寸图像输入。解码器同样仅包含卷积层,避免池化操作导致的特征信息损失,为对比学习与重构提供了完整的特征表示,具体如表 2 所示。

表 2 卷积解码器各层参数设置

Tab. 2 Parameter settings for each layer of the convolutional decoder

结构名称	卷积核	步长	输出尺寸($H \times W \times d$)
转置卷积层 1	(5, 5)	1	$128 \times 128 \times 256$
转置卷积层 2	(5, 5)	1	$128 \times 128 \times 128$
转置卷积层 3	(5, 5)	1	$128 \times 128 \times 64$
转置卷积层 4	(5, 5)	1	$128 \times 128 \times 3$

1.3 损失函数设计

为解决现有方法在捕捉局部细节与全局信息上的不足,本文提出了融合局部对比学习与全局重构的综合损失函数。该策略可确保编码器有组织地将相似的像素中心补丁映射至特征空间。

1.3.1 局部补丁对比损失

局部补丁对比损失作为损失函数的核心,其独特之处在于实现了图像内部的局部对比学习策略。与现有大规模预训练模型在图像级别进行预训练的方式不同,该策略在图像内部进行局部级别的对比学习,这一创新显著节省了数据资源,尤其适用于数据稀缺且标注困难的医学图像领域。

局部补丁对比损失利用邻近启发式和图像结构相似性(Structural Similarity Index, SSIM)区分正样本与负样本。正样本是与锚点补丁 z_{ij} 相似度较高的邻域补丁,其相似度高于预设阈值,而负样本是与锚点补丁距离较远或未通过相似性检查的补丁。在嵌入空间中进行对比学习,比较每个锚点补丁的嵌入向量 z_{ij}^+ 和 z_{ij}^- 。进而生成正、负样本对,并计算对比损失。该损失函数在嵌入空间中进行对比学习,具体计算如公式(1)所示。

$$l_{\text{contrastive}} = -\log \frac{\sum_{z_{ij}^+ \in \Omega^+} e^{\text{sim}(z_{ij}^+, z_{ij}^+)/\tau}}{\sum_{z_{ij}^- \in \Omega^-} e^{\text{sim}(z_{ij}^+, z_{ij}^-)/\tau}} \quad (1)$$

式中, $l_{\text{contrastive}}$ 为对比损失函数; z_{ij} 为位于 (i, j) 位置的锚点补丁的嵌入向量; z_{ij} 为分母中的一个样本; z_{ij}^+, z_{ij}^- 分别为 z_{ij} 的正样本、负样本; Ω 为所有正、负样本的完整集合; Ω^+ 为 z_{ij} 的正样本集合; τ 为温度参数; $\text{sim}(\cdot)$ 为相似度函数; e 为自然常数。

最小化损失函数的目的是鼓励网络将正样本集合 Ω^+ 嵌入到靠近锚点的位置,同时根据负样本集合 Ω^- 与锚点的距离调整靠近的程度。最小化损失函数促使正样本对更相似,而负样本则更远离锚点。温度参数 τ 在对比损失中用于控制距离的缩放方式,本文在所有实验中将其取值为 0.5。相似度函数 $\text{sim}(\cdot)$ 则设定为余弦相似度。

1.3.2 像素补丁对比损失

为了确保中间特征向量 z_{ij} 充分保留原始图像补丁 $\mathcal{P}_{ij} \in R^{5 \times 5 \times c}$ 及其像素的精细特征,本文量化了原始补丁与解码器重构后特征向量 $f_{\text{recon}}(z_{ij})$ 的差异。该方法通过损失函数的差异最小化,促使网络对补丁局部细节的精确重构,从而实现局部补丁特征向量的高效信息编码。该编码过程为后续的对比如学习和扩散凝聚算法^[13-14]提供了更丰富且精确的特征表示。

$$l_{\text{recon}} = \| \mathcal{P}_{ij} - f_{\text{recon}}(z_{ij}) \|^2 \quad (2)$$

式中, l_{recon} 为重建损失; $f_{\text{recon}}(\cdot)$ 为解码器网络,负

责从 z_{ij} 重建回原始补丁; $\|\cdot\|_2^2$ 表示平方 L^2 范数,用于衡量原始图像与重构图像之间的差异。

1.3.3 全局图像重构损失

在训练过程中,为了全面捕捉图像的纹理、结构信息,确保解码器能从编码器学到的特征向量中还原出原始图像的整体视觉内容,本文创新性地设计、引入全局图像重构损失。具体来说,本文利用解码器 $f_{\text{global}}(\cdot)$ 从实数空间 $R^{H \times W \times c}$ 转换为 $R^{H \times W \times d}$,进而处理编码器输出的全局特征图,见式(3)。

$$l_{\text{global}} = \| x - f_{\text{global}}(x) \|^2 \quad (3)$$

式中, l_{global} 为全局重构损失; x 为原始的、完整的输入图像; $f_{\text{global}}(\cdot)$ 为全局重构解码器网络,负责将特征图映射回图像空间。

1.3.4 综合损失

本文的网络目标函数由局部补丁对比损失、像素补丁对比损失和全局图像重构损失组成,并通过调整权重系数 λ ($\lambda \in [0, 1]$) 平衡各部分之间的相对大小,见式(4)。

$$\text{loss} = \lambda \cdot l_{\text{contrastive}} + (1 - 2 \cdot \lambda) \cdot l_{\text{recon}} + \lambda \cdot l_{\text{global}} \quad (4)$$

式中, loss 为最终的综合损失; λ 为用于平衡不同损失项的权重系数,经多次实验,本文将 λ 设置为 0.01。

1.4 扩散凝聚

经过卷积编码器的训练,输入图像中每个以像素为中心的补丁均被映射至维度为 R^d 的向量空间。这些嵌入向量经过扩散凝聚算法处理,实现图像分割。具体来说,对于坐标为 (i, j) 的每个图像补丁 \mathcal{P}_{ij} ,卷积编码器为其分配一个 d 维向量 z_{ij} ($z_{ij} \in R^d$)。对于尺寸为 $W \times H$ 的图像,嵌入空间中将存在 $W \times H$ 个 d 维向量。本文采用聚类算法 $\text{cls}(\cdot)$,将每个 d 维的特征向量映射为一个标量簇标签,从而将这些向量分配至 n 个不同的簇 $\{c_1, c_2, \dots, c_n\}$ 中。由此,本文构建出一个标签映射 L ($L \in R^{H \times W}$),其中 $L_{ij} = \text{cls}(z_{ij})$ 。最终的分割结果即为标签图 L 。 $\text{cls}(\cdot)$ 在整个扩散凝聚过程中有效变化,可以生成一组丰富的标签 L ,每个标签都是特定粒度的分割。多粒度的扩散凝聚是一个动态过程,可探索各种粒度级别,识别数据的自然分组^[13-14]。通过迭代扩散过程,将数据压缩到相邻的数据点,其速率由点之间的扩散概率定义。与大多数聚类方法不同,扩散凝聚构建了从粗到

细粒度的完整层次结构,其中每个粒度的聚类数量不是任意设置的,而是从数据的底层结构推断出来的。具体来说,根据具有 N 个观测值(图像中的像素数 $N = W \times H$)和 d 个特征的数据矩阵 $X^{N \times d}$,构建每个观测对之间的局部相似性 $(m, n) \in (1, \dots, N)$ 。本文采用高斯核函数计算,见式(5)。

$$K(x_m, x_n) = e^{-\frac{x_m - x_n^2}{\epsilon}} \quad (5)$$

式中, K 是一个 $N \times N$ 的 Gram 矩阵,其 (m, n) 项表示为 $K(x_m, x_n)$,强调对数据矩阵 X 的依赖性; ϵ 为带宽参数,控制邻域大小; e 为自然常数。

X_m 和 X_n 的维度均为 R^d ,给定 Gram 矩阵 K ,扩散算子 P 定义为

$$P = D^{-1}K \quad (6)$$

$$D(x_m, x_m) = \sum_n K(x_m, x_n) \quad (7)$$

式中, D 为对角线矩阵,其对角线元素 $D(x_m, x_m)$ 由公式(7)给出,通过对亲和矩阵 K 的第 m 行求和,得到代表数据点 x_m 的总亲和度; D^{-1} 为 D 的逆矩阵; P 为扩散算子,定义数据扩散过程的单步转移概率,可以视为马尔可夫随机游走。

为了执行多步扩散,通过将扩散算子 P 提升到 t 次幂来模拟时间均匀扩散过程,即 $X_t = P^t X$ 。另外,可以通过式(8)的迭代计算扩散算子 P 和数据矩阵,模拟时间非均匀扩散过程。

$$\begin{aligned} X_0 &\leftarrow X \\ \text{for } k \in [1, \dots, t]: \\ K_{k-1} &\leftarrow K(X_{k-1}) / \text{使用等式 5} / \\ D_{k-1} &\leftarrow D(K_{k-1}) / \text{使用等式 7} / \\ P_{k-1} &\leftarrow D_{k-1}^{-1} K_{k-1} / \text{使用等式 6} / \\ X_k &\leftarrow P_{k-1} X_{k-1} \end{aligned} \quad (8)$$

式中, t 为设定的总迭代次数; $k(k \in [1, \dots, t])$ 为当前迭代步数; X_0 为初始数据矩阵 X ; X_{k-1} 、 K_{k-1} 、 D_{k-1} 和 P_{k-1} 分别代表在第 $k-1$ 次迭代时的数据矩阵、亲和矩阵、对角线矩阵和扩散算子。

扩散凝聚的过程可以概括为以下两个步骤之间的交替:(1)根据迭代 t 次的数据计算时间非均匀扩散算子;(2)将此算子应用到数据上,将点移向局部重心,从而在迭代 $t+1$ 处形成数据。有关扩散凝聚的详细信息参考文献[13-14]。

2 实验

2.1 数据集

本研究聚焦于 GA 区域的识别与分割,采用公

开的彩色眼底图像数据集开展研究^[12]。该数据集的标注专注于 GA-AMD 晚期的关键症状,其特征是黄斑的逐渐退化。GA 识别面临多重挑战,包括 GA 颜色与周围组织的相似性、视神经和视盘周围萎缩易混淆以及血管结构的复杂性。这些周围组织的特征粒度差异显著影响分割算法的有效性。数据集包含 56 张已精确标注的图像,包括 GA 病变分割、视盘轮廓勾画及中心凹标记,所有标注均经分级人员初步评估并由视网膜专家复核。

在数据集划分上,本研究遵循文献[12]的建议,采用无监督学习方法特有的划分策略。因模型训练不依赖真实标签,为进一步确认该划分策略的稳健性,本文按 8:1:1 的比例开展验证实验,即将数据集随机划分为 80% 的训练集、10% 的验证集和 10% 的测试集,划分策略与文献[12]的建议方法相近。

2.2 实验处理

2.2.1 实验环境

模型基于 Python 3.8 和 PyTorch 实现,适用于深度学习任务。训练和测试在一台服务器上进行,该服务器配备两张 16 GB 显存的 DCU 加速卡和 40 GB 内存,为模型训练提供了有力的硬件支持。学习速度设置为 0.001,批量大小设置为 16,表示每次模型训练时,使用 16 个样本计算损失并进行权重更新。模型共训练 200 轮,每轮包含整个训练数据集的一次迭代。

2.2.2 评价指标

为全面评估模型的表现,本文选用 Dice 系数、Hausdorff 距离、SSIM 和均方根误差 (Root Mean Square Error, RMSE) 等评价指标。Dice 系数用于比较样本间的相似性,Hausdorff 距离用于衡量点集间的距离,SSIM 关注图像的结构相似性, RMSE 主要用于评估模型的预测能力。在不同的场景和应用需求下,以上指标的选取和计算方法可灵活调整。

2.3 实验结果

2.3.1 模型比较

在对比实验中,本文选取当前无监督分割领域的六大主流方法作为参照,这些方法涵盖了不同技术路径和应用场景:传统的基于图像处理的 Watershed 方法^[15]和 Felzenszwalb 算法^[16],侧重于利用手工特征进行图像分割;SLIC 超级像素^[17]则通过高效的聚类生成规整的图像基元,在手工特

征与深度特征之间构成了重要的过渡;基于大规模数据预训练的可微分特征聚类(Differentiable Feature Clustering, DFC)^[18]和STEGO^[19]方法,其在利用大量数据进行无监督特征学习方面具有优势;专为小数据量及医学图像设计的最先进的SOTA(state-of-the-art)无监督分割方法,即对比与无监督分割训练(Contrastive and Unsupervised Training for Segmentation, CUTS)^[12]。在成功复现上述方法的基础上,本研究提取各方法原论文中针对本文数据集的代表性实验结果,将其与本文方法的结果进行对比,如表3所示。

表3 视网膜眼底图像数据集对比实验结果
Tab.3 Compared experimental results of retinal fundus image dataset

方法	Dice 系数	Hausdorff 距离	SSIM	RMSE
Watershed ^[15]	0.187	61.600	0.177	0.825
Felzenszwalb ^[16]	0.222	64.050	0.164	0.833
SLIC ^[17]	0.567	28.760	0.779	0.323
DFC ^[18]	0.300	46.470	0.400	0.641
STEGO ^[19]	0.649	34.120	0.842	0.250
CUTS ^[12]	0.735	25.280	0.874	0.203
本文方法	0.772	23.881	0.899	0.175

由表3对比分析可知,基于大规模数据预训练的DFC与STEGO分割方法,在面对仅包含56张小规模视网膜眼底图像的分割任务时,均表现出明显的不适应性。CUTS虽然在小规模数据集上略有优势,但因未能充分捕捉眼底图像的全局信息及特征细节,其性能仍然受限。而本文提出的方法在小规模数据集上展现出显著优势,与CUTS相比,Dice系数提升了3.7%,Hausdorff距离减少了1.399%,SSIM指数增加了2.5%,RMSE指数降低了2.8%。

视网膜眼底图像示例分割图见图2,从两个方面展示了不同方法的效果:(1)输入图像经过特征学习后,采用不同方法处理得到的分割结果(奇数行);(2)相应的重构图像及其特征的可视化展示(偶数行)。

通过比较分割结果可以看出,k-means算法与扩散凝聚多粒度分割方法在性能和效果上存在显著差异。特别是,通过扩散凝聚多粒度分割方法得到了从细粒度到粗粒度的一系列分割结果,实现了视网膜结构的多粒度精准分割。该方法在粗粒度上可有效区分视网膜与背景,在细粒度上可清晰呈现视神经盘、地理萎缩区域等关键结构,

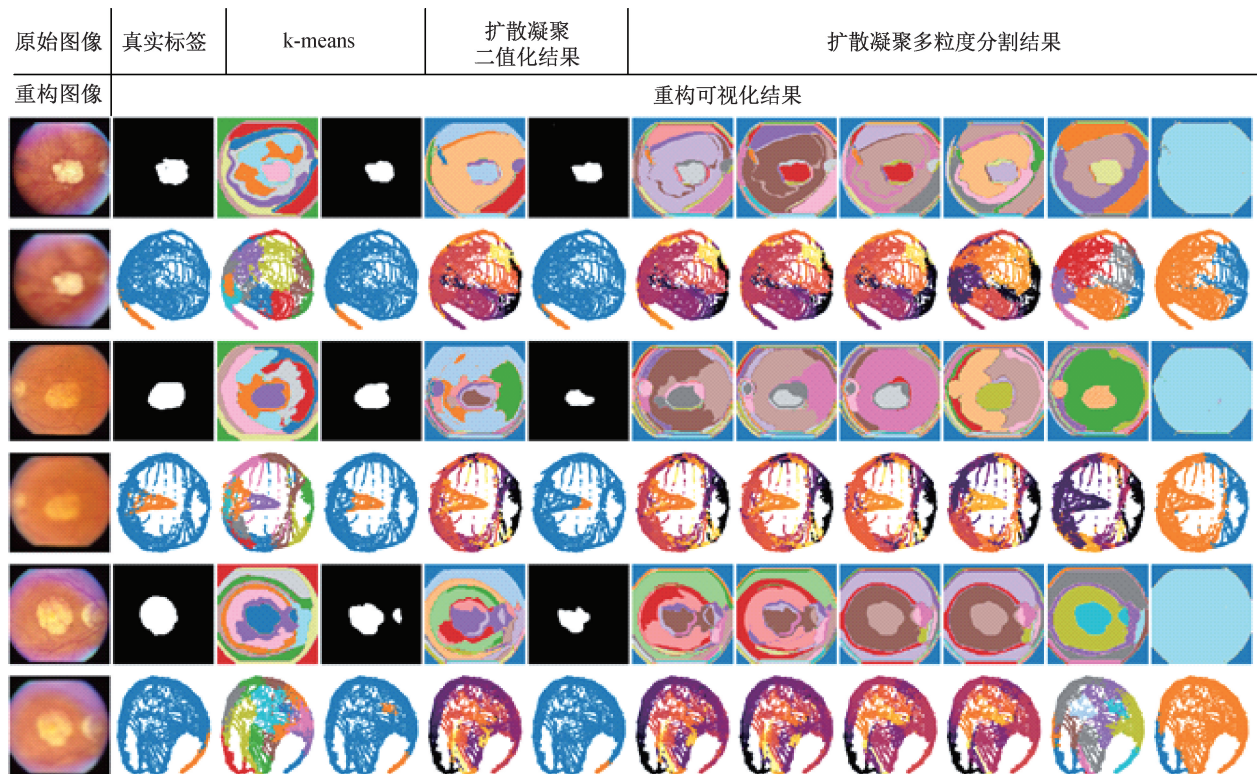


图2 视网膜眼底图像示例分割图

Fig. 2 Example segmentation diagram of retinal fundus image

表 4 视网膜眼底图像数据集上的消融研究分割结果

Tab. 4 Ablation study segmentation results on the retinal fundus image dataset

方法	$l_{\text{contrastive}}$	l_{recon}	l_{global}	扩散凝聚	Dice 系数	Hausdorff 距离	SSIM	RMSE
基准方法	✓				0.498	39.964	0.783	0.284
A	✓			✓	0.589	34.874	0.844	0.236
B		✓			0.641	31.489	0.867	0.231
C		✓		✓	0.736	24.507	0.885	0.195
D			✓		0.637	29.433	0.863	0.231
E			✓	✓	0.596	30.680	0.767	0.318
F	✓	✓			0.667	24.906	0.87	0.223
G	✓	✓		✓	0.742	20.604	0.883	0.198
H	✓	✓	✓		0.663	28.853	0.871	0.223
I	✓	✓	✓	✓	0.772	23.881	0.899	0.175

从而为自动化测量及临床干预规划提供了有力支持。另外,该方法在颜色相似性挑战下仍能保持高准确性,能准确识别萎缩中心区域,且受血管结构干扰极小,便于医生选取针对性的最优结果。

在重构图像及特征可视化方面,本文采用的全卷积网络架构结合综合损失函数有效提升了图像重构质量及特征空间分布的合理性。重构图像纹理细腻清晰,学习后的特征空间中形成明显的簇结构,为后续分割任务奠定了坚实基础。总体而言,以上结果证明了本文方法在视网膜图像分割任务中的有效性及其在多粒度下捕捉视网膜结构的能力。

2.3.2 消融实验

为了验证了本文提出的基于对比与重构的无监督多粒度分割算法各模块的有效性,本节设计了消融实验,结果见表 4。实验在相同条件下进行,采用编码器与局部补丁对比损失进行特征提取,并在第二阶段采用 k-means 进行分割。在此基础上,依次进行如下改进:(1)将 k-means 替换为扩散凝聚算法(改进 A);(2)将损失函数替换为像素补丁对比损失(改进 B);(3)在 B 的基础上,第二阶段采用扩散凝聚算法(改进 C);(4)将基准网络中的算法替换为全局重构函数(改进 D);(5)在 D 的基础上,第二阶段同样采用扩散凝聚算法(改进 E);(6)将损失函数替换为像素补丁对比和局部像素对比函数(改进 F);(7)在 F 的基础上,第二阶段采用扩散凝聚算法(改进 G);(8)基准网络同时采用三种损失(改进 H);(9)在 H 的基础上,第二阶段采用扩散凝聚算法即本文算法(改进 I)。实验结果表明,本文所提出的算法展现出卓越性能,有效验证了其内部结构的合理性,成功达成了预期研究目标。

3 结论

本文提出了一种基于对比与重构的无监督多粒度视网膜眼底图像分割模型。该模型通过全卷

积网络融合像素级补丁对比、表征级对比学习与全局重构进行优化,并引入扩散凝聚算法以实现多粒度的分割输出,得到以下结论:

1)模型在无监督条件下,Dice 系数较主流无监督模型平均提升 3.7%,在分割细节保真度与结构一致性上均表现出显著优势。

2)该方法能够有效实现眼底图像的高精度、多粒度分割,为后续视网膜疾病的自动识别与病灶分析提供了更可靠的图像基础支撑,展现出较高的临床应用潜力。

参考文献:

- [1] 晏鑫,李凌,关瑞娟. 年龄相关性黄斑变性中西医结合治疗新进展[J]. 中国老年学杂志, 2022, 42(17): 4379-4383.
- [2] YAN X, LI L, GUAN R J. New advances in the treatment of age-related macular degeneration with integrated Chinese and western medicine[J]. Chinese Journal of Gerontology, 2022, 42(17): 4379-4383.
- [3] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [4] ESTEVA A, ROBICQUET A, RAMSUNDAR B, et al. A guide to deep learning in healthcare[J]. Nature Medicine, 2019, 25(1): 24-29.
- [5] HAQUE I R I, NEUBERT J. Deep learning approaches to biomedical image segmentation[J]. Informatics in Medicine Unlocked, 2020, 18: 100297.
- [6] MINAEE S, BOYKOV Y, PORIKLI F, et al. Image segmentation using deep learning: a survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(7): 3523-3542.
- [7] CHEN T, KORNBLITH S, NOROUZI M, et al. A simple framework for contrastive learning of visual representations[C]//Proceedings of the 37th International Conference on Machine Learning. PMLR, 2020: 1597-1607.
- [8] CARON M, MISRA I, MAIRAL J, et al. Unsupervised learning of visual features by contrasting cluster assignments[C]//Proceedings of Advances in Neural Informa-

- tion Processing Systems(NeurIPS), 2020: 1-10.
- [8] HE K M, FAN H Q, WU Y X, et al. Momentum contrast for unsupervised visual representation learning[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 9726-9735.
- [9] GRILL J B, STRUB F, ALTCHÉ F, et al. Bootstrap your own latent-a new approach to self-supervised learning[J]. *Advances in Neural Information Processing Systems*, 2020, 33: 21271-21284.
- [10] ZBONTAR J, JING L, MISRA I, et al. Barlow twins: self-supervised learning via redundancy reduction[C]//Proceedings of the 38th International Conference on Machine Learning. PMLR, 2021: 12310-12320.
- [11] CHEN X L, HE K M. Exploring simple Siamese representation learning[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 15745-15753.
- [12] LIU C, AMODIO M, SHEN L, et al. CUTS: a deep learning and topological framework for multigranular unsupervised medical image segmentation[C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2024. Cham: Springer, 2024: 155-165.
- [13] BRUGNONE N, GONOPOLSKIY A, MOYLE M W, et al. Coarse graining of data via inhomogeneous diffusion condensation[C]//2019 IEEE International Conference on Big Data, 2020: 2624-2633.
- [14] KUCHROO M, DISTASIO M, SONG E, et al. Single-cell analysis reveals inflammatory interactions driving macular degeneration[J]. *Nature Communications*, 2023, 14(1): 2589.
- [15] VINCENT L, SOILLE P. Watersheds in digital spaces: an efficient algorithm based on immersion simulations[J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 1991, 13(6): 583-598.
- [16] FELZENSZWALB P F, HUTTENLOCHER D P. Efficient graph-based image segmentation[J]. *International Journal of Computer Vision*, 2004, 59: 167-181.
- [17] ACHANTA R, SHAJI A, SMITH K, et al. SLIC superpixels compared to state-of-the-art superpixel methods[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(11): 2274-2282.
- [18] KIM W, KANEZAKI A, TANAKA M. Unsupervised learning of image segmentation based on differentiable feature clustering[J]. *IEEE Transactions on Image Processing*, 2020, 29: 8055-8068.
- [19] HAMILTON M, ZHANG Z T, HARIHARAN B, et al. Unsupervised semantic segmentation by distilling feature correspondences[PP/OL]. arXiv(2022-03-16) [2024-06-28]. <https://doi.org/10.48550/arXiv.2203.08414>. (责任编辑 周雪梅)
- (上接第 104 页)
- [17] BERRUTI G, CONSALES M, GIORDANO M, et al. Radiation hard humidity sensors for high energy physics applications using polyimide-coated fiber Bragg gratings sensors[J]. *Sensors and Actuators B: Chemical*, 2013, 177(11): 94-102.
- [18] DING F X, WANG L T, FANG N, et al. Experimental study on humidity sensing using a FBG sensor with polyimide coating[C]//Asia Communications and Photonics Conference and Exhibition. IEEE, 2010: 280-281.
- [19] 国际法制计量组织. 饱和盐溶液标准相对湿度值; OIML R121[S]. 北京: 全国物理化学计量技术委员会, 2000.
International Organization of Legal Metrology. The standard humidity value relatively of saturated salt solution; OIML R121[S]. Beijing: Metrology Technical Committee on Physical Chemistry (MTC17), 2000.
- [20] HUANG X F, SHENG D R, CEN K F, et al. Low-cost relative humidity sensor based on thermoplastic polyimide-coated fiber Bragg grating[J]. *Sensors and Actuators B*, 2007, 127(2): 518-524.
- [21] 张向东, 李育林, 彭文达, 等. 光纤光栅型温湿度传感器的设计与实现[J]. *光子学报*, 2003, 32(10): 1166-1169.
ZHANG X D, LI Y L, PENG W D, et al. Design and realization of temperature and relative humidity sensor based on FBG[J]. *Acta Photonica Sinica*, 2003, 32(10): 1166-1169.
- [22] 向光华, 忽满利, 乔学光, 等. 基于聚酰亚胺材料的 FBG 湿度传感特性研究[J]. *光电子·激光*, 2012, 23(1): 41-45.
XIANG G H, HU M L, QIAO X G, et al. Study on the characteristics of FBG humidity sensor based on polyimide[J]. *Journal of Optoelectronics · Laser*, 2012, 23(1): 41-45.
- [23] 曹红葵. 聚酰亚胺性能及合成方法[J]. *化学推进剂与高分子材料*, 2008, 6(3): 24-26.
CAO H K. Properties of polyimide and its synthesis methods[J]. *Chemical Propellants & Polymeric Materials*, 2008, 6(3): 24-26.
- [24] 丁永红. 聚酰亚胺的合成方法[J]. *化工中间体*, 2011, 8(5): 36-43.
DING Y H. Polymerization approach for polyimide preparation[J]. *Chemical Intermediate*, 2011, 8(5): 36-43.
- [25] YEO T L, SUN T, GRATTAN K T V, et al. Polymer-coated fiber Bragg grating for relative humidity sensing[J]. *IEEE Sensors Journal*, 2005, 5(5): 1082-1089.
- [26] 崔千红. 聚酰亚胺电容型湿度传感器的研制[D]. 兰州: 兰州大学, 2010.
CUI Q H. Development and fabrication of polyimide capacitive humidity sensors[D]. Lanzhou: Lanzhou University, 2010. (责任编辑 周雪梅)