

基于多监督注意力机制神经网络的脑胶质瘤循环肿瘤细胞分割算法

袁红杰¹, 杨艳¹, 张东¹, 杨双^{1,2}

1. 武汉大学物理科学与技术学院, 湖北 武汉 430072; 2. 桂林航天工业学院电子信息与自动化学院, 广西 桂林 541004

【摘要】为了提升脑胶质瘤循环肿瘤细胞的分割准确率,解决人工分割中肉眼分辨边界困难、目标占比小和操作流程繁琐等问题,提出一种端到端的像素级分割算法。针对数据特征,提出一种基于多监督机制的混合损失函数用以提升预测区域与目标区域的交并比,同时训练网络向预测正确目标个数的方向收敛;其次,在网络中逐层加入卷积块注意力机制模块,使得网络能在空间、通道层面重点学习数据特征,进一步提升预测准确率;最后,通过采用混合训练的方式,只需一个网络模型就能直接分割出细胞核、细胞质区域,缩减训练流程。实验结果表明,此分割算法对比U-Net网络在召回率、精确率以及Dice系数方面均有显著提升,在细胞核分割方面,分别达到92.20%、86.56%、88.27%;在细胞质分割方面,分别达到89.33%、85.31%、86.33%。

【关键词】脑胶质瘤;循环肿瘤细胞;多监督;卷积块注意力机制模块;小目标分割

【中图分类号】R318;TP391

【文献标志码】A

【文章编号】1005-202X(2022)07-0828-06

Neural network-based multi-level supervision and attention mechanism algorithm for brain glioma CTC segmentation

YUAN Hongjie¹, YANG Yan¹, ZHANG Dong¹, YANG Shuang^{1,2}

1. School of Physics and Technology, Wuhan University, Wuhan 430072, China; 2. School of Electronic Information and Automation, Guilin University of Aerospace Technology, Guilin 541004, China

Abstract: In order to improve the segmentation accuracy of glioma circulating tumor cells (CTC) and solve the problems of difficulties in distinguishing boundaries with the naked eye, small target ratio and cumbersome operation process in manual segmentation, an end-to-end pixel-level segmentation algorithm is proposed. In view of data features, the proposed algorithm utilizes a hybrid loss function based on a multi-level supervision mechanism to improve the intersection-over-union between the prediction area and the ground truth area, and iterate the network converging in the direction of predicting the right numbers of targets. Then, the convolutional block attention module is put in every layer of the proposed network enables the network to focus on learning data features at the spatial and channel dimensions, thereby further improving the prediction accuracy. Finally, the proposed algorithm can segment the nucleus and cytoplasm by one network model through hybrid training, which simplifies the process of training. The experimental results showed that compared with U-Net network, the proposed segmentation algorithm has improved in terms of recall rate, precision and Dice coefficient. The above-mentioned indexes are 92.20%, 86.56%, 88.27% for cell nucleus segmentation, and 89.33%, 85.31%, 86.33% for cytoplasm segmentation.

Keywords: glioma; circulating tumor cell; multi-level supervision; convolutional block attention module; small target segmentation

前言

1869年, Ashwonh^[1]对1例癌症死亡患者进行尸检时发现其外周血存在类似肿瘤细胞,首次提出了循环肿瘤细胞(Circulating Tumor Cells, CTC)的概念。CTC是指由癌灶组织细胞脱离原发灶或转移灶释放进入外周血液循环的肿瘤细胞,存在于外周血

【收稿日期】2022-02-26

【基金项目】国家重点研发计划973项目(2011CB707900)

【作者简介】袁红杰, 硕士, 研究方向: 医学图像处理, E-mail: hongji_yuan@163.com

中的各类肿瘤细胞的统称^[2]。CTC 的检测可有效地应用在许多方面,包括化疗药物的快速评估、个体化治疗的耐药性检测、肿瘤复发的监测以及肿瘤新药开发等^[3]。近年来的研究表明脑胶质瘤 CTC 可以作为脑胶质瘤诊断的特异性分子标志物,并在评估放化疗疗效、鉴别肿瘤复发和假性进展、筛选分子靶向药物等方面具有极大的潜在临床价值。2020 年, Qi 等^[4]首次使用基于滤孔过滤的分离技术对脑胶质瘤 CTC 成功进行染色荧光标记。该标记技术优点颇多,但对于标记结果的判读却流程繁琐,其染色标记后的识别主要依赖于医师的专业经验,需要在显微镜下识别 CTC 的同时手动画出细胞核细胞质的区域。此类方法需要医师操作时注意力高度集中,且对边界模糊的 CTC 识别率低。为了更好地辅助医师,利用计算机学方法自动识别、分离、计算出 CTC 细胞核、细胞质区域面积成为当前研究的首要任务。

目前医学图像分割算法研究可以分为两大类。第一类是基于概率统计学的传统算法,例如阈值分割法、基于区域的分割法、基于边缘的分割方法、基于形态学的分割方法等,这类算法的鲁棒性和泛化性较差且分割结果极易受到噪声干扰^[5]。第二类是近些年较主流的基于机器学习的方法,例如监督学习、无监督学习和强化学习等,这些算法均需要搭建合适的神经网络框架,投入大量的数据,学习数据的共有特征,再通过优化自身网络参数来减小预测值和实际值的差距,以达到分割新数据的目的^[6-11]。2015 年, Ronneberger 等^[12]提出 U-Net 网络架构,该架构是基于 Encoder-Decoder 的经典思想,很好地解决了医学图像分割任务中因数据集收集困难导致的分割精度较差的难题。2018 年, Woo 等^[13]提出卷积块的注意力机制模块,该模块是在通道和空间维度上进行重点提取,使得神经网络更集中于目标物体识别,提高分割准确率。

1 方法

专业医师对 CTC 进行识别需要按照以下步骤: 第一步需要在 DAPI 染色(蓝色)层面,确定细胞核位置,并且细胞核大小需要大于滤孔直径且无分叶;第二步需要在 STEAM 染色(绿色)层面,确定细胞核周围胞质表达目标抗原的区域(即 CTC 细胞质绿染区域);第三步在 CD45 染色(红色)层面,确定 CTC 是否表达白细胞抗原(即红染区域),如果细胞红染、绿染面积比大于 0.3,则认为是白细胞,予以排除,不计入 CTC。按照以上要求,如果利用传统算法方式分割 CTC,需要按照上述 3 步依次进行,流程复杂且分割效果极易受到噪声干扰。因此,本研究选择机器学习

方法实现分割 CTC^[14]。按照常规思维,分割细胞质、细胞核属于两个分割任务,需要搭建两个独立的网络分割。本研究基于数据特点选择使用混合训练的方法,一次投入 3 通道的混合数据,网络直接输出 2 通道预测分割图。

1.1 算法流程

针对图像质量参差不齐的脑胶质瘤 CTC 免疫荧光图像,将根据原始数据的特点搭建合适的网络架构以满足最终的分割需求。本研究整体算法流程大致可以分为 3 个步骤:(1)数据集制作,包含图像预处理、标签制作、数据集增强等操作^[15];(2)搭建基于注意力机制的分割网络,并选择合适的损失函数;(3)投入测试数据集,分析预测结果。整体流程如图 1 所示。

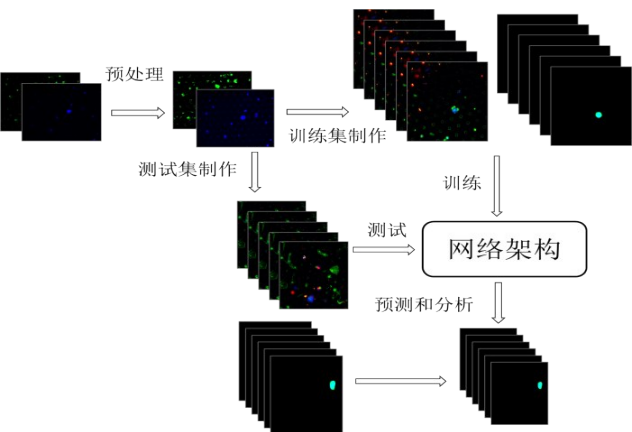


图1 算法流程
Figure 1 Algorithm flowchart

1.2 网络结构

设计的网络构架如图 2 所示,该网络架构主要由两个部分构成,分别为左边的 Encoder 部分和右边的 Decoder 部分,也就是比较经典的编码-解码模型^[16]。该网络主要由卷积模块、卷积块注意力机制(Convolutional Block Attention Module, CBAM)模块、下/上采样模块、计数模块构成,同层编码器的输出联结前一层的上采样输出作为解码器卷积模块的输入,以防止随着网络深度的增加输入特征丢失。网络的输入为 3 通道的混合图片(DAPI 蓝色、STEAM 绿色、CD45 红色),而输出为 2 通道的二值图片(即细胞核、细胞质的预测分割图)。

卷积模块(Convolutional Block, Conv Block)如图 3 所示,包含两个部分,它们均是由卷积函数、批标准化函数、Dropout 函数、线性整流激活 Relu 函数构成^[17-20]。其中卷积核选择为 3,步长为 1,为反射填充,Dropout 的比例为 0.3。

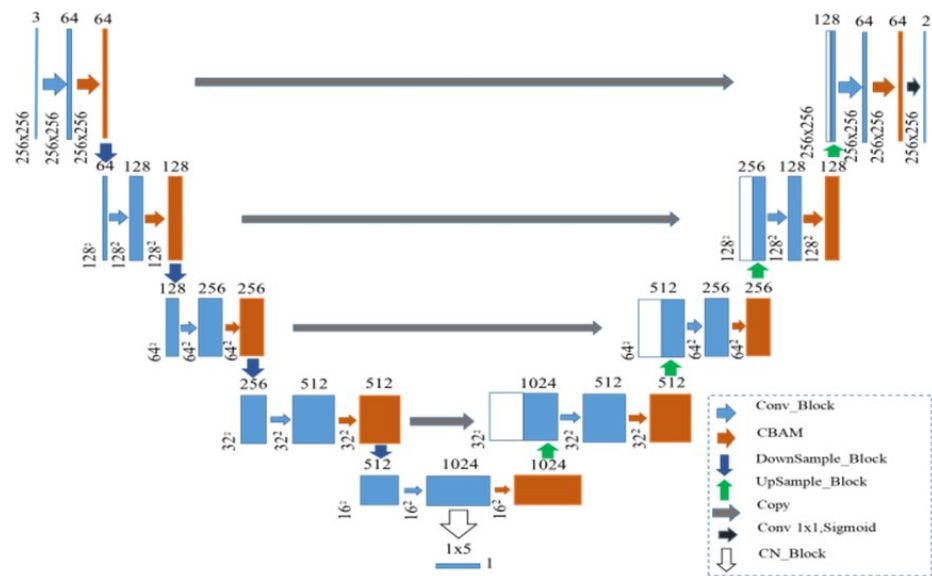


图 2 多监督注意力机制神经网络结构

Figure 2 Structure of neural network-based multi-level supervision and attention mechanism

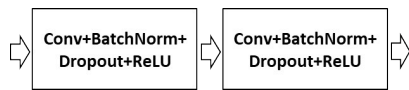


图 3 卷积模块

Figure 3 Convolutional block

CBAM 模块结构如图 4a 所示,主要包含两个部分:第一部分为通道注意力机制^[21];第二部分为空间注意力机制^[22]。CBAM 模块中的两部分串行处理能够很好地帮助网络学习输入特征的空间和通道维度的权重分布。输入特征图 F (Shape 为 $H \times W \times C$) 通过通道注意力模块 Channel Attention Module,得到 $1 \times 1 \times C$ 的通道注意力特征图 M_c 。再将 M_c 和原始特征图 F 做矩阵乘法,生成 F_c (Shape 为 $H \times W \times C$)。 F_c 再通过空间注意力模块 Spatial Attention Module,得到 $H \times W \times 1$ 的空间注意力特征 M_s 。再将 M_s 和特征图 F_c 做矩阵乘法,生成 F' (Shape 为 $H \times W \times C$)。通道注意力机制结构如图 4b 所示。主要是将输入的特征图 F 分别经过高为 H 和宽为 W 的全局最大池化(Global max pooling)和全局平均池化(Global average pooling),分别得到 1 个 $1 \times 1 \times C$ 的特征图,接着,再将它们分别送入独立、隐藏层为 1 的全连接神经网络(Multilayer Perceptron, MLP)^[23],其中激活函数设置均为 Relu。下一步将 MLP 输出的特征进行基于以通道维度为基准的矩阵加法操作。最后,将混合后的特征矩阵投入 Sigmoid 函数激活,生成最终的 Channel

attention feature,即 M_c 。空间注意力机制结构如图 4c 所示。主要是将输入的特征图 F_c 首先做一个基于通道维度为基准的 Global max pooling 和 Global average pooling,分别得到 1 个 Shape 为 $H \times W \times 1$ 的特征图,然后将这 2 个特征图做通道拼接操作。再经过一个卷积核为 7、步长为 1 的卷积操作,并把通道维数降为 1,Shape 即为 $H \times W \times 1$ 。投入 Sigmoid 函数激活生成 Spatial attention feature,即 M_s 。

下采样 (DownSample Block) 和上采样

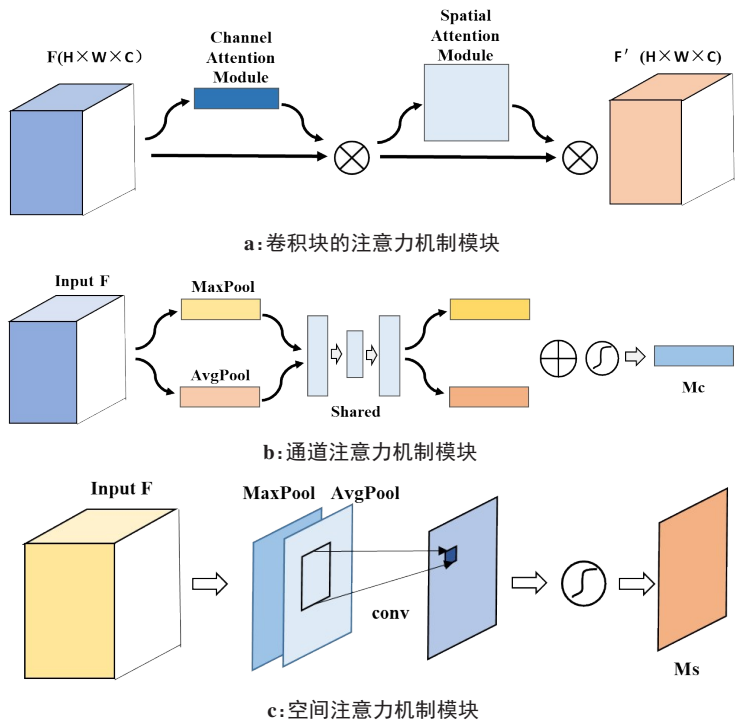


图 4 卷积块注意力机制模块

Figure 4 Convolutional block attention module

(UpSample Block)模块的作用是使特征图升维或降维以满足网络在不同尺度下的特征提取和传播^[24]。下采样模块由一个步长为2的卷积函数、批标准化函数、Relu激活函数构成。上采样模块为插值函数、批标准化函数、Relu激活函数构成。

计数模块(Counting Number Block, CN Block),其结构如图5所示。输入特征分别经过全局 Global max pooling、Global average pooling得两个 $1 \times 1 \times C$ 的特征矩阵,两特征矩阵做矩阵加法,再通过Relu激活函数和隐藏层为1的MLP,得到分割目标个数的预测输出。

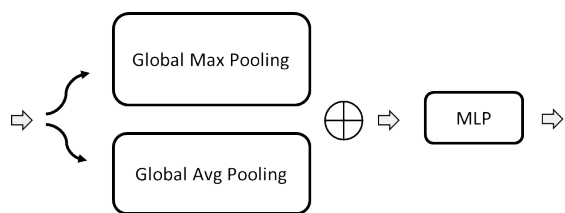


图5 计数模块

Figure 5 Counting number block

1.3 损失函数

损失函数是评判预测值和目标值之间差距程度的函数,选择一个合适的目标函数有助于网络模型预测效果、鲁棒性和收敛速度的提升,同时防止在反向传播时梯度消失和梯度爆炸问题。本研究提出自定义 Loss 作为损失函数,本文 Loss 主要由两部分构成, Loss 表达式如式(1)所示。第一部分为 l_1 (Dice Loss),如式(2)所示,主要目的是最大化目标矩阵和预测矩阵之间的交并比;第二部分为 l_2 (Cross Entropy Loss, 交叉熵损失函数),如式(3)所示,主要帮助网络向预测正确目标个数的方向收敛。

$$\text{Loss} = l_1 + l_2 \quad (1)$$

$$l_1(T, P) = 1 - \frac{2|P \cap T|}{|P| + |T|} \quad (2)$$

$$l_2(T_i, P_i) = - \sum_{i=1}^n T_i * \log(\text{softmax}(P_i)) \quad (3)$$

其中, P 表示预测矩阵, T 表示真实值矩阵, n 为预测类别个数, T_i 为目标个数为 i 的真实概率, P_i 为目标个数为 i 的预测概率。可以看到在几乎不用增加模型复杂度的情况下,使用混合自定义 Loss 函数,能在更多维度去监督模型的学习以提升模型的拟合能力。

2 实验与结果

本研究所有网络框架的运行环境均为 Anaconda3-2021.11-Linux-x86_64,均使用 Pytorch 网络框架(其中 Cuda 版本为 10.2、Python 版本为 3.8.12),均使用同一组自制训练/测试数据集(其中训

练集与测试集比为 10:1)。数据集的原始图片尺寸为 720×720 ,为了满足网络的输入输出尺寸要求,均压缩至 256×256 的尺寸(使用 INTER_AREA 插值方式,即像素区域关系重新采样)。所有实验网络均使用适应性较强的 Adam 优化器进行网络优化,其中动量值、衰减参数、学习率等参数均选择默认值^[25]。所有网络均训练 2000 轮,并保存训练过程最优模型。

2.1 实验数据

本文使用 Qi 等^[4]提供的基于滤孔过滤技术拍摄的 CTC 荧光染色图。数据集制作:第一步需要对 DAPI 蓝色、STEAM 绿色、CD45 红色原图进行合并,分别选择其 B、G、R 通道数据置于混合 B、G、R 图片的 3 通道中;第二步针对绿色和蓝色的图片进行人工标注,标出细胞核和细胞质的面积。由于原图亮度、对比度不一致,大部分数据难以凭肉眼分辨边界,需要对原图进行自适应的亮度和对比度增强。接着在 Qi 团队专业医师的指导下对改善后的图像进行标注,生成目标和背景的二值图。由于网络输入长宽比为 1:1(原图为 4:3),为防止输入长宽比压缩不一致,需要对原图进行裁减,流程图如图 6 所示。完成上述操作后对二值图进行归一化生成值为 0、1 的标签矩阵,共生成 189 组数据。由于数据量较小,需要对数据进行增强以扩充数据集来满足训练需求。为了防止数据集中每张图片目标个数和目标大小不一致对训练结果产生负面影响,本文采用分组增强数据集的方式进行数据扩充。按照分割目标为 1 个和多个的标准把数据分为 A、B 两组。A 组包含 139 张, B 组包含 50 张。为了使测试集和训练集中的原始数据占比一致,通过计算,本文选择从 A、B 两组分别随机取出 15、5 张,分别扩充至 30 张,组成 60 张的测试数据。将 A、B 每组剩余部分分别扩充至 300 张,总数 600 张作为训练数据。数据增强的方式为对图像随机进行旋转、缩放、裁减、亮度修改、对比度修改等混合方式,尽量减少人工干预对数据集扩充的影响。

2.2 评价指标

为了客观分析算法模型的预测性能,选择召回率(Recall)、精确率(Precision)以及 Dice 系数 3 类评价指标来衡量分割出的细胞核、细胞质和实际人为标注区域的差异。计算公式如式(4)~(6)所示:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

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

$$\text{Dice} = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (6)$$

其中,真阳性(True Positives, TP)为预测正样本,实际为正样本;假阳性(False Positives, FP)为预测正样

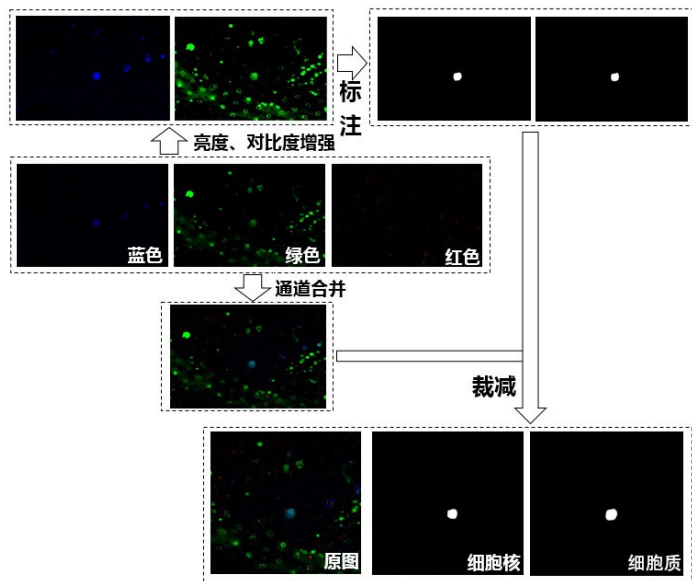


图 6 数据集处理流程图
Figure 6 Dataset processing flowchart

本,实际为负样本;真阴性(True Negatives, TN)为预测负样本,实际为负样本;假阴性(False Negatives, FN)为预测负样本,实际为正样本。

2.3 实验结果与分析

为了定量分析本文算法模型,实验对比了 U-Net 模型和 U-Net+Diceloss 模型,分别预测、计算、对比 3 类指标的平均值,结果如表 1、表 2 所示。在分割 CTC 细胞核、细胞质任务上,可以看到本研究算法在对比 U-Net+Diceloss 模型方面,可以在不牺牲预测精确率的情况下,在召回率和 Dice 系数方面均有明显提升:细胞核分割方面,分别提升 4.40%、2.51%;细胞质分割方面,分别提升 5.59%、3.17%。而对比 U-Net 模型,可以看到本研究算法 3 项指标均有明显提高,细胞核分割方面,分别提升 10.04%、1.02%、7.13%;细胞质分割方面,分别提升 8.85%、0.51%、7.38%。

表 1 细胞核分割对比定量表现(%)
Table 1 Quantitative results of different algorithms for nucleus segmentation (%)

算法模型	召回率	精确率	Dice 系数
U-Net	82.16	85.54	81.14
U-Net+Diceloss	87.80	86.88	85.76
本文算法	92.20	86.56	88.27

为定性分析本文算法优势,对比人工标注、U-Net 和本文算法分割结果图,可以看出当针对单个目标的分割,两者表现不分上下,如图 7a 所示。而在分割亮度、对比度较差的图片时,U-Net 只找到部分目标区域,而本文算法可以找到全部区域且分割效果较

表 2 细胞质分割对比定量表现(%)
Table 2 Quantitative results of different algorithms for cytoplasm segmentation (%)

算法模型	召回率	精确率	Dice 系数
U-Net	80.48	84.80	78.95
U-Net+Diceloss	83.74	85.89	83.16
本文算法	89.33	85.31	86.33

优,如图 7b 所示。而针对多目标分割的任务,在对比度较低的条件,本文算法依然能够找到多个分割目标,表现也明显优于 U-Net,如图 7c 所示。

3 结 语

本研究提出一种基于多监督注意力机制神经网络的脑胶质瘤 CTC 细胞核、细胞质分割算法,解决了人工分割脑胶质瘤 CTC 免疫荧光图像的边界分辨困难、操作流程复杂等问题。针对数据样本分割目标个数不一致的问题,本研究选择按比例增强数据,配平数据比。针对网络算法方面,首先提出混合训练的方式,只需一个网络模型,一次输入混合图片,网络直接预测细胞核、细胞质的分割图,缩减训练流程;其次提出一种混合 Loss,最大化分割区域与目标区域的交并比,并训练网络向预测正确目标个数的方向收敛;最后,本研究加入 CBAM 注意力机制模块,使得网络能在空间、通道层面重点学习数据特征。通过在同等条件下与 U-Net 算法的对比,确定本研究的算法优势,为后续计算脑胶质瘤 CTC 细胞核质比用于评估放化疗疗效、鉴别肿瘤复发和假性进展、筛选分子靶向药物等方面提供帮助。

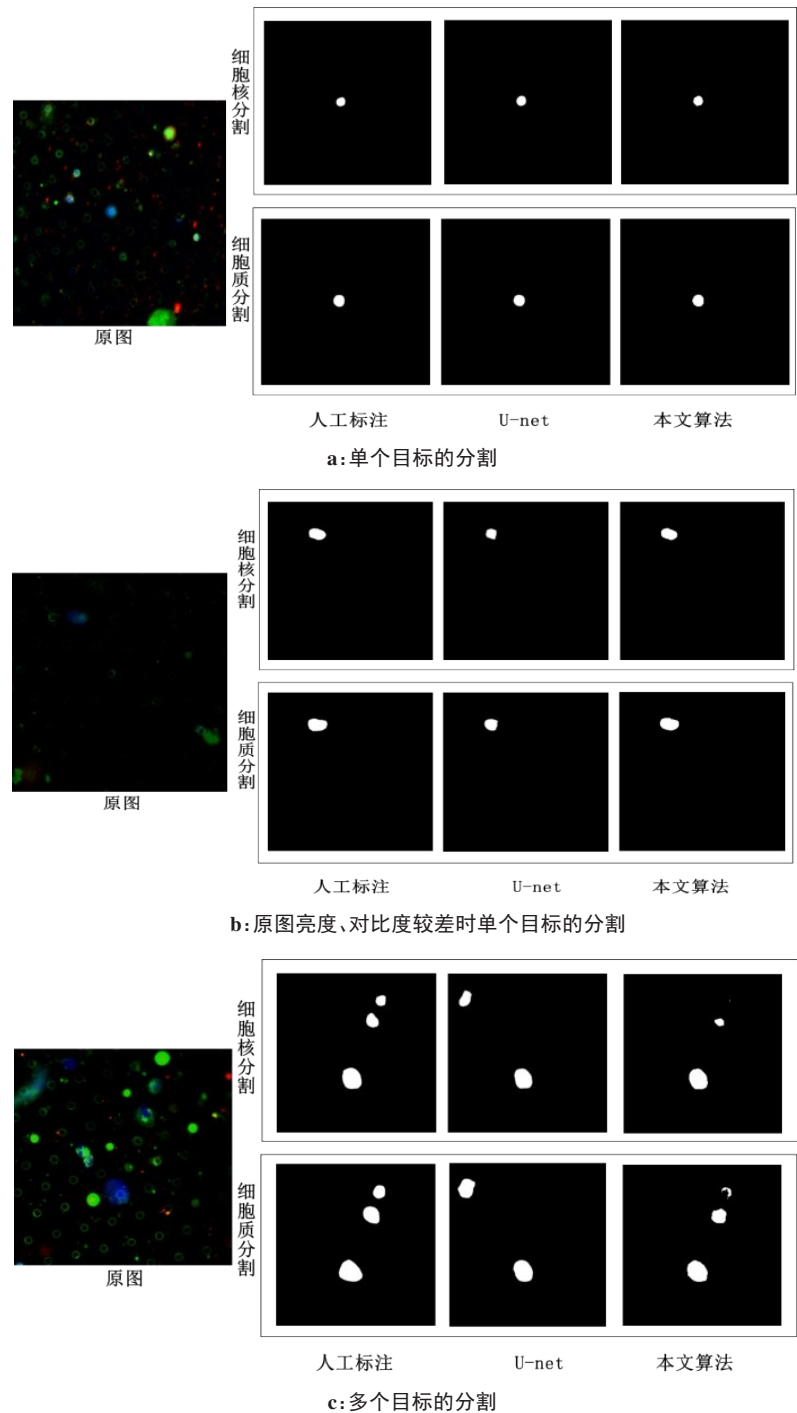


图7 预测分割图

Figure 7 Predicted segmentation

【参考文献】

[1] Ashwonh TR. A case of cancer in which cells similar to those in the tumors were seen in the blood after death[J]. Aust Med J, 1869, 14(1): 146-149.

[2] 王建, 邵荣金, 龚伟达. 循环肿瘤细胞研究进展[J]. 中国肿瘤外科杂志, 2019, 11(2): 141-145.

Wang J, Shao RJ, Gong WD. Advancement of research on circulating tumors[J]. Chinese Tumor Surgery, 2019, 11(2): 141-145.

[3] 王瑞玲, 李雪, 孙鹏, 等. 结直肠癌循环肿瘤细胞的研究进展[J]. 中华临床医师杂志(电子版), 2014, 16: 3025-3030.

Wang RL, Li X, Sun P, et al. Colorectal cancer circulation aneurysm detailed research development [J]. Chinese Journal of Clinicians (Electronic Edition), 2014, 16: 3025-3030.

[4] Qi YZ, Sun Q, Deng G, et al. Identifying circulating glioma cells and their clusters as diagnostic markers by a novel detection platform[J]. Clin Transl Med, 2021, 11(2): e318(1-7).

[5] Tarachandani A, Boltz D. Review of the basic image processing and segmentation techniques for biological images[J]. J Imaging Sci Technol, 2006, 50(3): 233-242.

[6] Somasundaram SK, Alli P. A review on recent research and implementation methodologies on medical image segmentation[J]. J Comput Sci, 2012, 8(1): 170-174.

[7] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation[J]. IEEE Trans Pattern Anal Mach Intell, 2015, 39(4): 640-651.

[8] Zhao B, Yu F, Sun J, et al. Glandular cell segmentation method combined with dense connective blocks and self-attention mechanism [J]. J Comput Desg Comput G, 2021, 33(7): 991-999.

[9] Zhu Y, Erik M. Automatic improvement of deep learning-based cell segmentation in time-lapse microscopy by neural architecture search [J]. Bioinformatics, 2021, 37(24): 4844-4850.

[10] Liu Z, Lin Y, Cao Y, et al. Swin transformer: hierarchical vision transformer using shifted windows [J]. arXiv preprint, 2021, arXiv: 2103.14030.

[11] Huang X, Deng Z, Li D, et al. MISSFormer: an effective medical image segmentation transformer [J]. arXiv preprint, 2021, arXiv: 2109.07162.

[12] Ronneberger O, Fischer P, Brox T, et al. U-Net: convolutional networks for biomedical image segmentation [J]. arXiv preprint, 2015, arXiv: 1505.04597.

[13] Woo S, Park J, Lee JY, et al. CBAM: convolutional block attention module [J]. arXiv preprint, 2018, arXiv: 1807.06521.

[14] Yu Y, Li M, Liu L, et al. Clinical big data and deep learning: applications, challenges, and future outlooks [J]. Big Data Min Anal, 2019, 2 (4): 70-75.

[15] Shorten C, Khoshgoftaar TM. A survey on image data augmentation for deep learning [J]. J Big Data, 2019, 6(1): 1-48.

[16] Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation [J]. IEEE Trans Pattern Anal Mach Intell, 2017, 39(12): 2481-2495.

[17] Lecun Y, Bottou L. Gradient-based learning applied to document recognition [J]. Proc IEEE, 1998, 86(11): 2278-2324.

[18] Carl M, Robert H, Tracy F, et al. Understanding the normalization of telemedicine services through qualitative evaluation [J]. JAMIA Open, 2003, 10(6): 596-604.

[19] Wahlbeck K, Tuunainen A, Ahokas A, et al. Dropout rates in randomised antipsychotic drug trials [J]. Psychopharmacology, 2001, 155(3): 230-233.

[20] Guarnieri S, Piazza F. Multilayer feedforward networks with adaptive spline activation function [J]. IEEE Trans Neural Netw, 1999, 10(3): 672-683.

[21] Haut JM, Fernandez-Beltran R, Paoletti ME, et al. Remote sensing image superresolution using deep residual channel attention [J]. IEEE Geosci Remote Sens, 2019, 57(11): 9277-9289.

[22] Du W, Wang Y, Yu Q. Recurrent spatial-temporal attention network for action recognition in videos [J]. IEEE Trans Image Process, 2017, 27 (99): 1347-1360.

[23] Yan H, Jiang Y, Zheng J, et al. A multilayer perceptron-based medical decision support system for heart disease diagnosis [J]. Expert Syst Appl, 2006, 30(2): 272-281.

[24] Zhang YB, Zhao DB, Zhang J, et al. Interpolation-dependent image downsampling [J]. IEEE Trans Image Process, 2011, 20(11): 3291-3296.

[25] Chandriah KK, Naraganahalli RV. RNN/LSTM with modified Adam optimizer in deep learning approach for automobile spare parts demand forecasting [J]. Multimed Tools Appl, 2021, 80(17): 26145-26159.

(编辑:陈丽霞)