

# 基于 CNN 的相衬显微图像序列的癌细胞多目标跟踪

胡海根<sup>1</sup> 周莉莉<sup>1</sup> 周乾伟<sup>1</sup> 陈胜勇<sup>1,2</sup> 张俊康<sup>1</sup>

(浙江工业大学计算机科学与技术学院 杭州 310024)<sup>1</sup>

(天津理工大学计算机科学与工程学院 天津 300384)<sup>2</sup>

**摘要** 检测与跟踪相衬显微图像序列下的癌细胞对于分析癌细胞的生命周期以及开发抗癌新药具有非常重要的意义。传统的目标跟踪方法大多应用于刚性目标跟踪或单目标跟踪,而癌细胞是非刚性且不断裂变的多目标,这就大大增加了跟踪的难度。文中以相衬显微图像序列中的膀胱癌细胞为研究对象,提出一种基于卷积神经网络(Convolutional Neural Network, CNN)的癌细胞多目标跟踪方法。该算法采用基于检测的多目标跟踪方法,首先利用深度学习检测框架 Faster R-CNN 卷积神经网络实现癌细胞的检测,初步获得待跟踪的癌细胞;再利用扫描圆算法(Circle Scanning Algorithm, CSA)实现黏连细胞的检测优化,进一步提高黏连区域的细胞检测精度;最后提取综合特征描述子,对卷积特征、尺寸特征和位置特征进行加权求和,实现跟踪目标的综合描述,从而实现不同帧癌细胞间的高效关联匹配,最终实现癌细胞的多目标跟踪。一系列实验结果表明,相较于传统方法,所提方法不仅在癌细胞的检测和跟踪上性能有较大的提升,而且可以有效处理目标的遮挡问题。

**关键词** 卷积神经网络,细胞检测,Faster R-CNN,扫描圆算法,多目标跟踪

**中图分类号** TP183 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2019.05.043

## Multi-target Tracking of Cancer Cells under Phase Contrast Microscopic Images Based on Convolutional Neural Network

HU Hai-gen<sup>1</sup> ZHOU Li-li<sup>1</sup> ZHOU Qian-wei<sup>1</sup> CHEN Sheng-yong<sup>1,2</sup> ZHANG Jun-kang<sup>1</sup>

(College of Computer Science & Technology, Zhejiang University of Technology, Hangzhou 310024, China)<sup>1</sup>

(School of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China)<sup>2</sup>

**Abstract** Detecting and tracking cancer cells under phase contrast microscopic images plays a critical role for analyzing the life cycle of cancer cells and developing new anti-cancer drugs. Traditional target tracking methods are mostly applied to rigid target tracking or single target tracking, while cancer cells are non-rigid multiple targets with constant fission, and it makes tracking more challenging. Taking bladder cancer cells in the sequence of phase contrast micrographs images as research object, this paper proposed a multi-target tracking method of cancer cells based on convolutional neural network. Firstly, through making use of detection-based multi-target method, the proposed algorithm adopted the deep learning detection framework Faster R-CNN to detect the bladder cancer cells and preliminarily obtain the cancer cells to be tracked. Then CSA (circle scanning algorithm) was utilized to further optimize the detection of adhesion cancer cells, and further improve the detection accuracy of cells in adhesion area. Finally, it integrated the features of convolution, size and position into a synthetic feature descriptor by using weighting methods, thus tracking multiple cancer cells by achieving the efficient correlation and matching of different frames of cancer cells. The results of a series of experiments show that this method can not only improve the accuracy of detecting and tracking cancer cell, but also deal with the occlusion problem effectively.

**Keywords** Convolutional neural network, Cell detection, Faster R-CNN, Circle scanning algorithm, Multi-target tracking

## 1 引言

在生物与医学领域,通过观察显微图像序列中癌细胞在

药物作用下的生命活动,描述癌细胞的完整运动轨迹,实现癌细胞在药物作用后生命周期的自动分析对研制与开发相应的抗癌药物有着重要的意义。早期癌细胞生命周期的分析是通

收稿日期:2018-06-30 返修日期:2018-08-10 本文受浙江省自然科学基金(LY18F030025),国家自然科学基金(61802347, U1509207, 31640053),中国微系统技术重点实验室基金(6142804010203)资助。

胡海根(1977—),男,博士,副教授,CCF会员,主要研究方向为深度学习、计算机视觉、进化计算以及温室环境智能化控制等,E-mail: hgghu@zjut.edu.cn(通信作者);周莉莉(1994—),女,硕士生,主要研究方向为深度学习,E-mail: 2352535359@qq.com;周乾伟(1986—),男,博士,主要研究方向为深度学习、信号处理及优化;陈胜勇(1973—),男,博士,教授,博士生导师,主要研究方向为计算机视觉;张俊康(1993—),男,硕士,主要研究方向为深度学习。

过人工肉眼观察癌细胞的活动来实现,因为存在人眼疲劳及其他不可控因素,使得跟踪准确率过低。随着计算机视觉的广泛应用<sup>[1]</sup>,目前传统细胞跟踪算法主要分为3类:1)使用模型演变的方法;2)基于目标检测与分割的方法;3)基于概率论的预测方法。这些方法在实现癌细胞的跟踪过程中面临诸多挑战,例如:不同癌细胞间形态存在较大的相似性;癌细胞属于非刚性特性物体,前后帧细胞有较大的形态变化;癌细胞在运动过程中存在较大的重叠遮挡。这些挑战大大降低了癌细胞生命周期分析的准确性,在一定程度上限制了生物及医学的发展。

近年来,深度学习已被应用于图像处理的多个领域<sup>[2-4]</sup>,从最初的图像识别与分类,发展到图像的检测与跟踪等领域,且取得了显著效果。基于深度学习的跟踪算法<sup>[2-3]</sup>也有所突破,特别是在有大量训练数据的前提下,深度学习这种监督学习方法能够超越传统方法,且效果显著。然而,当前基于深度学习的跟踪算法大多应用于单目标的跟踪,在细胞快速移动和细胞分裂的情况下,跟踪性能会明显下降,特别是当癌细胞受到抗癌药物的抑制时,生长状态发生不规则的变化,从而导致细胞的检测与跟踪更加复杂。为处理这些问题与挑战,本文提出一种基于卷积神经网络(Convolutional Neural Network, CNN)的癌细胞多目标跟踪方法,可以实现对显微图像序列中癌细胞的自动检测<sup>[5]</sup>与跟踪匹配,描述癌细胞的完整运动轨迹,进而通过研究癌细胞的行为状态来自动分析癌细胞的生命周期,对研制相应的抗癌药物具有重要意义。

## 2 相关工作

在生物及医学研究领域,细胞药效的研究及新药的开发生都离不开对细胞生命活动的观察,对细胞进行相应的检测和跟踪之后才能对细胞运动行为及形态变化进行研究。本文算法提高了细胞的检测与跟踪准确度,有利于提高生物学研究及制药水平。

在目标检测方面,传统检测方法的基本步骤为:先提取目标特征,再根据特征并利用分类器实现目标的识别。例如:Opelt等<sup>[6]</sup>提出基于Boosting方法实现目标的检测与识别,Ding等<sup>[7]</sup>提出基于HOG特征和SVM方法实现目标的检测。但HOG等传统特征难以全面表达目标特征,且Boosting等分类器受噪声的影响较大,这些因素都严重制约了检测算法的发展。随着深度学习在语音识别、机器翻译及计算机视觉等领域得到广泛使用,目前目标的分割、分类<sup>[8]</sup>及检测问题多考虑使用深度学习算法。相比于传统检测方法,深度学习可以大幅度提高目标检测的精度。基于深度学习的检测方法主要是利用CNN方法,具体可分为:基于Region Proposal建议区域进行分类(例如以Faster R-CNN<sup>[9]</sup>为代表的R-CNN<sup>[10]</sup>及Fast R-CNN<sup>[11]</sup>)和基于回归对目标位置进行拟合(例如YOLO<sup>[12]</sup>及SSD<sup>[13]</sup>)。本文利用基于建议区域的Faster R-CNN框架进行癌细胞的初检测。在Faster R-CNN中,作者提出了RPN网络来产生建议区域,相比于R-CNN中利用的Selective Search<sup>[14-15]</sup>方法,RPN网络可提高提取的建议区域的质量及检测准确度;同时RPN与分类网络Fast R-CNN共享卷积层的建议区域,从而使计算量大大降低,训练速度加快,可实现检测的实时性。

目标跟踪一直是计算机视觉领域的一个重要分支,在生物及医学领域,国内外众多学者在序列图像对目标的跟踪与分析方面展开了研究,提出的算法在各自的应用领域也各具优缺点。传统的目标跟踪算法如卡尔曼滤波<sup>[16]</sup>、粒子滤波<sup>[17]</sup>等结合了概率的方法,这类算法的优点是追踪器可使用上一帧的位置预测其在下一帧的位置,能够对单目标进行快速跟踪,但缺点是误差会传递,从而产生累计误差,同时跟踪过程中目标会出现较大幅度的形变且遮挡适应性较差,一旦目标丢失,追踪器将永久失效,因此该算法只适用于背景简单的跟踪。类似于TLD<sup>[18]</sup>这种集检测、学习、跟踪算法为一体的跟踪算法在目标丢失时,使用目标检测的方法重新在图像中检测出目标位置,然后重新初始化追踪器。但是该类算法在目标的检测与跟踪过程中依然利用传统特征对目标进行跟踪表达,同时多目标跟踪任务中,该类算法的搜索时间往往随目标数量的增多呈指数级增长,不能很好地解决多目标跟踪场景中新目标的出现、旧目标的消失以及目标的分裂或合并造成的目标数目变化等问题<sup>[19]</sup>。

癌细胞在运动过程中伴随较大幅度的形变或遮挡,同时数据集中黏连细胞的训练数据较少,即使利用Faster R-CNN网络进行检测也依旧存在一定的误差。为此,本文设计了基于阈值分割和扫描圆的再检测算法CSA<sup>[20]</sup>,以实现黏连细胞的检测优化,为之后在细胞黏连区域取得较好的跟踪效果奠定基础。同时,癌细胞在运动过程中往往存在新目标的出现及旧目标的消失,还可能伴随因细胞分裂造成目标数目变化<sup>[19]</sup>等问题,传统单目标跟踪算法难以实现实时跟踪,甚至会丢失目标,因此本文设计了高效的帧间目标匹配算法,同时基于深度学习检测框架,实现了多目标跟踪<sup>[21]</sup>,从而实现癌细胞高精度的检测与跟踪。

高效的帧间目标匹配算法<sup>[22]</sup>是基于多目标融合的思想,依据癌细胞的形态特点以及使用的研究数据情况,提出能综合描述癌细胞特征的特征描述子。本文为了提高黏连区域的细胞跟踪精度,提出了综合特征描述子,将从目标检测网络输出的卷积特征及位置信息进行加权融合,实现帧间癌细胞的高效匹配。同时在多个癌细胞进行匹配跟踪时,其采用半全局跟踪策略<sup>[23]</sup>,使用当前帧的前 $n$ 帧作为一个匹配窗口,既兼顾实时性的需求,又可以利用匹配窗口中的信息对可能出现的错误进行修正,提高了跟踪算法对可能出现的检测误差的鲁棒性。

## 3 方法

本实验以剑桥大学肿瘤研究所提供的一组膀胱癌T24相衬显微细胞图像序列作为研究对象。该数据集含有150帧相衬显微癌细胞的序列图像,共包含4000多个癌细胞目标。其中,对包含2600个癌细胞目标的前100帧显微细胞图像进行手工标注,只标注细胞的边界框来训练本文的Faster R-CNN检测器;对于后50帧,不但标注细胞边界框,还标注细胞的中心位置信息,作为跟踪的测试数据。

### 3.1 癌细胞的初检测

本文使用基于建议区域的Faster R-CNN框架进行癌细胞的初检测,Faster R-CNN的整体网络结构如图1所示。

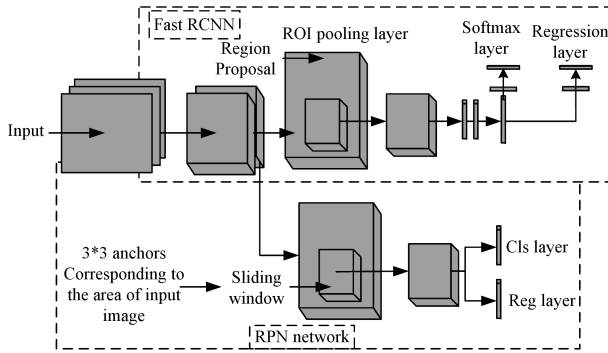


图 1 Faster R-CNN 的整体网络结构图  
Fig. 1 Overall network structure of Faster R-CNN

Faster R-CNN 提出一种新的建议区域提取方法(RPN),能解决提取建议区域开销大的问题。Faster R-CNN 为提取建议区域训练单独的深度卷积神经网络,可大大减少提取建议区域的时间消耗,基本解决了建议区域提取对目标检测系统效率的限制。在训练 Faster R-CNN 的过程中分别训练 RPN 和 Fast RCNN 网络,同时实现 RPN 与 Fast RCNN 网络共享卷积层,大大减少了网络的计算量。

首先训练一个用来提取建议区域的 RPN 网络,RPN 网络结构示意图如图 2 所示。

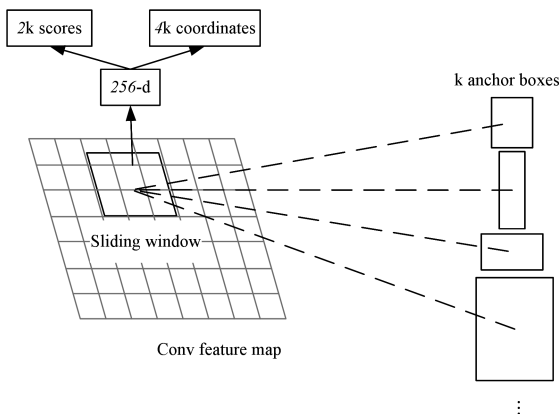


图 2 RPN 网络结构示意图  
Fig. 2 Structure diagram of RPN network

由图 2 可见,RPN 网络通过一个  $3 \times 3$  的滑动窗口(Sliding Windows, SW)与 ZF 卷积网络的最后一个卷积特征图连接,以最后的卷积特征图作为输入。为了使 RPN 有更好的尺度不变性,每一个滑动窗口对应原图上不同尺度的 9 个 anchors<sup>[9]</sup>(3 种大小,3 种长宽比),每一个 anchor 可以被映射到原图上一个固定大小的感受野,对应输入图像上的一个固定位置的候选区域,这种设计可以提高 RPN 网络对目标的尺寸变化和长宽比率变化的鲁棒性。训练 RPN 的正负样本根据 anchors 与目标标签的  $IoU$ (Intersection Over Union)是否大于阈值来选取,选取  $IoU$  大于设定阈值的 anchor 作为正样本, $IoU$  小于设定阈值的 anchor 作为负样本。在 RPN 网络中,首先利用  $3 \times 3$  的滑动窗口在卷积特征图上滑行,得到每个窗口对应的特征;再将每个特征映射成一个 256 维的向量;最后将这个向量作为 Cls layer 分支和 Reg layer 分支的输入。而在 Cls layer 分支上使用 Softmax<sup>[24]</sup> 分类器对候选 anchors 进行评估,得到候选 anchors 属于正样本和负样本的概率;

Reg layer 分支负责对建议区域坐标的偏移量进行回归,以优化预测的建议区域精确度。

同时,提取建议区域的 RPN 和对提取的建议区域进行分类的 Fast R-CNN 网络共用卷积层。提取建议区域之后,Fast R-CNN 网络判断提取的建议区域是否包含细胞的合适目标框。Fast R-CNN 网络的训练数据为 RPN 提取出的建议区域,将大于标签  $IoU$  阈值的建议区域作为正样本,小于阈值的作为负样本。如图 1 所示,将 RPN 网络产生的建议区域和卷积网络最后一层的卷积特征共同作为 Fast R-CNN 网络中 RoI pooling 层的输入,RoI pooling 层将每个建议区域映射位置对应的卷积特征映射成固定维度的特征作为下一层的输入。与 RPN 类似,Faster R-CNN 网络的末端同样连接两个分支,一个分类分支负责对建议区域是否为真实的目标进行分类,另一个回归分支对目标区域的坐标进行优化。

在实验中发现单独使用 Faster R-CNN 对黏连细胞进行检测时会出现错误,如图 3(c)所示。

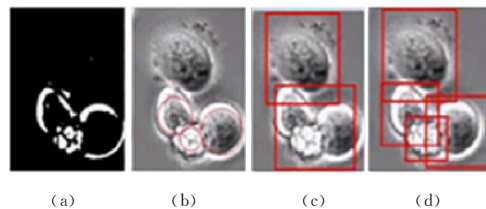


图 3 对黏连细胞区域再检测的效果图  
Fig. 3 Results of detecting adhesion cells

### 3.2 癌细胞的检测优化

对于图 3 中的错误检测,推测原因如下:一方面是粘连细胞训练数据较少,使得训练的深度检测器对细胞的粘连情况训练不充分;另一方面粘连细胞与本身细胞之间存在明显的重叠,加大了检测的难度。为此,针对粘连细胞区域,本文借鉴霍夫圆算法思想设计了扫描圆检测算法(Circle Scanning Algorithm, CSA),其伪代码如算法 1 所示。

#### 算法 1 CSA

1. Using radius  $r$  to extend binary image BW to EBW
2. While  $i \leq \text{row size of BW} + r$  do
3. Parameter of each row: count2=0;
4. While  $j \leq \text{column size of BW} + r$  do
5. Parameter of each column: count1=0;
6. for  $\Theta = 3, 6, \dots, 360$  do
7. if  $EBW(i, j, r, \Theta) > 0$  then
8. count1 = count1 + 1;
9. end
10. end
11. Parameter plane:  $BW_{para}(i-r, j-r) = \text{count1}$ ;
12. count2 = count2 + count1;  $j = j + 1$  or  $j + r$ ;
13. end
14.  $i = i + 1$  or  $i + r$ ;
15. end

CSA 算法首先根据细胞之间的距离来确定细胞粘连区域;再使用阈值分割<sup>[25]</sup>的方法得到粘连区域的二值图像(binary image)(见图 3(a));然后使用 CSA 算法对二值图像进行操作,得到一个类似于霍夫圆算法的参数平面( $BW_{para}$ ),同时在  $BW_{para}$  中寻找局部极大值点,将局部极大值点周围距

离小于阈值  $r$  的所有参数置零,重复这个操作,直到参数平面中的非零点只有局部极大值点;最后假设这些局部极大值点就是细胞的中心点,综合判断深度检测器和设计的 CSA 算法的结果。细胞检测操作流程如图 4 所示。

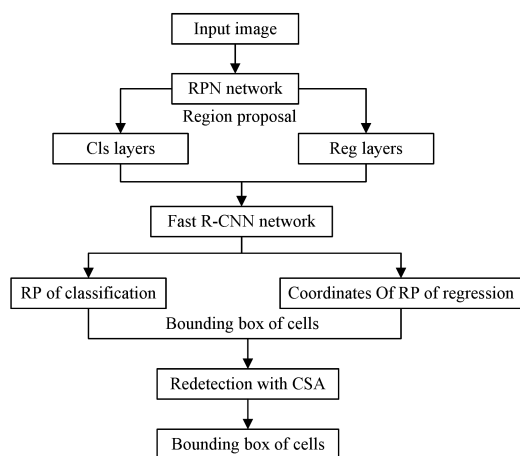


图 4 细胞检测操作流程

Fig. 4 Flowchart of cell detection operation

最终检测效果如图 3(d)所示,可以看出,对于图像中的分散细胞和黏连细胞,综合使用深度检测器和 CSA 再检测算法能获得优异的检测效果。

### 3.3 基于深度学习癌细胞的多目标跟踪

通过对癌细胞的跟踪以实现每个细胞轨迹的描述,是对每个细胞进行单独分析的基本需求,对实现自动化检测细胞形态、观测癌细胞在不同药物下所产生的反应、开发抗癌药物、帮助人类战胜癌症有着非常重要的意义。本文使用基于检测的方法实现对癌细胞的多目标跟踪。

传统特征描述方法,如颜色特征或纹理特征,都没有考虑到目标的空间位置信息,容易受到背景的干扰。本文在当前帧中选取固定数目的观测值与待匹配目标进行匹配,以保证跟踪系统的效率。同时本文设计了包括卷积特征在内的综合特征描述子<sup>[26]</sup>,使用 CNN 网络提取出每个细胞的卷积特征,相比于传统特征,深度神经网络所提取的特征同时包含目标的轮廓和颜色信息,并且对目标的微小形变更加鲁棒。除去卷积特征,本文还使用尺寸和位置特征综合描述目标,对 3 种特征的匹配代价进行加权累加。最后的代价函数如式(1)一式(4)所示。

$$\text{cost} = \text{para1} * \text{Conv} + \text{para2} * \text{Position} + \text{para3} * \text{Size} \quad (1)$$

$$\text{Conv} = \| FC \cdot f_{ced} \| \quad (2)$$

$$\text{Position} = \sqrt{(X - x_0)^2 + (Y - y_0)^2} \quad (3)$$

$$\text{Size} = (W - w_0) * (H - h_0) \quad (4)$$

代价函数中综合卷积特征代价为  $\text{Conv}$ ,位置特征代价为  $\text{Position}$ ,尺寸特征代价为  $\text{Size}$ 。其中, $FC$ 代表当前检测帧中所有检测到的卷积特征, $f_{ced}$ 表示匹配细胞的卷积特征,卷积特征代价即  $FC$ 和  $f_{ced}$ 的欧氏距离。 $X, Y$ 和  $x_0, y_0$ 分别代表当前检测帧中的细胞和匹配细胞的纵横坐标,位置特征代价即距离代价,这里直接使用两点间的空间距离。 $W, H$ 和  $w_0, h_0$ 分别代表当前检测帧中细胞和匹配细胞的长宽尺寸,尺寸特征代价即两者差值的乘积,表示细胞之间尺寸的差

异。由于相邻两帧同一个癌细胞的形态变化可能非常大,将观测帧的细胞与前一帧的卷积特征进行对比,然而细胞的尺寸相对比较稳定,因此为防止对某一个细胞尺寸检测失误而造成误匹配,使用半全局跟踪策略。半全局跟踪策略采用当前帧的前  $n$  帧作为一个匹配窗口,既兼顾实时性的需求,又可以利用匹配窗口中的信息对可能出现的错误进行修正。本文使用匹配帧的前 3 帧尺寸特征与观测帧计算匹配代价,从而提高综合特征描述的实时性、准确度及鲁棒性。

帧间目标匹配算法的伪代码如算法 2 所示,通过该算法可实现数据关联。

### 算法 2 多目标帧间匹配

1. Init trajectory, dist\_mat, min\_distN, fc\_ed, boxes, fc
2. For  $i=1; \text{length}(\text{trajectory})$
3. Cost\_dist\_mat = dist\_mat(min\_distN);
4. Cost\_fc\_mat =  $\| fc * fc\_ed \|$ ;
5. Cost\_size\_mat =  $(W - w_0) * (H - h_0)$ ;
6. end
7. normalize Cost\_dist\_mat, Cost\_fc\_mat, Cost\_size\_mat
8. cost\_mat = para1 \* Cost\_dist\_mat + para2 \* Cost\_fc\_mat + para3 \* Cost\_size\_mat
9. for  $i=1; \text{length}(\text{trajectory})$
10. if no matched trajectory(i)
11. continue
12. end
13. k = search(cost\_mat);
14. add cost\_mat(k) to trajectory
15. end
16. if(there is no matched cells in boxes)
17. cell split judge
18. end

算法 2 中,  $\text{trajectory}$  为细胞的轨迹;  $\text{dist\_mat}$  为待匹配细胞与检测细胞的距离;  $\text{min\_distN}$  为与待匹配细胞最近的  $N$  个检测细胞索引;  $\text{fc\_ed}$  为待匹配帧的细胞卷积特征;  $\text{boxes}$  为检测出的细胞边界框;  $\text{fc}$  为检测帧的卷积特征;  $\text{cost\_size\_mat}$  表示待匹配目标与观测值的尺度代价匹配矩阵;  $\text{cost\_dist\_mat}$  表示待匹配目标与观测值之间的位置差异匹配矩阵;  $\text{cost\_fc\_mat}$  表示待匹配目标与观测值之间的卷积特征的欧氏距离矩阵。相比普通细胞的分裂过程,癌细胞的分裂更为复杂,而且会有多种不同的分裂表现。当检测器检测到细胞数目增加时,依据上一帧对应最小距离位置的细胞形态是否处在分裂期来判断增加的细胞是否为新分裂出的细胞。本文先对待判细胞进行二值化处理,然后根据像素点的分布来判断细胞是否处于分裂期。

### 3.4 跟踪评价指标

追踪准确度(MOTA)用于统计跟踪中的误差累计情况,包括缺失数、误判数、误配数等。MOTA 的定义为:

$$\text{MOTA} = 1 - \frac{\sum_t (m_t + fp_t + mme_t)}{\sum_t g_t} \quad (5)$$

其中,  $m_t, fp_t$  及  $mme_t$  分别表示在  $t$  时刻丢失、误判以及误匹配的数量。

追踪精确度(MOPT)用于衡量目标位置确定的精确程度,其定义为:

$$MOPT = \frac{\sum_i d_i}{\sum_i c_i} \quad (6)$$

其中,每一帧中目标在检测之后得到的预测位置与目标实际位置之间的距离为  $d_i$ ,并用  $c_i$  表示第  $t$  帧的匹配个数。

缺失率(Loss Rate)表示序列中追踪丢失的比率,其定义为:

$$\bar{m} = \frac{\sum_i m_i}{\sum_i g_i} \quad (7)$$

误判率定义为:

$$f_p = \frac{\sum_i f p_i}{\sum_i g_i} \quad (8)$$

误匹配率定义为:

$$mme = \frac{\sum_i m m e_i}{\sum_i \sum_i g_i} \quad (9)$$

其中,每一帧中的缺失数用  $m_i$  表示,误判数用  $f p_i$  表示,误匹配数用  $m m e_i$  表示,并用  $g_i$  表示第  $t$  帧跟踪目标的个数。

## 4 实验结果与分析

### 4.1 癌细胞的检测实验结果

由于硬件的限制,本文没有采用诸如 VGG<sup>[27]</sup> 等特别大的网络结构,而是采用 ZF Net 网络<sup>[28]</sup> 作为实验中 RPN 和 Fast R-CNN 的基础网络结构。在训练深度卷积神经网络时,参数的设置会对结果产生明显的影响,本文通过大量实验选取了合适的参数。由于细胞的尺寸相对较小,为避免在卷积后进一步损失细胞的特征,本文使用没有经过缩放的原始尺寸(1200×1600)图片作为网络的输入。在训练 RPN 时,选取与标签交叉重叠率大于 0.75 的 anchors 作为正样本,交叉重叠率在 0~0.35 之间的 anchors 作为负样本,该交叉重叠率可以较好地筛选正负 anchors,有利于后续建议区域的正确分类。在训练 Fast R-CNN 时,选取与标签的交叉重叠率大于 0.6 的建议区域作为正样本,交叉重叠率在 0.1~0.5 之间的作为负样本。将正样本的交叉重叠率降低到 0.6 能保证更多有效的建议区域被保留下来,提高负样本的交叉重叠率能尽可能多地删除干扰样本。在测试阶段,使用非极大值抑制(NMS)<sup>[29]</sup> 的方法处理 RPN 产生的 2000 个区域,因为本数据集每帧图片的平均样本数量为 28,这里选取分数最高的 300 个作为生成的建议区域用作 Fast R-CNN 的输入。同时因为本数据集存在部分遮挡数据,本文选取阈值为 0.3 对 Fast R-CNN 的结果进行 NMS 操作,另外考虑到在测试阶段保留更多的候选区域,选取结果中置信度大于 0.6 的作为最终的细胞边界框。在深度学习检测器的检测效果上,针对容易出现检测错误的黏连细胞区域,采用设计的扫描圆检测算法对粘连达到 4 个的细胞区域的检测结果进行优化,以进一步提升整体的检测效果。在 Faster R-CNN 框架下,有无使用 CSA 检测优化方法的检测结果对比如表 1 所列。

采用准确率(precision)、召回率(recall)、平均精度 AP(average precision)和 AUC(area under roc curve)等多种指标来评测检测算法。

从表 1 可知,深度检测器对测试数据中的细胞获得了 0.974 的检测准确率,其他指标也都较好。加上 CSA 算法之后,检测准确率提升到 0.979,其他指标也都略有提升。其原

因一方面是黏连细胞本身在测试数据中较少;另一方面是评测算法中将检测框与标签  $IoU$  大于 0.5 的结果判断为检测正确,然而这样对较小的细胞(如图 3(c)所示)并不公平,导致加入 CSA 算法之后,检测性能看似并没有明显的提升。

表 1 Faster R-CNN 和 Faster R-CNN+CSA 训练结果的性能指标对比

Table 1 Performance comparison of detection results with Faster R-CNN and Faster R-CNN+CSA

Methods	Precision	Recall	AP	AUC
Faster R-CNN	0.974	0.961	0.907	0.922
Faster R-CNN+CSA	0.979	0.989	0.908	0.924

### 4.2 癌细胞的多目标跟踪与匹配结果

在跟踪实验过程中,将检测帧中距离目标最近的 5 个细胞与目标进行匹配,先根据距离排除每个目标最近的 5 个细胞中的异常细胞,再进行全局归一化。在计算细胞尺寸特征时,距离检测帧越近的帧其权重越大,选取前 3 帧检测帧的权重为[0.5 0.3 0.2],因为本文数据集的采样频率高,帧间细胞相对移动较缓慢,位置变化较小,所以使距离最近的检测目标的权重最大。这里使用欧氏距离来计算检测帧与待匹配帧的卷积特征差异。同时经过多次实验,将跟踪效果最好时的距离特征、卷积特征与尺寸特征的权值设置为[0.3 0.3 0.4]。

本文所使用的基于检测的跟踪算法中,细胞检测的结果直接影响目标跟踪的结果,因此对检测算法与跟踪算法都进行全面的评估。跟踪算法的量化评测结果如表 2 所列。

表 2 跟踪算法的量化评测对比

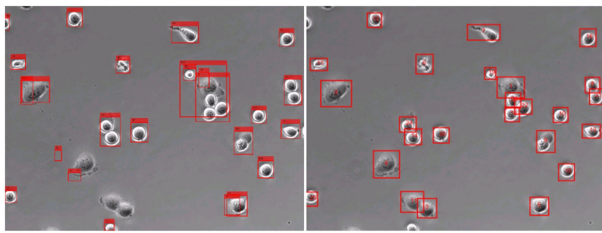
Table 2 Quantitative evaluation comparison of tracking algorithms

Methods	MOTA	MOTP	AUC	Loss Rate	FAR
DMOT	0.939	3.059	0.753	0.033	0.028
BMOT	0.853	5.800	0.701	0.044	0.054

由表 2 可见,不像单目标跟踪算法的评价那么统一,在多目标跟踪算法的评价中往往需要根据实际的应用场景设计具体的评价指标。本文参考 Wu 等<sup>[30]</sup> 和 Bernardin 等<sup>[31]</sup> 设计的评价方法,使用 MOTA(Multiple Object Tracking Accuracy)、MOTP(Multiple Object Tracking Precision)、AUC、缺失率(Loss Rate)和误判率(FAR)<sup>[32]</sup> 等多种指标来定量评测多目标跟踪算法的性能。

为验证本文所提出方法的先进性,将本文算法与基于深度学习的贝叶斯多目标跟踪算法(BMOT)<sup>[33]</sup> 和基于高斯混合模型的卡尔曼多目标跟踪算法(GMM+Kalman)<sup>[34]</sup> 的跟踪结果进行对比。对于 BMOT 跟踪算法,实验中使用相同的目标尺寸与位置特征描述细胞目标,用传统颜色直方图来代替本文方法中的卷积特征。由表 2 对比结果可知,基于卷积特征的跟踪算法性能更好。GMM+Kalman 跟踪算法是基于模型的目标跟踪算法,将该类算法应用于本数据集有三大难点:1)数据集本身存在遮挡,在 Kalman 滤波之后使用匈牙利匹配算法不能得到很好的跟踪效果;2)数据集采集的频率很高,帧间细胞变化过低,大多数细胞基本处于静止状态,不利于 Kalman 对下一帧细胞的位置预测;3)数据集不够多,增加了训练一个合适的高斯混合模型的难度,难以进行后续细胞的跟踪。GMM+Kalman 跟踪结果如图 5 所示,比较实验结果可

知,基于模型的多目标跟踪算法不适用于多目标且存在遮挡的情况,本文提出的基于检测的跟踪算法具有一定的优越性。

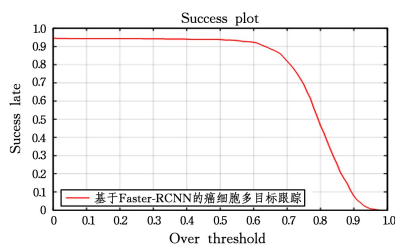


(a)GMM+Kalman 跟踪的结果 (b)DMOT 跟踪的结果

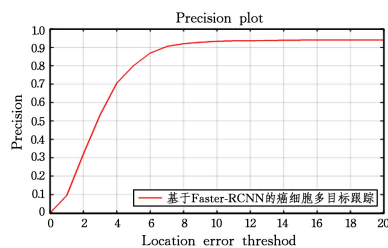
图 5 GMM+Kalman 和 DMOT 跟踪结果的对比

Fig. 5 Results comparison between GMM+Kalman tracking algorithm and DMOT

为了更全面且直观地描述跟踪性能,图 6 给出 IoU 阈值在 0~1 范围变化时成功率的变化,以及距离阈值在 0~20 范围内时跟踪精度的变化。从图 6 可以看出,跟踪算法在跟踪精度和准确度等各种指标上都有非常不错的表现。



(a)DMOT 的跟踪成功率



(b)DMOT 跟踪算法的精确度

图 6 DMOT 跟踪算法的跟踪成功率和跟踪精确度

Fig. 6 Success rate and accuracy of tracking using DMOT

本实验的检测可视化图如图 7 所示,跟踪效果如图 8 所示。图 7 为连续 4 帧的整体跟踪效果图,图 8 为黏连细胞区域连续 10 帧的跟踪效果。

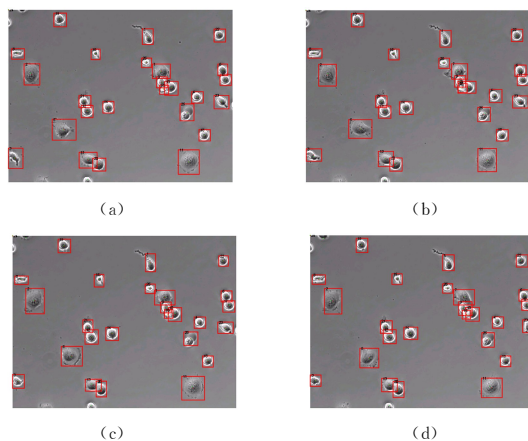


图 7 显微序列连续 4 帧的整体跟踪效果(电子版为彩图)

Fig. 7 Overall tracking effect with four consecutive frames of microscopic sequence

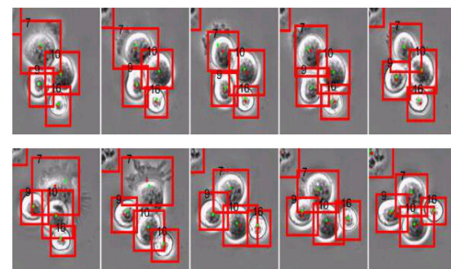


图 8 黏连细胞区域连续 10 帧的跟踪效果图(电子版为彩图)

Fig. 8 Tracking effect with 10 consecutive frames of adhesion cells

其中矩形框标识检测器输出的检测效果,绿色点为手工标注的细胞中心位置,红色点为算法预测的细胞的中心位置。由图 7 和图 8 可知,跟踪算法对大部分细胞都实现了准确且连续的跟踪。

结束语 本文将目前效果较好的深度学习检测算法 Faster R-CNN 应用到细胞的自动检测与跟踪中。首先采用 Faster R-CNN 对膀胱癌细胞进行初步检测,并提出 CSA 算法对细胞粘连区域进行进一步的检测优化,从而实现对癌细胞的有效检测。在对细胞准确检测的基础上,采用基于检测的多目标跟踪算法,设计包括卷积特征在内的癌细胞综合特征描述子,并设计匹配代价函数,同时实现高效的癌细胞多目标跟踪,对大部分细胞进行准确跟踪。

本文的主要贡献与创新在于:1)将深度学习与传统方法相结合,提出 CSA 细胞粘连区域检测算法,实现相位显微镜下序列图像的细胞检测;2)使用基于深度学习检测的方法处理多目标跟踪问题,取得了非常好的效果;3)设计包含卷积特征在内的综合特征描述子,使得对目标的描述更加准确和鲁棒;4)完成高效的目标匹配,实现相衬显微镜下癌细胞的多目标跟踪。

但本文方法不能实现目标的在线实时跟踪,未来可改进跟踪算法,将跟踪算法融入到检测模块,实现基于检测的目标在线实时跟踪。

致谢 衷心地感谢剑桥大学肿瘤研究所为本研究提供实验数据。

参考文献

[1] MÖLLER M,BURGER M,DIETERICH P,et al. A framework for automated cell tracking in phase contrast microscopic videos based on normal velocities[J]. Journal of Visual Communication & Image Representation,2012,25(2):396-409.
[2] LI H,LI Y,PORIKLI F. Robust Online Visual Tracking with a Single Convolutional Neural Network [M]//Computer Vision -- ACCV 2014. Springer International Publishing,2014:194-209.
[3] WANG N,YEUNG D Y. Learning a deep compact image representation for visual tracking [C]// International Conference on Neural Information Processing Systems. 2013:809-817.
[4] HU H G,LUO C,GUAN Q,et al. A fast online multivariable identification method for greenhouse environment control problems [J]. Neurocomputing,2018,312:63-73.
[5] GIRSHICK R,DONAHUE J,DARRELL T,et al. Region-Based

- Convolutional Networks for Accurate Object Detection and Segmentation [J]. *IEEE Trans Pattern Anal Mach Intell*, 2016, 38(1):142-158.
- [6] OPELT A, FUSSENEGGER M, PINZ A, et al. Weak Hypotheses and Boosting for Generic Object Detection and Recognition [C]// *European Conference on Computer Vision*. Springer Berlin Heidelberg, 2004:71-84.
- [7] DING S, LIU Z, LI C. AdaBoost learning for fabric defect detection based on HOG and SVM [C]// *International Conference on Multimedia Technology*. IEEE, 2011:2903-2906.
- [8] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. *Advances in Neural Information Processing Systems*, 2012, 25(2): 1097-1105.
- [9] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, 1: 91-99.
- [10] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C]// *IEEE Conference on Computer Vision and Pattern Recognition*. 2014:580-587.
- [11] GIRSHICK R. Fast r-cnn [C]// *Proceedings of the IEEE International Conference on Computer Vision*. 2015:1440-1448.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection [J]. *Computer Science*, 2016:779-788.
- [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector [C]// *European Conference on Computer Vision*. Springer International Publishing, 2016:21-37.
- [14] UIJLINGS J R R, SANDE K E A V D, GEVERS T, et al. Selective Search for Object Recognition [J]. *International Journal of Computer Vision*, 2013, 104(2): 154-171.
- [15] HU H G, GUAN Q, CHEN S Y, et al. Detection and Recognition for Life State of Cell Cancer Using Two-Stage Cascade CNNs [J]. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2017:1.
- [16] CHU H, WANG K. Target tracking based on mean shift and improved kalman filtering algorithm [C]// *IEEE International Conference on Automation and Logistics*. IEEE, 2009:808-812.
- [17] MA L, CHANG F L, QIAO Y Z. Target tracking based on mean shift algorithm and particle filtering algorithm [J]. *Pattern Recognition & Artificial Intelligence*, 2006, 19(6):787-793.
- [18] KALAL Z, MIKOLAJCZYK K, MATAS J. Tracking-learning-detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 34(7):1409-1422.
- [19] JIANG L H, GAN Z H, JIANG M. Review of Multi-Target Tracking [J]. *Computer Systems & Applications*, 2010, 19(12): 271-275.
- [20] LAURITSCH G, DENNERLEIN F. Method for reconstructing a CT image using an algorithm for a short-scan circle combined with various lines: US, US 7359477 B2 [P]. 2008.
- [21] TAKALA V, PIETIKAINEN M. Multi-Object Tracking Using Color, Texture and Motion [C]// *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. DBLP, 2007:1-7.
- [22] WANG L, WANG J, CHEN D, et al. An Algorithm of Target Tracking and Detecting Synchronously Based on Inter-Frame Dual Match [C]// *2009 Symposium on Photonics and Optoelectronics*. 2009:1-4.
- [23] CHEN L, ZHAO Z, YAN H. A Probabilistic Relaxation Labeling (PRL) Based Method for *C. elegans* Cell Tracking in Microscopic Image Sequences [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2016, 10(1):185-192.
- [24] CHOI W. Near-online multi-target tracking with aggregated local flow descriptor [C]// *Proceedings of the IEEE International Conference on Computer Vision*. 2015:3029-3037.
- [25] MEMISEVIC R, ZACH C, HINTON G, et al. Gated Softmax classification [C]// *International Conference on Neural Information Processing Systems*. Curran Associates Inc. 2010: 1603-1611.
- [26] TAN Y. Study on applied technology arithmetic of image threshold segmentation [J]. *Microcomputer Information*, 2007, 23(24):298-300.
- [27] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition [J]. *Computer Science*, arXiv:1409.1556.
- [28] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks [C]// *European Conference on Computer Vision*. Springer International Publishing, 2014:818-833.
- [29] LARKIN J, GUNTHER B, HOLIERHOEK H. NMS with multi-server change requests processing: US, US 8028052 B2 [P]. 2011.
- [30] WU Y, LIM J, YANG M H. Object Tracking Benchmark [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, 37(9):1834-1848.
- [31] BERNARDIN K, STIEFELHAGEN R. Evaluating multiple object tracking performance: the CLEAR MOT metrics [J]. *Eurasip Journal on Image & Video Processing*, 2008, 2008(1): 246309.
- [32] 张剑华, 刘儒瑜, 邹伟杰. 显微图像序列中癌细胞轨迹追踪与关联 [OL]. [http://www. paper. edu. cn/releasepaper/content/201704-582](http://www.paper.edu.cn/releasepaper/content/201704-582).
- [33] JU H Y, YANG M H, LIM J, et al. Bayesian Multi - object Tracking Using Motion Context from Multiple Objects [C]// *IEEE Winter Conference on Applications of Computer Vision*. IEEE Computer Society, 2015:33-40.
- [34] FAZLI S, POUR H M, BOUZARI H. Multiple object tracking using improved GMM-based motion segmentation [C]// *International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, Ecti-Con*. IEEE, 2009:1130-1133.