

DOI:10.16136/j.joel.2022.09.0754

适于多尺度宫颈癌细胞检测的改进算法

郑 雯¹, 张标标², 吴俊宏², 马仕强², 任 佳^{1*}

(1. 浙江理工大学 机械与自动控制学院, 浙江 杭州 310018; 2. 浙江远图互联科技股份有限公司, 浙江 杭州 310012)

摘要:深度学习技术因其强大的特征提取能力而被广泛应用于目标检测任务中。针对多尺度宫颈癌细胞的识别准确率不均衡、检测效率低等问题,本文提出一种基于 YOLO v3 模型的改进识别算法 mo-YOLO v3 (mini-object-YOLO v3)。选用 20 倍数字扫描仪下采集的宫颈细胞图像作为数据集,为提高算法的鲁棒性,引入对比度增强、灰度图、旋转和翻转等多种数据增强策略扩充数据集;模型以 Darknet53 网络结合注意力机制作为主干模块,针对宫颈癌细胞尺寸差异大的特点,提出一种多尺度特征融合算法来优化模型结构;针对小目标检测精度低的问题,提出一种改进的损失函数,采用相对位置信息的方法减弱物体框对检测结果的影响。测试结果表明,本文所提的 mo-YOLO v3 模型不仅在总体识别精度上有明显的优势,同时大大提高了小尺寸宫颈癌细胞的定位精度。该模型对宫颈癌细胞识别的准确率达到 90.42%,查准率达到 96.20%,查全率达到 93.77%,相似指数 ZSI 为 94.97%,高于同类算法。

关键词:宫颈癌细胞检测;深度学习;YOLO v3 网络;多尺度特征融合;注意力机制

中图分类号: TP183 **文献标识码:** A **文章编号:** 1005-0086(2022)09-0948-11

Improved algorithm of multi-scale cervical cancer cells detection

ZHENG Wen¹, ZHANG Biaobiao², WU Junhong², MA Shiqiang², REN Jia^{1*}

(1. School of Mechanical Engineering and Automation, Zhejiang Sci-Tech University, Hangzhou, Zhejiang 310018, China; 2. Zhejiang Yuantu Internet Technology Co., Ltd., Hangzhou, Zhejiang 310012, China)

Abstract: Deep learning technology is widely used in target detection tasks because of its powerful feature extraction capabilities. Aiming at the problems of uneven recognition accuracy and low detection efficiency of multi-scale cervical cancer cells, this paper proposes an improved recognition algorithm, mini-object-YOLO v3 (mo-YOLO v3) based on the YOLO v3 model. The cervical cell images collected under a 20× digital scanner are selected as the data set. In order to improve the robustness of the algorithm, multiple data enhancement strategies such as contrast enhancement, grayscale image, rotation and flipping are introduced to expand the data set; the model takes Darknet53 network combined with attention mechanism as the backbone module, for the large difference in the size of cervical cancer cells, a multi-scale feature fusion algorithm is proposed to optimize the model structure. In order to solve the problem of low detection accuracy of small targets, an improved loss function is proposed, adopting the relative position information method to reduce the influence of the object frame on the detection result. The test results show that the mo-YOLO v3 model proposed in this paper not only has obvious advantages in overall recognition accuracy, but also greatly improves the positioning accuracy of small-size cervical cancer cells. The model has an accuracy rate of 90.42% for identification of cervical cancer cells, a precision rate of 96.20%, a recall rate of 93.77%, and a similarity index ZSI of 94.97%, which is higher than similar algorithms.

Key words: cervical cancer cells detection; deep learning; YOLO v3; multi-scale feature fusion; attention mechanism

* E-mail: jren@zstu.edu.cn

收稿日期: 2021-11-04 修订日期: 2021-12-15

基金项目: 浙江省公益技术研究项目(LGG20F030007)资助项目

1 引言

宫颈癌是女性高发癌症,若早期发现,其治愈率可接近100%,到晚期,则治愈率只有20%—50%。因此,早期筛查对宫颈癌的防范和治疗非常重要,可以降低其发病率和死亡率^[1]。然而,通过人工筛查找出癌变细胞是一项既繁琐又耗时的工作,随着计算机技术和人工智能的发展,基于计算机视觉的癌细胞自动检测逐渐成为医学图像处理中的一个研究热点。目前深度学习在众多领域都取得了很大的进展,如人脸识别^[2]、指纹对比^[3]、路牌检测^[4]以及无人驾驶等^[5],其优势在于通过构建多隐含层的机器学习模型和海量训练数据来自动提取有效特征^[6],从而提高分类或预测的准确率。这些特性使得基于深度学习进行宫颈癌细胞的自动检测成为可能。

近年来,许多研究者将深度学习算法引入到宫颈癌细胞的分割检测任务中:XIANG等^[7]基于YOLO v3模型层叠了一个特定于任务的附加分类器,通过研究数据集中存在的噪声,提出了一种通过平滑其分布来减弱噪声影响的算法,测试过程中,将子宫颈细胞检测器输出的图像作为级联分类器的输入,以实现更高的准确性;HARANGI等人^[8]使用FCN-8(full convolutional network-8)和FCN-16的组合模型,计算验证集中每个图像上所有像素之间的差异,合并不同分割方法的输出以减少输入噪声的影响,为了进一步提高分割的最终精度,作者还实现了传统的基于超像素的细胞分割算法;HYEON等^[9]设计并训练了卷积神经网络VGGNet(visual geometry group network)和一些机器学习分类器,从显微图像中对宫颈细胞的正常/异常状态进行自动分类,通过将原始单元图像分成多块 224×224 尺寸从而减轻分辨率差异,并计算图像特征向量的平均值,获得代表细胞图像的特征向量;SOMPAWONG等^[10]采用Mask R-CNN(mask region-based convolutional neural network)开发了一种算法,找到细胞核的质心并填充细胞质的某些部分以便裁剪重叠的单元格,通过分析裁剪后的图像来确定单元格类,并通过在特征图上滑动小型网络生成区域建议,其中每个滑动窗口都将映射到较低维的特征。

综上所述,现有算法为宫颈癌细胞的自动检测提供了多种可借鉴的实现思路,但依然存在以下问题有待研究:已有算法研究成果大多集中在对癌变细胞的分割上,而忽略了对癌细胞位置信息的研究;针对宫颈癌细胞分布状态多样且尺寸差异明显的现象,多数算法对于小目标的特征提

取效果有限。基于此,本文首先在数据预处理部分,引入数据增强技术扩充数据集,将样本做颜色和角度的转变以提高模型的鲁棒性,同时提出一种适用于多尺寸宫颈癌细胞检测的深度神经网络模型mo-YOLO v3(mini-object-YOLO v3),主要改进内容如下:

1) 基于主干网络Darknet53模块,引入CBAM(convolutional block attention module)注意力机制,以改善权重分布并辅助网络更有效地专注于目标的位置信息进行检测和识别;

2) 结合宫颈癌细胞尺寸差异大的特点,为防止小尺寸目标被遗漏,提出多尺度特征融合算法,并将拼接前后的特征图进行融合,大大提高了模型预测的置信度;

3) 改进了YOLO v3模型中的损失函数,利用相对误差减弱小目标预测框对预测结果的影响,提高了宫颈癌细胞位置的定位精度。

2 数据集与预处理

2.1 数据集介绍

宫颈癌中常见的是鳞状上皮细胞癌,其癌变细胞有如下特征:细胞畸形,有纤维形、蝌蚪形、菱形、不规则形;核畸形,细胞核变大,有凹形、长芽、分叶状、桑椹状或弯月形等;群体之间细胞及核大小相差悬殊,排列紊乱,染色深浅不均。实例图片如图1所示。

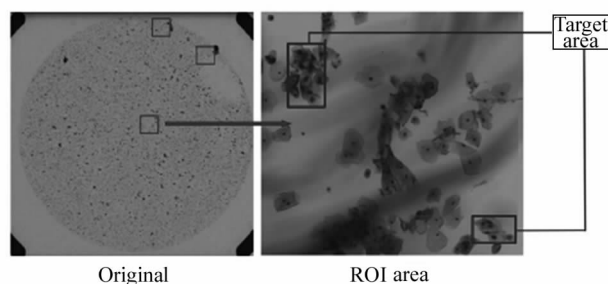


图1 数据集样本示例

Fig. 1 Sample example of dataset

本文数据集来自天池竞赛(<https://tianchi.aliyun.com>)提供的医学图像数据集,共800张宫颈癌细胞图片,其中阳性(pos)500张,阴性(neg)300张,阳性图片提供了多个ROI(region of interest)区域,且在ROI区域里面标明异常鳞状上皮细胞位置,阴性图片不包含异常鳞状上皮细胞,无标注。数据集为kfb格式,需要使用大赛指定SDK读取,实例图片如图1所示。每张数据在20倍数字扫描仪下获取,大小约300—400 M,像素实际物理间隔 $0.5 \mu\text{m}$ 。

2.2 数据增强

数据增强不仅可以增加数据集的规模,而且可以提高模型的鲁棒性,避免过拟合。本文以阳性图片提供的 1 212 个 ROI 区域作为网络输入,采用对比度增强、灰度图、旋转和翻转的增强方法^[11]对宫颈癌

数据集进行了扩充,数据集由原来的 1 212 张扩充到 18 180 张。

其中对比度增强以及灰度图的操作是对图片色彩进行改变,而经过旋转和翻转的操作后,图片的标注也会随着图片的转变而变化。文中使用的数据增强手段及效果示例如图 2 所示。

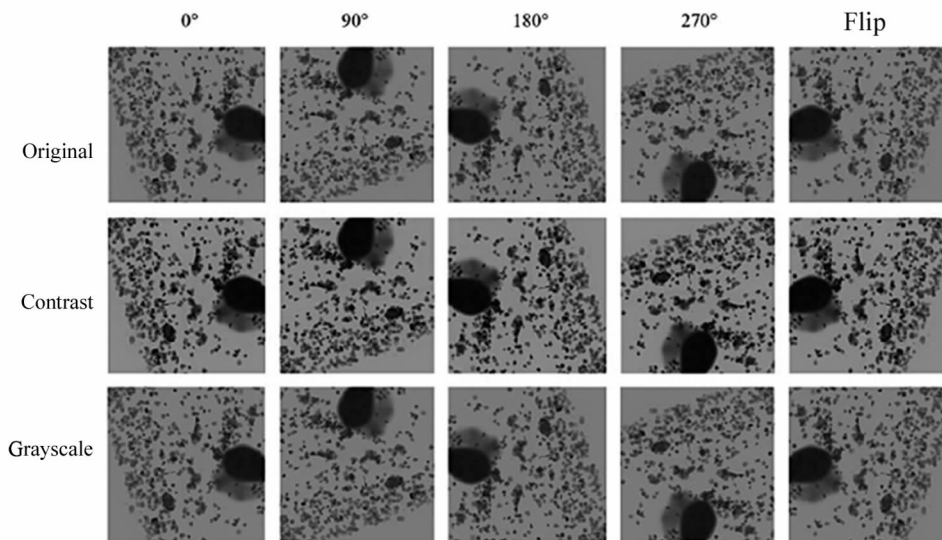


图 2 数据增强手段及效果示例

Fig. 2 Examples of data enhancement methods and effects

3 mo-YOLO v3 网络算法及实现

YOLO 模型是近年来在目标检测领域备受关注和好评的一种深度学习模型,本文采用 YOLO v3^[12]为基线模型,与 YOLO v2^[13]相比,它在 backbone、损失函数、先验框以及预测框等方面均进行了改进,对目标检测任务的效果有明显改善,但对于小目标的识别检测能力仍有待提升。针对多尺度宫颈癌细胞检测任务,本研究提出了一种基于 YOLO v3 的改进模型算法 mo-YOLO v3,其结构如图 3 所示。

基线模型中的主体网络 Darknet53 借用了 Res-Net 的思想,在网络中加入了残差模块,图 3 蓝框中所标注的 1,2,8,8,4 为残差模块的个数,整个 v3 结构里面,没有池化层和全连接层,网络的下采样通过卷积实现,而每个卷积层的实现包含“卷积+BN+Leaky_ReLU”。本文模型的主干网络 A-Darknet53 在最后两层残差模块前引入了 CBAM 注意力模块,该机制可以寻找特征图中关键的信息使网络关注更有价值的特征。将 416×416 尺寸的图片输入 A-Darknet53 后将最后 3 层输出的 52×52 , 26×26 , 13×13 尺寸的特征图进行堆叠。

针对宫颈癌细胞大小不一且分布杂乱无章等特点,mo-YOLO v3 模型将模型中堆叠前和堆叠后的特征图进行了融合(详细结构可见图 7),其中图 3 右侧模块与图 7 中模块颜色及功能一一对应,改进后的模型对小目标检测效果大大提升。此外,针对小目标定位不精准的现象,本文对损失函数进行了改进,利用相对误差改进模型的损失函数将位置参数的误差与实际值的比值做平方和运算,改进后的损失函数对宫颈癌细胞的定位有明显改善。

3.1 注意力机制的引入

深度学习中的注意力机制从本质上讲和人类的选择性视觉注意力机制类似,核心目标是从众多信息中选择出对当前任务目标更关键的信息。目前,注意力模型(attention model, AM)被广泛使用在图像识别、语义分割、目标检测等各种场景中,其最主要的作用就是从卷积层提取到的各种特征中找出更为关键的特征信息,从而辅助网络对目标的检测。

YOLO v3 网络由一系列卷积层、非线性层和下采样层构成,整个网络结构中没有池化层和全连接层,网络的下采样是通过设置卷积的 stride 为 2 来实现的。作为网络核心的卷积核是局部感受野上空间

(spatial)信息和特征维度(channel-wise)信息的聚合体,它能够从全局感受野上去捕获图像的特征来进行描述。因此,为改善权重分布并使网络更有效地专注于目标位置,通过实验对比注意力机制嵌入在

各层中的效果,本文在 A-Darknet53 的 Layer3 和 Layer4 之后分别引入 CBAM 注意力机制^[14]可以更好地分析模型提取到的特征,同时模型的检测速度基本保持原本的效率。其结构如图 4 所示^[15]。

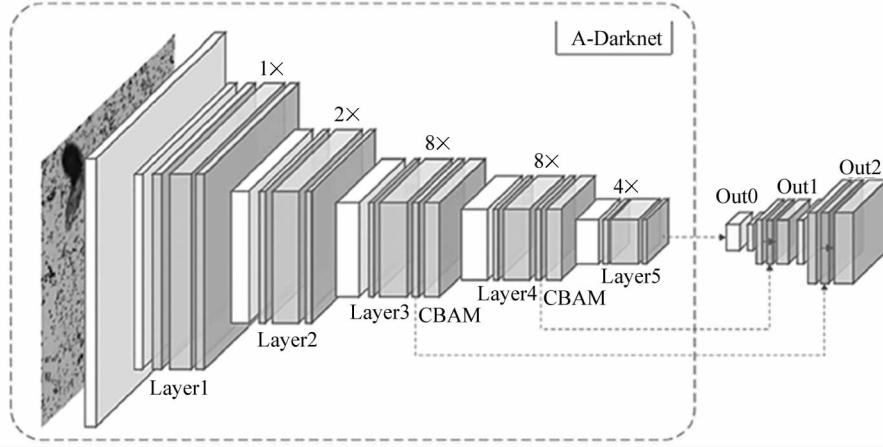


图3 本文所提 mo-YOLOv3 网络模型架构

Fig.3 The structure of mo-YOLO v3 network model

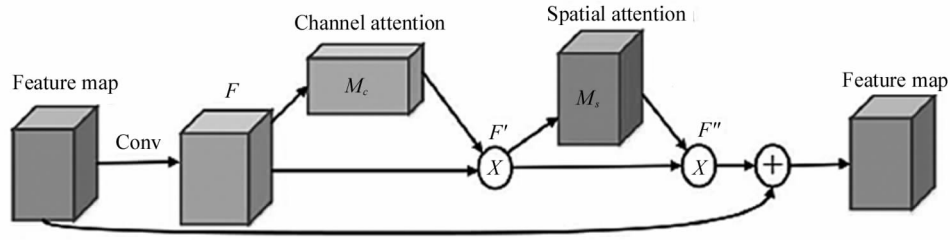


图4 CBAM 注意力机制结构图

Fig.4 Structure of CBAM attention mechanism

CBAM 注意力机制从通道和空间两个维度计算特征图的注意力权重,然后将获得的权重与输入的特征图依次相乘来进行特征的自适应学习,运算过程可见式(1),其中, F 表示特征图, $\otimes$ 表示逐元素相乘。实验表明,将通道注意力模块和空间注意力模块串行放置且通道模块在前,优于其他组合效果,相比于只关注通道的注意力机制 SENet,可以取得更好的效果。

$$\begin{aligned} F' &= M_c(F) \otimes F, \\ F'' &= M_s(F') \otimes F'. \end{aligned} \quad (1)$$

通道注意力模块结构如图 5 所示,输入是一个 $H \times W \times C$ 的特征图 F ,最大池化编码了最显著的部分,而平均池化编码了全局的统计信息,为有效计算通道注意力,该模块同时使用平均池化和最大池化得到两个 $1 \times 1 \times C$ 的通道描述,再将它们分别送入一个两层的神经网络,第一层神经元个数为 C/r ,第二层神经元个数为 C ,其中 r 为常数,本文 r 设置为

16。最后,将得到的两个特征相加后经过一个 Leaky_ReLU 激活函数^[16]得到权重系数 M_c ,计算过程可见式(2)。其中 $W_0 \in \mathbb{R}^{C/r \times C}$, $W_1 \in \mathbb{R}^{C \times C/r}$, σ 表示 LeakyReLU 函数。

$$\begin{aligned} M_c(F) &= \sigma(\text{MLP}(\text{AvgPool}(F)) + \\ &\quad \text{MLP}(\text{MaxPool}(F))) = \\ &\quad \sigma(W_1(W_0(F_{\text{avg}}^c)) + W_1(W_0(F_{\text{max}}^c))). \end{aligned} \quad (2)$$

CBAM 模块将通道注意力权重系数 M_c 和原来的特征 F 相乘即可得到缩放后的新特征,即 $F' = M_c(F) \otimes F$,生成空间 AM 需要的输入特征 F' 。与通道注意力相似,空间 AM 结构如图 6 所示。输入为上一模块所得大小为 $H \times W \times C$ 的特征图 F' ,为计算空间注意力,该模块沿着通道轴应用平均池化和最大池化操作得到两个 $H \times W \times 1$ 的通道描述图,再将它们通过 concatenation 操作合并特征通道生成一个有效的特征层。通过一个 7×7 的卷积层和 LeakyReLU 激活函数生成大小为 $H \times W \times 1$ 的空间

注意力图 M_s , 该空间注意力图编码了需要关注或抑制的位置, 计算过程见式(3):

$$M_s(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) = \sigma(f^{7 \times 7}([F_{\text{avg}}^s; F_{\text{max}}^s])), \quad (3)$$

式中, F_{avg}^s 大小为 $1 \times H \times W$, F_{max}^s 大小为 $1 \times H \times W$, $f^{7 \times 7}$ 表示滤波 (avg) 器大小为 7×7 的卷积运算, σ 表示 LeakyReLU 函数。将权重系数 M_s 和特征 F' 相乘即可得到缩放后的新特征, 即 $F'' = M_s(F') \otimes F'$ 。

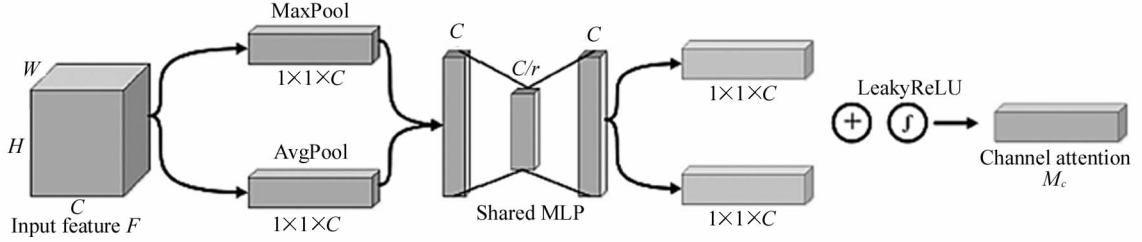


图5 通道注意力模块结构

Fig. 5 Structure of the channel attention module

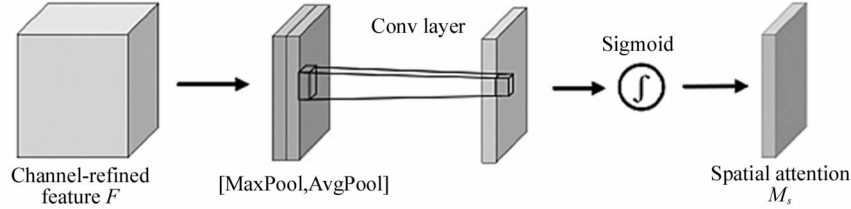


图6 空间注意力模块结构

Fig. 6 The structure of the spatial attention module

3.2 多尺度特征融合算法

在目标检测任务中, 特征图的分辨率直接反映了构成目标的像素数量, 像素数量越多, 对目标的细节表现就越具体。一般情况下, 在提取目标的特征图时仅需要提供必要的像素信息, 但对小目标而言, 需要结合大尺度特征图的像素信息才能更好地提取特征。

YOLO v3 采用 3 个不同尺度的特征图来进行对象检测, 13×13 特征图的感受野比较大, 适合检测大

尺寸对象, 26×26 适合检测中等尺寸对象, 52×52 则适合检测小尺寸对象。YOLO v3 模型通过 concatenation 操作^[17]将 A-Darknet 模块的中间层和对应下一层的上采样特征图进行通道拼接, 再将得到的特征图进行卷积操作以提取高层特征, 从而进行多尺度的特征图检测^[18]。

为进一步提高小目标特征检测的能力, 本文将拼接前后的特征图进行 element-wise summation 操作, 如图 7 所示, mo-YOLO v3 模型在 v3 的基础上

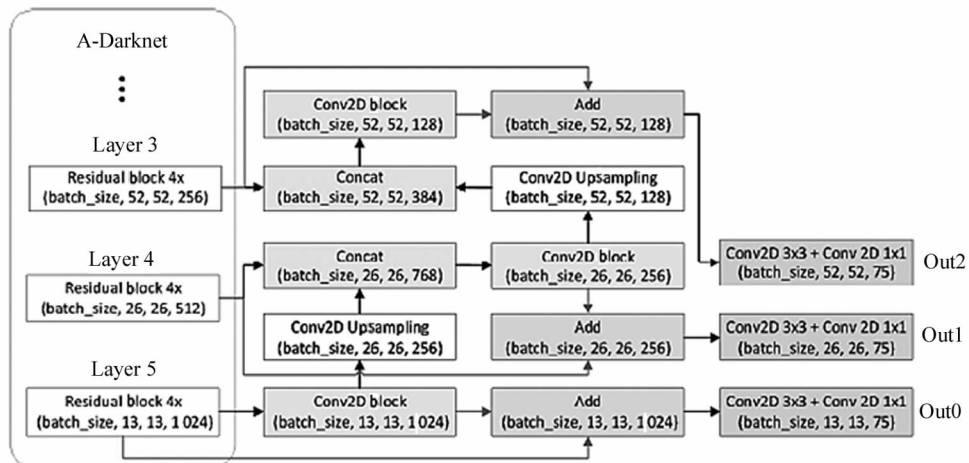


图7 多尺度特征融合算法结构

Fig. 7 Structure of multi-scale feature fusion algorithm

添加了蓝色连线区域以及 add 模块,即在高层卷积之后分别将其与 A-Darknet 的输出层对应的元素相加,所得输出与图 3 中的 out0、out1、out2 分别对应,综合考虑到信息丢失的可能性,提高了模型对目标检测的精度。同时,mo-YOLO v3 模型延续了 v3 中使用 K-means 聚类^[19]得到先验框尺寸的方法,为每种下采样尺度设定 3 种先验框,共聚类出 9 种尺寸的先验框。本文针对宫颈癌数据集设置的 9 个先验框尺寸如下:\$(7 \times 11)\$, \$(12 \times 20)\$, \$(16 \times 25)\$, \$(30 \times 62)\$, \$(47 \times 64)\$, \$(60 \times 95)\$, \$(95 \times 120)\$, \$(145 \times 208)\$, \$(325\$

\$\times 373)\$。

3.3 损失函数改进

损失函数是设计神经网络中的一个关键因素,其反映了模型对数据的拟合程度。YOLO v3 中采用的损失函数见式(5),包括中心坐标误差 ①,宽高坐标误差 ②,置信度误差(包括含目标 ③ 和不含目标 ④)以及分类误差 ⑤。其中,交叉熵函数形式为:\$C = - \sum y_i \ln a_i\$,将其求导可得:

$$\dot{C} = -y_i + a_i \sum y_j. \quad (4)$$

$$\text{loss(object)} = \lambda_{\text{coord}} \sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{\text{obj}} (2 - w_i \times h_i) [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \quad ①$$

$$\lambda_{\text{coord}} \sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{\text{obj}} (2 - w_i \times h_i) [(w_i - \hat{w}_i)^2 + (h_i - \hat{h}_i)^2] - \quad ②$$

$$\sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{\text{obj}} [\hat{C}_i \log(C_i) + (1 - \hat{C}_i) \log(1 - C_i)] - \quad ③$$

$$\lambda_{\text{noobj}} \sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{\text{noobj}} [\hat{C}_i) \log(C_i) + (1 - \hat{C}_i) \log(1 - C_i)] - \quad ④$$

$$\sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{\text{obj}} \sum_{c \in \text{classes}}^C [\hat{p}_i(c) \log(p_i(c)) + (1 - \hat{p}_i(c)) \log(1 - p_i(c))], \quad ⑤$$

$$\lambda_{\text{coord}} \sum_{i=0}^{K \times K} \sum_{j=0}^M I_{ij}^{\text{obj}} (2 - w_i \times h_i) \left[\left(\frac{w_i - \hat{w}_i}{\hat{w}_i} \right)^2 + \left(\frac{h_i - \hat{h}_i}{\hat{h}_i} \right)^2 \right], \quad (6)$$

式中,\$I_{ij}^{\text{obj}}\$为判断第\$i\$个网格中的第\$j\$个物体预测框是否包含该目标;\$x_i\$和\$y_i\$、\$\hat{x}_i\$和\$\hat{y}_i\$分别表示预测的物体框和真实区域的左上角对于第\$i\$个网格中心在\$x\$、\$y\$方向上的偏移量;\$w_i\$和\$h_i\$、\$\hat{w}_i\$和\$\hat{h}_i\$分别表示第\$i\$个网格中预测的物体框和真实区域的宽与高;\$C_i\$和\$\hat{C}_i\$分别表示第\$i\$个网格是否包含物体与物体框的置信度分数;\$p_i(c)\$和\$\hat{p}_i(c)\$分别表示第\$i\$个网格中物体预测的类别和实际类别的概率分数;\$\lambda_{\text{coord}}\$为物体位置参数\$x\$、\$y\$、\$w\$、\$h\$4个参数的损失权重;\$\lambda_{\text{noobj}}\$为不包含物体的网格的置信度误差的损失权重。

训练过程中,不同尺寸的目标预测出的物体框的大小是不同的,显然,对于相同的宽度和高度误差,相比于大物体框,小物体框的预测偏差对检测结果影响更大。原始模型中,对\$w\$和\$h\$采用计算误差平方和的方法,但是这种方法对定位精度的提升效果有限。本文考虑既要尽量消除位置误差的影响,又要保证得到有效的损失函数反馈,因此在损失函数中针对宽高坐标误差部分,即式②,采用了式(6)的形式。分别将位置参数\$w\$和\$h\$的误差与实际值的比值进行平方和计算,利用相对误差减弱物体框对预测带来的影响,从而提高了宫颈癌细胞位置的定位精度。

4 实验结果与分析

本文实验环境为 Windows10 操作系统,服务器 CPU i7 9700K,显卡 GPU GTX 1080 Ti,运行内存 16 G,编程环境为 Python,深度学习框架为 Pytorch。本文采用 Adam 作为 mo-YOLO v3 模型的优化算法,使用等间隔调整学习率 StepLR,初始学习率设为 \$1 \times 10^{-3}\$,参数 beta1=0.9, beta2=0.999,梯度 gamma=0.95, batch_size 设置为 64,模型初始值为 YOLO v3 在 ImageNet 上预训练后的参数。

4.1 评价标准

本文采用准确率(Accuracy)、查准率(Precision)、查全率(Recall)、相似指数 ZSI 以及精度检测指标 mAP (mean average precision) 来评价算法性能,其公式分别见式(7)~(10)。

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \quad (7)$$

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (8)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (9)$$

$$\text{ZSI} = \frac{2TP}{2TP + FP + FN}, \quad (10)$$

式中,变量 \$TP\$ 表示正确预测正样本数量,\$FP\$ 表示

将正样本错误判断为负样本数量, TN 表示正确预测负样本数量, FN 表示将负样本错误判断为正样本数量。其中 ZSI 考虑了 TP 、 FP 和 FN 各个方面, 如果 ZSI 分数高于 0.7, 则表明分割结果与标签高度匹配; mAP 表示均值平均精度, 作为目标检测中衡量检测精度的指标, 以 $Precision$ 和 $Recall$ 为坐标下的面积即为 mAP , 如图 8 所示。

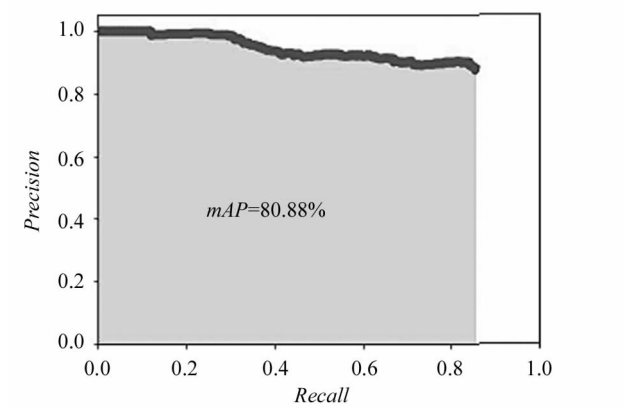


图 8 P-R 曲线

Fig. 8 P-R curve

4.2 实验结果

文中实验将验证集中的 ROI 区域图片放入 mo-YOLO v3 模型进行测试, 根据数据集提供的标注共存在 594 个异常鳞状上皮细胞(簇), 本文算法

性能如表 1 所示, 在 594 个样本中能够正确识别出 557 个阳性细胞。根据上述公式计算, mo-YOLO v3 模型算法准确率达到 90.42%, 查准率达到 96.20%, 查全率达到 93.77%, 相似指数 ZSI 为 94.97%。

表 1 mo-YOLO v3 模型算法性能

Tab. 1 Algorithm performance of mo-YOLO v3 model

Test picture	Diseased sample	Check out
106	594	579
Correct detection	False detection	Missed
557	22	37

本文算法 mo-YOLO v3 模型对癌细胞的检测效果如图 9 所示, 所测均为验证集中部分图片, 其中第一排为数据集提供的 label 真实标注图, 第二排为本文模型的检测结果标注图。分别进行纵向对比可以发现: mo-YOLO v3 模型对不同尺寸的异常鳞状上皮细胞均能实现准确识别及标注, 即使对于较难检测的小尺寸癌变细胞, 本文所提模型依然能够精确地检测出来。

由于检测图片均在 20 倍数字扫描仪下获取, 图 10 给出了模型改进前后检测效果对比的放大图。其中, 图(a1)、(a2)均为数据集提供的病变细胞真实标注图, 其对应图(b1)、(c1)、(b2)、(c2)分别为 YOLO v3模型与mo-YOLOv3模型对图(a1)、(a2)所框区域

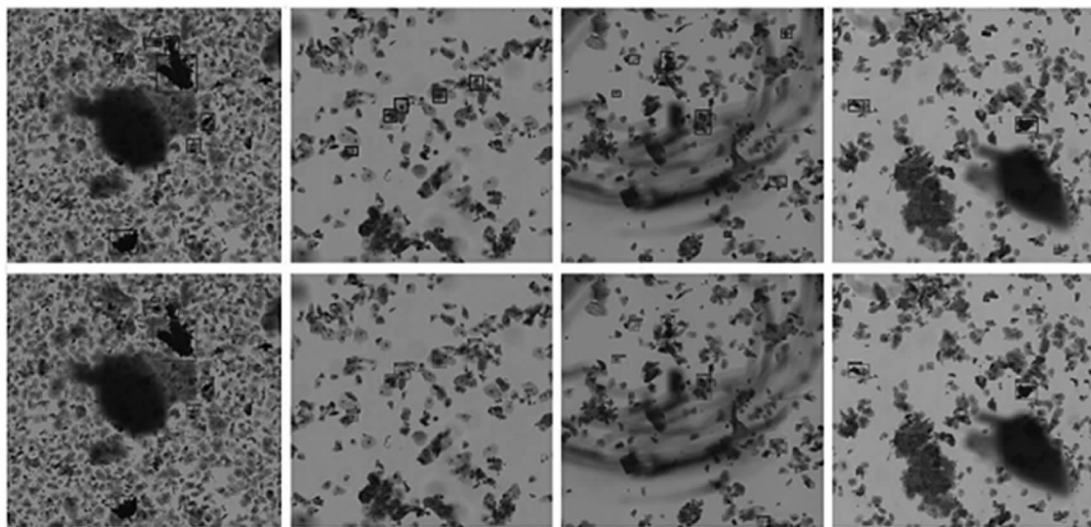


图 9 Mo-YOLO v3 模型测试结果

Fig. 9 Test results of mo-YOLOv3 model

的检测效果, 检测所得图像中不仅标注了被模型判定为异常鳞状上皮细胞的位置, 同时给出了相应的置信度。

从图 10(b1)、(c1)中上方图像的检测结果对比可见, 本文提出的 mo-YOLO v3 模型可以将 YOLO v3 模型遗漏的两个约 60×60 pixel 的小尺寸目标准

确地识别出来,同时模型所标出的位置信息与 label 信息完全一致;而图(b2)、(c2)的检测结果同样证实了这一点,本文模型同样检测出了 YOLO v3 模型遗漏的目标,并且模型的预测置信度也远高于 YOLO v3 模型。因此,mo-YOLO v3 模型能够以更高的准确率和更优的置信度识别出不同尺寸的病变细胞。

4.3 与其他算法的对比实验结果

本文将多种主流模型与本文所提算法在相同数据、相同实验环境下进行了对比测试,包括 R-CNN、Faster-R-CNN、U-Net、YOLO v4,实验结果见表 2。

表 2 不同网络模型在宫颈癌数据集上的对比结果

Tab. 2 Comparison results of different models on the cervical cancer dataset

Methods	Accuracy	Precision	Recall
R-CNN	75.43%	91.65%	80.99%
Faster-R-CNN	75.65%	92.21%	80.81%
YOLO v4	77.05%	93.81%	81.18%
U-Net	79.89%	94.77%	83.58%
YOLO v3	83.30%	95.71%	86.53%

由表 2 的数据分析可知:在主流模型中,针对宫颈癌数据集的检测,YOLO v3 模型效果最佳,因此本文采用 YOLO v3 作为基线模型。

本文算法在 YOLO v3 模型基础上,针对多尺度宫颈癌检测特点,引入 3 个模块对算法进行改进:引入了注意力机制(YOLO v3 & CBAM),引入了特征融合(YOLO v3 & feature_fusion),改进了损失函数(YOLO v3 & improved_loss_fun)。下面通过消融实验来分析各个模块对算法整体性能的贡献度 C_{index} ,其计算式为:

$$C_{\text{index}} = \frac{I_{\text{proposed}} - I_{\text{proposed-1}}}{I_{\text{proposed-1}}} \times 100\%, \quad (11)$$

式中, I_{proposed} 为使用该模块改进模型后的各项指标, $I_{\text{proposed-1}}$ 为基线模型 YOLO v3 的各项指标,计算指标的增长率 C_{index} 作为对应指标下该模块的贡献率,不同指标给模型带来不同程度的优化。

实验结果分别见表 3 及图 11。从表 3 可以明显看出:mo-YOLO v3 比改进前的 YOLO v3 模型的准确率整体提高了 7.12%,同时从图 11 可以看到,每一项指标下贡献度最高的组件都是特征融合模块,因此特征融合是对模型优化最具优势的模块。

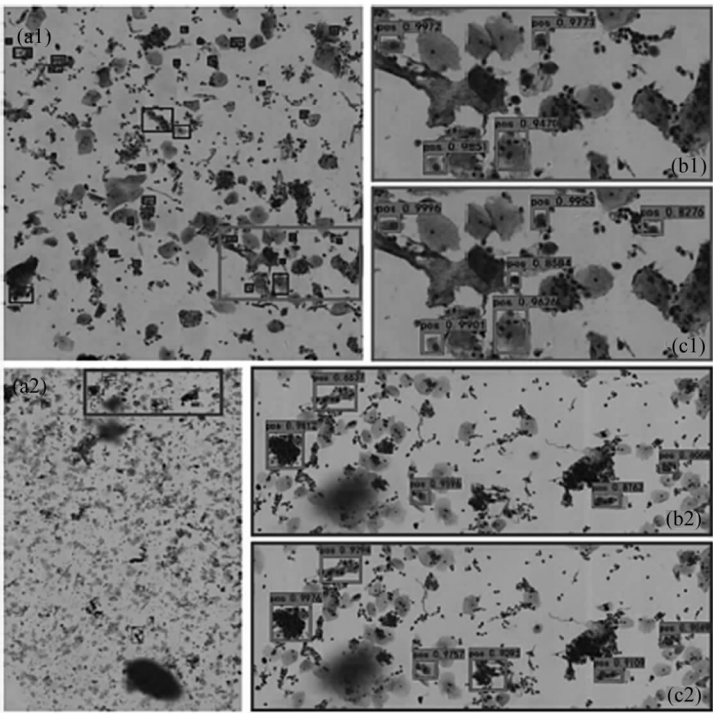


图 10 YOLO v3 模型与 mo-YOLO v3 模型的宫颈癌细胞检测效果对比:(a1)、(a2) 病变细胞标注图;
(b1)、(b2) YOLO v3 模型检测效果;(c1)、(c2) mo-YOLO v3 模型检测效果
Fig. 10 Comparison of cervical cancer cell detection effects between YOLO v3 model and mo-YOLO v3 model;
(a1), (a2) Labeling diagram of pathological cells; (b1), (b2) Results of YOLO v3 model detection;
(c1), (c2) Results of mo-YOLO v3 model detection

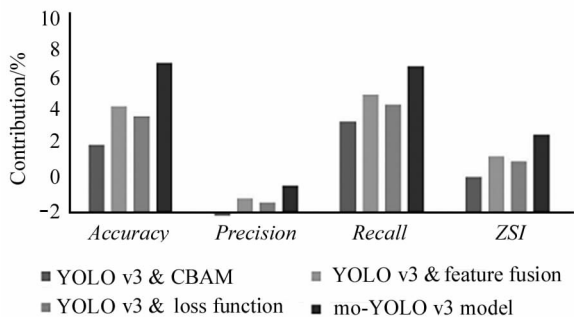


图 11 Mo-YOLO v3 模型各组件贡献度对比结果

Fig. 11 Comparison results of the contribution of each component of the mo-YOLO v3 model

表 3 mo-YOLOv3 模型的消融实验

Tab. 3 Ablation experiments of mo-YOLOv3 model

Methods	Accuracy	Precision	Recall	ZSI
YOLO v3	83.30%	94.71%	86.53%	90.89%
YOLO v3 & CBAM	86.56%	94.58%	91.07%	92.79%
YOLO v3 & feature_fusion	88.40%	95.51%	92.42%	93.84%
YOLO v3 & improved_loss_fun	87.92%	95.28%	91.91%	93.57%
mo-YOLO v3	90.42%	96.20%	93.77%	94.97%

表 4 与其他论文方法比较结果

Tab. 4 Comparison results with other methods

No.	Methods	Accuracy	Precision	Recall (Sensitivity)	mAP(0.5)
1	YOLO v3+[7]	66.78%	67.8%	97.5%	63.4%
2	FCNN[8]	86.67%	96.15%	87.7%	—
3	VGG-Net+[9]	88.18%	90.1%	96.8%	78%
4	Mask R-CNN[10]	89.80%	94.3%	72.5%	57.8%
5	mo-YOLO v3	90.42%	96.20%	93.77%	80.88%

5 结 论

本文针对多尺寸宫颈癌细胞的自动识别问题,提出了一种改进的深度学习目标检测算法模型 mo-YOLO v3。通过策略性的数据增强技术丰富宫颈癌细胞数据形态,为加强网络性能,在模型结构中引入 CBAM 注意力机制模块,并提出一种多尺度特征融合算法提升网络检测精度。此外,本文还研究设计了一项改进的损失函数,实现了模型对宫颈癌细胞尤其是小尺寸癌细胞的准确识别,大大减少了小尺寸癌细胞遗漏的概率。实验结果也充分表明,本文

同时,将本文所提算法与多种专门针对宫颈癌细胞的分割检测算法进行了对比研究。对比算法基于不同的数据集,本文使用的数据集基于多个目标的 ROI 区域,而对比算法大多采用一个细胞作为一个数据对象。实际的医学图像识别任务往往是采集细胞区域让模型标识出阳性细胞,因此本文采用的数据集更具实用性。

对比测试结果见表 4,其中预测框与真实标签框的 IOU(intersect over union)阈值为 0.5。虽然测试结果基于不同的数据集,没有完全相同的比较标准,但从对比结果还是可以看出:本文所提模型在同类模型中具有较好的识别效果和较强的竞争力。

所提 mo-YOLO v3 模型相比其他主流算法,具有明显的优势,对宫颈癌细胞的自动检测技术的临床实现具有一定的指导意义。

参考文献:

- [1] RAHAMAN M M, LI C, WU X, et al. A survey for cervical cytopathology image analysis using deep learning[J]. IEEE Access, 2020, 8: 61687-61710.
- [2] ZENG F, HU S, XIAO K. Research on partial fingerprint recognition algorithm based on deep learning[J]. Neural Computing and Applications, 2019, 31(9): 4789-4798.

- [3] ULIYAND M, SADEGHI S, JALAB H A. Anti - spoofing method for fingerprint recognition using patch based deep learning machine[J]. Engineering Science and Technology, 2020, 23(2): 264-273.
- [4] MENG L, SUN X Y, ZHAO B, et al. An identification method of high-speed railway sign based on convolutional neural[J]. Acta Automatica Sinica, 2020, 46(3): 518-530.
孟 磊, 孙霄宇, 赵滨, 等. 基于卷积神经网络的铁轨路牌识别方法[J]. 自动化学报, 2020, 46(3): 518-530.
- [5] LIU T, ABD-ELRAHMAN A, MORTON J, et al. Comparing fully convolutional networks, random forest, support vector machine, and patch-based deep convolutional neural networks for object-based wetland mapping using images from small unmanned aircraft system[J]. G I Science & Remote Sensing, 2018, 55(2): 243-264.
- [6] LI D M, LI J, LIANG D C, et al. Multiple salient objects detection using multi-scale prior and deep features[J]. Acta Automatica Sinica, 2019, 45(11): 2058-2070.
李东民, 李静, 梁大川, 等. 基于多尺度先验深度特征的多目标显著性检测方法[J]. 自动化学报, 2019, 45(11): 2058-2070.
- [7] XIANG Y, SUN W, PAN C, et al. A novel automation-assisted cervical cancer reading method based on convolutional neural network[J]. Biocybernetics and Biomedical Engineering, 2020, 40(2): 611-623.
- [8] HARANGI B, TOTH J, BOGACSOVICS G, et al. Cell detection on digitized Pap smear images using ensemble of conventional image processing and deep learning techniques[C]//Proceedings of 2019 11th International Symposium on Image and Signal Processing and Analysis (ISPA), September 23-25, 2019, Dubrovnik, Croatia. New York: IEEE, 2019: 19078068.
- [9] HYEON J, CHOI H J, LEE K N, et al. Automating papanicolaou test using deep convolutional activation feature [C]//18th IEEE International Conference on Mobile Data Management (MDM), May 29-June 1, 2017, Daejeon, Korea (South). New York: IEEE, 2017: 382-385.
- [10] SOMPAWONG N, MOPAN J, POOPRASERT P, et al. Automated pap smear cervical cancer screening using deep learning[C]//41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), July 23-27, 2019, Berlin, Germany. New York: IEEE, 2019: 7044-7048.
- [11] SARWAR A, SHARMA V, GUPTA R. Hybrid ensemble learning technique for screening of cervical cancer using papanicolaou smear image analysis[J]. Personalized Medicine Universe, 2015, 4: 54-62.
- [12] ZHANG G R, CHEN X, ZHAO Y, et al. Lightweight YOLOv3 algorithm for small object detection[J/OL]. Laser & Optoelectronic Process. (2021-07-19) [2021-10-21]. <https://kns.cnki.net/kcms/detail/31.1690.TN.20210716.1547.036.html>.
张官荣, 陈相, 赵玉, 等. 面向小目标检测的轻量化 YOLOv3 算法[J/OL]. 激光与光电子学进展. (2021-07-19). [2021-10-21]. <https://kns.cnki.net/kcms/detail/31.1690.TN.20210716.1547.036.html>.
- [13] ZHANG J, HUANG M, JIN X, et al. A real-time Chinese traffic sign detection algorithm based on modified YOLOv2[J]. Algorithms, 2017, 10(4): 127.
- [14] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European Conference on Computer Vision (ECCV), September 8-14, 2018, Munich, Germany. Berlin: Springer, 2018: 3-19.
- [15] ZHENG W, SHENG Q H, REN J. Recognition and classification of diabetic retinopathy based on Improved DR-Net algorithm [J]. Acta Optica Sinica, 2021, 41 (22): 2210002.
郑雯, 沈琪浩, 任佳. 基于 Improved DR-Net 算法的糖尿病视网膜病变识别与分级[J]. 光学学报, 2021, 41(22): 2210002.
- [16] ZHANG X, ZOU Y, SHI W. Dilated convolution neural network with LeakyReLU for environmental sound classification[C]//2017 22nd International Conference on Digital Signal Processing (DSP), August 23-25, 2017, London, UK. New York: IEEE, 2017: 2165-3577.
- [17] LIAN Q S, FU L P, CHEN S Z, et al. A compressed sensing algorithm based on multi-scale residual reconstruction network[J]. Acta Automatica Sinica, 2019, 45(11): 2082-2091.
练秋生, 富利鹏, 陈书贞, 等. 基于多尺度残差网络的压缩感知重构算法[J]. 自动化学报, 2019, 45(11): 2082-2091.
- [18] JIANG W T, ZHANG C, ZHANG S C, et al. Multiscale feature map fusion algorithm for target detection[J]. Journal of Image and Graphics, 2019, 24(11): 1918-1931.
姜文涛, 张驰, 张晟翀, 等. 多尺度特征图融合的目标检

测[J]. 中国图象图形学报, 2019, 24(11): 1918-1931.

[19] YANG H, CHEN L, CHEN M, et al. Tender tea shoots recognition and positioning for picking robot using improved YOLO-V3 model [J]. IEEE Access, 2019, 7: 180998-181011.

[20] MA Y J, LIU P P. Convolutional neural network based on DenseNet evolution for image classification algorithm[J]. Laser & Optoelectronic Progress, 2020, 57 (24): 241001.

马永杰, 刘培培. 基于 DenseNet 进化的卷积神经网络图像分类算法[J]. 激光与光电子学进展, 2020, 57(24): 241001.

[21] QIAO L C. SAR image target recognition method combining multi-resolution representation and complex domain CNN[J]. Laser & Optoelectronic Progress, 2020, 57 (24): 241007.

乔良才. 结合多分辨率表示和复数域 CNN 的 SAR 图像目标识别方法[J]. 激光与光电子学进展, 2020, 57(24): 241007.

作者简介:

任 佳 (1977—), 女, 博士, 副教授, 主要从事模式识别、智能故障诊断方面的研究.