

# 一种具有抗噪性能的圆形目标检测器

蔡钟晟, 陈 飞, 曾勋勋

(福州大学数学与计算机科学学院, 福建 福州 350108)

[摘要] 目前许多目标检测模型在噪声环境中都会出现精度下降. 为了提高目标检测模型在噪声环境中的精度, 本文从两方面提高检测精度. 首先, 提出一个基于残差结构的抗噪特征提取模块, 为后续的网络模块提供支撑. 其次, 利用目标的先验信息, 针对性地改进模型的锚框设计和损失函数设计. 根据目标的形状先验信息设计锚框的形状. 将 IOU 损失函数作为模型的 Bbox 损失函数, 其中 IOU 损失项及最小闭包区域根据目标形状先验信息计算. 实验数据集为血细胞数据集和下旁腺数据集, 基准对照模型为 YOLOv3 和 RetinaNet, 同时也可移植到其他检测模型. 在血细胞数据集环境中, 比较 YOLOv3 的精度由 62.7 提高到 75.7. 在下旁腺数据集中同样有所提升.

[关键词] 噪声图像, 目标检测, IOU 损失函数, 形状特征, 先验信息

[中图分类号] TP751 [文献标志码] A [文章编号] 1001-4616(2021)04-0085-09

## A Circular Object Detector with Anti-Noise Performance

Cai Zhongsheng, Chen Fei, Zeng Xunxun

(College of Mathematics and Computer Science, Fuzhou University, Fuzhou 350108, China)

**Abstract:** At present, many of the target detection models will suffer from precision degradation in noisy environment. In order to improve the accuracy of target detection model in noisy environment, two methods are proposed. Firstly, an anti-noise feature extraction module based on residual structure is proposed to support the subsequent network modules. Secondly, the anchor frame design and loss function design of the model are improved by using the prior information of the target. The shape of the anchor frame is designed according to the shape prior information of the target. The IOU loss function is taken as the Bbox loss function of the model, in which the IOU loss term and the minimum closure area are calculated according to the prior information of the target shape. The experimental data sets were blood cell data sets and inferior parathyroid gland data sets, and the benchmark control models were YOLOv3 and RetinaNet, which can also be transplanted to other detection models. In the blood cell dataset environment, the accuracy of YOLOv3 is 62.7, which is improved to 75.7, which is a great improvement. In the inferior parathyroid gland data set, there was also an improvement.

**Key words:** noisy image, object detection, IOU loss, shape feature, prior information

图像质量对基于深度学习的目标检测方法有着较大影响<sup>[1]</sup>, 但是生产环境中的复杂环境通常难以确保图像质量. 从而造成许多模型的精度大幅下降. 例如, 噪声环境中经典的 YOLOv3<sup>[2]</sup> 模型精度出现大幅下降, 效果如图 1 左所示. 因此本文引入深度残差网络提取图像抗噪特征, 同时结合目标先验形状信息, 实现在噪声环境下目标的精确识别. 本文的研究目标主要包括血细胞和下旁腺.

因为公共数据集通常目标种类繁多且形状不固定, 所以现有基于深度学习的目标检测器通常都是用矩形框检测目标, 但血细胞和下旁腺都是近似圆形的目标. 针对这种形状相对较固定的目标, 矩形不是最好的选择. 本文基于血细胞和下旁腺圆形形状的先验信息, 提出了圆形锚框和圆形 IOU 损失函数. 充分利用目标的先验信息, 更好地提取目标特征. 同时本文的方法并不局限于特定模型, 可以在其他模型使用.

收稿日期: 2021-05-16.

基金项目: 国家自然科学基金项目(61771141)、福建省教育厅中青年教师教育科研项目(JAT190020).

通讯作者: 陈飞, 博士, 副教授, 研究方向: 图像处理与机器学习. E-mail: chenfei314@fzu.edu.cn

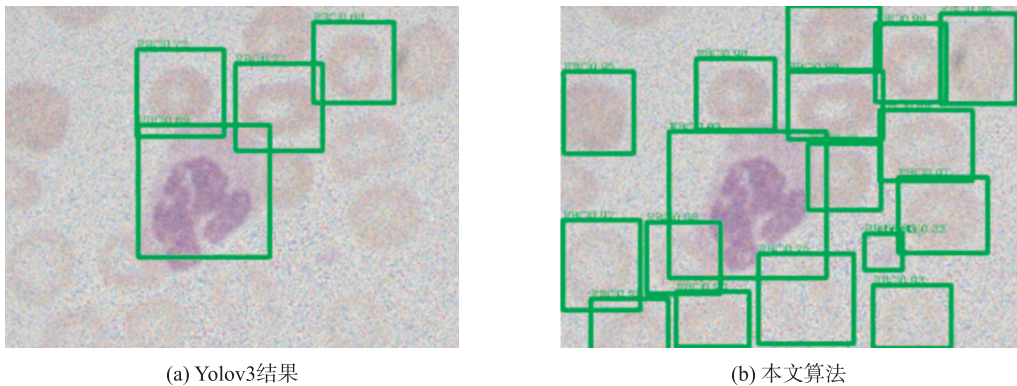


图 1 血细胞检测结果比较  
Fig. 1 Blood cell test results

# 1 相关工作

目前基于深度学习的目标检测算法主要通过两种方式检测目标<sup>[3]</sup>. 一种是基于锚框 Anchor 的模型, 这种模型又可分为一阶段模型和两阶段模型. 一阶段经典模型有 YOLOv3, 也就是 Yolo, You Only Look Once 的第三版续作. 两阶段经典模型有 Faster R-CNN<sup>[4-5]</sup> 和 Mask R-CNN<sup>[6]</sup>, 两个模型都是 RCNN 的续作, 也就是 Region With CNN Features 的续作. 另一种是 Anchor-free 的模型, 也就是无锚框的模型, 通过寻找目标特征点来确定目标的中心点位置和宽高. 经典模型有 FCOS<sup>[7]</sup> 和 CornerNet<sup>[8]</sup>, 也就是 Fully Convolutional One-Stage Object Detection 和角点检测网络.

锚框概念在 YOLO9000<sup>[9]</sup> 中提出, 之后在基于图像的深度学习任务中获得广泛应用<sup>[10]</sup>. YOLO9000 为 YOLO 的续作, 也就是 YOLOv2. 在 YOLO9000 中锚框的宽和高为根据经验设置的超参数. 之后提出的 YOLOv3 中锚框的宽和高通过 K-means<sup>[11]</sup> 聚类, 也就是 K 均值聚类获得, 锚框的设置主要依靠目标数据集的先验信息. 通过修改锚框的设置可以让目标检测召回率有所提升<sup>[9]</sup>.

关于先验信息的研究有很多. 文献[12]提出稀疏字典学习对背景特征图进行统计分析和估计, 最后综合颜色对比度特征和背景特征生成最终的鲁棒显著性图. 文献[13]提出利用背景先验知识, 用相似性度量计算得到背景显著性图, 后用能量函数对显著性图进行优化, 提高了算法检测的质量. 文献[14]提出将医学先验信息融入图像增强. 通过医学知识判断目标区域的亮度阈值, 以此阈值来对图像进行处理, 得到增强后的图片, 从而提高模型精度. 文献[15]提出结合霍夫变换<sup>[16]</sup>和 YOLO9000 的方法. 利用霍夫变换和图像旋转确定目标检测最佳识别角度, 解决 YOLO9000 旋转不变性较差的问题. 文献[17]提出了一种结合 KL 散度和形状约束的 Faster R-CNN 检测方法. KL 散度也就是相对熵. 在 Faster R-CNN 模型的基础上, 增加对目标边界框分布的预测, 同时使用 KL 散度量目标坐标预测分布与真值分布之间的距离, 并将其作为损失函数优化模型参数, 进一步将数据集中不同类别目标的形状特征作为约束加入损失函数中, 以提高模型的边界框回归精度.

本文提出的圆形锚框, 针对特定目标, 让目标形状特征和锚框形状一致, 可以更好地利用目标的形状先验信息, 提取目标特征. 同时, 增加抗噪特征提取网络, 有助于模型在噪声环境中保持精度稳定.

# 2 圆形锚框及 IOU 损失函数

## 2.1 圆形锚框

模型在训练的初始分配正负样本. 正负样本的分配与锚框有密不可分的关系. 正负样本通过锚框与标注框之间的 IOU 进行分配. IOU 即为交并比, 公式如式(1)所示. 其中  $S_i$  为标注框与预测框的交集面积,  $S_u$  为标注框与预测框的并集面积. 通常根据设置的 IOU 阈值进行判定, 超过阈值判定为正样本, 否则判定为负样本. 锚框的形状会对 IOU 的计算造成影响. 两个圆形框计算出的 IOU 和两个矩形框计算出的 IOU 不一定相同. 在相同的回归偏差下, 圆形框的 IOU 计算可能会得到更小的 IOU, 进而得到更小的 mAP. mAP 为 Mean Average Precision, 即为各类别 AP 的平均值, 是目标检测的评价指标. Precision 公式如式(2)所示. 其中 TP 为检测器判定为正样本且标签也为正样本的样本数据. TP 与 FP 的和即为所有检测

器判定为正样本的样本数据.

$$IOU = \frac{S_i}{S_u}, \quad (1)$$

$$precision = \frac{TP}{TP+FP}, \quad (2)$$

矩形框网络需要回归的参数有中心点 $(x_r, y_r)$ , 还有矩形框的高 $h$ 和宽 $w$ . 其中 $x, y$ 为轴坐标. 圆形框网络需要回归的参数同样包含中心点 $(x_c, y_c)$ , 以及圆形框的直径 $d$ . 因为矩形框需要回归4个参数, 而圆形框只需要回归3个参数, 不同参数量难以比较. 因此假设矩形框的宽 $w$ 和高 $h$ 相等, 即矩形框都为正方形. 那么矩形框和圆形框一样, 只需要回归3个参数. 以下假设两种情况.

第1种情况如图2所示: 固定中心点 $(x_r, y_r)$ 和 $(x_c, y_c)$ , 只有圆形框的直径 $d$ , 矩形框的宽 $w$ 和高 $h$ 有误差. 也就是预测的中心点位置和标签的中心点位置一样, 只有预测的 $w, d$ 参数和标签值不一样. 第2种情况固定宽 $w$ , 高 $h$ 和直径 $d$ 参数, 只有圆形框的中心点 $(x_c, y_c)$ 和矩形框的中心点 $(x_r, y_r)$ 有误差. 也就是预测的 $w, d$ 参数和标签值一样, 只有预测的中心点位置和标签的中心点位置不一样.

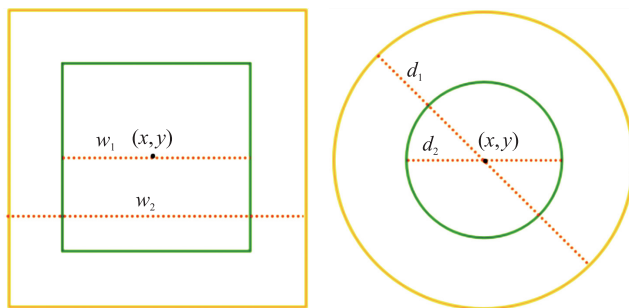


图2 情况1:缩放偏差

Fig. 2 Case 1: scaling deviation

图2左中绿色框为标注框 $k_1$ , 橙色框为预测框 $k_2$ .  $k_1$ 的宽为 $w_1$ ,  $k_2$ 的宽为 $w_2$ ,  $w_2 = 2w_1$ .  $k_1$ 和 $k_2$ 有共同的中心点 $(x, y)$ . IOU计算公式如式(1)所示. 根据IOU计算公式, 此时左的IOU为0.25. 右中绿色框为标注框 $k_1$ , 橙色框为预测框 $k_2$ .  $k_1$ 的直径为 $d_1$ ,  $k_2$ 的直径为 $d_2$ ,  $d_2 = 2d_1$ .  $k_1$ 和 $k_2$ 有共同的中心点 $(x, y)$ . 根据IOU计算公式可得右的IOU为0.25. 假设 $d_1 = w_1$ ,  $d_2 = w_2$ . 同时假设左的 $x$ 等于右的 $x$ , 左的 $y$ 等于右的 $y$ . 即左和右的标签值相同, 预测值相同, 偏差也相同. 可见在第1种情况下, 即预测的中心点 $(x, y)$ 和标签相同, 而预测的 $w, d$ 参数和标签不同. 这种情况下IOU是相同的. 第2种情况如图3所示.

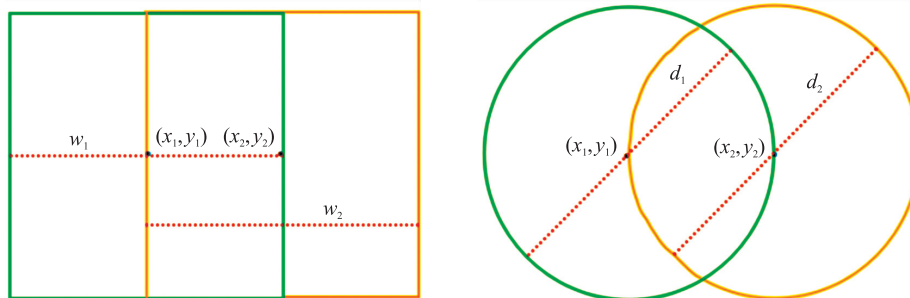


图3 情况2:移位偏差

Fig. 3 Case 2: displacement deviation

图3左中绿色框为标注框 $k_1$ , 橙色框为预测框 $k_2$ .  $k_1$ 的宽为 $w_1$ ,  $k_2$ 的宽为 $w_2$ ,  $w_2 = w_1$ .  $k_1$ 的中心点为 $(x_1, y_1)$ ,  $k_2$ 的中心点为 $(x_2, y_2)$ .  $x_2 = x_1 + w_1/2$ ,  $y_2 = y_1$ . 根据IOU计算公式可得左侧的IOU为0.333. 右中绿色框为标注框 $k_1$ , 橙色框为预测框 $k_2$ .  $k_1$ 的直径为 $d_1$ ,  $k_2$ 的直径为 $d_2$ ,  $d_2 = d_1$ .  $k_1$ 的中心点为 $(x_1, y_1)$ ,  $k_2$ 的中心点为 $(x_2, y_2)$ .  $x_2 = x_1 + d_1/2$ ,  $y_2 = y_1$ . 根据IOU计算公式可得右侧的IOU为0.243. 假设 $d_1 = w_1$ ,  $d_2 = w_2$ . 同时假设左的 $(x_1, y_1)$ 等于右的 $(x_1, y_1)$ , 左的 $(x_2, y_2)$ 等于右的 $(x_2, y_2)$ . 即左和右的标签值相同, 预测值相同, 偏差也相同. 可见在第2种情况下, 即预测 $w, d$ 参数和标签相同, 而预测的中心点 $(x, y)$ 和标签不相同. 这种情况下IOU是不同的, 且圆形框的IOU小于矩形框的IOU. 两种情况如表1所示.

根据表1, 可以看出圆形框的IOU算法本身不会

表1 圆形框IOU与矩形框IOU的对比

Table 1 Comparison of circular IOU and rectangular IOU

IOU	Round box	Rectangular box
scaling deviation	0.250	0.250
displacement deviation	0.243	0.333

带来 IOU 的提升. 而 mAP 的计算与 IOU 直接相关, IOU 的降低会直接导致 mAP 的降低. 在相同的负样本阈值下, 更低的 IOU 会使得负样本增多, 介于正样本和负样本之间的忽略样本数量变少. 同时因为 IOU 可能的下降, 圆形框之间的 IOU 算法也不会直接引起 mAP 的提升, 在第 2 种情况下甚至会降低 mAP. 因此如果模型性能有所提升, 则是因为圆形锚框能更好地提取目标的特征.

## 2.2 IOU 损失函数

IOU 损失函数是由 Rezatofighi 和 Zheng 等提出的一系列基于交并比 IOU 的边界框损失函数, 包括 GIOU, Generalized Intersection Over Union 以及 DIOU<sup>[18-19]</sup>, Distance-IOU Loss 等. 本文基于矩形框的 DIOU<sup>[19]</sup> 损失函数提出了圆形框的 IOU 损失函数. 原本的 DIOU 损失函数为:

$$loss = 1 - iou + \frac{\rho^2(b, b^{gt})}{c^2}, \quad (3)$$

式中,  $iou$  表示目标检测中的交并比. 最后一项中的  $b$  表示预测框的中心点,  $b^{gt}$  表示标注框的中心点.  $\rho^2$  表示欧式距离的平方.  $c$  表示同时包含预测框和标注框的最小闭包矩形区域的对角线距离.

因为  $iou$  与目标检测任务的指标 mAP 直接相关, 所以用  $iou$  直接作为损失函数相较传统损失函数更直观, 并且可以解决一些 Loss 相同, 但  $iou$  值却完全不同的问题<sup>[18]</sup>. 但为矩形框设计的 DIOU 不能在其他形状使用, 因为只有矩形有对角线, 而圆形或其他形状没有.

本文对 DIOU 损失函数进行改进. 首先是将  $iou$  的计算方法由矩形之间的  $iou$  计算改变为圆形之间的  $iou$  计算. 其次是将损失函数最后一项中的最小闭包区域矩形的对角线距离  $c$  改为最小闭包区域圆形的直径  $d$ . 损失函数如下所示:

$$loss = 1 - iou + \frac{\rho^2(b, b^{gt})}{d^2}, \quad (4)$$

损失函数也可采用 YOLOv3 原本的 Bbox 损失函数, 也就是 Bound Box 损失函数. 但损失函数中的矩形框宽  $w$  和高  $h$  的回归要改成圆形框直径  $d$  的回归.

## 3 抗噪特征提取模块

因为实际生产环境的复杂条件, 使图像质量无法保证. 目前较为经典的模型, 都会在一定程度上受到噪声的影响, 造成精度下降. 针对噪声环境, 本文构建抗噪特征提取模块.

由于真实噪声图像的稀缺, 本文通过生成随机标准差的噪声图像, 与原始图像相加获得噪声图像. 由于实际生产环境中噪声的不确定性, 随机标准差的噪声可以更好地模拟真实环境. 噪声模型采用高斯分布. 均值为 0, 标准差为 0~55 的随机数. 噪声图像表达式如式(5)所示. 其中  $x$  表示原始图像,  $v$  表示噪声图像.

$$y = x + v, \quad (5)$$

抗噪特征提取模块采用残差网络<sup>[20-21]</sup>. 残差结构公式如式(6)所示. 因此网络只需要学习  $F(x, w)$  部分中的  $w$ , 也就是  $y$  与  $x$  之间的差距. 残差结构的式(6)与图像表达式(5)相似, 可以通过训练  $w$  使得模型学习到噪声的分布.

$$y = x + F(x, w), \quad (6)$$

抗噪特征提取模块中, 肤色区域为  $3 \times 3$  的卷积层, 绿色区域为激活函数, 土黄色区域为 BN 层. ● 为分辨率相同的特征图相减. 网络整体为一个残差结构, 最终的输出为输入图像减网络输出. 如图 4 所示.

模型头部的网络只有两层, 一层  $3 \times 3$  的卷积和一层 Relu 激活函数. 尾部网络则只有一层  $3 \times 3$  的卷积. 中间部分重叠的网络单元数可以根据情况调整, 本文中中间的重叠单元设置了 20 层.

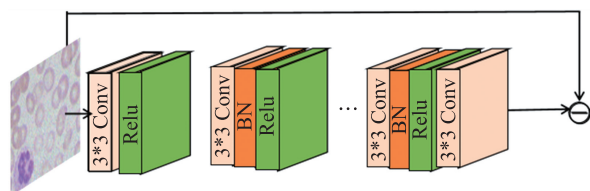


图 4 抗噪特征提取模块

Fig. 4 Anti-noise feature extraction module



4 整体框架及实现算法

以 Yolov3 为例,模型整体框架如图 5 所示. Yolov3 模型的流程如下:预处理输入图像,将图像变成固定大小的输入,并进行数据增强操作. 然后将图像输入模型,模型的 Backbone 采用 Darknet53<sup>[2]</sup>. 输出采用多尺度的目标检测方法,即采用 3 个不同尺度的输出,输出不同尺度的特征图来检测不同尺度的目标. 每个尺度的特征图上设置对应尺度的锚框. 输出特征图的通道数为  $(\text{num\_class}+5)\times 3$ . 其中 3 为每个尺度的特征图上的 Anchor 的数量, num\_class 为检测的目标的类别数目,5 为每个锚框的前后景置信度,以及对目标框的 xywh 微调的参数. 损失函数分为 3 个部分,分别是目标分类损失函数,前后景分类损失函数以及 xywh 的损失函数.

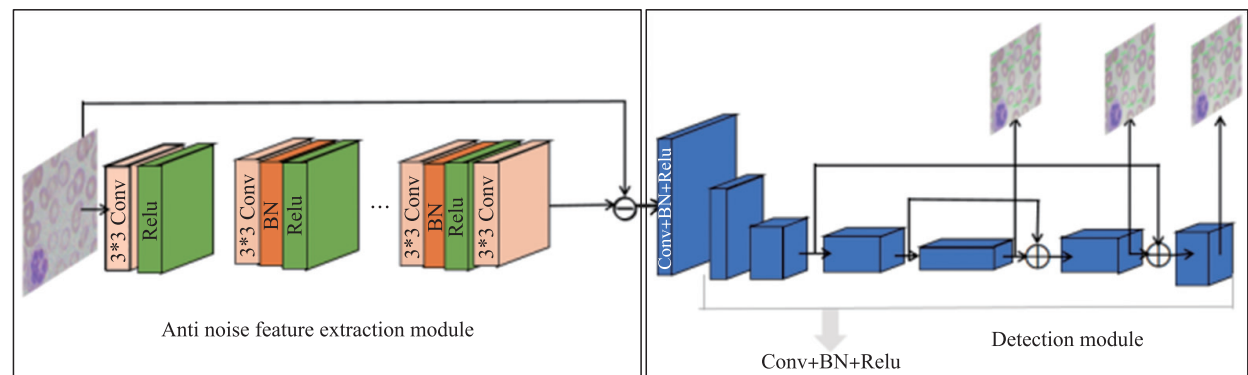


图 5 网络模型的整体框架

Fig. 5 The overall framework of network model

本文在 Yolov3 的主干网络之前设计了抗噪特征提取模块,为后续低质量图像目标识别提供支撑. 同时利用目标的形状信息,将锚框设置为圆形. 损失函数的 Bbox 调整部分的损失函数改为圆形 IOU 损失函数. 整体网络结构如图 5 所示.

其中肤色区域为  $3\times 3$  的卷积层,绿色区域为激活函数,土黄色区域为 BN 层,蓝色区域为卷积层 Conv, 批量归一化层 BN 和激活函数 Relu 构成的网络单元.  $\ominus$  为分辨率相同的特征图进行的对应点相减的操作,  $\oplus$  为通道拼接操作.

5 实验与结果

本节实验 Baseline 为 Yolov3. 实验数据集为血细胞数据集和下旁腺数据集. 下旁腺数据集只有下旁腺一个类别,来源于手术,全腔镜胸乳入路左侧甲状腺叶加峡部切除加中央区淋巴结清扫术的视频,共有 2 359 张图像. 血细胞数据集包括 3 个类别,其中大部分为红细胞,小部分为白细胞和血小板,数据集来源于 [https://github.com/Shenggan/BCCD\\_Dataset](https://github.com/Shenggan/BCCD_Dataset). 共有 364 张图像. 如图 6 所示.

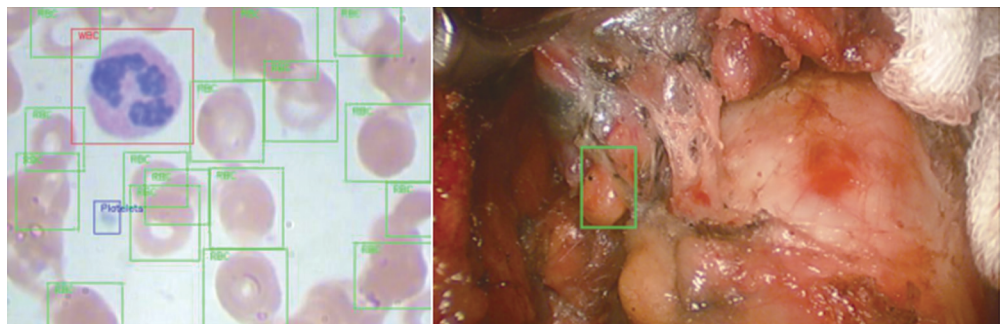


图 6 血细胞和下旁腺

Fig. 6 Blood cell and inferior parathyroid gland

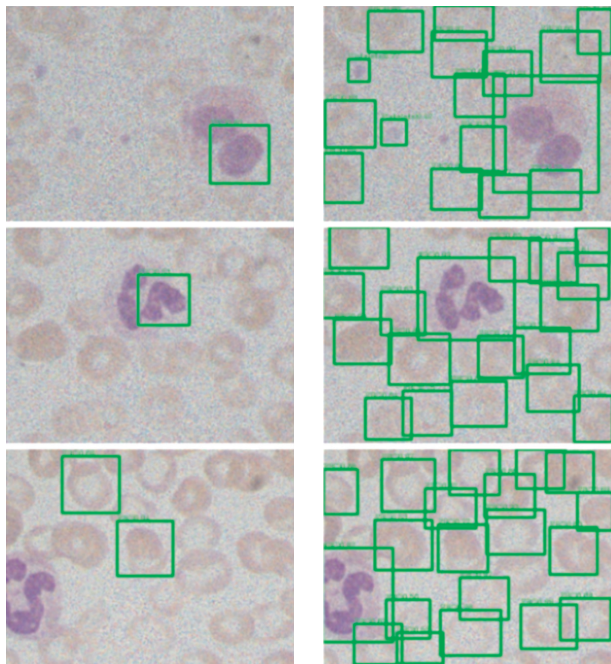
5.1 抗噪特征提取模块

在实验中,首先用对抗噪特征提取模块进行预训练,预训练结束后,将预训练参数文件作为初始化文

件输入抗噪特征提取模块,训练整个模型. 预训练共 20 轮,在 CBSD68 数据集上进行训练,噪声标准差设置为 0~55 的随机数. 预训练结束后,将抗噪特征提取模块和后续网络模块联合训练. 共训练 150 轮,分为两个阶段. 第一阶段,也就是前 100 轮,冻结抗噪特征提取网络的参数进行训练. 第二阶段,后 50 轮,解冻抗噪特征提取模块的参数进行训练.

Yolov3 模型中,数据集为无噪声的下旁腺数据集时,mAP 为 64.9. 输入 0~55 的随机噪声,mAP 为 15.8,下降 49.1. 数据集为无噪声的血细胞数据集时,mAP 为 62.7. 输入 0~55 的随机噪声,mAP 为 22.7,下降 40. 可见 Yolov3 在噪声环境精度会大幅下降,其他经典模型也相同.

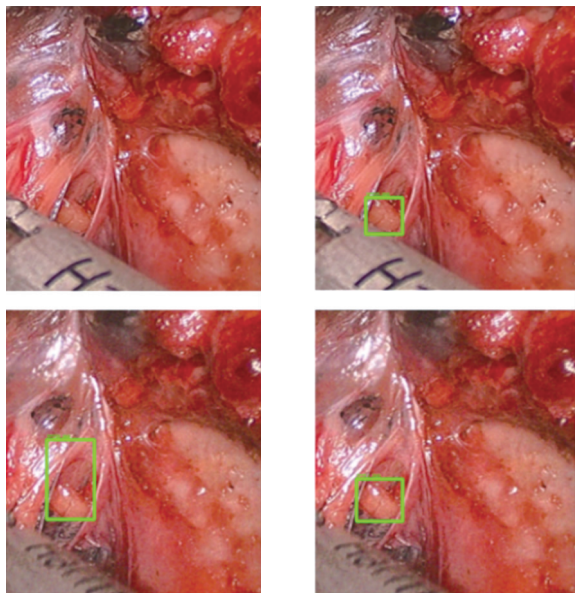
结合抗噪特征提取模块后,Yolov3 在噪声的下旁腺数据集中 mAP 为 49.2,下降 15.7. 在噪声的血细胞数据集中 mAP 为 66.6,提高 3.9. 可见抗噪特征提取模块帮助 Yolov3 在噪声环境精度中更加稳定. 实验结果如图 7 和图 8 所示.



(a) Baseline结果 (b) 添加抗噪特征提取模块后的结果

图 7 噪声环境中血细胞检测结果比较

Fig. 7 Comparison of blood cell detection results in noisy environment



(a) Baseline结果 (b) 添加抗噪特征提取模块后的结果

图 8 噪声环境中下旁腺检测结果比较

Fig. 8 Comparison of detection results of inferior parathyroid gland in noisy environment

5.2 圆形锚框和 IOU 损失函数

血细胞数据集和下旁腺数据集中目标形状都接近圆,标注框接近正方形,因此取矩形锚框的宽高平均值作为圆形锚框的直径. 这使得圆形锚框接近于矩形锚框的内接圆,相当于缩小了目标的面积. 因此对圆形锚框检测是较为不利的.

Bbox 损失函数为 IOU 损失函数. 实验调整了 Bbox 回归损失在损失函数中的权重,因为圆形 IOU 损失函数本身非常小. 损失函数包括 3 项,第 1 项常数 1,第 2 项 IOU 的值域为[0,1]. 第 3 项可视为预测框与标注框圆心距离和最小闭包区域圆形直径的比值的平方. 当圆形框不为内含关系时,最小闭包区域直径  $d$  和圆心距离的关系为式(7). 当圆形框为内含关系时,最小闭包区域直径  $d$  表达式为式(8),式(8)中  $r$  为标注框和预测框中较大圆形框的半径. 因此第 3 项的值域也为[0,1). 综合的损失函数式(4)值域为[0,2).

$$d=r_{gt}+\rho(b,b_{gt})+r, \tag{7}$$

$$d=2r. \tag{8}$$

因此相较分类损失,如果权重维持不变,Bbox 回归部分损失将过于微小. 损失函数完全由分类任务主导,分类与回归损失会失衡. 上述参数设置下在血细胞数据集中 mAP 为 75.7,相较 Baseline 的 62.7 提高了 11.4. 同样的设置在下旁腺数据集上进行测试,得到最高 mAP 为 75.3,相比较原本 Baseline 的 64.9,提高



了 10.4. 实验结果如图 9 和图 10 所示.

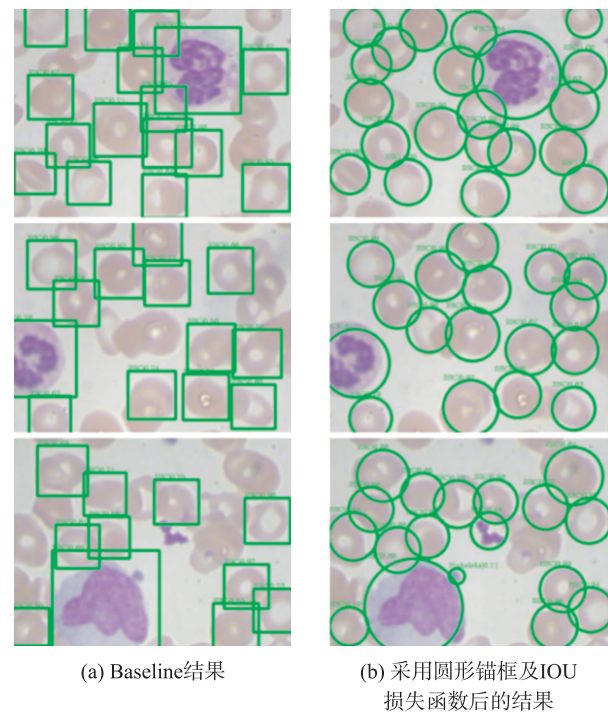


图 9 血细胞检测结果比较

Fig. 9 Blood cell test results were compared

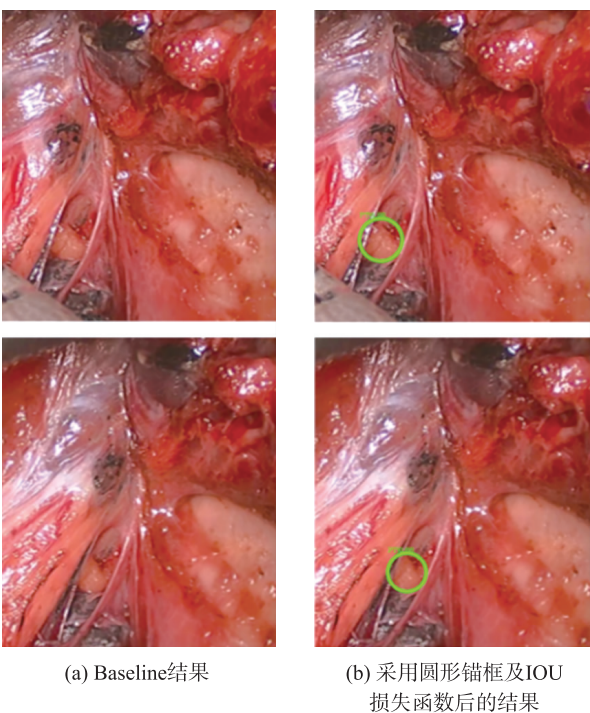


图 10 下旁腺检测结果比较

Fig. 10 Inferior parathyroid gland test results were compared

5.3 消融实验和对比实验

消融实验如表 2 和表 3 所示. 从表中可以看出抗噪特征提取模块能帮助模型在噪声环境中保持精度稳定,并且抗噪特征提取模块在干净环境中也不会影响精度. 圆形锚框和 IOU 损失函数在两个数据集中都有效提高了模型精度. 采用 Yolov3 在下旁腺数据集中训练一轮耗时约为 263 s,训练 100 轮,耗时约为 438 min. 在血细胞数据集中训练一轮耗时约为 27 s,训练 100 轮,耗时约为 45 min. 对比实验如表 4 和表 5 所示. 从表中可以看出,本文将方法应用在 RetinaNet 和 Yolov3 上,相比较原本模型,都取得了效果的提升.

表 2 血细胞数据集中 Yolov3 的实验结果

Table 2 Experimental results of Yolov3 in blood cell dataset

mAP	Baseline	Anti-noise feature extraction network	Anti-noise feature extraction network+ Round box+IOU loss
Clean image	62.7	59.8	75.7
Noisy image	22.7	66.6	73.0

表 3 下旁腺数据集中 Yolov3 的实验结果

Table 3 Experimental results of Yolov3 in inferior parathyroid gland dataset

mAP	Baseline	Anti-noise feature extraction network	Anti-noise feature extraction network+ Round box+IOU loss
Clean image	64.9	66.2	75.3
Noisy image	15.8	49.2	73.3

表 4 血细胞数据集中不同模型对比实验结果

Table 4 The comparative experimental results of different models in blood cell data set

mAP	FasterRCNN	GA <sup>[22]</sup>	Atss <sup>[23]</sup>	Foveabox <sup>[24]</sup>	LibraRCNN <sup>[25]</sup>	RetinaNet <sup>[26]</sup>	Our+RetinaNet	Yolov3	Our+Yolov3
Clean image	78.0	79.6	77.8	75.1	73.3	75.7	81.2	62.7	75.7
Noisy image	52.5	59.1	54.5	50.7	65.7	48.1	72.7	22.7	73.0

表 5 下旁腺数据集中不同模型对比实验结果

Table 5 The comparative experimental results of different models in blood cell data set

mAP	FasterRCNN	Ssd <sup>[27]</sup>	Fcos	Foveabox	Dconv <sup>[28]</sup>	RetinaNet	Our+RetinaNet	Yolov3	Our+Yolov3
Clean image	74.6	66.5	64.6	79.8	71.4	74.7	80.5	64.9	75.3
Noisy image	55.7	49.4	38.5	68.1	45.9	38.4	89.3	15.8	73.3

## 6 结论

本文提出了基于残差结构的抗噪特征提取模块,以及更适合检测圆形目标的圆形锚框和 IOU 损失函数. 结合抗噪特征提取模块后,在噪声环境中网络的检测精度相比较传统模型更加稳定. 检测圆形目标时,采用圆形锚框及 IOU 损失函数可以提高模型的精度. 本文的方法在下旁腺数据集和血细胞数据集上进行验证. 当数据集带有随机噪声,抗噪特征提取模块可以帮助网络保持精度基本稳定,只比在干净环境下的精度略低. 实验的两个数据集目标形状都接近圆形,在其他形状近似圆形的目标也适用该模型. 大部分的水果检测分类以及相对接近圆形的人脸检测都可以使用. 在两个数据集上采用圆形锚框及 IOU 损失函数,模型精度都有所提高. 以上实验基于 YOLOv3 和 RetinaNet 都有较好的效果,同时该方法并不局限于 YOLOv3 和 RetinaNet,可用于其他检测模型.

## [参考文献]

- [1] ZHAO X, SATOH Y, TAKAUJI H, et al. Robust adapted object detection under complex environment[C]//2011 8th IEEE International Conference on Advanced Video and Signal Based Surveillance. Washington DC:IEEE,2011:261-266.
- [2] REDMON J, FARHADI A. YOLOv3:an incremental improvement[C]//Computer Vision and Pattern Recognition. Guangzhou, 2018.
- [3] SUHAIL A, JAYABALAN M, THIRUCHELVAM V. Convolutional neural network based object detection;a review[J]. Journal of critical reviews,2020,7(11):2020.
- [4] 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,2017,39(6):1137-1149.
- [5] 张中宝,王洪元,张继,等. 基于 Faster-RCNN 的遥感图像飞机检测算法[J]. 南京师大学报(自然科学版),2018,41(4):79-86.
- [6] HE K, GKIOXARI G, DOLLAR P, et al. Mask R-CNN[J]. IEEE transactions on pattern analysis & machine intelligence, 2017,38(5):2961-2969.
- [7] TIAN Z, SHEN C, CHEN H, et al. FCOS:fully convolutional one-stage object detection[C]//2019 IEEE/CVF International Conference on Computer Vision. Seoul, Korea:IEEE,2020.
- [8] LAW H, DENG J. CornerNet:detecting objects as paired keypoints[J]. International journal of computer vision,2018,31(10):734-750.
- [9] REDMON J, FARHADI A. YOLOv3: better, faster, stronger[J]. IEEE conference on computer vision and pattern recognition,2017:6517-6525.
- [10] ZHAO Z Q, ZHENG P, XU S, et al. Object detection with deep learning;a review[J]. IEEE transactions on neural networks and learning systems,2019,30(11):3212-3232.
- [11] WONG J A, HARTIGAN M A. Algorithm AS 136;a K-means clustering algorithm[J]. Journal of the royal statistical society, 1979,28(1):100-108.
- [12] JIAN M, WANG J, YU H, et al. Visual saliency detection by integrating spatial position prior of object with background cues[J]. Expert systems with applications,2020,168(11):1142-1153.
- [13] 夏晨星. 基于背景先验和目标信息的自底向上显著性检测方法研究[D]. 长沙:湖南大学,2019.
- [14] ZHAO M, CHE X, LIU H, et al. Medical prior knowledge guided automatic detection of coronary arteries calcified plaque with cardiac CT[J]. Electronics,2020,9(12):2122-2134.
- [15] ZHAO H, ZHANG Z. Improving neural network detection accuracy of electric power bushings in infrared images by hough transform[J]. Sensors,2020,20(10):2931-2941.
- [16] BALLARD D H. Generalizing the hough transform to detect arbitrary shapes[J]. Pattern recognition,1981,13(2):111-122.
- [17] 赵振兵,李延旭,甄珍,等. 结合 KL 散度和形状约束的 Faster R-CNN 典型金具检测方法[J]. 高电压技术,2020,46(9):3018-3026.
- [18] REZATOFIGHI H, TSOI N, GWAK J Y, et al. Generalized intersection over union;a metric and a loss for bounding box regression[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Xi'an:IEEE,2019.
- [19] ZHENG Z, WANG P, LIU W, et al. Distance-IoU loss:faster and better learning for bounding box regression[C]//AAAI



- Conference on Artificial Intelligence. New York, 2020.
- [20] HE K M, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas, America; 2016.
  - [21] ZHANG K, ZUO W, CHEN Y, et al. Beyond a gaussian denoiser: residual learning of deep CNN for image denoising[J]. IEEE transactions on image processing, 2016, 26(7): 3142–3155.
  - [22] WANG J Q, CHEN K, YANG S, et al. Region proposal by guided anchoring[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Xi'an, 2019.
  - [23] ZHANG S, CHI C, YAO Y, et al. Bridging the gap between anchor-based and anchor-free detection via adaptive training sample selection[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, America; IEEE, 2020.
  - [24] KONG T, SUN F, LIU H, et al. FoveaBox: beyond anchor-based object detection[J]. IEEE transactions on image processing, 2020, 29: 7389–7398.
  - [25] PANG J, CHEN K, SHI J, et al. Libra R-CNN: towards balanced learning for object detection[C]//Conference on Computer Vision and Pattern Recognition. Seattle, America, 2020.
  - [26] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE transactions on pattern analysis & machine intelligence, 2020, 42(2): 318–327.
  - [27] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European Conference on Computer Vision. Amsterdam Netherlands, 2016.
  - [28] FERGUS R, TAYLOR G W, ZEILER M D. Adaptive deconvolutional networks for mid and high level feature learning[C]//International Conference on Computer Vision. Shenzhen; IEEE Computer Society, 2011.

[责任编辑:陆炳新]