

文章编号: 2095-2163(2020)05-0030-06

中图分类号: TP399

文献标志码: A

# 一种新的基于形态学与 Yolov3 算法的显微镜图像中孢子识别方法

李鑫铭, 赵磊, 邵宝民, 王栋

(山东理工大学 计算机学院, 山东 淄博 255000)

**摘要:** 针对国内对显微镜下致病菌种识别以及细胞活性检验等问题, 提出以 darknet-53 网络模型与 Yolov3 算法取出疑似目标, 再配以传统形态学算法进行筛选, 从而达到准确、快速、智能化的诊断方式。Yolo 算法检测速度较传统检测算法提高上百倍, 可满足医学诊断的高效需求。再加入形态学算法, 对目标做形状椭圆拟合, 对色值、轮廓和大小等维度做判断, 从而进一步提升检测正确率与检出率。经过自建样本数据训练以及大量实验表明, 加入传统形态学处理后, 对孢子的识别算法准确率高达 94% 以上, 检出率达 82% 以上。本文算法将传统图像处理方法与深度学习方法相结合, 应用于实际检测中, 提高了检测准确率, 检测速度是专业医务人员检测速度的两倍以上。

**关键词:** 菌种识别; Darknet-53; Yolov3 算法; 形态学算法; 椭圆拟合

## A new method for spore identification in microscope images based on morphology and Yolov3 algorithm

LI Xinming, ZHAO Lei, SHAO Baomin, WANG Dong

(School of Computer, Shandong University of Technology, Zibo Shangdong 255000, China)

**[Abstract]** Aiming at domestic problems such as identification of pathogenic bacteria under microscope and cell viability test, it is proposed to use darknet-53 network model and Yolov3 algorithm to extract suspected targets, and then use traditional morphological algorithms for screening to achieve accuracy, speed and intelligence diagnostic methods. The detection speed of the Yolo algorithm is hundreds of times higher than that of traditional detection algorithms to meet the high efficiency of medical diagnosis. then add the morphological algorithm to shape the target and judge the color value, contour, size and other dimensions to further improve the detection accuracy and detection rate. After self-built sample data training and a large number of experiments, the accuracy rate of the spore recognition algorithm after adding traditional morphological processing is as high as 94% and the detection rate is more than 82%. The algorithm in this paper combines traditional image processing methods and deep learning methods. When applied to actual detection, it improves the detection accuracy and is more than twice the detection speed of professional medical personnel.

**[Key words]** Identification of pathogenic bacteria; Darknet-53; Yolov3 algorithm; Morphological algorithms; Ellipse fitting

### 0 引言

微生态下的菌种检测是人体医疗学科的分支, 研究发现人体内分泌物中存在超过 200 种微生物, 其中不乏致病菌, 或出现某种菌类数量失衡从而致病情况。目前国内普遍采用医务人员肉眼进行显微镜下观察分析的方式, 诊断患者是否患病, 此种方式效率低下, 且易因视觉疲劳, 出现误诊、漏诊等现象。

对于目标检测问题, 当前可分为两大类, 传统检测算法诸如 Rcnm, Fast-rcnn, Faster-rcnn<sup>[1-2]</sup> 等一般分为 3 个阶段: (1) 在给定图像上选择候选区域。(2) 对区域进行特征提取。(3) 使用分类器进行分类。这类算法普遍耗时长, 磁盘空间占用大, 且算法冗余较多。Yolo (包括 Yolo, Yolov2, Yolov3<sup>[3-4]</sup>) 算

法不同于上面几种算法, 此类算法将检测问题作为回归问题处理, 即给定图像后, 直接在图像多个位置上预测出这个位置的边框以及目标类别。因此, 在检测准确率大致相同的情况下, 减少了提取候选区域等步骤的冗余计算, 大大提高了检测速度。

针对此类菌类检测问题, 熊应平、王阿川等人 (2016) 曾采用滤波, 分水岭算法对松杨栅锈菌孢子进行分割识别<sup>[5]</sup>, 但其检测背景较为简单, 且检测速度较慢。郑欣、田博、李晶晶 (2018) 等人曾采用 Yolov2 模型对宫颈细胞簇团进行智能识别<sup>[6]</sup>, 但其正确率 (判断类/实际类) 只能达到 90% 左右, 检出率 (检出数目/实际数目) 也有 83% 左右。廖欣、郑欣、邹娟 (2018) 等提出基于神经网络集成模型的宫

**作者简介:** 李鑫铭 (1993-), 男, 硕士研究生, 主要研究方向: 图像处理; 赵磊 (1964-), 男, 博士, 教授, 硕士生导师, 主要研究方向: 人工智能、自动化检测与控制; 邵宝民 (1982-), 男, 博士, 讲师, 主要研究方向: 图像分析与对象识别、计算机仿真、软件工程; 王栋 (1994-), 男, 硕士研究生, 主要研究方向: 自然语言处理。

收稿日期: 2020-03-13

哈尔滨工业大学主办 ● 学术研究与应用

颈细胞病理计算机辅助诊断方法<sup>[7]</sup>,但其时间效率差,不能满足患者实时诊断的需求。因此提出采用基于 darknet-53 网络模型的 Yolov3 算法与传统形态学相结合,实现智能识别各种致病菌类的解决办法。

由于显微镜下菌种的多样性,且背景极为复杂,干扰因素多。尤其是在 100 倍显微镜下,采集图像时,由于焦点无法兼顾多个图层,会出现景深问题,导致分层图像模糊,本文采用了一种比较快捷的景深融合方法成功解决了此问题,又在 yolov3 检测基础上加入传统图像处理方法,将整个检测过程分为两大部分,并取其交集。在经过 1 530 张图片测试后,在保证单张图片测试时间在 300 ms 左右的情况下,其准确率达 94.7%,检出率达 82.1%。经临床数据证明,资深医务人员对试片进行人工筛选时,准确率在 94% 左右。说明此算法效果可以达到实际使用的精度需求。

## 1 技术介绍

### 1.1 YOLOv3 简介

Yolov3 采用了锚点机预测来更高效的预测不同尺度与宽高比的物体边界框。YOLO3 为每种 FPN 预测特征图 (13 \* 13, 26 \* 26, 52 \* 52) 设定 3 种锚点框,总共聚类出 9 种尺寸的锚点框。锚点的设计方式使用 K-means 聚类,得到 9 个聚类中心,将其按照大小均分给 3 个尺度。尺度 1: 在基础网络之后添加一些卷积层,再输出锚点框信息。尺度 2: 从尺度 1 中的倒数第二层的卷积层上采样 (x2), 再与最后一个 16 \* 16 大小的特征图相加,再次通过多个卷积后输出锚点框信息,相比尺度 1 增大两倍。尺度 3: 与尺度 2 类似,使用了 32 \* 32 大小的特征图进行多个卷积输出锚点框信息。

距离公式(1)定义如下:

$$d(box, centroid) = 1 - IoU(box, centroid). \quad (1)$$

Yolov3 在通过训练后,对每个预选框预测 4 个坐标值 ( $t_x, t_y, t_w, t_h$ ) (其中  $t_x, t_y$  是预测的坐标偏移值,  $t_w, t_h$  是尺度缩放), 其预测过程与 Faster-rcnn 类似。可通过公式(2)~公式(5)计算得到:

$$t_x = (G_x - p_x) / p_w, \quad (2)$$

$$t_y = (G_y - p_y) / p_h, \quad (3)$$

$$t_w = \log(G_w - p_w), \quad (4)$$

$$t_h = \log(G_h - p_h). \quad (5)$$

其中,  $G_x, G_y, G_w, G_h$  为是目标真实框在 feature map 的 4 个坐标。  $P_w, P_h$  是预设的锚点框在 feature map 上的宽和高,预选框预测流程如图 1 所示。

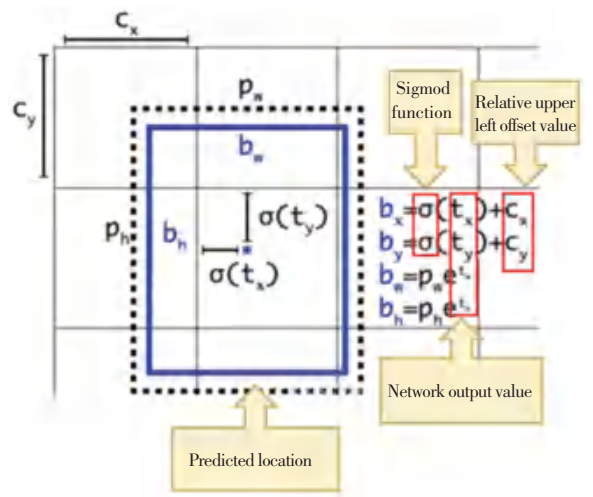


图 1 预选框预测流程图

Fig. 1 Pre-selection box prediction flowchart

对于预测预选框的 Cell (一幅图划分成 S \* S 个网格 cell), 可以对边框按图 1 所示进行预测。

预测边框时  $c_x, c_y$  为相对于左上角的位置, 每个小格的长度为 1, 图 1 中此时  $c_x = 1, c_y = 1$ 。  $t_x$  和  $t_y$  分别经过 sigmoid 函数处理, 与  $c_x, c_y$  相加后得到预选框中心点  $b_x, b_y$ 。  $t_w, t_h$  分别与  $p_w, p_h$  作用后, 最终得到预选框的宽和高  $b_w, b_h$ 。 sigmoid 函数公式(6):

$$\theta(x) = \frac{1}{1 + e^{-x}}. \quad (6)$$

除去预选框位置的四个值外, 每个网格还会预测一个框内有无待测目标的置信度, 以及 N 个待检测类的特定置信度。

预测有无待测置信度定义为公式(7):

$$\Pr(Class_i) * IOU_{pred}^{truth}, \quad (7)$$

其中,  $\Pr(Object) \in \{0, 1\}$ 。

此时, 将每个网格预测的种类信息和预选框区域预测的置信度信息相乘, 得到每个预选框区域的特定类型置信度, 公式(8):

$$\Pr(Class_i | Object) * \Pr(Object) * IOU_{pred}^{truth} = \Pr(Class_i) * IOU_{pred}^{truth}. \quad (8)$$

等式左边第一项是每个网格预测的类别信息, 第二、三项是每个预选框预测的置信度。这个乘积即代表了预测的区域属于某一类的概率, 也有该区域准确度的信息。

关于损失函数, Yolov3 使用均方和差作为 loss 函数, 由三部分组成: 坐标误差、IOU 误差 (预选框与实际框的交并比)、和分类误差, 公式(9)。

$$loss = \sum_{i=0}^{s^2} coordErr + iouErr + clsErr. \quad (9)$$

对于每种 loss 的贡献率, Yolo 给 coordErr 设置

权重  $\lambda_{coord} = 5$ 。在计算 IOU 误差时, YOLO 使用  $\lambda_{noobj} = 0.5$  以修正 IOUErr。采用了将物体大小的信息项(宽和高)进行求平方根的方法来改进对于大物体误差对检测的影响应小于小物体误差对检测的影响问题。

综上, Yolo 在训练过程中 Loss 计算如公式(10)所示:

$$\begin{aligned} \text{loss} = & \lambda_{coord} \sum_{i=0}^{s^2} \sum_{i=0}^B I_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] + \\ & \lambda_{coord} \sum_{i=0}^{s^2} \sum_{i=0}^B I_{ij}^{obj} [(\sqrt{x_i} - \sqrt{\hat{x}_i})^2 + \\ & (\sqrt{y_i} - \sqrt{\hat{y}_i})^2] + \sum_{i=0}^{s^2} \sum_{i=0}^B I_{ij}^{obj} (C_i - \hat{C}_i)^2 + \\ & \lambda_{coord} \sum_{i=0}^{s^2} \sum_{i=0}^B I_{ij}^{noobj} (C_i - \hat{C}_i)^2 + \\ & \sum_{i=0}^{s^2} I_{ij}^{noobj} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2. \quad (10) \end{aligned}$$

注:公式中 1-2 行为对预测中心以及预测边框宽高做损失,第 3-4 行分别为预测类别,第 5 行为对预测置信度做损失。

## 1.2 形态学筛选

在检测问题中,传统图像处理技术应用较少。目前多种深度学习检测算法更多注重于图像深层特征的提取与分类,而忽略了图像直观的内容。因此,在利用深度学习检测算法的同时,辅以形态学的条件进一步筛选,可以有效提高检测精度。传统图像处理方法具有运算速度快,图像操作较为直观等优点,尤为适合医学范围内的实时检测,其缺点是精度较差。而形态学与 YOLOv3 算法相结合,在保留时间效率的同时极大的提高了检测精度。针对孢子这种致病菌的检测,本文提出以加入滤波,灰度化,自适应阈值划分,轮廓提取等传统图像方法<sup>[10]</sup>,并针对孢子形状对其进行椭圆拟合,来对 YOLOv3 算法筛选出来的疑似区域进行再确定。

根据得到的图片特性,首先进行灰度化,自适应阈值划分,图像锐化等操作,采用 Canny 算子对其进行轮廓查找。针对孢子的形状,采用了基于最小二乘法椭圆拟合<sup>[11-12]</sup>的办法来对孢子进行准确检测。

已知椭圆方程

$$x^2 + Axy + By^2 + Cx + Dy + E = 0. \quad (11)$$

对于公式(11),方程中有 5 个未知数分别为 A, B, C, D, E, 则最少需要采样点为 5,对采样点进行椭圆非标准方程拟合可得

$$\min \| x^2 + Axy + By^2 + Cx + Dy + E \| = 0. \quad (12)$$

即:

$$F(A, B, C, D, E) = \sum_{i=1}^N (x_i^2 + Ax_i y_i + By_i^2 + Cx_i + Dy_i + E)^2. \quad (13)$$

其中, N 表示采样点个数。经过多次测试并结合孢子图像在图片中的像素大小,其轮廓上的拟合点定为 35-38 时,拟合效果最好。

## 2 实验与测试

### 2.1 图像集及其预处理

本文所用训练图像数据集与测试集均为患者真实病例试片在 100 倍显微镜下采集的,含有致病菌的视野,采集像素大小为 1600 \* 1200(后归一化为 608 \* 608)。得到清晰图像如图 2 所示。

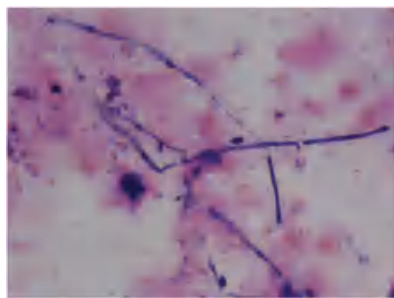


图 2 清晰孢子菌丝图像

Fig. 2 Clear spore hyphae image

其中,深色椭圆状细菌为孢子,深色长条形细菌为菌丝。得到清晰的图像后,再由专业医务人员对图像进行标注。

标注好的图像如图 3 所示:

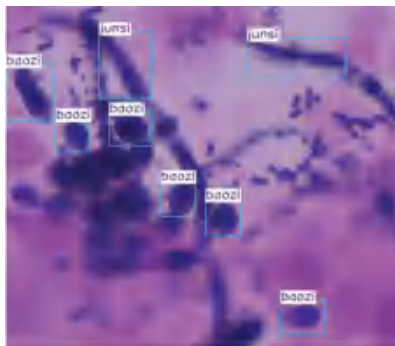


图 3 孢子菌丝标注图

Fig. 3 Spore hyphae labeling diagram

### 2.2 实验

实验基本流程为:选取部分待测图像(目前为 50 000 张),在待测图像中将识别目标做好标注,将其组成训练集放入网络训练,得到模型。再将要识别的图像(图像大小为 1600 \* 1200,共 1 530 张)放入测试程序,调用训练好的模型进行测试,得到模型初筛结果,并在原图中标出。对该疑似区域进行形



态学的判断,最终通过取交集的方法得到较为准确的识别结果。实验流程如图 4 所示。

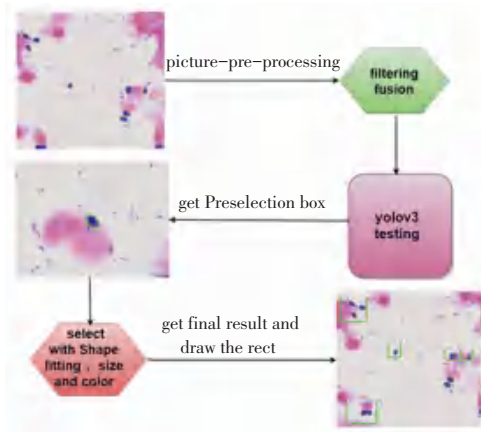


图 4 实验流程

Fig. 4 Experimental flow

(1) 基于 darknet 框架下 yolov3 模型的首筛。测试过程: 经过模型训练后, 将待测图像导入进模型网络, 得到每个区域的特定类型置信度, 设置阈值, 滤掉得分低的区域, 对保留的区域进行 NMS(Non-Maximum Suppression, 非极大值抑制) 处理<sup>[13]</sup>, 得到 Yolov3 的检测结果。

(2) 基于图像形态学处理。基于图像形态学处理与识别的流程如图 5 所示。

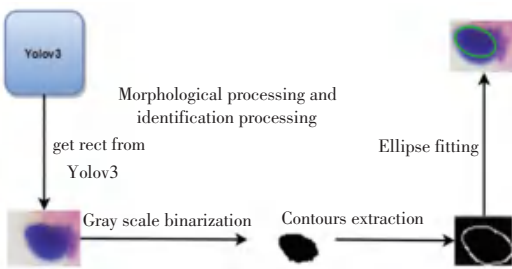


图 5 形态学处理与识别流程

Fig. 5 Morphological processing and identification flow

得到清晰的 Yolov3 算法筛选过的疑似区域之后, 对图像进行边缘增强及锐化。再根据孢子的色值做灰度化以及二值化分割, 将孢子与浅色背景分割开来, 通过 Canny 算子对其进行轮廓提取。得到清晰轮廓后, 在轮廓上随机取点进行椭圆拟合, 若拟合椭圆区域与二值化后孢子所在拟合椭圆的区域重叠部分超过 95% 以上, 则可以大致判断此区域内的形状大致为椭圆。即该物体符合孢子的形状特点, 可将其判定为孢子。图 6 为杆菌做拟合情况, 图 7 为孢子拟合情况。

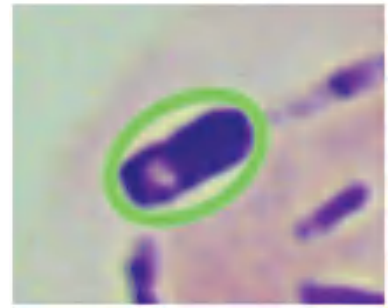


图 6 杆菌椭圆拟合

Fig. 6 Bacterial ellipse fitting

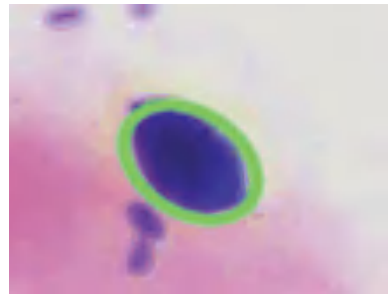


图 7 孢子椭圆拟合

Fig. 7 Spore ellipse fitting

由图 6, 图 7 可以看到, 杆菌形状特征并不满足筛选条件, 可以轻松被筛选掉。而细胞核较孢子大小要大, 孢子在 1600 \* 1200 大小的图片上时, 一般最小为 45 \* 45, 且其长宽比近似 1 : 1, 而细胞核大小一般最大为 20 \* 30, 长宽比接近 1 : 1.5。因此, 此时将大小阈值设为 30 \* 30, 长宽比 1 : 1.3 时可以通过轮廓的旋转矩形大小有效筛选掉细胞核与杆菌。

### 2.3 测试

为验证算法性能, 只以孢子识别为例, 在上文自制测试集中进行了测试(置信度阈值设置为 75%, 高于此阈值时标出)。包括本综合算法的时间性能, 准确率(判断正确总数/进行判断总数)与单纯形态学算法, 深度学习 Faster-RCNN, SSD<sup>[14]</sup>, YOLOv2 等算法在 GPU 运行环境性能对比如表 1 所示。

表 1 孢子识别方法性能对比表

Tab. 1 Spore identification method performance comparison table

Method	FpsonGPU	Accuracy/%	RecallRate/%
Imagemorphology	148.22	60.91	65.77
Fast-remn	0.61	69.33	81.33
Faster-remn	8.36	87.31	85.11
SSD300	31.33	82.21	85.91
SSD512	18.77	84.19	89.71
YoloV2	94.43	72.54	78.14
YoloV3	41.60	88.93	85.37
<b>New Algorithm</b>	<b>37.53</b>	<b>94.70</b>	<b>82.12</b>

注: 表中第 1-4 列分别为方法名, GPU 环境测试速度, 准确率, 召回率。

由上表1可见,单纯采用图像形态学进行孢子检测时,在GPU环境下检测速度可达 $148\text{ p/s}^{-1}$ ,但其准确率非常差,难以适应微生态中复杂的背景图像,在1530张测试图片中只有69.91%的准确率,召回率(检测出目标数目/所含目标总数)只有65.77%。而基于深度学习的Fast-rcnn与Faster-rcnn,SSD,Yolov3等算法采用卷积加权提取特征,并用不同分类器对其进行分类识别的方法较传统识别方法提升较大,具有一定可行性。但Fast-rcnn,Faster-rcnn等算法作为传统two-stage检测算法,在GPU环境下检测速度分别只有 $0.61\text{ p/s}^{-1}$ , $8.36\text{ p/s}^{-1}$ 。关于准确率,Faster-rcnn虽然可以达到87%,较Fast-rcnn提升较大,但是也只有87.31%的准确率与85.11%的召回率。而作为检测速度较快的SSD与Yolov2两种算法,其速度分别可达 $31.33\text{ p/s}^{-1}$ 与 $94.43\text{ p/s}^{-1}$ 。虽然这两种算法检测速度较快,但其精度仍较实际需求差距很大。Yolov3在检测速度与精度上都较上述算法有很大提升,但并不能满足在实际临床医学中的精度与效率需求。图8为采用Yolov2,Faster-rcnn,与本文算法对同一张图的识别情况对比。

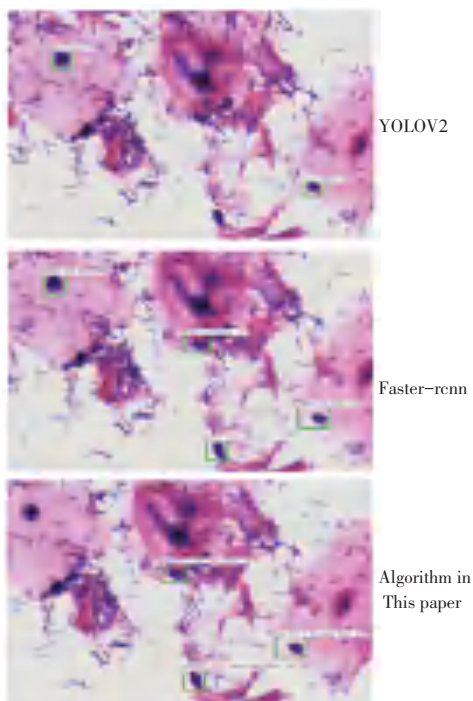


图8 算法效果对比图

Fig. 8 Comparison of algorithm effects

由图8可以看出:精度方面,Yolov2无论检出率与召回率都远不及后面两种算法,而Faster-rcnn在实际测试时虽然检测精度稍高,但是存在将细胞核误检的情况。而本文算法则不存在此类情况,鲁棒

性很强。因此,本文提出的Yolov3与传统图像处理方式相结合的方法,不仅检测速度与Yolov3相近,其准确率更是大大提升,完全可以达到实际临床检测的需求。

### 3 结束语

本文提出了一种将传统图像处理方法与深度学习算法Yolov3相结合,应用于临床医学检验微生态下孢子等致病菌的方法。针对孢子等致病菌识别的问题,采用了Darknet-53作为特征提取网络,并利用Yolov3算法进行疑似区域初筛,提取疑似区域,再利用孢子菌丝等菌类的形态学特点进一步筛选,从而去除杆菌堆簇、细胞核、杂质等造成的误检。针对孢子识别精度可达94%以上,召回率82%以上。对于数据的前期处理阶段,针对显微镜图像出现的景深问题,提出了一种简单的分割拼接方法,从而得到了高质量的训练数据集,使模型训练精度大大提升。但是此算法还存在一定的缺陷,在召回率上,跟Faster-rcnn与Yolov3有些差距,主要是因为筛选过程中条件增多后一些不规则目标被筛除了,识别出来的多为典型目标,在此问题上还需要继续深入研究,目前暂定两种解决方法来提高召回率。第一种就是不断增加训练数据,目前50000条训练数据还是偏少,需要标注人员不断进行标注工作来扩充;第二种就是针对待测图像修改网络框架,适当减少或增加特征提取网络层数以及修改相关参数,通过不断实验继续改进识别结果。

### 参考文献

- [1] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [C]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017: 1-14.
- [2] 刘博,王胜正,赵建森,等.基于Darknet网络和YoloV3算法的船舶跟踪识别[J].计算机应用,2019,39(6):1663-1668.
- [3] Redmon J, Farhadi A. Yolo9000: Better, Faster, Stronger [C]//Conference on Computer Vision and Pattern Recognition. Hawaii, USA: IEEE, 2018: 3-12.
- [4] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real time object detection [C]//Computer Vision and Pattern Recognition. Las Vegas, USA: IEEE, 2015
- [5] 熊应平,王阿川.松杨栅锈菌担孢子显微图像识别方法[J].东北林业大学学报.2016.07:116-119
- [6] 郑欣,田博.基于YOLO模型的宫颈细胞团簇智能识别方法[J].液晶与显示.2018.11:347-355
- [7] 廖欣,郑欣,邹娟,等.基于神经网络集成模型的宫颈细胞病理计算机辅助诊断方法[J].液晶与显示.2018.04:965-971
- [8] 黄震.外周血白细胞分割和识别的算法研究[D].中国计量学,2015.
- [9] 戴伟聪,金龙旭,李国宁,等.遥感图像中飞机的改进YoloV3实时检测算法[J].光电工程,2018,45(12):81-89.