

基于改进 YOLOv3 的瞳孔屈光度检测方法

李岳毅¹, 丁红昌^{1,2}, 张雷^{3,4}, 赵长福¹, 张士博¹, 王艾嘉¹

(1. 长春理工大学 机电工程学院, 吉林 长春 130022; 2. 长春理工大学 重庆研究院, 重庆 401135;

3. 郑州轻工业大学 电气信息工程学院, 河南 郑州 450002; 4. 中国烟草总公司 郑州烟草研究院, 河南 郑州 450001)

摘要: 针对瞳孔区域屈光度识别准确率低、检测效率低等问题, 本文提出一种基于改进 YOLOv3 深度神经网络的瞳孔图像检测算法。首先构建用于提取瞳孔主特征的二分类检测网络 YOLOv3-base, 强化对瞳孔特征的学习能力。然后通过迁移学习, 将训练模型参数迁移至 YOLOv3-DPDC (Deep Pupil Diopter Classify), 降低样本数据分布不均衡造成的模型训练困难以及检测性能差的难题, 最后采用 Fine-tuning 调参快速训练 YOLOv3 多分类网络, 实现了对瞳孔屈光度快速检测。通过采集的 1200 张红外瞳孔图像进行实验测试, 结果表明本文算法屈光度检测准确率达 91.6%, 检测速度可达 45 fps, 优于使用 Faster R-CNN 进行屈光度检测的方法。

关键词: 瞳孔屈光度检测; 深度学习; YOLOv3 网络; 多尺度特征; 机器视觉

中图分类号: TP391 文献标识码: A 文章编号: 1001-8891(2022)07-0702-07

Pupil Diopter Detection Approach Based on Improved YOLOv3

LI Yueyi¹, DING Hongchang^{1,2}, ZHANG Lei^{3,4}, ZHAO Changfu¹, ZHANG Shibo¹, WANG Ajia¹

(1. College of Mechanical and Electrical Engineering, Changchun University of Science and Technology, Changchun 130022, China;

2. Chongqing Research Institute, Changchun University of Science and Technology, Chongqing 401135, China;

3. School of Electrical Information Engineering, Zhengzhou University of Light Industry, Zhengzhou 450002, China;

4. Zhengzhou Tobacco Research Institute China National Tobacco Corporation, Zhengzhou 450001, China)

Abstract: To address the problems of low diopter recognition accuracy and low detection efficiency in the pupil area, this paper proposes a pupil image detection algorithm based on an improved YOLOv3 deep neural network. First, a two-class detection network YOLOv3 base for extracting the main features of the pupil is constructed to strengthen the learning ability of the pupil characteristics. Subsequently, through migration learning, the training model parameters are migrated to YOLOv3-DPDC to reduce the difficulty of model training and poor detection performance caused by the uneven distribution of sample data. Finally, fine-tuning is used to quickly train the YOLOv3 multi-classification network to achieve accurate pupil diopter detection. An experimental test was performed using the 1200 collected infrared pupil images. The results show that the average accuracy of diopter detection using this algorithm is as high as 91.6%, and the detection speed can reach 45 fps; these values are significantly better than those obtained using Faster R-CNN for diopter detection.

Key words: pupil diopter detection, deep learning, YOLOv3, multi-scale features, machine vision

0 引言

随着移动互联网、5G 技术不断发展, 手机、平板等电子产品在日常生活中普及, 视力衰退已成为广

大青少年面临的常见问题, 我国青少年近视呈现高发、低龄化趋势, 严重影响儿童身心健康。因此, 研究瞳孔屈光度检测方法对眼睛近视及早治疗与保护具有重要意义。

收稿日期: 2021-08-27; 修订日期: 2021-11-29。

作者简介: 李岳毅(1994-), 男, 硕士研究生, 主要从事图像处理、目标检测和计算机视觉方面的研究。E-mail: lyy642668743@163.com。

通信作者: 丁红昌(1980-), 男, 博士, 副教授, 博士生导师, 吉林省第七批拔尖创新人才。主要从事在线检测、模式识别和机器视觉方面的研究。

E-mail: custjdgc@163.com。

基金项目: 吉林省科技发展计划重点研发项目(20200401117GX); 河南省科技攻关计划(212102210155)。

目前,瞳孔屈光度的测量方法分为主观和客观两大类,在临床实践中普遍采用主观方法,依赖医生通过被测者在瞳孔计或者瞳孔对照表上的测试结果,获得屈光度数值。然而该方法对患者配合度要求高,检测流程长,且存在主观较强、效率低等问题。另一类方法主要以数字摄像为主,其基本思想是根据红外瞳孔图像中光亮区域的大小作为屈光度测量标准。

薛峰等^[1]基于采集到的红外图像,提出利用曲波变换和最小二乘法测量瞳孔屈光度,有效解决了图像光亮区域的渐晕现象。但受图像采集设备、拍摄角度、光照条件和环境变化等因素影响,所采集的图像具有不同的质量,决定了图像处理的难易程度不同。胡志轩采用连续的 Otsu 算法^[2]分割图像,虽然硬件上有所改进,但该方法占用内存多且耗时长,对硬件处理设备的要求较高。

近年来,深度学习技术在图像分类、目标检测、图像语义分割等领域取得了突破性进展。在医学图像领域,研究基于深度学习的目标检测方法是突破制约瞳孔检测的必然趋势。目前基于深度学习目标检测方法^[3-7]主要包括 Faster R-CNN、YOLO (You only look once)。文献[8]提出一种利用形态学对采集到的瞳孔图像进行清晰化处理,先利用 Faster R-CNN 算法对瞳孔图像进行识别,最后利用阈值分割,根据采集到的数据得出屈光度的数值。Faster R-CNN 网络采用 Two-stage 架构,相比于 One-stage 架构的 YOLOv3 目标检测模型漏检率低,该方法分两步得出瞳孔屈光度数值,识别效率低,难以进行快速检测。

综上所述,为了提高算法在瞳孔屈光度检测方面的适用性和准确性,本文在 YOLOv3 算法的基础上进行改进,提出两阶段训练框架,根据瞳孔屈光度数值将数据集分成 3 类,采用滑动窗口算法扩充瞳孔数据集,并用迁移学习训练网络来解决训练效率低、数据集小易拟合等问题。最后利用真实红外瞳孔图像进行实验测试,以验证本文算法准确性与实时性。

1 YOLOv3 目标检测原理

2018 年 Redmon 等在 YOLOv2 模型基础上,融合特征金字塔、残差结构、多尺度预测模块,设计了一种新的 YOLOv3 网络目标检测算法,在保持检测速度同时提升了目标检测的准确率^[9]。其基本步骤为,首先输入图像被划分成 $S \times S$ 网格,当检测对象处于某一网格时,负责对该对象进行预测,同时产生 B 个边界框 (bounding box),以及给出边界框的预测信息(x, y, w, h)和置信度 (confidence)。其中边界框中包含目标的概率 P_r 以及边界框的准确度 IOU 共同决

定置信度的大小,计算公式为:

$$\text{confidence} = P_r(\text{object}) * \text{IOU}_{\text{pred}}^{\text{truth}} \quad (1)$$

式中: $P_r(\text{object})$ 代表网格中包含某个检测目标的概率,是 $\text{IOU}_{\text{pred}}^{\text{truth}}$ 预测框与真实标注框之间的比值,当 $\text{IOU}=1$ 时,说明两者完全重合。边界框中目标属于某个类别的置信度为:

$$h = P_r(\text{Class}_i | \text{object}) * P_r(\text{object}) * \text{IOU}_{\text{pred}}^{\text{truth}} \quad (2)$$

式中: i 代表检测的类别数。

综上所述,瞳孔屈光度等级识别本质上属于一类目标检测问题。整体模型框架如图 1 所示,本文提出的瞳孔图像检测模型在第一阶段以 YOLOv3 网络模型为基础,保持其网络结构不变,采用迁移学习的方法,对瞳孔目标和背景进行分离训练,增强对瞳孔特征学习的能力。第二阶段以 YOLOv3-base 网络模型为基础,保持其特征提取网络模块参数不变,重新构建 FPN 网络,将原有的 3 个检测尺度提升为 4 个,充分利用浅层高分辨率和深层高语义信息,提高对 3 个不同屈光度等级的瞳孔定位识别能力。

2 基于 YOLOv3 的红外瞳孔识别算法

针对瞳孔屈光度快速识别难题,本文提出一种基于改进 YOLOv3 网络的瞳孔图像检测算法,主要包括主特征学习和模型迁移两个阶段。
① 建立 YOLOv3-base 网络提取瞳孔特征;
② 采用迁移学习方法,构建 YOLOv3-DPDC 多分类网络。

2.1 构建 YOLOv3-base 网络

首先,考虑数据集中 3 种不同屈光度等级的瞳孔数量分布不均匀以及相似性的问题,设计出基于 YOLOv3 的二分类检测网络,如图 1 所示。将数据集中的瞳孔图像都归为一类 “Eye”,然后利用 YOLOv3-base 网络对瞳孔目标和背景进行训练,该方法能够避免瞳孔图像数据量的缺乏以及 3 种等级数据集不均衡问题(数据集中的 R_1 等级的瞳孔数量有 120 个,而 R_3 等级的瞳孔数量有 780 个)所产生的训练问题,可以加强模型对瞳孔图像的识别能力。

为了提升模型的检测精度以及减少训练时间,第一阶段采用迁移学习^[10]的方法,训练过程如图 2 所示。采用一个在 ImageNet 上已经使用 VOC2007 训练完毕的 YOLOv3 模型作为预训练模型。

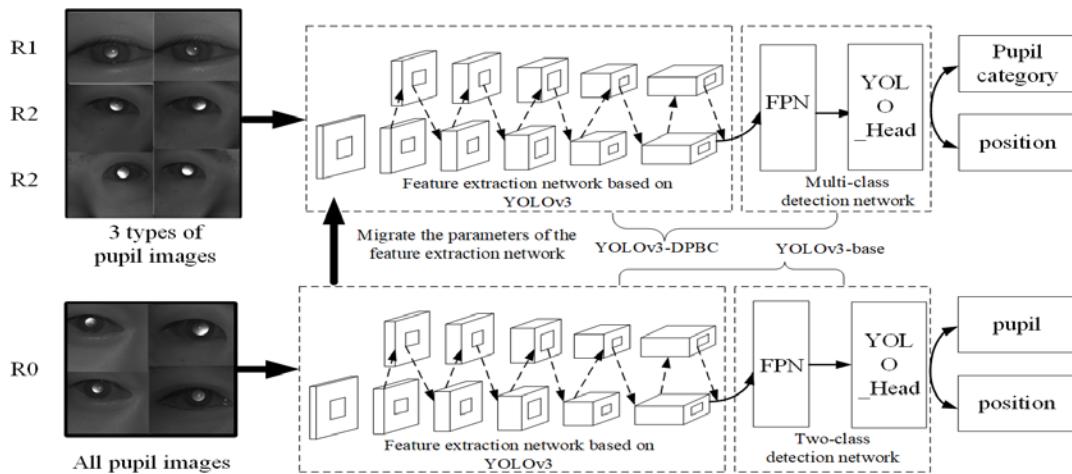


图1 本文提出的瞳孔识别算法示意图

Fig.1 Schematic diagram of the infrared pupil recognition algorithm proposed in this article

在迁移训练时,第一步先冻结前 a 个残差块,利用DarkNet-53网络对图像特征进行提取,使用大学习率对后面的参数进行训练;第二步将前面的残差块解冻投入训练,对参数进行微调变换学习率。训练结束后,提前将训练得到的权重文件进行替换,然后对瞳孔目标进行检测。由于第一阶段采用的是二分类模型,识别结果只有瞳孔和背景两个相互独立的类别,所以最后采用Soft-max函数^[11-12],通过输出张量输出红外瞳孔图像和背景。

2.2 构建 YOLOv3-DPDC 网络

首先,在YOLOv3-base的基础上构建YOLOv3-DPDC网络框架,第二阶段包括YOLOv3特征提取网络和多分类检测网络,由于在第一阶段训练出来的模型对瞳孔已有较好的提取能力,为了更好地识别三类瞳孔细粒度图像以及充分利用浅层高分辨率和深层

高语义信息,第二阶段的多分类检测网络在原有的基础上增加一个尺度特征为 104×104 的特征层,将低阶特征和高阶特征有机地融合起来,通过Logistic多标签分类器对3个等级的屈光度图像进行分类优化,形成完整的YOLOv3-DPDC模型。

整体的网络模型如图3所示,首先对图像进行归一化,将 416×416 大小的瞳孔图像作为输入。然后利用Darknet-53网络进行瞳孔特征提取,交替使用 3×3 和 1×1 的卷积以及上采样操作,最终得到 13×13 、 26×26 、 52×52 、 104×104 四个检测尺度。

本文模型构建了4个检测尺度,在没有明显增加网络深度的前提下,增强了网络结构的表征能力,通过多尺度特征分析解码,可以更好地适应对红外瞳孔图像分类识别能力,从而提高模型识别的准确率。

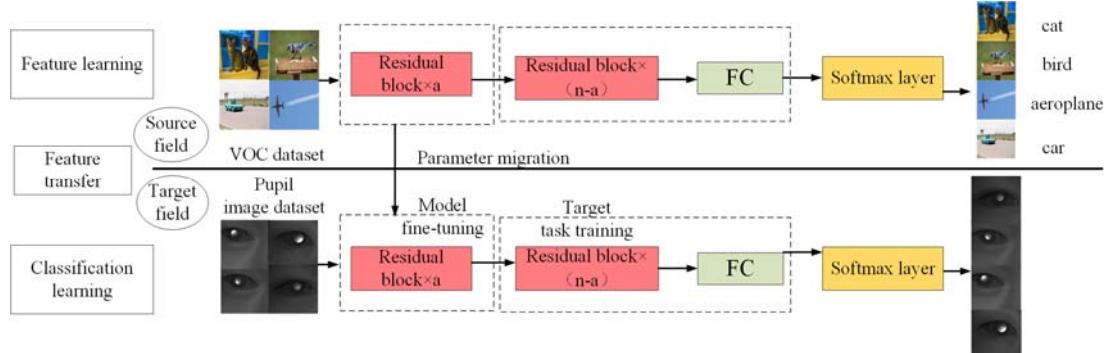


图2 第一阶段迁移学习处理过程

Fig.2 The first stage transfer learning process

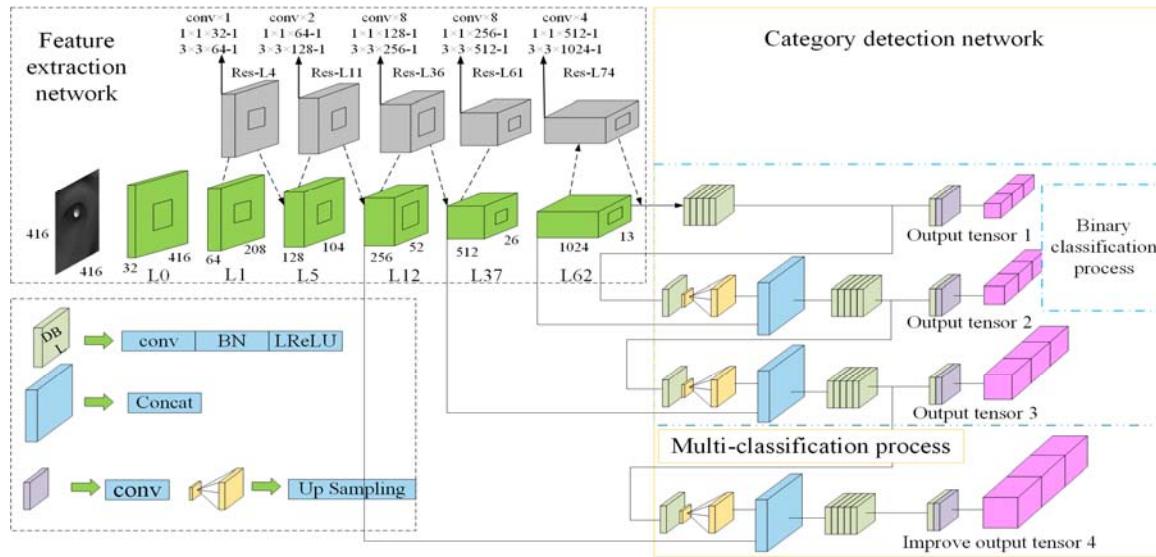


图3 本文所提算法的网络模型与参数说明

Fig.3 The network model and parameter description of the proposed algorithm in this article

2.3 损失函数分析

损失函数一般作为评估最终模型真实值与预测值误差的大小，最终会影响模型的收敛效果。为了提升模型的检测精度，减少候选框与真实框之间的误差以及类别间的误差，本文算法中损失函数由边界框损失、置信度损失和分类损失3部分组成。计算公式为：

$$\begin{aligned} \text{Loss} = & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{\text{obj}} [(x_i - \hat{x}_i^j)^2 + (y_i - \hat{y}_i^j)^2] + \\ & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{\text{obj}} [(\sqrt{w_i^j} - \sqrt{\hat{w}_i^j})^2 + (\sqrt{h_i^j} - \sqrt{\hat{h}_i^j})^2] - \\ & \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{\text{obj}} [\hat{C}_i^j \log(C_i^j) + (1 - \hat{C}_i^j) \log(1 - C_i^j)] - \\ & \lambda_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{i,j}^{\text{noobj}} [\hat{C}_i^j \log(C_i^j) + (1 - \hat{C}_i^j) \log(1 - C_i^j)] - \\ & \sum_{i=0}^{S^2} I_{i,j}^{\text{obj}} \sum_{c \in \text{classes}} [\hat{p}_i^j \log(p_i^j) + (1 - \hat{p}_i^j) \log(1 - p_i^j)] \end{aligned} \quad (3)$$

式中： S^2 表示 13×13 、 26×26 以及 52×52 三种网格的大小； B 表示预测框的数量； $I_{i,j}^{\text{obj}}$ 表示预测框在 i 、 j 点出现待检测目标的概率，取值为 0 或 1， $I_{i,j}^{\text{noobj}}$ 表示预测框在 i,j 点不出现待检测目标的概率。 λ_{coord} 表示边界框损失的权重， λ_{noobj} 表示置信度损失的权重。

3 基于人工扩增的数据集处理

基于深度学习的图像识别算法，往往需要大量带类别标签的瞳孔图像作为训练集。针对该问题，在采集 1200 张瞳孔图像基础上，本文提出一种基于滑动窗口算法扩充瞳孔数据。

首先，采集到的瞳孔图像归一化为 1024×1024 分辨率，接着将原图像用 32×32 像素的网格进行划分，然后用 $L \times H$ 固定大小的窗口在瞳孔图像上进行移动，窗口每移动到一个位置，就将其覆盖下的瞳孔图像当作一个 ROI 感兴趣区域。其中， $L \times H = 416$ 像素，移动窗口的右上角坐标 $A = (\text{img}R_x, \text{img}R_y)$ ，左下角坐标为 $B = (\text{img}L_x, \text{img}L_y)$ 计算公式如式(4)所示：

$$\begin{cases} \text{img}R_x = i \times L + L \\ \text{img}R_y = j \times H \\ \text{img}L_x = i \times L \\ \text{img}L_y = j \times H + H \end{cases} \quad (4)$$

式中： i 和 j 的取值范围为 $(0, 1, 2, 3, \dots, (\text{dstImg}_L - 416)/32)$ 和 $(0, 1, 2, 3, \dots, (\text{dstImg}_H - 416)/32)$ ， dstImg_L 和 dstImg_H 分别为被窗口滑动的瞳孔图像的宽和高，本文 $\text{dstImg}_L = \text{dstImg}_H = 1024$ 像素，最终整个数据集扩充的步骤如图 4 所示。

将采集的 1200 张红外瞳孔图像划分成 A 和 B 集合，然后将 1000 图像放入集合 A 作为训练集，用于模型训练，剩余 200 张图像放入集合 B 作测试集。集合 A 中的图像按屈光度的不同分成 3 类，屈光度范围 $0^\circ \sim 150^\circ$ 为 R_1 ， $150^\circ \sim 350^\circ$ 为 R_2 ， $350^\circ \sim 600^\circ$ 为 R_3 。对 A 集合中的瞳孔图像采用人工扩增数据集的方式进行扩增，扩增完之后手动挑选出一定数量且符合检测条件的图像放入集合 C 中，并且按照一定的比例划分为训练集和验证集，划分结果如表 1 所示。

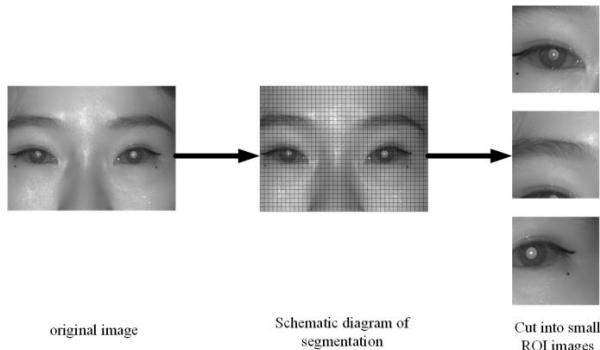


图4 瞳孔图像数据集人工扩增示意图

Fig.4 Schematic diagram of artificial augmentation of pupil image data set

表1 瞳孔数据集C的划分

Table 1 The division of pupil data set C

Diopter grade	Initial number of samples	Number of samples after amplification	Training samples	Validation samples
R_1	120	2000	1600	400
R_1	780	5000	4000	1000
R_3	300	4000	3200	800

4 实验与结果分析

4.1 实验说明

本文实验采用真实的红外瞳孔图像。实验软硬件配置为: AMD Ryzen 7 4800H, NVIDIA GeForce RTX 2060 GPU, Pytorch 版本为 1.2, CUDA 版本 10.0, CUDNN 版本 7.4.1.5, 内存为 16 GB, Python 3.6。

4.2 训练过程与分析

借鉴原始 YOLOv3 模型的网络参数, 根据本文瞳孔图像进行调参后, 第一阶段模型训练参数如表 2 所示。

表2 实验参数设置

Table 2 Experimental parameter settings

Parameter	Parameter value
Training image size/pixel	416×416
Number of iterations	9000
GPU quantity	1
IOU	0.6
The initial learning rate of the previous stage	0.001
The initial learning rate of the latter stage	0.0001

采用修改后的参数进行训练, 一共进行了 9000 次迭代训练, Loss 变化曲线如图 5 所示。

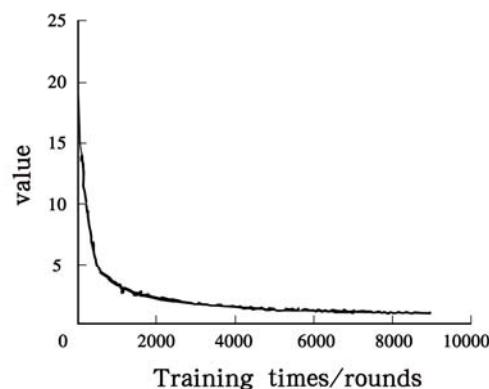


图5 Loss 值变化示意图

Fig.5 Schematic diagram of Loss value change

由图 5 可知, 前 2500 次迭代损失函数曲线快速衰减, 在 8000 次迭代以后, 损失函数曲线走势基本稳定到最低位置, 其中每迭代 100 次就输出一个模型, 共共得到 90 个模型。

4.3 分析与讨论

为验证模型的检测精度和稳定性, 本文采用 Precision (精确率) 和 Recall (召回率) 两个评价指标衡量瞳孔检测的效果。Precision 和 Recall 定义为:

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

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

准确率表示某一类别预测目标中预测正确占总正确样本的比例, 召回率表示预测目标正确占总预测样本的比例。TP 为真正例; FP 为假正例; FN 为假负例。

为了证明本文 YOLOv3 识别算法有效性, 实验选取 3 种经典模型进行对比测试, 包括 Faster R-CNN^[13]、SSD (Single Shot MultiBox Detector)^[14] 以及 RetinaNet。进行对比试验的模型都采用了同样的数据集进行训练。从测试集 C 中抽取 R_1 、 R_2 、 R_3 三类样本各 20 张用于测试。采用 Precision(精确率) 和 Recall(召回率) 作为模型的参考标准, 为了进一步对目标检测算法进行评估, 还引入 F_1 -score 指标, 定义为:

$$F_1\text{-score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}} \quad (7)$$

F_1 -score 是对精确率和召回率的调和平均, 可以更准确地反映模型的好坏。实验结果如表 3 所示。

由表 3 结果可知, 本文 YOLOv3 算法的 F_1 -score 的结果要高于 SSD、RetinaNet 和 Faster R-CNN, 在瞳孔图像检测时间方面, 由于本文提出的改进 YOLOv3 网络结构加深, 速度要略慢于 SSD。但是要

快于 Faster R-CNN。在表 4 数据中, Accuracy 是用传统的曲波变换识别出瞳孔的正确率。

表 3 不同算法数据对比

Table 3 Data of different algorithms

Algorithm model	Precision/%	Recall/%	Time/ms	F_1 -score/%
SSD	82.5	84.7	20.3	84.7
RetinaNet	84.7	82.4	32.6	82.4
Faster R-CNN	86.8	88.9	109.6	88.9
Our algorithm	91.6	90.5	22.3	90.5

表 4 传统算法数据统计

Table 4 Statistics of traditional algorithms

Algorithm model	Accuracy/%	Time/ms
Traditional algorithm	71.5	130

本文提出的模型在识别出瞳孔目标的同时还识别出了屈光度的数值范围, 相对于文献[8]先识别出瞳孔再进行瞳孔屈光度数据识别而言, 可以更快速地完成瞳孔屈光度识别任务, 优于其他对比实验。

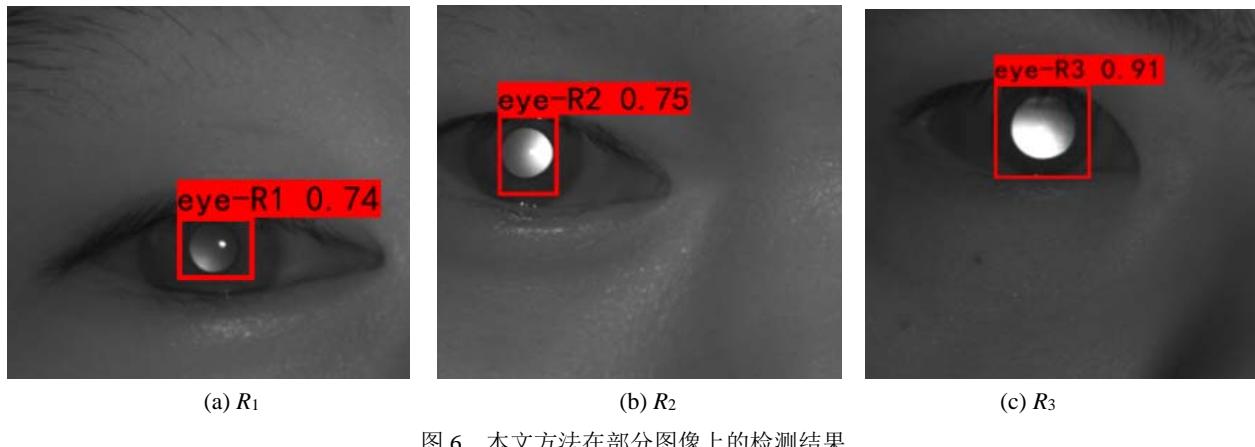


图 6 本文方法在部分图像上的检测结果

Fig.6 The detection results of the method in this paper on some images

5 结论

针对瞳孔屈光度识别准确率低、检测效率低的难题, 本文提出一种基于改进YOLOv3瞳孔屈光度检测算法, 所提算法使用两个阶段强化学习, 增加了基层卷积层对特征的提取能力, 提升了网络检测的准确率, 利用迁移学习的思想, 在任务中有效解决了数据样本不平衡与匮乏的问题, 在对比实验中取得了较好的检测结果。最终实验结果表明本文方法能实现屈光度高效率检测, 对于三类屈光度的瞳孔图像预测准确

为了进一步说明本文提出的改进算法的识别效果, 根据这 3 类样本进行实验, 生成的混淆矩阵如表 5 所示, 其中第一行代表真实值, 第一列代表预测值。由表 4 可知, 各类算法精确率按顺序分别为 95%, 85% 和 95%, 混淆矩阵中 R_1 和 R_3 种类的识别率要高于 R_2 的识别率, 分析原因一是由于屈光度数值过大或者过小的瞳孔图像特征较为明显, 容易被识别, 二是因为 R_2 与 R_3 个体之间差异很小, 特征较为接近, 最终导致错误分类。

表 5 本文算法得出的分类混淆矩阵

Table 5 The classification confusion matrix obtained by the algorithm in this paper

	R_1	R_2	R_3
R_1	19	1	0
R_2	1	17	2
R_3	0	1	19

此外, 为更直观显示本文算法瞳孔检测效果, 图 6 展示了本文模型在随机挑选的数据集上的检测结果, 包含 3 类不同屈光度等级的瞳孔图像。由检测结果可知, 本文提出的算法很好地识别出 3 类红外瞳孔图像, 并且对于弱特征样本 R_1 也有较好的识别效果。

率为 91.6%。在下一步工作中, 重点研究小样本学习方式继续提高算法检测准确性。

参考文献:

- [1] 薛烽, 李湘宁. 一种基于图像处理的屈光度测量方法[J]. 光电工程, 2009, 36(8): 62-66, 74.
XUE Feng, LI Xiangning. A method of diopter measurement based on image processing[J]. Photoelectric Engineering, 2009, 36(8): 62-66, 74.
- [2] 胡志轩. 基于红外图像的眼视力屈光度检测系统[D]. 武汉: 华中师范大学, 2018.

- HU Zhixuan. An Eye Vision Diopter Detection System Based On Infrared Image[D]. Wuhan: Center China Normal University, 2018.
- [3] 谢娟英, 侯琦, 史颖欢, 等. 蝴蝶种类自动识别研究[J]. 计算机研究与发展, 2018, 55(8): 1609-1618.
- XIE Juanying, HOU Qi, SHI Yinghuan, et al. Research on automatic recognition of butterfly species[J]. *Computer Research and Development*, 2018, 55(8): 1609-1618.
- [4] Redmon J, Divvala S, Girshick R, et al. You only look once:unified real-time object detection[C]//*Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2016: 779-788.
- [5] Kaur P, Khehra B S, Pharwaha A. Deep transfer learning based multi way feature pyramid network for object detection in images[J]. *Mathematical Problems in Engineering*, 2021, 2021: 1-13.
- [6] Noh H, Hong S, Han B. Learning deconvolution network for semantic segmentation[C]//*IEEE International Conference on Computer Vision*, 2015: 5119-5127.
- [7] 刘智, 黄江涛, 冯欣. 构建多尺度深度卷积神经网络行为识别模型[J]. 光学精密工程, 2017, 25(3): 799-805.
- LIU Zhi, HUANG Jiangtao, FENG Xin. Building a multi-scale deep convolutional neural network behavior recognition model[J]. *Optics and Precision Engineering*, 2017, 25(3): 799-805.
- [8] 王娟, 刘嘉润, 李瑞瑞. 基于深度学习的红外相机视力检测算法[J]. 电子测量与仪器学报, 2019, 33(11): 36-43.
- WANG Juan, LIU Jiarun, LI Ruirui. Infrared camera vision detection algorithm based on deep learning[J]. *Journal of Electronic Measurement and Instrument*, 2019, 33(11): 36-43.
- [9] LIU Y P, JI X X, PEI S T, et al. Research on automatic location and recognition of insulators in substation based on YOLOv3[J]. *High Voltage*, 2020, 5(1): 62-68.
- [10] Murugan A, Nair S, Kumar K. Detection of skin cancer using SVM, random forest and KNN classifiers[J]. *Journal of Medical Systems*, 2019, 43(8): 683-686.
- [11] ZHU S G, DU J P, REN N, et al. Hierarchical-Based object detection with improved locality sparse coding[J]. *Chinese Journal of Electronics*, 2016, 25(2): 290-295.
- [12] 温江涛, 王涛, 孙洁娣, 等. 基于深度迁移学习的复杂环境下油气管道周界入侵事件识别[J]. 仪器仪表学报, 2019, 40(8): 12-19.
- WEN Jiangtao, WANG Tao, SUN Jiedi, et al. Intrusion event identification of oil and gas pipeline perimeter in complex environment based on deep migration learning[J]. *Chinese Journal of Scientific Instrument*, 2019, 40(8): 12-19.
- [13] 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 and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [14] LIU W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[C]//*Proceedings of European Conference on Computer Vision*, 2016: 21-37.