

DOI: 10.13382/j.jemi.B2003332

引入视觉机制的 SSD 网络在摩托车头盔佩戴检测中的应用*

刘琛¹ 王江涛^{1,2} 王明阳¹

(1. 淮北师范大学 物理与电子信息学院 淮北 235000; 2. 淮北师范大学 信息学院 淮北 235000)

摘要:近年来摩托车驾驶的安全问题越来越受到关注,摩托车手头盔的佩戴对自身的安全十分重要。针对于摩托车手是否佩戴头盔的检测问题,为提高检测网络的准确率和鲁棒性,对主流的一步检测网络 SSD-Net 引入类似视觉机制的模块,对网络的特征图在通道和空间上进行了权重的重新选择。并增加了类似人类视觉偏心率机制的 RFB 模块。通过使用 Mosaic 方法进行数据增强,并采用余弦衰减学习率来优化网络。实验结果表明改进后的网络对摩托车头盔佩戴的检测结果比原始 SSD-Net 提升了 4% 的 MAP,具有更好的应用效果。

关键词:深度学习;目标检测;视觉机制;SSD-Net;头盔检测

中图分类号: TP391 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Application of SSD network with visual mechanism in motorcycle helmet wearing detection

Liu Chen¹ Wang Jiangtao^{1,2} Wang Mingyang¹

(1. College of Physics and Electronic Information, Huaibei Normal University, Huaibei 235000, China;

2. School of Information, Huaibei Normal University, Huaibei 235000, China)

Abstract: In recent years, more and more attention has been paid to the safety of motorcyclists. Wearing helmets is very important for their own safety. In order to improve the accuracy and robustness of the detection network, in this paper, the mainstream one-step detection network SSD net is introduced with similar visual mechanism module, and the weight of network feature map is reselected in channel and space. The RFB module is also added to the network, which is similar to the human visual eccentricity mechanism. We also use Mosaic method for data enhancement and cosine attenuation learning rate to optimize the network. The experimental results show that the MAP value of the improved network is about 4% higher than that of the original SSD net. And it has better application effect.

Keywords: deep learning; target detection; visual mechanism; SSD-Net; helmet detection

0 引言

持续的交通机动化导致全球道路相关伤亡人数持续增加。为了应对这种局面,各国政府正把重点放在安全和安全方面在交通中遵守行为。然而,对于摩托车的安全驾驶规范,安全的全面的头盔使用的标准十分重要。我国城市道路陌陌交通,在许多城市若路途较近,人们多会选择摩托车作为代步工具上下班以及接送孩子上下学。同时,如今的外卖产业兴起,外卖员多是驾驶摩托车

进行配送。在这些情况下摩托车驾驶人佩戴头盔对预防伤害,保护他人和自身安全至关重要。因此,将人工智能应用于道路安全中来检测摩托车手是否佩戴头盔有十分重要的意义。

近年来,目标检测方向的深度学习算法可以大体上划分为两阶段(two-stage)算法和单阶段(one-stage)算法。Faster R-CNN^[1]是两阶段算法的代表网络。此网络先通过一些预处理方法在预测图片上生成若干目标候选框,再使用一些常用的卷积神经网络如 VGG-16^[2]训练参数进行样本的分类。一般使用两阶段算法的网络进行

收稿日期:2020-07-14 Received Date: 2020-07-14

* 基金项目:国家自然科学基金(61203272)、安徽省高校自然科学研究重大项目(KJ2018ZD038)、安徽省高等学校省级质量工程项目(2017kfk043,2019jxtd142)资助

训练准确率更高,但计算成本较大,所以影响了网络的训练速度。后来推出的如 YOLO (you only look once)^[3], SSD (single shot multibox detector)^[4]等单阶算法是将目标边框定位的问题转化为回归问题,可以进行端到端检测。虽然牺牲了一定的准确率,但是没有候选框,速度快,实时性好的优势,使之被大量推广于工程应用中。因此单阶段的一步检测方法成为当下目标检测的主流方法^[5]。

随着检测网络的不断进化,人们开始对视觉机制在目标检测的效果和联系进行研究。因为人类在观测事物的时候,对不同物体的灵敏度会有区别。这种视觉的注意力机制对检测网络检测特征图中的信息也会有所影响。大量研究表明,在网络中加入类似视觉的注意力模块使检测网络的准确度得到了提升^[6]。

目前检测网络中的视觉机制模型中有在通道上和空间上进行权值激活或者在空间上通过空洞卷积实现模拟视觉的群体感受野机制。本文在一步检测网络 SSD-Net 的基础上通过添加通道注意力模块^[7],空间注意力

模块^[8],群体感受野 RFB (receptive field block) 模块^[9]三种视觉注意力机制的模块进行实验。并且对网络的数据增强方法和学习率进行了改进。通过验证,本文算法在摩托车手佩戴头盔的检测效果上取得了更好地效果。

1 SSD 模型

1.1 SSD 网络结构

SSD 网络使用了尺寸大小逐层递减的特征提取层,可以产生不同尺度下检测的预测值结果,使用这些尺度下的检测结果来预测目标的类别以及物体与默认边界框的偏移量,并且通过 NMS^[10]的方法来产生最终的检测框。SSD 的基础的训练网络一般使用 VGG-16 Net,并将输入图像归一化为 300×300。网络结构中会有 6 个特征层依次输出不同大小和维度的图片进行检测和分类。网络在训练过程中会在每一特征层上检测特定尺寸的预测目标。图 1 所示为 SSD 网络结构。

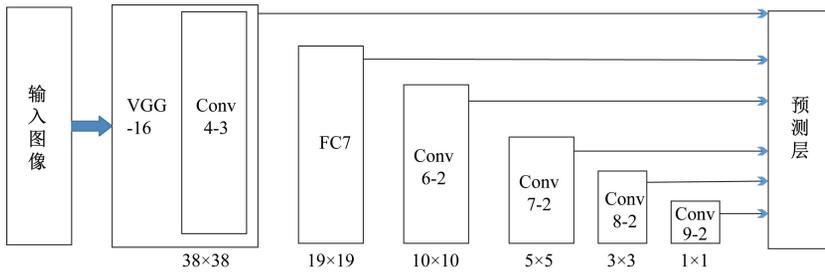


图 1 SSD 网络结构

Fig. 1 SSD network structure

1.2 SSD 网络候选框生成

SSD 中的候选框类似于两阶段网络中的锚点机制,而且在网络的不同层中都存在着相应尺度的候选框。对于 SSD 中的 m 个特征图的默认框尺寸 g_i 如式(1)所示。

$$g_i = g_{\min} + \frac{g_{\max} - g_{\min}}{n - 1}(i - 1) \quad i = 1, 2, \dots, m \quad (1)$$

式中: g_{\min} 表示最低层尺度; g_{\max} 最高层尺度; m 表示层数及特征图的个数; g_i 表示第 i 个特征图的尺度, $i \in [1, m]$ 。对默认框设置的不同宽高比可以用 $\alpha_r = \{1, 2, 3, \frac{1}{2}, \frac{1}{3}\}$ 表示,默认框的宽 w_i^a 和高 h_i^a 可由式(2)得到。

$$\begin{aligned} w_i^a &= g_i \sqrt{\alpha_r} \\ h_i^a &= g_i / \sqrt{\alpha_r} \end{aligned} \quad (2)$$

通过上述机制,对于每个特征单元就都会产生不同形状和尺寸的候选框,通过对 6 个不同维度的特征图进行候选框的生成,SSD 网络一共有 8 732 个框,通过这些候选框便可对检测物体进行分类和回归^[11]。

2 类视觉机制原理和实现

本文通过对视觉机制模型中的利用空洞卷积实现模拟视觉的群体感受野机制以及在通道和空间上进行权值重新激活的注意力机制进行研究,将深度学习中的类视觉机制与网络进行结合与应用。

2.1 RFB 模块

群体感受野是人类视觉研究学中的一种概念,有研究发现它的范围大小与离心率在视网膜的定位图中呈正比关系。这中机制突出了在视网膜中心区域目标的重要性。RFB 模块模拟了人类视网膜的这一能力,加强了轻量级卷积神经网络学习深层特征的能力。RFB 模块由多分支的卷积层结构和空洞卷积层构成^[12]。具体的模拟效果如图 2 所示。

RFB 模块中采用的多分支的结构,在每个分支中分别采用不同大小的卷积核卷积来获取不同比例的感受野。RFB 模块通过瓶颈 (bottle-neck) 结构减少了训练参

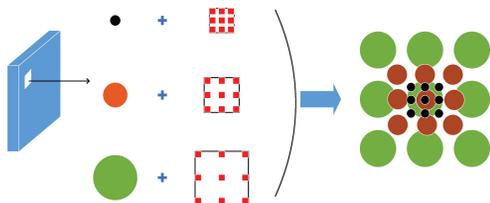


图 2 RFB 模块模拟的视觉感受野

Fig. 2 Visual field simulated by RFB module

数,使网络更轻量化。并且还运用了 Res-Net^[13] 中的直连结构,解决了网络层数过深时可能导致的梯度发散问题。而 RFB-S 模块的分支更多而且卷积核尺寸更小,类似于视网膜中较小的感受野,使模块更接近人类的浅层视网膜模型。两个网络模块的结构如图 3(a)和(b)所示。

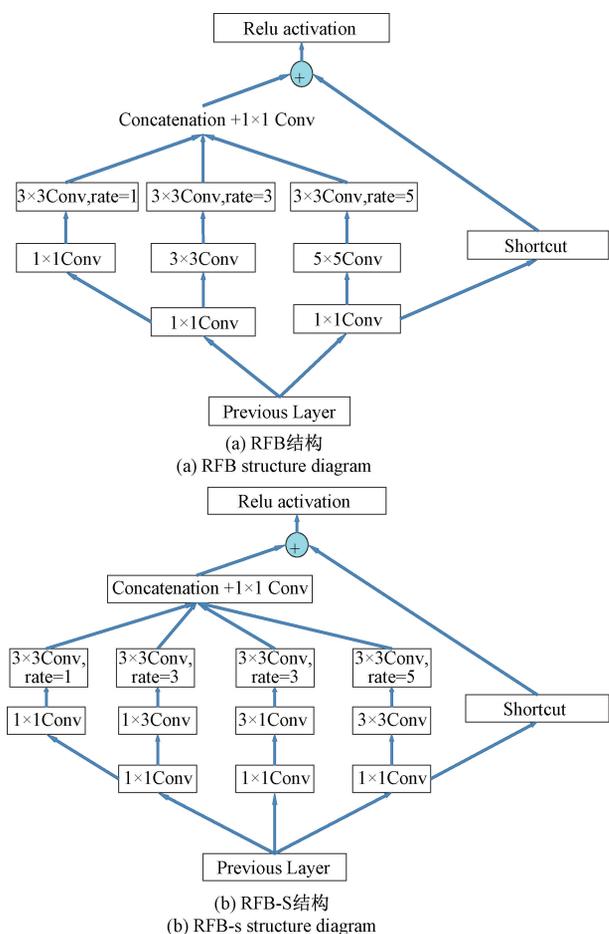


图 3 RFB 和 RFB-S 结构

Fig. 3 Structure of RFB and RFB-S

空洞卷积本质上是在普通卷积运算上的一种推广。在空洞卷积中有一个 dilation rate 参数,可通过此参数来扩张卷积核并用 0 填充卷积核中多出的部分^[14]。这个参数代表卷积核的膨胀系数。空洞卷积实际卷积核大小

如下:

$$N = n + (n - 1)(r - 1) \quad (3)$$

式中: n 为原始的卷积核大小; r 为空洞卷积参数及膨胀系数。空洞卷积的运算方法是通过卷积核与设定间隔的像素点做相应的乘法运算。这种卷积方法在没有增加过多复杂设计的情况下提高了感受野的范围,RFB 模块正是通过空洞卷积层来模拟实现了人类视觉皮质中离心率的效果^[14]。

2.2 通道注意力模块

注意力模块主要分为通过通道和空间两个方面对特征图的权值进行重新激活。其中通道注意力模块通过关注检测网络特征图通道中的信息情况进行权值的重新分配,使用通道对特征进行响应达到自适应的校准^[15]。具体结构如图 4 所示。

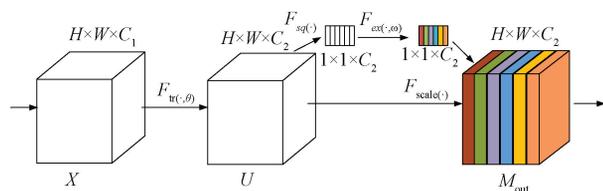


图 4 通道注意力模块结构

Fig. 4 Structure of channel attention module

首先对输入大小为 $H \times W \times C$ 的特征图使用设定好的卷积操作变换得到一个通道数为 C_2 的新特征图 U ,之后在空间维度也就是在通道上进行特征的压缩操作 F_{sq} 。具体步骤进行了一个全局空间的最大池化,获得了一个 $1 \times 1 \times C_2$ 的矩阵。这个矩阵具有原始特征图的全局感受野。通过门机制,给予一定的阈值对每个通道中的值进行激活,经过这种通道的激励变换得到不同的权重系数即图中 F_{ex} 的过程。最后用具有权值的模块与原来的特征图相乘得到拥有视觉强弱的新特征图 M_{out} 。具体过程可用式(4)表示。

$$u_c = v_c \cdot X = \sum_{s=1}^c v_c^s \cdot x^s$$

$$z_c = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (4)$$

$$s_c = F_{ex}(z_c, W) = \sigma(g, (z_c, W))$$

$$M_{OUT} = F_{scale}(u_c, s_c) = u_c \cdot s_c$$

式中: u_c 表示特征图 U 中第 c 个二维矩阵,下标 c 表示通道。

2.3 空间注意力模块

空间注意力模块与通道注意力模块的实现有所不同,它通过关注检测网络特征图的空间中的信息情况进行权值的重新分配。具体结构如图 5 所示。

对输入大小为 $H \times W \times C$ 的特征图,先对通道使用平

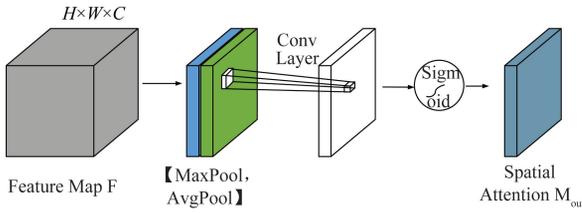


图5 空间注意力模块结构

Fig. 5 Structure of spatial attention module

均池化和最大池化来简化通道数,将特征更广泛的分布在空间之中,将得到两个 $H \times W \times 1$ 的特征描述图按通道

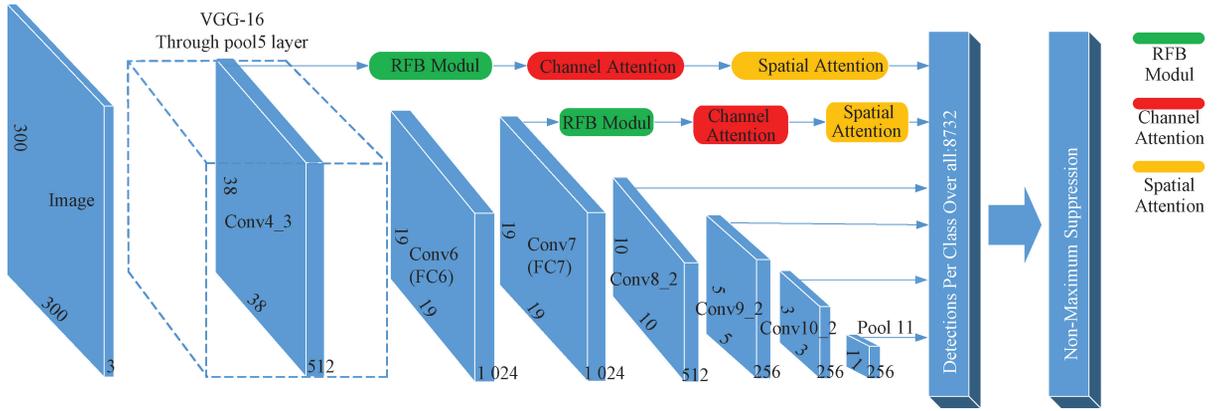


图6 添加视觉机制的SSD网络结构

Fig. 6 SSD network structure with visual mechanism added

在SSD网络主干中分别加入了不同的视觉模块在Conv4_3、Conv6两个低层的特征图之后,对特征图进行修正。这两个特征层的维度高特征明显,生成候选框最为密集的两层,也是主要用来检测小目标物体的特征层。先通过视觉感受野模增加特征图的丰富度再使用通道和空间的注意力模块,对特征图进行了新的权重分配。最后通过卷积分类器对新的特征图进行分类和回归。

网络训练时载入SSD网络的预训练参数,并将视觉模块中的权值随机初始化,整体网络多次训练迭代进行训练调参以得到最终模型。

2.5 数据增强和学习率的改进

为了使网络在不同情况下有更好的鲁棒性,通常会对训练图像进行数据增强的预处理。本文通过使用Mosaic数据增强的方法,来丰富训练集^[17]。此方法分别对4张图片进行翻转、缩放、色域变化等,并且按照4个方向位置摆好并进行拼接。图7所示为所使用的数据集图片使用Mosaic方法后的效果图。

在卷积神经网络的训练过程中,学习率的设定是一个十分重要的因素^[18]。本文通过改进使用了余弦衰减的学习率进行训练,使得学习率下降更加平缓 and 均匀,取

拼接。然后通过一个卷积层和激活函数对单通道的特征图进行处理,经过这种对空间的激励变换得到不同的权重系数。最后拿得到的权值与原始特征图相乘得到最终具有注意力特征的新特征图^[16],可用式(5)表示。

$$M_{OUT}(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPoll(F)])) \quad (5)$$

式中: σ 为sigmoid操作,7x7表示卷积核的大小。

2.4 添加类视觉模块的SSD网络结构设计

针对本文介绍的类视觉改进机制,本文分别将这些模型加入SSD检测网络中,进行了新的网络设计,结构如图6所示。

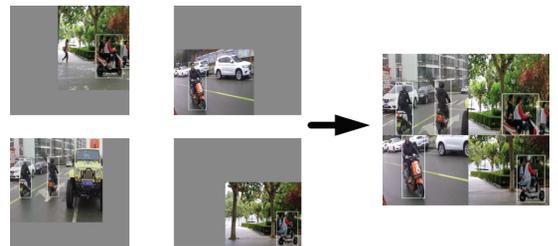


图7 Mosaic数据增强实例

Fig. 7 Example of Mosaic data enhancement

得了不错的训练效果。余弦衰减就是通过余弦的相关形式对学习率进行衰减,此种学习率的衰减图和余弦函数相似。学习率的更新机制如下:

$$\begin{aligned} g_step &= \min(g_step, d_step) \\ \cos_decay &= 0.5 \cdot (1 + \cos(\pi \cdot g_step/d_steps)) \\ decayed &= (1 - l_min) \cdot \cos_decay + l_min \\ decayed_learning_rate &= l_rate \cdot decayed \end{aligned} \quad (6)$$

式中: l_rate 是初始设定的学习率的; g_step 表示衰减计算过程中的全局步数; d_steps 为衰减步数; l_min 为设定的最小学习率。标准的余弦衰减曲线如图8所示,学习

率从初始值光滑下降到最低的学习率后维持不变。余弦衰减的学习率的下降十分平缓,不会产生较大的波动,对于一些网络的训练有意想不到的提升。

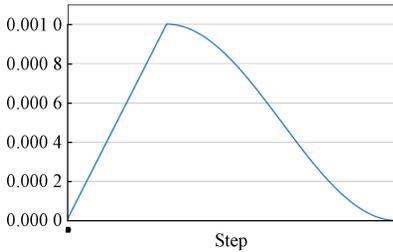


图 8 余弦衰减曲线

Fig. 8 Cosine attenuation curve

3 实验结果及分析

3.1 实验准备

测试网络使用的为 Win10 操作系统, GPU 为 NVIDIA GTX 1080Ti, 运行内存为 16G, 选择 Pytorch1.1.0 为深度学习框架。在一步检测网络 SSD-Net 的前两层特征图后通过添加通道注意力模块、空间注意力模块、群体感受野模块 RFB 来进行对比。实验中使用的数据集是自行拍摄收集的, 内容为安徽省淮北市市区道路和淮北高校校园主干道上的一些行驶的摩托车, 车辆为不同类型的电瓶车和摩托车。实验中选取的样本包括了较远车辆、部分遮挡和车辆行驶较快图像模糊以及各角度的情况。并收集了多云以及小雨下的天气环境, 为本研究的算法训练、验证和测试奠定了基础^[19]。

从收集的视频素材中每 3 帧抽取一张图片, 一共保存 2 000 张 1 080×720 的有效图片作为训练集, 其中测试集与训练集采用 1:4 划分, 使用标记工具 LabelImg 对头盔佩戴情况和车辆位置进行标注, 并生成 VOC 格式的训练文件。给出部分角度下收集的图像, 如图 9 所示, 图 9(a)、(b) 为数据集 A 模拟手持执勤设备对道路抓拍的车辆, 图 9(c) 为数据集 B 模拟违章监控拍摄的市区道路行驶车辆^[20]



(a) 图例1
(a) Key 1

(b) 图例2
(b) Key 2

(c) 图例3
(c) Key 3

图 9 数据集实例

Fig. 9 Dataset instance

3.2 加入不同模块的结果分析

本文方法中在低层特征层后添加了视觉机制模块, 有利于对特征主次的提取。为验证各模块的作用, 针对文中介绍的类视觉改进机制, 本文分别将这些模型加入 SSD 检测网络中。先通过 VOC 数据集进行训练和测试, 将原始的 SSD 再训练模型和本文模型进行对比。

训练完毕后对网络进行测试得出各种目标的 AP 值和 MAP 值并与原始 SSD Net 的上述指标进行记录对比, 其 MAP 结果如表 1 所示。通过表 1 可以观察到类视觉机制模块对检测网络的准确率确实得到了提升。尤其是在使用了空间和通道的注意力机制和 RFB 模块相互配合的情况下效果最好, 在 VOC-07 测试集上有 6% 的 MAP 值的提升。

表 1 添加不同视觉模块在 VOC-07 测试集上的性能对比

Table 1 Performance comparison of different visual modules on voc-07 test set

算法网络	SSD-300	RFB Net	RFB-Channel Net	RFB-Channel-Spatial-Net
MAP/%	73.58	75.39	76.16	79.49

之后, 再在效果最好的网络基础上使用 Mosaic 方法进行数据增强, 并采用余弦衰减学习率来优化网络。使用此网络对模拟违章摄像头对道路抓拍情况的数据集 A 进行训练。并在此基础上删除部分模块进行对比。训练完毕后对不同网络进行测试得到目标的 MAP 值并指标进行记录对比, 其 MAP 值结果如表 2 所示。通过表 2 可以观察到, 每个改进的部分对网络的性能都有 1% 的提高。本文改进后的 SSD 算法相比原始 SSD 网络模型对检测摩托车手是否佩戴了头盔的 MAP 值有 4% 的提高。

表 2 不同识别算法对头盔检测的性能对比

Table 2 Performance comparison of different recognition algorithms for helmet detection

模型	Mosaic	Cos-lr	RFB-Channel-Spatial-Net	MAP/%
SSD-1				88.46
SSD-2		✓	✓	90.52
SSD-3	✓		✓	91.43
SSD-4	✓	✓		89.21
本文 SSD	✓	✓	✓	92.73

为防止在校园道路上较多自行车对检测存在的干扰, 并验证本文网络的鲁棒性。在数据集 A 中加入了大量自行车存在的样本并对自行车进行标注。并通过本文网络进行训练。训练后网络对车辆的错分概率在 1% 左右, 体现了本文网络的优越性。自行车和电动车的检测结果如图 10 所示, 图 10(a) 为检测目标的类别和置信度, 图 10(b) 为检测框与真实框的效果, 绿色则表示检测类别正确。

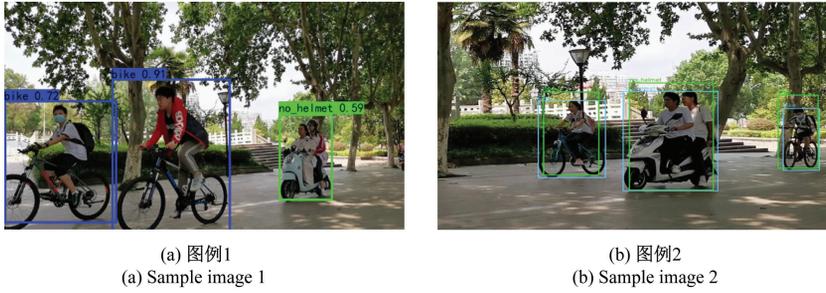


图 10 自行车和电动车的检测结果

Fig. 10 Test results of bicycles and electric vehicles

为了更清晰的展示改进后网络的训练效果和应用能力,对本文网络使用在道路天桥处模拟违章摄像头进行拍摄的数据集 B 进行训练,并通过错分概率、漏检概率以及每帧图片检测时间的性能指标进行评估,并与原始SSD网络对比。错分概率、漏检概率如式(7)所示,结果如表3所示。

$$\begin{aligned} \text{错分概率} &= \frac{N_{cf}}{N} \\ \text{漏检概率} &= \frac{N_{df}}{N_{gt}} \end{aligned} \quad (7)$$

式中: N_{cf} 为测框中分类错误的数量; N 为总预测框数; N_{gt} 为真实框总数; N_{df} 为未检测出的预测框数量。

表 3 不同识别算法对数据集 B 头盔检测的性能指标对比

Table 3 Performance comparison of different recognition algorithms for helmet detection in dataset B

算法网络	MAP/%	错分概率/%	漏检概率/%	检测时间/ms
SSD-300	89.82	2.34	3.48	18.43
本文 SSD	93.03	1.12	1.89	20.02

由表3可以观察到,本文通过改进的算法网络相较于原始的SSD网络,在MAP上有明显提高外,对目标的错检和漏检概率都有明显的降低,并且检测时间增加不大,可以满足实时检测的要求。部分数据集在改进前后的测试结果对比如图11所示。通过图11对比可看出,在改进前后的SSD网络训练结果下测试图11(a)中左侧的目标AP值有一定提高,图11(b)中本文改进后的网络成功检测出了在原始网络中未成功检测出的较远目标,图11(c)中右侧目标有一定程度错检的红色框,在改进后的网络中成功得到了修正。可见,本文改进的网络取得了比较明显的效果,有效减少了漏检错检的情况,体现出了改进网络的优越性。

4 结论

本文通过引入视觉机制,利用SSD网络对摩托车驾驶人头盔佩戴情况进行检测。针对传统单阶段算法检测网络准确率较低的情况,本文通过添加视觉机制在传统

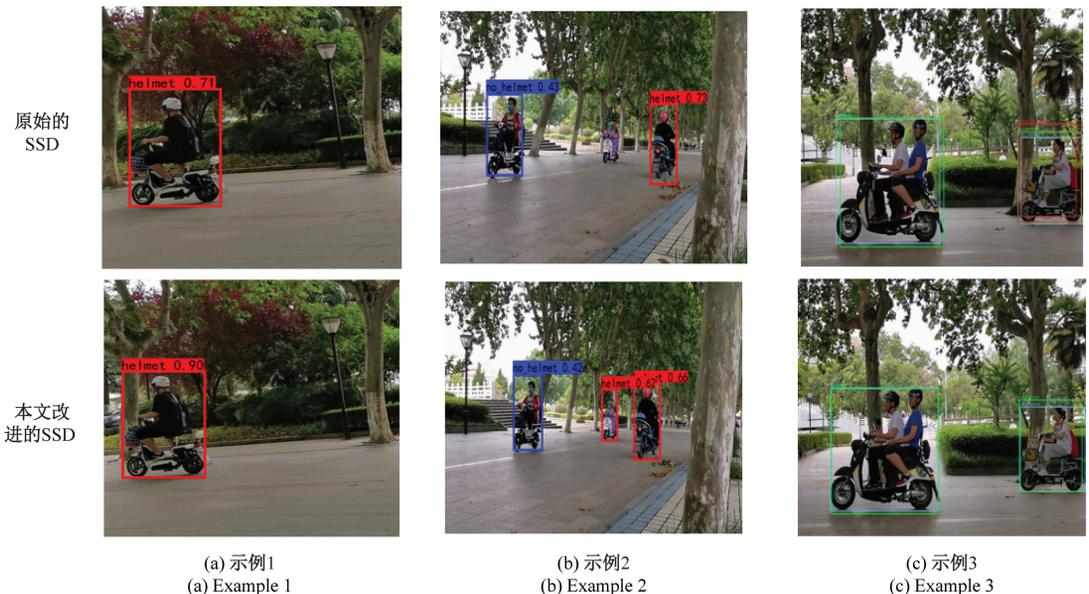


图 11 两种网络的测试结果对比

Fig. 11 Comparison of test results of two networks

的 SSD 网络之中,改进特征层的输出。并通过使用 Mosaic 方法进行数据增强和余弦衰减学习率来优化网络,使网络对摩托车驾驶人是否佩戴头盔的检出率的正确性上都得到了明显的提升。之后的工作中,将继续进行视觉机制的研究,让更多的类视觉机制与检测网络融合,提升检测网络的性能。

参考文献

- [1] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks[C]. Advances in Neural Information Processing Systems, 2015: 91-99.
- [2] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. Computer Science, arXiv preprint arXiv:1409.1556, 2014.
- [3] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.
- [4] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]. European Conference on Computer Vision, Springer, 2016: 21-37.
- [5] 谢晓蔚,史健芳. 弱监督卷积神经网络的多目标图像检测研究[J]. 电子测量与仪器学报, 2019, 33(6): 31-37.
- XIE X W, SHI J F. Research of convolutional neural networks with weakly-supervised learning on multi-object image detection [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(6): 31-37.
- [6] 孙萍,胡旭东,张永军. 结合注意力机制的深度学习图像目标检测[J]. 计算机工程与应用, 2019, 55(17): 180-184.
- SUN P, HU X D, ZHANG Y J. Object detection based on deep learning and attention mechanism [J]. Computer Engineering and Applications, 2019, 55(17): 180-184.
- [7] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [8] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[J]. ECCV, 2018.
- [9] LIU S, HUANG D. Receptive field block net for accurate and fast object detection [C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 385-400.
- [10] NEUBECK A, GOOL L J V. Efficient non-maximum suppression [C]. 18th International Conference on Pattern Recognition (ICPR 2006), IEEE Computer Society, 2006.
- [11] 孙锐,王慧慧,叶子豪. 融合深度感知特征与核极限学习机的行人检测[J]. 电子测量与仪器学报, 2019, 33(2): 39-47.
- SUN R, WANG H H, YE Z H. Pedestrian detection based on combining depth perception features with kernel extreme learning machine [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(2): 39-47.
- [12] 刘相云,郭呈渊,龚志辉,等. 一种改进的 RFB Net 遥感影像目标识别算法[J]. 测绘科学技术学报, 2019, 36(2): 179-184.
- LIU X Y, GUO CH Y, GONG ZH H, et al. Object detection in remote sensing image with improved RFB net [J]. Journal of Geomatics Science and Technology, 2019, 36(2): 179-184.
- [13] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [14] 王永利,曹江涛,姬晓飞. 基于卷积神经网络的 PCB 缺陷检测与识别算法[J]. 电子测量与仪器学报, 2019, 33(8): 78-84.
- WANG Y L, CAO J T, JI X F. PCB defect detection and recognition algorithm based on convolutional neural network [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(8): 78-84.
- [15] 王培森,宋彦,戴礼荣. 基于多通道视觉注意力的细粒度图像分类[J]. 数据采集与处理, 2019, 34(1): 157-166.
- WANG P S, SONG Y, DAI L R. Fine-grained image classification with multi-channel visual attention [J]. Journal of Data Acquisition and Processing, 2019, 34(1): 157-166.
- [16] 王培森. 基于注意力机制的图像分类深度学习研究方法[D]. 合肥:中国科学技术大学, 2018.
- WANG P S. Research on attention based image classification with deep learning [D]. Hefei: University of Science and Technology of China, 2018.
- [17] 吴雪,宋晓茹,高嵩,等. 基于数据增强的卷积神经网络火灾识别[J]. 科学技术与工程, 2020, 20(3): 1113-1117.
- WU X, SONG X R, GAO S, et al. Convolution neural

network based on data enhancement for fire identification [J]. Science Technology and Engineering, 2020, 20 (3): 1113-1117.

- [18] 王古森,高波. 基于自适应学习率BP神经网络的火点定位模型[J]. 计算机系统应用, 2019, 28(3): 250-254.

WANG G S, GAO B. Fire location model based on adaptive learning rate BP neural network [J]. Computer Systems & Applications, 2019, 28(3): 250-254.

- [19] 李晖晖,周康鹏,韩太初. 基于CReLU和FPN改进的SSD舰船目标检测[J]. 仪器仪表学报, 2020, 41(4): 183-190.

LI H H, ZOU K P, HAI T CH. Ship object detection based on SSD improved with CReLU and FPN [J]. Chinese Journal of Scientific Instrument, 2020, 41(4): 183-190.

- [20] 王卫东,程丹. 监控场景下的实时车辆检测方法[J]. 电子测量与仪器学报, 2018, 32(7): 83-88.

WANG W D, CHEN D. Real-time vehicle detection method for video surveillance [J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(7): 83-88.

作者简介



刘琛, 2019年于淮北师范大学获得学士学位,现为淮北师范大学硕士研究生,主要研究方向为数字图像处理、深度学习。

E-mail: 1010981639@qq.com

Liu Chen received B. Sc. degree from HuaiBei Normal University in 2019. He is currently a M. Sc. candidate at the HuaiBei Normal University. His main research interests include Digital image processing and deep learning.



王江涛, 2002年于青岛大学获得学士学位, 2005年于青岛大学获硕士学位, 2008年于南京理工大学获得博士学位, 现为淮北师范大学教授, 主要研究方向为图像处理、模式识别、计算机视觉。

E-mail: jiangtaowang@chnu.edu.cn

Wang Jingtiao received B. Sc. degree from Qingdao University in 2002, M. Sc. degree from Qingdao University in 2005, and Ph. D. degree from Nanjing University of Technology in 2008, respectively. Now he is a professor at HuaiBei Normal University. His main research interests include image processing, pattern recognition, computer vision.