

深圳大学学报(理工版)

Journal of Shenzhen University(Science and Engineering)

ISSN 1000-2618,CN 44-1401/N

《深圳大学学报(理工版)》网络首发论文

题目: 基于 CA-BIFPN 的交通标志检测模型
作者: 郎斌柯, 吕斌, 吴建清, 吴瑞年
收稿日期: 2022-07-16
网络首发日期: 2023-02-25
引用格式: 郎斌柯, 吕斌, 吴建清, 吴瑞年. 基于 CA-BIFPN 的交通标志检测模型[J/OL]. 深圳大学学报(理工版).
<https://kns.cnki.net/kcms/detail//44.1401.N.20230223.1413.004.html>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

【交通物流 / Transportation Logistics】

基于CA-BIFPN的交通标志检测模型

郎斌柯¹, 吕斌¹, 吴建清², 吴瑞年³

1) 兰州交通大学交通运输学院, 甘肃兰州 730070; 2) 山东大学齐鲁交通学院, 山东济南 250002;
3) 兰州市大数据管理局, 甘肃兰州 730000

摘要: 正确、快速的交通标志检测可为自动驾驶领域的环境感知提供重要信息。针对目前交通标志检测识别率低及多种交通标志检测存在的误检漏检等问题, 提出一种协调注意力-双向特征金字塔网络(coordinate attention-bidirectional feature pyramid network, CA-BIFPN)交通标志检测模型。该模型将YOLO(you only look once)v5模型和协调注意力(coordinate attention, CA)机制相结合, 引入双向特征金字塔网络(bidirectional feature pyramid network, BIFPN), 通过跳连特征融合提高模型的多尺度语义特征利用效率, 在提高小目标物体检测效率的同时, 也使交通标志的检测精度得到提高。以交通标志数据集TT100K为测试对象进行实验验证, 结果表明, 与SSD(single shot multibox detector)模型和YOLOv5模型相比, CA-BIFPN交通标志检测模型的检测准确率分别提高4.5%和1.3%, 验证了模型有效。

关键词: 人工智能; 深度学习; 交通标志检测; 小目标检测; 协调注意力; 双向特征金字塔网络
中图分类号: U495; TP183 **文献标志码:** A

A traffic sign detection model based on attention mechanism and spatial pyramid

LANG Binke¹, LÜ Bin¹, WU Jianqing², and WU Ruinian³

1) School of Traffic and Transportation, Lanzhou Jiaotong University, Lanzhou 730070, Gansu Province, P. R. China
2) School of Qilu Transportation, Shandong University, Jinan 250002, Shandong Province, P. R. China
3) Big Data Administration of Lanzhou, Lanzhou 730000, Gansu Province, P. R. China

Abstract: In the field of autonomous driving, the correct detection of traffic signs can provide important information for driving. Aiming at the low recognition rate of current traffic sign detection and the misdetection and missed detection of various traffic signs, a coordinate attention-bidirectional feature pyramid network (CA-BIFPN) traffic sign detection model is proposed. In this traffic sign detection model, a combination of you only look once version 5 (YOLOv5) network and coordinate attention (CA) mechanism is adopted, and a bidirectional feature pyramid network (BIFPN) is introduced to improve the utilization efficiency of multi-scale semantic features of the model through skip-connection feature fusion. The detection efficiency of small target objects is improved, and the detection accuracy of traffic signs is increased. The traffic sign data set TT100K is used as the test object to conduct comparative experiments, and the results of single shot multibox detector (SSD) and YOLOv5 model detection are compared and

Received: 2022-07-16; **Accepted:** 2022-11-23

Foundation: Higher Education Innovation Foundation of Department of Education of Gansu Province(2020B-108); Double-First Class Major Scientific Research Foundation of Department of Education of Gansu Province (GSSYLXM-04); Natural Science Foundation of Jiangsu Province (BK20200226); Science and Technology Program of Suzhou (SYG202033); Open Project of Key Laboratory of Ministry of Public Security for Road Traffic Safety (2021ZDSYSKFKT07); Research Program of Department of Transportation of Shandong Province (2020BZ01-03)

Corresponding author: Professor LÜ Bin. E-mail: jdlbxx@mail.lzjtu.cn

Citation: LANG Binke, LÜ Bin, WU Jianqing, et al. A traffic sign detection model based on attention mechanism and spatial pyramid [J]. Journal of Shenzhen University Science and Engineering, XXXX, XX(XX): 1-9. (in Chinese)

analyzed. The detection accuracy of CA-BIFPN traffic sign detection model is improved by 4.5% and 1.3%, which verifies the effectiveness of the model.

Key words: artificial intelligence; deep learning; traffic sign detection; small target detection; coordinate attention; bidirectional feature pyramid network

交通标志是用汉字或符号传达导向、控制、警戒或指挥等信号的道路设施,能够熟练解读标志含义是每个驾驶员的基本要求,对于自动驾驶也是如此.交通标志检测是自动驾驶及辅助驾驶系统的重要研究方向,正确识别道路交通标志关系着智能驾驶任务的成败.

有关交通标志检测的优化模型包括基于人力的交通标志检测优化模型和基于深度神经网络的交通标志检测模型两类^[1-2].运用深度神经网络的交通标志检测优化模型主要有1阶段目标检测模型和2阶段目标检测模型.2阶段目标检测模型需从图片中产生候选区(运行后续模型的子区域),再从候选区域中生成最终的物体检测边框^[3].与2阶段目标检测模型相比,1阶段目标检测模型直接得到目标物体的分类概率和方位坐标值,并通过此环节直接获得最后的检测结果.1阶段目标检测模型包括SSD(single shot multibox detector)^[8]与YOLO(you only look once)^[9-11]系列等.

2阶段目标检测模型的精度一般高于1阶段检测模型,但2阶段目标检测模型的速率较慢,无法满足实际应用中实时性的需要.1阶段目标检测模型同时确保了精度和处理速度,因而其应用更广泛^[12].但是由于交通标志目标通常为小目标(物体相对图像占比小于0.58%^[13])物体,所以,上述检测模型的精度均不能满足实际需求.

注意力机制可以提取图像中不同位置的重要程度,从而使神经网络专注于图像中更重要的部分^[14].因此引入在检测模型中引入注意力机制将有助于提高模型对于小目标物体的检测.基于此,考虑检测实时性和精度的需要,本研究选择使用YOLOv5(you only look once version 5)为基础检测模型,以此为骨干网络,并在骨干网络中添加协调注意力(coordinate attention, CA)机制,聚焦于检测小目标交通标志检测.同时在模型特征融合部分引入加权双向特征金字塔,增强不同尺度特征图融合效果,进一步提升模型检测精度.

1 CA-BIFPN 交通标志检测模型

在不降低实时检测精度的条件下,为提高YOLOv5网络模型对交通标志的识别精度,本研究提出基于协调注意力-双向特征金字塔网络(coordinate attention-bidirectional feature pyramid network, CA-BIFPN)的交通标志检测模型,其网络结构如图1.可见,基于CA-BIFPN的交通标志检测模型网络结构主要由输入、骨干网络、Neck特征融合网络及预测网络构成.采用YOLOv5 6.0版本神经网络为基础网络,在其骨干网络中加入CA注意力,在添加少量训练参数的同时还能增加网络特征提取能力.相关注意力对比实验^[15]表明,虽然通道注意力能够显著提高模型性能,但其容易忽视与空间选择性注意力产生密切相关的位置信息,在CA注意力机制中将位置信息放入通道注意力中,则能够较好规避这一问题.YOLOv5与注意力结合机制结合主要有两种方法^[16]:①将注意力机制与Neck特征融合网络融合,替换骨干网络中所有跨阶段局部网络(cross stage partial networks, CSPN)模块;②在骨干网络后单独加入注意力机制.本研究采用后种方法,将CA注意力机制加入至空间金字塔前,随后导入Neck特征融合网络,在特征融合网络中混合不一样的特征图.选用加权双向特征金字塔网络,可以将交通标志图像的各个特征合理融合,有效解决各尺度特征信息不一致以及其他特征金字塔计算量大的问题.以下分别说明检测模型各优化模块的组成.

1.1 CA-BIFPN 注意力机制

注意力机制的本质是寻找特定的数据信号,操纵无用信息,结果一般以几率图和概率特征空间向量形式表达,利于增强神经网络对于小目标物体的检测精度.注意力机制模型可分为空间注意力(spatial attention)模型、通道注意力(channel attention)模型及混合注意力(mixed attention)模型^[17].

空间注意力偏向全部通道,在二维平面图中训练尺寸为 $H \times W$ 的特征图权重值矩阵,并为全部图像单元获得1个权重值.该权重表示对某个空间位

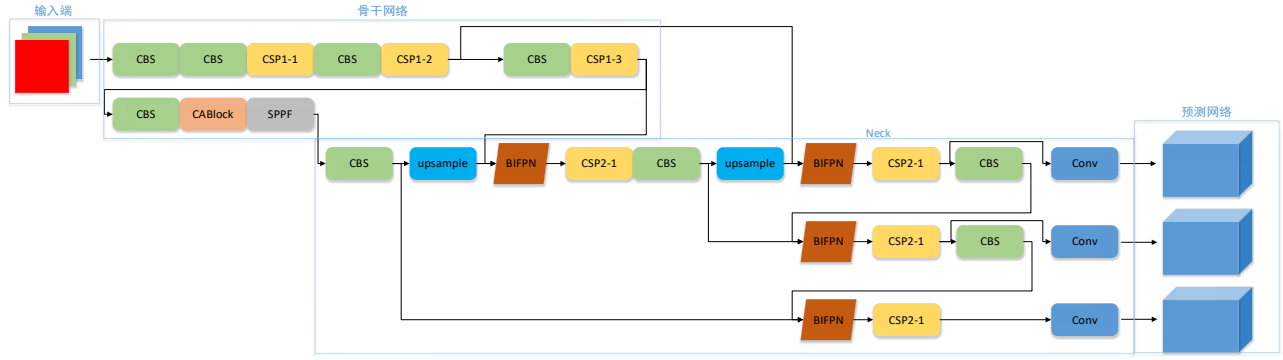


图1 CA-BIFPN模型网络结构

Fig. 1 (Color online) Network structure of CA-BIFPN model.

置信息的重要程度, 将空间注意力矩阵附着到原始特征图像上, 以增加可用特征, 减少无效特征, 进而达到特征筛选与增强得目的.

与空间注意力不同, 通道注意力类似于在各个通道的特征图中都给出1个权重值, 表明方式和重要信息之间的关联性, 同时该权重值与相关性呈正相关^[18]. 在深度神经网络中, 层面越高特征度的尺寸越小, 但通道数量更多. 通道还体现全部图形的特点信息, 信息过多时神经网络难以挑选出有效信息, 此时利用通道注意力就可以使网络判断出最重要信息, 这也是利用通道注意力进行目标检测时效果更优的原因^[19]. CA注意力机制, 为混合注意力, 其包含了通道注意力和空间注意力, 也继承了二者的优点, 使CA注意力在获得跨通道信息的同时, 获得方向感知和位置感知信息, 因此, 检测模型可以对目标进行更为精准地定位与识别. CA注意力机制模型如图2. CA注意力应用精确的部位信息对通道关联信息和长期性依靠信息开展编码, 编码过程包括Coordinate信息嵌入和CA注意力生成^[20].

在Coordinate信息嵌入过程中, 全局池化方法通常用作对空间内容的整体编码, 但由于该方式将去全局空间信息内容压缩在通道描述中, 无法表达信息的位置内容. 为使控制模块可以更好捕获具备精确部位信息的远程空间交互, 利用式(1)计算全局池划分, 并将其变换为一维特征编码.

$$z_c = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (1)$$

其中, z_c 是变量通道 c 输出; 特征图的高度和宽度分别为 H 和 W ; 通道 c 的坐标值为 $x_c(i, j)$. 使用长、宽为 $(H, 1)$ 、 $(1, W)$ 的池化核在水平和竖直坐标方位对每个通道进行编码, 相对高度为 h 的输

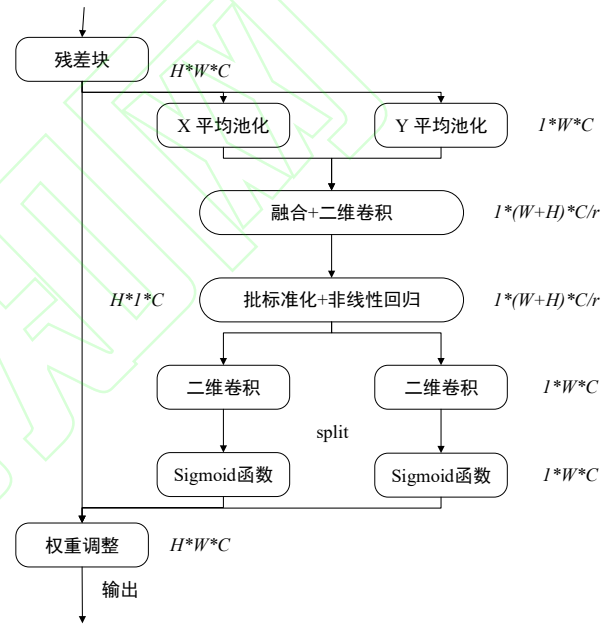


图2 CA注意力机制示意

Fig. 2 Schematic diagram of CA attention mechanism.

出可表述为

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < w} x_c(h, j) \quad (2)$$

其中, $z_c^h(h)$ 是通道 c 中相对高度为 h 的输出. 与通道注意力中转化成单独特征空间向量的控制模块方式不同, 式(1)和式(2)两种转换方式各自在2个空间上开展特征融合, 获得1对空间方位上感知的特征图. 这两种转换方式能够使实体模型捕获另一个空间方位的互相联系, 并维持从另一个空间方位的精准位置信息, 进而根据网络寻找合适的总体目标^[21-22], 可以取得比全局编码更确切的信息.

式(1)和式(2)两种输出结果, 应用到另一种转换, 完成CA注意力生成. CA注意力设计准则包括: ① 新的转化全过程应尽量简单; ② 注意力机制应能灵活运用捕获的信息, 使感兴趣的区域信息

可以被精确捕捉；③设计还需合理捕捉信息通道间的关联。特征信息转换完成后进行 Concatenate 变换，随后根据 1×1 卷积神经网络变换 F_1 函数公式进行变换，为

$$f = \sigma(F_1([z^h, z^w])) \quad (3)$$

其中， F_1 为卷积变换函数； $[\cdot, \cdot]$ 为 Concatenate 变换； σ 为空间信息在水平和垂直方位编码的中间特征投射，可用来操纵 CA 的控制模块尺寸减缩率。

1.2 CA-BIFPN 的特征金字塔网络

通常尺度较大的特征图更适于检测小目标物体，尺度较小的特征图更适合检测大目标物体，而将两者结合的特征金字塔则能更好兼具大小目标的检测。最早的特征金字塔网络(feature pyramid network, FPN)^[23]是一个自上而下的单项特征融合的金子塔结构，如图3。运算过程中，FPN网络的p6—p3输出层先将上一次特征图像的上采样结果进行特征融合，再经过与p7相同的运算流程，最终形成5个不同尺寸的特征图像，供后续网络检测。

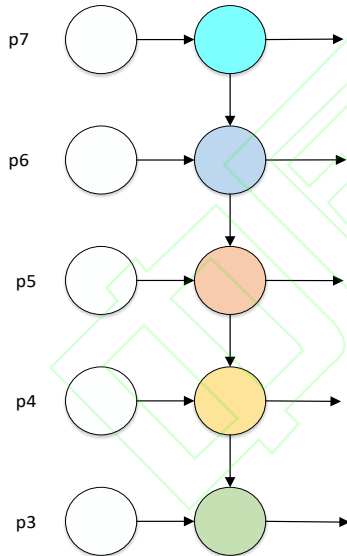


图3 FPN 网络结构

Fig. 3 (Color online) FPN network structure.

FPN可通过由顶向下的传递路径来获得语义更加丰富的特征，并将特征和更准确的位置信息结合。但针对由大目标产生的掩码，空间信息可能要传递数百层，导致传递路径过于冗长，使网络高层信息与低层信息无法有效结合，出现信息丢失。

路径聚合网络(path aggregation network, PANet)是一种改进的 FPN 网络，该网络通过使用由底层至顶层的横向连线来缩短传输路径，引入另一种由底向上的传输途径^[24]。PANet 相对于 FPN 具有更

好的网络精度，但其网络规模较大、参数量多，计算效率较低。TAN 等^[25]提出的 BIFPN 引入了跳跃连接，即在相同尺度的输入节点到输出节点中间再增加 1 个跳跃连接，因为运算处于相同层，该方法可在参数较少的同时结合更多特征。BIFPN 将每一条双向路径视作 1 个特征网络层，并对同一层参数进行多次计算，以实现更多的特征融合。

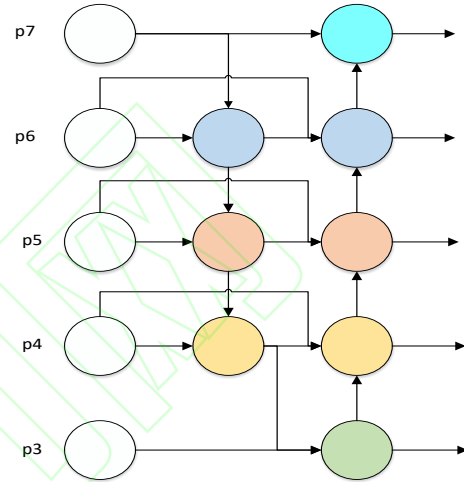


图4 BIFPN 网络结构

Fig. 4 (Color online) BIFPN network structure.

由于不同输入特征具有不同分辨率，因此，特征对于网络的影响也不同。针对无界融合、Softmax 基础融合以及快速归一化融合方法的比较结果表明，无界融合训练不稳定，需要使用权重归一化进行约束；Softmax 基础融合能表示出每个输入的重要程度，但会导致计算量增加；快速归一化融合与 Softmax 基础融合类似但没有将 Softmax 函数应用到参数中，因此，其运算准确度与 Softmax 基础融合类似，而计算速度较前者更快^[25]。本研究选择快速归一化融合方法作为 BIFPN 的融合算法。

2 实验和结果

2.1 实验数据集

本研究采用由腾讯公司与清华大学联合制定的数据集 TT100k，该数据集包含 9 176 张道路与交通标志牌图片。其中，6 105 张为训练集；3 071 张为测试集。每张图片都包括道路交通标志的一些案例，其中，训练集包括 16 527 个案例；测试集包括 8 190 个案例。涉及的道路交通标志牌类型有 221 种。由于 TT100K 中 221 种交通标志分布不均衡，部分交通标志如山体滑坡等相对罕见，为提高训练

效果, 本研究将所有数据进行清洗, 去除无标签和出现频次较少的交通标志种类, 筛选出 7 972 张超过 100 张图像的 45 种交通标志. 对清洗图像采用随机旋转和拓展数据, 包含放缩和增加高斯噪声^[26], 处理后的最终数据集共包括 23 916 张图像, 其中, 训练集 15 873 张; 测试集 8 043 张.

为测试模型在有雾环境下的检测效果, 对最终数据集随机使用红绿蓝(red green blue, RGB)通道合成雾、中心点合成雾及随机合成雾 3 种算法进行图像加雾处理, 制成雾天数据集进行辅助实验.

2.2 实验环境和评价指标

本实验操作系统为 Windows10, 训练使用的显卡为 NVIDIA RTX2070, 显存为 8 GByte, 深度学习框架为 Python3. 7、TensorFlow2. 0 和 Pytorch1. 8. 1. 为保证对比结果的客观性, 实验对不同网络进行测试时应用相同的超参数, 如表 1.

表1 超参数设置

Table 1 Hyperparameter settings

超参数	参数值
训练次数	700
初始学习率	0. 01
最终学习率	0. 2
优化器动量	0. 937
优化器权重衰减	0. 000 5

采用平均精度均值 mAP (mean average precision) 作为模型评价指标. mAP 通过统计模型成功检测数量与目标总数比值的高低来评价模型优劣, mAP 值越高, 表明目标检测模型在给定数据集上的检测效果越好. 通过选择不同的交并比 (预测框与真实框重叠面积与总面积之比) 阈值来统计不同标准下模型的精度均值, 常用的有 mAP0. 5 (交并比阈值为 0. 5) 和 mAP0. 5: 0. 95 (交并比阈值从 0. 5 到 0. 95, 步长为 0. 05). mAP 的计算过程为

$$\text{Precision}_{C_n} = \frac{N(\text{TruePositives})_{C_n}}{N(\text{TotalObjects})_{C_n}} \quad (4)$$

$$\text{AP}_c = \frac{\sum_{n=1}^{N(\text{TotalImages})_c} \text{Precision}_{C_n}}{N(\text{TotalImages})_c} \quad (5)$$

$$\text{mAP} = \frac{\sum_{c=1}^{N(\text{Classes})} \text{AP}_c}{N(\text{Classes})} \quad (6)$$

其中, Precision_{C_n} 表示图像 n 中, 模型对于类别为 C 的目标检测准确度; $N(\text{TruePositives})_{C_n}$ 为图像 n 中, 类别为 C 正确检测的目标数量; $N(\text{TotalObjects})_{C_n}$

为图像 n 中, 类别为 C 的目标数量; AP_c 为模型对类别为 C 的平均检测准确度; $N(\text{TotalImages})_c$ 表示存在类别为 C 的目标的图像数量, 如总共 20 张图像中, 有 10 张图像有类别为 C 的目标, 则 $N(\text{TotalImages})_c = 10$; $N(\text{Classes})$ 为类别数量.

2.3 实验结果分析

分别使用 SSD、YOLOv5、YOLOv5+CA 模块、YOLOv5+BIFPN 模块及 CA-BIFPN 模型的 mAP 作为实验评价指标值, 进行检测模型性能对比, 结果见表 2. 可见, 在相同参数训练条件下, 分别加入 CA 和 BIFPN 模块的 YOLOv5 模型 mAP0. 5 上升 0. 5% 和 0. 3%; CA-BIFPN 模型的 mAP0. 5 相比 SSD 和 YOLOv5 分别提升 4. 5% 和 1. 3%. CA-BIFPN 模型的 mAP0. 5: 0. 95 与 SSD 和 YOLOv5 相比分别提升 4. 1% 和 1. 2%. 在模拟雾天环境下, CA-BIFPN 模型精度为 80. 5%, 高于原模型的 80. 3%, 同时 CA-BIFPN 的 mAP0. 5: 0. 95 与原模型相比提升了 0. 6%.

表2 检测模型性能结果

Table 2 Model performance results

检测模型	mAP0. 5	mAP0. 5: 0. 95
SSD	79. 4	58. 2
YOLOv5	82. 6	61. 1
YOLOv5+CA	83. 1	62. 1
YOLOv5+BIFPN	82. 9	61. 5
CA-BIFPN	83. 9	62. 3
SSD(雾)	72. 1	55. 3
YOLOv5(雾)	80. 3	59. 5
CA-BIFPN(雾)	80. 5	60. 1

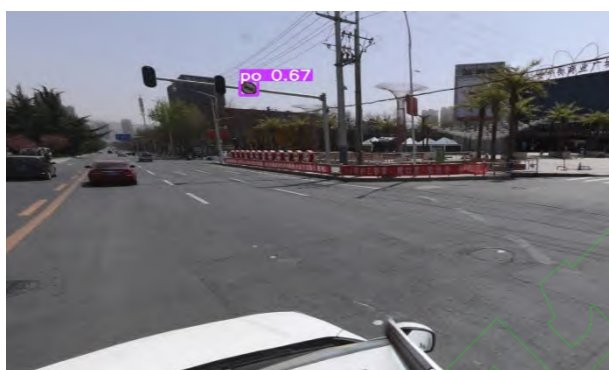
在 YOLOv5 中引入 BIFPN 模块并加入注意力机制后, 研究模型对数据集中小目标物体的检测精度, 从 45 类交通标志中分别选取 mAP0. 5 最高和最低的 2 类进行对比, 结果见表 3. 可见, 与 YOLOv5 相比, CA-BIFPN 模型各类小目标交通标志精度均有所提高, 表 3 中 4 种交通标志的 mAP0. 5 分别提高 2. 2%、2. 5%、0. 3%、0. 5%.

为进一步验证 CA-BIFPN 模型的可靠性, 从某市和某高速公路随机选取 5 张交通标志图像进行检测, 结果如图 5, 图中检测框上侧的字母和数字分别表示目标分类类型和概率. 其中, 图 5(a) 和 (b) 分别为 CA-BIFPN 模型城市道路中远、近目标检测结果; 图 5(c)、(e)(g) 为 CA-BIFPN、YOLOv5、SSD 模型高架道路中多目标远距离检测结果; 图 5

表3 TT100K中mAP0.5最高和最低的4类交通标志
Table 3 Value of mAP0.5 for the highest and lowest 4 types of traffic signs inTT100K %

检测模型	交通标志			
	禁止驶入	最低限速 100 km/s	限重 30 t	交通监控设备
SSD	92.1	91.3	61.4	60.8
YOLOv5	94.6	95.1	64.8	64.3
YOLOv5+CA	96.1	96.7	64.8	64.7
YOLOv5+BIFPN	94.9	96.4	65.0	64.3
CA-BIFPN	96.8	97.6	65.1	64.8

(d)、(f)、(h)为CA-BIFPN、YOLOv5、SSD模型高速公路中多目标近距离检测结果；图5(i)、(j)、(k)为CA-BIFPN、YOLOv5、SSD模型雾天交通标志检测检测结果。图5(a)和(b)表明，交通标志位于较远处时仍能被正确检测出；图5(d)、(f)、(h)表明，SSD发生了1处误检，CA-BIFPN模型与YOLOv5在近距离检测中无较大差距；图5(c)、(e)和(g)表明，仅CA-BIFPN模型实现了对小目标的检测，YOLOv5和SSD均出现不同程度漏检；图5(i)、(j)和(k)表明，本模型在部分雾天环境下也能实现对交通标志的有效检测。



(a)



(b)



(c)



(d)



(e)



(f)



图5 不同道路环境下 CA-BIFPN、YOLOv5 及 SSD 的交通标志检测结果 (a)CA-BIFPN 城市交通标志远距离检测;(b)CA-BIFPN 城市交通标志近距离检测;(c)CA-BIFPN 高架交通标志远距离多目标检测;(d)CA-BIFPN 高架交通标志近距离多目标检测;(e)YOLOv5 高架交通标志远距离多目标检测;(f)YOLOv5 高架交通标志近距离多目标检测;(g)SSD 高架交通标志远距离多目标检测;(h)SSD 高架交通标志近距离多目标检测;(i)CA-BIFPN 雾天交通标志检测;(j)YOLOv5 雾天交通标志检测;(k)SSD 雾天交通标志检测

Fig. 5 (Color online) CA-BIFPN, YOLOv5 and SSD traffic sign detection results in different road environments. (a)CA-BIFPN long-range detection of urban traffic signs, (b)CA-BIFPN proximity detection of urban traffic signs, (c)CA-BIFPN long-range multi-target detection of traffic signs on elevated roads, (d)CA-BIFPN proximity multi-target detection of traffic signs on elevated roads, (e)YOLOv5 long-range multi-target detection of traffic signs on elevated roads, (f)YOLOv5 proximity multi-target detection of traffic signs on elevated roads, (g)SSD long-range multi-target detection of traffic signs on elevated roads, (h)SSD proximity multi-target detection of traffic signs on elevated roads, (i)CA-BIFPN detection of traffic signs in foggy, (j)YOLOv5 detection of traffic signs in foggy, and (k)SSD detection of traffic signs in foggy.

结 语

交通标志检测对于智能交通意义重大,本研究提出 CA-BIFPN 交通标志检测模型,以 YOLOv5 为骨干网络,引入 CA 注意力机制聚焦小目标物体,在 Neck 网络中使用 BIFPN 加权双向特征金字塔,进一步提高了网络融合效果。基于 TT100K 数据集,将本模型与 SSD 和 YOLOv5 等经典目标检测模型进行比较检测实验,结果显示本模型的平均检测精度 mAP0.5 和 mAP0.5: 0.95 分别为 83.9% 和 62.3%,相比 YOLOv5 网络提高了 1.3% 和 1.2%,在交通标志检测中更可靠,在除极端雨雪气候及光照变化影响外的场景中均有较好的检测效果。本模型虽然提高了交通标志检测精度,但同时引入 CA 注意力机制以及 BIFPN 增加了模型的训练参数,降低了模型的运行速率,同时本模型缺少对于极端天气情况下的特殊检测优化,因此,后续研究中仍需要针对以上问题进一步改进,以提高模型检测效率。

基金项目: 甘肃省教育厅高等学校创新基金资助项目 (2020B-108); 甘肃省教育厅双一流重大科研资助项目 (GSSYLXM-04); 江苏省自然科学基金资助项目 (BK20200226); 苏州市科技计划资助项目 (SYG202033); 公安部道路交通安全重点实验室开放研究资助项目 (2021ZDSYSKF07); 山东省交通厅资助项目 (2020BZ01-03)

作者简介: 郎斌柯 (1997—), 兰州交通大学硕士研究生。研究方向: 交通信息收集与处理。E-mail: 854657021@qq.com

引 文: 郎斌柯, 吕斌, 吴建清, 等. 基于 CA-BIFPN 的交通标志检测模型 [J]. 深圳大学学报理工版, XXXX, XX (XX): 1-9.

参考文献 / References:

- [1] 李铭兴, 徐成, 李学伟. 交通标志识别研究综述[C]//中国计算机用户协会网络应用分会 2021 年第二十五届网络新技术与应用年会论文集. 北京: 中国计算机用户协会网络应用分会, 2021: 196-199.
LI Mingxing, XU Cheng, LI Xuewei. Summary of research on traffic sign recognition in urban traffic scenes [C]// Proceedings of the 25th Annual Conference on New Network Technologies and Applications, 2021, Network Application Branch of China Computer Users Association. Beijing: Network Application Branch of China Computer Users Association, 2021: 196-199. (in Chinese)
- [2] 陈飞, 刘云鹏, 李思远. 复杂环境下的交通标志检测与识别方法综述[J]. 计算机工程与应用, 2021, 57 (16): 65-73.
CHEN Fei, LIU Yunpeng, LI Siyuan. Survey of traffic sign detection and recognition methods in complex environment [J]. Computer Engineering and Applications, 2021, 57(16): 65-73. (in Chinese)
- [3] 杨晓玲, 江伟欣, 袁浩然. 基于 yolov5 的交通标志识别检测[J]. 信息技术与信息化, 2021(4): 28-30.
YANG Xiaoling, JIANG Weixin, YUAN Haoran. Recognition and detection of traffic signs based on yolov5 [J]. Information Technology and Informatization, 2021 (4): 28-30. (in Chinese)
- [4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2014: 580-587.
- [5] GIRSHICK R. Fast R-CNN [C]// 2015 IEEE International Conference on Computer Vision (ICCV). Piscataway, USA: IEEE, 2015: 1440-1448.
- [6] REN Shaoqing, HE Kaiming, 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.
- [7] CAI Zhaowei, VASCONCELOS N. Cascade R-CNN: delving into high quality object detection [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2018: 6154-6162.
- [8] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: single shot MultiBox detector [C]//The 14th European Conference Computer Vision - ECCV 2016. Amsterdam, The Netherlands: Springer, 2016: 21-37.
- [9] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection [C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, USA: IEEE, 2016: 779-788.
- [10] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, USA: IEEE, 2017: 6517-6525.
- [11] REDMON J, FARHADI A. YOLOv3: an incremental improvement [EB/OL]. [2022-05-10]. <https://doi.org/10.48550/arxiv.1804.02767>.
- [12] 寇大磊, 权冀川, 张仲伟. 基于深度学习的目标检测框架进展研究[J]. 计算机工程与应用, 2019, 55 (11): 25-34.
KOU Dalei, QUAN Jichuan, ZHANG Zhongwei. Research on progress of object detection framework based on deep

- learning [J]. Computer Engineering and Applications, 2019, 55(11): 25-34. (in Chinese)
- [13] CHEN Chenyi, LIU Mingyu, TUZEL O, et al. R-CNN for small object detection [C]// Asian Conference on Computer Vision - ACCV 2016. Taipei, China: Springer, 2017: 214-230.
- [14] 张宸嘉, 朱磊, 俞璐. 卷积神经网络中的注意力机制综述[J]. 计算机工程与应用, 2021, 57(20): 64-72. ZHANG Chenjia, ZHU Lei, YU Lu. A review of attention mechanisms in convolutional neural networks [J]. Computer Engineering and Applications, 2021, 57(20): 64-72. (in Chinese)
- [15] 鞠默然, 罗江宁, 王仲博, 等. 融合注意力机制的多尺度目标检测算法[J]. 光学学报, 2020, 40(13): 126-134. JU Moran, LUO Jiangning, WANG Zhongbo, et al. Multi-scale target detection algorithm based on attention mechanism [J]. Acta Optica Sinica, 2020, 40(13): 126-134. (in Chinese)
- [16] 徐诚极, 王晓峰, 杨亚东. Attention-YOLO: 引入注意力机制的YOLO检测算法[J]. 计算机工程与应用, 2019, 55(6): 13-23. XU Chengji, WANG Xiaofeng, YANG Yadong. Attention-YOLO: YOLO detection algorithm that introduces attention mechanism [J]. Computer Engineering and Applications, 2019, 55(6): 13-23. (in Chinese)
- [17] 任欢, 王旭光. 注意力机制综述[J]. 计算机应用, 2021, 41(S1): 1-6. REN Huan, WANG Xuguang. Review of attention mechanism [J]. Journal of Computer Applications, 2021, 41(S1): 1-6. (in Chinese)
- [18] 刘玉红, 陈满银, 刘晓燕. 基于通道注意力的多尺度全卷积压缩感知重构[J/OL]. 计算机工程: 1-8 [2022-04-29]. DOI:10.19678/j.issn.1000-3428.0063546. LIU Yuhong, CHEN Manyin, LIU Xiaoyan. Multi-scale full convolution compressed perceptual reconstruction based on channel attention [J/OL]. Computer Engineering: 1-8 [2022-04-29]. DOI: 10.19678/j.issn.1000-3428.0063546. (in Chinese)
- [19] 许德刚, 王露, 李凡. 深度学习的典型目标检测算法研究综述[J]. 计算机工程与应用, 2021, 57(8): 10-25. XU Degang, WANG Lu, LI Fan. Review of typical object detection algorithms for deep learning [J]. Computer Engineering and Applications, 2021, 57(8): 10-25. (in Chinese)
- [20] HOU Qibin, ZHOU Daquan, FENG Jiashi. Coordinate attention for efficient mobile network design [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, USA: IEEE, 2021: 13708-13717.
- [21] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2018: 7132-7141.
- [22] 刘学平, 李珂乾, 刘励, 等. 嵌入SENet结构的改进YOLOV3目标识别算法[J]. 计算机工程, 2019, 45(11): 243-248. LIU Xueping, LI Yuqian, LIU Li, et al. Improved YOLOV3 target recognition algorithm with embedded SENet structure [J]. Computer Engineering, 2019, 45(11): 243-248. (in Chinese)
- [23] KIM S W, KOOK H K, SUN J Y, et al. Parallel feature pyramid network for object detection [C]// European Conference on Computer Vision - ECCV 2018. Munich, Germany: Springer, 2018: 239-256.
- [24] LIU Shu, QI Lu, QIN Haifang, et al. Path aggregation network for instance segmentation [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, USA: IEEE, 2018: 8759-8768.
- [25] TAN Mingxing, PANG Ruoming, LE Q V. EfficientDet: scalable and efficient object detection [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, USA: IEEE, 2020: 10778-10787.
- [26] 杨真真, 匡楠, 范露, 等. 基于卷积神经网络的图像分类算法综述[J]. 信号处理, 2018, 34(12): 1474-1489. YANG Zhenzhen, KUANG Nan, FAN Lu, et al. Review of image classification algorithms based on convolutional neural networks [J]. Journal of Signal Processing, 2018, 34(12): 1474-1489. (in Chinese)

【中文责编: 方圆; 英文责编: 淡紫】