

基于改进 YOLOv5s 算法的电动车头盔佩戴检测

陆万浩, 黄孙港[†], 饶兴昌, 孙曾阳

(南京信息工程大学电子信息与工程学院, 江苏 无锡 214105)

摘 要:为高效检测出路上电动车骑行者不按规定佩戴头盔的情况,提出了一种基于 YOLOv5s 模型的电动车头盔佩戴检测的改进算法。首先,该算法以 YOLOv5s 模型为基础,以 C3Ghost 替代原有的 C3 模块使得改进后的模型更加轻量化以减少参数量;其次通过嵌入 Biformer 注意力机制,并且将颈部网络中的 Conv 模块用 GSconv 代替;最后引入改进后的损失函数 EIOU。通过消融实验对比后发现,改进后的模型算法平均精度均值(mAP)提高到了 95.2%,精确率(Precision)提高到了 95.6%。与 YOLOv5s 相比提升了 16.9%和 17%,同时浮点运算量下降 25.5%左右。实验结果说明,改进后的 YOLOV5s 算法可以有效提升电动车骑行者头盔佩戴情况的检测性能。

关键词:YOLOv5; 电动车头盔检测; 轻量化; Biformer

中图分类号:TP39

文献标识码:A

Research on Helmet Wearing Detection of Electric Vehicle with Improved YOLOV5s Algorithm

LU Wanhao, HUANG Sungang[†], RAO Xingchang, SUN Zengyang

(School of Electronic Information Engineering, Nanjing University of Information Science and Technology, Wuxi, Jiangsu 214105, China)

Abstract: In order to efficiently detect the situation of electric vehicle riders not wearing helmets according to regulations, an improved algorithm based on YOLOv5s model was proposed for helmet wearing detection of electric vehicles. Firstly, The algorithm is based on YOLOv5s model and replaces the original C3 module with C3Ghost to make the improved model more lightweight and reduce the number of parameters. Secondly, by embedding Biformer attention mechanism and replacing Conv module in neck network with GSconv. Lastly, the improved loss function EIOU is introduced. Through the comparison of ablation experiments, it is found that the average accuracy (mAP) and Precision of the improved model algorithm are increased to 95.2% and 95.6% respectively. Compared with YOLOv5s, it has improved by 16.9% and 17%, while floating point arithmetic has decreased by about 25.5%. The experimental results show that the improved YOLOV5s algorithm can effectively improve the detection performance of helmet wearing of electric vehicle riders.

Key words: YOLOv5; electric vehicle helmet detection; lightweight; Biformer

众所周知,电动车在我们现实生活中扮演着重要角色,作为成年人可以使用的出行工具,其给人们的出行提供了巨大的便利。但由于骑行电动车不需要经过系统的考核、培训和驾照证明,骑行者

的交通安全意识和规则意识普遍不高,这也导致了许多交通事故的发生。据统计,80%的电动车事故死亡原因都是颅脑损伤,而较低的头盔佩戴率则是导致颅脑损伤的关键原因,佩戴头盔可以有效降

低交通事故的伤害。随着“一盔一带”主题活动的开展,公安部门加强执法管理,对骑行电动车、摩托车不戴头盔的人依法查纠。然而在路口设置交警进行人工检查的方式,用检查数以万计的电动车骑行者是否佩戴头盔费时费力,效率低的同时也很难有显著的效果。因此,将深度学习和目标检测算法应用于道路交通监管中检查骑行者头盔佩戴情况具有很理想的现实意义。

目标检测算法主要分为两类:一是以 R-CNN 为代表的二阶段目标检测算法^[1];第二类是基于回归框架的一阶段检测算法,如 SSD、YOLO 算法等。一阶段目标检测算法只需要提取一次特征即可实现目标检测,检测速度更快且算法结构简单,能够实现目标的实时监测。由于本文的头盔佩戴检测系统对速度和实时性有较高的要求,因此选择了检测速度更快的 YOLOV5 算法。

1 YOLOv5 算法原理

YOLOv5 是一种检测速度较快、精度高同时还具有实时性的一阶段目标检测算法,它主要由 input、backbone、neck 组成,而 backbone 和 neck 是 YOLOV5 网络结构中的重要组成部分。YOLOV5 中各个部分的主要功能结构都不同。输入端采用 Mosaic 数据增强^[2]的方式丰富数据集,同时也减少了 GPU 的使用,以此获得更好的训练效果。在主干网络中采用了 Focus 和 CSP 结构。Focus 结构对普通卷积操作进行了优化,通过降低计算成本的方式实现了下采样,降低了参数量的同时提升了速度。在 Backbone 主干网络中的 CSP 结构是 CSP1_X 结构,卷积核大小为 3×3 ,步进值为 2。假如输入的图像尺寸为 640×640 ,那么经过一系列的操作后,输出的特征图像尺寸为 20×20 。此外在 neck 部分还添加了 CSP2_X 结构,极大地减少了运算量。在保持了轻量化的基础上,提高了准确性。在 neck 网络部分采用了 FPN 与 PAN 结构。FPN 通过上采样,自上而下地对不同层的输出进行特征聚合。PAN 则是通过自下而上的特征金字塔传递定位信息。最后在输出层中主要改进了训练时的损失函数。YOLOv5 的结构图如图 1 所示。

2 改进的 YOLOV5 网络模型

2.1 整体网络结构

YOLOv5 具有较强的灵活性,其具有 4 种不

同的网络模型可供选择。本文选择 YOLOV5s 网络模型作为基础网络,其是 4 种网络模型中网络最小、速度最快、最轻量化的模型。在此基础上提出了一种改进的电动车头盔检测算法。改进后的 YOLOV5s 模型结构图如图 2 所示。选择的电动车头盔佩戴检测数据集是由 MS COCO 大型目标检测数据集和网络中收集的非机动车道交通情况的数据集组合而成。由于原有的网络存在大量普通卷积,会严重影响对目标特征信息的提取,因此选择在主干网络中融入 C3Ghost 模块,加入后使得模型轻量化的同时也不会对检测精度造成过大的影响,同时对网络嵌入 Biformer 注意力机制^[3],作为一种双层路由注意力机制,它可以完成对远距离目标的捕获和在对小目标物体检测上有明显的提升。接着将 GSconv 替换颈部网络中的 conv,进一步提升模型的检测速度。最后加入改进的损失函数 EIOU,考虑了边界框回归的重叠面积、中心点距离和长宽边长真实差,能够解决图像中存在的遮挡、漏检等问题。

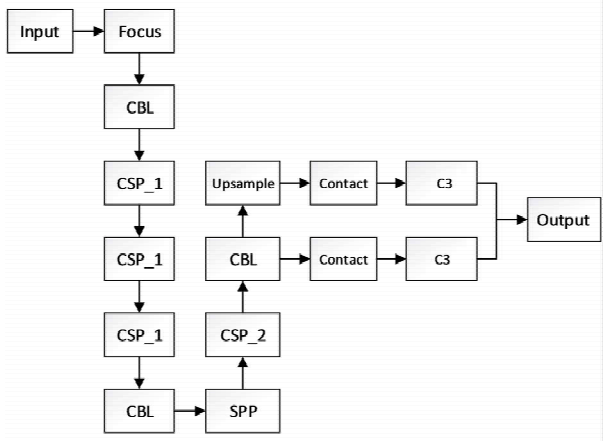


图 1 YOLOv5 整体模型结构图

2.2 Ghost 模块

本文中选用的 Ghost 是 C3ghost 模块,作为一种即插即用的模块,它能够快速配合组合设计出新的模块。其结构图如图 3 所示。主要目的是加入 Ghost 模块后使得模型更加轻量化。Ghost 模块的原理是对卷积操作进行改进,它可以用来替换任何经典 CNN 网络以得到更多的特征图,消除冗余特征,使得模型更轻量化。本文将原始 YOLOv5s 网络中的 C3 模块替换成 Ghost 模块,融合 Ghost 模块后,训练模型的浮点数和参数量均有所下降。

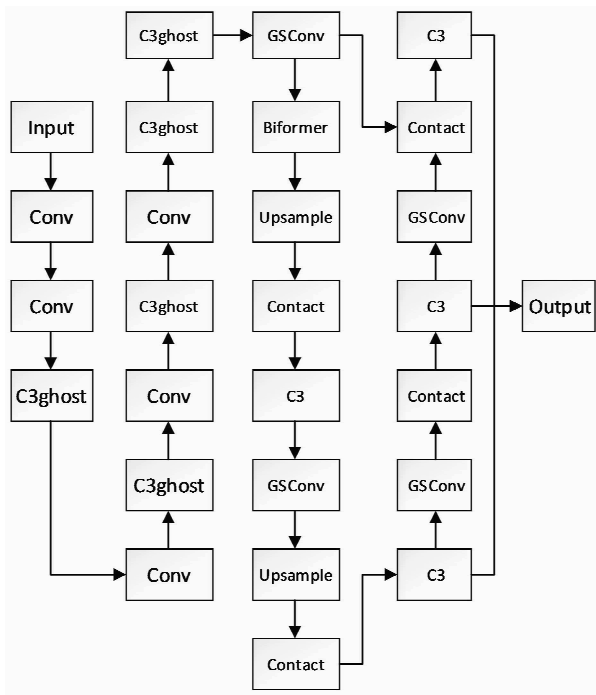


图 2 改进后的 YOLOV5s 模型结构图

2.3 Biformer 注意力机制

注意力机制是一种能够对各个位置的重要性进行分析和排位,选择出相对关键的信息同时将更多的计算资源投入关键的信息中的一个模块。当不添加注意力机制的时候,如果有输入图片进入到该网络模型中,而图

片中的部分数据与待检测的目标关系不大,但是却会占据一样的资源,这会导致运算效率和精度的降低。因此就需要添加注意力机制来改善这一问题。传统注意力机制有压缩激励网络 (SENet)^[4], SENet 为通道注意力机制,它是一种传统意义上的注意力机制,用最简单的方式来获得更好的检测效果。Hou 等^[5]提出的注意力机制 (Coordinate Attention, CA),在 SENet 的基础上获得了改进,提高了定位能力。加入该注意力机制后,能够使得模型更快更高效地找出关键信息的位置,用最直接的方式提高了检测效果。卷积注意力模块 (Convolutional Block Attention Module, CBAM)^[6]是由 Woo 等提出,相比于 SE 注意力模块增加了空间注意力机制,在结构上完美实现了通道注意力和空间注意力机制的融合。由于 YOLOv5s 网络中输入的数据集可能会存在模糊、分辨率低的情况,因此为了提高检测精度,对原网络嵌入了 Biformer 注意力机制。本文选用的 Biformer 注意力机制是在 2023 年 CVPR 推出的新注意力机制,作为一种双层路由注意力机制,它可以通过动态、查询感知的方式实现计算的有效分配,它是通过首先构建和修剪区域级有向图,然后在路由区域的联合中应用细粒度的 Token-to-Token 注意力来实现的。

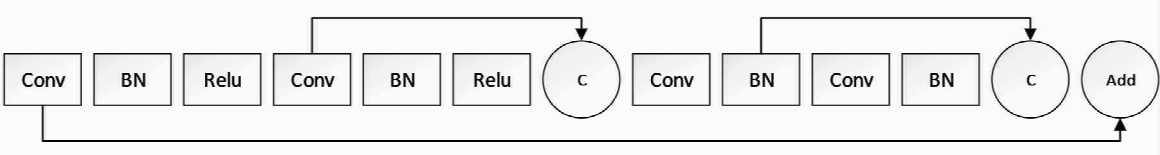


图 3 Ghost 模块结构图

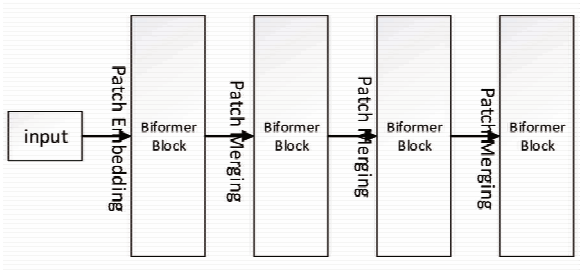


图 4 Biformer 结构

如图 4 所示,Biformer 注意力机制是 Biformer 第一阶段使用重叠块嵌入,在第二到第四阶段使用块合并来降低输入空间分辨率,同时增加通道数,然后采用连续的 Biformer 块做特征变换。

2.4 GSConv 模块

GSConv^[9]是一种能够降低模型计算量的同时还能保持甚至提高检测精确度的模块。在 Neck 网络中使用 GSConv 能够使得预测的结果更加迅速,实时性更强。本文选择 GSConv 模块以及 C3ghost 模块分别添加在 neck 和 backbone 网络中,提高了模型的轻量化和实时性,同时也保持了模型本身的精度,以此能够更好实现对电动车骑行者佩戴头盔的检测情况的实时关注。

2.5 损失函数改进

通过对各个不同领域的文献进行阅读和分析后,发现大部分 YOLOv5 中选择采用 GIoU 损失函数,本文中由于涉及有检测目标遮挡的问题,而 GIoU 无法判断目标框与预测框重叠的情况。因

此本文选择 EIoU^[10] 损失函数,以解决图像检测中存在的遮挡、漏检等问题。EIoU 是在 CIoU 的基础上分别计算宽高的差异值取代了纵横比,通过这种方式使得预测框和真实框的宽度差值最小,从而使得收敛速度更快。EIou 计算式为:

$$L_{\text{EIou}} = 1 - \text{IoU} + \frac{p^2(b, b^{\text{gt}})}{c^2} + \frac{p^2(w, w^{\text{gt}})}{c_w^2} + \frac{p^2(h, h^{\text{gt}})}{c_h^2}$$

式中, p 表示两点之间的欧氏距离; c_w 和 c_h 分别为覆盖两个框的最小外接框的宽度和高度; b, w, h 分别为预测框的中心位置、宽、高; $b^{\text{gt}}, w^{\text{gt}}, h^{\text{gt}}$ 分别为真实框的中心位置、宽、高。改进后的 EIoU 函数考虑了损失函数能够关注到边界框回归的重叠部分面积、目标点到中心点距离和长宽边长真实差,能够使模型获得更好的训练效果。

3 实验设备与结果分析

3.1 实验环境

主机操作系统选用的是 Windows10 专业版系统,电脑型号为戴尔 G15 5510,其中 CPU 处理器为 11th Gen Intel(R) Core(TM) i5-11260H @ 2.60GHz , GPU 配备的是 NVIDIA GeForce RTX3050Ti Laptop GPU,选用的编程语言为 Python3.8、Python 框架为 Pytorch2.0.0。本实验中选用的数据集是通过网络选取、道路拍摄以及视频剪裁等方式选出符合要求的图片共计 1695 张。通过使用 labelimg 标注工具对数据集集中的图片进行

标注,标注类别分为 helmet、without helmet 以及 two-wheeler。

3.2 评价指标

由于加入不同模块会对模型整体产生不同的效果,为了体现出这些模块对实验结果的具体影响,选择进行消融实验。通过消融实验可以从数据上更直观体现该改进模型的效果。综上所述,本实验选择通过比较准确率(Precision)、平均准确率均值(mAP)来反映实验效果。准确率 Precision 的公式如(1)所示,平均准确率均值 mAP 公式如(3)所示。其中,TP 表示正确检出的正样本,FP 表示错误检出的正样本,num 表示检测类别总数。同时将每秒浮点运算次数(FLOPs)作为衡量模型大小和计算量的指标。

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{1}$$

$$\text{AP} = \int_0^1 P dR \tag{2}$$

$$\text{mAP} = \frac{\sum_{i=1}^n \text{AP}_i}{mm} \tag{3}$$

3.3 消融实验

在本实验中,对改进后的 YOLOv5s 模型进行消融实验,将得到的数据进行分析,考虑不同模块加入不同位置后会对模型整体产生的影响。本消融实验以 YOLOv5s 网络作为基础网络,将 C3ghost 模块、Biformer 注意力模块和 GSConv 模块分别加入网络模型,最后得到的消融实验结果见表 1。

表 1 消融实验结果

模型	mAP	Precision	Time	FLOPs
YOLOv5s	81.4%	81.7%	1.93h	18.4G
YOLOv5s+C3ghost	80.7%	81.1%	1.69h	14.0G
YOLOv5s+CG+BIF	90.6%	90.8%	1.73h	15.6G
YOLOv5s+CG+BIF+GS	94.1%	94.3%	1.43h	13.7G
Previous+EIOU	95.2%	95.6%	1.43h	13.7G

由表 1 可以看出,加入轻量化网络 Ghost 模块后,浮点数和参数量均有所下降。在 YOLOv5s 网络的基础上用 C3ghost 替换 C3 模块后,C3ghost 模块能够在降低参数量和 FLOPS 的同时,保持精度不下降,提高了模型的轻量化。

在 Backbone 部分中加入 Ghost 模块使得模型的 mAP 仅下降 0.7%,与原网络检测效果基本持平。在此基础上嵌入 Biformer 注意力机制后,mAP 对比 ghost 轻量化网络模型增加了 10 个百

分点,说明加入 Biformer 注意力机制能够提升模型检测效果和精度。在此基础上,在颈部网络中将 GSConv 模块替换原来的 Conv 模块后,在精确率和 mAP 上对比原网络提升了 12 个百分点,且浮点数降低了 25.5%。在此基础上改进损失函数 EIoU,可以发现精确率提高 1.3%,mAP 提高 1.1%,且浮点运算量不变,可以发现改进损失函数并没有增加模型的运算量。

3.4 实验结果展示

为了检测改进后的 YOLOV5s 模型的实际应用效果,本文选择在马路上拍摄视频并截取部分片段进行测试,见图 5。

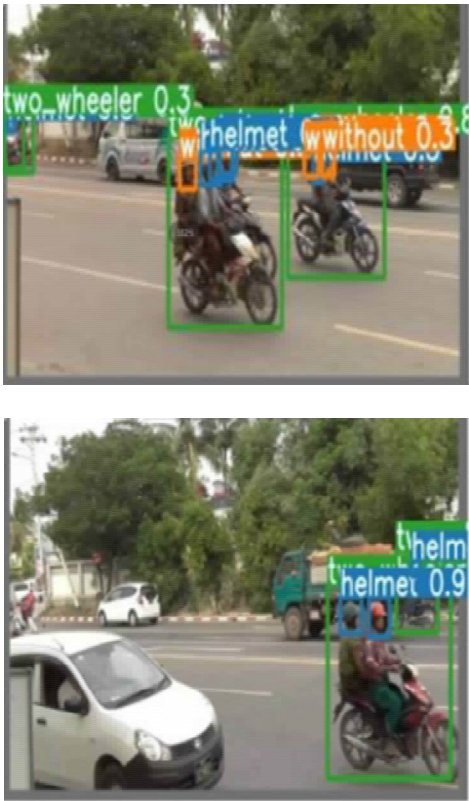


图 5 道路拍摄检测结果

如图 6 所示,如果图片中存在多个目标,模型也可以完成对目标的精确检测。左侧图片中,被遮挡的目标也可以检出,右侧图片中,远处的目标人物也可以被检出。改进后的 YOLOV5 模型中加入了 Biformer 注意力机制,能够很好地识别小目标物体。同时加入了改进的损失函数 EIOU,关注到边界框回归的重叠部分面积、目标点到中心点距离和长宽边长真实差,能够解决图像中存在的遮挡、漏检等问题。根据图 6 检测情况对比,可以进一步说明了改进模型的优越性。



(a)改进前



(b)改进后



(c)改进前



(d)改进后

图 6 检测情况对比

4 结 论

考虑到电动车骑行者头盔佩戴检测任务难度,选用了检测速度快、精度高且具有实时性的 YOLOv5s 作为基准模型。同时针对实际生活中会出现的人物遮挡以及近大远小的情况,提出了一种改进的 YOLOv5s 网络模型。在原始的 YOLOv5s 网络中用 C3ghost 替代原始网络中的 C3 模块使得模型更轻量化,同时加入 Biformer 注意力机制来提高模型检测的精确度,最后加入损失函数 EIou,在不增加计算量的基础上提高了训练精确率。经过改进后的 YOLOv5s 模型 mAP 提高 16.9%,精度提高 17%,浮点运算量降低 25.5%,能够完成现实生活中对电动车骑行者头盔佩戴情况的检测,有助于配合交警管理以及“一盔一带”主题活动的开展。

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[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137—1149.

[2] CHEN J, BAO E, PAN J Y. Classification and positioning of circuit board components based on improved YOLOv5[J]. Procedia Computer Science, 2022, 208: 613—626.

[3] ZHU L, WANG X J, KE Z H, et al. BiFormer: vision transformer with bi-level routing attention[C]//Conference on Computer Vision and Pattern Recognition, 2023.

[4] SOEKARNO I, HADIHARDAJA I K, CAHYONO M. A study of hold-out and k-fold cross validation for accuracy of groundwater modeling in tidal lowland reclamation using extreme learning machine[C]//Proceedings of the 2014 2nd International Conference on Technology, Informatics, Management, Engineering & Environment. Piscataway, NJ: IEEE, 2014: 228—233.

[5] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[C]//Computer Vision and Pattern Recognition, 2021: 21.

[6] WOO S, PARK J, LEE J Y, et al. Cbam: convolutional block attention module[C]//Proceedings of the European conference on computer vision (ECCV). Berlin: Springer, 2018: 3—19.

[7] 王德雨, 石伟, 张元良. 一种融合 CBAM 注意力机制和金字塔卷积的经济鱼类识别方法[J]. 江苏海洋大学学报(自然科学版), 2023(3): 1—8.

[8] ZHANG K P, ZHANG Z P, QIAO L Y. Joint face detection and alignment using multi-task cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23(10): 1499—1503.

[9] 刘雅洁, 伊力哈木·亚尔买买提, 席凌飞, 等. 改进 YOLOv5s 的安全帽佩戴检测算法研究[J]. 计算机工程与应用, 2023, 59(20): 184—191.

[10] ZHANG F, REN W Q, ZHANG Z, et al. Focal and efficient IOU loss for accurate bounding box regression[J]. Horizon Robotics, 2022: 1—10.

(上接第 131 页)

参考文献

[1] 施晓勇, 蓝贵文, 陈看, 等. 基于需求驱动的输变电工程多尺度场景建模[J]. 科学技术与工程, 2023, 23(5): 1839—1845.

[2] 李一鹏, 王亚军, 栗维勋, 等. 基于极限学习机的变电站监控系统负荷异常信息的辨识方法[J]. 电子器件, 2022, 45(5): 1219—1224.

[3] 王周虹, 屠雨夕, 王海园, 等. 基于全通信链路的变电站调控交互数据监测方法[J]. 南京理工大学学报, 2022, 46(4): 451—459.

[4] ZHAO J D, YUE X Z. Condition monitoring of power transmission and transformation equipment based on industrial internet of things technology[J]. Computer Communications, 2020, 157(5): 204—212.

[5] 祝鹏, 郭艳光. 基于 K-medoids 聚类算法的多源信息数据集成算法[J]. 吉林大学学报(理学版), 2023, 61(3): 665—670.

[6] 张亚平, 马唯婧, 张宸硕, 等. 基于图像识别与 CAE 仿真技术的输变电塔一体化分析[J]. 山东大学学报(工学版),

2023, 53(6): 122—130.

[7] 张建功, 鲁浩男, 万皓, 等. 特高压交流输变电工程工频电场测量及不确定度分析[J]. 水电能源科学, 2023, 41(6): 202—205+187.

[8] 单瑞卿, 盛阳, 苏盛, 等. 考虑攻击方身份的电力监控系统网络安全风险分析[J]. 电力科学与技术学报, 2022, 37(5): 3—16.

[9] 张华兵, 周英耀, 徐磊, 等. 基于微服务架构的电力信息系统全链路监控技术[J]. 沈阳工业大学学报, 2022, 44(4): 409—414.

[10] 贾鹏, 王平辉, 陈品安, 等. 基于无监督学习的智能数据中心电力拓扑系统[J]. 清华大学学报(自然科学版), 2023, 63(5): 730—739.

[11] 逄宝中, 李庚银, 王剑晓, 等. 计及监测与控制功能的电力信息物理系统关键输电线路辨识方法[J]. 中国电机工程学报, 2022, 42(7): 2556—2566.

[12] 钟启迪, 蒲莹, 孙建锋, 等. 适用于柔性直流电网的 500kV 级直流电缆与架空线路并联运行控制及保护策略[J]. 电力建设, 2022, 43(7): 80—86.

[13] 张朝辉, 罗炜, 林康熙, 等. 基于时序分析的电网合并单元电平预测[J]. 数据采集与处理, 2022, 37(5): 1169—1178.

[14] 肖异, 刘超, 马怡晴, 等. 海上风电柔性直流送出一体化监控系统方案设计[J]. 南方电网技术, 2022, 16(12): 126—133.