

李波,周孙源,杨茜. 基于深度学习的“雷视一体化”技术研究[J]. 智能计算机与应用,2025,15(2):111-115. DOI:10.20169/j. issn. 2095-2163. 250217

基于深度学习的“雷视一体化”技术研究

李波¹,周孙源²,杨茜¹

(1 西安雷通科技有限责任公司,西安 710119; 2 西北农林科技大学理学院,陕西 咸阳 712100)

摘要:在智能车辆设计和智能交通系统建设过程中,对车辆目标进行检测与识别至关重要。本文使用基于CFAR算法的目标检测,基于YOLOv5神经网络的目标检测与识别,并依据多传感器决策级融合原理,对毫米波雷达数据与视觉传感器数据进行空间融合、时间融合,最后基于全局最近邻的数据关联算法(GNN)完成了两种传感器的数据关联,实现了时空以及数据层面上的“雷视一体化”,从而完成了毫米波雷达与视觉传感器融合的车辆检测系统的搭建。为使对比结果更加直观,该文选择KITTI数据集作为目标检测算法的训练集,并采集多种不同条件下的车辆行驶图片进行测试。实验结果表明,相较于单传感器检测,“雷视一体化”技术可在天气条件较好时提供更详细的前方车辆信息,在天气条件差(如大雾天气)时提供视觉传感器无法检测到的前方车辆信息,由此体现出多传感器融合技术具有极大优势以及进一步研究发展的必要性。

关键词:雷视一体化;YOLOv5;数据融合;目标检测与识别;智能交通

中图分类号:TN957.51

文献标志码:A

文章编号:2095-2163(2025)02-0111-05

Study on "Radar and camera fusion" technology based on deep learning

LI Bo¹,ZHOU Sunyuan²,YANG Xi¹

(1 Xi'an LeiTong Science and Technology Co., Ltd., Xi'an 710119, China;

2 College of Science, Northwest A&F University, Xianyang 712100, Shaanxi, China)

Abstract: In the process of intelligent vehicle design and intelligent transportation system construction, the detection and identification of vehicle targets is very important. This paper uses target detection based on CFAR algorithm, target detection and recognition based on YOLOv5 neural network, and spatial fusion and time fusion of millimeter wave radar data and visual sensor data according to the principle of multi-sensor decision-level fusion, and finally completes the data association of the two sensors based on the data association algorithm (GNN) of the global nearest neighbor, realizing the "Radar and camera fusion" at the spatiotemporal and data level. Thus, the construction of a vehicle detection system that integrates millimeter-wave radar and visual sensors was completed. In order to make the comparison results more intuitive, this paper selects the KITTI dataset as the training set of the target detection algorithm, and collects a variety of vehicle driving pictures under different conditions for testing. The experimental results show that compared with single-sensor detection, "Radar and camera fusion" technology can provide more detailed information about the vehicle in front when the weather conditions are good, and provide the information of the vehicle in front that cannot be detected by the visual sensor when the weather conditions are poor (such as foggy weather), which reflects the great advantages of multi-sensor fusion technology and the necessity of further research and development.

Key words: Radar and camera fusion; YOLOv5; data fusion; object detection and recognition; intelligent transportation

0 引言

近年来,随着人民生活水平的不断提高,车辆也不断增多,并成为了交通道路上的重要元素,因此车辆目标检测与识别便成为了智能交通系统中不可或缺的一环,建设更加全面、精准的智能交通系统至关

重要。目前,在车辆目标检测与识别领域,视频监控发挥着主力作用,视频监控能获取到车辆的颜色、牌照等物理特征信息,但在现实情况中,由于透视效应、天气变化和恶劣环境等不利因素,视频监测效果不佳,可能存在大量的误报、漏报情况。相比于视频监控,雷达的监测范围更广,且不受透视效应和天气

作者简介:周孙源(2000—),女,硕士研究生,主要研究方向:数论,解析论,密码学等;杨茜(1988—),女,硕士,高级工程师,主要研究方向:毫米波雷达系统,测速雷达系统,智慧安防系统等。

通信作者:李波(1973—),男,学士,高级工程师,主要研究方向:要地安防及边境管控,无人机侦测反制。Email:578084126@qq.com。

收稿日期:2023-08-18

哈尔滨工业大学主办 ◆ 专题设计与应用

环境等因素的影响,能更准确地监测目标的位置、距离、速度等,但雷达监测横向精度低,不能精准识别目标的物理特征信息。

随着深度学习技术的不断发展与创新,如何开发一种监控数据全面、精确且易于安装和维护的“雷视一体化”道路车辆监测系统的问题亟待解决。“雷视一体化”技术将视频监控和雷达监测融合应用,在识别目标物理特征的同时,精准监测目标的位置、速度等,最大程度实现了这两大技术方案的优势互补,打开了监测设施升级的新思路。

本文通过毫米波雷达得到车辆的位置、速度等数据,并进行进一步的解析及过滤,得到有效目标;接着利用基于深度学习的目标检测算法,对车辆进行目标检测与识别;再通过多种坐标系的转换选取适宜的采样周期,完成毫米波雷达与视觉传感器数据的时空融合;最后把处理过的数据进行关联送入决策终端,输出得到车辆目标的定位及物理特征信息。

1 研究方法

“雷视一体化”是一种将雷达与视频监控技术融合起来的新兴科技,在近几年备受关注。人类主要通过视觉、听觉等感官来感知世界,而对机器来说要完成对视觉、听觉的理解,主要依靠深度学习技术^[1],而图像处理主要采用卷积神经网络。Tristan V^[2]通过研究表明,毫米波雷达通过检测获取并处理后的研究数据所形成的毫米波图像,经过卷积神经网络后,可以用于目标分类;张暄^[3]提出一种基于卷积神经网络的雷达目标检测方法,以雷达回波信号数据处理后得到的距离-多普勒图像作为模型的训练集和测试集,设计了基于 Faster R-CNN^[4]结构的雷达目标检测模型;董美琳^[5]将 DeepSORT^[6]算法中的表观特征提取方法修改为 HOG 特征提取方法,将 YOLOv4 深度目标检测网络与 DeepSORT 算法结合,最终对高速公路上的车辆出现的逆行、超速、行人闯入等交通违法行为进行识别;Chen B^[7]以雷达和摄像机为研究对象,采用目标级融合层次,利用区域碰撞关联和加权航迹关联设计两级关联结构,完成两个传感器局部航迹之间的关联,并使用非重置联邦滤波器估计融合航迹的状态。

目前,适用于图像目标检测算法的卷积神经网络有很多,主要分为两阶段和一阶段两大类。“两阶段”算法首先将所有物体框出来,再利用基于卷积神经网络的图像分类算法对物体进行分类,Ross Girshick 等^[8]提出的 R-CNN 是第一篇利用卷积神

经网络并基于候选区域的二阶段目标检测算法,除此之外二阶段算法在 R-CNN 基础上进行改进的 Fast-RCNN^[9]、Faster-RCNN 和 Mask-RCNN^[10]等;“一阶段”算法则直接对图片进行处理,同时输出物体的位置和类别,例如 YOLO 系列算法以及 SSD 系列算法^[11]等。其中,YOLO 系列算法经过不断地研究与改进,其精度与泛用性也在不断提升。YOLOv3^[12]在 YOLOv2^[13]的基础上,改良了网络的主干,利用多尺度特征图进行检测,改进了多个独立的 Logistic regression 分类器,取代 Softmax 来预测类别分类;而 YOLOv5 与之前的 YOLOv3、YOLOv4^[14]不同,其是该系列中深度最小,并且特征图的宽度也最小的网络。

雷达通过天线发射并接收从目标反射回来的电磁波,利用时间差来计算与目标之间的距离或其他参数。目前,有多种雷达可应用于车辆监测,如 K 波段雷达、毫米波雷达、激光雷达等^[15]。由于毫米波雷达稳定性高且成本低,具有较高的性价比,所以国内大部分采用毫米波雷达作为传感器。魏宏安^[16]提出基于毫米波雷达的车辆行为检测,利用毫米波雷达获得感兴趣区域内车辆方位和速度信息后,依据 ROI 内车辆行驶轨迹分布和速度状态判定车辆驾驶行为,制定合理的检测算法;高淑雅^[17]提出,对雷达回波数据进行二维傅里叶变换得到距离-多普勒图像,再以距离-多普勒图集作为数据集训练神经网络,得到能够完成雷达目标识别的网络模型。

“雷视一体化”可以提高目标检测、跟踪、分类等能力,在不同环境条件下都有较好的表现。吴建波^[18]以延庆至崇礼段高速为依托设计的高速公路路况感知系统,虽同时使用了雷达与视频对道路信息进行采集,但未对采集数据进行融合,易出现数据孤岛化问题。

关于毫米波雷达信息与视觉传感器信息的融合技术^[19],主要有特征级融合、数据级融合和决策级融合 3 种方法。其中,特征级融合虽然降低了计算量,但由于舍弃了部分数据,使得准确性有所下降;像素级融合的检测结果易受雷达本身的影响,若毫米波雷达出现漏检情况,最终数据也就缺失了这一目标信息;而决策级有容错性强、计算量小、兼容性好、应用范围广等优点,是最高层次的融合。

综合上述问题,本文选取毫米波雷达与视觉传感器组成主要环境感知系统,并通过基于 CFAR 算法的目标检测、基于 YOLOv5 神经网络的目标检测与识别、基于毫米波雷达和视觉传感器决策级融合等技术,建设了“雷视一体化”车辆检测与识别系统。

2 基于“雷视一体化”的融合算法

本文应用决策级融合将毫米波雷达与视觉传感器进行数据融合。数据融合中最重要的是多传感器之间的空间融合、时间融合以及数据关联。

2.1 空间融合

毫米波雷达与视觉传感器的空间融合主要通过两种传感器数据的坐标转换来实现,其中包括雷达坐标系与世界坐标系的转换、视觉传感器坐标系与世界坐标系的转换、视觉传感器坐标系与图像坐标系的转换以及图像坐标系与像素坐标系之间的转换等。坐标系之间的关系如图 1 所示。

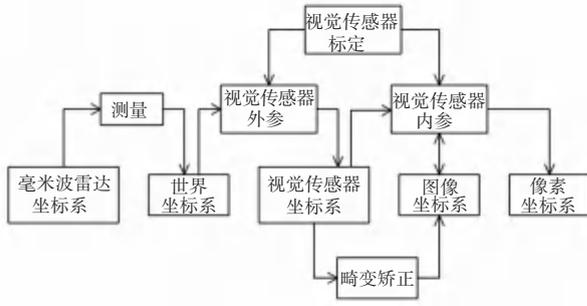


图 1 各坐标系关系图

Fig. 1 Relation diagram of each coordinate system

首先,通过公式(1)将毫米波雷达数据转换至世界坐标系中,再根据公式(2)将视频数据转换至世界坐标系中。由于图像坐标系为二维坐标系,与视觉传感器坐标单位不统一,故需依据公式(3)完成视觉传感器坐标系与图像坐标系之间的坐标转化,再根据公式(4)完成图像坐标系与像素坐标系之间的坐标转化。根据以上得到的视觉传感器坐标系与世界坐标系的转换公式,视觉传感器坐标系与图像坐标系的转换公式以及图像坐标系与像素坐标系的转换公式,可得到如公式(5)的世界坐标系与像素坐标系间的转换关系。考虑到视觉传感器图像畸变的问题,依据公式(6)对图像畸变进行校正,最终实现视觉传感器与毫米波雷达联合标定,实现两种传感器的空间融合。

$$\begin{cases} X = R \times \sin \alpha \\ Y = -H \\ Z = Z_0 + R \times \cos \alpha \end{cases} \quad (1)$$

其中, R 为径向距离, α 为偏航角。

$$\begin{pmatrix} \hat{e}^X_c \\ \hat{e}^Y_c \\ \hat{e}^Z_c \\ \hat{e}^1 \end{pmatrix} = \begin{pmatrix} \hat{e}^R \\ \hat{e}^0 \\ \hat{e}^0 \\ \hat{e}^1 \end{pmatrix} \begin{pmatrix} T \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} \hat{e}^X_w \\ \hat{e}^Y_w \\ \hat{e}^Z_w \\ \hat{e}^1 \end{pmatrix} = M_1 \begin{pmatrix} \hat{e}^X_c \\ \hat{e}^Y_c \\ \hat{e}^Z_c \\ \hat{e}^1 \end{pmatrix} \quad (2)$$

其中, $X_c Y_c Z_c - O_c$ 为视觉传感器坐标系; $X_w Y_w Z_w - O_w$ 为世界坐标系; 矩阵 R 表示两个坐标系之间的旋转关系; 矩阵 T 表示两个矩阵之间的平移关系; M_1 为视觉传感器坐标系与世界坐标系的转换矩阵。

$$\begin{pmatrix} \hat{e}^X_c \\ \hat{e}^Y_c \\ \hat{e}^Z_c \\ \hat{e}^1 \end{pmatrix} = \begin{pmatrix} \hat{e}^f & 0 & 0 & 0 \\ \hat{e}^0 & f & 0 & 0 \\ \hat{e}^0 & 0 & 1 & 0 \\ \hat{e}^0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{e}^X_u \\ \hat{e}^Y_u \\ \hat{e}^Z_u \\ \hat{e}^1 \end{pmatrix} \quad (3)$$

其中, f 为视觉传感器焦距。

$$\begin{pmatrix} \hat{e}^u \\ \hat{e}^v \\ \hat{e}^1 \end{pmatrix} = \begin{pmatrix} \hat{e}^1/dx & 0 & u_0 \\ \hat{e}^0 & 1/dy & v_0 \\ \hat{e}^0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{e}^X_u \\ \hat{e}^Y_u \\ \hat{e}^Z_u \\ \hat{e}^1 \end{pmatrix} \quad (4)$$

其中, (u_0, v_0) 为图像坐标系原点在像素坐标系中的位置,单个像素大小为 $dx \times dy$ 。

$$\begin{pmatrix} \hat{e}^u \\ \hat{e}^v \\ \hat{e}^1 \end{pmatrix} = \begin{pmatrix} \hat{e}^1/dx & 0 & u_0 \\ \hat{e}^0 & 1/dy & v_0 \\ \hat{e}^0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{e}^X_u \\ \hat{e}^Y_u \\ \hat{e}^Z_u \\ \hat{e}^1 \end{pmatrix} = \begin{pmatrix} \hat{e}^X_w \\ \hat{e}^Y_w \\ \hat{e}^Z_w \\ \hat{e}^1 \end{pmatrix} = \begin{pmatrix} \hat{e}^f & 0 & 0 & 0 \\ \hat{e}^0 & f & 0 & 0 \\ \hat{e}^0 & 0 & 1 & 0 \\ \hat{e}^0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{e}^R \\ \hat{e}^0 \\ \hat{e}^0 \\ \hat{e}^1 \end{pmatrix} \begin{pmatrix} T \\ 1 \\ 1 \\ 1 \end{pmatrix} = M_1 M_2 M_w = M M_w \quad (5)$$

$$\begin{cases} x' = x + (k_1 x r^2 + k_2 x r^4 + k_3 x r^6) \\ \quad + (p_2 (2x^2 + r^2) + 2p_1 xy) \\ y' = y + (k_1 y r^2 + k_2 y r^4 + k_3 y r^6) \\ \quad + (p_2 (2y^2 + r^2) + 2p_1 xy) \end{cases} \quad (6)$$

其中, (x', y') 为校正后的图像坐标系中的坐标, (u', v') 为像素坐标系中的坐标。

2.2 时间融合

在传感器融合系统中,不同传感器在时间维度上也要实现同步对准,以保证多传感器融合系统的有效性。毫米波雷达与视觉传感器通过间隔采样的方式,完成传感器数据在时间维度上的同步。为保证两传感器时间同步,应该选用两种传感器采样周期的最小公倍数作为系统的采样周期。

2.3 基于数据关联的车辆目标检测方法

在传感器信息融合系统中,除时空融合外,另一个至关重要的问题就是数据关联问题,即要在两个或多个数据集合之间建立关系或者联接。数据关联算

法的主要步骤为:生成数据关联假设,评估数据关联假设,确定数据关联假设。本文选用在工程中应用较为广泛的全局最近邻算法(GNN)进行数据关联。依据毫米波雷达在图像坐标系中的投影以及视觉目标检测算法检测结果,使用全局最近邻数据关联算法进行数据关联,将关联后的毫米波雷达数据与视觉目标检测结果联合显示于该帧图片中。若在一个融合周期内存在未关联的雷达目标,则以雷达数据在图像坐标系的投影点为中心,根据目标的相对距离画出目标预测框,并在目标框下显示目标距离及速度信息;若一个融合周期内存在未关联的视觉目标,则仅显示其视觉检测结果。图2为算法总体示意图。

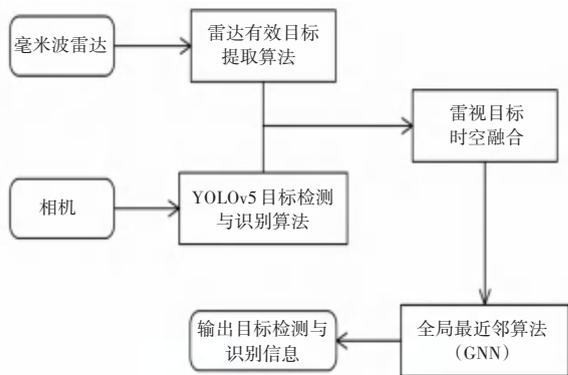


图2 算法总体示意图

Fig. 2 Overall schematic diagram of the algorithm

3 实验结果与分析

3.1 数据集

在训练视觉目标检测算法模型之前,要选取合适的数据集,由于该文所研究的模型与方法主要应用于交通场景,而KITTI数据集^[20]是目前国际上最大的自动驾驶场景下的计算机视觉算法评测数据集,故选择KITTI数据集作为目标检测算法的训练集,并依据模型要求修改数据集相关标注格式。下载YOLOv5算法的预训练权重names文件、data文件以及cfg文件,对其进行适应性修改。其中,由于该文的研究内容为交通场景下的车辆检测,故将检测类别设置为car、bus以及truck 3类。将所有设置调整完成,开始训练YOLOv5目标检测模型。

3.2 实验环境

实验所使用的是Windows 11操作系统,PyTorch深度学习开发框架,用Python作为开发语言;采用的CPU为Intel(R) Core(TM) i7-8565U CPU @ 1.80 GHz, GPU为NVIDIA GeForce MX150;选用BMR-1近距离管控雷达。实验所用设备参数见表1。

表1 设备参数

Table 1 Device parameters

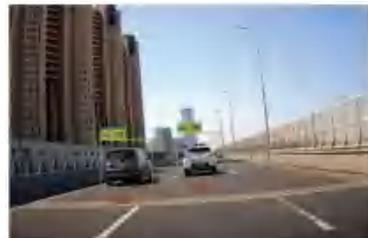
参数	数值
毫米波雷达探测距离(长距离)/m	1 500
毫米波雷达探测距离(中距离)/m	300
毫米波雷达探测方位角(长距离)/(°)	90
毫米波雷达探测方位角(中距离)/(°)	90
相机分辨率/像素 $\times 10^4$	400
相机帧率/ fps	25
相机像素尺寸/ μm	2 560 \times 1 440
相机焦距/mm	6.5~240

3.3 车辆目标检测实验验证与结果分析

图3~图5分别展示了单传感器和多传感器分别在白天、夜间和雾天的检测对比结果。



(a) 单传感器在白天的检测结果



(b) 多传感器进行数据融合后在白天的检测结果

图3 白天检测对比

Fig. 3 Comparison of daytime detection



(a) 单传感器在夜间的检测结果



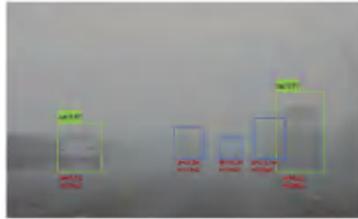
(b) 多传感器进行数据融合后在夜间的检测结果

图4 夜间检测对比

Fig. 4 Comparison of nighttime detection



(a) 单传感器在雾天的检测结果



(b) 多传感器进行数据融合后在雾天的检测结果

图5 雾天检测对比

Fig. 5 Comparison of foggy detection

通过在不同交通场景中的测试,相较于单传感器检测,“雷视融合”技术可在天气条件较好时提供更详细的前方车辆信息,在天气条件差(如大雾天气)时,提供视觉传感器无法检测到的前方车辆信息。由此体现出了多传感器融合系统的优势以及该项技术发展的必要性。

4 结束语

该文设计了一种基于“雷视一体化”的联动目标检测与识别方法,该方法充分结合了雷达和视频进行目标检测时的技术优势,采用深度学习的目标检测与识别算法,利用决策级融合对数据信息进行融合,在智能车辆避障、智能交通系统建设等领域发挥着至关重要的作用。此外,针对毫米波雷达在检测过程中由于内外部因素产生的干扰数据,该文设置了筛选条件完成了对有效数据的筛选工作,提高了模型的有效性和准确度。

由于实际交通情况比实验情况更加复杂多变,如何进一步提高模型的精确度和运算速度仍是我们努力的方向。

参考文献

[1] 陈云霁. 智能计算系统[M]. 北京:机械工业出版社, 2020.
 [2] TRISTAN V, ANDREJ S, HASCH J, et al. Classification of objects in polarimetric radar images using CNNs at 77 GHz[C]// Proceedings of 2017 IEEE Asia Pacific Microwave Conference (APMC). Piscataway, NJ:IEEE, 2017: 356-359.

[3] 张暄,高跃清. 基于卷积神经网络的雷达目标检测方法[J]. 计算机测量与控制, 2021, 29(2): 49-52,57.
 [4] 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.
 [5] 董美琳,任安虎. 基于深度学习的高速公路交通事件检测研究[J]. 国外电子测量技术,2021,40(10):9.
 [6] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric[C]//Proceedings of the IEEE International Conference on Image . Piscataway, NJ: IEEE, 2017: 3645-3649.
 [7] CHEN B, PEI X, CHEN Z. Research on target detection based on distributed track fusion for intelligent vehicles[J]. Sensors, 2020, 20(1): 56.
 [8] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2014:580-587.
 [9] GIRSHICK R. FAST R-CNN[C]//Proceedings of the IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE,2015: 1440-1448.
 [10] HE K, GKIOXARI G, DOLLÁR P, et al. Mask R-CNN[C]// Proceedings of the IEEE International Conference on Computer Vision. Piscataway, NJ:IEEE,2017: 2961-2969.
 [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector[C]// Proceedings of European Conference on Computer Vision. Cham: Springer, 2016: 21-37.
 [12] REDMON J, FARHADI A. YOLOv3: An incremental improvement [J]. arXiv preprint arXiv, 1804.02767. 2018.
 [13] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]// Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2017: 6517-6525.
 [14] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: Optimal speed and accuracy of object detection [EB/OL]. <https://arxiv.org/abs/2004.10934>.
 [15] 孙朋朋. 城市环境下智能车行车环境精确感知关键技术研究[D]. 西安:长安大学, 2019.
 [16] 魏宏安, 兰金伙, 赖黄政,等. 毫米波雷达的车辆行为监控技术研究与应用[J]. 单片机与嵌入式系统应用, 2020, 20(12): 35-39.
 [17] 高淑雅, 高跃清. 一种基于卷积神经网络的雷达目标分类方法[J]. 信息技术, 2020,94(1): 91-94, 100.
 [18] 吴建波. 延庆至崇礼高速公路雷达路况感知系统[J]. 中国交通信息化, 2021,252(1):105-107.
 [19] LU J L, TANG S M, WANG J Q, et al. A review on object detection based on deep convolutional neural networks for autonomous driving [C]// Proceedings of 2019 Chinese Control and Decision Conference (CCDC). Piscataway, NJ:IEEE, 2019: 5301-5308.
 [20] GEIGER A. Are we ready for autonomous driving? The KITTI vision benchmark suite [C]// Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ:IEEE, 2012: 3354-3361.