doi:10.12068/j.issn.1005-3026.2026.20240170

不利天气条件下目标检测的传感器融合方法

唐传茵[1]，吴龙杰[1]，於 涛[2]，付 博[1]
（1.东北大学 机械工程与自动化学院，辽宁 沈阳 110819；2.沈阳汽车集团有限公司，辽宁 沈阳 110044）

摘 要：针对相机和2D激光雷达传感器在低光照、雨、雪天气等恶劣环境下感知困难，目标检测精度低且存在漏检等问题，采用图像和2D激光雷达点云融合的目标检测方案，提出一种基于模糊理论的动态权重传感器融合方法，通过识别不同的天气和光照条件，动态分配融合权重以实现最佳的检测效果．针对雨雪天气以及低光照环境，设计并开展了雨、雪、低光照、对向来车4种场景实验．结果表明，2D激光雷达和相机融合算法有效降低了不利天气下自动驾驶车辆环境感知的漏检数量，提高了车辆在雨、雪、夜间、对向来车以及遮挡频繁环境下的感知能力．验证了本文提出的基于模糊理论的动态权重传感器融合方法的有效性．

关 键 词：环境感知；目标检测；传感器融合；模糊理论；动态权重
中图分类号：U 489 **文献标志码**：A **文章编号**：1005-3026(2026)02-0058-08

Target Detection Sensor Fusion Method Under Adverse Weather Conditions

TANG Chuan-yin[1], *WU Long-jie*[1], *YU Tao*[2], *FU Bo*[1]
(1. School of Mechanical Engineering & Automation, Northeastern University, Shenyang 110819, China; 2. Shenyang Automobile Group Co., Ltd., Shenyang 110044, China. Corresponding author: FU Bo, E-mail: 1940952561@qq.com)

Abstract: To address the challenges faced by cameras and LiDAR sensors in low-light, rainy, and snowy conditions, such as difficulties in perception, low target detection accuracy, and missed detections, a target detection scheme that fuses images and LiDAR point clouds was adopted. A dynamic weight sensor fusion method based on fuzzy theory was introduced, which dynamically allocates fusion weights by recognizing different weather and lighting conditions to achieve optimal detection performance. For rainy, snowy, and low-light environments, four experimental scenarios were designed and conducted: rain, snow, low light, and oncoming vehicles. The results show that the LiDAR and camera fusion algorithm effectively reduces the number of missed detections in unfavorable weather, improving the vehicle's perception capability under rain, snow, nighttime conditions, as well as in environments with oncoming vehicles and frequent obstacles. The results also validate the effectiveness of the proposed fuzzy theory based dynamic weight sensor fusion method.

Key words: environmental perception; target detection; sensor fusion; fuzzy theory; dynamic weight

智能驾驶汽车使用多种传感器感知环境．相机提供丰富的图像信息，但在低光照环境和识别某些反光物体时效果不佳．与相机相比，激光雷达是一种能够提供距离和三维信息的传感器．然而，激光雷达存在成本高、对天气环境敏感以及对某些表面的反射率较低等问题．2D激光雷达更少的数据量以及更低的价格，使其优势更为突出．利用多个传感器之间的互补性，可以进一步提高自动驾驶系统对于各种复杂场景的适应能力[1-2]．

Bijelic等[3]提出了一种适应性深度多模态融合网络，实现了在复杂天气条件下实时、准确的

收稿日期：2024-09-19
作者简介：唐传茵，(1979—)，女，辽宁沈阳人，东北大学副教授．
通信作者：付 博，E-mail:1940952561@qq.com.

目标检测.Mendez等[4]提出了一种面向网络边缘的目标检测模型,在保持高精度的同时降低了内存消耗.Song等[5]提出了快速反应的YOLO模型(MS-YOLO),有效提高了在光线不足条件下的检测率和远距离小目标的检测准确性.

上述工作,主要针对传感器融合方法进行研究,但并未充分利用不同传感器的优势来增强算法在不利环境下的感知能力.

为了解决不利天气下单一传感器固有的缺点,本文提出了基于模糊理论的动态权重融合方案,通过不同的天气和光照度动态分配融合权重,实现最佳的检测效果.设计4种场景的实验,提出模糊动态权重融合方法,以期降低不利天气下自动驾驶车辆环境感知的漏检数量,提高感知能力.

1　实验设备及数据

传感器采用M10二维测距激光雷达和摄像头.2D激光雷达可以直接输出目标序列,输出数据包括x,y坐标、目标距离、方位角、反射强度.实验平台主要传感器及相关部件如图1所示.使用图1所示智能小车实验平台进行数据采集,数据包括:①包含不同光照度、不同天气、不同人群密度等复杂环境;②包含但不限于行人、车辆、自行车等类别数据;③包含不同尺寸和遮挡程度的目标.自建数据集(3 567张图像),对图像算法进行微调.数据集部分图像如图2所示.

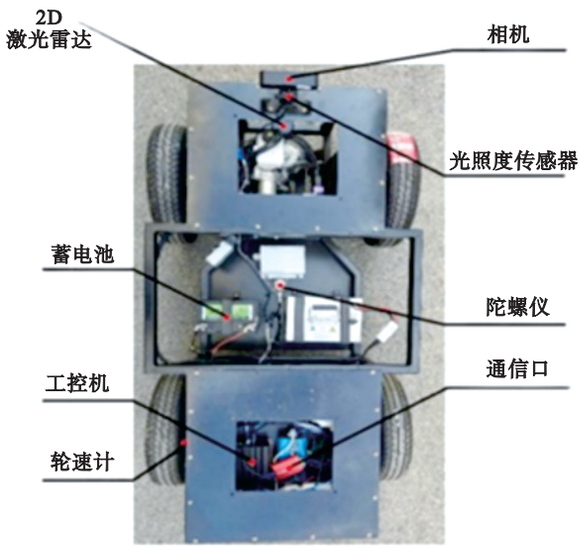


图1　智能小车实验平台

Fig. 1　Intelligent trolley experiment platform

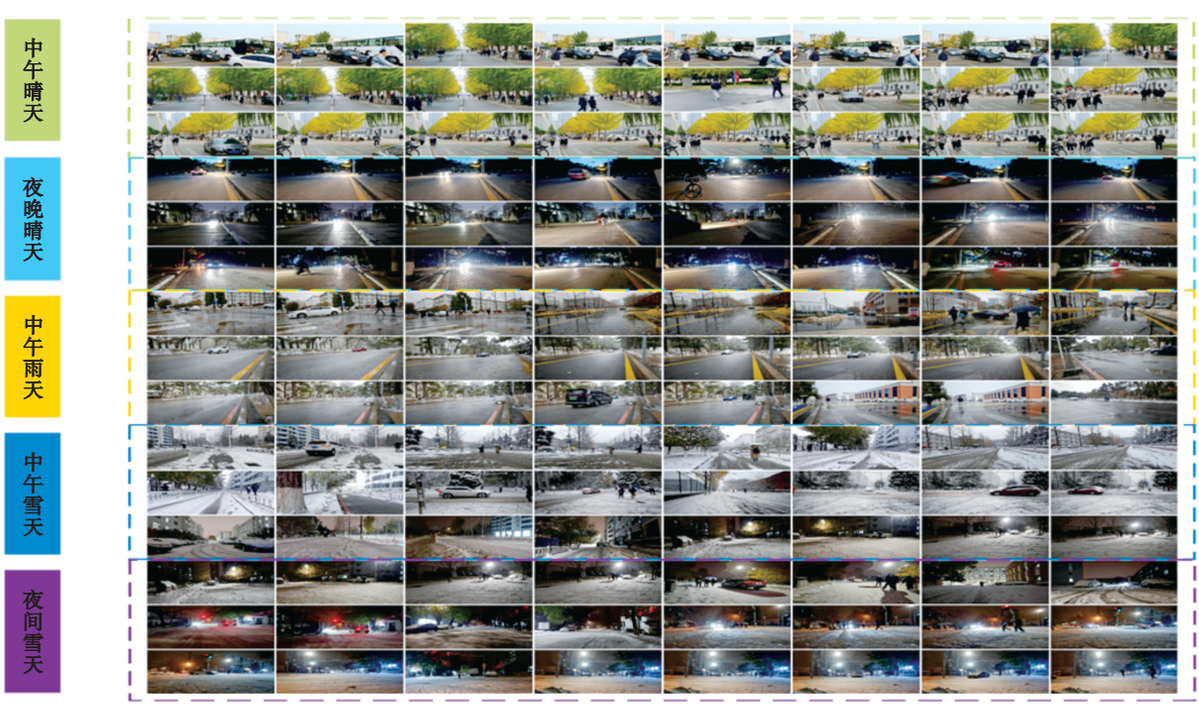


图2　自建数据集部分图像

Fig. 2　Partial images of the self-constructed dataset

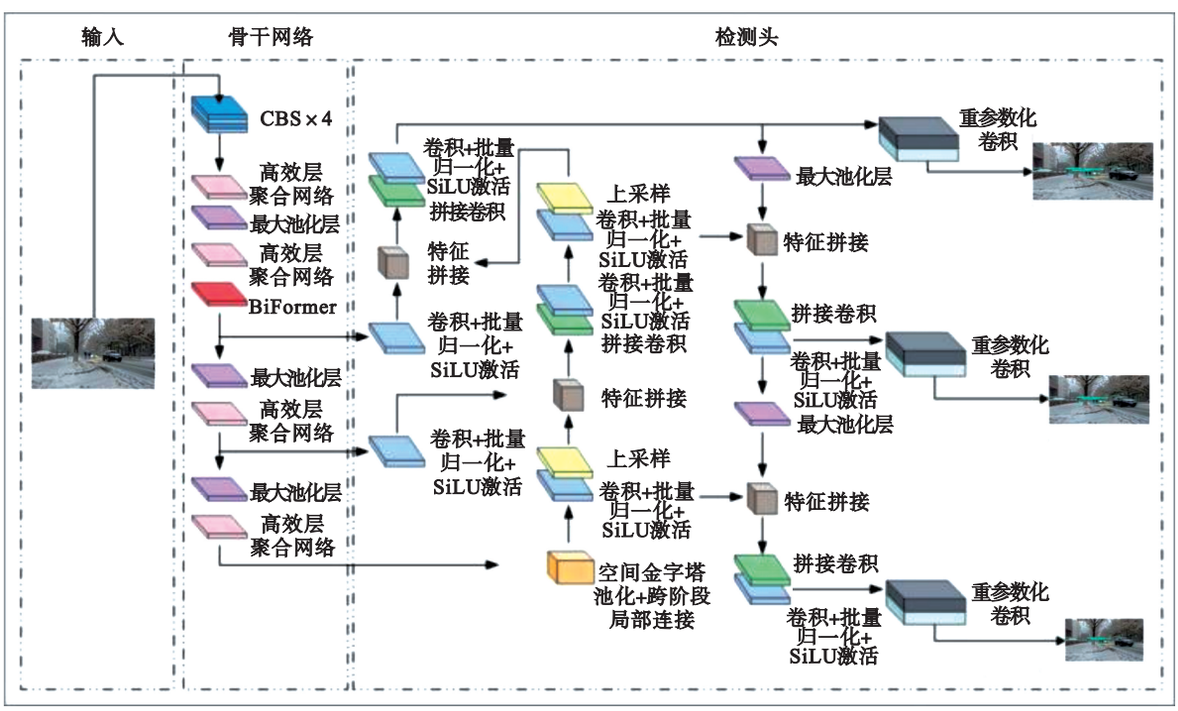
光照度是指示光照的强弱和物体被照明程度的量,单位勒克斯(lx).JIS标准[6]提供了日常各种环境下的照度值.黑夜:0.001~0.020 lx;月夜:0.020~0.200 lx;夜间路灯:5 lx;阴天室外:50~10 000 lx;高速公路隧道:150 lx以上;正常公路隧道:15~70 lx;晴天室外:30 000~100 000 lx.

1.1 图像目标检测算法

YOLOv7[7]相较于YOLOv3,其参数量减少了75%,计算量减少了36%,平均精度提高了1.5%.在小模型性能方面,YOLOv7-Tiny减少了39%的参数量和49%的计算量.YOLOv7算法在5~160帧/s时超越了当前所有的目标检测器.自注意力机制的Transformer模块相较于传统的卷积神经网络,在一定程度上提升了性能,但也存在内存占用大和计算代价高的问题.为了解决这一问题,视觉变换器的改进工作主要分为限制视野窗口、增加横向条纹、扩张窗口.这些改进的方法实质上都是通过参数限制Transformer模块的注意力范围,从而减轻其计算压力.由于Transformer直接在全局范围内进行操作,特别容易受到计算复杂性和内存占用方面的限制.

本文算法使用改进的双层路由注意力BiFormer-YOLO算法进行目标检测任务.BiFormer-YOLO算法利用卷积神经网络对图像进行特征提取.本质上,后续的分类预测结果是二分的,主要区分的是背景和目标,因此在Transformer模块中希望对感兴趣区域进行注意力关联.双层动态Transformer模块的结构首先将提取出的特征图进行粗略的感兴趣区域提取,然后将提取出的区域进行细致的自注意力特征提取,最后将提取好的特征发送至头部网络进行分类检测.在粗略提取感兴趣区域的过程中将特征图划分为$S\times S$份相对较大的区域(S为区域数量),然后进行Query(Q),Key(K)键值正确的点积阈值过滤得到感兴趣区域,但由于遮挡、背景与目标相似性等原因,导致包含目标的区域被过滤掉或者未包含目标的区域被选中.对此,本文希望借助上一时刻的融合结果对感兴趣区域进行标记.

首先在原有的YOLOv7算法中加入了本文设计的BiFormer模块.BiFormer模块首先将特征图划分为$S\times S$份补丁,再通过线性投影将补丁处理为一维的词元,然后计算一维词元的Q,K,根据Q,K的点积及上一时刻融合结果信息对词元进行过滤,并对过滤后的词元进行再细分,同时对细分后的词元嵌入位置编码,最后进行Transformer模块中多头自注意力的计算.BiFormer-YOLO算法结构如图3所示.



注:CBS×4为卷积、批量归一化和SiLU激活函数组成的复合模块重复使用4次.

图3 BiFormer-YOLO算法结构示意图

Fig. 3 Schematic structure of BiFormer-YOLO

本文算法采用自建数据集对KITTI数据集进行数据不均衡改进.首先使用KITTI数据集进行训练;然后使用自建数据集对训练参数进行微调,选取本文改进的算法以及一些目标检测算法在改进后的数据集上进行验证;最后进行实车实验.

图4为BiFormer-YOLO算法和YOLOv7算法在5种场景下的检测情况,包括中午晴天、夜晚晴天、中午雨天、中午雪天和夜间雪天.可以看

出,这些常见场景对改进前算法的检测效果都有不同程度的影响.这些信息与目标检测任务并不相关,且影响待检测目标的检测精度.

表 1 为对照组算法 YOLOv3, YOLOv5, YOLOv7,更快的区域卷积神经网络(Faster-RCNN)和单次多边框检测(SSD)的最终检测结果的比较.如表1所示,本文算法的精度和召回率在这些类型的算法中表现最好.YOLOv3 提高了小目标的检测能力,但整个框架缺乏正负样本的平衡机制;YOLOv5 虽然对正负样本平衡问题作出了修正,但特征提取能力没有 YOLOv7 强.Faster-RCNN 算法中的候选区域生成网络需要对整个输入图像进行滑动窗口操作,这导致算法的计算复杂度较高,且由于小目标的物体边界模糊,很容易受到滑动窗口尺度和分辨率限制的影响,从而导致漏检或误检.

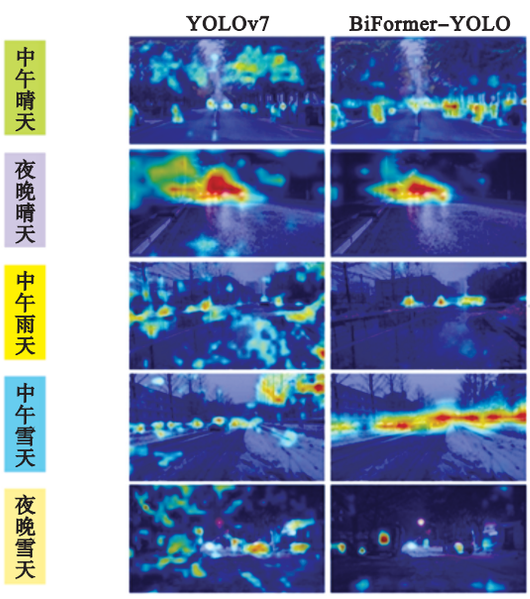


图4　YOLOv7 与 BiFormer-YOLO 区域提取对比

Fig. 4　Comparison of region extraction between YOLOv7 and BiFormer-YOLO

表 1　算法性能

Table 1　Algorithm performance

算法	精度	召回率	综合评分	帧率/(帧·s^{-1})
YOLOv3	0.813	0.821	0.815 2	30
YOLOv5	0.868	0.879	0.869 2	59
YOLOv7	0.902	0.918	0.910 8	68
Faster-RCNN	0.870	0.881	0.877 4	37
SSD	0.883	0.890	0.895 2	50
BiFormer-YOLO	**0.921**	**0.933**	**0.928 5**	**69**

1.2　基于模糊理论的动态权重融合方法

本节介绍基于模糊理论的动态权重融合方法,将2D激光雷达和相机信息融合.2D激光雷达能提供目标的位置和角度信息,而相机使用改进的BiFormer-YOLO算法检测图中的目标并输出相关数据.两个目标序列融合后,输出完整的信息.2D激光雷达与相机融合方法如图5所示.

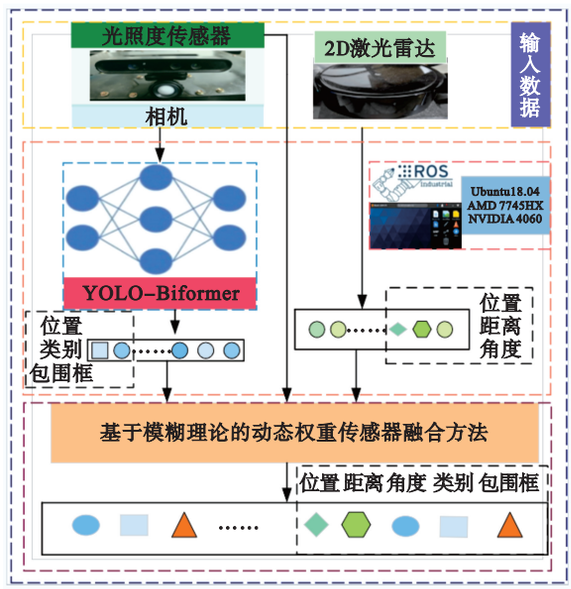


图5　2D激光雷达-相机融合方法

Fig. 5　2D LiDAR-camera fusion approach

设传感器检测目标离散时间状态向量序列为 $\boldsymbol{x}_k(k=0,1,2,3,\cdots)$,非线性函数为 f_k,控制输入为 $\boldsymbol{u}_{k-1}$,系统噪声为 $\boldsymbol{v}_k(k=0,1,2,3,\cdots)$.则递推估计式如下:

$$\boldsymbol{x}_k=f_k(\boldsymbol{x}_{k-1},\boldsymbol{u}_{k-1},\boldsymbol{v}_{k-1}). \tag{1}$$

递推估计式(1)中状态向量序列作为识别期望结果.

本文中目标匹配方法引入了马氏距离,该距离广泛用于计算两点之间的协方差距离.在使用马氏距离进行目标匹配时,需要判断图像目标序列和点云目标序列之间的哪些目标是同一目标,然后将2个目标的值匹配.在马氏距离中,相机目标与2D激光雷达目标之间的距离定义为

$$d_m=\sqrt{\left(\boldsymbol{X}_{\mathrm{C}}^{i}(x,y)-\boldsymbol{X}_{\mathrm{L}}^{n}(x,y)\right)^{\mathrm{T}}\boldsymbol{S}_k^{-1}\left(X_{\mathrm{C}}^{i}(x,y)-\mathrm{X}_{\mathrm{L}}^{n}(x,y)\right)}. \tag{2}$$

式中:$X_{\mathrm{C}}^{i}(x,y)$是相机第i个目标的位置信息;$X_{\mathrm{L}}^{n}(x,y)$是2D激光雷达目标序列中第n个目标的位置信息;当d_m取最小值的时候,相机第i个目标和2D激光雷达的第m个目标被认为是匹配的;$\boldsymbol{S}_k$是2个样本之间的协方差矩阵.

基于模糊理论的动态权重融合方法定义为

$$X^i=\alpha X_{\mathrm{C}}^{i}+\beta X_{\mathrm{L}}^{i}+\begin{bmatrix}\boldsymbol{O}\\D\end{bmatrix}^i+\begin{bmatrix}\boldsymbol{O}\\C\end{bmatrix}^i. \tag{3}$$

式中:X^i为第i个目标信息;α为相机融合权重,β为2D激光雷达融合权重,且$\alpha+\beta=1$;$\boldsymbol{O}$为空矩阵;D为第i个目标2D激光雷达检测到的目标距离;C为第i个目标相机检测到的类别信息.

对于天气隶属度函数的划分,本文基于ResNet50训练视觉天气识别模块,从加拿大恶劣驾驶条件数据集、降雨图像数据集、遥感视觉定位基准数据集中人工挑选数据对ResNet50算法进行训练,取前10帧识别结果的均值作为当前时刻的天气强度归一化值.

在模糊推理系统中,隶属度函数是整个模糊推理系统中极为重要的概念,它与特定的模糊集存在着紧密的对应关系.隶属度函数能够以量化的形式来描述输入值与模糊集之间的关联程度.这并非简单的线性关系,而是通过复杂的数学模型和计算方法,将输入值映射为1个表示其属于对应模糊集可能性的数值.

在模糊规则的制定中,充分考虑了多种影响因素.模糊规则的数量依据不同场景划分,如光照度隶属度函数划分为4部分,对应黑夜(含局部照明)、隧道、傍晚、白天等场景.同理,天气模糊规则是根据研究场景的数量及特性设计,从而充分发挥传感器的不同特性.隶属度函数如图6和图7所示.

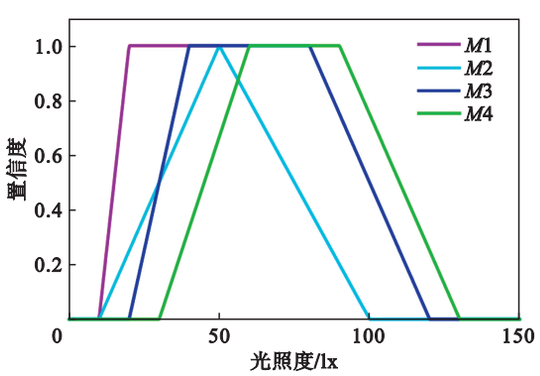


图6 光照度隶属度函数

Fig. 6 Illuminance membership function

将天气识别模块的训练类别与降雨或降雪强度值进行对应.其中,小雨、中雨、大雨的强度值分别为−0.2,−0.5,−1.0;晴天对应的强度值为0;小雪、中雪、大雪的强度值对应为0.2,0.5,1.0.根据划分的模糊集以及光照度对相机精度的影响,对融合权重进行模糊规则划分,并对相应权重进行空间曲面拟合.$M1$,$M2$,$M3$,$M4$分别为根据光照度划分的模糊集,对应的光照度范围分别0~5 lx,5~150 lx、150~10 000 lx,10 000~100 000 lx.$W1$,$W2$,$W3$,$W4$,$W5$,$W6$,$W7$分别为天气条件划分的模糊集,对应的天气条件分别为小雨、中雨、大雨、晴天、小雪、中雪、大雪.权重曲面如图8所示.垂直方向权重比为$\omega=\dfrac{\alpha}{\beta}$.

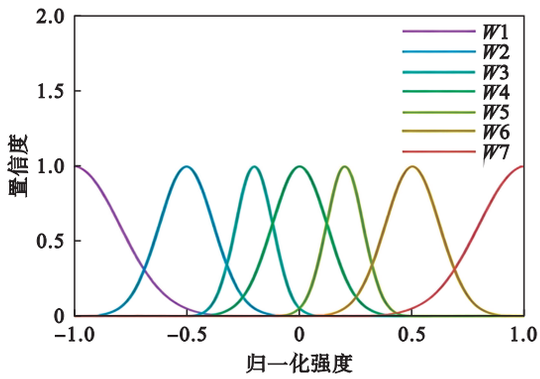


图7 天气条件隶属度函数

Fig. 7 Weather conditions membership function

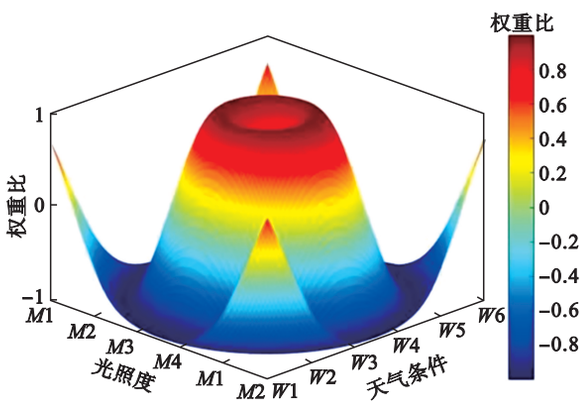


图8 模糊规则下的权重空间

Fig. 8 Weight space under fuzzy rules

2 实验评价

为了验证2D激光雷达与相机的融合结果,进行了不同场景的实车实验.实验包括4种场景,分别为夜晚晴天对向来车场景、雨天阴暗场景、雪天阴暗场景、夜间雪天十字路口场景.实验平台采用AMD7745HX处理器,16 G内存,操作系统采用Ubuntu 18.04,整个融合算法基于机器人操作系统(ROS)平台搭建.

为验证本文提出的融合方法的效果,设计了4种车辆在不利天气条件下的目标检测实验,检测结果如图9~12所示.

1)夜晚晴天对向来车环境,如图9所示,此时光照度为0~40 lx并伴随对向来车车灯的强烈灯光,算法权重及时调整,以达到更好的检测效果.融合算法的行人检测精度为0.92,黑色车辆检测精度为0.85,远处车辆检测精度为0.90.

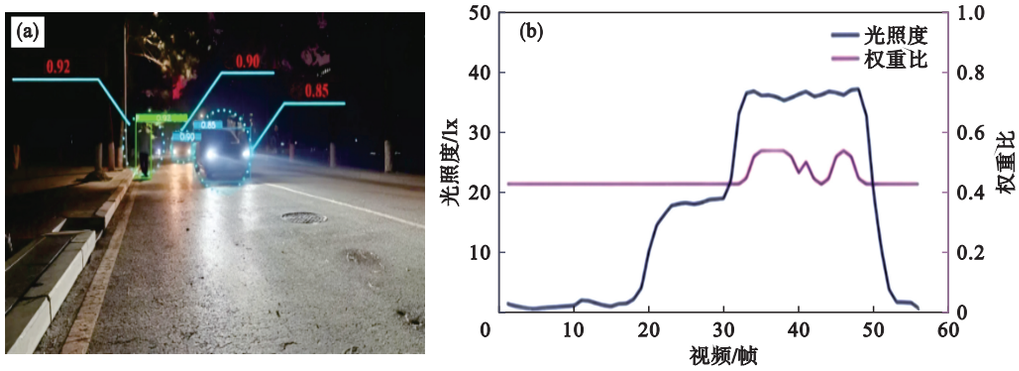


图 9　夜晚晴天对向来车场景检测结果

Fig. 9　Detection results of the oncoming traffic scene on a sunny night

(a)—融合算法检测结果;(b)—权重比与光照度.

2) 雨天阴暗场景如图 10 所示,光照度为 1 800~2 500 lx.融合算法左侧路灯和树木遮挡行人检测精度为 0.93,前方白色车辆检测精度为 0.96,远处行人检测精度分别为 0.90 和 0.85.此场景下光照度波动平缓,且由于光照度满足图像算法的检测精度,因此权重并未发生波动.

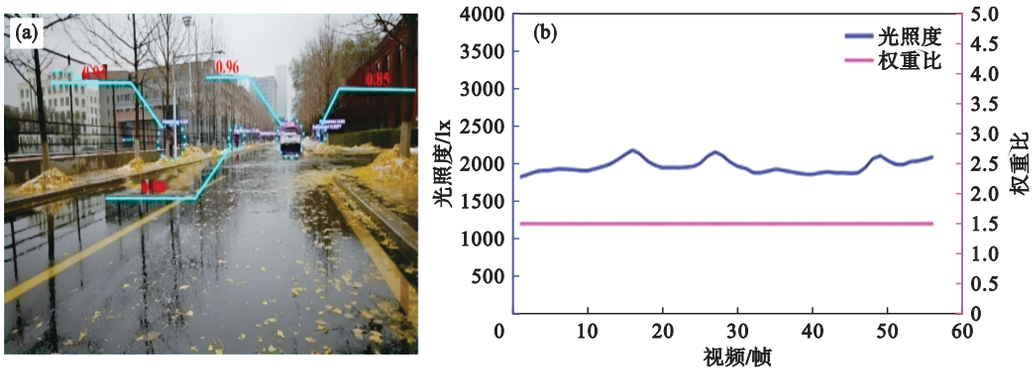


图 10　雨天阴暗场景检测结果

Fig. 10　Rainy and dark scene detection results

(a)—融合算法检测结果;(b)—权重比与光照度.

3) 雪天阴暗场景如图 11 所示,光照度在 2 000~3 000 lx 波动.融合算法对左侧被灌木丛遮挡的白色车辆检测精度为 0.92,左侧行人检测精度为 0.90,前方车辆检测精度为 0.96,右侧被遮挡行人检测精度为 0.93.雪天光照度波动较雨天波动更明显,这可能与积雪带来的漫反射有关.

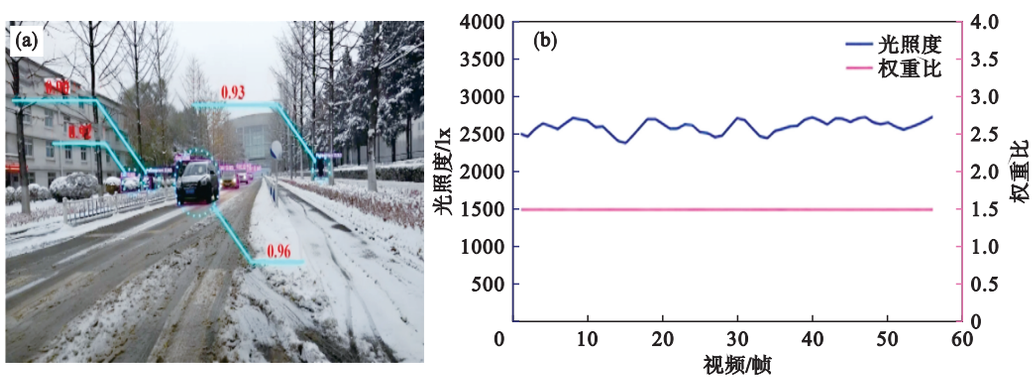


图 11　雪天阴暗场景检测结果

Fig. 11　Snowy and dark scene detection results

(a)—融合算法检测结果;(b)—权重比与光照度.

4) 夜晚雪天十字路口环境,如图 12 所示,光照度在 0~30 lx 波动.融合算法对左侧卡车检测精度为 0.96,对向远光灯车辆检测精度为 0.85,远处车灯处行人检测精度为 0.62,人行道行人检测精度为 0.93,白色车辆检测精度为 0.98,右侧被遮挡行人检测精度为 0.88.十字路口中稳定的强光照

射给图像算法检测带来了极大的难度,但本文的动态权重融合算法仍可以检测到车灯强光处的行人,且检测精度达到0.62.

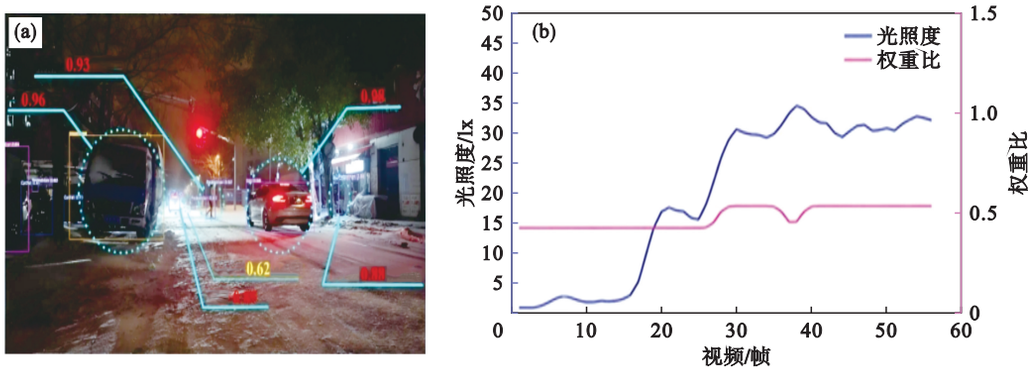


图 12 夜晚雪天十字路口场景检测结果

Fig. 12 Snowy night intersection scene detection results

(a)—融合算法检测结果;(b)—权重比与光照度.

如图13所示,在夜晚光线不充足的环境条件下,相机的检测性能会大大下降,反观2D激光雷达,在雨雪天气的影响下,漏检数量会明显增多.对于本文的融合算法而言,即使在雨雪天气的夜晚环境,也可以保证更低的漏检数量,有更好的检测效果.

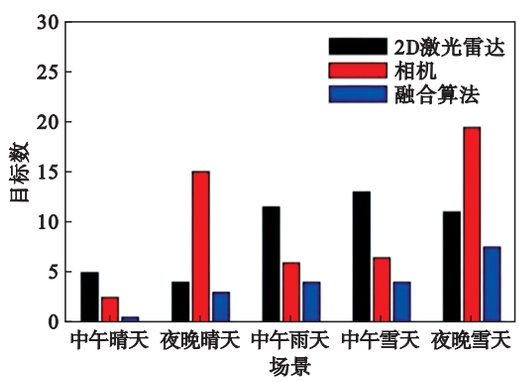


图 13 漏检数量对比

Fig. 13 Comparison of the number of missed inspections

表2为最终检测结果的比较,其中融合方法相较于BiFormer-YOLO算法精度提升0.028%,帧率相较于YOLOv7算法下降6帧/s^{-1},但仍然可以满足智能小车实时目标检测任务.

表 2 算法性能

Table 2 Algorithm performance

模型	精度	召回率	综合评分	帧率/(帧·s^{-1})
YOLOv3	0.813	0.821	0.815 2	30
YOLOv5	0.868	0.879	0.869 2	59
YOLOv7	0.902	0.918	0.910 8	68
Faster-RCNN	0.870	0.881	0.877 4	37
SSD	0.883	0.890	0.895 2	50
BiFormer-YOLO	0.921	0.933	0.928 5	69
融合方法	0.949	0.965	0.952 2	62

3 结 语

本文提出一种基于模糊动态权重的传感器融合方法,以改善车辆在不利环境下的感知能力.结合动态区域Transformer图像筛选算法改进,并利用智能小车平台采集数据制作3 567张自建图像数据集微调算法模型,提升不同天气下目标检测性能,最终BiFormer-YOLO算法精度达0.921.基于模糊理论,提出了一种模糊动态权重的融合方案,在夜晚晴天对向来车环境、阴暗雨天环境、阴暗雪天环境、夜晚雪天十字路口4种场景进行实验.结果表明,融合方法相较于单一图像算法精度提升0.028,有效降低了不利天气下自主车辆环境感知的漏检数量,提高了环境感知能力.

参考文献:

[1] Xiang C, Feng C, Xie X P, et al. Multi-sensor fusion and cooperative perception for autonomous driving: a review [J]. *IEEE Intelligent Transportation Systems Magazine*, 2023, 15(5): 36-58.

[2] Liu Z, Cai Y F, Wang H, et al. Robust target recognition and tracking of self-driving cars with radar and camera information fusion under severe weather conditions [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2022, 23(7): 6640-6653.

[3] Bijelic M, Gruber T, Mannan F, et al. Seeing through fog without seeing fog: deep multimodal sensor fusion in unseen adverse weather [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, 2020: 11679-11689.

[4] Mendez J, Molina M, Rodriguez N, et al. Camera-LiDAR multi-level sensor fusion for target detection at the network edge[J]. *Sensors*, 2021, 21(12): 3992.

[5] Song Y Y, Xie Z Y, Wang X W, et al. MS-YOLO: object detection based on YOLOv5 optimized fusion millimeter-

wave radar and machine vision[J]. *IEEE Sensors Journal*, 2022, 22(15): 15435-15447.

[6] 日本工业标准调查会. 照度基准总则：JIS Z 9110[S]. 东京：日本工业标准调查会, 2024.
(Japanese Industrial standards Commitee. General rules for illuminance standards：JIS Z 9110[S]. Tokyo：Japanese Industrial standards Committee. 2024.)

[7] Wang C Y, Bochkovskiy A, Liao H M. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [C]//2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, 2023: 7464-7475.

(上接第 39 页)

[5] Mountney P, Stoyanov D, Yang G Z. Three-dimensional tissue deformation recovery and tracking[J]. *IEEE Signal Processing Magazine*, 2010, 27(4): 14-24.

[6] Joseph J, George S N, Raja K. Parameter-free matrix decomposition for specular reflections removal in endoscopic images [J]. *IEEE Journal of Translational Engineering in Health and Medicine*, 2023, 11: 360-374.

[7] Godard C, Mac Aodha O, Firman M, et al. Digging into self-supervised monocular depth estimation[C]//2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, 2019: 3827-3837.

[8] Kumar S P, Panchami B. Optical flow glottovibrogram for the examination of vocal fold pathology [C]//2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). Sydney, 2023: 1-4.

[9] Zhu H W, Bai S Q, Shi J L, et al. IOFusion: instance segmentation and optical-flow guided 3D reconstruction in dynamic scenes[J]. *Visual Computer*, 2025, 41(2): 799-813.

[10] Shen T C, Munkberg J, Hasselgren J, et al. Flexible isosurface extraction for gradient-based mesh optimization [J]. *ACM Transactions on Graphics*, 2023, 42(4): 1-16.

[11] Zhou Y F, Zhang M C, Zhu J L, et al. A randomized block-coordinate adam online learning optimization algorithm[J]. *Neural Computing & Applications*, 2020, 32(16): 12671-12684.

[12] Turan M, Almalioglu Y, Araujo H, et al. Deep EndoVO: a recurrent convolutional neural network (RCNN) based visual odometry approach for endoscopic capsule robots[J]. *Neurocomputing*, 2018, 275: 1861-1870.

[13] Parashar S, Pizarro D, Bartoli A. Isometric non-rigid shape-from-motion with Riemannian geometry solved in linear time [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(10): 2442-2454.

[14] Ozyoruk K B, Gokceler G I, Bobrow T L, et al. EndoSLAM dataset and an unsupervised monocular visual odometry and depth estimation approach for endoscopic videos[J]. *Medical Image Analysis*, 2021, 71: 102058.

[15] Shao S W, Pei Z C, Chen W H, et al. Self-supervised monocular depth and ego-motion estimation in endoscopy: appearance flow to the rescue[J]. *Medical Image Analysis*, 2022, 77: 102338.

[16] Cao L, Zhuang S B, Tian S, et al. A global structure and adaptive weight aware ICP algorithm for image registration [J]. *Remote Sensing*, 2023, 15(12): 3185.

[17] Splietker M, Behnke S. Rendering the directional TSDF for tracking and multi-sensor registration with point-to-plane scale ICP [J]. *Robotics and Autonomous Systems*, 2023, 162: 104337.

[18] Marchisotti D, Zappa E. Virtual simulation benchmark for the evaluation of simultaneous localization and mapping and 3D reconstruction algorithm uncertainty [J]. *Measurement Science and Technology*, 2021, 32(9):095404.