

doi: 10.3969/j.issn.1002-0268.2023.02.022

基于特征加权视觉增强的雷视融合 车辆检测方法

李晓欢^{1,2}, 霍科辛¹, 颜晓凤³, 唐欣^{1,2}, 徐韶华⁴

(1. 桂林电子科技大学, 广西 桂林 541004, 广西综合交通大数据研究院, 广西 南宁 530000;

2. 广西交通科技集团有限公司, 广西 南宁 530000;

3. 广西北投信创科技投资集团有限公司, 广西 南宁 530000)

摘要: 为了有效提升高速公路低光照、远距离检测需求条件下的车辆检测准确度, 提出了一种基于视觉增强和特征加权的雷视融合车辆目标检测方法。首先从雷视数据层融合入手, 基于毫米波雷达对潜在目标空间位置进行表征, 并将表征结果用于视觉图像中远距离目标区域划分。其次对该划分区域图像进行重构、检测和还原, 以提升远距离目标的视觉检测精度。再次对雷视检测特征层融合进行建模。考虑不同层对特征检测贡献度差异, 通过模型训练获取不同特征图的权重参数, 并按照权重将不同层特征进行融合计算, 以增强目标的特征信息。其次, 增加分支网络, 利用不同尺寸的卷积层提取特征图中不同的感受野信息, 分支输出结果融合获得更强的图像表征能力, 达到提高低光照下检测精度的目标。最后, 结合特征加权雷视框架和基于毫米波雷达空间预处理的视觉增强思路, 基于YOLOv4-tiny框架设计了雷视融合检测网络, 完成验证系统搭建。结果表明: 在低光照环境下提出的算法与YOLOv4相比, 平均精度AP提高了20%, 与雷视融合典型算法RVNet相比AP值提高了5%; 在针对不同距离下的检测性能测试试验中, 本研究算法在检测120 m目标时, 平均精度值相较YOLOv4算法提高了73%, 相较于RVNet提高了63%, 提高了智能交通系统车辆检测的覆盖距离和低光照的检测精度。

关键词: 交通工程; 雷视融合; 特征加权; 视觉增强; 深度学习

中图分类号: U495

文献标识码: A

文章编号: 1002-0268 (2023) 02-0182-08

A Method for Radar-camera Fusion Vehicle Detection Based on Feature Weighted and Visual Enhancement

LI Xiao-huan^{1,2}, HUO Ke-xin¹, YAN Xiao-feng³, TANG Xin^{1,2}, XU Shao-hua⁴

(1. Guilin University of Electronic Technology, Guilin Guangxi 541004, China;

2. Guangxi Comprehensive Transportation Big Data Research Institute, Nanning Guangxi 530000, China;

3. Guangxi Transportation Science and Technology Group Co., Ltd., Nanning Guangxi 530000, China;

4. Guangxi Beitou IT Innovation Technology Investment Group Co., Ltd., Nanning Guangxi 530000, China)

Abstract: To improve the vehicle detection accuracy under the condition of low light and long distance detection requirements on expressway, a radar-camera fusion vehicle target detection method based on visual enhancement and feature weighting is proposed. First, starting with the fusion of radar-camera data layer, the spatial location of the potential target based on millimeter-wave radar is characterized, and the characterization result is used for the division of long-distance target areas in visual images. Then, the images of the divided area are reconstructed, detected, and restored to improve the visual detection accuracy of long-

收稿日期: 2022-04-01

基金项目: 国家自然科学基金项目 (61762030); 广西重点研发计划项目 (AB21196021, AB21196032); 广西杰出青年基金项目 (2019GXNSFFA245007)

作者简介: 李晓欢 (1984-), 男, 重庆人, 教授. (lxhguet@guet.edu.cn)

distance target. Next, the radar-camera detection feature layer is fused and the modeling is conducted. Considering the difference in the contribution of different layers to feature detection, the weight parameters of different feature maps are obtained through model training, and the features of different layers are fused and calculated according to the weight to enhance the feature information of the target. Next, branch network is added, and different receptive field information in the feature map is extracted by using convolutional layers of different sizes, and the branch output results are fused to obtain stronger image representation ability, which can achieve the goal of improving the detection accuracy in low light. Finally, combining with the feature-weighted radar-camera framework and the vision enhancement based on millimeter wave radar spatial preprocessing, the radar-camera fusion detection network based on YOLOv4-tiny is designed and the verification system is built. The result shows that (1) the average precision (AP) of the proposed algorithm in the low-light environment increased by 20% compared with that of YOLOv4, and the AP value increased by 5% compared with that of the radar-camera fusion algorithm RVNet; (2) in the test of detection performance at different distances, when detecting a 120-meter target, the AP value of the proposed algorithm is 73% higher than that of the YOLOv4 algorithm, and 63% higher than that of RVNet, which improved the coverage distance and low-light detection accuracy of the vehicle detection in ITS.

Key words: traffic engineering; radar and camera fusion; feature weighting; visual enhancement; deep learning

0 引言

在交通管理智能化的需求牵引下^[1-2], 道路上车辆目标检测得到了高速发展, 但是在低光照等恶劣环境下, 会造成视线不佳, 导致基于纯视觉图像的车辆目标检测性能受到诸多限制^[3]; 与此同时, 公路机电建设的实际工程部署要求设备检测距离覆盖远以减少检测设备部署的数量, 达到降本增效的效果。但当目标距离过远时, 由于其在图像中所占像素过少, 使得检测的精度大大降低。综上所述, 对于远距离以及低光照下的车辆目标检测是智能交通系统目标检测的巨大挑战。

现有智能交通系统中车辆检测常用到摄像头、毫米波雷达等。其中摄像头采集到的图像能提供物体以及环境详细的纹理信息, 但在低光照环境下检测效果不佳。而毫米波雷达虽不会受到光线影响, 抗干扰能力强、探测距离远, 但无法获取物体的纹理信息, 因此, 摄像头和毫米波互补的雷视融合方式有利于提高检测系统的精度和鲁棒性^[4]。近年来, 雷视融合方法的研究逐步增多^[5], 从融合的方式来看, 主要分为决策级融合、数据级融合和特征级融合。雷视决策级融合思路是传感器先检测得到初步的数据, 然后将各个传感器的数据送入决策终端, 通过对多组数据进行分析, 最后输出得到目标的信息。文献[6]提出一种决策级的雷视决策级融合方法。其将雷达测量的目标距离与图像中相同目标关

联, 为图像检测提供额外的距离信息, 但并未考虑基于雷达信息提升摄像头检测低光照和远距离条件下精度不高的问题。雷视数据级融合思路是利用毫米波雷达生成目标框列表, 列表中的目标框是对存在目标的假设, 再用视觉检测系统验证假设。文献[7]提出了一种基于 SSD^[8] 的雷视数据级融合车辆检测方法。首先对雷达数据进行处理, 确定合适的阈值参数以得到有效目标的信息, 通过有效目标生成目标假设, 最后通过 SSD 算法验证毫米波雷达检测到的目标是否为车辆。但该方法易受雷达漏检影响, 且雷达数据缺少目标的高度、大小等信息, 生成的检测候选区域与目标实际区域易偏差, 而影响图像检测精度。

雷视特征级融合思路是将毫米波雷达和视觉传感器数据都送入网络中, 结合两者的特征输出检测结果。决策级和数据级融合往往只能检测毫米波雷达与摄像头都可识别到的目标, 而特征级融合的方法可以从毫米波雷达与摄像头数据中提取特征, 让检测模型学习毫米波雷达与视觉数据之间的关联。即使遇到某一传感器检测效果不佳的情况, 也可以用另一传感器的数据进行补充, 是同时利用毫米波雷达和摄像头信息最有效的方法。文献[9]提出一种基于毫米波雷达数据的分割网络与视觉检测结果融合的方法。该方法提高了在低光照条件下的检测精度, 但没有考虑远距离车辆目标检测问题。文献[10]通过将毫米波雷达检测信息投影成图像, 然后

将毫米波雷达图像与视觉图像输入到 SSD 网络中进行检测,同时通过自制的远目标数据集和对 SSD 检测网络进行改进,提高对远处车辆的识别率,但该方法仅通过改进视觉检测网络来提高远距离目标识别率,没有有效利用到毫米波雷达信息。文献 [11] 提出了结合多帧毫米波雷达数据加强远处车辆特征的方法,实现对远处车辆检测,但对于中近距离的物体,多帧叠加形成的毫米波图像中目标点的位置与实际图像中目标的位置会有偏差,在训练中,不利于检测毫米波特征与图像特征的关联,而降低检测性能。

针对以上问题,研究了一种基于视觉增强特征加权的雷视融合车辆目标检测方法。与文献 [10] 不同,本研究在将图片送入网络检测前,先对图片进行处理,提出了一种基于雷达空间预处理的视觉增强方法,利用毫米波雷达对潜在目标空间位置进行表征,并将表征结果用于摄像头采集图像中远距离目标区域划分。进一步,对该划分区域图像进行重构,以提高远距离目标的视觉检测精度。同时改

进检测网络中毫米波雷达特征图与视觉特征图的融合方法,不同于文献 [10] 中的拼接融合操作,本研究提出一种基于特征权重进行融合的雷视特征图加权融合方法,提高系统在低光照条件下的检测精度。

1 雷视融合车辆目标检测方法

1.1 基于雷达空间预处理的视觉增强方法

远距离车辆在图像视野中占比较小,其可视化信息少,难以提取到具有鉴别力的特征。与此同时,随着网络的加深,下采样率变大,其细节信息可能会完全消失,导致纯视觉方法对远距离目标检测准确度低。针对以上问题,提出一种基于毫米波雷达空间预处理的视觉增强方法。该方法首先从数据层面入手,基于毫米波雷达对潜在目标空间位置进行表征,并将表征结果用于摄像头采集图像中远距离目标区域划分。进一步,对该划分区域图像进行重构、检测和还原,提升远距离目标的视觉检测精度。具体流程如图 1 所示。

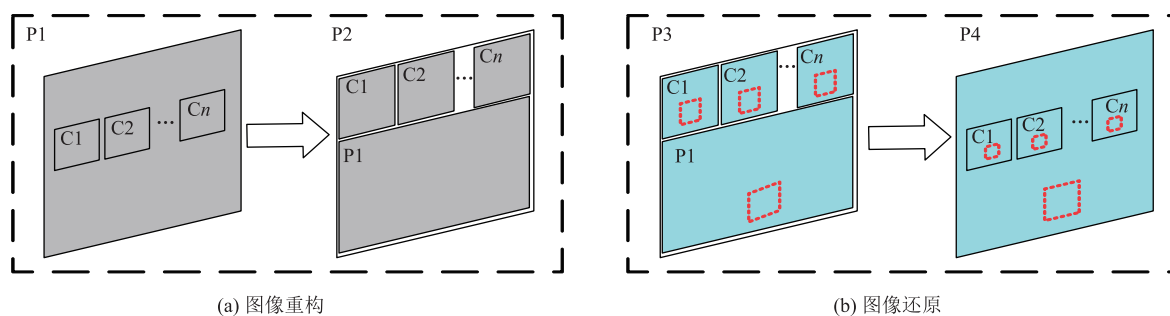


图 1 基于毫米波雷达空间预处理视觉增强方法流程

Fig. 1 Flowchart of visual enhancement method based on millimeter wave radar spatial preprocessing

基于毫米波雷达空间预处理视觉增强方法主要有图像重构与图像还原两个步骤。图 1 (a) 展示了图像重构过程,图 1 (b) 展示了图像还原的过程。其中 P1 表示输入原始图像, C1, C2, ..., Cn 表示经过毫米波雷达预检测得到的目标,映射到图像中并划分出的目标候选区域。目标区域坐标转换与映射方法如下:

$$w = \frac{(f \cdot \lambda)}{d}, \quad (1)$$

$$h = \frac{(f \cdot \vartheta)}{d}, \quad (2)$$

式中, w 为图片中目标候选框的宽度; λ 为车辆实际的宽度; h 为图片中目标候选框的高度; ϑ 为车辆实际的高度; f 为摄像头的焦距; d 为毫米波雷达检测到目标的距离。设毫米波雷达离地面的垂直

高度为 μ , 毫米波雷达高度与车辆高度之比 $\phi = \mu/\vartheta$, 可以通过式 (3) 求出目标候选区域在图像中的坐标:

$$\begin{cases} x_{\min, i}^o = x_r - w \cdot \frac{1}{2} \\ y_{\min, i}^o = y_r - h \cdot (1 - \phi) \\ x_{\max, i}^o = x_r + w \cdot \frac{1}{2} \\ y_{\max, i}^o = y_r + h \cdot \phi \end{cases}, \quad (3)$$

式中, $x_{\min, i}^o$, $x_{\max, i}^o$, $y_{\min, i}^o$, $y_{\max, i}^o$ 分别为目标区域在图像上的坐标; x_r , y_r 为毫米波雷达目标映射到图像上的位置坐标。进一步,考虑到车流量大的情况,获取的目标区域间会存在重叠,因此实际处理中可采用非极大值抑制 (Non-Maximum Suppression, NMS) 的方法,去除冗余的目标区域框。

接下来, 将放大尺寸的目标候选区域与缩小尺寸的输入 P1 进行拼接后形成新的图片 P2。然后, 将 P2 作为输入由图像检测网络进行运算, 检测后包含图像检测特征的结果为 P3。这一步骤主要是为了提高远距离目标在新图片 P2 中的占比, 使远距离目标更容易识别。尺寸变换如式 (4) 所示:

$$\begin{cases} \text{cut. size} = \frac{\text{img. size}}{n} - \nu \\ \text{origin. size} = \text{img. size} - \text{cut. size} \end{cases}, \quad (4)$$

式中, img. size 为输入图像的尺寸, 目标候选区域放大后的尺寸为粗体; cut. size 为变化后目标候选区的长宽; n 为目标候选目标框的数量。进一步, 考虑到检测网络易将两个相邻候选区域中不同物体识别为同一个, 为了防止最后输出的检测框跨越两个候选区域, 将相邻候选区域相隔一定距离, 距离设置为 ν ; 原图缩小后的尺寸为 origin. size。

为了不影响检测网络的速度, 设定拼接后的图像尺寸与原图相同。进一步, 创建与原图同尺寸的空白图, 并将目标候选框与原图进行位置变换放置在空白图上。位置变换计算如式 (5)~(6) 所示:

$$\begin{cases} x_{\min, i}^c = (\text{cut. size} + \nu) \cdot i \\ x_{\max, i}^c = x_{\min, i}^c + \text{cut. size} \\ y_{\min, i}^c = 0 \\ y_{\max, i}^c = \text{cut. size} \end{cases}, \quad (5)$$

$$\begin{cases} x_{\min}^r = 0 \\ x_{\max}^r = \text{origin. size} \\ y_{\min}^r = \text{cut. size} + \nu \\ y_{\max}^r = \text{cut. size} + \nu + \text{origin. size} \end{cases}, \quad (6)$$

式中 $x_{\min, i}^c$, $x_{\max, i}^c$, $y_{\min, i}^c$, $y_{\max, i}^c$ 为拼接图片中目标候选区域的坐标; i 为目标候选区域的序号, 其从 0 开始计数, 将目标候选区域排序后, 按序号依次从图片左上角拼接; $x_{\min, i}^r$, $x_{\max, i}^r$, $y_{\min, i}^r$, $y_{\max, i}^r$ 为原图通过尺寸缩小后在拼接图片中的位置。

将图像进行上述坐标转换与映射后输入图像检测网络进行运算, 得到包含图像目标检测框结果。由于以上目标检测框结果是基于坐标转换与映射后的图像生成, 所以需要将检测结果进行还原。图 1 (b) 展示了图像还原的过程。P3 中虚线框表示目标检测框, 图像还原是将 P3 中的检测框位置通过坐标转换重新复原到原始图像中, 并输出最终结果 P4。首先判断检测框所在的目标候选区域, 再根据目标候选区域对应原图中的坐标位置, 将检测框还原。

转换如式 (7)~(8) 所示:

$$\begin{cases} x_{\min, i}^{b-o} = \frac{x_{\min, i}^{b-c} - i \cdot (\text{cut. size} + \nu)}{\text{cut. size}} \cdot \\ \quad (x_{\max, i}^o - x_{\min, i}^o) + x_{\min, i}^o \\ x_{\max, i}^{b-o} = \frac{x_{\max, i}^{b-c} - i \cdot (\text{cut. size} + \nu)}{\text{cut. size}} \cdot \\ \quad (x_{\max, i}^o - x_{\min, i}^o) + x_{\min, i}^o \\ y_{\min, i}^{b-o} = \frac{y_{\min, i}^{b-c}}{\text{cut. size}} \cdot (y_{\max, i}^o - y_{\min, i}^o) + y_{\min, i}^o \\ y_{\max, i}^{b-o} = \frac{y_{\max, i}^{b-c}}{\text{cut. size}} \cdot (y_{\max, i}^o - y_{\min, i}^o) + y_{\min, i}^o \end{cases}, \quad (7)$$

$$\begin{cases} x_{\min, i}^{b-o} = \frac{x_{\min, i}^{b-r}}{\text{origin. size}} \cdot \text{img. size} \\ x_{\max, i}^{b-o} = \frac{x_{\max, i}^{b-r}}{\text{origin. size}} \cdot \text{img. size} \\ y_{\min, i}^{b-o} = \frac{y_{\min, i}^{b-r} - (\text{cut. size} + \nu)}{\text{origin. size}} \cdot \text{img. size} \\ y_{\max, i}^{b-o} = \frac{y_{\max, i}^{b-r} - (\text{cut. size} + \nu)}{\text{origin. size}} \cdot \text{img. size} \end{cases}, \quad (8)$$

式 (7) 为当检测框落在目标候选区域时, 转换到原始图中的坐标关系; $x_{\min, i}^{b-o}$, $x_{\max, i}^{b-o}$, $y_{\min, i}^{b-o}$, $y_{\max, i}^{b-o}$ 为从拼接图中还原到原始图中的检测框坐标; $x_{\min, i}^{b-c}$, $x_{\max, i}^{b-c}$, $y_{\min, i}^{b-c}$, $y_{\max, i}^{b-c}$ 为在目标候选区域内的检测框坐标; 确定检测框坐标与目标区域框的长宽的比例, 将其与原图中目标区域框的长宽相乘再加上 $x_{\min}^o$, $y_{\min}^o$, 即可得到检测框在原图上的坐标; 式 (8) 为当检测框落在缩小后的原图区域时, 转换到原始图中的坐标关系; $x_{\min, i}^{b-r}$, $x_{\max, i}^{b-r}$, $y_{\min, i}^{b-r}$, $y_{\max, i}^{b-r}$ 为在缩小后的原图区域内的检测框坐标。

1.2 特征加权雷视融合方法

毫米波雷达检测受光照影响小, 图像检测纹理信息多, 将两者的特征优势进行加权融合, 有利于提升系统低光照环境下的检测精度。但毫米波雷达和摄像头的结构机理、视角等差异会带来尺寸和空间位置不同。与此同时, 对目标检测而言, 不同层的特征贡献度存在差异, 如果简单将毫米波雷达特征图与视觉特征图进行空间变换, 并调整为相同尺寸后叠加, 将无法充分利用不同特征图的特性^[12]。为充分利用雷达和视觉的特征, 本研究提出一种基于特征权重^[13]进行分层融合的雷视特征图加权融合方法。该方法首先通过模型训练获取不同特征图的

权重参数, 然后按照权重将不同层特征进行融合, 其基本结构如图 2 所示。

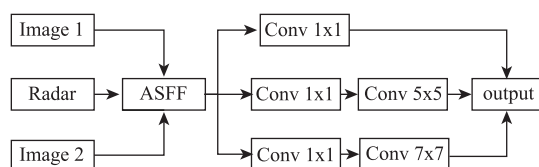


图 2 雷视加权融合框架 (A-Net)

Fig. 2 Radar-camera weighted fusion framework (A-Net)

图 2 中, Image1 和 Image2 表示低层视觉特征图和深层视觉特征图, 检测网络将两者融合以获取更多语义信息, Radar 表示通过毫米波雷达特征提取分支网络获得的毫米波雷达特征图。Conv 表示卷积层, 数字表示卷积核的尺寸, ASFF 表示特征加权模块。当特征图输入到 ASFF 时, 为便于计算, 首先统一各个特征图尺寸和通道数。利用 1×1 大小的卷积层将 Image1, Image2, Radar 调整成相同的通道数。接下来, 通过上采样或者下采样操作调节尺寸, 得到的特征图记为 resize_Image1 , resize_Image2 , resize_Radar 。再将 resize_Image1 , resize_Image2 , resize_Radar 通过 1×1 卷积后得到权重参数 α , β , γ 最后将 resize_Image1 , resize_Image2 , resize_Radar 分别与 α , β , γ 相乘即可得到加权融合的特征图。权重参数计算和加权融合计算如式 (9) 所示:

$$y = \alpha \cdot x^1 + \beta \cdot x^2 + \gamma \cdot x^3, \quad (9)$$

式中, y 为加权融合后得到的新特征图; x^1 , x^2 , x^3 为参与融合的不同特征图。其中权重参数经过 softmax 函数处理使其满足式 (10):

$$\alpha + \beta + \gamma = 1 \quad (\alpha, \beta, \gamma \in [0, 1]). \quad (10)$$

通过增加网络宽度和深度的方法进一步提高检测精度, 将加权融合后输出通过不同的分支网络, 分支上通过不同尺寸的卷积对特征图进行处理, 提取图中不同感受野的信息, 最后将 3 个分支的输出结果合并, 以获得更强的图像表征能力。如图 2 所示, 结构中包含 3 个不同的分支, 第 1 个分支 1×1 卷积组成, 第 2 个分支由 1×1 卷积和 5×5 卷积组成, 第 3 分支由 1×1 卷积和 7×7 , 分支的卷积层尺寸选取根据检测性能进行调整。进一步, 在 5×5 和 7×7 卷积层前加入 1×1 卷积层, 用于减少输入通道数, 减少网络参数的数量。3 个分支进行了不同尺度的特征提取, 增加了网络对于不同尺度的适应性。

2 雷视融合检测网络设计

基于以上视觉增强算法和特征加权融合方法的

思路, 进行雷视融合检测网络的设计。包含毫米波雷达与摄像头融合标定和网络结构设计两部分。

检测网络中视觉增强和特征加权计算都需要获取原始雷达和摄像头采集数据的空间坐标关系, 因此首先要对毫米波雷达与摄像头进行融合标定, 实现毫米波雷达坐标系的目标点与图像坐标系的映射^[14]。接下来, 对毫米波雷达数据进行处理, 生成毫米波图像。其中, 毫米波雷达检测的距离 D 、速度 V 和 RCS 分别被转换为不同通道 (R , G , B) 的像素值, 转换关系如式 (11) 所示:

$$\begin{cases} R = \frac{RCS - RCS_{\min}}{RCS_{\max} - RCS_{\min}} \cdot 255 \\ G = \frac{V - V_{\min}}{V_{\max} - V_{\min}} \cdot 255 \\ B = \frac{D - D_{\min}}{D_{\max} - D_{\min}} \cdot 255 \end{cases}, \quad (11)$$

式中, $V_{\max}$ 为速度最大值, 为当前道路的限速值; $V_{\min}$ 为速度最小值; $RCS_{\max}$ 为多次测量中毫米波雷达输出 RCS 最大值的平均值; $RCS_{\min}$ 为 RCS 最小值; $D_{\max}$ 为毫米波雷达能检测到的最大距离; $D_{\min}$ 为检测到的最小距离。

在单阶段目标检测算法中, YOLO 因检测速度快的特点被广泛应用于工业领域。其中 YOLOv4-tiny 作为 YOLOv4 的精简版本, 虽然检测精度略有下降, 但检测速度大幅提升。本研究综合考虑检测精度和速度, 认为 YOLOv4-tiny 相比其他轻量级检测网络, 更加适合部署在计算资源受限的嵌入式平台中。因此, 本研究基于 YOLOv4-tiny 基础框架设计雷视融合检测网络。网络主要包括毫米波雷达特征提取分支 (Radar _ Net)、视觉图像特征提取分支 (CSPDarknet53-tiny) 以及雷视加权融合框架 (A-Net) 3 部分, 结构如图 3 所示。

上文所述的毫米波图像与摄像头采集的视觉图像经过基于雷达空间预处理的视觉增强方法中的图像重构后, 输入到雷视融合检测网络中。Image Input 表示输入视觉图片, Radar Input 表示输入毫米波图像。毫米波特征提取分支 Radar-Net 由基础卷积单元 CBLM 和 CBL 构成, 而视觉提取分支 CSPDarknet53-tiny 主要由 CBL 单元和 Resblock 单元构成。其中 CBLM 由卷积层、批归一化处理、激活函数和最大池化层组成, CBL 由卷积层、批归一化处理、激活函数组合, Resblock 将 4 个 CBL 进行残差嵌套组合, 再经过最大池化层处理。

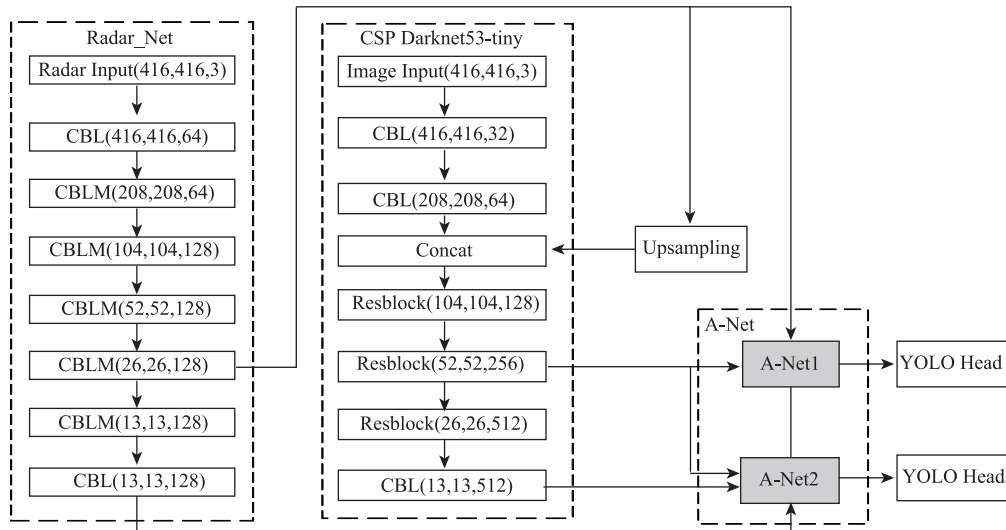


图3 融合检测网络结构

Fig. 3 Fused detection network structure

为了将毫米波雷达与视觉特征图融合, 以提高检测精度, 首先将 Radar_Net 中尺寸为 26×26 像素的特征图 Layer1 进行上采样后与 CSPDarknet53-Tiny 中尺寸为 52×52 像素的特征图 Layer2 进行 Concat 融合操作后输入到后续的视觉提取网络中。接着将 CSPDarknet53-Tiny 中尺寸为 26×26 像素的特征图 Layer3、尺寸为 13×13 像素的特征图 Layer4 和 Layer1 输入到 A-Net 框架中, 输出结果送入 YOLO Head 完成检测。

3 平台设计与试验

3.1 试验平台设计与搭建

试验平台采用 Nvidia AGX Xavier 嵌入式设备作为计算平台, 通过 Xavier 的 GPIO 扩展口与 CAN 收发器相连接, CAN 收发器再与毫米波雷达相连进行通信。摄像头和 Xavier 内置的 RJ45 型千兆以太网口, 通过网线连接完成两者的通信。为了方便采集数据, 将毫米波雷达与摄像头拼接在一起组成集成系统。

模型训练采用深度学习框架 Pytorch, 梯度优化算法为 Adam, 训练 Epoch 为 100, 初始学习率为 0.001, 输入图片的尺寸为 416×416 像素。

为了将毫米波雷达坐标系中的目标坐标转换到图像像素坐标系中, 需要进行毫米波雷达与摄像头外参标定。本研究采用单应性变换算法进行毫米波雷达与摄像头标定^[15]。

3.2 试验结果和分析

为了测试本研究算法的应用效果, 将本研究算法与纯视觉算法 YOLOv4^[16], YOLOv4-tiny, 雷视融

合典型算法 RVNet^[17] 进行测试对比。测试分别在 Nuscenes^[18] 公开数据集以及项目组自采集的黑夜低光照条件下城市道路的数据集进行。

Nuscenes 测试集上测试模型在晚上低亮度和白天正常亮度两个场景下的检测情况如表 1~2 所示。从表 2 可以看出, 在正常亮度场景测试下雷视融合网络 AP 值仅比纯视觉网络提升 5%, 但在低亮度场景下, 表 1 显示本研究算法比 YOLOv4 的 AP 值高出了 10%, 与 RVNet 相比提高了 7%, 结果表明, 本研究算法在夜晚低亮度和白天对车辆目标的检测性能都有所提升。

表 1 不同网络在 Nuscenes 夜晚数据集性能比较

Tab. 1 Comparison of performance of different networks in Nuscenes night dataset

算法类别	AP/%	检测速度/(帧·s ⁻¹)
YOLOv4-tiny	47.01	33
YOLOv4	57.86	10
RVNet	60.70	14
Ours	67.26	12

表 2 不同网络在 Nuscenes 白天数据集性能比较

Tab. 2 Comparison of performance of different networks in Nuscenes day dataset

算法类别	AP/%	检测速度/(帧·s ⁻¹)
YOLOv4-tiny	74	33
YOLOv4	78.17	10
RVNet	82.28	14
Ours	83.06	12

为进一步说明本研究算法检测车辆目标在实际场景下的检测效果, 在自制数据集上进行评估验证, 试验结果如表 3 所示。

表 3 不同网络在自制数据集性能比较

Tab. 3 Comparison of performance of different networks in self-made dataset

算法类别	AP/%	检测速度/(帧·s ⁻¹)
YOLOv4-tiny	41.29	33
YOLOv4	50.52	10
RVNet	65.29	14
Ours	70.60	12

基于 YOLOv4-tiny 和 YOLOv4 的检测算法由于低光照环境的干扰,识别率不高。相对同样是雷视融合的 RVNet,虽然本研究算法在帧数方面比 RVNet 稍低,但通过加权融合和预处理的方法,使本研究算法在黑夜环境下的鲁棒性更高,AP 提升了 5%。

表 4 展示的是各算法在检测不同距离目标时的效果。通过实地测量,标记下距离检测设备 40 m,距离设备 80 m,距离设备 120 m 的位置,让试验车辆在指定位置附近行驶,采集不同距离下的车辆目标测试集,利用检测精度高的 YOLOv4 算法和 RVNet 算法与本研究算法进行对比测试,而 YOLOv4-tiny 算法在检测远距离目标不具有优势,故不进行对比。在距离目标 40 m 时各算法都能有较好的检测效果,随着距离增加 YOLOv4 与 RVNet 算法的检测效果逐渐下降,而当距离增加到 120 m 以上,仅有本研究算法可以检测出目标,其利用了毫米波距离信息进行图像重构的方法,相比另外两种算法,在检测远距离目标车辆目标时更有优势。

表 4 不同距离下车辆检测的平均精度

Tab. 4 Average precision of vehicle detection at different distances

算法类别	距离/m		
	40	80	120
Yolov4	100.00	57.13	5.14
RVNet	100.00	60.73	15.38
Ours	100.00	79.85	78.67

表 5 是在雷视加权融合框架中使用不同的卷积层组合进行性能测试结果。雷视加权融合框架多支路部分中,大小为 3×3, 5×5, 7×7 的卷积层前面都添加一个尺寸为 1×1 的卷积层,通过这个卷积层将输入特征图的通道数减少到一定数值,以此来减少后续卷积操作的计算量。为了方便表示,表 5 中不同卷积层配置皆省略掉这个卷积层。如表 5 测试结果所示,3 个卷积层组合中 1×1, 5×5, 7×7 的效果更优,而 4 个卷积层组合的检测结果次之。因此使用 1×1, 5×5, 7×7 卷积层组成雷视加权融合框架。

表 5 不同卷积层配置下的检测性能测试

Tab. 5 Detection performance test under different convolutional layer configurations

算法类别	卷积层大小				AP/%
	1×1	3×3	5×5	7×7	
A-Net	√	√	√	—	65.10
	√	—	√	√	66.54
	—	√	√	√	63.60
	√	√	√	√	64.32

表 6 是测试毫米波雷达特征图与不同尺寸视觉特征图加权融合的检测效果,分别在 YOLOv4-Tiny 骨干特征提取网络中选取尺寸为 104, 52, 26 的特征图与毫米波雷达特征图融合,发现在视觉图像特征图尺寸为 52 的位置融合更优。

表 6 不同尺寸视觉特征图与毫米波雷达加权融合的检测性能测试

Tab. 6 Detection performance test under different sized visual characteristic drawing and millimeter wave radar weighted fusion

算法类别	图像加权融合特征图尺寸	AP/%
Yolov4-tiny+	26	65.43
	52	66.60
radar_ branch	104	62.14

为了测试雷视骨干网络融合、雷视加权融合框架,以及视觉增强对检测效果的影响,进行了消融试验,模型在相同的训练集和测试集进行训练与测试,消融试验结果如表 7 所示。可以看出 3 种改进方法都优化了检测效果,3 种方法同时使用的检测效果最优。

表 7 消融试验

Tab. 7 Fusion Elimination experiment

雷视骨干网络融合	雷视加权融合框架	视觉增强	AP/%
√	—	—	66.60
—	√	—	66.54
√	√	—	67.20
√	√	√	70.60

5 结论

本研究主要研究了面向智能交通的车辆目标检测,主要工作如下:(1)提出了基于毫米波雷达远距离目标的图片重构方法,提高了远距离目标的识别率。(2)提出了基于特征加权的雷视融合车辆目标检测方法,对毫米波雷达特征图与视觉特征图进行加权融合,提高了低光照环境下的检测精度。(3)完成了智能交通车辆检测系统搭建,并在公开数据

集以及试验环境下进行测试, 试验结果表明, 本研究算法提高了智能交通系统车辆检测的覆盖率和检测精度。

参考文献:

References:

- [1] 李斌, 候德燥, 张纪升. 论智能车路协同的概念与机理 [J]. 公路交通科技, 2020, 37 (10): 134-141.
LI Bin, HOU De-zao, ZHANG Ji-sheng, et al. Study on Conception and Mechanism of Intelligent Vehicle-infrastructure Cooperation [J]. Journal of Highway and Transportation Research and Development, 2020, 37 (10): 134-141.
- [2] 岑晏青, 宋向辉, 王东柱, 等. 智慧高速公路技术体系构建 [J]. 公路交通科技, 2020, 37 (7): 111-21.
CEN Yan-qing, SONG Xiang-hui, WANG Dong-zhu, et al. Establishment of Technology System of Smart Expressway [J]. Journal of Highway and Transportation Research and Development, 2020, 37 (7): 111-121.
- [3] MIMOUNA A, ALOUANI I, KHALIFA A B, et al. OLIMP: A Heterogeneous Multimodal Dataset for Advanced Environment Perception [J]. Electronics, 2020, 9 (4): 560.
- [4] 李原. 毫米波雷达在车路协同系统中的应用研究 [J]. 工业控制计算机, 2020, 33 (1): 44-46, 50.
LI Yuan. Research on Application of Millimeter Wave Radar in Vehicle-road Coordination System [J]. Industrial Control Computer, 2020, 33 (1): 44-46, 50.
- [5] 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] // 2019 Chinese Control and Decision Conference (CCDC). Nanchang: IEEE, 2019: 5301-5308.
- [6] JHA H, LODHI V, CHAKRAVARTY D. Object Detection and Identification Using Vision and Radar Data Fusion System for Ground-based Navigation [C] // 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN). Noida: IEEE, 2019.
- [7] 高继东, 焦鑫, 刘全周, 等. 机器视觉与毫米波雷达信息融合的车辆检测技术 [J]. 中国测试, 2021, 47 (10): 33-40.
GAO Ji-dong, JIAO Xin, LIU Quan-zhou, et al. Research on Vehicle Detection Based on Data Fusion of Machine Vision and Millimeter Wave Radar [J]. China Measurement & Test, 2021, 47 (10): 33-40.
- [8] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single Shot Multibox Detector [C] // European Conference on Computer Vision. Amsterdam: Springer, 2016: 21-37.
- [9] KOWOL K, ROTTMANN M, BRACKE S, et al. YODar Uncertainty-based Sensor Fusion for Vehicle Detection with Camera and Radar Sensors [C] // Proceedings of the 13th International Conference on Agents and Artificial Intelligence. Vienna: [s. n.], 2021.
- [10] CHADWICK S, MADDERN W, NEWMAN P. Distant Vehicle Detection Using Radar and Vision [C] // 2019 International Conference on Robotics and Automation (ICRA). Montreal: IEEE, 2019.
- [11] CHENG Y W, HU X, LIU Y M. Robust Small Object Detection on the Water Surface Through Fusion of Camera and Millimeter Wave Radar [C] // 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal: IEEE, 2021: 15263-15272.
- [12] 罗禹杰, 张剑, 陈亮, 等. 基于自适应空间特征融合的轻量化目标检测算法 [J]. 激光与光电子学进展, 2022, 59 (4): 310-320.
LUO Yu-jie, ZHANG Jian, CHEN Liang, et al. Lightweight Target Detection Algorithm Based on Adaptive Spatial Feature Fusion [J]. Laser & Optoelectronics Progress, 2022, 59 (4): 310-320.
- [13] LIU S T, HUANG D, WANG Y H. Learning Spatial Fusion for Single-shot Object Detection [R]. [S. l.]: arXiv e-prints, 2019.
- [14] 罗逍, 姚远, 张金换. 一种毫米波雷达和摄像头联合标定方法 [J]. 清华大学学报 (自然科学版), 2014, 54 (3): 289-93.
LUO Xiao, YAO Yuan, ZHANG Jin-huan. Unified Calibration Method for Millimeter-wave Radar and Camera [J]. Journal of Tsinghua University (Science and Technology Edition), 2014, 54 (3): 289-293.
- [15] WANG T, XIN J M, ZHENG N N. A Method Integrating Human Visual Attention and Consciousness of Radar and Vision Fusion for Autonomous Vehicle Navigation [C] // 2011 IEEE Fourth International Conference on Space Mission Challenges for Information Technology. Palo Alto: IEEE, 2011: 192-197.
- [16] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal Speed and Accuracy of Object Detection [R]. [S. l.]: arXiv e-prints, 2020.
- [17] JOHN V, MITA S. RVNet: Deep Sensor Fusion of Monocular Camera and Radar for Image-based Obstacle Detection in Challenging Environments [C] // 2019 Pacific-rim Symposium on Image and Video Technology, Sydney: Springer, 2019: 351-64.
- [18] CAESAR H, BANKITI V, LANG A H, et al. NuScenes: A Multimodal Dataset for Autonomous Driving [C] // 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle: IEEE, 2020: 11621-11631.