

DOI:10.19651/j.cnki.emt.2314998

基于双目视觉的拖车钩检测与定位方法研究^{*}

李冰 王豪伟 韩宇辰 胡钧涛 翟永杰

(华北电力大学自动化系 保定 071003)

摘要: 在某些危险环境下需要拖车实施救援时,救援人员难以靠近,救援人员可以通过遥控操作拖车杠来完成拖车钩的挂装。针对被救援车辆拖车钩的检测与定位问题,提出了一种拖车钩检测与定位方法 ECSA-YOLOv5,首先改进 YOLOv5 算法,设计了高效注意力模块 ECSA,将其替换掉空间金字塔池化模块上一层的模块,并增加一个大小为 160×160 的小目标检测层,能够更准确的获得拖车钩在图像中的像素坐标;通过在 SGBM 立体匹配算法预处理阶段加入引导滤波、后处理阶段引入加权最小二乘法 WLS 滤波与异常值处理,从而获得更优化的视差图,得到更为准确的目标深度信息,提高拖车钩位置信息计算的精确度。基于 Jetson Agx Xavier 开发板进行了实验验证,实验结果表明, ECSA-YOLOv5 模型较 YOLOv5s 模型 AP 值提升了 5.8%,达到了 99.0%,平均实时检测帧率为 14 fps,定位测距在 3 m 内时,误差在 3.5% 以下,能够满足拖车钩的检测与定位的准确性和实时性的要求。

关键词: YOLOv5;目标检测;定位测距;Agx Xavier

中图分类号: TP391.4 **文献标识码:** A **国家标准学科分类代码:** 510.4050

Research on tow hook detection and location method based on binocular vision

Li Bing Wang Haowei Han Yuchen Hu Juntao Zhai Yongjie

(Department of Automation, North China Electric Power University, Baoding 071003, China)

Abstract: When towing for rescue in certain hazardous environments, it is difficult for rescue personnel to approach. Rescue personnel can use remote control to operate the trailer bar to complete the installation of the trailer hook. This paper proposes a trailer hook detection and positioning method ECSA-YOLOv5 for rescue vehicles. Firstly, the YOLOv5 algorithm is improved by designing an efficient attention module ECSA, which replaces the module on the previous layer of the spatial pyramid pooling module. Additionally, a small object detection layer of 160×160 is added to obtain the pixel coordinates of the trailer hook in the image more accurately; By incorporating guided filtering in the preprocessing stage of the SGBM stereo matching algorithm and introducing weighted least squares (WLS) filtering and outlier handling in the post-processing stage, a more optimized disparity map can be obtained, resulting in more accurate target depth information and improving the accuracy of trailer hook position information calculation. Experimental verification was conducted based on the Jetson Agx Xavier development board, and the results showed that the ECSA-YOLOv5 model improved the AP value by 5.8% compared to the YOLOv5s model, reaching 99.0%. The average real-time detection frame rate was 14 fps, and when the positioning distance was within 3 meters, the error was below 3.5%, which can meet the accuracy and real-time requirements of trailer hook detection and positioning.

Keywords: YOLOv5; target detection; positioning and ranging; Agx Xavier

0 引言

拖车作为一种常见的救援方式,在日常生活和商业领

域中扮演着重要的角色。但是在一些危险环境下,例如车辆故障发生在狭窄的道路上,人员下车挂拖车钩会面临空间狭小、操作困难等问题;故障发生在陡峭的山坡上,人工

收稿日期:2023-11-15

^{*} 基金项目:国家自然科学基金面上项目(62373151)、国家自然科学基金联合项目(U21A20486)、中央高校基本科研业务费项目(2023JC006)、河北省自然科学基金(F2020502009, F2021502008)项目资助

挂拖车钩可能会面临安全风险;故障发生在化学品泄漏、有毒气体泄露等污染环境下,人工操作可能会导致中毒或其他化学伤害等,这些情况下,通过一个液压控制拖车杠装置,救援人员在车内遥控操作拖车杠进行拖车钩挂装,利用视觉识别和定位的技术,可以对视频中获取的信息进行分析和处理^[1],帮助本文更准确地检测和定位目标,从而实现更精确的拖车操作。

随着人工智能的发展和计算机硬件设备的不断升级迭代,对目标定位检测的研究,正逐渐从传统的特征提取方法逐渐发展为利用深度学习的技术^[2-3],近年来出现了许多深度学习的目标检测算法,两阶段方法包括 Faster R-CNN^[4]等,一阶段方法包括 SSD、YOLO 系列等。其中 YOLOv5 在检测速度与检测精度上都有不错的表现。

目前对相机视觉测距主要有单目测距和双目测距两种方法,单目测距成本低,实现方式简单,但是存在误差大、精度低等问题^[5],双目测距是基于人眼的感知距离,根据两个成像平面上目标图像的差异,计算目标的距离信息,精度较高^[6-7],其中立体匹配算法起着核心作用。

国内外已有很多研究人员对目标的检测与定位进行了研究,张鹏等^[8]根据几何透视变换原理,提出了海上单目测距方法,但单目测量误差较大。Zhang 等^[9]提出了一种基于 Mask R-CNN 的苹果双目定位方法,能够实现对苹果的检测与定位,但基于 Mask R-CNN 的目标检测耗时较长。梁纲等^[10]在 YOLOv5 的 Backbone 中添加 Transformer 模块有效捕获全局信息和目标图像的上下文内容信息,同时在 Neck 中使用 BiFPN 模块,能有效挖掘小目标图像内容信息及深层图像语义信息;最后使用 SIOU 损失函数,可以更准确定位目标框的位置信息。但是其参数量和推理时间大大增加,不利于部署。刘狮豪等^[11]用 GELU 激活函数替换 MobileNetV3 深层网络的 hard-swish 激活函数,将改进的 MobileNetV3 网络替换 YOLOv5 主干网络,降低了计算量,提高了模型检测速度,增加 BiFPN 特征金字塔结构与不同特征层融合,同时添加 CBAM 注意力机制关注重要特征抑制不显著特征,虽然模型参数和大小降低了,但是其检测准确率下降了。Zheng 等^[12]将 SE 模块添加到了 YOLOv5 的主干网络中,并将 ACON 激活函数取代了原有的 ReLU 激活函数,以此来提高织物的缺陷检测能力,但提高织物缺陷检测准确率的方法并不适用于对拖车钩的检测。毛先胤等^[13]在 SGBM 立体匹配算法引入 Census 变换,加快了立体匹配速度,但是其得到的视差图仍不够平滑,匹配效果较差。乔超等^[14]将 Roberts 算法与 SGBM 立体匹配算法相结合,更准确地匹配了边缘特征,但 Roberts 算法对噪声比较敏感,可能会导致立体匹配结果受到干扰,同时过度依赖边缘信息可能导致对非边缘区域的匹配精度下降。

上述算法虽然取得了一定的效果,但拖车钩目标不仅需要提升目标检测和立体匹配的准确率,还需要提升运行

速度来保证部署后的实时性,因此本文提出的基于双目视觉的拖车钩检测与定位方法,首先设计了一种基于 YOLOv5 的改进算法 ECSA-YOLOv5,将高效注意力模块 (efficient channel spatial attention module, ECSA) 替换空间金字塔池化模块 (spatial pyramid pooling fast, SPPF) 上一层的 C3 模块,并增加一个尺寸大小为 160×160 的小目标检测层,实现对拖车钩的识别和初步定位;对 SGBM 立体匹配算法进行改进,在预处理阶段加入引导滤波、后处理阶段引入加权最小二乘法 (weighted least squares estimation, WLS) 滤波与异常值处理,得到目标在两幅图像上的视差,根据相似三角形原理计算出目标距离。

实验结果表明所提目标检测算法可以有效提高目标检测的精度,所提改进 SGBM 算法能够降低图像噪声,在对视差图进行平滑处理的同时保留边缘信息,提高了立体匹配的质量和准确性。

1 改进 YOLOv5s 目标检测算法

1.1 YOLOv5s 网络结构

YOLOv5 作为一种基于深度学习的目标检测算法,兼顾了检测速度快与精度高的优点,能够做到实时检测目标。为了保证算法的运行流畅性,选择了网络深度和宽度较小、计算资源占用少且运行速度较快的 YOLOv5s 进行研究。

如图 1 所示, YOLOv5s 网络结构由骨干网络 (Backbone)、颈部网络 (Neck) 和头部网络 (Head) 3 个部分组成。其中 Backbone 负责对输入图像进行特征提取, Neck 负责对特征图进行多尺度特征融合,并把这些特征传递给预测层, Head 进行最终的回归预测。

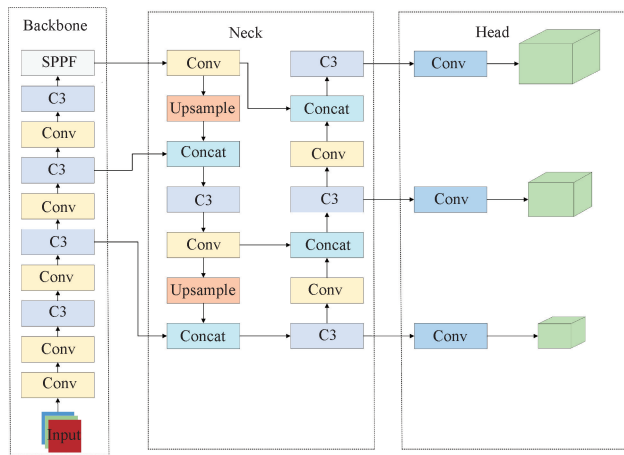


图 1 YOLOv5s 网络结构

1.2 ECSA-YOLOv5 算法

根据在拖车场景中,在相机距离拖车钩目标较远时,拖车钩在相机中成像较小,本文设计了基于 YOLOv5s 的 ECSA-YOLOv5 模型,提出一种高效注意力模块 ECSA 替换 SPPF 上一层的 C3 模块,并在头部网络 Head 添加一个尺寸大小为 160×160 的小目标检测层来提高对目标的检

测精度。

目标检测算法的部署,往往需要在提升检测精度的同时,还需要兼顾算法的高效性,利用添加基于卷积块的注意力模块(convolutional block attention module, CBAM)来改进算法,虽然增强了网络对重要特征的感知能力,但由于 CBAM 模块相对复杂、计算量大,会导致检测的时间增加;利用添加高效通道注意力模块(efficient channel attention, ECA)来改进算法,虽然增强了特征表达能力和减少了参数量,但其主要关注通道间的依赖关系,对于空间信息的利用相对较弱^[15]。

针对上述两个模块无法兼顾检测精度与检测速度的问题,所提 ECSA 注意力模块将两者进行了融合,即保留了 CBAM 的空间注意力模块,将 CBAM 的通道注意力模块替换为 ECA 注意力模块。ECA 模块相比于 CBAM 的通道注意力机制,参数量与计算复杂度更低,保留了 CBAM 的空间注意力机制,使得模型仍然可以利用空间注意力机制来捕捉输入特征图中的空间信息,利用 ECA 通道注意力模块的高效性优势,进一步提升网络的性能。

ECSA 注意力模块如图 2 所示。

ECSA 计算过程如下:

$$F' = M_c(F) \otimes F \quad (1)$$

$$F'' = M_s(F') \otimes F' \quad (2)$$

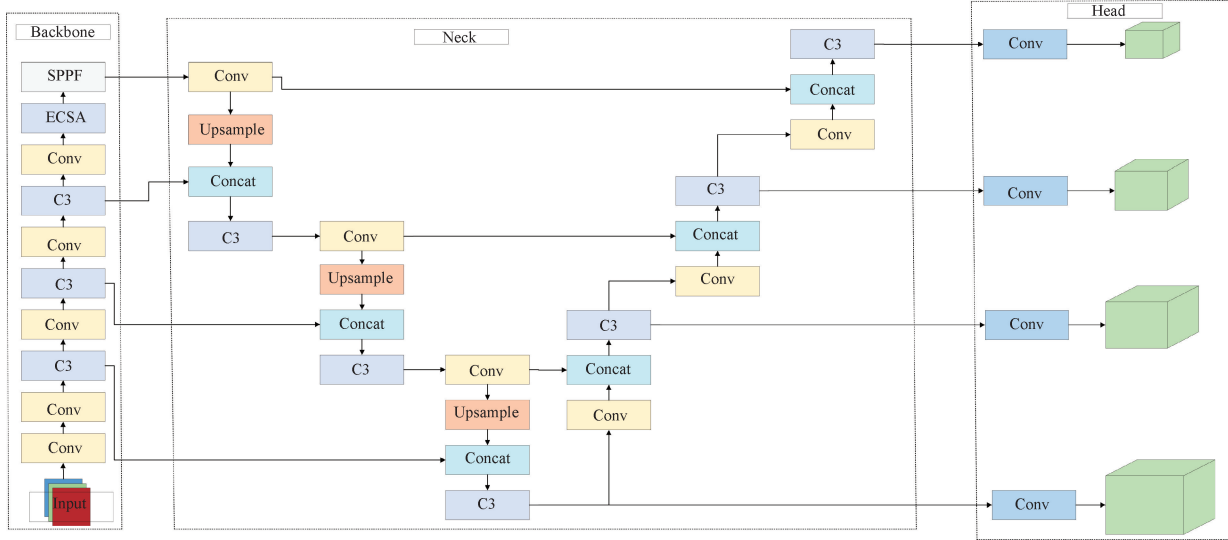


图 3 ECSA-YOLOv5 网络结构

2 目标精确定位

双目相机利用两个摄像头之间的视角差异,通过匹配算法获得对应的图像点,计算目标空间点在两幅图像中的视差,从而得到目标的距离^[16-17]。

2.1 立体匹配常用算法

立体匹配是计算机视觉中用于从一对图像中找到对应点以获得深度信息的技术^[18],其目的是为了得到视差

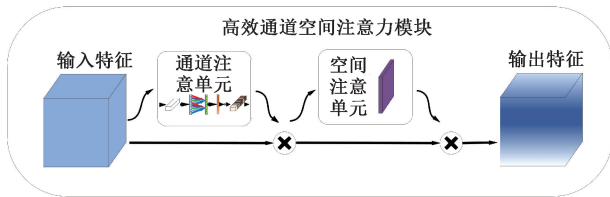


图 2 ECSA 模块

式中: F 为输入特征, M_c 为 ECA 注意力模块, M_s 为 CBAM 的空间注意力模块。

由于拖车钩目标在距离相机较远的情况下,目标在相机中成像较小,在经过多次下采样后,特征图中的目标信息不断变少,而 YOLOv5 输出 3 个检测层,通过检测层生成初始预测框,3 个层的下采样倍率分别是 8、16、32,如输入图像为 640×640 ,那么输出的检测层分别为 80×80 、 40×40 、 20×20 ,分别用于检测小、中、大目标,当目标更小时,原有的 3 个检测层无法满足检测精度,因此在原有 3 个目标检测层上新增加了一个尺寸大小为 160×160 的小目标检测层。该检测层可以更好的获取到拖车钩小目标的细节特征和位置信息,提升了对远距离拖车钩目标的识别准确率。添加小目标检测层和 ECSA 注意力模块后的改进的 ECSA-YOLOv5 网络结构如图 3 所示。

图。双目视觉测量精度与左、右视图立体匹配效果相关^[19],目前常用且效果较好的立体匹配算法有块匹配方法(block matching, BM)和半全局块匹配方法^[20](semi-global block matching, SGBM)。

BM 立体匹配算法是一种全局、基于窗口的匹配算法,采用设定的窗口由左至右提取特征向量遍历搜索,遍历过程中比较两张同步图像上窗口之间的相似性^[21]。

SGBM 算法引入了全局优化的思想,首先为每个像素

点选择视差值以形成初始的视差图,然后设定一个与视差图相关的全局能量函数,通过求解该能量函数的最小值来得到各像素的最佳视差。

SGBM 立体匹配算法的能量函数为:

$$E(D) = \sum_p \left\{ \begin{aligned} & C(p, DP) + \\ & \sum_{q \in N_p} \sum_{q \in N_p} p1I[|Dp - Dq| = 1] + \\ & \sum_{q \in N_p} \sum_{q \in N_p} p2I[|Dp - Dq| > 1] \end{aligned} \right\} \quad (3)$$

其中, $E(D)$ 指视差图 D 对应的全局能量函数, p, q 为图像中两个像素点, $C(p, DP)$ 指 p 的视差值是 DP 时的像素点代价 p_1, p_2 代表惩罚系数。

2.2 基于 SGBM 改进的立体匹配算法

SGBM 立体匹配算法得到的视差图存在很多噪声和不连续的区域,缺乏对目标边缘信息的处理^[22],得到的视差图质量较差,会导致深度估计的结果不准确。

当前已有很多针对 SGBM 的改进方法,例如在 SGBM 立体匹配算法引入 Census 变换^[13],但由于 Census 变换是一种基于局部邻域的像素亮度比较的方法,它将一个像素与其邻域中的每一个像素进行比较,然后根据比较结果生成一个二进制字符串,这个过程对噪声非常敏感,因为噪声可以改变像素的亮度,从而影响比较结果;将 Roberts 算法与 SGBM 立体匹配算法相结合的方法^[14],Roberts 算子是一种用于边缘检测的简单算子,它通过计算图像的局部梯度来检测边缘,可能会因为噪声而产生错误的边缘检测结果,为了得到效果更好的视差图,消除噪声带来的影响,提高立体匹配的质量和准确性,本文提出基于 SGBM 改进的立体匹配算法,主要包括:

1) 在 SGBM 算法预处理阶段加入引导滤波^[23],降低图像中的噪声,并增强图像的细节。

2) 在 SGBM 算法后处理阶段引入加权最小二乘法(WLS),对视差图进行平滑处理,同时保留边缘信息,进一步提高立体匹配的质量。

3) 对视差图进行异常值处理。

在将双目图像转换成灰度图,用于消除畸变和进行双目的立体校正后,双目图像会存在一些噪声,丢失一些细节信息等问题,为了降低图像中的噪声,并增强图像的细节信息,本文在 SGBM 算法对图像的预处理阶段加入引导滤波。

引导滤波是一种线性时间复杂度的边缘保持滤波器,它可以在大窗口上进行高效的滤波,从而在保持边缘的同时平滑大面积的同质区域。其基本思想是在每个像素处,通过考虑其邻域内的像素值,来预测输出像素的值。

具体来说,在一个以 i, k 为像素索引,半径为 r 的中心邻域 ω_k 中,对于输入图像和引导图像,引导滤波的输出 q 满足以下线性模型:

$$q_i = a_k I_i + b_k \quad \forall i \in \omega_k \quad (4)$$

其中, I_i 是输入图像的点, ω_k 是以 I_i 在引导图像中的

对应点的中心邻域, q_i 是 I_i 在输出图像中对应的点, a_k 和 b_k 是对应的系数,是在窗口 ω_k 中通过最小化以下优化问题得到的:

$$E(a_k, b_k) = \sum_{i \in \omega_k} [a_k I_i + b_k - p_i]^2 + \epsilon a_k^2 \quad (5)$$

其中, ϵ 是一个正则化参数,用于控制 ϵa_k^2 的大小, a_k 和 b_k 的计算公式为:

$$a_k = \frac{1}{|\omega|} \sum_{i \in \omega_k} I_i P_i - \mu_k P_k \quad (6)$$

$$b_k = \bar{P}_k - a_k \mu_k \quad (7)$$

其中, μ_k 和 σ_k^2 为引导图像在以像素 k 为中心,半径为 r 的均值和方差, $\bar{P}_k$ 为输入图像的均值。

在 SGBM 算法进行代价计算、动态规划后的后处理阶段引入加权最小二乘法(WLS)边缘保持滤波,得到高质量的视差图。加权最小二乘法(WLS)滤波使输出图像 u 与输入图像 p 经过平滑后尽量相似,在边缘部分尽量保持原状。该过程用数学公式表示为寻找式(8)的最小值。

$$\sum_p (u_p - g_p)^2 + \lambda (a_{x,p}(g) (\frac{\partial u}{\partial x})_p^2 + a_{y,p}(g) (\frac{\partial u}{\partial y})_p^2) \quad (8)$$

其中, g 表示输入图像, u 表示输出图像, p 表示像素位置。第 1 项 $(u_p - g_p)$ 代表输入图像和输出图像的差值。插值越小,边缘结构越相似,第 2 项参数 λ 用于控制图像的平滑程度,数值越大,平滑效果越好。 $a_{x,p}(g)$ 和 $a_{y,p}(g)$ 是两个平滑权重系数。具体表示为:

$$a_{x,p}(g) = (|\frac{\partial l}{\partial x}(p)|^a + \epsilon)^{-1} \quad (9)$$

$$a_{y,p}(g) = (|\frac{\partial l}{\partial y}(p)|^a + \epsilon)^{-1} \quad (10)$$

其中, l 表示对输入图像取对数, α 表示梯度的敏感程度, ϵ 是为了防止分母为 0,一般取一个极小值。

通过加权最小二乘法(WLS)滤波,改善了视差图的质量,保留了边缘信息,很大程度上消除了噪声、纹理缺失等因素的影响。

由于噪声、光照变化等因素,可能会导致视差图中出现一些异常值,这些异常值可能会对视差计算产生不良的影响,因此,采用对视差图进行异常值处理可以有效地提高视差计算的准确性和稳定性。

视差图的均值和标准计算如下:

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (11)$$

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2} \quad (12)$$

其中, μ 表示视差图的均值, σ 表示视差图的标准差, x_i 表示视差图中的第 i 个像素值, N 表示视差图中像素点的总数量,也就是视差图的面积。

过低的视差值视为异常值,将视差图中小于均值减去两倍标准差的像素值,替换为均值。

$$(x_i < \mu - 2\sigma) = \mu \tag{13}$$

过高的视差值视为异常值,将视差图中大于均值加上两倍标准差的像素值,替换为均值。

$$(x_i > \mu + 2\sigma) = \mu \tag{14}$$

基于 SGBM 改进的立体匹配算法减少了由消除畸变和双目立体校正后的噪声,对立体匹配生成的视差图进行了平滑处理,保留了边缘信息,同时对异常值进行了处理,进一步提高了立体匹配的质量和准确性。

2.3 目标测距与相对角度计算

由立体匹配得到视差值后,根据相似三角形原理可以求出拖车钩目标到相机的距离。为了满足后续遥控操作拖车杠伸向车辆去挂装拖车钩的要求,还需要计算出两个拖车钩目标以相机中心点为圆心的相对角度。具体计算如下:

设经过深度信息得到的左拖车钩在坐标系下的向量为 $\mathbf{V}_1 = (X_1, Y_1, Z_1)$, 右拖车钩在坐标系下的向量为 $\mathbf{V}_2 = (X_2, Y_2, Z_2)$ 。

设 d 为两个向量的点积,则:

$$\text{dot} = \mathbf{V}_1 \cdot \mathbf{V}_2 = X_1 \cdot X_2 + Y_1 \cdot Y_2 + Z_1 \cdot Z_2 \tag{15}$$

设 L_1 为向量 $\mathbf{V}_1$ 的向量长度, L_2 为向量 $\mathbf{V}_2$ 的向量长度,则:

$$L_1 = \|\mathbf{V}_1\| = \sqrt{X_1^2 + Y_1^2 + Z_1^2} \tag{16}$$

$$L_2 = \|\mathbf{V}_2\| = \sqrt{X_2^2 + Y_2^2 + Z_2^2} \tag{17}$$

用 $\cos\theta$ 表示两个向量夹角的余弦值,则:

$$\cos\theta = \frac{(\mathbf{V}_1 \cdot \mathbf{V}_2)}{(L_1 \cdot L_2)} = \frac{(\mathbf{V}_1 \cdot \mathbf{V}_2)}{\|\mathbf{V}_1\| \cdot \|\mathbf{V}_2\|} = \frac{X_1 \cdot X_2 + Y_1 \cdot Y_2 + Z_1 \cdot Z_2}{\sqrt{X_1^2 + Y_1^2 + Z_1^2} \cdot \sqrt{X_2^2 + Y_2^2 + Z_2^2}} \tag{18}$$

其中, θ 为两个拖车钩的相对角度。

3 实验与分析

3.1 开发平台

本实验采用 Jetson Agx Xavier 硬件平台,搭载 Ubuntu18.04 操作系统,双目相机选择汇博捷视的 HBV-1714-2 S2.0 双目相机,实验所用 Jetson Agx Xavier 硬件平台和双目相机的主要参数如表 1 和 2。

3.2 目标检测实验

1) 数据集获取与模型检测结果

本实验所用对象为红色与黑色相间的拖车钩,长度为 110 mm,宽度为 80 mm,厚度为 21 mm,可承受 4.7 吨的拉力,该拖车钩具有形状稳定、单一的特点。训练的原始图像为双目相机拍摄的彩色图像和网络搜集真实车辆上的拖车钩图像,共计 1 000 张。

实验训练环境:本文采取的实验所使用的服务器操作系统为 Ubuntu18.04, GPU 为 NVIDIA RTX 3090ti 显卡,

表 1 Jetson Agx Xavier 主要参数

参数	Jetson Agx Xavier
AI 性能	32TOPS
DL 加速器	2x NVDLA
视觉加速器	2x PVA
GPU	512 CUDA Core 64 Tensor Core
CPU	8 ARMv8.2 CPU Core
显存	32 G
存储	32 G
尺寸	105 mm×105 mm×65 mm

表 2 HBV-1714-2 S2.0 双目相机主要参数

参数	HBV-1714-2 S2.0 双目相机
像素	130×10 ⁴
视场角	80°无畸变
焦距	3.0 mm
分辨率	2 560×960
尺寸	156 mm×36 mm×25 mm

深度学习框架为 Pytorch1.13.1, Torchvision0.14.1, CUDA 11.6, Python3.10。训练时参数设置:优化器为随机梯度下降,初始学习率为 0.01,最终学习率为 0.000 1,共训练 200 个 epochs,按照 4:1 的比例完成训练集和验证集的划分。

通过模型训练后,开始对拖车钩目标进行实时与离线检测,为了方便在不损坏车辆,更好的去控制实验环境和参数情况下去检测拖车钩,本实验实时检测采用拖车钩放在灰色、白色背景下的实验场景,去模拟真实场景的情况;离线检测的对象为网络搜集真实车辆上的拖车钩图像。拖车钩目标检测部分检测结果如图 4 所示。



图 4 拖车钩目标检测部分检测结果

2) 对比试验

本实验针对同一数据集,分别用原 YOLOv5s 网络, YOLOv7^[24] 网络,分别加入 ECA、CBAM 注意力模块的 YOLOv5s 网络,本文加入 ECSA 注意力模块和小目标检测层的网络上进行实验,采用平均精确度 (average precision, AP) 和每秒检测帧率 (frames per second, FPS) 作为评价指标,得到的实验数据如表 3。

表 3 对比实验结果

模型	Image Size	AP/%	FPS/fps
YOLOv5s	720	93.2	65
YOLOv7	720	97.1	25
ECA-YOLOv5s	720	96.8	63
CBAM-YOLOv5s	720	98.6	52
ECSA-YOLOv5s	720	99.0	58

由实验结果可知,由于在训练集和验证集中包含有很多小目标,添加 ECSA 注意力模块和在头部网络添加一个尺寸大小为 160×160 的小目标检测层后的 ECSA-YOLOv5 的检测精度相较于其他算法有更高的准确率,高达 99.0%,在 Jetson Agx Xavier 硬件平台的检测帧率为 58 fps。

3) 消融实验

为了证明本文提出的 ECSA 注意力模块和小目标检测层对模型检测能力的提升,将这两个不同的改进方式对模型检测的效果做消融实验进行评估。得到的实验数据见表 4。

表 4 消融实验结果

模型	Image Size	AP/%	FPS/fps
YOLOv5s	720	93.2	65
+ECSA	720	97.6	55
+小目标检测层	720	95.8	61
+ECSA+小目标检测层	720	99.0	58

由实验可知,ECSA 注意力模块和添加小目标检测层都对模型检测目标的准确率得到了提升,加入了 ECSA 注意力模块后,AP 提升了 4.4%,说明 ECSA 帮助模型更好的关注了重要特征。添加了小目标检测层后,AP 提升了 3.6%,说明在相机距离目标较远时的准确率得到了提升。最后将 ECSA 注意力模块和小目标检测层都作用于 YOLOv5s,虽然每秒检测帧率略有下降,但是 AP 提升了 5.8%,达到了 99.0%,极大地提升了检测准确率。

3.3 立体匹配实验

1) 对比试验

为了更加直观地观察本文所提算法的有效性,对拖车钩模拟场景图进行了实验,并与 BM 立体匹配算法、目前其他对 SGBM 的改进算法进行比较,拖车钩模拟场景如图 5 所示,对比实验结果图如图 6 所示,运算时间如表 5。



图 5 拖车钩模拟场景图

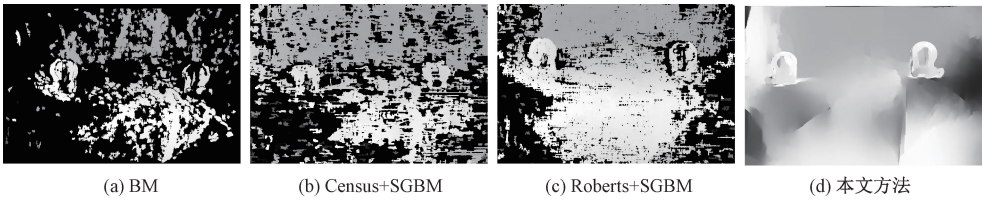


图 6 各立体匹配算法对比实验视差图

表 5 对比实验各立体匹配方法的运算时间

立体匹配算法	运算时间/ms
BM	41.5
Census+SGBM	58.5
Roberts+SGBM	75.3
本文方法	81.6

实验结果表明,BM 算法的处理速度最快,但在处理精度上,得到的视差图上表面粗糙,轮廓模糊,深度信息欠缺;采用在 SGBM 立体匹配算法引入 Census 变换,匹配速度较快,但是匹配结果无法消除噪声带来的影响,失真比较严重;将 Roberts 算法与 SGBM 立体匹配算法相结合的方法,对目标边缘信息匹配效果较好,而非边缘区域的信息丢失;本文采用的方法降低了图像中的噪声,对视差图

进行了平滑处理,保留了边缘信息,同时对异常值进行了处理,效果更好,更有利于深度估计的计算。

2) 消融实验

对双目校正后得到的双目图像分别进行 SGBM、引导滤波+SGBM、WLS+SGBM 和本文方法得到的目标视差图如图 7 所示,运算时间如表 6。

由匹配结果视差图可知,SGBM 算法得到的视差图有很多噪声影响,不够平滑,容易出现特征点匹配效果不好导致定位不到目标的情况。引导滤波+SGBM 的方法消除了图像中的噪声,但不够平滑。WLS+SGBM 的方法对视差图进行了平滑处理,但图像细节展现的略差。本文提出的引导滤波+异常值处理+WLS+SGBM 方法降低了图像中的噪声,对视差图进行了平滑处理,保留了边缘信息,对异常值进行了处理,提高了立体匹配的质量和准确性。

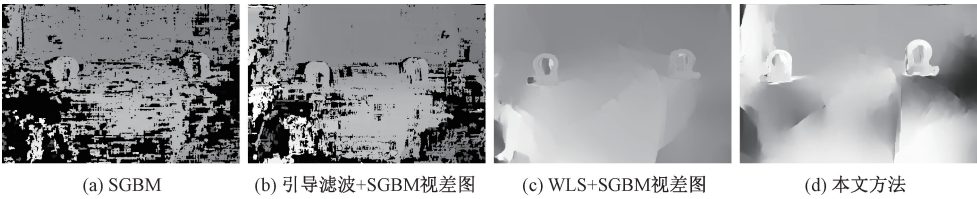


图7 各立体匹配算法消融实验视图

表6 各立体匹配方法的运算时间

立体匹配算法	运算时间/ms
SGBM	69.2
引导滤波+SGBM	73.8
WLS+SGBM	77.9
本文方法	81.6

3.4 目标定位实验

得到视差图后,便可以对拖车钩目标进行精确定位。现实场景中进行实时的拖车钩目标检测与定位如图8所示。



图8 实时的目标检测与定位

实验1:本文进行了6组实验,先用卷尺测量出相机距右拖车钩2000mm的位置,以这个位置为定点,然后在这个位置分别用SGBM算法、引导滤波+SGBM算法、WLS+SGBM算法和本文所提的引导滤波+异常值处理+WLS+SGBM算法进行消融实验,得到距离计算结果如图9所示。

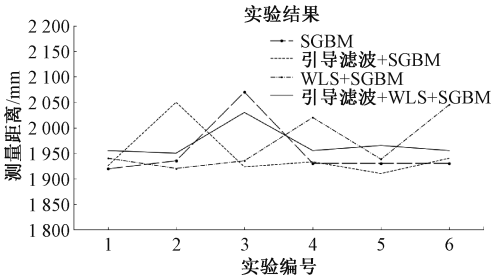


图9 不同算法在2000mm远的距离测量折线

由实验结果可以看出,本文提出的引导滤波+异常值处理+WLS+SGBM改进的立体匹配算法的误差最小,测

量距离最为准确。
实验2:本文进行了3组实验,用不同的方法分别在右拖车钩目标距离相机1000、2000、3000mm远处进行实验,先用卷尺测出相机与右拖车钩目标的实际距离,再用各改进的算法和本文所提出的方法分别去计算距离,进行对比试验,测量误差。其中方案(1)为YOLOv5结合SGBM算法,方案(2)为YOLOv5结合文献[13]的Census+SGBM算法,方案(3)为YOLOv5结合文献[14]的Roberts+SGBM算法,方案(4)为本文方法,得到的实验具体数据如表7~9。

表7 真实距离1000mm时的实验数据

方案序号	实际值/mm	测量值/mm	误差/mm	相对误差/%	FPS/fps
(1)	1000	1031	31	3.10	16
(2)	1000	1033	33	3.30	18
(3)	1000	1025	25	2.50	15
(4)	1000	1020	20	2.00	15

表8 真实距离2000mm时的实验数据

方案序号	实际值/mm	测量值/mm	误差/mm	相对误差/%	FPS/fps
(1)	2000	1922	78	3.90	16
(2)	2000	1917	83	4.15	17
(3)	2000	1948	52	2.60	15
(4)	2000	1955	45	2.25	14

表9 真实距离3000mm时的实验数据

方案序号	实际值/mm	测量值/mm	误差/mm	相对误差/%	FPS/fps
(1)	3000	2857	143	4.77	15
(2)	3000	2810	190	6.33	16
(3)	3000	2869	131	4.37	15
(4)	3000	2901	99	3.30	13

实验结果表明,本文所提改进的SGBM立体匹配算法,在3m以内误差可控制在3.5%以内,测量误差更小,说明了本文所提基于SGBM改进的立体匹配算法相对于其它改进算法精度更高,平均实时帧率为14fps,满足了遥控拖车时,拖车钩的实时检测与定位的准确性和实时性的要求。

4 结 论

本文针对某些危险环境下需要拖车实施救援、救援人员难以靠近时,需要遥控操作拖车杠来完成拖车钩挂装问题,提出了基于双目视觉的拖车钩检测与定位方法。首先用双目相机实时获取图像,然后利用 ECSA-YOLOv5 目标检测算法去检测拖车钩,利用改进 SGBM 立体匹配算法去获取深度信息,并计算视差,进而计算出拖车钩的距离和相对角度,并在 Jetson Agx Xavier 开发板上进行了实验验证。实验结果表明双目相机能够准确的检测到拖车钩目标,在 3 m 以内误差可控制在 3.5% 以内,平均实时检测帧率为 14 fps,能够满足拖车钩遥控挂装时检测与定位的准确性和实时性的要求。

参考文献

- [1] 刘丽丽. 机器视觉技术在汽车制造行业中的应用研究[J]. 汽车测试报告, 2023(7): 88-90.
- [2] ZHANG P, HOU W, WU D, et al. Real-time detection of small targets for video surveillance based on MS-YOLOv5[C]. 2023 6th International Conference on Artificial Intelligence and Big Data (ICAIBD), IEEE, 2023: 690-694.
- [3] LO E. Target detection algorithms in hyperspectral imaging based on discriminant analysis[J]. Journal of Image and Graphics, 2019, 7(4): 140-144.
- [4] REN S Q, HE K M, 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] 符强, 孔健明, 纪元法, 等. 基于双目视觉的无人机实时测距算法[J]. 电光与控制, 2023, 30(4): 94-99.
- [6] YE Q, CHENG Y, ZHANG M, et al. Research on flame location and distance measurement method based on binocular stereo vision[C]. 2020 Chinese Automation Congress(CAC), IEEE, 2020: 4089-4094.
- [7] FU P, CHEN X D, WU L. Research on absolute positioning of binocular vision based on corner[J]. Journal of Electronic Measurement and Instrument, 2018, 32(3): 1-8.
- [8] 张鹏, 黄亮, 杨露菁, 等. 基于海天线标定的海上单目测距方法[J]. 电光与控制, 2023, 30(9): 36-41, 73.
- [9] ZHANG H, TANG C, SUN X, et al. A refined apple binocular positioning method with segmentation-based deep learning for robotic picking[J]. Agronomy, 2023, 13(6): 1469.
- [10] 梁纲, 栗晓政, 饶宇飞, 等. 一种 Transformer 引导 YOLOv5 对高空吊钩违规操作识别[J]. 电工技术, 2023(10): 1-4.
- [11] 刘翀豪, 潘理虎, 杨帆, 等. 改进 YOLOv5 的轻量化口罩检测算法[J]. 计算机工程与应用, 2023, 59(7): 232-241.
- [12] ZHENG L, WANG X, WANG Q, et al. A fabric defect detection method based on improved yolov5[C]. 2021 7th International Conference on Computer and

Communications(ICC), IEEE, 2021: 620-624.

- [13] 毛先胤, 邢懿, 罗国强, 等. 基于改进 SGBM 算法的输电线路覆冰厚度测量[J]. 自动化与仪器仪表, 2021(11): 23-26, 31.
- [14] 乔超, 李帅康, 洪荣晶, 等. 基于 SGBM 算法的可转位刀片几何参数测量[J]. 工具技术, 2023, 57(9): 162-165.
- [15] QI S H, XU S S, XIE T, et al. AAM: An advanced attention module in convolutional neural networks[C]. 2020 2nd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI), IEEE, 2020: 371-374.
- [16] ZHANG Y P, SHAO Z H. Research on binocular forest fire source location and ranging system[C]. 2020 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS), IEEE, 2020: 199-202.
- [17] LUO D, LIAO Z X. Binocular stereo vision ranging based on OpenCV [J]. Popular Science and Technology, 2011, 4: 52-53.
- [18] 颜麟, 曹守启. 基于双目视觉的无人补料装置测距技术[J]. 上海海洋大学学报, 2023, 32(5): 1006-1014.
- [19] 王正家, 景嘉宝, 王思宇. 基于双目视觉的车辆外廓尺寸测量方法[J]. 电子测量技术, 2023, 46(12).
- [20] HIRSCHMULLER H. Stereo processing by semiglobal matching and mutual information [J]. IEEE Transactions on pattern analysis and machine intelligence, 2007, 30(2): 328-341.
- [21] 赵鹏翔, 冯培, 欧建鑫, 等. 基于双目立体视觉的纱线运动参数识别[J]. 河南理工大学学报(自然科学版), 2023, 42(3): 103-110.
- [22] ZHENG G X, YANG Y, LI Y, et al. A binocular location method based on feature point matching and region pixel matching [C]. 2021 2nd International Conference on Artificial Intelligence and Computer Engineering(ICAICE), IEEE, 2021: 637-641.
- [23] HE K M, SUN J, TANG X O. Guided image filtering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 35(6): 1397-1409.
- [24] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475.

作者简介

李冰, 副教授, 主要研究方向为模式识别与计算机视觉。
E-mail: li_bing_hb@126.com

王豪伟, 硕士研究生, 主要研究方向为模式识别与计算机视觉。
E-mail: 220222216165@ncepu.edu.cn

韩宇辰, 硕士研究生, 主要研究方向为模式识别与计算机视觉。
E-mail: 220222216105@ncepu.edu.cn

胡钧涛, 硕士研究生, 主要研究方向为模式识别与计算机视觉。
E-mail: 220222216157@ncepu.edu.cn

翟永杰(通信作者), 教授, 主要研究方向为模式识别与计算机视觉。
E-mail: zhaiyongjie@ncepu.edu.cn