

基于偏振图像的低照度场景多目标检测算法

寻华生^{1,2}, 张晶晶^{1,2,3}, 刘晓³, 李腾^{1,2}, 年福东^{1,4}, 张馨^{1,2}

(1. 安徽大学 计算智能与信号处理教育部重点实验室, 安徽 合肥 230601;

2. 偏振光成像探测技术安徽省重点实验室, 安徽 合肥 230031; 3. 中国科学院

通用光学定标与表征技术重点实验室, 安徽 合肥 230031; 4. 合肥学院 先进制造工程学院, 安徽 合肥 230601)

摘要: 偏振光反射信息可直接反演目标本征特性, 且在传输过程中具备较强的抗干扰特性, 因此偏振成像技术可适用于多种复杂环境中的智能监控、交通监察领域。近年来使用深度学习判读图像检测目标的方法迅速发展, 已经广泛应用于图像处理的各个领域。本文提出了一种基于偏振图像与深度神经网络算法的行人、车辆多目标检测算法 YOLOv5s-DOLP。首先, 通过实时获取到偏振图像进行偏振信息解析, 获取目标偏振度图像。其次, 为增强偏振度图像中检测目标与背景存在高对比度的特性, 在主干网络中引入通道注意力与空间注意力, 提升网络特征进行自适应学习的能力。此外, 使用 K-means 算法对目标位置信息进行聚类分析, 加快网络在偏振度图像的学习速度, 提升目标检测精度。实验结果显示, 该算法结合了偏振成像和深度学习目标检测的优势, 对于低照度复杂场景中的车辆、行人目标检测效果好、检测速度快, 对于道路车辆的目标检测、识别与跟踪具有一定的应用价值。

关键词: 偏振成像; 神经网络; YOLOv5s; 多目标检测; 注意力机制

中图分类号: TP183 **文献标识码:** A **文章编号:** 1001-8891(2022)05-0483-09

Multi-Target Detection of Low-Illuminance Scene Based on Polarization Image

XUN Huasheng^{1,2}, ZHANG Jingjing^{1,2,3}, LIU Xiao³, LI Teng^{1,2}, NIAN Fudong^{1,4}, ZHANG Xin^{1,2}

(1. Key Laboratory of Intelligent Computing and Signal Processing of Ministry of Education, Anhui University, Hefei 230601, China;

2. Anhui Province Key Laboratory of Polarization Detection Technology, Hefei 230031, China;

3. Key Laboratory of Optical Calibration and Characterization, Chinese Academy of Sciences, Hefei 230031, China;

4. School of Advanced Manufacturing Engineering, Hefei University, Hefei 230601, China)

Abstract: Polarized light reflection information can directly invert the intrinsic characteristics of a target and has strong anti-interference characteristics in the transmission process. Thus, polarization imaging technology can be applied to the fields of intelligent monitoring and traffic monitoring in various complex environments. In recent years, deep-neural-network methods for interpreting image detection targets have been developed rapidly and widely used in various fields of image processing. In this study, a vehicle multi-target detection algorithm based on polarized images and deep learning is proposed. First, the target polarization degree image can be obtained by acquiring the polarization image in real time and analyzing the polarization information. Second, to enhance the high contrast between the detection targets and the background in the polarization image, channel attention and spatial attention are introduced into the backbone network to improve the ability of the network features to perform adaptive learning. In addition, the K-means algorithm is used to perform clustering analysis on the target location information, thereby increasing the network's learning speed in the polarization image and improving the progress of target detection. The experimental results show that this method is effective and fast for vehicle detection in complex scenes with low illumination. This method

收稿日期: 2021-06-15; 修订日期: 2021-08-03.

作者简介: 寻华生 (1996-), 男, 硕士研究生, 主要研究方向为图像处理、深度学习。E-mail: Z19201027@stu.ahu.edu.cn.

通信作者: 张晶晶 (1974-), 女, 副教授, 博士, 主要研究方向为图像处理、遥感信息处理和模式识别。E-mail: fannyzj@ahu.edu.cn.

基金项目: 中国科学院通用光学定标与表征技术重点实验室开放研究基金项目; 偏振光成像探测技术安徽省重点实验室开放基金项目; 国家自然科学基金青年科学基金项目 (61902104); 安徽省自然科学基金项目 (2008085QF295); 安徽高校自然科学研究项目 (KJ2020A0651); 安徽省自然科学基金项目 (1808085MF).

combines the advantages of polarization imaging and deep-learning target detection and has substantial application scope in road vehicle target detection, recognition, and tracking.

Key words: polarization image, neural network, YOLO v5s, multi-target detection, attention mechanism

0 引言

在传统的监控领域,图像判读处理和目标识别一直为研究热点。一直以来,场景监控使用可见光图像居多,而光线角度的影响,雨雾等散射介质对成像的影响会造成成像效果不佳,设备有效探测距离不远。偏振成像与可见光相比,偏振成像反映的是物体的偏振信息,偏振成像受成像环境因素的影响较小,探测距离相对较远,目标的轮廓对比度强。偏振成像对于复杂环境下的监控等应用场景有较高的应用价值。

目前的偏振成像探测设备具备的功能除图像显示、人眼直接判读外,很少具有自动检测与识别等功能,对于低照度场景的目标检测识别智能化程度较低。于洵等人^[1]利用光谱信息和偏振信息相融合的目标识别方法,提高了伪装目标识别的准确率。李从利等人^[2]在雾天条件下利用偏振信息评估图像。李小明^[3]用偏振信息检测沙漠背景下的目标,该方法能有效提取沙漠背景下的金属靶信息。目前,偏振探测已经应用在多个研究领域中,并已验证了偏振信息的独特性以及在目标检测识别中的可行性。基于深度学习^[4]的目标检测^[5-6]和目标跟踪算法被提出。迄今为止,基于深度学习的目标检测算法大致分为两类:①双步目标检测算法,如 Fast R-CNN (Region-convolutional Neural Network)^[7]、Faster R-CNN^[8]、Mask R-CNN^[9],其将检测分为两步,先使用区域候选网络(region proposal network, RPN)来提取目标特征信息,其次运用检测网络针对候选目标的位置等信息的预测及识别;②单步目标检测算法,如 SSD (single shot Multibox detector)^[10]、YOLO (you only look once) 系列^[11-14]等,此类算法不使用 RPN,直接通过网络来产生目标的位置和类别信息,是一种端到端的目标检测算法。在复杂环境,例如低照度等,上述目标检测算法针对可见光图像的检测效果不佳。基于此,本文在 YOLO v5 中最轻便的 YOLO v5s 的基础上进行改进,提出基于偏振成像的 YOLO v5s 的目标检测算法 YOLO v5s-DOLP。数据层面,首先,本文所构建的偏振数据与可见光相比,偏振成像可以通过反映物体的偏振信息来减少环境因素的影响,远距离的探测和检测目标轮廓与背景高对比度特性,可以在低照度下增强算法在偏振成像中的目标检测能力。其次,对本文构建的偏振成像数据使用 K-means 算法重新获取目标的位置等信

息,提升目标检测精度。网络层面,首先,为了减少 Focus 操作对于偏振度图像中检测目标与背景相交边缘的灰度值梯度变化和背景深处的小目标轮廓被切割的情况,YOLO v5s-DOLP 先将输入数据进行卷积处理,保持检测目标与背景的高对比度特性和小目标的完整轮廓特征,对 Focus 模块结果进行信息补充。其次,在主干网络中,使用嵌入注意力模块卷积块代替标准卷积操作,提升网络在偏振成像数据特征中的自适应学习能力,抑制背景信息,更有效地检测到目标,减少误检和漏检问题。最后,受图像多尺度表达思想影响,将偏振度图像变换尺寸,经过轻量级卷积操作后,将输出与对应尺寸的特征图通道叠加,融合不同感受野的特征信息。

1 偏振成像数据的构建与预处理

1.1 偏振成像数据的构建

光的偏振信息采用斯托克斯 (Stokes) 矢量 (I, Q, U, V) 描述。其中 I 表示光的总强度; Q 和 U 表示线偏振光分量之差; V 表示右旋与左旋圆偏振光分量之差,在自然界大气背景及目标物对太阳光反射的偏振效应中,圆偏振分量极小,设定 $V=0$,同时利用 Stokes 矢量可以计算出场景的线偏振度 (degree of linear polarization, DOLP) 和偏振角参数 (angle of polarization, AOP)。确定0参考方向后,选取 0° 、 60° 、 120° 偏振方向的图像 $I(0^\circ)$ 、 $I(60^\circ)$ 、 $I(120^\circ)$ 后,由公式表达为:

$$\begin{cases} I = \frac{2}{3} [I(0^\circ) + I(60^\circ) + I(120^\circ)] \\ Q = \frac{2}{3} [2I(0^\circ) - I(60^\circ) - I(120^\circ)] \\ U = \frac{2}{\sqrt{3}} [I(60^\circ) - I(120^\circ)] \end{cases} \quad (1)$$

$$\text{AOP} = \frac{1}{2} \arctan \left(\frac{U}{Q} \right) \quad (2)$$

$$\text{DOLP} = \frac{\sqrt{Q^2 + U^2}}{I} \quad (3)$$

通过式(1)可得斯托克斯参量 I 、 Q 、 U ,通过公式(2)、(3)偏振角 AOP 与解析线偏振度 DOLP。

图 1 所示为偏振图像采集偏振参量流程,光线经过 3 个方向的偏振分析器、3 个探测器分别成像。采集到的数字图像经过计算机的处理,各对像素经运算

后得到的即为偏振信息图像的像素。

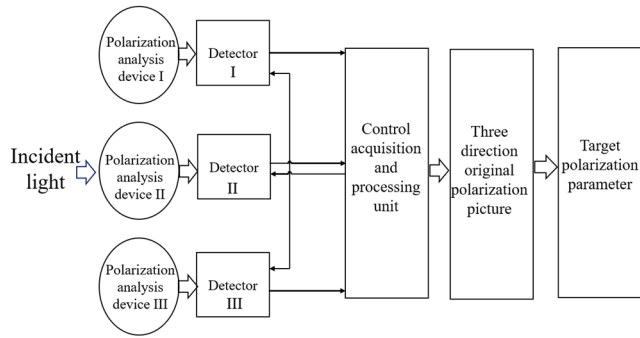


图1 偏振图像采集与偏振参量解析流程

Fig.1 Polarization image acquisition and polarization parameter analysis process

图2(a)~(c)为 0° 、 60° 和 120° 三个方向的偏振原图, (d)为 I 图, (e)、(f)依次为 Q 图和 U 图, (g)、(h)分别为偏振角图像和偏振度图像。通过对比 I 图、AOP图、DOLP图, 可知背景图像的偏振度相较被检测物体的偏振度较低, 如车辆等。其原因为检测目标不属于自然目标, 被测物比较光滑, 对比自然目标, 会以镜面反射为主, 地表等背景图像表面纹理不均匀, 以漫反射为主。故在检测目标与背景相交的边缘, 偏振度和偏振角产生较为明显的变化, 增强检测目标的边

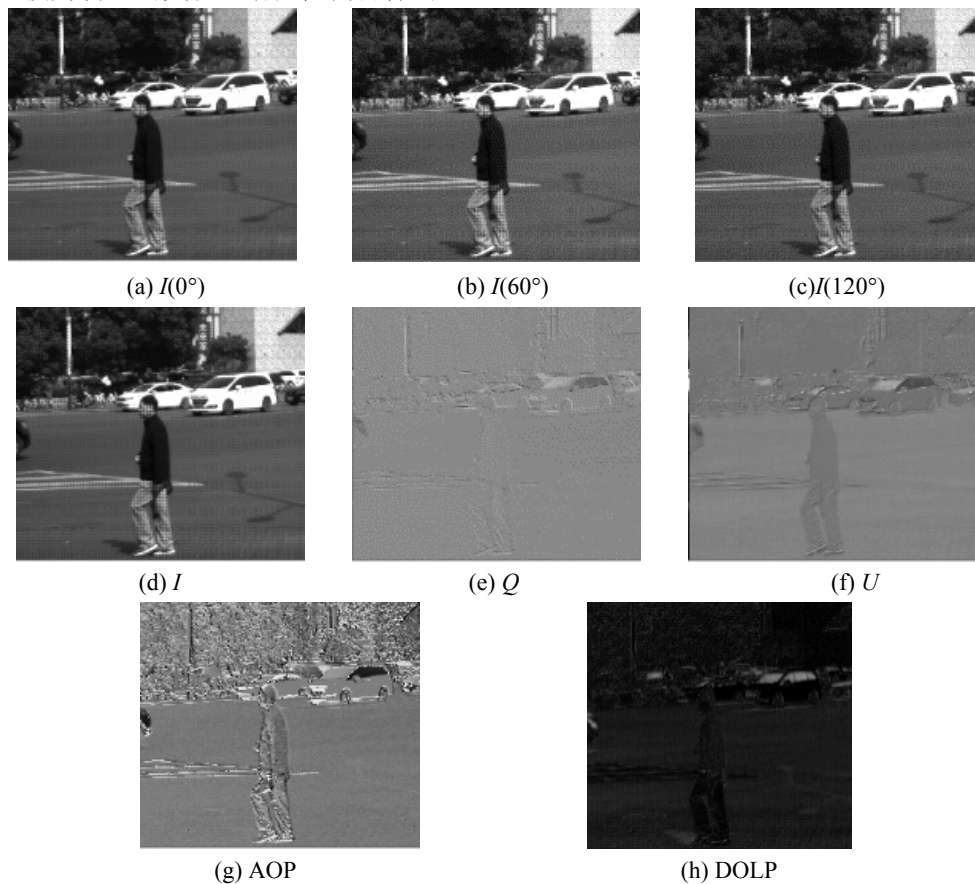


图2 偏振参量解析示意图

Fig.2 Process of polarization parameter analysis

缘轮廓, 提升目标的细节特性。

1.2 偏振度图像的锚点修正

YOLO v5s 网络初始锚点是在 COCO 以及 Pascal VOC 两种可见光图像上使用 K-means 算法对目标框位置信息进行聚类分析而得。本文做构建的偏振度图像数据经数据扩容, 检测目标种类和目标位置分布等参数与 YOLO v5s 网络的初始锚点有所差异, 为了提升网络对于偏振数据特征的学习能力, 提升目标检测精度, 通过 K-mean 算法重新对偏振成像数据目标框位置信息进行聚类分析。偏振度图像的大小为 608×608 , 新的锚点分别为: (60, 43), (100, 60), (73, 143), (29, 64), (182, 82), (27, 22), (11, 30), (16, 40), (40, 30)。

2 YOLO v5s 模型改进

图3为基于偏振度图像目标检测所提出的算法流程图。由图1的流程得到3个角度的偏振原图, 通过公式(1)得到 I 图、 Q 图、 U 图, 有公式(3)求解得到偏振度图像 DOLP 图。本文将偏振度图像输入到 YOLO v5s-DOLP 中训练, 获得 YOLO v5s-DOLP 模型, 最终获取检测结果。

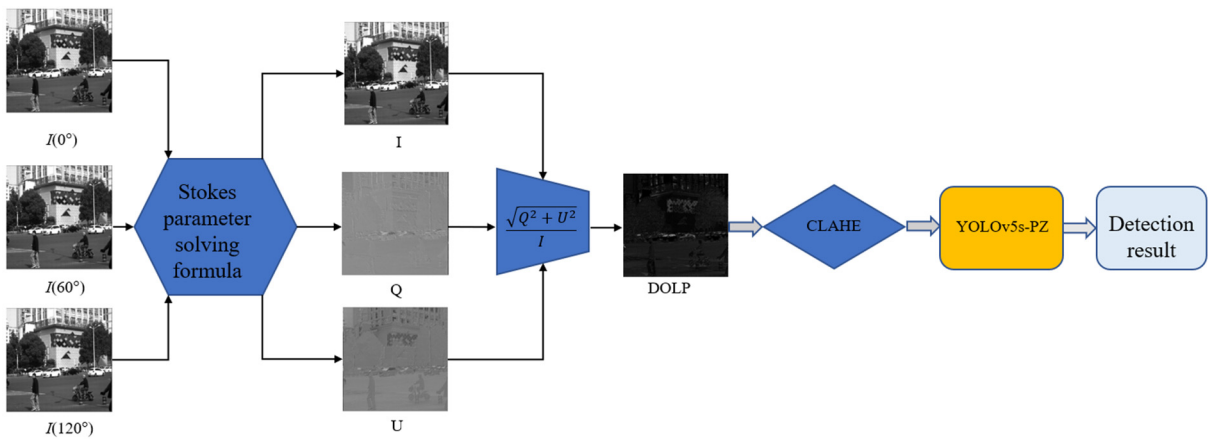


图3 算法流程图

Fig.3 Algorithm flowchart

2.1 针对 Focus 模块的特征融合

Focus 模块是 YOLOv5 对于输入图片进行切片的操作，如图 4 所示。具体操作是在输入尺寸为 $H \times W$ 的图像中每隔一个像素取一个值，遍历整张图像后得到 4 张尺寸为 $H/2 \times W/2$ 图像。其目的是将 W 、 H 的信息集中到通道空间，再通过 3×3 的卷积对其特征进行充分提取。此操作虽然在一定程度上帮助网络进一步提取图片特征，但对于偏振度图像，此操作不仅破坏了检测目标与背景相交边缘的灰度值梯度变化，削弱偏振度图像中检测目标对于背景的高对比度，如图 5 所示，而且对于图像背景深处的小目标容易造成进一步模糊或者分割等问题，造成漏检的几率增加。基于此，本文在主干网络开启分支，对于输入数据进行卷积操作，保留检测目标与背景相交边缘的灰度值的梯度变化和小目标完整的轮廓特征，再与 Focus 模块的输出进行信息交互，补充检测目标的特征细节。

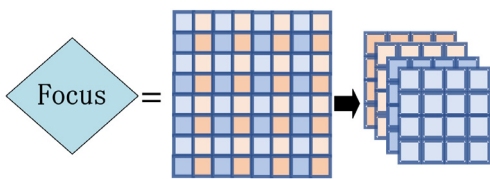


图4 Focus 操作示意图

Fig.4 Focus operation diagram

图 5 中(a)、(b)分别为拉普拉斯算子对原始偏振度图像、Focus 结果的图像边缘检测。整体观测，(a)列的边缘梯度比(b)列的边缘梯度变化清晰、明了。细节观测，如图中矩形框部分所示，相比较(a)列，Focus 结果的车辆及行人的梯度变化明显减弱，降低了偏振度图像的特性。

2.2 基于注意力机制的卷积操作

在偏振度图像中，检测目标与背景有着较高的对比度，即检测目标的轮廓较为清晰。为了进一步提高

图像信息的处理效率与检测精度，本文在主干网络引入了注意力机制。类似于人类在视线范围内聚焦于某一片区域，注意力机制通过快速扫描图像的整体信息，加大重点关注区域的权重，提升网络筛选信息的效率，抑制不相关信息。其中，基于通道注意力与空间注意力的卷积块注意力模块 (convolutional block attention module, CBAM)^[16]就是为卷积网络设计的轻量级、有效的注意力模块。对于主干网络的特征图 $F \in R^{C \times H \times W}$ ，其中 C 、 H 、 W 依次为特征图通道数、高度、长度，融入注意力机制会将特征 F 依次通道注意力机制 $M_c \in R^{C \times 1 \times 1}$ ，空间注意力机制 $M_s \in R^{1 \times H \times W}$ ，即：

$$F' = M_c(F) \otimes F \tag{4}$$

$$F'' = M_s(F') \otimes F' \tag{5}$$

式中： \otimes 为矩阵逐元素相乘，其过程如图 6 所示。



(a) 原图的边缘检测 (b) Focus 后的边缘检测

(a) Edge detection of image (b) Edge detection after focus

图5 拉普拉斯算子^[15]检测结果

Fig.5 Detecting result of Laplace

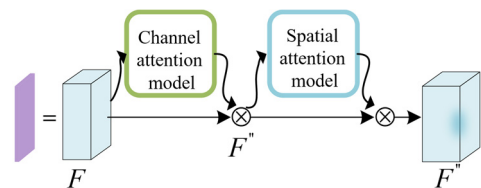


图6 卷积块注意力模块

Fig.6 Convolutional block attention module

通道注意力模块将特征图在空间维度上进行压缩, 得到一维矢量, 如图7所示。其中最大池化(max pooling, MaxPool)和平均池化(average pooling, AvgPool)用来聚合特征映射的空间信息, 再将信息输入到共享网络, 压缩输入特征图的空间维数, 得到 F_{max}^c 和 F_{avg}^c , 并逐元素求和, 产生通道注意力图:

$$M_c(F) = \sigma\{MLP(AvgPool(F)) + MLP(MaxPool(F))\} = W_1(W_0(F_{max}^c) + W_1(W_0(F_{avg}^c))) \quad (6)$$

式中: σ 为激活函数; MLP为共享网络; W_0 、 W_1 分别为共享网络的第一层和第二层。

空间注意力模块如图8所示, 将通道注意力模块的特征图作为输入, 在通道维度分别进行了平均值池化和最大值池化, 得到 F_{avg}^s 和 F_{max}^s , 将它们进行通道叠加, 再依次经过卷积和激活函数, 产生空间注意力图:

$$M_s(F) = \sigma\{f^{7 \times 7}([AvgPool(F); MaxPool(F)])\} = \sigma\{f^{7 \times 7}([F_{avg}^s; F_{max}^s])\} \quad (7)$$

式中: σ 为激活函数; $f^{7 \times 7}$ 代表卷积核尺寸 7×7 的卷积层。

对于一个中间特征图, 卷积块注意力模块会沿着两个独立的维度依次推断注意力图, 然后将注意力图与输入的特征图相乘以进行自适应特征优化。CBAM是轻量级的注意力模块, 因此可以用较小的开销而将其无缝集成到YOLOv5s-DOLP网络中的卷积层中, 进行目标检测模型的训练。在本文中, 嵌入注意力模块的卷积层可以抑制特征图中的干扰信息, 聚焦有效信息, 从而提升目标检测精度。

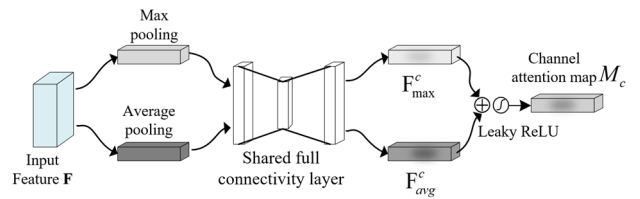


图7 通道注意力模块

Fig.7 Channel attention module

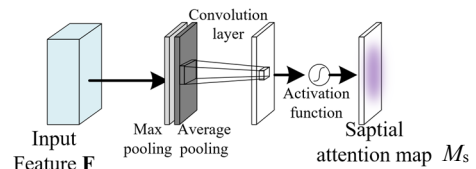


图8 空间注意力模块

Fig.8 Spatial attention module

2.3 多尺度偏振度图像的特征融合

在本文中, 被检测物的表面偏振信息通过偏振成像技术获得, 通过背景与被检测物的偏振特性不一致性, 进一步提高被检测物体与背景图像的对比度, 提升被检测物体与背景的边缘分离度, 从而使目标检测的精确度上升。受图像多尺度表达思想启发, 针对偏振成像受成像环境因素的影响较小, 探测距离相对较远, 目标的轮廓对比度强等特点, 本文将偏振度图像变换尺寸, 通过轻量级卷积模块后, 将输出与对应尺寸的特征图通道叠加, 进行特征融合, 如图9所示。实验表明, 偏振度图像中目标与背景的高对比度可以有效地提高检测目标准确率。

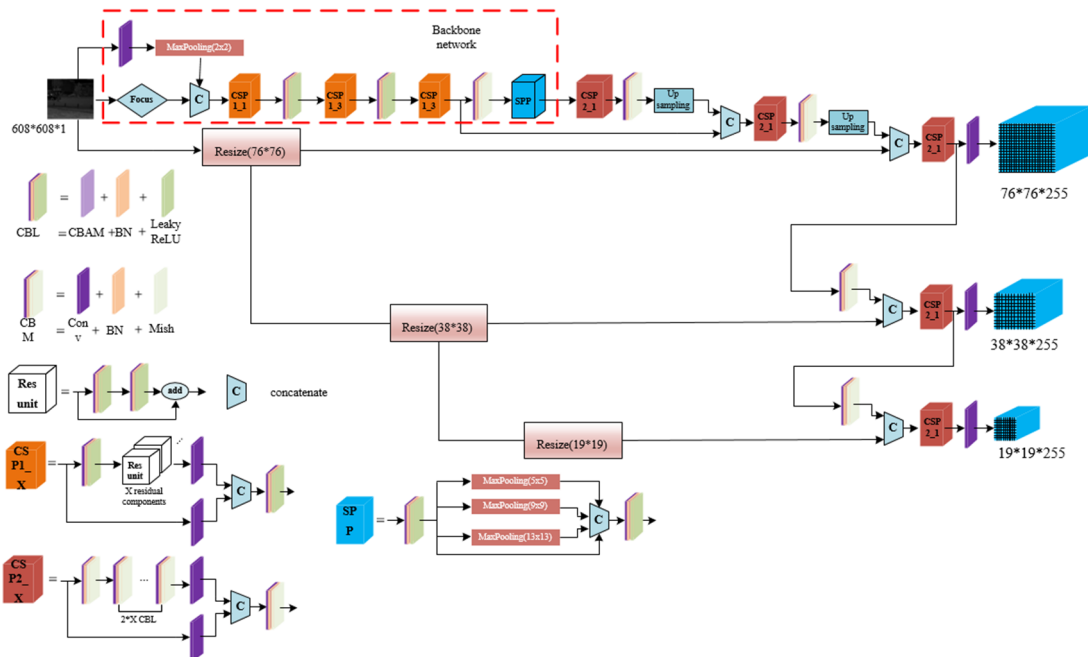


图9 YOLO v5s-DOLP网络

Fig.9 YOLO v5s-DOLP network

3 实验过程与结果分析

3.1 实验评价指标与环境配置

由于目前没有公开的关于偏振度图像的目标检测数据集,故本文在数据准备阶段,选取低照度环境,行人和车辆较为密集的路口采集若干批偏振成像数据。经过筛选,最终获取4000张偏振度图像,使用图像数据标注工具labelme对获取的偏振度图像进行标注。数据按照8:1:1比例进行随机分配,即训练集3200张、验证集400张、测试集400张。实验使用的是分光型偏振成像设备,具体型号参数如表1所示。

表1 分光型偏振成像设备

Max resolution	Frame	Image sensor	Pixel size/ μm	ADC
2488×2048	65	Sony IMX250, CMOS, 2/3 Global shutter	3.45	10 bit/ 12 bit

为了验证本文工作针对偏振度成像的目标检测的有效性,客观地衡量目标检测模型的性能,本文引入了精确率 Precision、召回率 Recall、平均精度 AP 及均值平均精度 mAP 作为检测训练模型的评价指标。首先根据真实标签将检测结果分为真正例 TP (true positive)、真反例 TN (true negative)、假正例 FP (false positive), 假反例 FN (false negative)。Precision 表示左右被预测为正的样本中实际为正样本的概率, Recall 表示为正的样本中被预测为正样本的概率。当 IOU 阈值和物体置信度阈值设置较高时,计算出的 Precision 较高, Recall 较低,反之亦然。AP 为某类 Precision-Recall 曲线下的面积,面积越大,AP 的数值越高,模型对某类目标的检测效果越好。mAP 代表所有种类 AP 的均值,从整体上反映检测模型性能。

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP}) \quad (8)$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN}) \quad (9)$$

实验环境配置如表2所示。算法使用 Pytorch 实现,并在2块 NVIDIA 1080Ti 显卡上运行。模型训练均采用 Adam 优化器,单个训练批次的样本数量根据 GPU 的显存上限设为64。学习率设为 $1 \times e^{-5}$,模型训练轮数 epoch 设为50。训练集和测试集分别包含3200和400张,尺寸大小为 608×608 。

3.2 实验结果与分析

实验分为两个阶段,第一阶段,将 YOLOv5s-DOLP 于目前主流的检测算法在偏振数据上进行精度与帧率的测量。第二阶段为消融实验验证本文提出的

YOLOv5s-DOLP 的有效性。其中 Faster R-CNN 的候选框数目为256,各算法锚点均为重新获取锚点数值。

表2 网络训练环境

Name	Configure
CPU	Intel Xeon E5-2630
GPU	NVIDIA 1080Ti * 2
Operating system	Ubuntu 18.04
Parallel computing library	Cuda10 + Cudnn7.4
Image processing	Python3.6、Opencv3.4.0
Deep learning framework	Pytorch

其中, FasterR-CNN 的 IOU 阈值为0.5, YOLOv4 的 IOU 阈值为0.5, YOLOv5s-DOLP 的 IOU 阈值为0.5。

第一阶段,由表3可知,各类算法对于 Car 类的检测精度均有较高的精度,验证了1.1节中所述,车辆作为非自然目标,其表面具有较强的镜面反射,在车辆与其他自然物体相交的边缘,偏振度和偏振角产生较为明显的变化,强化了车辆的特征,故 Car 类的 AP 值较高。在 Person 类检测中, Faster R-CNN(Res 50), YOLO v4(Res 50)的 AP 值相同, YOLO v5s-DOLP 的 AP 值相较其他两种算法有明显提升,证实了网络在嵌入注意力机制和对 Focus 模块进行特征融合等操作的有效性。在检测速度方面, YOLO 系列算法要优于 FasterR-CNN 数倍, YOLO v5s-DOLP 的检测速度虽然相较于 YOLO v4 略慢,但是依旧满足实时检测的要求。图10为上述算法部分检测结果示意图。

表3 目标检测算法检测结果

Methods	AP		mAP	Test time/s
	Person	Car		
Faster R-CNN(Res 50)	91.6	97.5	94.6	0.103
YOLOv4(Res 50)	91.6	97.3	94.5	0.029
YOLOv5s-DOLP	94.8	98.9	96.9	0.039

第二阶段,现通过消融实验验证本文中的 YOLOv5s-DOLP 在低照度环境下结合偏振度图像进行多目标检测的有效性,通过逐步添加各模块及修改网络结构,详细地证明各模块及操作的作用。实验过程均采用相同的实验环境及配置,以 YOLOv5s 作为基础网络结构。为了便于表示,将针对 Focus 模块的融合操作视为 A,添加 CBAM 的操作视为 B,将原图像进行多尺度缩放,经过卷积后与不同感受野特征图的融合操作视为 C。

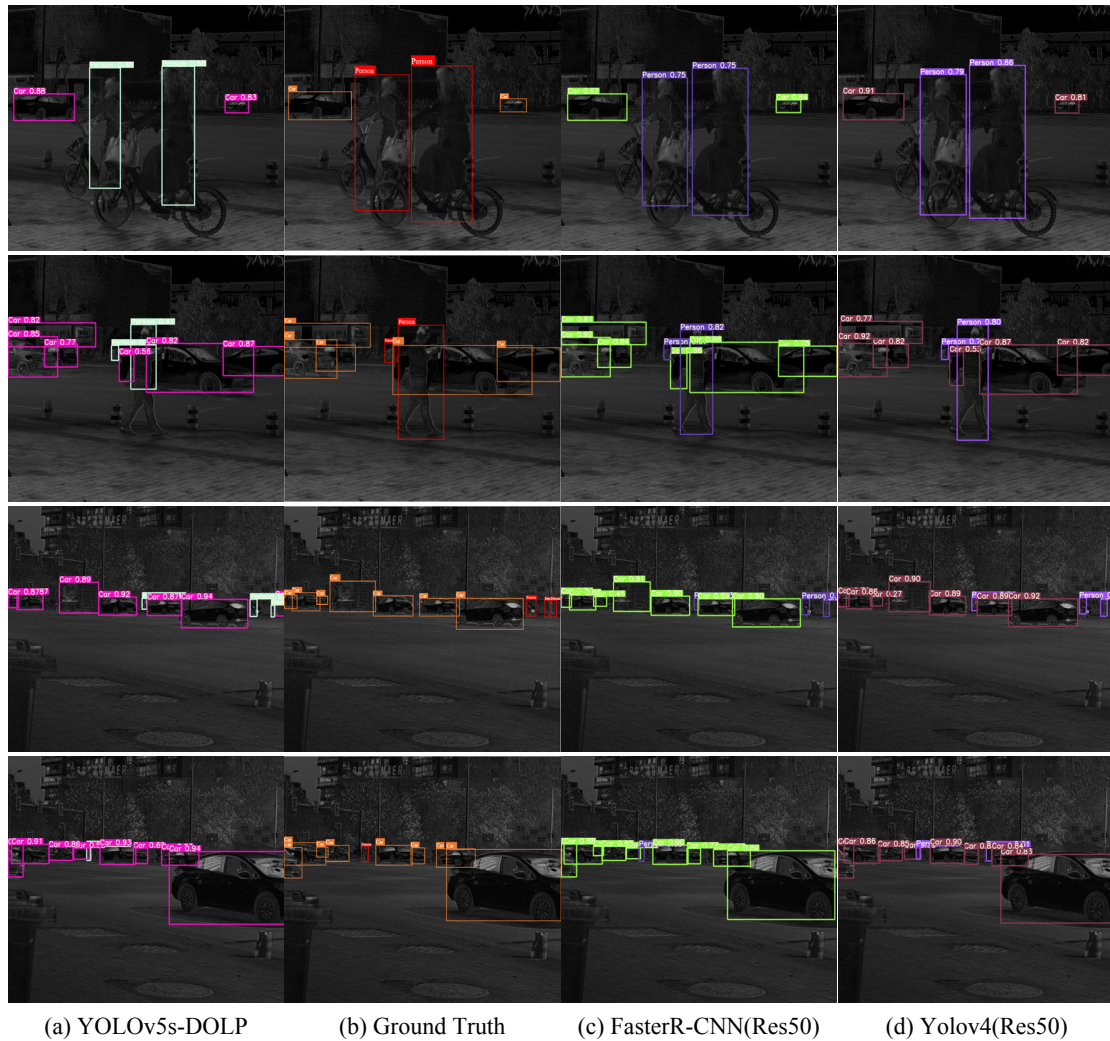


图 10 检测结果示例

Fig.10 Detect results of samples

在图 10 中, (a)表示基础结构 YOLOv5s 的 Precision-Recall 曲线, (b)、(c)、(d)依次表示经过 A 操作、B 操作、C 操作后的 Precision-Recall 曲线。图中各曲线分别代表行人、车辆和两者综合。通过图中各曲线包裹的面积可知添加使用注意力模块卷积块替换标准卷积块对于目标检测性能提升最为明显。相较于基础结构 YOLOv5s, 虽然在 Car 类的检测精度提升不大, 但在 Person 类的检测精度上有不小的提升, 证实了注意力机制对于提升目标检测网络的性能的有效性。

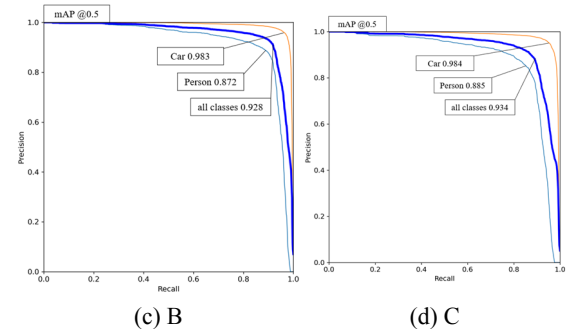
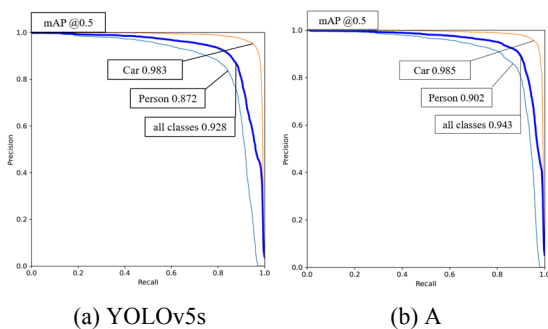


图 11 Precision-Recall 曲线

Fig.11 Precision-Recall curves

相比于表 3 中的 YOLOv5s-DOLP 模型, 表 4 和图 11 中 A、B、C 操作的参数量减少, 在检测速率上会提高。其中, 图 11(b)操作旨在弥补网络容易漏检背景深处的小目标, C 操作利用图像多尺度表达的思想进一步整合不同感受野层次的特征图, 两者对于目标检测精度都有提升, 且检测速度与基础结构 YOLOv5s 几乎持平, 检测一张图片的时间均在 0.027s 左右。B 操作添加 CBAM 导致参数量增多, 检测速度下降幅

度为 26%，但是检测精度提升幅度较大，相较于基础结构 YOLOv5s，AP 提升 4.8。提升较为明显的原因在于注意力模块卷积块在网络提取特征过程加速聚焦检测目标，提高网络对于图像特征的自适应学习能力，抑制背景等不相关信息，较少环境等因素对于目标检测的干扰，从而提高检测精度。图 12 为消融实验部分检测结果示例。

表 4 消融实验结果

Table 4 Results of ablation experiment

Methods	AP		mAP	Test time/s
	Person	Car		
YOLOv5s	87.2	98.3	92.8	0.027
A	90.2	98.5	94.3	0.029
B	92.0	98.8	95.4	0.034
C	88.5	98.4	93.4	0.027

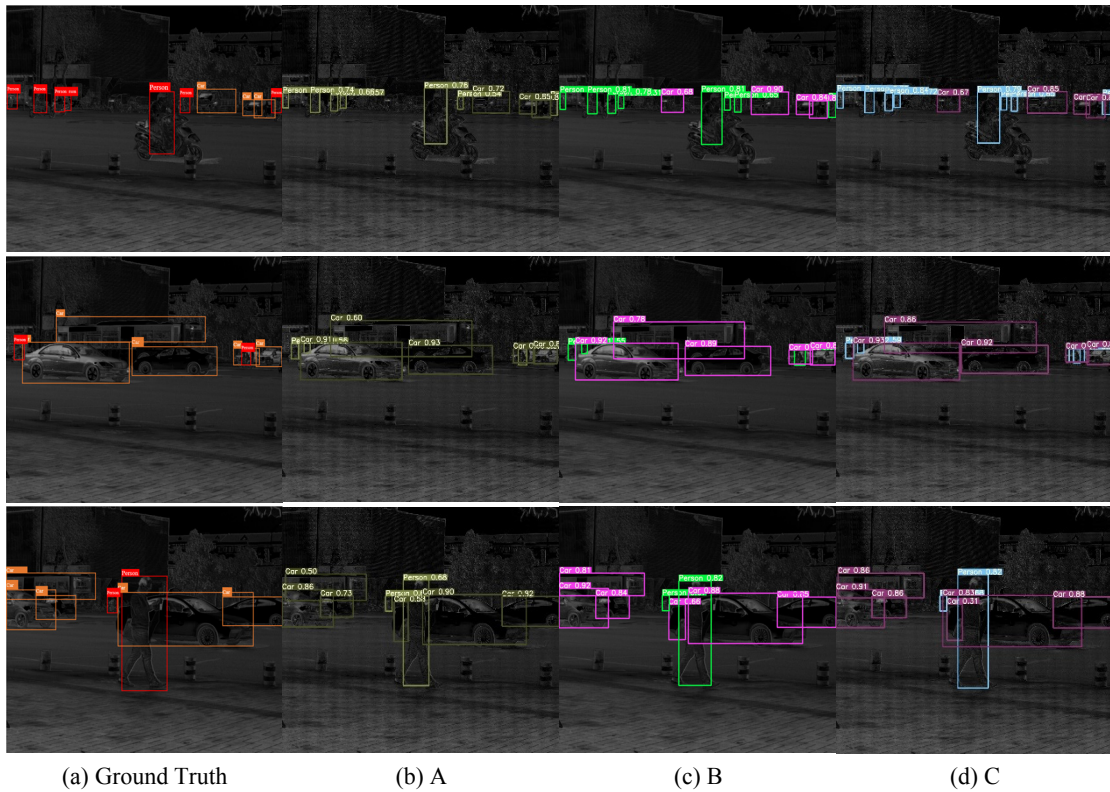


图 12 消融实验检测结果示例

Fig. 12 Detect results of samples of ablation experiment

4 结语

本文针对低照度下的多目标检测任务，提出基于偏振度图像的神经网络检测算法 YOLOV5s-DOLP。数据方面，本文构建的偏振成像数据在低照度环境下通过反映物体的偏振信息来减少环境因素的影响，通过检测目标与背景的高对比度，凸显目标的轮廓特征，提升目标检测进度。此外，本文根据偏振成像数据中被测物体的尺寸及位置信息重新获取锚点。网络层面，针对 Focus 操作降低检测目标与背景相交边缘的灰度值梯度的变化、破坏背景深处的小尺寸目标的完整轮廓等问题，本文在数据输入端开启分支，对原始数据进行卷积操作，再与 Focus 模块的输出进行信息融合，减少信息的丢失。其次，在主干网络内使用注意力模块卷积块代替标准卷积，强化网络

的自适应学习能力，从图像中快速选择行人等检测目标的信息，抑制背景信息。为了进一步整合网络深层的语义信息与浅层的细节特征，受图像多尺度表达思想启发，本文将原始的偏振度图像进行多尺度缩放，经过卷积操作后，使其与对应尺寸的特征图信息交互，融合不同感受野的特征信息。根据测试结果，该方法在某些情况下仍然有着改进的空间，例如在重叠、遮挡等干扰因素的结果存在一定偏差，这也是本文算法进一步优化的方向。

参考文献:

[1] 于洵, 杨焯, 姜旭, 等. 基于偏振光谱成像的目标识别方法研究[J]. 应用光学, 2016, 37(4): 537-541.
 YU Xun, YANG Ye, JIANG Xu, et al. Recognition of camouflage targets by polarization spectral imaging system[J]. *Journal of Applied Optics*, 2016, 37(4): 537-541.

- [2] 李从利, 薛松, 陆文骏, 等. 雾天条件下偏振解析成像质量评价[J]. 中国图象图形学报, 2017, 22(3): 366-375.
LI Congli, XUE Song, LU Wenjun, et al. Quality assessment of polarization imaging under foggy[J]. *Journal of Image and Graphics*, 2017, 22(3): 366-375.
- [3] 李小明, 黄勤超. 沙漠背景下红外偏振图像目标检测方法[J]. 红外技术, 2016, 38(9): 779-782, 792.
LI Xiaoming, HUANG Qinchao. Target detection for infrared polarization image in the background of desert[J]. *Infrared Technology*, 2016, 38(9): 779-782, 792.
- [4] 张荣, 李伟平, 莫同. 深度学习研究综述[J]. 信息与控制, 2018, 47(4): 385-397.
ZHANG Rong, LI Weiping, MO Tong. Review of deep learning[J]. *Information and Control*, 2018, 47(4): 385-397.
- [5] 王文秀, 傅雨田, 董峰, 等. 基于深度卷积神经网络的红外船只目标检测方法[J]. 光学学报, 2018, 38(7): 0712006.
WANG Wenxiu, FU Yutian, DONG Feng, et al. Infrared ship target detection method based on deep convolution neural network[J]. *Acta Optica Sinica*, 2018, 38(7): 0712006.
- [6] 罗海波, 许凌云, 惠斌, 等. 基于深度学习的目标跟踪方法研究现状与展望[J]. 红外与激光工程, 2017, 46(5): 0502002-0502002(7).
LUO Haibo, XU Lingyun, HUI Bin, et al. Status and prospect of target tracking based on deep learning[J]. *Infrared and Laser Engineering*, 2017, 46(5): 0502002-0502002(7).
- [7] Ross Girshick. Fast R-CNN[J]. *Computer Science*, 2015, 6: 1440-1448.
- [8] REN Shaoqing, HE Kaiming, Ross Girshick, 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.
- [9] HE Kaiming, Georgia Gkioxari, Piotr Dollar, et al. Mask R-CNN[C]// *Proceedings of the IEEE Computer Society Conference on Computer Vision*, 2017: 2961-2969.
- [10] LIU Wei, Dragomir Anguelov, Dumitru Erhan, et al. SSD: Single shot multibox detector[C]// *European Conference on Computer Vision*, 2016, 6: 21-27.
- [11] Joseph Redmon, Santosh Divvala, Ross Girshick, et al. You Only Look Once: unified, real-time object detection[C]// *Conference on Computer Vision and Pattern Recognition*, 2016, 6: 779-788.
- [12] Joseph Redmon, Ali Farhadi. YOLO9000: better, faster, stronger[C]// *Conference on Computer Vision and Pattern Recognition*, 2017, 7: 6517-6525.
- [13] Joseph Redmon, Ali Farhadi. YOLOv3: an incremental improvement [J/OL]. arXiv: 1804.02767, 2018.
- [14] Alexey Bochkovskiy, Chien-Yao Wang, Hong-Yuan Mark Liao. YOLOv4: optimal speed and accuracy of object detection[J/OL]. arXiv: 2004.10934, 2020.
- [15] 冈萨雷斯. 数字图像处理[M]. 第三版, 北京: 电子工业出版社, 2011: 26-29.
Rafael C. Gonzalez. *Digital Image Processing*[M]. 3th, Beijing: Publishing House of Electronics Industry, 2011: 26-29.
- [16] Woo S, Park J, Lee J Y, et al. Cbam: Convolutional block attention module[C]// *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018: 3-19.
- [17] 张汝榛, 张建林, 祁小平, 等. 复杂场景下的红外目标检测[J]. 光电工程, 2020, 47(10): 200314.
ZHANG R Z, ZHANG J L, QI X P, et al. Infrared target detection and recognition in complex scene[J]. *Opto-Electron Eng.*, 2020, 47(10): 200314.
- [18] 官剑, 吕俊伟, 刘亮, 等. 红外偏振图像的舰船目标检测[J]. 光谱学与光谱分析, 2020, 40(2): 586-594.
GONG Jian, LYU Junwei, LIU Liang, et al. Ship target detection based on infrared polarization image[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(2): 586-594.
- [19] 游江, 刘鹏祖, 容晓龙, 等. 基于暗通道先验原理的偏振图像去雾增强算法研究[J]. 激光与红外, 2020, 50(4): 493-500.
YOU Jiang, LIU Pengzu, RONG Xiaolong, et al. Dehazing and enhancement research of polarized image based on dark channel priori principle[J]. *Laser & Infrared*, 2020, 50(4): 493-500.
- [20] 王美荣, 徐国明, 袁宏武. 显著性偏振参量深度稀疏特征学习的目标检测方法[J]. 激光与光电子学进展, 2019, 56(19): 191101.
WANG Meirong, XU Guoming, YUAN Hongwu. Object detection by deep sparse feature learning of salient polarization parameters[J]. *Laser & Optoelectronics Progress*, 2019, 56(19): 191101.
- [21] 李慕锴, 张涛, 崔文楠. 基于YOLOv3的红外行人小目标检测技术研究[J]. 红外技术, 2020, 42(2): 176-181.
LI Mukai, ZHANG Tao, CUI Wennan. Research of infrared small pedestrian target detection based on YOLOv3[J]. *Infrared Technology*, 2020, 42(2): 176-181.