

基于 YOLOv5s 的轻量化森林火灾检测算法研究

皮骏, 刘宇恒, 李久昊

(中国民航大学交通科学与工程学院, 天津 300300)

摘 要: 针对过去无人机搭载嵌入式设备巡检森林火灾精确率低、适应性差和软硬件限制高等问题, 提出一种基于 YOLOv5s 的轻量化森林火灾目标检测算法。通过将 YOLOv5s 的骨干网络替换为轻量化网络 Shufflenetv2, 并以通道重组的思想, 让骨干网络对图片信息的提取效率变得更快, 在保持网络精度的同时保证检测速度; 接着在 Backbone 与 Neck 的连接处加入为轻量化网络设计的 CA 位置注意力模块, 可将图片不同的位置信息聚合到通道中, 使被检对象关注度得以提高; 最后在预测部分使用 CIOU 损失函数, 能够更好的优化矩形框的长宽比和更快加速模型收敛。算法部署在嵌入式系统 Jetson Xavier NX 上的结果显示, 改进后的网络模型大小与对比实验方法相比, 最多减少了 98%, 准确率(Precision)达到 92.6%, 精确率(AP)达到 95.3%, 帧率(FPS)提升到 132 帧每秒, 能满足在白天、黑夜或视野良好等情况下对森林火灾的实时性预防与检测, 并具有良好的准确率和鲁棒性。

关 键 词: 目标检测; YOLOv5s; 轻量化; 位置注意力模块; 森林火灾检测

中图分类号: TP 391

DOI: 10.11996/JGJ.2095-302X.2023010026

文献标识码: A

文章编号: 2095-302X(2023)01-0026-07

Research on lightweight forest fire detection algorithm based on YOLOv5s

PI Jun, LIU Yu-heng, LI Jiu-hao

(Institute of Traffic Engineering, Civil Aviation University of China, Tianjin 300300 China)

Abstract: A new algorithm for light-weight forest fire object detection was proposed based on YOLOv5s to address the low accuracy, poor flexibility, and high software and hardware limitations of the previous UAV-embedded equipment for forest fire inspection. The proposed algorithm replaced the backbone of YOLOv5s with the light-weight network Shufflenetv2, employed the idea of channel recombination to improve the speed of the backbone network in picture information extraction, and maintained both high accuracy and fast detection speed. Then, a coordinate attention (CA) positional attention module specially designed for light-weight network was added to the connection between Backbone and Neck, which could aggregate different position information of pictures into the channel, thus improving the attention of the detected object. Finally, the CIOU loss function was utilized in the prediction part to better optimize the ratio of length to width of the rectangular frame and accelerate the model convergence. The results of the algorithm deployed on Jetson Xavier NX show that compared with the Faster-RCNN, SSD, YOLOv4, and YOLOv5s experimental methods, the improved network model size was reduced by up to 98%, increasing the precision to 92.6%, accuracy rate to 95.3%, and FPS to 132 frames/s. It can effectively achieve the real-time prevention and detection of forest fire in daylight, darkness, or good visibility, exhibiting good accuracy and robustness.

Keywords: object detection; YOLOv5s; light-weight; positional attention module; forest fire detection

收稿日期: 2022-06-18; 定稿日期: 2022-09-02

Received: 18 June, 2022; Finalized: 2 September, 2022

第一作者: 皮骏(1973-), 男, 副教授, 博士。主要研究方向为模式识别、目标检测及人工智能。E-mail: jpi@cauc.edu.cn

First author: PI Jun (1973-), associate professor, Ph.D. His main research interests cover image recognition and artificial intelligence.

E-mail: jpi@cauc.edu.cn

现在许多的目标检测方法,在森林火灾检测方面仍然存在网络复杂度高、泛化能力弱、模型较大不易部署、数据样本量过少影响模型精确率等问题。但近些年随着人工智能的快速发展,在配备了更大存储容量的图形处理器后,使研究者可完成更多的视觉任务,并对模型建立和目标检测更加有利。

在目前的研究中,张彬彬和帕孜来·马合木提^[1]利用改进的 CIOU 和 Focal loss 作为新的损失函数提高了火焰检测精确率,并解决了正负样本数量不一致的问题,但训练速度和检测速度均有所下降;缪伟志等^[2]基于计算机视觉改进 YOLOv4 算法并部署在嵌入式设备上检测森林火灾,实验漏检率较低但检测速度较慢;文献[3-5]基于 YOLOv5 的轻量化检测方法,进行了分类实验,用 Shufflenet 替换主干网络使参数量有不同程度地降低,且改进的模块也有利于设备的部署,但实验结果的精确度不高;刘凯等^[6]利于支持向量机(support vector machine, SVM)算法,实现监测系统可全天候自动检测森林火灾,但过程设计较复杂;将改进的轻量化算法部署在嵌入式计算机视觉设备上,可使算法

在不牺牲精确率的情况下能够检测森林火灾,且保证更小的模型和更快的检测速度^[7-9]。接入视频流还可进行远程检测^[10],并且 YOLO^[11](you only look once)作为一种 one-stage 目标检测网络,从主干网络提取特征后,在特征图上以锚框的方式预测出物体的类别、置信度和位置信息,一次检测即可得到最终结果,检测速度通常是优势^[12-14]。

综上所述,针对森林火灾检测要求深度学习网络具有较高的检测速度和准确率等,且实际情况和实际部署的限制要求又较苛刻,本文在 YOLOv5s 算法上一步步进行改进,最后通过模型测试结果及一系列评价指标评估所训练模型的优越性,并在嵌入式设备上实验。

1 YOLOv5 网络模型基本结构

目前 YOLOv5 共有 4 个不同网络宽度和深度的版本(S,M,L,X),区别在于控制模型宽度和深度的系数不同。若进行嵌入式设备实验,应首选参数量少的 YOLOv5s 部署实验,结构如图 1 所示。

与 YOLOv4 相比, YOLOv5 做了更多的细节调整和优化,但大致结构依然保持不变,由输入端

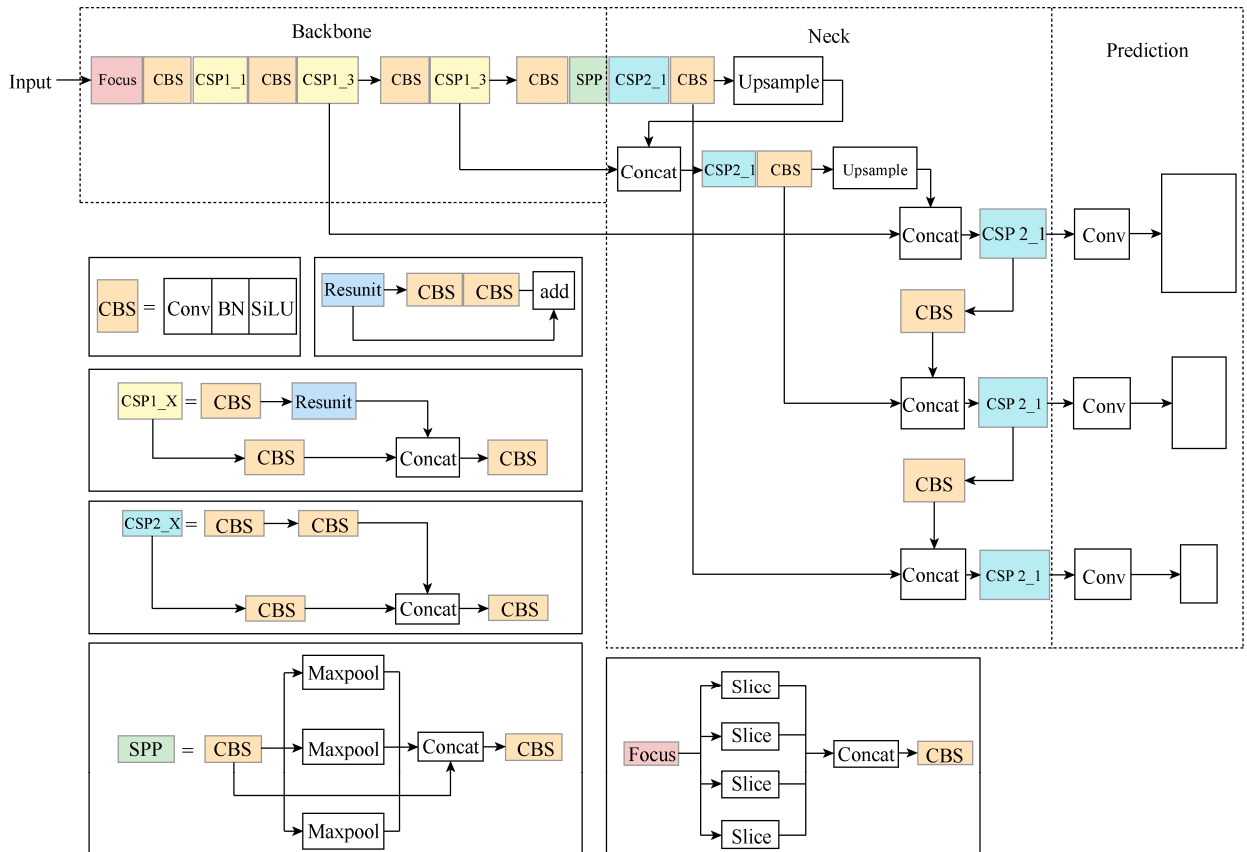


图 1 YOLOv5 基本结构

Fig. 1 The basic structure diagram of YOLOv5

(Input)、主干网络(Backbone)、颈部网络(Neck)和预测输出端(Prediction)构成^[15-16]。输入端主要包含 Mosaic 数据增强、自适应锚框计算和自适应图片缩放;主干网络包括 Focus 输入切分处理、CBL 模块处理、CSP3 结构和空间金字塔池化(spatial pyramid pooling, SPP); Neck 沿用了 YOLOv4 的路径聚合网络,一种从底至上的路径增强,其包含可通过管道聚合的信息组件,将 Backbone 所提取的不同阶段特征进行上采样和池化,使用低层和顶层的横向连接,缩短信息传播路径,运用增强路径的手段使不同层级的特征信息得到丰富。预测部分默认采用 GIOU_Loss 作为损失函数且分别有 3 种不同尺度的输出,对应大、中、小 3 种不同的目标预测^[17]。

2 改进的轻量化 YOLO 网络

2.1 Shufflenetv2 轻量级网络

由于目前的嵌入式设备更倾向于轻量化、低负载和多功能,所以网络模型大小是一个需要持续改进的问题。轻量级网络如 Mobilenet 和 Squeezenet 等^[18]主要关注的是一个间接量——以 FLOPS 为衡量的网络复杂度,即轻量级网络模型的快慢可通过浮点运算量来描述。但在小型计算机的设备中,设

备运行速度不仅要考虑浮点计算量,还要顾及内存访问成本(Memory access cost)和平台特征(Platform characteristics)。所以提出使用 Shufflenetv2 替换 YOLOv5s 的主干网络,通过控制环境来检测设备在网络运行的速度,以得到更优的速度和更低的延迟。

因此本文在改进的 Shufflenet 中,提出了通道改组(Channel shuffle)的基本思想。先将不同维度的特征矩阵进行融合,然后将每一组所提取的特征划分为若干份,再将对应位置的特征重新分配至同一组,并通过通道改组后的特征矩阵进行组卷积(Group convolution),以融合不同组的通道信息,再进行下一步的卷积。

整个流程是首先将输入的图像在经过 3×3 、步距为 2 的卷积核卷积提取特征向量,又通过 BN 和 ReLU 层操作后,为输入特征的通道提供二维的最大池化操作。此操作并未改变通道的维度,但不同通道的操作是一致的,但需对每个通道结果进行拼接。特征拼接之后送入 Shuffle_block 模块进行通道改组,并在 2 个步距不同的 Unit 单元继续特征的卷积和拼接,最后完成信息融合输出结果。网络结构如图 2 所示。

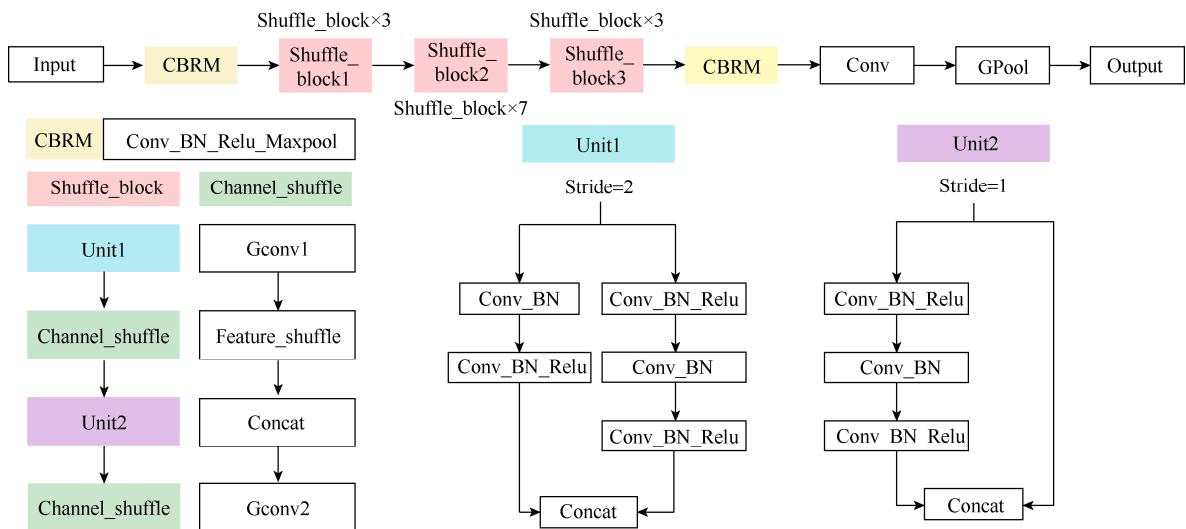


图 2 Shufflenet 网络框架

Fig. 2 Shufflenet framework

Shufflenetv2 网络对图像特征的处理过程相较于 YOLOv5 的主干 CSPDarknet, 少了许多卷积层的卷积过程并对特征处理的先后次序和方式进行了调整。通过改组融合特征信息的方式减少了卷积运算和参数量,使得后续在进行部署实验时能让嵌入式设备对输入图像的处理速度增加。从而维持一

定精度的情况下,减少网络模型大小,提升检测速度,便于嵌入式部署。CSPDarknet 与 Shufflenetv2 参数量对比见表 1。

2.2 CA 改进的位置注意力机制

轻量级网络常关注通道注意力,但通道注意力时常会忽视对生成特征图重要的位置信息^[19]。CA

(coordinate attention)的特点是在通道注意力中融合了位置信息, 特为 Shufflenet 等轻量化的网络而设计。不同于其他机制将网络提取的特征通过 2D 全局池化后得到单个向量, CA 倾向于利于 1D 编码并聚合 2 个方向的特征。在 2 个空间方向 h 和 w 上均进行自适应的二维平均池化, 使得一个方向为后面的网络预测保留了十分精确的位置信息, 另一个方向使卷积可以获得长距离的依赖关系。

表 1 主干网络参数量对比

Table 1 Comparison of parameter quantity				
Model summary	Layers	Parameters	Gradients	GFLOPS
CSPDarknet	272	7 233 211	7 233 211	16.4
Shuffnetv2	224	3 608 302	3 608 302	8.1

将 CA 放在每一个通道改组模块之后、且 Neck 部分拼接之前。特征图通过主干网络的特征提取和不同组信息融合之后, 能使得到的信息变得多元, 有利于 Neck 进一步地提取和后续的预测。基本网络结构如图 3 所示。

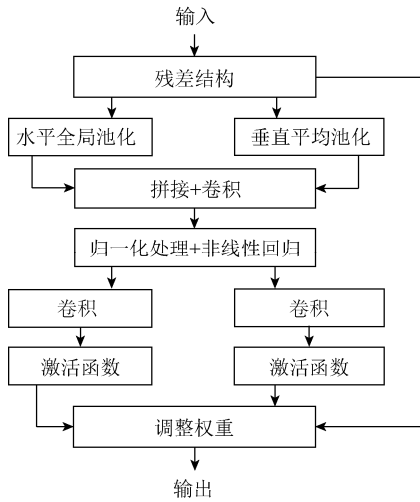


图 3 位置注意力机制流程图

Fig. 3 The flow chart of coordinate attention

2.3 快速空间金字塔池化

快速空间金字塔池化(spatial pyramid pooling-fast, SPPF)是在 SPP^[20]上进行地改进。虽然两者的作用相同, 但前者的效率更高。有实验表明, 两者拥有相同的计算结果, 但前者的计算速度是后者的 2 倍。SPPF 是将输入的特征图依次输入到 3 个 5×5 大小的最大池化层, 再将每一层的结果进行拼接, 并位于主干网络结构的末尾。在主干网络提取的特征向量图送入颈部网络进行操作之前, 需再一次地对特征图进行尺度统一, 使每幅图中的目标位置和大小始终保持相同, 以更好地保证图像特征的准确性, 如图 4 所示。

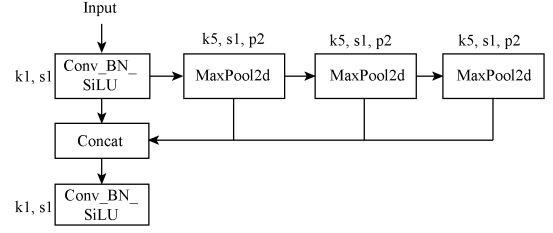


图 4 SPPF 结构

Fig. 4 The structure of SPPF

2.4 损失函数改进

YOLOv5s 是由分类损失(Classification_loss)、定位损失(Localization_loss)、置信度损失(Confidence_loss) 3 部分进行加权相加构成的网络总损失。而边界框的位置预测是目标检测中一个最主要的任务, 可通过对边界框具体的位置数据进行预测框输出需要检测的目标。过去多使用的是平方差方法, 即利用预测框的 (X, Y, W, H) 等数据计算损失。无法通过关注预测框和真实框重叠区域面积的边界框进行衡量计算, 因此提出了 IOU 的损失计算方式及后续的改进方法。 B 和 B_{gt} 分别表示预测框和真实框, IOU 则是表示 2 框交集与并集的比值。但仍存在很多问题, 如 2 框没有交集则损失为定值, 其无法反映 2 框位置关系; 或 2 框属于包含关系, 即一个框在另一框内的位置不能确定, 同样无法反映位置情况, 导致影响损失函数的计算, 即

$$IOU(B, B_{gt}) = \frac{|B \cap B_{gt}|}{|B \cup B_{gt}|} \quad (1)$$

$$L_{IOU} = 1 - IOU(B, B_{gt}) \quad (2)$$

为了解决 IOU 本身存在的问题, GIoU 则应运而生。其在 IOU 的基础上新增了一个最小矩形框, 并要求能同时框住 B 和 B_{gt} , 即

$$GIoU(B, B_{gt}) = IOU(B, B_{gt}) - \left| \frac{C - (B \cup B_{gt})}{C} \right| \quad (3)$$

$$L_{GIoU}(B, B_{gt}) = 1 - GIoU(B, B_{gt}) \\ = 1 - IOU(B, B_{gt}) + \frac{C - (B \cup B_{gt})}{C} \quad (4)$$

其中, C 为最小矩形框的面积。

在计算中发现, B 与 B_{gt} 的距离越远, C 值越大, 导致最后结果趋近于 1, 这解决了 2 框完全分离时 IOU 无法计算的问题。在解决 2 框完全重叠无法计算损失时, 有学者提出了 CIOU_Loss 的计算损失方法。即在 CIOU 的基础上继续改进了预测框与真实框的长宽比, 并新增了 α 和 v 2 个参数用来衡量损失。因此, 将改进的 CIOU 损失替换原 GIoU 损失

计算,使之更好地预测目标和观察损失值

$$CIOU(B, B_{gt}) = -\alpha v - \frac{\rho^2(B, B_{gt})}{C^2} + IOU(B, B_{gt}) \quad (5)$$

$$v = \frac{4}{\pi} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (6)$$

$$\alpha = \frac{v}{1 - IOU(B, B_{gt}) + v} \quad (7)$$

其中, w 和 h 为预测矩形框的宽和高; v 为归一化了预测框和真实框的比例差值; α 为权衡长宽比例

的平衡因子; ρ 为预测框与真实框中心点的距离。

2.5 改进后的森林火灾轻量化检测网络

本文检测模型如图 5 所示。输入森林火灾图像, 首先进入 Shufflenetv2 的主干网络模块, 完成特征提取及位置注意力机制对不同组特征进行信息融合, 在更快的 SPP 层将之前提取得到的特征图进行尺度统一并保证特征的准确性, 然后送入颈部网络的不同特征层再次进行特征的处理和融合, 最后传递给头部网络进行目标检测。

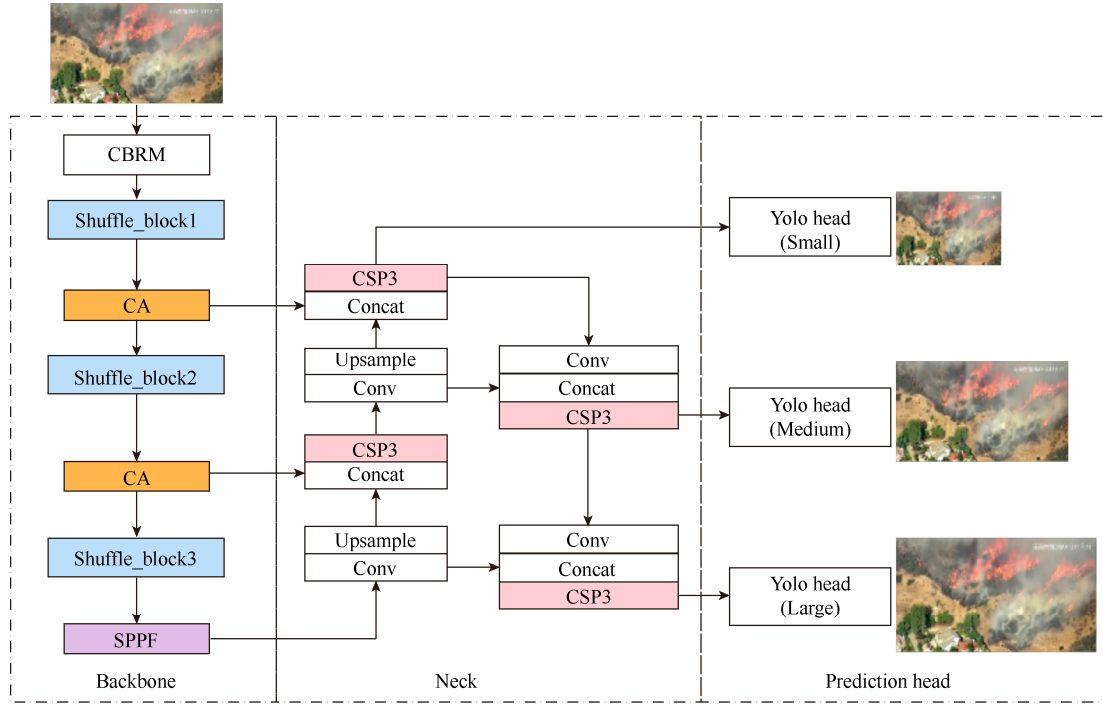


图 5 改进后的森林火灾轻量化检测网络

Fig. 5 Improved forest fire light weight detection network

3 实验结果及分析

3.1 数据增强及划分

本文通过网络爬取航拍森林火灾视频逐帧成图并制作成数据集来训练和评估模型, 所有视频均由无人机航拍而成, 共包括 55 个视频片段。经过抽取和反复挑选, 再由 python 批量进行图片亮度增强、对比度增强、旋转角度和翻转一共获得 4 053 张图像, 将此以 7:2:1 的比例分为训练集、验证集和测试集, 共有 21 122 个标注, 覆盖了国内的各种规模森林火灾。

3.2 实验环境

采用 PyTorch 深度学习框架, python 3.8 编程语言, Windows 10 系统, NVIDIA Geforce RTX 3080 10 G 显存。预训练权重使用的是官方的 YOLOv5s, 训练世代数设置为 300 epochs, Batch-size 设置的

为适应显存的 64。初始学习率设置为 0.01, SGD 动量为 0.937。

3.3 实验结果及分析

根据对 YOLO 模型持续地改进, 实验结果由准确率(Precision), 召回率(Recall), 精度(accuracy rate, AP)和每秒传输的帧数(frames per second, FPS)等指标进行评价, 因为本实验是单目标实验, 所以平均精确率由 AP 表示, 即

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

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

$$AP = \int_0^1 Precision(t) dt \quad (10)$$

其中, TP 为正样本被检测正确的数量; FP 为负样本被检测为正样本的数量; FN 为背景被错误检测为正样本的数量; FPS 为每秒算法可以处理图片的

数量,不同的嵌入式设备软硬件条件也会影响传输帧数的数值。消融实验见表 2。

从表 2 中可以看出,网络在陆续修改主干网络、损失函数和添加空间金字塔池化、位置注意力机制并进行了不同组合的实验对比后,精确率相比其余消融方法提高 5.1%,召回率虽有波动,但相比其余消融方法也有 4.9%的提高,AP 相比其余消融方法提升了 4.9%,以证明检测网络性能的优越性。

为了验证在实际检测过程中的效果,将已实验好的算法移植进嵌入式设备,实现在 NVIDIA 嵌入式开发板系统——Jetson Xavier NX 上进行测试和

实时的目标检测(图 6)。由于 NX 比 Nano 有更好的性能和配置,能够适应更加复杂的网络和严格的环境,其具体配置和内容见表 3。

4 部署实验

表 4 为各算法对比,可见本文模型大小较 YOLOv4 减少了 98%。且在准确度、召回率、精度和检测用时上均有一定的提升,FPS 达到了 132 帧每秒,能够使无人机搭载嵌入式设备巡检森林火灾时做到更加实时地反馈和流畅地检测,是一种更加轻量化、低延时和高精度的森林火灾检测算法,实验效果如图 7 所示。

表 2 消融实验

Table 2 Ablation experiments

Number	Shuffle (Backbone)	CIOU_Loss	SPPF	CA	Precision	Recall	AP
1	√	-	-	-	0.890	0.924	0.935
2	√	√	-	-	0.908	0.889	0.936
3	√	-	√	-	0.907	0.880	0.929
4	√	-	-	√	0.920	0.911	0.931
5	√	√	√	-	0.927	0.912	0.945
6	-	√	√	-	0.912	0.921	0.911
7	-	√	-	√	0.932	0.919	0.942
8	-	√	√	√	0.936	0.932	0.950
9	-	-	√	√	0.922	0.928	0.949
Ours	√	√	√	√	0.941	0.929	0.960

注: 加粗数据为最优值



图 6 嵌入式设备 NVIDIA Jetson Xavier NX ((a)整体外观; (b)内部结构; (c)俯视图)

Fig. 6 The embedded device of NVIDIA Jetson Xavier NX ((a) Overall appearance; (b) Interior structure; (c) Planform)

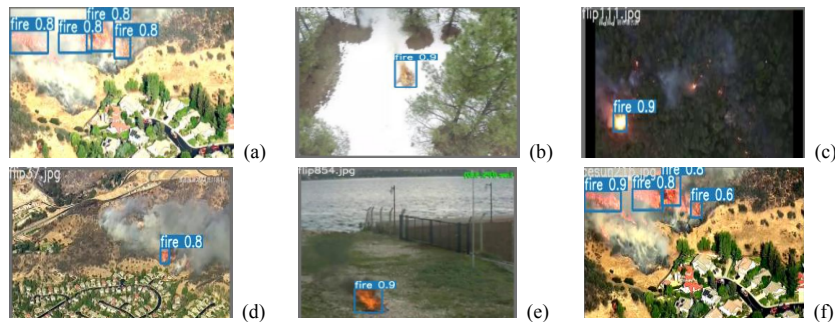


图 7 实验效果图((a)高亮度; (b)雪地环境; (c)黑夜环境; (d)烟雾环境; (e)弱光环境; (f)原版)

Fig. 7 Experimental renderings ((a) High brightness; (b) Snowy environment; (c) Dark environment; (d) Smoke environment; (e) Low light environment; (f) Original)

表 3 NVIDIA Jetson Xavier NX 具体配置

Table 3 The specific configuration of NVIDIA Jetson Xavier NX

配置	系统	开发环境	CPU	GPU	RAM	CSI 摄像头
细节	Ubuntu18.04	Python=3.7, C++, JAVA, CMake, OpenCV-4.0	6 核 NVIDIA CarmelARM 6MBL2+4MBL3	384 核 NVIDIA Volta	8 GB 128 位 LPDDR4x5 1.2 GB/s	12 通道 MIPI CSI-2DHY 1.2

表 4 对比实验
Table 4 Contrast experiment

算法	Backbone	模型大小(MB)	Precision	Recall	AP	FPS	检测用时(ms)
YOLOv4	CSPDarknet53	241.2	0.859	0.758	0.852	35	11.2
YOLOv4-tiny	CSPDarknet53-Tiny	25.3	0.869	0.792	0.843	42	4.5
YOLOv5s	CSPDarknet	11.5	0.923	0.909	0.944	79	9.8
YOLOv5s-P6	CSPDarknet	15.4	0.891	0.881	0.932	62	10.1
YOLOv5s-P7	CSPDarknet	15.5	0.906	0.892	0.939	65	11.3
YOLOv5s-Transformer	Transformer	20.1	0.915	0.911	0.950	70	9.2
SSD	VGGnet	180.3	0.613	0.521	0.631	22	60.7
Faster-RCNN	Resnet	212.3	0.732	0.612	0.756	28	53.1
Ours	Shufflenetv2	5.1	0.926	0.920	0.953	132	3.6

5 结束语

本文通过对森林火灾检测缺点提出的一种基于改进 YOLOv5 的轻量级目标检测算法, 自制数据集训练并搭建目标检测网络, 使之能够部署到嵌入式设备上, 实现轻量化高精度检测森林火灾并起到一定的预防作用。通过对比实验证明, 搭建的网络模型准确率达 92.6%, 召回率达到 92.0%, 精确率达 95.3%。在无人机需要搭载更轻量型的嵌入式设备检测森林火灾的今天, 在设备软硬件条件的限制下, 目标检测模型的轻量化和高精度是非常必要的。通过本文的实验为检测提供了方法和可靠的数据及模型。

本文的数据集目前只包含白天和黑夜 2 种情况。因此在面对不同天气及环境色产生影响的情况是否可行, 高分辨率图像对轻量级网络精度的影响等问题还有待继续研究。

参考文献 (References)

- [1] 张彬彬, 帕孜来·马合木提. 基于 YOLOv3 改进的火焰目标检测算法[J]. 激光与光电子学进展, 2021, 58(24): 289-296.
- [2] ZHANG B B, PAZILAI M H M T. Improved flame target detection algorithm based on YOLOv3[J]. Laser & Optoelectronics Progress, 2021, 58(24): 289-296 (in Chinese).
- [3] 缪伟志, 陆兆纳, 王俊龙, 等. 基于视觉的火灾检测研究[J]. 森林工程, 2022, 38(1): 86-92, 100.
- [4] MIAO W Z, LU Z N, WANG J L, et al. Fire detection research based on vision[J]. Forest Engineering, 2022, 38(1): 86-92, 100 (in Chinese).
- [5] DONG X D, YAN S, DUAN C Q. A lightweight vehicles detection network model based on YOLOv5[J]. Engineering Applications of Artificial Intelligence, 2022, 113: 104914.
- [6] CHEN Z C, YANG J, CHEN L F, et al. Garbage classification system based on improved ShuffleNet v2[J]. Resources, Conservation and Recycling, 2022, 178: 106090.
- [7] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [8] 刘凯, 魏艳秀, 许京港, 等. 基于计算机视觉的森林火灾识别算法设计[J]. 森林工程, 2018, 34(4): 89-95.
- [9] LIU K, WEI Y X, XU J G, et al. Design of forest fire identification algorithm based on computer vision full text

- replacement[J]. Forest Engineering, 2018, 34(4): 89-95 (in Chinese).
- [7] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [8] PERRONNIN F, SÁNCHEZ J, MENSINK T. Improving the fisher kernel for large-scale image classification[M]//Computer Vision - ECCV 2010. Cham: Springer International Publishin, 2010: 143-156.
- [9] SZEGEDY C, LIU W, JIA Y Q, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2015: 1-9.
- [10] 仲亨玉, 刘文萍, 刘鹏举. 基于分数阶微分视频融合的森林烟火检测算法[J]. 北京林业大学学报, 2017, 39(3): 24-31.
- [11] ZHONG T Y, LIU W P, LIU P J. A forest fire smoke detection algorithm based on fractional-calculus video fusion[J]. Journal of Beijing Forestry University, 2017, 39(3): 24-31 (in Chinese).
- [12] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[C]//2017 IEEE International Conference on Computer Vision. New York: IEEE Press, 2017: 2999-3007.
- [13] AVAZOV K, MUKHIDDINOV M, MAKHMUDOV F, et al. Fire detection method in smart city environments using a deep-learning-based approach[J]. Electronics, 2021, 11(1): 73.
- [14] LIU S T, ZHANG N N, YU G. Lightweight security wear detection method based on YOLOv5[EB/OL]. [2022-02-20]. <https://dl.acm.org/doi/10.1155/2022/1319029>.
- [15] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [16] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [17] 张苗, 李璞, 杨漪, 等. 基于目标检测卷积神经网络的图像型火灾探测算法[J]. 消防科学与技术, 2022, 41(6): 807-811.
- [18] ZHANG M, LI P, YANG Y, et al. Image fire detection algorithms based on object detection convolutional neural networks[J]. Fire Science and Technology, 2022, 41(6): 807-811 (in Chinese).
- [19] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[J]. Proceedings of Computer Vision and Pattern Recognition, 2005, 1: 886-893.
- [20] PERRONNIN F, SÁNCHEZ J, MENSINK T. Improving the fisher kernel for large-scale image classification[M]//Computer Vision - ECCV 2010. Cham: Springer International Publishin, 2010: 143-156.
- [21] HOU Q B, ZHOU D Q, FENG J S. Coordinate attention for efficient mobile network design[EB/OL]. [2022-01-12]. <https://arxiv.org/abs/2103.02907>.
- [22] HE K M, ZHANG X Y, REN S Q, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.