

基于改进YOLOv5s的森林烟火检测算法

冯艳玲, 韩毓莹, 余智美, 朱珉慧, 朱雨荷, 孙庆华

云南大学信息学院, 云南 昆明

收稿日期: 2024年3月21日; 录用日期: 2024年4月21日; 发布日期: 2024年4月29日

摘要

为了解决传统火焰烟雾检测算法在光照条件恶劣和山林云雾影响的条件下存在漏检误检、准确性下降的缺陷, 提出了一种基于YOLOv5s目标检测框架的森林烟火检测算法。首先, 针对火焰烟雾目标特征复杂的问题, 在C3模块中融合Res2Net, 增强了网络在不同尺度下的特征表示能力。然后, 在主干检测网络加入SE注意力模块, 达到抑制干扰信息, 提升模型表征能力的效果。最后, 通过集成GIOU优化损失函数, 进一步提高了检测精度。改进后的算法相比于传统算法, mAP50值提高了1.8%, P值提高了0.9%, R值提高了0.6%。

关键词

目标检测, 火焰烟雾检测, Res2Net, SE注意力机制

Forest Smoke and Fire Detection Algorithm Based on Improved YOLOv5s

Yanling Feng, Yuying Han, Zhimei Yu, Minhui Zhu, Yuhe Zhu, Qinghua Sun

School of Information Science & Engineering, Yunnan University, Kunming Yunnan

Received: Mar. 21st, 2024; accepted: Apr. 21st, 2024; published: Apr. 29th, 2024

Abstract

In order to solve the defects of traditional flame smoke detection algorithms in terms of leakage and misdetection and accuracy degradation under the conditions of poor lighting conditions and the influence of clouds and fog in mountain forests, a forest smoke and fire detection algorithm based on the YOLOv5 target detection framework is proposed. Firstly, to address the problem of complex features of flame smoke targets, Res2Net is fused in the C3 module, which enhances the feature representation ability of the network at different scales. Then, the SE attention module is added to the backbone detection network to achieve the effect of suppressing the interference in-

文章引用: 冯艳玲, 韩毓莹, 余智美, 朱珉慧, 朱雨荷, 孙庆华. 基于改进 YOLOv5s 的森林烟火检测算法[J]. 计算机科学与应用, 2024, 14(4): 290-297. DOI: 10.12677/csa.2024.144098

formation and enhancing the model representation ability. Finally, the detection accuracy is further improved by integrating GIOU to optimize the loss function. The improved algorithm continues to improve the mAP50 value by 1.8%, the P value by 0.9%, and the R value by 0.6% compared with the traditional algorithm.

Keywords

Target Detection, Flame Smoke Detection, Res2Net, SE Attention Mechanism

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

森林火灾是一种破坏性极强的森林灾害，具有突发性强、破坏性大、难以控制等特点，是全球发生最频繁、处置最困难、危害最严重的自然灾害之一[1]。自 2024 年春节以来，贵州全省共发生森林火情 221 起，是上年同期 3 倍，已核实为森林火灾计 11 起，造成 2 人员伤亡，森林火情大幅增加。2024 年 3 月 16 日 18 时 10 分许，云南省临沧市临翔区圈内乡斗阁村发生一起森林火灾，过火面积约 5.33 公顷，3 名扑火人员牺牲，1 人受伤。

对以上及相关森林火灾事故分析，了解到火焰和烟雾是森林火灾发生前的重要特征。随着深度学习的发展，可以利用无人机搭载深度学习平台，代替人工巡检森林，如果在火焰烟雾检测方面提高准确率，并降低检测的误报率，就可以在森林火灾发生的初始阶段及时报警。

现有的基于目标检测算法的森林火灾特征提取方法大都是针对可见光，通过多种特征增强方法和模型训练策略，使模型学习到烟雾和火焰细粒度特征，进而提升模型性能。Zheng Xin 等[2]。评估了几个主流的目标检测算法(Faster R-CNN [3], Efficient Det [4], YOLOv3 [5], SSD [6])在森林火灾数据集的性能，研究了该类算法检测森林火灾的可行性。谢书翰等[7]提出基于改进 YOLOv4 的火灾检测模型，在检测头为增强模型的烟雾特征提取能力，融合了通道注意力机制，同时对烟雾特征对损失函数和先验框进行优化，但该模型且易受山林云雾影响，产生误检。

此外，由于森林防火的巡检面积大、巡检困难，采用固定摄像头和人工巡检效率极低，因此 Treneska 等人[8]开发了一种基于迁移学习的卷积神经网络，利用无人机捕捉的视频流进行早期火灾特征提取，取得了良好的准确度和实时性。但是，该方法仅采用较好的光照条件下的可见光图像对模型进行训练和测试，在夜晚和光照条件恶劣的实际情况下，网络性能较差。Jiao 等人[9]提出一种基于小尺度卷积神经网络和无人机航拍进行森林火灾特征提取和检测的算法，利用无人机搭载红外或者可见光探头进行航拍监测，通过 YOLOv3 目标检测算法对图像进行特征提取，并通过无人机通信将图像传输到地面站。该方法在大面积森林火灾检测和报警应用中取得了优异的性能，但是由于训练集缺乏大量可用的初始森林火灾数据，该网络的小规模火灾检测性能差，准确度低，无法满足早期森林火灾的监测和预警需求。

综上所述，现有基于目标检测算法的森林火灾特征提取方法的研究在光照条件恶劣和山林云雾影响的条件下存在漏检误检、准确性下降的劣势，加之缺乏大量可用的初始森林火灾图像。为解决此问题，提出改进基于改进 YOLOv5s 的森林烟火检测算法，通过融合 Res2Net 并嵌入 SE 注意力机制优化 YOLOv5s 在森林火灾特征提取上的准确性。实验采用遥感森林火灾探测的多面合成数据集 (<https://github.com/Philharmy-Wang/M4SFWD>)作为训练集，提高模型泛化能力。改进后的算法相比于传统

算法，在保证实时性的基础上提升了准确率，使森林火焰烟雾的实时检测更加准确、有效。

2. 模型改进

相比于传统 yolov5s 网络结构，提出了一个新的模块——C3_Res2Block。C3_Res2Block 由主干检测网络的 C3 模块融合 Res2Net 得到的，能够在不增加网络深度和参数量的情况下，更好地捕捉图像中的细节和感兴趣的目标。为了进一步提升模型的重要特征表示能力，在主干检测网络中嵌入了 SE 注意力机制抑制无关特征的响应，并通过优化损失函数进一步提高了算法的准确性。经过改进的 yolov5s 算法模型解决了检测精度低、目标漏检的问题，具体网络如图 1 所示。

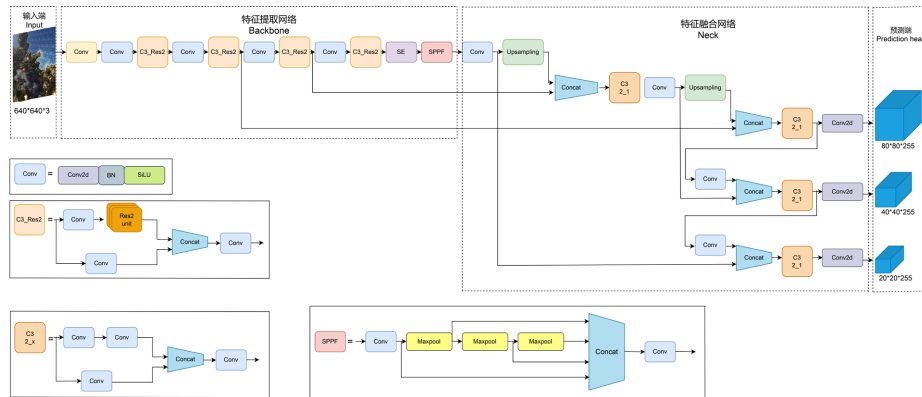


Figure 1. Structure of the improved yolov5 network model
图 1. 改进后的 yolov5 网络模型结构图

2.1. C3_Res2Block 模块的构成

原 YOLOv5 网络中，基础核心模块 C3 使用了以 ResNet 为核心算法的残差网络模块，可以使网络的学习能力大大提高。但是传统的 ResNet 块只使用单一的残差连接来传递信息，这意味着网络可能无法有效地捕获不同层次的特征。然而在森林火灾发生初期，火焰目标比较小而烟雾一般范围更大，实际情况还因不同地形和天气又不同，目标的尺度大小不规律，原 YOLOv5 网络在森林火灾特征提取场景下存在精度不高。

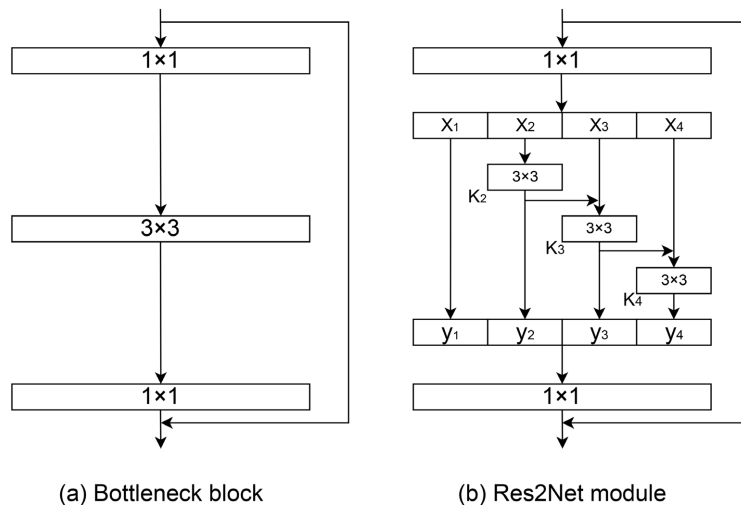


Figure 2. Res2Net network structure
图 2. Res2Net 网络结构图

2020年 Gao 等人[10]提出的 Res2Net, 在残差网络的基础上构建分层的残差连接, 在更细粒度上表示多尺度特征信息, 并增加了网络层的感受野。如图 2 是 Res2Net 的网络结构图。输入特征首先经过初始卷积层, 用于低级别的特征表示, 接下来, 输入特征被送入多个 Res2Net 块, 每个块都包含了多个分支, 每个分支处理不同分辨率的特征, 将每个分支的输出级联在一起形成块的最终输出。

根据 Res2Net 网络设计思想, 将残差网络 Res2Net 与 C3 模块融合, 创建了 C3_Res2Block 网络结构。C3_Res2Block 网络结构设计可以使得神经网络融合不同尺度下的特征信息, 更好地捕捉不同尺度的特征信息, 提高网络性能。

2.2. 嵌入 SE [11]注意力机制

注意力机制模仿了人类的视觉系统, 融合局部视觉结构, 获取场景特征, 并将注意力集中于特征较为显著的区域。在处理大量的输入信息时, 注意力机制会选择一些关键的输入信息进行处理, 忽略与目标无关的噪声数据, 从而提高神经网络模型的效果。

SENet 作为轻量级的注意力机制, 具有即插即用、结构简单的特点[12], 仅会小幅度增加网络的运算量。SENet 的网络结构如图 3 所示, 其中, SENet 包括两个重要的操作: Squeeze 和 Excitation。Squeeze 阶段涉及通过全局平均池化操作将输入特征图压缩成单个向量, 然后将这个向量映射到一个较小的向量。

Squeeze 阶段涉及通过全局平均池化操作将输入特征图压缩成单个向量, 然后将这个向量映射到一个较小的向量。具体是通过每个通道的信息($W \times H$)转化成 1 个像素(1×1), 于是每一个数值都具有该通道的全局感受野。Squeeze 操作的计算公式:

$$Z_C = F_{sq}(u_c) = \frac{1}{H \times W} \sum_{i=1}^W \sum_{j=1}^W u_c(i, j) \quad (1)$$

其中, H 、 W 、 C 是图片的长、宽和维度, u_c 是给定的输入 $H \times W \times C$, Z_C 是挤压后得到 $1 \times 1 \times C$ 的输出。

Excitation 阶段则是利用 Sigmoid 函数将上述小向量中的每个元素压缩到 0 到 1 之间, 并与原始输入特征图相乘, 从而为每个特征通道配置一个注意力权重。Excitation 阶段的计算公式:

$$s = F_{ex}(z, W) = \sigma(W_2 \delta(W_1 z)) \quad (2)$$

其中 W_1 为降维矩阵, W_2 为升维矩阵, δ 、 σ 分别为 ReLU 激活函数、Sigmoid 激活函数。

因此, SeNet 相对于其他注意力机制, 参数数量增加相对较少, 使神经网络负担较轻, 在训练时, 对神经网络特征图加权, 使模型注意力更加有效聚焦在目标个体上, 抑制周围环境对目标个体检测的干扰[13]。

改进后的 YOLOv5s 在主干检测网络的第 10 层加入 SE 注意力模块, 使网络模型在识别检测时进行通道级别的识别和权重分配, 达到抑制干扰信息, 提升模型表征能力的效果[14]。从而更准确的定位火焰烟雾目标的位置信息和特征。

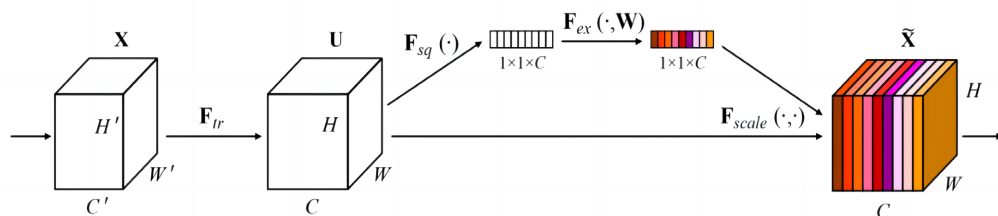


Figure 3. SE attention mechanism
图 3. SE 注意力机制

2.3. 损失函数优化

损失函数的作用是计算模型预测值和实际值之间的差距，是模型训练优化时需要重点关注。YOLOv5的损失函数由定位损失函数、目标置信度损失函数、分类损失函数三部分组成，总体的损失为以上三者的加权相加，通过改变权值可以调整对三者损失的关注度。

原 YOLOv5 使用 $CIOU$ 作为损失函数， $CIOU$ 通过引入边界框之间的中心距离以及宽度和高度之间的比率来考虑形状偏移，使得 $CIOU$ 在一定程度上能够更准确地反映物体之间的相似性。当检测宽高比有一定规律的目标时，能获取目标真实框的宽高比，从而可提高目标检测的 AP 值。

$CIOU$ 计算公式：

$$L_{CIOU} = 1 - IOU(A, B) + \frac{\rho^2(A_{ctr} + B_{ctr})}{c^2} + \alpha v \quad (3)$$

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

$$\alpha = \frac{v}{(1 - IOU) + v} \quad (5)$$

其中， A 是预测框， B 是真实框， A_{ctr} 和 B_{ctr} 代表 A 和 B 的中心点， ρ 为预测框和真实框中心点的欧氏距离， v 可度量预测框和真实框的长款比相似性， ω^B 、 h^B 分别为真实框的宽度、高度， ω 、 h 分别为预测框的宽度与高度。

但是在森林火灾特征提取场景下，火焰烟雾特征比较复杂，且目标的宽高比例不固定。在此场景下，网络只需要关注边界框之间的一般重叠度量，而不需要考虑中心点距离等额外因素，因此使用 $GIOU$ 作为损失函数。

$GIOU$ 计算公式：

$$L_{GIOU} = 1 - IOU + \frac{A \cap B}{C} \quad (6)$$

其中， C 是最小外接矩形， $A \cap B$ 是预测框和真实框的交集。

3. 实验结果及分析

3.1. 硬件条件及参数设置

本实验的环境配置见表 1，对比实验的硬件配置与该环境的配置相同。

Table 1. Operating environment

表 1. 运行环境

CPU	RAM	GPU	操作系统	Cuda	CudNN	Pytorch	python
I3-12100F	16.0G	GTXTITANX12G	Windows11	10.2	8.5.0	1.9.1	3.9.18

数据集来自于 GitHub 平台[15]中 yolov8 版本注释格式的数据，共有 3902 张图片，将数据集按照训练集：测试集：验证集 = 7:1:2 的比例进行划分。本实验 YOLOv5 经过大量实验选取最优参数，对部分参数进行了调整。实验批量大小设置为 32，迭代次数为 200 次。

3.2. 评价指标

与传统 YOLOv5s 进行比较，采用平均精度均值(mAP)、准确率(P)、召回率(R)来评价训练结果。

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}} \quad (7)$$

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (8)$$

$$AP = \int_0^1 P(R) dR \quad (9)$$

$$mAP = \frac{1}{c} \sum_{i=1}^c AP_i \quad (10)$$

其中, N_{TP} 为模型正确检测目标个数; N_{FP} 为模型错误检测目标个数; N_{FN} 是漏检的目标个数。 P 是目标检测的准确率, R 为目标检测的召回率, 可评估模型是否有漏检。 AP 为 P-R 曲线面积, mAP 则表示多个类别的 AP 平均值, 对检测目标的平均准确度进行综合评价。

3.3. 对比实验结果和消融实验结果

3.3.1. 嵌入 SE 注意力机制

为了验证使用 SE 注意力机制的有效性, 采用 CBAM [16]、CA [17]、EMA [18]、SE 四种注意力机制, 调整不同注意力机制嵌入位置和数量, 获得最佳性能进行对比。表 2 是不同注意力机制的客观指标对比, 嵌入注意力机制后各项指标均有所提升, 其中嵌入 SE 注意力机制在火焰烟雾数据集下性能提升最明显, P 值提高了 0.6%, R 值提高了 0.9%, $mAP50$ 提高了 0.9%, 检测平均精度最佳。

Table 2. Comparative experiments with different attention mechanisms

表 2. 不同注意力机制的对比实验

Models	P	R	$mAP50$	$mAP50-95$
Yolov5s	84.8	82.5	86.4	47.3
+CBAM	83.8	82.6	86.7	47.4
+CA	84.6	81.6	86.8	47.2
+EMA	85.1	82.4	87.3	47.1
+SE (ours)	85.3	83.4	87.3	47.5

3.3.2. 损失函数优化

Table 3. Effect of different loss functions on model performance (after adding C3_Res2Block)

表 3. 不同损失函数对模型性能的影响(加入 C3_Res2Block 后)

Models	P	R	$mAP50$	$mAP50-95$
Yolov5s+C3_Res2Block (baseline)	85.7	82.9	87.6	47.7
+DIOU	84.8	82.7	86.9	47.4
+EIOU	85.5	82.9	87.3	47.5
+SIOU	85.0	82.2	86.8	47.5
+WIOU	85.0	82.2	86.9	47.3
+GIOU (ours)	86.2	82.9	87.9	47.7

为了寻找最佳准确率的损失函数, 在融合 Res2Net 的基础上, 对的 $CIOU$ (传统 YOLOv5 使用的损失函数)、 $DIOU$ 、 $GIOU$ 、 $EIOU$ 、 $SIOU$ 、 $WIOU$ 进行对比实验。表 3 是不同损失函数的客观指标对比, 原

始网络结构的 P 值为 85.7%， R 值为 82.9%， $mAP50$ 为 87.6%， $mAP50-95$ 为 47.7。使用 $GIOU$ 作为损失函数后， P 值和 $mAP50$ 的提升最为明显，相比于传统算法， P 值提高了 0.5%， $mAP50$ 提高了 0.3%。

3.3.3. 消融实验

为了验证改进后算法的有效性以及检测性能提升程度，设计了 5 组消融对比实验，每次实验训练次数都设置为 200。

Table 4. Ablation experiments

表 4. 消融实验

experiments	C3_Res2Block	$GIOU$	SE	P	R	$mAP50$	$mAP50-95$
0				84.8	82.5	86.4	47.3
1	✓			85.7	82.9	87.6	47.7
2		✓		84.9	83.4	87.1	47.3
3			✓	85.3	83.4	87.3	47.5
4	✓	✓		86.2	82.9	87.9	47.7
5	✓	✓	✓	85.7	83.1	88.2	48.6

表 4 为消融实验的客观指标对比，融合 Res2Net 模块后， $mAP50$ 值提高了 1.2%， P 和 R 值均有所提高，修改损失函数为 $GIOU$ 后， $mAP50$ 值继续提高 0.3%， P 值有所提高， R 值保持不变。添加 SE 注意力后， $mAP50$ 值继续提高了 0.3%， P 值下降了 0.5%， R 值提高了 0.2%。改进后的 yolov5 相比于传统 yolov5 在检测森林火焰烟雾的场景下，准确性由所提升。

3.3.4. 检测效果与分析

选取 YOLOv5s 与改进后的算法检测进行对比，检测效果如图 4 所示。原始的 YOLOv5 在光照条件恶劣和存在山林云雾情况下存在漏检和准确率低的问题，比如(f)、(g)中，漏检了烟雾目标，(e)和(h)中漏检了火焰目标而且精度比较低，而经过改进的模型则能够准确地检测出图像中所有的烟雾目标，与传统算法相比，准确率也有所提升，成功地解决了漏检的问题。

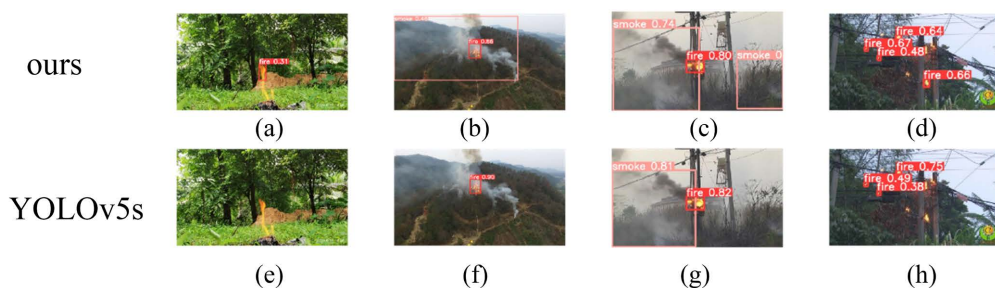


Figure 4. Comparison of test results

图 4. 检测结果对比

4. 结论

为有效解决传统火焰烟雾检测算法在光照条件恶劣和山林云雾影响的条件下存在漏检误检、准确性下降的缺陷，提出了基于改进 YOLOv5s 的森林烟火检测算法。所提算法的使用的 C3_Res2Block 增强了网络在不同尺度下的特征表示能力。嵌入注意力机制后达到了抑制干扰信息，提升模型表征能力的效果。

使用公开数据集对所提算法进行定性分析和定量比较。实验结果表明, 所提出的算法在火焰烟雾检测任务中表现出优越的性能, 提高了平均精度并解决了漏检的问题, 可为森林火灾的预防和控制提供更加可靠和高效的技术支持。

参考文献

- [1] 许嘉庆. 基于智慧林业的森林防火方案研究[J]. 林业勘查设计, 2021, 50(6): 39-41.
- [2] Zheng, X., *et al.* (2022) Real-Time Detection of Full-Scale Forest Fire Smoke Based on Deep Convolution Neural Network. *Remote Sensing*, **14**, Article 536. <https://doi.org/10.3390/rs14030536>
- [3] Ren, S.Q., *et al.* (2015) Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. In: Cortes, C., Lawrence, N., Lee, D., Sugiyama, M. and Garnett, R., Eds., *Advances in Neural Information Processing Systems* 28, 7-12 December 2015, Montreal.
- [4] Tan, M.X., Pang, R.M. and Le, Q.V. (2020) Efficientdet: Scalable and Efficient Object Detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, 13-19 June 2020, 10778-10787. <https://doi.org/10.1109/CVPR42600.2020.01079>
- [5] Redmon, J. and Ali, F. (2018) Yolov3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.
- [6] Liu, W., *et al.* (2016) SSD: Single Shot Multibox Detector. In: Leibe, B., Matas, J., Sebe, N. and Welling, M., Eds., *European Conference on Computer Vision*, Springer, Cham, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2
- [7] 谢书翰, 张文柱, 程鹏, 等. 嵌入通道注意力的 YOLOv4 火灾烟雾检测模型[J]. 液晶与显示, 2021, 36(10): 1445-1453.
- [8] Treneska, S. and Risteska Stojkoska, B. (2021) Wildfire Detection from UAV Collected Images Using Transfer Learning. *18th International Conference on Informatics and Information Technologies*, Skopje, North Macedonia.
- [9] Jiao, Z., *et al.* (2019) A Deep Learning Based Forest Fire Detection Approach Using UAV and YOLOv3. *2019 1st International Conference on Industrial Artificial Intelligence (IAI)*, Shenyang, 23-27 July 2019, 1-5. <https://doi.org/10.1109/ICIAI.2019.8850815>
- [10] Gao, S.H., Cheng, M.-M., Zhao, K., *et al.* (2019) Res2Net: A New Multi-Scale Backbone Architecture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **43**, 652-662.
- [11] Hu, J., Shen, L. and Sun, G. (2017) Squeeze-and-Excitation Networks. arXiv preprint, arXiv:1709.01507.
- [12] 武慧, 杨玉竹, 卜显峰, 等. 基于改进 YOLOv5 的城市火灾检测算法研究[J/OL]. 无线电工程: 1-10 <http://kns.cnki.net/kcms/detail/13.1097.TN.20240111.1703.006.html>, 2024-03-20.
- [13] 肖蕾, 蓝宗苗. 基于注意力机制的污水微型动物识别方法[J]. 激光与光电子学进展, 2023, 60(2): 249-256.
- [14] 陈富荣, 刘俊, 潘德伟. 基于改进 YOLOv5s 的造船起重机壁面缺陷检测算法[J]. 上海电机学院学报, 2023, 26(6): 344-349.
- [15] Philharmy-Wang (2024) M4SFWD. GitHub. <https://github.com/Philharmy-Wang/M4SFWD/>
- [16] Woo, S., Park, J., Lee, J.-Y. and Kweon, I.S. (2018). CBAM: Convolutional Block Attention Module. *Proceedings of the European Conference on Computer Vision (ECCV)*, Tel Aviv, 23-27 October 2022, 3-19. https://doi.org/10.1007/978-3-030-01234-2_1
- [17] Liu, B., Jiang, Z., He, R. and Sun, Z. (2021) Coordinate Attention for Efficient Mobile Network Design. arXiv preprint arXiv:2106.03186.
- [18] Zhang, Q., Wu, Y., Xie, L. and Lin, L. (2021) Efficient Multi-Scale Attention Module with Cross-Spatial Learning. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, 18-24 June 2022, 8977-8986.