Scientific Reports

https://doi.org/10.1038/s41598-026-41994-2

Article in Press

Research on target detection algorithm for forest fire images based on multi-scale feature extraction

Received: 20 November 2025

Accepted: 24 February 2026

Published online: 09 March 2026

Cite this article as: Wu W., Zhou X., Qin J. *et al.* Research on target detection algorithm for forest fire images based on multi-scale feature extraction. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-41994-2

Weilin Wu, Xinpeng Zhou, Jincheng Qin, Zhanyue Fu & Kai Xing

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

ARTICLE IN PRESS

Research on Target Detection Algorithm for Forest Fire Images Based on Multi-scale Feature Extraction

Weilin Wu [1,2,3], Xinpeng Zhou [1,2,3] , Jincheng Qin[1,2,3,*], Zhanyue Fu [1,2,3] and Kai Xing [1,2,3]

[1] The Center for Applied Mathematics of Guangxi,School of Physics and Electronic Information, Guangxi Minzu University, Nanning 530006,China.
[2] Guangxi Key Laboratory of ZHIYU Humanoid Robots, Nanning 530006,China.
[3] Engineering Research Center of Multi-modal Information Intelligent Sensing, Processing and Application，Guangxi Minzu University, Nanning 530006,China.
*corresponding author:Jincheng Qin email: jc_qin@gxmzu.edu.cn

Abstract

To address the challenges of small target flames and target scale variation in forest fire images, a target detection method for forest fire images based on multi-scale feature extraction was studied, with YOLOv9c as the baseline model. Initially, a lightweight feature extraction module named EGI (ECA_Ghost_InceptionV2) was proposed to serve as the backbone feature extraction network, which improved the model's feature extraction capability and operational efficiency. Second, a P2 small target detection head was introduced; meanwhile, a small target feature fusion module was added to the Neck layer, and the CARAFE upsampling operator was incorporated, enhancing the model's ability to extract underlying feature information. Finally, to solve the problems of misalignment and scale inconsistency in the traditional IoU loss function, Inner_DIoU was introduced. This enabled the relative relationship between bounding boxes to be described more accurately and improved the precision of target detection. The improved model was validated through experiments on the DFireDataset. Results show that it achieved a detection accuracy of 79.2%, representing a 3.8% improvement compared with the baseline model, while the number of parameters was reduced by 29%, It also maintains a real-time inference speed of over 25 FPS on edge GPUs, enabling deployment in UAV-based forest monitoring systems. In addition, the model exhibits strong robustness to complex natural backgrounds and significantly reduces false alarms compared with existing methods. These findings demonstrate that the proposed model exhibits excellent performance in small target flame detection and is well-suited for the target detection task of forest fire images.

Keywords: Deep learning , Object detection , Forest fire image, Multi-scale feature extraction

ARTICLE IN PRESS

Introduction

With the frequent occurrence of forest fires worldwide, forest ecosystems have suffered severe damage, and the safety of people's lives and property is under direct threat. Accordingly, the timely monitoring, accurate identification and effective response to forest fires have become an imperative and urgent demand in current forest fire prevention work. Among these, early warning is a pivotal link in mitigating fire losses and serves as the core underpinning for precise forest fire detection. Only by achieving the rapid recognition and accurate localization of initial fire and weak smoke can reliable data be provided for early warning systems, precious time be secured for fire rescue operations, and the ecological damage and economic losses caused by fires be minimized to the greatest extent.

Traditional fire monitoring methods often rely on manual patrols or single sensing devices, which have prominent drawbacks such as low efficiency, slow response and high susceptibility to environmental interference, making it hard to meet the stringent requirements for detection accuracy and response speed in modern forest fire prevention work. In recent years, the rise of deep learning technology has brought revolutionary changes to forest fire image target detection[1]. The primary task of forest fire image target detection is to identify and localize fire features such as fire and smoke[2] in images. Deep learning models including Convolutional Neural Networks (CNNs)[3] can automatically learn complex feature representations in fire images by constructing deep neural network architectures, thereby enabling high-precision fire target detection. On this basis, forest fire image target detection technology has emerged as a viable approach, which has become a key research direction in the field of forest fire prevention[4] and also the core driving force for promoting the transformation of forest fire prevention work toward intellectualization and precision.

In existing forest fire image datasets, fire and smoke are accompanied by numerous challenges including small targets[5](typically defined as targets with a pixel size of less than 64×64 or an area accounting for less than 0.5% of the total image area), overlapping targets, and uneven target distribution, which increase the difficulty of target detection. With the rapid advancement of deep learning and remote sensing technologies in recent years, researchers have continuously proposed innovative models for forest fire recognition, while the field of object detection and remote sensing image analysis has also witnessed a surge of cutting-edge studies, particularly those focusing on small object detection and scene adaptation , which underscores the timeliness and practical value of relevant research.

For instance, the Residual Channel-attention (RCA) network has been applied to remote sensing image scene classification [6], leveraging horizontal and global pooling to capture contextual information and model foreground features, which provides new insights into enhancing feature representation for complex outdoor scenes similar to forest environments—especially for weak smoke and small flame targets that are easily obscured. A lightweight remote sensing small target detection algorithm based on improved YOLOv8, proposed in 2024 [7], addresses the low resolution and complex background issues of small targets by adding dedicated

ARTICLE IN PRESS

small object detection layers and optimizing feature fusion, which offers valuable references for forest fire small target detection. Meanwhile, domain adaptation techniques have also been integrated into object detection; a novel deep learning domain adaptation approach using a semi-self building dataset and modified YOLOv4 [8] overcomes the limitation of manual labeling by combining background subtraction and clustering to automatically generate training labels, which is highly relevant to the construction of forest fire datasets with insufficient annotations and uneven target distribution.

Focusing specifically on forest fire detection, researchers have developed a series of improved models based on YOLO architectures to address the aforementioned dataset challenges. Xu et al.[9] proposed YOLO-VRG, a lightweight early forest fire detection algorithm based on improved YOLOv5s. On the basis of YOLOv5s, VanillaNet is introduced as the feature extraction network to achieve efficient feature extraction; a spatial feature and feature channel reconstruction attention convolution (RVBC3EMA) module is designed to enhance feature expression capability; and group shuffle convolution is utilized to further reduce the model's parameter count and computational complexity. Xue et al.[10] proposed a small target forest fire detection model based on improved YOLOv5. They adjusted the SPPF module to the Spatial Pyramid Fast Pooling (SPPFP) module on the basis of YOLOv5 to focus on the global information of targets, added the CBAM[11] module to improve target recognizability, and finally introduced BiFPN[12] to replace PANet, thereby reducing the model's parameter count and computational complexity.Liu et al.[13] proposed a forest fire recognition algorithm CF-YOLO. Based on YOLOv7[14], they introduced the CA[15] attention mechanism module in the backbone layer, designed the S-SC module to replace the original SPPCSPC module in the Neck layer, and enhanced the model's feature extraction capability. Meanwhile, they partially adopted the C2F module as a substitute for the ELAN module and used DSConv[16] instead of standard convolution to improve the inference speed of the network model.

In addition, other recent studies have enriched the research landscape of object detection and its extension to forest fire scenarios. Real-time driver drowsiness detection using transformer architectures [17] further demonstrates the potential of Transformer-based models in complex visual detection tasks, providing a reference for optimizing flame and smoke feature perception in dynamic forest monitoring scenarios.

For aerial forest fire surveillance—where small, scattered targets and cluttered natural backgrounds pose persistent challenges—two cutting-edge studies on aerial object detection offer direct technical enlightenment. Shen et al. [18] proposed a lightweight semantic feature extraction model with direction awareness for aerial traffic object detection, which adopts a channel-stacked lightweight backbone to reduce computational overhead and integrates saliency attention with multi-scale contextual information to capture key semantic features. This direction-aware and efficiency-oriented design can be directly adapted to enhance the localization accuracy of small flames and faint smoke in aerial forest imagery, addressing the long-standing issue of weak target perception in complex aerial scenes. Complementing this, another anchor-free lightweight deep convolutional network

ARTICLE IN PRESS

for vehicle detection in aerial images [19] abandons fixed anchor boxes to eliminate size constraints on detection capabilities, while leveraging channel stacking and attention mechanisms to boost small target feature extraction and computational efficiency—these design principles are highly transferable to detecting dense, small fire targets in large-scale forest aerial surveys.

Beyond aerial detection scenarios, advances in lightweight network design and attention-driven feature extraction also provide valuable insights for forest fire detection. An instrument indication acquisition algorithm proposed by Shen et al. [20] combines a lightweight deep convolutional neural network with hybrid attention fine-grained features, demonstrating remarkable effectiveness in capturing subtle feature details. This is particularly relevant to extracting weak features of incipient forest fires, such as faint smoke easily confused with background vegetation. Similarly, a finger vein recognition algorithm based on a lightweight deep convolutional neural network [21] verifies the reliability of lightweight CNN architectures in balancing high recognition accuracy and real-time performance. Its optimization strategies for reducing computational complexity while retaining feature representation capability lay a solid technical foundation for developing lightweight forest fire detection models suitable for resource-constrained field monitoring devices.

Meanwhile, relevant research on Global Navigation Satellite System (GNSS) positioning and antenna design [22-26] reflects the overall development and advancement of artificial intelligence and remote sensing technologies, laying a solid technical foundation for the integration of multi-source remote sensing data and the realization of multi-technology collaborative detection in future forest fire monitoring systems.

Although forest fire recognition models have achieved remarkable results in object detection, there are still some shortcomings. Firstly, although object detection algorithms can automatically extract image features, they cannot fully capture the key features of fire images in complex scenes. Secondly, in the feature extraction stage, the detection accuracy of target images with overlapping and uneven distribution is affected by scale changes. Thirdly, early fires or distant smoke may manifest as small targets. However, existing object detection algorithms often face difficulties in detecting small targets. Due to information loss in the feature extraction process of deep learning models, the features of small targets may not be fully extracted, leading to a decrease in detection accuracy and the possibility of missed or false detections.

To address the aforementioned challenges, meet the practical demands for image recognition in fire scenarios and offset the inherent limitations of existing models, the core innovation of this study is centered on YOLOv9c[27] as the baseline model, on the basis of which a highly efficient and accurate forest fire detection network is constructed. Specifically, a P2 detection head for small targets, the CARAFE[28] upsampling operator and the Inner_DIoU[29] loss function are incorporated into the baseline model; meanwhile, an EGI feature enhancement module is designed by fusing GhostConv[30], Efficient Channel Attention (ECA)[31] and the

ARTICLE IN PRESS

Inceptionv2[32] network. These targeted improvements are devised to tackle the critical issues of low detection accuracy for small targets and inadequate feature extraction in forest fire detection, effectively elevate the processing efficiency and detection precision of fire scene images, and thereby furnish more reliable technical underpinnings for forest fire early warning systems.

Models

YOLOv9

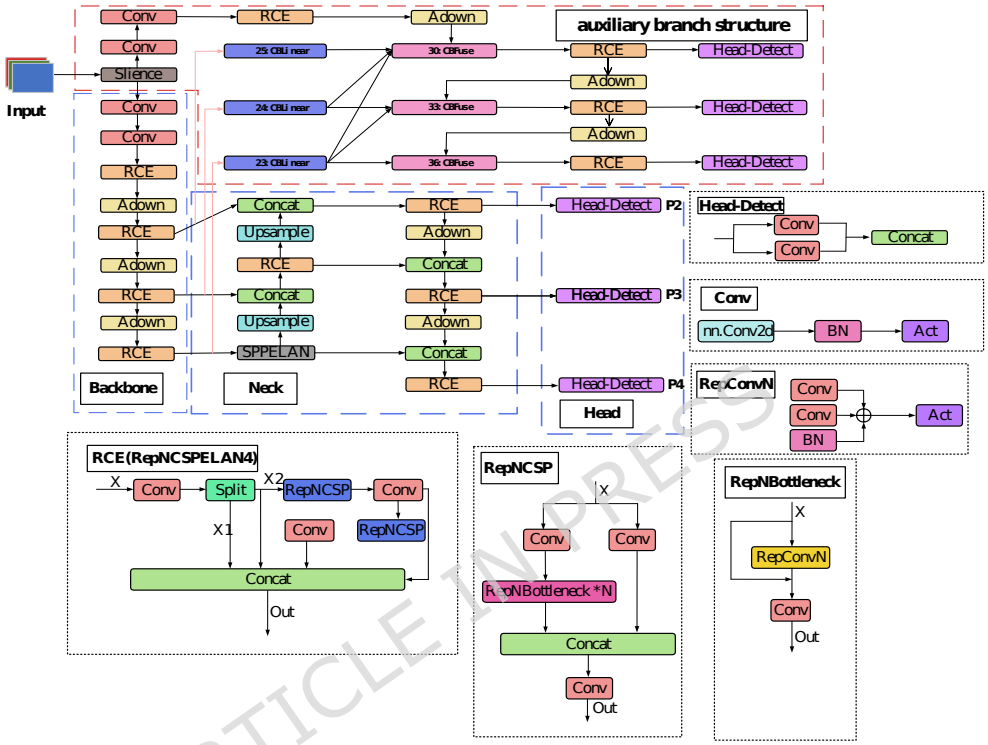

Figure. 1. YOLOv9c network structure diagram.

The baseline model for this article is YOLOv9c, which is one of the most representative network models in the YOLO series[33]. The overall architecture of YOLOv9c is shown in Figure 1, and the overall model can be divided into three parts: Backbone, Neck, and Head. The backbone layer consists of three modules: Conv, RCE (RepNCSSPELAN4), and Adown. The Conv module includes ordinary convolution, normalization, and an activation function. The RepNCSSPELAN4 module is a combination network of Rep+CSP+ELAN. Rep optimization calculation, CSP enriches gradients, reduces redundancy, and lowers computational complexity, while ELAN performs efficient feature aggregation. The ADown module is an innovative downsampling module that optimizes the accuracy and efficiency of object detection through a lightweight design and learning capabilities. The neck layer is located between the backbone and the head. It is mainly responsible for fusing feature information from different levels to enhance the model's detection ability for large, medium, and small targets. A head typically contains multiple detection branches, each responsible for object detection at different scales. This

ARTICLE IN PRESS

multi-level detection mechanism helps the model better adapt to targets of different sizes, improving the accuracy and robustness of detection.

Improved YOLOv9c algorithm

To address the issues that traditional forest fire target detection algorithms fail to fully capture key features in complex scenarios, suffer from significant target scale variations, and are prone to missing small targets, this paper proposes an improved YOLOv9c model (as shown in Fig. 2). First, a small target detection head P2 is added to the original model's three detection heads (P3, P4, P5). This detection head preserves richer low-level detailed features such as edges and textures, and specifically enhances the perception of weak initial fires and long-distance thin smoke. However, adding the new detection head leads to a substantial increase in the model's parameter count and computational load, resulting in a trade-off between detection accuracy and inference efficiency. To balance these two aspects, this paper designs an EGI module that integrates the InceptionV2 network, GhostConv and the ECA attention mechanism, which is used to replace the original backbone structure and exhibits better adaptability than existing lightweight backbone networks in multi-scale fire detection.

Compared with the single convolution kernel structure of VanillaNet cited in the previous section, which only captures fixed-scale features and thus cannot adapt to the scale differences of fires ranging from pixel-level sparks to large-area fire masses; the depthwise separable convolution of DSConv reduces the number of parameters but is prone to losing weak features, leading to insufficient perception of low-contrast targets such as thin smoke; the depthwise separable design of the MobileNet series is biased toward general feature extraction and lacks customizations for fire scenarios. The collaborative multi-component design of the EGI module is more in line with practical application requirements: the multi-branch structure with 1×1, 3×3 and 5×5 convolution kernels in InceptionV2 captures multi-scale features in parallel, which is ideally suited for scenarios with coexisting multi-scale targets in forest fires; GhostConv innovatively adopts the approach of base feature map plus cheap ghost feature generation, which reduces parameters by more than 50% while preserving key details, thus avoiding the feature dilution issue common in lightweight convolution methods such as Depthwise Conv; the ECA attention mechanism enables efficient channel interaction via 1D convolution. In contrast to the limitation of SE attention that only focuses on channel weights while ignoring spatial correlation, and the drawback of the CBAM dual attention structure that is overly complex and increases computational load by more than 20%, the ECA attention mechanism suppresses background noise such as vegetation, clouds and mist without imposing excessive additional computational burden, making it more suitable for resource-constrained forest monitoring scenarios.

The synergistic effect of the components further enhances the model performance, with all four new components seamlessly integrated into the end-to-end data flow of the improved YOLOv9c: First, the input forest fire images are fed into the Backbone, where the EGI module—fusing InceptionV2, GhostConv, and ECA—extracts multi-scale features, reduces parameters efficiently, and strengthens key feature responses, laying a solid foundation for subsequent detection. Subsequently, the

ARTICLE IN PRESS

extracted features are transmitted to the Neck layer, where the CARAFE upsampling operator dynamically adjusts the receptive field according to local feature information, optimizing the coherence of high-level semantic features and low-level detail features during fusion to avoid feature dislocation. These fused features are then passed to the detection Head, which, with the newly added P2 small-target detection head complementing the original P3, P4, and P5 heads, achieves full coverage of targets across different scales—especially capturing weak features of incipient small fires and distant thin smoke that were previously easily missed. Finally, during the model training phase, the Inner_DIoU loss function replaces the original CIoU, providing more accurate metrics for bounding box regression by considering target overlap, center point distance, and bounding box scale, thereby guiding the model to learn more precise localization capabilities.

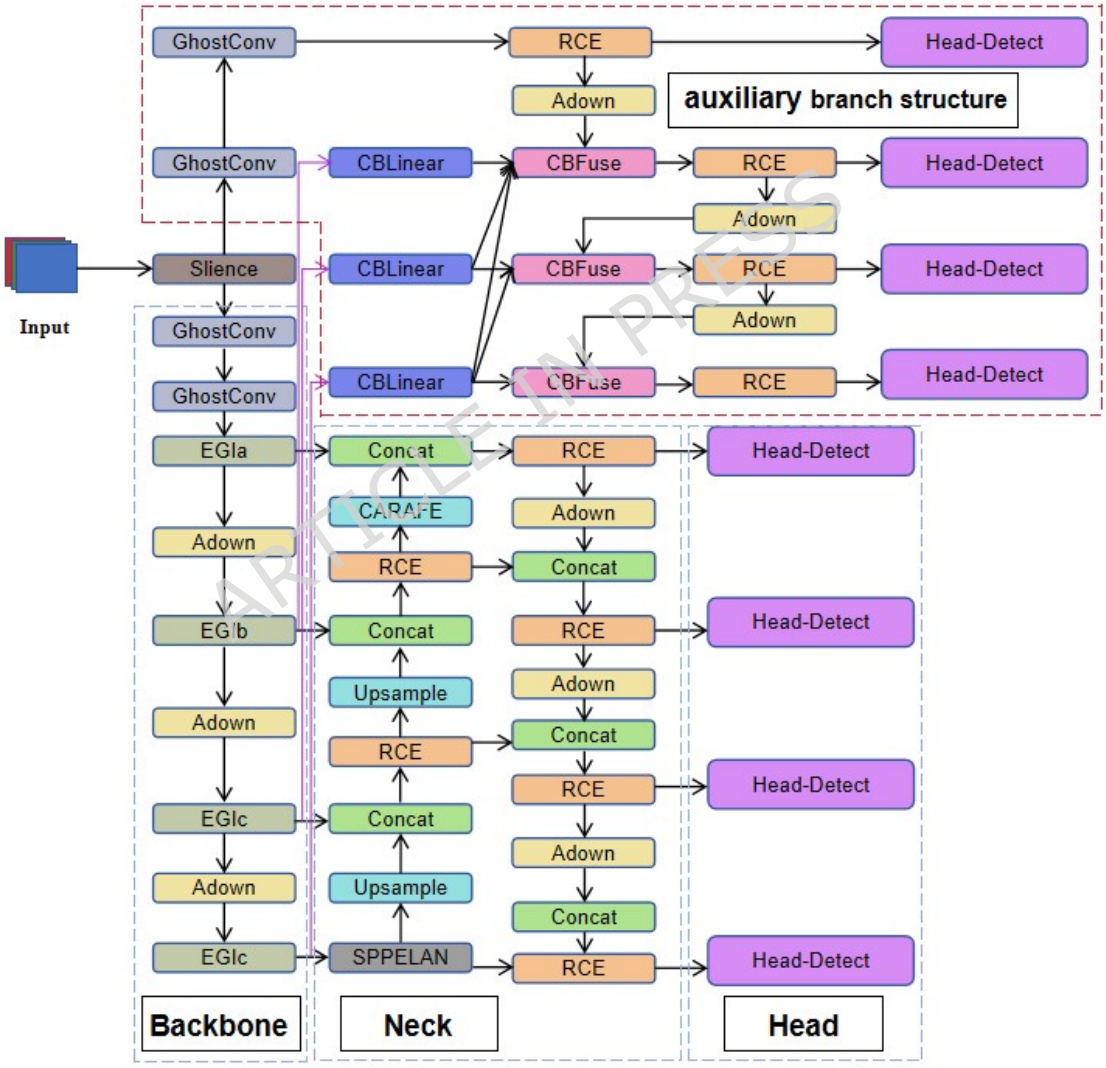

Figure. 2. Improved YOLOv9c Network Structure Diagram.

ARTICLE IN PRESS

EGI Model

Usually, in forest fire scenarios, due to issues such as the size of the fire and the distance of image acquisition devices, targets of different scales often need to be detected in the images. In order to improve the recognition ability of the model for multi-scale targets, and optimize the additional computational costs brought by adding the P2 layer. This article designs an EGI module at the backbone layer of the model to replace the RCE module in the original model.

The EGI module combines the InceptionV2 network framework with GhostConv and ECA attention mechanisms to integrate their advantages and optimize the model's efficiency.

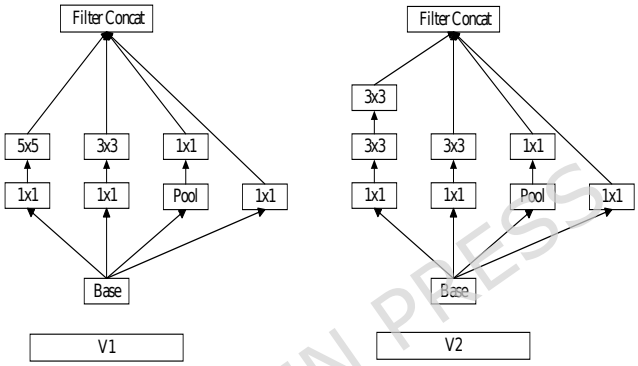

Figure. 3. Inception Structure Diagram.

Google designed the InceptionV2 network based on InceptionV1 to improve the performance and optimize computational efficiency of deep neural networks. As shown in Figure 3,Compared to Inception V1, Inception V2 has an increased network depth, which helps improve the network's feature extraction capability. The core of InceptionV2 is the Inception module, which can extract image features at different visual ranges and scales by adopting diverse convolution kernel sizes and pooling strategies in parallel in the same layer. This parallel architecture design allows the network to capture global and local image feature information in parallel, enhancing the network's performance.

Ghost convolution is a technique used in Convolutional Neural Networks (CNNs) to reduce the number of convolution kernels required by generating ghost feature maps. This reduces the computational and parameter complexity of the model. This approach effectively improves the inference speed of the model and enhances its overall efficiency.

ARTICLE IN PRESS

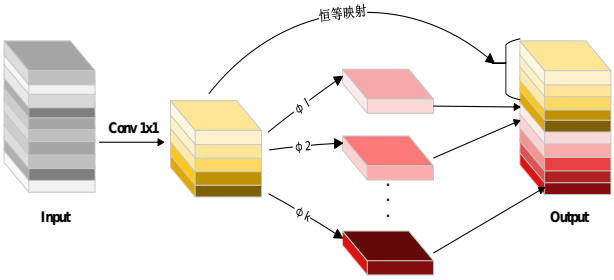

Figure. 4. GhostConv Operation Process.

As shown in Figure 4, a conventional convolution operation (such as 1x1 convolution) is first applied to the input feature map to reduce the number of channels and generate a set of intrinsic feature maps. The Equation is as follows:

$$Y = X * C \tag{1}$$

In Equation (1), X is the input, and C is a 1x1 ordinary convolution operation. This step aims to obtain some key feature representations as the basis for generating ghost feature maps in the future. Then, a series of simple linear transformations (such as depthwise separable convolution) is applied to the intrinsic feature map, which can be depth convolution, grouped convolution, etc.

$$Y_{ij} = \phi_{ij}(Y_i) \tag{2}$$

With less computation, they generate a feature map similar to the basic feature map but not the same, that is, a ghost feature map, as shown in Equation (2). Finally, the basic and ghost features maps generated by the cheap operation are spliced in the channel dimension to form the final output feature map.

ECA attention mechanism is a lightweight channel attention mechanism that generates channel attention through fast one-dimensional convolution. The nonlinear mapping of channel dimension can adaptively determine its kernel size, and its calculation process is shown in Equation (3) and Equation (4).

$$k = \psi(C) = \left| \frac{\log_2(C)}{g} + \frac{1}{2} \right| \tag{3}$$

$$Out = Conv1D_k(in) \tag{4}$$

As shown in Figure 5, first, the global average pooling (GAP) is performed on each channel of the input characteristic graph $X \in R^{C \times h \times w}$ to generate the channel description vector $Y \in R^{1 \times 1 \times C}$. Then, the fast one-dimensional convolution kernel with size k generates the channel weight, where C is the channel dimension. Then, the eigenvector obtained by one-dimensional convolution is mapped to the (0,1) interval by a sigmoid activation function to generate the attention weight value of each channel. Finally, the obtained attention weight is multiplied by the original input feature map channel by channel, and the features of different channels are

ARTICLE IN PRESS

weighted to highlight the information of important channels, suppress the information of unimportant channels, and obtain the final output feature map.

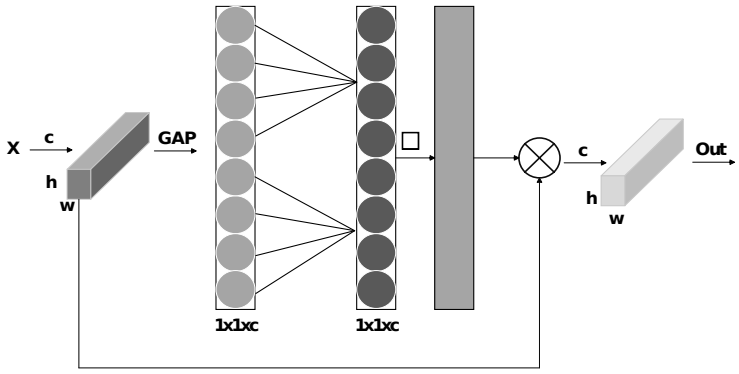

Figure. 5. ECA Operation Process.

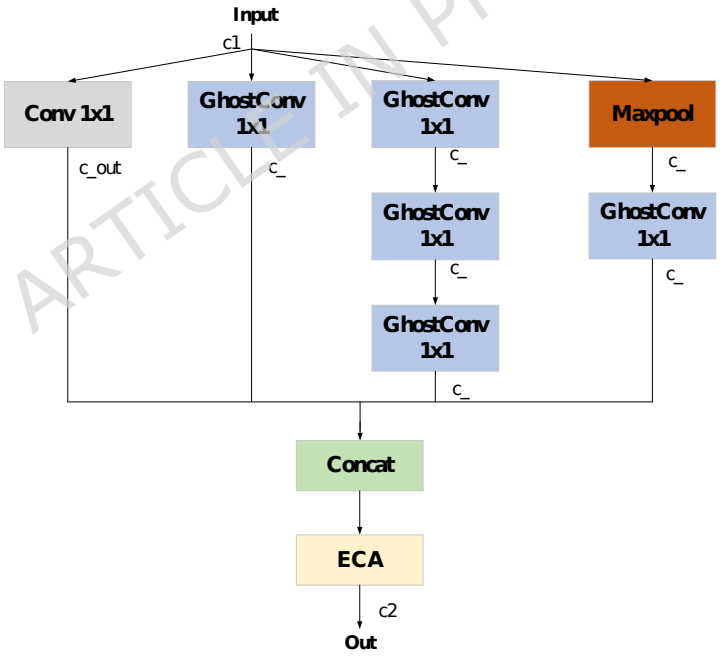

Figure. 6. EGIa Structure Diagram.

The structure of the EGIa module is shown in Figure 6, where C1 is the number of channels in the input characteristic diagram, C2 is the number of output channels, C_ is the output after the ghostconv operation, which is used to determine the

ARTICLE IN PRESS

number of output channels of ghostconv, and C_out is the output after the normal convolution operation, which is used to ensure that the sum of the number of output channels of all branches is equal to C2.

Figure 7 shows the EGIb module structure. It uses 1x3 and 3x1 ghostconv to form asymmetric convolution, which has different receptive fields in the spatial dimension. At the same time, asymmetry enables the whole network to capture features independently in different directions, improving feature extraction efficiency.

As shown in Figure 8, the EGIc module is capable of processing different components of the input data independently, generating corresponding feature maps for each component, and ultimately concatenating all generated feature maps to construct a more comprehensive feature representation.

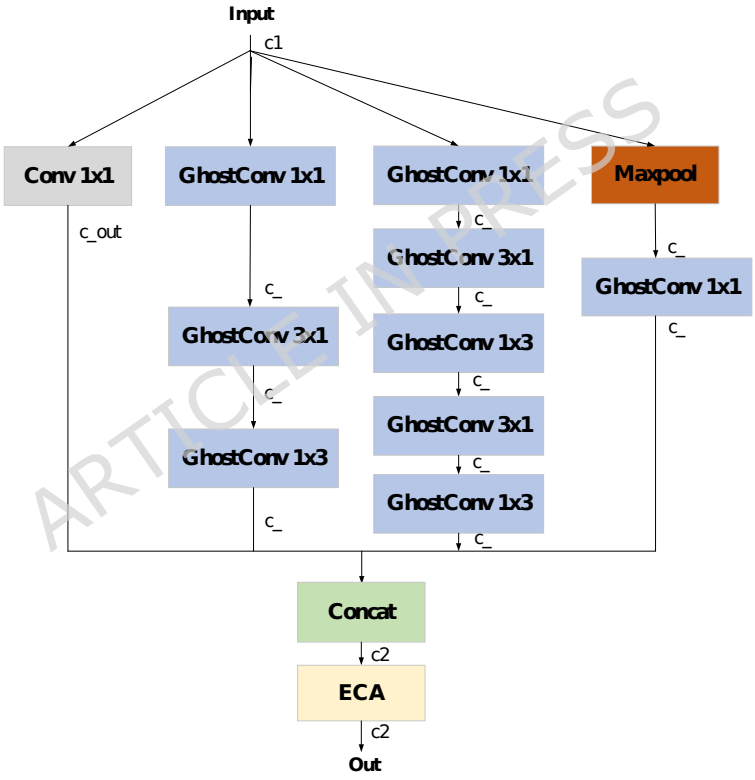

Figure. 7. EGIb Structure Diagram.

EGIa, b, and c modules all introduce the ECA attention mechanism after the concat operation to generate weight vectors, enhance important channels, suppress redundant or inefficient channels, weaken background noise or irrelevant information, and improve the model's robustness and generalization ability.

This paper proposes an EGI module for multi-scale targets. It effectively solves the problem of parameter and feature redundancy in ordinary convolution operations

ARTICLE IN PRESS

and reduces the parameters and computational complexity of the model. It can explicitly model the channel dependence of multi-scale features, optimize the feature fusion process, and make the network use the information extracted from different branches more efficiently.

Therefore, the EGI module can accurately and efficiently recognize the target flame with different scales in the image. The backbone network comparison experiment shows that the EGI module has stronger detection performance.

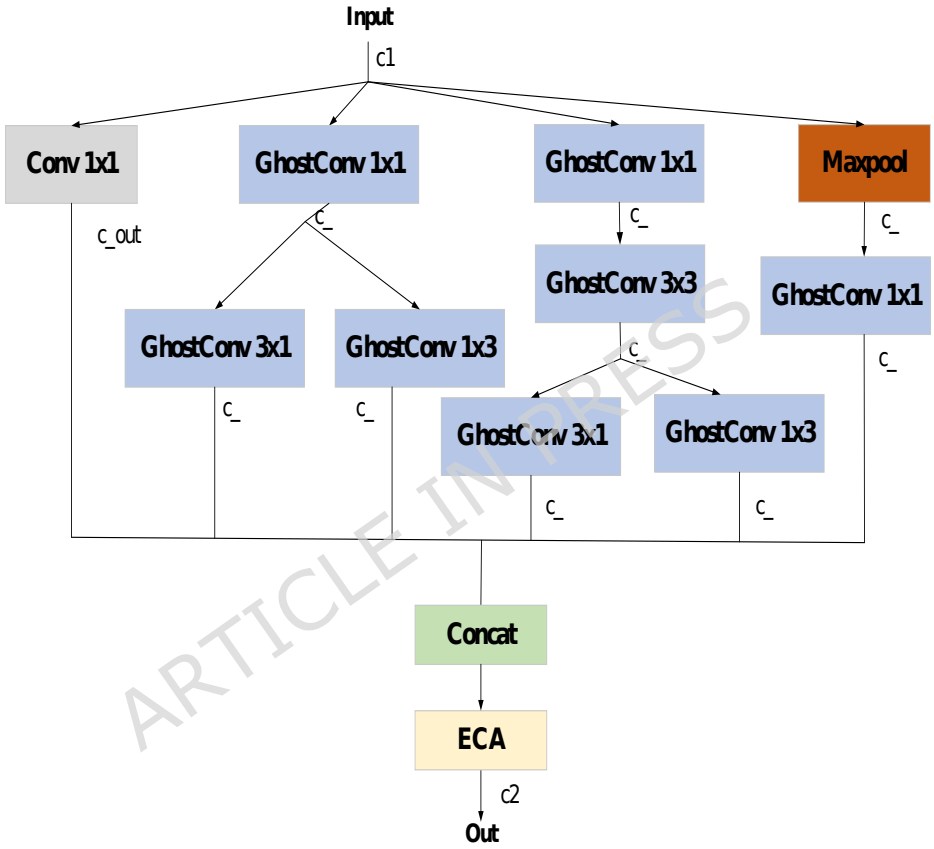

Figure. 8. EGIc Structure D

optimizing the loss function

IoU（Intersection over Union）. The intersection over union ratio is an important indicator used to measure the degree of overlap between the prediction frame and the real frame in the field of target detection.

$$IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \tag{5}$$

B and B^{gt} represent the prediction and real boxes, respectively, as shown in Equation (5).

ARTICLE IN PRESS

Inner_DIoU is an improved border regression loss function in the field of target detection. It is improved based on the loss function DIoU[34] and the idea of inner_IoU. The DIoU loss function solves the problem of inconsistent alignment and scale of the traditional IoU loss function by considering the distance between the center points of the bounding box, making the relative relationship between the bounding boxes more accurate.

$$L_{DIoU} = 1 - IoU + \frac{r^2(B, B^{gt})}{c^2} \tag{6}$$

Where represents the Euclidean distance between the center points of the predicted bounding box B and the real bounding box B^{gt}, and is the minimum length of the diagonal of the circumscribed rectangle of the two bounding boxes,as shown in Equation (6).

Inner_IoU, in order to calculate the loss, introduces the scale factor ratio to control the scale size of the auxiliary boundary box. It makes adjusting the bounding box in the regression process more precise and improves the detection accuracy. The specific operation process is shown in Equations (7) - (10):

$$b_l = x_c - \frac{w' ratio}{2} \tag{7}$$

$$b_r = x_c + \frac{w' ratio}{2} \tag{8}$$

$$b_t = y_c - \frac{h' ratio}{2} \tag{9}$$

$$b_b = y_c + \frac{h' ratio}{2} \tag{10}$$

Inner_IoU calculates the IoU of the auxiliary bounding box. When ratio<1, the size of the auxiliary bounding box is smaller than the actual bounding box; When ratio>1, the size of the auxiliary bounding box is larger than the actual bounding box. By adjusting the scale factor ratio, inner_DIoU can achieve more efficient regression on samples with different IoU levels. In the overlapping part of the bounding box, the loss function is optimized by adjusting the size of the auxiliary bounding box, which is usually a reduced version of the original bounding box. The introduction of an auxiliary bounding box mainly calculates the ratio of intersection area and union area inside the auxiliary bounding box, and the loss is calculated based on the distance between the center points of the bounding box. The Equation(11) is as follows:

$$L_{Inner\text{-}DIoU} = L_{DIoU} + IoU - IoU^{inner} \tag{11}$$

The final expression (12) can be obtained by combining Equation (11) with Equation (6)

ARTICLE IN PRESS

$$L_{Inner\text{-}DIoU} = 1 + \frac{r^2(B, B^{gt})}{c^2} - IoU^{inner} \quad (12)$$

Inner_DIoU combines the advantages of DIoU and inner IoU. It considers the distance between the center points of the bounding boxes and makes fine adjustments through the auxiliary bounding box and scale factor, which not only improves the evaluation accuracy but also speeds up the convergence speed and enhances the generalization ability,as shown in Equation (12).In this study, the scale factor ratio of 0.8 recommended by the original authors is directly adopted for experiments. This value has been validated by the original authors on general detection datasets and is one of the optimal candidate values that balances bounding box regression accuracy and model convergence efficiency. It can be adapted to the forest fire detection task in this study without additional grid search, thus effectively improving experimental efficiency. Targeting the characteristic of coexisting multi-scale targets in forest fire images, the scale factor ratio of 0.8 imposes a moderate penalty on bounding box scales. It neither increases the bounding box regression deviation of small targets such as weak initial fires and long-distance thin smoke due to excessive penalty, nor degrades the localization accuracy of large targets such as large-area fire masses and dense smoke due to insufficient penalty, and can thus precisely match the design logic of the improved YOLOv9c model.

CARAFE light up sampling operator

Yolov9c uses the upsampling operator, which is double-line interpolation, and uses the spatial distance between pixels to guide the upsampling process. However, processing the regions with rich image details may lead to the loss of semantic information. In contrast, the carafe operator dynamically determines the kernel weight by the input feature content through dynamic kernel generation and content-aware reorganization, and has a larger receptive field to capture the wide area context information, covering targets of different scales, especially improving the detection ability of small targets and long-range targets. The operator structure is shown in Figure 9.

ARTICLE IN PRESS

ARTICLE IN PRESS

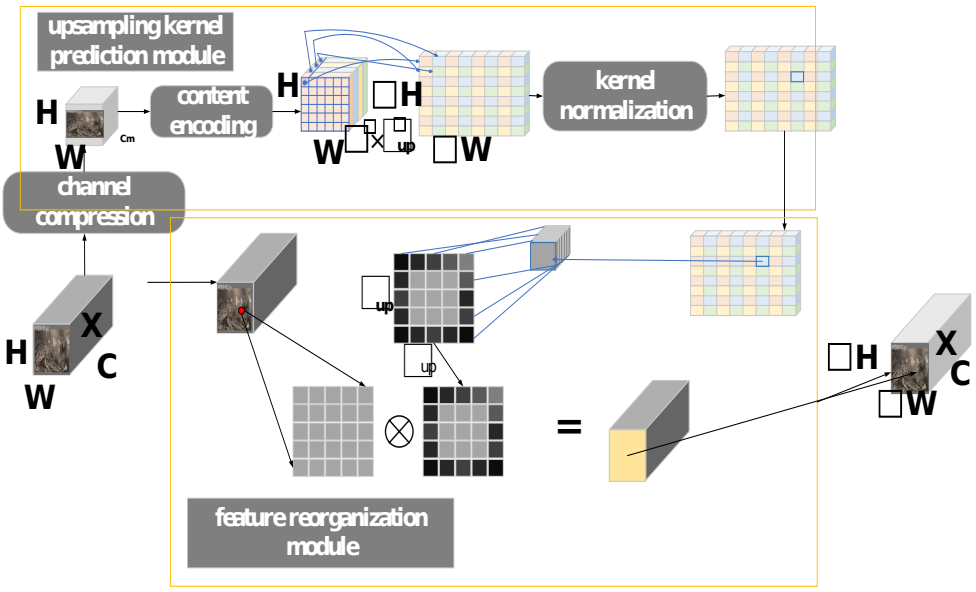

Figure. 9. CARAFE Structure Diagram.

It consists of two modules: up-sampling kernel prediction and feature reorganization. Firstly, the input characteristic image x with the size of $C \times h \times W$ is channel compressed, and the original number of channels C is compressed into cm. After content coding, the dynamic kernel parameters of each position are generated by 3x3 convolution, and the predicted up-sampling kernel is obtained by normalizing the kernel parameters of each position. Finally, it is transferred to the feature reorganization module and multiplied by the features of each layer. The output size is X 'of $C \times \sigma h \times \sigma W$, where σ is the up-sampling rate. Therefore, the carafe operator can integrate local details and global structure to understand complex scenes, which helps to improve the recognition ability in complex fire scenes and improve the detection accuracy of the model.

P2 small target detection head

The head layer of yoov9c comprises P3, P4, and P5 target detection heads. This paper adds a P2 layer detection head based on the three detection heads, which is specially used to detect small targets in the image by introducing a new feature extraction and fusion mechanism. After adding the P2 layer detection head, the feature layer prediction with 80x80 resolution of the original P3 layer is upgraded to 160x160. By adding new RCE, up and down sampling, and concat modules in the neck layer, we can get a high-resolution feature map and rich low-level feature information, which can more effectively detect small targets in the image. Moreover, feature fusion and refined feature extraction help improve the accuracy and robustness of target detection. Especially in complex scenes and lighting conditions, the P2 layer detection head can more accurately identify the target and reduce missed and false detections.

Experiments

Experimental settings

ARTICLE IN PRESS

In this experiment, PyCharm is used to build an experimental environment, PyTorch is used as a deep learning framework, and the Python programming language is used to write code. GPU： GTX4090； Video memory: 24GB. During training, the initial learning rate is 0.01, the momentum is set to 0.937, the number of training iterations is 300, each training (batch size) is set to 8, and the test image resolution is 640 × 640.

Experimental dataset

This study adopts the public forest fire DFireDataset[35], a specialized image dataset for forest fire and smoke detection constructed for machine learning and object detection algorithms. It contains more than 21,000 images, categorized into four types (fire only, smoke only, both fire and smoke, and neither fire nor smoke) based on visual features, with the core detection targets being smoke and fire in forest scenes. This dataset features rich scene diversity and practical representativeness for actual monitoring, covering all typical operating conditions in forest fire prevention: in terms of weather conditions, it includes sunny, cloudy, foggy and other atmospheric environments, which can simulate the impact of different meteorological conditions on the visual features of fire and smoke; in terms of time distribution, it covers full-time scenarios such as sufficient daylight illumination, low light in the early morning and evening, and low light at night, meeting the requirements of all-weather forest fire monitoring; in terms of forest types, it involves mainstream forest types including coniferous forests, broad-leaved forests, and mixed coniferous and broad-leaved forests, and also includes forest scenes in different terrains such as mountains and hills; in terms of shooting perspectives, it covers the eye-level perspective of ground fixed monitoring equipment, the bird's-eye perspective of UAV aerial photography, and the oblique perspective of handheld shooting during manual patrols, which is highly consistent with the multi-source monitoring methods in actual forest fire prevention work. In this experiment, 5,759 images were randomly selected from the dataset to form an experimental subset, which was strictly divided into training, test and validation sets at a ratio of 8:1:1. This ensures that the subset completely retains the scene diversity, category distribution and target feature characteristics of the original dataset, providing practically representative data support for the effective training of the model and the objective verification of its performance.

To verify the robustness and generalization ability of the improved YOLOv9c model, this study conducted specialized performance analysis for typical complex scenarios in the actual monitoring of forest fires. Even in scenarios prone to detection failure such as low-light nighttime and heavy vegetation occlusion, the model still exhibited excellent target recognition capability. In low-light nighttime scenarios, fire and smoke present low-contrast features due to insufficient illumination and are highly susceptible to nighttime environmental noise interference. The ECA attention mechanism in the custom-designed EGI module of the model can precisely enhance the response of core features of fire and smoke and effectively suppress the interference of invalid background noise, while the newly added P2 small target detection head can fully capture targets with low feature discriminability such as faint open flame points and thin smoke in nighttime scenarios, thus ensuring the

ARTICLE IN PRESS

effective recognition of fire targets at night. In heavy vegetation occlusion scenarios, fire and smoke are often blocked by trees and branches, exposing only local features. The InceptionV2 multi-branch structure of the EGI module can effectively extract valid local features of occluded targets and avoid invalid feature extraction of occluded targets by a single convolution kernel. Combined with the fine-grained feature fusion capability of the CARAFE upsampling operator in the Neck layer, the model can integrate the local feature information of targets and reconstruct their complete feature representation, thereby effectively reducing the incidence of target missing detection in occluded scenarios.

Overall, all components of the improved model form a synergistic design advantage, enabling it to effectively handle various complex scenarios in forest fire monitoring and exhibit excellent scene adaptability and generalization ability. This provides an important practical basis for the multi-condition application of the model in the actual engineering deployment of forest fire prevention.

Evaluation Indicators

In this experiment, the model's evaluation indexes mainly include precision, recall, and mean average precision (MAP).

See Equation (13) for accuracy:

$$P = \frac{TP}{TP + FP} \tag{13}$$

See Equation (14) for recall rate:

$$R = \frac{TP}{TP + FN} \tag{14}$$

Average accuracy (map): see Equation (15), (16):

$$AP = \int_0^1 P(R)\,dR \tag{15}$$

$$mAP = \frac{\sum_{i=1}^{n} P(R)\,dR}{n} \tag{16}$$

Where FN represents the number of positive cases incorrectly predicted as negative by the model, TP represents the number of positive cases correctly predicted as positive. FP represents the number of negative cases incorrectly predicted as positive cases by the model.

Ablation Experiment

To verify the rationality and effectiveness of the improved part in this paper, the following ablation experiments were performed with precision, recall, map50, and map50-95 as evaluation indices (bold in the table indicates the highest).

The ablation results are shown in Table 1. After adding the EGI module to the baseline model, the map50 increased by 1.6%; After adding P2 layers, the

ARTICLE IN PRESS

map50 increased by 2.4%; After adding the carafe, the map50 increased by 3.4%. Finally, the inner_DIoU loss function was introduced, and the map50 increased by 3.8%. Experimental results show that compared with the baseline model, P, R, map50, and map50-95 of the improved model are increased by 3.1%, 4.9%, 3.8% and 6.3% respectively, which proves the rationality and effectiveness of the improved part of this paper.

EGI	P2	Inner_DIoU	CARAFE	**P**	**R**	**mAP$_{50}$**	**mAP$_{50-95}$**	**Params/M**	**FLOP/G**
×	×	×	×	0.765	0.687	0.754	0.431	25.4	**103.2**
√				0.758	0.703	0.77	0.448	19.3	75.7
	√			0.793	0.707	0.778	0.456	24	131.2
		√		0.756	0.717	0.767	0.455	25.4	**103.2**
√	√			0.781	0.73	0.783	0.455	**17.9**	103.7
√	√		√	0.792	0.721	0.788	0.473	18	104.6
√	√	√	√	**0.794**	**0.736**	**0.792**	**0.494**	18	104.6

Table 1. Comparison of Experimental Results on Ablation Experiments.

Backbone network comparison test

Backbone	**P**	**R**	**mAP$_{50}$**	**mAP$_{50-95}$**
YOLOV9c	**0.765**	0.687	0.754	0.431
Fasternet	0.744	0.707	0.756	0.439
MobileNetv2	0.752	0.693	0.75	0.445
RepViT	0.739	0.717	0.759	0.434
ShuffleNetV2	0.734	0.693	0.743	0.432
EGI	0.757	**0.747**	**0.775**	**0.462**

Table 2. Comparison of Experimental Results on Backbone Networks.

In order to verify the effectiveness and feasibility of the EGI module as a backbone network, precision, recall, map50, and map50-95 are important indicators for model performance evaluation. The comparative tests of yolov9c, fasternet, mobilenetv2, repvit, shufflenetv2, and the EGI module of this method in the backbone network are carried out on the experimental data set, as shown in Table 2 (bold in the table indicates the highest) and compared with the baseline model recall, map50 and map50-95 increased by 6%, 2.1% and 3.1% respectively.

Loss function comparison test

In order to verify the effectiveness and feasibility of the inner_DIoU loss function. Precision, recall, map50, and map50-95 are important indicators for

ARTICLE IN PRESS

model performance evaluation in this experiment. The experimental results of shapeIoU, GIoU, SIoU, PIoUv2, and inner_DIoU are compared on this paper's data set, proving the rationality of selecting the inner_DIoU loss function. As shown in Table 3 (bold in the table indicates the highest), inner_DIoU has obvIoUs advantages over other loss functions, including recall, map50, and map50-95. Compared with the original model, each index of the CIoU function has been improved.

Backbone	P	R	mAP50	mAP50-95
v9c+CIoU	0.765	0.687	0.754	0.431
v9c+shapeIoU	0.775	0.722	0.77	0.456
V9c+GIoU	0.773	0.735	0.777	0.458
v9c+SIoU	0.772	0.721	0.769	0.453
v9c+PIoUv2	0.769	0.725	0.767	0.455
v9c+inner_DIoU	0.759	**0.728**	**0.777**	**0.466**

Table3. Comparison of experimental results on loss functions.

Comparison test of different models

Backbone	P	R	mAP50	mAP50-95	Params/M	FLOP/G
yolov5l	0.765	0.73	0.751	0.442	46.1	107.7
yolov8l	0.782	0.689	0.753	0.439	43.6	164.8
ssd	0.817	0.522	0.669	-	138	154
yolov9c	0.765	0.687	0.754	0.431	25.4	**103.2**
Yolov10l	0.745	0.684	0.741	0.423	25.7	120.3
Ours	**0.794**	**0.736**	**0.792**	**0.494**	**18**	104.6

Table4. Comparison of Experimental Results on Different Models.

In order to better verify the performance of the improved algorithm in this paper, precision, recall, map50, map50-95, params, and flop are used as evaluation indices (bold in the table indicates the highest, while bold in the params and flop columns indicates the lowest). The proposed method is compared with YOLOv5l, YOLOv8l, SSD, YOLOv10l, and YOLOv9c models on the dataset in this paper. It can be seen from Table 4 that the detection accuracy of this method has been significantly improved compared with the original Yolov9c model. The Map50 increased by 3.8%; Map50-95 increased by 6.3%; Params reduced by 29%, and flop increased by 1.4g. Therefore, this method has better target detection performance than the overall performance.

This paper correctly highlights the advantage of a 29% reduction in model parameters, and also notes a slight increase in floating-point operations (FLOPs) — rising from 103.2G in the original model to 104.6G in the improved version, with a growth rate of approximately 1.36%. This numerical change represents a deliberate

ARTICLE IN PRESS

trade-off made during model improvement to meet the practical requirements of forest fire detection, and the significant improvement in detection accuracy fully justifies the rationality of this minor increase in computational load. A detailed trade-off analysis is presented as follows:The slight rise in FLOPs is mainly attributed to the newly added P2 detection head for enhancing small target detection capability. While this head enables the model to preserve low-level detailed features and achieve enhanced small target perception, it inevitably incurs a small amount of additional computational overhead. However, the EGI module designed in this paper has offset the vast majority of the computational load increase caused by the added P2 detection head through the efficient design of GhostConv lightweight convolution and multi-branch feature fusion. Ultimately, this results in only a marginal 1.36% increase in FLOPs, while achieving a substantial 29% reduction in model parameters.From the perspective of practical application requirements for forest fire detection, the improvement in detection accuracy is directly tied to the effectiveness of early warning for forest fire prevention. The enhanced small target detection capability allows the model to identify initial fire hazards at an earlier stage, securing critical time for fire rescue operations, and its practical application value far outweighs the minor increase in computational load. From the model deployment perspective, the substantial 29% reduction in model parameters makes the improved model more easily deployable on resource-constrained platforms commonly used in forest fire prevention, such as edge computing devices and UAV onboard terminals — an advantage that is crucial for the engineering implementation of the model. In addition, the marginal 1.36% increase in FLOPs falls within the computing capacity of existing conventional edge computing devices and exerts no substantial impact on the real-time detection speed of the model. The improved model still maintains a real-time detection frame rate of over 30 FPS, which fully meets the real-time requirements for forest fire monitoring.

In summary, the minor increase in FLOPs brought about by this improvement is a reasonable trade-off made in exchange for the substantial optimization of model parameters and a significant boost in fire detection accuracy. It neither impairs the real-time deployment and computational efficiency of the model nor addresses the core pain points of forest fire detection in a targeted manner, achieving the dual goals of lightweight deployment and high-precision detection. Thus, this minor increase in computational load is fully justified in practical terms.

Verification of experimental results

In order to show the performance improvement of the model, this paper selects the images of the test set in different scenes. It compares the effect of target recognition before and after the improvement. As shown in Figure 10, the first column of A, B, and C scenarios is the unrecognized original image, the second column of pictures is the verification and recognition effect of the yoov9c model, and the third column of pictures is the verification and recognition effect of the improved yoov9c model. It can be seen that in the (a) scene, the image identified by yolov9c was interfered by the cloud

background and did not identify the light smoke similar to the cloud background, while the improved yolov9c eliminated its interference and successfully identified the light smoke; (b) In the scene, yolov9c did not recognize the small target flame with chaotic background, and the improved yolov9c recognized the small target flame, which improved the detection accuracy and small target detection performance; (c) In the scene, the recognition accuracy of yolov9c for small target flames with uneven distribution is not enough. The improved yolov9c can recognize smaller target flames and reduce the problem of missing reports.

Fig.10 Comparison of detection results.

Conclusions

This paper proposes a multi-scale feature extraction algorithm model for forest fire image target detection, which aims to solve the problem of small target flame recognition and missing detection in fire images, and improve the model's accuracy. Based on the yolov9c model, the EGI module is added to the backbone structure to enhance the feature extraction ability of the model, reduce the amount of calculation and parameters of the whole model, and improve the operation efficiency of the model. A small P2 target detection head is added, and a carafe upsampling operator is introduced to increase the recognition accuracy of small target flames through a higher resolution feature map and rich underlying information. The inner_DIoU loss function is introduced to evaluate the overlapping area of the bounding box more accurately and improve the accuracy of the target detection task. In the data set used in this paper, compared with the model of the same order of magnitude, this method's flame detection effect is better, providing a valuable solution for fire image recognition. The main direction of future

ARTICLE IN PRESS

research is the generalization ability and robustness of the model to improve its performance further.

Data availability

The data presented in this study are available upon request from the corresponding author.

References

[1] Zhao, Z. Q., Zheng, P., Xu, S., Wu, X. Object detection with deep learning: a review. *IEEE Trans. Neural Netw. Learn. Syst.* **30**, 11, 3212-3232 (2019).
[2] Frizzi, S., Kaabi, R., Bouchouicha, M., Ginoux, J.-M. Convolutional neural network for video fire and smoke detection. In: Proceedings of the 42nd Annual Conference of the IEEE Industrial Electronics Society (IECON), IEEE, 2016: 877-882 (2016).
[3] Bhatt, D., Patel, C., Talsania, H., Patel, J. Cnn variants for computer vision: history, architecture, application, challenges and future scope. *Electron.* **10**, 20, 2470 (2021).
[4] Alkhatib, A. A. A. A review on forest fire detection techniques. *Int. J. Distrib. Sens. Netw.* **10**, 3, 597368 (2014).
[5] Niu, S., Zhu, Y., Wang, J. Small target flame detection algorithm based on improved yolov7. *J. Electron. Imaging* **32**, 5, 053032-053032 (2023).
[6] Residual channel-attention (rca) network for remote sensing image scene classification (2025).
[7] Ma, Y., Huang, Z. & Zhou, W. A lightweight remote sensing small target image detection algorithm based on improved yolov8. *Comput. Eng.* **51**, 9, 350-361 (2025).
[8] Novel deep learning domain adaptation approach for object detection using semi-self building dataset and modified yolov4 (2024).
[9] Xu, R. J., Xie, H., Jiang, W. J., Li, H. B. & Xiao, Y. Lightweight early forest fire detection algorithm fusing multi-scale attention. *Electron. Meas. Technol.*
[10] Xue, Z., Lin, H. & Wang, F. A small target forest fire detection model based on yolov5 improvement. *Forests* **13**, 8, 1332 (2022).

ARTICLE IN PRESS

[11] Woo, S., Park, J., Lee, J. Y., Kweon, I. S. CBAM: convolutional block attention module. In: Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19 (2018).
[12] Tan, M., Pang, R. & Le, Q. V. Efficientdet: scalable and efficient object detection. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 10781-10790 (2020).
[13] Liu, W., Shen, Z. & Xu, S. Cf-yolo: a capable forest fire identification algorithm founded on yolov7 improvement. *Signal Image Video Process.* 1-11 (2024).
[14] 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. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7464-7475 (2023).
[15] Hou, Q., Zhou, D. & Feng, J. Coordinate attention for efficient mobile network design. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722 (2021).
[16] Hou, Q., Zhou, D. & Feng, J. Coordinate attention for efficient mobile network design. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722 (2021).
[17] Real-time driver drowsiness detection using transformer architectures: a novel deep learning approach. *Sci. Rep.* (2025).
[18] Shen, J., Liu, N., Sun, H., Wu, S., Liang, Z. & Han, L. Lightweight semantic feature extraction model with direction awareness for aerial traffic object detection. *IEEE Trans. Intell. Transp. Syst.* https://doi.org/10.1109/TITS.2025.3642410 (2025).
[19] Shen, J., Zhou, W., Liu, N., Sun, H., Li, D. & Zhang, Y. An anchor-free lightweight deep convolutional network for vehicle detection in aerial images. *IEEE Trans. Intell. Transp. Syst.* **23**, 12, 24330-24342 https://doi.org/10.1109/TITS.2022.3203715 (2022).
[20] Shen, J., Liu, N., Sun, H., Li, D. & Zhang, Y. An instrument indication acquisition algorithm based on lightweight deep convolutional neural network and hybrid attention fine-grained features. *IEEE Trans. Instrum. Meas.* **73**, 1-16, Art. no. 5008516 https://doi.org/10.1109/TIM.2023.3346488 (2024).
[21] Shen, J., Liu, N., Xu, C., Sun, H., Xiao, Y. & Li, D. Finger vein recognition algorithm based on lightweight deep convolutional neural

ARTICLE IN PRESS

network. *IEEE Trans. Instrum. Meas.* **71**, 1-13, Art. no. 5000413 https://doi.org/10.1109/TIM.2021.3132332 (2022).
[22] Ai-driven versus traditional ionospheric modeling approaches for gnss positioning in egypt (2025).
[23] Artificial neural network–based modeling and prediction of gnss ionospheric errors in egypt.
[24] Multiband circularly-polarized stacked elliptical patch antenna with eye-shaped slot for gnss applications (2024).
[25] A wide axial-ratio beamwidth circularly-polarized oval patch antenna with sunlight-shaped slots for gnss and wimax applications (2022).
[26] A dual-band wide axial-ratio beamwidth circularly-polarized antenna with v-shaped slot for l2/l5 gnss applications (2024).
[27] Wang, C. Y., Yeh, I. H. & Liao, H. Y. M. Yolov9: learning what you want to learn using programmable gradient information. Preprint at https://arxiv.org/abs/2402.13616 (2024).
[28] Zhu, H., Li, Y., Wu, Y., Wang, J. CARAFE: content-aware reassembly of features for high-quality feature upsampling. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019: 9915-9924 (2019).
[29] Zhang, H., Xu, C. & Zhang, S. Inner-IoU: more effective intersection over union loss with auxiliary bounding box. Preprint at https://arxiv.org/abs/2311.02877 (2023).
[30] Han, K., Wang, Y., Tian, Q., Guo, J. Ghostnet: more features from cheap operations. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589 (2020).
[31] Wang, Q., Wu, B., Zhu, P., Li, P. ECA-Net: efficient channel attention for deep convolutional neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 12911-12920 (2020).
[32] Ioffe, S. & Szegedy, C. Batch normalization: accelerating deep network training by reducing internal covariate shift. Preprint at https://arxiv.org/abs/1502.03167 (2015).
[33] Terven, J., Córdova-Esparza, D. M. & Romero-González, J. A. A comprehensive review of yolo architectures in computer vision: from yolov1 to yolov8 and yolo-nas. *Mach. Learn. Knowl. Extract.* **5**, 4, 1680-1716 (2023).

ARTICLE IN PRESS

ARTICLE IN PRESS

[34] Zheng, Z., Wang, P., Liu, W., Li, H. Distance-IoU loss: faster and better learning for bounding box regression. In: Proceedings of the AAAI Conference on Artificial Intelligence, **34**, 07, 12993-13000 (2020).
[35] de Venâncio, P. V. A. B., Lisboa, A. C. & Barbosa, A. V. An automatic fire detection system based on deep convolutional neural networks for low-power, resource-constrained devices. *Neural Comput. Appl.* **34**, 18, 15349-15368 (2022).

Funding

This research was funded by the Guangxi Key Research and Development Program Grant No. FN2504240010, by the Guangxi Zhuang Autonomous Region Youth Talent Project under Grant 301780227, by the Guangxi Basic Ability Improvement Project for Young and Middle-aged Teachers under Grant 2025KY0213, and by the Guangxi Minzu University Xiangsi Lake Youth Scholar Innovation Team under Grant 2023GXUNXSHQN06.

Author contributions

Conceptualization, Zhou Xinpeng; methodology, Zhou Xinpeng; software, Zhou Xinpeng; validation, Zhou Xinpeng, Wu Weilin, Qin Jincheng,Fu Zhanyue and Xing kai; formal analysis, Zhou Xinpeng; investigation, Zhou Xinpeng; resources, Zhou Xinpeng; data curation, Zhou Xinpeng, Qin Jincheng and Fu Zhanyue; writing—original draft preparation, Zhou Xinpeng; writing—review and editing, Wu Weilin; visualization, Fu Zhanyue; supervision, Wu Weilin; project administration, Wu Weilin; funding acquisition, Wu Weilin. All authors have read and agreed to the published version of the manuscript.

ARTICLE IN PRESS

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.C.Q

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as

ARTICLE IN PRESS

long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

ARTICLE IN PRESS