



Article

SIMCB-Yolo: An Efficient Multi-Scale Network for Detecting Forest Fire Smoke

Wanhong Yang , Zhenlin Yang, Meiyun Wu, Gui Zhang, Yinfang Zhu and Yurong Sun * 

College of Computer and Mathematics, Central South University of Forestry & Technology, Changsha 410004, China; 20211685@csuft.edu.cn (W.Y.); 20212692@csuft.edu.cn (Z.Y.); 20212763@csuft.edu.cn (M.W.); csfu3s@163.com (G.Z.); yinfangzhu@csuft.edu.cn (Y.Z.)

* Correspondence: yurongsun@csuft.edu.cn

Abstract: Forest fire monitoring plays a crucial role in preventing and mitigating forest disasters. Early detection of forest fire smoke is essential for a timely response to forest fire emergencies. The key to effective forest fire monitoring lies in accounting for the various levels of forest fire smoke targets in the monitoring images, enhancing the model's anti-interference capabilities against mountain clouds and fog, and reducing false positives and missed detections. In this paper, we propose an improved multi-level forest fire smoke detection model based on You Only Look Once v5s (Yolov5s) called SIMCB-Yolo. This model aims to achieve high-precision detection of forest fire smoke at various levels. First, to address the issue of low precision in detecting small target smoke, a Swin transformer small target monitoring head is added to the neck of Yolov5s, enhancing the precision of small target smoke detection. Then, to address the issue of missed detections due to the decline in conventional target smoke detection accuracy after improving small target smoke detection accuracy, we introduced a cross stage partial network bottleneck with three convolutional layers (C3) and a channel block sequence (CBS) into the trunk. These additions help extract more surface features and enhance the detection accuracy of conventional target smoke. Finally, the SimAM attention mechanism is introduced to address the issue of complex background interference in forest fire smoke detection, further reducing false positives and missed detections. Experimental results demonstrate that, compared to the Yolov5s model, the SIMCB-Yolo model achieves an average recognition accuracy (mAP50) of 85.6%, an increase of 4.5%. Additionally, the mAP50-95 is 63.6%, an improvement of 6.9%, indicating good detection accuracy. The performance of the SIMCB-Yolo model on the self-built forest fire smoke dataset is also significantly better than that of current mainstream models, demonstrating high practical value.

Keywords: SIMCB-Yolo; forest fire smoke detection; SimAM; Swin transformer



Citation: Yang, W.; Yang, Z.; Wu, M.; Zhang, G.; Zhu, Y.; Sun, Y. SIMCB-Yolo: An Efficient Multi-Scale Network for Detecting Forest Fire Smoke. *Forests* **2024**, *15*, 1137. <https://doi.org/10.3390/f15071137>

Academic Editors: Rafael Coll Delgado and Rafael De Ávila Rodrigues

Received: 17 June 2024
Revised: 24 June 2024
Accepted: 25 June 2024
Published: 29 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Forests are often referred to as the “lungs of the Earth”, serving not only significant economic purposes [1] but also providing vital ecosystem services and holding an indispensable position within global ecosystems [2]. However, in recent years, as a result of climate change, the scale, intensity, and duration of forest fires have escalated [3], posing significant challenges to the conservation and protection of forests.

Forest fires are difficult to detect in their early stages due to dense forest vegetation cover. However, under the influence of forest environmental humidity, the insufficient combustion of combustibles can produce a small amount of smoke. Hence, smoke detection in forest fires is crucial for enhancing early warning capabilities. Traditional methods, such as manually constructed watchtowers, rotating patrols, and real-time observation by forest rangers, have limited detection range and duration and require significant human and material resources. Consequently, detecting and identifying forest fire smoke through images or videos captured by various cameras [4] proves to be an effective approach.

Early algorithms for forest fire detection primarily identified fires by analyzing the texture and color characteristics of smoke. Vicente et al. [5] developed an automated smoke detection algorithm for early forest fires that combined pixel and spectral analysis. Toreyin et al. [6] utilized wavelet transforms to compute wavelet energy based on smoke texture features, thus improving the accuracy of fire detection. This method significantly enhanced the efficiency of automatic monitoring. Huang et al. [7] developed a texture analysis algorithm using texture image classification (LBP) and a graph pyramid model that took into account the impact of mountain clouds and fog, effectively lowering the false alarm rate in forest fire detection. While these early algorithms for forest fire detection yielded satisfactory outcomes, they all required sophisticated image processing methods, leading to suboptimal generalization and robustness [8].

In recent years, significant advancements in deep learning technology for image processing have led to its widespread application in target detection. Based on the nature of their calculation methods, deep learning technologies can be classified into two categories: two-stage regression detection methods, including Faster R-CNN [9], Libra R-CNN [10], and R-FCN [11], and single-stage regression detection methods, such as RetinaNet [12], the Yolo series [13–15], and EfficientDet [16]. Building on these methods, researchers have developed numerous algorithms for detecting forest fire smoke, achieving impressive results. Zhao et al. [17] utilized Fast R-CNN as the foundational framework and introduced a forest fire smoke recognition method that integrates environmental information by constructing a multi-level region of interest pool structure. This approach significantly lowered the false alarm rate for forest fire smoke. Khan et al. [18] proposed a forest fire detection model utilizing a convolutional neural network (CNN) based on the MobileNet model. They incorporated the symmetrical feature image extracted by the fully connected layer for enhanced forest fire detection. Their model's detection accuracy was significantly improved compared to other CNN models. Pang et al. [19] introduced an anchor-free target detection algorithm based on an encoder–decoder structure. They designed the remaining effective channel attention block as the decoding unit, enhanced it using an attention-based adaptive fusion residual module, and incorporated the CA attention mechanism. This algorithm demonstrates high accuracy and recall rates, significantly improving the detection accuracy of early forest fire smoke. Qian et al. [20] combined two independent weak supervision models, Yolov5 and EfficientDet, and applied the weighted box fusion algorithm (WBF) to process the prediction results. This approach enhanced both the speed and accuracy of forest fire smoke detection. These methods have performed well on self-made or public datasets, proving to be highly significant for early forest fire detection.

Given the sudden onset of forest fires, detection equipment captures forest fire smoke targets of varying sizes. While existing models have achieved high accuracy in detecting conventional smoke targets, there is still a need to improve the detection accuracy for small smoke targets. To address the issue of low precision in detecting small smoke targets, many scholars have developed specialized algorithms for small target smoke detection. Zhao et al. [21] extended the feature extraction network of Yolov3 into three dimensions and enhanced the model's prediction performance using a feature pyramid. They proposed a target detection model specifically for small target forest fires that demonstrated significant advantages in detecting small forest fire targets. Guo et al. [22] introduced a more accurate small target detection mechanism (MASD) combined with a multi-scale feature fusion (MCF) path. This approach fused image semantic features with location information and incorporated the SimAM attention mechanism. They designed a small target detection model for UAV aerial forest fire images, significantly enhancing the detection accuracy and speed for small target forest fire smoke. Hui et al. [23] designed a small target forest fire smoke detection model that leverages the strengths of the Yolov5 structure. They incorporated a spatial graph convolution operation based on the message passing neural network mechanism and introduced the multi-head attention mechanism before each detection head of Yolov5. This model demonstrated excellent accuracy using a self-made

small target forest fire smoke dataset. These algorithms for detecting small target forest fire smoke offer high detection accuracy and practical value.

Currently, scholarly research on forest fire smoke detection primarily focuses on enhancing the accuracy of small target smoke detection, as well as improving the overall detection accuracy, speed, and anti-interference capabilities of forest fire smoke detection methods. Research on balancing the detection accuracy of forest fire smoke for targets of various sizes and reducing false positives and missed detections remains insufficient. Based on this, this paper proposes an improved algorithm based on Yolov5s called SIMCB-Yolo. This algorithm enhances the detection accuracy of smoke targets at various levels and improves the model's anti-interference capability in complex natural landscapes, such as clouds and fog, thereby reducing false positives and missed detections of forest fires. The specific contributions of this article are as follows:

- This study presents a multi-tiered forest fire smoke detection model based on Yolov5s that is designed to accurately detect small target smoke in the early stages of forest fires. It offers new insights and approaches for fire detection and monitoring.
- To improve the model's accuracy in detecting small target smoke, we incorporated the Swin transformer module into the neck of Yolov5s, blending its output feature maps into the detection layer to bolster the model's capability to extract pixel-level feature information, thereby improving its detection accuracy for small target smoke and significantly reducing missed detections.
- To balance the model's accuracy in detecting general target smoke, we incorporated a C3 block and CBS block into the main body of Yolov5s, which not only expanded the model's receptive field but also enriched the information it extracted.
- To bolster the model's resilience against complex natural backgrounds, such as mountain clouds and fog, we incorporated the SimAM attention mechanism at the end of the backbone network. This enhancement enables the model to more effectively identify regional features resembling smoke, mitigating interference from complex environmental backgrounds, and thereby significantly improving detection accuracy while reducing false positives.

2. Data and Methods

2.1. Data

Forest fire smoke recognition requires adequate data support. To this end, we utilized Python web scraping techniques to obtain 2000 images related to forest fires from the internet, along with an additional 400 images of backgrounds in forest areas. Among these images are also 100 pictures of fog that are easily confused with smoke from fires. Some samples of the dataset are shown in Figure 1.

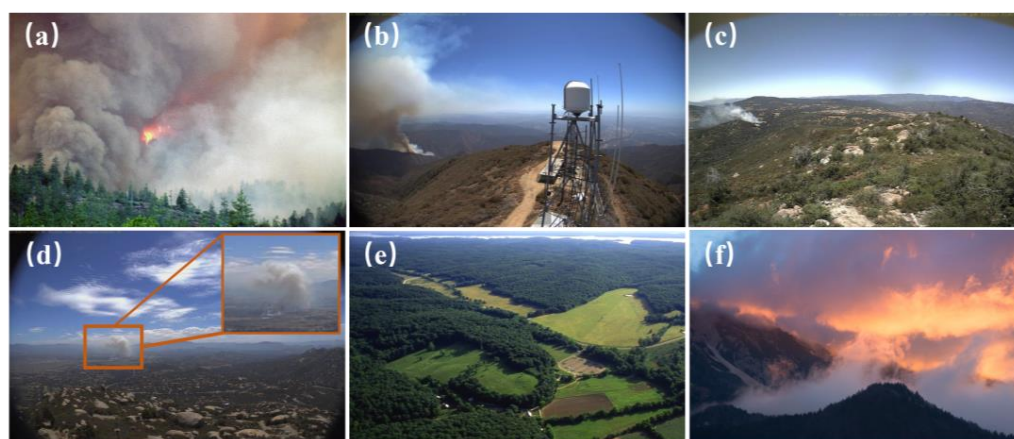


Figure 1. Samples from the smoke dataset. (a) Close-up of large-scale smoke. (b–d) Close-up, mid-range, and distant small target smoke. (e) Forest areas without fire occurrence. (f) Mountain clouds and fog that are easily confused with forest fire smoke.

To further expand the dataset, we applied data augmentation techniques, including rotation, blurring, and noise addition, to the collected images [24], aiming to simulate the complex interference backgrounds likely encountered in practical forest fire smoke detection scenarios (as show in Figure 2). The final dataset comprises 7200 images. Prior to conducting the experiment, we reviewed key literature on dataset partitioning [25,26]. We chose to divide the dataset into an 8:1:1 ratio, allocating 80% of the data for training, 10% for validation, and the remaining 10% for testing.

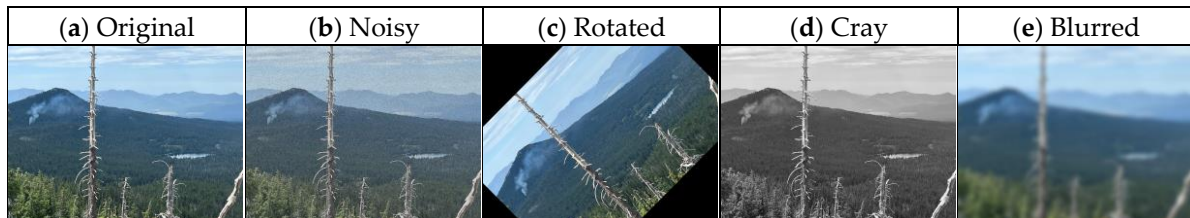


Figure 2. Schematic diagram of data augmentation. (a) Image without any operation (b) Add 10% salt and pepper noise to the image. (c) Rotate the image counterclockwise/counterclockwise by 45 degrees. (d) Grayscale the image. (e) Blur the image using a fuzzy convolution with a size of 255.

2.2. Methods

2.2.1. Yolov5s

The improvements in this article are based on the lightweight Yolov5s model from the Yolov5 series. Yolov5, developed by Ultralytics in June 2020, is a deep learning model widely used in the field of image processing [13,14]. Due to its excellent detection efficiency and accuracy, Yolov5 was widely adopted and improved upon shortly after its release. The Yolov5 model primarily comprises four components: input, backbone, neck, and prediction.

In the Yolov5s model, the input part employs adaptive image scaling to reduce interference from image noise and other irrelevant information. The backbone primarily consists of C3 [15], SPPF, and CBS [27] components. CBS, as the basic convolution unit in Yolov5s, comprises a 2D convolution layer (Conv), batch normalization layer (BN), and SiLU activation function. The C3_1_x and C3_2_x, based on CSPNet [27], are mainly used for feature fusion and downsampling, enhancing the feature extraction capabilities of the backbone while reducing the number of parameters. The spatial pyramid pooling improvement (SPPF) is located at the end of the backbone and extracts global features. It evolved from SPP but is more efficient than its predecessor. The neck uses a PANet structure to merge features extracted from the backbone with those from the detection layers, thereby obtaining richer and more detailed features. The final prediction part consists of three detection layers for object detection. A detailed diagram of the Yolov5s network structure is shown in Figure 3.

The backbone and neck shown in Figure 3 play an extremely important role in feature extraction and fusion. The backbone utilizes a deep convolutional network to extract low-level, intermediate, and high-level features from the input image. Among them, low-level features include edges and textures, intermediate features include the shape and contour of objects, and high-level features involve the complex patterns and semantic information of objects. The neck provides rich feature representation for the head by further integrating and processing these features. Specifically, the neck employs the feature pyramid network (FPN) and path aggregation network (PaNet) to achieve multi-scale feature fusion and context information aggregation. This approach enhances the model's ability to detect objects of various sizes, thereby improving detection accuracy and robustness. Through the synergy of the trunk and neck, Yolov5s can efficiently extract and utilize multi-level features from images to achieve accurate target detection.

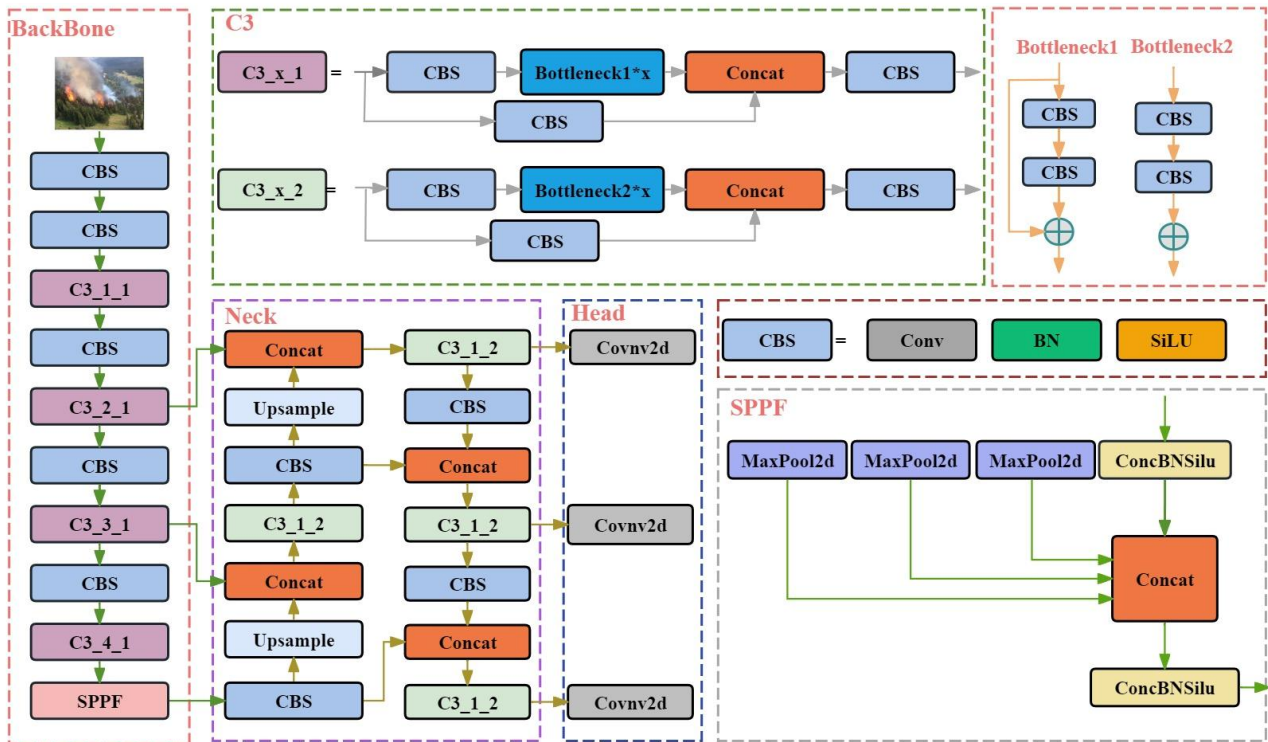


Figure 3. YOLOv5s Network architecture diagram. The whole network is divided into the backbone, neck and head. The * sign in the figure indicates that the module can be repeated many times.

2.2.2. Swin Transformer

In today’s image processing tasks, transformers play an essential role. The Swin transformer, proposed by Liu et al. in 2021, is a compact visual transformer [28] that, compared to the original transformers used in natural language processing [29], adopts a hierarchical structure to avoid the high computational overhead associated with global self-attention.

The Swin transformer block consists of four main components: a multi-layer perceptron (MLP), window-based multi-head self-attention (W_MSA), shifted window multi-head self-attention (SW_MSA), and layer normalization (LN). The backbone structure is depicted in Figure 4.

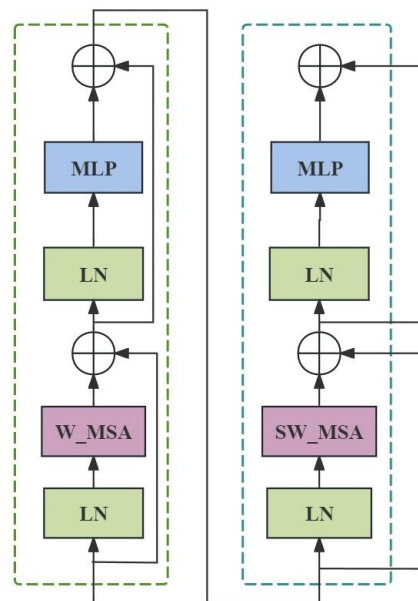


Figure 4. Swin transformer block [30].

Within the Swin transformer, input features are first normalized and subsequently processed through a shifted window attention mechanism. This process involves direct residual computations, followed by subsequent learning and residual operations within the perceptron layer and the window attention matrix layer. The computed outputs from these operations are as follows:

$$\hat{z}^l = W_MSA\left(LN\left(z^{l-1}\right)\right) + z^{l-1} \quad (1)$$

$$z^l = MLP\left(LN\left(\hat{z}^l\right)\right) + \hat{z}^l \quad (2)$$

$$\hat{z}^{l+1} = SW_MSA\left(LN\left(z^l\right)\right) + z^l \quad (3)$$

$$z^{l+1} = MLP\left(LN\left(\hat{z}^{l+1}\right)\right) + \hat{z}^{l+1} \quad (4)$$

where z^l represents the features extracted from forest fire smoke images by various modules.

In this study, we incorporate the Swin transformer as a standalone module into the neck structure of Yolov5s. By integrating this multi-head attention mechanism, we enhance the extraction of correlated features among neighboring pixels in small target smoke, thereby improving the model's capability to detect smoke from distant forest fires.

When the input feature passes through the swing transformer block shown in Figure 4, it first passes through LN and then calculates the attention score using W_MSA to achieve efficient feature extraction. The further optimized features are input into SW_MSA, and the context information exchange between features is enhanced through window offset. The enhanced features are further processed and transformed using MLP. There are residual connections behind each main module (W_MSA and MLP) to retain input information, promote gradient transmission, and ultimately achieve the purpose of improving the expression ability and training stability of the model. These steps enable the Swin transformer block to efficiently extract and fuse multi-level features, thereby enhancing the model's representation capability.

2.2.3. SimAM

Image datasets of forest fire smoke typically contain a large amount of background interference, such as sun glow and mountain fog. These factors increase the difficulty of target monitoring. When employing methods such as adding C3 and CBS blocks to deepen the depth of the network backbone to capture more features, this also makes the network susceptible to these interference signals. Moreover, adding more network layers could result in some overfitting, potentially impacting the model's effectiveness in smoke detection. To tackle this issue, we introduced the SimAM attention mechanism [31]. This mechanism constructs three-dimensional weights without adding parameters, enhances the unit's downsampling capability, and thereby improves the model's efficiency in extracting features related to the similarity between smoke instances. The relevant computation formula is as follows:

$$e_t(w_t, b_t, y, x_i) = (y_t - \hat{t})^2 + \frac{1}{M-1} \sum_{i=1}^{M-1} (y_0 - \hat{x}_i)^2 \quad (5)$$

Here, $\hat{t} = w_t + b_t$ and $\hat{x}_i = w_i x_i + b_t$ are linear transformations of t , while t and x_i represent the target neurons and other neurons within a single channel of the input features, respectively. M is the number of neurons in the channel, and w_t and b_t are the weights required for the linear transformation. The specific calculation formulas are as follows:

$$w_t = -\frac{2}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} (t - \mu t) \quad (6)$$

$$b_t = -\frac{1}{2}(t + \mu t)w_t \quad (7)$$

$$\hat{\mu} = \frac{1}{M-1} \sum_{i=1}^{M-1} x_i \quad (8)$$

$$\hat{\sigma}^2 = \frac{1}{M-1} \sum_{i=1}^{M-1} (x_i - \hat{\mu})^2 \quad (9)$$

Here, $\hat{\mu}$ and $\hat{\sigma}^2$ represent the mean and variance, respectively, calculated for the neurons in the channel, excluding t . The distribution patterns of the two follow a normal distribution. After processing with SimAM, the complete three-dimensional image weights are obtained, as illustrated in Figure 5.

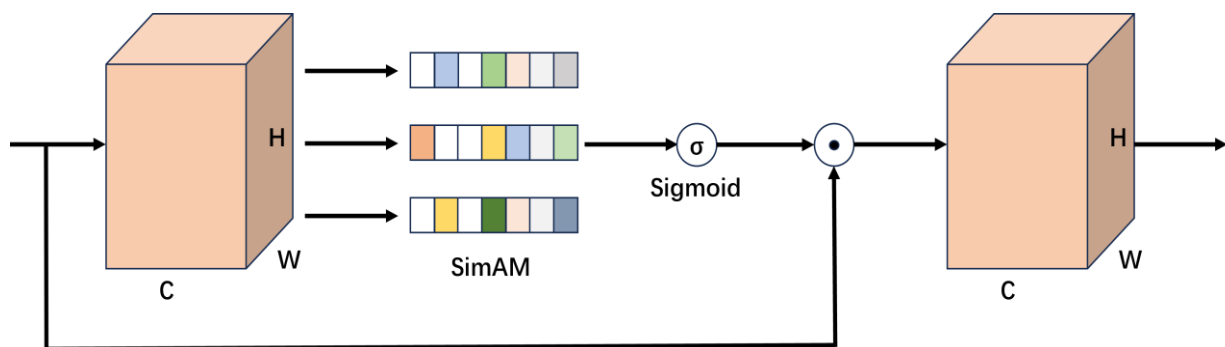


Figure 5. The structure diagram of SimAM [32].

This module converts feature x into three-dimensional weights, which are then multiplied by the target neuron's weights and the initial feature map's features to generate the final output feature map. In each subfigure, the consistent color signifies that one scalar is assigned to each channel, spatial location, or individual point within those features [31].

2.2.4. The Improved Forest Fire Smoke Recognition Model

By incorporating the Swin transformer [28] and the SimAM attention mechanism [31], we have significantly enhanced the backbone and neck of the Yolov5s model. This integration has resulted in the creation of an innovative forest fire smoke detection model, as depicted in Figure 6.

At the end of the backbone, we introduced a CBS block [27] and a C3 block [15] to enhance the extraction of smoke features. To avoid extracting excessive redundant features unrelated to forest fire smoke detection, we incorporated a SimAM attention mechanism after the final C3 block. This mechanism helps suppress irrelevant features and enhances the network's ability to extract smoke-related information, as shown in Figure 6.

In the neck, we added an upsampling and concatenation module at the entry point. This module links the feature map output from the final C3 module (eighth layer) in the backbone to the upsampled feature map, ensuring that the added C3 and CBS modules do not introduce dimensional computation errors. Prior to the final concat block, we integrated an extra concat block and upsample block to link the feature maps drawn from the newly incorporated C3 and CBS modules in the backbone to the neck. Ultimately, we added a Swin transformer block at the neck's end, using feature maps from layers 24, 27, 30, and 33 as inputs for the smoke detection layer. By preserving the original Yolov5s detection structure and incorporating the output from the Swin transformer block as new detection feature maps, the model's accuracy in detecting small target smoke is significantly enhanced.

In Figure 6, the red box highlights several crucial aspects of network enhancement: ①~② Upgrading the neck region to capture necessary feature maps for small object detection, ③ Enhancing the neck region to connect the feature maps, and ④~⑤ Increasing the

backbone capacity to extract smog features more efficiently. These enhancements have significantly bolstered the model’s capability to efficiently detect and analyze smoke features. We call this improved Yolov5s model SIMCB-Yolo.

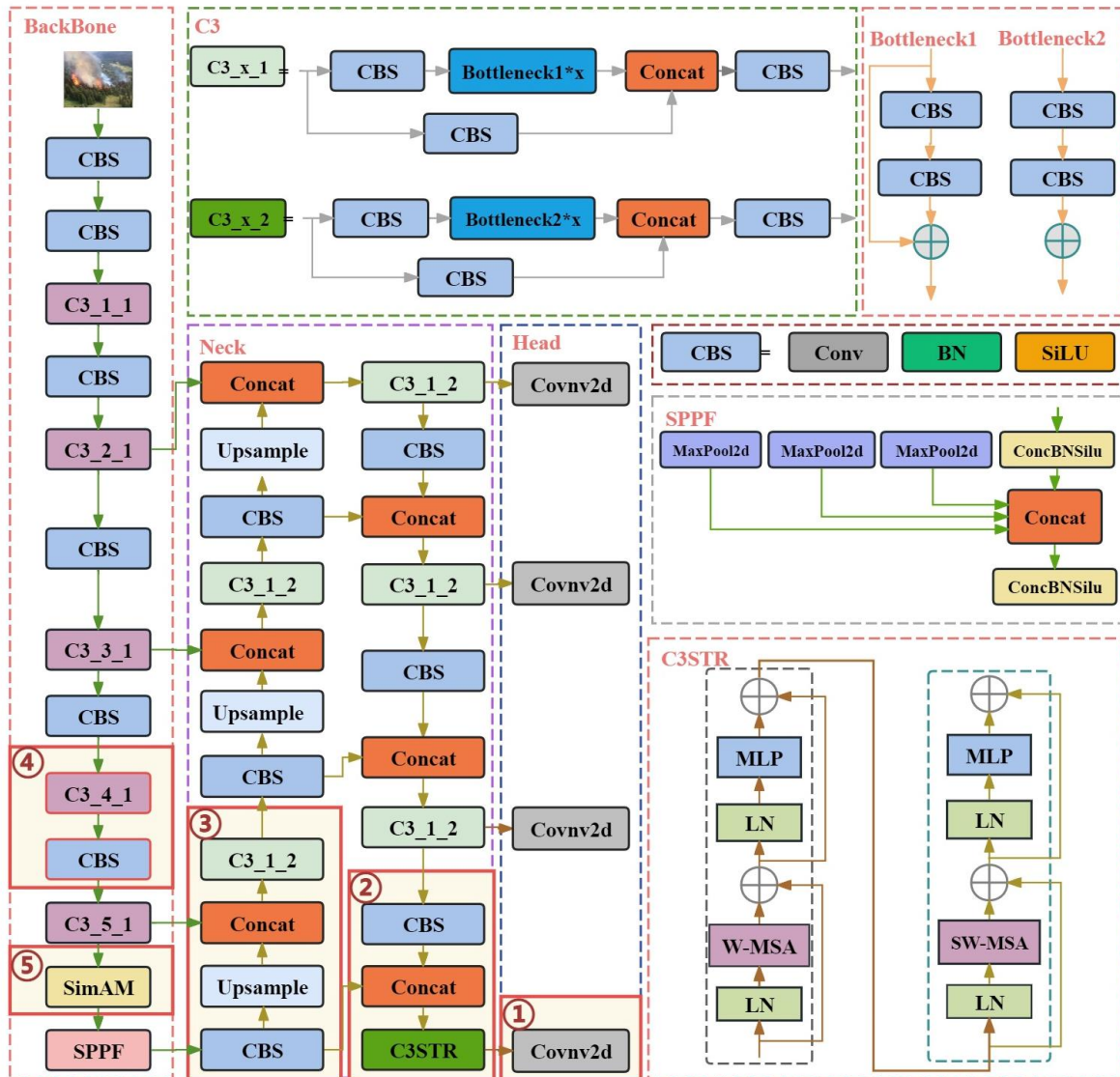


Figure 6. The improved model features several key enhancements. The * sign in the figure indicates that the module can be repeated many times.

3. Results

3.1. Implement Details

Table 1 displays the experimental setup for our work, while Table 2 lists some of the hyperparameters utilized during training. Additionally, we employed the Yolov5s model as our baseline model to assess the advantages of model enhancements throughout the training phase.

Table 1. Parameters of study resources.

Train Environment	Details
Programming language	Python3.9.18
Operating system	Centos7
Deep learning framework	Pytorch1.12.0+cu102
Running device	Tesla V100

Table 2. Initial training hyperparameters.

Training Parameters	Details
Epochs	300
Batch size	32
Image size	640
Learning rate	0.01
Optimizer	SGD
Patience	True

3.2. Evaluation Metrics

Given that Yolo series models employ intersection over union (IOU) to gauge training effectiveness, this study assesses the final recognition performance of the model using metrics such as precision (P), recall (R), mean average precision (mAP@0.5), and frames per second (FPS).

Before outlining these evaluation metrics, the study first defines a set of acronyms as shown in Table 3.

Table 3. Metrics used to evaluate the models.

Symbols	Meanings
<i>TP</i>	The count of samples where the model accurately detects smoke.
<i>FP</i>	The quantity of samples where the model incorrectly identifies background as smoke.
<i>FN</i>	The count of samples where the model identifies smoke as background.
<i>TN</i>	The quantity of samples where the model correctly identifies background.
<i>AP</i>	The area beneath the precision–recall curve, indicative of the average precision in smoke detection.

Precision represents the proportion of correctly identified smoke samples relative to the total number of samples detected as smoke by the model. This metric assesses the model's accuracy in smoke detection. The formula used to calculate precision is as follows:

$$P = \frac{TP}{FP + TP} \quad (10)$$

Recall quantifies the percentage of smoke samples accurately detected by the model out of all smoke samples, assessing the model's smoke recognition capabilities. The formula for this calculation is outlined below:

$$R = \frac{TP}{TP + FN} \quad (11)$$

mAP@0.5 represents the average detection rate when setting the IOU threshold at 0.5 and is typically derived using precision and recall. The formula for computation is outlined as follows:

$$AP = \int_0^1 P(r)dr \quad (12)$$

$$mAP = \sum_{i=1}^C \frac{AP_i}{C} \quad (13)$$

where C is the number of samples.

FPS indicates the quantity of images processed by the model per second, assessing the operational efficiency of the model. The formula for this calculation is detailed below:

$$FPS = \frac{N}{T} \quad (14)$$

where N represents the number of images, and T denotes the time required to process these images.

3.3. Impact of Different Swin Transformer Improvements

There are two methods for incorporating a Swin transformer block. One approach involves using the Swin transformer block as the main unit for feature extraction, substituting all C3 blocks within the model. The other involves incorporating the Swin transformer as a standalone module at the end of the neck, as depicted in Figure 6. To ascertain which strategy most significantly enhances detection accuracy, we conducted separate experiments for each approach and documented their impact on forest fire smoke detection accuracy. The findings are displayed in Table 4.

Table 4. Comparison of different Swin transformer improvement methods.

Improvement Methods	Precision	Recall	mAP50	mAP50-95	FPS _{bs=1}
Yolov5s	0.895	0.793	0.811	0.567	112
Yolov5s+Swin	0.859	0.784	0.836	0.594	86
Yolov5s+Swin_RE_C3	0.846	0.711	0.755	0.468	71

Table 4 shows that Yolov5s+Swin outperforms Yolov5s+Swin_RE_C3 in terms of performance metrics. While the Yolov5s+Swin shows a slight reduction in detection precision and recall compared to the benchmark model, the two critical metrics, mAP@50 and mAP@50-95, have improved by 2.5% and 2.7% respectively. We contend that integrating the Swin transformer block improves the model's capability to extract smoke features from small targets. In this scenario, the model might incorrectly identify smoke with features similar to clouds, fog, etc., as background. This misrecognition diminishes the model's capability to accurately recognize such features, consequently impacting the overall precision and recall of the model. However, mAP@50 and mAP@50-95 shows that the object detection ability of the model has been enhanced, and the recognition accuracy of multi-level forest fire smoke has been improved. It should be noted that although the decrease in precision and recall may appear unfavorable, in application, mAP is often seen as a more comprehensive and important evaluation indicator because it comprehensively considers the performance of precision and recall. Therefore, although there has been a decrease in certain indicators, the improvement of Yolov5s+Swin in mAP indicates an overall improvement in its performance in small target smoke detection tasks.

3.4. Impact of Different Attention Mechanisms

In order to verify whether the precision and recall reduction of Yolov5s+Swin are indeed due to the insufficient ability to extract smoke features from large targets, we made targeted improvements to the backbone of the model. We incorporated the C3 block and CBS block to bolster the model's capacity to extract smoke features from large targets and introduced an attention mechanism to enhance the model's resilience against interference from cloud and fog backgrounds. We designed and implemented experiments to combine several commonly used attention mechanisms with the benchmark model Yolov5s to explore their impact on model performance. The results are shown in Table 5.

Table 5. Comparative analysis of various attention mechanisms.

Improvement Methods	Precision	Recall	mAP50	mAP50-95	FPS _{bs=1}
Yolov5s	0.895	0.793	0.811	0.567	112
Yolov5s+SE	0.884	0.753	0.797	0.545	113
Yolov5s+CA	0.903	0.794	0.816	0.570	116
Yolov5s+SimAM	0.892	0.764	0.798	0.577	111
Yolov5s+CBAM	0.906	0.753	0.810	0.560	114

From Table 5, it is evident that incorporating the squeeze and excitation (SE) [33] and convolutional block attention module (CBAM) [34] attention mechanisms did not substantially

enhance the model's detection performance, though there was a minor increase in processing speed. The detection performance of models with added coordinate attention (CA) [35] and SimAM Attention has been improved. This proves that these two attention mechanisms have a positive effect on model improvement. Therefore, when selecting attention mechanisms, priority should be given to adding CA and SimAM attention mechanisms.

Figure 7 depicts how the model's focus shifts in the task of detecting forest fire smoke under the influence of different attention mechanisms. By observing the changes in detection focus after integrating different attention mechanisms into the model, we can more intuitively understand the role of these mechanisms in enhancing the model's multi-level forest fire smoke detection accuracy.

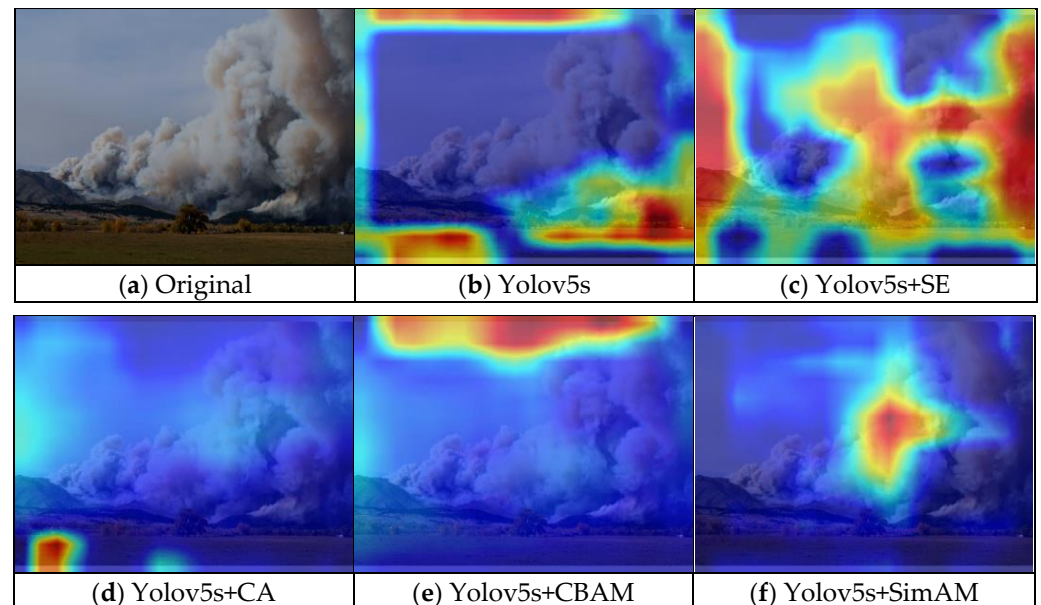


Figure 7. Focus of attention of different attention mechanisms.

As shown in Figure 7, introducing the SE attention mechanism or SimAM attention mechanism into the model results in the focus of attention being highly consistent with the distribution of smoke. In particular, the model using SE attention mechanism shows a wider range of areas of concern. However, in the detection task involving minimal targets, such a wide range of areas of concern may lead to inaccurate positioning of minimal targets [36]. On the other hand, the model with CA or CBAM attention mechanism focuses on the edge of the image or the edge of the smoke. This primarily occurs because these mechanisms focus on vital information using channel attention and employ spatial attention to accentuate key spatial regions on the feature map, thereby overstating the significance of image and target edge detection. Consequently, in this study, we opted for the SimAM attention mechanism to more precisely extract characteristics of small target smoke.

3.5. Ablation Experiment

In the preceding sections, we structured experiments from two perspectives: enhancing the neck network and upgrading the backbone network. This was done to thoroughly investigate the model's detection performance on both large and small smoke targets. Table 4 indicates that enhancing the neck of Yolov5s with Swin improves the model's capability to detect small target smoke. Table 5 reflects that the improvement of the backbone and the attention mechanism combination effectively enhance the detection effect on large target smoke. However, combining the experimental results from Tables 4 and 5, it can be inferred that Yolov5+Swim+SimAM and Yolov5+Swim+CA may achieve excellent model performance. However, it cannot be ruled out that performance degradation may occur when two modules are combined. Therefore, we further designed ablation experiments. Four design schemes were obtained by combining four attention mechanisms

with the Yolo5s+Swin model. Subsequently, we will compare these four schemes with the benchmark module Yolo5s. The experimental results are shown in Table 6.

Table 6. Ablation experiment comparison results.

Improvement Methods	Precision	Recall	mAP50	mAP50-95	FPS _{bs=1}
Yolo5s	0.895	0.793	0.811	0.567	112
Yolo5s+SE+Swin	0.856	0.763	0.803	0.58	87
Yolo5s+CA+Swin	0.886	0.773	0.815	0.592	73
Yolo5s+SimAM+Swin	0.915	0.794	0.856	0.636	85
Yolo5s+CBAM+Swin	0.827	0.739	0.796	0.594	80

From Table 6, it can be seen that among the four alternative schemes, the Yolo5s+SimAM+Swin scheme has higher values in all other indicators except for slightly lower FPS (frame rate per second), which fully indicates that this scheme is the best combination choice. Compared with the benchmark model Yolo5s, although its FPS has decreased, it shows superiority in other indicators. Specifically, mAP@50 increased by 4.5%, while mAP@50-95 increased by 6.9%, significantly improving the model's smoke detection performance.

In order to observe the dynamic changes in smoke detection performance of the benchmark model Yolo5s and the combination schemes Yolo5s+SimAM+Swin and Yolo5s+CA+Swin, we plotted the relationship between the number of iterations and the mAP50 index, as shown in Figure 8.

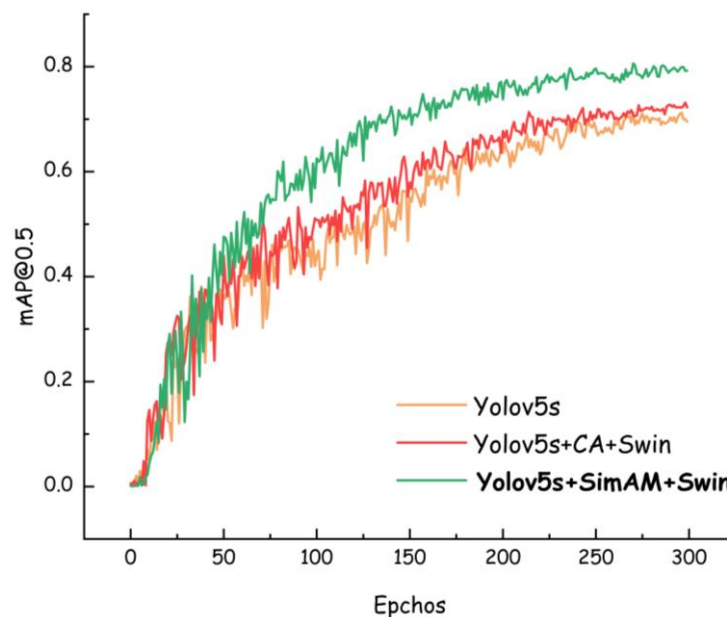


Figure 8. Comparison of ablation experiment trends in mean average precision (mAP@0.5).

From Figure 8, it can be seen that the performance indicators of the Yolo5s, Yolo5s+CA+Swin, and Yolo5s+SimAM+Swin models during forest fire smoke detection are as follows: mAP@50 values of all models show an upward trend with the increase in iteration times. Among them, the performance of the Yolo5s and Yolo5s+CA+Swin models is basically similar. However, when the number of iterations reaches 140 or more, the Yolo5s+SimAM+Swin model exhibits significant advantages. This result fully demonstrates that the Yolo5s+SimAM+Swin model performs the best in detecting multi-level large, medium, and small target smoke with natural landscape interference background and can meet the needs of our fire smoke detection targets.

3.6. Comparison

To further evaluate the performance of the SIMCB-Yolo model, we designed a comparative experiment. We compared the improved SIMCB-Yolo model with five current advanced target detection models, including YOLOv5s, YOLOv8s, SSD, YOLOv4 tiny, and Faster R-CNN, to explore the differences in detection effectiveness for various levels of forest fire smoke. The comparison results are presented in Table 7.

Table 7. Comparison of forest fire smoke detection effects of different models.

Model	mAP50	mAP50-95	FPS _{bs=1}
YOLOv5s	0.811	0.567	112
YOLOv8s [37]	0.728	0.411	101
SSD [38]	0.773	0.423	103
YOLOv4 tiny [15]	0.703	0.416	88
Faster R-CNN [9]	0.847	0.502	54
SIMCB-Yolo	0.856	0.636	85

The results in Table 7 show that the SIMCB-Yolo model outperforms these existing models in terms of average recognition accuracy mAP50 and mAP50-95. In particular, the mAP50-95 index is significantly better than other models; for example, it is 22.5% higher than the YOLOv8s model. This shows that the SIMCB-Yolo model has high detection accuracy.

In order to more intuitively compare and display the differences between the SIMCB-Yolo model and the existing advanced target detection model, we draw the focus heat map of the model that performs better in the detection and recognition comparative experiments, so as to observe the detection effect of different models (Figure 9).

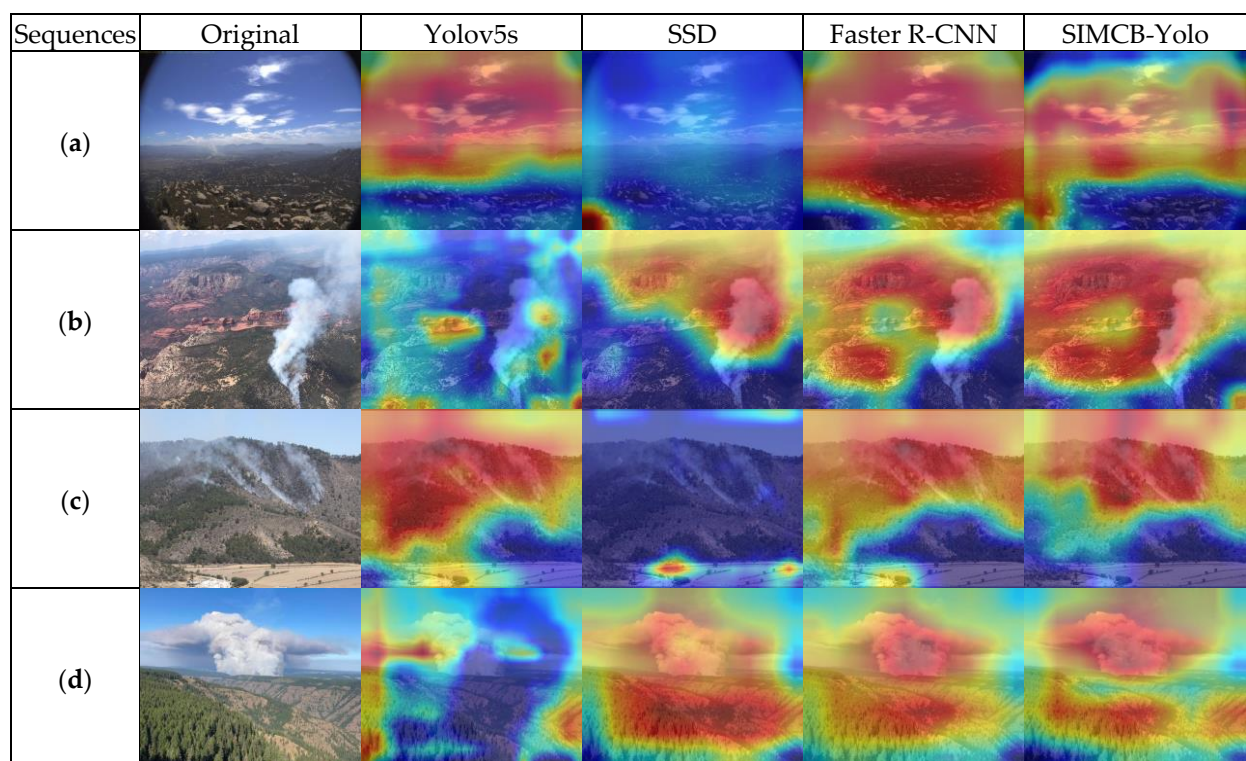


Figure 9. Heat maps of different model detection concerns. From (a–d), the target of forest fire smoke continues to increase, and the range of smoke continues to expand.

Figure 9 illustrates how each model's detection focus on forest fire smoke shifts under varying smoke target conditions. Line (a) clearly shows that when the detection target is very small, the SSD model's focus does not encompass the smoke location, whereas the

SIMCB-Yolo model, YOLOv5s, and Faster R-CNN models exhibit similar detection focuses, completely covering the forest fire smoke location. It is worth noting that with the gradual increase in the size of the detection target, the focus of each model also began to change, and there are varying degrees of errors between the focus coverage and the range of the actual target. Conversely, the SIMCB-Yolo model effectively covers detected targets across various levels in forest fire smoke detection, further underscoring its outstanding performance in identifying forest fire smoke at multiple levels.

To further assess the performance of the SIMCB-Yolo model, we selected three types of forest fire smoke images from real-world scenes that are closely related to our improvements for detection. We subsequently compared the results with those from YOLOv5s, SSD, Fast R-CNN, and other models. The results are shown in Figure 10.

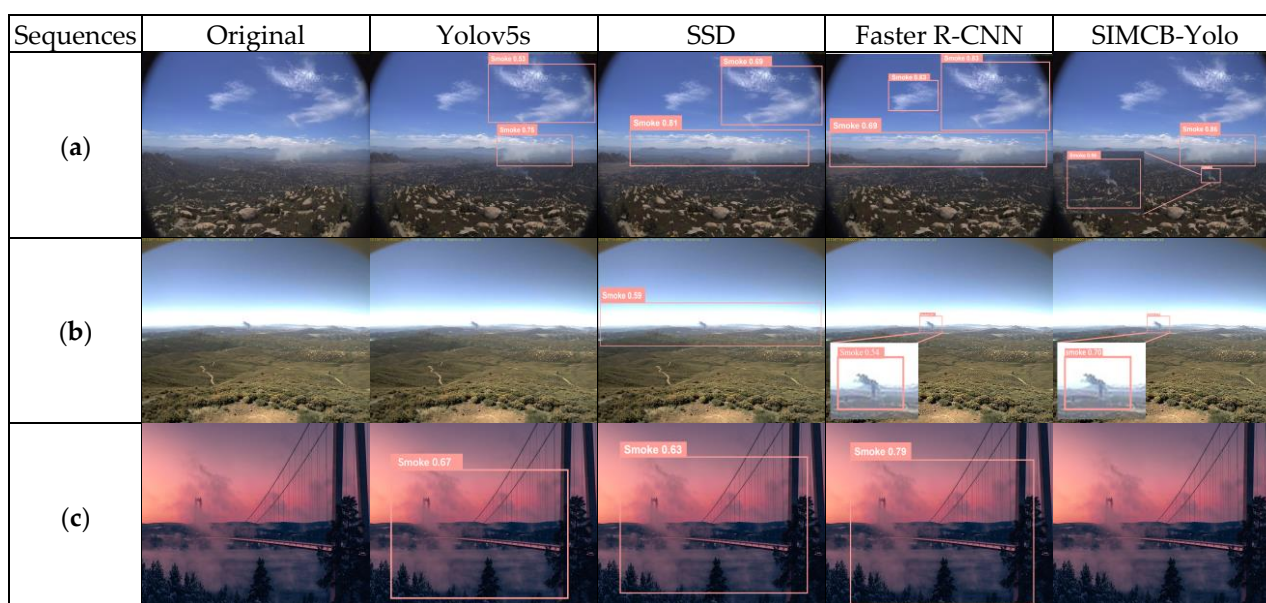


Figure 10. The detection effect of different models in the actual scene. (a) Forest fire smoke detection of different sizes under cloud interference. (b) Small target smoke detection with light interference. (c) Smoke detection under the interference of natural landscape similar to smoke.

As can be seen from Figure 10, the four models show significant differences in the detection effect of forest fire smoke in the actual scene. In Figure 10a, the YOLOv5s, SSD and Faster R-CNN models mistakenly recognize clouds as smoke, while the SIMCB-Yolo model accurately eliminates the interference of clouds. In Figure 10b, the YOLOv5s model missed the detection of small target smoke. Although the SSD model detected small targets, there was a large error between the marked forest fire smoke location and the actual forest fire location. Although the Fast R-CNN model also detects small target smoke, its confidence is lower than that of the SIMCB-Yolo model. In Figure 10c, we selected natural landscape images that were very similar to forest fire smoke for testing. The results showed that the YOLOv5s, SSD and Faster R-CNN models misjudged mist as forest fire smoke, while the SIMCB-Yolo model was not affected by this interference. By comparing the detection effect in the actual scene, the superior performance of the SIMCB-Yolo model in the detection of different levels of forest fire smoke detection is further highlighted.

4. Discussion

Major forest fires often originate from small sources in their early stages, making precise and rapid detection crucial [26]. Given the damp conditions and dense vegetation in forests, combustibles burn inefficiently in the early stages of a fire, resulting in only a small amount of smoke. Furthermore, the detection of this smoke is challenging, compounded by the interference of natural features like clouds and fog. Consequently, the task of early smoke

detection is highly challenging. Traditional methods, such as ground patrols, observation towers, and satellite remote sensing, have limitations in terms of monitoring accuracy, real-time capabilities, and coverage. In recent years, deep learning's outstanding performance in image feature extraction has increasingly been utilized in forest fire monitoring.

Common deep learning algorithms are categorized into two-stage and single-stage algorithms. Consequently, two approaches for forest fire image detection have emerged. Two-stage monitoring models like Faster R-CNN, with their large parameter counts and lengthy computation times, are not conducive to real-time detection tasks. On the other hand, single-stage models like the Yolo series excel in real-time performance but suffer from somewhat lower accuracy [15]. Therefore, the choice of model often depends on the specific needs of the task.

To improve the accurate detection of smoke, flames, and other targets in forest fires and to enhance early warning capabilities, scholars have developed various deep learning models focused on forest fire smoke imagery. These models aim to address the detection challenges present in real-world scenarios [39–41].

Given the issues of false positives stemming from the resemblance of natural landscapes like clouds and fog to smoke in open environments and the challenges of missed detections due to the scarcity of features in small targets within the context of image hierarchy, this study has implemented improvements using YOLOv5s as the benchmark model.

(1) To address the issue of missed detection of smaller target smoke in the benchmark YOLOv5s model, the Swin transformer module was integrated into the neck region of the model in this study. This enhancement improved the model's accuracy in detecting small forest fire smoke targets; however, we observed a decline in its precision and recall performance. The reason for this is that under such conditions, the model might mistakenly classify smoke with similar features, such as clouds and fog, as the background, leading to a lack of recognition for these features and subsequently impacting the model's overall accuracy and recall. (2) To address this issue, a CBS block and C3 block were introduced into the main body of the network to further extract features of smoke at different levels of the image. This, however, also led to an increase in the extraction of information from natural landscape interference factors such as clouds and fog. Consequently, a SimAM attention mechanism was added to the end of the network backbone to mitigate the impact of these interference factors on model performance and, to a certain extent, mitigate false detections due to interference factors. The ablation experiment results demonstrate that the model's overall performance was enhanced. This, in turn, partially resolves the issue of small target smoke detection in the presence of natural landscape interference, such as clouds and fog.

While the SIMCB-YOLO has significantly enhanced accuracy, it exhibits a decrease in FPS. Moving forward, we will delve deeper into methods to elevate the model's FPS performance. Lightweighting the model further by integrating strategies such as knowledge distillation [42,43], model pruning [44,45], and lowrank decomposition [46,47] is also a focus of our research. Furthermore, forest fire smoke monitoring tasks in natural environments present highly intricate scenarios, making dataset construction a crucial aspect. There are shortcomings in datasets utilized in existing studies. Consequently, in subsequent research, we will also address the issue of enhancing dataset quality.

SIMCB-YOLO remains in the experimental phase, and, in the future, we will contemplate how to integrate it with electronic sensing devices, such as small servers or drones, to facilitate the transition from theoretical research to practical applications.

5. Conclusions

Forest fires are characterized by their suddenness, destructiveness, and danger [48]. Once they break out, they pose a significant threat to human society [49]. Therefore, it is crucial to detect and respond promptly to forest fires in their early stages.

Given that smoke often serves as a crucial indicator for early warning of forest fires, developing models using deep learning technology for detecting forest fire smoke holds

significant practical importance [50]. However, images captured in open environments are frequently disturbed by natural landscapes such as clouds and fog, inevitably leading to false positives in model detection. Additionally, the initial fine smoke, due to its small size and unclear features, also makes detection more challenging, leading to missed detections. To address these issues, this paper proposes an enhanced Yolov5s model, SIMCB-Yolo, based on the concept of image hierarchy. The goal is to enhance the recognition accuracy of multi-level forest fire smoke in images disturbed by natural landscapes, while minimizing the occurrence of false positives and missed detections. This model aims to develop a robust and well suited smoke detection model for forest fires, thereby improving the accuracy and reliability of forest fire warnings.

Specifically, we added a Swin transformer block to the neck of the Yolov5s model to improve its accuracy in detecting small target smoke. Following this, we introduced an additional C3 block and CBS block to the backbone to augment its feature extraction capabilities and acquire more detailed image feature information. Ultimately, the SimAM attention mechanism was integrated at the end of the backbone network to boost the model's ability to resist interference from natural elements like clouds and fog.

Relative to the Yolov5s model, the enhanced model demonstrates a notable improvement in detection accuracy, with increases of 4.5% in mAP@50 and 6.9% in mAP@50-95. Furthermore, the model's robustness and generalization capabilities are significantly improved. While the detection speed of the enhanced model may theoretically decrease, it stays within an acceptable limit, rendering it appropriate for use on small devices for real-time forest fire monitoring. The natural landscape in open environments is extremely complex, and the information in smoke images is highly variable, making forest fire smoke detection a highly challenging problem with vast application potential. In future research, we will conduct practical application tests using SIMCB-Yolo.

Author Contributions: Conceptualization, W.Y. and Y.S.; methodology, W.Y.; software, W.Y.; validation, W.Y., Z.Y. and Y.S.; formal analysis, Y.Z.; investigation, G.Z. and Y.S.; resources, Y.S.; data curation, W.Y.; writing—original draft preparation, W.Y.; writing—review and editing, Y.S. and Y.Z.; visualization, M.W.; supervision, Y.S.; project administration, G.Z.; funding acquisition, G.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China under Grant 32271879.

Data Availability Statement: The data presented in this study are available upon request from the first author. The dataset and code cannot be shared due to specific reasons.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Li, Y.; Mei, B.; Linhares-Juvenal, T. The economic contribution of the world's forest sector. *Forest Policy Econ.* **2019**, *100*, 236–253. [[CrossRef](#)]
2. Sahoo, G.; Wani, A.; Rout, S.; Sharma, A.; Kar, S.; Prusty, A. Impact and Contribution of Forest in Mitigating Global Climate Change. *Des. Eng.* **2021**, *4*, 667–682.
3. Arteaga, B.; Diaz, M.; Jojoa, M. Deep Learning Applied to Forest Fire Detection. In Proceedings of the 2020 IEEE International Symposium on Signal Processing and Information Technology, Abu Dhabi, United Arab Emirates, 7–10 December 2020.
4. Lin, Q.; Li, Z.; Zeng, K.; Fan, H.; Li, W.; Zhou, X. Fire Match: A semi-supervised video fire detection network based on consistency and distribution alignment. *Expert Syst. Appl.* **2024**, *248*, 123409. [[CrossRef](#)]
5. Vicente, J.; Guillemant, P. An image processing technique for automatically detecting forest fire. *Int. J. Therm. Sci.* **2002**, *41*, 1113–1120. [[CrossRef](#)]
6. Toreyin, B.U.; Dedeoglu, Y.; Cetin, A.E. Contour based smoke detection in video using wavelets. In Proceedings of the European Signal Processing Conference, Florence, Italy, 4–8 September 2006; pp. 1–5.
7. Huang, J.; Zhao, J.; Gao, W.; Long, C.; Xiong, L.; Yuan, Z.; Han, S. Local Binary Pattern Based Texture Analysis for Visual Fire Recognition. In Proceedings of the 2010 3rd International Congress on Image and Signal Processing, Yantai, China, 16–18 October 2010; pp. 1887–1891.
8. Xiao, G.; Yichao, C.; Tongxin, H. An Efficient and Lightweight Detection Model for Forest Smoke Recognition. *Forests* **2024**, *15*, 210. [[CrossRef](#)]

9. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)]
10. Pang, J.; Chen, K.; Shi, J.; Feng, H.; Ouyang, W.; Lin, D. Libra R-CNN: Towards Balanced Learning for Object Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 821–830.
11. Dai, J.; Li, Y.; He, K.; Sun, J. R-FCN: Object Detection via Region-Based Fully Convolutional Networks. In Proceedings of the NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; pp. 379–387.
12. Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. *IEEE Trans. Pattern. Anal. Mach. Intell.* **2020**, *42*, 318–327. [[CrossRef](#)] [[PubMed](#)]
13. Ultralytics-YOLOv5. Available online: <https://github.com/ultralytics/YOLOv5> (accessed on 2 June 2024).
14. Joseph, R.; Santosh, D.; Ross, G.; Ali, F. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
15. Bochkovskiy, A.; Wang, C.-Y.; Liao, H.-Y.M. YOLOv4: Optimal Speed and Accuracy of Object Detection. *arXiv* **2020**, arXiv:2004.10934.
16. 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, Seattle, WA, USA, 13–19 June 2020; pp. 10781–10790.
17. Zhao, E.; Liu, Y.; Zhang, J.; Tian, Y. Forest Fire Smoke Recognition Based on Anchor Box Adaptive Generation Method. *Electronics* **2021**, *10*, 566. [[CrossRef](#)]
18. Khan, S.; Khan, A. FFireNet: Deep Learning Based Forest Fire Classification and Detection in Smart Cities. *Symmetry* **2022**, *14*, 2155. [[CrossRef](#)]
19. Pang, Y.; Wu, Y.; Yuan, Y. FuF-Det: An Early Forest Fire Detection Method under Fog. *Remote Sens.* **2023**, *15*, 5435. [[CrossRef](#)]
20. Qian, J.; Lin, H. A Forest Fire Identification System Based on Weighted Fusion Algorithm. *Forests* **2022**, *13*, 1301. [[CrossRef](#)]
21. Zhao, L.; Zhi, L.; Zhao, C.; Zheng, W. Fire-YOLO: A Small Target Object Detection Method for Fire Inspection. *Sustainability* **2022**, *14*, 4930. [[CrossRef](#)]
22. Guo, J.; Liu, X.; Bi, L.; Liu, H.; Lou, H. UN-YOLOv5s: A UAV-Based Aerial Photography Detection Algorithm. *Sensors* **2023**, *23*, 5907. [[CrossRef](#)] [[PubMed](#)]
23. Yuan, H.; Lu, Z.; Zhang, R.; Li, J.; Wang, S.; Fan, J. An effective graph embedded YOLOv5 model for forest fire detection. *Comput. Intell.* **2024**, *40*, e12640. [[CrossRef](#)]
24. Fraser, D.; Schowengerdt, R.A. Avoidance of additional aliasing in multipass image rotations. *IEEE Trans. Image Process.* **1994**, *3*, 6. [[CrossRef](#)] [[PubMed](#)]
25. Zhan, J.; Hu, Y.; Zhou, G.; Wang, Y.; Cai, W.; Li, L. A high-precision forest fire smoke detection approach based on ARGNet. *Comput. Electron. Agric.* **2022**, *196*, 106874. [[CrossRef](#)]
26. Xiao, Z.; Wan, F.; Lei, G.; Xiong, Y.; Xu, L.; Ye, Z.; Liu, W.; Zhou, W.; Xu, C. FL-YOLOv7: A Lightweight Small Object Detection Algorithm in Forest Fire Detection. *Forests* **2023**, *14*, 1812. [[CrossRef](#)]
27. Wang, C.-Y.; Liao, H.-Y.M.; Yeh, I.-H.; Wu, Y.-H.; Chen, P.-Y.; Hsieh, J.-W. CSPNet: A New Backbone that can Enhance Learning Capability of CNN. *arXiv* **2019**, arXiv:1911.11929.
28. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows. *arXiv* **2021**, arXiv:2103.14030.
29. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 2.
30. Chollet, F. *Deep Learning with Python*, 2nd ed.; Manning Publications: Shelter Island, NY, USA, 2021; pp. 298–310.
31. Wang, Y.; Sun, Q.; Liu, Z.; Tan, T. SimAM: A Simple but Effective Attention Module for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 6292–6301.
32. Fu, G. *Artificial Intelligence Attention Mechanism: System, Model, and Algorithm Analysis*; Mechanical Industry Press: Beijing, China, 2024; pp. 128–150.
33. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
34. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. CBAM: Convolutional Block Attention Module. In *Computer Vision—ECCV 2018*; Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2018; Volume 11211.
35. 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, Nashville, TN, USA, 20–25 June 2021; pp. 13713–13722.
36. Ge, Q.; Li, J.; Wang, X.; Deng, Y.; Zhang, K.; Sun, H. LiteTransNet: An interpretable approach for landslide displacement prediction using transformer model with attention mechanism. *Eng. Geol.* **2024**, *331*, 107446. [[CrossRef](#)]
37. Jocher, G.; Chaurasia, A.; Qiu, J. Ultralytics YOLO (Version 8.0.0) [Computer Software]. 2023. Available online: <https://github.com/ultralytics/ultralytics> (accessed on 15 March 2024).
38. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.-Y.; Berg, A.C. SSD: Single Shot MultiBox Detector. *arXiv* **2016**, arXiv:1512.02325.

39. Li, R.; Hu, Y.; Li, L.; Guan, R.; Yang, R.; Zhan, J.; Cai, W.; Wang, Y.; Xu, H.; Li, L. SMWE-GFPNet: A high-precision and robust method for forest fire smoke detection. *Knowl. Based Syst.* **2024**, *289*, 111528. [[CrossRef](#)]
40. Yang, X.; Hua, Z.; Zhang, L.; Fan, X.; Zhang, F.; Ye, Q.; Fu, L. Preferred vector machine for forest fire detection. *Pattern Recognit.* **2023**, *143*, 109722. [[CrossRef](#)]
41. Xue, Z.; Lin, H.; Wang, F. A small target forest fire detection model based on YOLOv5 improvement. *Forests* **2022**, *13*, 1332. [[CrossRef](#)]
42. Hinton, G.; Vinyals, O.; Dean, J. Distilling the knowledge in a neural network. *arXiv* **2015**, arXiv:1503.02531.
43. Zhou, M.; Wu, L.; Liu, S.; Li, J. UAV forest fire detection based on lightweight YOLOv5 model. *Multimed. Tools Appl.* **2023**, *2*. [[CrossRef](#)]
44. Jiang, Y.; Wang, S.; Valls, V.; Ko, B.J.; Lee, W.-H.; Leung, K.K. Model Pruning Enables Efficient Federated Learning on Edge Devices. *IEEE Trans. Neural Netw. Learn. Syst.* **2023**, *34*, 12. [[CrossRef](#)]
45. Zheng, Y.; Sun, P.; Ren, Q.; Xu, W.; Zhu, D. A novel and efficient model pruning method for deep convolutional neural networks by evaluating the direct and indirect effects of filters. *Neurocomputing* **2024**, *569*, 127124. [[CrossRef](#)]
46. Peng, Y.; Ganesh, A.; Wright, J.; Xu, W.; Ma, Y. RASL: Robust Alignment by Sparse and Low-Rank Decomposition for Linearly Correlated Images. *IEEE Trans. Pattern Anal. Mach. Intell.* **2012**, *34*, 11.
47. Ong, F.; Lustig, M. Beyond Low Rank + Sparse: Multiscale Low Rank Matrix Decomposition. *IEEE J. Sel. Top. Signal Process.* **2016**, *10*, 4. [[CrossRef](#)] [[PubMed](#)]
48. Ying, L.; Han, J.; Du, Y.; Shen, Z. Forest fire characteristics in China: Spatial patterns and determinants with thresholds. *For. Ecol. Manag.* **2018**, *424*, 345–354. [[CrossRef](#)]
49. Kumar, A. Preserving life on earth. In *Adaptation, Ecosystem-Based*; Elsevier: Amsterdam, The Netherlands, 2022; pp. 105–111.
50. Šerić, I.; Stipančević, D.; Štula, M. Observer network and forest fire detection. *Inf. Fusion* **2011**, *12*, 160–175. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.