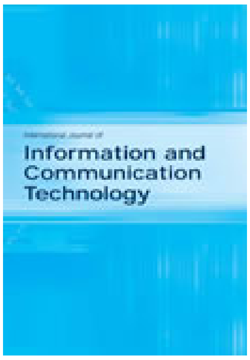


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

A wildfire smoke detection based on improved YOLOv8

Jieyang Zhou, Yang Li, Pengfei Yin

Article History:

Received:	13 June 2024
Last revised:	30 July 2024
Accepted:	07 August 2024
Published online:	12 September 2024

A wildfire smoke detection based on improved YOLOv8

Jieyang Zhou, Yang Li and Pengfei Yin*

College of Computer Science and Engineering,
Jishou University,
Jishou, 416000, China
Email: 2021404285@stu.jsu.edu.cn
Email: 2022405571@stu.jsu.edu.cn
Email: pppypf@163.com
*Corresponding author

Abstract: A novel FPN network architecture, which designed to modify and improve the performance of the original YOLOv8 mode to overcome the challenges associated with diminished detection accuracy and sluggish wildfire smoke detection, is introduced in this paper. The architecture integrates the prediction layer, downsampled four times in its feature map size from the original image, in tandem with ODConv. These augmentations aim to improve the network's feature fusion capability and predictive proficiency. Initially, the model replaces the traditional convolution in the intermediate layer with ODConv, resulting in a substantial performance enhancement. Acknowledging the non-rigid nature of smoke and the considerable variation in target size, especially prevalent in real-world settings with smaller targets, the addition of the prediction layer, subsampled four times from the original image's feature map size, enhances the model's ability to capture shallow feature information. Experimental verification underscores the efficacy of the improved YOLOv8 in smoke detection, demonstrating the precision attains 91.37% while recall improves to 87.67% and the mean average precision (mAP) of 95.18% for mAP50 and 67.43% for mAP50-95, signifying enhancements of 1.29%, 4.61%, 2.15%, and 3.86% compared to the original YOLOv8.

Keywords: wildfire detection; wildfire; ODConv; YOLOv8; deep learning; mean average precision; mAP.

Reference to this paper should be made as follows: Zhou, J., Li, Y. and Yin, P. (2024) 'A wildfire smoke detection based on improved YOLOv8', *Int. J. Information and Communication Technology*, Vol. 25, No. 6, pp.52–67.

Biographical notes: Jieyang Zhou is an undergraduate majoring in Data Science and Big Data Technology in 2021 from College of Computer Science and Engineering, Jishou University.

Yang Li is an undergraduate majoring in Data Science and Big Data Technology in 2022 from College of Computer Science and Engineering, Jishou University.

Pengfei Yin is a Doctor of Philosophy, Assistant Researcher and Deputy Dean of the Graduate School of Jishou University, postgraduate tutor. His main research direction is the recommendation algorithm and the search algorithm.

1 Introduction

Over the past several years, the escalating global warming trend has given rise to increasingly conspicuous phenomena, most notably El Niño and La Niña. These climatic developments have led to a notable escalation in the frequency, intensity, and duration of various extreme weather events, including but not limited to severe droughts and extreme heatwaves, on a global scale. The comprehensive statistical data jointly disseminated by reputable organisations such as the International Federation of Red Cross and Red Crescent Societies (IFRC) through the United Nations Office for Disaster Risk Reduction (UNDRR) reveals that from 2020 to 2022, the world experienced a total of forty wildfire disasters. These calamitous events have exacted a significant human toll, resulting in 274 recorded fatalities, and profoundly impacting nearly 1.1 million individuals. Furthermore, the economic repercussions of these disasters have been nothing short of staggering, resulting in cumulative losses amounting to a remarkable 21.502 billion US dollars (IFRC, 2020; Ministry of Emergency Management, Beijing Normal University, National Disaster Reduction Centre of China, IFRC, 2022; Academy of Disaster Reduction and Emergency Management, National Disaster Reduction Centre of China, IFRC, Beijing Normal University, 2023).

Research conducted by Zheng's team at the Graduate School of Shenzhen, Tsinghua University, indicates a progressive increase in the prevalence of extreme wildfires. When wildfires encroach upon carbon-rich ecosystems, such as peatlands and forests, which serve as significant carbon sinks, they not only result in substantial carbon emissions but also trigger severe natural disasters, including peatland fires, deforestation, and forest degradation. This, in turn, obstructs the complete absorption of carbon released during the wildfire combustion process and may impede the swift recovery and the restoration of ecosystems. Consequently, it diminishes the carbon sequestration capacity of terrestrial ecosystems and further obstructs agricultural development (Zheng et al., 2023). As a result, the prevention, early detection, and containment of wildfires have consistently held a strategic position in the development of public infrastructure in numerous countries. In cases of wildfire outbreaks, smoke typically precedes the appearance of open flames and can serve as an early indicator of fire incidents (Hu et al., 2022; Mukhiddinov et al., 2022; Avazov et al., 2021).

Although research in the domain of wildfire smoke detection in natural environments has affirmed the effectiveness of diverse detection models and yielded positive results, the complex attributes of outdoor settings and the intricacies involved in feature extraction of wildfire smoke have presented challenges in early detection. These challenges encompass issues related to elements like clouds, water surfaces, and fog. These variables introduce difficulties in distinguishing smoke and are further compounded by variations in natural lighting, causing alterations in image attributes that impact sub-sequent processes of feature extraction and recognition. This paper introduces advanced enhancements rooted in the YOLO model (Redmon et al., 2016) for wildfire smoke monitoring. The approach, which entails the utilisation of pre-trained weights as fundamental parameters for the base network, followed by the fine-tuning of network architecture parameters to enhance the efficiency of the traditional YOLOv8, and by incorporating this refined network architecture into datasets associated with wildfire smoke, the precise identification of hazardous emissions is achieved, such as smoke of wildfire.

The main research contents and principal contributions of this paper are summarised as follows:

- In this study, we implemented an upgrade within the backbone, substituting traditional convolution kernels with ODConv (Li et al., 2022), leading to the establishment of a swift pyramid pooling module. This alteration notably enhances the model's Precision and Recall under identical training conditions, achieving this improvement without significantly enlarging the size of the improved YOLOv8, which compared to the original YOLOv8 that the updated model saw an observably increase in Precision from 90.08% to 91.37% and in Recall from 83.07% to 86.41%.
- This research undertakes a comparative analysis between the enhanced YOLO architecture and the original YOLOv8 structure, illustrating the method's superiority through metric assessments, including mAP50 and mAP50-95. Under similar training conditions and without a substantial increase in model size, the original YOLOv8 achieved mAP50 scores of 93.03% and 94.18% for the new model, with mAP50-95 scores of 63.57% and 66.43%s.

2 Related works

Image-based fire detection represents a multidisciplinary research field encompassing domains such as image processing, pattern recognition, and deep learning. Currently, it can be classified into two primary domains of wildfires detection: methods based on digital image processing and methods grounded in deep learning. Fire detection approaches rooted in traditional digital image processing, which have an earlier inception and a relatively mature technology, while fire detection methods grounded in deep learning emerged later and hold greater potential for development.

2.1 Based on deep learning

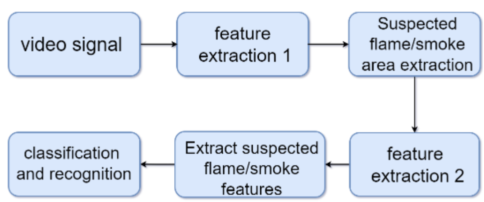
These significant challenges, which traditional fire detection methods relying on digital image processing faced in extracting intricate semantic information related to flames and smoke, often lead to an increased incidence of both false positives and false negatives. Convolutional neural networks (CNNs) possess exceptional self-learning capabilities, allowing them to proficiently uncover profound semantic insights within images. With the continuous evolution of CNNs and the emergence of high-performance computing devices, fields like the image classification and the object detection have undergone substantial improvements. This progress has resulted in the successive introduction of efficient single-stage object detection networks, including YOLO (Redmon et al., 2016), scalable, and the single shot multibox detector (Liu et al., 2016) and efficient object detection (Tan et al., 2020).

As a consequence of these advancements, fire detection methods rooted in deep learning have garnered widespread adoption within the domain of fire detection. The fundamental process can be delineated into two phases: training and inference. In the training phase, the initial step involves amassing a significant volume of fire-related images to construct a comprehensive training dataset. Subsequently, the data is subject to annotation or pre-processing to adapt it for the specific task. Following this, the dataset undergoes pre-processing designed to enhance data diversity. The processed data is then utilised as input for a convolutional neural network to execute feature extraction. Finally, the features are classified, and this sequence of steps is iteratively executed during training to attain the optimal model.

2.2 Based on digital image processing

Fire detection, which based on traditional digital image processing primarily focuses on flames or smoke as its core research subjects and employs digital image processing algorithms to extract features, thereby facilitating the determination of the presence of flames or smoke in images or videos. The fundamental operational principle is illustrated in Figure 1. Initially, feature extraction is applied to video signals to identify suspected areas containing flames or smoke within the signal. Subsequently, another round of feature extraction is carried out on the suspected flame or smoke areas to confirm whether these regions indeed contain flames or smoke. Feature extraction 1 and feature extraction 2 are typically employed to extract dynamic and static features related to flames or smoke.

Figure 1 Flow chart of wildfire smoke detection based on digital image processing (see online version for colours)

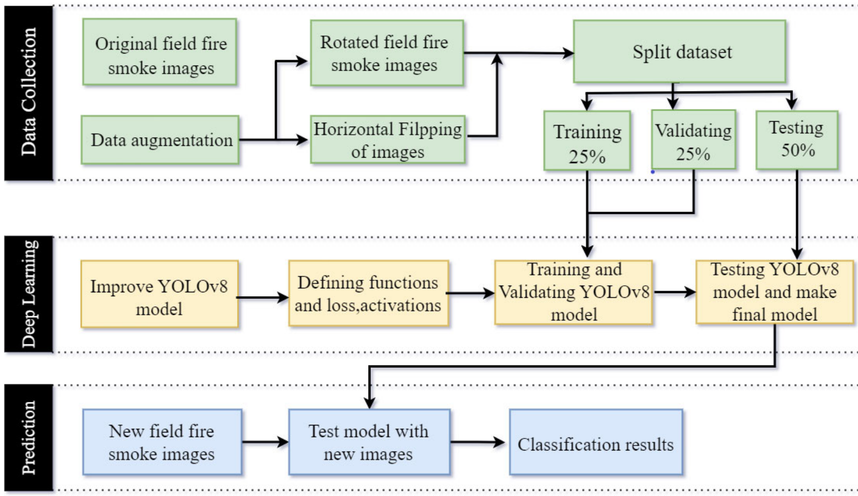


3 Materials and methods

This paper conducts a study that leverages images captured by UAVs and surveillance equipment, to improve the precision and efficiency of early detection of the wildfire smoke in complex environmental settings, in conjunction with deep learning models (Saydirasulovich et al., 2023). The introduction of ODConv is a pivotal component of this research, as it aims to optimise the original YOLOv8 model, thus resulting in marked improvements in its accuracy and convergence speed.

Upon the collection of these images, it becomes imperative to execute pre-processing procedures designed to optimise and enhance their features. This phase encompasses the crucial task of segregating the pixels of interest from the surrounding background, an operation of paramount importance. Extracting details related to smoke and fire in images relies heavily on capturing them under particular daytime and lighting situations. Elements such as edges, corner points, colour characteristics, and intensities play crucial roles in this extraction process. Analysing segmented images thoroughly involves a sequence of operations connected to extracting these features in order to pinpoint the most critical areas. Once these images are processed, they are fed into a trained model to confirm their accuracy. Figure 2 illustrates a detailed breakdown of this particular process.

Figure 2 Flow chart of the wildfire smoke detection model training (see online version for colours)



3.1 *Original YOLOv8*

The YOLO model represents an object recognition and localisation algorithm grounded in deep neural networks. Its most prominent attribute lies in its high-speed operational capabilities, rendering it particularly suited for real-time systems. This model has garnered significant acclaim and has been extensively employed in the domain of computer vision. Expanding upon this foundational framework, a multitude of scholars have engaged in research efforts and instituted improvements, culminating in the development of diverse novel enhancement modules and methodologies. Consequently, this evolutionary process has given rise to numerous classical models that exhibit substantial enhancements. YOLOv8, unveiled by Ultralytics on January 10, 2023, signifies a momentous leap in this trajectory of development. Distinguishing itself from its forerunners, such as YOLOv5 and YOLOv7, it now claims the position of the most innovative computer vision model globally and stands as a highly adaptable platform for customisation. The network's fundamental structure primarily encompasses three key components: the head, backbone, and neck (Jocher, 2023).

In YOLOv8, the enhanced CSPDarknet53 (Redmon and Farhadi, 2018) functions as the backbone network, facilitating the generation of five discrete scale features through a sequence of five successive down sampling stages. Within the preceding backbone architecture, the process of feature extraction is characterised by the application of cross stage partial network (CSPNet) in conjunction with the principles underpinning residual structures. This amalgamation gives rise to the C3 block, which serves as a distinctive feature. Within this framework, the cross stage partial (CSP) primary branch gradient module assumes the form of the BottleNeck module, commonly recognised as the residual module. Notably, the backbone structure of YOLOv8 exhibits refinements to the C3 module, informed by the ELAN concept, culminating in the development of the C2f module. This strategic evolution ensures that the model retains its lightweight characteristics while concurrently enhancing the flow of gradient information.

Furthermore, a seminal design innovation within the YOLO framework is the incorporation of the spatial pyramid pooling (SPP) structure (He et al., 2015). The spatial pyramid pooling fast (SPPF) module is efficiently configured. It dynamically generates data of uniform dimensions through three consecutive maximum pooling layers in the latter PHASES of the backbone network. This process pools the input feature maps. Compared to the conventional SPP structure, the SPPF substantially augments computational efficiency and mitigates latency.

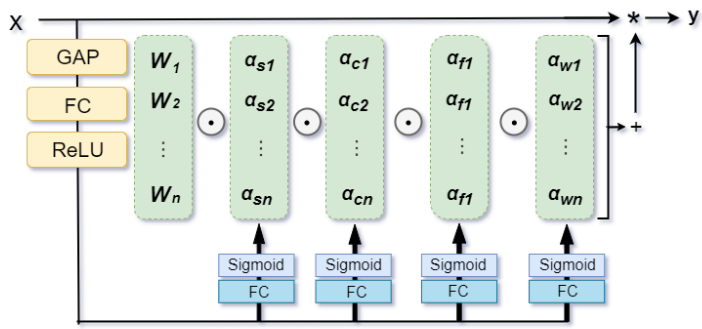
Inspired by path aggregation network (PANet) (Liu et al., 2018), the PAN-FPN architecture is integrated into the Neck component of YOLOv8, which in contrast to the neck designs in past version, such as YOLOv5 and YOLOv7, a modification is made by YOLOv8 within the PANet structure, following up-sampling eliminating the convolution operation, thereby observably streamlines the configuration without compromising the model's initial performance. A comprehensive network structure that harmonises both top-down and bottom-up components are created in this approach, which through feature fusion blends surface-level positional insights with deep semantic details, thereby enriching the breadth and depth of features.

Within the head structure of the YOLOv8, it embraces a decoupled-head and incorporates the distribution focal loss (DFL) (Li et al. 2020) concept for object classification and bounding box regression predictions. This version refines loss functions, using the vertical federated learning (VFL) loss for classification and the complete intersection over union (CIOU) loss alongside the DFL for regression, each offering distinct features. The DFL focuses on modelling box positions, allowing the network to quickly zero in on positions close to the target and the classification tasks are supported by the binary cross-entropy loss (BCE loss). All these choices aim to boost speed up and detection accuracy the learning of the model. In addition, anchor-free detection of the YOLOv8 simplifies distinguishing between positive and negative samples, and the task-aligned one-stage object detection (TOOD) (Feng et al., 2021) is integrated for better sample allocation, which improve both the robustness and detection accuracy of the model.

3.2 Omni-dimensional dynamic convolution (ODConv)

ODConv, whose outstanding performance has been vividly demonstrated by empirical verification through ImageNet classification and the COCO detection tasks, employs a multi-dimensional attention mechanism that encompasses four dimensions within the kernel space and employs a parallel strategy to capture complementary attention, not only enhances the efficacy of extensive models but also confers advantages upon lightweight models and it serving as a 'plug-and-play' operation and seamlessly integrates into existing the CNN networks. In terms of the prevailing view, the ODConv is designated as the full-dimensional dynamic convolution because to a certain extent, the ODConv can be conceptualised as an extension of the CondConv, which expands upon the dynamic attributes inherent in the CondConv within a singular dimension, while concurrently accounting for the dynamics spanning spatial, the input channel, the output channel, and other dimensions.

Figure 3 The architecture of ODConv (see online version for colours)



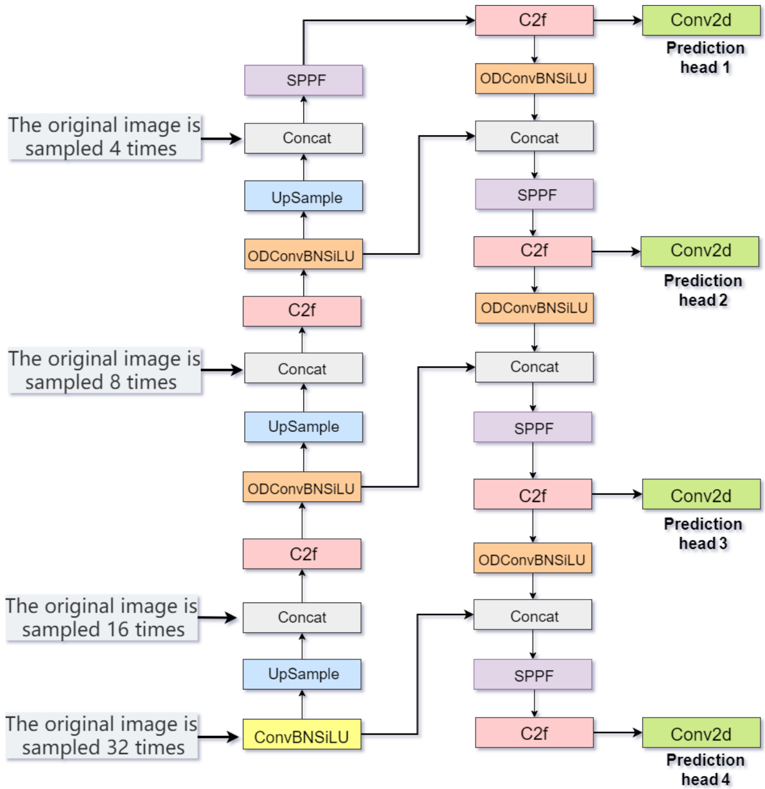
ODConv using a parallel approach to cultivate versatile attention across four dimensions inherent to the convolutional kernel space introduces a multi-dimensional attention mechanism, which is mutually complementary in principle, and its architecture is presented in Figure 3. The gradual multiplication of convolutional, which is specifically manifested as that each parameterised differently across dimensions such as the channel, the position, the kernel, and the filter, enhances the convolution operation's sensitivity to input disparities across these diverse dimensions and fortifies its potential to capture intricate contextual information, and in the issue that, ODConv significantly enhances the feature extraction capacity within convolution (Li et al., 2022). What is more important that, even when equipped with fewer convolutional kernels in it, ODConv attains comparable or even superior performance when juxtaposed with the CondConv and the DyConv, and it is worth underscoring that, thanks to its enhanced feature extraction capabilities, ODConv can achieve performance levels on par with, or surpassing, existing multi-kernel dynamic convolutions when employing a solitary convolutional kernel.

3.3 The design of feature fusion network structure based on YOLOv8

To enhance the feature fusion capabilities of the feature pyramid networks (FPN) and the network's predictive capacity, this paper introduces an innovative FPN structure. Specifically, as depicted in Figure 4, it integrates the ODConv as elucidated in Figure 3. This study substitutes the original network's convolutions with ODConv, which reduces the model's complexity. By utilising a multi-dimensional attention mechanism and a parallel strategy to capture complementary attention, this approach significantly improves the model's recognition accuracy and convergence speed.

Moreover, recognising that smoke in real-world scenarios often manifest as non-rigid objects with variable sizes, this design caters to situations involving small targets. Furthermore, the structure developed in this paper incorporates a prediction layer with feature map size subsampled four times from the original image (4x subsampled prediction layer). This improvement introduces an input layer with a fourfold sampling rate to preserve more of the original image features. This approach increases the feature information for small objects and enhances the input of shallow object texture information. This enhancement fortifies its aptitude for capturing and leveraging shallow feature information, particularly in scenarios characterised by small objects.

Figure 4 The structure of FPN layer (see online version for colours)



4 Experimental results

The present section offers a comprehensive overview, which encompasses the delineation of hyperparameter configurations, delineation of specific training, test datasets employed, an elaborate illustration of the experimental hardware setup, and a thorough explication of the validation procedures that applied to gauge the efficacy of the improved YOLOv8 model in identifying instances of wildfire smoke within the dataset.

Table 1 Descriptions of hardware and software

Items	*Descriptions*
Storage	SSD: 2TB
RAM	DDR5 32GB
Motherboard	MPG Z790 EDGE TI MAX WIFI
GPU	GeForce 4060Ti 16GB
CPU	Intel core i7 processor 14700K
OS	Windows 11 Pro:22H2
Storage	SSD: 2 TB

In order to make sure methodological robustness that all experiments were rigorously executed under uniform hardware conditions. The experimentation was conducted employing a custom-assembled computer systems, which features specific and finely detailed specifications, meticulously elucidated in Table 1, inclusive of an Nvidia GeForce 4060Ti graphics processing unit endowed with 16 GB of memory, supplemented by 32 GB of RAM, and powered by an Intel Core i7 processor 14700K running at 5.6 GHz across 20 cores. The input images, which are used for the training of improved YOLOv8 model, were extracted from a specialised dataset focused on the analysis of wildfire smoke (Zhang et al., 2018). The comprehensive evaluation entails an expansive array of dimensions, a comprehensive performance analysis of the YOLOv8 model, a meticulous assessment of methodological impact, encompassing the experimental setup and its design, comparative analyses of different models, in-depth ablation studies, and discernible visualisation outcomes. The specific parameters instrumental in the training process aimed at the detection of wildfire smoke are meticulously documented in Table 2.

Table 2 Hyperparameters for training

Train hyperparameters	*Details*
Learning rate	0.01
Batch size	16
Epoch	400
Image size	640

4.1 Evaluation metrics

The research relied exclusively on the publicly available wildfire smoke dataset provided by the State Key Laboratory of Fire Science (SKLFS), University of Science and Technology of China (USTC). The evaluation of the detector's accuracy commonly involved the assessment of its ability to correctly identify instances. In contrast, the recall rate of the model represents the ratio of its accurate predictions to the total count, thereby acting as a pivotal measure of its competency in identifying occurrences. Models exhibiting high recall rates proficiently detect a substantial number of fire-related images while maintaining precision by focusing on pertinent targets. The calculations for accuracy and recall in this study are based on equations (1) and (2) and are utilised to appraise the accuracy and recall of the smoke detection method.

$$Precision_{C_{ij}} = \frac{TP_{C_{ij}}}{TP_{C_{ij}} + FP_{C_{ij}}} \tag{1}$$

$$Recall_{C_{ij}} = \frac{TP_{C_{ij}}}{TP_{C_{ij}} + FN_{C_{ij}}} \tag{2}$$

The count of accurately identified smoke regions is labelled as $TP_{C_{ij}}$ (true positives), while instances where non-smoke areas are erroneously classified as smoke are labelled as $FP_{C_{ij}}$ (false positives). $FN_{C_{ij}}$ (false negatives) occur when genuine smoke regions are incorrectly categorised as non-smoke areas. Assessing these values involves the computation of the average precision (AP) and mean average precision (mAP) using equations (3) and (4).

$$AP_{C_{ij}} = \frac{1}{m}\sum_{j=1}^{m} Precision_{C_{ij}}, \tag{3}$$

$$mAP = \frac{1}{C}\sum_{j}^{c} AP_j, \tag{4}$$

4.2 Quantitative comparison

The quantitative evaluation within this article aimed to assess the efficacy of the proposed methodology comprehensively. Various metrics, including precision, recall, mAP50 and mAP50-95, were computed through equations (1) to (4). In addressing the dataset's diversity in smoke instances, ranging across different distances, and encompassing both small and large areas, systematic and thorough testing occurred among diverse YOLO series models. The primary goal was to identify a robust approach capable of accurately detecting smoke within wildfire contexts.

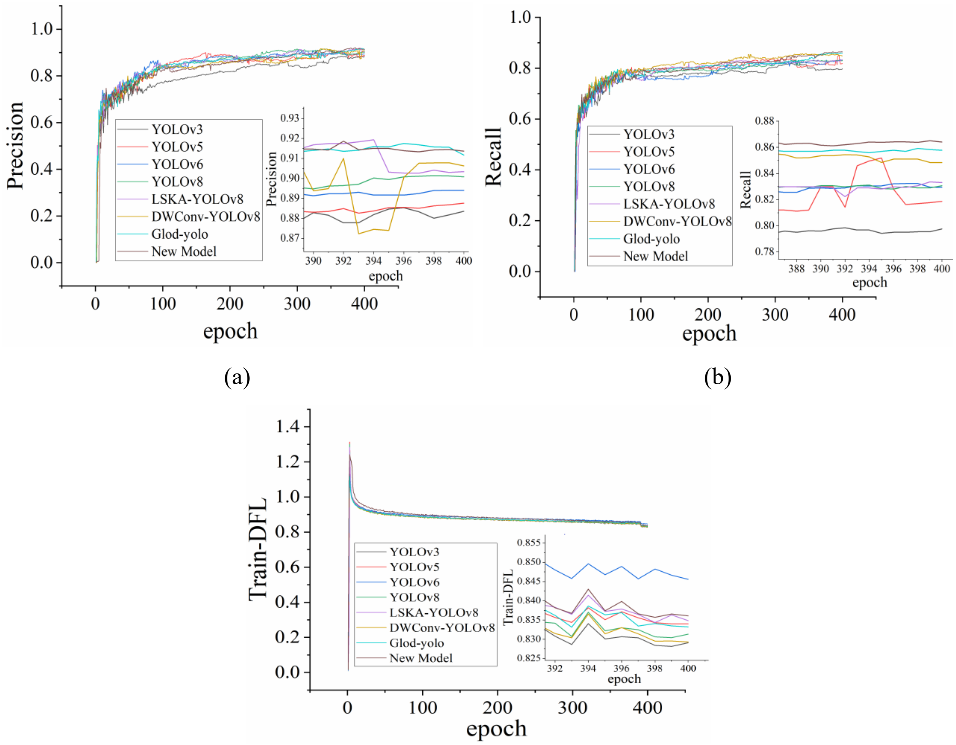
This study primarily focuses on the utilisation of deep learning models for wildfire smoke detection, with a specific emphasis on mitigating impacts on forest ecosystems and ensuring human safety. YOLOv8 was selected as the primary framework after a thorough dataset evaluation due to its efficiency in swiftly detecting smoke instances of diverse sizes and orientations. Single-stage detectors like YOLOv8 proved more suitable for immediate and real-time applications compared to the more complex multi-stage object detectors typically used in this field. The proposed forest smoke detection model, built upon YOLOv8, showcases substantial enhancements across multiple performance metrics – such as mAP5, mAP50-95, recall, precision, and convergence speed – when contrasted with alternative object detection methods.

To thoroughly assess the effectiveness of the proposed approach, the analysis encompassed various object detection technologies and their modifications, including YOLOv3 (Jocher, 2023), YOLOv5 (Jocher, 2022), YOLOv6 (Li, 2022), YOLOv8 (Jocher, 2023), Glod-YOLO (Wang, 2023), YOLOv8 enhanced using DWConv (Howard, 2017), and YOLOv8 incorporating large-separable convolutional attention (LSKA) (Lau et al., 2023). The comparative performance analysis between the improved YOLOv8 model and various object detectors concerning wildfire smoke dataset is detailed in Table 3, and Figure 5.

Table 3 Model quantitative analyses index table

Models	*YOLOv3*	*YOLOv5*	*YOLOv6*	*YOLOv8*	*LSKA-YOLOv8*	*DWConv-YOLOv8*	*Glod-yolo*	*New model*
mAP50 (%)	91.05	92.23	92.86	93.03	93.35	93.68	94.01	95.18
mAP50-95 (%)	58.42	61.13	63.03	63.57	64.09	65.04	65.97	67.43
GFLOPs	28.3	7.2	11.9	8.2	7.8	8.0	12.0	8.1

Figure 5 Graph of the change curve of the model evaluation index, (a) precision curve graph, (b) recall curve graph, (c) train-DFL curve graph (see online version for colours)



(a) (b)

(c)

The recently developed model, as portrayed in Figure 5 and Table 3, has augmented its convergence speed without a substantial escalation in the size of the model. This advancement has notably amplified both model recall and precision, resulting in superior performance relative to analogous models evaluated within the identical dataset and training circumstances. Specifically, in Table 3, the original YOLOv8 obtained mAP50 scores of 93.03% and 95.18% for the new model, while registering mAP50-95 scores of 63.57% and 67.43%, respectively. Despite its reduced dimensions, the fresh model highlights improvements compared to the Glod-YOLO model, which secured the second position in mAP50 and mAP50-95 scores. Additionally, as depicted in Figure 5, the revised model has elevated its recall from 83.07% to 87.67% and its precision from 90.08% to 91.37% in contrast to the original YOLOv8. While maintaining a precision akin to the second-ranked Glod-YOLO, it manifests higher recall. When juxtaposed with other enhancements based on YOLOv8, including LSKA and DWConv, the new model excels in every aspect except model size.

4.3 Qualitative evaluation

Additionally, using quantitative assessment methods to test the efficiency of wildfire smoke detection methods, a qualitative analysis was conducted in this paper. Twenty images were specifically chosen from the dataset, and sixteen of them portrayed smaller and abrupt smoke plumes caused by wildfires, while the remaining four of them exhibited substantial smoke plumes caused by wildfires. The reliable results for both categories were consistently delivered by the improved YOLOv8 model, which as evidenced in Figure 6 that a variety of scenarios and conditions were portrayed in these images and instances of smoke dispersal in different directions were also displayed.

Figure 6 Examples of wildfire smoke detection results (see online version for colours)

In response to the difficulties that numerous methodologies delineated in existing literature in accurately discerning smoke arising from minor-scale wildfires in images, with the intent of augmenting the dataset and enhancing the precision of smoke detection, we carefully selected and exhibited a set of photographs as evidenced in Figure 6, which capture the scale of smoke from wildfires on multiple scales. Figure 6 serves as a visual representation of the efficacy of the proposed approach for detecting wildfire smoke in various forested settings, which employed an improved YOLOv8 model, and the robustness of method was substantiated through assessments encompassing both minor and substantial wildfires smoke images.

The expeditious detection of smoke holds paramount importance in the sphere of wildfire prevention and management because that even minor trace of smoke, if without promptly treatment, will carry the potential to trigger catastrophic wildfires and might place ecosystems in jeopardy, natural resources, and human lives. Furthermore, the proposed method demonstrates exceptional precision in the identification of small smoke patches within image data. The research findings underscore the proficiency of the proposed methodology in diminishing instances of erroneous smoke detections. This efficacy, regardless of their size, composition, or spatial orientation, translates into the ability to promptly address and respond to diverse wildfire smoke scenarios.

4.4 Ablation study

This study incorporates a sequence of ablation experiments focused on evaluating the impact of the ODConv module and the 4x subsampled prediction layer on refining the accuracy of the proposed YOLOv8 model for smoke detection. More precisely, four specific ablation experiments were undertaken in this research, which comprised YOLOv8, YOLOv8 with ODConv, YOLOv8 with the 4x subsampled prediction layer and YOLOv8 with both the 4x subsampled prediction layer and ODConv. The findings of these experiments are comprehensively outlined in Table 4, providing a detailed assessment of the potential improvements in the YOLOv8 model's performance attributed to these modifications. These evaluations are based on a spectrum of metrics, encompassing mAP50, mAP50-95, precision, and recall.

Table 4 Results of the ablation study

Models	*ODConv*	*4x subsampled prediction layer*	*mAP50 (%)*	*mAP50-95 (%)*	*Precision (%)*	*Recall (%)*
YOLOv8	×	×	93.03	63.57	90.08	83.06
	×	√	93.83	65.56	90.67	85.52
	√	×	93.97	66.38	89.18	85.86
	√	√	95.18	67.43	91.37	87.67

The ablation study illustrates that, while the YOLOv8 object detection model maintains robust performance, its effectiveness may not reach its optimal potential in particular scenarios. These findings suggest that the incorporation of ODConv and the 4x subsampled prediction layer into network architecture of the YOLOv8 holds the potential to markedly improve model accuracy in diverse contexts.

5 Conclusions

This research introduces an improved YOLOv8 model tailored for the detection of wildfire smoke in intricate forest environments. The challenge faced by wildfire smoke detection algorithms in maintaining consistent performance arises from the inadequate availability of appropriate training images, resulting in issues like overfitting and data imbalance. The incorporated elements, as illustrated in Table 4, encompass the integration of the ODConv module and the 4x subsampled prediction layer, showcasing considerable outcomes: precision at 91.37%, recall at 87.67%, and mAP50 and mAP50-95 at 95.18% and 67.43%, respectively. These enhancements notably amplify precision, recall, mAP50, and mAP50-95, reflecting the respective increases of 1.29%, 4.61%, 2.15%, and 3.86%. Experimental findings highlight that the optimised YOLOv8 model, without significant alterations to its size, outperforms all YOLO series models and exceeds several improved versions of YOLOv8 incorporating diverse modules. The improved YOLOv8 approach successfully addresses limitations seen in traditional wildfire smoke sensors related to their coverage and concurrent fire detection capabilities, allowing for wildfire smoke detection that accounts for material and geographic attributes.

Enriching the diversity of wildfire smoke images holds paramount importance for enhancing detection efficacy in natural settings. In order to effectively deal with the problem that limited data sets may lead to poor generalisation and low detection accuracy, our forthcoming research endeavours will be dedicated to the comprehensive collection of a diverse array of wildfire smoke images, coupled with the implementation of specialised methodologies to refine their quality. The integration of varied data sources, such as satellite imagery and meteorological data, has the potential to substantially augment the precision and dependability of detection systems. In addition, we will also further optimise the model by introducing the latest attention mechanisms or further optimise the network structure by referring to other network design paradigms. The primary objective is to expedite the detection process while maintaining model compactness without compromising accuracy. The development of robust algorithms capable of real-time operation in diverse environmental conditions is of critical significance. Emphasising these considerations is poised to not only advance early wildfire detection, but also contribute significantly to more efficacious disaster management, thereby safeguarding both ecosystems and human life.

Acknowledgements

The funding of this research is as follows:

1 Jishou University introduced personnel and returned doctoral research funding project: Key Technologies on Collaborative Information Retrieval for Big Data.

2 Hunan student's innovation and entrepreneurship training program (No. S202310531041).

3 Hunan student's innovation and entrepreneurship training program (No. S202310531073).

4 Hunan student's innovation and entrepreneurship training program (No.S202410531098S).

5 Hunan student's innovation and entrepreneurship training program (No. S202410531008).

References

Academy of Disaster Reduction and Emergency Management, National Disaster Reduction Centre of China International Federation of Red Cross and Red Crescent Societies (IFRC), Beijing Normal University (2023) *2022 Global Natural Disaster Assessment Report* [online] https://www.preventionweb.net/publication/2022-global-natural-disaster-assessment-report (accessed 12 October 2023).

Avazov, K., Mukhiddinov, M., Makhmudov, F. and Cho, Y.I. (2021) 'Fire detection method in smart city environments using a deep-learning-based approach', *Electronics*, Vol. 11, No. 1, p.73.

Feng, C., Zhong, Y., Gao, Y., Scott, M.R. and Huang, W. (2021) 'TOOD: task-aligned one-stage object detection', in *Proceedings of the 2021 IEEE International Conference on Computer Vision (ICCV)*, Montreal, QC, Canada, 10–17 October, pp.3490–3499.

He, K., Zhang, X., Ren, S. and Sun, J. (2015) 'Spatial pyramid pooling in deep convolutional networks for visual recognition', *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 37, No. 9, pp.1904–1916.

Howard, A.G. (2017) *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*, arXiv 2017, arXiv:1704.04861.

Hu, Y., Zhan, J., Zhou, G., Chen, A., Cai, W., Guo, K., Hu, Y. and Li, L. (2022) 'Fast forest fire smoke detection using MVMNet', *Knowl.-Based Syst.*, Vol. 241, p.108219.

International Federation of Red Cross and Red Crescent Societies (IFRC) (2020) *2020 Global Natural Disaster Assessment Report*, National Disaster Reduction Centre of China, Academy of Disaster Reduction and Emergency Management [online] https://www.preventionweb.net/publication/2020-global-natural-disaster-assessment-report (accessed 20 October 2021).

Jocher, G. (2022) *yolov5*, Github [online] https://github.com/ultralytics/yolov5 (accessed 22 November 2022).

Jocher, G. (2023) *YOLOv8*, Github, Ultralytics [online] https://github.com/ultralytics/ultralytics (accessed 10 May 2023).

Lau, K.W., Po, L-M. and Rehman, Y.A.U. (2023) *Large Separable Kernel Attention: Rethinking the Large Kernel Attention Design in CNN*, arXiv 2023, arXiv:2309.01439.

Li, C. (2022) *YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications*, arXiv 2022, arXiv: 2209.02976.

Li, C., Zhou, A. and Yao, A. (2022) *Omni-Dimensional Dynamic Convolution*, arXiv 2022, arXiv: 2209.07947.

Li, X., Wang, W., Wu, L., Chen, S., Hu, X., Li, J., Tang, J. and Yang, J. (2020) *Generalized Focal Loss: Learning Qualified and Distributed Bounding Boxes for Dense Object Detection*, arXiv, arXiv:2006.04388.

Liu, S., Qi, L., Qin, H., Shi, J. and Jia, J. (2018) 'Path aggregation network for instance segmentation', in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 18–23 June, pp.8759–8768.

Liu, W., Anguelov, D., Erhan, D. et al. (2016) 'SSD: single shot multibox detector', *European Conference on Computer*, *Vision*, Springer, Cham, pp.21–37.

Ministry of Emergency Management, Beijing Normal University, National Disaster Reduction Centre of China, International Federation of Red Cross and Red Crescent Societies (IFRC) (2022) *2021 Global Natural Disaster Assessment Report* [online] https://www.preventionweb.net/publication/2021-global-disaster-assessment-report (accessed 14 October 2022).

Mukhiddinov, M., Abdusalomov, A.B. and Cho, J. (2022) 'Automatic fire detection and notification system based on improved YOLOv4 for the blind and visually impaired', *Sensors*, Vol. 22, No. 9, p.3307.

Redmon, J. and Farhadi, A. (2018) *YOLOv3: An Incremental Improvement*, arXiv 2018, arXiv: 1804.02767.

Redmon, J., Divvala, S., Girshick, R. et al. (2016) 'You only look once: unified, real-time object detection', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.779–788.

Saydirasulovich, S.N., Mukhiddinov, M., Djuraev, O., Abdusalomov, A. and Cho, Y-I. (2023) 'An improved wildfire smoke detection based on YOLOv8 and UAV images', *Sensors*, Vol. 23, No. 20, p.8374.

Tan, M., Pang, R. and Le, Q.V. (2020) 'Efficientdet: scalable and efficient object detection', *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp.10781–10790.

Wang, C. (2023) *Gold-YOLO: Efficient Object Detector via Gather-and-Distribute Mechanism*, arXiv 2023, arXiv: 2309.11331.

Zhang, Q-X., Lin, G-H., Zhang, Y-M., Xu, G. and Wang, J-J. (2018) 'Wildland forest fire smoke detection based on faster R-CNN using synthetic smoke images', *Procedia Engineering*, Vol. 211, pp.441–446.

Zheng, B. et al. (2023) 'Record-high CO2 emissions from boreal fires in 2021', *Science*, Vol. 379, pp.912–917, DOI: 10.1126/science.ade0805.