Opium poppy image detection based on improved YOLOV5

Xinyun Zhang
Nanning Guidian Electronic Technology Co.
Nanning, China
wxqys178@163.com

Dawei Luo
School of Computer Science and Information Security
Guilin University of Electronic Technology
Guilin, China
dived_luo@163.com

Li Zhou
School of Computer Science and Information Security
Guilin University of Electronic Technology
Guilin, China
lizhou2022@gmail.com

Dejiangmo
School of Mathematics and Computer Science
Northwest Minzu University
Lanzhou, China
dqwm2000403@163.com

2024 International Symposium on Digital Home (ISDH) | 979-8-3315-0987-3/24/$31.00 ©2024 IEEE | DOI: 10.1109/ISDH64927.2024.00028

***Abstract*—In the past, the use of drones to collect poppy images and perform detection and recognition mainly relied on remote sensing technology, but the limitations of this method are the complex computational process and low recognition efficiency. In recent years, the rise of deep learning technology has greatly boosted the scope of image detection and recognition applications. However, there are certain challenges in applying deep learning to remote sensing images of poppies taken aerially by unmanned aerial vehicles(UAVs), including problems such as insufficient information on image details and interference from diverse planting backgrounds. Therefore, how to form an efficient poppy detection method has become an urgent and exploratory problem. In this paper, we take poppy image recognition in complex planting environments as the research object, based on the yolov5 model, we propose a multiscale attention mechanism applicable to small target detection, through the multiscale convolution operation in the channel dimension, the model can capture the information of multiple semantic levels at the same time, so as to better adapt to the target's changes in different scales and complexity, and then improve the accuracy of target detection.**

***Keywords*-Image detection; opium poppy; deep learning; attention mechanism**

I. Introduction

As the core raw material of traditional drugs, opium poppy not only endangers the physical and mental health of individuals, but also poses a potential risk to socio-economic stability; opium poppy cultivation often chooses hidden locations in order to circumvent the supervision of drug control authorities, and drug control efforts are centred on curbing illicit opium poppy cultivation at the source, which has become a key task for drug law enforcement agencies. Currently, the detection of opium poppy relies mainly on field image collection and manual field surveys, a process that is cumbersome and resource-consuming. In addition, traditional target detection techniques have obvious shortcomings in terms of efficiency and accuracy, making it difficult to effectively detect and recognize areas under opium poppy cultivation[2,3].

In the past, remote sensing techniques were commonly used to collect and identify opium poppy images, but this method is computationally complex, inefficient and has certain limitations. In contrast, unmanned aerial vehicles (UAVs) have a higher spatial resolution of images and are more flexible, capable of capturing and observing finer poppy features close to the ground under variable environmental conditions, and their high resolution and low cost make UAVs an effective alternative to satellite remote sensing, especially in monitoring illicit opium poppy cultivation. In recent years, the development of deep learning technology has greatly facilitated the range of applications for image detection and recognition. However, when applying this technique to poppy images taken aerially by UAVs, we encountered challenges including insufficient information on image details and complex and diverse cultivation environments, so it is particularly important to explore an efficient poppy detection method. The core research of this paper is the key feature detection of poppy images.

Moshia et al[4] studied poppy cultivation in Mexico through deep learning techniques.Liu et al[5] used satellite images to collect poppy remote sensing image dataset and used SSD model for poppy detection. However, their method could not effectively detect low-density poppy cultivation. Illicit opium poppy cultivation still exists in some rural areas of China and is difficult to be detected by remote sensing satellites due to its scarcity, low density and interference from other plants. Unmanned aerial vehicles (UAVs) are more flexible and manoeuvrable than remote sensing satellites, and their high-resolution imagery can help to detect opium poppy in areas that are difficult to detect by satellite remote sensing. However, poppies at different growth stages are extremely similar in shape to the surrounding vegetation,

2769-8823/24/$31.00 ©2024 IEEE
DOI 10.1109/ISDH64927.2024.00028

so it is still challenging to detect and eradicate illicit poppy cultivation using UAV technology in complex growth environments, and the pursuit of lightweight and low-power models has been the focus of research on detecting and combating poppy cultivation.

In this paper, we propose an algorithm based on small-target image detection: in order to address the challenges posed by the size variation of poppies at different growth stages and its impact on the detection performance, we introduce a new multiscale attentional mechanism, whereby by performing multiscale convolutional operations on the channel dimensions, the model can capture information at multiple semantic levels simultaneously, thus better adapting to the target at different scales, complexity, and thus improve the accuracy of small target detection.

II. Related Work

A. UAV remote sensing image detection

With the development of UAV remote sensing technology, researchers have used UAVs to solve some problems in various fields such as agricultural production, disaster relief, etc.[8-10] Feng et al[11] used UAV remote sensing technology to achieve differentiation of urban vegetation mapping.Alvarez-Vanhard et al[13] concluded that the synergy between UAVs and remote sensing satellites is potentially valuable for the surface dynamics. Maes and Steppe[14] analysed in detail the application of UAV remote sensing technology in precision agricultural production. In the field of plant disease detection, Ye et al[12] used UAV red-edge band remote sensing to achieve the detection and identification of banana blight, and Wang et al[15] concluded that UAV remote sensing technology is an effective means to delineate the infected area of cotton root rot; Bharathiraja et al [16] proposed a drone-implemented system for controlling rice diseases relying on Internet of Things (IoT) architecture; Thirumurugan et al [17] explored the use of robots and drones in agriculture and management of foliar diseases in vegetable plants. The above cases show that UAV remote sensing technology has been widely used in several fields with good results, especially in the development of precision agriculture. In recent years, with the development of deep learning technology, target detection algorithms have been applied to a number of fields [18-20], including smart agriculture, medical image processing, etc.

B. Yolo-based poppy image detection

In recent years, with the improvement of computer hardware performance, deep learning algorithms have gradually become a mainstream method for performing plant image detection.Wang et al [21] proposed a detection method based on the YOLOv3 algorithm, which improves the accuracy of poppy detection by adding larger detection frames.Zhou et al [3] proposed a detection method based on the YOLOV3 SPP-GIoU-YOLOV3 -MN, Spatial Pyramid Pooling (SPP) unit and Generalised Interconnection (GIoU) model for UAV poppy detection, achieving better performance than the general YOLOv3. Wang et al [22] proposed a poppy image detection method based on YOLOV5s and DenseNet121 , which reduced the number of misdetected images by 73.88 percent and greatly reducing the workload of manually screening poppy images.Pérez-Porras et al [25] proposed a method for early detection of poppy in wheat and conducted a number of comparative experiments using the YOLOV3, V4, and V5 frameworks as the base model, and training the model with approximate RGB images early in the crop's growth.

YOLO model is one of the most popular target detection methods. Many scholars have improved the detection capability of YOLO model by designing suitable feature extraction modules and insertion of attention mechanisms to meet the needs of different task scenarios. In this paper, a reasonable target detection model still needs to be designed for the small target model of poppy images collected by UAV.

III. Network of this paper

A. Basic YOLOV5 network

The network structure of Yolov5 is shown in the figure below, which mainly consists of Backbone backbone network, Neck neck structure and Head head structure, which are responsible for the preprocessing of input image, feature extraction, feature fusion and output of detection information, respectively.

The Backbone network is responsible for extracting image features, and Neck's network architecture is designed using a FPN+PAN structure, where the FPN layer delivers strong semantic features from the top down and the PAN towers deliver localisation features from the bottom up. The main part of the Head network consists of three detectors that perform object detection on feature maps of different scales using grid-based anchors.

B. Our proposed model

This paper proposes an improved YOLOV5 model, which we have named MAM-yolov5. Specifically, we improve the backbone module of the base model by introducing a new attentional mechanism designed for small-target detection: the Multiscale Attention Mechanism(MAM). This attention mechanism module is inserted between the SSPF module and CSP1_1, as shown in Figure 2. By incorporating this attention mechanism, we aim to enhance the model's ability to extract features from small objects, thereby optimizing the model's performance and achieving more accurate detection of poppies.

First, the input feature tensor:

$$X \in \mathbb{R}^{C\times H\times W} \tag{1}$$

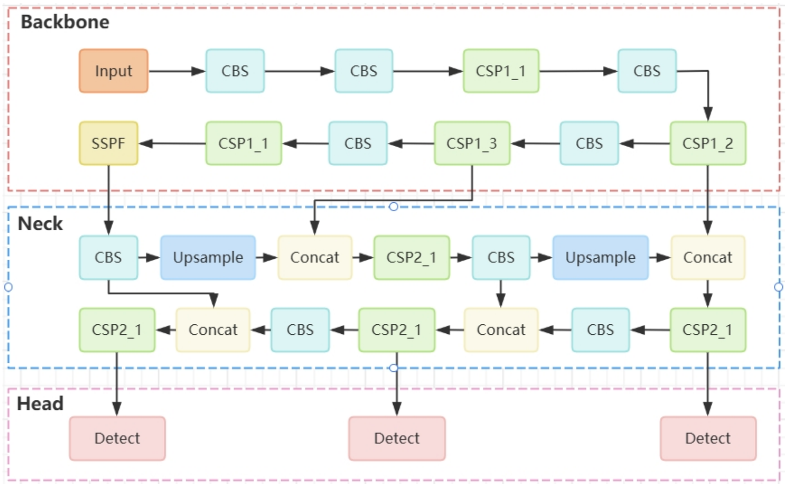

Figure 1. Basic YOLOV5 network structure

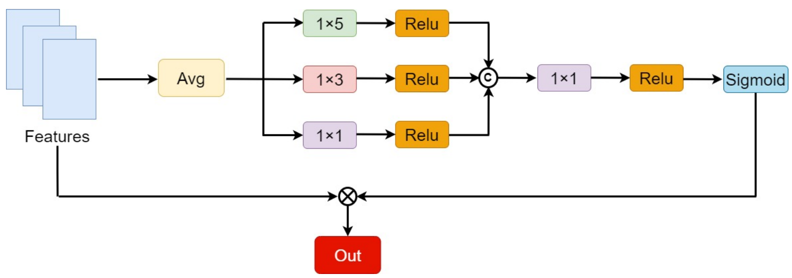

Figure 2. N-attention model

Where C is the number of channels, H and W are the height and width of the feature map respectively.

Then global average pooling is performed to compress the spatial dimension and integrate the global spatial information to better capture the global characteristics of the feature map.

$$z = F_{gap}(X) = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} X_{c,i,j}, \quad z \in \mathbb{R}^C \tag{2}$$

Then multi-scale feature extraction is performed.

1x1 convolution:

$$f_1(z) = \delta(W_1 * z + b_1) \tag{3}$$

1x3 Convolutional:

$$f_3(z) = \delta(W_3 * z + b_3) \tag{4}$$

1x5 convolutional:

$$f_5(z) = \delta(W_5 * z + b_5) \tag{5}$$

where δ denotes the ReLU activation function. And the features are fused so as to merge the features of different scales:

$$u = f_1(z) + f_3(z) + f_5(z), \quad u \in \mathbb{R}^C \tag{6}$$

Attentional weights are generated:

$$s = \sigma(W_s * \delta(W_r * u + b_r) + b_s), \quad s \in \mathbb{R}^C \tag{7}$$

where W_r and b_r are the middle 1x1 convolutional layer parameters, W_s and b_s are the last 1x1 convolutional layer parameters, and σ is the Sigmoid activation function.

Finally, applying attention:

$$Y = X \otimes s \tag{8}$$

where $\otimes$ denotes matrix multiplication. The resulting output:

$$Y \in \mathbb{R}^{C \times H \times W} \tag{9}$$

IV. Experimental Results and Analysis

A. Datasets

The dataset consists of 2358 poppy images, which are all from web crawling, with a size of 640×640 pixels as shown in Fig. 3, and the images are labelled manually using the open source tool Labeling. The dataset is divided into a training set and a test set with 2079 and 279 images, respectively. In the training phase, we divide the training set into training part and validation part, and the test set is only used to test the generality of the model.

Figure 3. Poppy Images

B. Data Enhancement

In YOLOV5, we use Mosaic Data Enhancement, by which the poppy images are processed by this data enhancement method, four poppy images are captured and stitched together by random scaling, random cropping and random arranging. As shown in Fig. 4. the Mosaic data enhancement has the following advantages:

(1)Enriching the dataset: randomly using 4 images, randomly scaling, and then randomly distributing them for splicing greatly enriches the detection dataset, especially the random scaling adds many small targets to make the network more robust;

(2)Reduced GPU memory: directly calculates the data of 4 images, so that the mini-batch size does not need to be very large to achieve better results.

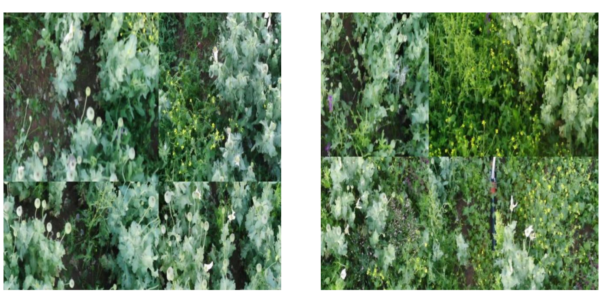

Figure 4. Poppy Mosaic Data Enhancement

C. Parameter setup

Experiments were conducted using Nvidia GeForce GTX 3090, based on Python 3.8 development language, Pytorch 1.11.0 deep learning framework, and Cuda 11.3.0 to build the network model for training. The model was built using the Adam optimiser, with batch size set to 16, epoch set to 300, and learning rate set to 0.001.

D. Evaluation indicators

Poppy target detection recognition can be regarded as a binary classification task, therefore, it can be measured using the common Precision, Recall, and F1 scores, which are evaluated using the formulas, respectively:

- Precision: Precision is defined as the ability of the model to accurately identify relevant objects. It is calculated as the ratio of the number of correctly classified positive instances to the total number of instances classified as positive.

$$Presion = \frac{TP}{TP + FP} \tag{10}$$

- Recall: Also known as sensitivity, it measures the ability of a model to accurately detect true positives. Mathematically, it represents the ratio of correctly identified positive instances to all actual positive instances.

$$Recall = \frac{TP}{TP + FN} \tag{11}$$

- F1 Score: The F1 score is an evaluation metric used to measure the overall accuracy of the model. It takes into account both the precision and recall of the model's predictions.

$$F1 = \frac{2TP}{2TP + FP + FN} \tag{12}$$

where TP, TN, FP and FN denote the true positive (TP), true negative (TN), false positive (FP) and false negative (FN) produced by each pixel in the segmentation result, respectively.

E. Comparative experiments and analyses

In this paper, we perform comparative experiments on YoloV5-based related algorithms. In this paper, we use pre-training weights yolov5s.pt to conduct experiments, on this basis, we conduct comparison experiments on multiple attention mechanisms and compare them with the multi-scale attention mechanism proposed in this paper, and we present the comparison results through the relevant parameters, and the results of the comparison experiments are shown in Table I:

Table I
COMPARISON OF EXPERIMENTAL RESULTS

Model	*FLOPS*	*Param.*	*P*	*R*	*F1*	*MAP*
yolov5s	15.9 G	7.01 M	0.888	0.857	0.872	0.913
yolov5_C3CA	49.7 G	21.0 M	0.838	0.848	0.843	0.891
_C3CBAM	49.7 G	21.1 M	0.863	0.838	0.850	0.893
_ECA	48.8 G	21.2 M	0.892	0.854	0.872	0.911
_SE	48.8 G	21.3 M	0.896	0.830	0.862	0.901
_CA	15.1 G	7.04 M	0.897	0.850	0.873	0.913
MAM-yolov5	17.6 G	7.00 M	0.903	0.865	0.884	0.914

From the above table, we can see: the algorithm proposed in this paper is slightly higher than the base model yolov5 and the _CA model with added attention mechanism in terms of complexity less, but the presion is improved by 1.5% and 0.6%, and the F1 score is improved by 1.2% and 1.1%; comparing with the other attention mechanism added, the multi-scale attention mechanism proposed in this paper is in terms of both model complexity and presion Improvement.

V. Conclusion

Compared with the models after introducing other attention mechanisms, the multiscale attention mechanism proposed in this paper demonstrates superior performance characteristics. Specifically, MAM-YOLOv5 shows significant improvement in several key evaluation metrics such as presion and MAP. Notably, the computational complexity of the model remains relatively low despite the improved performance. This result fully demonstrates the advantages of MAM-yolov5 in terms of model robustness and generalisation ability. Taken together, our improvement strategy not only effectively enhances the overall performance of the model, but also achieves a good balance between computational efficiency and detection presion.

Acknowledgment

This work is supported by the Yongjiang Talent Project (No. 2022003) and the Guangxi Science and Technology Major Program (No. AA22068059).

References

[1] Zhang Z, Xia W, Xie G, et al. Fast opium poppy detection in Unmanned Aerial Vehicle (UAV) imagery based on deep neural network[J]. Drones, 2023, 7(9): 559.

[2] Demir S, Başayiğit L. Determination of opium poppy (Papaver somniferum) parcels using high-resolution satellite imagery[J]. Journal of the Indian Society of Remote Sensing, 2019, 47(6): 977-987.

[3] Zhou J, Tian Y, Yuan C, et al. Improved UAV opium poppy detection using an updated YOLOv3 model[J]. Sensors, 2019, 19(22): 4851.

[4] Moshia M E, Newete S W. Mexican poppy (Argemone mexicana) control in cornfield using deep learning neural networks: A perspective[J]. Acta Agriculturae Scandinavica, Section B—Soil & Plant Science, 2019, 69(3): 228-234.

[5] Liu X, Tian Y, Yuan C, et al. Opium poppy detection using deep learning[J]. Remote Sensing, 2018, 10(12): 1886.

[6] Iqbal F, Lucieer A, Barry K, et al. Poppy crop height and capsule volume estimation from a single UAS flight[J]. Remote Sensing, 2017, 9(7): 647.

[7] Luo Z, Yang W, Gou R, et al. TransAttention U-Net for Semantic Segmentation of Poppy[J]. Electronics, 2023, 12(3): 487.

[8] Tijtgat N, Van Ranst W, Goedeme T, et al. Embedded real-time object detection for a UAV warning system[C]//Proceedings of the IEEE international conference on computer vision workshops. 2017: 2110-2118.

[9] Bejiga M B, Zeggada A, Melgani F. Convolutional neural networks for near real-time object detection from UAV imagery in avalanche search and rescue operations[C]//2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2016: 693-696.

[10] Bao W, Zhu Z, Hu G, et al. UAV remote sensing detection of tea leaf blight based on DDMA-YOLO[J]. Computers and Electronics in Agriculture, 2023, 205: 107637.

[11] Feng Q, Liu J, Gong J. UAV remote sensing for urban vegetation mapping using random forest and texture analysis. Remote Sensing. 2015, 7, 1074–1094.

[12] Ye H, Huang W, Huang S, et al. Recognition of banana fusarium wilt based on UAV remote sensing[J]. Remote Sensing, 2020, 12(6): 938.

[13] Alvarez-Vanhard E, Corpetti T, Houet T. UAV & satellite synergies for optical remote sensing applications: A literature review[J]. Science of remote sensing, 2021, 3: 100019.

[14] Maes W H, Steppe K. Perspectives for remote sensing with unmanned aerial vehicles in precision agriculture[J]. Trends in plant science, 2019, 24(2): 152-164.

[15] Wang T, Thomasson J A, Yang C, et al. Automatic classification of cotton root rot disease based on UAV remote sensing[J]. Remote Sensing, 2020, 12(8): 1310.

[16] Bharathiraja N, Pradeepa K, Sheela I J, et al. Prediction of Plant Leaf Diseases using Drone and Image Processing Techniques[C]//2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT). IEEE, 2023: 1723-1727.

[17] Thirumurugan V, Thirumal S, Kumar N, et al. Cost Effective Multifunctional Agri Robotrone Technology for Vegetable Plant Leaf Disease Identification and Prevention[C]//2024 10th International Conference on Communication and Signal Processing (ICCSP). IEEE, 2024: 1617-1620.

[18] Jiang H, Diao Z, Shi T, et al. A review of deep learning-based multiple-lesion recognition from medical images: classification, detection and segmentation[J]. Computers in Biology and Medicine, 2023, 157: 106726.

[19] Abbas A, Zhang Z, Zheng H, et al. Drones in plant disease assessment, efficient monitoring, and detection: a way forward to smart agriculture[J]. Agronomy, 2023, 13(6): 1524.

[20] Chen S, Zou X, Zhou X, et al. Study on fusion clustering and improved YOLOv5 algorithm based on multiple occlusion of Camellia oleifera fruit[J]. Computers and Electronics in Agriculture, 2023, 206: 107706.

[21] Wang C, Wang Q, Wu H, et al. Low-altitude remote sensing opium poppy image detection based on modified yolov3[J]. Remote sensing, 2021, 13(11): 2130.

[22] Wang Q, Wang C, Wu H, et al. A two-stage low-altitude remote sensing papaver somniferum image detection system based on YOLOv5s+ DenseNet121[J]. Remote Sensing, 2022, 14(8): 1834.

[23] Rominger K, Meyer S E. Application of UAV-based methodology for census of an endangered plant species in a fragile habitat[J]. Remote Sensing, 2019, 11(6): 719.

[24] He Q, Zhang Y, Liang L. Identification of poppy by spectral matching classification[J]. Optik, 2020, 200: 163445.

[25] Pérez-Porras F J, Torres-Sánchez J, López-Granados F, et al. Early and on-ground image-based detection of poppy (Papaver rhoeas) in wheat using YOLO architectures[J]. Weed Science, 2023, 71(1): 50-58.