

Enhancement Dataset for Low Altitude Unmanned Aerial Vehicle Detection

WANG Zhi^{1,2}, HU Wei³, WANG Ershen^{3*}, HONG Chen⁴, XU Song³, LIU Meizhi⁵

1. Zhejiang Jiande General Aviation Research Institute, Jiande 311612, P. R. China;

2. Department of General Aviation, Civil Aviation Management Institute of China, Beijing 100102, P. R. China;

3. School of Electronic and Information Engineering, Shenyang Aerospace University, Shenyang 110136, P. R. China;

4. College of Robotics, Beijing Union University, Beijing 100101, P. R. China;

5. School of Artificial Intelligence, Shenyang Aerospace University, Shenyang 110136, P. R. China

(Received 28 June 2021; revised 25 October 2021; accepted 21 November 2021)

Abstract: In recent years, the number of incidents involved with unmanned aerial vehicles (UAVs) has increased conspicuously, resulting in an increasingly urgent demand for developing anti-UAV systems. The vast requirements of high detection accuracy with respect to low altitude UAVs are put forward. In addition, the methods of UAV detection based on deep learning are of great potential in low altitude UAV detection. However, such methods need high-quality datasets to cope with the problem of high false alarm rate (FAR) and high missing alarm rate (MAR) in low altitude UAV detection, special high-quality low altitude UAV detection dataset is still lacking. A handful of known datasets for UAV detection have been rejected by their proposers for authorization and are of poor quality. In this paper, a comprehensive enhanced dataset containing UAVs and jamming objects is proposed. A large number of high-definition UAV images are obtained through real world shooting, web crawler, and data enhancement. Moreover, to cope with the challenge of low altitude UAV detection in complex backgrounds and long distance, as well as the puzzle caused by jamming objects, the noise with jamming characteristics is added to the dataset. Finally, the dataset is trained, validated, and tested by four mainstream deep learning models. The results indicate that by using data enhancement, adding noise contained jamming objects and images of UAV with complex backgrounds and long distance, the accuracy of UAV detection can be significantly improved. This work will promote the development of anti-UAV systems deeply, and more convincing evaluation criteria are provided for models optimization for UAV detection.

Key words: unmanned aerial vehicle (UAV); UAV dataset; object detection; deep learning

CLC number: TP391 **Document code:** A **Article ID:** 1005-1120(2021)06-0914-13

0 Introduction

The wide popularization and use of low altitude unmanned aerial vehicles (UAVs) facilitate people's lives, while bringing significant challenges to the safety of critical infrastructures such as airports, power stations, water treatment plants, and so on. In the urban environments, the need of detecting UAVs is more frequent than ever before. In 2020 alone, 19 cases of the black flight of UAVs were found in China. One of recent high-profile events

connected to UAVs is the notable Guangzhounan Railway Station drone incident. In complex low altitude areas, some specialized detection devices such as radar cannot detect UAVs effectively. To reduce the black flight of UAVs, it is necessary to develop anti-UAV systems that can detect UAV accurately in low altitude areas^[1].

In recent years, many countermeasures against UAVs have been proposed. The early detection methods of UAVs mainly depend on radar. Pisa et al.^[2] evaluated the radar cross sections of

*Corresponding author, E-mail address: wanges_2016@126.com.

How to cite this article: WANG Zhi, HU Wei, WANG Ershen, et al. Enhancement dataset for low altitude unmanned aerial vehicle detection[J]. Transactions of Nanjing University of Aeronautics and Astronautics, 2021, 38(6):914-926.

<http://dx.doi.org/10.16356/j.1005-1120.2021.06.002>

the commercial UAVs for anti-UAV passive radar source selection, whose values are between -18 dB and -1 dB in the 1–4 GHz frequency range. This preliminary analysis set a solid ground for the selection of anti-UAV radar carrier frequency when using both conventional and passive radars. However, radar cannot detect UAVs accurately in complex urban environments, so UAV detection methods based on computer vision have been proposed to address the issue of UAV detection. Wang et al.^[3] proposed a detection method to address the problems of visual surveillance for flying small UAVs, in which the Gaussian mixture background modeling was used in a compressive sensing domain and low-rank and sparse matrix decomposition of local image.

At present, how to improve the accuracy of UAV detection and meanwhile reduce the false alarm rate (FAR) and missing alarm rate (MAR) is still an open problem. Benefiting from the ability of some powerful deep learning models, the increasing availability of computing power, and the large-scale label datasets, deep learning has achieved great success in the field of object detection^[4-5], such as the well-known YOLO^[6-9], EfficientDet^[10], etc. Huge and high-quality datasets are the fundamental guarantees for deep learning to solve computer vision tasks. Some datasets, such as Microsoft COCO^[11], the pattern analysis statistical modeling and computational learning visual object classes (PASCAL VOC)^[12], and ImageNet^[13] datasets, can achieve high accuracy, low FAR and MAR for most object classifications. Microsoft COCO, PASCAL VOC, and ImageNet are valuable datasets and benchmarks for common objects classification, location and detection. Unfortunately, special high-quality low altitude UAV detection dataset is still lacking.

A dataset named ATAUAV was proposed in 2020 for UAV detection, in which 10 000 different UAV images were collected from videos and google image search^[14]. When using YOLOv3-Tiny^[7] for UAV detection, the highest accuracy for the dataset is 82.7%. When the dataset is used for actual detection on the platform built by the authors' team, the

detection speed is 29.6 frames/s. However, the dataset is only used for the development of airborne anti-UAV system by the authors' team, and authors' team stated categorically that they refuse to share the dataset with others.

In the same year, the dataset RWQUAV (Real world object detection dataset for quadcopter unmanned aerial vehicle detection) was proposed, which consists of 51 446 training set images and 5 375 test set images^[15]. The highest accuracy for UAV detection of the dataset is only 55%. Since UAV images with long distance and complex backgrounds are very scarce in the dataset, resulting in high FAR and MAR. At the same time, there are too many UAV images from video output in the dataset, and these UAV images only have slight differences, which leads to the large number of images in the dataset but poor actual performance. Moreover, it will take a lot of manpower and material resources to obtain UAV images only by taking real world images, and updating later datasets will be difficult. It is necessary to try and explore more simple and effective methods to obtain UAV images, such as data enhancement.

Inspired by the literatures mentioned above, the main contributions of this paper are as follows:

(1) A novel UAV detection dataset is presented and shared to deal with the problem of no public dataset available in the field of UAV detection and no high-quality datasets for uniform evaluation of algorithm innovation for UAV detection.

(2) Multiple methods of data enhancement are first used in the dataset for UAVs and have been proven to improve dataset performance. An effective way to obtain UAV images is provided, and it dramatically reduces the difficulty and cost of UAV image acquisition.

(3) Images of the main jamming targets for UAV detection, such as birds and kites, are first added to the dataset and have proven to be effective in reducing the FAR and MAR for UAV detection.

(4) The objects that interfere with UAV detection in a specific scene, including airliners, bats, balloons and airships, are added to the dataset,

which further improves the practical application value of the dataset for UAV detection, so that the models can obtain better performance in a specific scene.

(5) To fully demonstrate the successful application of the proposed dataset to the challenging problem of drone detection, we use the dataset to train and validate the YOLOv4-Tiny, CenterNet, YOLOv4 and EfficientDet models, and perform a direct side-by-side comparison of the detection results.

The theoretical contribution in this paper is as follows. Noise containing jamming objects is introduced into the generating process of the UAV dataset, avoiding the problem of over-fitting when the models are used in a single-object dataset.

In this paper, the method contribution can be summarized as follows. Many data enhancement methods such as geometric texture enhancement, optical enhancement, and image mixing enhancement are introduced into the generating process of the UAV dataset, further reducing the FAR and MAR of UAV detection.

1 Dataset Preparation

1.1 Real world images generation process

1.1.1 Real world images of UAVs

Real world images of UAVs can be obtained in two ways. The first is the videos taken for UAVs with various types in a variety of complex backgrounds and long distance, and the videos are output as images in 1 frame/s by Free Studio software. The purpose of taking videos from a wide range of complex backgrounds and long distance for UAVs is to extend the UAV image samples with these two features because there are very few of them on the internet. The reason of the more considerable time interval setting is that if the time interval of image output is too small, the flight attitude of UAVs will be similar, and the backgrounds will be too similar. An example diagram is given in Fig.1.

The other way to obtain real world images of UAVs is by web crawler from major mainstream picture websites. The background of images is more



Fig.1 An example of UAV images from real world shooting

affluent and includes various types of UAVs. A part of them are high-definition images of UAVs in close shot, containing more appearance information of UAVs. An example diagram is given in Fig.2.



Fig.2 An example of UAV images from web crawler

1.1.2 Real world images of birds and kites

Real world images of birds and kites come from web crawler, and over 70% of the images are with long distance. Two example diagrams are given in Figs.3 and 4, respectively.



Fig.3 An example of bird images from web crawler



Fig.4 An example of kite images from web crawler

1.1.3 Real world images of secondary jamming objects

Secondary jamming objects include airliners, bats, balloons and airships. Although they are not the main reason for the high MAR of UAV detection, they will cause interference in specific detection scenarios. Adding these images will further enhance the generalization ability of the models and increase the usefulness of the dataset.

Real world images of secondary jamming objects come from web crawler. Four example diagrams are given in Figs.5—8, respectively.

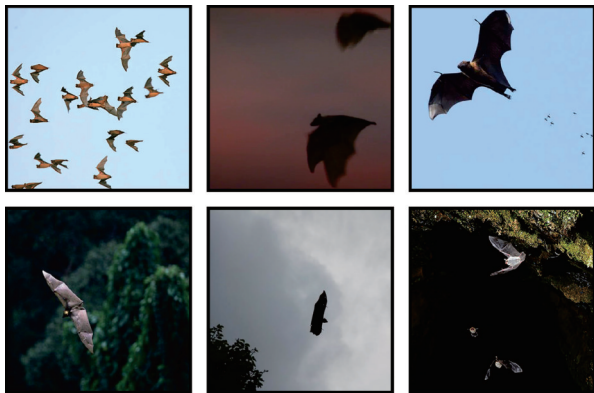


Fig.5 An example of bat images from web crawler



Fig.6 An example of airship images from web crawler



Fig.7 An example of airliner images from web crawler



Fig.8 An example of balloon images from web crawler

1.2 Data enhancement for UAV images

The methods of data enhancement used in datasets can be divided into three main categories: Geometric texture enhancement, optical enhancement, and image mixing enhancement. Concretely speaking, the methods of geometric texture enhancement include flipping or rotation, Gaussian or Salt Pepper noise, Gaussian blur, and affine transformation. The methods of optical enhancement include changes in brightness, contrast, and sharpening. The methods of image mixing enhancement include Mix-Up, CutMix, Mosaic, and DCGANs. Such methods of data enhancement used in datasets are shown intuitively in Fig.9.

1.2.1 Geometric texture enhancement

Geometric texture enhancement is an enhancement operation in the geometric space of images. The main methods used in this paper include randomly flipping along the x- or y-axis and rotating at clockwise 90° , 180° , 270° , or 360° , adding Gaussian noise or Salt Pepper noise with a mean value of 0 and a variance of 0.001 to 0.2, adding Gaussian

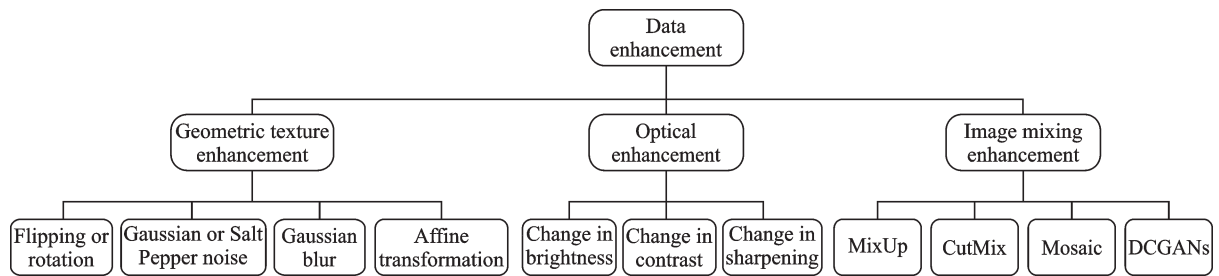


Fig.9 Data enhancement methods used in datasets

blur, and adding affine transformation^[16-18]. Examples of flipping and rotation, adding Gaussian or Salt Pepper noise, adding Gaussian blur, and adding affine transformation are given in Figs.10—13, respectively.



Fig.10 An example of flipping or rotation from geometric texture enhancement



Fig.11 An example of adding Gaussian or Salt Pepper noise from geometric texture enhancement

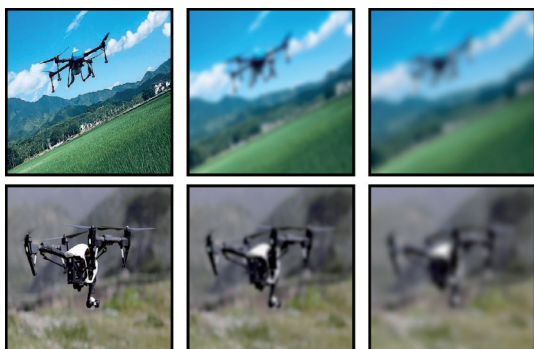


Fig.12 An example of adding Gaussian blur from geometric texture enhancement



Fig.13 An example of affine transformation from geometric texture enhancement

1.2.2 Optical enhancement

Optical enhancement is an enhancement operation by adjusting the visual space of images. The main methods used in this paper are random brightness change, contrast change, and sharpening change^[19-20]. Examples of the change in brightness, contrast and sharpening are given in Figs.14—16, respectively.

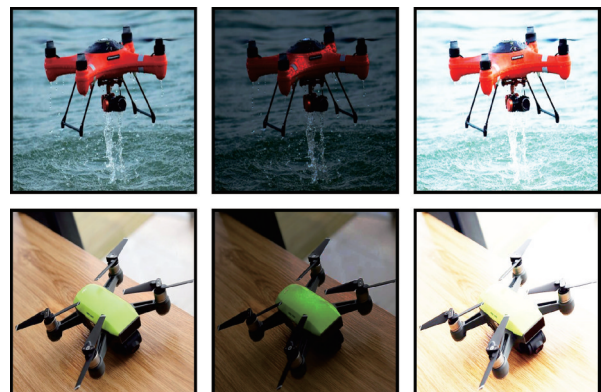


Fig.14 An example of change in brightness from optical enhancement

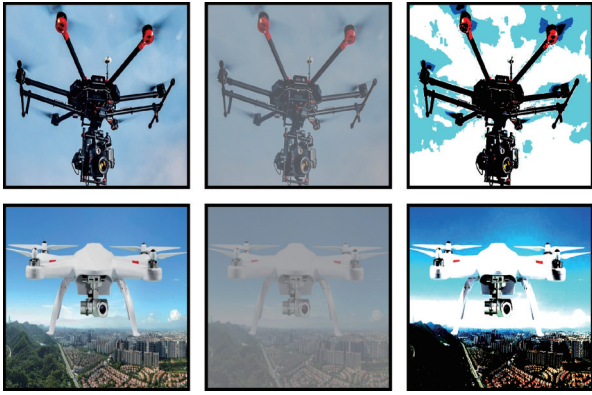


Fig.15 An example of change in contrast from optical enhancement

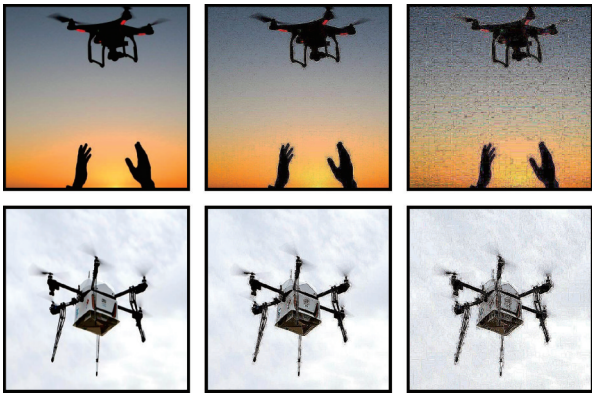


Fig.16 An example of change in sharpening from optical enhancement

1.2.3 Image mixing enhancement

Image mixing enhancement combines multiple image samples from a dataset to synthesize new image samples. The image blending augmentation method has the following characteristics. First, it requires two or more image samples to participate in the process of image augmentation. Then the semantic information of new image samples generated after blending enhancement depends on the semantics of several participating enhancement samples. The enhanced image samples often do not have human visual understanding characteristics. The main methods used in this paper are MixUp^[21], CutMix^[22], Mosaic^[9] and DCGANs^[23], whose examples are given in Figs.17—20, respectively.

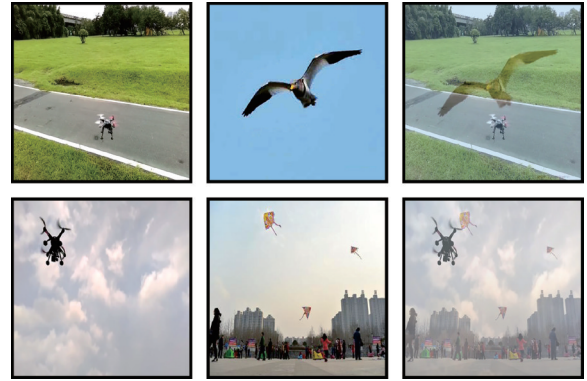


Fig.17 An example of MixUp from image mixing enhancement

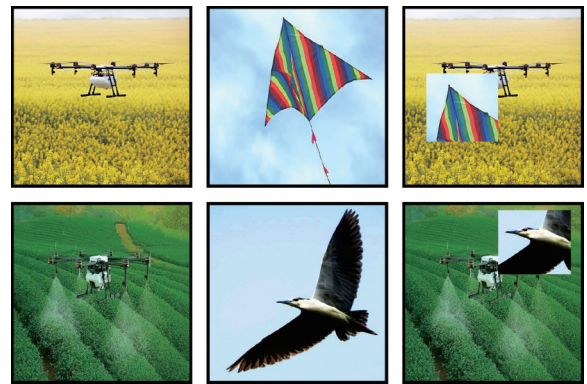


Fig.18 An example of CutMix from image mixing enhancement

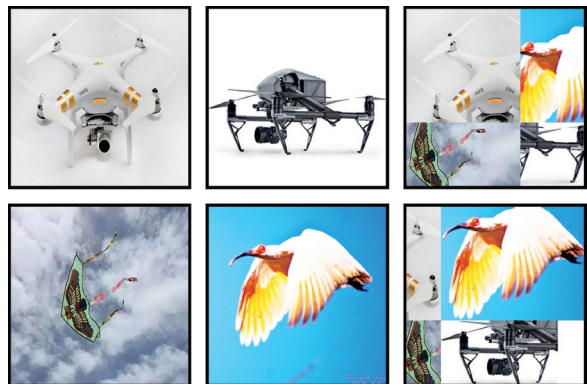


Fig.19 An example of Mosaic from image mixing enhancement

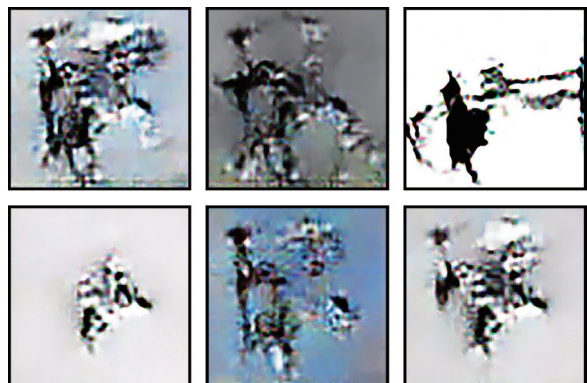


Fig.20 An example of DCGANs from image mixing enhancement

2 Dataset Establishment

The structure of the dataset is given in Fig.21.

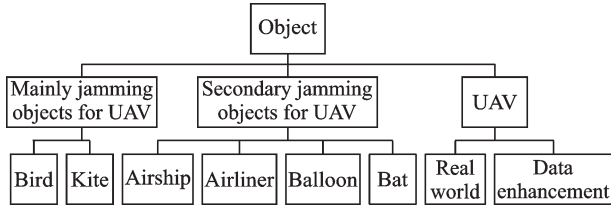


Fig.21 Structure of dataset

2.1 Parameters of dataset

The image format of the dataset is JPG, and the image dimension is $628 \times 628 \times 3$, namely RGB three-channel image. There are seven classes, including 46 571 UAV images, 43 001 bird images, 44 340 kite images, 12 305 airship images, 11 321 airliner images, 10 587 bat images, and 10 363 balloon images, with a total of 200 654 annotations.

2.2 Label of dataset

The annotations are in XML format, and the specific labeling process is as follows: First, we label part of the whole class by using Labelimg software on a small scale manually, and then train with the EfficientDet to get a good accuracy model. Then we input all the pictures into the EfficientDet and outputted the location with the prediction box and prediction box. The use of the model with training significantly improves the efficiency and accuracy of the labeling. Of course, not all images can be labeled accurately. We manually inspect the image information of all model automatic labeling and modify those relatively unreasonable labels, which do not take us too much time. Due to the use of MixUp, CutMix, and Mosaic enhancements, each image may contain one or more classifications or has multiple detections of the same type, which may result in multiple tags in the XML file corresponding to each image and should be noted by the reader when using it.

2.3 Partition of dataset

Referring to the holdout cross-validation method in the cross-validation, that is, according to a fixed proportion, the dataset is statically divided,

ensuring that the train set, verification set, and test set are completely unrelated, and avoiding the virtual high classification accuracy in the actual test due to their overlapping, so that the follow-up experimental results are more scientific. Training set: verification set: test set = 80%:10%:10%. This proportion division comprehensively refers to the division rules of major public datasets and the scale of our dataset.

3 Dataset Verification

Methods of data enhancement used in this paper have been proven by their authors to improve the detection accuracy to a certain extent in most classifications on public datasets such as Microsoft COCO and PASCAL VOC. However, whether the accuracy of UAV detection can be improved still needs relevant experiments to prove its validity, especially after UAV images with complex backgrounds and long distance are deliberately added.

It is important to note that datasets published by the other teams mentioned earlier for UAV are not publicly available and cannot be obtained. Therefore, the accuracy of UAV detection for the above datasets is extracted from the corresponding papers. The average precision of AP50 of ATA UAV and RWQUAV datasets for UAV detection is given in Table 1.

Table 1 AP50 of ATA UAV and RWQUAV datasets for UAV detection %

Dataset	AP50	FAR & MAR (Remote)	FAR & MAR (Complex backgrounds)
ATA UAV	82.7	—	—
RWQUAV	55.0	—	—

3.1 Experimental conditions and parameter setup

These experiments are divided into Experiment 1, Experiment 2, and Experiment 3. The purpose of Experiment 1 is to obtain the accuracy of UAV detection under pure complex backgrounds and long distance, especially after providing a dataset with huge number of images, so as to provide a

benchmark for future research. However, in the previous literatures, there is no emphasis on expanding this kind of image, and it is only mentioned that in the case of complex backgrounds and long distance, the accuracy of UAV detection will be reduced, and the detailed data cannot be given. The purpose of Experiment 2 is to verify the effectiveness of adding images of jamming objects such as birds and kites in improving the accuracy and reducing the FAR and MAR. The purpose of Experiment 3 is to verify the effectiveness of data enhancement methods adopted for UAV images in improving the accuracy and reducing the FAR and MAR. The training, verification, and testing of all experiments are completed on NVIDIA RTX2080Ti graphics card matrix, and the dataset is divided according to the static of holdout cross-validation, in which the train set: verification set: test set = 80%: 10%: 10%, the training epoch is set to 200, the learning rate set to 0.000 1, and the training batch size is 32. Four main models are used in the experiment, YOLOv4-Tiny with CSPDarknet53-Tiny backbone, CenterNet^[24] with Hourglass backbone, YOLOv4 with CSPDarknet-53 backbone, and EfficientDet with D7 backbone.

3.1.1 Experiment 1

The dataset used in Experiment 1 consists of 7 010 original UAV images without data enhancement. Moreover, these images are pure UAV images in long distance and complex backgrounds. The precision-recall (PR) curves of the four models are shown in Figs.22(a—d). Finally, the detection accuracy of UAV, the FAR, and the MAR under long distance and complex backgrounds are shown in Table 2.

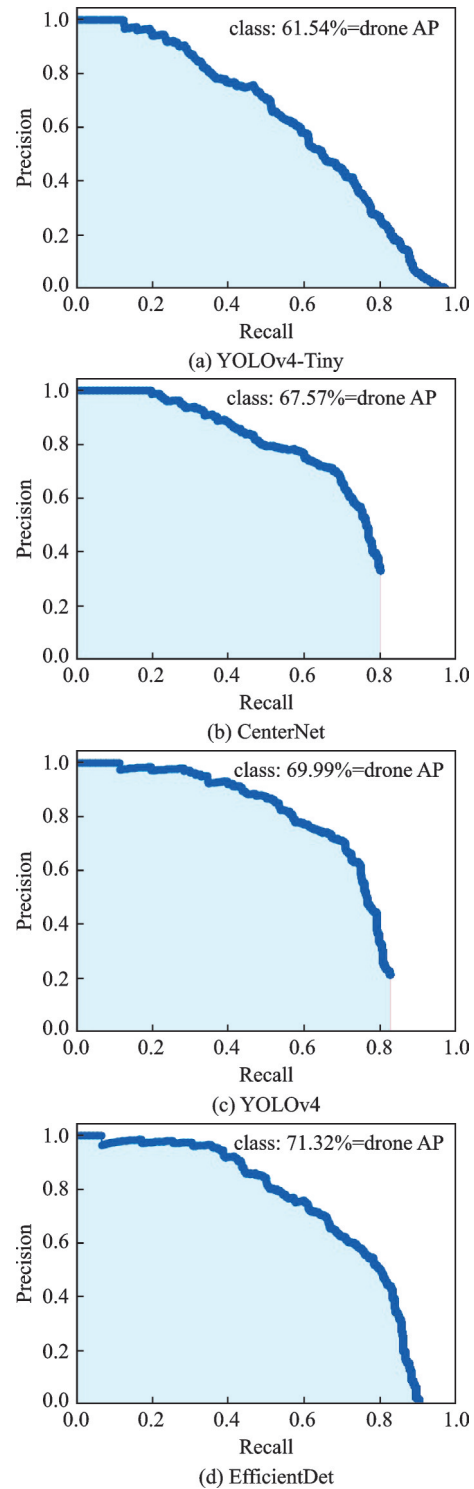


Fig.22 PR curves obtained from Experiment 1

Table 2 AP50, FAR & MAR (Remote), and FAR & MAR (Complex backgrounds) of four models for UAV detection in Experiment 1 %

Model	AP50	FAR & MAR (Remote)	FAR & MAR (Complex backgrounds)
YOLOv4-Tiny	61.54	35.11	36.53
CenterNet	67.57	27.66	28.47
YOLOv4	69.99	24.85	26.26
EfficientDet	71.32	21.43	24.62

3.1.2 Experiment 2

The dataset used in Experiment 2 consists of 7 010 original UAV images the same as Experiment 1, 7 000 bird images, and 7 000 kite images. The PR curves of the four models are shown in Fig.23.

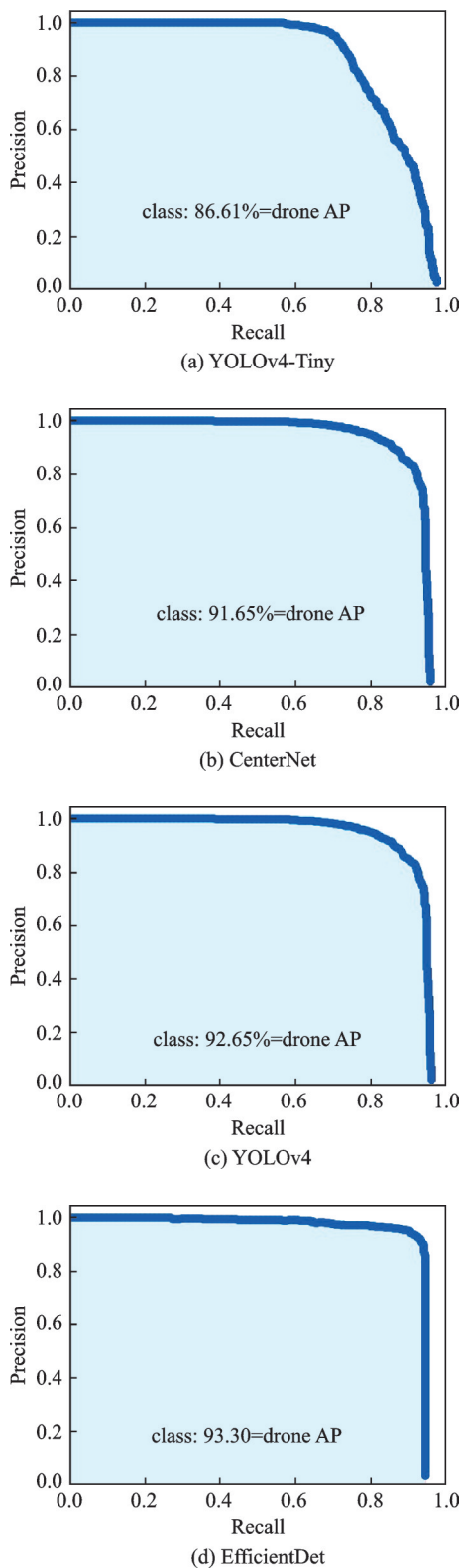


Fig.23 PR curves obtained from Experiment 2

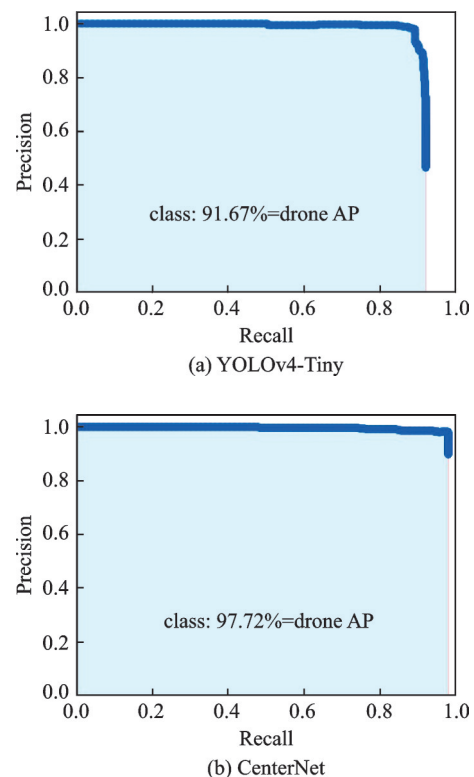
Finally, the detection accuracy of UAV, the FAR, and the MAR under long distance and complex backgrounds are shown in Table 3.

Table 3 AP50, FAR & MAR (Remote), and FAR & MAR (Complex backgrounds) of four models for UAV detection in Experiment 2 %

Model	AP50	FAR & MAR (Remote)	FAR & MAR (Complex backgrounds)
YOLOv4-Tiny	86.61	18.72	26.78
CenterNet	91.65	11.36	13.66
YOLOv4	92.65	10.65	12.76
EfficientDet	93.30	5.62	6.55

3.1.3 Experiment 3

The dataset used in Experiment 3 consists of 46 571 UAV images composed of original UAV images and UAV images generated by data enhancement, 43 001 bird images, and 44 340 kite images. The PR curves of the four models are shown in Fig.24. Finally, the detection accuracy of UAV, the FAR, and the MAR under long distance and complex backgrounds are shown in Table 4.



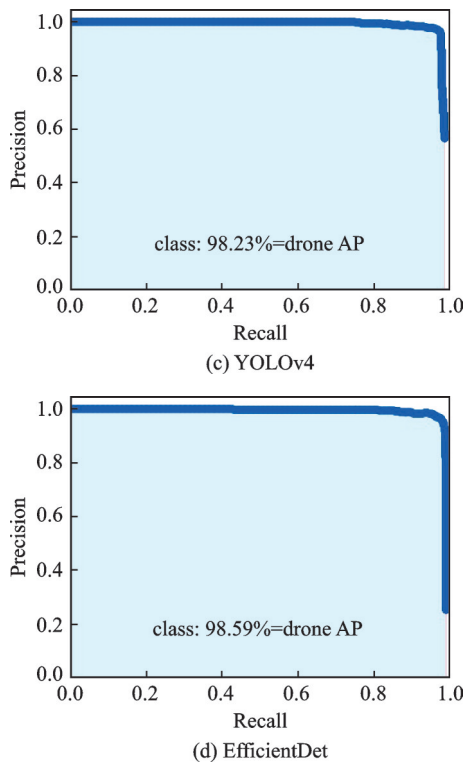


Fig.24 PR curves obtained from Experiment 3

Table 4 AP50, FAR & MAR (Remote), and FAR & MAR (Complex backgrounds) of four models for UAV detection in Experiment 3 %

Model	AP50	FAR & MAR (Remote)	FAR & MAR (Complex backgrounds)
YOLOv4-Tiny	91.67	15.33	19.67
CenterNet	97.72	8.63	10.34
YOLOv4	98.23	7.75	8.71
EfficientDet	98.59	4.65	4.92

3.2 Experimental results and analysis

The accuracy of ATAUAV and RWQUAV datasets for UAV detection is given in Table 1. The data are extracted from the corresponding papers.

3.2.1 Experiment 1

It can be seen from Fig.22 and Table 2 that in long distance and complex backgrounds, the accuracy of UAV detection is not quite high, and the highest AP50 is from EfficientDet, about 71.32%. However, it is not particularly low as compared with Refs.[14-15]. In fact, the AP50 is even higher than that in Ref.[15] under normal circumstances.

3.2.2 Experiment 2

It can be seen from Fig.23 and Table 3 that the AP50 of four models for UAVs is much higher than

that of Experiment 1. It is because a large number of bird and kite images are added, which are highly similar to UAVs in long distance and complex backgrounds, and the FAR and MAR of UAV detection in long distance and complex backgrounds are significantly reduced. Although the dataset used is not expanded by data enhancement, the above four models achieve higher AP50 of UAV detection than those of ATAUAV and RWQUAV datasets, and significantly reduce the FAR and MAR of UAV detection in long distance and complex backgrounds. Even YOLOv4-Tiny, the simplest of the four models, still reaches the AP50 of 86.61% for UAV detection.

3.2.3 Experiment 3

From Fig.24 and Table 4, it can be seen that the AP50 of the four models for UAV detection is further improved by using the dataset after the use of data enhancement. By comparing the results of Experiment 2, YOLOv4-Tiny, CenterNet, and YOLOv4 are enhanced by about 5%. It shows that after data enhancement, the three models are learned more image detail features, which helps these models better detect UAVs in long distance and complex backgrounds. While the improvement of EfficientDet model is not obvious, about 0.16%, showing that the data enhancement effect is not obvious in the model with complex network structure, such as EfficientDet. In other words, from the experimental results, for the specific target of UAV, when using the model with simple network structure, the use of data enhancement can indeed improve the accuracy to a certain extent, and reduce the FAR and MAR. However, when the model network is complex and large enough, the effect of data enhancement is not significant.

The successful application of data enhancement methods essentially increases data by introducing prior knowledge, which reduces the structural risk of the model and improves its generalization ability. Specifically, various methods of data enhancement are used to express many images correlating with the original UAV images, which cannot be obtained by real world shooting.

For example, flipping and rotation can help the model better detect UAVs flying in abnormal pos-

tures. Random Gaussian or Salt Pepper noise can more closely mimic the image of UAVs flying on rainy days. Gauss blur can improve the model's accuracy in detecting UAVs when camera lens blurring occurs with irresistible factors. In affine transformations, its random rotation and the compression and pulling of the image at random latitudes can significantly change UAV images appearance and better help the model detect UAVs based on local features when the part of UAVs is not clear. The in-

tensity and contrast enhancements are essentially based on gray-scale changes in UAV images, which can help the model better exclude the impact of gray-scale changes in images caused by lighting and weather factors on the accuracy.

3.3 Visualization results

The EfficientDet model with D7 backbone is used to randomly select some original or enhanced images, including some or all of classification. The detection effect is shown in Fig.25.



Fig.25 Final object detection results of using EfficientDet

4 Conclusions

A dataset dedicated to UAV detection is built with high standard and performance. By incorporating a large number of UAV images in long distance and complex backgrounds, the problem of high

FAR and MAR in the existing UAV datasets can be avoided. The addition of noise with secondary jamming objects can further improve the generalization ability of models, and avoid the problem of over-fitting of models caused by the existing UAV datasets containing only single target UAV images, and the

accuracy of UAV detection is improved. Moreover, an effective way to obtain UAV images is provided, and the difficulty and cost of UAV image acquisition are dramatically reduced. Through the mainstream target detection models and a large number of comparative experiments, the effectiveness of the proposed dataset is verified. So in the field of UAV detection in the future, there is a truly usable high standard and high-performance special dataset. With the proposed new methods of data enhancement, we believe that our dataset can be further expanded and will achieve better performance in deep learning models.

References

- [1] SHI X, YANG C, XIE W, et al. Anti-drone system with multiple surveillance technologies: Architecture, implementation, and challenges[J]. *IEEE Communications Magazine*, 2018, 56(4): 68-74.
- [2] PISA S, PIUZZI E, PITTELLA E, et al. Evaluating the radar cross section of the commercial IRIS drone for anti-drone passive radar source selection[C]//*Proceedings of 2018 22th International Conference on Microwave and Radar Conference (MIKON)*. Poznan, Poland: IEEE, 2018: 699-703.
- [3] WANG C, WANG T, WANG E, et al. Flying small target detection for anti-UAV based on a Gaussian mixture model in a compressive sensing domain[J]. *Sensors (Basel, Switzerland)*, 2019, 19(9): 2168-2183.
- [4] GARCIA G A, ORTS E S, OPREA S, et al. A review on deep learning techniques applied to semantic segmentation[J]. *arXiv: 1704.06857*, 2017: 1-23.
- [5] SUN C, SHRIVASTAVA A, SINGH S, et al. Revisiting unreasonable effectiveness of data in deep learning era[C]//*Proceedings of 2017 IEEE International Conference on Computer Vision (ICCV)*. Venice, Italy: IEEE, 2017: 843-852.
- [6] JOSEPH R, SANTOSH D, ROSS G, et al. You only look once: Unified, real-time object detection[C]//*Proceedings of 2016 IEEE International Conference on Computer Vision and Pattern Detection (CVPR)*. Las Vegas, NV, USA: IEEE, 2016: 779-788.
- [7] REDMON J, FARHAD A. YOLOv3: An incremental improvement[C]//*Proceedings of 2018 IEEE International conference on Computer Vision and Pattern Detection (CVPR)*. Washington, USA: IEEE, 2018: 2767-2773.
- [8] JOSEPH R, ALI F. YOLO9000: Better, faster, stronger[C]//*Proceedings of 2017 IEEE International Conference on Computer Vision and Pattern Detection (CVPR)*. Honolulu, HI, USA: IEEE, 2017: 6517-6525.
- [9] BOCHKOVSKIY A, WANG C Y, LIAO H. YOLOv4: Optimal speed and accuracy of object detection[C]//*Proceedings of 2020 IEEE/CVF International Conference on Computer Vision and Pattern Detection (CVPR)*. Taiwan, China: IEEE, 2020: 10934-10951.
- [10] TAN M, PANG R, LE Q V. EfficientDet: Scalable and efficient object detection[C]//*Proceedings of 2020 IEEE/CVF International Conference on Computer Vision and Pattern Detection (CVPR)*. Seattle, WA, USA: IEEE, 2020: 10781-10790.
- [11] LIN T Y, MAIRE M, BELONGIE S, et al. Microsoft COCO: Common objects in context[C]//*Proceedings of European Conference on Computer Vision*. [S.l.]: Springer, 2014: 740-755.
- [12] EVERINGHAM M, GOOL L V, WILLIAMS C, et al. The pascal visual object classes (VOC) challenge[J]. *International Journal of Computer Vision*, 2010, 88(2): 303-338.
- [13] DENG J, DONG W, SOCHER R, et al. ImageNet: A large-scale hierarchical image database[C]//*Proceedings of 2009 IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.]: IEEE, 2009: 248-255.
- [14] CINTAS E, OZYER B, SIMSEK E. Corrections to "vision-based moving UAV tracking by another UAV on low-cost hardware and a new ground control station"[J]. *IEEE Access*, 2021, 9: 6888-6889.
- [15] PAWELCZYK M L, WOJTYRA M. Real world object detection dataset for quadcopter unmanned aerial vehicle detection[J]. *IEEE Access*, 2020, 8: 174394-174409.
- [16] HUSSAIN Z, GIMENEZ F, YI D, et al. Differential data augmentation techniques for medical imaging classification tasks[J]. *AMIA Annual Symposium Proceedings*, 2017: 979-984.
- [17] CUBUK E D, ZOPH B, MANE D, et al. AutoAugment: Learning augmentation strategies from data[C]//*Proceedings of 2019 IEEE/CVF International Conference on Computer Vision and Pattern Detection (CVPR)*. Long Beach, CA, USA: IEEE, 2019: 113-123.
- [18] YU X, WU X, LUO C, et al. Deep learning in remote sensing scene classification: A data augmentation enhanced convolutional neural network framework[J]. *GIS Science and Remote Sensing*, 2017, 54(5): 741-758.
- [19] ZHANG X X, YANG Y M. Common color space and its conversions in color image project[J]. *Computer*

- Engineering and Design, 2008, 29(5): 1210-1212.
- [20] GALDRAN A, ALVAREZ G A, MEYER M I, et al. Data-driven color augmentation techniques for deep skin image analysis[C]//Proceeding of 2017 IEEE International Conference on Computer Vision and Pattern Detection (CVPR). Potro, Spain: IEEE, 2017: 3702-3706.
- [21] ZHANG H, CISSE M, DAUPHIN Y N, et al. mix-up: Beyond empirical risk minimization[C]//Proceedings of International Conference on Learning Representations.[S.l.]:[s.n.], 2018: 9412-9425.
- [22] YUN S, HAN D, OH S J, et al. Cutmix: Regularization strategy to train strong classifiers with localizable features[C]//Proceedings of the IEEE International Conference on Computer Vision.[S.l.]: IEEE, 2019: 6023-6032.
- [23] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[C]//Proceedings of International Conference on Learning Representations.[S.l.]:[s.n.], 2015: 6434-6450.
- [24] KAI W D, SONG B, LING X X, et al. CenterNet: Keypoint triplets for object detection[C]//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision (ICCV). Seoul, Korea: IEEE, 2019: 6569-6578.

Acknowledgements The work was supported by the National Natural Science Foundation of China (No. 62173237), the National Key R&D Program of China (No.2018AAA0100804), the Zhejiang Key laboratory of General Aviation Operation technology (No.JDGA2020-7), the Talent Project of Revitalization Liaoning (No. XLYC1907022), the Key R & D Projects of Liaoning Province (No. 2020JH2/10100045),

the Natural Science Foundation of Liaoning Province (No. 2019-MS-251), the Scientific Research Project of Liaoning Provincial Department of Education (No.JYT2020142), the High-Level Innovation Talent Project of Shenyang (No.RC190030), the Science and Technology Project of Beijing Municipal Commission of Education (No. KM201811417005), and the Academic Research Projects of Beijing Union University (No.ZB10202005).

Authors Dr. WANG Zhi received the Ph. D. degree in mechanical engineering from Beijing University of Technology in 2020. He is currently a lecturer in Civil Aviation Management Institute of China. His current research interests mainly concern unmanned aerial vehicle operation and route optimization.

Prof. WANG Ershen received the Ph.D. degree from Dalian Maritime University in 2009. He is a Professor of College of Electronic and Information Engineering in Shenyang Aerospace University. His research currently focuses on global navigation satellite system (GNSS) positioning algorithms, and UAV surveillance technology.

Author contributions Dr. WANG Zhi designed the dataset and wrote the manuscript. Mr. HU Wei contributed to the background and discussion of the results. Prof. WANG Ershen contributed to the improvement of the model and interpretation of the detection results. Dr. HONG Chen contributed to proposing the detection model. Mr. XU Song contributed to data analysis. Ms. LIU Meizhi contributed to the background. All authors commented on the manuscript draft and approved the submission.

Competing interests The authors declare no competing interests.

(Production Editor: WANG Jing)

低空无人机探测增强数据集研究

王志^{1,2}, 胡威³, 王尔申³, 宏晨⁴, 徐嵩³, 刘梅芷⁵

(1. 浙江建德通用航空研究院, 建德311612, 中国; 2. 中国民航管理干部学院通用航空系, 北京100102, 中国; 3. 沈阳航空航天大学电子信息工程学院, 沈阳110136, 中国; 4. 北京联合大学机器人学院, 北京100101, 中国; 5. 沈阳航空航天大学人工智能学院, 沈阳110136, 中国)

摘要:近年来,与无人机有关的事故显著增加,使得对反无人机系统研究有着迫切的需求。在对低空无人机探测的方法中深度学习算法应用具有较大潜力,但目前缺乏高质量的低空无人机探测数据集。本文研究建立了一种包含无人机和其他干扰目标的综合增强高质量数据集。通过实际场景拍摄、网络爬虫和数据增强方法获得大量高清无人机图像。并且,为解决复杂场景和远距离的低空无人机探测以及干扰目标所造成的较高误检率和漏检率问题,在数据集中增加了具有干扰特征的目标。最后,通过4种主流深度学习模型对所建立的数据集进行训练、验证和测试。结果表明,建立的高质量数据集可显著提高无人机探测的准确率。该研究为反无人机系统研究以及无人机探测模型优化提供了一定的验证数据集。

关键词:无人机; 无人机数据集; 目标探测; 深度学习