

## Research Article

# A UAV Detection and Tracking Algorithm Based on Image Feature Super-Resolution

Bin Li <sup>1</sup>, Shi Qiu <sup>2</sup>, Wei Jiang <sup>1</sup>, Wei Zhang <sup>1</sup> and Mingnan Le <sup>1</sup>

<sup>1</sup>School of Information Science and Technology, Northwest University, Xi'an 710127, China

<sup>2</sup>Key Laboratory of Spectral Imaging Technology CAS, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China

Correspondence should be addressed to Shi Qiu; [qiushi215@163.com](mailto:qiushi215@163.com)

Received 10 November 2021; Revised 7 January 2022; Accepted 21 January 2022; Published 9 February 2022

Academic Editor: Kun Wei

Copyright © 2022 Bin Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

UAV is difficult to detect by visual methods at a long distance, so a UAV detection and tracking algorithm is proposed based on image super-resolution. Firstly, a saliency transformation algorithm is built to focus on the suspected area. Then, a generative adversarial network is established on the basis of ROI to realize the super-resolution of weak targets and restore the high-resolution details of target features. Finally, the cooperative attention module is built to recognize and track UAV. Our experiments show that the proposed algorithm has strong robustness.

## 1. Introduction

Small UAV has the characteristics of portability and strong mobility and also has broad application space in unmanned investigation. However, for long-distance UAV, only a limited number of pixels are displayed on the image, which is of great significance for its accurate discrimination [1]. The research on UAV mainly focuses on target tracking. Ibrahim et al. [2] use UAV to track moving targets. Zhou et al. [3] construct a Kalman filter to realize UAV tracking. Nodland et al. [4] propose the track optimization strategy. Liu [5] introduces prediction points to assist UAV track detection. Ragi and Chong [6] propose an algorithm to dynamically realize multi-UAV tracking. Yoo and Hong [7] realize UAV detection and tracking by visual means. Kadouf and Mustafah [8] analyze the color characteristics of UAV to realize UAV tracking. Yu et al. [9] propose a new coordinate system to analyze the attitude of UAV. Teuliere et al. [10] build a 3D model to track the UAV flying indoors. Choi and Kim [11] use monocular to analyze the UAV track. Quintero et al. [12] use the output-feedback model to predict UAV flight trajectory. Santos et al. [13] build a ground visual tracking system to detect UAV. Zhou et al. [14] construct a Hough transform to detect and track UAV. Vetrella et al. [15] realize

dynamic navigation through a multi-UAV network. Elloumi et al. [16] propose a low-power tracking algorithm from the perspective of UAV energy. Greatwood et al. [17] use parallel means to realize rapid detection and tracking of UAV. Santos et al. [18] propose a 3D model for UAV positioning. Zhang et al. [19] use deep learning to build a coarse to fine detection algorithm to realize UAV detection and tracking. Huang et al. [20] build correlation filters based on deep learning to realize UAV tracking. Rabah et al. [21] build a model based on fuzzy set theory to realize target tracking. Kokunko and Krasnova [22] propose variable constraint mechanism to realize UAV tracking. Li et al. [23] propose augmented memory for correlation filters to realize UAV tracking. Li et al. [24] use deep learning network to extract features and realize UAV object tracking. Moon et al. [25] realize multi-UAV tracking based on deep learning network.

Through the above analysis, main problems of many research studies on UAV detection and tracking are as follows. (1) Traditional target detection algorithms cannot be effectively applied to UAV due to the small size of UAV. (2) The number of pixels on the image of UAV is limited, and its features are not obvious. (3) The flight uncertainty of UAV leads to the construction difficulty of the unified model.

Therefore, on the basis of images, (1) a complete UAV detection and tracking process is proposed, (2) the enhancement algorithm is built to improve the spatial resolution of the target and a new idea of UAV detection is proposed, and (3) according to the principle of visual perception, a depth-based attention model is built to focus on the area, where UAV is located to realize tracking.

## 2. Algorithm

According to the characteristics of UAV images, a UAV detection and tracking algorithm is proposed based on image super-resolution, as shown in Figure 1. Firstly, the image composition is analyzed and the suspected area extraction module is constructed. Then, a super-resolution model is constructed to highlight local information and increase signal strength. Finally, a deep learning module is constructed to realize UAV tracking based on the attention mechanism.

$$E_i(C; M_t^i) = - \left[ \sum_{j \in X^+} \log I_f(y_j = 1 | C; M_t^i) + \sum_{j \in X^-} \log I_f(y_j = 0 | C; M_t^i) \right], \quad (2)$$

where  $X^+$  and  $X^-$  represent the edge and background pixels of the salient target area, respectively. Image segmentation is realized through the edge of the salient target area, but the resolution, affected by the multilayer transmission architecture, will gradually decrease with the change of layers, which needs to be further strengthened:

$$C_i = O_b [C_{e,i}; \theta] + f_h(C_{i+1}), \quad (3)$$

where  $f_h(C_{i+1})$  represents the enhancement processing of the features of the previous layer. The cumulative loss function between the corresponding estimation result and the real image is

$$E_T = \sum_i \{E_i(C_{e,i}; M_t^i) + E_i(C_i; M_t^i)\}. \quad (4)$$

The idea of edge enhancement is used to suppress the weakening caused by multiple layers in the process of feature fusion, in order to achieve more accurate salient target area estimation. During the convolution operation, the feedback mechanism is introduced to continuously input the edge features and salient target area features of the image into the convolution operation to obtain a new estimation graph function:

$$I_t = \begin{cases} U(R_1(I(x, y); \theta), R_2(I^P(x, y); \theta); q), & t = 1, \\ U(R_1(I(x, y); \theta), R_2(I^{t-1}(x, y); \theta); q), & t > 1, \end{cases} \quad (5)$$

where  $U(\cdot)$  is deconvolution,  $I^P(x, y)$  is a priori graph, and  $q$  is convolution parameter. Divergence occurs during feedback correction. In order to solve this phenomenon, any two pixels adjacent to each other can be used as a path in  $I(x, y)$ . Let all paths corresponding to  $n$  pixels be expressed as  $q =$

*2.1. ROI Extraction.* The UAV is usually at a long distance, so it presents a limited number of pixels on the image, which leads to difficult detection from the spatial domain. Therefore, the salient area is introduced to extract ROI. First of all, the features are obtained by linear difference:

$$C_f = p_1 + O_b [f_r(R(f_\theta))], \quad (1)$$

where  $p_1$  is the characteristic,  $O_b$  is the gradient third-order matrix,  $f_r$  is the activation function, and  $R(f_\theta)$  is the convolution calculation performed by  $\theta$ . The salient target area model is constructed by supervised learning, and the image is described as  $I_f(y_j = 1 | M_t^i)$ ; then,  $y_j$  represents the reliability of pixel  $j$ . The corresponding cross-entropy loss function is

$\{q_0, \dots, q_n\}$ . Then, after obtaining the salient target edge, the loss function is calculated as

$$L(x) = \min_{q,x} \left\{ \sum_{i=1}^n |C(p_{i-1}) - C(p_i)| \right\}. \quad (6)$$

Let the width and height of  $L(x)$  be  $w$  and  $h$ . The mapping matrix  $M = [m_{ij}]_{w \times h}$  can be constructed:

$$m_{ij} = 1 - \frac{\sqrt{(i - 0.5w)^2 + (j - 0.5h)^2}}{\sqrt{(0.5w)^2 + (0.5h)^2}}. \quad (7)$$

According to the mapping matrix, the position of the salient target can be determined. On the basis of a priori information feedback correction, the position of the edge of the salient target can be distinguished to prevent the performance degradation of the feedback correction and realize the suspected area extraction.

*2.2. ROI Super-Resolution Reconstruction.* The motion area has been extracted in the previous section. On this basis, the research on ROI super-resolution reconstruction is carried out in this section to further determine whether UAV is included. Because the target area is smaller than the background area, when the target area is enlarged, the background is enlarged at the same time, resulting in coarse-grained information after over division. Therefore, we design the feature super-resolution GAN, transform the weak target features into super-resolution features through super-resolution processing in the feature space, and enhance the feature representation of the weak target. The structure is shown in Figure 2.

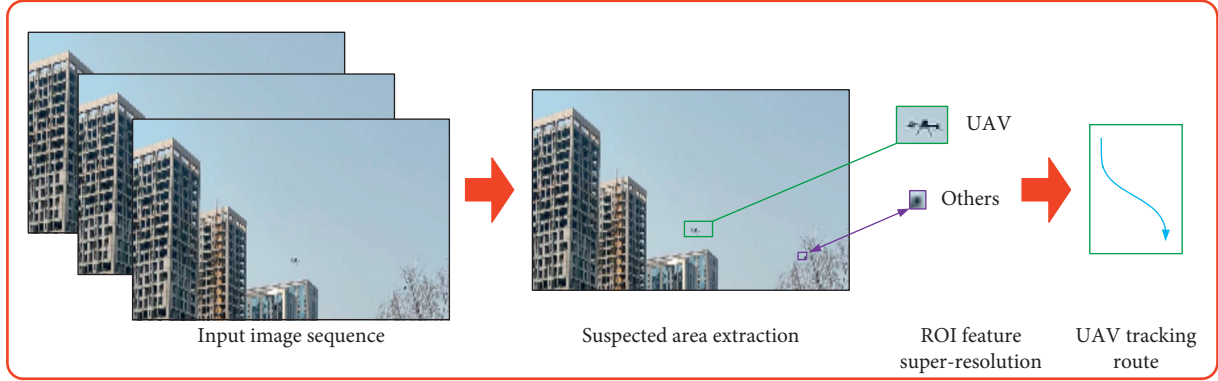


FIGURE 1: Algorithm flowchart.

Let the original input image be  $\mathbf{X}_I$ . The image  $\mathbf{X}_{0.5I}$  is obtained by 2 times downsampling to obtain a pair of low-resolution and high-resolution target features. The output features  $\mathbf{F}_I = \{\mathbf{F}_I^1, \mathbf{F}_I^2, \dots, \mathbf{F}_I^q\}$  and  $\mathbf{F}_{0.5I} = \{\mathbf{F}_{0.5I}^1, \mathbf{F}_{0.5I}^2, \dots, \mathbf{F}_{0.5I}^q\}$  are obtained through the feature network. The weak target features  $\mathbf{F}_{0.5I}^q$  are iteratively generated into super-resolution features  $\mathbf{S}_{0.5I}^q$  so that the super-resolution features are similar to the features  $\mathbf{T}_I^q$  output by the supervisor as much as possible. The feature loss function is defined as

$$L_F = \sum_{q=1}^C \|\mathbf{T}_I^q - \mathbf{S}_{0.5I}^q\|_2^2. \quad (8)$$

The generator loss function is defined as

$$L_G = - \sum_{q=1}^C \log D(\mathbf{S}_{0.5I}^q). \quad (9)$$

The feature super-resolution discriminator adopts a three-layer perceptron to train and distinguish  $\mathbf{S}_{0.5I}^q$  and  $\mathbf{T}_I^q$ . The discriminator loss function is defined as

$$L_D = - \sum_{q=1}^C [\log D(\mathbf{T}_I^q) + \log(1 - D(\mathbf{S}_{0.5I}^q))]. \quad (10)$$

The feature super-resolution supervisor extracts the high-resolution target feature  $\mathbf{T}_I$  similar to the low-resolution input feature as the supervision signal for super-resolution model training, in order to enhance the stability of training and improve the quality of super-resolution. In order to avoid the inconsistency between the receptive fields of low-resolution features and high-resolution features, a feature extraction backbone-shared parameter network is designed to extract the  $q$ th feature  $\mathbf{T}_I^q$  which is more suitable for training the super-resolution model without adding parameters.

**2.3. UAV Tracking.** Through the online recognition network, we can enhance the discriminative power of the classifier to distinguish the target from other interfering objects in the background, minimize the false detection rate, and complete the rough positioning of the target. We use the

depth regression structure to construct the network structure, as shown in Figure 3, which uses two-layer convolution neural network:

$$f(x, w) = \varphi_2\{w_2 * \varphi_1(w_1 * x)\}, \quad (11)$$

where  $x$  is the countermeasure network feature graph generated by feature super-resolution,  $w_1$  and  $w_2$  represent the weight of the convolution layer,  $*$  is a convolution operation, and  $\varphi$  is the activation function. The loss function is defined as

$$L(w) = \sum_{j=1}^m \gamma_j \|f(x_j; w) - y_j\|^2 + \sum_k \lambda_k \|w_k\|^2, \quad (12)$$

where  $m$  is the total number of feature graph samples,  $\gamma_j$  is the learning weight,  $y_j$  is the regression classification confidence of every feature sample  $x_j$ , and  $\lambda_k$  is a regular term. Gauss-Newton algorithm is used to solve the problem.

We transform the target tracking task into a similarity measurement problem and take the first frame  $z$  and the candidate region  $x$  of subsequent frames as the input images of template branch and detection branch, respectively. Feature extraction network based on weight sharing  $\varphi(\cdot)$  maps to the feature space, and the metric function  $f(z, x)$  is learnt to compare the similarity between the template image and the candidate area search image. Finally, return the response graph. In order to highlight the importance of different spaces, a spatial collaborative attention module is designed.

Channel attention models the dependencies between channels, learns the association between features from the semantic level, optimizes features, activates feature channels more related to the target, and removes redundant features. Suppose that the feature diagram extracted by MobileNetV2 network template branch and detection branch is  $\varphi(z)$  and  $\varphi(x)$ . A typical MobileNetV2 network template is shown in Figure 4. The global information of each channel is obtained by global average pooling and condensing spatial dimensions to provide salient target features. Input the results to the input layer, hidden layer, and output layer. We reduce the number of channels in the hidden layer to 1/16 of the input layer. Output channel attention weight  $A_c\{\varphi(z)\}$  and  $A_c\{\varphi(x)\}$ . Finally, the channel attention feature graph is

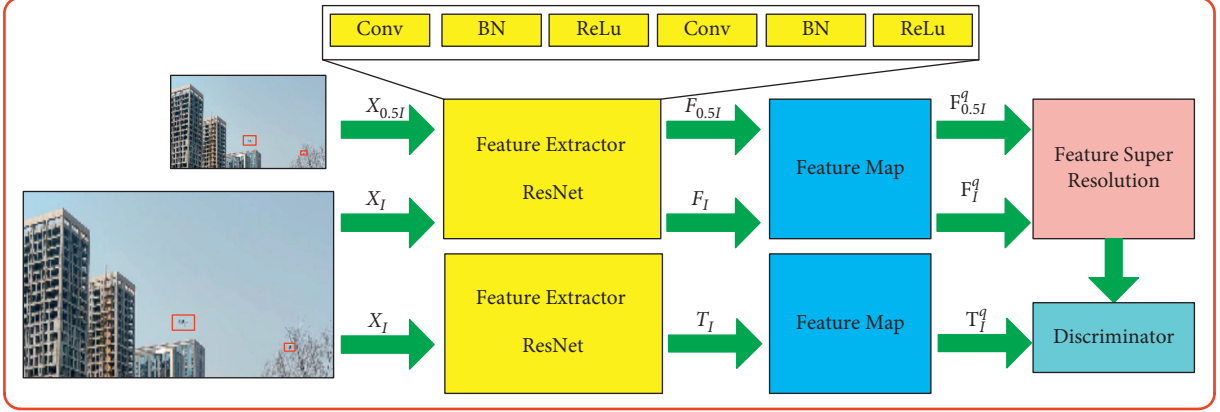


FIGURE 2: Feature super-resolution countermeasure GAN.

obtained by point multiplication with the input feature  $\varphi_c(z)$  and  $\varphi_c(x)$ .

In the collaborative attention module, each branch code is integrated into another branch to make full use of the background information. In order to facilitate matrix multiplication with features, the output collaborative attention weight  $A\{\varphi(x)\}$  and  $A\{\varphi(z)\}$ . After passing through the channel attention and collaborative attention modules, the weights of the two branches are fused to obtain  $\varphi'(z)$  and  $\varphi'(x)$ .

Spatial attention focus is used to describe the position, which can construct the relationship between different positions in the feature graph, and supplement the channel attention through position weighted fusion. Feature graph  $\varphi'(\cdot)$  compresses along the channel dimension to obtain the spatial attention weight  $As(\cdot)$  and then obtains the final attention feature graph.

The feature graphs of each layer output by template branch and detection branch in the network are normalized by adjusting the convolution operation of layers, in order to make the feature graph with uniform resolution and the same number of channels.

Let the  $q$ th layer adjusted feature graph of classified branch input be  $\varphi_{cls}^q(z)$  and  $\varphi_{cls}^q(x)$ . The adjusted feature diagrams of the  $q$ th layer of regression branch input are  $\varphi_{reg}^q(z)$  and  $\varphi_{reg}^q(x)$ . Finally, the output is weighted and fused:

$$\begin{cases} A_{cls}^q = \varphi_{cls}^q(x) * \varphi_{cls}^q(z), \\ A_{reg}^q = \varphi_{reg}^q(x) * \varphi_{reg}^q(z). \end{cases} \quad (13)$$

When classifying the foreground or background of each candidate area, the same target may exist in multiple overlapping rectangular boxes at the same time, so non-maximum suppression (NMS) is used for elimination to accurately track the target.

### 3. Experiment and Result Analysis

The experimental data include 20 groups of UAV visible light data from far to near, as shown in Figure 5, with an image resolution of  $1024 \times 1024$ .

Based on the network structure, the image is normalized to  $512 \times 512$  in order to ensure that small targets are not lost. On Win10 operating system with Intel<sup>®</sup> Core<sup>™</sup> I5-6500 CPU, 3.20ghz system, the proposed program is run by 8 frames/s, which cannot meet practical requirement. However, due to the continuity of the target, the images can be processed at an interval of 1 frame, which can achieve near real-time operation speed.

*3.1. ROI Extraction Algorithm Performance.* We introduce the following indicators to measure the algorithm performance [26], as shown in Table 1:

$$\begin{cases} SEN = \frac{TP}{TP + FN}, \\ SPE = \frac{TN}{TN + FP}, \\ ACC = \frac{TP + TN}{TP + FP + TN + FN}, \\ FPF = 1 - ACC, \end{cases} \quad (14)$$

where SEN reflects the detection performance of the algorithm for real targets, SPE reflects the detection performance of the algorithm for false targets, ACC reflects the ratio of correct test results to all samples in the test results, and FPF reflects the ratio of false test results diagnosed as true targets.

The experimental results are shown in Table 2. The color model built in [8] has a good detection effect on UAVs with close range and obvious color characteristics. However, for long-distance UAVs, it is difficult to obtain color information and results in poor effect. Based on the target composition structure, Zhou et al. [14] construct the model through texture features, which is more stable than the color model, and the effect is significantly improved. However, for specific UAVs, the detection rate is limited. Zhang et al. [19] construct a deep learning network based on the deep reinforcement learning (DRL) model to realize ROI detection. It is the current mainstream target detection algorithm, and the effect and performance are further improved. However,

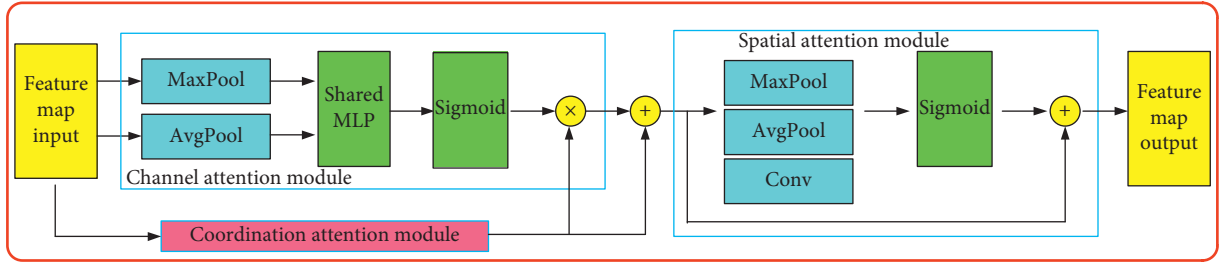


FIGURE 3: Spatial collaborative attention module.

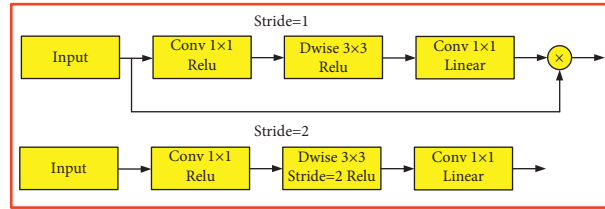


FIGURE 4: Convolution module block diagram.

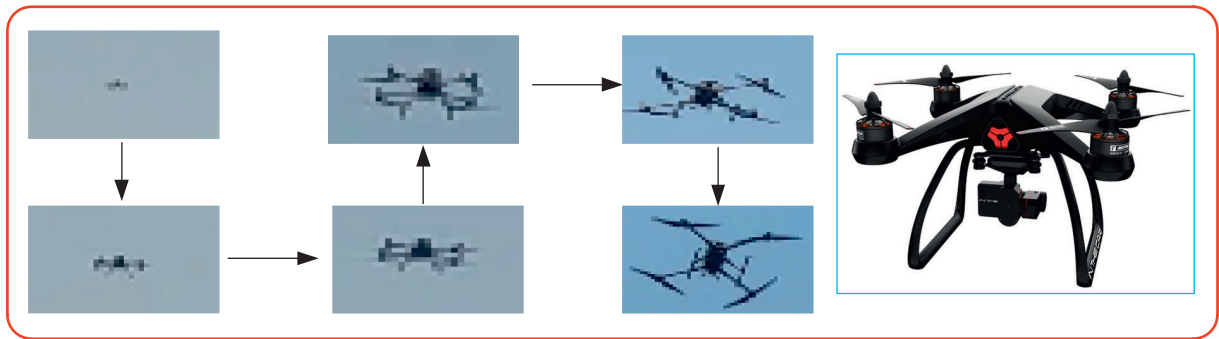


FIGURE 5: UAV data.

TABLE 1: Test result indicators.

TP (True positive)	FP (False positive)
FN (False negative)	TN (True negative)

this process needs to traverse the global image, and the computational cost is high and cannot meet the detection of targets with different scales. Our proposed algorithm introduces saliency region (SR) and focuses on the region of interest step by step. It conforms to the principle of visual perception and uses a priori knowledge to extract ROI. ACC has reached 95. For areas with too small area, there is still a risk of missing detection. On the basis of SR super-resolution module is added to establish the relationship between low-resolution and high-resolution target features, which further improves the detection of small targets.

**3.2. UAV Tracking Algorithm Performance.** We introduce the tracking success rate curve to intuitively show the algorithm performance, as shown in Figure 6. For short-range UAV tracking, all algorithms have achieved good results because of the high-resolution of the target displayed on the image. With the increase of distance, the performance of the

TABLE 2: Detection effect.

Algorithm	SEN	SPE	ACC	FPF
Color	73	21	76	24
Texture	82	20	86	14
DRL	88	12	90	10
SR	91	9	92	8
SR + UR	93	5	95	5

algorithm decreases. Kalman filter algorithm is the most obvious. Because the size of the target changes greatly in the image, the tracking is easy to be affected by the surrounding environment. According to the difference between UAV and background characteristics, fuzzy set constructs a segmentation algorithm to realize target tracking and has certain robustness to target size. Due to the UAV flying at low speed and uniform speed, the model updating is stable, and all algorithms have good tracking effect. However, in the face of turning flight or sudden acceleration or deceleration, Kalman filter and fuzzy set algorithm will not track due to the limitation of model updating speed. Augmented memory for correlation filters has achieved good results in analyzing short-time flight states. In depth network, UAV information

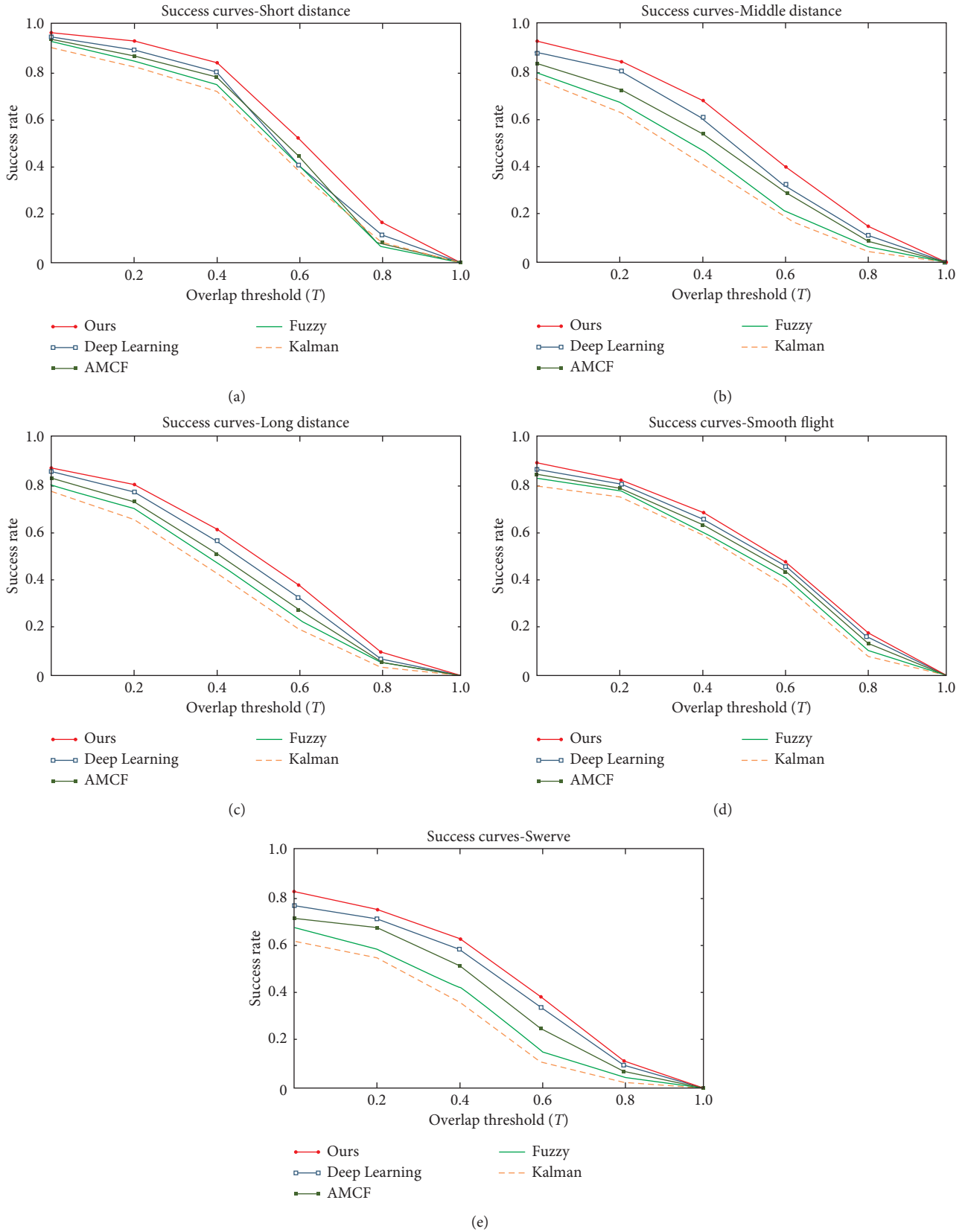


FIGURE 6: Successful ratio curves of different algorithms.

is obtained through a large number of training samples to build the model, but the time dimension information is not used, resulting in some limitations of the algorithm. However, the proposed algorithm introduces the spatial collaborative attention module to focus the target hierarchically to achieve target tracking, which decreases slowly, and the performance is better than other algorithms.

#### 4. Conclusion

Aiming at the difficulty of visual detection and tracking of long-distance UAV, a complete set of weak and small UAV detection process is proposed from the perspective of visual cognition. The ROI area is focused step by step to establish and generate the GAN according to the idea of image super-resolution. The target details are restored to highlight the characteristics of weak targets, and a collaborative attention module is built to identify and track UAVs. The algorithm can be applied to fixed cameras, and the region of UAV can be further determined by the difference between frames. However, the proposed algorithm can also be applied to mobile cameras to focus the UAV area according to the saliency area. It can provide a new idea for the detection and recognition of weak and small targets.

#### Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

#### Conflicts of Interest

The authors declare that they have no conflicts of interest.

#### Acknowledgments

This work was supported by the China Postdoctoral Science Foundation (program no. 2020M683696XB), Natural Science Basic Research Plan in Shaanxi Province of China (program no. 2021JQ-455), and Science and Technology Rising Star of Shaanxi Youth (program no. 2021KJXX-61).

#### References

- [1] C. Fu, F. Ding, Y. Li, J. Jin, and C. Feng, "DR 2 track: towards real-time visual tracking for UAV via distractor repressed dynamic regression," in *Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1597–1604, IEEE, Las Vegas, NV, USA, October 2020.
- [2] A. Ibrahim, P. Ching, G. Seet, W. Lau, and W. Czajewski, "Moving objects detection and tracking framework for UAV-based surveillance," in *Proceedings of the 2010 Fourth Pacific-Rim Symposium on Image and Video Technology*, pp. 456–461, IEEE, Singapore, November 2010.
- [3] Q. Zhou, Y. Zhang, Y. Qu, and C. Rabbath, "Dead reckoning and kalman filter design for trajectory tracking of a quadrotor UAV," in *Proceedings of the 2010 IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications*, pp. 119–124, IEEE, QingDao, China, July 2010.
- [4] D. Nodland, H. Zargarzadeh, and S. Jagannathan, "Neural network-based optimal control for trajectory tracking of a helicopter UAV," in *Proceedings of the 2011 50th IEEE Conference on Decision and Control and European Control Conference*, pp. 3876–3881, IEEE, Orlando, FL, USA, December 2011.
- [5] W. Liu, "A UAV route tracking algorithm based on the previewing point and fuzzy PID control," in *Proceedings of the 30th Chinese Control Conference*, pp. 2835–2839, IEEE, Yantai, China, July 2011.
- [6] S. Ragi and E. Chong, "Dynamic UAV path planning for multitarget tracking," in *Proceedings of the 2012 American Control Conference (ACC)*, pp. 3845–3850, IEEE, Montreal, QC, Canada, June 2012.
- [7] M.-G. Yoo and S.-K. Hong, "Target tracking control of a quadrotor UAV using vision sensor," *Journal of the Korean Society for Aeronautical & Space Sciences*, vol. 40, no. 2, pp. 118–128, 2012.
- [8] H. H. A. Kadouf and Y. M. Mustafah, "Colour-based object detection and tracking for autonomous quadrotor UAV," *IOP Conference Series: Materials Science and Engineering*, vol. 53, no. 1, Article ID 012086, 2013.
- [9] Y. Yu, X. Ding, and J. J. Zhu, "Attitude tracking control of a quadrotor UAV in the exponential coordinates," *Journal of the Franklin Institute*, vol. 350, no. 8, pp. 2044–2068, 2013.
- [10] C. Teuliere, E. Marchand, and L. Eck, "3-D model-based tracking for UAV indoor localization," *IEEE Transactions on Cybernetics*, vol. 45, no. 5, pp. 869–879, 2014.
- [11] H. Choi and Y. Kim, "UAV guidance using a monocular-vision sensor for aerial target tracking," *Control Engineering Practice*, vol. 22, pp. 10–19, 2014.
- [12] S. Quintero, D. Copp, and J. Hespanha, "Robust UAV coordination for target tracking using output-feedback model predictive control with moving horizon estimation," in *Proceedings of the 2015 American Control Conference (ACC)*, pp. 3758–3764, IEEE, Chicago, IL, USA, July 2015.
- [13] N. Santos, V. Lobo, and A. Bernardino, "A ground-based vision system for uav tracking," in *Proceedings of the OCEANS 2015-Genova*, pp. 1–9, IEEE, Genova, Italy, May 2015.
- [14] C. Zhou, Y. Liu, and Y. Song, "Detection and tracking of a UAV via hough transform," in *Proceedings of the 2016 CIE International Conference on Radar (RADAR)*, pp. 1–4, IEEE, Guangzhou, China, October 2016.
- [15] A. Vetrella, G. Fasano, D. Accardo, and A. Moccia, "Differential GNSS and vision-based tracking to improve navigation performance in cooperative multi-UAV systems," *Sensors*, vol. 16, no. 12, p. 2164, 2016.
- [16] M. Elloumi, B. Escrig, R. Dhaou, and H. Idoudi, "Saidane Designing an energy efficient UAV tracking algorithm," in *Proceedings of the 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pp. 127–132, IEEE, Valencia, Spain, June 2017.
- [17] C. Greatwood, L. Bose, T. Richardson et al., "Tracking control of a UAV with a parallel visual processor," in *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 4248–4254, IEEE, Vancouver, BC, Canada, September 2017.
- [18] N. Santos, V. Lobo, and A. Bernardino, "3d model-based estimation for uav tracking," in *Proceedings of the OCEANS 2018 MTS/IEEE Charleston*, pp. 1–9, IEEE, Charleston, SC, USA, October 2018.
- [19] W. Zhang, K. Song, X. Rong, and Y. Li, "Coarse-to-fine uav target tracking with deep reinforcement learning," *IEEE*

- Transactions on Automation Science and Engineering*, vol. 16, no. 4, pp. 1522–1530, 2018.
- [20] Z. Huang, C. Fu, Y. Li, F. Lin, and P. Lu, “Learning aberrance repressed correlation filters for real-time uav tracking,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 2891–2900, Seoul, Republic of Korea, November 2019.
- [21] M. Rabah, A. Rohan, S. A. S. Mohamed, and S.-H. Kim, “Autonomous moving target-tracking for a UAV quadcopter based on fuzzy-PI,” *IEEE Access*, vol. 7, pp. 38407–38419, 2019.
- [22] Y. Kokunko and S. Krasnova, “Synthesis of a tracking system with restrictions on UAV state variables,” *Mathematics in Engineering, Science & Aerospace (MESA)*, vol. 10, no. 4, 2019.
- [23] Y. Li, C. Fu, F. Ding, Z. Huang, and J. Pan, “Augmented memory for correlation filters in real-time uav tracking,” in *Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1559–1566, IEEE, Las Vegas, NV, USA, October 2020.
- [24] Y. Li, C. Fu, Z. Huang, Y. Zhang, and J. Pan, “Intermittent contextual learning for keyfilter-aware uav object tracking using deep convolutional feature,” *IEEE Transactions on Multimedia*, vol. 23, pp. 810–822, 2020.
- [25] J. Moon, S. Papaioannou, C. Laoudias, P. Kolios, and S. Kim, “Deep reinforcement learning multi-UAV trajectory control for target tracking,” *IEEE Internet of Things Journal*, vol. 8, no. 20, 2021.
- [26] S. Qiu, J. Luo, S. Yang, M. Zhang, and W. Zhang, “A moving target extraction algorithm based on the fusion of infrared and visible images,” *Infrared Physics & Technology*, vol. 98, pp. 285–291, 2019.