



Multi-Dimensional Visual Feature Fusion for UAV Maneuver Recognition in Power Transmission Line Protection

Liangfeng Chen¹, Ruiqi Xu¹, Jinyu Ma¹, Zheng Wang¹ and Han Li^{1,*}

¹School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

Abstract

The increasing popularity of civilian unmanned aerial vehicles (UAVs) poses serious safety challenges to low-altitude airspace near power transmission and distribution lines. Existing monitoring technologies rely heavily on expensive radar or radio frequency devices that are susceptible to interference, while mainstream vision schemes are typically limited to object detection and fail to analyze flight intentions. To address these problems, this paper proposes a UAV maneuver recognition method for transmission and distribution line protection, providing precise semantic support for action-based risk warnings through multi-dimensional visual feature fusion. First, a YOLO object detection model extracts real-time UAV bounding boxes and trajectories from video streams to construct raw motion state sequences. Second, to suppress detection noise while preserving critical maneuvering features,

a feature-preserving trajectory reconstruction algorithm based on median absolute deviation (MAD) enhancement (FP-SGM) is proposed, which combines a Savitzky-Golay (S-G) filter with an adaptive statistical threshold. On this basis, a multi-dimensional feature space comprising trajectory and scale features is constructed to design a cascade decision strategy. By fusing these features, the lateral and radial motions of UAVs are effectively decoupled, achieving robust recognition of four typical maneuvers: 'Hovering', 'Circling', 'Approaching', and 'Receding'. Simulation experiments on the Gazebo platform demonstrate a 95% accuracy rate in maneuver recognition. The results indicate that this low-cost method presents a viable potential for integration into existing surveillance terminals, laying a foundation for intelligent power line defense systems.

Keywords: UAV maneuver recognition, object detection, feature fusion, transmission and distribution line protection.



Academic Editor:

Fuyuan Xiao

Submitted: 10 November 2025

Revised: 18 June 2026

Accepted: 30 June 2026

Published: 27 September 2026

Vol. 3, No. 3, 2026.

10.62762/CJIF.2026.661309

*Corresponding author:

✉ Han Li

hanli@usst.edu.cn

1 Introduction

With the rapid development of UAV technology, vision-based detection of UAVs in complex

Citation

Chen, L., Xu, R., Ma J., Wang, Z. & Li, H. (2026). Multi-Dimensional Visual Feature Fusion for UAV Maneuver Recognition in Power Transmission Line Protection. *Chinese Journal of Information Fusion*, 3(3), 189–208.



© 2026 by the Authors. Published by Institute of Central Computation and Knowledge. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

backgrounds has become an active research topic, attracting growing interest from the computer vision community [1]. By 2027, the number of commercial drones will increase to about 955,000, while the number of consumer drones will reach 1.82 million [2]. However, with the widespread adoption of drones in agricultural plant protection [3], logistics delivery [4], power inspection [5], and other fields, the challenge to the safety of the distribution network is becoming increasingly prominent. UAVs may impact distribution lines or key power facilities due to operational errors, system failures, or intentional acts, potentially causing short circuits, equipment damage, or even large-scale power outages, seriously affecting people's lives and economic growth. Therefore, to ensure the safe operation of transmission and distribution lines, it is crucial to monitor UAVs in the vicinity of power lines in real time using key technologies.

Over the past decade, researchers have extensively explored various UAV monitoring technologies. Existing mainstream approaches primarily encompass four techniques: radar, radio frequency (RF), acoustics, and computer vision [6, 7]. Radar detection is a mature and widely utilized method for airspace surveillance. However, it faces significant limitations in identifying drones due to their weak radar cross-sections. Furthermore, its efficacy is often constrained in complex environments, and the cost of enhancing its performance remains prohibitive [8]. RF-based detection identifies UAVs by analyzing communication signals between the aircraft and the controller; however, it is highly susceptible to electromagnetic interference. For example, signal clutter in the high-radio-flow area reduces detection accuracy [9], rendering it ineffective for autonomous or encrypted UAVs. Therefore, it is impossible to locate or identify such objects reliably. Acoustic detection, which relies on capturing propeller noise, suffers from limited detection ranges, elevated false-alarm rates in noisy environments, and lower cost-effectiveness [10].

In contrast, vision-based UAV detection technology is increasingly gaining traction among researchers due to its notable advantages, including cost-efficiency, high precision, and ease of deployment [11, 12]. It is worth noting that the vision-based monitoring technology has a good foundation in practical applications. Taking China as an example, the security monitoring camera network has been widely deployed along transmission and distribution lines, providing a ready-made hardware foundation for visual solutions. However,

existing advanced visual detection equipment that can realize depth perception or behavior analysis, such as multi-sensor fusion devices like depth cameras and lidars, although it has high detection accuracy, is expensive and difficult to deploy at scale on transmission and distribution lines spanning tens of thousands of kilometers. In contrast, current visual research and applications based on ordinary cameras mainly focus on drone object detection; that is, they can only judge whether a drone is in the field of view and cannot analyze the object's depth information and behavioral intention. To truly achieve threat assessment, the key lies in the understanding of UAV maneuvering behavior. Therefore, given the existing security camera equipment, building a low-cost, automated threat early-warning system for transmission and distribution lines to achieve accurate identification of UAV maneuvers is a prerequisite.

In summary, to fully leverage the application potential of security equipment in existing transmission and distribution lines and to realize automatic monitoring of dangerous behavior in these lines, this study proposes a UAV maneuver recognition method based on a camera video stream. This approach leverages the inherent advantages of visual sensors to enable real-time recognition of UAV maneuvers by fusing dynamic trajectory data with real-time object detection. Specifically, this study first employs the YOLO object detection model to extract the centroid trajectory and bounding box area sequences of UAVs from video streams. To enhance data quality, an outlier filtering algorithm based on the MAD of smooth trajectory deviation is introduced to eliminate path drift caused by detection noise. Finally, accurate UAV maneuver recognition is achieved by fusing trajectory and bounding box size information. This method does not require additional sensors and has a low deployment cost. It has good applicability in complex power line environments and can be directly integrated with existing security cameras on transmission and distribution lines. This study verifies the method's overall effectiveness and robustness through experiments. The main contributions are as follows:

- (1) A low-cost visual perception UAV maneuver recognition framework for transmission and distribution line protection is proposed. Aiming to address the pain point of limited deep perceptual ability in the existing visual scheme, this study designs a recognition strategy based on a cascade decision. By combining the characteristics of two-dimensional

trajectory and object scale variation, the lateral maneuver and radial motion of UAV are effectively decoupled, and the accurate recognition of four typical maneuvers of 'Hovering', 'Circling', 'Approaching', and 'Receding' is realized under the condition of relying only on a single visual sensor.

(2) A feature-preserving trajectory reconstruction algorithm based on MAD enhancement (FP-SGM) is designed. To address the problems of large trajectory noise in visual detection and excessive smoothing of key maneuvering features by traditional filtering algorithms, this study proposes combining the Savitzky-Golay filter with an adaptive statistical threshold based on MAD. The algorithm can retain trajectory inflection point information that reflects the UAV's violent maneuver to the greatest extent while suppressing noise, thereby significantly improving the robustness of subsequent maneuver recognition.

(3) A multi-dimensional feature fusion recognition index combined with a heuristic algorithm optimization is proposed. The index is optimized using the APO algorithm. The fusion weights and decision thresholds for UAV trajectory and scale features are dynamically optimized globally, and a decision-making mechanism for decoupling lateral and radial motion is constructed. Experiments show that this index can effectively address motion semantic ambiguity in visual recognition. In complex environments, the recognition accuracy for typical UAV maneuvers such as hovering, circling, approaching, and receding is significantly improved, with strong environmental adaptability and robustness.

2 Related Work

2.1 Vision-based UAV object detection method

In recent years, object detection, as a core task in the field of computer vision, has been devoted to identifying and locating specific objects in digital images. With the rapid development of deep learning, the method based on deep learning has become the mainstream technology in this field [13].

In the aspect of UAV recognition, the object detection method based on deep learning shows important application value, and its performance depends largely on the support of high-quality datasets. However, UAV detection still faces many challenges in practical applications, mainly including small object size and complex background [14]. To cope with these challenges, researchers have

constructed a number of public datasets for UAV detection, such as USC-GRAD-STDdb [15], Aircraft Context Dataset [17], DUT Anti-UAV [18], and MMFW-UAV [19], which provide a basis for algorithm training and evaluation. In addition, air-to-air visual detection of micro-UAVs using deep learning has also been investigated to address detection in low-contrast and dynamic backgrounds [16]. At the same time, in order to improve the performance of small object detection, the related object detection algorithms also continue to evolve. In 2019, Hu et al. [20] proposed an improved small object detection algorithm based on YOLO, which significantly improved the performance of small object detection. In 2020, Zhizhong et al. [21] proposed a fast UAV object detection algorithm based on improved CenterNet. This algorithm can extract depth features from images and use a Gaussian kernel function to generate key point heat maps. While achieving high-efficiency detection, it also maintains high recognition accuracy. In 2024, Misbah et al. [22] proposed MSF-GhostNet, a computationally efficient YOLO-based framework incorporating multi-scale feature fusion for detecting drones under low-light conditions, which significantly improved detection accuracy while reducing computational overhead. In 2025, Du et al. [23] proposed the F-YOLOv7 algorithm, a fast and robust real-time detection method for drones. This method combines background difference preprocessing and a lightweight Fast-YOLOv7 network to achieve real-time and robust drone detection.

Although the above methods have made significant progress in detection accuracy and efficiency, they mainly focus on category recognition and spatial positioning, and cannot further infer the UAV's maneuvering intention or its threat level. In actual scenarios, such as transmission and distribution line protection, it is difficult to support timely and effective risk early warning only by relying on detection technology to judge the existence of drones. It is still necessary to combine other multi-dimensional information for deeper action analysis.

2.2 Vision-based trajectory analysis method

The trajectory is the core data to describe the motion state of the object, and the trajectory points can be obtained by the tracking algorithm to correlate the position points of the object in the time domain. Trajectory analysis has become an indispensable part of various UAV monitoring and behavior understanding tasks, and its application in UAV intention recognition and maneuver analysis

has received extensive attention from the research community over the past decade [24, 25].

With the continuous progress of this field, a large number of effective tracking algorithms have been proposed. The core method is to continuously locate moving objects and connect this location information in time series to form a complete motion trajectory. In 2019, Shuaau et al. [26] used a Gaussian mixture model and a Kalman filter for trajectory extraction and analysis, and designed a set of regular anomaly detection methods based on trajectory analysis to realize the automatic recognition of abnormal behaviors such as breaking into the forbidden zone and running in surveillance video. In 2025, Luan et al. [27] proposed a fast identification and detection algorithm for maneuverable unmanned aircraft based on multimodal data fusion, which combines an improved YOLOv11 architecture with multi-source sensor information to enhance the detection robustness of highly maneuverable UAVs in complex backgrounds. In 2024, Gao et al. [28] proposed a hybrid dual-scale neural network model based on the generalized regression multi-model and cubature information filter framework for tracking complex maneuvering UAVs, which effectively handles the dynamic and unpredictable motion patterns of highly maneuverable aerial targets.

The above research reflects the successful application of trajectory analysis in anomaly detection, intelligent monitoring, UAV behavior understanding, and other fields. However, in the UAV application scenario, most of the existing research focuses on trajectory prediction, and few studies have conducted an in-depth analysis of the UAV flight trajectory to achieve the identification of its specific maneuvers and assess potential threats. In order to fill the gap in the existing research on the behavior analysis of UAV trajectories, this study proposes a UAV trajectory analysis method based on visual perception, which aims to achieve accurate recognition of maneuvers by fusing trajectory information and object detection information.

3 Methodology

3.1 Overall framework

As shown in Figure 1, the UAV maneuver recognition method proposed in this paper is a top-down cascade processing framework, which aims to achieve robust recognition of UAV maneuvers relying only on a single visual sensor. The framework mainly includes four core processing modules: YOLO

object detection module, visual information extraction module, trajectory processing module, information fusion and solution module.

Firstly, the system receives the real-time video stream of the transmission and distribution line security camera. In the stage of visual information extraction, the object detection network based on YOLOv11 is used to infer each frame of the image [29], and the two-dimensional bounding box of the UAV is extracted. At this time, the geometric center of the bounding box is mapped to the original trajectory point sequence, and the bounding box area sequence is retained as the depth proxy variable for the relative distance change between the object and the camera. However, limited by complex background interference and detector jitter, the original trajectory data often contain noise points. Therefore, this paper introduces a trajectory cleaning mechanism based on statistics. The mechanism uses the Savitzky-Golay filter to construct a local smooth path and combines the MAD algorithm to dynamically eliminate the abnormal detection points that deviate significantly from the motion trend, thereby generating a high-confidence pure trajectory flow. On this basis, the framework enters the stage of feature solution fusion. It should be noted that the algorithm does not directly use coordinate data, but transforms it into kinematic features with physical semantics. Specifically, this method calculates five key indicators from the two dimensions of 'trajectory information' and 'scale information': angular velocity amplitude based on virtual anchors, directional oscillation rate, angular monotonicity, area change rate, and area linear fitting slope. These five indicators decouple the lateral maneuvering characteristics and radial distance characteristics of the object, respectively. Finally, in the action reasoning decision-making stage, this paper designs a rule-based cascade fusion strategy. This strategy gives high priority to scale trend features to compensate for the lack of depth perception of visual sensors, and maps the fusion results to four typical maneuvers of 'Hovering', 'Circling', 'Approaching', and 'Receding', so as to realize the recognition of UAV maneuvers in transmission and distribution lines.

It can be seen that the core advantage of the framework lies in its progressive processing mechanism. That is, based on visual perception data, by fusing the trajectory information and scale information of the object detection frame, the robust recognition of maneuvers is realized in a complex transmission and distribution environment. The algorithm can not only be directly applied to the security camera of the current

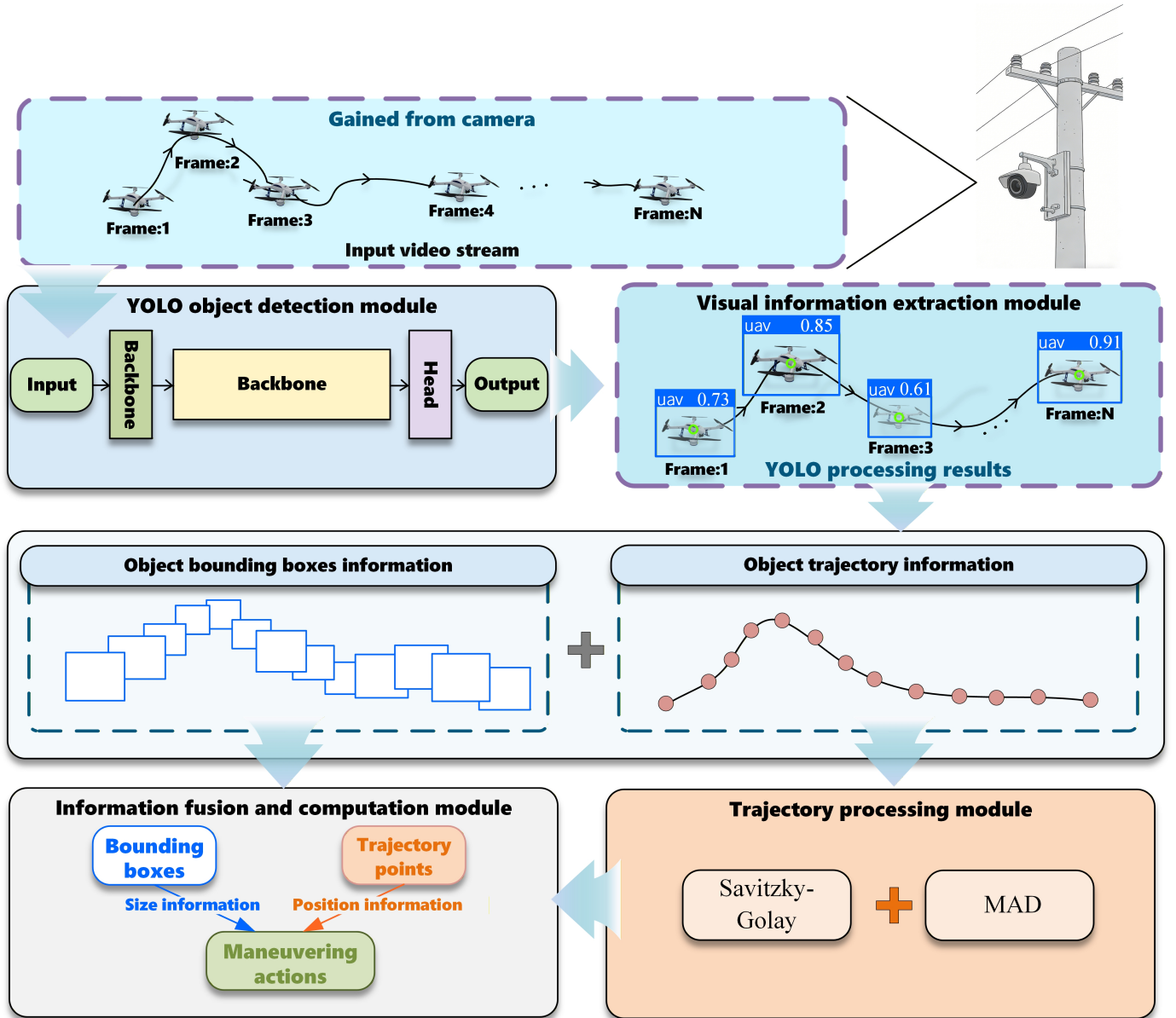


Figure 1. The framework of UAV maneuver recognition.

transmission and distribution lines to achieve low-cost intelligent monitoring, but also lay a solid foundation for the subsequent construction of an automated threat assessment system based on visual perception.

3.2 Visual object detection and state extraction

In order to realize the continuous tracking and maneuvering analysis of the invading UAV, the system first needs to extract the motion state of the UAV in real time from the video stream. Considering the sensitivity of the transmission and distribution line inspection scene to real-time performance and deployment cost, this paper adopts the visual perception scheme and uses the object detection model based on deep learning as the front-end core of motion state acquisition.

It is assumed that the video input stream of the transmission and distribution line protection camera is a sequence $F = \{f_1, f_2, \dots, f_t\}$, where f_t denotes the t -th frame image. In the detection stage, the YOLOv11 object detection algorithm is used to perform forward reasoning on each frame of the image. Compared with the two-stage detection algorithm, the YOLO algorithm directly predicts the category probability and position coordinates of the object through a single convolutional neural network, which can meet the needs of high frame rate processing.

For any frame f_t , the detector outputs a two-dimensional bounding box of the UAV, denoted by vector B_t :

$$B_t = [x_{1,t}, y_{1,t}, x_{2,t}, y_{2,t}] \quad (1)$$

where $(x_{1,t}, y_{1,t})$ and $(x_{2,t}, y_{2,t})$ represent the pixel coordinates of the upper left corner and the lower right corner of the bounding box.

Based on the bounding box information, this paper proposes two key types of state information, trajectory information and scale information, to describe the instantaneous motion characteristics of the UAV. The geometric center point of the bounding box is taken as the position representation on the UAV image. The geometric center point is the centroid point, which is recorded as P_t . This point constitutes the basic trajectory point for subsequent trajectory analysis :

$$P_t(u_t, v_t) = \left(\frac{x_{1,t} + x_{2,t}}{2}, \frac{y_{1,t} + y_{2,t}}{2} \right) \quad (2)$$

where (u_t, v_t) is the center pixel coordinates of the object in the t frame. The time series set composed of P_t is the original flight trajectory $\Gamma_{raw} = \{P_1, P_2, \dots, P_t\}$ of the UAV.

Due to the lack of direct depth information, the visual sensor cannot directly measure the physical distance between the object and the camera. According to the pinhole camera imaging model, the imaging size of the object is inversely proportional to its physical distance when the focal length is fixed. Therefore, this paper introduces the bounding box area as the proxy variable of depth :

$$S_t = |x_{2,t} - x_{1,t}| \times |y_{2,t} - y_{1,t}| \quad (3)$$

The sequence $S_{raw} = \{S_1, S_2, \dots, S_t\}$ reflects the trend of the object near or far away in the direction of the line of sight.

In summary, the visual detection module maps the high-dimensional video pixel space to a low-dimensional motion state space $M_t = \{u_t, v_t, S_t\}^T$. However, limited by the complex background interference and the positioning jitter of the detector itself, the original state sequences Γ_{raw} and S_{raw} often contain noise points. If they are directly used for subsequent differential calculations, it will lead to serious error amplification. Therefore, before determining the maneuver, the original state data must be robustly smoothed and denoised.

3.3 Robust trajectory filtering and refinement

Because the original trajectory of visual detection output is susceptible to complex background

interference, it often contains noise points that deviate significantly from the real path. In order to effectively suppress noise while maximizing the retention of key motion features in the rapid maneuvering process of UAVs, this study proposes FP-SGM algorithm. The algorithm combines the edge-preserving characteristics of the S-G filter and the robust statistical characteristics of the MAD. By processing the trajectory data step by step, the noise is removed while ensuring that the key maneuvering inflection points are retained to avoid being excessively smoothed.

3.3.1 Local reference path construction

In order to identify abnormal noise points, the algorithm first needs to construct a reference path that can not only reflect the real motion trend of the UAV, but also retain the key maneuvering inflection point. In this study, the S-G filter is used to perform local polynomial fitting on the original trajectory coordinates. The essence of the S-G filter is a time-domain smoothing algorithm based on local polynomial least squares fitting [30]. Common linear smoothing methods, such as Gaussian filtering and mean filtering, often inevitably weaken the extreme value of the signal while suppressing high-frequency noise, resulting in reduced trajectory peaks and broadened waveforms, thereby losing key motion details [31]. Especially in the identification of UAV trajectories, important inflection point information is easily lost. Therefore, compared with the traditional linear smoothing method, the S-G filter can better retain the high-order moments of the signal and is more suitable for processing UAV trajectory data with maneuvering characteristics. The specific implementation method is as follows.

For any point $P_t(u_t, v_t)$ in the trajectory sequence, the smoothed reference coordinate $\hat{P}_t(\hat{u}_t, \hat{v}_t)$ is defined by fitting $2m + 1$ data points in the sliding window centered on :

$$\hat{u}_t = \sum_{j=-m}^m c_j u_{t+j} \quad (4)$$

$$\hat{v}_t = \sum_{j=-m}^m c_j v_{t+j} \quad (5)$$

where m is the half window width, c_j is the convolution coefficient determined by the least square

method, and its value depends on the order of the selected fitting polynomial D . In this study, a set of fixed c_j is calculated in advance by the least square method, which converts polynomial fitting into an efficient time-domain convolution operation. Through this process, a smooth reference path $\Gamma_{smooth} = \{\hat{P}_1, \hat{P}_2, \dots, \hat{P}_t\}$ is finally obtained.

The final smooth path $\Gamma_{smooth} = \{\hat{P}_1, \hat{P}_2, \dots, \hat{P}_t\}$ is not directly used to eliminate noise points, but as a measurement benchmark to calculate the deviation between the original detection point and the path. The deviation is regarded as a key indicator to measure the detection error and is input into the subsequent MAD anomaly determination link.

3.3.2 Robust deviation metric and adaptive thresholding

Although the S-G filter can effectively construct a local reference path that reflects the motion trend of the UAV, it cannot directly identify and eliminate noise points that significantly deviate from the real trajectory. To this end, this study proposes an adaptive anomaly detection mechanism based on Trajectory Deviation to effectively eliminate noise points. The mechanism first calculates the Euclidean distance between each original observation point and its corresponding reference path point to form a trajectory deviation sequence, and constructs a robust adaptive decision threshold accordingly. In view of the fact that the traditional mean and standard deviation are easily affected by individual extreme noise points, resulting in threshold distortion and misjudgment. Therefore, this paper uses the median absolute deviation method to construct an adaptive threshold. The specific implementation process is as follows.

Firstly, for each time step, the Euclidean distance deviation between the original observation point and the reference path point is calculated :

$$d_t = \|P_t - \hat{P}_t\|_2 = \sqrt{(u_t - \hat{u}_t)^2 + (v_t - \hat{v}_t)^2} \quad (6)$$

Then calculate the median of the deviation of the whole sequence μ_d :

$$\mu_d = median(d_1, d_2, \dots, d_N) \quad (7)$$

Then the absolute deviation median of the deviation σ_{MAD} is calculated :

$$\sigma_{MAD} = median(|d_t - \mu_d| : t = 1, 2, 3, \dots, N) \quad (8)$$

Finally, the dynamic anomaly determination threshold τ is defined :

$$\tau = \mu_d + \lambda \cdot \kappa \cdot \sigma_{MAD} \quad (9)$$

where λ is the scale factor, theoretically set to 1.4826 for consistency in standard deviation estimation under a normal distribution [34], and κ is the sensitivity multiplier, set to 3.0 following the classic Hampel identifier for robust outlier filtering [35].

By introducing the MAD statistic, the threshold can dynamically adjust the tolerance range according to the change of the maneuvering amplitude of the UAV, and avoid the misjudgment of the fixed threshold in severe maneuvering. This robust adaptive threshold will be used as a key criterion for the subsequent iterative cleaning process.

3.3.3 Iterative cleaning strategy

In order to avoid the impact of a few noise points on the overall reference path $\hat{P}_t$, the algorithm uses an iterative optimization mechanism. In the k -th iteration, the following discriminant logic is executed :

$$M_t^{(k)} = \begin{cases} 0, & \text{if } d_t > \tau^{(k)} \text{ and } d_t > \varepsilon_{min} \\ 1, & \text{other} \end{cases} \quad (10)$$

where ε_{min} is the minimum deviation bottom noise, which is used to prevent the false removal of normal points due to $\sigma_{MAD} = 0$ when the trajectory is extremely smooth.

The points whose $M_t^{(k)}$ takes 0 will be removed from the sequence, and the remaining point set will enter the next round of the loop to recalculate the S-G reference path and MAD threshold. The process is repeated until no new outliers are detected or the maximum number of iterations K_{max} is reached. The final output trajectory Γ_{clean} after cleaning will be used as the input for subsequent feature extraction.

Through this cyclic iteration mechanism that combines the S-G filter and the MAD method, the local distortion introduced by the noise points in the path can be gradually corrected, and finally, a high-fidelity motion trajectory of the UAV that can retain the key inflection point information after denoising is generated.

3.4 Spatio-temporal motion modeling and maneuver recognition

In order to identify the specific maneuver of the UAV, this study constructs a multi-dimensional feature space containing angle features and scale features, and designs a rule-based cascade decision machine for final judgment.

3.4.1 Analysis of angular evolution metrics

In order to quantify the steering behavior of the object, the algorithm needs to first construct a reference point O outside the image plane. The reference point is extended along the diagonal direction of the image, aiming to map the two-dimensional plane trajectory Γ_{clean} to a one-dimensional angular displacement sequence relative to O . Its location is defined as :

$$O = C_{img} + n \cdot L_{diag} \cdot \eta \quad (11)$$

where C_{img} is the image center, n is the normalized direction vector, L_{diag} is the diagonal length of the image, and η is the expansion factor.

For the trajectory point P_t , the azimuth angle Θ_t relative to O is calculated, and then the angular velocity sequence $\Omega = \{\omega_1, \omega_2, \dots, \omega_t\}$ is obtained. The azimuth angle and the angular velocity are defined as follows :

$$\theta_t = \arctan\left(\frac{v_t - v_0}{u_t - u_0}\right) \quad (12)$$

$$\omega_t = \frac{d\theta}{dt} \quad (13)$$

Based on the angular velocity ω_t , the three key evaluation indices of angular velocity amplitude V_{mad} , oscillation ratio R_{osc} , and angle monotonicity M can be calculated, which are defined as follows :

$$V_{mad} = \text{median}(|\omega_t - \text{median}(\omega)|) \quad (14)$$

where the angular velocity amplitude V_{mad} is mainly used to measure the activity of the object motion.

$$R_{osc} = \frac{1}{N-1} \sum_{t=2}^{N-1} \parallel [\text{sgn}(\omega_t) \neq \text{sgn}(\omega_{t+1})] \quad (15)$$

where $\parallel$ is the indicator function. The oscillation ratio R_{osc} is used to count the frequency of the

angular velocity symbol flipping, which is mainly used to emphasize the stability of the motion direction. The larger the value, the more unstable the motion direction of the UAV.

$$M = \max\left(\frac{N(\omega_t > 0)}{N_{total}}, \frac{N(\omega_t < 0)}{N_{total}}\right) \quad (16)$$

Angle monotonicity M is mainly used to measure the consistency of motion direction. If the value of M is close to 1, it shows that the motion direction of the UAV is basically the same; if the value of M is close to 0.5, it indicates that the UAV may be doing a reciprocating motion.

3.4.2 Scale evolution metrics

In the transmission and distribution line protection scenario, it is difficult to effectively perceive the depth change of the UAV only by relying on the two-dimensional trajectory information, and the size scaling of the object detection frame is closely related to the radial distance of the object. Therefore, this paper introduces the scale change of the detection frame as the proxy variable of the depth information, and derives the area change rate ρ and the area linear fitting slope K_{slope} to infer the dynamic change of the UAV in the radial distance.

The area change rate ρ refers to the size change rate of the object detection frame of the first frame of the UAV in the picture and the last frame in which the UAV remains visible. It is defined as follows:

$$\rho = \frac{S_{end} - S_{start}}{S_{start}} \quad (17)$$

where S_{start} is the first frame of the UAV in the picture, and S_{end} is the last frame of the UAV in the picture.

The area linear fitting slope K_{slope} refers to the average rate at which the area of the detection box changes with the frame sequence. If the value is greater than 0, it indicates that the overall trend of the area of the detection box is increasing, indicating that the UAV may be approaching the camera. If the value is less than 0, it indicates that the overall trend of the detection frame area is decreasing, indicating that the drone may show a trend away from the camera; if the value is equal to 0, it means that there is no obvious change trend in the area of the detection box, indicating that the drone has no obvious tendency to approach or stay away.

$$K_{slope} = \frac{n \sum_{k=1}^n (k \cdot S_k) - (\sum_{k=1}^n k)(\sum_{k=1}^n S_k)}{n \sum_{k=1}^n k^2 - (\sum_{k=1}^n k)^2} \quad (18)$$

where S_k denotes the area of the drone in frame k .

3.4.3 Cascade decision strategy

In order to realize the robust recognition of UAV maneuvers under the visual condition of lack of depth information, this paper constructs a cascade decision-making reasoning mechanism based on the aforementioned trajectory information characteristics and scale change characteristics. The framework constructs a mapping rule from static state to complex maneuvering behavior by fusing trajectory information features and scale information features. The typical spatio-temporal evolution process of various maneuvering actions is shown in Figure 2. In the figure, the blue box is the object detection box, and the red dot is the trajectory point.

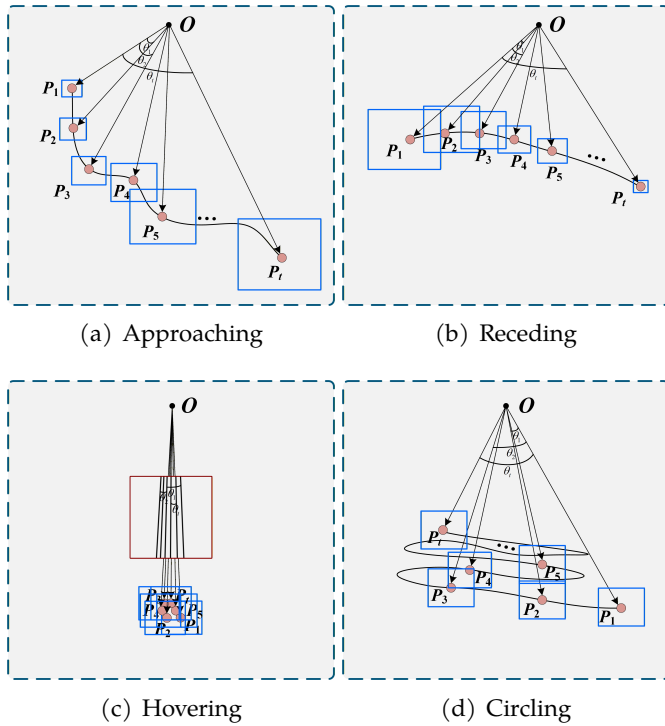


Figure 2. Spatio-temporal evolution process of four typical UAV maneuvering actions.

Firstly, the algorithm determines the high priority of the radial maneuver mode with significant risk characteristics. As shown in Figures 2(a) and 2(b), the core feature of such maneuvers is the monotonic evolution of the detection frame area along the time axis. In order to quantify the overall linear trend of the area sequence more robustly, this paper uses the area linear fitting slope K_{slope} and the area change rate ρ as

the core indicators to measure the radial motion. When obvious scale expansion or contraction is detected, that is, $|\rho|$ is greater than the radial change threshold of the area change rate, the system will trigger the radial maneuver judgment: if $K_{slope} > 0$, it is judged to be Approaching maneuver, indicating that the UAV is approaching the power facility. If $K_{slope} < 0$, it is judged to be Receding maneuver, indicating that the drone is far away from the power facilities. This logic ensures the accuracy of maneuver recognition in complex scenes through the statistical advantages of linear fitting. The specific manifestations are as follows:

$$\begin{cases} |K_{slope}| > T_{slope} \\ K_{slope} > 0 \\ |\rho| > T_{area} \\ \rho > 0 \end{cases} \rightarrow Action = Approaching \quad (19)$$

$$\begin{cases} |K_{slope}| > T_{slope} \\ K_{slope} < 0 \\ |\rho| > T_{area} \\ \rho < 0 \end{cases} \rightarrow Action = Receding \quad (20)$$

where T_{slope} is the radial change threshold of the area linear fitting slope, and T_{area} is the radial change threshold of the area change rate.

Subsequently, the system identifies the hovering mode by comprehensively evaluating the angular velocity amplitude V_{mad} , oscillation ratio R_{osc} , and area linear fitting slope K_{slope} . As shown in Figure 2(c), when the drone is in a hover state, affected by the positioning jitter of the detector, the trajectory point shows a high random oscillation characteristic in a small range, but its angular velocity amplitude relative to the reference pole is extremely low. At the same time, because the radial distance between the object and the camera is basically constant, the scale of the detection frame shows extremely high stability, and its area change trend will be close to 0. Therefore, the criterion of hovering action is defined in this paper: the amplitude of angular velocity V_{mad} and the linear fitting slope of the area K_{slope} are lower than the preset threshold, and the oscillation ratio R_{osc} is higher than the preset threshold. This criterion effectively uses the scale stationarity as an auxiliary criterion and eliminates the misjudgment caused by the long-distance and low-speed drift of the object. The specific judgment form is as follows :

$$\begin{cases} V_{mad} < T_{mad} \\ |K_{slope}| < T_{slope} \\ R_{osc} > T_{osc} \end{cases} \rightarrow Action = Hovering \quad (21)$$

where T_{osc} is the directional oscillation limit of the oscillation ratio, and T_{mad} is the motion significance threshold of the angular velocity amplitude.

Finally, for the object that does not have significant scale variation characteristics but shows a continuous motion state, the algorithm identifies the circling mode by analyzing the symmetry and periodicity of its trajectory. As shown in Figure 2(d), the circling maneuver is usually represented by the reciprocating motion of the object within a specific angular displacement interval. At the feature level, although this kind of action has a certain angular velocity amplitude V_{mad} , the monotonicity of its trajectory is weak, showing a very high direction flip rate. Therefore, when the object satisfies the active motion condition but does not trigger the radial warning, and the angle monotonicity M is less than the threshold, it is determined as the circling maneuver. The specific manifestations are as follows:

$$\begin{cases} V_{mad} > T_{mad} \\ M < T_M \end{cases} \rightarrow Action = Circling \quad (22)$$

where T_M is the angle monotonicity threshold.

Through this progressive cascade strategy, this method can effectively decouple the complex three-dimensional maneuvering motion under the constraint of a lack of depth information, and realize the accurate identification of the UAV maneuvering motion around the transmission and distribution lines.

4 Experiment

4.1 Simulation environment and experimental setup

In order to verify the effectiveness of the proposed method, this study builds a simulation environment based on the XTDrone platform, which integrates Robot Operating System (ROS), PX4, and Gazebo for the test and verification of the method [32]. The simulation experiment is carried out on an Ubuntu 20.04.4 operating system, which is equipped with a 12th Gen Intel Core i9-12900H processor and 16 GB of RAM.

In this experiment, the built-in UAV model of the simulation environment was utilized, and the path planning and flight control are carried out by QGroundControl (QGC). QGC is a powerful ground control station software that has good support for the PX4 system. It can easily plan the route and adjust the important parameters such as flight height and speed [33]. As shown in Figure 3, it is the path planning interface of QGC. Figure 4 is a four-rotor UAV flying according to the planned path in the simulation environment.

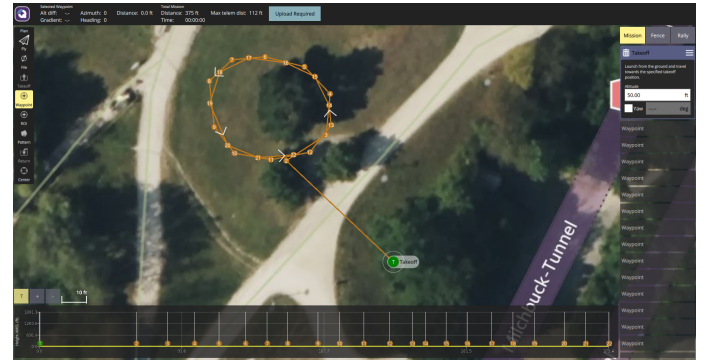


Figure 3. Interface for UAV trajectory planning in QGC.

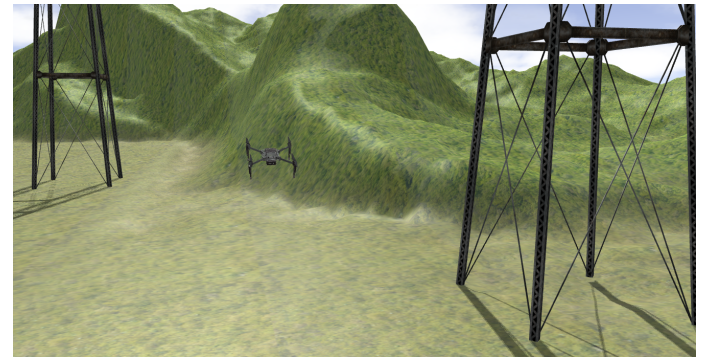


Figure 4. UAV flying along the planned trajectory in the simulation environment.

To obtain the visual information of the simulation environment, we set up a virtual camera fixed in the world coordinate system in Gazebo, which can capture the video stream taken in the virtual environment in real time. The camera is implemented by a custom SDF model file, and its specific configuration is shown in Table 1.

In the experiment, firstly, QGC is used to plan the route of diversified UAVs to simulate various flight maneuvers in real power transmission and distribution scenarios. At the same time, the real-time video stream is synchronously collected by the virtual camera in the simulation environment as the system input, and the object motion state in the video stream is

Table 1. Detailed configuration of the virtual camera in the simulation environment.

Parameter	Value
FOV ($^{\circ}$)	90
FPS (Hz)	30
Resolution (pixels)	640×480
Near Clip Distance (m)	0.1
Far Clip Distance (m)	100

monitored and analyzed in real time by using the method proposed in this paper, and finally, the specific maneuver category of the UAV is output.

In addition to the nominal simulation environment, an environmental robustness test set was further constructed to analyze the influence of external environmental disturbances. Since real-world UAV monitoring around transmission lines is often affected by illumination variation and wind-induced flight disturbance, this study introduces two representative environmental factors, namely illumination condition and wind speed. The illumination condition was divided into daytime and nighttime settings, and the wind disturbance was modeled as horizontal crosswind with different wind speeds. Other camera parameters, including FOV, FPS, resolution, and clipping distance, were kept the same as those in the nominal simulation environment.

4.2 UAV object detection based on YOLOv11

This study uses the publicly available UAV dataset S-UAV-T. Compared with other existing UAV datasets, this dataset has significant advantages in terms of UAV model, environmental diversity, and data size. Specifically, the dataset is constructed based on 10 different multi-rotor aircraft models, 32 atypical textures, and a variety of poses, backgrounds, and perspectives. It contains a total of 52,500 images of 604×604 pixels, with standard YOLO format annotation files. Since the original data is arranged in the order of the scene, in order to improve the generalization ability of the model training and avoid the order deviation, this study first randomly disrupts the dataset before training the model, and then divides it into training set, verification set and test set according to the ratio of 8:1:1. During the model training process, we follow the default parameter settings in the YOLOv11 model. In particular, we use stochastic gradient descent (SGD) as the optimizer, with a momentum of 0.937 and an initial learning rate of 0.01. The model is trained for 100 epochs with a batch size of 16. The specific performance of the trained model is shown in Table 2.

Table 2. Performance of the object detection model.

	Precision	Recall	mAP50	mAP50:95
	(%)	(%)	(%)	(%)
Value	99.63	98.90	99.51	92.00

4.3 Comparative ablation experiment and qualitative analysis of filtering algorithm

4.3.1 Contrast ablation experiment

To verify the feasibility and superiority of the FP-SGM algorithm, a comparative ablation experiment was carried out in this study. The proposed algorithm was qualitatively evaluated with the original detection data, only relying on the S-G filter and the classical Kalman filter algorithm. In the experiment, 400 UAV flight trajectories were selected as test samples. These samples evenly cover four typical maneuvers of hovering, circling, approaching, and receding, so as to ensure the fairness of evaluation under different motion characteristics. Specifically, we use Root Mean Square Error (RMSE), Smoothness, Average Corner Error (ACE), and Latency to test the original data and three filtering methods, respectively. *RMSE* is used to measure the global position deviation between the filtered trajectory and the real trajectory. The calculation formula is as follows.

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N \|P_{est}(t) - P_{gt}(t)\|^2} \quad (23)$$

where N is the total number of frames of the trajectory sequence, $P_{est}(t)$ is the position coordinate (x_{est}, y_{est}) after filtering at time t , $P_{gt}(t)$ is the actual real position coordinate (x_{gt}, y_{gt}) at time t , and $\|\bullet\|$ represents the Euclidean distance.

Smoothness is used to measure the smoothness of the curve, and its calculation formula is as follows.

$$Smoothness = \frac{1}{N} \sum_t (\|P(t+1) - 2P(t) + P(t-1)\|^2) \quad (24)$$

ACE is specifically used to evaluate the feature retention ability of the algorithm at the key inflection point of the trajectory. To calculate *ACE*, it is first necessary to define the set of moments when the curvature exceeds the set threshold $\kappa_t h$ in the real trajectory as the inflection point set T_{corner} .

$$T_{corner} = \{t | \kappa_{gt}(t) > \kappa_{th}\} \quad (25)$$

where $\kappa_{gt}(t)$ is the curvature of the real trajectory at time t , and κ_{gt} is the curvature threshold. After obtaining T_{corner} , ACE can be calculated according to the following formula.

$$ACE = \frac{1}{T_{corner}} \sum_{t \in T_{corner}} ||P_{est}(t) - P_{gt}(t)|| \quad (26)$$

Latency reflects the processing time of the filtering algorithm. In the scenarios with extremely high real-time requirements, such as transmission and distribution line threat assessment, the processing delay of the filtering algorithm is a key indicator to measure its computational efficiency. Low latency directly determines the ability of the system to respond in real time. The specific calculation formula is as follows.

$$Latency = \frac{1}{M} \sum_{j=1}^M T_j \quad (27)$$

where M is the total number of test trajectories, and T_j is the single processing time of the j -th trajectory.

Based on the above evaluation criteria, we summarize the test results of the trajectories processed by different filtering algorithms on various indicators as shown in Table 3.

Table 3. Quantitative performance comparison of different filtering algorithms.

Filter Type	RMSE (pixels)	Smoothness (pixels/frame)	ACE (pixels)	Latency (ms)
Raw	18.45	8.12	14.20	/
Kalman	9.42	2.15	28.34	12.65
S-G	13.78	4.21	12.50	1.25
FP-SGM	5.32	4.89	4.12	2.45

It can be seen from Table 3 that although the Kalman filter performs well in trajectory smoothness, its ACE is high, which indicates that the algorithm excessively suppresses the signal characteristics in the smoothing process, resulting in the loss of key maneuvering inflection point information. In addition, its high calculation delay also makes it difficult to meet the strict real-time requirements of transmission and distribution line threat assessment. In contrast, although S-G filtering is superior to Kalman filtering in retaining inflection point features, from the perspective of *Smoothness*, its denoising ability is relatively limited, and the trajectory is still disturbed by extreme

noise points. In contrast, although the FP-SGM algorithm proposed in this study is slightly inferior to the Kalman filter in absolute smoothness, it achieves the best balance between noise suppression and feature retention, especially in inflection point retention and low latency. The characteristics are significantly better than the above two methods, and these two characteristics are crucial for the highly dynamic UAV threat determination scenario.

4.3.2 Visual analysis

The effect of trajectory reconstruction under different processing methods is shown in Figure 5. In the figure, the red dot represents the trajectory point, the green dot represents the starting trajectory point, and the blue dot represents the trajectory end point.

From the first line of Figure 5, it can be seen that due to the false detection of the object detection model in complex scenes, there will be obvious noise points deviating from the trajectory in the original trajectory data, especially in the original images of the ‘Approaching’, ‘Receding’, and ‘Hovering’ paths in the figure. This kind of noise point generated by false detection seriously damages the continuity of the trajectory. If it is not processed, it will easily lead to a large deviation in the subsequent calculation, which will lead to the wrong recognition of the final maneuver.

In view of the above problems, the last three lines of Figure 5 show the processing effect of the three filtering algorithms. The comparative analysis shows that although the Kalman filter can effectively suppress noise points and generate a smoother path, it will smooth the key inflection points in the UAV trajectory. Although relying only on the S-G filter can make the trajectory smooth, it will still be disturbed by noise. As the trajectory of the ‘Receding’ action in Figure 5 shows, the Kalman filter smoothes the original sharp maneuvering inflection point into a rounded arc, which fails to truly reflect the motion path of the object. In the specific application scenario of distribution network threat assessment, the geometric shape of the object motion path is an important basis for subsequent feature extraction, and the smoothing process of the Kalman filter will suppress the path inflection point, resulting in the loss of key feature information. In contrast, the FP-SGM algorithm proposed in this study shows excellent edge preservation characteristics due to the polynomial fitting characteristics of the S-G filter and the adaptive constraint of the MAD statistical threshold. When

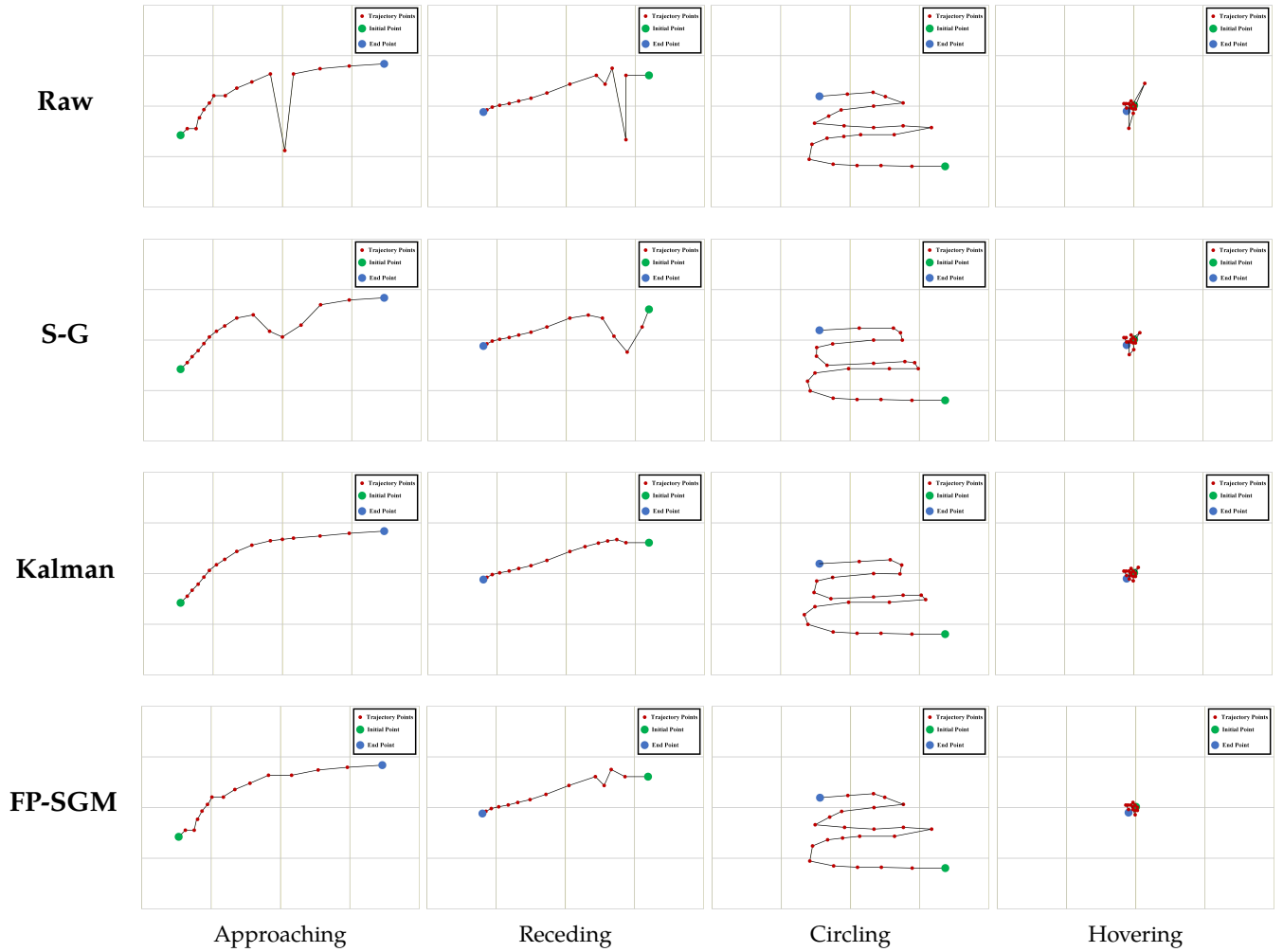


Figure 5. Visualization comparison of trajectory under four typical maneuvering actions.

dealing with the 'Approaching' trajectory, the FP-SGM algorithm not only robustly eliminates the noise points that deviate from the trajectory, but also retains the inflection point characteristics of the trajectory more accurately, and there is no phenomenon that the Kalman filter algorithm filters out the inflection point. The experimental results fully prove that the algorithm effectively overcomes the problem of the loss of inflection point information in the dynamic scene of the traditional Kalman filter while efficiently filtering the observation noise, and ensures the high fidelity reproduction of the reconstructed trajectory to the real motion intention of the UAV, which provides a reliable data basis for the subsequent feature calculation.

4.4 Threshold optimization algorithm experiment and analysis

In the UAV maneuver recognition method based on visual trajectory analysis proposed in this paper, the final judgment result depends on a set of judgment thresholds related to trajectory dynamic characteristics

and scale change behavior. These thresholds together constitute the core parameter configuration of the rule discrimination module. Although each threshold has a clear physical meaning and has been defined by feature modeling in the previous section, its optimal value is highly sensitive to factors such as imaging angle, motion amplitude, and video sampling conditions in practical applications. If the threshold is set only based on artificial experience, it is not only easy to introduce subjective bias, but also difficult to maintain the consistency of recognition performance in different scenarios or different maneuvering modes. At the same time, the thresholds are not independent of each other, but there is a significant coupling relationship. The adjustment of a single threshold often affects the overall decision boundary, thereby reducing the stability and generalization ability of the rule-based method. Therefore, it is difficult to obtain reliable and reproducible maneuver recognition results by relying solely on artificial experience for threshold setting. Based on the above analysis, this paper will introduce

a heuristic optimization algorithm to automatically optimize the threshold set, so as to obtain the threshold configuration with the best performance and good robustness in the overall sense.

On this basis, this paper selects five representative heuristic optimization algorithms for threshold optimization experiments and comparative analysis, namely artificial protozoan optimization algorithm (APO), cheetah optimization algorithm (CO), particle swarm optimization algorithm (PSO), pigeon swarm optimization algorithm (PIO), and goose swarm optimization algorithm (GOOSE). The above algorithms cover different search mechanisms and exploration–exploitation balance strategies, including classical continuous optimization methods based on group collaboration, behavior-driven biological heuristic algorithms, and optimization models with phased convergence characteristics. In order to ensure the fairness of the comparison between different algorithms, all heuristic algorithms are run under a unified optimization framework and experimental conditions. The parameters are set as shown in Table 4. The population size is set to 50, the maximum number of iterations is set to 100, and the remaining parameters are set by the conventional settings of each algorithm. By comparing a variety of heuristic algorithms under the same parameter configuration, the applicability and stability of different search strategies in the threshold optimization task are systematically analyzed, so as to avoid the result bias caused by a single optimization method or parameter setting.

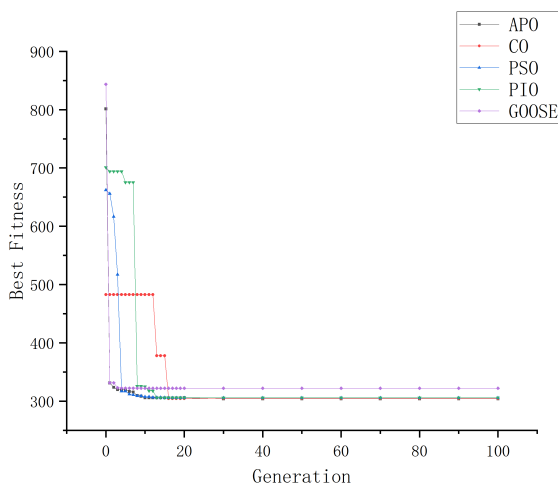


Figure 6. Convergence curves of different optimization algorithms.

As can be seen from the comparison and analysis of

the convergence curve in Figure 6, in the optimization process, the CO and PIO algorithms have experienced a long stagnation period before the fitness value is greatly reduced, and the final convergence accuracy is low. Although the GOOSE algorithm has a certain downward trend in the initial stage, it is easy to fall into a local optimum, resulting in the worst performance of its final convergence value among the five algorithms. In the comparison algorithm, the PSO algorithm performs relatively well, and its convergence speed and accuracy are better than the first three. Among all the algorithms, the APO algorithm is the most prominent. It not only shows a strong explosive force of optimization at the initial stage of iteration, but also the convergence speed is significantly faster than PSO, and the final fitness value is lower, showing better optimization accuracy and global search stability. In summary, the APO algorithm has obvious performance advantages in processing UAV maneuver recognition tasks.

After the global optimization of the APO algorithm, the final optimal parameter combination is: the angular velocity amplitude threshold is 0.0050, the area change rate threshold is 0.0500, the area linear fitting slope threshold is 8.7464, the oscillation ratio threshold is 0.3784, and the angle monotonicity threshold is 0.6717. The best fitness value obtained at this time is 304.4040. Based on this set of optimal parameter configurations, the accuracy data analysis of UAV maneuver recognition is further carried out to verify the effectiveness of the optimized model in recognition tasks at different distances.

4.5 Performance evaluation of maneuver recognition

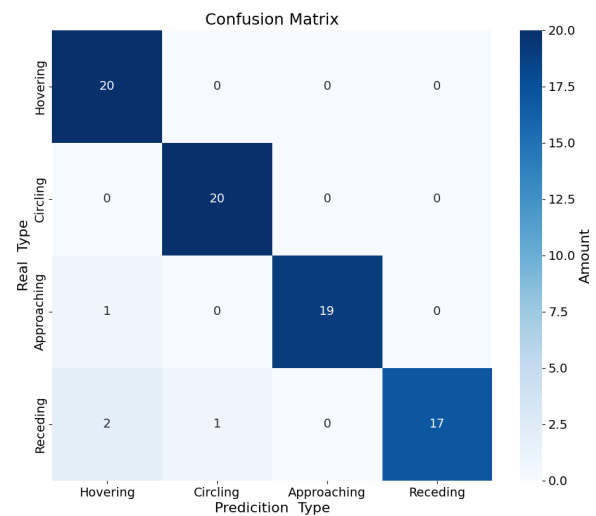
In order to verify the performance of the proposed method in UAV maneuver recognition, this study constructed a dataset containing 400 video samples. The dataset covers four types of typical maneuvers, including 'Hovering', 'Circling', 'Approaching' and 'Receding', and the number of four types of action samples is balanced, all of which are 100 segments, so as to avoid the deviation of model training caused by class imbalance. During the experiment, the dataset is partitioned into a training set, a verification set and a test set via a manual stratified partitioning strategy, with 60% of the data used for parameter optimization and model training of the algorithm; 20% as the verification set, which is used to monitor the recognition accuracy and adjust the hyperparameters in real time during the optimization process. The

Table 4. Parameters of APO, CO, PSO, PIO, and GOOSE algorithm.

Algorithms	Parameters	Meanings	Values
APO	N_p	Population size	50
	$N_{c,max}$	Maximum number of iterations	100
	σ_f	Exploration step size in foraging behavior	0.1
	P_f	Probability of random foraging	0.5
CO	N_p	Population size	50
	$N_{c,max}$	Maximum number of iterations	100
	ω	Mass factor	0.7
	d_{th}	Distance threshold between cheetah and prey	1.0
PSO	c_1, c_2	Learning factor	1.5
	$N_{c,max}$	Maximum number of iterations	100
	N_p	Population size	50
	ω	Mass factor	0.7
PIO	$N_{c1,max}$	Maximum number of iterations of map and compass operators	50
	$N_{c2,max}$	Maximum number of iterations of landmark operator	50
	N_p	Population size	50
	R	Decay coefficient	0.05
GOOSE	N_p	Population size	50
	$N_{c,max}$	Maximum number of iterations	100
	$f_{exchange}$	Information Exchange Frequency	5
	$P_{exchange}$	Individual Exchange Probability	0.2

remaining 20% is used as an independent test set to finally evaluate the generalization ability and robustness of the maneuver recognition system in unknown scenarios.

Based on the optimized optimization parameters, the performance of the maneuver recognition system is evaluated on the test set. The experimental results are shown in Table 5. The system shows excellent recognition performance. The overall recall reaches 95.00%, and the F1-score is 94.98%, which proves the balance of the algorithm in multi-category tasks. A separate analysis of the four maneuvers shows that the recognition accuracy of the circling and approaching maneuvers is extremely high, and the F1-score reaches 97.56% and 97.44%, respectively. Among them, the circling action achieves a recall rate of 100%, indicating that the system can capture the maneuvering trajectory with such obvious characteristics without omission. Although the precision rate of hovering action is 86.96%, the recall rate reaches 100%, which means that the system misjudges some other actions as hovering, but does not miss any real hovering state. In contrast, the recall rate of the receding maneuver is 85%, and there is a small amount of missed detection. Figure 7 is the confusion matrix diagram of four typical maneuver recognition results, and the visualization results of the experimental process are shown in Figure 8.

**Figure 7.** Confusion matrix of the maneuver recognition results.

On the whole, the algorithm shows high robustness in four typical maneuvers and can effectively deal with the complex visual changes in the dynamic flight process of the UAV.

In order to further evaluate the generalization ability and robustness of the algorithm, this paper extends the detection distance from 0-20 m to 20-40 m and 40-60 m, and analyzes the recognition performance of the

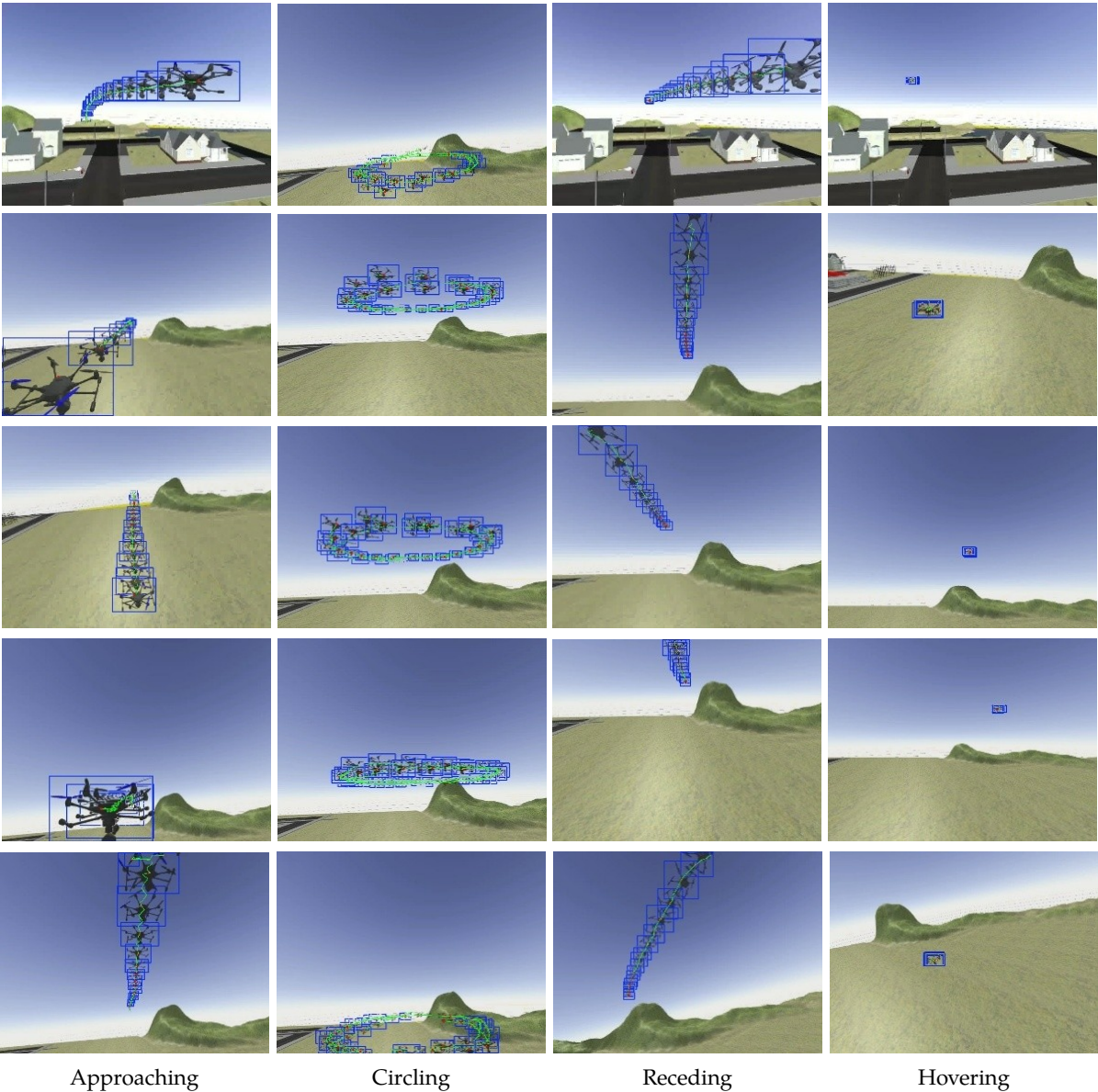


Figure 8. Visualization results of the recognition experiments for four typical maneuvering actions.

Table 5. Performance evaluation of different maneuver types recognition.

Maneuver Types	Precision (%)	Recall (%)	F1-score (%)
Hovering	86.96	100.00	93.02
Circling	95.24	100.00	97.56
Approaching	100.00	95.00	97.44
Receding	100.00	85.00	91.89
Total	95.55	95.00	94.98

In the range of 20-40 m, the algorithm can still maintain relatively robust recognition performance, and the Precision is 88.42%. The slight decrease in performance at this time is mainly due to the decrease in the proportion of pixels in the imaging picture of the drone, resulting in a slight jitter in the YOLO detection frame, which in turn introduces a small noise to the area slope feature.

Table 6. Performance analysis of maneuver recognition across different distance ranges.

Distance Range (m)	Precision (%)	Recall (%)	F1-score (%)
0-20	95.55	95.00	94.98
20-40	88.42	86.50	87.50
40-60	76.15	70.24	73.11

algorithm at different distances. As shown in Table 6, the experimental results show that with the increase of observation distance, the performance of maneuver recognition shows obvious degradation characteristics.

However, when the distance is further extended to 40-60 m, the performance degradation trend is intensified, and the predicted recall is reduced to 70.24%. This phenomenon is caused by two factors: First, at a fixed focal length, a long-range small object will lead to a decrease in the confidence of the YOLO detector, resulting in discontinuous detection results, which destroys the timing integrity required for maneuvering analysis; secondly, because the area characteristics of the object near and far from the action tend to be flattened at a long distance, it is difficult for the algorithm to effectively distinguish the subtle displacement changes.

Although the above results demonstrate the effectiveness of the proposed method under nominal simulation conditions and different observation distances, real-world deployment may also be affected by external environmental disturbances. Therefore, an additional robustness evaluation under illumination and wind-speed variations is conducted in the following section.

4.6 Robustness evaluation under external environmental disturbances

The experiments in the previous sections were mainly conducted under nominal Gazebo simulation conditions. However, in practical transmission line protection scenarios, the visual monitoring system may be affected by external environmental factors, particularly illumination changes and wind-induced deviations in UAV trajectories. Since large-scale real-world UAV flight tests around operating transmission lines are constrained by safety, site, and operational considerations, this section introduces controlled environmental disturbance experiments in a simulation environment.

For consistency with the baseline experiment in Section 4.5, all samples in the environmental robustness test set were collected within the observation distance range of 0-20 m. This setting ensures that performance differences across conditions are primarily due to illumination and wind-speed variations, rather than to changes in observation distance.

To avoid ambiguity with the original dataset used for threshold optimization and baseline evaluation, an independent environmental robustness test set was constructed. The test set contains six environmental conditions: one baseline and five disturbed. Each condition contains 80 video samples, with 20 samples for each maneuver type. Therefore, the environmental

robustness test set contains 480 video samples. All samples in this test set were used only for robustness evaluation. The YOLO detector, the FP-SGM trajectory reconstruction module, and the optimized decision thresholds were left unchanged; no retraining or threshold re-optimization was performed.

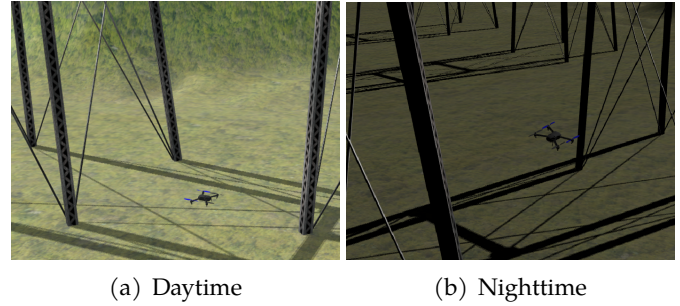


Figure 9. Visual examples of daytime and nighttime illumination conditions.

The wind direction was fixed as horizontal crosswind in all wind-disturbance experiments. This setting was adopted because crosswind can introduce lateral deviations in the image plane, which directly affect the trajectory features used by the proposed maneuver recognition method. The three wind-speed levels, 0 m/s, 3 m/s, and 6 m/s, correspond to no-wind, moderate-wind, and strong-wind conditions, respectively. The nighttime condition was used to simulate low-illumination visual monitoring, while keeping the UAV object observable in the camera view. Figure 9 shows visual examples of the daytime and nighttime simulation environments.

Table 7. Performance analysis under different illumination and wind-speed conditions.

Illumination	Wind speed (m/s)	Precision (%)	Recall (%)	F1-score (%)
Daytime	0	95.55	95.00	94.98
Daytime	3	93.42	92.51	92.96
Daytime	6	89.86	88.73	89.29
Nighttime	0	91.28	90.17	90.78
Nighttime	3	87.31	86.25	86.77
Nighttime	6	82.74	81.88	82.31

As shown in Table 7, recognition performance decreases gradually as wind speed increases under both daytime and nighttime illumination. Under daytime conditions, the F1-score decreases from 94.98% at 0 m/s to 89.29% at 6 m/s, indicating that the proposed method can still maintain stable performance under moderate wind disturbance. Compared with daytime conditions, nighttime illumination results in more pronounced performance degradation. The lowest recall appears under nighttime illumination

with 6 m/s crosswind, where the F1-score decreases to 82.31%. This is mainly because low illumination reduces the stability of UAV detection, while strong crosswinds increase trajectory fluctuations.

5 Conclusion

In this paper, a low-cost UAV maneuver recognition method based on visual perception is proposed to address the urgent need for automatic, intelligent monitoring of UAVs in transmission and distribution line protection scenarios. This method aims to address the problem that existing visual perception schemes cannot perform deep perception and action analysis. Fusing multi-dimensional motion features from the video stream provides accurate action semantic support for the subsequent construction of a fully automatic threat assessment and risk early warning system.

Specifically, this method first uses the YOLOv11 model to extract the two-dimensional bounding box of the UAV from the video stream, frame by frame. It constructs the original state sequence containing location and scale information. To address the problem of large noise in the original detection data and the tendency of traditional filtering to lose key maneuvering features, this paper proposes a feature-preserving trajectory reconstruction algorithm based on MAD enhancement. The contrast ablation experiment on the filtering algorithm demonstrates that it achieves the best balance between noise suppression and feature preservation. Compared with traditional filtering algorithms such as the Kalman filter, the proposed algorithm can effectively filter out observation noise and accurately retain trajectory inflection point information, thereby overcoming the problem of excessive smoothing in traditional methods that leads to the loss of key motion features. These filtered and reconstructed high-quality trajectories lay a reliable data foundation for the subsequent accurate analysis of the UAV's maneuvers.

Based on obtaining high-quality trajectories, this study constructs a maneuver recognition method based on a cascade decision strategy, and realizes the robust recognition of four typical maneuvers of 'Hovering', 'Circling', 'Approaching', and 'Receding' by fusing trajectory information and bounding box size information. Since the recognition method relies on multiple decision thresholds, to overcome the subjectivity and limitations of manual setting, we introduce an adaptive threshold optimization method to obtain a globally optimal threshold combination.

The final experimental results show that the proposed method not only realizes the robust recognition of four typical maneuvers of hovering, circling, approaching, and receding, but also maintains good generalization performance under different observation distances, which verifies the feasibility of relying only on a single visual sensor to realize complex three-dimensional maneuver analysis.

In summary, this study successfully developed a UAV maneuver recognition scheme based exclusively on a single visual sensor. The scheme has significant advantages, including low cost and ease of deployment, and can be directly integrated into the existing power transmission and distribution security system. Through the analysis of UAV maneuvers, a solid foundation is laid for the future realization of intelligent line defense and countermeasure decision-making. However, due to experimental limitations, current validation is primarily conducted in high-fidelity simulation environments, and robustness in extremely complex scenarios remains to be verified. Furthermore, current maneuver recognition relies heavily on changes in the area of 2D bounding boxes; however, during complex non-radial maneuvers, the abrupt shifts in trajectory points caused by drastic changes in UAV attitude can easily interfere with scale-based feature extraction, leading to misclassification. In the future, we will incorporate 3D pose estimation or monocular depth prediction to mitigate interference caused by attitude variations, further verify the scheme in multi-UAV cooperative scenarios and real complex meteorological environments, and explore the deep integration of the recognition module with high-level threat assessment models to achieve end-to-end automated analysis from maneuver recognition to threat determination.

Data Availability Statement

Data will be made available on request.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest.

AI Use Statement

The authors declare that ChatGPT-5 was used for language editing of the manuscript. The authors

have carefully reviewed, revised, and verified the AI-assisted output and take full responsibility for the content of the manuscript.

Ethical Approval and Consent to Participate

Not applicable.

References

- [1] Lin, X., Niu, Y., Yu, X., Fan, Z., Zhuang, J., & Zou, A. M. (2025). Paying more attention on backgrounds: Background-centric attention for UAV detection. *Neural Networks*, 185, 107182. [CrossRef]
- [2] Shakhathreh, H., Sawalmeh, A. H., Al-Fuqaha, A., Dou, Z., Almaita, E., Khalil, I., ... & Guizani, M. (2019). Unmanned aerial vehicles (UAVs): A survey on civil applications and key research challenges. *IEEE access*, 7, 48572-48634. [CrossRef]
- [3] Nafar, N., Fatemi, M., & Rezaei-Moghaddam, K. (2025). Decoding drone adoption in agriculture: A comparative analysis of behavioral models. *Information Processing in Agriculture*. [CrossRef]
- [4] Adeoye, Y., Onotole, E., Ogunyankinnu, T., Aipoh, G., Osunkanmibi, A., & Egbemhenghe, J. (2025). Artificial intelligence in logistics and distribution: The function of AI in dynamic route planning for transportation including self-driving trucks and drone delivery systems. *World Journal of Advanced Research and Reviews*, 25(02), 155-167. [CrossRef]
- [5] Xu, B., Zhao, Y., Wang, T., & Chen, Q. (2023). Development of power transmission line detection technology based on unmanned aerial vehicle image vision. *SN Applied Sciences*, 5(3), 72. [CrossRef]
- [6] Tian, J., Wang, C., Cao, J., & Wang, X. (2024). Fully convolutional network-based fast UAV detection in pulse Doppler radar. *IEEE Transactions on Geoscience and Remote Sensing*, 62, 1-12. [CrossRef]
- [7] Lee, H., Han, S., Byeon, J. I., Han, S., Myung, R., Joung, J., & Choi, J. (2023). CNN-based UAV detection and classification using sensor fusion. *IEEE Access*, 11, 68791-68808. [CrossRef]
- [8] Wang, J., Liu, Y., & Song, H. (2021). Counter-unmanned aircraft system (s)(C-UAS): State of the art, challenges, and future trends. *IEEE Aerospace and Electronic Systems Magazine*, 36(3), 4-29. [CrossRef]
- [9] Chamola, V., Kotesch, P., Agarwal, A., Gupta, N., & Guizani, M. (2021). A comprehensive review of unmanned aerial vehicle attacks and neutralization techniques. *Ad Hoc Networks*, 111, 102324. [CrossRef]
- [10] Jang, B., Seo, Y., On, B., & Im, S. (2018, January). Euclidean distance based algorithm for UAV acoustic detection. In *2018 International conference on electronics, information, and communication (ICEIC)* (pp. 1-2). IEEE. [CrossRef]
- [11] Seidaliyeva, U., Ilipbayeva, L., Taissariyeva, K., Smailov, N., & Matson, E. T. (2023). Advances and challenges in drone detection and classification techniques: A state-of-the-art review. *Sensors*, 24(1), 125. [CrossRef]
- [12] Ding, G., Ren, Y., Liu, Y., Zhao, Q., & Li, S. (2025). Vision-based anti-Unmanned aerial technology: Opportunities and challenges. *IEEE Geoscience and Remote Sensing Magazine*, 3(4), 382-405. [CrossRef]
- [13] Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). Object detection in 20 years: A survey. *Proceedings of the IEEE*, 111(3), 257-276. [CrossRef]
- [14] Liu, Z., An, P., Yang, Y., Qiu, S., Liu, Q., & Xu, X. (2024). Vision-based drone detection in complex environments: A survey. *Drones*, 8(11), 643. [CrossRef]
- [15] Bosquet, B., Mucientes, M., & Brea, V. M. (2020). STDnet: Exploiting high resolution feature maps for small object detection. *Engineering Applications of Artificial Intelligence*, 91, 103615. [CrossRef]
- [16] Zheng, Y., Chen, Z., Lv, D., Li, Z., Lan, Z., & Zhao, S. (2021). Air-to-air visual detection of micro-uavs: An experimental evaluation of deep learning. *IEEE Robotics and automation letters*, 6(2), 1020-1027. [CrossRef]
- [17] Steininger, D., Widhalm, V., Simon, J., Kriegler, A., & Sulzbachner, C. (2021). The aircraft context dataset: Understanding and optimizing data variability in aerial domains. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 3823-3832). [CrossRef]
- [18] Zhao, J., Zhang, J., Li, D., & Wang, D. (2022). Vision-based anti-UAV detection and tracking. *IEEE Transactions on Intelligent Transportation Systems*, 23(12), 25323-25334. [CrossRef]
- [19] Liu, Y., Sun, Z., Xi, L., Zhang, L., Dong, W., Chen, C., ... & Deng, F. (2025). MMFW-UAV dataset: multi-sensor and multi-view fixed-wing UAV dataset for air-to-air vision tasks. *Scientific data*, 12(1), 185. [CrossRef]
- [20] Hu, Y., Wu, X., Zheng, G., & Liu, X. (2019, July). Object detection of UAV for anti-UAV based on improved YOLO v3. In *2019 Chinese Control Conference (CCC)* (pp. 8386-8390). IEEE. [CrossRef]
- [21] Zhizhong, X., Jingen, W., Zhenghao, H., & Yuhui, S. (2020, October). Research on multi UAV target detection algorithm in the air based on improved CenterNet. In *2020 International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE)* (pp. 367-372). IEEE. [CrossRef]
- [22] Misbah, M., Khan, M. U., Kaleem, Z., Muqabel, A., Alam, M. Z., Liu, R., & Yuen, C. (2024). MSF-GhostNet: computationally efficient YOLO for detecting drones in low-light conditions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 18, 3840-3851. [CrossRef]
- [23] Du, Y., Wu, T., Dai, Z., Xie, H., Hu, C., & Wei, S. (2025).

- F-yolov7: Fast and robust real-time UAV detection. *Computing*, 107(1), 50. [CrossRef]
- [24] Wang, Z., Chen, J., Wang, J., & Shen, Q. (2023). Intention recognition of UAV swarm with data-driven methods. *Aerospace Systems*, 6(4), 703-714. [CrossRef]
- [25] Perrusquía, A., Guo, W., Fraser, B., & Wei, Z. (2024). Uncovering drone intentions using control physics informed machine learning. *Communications Engineering*, 3(1), 36. [CrossRef]
- [26] Shuaau, M., Thang, K. F., & Lai, N. S. (2019). Autonomous abnormal behaviour detection using trajectory analysis. *International Journal of Electrical and Computer Engineering (IJECE)*, 2. [CrossRef]
- [27] Luan, T., Zhou, S., Zhang, Y., & Pan, W. (2025). Fast identification and detection algorithm for maneuverable unmanned aircraft based on multimodal data fusion. *Mathematics*, 13(11), 1825. [CrossRef]
- [28] Gao, Y., Gan, Z., Chen, M., Ma, H., & Mao, X. (2023). Hybrid dual-scale neural network model for tracking complex maneuvering UAVs. *Drones*, 8(1), 3. [CrossRef]
- [29] Khanam, R., & Hussain, M. (2024). Yolov11: An overview of the key architectural enhancements. *arXiv preprint arXiv:2410.17725*. [CrossRef]
- [30] Schmid, M., Rath, D., & Diebold, U. (2022). Why and how Savitzky–Golay filters should be replaced. *ACS Measurement Science Au*, 2(2), 185. [CrossRef]
- [31] Zhu, F., Chang, C., Li, Z., Li, B., & Li, L. (2024). A generic optimization-based enhancement method for trajectory data: Two plus one. *Accident Analysis & Prevention*, 200, 107532. [CrossRef]
- [32] Xiao, K., Tan, S., Wang, G., An, X., Wang, X., & Wang, X. (2020, June). XTDrone: A customizable multi-rotor UAVs simulation platform. In *2020 4th International Conference on Robotics and Automation Sciences (ICRAS)* (pp. 55-61). IEEE. [CrossRef]
- [33] Meier, L., Honegger, D., & Pollefeys, M. (2015, May). PX4: A node-based multithreaded open source robotics framework for deeply embedded platforms. In *2015 IEEE international conference on robotics and automation (ICRA)* (pp. 6235-6240). IEEE. [CrossRef]
- [34] Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the median absolute deviation. *Journal of the American Statistical association*, 88(424), 1273-1283. [CrossRef]
- [35] Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of experimental social psychology*, 49(4), 764-766. [CrossRef]



Liangfeng Chen is currently pursuing the M.Eng. degree at the School of Mechanical Engineering, University of Shanghai for Science and Technology. His main research interests include low-altitude safety, artificial intelligence assisted decision-making and UAV intention assessment. (Email: 242241456@st.usst.edu.cn)



Ruiqi Xu is currently pursuing the B.Eng. degree at the School of Mechanical Engineering, University of Shanghai for Science and Technology. His main research interests include information fusion, low-altitude safety, artificial intelligence assisted decision-making and UAV intention assessment. (Email: 2235060327@st.usst.edu.cn)



Jinyu Ma is currently pursuing the M.Eng. degree at the School of Mechanical Engineering, University of Shanghai for Science and Technology. His main research interests include embedded systems, deep learning application and UAV intention assessment. (Email: 243431696@st.usst.edu.cn)



Zheng Wang is currently pursuing the M.Eng. degree at the School of Mechanical Engineering, University of Shanghai for Science and Technology. His main research interests include image processing, object detection and UAV inspection. (Email: 232241490@st.usst.edu.cn)



Han Li received the Ph.D. degree in flight dynamics and control from Fudan University, in 2024. Now, he is currently a lecturer and master's supervisor at the School of Mechanical Engineering, University of Shanghai for Science and Technology. His current research interests include information fusion and artificial intelligence assisted decision-making. (Email: hanli@usst.edu.cn)