Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Integrating the cat's eye effect and deep learning for low-altitude target detection*

Bin ZHOU[†‡1], Weiming WANG[2], Ning YAN[1], Linlin ZHAO[1], Chuanzhen LI[1]

[1]*School of Electronics and Electrical Engineering, Zhengzhou University of Science and Technology, Zhengzhou 450064, China*

[2]*Laboratory of Micro-Nano Sensor Technology, Xiong'an Institute of Innovation, Xiong'an 071800, China*

[†]E-mail: whelmmail@126.com

Received July 22, 2025; Revision accepted Oct. 31, 2025; Crosschecked Nov. 19, 2025; Published online Dec. 12, 2025

Abstract: This paper addresses the urgent need to detect low, slow, and small (LSS) unmanned aerial vehicles (UAVs) in complex and critical environments, proposing an active low-altitude target detection method based on the cat's eye effect. The detection system incorporates a control module, a laser emission component, a co-optical path panoramic scanning optical mechanism structure, an echo reception component, target detection, and visualization processing to achieve small target detection. The light source is emitted by a near-infrared laser, and the scanning optical path is realized using micro-electro-mechanical system (MEMS) mirrors and servo mechanisms. The echo reception signal is received by an avalanche photodiode (APD) and the target detection module, which captures the reflected signal and distance information. The detection software integrates the local pyramid attention (LPA) module and the field pyramid network (FPN) through the UAV micro lens identification algorithm. It eliminates false alarms by incorporating SKNet21 and uses the APD to collect echo intensity and flight time, thereby reducing the false alarm rate. The results demonstrate the feasibility of the proposed target detection method, which achieves a mean average precision of 0.809 at an intersection over union (IoU) of 0.50, a mean average precision of 0.324 at an IoU of 0.50–0.95, and a throughput of 49.8 Giga floating-point operations per second (GFLOPs), indicating that it can address the current limitations in LSS target detection.

Key words: Low-altitude detection; Optical path detection; Cat's eye effect; SKNet21; Local pyramid attention; Average precision
https://doi.org/10.1631/FITEE.2500522 **CLC number:** TP391

1 Introduction

Low, slow, and small (LSS) targets, primarily represented by unmanned aerial vehicles (UAVs), typically refer to flying objects with low flight altitudes, slow movement speeds, and a small radar cross-section (Bishop and Tufariello, 2019). Currently, the primary methods for detecting UAVs include radar detection, radio spectrum detection, and acoustic wave detection (Farlik et al., 2019; Liaquat et al., 2024; Zakaria et al., 2024; Wang XW et al., 2025). Radar can detect targets in various weather conditions, is less susceptible to electromagnetic interference, and can track and locate multiple targets over long distances. However, radar detection of low-altitude micro UAVs necessitates the effective filtering of clutter interference or suppression of the multipath effect due to these vehicles' small radar cross-section, significant electromagnetic wave echo interference, and severe multipath effect (Bao et al., 2025; Figueiredo et al., 2025). The echo signals also need to be processed in a refined manner. Radio spectrum detection technology can distinguish a UAV's model by analyzing its unique radio frequency characteristics, and can determine its intentions when combined with artificial intelligence (AI). This technology can effectively

‡ Corresponding author

* Project supported by the Science and Technology Research Project of Henan Province (Nos. 242102211110 and 252102211060) and the Science and Technology Innovation Project of Hebei Province (No. SJMYF202312)

ORCID: Bin ZHOU, https://orcid.org/0009-0000-2770-5527

reduce the false alarm rate, but it relies on the communication between UAVs, resulting in a significant reduction in the detection rate. Additionally, its detection range is limited and significantly affected by the surrounding electromagnetic environment. Thus, the use of radio spectrum detection technology should be coordinated with radar and electro-optical systems to form a "detection–identification–attack" closed loop (Zhang WH et al., 2025). Finally, acoustic wave detection technology for low-altitude UAVs can achieve passive detection; however, the detection range is limited to small areas with low accuracy (Goldman, 2016). Currently, UAVs are used in both military and civilian contexts due to their economic benefits and flexibility. While they have benefited the public, they have also led to numerous illegal data collection and criminal activities, making it necessary to develop a method for detecting low-altitude LSS UAVs.

Numerous algorithms have been proposed in the field of target detection and recognition, particularly for detecting UAVs (Li et al., 2021; Liu YC et al., 2021; Liu BL and Luo, 2022; AlKhonaini et al., 2024; Rahman et al., 2024; Randieri et al., 2025; Zheng et al., 2025). Anti-UAV technology can be classified as UAV recognition, positioning, or interception technologies (Xu and Luo, 2025). In recent years, Zhang QQ et al. (2024) have proposed an improved small target detection method based on Picodet to address the real-time issue of UAV detection. They enhanced detection accuracy by refining the feature pyramid structure. Wang XW et al. (2025), focusing on the limitations of the YOLOv8s target detection algorithm, introduced an improved target detection approach. By incorporating AKConv, the method achieves more efficient feature extraction, enabling the system to adaptively capture the LSS targets at various positions. AlKhonaini et al. (2024) proposed a hierarchical reinforcement learning technique, primarily based on radio frequency signals, to identify invading UAVs. This approach improves overall detection accuracy by using hierarchical training sets. Liu BL and Luo (2022) replaced the algorithm backbone of Yolov5 with efficient net-lite, reducing the number of parameters in the model and improving accuracy in recognizing multi-rotor UAVs. While the above-mentioned algorithms have all improved and optimized the recognition and detection of UAVs, challenges remain due to the complexity of flight environments. Factors such as flight speed and the angle of targets often result in low precision and reduced stability of the collected image data. Furthermore, while some deep learning algorithms demonstrate high recognition accuracy within distances of less than 500 m, their performance is limited when applied to targets beyond 10 000 m. Although infrared waves are employed to mitigate distance-related issues, there is still a need to develop algorithms and technologies that are specifically tailored for low-altitude environments to enhance detection capabilities. Given the limitations inherent in individual detection technologies, such as the need for clutter processing in radar, dependence on environmental conditions in photoelectric systems, and reliance on communication signals in the radio frequency spectrum, it is challenging for a single detection method to meet the demands of diverse scenarios. Consequently, integrating multiple technologies to establish a comprehensive system for detecting small targets has become an essential development need.

The cat's eye effect occurs when a laser scans the space where a target is located. If the laser illuminates the target's photoelectric equipment, a reflected echo forms, which returns along the original path of the laser. The intensity of the echo is 2–3 orders of magnitude higher than that of the diffuse reflection target. Systems producing this cat-eye effect are called cat-eye systems (Fan et al., 2020; Lv et al., 2022; Du et al., 2024). Addressing the urgent need for LSS UAVs to detect complex key areas, and in light of the common shortcomings of existing detection algorithms, this paper proposes a low-altitude target detection method that integrates the "cat's eye effect" with deep learning. By incorporating laser radar and visible light imaging through a shared optical path, the detection range for LSS targets is expanded. The SKNet21 dual-network architecture enhances both visible light images and echoes, ultimately establishing an efficient, multidimensional detection mechanism under the control of the main control module. The primary contributions of this work are as follows:

1. UAV detection relies on leveraging the imaging characteristics of drone's micro lens and laser echo properties, particularly how changes in incident angle and diffraction phenomena affect laser reflection and diffraction. Analyzing the echo characteristics of the "cat's eye effect" in drone micro-lenses can therefore

significantly enhance the detection system's ability to identify UAVs.

2. To address the limitations of various radar technologies, such as the limited detection range of single-line scanning radar, the complexity of multi-line radar (which requires multiple transmit/receive arrays and intricate installation/calibration), and the inability of solid-state radar to achieve omnidirectional scanning, this paper proposes an omnidirectional scanning system with a shared optical path for laser and visible light. By combining micro-electro-mechanical system (MEMS) mirrors with precision servo mechanisms, the proposed system enables high-precision, omnidirectional detection.

3. By capturing both echo and visible light images of "cat's eye targets" at varying distances, this approach combines convolutional neural network (CNN) models with attention mechanisms to enhance the resolution of both image types. This process enriches the detailed information about "cat's eye targets," and performance is evaluated using metrics such as mean accuracy, recall rate, and model parameter count.

2 System design and theoretical calculation

Fig. 1 depicts a schematic diagram of our method, which comprises a laser transmission module, an echo reception module, and optical path scanning and image acquisition processing modules. This configuration enables co-optical-path panoramic image acquisition and the reception of diffracted echoes from the "cat's eye target," followed by echo image analysis and processing.

2.1 Design of the laser emission module

The launch module comprises a near-infrared laser, a MEMS scanning mirror, and a front-end collimating optical system. The infrared laser not only receives information in the visible light spectrum but also detects laser echoes reflected from the "cat's eye target." The receiving assembly comprises a visible light echo reception channel and a laser echo reception channel. The former captures echo image information from the target area, simultaneously feeding these data to the target detection and visualization module for image analysis and processing. Concurrently, it transmits this information alongside the frame synchronization signal to the central control module. The laser echo reception channel receives reflected signals from near-infrared laser pulses, ultimately conveying distance information to the target detection and visualization module. The detailed parameters are listed in Table 1.

2.2 Design of the detection receiving circuit

The receiving end detects the echo signals from the target area, achieving the amplification, collection, and analysis of the echo signals, thereby completing

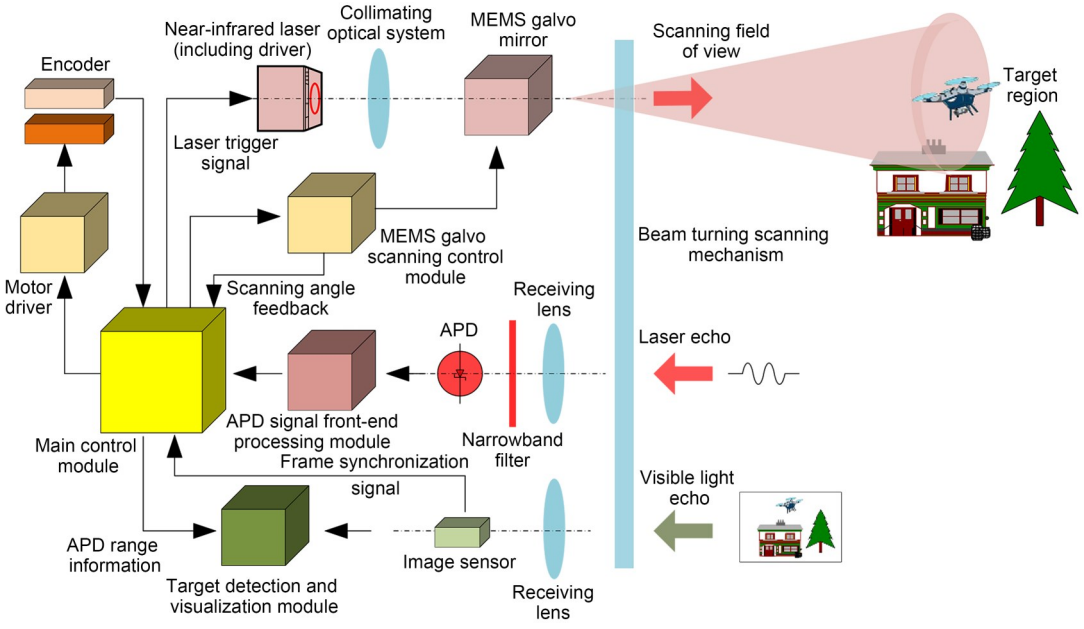

Fig. 1 Schematic diagram of the system

Table 1 List of parameters

Parameter	Unit	Value
CCD detector		WAT-902B
CCD operating voltage	V	12
CCD (SNR)	dB	50
Laser wavelength	μm	1.06
MEMS tilt angle	°	≤28
MEMS power	W	125

CCD: charge-coupled device

recognition of the "cat's eye target." It is composed of a front-end convergent optical system, the APD unit detector, an amplification circuit, a power supply circuit, an automatic gain control (AGC) circuit, and a high-speed data acquisition and processing board card, as shown in Fig. 2.

The power supply circuit provides working voltage for the APD unit detector, including a 300–600 V high-voltage direct current. The "cat's eye target" echo is converged onto the APD detector by the receiving optical system. The detector converts the received optical signal into the corresponding electrical signal, which is amplified by the amplification circuit and then serves as the input signal for the high-speed data acquisition and processing board card. The "cat's eye target" is detected based on the difference in response signal strength. Additionally, the data acquisition and processing board card can provide the data acquisition card trigger signal and AGC signal to generate distance gain control.

2.3 Optical path sharing panoramic scanning optical engine structure design

To address the limitations in the detection distance and range of the existing detection system, this paper proposes an opto-mechanical structure that combines MEMS scanning with mechanical scanning, as shown in Fig. 3. During the scanning process, rotating images undergo synchronous forward correction to ensure a direct mapping relationship between the real object space and its imaging space, thereby reducing the difficulty of target recognition. The optical-mechanical structure for co-path joint detection is illustrated in Fig. 3.

This system integrates MEMS mirrors and precise servo mechanisms, achieving all-around and high-precision wide-angle detection capabilities. It is also equipped with a de-chirp mechanism, enabling collaborative detection and precise positioning of specific

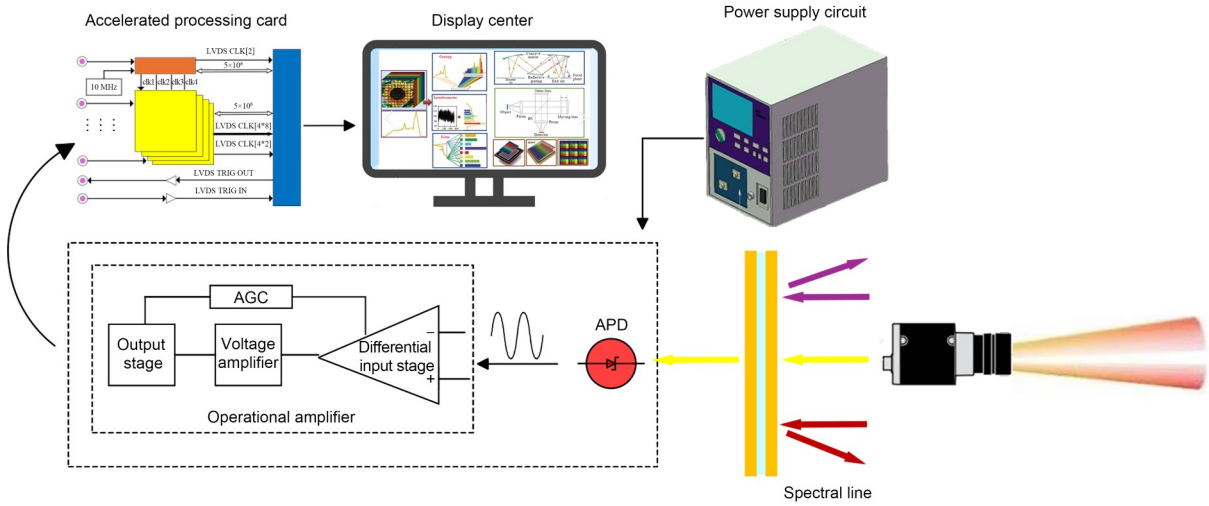

Fig. 2 Block diagram of the detection receiving end

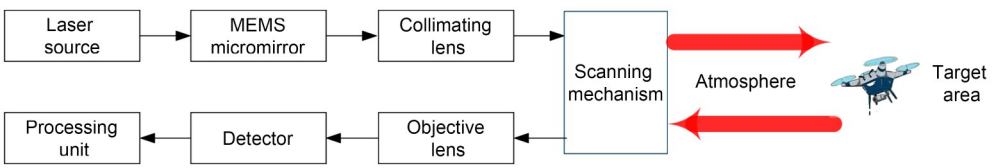

Fig. 3 Simplified diagram of the laser panoramic scanning system

targets using both visible light and laser. The optical-mechanical structure mainly consists of a target information detection and processing component and a wide-angle scanning receiving component that communicates with it. The scanning receiving component uses a laser as the emission unit, with individual detectors and image detectors serving as the receiving elements. Through the coordinated action of the servo motor and MEMS mirrors, the scanning optical machine realizes all-around scanning and detection of the target area, as shown in Fig. 4.

The system employs the principle of the panoramic periscope to eliminate image rotation during the imaging process, as shown in Fig. 5a. The basic structure is periscope-like, consisting of an upper reflector, an objective lens, a lower reflector, and an eyepiece. The light is reflected by the upper reflector into the vertical objective lens and then reflected by the lower reflector into the eyepiece for human observation. When the upper reflector rotates around its vertical axis (at the center of the vertical objective lens), the surrounding environment can be observed. According to optical principles, the rotation of the upper mirror induces the rotation of the image in the eyepiece. A Dove prism is introduced to counteract this, which rotates strictly in a 1:2 ratio, following the upper reflector in the same or opposite direction, to eliminate image rotation (Yasir and Goyal, 2021). The panoramic lens of the panoramic scanning receiving component and the reflective surface of the Dove prism for image rotation elimination can be coated with aluminum film or other high-reflection films. Anti-reflection films can also be applied to the light beam entry and exit surfaces to increase the light transmission rate. The calculation formula for the length of the Dove prism is:

$$2l=\frac{\sqrt{2}\,h\sin\left[\arcsin\left(\frac{\sqrt{2}}{2n}\right)+\frac{\pi}{2}\right]}{\sin\left[\frac{\pi}{4}-\arcsin\left(\frac{\sqrt{2}}{2n}\right)\right]}, \tag{1}$$

where n represents the refractive index of the glass, h is the length of the side of the right-angle prism, and $2l$ is the side length of the Dove prism.

To achieve effective APD reception, the beam-splitting prism and the filter are introduced. The beam-splitting prism is used to split the received light beam, which is then output to the visible light image sensing and laser unit detection systems, respectively, realizing signal conversion and providing data support for subsequent processing. The filter only allows near-infrared band lasers to pass through. Due to the complex optical path and the fact that the upper optical lens and the Dove prism rotate at a differential speed of 1:2 during panoramic scanning, requiring strict synchronization, the design of the common optical path structure is complex. Three core-less direct current (DC) motors are adopted for synchronous motion control, effectively reducing the difficulty of subsequent recognition algorithms. Fig. 5b shows the optical-mechanical structure of the common optical path panoramic scanning.

2.4 Analysis of the algorithm

2.4.1 Analysis of echo diffraction with the micro lens

Under low-altitude, short-range detection conditions, the diameter of the expanded and collimated laser beam exceeds the aperture diameter of the lens. Based on the structure of the miniature lens and the principle of optical path reversibility, the optical system can be equivalently modeled as a combination of a

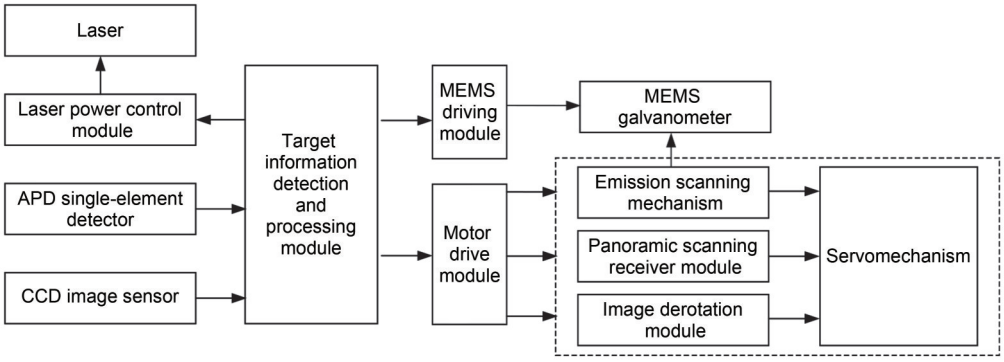

Fig. 4 Structural diagram of the scanning optical machine

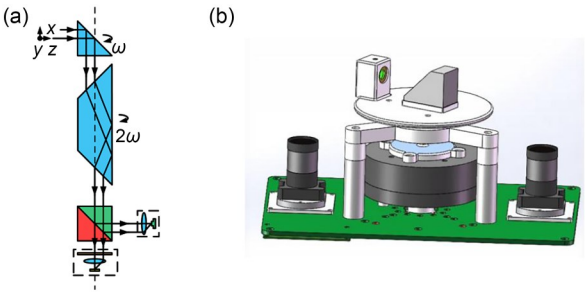

Fig. 5 Collimated optical path omnidirectional scanning diagram (a) and optical-mechanical simulation structure diagram (b)

single lens and a highly reflective surface. It is assumed that the diameter of the incident laser beam is larger than the aperture diameter of the miniature lens, and that the center of the incident beam coincides with the center of the miniature lens's surface. According to the structure of the miniature lens and the principle of optical path reversibility, it can be simplified into a combined model of a single lens and a highly reflective surface, as shown in Fig. 6a. The light path propagation process of the beam through the "cat's eye target" can be expanded using Fourier transform theory symmetrically along the reflective surface, as shown in Fig. 6b. The focusing lens can be regarded as the entrance pupil of the "cat's eye target" optical system, and the reflective surface is viewed as the focal plane of the focusing lens.

In Fig. 6b, lenses 1 and 2 are micro lenses with aperture sizes in the millimeter scale. The lens focal plane is the reflective surface. L and D represent the detection distance and the diaphragm diameter, respectively. The main optical axis of the system is the z axis; the direction perpendicular to the paper surface is the y axis, and the direction perpendicular to the y axis and z axis is the x axis. $U_i(x_i, y_i)$ represents the two-dimensional (2D) distribution of the light field at the z=z_i (i=0, 1, 2, 3) plane, and f is the lens focal

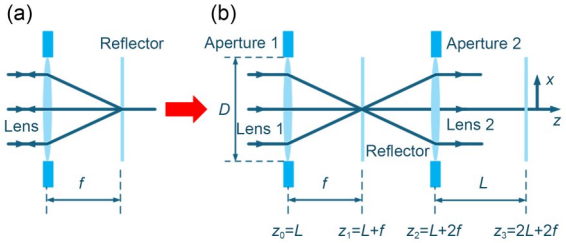

Fig. 6 Ideal model diagram (a) and equivalent model diagram (b)

length. The laser enters from the left, passes through lens 1, and converges at the focal plane before reflecting. When the detection laser reaches the micro lens, two diffractions will occur. The first scenario is when the laser passes through the lens, reaches the focal plane, and is reflected. In contrast, the second is when the laser passes through the lens and propagates directly to the receiving screen. Since the detection distance is longer than the lens focal length, the diffraction phenomenon at the latter stage is significantly stronger than that at the former stage.

Since the phase transformation effects of lenses 1 and 2 on the light waves cancel each other out during the transmission process, the detection laser between these two lenses is equivalent to being transmitted in the form of parallel light. At the same time, the convergence point at the focal plane can be enlarged. Thus, the phase transformation effect of the lenses on the light waves can be ignored. After the detection laser reaches lens 1, it receives the spatial modulation effect of the aperture first. Only the central part of the light can pass through the aperture. This part of the light converges onto the reflective surface after passing through lens 1. According to the theory of angular spectrum diffraction, before the detection laser reaches lens 2, the expression of the light field is as follows:

$$U_2(x_2,y_2)=\mathscr{F}^{-1}\left\{\mathscr{F}\left[U_0(x_0,y_0)\operatorname{circ}\left(\frac{\sqrt{x_0^2+y_0^2}}{r_0}\right)\right]\cdot\exp\left[\mathrm{j}\frac{4\pi f}{\lambda}\sqrt{1-\left(\lambda f_{x0}\right)^2-\left(\lambda f_{y0}\right)^2}\right]\right\}. \quad (2)$$

In Eq. (2), $\mathscr{F}$ and $\mathscr{F}^{-1}$ represent the forward and inverse 2D Fourier transforms, respectively, $\operatorname{circ}\left(\sqrt{x_0^2+y_0^2}/r_0\right)$ is a circular domain function that characterizes the influence of the micro-lens equivalent aperture, f_{x0} and f_{y0} are the components of the light field U_0 in the frequency domain, λ is the laser wavelength, and r_0 is the aperture radius. After the detection laser passes through lens 2, due to the influence of the equivalent aperture, the expression of the light field is

$$U_2'(x_2,y_2)=C_2(x_2,y_2)U_2(x_2,y_2), \quad (3)$$

where $C_2(x_2, y_2)$ is a function determined by the lens size and the incident angle.

When the laser is incident at an angle, the ray diagram is illustrated in Fig. 7, where θ is the angle between the incident beam in the z direction and the lens principal axis.

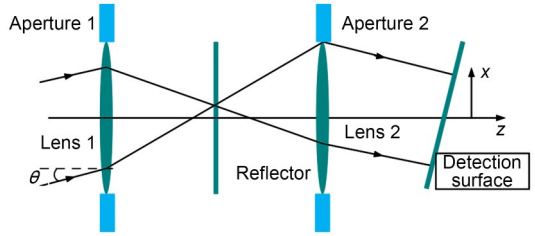

Fig. 7 Schematic diagram of the transmission of an obliquely incident light beam

Considering the existence of the incident angle θ, the optical field of the Gaussian beam before lens 1 is modified as

$$U_0'(x_0,y_0)=\frac{c}{\omega(z')}\exp\left\{-\frac{x_0'^2+y_0'^2}{\omega^2(z')}-\mathrm{j}k\left[z'-z_0+\frac{x_0'^2+y_0'^2}{2R_g(z')}\right]\right\}, \tag{4}$$

$$x_0'=x_0\cos\theta, \tag{5}$$

$$y_0'=y_0\cos\theta, \tag{6}$$

$$z'=z+x_0\sin\theta, \tag{7}$$

where ω represents the variation of the waist radius of the Gaussian beam with the propagation distance z', k is a correlated Gaussian beam, which describes the phase variation of the light wave along the propagation direction, and R_g denotes the curvature radius of the Gaussian beam.

The angle between the incident light and the optical axis, as well as the defocus amount of the reflecting surface, will prevent some of the light from reaching lens 2 during the transmission process, distorting the shape of the detected laser spot and subsequently affecting the function expression of C_2, that is

$$C_2(x_2,y_2)=\mathrm{circ}\left(\frac{\sqrt{x_2^2+y_2^2}}{r_0}\right)\mathrm{circ}\left[\frac{\sqrt{(x_2-2f\tan\theta)^2+y_2^2}}{r_0(1-2d/f)}\right], \tag{8}$$

where d represents the defocus amount of the reflecting surface. The presence of the defocus amount will alter the equivalent radius of lens 2. Combined with the process at vertical incidence, the light field distribution at the far-field receiving end can be obtained in Eq. (9) at the bottom of this page.

This represents the diffraction effect light field distribution of a Gaussian laser beam upon irradiating a micro lens. The light intensity distribution is

$$I(x_3,y_3)=U_3(x_3,y_3)U_3^*(x_3,y_3), \tag{10}$$

where U_3 is the size of the light field distribution at the receiving end, and U_3^* is its conjugate representation.

When the incident laser beam detects the micro lens, the optical field is analyzed in the frequency domain based on the angular spectrum diffraction theory. The numerical simulation analysis is conducted in MATLAB.

2.4.2 Micro lens recognition algorithm

Collaborative detection and recognition based on echo intensity images and distance information involves two strategies. One is integrating the local pyramid attention (LPA) module into the field pyramid network (FPN), combined with SKNet21 to identify and eliminate false alarm targets; the other is using the high-performance APD detector, combined with echo intensity and flight time, to achieve collaborative recognition of "cat's eye targets" and reduce the false alarm rate.

The basic network module we proposed is based on an improved residual network, SKNet. It can adjust the receptive field size adaptively according to the

$$U_3(x_3,y_3)=\mathscr{F}^{-1}\left\{\mathscr{F}\left\{C_2(x_2,y_2)\mathscr{F}^{-1}\left\{\mathscr{F}\left[U_0'(x_0,y_0)\mathrm{circ}\left(\sqrt{x_0^2+y_0^2}/r_0\right)\right]\exp\left[\mathrm{j}\frac{4\pi f}{\lambda\cos\theta}\sqrt{1-(\lambda f_{x0})^2-(\lambda f_{y0})^2}\right]\right\}\right.$$
$$\left.\cdot\mathrm{circ}\left(\sqrt{x_2^2+y_2^2}/r_0\right)\mathrm{circ}\left[\frac{\sqrt{(x_2-2f\tan\theta)^2+y_2^2}}{r_0(1-2d/f)}\right]\right\}\exp\left[\mathrm{j}\frac{2\pi L}{\lambda}\sqrt{1-(\lambda f_{x2})^2-(\lambda f_{y2})^2}\right]\Bigg\}. \tag{9}$$

multi-scale input information and select information captured by the receptive field that is effective for classification and extraction, as shown in Fig. 8.

The split stage of SKNet involves passing the input $\boldsymbol{I}_{\text{in}}$ through 3×3 grouped convolution and 5×5 dilated convolution with a receptive field, generating two different feature maps. At the fuse stage, two feature maps are summed in an element-wise manner to generate $\boldsymbol{U}$, which is then subjected to adaptive global average pooling to produce feature map $\boldsymbol{S}$. $\boldsymbol{S}$ is passed through two fully connected layers to generate vector $\boldsymbol{Z}$. $\boldsymbol{Z}$ is divided into two weights with different values a and b by applying the Softmax function, where the sum of the values of a and b is 1. The two feature maps are respectively multiplied by the obtained a and b. The two weighted feature maps are fused to obtain the output feature $\boldsymbol{I}_{\text{out}}$ at the select stage. The output of the convolution module SK-conv, which incorporates various receptive field combinations, significantly enhances the representation of feature information. Therefore, the information obtained from different receptive fields can expand the perception range of neurons, which is more conducive to the network's distinction between "cat's eye targets" and background clutter, providing better recognition effects for different-sized "cat's eye targets."

To achieve efficient algorithm processing while reducing network parameters and computational load, the SKNet module is constructed as a 21-layer network, known as the SKNet21 twin network. The network takes 3-channel candidate target regions as its input. After passing through two SKNet21 networks, the obtained two multi-dimensional features are expanded into one-dimensional (1D) vectors. These two 1D vectors are subtracted, and the absolute value is summed, which is equivalent to calculating the L_1 norm of the distance of the two feature vectors. Then, the result is processed by two fully connected layers, and the output is a neuron. Finally, the result is subjected to a sigmoid function, making the value range from 0 to 1, which indicates the degree of similarity between the "cat's eye target" sample and the image to be detected.

The LPA module is integrated into multi-scale FPN outputs of the SKNet21 Siamese network to enhance the initial detection of small UAV targets by emphasizing target–background relationships and incorporating high-level semantic information across channels, thereby improving the accuracy of preliminary target screening, as illustrated in Fig. 9.

The LPA structure employed uses both channel attention and spatial attention mechanisms to enhance contextual modeling capabilities. The channel attention explores dependencies across different feature channels, amplifying relevant features while suppressing irrelevant ones. The spatial attention models dependencies between distant pixels across various channels. When applying the channel attention module, the input features are processed through a 3×3 pooling kernel with a stride of 2, followed by a max pooling layer. This is then followed by a 2×2 pooling kernel with a stride of 2, and finally, an average pooling layer. This yields aggregated feature images from different channels. Subsequently, two 3×3 convolutional kernels with a stride of 1 are applied to perform feature interaction on the two pooled channel information. Finally, the maximum and average pooled feature maps are summed using a sigmoid activation function to generate the channel attention map.

After channel attention calibration, the LPA performs spatial attention modeling on the calibrated feature images. It employs dilated convolutions with varying dilation rates to extract multi-scale spatial features, and then merges these features across different scales. Following a pixel shuffle operation, pixels from different scales are grouped and rearranged to form

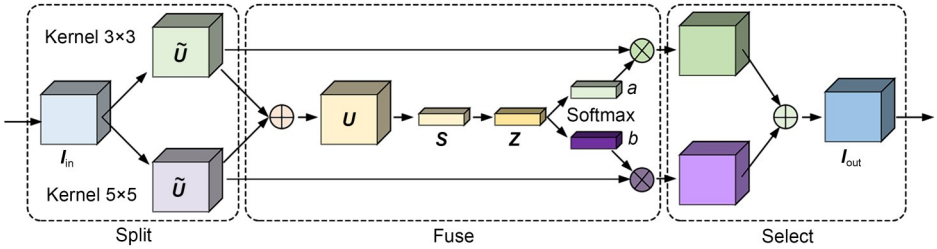

Fig. 8 Schematic diagram of SKNet

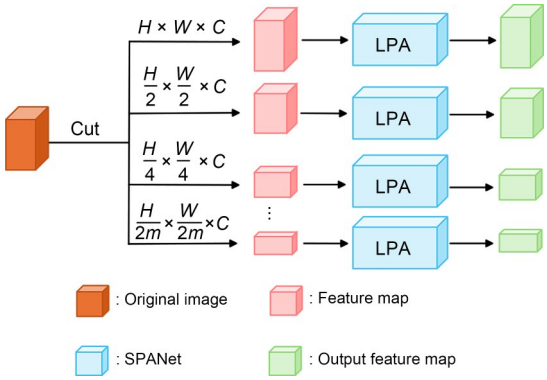

Fig. 9 Diagram of the target area filtering network framework. *H*: the height dimension; *W*: the width dimension; *C*: the channel dimension; *m*: the number of segmented images

feature sets. A 1D convolution then models correlations between pixels at different scales. A sigmoid activation function is applied to generate the spatial attention map. Finally, the feature map calibrated by the channel attention mechanism is combined with that calibrated by the spatial attention mechanism to produce the final feature map, as shown in Fig. 10.

The circularity characteristics of the "cat's eye effect" echo from the micro lens will be damaged to varying degrees, according to the above analysis. Therefore, in addition to the gray value, one should check whether the candidate area conforms to other optical characteristics of the "cat's eye target" echo. By observing the brightness characteristics of the "cat's eye target" echoes, the shape of the light spot in the image can be analyzed, appearing as solid, approximately circular, elliptical, or elongated. The real target can be identified based on the characteristics of the reflected light spot. A comprehensive shape measurement value Δ will be adopted as the criterion, which can be expressed as

$$\begin{cases} R_0 = S^2/(4\pi A), \\ \Delta = |1 - R_0|, \end{cases} \tag{11}$$

where S represents the perimeter of the echo spot, and A and R_0 ($R_0 \in [0,1]$) denote the area and the circularity, respectively. After obtaining the refined candidate area of the target, a twin network can be used to extract the target from this area.

3 Experimental results and analysis

3.1 Construction of the dataset

The dataset construction requires consideration of the UAV's flight altitude and speed, with a flight altitude ranging from 50 to 200 m and a flight speed

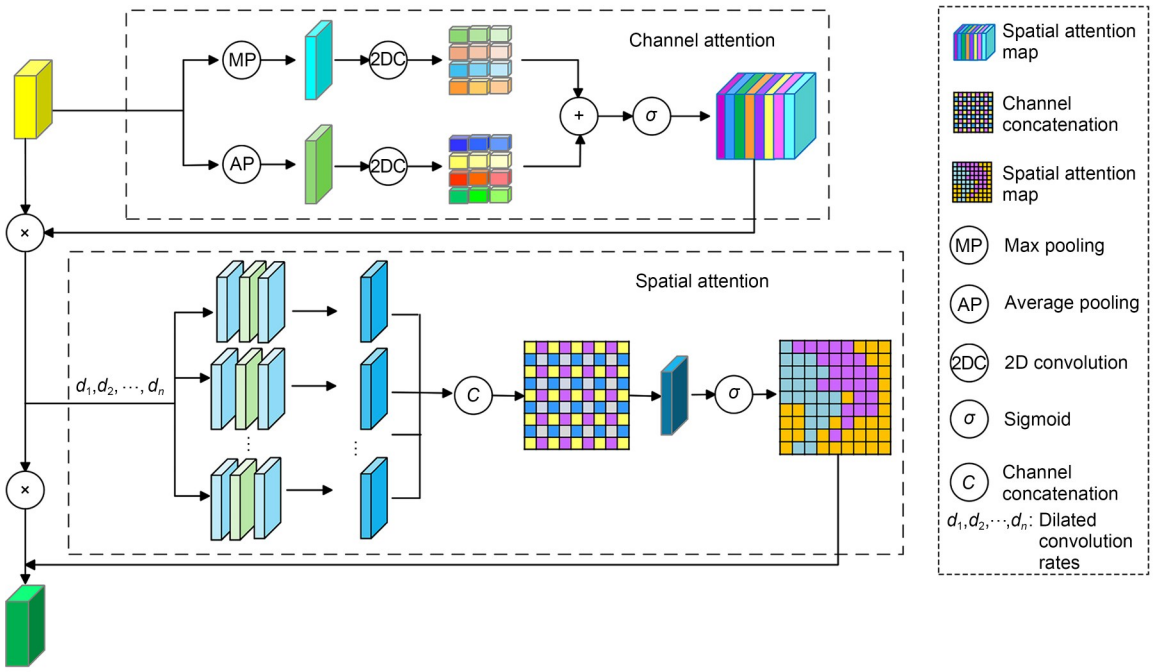

Fig. 10 Schematic diagram of the LPA module

maintained between 10 and 20 km/h. By collecting 200 data points across various time periods throughout the day, we capture the "cat's eye effect" characteristics under varying light intensities and angles. Images with clarity scores below 0.5 are discarded. Additionally, images in which the "cat's eye effect" area constitutes less than 1% of the total image area are deemed invalid, ensuring that the retained images exhibit distinct features. Because of the limited field of view of the miniature lens during actual sampling, detection becomes challenging at larger angles of inclination. To collect more positive samples, the "cat's eye target" is mounted on a bracket at a certain distance to acquire partial data. Each echo image in the dataset has a resolution of 993×662 pixels, 24-bit color depth, and a file size of 42.5 KB. Three representative airborne scenarios are selected for testing. The dataset characteristics are described in Table 2.

Table 2 Dataset characteristics description

Labeling	Total number of frames	Scene characteristics
Scene 1	190	Close range, single target, sky background
Scene 2	150	Close range, single target, ground–air interface background
Scene 3	210	Close range, multiple targets, sky background

The training dataset requires both positive and negative sample sets. Generally, to minimize interference during target detection and reduce computational overhead, positive samples should contain only the target's image information, while negative samples must exclude the target and avoid duplication. To construct the positive sample library, the detection algorithm described above is applied to detect small objects within the sample training dataset. The negative sample library is established by randomly cropping samples from the test dataset. Partial screenshots of the dataset are shown in Fig. 11.

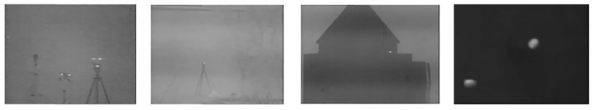

Fig. 11 Dataset selection diagram

3.2 Performance evaluation indices

The experimental setup consists of a computer equipped with an Intel® Core™ i5-8250U processor (CPU @1.60 GHz), 12 GB RAM, and an NVIDIA GeForce GTX 1080 Ti graphics card, running on the Linux operating system with Python 3.1.1.

This paper evaluates model performance using the following metrics: precision (P), recall (R), average precision (AP), mean average precision (mAP), model parameters, and the number of floating-point operations (Giga floating-point operations per second, GFLOPs) performed by the model.

$$P = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}, \tag{12}$$

$$R = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \tag{13}$$

$$\mathrm{AP} = \int_0^1 P(R)\,\mathrm{d}R, \tag{14}$$

$$\mathrm{mAP} = \frac{\sum_{t=1}^{M} \mathrm{AP}(t)}{M}. \tag{15}$$

Here, P refers to the proportion of samples correctly predicted as positive among those predicted as positive, R is the proportion of all positive samples successfully predicted as positive, TP is the true positive examples, which are the cases where the model correctly identifies the target, FP is the false positive examples, which are the cases of incorrect detection, FN is the false negative example, that is, the cases where the model wrongly predicts the sample which is actually of the positive class as the negative class.AP is the average of accuracy values at different recall points. The mAP represents the average value of the AP scores across all classes detected by the model. M denotes the total number of categories in the detection task, and t indicates the t^{th} category. Since this paper focuses solely on detecting small low-altitude targets, the AP and mAP values are identical. mAP@0.50 is used as an indicator for evaluating model accuracy, and mAP@0.50∶0.95 represents the mean of the mAP value at an intersection over union (IoU) of 0.50–0.95. The stride is 0.05. The number of parameters refers to the total number of parameters in the model—a smaller value indicates lower model complexity. A greater number of floating-point operations performed by the model indicates a better graphics processing unit (GPU) performance.

3.3 Analysis of the "cat's eye effect" echo characteristics

3.3.1 Simulation analysis

In the numerical simulation analysis using MATLAB, the initial parameters of the detection laser and miniature lens are assumed to be λ=1064 nm, ω_0=4 mm, f=1.6 mm, d=0, with a sampling grid of 800×800. The echo diffraction patterns of the laser incident perpendicularly on the ideal "cat's eye target" at different distances can be obtained by using the Fourier transformation. Fig. 12 shows the distribution of the diffraction light intensity at 500 and 5000 m.

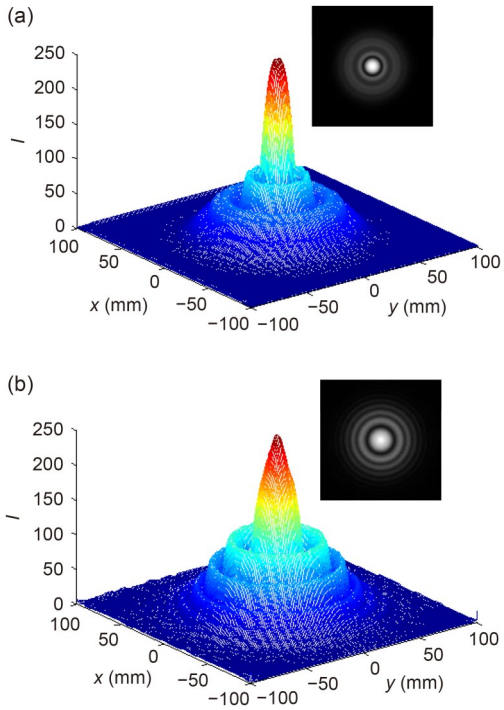

Fig. 12 Three-dimensional (3D) intensity distribution and 2D spot pattern of the echo from an ideal "cat's eye target" under normal incidence conditions: (a) L=500 m; (b) L= 5000 m. I is the gray value

As shown in Fig. 12, under normal incidence conditions, the distribution of echo intensity changes significantly with an increase in the transmission distance. The farther the transmission distance, the larger the echo's diffraction light spot, with the light intensity distribution presenting a multi-layer diffraction circular pattern. While the energy is mainly concentrated within the central peak at a detection distance of greater than 500 m, the energy of the side lobes cannot be ignored, which leads to difficulty defining the edge of the circular feature of the echo light spot.

When setting the laser oblique incidence condition as θ=30° or θ=60°, the corresponding 3D diffraction light field distribution and 2D light spot distribution patterns (50 m) are obtained, as shown in Fig. 13. It can be seen that oblique incidence causes the light spot to deform, and the diffraction effect is enhanced. As the inclination angle increases, the degree of light spot deformation becomes more obvious. This is due to the diffraction aperture gradually becoming smaller, resulting in a steadily reduced contour of the light field distribution and a gradually enhanced diffraction effect.

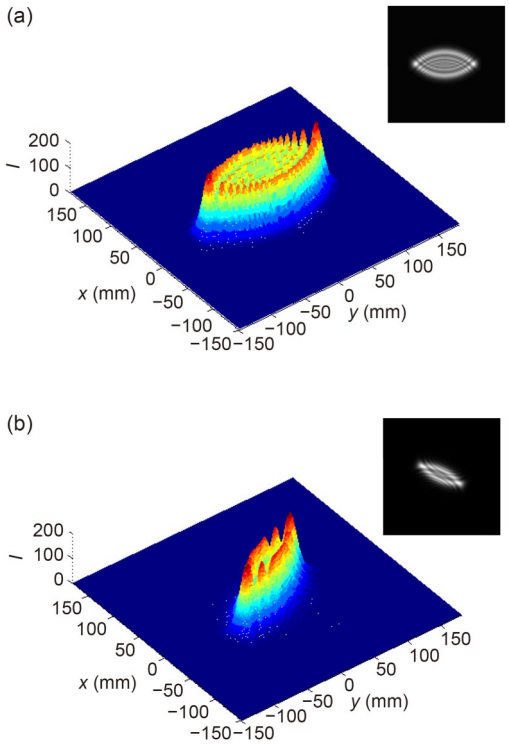

Fig. 13 3D intensity distribution and 2D spot pattern of the outgoing beam of the ideal "cat's eye target" under the condition of laser oblique incidence: (a) θ=30°; (b) θ=60°. I is the gray value

The simulations reveal that when detecting the micro lenses carried by UAVs, the characteristic parameters that represent the circularity of the echo spots, such as circularity, eccentricity, and eccentric distance, are challenging to extract. Using the conventional method of judging the spot's shape for detection and recognition has significant limitations.

3.3.2 Analysis of physical echoes

Based on the echo information of the "cat's eye target" collected under conditions of direct incidence at 50 and 2500 m, as shown in Fig. 14, it can be seen that the near-infrared laser collected by the image sensor produces relatively bright light spots on the image. However, such images have low resolution and blurred target details, which limit subsequent target recognition (Elyousseph and Altamimi, 2024; Wang G et al., 2024; Nguyen et al., 2025). Therefore, different deep learning methods are adopted for comparative analysis.

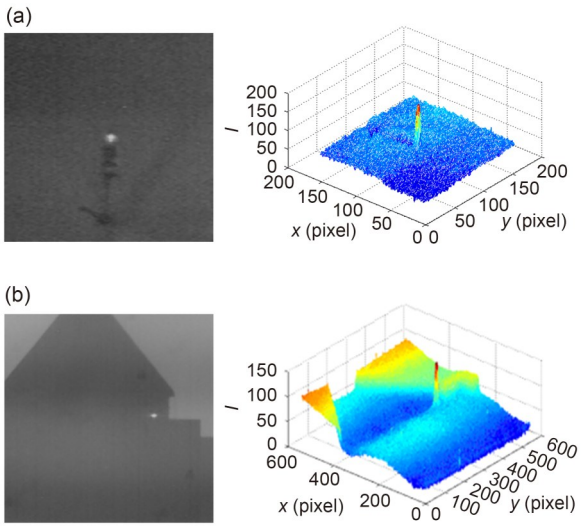

Fig. 14 Echo patterns of "cat's eye targets" at different distances: (a) 50 m; (b) 2500 m. *I* is the gray value

Five images are selected from the captured "cat's eye targets" dataset for validation analysis, as shown in Fig. 15. The algorithms employed include YOLOv8 (Chang and Wang, 2024), YOLOv10 (Cai et al., 2025), YOLOv8-UD (Huangfu and Li, 2023), as well as the algorithm proposed in this paper. The YOLOv8 network architecture is a three-stage structure comprising a backbone network, a feature fusion neck, and a detection head, which demonstrates high stability and maturity for small object detection. YOLOv10 is an enhanced version of YOLOv8, offering improved accuracy and processing speed for small objects, as well as strong detection capabilities in complex backgrounds. The YOLOv8-UD architecture incorporates a small object detection head into the YOLOv8 structure to enhance small object feature detection. Table 3 compares metrics across different algorithms, indicating that our algorithm outperforms other methods in terms of

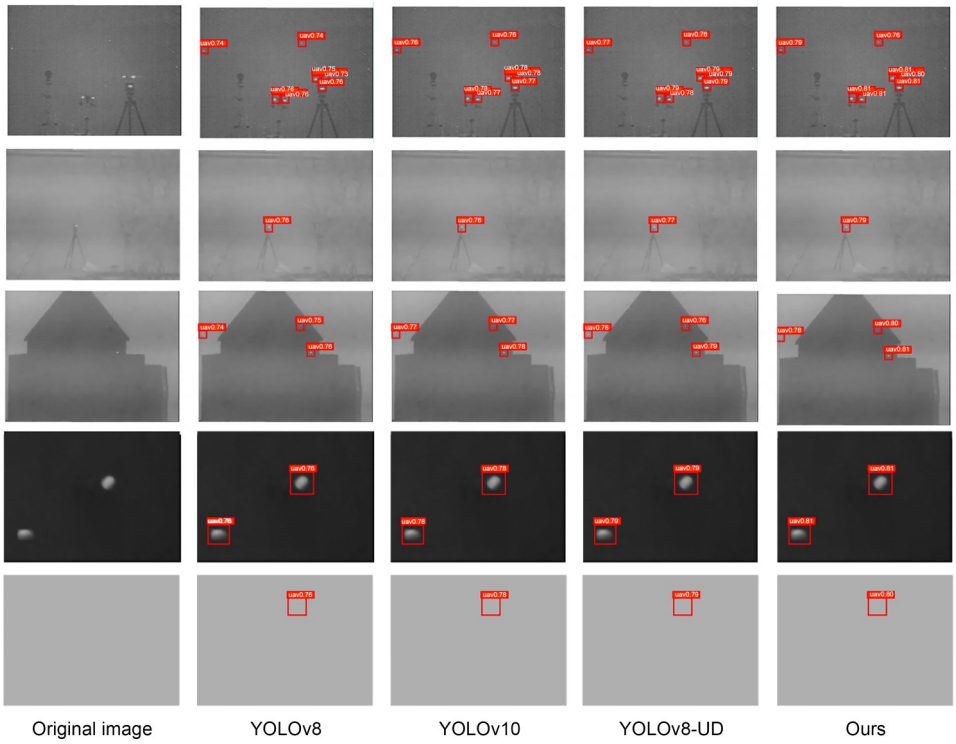

Fig. 15 Different algorithms' processing of the comparison image

Table 3 Comparison of different algorithm metrics

Method	mAP (0.50)	mAP (0.50:0.95)	Parameter efficiency	GFLOPs
YOLOv8	0.758	0.273	1.245×10^7	30.1
YOLOv10	0.773	0.289	1.228×10^7	32.5
YOLOv8-UD	0.786	0.311	**1.142×10^7**	36.7
Ours	**0.809**	**0.324**	1.155×10^7	**49.8**

Best results are in bold

mAP (0.50), mAP (0.50:0.95), and GFLOPs. Notably, the YOLOv8-UD algorithm demonstrates superior model parameter efficiency, primarily due to its implementation of the cross stage partial (CSP) module, which enhances detection speed while preserving feature extraction capabilities.

3.3.3 Ablation experiment

Ablation studies are conducted to evaluate the contribution of each proposed module to the performance of the original SKNet. The experimental results are shown in Table 4.

The ablation experiments are divided into four groups. Experiment 1 employs the original SKNet algorithm, achieving an mAP (0.50) of 0.754 and an mAP (0.50:0.95) of 0.268. Experiment 2 modifies the original model into a dual-network SKNet architecture, yielding only a marginal improvement in mAP (0.50). Experiment 3 incorporates an LPA structure into the SKNet dual-network architecture. The data demonstrate that this network structure elevates mAP (0.50) to 0.796 and mAP (0.50:0.95) to 0.301, effectively enhancing small object detection rates. Experiment 4 builds upon Experiment 3 by incorporating an APD ensemble module to enhance cat-eye object recognition. This ultimately increases mAP (0.50) to 0.809 and mAP (0.50:0.95) to 0.324. Compared to Experiment 1, these represent improvements of 5.5 percentage points (PPs) and 5.6 PPs, respectively, demonstrating the effectiveness and superiority of the proposed algorithm.

3.4 Comparison of echo waveforms of "cat's eye targets"

Once the laser has scanned the target area and detected the "cat's eye target" optical window, the reflected echo is focused by the receiving optical system onto the APD unit detector. The avalanche operation of the APD converts the echo signal intensity into an electrical signal, which is sent to the high-speed data acquisition and processing system for processing. This echo appears as a distinct voltage peak in the electrical signal. Based on the results of short-range (300 and 500 m) experiments, when the average output power of the high-repetition-frequency laser (20 kHz) is 0.2 W, the echo waveform can be obtained, as shown in Fig. 16. The echo intensities of both the near-range "cat's eye target" and the diffuse reflection target reach very high amplitudes after being amplified by the amplifier connected after the APD detector, and there is a significant broadening in the pulse width. It is impossible to distinguish the "cat's eye target" from the diffuse reflection background based on the peak values.

In low-altitude detection, the laser output power is first calibrated and adjusted experimentally to ensure a distortion-free output waveform. The system's medium- and long-range detection capabilities are subsequently validated. The echo waveform carries the target's spatial information, while the distance is derived by calculating the time of flight through a delay line array.

Table 4 Melting results

SKNet	SKNet21	LPA	APD	mAP (0.50)	mAP (0.50:0.95)	Parameter efficiency	GFLOPs
√				0.754	0.268	1.113×10^7	29.3
√	√			0.768	0.285	1.113×10^7	29.3
√	√	√		0.796	0.301	1.047×10^7	27.1
√	√	√	√	0.809	0.324	1.135×10^7	49.8

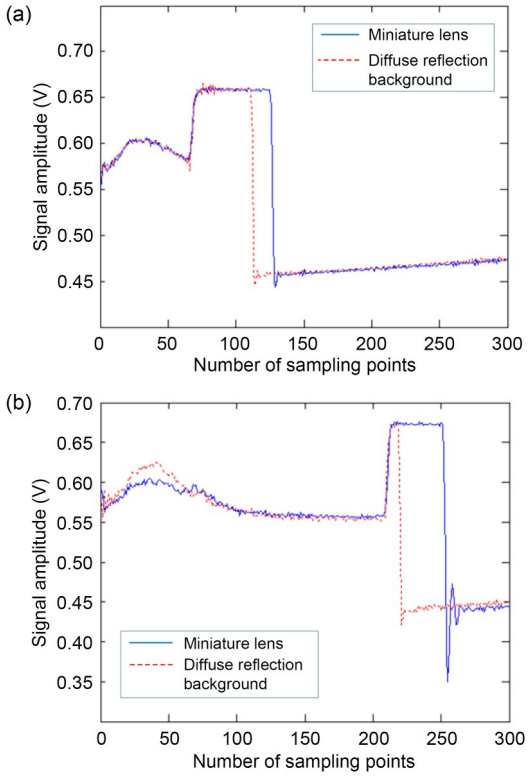

Fig. 16 "Cat's eye target" and reflection target echo waveforms: (a) L=300 m; (b) L=500 m

4 Conclusions

This paper overcomes multidimensional technical bottlenecks in "cat's eye target" detection by establishing an innovative methodology that spans hardware collaboration to algorithm optimization. By designing laser transmission and dual-channel reception modules, this technique can achieve the simultaneous acquisition and fusion of image and distance information. Employing Dove prism racemization and differential speed synchronization control resolves image rotation interference during omnidirectional scanning. An oblique-incidence echo model based on angular spectrum diffraction theory quantifies the influence of incidence angle and defocuses on spot distortion. Algorithmically, it integrates LPA mechanisms with SKNet21 twin networks to enhance the accuracy of small target screening. By introducing comprehensive shape metrics, this approach overcomes the limitations of traditional recognition methods, which rely on grayscale features, thereby reducing false alarms and missed detections. Research will advance in two key areas: first, by promoting hardware-algorithm co-evolution through the development of adaptive optical systems with real-time wavefront correction to address oblique-incidence spot distortion, and second, by expanding multi-scenario adaptability via constructing a universal dataset that spans various distances and lighting conditions. Transfer learning will optimize model generalization to enable the tracking and recognition of violating vehicles in complex urban environments, further enhancing the anti-interference capabilities of small target detection and paving new pathways for drone-based "cat's eye target" detection.

Contributors

Bin ZHOU conceptualized the innovations and drafted the paper. Weiming WANG and Ning YAN processed the data and revised the paper. Linlin ZHAO and Chuanzhen LI helped organize and finalized the paper.

Conflict of interest

All the authors declare that they have no conflict of interest.

Data availability

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

References

AlKhonaini A, Sheltami T, Mahmoud A, et al., 2024. UAV detection using reinforcement learning. *Sensors*, 24(6):1870. https://doi.org/10.3390/S24061870

Bao L, Guo ZT, Gao XZ, et al., 2025. Stealth UAV path planning based on DDQN against multi-radar detection. *Aerospace*, 12(9):774. https://doi.org/10.3390/aerospace12090774

Bishop BB, Tufariello J, 2019. Overview of acoustic detection for counter unmanned aircraft systems. *J Acoust Soc Am*, 145(3):1661. https://doi.org/10.1121/1.5101099

Cai C, Wang YW, Ma PF, et al., 2025. Improvement and implementation of UAV target detection algorithm based on YOLOv10. *Acad J Comput Inform Sci*, 8(7):72-78. https://doi.org/10.25236/AJCIS.2025.080709

Chang H, Wang ZS, 2024. UAV vehicle detection system based on YOLOv8. *J Phys Conf Ser*, 2872(1):012019. https://doi.org/10.1088/1742-6596/2872/1/012019

Du XC, Wang CR, Zhou HR, et al., 2024. Progress in laser active detection technology based on the cat's eye effect. *Results Phys*, 67:108030. https://doi.org/10.1016/j.rinp.2024.108030

Elyousseph H, Altamimi M, 2024. Robustness of deep-learning-based RF UAV detectors. *Sensors*, 24(22):7339. https://doi.org/10.3390/s24227339

Fan YC, Zhang LX, Liu BL, et al., 2020. The influence of cat's eye effect on CEMRR FSO gain under defocus oblique

incidence. Proc SPIE Advanced Laser Technology and Application, Article 115620M. https://doi.org/10.1117/12.2579011

Farlik J, Kratky M, Casar J, et al., 2019. Multispectral detection of commercial unmanned aerial vehicles. *Sensors*, 19(7):1517. https://doi.org/10.3390/s19071517

Figueiredo A, Amaral J, Mendes M, et al., 2025. Experimental dataset of video and radar detection for cooperative perception in urban environment. *Data Brief*, 63:112091. https://doi.org/10.1016/j.dib.2025.112091

Goldman GH, 2016. Acoustic detection results for a small unmanned aircraft system extrapolated over range. *J Acoust Soc Am*, 140(S4):3120. https://doi.org/10.1121/1.4969771

Huangfu ZM, Li SQ, 2023. Lightweight you only look once v8: an upgraded you only look once v8 algorithm for small object identification in unmanned aerial vehicle images. *Appl Sci*, 13(22):12369. https://doi.org/10.3390/app132212369

Li J, Ye DH, Kolsch M, et al., 2021. Fast and robust UAV to UAV detection and tracking from video. *IEEE Trans Emerg Top Comput*, 10(3):1519-1531. https://doi.org/10.1109/TETC.2021.3104555

Liaquat S, Faizan M, Chattha JN, et al., 2024. A framework for preventing unauthorized drone intrusions through radar detection and GPS spoofing. *Ain Shams Eng J*, 15(5): 102707. https://doi.org/10.1016/j.asej.2024.102707

Liu BL, Luo H, 2022. An improved Yolov5 for multi-rotor UAV detection. *Electronics*, 11(15):2330. https://doi.org/10.3390/ELECTRONICS11152330

Liu YC, Liao LC, Wu H, et al., 2021. Trajectory and image-based detection and identification of UAV. *Visual Comput*, 37(7):1769-1780. https://doi.org/10.1007/s00371-020-01937-y

Lv Z, Du XC, Zhang K, et al., 2022. Three-dimensional echo distribution analysis of multi-band and oblique Gaussian beams propagating through cat-eye optical system. *Opt Laser Technol*, 151:108044. https://doi.org/10.1016/j.optlastec.2022.108044

Nguyen PT, Nguyen GL, Bui DD, 2025. LW-UAV–YOLOv10: a lightweight model for small UAV detection on infrared data based on YOLOv10. *Geomatica*, 77(1):100049. https://doi.org/10.1016/J.GEOMAT.2025.100049

Rahman MH, Sejan MAS, Aziz MA, et al., 2024. A comprehensive survey of unmanned aerial vehicles detection and classification using machine learning approach: challenges, solutions, and future directions. *Remote Sens*, 16(5):879. https://doi.org/10.3390/rs16050879

Randieri C, Ganesh SV, Raj RDA, et al., 2025. Aerial autonomy under adversity: advances in obstacle and aircraft detection techniques for unmanned aerial vehicles. *Drones*, 9(8): 549. https://doi.org/10.3390/DRONES9080549

Wang G, Zhang YM, Ai Q, 2024. Lightweight aerial target detection algorithm with enhanced small target perception. *IAENG Int J Comput Sci*, 51(12):2123-2134.

Wang XW, Hu Y, Liang Q, et al., 2025. An improved YOLOv8s-based UAV target detection algorithm. *PLoS ONE*, 20(8): e0327732. https://doi.org/10.1371/journal.pone.0327732

Xu L, Luo ZQ, 2025. Anti-UAV detection and identification technology: fundamentals, methods and challenges. *Phys Commun*, 71:102676. https://doi.org/10.1016/j.phycom.2025.102676

Yasir PAA, Goyal SK, 2021. Polarization selective Dove prism. *Opt Express*, 29(10):14917-14930. https://doi.org/10.1364/OE.420891

Zakaria NAZ, Rozaimi FNSM, Ab Rahim SAE, et al., 2024. Signal analysis for drone detection and characterization using acoustic radar. *J Phys Conf Ser*, 2922:012003. https://doi.org/10.1088/1742-6596/2922/1/012003

Zhang QQ, Zhou L, An JS, 2024. Real-time recognition algorithm of small target for UAV infrared detection. *Sensors*, 24(10):3075. https://doi.org/10.3390/S24103075

Zhang WH, Li ZH, Wang YR, et al., 2025. Sensitive micro UAV detection based on a high-repetition-rate single-photon LiDAR. *Opt Express*, 33(8):18102-18111. https://doi.org/10.1364/OE.554494

Zheng CW, Liu LM, Fu Q, et al., 2025. YOLO-DD: a lightweight framework for UAV detection in complex environments via boundary-aware fusion. *EURASIP J Adv Signal Process*, 2025(1):44. https://doi.org/10.1186/s13634-025-01253-4