




## SURVEY PAPER

# Detection and recognition of UA targets with multiple sensors

W.S. Chen<sup>1,\*</sup> , X.L. Chen<sup>2</sup>, J. Liu<sup>3</sup>, Q.B. Wang<sup>1</sup>, X.F. Lu<sup>1</sup> and Y.F. Huang<sup>1</sup>

<sup>1</sup>China Academy of Civil Aviation Science and Technology, Beijing 100028, China, <sup>2</sup>Naval Aviation University, Yantai 264001, China and <sup>3</sup>Beihang University, Beijing 100191, China

\*Corresponding author. Email: [wishchen@buaa.edu.cn](mailto:wishchen@buaa.edu.cn)

**Received:** 25 February 2022; **Revised:** 18 April 2022; **Accepted:** 19 April 2022

**Keywords:** UA; Detection; Deep learning; Multi-sensor; Information fusion

### Abstract

Modern low-altitude unmanned aircraft (UA) detection and surveillance systems mostly adopt the multi-sensor fusion technology scheme of radar, visible light, infrared, acoustic and radio detection. Firstly, this paper summarises the latest research progress of UA and bird target detection and recognition technology based on radar, and provides an effective way of detection and recognition from the aspects of echo modeling and micro motion characteristic cognition, manoeuvre feature enhancement and extraction, motion trajectory difference, deep learning intelligent classification, etc. Furthermore, this paper also analyses the target feature extraction and recognition algorithms represented by deep learning for other kinds of sensor data. Finally, after a comparison of the detection ability of various detection technologies, a technical scheme for low-altitude UA surveillance system based on four types of sensors is proposed, with a detailed description of its main performance indicators.

### Nomenclature

UA	unmanned aircraft
DNNs	deep neural networks
LSS	Low Slow Small
CFAR	constant false alarm rate
MTI	moving target indicator
RCS	radar cross section
MIMO	multi input multi output
FMCW	frequency modulated continuous wave
DTMB	digital terrestrial multimedia broadcast
m-D	micro-Doppler
STFT	short time Fourier transform
SVM	support vector machine
NBC	naive Bayes classifier
FFT	Fast Fourier transform
SVD	singular value decomposition
CVD	cadence velocity diagram
EMD	empirical mode decomposition
PCA	principal component analysis
LTCI	long time coherent integration
NUSP	non-uniform sampling order reduction
RFT	Radon Fourier Transform
CNNs	convolutional neural networks
SNR	signal to noise ratio

MLP	multi-layer perceptron
YOLO	You Only Look Once
MFCCs	Mel-frequency cepstral coefficients
RDF	Radio Direction Finding
TDOA	time difference of arrival

## 1.0 Introduction

At present, unmanned aircraft (UA) are widely used around the world, in fields including aerial photography, express transportation, emergency rescue, electric power inspection, agricultural plant protection, border monitoring, mapping, fire monitoring and environmental protection. However, the rapid popularisation of UA also brings serious security problems [1]. In recent years, the media have reported dozens of public security incidents caused by UA intrusion, whose main targets are airports, prisons, public buildings and other sensitive places [2].

UA detection systems generally use radar as the core, supplemented by photoelectric (visible and infrared), acoustic, radio detection and other sensors, aiming at achieving all-round situation awareness in the guarantee of major events and the defense of important places. When multiple sensors are deployed, both visible and infrared devices have some classification ability and have more accurate positioning and ranging functions. Visible-light cameras are usually cheap, while infrared cameras are expensive, but both are sensitive to environmental conditions. In addition, although an acoustic sensor is not as sensitive to the environments, its limited detection range limits its application. Radio detection technology is sensitive to complex electromagnetic environments, and is ineffective to electromagnetically silent UA. Therefore, in view of radar's accurate positioning ability and large detection range, as well as its better target classification ability and environmental adaptability, it has become the most common UA detection means.

Since current multi-sensor UA detection systems cannot realise automatic operation, it is necessary to confirm the detected target manually. For example, the early detection results from the radar might guide the operator to observe the direction of the target through the optical camera for confirmation [3]. In recent years, the demand for multi-source information fusion in various applications is increasing day by day, which makes the data fusion technology widely important. The goal of data fusion is to make up for the weakness of a single sensor, so as to obtain more accurate detection results. On the other hand, artificial intelligence and deep neural networks (DNNs) have become a very attractive data processing method, which can find high-order abstract features that are difficult to find by typical feature extraction methods. Therefore, they are widely used in massive multi-source data processing, and have achieved good results in UA detection and classification.

This paper focuses on the research results of UA detection and classification using radar, visible light, infrared, acoustic sensors and radio detection as the data acquisition tools and deep learning as the main data analysis tool. In recent years, there has been a growing demand for multi-source information fusion in various applications, and data fusion technology has been widely considered [4]. The sensors selected in this paper have excellent performance in the typical monitoring system, and complement each other in the multi-sensor information fusion scheme. Radar is generally considered as reliable detection equipment, but in most cases, it is still difficult to distinguish small UA from birds [5]. Taking the airport clearance area as an example, birds and UA are the two main dangers that threaten the safety of flight during takeoff and landing. The airport needs to take different countermeasures after discovering the target. After the UA target is found, the airport will first issue an early warning, guide the takeoff and landing flights to avoid the target and coordinate with the airport and local public security for disposal. After a bird target is found, the airport will use a variety of bird repulsion equipment to drive it away from the dangerous area or take certain avoidance measures based on a bird driving strategy. Therefore, it is necessary to identify both UA and birds. Due to the wide variety of UAV and bird targets and their different size, shape and motion characteristics, the targets have different radar scattering and Doppler characteristics. Therefore, they belong to typical Low Slow Small (LSS) targets

with low observability. At present, the detection methods of LSS targets mainly involve constant false alarm rate (CFAR), moving target indicator (MTI), moving target detection (MTD), coherent accumulation, feature detection and other methods. When using the existing technology to detect LSS targets, due to the complex environment, many false targets, strong target mobility, short effective observation time and many other problems, it leads to low target detection probability, high-clutter false alarm, low accumulation gain and unstable tracking, which bring great difficulties to target detection and judgment, making the detection of birds and UA a worldwide problem.

This paper summarises the research progress of unmanned rotorcraft and bird target detection and recognition technology based on multi-source sensors in complex schemes in recent years, and points out its development trend. The paper is organised as follows. Section 2 introduces the radar detection technology for UA and bird targets and introduces the relevant system applications. Section 3 focuses on the technical methods of UA radar target detection and classification. Firstly, from the aspects of target characteristic cognition and feature extraction, echo modeling and micro-motion characteristic cognition methods are introduced to realise the fine feature description of the target. The method of distinguishing UA from birds according to the difference of motion trajectory and polarisation characteristics is discussed. Finally, combined with machine learning or deep learning methods, the effective technical approaches and related achievements of target intelligent recognition are summarised. Section 4 introduces multi-source detection methods of non-cooperative UA targets, such as visible light, infrared, acoustic and radio detection, and discusses the target detection and classification methods based on various technical means. On the basis of comparing the detection ability and performance of various sensors, Section 5 introduces several typical cases of multi-sensor fusion detection systems, and recommends a construction scheme of UA detection systems based on four types of sensors. Section 6 summarises the full text and proposes a research and development trend in future.

## 2.0 Radar technology

As the main means of air and sea target surveillance and early warning, radar is widely used in the field of national defense and public security. Compared with other technologies, radar is actually the only technology that can realise long-range detection from several to tens of kilometres, and is almost unaffected by unfavourable light and weather conditions. However, traditional surveillance radar is generally used to detect moving targets with relatively large radar cross section (RCS) and high speed, and it is not suitable for detecting low-altitude flying targets such as UA with very small RCS and slow speed (LSS) targets. In addition, the similarity between UA and birds is very high, so reliable classification of two kinds of targets is another problem. Therefore, special design is required for avian radar according to the above demanding requirements [6]. The scheme based on deep learning arranges the original data into a specific data structure, greatly reduces the workload of manual annotation through fusion processing and ensures the recognition effect of the system under the condition of low amount of data and low annotation [7].

### 2.1 UA detection system

The traditional surveillance radar usually uses a mechanically scanning antenna to search the whole airspace, so that the residence time on a single target is short. It can only obtain limited target data such as the azimuth, angle, radial velocity and RCS, so it is difficult to detect LSS targets. However, with the continuous development of radar technology in recent years, the ability to detect weak targets is steadily improving, which provides a new way to detect and recognise LSS targets. In recent years, new radar systems for weak target such as UA have become a research focus [8].

Some research institutions and enterprises, such as the University of London, Warsaw University of Technology, Fraunhofer-Gesellschaft, French National Aerospace Research Center, Robin Radar of the Netherlands, Aveillant Radar of the UK and the Wuhan University, conduct a series of UA detection and identification research projects with advanced radar technologies of active phased array, multi input

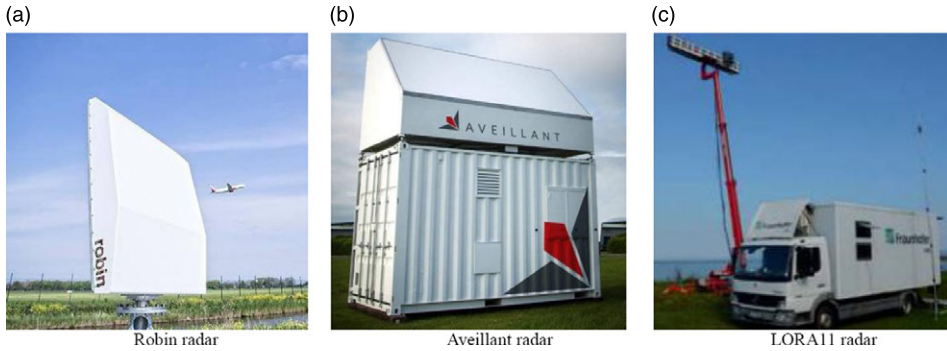


Figure 1. Typical UA detection radar system.

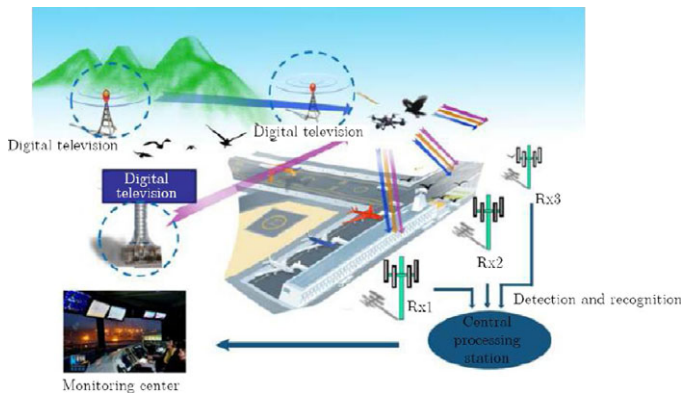


Figure 2. UA target surveillance application based on DTMB signal passive radar [9].

multi output (MIMO), frequency modulated continuous wave (FMCW), holographic and passive radar, which have developed preliminary engineering capacity. Figure 1 shows the Robin phased array radar, Aveillant holographic radar and Fraunhofer-Gesellschaft LORA11 passive radar for UA detection.

Active radar requires frequency allocation and strong electromagnetic radiation. However, in some application scenarios, such as airports, strict electromagnetic access systems are in place. The management department is usually cautious about the use of active radar, where even scientific experiments need to go through a strict approval process. However, passive radar has the advantages of low power consumption, concealment and good coverage. As an economical and safe means of non-cooperative target detection, it has attracted extensive attention in academic circles in recent years. Especially with the improvement of hardware performance and the development of available external radiation sources such as communication and wireless network, passive radar has attracted universal attention and extensive research in various countries. The radio wave propagation laboratory of Wuhan University has carried out a series of research projects on UA detection using passive radar [9–11]. It uses digital terrestrial multimedia broadcast (DTMB) signals to carry out UA detection experiments, and studies time-varying strong clutter suppression in urban complex environment. Experimental results show that after clutter suppression, the detection distance is more than 3km for a typical DJI Phantom 3. Figure 2 shows an example of a UA target surveillance application based on DTMB signal passive radar.

### 2.2 Avian radar system

Since 2000, a number of avian radar systems have been developed, the most representative of which are Merlin radar from the United States, Accipiter radar from Canada and Robin radar from the Netherlands.

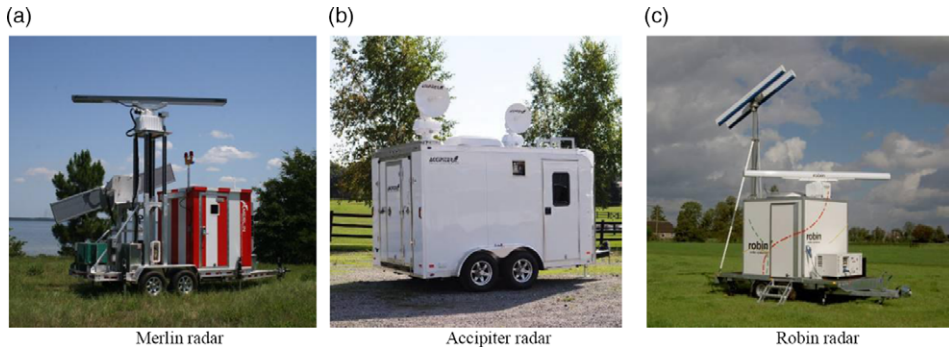


Figure 3. Typical avian radar system.

Table 1. Comparison of typical avian radars

System name	Technical characteristics and deployment modes
Merlin	Combination of horizontal and vertical scanning radar. The horizontal scanning radar is usually deployed near the center of the airport, which is responsible for low altitude early warning around the airport; the vertical scanning radar is usually deployed in the center of each runway to monitor the flight take-off and landing channel.
Accipiter	Combination of horizontal and vertical scanning radar with additional parabolic antenna. The deployment mode is similar to Merlin.
Robin	In the early stage, the horizontal and vertical scanning radars were combined. In a later stage, they were upgraded to FMCW. The latest generation employed phased arrays. The deployment mode of the early product is similar to Merlin. The latest product is usually deployed in the center of the airport.

Merlin radar was the first avian radar in the world. It is equipped with two navigation radars of different wavebands, one being S-band horizontal scanning, the other being X-band vertical scanning, both of which adopt standard T-shaped waveguide array antennae [12]. This dual-radar system has also become the most typical technical scheme for other avian radar systems. Figure 3(a) shows a typical Merlin radar system. For large birds or flocks, the detection range of the system can reach 4 to 6nmi (7 to 10km) horizontally and 15,000feet (4,500m) vertically. For small- and medium-sized birds, the detection range of the system can reach 2 to 3nmi (4 to 6km) horizontally and 7,500feet (2,300m) vertically.

The Accipiter radar has also developed a similar avian radar system with horizontal and vertical scanning radars, which has been applied in some airports in the United States and Canada. In order to obtain 3D target information, some Accipiter radars replace the waveguide slot antenna with a parabolic antenna, and integrate the data processor, control system, GPS, mobile power supply and other auxiliary operation systems on a mobile trailer, as shown in Fig. 3(b). The X-band parabolic antenna used in the Accipiter radar is usually placed on top of the trailer. However, compared with the waveguide slot antenna with larger beam width, the detection efficiency of this kind of antenna is lower. In order to make up for the deficiency of parabolic antenna detection efficiency, Accipiter has developed a variety of avian radar antennae, such as multi-beam parabolic antennae, dual axis scanning parabolic antennae and two parabolic antennae with different tilt angles, to improve the coverage of parabolic antennae in the vertical plane [13, 14].

The first generation Robin avian radar, which appeared after 2010, continues the dual-radar scheme of S-band horizontal scanning radar and X-band vertical scanning radar adopted by Merlin [15]. The antenna length of the S-band horizontal scanning radar is 3.6m. The rotating speed reaches 45rpm, which is the fastest among the existing avian radars. The S-band radar can detect large birds within a radial distance of 10km and a vertical height of 2km. The antenna of the X-band vertical scanning radar, which is 8feet (2.4m) long, can scan the airspace vertically and rotate at 24rpm, detecting small birds within 2.5km and large birds and flocks within 5km. The information of the overlapped coverage area of the two radars can be associated and integrated to obtain three-dimensional target information in this area. In addition, Robin has developed a mechanically scanned FMCW radar to replace the X-band vertical scanning radar with an average output power of only 400mW, as shown in Fig. 3(c). Different from pulse radar with only one antenna, FMCW avian radar adopts two separate antennae. One antenna continuously transmits electromagnetic waves of different frequencies, and the other continuously receives echo information. FMCW radar has more powerful bird detection ability with three working modes of scanning, tracking and staring, which is especially suitable for professional ornithological research. The latest generation of Robin radar shown in Fig. 1(a) adopts technologies of X-band FMCW and phased arrays, with an average output power of 20W. It can obtain the three-dimensional information of the target in the whole airspace within 15km, and the target trajectory update rate is 1s. At the same time, it has the ability to distinguish between flying birds and UA based on the micro-Doppler (m-D) feature.

Table 1 compares the technical characteristics and deployment modes of three typical avian radars. In the remainder of this section, the latest research results of UA detection and classification with radar will be introduced, including the traditional target feature-based classification method and the latest deep learning method.

### **3.0 Radar target detection and recognition method**

#### **3.1 M-D feature extraction**

The micro-motion of moving target components will introduce modulation sidebands near the radar Doppler signal generated by the motion of the target body, resulting in Doppler spectrum broadening, which is called m-D signal [16]. At present, as the most commonly used radar fine signal feature in automatic target classification, the m-D feature has been widely used in automatic target classification, such as ground moving target classification, ship detection, human gait recognition and human activity classification, and has become a very popular research direction in UA radar detection application in recent years [17].

The internal motion of the target, including blade rotation of the rotor UA or helicopter, the turbine propulsion of a jet, the wing flapping of a bird and so on, can be statistically described by m-D features [18]. The m-D feature could be generated by the flapping of bird wings and the rotation of a UA rotor, which is the main technical means of radar target recognition [19]. Figure 4 shows the simulation results of m-D characteristics of bird and rotor UA at X-band. It can be seen that the m-D of bird targets is generally concentrated in the low-frequency region with long period, which is closely related to the flapping motion characteristics of different types of bird wings, while the m-D characteristics of UA targets show obvious periodic characteristics, significantly different from birds. The echo of a rotor UA is the superposition of Doppler signals of the main body and rotor components, therefore, the m-D characteristics of UA with different types and numbers of rotors are also different.

The modeling of m-D characteristics and fine characteristic cognition of rotor UA are the pre-conditions for subsequent detection and target classification. The echo signal of the rotor UA is the superposition of the Doppler signal of the main body of the UA and the m-D signal of the rotor components, as shown in Fig. 5 [20].

Firstly, the main body motion and micro-motion of UAV target with micro-motion components are modeled, analysed and parameterised, and then the corresponding relationship among Doppler frequency, rotation rate, number and size of blades is deduced [21]. For multi-rotor UAV, assuming that

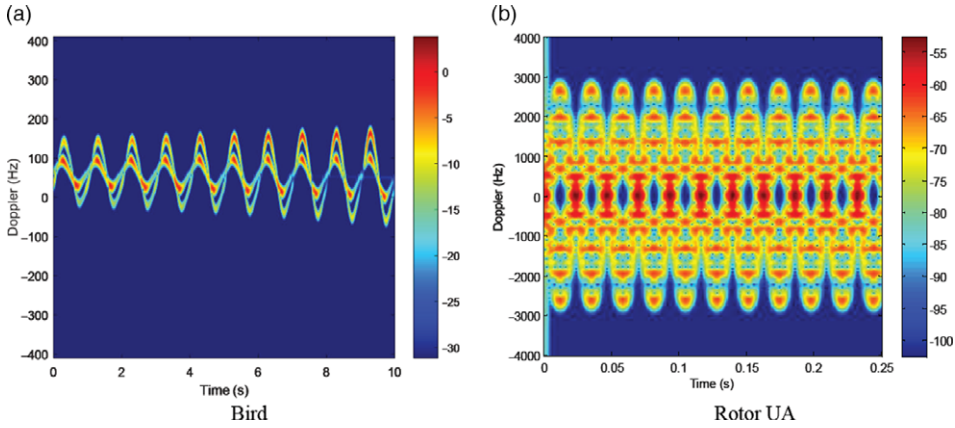


Figure 4. Simulation results of *m*-D characteristics of flying birds and rotor UA [19].

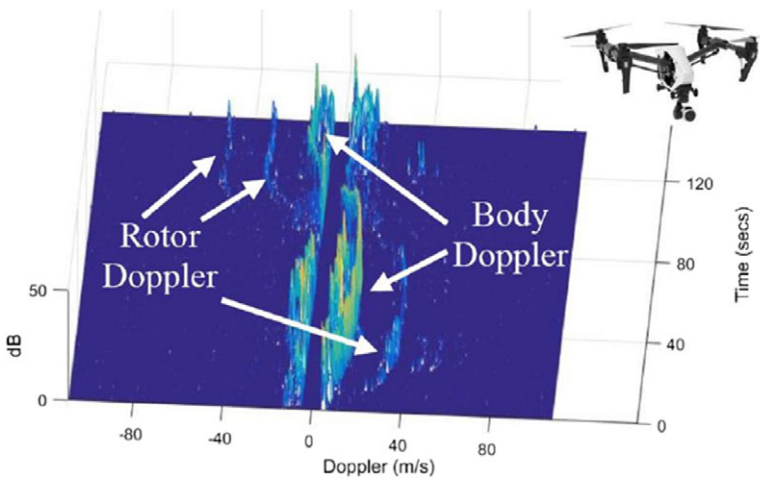


Figure 5. *m*-D characteristics of four-rotor UA [20].

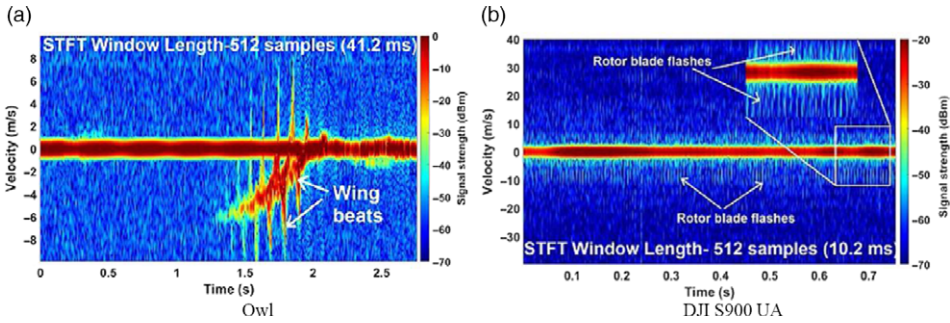
the RCS of all rotor blades is the same and the RCS value is 1, based on the helicopter rotor model, the target echo model of multi-rotor UAV is as follows [22]

$$s_{\Sigma}(t) = \sum_{m=1}^M L \exp \left\{ -j \frac{4\pi}{\lambda} [R_{0_m} + z_{0_m} \sin \beta_m] \right\} \cdot \sum_{k=0}^{N-1} \text{sinc} \left\{ \frac{4\pi}{\lambda} \frac{L}{2} \cos \beta_m \cos(\Omega_m t + \varphi_{0_m} + k2\pi/N) \right\} \exp \{ -j\Phi_{k,m}(t) \} \quad (1)$$

$$\Phi_{k,m}(t) = \frac{4\pi}{\lambda} \frac{L}{2} \cos \beta_m \cos(\Omega_m t + \varphi_{0_m} + 2k\pi/N), (k = 0, 1, 2, \dots, N - 1; m = 1, 2, \dots, M) \quad (2)$$

where,  $M$  is the number of rotors,  $N$  is the number of blades per rotor,  $L$  represents the length of blade,  $R_{0_m}$  is the distance from the radar to the centre of the  $m^{\text{th}}$  rotor,  $z_{0_m}$  represents the height of the  $m^{\text{th}}$  rotor blade,  $\beta_m$  is the pitch angle from the radar to the  $m^{\text{th}}$  rotor,  $\Omega_m$  is the rotation angle frequency of the  $m^{\text{th}}$  rotor with the unit of rad/s, and  $\varphi_{0_m}$  is the initial rotation angle of the  $m^{\text{th}}$  rotor.

The instantaneous Doppler frequency of the echo signal can be obtained from the time derivative of the phase function of the signal, and the equivalent instantaneous *m*-D frequency of the  $k^{\text{th}}$  blade of the



**Figure 6.** Measurement results of *m*-*D* characteristics of flying birds and UA targets [23].

$m^{\text{th}}$  rotor is obtained as

$$f_{m,k}(t) = -\frac{L\Omega_m}{\lambda} \cos \beta_m \sin(\Omega_m t + \varphi_{0m} + 2k\pi/N) \quad (3)$$

It can be seen from the above formula that the *m*-*D* characteristic of a multi-rotor UA is composed of  $M \times N$  sinusoidal curves and is affected by carrier frequency, number of rotors, rotor speed, number of blades, blade length, initial phase and pitch angle. Among them, the carrier frequency, blade length and pitch angle are only related to the amplitude of the *m*-*D* frequency, while the number of rotors, rotor speed, blade number and initial phase will affect the amplitude and phase of the *m*-*D* characteristic curve.

Due to the influence of environmental noise, it is difficult to measure the *m*-*D* characteristics in the outfield, which requires the radar to have strong accumulation ability of target Doppler echo information. Figure 6 compares the field measurement and data processing results of the *m*-*D* characteristics of an owl and a DJI S900 UA. Since the wing width of an owl is large, its flapping frequency is significantly lower than the rotor speed of the UA, resulting in the *m*-*D* intensity of UAV is significantly weaker than that of owl, while the *m*-*D* period is significantly faster than that of owl [23]. Therefore, the obvious difference of *m*-*D* characteristics between them lays a theoretical foundation for the establishment of a classification and recognition algorithm for birds and UA, which has good discrimination and high reliability without prior information. At present, the main difficulties of using *m*-*D* features in avian radar to recognise and classify such LSS targets include weak echo and low time-frequency resolution, and the unclear corresponding relationship between *m*-*D* features and the micro motion of target components.

In Refs [24, 25], the *m*-*D* feature was used to classify UA, and the short time Fourier transform (STFT) was used to extract *m*-*D* features, where the method of extracting key features such as speed, tip speed, rotor diameter and rotor number from radar signals was studied, so as to classify different rotor UA. Following a similar method, Molchanov et al. [26] used STFT to generate *m*-*D* features, which extracted feature pairs from the correlation matrix of *m*-*D* features, and trained three classifiers, including a linear and a nonlinear support vector machine (SVM) classifier, and a naive Bayes classifier (NBC), and realised the classification of 10 types of rotorcraft and birds. Chinese Academy of Sciences [27] proposed a *m*-*D* feature extraction algorithm based on instantaneous frequency estimation of a Gabor transform and Fast Fourier transform (FFT) to estimate the number of rotors, speed and blade length of UA.

DeWit et al. [28] followed a signal processing flow similar to Ref. [26] before applying singular value decomposition (SVD) to spectra. In order to achieve fast classification, three main features of target speed, spectrum periodicity and spectrum width were extracted. Similarly, in Ref. [29], the author used three common signal characterisation methods to generate *m*-*D* signals, namely STFT, cepstrum and cadence velocity diagram (CVD), and then used the combination of SVD feature extraction and an SVM classifier to classify the measured data of fixed wing, rotor UA and flying birds. The Chinese Academy of Sciences [30] proposed a method of estimating UA rotor rotation frequency based on autocorrelation



cepstrum joint analysis, which can estimate the UAV rotor rotation frequency more effectively through weighted equalisation.

In order to use the phase spectrum in the m-D feature extraction process, Ren et al. [31] proposed a robust signal representation method, namely two-dimensional regularised complex logarithm Fourier transform and object-oriented dimension reduction technology, where the subspace reliability analysis was designed for a binary UAV classification problem to distinguish UA and bird targets. Another m-D label extraction algorithm was proposed in Ref. [32], where the author used empirical mode decomposition (EMD) to classify UA automatically. On the basis of Ref. [32], Ma et al. [33] studied the feasibility of six entropies of a group of intrinsic mode functions extracted from EMD for UA classification, and proposed to fuse the extracted features into three kinds of entropies, and obtained the features through signal down sampling and normalisation as the input of a nonlinear SVM classifier. In Ref. [34], the problem of UA wing type and UA positioning was studied. They treated the positioning as a classification problem and expanded the number of different categories according to a group of positions of each UA. The method combined EMD and STFT to generate m-D features, which were analysed by principal component analysis (PCA), and the problem of UA target classification and location based on nearest neighbor, random forest, NBC and SVM was studied.

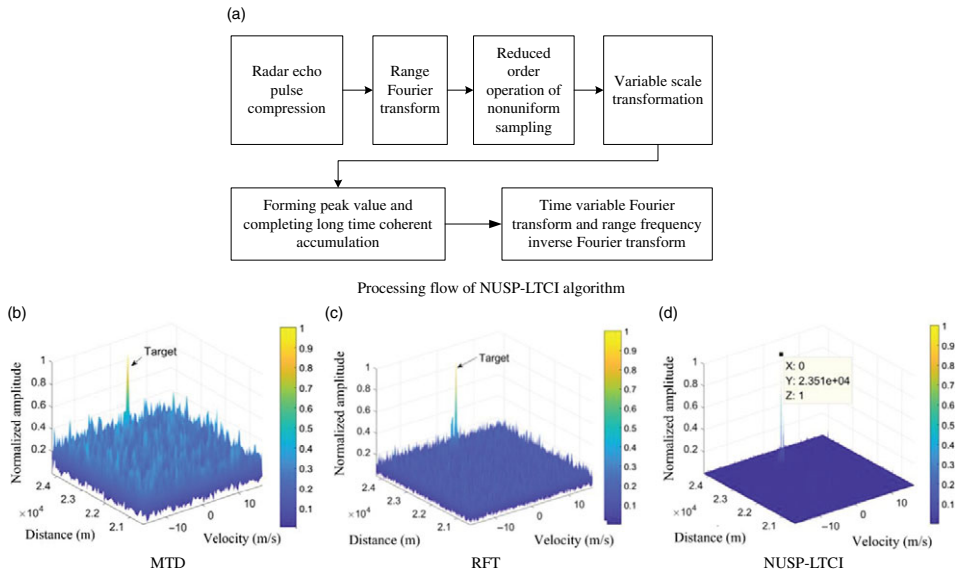
In addition to the typical radar with single antenna, some studies consider multi-antenna, multi-station radar and passive radar. In Ref. [35], the author proposed an NBC and discriminated analysis classifier based on Doppler and bandwidth centeed features of m-D signal, where the experiment considered the actual measurement of rotorcraft, including the loading and unloading of potential payload. In a similar study, Hoffman et al. [36] proposed a UA detection and tracking method with multi-station radar, which combines the m-D features with constant false alarm rate (CFAR) detector to improve the detection effect of UA, using an extended Kalman filter for tracking.

In Ref. [37], the author uses two radars of different bands to extract and fuse m-D features, so as to classify three kinds of rotorcraft. Tsinghua University [38] proposed a UA recognition method based on micro-motion feature fusion with multi-angle radar observation. Wuhan University [39] carried out the m-D effect experiment of multi-rotor UA using multi-illuminator-based passive radar, which confirmed the technical feasibility of micro-motion feature extraction of UA with passive radar.

In recent years, with the continuous development of phased array, holographic, MIMO and multi-static radar technology, the gain and velocity resolution of the target signal can be improved by prolonging the observation accumulation time, and the high-precision extraction and description of complex m-D features can be realised [40–42]. Traditional MTD methods usually use the combination of range migration compensation and Doppler migration compensation, but the algorithm is complex and cannot achieve long-term and fast coherent accumulation of target echoes [43]. The long-time coherent integration (LTICI) method determines the search range based on the multi-dimensional motion parameters of the target, and accumulates the observed values of the extracted target by selecting specific transformation parameters. This technique involves a large amount of calculation and is difficult to apply [44]. On the basis of LTICI, the target long-time coherent integration method based on non-uniform sampling order reduction and variable scale transformation (NUSP-LTICI) is adopted to reduce the high-order signal to first-order, which greatly reduces the amount of calculation and creates conditions for engineering application, whose algorithm flow is shown in Fig. 7(a) [45]. Figure 7(b-d) compares the low-altitude radar target coherent integration results of different algorithms, where the NUSP-LTICI algorithm has higher parameter estimation accuracy and stronger clutter suppression ability, compared with the traditional MTD algorithm and the classical Radon Fourier Transform (RFT) algorithm.

### 3.2 Feature extraction of target echo and motion

Traditional surveillance radar usually uses mechanical scanning antennae to detect and track multiple targets, whose design purpose is to find new targets by constantly searching the space. Since this kind of radar is always searching in all directions, the time of focusing on a single target is usually very short,



**Figure 7.** Processing flow of NUSP-LTCI algorithm and comparison of low-altitude radar target coherent accumulation results [45].

which makes it difficult to extract m-D features. Therefore, only the RCS and other echo signal features or motion features of the target can be used to classify detected targets [46].

Chen et al. proposed a target classification method based on motion model, using traditional surveillance radar data to classify UA and birds [47]. Based on Kalman filter tracking, a smoothing algorithm is proposed to enlarge the difference between the conversion frequency estimation of bird and UA target models. The method is based on the assumption that birds have higher manoeuvrability. Figure 8 shows the flow of the algorithm and an example of target classification in low-altitude clutter environment. Firstly, multiple motion models are established to track the target trajectory, and then the manoeuvrability characteristics of the target are extracted by estimating the conversion frequency of the target model, and finally the classification of bird and UA targets is realised. It is shown in Fig. 8(b) that the flight path of a UA target is relatively straight and stable, while the trajectory of a flying bird target is shorter and more variable. It should be noted that there are also some special cases that UA manoeuvres along the curve while bird keeps flying in a straight line, which inevitably results in some limitations of this method.

Messina and Pinelli [48] studied the problem of UA classification using two-dimensional surveillance radar data. The method was divided into two steps. Firstly, the UA and bird targets were classified, and then rotary wing and aircraft were classified. By creating a set of feature sets based on the target RCS, signal-to-noise ratio, tracking trajectory and speed information, SVM was selected as the classifier, and a subset of 50 features was established to achieve high classification accuracy.

Polarisation feature is another effective method to distinguish birds and UA. Figure 9(a) shows the periodic echo maps of a DJI UA under HH, HV and VV polarisation conditions, where the HH polarisation has the strongest fluctuation characteristics. Figure 9(b) shows the classification results for four types of targets (two types of UA, gliding bird and wing-flapping bird) with different polarisation parameters and sampling intervals [49].

### 3.3 UA target detection and classification based on deep learning

In recent years, deep learning methods have been successfully applied in audio and video data processing, but the application on radar data has just started. Typical deep learning schemes usually need to

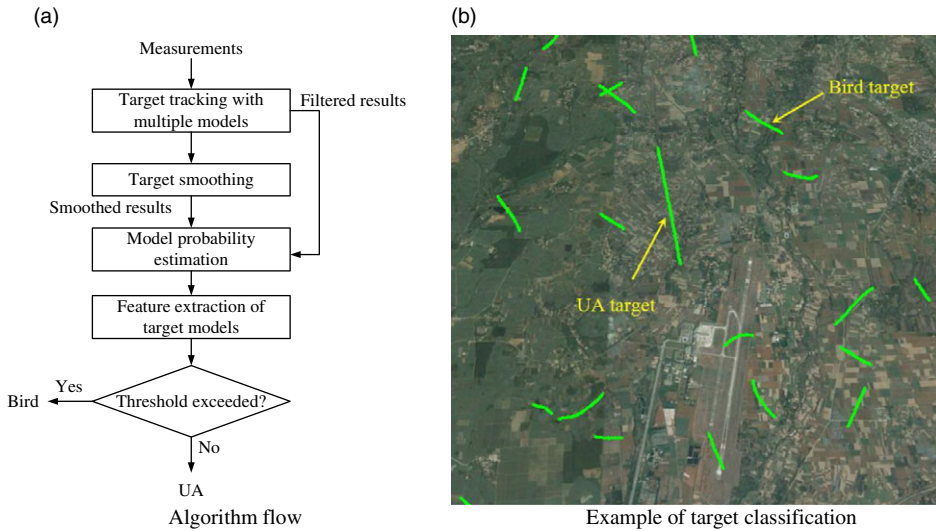


Figure 8. Classification of flying birds and UA based on flight trajectory characteristics [47].

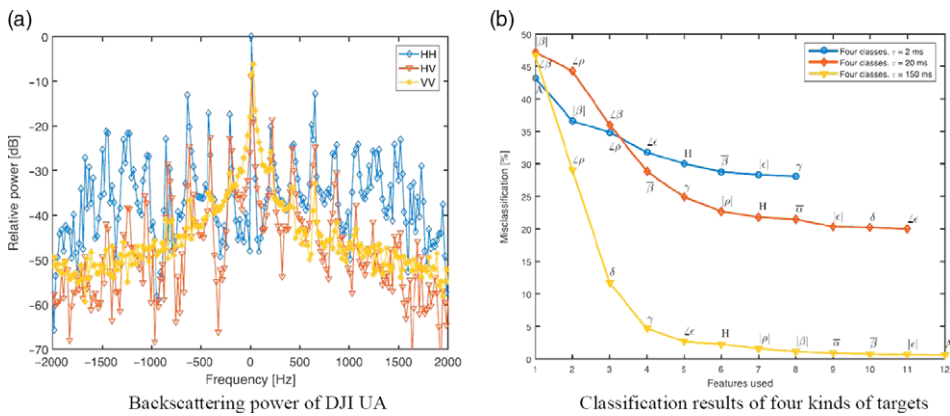


Figure 9. Examples of target scattering characteristics and classification results based on polarisation characteristics [49].

label a large amount of real data, which is easy to generate for audio and video data, but the labeling for radar data is still very limited. Despite the above difficulties, the work of combining radar data with deep learning has increased in recent years.

In Ref. [50], convolutional neural networks (CNNs) were first used to directly learn m-D characteristic spectra and classify UA. The spectra of UA were measured and treated by GoogleNet, where two UA were tested in indoor and outdoor environments. Mendis et al. [51] developed a UA classification algorithm based on deep learning for an S-band radar. They tested three different UA types of two rotary-wing and one aircraft, where the autocorrelation function (SCF) was used to identify the modulation components.

Wang et al. [52] proposed a range Doppler spectrum target detection algorithm based on CNNs, which was compared with the traditional CFAR detection method, achieving satisfactory results. The structure of the network was an eight-layer CNN trained with different range Doppler fixed windows under multiple signal to noise ratio (SNR) conditions. The detection problem was treated as the classification task of target and clutter, in which the fixed size window slid on the whole range Doppler matrix to check all units.

**Table 2.** Summary of radar based UA classification results in existing literatures

Target type	Type number (UAV + Bird)	Classification method	Reference
UA type vs. Bird	10 + 1	m-D + Nonlinear SVM	[26]
UA type vs. Bird		m-D (EMD) + SVM	[32]
UA type vs. Bird		m-D (EMD), Entropies of EMD + SVM	[33]
UA vs. Bird	1 + 1	m-D (SVD) + SVM	[29]
UA type	2		
UA vs. Bird	1 + 1	Fourier transform + Subspace reliability analysis	[31]
UA type + positioning	66	m-D (PCA) + Random forest	[34]
UA loading vs. unloading)	3	m-D + DAC	[35]
UA type	3	Dual radar m-D (PCA) + SVM	[36]
UA type	3	Dual radar m-D (PCA) + SVM	[37]
UA type vs. Bird	3 + 1	Radar polarization + Nearest neighbor	[49]
UA type vs. Bird	5 + 1	m-D (CVD) + CNN	[50]
UA type vs. Bird	2 + 1	SCF	[51]
UA type	2	IQ + MLP	[54]
UA type	3	Point cloud + MLP	[55]
UA vs. Bird	1 + 1	Motion model, speed and RCS + MLP	[56]
UA type vs. Bird	2 + 1	Motion model, speed and tracking + SVM	[48]

In Ref. [53], the range profile derived from the range Doppler matrix and the manually selected features were used as the input of DNNs to classify UA. Regev et al. [54] developed a multi-layer perceptron (MLP) neural network classifier and parameter estimator, which can determine the number of propellers and blades of UA and realise the diversification of micro-motion analysis of UA. This method was the first attempt to input the received complex signal directly into the learning network. The network architecture consisted of five independent branches, which received in quadrature (IQ), time, frequency and absolute data. There were two unique MLP classifiers, which first analysed the characteristics of propellers, then analysed the number of blades, and then fed them into the estimation algorithm. Simulation results showed that the classification accuracy was closely related to SNR.

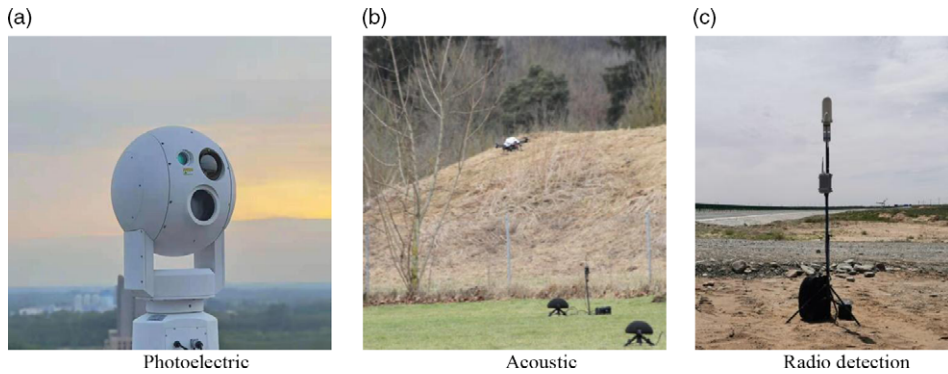
Habermann et al. [55] studied the classification of UA and helicopters using the point cloud features from radar. Based on the geometric differences between the point clouds, 44 features were extracted. Using artificial data training neural network to solve two classification problems, one is the classification of seven types of helicopters, the other is the classification of three types of rotor UA.

Mohajerin et al. [56] proposed a binary classification method, which used the data captured by surveillance radar to distinguish the trajectory of UA and birds. Twenty features based on motion, velocity and RCS are used. The feature manually extracted is combined with an MLP classifier, achieving high classification accuracy.

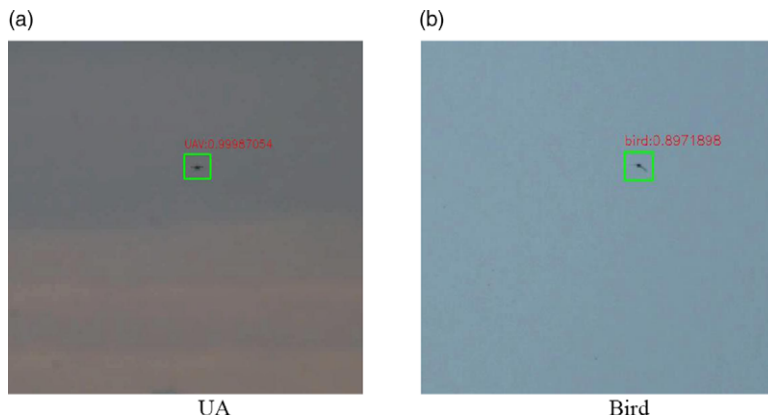
Most UA detection and classification methods based on m-D features are listed in Table 2, where the recognition rate of all methods is more than 90%, and some of them are close to or reach 100%. Unfortunately, because only part of the literature used the same data set, it is difficult to compare the performance of all methods directly.

#### 4.0 Other auxiliary detection technologies

At present, some advanced low-altitude surveillance radars have the ability of classification and recognition of UA, birds and other targets. The development of m-D feature extraction, deep learning and other



**Figure 10.** Multi-sensor detection technology of UA.



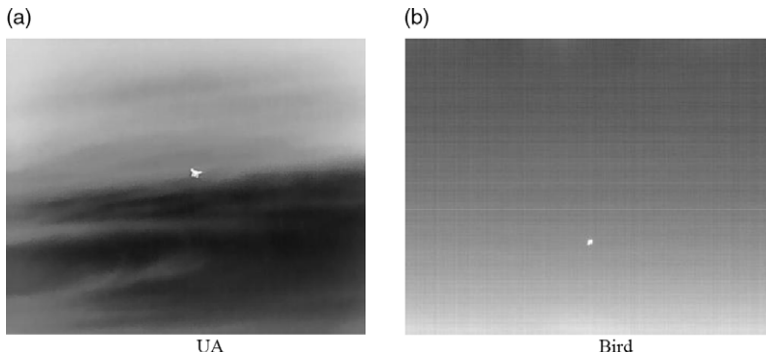
**Figure 11.** Visible image.

technologies further promotes progress in this field. At the same time, much experience shows that in engineering applications, the UA detection technology with radar as the main body, supplemented by photoelectric, acoustic, radio detection and other multi-source sensor fusion has become the mainstream, as shown in Fig. 10. In this section, the technical characteristics of various auxiliary sensors and the data processing methods based on deep learning are discussed.

#### 4.1 Visible light

Compared with radar, the target image information obtained by visible light detection equipment is more abundant, providing advantages in target recognition. However, a single visible light camera usually does not have target ranging capability, so it is necessary to obtain the target distance through multi-station cross positioning or adding laser ranging equipment. Optical cameras are also affected by environmental factors, and the detection distance is much lower than that of radar, which makes it difficult to use alone in general. Therefore, visible light cameras are usually only used as supplementary confirmation equipment for radar. Figure 10(a) shows a typical photoelectric system, which can obtain visible and infrared images of the target at the same time. Figure 11 shows an example of the images of UA and bird targets captured by a standard visible-light camera, with a detection range of about 2km.

With the development of neural network and deep learning algorithms, optical images have become one of the most valuable information sources in UA detection. Since the deep learning method was successfully applied to the image classification of the ImageNet dataset in 2012, the research around deep learning has been developing [57]. Most UA target recognition research using DNNs adopts the



**Figure 12.** *Infrared image.*

general target detection structure, and takes the powerful DNNs as the classification model. Most UA target recognition research using DNNs adopts the general target detection structure, and uses powerful DNNs as the classification model. In such research, DNNs are usually trained in advance on general data sets such as ImageNet, and then optimised with UA data to improve the recognition performance by adjusting the parameters.

Saqib et al. [58] tested the Faster-RCNN model for UA detection. They used VGG-16 [59], ZF-net [60] and other models to carry out experiments in the detection scheme. It is shown that VGG-16 performs the best. In addition, they trained birds as a separate category, which effectively reduced the false alarm rate. It has been proposed in recent studies to add a preprocessing module before the formal detection process to improve the detection effect. In one study, a U-net was inserted in front of the detector, which was a network that calculated the motion of continuous frames and generated a frame, which can contain the UA with a certain probability [61]. In other work, a super-resolution processing module was added in front of the detector to improve the detection ability of UA targets with fewer pixels in the image due to long detection distance [62].

Aker et al. [63] used a YOLO (You Only Look Once) [64] detector to detect UA targets quickly and accurately. In addition, they extracted UA targets from some public images and added them to the natural images with various complex backgrounds, and constructed a new artificial data-set to train UA deep learning models in different scales and backgrounds, so as to solve the scarcity of UA marked public data.

Rosantev et al. [65] established a detection framework based on spatiotemporal cube video data at different scales. This method used reinforcement trees and CNNs for classification. The experimental results showed that time domain information played an important role in small target detection such as UA, and CNNs performed better in accuracy recovery.

Gökçe et al. [66] proposed a UA detection method with an optical camera based on gradient histogram feature and cascade classification. The detection part evaluated the increasingly complex features at different stages through the cascade classification method, and it was regarded as detecting the target if it successfully passed through all levels of classifiers. In addition, the target distance was estimated by support vector regression.

## 4.2 Infrared

Compared with visible light cameras, the main advantage of infrared cameras is that they are not affected by light or weather conditions and can still operate normally even in complete darkness. Generally, the resolution of infrared camera imaging is low, and the cost is high. Therefore, it is only used in military applications, but with the progress of technology, its cost is steadily declining. Figure 12 shows image data of a UA and a bird target taken by a typical infrared camera at a distance of 2km.

In UA detection systems for security of prisons and other facilities, infrared cameras are generally placed on top of buildings or monitoring towers. In most multi-sensor systems, infrared cameras are

usually combined with visible light as auxiliary equipment for radar, but there are some public reports of using infrared sensors alone to detect, track or classify UA. Like visible light cameras, infrared cameras cannot measure the distance to the target, so it cannot locate the target. Anthony-Thomas et al. [67] proposed a method of UA target location by using the images from multiple infrared cameras.

At present, many UA detection systems integrated with infrared sensors have realised a UA target recognition function based on infrared images. The target recognition algorithm usually draws lessons from the existing deep learning model, but there is little literature on the topic. The rest of this section will review some published results of the application of infrared vision in other target detection, tracking and classification tasks, in order to provide reference for the application scenarios of UA target recognition.

Liu et al. [68] designed and evaluated different pedestrian target detection solutions based on Faster-RCNN architecture and multispectral input data. They started from different branches of each input data, and established a VGG-16 basic network at each branch, exploring the feature fusion technologies at low, medium and high levels of the network.

Cao et al. [69] adjusted the general pedestrian detector to the multispectral domain and used the complementary data captured by visible light and infrared sensors, which not only improved the pedestrian detection performance, but also generated additional training samples without manual labeling. Pedestrian detection was realised by the proposed network, while unsupervised automatic labeling was based on a new iterative method to mark moving targets from calibrated visible and infrared data.

John et al. [70] used CNNs to classify candidate pedestrian targets. Firstly, the infrared image was segmented by fuzzy clustering, and the candidate pedestrians were located. The candidate pedestrians were screened according to the human body posture characteristics and the second central moment ellipse. Then the cropped image block around each candidate object was adjusted to a fixed size and input into eight-layer CNNs for binary classification. Finally, the training was carried out based on the sample data set.

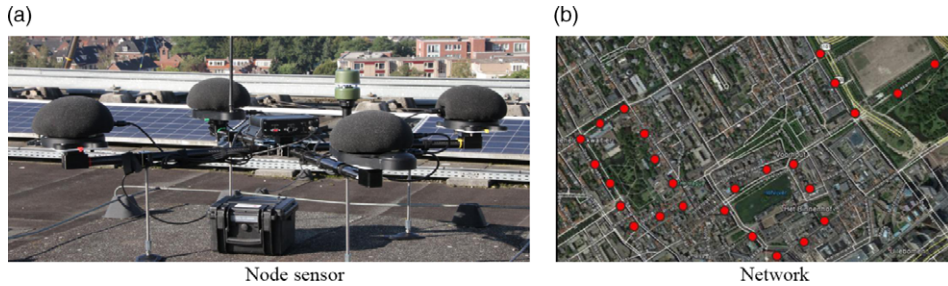
Ulrich et al. [71] realised human target detection based on the fusion of infrared image and m-D radar data based on dual stream neural network. Firstly, the Viola-Jones framework [72] was used to detect human targets in infrared images, and then the calculated sensor distance was used to perform associated processing between infrared and radar targets. Finally, the information obtained by each sensor was fused in the feature layer of a single joint classifier.

Liu et al. [73] redesigned the CNN training framework for infrared target tracking based on the relative entropy theory, and its performance was verified on two public infrared data sets. In this method, multiple weak trackers were constructed by using the features extracted from different convolution layers pre-trained on ImageNet, and the response diagram of each weak tracker was fused with the strong estimation of target position.

### 4.3 Acoustic

Acoustic technology uses the special “audio fingerprint” generated by UA flight to detect and identify. As a passive technology, acoustic technology has become an important technical means of multi-sensor fusion detection of UA, and can be a useful supplement to radar, photoelectric and other detection technologies, which does not interfere with other detectors.

Based on acoustic detection technology, relevant research institutions and companies in South Korea, France, Hungary, Germany, the Netherlands and the United States have developed corresponding UA detection systems and proposed relevant solutions. Figure 13 shows SkySentry, an acoustic system developed by MicroflownAvisa in Netherlands, which can accurately locate small UA in urban environments [74]. The core technology of SkySentry is an acoustic multi-mission sensor with an acoustic vector measurement function, which can track and locate low-altitude small UA by placing sensor array in a certain area. The capability is based upon either ground-based or vehicle-based node sensor that could be used stand-alone or networked. Depending upon the deployment of the sensors nodes, a perimeter can be sealed off against incoming intruder UA or a complete area can be covered.



**Figure 13.** *SkySentry system example.*

A general acoustic detection system consists of three main modules: detection, feature extraction and classification. The detection module captures the original audio signal of the target in the noise environment; the feature extraction module automatically extracts features from the original signal as the input of the classifier; the classification module assigns the extracted features to the corresponding classes. The ability of deep learning networks to extract unique features from raw data makes them highly suitable to acoustic detection. Lee et al. [75] produced one of the earliest research results of unsupervised learning of audio data using convolutional deep belief networks, where the features learned from the neural network correspond to the phonemes in the speech data one by one. Piczak [76] tested a simple CNN architecture with ambient audio data, whose accuracy is close to the most advanced classifier. Lane et al. [77] created a mobile application program that can perform very accurate speech binarisation and emotion recognition using deep learning. Recently, Wilkinson et al. [78] performed unsupervised separation of environmental noise sources, added artificial Gaussian noise to the pre-labeled signal, and used an automatic encoder for clustering. However, the background noise in environmental signals is usually non-Gaussian, which makes this method only suitable for specific data sets.

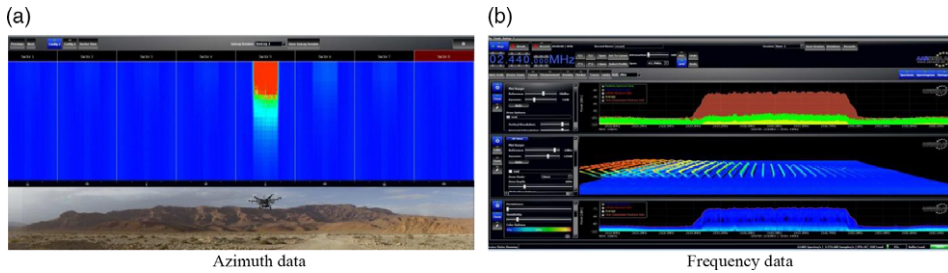
In UA detection, Park et al. [79] proposed a method of fusing radar and acoustic sensors to realize the detection, tracking and identification of rotor UA based on feed-forward neural networks. Liu et al. [80] used the Mel-frequency cepstral coefficients (MFCCs) and an SVM classifier to detect UAV. Kim et al. [81] introduced a real-time detection and monitoring system of UA based on microphones, which used nearest neighbor and graph learning algorithms to learn the characteristics of FFT spectra. Kim and Kim [82] expanded on their own work, and improved the classification accuracy of the system from 83% to 86% by using artificial neural networks. Jeon et al. [83] proposed a binary classification model, which used audio data to detect UA, used an F-score as performance measure, compared CNNs with Gaussian mixture model and recurrent neural network, and synthesised the original UA audio with different background noise. In addition, many research institutions used the improved MFCC acoustic feature extraction method to continuously improve the UA target recognition ability in different ways [84, 85].

#### **4.4 Radio detection**

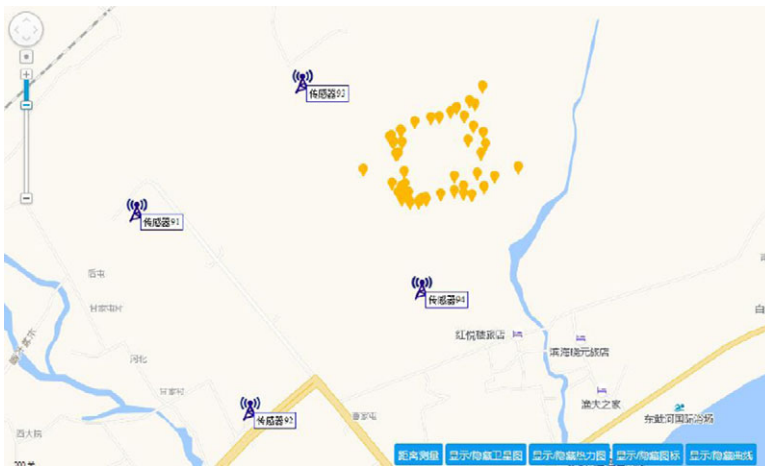
Radio detection uses radio frequency scanning technology to monitor, analyse and measure the data transmission signals, image transmission signals and satellite navigation signals of civil UAV in real time. Firstly, the spectrum characteristics of UA data and image transmission signals are extracted, and the UAV feature database is established to realise the discovery of UA; then, the UA positioning, tracking and recognition are realised by using the technology of Radio Direction Finding (RDF) [86] or the technology of time difference of arrival (TDOA) positioning [87].

In RDF, a 3D array antenna is usually used to receive the electromagnetic signal transmitted by UA in real time, and the received signal is analysed by spectrum analyser and monitoring software. At the same time, the spectrum analyser and monitoring software collect and analyse the frequency, bandwidth,





**Figure 14.** UA target data obtained by radio detection system.



**Figure 15.** UA target data obtained by TDOA system.

modulation mode, symbol rate and other data of the signal, and compare them with the established UA signal feature library to identify the signal features. Figure 14 shows the target azimuth and frequency data of UA obtained by a typical radio detection system.

The biggest advantage of radio detection systems is that they do not emit electromagnetic waves and belong to green passive technology. At present, a large number of companies and scientific research institutions have developed relevant systems using radio detection technology, and some of them have been widely used in the security of airports, prisons, stadiums, public buildings and major events. However, the biggest problem with radio monitoring technology is that it may not be able to detect “silent” UA that turn off data transmission and image transmission signals.

TDOA is a method of positioning by using time difference, which determines the location of a signal source by comparing the time difference of signals arriving at multiple monitoring stations. TDOA positioning systems have the characteristics of low cost and high accuracy, but usually need to place more than three stations, whose equipment includes antenna, receiver and time synchronisation module. Figure 2(c) shows typical TDOA station equipment and Figure 15 shows the UA target data obtained by a typical TDOA system.

Radio signal recognition is the core technology of radio detection system, which is essentially the same as the recognition technology for radar, visible light and infrared signals. It uses fixed rules or machine learning classifiers based on the extraction of key features of signals. In recent years, the application of deep learning theory has sprung up in the field of artificial intelligence. In 2016, O’Shea et al. took the lead in applying deep learning algorithm to radio signal recognition, which used the CNN framework to automatically recognise 11 kinds of modulation signals [88]. Compared with traditional machine learning algorithms, the recognition rate was greatly improved. Since then, there have been a

lot of research results of radio signal recognition based on deep learning. Karra et al. first distinguished analog modulation from digital modulation, and then identified the specific modulation type and order by using the cascade identification method of CNN networks [89]. West and O'Shea [90] further optimised the network structure by combining the CNN network with the signal timing characteristics. On the basis of the model proposed in Ref. [88], Arumugam et al. [91] used cyclostationary characteristic parameters as the input of CNN networks, which improved the recognition results at the expense of greater computational complexity. Schmidt et al. [92] identified three kinds of radio communication standard signals by using FFT results as the input to CNNs.

## 5.0 Multi-sensor system performance

The purpose of data fusion based on multiple sensors is to combine the data from different modes to produce an effect that cannot be achieved by a single sensor. In recent years, the application of artificial intelligence and deep neural networks in multi-source data processing has been widely pursued, providing a possible solution for complex multi-modal learning problems [93].

### 5.1 Performance analysis

The performance and working characteristics of various types of sensors are compared in Table 3. This section further summarises and analyses their advantages and disadvantages, potential limitations and development prospects.

Radar can effectively detect potential UA and track multiple targets in the whole airspace. Radar systems that can detect small UA usually work in X-band, but the cost of such systems is usually high, accompanied by a certain false alarm rate. Except for passive radars, most radars may cause frequency interference, whose deployment needs to be approved by the local radio regulator. Radar based UA target detection and tracking is mainly realised by classical radar signal processing techniques such as CFAR, Doppler processing and hypothesis testing. At present, the most commonly used radar signal feature in UA classification is the m-D feature, which is mainly affected by the number of rotor blades, incidence angle range, pulse repetition frequency and radar irradiation time [26]. In fact, most of the current work is carried out under the ideal condition of close range [26, 32, 33], and rarely on the original radar data, because the echo data of small UA targets is difficult to obtain under strong noise conditions in the field. Otherwise, the traditional surveillance radar uses mechanical scanning antennae, which have a short residence time for a single target and cannot produce m-D features, when it can only rely on other information sources such as the trajectory characteristics [47], RCS [48] and polarisation characteristics [49] of the target. However, compared with m-D methods, this kind of research is relatively rare and the scope of application is limited. The trajectory classification method based on deep learning has been successfully applied in general motion model applications, which may be transplanted to UAV trajectory classification and become a potential research direction [94].

Optical technology can image UA targets in the field of vision in visible or infrared light. There are visible light image public data sets for UA detection and classification [95, 96]. The common deep learning architectures for visible light image processing include Faster-RCNN [58], VGG-16 [59] and YOLO [64], whose performance is far superior to the traditional computer vision methods of manual calibration features [66]. Note that it is not enough to transplant the existing target detection algorithms directly to UA image data. It is necessary to improve these methods by studying the characteristics of UA target detection, so as to avoid false alarm while detecting small targets. Although existing UA detection systems are generally equipped with infrared cameras, such commercial cameras are unsuitable to detect small UA due to their low resolution, and the high price of high-resolution infrared equipment may be the direct reason for the small number of such research papers published. In fact, the processing mechanism of infrared image is similar to that of visible images. With the decreasing cost of high-resolution infrared equipment, combined with deep learning algorithms, the research prospects of UA target detection and classification using infrared data are promising.

**Table 3.** Comparison of main performances of various UA detection sensors

Sensor type	Range	Classification ability	Advantages	Disadvantages
Radar	3~5km	Average	It is not affected by UA type and has a long effective detection distance. Target classification can be realized based on m-D features.	There is a certain proportion of false alarm and a certain range of low altitude detection blind area; active radar may cause interference to the airport air traffic control equipment, so frequency licensing is required before use.
Visible light	1~2km	Good	It can capture all kinds of UA, realize real-time and visual monitoring, and has strong target classification ability.	Limited by light conditions, the detection distance is relatively short. It has no ranging function.
Infrared	1~2km	Good	It is not limited by lighting conditions, can achieve real-time, visual monitoring, and has strong target classification ability.	The cost is high and the field of view is small.
Acoustic	150m	Poor	Low cost, suitable for various types of UA, can achieve all day early warning detection.	The noise level of UA is very low. In noisy environments, the applicability of this technology is poor and the detection distance is short.
Radio detection	1~2km	Poor	It is suitable for all types of UA, and can effectively detect the UA operator.	“Silent” UA that do not transmit radio signals cannot be detected.

Acoustic sensors are light and easy to install. They can be used in mountainous areas or highly urbanised areas where other devices are not suitable. At the same time, the passive acoustic sensor does not interfere with the surrounding communication, and the power consumption is very low. However, in noisy environments, the detection range of a single acoustic sensor is less than 150m [83]. The research on multichannel audio shows that the detection performance can be significantly improved by using the most advanced beam-forming algorithms and fusing with other sensors such as optical cameras. Therefore, the establishment of a UA audio signal public database is of great significance to promote the research and development in this field. In addition, marking such data sets is prone to human errors, which mislead the machine learning algorithm. Therefore, in the field of environmental audio monitoring, unsupervised deep learning algorithms have attracted great attention in recent years, and its application in UA target detection and recognition worth pursuing [75].

Radio detection technology is a traditional tool of radio spectrum management, and has become an important technical means of UA detection and identification. Compared with radar, photoelectric, acoustic and other technical means, the biggest advantage of radio detection technology is that it can detect the position of the UA operator, and the biggest problem is the technology may fail for the “silent” UA without communication links. In addition, the complex electromagnetic environment could cause

interference to the radio detection technology, affecting its detection range and performance. Radio signal recognition technology based on deep learning can adapt to the complex electromagnetic environment and ensure a certain recognition rate under conditions of low signal to noise ratio (SNR), which represents the development direction of current radio detection technology [97]. At the same time, we notice that TDOA technology has been successfully applied in a UA detection and positioning system based on various sensors such as radio detection, MIMO radar and microphone arrays.

### 5.2 *Some existing cases*

In recent years, for UA target detection, recognition and tracking systems based on multi-sensor fusion, scientific research institutions and related enterprises have put forward a variety of construction schemes and data fusion methods.

Shanghai Jiao Tong University [98] proposed a low-altitude target monitoring and tracking method using radar and visible light. This method uses radar as the main tracker and a visible light camera as the detector to track the target through an interactive multiple model, and realises the monitoring and tracking simulation of rotor UA and other low-altitude targets.

Park et al. [79] proposed a small UA detection system combining radar and acoustic sensors. The system uses low-cost radar to scan the interesting area and an acoustic sensor array to confirm the target. The system uses a pre-trained deep learning framework composed of three MLP classifiers to vote whether the UA exists or not. The cost of the system is low, but the detection distance is only about 50m.

Liu et al. [80] proposed a solution of joint detection using a camera array and acoustic sensors. The system is composed of 30 cameras, three microphones, 8 workstation nodes and several network devices. The UA image and audio data are recognised by an SVM classifier. For the UA target with a height of 100m and horizontal distance of 200m, high detection accuracy is achieved.

Jovanoska et al. [99] built a UA detection, tracking and positioning system based on bearings-only multi-sensor data fusion, and adopted a centralised data fusion method based on multiple hypothesis tracker. The system first distinguishes the target and false alarm from different sensors, identifies the targets and estimates the contribution of each single sensor to the detection. It then uses a fusion algorithm to predict the target position and modify the target location by fusing the measurement results of different sensor systems.

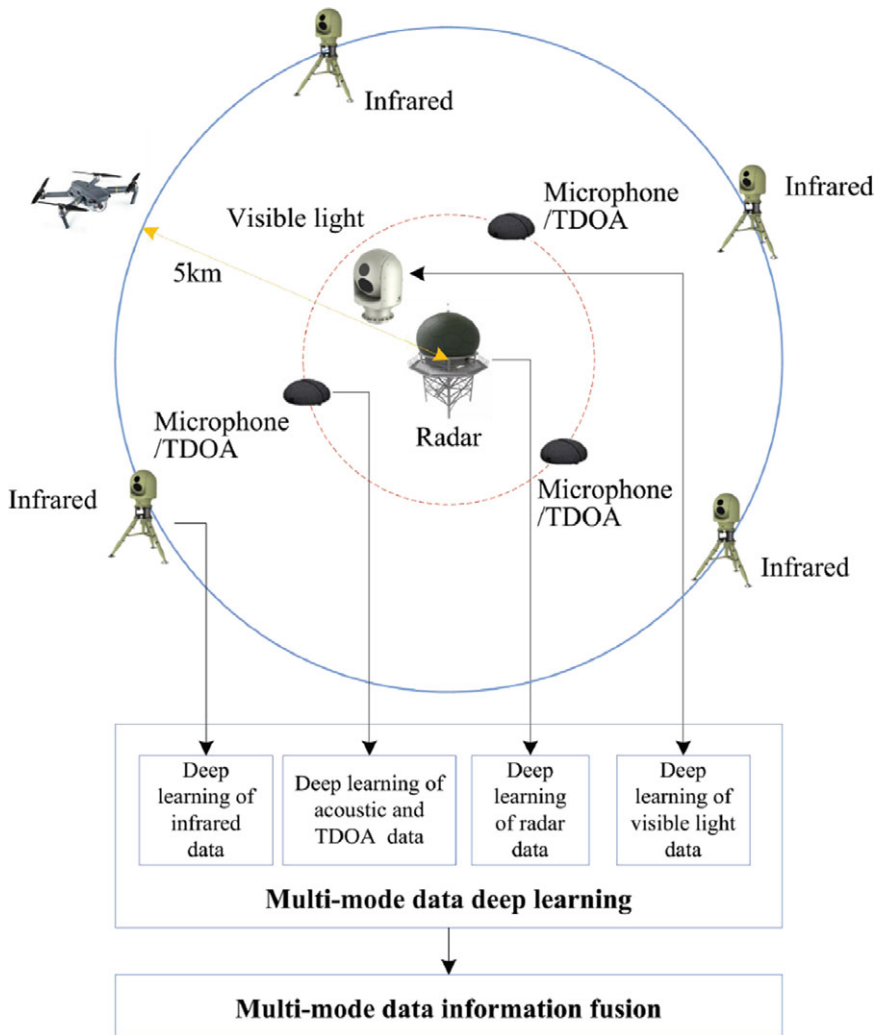
Hengy et al. [100] proposed a UA detection, positioning and classification system based on optical, acoustic and radar multi-sensor fusion. The system uses a multi-signal classification method to realise acoustic array positioning, uses radar technology to improve the detection rate and reduce the false alarm rate, and combines an infrared image with a visible image to realise easier and faster detection of UA targets under clutter, smoke or complex background conditions based on optical sensors.

Laurenzis et al. [101] studied the detection and tracking of multiple UA in an urban environment. The system consists of a distributed sensor network with static and mobile nodes, including optical, acoustic, lidar, microwave radar and other sensors. The experimental results show that the average error between the UA positioning results of the system and the real data is about 6m.

### 5.3 *Suggestions on system construction*

Due to the advantages and disadvantages of various kinds of sensors, it is almost impossible to provide the required situation awareness by using a single sensor for UA detection and identification. However, if all kinds of technologies can be mixed and complemented, it is possible to find an effective solution. Figure 16 recommends a construction scheme of UA detection system, which realises all-round situation awareness in a robust way by fusing the information of four types of sensors.

A long-range radar is placed in the centre of the sensor coverage area. Radar is a reliable early warning method, among which the holographic radar, which can obtain all-round target fine features with high data rate, is the best, so as to provide data support for extracting target m-D features. In order to



*Figure 16. Proposal of UA surveillance system.*

further correct the range and azimuth information of the detected target, multiple panoramic infrared cameras need to form complementary and cross validation at the edge of the coverage area to reduce the false alarm rate. Due to the sensitivity of infrared cameras in bad weather conditions, a visible light camera with rotation and zoom functions is placed near the radar to further improve the recognition ability. In addition, in order to avoid ground object occlusion, microphone arrays or TDOA stations can be distributed around the protected area to provide another complementary and alternative solution. In addition, the deployment of the TDOA array will help to complement other sensors and help trace the UA operator. Based on the deep learning method, the deep learning network for UA detection and recognition can utilise the recorded data of each sensor. Finally, the single source warning signals and features generated by each single source deep learning network are integrated with the deep learning network of multi-sensor fusion to improve the target recognition ability of the invading UA. The heterogeneity of multi-sensor data requires deep learning methods to construct joint representation data by using its internal relationship to effectively deal with the diversity of data representation. The simultaneous interpreting of different signals from different sensors can provide important knowledge aggregation

compared with single sensors, which is the advantage of this scheme using multi-sensor fusion deep learning.

## 6.0 Conclusions and prospects

In this paper, the research work of UA detection and recognition methods based on radar, photoelectric, acoustic, radio detection sensors and multi-sensor information fusion algorithm is reviewed. The m-D feature reflects the micro-motion of the target body and components, which shows a promising ability of detection and classification. Photoelectric detection is still an important means of UA target detection and recognition. In the UA detection mission, the deployment cost of microphone arrays is low, which will help to form a complementary and robust system framework when combined with other sensors. The radio detection and positioning technology represented by TDOA can lock the position of the UA operator, which plays an irreplaceable role in multi-sensor UA detection. In addition, the application of deep learning may lead to major breakthroughs in this field in the next few years.

With the rapid development of multi-source sensor performance and signal processing and fusion technology, the development trend of UA detection and recognition is mainly reflected in the following aspects:

- (a) The progress of radar hardware technology provides a solid foundation for the development of radar target detection refinement and intelligent information processing technology. At the same time, new radar systems such as digital array systems, external emitter radar, holographic radar, MIMO radar, software radar and intelligent radar expand the available signal dimension and provide a flexible software framework for the refinement of radar target detection and the integrated processing of detection and recognition. In addition, in the complex environment, the comprehensive use of the information of photoelectric, acoustic, radio and other multi-source sensors to make up for the limitations of radar can further improve the recognition efficiency and accuracy.
- (b) The fusion of signal and data features is an effective way to improve the accuracy of target classification. The long-time observation mode of staring radar make it possible to extract high-resolution m-D features, and can obtain the fine signal features of UA and birds. The m-D features of UA and bird targets are obviously different, which lays a foundation for classification and recognition. In addition, the motion trajectories of UA and birds have a certain degree of discrimination, and the mature target tracking algorithm approximates the real motion state of the target by establishing a variety of motion models. Therefore, the fusion of signal and data features of UA and birds can expand the difference in feature space and improve the classification probability.
- (c) Deep learning networks provide a new means for intelligent recognition of UA and bird targets. Using the idea of intelligent learning such as deep learning, through the construction of multilayer convolution neural network, the complex structure in high-dimensional data could be found, whose strong feature expression ability and high classification and recognition accuracy have been verified in the fields of image recognition and speech recognition. When classifying UA and bird signals by using only the m-D features obtained from a time-frequency map, the feature space is limited and the accuracy is low. Therefore, the idea of multi-feature and multi-channel CNN fusion processing can be adopted to input the echo sequence diagram, m-D time-frequency diagram, range Doppler diagram, transform domain diagram and motion trajectory diagram of the target into multi-channel CNNs, so as to make better use of CNNs for feature learning and target recognition.

**Acknowledgements.** The research work is jointly funded by the National Natural Science Foundation of China (NSFC) and Civil Aviation Administration of China (CAAC) (U2133216, U1933135), and also supported by BeiHang Zhuo Bai Program (ZG216S2182).

**Supplementary material.** To view supplementary material for this article, please visit <https://doi.org/10.1017/aer.2022.50>

## References

- [1] Chen, W., Wan, X. and Li, J. Detection technology of non-cooperative UAV targets in airport clearance area, *J Civil Aviat*, 2018, **2**, (5), pp 54–57.
- [2] Jin, W. and Shang, Y. The safety supervision of unmanned aircraft systems in China, *Sci Technol Rev*, 2019, **37**, (14), pp 66–77.
- [3] Luo, J. and Yang, Y. An overview of target detection methods based on data fusion, *Control Decis*, 2020, **35**, (1), pp 1–15.
- [4] Li, X., Zha, Y., Zhang, T., et al. Survey of visual object tracking algorithms based on deep learning, *J Image Graph*, 2019, **24**, (12), pp 2057–2080.
- [5] Chen, W., Liu, J., Chen, X., et al. Non-cooperative UAV target recognition in low-altitude airspace based on motion model, *J Beijing Univ Aeronaut Astronaut*, 2019, **45**, (4), pp 687–694.
- [6] Chen, W. and Li, J. Review on development and applications of avian radar technology, *Modern Radar*, 2017, **39**, (2), pp 7–17.
- [7] Li, S., Sun, R., Sun, J., et al. A study on the architecture of radar ATR based on deep learning, *Modern Radar*, 2019, **41**, (11), pp 57–61+84.
- [8] Luo, Q. Small moving target detection using sparse clutter modeling, *Modern Radar*, 2016, **38**, (2), pp 43–46+83.
- [9] Liu, Y., Yi, J., Wan, X., et al. Time varying clutter suppression in CP OFDM based passive radar for slowly moving targets detection, *IEEE Sens J Early Access*. doi: [10.1109/JSEN.2020.2986717](https://doi.org/10.1109/JSEN.2020.2986717).
- [10] Yi, Y., Wan, X., Yi, J., et al. Polarization diversity technology research in passive radar based on subcarrier processing, *IEEE Sens J*, 2019, **19**, (5), pp 1710–1719. doi: [10.1109/JSEN.2018.2881226](https://doi.org/10.1109/JSEN.2018.2881226).
- [11] Dan, Y., Yi, J., Wan, X., et al. LTE-based passive radar for drone detection and its experimental results, *J Eng*, 2019, **2019**, (20), pp 6910–6913. doi: [10.1049/joe.2019.0583](https://doi.org/10.1049/joe.2019.0583).
- [12] Anderson, R. Avian Radar Systems [EB/OL]. Panama City, Florida: [s. n.], 2007 [2020-08-31]. <http://www.detect-inc.com/downloads>
- [13] Weber, P., Nohara, T.J. and Gauthreaux, S.A. Affordable, real-time, 3-D avian radar networks for centralized North American bird advisory systems, *Proceedings of the Bird Strike Committee-USA/Canada*, 2005.
- [14] AC No: 150/5220-25 FAA Advisory Circular on Airport Avian Radar Systems, 2010.
- [15] Hoffmann, F., Ritchie, M., Fioranelli, F., et al. Micro-Doppler based detection and tracking of UAVs with multistatic radar, 2016 *IEEE Radar Conference (RadarConf)*, 2–6 May 2016, Philadelphia, PA, USA. doi: [10.1109/RADAR.2016.7485236](https://doi.org/10.1109/RADAR.2016.7485236).
- [16] Zhang, Q., Hu, J., Luo, Y., et al. Research progresses in radar feature extraction, imaging, and recognition of target with micro-motions, *J Radars*, 2018, **7**, (5), pp 531–547.
- [17] Li, X., Huang, Y., Yin, K., et al. Micro-motion signature and recognition of battlefield targets in airborne radar, *Modern Radar*, 2017, **39**, (2), pp 56–60.
- [18] Ritchie, M., Fioranelli, F., Griffiths, H., et al. Micro-drone RCS analysis, 2015 *IEEE Radar Conference*. IEEE Press, Arlington, VA, USA, 2015, pp 452–456.
- [19] Chen, V.C. *The Micro-Doppler Effect in Radar*. Artech House, Norwood, 2011.
- [20] Jahangir, M. and Baker, C.J. Extended dwell Doppler characteristics of birds and micro-UAS at l-band, 2017 *18th International Radar Symposium (IRS)*, Prague, 2017, pp 1–10, doi: [10.23919/IRS.2017.8008144](https://doi.org/10.23919/IRS.2017.8008144).
- [21] Singh, A.K. and Kim, Y. Automatic measurement of blade length and rotation rate of drone using w-band micro-doppler radar, *IEEE Sens J*, 2018, **18**, (5), pp 1895–1902. doi: [10.1109/JSEN.2017.2785335](https://doi.org/10.1109/JSEN.2017.2785335).
- [22] Song, C., Zhou, L., Wu, Y. and Ding, C. An estimation method of rotation frequency of unmanned aerial vehicle based on auto-correlation and cepstrum, *J Electron Inform Technol*, 2019, **41**, (2), pp 255–261. doi: [10.11999/JEIT180399](https://doi.org/10.11999/JEIT180399).
- [23] Rahman, S. and Robertson, D.A. Radar micro-Doppler signatures of drones and birds at K-band and W-band, *Sci Rep*, 2018, **8**, (1), pp 1–11. doi: [10.1038/s41598-018-35880-9](https://doi.org/10.1038/s41598-018-35880-9).
- [24] De Wit, J.J.M., Harmanny, R.I.A. and Premel-Cabic, G. Micro-Doppler analysis of small UAVs, 2012 *9th European Radar Conference*, IEEE Press, Amsterdam, The Netherlands, 2012, pp 210–213.
- [25] Harmanny, R., De Wit, J. and Cabitc, G.P. Radar micro-Doppler feature extraction using the spectrogram and the cepstrogram, 2014 *11th European Radar Conference*. IEEE Press, Cincinnati, OH, USA, 2014, pp 165–168.
- [26] Molchanov, P., Egiazarian, K., Astola, J.T., et al. Classification of small UAVs and birds by micro-Doppler signatures, *Int J Microwave Wirel Technol*, 2013, **6**, (3–4), pp 435–444.
- [27] Ma, J., Dong, Y., Li, Y., et al. Multi-rotor UAV's micro-Doppler characteristic analysis and feature extraction, *J Univ Chin Acad Sci*, 2019, **36**, (2), pp 235–243.
- [28] De Wit, J.J.M., Harmanny, R.I.A. and Molchanov, P. Radar micro-Doppler feature extraction using the singular value decomposition, 2014 *International Radar Conference*. IEEE Press, Lille, France, 2014, pp 1–6.
- [29] Fuhrmann, L., Biallawons, O., Klare, J., et al. Micro-Doppler analysis and classification of UAVs at Ka band, 2017 *18th International Radar Symposium (IRS)*. IEEE Press, Prague, Czech Republic, 2017, pp 1–9.
- [30] Song, C., Zhou, L., Wu, Y., et al. An estimation method of rotation frequency of unmanned aerial vehicle based on auto-correlation and cepstrum, *J Electron Inform Technol*, 2019, **41**, (2), pp 255–261.
- [31] Ren, J. and Jiang, X. Regularized 2D complex-log spectral analysis and subspace reliability analysis of micro-Doppler signature for UAV detection, *Pattern Recognit*, 2017, **69**, pp 225–237.

- [32] Oh, B.S., Guo, X., Wan, F.Y., et al. Micro-Doppler mini-UAV classification using empirical-mode decomposition features, *IEEE Geosci Rem Sens Lett*, 2017, **15**, pp 227–231.
- [33] Ma, X., Oh, B.S., Sun, L., et al. EMD-based entropy features for micro-Doppler mini-UAV classification, 2018 *24th International Conference on Pattern Recognition (ICPR)*, IEEE Press, Beijing, 2018, pp 1295–1300.
- [34] Sun, Y., Fu, H., Abeywickrama, S., et al. Drone classification and localization using micro-doppler signature with low-frequency signal, 2018 *IEEE International Conference on Communication Systems (ICCS)*. IEEE Press, Chengdu, 2018, pp 413–417.
- [35] Fioranelli, F., Ritchie, M., Griffiths, H., et al. Classification of loaded/unloaded micro-drones using multistatic radar, *Electron Lett*, 2015, **51**, pp 1813–1815.
- [36] Hoffmann, F., Ritchie, M., Fioranelli, F., et al. Micro-Doppler based detection and tracking of UAVs with multistatic radar, 2016 *IEEE Radar Conference (RadarConf)*. IEEE Press, Philadelphia, PA, USA, 2016, pp 1–6.
- [37] Zhang, P., Yang, L., Chen, G., et al. Classification of drones based on micro-Doppler signatures with dual-band radar sensors, 2017 *Progress in Electromagnetics Research Symposium-Fall (PIERS-FALL)*. IEEE Press, Singapore, 2017, pp 638–643.
- [38] Zhang, P., Li, G., Huo, C., et al. Classification of drones based on micro-doppler radar signatures using dual radar sensors, *J Radars*, 2018, **7**, (5), pp 557–564.
- [39] Liu, Y., Yi, J., Wan, X., et al. Experimental research on micro-doppler effect of multi-rotor drone with digital television based passive radar, *J Radars*, 2018, **7**, (5), 585–592.
- [40] Singh, A.K. and Kim, Y. Automatic measurement of blade length and rotation rate of drone using w-band micro-doppler radar, *IEEE Sens J*, 2018, **18**, (5), pp 1895–1902. doi: [10.1109/JSEN.2017.2785335](https://doi.org/10.1109/JSEN.2017.2785335).
- [41] Song, C., Zhou, L., Wu, Y. and Ding, C. An estimation method of rotation frequency of unmanned aerial vehicle based on auto-correlation and cepstrum, *J Electron Inform Technol*, 2019, **41**, (2), pp 255–261.
- [42] Rahman, S. and Robertson, D.A. Radar micro-Doppler signatures of drones and birds at K-band and W-band, *Sci Rep*, 2018, **8**, (1), pp 1–11. doi: [10.1038/s41598-018-35880-9](https://doi.org/10.1038/s41598-018-35880-9).
- [43] Chen, X., Guan, J., Liu, N., et al. Maneuvering target detection via Radon-fractional Fourier transform-based long-time coherent integration, *IEEE Trans Sig Process*, 2014, **62**, (4), pp 939–953. doi: [10.1109/TSP.2013.2297682](https://doi.org/10.1109/TSP.2013.2297682).
- [44] Chen, X., Guan, J., Chen, W., et al. Sparse long-time coherent integration-based detection method for radar low-observable manoeuvring target, *IET Radar Sonar Navigat*, 2020, **14**, (4), pp 538–546. doi: [10.1049/iet-rsn.2019.0313](https://doi.org/10.1049/iet-rsn.2019.0313).
- [45] Chen, X., Guan, J., Wang, G. et al. Fast and refined processing of radar maneuvering target based on hierarchical detection via sparse fractional representation, *IEEE Access*, 2019, **7**, pp 149878–149889. doi: [10.1109/ACCESS.2019.2947169](https://doi.org/10.1109/ACCESS.2019.2947169).
- [46] Patel, J.S., Fioranelli, F. and Anderson, D. Review of radar classification and RCS characterization techniques for small UAVs or drones, *IET Radar Sonar Navigat*, 2018, **12**, pp 911–919.
- [47] Chen, W., Liu, J. and Li, J. Classification of UAV and bird target in low-altitude airspace with surveillance radar data, *Aeronaut J*, 2019, **123**, pp 191–211.
- [48] Messina, M. and Pinelli, G. Classification of drones with a surveillance radar signal, *12th International Conference on Computer Vision Systems (ICVS)*. Springer, Thessaloniki, Greece, 2019, pp 1–10.
- [49] Torvik, B., Olsen, K.E. and Griffiths, H. Classification of birds and UAVs based on radar polarimetry, *IEEE Geosci Rem Sens Lett*, 2016, **13**, pp 1305–1309.
- [50] Kim, B.K., Kang, H.S. and Park, S.O. Drone classification using convolutional neural networks with merged Doppler images, *IEEE Geosci. Rem. Sens. Lett*, 2016, **14**, pp 38–42.
- [51] Mendis, G.J., Randeny, T., Wei, J., et al. Deep learning based doppler radar for micro UAS detection and classification, *MILCOM 2016-2016 IEEE Military Communications Conference*. IEEE Press, Baltimore, MD, USA, 2016, pp 924–929.
- [52] Wang, L., Tang, J. and Liao, Q. A study on radar target detection based on deep neural networks. *IEEE Sens Lett*, 2019, doi: [10.1109/LSENS.2019.2896072](https://doi.org/10.1109/LSENS.2019.2896072).
- [53] Samaras, S., Diamantidou, E., Ataloglou, D., et al. Deep learning on multi sensor data for counter UAV applications—A systematic review, *Sensors*, 2019, **19**, (22), pp 4837–4871.
- [54] Regev, N., Yoffe, I. and Wulich, D. Classification of single and multi propelled miniature drones using multilayer perceptron artificial neural network, *International Conference on Radar Systems (Radar 2017)*. IET Press, Belfast, UK, 2017, pp 1–5.
- [55] Habermann, D., Dranka, E., Caceres, Y., et al. Drones and helicopters classification using point clouds features from radar, 2018 *IEEE Radar Conference (RadarConf18)*. IEEE Press, Oklahoma City, OK, USA, 2018, pp 246–251.
- [56] Mohajerin, N., Histon, J., Dizaji, R., et al. Feature extraction and radar track classification for detecting UAVs in civilian airspace, 2014 *IEEE Radar Conference*. IEEE Press, Cincinnati, OH, USA, 2014, pp 674–679.
- [57] Krizhevsky, A., Sutskever, I. and Hinton, G.E. Imagenet classification with deep convolutional neural networks, *Adv Neural Inf Process Syst* 2012, **25**, pp 1097–1105.
- [58] Saqib, M., Khan, S.D., Sharma, N., et al. A study on detecting drones using deep convolutional neural networks, 2017 *14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE Press, Lecce, Italy, 2017, pp 1–5.
- [59] Simonyan, K. and Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.
- [60] Zeiler, M.D. and Fergus, R. Visualizing and understanding convolutional networks. *Proceedings of the European Conference on Computer Vision*, Zurich, Switzerland, 6–12 September 2014, pp. 818–833.
- [61] Craye, C. and Ardjoune, S. Spatio-temporal semantic segmentation for drone detection, 2019 *16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE Press, Taipei, China, 2019, pp 1–5.



- [62] Vasileios Magouliianitis, D.A., Anastasios Dimou, D.Z., Daras, P. Does deep super-resolution enhance UAV detection, 2019 *16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE Press, Taipei, China, 2019, pp 1–6.
- [63] Aker, C. and Kalkan, S. Using deep networks for drone detection, 2017 *14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE Press, Lecce, Italy, 2017, pp 1–6.
- [64] Redmon, J., Divvala, S., Girshick, R. and Farhadi, A. You only look once: Unified, real-time object detection, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 26 June–1 July 2016, pp 779–788.
- [65] Rozantsev, A., Lepetit, V. and Fua, P. Detecting flying objects using a single moving camera, *IEEE Trans Pattern Anal Mach Intell*, 2016, **39**, (5), pp 879–892.
- [66] Gökçe, F., Üçoluk, G., Sahin, E., et al. Vision-based detection and distance estimation of micro-unmanned aerial vehicles, *Sensors*, 2015, **15**, (9), pp 23805–23846.
- [67] Anthony Thomas, V.L., Antoine Cotinat, P.F., Gilber, M. UAV localization using panoramic thermal cameras, *Proceedings of the 12th International Conference on Computer Vision Systems (ICVS)*, Thessaloniki, Greece, 23–25 September 2019.
- [68] Liu, J., Zhang, S., Wang, S., et al. Multispectral deep neural networks for pedestrian detection, *British Machine Vision Conference 2016*. British Machine Vision Association, York, UK, 2016, pp 1–13.
- [69] Cao, Y., Guan, D., Huang, W., et al. Pedestrian detection with unsupervised multispectral feature learning using deep neural networks, *Inform Fusion*, 2019, **46**, pp 206–217.
- [70] John, V., Mita, S., Liu, Z., et al. Pedestrian detection in thermal images using adaptive fuzzy C-means clustering and convolutional neural networks, 2015 *14th IAPR International Conference on Machine Vision Applications (MVA)*, IEEE Press, Tokyo, Japan, 2015, pp 246–249.
- [71] Ulrich, M., Hess, T., Abdulatif, S., et al. Person recognition based on micro-Doppler and thermal infrared camera fusion for firefighting, 2018 *21st International Conference on Information Fusion (FUSION)*. IEEE Press, Cambridge, UK, 2018, pp 919–926.
- [72] Viola, P. and Jones, M. Robust real-time object detection, *Int J Comput Vis*, 2001, **4**, p 4.
- [73] Liu, Q., Lu, X., He, Z., et al. Deep convolutional neural networks for thermal infrared object tracking, *Knowl-Based Syst*, 2017, **134**, pp 189–198.
- [74] Microflown AviSa. Sky sentry: Acoustic MicroDrone localization: system in an urban environment 2015. Available online: <https://microflown-avisa.com/wp-content/uploads/2015/09/PRESSRELEASE-ACOUSTICDRONELOCALIZATIONSYSTEM150715.pdf>
- [75] Lee, H., Pham, P., Largman, Y., et al. Unsupervised feature learning for audio classification using convolutional deep belief networks, 22nd *International Conference on Neural Information Processing Systems*. IEEE Press, Vancouver, BC, Canada, 2009, pp 1096–1104.
- [76] Piczak K.J. Environmental sound classification with convolutional neural networks, 2015 *IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE Press, Boston, MA, USA, 2015, pp 1–6.
- [77] Lane, N.D., Georgiev, P. and Qendro, L. DeepEar: robust smartphone audio sensing in unconstrained acoustic environments using deep learning, 2015 *ACM International Joint Conference on Pervasive and Ubiquitous Computing*. IEEE Press, Osaka, Japan, 2015, pp 283–294.
- [78] Wilkinson, B., Ellison, C., Nykaza, E.T., et al. Deep learning for unsupervised separation of environmental noise sources, *J Acoust Soc Amer*, 2017, **141**, pp 3964–3964.
- [79] Park, S., Shin, S., Kim, Y., et al. Combination of radar and audio sensors for identification of rotor-type unmanned aerial vehicles (UAVs), 2015 *IEEE SENSORS*. IEEE Press, Busan, Korea, 2015, pp 1–4.
- [80] Liu, H., Wei, Z., Chen, Y., et al. Drone detection based on an audio-assisted camera array, 2017 *IEEE Third International Conference on Multimedia Big Data (BigMM)*. IEEE Press, Laguna Hills, CA, USA, 2017, pp 402–406.
- [81] Kim, J., Park, C., Ahn, J., et al. Real-time UAV sound detection and analysis system, 2017 *IEEE Sensors Applications Symposium (SAS)*. IEEE Press, Glassboro, NJ, USA, 2017, pp 1–5.
- [82] Kim, J. and Kim, D. Neural network based real-time UAV detection and analysis by sound, *J Adv Inform Technol Conver*, 2018, **8**, (1), pp 43–52.
- [83] Jeon, S., Shin, J.W., Lee, Y.J., et al. Empirical study of drone sound detection in real-life environment with deep neural networks, 2017 25th *European Signal Processing Conference (EUSIPCO)*. IEEE Press, Kos, Greece, 2017, pp 1858–1862.
- [84] Ma, C., Zhang, J. and Bao, M. UAV acoustic detection based on non-parametric fusion of spatial-frequency characteristics of sound field, *J Sig Process*, 2019, **35**, (9), pp 1590–1598.
- [85] Wang, W., An, T. and Ou, J. Research on audio detection and recognition of UAV, *Tech Acoust*, 2018, **37**, (1), pp 89–93.
- [86] Zhu, Y., Tu, D., Li, X. Research on the field test system of radio monitoring and direction finding facilities, *China Radio*, 2020, **2020**, (1), pp 38–41.
- [87] Wang, Z., Zhou, X. and Xu, C. Research on detection and control method of civilian UAV based on radio technology in prison environment, *Comput Sci Appl*, 2018, **8**, (9), pp 1407–1415.
- [88] O’Shea, T., Corgan, J. and Clancy, T.C. Convolutional radio modulation recognition networks, *International Conference on Engineering Applications of Neural Networks*. Springer, Aberdeen, UK, 2016, pp 213–226.
- [89] Karra, K., Kuzdeba, S. and Pertersen, J. Modulation recognition using hierarchical deep neural networks, 2017 *IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. IEEE Press, Piscataway, New Jersey, USA, 2017, pp 1–3.
- [90] West, N. and O’Shea, T. Deep architectures for modulation recognition, 2017 *IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. IEEE Press, Piscataway, New Jersey, USA, 2017, pp 1–6.

- [91] Arumugam, K., Kadampot, I., Tahmasbi, M., et al. Modulation recognition using side information and hybrid learning, 2017 *IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. IEEE Press, Piscataway, New Jersey, USA, 2017, pp 1–2.
- [92] Schmidt, M., Block, D. and Meier, U. Wireless interference identification with convolutional neural networks, 2017 *IEEE 15th International Conference on Industrial Informatics (INDIN)*. IEEE Press, Emden, Germany, 2017, pp 180–185.
- [93] Baltrušaitis, T., Ahuja, C. and Morency, L.P. Multimodal machine learning: A survey and taxonomy, *IEEE Trans Pattern Anal Mach Intell* 2018, **41**, pp 423–443.
- [94] Endo, Y., Toda, H., Nishida, K. and Ikedo, J. Classifying spatial trajectories using representation learning, *Int J Data Sci Anal*, 2016, **2**, pp 107–117.
- [95] Bounding Box Detection of Drones. 2017. Available online: <https://github.com/creiser/drone-detection/>
- [96] MultiDrone Public DataSet. 2018. Available online: <https://multidrone.eu/multidrone-public-dataset/>
- [97] Yuan, B.Q., Wang, Y.S and Zheng, L.G. A survey of deep learning applied on radio signal modulation recognition, *Appl Electron Tech*, 2019, **45**, (5), pp 1–4.
- [98] Zhang, Y. and Hu, S. Research on radar/visible light surveillance and tracking methods of target in low-altitude airspace, *Comput Eng Appl*, 2018, **54**, (6), pp 234–240.
- [99] Jovanoska, S., Brötje, M. and Koch, W. Multisensor data fusion for UAV detection and tracking, *Proceedings of the 2018 19th International Radar Symposium (IRS)*, Bonn, Germany, 20–22 June 2018, pp. 1–10.
- [100] Hengy, S., Laurenzis, M., Schertzer, S., Hommes, A., Kloeppe, F., Shoykhetbrod, A., Geibig, T., Johannes, W., Rassy, O. and Christnacher, F. Multimodal UAV detection: Study of various intrusion scenarios. In *Proceedings of the Electro-Optical Remote Sensing XI International Society for Optics and Photonics*, Warsaw, Poland, 11–14 September 2017, Volume **10434**, p. 104340P.
- [101] Laurenzis, M., Hengy, S., Hammer, M., Hommes, A., Johannes, W., Giovanneschi, F., Rassy, O., Bacher, E., Schertzer, S. and Poyet, J.M. An adaptive sensing approach for the detection of small UAV: First investigation of static sensor network and moving sensor platform, *Proceedings of the Signal Processing, Sensor/Information Fusion, and Target Recognition XXVII International Society for Optics and Photonics*, Orlando, FL, USA, 16–19 April 2018, Volume **10646**, p 106460S.