remote sensing

MDPI

Article

Ground and Low-Altitude Target Classification in Cluttered Radar Remote Sensing via Velocity-Aware Multi-Feature Fusion

Peilong Hu [1], Liyu Tian [2,*], Mengze Zhang [2] and Zhongshan Zhang [1]

[1] Beijing Institute of Technology, Zhuhai 519088, China; hupeilong0526@163.com (P.H.); zhangzs@bit.edu.cn (Z.Z.)
[2] School of Integrated Circuits and Electronics, Beijing Institute of Technology, Beijing 100081, China; zmzbit2022@163.com
* Correspondence: tianliyu@bit.edu.cn

Highlights

What are the main findings?

- A velocity-aware multi-feature fusion method was proposed for classifying ground and low-altitude targets using measured X-band pulse-Doppler radar echoes.
- By combining echo preprocessing and discriminative feature extraction, the proposed method achieved robust separation of humans, vehicles, and UAVs in cluttered outdoor environments.

What are the implications of the main findings?

- The proposed framework provides a practical approach for radar target classification under complex outdoor clutter conditions.
- The study supports the use of measured radar data and multi-feature fusion for ground surveillance and low-altitude target monitoring.

Abstract

Classification of ground and low-altitude targets in radar remote sensing is challenging because environmental clutter and noise can significantly degrade the discriminability of target echoes, especially under complex outdoor observation conditions. To improve the classification performance for humans, vehicles, and unmanned aerial vehicles (UAVs), this paper proposes a velocity-aware multi-feature fusion method based on measured radar echo data. First, radar echoes are preprocessed using a wavelet-decomposition-based strategy to suppress clutter and noise while preserving useful target information. Then, multiple complementary features, including wavelet packet energy distribution, spectral entropy, spectral standard deviation, temporal standard deviation, amplitude dispersion coefficient, and relative radar cross-section (RCS), are extracted to characterize the target echoes from different perspectives. Considering the influence of target velocity on Doppler distribution and class separability, the measured data are further divided into different velocity intervals for stratified classification. Based on the fused feature vectors, a long short-term memory (LSTM) network is employed to model feature relationships and perform target classification. Experiments conducted on real measured radar echo data demonstrate that the proposed method achieves classification accuracies of 97.82% for UAVs, 96.00% for vehicles, and a mean interval-level accuracy of 96.94%, indicating its effectiveness for ground and low-altitude target classification in cluttered radar remote sensing environments.

Academic Editor: Massimiliano Pieraccini

Received: 7 April 2026
Revised: 18 May 2026
Accepted: 26 May 2026
Published: 1 June 2026

Keywords: radar remote sensing; ground and low-altitude target classification; multi-feature fusion; clutter suppression; velocity-aware classification; long short-term memory (LSTM)

1. Introduction

Ground-based radar remote sensing plays an important role in the observation and classification of ground and low-altitude targets in complex outdoor environments. It has found wide applications in both military and civilian scenarios, including situational awareness, weather monitoring, and autonomous driving [1–4]. Compared with optical sensing systems, radar is less sensitive to illumination and weather conditions, which makes it more suitable for all-day and all-weather target observation. However, reliable classification of ground and low-altitude targets remains a challenging task in practical radar sensing systems, especially when targets are embedded in cluttered backgrounds and contaminated by environmental noise. In general, radar target classification consists of two stages: the training stage and the operational stage [4]. In the training stage, discriminative information is extracted through data acquisition, preprocessing, feature extraction, and classifier training to construct a recognition model; in the operational stage, the trained feature extractor and classifier are used to process incoming data and identify target categories. Therefore, classification performance depends strongly on both the robustness of extracted features and the effectiveness of the classification model [5].

Despite the continuous development of radar target recognition techniques, several challenges remain in real-world applications. First, in complex electromagnetic environments, target echoes are often affected by clutter, noise, and interference, which significantly degrades feature robustness and reduces classification reliability. This problem is particularly severe for low-velocity targets, whose Doppler components are more likely to overlap with strong clutter-dominated low-frequency regions. Second, the increasing diversity of target types and motion patterns makes it difficult for classification methods based on a single type of feature to maintain stable performance under varying observation conditions. In addition, although improving classification accuracy is always an important objective, practical systems also require efficient processing and good adaptability, which further increases the difficulty of method design. For this reason, developing a classification framework that is both robust and effective in complex observation conditions remains an important research topic.

Existing radar target classification methods mainly differ in their choice of target representation. One representative category is based on radar cross-section (RCS) characteristics. Since RCS describes the electromagnetic scattering capability of a target and can be obtained without stringent bandwidth requirements, it has long been used as an important feature for radar target recognition [6,7]. Han et al. [8] analyzed RCS and other characteristics of conical ballistic targets and constructed several deep learning models using electromagnetically simulated datasets. However, the lack of validation on real measured data limits the assessment of their practical performance. Reference [9] proposed an unmanned aerial vehicle (UAV) classification method using RCS sequences and Bayesian fusion in a multistatic radar system. Although RCS-based methods are physically meaningful, their discriminative ability may be affected by target aspect, material, and measurement conditions, making it difficult for them to provide sufficiently robust performance on their own.

Another important category is based on micro-Doppler characteristics. Because different targets exhibit different micro-motion mechanisms, the induced Doppler modulation can provide useful motion-sensitive information for classification [10]. Du et al. [11] employed time–frequency analysis to compare the micro-Doppler characteristics of walking humans, two-person walking targets, and moving wheeled vehicles, and then used the extracted features for classification. Reference [12] visualized human micro-Doppler features and incorporated an attention mechanism into AlexNet to classify different human motion states. In addition, polarization-based methods have also been investigated to exploit the structural and electromagnetic differences among targets [13,14]. These approaches can provide discriminative information from specific perspectives, but many of them either rely heavily on idealized conditions or lack sufficient validation using measured radar data. Moreover, directly using high-dimensional time–frequency or multi-polarization representations often increases computational complexity and reduces the practicality of the method in real radar sensing applications.

With the rapid development of machine learning, deep learning models such as convolutional neural networks (CNNs) and long short-term memory (LSTM) networks have also been widely introduced into radar target classification [15,16]. Reference [17] proposed a CNN-LSTM method with a multi-view attention mechanism for synthetic aperture radar (SAR) target recognition, achieving a recognition rate above 95.57%. Reference [18] introduced a one-dimensional CNN-based method for low-resolution ground surveillance radar target recognition, achieving 99% accuracy in human and vehicle classification. These methods have demonstrated strong representation capabilities when sufficient and consistent training data are available. However, end-to-end deep learning methods are often highly dependent on the quality, quantity, and distribution consistency of the training data. When the radar operating environment changes, the trained model may suffer from degraded generalization ability and may require retraining using newly collected data [15–18]. This limitation is particularly notable in complex outdoor radar sensing scenarios, where target echoes are influenced by background clutter, observation geometry, and motion variation.

To address the above issues, this paper proposes a velocity-aware multi-feature fusion method for ground and low-altitude target classification using measured radar echo data, as shown in Figure 1. First, radar echoes are preprocessed to suppress clutter and noise and to improve the signal-to-clutter ratio and signal-to-noise ratio. Then, multiple complementary features are extracted from different domains, including wavelet packet energy distribution, spectral entropy, spectral standard deviation, temporal standard deviation, amplitude dispersion coefficient, and relative radar cross-section (RCS). By combining these features, high-dimensional raw echo data are transformed into low-dimensional yet physically meaningful feature vectors, which improves target separability while reducing the input burden of the classifier. Furthermore, considering that target velocity directly affects Doppler distribution and classification difficulty, a velocity-stratified classification strategy is introduced to divide measured data into different velocity intervals for more refined recognition. On this basis, an LSTM network is employed to model the relationships among fused features and perform final classification.

The main contributions of this paper can be summarized as follows:

- A wavelet-based radar echo preprocessing framework is proposed for suppressing low-frequency clutter while preserving useful target information in measured radar echoes.
- A velocity-aware multi-feature fusion strategy is designed by combining complementary time-domain, frequency-domain, and wavelet-domain features to improve target separability.
- A velocity-stratified classification mechanism is introduced to reduce the influence of Doppler-spectrum overlap and improve classification performance under different motion conditions.

- Experimental results based on measured X-band pulse-Doppler radar data demonstrate that the proposed framework achieves effective classification performance for humans, vehicles, and UAVs in cluttered environments.

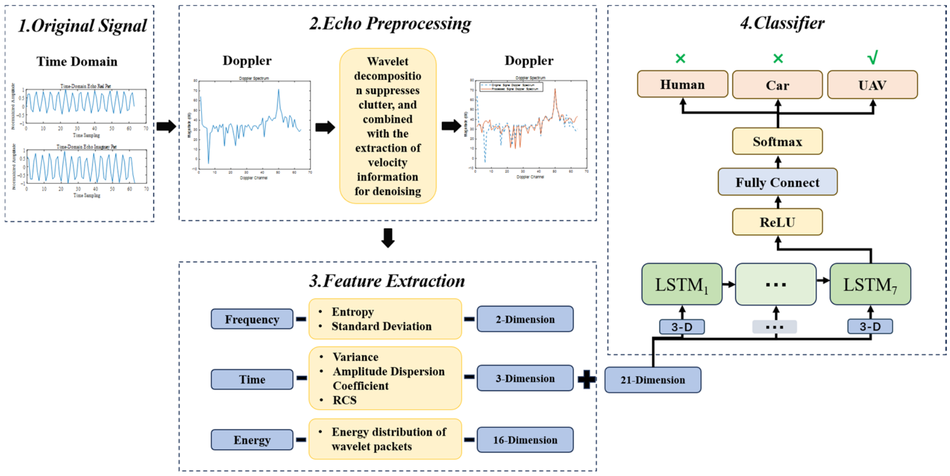


Figure 1. Flowchart of the Proposed Method in This Paper. In Step 2, echo preprocessing is illustrated. The left panel shows the unprocessed echo. In the right panel, the blue curve represents the unprocessed echo, while the orange curve represents the preprocessed echo.

The remainder of this paper is organized as follows. Section 2 introduces the proposed radar target classification framework, including preprocessing, feature extraction, and classification methods. Section 3 presents the experimental results and performance analysis. Section 4 discusses the effectiveness, applicability, and limitations of the proposed framework. Finally, Section 5 concludes the paper and outlines future research directions.

2. Materials and Methods

2.1. Observation Scenario and Measured Radar Dataset

The dataset used in this study was collected using a ground-based X-band pulse-Doppler radar system. Due to confidentiality restrictions associated with the radar hardware platform and deployment configuration, detailed radar-system parameters cannot be disclosed in this paper. Nevertheless, to improve the transparency of the experimental conditions, system-level parameter information characterizing the radar sensing capability is provided. Representative approximate values of several radar-system parameters are summarized in Table 1 to facilitate evaluation of the engineering applicability of the proposed method.

Table 1. Representative radar-system parameters of the measured radar system.

Parameter	Value
Operating band	X-band
Radar type	Pulse-Doppler radar
Waveform type	LFM chirp pulse
Range resolution	4 m
Velocity resolution	0.7 m/s
Maximum unambiguous velocity	21 m/s
Maximum detection range	35 km

The measurements were conducted in representative outdoor environments with complex background clutter. Major clutter sources mainly arose from ground surfaces, vegetation, buildings, and road-related structures, which increased the difficulty of reliable classification, especially for low-velocity targets.

The measured dataset includes radar echoes from humans, vehicles, and unmanned aerial vehicles (UAVs). During each experimental session, the radar was operated under a fixed sensing configuration, and the same signal processing procedure was applied to all samples. After conventional front-end processing, the range bin containing the target response was selected, and the corresponding slow-time signal within a coherent observation window was extracted for subsequent analysis.

2.2. Radar Echo Characteristics of Ground and Low-Altitude Targets

For the ground-based X-band pulse-Doppler radar considered in this study, the transmitted signal is a chirp waveform, which can be expressed as

$$S(t) = \omega(t)cos\left(2\pi f_c t + \pi\mu t^2\right) \tag{1}$$

where $\omega(t)$ is the window function, f_c denotes the carrier frequency, μ is the frequency modulation rate, and T_p is the pulse duration. After target reflection, digital down-conversion, and pulse compression, the range bin containing the target response is selected. The corresponding slow-time signal can be written as

$$x[m] = e^{j2\pi f_D m T_r} \sum_{l=1}^{L} A_l e^{j2\pi f_{md,l} m T_r} + c[m] + n[m] \tag{2}$$

where A_l denotes the complex amplitude of the l-th scatterer, f_D and $f_{md,l}$ represent the bulk Doppler and micro-Doppler components, respectively, $c[m]$ denotes clutter, $n[m]$ denotes noise, and T_r is the pulse repetition interval (PRI). For all target categories considered in this paper, the micro-Doppler component is determined by the projection of the local micro-motion onto the radar line of sight (LOS). Here, a unified LOS projection coefficient, denoted by η, is introduced to describe the radial component of local micro-motion. Although humans, vehicles, and UAVs exhibit different motion mechanisms, their echo models can be described within the same physical framework, namely, that the radial micro-motion observed by the radar is obtained from the LOS projection of the corresponding local motion.

For human targets, the bulk motion of the body produces the main Doppler shift, while the periodic swinging of the arms and legs introduces additional micro-Doppler modulation. The limbs can be approximately modeled as rotating scatterers with a stop-and-hop motion pattern. For the k-th limb, the instantaneous radial velocity can be written as

$$v_k[m] \approx L_k \omega_k cos(\omega_k m T_r + \phi_k)\eta_{H,k} \tag{3}$$

where L_k, ω_k, and ϕ_k denote the equivalent length, angular velocity, and initial phase of the k-th limb, respectively, and $\eta_{H,k}$ is the LOS projection coefficient of the corresponding limb motion. The corresponding micro-Doppler frequency is

$$f_{md,k}[m] = \frac{2v_k[m]}{\lambda} = \frac{2L_k \omega_k \eta_{H,k}}{\lambda} cos(\omega_k m T_r + \phi_k) \tag{4}$$

Thus, the human slow-time echo at the selected range bin can be expressed as

$$x_H[m] = e^{j2\pi f_D m T_r} \sum_{l \in (Limbs)} A_l e^{j2\pi f_{md,l}[m] m T_r} + c[m] + n[m] \tag{5}$$

This model indicates that human echoes usually exhibit relatively strong temporal fluctuation and dispersed Doppler energy due to periodic limb motion. Accordingly, the Doppler spectrum of a human target generally contains a body-related main component together with broader motion-induced side structures.

For vehicle targets, the dominant part of the echo is associated with the translational motion of the vehicle body, whereas the micro-Doppler component is mainly introduced by rotating wheels or other moving mechanical parts. For a scatterer located on the wheel rim, the radial displacement can be approximated as

$$\Delta R_l[m] = r sin(\omega m T_r + \phi_l)\eta_{V,l} \tag{6}$$

where r is the wheel radius, ω is the angular velocity, ϕ_l is the initial phase, and $\eta_{V,l}$ is the LOS projection coefficient of the local wheel motion. The corresponding micro-Doppler frequency is

$$f_{md,l}[m] = \frac{2d\Delta R_l}{\lambda dt} \approx \frac{2r\omega}{\lambda} cos(\omega m T_r + \phi_l)\eta_{V,l} \tag{7}$$

Hence, the vehicle echo can be written as

$$x_V[m] = e^{j2\pi f_D m T_r} = \sum_{l=1}^{N} A_l e^{j2\pi f_{md,l}[m] m T_r} + c[m] + n[m] \tag{8}$$

Compared with human targets, vehicle echoes are generally more stable and contain stronger body scattering. Therefore, their Doppler spectra often exhibit a more concentrated main component, while the wheel-induced modulation introduces relatively fine spectral broadening or sideband-like structures.

For UAV targets, the main Doppler component is generated by the overall motion of the platform, while the most distinctive micro-Doppler effect is caused by periodic rotor-blade rotation. For the p-th blade of the q-th rotor, the radial displacement can be approximated as

$$\Delta R_{q,p}[m] = R_{b,q}\eta_{U,q} sin(\omega_q m T_r + \phi_{q,p}) \tag{9}$$

where $R_{b,q}$ is the effective blade radius, ω_q is the angular velocity of the q-th rotor, $\eta_{U,q}$ is the LOS projection coefficient of the corresponding rotor-induced motion, and $\phi_{q,p}$ is the initial phase of the blade. The corresponding micro-Doppler frequency is

$$f_{md,q,p}[m] = \frac{2d\Delta R_{q,p}}{\lambda dt} \approx \frac{2R_{b,q}\omega_q}{\lambda} \eta_{U,q} cos(\omega_q m T_r + \phi_{(q,p)}) \tag{10}$$

By superimposing the contributions of all illuminated blades within the selected range bin, the UAV slow-time echo can be expressed as

$$x_U[m] = e^{j2\pi f_D m T_r} \sum_{q=1}^{Q} \sum_{p=1}^{N_b} B_{q,p} e^{j2\pi f_{md,q,p}[m] m T_r} + c[m] + n[m] \tag{11}$$

Because of the rotor-blade modulation, UAV echoes usually produce more pronounced and structured Doppler spreading than those of ground targets. As a result, the Doppler spectrum of a UAV often exhibits a more complex energy distribution than that of humans and vehicles.

In summary, although the echoes of humans, vehicles, and UAVs all contain bulk Doppler, clutter, and noise components, their local micro-motion mechanisms and scattering structures are significantly different. These differences are reflected in the temporal fluctuation, spectral distribution, and energy concentration of the echoes, and can be directly observed in their Doppler spectra. Therefore, the Doppler-spectrum characteristics, together with time-domain, frequency-domain, wavelet-packet-energy, and relative-RCS-related features, provide a reasonable basis for the multi-feature fusion strategy adopted in this paper.

To illustrate the above models, representative Doppler spectra of human, vehicle, and UAV targets were simulated based on Equations (5), (8), and (11), as shown in Figure 2. The three target categories exhibit distinct spectral spreading patterns and energy distributions, reflecting the effects of limb swinging, wheel rotation, and rotor-blade motion. These differences provide an intuitive basis for the feature extraction strategy adopted in the following sections.

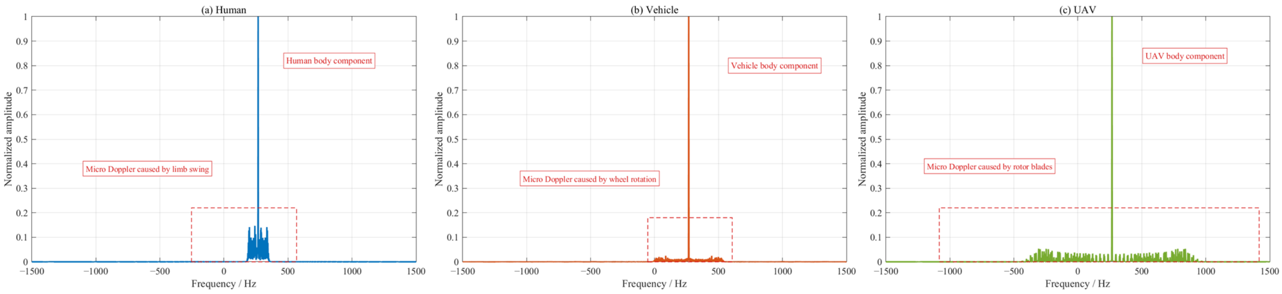


Figure 2. Simulated Doppler spectra of three target categories based on the echo models in Section 2.2: (**a**) human target; (**b**) vehicle target; (**c**) UAV target.

2.3. Proposed Velocity-Aware Multi-Feature Fusion Method

To improve the classification performance of ground and low-altitude targets in cluttered outdoor environments, a velocity-aware multi-feature fusion method is proposed in this paper. As illustrated in Figure 1, the method consists of four main stages: echo preprocessing, multi-domain feature extraction, feature fusion, and velocity-stratified classification. First, wavelet-decomposition-based preprocessing is applied to suppress clutter and noise while preserving useful target information. Then, complementary features are extracted from the wavelet packet domain, time domain, and frequency domain to characterize target echoes from different perspectives. After feature normalization and fusion, the samples are divided into different velocity intervals and classified using an LSTM network.

2.3.1. Velocity-Aware Echo Preprocessing

In practical radar target recognition, low-velocity target echoes often overlap with ground clutter in the Doppler domain, especially in complex outdoor environments. For ground-based radars, clutter energy is mainly concentrated near the zero-frequency region and may be significantly stronger than the target echo, which degrades subsequent feature extraction and classification [19]. Conventional methods such as the moving target indication (MTI) filter, CLEAN, and the generalized matched filter (GMF) suffer from limitations including signal distortion, prior-clutter dependence, and high computational cost [20]. To address these issues, this paper adopts a velocity-aware clutter and noise suppression method based on wavelet decomposition [21,22].

First, the number of wavelet decomposition levels is selected as 6. This choice provides a compromise between clutter separation capability and reconstruction distortion, and also matches the low-frequency distribution characteristics of ground clutter [23]. After six-level decomposition, the obtained coefficients are denoted as $CD1$, $CD2$, $CD3$, $CD4$, $CD5$, $CD6$, and $CA6$. Since $CA6$ mainly describes the low-frequency trend of the signal, it is used to represent the clutter-dominated component near zero Doppler. By setting $CA6$ to zero during reconstruction, the low-frequency clutter component can be effectively suppressed.

Second, the detail coefficients are processed adaptively according to target velocity. The unambiguous velocity of the radar is 21 m/s. To match the migration of target Doppler energy from lower-frequency to higher-frequency wavelet detail bands as the target velocity increases, two practical cutoff values are introduced at one-third and two-thirds of the unambiguous velocity, i.e., 7 m/s and 14 m/s. This partition provides a simple three-region description of low-, medium-, and high-velocity regimes and is used to determine which detail coefficients are more likely to contain target-related information. Specifically, when the target velocity is lower than 7 m/s, threshold denoising is applied to $CD1$ and $CD2$; when the target velocity lies between 7 m/s and 14 m/s, denoising is applied to $CD5$ and $CD6$; and when the target velocity exceeds 14 m/s, denoising is applied to $CD3$, $CD4$, $CD5$, and $CD6$. In this way, noise can be reduced while preserving the coefficients that are more likely to carry target micro-Doppler information.

Let $W_j(k)$ denote the wavelet coefficient at scale j and index k, and let $\widetilde{W}_j(k)$ denote the processed coefficient. The wavelet threshold denoising rule is written as

$$\widetilde{W_j}(k) = \begin{cases} \mathrm{W_j(k)}, & |\ \mathrm{W_j}(k)\ | \geq \lambda_\mathrm{j} \\ 0, & |\ \mathrm{W_j(k)}\ | < \lambda_\mathrm{j} \end{cases} \tag{12}$$

where λ_j is the threshold at scale j, defined as

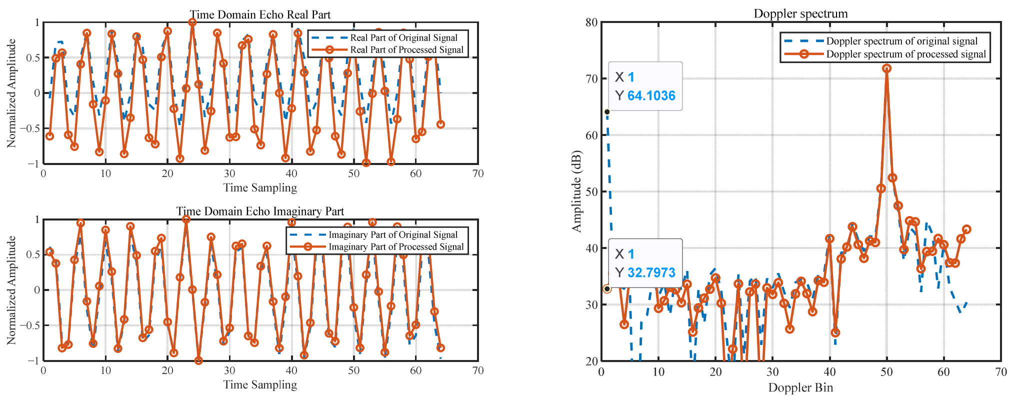
$$\lambda_j = \sigma_j\sqrt{2lnN},\ \sigma_j = \frac{median(|CD_j|)}{0.6745}$$

where N is the signal length and σ_j is the estimated noise standard deviation [24].

After processing $CA6$ and the selected detail coefficients, the denoised echo is reconstructed as

$$\hat{x}[m] = IDWT\left(\widetilde{CD1}, \widetilde{CD2}, \widetilde{CD3}, \widetilde{CD4}, \widetilde{CD5}, \widetilde{CD6}, \widetilde{CA6}\right) \tag{13}$$

where $\widetilde{CA6} = 0$. As shown in Figure 3, the clutter amplitude near zero Doppler is significantly reduced, while the useful target component is preserved.



(**a**) (**b**)

Figure 3. Time domain and frequency domain of a target before and after echo clutter suppression. (**a**) Time domain waveform of the echo signal. (**b**) Frequency domain waveform of the echo signal.

Note that the thresholds used in preprocessing and those used in classification are defined for different purposes. The former are used for adaptive wavelet-coefficient selection, whereas the latter are used for velocity-stratified classification; therefore, the two sets of thresholds are different.

2.3.2. Wavelet Packet Energy Feature Extraction

After echo preprocessing, wavelet packet decomposition is used to extract the energy distribution features of the target echo across different frequency bands [25]. Since different target velocities may shift the main Doppler component to different positions, the Doppler peak is first shifted to one quarter of the Doppler-frequency range to enhance feature comparability across samples. Then, a four-level wavelet packet decomposition is performed using the "db15" wavelet basis, yielding 16 sub-bands.

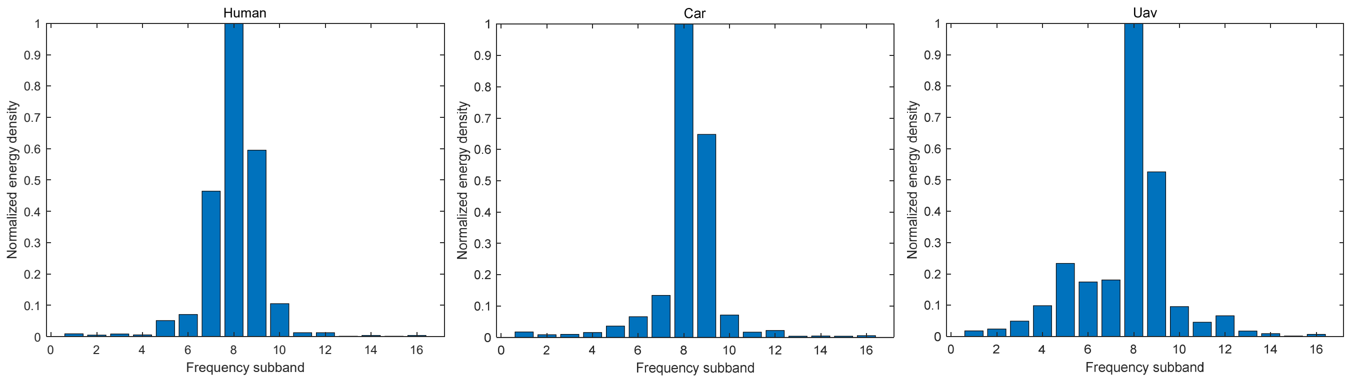
Let $c_k(n)$ denote the coefficient sequence of the k-th sub-band. The energy density of the k-th sub-band is defined as

$$E_k = \sum_{n=1}^{N_k} | c_k(n) |^2 \tag{14}$$

where N_k is the number of coefficients in the k-th sub-band. The 16-dimensional wavelet packet energy feature vector is then normalized as

$$f_{wp} = \left[\frac{E_1}{\sum_{k=1}^{16} E_k}, \frac{E_2}{\sum_{k=1}^{16} E_k}, \ldots, \frac{E_{16}}{\sum_{k=1}^{16} E_k} \right] \tag{15}$$

As shown in Figure 4, the wavelet packet energy distributions of humans, vehicles, and UAVs differ significantly: human and vehicle echoes tend to concentrate energy around several dominant bands, whereas UAV echoes usually exhibit a more dispersed distribution.



(**a**) (**b**) (**c**)

Figure 4. Wavelet packet energy feature of three types of targets. (**a**) Wavelet energy feature of a human echo. (**b**) Wavelet energy feature of a car echo. (**c**) Wavelet energy feature of a UAV echo.

2.3.3. Time-Domain Feature Extraction

Time-domain statistical features are extracted from the amplitude sequence of the slow-time echo to characterize temporal fluctuation and relative scattering intensity. Let the amplitude sequence within the target-containing range bin be denoted by $a(n)$, $n = 1, 2, \ldots, N$, and let its mean value be

$$\bar{a} = \frac{1}{N} \sum_{n=1}^{N} a(n)$$

The time-domain variance is defined as

$$\sigma_t^2 = \frac{1}{N} \sum_{n=1}^{N} (a(n) - \bar{a})^2 \tag{16}$$

which reflects the fluctuation of the target echo amplitude over slow time.

The amplitude dispersion coefficient is defined as

$$C_v = \frac{\sigma_t}{\bar{a}} \tag{17}$$

where σ_t is the standard deviation of $a(n)$. This feature measures the relative degree of dispersion of the echo amplitude.

In addition, a relative radar cross-section (RCS) feature is introduced. According to the radar equation, the received echo power P_r can be written as

$$P_r = \frac{P_t G A_e \sigma}{(4\pi)^2 R^4} \tag{18}$$

where P_t is the transmitted power, G is the antenna gain, A_e is the effective receiving area, σ is the target RCS, and R is the target range. For a fixed radar system, P_t, G, and A_e are constants, so Equation (19) can be simplified as

$$P_r = k_1 \frac{\sigma}{R^4} \tag{19}$$

Thus, the target RCS can be expressed as

$$\sigma = k_2 P_r R^4 \tag{20}$$

where k_1 and k_2 are constants. Since the absolute RCS is difficult to obtain in practice, this paper uses a relative RCS feature calculated from the echo power and range.

Let $S(n)$ denote the Doppler spectrum amplitude of the target echo. The echo power is estimated as

$$P_r = \sum_{n=1}^{M} \mid S(n) \mid^2 \tag{21}$$

Let n_r denote the range-bin index of the target. The target range is estimated by

$$R = n_r \Delta R, \qquad \Delta R = \frac{c}{2B} \tag{22}$$

where ΔR is the range resolution, c is the speed of light, and B is the radar bandwidth. Substituting the estimated P_r and R into Equation (20) yields the relative RCS feature.

2.3.4. Frequency-Domain Feature Extraction

By applying the FFT to the slow-time signal within the target range bin, the Doppler spectrum of the target can be obtained. Let $S(n)$ denote the Doppler spectrum amplitude and let the normalized spectral amplitude be

$$P_n = \frac{|S(n)|}{\sum_{n=1}^{M} \mid S(n) \mid^2}$$

where M is the number of Doppler bins.

The spectral entropy is defined as

$$H_s = -\sum_{n=1}^{M} p_n log p_n \tag{23}$$

which reflects the concentration or dispersion of the spectral energy distribution [20]. Smaller entropy indicates stronger energy concentration, whereas larger entropy indicates a more dispersed spectrum.

The spectral standard deviation is defined as

$$\sigma_j = \frac{1}{M} \sum_{n=1}^{M} \left(\mid S(n) \mid - \overline{\mid S \mid} \right)^2 \tag{24}$$

where $|\bar{S}|$ denotes the mean spectral amplitude. This feature reflects the degree of fluctuation in the Doppler spectrum amplitude.

2.3.5. Feature Normalization and Fusion

Because the extracted features have different physical meanings and numerical ranges, feature normalization is required before model training [26]. In this paper, the Min–Max normalization method is adopted:

$$x_i' = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{25}$$

where x_{max} and x_{min} denote the maximum and minimum values of the corresponding feature samples, respectively.

After normalization, the 16-dimensional wavelet packet energy feature vector, the 3-dimensional time-domain feature vector, and the 2-dimensional frequency-domain feature vector are concatenated into a 21-dimensional fused feature vector. This fused representation combines motion-related, statistical, and scattering-related information, thereby improving the discriminability of different target categories.

2.3.6. Velocity Interval Division

To further improve classification performance, especially for slow-moving targets, a velocity-stratified classification strategy is introduced. Based on the Doppler spectrum and the actual target velocity, the samples are divided into three velocity intervals: ultra-low velocity, low velocity, and high velocity. The schematic diagram of the velocity interval division is shown in Figure 5, and the corresponding target categories are listed in Table 2.

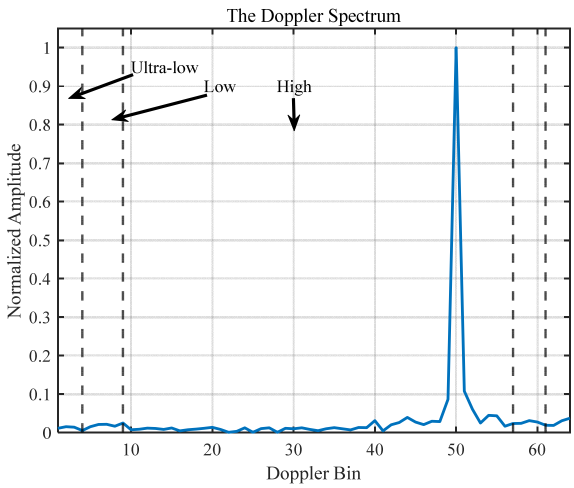


Figure 5. Schematic diagram of velocity interval division. The dashed vertical lines denote the boundaries of the ultra-low-, low-, and high-velocity intervals.

Table 2. Target type of each velocity interval.

Velocity Interval	Target Type	Velocity (m/s)
Ultra-low velocity	Human, Vehicle, UAV	0–1.4
Low velocity	Human, Vehicle, UAV	1.4–5
High velocity	Vehicle, UAV	5–22

In the high-velocity interval, the classification task mainly involves vehicles and UAVs. This design is closely related to the practical motion characteristics of human targets in the considered outdoor radar sensing scenario. In real outdoor environments, human targets are often affected by complex terrain conditions and equipment carrying, which limit sustained high-speed motion. Moreover, the radar measures radial velocity rather than absolute motion speed, and the radial component of human motion is usually smaller than the actual moving speed. As a result, human targets with radial velocities exceeding 5 m/s were rarely observed in the measured dataset.

In addition, the ultra-low-velocity interval and low-velocity interval were further separated because the target spectra in these two intervals exhibit different degrees of overlap with clutter-dominated low-frequency components. In the ultra-low-velocity interval, the overlap between target Doppler spectra and clutter spectra is more significant. Therefore, the target spectra after clutter suppression may still differ from those in the low-velocity interval, even for the same target category. Consequently, the ultra-low-velocity and low-velocity intervals were trained separately in this study, allowing the classifier to better learn the spectral characteristics corresponding to different velocity intervals.

It should be noted that Doppler-spectrum differences alone may not always provide sufficiently reliable target discrimination under strong clutter, noise interference, and varying target-motion conditions. Therefore, the proposed framework first applies clutter suppression and denoising to reduce the influence of low-frequency clutter and noise components. Furthermore, targets whose Doppler spectra exhibit stronger overlap with clutter-dominated low-frequency components are separately assigned to the ultra-low-velocity interval for independent training, thereby reducing the influence of spectral overlap on classification performance.

Therefore, the proposed velocity interval division not only reflects the physical characteristics of target motion and radar observation geometry, but also serves as an effective strategy for reducing clutter-induced feature inconsistency and improving classification performance.

2.3.7. LSTM-Based Classification Model

After feature fusion and velocity stratification, an LSTM network is employed as the classifier. As an improved form of recurrent neural network, LSTM can effectively capture feature dependencies through the input gate, forget gate, output gate, and memory cell [27]. In this paper, the LSTM network consists of an input layer, an LSTM layer, an activation layer, a fully connected layer, and an output layer, as shown in Figure 6.

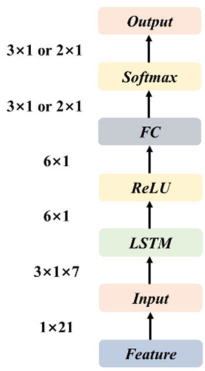


Figure 6. Network structure.

The 21-dimensional fused feature vector is reshaped into a three-dimensional input structure with a time step of 7. Since the fused feature vector contains 21 elements, it is reshaped into a short sequence with a time step of 7 so that the input can be organized into three successive feature groups. This arrangement allows the LSTM to model dependencies among adjacent feature groups while keeping the input structure compact. In our experiments, this setting provided a good trade-off between model complexity and classification performance. After processing by the LSTM layer, a 6-dimensional hidden feature vector is obtained. To enhance the nonlinear representation capability of the network, the hidden feature is further passed through a ReLU activation layer and then fed into the fully connected layer. Since the number of target categories differs across velocity intervals, the output dimension of the fully connected layer is set to 3 for the ultra-low-velocity and low-velocity intervals, and 2 for the high-velocity interval. Finally, the Softmax output layer is used to generate the target classification result.

During training, the cross-entropy loss function is used as the optimization objective. The network is trained for 1000 epochs. The initial learning rate is set to 0.01, and it is reduced to 0.001 after 750 epochs to improve convergence stability and final classification performance.

The learning-rate settings were empirically determined through repeated preliminary Monte Carlo experiments. Several candidate learning-rate values were evaluated according to the convergence stability and averaged classification performance over multiple independent trials. It was observed that excessively large learning rates could lead to unstable convergence and performance fluctuations, whereas excessively small learning rates slowed down the convergence process and increased training time. Therefore, the adopted learning-rate schedule provided a practical compromise between convergence stability and classification performance.

2.4. Experimental Settings and Evaluation Metrics

To evaluate the proposed method, experiments were conducted on measured radar echoes of humans, vehicles, and UAVs under different velocity intervals. Following the velocity-stratified strategy described in Section 2.3.6, all available samples were first grouped into ultra-low-, low-, and high-velocity intervals according to their measured radial velocities. Within each velocity interval, the samples of each target category were randomly shuffled in each Monte Carlo trial. In each trial, a fixed number of training samples and testing samples were randomly selected from all available samples of the corresponding target category and velocity interval, and the selected training and testing samples did not overlap. The remaining samples were not used in that trial. The detailed numbers of total available samples, training samples, and testing samples for each category and velocity interval are summarized in Table 3. Since the number of human samples in the low-velocity interval was relatively limited, 800 training samples and 100 testing samples were selected for each category in this interval, whereas 800 training samples and 200 testing samples were selected for each category in the ultra-low- and high-velocity intervals. In addition, the training and testing sets were independently regenerated in every Monte Carlo trial, and the final reported results were obtained by averaging the results over all repeated trials.

Table 3. Dataset statistics and sample partition under different velocity intervals.

Velocity	Class	Total Samples	Training Samples	Testing Samples
Ultra-Low	Human	1433	800	200
	Vehicle	1713	800	200
	UAV	1267	800	200
Low	Human	922	800	100
	Vehicle	1658	800	100
	UAV	1431	800	100
High	Vehicle	1334	800	200
	UAV	1011	800	200

To reduce the influence of randomness introduced by dataset partitioning, more than 20 Monte Carlo trials were carried out for each velocity interval. In every trial, the training and test sets were newly generated, and the test samples were not involved in the training process. The final results were obtained by averaging the outcomes of all repeated trials. The same data-partition protocol and feature-extraction procedure were adopted in all ablation, comparison, and robustness experiments to ensure fair and consistent evaluation conditions.

https://doi.org/10.3390/rs18111788

The proposed method was evaluated in terms of clutter suppression, feature separability, classification accuracy, and robustness. For clutter suppression, the signal-to-clutter ratio (SCR) was used as a quantitative indicator. For feature separability, Cohen's d was adopted [28]:

$$d = \frac{\mu_1 - \mu_2}{s_p} \tag{26}$$

where μ_1 and μ_2 denote the sample means of the two target classes, respectively, and s_p is the pooled standard deviation, defined as

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \tag{27}$$

where n_1 and n_2 are the sample sizes of the two classes, and s_1^2 and s_2^2 are the corresponding sample variances. In addition, t-SNE was used to visualize the clustering behavior of different target categories in the fused feature space.

For classification performance, both per-class accuracy and overall accuracy were used:

$$ACC_i = \frac{N_i^{cor}}{N_i^{test}} \tag{28}$$

$$ACC = \frac{\sum_i N_i^{cor}}{\sum_i N_i^{test}} \tag{29}$$

where N_i^{cor} and N_i^{test} denote the numbers of correctly classified and total test samples of the i-th class, respectively, and i is the class index.

3. Results

In this section, the effectiveness of the proposed method is evaluated from four aspects: clutter suppression performance, feature separability, classification results under different velocity intervals, and ablation and comparison experiments. First, the preprocessing performance is analyzed to verify the effectiveness of the proposed clutter suppression strategy. Then, the discriminability of the extracted features is examined. Finally, classification accuracy, feature contribution, and comparative performance are presented to validate the overall effectiveness of the proposed framework.

3.1. Clutter Suppression Performance

To evaluate the effectiveness of the proposed echo preprocessing method, experiments were first conducted to analyze the influence of the wavelet decomposition level on clutter suppression performance. Specifically, the approximation coefficients obtained from 5-level, 6-level, and 7-level wavelet decompositions were used to reconstruct the signal, and the reconstructed results were compared with the original echo. The corresponding results are shown in Figure 7. In the figure, the black dotted line represents the original signal, in which the clutter is mainly concentrated near the zero-Doppler region, while the target component is located in the neighboring Doppler channel.

As shown in the zoomed view of Figure 7d, compared with the 6-level decomposition, the 5-level decomposition causes an additional attenuation of approximately 3.5 dB at the target Doppler bin, indicating partial target-energy loss. In contrast, compared with the 7-level decomposition, the 6-level decomposition reduces the average residual amplitude in the low-frequency clutter-dominated region by approximately 6.0 dB. Therefore, the 6-level decomposition provides a better balance between target-energy preservation and clutter suppression.

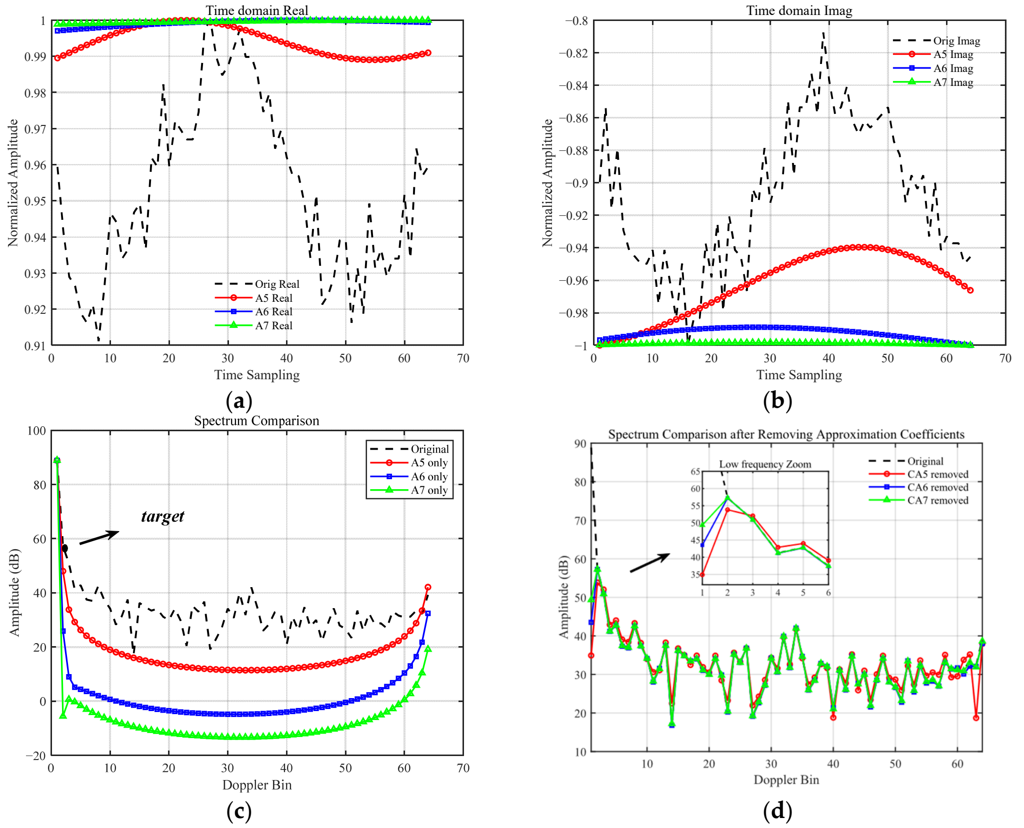


Figure 7. Comparison of decomposition layers. (**a**) Approximate coefficients: The real part of the original signal in the time domain. (**b**) Approximate coefficients: The imaginary part of the original signal in the time domain. (**c**) Approximate coefficients: The Doppler spectrum of the reconstructed signal versus the original signal. (**d**) Reconstructed signals after removing the approximation coefficients (CA5, CA6, and CA7): Comparison of residual Doppler spectra with a low-frequency zoomed view for evaluating target preservation and clutter suppression performance.

To further verify the effectiveness of the proposed method, the wavelet-decomposition-based clutter suppression approach was compared with two commonly used methods, namely singular value decomposition (SVD) and empirical mode decomposition (EMD) [29,30]. The comparison results are shown in Figure 8, and the average signal-to-clutter ratio (SCR) values of multiple frames of slow-target echoes are summarized in Table 4. As shown in Table 4, the proposed wavelet-based method achieves an SCR of 12.92 dB, which is higher than those of the SVD-based method (7.52 dB) and the EMD-based method (5.73 dB).

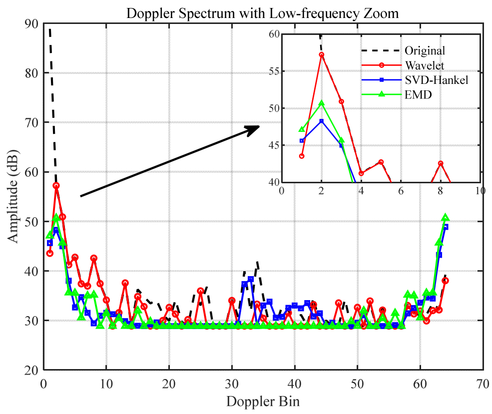


Figure 8. Comparison of the effects of clutter suppression methods.

Table 4. Comparison of clutter suppression methods.

Method	Wavelet	SVD	EMD
SCR(dB)	12.92	7.52	5.73

It should be noted that the comparative performance between the proposed method and the SVD-/EMD-based methods may be influenced by radar configuration, clutter environment, and observation scenario. Therefore, the results presented in this work should not be interpreted as universally optimal conclusions for all radar sensing systems. Nevertheless, the core ideas of the proposed framework, including low-frequency clutter suppression, multi-feature fusion, and velocity-aware classification, are not inherently restricted to a specific radar system and therefore exhibit potential applicability to other radar sensing scenarios.

3.2. Feature Separability Analysis

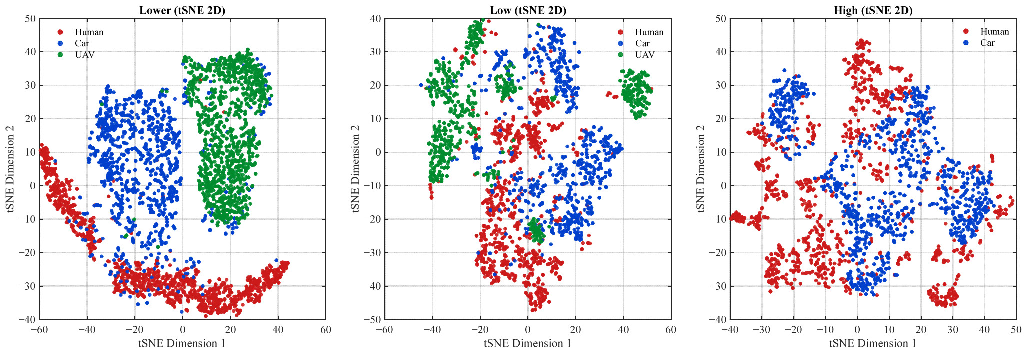
To further evaluate the discriminative capability of the extracted features, feature separability analysis was conducted for different target categories under different velocity intervals. In this study, Cohen's d was adopted as the main quantitative metric to measure the degree of separation between two classes in a given feature dimension [28]. A larger value of Cohen's d indicates stronger separability and better discriminative capability of the corresponding feature. Based on this criterion, representative handcrafted features were analyzed, including relative RCS, spectral entropy, spectral standard deviation, time-domain amplitude dispersion coefficient, and time-domain variance. The results are summarized in Table 5.

Table 5. Cohen's d results among features (Fea1: relative RCS; Fea2: spectral entropy; Fea3: spectral standard deviation; Fea4: amplitude dispersion coefficient; Fea5: time-domain variance).

		Fea1	Fea2	Fea3	Fea4	Fea5
U-Low	**Human–Car**	1.29	0.05	0.06	0.55	1.63
	Human–UAV	2.25	0.11	0.82	0.92	0.21
	Car–UAV	0.79	0.16	0.73	1.25	1.25
Low	**Human–Car**	0.83	0.87	0.78	0.39	1.58
	Human–UAV	2.90	3.48	2.52	2.02	1.49
	Car–UAV	2.09	1.68	1.97	1.84	0.14
High	**Car–UAV**	2.05	0.38	1.55	1.62	0.16

As shown in Table 5, the separability of individual features varies with both target category and velocity interval. In the ultra-low-velocity interval, some features show limited discrimination between certain class pairs, whereas others still maintain relatively strong separability. In the low-velocity interval, the differences between target classes become more evident for several representative features, especially for the human–UAV and vehicle–UAV pairs. In the high-velocity interval, although only vehicle and UAV targets are involved, some features still provide clear discrimination, while others contribute less. These results indicate that no single feature can provide uniformly strong discrimination for all target pairs and all velocity intervals, which justifies the need for feature fusion.

To provide a more intuitive illustration of feature separability, t-SNE was further employed to visualize the clustering behavior of different target categories in the fused feature space. The visualization results for the ultra-low-, low-, and high-velocity intervals are shown in Figure 9, respectively. It can be observed that the three target categories form distinguishable clusters in the fused feature space, although the degree of overlap differs across velocity intervals. In particular, the clustering effect in the high-velocity interval is more evident, which is consistent with the higher classification accuracy reported in the subsequent section. Overall, the Cohen's d analysis and t-SNE visualization together demonstrate that the extracted features have meaningful discriminative capability, while also confirming the importance of multi-feature fusion for robust target classification.



(**a**) (**b**) (**c**)

Figure 9. (**a**) Distribution map of t-SNE characteristics in the ultra-low velocity interval. (**b**) Distribution map of t-SNE characteristics in the low velocity interval. (**c**) Distribution map of t-SNE characteristics in the high velocity interval.

It should be noted that Cohen's d and t-SNE visualization are used in this work as auxiliary tools for feature-separability analysis rather than definitive criteria for evaluating classification effectiveness. Since t-SNE visualization may be influenced by parameter settings and dimensionality-reduction characteristics, the feature-analysis results were further validated through confusion matrices, classification performance, ablation experiments, and robustness analysis under different SNR conditions.

3.3. Classification Results in Different Velocity Intervals

To evaluate the classification performance of the proposed method under different motion conditions, the experimental results are reported separately for the ultra-low-, low-, and high-velocity intervals. The overall accuracies of the three velocity intervals are summarized in Table 6, while the detailed classification results are given in Tables 7–9. As shown in Table 6, the proposed method achieves overall accuracies of 95.78%, 96.87%, and 98.16% in the ultra-low-velocity, low-velocity, and high-velocity intervals, respectively. These results indicate that the proposed framework maintains high classification performance over different motion regimes.

Table 6. Overall classification accuracy in different velocity intervals.

Velocity Interval	**Overall ACC (%)**
Ultra-low velocity	95.78
Low velocity	96.87
High velocity	98.16

Table 7. Classification results in the ultra-low-velocity interval. (average confusion matrix entries over Monte Carlo trials).

		Prediction			**ACC**	**Overall ACC**
		Human	**Vehicle**	**UAV**		
Real	**Human**	191.8	7.4	0.8	95.92%	95.78%
	Vehicle	6.3	189.1	4.6	94.57%	
	UAV	1.5	4.8	193.7	96.85%	

Table 8. Classification results in the low-velocity interval (average confusion matrix entries over Monte Carlo trials).

		Prediction			ACC	Overall ACC
		Human	Vehicle	UAV		
Real	Human	97	2.7	0.3	97.00%	96.87%
	Vehicle	3.9	95.4	0.7	95.36%	
	UAV	0.5	1.2	98.3	98.25%	

Table 9. Classification results in the high-velocity interval (average confusion matrix entries over Monte Carlo trials).

		Prediction		ACC	Overall ACC
		Vehicle	UAV		
Real	Vehicle	195.5	4.5	97.75%	98.16%
	UAV	2.9	197.1	98.57%	

For the ultra-low-velocity interval, the classification results are presented in Table 7. In this interval, the method achieves classification accuracies of 95.92% for humans, 94.57% for vehicles, and 96.85% for UAVs, with an overall accuracy of 95.78%. Although all three target categories can still be effectively distinguished, the overall performance in this interval is slightly lower than that in the other intervals. This is mainly because ultra-low-velocity targets are more susceptible to clutter interference, and the Doppler differences among target categories are less pronounced under this condition.

For the low-velocity interval, the detailed results are listed in Table 8. The proposed method achieves classification accuracies of 97.00% for humans, 95.36% for vehicles, and 98.25% for UAVs, with an overall accuracy of 96.87%. Compared with the ultra-low-velocity interval, the classification performance is further improved, which indicates that the extracted features become more discriminative when the target motion characteristics are more evident and the overlap with near-zero-Doppler clutter is relatively reduced.

For the high-velocity interval, the classification results are shown in Table 9. Since only vehicles and UAVs are considered in this interval, the classification task becomes simpler than that in the other two intervals. In addition, targets in the high-velocity interval generally exhibit more distinct Doppler characteristics and reduced low-frequency clutter overlap, which further decreases classification difficulty. As a result, the proposed method achieves accuracies of 97.75% for vehicles and 98.57% for UAVs, with an overall accuracy of 98.16%, which is the highest among the three velocity intervals. Nevertheless, the subsequent results of the merged low-velocity interval experiment further indicate that the proposed velocity-stratified classification strategy itself also contributes positively to recognition performance, rather than the improvement arising solely from reduced class complexity.

Overall, the classification results in different velocity intervals show that the proposed method can effectively distinguish humans, vehicles, and UAVs under measured radar data. In particular, the superior performance in the high-velocity interval and the stable results in the low- and ultra-low-velocity intervals demonstrate the effectiveness of combining velocity interval division with multi-feature fusion for robust target classification in cluttered radar sensing environments.

To further evaluate the classification performance, the precision, recall, and F1-score results for different target categories and velocity intervals are summarized in Table 10. The results show that most target categories achieve consistently high precision and recall values above 93%, while the corresponding F1-scores also remain high across different velocity intervals. In particular, UAV targets exhibit relatively stable recognition performance in all velocity intervals, whereas the ultra-low-velocity interval remains more challenging due to stronger clutter influence and spectral overlap. These results further confirm the effectiveness and robustness of the proposed velocity-aware classification framework.

Table 10. Precision, recall, and F1-score results for different target categories under different velocity intervals.

Velocity	Class	Precision (%)	Recall (%)	F1-Score (%)
Ultra-Low	**Human**	96.09	95.92	96.01
	Vehicle	93.95	94.57	94.26
	UAV	97.32	96.85	97.08
Low	**Human**	95.59	97.00	96.29
	Vehicle	96.05	95.36	95.70
	UAV	98.99	98.25	98.62
High	**Vehicle**	98.56	97.75	98.15
	UAV	97.77	98.57	98.17

3.4. Ablation and Comparison Experiments

To further evaluate the contribution of different feature groups and the effectiveness of the overall framework, ablation and comparison experiments were conducted in this study. The corresponding results include module-level ablation analysis of the preprocessing stage and velocity-stratified classification mechanism, feature-ablation results in different velocity intervals, validation of the velocity interval division strategy, comparisons with different classifiers, comparisons with representative methods from the literature, and robustness analysis under different SNR conditions.

First, module-level and feature-level ablation experiments were performed to analyze the contribution of different components to classification performance. The module-level ablation results are summarized in Table 11. The results show that both the preprocessing stage and the velocity-stratified classification mechanism contribute positively to the final classification performance, demonstrating the effectiveness of the proposed framework design.

Table 11. Ablation analysis of preprocessing and velocity stratification.

Scenario	Preprocessing	Velocity Stratification	ACC
Full framework	Yes	Yes	96.94%
Without preprocessing	No	Yes	88.15%
Without velocity stratification	Yes	No	90.83%

The corresponding results for the ultra-low-, low-, and high-velocity intervals are summarized in Tables 12–14, respectively. In these experiments, Scenario A denotes the combination of time-domain features and wavelet packet energy features, Scenario B denotes the combination of frequency-domain features and wavelet packet energy features, and Scenario C denotes the combination of frequency-domain features and time-domain features. As shown in Tables 12–14, removing any feature group leads to a degradation in overall performance, although the degree of degradation differs across velocity intervals. In particular, the results indicate that the fused feature representation consistently outperforms partial feature combinations, which confirms that the different feature groups provide complementary information for target classification.

Table 12. Ablation results in the ultra-low-velocity interval.

Type	Scenario A	Scenario B	Scenario C
Human	81.38%	86.33%	77.74%
Vehicle	82.03%	94.15%	86.27%
UAV	93.36%	95.87%	91.16%
ACC	85.59%	92.12%	85.06%

Table 13. Ablation results in the low-velocity interval.

Type	Scenario A	Scenario B	Scenario C
Human	76.28%	92.63%	77.40%
Vehicle	71.45%	94.81%	70.57%
UAV	85.42%	96.36%	83.26%
ACC	77.72%	94.60%	77.08%

Table 14. Ablation results in the high-velocity interval.

Type	Scenario A	Scenario B	Scenario C
Vehicle	82.83%	90.91%	83.57%
UAV	95.87%	92.36%	91.14%
ACC	89.35%	91.64%	87.36%

To further verify the effectiveness of the proposed velocity-stratified classification strategy, an additional experiment was conducted by merging the ultra-low-velocity and low-velocity intervals into a single interval. The corresponding results are listed in Table 15. As shown in the table, the merged setting achieves an overall accuracy of 95.63%, which is lower than the accuracies obtained when the ultra-low-velocity interval and the low-velocity interval are treated separately. This result demonstrates that dividing low-speed samples into finer velocity intervals is beneficial for improving classification performance, especially in clutter-affected scenarios.

Table 15. Comparison results under the merged low-velocity setting.

Merging	95.63%
Ultra-Low velocity interval	95.78%
Low velocity interval	96.87%

Comparative experiments with several commonly used classifiers were also carried out to evaluate the suitability of the proposed LSTM-based framework. The compared classifiers include 1D-CNN, BP neural network, support vector machine (SVM), decision tree (DT), and K-nearest neighbors (KNN). The comparison results are shown in Figure 10. As observed in the figure, the proposed LSTM-based method achieves the best average classification performance among the compared classifiers, outperforming the second-best BP classifier by 0.56%. This result indicates that the LSTM network is effective in modeling the relationships among the fused features for the considered radar target classification task.

In addition, the proposed method was compared with several representative methods reported in the literature, including end-to-end deep learning approaches and a micro-Doppler-based method. Since the datasets and signal formats used in the original studies differ from the measured radar dataset adopted in this work, the competing methods listed in Table 16 were re-implemented according to the main architectures and methodological descriptions reported in the corresponding references and adapted to the input format and data characteristics of this study.

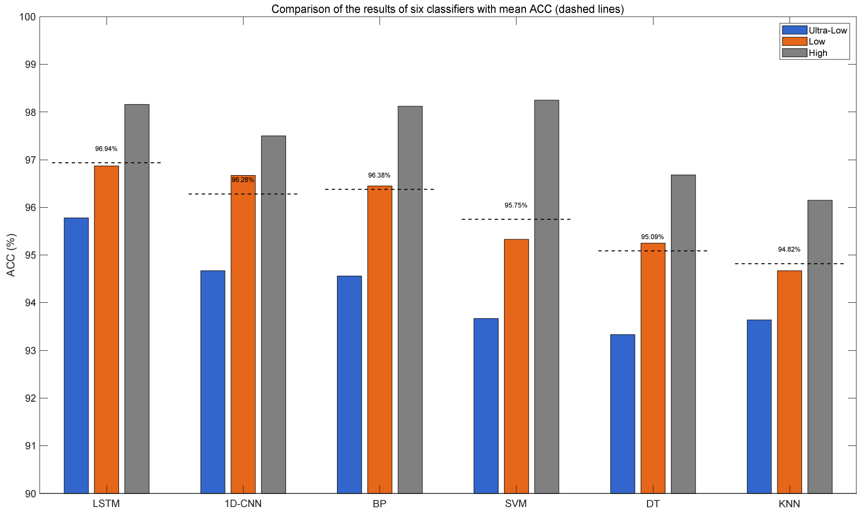


Figure 10. Performance comparison of six classifiers.

Table 16. The method proposed in this paper is compared with those in other papers.

Method	Feature/Input Format	Network Structure	Training Samples	ACC
Proposed	Proposed velocity-aware multi-feature fusion	LSTM classifier	Same Monte Carlo partition protocol	96.94%
Transformer	64-dimensional feature vector reshaped into 8 × 8 tokens	Positional encoding + multi-head self-attention + FFN + mean pooling + FC + Softmax	4 heads, $d_{model} = 8$, $d_{ff} = 32$, Adam, cosine learning-rate decay	87.50%
1D-CNN [18]	64-dimensional Doppler feature sequence	Conv1D + MaxPool + Conv1D + MaxPool + FC + Softmax	Kernels: 3 and 7, Adam, dropout	94.34%
2D-CNN [31]	64-dimensional feature vector transformed into 32 × 32 Toeplitz matrix	Three Conv2D + MaxPool blocks + FC + Softmax	Kernels: 7 × 7, 3 × 3, 3 × 3; Adam, dropout	84.92%
GRU-based [32]	Dynamic feature sequence adapted from the micro-Doppler feature-extraction idea in [32]	Two stacked GRU layers + FC + Softmax	64 hidden units, dropout, Adam, learning-rate decay	90.26%

For a fair comparison, all methods employed the same preprocessing procedure, dataset partition protocol, and Monte Carlo evaluation strategy. Specifically, the training and test sets were randomly regenerated in each Monte Carlo trial, and the final results were obtained by averaging the repeated experimental outcomes. The comparison results are summarized in Table 16. As shown in the table, the proposed method achieves an overall accuracy of 96.94%, which is higher than those of the compared Transformer-based method, the method in [18], the method in [31], and the method in [32]. These results suggest that, on the measured radar dataset used in this study, the proposed multi-feature fusion framework provides a more effective balance between discriminative capability and robustness than the compared methods.

Finally, the robustness of the proposed method was examined under different SNR conditions. The corresponding variation in recognition rate with SNR is shown in Figure 11. As the SNR decreases, the classification performance gradually degrades, but the proposed method still maintains a relatively stable recognition trend over a range of SNR conditions. This result further demonstrates that the proposed framework has good robustness to noise interference and is suitable for practical radar sensing scenarios.

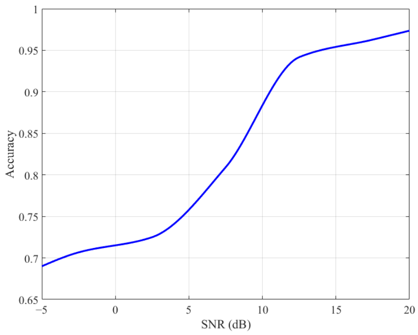


Figure 11. The relationship curve between recognition rate and SNR.

3.5. Embedded Implementation and Runtime Analysis

To evaluate the practical deployability of the proposed method, the core algorithm was further implemented on a TMS320C6678-based embedded platform (Texas Instruments, Dallas, TX, USA). Compared with CPU platforms, embedded DSP architectures are more suitable for radar signal-processing tasks with repetitive numerical operations and strict real-time requirements. In particular, DSP platforms provide hardware optimization for parallel multiply–accumulate operations, deterministic real-time execution, and multi-core embedded processing, which are important for radar echo preprocessing, clutter suppression, feature extraction, and target classification. In addition, embedded DSP platforms generally offer lower power consumption and easier deployment for practical radar sensing systems.

For the wavelet decomposition module, the real and imaginary parts of the complex echo were processed separately, and the computation of multiple wavelet coefficients was organized in parallel within each iteration. In addition, the decomposition procedure was implemented with a software-pipelined structure to improve the utilization of the DSP functional units and reduce repeated memory access overhead. A similar idea was also adopted in the reconstruction stage, where the filter operations were simplified according to the coefficient structure. These optimizations significantly reduced the computational burden of the preprocessing stage.

For the nonlinear activation functions in the LSTM network, a lookup-table-based implementation was adopted to replace direct function evaluation, thereby reducing runtime overhead. The execution time before and after optimization is summarized in Table 17. The total runtime was reduced from 1380 μs to 94 μs, corresponding to an overall improvement of 93.2%, which indicates that the proposed framework can satisfy the real-time processing requirements of practical embedded radar sensing applications.

Table 17. Comparison of Execution Time Before and After Optimization.

	Primitive (μs)	**Optimized (μs)**	**Enhanced**
Data Preprocessing	596	31.3	94.7%
Feature Extraction	723	42.5	94.1%
Classification	61	20.2	66.9%
ALL	1380	94	93.2%

4. Discussion

The experimental results demonstrate that the proposed velocity-aware multi-feature fusion framework is effective for classifying ground and low-altitude targets in cluttered radar sensing environments. First, the preprocessing results show that clutter suppression is essential for improving the separability of low-velocity target echoes. By suppressing low-frequency clutter and reducing noise, the wavelet-decomposition-based preprocessing method improves signal quality and provides a more reliable basis for subsequent feature extraction and classification.

Second, the feature analysis and ablation results indicate that the adopted feature groups provide complementary information. Wavelet packet energy features describe energy distribution, time-domain features characterize fluctuation and relative scattering strength, and frequency-domain features reflect spectral concentration and dispersion. Since no single feature can maintain strong discrimination for all target pairs and all velocity intervals, feature fusion is necessary for robust classification.

Third, the classification results confirm the importance of velocity interval division. The high-velocity interval achieves the best performance, while the ultra-low-velocity interval is relatively more challenging because of stronger clutter influence and weaker motion separability. The additional experiment with the merged low-speed setting further shows that finer velocity partitioning is beneficial for improving classification performance.

The current study is limited by the scene diversity of the measured dataset and the number of target categories considered. Future work will focus on expanding the dataset and further validating the proposed framework under more diverse radar systems and observation conditions.

5. Conclusions

This paper proposed a velocity-aware multi-feature fusion method for ground and low-altitude target classification using measured X-band pulse-Doppler radar echoes. By combining wavelet-decomposition-based preprocessing, multi-domain feature extraction, feature fusion, and velocity-stratified classification, the proposed framework effectively improves target discrimination under cluttered outdoor conditions. Experimental results on measured human, vehicle, and UAV radar echoes demonstrated that the method achieved high classification accuracy in different velocity intervals and showed good effectiveness and robustness in practical radar sensing scenarios.

Although the present experiments were conducted using measured X-band pulse-Doppler radar echoes, the core idea of the proposed framework is not limited to this specific radar system. In essence, the proposed method aims to suppress dominant clutter components, extract complementary target-related features, and perform velocity-aware classification according to the Doppler-distribution characteristics of the observed targets. This processing philosophy may also be extended to other radar sensing systems where weak target signatures are affected by strong clutter. For example, in high-frequency surface wave radar (HFSWR), maritime target detection is often limited by strong sea clutter and Bragg-line interference in the range-Doppler domain. Recent studies on high-resolution HFSWR processing have shown that improving range-Doppler representation and suppressing clutter-related interference are important for enhancing ship detectability [33–35]. Therefore, the proposed clutter-aware feature extraction and velocity-aware classification strategy has potential applicability to other radar remote sensing scenarios after appropriate parameter adjustment and system-specific calibration.

In future work, the proposed framework will be further extended in the following directions. First, small-sample learning and data augmentation strategies will be investigated to improve the classification performance when measured radar samples are limited.

https://doi.org/10.3390/rs18111788

Second, domain adaptation and transfer learning will be explored to enhance the generalization capability of the model across different radar systems, operating bands, deployment environments, and target observation conditions. Third, adaptive parameter selection will be studied for the preprocessing and velocity-stratification modules, so that the framework can be more flexibly applied to complex and time-varying clutter environments. Finally, further validation on additional measured datasets and different radar platforms will be conducted to evaluate the broader applicability of the proposed method.

Author Contributions: Conceptualization, P.H. and L.T.; methodology, P.H. and M.Z.; software, P.H.; validation, M.Z., L.T. and Z.Z.; formal analysis, P.H.; investigation, P.H.; resources, L.T.; data curation, P.H. and M.Z.; writing—original draft preparation, P.H.; writing—review and editing, L.T., M.Z. and Z.Z.; visualization, P.H.; supervision, L.T.; project administration, L.T.; funding acquisition, L.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to confidentiality restrictions associated with the project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rizik, A.; Tavanti, E.; Vio, R.; Delucchi, A.; Chible, H.; Randazzo, A.; Caviglia, D.D. Single target recognition using a low-cost FMCW radar based on spectrum analysis. In Proceedings of the 2020 27th IEEE International Conference on Electronics, Circuits and Systems (ICECS), Virtual Conference, Glasgow, UK, 23–25 November 2020; pp. 23–25.
2. Clemente, C.; Pallotta, L.; Gaglione, D.; De Maio, A.; Soraghan, J.J. Automatic target recognition of military vehicles with Krawtchouk moments. *IEEE Trans. Aerosp. Electron. Syst.* **2017**, *53*, 493–500. [CrossRef]
3. Abdu, F.J.; Zhang, Y.; Fu, M.; Li, Y.; Deng, Z. Application of deep learning on millimeter-wave radar signals: A review. *Sensors* **2021**, *21*, 1951. [CrossRef]
4. Jiang, W.; Wang, Y.; Li, Y.; Lin, Y.; Shen, W. Radar target characterization and deep learning in radar automatic target recognition: A review. *Remote Sens.* **2023**, *15*, 3742. [CrossRef]
5. Zhang, M.; Tian, L.; Sun, B.; Hu, P. A classification method for ground and low altitude moving targets based on multiple features. In Proceedings of the 2024 4th International Conference on Neural Networks, Information and Communication (NNICE), Guangzhou, China, 19–21 January 2024; pp. 677–684.
6. Ye, L.; Hu, S.; Yan, T.; Meng, X.; Zhu, M.; Xu, R. Radar target shape recognition using a gated recurrent unit based on RCS time series' statistical features by sliding window segmentation. *IET Radar Sonar Navig.* **2021**, *15*, 1715–1726. [CrossRef]
7. Wang, T.; Bi, W.; Zhao, Y.; Xue, W. Radar target recognition algorithm based on RCS observation sequence—Set-valued identification method. *J. Syst. Sci. Complex.* **2016**, *29*, 573–588. [CrossRef]
8. Han, L.; Feng, C.; Hu, X.; He, S.; Xu, X. Ballistic target recognition based on multiple data representations and deep-learning algorithms. *Chin. J. Aeronaut.* **2024**, *37*, 167–181. [CrossRef]
9. Potter, M.; Akcakaya, M.; Necsoiu, M.; Schirner, G.; Erdoğmuş, D.; Imbiriba, T. Multistatic radar RCS-signature recognition of aerial vehicles: A Bayesian fusion approach. *IEEE Trans. Aerosp. Electron. Syst.* **2025**, *61*, 151–161. [CrossRef]
10. Chen, V.; Li, F.; Ho, S.-S.; Wechsler, H. Micro-Doppler effect in radar: Phenomenon, model, and simulation study. *IEEE Trans. Aerosp. Electron. Syst.* **2006**, *42*, 2–21. [CrossRef]
11. Du, L.; Li, L.; Wang, B.; Xiao, J. Micro-Doppler feature extraction based on time-frequency spectrogram for ground moving targets classification with low-resolution radar. *IEEE Sens. J.* **2016**, *16*, 3756–3763. [CrossRef]
12. Vishwakarma, S.; Li, W.; Tang, C.; Woodbridge, K.; Adve, R.R.; Chetty, K. Attention-enhanced AlexNet for improved radar micro-Doppler signature classification. *IET Radar Sonar Navig.* **2016**, *17*, 652–664. [CrossRef]
13. Jiang, X.; Lai, Y.; Wang, S.; Song, Y. Using polarization-diverse wave-coefficients for aerospace target recognition. *IEEE Access* **2020**, *8*, 143033–143041. [CrossRef]
14. Ma, M.; Liu, K.; Luo, X.; Zhang, T. Multi-view polarization HRRP target recognition based on convolutional neural network. *Lect. Notes Inst. Comput. Sci. Soc.-Inform. Telecommun. Eng.* **2021**, *346*, 710–720.
15. Zhou, F.; Li, J.; Wang, Y. An improved CNN-LSTM network for modulation identification relying on periodic features of signal. *IET Commun.* **2023**, *17*, 2097–2106. [CrossRef]
16. Ma, J.; Yuan, Y. Dimension reduction of image deep feature using PCA. *J. Vis. Commun. Image Represent.* **2019**, *63*, 102578. [CrossRef]

17. Wang, C.; Liu, X.; Pei, J.; Huang, Y.; Zhang, Y.; Yang, J. Multiview attention CNN-LSTM network for SAR automatic target recognition. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2021**, *14*, 12504–12513. [CrossRef]
18. Xie, R.; Dong, B.; Li, P.; Rui, Y.; Wang, X.; Wei, J. Automatic target recognition method for low-resolution ground surveillance radar based on 1D-CNN. In Proceedings of the SPIE 11719, Twelfth International Conference on Signal Processing Systems, Shanghai, China, 20 January 2021; p. 1171907.
19. Richards, M.A. *Fundamentals of Radar Signal Processing*, 2nd ed.; McGraw-Hill Education: New York, NY, USA, 2014.
20. Zhu, L.; Zhang, S.; Ma, Q.; Zhao, H.; Chen, S.; Wei, D. Classification of UAV-to-ground targets based on enhanced micro-Doppler features extracted via PCA and compressed sensing. *IEEE Sens. J.* **2020**, *20*, 14360–14368. [CrossRef]
21. Daubechies, I. The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inf. Theory* **1990**, *36*, 961–1005. [CrossRef]
22. Ellonen, I.; Kaarna, A. Rain clutter filtering from radar data with discrete wavelet transform. In Proceedings of the 2006 International Radar Symposium, Krakow, Poland, 24–26 May 2006; pp. 1–4.
23. Mallat, S.; Peyré, G. *A Wavelet Tour of Signal Processing*; Elsevier: Amsterdam, The Netherlands; Academic Press: Amsterdam, The Netherlands, 2009.
24. Allabakash, S.; Yasodha, P.; Reddy, S.; Srinivasulu, P. Wavelet transform-based methods for removal of ground clutter and denoising the radar wind profiler data. *IET Signal Process.* **2015**, *9*, 440–448. [CrossRef]
25. Yan, R.; Gao, R.X.; Chen, X. Wavelets for fault diagnosis of rotary machines: A review with applications. *Signal Process.* **2014**, *96*, 1–15. [CrossRef]
26. Singh, D.; Singh, B. Investigating the impact of data normalization on classification performance. *Appl. Soft Comput.* **2020**, *97*, 105524. [CrossRef]
27. Hochreiter, S. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]
28. Lachenbruch, P.A. *Statistical Power Analysis for the Behavioral Sciences*; Routledge: New York, NY, USA, 1989.
29. Cheng, Q.; Wu, X.; Zhang, X.; Yang, Q. A novel sea clutter suppression method based on SVD-FRFT at low signal-to-clutter ratio. *Electron. Lett.* **2023**, *59*, e12874. [CrossRef]
30. Lv, M.; Zhou, C. Study on sea clutter suppression methods based on a realistic radar dataset. *Remote Sens.* **2019**, *11*, 2721. [CrossRef]
31. Li, Y.; Rui, Y.; Gao, Y.; Gao, J. Ground surveillance radar target classification based on 2D CNN. In Proceedings of the SPIE 11719, Twelfth International Conference on Signal Processing Systems, Shanghai, China, 20 January 2021; p. 117190N.
32. Wu, Z.; Liu, H.; Ma, C.; Liu, Z. Aerial target recognition with enhanced micro-Doppler dynamic features based on frequency-modulated continuous-wave radar. *IEEE Sens. J.* **2023**, *23*, 23119–23132. [CrossRef]
33. Golubović, D.; Erić, M.; Vukmirović, N. High-Resolution Doppler and Azimuth Estimation and Target Detection in HFSWR: Experimental Study. *Sensors* **2022**, *22*, 3558. [CrossRef]
34. Golubović, D.; Erić, M.; Vukmirović, N.; Orlić, V. High-Resolution Sea Surface Target Detection Using Bi-Frequency High-Frequency Surface Wave Radar. *Remote Sens.* **2024**, *16*, 3476. [CrossRef]
35. Wang, T.; Zhang, L.; Li, G. Shipborne HFSWR Target Detection in Clutter Regions Based on Multi-Frame TFI Correlation. *Remote Sens.* **2022**, *14*, 4192. [CrossRef]