Comparative Analysis of Deep Learning Methods for Unmanned Aerial Vehicles (UAVs) Recognition and Identification Using Micro-Doppler Signatures

Ala Agrebi
Royal Military Academy
Brussels, Belgium

Xavier Neyt
Royal Military Academy
Brussels, Belgium

***Abstract*—In this work, we introduce a novel methodology for small Unmanned Aerial Vehicles (UAVs) classification by leveraging their specific micro-Doppler signatures (mDs). The proposed approach involves the direct application of Recurrent Neural Network (RNN) techniques to temporal radar signals. Hence, we have constructed neural network architectures incorporating Gated Recurrent Unit (GRU) layers, achieving classification accuracies of 100%. To comprehensively evaluate our methodology, we compare our findings with two alternative approaches for UAVs classification: one utilizing Convolutional Neural Networks (CNN)-based networks where mDs are represented as spectrograms transformed into images, and another employing RNN-based methods applied to spectrograms of the mDs.**

***Index Terms*—Unmanned Aerial Vehicles (UAVs), Classification, Radar, micro-Doppler signatures (mDs), Gated Recurrent Unit (GRU).**

I. Introduction

Recent conflicts underscore the growing importance of Unmanned Aerial Vehicles (UAVs) on the battlefield and highlight the need to develop countermeasures requiring identification and recognition to effectively monitor and manage UAVs activities. A primary challenge in UAVs recognition lies in distinguishing them from flying birds, especially at lower elevations, where they share similar Radar Cross-Section (RCS) and velocity ranges. The identification of birds and UAVs has been extensively researched in the literature [1], [2], with various methods, including radar, visual, acoustic, and radio-frequency sensing systems, applied for UAVs classification. Radar technology offers advantages such as long-range surveillance capabilities, unaffected by lighting conditions, and robustness against weather conditions, making it an attractive option. Radar-based approaches rely mainly on the analysis of micro-Doppler signatures (mDs), which generally refer to Doppler shifts resulting from the motion of object components [3]. In the case of UAVs, mDs specifically include Doppler frequency modulations caused by mechanical vibrations or rotations of UAVs parts, including distinctive patterns such as the rotary motion of propeller blades, as well as the oscillations and vibrations of the UAVs structure [4]. This rotations and oscillations induce periodic variations in the amplitude and frequency of the radar signal, appearing as a periodic pattern superimposed on the main radar return signal. These can be signatures of UAVs that serve as unique identifiers, offering insights into the target's identity through its motion characteristics. These characteristics can vary based on parameters such as rotation speed, number of propellers, number of blades per propeller, blade length, and blade construction materials [5].

Existing research predominantly focus on addressing the challenge of radar-based UAVs classification by converting mDs representations—such as spectrograms, cepstrogams, and Cadence Velocity Diagrams (CVD) into image formats suitable for use in CNN based methods [6]–[10]. Unlike conventional images, which are characterized solely by spatial dimensions, the mDs of UAVs possess both temporal and frequency dimensions, marking a notable distinction from typical image data. It is imperative to acknowledge and accommodate this inherent specificity during the classification process. Considering this, Recurrent Neural Network (RNN) appear as a promising approach for analyzing mDs due to their sequential nature and ability to effectively capture temporal relationships. Different from feed-forward neural networks, RNN have recurrent connections between hidden states with a time delay. This characteristic allows RNN to model temporal correlations between events in input sequences through their ability to retain a form of 'memory', incorporating information from previous inputs to influence the outputs. This capability sets RNN apart from traditional neural networks, which treat inputs as independent entities, by allowing them to consider the sequence context in their processing.

Utilizing the benefits of RNN, we implemented a Gated Recurrent Unit (GRU) based neural network to develop an innovative method for classifying UAVs and birds based on radar signals. GRU networks, as a type of RNN, are capable at capturing temporal dependencies. To evaluate the efficacy of the proposed method, we conducted a comparative against two alternative methods. The first alternative uses a CNN-based approach transforming Doppler signals into spectrograms and then into images, as described in a separate study [12]. In contrast, the second method employs a RNN-based method trained on mDs representations, such as spectrograms, drawing from previous research [11].

The remainder of this paper is organized as follows: Section II offers a detailed explanation of the proposed method, along with the two other methods under comparison, while Section III delves into its implementation. Section IV showcases the classification results, and Section V presents the conclusions drawn from the work and and outlines potential future research directions.

II. Compared Methods

In this section, we will elaborate on the proposed method, referred to as the temporal signals-based method, along with the two alternative methods compared in this work: the spectrograms-based method and the CNN-based method for classifying UAVs and birds using mDs.

A. CNN-based Method

Much of the research in UAVs classification has predominantly focused on radar-based methods, involving transforming mDs into images suitable for CNN-based approaches. Hence, we opted to compare our approach with one of these established methods. CNN excel at processing image and video data by effectively extracting spatial features due to their convolution architecture, when the data shares the same dimensional characteristics. However, their capability to capture distinct dimensions, such as temporal and frequency features, may be limited, particularly with radar signals of UAVs, compared to RNN based methods.

For comparison, we selected the approach proposed by [12], as it not only presents a UAVs classification method based on CNN but also introduces an open-source dataset for radar signals from various types of UAVs, known as DIAT-μSAT, which will be detailed in Subsection III-A. This was motivated by the opportunity to directly compare the results of our method with those obtained using the same dataset, ensuring a fair and meaningful comparison. The approach outlined in [12] employs pre-trained VGG16 and VGG19 CNN models, primarily utilizing transfer learning, to classify small UAVs using spectrograms transformed into images.

B. Spectrograms-based Method

This method is inspired by the work of [11], which utilized an RNN-based approach to classify UAVs using mDs. In that study, spectrograms derived from radar signals were utilized as input data for the RNN. The spectrograms provided a time-frequency representation, where each time window of the spectrogram was treated as a distinct time step in the RNN. The spectrum samples points within each time window were used as feature vectors, allowing the RNN to capture the temporal and spectral characteristics of the UAVs' motion patterns.

While this method demonstrates promising performance, it may not be the ideal approach for UAVs classification applications because spectrograms rely only on the magnitude of the Fourier transform, ignoring the phase information of the radar signal. However, this phase component contains significant information about the behavior of UAVs flights and movements, which can complement the magnitude information and enhance classification accuracy [13].

The decision to compare the proposed method with this alternative approach was driven by the desire to determine whether the use of spectrograms, commonly employed in most research in this context, truly enhances the ability of RNN-based approaches to classify and recognize UAVs. Specifically, we aim to assess whether the time-frequency representation of radar signals can help RNN better capture temporal dependencies and relevant features that might be less apparent in temporal radar signal. We drew inspiration from [11] due to the limited availability of research utilizing spectrograms with RNN-based methods in this domain, making it one of the few studies to adopt such an approach.

C. Temporal Signals-based Method

In this subsection, we detail the proposed method for UAVs classification using mDs, as depicted in Fig.1. At the core of our approach lies the direct utilization of a GRU-based network, a type of RNN, to process the temporal radar signals. Unlike other methods that involve transforming or modifying the signal, our approach retains the raw temporal data, allowing the GRU network to capture intricate temporal features inherent to the mDs of radar signal.

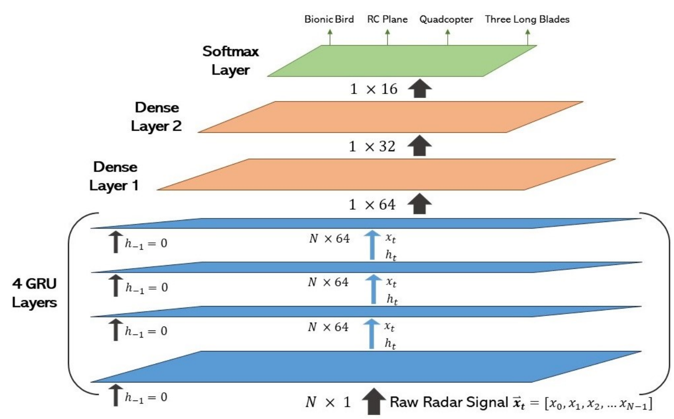


Fig. 1: Proposed method for UAVs classification.

Proposed in [14], the Gated Recurrent Unit (GRU) is a type of RNN that features two gating mechanisms—the reset gate and the update gate. These gates control the flow of information, enabling the GRU to retain important temporal dependencies over extended periods. A graphical representation of the structure of a GRU cell is shown in Fig. 2.

The reset gate $\boldsymbol{r_t}$ determines how much of the previous information should be discarded or reset when processing the current input. On the other hand, the update gate $\boldsymbol{z_t}$ controls the extent to which the current time step incorporates information from the previous memory state. Mathematically, for a given time step t, with $\boldsymbol{x_t}$ as input vector, and $\boldsymbol{h_{t-1}}$ as the hidden state vector from the previous time step, the two vectors, reset gate $\boldsymbol{r_t}$ and update gate $\boldsymbol{z_t}$ are computed as follows [14].

$$\begin{aligned} \boldsymbol{r_t} &= \sigma(W_r \boldsymbol{x_t} + U_r \boldsymbol{h_{t-1}} + \boldsymbol{b_r}) \\ \boldsymbol{z_t} &= \sigma(W_z \boldsymbol{x_t} + U_z \boldsymbol{h_{t-1}} + \boldsymbol{b_z}) \end{aligned} \tag{1}$$

Where σ is the logistic sigmoid function, W_r, W_z, U_r, and U_z are weight matrices and $\boldsymbol{b_r}, \boldsymbol{b_z}$ are bias vectors. The activation

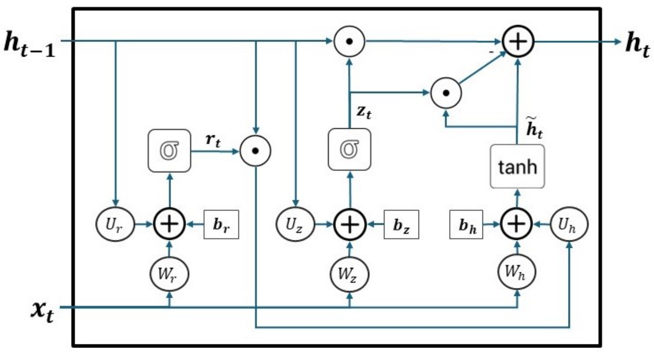


Fig. 2: GRU cell architecture

functions for both gates is sigmoid function, which outputs values between 0 and 1, regulating the flow of information. Additionally, $\tilde{\boldsymbol{h}}_{\boldsymbol{t}}$ denotes the candidate hidden state vector for each time step t, calculated using the reset gate $\boldsymbol{r_t}$ as follows [14].

$$\tilde{\boldsymbol{h}}_{\boldsymbol{t}} = tanh(W_h \boldsymbol{x_t} + U_h(\boldsymbol{r_t} \odot \boldsymbol{h_{t-1}}) + \boldsymbol{b_h}) \tag{2}$$

Where $\odot$ represents Hadamard (element-wise) multiplication, W_h and U_h are weight matrices and $\boldsymbol{b_h}$ is a bias vector. The candidate hidden state $\tilde{\boldsymbol{h}}_{\boldsymbol{t}}$ controls the influence of the previous states through the Hadamard multiplication of the reset gate $\boldsymbol{r_t}$ and the previous hidden state $\boldsymbol{h_{t-1}}$. When the reset gate values are close to 0, the candidate hidden state $\tilde{\boldsymbol{h}}_{\boldsymbol{t}}$ is primarily derived from the current input $\boldsymbol{x_t}$, effectively ignoring the previous hidden state $\boldsymbol{h_{t-1}}$. Conversely, when the reset gate $\boldsymbol{r_t}$ values are close to 1, the candidate hidden state $\tilde{\boldsymbol{h}}_{\boldsymbol{t}}$ incorporates both the current input $\boldsymbol{x_t}$ and the previous hidden state $\boldsymbol{h_{t-1}}$ [14].

Finally, the output hidden state $\boldsymbol{h_t}$ vector is calculated as in (3), where the update gate $\boldsymbol{z_t}$ is used to determine the extent to which the new hidden state incorporates the old hidden state $\boldsymbol{h_{t-1}}$ compared to the new candidate state $\tilde{\boldsymbol{h}}_{\boldsymbol{t}}$ [14].

$$\boldsymbol{h_t} = \boldsymbol{z_t} \odot \boldsymbol{h_{t-1}} + (1 - \boldsymbol{z_t}) \odot \tilde{\boldsymbol{h}}_{\boldsymbol{t}} \tag{3}$$

When the update gate $\boldsymbol{z_t}$ is close to 1, the hidden state simply retains the old state $\boldsymbol{h_{t-1}}$, effectively ignoring the information from $\boldsymbol{x_t}$ and skipping the current time step t in the dependency chain. Conversely, when $\boldsymbol{z_t}$ is close to 0, the new hidden state $\boldsymbol{h_t}$ closely matches the candidate hidden state $\tilde{\boldsymbol{h}}_{\boldsymbol{t}}$. Due to their gating architecture, GRU are particularly adept at processing sequential data, making them ideal for extracting valuable features from the temporal signals of radar reflections from UAVs.

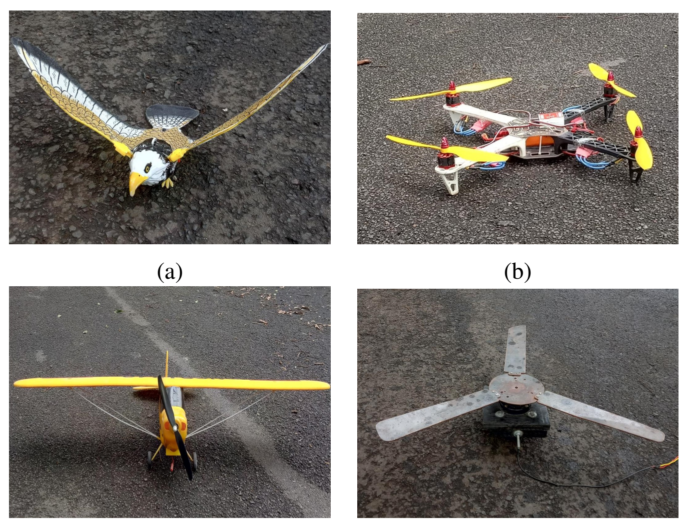
III. Implementation Details

This section provides an overview of the DIAT-μSAT dataset utilized for evaluating the three previously described methods, along with a comprehensive explanation of the experimental setup employed for this purpose.

A. Dataset

To conduct our experiments, we required a radar dataset with a large number of signals per class. Consequently, we opted to use the DIAT-μSAT dataset [12], which encompasses 4,849 radar signatures for six distinct aerial targets, including both bionic birds and various types of UAVs, amounting to approximately 800 radar signals per class. The radar system used to collect this dataset is a Continuous Wave (CW) radar operating in the X-band (10 GHz). Notably, all the UAVs targets operated at different speeds and orientations, including revolutions per minute (RPM) ranging from 200 to 1740, flapping rates of 2 to 4 flaps per second, azimuth angles spanning from 0° to 360° in step of 45°, and elevation angles ranging from 0° to 90° in step of 30°, with tilted target positions relative to the radar's boresight. The echo signals were collected over a period of 3 seconds with ranges of 10-15 meters between the radar and targets in open-field experiments, covering various activities such as flying, take-off, landing, turning, hovering, and maneuvering, at all elevation angles for all azimuth angles.

Figure 3 showcases four examples of aerial targets utilized to construct the DIAT-μSAT dataset, each representing a distinct class. These examples include a bionic bird, a quadcopter, an RC plane, and a UAVs featuring one rotor and three long blades.



(a) (b) (c) (d)

Fig. 3: Examples of aerial targets used to build DIAT-μSAT dataset: (a) Bionic bird, (b) Quadcopter, (c) RC plane and (d) Three long blades UAVs.

B. Experimental Setup

To compare and evaluate the three methods presented in Section II, we developed two neural network models: one for the proposed method, which relies on temporal signals, and another for the spectrograms-based method. For the CNN-based method, the results of experiments are already available and detailed in [12]. All models were designed to classify four types of aerial targets, including a bionic bird, as well as three specific types of UAVs: RC plane (fixed-wing UAVs), quadcopter, and a UAVs with three long blades. These models were trained using the same DIAT-μSAT dataset.

The model designed for the spectrograms-based method consists of three GRU layers, each with 64 cells, followed by a dense layer with 32 neurons. The final output layer, with a softmax activation, contains four neurons representing the expected classes. In this setup, the spectrograms were fed into the neural network, where the time windows of each spectrogram served as time steps, and the FFT points were used as feature vectors.

These spectrograms were generated from radar signals sourced from the DIAT-μSAT dataset using the Short-Time Fourier Transform (STFT) with 1024 FFT points across the entire 3-second signal duration (30000 samples). The Hamming window's length was set to 256 samples, equivalent to 25.6 milliseconds, with an overlap of 200 samples, roughly 78.12% overlap, corresponding to 20.0 milliseconds.

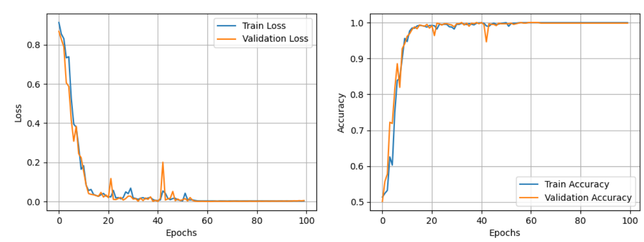
For training all models, we utilized the entire set of radar signals for each class from the DIAT-μSAT dataset, which comprises 800 signals per class. The data for the four classes was split randomly into 70%, 20%, and 10% for the training, validation, and testing phases, respectively. The optimization was performed using the Adam optimizer initialized with a learning rate of 0.001. A batch size of 16 samples was adopted for training, and the categorical cross-entropy loss function was employed to guide the adjustment of model weights during the training process. Finally, the two networks were trained from scratch over 100 epochs. The experiments were conducted on a shared cluster node equipped with a 10-core CPU and an NVIDIA Tesla V100 DGXS graphics card with 32 GB of GPU memory. We built and trained the models using Python 3.11.5, TensorFlow 2.10, and Keras 2.10.

The computational time required for training the model designed for temporal signals-based method was about 9 hours, while the model designed for spectrograms-based method required only 14 minutes for training. During the prediction phase, the time required to predict all samples reserved for testing using the proposed method's model was 15.14 seconds. In contrast, the model designed for the spectrograms-based method required more than double that time, taking 37.48 seconds. This duration includes both the calculation of spectrograms and the model's prediction process.

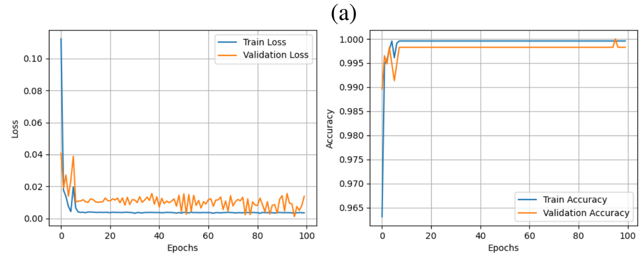
IV. Results

In this section, we delve into the results obtained from training three models using the DIAT-μSAT dataset and provide a detailed analysis. We compare the performance of the proposed method with the spectrograms-based method. Additionally, we evaluate how our results fare against the CNN-based models proposed in [12]. To assess the models' performance, we utilize standard metrics including accuracy, precision, recall, and F1-score. Accuracy provides a straightforward measure of how often the model correctly classifies radar signals, giving an overall sense of model performance. Precision focuses on the correctness of positive predictions, ensuring that the identified UAVs types are correctly classified [15]. Recall measures the model's ability to correctly identify all true UAVs [15]. The F1 score combines precision and recall into a single metric, providing a balanced view of model performance [15]. The classification reports for the model designed for the proposed method and the spectrograms-based method are presented in Table I and Table II, respectively. For comparison with the CNN-based models, the classification report is provided in Table III.

The evolution of loss and accuracy during the training process is depicted in Fig.4 for both the proposed method and the spectrograms-based method models. Examination of the accuracy and loss plots indicates no signs of overfitting or underfitting. The smooth decrease in loss and simultaneous increase in accuracy throughout training suggest effective model learning. Additionally, the convergence of the loss curve to a low value and the accuracy curve to a high value signify that the models have learned to generalize well on the training data. Furthermore, the absence of significant fluctuations or irregularities in the loss and accuracy curves suggests stable and consistent learning.



(a)



(b)

Fig. 4: Loss and accuracy plots of developed models during training: (a) for temporal signals-based method and (b) for spectrograms-based method.

Based on the results presented in Tables I, II, and III, all methods exhibit robust overall performance, with the temporal signals-based method slightly outperforming the other two methods in terms of accuracy, precision, F1-score, and recall. Specifically, it achieves 100% in accuracy, precision, recall and F1-score. Compared to the spectrograms-based method, it shows a marginal improvement of 0.2% in accuracy and 0.3% in precision, recall and F1-score. When compared to the CNN-based method, particularly the VGG19 model, the temporal signals-based method offers improvements of 3% in accuracy, 4% in precision, 1.3% in recall, and 2.8% in F1-score.

TABLE I: Classification report of temporal signals-based method.

	Accuracy	*Precision*	*Recall*	*F1-Score*
Bionic Bird	1	1	1	1
RC Plane	1	1	1	1
Quadcopter	1	1	1	1
Three Long Blades	1	1	1	1
Average	1	1	1	1

TABLE II: Classification report of spectrograms-based method.

	Accuracy	*Precision*	*Recall*	*F1-Score*
Bionic Bird	1	1	1	1
RC Plane	1	1	1	1
Quadcopter	0.997	0.988	1	0.993
Three Long Blades	0.997	1	0.987	0.993
Average	0.998	0.997	0.997	0.997

TABLE III: Classification report of CNN-based method (VGG19 CNN-based model) [12].

	Accuracy	*Precision*	*Recall*	*F1-Score*
Bionic Bird		1	1	1
RC Plane		0.97	0.95	0.96
Quadcopter		0.98	1	0.99
Three Long Blades		0.89	1	0.94
Average	0.97	0.96	0.987	0.972

Analyzing the confusion matrices shown in Tables IV and V, we observe that the temporal signals-based method achieves perfect classification accuracy, reaching 100% in the four classes. Similarly, the spectrograms-based method demonstrates near-perfect classification, reaching 100% in three classes, and maintaining a misclassification rate below 1.3% in the remaining class, resulting in an average accuracy of 99.67%. This represents a slight decrease of 0.33% compared to the temporal signals-based method. A comparison with the confusion matrix of the CNN-based methods proposed in [12] reveals that both the temporal signals-based and spectrograms-based methods outperform the CNN-based methods, with the temporal signals-based method surpassing the average accuracy of the CNN-based methods by 3%.

TABLE IV: Confusion matrices for temporal signals-based method

	Bionic Bird	*RC Plane*	*Quadcopter*	*Three Long Blades*
Bionic Bird	100%	0%	0%	0%
RC Plane	0%	100%	0%	0%
Quadcopter	0%	0%	100%	0%
Three Long Blades	0%	0%	0%	100%

V. Conclusion

This paper presents an innovative method for classifying UAVs and birds, utilizing GRU-based networks applied directly to temporal radar signals. The proposed approach takes advantage of the suitability of RNN for analyzing radar data due to their sequential nature and ability to capture temporal dependencies. Our findings demonstrate that effective and improved UAVs classification can be achieved by utilizing the temporal radar signals directly, without any modifications, outperforming both the spectrograms-based and CNN-based approaches.

TABLE V: Confusion matrices for spectrograms-based method

	Bionic Bird	*RC Plane*	*Quadcopter*	*Three Long Blades*
Bionic Bird	100%	0%	0%	0%
RC Plane	0%	100%	0%	0%
Quadcopter	0%	0%	100%	0%
Three Long Blades	0%	0%	1.3%	98,7%

Although the spectrograms-based method also performs well, it may not be efficient for real-time UAVs classification applications due to the significant computational time required for the calculation of spectrograms and the neural network's prediction process. Additionally, using spectrograms does not fully capture the information from both the magnitude and phase components of the radar signal, which can be a factor affecting the accuracy of the classification.

Through our comparison with an alternative method outlined in [12], where Doppler signals are represented as spectrograms transformed into images and analyzed using CNN, we observe that the conventional approach of using CNN to analyze spectrogram images may not be optimal. The characteristics of UAVs spectrograms, which encompass both time and frequency features, suggest that traditional image-based CNN methods may not be well-suited for capturing these multidimensional aspects.

For future research, we plan to work with an extended dataset that includes a broader range of UAVs for different categories, recorded in various flight conditions and weather scenarios. This will enable us to assess the performance of the proposed approach under more realistic conditions. Additionally, we aim to explore the impact of decreasing the signal-to-noise ratio (SNR) of the micro-Doppler signal on the efficacy of the proposed approach. We also intend to investigate how varying the length of the micro-Doppler signal affects the neural network's performance and its ability to capture temporal features over time.

References

[1] Taha, B., & Shoufan, A. (2019). Machine learning-based drone detection and classification: State-of-the-art in research. IEEE access, 7, 138669-138682.

[2] Guvenc, I., Koohifar, F., Singh, S., Sichitiu, M. L., & Matolak, D. (2018). Detection, tracking, and interdiction for amateur drones. IEEE Communications Magazine, 56(4), 75-81.

[3] Chen, V. C., Li, F., Ho, S. S., & Wechsler, H. (2006). Micro-Doppler effect in radar: phenomenon, model, and simulation study. IEEE Transactions on Aerospace and electronic systems, 42(1), 2-21.

[4] Patel, J. S., Fioranelli, F., & Anderson, D. (2018). Review of radar classification and RCS characterisation techniques for small UAVs or drones. IET Radar, Sonar & Navigation, 12(9), 911-919.

[5] Ciattaglia, G., Iadarola, G., Senigagliesi, L., Spinsante, S., & Gambi, E. (2023). UAV propeller rotational speed measurement through FMCW radars. Remote Sensing, 15(1), 270.

[6] Raval, D., Hunter, E., Lam, I., Rajan, S., Damini, A., Balaji, B. (2022, March). Detecting drones with radars and convolutional networks based on micro-Doppler signatures. In 2022 IEEE Radar Conference (RadarConf22) (pp. 1-6). IEEE.

[7] Hanif, A., & Muaz, M. (2021, December). Deep Learning Based Radar Target Classification Using Micro-Doppler Features. In 2021 Seventh International Conference on Aerospace Science and Engineering (ICASE) (pp. 1-6). IEEE.

[8] Park, D., Lee, S., Park, S., & Kwak, N. (2020). Radar-spectrogram-based UAV classification using convolutional neural networks. Sensors, 21(1), 210.

[9] Rahman, S., & Robertson, D. A. (2020). Classification of drones and birds using convolutional neural networks applied to radar micro-Doppler spectrogram images. IET radar, sonar & navigation, 14(5), 653-661.

[10] Kim, B. K., Kang, H. S., & Park, S. O. (2016). Drone classification using convolutional neural networks with merged Doppler images. IEEE Geoscience and Remote Sensing Letters, 14(1), 38-42.

[11] Sun, Y., Abeywickrama, S., Jayasinghe, L., Yuen, C., Chen, J., & Zhang, M. (2020). Micro-Doppler signature-based detection, classification, and localization of small UAV with long short-term memory neural network. IEEE Transactions on Geoscience and Remote Sensing, 59(8), 6285-6300.

[12] Kumawat, H. C., Chakraborty, M., Raj, A. A. B., & Dhavale, S. V. (2021). DIAT-μSAT: Small aerial targets' micro-Doppler signatures and their classification using CNN. IEEE Geoscience and Remote Sensing Letters, 19, 1-5.

[13] Dey, A., Cabanes, Y., Rajan, S., Balaji, B., Damini, A., & Chanchlani, R. (2023, July). Radar-Based Drone Detection Using Complex-Valued Convolutional Neural Network. In 2023 IEEE Sensors Applications Symposium (SAS) (pp. 1-5). IEEE.

[14] Cho, K., Van Merriënboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078.

[15] Rainio, O., Teuho, J., & Klén, R. (2024). Evaluation metrics and statistical tests for machine learning. Scientific Reports, 14(1), 6086.