A Framework and Method for Intention Recognition to Counter Drones in Complex Decision-Making Environments

Yiming Wang[1,†][a], Bowen Duan[1,†][b], Heng Wang[1,†][c], Yunxiao Liu[1,†][d], Han Li[2,†][e] and Jianliang Ai[1,*][f]

[1]*Department of Aeronautics & Astronautics, Fudan University, 220 Handan Road, Shanghai 200433, China*

[2]*Department of Electrical Engineering, University of Shanghai for Science and Technology, 516 Jungong Road, Shanghai 200093, China*

{*yimingwang23, bwduan21, hengwang21*}*@m.fudan.edu.cn, yunxiaoliu20@fudan.edu.cn, hanli@usst.edu.cn,*

Keywords: Counter Drones, Intention Recognition, BiLSTM, Decision-Making.

Abstract: With the large-scale application of drones in various industries and the rapid development of a low-altitude economy, a complex decision-making environment for counter-drones has been formed. Accurate drone intention recognition is of great significance for the defense of core assets such as power facilities. Therefore, based on the proposed intention space description model, this paper establishes a drone intention recognition framework in a complex decision-making environment and defines the key modules and processing procedures. In addition, to further solve the problem of uncertainty in the time window of drone intention recognition information acquisition in complex decision-making scenarios, this paper optimizes the algorithm by introducing a dynamic time window adaptive adjustment mechanism based on the BiLSTM (Bi-directional Long Short-Term Memory) network. The method described was validated through simulation experiments, confirming the effectiveness of the framework presented in this paper. It is capable of performing four types of intention classification, assessing threats, and scoring for offensive targets. The optimized BiLSTM method demonstrates high recognition accuracy. For drone targets with varying intentions, the recognition accuracy exceeds 96% after applying the time window division.

1 INTRODUCTION

Drones are now commonly utilized in various fields, including power inspection, agriculture, forestry, plant protection, geographical surveying and mapping, and transportation. However, this increased use has led to numerous challenges regarding drone safety supervision. As a result, there is an urgent need for effective intention recognition systems to protect critical facilities from potential drone threats. Currently, standard methods for early warning detection of drones primarily include radar detection (Yang et al., 2023), radio detection (Cai et al., 2024), and photoelectric detection (Wang et al., 2024). These techniques allow for the identification of a drone's position and movement. The goal of drone early warning detection is to assess whether a drone poses a threat to protected facilities by classifying and identifying it. To effectively evaluate the threat level of a drone, it is important to analyze its intentions based on the available information about the drone. For instance, in the context of substation defence, drones operating nearby may be engaged in a variety of tasks, such as inspecting transmission lines, spraying , conducting surveys, or capturing recreational aerial photography. This leads to a complex operating environment for drones. Traditional detection methods, which depend on detection equipment and management platforms, can only identify drone targets and provide information about their positions and movements without offering detailed insights into their potential intentions. Whether it poses a threat still needs to be further judged manually (Xue et al., 2024; Lofù

[a] https://orcid.org/0009-0002-5100-977X
[b] https://orcid.org/0009-0004-1101-3744
[c] https://orcid.org/0009-0003-5207-6832
[d] https://orcid.org/0009-0004-0593-9444
[e] https://orcid.org/0009-0007-0284-2083
[f] https://orcid.org/0000-0002-8982-8503

[†] means these authors contributed equally to this work and should be considered co-first authors; * means the corresponding author.

Wang, Y., Duan, B., Wang, H., Liu, Y., Li, H. and Ai, J.
A Framework and Method for Intention Recognition to Counter Drones in Complex Decision-Making Environments.
DOI: 10.5220/0013431800003929

In *Proceedings of the 27th International Conference on Enterprise Information Systems (ICEIS 2025) - Volume 1*, pages 913-920
ISBN: 978-989-758-749-8; ISSN: 2184-4992

et al., 2023; Zhang et al., 2023). The above defects pose severe challenges to defence in complex drone operating environments. There is an urgent need to research drone intention recognition-related technologies to achieve automated drone intention recognition and provide a reference basis for threat assessment.

The drone intention recognition process translates multidimensional features into intention output. These features may include the drone's position, movement state, mission area, and facilities to be protected. For the simplified drone intention recognition method, based on the observed proximity of the drone's position to a particular area, it can be judged whether the drone poses a threat. Bayesian reasoning method can better realize the inference of drone intrusion intention, but this method needs to obtain the drone state space for a long time (Liang et al., 2021; Yun et al., 2023b). This method makes it difficult to accurately predict the potential intention of the drone in a short time and results in poor real-time intention recognition performance. To improve the accuracy of drone intention recognition, an intention recognition framework based on the crucial waypoint and critical route area modelling was proposed (Kaza et al., 2024), and the motion process of the drone was used as a critical element in inferring intention (Yun et al., 2023a). In addition, the method of intention recognition based on drone flight trajectory prediction has also been studied to a certain extent (Fraser et al., 2023; Yi et al., 2024). The above methods have made specific contributions to drone intention recognition. However, modelling identifying intentions in complex drone missions and area types requires further research and testing.

In this paper, contribution could be summarized as follows:

1. An intention space description model of drones was established for the complex decision-making scenarios of counter drones under the operation of multiple types of drones, aiming at the problem of intention recognition. The mapping with four types of intentions was expressed based on the position, velocity, and area characteristics of drones. The model forms the basis for research on intention recognition of counter drones.

2. A drone intention recognition framework for different types of tasks in complex decision-making scenarios was proposed. It includes a data standardization module, a multidimensional intention recognition module, and an offensive targets threat ranking module that can realize intention recognition based on drone original detection information. This framework can further complete the discriminative output of offensive target sequences.

3. To address the uncertainty optimization problem of information collection time windows for drone intention recognition, a recognition method based on the BiLSTM neural network with a dynamically adapting time window is proposed. Different time windows are employed for intention recognition depending on the type of drone. Simulation experiments have confirmed the accuracy of this method.

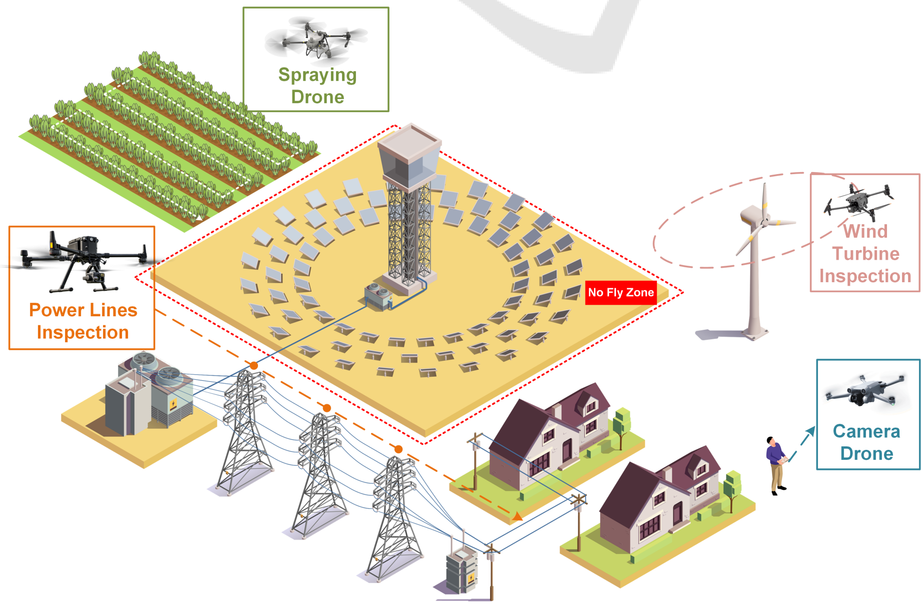

Figure 1: Substation defence scenario.

2 SYSTEM FRAMEWORK

This paper discusses the method using the defence scenario of important power facilities as an example. As shown in Fig. 1, the no fly zone marked area is an important power facility, which farmland, transmission lines, wind power generation facilities, and residential areas surround. There are many types of drones with different purposes running in the above-mentioned different types of areas. The main features of intention recognition classification can be extracted based on the mission area type and the flight status information of the drones. Therefore, this paper proposes a drone feature modelling method based on the fusion of mission area type and flight status information, which is used as an effective input for the intention recognition reasoning framework.

2.1 Framework Implementation

The proposed framework for countering drone intention recognition consists of three modules, as Fig. 2 shown. The Multi-dimensional intention recognition module adopts a multi-dimensional temporal intention recognition method based on BiLSTM to identify the intention of the drone through serialized drone feature information. It should be noted that there are 6 BiLSTM networks constructed in this paper, which can be adaptively selected according to the different Time windows of drone feature information to improve the accuracy of drone intention recognition results. The network needs to be trained and tuned before applying the BiLSTM network. The serialized feature information of the drone comes from the Data standardization module. In this module, the original information on the drone is processed by timeline snapping, and the division of the time window is solved according to the distance to closest point of approach (DCPA) and the time of closest point of approach (TCPA) information of the drone and the regional information where it is located. The Offensive targets identification module analyzes, solves and sorts the threat of offensive drone targets based on the results of the intention recognition module.

2.2 Intention Space

Some typical drone intentions, such as pesticide spraying and power line inspection, are related to some specific movement patterns of drones and the areas where their missions are located. Drone intention recognition is the process of inferring and predicting the target intention based on the target's feature information and intention recognition rules. For a single drone target, the feature information required for its intention recognition over some time can be represented by a multidimensional feature sequence $(Y_{t-n+1}, Y_{t-n+2}, Y_{t-n+3}, \cdots, Y_t)$, which describes the historical feature information of the target from time $t-n+1$ to the current time t, and n is the sequence length. $Y_t = \{y_t^1, y_t^2, y_t^3, \cdots, y_t^N\}$ represents the features of a single target at time t, and the number of features is N. Define its intention space as $M = \{m_1, m_2, \cdots, m_{|M|}\}$, and its number of intentions as $|M|$. Then, there is a mapping relationship:

$$m = f_{Y\to M}(Y_{t-n+1}, Y_{t-n+2}, Y_{t-n+3}, \cdots, Y_t) \in M \quad (1)$$

According to the scenario in Fig. 1 and the drone feature information that can be obtained in the actual anti-drone scenario, this paper established the target

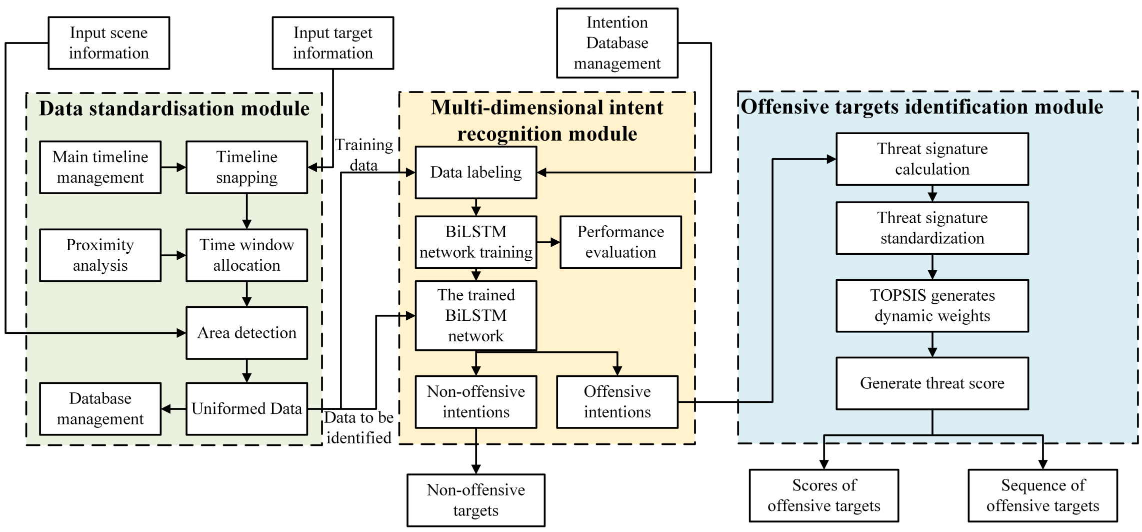

Figure 2: System framework.

feature space $Y = \{y^1, y^2, y^3, \cdots, y^7\}$ and target intention space $M = \{m_1, m_2, m_3, m_4\}$ as shown in Table 1:

Table 1: Target feature space and intention space set.

Space	Symbol	Value
Feature Y	y^1, y^2, y^3	Position information
	y^4, y^5, y^6	Velocity information
	y^7	Area information
Intention M	m_1	Spraying
	m_2	Orbit fly
	m_3	Power lines inspection
	m_4	Free flight

m_1, m_2, m_3 are characteristic intentions to perform specific tasks at specific locations. These targets are considered to be non-offensive targets. Free flight m_4 includes all intentions except m_1, m_2, m_3, including random flying targets with unfixed routes and offensive targets. Therefore, it is necessary to conduct a threat assessment on the targets with m_4 intentions to further confirm whether to take disposal measures against those targets and their disposal order.

3 METHODOLOGY

3.1 Multi-Dimensional Temporal Intention Recognition Based on BiLSTM

The motion trajectory of a drone is a continuous-time segment, which can convey more information than the state information at a single time point, and the track characteristics and possible intentions of a drone are more closely related to the motion trajectory of the drone. Compared with traditional regression algorithms, artificial recurrent neural networks (RNNs) with memory functions can predict the characteristic information of time series more accurately. (Teng et al., 2021) transformed the target intention recognition problem into a multi-classification problem based on time series recognition features. This paper uses the BiLSTM model to train multi-dimensional time series feature data. It uses a probability prediction layer (softmax layer) and a classification output layer to convert the output of the neural network into the intention classification recognition result.

A single-layer BiLSTM is composed of two LSTMs, one for processing the input sequence forward and the other for processing the sequence backwards. After processing, the outputs of the two LSTMs are concatenated to obtain the final BiLSTM output result. While retaining the advantage of the LSTM model in capturing the dependencies of data over a long period, it solves the problem of being unable to encode information from the back to the front (Siami-Namini et al., 2019). The manoeuvrability and motion pattern of a drone is a features that can be interpreted and judged forward and backward. Therefore, using BiLSTM can extract the motion features of the drone more accurately and identify the drone's intentions more accurately. The LSTM model consists of a forget gate, an input gate and an output gate. Its structure at time t is shown in Fig. 3.

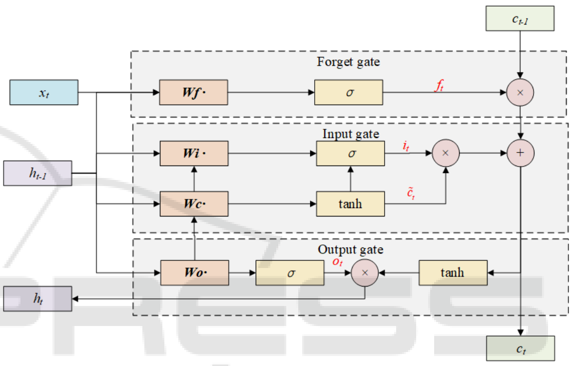

Figure 3: Structure of LSTM.

First, the forget gate can control which long-term memories of the LSTM layer are forgotten:

$$f_t = \sigma(W_f x_t + W_f h_{t-1} + b_f) \tag{2}$$

The input gate then computes the information obtained from the input:

$$\begin{aligned} i_t &= \sigma(W_i x_t + W h_{t-1} + b_i) \\ \tilde{c}_t &= \tanh(W_c x_t + W_c h_{t-1} + b_c) \\ c_t &= f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \end{aligned} \tag{3}$$

Finally, the output gate updates the current hidden layer state based on the input, the cell state, and the previous hidden state:

$$\begin{aligned} o_t &= \sigma(W_o x_t + W_o h_{t-1} + b_o) \\ h_t &= o_t \cdot \tanh c_t \end{aligned} \tag{4}$$

Where W and b represent weight and bias terms, respectively, and σ represents the sigmoid activation function.f_t, i_t, o_t are the results of the forget gate, input gate and output gate, and c_t is the variable transfers the long-term memory.

This paper establishes a drone intention recognition method based on the BiLSTM network, as shown

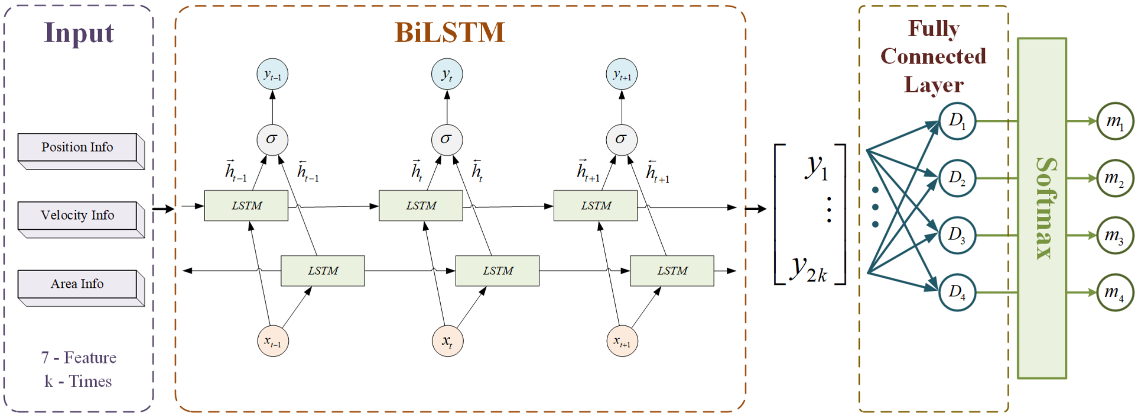

Figure 4: Schematic of Drones intention identification method based on BiLSTM.

in Fig. 4. The input layer of the network is the time series features of the drone. Among them, k is the time window, which determines the BiLSTM network for the subsequent feature information input. There are four types of intentions in the classification output layer of the network, and their definitions are the same as in Table 1. The feature information of the drone is classified and predicted by the BiLSTM network, and the targets with m_4 intention are screened out to further determine whether they have attack intention.

3.2 Time Window Division and Threat Ranking

The input to the BiLSTM network consists of a continuous feature sequence of the drone, which includes attributes such as position, speed, region, and other relevant features. The length of the time window, denoted as k, is a crucial parameter. Given the variations in region, flight speed, altitude, and heading of the drone being analyzed, the length of the feature sequence necessary for accurately determining its intention can vary. Selecting an appropriate time window k can significantly enhance the efficiency of drone intention recognition while minimizing the consumption of computing resources. It is worth noting that offensive targets often have the characteristics of high speed and clear direction. The time from discovering the target to disposing of the target is short. Selecting a smaller time window can identify the drone's attack intention more quickly.

To determine whether the target has an attack intention, this paper selects three typical motion features as the basis for target threat calculation and ranking, namely, DCPA between the target and the substation, TCPA and target height. Among them, drones with lower flying altitudes are more concealed and switch to attack behaviour faster and can be used as a supplementary factor in target threat calculations.

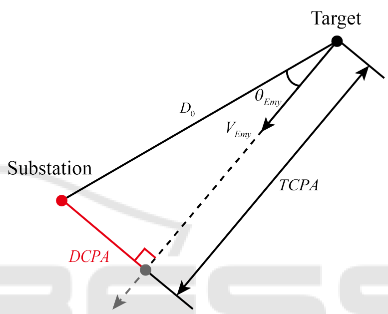

Figure 5: Schematic of DCPA and TCPA.

Fig. 5 shows the definitions of DCPA and TCPA, where the red dot represents the centre position of the substation, the black dot is the moving target, the target speed is V_{Emy}, and the angle between the target speed and the target relative to the substation is θ_{Emy}. The distance between the target and the substation during movement can be expressed by Eq. (5).

$$\begin{aligned} D(t) = & \left(V_{Emy}^2 - 2V_{Emy}\cos\theta_{Emy}\right)t^2 \\ & + 2\left(V_{Emy}\cos\theta_{Emy}\right)D_0 t + D_0^2 \end{aligned} \tag{5}$$

According to Eq. (5), the expressions of DCPA and TCPA can be obtained as shown in Eq. (8) and Eq. (9).

$$DCPA = \begin{cases} D_0 & \cos\theta_{Emy} < 0 \\ D_0 \sin\theta_{Emy} & \cos\theta_{Emy} \geq 0 \end{cases} \tag{6}$$

$$TCPA = \begin{cases} 0 & \cos\theta_{Emy} < 0 \\ \frac{D_0 \cos\theta_{Emy}}{V_{Emy}} & \cos\theta_{Emy} \geq 0 \end{cases} \tag{7}$$

Finally, the DCPA, TCPA, and height indicators of the target are normalized to obtain the threat grade of each indicator.

$$\xi_D(D) = 1 - \frac{DCPA}{\bar{D}} \tag{8}$$

$$\xi_T(T) = \frac{\overline{T}^2}{TCPA^2 + \overline{T}^2} \tag{9}$$

$$\xi_H(H) = \frac{\overline{H}^2}{H^2 + \overline{H}^2} \tag{10}$$

$\bar{D}$, $\bar{T}$, $\bar{H}$ are the normalized units of $DPCA$, $TCPA$, and H.Set W and Q as the evaluation standard of the time window and the threat assessment model. The expression of W and Q can be written as:

$$W = \frac{1}{2}\left[\xi_D(D) + \xi_T(T)\right] \tag{11}$$

$$Q = \frac{1}{3}\left[\xi_D(D) + \xi_T(T) + \xi_H(H)\right] \tag{12}$$

4 SIMULATION EXPERIMENTS

4.1 Simulation Scenario

To verify the feasibility of the intention recognition method and the correctness of the recognition results, we designed the actual scene of Fig. 1 as a two-dimensional plane scene as shown in Fig. 6. The whole scene contains four areas: the Spraying zone, No fly zone, Wind turbine zone and Power lines zone.

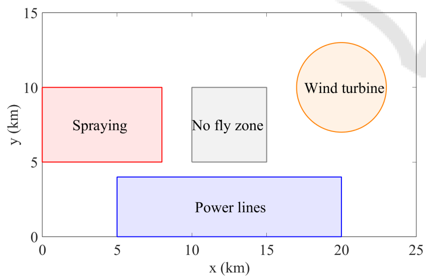

Figure 6: Experimental scenario description.

To clearly illustrate the trajectory of the drone, 100 drone targets were selected at random. The two-dimensional and three-dimensional position diagrams of these targets are displayed in Fig. 7 and Fig. 8.

4.2 Dataset Generation and Time Window Division

This paper presents a dataset consisting of 3,000 batches of drone targets, categorized by their intended purposes. The dataset includes 500 batches intended for spraying, 500 batches for orbit fly, 500 batches for power lines inspection, and 1,000 batches designated for free flight. A detailed breakdown of the dataset can be found in Table 2. The 3,000 sets of drone flight data were divided into training and validation sets in a 9:1 ratio to facilitate the development and evaluation of the intention recognition model.

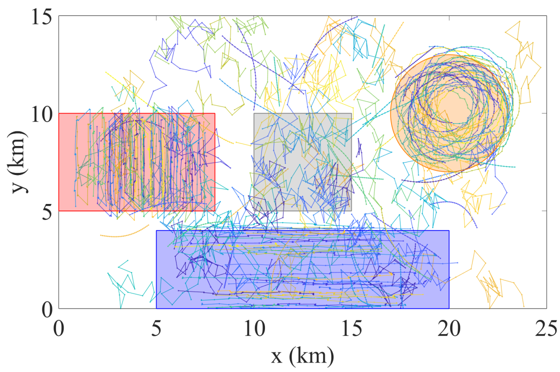

Figure 7: Horizontal trajectory of drones.

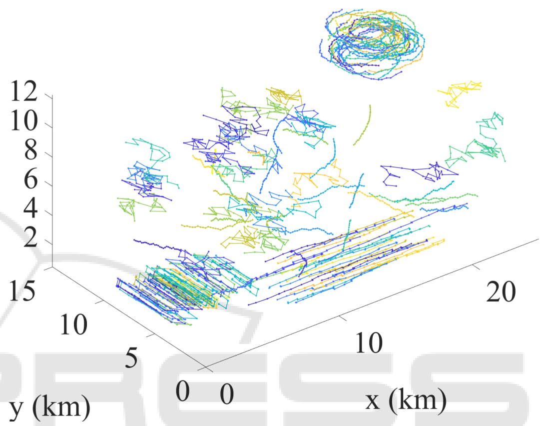

Figure 8: Three-dimensional flight trajectory of drones.

Table 2: Dataset detailed description.

Intention M	Dataset size	Description
Spraying	500	Flying in an S-shaped trajectory around farmland
Orbit fly	500	Surrounding wind turbines for inspection work
Power lines inspection	500	Movement along the power line direction
Free flight	1,000	Without the above three moving characteristics

Calculate W for each target using the Eq. (11), then divide the 3000 batches of targets into time windows according to the rules in Table 3. The six data sets, divided by time windows, are input into the BiLSTM network for training and tuning. This process results in a prediction classification network for each time window, which identifies the intention of the drones.

Table 3: Time Window division results.

W	Dataset size	Time Window	Intention type
$(0.8, 1]$	583	5 s	m_1, m_2, m_3, m_4
$(0.6, 0.8]$	538	10 s	m_1, m_2, m_3
$(0.4, 0.6]$	369	15 s	m_1, m_2, m_3
$(0.3, 0.4]$	183	20 s	m_1, m_2, m_3
$(0.2, 0.3]$	552	25 s	m_1, m_2, m_3, m_4
$[0, 0.2]$	285	30 s	m_1, m_2, m_3, m_4

4.3 BiLSTM and Threat Ranking

The recognition model shown in Fig. 4 is set for six time periods, respectively. The time window length determines the sequence length of the input layer and the BiLSTM output, and the number of input and output nodes of the fully connected layer and the softmax layer is determined by the total number of intentions in Table 3.

Table 4 shows the experiment's training environment and parameter settings based on the BiLSTM network constructed in this paper.

Table 4: Training environment and parameters.

Training parameters	Value
Environment	Intel Core i7-13650HX
Optimizer	ADAM
Learning rate	0.001
Hidden units of LSTM	100
Max epochs	100
Min batch size	25
Gradient threshold	1.0

The data set is trained according to the time window from short to long, and the accuracy curve and loss function curve of the intention recognition network under each time window are obtained as shown in Fig. 9 and Fig. 10.

We can observe that as the time window increases, more motion features are preserved, leading to a faster training convergence speed, higher training accuracy, and a loss function that approaches zero. Additionally, since there are only three intentions to identify when $k = 10, 15, 20$, the convergence for these three time windows is further enhanced.

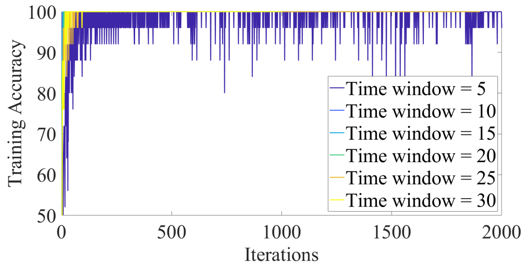

Figure 9: Accuracy curve of the training.

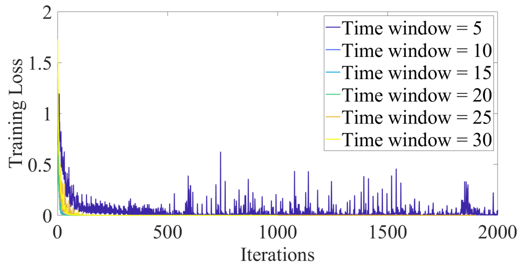

Figure 10: Loss curve of the training.

Subsequently, the validation set is used to test the accuracy of the intention recognition network, and the threat score of the validation set is calculated using Eq. (12) to obtain the recognition accuracy and average threat value of each validation set as shown in Table 5.

Table 5: Training results and average threat value.

Time window	Training time	Accuracy of validation	Average threat value
5 s	19 s	0.9655	0.8659
10 s	23 s	1.0000	0.5808
15 s	20 s	1.0000	0.5627
20 s	14 s	1.0000	0.5306
25 s	39 s	0.9818	0.5796
30 s	76 s	0.9643	0.3569

From Table 5, we can see that for the $k = 5, 25, 30$ time windows containing four intentions targets, the accuracy of intention recognition is above 96%, and the shorter the time window, the higher the final threat score. For the $k = 10, 15, 20$ time windows containing three intentions targets, the training time is shorter, the recognition accuracy reaches 100%, and the final threat score is relatively average.

5 CONCLUSION

This paper aims to solve the problem of drone defence in complex decision-making environments. It constructs an intention space model to establish the connection between target motion characteristics, regional information, and target intentions so as to test the ability of intelligent algorithms to recognize multi-dimensional intentions in complex environments. In addition, the dynamic time window adaptation proposed in this paper dynamically adjusts the time window of the BiLSTM network based on motion rules and the principle of reserving sufficient disposal time while ensuring rapid intention recognition of high-threat targets and accurate intention recognition of medium-threat and low-threat targets. Finally, based on a multi-dimensional intention dataset for complex decision-making environments, window division and intention recognition were performed based on BiLSTM and dynamic time window adaptation, verifying that the dynamic time window can be dynamically processed while ensuring the accuracy of intention recognition.

In future work, we would like to take the drones' types and features that can be obtained in reality to expand the feature space set. By establishing time window division rules and threat ranking indicators for different types of drones, we can make full use of detectable information for faster and more accurate intention recognition. How to dynamically adjust the time window division and feature selection according to the comprehensive motion state of all targets is also an important direction for future work. In addition, we would like to expand the intention space set based on more realistic complex decision-making environments, using manual reasoning and system-derived methods so that the system can more accurately identify the intentions of multiple types of targets in more real scenarios.

REFERENCES

Cai, Z., Wang, Y., Jiang, Q., Gui, G., and Sha, J. (2024). Toward intelligent lightweight and efficient uav identification with rf fingerprinting. *IEEE Internet of Things Journal*.

Fraser, B., Perrusquía, A., Panagiotakopoulos, D., and Guo, W. (2023). A deep mixture of experts network for drone trajectory intent classification and prediction using non-cooperative radar data. In *2023 IEEE Symposium Series on Computational Intelligence (SSCI)*, pages 1–6. IEEE.

Kaza, K., Mehta, V., Azad, H., Bolic, M., and Mantegh, I. (2024). An intent modeling and inference framework for autonomous and remotely piloted aerial systems. *arXiv preprint arXiv:2409.08472*.

Liang, J., Ahmad, B. I., Jahangir, M., and Godsill, S. (2021). Detection of malicious intent in non-cooperative drone surveillance. In *2021 Sensor Signal Processing for Defence Conference (SSPD)*, pages 1–5. IEEE.

Lofù, D., Di Gennaro, P., Tedeschi, P., Di Noia, T., and Di Sciascio, E. (2023). Uranus: Radio frequency tracking, classification and identification of unmanned aircraft vehicles. *IEEE Open Journal of Vehicular Technology*.

Siami-Namini, S., Tavakoli, N., and Namin, A. S. (2019). The performance of lstm and bilstm in forecasting time series. In *2019 IEEE International Conference on Big Data (Big Data)*, pages 3285–3292.

Teng, F., Guo, X., Song, Y., and Wang, G. (2021). An air target tactical intention recognition model based on bidirectional gru with attention mechanism. *IEEE Access*, 9:169122–169134.

Wang, C., Meng, L., Gao, Q., Wang, T., Wang, J., and Wang, L. (2024). A target sensing and visual tracking method for countering unmanned aerial vehicle swarm. *IEEE Sensors Journal*.

Xue, C., Li, T., and Li, Y. (2024). Radio frequency based distributed system for noncooperative uav classification and positioning. *Journal of Information and Intelligence*, 2(1):42–51.

Yang, Y., Yang, F., Sun, L., Xiang, T., and Lv, P. (2023). Echoformer: Transformer architecture based on radar echo characteristics for uav detection. *IEEE Sensors Journal*, 23(8):8639–8653.

Yi, P., Yang, Y., and Lin, Z. (2024). Cooperative opponent intention recognition based on trajectory analysis by distributed fractional bayesian inference. In *2024 6th International Conference on Electronic Engineering and Informatics (EEI)*, pages 1551–1555. IEEE.

Yun, J., Anderson, D., and Fioranelli, F. (2023a). Estimation of drone intention using trajectory frequency defined in radar's measurement phase planes. *IET Radar, Sonar & Navigation*, 17(9):1327–1341.

Yun, J., Shin, H.-S., and Tsourdos, A. (2023b). Perimeter intrusion prediction method using trajectory frequency and naive bayes classifier. In *2023 23rd International Conference on Control, Automation and Systems (ICCAS)*, pages 1610–1615. IEEE.

Zhang, H., Li, T., Li, Y., Li, J., Dobre, O. A., and Wen, Z. (2023). Rf-based drone classification under complex electromagnetic environments using deep learning. *IEEE Sensors Journal*, 23(6):6099–6108.