IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 23 October 2023, accepted 27 November 2023, date of publication 28 November 2023,
date of current version 4 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3337640

RESEARCH ARTICLE

Combat Intention Recognition of Air Targets Based on 1DCNN-BiLSTM

CHENHAO ZHANG[1], YAN ZHOU[1], HUI LI[1], FUTAI LIANG[1,2], ZIHAO SONG[1], AND KAI YUAN[1]
[1]Air Force Early Warning Academy, Wuhan 430014, China
[2]31121 PLA Troops, Nanjing, Jiangsu 210000, China

Corresponding author: Chenhao Zhang (zch96@foxmail.com)

This work was supported in part by the Graduate Student Sponsored Projects under Grant JY2020B149 and Grant JY2022C204.

ABSTRACT In air combat, target intent recognition is the premise and foundation of battlefield situation awareness and intelligent decision-making. Aiming at the problem that traditional intention recognition methods cannot deal with a large amount of continuous target data, an air target combat intention recognition model based on one-dimensional convolutional neural networks and bidirectional long short-term memory (1DCNN-BiLSTM) is proposed. First, the target data is divided into fixed-size continuous subsequences by time sliding window on the basis of determining the target feature space and intention space. Second, the convolution operation is performed on the target sequence through the 1DCNN module as a means of extracting the features of the target attributes in the time dimension, and at the same time reducing the dimensionality of the target sequence, so as to facilitate the subsequent processing of the target data. Then, the BiLSTM module is utilized to capture the dependencies on the longer distance of the target sequence from both forward and reverse directions simultaneously. Finally, the optimal model structure and hyperparameters of 1DCNN-BiLSTM are not only determined through experiments, but also the validity of each part of the model is verified. Compared with the traditional methods, the model proposed in this paper effectively improves the accuracy of combat intention recognition of air targets, and provides the essential basis and auxiliary support for the decision-making of the commanders.

INDEX TERMS Battlefield situation awareness, air targets, intention recognition, one-dimensional convolutional neural networks, bidirectional long short-term memory.

I. INTRODUCTION

In recent years, with the rapid development of science and technology and military doctrine, the performance of weapons and equipment in the air battlefield has increased dramatically, and the combat styles are complex and diverse, leading to rapid changes in the air battlefield situation. The intention of an air target refers to the task to be completed or the purpose to be achieved by the air target. Typically, target intention cannot be directly observed, but is obtained by observing the behavior and state of the target and analyzing it. Air targets combat intention recognition refers to the process of analyzing and speculating the air target data obtained from various information sources and combining them with a priori knowledge to finally get the target's operational plan, operational scenario, and so on. As a key step in battlefield situation assessment, target combat intention recognition can provide decision information for commanders.

In modern air warfare, intricate sensor systems provide comprehensive tracking and surveillance of targets, resulting in the acquisition of a vast amount of data. In addition, the rapidly evolving air battlefield situation is also generating information from moment to moment. In the face of massive air target data, merely relying on the commander's combat experience to recognize the target's intention can no longer meet the high-intensity, fast-paced operational requirements of modern air warfare. Therefore, how to efficiently and accurately utilize a large amount of battlefield situational data to identify the intention of an air target has become an urgent problem in battlefield situational assessment.

The associate editor coordinating the review of this manuscript and approving it for publication was Rosario Pecora.

In the existing research on target intention recognition, scholars at the beginning mainly used expert systems [1], template matching [2] and other methods, but such methods have to convert expert knowledge into intention recognition rules and construct recognition models, which rely heavily on a priori knowledge and cannot meet the requirements of high efficiency and high accuracy of the current air battlefield situation. In order to apply the air battlefield situation data more efficiently, the intention recognition methods based on Dempster-Shafer (D-S) Evidence Theory [3], [4] and Bayesian networks [5], [6] begin to receive attention from scholars. These two types of methods introduce probabilistic reasoning to reduce the impact of a priori knowledge on the recognition effect. The rapid development of machine learning provides a new path for the efficient recognition of target intention. After inputting the air battlefield situation data into the machine learning model, the model is trained to obtain the model. However, the existing recognition models mainly recognize the target intention through machine learning algorithms such as RNN [7], DNN [8], etc., which cannot perform effective feature extraction on long sequences of target data. In summary, in order to avoid relying too much on subjective factors, and at the same time to be able to deal with a large number of long sequences of target data and effectively extract the feature information, this paper proposes an air target combat intention recognition model based on 1DCNN-BiLSTM. The main contributions are summarized below.

(1) Aiming at the problem that traditional intention recognition methods cannot deal with long sequential target data, a target data segmentation method is proposed, which divides the air target data into multiple sub-sequences through a temporal sliding window, so as to facilitate the processing of the data by the subsequent model.

(2) In order to better extract the target features in the data, the 1DCNN module is introduced which not only accurately captures the features of the target attributes in the time dimension while maintaining the temporal structure of the target sequence, but also reduces the training overhead of BiLSTM by lowering the dimension of the input target data.

(3) Aiming at the problems of gradient explosion and gradient vanishing that tend to occur when training the target sequence data, the model better captures the dependencies of the target data over longer distances in the time dimension through BiLSTM in both forward and reverse directions to improve the model's intention recognition capability.

The rest of the paper is organized as follows. Section II analyzes and summarizes related work. Section III analyzes and discusses the problem of air target intent recognition, feature space, and intention space. Section IV builds an combat intention recognition model of air targets based on 1DCNN-BiLSTM. Section V conducts experimental validation of the proposed model and compares it with other classical algorithms. Section VI summarizes the work of this paper.

II. RELATED WORKS

Currently, many scholars have conducted research on target intention recognition, and the main methods include expert systems [1], template matching [2], D-S Evidence Theory [3], [4], Bayesian Networks [5], [9], neural networks [10], [11], and so on. In the case of less target data, it mainly relies on experience and knowledge to define the rules in advance, and then matches the input target data with the rules so as to realize the recognition of the target intention. Based on this idea, He et al. [1] proposed an air targets intention recognition method based on a confidence rule database, abstracting expert knowledge and related information into rules to form a confidence rule database, and then using Differential Evolution to optimize the initial confidence rule database, so as to reason and recognize the combat intention of air targets. However, this method relies heavily on manual experience and knowledge, and cannot recognize the types of intentions beyond the range of empirical knowledge. Moreover, when the requirements change, it needs to rely on manual modification and updating of the rule database, which increases the workload. Aiming at the problems of rule matching, the multi-layer blackboard model can not only deal with more complex problems through multi-level reasoning, but also can be extended and adjusted relatively easily. Li et al. [2] proposed a target intention recognition method based on the improved multi-layer blackboard model, which introduces the target threat degree into the classical multi-layer blackboard model to realize online recognition of target intention. This method splits the target intention recognition into multiple sub-problems, which can simplify the problem, but each sub-problem is independent of the other and lacks the mechanism of global information sharing and mutual collaboration, which may lead to the failure of obtaining the global optimal solution.

In addition, both rule matching and multi-layer blackboard models rely on manual experience and knowledge and are weak in dealing with uncertainty reasoning problems, whereas D-S Evidence Theory and Bayesian Networks, as mathematical models based on probabilistic theory, are more applicable to complex and changing air battlefield situation. Wang et al. [3] and Cao et al. [4] utilized the D-S Evidence Theory to construct an intent prediction model to achieve the recognition of target intention. In order to exclude the influence of subjective factors in determining the basic probability assignment (BPA), Zhang et al. [12] combined a deep learning network with the D-S evidence theory to determine the BPA through an LSTM network and a GAN network. this approach still fails to solve the problem of possible conflicting evidence in the D-S theory. Zhang et al. [13] combined Evidence Network (EN) and Belief Decision Trees and Random Forests (BDT-RF) to recognize the target intention individually before fusing the results. Although this method can solve the uncertainty existing in intent recognition to some extent, EN and BDT-RF introduce a large number of nodes and relationships in a complex problem.

Yang et al. [9] proposed a hierarchical recognition method of target intention based on Bayesian reasoning in accordance with the commander's decision logic, by decomposing the intention recognition into two layers, the first layer identifies two major types of air target intention, and the second layer specifically identifies the target intention on the basis of the previous layer. In order to adapt to the dynamically changing battlefield situation, Chai and Wang [5] proposed a dynamic Bayesian-based tactical situation estimation method, which constructs a dynamic Bayesian inference network by unfolding a static Bayesian network in the time dimension, so as to realize the dynamic recognition of the intention in the target's tactical situation. D-S Evidence Theory and Bayesian Network, as two types of probabilistic reasoning-based methods, need to determine the basic probability assignment and prior probability when recognizing target intention, which is more dependent on subjective assumptions. With the gradual quantization and high-dimensionalization of air target data, the above methods can no longer adapt to the realistic needs of intention recognition. Neural networks, as a method capable of handling nonlinear and high-dimensional air target data, have been continuously applied to the field of intention recognition. Xue et al. [14] In order to avoid over-reliance on a priori knowledge, they first predicted the destination of target movement by AGADESN (Deep Echo State Network Optimized by Adaptive Genetic Algorithm) model, and then utilized DBSCAN (Density-Based Spatial Clustering of Applications with Noise) algorithm for target behavior determination. Ahmed and Mohammed [10] proposed an intention recognition method using fuzzy min-max neural networks to classify attacks based on their characteristics, identify the motivation of an attack, and predict its intention through similarity metrics. Zhou et al. [11] designed a deep neural network-based target combat intention recognition model for the lack of expert knowledge by introducing the ReLU activation function and the Adam optimization algorithm to improve the effect of recognition. However, this method cannot effectively deal with a large number of target sequences.

III. PROBLEM ANALYSIS

A. ANALYSIS OF COMBAT INTENTION RECOGNITION OF AIR TARGETS

Air targets intention recognition is an important part of air battlefield situation assessment. By processing and analyzing air battlefield situation data and combining relevant laws and principles, we can infer and identify the combat purpose, plan and intention of the target, so as to help develop coping strategies and measures. In air combat, each sensor can obtain the data of the air target within its own detection range, and then the data fusion of the same target data obtained by different sensors can realize the continuous dynamic monitoring of the target within the range of the detection network [15]. Since the target will hide its own intention as much as possible during combat, it is impossible to accurately judge its combat intention only from the state of the target at a single moment or a few moments. However, the target will execute a series of tactical maneuvers when performing combat missions, and different tactical maneuvers can change the target's state. To summarize, the target's combat intention can be deduced from the target's successive moments of state data, that is, there is a mapping relationship between the target's combat intention and the target's battlefield situation data.

Assume that at time t, the battlefield situation data of an air target is $\boldsymbol{x}^{(t)} = (x_1^{(t)}, x_2^{(t)}, \ldots, x_n^{(t)})$, where $x_n^{(t)}$ refers to the n-th attribute feature of the target at time t. Then the battlefield situation data of the air target for m consecutive moments can be expressed as a matrix

$$\boldsymbol{X} = \begin{bmatrix} x_1^{(t_1)} & x_2^{(t_1)} & \cdots & x_n^{(t_1)} \\ x_1^{(t_2)} & x_2^{(t_2)} & \cdots & x_n^{(t_2)} \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{(t_m)} & x_2^{(t_m)} & \cdots & x_n^{(t_m)} \end{bmatrix} \tag{1}$$

where $x_n^{(t_m)}$ denotes the n-th attribute feature of the air target at time t.

Assume that the target intention is $I = \{i_1, i_2, \ldots, i_p\}$, where i_p denotes the p-th intention. Then the relationship between air target battlefield situation data $\boldsymbol{X}$ and target intent I is

$$I = f(\boldsymbol{X}) = f\left(\left(\boldsymbol{x}^{(t_1)}, \boldsymbol{x}^{(t_2)}, \ldots, \boldsymbol{x}^{(t_m)}\right)^{\mathrm{T}}\right) \tag{2}$$

where f denotes the mapping relationship between air target battlefield situation data and target intention.

Obviously, recognizing the combat intention of air targets through battlefield situation data requires an in-depth analysis of the data's distribution pattern, change pattern, etc., and extracting features from it. In this process, identifying air target characteristics and combat intention is the primary task.

B. FEATURE SPACE

Targets display different states when performing different tasks, and the description of the target state is generally portrayed from different angles by different target features. The target battlefield situation data after the fusion of different information sources contains rich and diverse target features, from which the features with high correlation and greater differentiation are selected to construct the feature space of air targets. Constructing feature space is the key step to recognize the combat intention of air targets, and it is also the process of dimensionality reduction of multi-source data, which can not only effectively remove redundant or noisy features to improve the stability and reliability of the model, but also help to reduce the likelihood of model overfitting. By selecting appropriate target features to construct the feature space, the key attributes and states of the target can be better captured, thus realizing more accurate intention recognition of air target.

The recognizition of air targets intention is mainly based on the motion state of target, sensor state, self-performance and other characteristics. Among them, the motion state is the description of the position, speed, height, direction and other attributes related to the movement of the target during the flight. The sensor state is the working state of the sensor itself when the target performs the task. Self-performance is an inherent property of the target itself and does not usually change over time. There are certain differences in the characteristics of targets when they perform different combat tasks. For example, the target's flight height, flight speed and route shape will be different when executing the tasks of attack, early warning and transportation. The sensor state of the target is also different in the task of reconnaissance and penetration. The types of targets are also different when carrying out patrol, bombing and other tasks. To sum up, we construct the feature space of air targets intention recognition from the aspects of motion state, sensor state and its own attributes.

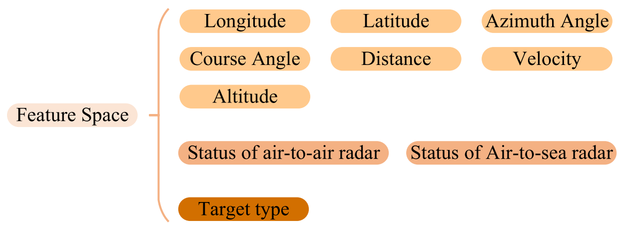


FIGURE 1. Feature space.

The target motion state refers to the motion state of the target in space, generally including its longitude, latitude, azimuth, heading Angle, distance, speed and height and other physical quantities [16]. Through the analysis of the motion state of the target, the motion law of the target in different tasks can be mined, so as to provide useful information for the recognition of the target intention. The sensor state mainly refers to the working state of air-to-air radar and air-to-sea radar. Air-to-air radar is mainly used to search, track and attack enemy aircraft, while air-to-sea radar is mainly used to search and track maritime targets. However, in order to reduce the possibility of detection by the other side's radar, the target will turn off the radar when carrying out low-altitude raid, electronic jamming and other tasks. The target's own attributes are mainly its own inherent characteristics, such as the length of the target, the maximum takeoff weight, the maximum flight speed, the maximum combat radius, the thrust-to-weight ratio and the amount of ammunition carried. Since some of these attributes are not directly related to the task performed, we choose the type of target as the feature of target intention recognition. Target types include large aircraft, small aircraft, helicopters and missiles, with small aircraft generally deployed for missions requiring high mobility and large aircraft deployed for transport, refueling and early warning. Therefore, the type of target is also one of the key elements to recognize the combat intention. The feature space of air target combat intention recognition is shown in Figure 1.

C. INTENTION SPACE

In essence, target intention recognition is a typical pattern recognition problem [6]. In order to recognize more accurately, it is necessary to define and describe the recognition framework completely and clearly, that is, to construct the intention space of the target. Different operational backgrounds, operational modes and operational objectives, the corresponding intention space is different. Therefore, when determining the combat intention space of air target, it is necessary to consider the operational background, operational mode and operational entity.

First of all, the enemy's objectives and actions will be different for different combat contexts. For example, in offensive air warfare, the targets of air attacks are usually fighters, bombers or other air targets, and the intention of the operation may be to shoot down the target, destroy the effectiveness of the other side, or prevent the other side from attacking; The main combat intention of defensive air warfare is to protect one's own base, military facilities, or personnel from air attacks. In electronic warfare, the goal of an air attack may be to interfere with the other side's communications equipment, radar or other electronic equipment, and the intention of the operation is primarily to interfere with the other side's communications and radar systems so that they cannot function properly. Secondly, for different combat styles, the main tasks performed by the target are different. For example, in a surprise attack, the enemy will usually use a quick attack to hit our targets; In defensive warfare, it is necessary to pay more attention to the ability to defend and counter enemy attacks. Therefore, it is necessary to identify the intention according to different characteristics in different styles of operations. Finally, the objectives and actions of the enemy will be different for different target entities. For example, for fighter aircraft, its main task is to strike air targets; For reconnaissance aircraft, it is necessary to pay more attention to the search and defense capabilities of the target. To sum up, according to different operational backgrounds, operational styles and operational entities, the combat intention space defined for air targets includes seven types of intentions, including attack, bombing, refueling, patrol, early warning, transport and scout, as shown in Figure 2.

The attack mainly relies on a variety of weapon systems, such as missiles, rockets, aircraft guns, etc., to carry out direct precision strikes on the other's targets, such as destroying enemy fighters, ground defense facilities and military bases. Bombing is primarily the dropping of bombs, missiles or other explosives on opposing targets to cause damage to their strategic and tactical objectives. Refueling is primarily used to extend the operational time and range of other warplanes through air tankers so that they can carry out missions or return to base for longer. Patrol is fighter aircraft, attack helicopters and other air targets through advanced electronic equipment and guided weapons in the easy to search, find, strike the target of the location of the guard patrol. Early warning is through the early warning aircraft equipped with advanced radar and surveillance equipment,

FIGURE 2. **Intention space.**

search, monitor each other's air or sea targets, to provide real-time intelligence and battlefield command support for their own side. Transportation is the provision of support or evacuation of equipment, material or personnel by means of transport planes or helicopters. Scout means that reconnaissance aircrafts collect information about the other side's troop deployment, position, key facilities, and terrain through radio sensors, optical sensors and other devices.

IV. AIR TARGET INTENTION RECOGNITION MODEL BASED ON 1DCNN-BiLSTM

A. DATA SEGMENTATION

Air target data is a collection of samples arranged in chronological order, where each sample represents the target state of the target at a certain point in time. The state of a target is generally portrayed by some attributes, which usually have specific distribution patterns in the time dimension and correlations between target samples at neighboring time points. In order to extract the features in the target time-series data and provide more key information for target intention recognition, the air target data is usually divided into multiple subsequences when recognizing the target intention by 1DCNN-BiLSTM.

Convolutional Neural Network (CNN) performs well on image processing tasks [17] because it exploits the spatial structure of relationships between pixels in an image. Similarly, the relationship between neighboring time points in time series data can be considered as a spatial structure. Therefore, CNN can be used to extract local features in time series. However, CNN models have a problem in processing time series data, that is, the length of the input data is fixed. Therefore, if the complete time series is directly used as input, it may lead to the problem of dimension mismatch. Moreover, the air target data is a kind of data with a relatively long time series, and the traditional CNN model may not be able to adequately capture the timing information of the whole target sequence. To solve this problem, this paper adopts time sliding window to divide the air target data into multiple subsequences. Time sliding window is a method to partition the air target data into consecutive subsequences of fixed size. With this division, the long sequence of air target data is sliced into multiple shorter subsequences, each of which contains a continuous segment of timing information from the original air target data.

Assuming that there are m target samples, the length of the sliding window is l, and the moving step of the sliding window is s, the total number of target subsequences can be divided into $\left[\frac{m-l}{s}+1\right]$, where $[\cdot]$ denotes the largest integer that does not exceed no more than itself. Then all the target sub-sequences are pooled to form a three-dimensional data structure of size $\left[\frac{m-l}{s}+1\right] \times l \times n$, where n denotes the dimension of a target sample. The process is shown in Figure 3.

Dividing air target data into multiple subsequences can not only reduce the computational overhead but also improve the model performance. First, target feature extraction can be performed on the divided target subsequence for localized regions, which helps to better capture the target temporal patterns in the sequence. Applying 1DCNN on each subsequence can extract localized features that may be locally relevant in the whole air target data. In this way, the 1DCNN-BiLSTM model can better capture the target features at different time points and learn a more meaningful representation from them. Second, segmenting the air target data into multiple subsequences can effectively reduce the number of parameters in the model. Long target sequences may require a larger model to handle, but by dividing into multiple sub-sequences, the target sequence length can be reduced, thus reducing the complexity and computational cost of the model. In addition, the temporal sliding window can help the model learn local and global patterns in the airborne target data. By extracting the local features of subsequences using 1DCNN, and then inputting these target feature sequences into BiLSTM for temporal modeling, the long-term dependencies and global patterns in the whole air target data can be captured. BiLSTM is able to memorize and utilize past information to better recognize the combat intention of air targets. In summary, the use of temporal sliding windows to divide the air target data into multiple sub-sequences is designed to better utilize temporal information, extract local features, and fully capture temporal patterns in the sequences. This segmentation method can improve the performance of the 1DCNN-BiLSTM model in recognizing the combat intention of air targets, and it improves the accuracy and generalization ability of the recognition while reducing the complexity of the model.

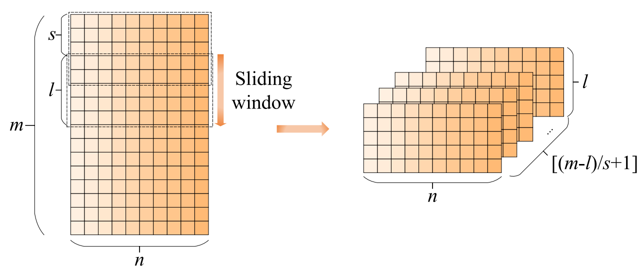


FIGURE 3. **Segmentation of target subsequences.**

B. 1DCNN LAYER

CNNs are a common class of deep learning neural networks that are widely used in computer vision [18], signal processing [19] and natural language processing [20]. There is correlation between neighboring moments in the air target data, and the one-dimensional convolution slides only in one dimension, which can not only maintain the temporal structure of the target sequence and more accurately capture the features of the target attributes in the temporal dimension, but also reduce the number of parameters of the model while maintaining the validity, thus reducing the computational cost. Therefore, we use One-dimensional Convolution Neural Network (1DCNN) to extract the target sequence features, and its operational formula is

$$\boldsymbol{Z}' = f_r\left(\sum_{w\in M} \boldsymbol{w}\cdot\boldsymbol{Z}+\boldsymbol{b}\right) \tag{3}$$

where $\boldsymbol{Z}'$ denotes the output of 1DCNN layer; f_r denotes the activation function; $\boldsymbol{w}$ denotes the convolutional kernel; M denotes the set of convolutional kernels; $\boldsymbol{Z}$ denotes the input of 1DCNN layer; and $\boldsymbol{b}$ denotes the bias.

Compared to the traditional two-dimensional convolution, one-dimensional convolution slides the convolution kernel in only one direction, so it is computationally more efficient. One-dimensional convolution is to multiply the input target sequence with the convolution kernel element by element and then sum, and slide the convolution kernel in a certain direction to loop the above operation, so as to obtain a new output sequence, as shown in Figure 4. The convolution kernel size and step size are two important hyperparameters when performing one-dimensional convolution operations [21]. The convolution kernel of 1DCNN is considered as a vector, and its size is the number of elements in the vector. The more elements in the convolution kernel, the more input sequence elements can be calculated at the same time, that is, the larger convolution kernel can extract the feature of the target in a larger field of view, while the smaller convolution kernel can extract the finer feature of the target. The step size is the displacement of the convolution kernel relative to the input target sequence during two adjacent convolution operations. If the step size of each move of the convolution kernel is too large, some target features cannot be extracted. On the contrary, if the step size is too small, the calculation amount of the model will be increased. Therefore, the size and step size of the convolution kernel affect the extraction of target features to a certain extent.

C. BiLSTM LAYER

Air target data is a kind of data arranged in chronological order. In order to capture the temporal dependency in it, Recurrent Neural Network (RNN) can be used to process it. Longer target sequences are usually input to provide more target information when recognizing the combat intention of air targets, which improves the accuracy and generalization ability of the model. However, RNNs are prone to gradient explosion or disappearance when processing longer time-series data [22] making the training process of the deep network converge slowly and unstable. In addition, since the past target information can hardly affect the current gradient update after many iterations, it also leads to the inability of the RNN to effectively learn long-distance dependencies. The Long Short-Term Memory (LSTM) network controls the flow of target information by introducing a gating mechanism [23], filters and retains important target information, and avoids unnecessary information transmission, thus effectively solving the problem of long-term dependence. Its structure is shown in Figure 5.

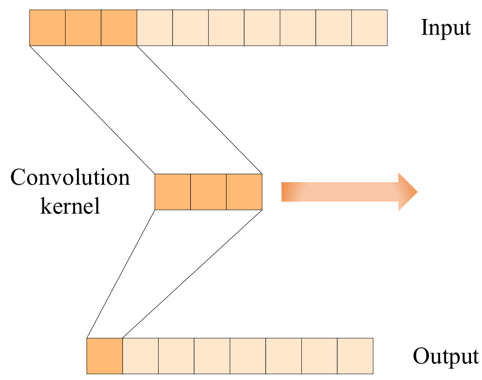


FIGURE 4. One-dimensional convolutional operation.

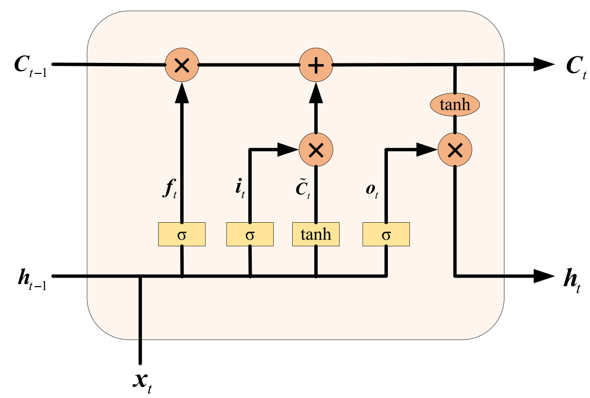


FIGURE 5. LSTM structure.

The forget gate is used to determine which target information in the cell state $\boldsymbol{C}_{t-1}$ of the previous moment needs to be discarded. The sigmoid function is utilized to map the output $\boldsymbol{h}_{t-1}$ of the previous moment and the input $\boldsymbol{x}_t$ of the current moment to the interval (0, 1), and the value indicates how much information is abandoned. The formula for the forget gate is

$$\boldsymbol{f}_t = \sigma\left(\boldsymbol{W}_f\boldsymbol{x}_t + \boldsymbol{U}_f\boldsymbol{h}_{t-1} + \boldsymbol{b}_f\right) \tag{4}$$

where $\boldsymbol{f}_t$ denotes the output of the forget gate; σ denotes the sigmoid function; $\boldsymbol{W}_f$ denotes the recurrent weight of the forget gate; $\boldsymbol{U}_f$ denotes the input weight of the forget gate; and $\boldsymbol{b}_f$ denotes the bias of the forget gate.

The input gate is used to determine what new information to add to the cell state, that is, what information in the candidate cell state $\tilde{C}_t$ needs to be saved. The sigmoid function is used to determine the information to be updated. The expression is

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right) \tag{5}$$

where i_t denotes the output of the input gate, W_i denotes the recurrent weight of the input gate; U_i denotes the input weight of the input gate; and b_i denotes the bias of the input gate.

The candidate cell state $\tilde{C}_t$ is obtained through the tanh function, and the expression is

$$\tilde{C}_t = \tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \tag{6}$$

where W_c denotes the recurrent weight of the layer; U_c denotes the input weight of the layer; and b_c is the bias of the layer.

The forget gate f_t is used to discard the cell state C_{t-1} at the previous moment, and the input gate i_t is used to determine the state of the candidate cells $\tilde{C}_t$ for preservation, and the cell state C_t at the current moment is obtained. The expression is

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{7}$$

where $\odot$ denotes point-wise multiplication.

The output gate is used to determine the output at the current time, and its expression is

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{8}$$

where o_t denotes the output of the output gate, W_o denotes the recurrent weight of the output gate, U_o denotes the input weight of the output gate; and b_o denotes the bias of the output gate.

The final output is obtained through the output gate o_t and the cell state C_t at the current moment with the expression

$$h_t = o_t \odot \tanh\left(C_t\right) \tag{9}$$

Through the above analysis, LSTM processes the target sequence in chronological order. However, in the battlefield, the state change of the target is a continuous process, and the target state at the current moment is not only related to the state at the previous moment, but also the target state at the moment afterward is a continuation of the state at the current moment. Therefore, the target intention is not only related to the data of the previous moment, but also correlated with the data of the subsequent moments, and unidirectional LSTM cannot accurately capture the feature information in the target sequence. To address this problem, we adopt Bidirectional Long Short-Term Memory (BiLSTM) to learn the dependencies between the attributes in the target sequence at the previous and subsequent moments, so as to extract the target features more comprehensively to recognize its combat intention. BiLSTM contains two LSTM layers for forward and reverse passes of the target sequence [24], and its structure is shown in Figure 6.

BiLSTM consists of forward LSTM and reverse LSTM. For any moment t, whose input target sequence is x_t, the update formula for the forward hidden state $\vec{h}_t$ and the reverse hidden state $\overleftarrow{h}_t$ is

$$\vec{h}_t = g\left(W_{xh}^{(f)} x_t + W_{hh}^{(f)} \vec{h}_{t-1} + b_h^{(f)}\right) \tag{10}$$

$$\overleftarrow{h}_t = g\left(W_{xh}^{(b)} x_t + W_{hh}^{(b)} \overleftarrow{h}_{t-1} + b_h^{(b)}\right) \tag{11}$$

where g denotes the activation function; $W_{xh}^{(f)}$ and $W_{xh}^{(b)}$ denote the weights from the input layer to the forward and reverse LSTM hidden layers, respectively; $W_{hh}^{(f)}$ and $W_{hh}^{(b)}$ denote the weights from the forward LSTM layer to the forward LSTM layer and from the reverse LSTM layer to the reverse LSTM layer, respectively; and $b_h^{(f)}$ and $b_h^{(b)}$ denote the bias of each part.

By connecting the forward hidden state $\vec{h}_t$ and the reverse hidden state $\overleftarrow{h}_t$, the hidden state h_t is obtained, and then it is input to the output layer to calculate the output y_t of BiLSTM, the formula is

$$y_t = W_{ho} h_t + b_o \tag{12}$$

where W_{ho} and b_o denote the weights and bias of the BiLSTM output layer, respectively.

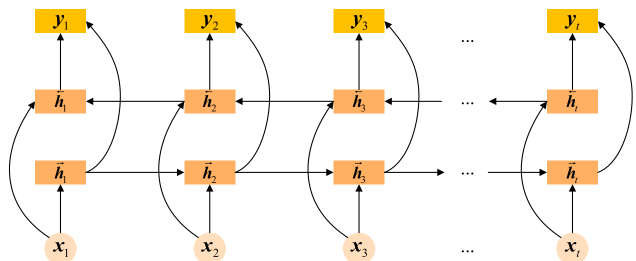


FIGURE 6. BiLSTM structure.

D. PROPOSED MODEL

Air target intention recognition (ATIR) is a process of analyzing target state data in order to identify the combat mission performed by the target. Among them, target state data are multi-dimensional data arranged in chronological order. In order to be able to extract the features in the target data more efficiently and to realize end-to-end intention recognition, we propose a combat intention recognition model of air targets based on 1DCNN-BiLSTM, whose flow is shown in Figure 7. Firstly, the target data is preprocessed, including coding, denoising and normalization. Secondly, the target data is divided into sub-sequences by sliding window, and based on this, the target sequences are divided into training set, validation set and test set respectively in the ratio of 6:2:2. Thirdly, the training and validation sets are input into the 1DCNN module, and the local features of the air target data can be extracted by convolution operation, and then the multi-dimensional target data is one-dimensionalized by Flatten layer, and the Repeeat Vector layer is to connect

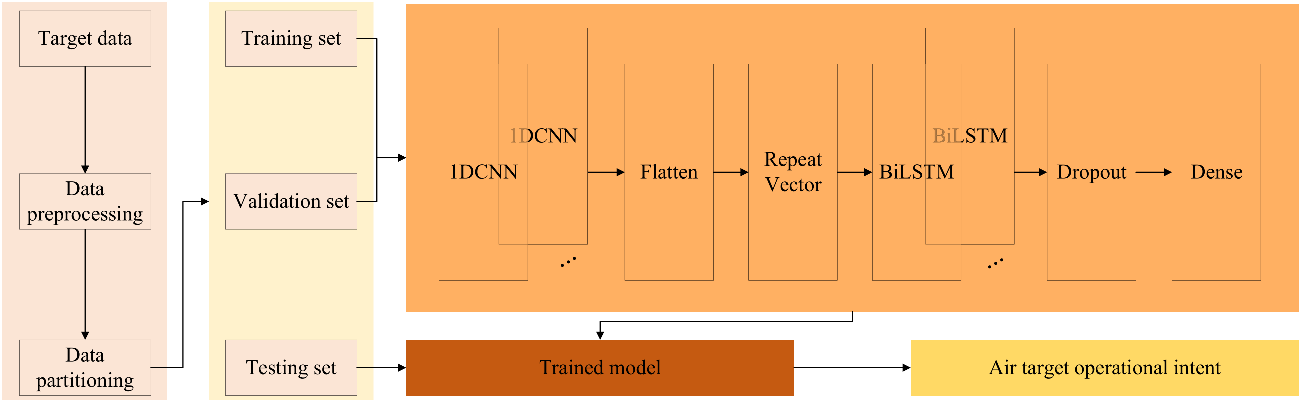


FIGURE 7. Air target intention recognition process based on 1DCNN-BiLSTM.

the 1DCNN and BiLSTM by repeatedly inputting the target sequences, where BiLSTM is able to capture the dependencies in the target data in both forward and reverse directions, followed by randomly deleting neurons through the Dropout layer to prevent overfitting of the model, using the Dense layer as the output. Finally, the test set is fed into the trained model to recognize the combat intention of air targets.

V. EXPERIMENTAL ANALYSIS

A. EXPERIMENTAL DATA AND ENVIRONMENT

The experimental data come from the air battlefield posture simulation system. Against the background of airborne multi-service joint operations between the two warring parties, various types of air targets, such as fighter planes, transport planes, early warning planes and reconnaissance planes, are simulated to carry out the seven types of combat missions identified in Section III in an air battle, so as to obtain the state of each target. The target dataset constructed is a two-dimensional time series data, where each row is a sample and each column contains 10 attributes such as heading angle, azimuth, altitude, speed, etc., and the dataset has a total of 23,530 samples. Among all the target data, the proportion of operational intent is attack is 14.72%, bombing is 14.68%, refueling is 13.20%, patrol is 14.23%, early warning is 14.45%, transport is 14.39%, and scout is 14.33%. The data was divided into training set, validation set and test set in the ratio of 6:2:2.

We programmed in Python 3.8 on a 64-bit Windows 10 computer with the Keras 2.3.1 deep learning framework, an 11th Gen Intel(R) Core(TM) i7-1165G7 @ 2.80 GHz processor, and 16.0 GB of RAM, and the specific experimental environment is shown in Table 1.

B. DETERMINING THE MODEL STRUCTURE

The 1DCNN-BiLSTM model is mainly composed of 1DCNN module and BiLSTM module, in order to completely build the structure of the model, it is necessary to determine the structure of the model, mainly including the number of layers of 1DCNN and its number of convolution kernels, the number of layers of BiLSTM and its number of neurons.

TABLE 1. Experimental environment.

Environment	Version
Operating system	Windows 10 x64
RAM	16.0 GB
CPU	i7-1165G7
Programming language	Python 3.8
Deep learning framework	Keras 2.3.1
Main libraries	NumPy 1.19.1, pandas 1.4.2, Scikit-learn 1.0.2

Firstly, the number of 1DCNN layers is determined, and the layers of 1DCNN are set as 1, 2, 3, 4 and 5 respectively on the basis of one BiLSTM layer. As shown in Figure 8(a), the complex patterns and features in the time series are not adequately captured when there is only one 1DCNN layer, which leads to the lowest accuracy in the test set; when the layers of 1DCNN are 2, 3 and 5 respectively, the accuracy is the same, and the model performance is enhanced in comparison to the one with only one 1DCNN layer; however, the over-fitting problem may occur when there are too many layers. The performance of the model is best when there are 4 1DCNN layers, so the number of 1DCNN layers is determined to be 4. Secondly, the number of convolutional kernels in each 1DCNN layer is determined, and 8, 16, 32, 64 and 128 convolutional kernels are set in each layer based on setting four 1DCNN layers, and the results are shown in Fig. 8(b), when the number of convolutional kernels is 8 and 16 the accuracy of the test set is comparable, and the model's performance is optimal when the number of convolutional kernels is 32. This is due to the fact that as the number of convolutional kernels increases the model is able to learn richer and more complex features, but then continue to increase the number of convolutional kernels is prone to overfitting phenomenon leading to a gradual decline in performance, so the number of convolutional kernels for each 1DCNN layer is determined to be 32. Thirdly, the number of BiLSTM layers is determined. On the basis

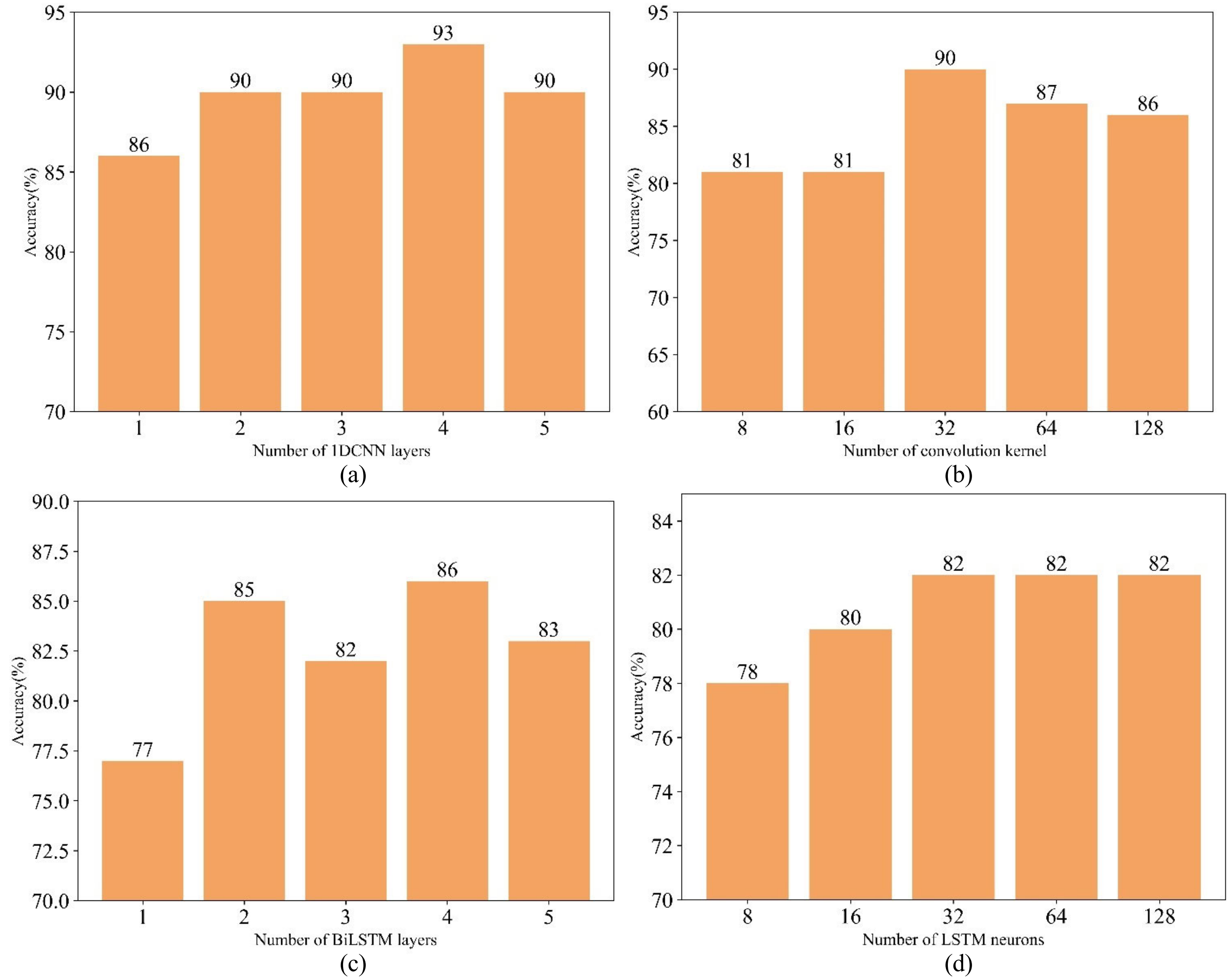


FIGURE 8. Performance of different model structures. (a) Performance of different 1DCNN layer numbers. (b) Performance of different number of convolutional kernels. (c) Performance of different BiLSTM layer numbers. (d) Performance of different numbers of LSTM neurons.

of setting four 1DCNN layers and 32 convolution kernels in each layer, 1, 2, 3, 4 and 5 BiLSTM layers are set respectively. The results are shown in Figure 8(c), the model's performance is the worst when there is only one BiLSTM layer, and the accuracy of the test set is the highest when there is a set up of four BiLSTM layers, so the number of BiLSTM layers is determined to be 4. Finally, the number of LSTM neurons in each layer of BiLSTM is determined, and the number of LSTM neurons in each layer is set as 8, 16, 32, 64 and 128 on the basis of the aforementioned hyperparameters. The results are shown in Figure 8(d). When the number of LSTM neurons is 32, 64 and 128, the performance of the model is the best and equivalent. Considering the operation cost, the number of LSTM neurons in each BiLSTM layer is determined to be 32. This is due to the fact that increasing the number of BiLSTM layers and the number of neurons in the LSTM can enhance the expressive ability of the model to learn the long-term dependencies in the target data, but when the number of layers and neurons is too high, it may lead to the disappearance of the gradient thus affecting the recognition performance of the model. In addition, the batch size is determined to be 32 according to the size of the dataset; in the above experiments when epoch is 60, the accuracy and loss value have converged, so epoch is determined to be 60; considering that recognizing the target's combat intention is a multi-label classification problem, categorical crossentropy is selected as the loss function, and softmax as the activation function of the output layer; in order to improve the convergence speed of the model, Adam is selected as the optimizer, Relu as the activation function of the 1DCNN layer, and the learning rate is 0.001; in order to avoid overfitting of the model, the Dropout layer is introduced and the probability is 0.5. The hyperparameter settings of the model were determined as shown in Table 2.

C. EVALUATION METRICS

In order to evaluate the intention recognition performance of the model, in addition to the Accuracy, we also adopt the Precision, Recall and F1-Score to measure the performance

TABLE 2. Model hyperparameter settings.

Parameters	Value
Number of 1DCNN layers	4
Number of convolutional kernel	32
Number of BiLSTM layers	4
Number of LSTM neurons	32
Batch Size	32
Epochs	60
Loss	Categorical crossentropy
Optimizer	Adam
Learning rate	0.001
Dropout	0.5
Activation(1DCNN)	Relu
Activation(output)	softmax

of the 1DCNN-BiLSTM model in recognizing the combat intention of air targets.

Accuracy is the proportion of all target samples for which the combat intention is correctly recognized, and is calculated as follows

$$\text{Accuracy} = \frac{a}{A} \times 100\% \tag{13}$$

where a denotes the number of target samples that are correctly identified with combat intention; A denotes the number of all target samples.

For target intention recognition, the target data can be divided into True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) based on its true combat intention and combat intention obtained from the recognition [25], and the confusion matrix is shown in Table 3.

TABLE 3. Confusion matrix.

Actual Intention	Predicted Intention	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

Precision [26] represents the proportion of the target samples identified as an intention whose true intention is indeed the intention, and is calculated as

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{14}$$

Recall [26] represents the proportion of samples in which an intention is correctly identified in all the intention samples. The formula is

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{15}$$

Generally, Precision and Recall are conflicting metrics with some negative correlation between them. However, it is necessary to accurately recognize each combat intention in air combat to reduce misjudgment. Therefore, F1-Score is introduced, which integrates the factors of both Precision and Recall [26], and is calculated as

$$\text{F1-Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{16}$$

In addition, the intention space contains 7 types of combat intentions, and in order to more accurately evaluate the recognition performance of the model for all combat intentions, macro-Precision, macro-Recall and macro-F1 are introduced [27]. The calculation formulas are respectively

$$\text{macro-Precision} = \frac{1}{I}\sum_{i=1}^{I} \text{Precision}_i \tag{17}$$

$$\text{macro-Recall} = \frac{1}{I}\sum_{i=1}^{I} \text{Recall}_i \tag{18}$$

$$\text{macro-F1} = \frac{2 \times \text{macro-Precision} \times \text{macro-Recall}}{\text{macro-Precision} + \text{macro-Recall}} \tag{19}$$

where I denotes the number of combat intentions.

D. COMPARATIVE EXPERIMENT

In order to verify the performance of the 1DCNN-BiLSTM model in recognizing the combat intention of air targets, a comparative experiment is designed. The radial basis function (RBF) neural network target intention recognition method proposed by Wei and Wang [28], the fully connected neural network (FCNN) intention recognition method proposed by Zhou et al. [8], [11], the RNN intention recognition method proposed by Wang [7], and its related variant models, one-dimensional convolutional neural network and recurrent neural network (1DCNN-RNN), bidirectional recurrent neural network (BiRNN), and 1DCNN-BiRNN, are compared with the methods proposed in this paper. The data set constructed in section V- subsection A is utilized to conduct comparative experiments on the above methods and the results are shown in Table 4.

TABLE 4. Results of comparative experiment.

Model	Accuracy (%)	Macro-Precision(%)	Macro-Recall(%)	Macro-F1(%)
RBF	63	63	62	61
FCNN	84	85	82	83
RNN	73	74	75	71
1DCNN-RNN	88	90	87	87
BiRNN	84	83	84	83
1DCNN-BiRNN	89	88	87	87
1DCNN-BiLSTM	**93**	**93**	**92**	**93**

As can be seen from Table 4, since RBF consists of only one layer of radial basis function and one output layer, its feature extraction capability is weak, so the performance of airborne target intent recognition based on RBF model is the worst among all the models, and its accuracy is only 63%. The performance for target intention recognition of the RNN model is higher than that of the RBF but lower than the

TABLE 5. Results of ablation experiment.

Model	Accuracy (%)	Macro-Precision(%)	Macro-Recall(%)	Macro-F1(%)
CNN	76	74	73	73
LSTM	83	81	81	81
BiLSTM	85	87	84	84
1DCNN-LSTM	87	88	87	87
1DCNN-BiLSTM	**93**	**93**	**92**	**93**

TABLE 6. Air target combat intention recognition results.

	Precision(%)	Recall(%)	F1(%)
Attack	90	96	93
Bombing	86	92	89
Refuel	91	99	95
Patrol	100	88	94
Early Warning	94	94	94
Transport	90	84	87
Scout	100	95	98
accuracy			93
macro avg	93	92	93

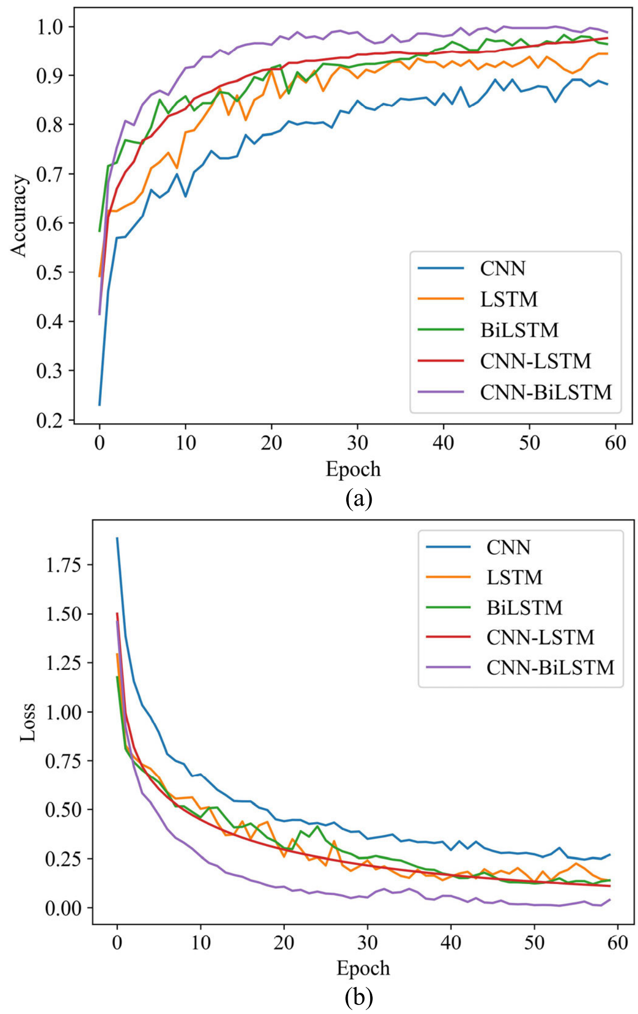


FIGURE 9. Changes in accuracy and loss values during training. (a) Variation curve of accuracy. (b) Loss curve of accuracy.

performance of the FCNN, which is due to the fact that the RNN suffers from gradient vanishing or gradient explosion when processing the target time-series data, which results in poorer performance. However, BiRNN, an improved model of RNN, is able to process two forward and reverse target information at the same time, so as to mitigate the effects of gradient disappearance and gradient explosion, and its performance is improved by about 14.03% on average compared to RNN in recognizing the operational intent of air targets. Another improved model of RNN, 1DCNN-RNN, extracts the features in the target sequence through convolution operation to reduce the processing difficulty of RNN, and can also improve the model's intention recognition performance by about 5.40% compared with BiRNN. Combining the two, the performance of the 1DCNN-BiRNN model is comparable to that of the 1DCNN-RNN. Combining 1DCNN-RNN and BiRNN, the performance of 1DCNN-BiRNN model is comparable to that of 1DCNN-RNN. Our proposed 1DCNN-BiLSTM model all outperforms the above models, with an average improvement of about 17.77% for Accuracy, 17.30% for Macro-Precision, 17.38% for Macro-Recall, and 20.22% for Macro-F1.

E. ABLATION EXPERIMENT

The 1DCNN-BiLSTM model proposed in this paper is compared with RBF, FCNN, RNN, 1DCNN-RNN, BiRNN, and 1DCNN-BiRNN in the above, and the result fully proves that the model has high performance in recognizing the combat intention of the air targets. In order to verify the effectiveness of each part of the model, ablation experiments are conducted here and the results are shown in Table 5. As can be seen from the table, the Accuracy of the 1DCNN-BiLSTM model proposed in this paper is 22.37%, 12.05%, 9.41% and 6.90% higher than that of CNN, LSTM, BiLSTM and 1DCNN-LSTM respectively. Macro-Precision is improved by 25.68%, 14.81%, 6.90% and 5.68%, and Macro-Recall is improved by 26.03%, 13.58%, 9.52% and 5.75%, respectively. Macro-F1 is 27.40%, 14.81%, 10.71% and 6.90% better than the above models. Although the CNN model can continuously reduce the length of the processed sequence by extracting the features of the target sequence through convolutional operations, it still cannot handle the forward and backward dependencies in the target sequence well, while LSTM has better performance in processing long target sequences, so its performance is worse than that of the LSTM model in recognizing the target intention. However, compared to LSTM, BiLSTM is able to capture the dependencies of the target sequence in both forward and reverse directions, and its performance is improved by 4.31% on average.

Combining 1DCNN with LSTM, 1DCNN-LSTM effectively extracts the target data features as well as captures the target sequence dependencies in both forward and backward directions, improving the performance of LSTM by 7.07% on average. The 1DCNN-BiLSTM model proposed in this paper combines 1DCNN and bi-directional structure with LSTM at the same time, and is able to capture more target sequence context information on top of 1DCNN-LSTM, which makes the model able to capture more complex target sequence features, and improves the performance of the model by an average of 13.81% over the traditional LSTM model.

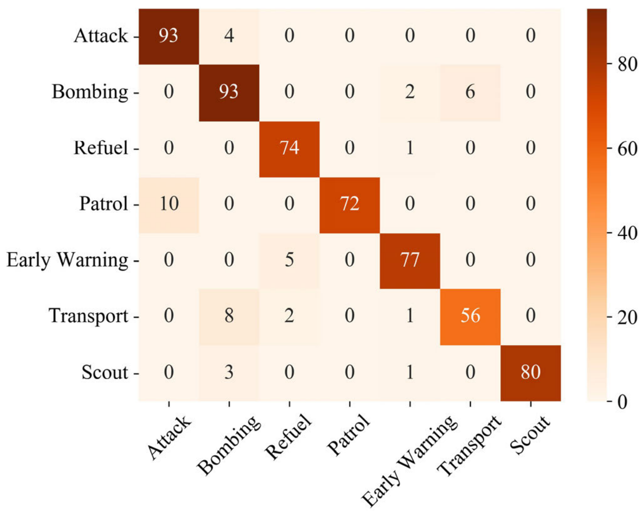


FIGURE 10. Confusion matrix for air target combat intention recognition results.

The variation of accuracy and loss value of the above models with Epoch during training is shown in Figure 9, from which it can be seen that the 1DCNN-BiLSTM model proposed in this paper has the optimal convergence value of both accuracy and loss value with Epoch, followed by 1DCNN-LSTM, BiLSTM and LSTM, and the CNN performs the worst. In addition, in terms of convergence speed, 1DCNN-BiLSTM starts to converge around Epoch of 15, which is faster than the other models.

F. COMBAT INTENTION RECOGNITION OF AIR TARGETS

The 1DCNN-BiLSTM model is used to recognize the combat intent of air targets, and the confusion matrix of the recognition results is shown in Figure 10. From the figure, it can be seen that it is easy to confuse when recognizing the two combat intentions of bombing and transportation, and the model is easy to misjudge as attack when the air target performs the patrol mission, which is mainly due to the fact that the aerial target's state is relatively close to that of the two when it performs the relevant combat mission.

VI. CONCLUSION

Air target combat intention identification is a key part of air battlefield situation assessment. We analyze the air target intent recognition problem in detail and construct its mathematical model. On this basis, combined with the actual situation of intention recognition and the characteristics of air target data analysis, the 10-dimensional feature space is determined and the intention space is encoded and packaged. In order to solve the current problem that long sequence target data can not be processed efficiently, for the air target data with the characteristics of long time series, trend and regularity, this paper proposes an end-to-end model based on 1DCNN-BiLSTM for the recognition of combat intention of air targets. Among them, the 1DCNN module captures the features of target attributes in the time dimension, while reducing the number of parameters of the model to lower the computational cost; BiLSTM is able to learn the dependencies between attributes in the target sequence in the previous and subsequent moments, so as to extract the target features more comprehensively. The ablation experiments also verify the performance of each module in the model. In the comparative experiments, the 1DCNN-BiLSTM model shows a great improvement in the recognition performance compared to other common models, and at the same time, the convergence speed is faster. Therefore, the air target combat intention recognition model we proposed is not only able to handle high-dimensional target data, but also has the advantages of good recognition performance, fast convergence and high reliability. However, due to the limitation of the architecture of the model proposed in this paper, this paper only aims at recognizing the intention of a single target, and cannot recognize the operational intention of a group of targets. The research of this paper can provide effective support for air battlefield situation assessment and auxiliary information for command and decision-making. In the next step, we will improve the model architecture based on the single-target intention recognized in this paper and combine it with other situational information to complete the recognition of air group targets.

REFERENCES

[1] Y. He, L. L. Chang, J. Jiang, and Y. J. Tan, "Intension identification in air defense based on belief rule base expert system under expert guidance," *Fire Control Command Control*, vol. 42, no. 9, pp. 7–12, 2017.

[2] D. L. Li, Y. Liu, H. L. Xu, Y. Zhang, and D. D. Li, "Under water target intent analysis method applied to target tracking decision," *Comput. Appl. Softw.*, vol. 34, no. 9, pp. 188–191, 2017.

[3] X. P. Wang, M. H. Xia, Q. Y. Lin, Z. Wang, and F. E. Kong, "Combat intent forecast based on D-S evidence theory before contacting the enemy," *Fire Control Command Control*, vol. 41, no. 9, pp. 185–188, 2016.

[4] S. Y. Cao, Y. A. Liu, and S. Xue, "Target tactical intention recognition method of improved high-dimensional data similarity," *Transducer Microsyst. Technol.*, vol. 36, no. 5, pp. 25–28, 2017, doi: 10.13873/j.1000-9787(2017)05-0025-04.

[5] H. M. Chai and B. S. Wang, "Application of dynamic Bayesian networks in tactical situation assessment," *Appl. Res. Comput.*, vol. 28, no. 6, pp. 2151–2153, 2011.

[6] H. P. Leng, X. F. Wu, J. G. Hu, and Y. Q. Yu, "Study on sequential recognition technique of marine targets' tactical intentions," *Syst. Eng. Electron.*, vol. 30, no. 3, pp. 462–465, 2008.

[7] C. Wang, "Research of situation assessment and threat assessment algorithms based on big data processing technology," M.S. thesis, Univ. Electron. Sci. Technol. China, Chengdu, China, 2019.

[8] W. W. Zhou, P. Y. Yao, Y. J. Zhang, X. Wang, and S. Wei, "Combat intention recognition for aerial targets based on deep neural network," *Acta Aeronautica et Astronautica Sinica*, vol. 39, no. 11, pp. 200–208, 2018.

[9] L. Yang, F. X. Liu, F. Zhu, and D. Guo, "Hierarchical recognition method of hostile air-targets in sea battlefields based on Bayesian deduction," *Fire Control Command Control*, vol. 43, no. 7, pp. 86–93, 2018.

[10] A. A. Ahmed and M. F. Mohammed, "SAIRF: A similarity approach for attack intention recognition using fuzzy min-max neural network," *J. Comput. Sci.*, vol. 25, pp. 467–473, Mar. 2018, doi: 10.1016/j.jocs.2017.09.007.

[11] W. W. Zhou, J. Y. Zhang, N. N. Gu, and G. Q. Yan, "Recognition of combat intention with insufficient expert knowledge," in *Proc. 3rd Int. Conf. Comput. Model., Simul. Appl. Math.*, Wuhan, China, 2018, pp. 316–321.

[12] Z. Zhang, H. Wang, J. Geng, W. Jiang, X. Deng, and W. Miao, "An information fusion method based on deep learning and fuzzy discount-weighting for target intention recognition," *Eng. Appl. Artif. Intell.*, vol. 109, Mar. 2022, Art. no. 104610, doi: 10.1016/j.engappai.2021.104610.

[13] Y. Zhang, F. Huang, X. Deng, M. Li, and W. Jiang, "Air target intention recognition and causal effect analysis combining uncertainty information reasoning and potential outcome framework," *Chin. J. Aeronaut.*, 2023. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1000936123003126

[14] X. Xue, S. Huang, and D. Wei, "Combined AGADESN with DBSCAN algorithm for cluster target motion intention recognition," *Int. J. Aerosp. Eng.*, vol. 2022, Oct. 2022, Art. no. e8220029, doi: 10.1155/2022/8220029.

[15] F. Teng, Y. Song, G. Wang, P. Zhang, L. Wang, and Z. Zhang, "A GRU-based method for predicting intention of aerial targets," *Comput. Intell. Neurosci.*, vol. 2021, Nov. 2021, Art. no. 6082242, doi: 10.1155/2021/6082242.

[16] Y. L. Zhang, N. N. Liu, and Z. W. Wang, "Target grouping and target motion state prediction based on neural network," *Shipboard Electron. Countermeasure*, vol. 43, no. 3, pp. 7–12, 2020, doi: 10.16426/j.cnki.jcdzdk.2020.03.002.

[17] D. Bhatt, C. Patel, H. Talsania, J. Patel, R. Vaghela, S. Pandya, K. Modi, and H. Ghayvat, "CNN variants for computer vision: History, architecture, application, challenges and future scope," *Electronics*, vol. 10, no. 20, p. 2470, Oct. 2021, doi: 10.3390/electronics10202470.

[18] W. Fang, P. E. D. Love, H. Luo, and L. Ding, "Computer vision for behaviour-based safety in construction: A review and future directions," *Adv. Eng. Informat.*, vol. 43, Jan. 2020, Art. no. 100980, doi: 10.1016/j.aei.2019.100980.

[19] Mustaqeem and S. Kwon, "A CNN-assisted enhanced audio signal processing for speech emotion recognition," *Sensors*, vol. 20, no. 1, p. 183, Dec. 2019, doi: 10.3390/s20010183.

[20] A. K. Sharma, S. Chaurasia, and D. K. Srivastava, "Sentimental short sentences classification by using CNN deep learning model with fine tuned Word2Vec," *Proc. Comput. Sci.*, vol. 167, pp. 1139–1147, Jan. 2020, doi: 10.1016/j.procs.2020.03.416.

[21] S. Huang, J. Tang, J. Dai, and Y. Wang, "Signal status recognition based on 1DCNN and its feature extraction mechanism analysis," *Sensors*, vol. 19, no. 9, p. 2018, Apr. 2019, doi: 10.3390/s19092018.

[22] A. Sherstinsky, "Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network," *Phys. D, Nonlinear Phenomena*, vol. 404, Mar. 2020, Art. no. 132306, doi: 10.1016/j.physd.2019.132306.

[23] Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: LSTM cells and network architectures," *Neural Comput.*, vol. 31, no. 7, pp. 1235–1270, Jul. 2019, doi: 10.1162/neco_a_01199.

[24] Z. Hameed and B. Garcia-Zapirain, "Sentiment classification using a single-layered BiLSTM model," *IEEE Access*, vol. 8, pp. 73992–74001, 2020, doi: 10.1109/ACCESS.2020.2988550.

[25] D. Chen, Y. Lu, and C.-Y. Hsu, "Measurement invariance investigation for performance of deep learning architectures," *IEEE Access*, vol. 10, pp. 78070–78087, 2022, doi: 10.1109/ACCESS.2022.3192468.

[26] B. Mirza, D. Haroon, B. Khan, A. Padhani, and T. Q. Syed, "Deep generative models to counter class imbalance: A model-metric mapping with proportion calibration methodology," *IEEE Access*, vol. 9, pp. 55879–55897, 2021, doi: 10.1109/ACCESS.2021.3071389.

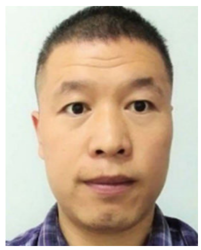
[27] S. Memis, S. Enginoglu, and U. Erkan, "Numerical data classification via distance-based similarity measures of fuzzy parameterized fuzzy soft matrices," *IEEE Access*, vol. 9, pp. 88583–88601, 2021, doi: 10.1109/ACCESS.2021.3089849.

[28] W. Wei and G. B. Wang, "Detection and recognition of air targets by unmanned aerial vehicle based on RBF neural network," *Ship Electron. Eng.*, vol. 38, no. 10, pp. 37–40, 2018.

CHENHAO ZHANG was born in Xiangyang, Hubei, China, in 1996. He received the M.S. degree in information and communication engineering from the Air Force Early Warning Academy, Wuhan, China, in 2020, where he is currently pursuing the Ph.D. degree in information and communication engineering. His current research interests include situational assessment, data mining, and artificial intelligence.

YAN ZHOU received the Ph.D. degree in control science and engineering from the Huazhong University of Science and Technology, Wuhan, China, in 2000. He is currently a Professor with the Air Force Early Warning Academy, Wuhan. His current research interests include pattern recognition, data fusion, and image processing.

HUI LI received the M.S. degree from the Beijing University of Technology, in 2007. She is currently a Professor with the Air Force Early Warning Academy, Wuhan. Her current research interests include target tracking and multi-source information fusion.



FUTAI LIANG received the master's degree in information and communication engineering from the Air Force Early Warning Academy, Wuhan, China, in 2020, where he is currently pursuing the Ph.D. degree in information and communication engineering. His current research interests include target detection, image recognition, and AI.

ZIHAO SONG received the master's degree in information and communication engineering from the Air Force Early Warning Academy, Wuhan, China, in 2021, where he is currently pursuing the Ph.D. degree in information and communication engineering. His current research interests include signal processing, information fusion, and AI.

KAI YUAN received the master's degree in information and communication engineering from the Air Force Early Warning Academy, Wuhan, China, in 2022, where he is currently pursuing the Ph.D. degree in information and communication engineering. His current research interests include signal processing, information fusion, and AI.