Learning Embedding Features Based on Multisense-Scaled Attention Architecture to Improve the Predictive Performance of Air Combat Intention Recognition

XUHUA WANG1, ZHIKAI LIN2, YAHUI HU2, AND JIANWEI LIU2

1Xidian University, Xi’an 710126, China
2Air Force Early Warning Academy, Wuhan 430014, China

Corresponding author: Yahui Hu (hyh5800@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61502522 and Grant 61502523.

ABSTRACT In modern air combat, acquiring the opponent’s air combat intention is one of essential prerequisites to evaluate the air combat situation effectively and master the battlefield initiative. On account of multi-dimensional and temporal characteristics of the target state, a recognition model is proposed to identify tactical intention of aerial target based on a multi-sense-scaled attention architecture. First of all, the multi-dimensional feature information, including target state attributes, battlefield environment, target attributes and so on, is constructed as target feature. Secondly, the non-numerical information, such as battlefield environment characteristics, enemy and friend attributes of targets, radar status, maneuver type, etc, is transformed into numerical data. For the purpose of subsequent data processing, the flight speed, altitude, RCS, etc, in the target status information are normalized into the same dimension. Furthermore, a target intention recognition model with multiple sense-scaled attention mechanism is designed to depict the target state, attributes and the information of the battlefield environment from multiple dimensions, which is convenient to be close to the actual combat. The BiLSTM neural network is used to learn the deep-seated information in the air combat intention feature vector, and the attention mechanism is used to adaptively allocate the network weights. The air combat feature information with different weights is introduced into the softmax function layer for intention recognition. Compared with the traditional air tactical target intention recognition model, the proposed model effectively improves the efficiency of air target tactical intention recognition as well as affords important theoretical significance and reference value for the auxiliary combat system.

INDEX TERMS Target intention recognition, deep learning neural network, multi-sense-scaled attention, long short-term memory neural network.

I. INTRODUCTION
In modern air combat, combat intention recognition is a complex battlefield situation recognition problem [1], which is always the focus of commanders at all levels, and the basis for deciding the next combat action. It needs to consider the uncertainty or discreteness of the environment, which is also an important realization mechanism of “know yourself as well as the enemy” [2].

In the increasingly complex air combat battlefield, target intention recognition is the key to anticipate the enemy’s initiative and control the change of air combat situation, and has guiding significance for target allocation, mobility decision and firepower decision. At present, scholars at home and abroad have done a lot of research on target intention prediction. The commonly used methods include Bayesian...
network [3], [4], evidential reasoning [5], discriminant analysis [6], expert system [7], [8], decision trees [9], D-S evidence theory [10], [11], time window [12], support vector machines [13], etc. Guo et al. [3] proposed a probabilistic model of information fusion for target recognition under the background of information fusion in electronic warfare with the basic structure of Naïve Bayes classifier and augmented Naïve Bayes classifier. Zhao et al. [5] used the expert system of belief rules to establish intention recognition rules, and evidential reasoning was mainly used for the reasoning synthesis of rules. Yin et al. [6] utilized the Fisher discriminant and Bayes discriminant to establish target intention recognition rules. Ben-Bassat and Freedy [7] constructed an expert decision support system to evaluate military situation. Zhou et al. [8] designed a model of insufficient expert knowledge with expert system. Niu et al. [9] applied the decision tree to the research of intention recognition of naval vessel. Xia [10] analyzed the situation repository and event relationship of intention assessment reasoning and designed an intention assessment reasoning frame based on D-S evidence theory. Ma et al. [12] introduced a time window index to model and analyze the time performance of surface ships to air combat deployment. Wang et al. [13] proposed a real-time detection algorithms based on support vector machine to recognize the target intention of unmanned aerial vehicle. Dai et al. [14] proposed a method of target tactical intention recognition based on interval grey relational degree in view of the confusing characteristics of the state attributes of the enemy aircraft at a certain time.

In fact, the target intention is implemented through a series of tactical actions in the battlefield, so the dynamic attributes of the target and the battlefield environment show dynamic and time-series characteristics. Therefore, it is incomplete to infer the enemy’s target intention based on the characteristic information at a single time. With the development of artificial intelligence [15], [16], data fusion [17], [18] and deep learning [19], [20], [21], [22], the computer can handle a large number of complex data with high performance and high speed.

Kong et al [21] proposed a sensor placement methodology of hydraulic control system to improve the performance of fault diagnosis of the hydraulic control system. Cai et al [22] proposed an artificial intelligence enhanced reliability assessment methodology to improve the accuracy of the reliability assessment of a test product. Zhou et al. [23] introduced the rectified linear unit activation function and the adaptive moment estimation optimization algorithm to improve the convergence speed of the model and effectively prevent the algorithm from falling into a local optimum. Li et al. [24] used the LSTM network to train the target feature vector, and enhanced the feature learning of the target with attention mechanism. Zhou et al. [25] designed an intention prediction method with LSTM network and decision tree to improve the effectiveness of air combat decision-making systems. Qiao et al. [26] constructed the mapping relationship between multi-domain operation tactical rule base and scene situation and achieved the group target intention recognition with a multi-entity hierarchical Bayesian network. Teng et al. [27] used a hierarchical method to construct the feature of air combat intention and improved the recognition efficiency of air target tactical intention. Xue et al. [28] designed a panoramic convolutional long short-term memory networks (PCLSTM) to improve the recognition ability. Zhou et al. [29] proposed an attention-based bidirectional long short-term memory networks to capture the most important semantic information in a sentence in the field of natural language processing. Zhang et al. [30] introduced an emotional integration through the Bi-directional neural network to improve the accuracy of sentiment classification.

These models have good effects on identifying the target’s combat intention, and conform to the time sequence characteristics and the logical relationship between the front and back in the battlefield situation information. However, the above methods still have shortcomings in time feature learning and knowledge representation. Aiming at the above limitations, in order to effectively identify the combat target intention, this paper proposes a multi-sense-scaled attention architecture to improve the predictive performance of air combat intention recognition. Specifically, the main contributions are three-fold:

1. Build the model space from three target feature information: battlefield environment, target attributes and target status. Among them, the battlefield environment of air combat mainly involves meteorological environment, information warfare environment, electromagnetic environment, and so on. Target attributes mainly include target radar status and type. Target status mainly includes velocity, height, radar cross-section (RCS), etc.

2. Preprocess the original data set. Normalize the flight velocity, height, RCS and other dimensional data. The non-numerical data such as the environment, motivation type, radar status, and so on, are transformed into numerical data.

3. Design a multi-sense-scaled attention model. Based on the LSTM network, the bidirectional circulation mechanism and attention mechanism are introduced to simulate the reasoning process of decision-makers for air combat.

The rest of this article is organized as follows. Section 2 briefly introduces the problem description of combat intention recognition including intention space description and target feature space. Section 3 reviews the overall framework of our model including data normalization, model framework, algorithm flow and so on. Section 4 further explores the simulation and analysis. Finally, we present concluding remarks in section 5.

II. PROBLEM DESCRIPTION

Target tactical intention identification is to analyze and identify the operational tasks performed by the target in the process of complex air combat, which belongs to the category of confrontation identification [23]. It is a process of inferring the operational intention of enemy targets from the real-time and confrontational environment by extracting the battlefield...
environment information in the corresponding space-time domain, analyzing the static attributes and real-time dynamic information of air combat enemy and our targets, and combining the corresponding military domain knowledge. Therefore, as shown in Fig.1.

A. TARGET FEATURE SPACE

According to the three target information feature sets of battlefield environment, target attribute and target status, the target feature space mainly involved in complex air combat countermeasure is shown in Fig.2.

Battlefield environment mainly involves electromagnetic environment, meteorological environment, information combat environment, etc. The electromagnetic environment involves the electromagnetic interference of the target and is closely related to the electronic countermeasure operation of the target. The meteorological environment can affect the selection and use of weapons. The information combat environment involves the mode and capability of cooperative combat, as shown in Table 1.

Target attributes mainly include target radar status and motivation type. Among them, the target radar status varies according to different tasks, such as turning on the air-to-air radar when the fighter is attacking, keeping the air-to-air radar or air-to-surface radar on when the reconnaissance aircraft is performing scout tasks, etc. Motivation type involves the different types including 8-type, low jumping, high speed shaking, S-type, turning around and so on, as shown in Table 2.

Target status mainly includes velocity, height, RCS, etc. RCS reflects the stealth ability of the target and determines the mission range that the target can carry out, as shown in Table 3.

B. INTENTION SPACE DESCRIPTION

The selection of intention space is closely related to the factors such as combat form, location and scale. There are

---

**FIGURE 1.** A multi-sense-scaled attention architecture.

**TABLE 1.** A set of battlefield environment data.

| Index       | Categorical value |
|-------------|-------------------|
| Sunshine    | 0                 |
| Rain        | 1                 |
| Thunderstorm| 2                 |
| Fog         | 3                 |
| High wind   | 4                 |

**FIGURE 2.** Instance of target feature space.
usually different combat intentions in different battlefield backgrounds. The enemy will choose different operational intentions according to the importance of the target, and will also adjust the operational intentions in real time according to the battlefield conditions.

According to different target characteristics, the possible tactical intentions of the target can be divided into attack, transportation, AWACS, penetration, scout, jamming and retreat.

The numerical label space transformation of the combat intention space is shown in Table 4.

### III. CONSTRUCTION OF ARCHITECTURE

Architecture construction is the process of mapping function from target feature space to intention feature. The following is a detailed description.

#### A. DATA NORMALIZED PREPROCESSING

Actually, the original data set is collected from sensors, and there are problems such as missing and inconsistent data. Therefore, data cleaning is required, that is, the process of re-examining and verifying the data. The purpose is to delete duplicate information, correct existing errors, and provide data consistency. For missing values, we use mean estimation, probability estimation and other methods to deal with them. Duplicate records are processed by consolidation. For a few abnormal and missing values, delete them directly. After a series of data cleaning mentioned above, we can get relatively clean data, laying a good foundation for the next step of data normalization.

The characteristics of the target are diverse and multidimensional. Before intention recognition, it is necessary to eliminate the influence of different dimensions on the characteristic data of the target, as well as non-numerical digitization, so as to construct the input that can be used for the recognition model, that is, data standardization, also known as data normalization.

In particular, the purpose of this process is to convert the original data into a dimensionless pure digital expression, narrow the relative scale of different magnitude indicators, and avoid the excessive adverse impact of "large value indicators" on modeling. Practice has proved that data standardization can accelerate the convergence speed of the model and improve the generalization ability of the model.

Assuming that the target characteristics are uniformly represented by $X$, which is obtained as follows,

$$
X = \begin{pmatrix}
  x_{i1}, x_{i2}, \ldots, x_{ij}, \ldots, x_{im} \\
  x_{21}, x_{22}, \ldots, x_{2j}, \ldots, x_{2m} \\
  \ldots \\
  x_{1j}, x_{12}, \ldots, x_{ij}, \ldots, x_{1m} \\
  \ldots \\
  x_{n1}, x_{n2}, \ldots, x_{nj}, \ldots, x_{nm}
\end{pmatrix}
$$

in which, $x_{ij}$ means information of the $i$-th dimension for the $j$-th aerial feature.

1) NORMALIZATION OF NUMERICAL DATA

In order to eliminate the influence of different dimensions and accelerate the network convergence, for example, the flight speed, altitude and RCS in the target state data need to be normalized. Here, we use maximum minimum normalization:

$$
x^\prime_{ij} = \frac{x_{ij} - \min_{1 \leq i \leq n} \{x_{ij}\}}{\max_{1 \leq i \leq n} \{x_{ij}\} - \min_{1 \leq i \leq n} \{x_{ij}\}}
$$

### TABLE 2. A set of target status data.

| Index       | Description                | Categorical value |
|-------------|----------------------------|-------------------|
| Radar status| air-to-air radar status    | attack            |
|             | air-to-surface radar status| attack            |
|             | air-to-air/surface radar status| scout            |
|             | silence                    | penetration/transportation |
| Motivation type | 8-type            | AWACS/jamming    |
|             | low jump                  | attack/transportation |
|             | high speed shake          | attack            |
|             | S-type                    | attack            |
|             | turning around            | retreat           |

### TABLE 3. A set of target attribute data.

| Index   | Range      | Unit       | Categorical value |
|---------|------------|------------|-------------------|
| Velocity| 600-850    | km/h       | attack/transportation |
|         | 750-950    |            | AWACS             |
|         | 735-1470   |            | attack            |
| Height  | 50-1000    | m          | penetration/transportation |
|         | 1000-6000  |            | attack            |
|         | 15000      |            | scout             |
|         | 10000-11000|            | penetration       |
| RCS     | radar reflection cross-sectional area | dBsm   | 5 |

### TABLE 4. Labeled intention space.

| Combat intention space | Label |
|------------------------|-------|
| attack                 | 0     |
| transportation         | 1     |
| AWACS                  | 2     |
| penetration            | 3     |
| scout                  | 4     |
| jamming                | 5     |
| retreat                | 6     |
in which, \( \max \{ x_{ij} \} \), \( \min \{ x_{ij} \} \) denote the maximum, and minimum values, respectively, of the \( i \)-th row data, \( x_{ij} \) represents the new normalized value of \( x_{ij} \). Max-min algorithm ensures that the new data value obtained falls on the interval \([0,1]\).

2) NORMALIZATION OF NON-NUMERICAL DATA
Meteorological environment data, radar status, maneuver type, enemy and friend attributes, etc. belong to non-numerical data. For the numerical processing of this kind of data, labels are used for marking, and then normalization processing is carried out, i.e

\[
x_{ij}^{new} = \frac{x_{ij}^{r}}{\max_{1 \leq i \leq n} \{ x_{ij}^{r} \}}
\]  

(3)

By mapping to the interval of \([0,1]\), sunny days, rainy days, thunderstorms, heavy fog and strong winds are mapped to 0, 1, 2, 3, and 4, respectively, and the normalized values are 0, 0.25, 0.5, 0.75, and 1, respectively.

B. CNN-BILSTM-ATTENTION MODEL
According to the above characteristics of air combat intention information, we designed a multi-sense-scaled attention architecture with BiLSTM including softmax function, RELU function, sigmoid function, Adam optimization algorithm to improve the predictive performance of air combat intention recognition.

1) CONVOLUTIONAL OPERATION
Air target intelligence information is composed of multiple time series data with different characteristics. If the data of each feature dimension on each timestamp is regarded as a pixel, the air target intelligence information can be transformed into an image. The height of the picture is equal to the dimension of the feature, and the width of the picture is equal to the dimension of time.

The main part of convolution neural network is convolution. The convolution operation is actually the multiplication of the dot product of the convolution kernel matrix and a small matrix in the corresponding input layer. Convolutional neural network (CNN) can be used for classification, retrieval, detection, segmentation, face recognition, font and license plate recognition, pose estimation, etc. Further details of CNN can be seen in Zhou et al. [23]. The convolution kernel extracts features by sliding in the input layer up, down, left and right according to the stride in the way of weight sharing, so as to map the features of the input layer as the output layer, which is as shown in Fig. 3.

In this example, we set the filter as 64, the convolution kernel size as 1, the activation function as ReLU, and the padding as same to ensure that the input and output data sizes are consistent. To dropout probability, we set input unit dropout probability as 0.3.

2) BILSTM
As a special recurrent neural network, the LSTM network simulates the forgetting mechanism and memory mechanism of human brain by introducing the idea of gated switch, so as to overcome the problems of gradient disappearance and gradient explosion in the process of long sequence training, as shown in Fig. 4. Further details of LSTM can be seen in Li et al. [24] and Xue et al. [28].

a. Forgetting mechanism

The forgetting gate is used to determine which target feature information in the cell \( Cell_{t-1} \) at the previous time needs to be discarded. The output \( h_{t-1} \) at the previous time and the input \( X_{t} \) at the current time (representing a line of time-series feature data) are mapped to the interval \((0,1)\) through the sigmoid function. The value is used to determine how much information is forgotten, that is

\[
D_{f}^{t} = \sigma (W_{f} [h_{t-1}, X_{t}^{r}] + b_{f})
\]

(4)

in which, \( D_{f}^{t} \) represents the forgetting gate, \( W_{f} \) represents the weight of connections between neurons in the forgetting gate, and \( b_{f} \) represents the bias of the forgetting gate.

b. Memory mechanism

The input gate \( i_{t} \) determines which new target information is added to the memory unit \( Cell_{t}^{i} \).

The sigmoid function determines the information to be updated:

\[
i_{t} = \sigma (W_{i} [h_{t-1}, X_{t}^{r}] + b_{i})
\]

(5)

Candidate cell units \( Cell_{t}^{i} \) can be obtained through tanh layer:

\[
Cell_{t}^{i} = \tanh (W_{c} [h_{t-1}, X_{t}^{r}] + b_{c})
\]

(6)
where, $W_c$ represents the weight of connections between neurons in this layer, and $b_c$ represents the bias of this layer.

c. Update mechanism

First, a part of the nuclear information at the previous time is discarded through the forgetting gate, and then a part of the candidate cell unit is added through the input gate to obtain the cell unit at the current time, that is

$$Cell_t^i = D_t^i \cdot Cell_{t+1}^i + i_t \cdot Cell_t^i \quad (7)$$

(1) Output layer

The output at the current time is determined through the output gate. The judgment conditions are obtained through sigmoid, i.e.

$$O_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o) \quad (8)$$

where, $O_t$ is the output of this layer, $W_o$ is the weight of connections between neurons in this layer, and $b_o$ is the offset of this layer.

From the above process description, it can be seen that the traditional LSTM network is an unidirectional neural network structure which information obtained is the historical information before the current time. In fact, in operational practice, not only the current time information should be analyzed according to historical data, but also the future information should be mutually verified. The Bi-directional Long Short-Term Memory (BiLSTM) network is composed of a forward LSTM network and a backward LSTM network, which can capture the characteristics of the front and rear information. The model is shown in Fig. 5. Further details of BiLSTM can be seen in Teng et al. [27] and Zhou et al. [29].

The hidden layer state $O_t$ of bilstm at time $t$ can be obtained by two parts: the forward hidden layer state $\vec{h}_t$ and the backward hidden layer state $\tilde{h}_t$. $\vec{h}_t$ is determined by the current input $X_t$ and the forward hidden layer state $h_{t-1}$ at time $t-1$. The backward hidden layer state $\tilde{h}_t$ is determined by the current input $X_t$ and the backward hidden layer state $\tilde{h}_{t+1}$ at time $t+1$.

The calculation formula is shown in equations (9) to (11), where $w_i$ ($i = 1, 2, \ldots, 6$) represents the weight from one cell layer to another.

$$\vec{h}_t = f(w_1 X_t + w_2 \vec{h}_{t-1}) \quad (9)$$

$$\tilde{h}_t = f(w_3 X_t + w_5 \tilde{h}_{t+1}) \quad (10)$$

$$O_t = g(w_4 \vec{h}_t + w_6 \tilde{h}_t) \quad (11)$$

3) ATTENTION MECHANISM

In order to make the network pay more attention to the key information in the target features and improve the recognition accuracy, the attention mechanism is introduced behind the BiLSTM layer. The attention mechanism simulates the human brain’s different attention to different objects in the same field of vision by giving a certain degree of correlation weight to the input sequence features. Further details of Attention mechanism can be seen in Qiao et al. [26], Teng et al. [27] and Zhou et al. [29].

In this paper, we first learn the importance of each feature, and then assign corresponding weights to each feature according to the importance. For example, when the enemy aircraft is intended to attack, the course angle, maneuver type and other characteristics will be assigned more weight by the attention mechanism to deepen the model memory. The basic structure of the attention mechanism model is shown in Fig. 6.

The calculation formula is as follows:

$$S = \sum_{i=1}^{t} \gamma_i \times h_i \quad (12)$$

$$\gamma_i = \frac{\exp(\delta_i)}{\sum_{a=1}^{t} \exp(\delta_a)} \quad (13)$$

$$\delta_i = v^T \gamma \tanh(W h_i + b_i) \quad (14)$$

in which, $S$ means that the vector $s$ represents the sum of the product of the weight of each hidden state in the new hidden state and the input hidden state; $\delta_i$ refers to the amount of information contained in the implicit state at the current time; $v^T$ and $W_i$ are weight vector.

Initializing the network and constantly updating the parameters in the formula make the attention state change accordingly. After sequence preprocessing, the target data can meet the input interface requirements of CNN bilstm attention network. In order to find the optimal network structure for intention recognition, it is necessary to select the number of network layers, the number of neurons in each layer and the network super parameters of specific CNN module and bilstm module. In addition, it is also necessary to prevent over fitting of the network model, so as to reduce the accuracy of the network to identify the target intention.

4) TRAINING AND TEST

After building the model, you need to input the training data into the network and train the network. For this model, one of
the cores of network construction is the selection of network layers. However, the selection of layers is closely related to which data set to process. Therefore, pre experiments can be carried out in advance to obtain better network layers, number of neurons and network super parameters for different target data.

C. ALGORITHM FLOW
There is a certain correlation between the target intention and the target state information. The aircraft can perform different combat tasks in different states. Therefore, the air combat state presented by the target to the pilot is different under different combat intentions. The status data of the target is obtained through the onboard sensor, and the data is deeply analyzed to mine the correlation between the tactical intention and the target air combat status. Finally, the network model with the ability of tactical intention recognition is obtained through the network learning and training.

The steps of intention identification are as follows:
Step 1: obtain the time sequence state data of the target at multiple times from multiple dimensions, and select multiple main target features from the feature space as the input for identifying the intention features;
Step 2: clean the original data and then normalize data set, including the numerical processing of non-numerical target features, and standardize the numerical data to a unified dimension;
Step 3: input the multi-dimensional initialized target intention features into the trained model for code recognition;
Step 4: feature attention is obtained through attention mechanism and BiLSTM network learning, target features are weighted and intention prediction is carried out.
Step 5: decode and identify the input features through a softmax layer and classify the intentions to obtain the most likely tactical intentions of the target at this time.

IV. SIMULATION AND ANALYSIS
A. DATA
The experimental data comes from simulation data, and its data composition includes azimuth, distance, velocity, direction, altitude, RCS, etc. as feature input.

The original experimental data were randomly divided into two parts. The training set and verification set accounted for 80% of the total data set, which were used to train and adjust the model parameters. The test set accounts for 20% of the total data set to verify the performance of the algorithm.

B. SOFTWARE AND HARDWARE
Our experiment is entirely conducted in Python 3.8.0, relying on the TensorFlow 2.0.0. The deep learning networks are trained on NVIDIA GPUs, all other models are trained on a CPU cluster, with specifics shown in Table 5.

C. PARAMETER SETTINGS
The relevant parameters are optimized to obtain better recognition performance of aerial targets combat intention. Parameters include hyper parameters play an important role in the classification performance of neural networks. In our model, it mainly includes epoch, activation functions, time steps, LSTM units, batch size, dropout, adam optimization and so on.

Combined with the relevant hyper parameters setting according to experience above, hyper parameters are set as follows in Table 6.

The error rate of the optimal training set and verification set obtained by each group of hyper parameters is shown in Table 7.

In our experiment, there are many hyper parameters that need to be set and adjusted according to the accuracy. Through multiple parameter adjustments, it is found that the number of hidden layer nodes is not the more the better. When a certain number of layers is reached, the effect of increasing the number of hidden layers will become worse, and even over fitting will occur. Other parameters are set in the above charts.

According to the section III, subsection B, we know batch size is like a sliding window. If there are 20000 rows of data and all data are trained after 100 times of training, then batch size = 20000/100 = 200. If too many rows of data are fed at a time, the model will easily suffer from indigestion like people.

LSTM unit is the size of hidden layer in an LSTM cell, and is also the dimension of the output space. Time step means that each input data is related to the first number of successively input data, which will not change after the model is built. While batch size is the training parameter.
During model training, it can be adjusted at any time according to the results of model training and loss to achieve the best.

After many experiments, the best result is batch size = 32, LSTM units = 32 and time steps = 10 or time steps = 15, which is shown in Fig. 7.

### D. EVALUATION METRICS

To evaluate the performance of our proposed model on targets intention recognition, Precision Rate (PR), Recall Rate (RR) and F1-Score are applied to estimate the classification performance.

To better understand the above evaluation metrics, confusion matrix is introduced which is shown in Table 8.

Each combat intention can be considered as a two-class problem. That is to say, the combat intention identification problem can be divided into two categories, either one class or other classes. If we correctly classify invasion, it’s called a True Positive (TP), otherwise it’s called a False Negative (FN) when we classify invasion as others. Similarly, if we misclassify other targets as invasion, it’s named False Positive (FP), and the rest of the case is called False Negative (FN).

Precision Rate is the probability that the samples predicted to be positive samples in the prediction results are correctly predicted to be positive samples.

\[
PR = \frac{TP}{TP + FP}
\]  

Recall Rate is the probability of being correctly predicted as a positive sample in the positive sample of the original sample.

\[
RR = \frac{TP}{TP + FN}
\]  

Ideally, it is best to have high precision and recall. However, in general, if the precision is high, the recall is low, if the recall is high, the accuracy is low. Therefore, F1-Score is the harmonic value of precision and recall.

\[
F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}
\]  

The diagonal line indicates the number of correct samples identified, as shown in Fig. 8. Confusion matrix is a standard format for expressing accuracy evaluation. Each column of the confusion matrix represents a prediction category, and the total number of each column represents the number of data predicted as the category. Meanwhile, each row represents the true belonging category of the data, and the total number of data in each row represents the number of data instances of the category. The data on the diagonal line indicates that the predicted result is correct, otherwise the predicted result is wrong. Therefore, the larger the value on the diagonal, the higher the accuracy of the prediction results.

It can be seen that our model has the high recognition accuracy for all seven intentions, especially retreat intention can reach 100%. Because the air combat characteristics corresponding to AWACS intention and scout intention are highly similar and deceptive, there are a few mutual recognition errors between each other, which are also in line with the actual situation.
Table 9 shows that the three evaluation indicators further verify the superiority of our model according to the table 8. Among them, the precision rate refers to the proportion of the number of correct positive samples to the number of positive samples determined by the model, the recall rate refers to the proportion of the number of correct positive samples identified by the model to the number of actual positive samples, and F1-score refers to the harmonic average value of the precision rate and the recall rate.

### E. MODEL EVALUATION

Since there is no public data set, it is difficult to directly compare with the existing literature on the same data set. Therefore, we take the typical methods of the existing literature as the benchmark model, and make a comparative analysis with our model on the same data set we collected.
achieved good training effect in both the training set and the test set, that is, fast convergence and high recognition accuracy.

By training the input features and historical data, our model obtains the attention of different features and time points, that is, the influence of different features and time points on intention recognition. Due to the randomness of the training results, the number of training iterations for the attention mechanism layer is set to 50, and finally the feature weights are obtained by averaging, as shown in Fig. 11. The attention weights of the target’s speed, altitude and radar state features are large, which means that they are more critical to target intention recognition, so they need to be given more “attention”. In general, the weight distribution obtained by the attention mechanism basically conforms to the analysis results of expert experience. This also shows that the importance of different features can be more accurately analyzed through weighted features, so as to improve the recognition accuracy of target intention.

V. CONCLUSION

Aiming at the problem of air combat target intention recognition, this paper proposed an air target tactical intention recognition model based on multi-scaled attention architecture. The following conclusions can be drawn through theoretical analysis and simulation experiments:

1) After preprocessing the data collected by multi-sensor through data cleaning and normalization, the multi-dimensional data including battlefield environment, target attribute and target state are used to construct the target feature space and intention space. As convolutional neural networks is actually an input-output mapping, it can learn a large number of mapping relations between input and output without any precise mathematical expression between input and output. As long as the convolution network is trained with a known pattern, the network has the mapping ability between input and output pairs, thus reducing the dependence on expert knowledge and experience. The BiLSTM model can not only train the features of historical time and future time information, but also extract deeper features. And different target state characteristics have different influence on the intention analysis of the target. Therefore, the introduction of attention mechanism can further analyze the importance of different target combat characteristics to achieve the purpose of more accurate identification of intention.

2) Our model can effectively distinguish the influence degree of different characteristic indexes on air combat target intention recognition, with fast convergence speed, high recognition accuracy and good robustness.

3) Our model is more suitable for multi-dimensional and special air combat data, and also has better recognition performance.

In conclusion, our model has great advantage for recognition of aerial targets combat intention. It is of great significance to improve the recognition ability of aerial targets combat intention, and has theoretical significance and reference value for command decision. Furthermore, our model in this paper is still insufficient for the intention recognition of target clusters with high similarity and strong deception, which will be the focus of the next research.

REFERENCES

[1] T. Sun, Sun Tzu’s Art of War Strategy and Attack. Beijing, China: Zhonghua Book Company, 2011.
[2] D. F. Noble, “Schema-based knowledge elicitation for planning and situation assessment aids,” IEEE Trans. Syst., Man, Cybern., vol. 19, no. 3, pp. 473–482, May 1989.
[3] X. B. Guo, Z. Wang, and W. D. Hu, “Information fusion with Bayesian networks for target recognition,” J. Syst. Simul., vol. 17, no. 11, pp. 2713–2716, Nov. 2005.
[4] A. W. Yang, L. Zhan-Wu, X. An, X. Zhi-Fei, and C. Yi-Zhe, “Threat level assessment of the air combat target based on weighted cloudy dynamic Bayesian networks,” Flight Dyn., vol. 38, no. 4, pp. 87–94, Apr. 2020.
[5] F. J. Zhao, Z. J. Zhou, C. H. Hu, L. Wang, and T. Liu, “Aerial target intention recognition approach based on belief-rule-base and evidential reasoning,” Electron. Opt. Control, vol. 24, no. 8, pp. 15–19, Aug. 2017.
[6] X. Yin, M. Zhang, and M. Q. Chen, “Combination intention recognition of the target in the air based on discriminant analysis,” J. Projectile, Rocket, Missiles Guid., vol. 38, no. 3, pp. 46–50, Jun. 2018.
[7] M. Ben-Bassat and A. Freedy, “Knowledge requirements and management in expert decision support systems for (military) situation assessment,” IEEE Trans. Syst., Man, Cybern., vol. SMC-12, no. 4, pp. 479–490, Jul. 1982. doi: 10.1109/TSMC.1982.4308575.
[8] W. Zhou, J. Zhang, N. Gu, and G. Yan, “Recognition of combat intention with insufficient expert knowledge,” in Proc. 3rd Int. Conf. Comput. Modeling, Simulation Appl. Math., Wuhan, China, Sep. 2018, pp. 328–333.
[9] X. Niu, H. Zhao, and Y. Zhang, “Naval vessel intention recognition based on decision tree,” Ordinance Ind. Autom., vol. 29, no. 6, pp. 473–482, May/Jun. 2010.
[10] X. Xia, “The study of target intention assessment method based on D-S evidence theory,” M.S. thesis, Dept. Inf. Syst. Eng., Natl. Univ. Defense Technol., Changsha, China, 2006.
[11] Y. L. Sun and L. Bao, “Study on recognition technique of targets’ tactical intentions in sea battle field based on D-S evidence theory,” Ship Electron. Eng., vol. 32, no. 5, pp. 48–51, May 2012.
[12] L. Ma, F. Tian, and H. Y. Shan, “Optimization method of surface ship to air combat deployment based on time window,” J. Phys. Conf. vol. 1574, Jun. 2020, Art. no. 012151.
[13] W. Wang, H. Hong, Y. Zhang, S. Li, and J. Shi, “Realization of real-time detection algorithms for key parts of unmanned aerial vehicle based on support vector machine,” in Proc. Autom. Target Recognit. Navigat., 2020, pp. 116–122.
[14] G. L. Dai, W. Chen, and Z. J. Liu, “Method of target tactical intention recognition based on interval grey relational degree,” Mathematics Pract., vol. 34, no. 20, pp. 198–207, Oct. 2014.
[15] Q. K. Yao, S. J. Liu, X. Y. He, and W. Ou, “Research and prospect of battlefield target operational intention recognition,” J. Command Control, vol. 3, no. 2, pp. 127–131, Feb. 2017.
[16] B. Li, W. Pei, J. Li, X. Zhang, and Y. Liu, “ISAR target recognition system design based on artificial intelligence,” in Proc. Pattern Recognit. Comput. Vis., Feb. 2020, pp. 149–155.
[17] Y. Gong, Z. Ma, M. Wang, X. Deng, and W. Jiang, “A new multi-sensor fusion target recognition method based on complementarity analysis and neutrosophic set,” Symmetry, vol. 12, no. 9, p. 1435, Aug. 2020.
[18] A. Li, B. Zheng, and L. Li, “Intelligent transportation application and analysis for multi-sensor information fusion of Internet of Things,” IEEE Sensors J., vol. 21, no. 22, pp. 25035–25042, Nov. 2021.
[19] J. Pei, W. Hao, C. Wang, Y. Huang, Y. Zhang, J. Wu, and J. Yang, “Multi-view deep feature learning network for SAR automatic target recognition,” Remote Sens., vol. 13, no. 8, p. 1455, Apr. 2021.
[20] H. Chen, Q. L. Ren, and Y. Hua, “Fuzzy neural network based tactical intention recognition for sea targets,” Syst. Eng., Electron., vol. 38, no. 8, pp. 1847–1853, Aug. 2016.
[21] X. Kong, B. Cai, Y. Liu, H. Zhu, Y. Liu, H. Shao, C. Yang, H. Li, and T. Mo, “Optimal sensor placement methodology of hydraulic control system for fault diagnosis,” Mech. Syst. Signal Process., vol. 174, Jul. 2022, Art. no. 109069.
[22] B. Cai, C. Sheng, C. Gao, Y. Liu, M. Shi, Z. Liu, Q. Feng, and G. Liu, “Artificial intelligence enhanced reliability assessment methodology with small samples,” IEEE Trans. Neural Netw. Learn. Syst., early access, Nov. 25, 2021, doi: 10.1109/TNNLS.2021.3128514.

[23] W. W. Zhou, P. Y. Yao, J. Y. Zhang, X. Wang, and S. Wei, “Combat intention recognition for aerial targets based on deep neural network,” Acta Aero-Nauticaet Astronautica Sinica, vol. 39, no. 11, Aug. 2018, Art. no. 322468.

[24] Z. W. Li. (2022). Air Combat Intention Recognition Method of Target Based on LSTM Improved by Attention Mechanism. [Online]. Available: https://kns.cnki.net/kcms/detail/41.1227.TN.20220321.1355.002.html

[25] T. Zhou, M. Chen, Y. Wang, J. He, and C. Yang, “Information entropy-based intention prediction of aerial targets under uncertain and incomplete information,” Entropy, vol. 22, no. 3, p. 279, Feb. 2020.

[26] D. F. Qiao. (2021). Recognition and Prediction of Group Target Intention in Multi-Domain Operations. [Online]. Available: https://kns.cnki.net/kcms/detail/11.2422.TN.20211214.2033.028.html

[27] F. Teng, “BiLSTM-attention: An air target tactical intention recognition model,” Aero Weaponry, vol. 28, no. 5, pp. 24–32, Oct. 2021.

[28] J. J. Xue and J. Y. Zhu. “Panoramic convolutional long short-term memory networks for combat intention recognition of aerial targets,” IEEE Access, vol. 8, pp. 183312–183323, 2020.

[29] P. Zhou, W. Shi, J. Tian, Z. Qi, B. Li, H. Hao, and B. Xu, “Attention-based bidirectional long short-term memory networks for relation classification,” in Proc. 54th Annu. Meeting Assoc. Comput. Linguistics, Berlin, Germany, vol. 1, Aug. 2016, pp. 207–212.

[30] T. Zhang, X. L. Liu, and Y. P. Gao, “Sentiment analysis of attention mechanism based on convolutional neural network and bidirectional gated recurrent unit network.” Sci. Technol. Eng., vol. 21, no. 1, pp. 269–274, Jan. 2021.

XUHUA WANG was born in Longyou, Zhejiang, China, in 1984. He received the Ph.D. degree in electronic science and technology from Air Force Engineering University. He is currently a Lecturer with Xidian University. His research interests include application of UAV swarm operation and anti-UAV technology.

ZHIKAI LIN was born in Zhoukou, Henan, China, in 1993. He received the B.S. degree in measurement control technology and instruments from the Huazhong University of Science and Technology, Wuhan, China, in 2015, and the M.S. degree in control science and engineering from the National University of Defense Technology, Changsha, China, in 2017. His research interests include air traffic control and aviation cluster collaboration.

YAHUI HU was born in Houma, Shanxi, China, in 1980. She received the Ph.D. degree in computer science from Wuhan University, China, in 2016. She is currently an Associate Professor at the Air Force Early Warning Academy. Her research interests include air combat planning, aviation mission planning, big data, and artificial intelligence.

JIANWEI LIU was born in Xintai, Shandong, China, in 1985. He received the M.S. degree in signal and information processing from the Air Force Radar Academy, China, in 2010. He is currently a Lecturer with the Air Force Early Warning Academy. His research interests include array signal processing, radar equipment support, and secondary surveillance radar.

**VOLUME 10, 2022**