Journal of Systems Engineering and Electronics
Vol. 35, No. 2, April 2024, pp.406 – 416

CBA: multi source fusion model for fast and intelligent target intention identification

WAN Shichang[1,2], LI Qingshan[1,*], WANG Xuhua[1], and LU Nanhua[1]

1. School of Computer Science and Technology, Xidian University, Xi'an 710071, China;
2. School of Mathematics and Computer Application, Shangluo University, Shangluo 726000, China

Abstract: How to mine valuable information from massive multi-source heterogeneous data and identify the intention of aerial targets is a major research focus at present. Aiming at the long-term dependence of air target intention recognition, this paper deeply explores the potential attribute features from the spatiotemporal sequence data of the target. First, we build an intelligent dynamic intention recognition framework, including a series of specific processes such as data source, data preprocessing, target space-time, convolutional neural networks-bidirectional gated recurrent unit-atteneion (CBA) model and intention recognition. Then, we analyze and reason the designed CBA model in detail. Finally, through comparison and analysis with other recognition model experiments, our proposed method can effectively improve the accuracy of air target intention recognition, and is of significance to the commanders' operational command and situation prediction.

Keywords: intention, massive data, deep network, artificial intelligence.

DOI: 10.23919/JSEE.2024.000023

1. Introduction

As the modern battlefield environment and tasks become more and more complicated, the opportunity for fighting flit now is fleeting. It is an important prerequisite for defeating the enemy to be able to quickly mine valuable information from massive data, make reasonable inferences, and obtain accurate enemy combat intentions. Therefore, the reasoning and recognition of combat intention are current research foci of battlefield situation analysis.

At present, the commonly used methods mainly include expert system, template matching, evidence theory based methods, Bayesian network, and neural network. Expert system is the process of building a priori knowledge base by domain experts, designing the formulation of corresponding rules between battlefield situation and air combat intention, and then recognizing the target intention through the inference mechanism [1,2]. Template matching mainly refers to building a prior knowledge base of battlefield situation based on the field expert experience, and matching the extracted target tactical features to obtain the target tactical intention [2]. Bayesian network is to use network nodes to describe the states of the target, determine the corresponding node parameters through a prior knowledge base, and then recognize the combat intention to achieve uncertain goals [3–8].

These above-mentioned methods generally have problems including excessive reliance on prior knowledge, lower recognition accuracy, lower training efficiency and poor generalization ability, which are not suitable for complex and changeable battlefield environment. At the same time, many studies have focused on the past information and failed to effectively use the future information for intention recognition.

With the in-depth development and extensive application of artifical intelligence (AI) technology, researchers have found that the deep learning network technology can simulate the cognitive model and reasoning process of commanders [9–12], which provides a good theoretical basis for target intention recognition. For example, Xue et al. constructed a panoramic convolution long and short term memory (PCLSTM) neural network model to reduce the neural network parameters from the time series, thus improving the intention recognition accuracy [13]. Li et al. regarded intention recognition as a multi-classification problem, and designed a deep learning network based on the attention mechanism to improve recognition accuracy [14]. Qiao et al. took the identification and prediction of aircraft carrier formation intention in multi-domain warfare as an example, and gave a cluster target

Manuscript received March 30, 2023.
*Corresponding author.
This work was supported by the National Natural Science Foundation of China (61502523).

intention reasoning scheme [15]. Teng et al. introduced a model of air target tactical intent recognition based on bidirectional long short-term memory (BiLSTM)-attention, which effectively improved the efficiency of air target tactical intent recognition and had important theoretical significance and reference value for auxiliary combat systems [16]. To solve the problem that the strong hypothesis of the classical hidden Markov model (HMM) is not applicable in the real battlefield, Dai et al. defined a framework of intention recognition based on the maneuver intention in the tactical intention and solved the probability calculation, parameter learning and sequence labeling problems of the model [17]. These models have good results in identifying target combat intentions and are consistent with the temporal characteristics and logical relationships in battlefield situation information. However, the accuracy and efficiency of recognition still need to be improved.

To sum up, it is a process of obtaining information through various technical means and recognizing the combat intention based on relevant domain knowledge in a real-time dynamic and antagonistic battlefield environment. Therefore, it is more scientific to infer the operational intention of the identified target from the characteristic information of multiple consecutive moments.

2. Motivation

With the development of modern information war, the combat mode is increasingly diversified and multi-domain. Air target tactical intention is often hidden in a spatiotemporal sequence data. We mainly focus on the effectiveness and accuracy of target intention recognition, and use the deep neural network as a technical means to mine the potential feature association from the continuous time and space information, reasoning the internal mechanism between training dataset and test dataset to identify their real intention. In summary, the main contributions of this paper are:

(i) Propose a framework of target intention recognition. In order to describe the battlefield data more truly and objectively, we describe the battlefield information in detail from multiple space-time dimensions, preprocess the data to obtain standardized data information, construct the target space-time space and design a deep learning network to achieve the reasoning of the target's tactical intention.

(ii) Design a convolutional neural network-bidirectional gated recurrent unit-attention (CBA) model. Among it, convolution neural network (CNN) is to extract the data features in the target space-time space. Bidirection gated recurrent unit (BiGRU) is mainly used to discover the change trend from the time and space data, and learn the depth information in the feature vector. In order to select the features closely related to intention recognition, we introduce the attention mechanism to assign different weights to different features, thus reducing the data dimension and improving the accuracy and efficiency of target recognition.

(iii) Evaluate our model. Compared with other classical models, the effectiveness and feasibility of our method are verified.

3. Framework

Target tactical intention recognition is to extract spatiotemporal information (including battle environment information, target static attribute, and target dynamic attribute) from real-time and confrontation environment. In order to accurately, timely and continuously obtain the battlefield data and provide scientific and powerful information support for command and decision-making, this paper proposes a framework, which is from data collection, target spatiotemporal space construction, to intelligent intention recognition and reasoning, as shown in Fig. 1.

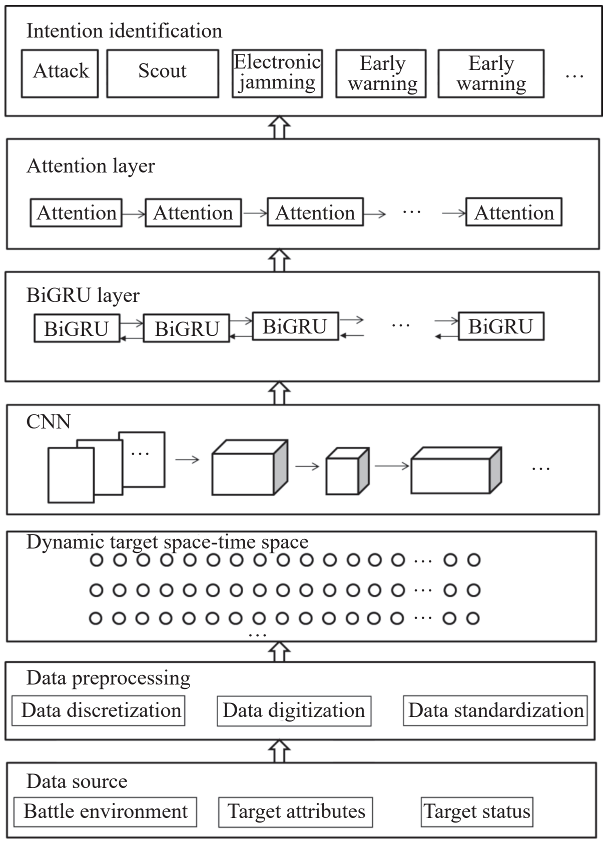


Fig. 1 Framework

Data source: describe the battlefield spatial-temporal information from multiple dimensions such as battle environment, target static attribute, and target dynamic attribute.

Data preprocessing: normalize the original data. For example, convert non-numerical data such as radar status and maneuver type into numerical data. Standardize the data into the same dimension, such as the speed, altitude, and radar cross section (RCS). The smaller RCS means that the target has better ability of stealth and performing remote tasks.

Construct target space-time space: each column of data represents battlefield environment information, target static and dynamic attributes. Each row of data means the data value of each battlefield information at different times.

Design a CBA model: in order to fully grasp the history and future time state target information, we extract the spatiotemporal and spatial feature data with the CNN technology, introduce the bidirectional mechanism and attention mechanism on the basis of gated recurrent unit (GRU), and assign weights to each depth information, so as to improve the capability of capturing more distinctive features.

Target intention identification: the target intention is not only related to the combat scenario, but also closely related to the factors such as the combat location, scale and mode. At the same time, the opponent can quickly change their intention after analyzing the current situation. After many times of training and testing, we predict and evaluate the target intention, including attack, detection, and jamming.

3.1 Data source

When the target performs a task, some of its characteristics must meet certain requirements to ensure the completion of the given task. In other words, intention is specifically reflected in the operational action and target status. For example, the target's flight altitude, azimuth, velocity, flight action and battlefield meteorological conditions can provide a basis for identifying the target's intention.

We classify the above information. Combined with [14,18–22], the information is summarized as battle environment, target static and dynamic attributes, as shown in Fig. 2.

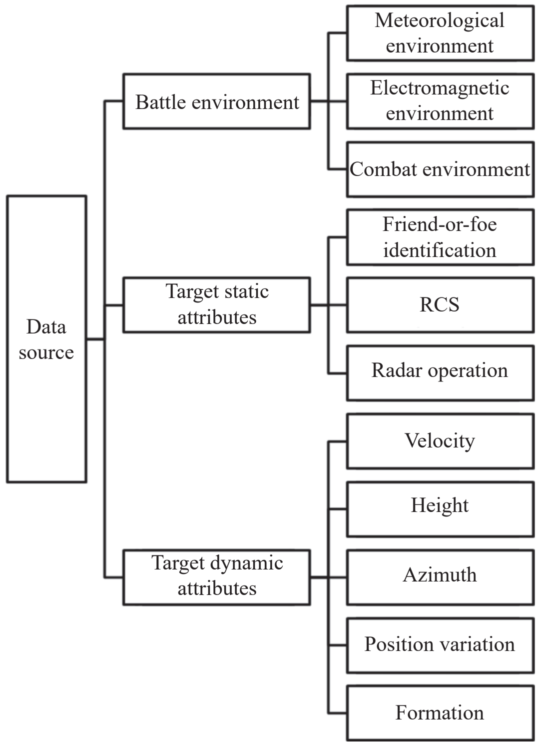


Fig. 2 Data source

Here, the battle environment information refers to the temperature, humidity, wind speed, cloud, fog and other meteorological environments that affect the flight of the target, as Table 1 shows. Information operation environment means the cooperative operation mode and capability involving multiple formations according to the operation needs. The electromagnetic environment involves electromagnetic interference, electromagnetic waves released by weapons and equipment, electromagnetic phenomena generated in nature, etc. These factors are closely related to the electronic warfare operations carried out by the target.

Table 1 Meteorological environment

Environment	Label
Sun	1
Rain	2
Thunderstorm	3
Fog	4
Gale	5

Target static attributes mainly include radar status and maneuvering type. Radar can provide information for targeted bombing, guided air-to-ground missiles, navigation, etc. The status includes whether air-to-air and air-to-ground radars are turned on. The maneuver type is used to indicate that the target achieves its tactical intention through certain tactical maneuvers, including 8-shaped, Sshaped, low jump, high speed shake, and turn around, as Table 2 shows.

Table 2 Target static attributes

Category	Description	Label
Radar status	Air-to-air radar power-on status	−2
	Ground radar power-on status	−1
	Air-to-surface radar power-on status	0
	Silence	1
Motivation type	8-type	0
	Low jump	1
	High speed shake	2
	S-type	3
	Turn around	4
Friend-or-foe attribute	Hostile plane	0
	Friendly aircraft	1
	Unidentified aircraft	2
Target size	Big target	0
	Middle target	1
	Small target	2

Target dynamic attributes mainly include flight speed, flight altitude, RCS, position change, formation, and azimuth, as Table 3 shows.

Table 3 Target dynamic status

Category	Range
Velocity/(km/h)	600−850
	750−950
	735−1 470
Height/m	50−1 000
	1 000−6 000
	15 000
	10 000−11 000
RCS/dBsm	Radar reflection cross-sectional area
Position change	−0.5
	−0.4
	−0.3
	−0.2
	−0.1
	0.1
	0.2
	0.3
	0.4
	0.5
Formation	Cuneiform (0)
	Trapezoid (1)
	Herringbone (2)
	Diamond (3)
	Longitudinal shape (4)
Azimuth/mil	0−6 400

3.2 Data preprocessing

Data preprocessing is the preparation before data analysis, and is also an indispensable part of data analysis. It mainly uses a series of methods to process "dirty" data, accurately extract data, and adjust the format of data, so as to obtain a group of high-quality data that conforms to accurate, complete, concise and other standards, and ensure that the data can better serve data analysis.

For data from different scenarios, it is necessary to obtain unified and standardized data through discretization, numerization, normalization and other processing to improve the convergence speed and generalization ability of the mode [23].

Assume that a_{ij} represents the ith attribute value of target a at the jth moment, there are the following preprocessing operations:

(i) Unified dimension, namely normalization. The speed, height, RCS and other values of the target need to be normalized. Here, the normalization method is as follows:

$$a'_{ij} = \frac{a_{ij} - \min\limits_{1\leqslant i\leqslant n}\{a_{ij}\}}{\max\limits_{1\leqslant i\leqslant n}\{a_{ij}\} - \min\limits_{1\leqslant i\leqslant n}\{a_{ij}\}} \tag{1}$$

where $\max\limits_{1\leqslant i\leqslant n}\{a_{ij}\}$ is the maximum value of attribute i at the jth moment. Otherwise, $\min\limits_{1\leqslant i\leqslant n}\{a_{ij}\}$ means the minimum value. a'_{ij} represents the new value after normalization. Through the max-min normalization calculation method, the new data are normalized to the interval [0,1].

(ii) Numerization of non-numeric data. In this paper, the non-numerical data include radar status, enemy and friend attributes, meteorological environment, and maneuver type. We use digital tags to mark them, and then normalize them.

a''_{ij} means the new value after normalization.

$$a''_{ij} = \frac{a_{ij}}{\max\limits_{1\leqslant i\leqslant n}\{a_{ij}\}} \tag{2}$$

3.3 Target spatial-temporal space

Suppose that target spatial-temporal space is expressed in $\boldsymbol{C}$, the column data represents the value of different attributes of target a at each time, and the row data represents the value of each attribute of target a at different time, that is

$$\boldsymbol{C} = \begin{bmatrix} c_{11}, c_{12}, \cdots, c_{1j}, \cdots, c_{1m} \\ c_{21}, c_{22}, \cdots, c_{2j}, \cdots, c_{2m} \\ \vdots \\ c_{i1}, c_{i2}, \cdots, c_{ij}, \cdots, c_{im} \\ \vdots \\ c_{n1}, c_{n2}, \cdots, c_{nj}, \cdots, c_{nm} \end{bmatrix} \tag{3}$$

where c_{ij} indicates the normalized value of the ith attribute of target a at time j.

3.4 CNN

After the above processing, the input interface requirements that meet the model network can be obtained. It can be seen from the target spatial-temporal space that the target information is composed of multiple attributes according to the time series [24,25]. If $\boldsymbol{C}$ is regarded as an image, c_{ij} is a pixel of this image, n is the height of the image and m is its width.

With $\boldsymbol{C}$ as the input, convolution operation is a dot product multiplication calculation process, which is multiplied by the convolution kernel matrix and the small matrix in the corresponding input layer. After several convolution layers and pooling layers, the important features of the target can be extracted by sliding up, down, left and right according to the step.

3.5 Deep network

3.5.1 GRU

GRU is a deep neural network developed from LSTM [26,27] and has optimized the gate design with integrating the forgetting gate and input gate in LSTM into an update gate. In other words, the original unit structure composed of three doors is optimized to the unit structure composed of two doors. At the same time, it also has the "memory" function for the time series data, which can make intelligent prediction and effectively alleviate the problem of gradient disappearance and gradient explosion.

According to the previous analysis, the state and behavior of the target is in time seires. Recent trend changes have an impact on the trend of future trends. These characteristics are applicable to the GRU model, as shown in Fig. 3.

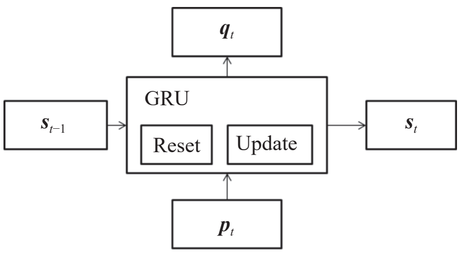


Fig. 3 GRU

$\boldsymbol{p}_t$ is the input of time t. $\boldsymbol{s}_{t-1}$ means the hidden layer state of the time $t-1$ including the relevant information of the preceding node. $\boldsymbol{s}_t$ means the next hidden layer state. And $\boldsymbol{q}_t$ represents the output at the time t.

Two gate control states are from the input $\boldsymbol{s}_t$ at the time t and the previous state $\boldsymbol{s}_{t-1}$, namely, reset gate f_t and update gate g_t.

Reset gate:

$$f_t = \alpha(\boldsymbol{p}_t \boldsymbol{W}_1 + \boldsymbol{s}_{t-1} \boldsymbol{W}_2 + b_f). \tag{4}$$

Update gate:

$$g_t = \alpha(\boldsymbol{p}_t \boldsymbol{W}_3 + \boldsymbol{s}_{t-1} \boldsymbol{W}_4 + b_g) \tag{5}$$

where α represents the sigmoid function. Through this function, all data can be converted to 0 to 1, which is used as a gating signal. $\boldsymbol{W}_1$, $\boldsymbol{W}_2$, $\boldsymbol{W}_3$, $\boldsymbol{W}_4$, b_f, b_g are the parameters.

Reset the data after the gate is reset:

$$\boldsymbol{s}'_{t-1} = \boldsymbol{s}_{t-1} \odot w_t \tag{6}$$

where w_t represents the parameter at the time t. $\odot$ is the multiplication by element.

Candidate hidden layer status

$$\tilde{\boldsymbol{s}}_t = \tanh(p_{ij} \boldsymbol{W}_f + w_t \odot \boldsymbol{s}_{t-1} \boldsymbol{W}_g + b_s) \tag{7}$$

where $\boldsymbol{W}_f$ and $\boldsymbol{W}_g$ are the weight matrix.

When the parameter w_t approaches 0, the model will lose the past information and only retain the current information. When w_t approaches 1, the model will consider that the past information has a certain role, so it will be included in the current information. $\boldsymbol{s}_{t-1}$ includes the past information and $\tilde{\boldsymbol{s}}_t$ is the candidate hidden state. The final hidden state is

$$\boldsymbol{s}_t = (1-\gamma_t) \odot \boldsymbol{s}_{t-1} + \gamma_t \odot \tilde{\boldsymbol{s}}_t \tag{8}$$

where the part before the plus sign represents the selective "forgetting" of the original hidden state, and the part after the plus sign represents the selective "memory" of the candidate hidden state of the current node. The range of γ_t is 0−1.

When γ_t is closer to 0, it means that the past data is "forgotten", that is, some unimportant information in the hidden information is forgotten. When γ_t is closer to 1, it means that the past data is "remembered".

Therefore, the reset gate is used to combine the new input information with the previous information, and the update gate is used to indicate which information can be retained in the current information. In addition, there are many parameters and super parameters in the whole process that need to be trained and tested several times before they can be determined. These contents are analyzed one by one in the experimental analysis.

3.5.2 BiGRU

GRU is a unidirectional transmission network. It only uses the target information in the past for judgment, but does not fully use the future information. BiGRU can compensate for the disadvantage of GRU, as shown in Fig. 4.

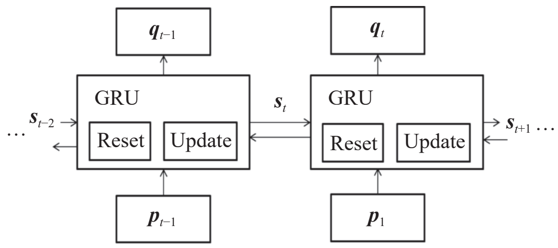


Fig. 4 BiGRU

The hidden layer state s_t at time t is devided into two segments: the forward hidden layer state $\overrightarrow{s}_t$ and the backward hidden layer state $\overleftarrow{s}_t$. $\overrightarrow{s}_t$ is determined by p_i^t and the forward hidden layer state $\overrightarrow{s}_{t-1}$ at time $t-1$. The backward hidden layer state $\overleftarrow{s}_t$ is determined by p_i^t and the backward hidden layer state $\overleftarrow{s}_{t+1}$ at the time $t+1$.

3.6 Attention mechanism

In order to focus on the key feature information of the target in the model, we introduce the attention mechanism behind the BiGRU layer [28,29].

In our daily life, we quickly obtain a small part of useful information through eyes, ears, mouth, hands, etc., and ignore other relatively unimportant information. This is actually attention. For example, when a scene enters the human vision, it often pays attention to the bright colors and dynamic points in the scene, while the remaining scenes are often temporarily ignored. Therefore, the attention mechanism can give weights to the input features, simulate the human brain's ability to focus on one or more features, and realize the efficient allocation of information processing resources.

Therefore, the attention mechanism can achieve different attention to different objects by giving different weights to features, thus realizing rapid processing of complex information.

In this paper, the corresponding attribute feature weights of different target intentions are also different. For example, when the target intention is surveillance, its maneuver types and other characteristics will be assigned more weights to deepen the model memory, as shown in Fig. 5.

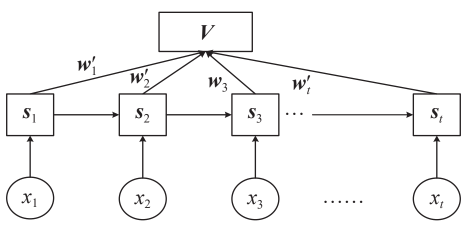


Fig. 5 Attention

$$V = \sum_{i=1}^{t} w_i' \times s_i \tag{9}$$

$$w_i' = \frac{\exp(\lambda_i)}{\sum_{i=1}^{t} \exp(\lambda_i)} \tag{10}$$

$$\lambda_i = \eta_\gamma^{\mathrm{T}} \tanh(W_i s_i + b_i) \tag{11}$$

The vector V is the sum of the weight of the hidden state and the product of the input hidden state. λ_i is the information contained in the implicit state at the current moment. w_i' and W_i are the weight vectors.

3.7 Target intention recognition

The target intention is different in different forms of operations. Define the intention of the target according to the corresponding operational form, environment, target status and other information.

According to the target attribute characteristics, there are six types of target intentions: scout, attack, electronic jamming, transport, early warning detection and retreat.

For example, Yang et al. [21] defined the intention of underwater targets to attack, retreat, patrol, etc., Wang et al. [18] and [22] believed that the target intention includes attack, early warning detection, electronic jamming, scout, and retreat, as shown in Fig. 6.

Intention	Label
Attack	0
Scout	1
Electronic jamming	2
Transport	3
Early warning detection	4
Retreat	5

Fig. 6 Intention category

For example, if the output of our model is 0, it means that the target's intention is an attack during the observed period of time. By coding the combat intent, the abstract intent can be embodied and more easily represented by models.

4. Experiments

4.1 Data set

Our data is analog data. All data are divided into training set and test set according to the percentage of 8:2. The training dataset is to train our model's parameters. The test dataset is to evaluate the accuracy and error rate of intention recognition.

4.2 Experimental environment

The main configurations of hardware and software involved in this experiment are shown in Table 4.

Table 4 Configuration

Name	Version/configuration
Operating system	Windows10
Computer memory	8GB
CPU	i7-9700 CPU
Programming language	Python 3.8.0
Compilation environment	Anaconda 3
Deep learning network	TensorFlow 2.0.0，keras2.8.0
Library	Os,numpy,pandas, etc.

4.3 Evaluation criteria

4.3.1 Metrics

In order to analyze the effectiveness and feasibility of our method in detail, loss function and accuracy are used to measure [30–32].

(i) Loss function

The loss function is to evaluate the distance between the predicted value and the real value of our model. This is a common method for deep learning and training models. It updates the parameters by calculating the loss function, and then reduces the optimization error until the loss function value falls to the target value or reaches the training number.

In the case of no fitting, the smaller the loss, the better. The smaller the loss value, the closer the predicted value is to the true value. Measurement methods include mean square error (MSE), root mean square error (RMSE) and mean absolute error (MAE).

MSE is to calculate the square sum of the difference between the predicted value and the real target value:

$$\mathrm{MSE}=\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2. \tag{12}$$

The smaller the MSE value, the higher the accuracy of our model.

$$\mathrm{RMSE}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2} \tag{13}$$

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}|y_i-\hat{y}_i| \tag{14}$$

(ii) Accuracy

Accuracy is the propotion of the number of samples correctly classified by the model in the total number of samples for the given test dataset. Generally, the higher the accuracy, the better the model, as shown in Table 5.

Table 5 Confusion matrix

Confusion matrix		Result of intention recognition: 1	Result of intention recognition: 0	Total
Real label	1	TP	FN	Actual positive (TP+FN)
	0	FP	TN	Actual negative(FP+TN)
Total		Positive recognition (TP+FP)	Negative recognition (FN+TN)	TP+FN+ FP+TN

$$\mathrm{Accuracy}=\frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{TN}+\mathrm{FP}+\mathrm{FN}} \tag{15}$$

TP: true positive. The sample’s real label is positive, and the experiment’s result is also positive.

TN: true negative. The sample’s true label is negative, and the experiment’s result is also negative.

FP: false positive. The sample’s true label is negative, and the experiment’s result is positive (i.e., false positive).

FN: false negative. The sample’s true label is positive, and the experiment’s result is negative (i.e., missing).

4.3.2 Parameter training

The model involves a lot of hyper-parameter settings, such as convolution kernel, number of iterations, training batch, and selection of optimization algorithm. The setting of these parameters requires multiple training, iterations and multiple optimization algorithm adjustments, as well as measurement and comparison analysis of loss rate and accuracy rate, before the model parameters applicable to this experiment can be determined.

The number of iterations means that in order to make the predicted value as close as possible to the real value, the model needs to be trained and tested several times before it can be trained.

Batch means that according to the CPU and memory configuration of the computer hardware and the current CPU usage, the computer can train the model with better performance, read the data source into the memory in batches and execute the corresponding program commands.

A cost function will be defined before model training. The purpose of our model training is to minimize the cost function, which is an optimization problem. There are many commonly used optimization algorithms, such as gradient descent, and adaptive optimizer.

Regularization is to introduce additional information in order to prevent the model from over-fitting, and to intervene and modify the model externally to improve the model's generalization capability. The commonly used regularization methods include L2 regularization and Dropout regularization.

The basic principle of Dropout regularization is to impose a discard probability on neurons in some layers of the neural network. Each model training can only be completed by a part of neurons. The neurons discarded in each iteration are randomly selected. This mechanism makes the model be randomly simplified, and on the other hand makes different neurons be randomly valued. Its effect is helpful to reduce the over-fitting and make the model more robust and general.

Parameter values such as filter and kernel can be determined after multiple training and iteration, which is also the process of neural network learning data characteristics, as shown in Table 6, Table 7 and Table 8. When filter number is 128, kernel size is 1 and dropout is 0.3, the result shows that our model has the highest accuracy and lowest loss values.

Table 6 Comparison of different filter numbers

Filter number	Loss	MAE
8	0.1523	0.3684
16	0.1038	0.3219
32	0.0667	0.2580
64	0.0532	0.2316
128	0.0268	0.1632

Table 7 Comparison of different kernel sizes

Kernel size	MAE	R^2
1	0.1573	1.246 9e+06
2	0.1605	1.299 4e+06
3	0.1988	1.990 3e+06
4	0.1725	1.539 7e+06

Table 8 Comparison of different dropouts

Dropout	Loss	R^2
0.1	0.0525	2.625 9e+06
0.2	0.0305	1.523 1e+06
0.3	0.0249	1.220 4e+06

Table 9 is the important hyper parametric value obtained after the model has been trained and tested for many times, and after the analysis and comparison of loss rate and accuracy rate. Among them, the number of iterations is set to 50 and the batch size is adjusted to 32, the optimization method is the Adam algorithm, and the error rate is measured by variance.

Table 9 Parameter setting

Parameter	Value
Filter	128
CNN activation	Relu
CNN dropout	0.3
Dense activation	Sigmoid
Time steps	15
Loss	MSE
Optimizer	Adam
Metrics	MAE
Epoch	50
Batch size	32
Dropout	0.3

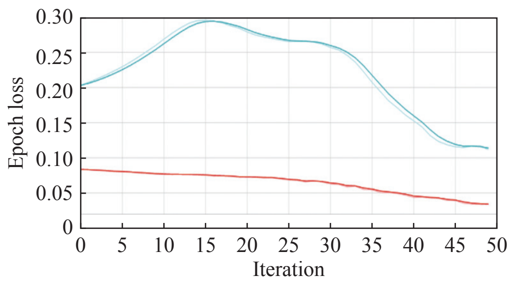
After many training tests, important parameters in the CBA model are determined, as shown in Table 10.

Table 10 Construction of network

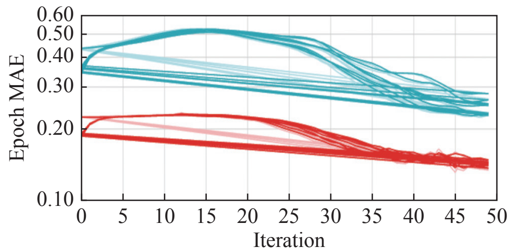
Layer	Value
Convolution	Filers=64, kernel-size=2, activation=relu
Dropout	0.3
BiGRU	Unit=32
Dropout	0.3
Attention	Attention_3d_block2, flatten
Dense	Activation=sigmoid

4.3.3 Loss function analysis

As the number of iterations increases, the error rate of training dataset and verification dataset gradually decreases and tends to be stable. This method is suitable for intention identification analysis. Red line is training, and blue line is test, as shown in Fig. 7.



(a) Epoch loss



(b) Epoch MAE

Fig. 7 Loss function analysis example

4.4 Model evaluation

In order to further illustrate our model performance, it is compared with three identification networks based on GRU, LSTM and LSTM-attention, as shown in Fig. 8, Table 11 and Table 12. These algorithms are used for intention recognition after the algorithm model is trained. The accuracy of our method is better than other algorithms in target intention recognition. And the convergence speed is the fastest, which is also the biggest advantage of our model. The LSTM model is similar to the GRU model, but their internal structure of the cycle is different. Therefore, the intention recognition results of the two models are the closest.

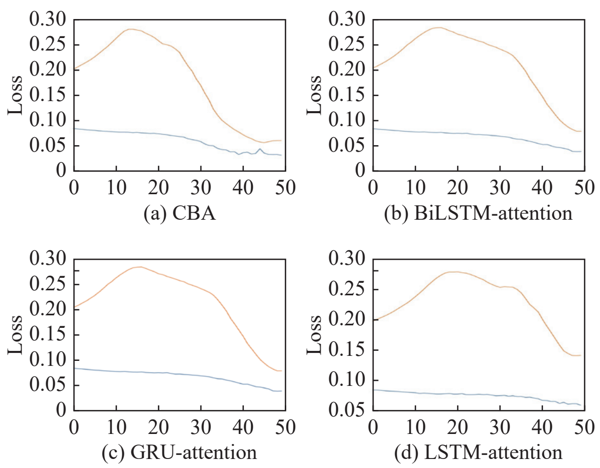


Fig. 8 Training loss vs test loss

Table 11 Comparison of algorithm performance

Name	Loss	MAE	R^2	Epoch/s	Step/s
GRU	0.1757	0.4173	−8.786 6e+06	39	39
LSTM	0.1708	0.4048	−8.52 4e+06	31	31
CBA	0.0303	0.1738	−1.514 4e+06	56	56
BiLSTM	0.0591	0.2429	−2.956 4e+06	66	66

Table 12 Comparison of loss

Parameter	LSTM	CBA	BiLSTM	GRU
Max	0.280 811 012	0.273 370 922	0.280 811 012	0.276 434 325
Min	0.055 925 872	0.052 450 593	0.055 925 872	0.073 835 136
Average	0.180 188 963	0.160 373 669	0.180 188 963	0.175 134 731

Table 13 shows three evaluation indicators measuring our model. Our model has a high accuracy of intention recognition. In air combat, the similarity of air combat characteristics corresponding to early warning detection and scout is especially high, thus there is a small number of recognition errors, which is also in line with the actual situation.

Table 13 Performance measurement %

Intention	Accuracy	Recall	$F1$-score
Attack	93.8	92.5	93.2
Transport	95.3	86.3	90.4
Electronic jamming	90.8	97.4	94.0
Early warning detection	89.2	97.5	93.2
Scout	96.3	97.5	96.7
Retreat	94.7	100	97.3

5. Conclusions

In this paper, a multi-source fusion CBA model is proposed to solve the problem that it is difficult to obtain the long-term dependence of air target intention recognition. We use neural network to mine internal feature association from continuous multi-time target state sequence, and build a framework and model of intelligent dynamic intention recognition. The comparison test with other models is completed on the data set built by the simulation system. The experiment shows that the proposed model has advantages in recognition accuracy. In the future, CBA model will be optimized in more scenarios to further improve the accuracy of target intention recognition in air combat.

In summary, we focus on the urgent military demand for battlefield situational awareness in the information age, and combine deep learning in the field of artificial intelligence to comprehensively analyze the target tactical intent. Although the application of deep learning in the military field has only just begun, once the potential of deep learning is unleashed, it will inevitably lead to significant changes in the military field. Therefore, we can continue to conduct relevant research in scenarios such as group targets and multi domain operations, laying a solid foundation for winning future intelligent warfare.

References

[1] BEN-BASSAT M, FREEDY A. Knowledge requirements and management in expert decision support systems for (military) situation assessment. IEEE Trans. on Systems, Man, and Cybernetics, 1982, 12(4): 479–490.

[2] ZHOU W W, ZHANG J Y, GU N N, et al. Recognition of combat intention with insufficient expert knowledge. Proc. of the 3rd International Conference on Computational Modeling, Simulation and Applied Mathematics, 2018: 316−321.

[3] GAO X G, YANG Y. Threat assessment for warship air defense based on Bayesian network. Tactical Missile Technology, 2020(4): 47–57. (in Chinese)

[4] HOSSEINI S, IVANOV D. Bayesian networks for supply chain risk, resilience and ripple effect analysis: a literature review. Expert Systems with Applications, 2020, 161: 113649.

[5] YANG A W, LI Z W, XU A, et al. Threat level assessment of the air combat target based on weighted cloudy dynamic Bayesian networks. Flight Dynamics, 2020, 38(4): 87–94.

[6] GE S. Research on tactical intention recognition based on rule discovery and Bayesian reasoning. Harbin: Harbin Engineering University, 2015. (in Chinese)

[7] GE S, XIA X Z. Study of optimal decision model based on BN. Applied Mechanics and Materials, 2014, 687: 1552–1556.

[8] YANG J, GAO W Y, LIU J. Threat assessment method based on Bayesian network. Journal of PLA University of Science and Technology, 2010, 11(1): 43–48.

[9] YAO Q K, LIU S J, HE X Y, et al. Research and prospect of battlefield target operational intention recognition. Journal of Command and Control, 2017, 3(2): 127–131.

[10] LI A, ZHENG B Y, LI L. Intelligent transportation application and analysis for multi-sensor information fusion of Internet of Things. IEEE Sensors Journal, 2020, 21: 25035–25042.

[11] GONG Y M, MA Z Y, WANG M J, et al. A new multi-sensor fusion target recognition method based on complementarity analysis and neutrosophic set. Symmetry, 2020, 12(9): 1435−1435.

[12] LI B N, PEI W, LI J Q, et al. ISAR target recognition system design based on artificial intelligence. Proc. of the 11th International Symposium on Multispectral Image Processing and Pattern Recognition, 2020: 149−155.

[13] XUE J J, ZHU J, XIAO J Y, et al. Panoramic convolutional long short-term memory networks for combat intension recognition of aerial targets. IEEE Access, 2020, 8: 183312–183323.

[14] LI Z W, LI S Q, PENG M Y, et al. Air combat intention recognition method of target based on LSTM improved by attention mechanism. Electronics Optics & Control, 2023, 30(3): 1–9.

[15] QIAO D F, LIANG Y, MA C X, et al. Recognition and prediction of group target intention in multi-domain operations. Systems Engineering and Electronics, 2022, 44(11): 3403–3412. (in Chinese)

[16] TENG F, LIU S, SONG Y F. BiLSTM-Attention: an air target tactical intention recognition model. Aero Weaponry, 2021, 28(5): 24–32.

[17] DAI F H. Research on intention recognition technology in battlefield based on improved Markov model. Changsha: National University of Defense Technology, 2019. (in Chinese)

[18] WANG X H, LIN Z K, HU Y H, et al. Learning embedding features based on multisense-scaled attention architecture to improve the predictive performance of air combat intention recognition. IEEE Access, 2022, 10: 104923–104933.

[19] FAN H. Air target intention estimation for airborne situation awareness system. Ordnance Industry Automation, 2022, 41(4): 14–18. (in Chinese)

[20] ZHANG J T, ZHOU W N, YAN X F, et al. Intelligent prediction for behavioral intent of marine targets based on LSTM and fuzzy reasoning. Journal of China Academy of Electronics and Information Technology, 2022, 17(9): 897–904. (in Chinese)

[21] YANG Y T, YANG J, LI J G. Research on air target tactical intention recognition based on EMEBN. Fire Control & Command Control, 2022, 47(5): 163–170. (in Chinese)

[22] HU Z Y, LIU H L, GONG S J, et al. Target intention recognition based on random forest. Modern Electronics Technique, 2022, 45(19): 1–8. (in Chinese)

[23] CAI Q Y. Research on power behavior analysis method based on load data preprocessing. Nanjing: Nanjing University of Posts and Telecommunications, 2022. (in Chinese)

[24] LI Z L, XU K, XIE J F, et al. Deep multiple instance convolutional neural networks for learning robust scene representations. IEEE Trans. on Geoscience and Remote Sensing, 2020, 58(5): 3685–3702.

[25] LIU C, TANG X, MA J J, et al. Remote sensing images feature learning based on multibranch networks. Proc. of the IEEE International Geoscience and Remote Sensing Symposium, 2020: 2057−2060.

[26] WU J X. Introduction to convolutional neural networks. National Key Lab for Novel Software Technology, 2017, 5(23): 495–495.

[27] ZHANG T, LIU X L, GAO Y P, et al. Sentiment analysis of attention mechanism based on convolutional neural network and bidirectional gated recurrent unit network. Science Technology and Engineering, 2021, 21(1): 269–274.

[28] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need. Proc. of the 31st Conference on Neural Information Processing System, 2017: 6000−6010.

[29] BELLO I, ZOPH B, VASWANI A, et al. Attention augmented convolutional networks. Proc. of the IEEE/CVF International Conference on Computer Vision, 2019: 3286−3295.

[30] JIANG Q, HUANG R M, HUANG Y C, et al. Application of BP neural network based on genetic algorithm optimization in evaluation of power grid investment risk. IEEE Access, 2019, 7: 154827–154835.

[31] ZHONG Y W, WU X L. Effects of cost-benefit analysis under back propagation neural network on financial benefit evaluation of investment projects. PloS ONE, 2020, 15(3): e0229739.

[32] XUE J K, SHEN B. A novel swarm intelligence optimization approach: sparrow search algorithm. Systems Science & Control Engineering, 2020, 8(1): 22–34.

Biographies

WAN Shichang was born in 1984. He received his M.S. degree from Shaanxi Normal University, Xi'an, China, in 2010. He is currently pursuing his Ph.D. degree in the School of Computer Science and Technology, Xidian University. His research interests include UAV swarm combat application, multi-source data fusion, and intelligent data analysis.

E-mail: wanshichang@126.com

LI Qingshan was born in 1973. He received his B.S., M.S. and Ph.D. degrees from Xidian University, Xi'an, China, in 1995, 1999 and 2003, respectively. He is a professor in Xidian University. His research interests include agent-oriented software engineering, intelligent data analysis, and decision support system.
E-mail: qshli@mail.xidian.edu.cn

WANG Xuhua was born in 1984. He received his B.S., M.S. and Ph.D. degrees from Air Force Engineering University, Xi'an, China, in 2007, 2009 and 2013, respectively. He is a lecturer in Xidian University. His research interests include radar network modeling, UAV swarm combat application, and anti-UAV technology.
E-mail: daleiwxh@163.com

LU Nanhua was born in 1999. He received his B.S. degree from Wuhan University of Technology, Wuhan, China, in 2021. He is studying for his M.S. degree at Xidian University. His research interests include multi-UAV mission planning, radar network modeling, and 3D scene simulation.
E-mail: lunanhua@qq.com