IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received August 3, 2020, accepted September 16, 2020, date of publication September 22, 2020, date of current version October 19, 2020.

*Digital Object Identifier 10.1109/ACCESS.2020.3025926*

# Panoramic Convolutional Long Short-Term Memory Networks for Combat Intension Recognition of Aerial Targets

**JUNJIE XUE, JIE ZHU, JIYANG XIAO, SHENG TONG, AND LING HUANG**
Air Traffic Control and Navigation College, Air Force Engineering University, Xi'an 710051, China

Corresponding author: Junjie Xue (xuejunjiekjgcdx@163.com)

This work was supported in part by the Domain Foundation of China under Grant 61400010304.

**ABSTRACT** Aiming at the limitation that traditional methods for combat intention recognition of aerial targets are difficult to effectively capture the essential characteristics of intelligence information, we design a novel deep learning method, Panoramic Convolutional Long Short-Term Memory networks (PCLSTM), to improve the recognition ability. First, based on the characteristics of aerial target intelligence information, a panoramic convolutional layer is designed to extract the loosely coupled characteristics of intelligence information, and a time series pooling layer is designed to reduce the scale of neural network parameters on a large scale. Then, the temporal feature extraction capability of the LSTM layer and the depth feature mining capability of the traditional deep learning layer are combined to construct the PCLSTM neural network. Subsequently, the recognition performance of PCLSTM is analyzed by simulation experiments compared with standard deep net, convolutional neural network and LSTM network as benchmark models. Finally, PCLSTM was used to carry out simulation tests on different truncated data sets of original intelligence information, to analyze the optimal length of truncated data for different combat intention recognition. And then a reasonable aerial target combat intention recognition method is designed. The simulation results show that the method presented in this paper has theoretical significance and reference value for command decision-making.

**INDEX TERMS** Aerial targets, combat intension recognition, deep learning, panoramic convolutional long short-term memory neural network.

## I. INTRODUCTION

Recognition for combat intention of enemy aerial targets is an important prerequisite to modern air combat command decision, which directly determines the outcome of air combat and even the whole campaign [1]. With the continuous development of military science and technology, the complexity and dynamic characteristics of modern battlefield increase dramatically. Faced with the vast and complicated battlefield information, it becomes very difficult for commanders to quickly and accurately recognize enemy's combat intension, especially to the aerial targets of large-scale systematic confrontation [2]. Therefore, the construction of computer combat assistance support system to improve the accuracy and speed of air target combat intention recognition, has gradually become one of the research hotspots in modern air combat.

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

Combat intention recognition tasks on aerial targets time series are notoriously difficult, primarily driven by the high degree of noise and the operational art of the commander [3]. Yet, combat intention of an aerial target is always closely related to its equipment performance and the combat mission to be carried out. That is to say, it is possible to extract the typical characteristics of aerial targets from the chaotic intelligence information, and then recognize their combat intentions. The existing research methods of targets intention recognition mainly include template matching method [4], expert system method [5]–[7], decision tree [8], [9], Bayesian network [10]–[12] and neural network [13]–[15]. Floyd *et al.* [4] applies the template matching method to combat intention recognition in beyond-visual-range air combat. Ben-Bassat and Freedy [5] evaluates military situation by integrating knowledge requirements and management into expert decision support systems. In response to the rapid iteration of expert knowledge in military issues, Carling [6] designed a real-time expert system to assess naval combat situations.

Zhou *et al.* [7] modified the expert system to accommodate combat intention recognition problems with insufficient expert knowledge. In decision tree, Niu *et al.* [8] has applied it to the study of intention recognition of naval vessel. By integrating information entropy into the decision tree, Zhou *et al.* [9] improved the effectiveness of combat intention recognition of aerial targets. Considering the advantage of Bayesian network which is easy to calculate, Jin *et al.* [10] tries to use it to solve the problem of intension recognition of aerial targets. Chen and Wu [11] designed a layered intention model to represent the uncertain elements relating to adversarial intention and their uncertain relations in naval battlefield domain. In order to verify the effectiveness of Bayesian algorithm, Xu *et al.* [12] used ACMI measured data to assess situation of the beyond-visual-range air combat. However, with the increasing complexity of battlefield environment and the emergence of new combat platforms and combat styles, it is more difficult for field experts to grasp the mapping relationship between target information characteristics and combat intention in a short time. Therefore, the above methods are not able to capture complex non-linear dependencies between intelligence information and combat intension.

Aiming at the above limitations, researchers introduce neural network into the field of target intention recognition. Ahmed and Mohammed [13] designed fuzzy min-max neural network to automatically extract features and rules from training data, and memorized recognition rules in network weights for later intention prediction. By introducing Rectified Linear Unit (ReLU) activation function and adaptive torque estimation (Adam) optimization algorithm, Zhou *et al.* [14] designed a combat intention recognition model based on deep neural network, which can improve the convergence speed of the model and effectively prevent it from falling into local optimum. In fact, the mission implementation of air targets is determined by a set of dynamic changes, one part of which is contained in the static information, and the other part is contained in the change of timing information. Obviously, the maneuverability of an airplane is a typical example of how difficult it is to describe static information at a certain moment. Unfortunately, none of the above neural networks has effectively considered the timing characteristics of the target intelligence information. Although there is no literature on identifying combat intention of air targets through timing sequence information at this stage, the deep learning method has made good achievements in detecting and identifying high-speed trains. For example, Yao *et al.* [15] presents intelligent inspection model for exterior substance inspection.

In order to effectively recognize the combat intention of aerial targets, we design the following strategies based on the characteristics of aerial targets intelligence information. 1) According to the smooth transition characteristics of the time dimension, time series pooling is carried out to reduce the parameter scale of the neural network in a large scale. 2) According to the loose coupling of feature dimensions, panoramic convolution operation of all feature dimensions is carried out to abstract the essential features of information. 3) According to the characteristics of the temporal progression of intelligence information, Long Short-Term Memory net is used to obtain useful information between the temporal sequences. On the basis of the above strategy, combined with Adam optimization algorithm and Sigmoid activation function, Panoramic Convolutional Long Short-Term Memory Neural Network is construct to recognize combat intention of aerial targets.

The remainder of this paper is organized as follows. Section 2 briefly covers the data sample, software packages, and hardware. Section 3 introduce deep learning models and compare our model architectures, including the generation of training and trading sets, the construction of input sequences, the model architecture and training as well as the classifying steps. Section 4 presents the results and discusses our most relevant findings in light of the existing literature. Finally, Section 5 concludes.

## II. DATA, SOFTWARE, HARDWARE

### A. CHARACTERISTIC VARIABLES

For the empirical application, we construct our aerial targets combat intension database according to the actual characteristics of the air combat data from radar station. The information data of aerial target includes two parts, numerical characteristic data and non-numerical characteristic data. The numerical characteristic data includes azimuth angle, distance, heading angle, velocity and height of the aerial targets relative to our side. The non-numerical characteristic data includes air-to-air radar status, air-to-surface radar status, electronic-jamming status, level of Radar Cross-Section (RCS) and so on. To air-to-air radar status, 1 indicates that the target is using radar to search for airborne targets or to guide air-to-air missiles and 0 means that this status is not occurred. Air-to-surface radar status is similar to air-to-air radar status, with 1 indicating that the target is searching the ground and the sea surface, and 0 indicating that this status has not occurred. 1 in the electronic-jamming status means that the target is using related equipment to carry out the electronic jamming task. To level of RCS, 1 represents stealth aircraft whose RCS always below 1, 2 represents medium aircraft whose RCS is always between 1 and 10, and 3 represents large aircraft whose RCS is above 10. A set of aerial target intelligence data is proposed in this paper, with specifics shown in Table 1 and Table 2.

Let $I^s = \left(I^s_{i,j}\right)_{i \in N, j \in T}$ be defined as the time series information of aerial target s, and $I^s_{i,j}$ means information of the $j$-th dimension for the $s$-th aerial target at time $i$. Set $N$ is determined by the dimension of aerial targets intelligence information, denote as $N = \{1, 2, \ldots j, \ldots n\}$. There are 9 types of aerial targets intelligence information in this paper, that is to say, n is 9 here. The time set $T$ is determined by the time length of the acquired aerial target intelligence information, or the shortest combat intention recognition time after finding the target, we denote it as $T = \{1, 2, \ldots i, \ldots m\}$.

**TABLE 1.** A set of numerical air target intelligence data.

| Index | Description | unit | Numerical value |
|---|---|---|---|
| Azimuth angle | the angle between the vertical plane containing it and the plane of the meridian | *mil* (Russia) | 2230.0 |
| Distance | the distance between enemy planes and ours | *km* | 310.0 |
| Heading angle | the angle between the earth's North Pole and the projection of the plane's speed axis onto ground | ° | 12.0 |
| Velocity | the velocity of aerial target | *m/s* | 220.0 |
| Height | the height of aerial target | *km* | 15.8 |

**TABLE 2.** A set of non-numerical air target intelligence data.

| Index | Description | Categorical value |
|---|---|---|
| air-to-air radar state | whether the air-to-air radar is switched on | 1 |
| air-to-surface radar state | whether the air-to-surface radar is switched on | 0 |
| electronic-jamming status | whether the electronic-jamming device is turned on | 1 |
| level of RCS | the size of the radar reflection cross-sectional area | 2 |

Due to the difference of measurement units, the distribution range of the original characteristics of aerial targets intelligence information is often very different. Features with a wide range of values play a leading role, so the core features cannot be accurately obtained from the raw data. Therefore, data preprocessing is needed to reduce the impact of data dimensional differences and improve the convergence efficiency of the network. In this paper, the maximum-minimum standardized model is adopted to normalize 9 types of attribute data such as azimuth, distance, heading angle, velocity, height, air-to-air radar state, air-to-surface radar state, disturbing state and level of RCS.

By mapping to the interval of [0,1], information about row data in the $i$-th dimension, $X_i^r = \left[x_{i,1}^r, x_{i,2}^r, \ldots x_{i,j}^r, \ldots x_{i,m}^r\right]$, ($i = 1, 2, \ldots, 9$, m is the dimension of time set $T$), is normalized to $X_i = \left[x_{i,1}, x_{i,2}, \ldots x_{i,j}, \ldots x_{i,m}\right]$, which is obtained as follows,

$$x_{i,j} = \frac{x_{i,j}^r - \min_j x_{i,j}^r}{\max_j x_{i,j}^r - \min_j x_{i,j}^r} \tag{1}$$

in which $\max_j x_{i,j}^r$, $\min_j x_{i,j}^r$ respectively denotes the maximum, minimum value of the $i$th row data $X_i^r$.

### B. SOFTWARE AND HARDWARE

Data preparation and handling is entirely conducted in Python 3.7.0, relying on the packages numpy 1.18.1 and pandas. Our deep learning networks are developed with Google TensorFlow 2.0.0, a powerful library for large-scale machine learning on heterogenous systems. The deep learning networks are trained on NVIDIA GPUs, all other models are trained on a CPU cluster.

## III. MODEL

This section consists of three steps. First, we split our raw data in study periods, composed of training sets (for in-sample training) and trading sets (for out-of-sample classifications). Second, we design a novel deep learning model, named Panoramic Convolutional Long Short-Term Memory Neural Network (PCLSTM), for combat intension recognition of aerial targets. Third, we briefly describe the deep net, Convolutional Neural Network and LSTM model we apply.

### A. TRAINING AND TESTING SETS

We divide the raw data set into two complementary parts randomly. The first part is training and verification set, accounting for 80% of the raw data set. Training and verification set is used for training and adjusting relevant parameters of neural network. The remaining part is testing set, which is used to verify the algorithm performance.

On the basis of the above separation, the training and verification set is further divided into two complementary subsets for each other, the training set and the verification set. Training subset, which accounting for 80% of training and verification set, is used to train neural network parameters. The rest of the data used to adjust the neural network hyperparameters is validation subsets.

### B. RESPONSE VARIABLES

Depending on the difference of combat response strategy, combat intention of aerial target can be preliminarily denoted as 8 categories, such as penetration, attack, jamming, transportation, refueling, Airborne Warning and control system (AWACS), Scout and civil flight. Therefore, we can solve the problem of combat intention recognition of aerial targets by defining an eight-category classification problem. i.e., the response variable $Y_{s,t+1}$ for each time series $s$ and time $t$ can take on eight different probability values, and the combat intension corresponds to the maximum probability means the aerial target.

### C. MODEL OF PCLSTM NEURAL NETWORK

According to the characteristics of aerial target intelligence information, we designed the Panoramic Convolutional Layer to extract the loosely coupled characteristics of intelligence information, Time Series Pooling Layer to reduce the parameter scale of neural network. Combined with LSTM layer, batch normalization layer, soft max layer, Adam optimization algorithm and sigmoid activation function, Panoramic Convolutional Long Short-Term Memory Neural Network is constructed to recognize the combat intention of aerial targets.

#### 1) PANORAMIC CONVOLUTIONAL LAYER

In the deep learning problem of static image, the relationship between pixels in the image is closely related to their distance. The closer the pixels are, the higher the correlation is. Therefore, convolution kernel can be used in typical static image recognition problems to explore the feature information in the small receptive field. However, when it comes

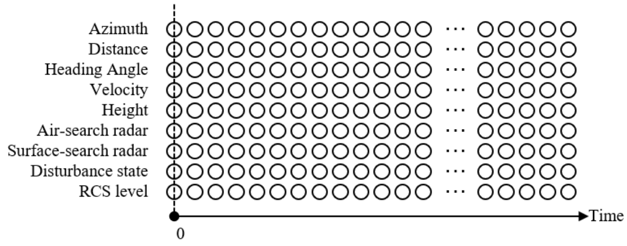



**FIGURE 1. Schematic of aerial targets intelligence information.**

to the recognition of aerial targets combat intention, there is no fixed position relationship between the nine types of typical information. That is to say, although there is a certain coupling between each type of aerial targets intelligence information, it presents obvious characteristics of loose coupling. At the same time, the air target intelligence information is a data whose length is much larger than the width, the traditional convolutional neural network layer is difficult to capture data characteristics effectively.

In order to obtain the essential characteristics of aerial targets intelligence information effectively, we designed a panoramic convolutional layer. In a small range of time, the panoramic convolution kernel acquires the essential characteristics of the temporal receptive field by covering all dimensional information. For a set of information with time length $t$ and information dimension $n$, it can be denoted as $X = (x_{i,j})_{i\in\{1,2,\ldots,n\},j\in\{1,2,\ldots,t\}}$ and expressed as Fig. 1.

According to the dimension of the aerial target information, the convolution kernel which can cover all the information dimensions is designed, and then slide through the time dimension to obtain the feature information. Assuming that the input dimension of the first layer (or the output of the previous layer) is $n$ times $t$, then the panoramic convolution kernel can be designed as $W^{pc} = (w_{i,j})_{i\in\{1,2,\ldots,n\},j\in\{1,2,\ldots,k\}}$, in which $n$ is the characteristic dimension of input information, and $k$ is the width of convolution kernel. In general, width of convolution kernel $k$ is much less than the time length $t$, so that the convolution kernel may slide on the input information matrix.

Correspondingly, the shared bias of panoramic convolution layer is denoted as $b_{pc}$. Using the panoramic convolution kernel to slide in the time dimension, the output of panoramic convolution can be obtained. Panoramic convolution is calculated in the same way as a typical convolutional neural network, that is, the corresponding elements are multiplied, and then summed. For input information $X$ at time $l$, its panoramic convolution output is:

$$o_l = b_{pc} + \sum_{i=1}^{n}\sum_{j=1}^{k} w_{i,j} \cdot x_{i,j+l} \tag{2}$$

By designing multiple panoramic convolution kernels, the essential features of input information can be extracted in many aspects, and the output feature under multiple mappings can be obtained. The schematic diagram of panoramic convolutional layer is shown in Fig. 2.

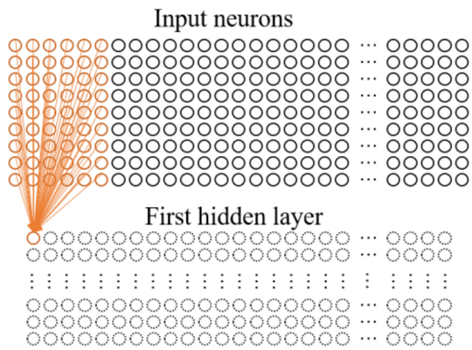



**FIGURE 2. Schematic of panoramic convolution operation.**

It can be seen that the training parameters of the panoramic convolutional layer are jointly determined by the characteristic dimension of input information $n_i$, the characteristic dimension of output information $n_o$ and the width of the panoramic convolutional kernel $n_h$. The number of training parameters of a panoramic convolutional layer can be calculated as follows,

$$n_{pc} = n_i \times n_h \times n_o + 1 \tag{3}$$

### 2) TIME SERIES POOLING LAYER

For a typical convolutional neural network, the pooling operation is to simplify the information in the fixed pooling field by obtaining the maximum or average value. Because the aerial targets information does not have obvious position relation in the characteristic dimension, and shows obvious smooth change characteristic in the time dimension, the traditional pooling operation is not suitable for the aerial targets combat intention recognition. Therefore, we design a time series pooling strategy for discrete data characteristics. On the one hand, time series pooling retains information of every feature by performing on each feature dimension. Considering the smooth change of time series information, the partially overlapping view sliding is adopted to simplify information on time dimension. The value of time series pooling is taken as the average of the information on the pooling field of vision. The schematic diagram of average time series pooling is shown in Fig. 3.

### 3) LSTM LAYER

LSTMs have been introduced by Hochreiter and Schmidhuber in [16], and then were further refined in the following years in [17], [18]. It belongs to the class of recurrent neural networks (RNNs), i.e., neural networks whose "underlying topology of inter-neuronal connections contains at least one cycle". LSTM networks are specifically designed to learn long-term dependencies and are capable of overcoming the previously inherent problems of RNNs, i.e., vanishing and exploding gradients in [19].

LSTM networks are composed of an input layer, one or more hidden layers, and an output layer. The number of neurons in the input layer is equal to the number of explanatory variables (feature space). The number of neurons in the output layer reflects the output space, i.e., eight neurons in

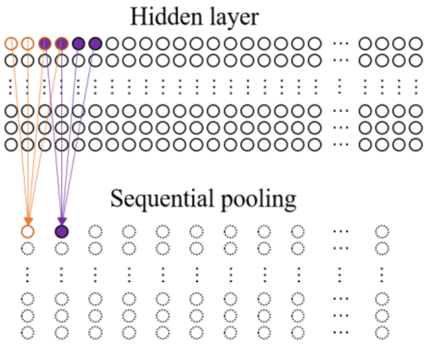



**FIGURE 3.** Schematic diagram of average time series pooling.

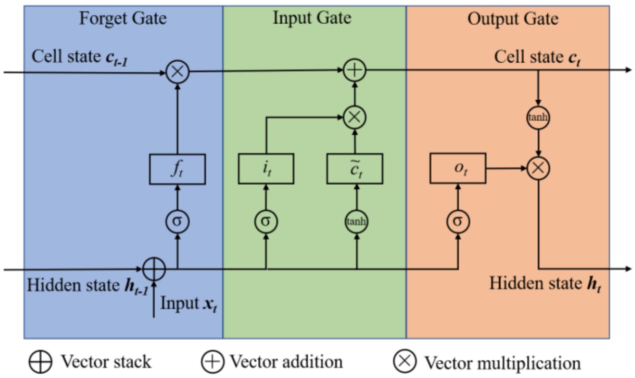



**FIGURE 4.** Architecture of typical LSTM memory cell.

our case mapping eight types of combat intention. The main characteristic of LSTM networks is contained in the hidden layer(s) consisting of memory cells.

Structure of typical LSTM memory cell is shown as Fig. 4, in which $c_{t-1}$ means the memory cell states at time step $t-1$, $c_t$ means the cell states at time step $t$, $\tilde{c}_t$ is the candidate values at time step $t$, $h_{t-1}$ is output vector of memory cell at time step $t$-1, $h_t$ is output value of memory cell at time step $t$, $x_t$ is vector of input sequence at time step $t$, and $f_t$ is the forget gate, $i_t$ is the input gate, $o_t$ is the output gate.

As we can see, an LSTM cell contains three gates inside, forget gate $f_t$, input gate $i_t$, and output gate $o_t$. The design of gating mechanism can control the path of information transmission. The value of "gate" control is between (0, 1), indicating that the proportion of information could pass through.

The forget gate $f_t$ defines which information is removed from the cell state $c_{t-1}$.

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{4}$$

in which $W_f$, $U_f$, $b_f$ are respectively input weight matrics, recurrent weight matrics, bias vector for forget gate, and $\sigma(\cdot)$ is Logistic function.

The input gate $i_t$ specifies which information is added to the cell state.

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

in which $W_i$, $U_i$, $b_i$ are respectively input weight matrics, recurrent weight matrics, bias vector for input gate.

**TABLE 3.** Training parameters of LSTM layer.

| Unit | Symbol | Function |
|---|---|---|
| forget gate | $W_f$ | input weight matrics |
| | $U_f$ | recurrent weight matrics |
| | $b_f$ | bias vextor |
| input gate | $W_i$ | input weight matrics |
| | $U_i$ | recurrent weight matrics |
| | $b_i$ | bias vextor |
| output gate | $W_o$ | input weight matrics |
| | $U_o$ | recurrent weight matrics |
| | $b_o$ | bias vextor |
| memory cell | $W_c$ | input weight matrics |
| | $U_c$ | recurrent weight matrics |
| | $b_c$ | bias vextor |

The output gate $o_t$ specifies which information from the cell state is used as output.

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{6}$$

in which $W_o$, $U_o$, $b_o$ are respectively input weight matrics, recurrent weight matrics, bias vector for output gate.

In order to reduce the input information loss caused by single activation function, candidate state $\tilde{c}_t$ is designed to improve the information acquisition ability of the input gate. The calculation formula of candidate states is as follows,

$$\tilde{c}_t = \tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \tag{7}$$

in which $W_c$, $U_c$, $b_c$ are respectively input weight matrics, recurrent weight matrics, bias vector into the LSTM cell.

By integrating forgetting gate information $f_t$, input gate information $i_t$ and candidate state information $\tilde{c}_t$, LSTM memory cell states $c_t$ can be obtained as follows,

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \tag{8}$$

in which $\odot$ is the Hadamard (elementwise) product.

Then, output value is obtained by synthesizing input gate value $i_t$ and memory cell states $c_t$.

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

To the entire LSTM layer, its input information contains input sequence $X = \left\{x_1, x_2, \ldots x_t, \ldots, x_{n_i}\right\}$, initial memory cell states $c_0$, initial output value of memory cell $h_0$. The input sequence is determined by the output of the previous layer. Initial memory cell states c0 and initial output value of memory cell $h_0$ is obtained by random generation.

To sum up, training parameters for the LSTM layer include input weight matrics, recurrent weight matrics, bias vector for forget gate, input gate, output gate and memory cell. All the training parameters of LSTM layer are shown as Table 3.

The number of weights and bias terms being trained is calculated as follows, let $n_c$ denote the number of memory cell of the LSTM layer, and $n_i$ the number of input features, then the number of parameters of the LSTM layer that needs to be trained is,

$$n_{ls} = 4\left(n_h \times n_i + n_h + n_h \times n_h\right) \tag{10}$$

Hereby $n_h \times n_i$ refers to the dimensions of the input weight matrices, $n_h$ refers to the dimensions of the bias vectors,

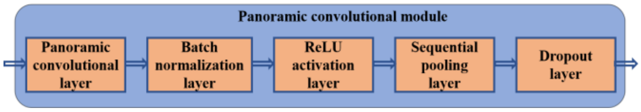



**FIGURE 5.** Architecture of panoramic convolutional module.

$n_h \times n_h$ corresponds to the dimensions of the recurrent weight matrices. Since the input weight matrix, cycle weight matrix and bias vector dimensions of the four gates are the same, the total number of training parameters is 4 times of $n_h \times n_i + n_h + n_h \times n_h$.

#### 4) ARCHITECTURE OF PCLSTM NEURAL NETWORK

Based on the design of panoramic convolutional layer, time series pooling layer and LSTM layer, Panoramic Convolution Long-Short time memory network for aerial targets combat intention recognition is construct by combining typical network layers and related functions of deep learning. The functional modules of PCLSTM mainly include panoramic convolution module, LSTM module and fully connected output module.

The panoramic convolution module aims to explore the essential characteristics of data information at feature dimension. On the basis of panoramic convolution layer and time series pooling layer, batch standardization layer, ReLU activation layer and Dropout layer are added to improve its deep learning performance. And the panoramic convolution module is built with panoramic convolutional layer, batch standardization layer, activation layer, timing pooling layer and Dropout layer in sequence. The architecture of panoramic convolutional module is shown in Fig. 5.

The LSTM module aims to explore the essential characteristics of data information at time dimension. On the basis of LSTM layer, LSTM module improves deep learning ability by stacking dropout layers. LSTM modules are built in the order of LSTM layer and Dropout Layer.

The fully connected output layer aims to transform feature information into the dimension of output classification. And It is formed by stacking the full connection layer and the softmax layer in turn.

According to the specific requirements of aerial target intention recognition, the Panoramic Convolutional LSTM neural network is formed by stacking 4 panoramic convolution modules, 2 LSTM modules and 1 fully connected module. The architecture of PCLSTM is shown in Fig. 6.

### D. BENCHMARK MODELS

For benchmarking the PCLSTM, we choose a standard deep net, i.e., a standard classifier as baseline, a Convolutional Neural Network, i.e., to compare time-domain analysis capabilities with PCLSTM, and a LSTM, i.e., for showing the advantage of PCLSTM.

#### 1) STANDARD DEEP NET

We deploy a standard deep neural network to be a standard classifier as baseline, we use a feed forward neural network with 9 input neurons, 10 neurons in the first, 20 in the second, 20 in the third hidden layer, and 8 neurons in the output layer.

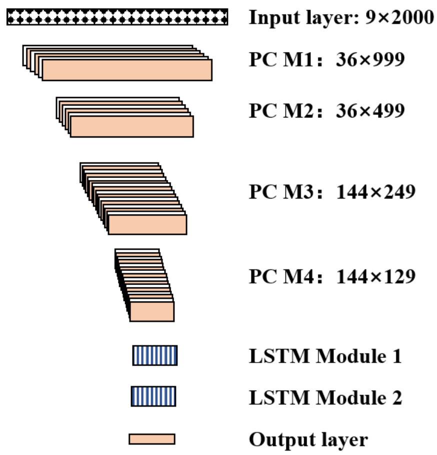



**FIGURE 6.** Architecture of PCLSTM.

The activation function is ReLU, and softmax in the output layer – see ZHOU *et al.* for further details in [14].

#### 2) CONVOLUTIONAL NEURAL NETWORK

The convolutional neural network is the most successful deep learning model. Therefore, it is selected as one of the benchmark models. It is well known that aerial targets intelligence information is composed of multiple time-series data with different characteristics. If we treat the data of each feature dimension on each timestamp as a pixel, the aerial targets intelligence information can be turned into a picture. And the height of the picture is equal to the dimension of characteristic, and the width of the picture is equal to the dimension of time. Therefore, standard convolutional neural networks can be used for classification, architecture and further details of standard convolutional neural networks can be seen in Lecun and Bottou [20].

The convolutional neural network is composed of three convolution modules and three fully connected modules. The convolution module is composed of the convolution layer, batch standardization layer, sigmoid activation layer, pooling layer and dropout layer. The fully connected module consists of the fully connected layer, softmax layer and dropout layer stacked one by one. The last fully connected module, which is the output module, only contains fully connected layer and softmax activation layer. The architecture diagram of convolutional neural network is shown as Fig. 7.

#### 3) LONG SHORT-TERM MEMORY NETWORK

Note that standard deep net and convolutional neural network are both memory-free methods, we choose LSTM as the benchmark model to show the advantage of PCLSTM.

LSTM network is stacked successively with LSTM Layer, dropout layer, fully connected layer, softmax layer, and the optimizer is RMSprop – see Thomas *et al.* for further details in [21].

## IV. SIMULATION AND ANALYSIS

Our results are presented in three stages. First, we analyze the original intelligence information of aerial targets for combat

**TABLE 4.** Confusion matrix instance for a eight-class problem.

| Predicted / actual | penetration | attack | jamming | transportation | refueling | civil flight | AWACS | scout |
|---|---|---|---|---|---|---|---|---|
| penetration | 921 | 53 | 24 | 0 | 0 | 0 | 0 | 2 |
| attack | 42 | 936 | 16 | 0 | 0 | 0 | 0 | 6 |
| jamming | 85 | 116 | 785 | 0 | 0 | 0 | 0 | 14 |
| transportation | 0 | 0 | 2 | 763 | 5 | 195 | 8 | 27 |
| refueling | 0 | 0 | 0 | 13 | 893 | 24 | 63 | 7 |
| civil flight | 0 | 0 | 0 | 215 | 5 | 768 | 2 | 10 |
| AWACS | 0 | 0 | 0 | 3 | 52 | 6 | 927 | 12 |
| scout | 3 | 2 | 9 | 12 | 6 | 2 | 62 | 904 |

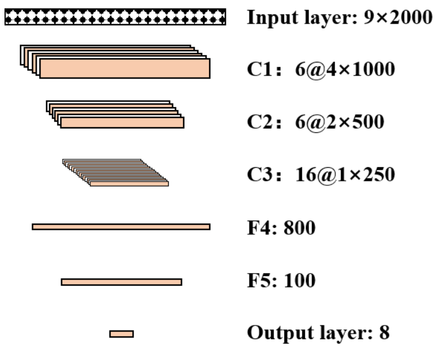



**FIGURE 7.** Architecture of CNN.

**TABLE 5.** Outcome labels of confusion matrix for two-class problem.

| Predicted / actual | Invasion | others |
|---|---|---|
| Invasion | True Positive (TP) | False Negative (FN) |
| others | False Positive (FP) | Ture Negative (TN) |

intention recognition, and contrast the performance of the LSTM network against the standard deep net, convolutional neural network, and the LSTM. Second, in order to find a more reasonable time span of aerial targets combat intention data, we truncate the original data to obtain the intelligence data with different length.

### A. EVALUATION METRICS

To evaluate the performance of our proposed PCLSTM model on aerial targets combat intention recognition, Error Rate (ER), Accuracy Rate (AR) and Recall Rate (RR) are applied to estimate the classification performance.

To better understand the above evaluation metrics, confusion matrix is introduced first. A confusion matrix instance for an eight-class problem involving predicting aerial targets combat intension is shown in table 4.

In the previous eight-class problem, confusion matrix can be viewed as two-class problem by packaging all classes other than one class into others. To each combat intension, such as invasion, the outcome labels of packaging two-class confusion matrix is given in Table 5. If we correctly classify invasion, it's called a True Positive (TP), and it's called a False Negative (FN) when we classify invasion as others. On the contrary, if we misclassify something as invasion, it's outcome label called False Positive (FP), and the rest of the case is called Ture Negative (FN).

On the basis of confusion matrix, ER, AR and RR can be expressed simply. Error Rate is defined as the number of misclassifications divided by the number of test executions, which can be expressed as follows,

$$ER = 1 - \frac{\sum_{i=1}^{n_{class}} TP_i}{n_{total}} \tag{11}$$

in which $n_{class}$ is the number of combat intension type, $TP_i$ is the number of True Positive for the $i$th type of combat intension, $n_{total}$ is the total number of instances tested.

AR tells us the fraction of records that were positive from the group that the classifier predicted to be positive,

$$AR = \frac{TP}{TP + FP} \tag{12}$$

Recall rate measures the fraction of positive examples the classifier got right.

$$RR = \frac{TP}{TP + FN} \tag{13}$$

The error rate is used to evaluate the performance of the classifier over the entire aerial targets combat intention recognition data set. The accuracy rate and recall rate are more concerned with one or more combat intentions in the classification problem. Because accuracy rate pays more attention to the proportion of true positive value, it is more applicable to the combat intention that poses little threat to us but our countermeasures cost more. Recall rate is more concerned with recognizing the proportion of such intentions, it is more applicable to the combat intention which poses more threat to us.

The larger the AR index is, the better the model performance will be, while the smaller the ER and RR index is, the better the model performance will be. In order to facilitate the performance result demonstration in the following text, we design reverse-accuracy rate (R-Acc), it can be calculated as follows,

$$R - Acc = 1 - AR \tag{14}$$

### B. MODEL EVALUTION

In this subsection, hyperparameters of PCLSTM are designed through simulation experiment. Then, the performance of PCLSTM network designed in this paper is verified by comparing with standard deep network, standard convolutional

**TABLE 6.** Minimum error rate on validation set (Epoch = 50).

| batch size / learning rate | $n_b$=8 | $n_b$=16 | $n_b$=32 |
|---|---|---|---|
| $lr$=0.01 | 13.32% | 8.54% | 7.76% |
| $lr$=0.05 | 9.91% | 3.15% | 4.58% |
| $lr$=0.1 | 4.12% | 2.87% | 3.47% |

**TABLE 7.** Minimum error rate on validation set (Epoch = 100).

| batch size / learning rate | $n_b$=8 | $n_b$=16 | $n_b$=32 |
|---|---|---|---|
| $lr$=0.01 | 7.91% | 5.10% | 5.32% |
| $lr$=0.05 | 5.33% | 4.41% | 4.54% |
| $lr$=0.1 | 4.51% | 3.43% | 3.91% |

neural network and LSTM network on the aerial targets combat intention recognition data set. Due to the confidentiality of the information, there is no public data set for aerial target intelligence data, so we use the data collected by ourselves for analysis. The basic information of our aerial target intelligence data is as same as described in Section II. Subsection A. Data, data set with 8,000 time-series data are obtained by normalizing the raw data. And then we can get training subset, verification subset and testing subset through the subset partitioning method described in Section III. Subsection A. Training and testing sets.

#### 1) EXPERIMENTAL DESIGN

Hyperparameters have great influence on the classification performance of neural networks. Therefore, the relevant hyperparameters of PCLSTM neural network are optimized to obtain better recognition performance of aerial targets combat intention. Hyperparameters of PCLSTM neural network mainly include epoch $n_e$, batch size $n_b$, learning rate $lr$, dropout probability $p$ of each layer, and hyperparameters in Adam algorithm, such as exponential decay rate of first order moment estimation $\beta_1$, exponential decay rate of the second moment estimation $\beta_2$. According to the design experience of neural network, hyperparameters $\beta_1$ of Adam optimization algorithm is set as 0.9 and $\beta_2$ as 0.99. To dropout probability, we set input unit dropout probability as 0.8 and hidden unit dropout probability $p$ as 0.5.

That is to say, the hyperparameters that are difficult to set directly include epochs, batch size and learning rate $lr$. According to the design experience of deep neural network, the alternative set of three hyperparameters was set as $n_e = \{50, 100\}$, $n_b = \{8, 16, 32\}$ and $lr = \{0.01, 0.05, 0.1\}$. Since hyperparameters optimization is mainly aimed at the overall performance of the classifier PCLSTM, the error rate ER is used to evaluate the classification and recognition performance of each group of hyperparameters in the test set and verification set. The error rate of the optimal verification set obtained by each group of hyperparameters is shown in Table 6 and Table 7.

It can be seen that, the best classification performance is for parameters $n_e = 50$, $n_b = 16$ and $lr = 0.1$, namely 2.87%. The evolution curve of PCLSTM with previous hyperparameter combination is shown in Fig. 8.

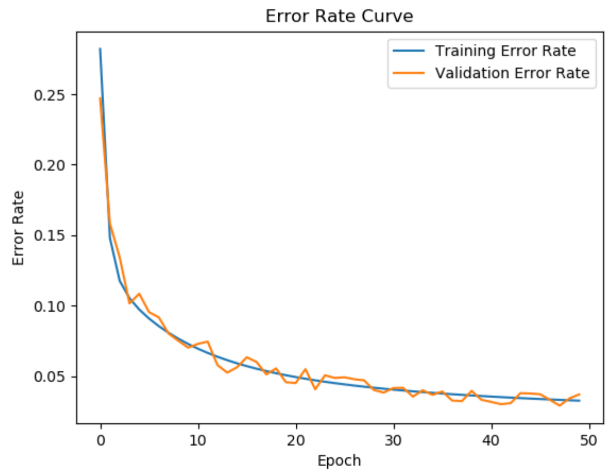



**FIGURE 8.** Error Rate of PCLSTM on validation data set and training data set with hyperparameter epoch = 50, batch size = 16 and $lr$ = 0.1.

**TABLE 8.** Hyperparameters of PCLSTM.

| hyperparameter | symbol | value |
|---|---|---|
| epoch | $n_e$ | 50 |
| batch size | $n_b$ | 16 |
| learning rate | $lr$ | 0.1 |
| exponential decay rate of first order moment estimation | $\beta_1$ | 0.9 |
| exponential decay rate of second order moment estimation | $\beta_2$ | 0.99 |
| dropout probability of input unit | $p_i$ | 0.8 |
| dropout probability of hidden unit | $p_h$ | 0.5 |

Combined with the relevant hyperparameters setting according to experience above, hyperparameters of PCLSTM are set as follows in Table 8.

#### 2) PERFORMANCE ANALYSIS OF PCLSTM

When recognizing the combat intention of aerial targets, we should not only care about the classification performance of the classifier on the whole data set (judged by the error rate), but also make targeted evaluations according to the particularity of some combat intentions. To aerial targets threat us more, such as penetration and attack, recall rates are more focused on unrecognized threats to improve our combat response capability. And the accuracy rate is more important for civil aviation flight, so as to avoid the accidental attack and injury of civil aviation flight as much as possible.

Considering the large amount of randomness in neural network, we take several independent simulation experiments to improve the objectivity of the type performance test. However, neural network analysis requires a large amount of computational cost, so the number of independent simulation tests is often not very large. Since there is no published data set, it is difficult to make a direct comparison with existing literature on the same data set. Therefore, we take the method of the existing typical literature as the benchmark model and make a comparative analysis with PCLSTM on the data set we collected. Further details of standard deep network, standard CNN and LSTM can be seen in Section III.

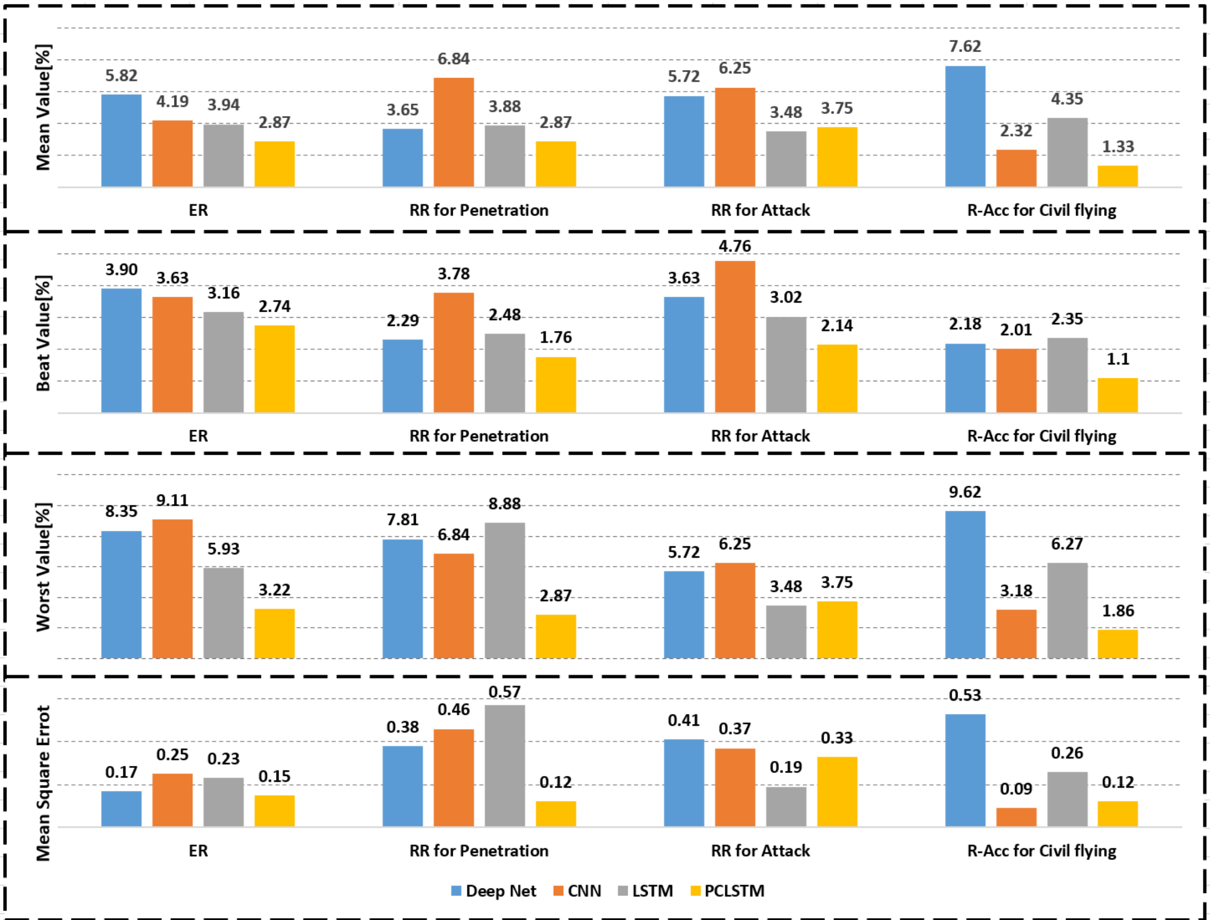



**FIGURE 9.** Performance of four networks in different evaluation metrics.

Subsection D. benchmark models. According to the computational cost of each neural networks in this paper, 10 independent simulation tests are conducted for each neural network. Performance of PCLSTM, standard deep network, standard CNN and LSTM are shown in Fig. 9.

Simulation results show that, PCLSTM has the optimal ER on the overall data set, and it reduces the error rate by about half than the other classifiers. Moreover, PCLSTM has a lower recall rate in the recognition accuracy rate of penetration and attack. Recall rate of penetration is reduced more than 20%, and attack 40%. In terms of the accuracy rate of civil flying, PCLSTM also has a better recognition performance. The accuracy rate of PCLSTM is 1.32% higher than CNN, which is the second-best classifier. The improvement of the accuracy rate will effectively reduce the mis-hit and mis-hurt to civilian aircraft, and better safeguard the safety of people's lives in the face of enemy threats.

To sum up, PCLSTM classifier not only has better performance on the error rate of overall recognition, but also has a good performance in specific combat intention recognition such as penetration, attack, civil flying and so on.

### C. OPTIMIZATION OF DATA LENGTH TRUNCATION

The previous recognition of aerial targets combat intention are all based on the complete data acquiring from aerial targets. However, in the process of carrying out the aerial targets combat intention recognition mission, it is not expected to analyze after obtaining the full-time data. When the aerial target enters radar coverage area, combat intention should be recognized with as little data as possible. In other words, aerial target combat intention is expected to be effectively recognized within a short time range. Therefore, this subsection intends to truncate the intelligence information data into different lengths, and select the optimal data length truncation for combat intention recognition by comparing performance of truncated intelligence data with different length.

#### 1) DATA TRUNCATION

Without loss of generality, it is assumed that the data length after truncation is $t_1$. The way of data truncation is as follows. For each piece of data in the original data set, the $t_1$ length fragment is truncated in a random way. The truncated data is put into the new dataset, and the remaining two data fragments are returned to the original dataset. The above truncation is repeated on the original data set until the length of the remaining data fragments in the original data set are all less than $t_1$. Thus, the truncated data set of $t_1$ length is obtained.

#### 2) SIMULATION TEST

In order to compare the combat intention recognition capability of different duration data, we set the truncated length

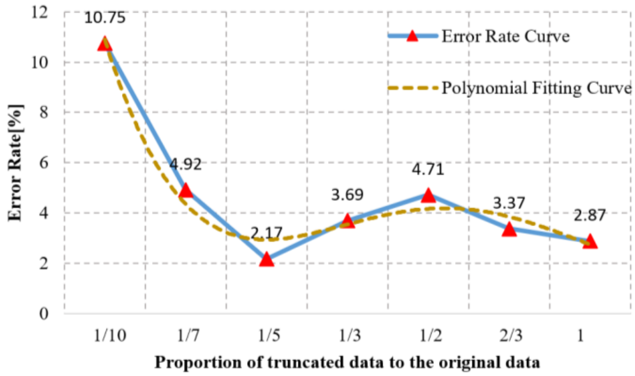



**FIGURE 10.** Error rate on each truncated data set.

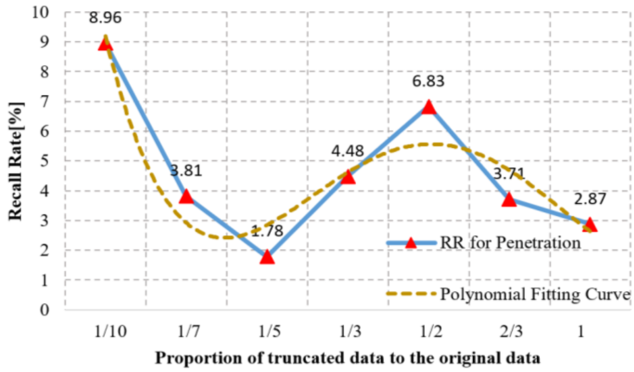



**FIGURE 11.** Recall rate for penetration on each truncated data set.

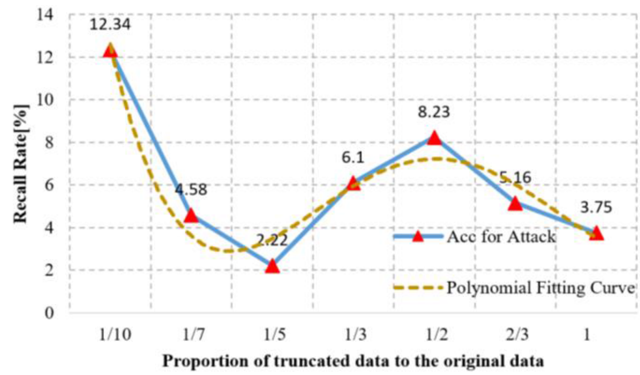



**FIGURE 12.** Recall rate for attack on each truncated data set.

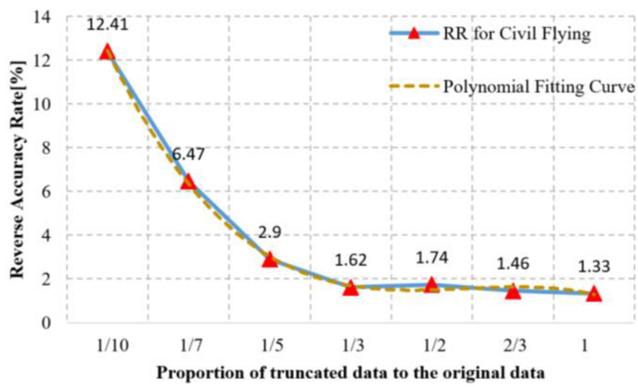



**FIGURE 13.** Reverse Accuracy rate for civil flying on each truncated data set.

as 1/10, 1/7, 1/5, 1/3, 1/2 and 2/3 of the original data length respectively, and denote them as $t_1$, $t_2$, $t_3$, $t_4$, $t_5$ and $t_6$. The number of data in the truncated data set obtained by different truncation duration is different. In a random simulation test, the sizes of each truncated data set are 58,462, 41,973, 30,868, 20,023, 8,000 and 8,000, and the corresponding time lengths are $t_1$, $t_2$, $t_3$, $t_4$, $t_5$ and $t_6$ respectively.

According to the training set and test set allocation method introduced in Section III. subsection A. training and testing sets, randomly select 80% of each truncated data set as its training set. And then train the PCLSTM network with hyper-parameters in Table 8. After the trained PCLSTM neural network is obtained, simulation experiments are carried out on each truncated test sets respectively. The performance results on each truncated test sets are shown in Fig. 10, 11, 12 and 13 in terms of the overall error rate, recall rate of penetration, recall rate of attack, Accuracy rate of civil flying respectively.

According to Fig. 10, the error rate first decreases and then increases with the growth of truncation time. The lowest rate of recognition error occurs at 1/5 of the original length. To recall rate of penetration, the curve shows a trend of first decreasing and then increasing in Fig. 11, and the lowest recall rate occurs at 1/5 of the original data. As shown in Fig. 12, recall rate of attack is very similar to penetration. To accuracy rate (or reverse accuracy rate) of civil flying, the curve in Fig. 13 shows a trend of first decreasing, then a small increase, and then continued to fall down. The optimal reverse accuracy rate of civil flying is the test based on the original data set.

The recognition performance of classifier is closely related to the size of data set and the information amount of a single truncated data. As the size of the data set decreases, the recognition performance gradually decreases. At the same time, the recognition performance will continuously improve with the growth of truncation time. Therefore, the overall recognition performance is bound to increase first and then decrease, which is mutually verified with the results of our simulation test.

According to the classifier recognition performance trend obtained above, the overall error rate, recall rate of penetration and attack all show the optimal performance at $t_3$, and accuracy rate of civil flying shows good performance at data sets more than 1/3 of the original data set. By sorting the optimal data duration for different combat intentions, the following combat intention recognition method is designed. The aerial targets combat intention recognition is carried out with two truncated length, $t_3$ and $t_4$ respectively. When the aerial target is detected after $t_3$, the first combat intention recognition is carried out, and countermeasures are directly implemented for penetration and attack. To other combat intensions, we should prepare the corresponding countermeasures, but not carry out countermeasures for the time being.

And then, comprehensive judgment should be made based on the secondary recognition results on truncated data set of $t_4$. If it is recognized as civil flying by secondary recognition, the aerial target combat intension is determined as civil flying. If it is not recognized as civil flight, the combat intention is determined according to the recognition result on $t_3$ truncated data set.

## V. CONCLUSION

In view of the important role of aerial targets combat intention recognition in modern air combat, we design a novel deep learning method to improve the recognition ability. Firstly, the panoramic convolutional layer, time series pooling layer and LSTM layer are designed according to the characteristics of aerial targets intelligence information. Then the Panoramic Convolution LSTM neural network is constructed by stacking panoramic convolution module, LSTM module and fully connected output module. On this basis, simulation test is carried out with performance comparison indexes error rate of overall data set, recall rate of penetration, recall rate of attack and accuracy rate of civil flying. In the simulation experiment, benchmark models include standard deep net, standard convolutional neural network and LSTM neural network.

Simulation results show that, PCLSTM not only has the optimal error rate, but also has obvious advantages in critical combat intention recognition than any other benchmark models. Therefore, the PCLSTM designed in this paper can better complete the task of aerial target combat intention recognition.

In terms of truncation length optimization of intelligence information, we compare and analyze the performance of several combat intentions on each truncation data sets, and design the following countermeasures. First, determine the combat intention of penetration and attack within 1/5 of the cut-off time, and take corresponding countermeasures directly. And the rest of the combat intentions can only be used as references to prepare corresponding countermeasures, but not to carry out counter-strikes for the time being. Third, determine the combat intention (other than penetration and attack) within 1/3 of the cut-off time. If it is recognized as civil flying in this step, the aerial target combat intension is determined as civil flying. If it is not recognized as civil flight, the combat intention is determined according to the recognition result in first step.

In conclusion, we successfully design the PCLSTM classifier which has great advantage for recognition of aerial targets combat intention. Moreover, the specific determination methods for different combat intentions are obtained by optimizing the truncation data length. 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.

## REFERENCES

[1] K. Zhou, R. Wei, Q. Zhang, and Z. Xu, "Learning system for air combat decision inspired by cognitive mechanisms of the Brain," *IEEE Access*, vol. 8, no. 1, pp. 8129–8155, Jan. 2020.

[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] W. Jiang, D. Han, X. Fan, and D. Duanmu, "Research on threat assessment based on dempster-shafer evidence theory," in *Proc. GCN-LNEE*, Dordrecht, The Netherlands, vol. 113, 2012, pp. 975–984.

[4] M. W. Floyd, J. Karneeb, and D. W. Aha, "Case-based team recognition using learned opponent models," in *Proc. 25th Int. Conf. ICCBR*, Trondheim, Norway, Jun. 2017, pp. 123–138.

[5] 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. 12, no. 4, pp. 479–490, Jul. 1982, doi: 10.1109/TSMC.1982.4308852.

[6] R. L. Carling, "Naval situation assessment using a real-time knowledge-based system," *Nav. Eng. J.*, vol. 111, no. 3, pp. 44–46,53, Jun. 2010.

[7] 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.

[8] X. Niu, H. Zhao, and Y. Zhang, "Naval vessel intention recognition based on decision tree," *Ordnance Ind. Autom.*, vol. 29, no. 6, pp. 473–482, May/Jun. 2010.

[9] 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, pp. 1–19, Feb. 2020, Art. no. 2600111.

[10] J. Qing, G. Xiantai, J. Weidong, and W. Nanfang, "Intention recognition of aerial targets based on Bayesian optimization algorithm," in *Proc. 2nd IEEE Int. Conf. Intell. Transp. Eng. (ICITE)*, Piscataway, NJ, USA, Sep. 2017, pp. 356–359.

[11] Z. G. Chen and X. F. Wu, "A novel multi-timescales layered intention recognition method," *Appl. Mech. Mater.*, vols. 644–650, pp. 4607–4611, Sep. 2014, doi: 10.4028/www.scientific.net/AMM.644-650.4607.

[12] X. M. Xu, R. N. Yang, and Y. Fu, "Situation assessment for air combat based on novel semi-supervised naive Bayes," *J. Syst. Eng. Electron.*, vol. 29, no. 4, pp. 768–779, Aug. 2018.

[13] 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.

[14] W. W. Zhou *et al.*, "Combat intention recognition for aerial targets based on deep neural network," (in Chinese), *Acta Aero-Nauticaet Astronautica Sinica*, vol. 39, no. 11, pp. 1–9, Aug. 2018, Art. no. 322468.

[15] Z. Yao, D. He, Y. Chen, B. Liu, J. Miao, J. Deng, and S. Shan, "Inspection of exterior substance on high-speed train bottom based on improved deep learning method," *Measurement*, vol. 163, pp. 1–12, May 2020, Art. no. 108013.

[16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[17] A. Graves, "Generating sequences with recurrent neural networks," 2013, *arXiv:1308.0850*. [Online]. Available: http://arxiv.org/abs/1308.0850

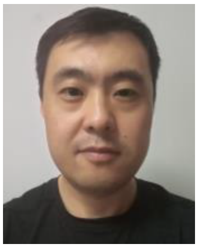
[18] C. Olah. (2015). *Understanding LSTM Networks*. [Online]. Available: http://colah.github.io/posts/2015-08-Understanding-LSTMs/

[19] H. Sak, A. Senior, and F. Beaufays, "Long short-term memory based recurrent neural network architectures for large vocabulary speech recognition," 2014, *arXiv:1402.1128*. [Online]. Available: http://arxiv.org/abs/1402.1128

[20] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[21] T. Fischer and C. Krauss, "Deep learning with long short-term memory networks for financial market predictions," *Eur. J. Oper. Res.*, vol. 270, no. 2, pp. 654–669, Oct. 2018.

**JUNJIE XUE** was born in Zhouzhi, Xi'an, Shannxi, China, in 1988. He received the B.S. degree, the M.S. degree in aerospace science and engineering, and the Ph.D. degree in systems engineering form Air Force Engineering University, China, in 2010, 2012, and 2016, respectively. He is currently a Lecturer with Air Force Engineering University. His research interests include military systems engineering, military operation research, air combat planning, data mining, and machine learning.



**JIE ZHU** was born in Hubei, China, in 1980. He received the M.S. degree in military traffic engineering form Air Force Engineering University, China, in 2004. He is currently an Associate Professor with Air Force Engineering University. His research interests include air traffic control and aviation cluster collaboration.

**JIYANG XIAO** was born in Changde, Hunan, China, in 1988. He received the Ph.D. degree in science of command from Air Force Engineering University, China, in 2016. He is currently an Associate Professor with Air Force Engineering University. His research interests include military operation research, air combat planning, aviation mission planning, data mining, and machine learning.

**SHENG TONG** was born in Xi'an, Shanxi, China, in 1990. He received the M.S. degree in systems engineering from Air Force Engineering University, China, in 2015, where he is currently pursuing the Ph.D. degree. His research interests include air combat planning, aviation mission planning, data mining, and machine learning.

**LING HUANG** was born in Wushan, Chongqing, China, in 1993. He received the M.S. degree in applied physics from the National University of Defense Technology, China, in 2016. He is currently a Lecturer with Air Force Engineering University. His research interests include air combat planning, air command guidance, and machine learning.