ARTICLE

KGTLIR: An Air Target Intention Recognition Model Based on Knowledge Graph and Deep Learning

Bo Cao[1,*], Qinghua Xing[2], Longyue Li[2], Huaixi Xing[1] and Zhanfu Song[1]

[1]Graduate School, Air Force Engineering University, Xi'an, 710051, China

[2]Air Defense and Antimissile School, Air Force Engineering University, Xi'an, 710051, China

*Corresponding Author: Bo Cao. Email: nuyoah_bc@163.com

ABSTRACT

As a core part of battlefield situational awareness, air target intention recognition plays an important role in modern air operations. Aiming at the problems of insufficient feature extraction and misclassification in intention recognition, this paper designs an air target intention recognition method (KGTLIR) based on Knowledge Graph and Deep Learning. Firstly, the intention recognition model based on Deep Learning is constructed to mine the temporal relationship of intention features using dilated causal convolution and the spatial relationship of intention features using a graph attention mechanism. Meanwhile, the accuracy, recall, and F1-score after iteration are introduced to dynamically adjust the sample weights to reduce the probability of misclassification. After that, an intention recognition model based on Knowledge Graph is constructed to predict the probability of the occurrence of different intentions of the target. Finally, the results of the two models are fused by evidence theory to obtain the target's operational intention. Experiments show that the intention recognition accuracy of the KGTLIR model can reach 98.48%, which is not only better than most of the air target intention recognition methods, but also demonstrates better interpretability and trustworthiness.

KEYWORDS

Dilated causal convolution; graph attention mechanism; intention recognition; air targets; knowledge graph

1 Introduction

With the introduction of new operational concepts such as Decision-Centric Warfare and Joint All-Domain Operation, as well as the development of modern weapons, the intensity, timeliness, and complexity of confrontation in the air and space domain have increased dramatically. This requires us to quickly obtain and process battlefield information, infer the enemy's target's operational intention, realize clear and effective battlefield situation assessment, and implement reasonable and favorable decision-making [1].

As early as the 1970s, some countries had begun to use information systems to aid in intent recognition. With the development of decision support systems, more and more researchers have begun to study the problem of tactical intent recognition of air targets. Existing methods for target intention

recognition are mainly based on statistical theory (evidence theory [2–4], Bayesian networks [5–7]), cognitive modeling (template matching [8], expert systems [9]), and artificial intelligence [10–14].

Xia et al. used a gray Markov chain to analyze and predict four factors of enemy UAVs: speed, angle, attack, and detection, while combining the predictors with rules provided by rough sets. The proposed model can infer the intentions of enemy drones in a short time in the future [2]. Sun et al. used a similarity function in a high-dimensional data space to measure the degree of support of a target's state for its intention, and combined the D-S evidence theory to combine the support of each temporal sequence to form a sequential identification of the target's tactical intention [3]. Jun et al. combined confidence rule base and evidence theory to identify the intentions of aerial targets [4]. Xu et al. applied the dynamic sequential Bayesian network to the problem of target intention recognition, and optimized the algorithm through the information entropy theory, and achieved better results [5]. Meng et al. proposed an improved algorithm based on data classification confidence based on a semi-supervised plain Bayesian classifier to achieve effective identification of air combat target intention [6]. Chen et al. constructed a posture template based on expert experience and used D-S evidence theory to construct an inference model for intention recognition [8]. Yin et al. used a statistical approach to the a priori knowledge to obtain the target operational intent knowledge and rule base and proposed an intention recognition method based on discriminant analysis [9]. However, the real battlefield environment has a high degree of complexity and uncertainty. Evidence theory methods have problems in application such as limited ability to deal with conflict evidence and difficulty in constructing basic probability assignment functions. Bayesian network methods have problems such as difficulty in parameter estimation and poor adaptability to the dynamically changing battlefield environment. Template matching methods have problems such as difficulty in constructing and updating template libraries and limited ability to handle complex battlefield information. The expert system has problems such as difficulty in knowledge acquisition, limited real-time, and practicality.

In response to the above problems, researchers continue to try new methods such as deep learning and reinforcement learning. Zhou et al. proposed an air target intention recognition method based on a deep neural network, which optimizes the backpropagation algorithm using ReLU (Rectified Linear Unit) function and Adaptive Moment Estimation (Adam) algorithm to improve the model recognition effect [10]. To address the temporal nature of target intention features, Teng et al. combined a Temporal Convolutional Network (TCN) and a Bidirectional Gated Recurrent Unit (BiGRU) to extract the temporal characteristics of the features and introduced an attention mechanism to assign different weights to the features [11]. Aiming at the limitations of the existing algorithms such as relying on empirical knowledge, having difficulty to extract full temporal features, and being unable to meet the requirements of real combat, Wang et al. proposed a target tactical intention recognition algorithm based on bidirectional long-short temporal memory (BiLSTM), with a recognition accuracy rate of up to 92% [12]. Aiming at the deficiencies of existing methods in terms of temporal order and interpretability, Wang et al. designed an air target intention recognition model based on BiGRU and Conditional Random Field (CRF), which achieved high recognition accuracy [13]. Qu et al. proposed an air target intention recognition method based on fully connected neural networks, convolutional neural networks, and recurrent neural networks, which realizes the target intention recognition function based on real-time situational information, and the proposed method has good robustness [14].

Although the above methods have achieved better recognition results, there are the following problems in the modeling process. Firstly, most models only consider the temporal characteristics of air combat data, ignoring the influence of its spatial characteristics on the recognition results. Secondly, although most models have high recognition accuracy, there is still room for improvement in accuracy.

In the process of combat, once the intention of enemy targets is incorrectly recognized, it will bring high combat costs or even lead to the failure of the whole battle. Third, most models ignore the effects of misclassification in the process of recognizing target intent. For example, there is a great similarity between the features of attack intention and feint intention, which can be easily misidentified.

To address the above problems, this paper proposes an air target intention recognition method (KGTLIR) based on Knowledge Graph and Deep Learning, where K denotes the Knowledge Graph, G denotes the graph attention mechanism, T denotes the temporal module based on dilated causal convolution, L denotes the improved cross-entropy loss function, and IR denotes intention recognition.

The main contributions of this paper are as follows.

(1) The problem of air target intention recognition is analyzed in detail, the air target intention space and intention features set are constructed and coded uniformly, and the decision maker's empirical knowledge is encapsulated into intention labels.

(2) The temporal module based on dilated causal convolution is constructed to solve the temporal feature extraction problem of air combat data; meanwhile, the graph attention mechanism is introduced to solve the spatial feature extraction problem of air combat data. By combining the two, the feature extraction capability of the model is greatly improved.

(3) Sample weights are introduced to improve the cross-entropy loss function, and at the same time, combined with the attention mechanism module in GAT, to realize the dynamic adjustment of sample category weights and reduce the misclassification probability of the model.

(4) Construct an air target intention recognition model based on Knowledge Graph, which mainly consists of a data level and a schema level. The original data is stored in the data level, and the probability distribution of different intentions of the target is predicted through the schema level. The construction of the knowledge graph model can fully explore the potential relationship of the data, which makes the model better understand the expert experience and improves the credibility of the model.

(5) Evidence theory is used to fuse the recognition results of the knowledge graph model and the GTLIR model, after which the target intention is decided. The accuracy of the model is improved to 98.48%.

The rest of this paper is organized as follows. The second part describes in detail the problem of air target operational intention recognition and the construction process of air target intention space and intention features set. The third part describes in detail the air target intention recognition model based on Deep Learning and the air target intention recognition model based on Knowledge Graph. The fourth part presents the experimental results and experimental analysis. The fifth part summarizes the whole paper and points out the direction of the research.

2 Description of Air Target Intention Recognition

2.1 Air Target Intention Recognition Problem

For the intention recognition task in air defense operations, firstly, extract the battlefield information in time and airspace from the air defense battlefield environment as well as the static and real-time dynamic attributes of air targets, and then analyze these elements to reason the enemy's air targets' operational intention [15]. The specific process is shown in Fig. 1.
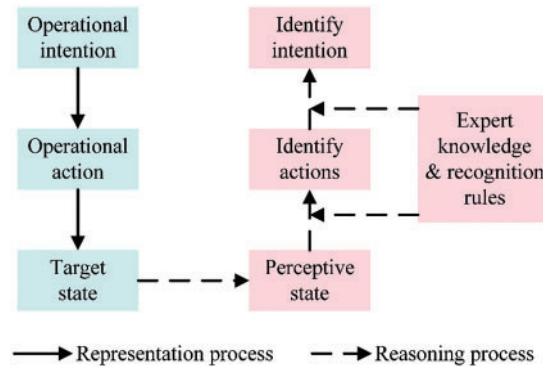
Figure 1: Reasoning process of the operational intention of air targets in air defense operations

The air target intention recognition problem is a specific mapping of target intention features to target intention types. Target intention recognition in air combat occurs in a real-time, complex, and high-confrontation battlefield environment. There are certain limitations in the ability to obtain and analyze information, which usually makes it difficult for our commanders to accurately obtain the combat intention of enemy targets. Therefore, in this paper, we use the probability that an intention may occur to represent the outcome of intention recognition at that moment. In the actual combat environment, enemy targets may conceal their tactical intention through a variety of means, and there is a certain degree of one-sidedness and falsity in judging the target's tactical intention based only on the battlefield data of a single moment. Therefore, in this paper, we choose to extract target features from battlefield data obtained at consecutive moments and thus recognize their tactical intentions. Define the set of temporal features of the target $I_t = \{I_{t_1}, I_{t_2}, \ldots, I_{t_n}\}$, I_t consists of consecutive air combat real-time features from moment t_1 to moment t_n. Define Y_t as the tactical intent space of the target. The air target intention recognition problem can then be described as a mapping process from I_t to Y_t.

$$Y_t = f(I_t) = f\left(I_{t_1}, I_{t_2}, \ldots, I_{t_n}\right) \tag{1}$$

In this paper, we build a KGTLIR model by training the data obtained from the battlefield simulation system, so as to implicitly establish the mapping relationship between features and intentions. The air target intention recognition process is shown in Fig. 2.

2.2 Air Target Intention Space

For different battlefield environments, different combat styles, and different combat objectives, the intention space has large differences, so it needs to be defined in conjunction with the specific combat context. This paper mainly focuses on air targets in air defense operations, such as reconnaissance planes, fighter planes, early warning planes, etc., and determines that the intention space of enemy air targets is {attack, penetrate, interference, feinting, surveillance, reconnaissance, retreat} seven kinds of intention. The details of the different intentions are shown in Table 1.

The intention recognition problem is also a multi-classification problem, this paper establishes a coding and decoding mechanism as shown in Fig. 3, which facilitates the training of the model by coding the intention of the enemy target.
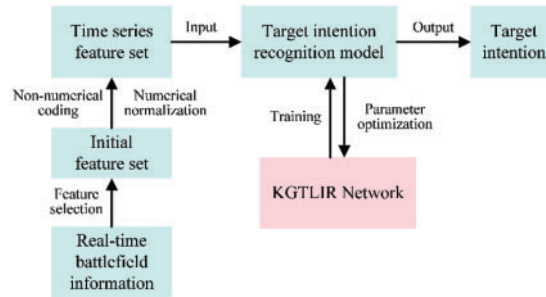
Figure 2: KGTLIR-based air target intention recognition process. The model consists of three main phases. First, the real-time battlefield information is feature-selected to obtain the initial dataset. Secondly, the numerical data in the dataset is normalized and the non-numerical data is coded to obtain the time series dataset. Finally, the processed data is fed into the intention recognition model (KGTLIR) to get the target intention

Table 1: Detailed description of the operational intention of air targets

Intention	Detailed description
Attack	Firing of missiles, bombs, etc., by enemy air targets against air defense sites to achieve destructive effects
Penetrate	Actions by enemy air targets to break through our air defenses and enter the intended target area
Interference	Jamming of our radar and communications systems by enemy air targets through electronic equipment
Feinting	Enemy air targets to conceal combat attempts and to deceive and confuse our side by attacking with a certain number of troops to create a false impression of combat operations
Surveillance	Passive surveillance of the area by enemy air targets
Reconnaissance	Active detection activities by enemy airborne targets against our air defenses and strongholds
Retreat	Enemy air targets departed from the current area of operations
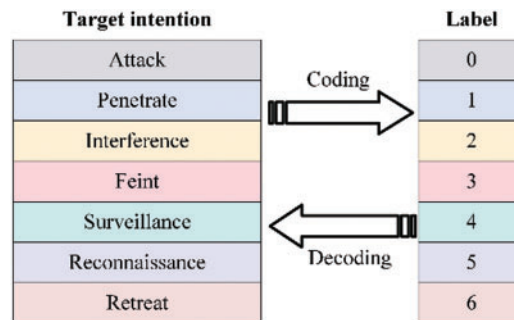
Figure 3: Air target operational intention coding and decoding method

2.3 Air Target Intention Feature Selection

Once the target intention space has been determined, the feature information needed for the model can be determined based on the relationship between the target attribute features and the intention. From the perspective of the combat mission attempted by the enemy, there are certain differences in the characteristic information displayed by enemy airplanes when they perform different combat missions [16].

(i) As far as the flight speed of the target is concerned, the flight speed of fighter planes is generally 735 to 1470 km/h when they are engaged in aerial combat. The flight speed of bombers and transport planes is generally 600 to 850 km/h when they are carrying out their missions. The flight speed of early warning planes is generally 750 to 950 km/h when they are carrying out surveillance missions.

(ii) As far as the flight altitude of the target is concerned, when the enemy aircraft performs the mission of surprise defense, the low altitude of surprise defense ranges from 50 to 200 m, and the high altitude of surprise defense ranges from 10,000 to 11,000 m. The altitude of the fighter aircraft in the aerial combat is generally from 1000 to 6000 m. When the enemy aircraft performs the mission of reconnaissance, the flight altitude of the low altitude reconnaissance ranges from 100 to 1000 m, and that of the ultra-high altitude reconnaissance ranges from 15,000 m or more.

(iii) As far as the radar status of the target is concerned, fighter planes usually keep their air-to-air radar on when they are engaged in aerial combat. Transport planes generally keep their radar in a silent state when they are performing transport missions. Bombers keep their air-to-air radar or sea-to-sea radar on only when they are performing bombing missions. Planes keep their air-to-air radar and sea-to-sea radar on when they are performing reconnaissance missions.

In addition, considering the limitations of information technology, there are target features that cannot be directly acquired, such as target type, shape, etc., this paper chooses to use radar one-dimensional distance image and radar reflection cross-sectional area instead.

In summary, the air defense target intention feature set constructed in this paper is {altitude, velocity, acceleration, heading angle, azimuth, distance, radar one-dimensional distance image, radar reflective cross-sectional area, air-to-air radar status, ground-to-ground radar status, interference state, and target type} in 12 dimensions. The first eight are numerical features, and the last four are non-numerical features. The detailed descriptions of the features are shown in Table 2.

Table 2: Air target intention feature set information

Feature name	Description	Data type	Unit
Altitude	Flight altitude of enemy air targets	Numerical	km
Speed	Flight speed of enemy air targets	Numerical	m/s
Acceleration	Flight acceleration of enemy air targets	Numerical	m/s^2
Heading angle	Direction of flight of enemy airborne targets	Numerical	°
Azimuth	The angle between the direction of flight of enemy airborne targets and my air defenses	Numerical	mil

(Continued)

Table 2 (continued)

Feature name	Description	Data type	Unit
Distance	Distance of enemy air targets from the center of my air defense positions	Numerical	km
1D range profile	Distance between the nearest and farthest peaks in a one-dimensional distance distribution	Numerical	m
RCS	The size of the radar cross-sectional area	Numerical	m^2
Air-to-air radar status	Air-to-air radar is on or off.	Non-numerical	Coding
Ground penetrating radar status	Ground penetrating radar is on or off.	Non-numerical	Coding
Interference state	Enemy aircraft electronic jamming devices are on or off	Non-numerical	Coding
Target type	Types and models of enemy air targets	Non-numerical	Coding

For the eight numerical types of data in the table, the Min-Max normalization method was used to map them to the interval [0,1]. The specific calculations are as follows:

$$x' = \frac{x - \min}{\max - \min} \tag{2}$$

where x represents the eigenvalue of a numerical type of feature, x' represents the normalized result of this numerical type of feature, min represents the minimum value of this feature value in the set and max is the maximum value of this feature value in the set.

The four kinds of non-numerical data in the table are all categorical data. In order to facilitate the learning of the neural network, it is necessary to numericalize these four kinds of data, mapping them into the interval between [0,1]. The specific calculations are as follows:

$$x = \frac{i - 1}{j} \tag{3}$$

where j represents the size of the classification space and x is the value of the original i-th class non-numeric feature after mapping it to the interval [0,1].

3 Air Target Intention Recognition Model

Aiming at the problems of insufficient feature extraction and misclassification mentioned in the introduction, this paper proposes an intention recognition model called KGTLIR, as shown in Fig. 4. The raw data collected from the battlefield environment is simultaneously fed into an intention recognition model based on deep learning (GTLIR) and an intention recognition model based on Knowledge Graph to obtain the probability distributions p_1 and p_2 for different intentions. Afterward, p_1 and p_2 are fused through evidence theory to get the final intention of the target.
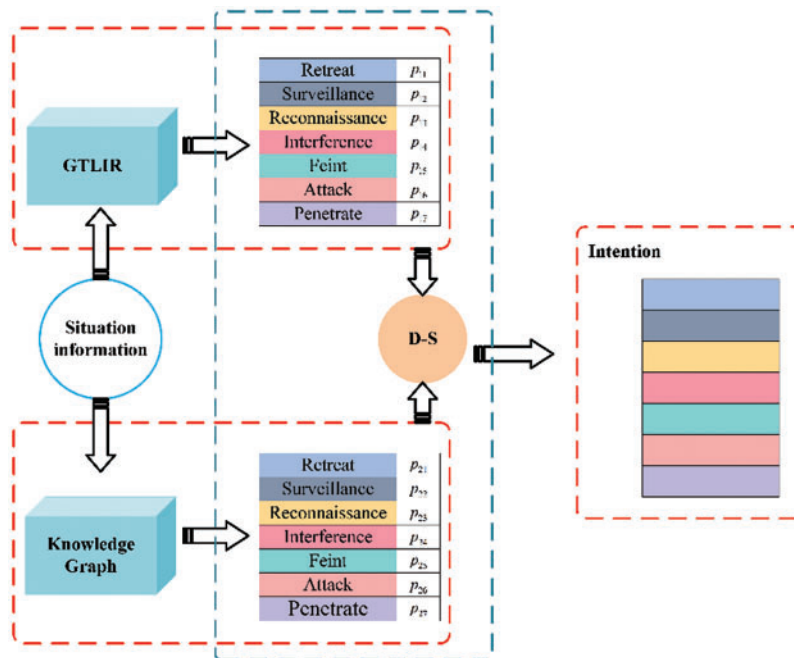
Figure 4: Overall framework of the KGTLIR model. The input of the model is situational information. Input situational information into the GTLIR model to obtain the probability p_{1i} of different intentions of the target. Input situational information into the Knowledge Graph model to obtain the probability p_{2i} of different intentions of the target. The evidence theory is utilized to fuse p_1 and p_2 to get the final target intention

3.1 Air Target Intention Recognition Model Based on Deep Learning

This section focuses on the intention recognition model (Graph Temporal Net Intention Recognition, GTLIR). Its steps are shown as follows:

First, data preprocessing. Raw data collected from the battlefield environment is characterized by real-time, complexity and diversity. This paper transforms non-numerical features into numerical ones by the method mentioned in Section 2.3, after which the features are normalized to form a standard dataset.

Second, temporal feature extraction. To effectively extract the temporal features of the data, inflated causal convolution is introduced in the paper to model the links between inputs at different moments and mine the information embedded in the intention data.

Third, spatial feature extraction. To effectively extract the spatial features of the data, the graph attention mechanism is introduced in the paper, which has a powerful feature extraction capability to capture the connection between multiple features. The raw inputs are fused after the features are extracted by one-dimensional convolution and inflated causal convolution, respectively, after which they are transported to the graph attention mechanism to realize the extraction of temporal and spatial features.

Fourth, intention recognition. The output of the graph attention mechanism will be delivered to the SoftMax layer for classification, and finally get the result of target intention recognition. The model is trained with an improved cross-entropy loss function to calculate the loss value.

The overall framework of the model is shown in Fig. 5. $V \in R^m$ denotes the processed standard feature matrix, X denotes the output of the one-dimensional convolutional layer, S denotes the output of the Temporal layer, H denotes the output of the fusion of X and S, H′ denotes the output of the Graph Attention Mechanism layer, and Y denotes the final output of the target intent.
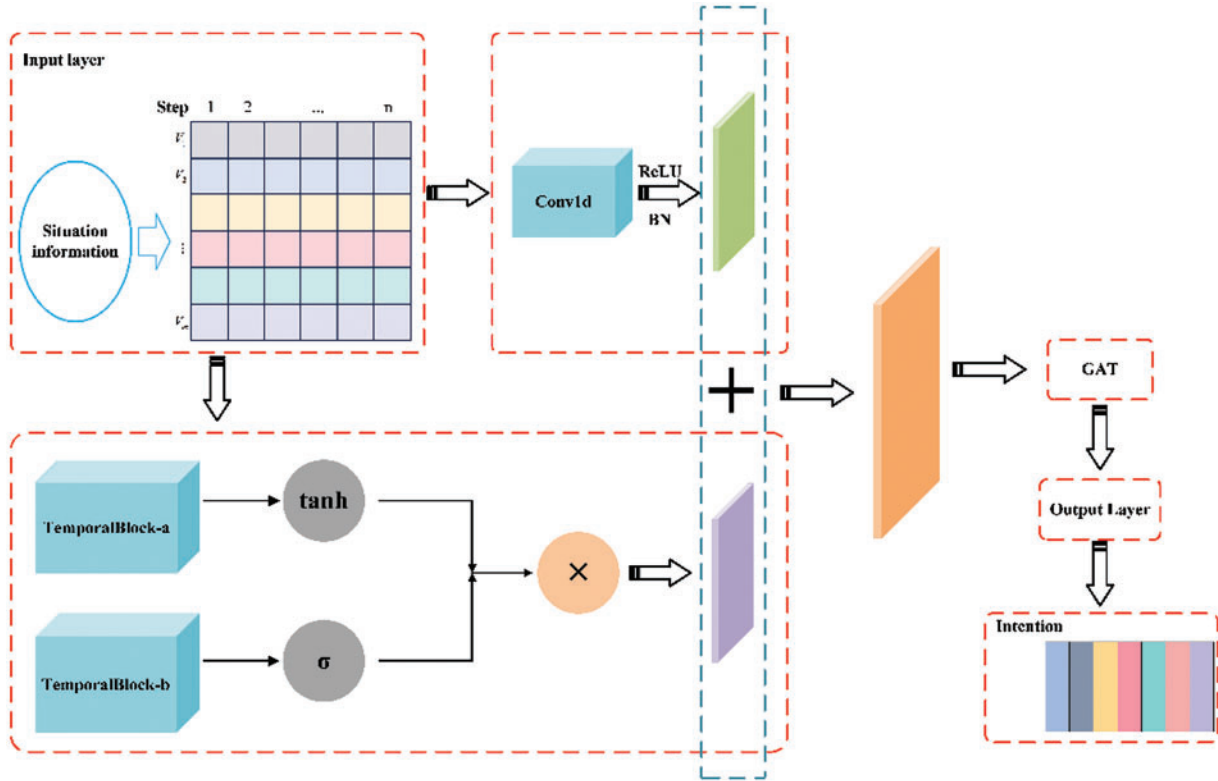

**Figure 5:** Overall framework of the GTLIR model

3.1.1 Temporal Block

Dilated Causal Convolution is a network structure that can handle time series data as proposed by Bai et al. in 2018 in Temporal Convolution Neural Network [17]. Compared to CNN, causal convolution can predict y_t based on $x_1, x_2, \ldots, x_t$ and $y_1, y_2, \ldots, y_{t-1}$, making y_t close to the true value. The computational process of causal convolution is shown in Fig. 6.

Dilated causal convolution adds a dilated factor to causal convolution [18]. It can realize the exponential expansion of the sense field without increasing the parameters and model complexity, which makes it able to deal with longer time series data, as shown in Fig. 7.

Given the model input V, $f \in R^k$ represents the size of the one-dimensional inflated causal convolution kernel, the output after the inflated causal convolution operation is shown in Eq. (4).

$$F(V, s) = (V * df)(s) = \sum_{i=0}^{k-1} f(i) \cdot V_{s-d \cdot i} \tag{4}$$

where d represents the expansion factor, k represents the size of the convolution kernel and $s - d \cdot i$ represents the position point corresponding to the input sequence. The exponential expansion of the inflated causal convolutional sensory field enables it to cover all valid inputs of the input time series, which leads to better fusion of information and effective modeling of long-term patterns in the series.
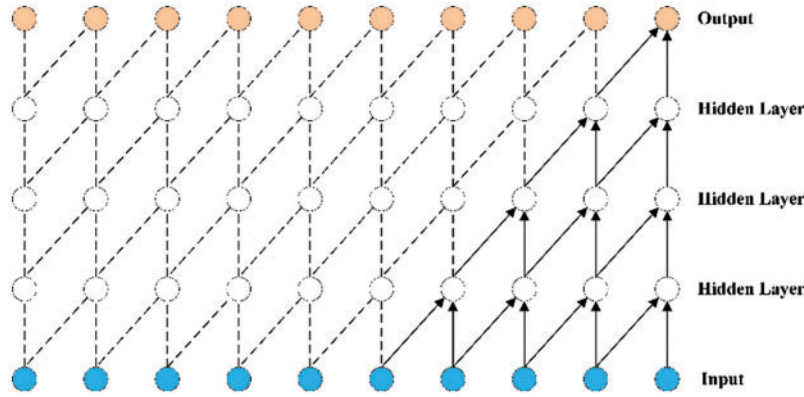
Figure 6: Causal convolution
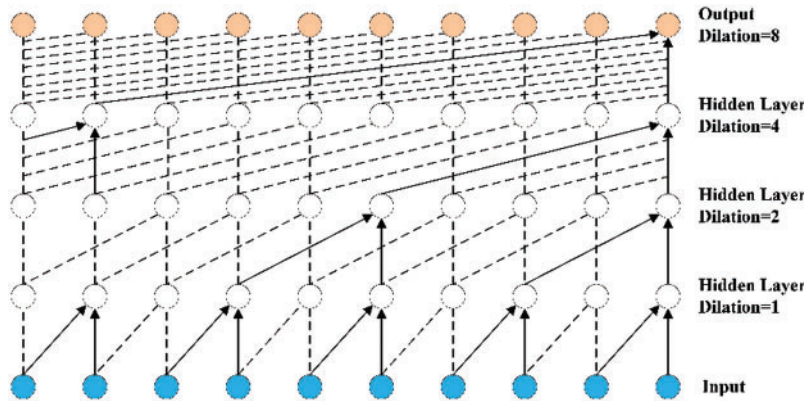
Figure 7: Dilated causal convolution

In order to fully extract the temporal features of the posture data, this paper introduces the idea of Inception to construct the Temporal Block as shown in Fig. 8, which consists of 4 parts, each of which has the same size of the inflated causal convolution kernel and different inflation factors.

Given the processed input data $V = \{V_1, V_2, \ldots, V_m\}$, activated by the sigmoid function and tanh function respectively after Temporal Block, the output features S_a and S_b are obtained.

$$S_a = \tanh \left(F \left(V, s \right) \right) \tag{5}$$

$$S_b = \sigma \left(F \left(V, s \right) \right) \tag{6}$$

$$S = S_a \times S_b \tag{7}$$

3.1.2 Graph Attention Network

Graph Attention Network is an attention-based node classification network proposed by Velickovic et al. [19], whose basic idea is to update the node representation based on each node's

attention to its neighboring nodes. At the input layer, GAT receives the node features of the graph data and the topology of the graph. For the KGTLIR model constructed in this paper, the node features of the graph are the feature vectors extracted by fusing Temporal Block and Conv1d, and the topology is the adjacency matrix constructed by the similarity between the feature vectors.
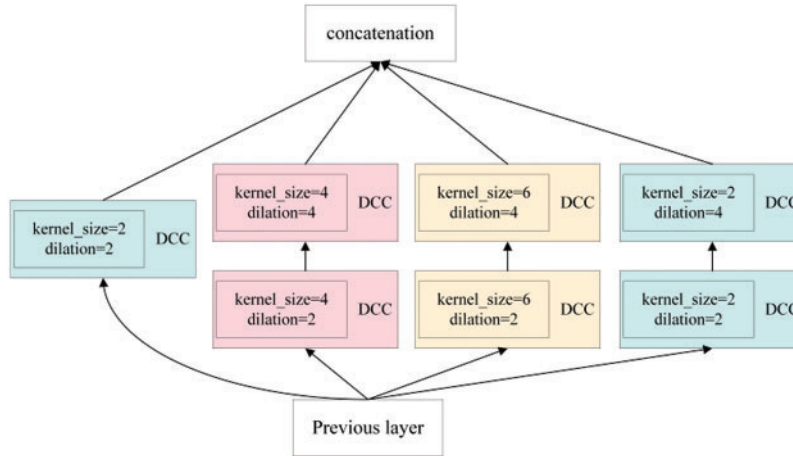
Figure 8: Specific framework of temporal block

The inputs are the node features $H = \left\{ \vec{h}_1, \vec{h}_2, \ldots, \vec{h}_N \right\}, \vec{h}_i \in R^F$, where N represents the number of nodes and F represents the feature dimension of the nodes. In order to obtain sufficient expressive power, GAT uses a shared linear transform implementation parameterized by a weight matrix to transform the input features into higher-level features. At the same time, a shared attention mechanism a is introduced in the paper to calculate the importance e_{ij} of node j's features with respect to node i, as shown in Eq. (8).

$$e_{ij} = a\left(W\vec{h}_i, W\vec{h}_j \right) \tag{8}$$

where the shared attention mechanism $a: R^{F'} \times R^F \to R$ is a single-layer feed-forward neural network and $W \in R^{F' \times F}$ denotes the weights applied to each node of the graph.

To make the coefficients between different nodes easy to compare, the paper is normalized using the SoftMax function to obtain the normalized attention coefficients.

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(\vec{a}^T \left[W\vec{h}_i || W\vec{h}_j \right] \right) \right)}{\sum_{k \in D_i} \exp\left(\text{LeakyReLU}\left(\vec{a}^T \left[W\vec{h}_i || W\vec{h}_k \right] \right) \right)} \tag{9}$$

where $\vec{a} \in R^{2F'}$ is the initialized weight vector of the attention mechanism a, LeakyReLU is used to perform nonlinear operations, D_i is some domain of node i in the graph, and || represents the splicing operation.

Finally, normalized attention coefficients are used to compute linear combinations of the corresponding features and the output features of the nodes are obtained by means of a nonlinear activation

function σ.

$$\vec{h}'_i = \sigma \left(\sum_{j \in N_i} \alpha_{ij} \mathbf{W} \vec{h}_j \right) \tag{10}$$

To stabilize the learning process of the self-attention mechanism, the paper extends the attention mechanism to a multi-attention mechanism, where K attention mechanisms run independently, and the obtained output features are spliced to obtain the final feature.

$$\vec{h}'_i = \|_{k=1}^{K} \sigma \left(\sum_{j \in N_i} \alpha_{ij}^k \mathbf{W}^k \vec{h}_j \right) \tag{11}$$

3.1.3 Introduction of Sample Weights

The GTLIR model uses a cross-entropy loss function to compute the loss values during the training process. For the multi-classification task, the model's loss function over the entire training set is as follows:

$$Loss\,(p, q) = -\frac{1}{N} \sum_{n=1}^{N} \sum_{m=1}^{M} o\left(x_{n,m}\right) \ln q\left(x_{n,m}\right) \tag{12}$$

where N denotes the total number of training samples in the training set, M denotes the total number of labeled categories in the training set, $o(x_{n,m})$ denotes the probability of the n-th sample in the training set for the true labeled category m, and $q(x_{n,m})$ denotes the probability of the output value of the n-th sample of the GTLIR model for the m-th category.

In the actual training process, the cross-entropy loss function is mainly used to quantify the difference between the GTLIR model's predicted result q and the true intention p. The closer the two are, the smaller the value of the loss, and vice versa. This affects the model's recognition results for target intention to some extent, but it does not improve the misclassification problem well, and the GTLIR model can only achieve 96.45% recognition accuracy for the test set.

Therefore, to address the problem of the cost associated with intention recognition misclassification mentioned in the introduction, the paper introduces sample weights to give samples different misclassification costs. The formula for calculating the sample weights is as follows:

$$G_i = \frac{N_i}{\sum_{\substack{j=0 \\ j \neq i}}^{M} N_j} \tag{13}$$

where G_i denotes the weight of the i-th sample; N_i denotes the number of the i-th sample, and M denotes the total number of label types.

The sample weights are introduced into the cross-entropy loss function as misclassification cost indicators, and the improved cross-entropy loss function is obtained as follows:

$$Loss\,(p, q) = -\frac{1}{N} \sum_{n=1}^{N} \sum_{m=1}^{M} G_m o\left(x_{n,m}\right) \ln q\left(x_{n,m}\right) \tag{14}$$

During each round of iteration, the weights G_i are further updated based on the accuracy, recall, and F1-score after iteration, and the updated G_i and loss function are calculated as follows:

$$G_i' = G_i \cdot \exp\left(\frac{\text{Accuracy}}{M}\right) \cdot \exp\left(\frac{\text{Recall}}{M}\right) \cdot \exp\left(\frac{\text{F1}}{M}\right) \tag{15}$$

$$Loss\,(p, q) = -\frac{1}{N} \sum_{n=1}^{N} \sum_{m=1}^{M} G_m{}' p\,(x_{n,m}) \ln q\,(x_{n,m}) \tag{16}$$

where Accuracy denotes the accuracy after iteration, Recall denotes the recall after iteration, and F1 denotes the F1-score after iteration.
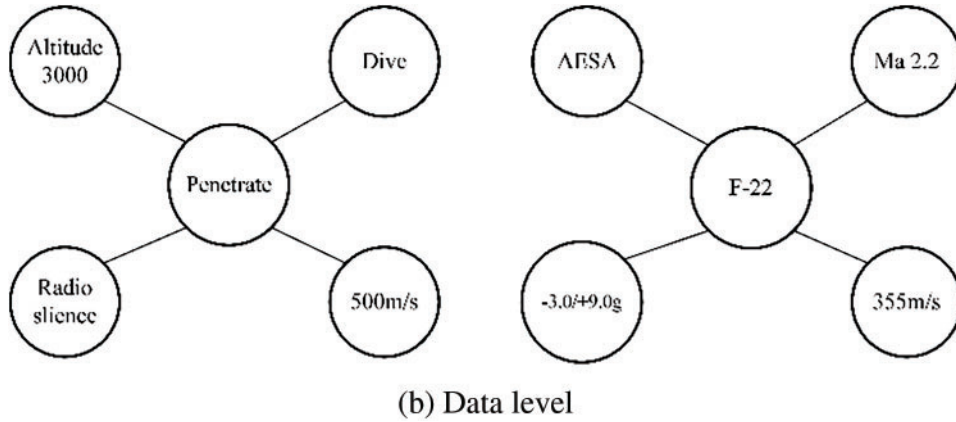
3.2 Intention Recognition Model Based on Knowledge Graph

After the discovery of enemy air targets, the enemy target information obtained through multiple ways is fused through the Knowledge Graph to recognize their operational intention. In this paper, we refer to Literature [20] to establish an air target intention knowledge graph model, which is mainly composed of two parts, the schema level, and the data level, as shown in Fig. 9. The schema level refers to the description of schema information such as relevant formulas, rules, and association relations through formalization, which mainly consists of the motion state, entity attribute, and sensor state of the air target, as shown in Fig. 9a. The data level contains the association relationships between instances and is a concrete representation of the schema level, as shown in Fig. 9b. The data level is the foundation, which is mainly used to store the feature data extracted based on the ACMG system. The schema level is built on top of the data level, which firstly inputs the target entity attributes, sensor states, motion states, and other information from the data level, after which the probabilities of different intentions of the target are obtained through the relational constraints and the intention analysis, and finally the probability with the largest probability is selected to get the intention as the current state of the target. For the whole KGTLIR model, the original data is stored in the data level, the probability distribution p_2 of different intentions in the current state is obtained through the schema level, and the probability distribution of the final target's intention is obtained by fusing p_2 with p_1 using evidence theory, and the one with the largest probability is selected as the intention.
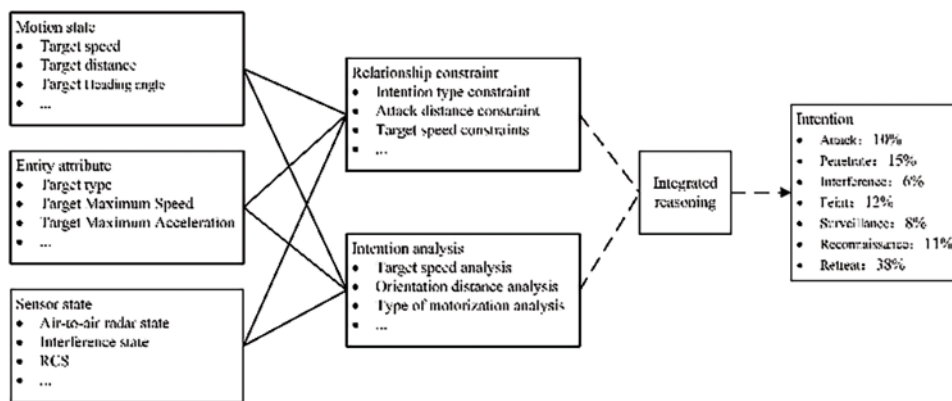
(a) Schema level

Figure 9: (Continued)

(b) Data level

Figure 9: Intention recognition model based on Knowledge Graph

For the air target intention recognition task, each target has other entities in the knowledge graph that are connected to it through direct or indirect relationship paths. These relationships include affiliation, inclusion, synergy, etc. These entities include the target's physical attributes, motion state, and sensor state, which can reflect the target's tactical intention tendencies to some extent.

To fully utilize this knowledge information, the collection of entities and relationship paths is defined in the paper to represent the propagation information of the target entity, which is formulated as follows:

$$S_i^{(l)} = \left\{ \left(p^l = p^{l-1} || < r, > t \right) \mid \left(p^{l-1}, h \right) \in S_i^{(l-1)} \wedge (h, r, t) \in S \right\} \tag{17}$$

where $||$ denotes the new path formed by placing the relation r on the path p^{l-1}, p^l denotes the relational path that propagates from target entity i to the j-th hop. $S_i^{(0)} = \{(<>, i)\}$ is the initial set of entity and path relationships, $S_i^{(1)} = \{(< r_2, e_3 >, i), (< r_3, e_4 >, i)\}$, and so on.

The attentional weights between goals and relationship paths are calculated by Eq. (18). The attentional weights between goals and entity paths are calculated by Eq. (19). Normalization is done to get the final weights $\tilde{\pi}_{(p^l,t)}^u$.

$$\alpha_{p^l}^u = \sigma \left(\delta \left(W_1 u + b_1 \right)^T p^l \right) \tag{18}$$

$$\beta_{(p^l,t)}^u = \sigma \left(\delta u^T diag \left(p^l \right) e_l \right) \tag{19}$$

$$\tilde{\pi}_{(p^l,t)}^u = \text{softmax} \left(\alpha_{p^l}^u \cdot \beta_{(p^l,t)}^u \right) = \frac{\exp \left(\alpha_{p^l}^u \cdot \beta_{(p^l,t)}^u \right)}{\sum_{(p^l,t) \in S_i^l} \exp \left(\alpha_{p^l}^u \cdot \beta_{(p^l,t)}^u \right)} \tag{20}$$

The normalized weights are linearly combined, which in turn yields the set Γ^i of all L-hop entity embedding information.

$$e_{S_i^l}^u = \sum_{(p^l,t) \in S_i^l} \tilde{\pi}_{(p^l,t)}^u e_t \tag{21}$$

$$\Gamma^i = \left\{ e_{S_i^0}^u, e_{S_i^1}^u, \ldots, e_{S_i^L}^u \right\} \tag{22}$$

The new embedding e_i^* can be obtained by performing a summation operation on $e_{S_i^l}^u$ in the set Γ^i.

$$e_i^* = agg_i^u = \sum_{e_{S_i^l}^u \in \Gamma_i} e_{S_i^l}^u \tag{23}$$

The embedding u^* of target u can be obtained by combining the collaborative information between targets with the SVD++ method.

$$u^* = u + \frac{1}{|T(u)|} \sum_{e_{i_k} \in T(u)} e_{i_k} \tag{24}$$

where $T(u)$ is denoted as the set of intentions that have interacted with target u.

Finally, performing inner product operation on u^* and e_i^* can obtain the probability score of the predicted target's intention.

$$\hat{y}_{ui} = u^{*T} e_i^* \tag{25}$$

4 Experimental Analyses

4.1 Experimental Data and Experimental Environment

The experimental data were obtained from the Air Combat Maneuvering Generator (ACMG). During the data acquisition process, multiple air combat scenarios were set up, such as air defense operations in different terrains such as plains and mountains. The temporal characterization data of the target and the initially set intention data are output through the system interface, after which the target intention is revised by experts in the field of air combat. The data under different scenarios are fused to obtain a total of 10,560 samples, including 8448 training samples and 2112 test samples, with a time cloth length of 12 sampling cycles. Considering that the enemy air targets perform different combat missions and our information technology, the paper selects {altitude, speed, acceleration, heading angle, azimuth, distance, radar one-dimensional distance image, radar reflective cross-sectional area, air-to-air radar status, ground-to-ground radar status, jamming status, and enemy identification response} as the 12-dimensional features. The first eight numerical features are normalized, and the last four non-numerical features are encoded to obtain the final air target intention dataset.

The experiments performed in the paper are in Python language, version 3.8, accelerated by NVIDIA GeForce RTX2080 GPUs and CUDA 12.2, and using pytorch deep learning framework.

4.2 Evaluation Indicators

To validate the performance of the proposed air target operational intention recognition method in this paper, six metrics, namely Accuracy, Precision, Recall, F1-score, Loss, and Total_error, were used to evaluate the classification of the network. They are calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{26}$$

$$Precision = \frac{TP}{TP + FP} \tag{27}$$

$$Recall = \frac{TP}{TP + TN} \tag{28}$$

$$F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{29}$$

$$\text{Total\_error} = FN + FP \tag{30}$$

where TP is the true class, FN is the false negative class, FP is the false positive class, and TN is the true negative class.

4.3 Model Parameter Setting

The hyperparameter settings of the GTLIR model have a significant impact on its classification performance. In this paper, multiple sets of experiments are evaluated to select the hyperparameters with optimal model performance. The hyperparameters of the GTLIR model mainly contain the optimizer, batch size, epoch, learning rate, convolution kernel size, and dilated factor for dilated causal convolution in Temporal Block, and the number of nodes for GTA.

For optimizers, this paper compares five candidate algorithms, Stochastic Gradient Descent (SGD) [21], Root Mean Square prop (RMSprop) [22], Adaptive Moment Estimation (Adam) [23], Nadam, and Adamax. The resulting test set accuracy and loss values are shown in Table 3. In the paper, Adam, which has the highest accuracy, is chosen as the optimizer for the GTLIR model.

Table 3: Accuracy of test set obtained by different optimizers

Optimizer	Accuracy (%)	Loss	Total_error
Adam	97.82	0.101	46
SGD	80.92	0.714	403
RMSprop	96.54	0.139	73
NAdam	97.02	0.115	63
Adamax	97.16	0.111	60

In this paper, we compare the accuracy of the model on the test set for different batch sizes and learning rates, and the experimental results are shown in Table 4. The model achieved the highest accuracy when the batch size was set to 64 and the learning rate was set to 0.001. The model hyperparameter settings for GTLIR are shown in Table 5.

Table 4: Accuracy for different parameters

	lr = 0.1	lr = 0.01	lr = 0.001	lr = 0.0001
Batch_size = 32	90.00%	96.59%	97.10%	94.74%
Batch_size = 64	94.24%	96.63%	97.82%	92.44%
Batch_size = 128	86.35%	96.31%	96.82%	86.12%
Batch_size = 256	91.38%	97.05%	97.23%	85.71%

Table 5: GTLIR model hyperparameter setting

Hyperparameter	Value
Optimizer	Adam
Batch size	64
Epoch	500
Learning rate	0.001
Convolution kernel size in temporal block	2,4,6
Dilated factor in temporal block	2,4
The number of nodes for GAT	12

4.4 Analysis of Experimental Results of Different Loss Function

The effect of Eqs. (12) and (14) in the GTLIR model is compared experimentally, and the results are shown in Fig. 10.
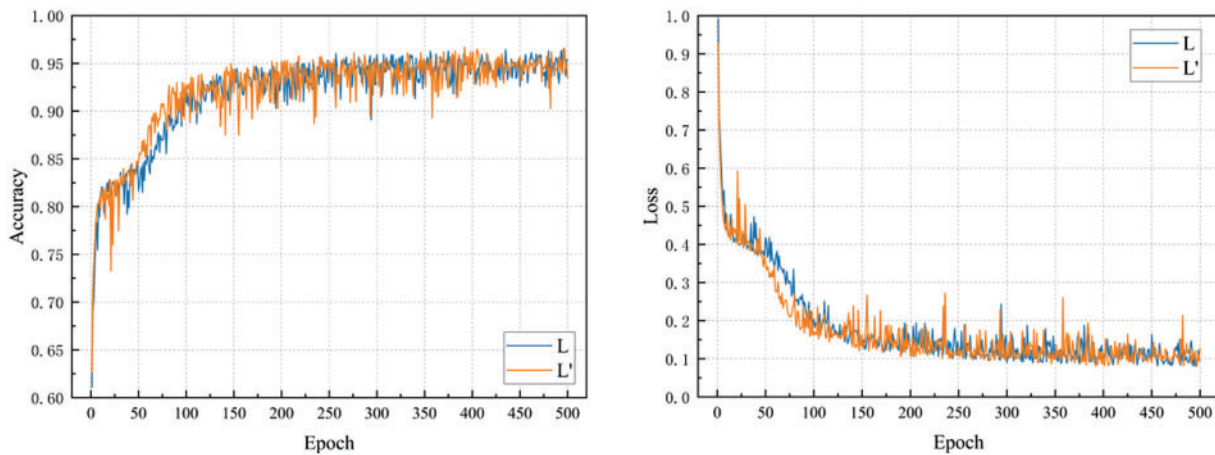
Figure 10: Comparison of loss function effects

In Fig. 10, L denotes the experimental accuracy and loss value variation curves for the loss function in Eq. (12), and the GTLIR model has an accuracy of 96.45% on the test set. L′ denotes the experimental accuracy and loss value change curve of the loss function in Eq. (14), and the GTLIR model has an accuracy of 96.71% on the test set. It can be found that after the introduction of the sample weights, the accuracy of the model is then improved, but the improvement is not obvious, only 0.26%.

The effect of Eqs. (12) and (16) in the GTLIR model is compared experimentally, and the results are shown in Fig. 11.

In Fig. 11, L denotes the experimental accuracy and loss value variation curves for the loss function in Eq. (12), and the GTLIR model has an accuracy of 96.45% on the test set. L′denotes the experimental accuracy and loss value change curve of the loss function in Eq. (16), and the GTLIR model has an accuracy of 97.82% on the test set. It can be found that the introduction of dynamically updated sample weights improves the accuracy of the model by 1.37%.

Through the above two comparison experiments, we can clearly find that the improved loss function in this paper has a beneficial effect on reducing the probability of misclassification.
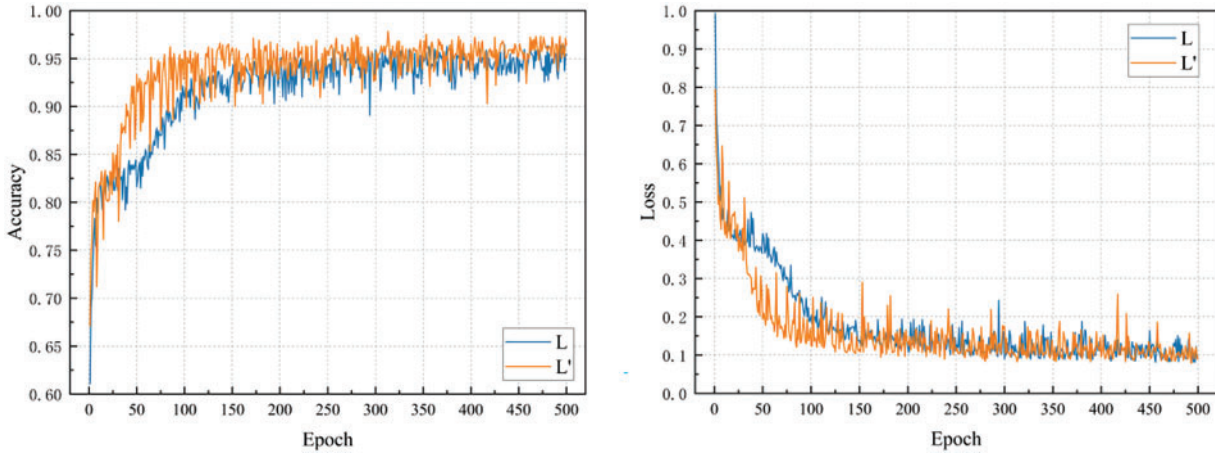

Figure 11: Comparison of loss function effects

4.5 Analysis of Experimental Results of the GTLIR Model

4.5.1 Analysis of Intention Recognition Results

The GTLIR model achieved an intention recognition accuracy of 97.82%, a precision of 97.85%, a recall of 97.92%, an F1-score of 97.87%, a loss value of 0.101, and a total number of misclassified samples of 46 on the test set of 2112 samples. The number of samples for each intention of the target in the test set varies is different. In the paper, the confusion matrix for the test set is generated, as shown in Fig. 12, as well as the precision, recall, and F1-score for different intentions, as shown in Table 6.
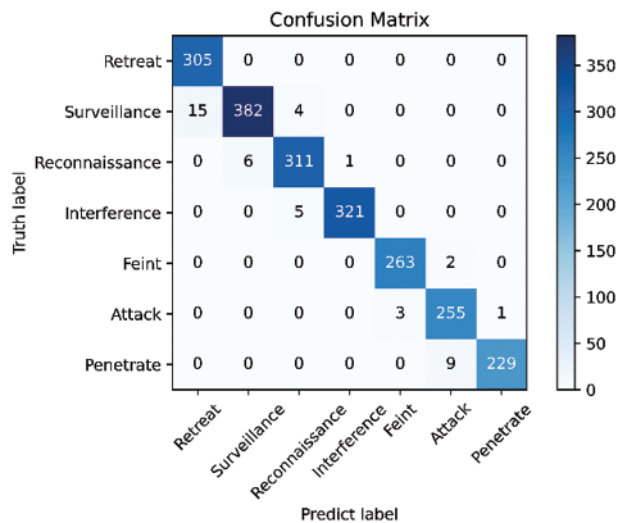

Figure 12: Confusion matrix for the GTLIR model

Table 6: Detection results of different intentions of the GTLIR model

Intention	Precision (%)	Recall (%)	F1-score (%)
Retreat	95	100	98
Surveillance	98	95	97
Reconnaissance	97	98	97
Interference	100	98	99
Feint	99	99	99
Attack	96	98	97
Penetrate	100	96	98

From Fig. 12 and Table 6, it can be found that the GTLIR model has high recognition accuracy for all seven intentions in the test set, especially for the retreat intention, which can reach 100%. The remaining six intents were recognized with over 95% accuracy. The reason for this analysis is that when the intention is to retreat, the target's maneuvering characteristics are relatively special, and the distance between the target and the center of our air defense strongholds is increasing, thus making it easy to identify. There are some cases of misclassification between attack intention and feint intention. The target performing a feint mission, in order to achieve the purpose of confusing the opponent, has state characteristics that are very close to those of the target performing an attack mission. Thus, there is a certain possibility of miscalculation between the two. Targets performing surveillance intention and reconnaissance intention have certain similarities in characteristics such as altitude, speed, acceleration, heading angle, azimuth, distance, etc., and thus are also susceptible to misjudgment.

4.5.2 Ablation Experiment

In order to further validate the effectiveness of the GTLIR model on the task of air target intention recognition, ablation experiments were conducted in the paper on the same test set. The experimental results are shown in Table 7.

**Table 7:** Comparative results of ablation experiments

Model composition structure			Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)	Total_error
Temporal block	GAT	Wi					
√			95.55	95.68	95.78	95.68	94
√		√	96.35	96.58	96.63	96.57	77
√	√		96.45	96.35	96.49	96.39	75
√	√	√	97.82	97.85	97.92	97.87	46

The experiments show that all evaluation indexes of the GTLIR model are better than the other three models. The introduction of sample weights improves the accuracy of Temporal Block by 0.8% and TemporalBlock_GAT by 1.37%. The introduction of GAT improves the accuracy of the Temporal Block by 0.9%.

Considering the different sample sizes of each intention in the test set, this paper uses precision, recall, and F1-score to reflect the recognition results of these four models, as shown in Table 8, where ① denotes the Temporal Block model, ② denotes the TemporalBlock_Wi model, ③ denotes the TemporalBlock_GAT model, and ④ denotes the GTLIR model. The confusion matrix for whether the GTLIR model introduces sample weights is shown in Fig. 13.

Table 8: Comparison of different intention recognition evaluation metrics

Intention	Precision (%)				Recall (%)				F1-score (%)			
	①	②	③	④	①	②	③	④	①	②	③	④
Retreat	91	92	98	95	98	99	100	100	94	96	99	98
Surveillance	94	95	98	98	93	93	96	95	93	94	97	97
Reconnaissance	98	97	96	97	94	94	96	98	96	95	96	97
Interference	99	98	97	100	96	96	96	98	98	97	97	99
Feint	93	98	96	99	98	97	91	99	96	98	94	99
Attack	97	97	91	96	93	98	98	98	95	98	95	97
Penetrate	97	99	98	100	98	98	99	96	98	99	98	98

(a) No sample weights introduced   (b) Introducing sample weights
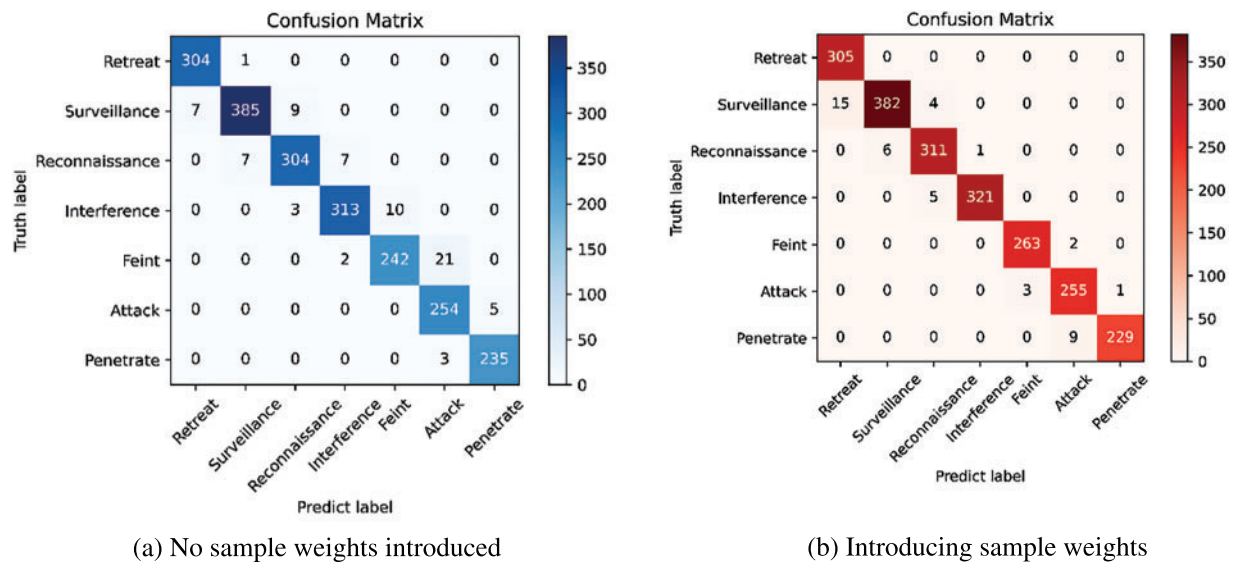
Figure 13: The confusion matrix for whether the GTLIR model introduces sample weights

As can be seen from Tables 7 and 8, the GTLIR model has a high recognition accuracy for each class of intention, and the misclassification of intention on the test set improves greatly after the introduction of sample weights, with the lowest total number of misclassified samples at only 46. From Fig. 13, after introducing sample weights, the model has significantly improved the misclassification of feint intention as attack intention and reconnaissance intention as surveillance intention. It is further demonstrated that with the introduction of graph attention mechanism and sample weights, the recognition accuracy of the GTLIR model is greatly improved and the effect of intention recognition is better.

4.5.3 Comparative Analysis Experiment

To verify the superiority of the model, the paper first compares the GTLIR model with TCN, GRU, LSTM, and their combinations, which are three neural networks commonly used for feature extraction of time-series data. The intention recognition accuracy, total number of misclassified samples, and runtime of the models are compared under the same test set, as shown in Table 9.

Table 9: Comparative experiments of different models

Model	Accuracy (%)	Total_error	Time (ms)
GTLIR	97.82	46	24.15
TCN	90.91	192	29.31
GRU	85.42	308	20.36
LSTM	84.99	317	21.54
TCN-GRU	94.93	107	30.53
TCN-BiGRU	95.27	100	43.51
TCN-LSTM	94.98	106	37.41
TCN-BiLSTM	95.31	99	48.85

From Table 9, it can be seen that GTLIR achieves an intention recognition accuracy of up to 97.82% on the test set with the least number of total misclassified samples. The comparison experiments between LSTM and GRU, TCN-LSTM and TCN-GRU, TCN-BiLSTM and TCN-BiGRU revealed that the recognition accuracy of the three groups of models did not differ much, but the time to recognize the samples was significantly longer for LSTM compared to GRU, TCN-LSTM compared to TCN-GRU, and TCN-BiLSTM compared to TCN-BiGRU. This is mainly due to the different structure and number of parameters of LSTM and GRU. The GTILR model improves the accuracy by at least 2.89%, the total number of misclassified samples is reduced by at least 53, and the sample identification time is reduced to some extent compared to the four models, namely, TCN-LSTM, TCN-GRU, TCN-BiLSTM, and TCN-BiGRU. The analysis found that the GTLIR model introduces dilated causal convolution and GAT, which can effectively extract the temporal and spatial features of the intent data. At the same time, due to the introduction of the GAT, the representation of the model is improved, and the output features can be computed in parallel on all nodes, improving the sample identification time of the model.

In addition, the paper compares the GTLIR model with seven existing air target intention recognition models, namely LSTM-Attention [2], (Deep BP Neural Networks) DBP [10], Attention-TCN-BiGRU [11], STABC-IR [13], BiGRU-Attention [24], PCLSTM [25], SVM. The intention recognition accuracy and loss values of the models are compared under the same test set as shown in Table 10. It can be found that the GTLIR model has the highest accuracy, which fully proves the high efficiency and importance of the model in recognizing the intention of air targets.

Table 10: Comparison experiment between GTLIR model and existing model

Model	Accuracy (%)	Loss
GTLIR	97.82	0.101
STABC-IR	95.76	0.146
Attention-TCN-BiGRU	94.63	0.138
LSTM-attention	93.79	0.173
BiGRU-attention	93.54	0.180
PCLSTM	88.73	0.338
SVM	84.65	0.385
DBP	81.56	0.414

4.6 Analysis of Results by Fusing the Knowledge Graph Model and GTLIR Model

To further improve the accuracy of intention recognition, this paper comprehensively analyzes the results based on the Knowledge Graph model and the results based on the GTLIR model to get the final target intention. Given a set of data, feeding it into the Knowledge Graph model will result in a probability distribution of different intentions p_1, and feeding it into the GTLIR model will result in a probability distribution of different intentions p_2. Afterward, the two are fused through evidence theory to obtain a new probability distribution p. The type of intention with the largest probability value is selected as the final intention of the target.

In order to illustrate the intention recognition process of fusing the Knowledge Graph model and GTLIR model, this paper is presented in two cases.

Case 1, the predicted results of the target under both recognition models are consistent.

The predicted probability distribution p_1 of the target under the GTLIR model is as follows:

$$p_1 = \{0.0001,\ 0.02,\ 0.9795,\ 0.0001,\ 0.0001,\ 0.0001,\ 0.0001\} \tag{31}$$

The predicted probability distribution p_2 of the target under the Knowledge Graph model is as follows:

$$p_2 = \{0.08,\ 0.10,\ 0.61,\ 0.06,\ 0.03,\ 0.05,\ 0.07\} \tag{32}$$

The value of K is calculated according to the orthogonal sum formula as follows:

$$K = \sum_{x \cap y \neq \phi} p_1(x) \times p_1(y) = 0.5995 \tag{33}$$

The fused probability distribution of intentions P can be obtained from Eq. (34).

$$P(A) = \frac{1}{K} \times \sum_{x \cap y = A} p_1(x) \times p_1(y) \tag{34}$$

$$P = \{0.00001,\ 0.003301,\ 0.996656,\ 0.00001,\ 0.000005,\ 0.000008,\ 0.00001\} \tag{35}$$

The intention with the largest probability value is chosen as the target's final intention. So, the target's intention was Reconnaissance.

Case 2, the predicted results of the target under both recognition models are inconsistent.

The predicted probability distribution p_1 of the target under the GTLIR model is as follows:

$$p_1 = \{0.0001,\ 0.0001,\ 0.0001,\ 0.0001,\ 0.6532,\ 0.3463,\ 0.0001\} \tag{36}$$

The predicted probability distribution p_2 of the target under the Knowledge Graph model is as follows:

$$p_2 = \{0.01,\ 0.01,\ 0.01,\ 0.01,\ 0.13,\ 0.48,\ 0.35\} \tag{37}$$

The value of K is calculated as follows:

$$K = 0.2512 \tag{38}$$

The fused probability distribution of intentions P can be obtained from Eq. (34).

$$P = \{0.000004,\ 0.000004,\ 0.000004,\ 0.000004,\ 0.338041,\ 0.661720,\ 0.000223\} \tag{39}$$

The intention with the largest probability value is chosen as the target's final intention. So, the target's intention was Attack.

For the test set of 2112 samples, the recognition accuracy using the above method can reach 98.48%, which significantly improves the recognition results of the model, reduces the probability of misclassification, and increases the interpretability and trustworthiness of the proposed air target intention recognition model.

5 Conclusions

Air target intention recognition plays an important role in modern air operations as a core part of situational awareness. Aiming at the problems of insufficient feature extraction and misclassification in intention recognition, this paper designs an air target intention recognition method based on dilated causal convolution and graph attention mechanism. Firstly, relevant intention features are selected and coded by analyzing the characteristics of the air target intention recognition task. Secondly, the Temporal Block module composed of dilated causal convolution is used to mine the temporal relationship of intention features, and the graph attention mechanism is used to mine the spatial relationship of intention features. Thirdly, the iterated evaluation metrics are introduced to design the sample weights to reduce the possibility of misclassification. Fourthly, the Intention recognition model based on Knowledge Graph is constructed to predict the probability of different intentions. Fifthly, the final intent determination result is given by comprehensively considering the intent recognition probability distribution of the two models. Finally, the performance of the proposed model in this paper is verified through simulation experiments, ablation experiments, and comparative analysis experiments, which show that the KGTLIR model has a high accuracy of intention recognition. Besides, the interpretability and trustworthiness of the model are both improved to some extent.

During the study, it was found that there are still many problems with the air target intention recognition task, and the following areas will be focused on in the next study.

Firstly, the costs of war caused by miscalculation of intention in actual combat are beyond the means of either side. Although the existing intent recognition models have achieved high recognition accuracy, the performance and interpretability of the models still need to be improved.

Secondly, the existing intentions are labeled by experts based on empirical knowledge. In actual combat, there is a large similarity between the intentions, and it is difficult to make a strict distinction between different intentions.

Thirdly, the battlefield data collected in the actual combat is less, and the existing model is more dependent on the support of data, how to obtain more high-quality intent data is also a future research direction.

Fourthly, analyzing from the perspective of the OODA ring, if we can predict the enemy target's combat intention in advance, we can lay the foundation for accelerating the closure of the OODA ring, and then we can obtain the prerequisites for the victory of the war.

Author Contributions: Study conception and design: Qinghua Xing; data collection: Huaixi Xing, Zhanfu Song; analysis and interpretation of results: Bo Cao, Longyue Li; draft manuscript preparation: Bo Cao. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: The datasets used or analyzed during the current study are available from the corresponding author Bo Cao on reasonable request.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] A. Fusano, H. Sato, and A. Namatame, "Multi-agent based combat simulation from OODA and network perspective," in *2011 UkSim 13th Int. Conf. Comput. Model. Simul.*, Cambridge, UK, Mar. 2011, pp. 249–254.

[2] P. Xia, M. Chen, J. Zou, and X. Feng, "Prediction of air target intention utilizing incomplete information," in *Proc. 2016 Chin. Intell. Syst. Conf.*, Singapore, Sep. 2016, pp. 395–403.

[3] Y. L. Sun and L. Bao, "Research on tactical intent recognition technology for targets in the maritime battlefield based on D-S evidence theory," (in Chinese), *Ship Electron. Eng.*, vol. 32, no. 5, pp. 48–51, 2012.

[4] Z. F. Jun, Z. Z. Jie, H. C. Hua, W. Li, and L. T. Yuan, "Aerial target intention recognition approach based on belief-rule-base and evidential reasoning," *Electron. Optics Control*, vol. 24, no. 8, pp. 15, 2017. doi: 10.3969/j.issn.1671-637x.2017.08.004.

[5] Y. Xu, S. Cheng, H. Zhang, and Z. Chen, "Air target combat intention identification based on IE-DSBN," in *2020 Int. Workshop Electron. Commun. Artif. Intell. (IWECAI)*, Jun. 2020, pp. 36–40.

[6] X. X. Meng, Y. R. Nong, and F. Ying, "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. doi: 10.21629/JSEE.2018.04.11.

[7] J. Qing, G. X. Tai, J. W. Dong, and W. N. Fang, "Intention recognition of aerial targets based on Bayesian optimization algorithm," in *2017 2nd IEEE Int. Conf. Intell. Transport. Eng. (ICITE)*, Singapore, Sep. 2017, pp. 356–359.

[8] Y. M. Chen and C. Y. Li, "Target tactical intention recognition simulation based on knowledge graph," (in Chinese), *Comput. Simul.*, vol. 36, pp. 1–4+19, 2019.

[9] X. Yin, M. Zhang, and M. Q. Chen, "Air target combat intent recognition based on discriminant analysis," (in Chinese), *J. Missile Guidance*, vol. 38, pp. 46–50, 2018.

[10] W. W. Zhou, P. Y. Yao, J. Y. Zhang, X. Wang, and S. Wei, "Combat intention recognition for aerial targets based on deep neural network," (in Chinese), *J. Aeronaut.*, vol. 39, pp. 200–208, 2018.

[11] F. Teng, Y. Song, and X. Guo, "Attention-TCN-BiGRU: An air target combat intention recognition model," *Mathematics*, vol. 9, no. 19, pp. 2412, Sep. 2021. doi: 10.3390/math9192412.

[12] X. Wang, Z. Yang, G. Zhan, J. Huang, S. Chai and D. Zhou, "Tactical intention recognition method of air combat target based on BiLSTM network," in *2022 IEEE Int. Conf. Unmanned Syst. (ICUS)*, Guangzhou, China, Oct. 2022, pp. 63–67.

[13] S. Wang, G. Wang, Q. Fu, Y. Song, J. Liu and S. He "STABC-IR: An air target intention recognition method based on bidirectional gated recurrent unit and conditional random field with space-time attention mechanism," *Chin. J. Aeronaut.*, vol. 36, no. 3, pp. 316–334, Mar. 2023. doi: 10.1016/j.cja.2022.11.018.

[14] C. Qu, Z. Guo, S. Xia, and L. Zhu, "Intention recognition of aerial target based on deep learning," *Evol. Intell.*, vol. 17, pp. 303–311, May 2022. doi: 10.1007/s12065-022-00728-9.

[15] C. H. Zhang, Y. Zhou, Y. C. Cai, and J. Q. Guo, "A review of research on air target combat intent recognition," (in Chinese), *Modern Defense Technol.*, vol. 51, no. 4, pp. 1–16, 2023.

[16] M. G. Lei, Z. R. Nan, W. Biao, Z. M. Zhe, W. Yu and L. Xiao, "Target tactical intention recognition in multiaircraft cooperative air combat," *Int. J. Aerosp. Eng.*, vol. 2021, pp. 1–18, Nov. 2021. doi: 10.1155/2021/9558838.

[17] S. Bai, J. Z. Kolter, and V. Koltun, "An empirical evaluation of generic convolutional and recurrent networks for sequence modeling," Apr. 19, 2018. doi: 10.48550/arXiv.1803.01271.

[18] A. Ayodeji *et al*., "Causal augmented ConvNet: A temporal memory dilated convolution model for long-sequence time series prediction," *ISA Trans.*, vol. 123, pp. 200–217, Apr. 2022. doi: 10.1016/j.isatra.2021.05.026.

[19] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò and Y. Bengio, "Graph attention networks," Feb. 4, 2018. doi: 10.48550/arXiv.1710.10903.

[20] Y. Zuo, G. L. Zhang, W. Wu, J. Wang, X. D. Li and H. J. Wang, "Application of knowledge graph for battlefield sea and air target recognition," (in Chinese), *Command Inf. Syst. Technol.*, vol. 10, pp. 1–5+24, 2019.

[21] R. Archibald, "A stochastic gradient descent approach for stochastic optimal control," *EAJAM*, vol. 10, no. 4, pp. 635–658, Jun. 2020. doi: 10.4208/eajam.190420.200420.

[22] D. P. Xu, S. D. Zhang, H. S. Zhang, and D. P. Mandic, "Convergence of the RMSProp deep learning method with penalty for nonconvex optimization," *Neural Netw.*, vol. 139, pp. 17–23, 2021. doi: 10.1016/j.neunet.2021.02.011.

[23] W. Zhang, L. Niu, D. Zhang, G. Wang, F. U. D. Farrukh and C. Zhang, "HW-ADAM: FPGA-based accelerator for adaptive moment estimation," *Electronics*, vol. 12, no. 2, pp. 263, Jan. 2023. doi: 10.3390/electronics12020263.

[24] F. Teng, X. Guo, Y. Song, and G. Wang, "An air target tactical intention recognition model based on bidirectional GRU with attention mechanism," *IEEE Access*, vol. 9, pp. 169122–169134, 2021. doi: 10.1109/ACCESS.2021.3135495.

[25] J. Xue, J. Zhu, J. Xiao, S. Tong, and L. Huang, "Panoramic convolutional long short-term memory networks for combat intension recognition of aerial targets," *IEEE Access*, vol. 8, pp. 183312–183323, Jan. 2020. doi: 10.1109/ACCESS.2020.3025926.