

# 面向空中目标意图识别的时空 Transformer 模型设计

王科, 李成海, 宋亚飞, 王鹏, 李乐民

(空军工程大学 防空反导学院, 陕西 西安 710051)

**摘要:**信息化条件下的战争环境瞬息万变,战场态势数据呈现海量、多样化的特点,导致利用专家经验识别空中目标作战意图的难度越来越高。结合目前先进的智能化方法,对 Transformer 模型进行研究并将其引入空中目标意图识别领域,设计出一种新的意图识别方法 Temporal-Spatial Transformer (TST),可以有效地挖掘战场态势数据中时间域和空间域的深层特征信息,提高空中目标作战意图识别准确率。同时,对4种目前较为先进的神经网络意图识别方法进行效果对比,结果表明,TST模型在各类指标上都取得了突出的效果,优于所有对比的神经网络模型。TST模型不仅有优异的准确率,而且收敛速度极快,可以迅速地抓取战场态势数据中的关键信息进行意图识别。

**关键词:**Transformer;时空融合;空中目标;意图识别;自注意力机制

**中图分类号:**V219;TP173;E917 **文献标志码:**A **文章编号:**1000-2758(2024)04-0753-11

随着科技进步和时代变革,当今时代的战场态势信息数据具有海量、高增长率、多样化的特点,导致现代化战场环境的复杂性和信息不对称性呈指数级增长,未来的战争必然会朝着科技化、信息化不断转变。同时,战场态势数据信息的规模和维度不断增加,使得军事领域专家仅仅依靠经验难以从海量的信息数据中及时准确地识别出敌方目标的作战意图,这一现象加快了智能化技术在态势感知领域中的应用。因此,我国迫切需要利用前沿智能技术来提高空中目标作战意图识别的准确率。

为了提高对战场态势的信息挖掘能力和军事专家的决策效果,国内外众多学者针对目标意图识别问题开展了大量研究,空中目标作战意图识别已成为空战领域的研究热门。传统的空中目标意图识别方法主要有模糊理论、D-S 证据理论、模板匹配、模糊推理、贝叶斯网络等方法<sup>[1-2]</sup>。模糊理论主要分为模糊数学、模糊逻辑与人工智能、模糊系统、不确

定性和信息、模糊决策5个分支,但这5个分支之间有着紧密的联系<sup>[3-4]</sup>。D-S 证据理论属于不确定性推理方法,最早应用于专家系统中辅助决策<sup>[5-6]</sup>。模板匹配是最原始的模式识别方法,通过知识提取得到数据结构,进而构造模板,利用模板识别出与之具有相同特征的目标<sup>[7]</sup>。模糊推理以模糊逻辑为基础制定从输入到输出的映射过程,是不确定推理方法的一种<sup>[8]</sup>。贝叶斯网络是人工智能领域处理不确定性的主要方法之一,结合专家知识对目标意图进行推理<sup>[9-12]</sup>。文献[13]提出了一种模糊置信规则库(BRB)信息处理方法,根据数据集的分布特点设计了模糊集分割方法,并建立规则权重和特征权重优化模型,采用置信度最大原则对目标意图进行识别。但是,这些传统智能化方法很难用于处理现代战争中规模巨大、指数式增长、来源多样的态势信息数据。目前,一些学者在多源信息融合和量子决策领域展开了研究,解决了大规模数据难以处理、分析的问题<sup>[14-15]</sup>。

近年来,随着计算机性能不断提升和硬件的迭代更新,人工智能实现飞跃式发展,机器学习、深度学习等已成为智能化方法的主要代表。不同的机器学习方法<sup>[16-19]</sup>已经被应用到空中目标作战意图识

收稿日期:2023-06-30

基金项目:国家自然科学基金(61806219, 61703426, 61876189)与陕西省自然科学基金(2021JM-226)资助

作者简介:王科(1999—),硕士研究生

通信作者:宋亚飞(1988—),副教授 e-mail:Yafei\_song@163.com

别领域之中,通过机器学习模拟人脑特性的特点,结合数据驱动,从战场原始态势数据中提取关键信息,进一步预测目标意图行为,克服了传统模型在知识表达方面所面临的困难<sup>[20]</sup>。文献[21]提出一种 GASVM-HMM 模型,能够通过飞行员某一动作识别其当前意图以及下一个意图所对应的动作及动作序列。上述文献提出的方法虽然一定程度上解决了面临海量数据场景的难题,但是机器学习部分操作需要人为干预,而且算法复杂度比较高,预测耗时比较长,因此,难以很好地处理战场实时变化的复杂信息数据。文献[22]针对机器学习的缺陷,利用反向传播算法(back propagation, BP),同时引入 ReLU(recified linear unit)激活函数和自适应矩估计(adaptive moment estimation, Adam)优化算法,设计了一种基于深度神经网络的空中目标作战意图识别模型,加快了模型收敛速度,同时解决了模型陷入局部最优的问题。文献[23]提出一种基于长短时记忆神经网络(LSTM)的空中目标意图识别方法,对战场数据进行清洗及滑窗处理,相较于传统方法取得了更好的效果。文献[24]以循环神经网络(recurrent neural network, RNN)为基础,融合注意力机制,对每一个时间步长的隐层单元输出分配差异化权重,进一步提高模型意图识别的准确率。文献[25]使用注意力机制提取特征信息,利用时序卷积网络(temporal convolutional network, TCN)深度挖掘数据中潜在的信息,结合双向门控循环单元(bi-gate recurrent unit, BiGRU)减少历史信息的丢失,设计了一种基于时序卷积神经网络的空中目标作战意图识别模型,很好地解决了传统方法难以有效捕获动态战场态势数据的特征信息问题。文献[26]根据作战意图的特点,设计了一种层次聚合模型,模型底层利用卷积神经网络感知行为特征,中间层使用双向长短时记忆网络表达意图内部之间相互依赖关系,最顶层结合注意力机制提取识别任务中的关键特征,解决了传统方法无法表达意图的层次性及意图之间长时依赖关系的问题。

根据上述文献发现,大量的意图识别研究已结合卷积神经网络和循环神经网络的深度学习方法展开,然而大多卷积和循环神经网络因为其复杂的架构,所以模型收敛速度较为缓慢。目前,Transformer 模型在各领域越演越热,凭借其出色的性能和简单的架构已成为各领域的研究热门<sup>[27]</sup>。本文将 Transformer 方法应用到空中目标意图识别领域,并

提出 TST 意图识别模型。TST 中的自注意力机制(self-attention)能够显著增加对复杂数据的特征信息提取能力,同时,TST 模型拥有从整个输入数据获取信息的能力,可以准确地建模长距离的依赖关系。而且,TST 不仅可以实现并行运算,能够更好地处理高维态势数据,提升意图识别的准确率,同时还使用了一种新的正则化方法——零范数正则化,将模型中一些不需要的参数压缩,从而提高模型的泛化能力。由于双塔式架构的特点,TST 可以提取时间域和空间域的深层全局特征,并通过本文提出的 Temporal-Spatial Fusion 融合机制实现高效融合,深度挖掘态势数据中的高价值潜藏信息。最后,通过大量实验结果证明,TST 在空中目标作战意图识别领域取得了优异的效果,超越了目前该领域中一些较为先进的意图识别方法,对作战辅助系统有较高的参考价值。

## 1 空中目标意图识别问题描述

空中目标作战意图识别是通过分析战场环境中空中目标动态的、复杂的实时数据,包括目标的相对距离、速度、高度等,判定目标状态,如飞行、逃逸等,结合深度学习技术、先验知识、相关作战领域专家经验等高效、精准地推理出敌方空中目标状态及目标意图。

定义向量  $\mathbf{S}_T$  为连续时间  $T$  秒内的空中目标作战实时特征集,由  $t_1$  到  $t_T$  时刻的实时特征信息组成,可以表示为

$$\mathbf{S}_T = (S_{t_1}, S_{t_2}, \dots, S_{t_T}) \quad (1)$$

定义向量  $\mathbf{Q}$  为空中目标作战意图空间集,在实际空中战场环境中,敌方目标的行动受到多种因素影响,很容易进行误导和欺骗,因此单一时刻识别敌方作战意图可能不够准确和可靠。相比较而言,通过分析敌方空中目标在多个连续时刻的特征信息,推测其作战意图的准确度和科学性会更高<sup>[20]</sup>。因此,可以确定空中目标作战特征集到空中目标作战空间集的映射函数如(2)式所示

$$\mathbf{Q} = f(\mathbf{S}_T) = f(S_{t_1}, S_{t_2}, \dots, S_{t_T}) \quad (2)$$

显然,如果要准确识别空中目标作战意图,需要结合空战专家的军事知识和经验,进行复杂的信息提取、分析和推理等思维活动,这个过程不是简单的数学运算,无法用单一的数学公式来建立目标特征和目标意图之间的映射关系<sup>[28]</sup>。为此,本文通过搭

建 TST 神经网络模型并加入空中目标作战意图特征集来训练,建立了目标特征与目标意图之间的映射关系。空中目标意图识别整体流程如图 1 所示。

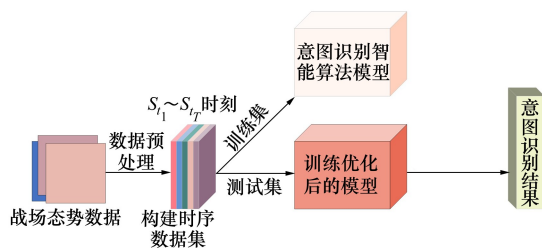


图 1 空中目标作战意图识别过程

### 1.1 空中目标作战意图空间描述

不同的作战环境、作战形式、敌方作战目标所对应的目标意图空间都有不同的特点。如陆光宇等<sup>[29]</sup>根据水下目标所具备的属性以及作战环境等定义目标意图空间为{撤离,攻击,巡逻};陈浩等<sup>[4]</sup>针对海上战舰编队,建立的目标意图空间为{撤退,掩护,攻击,侦察};马钰棠<sup>[30]</sup>根据空中目标作战特点建立的目标意图空间为{攻击,佯攻,撤退,侦察,监视,电子干扰}。本文结合空中目标作战环境、空中目标作战的属性,定义目标意图空间为{进攻,突防,警戒,侦察,佯攻,后撤,电子干扰}。

### 1.2 空中目标作战意图识别特征描述

通常,敌方空中目标执行的任务、任务针对的对象以及一些特定的战术动作都与目标的意图密切相关。为了清晰直观地表现出敌方目标的行为意图,本文根据空战领域专家经验和与目标意图相关联的条件选取空中目标作战特征。

从敌方空中目标作战执行的任务角度进行分析,空中目标作战执行不同任务时,目标的行为、所具备的条件都有所不同。例如,当歼击机、运输机和轰炸机执行突防攻任务时,其飞行速度都存在差异,歼击机通常会以大速度接敌,飞行速度一般为 735~1 470 km/h;运输机和轰炸机的飞行速度为 600~850 km/h<sup>[22]</sup>。同样以执行突防任务为例,空战目标一般有低空突防和高空突防 2 种形式,低空突防一般为 50~200 m,高空突防一般为 10 000~11 000 m;如果目标执行侦察任务,目标的侦察形式分为低空侦察和超高空侦察,飞行高度分别为 100~1 000 m 及 15 000 m 以上<sup>[22]</sup>。

此外,空中目标作战执行攻击、突防任务时,通常会开启电子干扰及对空雷达;执行侦察任务时,通

常会开启对空雷达及对海雷达<sup>[31]</sup>。在战场环境不断变化和目前对空战目标类型识别的研究较为薄弱的情况下,本文采用雷达的反射截面积大致判断空战目标类型,若反射截面积小于 1 m<sup>2</sup>,空战目标可以大致确定为小型战机或隐身战机;若反射截面积为 1~10 m<sup>2</sup>,则为中型战机;大型战机的雷达反射截面积通常大于 10 m<sup>2</sup>。

结合空战领域专家经验,在战场环境中能实时采集到的数据主要有空中作战目标飞行速度、高度、加速度、航向角、方位角及相对距离。

除此之外,空战能力因子及携带武器特征也是不容忽略的外部特征,对于空战目标的作战能力,本文结合李战武和欧爱辉等<sup>[32-33]</sup>对空战能力的评估方法,构建出空战能力因子评估函数

$$P = \alpha_1 \alpha_2 \alpha_3 \alpha_4 [\ln \beta_1 + \ln(\beta_2 + 1) + \ln(\beta_3 + 1)] \quad (3)$$

式中: $\alpha_1, \alpha_2, \alpha_3, \alpha_4$  分别表示战机操作性能、战机航行能力、战机生存能力及电子干扰对抗能力; $\beta_1, \beta_2, \beta_3$  分别表示战机机动能力、机载探测器性能、机载武器能力。不同战机的空战能力因子也有所不同,可以将相同战机的空战能力因子计算后保存,并根据情报数据定期更新<sup>[34]</sup>。

空战目标携带武器特征主要以识别是否携带武器为主,携带武器的空战目标则可以划分为:①高威胁目标,可能有攻击意图;②未携带武器的目标,以侦察、电子干扰意图为主。

综上,本文共选取 12 种关键意图特征,包括 4 种非数值型意图特征和 8 种数值型意图特征,如表 1~2 所示。

表 1 数值型空中目标特征

特征	描述
相对距离/km	空中目标与我方建筑物的距离
速度/(m·s <sup>-1</sup> )	空中目标飞行速度
高度/km	空中目标相对海平面的飞行高度
雷达反射面积/m <sup>2</sup>	空中目标在雷达上反射波的大小
方向角/(°)	空中目标与我方建筑的方位角
航向角/(°)	空中目标的飞行方向
加速度/(m·s <sup>-2</sup> )	空中目标的加速度
空战能力因子	空中目标的空中作战能力

其中,方向角和航向角都以正北方向均设为 0°。

表 2 非数值型空中目标特征

特征	描述	编码值
电子干扰	电子干扰装置开启状态	1
对空雷达状态	对空雷达开启状态	1
对地雷达状态	对地雷达关闭状态	0
携带武器特征	目标挂载武器	1

## 2 意图识别模型描述

本节详细地介绍了本文提出的 TST 模型的总体结构。TST 是基于 Transformer 模型的时空融合意图识别方法,Transformer 模型最近在自然语言处理和计算机视觉领域的各种任务都取得了前沿性的优势。受 Transformer 模型在众多领域表现优异的启发,本文将 Transformer 引入到空中目标作战意图识别领域,并设计为双塔式架构分别提取态势数据中的深层时间特征和空间特征,再利用时空融合模块进一步提取特征中的有效信息,提高意图识别性能并取得了优异的效果。

### 2.1 TST 模型整体框架

本文提出的 TST 模型主要由三部分组成,即 Temporal-Encoder 模块、Spatial-Encoder 模块和 Temporal-Spatial Fusion 模块,具体架构如图 2 所示。TST 使用 Transformer 框架中的编码器(Encoder)部分提取时间序列特征; Temporal-Transformer 和 Spatial-Transformer 都是由堆叠的编码器组成,分别用以提取时间域和空间域的深层全局特征,然后由 Temporal-Spatial Fusion 模块将时间特征和空间特征进行融合和识别。

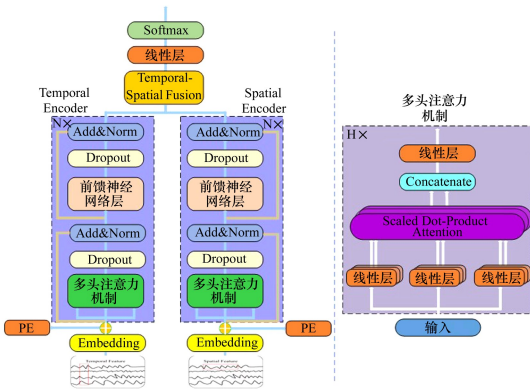


图 2 TST 模型框架图

### 2.2 Encoder 架构

Encoder 是 Transformer 的重要组成部分,通常

由多头注意力机制(multi-head attention)、前馈网络(feed forward network)、归一化(layer normalization)和残差连接(residual connections)组成。

图 2 右侧部分展示了多头注意力机制的结构。模型通过多头注意力机制可以基于相同的注意力机制学习到  $H$  组不同的线性投影,然后并行地传入注意力机制,最后融合到一起以产生最终的输出。其过程描述如下。

输入序列  $Y$  被映射为查询向量矩阵  $Q$ 、键向量矩阵  $K$  和值向量矩阵  $V$  经过注意力机制得到的权重矩阵  $W \in \mathbf{R}^{n \times n}$ ,这个过程被定义为

$$W = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_H}}\right) \quad (4)$$

式中,  $d_H$  表示多头注意力机制的维度  $H$ 。

通过注意力权重矩阵对值向量矩阵  $V$  进行加权后,经过单个注意力机制得到的输出为

$$H(Q, K, V) = WV \quad (5)$$

多头注意力机制将每一个注意力机制的输出并在一起,其过程如下

$$I_{MH}(Q, K, V) = [H_1(Q, K, V), H_2(Q, K, V), \dots, H_H(Q, K, V)]W_H \quad (6)$$

式中:  $W_H \in \mathbf{R}^{H \times d_H \times d_m}$  为权重矩阵;  $d_m$  表示向量的维度;  $I_{MH}(Q, K, V)$  表示多头注意力机制的输出。

多头注意力机制通过线性映射将嵌入有位置信息的高维向量序列  $Y$  映射为  $Q, K, V$ 。然后,通过  $Q$  和  $K$  计算特征之间的相关性,得到注意力权重矩阵。再利用注意力矩阵对  $V$  进行加权,得到单头注意力机制的输出,最后对单头注意力机制的输出进行合并,得到多头注意力机制的输出。

多头注意力机制得到的输出经过 Dropout 层、残差连接和归一化处理后与输入融合作为前馈网络层的输入。通常,前馈网络包含 3 层,第 1 层的作用是将输入映射到高维空间,第 2 层通过非线性激活层增强特征的非线性表示,第 3 层则进行降维,过程描述为

$$O_{FFN}(x) = (\text{Relu}(\text{LN}(I_{MH} + Y)W_1))W_2 \quad (7)$$

式中,  $W_1 \in \mathbf{R}^{d_m \times d_{FFN}}$ ,  $W_2 \in \mathbf{R}^{d_{FFN} \times d_m}$  表示 2 层维度变换的权重矩阵。LN 则表示归一化函数,  $O_{FFN}$  为前馈神经网络层的输出。

FFN 得到的输出再经过 Dropout 层、残差连接和归一化处理后,得到 Encoder 的输出,其过程为

$$E_{\text{output}} = \text{LN}(O_{FFN} + (\text{LN}(I_{MH} + Y))) \quad (8)$$

以上是单个编码器的输出,堆叠的编码器是串

连接的,将前一个编码器的输出作为输入,最终的输出就是输入序列的深层全局特征。

### 2.2.1 Temporal-Encoder 模块

为了提取输入序列中的时间特征,TST 使用带有掩码的自注意力,通过计算所有时间步长之间的成对注意力权重来关注所有通道上的每个点,同时,在多头注意力层中所有时间步长上形成注意力矩阵。与原始的 Transformer 架构一样,位置编码的全连接前馈网络层堆叠在每个多头注意力层上,用于增强特征提取。2 个子层中的残差连接引导信息和梯度流,随后进行层归一化。Temporal-Encoder 通过提取动态时间变化信息来有效地表示全局时间关系,而且 Temporal-Transformer (T-T) 由  $N$  层编码器组成,可以挖掘出序列的深层特征。总之,Temporal-Transformer 可以有效提取时间特征的深层全局信息,同时作为空间特征信息的补充,提高目标意图识别准确率。

### 2.2.2 Spatial-Encoder 模块

类似地,Spatial-Encoder 可以跨所有时间步长计算不同空间特征之间的注意力权重。通常,真实场景中提取得到的态势数据信息包含大量的噪声区域,Spatial-Encoder 可以有效地提取序列中的全局空间特征信息,并大幅降低序列中噪声区域的影响。Spatial-Encoder 通过计算距离单元之间的长距离相关性来对距离单元之间的空间相关性进行建模,从而自适应地提取目标区域的空间信息并抑制噪声区域的影响。此外,Spatial-Transformer (S-T) 同样包括  $N$  层编码器,可以有效地挖掘序列中的空间特征信息。

### 2.2.3 Temporal-Spatial Fusion 模块

为了融合 Temporal-Encoder 和 Spatial-Encoder 提取的特征信息,本文使用 Add & Concatenate 融合方法作为 Temporal-Spatial Fusion (T-S F) 模块。

输入的高维向量序列  $Y$  通过堆叠的 Encoder 可以提取其中的深度全局特征,其中 Temporal-Encoder 输出时间全局特征  $\Theta_T$ , Spatial-Encoder 输出空间全局特征  $\Theta_S$ ,然后再通过 Temporal-Spatial Fusion 实现时间特征  $\Theta_T$  和空间特征  $\Theta_S$  的融合,其通过注意力机制权重的计算过程如下

$$A_F = \text{softmax}([\Theta_T, \Theta_S]W_F) \quad (9)$$

为了实现时空融合,使用注意力权重矩阵  $A_F$  对时间和空间特征加权,然后融合的特征为

$$\Theta_F = [A_T, A_S]A_F \quad (10)$$

虽然时间和空间的全局特征信息都能用于识别任务,但仅通过简单的方法连接 2 个 Transformer 的特征,反而会损害两者的性能。本文的融合方法加入了注意力机制,可以有效地学习 2 个特征之间的相关性,从而可以自适应地为时间和空间全局特征分配权重,实现更高效的融合。

## 3 实验分析

### 3.1 性能评价指标

为评价本文提出的 TST 空中目标作战意图识别模型性能优劣,采用准确率 (accuracy),精确率 (precision),召回率 (recall),F1-分数 (F1-score),损失值 (Loss) 作为评价指标。

在目标意图识别任务中,计算某一意图行为的准确率、精确率、召回率、F1 分数、损失值时,该意图行为所有的标注样本为正例,其他意图为负例。准确率、精确率、召回率、F1-分数、损失值的计算方式分别为

$$A = \frac{P_T + N_T}{P_T + P_F + N_T + N_F} \quad (11)$$

$$P = \frac{P_T}{P_T + P_F} \quad (12)$$

$$R = \frac{P_T}{P_T + N_T} \quad (13)$$

$$F_1 = \frac{2 \times P \times R}{R + P} \quad (14)$$

$$L = - \sum p_i \times \log(q_i) \quad (15)$$

式中:  $P_T$  为正确分类的正例;  $N_F$  为错误分类的负例;  $P_F$  为错误分类的正例;  $N_T$  为正确分类的负例。(15) 式中,  $p_i$  表示真实标签的概率,  $q_i$  表示预测标签的概率。

### 3.2 实验数据与环境

实验中所使用的数据集来源于 Air Combat Maneuvering Instrumentation (ACMI) 作战仿真系统,由空战领域专家对其标签进行修订。数据集包括 8 400 个训练样本和 2 100 个测试样本,共 10 500 个样本,每个样本数据包含连续 12 帧信息,每帧信息包括 12 个维度的特征信息。实验使用 Python 3.7 语言,在 GPU 和 CUDA 12.1 加速环境下运行,采用 PyTorch 1.12.1 深度学习框架,实验中使用的操作系统

是 Windows11, 硬件配置是 Inter ® Xeon ® Silver 4214R CPU、NVIDIA A100 80GB PCIe 显卡、128G 内存。

### 3.3 数据输入

由于战场态势数据多样复杂, 传感器捕捉到的数据来自不同机型的不同特征状态, 例如, 一些战机攻击意图的飞行高度为 1 000~6 000 m。不同数值型特征的单位 and 数值差异较大, 会影响 TST 模型的收敛效率和准确率, 为此, 本文对数据进行了 min-max 标准化处理

$$X'_{in} = \frac{x_{in} - \min X_i}{\max X_i - \min X_i} \quad (16)$$

式中:  $X_i$  表示第  $i$  种数值型特征;  $x_{in}$  表示第  $i$  种数值型特征的第  $n$  个原始值,  $\max X_i$  表示第  $i$  种数值型特征的最大值。经过 min-max 标准化处理后, 每种数值型特征的原始值都被映射到区间  $[0, 1]$ , 显著降低了数值型特征由于方差大导致的影响。

### 3.4 实验结果与分析

#### 3.4.1 TST 模型结果分析

TST 模型测试集和训练集的准确率和损失值结果如图 3 所示。在训练次数为 20 左右时, 模型达到收敛状态, 此时训练集的准确率达到 100% 并保持稳定; 训练集的损失值在 0.000 03 上下浮动, 最低可达 0.000 026; 测试集的准确率在 99% 上下浮动, 最高可达 99.15%; 测试集的损失值在 0.007 上下浮动, 最低可达 0.005 7。

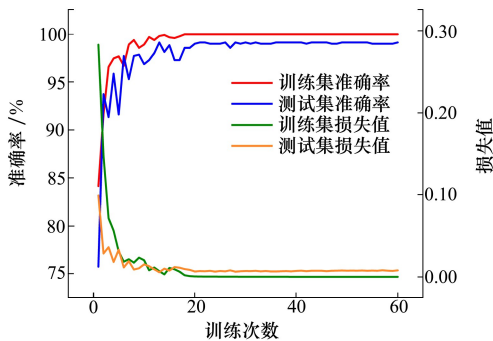


图 3 TST 模型训练和测试的准确率和损失值

为了进一步分析模型的意图识别结果在各意图下的具体情况, 绘制了测试集的混淆矩阵, 如图 4 所示, 混淆矩阵的对角线表示识别的正确率。根据图中的结果, 可以看出本文所提模型对 7 类意图的识别准确率都较高, 但在识别进攻/佯攻/突防的意图类别上, 模型将小部分样本识别错误, 导致错分情况

的原因是这 3 类意图对应的样本特征相似度较高。总体来看, 模型对后撤的意图识别精度最高, 对其余 6 类意图也有较高的识别准确率, 而且准确率已经超过了现有的识别方法, 能在一定程度上减少战时指挥员的压力。

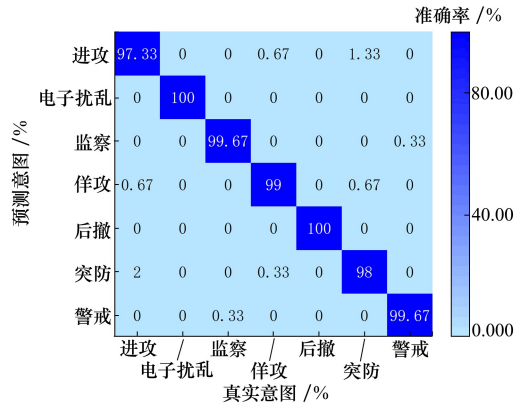


图 4 TST 模型意图识别结果的混淆矩阵

#### 3.4.2 模型消融实验分析

消融实验结果如表 3 所示, 其中 Add 是指仅通过简单的连接来融合 Temporal-Transformer 和 Spatial-Transformer 的输出。同时通过计算各模型的精确率、召回率及 F1-分数 3 种评价指标对比分析每类意图的反应情况, 如表 4 所示。

表 3 消融实验结果

模型组成结构				准确率/%	损失值
T-T	S-T	Add	T-SF		
✓	✓		✓	99.15	0.0057
✓	✓	✓		97.73	0.0814
✓			✓	95.45	0.3473
	✓		✓	99.01	0.0071

表 4 消融实验中各评价指标结果

模型	精确率/%	召回率/%	F1-分数/%
TST	99.28	99.32	98.85
T-T	91.92	91.78	90.71
S-T	98.47	98.89	98.65
TST-Add	96.16	96.67	96.16

根据上述结果可以分析出, TST 模型性能十分优越, 在准确率和损失值的表现也非常突出。同时, TST 在精确率、召回率及 F1-分数的表现相比融合前的 Transformer 更加优异, 进一步证明了通过 Temporal-Spatial Fusion 将 Temporal-Transformer 和

Spatial-Transformer 结合得到的 TST 能够在意图识别领域发挥独特的优势。

从消融实验结果能够看出, Temporal-Transformer 的表现不如 Spatial-Transformer, 为了进一步验证时间步长对 TST 的影响, 本文根据数据集的特点, 选取不同时间步长作为对比, 结果如表 5 所示。

表 5 时间步长对 TST 准确率的影响

时间步长	准确率/%		
	TST	T-T	S-T
12	99.15	95.45	99.01
6	98.58	93.18	97.59
1	83.99	79.89	83.43

将时间步长分别调整为 12, 6 和 1 来验证其对准确率的影响。通过表 5 结果可以发现, 时间步长的缩短对 TST 准确率的影响比较小。同时, 说明样本空间特征的影响大于时间特征对 TST 准确率的影响。

### 3.4.3 各模型对比实验分析

为了证明 TST 模型性能的优越性, 现使用准确率较高的多种先进模型对比意图识别结果的准确率及损失值。对比模型包括文献 [35] 所提的基于 GRU 的 FCN 空中目标意图识别方法; 文献 [36] 的 BiLSTM-Attention 空中目标意图识别模型; 文献 [37] 的基于 BiGRU 模块结合注意力机制的 TCN 意图识别模型; 文献 [38] 提出了基于 LSTM 的空中目标意图识别模型, 本文在此基础上加入了 FCN 神经网络, 以提高可比性; 为了进一步验证 TST 时空自注意力机制的效果, 本文引入了循环神经网络 (RNN) 结合时空注意力机制进行对比, 对比结果如表 6 所示。

表 6 各对比模型意图识别准确率和损失值

模型	准确率/%	损失值
TST	99.15	0.005 7
BiGRU-FCN	98.72	0.028 0
BiLSTM-Attention	98.58	0.007 9
LSTM-FCN	98.30	0.033 0
GRU-TCN	98.01	0.011 0
RNN-Attention	97.16	0.068 0

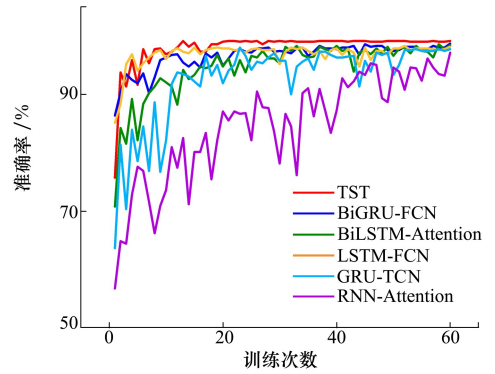


图 5 各对比模型准确率结果

通过以上实验结果, 可以观测出 TST 在空中目标意图识别领域的优越性, 该模型的准确率不仅超越了目前一些先进的空中意图识别模型, 还能迅速高效地收敛得到最优结果。同样, 如表 7 所示, 本文所提的模型在损失值、精确率、召回率、F1-分数的表现上也优于各对比模型, 不仅能快速降低损失值, 最低能达到 0.005 7, 而且在精确率、召回率、F1-分数都表现优异。

表 7 对比实验中各评价指标结果

模型	精确率/%	召回率/%	F1-分数/%
TST	99.28	99.32	98.85
BiGRU-FCN	98.18	97.89	97.64
BiLSTM-Attention	98.28	97.61	97.18
LSTM-FCN	97.79	98.17	98.59
GRU-TCN	97.00	95.80	94.90
RNN-Attention	96.18	96.64	96.40

根据图 5 可以发现, TST 相较于其他对比模型收敛速度最快, 为了进一步证明 TST 收敛的迅速性, 本文将各对比模型的参数量与收敛所需时间同 TST 进行了对比, 结果如表 8 所示。

表 8 模型参数量、收敛轮数及训练时长比较

模型类别	模型参数量/Mb	收敛所需轮数	训练时长/s
TST	8.03	20	265
BiGRU-FCN	2.32	121	312
BiLSTM-Attention	4.75	236	694
LSTM-FCN	2.20	372	749
GRU-TCN	2.41	147	472
RNN-Attention	1.26	217	521

表 8 结果都是在与 TST 相同环境和参数设置下得出的, 因此, 部分模型在此条件下收敛速度较为缓

慢。从表 8 中可以看出, TST 模型参数量相比其他模型规模更大, 但是经过 20 轮的训练就能够达到收敛状态, 同时训练耗费的时间也是最短的。总体来说, TST 模型表现出了优越的性能。

## 4 结 论

1) 针对传统卷积和循环神经网络模型架构复杂、处理时序问题训练时间过长、模型收敛速度缓慢、容易陷入局部最优, 以及在处理长序列数据时存在记忆衰退现象, 容易丢失序列中重要信息的问题, 本文将 Transformer 方法引入到空中目标意图识别领域, 并设计了一种 TST 模型, 该模型可以实现并行运算, 能够更好地处理高维态势数据, 提高意图识别的准确率, 而且相比其他模型能够迅速收敛, 达到最佳效果。

2) 由于具有双塔式架构, TST 可以有效地挖掘战场态势数据中时间域和空间域的深层特征信息, 能够进一步提取战场态势数据中蕴藏的目标真实意图的关键特征。同时, 本文通过实验与目前比较先进的卷积神经网络和循环神经网络模型进行效果对比, 并进行了自身的消融实验验证, 结果表明, TST 在各方面指标都能够取得优异的效果, 对辅助指挥员进行指挥决策及应对突发情况具有参考意义。

在未来的工作中, 将对模型进行优化使其能够有效辨别欺骗性意图和相似意图, 并将其设计为轻量化的 TST, 在不影响性能的同时, 缩减其参数量, 进一步提升训练速度。此外, 大多现有研究都是基于有监督的学习方式来训练模型, 数据标签大多由空战领域专家所制订, 可能会存在误判的情况, 导致训练的模型也会识别错误, 考虑无监督或者半监督学习方式可能会使意图识别结果更加准确。

## 参考文献:

- [1] 王海旺, 史红权, 赵晓哲. 目标意图识别方法综述[C]//2020 中国系统仿真与虚拟现实技术高层论坛, 北京, 2020: 3  
WANG Haiwang, SHI Hongquang, ZHAO Xiaozhe. A summary of target intent identification methods[C]//2020 China System Simulation and Virtual Reality Technology Forum, Beijing, 2020: 3 (in Chinese)
- [2] 龚诚, 衣军波, 蔡林恭, 等. 基于知识图谱嵌入学习的目标意图智能识别方法[J]. 舰船电子工程, 2023, 43(1): 32-35  
GONG Cheng, YI Junbo, CAI Lingong, et al. Intelligent recognition method of target intention based on knowledge graph embedded learning[J]. Ship Electronic Engineering, 2023, 43(1): 32-35 (in Chinese)
- [3] 张肃, 程启月, 解瑶, 等. 不确定空情信息条件下的意图识别方法[J]. 空军工程大学学报, 2008(3): 50-53  
ZHANG Su, CHENG Qiyue, XIE Yao, et al. A method of inference intention with uncertain aerial information[J]. Journal of Air Force Engineering University, 2008(3): 50-53 (in Chinese)
- [4] 陈浩, 任卿龙, 滑艺, 等. 基于模糊神经网络的海面目标战术意图识别[J]. 系统工程与电子技术, 2016, 38(8): 1847-1853  
CHEN HAO, REN Qinglong, HUA Yi, et al. Fuzzy neural network based tactical intention recognition for sea targets[J]. System Engineering and Electronics, 2016, 38(8): 1847-1853 (in Chinese)
- [5] 孙越林, 包磊. 基于 D-S 证据理论的海战场目标战术意图识别技术研究[J]. 舰船电子工程, 2012, 32(5): 48-51  
SUN Yuelin, BAO Lei. Study on recognition technique of targets' tactical intentions in sea battlefield based on D-S evidence theory[J]. Ship Electronic Engineering, 2012, 32(5): 48-51 (in Chinese)
- [6] 张瑜, 邓鑫洋, 李明烜, 等. 基于证据网络因果分析的空中目标意图识别[J]. 航空学报, 2022, 43(增刊): 143-156  
ZHANG Yu, DENG Xinyang, LI Mingda, et al. Air target intention recognition based on evidence-network causal analysis[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(suppl): 143-156 (in Chinese)
- [7] 夏曦. 基于模板匹配的目标意图识别方法研究[D]. 长沙: 国防科学技术大学, 2006  
XIA Xi. The study of target intention assessment method based on the template-matching[D]. Changsha: National University of Defense Technology, 2006 (in Chinese)
- [8] 徐建平, 张立凡, 韩德强. 基于模糊推理的空中目标意图识别[J]. 指挥信息系统与技术, 2020, 11(3): 44-48  
XU Jianping, ZHANG Lifan, HAN Deqiang. Air target intention recognition based on fuzzy inference[J]. Command Information System and Technology, 2020, 11(3): 44-48 (in Chinese)
- [9] 耿振余, 张建康. 基于贝叶斯网络的空中目标作战企图识别研究[J]. 现代防御技术, 2008, 36(6): 40-44



- GENG Zhenyu, ZHANG Jiankang. Research on air target combat intention inference based on bayesian networks[J]. Modern Defence Technology, 2008, 36(6): 40-44 (in Chinese)
- [10] 王昊冉. 基于多实体贝叶斯网络的空中目标意图识别方法研究[D]. 长沙:国防科学技术大学, 2011  
WANG Hongran. Research on air target intention recognition based on multi-entities bayesian networks[D]. Changsha: National University of Defense Technology, 2011 (in Chinese)
- [11] 余振翔, 胡笑旋, 夏维. 基于模糊动态贝叶斯网空战敌方作战企图识别[J]. 合肥工业大学学报, 2013, 36(10): 1210-1216  
YU Zhenxiang, HU Xiaoxuan, XIA Wei. Foe intention inference in air combat based on fuzzy dynamic Bayesian network[J]. Journal of Hefei University of Technology, 2013, 36(10): 1210-1216 (in Chinese)
- [12] ZHANG W, YANG F, LIANG Y. A Bayesian framework for joint target tracking, classification, and intent inference[J]. IEEE Access, 2019, 7: 66148-66156
- [13] 王海滨, 关欣, 衣晓, 等. 基于改进模糊置信规则的意图识别方法[J]. 电子与信息学报, 2023, 45(3): 941-948  
WANG Haibin, GUAN Xin, YI Xiao, et al. An intention recognition method based on fuzzy belief-rule-base[J]. Journal of Electronics & Information Technology, 2023, 45(3): 941-948 (in Chinese)
- [14] 徐艺博, 于清华, 王炎娟, 等. 基于多源信息融合的巡飞弹对地目标识别与毁伤评估[J]. 系统仿真学报, 2023, 36(2): 511-521  
XU Yibo, YU Qinghua, WANG Yanjuan, et al. Ground target recognition and damage assessment of patrol missiles based on multi-source information fusion[J]. Journal of System Simulation, 2023, 36(2): 511-521 (in Chinese)
- [15] FUYUAN X, WITOLD P. Negation of the quantum mass function for multisource quantum information fusion with its application to pattern classification[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2022, 45(2): 2054-2070
- [16] 胡智勇, 刘华丽, 龚淑君, 等. 基于随机森林的目标意图识别[J]. 现代电子技术, 2022, 45(19): 1-8  
HU Zhiyong, LIU Huali, GONG Shujun, et al. Target intention recognition based on random forest[J]. Modern Electronic Technology, 2022, 45(19): 1-8 (in Chinese)
- [17] 杨雨田, 杨杰, 李建国. 基于 EMEBN 的空中目标战术意图识别方法[J]. 火力与指挥控制, 2022, 47(5): 163-170  
YANG Yutian, YANG Jie, LI Jianguo. Research on air target tactical intention recognition based on EMEBN[J]. Fire Control & Command Control, 2022, 47(5): 163-170 (in Chinese)
- [18] 刘施彤. 基于机器学习的空中目标意图识别技术研究与应用[D]. 成都:电子科技大学, 2022  
LIU Shitong. Research and application of air target intent recognition technology based on machine learning[D]. Chengdu: University of Electronic Science and Technology, 2022 (in Chinese)
- [19] 吴广宇, 史红权, 邱楚楚. 基于 SSA-SVM 的空中目标意图识别方法[J]. 舰船电子工程, 2022, 42(3): 29-34  
WU Guangyu, SHI Hongquan, QIU Chuchu. Intention recognition method of air target based on SSA-SVM[J]. Ship Electronic Engineering, 2022, 42(3): 29-34 (in Chinese)
- [20] 欧微, 柳少军, 贺筱媛, 等. 战场对敌目标战术意图智能识别模型研究[J]. 计算机仿真, 2017, 34(9): 10-14  
OU Wei, LIU Shaojun, HE Xiaoyuan, et al. Study on the intelligent recognition model of enemy target's tactical intention on battlefield[J]. Computer Simulation, 2017, 34(9): 10-14 (in Chinese)
- [21] 江佳运, 孙有朝, 晏传奇. 基于 GASVM-HMM 算法的飞行员操控意图识别[J]. 航空计算技术, 2023, 53(1): 43-47  
JIANG Jiayun, SUN Youchao, YAN Chuanqi. Pilot manipulation intention recognition based on GASVM-HMM algorithms[J]. Aeronautical Computing Technique, 2023, 53(1): 43-47 (in Chinese)
- [22] 周旺旺, 姚佩阳, 张杰勇, 等. 基于深度神经网络的空中目标作战意图识别[J]. 航空学报, 2018, 39(11): 200-208  
ZHOU Wangwang, YAO Peiyang, ZHANG Jieyong, et al. Combat intention recognition for aerial targets based on deep neural network[J]. Acta Aeronautica et Astronautica Sinica, 2018, 39(11): 200-208 (in Chinese)
- [23] 钱钊, 刘钦, 鹿瑶, 等. 基于长短时神经网络的目标意图识别[J]. 太赫兹科学与电子信息学报, 2022, 20(11): 1156-1162  
QIAN Zhao, LIU Qin, LU Yao, et al. Identification of target's combat intention based on long short term memory network[J]. Journal of Terahertz Science and Electronic Information Technology, 2022, 20(11): 1156-1162 (in Chinese)
- [24] 吴南方. 基于循环神经网络的集群意图识别方法研究[D]. 成都:西南交通大学, 2019  
WU Nanfang. Research on cluster intention recognition method based on recurrent neural network[D]. Chengdu: Southwest Jiao-

- tong University, 2019 (in Chinese)
- [25] 滕飞, 邢千里, 王龙飞, 等. 基于时序卷积网络的空中目标作战意图识别模型[C]//2022 中国自动化大会, 厦门, 2022: 8  
TENG Fei, XING Qianli, WANG Longfei, et al. A model for aerial target intention recognition based on TCN[C]//2022 China Automation Conference, Xiamen, 2022: 8 (in Chinese)
- [26] 李颖, 武君胜, 李伟刚, 等. 一种识别作战意图的层次聚合模型[J]. 西北工业大学学报, 2023, 41(2): 400-408  
LI Ying, WU Junsheng, LI Weigang, et al. A hierarchical aggregation model for identifying operational intent[J]. Journal of Northwestern Polytechnical University, 2023, 41(2): 400-408 (in Chinese)
- [27] SINGH S, MAHMOOD A. The NLP Cookbook: modern recipes for transformer based deep learning architectures[J]. IEEE Access, 2021, 9: 68675-68702
- [28] 欧微, 柳少军, 贺筱媛, 等. 基于时序特征编码的目标战术意图识别算法[J]. 指挥控制与仿真, 2016, 38(6): 36-41  
OU Wei, LIU Shaojun, HE Xiaoyuan, et al. Tactical intention recognition algorithm based on encoded temporal features[J]. Command Control and Simulation, 2016, 38(6): 36-41 (in Chinese)
- [29] 陆光宇, 丁迎迎. 水下平台对敌意图识别技术研究[J]. 指挥控制与仿真, 2012, 34(6): 100-102  
LU Guangyu, DING Yingying. Study on intention recognition to foe of underwater platform [J]. Command Control and Simulation, 2012, 34(6): 100-102 (in Chinese)
- [30] 马钰棠, 孙鹏, 张杰勇, 等. 样本不平衡下的空中群组意图识别方法[J]. 系统工程与电子技术, 2022, 44(12): 3747-3755  
MA Yutang, SUN Peng, ZHANG Jieyong, et al. Air group intention recognition method under imbalance samples[J]. Systems Engineering and Electronics, 2022, 44(12): 3747-3755 (in Chinese)
- [31] 刘钻东, 陈谋, 吴庆宪, 等. 非完备信息下无人机空战目标意图预测[J]. 中国科学:信息科学, 2020, 50(5): 704-717  
LIU Zhuandong, CHEN Mou, WU Qingxian, et al. Prediction of unmanned aerial vehicle target intention under incomplete information[J]. Scientia Sinica Informationis, 2020, 50(5): 704-717 (in Chinese)
- [32] 李战武, 常一哲, 杨海燕, 等. 基于动态威力场的协同空战态势评估方法研究[J]. 系统仿真学报, 2015, 27(7): 1584-1590  
LI Zhanwu, CHANG Yizhe, YANG Haiyan, et al. Situation assessment method for cooperative air combat based on dynamic combat power field[J]. Journal of System Simulation, 2015, 27(7): 1584-1590 (in Chinese)
- [33] 欧爱辉, 朱自谦. 基于多属性决策和态势估计结果的空战威胁评估方法[J]. 火控雷达技术, 2006(2): 64-67  
OU Aihui, ZHU Ziqian. A Method of threat assessment based on MADM and results of situation assessment in air to air combat [J]. Fire Control Radar Technology, 2006(2): 64-67 (in Chinese)
- [34] 奚之飞, 徐安, 寇英信, 等. 基于 PCA-MPSO-ELM 的空战目标威胁评估[J]. 航空学报, 2020, 41(9): 216-231  
XI Zhifei, XU An, KOU Yingxin, et al. Target threat assessment in air combat based on PCA-MPSO-ELM algorithm[J]. Acta Aeronautica et Astronautica Sinica, 2020, 41(9): 216-231 (in Chinese)
- [35] 丁鹏, 宋亚飞. 代价敏感的空中标意图识别方法[J]. 航空学报, 2023, 44(24): 176-191  
DING Peng, SONG Yafei. A cost-sensitive method for aerial target intention recognition [J]. Acta Aeronautica et Astronautica Sinica, 2023, 44(24): 176-191 (in Chinese)
- [36] 滕飞, 刘曙, 宋亚飞. BiLSTM-Attention:一种空中目标战术意图识别模型[J]. 航空兵器, 2021, 28(5): 24-32  
TENG Fei, LIU Shu, SONG Yafei. BiLSTM-Attention: an air target tactical intention recognition model[J]. Aero Weaponry, 2021, 28(5): 24-32 (in Chinese)
- [37] TENG F, SONG Y F, GUO X P. Attention-TCN-BiGRU: an air target combat intention recognition model[J]. Mathematics, 2021, 9(19): 2412
- [38] XUE J, ZHU J, XIAO J, et al. Panoramic convolutional long short-term memory networks for combat intension recognition of aerial targets[J]. IEEE Access, 2020, 8: 183312-183323

# Design on temporal-spatial Transformer model for air target intention recognition

WANG Ke, LI Chenghai, SONG Yafei, WANG Peng, LI Lemin

(Air and Missile Defense College, Air Force Engineering University, Xi'an 710051, China)

**Abstract:** The battlefield changes rapidly in information warfare. The battlefield situation data presents massive and diversified characteristics, which makes it increasingly difficult to identify the operational intent of air targets based on expert experience. In combination with the current state-of-the-art intelligent methods, the Transformer model is studied and introduced into the field of air target intent recognition for the first time, and a new intent recognition method temporal-spatial transformer(TST) is designed, which can effectively mine the deep feature information in the temporal and spatial domains of battlefield situational data to improve the accuracy of air target combat intention recognition. In addition, a comparative study of the four currently advanced neural network intention recognition methods shows that TST achieved outstanding performance in all indicators and outperformed all compared neural network models. TST method not only has excellent accuracy but also extremely fast convergence rate, which allows it to quickly capture key information from battlefield situation data for intention recognition.

**Keywords:** Transformer; temporal-spatial fusion; air target; inten recognition; self-attention

**引用格式:**王科, 李成海, 宋亚飞, 等. 面向空中目标意图识别的时空 Transformer 模型设计[J]. 西北工业大学学报, 2024, 42(4): 753-763

WANG Ke, LI Chenghai, SONG Yafei, et al. Design on temporal-spatial Transformer model for air target intention recognition[J]. Journal of Northwestern Polytechnical University, 2024, 42(4): 753-763 (in Chinese)