空中对抗场景下对比学习驱动的弱监督机动识别方法

朱龙俊[①②] 袁伟伟[③] 门雪峰[④] 童 伟[⑤] 吴 奇*[⑥]

①(上海交通大学自动化系 上海 200240)

②(上海师范大学天华学院 上海 201815)

③(南京航空航天大学计算机科学与技术学院 南京 211106)

④(北京飞宇星科技有限公司 北京 100048)

⑤(南京邮电大学自动化学院 南京 210023)

⑥(上海交通大学计算机科学与工程系 上海 200240)

摘 要：针对空中对抗场景中飞行机动标注数据获取困难、时序特征提取不充分等问题，该文提出一种基于对比学习的弱监督机动识别方法，旨在提升机动识别性能。通过将视觉表征对比学习的简单框架(SimCLR)创新性地扩展至时间序列分析，设计针对时间序列的数据增强策略，构建具有时序不变性的特征空间。进而结合对比学习机制，在特征空间内形成正负样本组的竞争关系，有效抑制伪标签噪声干扰。最后结合微调技术，在DCS World飞行模拟数据上进行实验验证。结果表明，该方法能有效利用时间序列数据潜在信息，在缺乏标注数据情况下展现出良好性能，为空中对抗机动识别及时间序列分析领域提供了新的思路与方法。

关键词：对比学习；机动识别；时间序列；SimCLR；数据增强

中图分类号：TN911.7; TP181.8; TP391.41 **文献标识码**：A **文章编号**：1009-5896(2025)11-4504-11

DOI: 10.11999/JEIT250495 **CSTR**: 32379.14.JEIT250495

1 引言

在现代空战中，战机的机动行为复杂多变，准确识别战机的机动行为对于空战态势感知、制定有效的空战策略、保障飞行安全等至关重要。传统的机动识别方法大多依赖于大量的标注数据进行模型训练，然而在实际应用中，获取大规模高质量的标注数据往往面临成本高、效率低以及主观性强等诸多挑战。随着深度学习技术的发展，自监督学习和对比学习等新兴方法逐渐崭露头角，为解决数据标注难题提供了新的途径。对比学习的核心思想是通过拉近正样本对的特征距离，同时推远负样本对的特征距离，从而学习到具有判别性的特征表示[1,2]。其核心目标是将相似样本(即正样本)映射到嵌入空间的同一区域，而将不同样本(即负样本)映射到不同区域，最终形成一个可有效区分类别的特征空间。

对比学习已成为自我监督学习领域最成功的方法之一，在许多任务和研究领域都取得了进展[3–5]。例如，文献[6]提出一个用于视觉表征学习的简单对比学习框架(Simple framework for Contrastive Learning of visual Representations, SimCLR)，其通过最大化同一对象不同视图之间的相似性，最小化不同对象之间的相似性，成功从无标注数据中学习到高质量的特征表示。文献[7]提出一种用于自监督视频表征学习的视频级对比学习方法(Video-level Contrastive LeaRning, VCLR)。他们基于视频片段定义正样本对以捕捉全局上下文信息，并引入了时间顺序正则化项。文献[8]提出一种基于上下文对比学习的语义分割框架(Contextual contrastive learning for semantic segmentation, Contextrast)，其核心创新在于多尺度特征融合机制与边界感知负采样策略，实现更好的上下文理解和更精确的分割，为语义分割任务提供了新的解决方案。文献[9]提出一种用于自然语言处理中句子表征学习的对比学习框架(Contrastive LEArning for sentence Representation, CLEAR)，利用对比学习和独特的增强策略，提升了语言模型的性能，为自然语言处理的句子表征学习提供了新的视角。文献[10]将对比学习扩展到音乐领域，通过利用独特的音频增强策略，提供了一种无须标签即可学习音乐表征的新方法。然而，在时间序列分析领域，对比学习的应用尚处于起步阶段。这源于时间序列数据具有独特的时序性和多维稀疏性，传统的计算机视觉数据增强方法难以直接应用，需要设计专门针对时间序列数据特性的方法，以充分发挥对比学习在时间序列分析中的潜力。

收稿日期：2025-06-03；改回日期：2025-08-29；网络出版：2025-09-08

*通信作者： 吴奇 edmondqwu@163.com

基金项目：国家自然科学基金(T2325018, 62171274)，江苏省自然科学基金(BK20240641)

Foundation Items: The National Natural Science Foundation of China (T2325018, 62171274), Natural Science Foundation of Jiangsu Province (BK20240641)

本研究旨在提出一种基于对比学习的弱监督机动识别方法，通过构建有效的对比学习框架，实现对时间序列形式的飞行数据进行高效分析，从而准确识别战机的机动行为，减少对大量标注数据的依赖，提高机动识别的性能和泛化能力。本文主要贡献如下：

(1)提出一种创新的时间序列对比学习框架，将SimCLR框架扩展至飞行机动识别场景，设计了时序感知的时间压缩、缩放、排序、掩蔽、翻转5种数据增强方法，以生成同一实例的不同视图。

(2)基于弱监督特征对齐机制，引入自对比损失函数NT-Xent，通过动态调整正负样本对解决类别不平衡问题，助力模型在弱监督机动识别任务中更好地学习特征，提升了识别准确率与泛化能力。

(3)采用端到端微调策略，结合预训练特征与任务特定分类头，验证了提出的框架在长序列分类中的建模能力。

2　相关工作

现有飞机机动识别方法主要分为规则驱动与数据驱动两类技术路线。基于规则的方法依据领域专家的经验和知识，预先设定机动规则与特征阈值来识别机动行为，通过分析飞行参数变化范围和特征进行判断。在简单场景且规则明确时，能快速判断，结果也具有可解释性。例如，文献[11]利用飞行参数变化特征和专家先验知识，运用产生式规则构建了飞行动作识别知识库和正向精确推理机制。文献[12]提出基于动态贝叶斯网络的飞机机动动作识别模型用于飞行训练评估，但其模型参数最初依赖先验知识和专家经验设定，在参数学习方面有待完善。文献[13]通过多参数时程提取飞行机动，建立特征数据库，借助模式识别理论编程实现高精度自动机动识别。这类基于规则的方法存在明显不足，构建知识库难度大，难以涵盖所有复杂情况。飞行过程存在不确定性，机动动作多样，规则难以保证准确和完整，遇到非标准或复杂机动时，容易误判或无法识别。

数据驱动方法凭借对大量数据的学习，能挖掘潜在规律，在飞行机动识别上准确率较高，比基于规则的方法更擅长处理复杂数据和模式识别。例如，文献[14]将主成分分析(Principal Component Analysis, PCA)降维与改进支持向量机(Support Vector Machine, SVM)算法结合，用于识别和评估5种基本飞行机动。文献[15]提出在线集成半监督分类框架(Online Ensemble Semi-supervised Classification Framework, OESCF)，融合在线学习、集成学习等多种算法，提升空战目标机动识别的准确性和实时性。文献[16]为解决无人作战飞机(Unmanned Combat Air Vehicle, UCAV)飞行机动时空对齐问题，提出多策略仿射规范时间规整(Multi-strategy Affine Canonical Time Warping, MACTW)方法，结合动态时间规整(Dynamic Time Warping, DTW)与进化优化。该方法虽在UCR时间序列数据集和UCAV数据集上表现优于传统DTW变体，但计算复杂度高，影响实时应用。文献[17]提出结合无监督聚类和相空间重构的通用框架，用动态聚类分割序列、近似熵分析表征复杂性，在克服模型泛化和人工标注局限方面有竞争力。综上所述，这类数据驱动方法同样存在局限性。不仅对领域知识有一定依赖，特征工程阶段需人工处理数据特征，而且模型性能受数据质量影响大，数据若存在噪声、缺失或不平衡等问题，模型泛化能力会下降，识别效果不稳定。

随着深度学习发展，基于神经网络的机动识别方法成为研究热点，其中循环神经网络(Recurrent Neural Network, RNN)及其变体长短期记忆网络(Long Short-Term Memory, LSTM)表现突出。它们能有效处理时间序列数据，自动提取复杂特征，在高维度、非线性飞行数据处理上优势明显，且数据量增加时性能提升显著。文献[18]提出基于可解释LSTM的视距内空战机动决策框架。借助SHAP分析量化特征贡献，揭示基本战斗机机动(Basic Fighter Maneuvers, BFMs)决策关键特征，提升了飞行员在视距内空战的胜率和优势时间比。文献[19]构建了用于超视距空战机动意图识别的注意力增强金枪鱼群优化并行双向门控循环单元网络模型(Attention enhanced Tuna Swarm Optimization-Parallel Bidirectional Gated Recurrent Unit network, A-TSO-PBiGRU)。该模型采用并行双向门控循环单元结构，灵活调整前后向网络比例，融合前后向轨迹数据，以克服传统方法依赖单一方向数据的缺陷。文献[20]利用深度变分自编码器网络(Deep Variational Autoencoder, DVA)和自适应动态时间规整算法(Adaptive Dynamic Time Warping, ADTW)，构建飞行机动识别与质量评估集成系统，然后结合全局和局部特征评分算法评估飞行质量，但识别精度受数据质量和数量影响。可见，基于深度学习的机动识别方法存在的问题主要为模型结构复杂、可解释性差；训练依赖大量标注数据，成本高且耗时；计算资源消耗大，对硬件要求高，容易出现过拟合现象等。

上述研究表明，现有飞机机动识别方法各有局

限。本文旨在探索更高效、准确且适应性更强的方法：利用对比学习在弱监督条件下挖掘数据的相似性与差异性，提升模型特征学习能力，从而减少对大量标注数据的依赖、降低标注成本，为空战决策等领域提供更可靠的技术支持。

3 系统模型

3.1 模型概述

本文提出一种基于对比学习的弱监督机动识别方法，通过改进SimCLR框架(Simple framework for Contrastive Learning of visual Representations)[6]，实现时间序列数据分析的优化，模型架构如图1所示。整体框架包含数据增强(Data Augmentation)、编码器(Base Encoder)、投影头(Projection Head)，对比损失函数(Contrastive Loss Function)和微调(Fine-tuning)5个部分。此方法能有效挖掘时间序列潜在信息，在标注数据稀缺时性能优越，应用潜力大[6,10]。

3.2 时间序列数据增强

数据增强是对比学习的第1步，其目的是通过对原始数据应用各种变换，来生成同一实例的不同视图[21,22]。这些增强视图作为正样本对输入模型，来自不同样本的实例则作为负样本对。由于时间序列数据具有显著的时间和变量依赖性，标准的数据增强方法无法直接应用于时间序列分析，因此需要设计专门的增强方法以适应时间序列的特性。设矩阵$\boldsymbol{X}=[\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_n]$为机动时序数据，其中$\boldsymbol{x}_i\in R^m$ $(i=0,1,\cdots,n)$为节点t_i处的值，m表示特征维度，n表示时间步。为了更好地处理时间序列数据的特性，本文采用以下5种增强方法。

(1)时间压缩(Time Compression)：采用3次样条插值，实现0.5～2.0倍时间尺度变换。通过区域插值方法沿时间轴调整矩阵的长度，模拟时间序列在不同时间尺度下的变化。设压缩后的时间序列为$\tilde{\boldsymbol{x}}_i$，在区间$[t_i,t_{i+1}]$内的插值公式为：$\tilde{\boldsymbol{x}}_i=\boldsymbol{a}_i(\alpha t-t_i)^3+\boldsymbol{b}_i(\alpha t-t_i)^2+\boldsymbol{c}_i(\alpha t-t_i)+\boldsymbol{d}_i$。其中，$\alpha$表示压缩因子，控制时间序列在不同时间尺度下的变化程度；系数$\boldsymbol{a}_i$, $\boldsymbol{b}_i$, $\boldsymbol{c}_i$, $\boldsymbol{d}_i$可由原始时间序列节点处的函数值和1阶导数计算得出。

(2)缩放(Scaling)：随机选择10%～30%特征维度进行0.8～1.2倍缩放。对矩阵中选定比例的条带应用不同的缩放因子，增强模型对幅度变化的鲁棒性。设缩放因子矩阵$\boldsymbol{B}$为一个$m\times n$的矩阵，其中β_{ij}表示对矩阵$\boldsymbol{X}$中第i行第j列元素的缩放因子。则缩放后的矩阵$\tilde{\boldsymbol{X}}$可通过矩阵乘法表示为：$\tilde{\boldsymbol{X}}=\boldsymbol{X}\circ\boldsymbol{B}$。其中$\circ$表示Hadamard积，即2个同维度矩阵对应元素相乘。

(3)排列(Permutation)：将每个机动数据序列分割为5～10个片段随机重组，迫使模型在不依赖时间顺序的情况下学习时间序列中的潜在关系。假设排列操作将时间序列分割为k个子序列$\boldsymbol{S}=\boldsymbol{S}_1,\boldsymbol{S}_2,\cdots,\boldsymbol{S}_k$，随机排列后的片段顺序为$\pi=\pi_1,\pi_2,\cdots,\pi_k$，则排列后的数据$\tilde{\boldsymbol{X}}$为：$\tilde{\boldsymbol{X}}=[\boldsymbol{S}_{\pi1},\boldsymbol{S}_{\pi2},\cdots,\boldsymbol{S}_{\pi k}]$。

(4)掩蔽(Masking)：随机掩盖15%～25%时间

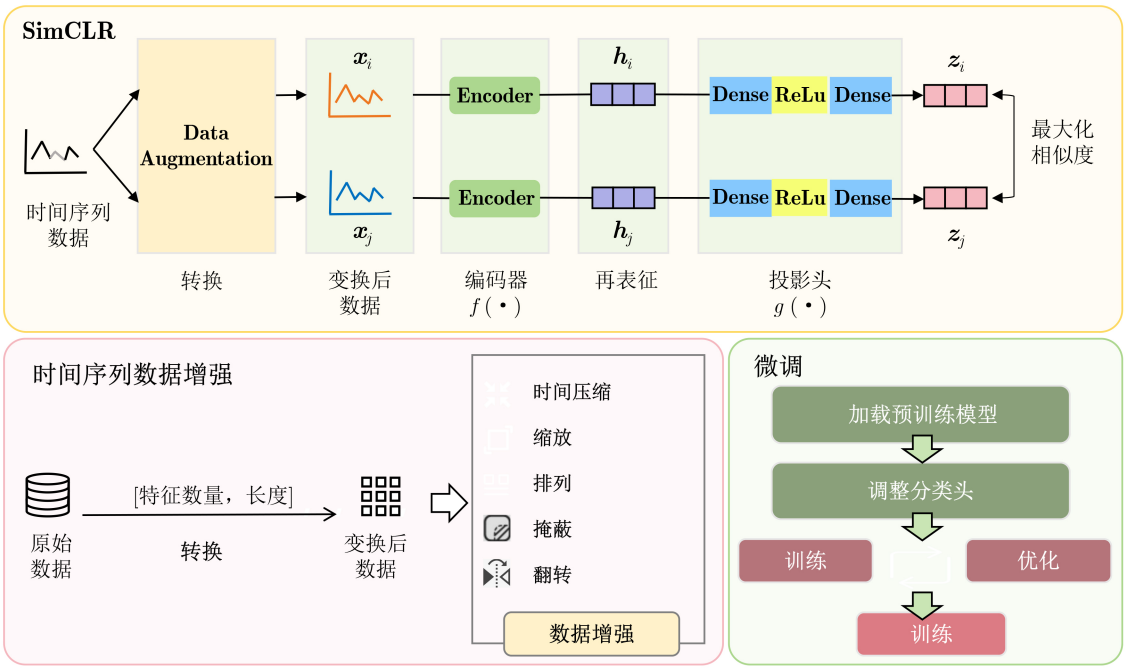

图 1 基于对比学习的弱监督机动识别框架

窗口，即在矩阵的随机条带上应用遮罩，模拟数据缺失的情况，促使模型利用上下文信息进行推断。随机选择x_i中的部分元素并替换为缺失值(如零或平均值)，得到新的序列

$$\tilde{\boldsymbol{x}}_i = \begin{cases} 0, \ t \in \text{掩码区域} \\ \boldsymbol{x}_i, \ \text{其他} \end{cases} \tag{1}$$

(5)翻转(Flipping)：采用100%概率进行时序反转，增强模型对时间方向变化的鲁棒性。翻转后的序列$\tilde{\boldsymbol{X}}$为：$\tilde{\boldsymbol{X}} = [\boldsymbol{x}_n, \boldsymbol{x}_{n-1}, \cdots, \boldsymbol{x}_1]$。

本文针对每个机动样本执行2次相互独立的数据增强变换，从而生成2个不同视图的样本。各项增强操作均以0.5的概率随机组合运用，通过多样化的增强组合，模型得以学习到更为广泛的特征，有力地提升了自身的泛化能力。

3.3 编码器

对比学习的编码器通常是一个深度神经网络，本文采用ResNet-18架构作为编码器对增强后的机动数据样本进行特征提取。由于空中对抗飞行机动时间序列数据具有强时序关联性与多维度耦合特性，为此本文对ResNet-18架构进行了适应性改进。考虑到数据的单通道特性，将第1个卷积层替换为输入通道为1、卷积核大小为3×3、步长为1、填充为1的卷积层；移除原始架构中的首个最大池化层，避免对时序序列进行空间下采样，从而保留机动状态转换过程中相邻时间步的细微特征变化，如俯仰角、滚转角的连续波动等；保留ResNet-18的残差块结构，去除末端的全连接分类层。

输入特征提取器得到特征，经过squeeze操作后输入到投影头MLP中。投影头MLP由2个全连接层组成，用于将特征提取器的输出投影到指定维度，把编码器生成的表示向量进一步映射到对比学习的隐空间中，优化特征表示。

3.4 对比损失函数

损失函数是对比学习框架的核心组件之一，能够有效地度量正样本对和负样本对之间的相似性[23–25]。本文采用了归一化温度缩放交叉熵损失函数(Normalized Temperature-scaled cross-ENTropy, NT-Xent)[6]。在给定样本集合$\{\tilde{\boldsymbol{x}}_k\}$中的正样本对$\tilde{\boldsymbol{x}}_i$和$\tilde{\boldsymbol{x}}_j$的情况下，模型的目标是学习一个特征空间，使得在给定$\tilde{\boldsymbol{x}}_i$的情况下，能够正确识别其对应的正样本$\tilde{\boldsymbol{x}}_j$，同时将负样本的表示推远。损失函数NT-Xent的计算公式为

$$L_{i,j} = -\ln \frac{\exp(\text{sim}(\boldsymbol{z}_i, \boldsymbol{z}_j)/\tau)}{\sum_{k=1}^{2N} 1_{[k\neq i]} \exp(\text{sim}(\boldsymbol{z}_i, \boldsymbol{z}_k)/\tau)} \tag{2}$$

其中，N是批量大小；$\boldsymbol{z}_i$和$\boldsymbol{z}_j$是正样本对的特征表示，$\boldsymbol{z}_k$表示批次中第k个样本的表示向量；τ是温度参数，用于缩放相似度得分，控制模型对不同样本对的区分程度；$\text{sim}(\boldsymbol{u}, \boldsymbol{v}) = \dfrac{\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}}{\|\boldsymbol{u}\| \, \|\boldsymbol{v}\|}$表示归一化后的向量$\boldsymbol{u}$和$\boldsymbol{v}$之间的余弦相似度；$1_{[k\neq i]}$是一个指示函数，当$k \neq i$时取值为 1，否则为 0。整个批量的NT-Xent损失为

$$L_{i,j} = \frac{1}{2N} \sum_{i=1}^{N} (L_{i,i+N} + L_{i+N,i}) \tag{3}$$

在前向传播中，先拼接2个视图特征表示，计算相似度矩阵，提取正、负样本对相似度得分并拼接，经温度参数缩放后，用交叉熵损失函数计算并归一化得到最终的NT-Xent损失。模型训练过程中采用LARS(Layer-wise Adaptive Rate Scaling)优化器，通过层自适应学习率缩放机制，提升训练效率与稳定性，助力模型在弱监督机动识别任务中更好地学习特征，提升识别准确率与泛化能力。

3.5 模型微调

为提升模型在时间序列分类任务中的性能，本文引入了微调技术：将对比学习阶段训练得到的模型参数作为初始值，此时模型已具备从大规模数据学习到的通用特征表示；根据机动识别任务需求，对模型最后一层进行调整。设模型的输出为$f(\boldsymbol{\theta})$，机动类别数为C，则输出层的调整可以通过修改最后一层的参数，使得输出维度从N变为C；在机动识别任务的标注数据集上进行训练，采用较小的学习率进行优化，以确保预训练阶段学到的通用特征不被大幅修改，保证模型稳定性和泛化能力。

优化目标是最小化机动识别任务特定损失函数，并加入正则化项避免过拟合。微调的目标函数为

$$L_{\text{FT}} = L_{\text{tk}}(f(\boldsymbol{\theta})) + \lambda \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|^2 \tag{4}$$

其中，$L_{\text{tk}}(f(\boldsymbol{\theta}))$是任务特定的损失函数，$L_{\text{FT}}$表示微调阶段对目标任务进行优化时的误差；$f(\boldsymbol{\theta})$是任务模型，$\boldsymbol{\theta}$是模型参数；$\boldsymbol{\theta}_0$是预训练阶段的模型参数，包含了从大规模数据集学习到的通用特征；λ是正则化参数，用于控制任务特定优化与预训练特征保留之间的平衡；$\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|^2$是正则化项，旨在惩罚过度调整预训练参数，以保持预训练阶段学到的通用特征。

4 仿真实验与分析

4.1 实验数据与应用场景

本研究所使用的飞行机动数据完全来源于高保真空战仿真平台DCS-World。该平台因其对F/A-18“黄蜂”战机的精准建模及多样化空战场景的高

度还原而被广泛应用于军事训练与科研。为获取具有代表性的机动响应序列，本文调用内置AI驾驶模块，分别在不同的初始高度、初始速度及预设机动指令条件下开展仿真飞行。数据覆盖多种典型飞行工况，并生成丰富的机动轨迹数据，为后续识别研究奠定基础。

数据采集阶段采用20 Hz高频雷达流(即每隔0.05 s记录1帧)，同步获取经度、纬度、海拔以及机体滚转角、俯仰角、偏航角等6维姿态信息，构成原始时空轨迹向量。考虑大地坐标系下的非正交性与传感器噪声，所有原始位置与姿态数据首先转换至东-北-天(ENU)局部参考系，以实现正交处理；随后引入扩展卡尔曼滤波器，利用递归估计方法抑制随机噪声与系统误差，最终获得高精度、低噪声的时序机动向量，确保数据在空间与时间维度上的一致性与可靠性。

针对空战中最具代表性的8种机动动作——半舵翻滚、俯冲、横滚、急转、盘旋、爬升、尾冲、旋降，构建了3级递进式数据集体系。基础数据集(D1～D3)由3个原始子集组成，每子集包含5类机动样本；融合数据集(D4～D6)在基础样本基础上分别增补2类机动，实现7种动作的覆盖；全量数据集(D7)整合全部8大典型机动，形成完整机动数据库(如表1所示)。各数据集均由300条、50帧采样的序列构成，并统一通过2.5 s的时间窗口标准化处理，以确保样本量平衡及时空特征对齐。该分层式设计不仅满足了对单一机动的精准识别需求，也显著提升了模型在复杂机动组合环境下的泛化能力，为仿真数据驱动的战场态势感知与决策支持系统研究提供了坚实的数据基础。

4.2 实验设置

本实验采用“对比学习预训练-微调检测”两阶段范式，参数设置综合考虑空战机动时序数据的强关联性与标注稀缺性。预训练阶段以改进后的ResNet-18作为基础编码器，基于SimCLR框架完成无监督特征学习，训练配置如下：批量大小(batch size)设为256；采用NT-Xent损失函数构建对比约束，温度参数τ取0.2以平衡正负样本区分度；优化器选用LARS，学习率0.05、动量0.9、LARS系数0.001，配合1e–4的权重衰减抑制过拟合，总计训练30个epoch。

微调检测阶段重点验证预训练编码器在机动识别任务中的迁移性能，采用“弱监督微调-线性分类-投票集成”的递进式方案。微调实验中，使用不同比例的标注训练集与验证集模拟弱监督场景，微调参数设置为：Adam优化器学习率0.0001，损失函数采用交叉熵结合L2正则化，正则化系数为0.001，batch size设为128，训练周期40个epoch，同时引入早停机制，连续5轮验证集准确率无提升则终止训练，以保障泛化性能。为提升长机动序列分类的准确性和鲁棒性，本文还在测试集上训练了一个线性分类器，采用投票方法对长机动序列的预测结果进行集成。具体而言，将50帧构成的机动单元划分为5个重叠片段，采用多数投票策略融合各片段预测结果。

表 1 机动数据集构成

数据集层级	数据集标识	包含的机动类型
基础层	D1	半舵翻滚、俯冲、横滚、盘旋、爬升
	D2	半舵翻滚、横滚、急转、爬升、旋降
	D3	半舵翻滚、俯冲、爬升、尾冲、旋降
融合层	D4	半舵翻滚、俯冲、横滚、急转、盘旋、爬升、旋降
	D5	半舵翻滚、俯冲、横滚、盘旋、爬升、尾冲、旋降
	D6	半舵翻滚、俯冲、横滚、急转、爬升、尾冲、旋降
全量层	D7	半舵翻滚、俯冲、横滚、急转、盘旋、爬升、尾冲、旋降

4.3 实验结果综合分析

微调阶段使用预训练参数初始化被用作下游任务的分类器，然后针对各种机动识别任务采用一定比例的机动数据，进行简单的训练。本研究采用了2%～30%的训练数据比例进行性能测试，验证弱监督框架在空战机动识别中的有效性。具体设置如表2所示。

表3–表5分别展示了不同微调比例下基础方案(Base Method, BM)和投票方案(Voting Method, VM)在7个数据集上的识别准确率。

表3所示极低数据场景下，基础方法Base在D1数据集上仅使用2%的标注数据，识别准确率即可达到35.6%，远超传统监督学习方法，后者需至少20%标注数据才能达到相近准确率。据表3所示，当微调比例提升至10%时，准确率跃升至83.0%，增幅达133%，证实了对比学习利用未标注数据的有效性。

在泛化能力方面，该框架在各类数据集上表现优异。例如，通过集成投票策略Voting，模型在D1数据集上展现出显著的低资源学习能力。在仅

表 2 微调比例划分情况

场景	数据范围(%)	间隔(%)	总测试点个数
极低数据场景	2,4,6,8,10	2	5
中低数据场景	12,14,16,18,20	2	5
低数据场景	22,24,26,28,30	2	5

表 3　极低数据场景识别准确率

微调比例(%)	D1		D2		D3		D4		D5		D6		D7	
	BM	VM	BM	VM	BM	VM	BM	VM	BM	VM	BM	VM	BM	VM
2	0.356	0.545	0.328	0.281	0.533	0.466	0.429	0.540	0.331	0.320	0.420	0.476	0.316	0.327
4	0.675	0.828	0.604	0.625	0.730	0.742	0.627	0.745	0.553	0.625	0.605	0.755	0.557	0.696
6	0.751	0.893	0.682	0.765	0.776	0.832	0.700	0.802	0.670	0.780	0.678	0.824	0.655	0.774
8	0.820	0.955	0.708	0.835	0.803	0.870	0.738	0.855	0.698	0.814	0.691	0.857	0.698	0.849
10	0.830	0.958	0.694	0.816	0.805	0.832	0.751	0.859	0.726	0.867	0.710	0.867	0.707	0.847

表 4　中低数据场景识别准确率

微调比例(%)	D1		D2		D3		D4		D5		D6		D7	
	BM	VM	BM	VM	BM	VM	BM	VM	BM	VM	BM	VM	BM	VM
12	0.843	0.937	0.758	0.956	0.838	0.917	0.763	0.859	0.713	0.857	0.718	0.886	0.763	0.908
14	0.862	0.980	0.763	0.886	0.840	0.917	0.767	0.859	0.736	0.886	0.735	0.901	0.757	0.861
16	0.867	0.980	0.775	0.909	0.867	0.981	0.768	0.831	0.755	0.901	0.755	0.930	0.770	0.873
18	0.871	0.958	0.779	0.886	0.890	0.999	0.781	0.859	0.749	0.915	0.759	0.930	0.775	0.896
20	0.875	0.980	0.803	0.979	0.879	0.959	0.768	0.816	0.759	0.915	0.764	0.930	0.773	0.896

表 5　低数据场景识别准确率

微调比例(%)	D1		D2		D3		D4		D5		D6		D7	
	BM	VM	BM	VM	BM	VM	BM	VM	BM	VM	BM	VM	BM	VM
22	0.890	0.977	0.783	0.862	0.888	0.979	0.775	0.830	0.771	0.912	0.765	0.926	0.783	0.916
24	0.883	0.979	0.781	0.859	0.889	0.999	0.798	0.855	0.753	0.897	0.761	0.915	0.772	0.895
26	0.882	0.976	0.798	0.885	0.886	0.998	0.815	0.898	0.770	0.915	0.773	0.926	0.782	0.918
28	0.892	0.977	0.795	0.905	0.887	0.999	0.817	0.902	0.771	0.912	0.775	0.928	0.780	0.918
30	0.895	0.979	0.791	0.861	0.870	0.977	0.816	0.915	0.772	0.915	0.768	0.928	0.789	0.928

使用2%标注数据的极端条件下，其分类准确率即可达到54.5%；当微调数据比例增至30%时，准确率进一步提升至97.9%，见表3和表5。这一结果相较于Base在同等条件下的准确率，体现了投票机制对小样本学习的显著增强作用。在模拟真实空战多源数据融合场景的数据集上，Base方法和Voting方法同样呈现显著性能差异。以融合数据集D6为例，综合表3～表5中的实验数据可以看出，随着微调比例从2%上升至30%，Base方法准确率从0.420逐步提升至0.768，增长缓慢且出现性能波动；而Voting方法凭借对比学习特征表示与投票集成策略，准确率从0.476快速提升至0.928，在低数据阶段增速较Base方法提升53%，且在14%微调比例时已达0.901，16%后稳定在0.930左右，标准差仅为0.012，显著低于Base方法的0.035。充分证明了该方法在利用未标注数据、抑制噪声干扰及提升复杂场景下的泛化与稳定性能方面的优越性。

图2呈现了基础数据集、融合数据集以及全量数据集场景下，Base方案与Voting方案的准确率对比情况。通过这些可视化图表，能够直观洞察到在各数据集不同微调比例条件下，两种方案准确率的动态变化趋势。

图3给出了各个数据集在不同微调比例下，分别采用Base方法和Voting方法下的平均准确率数值及对比。在基础数据集上，Voting方法凭借对关键特征的精准提取，显著超越Base方法。融合数据集里，尤其在复杂的全量数据集D7上，Voting方法的准确率标准差仅为0.011，而Base方法为0.015，优势依旧明显。可见Voting方法在平衡准确率、数据需求和模型复杂度上表现卓越，且提升幅度在低数据场景下时更具优势。

为量化各类数据增强操作对模型性能的具体贡献，本研究在D1, D2和D3 3个数据集上设计对比实验：分别单独应用时间压缩、缩放、排列、掩蔽、翻转5种时序增强方法，同时设置“随机组合”策略(每种增强以0.5概率独立叠加)作为对照。所有实验严格保持一致的技术框架(编码器+对比学习+微调流程+Voting方案)，以测试集识别准确

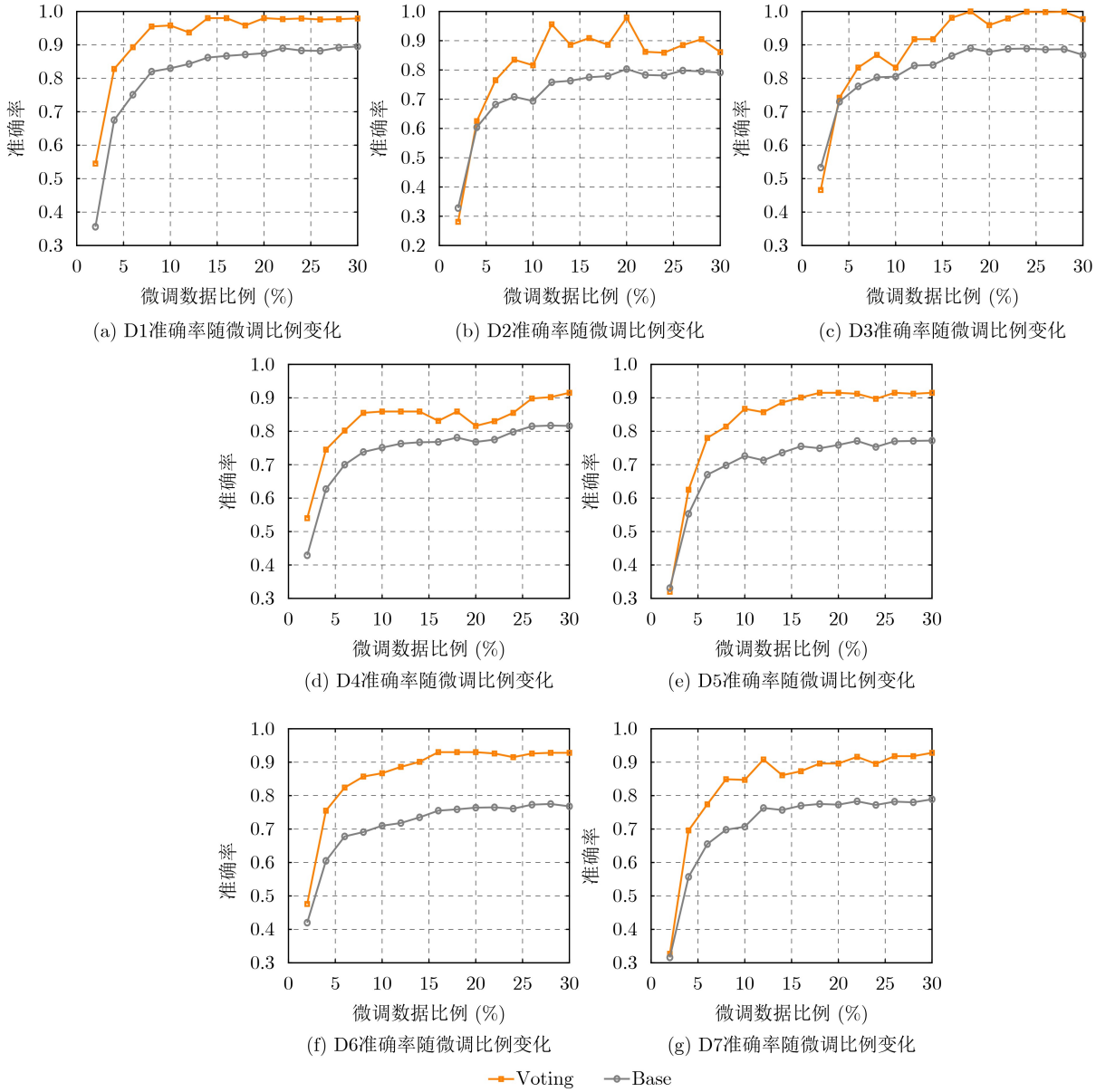

图 2 两种方案的准确率受微调比例影响对比

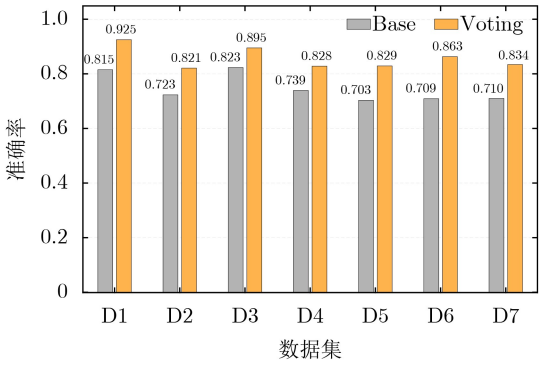

图 3 两种方案的平均准确率比较

率为核心评价指标，具体结果如表6所示。可以看出，“随机组合”策略在各数据集上的性能均优于单一增强方法：在D1数据集上，较表现最佳的“掩蔽”增强提升1.5%；在D2数据集上，较最优单一增强“掩蔽”提升1.3%；在D3数据集上，较最佳单一增强“掩蔽”提升1.2%。这一结果充分证明，多种增强操作并用能最大化丰富训练样本的时序分布，有效提升模型在弱监督机动识别任务中的鲁棒性和准确性。

为进一步验证本文所提出方法的有效性，实验选取7种涵盖传统机器学习与深度学习的经典模型作为基线对比，其与Voting方案的机动识别平均准确率见表7。结果显示，Voting方案在所有数据集上均表现出色。在D1数据集上，准确率达0.925，与表现最优的GRU(0.924)持平；在D2数据集，超越所有基线模型，较最优的TimesNet(0.820)略高；在D3数据集，以0.895高于各基线模型，包括

表 6　不同数据增强策略在各数据集上的识别准确率对比

数据集	时间压缩	缩放	排列	掩蔽	翻转	随机组合
D1	0.900	0.905	0.890	0.910	0.895	**0.925**
D2	0.805	0.800	0.790	0.808	0.795	**0.821**
D3	0.875	0.880	0.865	0.883	0.870	**0.895**

表 7　Voting方案与各基线模型的机动识别平均准确率对比

数据集	LSTM	GRU	T-Rep	XGBoost	Rocket	MLP	TimesNet	Voting
D1	0.900	0.924	0.915	0.898	0.876	0.912	0.899	**0.925**
D2	0.813	0.800	0.815	0.699	0.781	0.789	0.820	**0.821**
D3	0.859	0.851	0.888	0.758	0.802	0.851	0.885	**0.895**
D4	0.782	0.775	0.751	0.764	0.736	0.774	0.802	**0.828**
D5	0.752	0.776	0.731	**0.860**	0.742	0.781	0.794	0.829
D6	0.764	0.763	0.761	0.808	0.682	0.775	0.808	**0.863**
D7	0.782	0.794	0.738	0.803	0.730	0.761	0.766	**0.834**

表现较好的T-Rep(0.888)和TimesNet(0.885)；在D4, D6, D7等融合数据集上，也均优于所有基线模型，充分验证了本文方法的有效性。

实验结果显示，本文提出的对比学习驱动的弱监督机动识别框架，在不同复杂程度数据集以及不同微调比例条件下，均具备良好且稳定的有效性，为该框架在实际空战机动识别场景中的应用，提供了坚实的数据依据。

5　结论与展望

本文创新性地将SimCLR框架引入机动识别任务，提出基于对比学习的弱监督机动识别方法。通过设计针对性的数据增强手段，融合自监督学习与微调技术，充分挖掘时间序列数据中的潜在信息，在标注数据稀缺的情况下，实现了机动识别性能的显著提升。在模拟空战数据集上的实验表明，该方法在不同类别数据集上均能实现稳定的识别效果，且优于LSTM等经典的基线对比模型，为时间序列分类任务中的特征学习与模型优化提供了新的思路与实践参考。未来将从3个方向进一步展开研究：一是融合真实飞行数据，验证模型在实际场景中的泛化能力；二是探索动态自适应的数据增强策略，优化模型在复杂场景下的性能表现；三是考虑结合强化学习等技术，探索模型在动态环境中的自主决策能力提升路径，为飞行任务的智能化发展提供更多可能。

参 考 文 献

[1] TIAN Yonglong, SUN Chen, POOLE B, *et al*. What makes for good views for contrastive learning?[C]. The 34th International Conference on Neural Information Processing Systems, Vancouver, Canada, 2020: 573.

[2] CHUANG C Y, ROBINSON J, LIN Y C, *et al*. Debiased contrastive learning[C]. The 34th International Conference on Neural Information Processing Systems, Vancouver, Canada, 2020: 735.

[3] ZHENG Mingkai, WANG Fei, YOU Shan, *et al*. Weakly supervised contrastive learning[C]. IEEE/CVF International Conference on Computer Vision, Montreal, Canada, 2021: 10022–10031. doi: 10.1109/ICCV48922.2021.00989.

[4] KHOSLA P, TETERWAK P, WANG Chen, *et al*. Supervised contrastive learning[C]. The 34th International Conference on Neural Information Processing Systems, Vancouver, Canada, 2020: 1567.

[5] WU Linshan, ZHUANG Jiaxin, and CHEN Hao. VoCo: A simple-yet-effective volume contrastive learning framework for 3D medical image analysis[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2024: 22873–22882. doi: 10.1109/CVPR52733.2024.02158.

[6] CHEN Ting, KORNBLITH S, NOROUZI M, *et al*. A simple framework for contrastive learning of visual representations[C]. The 37th International Conference on Machine Learning, 2020: 149.

[7] KUANG Haofei, ZHU Yi, ZHANG Zhi, *et al*. Video contrastive learning with global context[C]. IEEE/CVF International Conference on Computer Vision Workshops, Montreal, Canada, 2021: 3188. doi: 10.1109/ICCVW54120.2021.00358.

[8] SUNG C, KIM W, AN J, *et al*. Contextrast: Contextual contrastive learning for semantic segmentation[C]. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, USA, 2024: 3732–3742. doi: 10.1109/CVPR52733.2024.00358.

[9] WU Zhuofeng, WANG Sinong, GU Jiatao, *et al*. CLEAR: Contrastive learning for sentence representation[J]. arXiv preprint arXiv: 2012.15466, 2020. doi: 10.48550/arXiv.2012.15466.

[10] SPIJKERVET J and BURGOYNE J A. Contrastive learning of musical representations[C/OL]. The 22nd International Society for Music Information Retrieval Conference, 2021: 673–681.

[11] 倪世宏, 史忠科, 谢川, 等. 军用战机机动飞行动作识别知识库的建立[J]. 计算机仿真, 2005, 22(4): 23–26. doi: 10.3969/j.issn.1006-9348.2005.04.007.
NI Shihong, SHI Zhongke, XIE Chuan, *et al*. Establishment of avion inflight maneuver action recognizing knowledge base[J]. *Computer Simulation*, 2005, 22(4): 23–26. doi: 10.3969/j.issn.1006-9348.2005.04.007.

[12] 孟光磊, 陈振, 罗元强. 基于动态贝叶斯网络的机动动作识别方法[J]. 系统仿真学报, 2017, 29(S1): 140–145. doi: 10.16182/j.issn1004731x.joss.2017S1020.
MENG Guanglei, CHEN Zhen, and LUO Yuanqiang. Maneuvering action identify method based on dynamic Bayesian network[J]. *Journal of System Simulation*, 2017, 29(S1): 140–145. doi: 10.16182/j.issn1004731x.joss.2017S1020.

[13] WANG Yongjun, DONG Jiang, LIU Xiaodong, *et al*. Identification and standardization of maneuvers based upon operational flight data[J]. *Chinese Journal of Aeronautics*, 2015, 28(1): 133–140. doi: 10.1016/j.cja.2014.12.026.

[14] LI Xiaokang, ZHU Tianyi, BIAN Zimu, *et al*. An improved algorithm for flight maneuver recognition and evaluation based on support vector machines[C]. 2024 International Conference on Cyber-Physical Social Intelligence (ICCSI), Doha, Qatar, 2024: 1–6. doi: 10.1109/ICCSI62669.2024.10799254.

[15] XI Zhifei, LYU Yue, KOU Yingxin, *et al*. An online ensemble semi-supervised classification framework for air combat target maneuver recognition[J]. *Chinese Journal of Aeronautics*, 2023, 36(6): 340–360. doi: 10.1016/j.cja.2023.04.020.

[16] WEI Zhenglei, DING Dali, ZHOU Huan, *et al*. A flight maneuver recognition method based on multi-strategy affine canonical time warping[J]. *Applied Soft Computing*, 2020, 95: 106527. doi: 10.1016/j.asoc.2020.106527.

[17] LU Jing, CHAI Hongjun, and JIA Ruchun. A general framework for flight maneuvers automatic recognition[J]. *Mathematics*, 2022, 10(7): 1196. doi: 10.3390/math10071196.

[18] WANG Can, TU Jingqi, YANG Xizhong, *et al*. Explainable basic-fighter-maneuver decision support scheme for piloting within-visual-range air combat[J]. *Journal of Aerospace Information Systems*, 2024, 21(6): 500–514. doi: 10.2514/1.I011388.

[19] LEI Xie, SHILIN D, SHANGQIN T, *et al*. Beyond visual range maneuver intention recognition based on attention enhanced tuna swarm optimization parallel BiGRU[J]. *Complex & Intelligent Systems*, 2024, 10(2): 2151–2172. doi: 10.1007/s40747-023-01257-3.

[20] TIAN Wei, ZHANG Hong, LI Hui, *et al*. Flight maneuver intelligent recognition based on deep variational autoencoder network[J]. *EURASIP Journal on Advances in Signal Processing*, 2022, 2022(1): 21. doi: 10.1186/s13634-022-00850-x.

[21] LUO Dongsheng, CHENG Wei, WANG Yingheng, *et al*. Time series contrastive learning with information-aware augmentations[C]. The 37th AAAI Conference on Artificial Intelligence, Washington, USA, 2023: 4534–4542. doi: 10.1609/aaai.v37i4.25575.

[22] CHEN Muxi, XU Zhijian, ZENG Ailing, *et al*. FrAug: Frequency domain augmentation for time series forecasting[J]. arXiv preprint arXiv: 2302.09292, 2023. doi: 10.48550/arXiv.2302.09292.

[23] HADSELL R, CHOPRA S, and LECUN Y. Dimensionality reduction by learning an invariant mapping[C]. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, USA, 2006: 1735–1742. doi: 10.1109/CVPR.2006.100.

[24] SCHROFF F, KALENICHENKO D, and PHILBIN J. FaceNet: A unified embedding for face recognition and clustering[C]. IEEE Conference on Computer Vision and Pattern Recognition, Boston, USA, 2015: 815–823. doi: 10.1109/CVPR.2015.7298682.

[25] SOHN K. Improved deep metric learning with multi-class n-pair loss objective[C]. The 30th International Conference on Neural Information Processing Systems, Barcelona, Spain, 2016: 1857–1865.

朱龙俊：女，博士生，副教授，研究方向为模式识别、机器学习和鲁棒控制等.

袁伟伟：女，教授，研究方向为数据挖掘、智能计算及人工智能的航空航天应用等.

门雪峰：男，工程师，研究方向为深度学习、脑认知技术应用.

童 伟：男，博士，研究方向为机器人视觉、人机交互、医学图像分析和脑认知.

吴 奇：男，教授，研究方向为深度学习、疲劳识别和人机交互等.

责任编辑：马秀强

Weakly Supervised Recognition of Aerial Adversarial Maneuvers via Contrastive Learning

ZHU Longjun[①②]　YUAN Weiwei[③]　MEN Xuefeng[④]　TONG Wei[⑤]　WU Qi[⑥]

[①](*Department of Automation, Shanghai Jiao Tong University, Shanghai* 200240, *China*)

[②](*School of Artificial Intelligence, Shanghai Normal University Tianhua College, Shanghai* 201815, *China*)

[③](*College of Computer Science and Technology, Nanjing University of Aeronautics and Astrnautics, Nanjing* 211106, *China*)

[④](*Beijing FISTAR TECHNOLOGY Co., Ltd., Beijing* 100048, *China*)

[⑤](*College of Automation, Nanjing University of Posts and Telecommunications, Nanjing* 210023, *China*)

[⑥](*Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai* 200240, *China*)

Abstract:

Objective　Accurate recognition of aerial adversarial maneuvers is essential for situational awareness and tactical decision-making in modern air warfare. Conventional supervised approaches face major challenges: obtaining labeled flight data is costly due to the intensive human effort required for collection and annotation, and these methods are limited in capturing temporal dependencies inherent in sequential flight parameters. Temporal dynamics are crucial for describing the evolution of maneuvers, yet existing models fail to fully exploit this information. To address these challenges, this study proposes a weakly supervised maneuver recognition framework based on contrastive learning. The method leverages a small proportion of labeled data to learn discriminative representations, thereby reducing reliance on extensive manual annotations. The proposed framework enhances recognition accuracy in data-scarce scenarios and provides a robust solution for maneuver analysis in dynamic adversarial aerial environments.

Methods　The proposed framework extends the Simple framework for Contrastive Learning of visual Representations (SimCLR) into the time-series domain by incorporating five temporal-specific data augmentation strategies: time compression, masking, permutation, scaling, and flipping. These augmentations generate multi-view samples that form positive pairs for contrastive learning, thereby ensuring temporal invariance in the feature space. A customized ResNet-18 encoder is employed to extract hierarchical features from the augmented time-series data, and a Multi-Layer Perceptron (MLP) projection head maps these features into a contrastive space. The Normalized Temperature-scaled cross-entropy (NT-Xent) loss is adopted to maximize similarity between positive pairs and minimize it between negative pairs, which effectively mitigates pseudo-label noise. To further improve recognition performance, a fine-tuning strategy is introduced in which pre-trained features are combined with a task-specific classification head using a limited amount of labeled data to adapt to downstream recognition tasks. This contrastive learning framework enables efficient analysis of time-series flight data, achieves accurate recognition of fighter aircraft maneuvers, and reduces dependence on large-scale labeled datasets.

Results and Discussions　Experiments are conducted on flight simulation data obtained from DCS World. To address the class imbalance issue, hybrid datasets (Table 1) are constructed, and training data ratios ranging from 2% to 30% are employed to evaluate the effectiveness of the weakly supervised framework. The results demonstrate that contrastive learning effectively captures the temporal patterns within flight data. For example, on the D1 dataset, accuracy with the base method increases from 35.6% with 2% labeled data to 89.5% when the fine-tuning ratio reaches 30% (Tables 3～5, Fig. 2(a)). To improve recognition of long maneuver sequences, a linear classifier and a voting strategy are introduced. The voting strategy markedly enhances few-shot learning performance. On the D1 dataset, accuracy reaches 54.5% with 2% labeled data and rises to 97.9% at a 30% fine-tuning ratio, representing a substantial improvement over the base method. On the D6 dataset, which simulates multi-source data fusion scenarios in air combat, the accuracy of the voting method increases from 0.476 with 2% labeled data to 0.928 with 30% labeled data (Fig. 2(f)), with a growth

rate in the low-data phase 53% higher than that of the base method. Additionally, on the comprehensive D7 dataset, the accuracy standard deviation of the voting method is only 0.011 (Fig. 2(g), Fig. 3), significantly lower than the 0.015 observed for the base method. The superiority of the proposed framework can be attributed to two factors: the suppression of noise through integration of multiple prediction results using the voting strategy and the extraction of robust features from unlabeled data via contrastive learning pre-training. Together, these techniques enhance generalization and stability in complex scenarios, confirming the effectiveness of the method in leveraging unlabeled data and managing multi-source information.

Conclusions This study applies the SimCLR framework to maneuver recognition and proposes a weakly supervised approach based on contrastive learning. By incorporating targeted data augmentation strategies and combining self-supervised learning with fine-tuning, the method exploits the latent information in time-series data, yielding substantial improvements in recognition performance under limited labeled data conditions. Experiments on simulated air combat datasets demonstrate that the framework achieves stable recognition across different data categories, offering practical insights for feature learning and model optimization in time-series classification tasks. Future research will focus on three directions: first, integrating real flight data to evaluate the model's generalization capability in practical scenarios; second, developing dynamically adaptive data augmentation strategies to enhance performance in complex environments; and third, combining reinforcement learning and related techniques to improve autonomous decision-making in dynamic aerial missions, thereby expanding opportunities for intelligent flight operations.

Key words: Contrastive learning; Maneuver recognition; Time series; Simple framework for Contrastive Learning of visual Representations (SimCLR); Data augmentation