

基于特征融合的飞行员进近着陆能力评价

杨泳^{1,2,3}, 吴冠辰^{1,3}, 徐开俊^{1,2,3}, 李东霖^{1,3}

(1. 中国民用航空飞行学院 飞行技术学院, 广汉 618307)

(2. 四川省民航飞行技术与飞行安全工程技术研究中心, 广汉 618307)

(3. 国产民机飞行与运行支持四川省工程研究中心, 广汉 618307)

摘要: 现有飞行员进近着陆操纵能力评价多依赖单一人工特征, 缺乏兼顾物理可解释性与复杂时序模式捕捉的特征体系, 且操纵行为与飞行品质的关联机制尚未明确。基于 105 例波音 737-800 机型快速存取记录器 (QAR) 数据, 构建“最后进近+着陆拉平”两阶段结果导向型评价体系并引入能量管理的评价, 融合人工特征与长短期记忆 (LSTM) 隐含特征表征操纵行为, 结合 SHAP 分析、K-means 聚类与统计检验, 揭示操纵特征对着陆品质的影响规律。结果表明: 评价结果与航司人工评估一致性达 90%, 核心特征与着陆品质呈显著相关, 并识别出两类典型操纵模式。

关键词: QAR 数据; 着陆操纵能力; 品质评价; 特征融合; SHAP 分析; LSTM

中图分类号: V323.11

文献标识码: AA

Pilot approach and landing capability evaluation based on feature fusion

YANG Yong^{1,2,3}, WU Guanchen^{1,3}, XU Kaijun^{1,2,3}, LI Donglin^{1,3}

(1. Flight Technology College, Civil Aviation Flight University of China, Guanghan 618307, China)

(2. Sichuan Provincial Engineering Research Center for Civil Aviation Flight Technology and Flight Safety, Guanghan 618307, China)

(3. Sichuan Provincial Engineering Research Center of Domestic Civil Aircraft Flight and Operation Support, Guanghan 618307, China)

Abstract: Current evaluations of pilots' approach and landing maneuvering capabilities mostly rely on single manual features, lacks a feature system that balances physical interpretability and complex temporal pattern capture. Moreover, the correlation mechanism between maneuvering behavior and flight quality remains unclear. Based on 105 sets of quick access recorder (QAR) data from Boeing 737-800 aircraft, this study constructs a two-stage, result-oriented evaluation system covering “final approach + landing flare” and introduces an energy management evaluation module. It integrates manual features and long short-term memory (LSTM) latent features to represent maneuvering behavior, and combines SHAP analysis, K-means clustering, and statistical tests to reveal how maneuvering features influence landing quality. Experiments show that the evaluation results achieve 90% consistency with manual assessments by airline experts, core features are significantly correlated with landing quality, and two typical maneuvering patterns are identified. The findings provide reliable data support for personalized pilot training and airline safety management.

Key words: QAR data; landing maneuvering capability; evaluation system; feature fusion; SHAP analysis; LSTM

收稿日期: 2026-04-02; 修回日期: 2026-06-23

基金项目: 国产民机飞行与运行支持四川省工程研究中心开放基金(MJCYZY202503); 2025 年度中央高校基本科研项目(26CAFUC01004, 26CAFUC01004, 25CAFUC04001); 四川省民航机场智慧运营与维护工程研究中心自主研究项目(JCZX2023ZZ07)

通信作者: 杨泳(1982-), 男, 博士, 副教授。E-mail: yangyong@cafuc.edu.cn

引用格式: 杨泳, 吴冠辰, 徐开俊, 等. 基于特征融合的飞行员进近着陆能力评价[J]. 航空工程进展.

YANG Yong, WU Guanchen, XU Kaijun, et al. Pilot approach and landing capability evaluation based on feature fusion[J]. Advances in Aeronautical Science and Engineering. (in Chinese)

0 引言

随着民航运输业快速发展,飞机自动化水平持续提升,但机组操纵失误仍是不安全事件的主要诱因^[1]。波音公司统计显示,2015-2024年47%的商业喷气式飞机致命事故发生在进近着陆阶段^[2],该阶段以飞行员手动操纵为主,其操纵行为直接决定飞行品质与运行安全。

近年来,机器学习算法在数据处理和模式识别方面展现了强大的能力,尤其是随机森林算法等算法兼具高效性和鲁棒性,常被用于分类、回归和特征选择研究。国内外研究者也逐步依托快速存取记录器(QAR)数据,分析飞行员的操纵行为,并建立对应量化评价体系。周鑫^[3]结合最后进近阶段对飞行员操作技能要求和关键胜任力形成飞行姿态控制、轨迹偏差监控与修正、轨迹安全管理和飞行指引系统监控4个评价维度、11项评价指标的评价体系,以QAR数据为基础运用时间序列检索算法完成指标的测量,并采用模糊综合评价理论构建评估模型;汪磊等^[4-5]基于QAR利用功率谱图对驾驶杆、方向舵的控制输入数据分析飞行员操作技能水平,依据平均功率频率、中值功率频率以及相应算法对飞行员操作输入的幅度和频率进行了评价,并针对飞行员着陆操作依据统计学原理量化出多个等级区间,综合考量关键着陆数据发生异常后所产生的风险。并以正态云相似度方法为核心,建立了一个对航线飞行员着陆操作风险进行评价的模型。实现了着陆操作风险评价指标向着陆操作风险等级的不确定性映射;叶亮^[6]使用主成分分析法(Principal Components Analysis, PCA)对QAR数据进行降维处理,使用特征空间表示的范数将多元时间序列数据转化为一维时间序列,使用距离函数评价时间序列数据的相似度,并提出基于偏离程度和偏离频率的飞行员安全绩效评价模型;刘胤甫^[7]引入EM-GMM((Expectation-Maximization Gaussian Mixture Model))算法聚类对起飞及着陆阶段飞行员驾驶杆操作整体特征进行了挖掘,并且将飞行操作风格与不安全事件预警进行结合,提高预测精度;祁明亮^[8]等人使用黄金分割法,利用QAR数据,获得对飞行超限事件的高风险数据区间的诊断和识别;曹海鹏^[9]利用神经网络自适应学习的特点,提升了对飞机重

着陆的诊断准确率和可靠性,并且展示了机器学习在飞行数据分析中的应用潜力;Cooper等^[10]首次系统化地将飞行员主观感受转化为工程可用的数据。使用量表的方式对飞行员的操纵水平进行评价,该量表后续也被大量应用于帮助工程师优化飞机操控特性;Wang^[11]等人选取了着陆距离、着陆垂直加速度和着陆俯仰角三个着陆参数作为评价飞行员着陆操作性能的指标,并建立了飞机着陆作业绩效评价体系。Puranik等^[12-13]分析了能量管理对于飞行安全的重要性,验证了能量指标在事后飞行数据分析场景中的潜在价值,其不仅能够量化飞行轨迹的能量特征,还可为事故风险溯源与安全决策提供量化支撑。Jesse C等^[14]应用聚类对飞行数据进行分析,使用聚类的方法获取飞行关键特征,并验证了方法的可行性;Payne^[15]以筛选出的12项关键技术指标为基础,构建飞行员操纵技能评价体系,从倾斜控制、升降控制、偏航控制、平衡能力、速度控制五大维度实现对操纵技能的多维度评估,并通过10个等级对行为参数的操纵水平进行技能分级;聂浩南等^[16]提出一种基于优劣解距离法(TOPSIS)辅助标注与改进多层感知机(Multi-Layer Perceptron, MLP)的自动化等级评估方法为飞行学员着陆阶段手动操纵能力进行评估研究,可为实际训练中海量数据提供快速准确的评估支撑;毛海涛等^[17]将飞行员抽象的操纵技能水平转化为飞行操纵量概率分布函数,基于蒙特卡洛仿真方法对飞行员进近操纵过程进行模拟仿真,实现了飞行员操纵技能对进近阶段飞行安全风险的定量评估。

现有研究存在三大不足:第一,飞行品质评价体系多混排操纵型指标与飞行结果型指标,缺乏针对性的结果性指标;第二,操纵特征提取多依赖单一人工特征或深度学习特征,难以兼顾可解释性与复杂模式捕捉;第三,操纵特征与飞行品质的关联性分析不够深入,未完成核心影响规律的量化。为此本文构建结果导向的两阶段多维度评价体系,新增能量管理能力指标;提出“人工特征+LSTM隐含特征”融合方案提取飞行员操纵特征,经完整性验证确保核心模式覆盖;建立多方法融合的关联性分析框架,系统性分析飞行员操纵特征对飞行安全品质指标的影响。

1 进近着陆品质评价体系构建

1.1 数据预处理

研究数据来源于国内某主流航空公司波音737-800机型的实际航班QAR数据,共105例样本。所有样本均执行一类盲降进近程序,涉及国内多个常规机场,研究剔除了复杂气象、特殊地形等干扰样本,保证数据集具备代表性与可比性。QAR数据包含驾驶盘、驾驶杆、油门杆等操纵参数,以及下滑道偏差、航向道偏差、高度、空速、航向、坡度等状态参数,采样频率为1 Hz。

为获取高质量的原始数据集,设计标准化预处理流程:目标区间截取,根据航班运行实际,截取最后进近阶段自动驾驶脱开(QAR数据中指示自动驾驶脱开)至着陆接地(空地电门指示地面)的飞行区间,该区间使飞行员手动完成进近、着陆动作的核心阶段;数据类别提取,从原始QAR数据中筛选出与操纵行为和飞行状态直接相关的13类核心参数(驾驶杆状态、驾驶盘状态、油门N1转速、无线电高度、气压高度、空速、目标空速、航向、下滑道偏差、航向道偏差、着陆载荷、平飘距离、坡度角),剔除低价值冗余数据,降低计算复杂度;数据清洗,采用线性插值法填充随机缺失值,确保时序数据的连续性,通过Z-score法($|Z| > 3$)识别异常值,结合飞行专业知识判定后采用线性插值修正,针对传感器采集噪声,采用Savitzky-Golay(SG)滤波进行平滑处理,辅以3点中心滚动平均作为兜底方案,确保数据保真。

1.2 评价指标选取

区别于现有评价体系,为实现评价体系与后续操纵特征分析模块解耦,本文在指标设计阶段刻意规避各类操纵型指标,仅选取能够反映飞行最终表现的结果型指标,搭建纯结果导向的两阶段评价体系,从根源上避免评价指标与后续操纵特征提取内容重叠冲突。

最后进近阶段聚焦轨迹控制与能量管理两大核心能力,选取水平轨迹控制能力(航向道偏差数据)、垂直轨迹控制能力(下滑道偏差数据)、能量管理能力(实际与理想总能量差值)3项结果性指标,如图1所示。

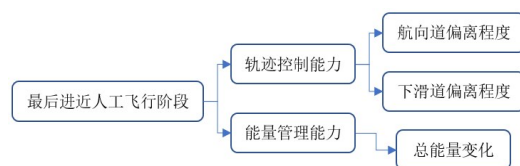


图1 最后进近阶段人工飞行评价体系

Fig. 1 Manual flight evaluation system in the final approach phase

能量管理能力指标通过计算飞机动能与重力势能的实际值与理想值差值获得。

$$E_{\Delta T} = \left(\frac{1}{2} m v^2 - \frac{1}{2} m v_{APP}^2 \right) + (mgH_{实际} - mgH_{理想}) \quad (1)$$

式中: $E_{\Delta T}$ 为飞机实际与理想总能量的差值; v 为飞机实际空速; v_{APP} 为飞机的理想进近速度; $H_{实际}$ 为飞机实际高度; $H_{理想}$ 为飞机理想高度; m 为飞机质量; g 为重力加速度。

理想下滑角度设为 3° (民航盲降进近通常下滑角度),以此角度与实际下滑角度的差值与飞机实际飞行高度根据三角函数关系获得飞机理想高度。理想空速采用QAR数据记录的目标进近速度 v_{APP} 。

着陆拉平阶段结合飞行安全要求与飞行知识,选取下沉控制能力、着陆载荷、平飘距离、接地位置偏差、接地方向偏差、接地坡度偏差6项结果性指标,全面反映飞机接地瞬间的安全性能与操纵精准度,如图2所示。



图2 着陆拉平阶段人工飞行评价体系

Fig. 2 Manual flight evaluation system in the landing flare phase

1.3 评价指标处理与评分区间划分

针对最后进近阶段时序数据,研究未采用DTW动态时间规整、欧式距离等传统曲线相似性分析方法。研究从时域、频域两个维度,选取26维具有强物理意义的数据特征(均值、方差、标准差、峰峰值、主导频率等),并通过计算各特征贡献度的方法赋权,用以刻画整体时序数据规律。

计算数据特征对数据的贡献率为无监督的计算模式,不适合使用LightGBM、随机森林等机器

学习模型,因此使用方差贡献、PCA 重构误差、特征冗余度的综合统计学方法,从冗余程度、信息保留、还原精度 3 个互补维度进行特征重要性计算,最终使用 CRITIC 客观赋权法对三种方法计算的贡献度数值进行客观赋权最终获得特征的重要性。按照重要性排名,提取累计贡献度超过 95% 的核心特征作为评价子指标。

以能量管理能力指标为例,如图 3 所示。图 3 中横坐标的特征名用于表示飞行员的能量管理能力数据规律且作为评价子指标,并依据各数据特征的贡献率为各个评价子指标赋权,最终获得时序数据评价指标最终得分情况。

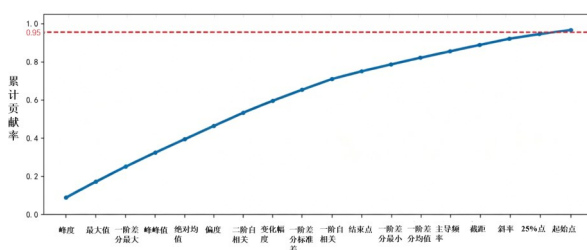


图3 能量管理能力核心特征累计贡献率图
Fig. 3 Cumulative contribution rate chart of core characteristics of energy management capability

着陆拉平阶段指标中下沉控制能力用于评价飞行员对拉平阶段下沉控制均匀减少的程度,因此选取飞机 50 ft 至接地区间内无线电高度变化的速度衰减系数用于本文研究,如式(2)所示。

除此之外,其余指标均使用飞机着陆接地时的单值数据,可直接从 QAR 数据中提取或简单差值计算得到,例如接地位置偏差指标使用飞机接地时的航向道偏差数据绝对值反映飞机位置偏离情况。

$$\lambda_{i,j} = -\frac{\ln\left(\frac{h_i}{h_j}\right)}{t_j - t_i} \quad (2)$$

式中: $\lambda_{i,j}$ 为飞行员下沉控制的指数衰减系数; h_i 为第*i*个时刻的无线电高度; t_i 为第*i*个时刻的时间; h_j 为第*j*个时刻的无线电高度; t_j 为第*j*个时刻的无线电高度。

评分区间划分方法综合考虑数据本身的最优聚类簇数与文章可解释性,采用 K-means++ 聚类算法设置簇为 5,划分 5 分制评分区间(5 分为优

秀,1 分为待提升),并通过轮廓系数验证聚类数 $k=5$ 时所有指标的聚类轮廓系数均大于 0.5,表明聚类效果良好。以能量管理能力指标为例,如图 4 所示。图中纵坐标为轮廓系数,横坐标为评价子指标的名称。

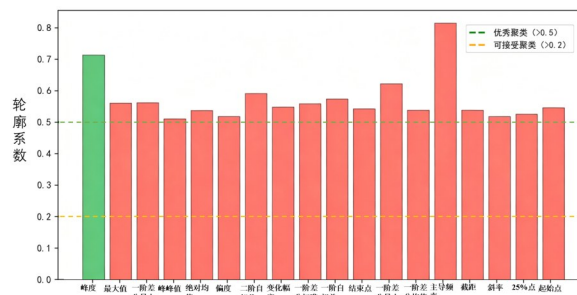


图4 能量管理能力核心特征聚类效果图
Fig. 4 Clustering effect diagram of core characteristics of energy management capability

以下沉控制能力指标为例,评分区间的聚类效果图,如图 5 所示。每一簇代表一类分值,按照指标极大值或极小值特性,根据簇的分布进行打分。在各评价指标或子指标中,仅峰度这一子指标为极大值指标,其余指标均为极小值指标,即峰度越大表明数据越集中于 0,因此峰度越大数据偏离越少,表明飞行越稳定。

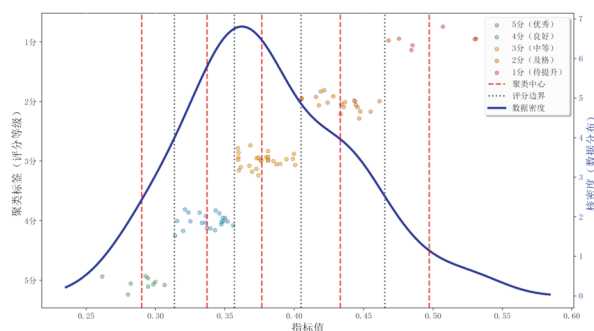


图5 下沉控制能力数据聚类效果图
Fig. 5 Data clustering effect diagram of descent control capability

依据上述评价指标选择和评分区间划分方法,对 QAR 数据样本中提取出的各评价指标于评价子指标进行评价打分。着陆拉平阶段评价指标评分直接输出,最后进近阶段的评价指标得分由评价子指标评分赋权计算并输出最终获得 9 维评分矩阵。

2 操纵特征提取与融合

2.1 人工特征提取

结合飞行理论知识,构建“时域+频域+跨通道协同”三维人工特征体系,共提取40维特征:时域特征,包括操纵幅度(平均幅度、最大幅度)、修正频率、方向切换率、平滑度、平均修正间隔等,反映操纵强度与稳定性;频域特征,通过 Welch 功率谱估计获取主频率与频率能量集中度,刻画操纵动作的主导振荡频率与能量分布;跨通道协同特征,包括横侧俯仰同步性、俯仰能量协同性等,描述驾驶盘、驾驶杆、油门杆的联动控制模式。所有人工特征均具有明确的物理意义,可直接关联飞行员操纵行为。

2.2 LSTM 隐含特征提取

为捕捉人工特征难以覆盖的复杂时序关联,设计轻量化 LSTM 模型提取隐含特征:1)模型结构,输入层接收驾驶盘、驾驶杆、油门杆的变长时序数据,经 Masking 层忽略填充值(填充值设为-999.0),LSTM 层(64个单元)返回全序列特征,通过全局平均池化层将变长序列转化为固定维度向量,最后经全连接层(ReLU 激活函数)输出64维隐含特征;2)训练设置,采用 Adam 优化器(学习率0.001),损失函数为均方误差(MSE),无监督训练方式避免对标注数据的依赖,确保特征提取的客观性。

2.3 特征筛选与完整性验证

将40维人工特征与64维 LSTM 隐含特征横向拼接,形成104维融合特征矩阵。基于 XGBoost 回归模型拟合融合特征集合与9项人工飞行评价指标的映射关系,并计算各特征对9项评价指标的博弈论的机器学习模型解释工具(SHAP)值,筛选每个评价指标的前20项核心特征,合并去重后得到核心特征集。

为验证人工特征的完整性,设计对比实验:分别采用纯人工特征集、融合特征集构建 XGBoost 模型,开展10轮交叉验证,对比两组模型的决定系数(R^2)与均方误差(MSE),并进行配对 t 检验($\alpha=0.05$),完成显著性分析。

2.4 统计关联性分析

计算核心特征与评价指标之间的 Pearson 线性相关系数(反映线性关联)与 Spearman 秩相关系数(反映单调关联),以 $|r| \geq 0.1$ 且 $p < 0.05$ 为有效关联阈值,筛选具有统计意义的特征-指标组合。同时,通过计算 $|\text{Pearson } r - \text{Spearman } r_s| \geq 0.2$ 识别强非线性关联,为后续分析提供依据。

采用 K-means 算法对标准化后的核心特征集进行无监督聚类。结合聚类中心特征值与飞行专业知识,分析各类操纵模式的典型特征,揭示不同操纵风格的差异。

基于训练好的 XGBoost 模型,计算核心特征的 SHAP 值,绘制 SHAP 依赖图,可视化特征值变化对着陆品质评分的影响规律,捕捉阈值效应与最优区间。针对 LSTM 隐含特征,采用相关性回溯法进行解码:计算 LSTM 特征与原始操纵时序数据的 Pearson 相关系数,选取关联性最强的时序数据类别与时间段,通过分组时序模式对比,将“黑箱”特征转化为可解释的物理行为。

3 实验过程

3.1 评价体系有效性验证

根据105例样本的评价结果显示,各指标评分呈合理正态分布,如图6所示。图6展示最后进近阶段105个样本中各78维(26维 $\times$ 3维评价指标)的评价子指标的评分情况,着陆拉平阶段随机3个样本数据的得分情况如表1所示。水平轨迹控制能力平均评分3.2分(标准差0.85),垂直轨迹控制能力平均评分3.1分(标准差0.79),能量管理能力平均评分2.9分(标准差0.92),着陆拉平阶段6项指标平均评分3.0分(标准差0.81)。

基于 Likert 5级量表设计问卷,将模型评价结果与航司资深飞行员主观评估结果进行对比,内容包含评价指标、体系的设计与选择及可视化 QAR 样本等相关问题。邀请10名资深现役飞行员进行主观评价,回收有效问卷10份。结果表明:8名飞行员认可本文指标与评价体系,9名的主观打分与模型计算结果一致,表明该评价体系能够客观反映飞行员的实际操纵能力水平,验证模型的可靠性。

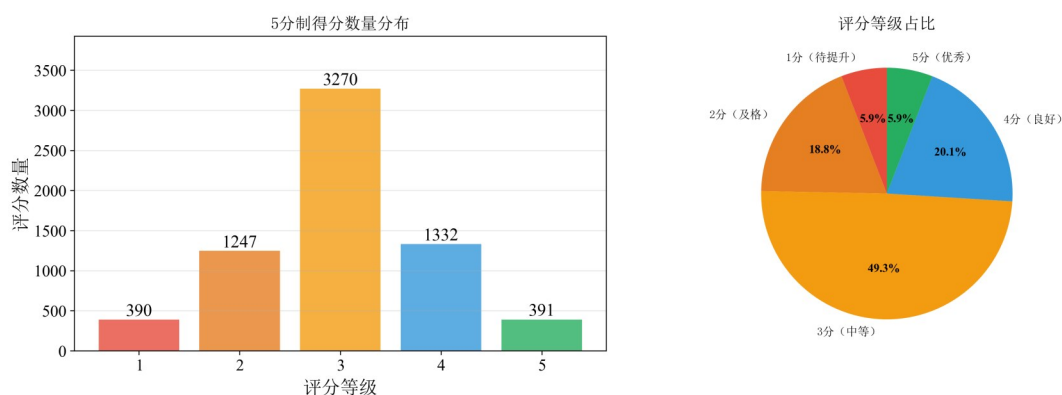


图6 最后进近阶段样本间全部78维特征评价子指标得分

Fig. 6 Scores of all 78-dimensional feature evaluation sub-indicators among samples during the final approach phase

表1 样本着陆拉平单值数据评分情况
Table1 Scoring of single-value data for sample landing flare

样本编号	坡度偏差	着陆载荷	平飘距离	下沉控制	位置偏差	航向偏差
1	4	3	4	5	4	3
2	2	2	2	5	5	3
3	1	4	3	3	1	5

3.2 操纵特征筛选结果

基于XGBoost机器学习模型,计算评分矩阵与融合操纵特征集之间的SHAP值,用以描述飞行员操纵特征对于飞行员人工飞行评价指标的映射关系,以筛选核心特征,实现特征降重。以水平轨迹控制评分为例,该评价指标的特征SHAP值摘要图如图7所示,SHAP摘要图可清晰展示各特征的SHAP值分布的离散程度,离散程度越大表明特征对预测结果的影响波动越明显。

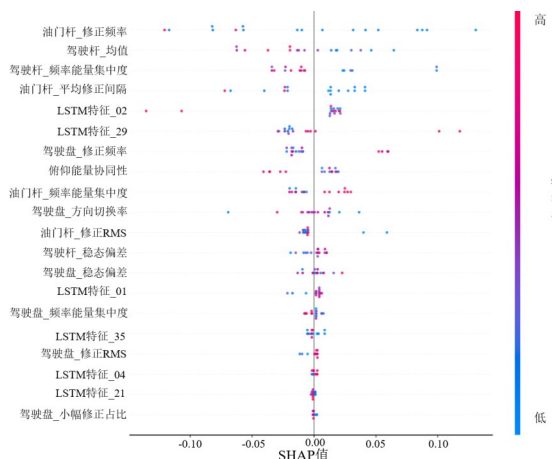


图7 水平轨迹控制能力特征SHAP值摘要图

Fig. 7 SHAP value summary plot of horizontal trajectory control capability features

各评价维度Top10特征SHAP重要性热力图如图8所示,用于反映飞行员操纵特征与各评价指标的关联性。并将9个评价指标的Top20特征集合进行合并去重,形成共85维的综合核心特征集。



图8 各评价维度Top10特征SHAP重要性热力图

Fig. 8 SHAP importance Heatmap of Top 10 features in each evaluation dimension

SHAP特征重要性分析表明,85维核心特征中,人工特征40个(占47.1%),LSTM特征45个(占52.9%),如图9所示。通过SHAP重要性分析,可有效区分冗余特征与核心特征,剔除对所有评价维度贡献度极低的特征,降低后续模型训练的计算复杂度,同时保留与着陆品质评价紧密相关的核心特征,为后续特征完整性验证与着陆品质评价模型优化奠定基础。

对上述核心特征集合进行完整验证,特征完整性检验是高维融合特征工程的关键环节,其核心目的是验证40维人工操纵特征与64维LSTM深度特征融合后,相较于单一人工特征集,是否能显著提升着陆品质各评价维度的预测性能,同时验证筛选后核心特征集的有效性、稳定性与冗余

剔除合理性,并验证所设计的人工操纵特征是否已完整覆盖着陆操纵的核心运动模式,开展特征完整性验证。

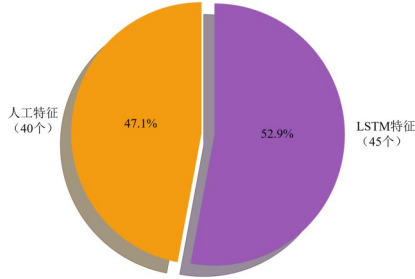


图9 核心特征集合人工特征与LSTM特征占比

Fig. 9 Proportion of manual features and lstm features in the core feature set

开展10轮独立验证实验,每轮实验均记录模型的决定系数(R^2)与均方误差(MSE)两个核心性能指标。结果显示,两组模型的 R^2 与MSE差异不显著($p>0.05$),表明融合的飞行员操纵特征对样本的预测效果进行了补充,但人工特征已完整覆盖着陆操纵的核心模式。这并不说明融合后特征集没有意义,后续研究表明LSTM特征主要捕捉复杂非线性关联,为关联性分析提供补充,提高模型的可解释性。各维度评价指标显著性提升图如图10所示。

从统计学的角度,对操纵特征—评价指标关联性进行分析,用于补充解释前门基于XGBoost

模型进行SHAP分析的局限性。机器学习模型的预测效果不完全等同于统计学中关联性,模型过度依赖机器学习模型,可能出现过拟合以及导致筛选出的特征只对XGBoost模型预测重要,而不一定真实重要。

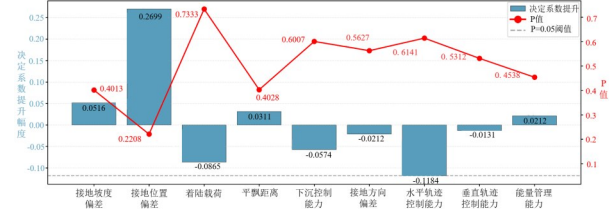


图10 各维度评价指标显著性提升图

Fig. 10 Significance improvement chart of evaluation indexes in each dimension

通过计算Pearson系数与Spearman系数并进行显著性 p 值检验,并对特征计算结果进行汇总,如图11所示(仅展示部分指标,其中下沉对应下沉控制评分,位置对应位置偏差评分,载荷对应着陆载荷评分,平飘对应平飘距离评分,坡度对应坡度偏差评分,方向对应方向偏差评分)。

结果显示,765对操纵特征—评价指标组合中,有效关联组合达712对,占总组合的93.1%,其中强线性关联($|r|\geq 0.3$)153对,中等关联($0.2\leq |r|<0.3$)204对,弱关联($0.1\leq |r|<0.2$)251对,验证了特征筛选的有效性。

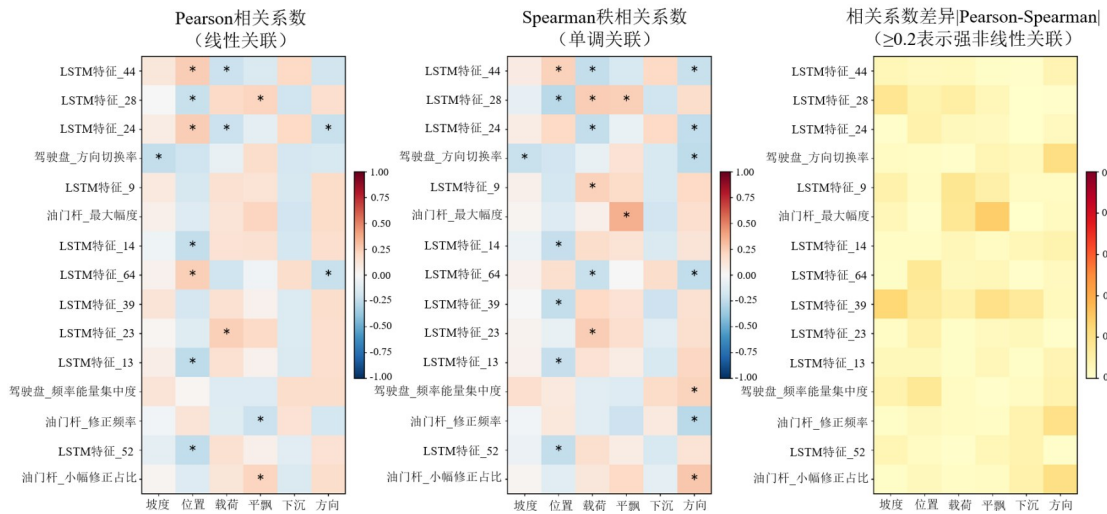


图11 部分操纵特征与评价指标的统计学对比相关性分析热力图

Fig. 11 Heatmap of statistical comparative correlation analysis between partial control characteristics and evaluation indexes

统计学分析关联性Top5的核心特征及其关联强度如表2所示。这些核心特征涵盖了驾驶盘、

驾驶杆、油门杆的关键操纵参数,为后续关联性分析提供了重点对象。

表2 Top5特征关联强度排名表
Table 2 Ranking of Top5 feature correlation strengths

排名	特征名称	特征类型	平均 $ r $	最强关联评价指标	最强 $ r $	有效指标数
1	驾驶杆—修正频率	人工特征	0.214	着陆载荷	0.287	2
2	油门杆—平均幅度	人工特征	0.206	能量控制能力	0.276	2
3	LSTM特征—63	LSTM特征	0.201	水平轨迹控制能力	0.271	2
4	驾驶盘—修正频率	人工特征	0.198	平飘距离	0.268	1
5	驾驶杆—平滑度	人工特征	0.195	接地坡度偏差	0.260	2

3.3 操纵模式识别结果

通过轮廓系数确定 K-means 聚类的最优聚类数 $k=2$, 识别出两类典型操纵模式, 如图 12。其特征差异与占比情况如下:

1) 激进型操纵模式, 占比 25.9% (22 例样本), 典型特征为驾驶杆修正频率高 (1.85 倍于平稳型)、平滑度低 (-2.0)、油门杆稳态偏差大 (0.34)、驾驶盘方向切换率高 (1.17 倍于平稳型), 横侧-俯仰协同性差, 操纵动作频繁且幅度大;

2) 平稳型操纵模式, 占比 74.1% (63 例样本), 典型特征为修正频率适中、操纵平滑度高、油门杆控制稳定、驾驶盘方向切换少, 横侧控制精准, 操纵动作连贯且幅度适中。

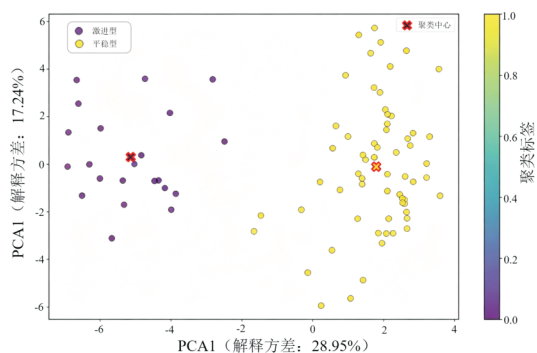


图 12 操纵模式聚类结果可视化 ($k=2$)

Fig. 12 Visualization of control mode clustering results ($k=2$)

虽然两种操纵模式的操纵风格具有显著差异, 但两类模式各样本的评分结果对比显示, 平稳型模式的平均评分相较于激进型模式平均评分并无显著差异。分析原因为本研究所选取的样本并未突破安全阈值, 且聚焦于正常航班数据。因此在正常航班数据中操纵风格对飞行安全品质影响不明显。但航司出于乘客乘坐舒适性、飞行风险管控的角度, 更倾向于培养操纵风格平稳的飞

行员。

3.4 关联性分析结果

经过前序研究, 分别从机器学习和统计学角度, 得到了关联性研究结果。

其中, 以部分核心特征与评价指标的统计显著关联结果为例, 显示如下: 驾驶杆修正频率与着陆载荷呈显著负相关 ($r=-0.287$), 表明驾驶杆修正越频繁, 着陆载荷越大, 不利于飞行安全与乘坐舒适性; 油门杆平均幅度与能量控制能力呈显著正相关 ($r=0.276$), 适当的油门杆调整幅度有助于维持飞机能量平衡; 驾驶盘方向切换率与接地方向偏差呈显著负相关 ($r=-0.235$), 减少驾驶盘频繁方向切换可改善接地方向精度; 驾驶杆平滑度与接地坡度偏差呈显著正相关 ($r=0.260$), 操纵动作越平滑, 接地坡度控制越精准。

此外, 识别出 104 对强非线性关联组合 (占有有效关联的 14.6%), 其中 LSTM 特征占 72.1%, 人工特征占 27.9%, 主要包括阈值型、指数型与倒 U 型三种关联类型, 表明 LSTM 对数据非线性特征的捕捉效果显著。

通过 SHAP 依赖图将非线性关联特性进行可视化操作, 进一步揭示核心特征的影响机制, 如图 13 所示, 驾驶杆_小幅修正占比呈现强负向影响着陆载荷, 随着占比从 0.1 增加到 0.8, SHAP 值从 0.4 下降到 -0.1。这表明低比例的小幅修正能有效降低着陆载荷; 如图 14 所示, 驾驶杆_平均幅度呈现倒 U 型关系与能量管理能力评分呈现微弱负相关, 在 0.04~0.08 区间, 在数值超过 0.10 后正向影响显著增长。

此外, 横侧俯仰同步性存在临界点 0.1, 超过该值后对着陆载荷的负向影响显著增强; 油门平均幅度的最优区间为 0.2~0.4, 过大或过小均会降低能量管理能力评分。

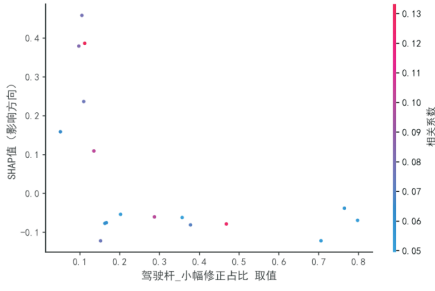


图13 驾驶杆_小幅修正与着陆载荷 SHAP 依赖图

Fig. 13 SHAP dependence plot of control column: small corrections and landing load

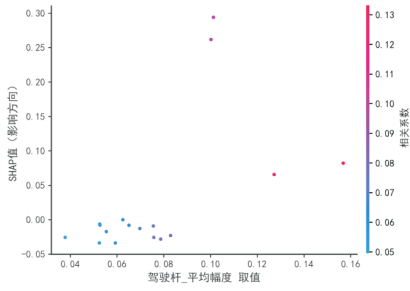


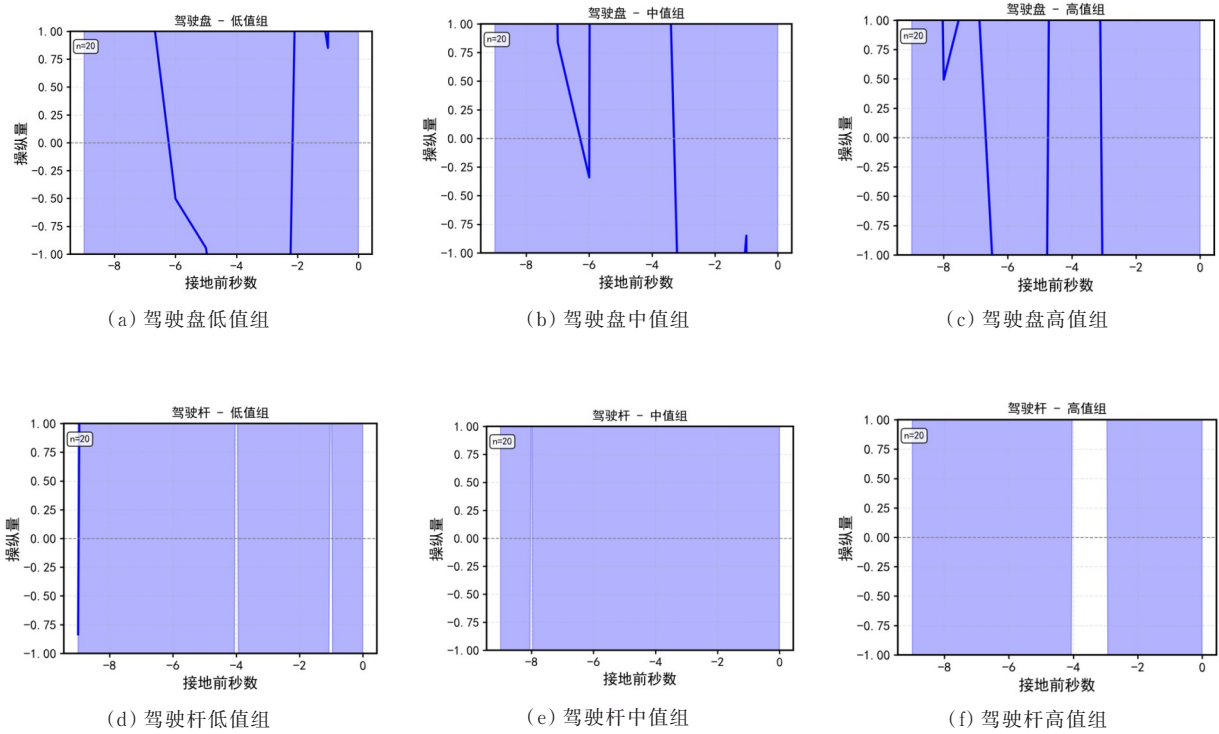
图14 驾驶杆_平均幅度与能量管理能力 SHAP 依赖图

Fig. 14 SHAP dependence plot of control column: average amplitude and energy management capability

最后,对深度学习的 LSTM 隐含特征进行解码,进一步将具有“黑箱”性的 LSTM 特征与评价指标的关联性进行研究。

以 LSTM 特征_63 为例,结合特征与原始操纵数据的 SHAP 值,划分低、中、高值组进行对照,如图 15 所示,并根据上述相关性检验结果显示,该特征与驾驶盘操纵数据强相关(平均 $|r|=0.072$),重点捕捉横测方向的非线性特征,解码结果显示该特征主要编码横侧操纵的修正频率与幅度:低值组,驾驶盘在接地前 6 s 出现一次大幅偏转,随后回弹,表现为单次剧烈横侧修正;中值组,驾驶盘在接地前 8 s 和 -2 s 出现两次中等幅度偏转,操纵精准性一般;高值组,驾驶盘在接地前多次出现剧烈偏转,横侧操纵最不稳定,航向偏差风险显著升高。

该解码结果为理解 LSTM 隐含特征的物理意义提供了明确依据。



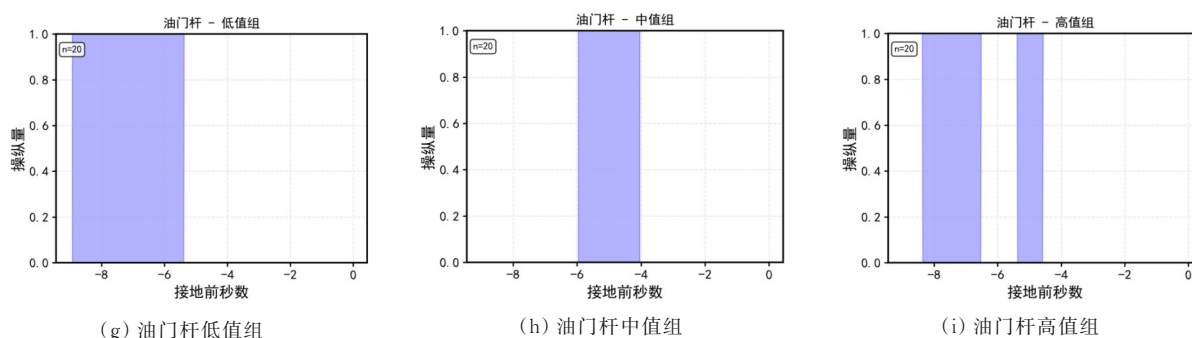


图15 LSTM特征_63分组时序模式可视化图

Fig. 15 Visualization of LSTM feature_63 grouped time series patterns

结合飞行力学理论与实验结果,关联性较强的部分核心特征的影响机制可解释如下:驾驶杆修正频率过高反映飞行员对俯仰姿态控制不精准,存在过量修正现象,导致飞机垂直方向受力波动增大,进而使着陆载荷升高;油门杆平均幅度在最优区间内时,飞行员能够通过适度调整油门维持飞机能量平衡,避免能量过高或过低导致的着陆困难;驾驶盘方向切换率过高会造成飞机横向姿态频繁波动,破坏水平轨迹稳定性,导致接地方向偏差增大;驾驶杆平滑度高表明操纵动作连贯柔和,能够有效控制飞机坡度变化,减少接地时的坡度偏差。这些机制揭示了飞行员操纵行为与着陆品质之间的内在联系,为飞行训练提供了明确的优化方向。

4 结 论

1) 构建的“最后进近+着陆拉平”两阶段评价体系具有良好的有效性与可靠性,与航司人工评估一致性达90%,新增的能量管理能力指标能够有效反映飞行员的能量控制水平,弥补了传统评价体系的短板;

2) “人工特征+LSTM隐含特征”融合方案兼顾物理可解释性与复杂模式捕捉能力,经完整性验证证实人工特征已覆盖核心操纵模式,筛选得到的85维核心特征与着陆品质高度相关(有效关联占比93.1%);

3) 多方法融合的关联性分析框架成功揭示了操纵特征相对于结果性评价指标的影响机制,并实现了LSTM隐含特征的物理解码,为飞行训练提供了客观数据支撑。

4) 识别出激进型和平稳型两类典型操纵模式,平稳型模式的着陆品质整体更优;驾驶杆修正频率、油门杆平均幅度、驾驶盘方向切换率等核心特征对着陆品质具有显著影响,且存在明确的阈值效应与最优区间。后续研究可进一步结合飞行工程实际对CRITIC赋权权重进行验证,并讨论与样本数据统计特征相关性;同时,多维特征集选取可进行共性和冗余性检测,选取更优特征集合。

基于QAR数据的飞行品质评价与操纵特征研究是民航数字化、智能化发展的重要方向,本研究的框架与成果为后续研究提供了参考,未来随着研究的深入与技术的迭代,相关成果将持续服务于航司飞行品质监控与飞行员训练优化,助力民航飞行安全治理能力的持续提升。

参考文献

- [1] 中国民航总局人为因素课题组. 民用航空人的因素培训手册[M]. 北京: 中国民航出版社, 2003.
Human Factors Research Group of the General Administration of Civil Aviation of China. Human factors training manual for civil aviation[M]. Beijing: China Civil Aviation Publishing House, 2003. (in Chinese)
- [2] Boeing. Statistical summary of commercial jet airplane accidents 2015-2024 [R]. Seattle: Boeing Commercial Airplanes, 2025.
- [3] 周鑫. 面向胜任力的飞行员手动操纵品质评价方法研究[D]. 广汉: 中国民用航空飞行学院, 2024.
Zhou Xin. 面向胜任力的飞行员手动操纵品质评价方法研究[D]. Guanghan: Civil Aviation Flight University of China, 2024. (in Chinese)
- [4] 汪磊, 董传亭, 崔美叶. 基于功率谱的飞行操作技能评价方法[J]. 飞行力学, 2018(4): 92-96.
Wang Lei, Dong Chuanting, Cui Meiyue. Evaluation method

- of flight operation skills based on power spectrum[J]. *Flight Dynamics*, 2018(4): 92-96. (in Chinese)
- [5] 汪磊, 郭世广, 任勇. 基于飞行数据正态云的着陆操作风险评价方法[J]. *安全与环境学报*, 2019, 19(5): 1555-1561.
- Wang Lei, Guo Shiguang, Ren Yong. Landing operation risk evaluation based on the normal cloud of the flight data [J]. *Journal of Safety and Environment*, 2019, 19(5): 1555-1561. (in Chinese)
- [6] 叶亮. 基于运行数据的飞行员安全绩效评估研究[D]. 广汉: 中国民用航空飞行学院, 2022.
- Ye Liang. Pilot safety performance evaluation based on operational data[D]. Guanghan: Civil Aviation Flight University of China, 2022. (in Chinese)
- [7] 刘胤甫. 基于QAR数据的飞行员操作风格识别方法研究[D]. 天津: 中国民航大学, 2023.
- Liu Yinfu. Research on pilot operation style recognition method based on QAR data [D]. Tianjin: Civil Aviation University of China, 2023. (in Chinese)
- [8] 祁明亮, 邵雪焱, 池宏. QAR超限事件飞行操作风险诊断方法[J]. *北京航空航天大学学报*, 2011, 37(10): 1207-1210.
- Qi Mingliang, Shao Xueyan, Chi Hong. Flight operations risk diagnosis method on quick-access-record exceedance [J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2011, 37(10): 1207-1210. (in Chinese)
- [9] 曹海鹏, 舒平, 黄圣国. 基于神经网络的民用飞机重着陆诊断技术研究[J]. *计算机测量与控制*, 2008, 16(7): 906-908.
- Cao Haipeng, Shu Ping, Huang Shengguo. Study of aircraft hard landing diagnosis based on neural network[J]. *Computer Measurement & Control*, 2008, 16(7): 906-908. (in Chinese)
- [10] Cooper G E, Harper R P. The use of pilot rating in the evaluation of aircraft handling qualities, NASA TN-D-5153 [R]. Moffett Field, CA: NASA Ames Research Center, 1969.
- [11] Wang L, Zhang J Y, Dong C T, et al. A method of applying flight data to evaluate landing operation performance[J]. *Ergonomics*, 2019, 62(2): 171-180.
- [12] Puranik T, Jimenez H, Mavris D. Energy-based metrics for safety analysis of general aviation operations[J]. *Journal of Aircraft*, 2017, 54(6): 2285-2297.
- [13] Puranik T G, Mavris D N. Anomaly detection in general-aviation operations using energy metrics and flight-data records [J]. *Journal of Aerospace Computing, Information, and Communication*, 2018, 15(1): 22-35.
- [14] Jesse C, Liu H H, Smart E, et al. Analysing flight data using clustering methods [M] // *Knowledge-based intelligent information and engineering systems*. Berlin, Heidelberg: Springer, 2008: 733-740.
- [15] Payne K H, Harris D. A psychometric approach to the development of a multidimensional scale to assess aircraft handling qualities[J]. *The International Journal of Aviation Psychology*, 2000, 10(4): 343-362.
- [16] 聂浩南, 钱宇, 李秀易, 等. 基于改进多层感知机的飞行学员着陆阶段手动操纵能力评估研究[J/OL]. *航空工程进展*, 1-10 [2026-07-02]. <https://link.cnki.net/urlid/61.1479.V.20260331.1248.004>.
- Nie Haonan, Qian Yu, Li Xiuyi, et al. Improved MLP-based assessment for pilot manual control performance [J/OL]. *Advances in Aeronautical Science and Engineering*, 1-10 [2026-07-02]. <https://link.cnki.net/urlid/61.1479.V.20260331.1248.004>. (in Chinese)
- [17] 毛海涛, 张柱柱, 胡文林, 等. 飞行员操纵技能对进近阶段飞行安全影响研究[J/OL]. *航空工程进展*, 1-9 [2026-07-02]. <https://link.cnki.net/urlid/61.1479.V.20250908.1128.008>.
- Mao Haitao, Zhang Zhuzhu, Hu Wenlin, et al. Research on the impact of pilot control skills on flight safety during approach stage[J/OL]. *Advances in Aeronautical Science and Engineering*, 1-9 [2026-07-02]. <https://link.cnki.net/urlid/61.1479.V.20250908.1128.008>. (in Chinese)

(编辑:丛艳娟)