

基于 BiGRU-MHA-KAN 的飞行训练 轨迹预测方法

张迪¹, 魏雪倩², 黎思杰²

(1.中国民航大学 工程技术训练中心, 天津 300300; 2.中国民航大学 电子信息与自动化学院, 天津 300300)

摘要:为提升飞行训练中轨迹预测的精确度,优化预测模型的可靠性,提出了一种融合双向门控循环单元(BiGRU)、多头注意力机制(MHA)与 Kolmogorov-Arnold 网络(KAN)的深度混合神经网络模型 BiGRU-MHA-KAN。该模型通过航迹数据预处理与重构,结合双向建模、注意力机制和 KAN 网络,强化时序特征提取与非线性动态建模。仿真实验系统分析了不同参数设置及历史数据量对模型性能的影响。结果表明,相较于其他轨迹预测模型,所提方法将预测准确提高了 4.81%~5.83%,且均方误差与均方根误差显著降低,在飞行训练场景下展现出更强的时序建模能力和稳定性。

关键词:轨迹预测;飞行训练;KAN 网络;双向门控循环单元;注意力机制

中图分类号:TP183; V271.6

文献标志码:A

文章编号:1000-2758(2026)01-0185-09

随着我国航空运输业的快速发展,行业对飞行员需求不断增长,飞行员的规范化和高水平训练成为保障航空安全的关键环节。目前,经验驱动型训练模式依赖于教员主观经验与仪表参数的离散化反馈,无法实现对学员飞行轨迹的实时预测和量化评估。传统轨迹预测训练模式主要基于数学模型和统计分析,在应对动态复杂的飞行场景时存在显著局限性。随着广播式自动相关监视技术全面覆盖,基于机器学习的轨迹预测方法得到了广泛关注。然而,该方法对于高质量标签数据的需求量较大,且在数据不足或特征复杂时预测模型性能下降^[1]。

为解决上述问题,深度学习方法提供了新的思路和路径。张宏鹏等^[2]利用门控循环单元(gated recurrent unit, GRU)预测轨迹。付茂洺等^[3]根据时间序列飞行数据的特点利用长短期记忆网络(long short-term memory, LSTM)建立预测模型。Hu 等^[4]提出一种结合时间序列卷积神经网络和 LSTM 的混合模型,预测高速飞机的位置和速度。方伟等^[5]结

合 LSTM 和残差神经网络预测轨迹。张宗腾等^[6]通过双向门控循环单元(bidirectional gated recurrent unit, BiGRU)网络提高轨迹信息利用率。然而,该网络在处理高度复杂的非线性数据时,仍存在一定的局限性。本研究提出 BiGRU-MHA-KAN 三重机制架构,通过 BiGRU 网络实现时空特征融合;结合多头注意力机制(multi-head attention, MHA)强化关键特征捕获能力;引入科尔莫哥罗夫-阿诺德(Kolmogorov-Arnold networks, KAN)网络^[7],建立非线性映射通路,有效解决复杂轨迹建模中的非线性表征难题,提升飞机轨迹预测精度。

1 基础理论研究

1.1 BiGRU 网络

鉴于飞行轨迹同时受历史状态与未来意图影响的特性,本研究采用 BiGRU 网络提取时序特征。该网络通过并行的正、反向 GRU 模块,分别捕获历史演进规律与潜在趋势特征,更全面地提取长序列数据的上下文关联信息^[8],同时保持 GRU 网络参数量少、训练效率高等优势。

更新门 z_t

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad (1)$$

收稿日期:2025-05-14

基金项目:国家自然科学基金(U1733119)与民航安全能力建设基金([2023]50)资助

作者简介:张迪(1979—),副教授

通信作者:张迪(1979—) e-mail:dzhang@cauc.edu.cn

重置门 r_t

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad (2)$$

候选隐藏状态 \tilde{h}_t

$$\tilde{h}_t = \tanh(W_h \cdot [r_t \cdot h_{t-1}, x_t] + b_h) \quad (3)$$

最终隐藏状态 h_t

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \quad (4)$$

z_t 调节 h_{t-1} 的保留比例和 \tilde{h}_t 的吸收程度, r_t 控制对 h_{t-1} 的影响程度。输入特征 x_t 经 r_t 筛选历史信息后, 与当前输入共同生成 \tilde{h}_t 。 h_t 通过 z_t 对 h_{t-1} 和 \tilde{h}_t 实现加权融合。模型参数包含可训练权重矩阵 (W_r, W_z, W_h) 及偏置向量 (b_r, b_z), 通过反向传播优化。激活函数配置为: σ (Sigmoid) 实现门控值 (0, 1) 归一化, \tanh 将候选状态约束至 (-1, 1) 区间。

第 n 层的正向隐藏状态 $\tilde{h}_1^n, \tilde{h}_2^n, \dots, \tilde{h}_T^n$ 与反向隐藏状态 $\tilde{h}_1^n, \tilde{h}_2^n, \dots, \tilde{h}_T^n$ 分别按时间正序和逆序计算。最终通过张量拼接实现时空特征表征

$$h_t^{\text{BiGRU}} = [\tilde{h}_t^n; \tilde{h}_t^n] \quad (5)$$

1.2 MHA 机制

针对传统 BiGRU 模型时序建模中存在的特征均等化缺陷, 本研究提出多头注意力增强架构。该架构采用特征空间解耦策略, 将输入序列通过 3 个

可学习权重矩阵 W^Q, W^K 和 W^V 经线性变换形成查询矩阵 Q 、键矩阵 K 和值矩阵 V , 沿特征维度拆分为 h 个正交子空间。在每个子空间中进行自注意力计算时, 通过缩放点积运算建立元素间关联。

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (6)$$

式中: d_k 为键向量的维度, 用于调节点积值的量纲。

将 h 个独立子空间的注意力输出进行特征融合

$$H_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \quad (7)$$

$$\text{Multihead}(Q, K, V) = \text{concat}(H_1, H_2, \dots, H_h)W^O \quad (8)$$

式中: W^O 为输出投影矩阵; $1 \leq i \leq h, h$ 表示子空间个数; concat 表示沿特征维的拼接操作。

1.3 KAN 网络

针对飞行训练轨迹的多参数输入高维非线性映射问题, 引入基于 Kolmogorov Arnold 表示定理^[9] 的新型 KAN 网络。如图 1 所示, 该网络依托定理核心特性^[10], 通过可学习样条基函数替代静态权重参数, 结合自适应函数组合策略, 以较少参数实现高维非线性关系的高效拟合。

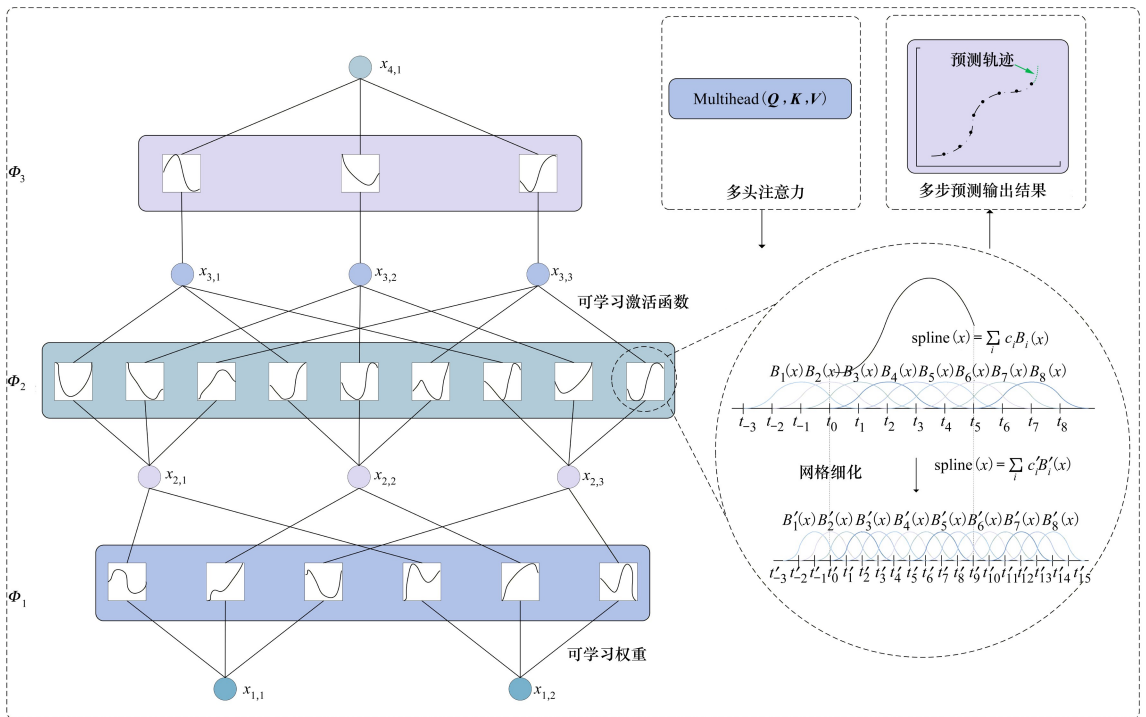


图 1 KAN 网络结构示意图

该网络的优越性已在多研究领域得到验证: 袁

立宁等^[11] 将其引入图深度学习领域, 有效规避了图

卷积网络对显式信息传递策略的依赖。Al-Qaness 等^[12]借助该网络提高了肌电信号分类精度。Ren 等^[13]采用样条函数网格细化策略,增强了函数逼近精度。Sidharth 等^[14]将逼近论的理论基础与切比雪夫多项式相结合,成功捕捉了分形结构中的复杂模式与变化。

KAN 网络是由内外函数组成的双层级联结构:内部函数层实现 n 维输入到 $2n + 1$ 维输出的非线性变换,外部函数层将 $2n + 1$ 维特征映射至 n 维输出空间。公式为

$$\text{KAN}(\mathbf{x}) = f(x_1, x_2, \dots, x_n) = \sum_{q=1}^{2n+1} \Phi_q \left(\sum_{p=1}^n \phi_{p,q}(x_p) \right) \quad (9)$$

式中: $\text{KAN}(\mathbf{x})$ 代表输入 x 的预测值; $\Phi_q(\cdot)$ 和 $\phi_{p,q}(\cdot)$ 为可训练函数; $\phi_{p,q}(x_p)$ 为第 q 与第 p 项的函数组合, x_p 为输入向量的第 p 个分量; Φ_q 为高维特征的合成映射。

为提升网络优化效率,采用残差激活策略。通过引入基函数 $b(x)$, 使可训练激活函数 $\phi(\mathbf{x})$ 表示为

$$\phi(\mathbf{x}) = w_b b(\mathbf{x}) + w_s \cdot \text{spline}(\mathbf{x}) \quad (10)$$

式中,样条函数 $\text{spline}(\mathbf{x})$ 为一维函数组合,表示为 B 样条曲线的线性组合; w_s 和 w_b 为可训练变量, $b(\mathbf{x})$ 为固定形式 SiLU 函数。其中

$$b(\mathbf{x}) = \text{SiLU}(\mathbf{x}) = \frac{\mathbf{x}}{1 + e^{-\mathbf{x}}} \quad (11)$$

$$\text{spline}(\mathbf{x}) = \sum_i c_i B_i(\mathbf{x}) \quad (12)$$

式中: $B_i(\mathbf{x})$ 为相关的样条函数; c_i 为权重系数。

样条函数受 2 个关键参数的影响:样条阶数和节点密度。其中,样条阶数指在控制点之间进行多项式插值的函数阶数,并决定了每段样条曲线的平滑程度和拟合能力。节点密度表示样条插值中控制点的数量,影响样条函数捕捉输入特征细节变化的能力^[15]。因此,样条阶数与节点密度共同决定了 KAN 的非线性建模能力,是影响模型性能的核心参数。

2 深度混合预测模型设计

2.1 飞行数据多维度时序重构

飞行训练轨迹可表示为一系列离散时序数据点,每个轨迹点包含经度、纬度、高度等信息。以地

面坐标系 O_{XYZ} 为基准,飞行轨迹 T 可表示为 t 个时序点,公式为

$$T = \{(x_t, y_t, z_t) \mid t = 1, 2, \dots, n\} \quad (13)$$

式中: (x_t, y_t, z_t) 为飞机在时刻 t 的坐标值。飞机的运动轨迹可在 O_{XYZ} 坐标系上沿着 X, Y, Z 轴正交分解,且相对独立。现以 X 轴为例(Y, Z 轴同理),其运动轨迹数据集 X 为

$$X = \{x_1, x_2, \dots, x_n\} \quad (14)$$

预测轨迹数据集 \hat{X} 为

$$\hat{X} = \{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_t\} \quad (15)$$

则飞行轨迹预测问题可描述为

$$\hat{T} = f(T) \quad (16)$$

式中, $f(\cdot)$ 表示预测模型函数映射关系。

如图 2 所示,时间序列数据重构流程为:在航迹数据间隔为 1 s 的条件下,从首个航迹点开始,将连续 w 个输入点与后续时间戳相邻的 p 个航迹点进行组合形成一个时间窗。该操作以滑动窗口方式重复执行,直至覆盖轨迹的第 p 个航迹点,其中 n 为选取轨迹样本的总长度。

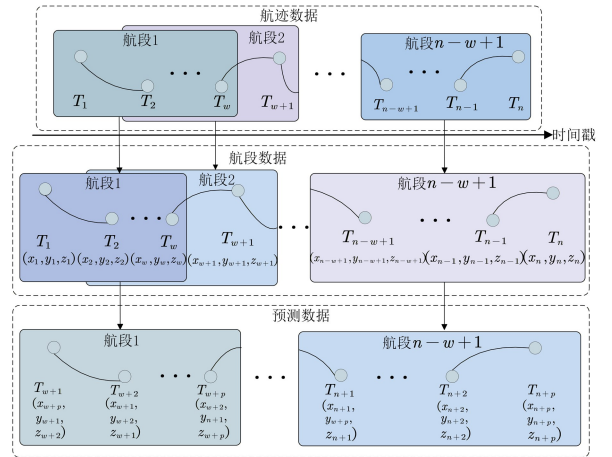


图 2 时间序列数据重构图

2.2 BiGRU-MHA-KAN 模型架构

针对 BiGRU 在时序特征学习中的特征均等化缺陷及 KAN 网络在时序依赖捕获的局限性,本研究提出三级递进式 BiGRU-MHA-KAN 融合架构,如图 3 所示。尽管 BiGRU-MHA 模块已具备时序特征捕捉能力,但传统全连接层在处理高维非线性飞行参数映射时仍存在函数逼近不足的缺陷。因此,引入 KAN 网络替代传统全连接层,利用其基于样条基函数的非线性映射能力对注意力加权特征进行非线性变换,提升复杂非线性时间序列预测的表征能力。

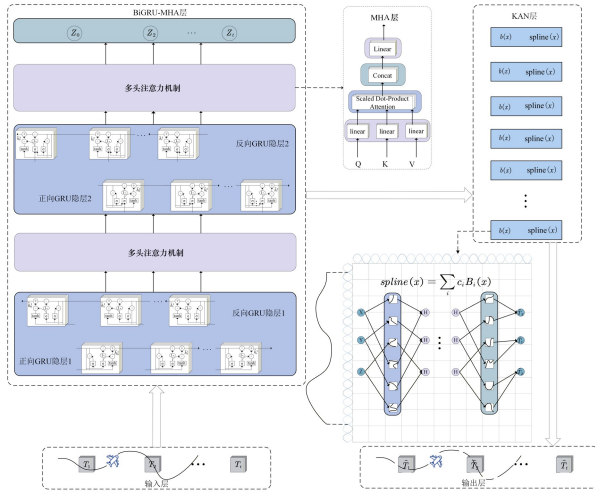


图 3 BiGRU-MHA-KAN 模型结构图

在该模型中,双层 BiGRU-MHA 作为时间序列特征提取模块:第一层 BiGRU-MHA 负责提取时间序列特征,第二层 BiGRU-MHA 深化轨迹特征关联

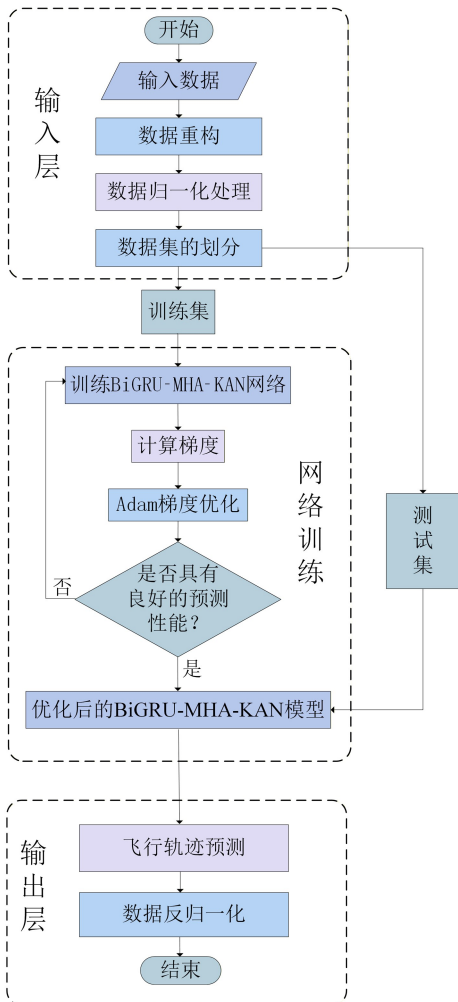


图 4 BiGRU-MHA-KAN 模型训练流程图

挖掘,学习更深层次的时间序列信息。而 KAN 则作为非线性映射模块,接收 BiGRU-MHA 的加权特征表示输出,通过 (10) ~ (13) 式的样条基函数实现非线性映射。其模型训练如图 4 所示。

3 仿真实验及结果分析

本研究利用真实的飞行数据对轨迹预测模型 BiGRU-MHA-KAN 进行仿真实验,并与 GRU、LSTM 等多种神经网络模型开展对比实验,以评估和验证该模型的可行性和优越性。

3.1 数据集构建与预处理

本研究采用国外 OpenSky 网络公开的 ADS-B 数据集。该数据集的采样频率为 2 Hz,主要包括 Unix 时间戳、飞机识别号、飞机位置、气压高度、应答传感器数量和测量数据等信息。

依据 Unix 标准化时间格式,将数据集时间戳的信息转化为真实时间。提取同一架飞机的经度、纬度和高度三维空间数据,按照真实时间的顺序,形成独立的 csv 文件。

生成的数据集的飞行数据量共 6 259,其中 80% 的数据量作为训练集,其余 20% 作为测试集。该模型训练采用 Adam 优化策略,学习率设定为 0.01。仿真运行环境为 Pytorch 平台。

航迹预测输出的结果包括预测位置点的经度、纬度和高度,因此,根据 (13) 式可得航迹预测点的输出数据结构为

$$\hat{T}_i = (\hat{x}_i, \hat{y}_i, \hat{z}_i) \quad (17)$$

为了优化模型对数据特征的学习效率,采用归一化方法,将数据映射到 [0, 1] 的范围内,以消除突变样本对模型的影响,公式为

$$\xi' = \frac{\xi - \xi_{\min}}{\xi_{\max} - \xi_{\min}} \quad (18)$$

式中: ξ 为航迹数据 \hat{T}_i 中的元素,包含经度、纬度以及高度; ξ_{\min} 为航迹数据集中的最小值; ξ_{\max} 为航迹数据集中的最大值; ξ' 为归一化后的航迹数据。

获得航迹预测数据后,需要经过反归一化处理以恢复数据的原始尺度,公式为

$$\xi = \xi'(\xi_{\max} - \xi_{\min}) + \xi_{\min} \quad (19)$$

3.2 对比实验

3.2.1 评价指标

本研究构建了包含预测偏差率、预测准确率、多

维度绝对误差、平均绝对误差 (mean absolute error, MAE)、均方根误差 (root mean square error, RMSE) 及推理时间和吞吐率的多维度量化评估体系。

预测偏差率采用逐点绝对误差进行量化表征, 公式为

$$E_p = \frac{1}{n} \sum_{i=1}^n \frac{|\hat{T}_i - T_i|}{|\hat{T}_i|} \quad (20)$$

式中: n 为超前预测的时刻点总数; \hat{T}_i 为 t 时刻轨迹的预测值; T_i 为 t 时刻轨迹的真实值。

预测准确率 P 的判定机制基于多维度误差阈值。首先计算经度、纬度及高度的三轴相对误差 E_i , 设定预设误差阈值 ε ; 当且仅当三轴误差均满足阈值条件时, 判定该轨迹点预测准确, 并累加至正确预测的轨迹点数量 \hat{N} , 公式为

$$E_i = \frac{|\hat{T}_i - T_i|}{T_i} \quad (21)$$

$$E_i \leq \varepsilon \quad (22)$$

$$P = \frac{\hat{N}}{N} \quad (23)$$

式中: N 为测试点总数量。

3.2.2 KAN 网络参数敏感性分析

在 BiGRU-MHA-KAN 模型中, KAN 网络通过样条函数实现非线性关系的精确建模, 其拟合能力直接影响模型的整体性能。因此, 需要通过敏感性分析, 探究 B 样条插值阶数 s (spline_order) 和节点密度 g (grid_size) 对模型性能的影响。根据文献[16-17]中选取数据长度的方式, 以 10 s 历史数据预测 2 s 未来数据为例, 利用平均准确率和偏差率评估

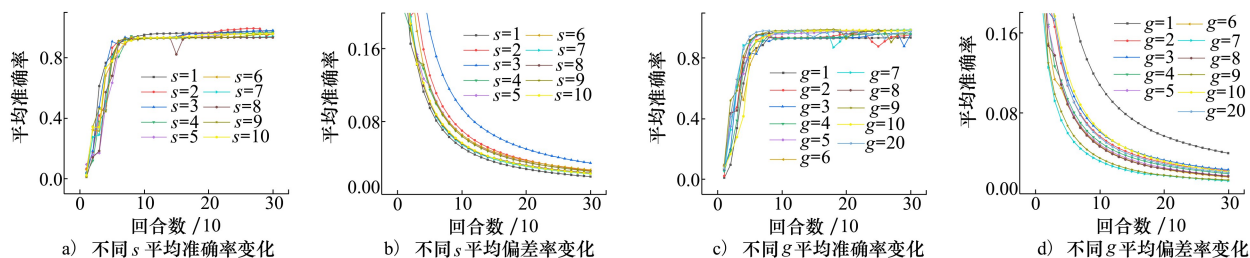


图 5 KAN 网络关键参数的敏感性实验结果

3.2.3 轨迹预测验证分析

为评估预测效果, 本研究构建包含 GRU、LSTM、BiGRU、BiLSTM 及 BiGRU-MHA-KAN 网络的对照实验, 各模型网络参数如表 1 所示。利用不同的历史飞行轨迹数据预测未来轨迹, 并每 10 回合计算准确率与偏差率的平均值。

KAN 网络参数中的 B 样条插值的阶数和节点密度对模型的影响。

1) 样条插值阶数影响

为了便于实验分析, 在设定节点密度为 5 的条件下, 探讨了不同样条插值阶数取值对模型性能的影响。从图 5a)~5b) 可以看出, 当样条插值阶数过大时, 模型在收敛过程中平均准确率反而会有所下降。

为确保模型维持较高准确率, 样条插值阶数的理想取值区间为 $[1, 4]$, 若超出该范围, 模型性能会出现显著下降。尽管高样条插值阶数有助于提升模型的灵活性, 但也易引发过拟合。因此, 样条插值阶数的合理选择是模型性能优化的关键因素。

2) 样条节点密度影响

控制样条插值中使用的控制点数量, 决定了样条节点密度和网格大小。 g 过小会导致平均准确率曲线上升缓慢、准确率偏低且偏差率较高; g 过大则会使迭代曲线呈现明显振荡。需注意的是, 图 5c) 中节点密度 $g = 10$ 的模型训练初期准确率提升较慢。这表明控制点数量增加虽能增强模型表达能力, 但会扩大优化空间复杂度, 进而延缓训练初期的收敛速度; 而 g 较小的模型因结构更简洁, 在训练初期更易收敛, 从而准确率提升更快。

由图 5c)~5d) 可以看出, 节点密度的理想取值区间为 $[5, 9]$ 。样条节点密度较小会导致模型细节捕捉能力不足, 降低预测性能; 节点密度较大虽能提升细节捕捉能力, 但易引发过拟合。因此, 节点密度的合理选取同样对模型性能优化至关重要。

1) 1 s 历史数据的短时轨迹预测

基于 1 s 历史飞行数据预测未来 1 s 的飞行训练轨迹。如图 6a)~6b) 所示, GRU、LSTM 网络预测效果欠佳。引入双向传播机制的 BiGRU、BiLSTM 网络因双向结构有效捕捉轨迹序列双向依赖关系表现更优。引入 KAN 网络与多头注意力机制后, MAE

表 1 网络参数取值表

模型	学习率	隐层维度	s	g	H	$\varepsilon/\%$
GRU	0.01	128				5
LSTM	0.01	128				5
BiGRU	0.005	128×2				5
BiLSTM	0.005	128×2				5
BiGRU-MHA-KAN	0.01	64×2	1	5	2	5

和 RMSE 显著降低,模型关键信息捕获与非线性表达能力均得到提升。

表 2 结果显示,BiGRU-MHA-KAN 模型平均准确率达 97.12%,平均偏差率为 1.78%,RMSE 与 MAE 也均为最低值。图 6c) 中该模型损失函数曲线稳定下降,收敛性良好,训练拟合效果佳,整体稳定性与精度更优。

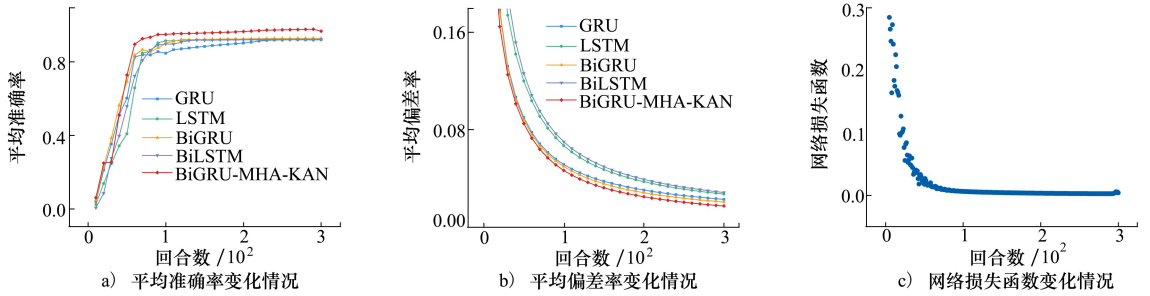


图 6 1 s 历史数据预测过程结果图

表 2 1 s 历史数据的轨迹预测评价指标

模型	平均准确率/%	平均偏差率/%	MAE/ 10^{-3}	RMSE/ 10^{-3}	推理时间/ μs	吞吐率/ s^{-1}
GRU	92.43	2.32	6.2	9.4	7.8	128 624
LSTM	92.67	2.75	7.1	10.6	18.4	54 485
BiGRU	93.23	2.11	5.3	8.6	16.8	59 665
BiLSTM	92.68	2.88	6.5	100	35.5	28 172
BiGRU-MHA-KAN	97.12	1.78	3.0	4.2	8.6	116 235

如图 7 所示,LSTM 网络误差波动显著、幅度较大。引入双向结构后各网络模型的预测效果均有所

改善,而融合 KAN 与多头注意力机制的 BiGRU 模型,不仅预测轨迹与真实轨迹重合度高、整体偏离最

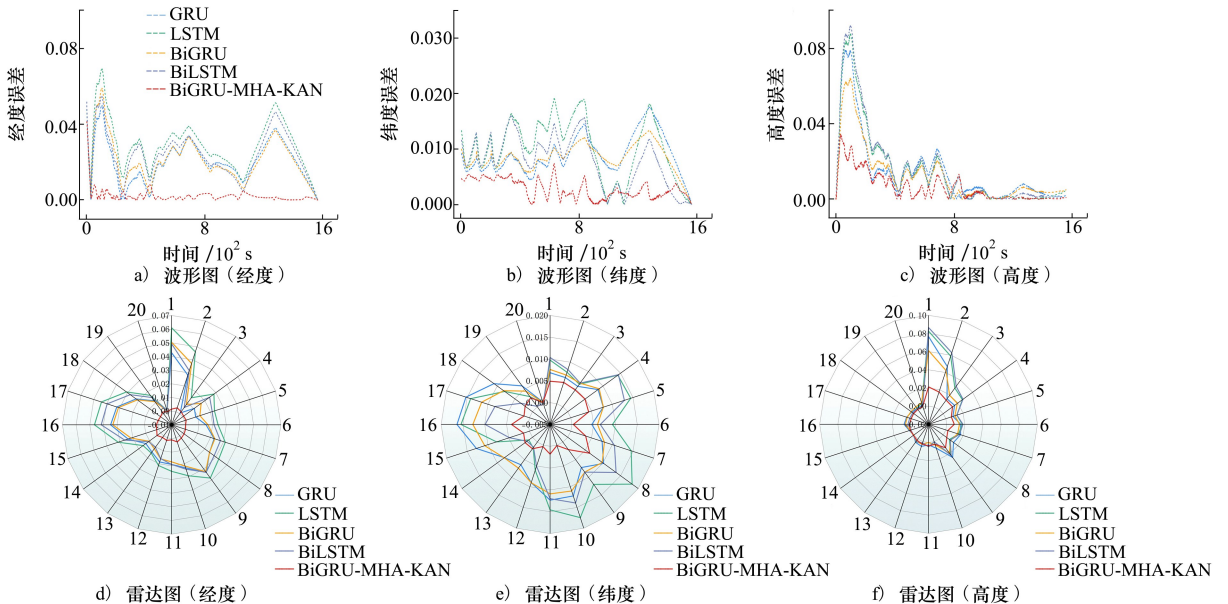


图 7 1 s 历史数据不同维度的预测误差对比图

小,预测误差也最小,预测性能最优,进一步验证了其在飞行轨迹预测任务中的优越性。实时性方面,该模型吞吐率 $116\ 235\ \text{s}^{-1}$ 、单样本推理时间 $8.6\ \mu\text{s}$,远低于 $0.5\ \text{s}$ 航迹点采样间隔,虽推理时间较传统 GRU 略有增加,但在精度大幅提升的同时仍保持良好实时性与计算效率。

2) 10 s 历史数据的短时轨迹预测

基于 10 s 历史飞行数据预测未来 1 s 的飞行训

练轨迹。表 3 结果同样显示 BiGRU-MHA-KAN 模型的平均准确率最高、平均偏差率及 MAE、RMSE 均为最低,表明其处理长历史信息时,仍能保持高精度输出与稳定性能。该模型单样本推理时间 $13.8\ \mu\text{s}$ 、吞吐率 $72\ 508\ \text{s}^{-1}$,进一步表明 BiGRU-MHA-KAN 模型在历史输入长度增加的情况下,仍具备良好的实时响应能力。

表 3 10 s 历史数据的轨迹预测评价指标

模型	平均准确率%	平均偏差率%	MAE/ 10^{-3}	RMSE/ 10^{-3}	推理时间/ μs	吞吐率/ s^{-1}
GRU	92.79	2.11	5.9	9.1	13.9	72 174
LSTM	93.43	3.05	6.5	9.2	28.4	35 195
BiGRU	93.57	1.78	4.0	6.3	25.9	38 481
BiLSTM	92.55	2.39	5.8	9.2	58.9	17 274
BiGRU-MHA-KAN	97.12	1.81	1.3	2.3	13.8	72 508

如图 8a)~8c) 所示,GRU 网络预测效果最差,LSTM 次之,同样 BiGRU、BiLSTM 网络的预测效果有所改善,再次验证了双向结构的有效性。与利用

1 s 历史数据的实验结果类似,引入多头注意力机制与 KAN 网络后,BiGRU-MHA-KAN 模型的平均准确率曲线上升,而平均偏差率曲线下降,网络损失函数

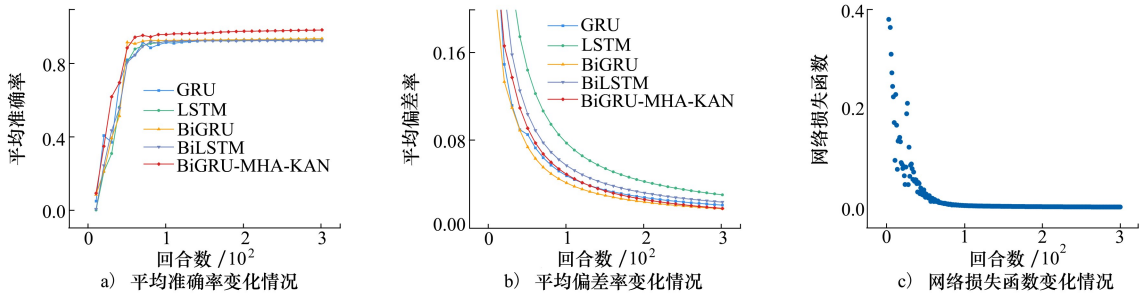


图 8 10 s 历史数据预测过程结果图

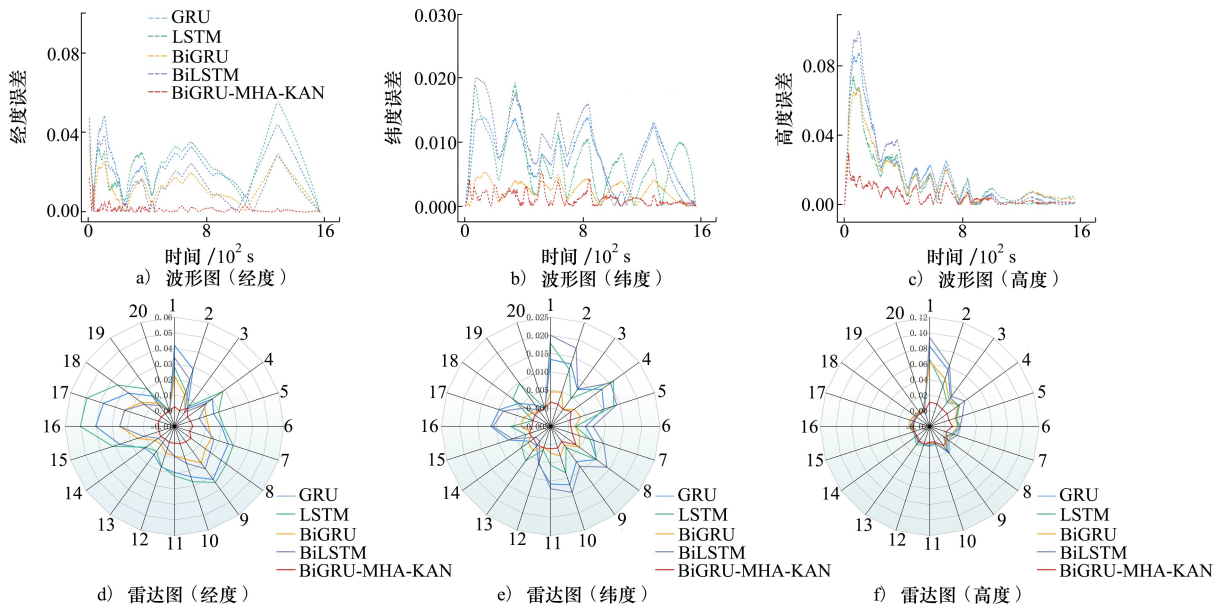


图 9 10 s 历史数据不同维度的预测误差对比图

曲线持续走低,有效解决了关键信息捕获均等化与非线性表达能力有限的问题,同样展现出良好的预测准确性。从图9中可以看出,BiGRU-MHA-KAN模型在不同维度的预测误差最小,进一步证明了该模型在飞行轨迹预测领域的性能优势更为突出。

4 结 论

本文利用融合 BiGRU 的时序建模能力、MHA 的特征选择机制,以及 KAN 可学习非线性逼近特

性,构建了 BiGRU-MHA-KAN 深度混合神经网络模型,实现了复杂飞行轨迹的高精度预测,解决了飞行训练轨迹预测中面临的高维非线性建模难题。该模型预测准确性与误差控制性能优于其他模型,且预测能力与历史信息长度呈显著正相关。KAN 的样条参数配置对模型性能起关键作用。未来可引入多源数据与强化学习机制,进一步提升模型在复杂环境中的实时预测能力,并推动其在飞行训练系统中的广泛应用。

参考文献:

- [1] Möller D P F. Machine learning and deep learning[M]. Cham: Springer Nature Switzerland, 2023: 347-384.
- [2] 张宏鹏, 黄长强, 唐上钦, 等. 基于卷积神经网络的无人作战飞机飞行轨迹实时预测[J]. 兵工学报, 2020, 41(9): 1894-1903.
Zhang Hongpeng, Huang Changqiang, Tang Shangqin, et al. CNN-based real-time prediction method of flight trajectory of unmanned combat aerial vehicle[J]. Acta Armamentarii, 2020, 41(9): 1894-1903. (in Chinese)
- [3] 付茂洛, 陈纪宗. 基于 LSTM 模型的机场航班飞行轨迹预测[J]. 现代计算机, 2022, 28(14): 55-60.
Fu Maoming, Chen Jizong. LSTM model-based flight trajectory prediction for airport flights[J]. Modern Computer, 2022, 28(14): 55-60. (in Chinese)
- [4] Hu Y C, Dai J, Yi J, et al. The high-speed aircraft trajectory prediction method based on deep learning network[J]. Journal of Physics, 2023, 2489(1): 012023.
- [5] 方伟, 汤森, 闫文君, 等. 基于残差神经网络和 LSTM 的飞行轨迹预测[J]. 舰船电子工程, 2023, 43(10): 42-46.
Fang Wei, Tang Miao, Yan Wenjun, et al. Flight trajectory prediction based on residual neural network and LSTM[J]. Ship Electronic Engineering, 2023, 43(10): 42-46. (in Chinese)
- [6] 张宗腾, 张琳, 汪文峰, 等. 一种基于双向 GRU 的 UAV 飞行轨迹预测方法[J]. 光电与控制, 2022, 29(3): 11-15.
Zhang Zongteng, Zhang Lin, Wang Wenfeng, et al. A method for UAV flight trajectory prediction based on bidirectional GRU[J]. Electronics Optics & Control, 2022, 29(3): 11-15. (in Chinese)
- [7] Liu Z, Wang Y, Vaidya S, et al. KAN: Kolmogorov-Arnold networks[EB/OL]. (2024-04-30)[2026-01-24]. <https://doi.org/10.48550/arXiv.2404.19756>.
- [8] Dai Y, Yu W, Leng M. A hybrid ensemble optimized BiGRU method for short-term photovoltaic generation forecasting[J]. Energy, 2024, 299: 131458.
- [9] Somvanshi S, Javed S A, Islam M M, et al. A survey on Kolmogorov-Arnold network[J]. ACM Computing Surveys, 2025, 58(2): 1-35.
- [10] Polar A, Poluektov M. A deep machine learning algorithm for construction of the Kolmogorov-Arnold representation[J]. Engineering Applications of Artificial Intelligence, 2021, 99: 104137.
- [11] 袁立宁, 冯文刚, 刘钊. 基于 Kolmogorov-Arnold 网络的节点分类算法[J]. 计算机科学与探索, 2025, 19(3): 645-656.
Yuan Lining, Feng Wengang, Liu Zhao. Node classification based on Kolmogorov-Arnold network[J]. Journal of Frontiers of Computer Science and Technology, 2025, 19(3): 645-656. (in Chinese)
- [12] Al-Qaness M A A, Ni S. TCNN-KAN: optimized CNN by Kolmogorov-Arnold network and pruning techniques for sEMG gesture recognition[J]. IEEE Journal of Biomedical and Health Informatics, 2025, 29(1): 188-197.
- [13] Ren D, Hu Q, Zhang T. EKLT: Kolmogorov-Arnold attention-driven LSTM with transformer model for river water level prediction[J]. Journal of Hydrology, 2025, 649: 132430.
- [14] Sidharth S S, Keerthana A R, Gokul R, et al. Chebyshev polynomial-based Kolmogorov-Arnold networks: an efficient architec-

- ture for nonlinear function approximation[J/OL]. (2024-05-12)[2026-01-24]. <https://doi.org/10.48550/arXiv.2405.07200>.
- [15] Vaca-Rubio C J, Blanco L, Pereira R, et al. Kolmogorov-Arnold networks (KANs) for time series analysis[C]//2024 IEEE Globecom Workshops, Cape Town, South Africa, 2024.
- [16] 张堃, 杜睿怡, 时昊天, 等. 基于 Mogrifier-BiGRU 的飞行器轨迹预测[J]. 兵工学报, 2024, 45(2): 373-384.
Zhang Kun, Du Ruiyi, Shi Haotian, et al. Prediction of aircraft trajectory based on Mogrifier-BiGRU[J]. Acta Armamentarii, 2024, 45(2): 373-384. (in Chinese)
- [17] 罗恺玮, 周激流. 基于大语言模型的飞行轨迹预测方法[J]. 计算机应用研究, 2025, 42(12): 3644-3650.
Luo Kaiwei, Zhou Jiliu. Flight trajectory prediction method based on large language models[J]. Application Research of Computers, 2025, 42(12): 3644-3650. (in Chinese)

BiGRU-MHA-KAN based flight training trajectory prediction

Zhang Di¹, Wei Xueqian², Li Sijie²

(1.Engineering Technology Training Center, Civil Aviation University of China, Tianjin 300300, China;
2.School of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China)

Abstract: To improve the accuracy of trajectory prediction in flight training and enhance the reliability of prediction models, a deep hybrid neural network model named BiGRU-MHA-KAN is proposed, in which the bidirectional gated recurrent unit (BiGRU), multi-head attention (MHA) and Kolmogorov-Arnold networks (KAN) are integrated. The model strengthens the temporal feature extraction and nonlinear dynamic modeling through trajectory data pre-processing and reconstruction, combining with the bidirectional modeling, attention mechanisms and KAN networks. Simulated experiments systematically analyze the effect of the different parameter settings and historical data volumes on the model performance. The results demonstrate that, comparing with the other trajectory prediction models, the present method achieves an improvement in prediction accuracy by 4.81%–5.83%, while significantly reducing the mean squared error and root mean squared error, demonstrating the stronger temporal modeling capability and stability in flight training scenarios.

Keywords: trajectory prediction; flight training; Kolmogorov-Arnold networks; bidirectional gated recurrent unit; attention mechanisms

引用格式: 张迪, 魏雪倩, 黎思杰. 基于 BiGRU-MHA-KAN 的飞行训练轨迹预测方法[J]. 西北工业大学学报, 2026, 44(1): 185-193.

Zhang Di, Wei Xueqian, Li Sijie. BiGRU-MHA-KAN based flight training trajectory prediction[J]. Journal of Northwestern Polytechnical University, 2026, 44(1): 185-193. (in Chinese)