



一种复值多域损失融合的时间序列预测网络

郝剑龙, 杜银焕, 薛令荣, 张卫, 李睿璞

引用本文:

郝剑龙, 杜银焕, 薛令荣, 等. 一种复值多域损失融合的时间序列预测网络[J]. 智能系统学报, 2026, 21(5): 1211–1220.

HAO Jianlong, DU Yinhuang, XUE Lingrong, et al. A complex-valued multi-domain loss fusion network for time series forecasting[J]. CAAI transactions on intelligent systems, 2026, 21(5): 1211–1220.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202511025>

您可能感兴趣的其他文章

基于分类差异与信息熵对抗的无监督域适应算法

Unsupervised domain adaptation algorithm based on classification discrepancy and information entropy
智能系统学报. 2021, 16(6): 999–1006 <https://dx.doi.org/10.11992/tis.202010020>

新冠肺炎疫情趋势预测模型

Epidemics trend prediction model of COVID-19
智能系统学报. 2021, 16(3): 528–536 <https://dx.doi.org/10.11992/tis.202008037>

一种基于深度学习目标检测的长时目标跟踪算法

A long-term object tracking algorithm based on deep learning and object detection
智能系统学报. 2021, 16(3): 433–441 <https://dx.doi.org/10.11992/tis.201910029>

基于数据增广和复制的中文语法错误纠正方法

Chinese grammatical error correction method based on data augmentation and copy mechanism
智能系统学报. 2020, 15(1): 99–106 <https://dx.doi.org/10.11992/tis.202001014>

融合朴素贝叶斯方法的复杂网络链路预测

Link prediction in complex networks with syncretic naive Bayes methods
智能系统学报. 2019, 14(1): 99–107 <https://dx.doi.org/10.11992/tis.201810025>

基于宽度学习方法的多模态信息融合

Multi-modal information fusion based on broad learning method
智能系统学报. 2019, 14(1): 150–157 <https://dx.doi.org/10.11992/tis.201803022>

DOI: 10.11992/tis.202511025

网络出版地址: <https://link.cnki.net/urlid/23.1538.tp.20260707.0938.006>

一种复值多域损失融合的时间序列预测网络

郝剑龙, 杜银焕, 薛令荣, 张卫, 李睿璞

(山西财经大学 信息学院, 山西 太原 030006)

摘要: 时间序列预测在电力、气象、交通等关键领域具有广泛的应用价值, 传统预测模型受限于损失函数单一、实数域相位信息丢失及固定权重优化失衡, 难以适配复杂时序特性。本文提出了一种基于复值多域损失融合的时间序列预测网络 (complex-valued multi-domain loss fusion time series forecasting network, CFuseTS), 利用复值卷积与共轭复值注意力方法联合捕捉时序数据的幅度与相位特征, 突破实数域建模的局限; 同时构建多域协同损失模块对多元时间序列进行联合约束, 依托动态梯度均衡算法实时调整各域权重, 解决传统固定权重导致的优化偏差。在 8 个基准数据集上的实验表明, CFuseTS 与现有先进方法相比, 在所有预测长度 (96/192/336/720 时间步) 上的均方误差平均降低了 6.51%, 最多降低了 14.78%, 充分验证了其有效性与泛化能力。

关键词: 时间序列预测; 复值神经网络; 损失函数; 复值注意力机制; 快速傅里叶变换; 相位感知; 时频融合; 动态梯度均衡

中图分类号: TP301.6 文献标志码: A 文章编号: 1673-4785(2026)05-1211-10

中文引用格式: 郝剑龙, 杜银焕, 薛令荣, 等. 一种复值多域损失融合的时间序列预测网络 [J]. 智能系统学报, 2026, 21(5): 1211-1220.

英文引用格式: HAO Jianlong, DU Yinhuang, XUE Lingrong, et al. A complex-valued multi-domain loss fusion network for time series forecasting[J]. CAAI transactions on intelligent systems, 2026, 21(5): 1211-1220.

A complex-valued multi-domain loss fusion network for time series forecasting

HAO Jianlong, DU Yinhuang, XUE Lingrong, ZHANG Wei, LI Ruipu

(School of Information, Shanxi University of Finance and Economics, Taiyuan 030006, China)

Abstract: Time series forecasting possesses extensive application value in critical fields such as power systems, meteorology, and transportation. Traditional models are constrained by single temporal domain loss functions, phase information loss in the real number domain, and fixed-weight optimization imbalance, making it difficult to adapt to complex temporal characteristics. This paper proposes a complex-valued multi-domain loss fusion time series prediction network (CFuseTS). It adopts complex-valued convolution and conjugate complex-valued attention to jointly capture the amplitude and phase features of time series data, breaking the inherent limitations of real-valued modeling. Meanwhile, a multi-domain synergistic loss (MDSL) module is constructed to implement joint constraints on multivariate time series, and the dynamic gradient equalizer (DGE) dynamically adjusts the weights of each domain in real time to mitigate optimization bias from conventional fixed-weight strategies. Experiments on 8 benchmark datasets show that CFuseTS reduces mean squared error (MSE) by an average of 6.51% and a maximum of 14.78% for all predicted lengths (96/192/336/720 time steps) compared to existing advanced methods, fully validating its effectiveness and generalization ability.

Keywords: time series forecasting; complex-valued neural networks; loss function; complex attention mechanism; fast Fourier transform; phase aware; time-frequency fusion; dynamic gradient balancing

收稿日期: 2025-11-19. 网络出版日期: 2026-07-07.

基金项目: 山西省基础科学研究计划自然科学面上项目 (202303021221185).

通信作者: 郝剑龙. E-mail: haojianlong2012@sxufe.edu.cn.

©《智能系统学报》编辑部版权所有

时间序列预测作为数据分析与机器学习领域的核心任务, 在电力^[1]、气象^[2-3]、交通^[4]等关键领域具有不可替代的应用价值。然而, 现实世界的时序数据往往呈现出复杂特性: 非平稳性^[5]表现

为突发波动,如极端天气导致的电力负荷骤增;多尺度周期性^[6]体现为日、周、年等周期叠加,如电力负荷的日周期,交通流量的周周期等;长程依赖^[7]则表现为跨时间步的关联影响,如政策调整对经济指标的持续作用。这些特性对预测模型的特征捕捉能力带来了严峻的挑战。

传统预测模型难以同时适配上述复杂特性,具体可归纳为三类技术局限:其一,损失函数设计单一化。多数模型仅采用时域损失均方差(mean squared error, MSE)约束数值误差,忽视频域中的周期成分与结构特征,导致预测结果数值出现偏差。其二,实数域建模缺失相位表征能力,难以适配多尺度周期叠加的数据特征。周期信号的时序对齐高度依赖相位信息。现有 Transformer^[8]类模型(Informer^[9]、PatchTST^[10]、TFFormer^[11])等主流时频融合模型均基于实数域进行特征提取,只能提取幅度信息,无法建模相位关联信息。其三,多域损失与多尺度融合策略存在问题。现有的多域特征建模模型^[12]多采用固定权重配置,无法适配不同特征域动态收敛特性,易造成优化偏向单一域;Mts-mixers^[13]虽能识别多周期尺度,但缺少动态融合机制,各尺度特征难以有效互补。

针对上述问题,本文提出了一种基于复值多域损失融合的时间序列预测网络(complex-valued multi-domain loss fusion time series forecasting network, CFuseTS),核心贡献包括:

1) 构建了相位感知复值滤波模块,同步提取时序幅度与相位特征,提升多周期信号建模能力,弥补实数域建模缺陷。

2) 设计了一种多尺度复值卷积-门控融合架构,同步捕获细粒度短期波动与粗粒度长期周期趋势,弥补现有模型多尺度特征无法动态互补的不足。

3) 提出了一种多域协同损失融合机制,融合多维度损失并通过动态梯度均衡算法(dynamic gradient equalizer, DGE)自适应调整权重。

1 相关工作

1.1 多元时序预测方法

多元时间序列预测(multivariate time series forecasting, MTSF)^[14]的目的是通过捕捉变量间耦合依赖与长程时序关联实现精准预测,现有方法可分为统计模型与深度学习模型两类,二者在建模范式上各有突破,但均存在对复杂多元时序特征刻画不充分的局限。

统计模型中 ARIMA^[15]通过差分平稳化序列建模线性依赖,适用于单变量平稳场景,无法适配非线性与多变量耦合;SARIMA^[16]引入季节性因子以增强周期适配性,但其线性假设仍难以适配电力、交通等复杂时序数据的复杂动态性。

深度学习模型中, RNN^[17]存在梯度消失问题,限制长序列建模能力; LSTM^[18]引入门控机制缓解梯度问题,但多元场景下效率低下,变量关联捕捉能力弱; Transformer 及其变体(Informer、TFFormer 等)基于自注意力机制捕获长程依赖,但多聚焦时域建模,未充分利用频域相位信息,实数域建模也无法刻画周期对齐所需的相位依赖; Mts-mixers、PatchTST、iTransformer^[19]等虽在效率或变量依赖上有所优化,却未涉及复值相位建模,难以充分挖掘周期对齐核心特征。

1.2 频域增强时序预测

频域方法在信号处理领域已有广泛应用,例如将 MSE 通过快速傅里叶变换(fast Fourier transform, FFT)转换至频域,以约束频谱特征。但现有研究多采用单一的频域损失,缺少时域、结构特征联合约束。Fedformer^[20]通过频域自注意力挖掘周期分量,但忽略了频率间的动态关联; FreTS^[21]融合时频信息,但仅基于实数建模,缺失相位表征; TimesNet^[22]采用时序分解提取频域特征,但多尺度融合不具备自适应性。现有频域增强模型普遍仅采用单一时域损失,未联合频域、结构特征共同优化,难以充分拟合复杂时序规律。

1.3 复值神经网络与相位建模

复值神经网络通过复数运算处理复值信号,在雷达^[23]、通信^[24]等领域广泛应用。如 Complex CNN^[25]将实数卷积扩展至复数域,通过哈密顿乘积融合虚实特征,但卷积核未适配多元时序输入窗口 L ; Complex Transformer^[26]通过复值自注意力捕捉相位依赖,缺少频域约束; C-LSTM^[27]通过复值门控机制处理复值时序数据,多用于信号重构而非预测任务。现有复值注意力时序研究^[28-29]未适配多变量维度,难以提取跨变量共享相位特征。

综上,现有方法在复数域建模、多域损失设计与多尺度融合方面仍存在一定不足。CFuseTS 模型通过相位感知复值滤波模块与多域协同损失机制的深度融合,实现了多元时序预测精度的提升。

2 模型方法

2.1 问题定义

针对多变量时序观测样本,挖掘序列时域动

态趋势、频域周期分量与跨变量联动关联, 捕捉幅值与相位耦合的内在协同特征, 实现未来时序的高精度预测。给定多元历史观测序列 $\mathbf{X}_{n,l} \in \mathbb{R}^{N \times L \times D}$, 其中:

- 1) N 为变量数, 如电力数据中的电压、电流、温度, 气象数据中的湿度、风速等;
- 2) L 为输入窗口长度;
- 3) $\mathbf{X}_{n,l} \in \mathbb{R}$ 为第 n 个变量在第 l 个时间步的值;
- 4) D 为特征维度。

预测目标为未来 H 个时间步的多元预测序列 $\hat{\mathbf{Y}}_{n,h} \in \mathbb{R}^{N \times H \times D}$, 其中 $\hat{\mathbf{Y}}_{n,h}$ 表示第 n 个变量在第 $T+h$ 个时间步 ($h=1,2,\dots,H$) 的预测值, T 为历史序列的总长度。

为生成训练样本, 采用滑动窗口^[30]策略: 以长度为 L 的窗口在历史序列上滑动, 每个窗口的输入样本为第 i 个窗口的多变量历史数据, 记作 $\mathbf{X}_i \in \mathbb{R}^{N \times L \times D}$; 输出样本为该窗口对应的未来 H 步多变量真实数据, 记作 $\mathbf{Y}_i \in \mathbb{R}^{N \times H \times D}$ 。

2.2 模型整体架构

CFuseTS 模型采用了“特征提取-损失约束”的双驱动架构, 解决了实数域模型难以刻画相位信息、频域模型无法动态跨变量关联的问题, 实现复数域特征提取与多域损失约束的协同优化, 模型整体架构如图 1 所示。模型由数据嵌入模块、相位感知复值滤波模块、投影层和多域协同损失模块组成。

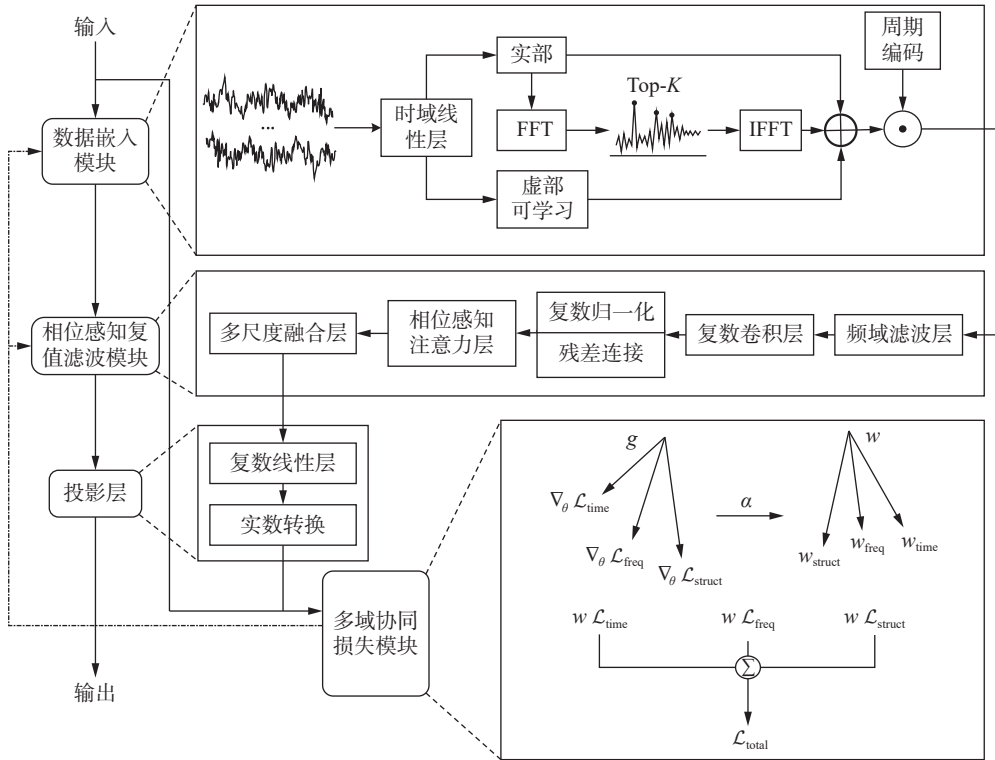


图 1 CFuseTS 模型整体架构

Fig. 1 CFuseTS model overall architecture

2.3 数据嵌入模块

数据嵌入模块将原始实数时序数据转换为含相位信息与多尺度周期特征的复数域特征, 为复数域特征提取提供输入基础。

2.3.1 可学习虚部基底

传统的复值嵌入常将虚部固定为零, 无法适应不同序列的相位偏移特性。本模块通过线性层构建虚部特征, 在初始阶段, 虚部被设置为零, 随着模型训练的推进, 能够自适应地捕捉序列特异性相位偏移, 从而为后续复数域处理提供更为丰富、准确的相位信息, 公式表示为

$$\mathbf{X}_{\text{imag}} = \text{Linear}_{\text{Im}}(\mathbf{X})$$

式中 $\text{Linear}_{\text{Im}} \in \mathbb{R}^{L \times D}$ 为虚部线性层, 输出 $\mathbf{X}_{\text{imag}} \in \mathbb{R}^{N \times L \times D}$ 。虚部初始为零的设计, 是为了让模型在训练初期依赖实部的数值特征, 随训练迭代逐步通过梯度下降学习相位偏移, 当输入序列存在周期性波动时, 虚部会自适应调整以匹配从数值变化到相位变化的映射关系, 实现实部载数值, 虚部载相位的复值特征表达。

2.3.2 频域增强实部

为显式捕捉时序数据的多尺度周期性, 模块先对输入数据进行实部处理, 通过表达式 $\mathbf{X}_{\text{real}} = \text{Linear}_{\text{Re}}(\mathbf{X})$ 获取实部特征, 对实部特征 $\mathbf{X}_{\text{real}}$ 执行快速傅里叶变换提取关键频率分量, 再经逆

傅里叶变换 (inverse fast Fourier transform, IFFT) 映射回时域并融合原始实部, 得到融合频率特征的实部 $\mathbf{X}_{\text{real-freq}}$, 公式表示为

$$\mathbf{X}_{\text{real-freq}} = \text{IFFT}(\text{Top} - K(\text{FFT}(\mathbf{X}_{\text{real}})))$$

式中 $\text{Top} - K$ 操作提取频率谱中的前 K 个关键频率分量。最终实部特征更新为 $\mathbf{X}_{\text{real}} = \mathbf{X}_{\text{real}} + \mathbf{X}_{\text{real-freq}}$, 确保实部同时承载原始数值与频域周期信息。

2.3.3 复值周期编码与融合

对每个时间步生成周期编码 $T_t = e^{i2\pi t/\tau}$, 其中 τ 为周期参数。该编码将离散时间戳转换为连续相位信号, 强化了时间特征与数值特征的融合, 提升对时序数据的理解与处理能力。

首先将实部与虚部进行融合, 得到输出结果 $\mathbf{X}_{\text{emb}} = \mathbf{X}_{\text{real}} + i \cdot \mathbf{X}_{\text{imag}} \in \mathbb{C}^{N \times L \times D}$ 完成从实数到复数域的嵌入转换。其次, 进行编码融合, 将周期编码映射到特征维度 D , 并与 $\mathbf{X}_{\text{emb}}$ 执行哈密顿乘积, 注入周期信息, 得到最终输出, 公式表示为

$$\mathbf{X}_{\text{emb}}^{\text{fusion}} = \mathbf{X}_{\text{emb}} \odot \text{Linear}_c(T)$$

式中: $\text{Linear}_c \in \mathbb{C}^{L \times D}$ 为复值线性层; $T \in \mathbb{C}^L$ 为周期编码; $\mathbf{X}_{\text{emb}}^{\text{fusion}} \in \mathbb{C}^{N \times L \times D}$; $\odot$ 为哈密顿乘积, 对于复数 $a = a_r + i \cdot a_i$ 和 $b = b_r + i \cdot b_i$, 乘积结果为 $(a_r b_r - a_i b_i) + i \cdot (a_r b_i + a_i b_r)$ 。

2.4 相位感知复值滤波模块

相位感知复值滤波模块 (phase aware complex-valued filtering, PACF) 通过复数域卷积与注意力机制, 提取时序数据的频率特征与长程相位依赖, 解决了传统实值滤波无法同时捕捉幅度与相位的问题。

2.4.1 频域滤波层

频域滤波作为前置操作, 先筛选输入特征 $\mathbf{X}_{\text{emb}}^{\text{fusion}}$ 中的关键频率, 再输入复值卷积层, 避免冗余频率干扰相位提取。为降低高频噪声与冗余频率分量, 采用“幅度排序 + 动态门控”策略, 先保留幅度前 K_{static} 的基础频率, 再通过动态门控筛选动态关键频率, 输出 $\mathbf{F}_{\text{filter}} \in \mathbb{C}^{N \times L \times D}$, 公式表示为

$$\mathbf{F}_{\text{filter}} = \sigma(\text{Linear}(\mathbf{X}_{\text{emb}}^{\text{fusion}})) \odot \mathbf{X}_{\text{emb}}^{\text{fusion}}$$

2.4.2 复值卷积层

基于哈密顿乘积的复值卷积, 同时捕捉时域波形与频域相位 / 幅度, 公式表示为

$$\begin{cases} \mathbf{F}_{\text{conv,real}} = \mathbf{W}_{\text{real}} * \mathbf{Z}_{\text{real}} - \mathbf{W}_{\text{imag}} * \mathbf{Z}_{\text{imag}} \\ \mathbf{F}_{\text{conv,imag}} = \mathbf{W}_{\text{real}} * \mathbf{Z}_{\text{imag}} + \mathbf{W}_{\text{imag}} * \mathbf{Z}_{\text{real}} \end{cases}$$

式中: $\mathbf{Z} = \mathbf{X}_{\text{emb}}^{\text{fusion}} \in \mathbb{C}^{N \times L \times D}$ 为输入复值特征, $\mathbf{W} \in \mathbb{C}^{K \times D \times D}$ 为复值卷积核, K 为核大小, $*$ 表示实数域滑动窗口卷积。 $\mathbf{W}_{\text{real}}$ 为卷积核的实部, $\mathbf{W}_{\text{imag}}$ 为卷积核的虚部; $\mathbf{F}_{\text{conv,real}}$ 和 $\mathbf{F}_{\text{conv,imag}}$ 分别为卷积输出的实部和

虚部。

为输出复值特征添加复值偏置 $b \in \mathbb{C}^D$, 增强模型表达能力, 输出最终复值特征, 公式表示为

$$\mathbf{F}_{\text{conv}} = (\mathbf{F}_{\text{conv,real}} + b_{\text{real}}) + i \cdot (\mathbf{F}_{\text{conv,imag}} + b_{\text{imag}})$$

式中输出 $\mathbf{F}_{\text{conv}} \in \mathbb{C}^{N \times (L-K+1) \times D}$ 。上述运算基于哈密顿乘积, 确保实部与虚部的交叉作用被捕捉, 是复值卷积能同时建模幅度与相位的核心。

2.4.3 复值归一化与残差连接

复值层归一化对实部和虚部分别归一化, 避免单一维度数值偏移, 公式表示为

$$\begin{cases} \mathbf{F}_{\text{norm,real}} = \frac{\mathbf{F}_{\text{conv,real}} - \mu_{\text{real}}}{\sqrt{\sigma_{\text{real}}^2 + \varepsilon}} \\ \mathbf{F}_{\text{norm,imag}} = \frac{\mathbf{F}_{\text{conv,imag}} - \mu_{\text{imag}}}{\sqrt{\sigma_{\text{imag}}^2 + \varepsilon}} \end{cases}$$

式中: μ 与 σ 分别为均值和方差, $\varepsilon = 10^{-6}$ 为数值稳定项。最终归一化特征 $\mathbf{F}_{\text{norm}} = \mathbf{F}_{\text{norm,real}} + i \cdot \mathbf{F}_{\text{norm,imag}} \in \mathbb{C}^{N \times (L-K+1) \times D}$ 。

复值残差连接是将归一化后的特征与输入复值特征相加, 保留原始相位与幅度信息, 缓解梯度衰减, 公式表示为

$$\mathbf{F}_{\text{res}} = \mathbf{F}_{\text{norm}} + \text{Pad}(\mathbf{Z})$$

式中: $\text{Pad}(\cdot)$ 为零填充操作, 输出 $\mathbf{F}_{\text{res}} \in \mathbb{C}^{N \times (L-K+1) \times D}$ 。

2.4.4 相位感知注意力层

相位感知注意力层通过共轭复数运算计算相似度, 保留多变量间的相位依赖, 避免梯度消失, 其示意图如图 2 所示。

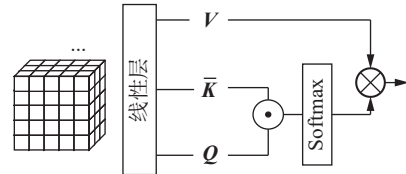


图 2 相位感知注意力层

Fig. 2 Phase-aware attention layer

生成复数的 Query($\mathbf{Q}$), Key($\mathbf{K}$) 和 Value($\mathbf{V}$): 将 $\mathbf{F}_{\text{res}}$ 通过 3 个独立复值线性层映射为注意力三要素, 公式表示为

$$\begin{cases} \mathbf{Q} = \text{Linear}_Q(\mathbf{F}_{\text{res}}) \\ \mathbf{K} = \text{Linear}_K(\mathbf{F}_{\text{res}}) \\ \mathbf{V} = \text{Linear}_V(\mathbf{F}_{\text{res}}) \end{cases}$$

式中: $\text{Linear}_Q, \text{Linear}_K, \text{Linear}_V \in \mathbb{C}^{D \times d}$ (d 为注意力头维度), 输出 $\mathbf{Q}, \mathbf{K}, \mathbf{V} \in \mathbb{C}^{N \times L' \times d}$ ($L' = L - K + 1$)。

相位感知相似度计算, 保持实部不变, 虚部取反, 确保分数为实数, 避免复数分数导致的优化不稳定, 同时保留相位依赖, 公式表示为

$$S_{\text{score}} = \frac{\text{Re}(\mathbf{Q} \odot \bar{\mathbf{K}})}{\sqrt{d}}$$

式中: $\bar{K}$ 为 K 的共轭复数; $\text{Re}(\cdot)$ 为取实部操作; $\sqrt{d}$ 为缩放因子, 缓解维度增长导致的分数膨胀。

注意力加权通过 Softmax 归一化相似度得分, 再与 V 的实部、虚部分别相乘得到注意力输出, 公式表示为

$$\begin{cases} \mathbf{F}_{\text{attn,real}} = \text{Softmax}(\mathbf{S}_{\text{score}}) \cdot \mathbf{V}_{\text{real}} \\ \mathbf{F}_{\text{attn,imag}} = \text{Softmax}(\mathbf{S}_{\text{score}}) \cdot \mathbf{V}_{\text{imag}} \end{cases}$$

式中 $\text{Softmax}(\cdot)$ 沿时间步维度归一化。最终输出复值特征为 $\mathbf{F}_{\text{attn}} = \mathbf{F}_{\text{attn,real}} + \mathbf{i} \cdot \mathbf{F}_{\text{attn,imag}} \in \mathbb{C}^{N \times L' \times d}$ 。

2.4.5 多尺度复值特征融合

针对时序数据中高频短时波动、中频趋势、低频长周期等不同频率成分, 采用多尺度卷积提取特征并自适应融合。

采用核大小为 3、5、7 的复值卷积提取多粒度特征, 公式表示为

$$\begin{cases} \mathbf{F}_{\text{scale1}} = \text{Conv}_{K=3}(\mathbf{X}_{\text{emb}}^{\text{fusion}}) \\ \mathbf{F}_{\text{scale2}} = \text{Conv}_{K=5}(\mathbf{X}_{\text{emb}}^{\text{fusion}}) \\ \mathbf{F}_{\text{scale3}} = \text{Conv}_{K=7}(\mathbf{X}_{\text{emb}}^{\text{fusion}}) \end{cases}$$

门控权重计算是 CFuseTS 模型实现多尺度特征自适应融合的核心机制, 通过门控融合自适应调整各尺度特征的贡献比例, 公式表示为

$$\begin{cases} \mathbf{G} = \sigma(\text{Linear}_G([\mathbf{F}_{\text{scale1}}; \mathbf{F}_{\text{scale2}}; \mathbf{F}_{\text{scale3}}])) \\ \mathbf{F}_{\text{fusion}} = \mathbf{G} \odot \mathbf{F}_{\text{scale1}} + (1 - \mathbf{G}) \odot (\mathbf{F}_{\text{scale2}} + \mathbf{F}_{\text{scale3}}) \end{cases}$$

式中: $\text{Linear}_G \in \mathbb{R}^{3D \times D}$ 为门控线性层, $\sigma(\cdot)$ 为 Sigmoid 函数, 对于长度不匹配的特征采用右零填充策略, 将注意力特征 $\mathbf{F}_{\text{attn}}$ 与多尺度融合特征相加, 得到 $\mathbf{F}_{\text{PACF}} = \mathbf{F}_{\text{attn}} + \text{Pad}(\mathbf{F}_{\text{fusion}})$, PACF 模块最终输出为

$$\mathbf{F}_{\text{PACF}} \in \mathbb{C}^{N \times L'' \times D} \quad (L'' = L - K + 1)$$

2.5 投影层

投影层将 PACF 模块输出的复值特征转换为实数预测序列, 完成从复值特征空间到实数预测空间的映射, 同时保留复数域学习到的关键信息, 其示意图如图 3 所示。

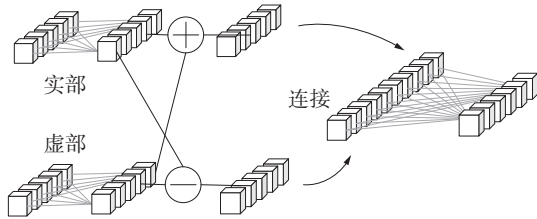


图 3 投影层
Fig. 3 Projection layer

2.5.1 复值投影

将 PACF 模块的复值特征 $\mathbf{F}_{\text{PACF}}$ 通过复值线性层映射到预测空间, 再通过实部与虚部的交叉运算, 确保映射过程中保留 PACF 模块提取的相位信息, 避免传统实数投影直接丢弃虚部导致的相

位丢失, 公式表示为

$$\begin{cases} \mathbf{Y}_{\text{complex,real}} = \text{Linear}_{\text{proj,Re}}(\mathbf{F}_{\text{PACF,real}}) - \text{Linear}_{\text{proj,Im}}(\mathbf{F}_{\text{PACF,imag}}) \\ \mathbf{Y}_{\text{complex,imag}} = \text{Linear}_{\text{proj,Re}}(\mathbf{F}_{\text{PACF,imag}}) + \text{Linear}_{\text{proj,Im}}(\mathbf{F}_{\text{PACF,real}}) \end{cases}$$

式中 $\text{Linear}_{\text{proj,Re}}$ 、 $\text{Linear}_{\text{proj,Im}} \in \mathbb{R}^{D \times C \times H}$ 为复值投影矩阵, 输出为

$$\mathbf{Y}_{\text{complex}} = \mathbf{Y}_{\text{complex,real}} + \mathbf{i} \cdot \mathbf{Y}_{\text{complex,imag}} \in \mathbb{C}^{N \times H \times C}$$

2.5.2 实数转换

取复值投影结果的实部作为最终预测结果, 保留复值学习得到的相位信息, 公式表示为

$$\hat{\mathbf{Y}} = \mathbf{Y}_{\text{complex,real}}$$

式中 $\hat{\mathbf{Y}} \in \mathbb{R}^{N \times H \times C}$, 实部保留输出符号信息。

2.6 多域协同损失模块

多域协同损失模块 (multi-domain synergistic loss, MDSL) 通过多维度约束预测误差, 动态调整权重并平衡各损失的贡献, 解决单一损失无法全面评估预测模型质量的问题, 该模块包含多域损失集成与动态梯度均衡器两部分。

2.6.1 多域损失集成

输入预测序列 $\hat{\mathbf{Y}} \in \mathbb{R}^{N \times H \times C}$ 与真实序列 $\mathbf{Y} \in \mathbb{R}^{N \times H \times C}$, 在传统 MSE 基础上, 引入频域损失、结构损失和平滑性损失, 实现对预测序列多维属性的全面约束。

时域损失的目标是平滑预测结果的数值拟合精度与异常值鲁棒性。MSE 对全局数值偏差敏感, 可高效反映整体偏离程度; MAE 对异常值鲁棒性, 可避免极端值主导损失优化。二者协同可全面约束时域误差, 公式表示为

$$\mathcal{L}_{\text{time}} = \text{MSE}(\hat{\mathbf{Y}}, \mathbf{Y}) + \text{MAE}(\hat{\mathbf{Y}}, \mathbf{Y})$$

频域损失聚焦于保证预测序列与真实序列的频谱一致性, 包括幅度谱损失与相位谱损失约束。首先对 $\hat{\mathbf{Y}}$ 与 $\mathbf{Y}$ 执行 FFT, 获取频域复值表示, 利用幅度谱损失确保预测序列与真实序列的频率强度一致。同时, 相位谱损失使用正弦转换避免相位模糊约束相对相位关系一致。最后, 依托 DGE 实现动态权重调整, 公式表示为

$$\begin{cases} \mathcal{L}_{\text{amp}} = \text{MSE}(|\hat{\mathbf{Y}}_{\text{ft}}|, |\mathbf{Y}_{\text{ft}}|) \\ \mathcal{L}_{\text{phase}} = \text{MSE}(\sin(\angle \hat{\mathbf{Y}}_{\text{ft}}), \sin(\angle \mathbf{Y}_{\text{ft}})) \\ \mathcal{L}_{\text{freq}} = w_{\text{amp}} \cdot \mathcal{L}_{\text{amp}} + w_{\text{phase}} \cdot \mathcal{L}_{\text{phase}} \end{cases}$$

式中: $|\cdot|$ 为复数模值运算; $\angle \cdot$ 表示复数相位角运算; $w_{\text{amp}}, w_{\text{phase}} \in [0, 1]$ 为动态权重, 满足 $w_{\text{amp}} + w_{\text{phase}} = 1$ 。

结构损失从趋势、相对幅度和极值点 3 个维度约束序列形态一致性, 依托 DGE 实现动态权重调整。趋势损失通过一阶差分刻画序列变化方

向, 约束预测与真实趋势一致, 公式表示为

$$\mathcal{L}_{\text{trend}} = \text{MSE}(\Delta \hat{Y}, \Delta Y)$$

式中 Δ 为一阶差分运算, $\Delta a_{n,h} = a_{n,h} - a_{n,h-1}$ ($h \geq 2$)。

极值点损失能约束预测与真实序列的最大值、最小值的匹配, 保留局部关键形态, 公式表示为

$$\mathcal{L}_{\text{extreme}} = \text{MSE}(\max(\hat{Y}), \max(Y)) + \text{MSE}(\min(\hat{Y}), \min(Y))$$

式中: $\max(a)_n = \max_{1 \leq h \leq H} a_{n,h}$, $\min(a)_n = \min_{1 \leq h \leq H} a_{n,h}$ 。

相对幅度损失通过均值归一化维持序列比例关系, 避免全局缩放导致的形态失真, 公式表示为

$$\mathcal{L}_{\text{amp-rel}} = \text{MSE}\left(\frac{\hat{Y}}{\bar{Y}}, \frac{Y}{\bar{Y}}\right)$$

由 DGE 赋予 3 个维度的自适应权重, 公式表示为

$$\mathcal{L}_{\text{struct}} = w_{\text{trend}} \cdot \mathcal{L}_{\text{trend}} + w_{\text{extreme}} \cdot \mathcal{L}_{\text{extreme}} + w_{\text{amp-rel}} \cdot \mathcal{L}_{\text{amp-rel}}$$

式中 $w_{\text{trend}}, w_{\text{extreme}}, w_{\text{amp-rel}} \in [0, 1]$ 为动态权重, 满足 $w_{\text{trend}} + w_{\text{extreme}} + w_{\text{amp-rel}} = 1$ 。

2.6.2 动态梯度均衡器

DGE 根据训练过程中的损失变化自适应调整各损失分量的权重, DGE 通过实时梯度贡献调整权重, 梯度大的损失项权重升高, 使模型在不同阶段聚焦于最需优化的特征, 提升训练的稳定性。

梯度范数计算, 计算每个损失对模型参数的梯度范数, 衡量损失对参数更新的影响, 公式表示为

$$g_i = \|\nabla_{\theta} \mathcal{L}_i\|_2, \quad i \in I, I = \{\text{time}, \text{freq}, \text{struct}\}$$

权重自适应更新, 通过温度参数 α 控制权重灵敏度, 梯度范数越大, 对应权重越高, 公式表示为

$$w_i = \frac{\exp(\alpha \cdot g_i)}{\sum_{j \in I} \exp(\alpha \cdot g_j)}, \quad i \in I$$

式中 α 过大会导致权重集中于单一损失, 过小则权重趋同。总损失函数: 整合时域、频域及结构域的损失, 公式表示为

$$\mathcal{L}_{\text{total}} = \sum_{i \in I} w_i \cdot \mathcal{L}_i$$

3 实验与结果分析

3.1 数据集

本文实验数据集采用 8 个公开的时序数据集:

1) 变压器温度 (ETT) 数据集, 分别从 2 个电力站的变压器采集的 ETTh1 和 ETTh2 数据集, 各含有 2 个不同采集频率的版本 ETTm(15 min) 和 ETTh(60 min)。

2) 电力 (Electricity) 数据集, 含 321 用户的用

电量数据, 反映了住宅和工业的使用情况。

3) 汇率 (Exchange) 数据集, 含 8 个不同国家的每日汇率。

4) 交通 (Traffic) 数据集, 含 862 个传感器捕捉交通流量和入住率的动态。

5) 天气 (Weather) 数据集, 含德国一年范围内的 21 个气象指标。

表 1 标明了 8 个数据集的特征细节, ETT 数据集依据 6:2:2 的比例划分训练集、验证集和测试集, 其他数据集按照 7:1:2 的比例划分。

表 1 8 个数据集的特征细节
Table 1 Feature details for 8 datasets

数据集	长度	数据维度	采样频率	领域
ETTh1/2	69 680	7	15 min	能源
ETTh1/h2	17 420	7	1 h	能源
Electricity	26 304	321	1 h	能源
Exchange	7 588	8	1 d	金融
Traffic	17 544	862	1 h	交通
Weather	52 696	21	10 min	天气

3.2 基准模型与评价指标

本文实验选取了 6 个基准模型作为对比, 包括 FilterTS^[31]、TimeMixer^[32]、iTransformer^[19]、PatchTST^[10]、TimesNet^[22]、FreTS^[21]。由于 FilterTS 达到了当前模型的最优性能, 本实验主要以 FilterTS 作为基准模型比较性能。

3.3 实现细节

实验中的深度学习网络是基于 Python3.8 环境、PyTorch 版本 1.13 和 CUDA 11.8 实现的, 并在 16 GB 内存的 NVIDIA GeForce GTX 3060 GPU 上进行训练。

实验的配置框架将回顾窗口设置为 336 个时间步, 对每个数据集预测的输出长度范围为 {96, 192, 336, 720}, 使用 Adam 优化器进行优化。每轮训练结束后学习率减半, 丢失率设置为 0.15; 实验统一采用 32 的固定批次大小与 10 轮总训练周期。利用系统的网格搜索来优化初始学习速率, 搜索集合为 {0.0001, 0.0005, 0.001}。

3.4 实验结果

表 2 给出了各模型在 8 个数据集上的 MSE 和 MAE 均值, 其中最佳结果以粗体突出显示, 次佳结果以下划线显示。实验结果表明, 在所有预测长度上的平均 MSE, CFuseTS 在 8 个数据集上的 7 个上取得了最优性能。在 Exchange 数据集中, 模型性能出现一定程度的下降, 这一现象主要归因于汇率数据的长程依赖特性更为复杂, 复

值滤波机制对这类强非线性长程关联的建模存在一定的局限性。相比之下, 像 FilterTS、PatchTST 这类的模型通过灵活的频率筛选或序列分块策略, 更适配汇率数据的长程交互场景, 而 iTransformer 等模型利用复杂注意力机制动态调整权重, 也能适应多变量间的复杂交互场景。尽管面临一定挑战, CFuseTS 的平均 MSE 比次优的基准模型 FilterTS 降低 2.43%, 验证了复值相位感知对强周期特征捕捉能力的提升。同时, CFuseTS 与通道独立模型的代表 PatchTST 相比平均 MSE 降

低了 6.66%, 这表明其对变量间共享信息的有效挖掘, 可以显著提高模型性能。其次, CFuseTS 在预测性能上优于采用 MLP 结构的 TimeMixer 和使用注意力机制的 iTransformer, 将平均 MSE 分别降低了 6.07% 和 3.95%, 这表明通过滤波机制提取并融合多变量间共有频率分量的方法比简单的融合策略更有效。此外, CFuseTS 在所有任务中超过 FreTS, 进一步表明其性能优势不仅来自其在频域的特征构建, 还来自其对冗余频率的有效过滤策略, 实现了频域信息的精准利用。

表 2 8 个数据集上的多元长时间序列预测结果
Table 2 Prediction results of multivariate long-term series on 8 datasets

数据集	输出长度	CFuseTS		FilterTS		TimeMixer		iTransformer		PatchTST		TimesNet		FreTS	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
ETTm1	96	0.298	0.349	<u>0.321</u>	<u>0.360</u>	0.333	0.371	0.334	0.368	0.329	0.367	0.338	0.375	0.355	0.371
	192	0.341	0.375	<u>0.363</u>	<u>0.382</u>	0.367	0.386	0.377	0.391	0.367	0.385	0.374	0.387	0.377	0.394
	336	0.378	0.396	<u>0.395</u>	<u>0.403</u>	0.397	0.408	0.426	0.420	0.399	0.410	0.410	0.411	0.413	0.418
	720	0.437	0.429	0.462	<u>0.438</u>	0.460	0.445	0.491	0.459	<u>0.454</u>	0.439	0.478	0.450	0.483	0.461
	平均	0.364	0.387	<u>0.385</u>	<u>0.396</u>	0.389	0.403	0.407	0.410	0.387	0.400	0.400	0.406	0.402	0.411
ETTm2	96	0.176	0.262	0.172	0.255	<u>0.174</u>	<u>0.258</u>	0.180	0.264	0.175	0.259	0.187	0.267	0.181	0.269
	192	0.233	0.299	<u>0.237</u>	0.299	0.239	<u>0.302</u>	0.250	0.309	0.241	0.302	0.249	0.309	0.249	0.322
	336	0.281	0.333	0.299	<u>0.398</u>	<u>0.296</u>	0.340	0.311	0.348	0.305	0.343	0.321	0.351	0.340	0.382
	720	0.375	0.382	0.397	<u>0.394</u>	<u>0.393</u>	0.397	0.412	0.407	0.402	0.400	0.408	0.403	0.449	0.455
	平均	0.266	0.319	<u>0.276</u>	<u>0.321</u>	<u>0.276</u>	0.324	0.288	0.332	0.281	0.326	0.291	0.333	0.305	0.357
ETTh1	96	0.386	0.406	0.374	0.391	<u>0.384</u>	<u>0.398</u>	0.386	0.405	0.414	0.419	0.384	0.402	0.390	0.404
	192	0.422	<u>0.427</u>	<u>0.424</u>	0.421	0.439	0.429	0.441	0.436	0.460	0.445	0.436	0.429	0.448	0.439
	336	0.444	0.441	<u>0.464</u>	0.441	0.487	0.455	0.487	0.458	0.501	0.466	0.491	0.469	0.501	0.470
	720	<u>0.479</u>	<u>0.477</u>	0.470	0.466	0.502	0.482	0.503	0.491	0.500	0.488	0.512	0.500	0.559	0.535
	平均	0.433	<u>0.438</u>	0.433	0.430	<u>0.453</u>	0.441	0.454	0.448	0.469	0.455	0.458	0.450	0.475	0.462
ETTh2	96	<u>0.293</u>	0.351	0.290	0.338	<u>0.293</u>	<u>0.344</u>	0.297	0.349	0.302	0.348	0.340	0.374	0.317	0.373
	192	0.362	<u>0.393</u>	<u>0.374</u>	0.390	<u>0.374</u>	0.396	0.380	0.400	0.388	0.400	0.402	0.414	0.427	0.442
	336	0.380	0.413	<u>0.406</u>	<u>0.420</u>	0.416	0.430	0.428	0.432	0.426	0.433	0.452	0.452	0.526	0.505
	720	0.412	<u>0.442</u>	<u>0.418</u>	0.437	0.440	0.452	0.427	0.445	0.431	0.446	0.462	0.468	0.684	0.591
	平均	0.362	<u>0.400</u>	<u>0.372</u>	0.396	0.381	0.406	0.383	0.407	0.387	0.407	0.414	0.427	0.489	0.478
Electricity	96	0.138	0.236	0.151	0.245	0.157	0.250	<u>0.148</u>	<u>0.240</u>	0.181	0.270	0.168	0.272	0.177	0.264
	192	0.156	0.252	0.163	0.256	0.177	0.263	<u>0.162</u>	<u>0.253</u>	0.188	0.274	0.184	0.289	0.183	0.271
	336	0.172	0.269	0.180	<u>0.274</u>	0.189	0.282	<u>0.178</u>	0.269	0.204	0.293	0.198	0.300	0.198	0.288
	720	0.210	0.302	0.224	<u>0.311</u>	0.233	0.316	0.255	0.317	0.246	0.324	<u>0.220</u>	0.320	0.234	0.322
	平均	0.169	0.265	0.180	0.271	0.189	0.278	<u>0.178</u>	<u>0.270</u>	0.205	0.290	0.193	0.295	0.255	0.300
Exchange	96	0.093	0.217	0.081	0.199	0.087	<u>0.205</u>	0.086	0.206	0.088	0.205	0.107	0.234	<u>0.085</u>	0.211
	192	0.193	0.314	0.171	0.294	0.179	0.300	0.177	<u>0.299</u>	<u>0.176</u>	<u>0.299</u>	0.226	0.344	0.181	0.312
	336	0.363	0.441	<u>0.321</u>	<u>0.409</u>	0.333	0.417	0.331	0.417	0.301	0.397	0.367	0.448	0.471	0.508
	720	0.968	0.737	0.837	0.688	0.912	0.719	<u>0.847</u>	<u>0.691</u>	0.901	0.714	0.964	0.746	0.858	0.695
	平均	0.404	0.427	0.352	0.397	0.378	0.410	<u>0.360</u>	<u>0.403</u>	0.367	0.404	0.416	0.443	0.399	0.432

续表 2

数据集	输出长度	CFuseTS		FilterTS		TimeMixer		iTransformer		PatchTST		TimesNet		FreTS	
		MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE	MSE	MAE
Traffic	96	0.383	0.274	0.448	0.309	0.485	0.323	<u>0.395</u>	<u>0.268</u>	0.462	0.295	0.593	0.321	0.512	0.328
	192	0.401	<u>0.282</u>	0.455	0.307	0.488	0.322	<u>0.417</u>	0.276	0.466	0.296	0.617	0.336	0.507	0.325
	336	0.416	<u>0.290</u>	0.472	0.313	0.507	0.321	<u>0.433</u>	0.283	0.482	0.304	0.629	0.336	0.521	0.331
	720	0.452	<u>0.312</u>	0.508	0.332	0.549	0.335	<u>0.467</u>	0.302	0.514	0.322	0.640	0.350	0.562	0.348
	平均	0.413	<u>0.290</u>	0.471	0.315	0.507	0.325	<u>0.428</u>	0.282	0.481	0.304	0.620	0.336	0.526	0.333
Weather	96	0.158	<u>0.211</u>	<u>0.162</u>	0.207	0.166	0.213	0.174	0.214	0.177	0.218	0.172	0.220	0.182	0.236
	192	0.209	0.256	0.209	<u>0.252</u>	0.209	0.251	0.221	0.254	0.225	0.259	0.219	0.261	0.219	0.217
	336	0.253	0.286	<u>0.263</u>	<u>0.292</u>	0.264	0.293	0.278	0.296	0.278	0.297	0.280	0.306	0.270	0.313
	720	0.324	0.336	0.344	0.344	<u>0.342</u>	<u>0.343</u>	0.358	0.349	0.354	0.348	0.365	0.359	0.348	0.380
	平均	0.236	0.272	<u>0.244</u>	<u>0.274</u>	0.245	0.275	0.258	0.278	0.259	0.281	0.259	0.287	0.255	0.300
最优结果的次数		31	20	10	16	0	1	0	5	1	1	0	0	0	0

注: 表中加粗数值代表最优结果, 下划线数值代表次优结果。

3.5 消融实验

为了验证 CFuseTS 模块设计的有效性, 本文在 ETTm1、Weather、Electricity 3 个数据集上进行了消融实验。具体来说, 评估了以下变体: w/o-PACF(删除相位感知复值滤波模块) 和 w/o-MDSL(删除多域协同损失模块)。

表 3 给出了消融实验的结果, 可以观察到, 删

除相位感知复值滤波模块后, 模型在 3 个数据集的平均 MSE 分别上升 0.3%、2.1%、1.3%, 平均性能下降约 1.23%, 这验证了复数域建模, 多尺度滤波方法的必要性; 删除多域协同损失模块后, 模型的平均 MSE 分别上升 1.0%、1.6%、1.5%, 平均性能下降约 1.37%, 证明了多域约束对模型预测精度的提升价值。

表 3 消融实验结果
Table 3 Ablation experimental results

数据集	输出长度	CFuseTS		w/o PACF		w/o MDSL	
		MSE	MAE	MSE	MAE	MSE	MAE
ETTM1	96	0.298	0.349	0.338	0.370	0.344	0.375
	192	0.341	0.375	0.345	0.370	0.345	0.370
	336	0.378	0.396	0.345	0.413	0.379	0.390
	720	0.437	0.429	0.438	0.423	0.428	0.423
	平均	0.364	0.387	0.367	0.394	0.374	0.390
Weather	96	0.158	0.211	0.172	0.215	0.192	0.232
	192	0.209	0.256	0.219	0.261	0.226	0.263
	336	0.253	0.286	0.268	0.301	0.266	0.296
	720	0.324	0.336	0.368	0.373	0.322	0.341
	平均	0.236	0.272	0.257	0.288	0.252	0.283
Electricity	96	0.138	0.236	0.165	0.264	0.190	0.270
	192	0.156	0.252	0.164	0.263	0.163	0.256
	336	0.172	0.269	0.186	0.284	0.172	0.268
	720	0.210	0.302	0.212	0.315	0.210	0.302
	平均	0.169	0.265	0.182	0.282	0.184	0.274

4 结束语

本文提出了一种基于复值相位感知与多域协

同损失的多元长时间序列预测模型 CFuseTS。它以复数域建模为核心, 通过多尺度复值滤波模块提取变量间的频率与相位特征, 结合时域、频域、

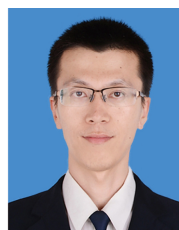
结构域的动态损失进行约束,实现对复杂时序数据多维度特征的协同优化,突破了传统实数域模型丢失相位特征的局限。在8个公开数据集上的实验表明了CFuseTS模型的有效性。未来研究将专注于轻量化模型架构设计与多变量预测扩展,以提升模型的实用性与泛化能力。

参考文献:

- [1] 李毅,彭晋卿,廖维,等.一种基于时间序列的集成电力负荷预测方法研究[J]. *建筑科学*, 2022, 38(10): 190–197.
LI Yi, PENG Jinqing, LIAO Wei, et al. Research on an integrated model for electrical load forecasting based on time series[J]. *Building science*, 2022, 38(10): 190–197.
- [2] 骆钊,吴谕侯,朱家祥,等.基于多尺度时间序列块自编码Transformer神经网络模型的风电超短期功率预测[J]. *电网技术*, 2023, 47(9): 3527–3537.
LUO Zhao, WU Yuhou, ZHU Jiaxiang, et al. Wind power forecasting based on multi-scale time series block auto-encoder transformer neural network model[J]. *Power system technology*, 2023, 47(9): 3527–3537.
- [3] 梁达,商越,王兆芬,等.青海省肺结核发病与气象因素的时间序列分析[J]. *中华疾病控制杂志*, 2021, 25(10): 1186–1193.
LIANG Da, SHANG Yue, WANG Zhaofen, et al. Time series analysis of tuberculosis incidence and meteorological factors in Qinghai Province[J]. *Chinese Journal of Disease Control and Prevention*, 2021, 25(10): 1186–1193.
- [4] 杨文忠,张志豪,吾守尔·斯拉木,等.基于时间序列关系的GBRT交通事故预测模型[J]. *电子科技大学学报*, 2020, 49(4): 615–621.
YANG Wenzhong, ZHANG Zhihao, WUSHOUER Silamu, et al. GBRT traffic accident prediction model based on time series relationship[J]. *Journal of University of Electronic Science and Technology of China*, 2020, 49(4): 615–621.
- [5] 冯强,赵建光,杨茸,等.时间序列中非平稳性和波动性的建模及预测[J]. *计算机科学与探索*, 2025, 19(5): 1313–1321.
FENG Qiang, ZHAO Jianguang, YANG Rong, et al. Modeling and predicting time series with non-stationarity and volatility[J]. *Journal of frontiers of computer science and technology*, 2025, 19(5): 1313–1321.
- [6] 徐任超,阎威武,王国良,等.基于周期性建模的时间序列预测方法及电价预测研究[J]. *自动化学报*, 2020, 46(6): 1136–1144.
XU Renchao, YAN Weiwu, WANG Guoliang, et al. Time series forecasting based on seasonality modeling and its application to electricity price forecasting[J]. *Acta automatica sinica*, 2020, 46(6): 1136–1144.
- [7] 李洁,林永峰.基于多时间尺度RNN的时序数据预测[J]. *计算机应用与软件*, 2018, 35(7): 33–37, 62.
LI Jie, LIN Yongfeng. Prediction of time series data based on multi-time scale rnn[J]. *Computer applications and software*, 2018, 35(7): 33–37, 62.
- [8] 陈嘉俊,刘波,林伟伟,等.基于Transformer的时间序列预测方法综述[J]. *计算机科学*, 2025, 52(6): 96–105.
CHEN Jiajun, LIU Bo, LIN Weiwei, et al. Survey of transformer-based time series forecasting methods[J]. *Computer science*, 2025, 52(6): 96–105.
- [9] ZHOU Haoyi, ZHANG Shanghang, PENG Jieqi, et al. In-former: beyond efficient transformer for long sequence time-series forecasting[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. Vancouver: AAAI, 2021: 11106–11115.
- [10] NIE Yuqi, NGUYEN N H, SINHTONG P, et al. A time series is worth 64 words: long-term forecasting with transformers[EB/OL]. (2023–03–05)[2025–11–19]. <https://doi.org/10.48550/arXiv.2211.14730>.
- [11] LI Chengjian, SHI Zhenghao, ZHOU Liang, et al. TF-Former: a time–frequency information fusion-based CNN-transformer model for OSA detection with single-lead ECG[J]. *IEEE transactions on instrumentation and measurement*, 2023, 72: 1–17.
- [12] JIN Kaiwen, XIONG Yifeng, LOU Shuya, et al. MAFD: multiple adversarial features detector for enhanced detection of text-based adversarial examples[J]. *Neural processing letters*, 2024, 56(6): 251.
- [13] LI Zhe, LI Xuanxuan, RAO Zhongwen, et al. MTS-mixers: multivariate time series forecasting via factorized temporal and channel mixing[C]//*2025 International Joint Conference on Neural Networks*. Rome: IEEE, 2025: 1–8.
- [14] BELTRAN A O, SUAZA-MEDINA M E, ZARAZAGA-SORIA F J, et al. Multivariate integration of time series with ML for corn price forecasting in Colombia[J]. *Expert systems with applications*, 2026, 299: 129822.
- [15] AASIM, SINGH S N, MOHAPATRA A. Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting[J]. *Renewable energy*, 2019, 136: 758–768.
- [16] KUMAR DUBEY A, KUMAR A, GARCÍA-DÍAZ V, et al. Study and analysis of SARIMA and LSTM in forecasting time series data[J]. *Sustainable energy technologies and assessments*, 2021, 47: 101474.
- [17] SHERSTINSKY A. Fundamentals of recurrent neural net-

- work (RNN) and long short-term memory (LSTM) network[J]. *Physica D: nonlinear phenomena*, 2020, 404: 132306.
- [18] YU Yong, SI Xiaosheng, HU Changhua, et al. A review of recurrent neural networks: LSTM cells and network architectures[J]. *Neural computation*, 2019, 31(7): 1235–1270.
- [19] LIU Yong, HU Tengge, ZHANG Haoran, et al. iTransformer: inverted transformers are effective for time series forecasting[EB/OL]. (2024–03–14)[2025–11–19]. <https://doi.org/10.48550/arXiv.2310.06625>.
- [20] ZHOU Tian, MA Ziqing, WANG Xue, et al. Fedformer: frequency enhanced decomposed transformer for long-term series forecasting[M]//*AI for Time Series*. Boca Raton: CRC Press, 2026: 10–34.
- [21] YI Kun, ZHANG Qi, FAN Wei, et al. Frequency-domain MLPs are more effective learners in time series forecasting[C]//*Proceedings of the 37th International Conference on Neural Information Processing Systems*. New Orleans: Curran Associates Inc., 2023: 76656–76679.
- [22] WU Haixu, HU Tengge, LIU Yong, et al. TimesNet: temporal 2D-variation modeling for general time series analysis[EB/OL]. (2023–04–12)[2025–11–19]. <https://doi.org/10.48550/arXiv.2210.02186>.
- [23] ZHANG J A, LIU Fan, MASOUIROS C, et al. An overview of signal processing techniques for joint communication and radar sensing[J]. *IEEE journal of selected topics in signal processing*, 2021, 15(6): 1295–1315.
- [24] CHEN Haotian, NATSUAKI R, HIROSE A. Robust symbol detection based on quaternion neural networks in wireless polarization-shift-keying communications[J]. *IEEE transactions on neural networks and learning systems*, 2024, 35(11): 16064–16075.
- [25] LIU Mengxi, CHAI Zhuoqun, DENG Haojun, et al. A CNN-transformer network with multiscale context aggregation for fine-grained cropland change detection[J]. *IEEE journal of selected topics in applied earth observations and remote sensing*, 2022, 15: 4297–4306.
- [26] YANG Muqiao, MA M Q, LI Dongyu, et al. Complex transformer: a framework for modeling complex-valued sequence[C]//*ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing*. Barcelona: IEEE, 2020: 4232–4236.
- [27] KIM T Y, CHO S B. Web traffic anomaly detection using C-LSTM neural networks[J]. *Expert systems with applications*, 2018, 106: 66–76.
- [28] TIAN Zhen, ZHAO W X, ZHANG Changwang, et al. EulerFormer: sequential user behavior modeling with complex vector attention[C]//*Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*. Washington DC: ACM, 2024: 1619–1628.
- [29] KOTHAPALLY V, HANSEN J H L. Complex-valued time-frequency self-attention for speech dereverberation[EB/OL]. (2022–11–22)[2025–11–19]. <https://doi.org/10.48550/arXiv.2211.12632>.
- [30] XIU Hao, HU Siran, CUI Yuanxin, et al. Sliding window regression method for mechanical meter reading recognition in time-series images[J]. *Engineering applications of artificial intelligence*, 2025, 158: 111240.
- [31] WANG Yulong, LIU Yushuo, DUAN Xiaoyi, et al. FilterTS: comprehensive frequency filtering for multivariate time series forecasting[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. Philadelphia: AAAI, 2025: 21375–21383.
- [32] WANG Shiyu, WU Haixu, SHI Xiaoming, et al. Time-Mixer: decomposable multiscale mixing for time series forecasting[EB/OL]. (2024–05–23)[2025–11–19]. <https://doi.org/10.48550/arXiv.2405.14616>.

作者简介:



郝剑龙, 副教授, 博士, 主要研究方向为深度学习、时间序列分析。主持国家自然科学基金青年项目 1 项, 参与省部级课题多项。E-mail: haojianlong2012@sxufe.edu.cn。



杜银焕, 硕士研究生, 主要研究方向为时间序列分析、人工智能。E-mail: duyinhuan2024@163.com。



薛令荣, 副教授, 博士, 主要研究方向为随机非线性系统控制。主持国家自然科学基金青年项目 1 项、省部级课题 2 项, 发表学术论文 20 余篇。E-mail: 20181058@sxufe.edu.cn。