

# 控制与决策

Control and Decision

## 基于隐连续性恢复的粗粒度离散标签混合时序预测方法

孙曦, 赵春晖, 张皓然

引用本文:

孙曦, 赵春晖, 张皓然. 基于隐连续性恢复的粗粒度离散标签混合时序预测方法[J]. *控制与决策*, 2026, 41(5): 1457–1467.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2025.0931>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 一种基于池计算的宽度学习系统

A broad learning system based on reservoir computing

控制与决策. 2021, 36(9): 2203–2210 <https://doi.org/10.13195/j.kzyjc.2019.1729>

#### 基于深度时序特征迁移的轴承剩余寿命预测方法

Remaining useful life prediction of bearing based on deep temporal feature transfer

控制与决策. 2021, 36(7): 1699–1706 <https://doi.org/10.13195/j.kzyjc.2019.1809>

#### 一种基于深度学习的时间序列预测方法

A time series prediction method based on deep learning

控制与决策. 2021, 36(3): 645–652 <https://doi.org/10.13195/j.kzyjc.2019.0809>

#### 考虑退化轨迹差异性与相似性的轴承RUL预测

Prediction of bearing remaining useful life involving difference and similarity of degradation trajectories

控制与决策. 2021, 36(11): 2833–2840 <https://doi.org/10.13195/j.kzyjc.2020.1028>

#### 随机变批次长度的反馈辅助PD型量化迭代学习控制

Feedback-assisted PD-type quantized iterative learning control with randomly iteration varying lengths

控制与决策. 2021, 36(10): 2569–2576 <https://doi.org/10.13195/j.kzyjc.2020.0273>

# 基于隐连续性恢复的粗粒度离散标签混合时序预测方法

孙曦<sup>1</sup>, 赵春晖<sup>1,2,3†</sup>, 张浩然<sup>2,3</sup>

- 华东交通大学 电气与自动化工程学院, 南昌 330013;
- 浙江大学 控制科学与工程学院, 杭州 310027;
- 浙江大学 工业控制技术全国重点实验室, 杭州 310027)

**摘要:** 受传感器精度、数据存储以及人工标注效率等因素限制, 工业场景中常常使用粗粒度的离散标签替代原本连续的值。然而, 离散标签通常会造成信息损失, 导致其背后隐含的连续动态过程难以准确建模; 此外, 离散变量与连续变量在分布特性和信息粒度上存在差异, 无法实现统一的预测。针对上述问题, 提出一种基于隐连续性恢复的粗粒度离散标签混合时序预测方法 (CDL-HCR)。首先, 所提出方法通过隐连续性恢复策略, 利用可观测的离散标签推断不可观测的连续变量, 从而实现原本不可测连续变量的估计。然后, 结合自监督学习机制, 将恢复得到的隐连续变量与原有连续变量在统一空间中进行特征融合, 以提升混合时序数据的预测精度。在线应用阶段, 模型可直接输入实时采集的混合数据进行预测。最后, 在真实的滚筒烘丝过程数据集上验证所提出方法, 验证结果显示其在混合时序数据预测任务中相比于传统方法具有更高的准确性。

**关键词:** 隐连续性恢复; 自监督学习; 混合时序预测; 高斯核嵌入

中图分类号: TP183 文献标志码: A

DOI: 10.13195/j.kzyjc.2025.0931

引用格式: 孙曦, 赵春晖, 张浩然. 基于隐连续性恢复的粗粒度离散标签混合时序预测方法 [J]. 控制与决策, 2026, 41(5): 1457-1467.

## Coarse-grained discretization label hybrid time series prediction method based on hidden continuity recovery

SUN Xi<sup>1</sup>, ZHAO Chun-hui<sup>1,2,3†</sup>, ZHANG Hao-ran<sup>2,3</sup>

- College of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China;
- College of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China;
- State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China)

**Abstract:** Due to limitations such as sensor accuracy, data storage capacity and the efficiency of manual labelling, coarse-grained discrete labels are often used instead of the original continuous values in industrial scenarios. However, discrete labels typically result in information loss, making it challenging to accurately model the underlying continuous dynamic processes. Additionally, differences in distribution characteristics and information granularity exist between discrete and continuous variables, preventing unified prediction. To address these issues, a hybrid time series prediction method called coarse-grained discretized labels-hidden continuity recovery (CDL-HCR) is proposed. Using the hidden continuity recovery strategy, the method infers unobservable continuous variables from observable discrete labels in order to estimate originally unmeasurable continuous variables. Then, combined with a self-supervised learning mechanism, the recovered hidden continuous variables and the original continuous variables are fused in a unified space to improve the prediction accuracy of mixed time series data. During the online application stage, the model can input real-time mixed data directly for prediction. Finally, the proposed method is verified using real drum drying process data, and the results show that it is more accurate than traditional methods for predicting mixed time series data.

**Keywords:** hidden continuity recovery; self-supervised learning; hybrid time series forecasting; Gaussian kernel embedding

## 0 引言

随着工业系统的复杂化和智能化, 传统传感器

技术在实时监测和数据采集方面面临许多局限<sup>[1]</sup>. 为应对这些挑战, 预测技术应运而生, 它通过对历史数

收稿日期: 2025-09-04; 录用日期: 2025-12-19.

基金项目: 江西省自然科学基金项目 (20242BAB27003); 国家自然科学基金原创探索项目 (62450020); 工业控制技术全国重点实验室浙大专项项目 (ICT2025C01); 工业控制技术全国重点实验室项目 (ICT2024A06).

†通信作者. E-mail: chzhao@zju.edu.cn.

据和实时信息的深度分析,能够预测那些难以直接测量或未来的关键变量<sup>[2]</sup>.作为智能化系统的重要组成部分,预测技术不仅提高了系统的决策能力,也推动了各行业在数字化转型中的发展<sup>[3]</sup>.

在实际工业应用中,由于传感器测量精度、数据存储能力和人工标注效率等因素的限制,许多原本连续变化的变量被压缩为粗粒度的离散标签形式(如高/中/低或开/关状态)<sup>[4-5]</sup>.这种离散化处理使得连续变量表现出离散特征,其本质上是通过对离散信号近似表示连续变量的状态变化<sup>[6]</sup>.如:在智能制造过程中,由于高温、高压等关键变量难以直接测量,通常以报警信号的形式指示是否超出设定的控制限,从而间接表征变量状态;在质量控制环节,产品性能指标通常以等级标签代替具体数值,既提升了标注效率,也降低了人为标注带来的一致性风险<sup>[7-9]</sup>.此类离散标签已广泛应用于工业数据采集和过程监测,并常与其他实测的连续变量共同构成混合数据.

尽管离散标签在工程实践中具有一定的应用价值,但是,其数据粒度过粗不可避免地会导致信息损失.一方面,粗粒度标签难以刻画连续过程中的细微动态变化;另一方面,离散标签与连续变量在混合建模中存在分布特性和信息粒度的差异,使得传统面向单一数据类型的预测模型难以有效处理此类混合数据<sup>[10]</sup>.特别是在混合时序预测中,离散标签的非连续性和稀疏性与连续变量的平滑性和时序依赖性间的差异,增加了模型预测的复杂度<sup>[11]</sup>.此外,在许多预测任务中,关键的离散标签通常不可直接观测,进一步加剧了混合时间序列建模与预测的复杂性和不确定性.

为解决工业场景中离散标签与连续变量共存且难以实现统一有效建模和预测的问题.传统方法如偏最小二乘回归(PLSR)<sup>[12]</sup>、自回归模型(AR)<sup>[13]</sup>以及支持向量回归(SVR)<sup>[14]</sup>等主要聚焦于连续变量的建模,通常将离散标签视为噪声或辅助特征,未能充分挖掘其潜在的结构信息.在处理连续变量时,传统方法表现出良好的建模性能;但是在面对离散标签时,往往会忽视其潜在的时序特性,从而导致关键信息的损失和预测性能的下降.为克服这一问题,近年来,越来越多的研究开始引入深度学习方法来处理复杂时间序列任务.神经网络具备强大的非线性建模能力和特征自动提取能力,受到了越来越多研究者的青睐,已被广泛应用于时序预测场景中,典型模型包括多层感知机(MLP)<sup>[15]</sup>、长短期记忆网络(LSTM)<sup>[16]</sup>和时域卷积网络(TCN)<sup>[17]</sup>等.然而,现有模型多针对连续变量设计,在处理离散标签时仍然存在建模机制单一、信息利用率低等问题.为实现离

散标签和连续变量在统一空间下的有效建模, Tan等<sup>[18]</sup>提出了一种基于变分自编码器(VAE)的方法,旨在用于信用数据合成,该方法通过构建潜在空间并从中采样生成数据,能够有效地同时处理离散特征和连续特征,但是,仍然依赖于强先验的指导以确保模型的有效性;Chen等<sup>[19]</sup>提出了一种MiTCD框架,旨在处理混合时序数据,缓解时空异质性问题,取得了优异的预测性能,然而,该方法依赖于较强的监督信号,在高噪声数据下泛化能力有限;Zhao等<sup>[20]</sup>提出了因果解耦方法,旨在改善混合时序数据在预测任务中的表现,然而,该方法依赖显式因果结构建模,难以充分刻画隐式的连续动态过程;Cheng等<sup>[21]</sup>提出了融合离散变量与连续变量模型的框架来提升混沌时序的长期预测能力,但是其方法依赖显式动力学结构,对于缺乏明确物理机理的场景适用性有限;Sedgewick等<sup>[22]</sup>提出了混合图模型来融合离散与连续变量,从而增强了多模态因果分析能力,但是,该方法面向静态混合数据设计,难以刻画时序依赖,对混合时序预测任务中的动态关联和噪声扰动适应性有限.

综上所述,现有方法未能充分挖掘离散标签背后隐含的连续过程,从而无法有效地对离散与连续变量统一建模.为此,本文提出一种基于隐连续性恢复的粗粒度离散标签混合时序预测方法(CDL-HCR).其核心思想是在统一的连续空间中通过多尺度高斯核嵌入和自监督学习等策略,实现离散标签与连续变量的联合建模.该方法利用可观测的离散标签信息推测不可观测的连续过程,通过隐连续性恢复直接实现对原不可测连续变量的估计,并将隐连续变量与连续变量进行特征融合,从而提升混合时序数据的预测能力.本文主要内容如下:

1) 针对离散标签无法直接反映连续动态过程的问题,本文提出一种基于上下文自适应高斯核嵌入的隐连续恢复机制.相比于传统单一建模方式,该机制在有效恢复离散标签背后的潜在连续过程,可实现对原不可观测连续变量的估计,从而提高模型在处理复杂动态数据时的准确性和稳定性.

2) 针对混合时序数据建模中离散标签与连续变量在分布特性和信息粒度上的差异,本文提出一种基于隐连续性恢复的混合时序统一预测方法.该方法能够同时实现离散标签有效恢复并预测连续变量,相较于传统方法,能够更好地实现对混合时序数据的联合建模和预测.

3) 针对连续变量标注数据稀缺和模型泛化能力不足的问题,本文设计一种自监督一致性约束学习机制.与现有方法依赖大量标注数据的方法不同,该

机制通过引入自监督一致性约束, 利用无标签数据的时序结构, 自主学习离散标签与潜在连续过程间的映射关系, 从而减少对标注数据的依赖, 并提升模型在混合时序数据预测中的学习能力和泛化性.

### 1 CDL-HCR: 基于隐连续性恢复的粗粒度离散标签混合时序预测方法

本节首先对所研究问题及其动机进行详细阐述,

然后介绍所提出 CDL-HCR 方法, 并展示其整体结构框架, 如图 1 所示. 该框架由隐连续性恢复模块和混合时序自监督预测器两部分组成. 接着, 将在后文第 1.2 节和第 1.3 节详细阐述 CDL-HCR 方法中的隐连续性恢复模块以及混合时序自监督预测器模块. 最后, 将在后文第 1.4 节说明整体算法的实现流程.

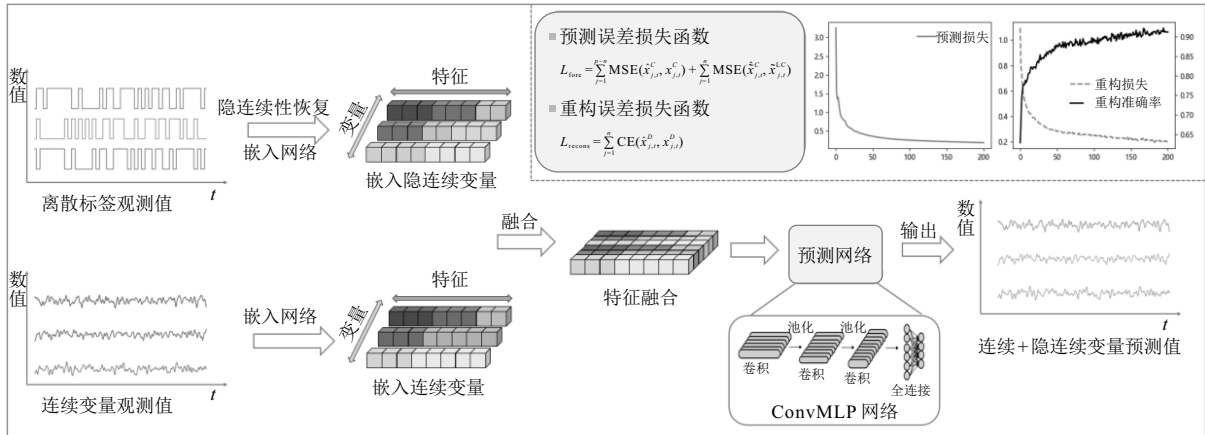


图1 CDL-HCR 整体结构框图

#### 1.1 问题阐述与研究动机

在工业与信息系统中, 由于传感器精度、存储能力和人工标注效率等限制, 许多连续变化的过程变量常被压缩为粗粒度的离散标签形式, 以降低数据采集和处理成本. 如: 在由  $p$  个观测到的时间序列所组成的非线性动态系统中, 时间序列的长度为  $T$ , 整个时间序列表示为  $X = \{x_1, x_2, \dots, x_p\} \in \mathbb{R}^{p \times T}$ . 其中: 包含  $p - n$  个连续变量  $X^C = \{x_1, x_2, \dots, x_{p-n}\} \in \mathbb{R}^{(p-n) \times T}$ , 以及  $n$  个离散标签  $X^D = \{x_1, x_2, \dots, x_n\} \in \mathbb{A}^{n \times T}$ . 根据已有研究, 离散标签是通过离散化过程从潜在的连续变量中获得的, 离散状态数量用  $N_{DLS}$  表示, 离散标签取值范围为  $0 \sim N_{DLS} - 1$  之间的整数, 即  $\mathbb{A} = \{0, 1, \dots, N_{DLS} - 1\}$ .

这种离散化方式虽然在一定程度上保留了连续过程的关键特征, 并在智能制造、质量控制以及系统监测等场景中得到了广泛应用, 但是, 不可避免地会造成信息损失, 难以刻画连续过程中的细微动态变化.

在混合数据场景下, 离散标签与连续变量在分布特性和信息粒度上的差异, 使得传统面向单一数据类型预测和建模方法难以直接适用. 离散标签通常表示粗粒度类别信息, 而连续变量则反映更细腻的动态变化. 由于这两者在信息表达上的差异, 传统方法难以同时有效处理这两种数据类型. 若仅依赖连续变量建模, 则可能会丢失关键的离散标签信

息, 导致预测误差和不稳定性. 如何有效融合离散标签与连续变量, 处理它们在信息粒度、分布特性以及时间序列依赖上的差异, 成为混合时序预测中的核心难点. 因此, 亟需面向混合数据特性的建模方法, 以充分融合离散与连续信息, 从而提升混合时序预测的准确性和可靠性.

#### 1.2 离散标签隐连续性恢复策略

离散标签通过对连续变量进行离散化获得, 能够在一定程度上保留连续过程的关键特征. 然而, 离散标签的粗粒度特性难以准确地反映连续过程中的细微变化和动态演变. 为有效恢复离散标签背后潜在的连续信息并弥补离散化过程中的细节丢失, 本文提出一种基于上下文自适应高斯核嵌入 (CAKGE) 的隐连续性恢复策略. 该策略通过高斯核的平滑特性, 将离散标签映射回连续过程的潜在状态, 从而恢复其连续性的动态特征.

高斯核函数<sup>[23]</sup>, 又称径向基函数, 是一类基于样本间距离构建相似性度量的核函数, 广泛应用于模式识别、机器学习、图像处理以及信号建模等领域<sup>[24-25]</sup>. 其核心思想是通过计算输入样本间的欧氏距离, 并采用指数衰减函数进行非线性变换, 在无需显式高维映射的前提下, 实现对输入空间的非线性建模, 借助“核技巧”, 高斯核可将数据隐式映射至高维特征空间, 从而增强模型对非线性关系的表达能力. 高斯核的数学表达式如下所示:

$$\kappa(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right). \quad (1)$$

其中:  $\|x - x'\|^2$  为输入样本点的欧氏距离平方;  $\sigma$  为核函数的带宽参数, 控制核函数对输入数据之间距离的响应敏感度.

为提升模型对不同尺度特征的建模能力, 多尺度激励高斯核引入多个带宽各异的高斯核函数, 对输入数据进行多层次、多粒度的特征表达<sup>[26]</sup>. 表达式为

$$\kappa_{\text{multi}}(x, x') = \sum_{i=1}^d \alpha_i(x, x') \exp\left(-\frac{\|x - x'\|^2}{2\sigma_i^2}\right). \quad (2)$$

其中:  $d$  为高斯核的尺度数量, 即使用了多少个不同带宽的核;  $\sigma_i$  为第  $i$  个尺度下的高斯核带宽;  $\alpha_i(x, x')$  为激励权重系数, 是由上下文学习得到的动态权重.

准确而言, 给定一个离散标签  $x_j^D$ , 每个非 0 值时间点可视为激励信号. 在每个激励点, 带有带宽  $\sigma$  的高斯核被叠加在该时间点周围, 作为对该时间点的影响. 计算公式如下所示:

$$f(\Delta t; \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\Delta t^2}{2\sigma^2}\right). \quad (3)$$

其中:  $\sigma$  为标准差, 也可理解为高斯核的带宽;  $\Delta t^2$  为与激励信号的时间距离. 该高斯函数能够平滑地扩展激励信号的影响, 确保时间上相邻的激励点对最终结果的贡献合理且连续.

考虑到所有激励时间点的影响, 连续值时间序列的恢复可通过叠加所有激励高斯核来实现. 具体而言, 恢复后的时间序列是所有激励信号叠加效应的结果, 其计算公式如下所示:

$$F_{x_j^D}(t; \sigma) = \sum_{t_i=1}^T x_{t_i, j}^D \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(t - t_i)^2}{2\sigma^2}\right). \quad (4)$$

其中:  $t_i$  为每个激励信号对应的时间点,  $x_{t_i, j}^D$  为该时间点的离散标签值, 叠加后的高斯核可以有效地恢复由离散标签表示的连续过程的动态特性.

基于时间序列中多尺度时间信息的固有特性和重要性, 本文提出一种多尺度高斯激励核的设计方法, 旨在有效地融合来自不同时间尺度的信息. 这些高斯激励核的带宽参数用于表征离散标签的多尺度时间特性, 不同尺度的高斯激励核能够捕捉不同时间尺度上的动态变化. 此外, 不同大小的高斯激励核通过相应的权重进行聚合, 以表征它们对恢复的潜在在连续变量产生的不同影响. 具体而言, 恢复的潜在在连续变量可通过下式计算:

$$x_{t, j}^{\text{LC}} = \sum_{i=1}^d \hat{\alpha}_i F_{x_j^D}(t; \sigma_i) + \varepsilon_t. \quad (5)$$

其中:  $\sigma_i = \{\sigma_1, \sigma_2, \dots, \sigma_j, \dots, \sigma_d\}$  为用于表征多尺度时间模式的带宽参数,  $d$  为多尺度激励高斯核的数量;  $\varepsilon_t$  为噪声项, 服从正态分布  $\varepsilon_t \sim N(0, 0.01)$ . 不失一般性, 带宽参数从小到大排列为  $\{\sigma_{\min} = \sigma_1 < \sigma_2 < \dots < \sigma_j \dots < \sigma_d = \sigma_{\max}\}$ . 具体表示为

$$\sigma_i = \sigma_{\min} + (i - 1) \frac{\sigma_{\max} - \sigma_{\min}}{d}. \quad (6)$$

这里:  $\sigma_{\min}$  和  $\sigma_{\max}$  为可学习参数, 用于控制带宽的上下界. 由于不同数据集的特征尺度和噪声水平差异较大, 带宽初值通常根据经验和数据尺度设定, 以保证初始带宽能够覆盖合理的多尺度特征范围<sup>[27-28]</sup>. 此外,  $\hat{\alpha}_i$  为可学习的权重系数, 对应于每个激励高斯核, 这些权重系数通过 Softmax 操作进行归一化, 计算公式如下所示:

$$(\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_d) = \text{Softmax}(\alpha_1, \alpha_2, \dots, \alpha_d). \quad (7)$$

在训练过程中, 权重系数  $\alpha$  初始化为 1, 且可通过反向传播自适应地调整, 用以表征多尺度高斯激励核的不同影响.

通过上述操作, 能够有效恢复离散标签背后的连续过程信息, 从而弥补工业场景中离散标签利用不足的问题.

### 1.3 面向混合时序预测的自监督策略

通过 CAKGE 方法, 可由下式从离散标签中恢复隐连续变量, 进而实现混合时序预测:

$$\hat{x}_j^{\text{LC}} = \text{CAKGE}(x_j^D; \{\sigma_1, \sigma_2, \dots, \sigma_d; \alpha_1, \alpha_2, \dots, \alpha_d\}). \quad (8)$$

下面将详细介绍基于混合时序预测的自监督策略.

首先, 对隐连续变量与连续变量进行特征融合, 具体融合方式为

$$\tilde{x}^{\text{LC}} \parallel x^C \in \mathbb{R}^{p \times T}, \quad (9)$$

其中  $\parallel$  表示拼接操作.

将融合后的特征进行进一步处理, 通过滑动窗口划分数据并输入至预测器. 预测器由卷积神经网络 (CNN) 与 MLP 相结合的 ConvMLP 网络组成, 图 2 为其基本结构. MLP 部分通过多个隐藏层逐步进行非线性变换, 以捕捉时序数据中的复杂模式; 而 CNN 部分由卷积层、池化层和全连接层等关键模块构成, 通过卷积操作提取局部特征, 从而增强模型对短期依赖关系的学习能力. 接着, 将隐连续变量的过去取值和连续变量的过去取值一并输入至预测器, 以预测连续变量的未来状态. 通过这种方式, 模型预测得到连续变量的值不仅可作为目标输出, 还可作为监督信号, 指导 CAKGE 参数的学习. 由下式分别得到连续变量和隐连续变量的预测结果为  $\hat{x}_{j, t}^C$ 、 $\hat{x}_{j, t}^{\text{LC}}$ :

$$\hat{x}_{j,t}^C = \text{Predictor}(\{x_{j,t-\tau:t-1}^C\}_{j=1}^{p-n}; \{x_{j,t-\tau:t-1}^{\text{LC}}\}_{j=1}^n), \quad (10)$$

$$\hat{x}_{j,t}^D = \text{Predictor}(\{x_{j,t-\tau:t-1}^C\}_{j=1}^{p-n}; \{x_{j,t-\tau:t-1}^{\text{LC}}\}_{j=1}^n), \quad (11)$$

其中 Predictor 为预测器。

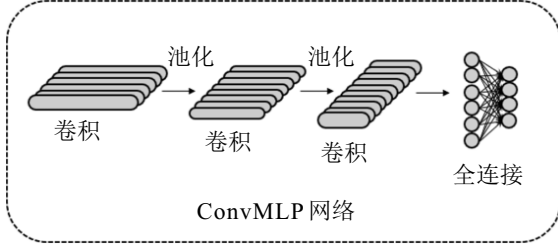


图2 ConvMLP 网络结构

引入隐连续变量预测使得模型能够同时建模隐连续变量的预测误差, 增强了对隐连续变量数据可预测性的处理能力, 以提升整体预测性能. 由此, 预测误差损失函数可表示为

$$L_{\text{fore}} = \sum_{j=1}^{p-n} \text{MSE}(\hat{x}_{j,t}^C, x_{j,t}^C) + \sum_{j=1}^n \text{MSE}(\hat{x}_{j,t}^C, \tilde{x}_{j,t}^{\text{LC}}), \quad (12)$$

其中 MSE 为均方误差损失函数. 同时, 隐连续变量的过去取值也被输入至离散标签解码器 (记作 Decoder), 用于重建原始的离散标签. 解码器的结构与预测器相同, 重建损失函数能够确保 CAKGE 的变换不会过度关注连续变量的预测, 而是依然保留原始离散值的信息. 随后, 将隐连续变量进行解码重构, 得到重构结果, 记为  $\hat{x}_{j,t}^D$ . 重构损失函数为

$$L_{\text{recons}} = \sum_{j=1}^n \text{CE}(\hat{x}_{j,t}^D, x_{j,t}^D). \quad (13)$$

这里:  $\hat{x}_{j,t}^D \leftarrow \text{Decoder}_j(\tilde{x}_{j,t-\tau:t}^{\text{LC}})$ , CE 为交叉熵损失函数. 最终, 总体损失函数可表示为

$$L_{\text{all}} = \lambda_1 L_{\text{fore}} + \lambda_2 L_{\text{recons}}, \quad (14)$$

其中  $\lambda_1$  和  $\lambda_2$  为加权系数.

#### 1.4 CDL-HCR 建模整体流程

为实现对粗粒度离散标签中所隐含连续信息的有效恢复, 并进一步提升混合时序数据的建模和预测能力, 本文提出 CDL-HCR 的算法流程, 具体步骤如算法 1 所示.

##### 算法1 CDL-HCR的算法流程.

输入: 连续变量  $x^C = \{x_1, x_2, \dots, x_{p-n}\} \in \mathbb{R}^{(p-n) \times T}$ , 离散标签  $x^D = \{x_1, x_2, \dots, x_n\} \in \mathbb{R}^{n \times T}$ ,  $\tau$ ,  $\lambda_1$  和  $\lambda_2$ ;

输出: 联合训练模型参数  $\text{CAKGE}_i = (\sigma_{\min}, \sigma_{\max}; \alpha_1, \dots, \alpha_2)$ ,  $\text{Decoder}_i (i = 1, 2, \dots, p)$ ,  $\text{Predict}_i (i = 1, 2, \dots, p)$ .

1. 初始化模型参数

2. 当模型未收敛时

3. //模型前向传播

4. for  $i$  in range ( $n$ ) do

$$\tilde{x}_i^{\text{LC}} \leftarrow \text{CAKGE}_i(x_i^D)$$

$$\hat{x}_i^D \leftarrow \text{Decoder}_i(\tilde{x}_i^{\text{LC}}, \tilde{x}_{i,t-\tau:t-1}^{\text{LC}})$$

$$\hat{x}_{i,t}^C \leftarrow \text{predict}_i(\{x_{j,t-\tau:t-1}^C\}_{j=1}^{p-n}; \{\tilde{x}_{j,t-\tau:t-1}^{\text{LC}}\}_{j=1}^n)$$

5. end for

6. for  $i$  in range ( $p - n$ ) do

$$\hat{x}_{i,t}^C \leftarrow \text{predict}_i(\{x_{j,t-\tau:t-1}^C\}_{j=1}^{p-n}; \{\tilde{x}_{j,t-\tau:t-1}^{\text{LC}}\}_{j=1}^n)$$

7. end for

8. //计算损失

9. 预测误差损失

$$L_{\text{fore}} \leftarrow \sum_{j=1}^{p-n} \text{MSE}(\hat{x}_{j,t}^C, x_{j,t}^C) + \sum_{j=1}^n \text{MSE}(\hat{x}_{j,t}^C, \tilde{x}_{j,t}^{\text{LC}})$$

10. 重构误差损失

$$L_{\text{recons}} \leftarrow \sum_{j=1}^n \text{CE}(\hat{x}_{j,t}^D, x_{j,t}^D)$$

11. 总体损失函数

$$L_{\text{all}} \leftarrow \lambda_1 L_{\text{fore}} + \lambda_2 L_{\text{recons}}$$

12. //参数更新

13. 使用Adam优化器最小化  $L_{\text{all}}$ , 更新模型参数

14. end while

15. 训练结束, 返回训练后的模型参数

该算法以连续变量和离散标签作为输入, 通过迭代优化联合训练的模型参数, 在自监督机制的指导下实现隐连续性恢复与时序预测任务的融合建模. 训练过程包括模型初始化、前向传播、损失计算以及参数更新等步骤, 采用 Adam 优化器最小化由预测误差与重构误差构成的总损失函数. 最终, 算法输出训练后的模型参数.

## 2 CDL-HCR 在线应用

本节主要介绍所提出 CDL-HCR 方法在在线应用阶段的具体流程, 其整体在线应用阶段如图 3 所示.

在完成离线训练后, 模型部署在实际工业系统中, 以烟丝滚筒烘丝机为典型应用场景, 进入在线应用阶段. 系统实时接收筒壁温度、热风温度等连续变量以及加工强度、蒸汽压力离散标签作为输入, 其中离散标签经隐连续性恢复模块映射为隐连续变量  $x^{\text{LC}}$ . 然后, 将隐连续变量与连续变量共同输入至自监督预测模块, 实现关键变量的动态预测. 根据预测结果与真实值间的差异, 模型计算预测误差和重构误差, 并构建总体损失函数用于性能评估. 为提升模

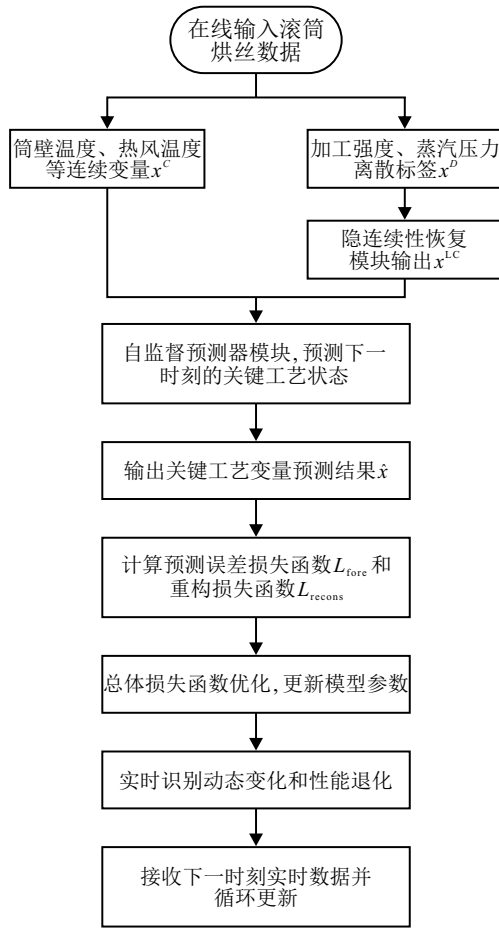


图3 在线应用流程

型适应性, 系统采用 Adam 优化器对模型参数进行在线更新, 使其持续适应工况变化, 实现稳定、高效的实时预测。

在线应用过程中, 模型对关键变量的预测结果能够反映设备运行状态的动态变化特征. 通过对预测误差的持续监测, 当误差出现系统性上升趋势时, 即可识别出设备性能逐渐退化的早期征兆, 从而实现对运行状态的实时跟踪和性能退化的早期识别。

### 3 案例分析

本节主要介绍实验中使用的数据集、评价指标和模型参数设定, 并对实验结果进行分析。

#### 3.1 实验数据集与评价指标

本文采用的数据集来源于实际烟草工业中的滚筒烘丝加工过程, 旨在刻画烟丝在加工过程中的动态变化. 该数据集包含 8 个连续变量和 2 个离散标签变量. 其中: 8 个连续变量通过传感器直接采集, 而 2 个离散标签变量则是将原本连续变量进行离散化处理得到的. 通常情况下, 这些离散标签变量由二值化报警信号构成. 数据集总共包含 1000 条记录. 其中: 800 条用于训练, 200 条用于测试. 为提高模型的训练效率和鲁棒性, 所有数据均进行归一化处理. 由于加工强度和蒸汽压力是从二值化报警信号转化

而来的离散标签变量 (记作 DL), 这两者无需进行归一化处理. 最终, 仅对 8 个连续变量 (记作  $CV_3 \sim CV_{10}$ ) 进行归一化处理, 将其值压缩至  $[0, 1]$  区间内.

为验证所提出方法的泛化能力, 本文采用一个公开可获取的功能磁共振成像 (fMRI) 数据集. 该数据集是神经影像领域常用的标准基准数据<sup>[29]</sup>, 因其多变量、长序列和复杂动态特征, 在工业时序建模相关研究中也常被频繁采用. 本文选取其中第 6 个子数据集, 该数据集包含 10 个变量和 1200 个样本, 其信号特性与工业过程数据存在显著相似, 能够有效检验模型的泛化能力. 为与滚筒烘丝数据集的变量形式保持一致, 本文保留数据集中原有的变量类型. 其中: 2 个离散标签变量 (记作  $DL_1$  和  $DL_2$ ), 其余 8 个变量为连续变量 (记作  $CV_3 \sim CV_{10}$ ). 数据预处理过程与滚筒烘丝数据集一致, 且样本数据按照 8:2 的比例划分为训练集和测试集.

为评估方法的预测性能, 引入性能评价指标, 包括平均绝对误差 (MAE)、均方根误差 (RMSE)、平均绝对百分比误差 (MAPE) 以及决定系数 ( $R^2$ ), 这些指标计算公式如下所示:

$$MAE = \frac{1}{T} \sum_{i=1}^T |\hat{x}_i - x_i|, \quad (15)$$

$$RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (\hat{x}_i - x_i)^2}, \quad (16)$$

$$MAPE = \frac{100\%}{T} \sum_{i=1}^T \left| \frac{x_i - \hat{x}_i}{x_i} \right|, \quad (17)$$

$$R^2 = 1 - \frac{\sum_{i=1}^T (x_i - \hat{x}_i)^2}{\sum_{i=1}^T (x_i - \bar{x})^2}. \quad (18)$$

其中:  $\hat{x}_i$  和  $x_i$  分别为预测值和实际值;  $\bar{x}$  为实际值的均值;  $T$  为样本数量;  $R^2$  取值范围通常在  $0 \sim 1$  之间, 值越接近 1, 模型拟合效果越好。

#### 3.2 模型参数设定

本文采用的 ConvMLP 模型结构包括两层 CNN: 第 1 层包含 10 个神经元, 使用  $5 \times 5$  的二维卷积核; 第 2 层包含 20 个神经元, 采用  $1 \times 1$  的一维卷积核. 每层均设置了激活函数为修正线性单元. 在模型训练过程中, 数据批量大小设定为 25, 时间窗口大小为 5, 学习率为 0.01, 训练轮数为 200. CAKGE 模块的带宽  $\sigma_{max}$  和  $\sigma_{min}$  分别设置为 4 和 0.3,  $\lambda_1$  和  $\lambda_2$  分别设置为 1 和 0.5. 带宽参数通过网格搜索实验确定. 其中:  $\sigma_{max}$  的网格搜索的范围为  $3 \sim 5$ ,  $\sigma_{min}$  的网格搜

索的范围为0.2 ~ 0.5, 搜索步长分别为 0.1 和 0.01. 实验中: 将样本数据按照 8 : 2 的比例划分为训练集和测试集, 在固定其他超参数的条件下, 通过网格搜索方式对目标参数进行遍历, 并根据模型在测试集上的预测性能和收敛稳定性综合评估最优参数组合. 最终, 将 CAKGE 模块的带宽参数 $\sigma_{\max}$ 和 $\sigma_{\min}$ 分别确定为 4 和 0.3.

采用 MiTCD<sup>[19]</sup>、cLSTM<sup>[25]</sup> 等方法作为对比模型, 并分别在滚筒烘丝数据集和 fMRI 数据集上进行实验评估. 为保证实验的公平性, 所有对比模型的初始学习率均设置为 0.01, 且训练轮数与时间窗口长度均保持一致.

### 3.3 结果分析与讨论

在本实验中, 采用 CAKGE 模块对加工强度和蒸汽压力的离散标签进行隐连续性恢复. 图 4 为由离散化连续变量得到的离散标签 (灰色曲线) 和恢复结果 (灰色虚线). 为验证隐连续性恢复方法的有效性, 采用由专家提供的实际隐连续变量数据 (黑色曲线) 进行对比. 该数据的采集难度大、成本高且稀缺, 因此, 作为验证标准具有重要意义.

图 4 实验结果表明, 隐连续性恢复方法在处理加工强度时存在较大波动, 未能精准捕捉实际隐连续变量的细节变化. 这主要是由于离散标签信号在激励点处频繁并剧烈地跃迁, 且每次跃迁的持续时间较短, 这使得高斯核难以有效捕捉这些快速变化的时刻. 此外, 实际工业场景中的离散化过程会导致信息丢失, 从而使得恢复得到的隐连续变量与实际隐连续变量间存在一定的误差. 相比之下, 蒸汽压力离散标签信号较为平稳, 隐连续性恢复方法能够较好地捕捉其周期性变化, 恢复结果与实际隐连续变量接近, 表明该方法在处理平稳信号时具有较强的恢复能力.

表 1 为隐连续性恢复方法在不同变量上的性能评价指标, 用于评估恢复结果与实际隐连续变量间的误差. 表 1 结果显示, 蒸汽压力的恢复效果优于加工强度, 在 MAE 和 MAPE 指标上均表现出更高的精度.

图 5 为模型在训练过程中, 损失值和准确率随训练轮数变化的曲线. 由图 5 可见, 随着训练轮数的增加, 预测损失和重构损失均实现了快速收敛. 与此同时, 重构准确率持续上升, 从初始的 0.6285 稳步提升至 0.9119, 表明模型在重构原始离散标签方面的能力不断增强. 结合图 4 和图 5 的实验结果, 可以验证该模型不仅能够有效恢复潜在的连续信息, 还能成功重构原始的离散标签变量, 并保留观测数据离散标签中的关键信息.

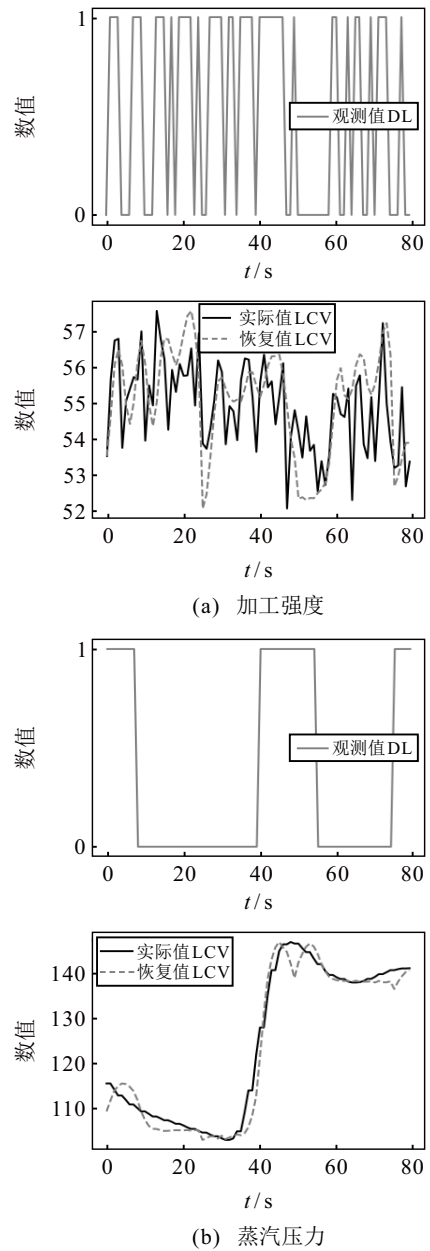


图 4 隐连续性恢复结果对比: 加工强度与蒸汽压力

表 1 隐连续性恢复方法在不同变量上的性能评价指标

变量名	MAE	RMSE	MAPE/%
加工强度	0.8930	1.1107	1.5924
蒸汽压力	0.5308	1.2698	0.4226

为进一步分析该模型在混合时序预测方面的有效性, 图 6 展示了 8 个连续变量的单步预测结果. 图 6 中: 第 1 排从左往右依次为 CV<sub>3</sub> ~ CV<sub>6</sub>, 第 2 排从左往右则为 CV<sub>7</sub> ~ CV<sub>10</sub>; 灰色虚线为模型的预测值, 黑色曲线为真实值. 由图 6 可见, 模型预测的估计值与真实值高度接近, 表明模型在捕捉数据变化趋势和预测未来时具有较高的准确性. 同时, 表 2 为这 8 个变量的预测指标, 采用 MAE、RMSE 和 R<sup>2</sup> 指标对预测效果进行评估. 其中: CV<sub>4</sub> 和 CV<sub>9</sub> 的 MAE 值较小, 分别为 0.0422 和 0.0494, 表明这些变量的

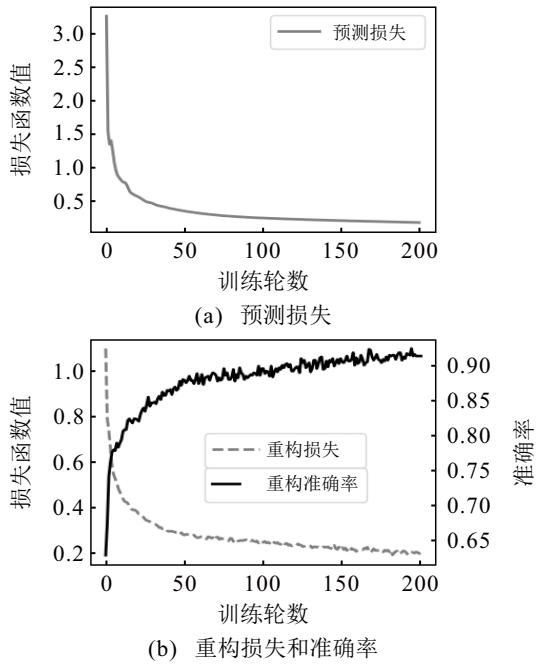


图5 损失值和准确率随训练轮数变化

预测误差非常小,模型能够很好地捕捉到数据的细微变化.通过对所有变量进行平均计算,结果显示,8个变量的平均 MAE 为 1.9707,平均 RMSE 为 1.2179,平均  $R^2$  为 0.9912.平均  $R^2$  的值接近 1,表明整体预测结果较为准确,进一步验证了该模型在混合时序预测中的优越性.

表2  $CV_3 \sim CV_{10}$  预测结果指标记录

变量名	MAE	RMSE	$R^2$
$CV_3$	0.1422	0.3362	0.9982
$CV_4$	<b>0.0422</b>	0.1160	<b>0.9987</b>
$CV_5$	0.2270	0.8349	0.9978
$CV_6$	3.9066	1.0626	0.9788
$CV_7$	4.1682	3.7666	0.9863
$CV_8$	7.1538	3.3258	0.9763
$CV_9$	0.0494	0.1941	0.9981
$CV_{10}$	0.0761	<b>0.1073</b>	0.9954
平均结果	1.9707	1.2179	0.9912

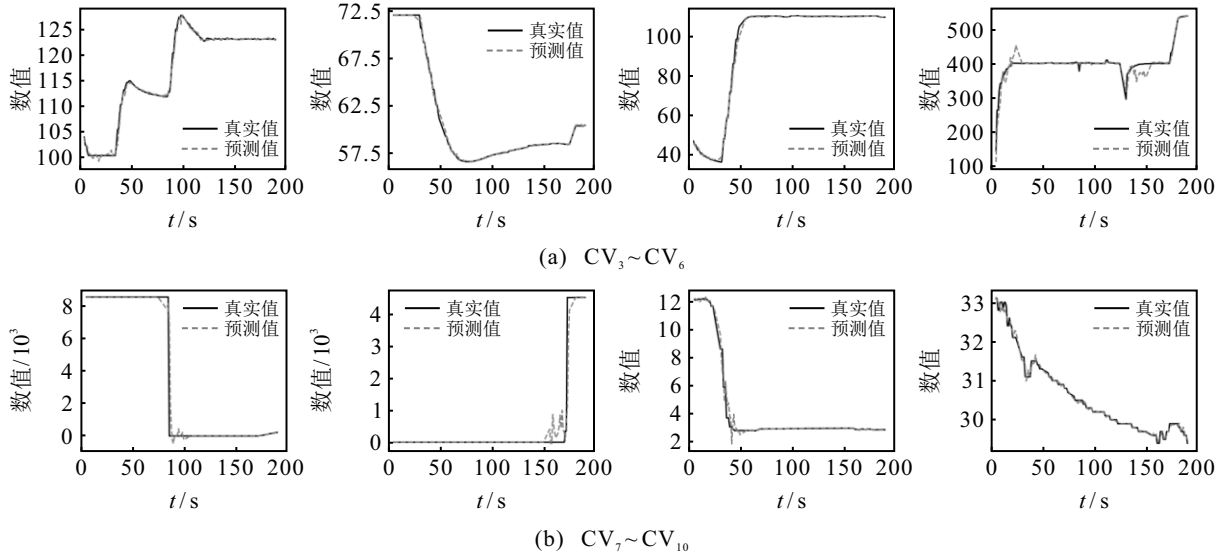


图6 滚筒烘丝数据集下混合时序预测结果展示

为验证所提出模型各组成模块的有效性,本文设计并开展消融实验,其结果如表3和表4所示.以完整的 CDL-HCR 结构为基准,分别移除 CAKGE 模块(记作  $CDL-HCR^{\ddagger}$ )和  $L_{recons}$  模块(记作  $CDL-HCR^{\S}$ ).通过比较各模型在 MAE、RMSE 和平均  $R^2$  等指标上的表现,评估各模块对于整体模型性能的具体贡献.结果表明,相较于移除  $L_{recons}$  模块,移除 CAKGE 模块对于离散标签变量的隐连续性建模的影响更为显著.此外,进一步分析连续变量的预测效果.表5结果显示,CDL-HCR 的平均  $R^2$  相较于  $CDL-HCR^{\ddagger}$  提升了 2.42%,相较于  $CDL-HCR^{\S}$  提升了 1.51%.这一结果表明:移除 CAKGE 模块对于连续变量预测性能的影响大于移除  $L_{recons}$  模块;同时,也表明离散标签变量在一定程度上对于提高连续变量预测具有

表3 不同变量下移除 CAKGE 模块的消融实验结果

变量名	MAE	RMSE	MAPE/%
加工强度	0.9925	<b>1.1457</b>	1.7679
蒸汽压力	<b>0.6944</b>	2.0585	<b>0.5581</b>

表4 不同变量下移除  $L_{recons}$  模块的消融实验结果

变量名	MAE	RMSE	MAPE/%
加工强度	0.9694	<b>1.2090</b>	1.7289
蒸汽压力	<b>0.5234</b>	1.2361	<b>0.4155</b>

表5 连续变量预测任务中的消融实验结果

模型	CAKGE	$L_{recons}$	平均 $R^2$
CDL-HCR	√	√	<b>0.9912</b>
$CDL-HCR^{\ddagger}$	×	√	0.9678
$CDL-HCR^{\S}$	√	×	0.9765

促进作用.

为进一步验证所提出模型的泛化能力, 本文采用 fMRI 数据集开展泛化实验分析. 为评估模型在 fMRI 数据集上的混合时序预测性能, 对 8 个连续变量进行单步预测实验, 预测结果如图 7 所示. 图 7 中: 第 1 排从左往右依次为  $C\bar{V}_3 \sim C\bar{V}_6$ , 第 2 排从左往右则为  $C\bar{V}_7 \sim C\bar{V}_{10}$ ; 灰色虚线为模型的预测值, 黑色曲线为真实值. 由图 7 可见, 模型预测曲线与真实值拟合较好, 能够较为准确地捕捉时间序列的变化趋势. 表 6 为模型在 fMRI 数据集中对两个离散标签变量的隐连续性恢复实验结果. 由 MAE、RMSE 和 MAPE 评价指标, 以及图 8 的对比可见,  $D\bar{L}_2$  的恢复

效果要优于  $D\bar{L}_1$ , 表明该模型在处理复杂时间信号特征方面具有较强的泛化能力. 同时, 表 7 记录了这 8 个连续变量在预测任务中的关键性能指标, 包括 MAE、RMSE 和  $R^2$ . 其中:  $C\bar{V}_7$  的 MAE、RMSE 值最小;  $R^2$  达到 0.7548, 表现最优. 将 8 个变量的结果进行平均计算后可得到 8 个变量的平均 MAE 为 1.1979, 平均 RMSE 为 1.4972, 平均  $R^2$  为 0.5759. 考虑到 fMRI 数据集本身具有较强的波动性和噪声特征, 模型在该数据集上依然保持良好的预测性能, 表明所提出方法在跨领域混合时序数据建模任务中具备较强的泛化能力和鲁棒性.

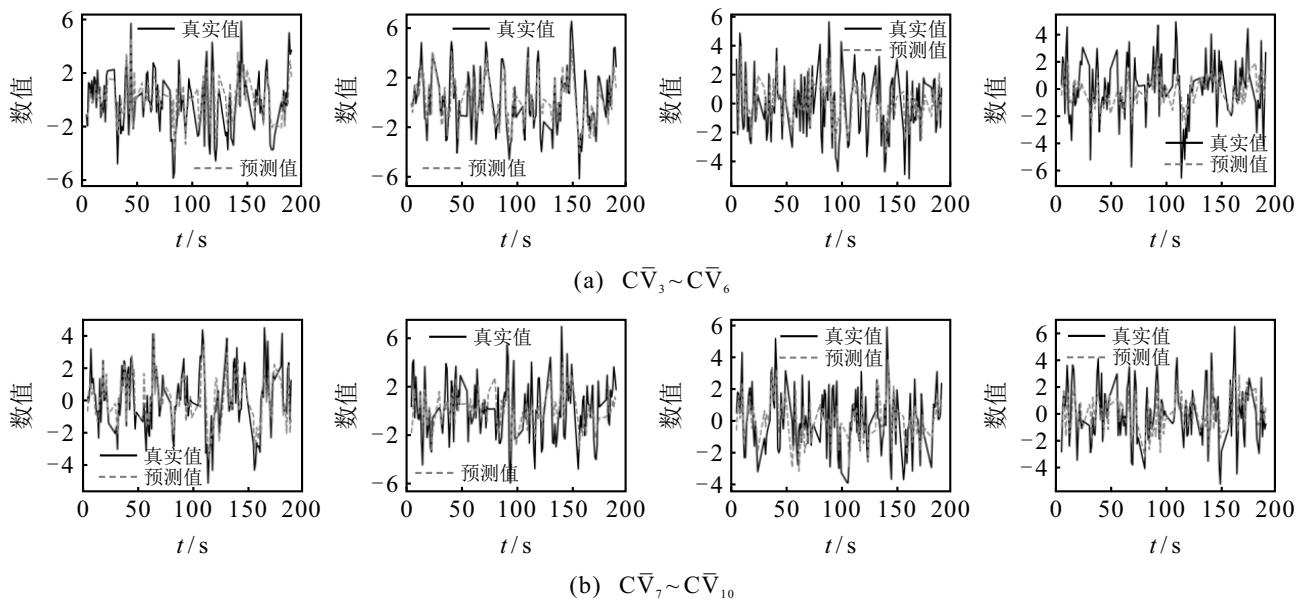


图7 fMRI 数据集下混合时序预测结果展示

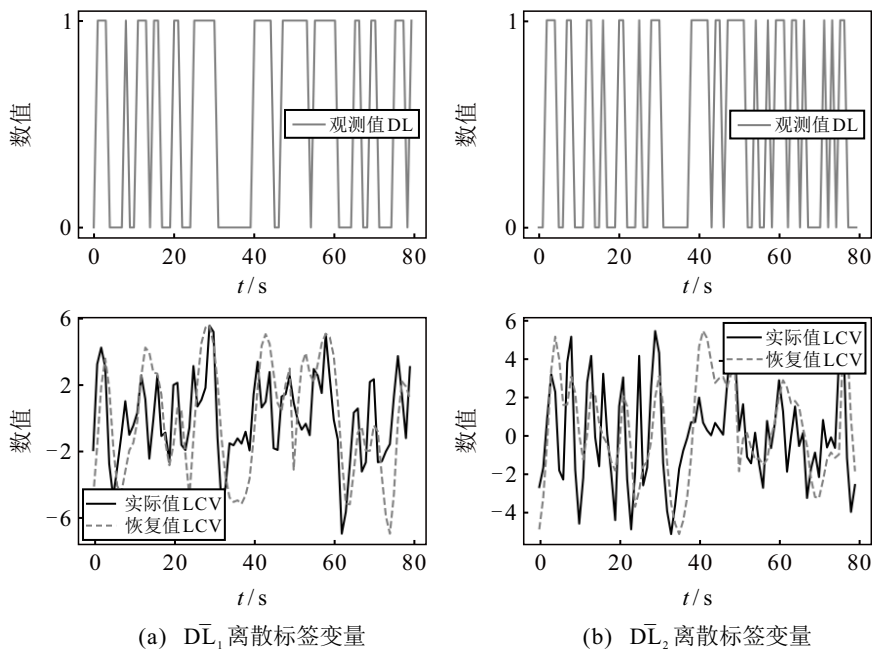


图8 隐连续性恢复结果对比:  $D\bar{L}_1$  与  $D\bar{L}_2$

表6 fMRI数据集下不同变量上的性能评价指标

变量名	MAE	RMSE	MAPE/%
$DL_1$	2.7512	3.2963	6.7621
$DL_2$	<b>2.2754</b>	<b>2.7412</b>	<b>3.8262</b>

表7 fMRI数据集下 $C\bar{V}_3 \sim C\bar{V}_{10}$ 预测结果指标记录

变量名	MAE	RMSE	$R^2$
$C\bar{V}_3$	1.1599	1.4603	0.6064
$C\bar{V}_4$	1.0364	1.3020	0.7030
$C\bar{V}_5$	1.3283	1.6348	0.4835
$C\bar{V}_6$	1.5494	1.9172	0.3282
$C\bar{V}_7$	<b>0.9541</b>	<b>1.2201</b>	<b>0.7548</b>
$C\bar{V}_8$	1.1441	1.4359	0.6248
$C\bar{V}_9$	1.1742	1.4817	0.5738
$C\bar{V}_{10}$	1.2366	1.5255	0.5330
平均结果	<u>1.1979</u>	<u>1.4972</u>	<u>0.5759</u>

表8为不同方法在各数据集下的对比实验结果。由表8可知,所提出CDL-HCR方法在两个数据集上均显著优于各类对比模型。在滚筒烘丝数据集上,CDL-HCR的平均 $R^2$ 相较于MiTCD、cLSTM与cMLP分别提升了2.60%、1.66%和8.76%;同时,平均MAE分别降低了7.37%、4.26%和27.62%,表明该方法在混合时序预测任务中具有更高的拟合能力和更低的预测误差。在fMRI数据集上,CDL-HCR同样取得了最优表现。与次优模型MiTCD相比,其平均 $R^2$ 提升了52.31%,平均MAE降低了12.38%。这一结果充分表明,CDL-HCR在处理复杂生理信号时具备较强的特征表达能力和泛化能力。

表8 不同方法在各数据集下的对比实验结果

数据集	模型	平均MAE	平均 $R^2$
滚筒烘丝数据集	cMLP	2.7224	0.9114
	cLSTM	2.0583	0.9750
	MiTCD	2.1275	0.9661
	CDL-HCR	<b>1.9707</b>	<b>0.9912</b>
fMRI数据集	cMLP	1.5458	0.2196
	cLSTM	1.4440	0.3097
	MiTCD	1.3672	0.3781
	CDL-HCR	<b>1.1979</b>	<b>0.5759</b>

通过上述实验和分析,本文验证了隐连续性恢复在混合时序预测中的有效性。该技术能够从可观测的离散标签信息中有效推测出隐含的连续变量,这些隐连续变量通常是无法直接测量的。通过这一过程,模型能够实现混合时序数据特征的有效融合,并为混合时序预测提供统一的连续的空间,从而在

预测任务中取得了显著的效果。

## 4 结论

本文提出了一种基于隐连续性恢复的离散标签混合时序预测方法CDL-HCR,并在实际工业场景中的滚筒烘丝数据集上进行了实验验证。该方法通过隐连续性恢复策略,将粗粒度的离散标签恢复为细粒度的隐连续变量,从而实现了原本不可测的连续变量进行准确估计。同时,将恢复后的隐连续变量与其他连续变量在同一连续空间下进行了特征融合,并结合自监督学习策略,高效预测了混合时序数据。考虑到隐连续变量与原始细粒度连续变量在时间动态特性和信息分辨率上的差异,未来可进一步设计特征融合机制,将这两类变量有效地整合到统一的表示空间中。通过这种融合机制,能够为复杂的工业应用场景提供更加精确的预测结果。

## 参考文献 (References)

- [1] 赵春晖,余万科,高福荣.非平稳间歇过程数据解析与状态监控——回顾与展望[J].自动化学报,2020,46(10):2072-2091.  
(Zhao C H, Yu W K, Gao F R. Data analytics and condition monitoring methods for nonstationary batch processes — Current status and futur[J]. Acta Automatica Sinica, 2020, 46(10): 2072-2091.)
- [2] 赵健程,冯良骏,岳嘉祺,等.从零样本学习理论模型到工业应用——动机、演变与挑战[J].控制与决策,2024,39(9):2833-2857.  
(Zhao J C, Feng L J, Yue J Q, et al. From zero-shot learning theoretical model to its industrial application: Motivation, evolution and challenges[J]. Control and Decision, 2024, 39(9): 2833-2857.)
- [3] Feng L J, Zhao C H, Li Y L, et al. Multichannel diffusion graph convolutional network for the prediction of endpoint composition in the converter steelmaking process[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 70: 3000413.
- [4] Zeng Y, Shimizu S, Matsui H, et al. Causal discovery for linear mixed data[C]. Proceedings of the Conference on Causal Learning and Reasoning. Cambridge, 2022: 994-1009.
- [5] 柴铮,汪嘉业,赵春晖,等.面向工业监控典型监督任务的深度迁移学习方法:现状、挑战与展望[J].中国科学:信息科学,2023,53(5):821-840.  
(Chai Z, Wang J Y, Zhao C H, et al. Deep transfer learning methods for typical supervised tasks in industrial monitoring: Status, challenges and prospects[J]. Science China Information Sciences, 2023, 53(5): 821-840.)
- [6] Li P, Shlezinger N, Zhang H Y, et al. Graph signal compression by joint quantization and sampling[J]. IEEE Transactions on Signal Processing, 2022, 70: 4512-4527.
- [7] Eylen T, Eren P E, Koçyiğit A. Data-driven alarm parameter optimization[J]. Computers & Chemical

- Engineering, 2025, 196: 109041.
- [8] 刘金培, 罗瑞, 陈华友, 等. 基于多尺度 3D-CNN-CBAM 的空气质量指数时空预测[J]. 控制与决策, 2025, 40(2): 404-412.  
(Liu J P, Luo R, Chen H Y, et al. Spatio-temporal prediction of air quality index based on multi-scale 3D-CNN-CBAM[J]. Control and Decision, 2025, 40(2): 404-412.)
- [9] Zhao C H. Perspectives on nonstationary process monitoring in the era of industrial artificial intelligence[J]. *Journal of Process Control*, 2022, 116: 255-272.
- [10] Li J X, Huang K K, Wu D H, et al. Hybrid variable dictionary learning for monitoring continuous and discrete variables in manufacturing processes[J]. *Control Engineering Practice*, 2024, 149: 105970.
- [11] Zhu K, Zhao C H, Huang B. Breaking information granularity heterogeneity: A mutual information-inspired causal discovery framework for multi-rate time series[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2025, 37(12): 7001-7015.
- [12] Qin S J. Survey on data-driven industrial process monitoring and diagnosis[J]. *Annual Reviews in Control*, 2012, 36(2): 220-234.
- [13] Yang H Q, Huang K Z, King I, et al. Localized support vector regression for time series prediction[J]. *Neurocomputing*, 2009, 72(10/11/12): 2659-2669.
- [14] Jiang H G, Zhang Y C, Muljadi E, et al. A short-term and high-resolution distribution system load forecasting approach using support vector regression with hybrid parameters optimization[J]. *IEEE Transactions on Smart Grid*, 2018, 9(4): 3341-3350.
- [15] Yi K, Zhang Q, Fan W, et al. Frequency-domain MLPs are more effective learners in time series forecasting[J]. *Advances in Neural Information Processing Systems*, 2023, 36: 8588-8601.
- [16] 李闯, 毛志忠, 欧阳, 等. 机理指导的 LSTM 网络及其在康斯迪电弧炉钢水连续温度预测中的应用[J]. 控制与决策, 2025, 40(10): 3055-3064.  
(Li C, Mao Z Z, Ou Y, et al. Mechanism-guided LSTM network for continuous temperature prediction in Consteel smelting process[J]. *Control and Decision*, 2025, 40(10): 3055-3064.)
- [17] Yuan X F, Qi S B, Wang Y L, et al. Quality variable prediction for nonlinear dynamic industrial processes based on temporal convolutional networks[J]. *IEEE Sensors Journal*, 2021, 21(18): 20493-20503.
- [18] Tan Y D, Zhu H B, Wu J, et al. DPTVAE: Data-driven prior-based tabular variational autoencoder for credit data synthesizing[J]. *Expert Systems with Applications*, 2024, 241: 122071.
- [19] Chen J W, Zhao C H. Addressing information asymmetry: Deep temporal causality discovery for mixed time series[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2025, 47(7): 5723-5741.
- [20] Zhao L, Li B X, Zhao C H. Addressing external distorted heterogeneity: Input-output disentangled causal representation for mixed time series[J]. *Journal of Process Control*, 2025, 154: 103521.
- [21] Cheng X Z, Liu Z L, Fang H B. Deep fusion of discrete-time and continuous-time models for long-term prediction of chaotic dynamical systems[J]. *Nonlinear Dynamics*, 2025, 113(11): 12545-12563.
- [22] Sedgewick A J, Buschur K, Shi I, et al. Mixed graphical models for integrative causal analysis with application to chronic lung disease diagnosis and prognosis[J]. *Bioinformatics*, 2019, 35(7): 1204-1212.
- [23] Ring M, Eskofier B M. An approximation of the Gaussian RBF kernel for efficient classification with SVMs[J]. *Pattern Recognition Letters*, 2016, 84: 107-113.
- [24] Miao X Y, Jiang A M, Liu T, et al. Kernel regression for matrix-variate Gaussian distributed signals over sample graphs[J]. *IEEE Transactions on Signal and Information Processing Over Networks*, 2022, 8: 726-738.
- [25] Chen J W, Zhao C H. Addressing spatial-temporal heterogeneity: General mixed time series analysis via latent continuity recovery and alignment[C]. *Advances in Neural Information Processing Systems 37*. Vancouver, 2024: 17910-17946.
- [26] Panigrahi L, Verma K, Singh B K. Ultrasound image segmentation using a novel multi-scale Gaussian kernel fuzzy clustering and multi-scale vector field convolution[J]. *Expert Systems with Applications*, 2019, 115: 486-498.
- [27] Bao J, Chen Y Y, Yu L, et al. A multi-scale kernel learning method and its application in image classification[J]. *Neurocomputing*, 2017, 257: 16-23.
- [28] Shi L, Tan J H, Wang J, et al. Robust kernel adaptive filtering for nonlinear time series prediction[J]. *Signal Processing*, 2023, 210: 109090.
- [29] Smith S M, Miller K L, Salimi-Khorshidi G, et al. Network modelling methods for FMRI[J]. *NeuroImage*, 2011, 54(2): 875-891.

## 作者简介

孙曦 (1997-), 男, 硕士生, 主要研究方向为工业软测量, E-mail: [sunxi1005@163.com](mailto:sunxi1005@163.com);

赵春晖 (1979-), 女, 教授, 博士, 博士生导师, 主要研究方向为工业大数据分析与应用、状态监测、故障诊断、软测量, E-mail: [chhzhao@zju.edu.cn](mailto:chhzhao@zju.edu.cn);

张皓然 (1996-), 男, 助理研究员, 博士, 主要研究方向为工业人工智能、数据驱动控制, E-mail: [0625495@zju.edu.cn](mailto:0625495@zju.edu.cn).