文章编号:1673-0291(2026)04-0001-12 DOI:10.11860/j.issn.1673-0291.20250196

基于SKInformer的列车部件加速度时序预测

彭珍瑞[1,2],武小晶[1],殷 红[1]
(1.兰州交通大学 机电工程学院,兰州 730070;2.兰州理工大学 机电工程学院,兰州 730050)

摘 要:针对列车运行时部件振动加速度时序呈现的非线性与非平稳性,导致Informer模型预测精度不佳的问题,提出一种将选择性内核机制(Selective Kernel mechanism,SKmechanism)与Informer架构结合的选择性内核机制Informer(Selective Kernel mechanism Informer,SKInformer)模型,以实现列车部件振动加速度时序预测.首先,通过序列维度转换实现一维时序与SKmechanism的二维图卷积之间的转换,旨在应用图卷积来挖掘一维时序的空间相关性,增强不同时序间的特征表达能力;其次,通过SKmechanism对卷积进行分裂、融合与选择操作,使网络能够依据输入动态调整感受野大小,增强其对多尺度特征的适应能力;再次,将SKmechanism嵌入Informer的编码器,通过动态选择感受野,同步捕捉时序中的短期突变特征与长期趋势信息;最后,将SKmechanism嵌入Informer的解码器,自适应调整对历史序列的关注范围,提取更具判别力的局部时序表征.建立受电弓数据集消融实验与LRV4313轻轨列车数据集对比实验,对设计的预测方法进行验证,结果表明:相较于基准Informer,消融实验中学习率取0.000 1时,SKInformer在测试集上的损失值下降34.18%;在对比实验中,SKInformer最优决定系数为0.99.研究证明所提方法对列车零部件加速度时序具有更好的预测精度与鲁棒性.

关键词:加速度时序;多变量预测;SKmechanism;编码器;解码器
中图分类号:TH113.1 **文献标志码**:A

Acceleration time-series prediction for train components based on SKInformer

Peng Zhenrui[1,2], *Wu Xiaojing*[1], *Yin Hong*[1]
(1.School of Mechanical Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China; 2.School of Mechanical and Electrical Engineering, Lanzhou University of Technology, Lanzhou 730050, China)

Abstract: The nonlinear and nonstationary characteristics of vibration acceleration time series from train components degrade the predictive accuracy of the Informer model. To address this issue, this study proposes a Selective Kernel Mechanism Informer (SKInformer) model that integrates the Selective Kernel mechanism (SKmechanism) into the Informer architecture for predicting vibration acceleration time series of train components. First, sequence-dimension transformation converts one-dimensional time series into the two-dimensional feature maps required by the SKmechanism, en-

收稿日期:2025-12-23;**修回日期**:2026-02-12
基金项目:国家自然科学基金(62161018);兰州市人才创新创业项目(2017-RC-66)
Foundation items: National Natural Science Foundation of China(62161018); Talent Innovation and Entrepreneurship Project in Lanzhou City (2017-RC-66)
第一作者:彭珍瑞(1972—),男,甘肃民勤人,教授,博士,博士生导师.研究方向为结构健康监测、机械故障诊断、结构动力学及模态分析.
email:pzrui@163.com.

引用格式:彭珍瑞,武小晶,殷红.基于SKInformer的列车部件加速度时序预测[J].北京交通大学学报,2026,50(4):1-12.
Peng Zhenrui, Wu Xiaojing, Yin Hong. Acceleration time-series prediction for train components based on SKInformer[J]. Journal of Beijing Jiaotong University, 2026, 50(4): 1-12. (in Chinese)

abling two-dimensional convolution to capture spatial correlations within the time series and enhance feature interactions across time steps. Second, the SKmechanism performs split, fusion, and selection operations on convolutional branches, allowing the network to adjust its receptive-field size dynamically according to the input and thereby improve its adaptability to multiscale features. Third, the SKmechanism is embedded in the Informer encoder. Through dynamic receptive-field selection, the model simultaneously captures short-term abrupt variations and long-term trends in the time series. Finally, the SKmechanism is incorporated into the Informer decoder to adaptively adjust the range of attention over the historical sequence and extract more discriminative local temporal representations. Ablation experiments on a pantograph dataset and comparative experiments on the LRV4313 light-rail vehicle dataset are conducted to validate the proposed method. Compared with the baseline Informer, SKInformer reduces the test-set loss by 34.18% at a learning rate of 0.000 1 in the ablation experiments and achieves a maximum coefficient of determination ($R^2=0.99$) in the comparative experiments. These results demonstrate that the proposed method provides greater accuracy and robustness in predicting acceleration time series for train components.

Keywords: acceleration time series; multivariate prediction; SKmechanism; encoder; decoder

列车部件在长时间、复杂工况下运行易出现性能退化甚至突发故障，对行车安全构成严重威胁[1-2]. 传统监测方法在实时感知及早期预警方面存在局限，而基于深度学习的时间序列预测技术正逐渐成为实现列车状态智能诊断与运维决策的重要技术支撑[3-4]. 尤其是对受电弓、车厢内壁等列车部件加速度序列的高精度预测，不仅能够反映列车运行平稳性与乘客舒适度，还能为异常检测、安全评估及维护优化提供依据[5].

深度学习中的时序预测方法主要可分为单变量[6]与多变量[7]两类. 前者依赖单一历史序列，结构相对简单，适用于变量独立性较高的场景；后者通过多个变量的联合建模，能够捕捉系统内部的耦合关系，更适合变量交互显著的任务. 因列车部件不同位置检测的加速度信号序列具有强时序关联、多变量耦合等特性，故更适合多变量预测，对加速度信号序列进行精准预测对于疲劳寿命评估、智能运维以及系统动力学行为分析均具有重要意义.

现有研究广泛采用卷积神经网络[8]、循环神经网络[9]及长短期记忆网络[10]对铁路领域进行时序预测，但以上网络在建模长程依赖和提升计算效率方面仍存在挑战. Transformer[11-12]依靠全局注意力机制增强长序列建模能力，文献[13]结合Transformer与自适应快速探索随机树算法优化铁路运行线路；文献[14]基于Transformer框架实现高速铁路接触网异物检测. Informer[15]在Transformer基础上引入稀疏自注意力机制降低计算复杂度，文献[16]基于SVMD-Informer进行隧道区间时间序列预测；文献[17]利用1DCNN-Informer实现无人机滚动轴承的故障检测. 基于卷积神经网络的改进架构[18](Selective Kernel Networks，SKNet)核心是动态调整卷积层的感受野大小，从而捕获不同尺度的输入信息. 文献[19]提出SPPELAN-SKNet以捕捉特征多样性，实现焊接过程中的焊缝缺陷检测；文献[20]设计Li-SKNet以解决采矿过程中错误卸载指令的问题. 然而，以上Informer研究依然面临注意力模式单一、局部与全局特征难以兼顾以及多尺度信息提取能力不足等问题.

针对以上问题，本文提出选择性内核机制Informer(Selective Kernel mechanism Informer，SKInformer). 该模型以Informer为基础架构，引入选择性内核机制(Selective Kernel mechanism，SKmechanism)，通过分裂、融合与选择操作实现感受野大小的动态自适应调整，并将SKmechanism分别嵌入Informer的编码器与解码器，以同时增强模型对局部细节与全局依赖的时序特征建模能力.

1 基础理论

1.1 Informer模型

Informer由编码器和解码器组成，如图1所示. 其整体结构继承自Transformer框架，采用编码器-解码器并行架构. 与Transformer依赖全局注意力计算不同，Informer通过引入稀疏自注意力机制降低时间复杂度，使其更适用于长时间序列预测任务. Informer的编码器和解码器以并行耦合的方式构建，可以在时间序列的不同部分之间直接交换信息，无需进行严格顺序递归处理，这一设计不仅延续了Transformer并行计算的优势，而且提升了Informer

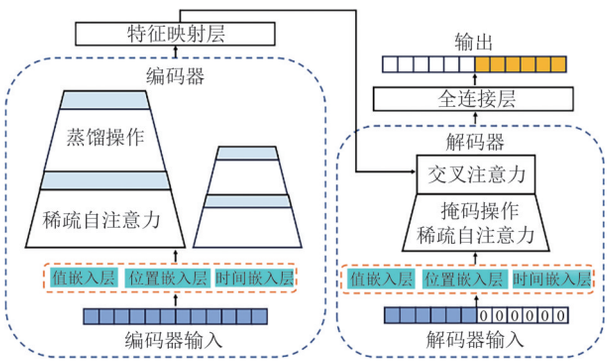


图 1　Informer 模型

Fig.1　Informer model

在时间序列任务中的训练速度与计算效率.

在输入结构方面,Informer延续了Transformer的双端输入模式,针对时间序列预测任务进行了专门改进.编码器输入侧重接收完整的历史时间序列,以学习全局依赖关系;解码器输入则由部分历史片段与目标预测步长对应的占位符共同组成,实现多步预测.编码器与解码器输入都经过结构相同的嵌入模块进行映射.该嵌入模块与Transformer中的嵌入层结构类似,均包括值嵌入与位置嵌入,但Informer进一步引入时间嵌入,用于刻画周期性特征及时间戳相关信息,从而增强模型对时间动态模式的表达能力.通过值嵌入、位置嵌入与时间嵌入的协同融合,模型能够在输入阶段充分表征原始序列的数值特征与时序结构信息,为后续注意力计算提供高质量的特征表示.

编码器与解码器的核心都为稀疏自注意力机制,该机制通过仅保留注意力贡献度最大的查询矩阵$\boldsymbol{Q}$,实现近似稀疏的注意力计算.稀疏自注意力机制为

$$\mathcal{A}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})=\mathrm{Softmax}\left(\frac{\bar{\boldsymbol{Q}}\boldsymbol{K}^{\mathrm{T}}}{\sqrt{d}}\right)\boldsymbol{V} \tag{1}$$

式中:$\mathcal{A}(\boldsymbol{Q},\boldsymbol{K},\boldsymbol{V})$为稀疏自注意力机制函数;$\boldsymbol{Q}$为查询矩阵;$\boldsymbol{K}$为键矩阵;$\boldsymbol{V}$为值矩阵;$\bar{\boldsymbol{Q}}$为筛选出的查询矩阵;Softmax(·)为把每个筛选出的查询矩阵对所有键矩阵的相似性分数转为归一化权重的函数;d为缩放因子.稀疏自注意力机制关键在于筛选最大贡献值的查询矩阵,表达式为

$$\bar{M}(\boldsymbol{q}_i,\boldsymbol{K})=\max_j\left(\frac{\boldsymbol{q}_i\boldsymbol{k}_j^{\mathrm{T}}}{\sqrt{d}}\right)-\frac{1}{N_{\mathrm{K}}}\sum_{j=1}^{N_{\mathrm{K}}}\frac{\boldsymbol{q}_i\boldsymbol{k}_j^{\mathrm{T}}}{\sqrt{d}} \tag{2}$$

式中:$\bar{M}(\boldsymbol{q}_i,\boldsymbol{K})$为筛选出的最大贡献值查询矩阵;$\boldsymbol{q}_i$为第$i$个查询向量;$\boldsymbol{k}_j^{\mathrm{T}}$为第$j$个键向量转置;$N_{\mathrm{K}}$为键矩阵的数量.

1.1.1　编码器

编码器负责从输入序列中提取深层特征并对其时序依赖建模,编码器结构如图2所示.编码器采用三级蒸馏式下采样多头自注意力堆叠架构,面对一维时序完成分层、多尺度特征提取.三级蒸馏式结构由双分支嵌入输入单元、三级串联多头注意力子模块、逐层1D卷积下采样单元、双路多尺度输出单元4部分构成:解码器输入分为两支,上支生成长度为L的时序特征,下支先下采样得到长度为$L/2$的轻量化处理后的时序特征,二者输入注意力子模块1;子模块1输出主干特征下采样长度为$L/2$时序特征并送入子模块2,旁路保留浅层时序特征;子模块2融合双路长度为$L/2$时序特征后再次下采样至$L/4$输入子模块3;子模块3提取全局时序特征,最终输出两组长度为$L/4$的多尺度时序特征.编码器的蒸馏操作表达式为

$$\boldsymbol{X}_{e+1}^{t}=\mathrm{MaxPool}\big(\mathrm{ELU}(\mathrm{Conv1d}(\boldsymbol{X}_e^{t}))\big) \tag{3}$$

式中:$\boldsymbol{X}_{e+1}^{t}$为t时刻的第$e+1$层加速度时序特征矩阵;MaxPool(·)为最大池化函数;Conv1d(·)为一维卷积函数;ELU(·)为激活函数.

1.1.2　解码器

解码器的整体结构和长度为L的时序特征处

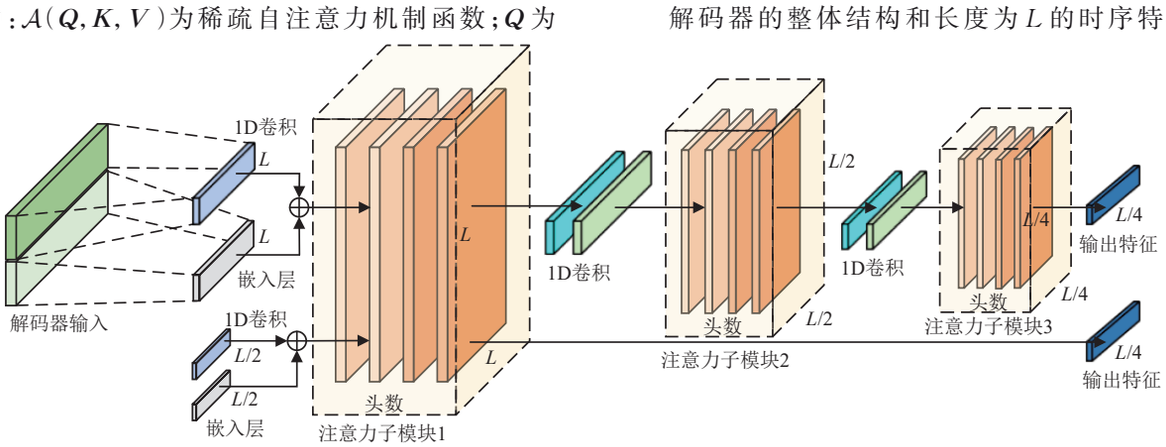


图 2　编码器结构

Fig.2　Encoder structure

理方式均与编码器相似，如图3所示．解码器的主要作用是将编码器处理后的长度为$L/4$时序特征（如长期依赖关系、全局模式和趋势信息）与解码器的目标序列（长度为L）相融合．通过将编码器和解码器相结合，生成更连贯的预测值，避免局部偏差．

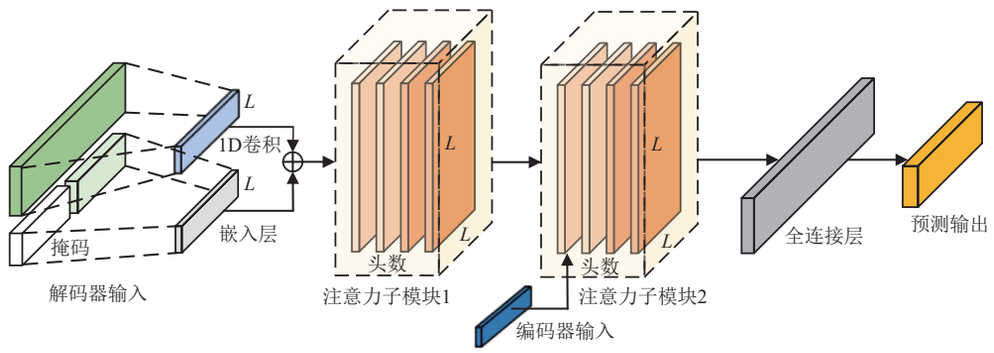


图3　解码器结构

Fig.3　Decoder structure

解码器与编码器之间最显著的区别在于，解码器的输入由两部分组成．第一部分是输入解码器自身的输入，包括目标序列和掩码；第二部分输入则是通过稀疏自注意力机制和蒸馏操作从编码器中获得的输出．解码器输入$\boldsymbol{X}_{\mathrm{de}}$为

$$\boldsymbol{X}_{\mathrm{de}}=\mathrm{Concat}(\boldsymbol{X}_{\mathrm{token}},\boldsymbol{X}_0)\tag{4}$$

式中：$\boldsymbol{X}_{\mathrm{token}}$为编码器的历史序列，长度为$L_{\mathrm{token}}$；$\boldsymbol{X}_0$为预测长度占位符，初始化为0，长度为$L_0$；Concat为沿时间步（序列长度）维度拼接的两个输入函数，拼接后的序列长度为$(L_{\mathrm{token}}+L_0)\times d_{\mathrm{model}}$，其中$d_{\mathrm{model}}$为通道维度，代表模型隐藏层维度．

1.2　SKmechanism

SKmechanism是一种用于图像特征提取的非线性注意力机制，能从多种卷积核中聚合有效信息并自适应调整感受野大小，适用于长期时序预测任务．SKmechanism结构如图4所示．图4中，分裂操作通过不同内核尺寸构建多条路径以覆盖多尺度感受野；融合操作对多路径特征进行融合以生成全局权重；选择操作根据权重自适应选择合适的卷积核以输出最终特征．

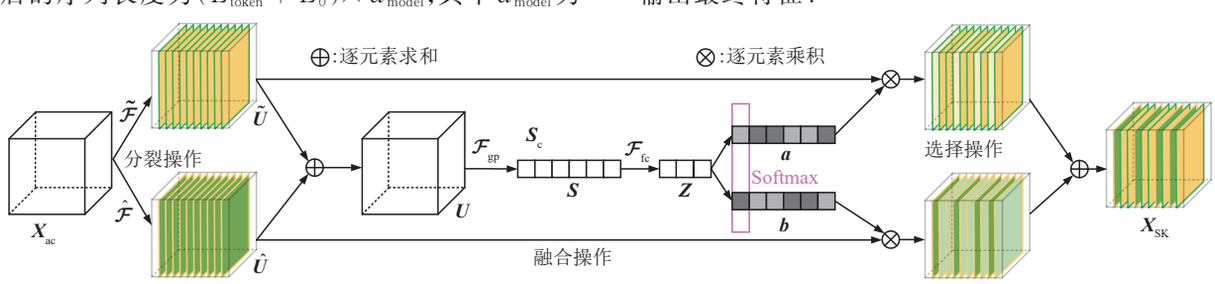


图4　选择性内核机制

Fig.4　Selective kernel mechanism

SKmechanism在时序预测任务中具备多项优势．第一，通过多尺度卷积核的选择性交互操作，能够整合不同层级的特征并增强表示能力；第二，分裂与融合操作可根据任务自动调整核尺度，提高模型的泛化性；第三，在多尺度建模的前提下，选择操作仅激活部分卷积核，以降低计算量并保持较高的推理效率．

1.2.1　分裂操作

分裂操作目的是将卷积展开，然后再从展开后的卷积获取特征．加速度时序$\boldsymbol{X}_{\mathrm{ac}}$分裂成两组互补子特征，分别为$\tilde{\mathcal{F}}:\boldsymbol{X}_{\mathrm{ac}}\to\tilde{\boldsymbol{U}}$与$\hat{\mathcal{F}}:\boldsymbol{X}_{\mathrm{ac}}\to\hat{\boldsymbol{U}}$．虽然卷积展开为两部分，但是两个特征变换的输出依然保持相同的维数．其中：$\tilde{\mathcal{F}}$代表将加速度时序$\boldsymbol{X}_{\mathrm{ac}}$切分为上分支子特征操作；$\tilde{\boldsymbol{U}}$代表分裂操作后所得上分支子特征；$\hat{\mathcal{F}}$代表将加速度时序$\boldsymbol{X}_{\mathrm{ac}}$切分为下分支子特征操作；$\hat{\boldsymbol{U}}$代表分裂操作后所得下分支子特征．

1.2.2　融合操作

融合操作的目的是将获得的上分支时序特征$\tilde{\boldsymbol{U}}$与下分支时序特征$\hat{\boldsymbol{U}}$进行特征融合，融合后的时序特征$\boldsymbol{U}$为

$$\boldsymbol{U}=\tilde{\boldsymbol{U}}+\hat{\boldsymbol{U}}\tag{5}$$

然后，将融合后的时序特征进行全局池化，$\boldsymbol{U}$通过全局平均池化生成$\boldsymbol{S}$来嵌入全局信息，其表达式为

$$\mathcal{F}_{\mathrm{gp}}(\boldsymbol{U})=\frac{1}{H\times W}\sum_{f=1}^{H}\sum_{g=1}^{W}\boldsymbol{U}(f,g)=\boldsymbol{S}\tag{6}$$

式中:$\mathcal{F}_{gp}(\cdot)$为全局平均池化函数;H为二维时序特征的高度;W为二维时序特征的宽度;$\boldsymbol{U}(f,g)$为选取时序特征$\boldsymbol{U}$中第f行、第g列的值;$\boldsymbol{S}$为经过全局平均池化的时序特征,其中$\boldsymbol{S}_c$为时序特征$\boldsymbol{S}$的第c元素的列向量,旨在通过将$\boldsymbol{U}$的第c个特征缩小$H\times W$倍来计算得到$\boldsymbol{S}$的第c个元素.全局池化的目的是整合来自所有分支的信息,使SKmechanism能够了解特征之间的关系并嵌入全局信息.为提高计算效率,采用连接层进行降维,降维表达式为

$$\mathcal{F}_{fc}(\boldsymbol{S})=\mathrm{ReLU}(\mathcal{B}(\boldsymbol{W}_S\boldsymbol{S}))=\boldsymbol{Z} \tag{7}$$

式中:$\mathcal{F}_{fc}(\cdot)$为压缩降维函数;ReLU(·)为激活函数;$\mathcal{B}(\cdot)$为批量归一化函数;$\boldsymbol{W}_S$为时序特征$\boldsymbol{S}$的权重矩阵;$\boldsymbol{Z}$为时序特征$\boldsymbol{S}$通过全连接层的压缩操作所得的列向量.

1.2.3　选择操作

选择操作目的是利用跨通道的软注意以自适应地选择不同的信息空间尺度,然后通过逐元素求和融合两路加权后的时序特征,得到经过SKmechanism处理后的输出$\boldsymbol{X}_{SK}$.对列向量$\boldsymbol{Z}$使用Softmax函数得到软注意力向量$\boldsymbol{a}$与$\boldsymbol{b}$,向量中的每一个数值对应一个通道的分数,代表其重要程度,同时再将软注意力向量$\boldsymbol{a}$与$\boldsymbol{b}$再次变回原来的维度,表达式为

$$a_c=\frac{\mathrm{e}^{\boldsymbol{A}_c\boldsymbol{Z}}}{\mathrm{e}^{\boldsymbol{A}_c\boldsymbol{Z}}+\mathrm{e}^{\boldsymbol{B}_c\boldsymbol{Z}}} \tag{8}$$

$$b_c=\frac{\mathrm{e}^{\boldsymbol{B}_c\boldsymbol{Z}}}{\mathrm{e}^{\boldsymbol{A}_c\boldsymbol{Z}}+\mathrm{e}^{\boldsymbol{B}_c\boldsymbol{Z}}} \tag{9}$$

式中:$\boldsymbol{A}$,$\boldsymbol{B}$分别为两个分支对应的不同通道的权重矩阵;$\boldsymbol{a}$与$\boldsymbol{b}$分别为$\tilde{\boldsymbol{U}}$与$\hat{\boldsymbol{U}}$的软注意力向量;$\boldsymbol{A}_c$为$\boldsymbol{A}$的第c行向量,a_c为$\boldsymbol{a}$向量的第c个元素;$\boldsymbol{B}_c$为$\boldsymbol{B}$的第c行向量,b_c为$\boldsymbol{b}$向量的第c个元素.

2　本文方法

2.1　加速度时序预处理

对实验所得受电弓数据集和LRV4313轻轨列车数据集进行[0,1]区间归一化处理.其表达式为

$$X_{norm}=\frac{X-X_{min}}{X_{max}-X_{min}} \tag{10}$$

式中:X_{norm}为归一化后的数据集;X为实验所得数据集,即原始时序;X_{min}为实验所得数据集中最小值;X_{max}为实验所得数据集中的最大值.

2.2　加速度时序维度转换

在加速度时序预测任务中,输入为三维张量[B, L, D],其中B为批量大小,D为特征维度.然而,直接使用长度为L的一维时序特征表示进行特征学习常面临一些局限:固定的三维张量结构[B, L, D]将预测加速度时序长度与特征维度明确分离,使得模型在捕获全局依赖与局部细节时存在约束;此外,在单一时间轴上,不同尺度的时间模式仅能通过常规操作进行融合,缺乏有效的空间化表征来提升特征的区分度.因此,引入加速度时序维度转换,通过重映射操作将三维张量输入[B, L, D]与二维空间形式的四维张量[B, D, H, W]相互转换,促进多尺度特征的提取与融合.该转换过程如图5所示.

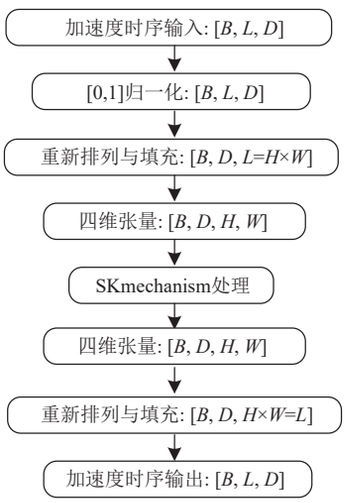


图5　加速度时序维度转换

Fig.5　Dimensional transformation of acceleration time series

2.3　SKInformer

SKInformer是将SKmechanism的分裂、融合与选择操作分别加入编码器与解码器中,通过在多尺度卷积路径上自适应选择最具代表性的特征,动态调整感受野,相较于Transformer与Informer,所提模型在继承全局依赖建模能力的同时,进一步增强了时序依赖局部建模能力并提升了关键特征的表达效率.

SKmechanism在编码器的应用位置是在稀疏自注意力机制之前,目的是通过通道级自适应,以重新标定输入特征的先验强化.该设计能够在时序建模阶段之前有效抑制冗余通道并增强判别性通道的响应,从而提升后续稀疏自注意力机制在建模跨时间依赖时的准确性和计算效率.SKmechanism与稀疏自注意力机制在功能上形成互补:稀疏自注意力机制主要关注时序维度上的长程关系进行建模,而SKmechanism则在前端执行跨通道的动态分裂、选择与融合操作,等效于在生成查询、键与值向量之前对特征通道进行重校准与增强,进而提升输入至稀疏注意力机制的特征表示质量,选择性内核机制在编码器中的应用如图6所示.

选择性内核机制在解码器的应用如图7所示.在

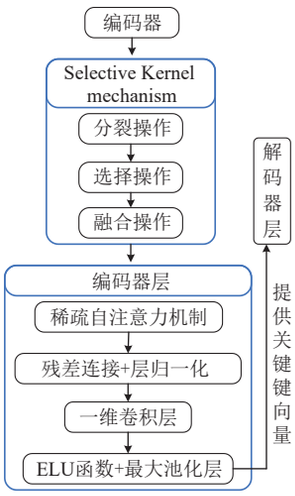


图 6 选择性内核机制在编码器中的应用

Fig.6 Application of the selective kernel mechanism in the encoder

解码器中引入 SKmechanism，作用于稀疏自注意力机制与交叉注意力机制之前的查询特征，通过通道级自适应重标定提升解码查询的通道质量．具体而言，对解码器查询序列进行通道重加权后，再送入后续注意力模块，使得交叉注意力能够聚焦于与目标通道高度相关的编码器上文信息．这一设计可减轻跨注意力的搜索负担：查询通道分布可减少在编码器输出的键值对上文信息进行无效匹配的可能，从而提升长时序或多变量场景下的相关性检索效率．

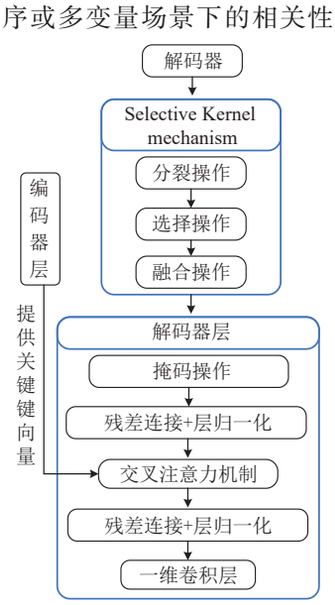


图 7 选择性内核机制在解码器中的应用

Fig.7 Application of the Selective Kernel mechanism in the decoder

2.4 评价指标

本文使用相对平方误差（Relative Squared Error，RSE）[25]、平均绝对误差（Mean Absolute Error，MAE）[26]、均方误差（Mean Squared Error，MSE）[25]、均方根误差（Root Mean Squared Error，RMSE）[26]和决定系数 R^2[27] 5种指标验证 SKInformer 的性能．指标的表达式分别为

$$\mathrm{RSE}=\frac{\sqrt{\sum_{t=1}^{N}(\hat{y}_t-y_t)^2}}{\sqrt{\sum_{t=1}^{N}(y_t-\bar{y})^2}} \tag{11}$$

$$\mathrm{MAE}=\frac{1}{N}\sum_{t=1}^{N}\left|\hat{y}_t-y_t\right| \tag{12}$$

$$\mathrm{MSE}=\frac{1}{N}\sum_{t=1}^{N}\left(\hat{y}_t-y_t\right)^2 \tag{13}$$

$$\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{t=1}^{N}(\hat{y}_t-y_t)^2} \tag{14}$$

$$R^2=1-\frac{\sum_{t=1}^{N}\left(y_t-\hat{y}_t\right)^2}{\sum_{t=1}^{N}\left(y_t-\bar{y}\right)^2} \tag{15}$$

式中：y_t 为加速度传感器在 t 时刻测得的真实值；$\hat{y}_t$ 为 SKInformer 得到的预测值；$\bar{y}$ 为加速度传感器测得的真实值的平均；N 为总样本点数．

3 实验验证

为验证所提方法的预测能力，选用实验室自建受电弓数据集进行消融实验；选用美国匹兹堡 DR-Train 公开数据集[21]中的 LRV4313 列车进行模型对比实验．

3.1 受电弓数据集

选取实验室搭建的受电弓实验平台作为研究对象．受电弓加速度数据集采样频率为 1 000 Hz，总采样点数为 25 050．受电弓装置如图 8 所示，在其关键位置共布置 7 个加速度传感器．在弓头处施加力锤激励以模拟外部扰动工况，各加速度传感器的空间分布以及力锤作用位置如图 9 所示．将此数据集划分为训练集、验证集和测试集，划分比例为 7∶1∶2，由于布置 7 个加速度传感器，故多变量输入和输出的通道数目均为 7．

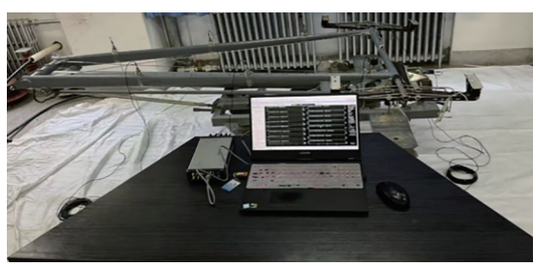


图 8 受电弓装置

Fig.8 Pantograph test rig

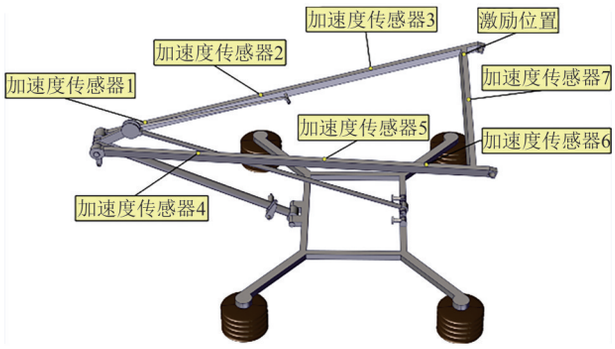


图 9　加速度传感器分布

Fig.9　Distribution of acceleration sensors

3.1.1　不同学习率下损失函数验证

为评估所提响应预测方法的有效性,降低实验过程中随机因素对结果的影响,在训练阶段设置3种不同的学习率,并通过消融实验方式验证Informer、仅编码器结合SKmechanism(Encoder+SKmechanism)、仅解码器结合SKmechanism(Decoder+SKmechanism)、SKInformer4种预测方法的效果.针对加速度传感器1~7采集的加速度响应数据,分别开展预测实验.学习率为0.001、0.000 1和0.000 01时各模型对应的训练损失、验证损失及测试损失的变化过程如图10~图12所示.

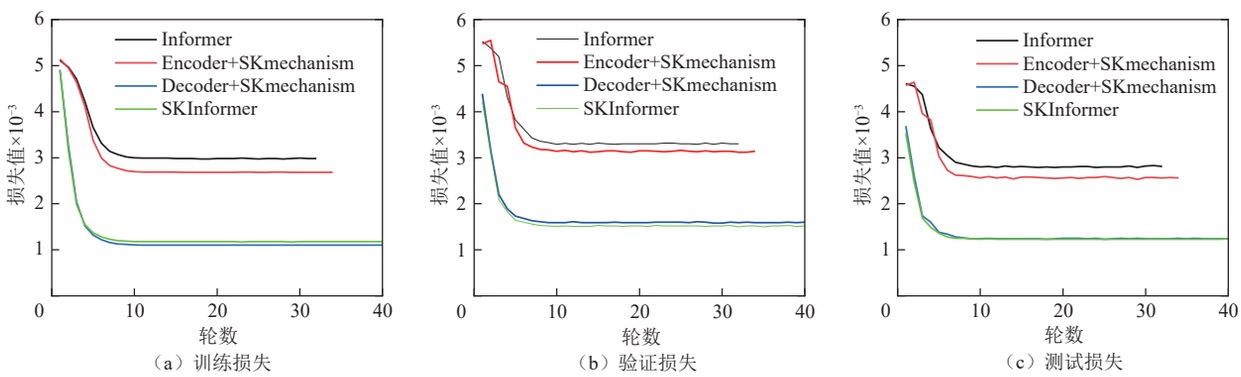


图 10　学习率为 0.001 时的损失曲线

Fig.10　Loss curves at a learning rate of 0.001

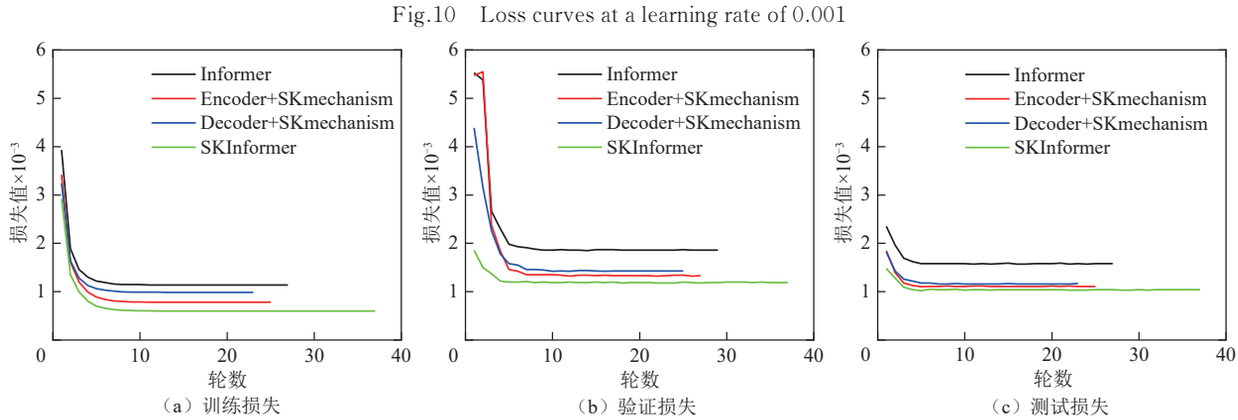


图 11　学习率为 0.000 1 时的损失曲线

Fig.11　Loss curves at a learning rate of 0.000 1

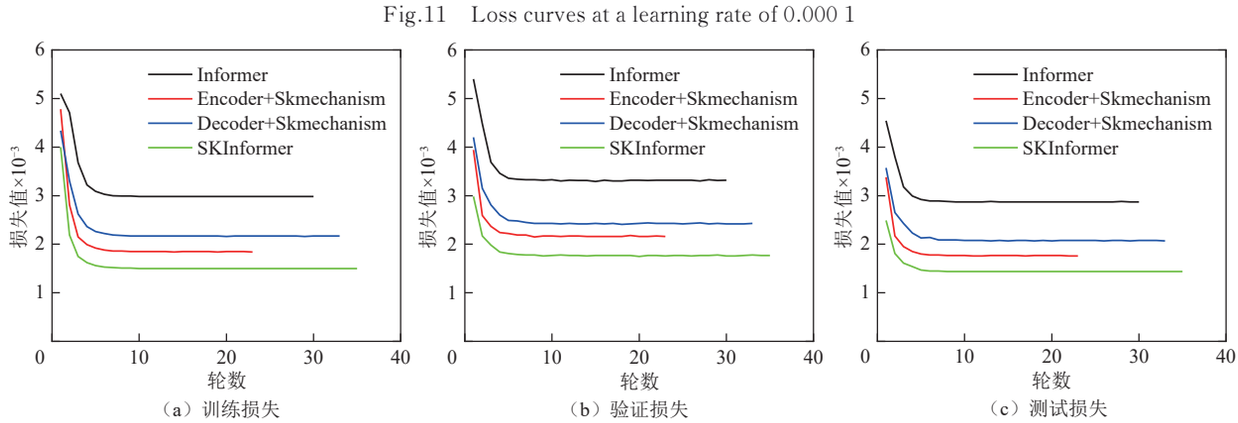


图 12　学习率为 0.000 01 时的损失曲线

Fig.12　Loss curves at a learning rate of 0.000 01

由图 10～图 12 可知：SKInformer 在 3 种学习率设置下的损失函数均在第 6 个训练周期左右开始收敛，并随后保持平稳；受益于早停机制的引入，部分实验在训练周期尚未达到 40 轮时即提前终止，有效避免了模型过拟合；进一步对比发现，在不同学习率条件下，SKInformer 在训练、验证及测试阶段均取得了低于其他 3 种对比模型的损失值，且 3 类损失曲线变化趋势一致、数值接近，表现出良好的稳定性.

实验结果表明，SKInformer 在受电弓加速度时序预测任务中具有较强的泛化能力，能够在不同学习率配置下有效抑制过拟合现象.

3.1.2　拟合效果验证

由 3.1.1 节可知，学习率为 0.000 1 时，SKInformer 具有良好的损失函数，故本节对 3 种学习率下的 SKInformer 与传感器所测得的真实值拟合效果进行对比. 当学习率为 0.001、0.000 1 和 0.000 01 时，SKInformer 模型预测结果与真实值的拟合对比如图 13～图 15 所示. 由图 13～图 15 可知，SKInformer 在学习率为 0.000 1 时，除个别极值点存在微小偏差外，7 个加速度时序通道的预测曲线在整体变化趋

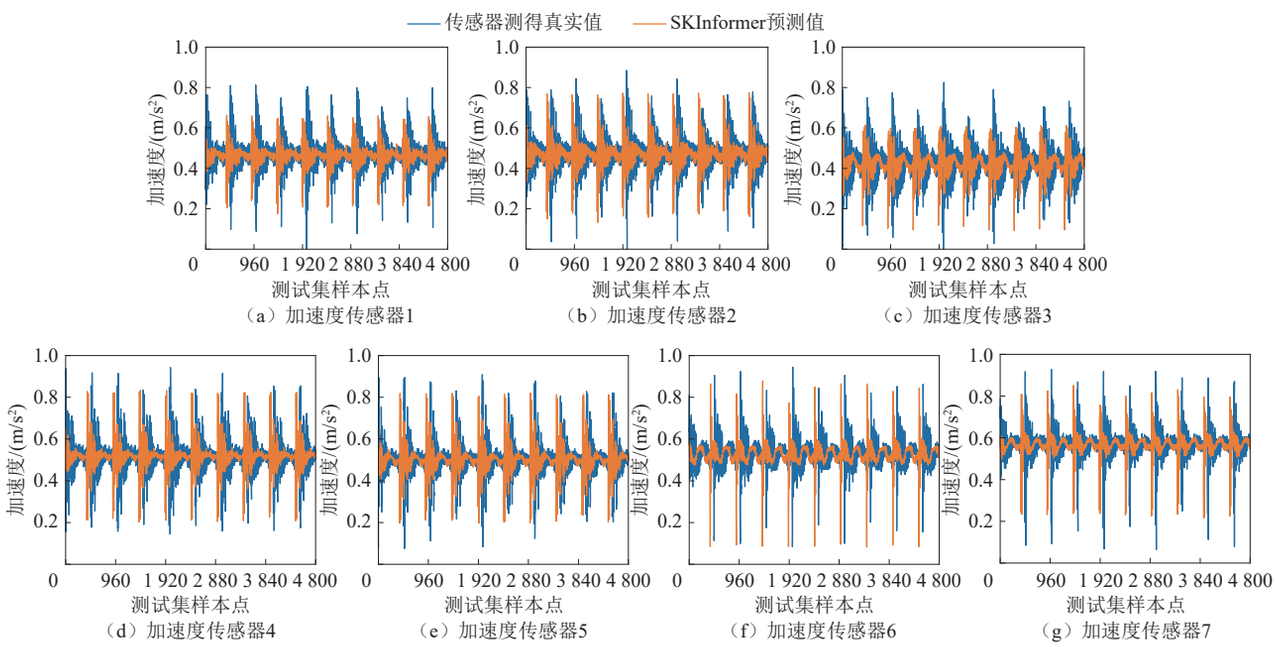


图 13　学习率为 0.001 时 7 种加速度预测值与实测值对比

Fig.13　Comparison of predicted and measured acceleration values for seven channels at a learning rate of 0.001

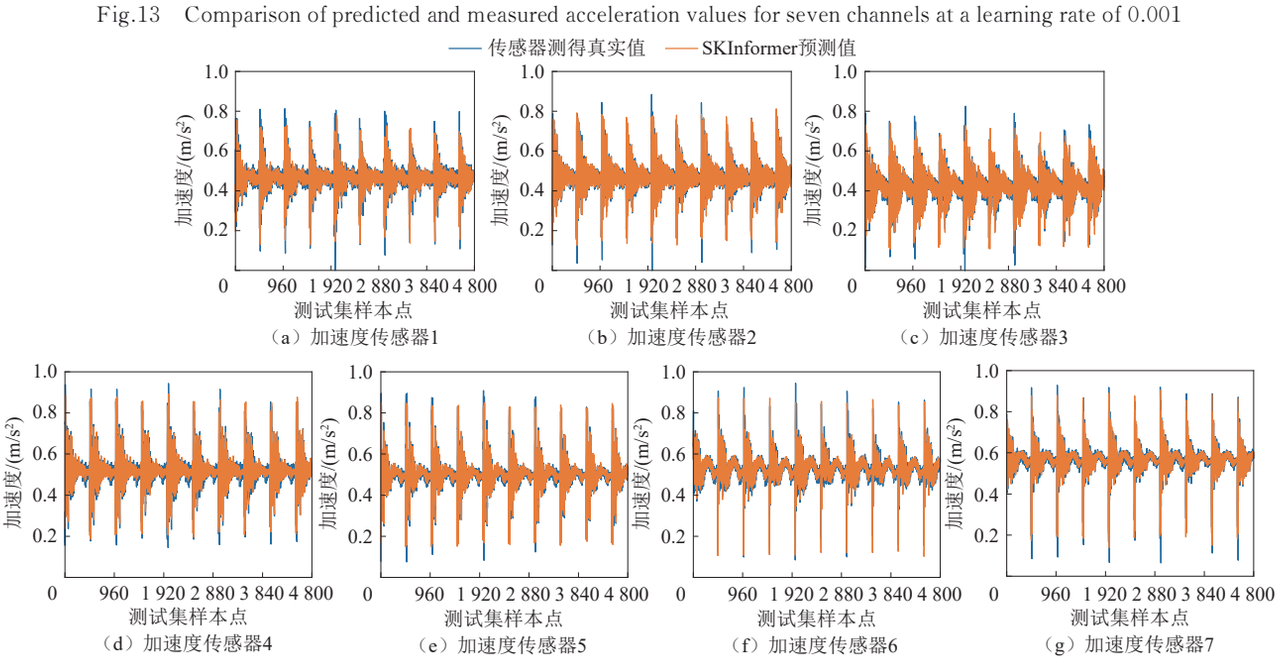


图 14　学习率为 0.000 1 时 7 种加速度预测值与实测值对比

Fig.14　Comparison of predicted and measured acceleration values for seven channels at a learning rate of 0.000 1

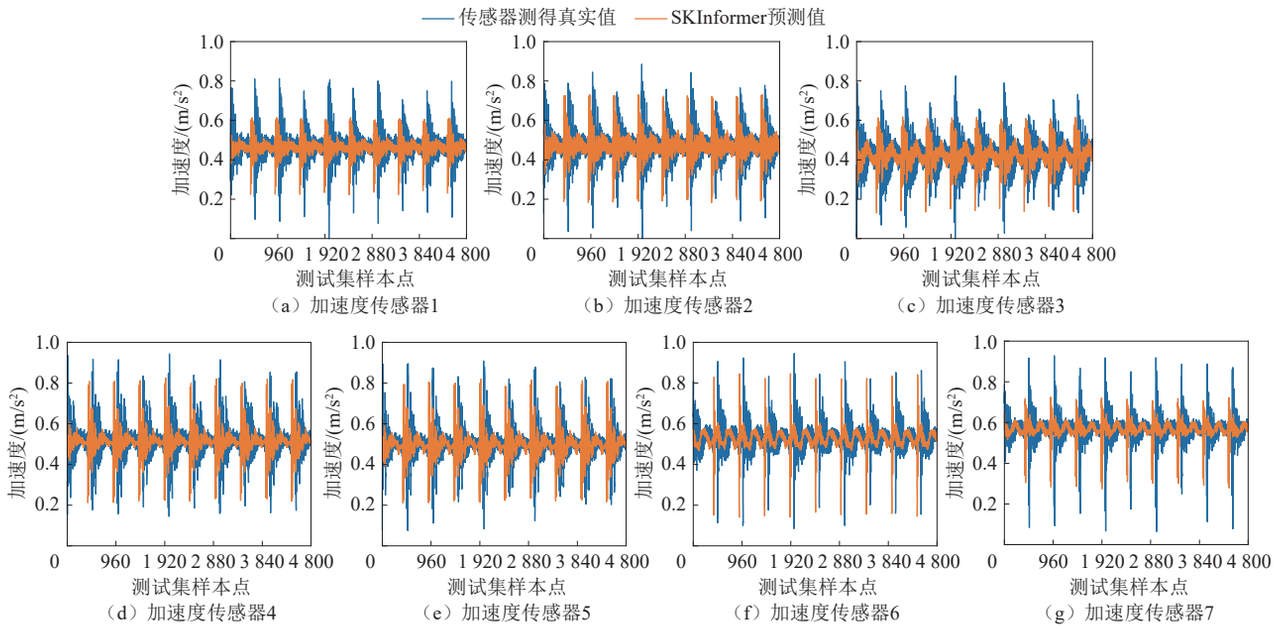


图 15　学习率为 0.000 01 时 7 种加速度预测值与实测值对比

Fig.15　Comparison of predicted and measured acceleration values for seven channels at a learning rate of 0.000 01

势上与真实信号基本一致.该结果表明:SKInformer在学习率为0.000 1时能够有效建立不同子空间输入信息之间的内在关联,从而挖掘时间序列的长期依赖特性;加速度时序中的关键动态特征得到了较好的刻画,模型能够表征7个加速度通道之间的映射关系,这进一步验证了所提模型在全局特征建模与局部细节捕捉方面的协同优势.

3.2　LRV4313列车数据集

LRV4313轻轨列车数据集采集了车厢内部的加速度时序信息,LRV4313轻轨列车如图16所示.在此数据集中,以列车行进方向定义为x轴正方向,加速度信号的采样频率为1 600 Hz,选取30 000个连续采样点用于实验分析.LRV4313轻轨列车共布置8个加速度测量通道,包括车内的两个单轴加速度传感器以及两个三轴加速度传感器,后者分别获得x轴、y轴与z轴三个方向的振动信息.LRV4313轻轨列车加速度传感器的具体安装位置与编号信息见表1.



图 16　LRV4313轻轨列车[21]

Fig.16　LRV4313 light -rail train[21]

为验证SKInformer在预测任务中的性能优势,选取具有公开源代码的Transformer[22]、Autoformer[23]及Reformer[24]作为对比模型.在此数据集中依次选前单轴加速度传感器、后单轴加速度传感器、前三轴加速度传感器-z、后三轴加速度传感器-x、后三轴加速度传感器-y和后三轴加速度传感器-z作为实验数据集,6个加速度数据集的训练集、验证集和测试集的数据量比例同样为7∶1∶2.由于选取6个加速度传感器,故多变量输入和输出通道数目都为6.

表 1　LRV4313轻轨列车加速度传感器安装位置

Tab.1　Locations of acceleration sensors on the LRV4313 light-rail train

数据集	安装位置
前单轴加速度传感器	车厢内壁靠窗位置
后单轴加速度传感器	车厢内壁座椅附近
前三轴加速度传感器(x轴,y轴,z轴)	车厢内壁车厢前端
后三轴加速度传感器(x轴,y轴,z轴)	车厢内壁车厢后端

3.2.1　对比实验评价指标验证

4种模型下的评价指标对比结果如图17所示.由图17可知:SKInformer的RSE、MAE、MSE和RMSE的指标值均最小,表明其在时序预测任务中的优势;随着预测序列长度的增加,SKInformer保持较高的预测精度;值得注意的是,当预测长度由510扩展至785时,SKInformer的RSE、MSE、MAE和RMSE并未出现性能退化,反而呈现出不同程度的提升,该现象主要得益于Informer框架与SK-mechanism的有效融合,该机制增强了模型对全局与局部时间依赖关系的建模能力;通过在多尺度特征空间中自适应地调整注意力分布与感受野范围,SKInformer能够更好地挖掘长期序列中的潜在时序特征.

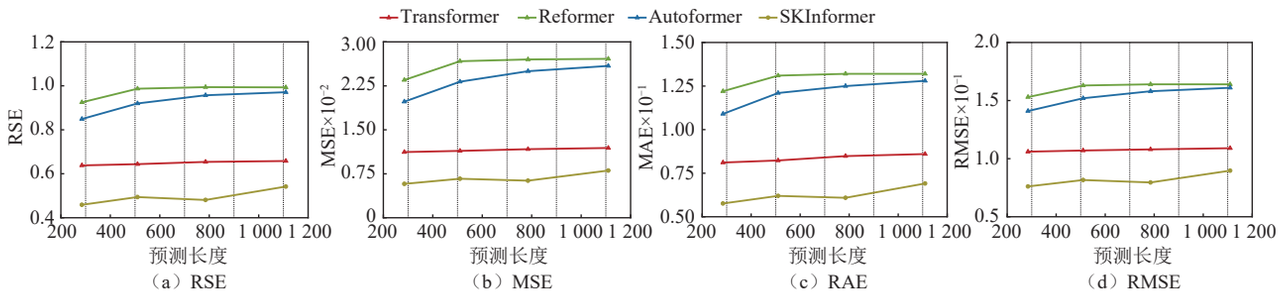


图 17　4种模型评价指标对比

Fig.17　Comparison of evaluation metrics for the four models

3.2.2　加速度时序决定系数验证

为进一步验证SKInformer的预测性能优势，评估6个传感器预测加速度时序为285时的决定系数R^2，如图18所示．图18中：横轴为实验数据集的原始值；纵轴为SKInformer的预测值；X为原始时序所对应的真实值．由图18可知：SKInformer在预测后单轴加速度时序方面，与实际观测值相比拟合性能最佳，而后三轴y方向加速度时序的预测结果与实际值的吻合度相对较差；其他4个传感器的预测结果与实际值均表现出良好的一致性．这些差异主要归因于6个加速度传感器的不同安装位置导致测量的加速度时序数据存在显著差异和较弱的相关性．结果表明，即使同时预测6个差异显著的加速度时序，SKInformer仍能保持较高的预测精度．

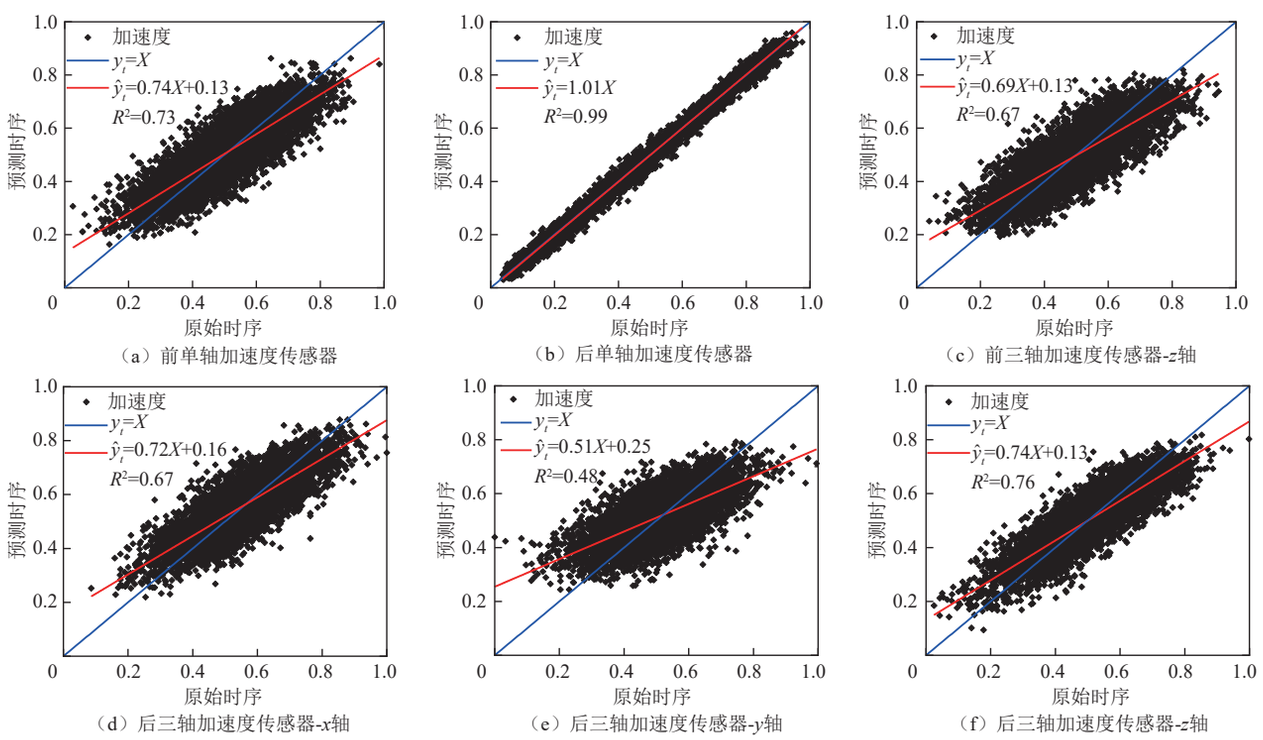


图 18　6个传感器测得的加速度 R^2值

Fig.18　Coefficients of determination (R^2) for acceleration measured by six sensors

4　结论

1）通过时序维度转换，将一维时序映射为适配SKmechanism的二维特征表示，实现从一维时序到二维图卷积的有效转换，以挖掘时序中潜在的空间相关结构，增强跨时间维度的特征交互并提升整体表征能力．

2）通过引入SKmechanism，在卷积计算中依次执行分裂、融合与选择操作，使网络可随输入特征自适应调节有效感受野尺度，从而增强多尺度表征建模能力与特征适应性．

3）将SKmechanism嵌入SKInformer编码器，通过对多分支卷积核的动态选择实现自适应感受野调节，使编码器在统一表示空间内同时建立不同时间尺度依赖，从而提升多尺度时序模式的表征能力．

4）将SKmechanism嵌入SKInformer解码器，

通过动态感受野选择自适应调节对历史序列的有效关注范围,强化对关键时间片段的响应并提取更具判别力的局部时序表征,从而提升预测性能.

研究成果借助SKInformer实现了列车零部件的加速度时序预测,通过将SKmechanism的分裂、融合和选择操作与Informer的编码器与解码器依次相结合,最终可提高SKInformer的多变量预测精度.

后续研究可在持续提升模型预测精度的基础上,进一步开展网络轻量化优化工作,以此提升模型推理效率,同时降低模型的计算开销与存储空间占用.

参考文献(References):

[1] 龚凯, 卞文君, 刘林芽, 等. 横风作用下350 km/h及以上高速列车行车安全性分析[J].北京交通大学学报, 2025, 49(2): 166－173.
Gong Kai, Bian Wenjun, Liu Linya, et al. Operational safety analysis of high-speed trains at 350 km/h and above under crosswind conditions[J]. Journal of Beijing Jiaotong University, 2025, 49(2): 166－173.(in Chinese)

[2] 龚飞杰, 上官伟, 宋鸿宇, 等. 基于深度强化学习优化LQR的虚拟重联高速列车间隔控制方法[J].铁道学报, 2025, 47(11): 116－126.
Gong Feijie, Shangguan Wei, Song Hongyu, et al. Deep Reinforcement Learning-based Optimization of LQR for Virtual Coupling Train Interval Control[J]. Journal of the China Railway Society, 2025, 47(11): 116－126.(in Chinese)

[3] Hu X, Tan L, Tang T. M2BIST-SPNet: RUL prediction for railway signaling electromechanical devices[J]. The Journal of Supercomputing, 2024, 80(12): 16744－16774.

[4] Nakiganda A, Lindahl M, Henderson C, et al. Quantification and Forecasting of Reserve Capacity from Electric Trains[J]. eTransportation, 2025: 100524.

[5] Gazafrudi S M M, Darvishpour Y, Farsi S. Hybrid ANFIS-BiLSTM-ANFIS model for remaining useful life prediction of Li-ion batteries in electric trains[J]. Results in Engineering, 2025: 108690.

[6] Wang Z, Wei A, Tang K, et al. Enhanced accuracy and interpretability of nitrous oxide emission prediction of wastewater treatment plants through machine learning of univariate time series: A novel approach of learning feature reconstruction[J]. Journal of Water Process Engineering, 2025, 71: 107263.

[7] Zhou Z, Yang J, Song J, et al. Research on shield parameter prediction method based on multivariate variable weight combination model[J]. Tunnelling and Underground Space Technology, 2026, 168(2): 107189.

[8] Mo Y, Wang X, Li S. Physics-informed Laplace neural operator-CNN hybrid model for wheel-rail force identification on high-speed railway bridges[J]. Engineering Structures, 2025, 345: 121546.

[9] Jiang L, Yi J, Wu L, et al. Rapid evaluation of seismic damage of regional railway lines based on GNN and genetic algorithm[J]. Structures, 2025, 79: 109548.

[10] Zhao H, Wei B, Zhang P, et al. Safety analysis of high-speed trains on bridges under earthquakes using a LSTM-RNN-based surrogate model[J]. Computers & Structures, 2024, 294: 107274.

[11] 刘树鑫,马志斌,邢朝健,等.基于TimesNet-Transformer的铁路继电器退化状态划分与识别方法[J/OL].铁道科学与工程学报,2026:1－14. doi:10.19713/j.cnki.43-1423/u.T20251915.(in Chinese)
Liu Shuxin, Ma zhibin, Xing zhaojian, et al. Method for division and identification of railway relay degradation states based onTimesNet-Transformer[J/OL]. Journal of Railway Science and Engineering, 2026: 1－14. doi: 10.19713/j.cnki.43-1423/u.T20251915.(in Chinese)

[12] 盛兴旺,欧志坚,郑纬奇,等.融合多头自注意力机制Transformer的震致桥上轨道变形预测方法[J].中南大学学报(自然科学版), 2025,56(12):5126－5137.
Sheng xingwang, Ou zhijian, Zheng weiqi, et al. A prediction method for seismic-induced track deformation on-bridges incorporating multi-head self-attention mechanism and Transformer model[J]. Journal of Central South University (Science and Technology), 2025, 56(12): 5126－5137.(in Chinese)

[13] Wan X, Pu H, Schonfeld P, et al. Transformer-guided optimization of railway alignments using an adaptive rapidly-exploring random tree algorithm[J]. Applied Soft Computing, 2025,185:113977.

[14] Chen Y, Zhou J, Chen C. An efficient foreign object detection model based on Transformer for high-speed railway catenary[J]. Engineering Applications of Artificial Intelligence, 2026, 164: 113313.

[15] Zhou H, Zhang S, Peng J, et al. Informer: Beyond efficient transformer for long sequence time-series forecasting[C]//Proceedings of the AAAI conference on artificial intelligence. Virtual,2021, 35(12): 11106－11115.

[16] Fu K, Xue Y, Qiu D, et al. Interval prediction of TBM tunneling performance under uncertainty using the successive variational mode decomposition (SVMD)-Informer model[J]. Automation in Construction, 2026, 181: 106656.

[17] Ma S, Shi S, Zhang Y, et al. A High-precision method

for detecting rolling bearing faultis in unmanned aerial vehicle based on improved 1DCNN-Informer model[J]. Measurement, 2025, 256: 118200.

[18] Li X, Wang W, Hu X, et al. Selective kernel networks [C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. Long Beach, 2019: 510－519.

[19] Liu L, Zhang M, Zhou X, et al. A weld defect detection algorithm designed based on YOLOv9[J]. Measurement, 2025, 257: 118630.

[20] Cui T, Pan Y, Bi Y, et al. A fuzzy decision-making method for unloading command at crushing stations based on deep learning and dynamic coal flow features slicing[J]. Engineering Applications of Artificial Intelligence, 2025, 162: 112617.

[21] Liu J, Chen S, Lederman G, et al. Dynamic responses, GPS positions and environmental conditions of two light rail vehicles in Pittsburgh[J]. Scientific Data, 2019, 6(1): 146.

[22] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need[C]//Proceedings of the 31st International Conference on neural information processing systems. Long Beach, 2017:6000－6010.

[23] Wu H, Xu J, Wang J, et al. Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting[J]. Advances in neural information processing systems, 2021, 34: 22419－22430.

[24] Kitaev N, Kaiser Ł, Levskaya A. Reformer: The efficient transformer[J]. arXiv preprint arXiv: 2001.04451, 2020.

[25] Shcherbakov M V, Brebels A, Shcherbakova N L, et al. A survey of forecast error measures[J]. World applied sciences journal, 2013, 24(24): 171－176.

[26] Karunasingha D S K. Root mean square error or mean absolute error? Use their ratio as well[J]. Information Sciences, 2022, 585: 609－629.

[27] Wang Z F, Peng Z R. Structural acceleration response reconstruction based on BiLSTM network and multi-head attention mechanism[J]. Structures, 2024, 64: 106602.