

DOI:10.19651/j.cnki.emt.2416606

单变量时间序列的 MHAGRU-MCCE 分类方法^{*}林泓¹ 刘桂雄¹ 戈燕红² 崔怀丰¹

(1. 华南理工大学机械与汽车工程学院 广州 510640; 2. 广东盈峰科技有限公司 佛山 528322)

摘要: 在单变量时间序列分类任务中,有效利用时间序列的多尺度特征、时间依赖特征对提高分类准确率至关重要。针对现有模型在综合利用多尺度特征、时间依赖特征方面局限,本文提出一种结合多尺度条件卷积增强模块(MCCE)与基于多头注意力机制的门控循环单元(MHAGRU)新型混合模型 MHAGRU-MCCE,MCCE 从不同尺度捕捉丰富时序特征,MHAGRU 侧重于提取时间序列数据中依赖关系。在 UCR 的 85 个公共数据集上,与 MACNN、AFFNet、OS-CNN、LITETime、MLP 和 LSTM-FCN 等 6 种主流基于深度学习时间序列分类模型相比,验证表明 MHAGRU-MCCE 在平均准确率(MA)上分别提升 0.66%、2.04%、3.45%、2.70%、12% 和 2.89%,并取得最高算术平均排名(AMR)=2.45、几何平均排名(GMR)=1.98,充分证明 MHAGRU-MCCE 在处理单变量时间序列分类问题上的有效性、优越性。

关键词: 单变量时间序列分类;多尺度卷积;门控循环单元;注意力机制

中图分类号: TP391.4; TN911.7 **文献标识码:** A **国家标准学科分类代码:** 460.99

Univariate time series classification approach using MHAGRU-MCCE

Lin Hong¹ Liu Guixiong¹ Ge Yanhong² Cui Huai Feng¹

(1. School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou 510640, China;

2. Guangdong Infore Technology Co., Ltd., Foshan 528322, China)

Abstract: In univariate time series classification tasks, effectively utilizing the multi-scale and time-dependent features of time series is crucial for enhancing classification accuracy. Addressing the limitations in existing models regarding the comprehensive use of multi-scale and time-dependent features, this paper introduces a new hybrid model MHAGRU-MCCE that combines the multi-scale conditional convolution and enhancement (MCCE) module with a multi-head attention mechanism based GRU (MHAGRU). MCCE captures rich temporal features across different scales, while MHAGRU focuses on extracting the dependency relationships within the time series data. On 85 public datasets from UCR, comparative validation with six mainstream deep learning-based time series classification models, including MACNN, AFFNet, OS-CNN, LITETime, MLP, and LSTM-FCN, demonstrates that MHAGRU-MCCE achieves respective improvements in mean accuracy (MA) of 0.66%, 2.04%, 3.45%, 2.70%, 12% and 2.89%. It also achieved the highest arithmetic mean rank (AMR) = 2.45 and geometric mean rank (GMR) = 1.98, fully demonstrating the effectiveness and superiority of MHAGRU-MCCE in handling univariate time series classification problems.

Keywords: univariate time series classification; multi-scale convolution; gated recurrent unit; attention mechanism

0 引言

时间序列作为一种常见的数据载体,按时间先后顺序进行排列的数据点集合,广泛存在于工业生产、环境监测、医疗健康等现实领域^[1-3],对其数据挖掘非常重要,时间序列分类(time series classification, TSC)是核心任务之一,

基于时序数据所包含多尺度特征、时间依赖特征分析将时间序列准确划分到预先定义类别标签中。目前单变量 TSC 方法主要包括传统机器学习方法^[4]、深度学习^[5-7]两类,其中 TSC 传统机器学习方法包括 k-近邻^[8]、支持向量机^[9]、随机森林^[10]等,它们均需人工选择与提取特征,提取特征复杂性、模型泛化能力不足以支撑其时序数据多噪

收稿日期:2024-08-04

^{*} 基金项目:广东省重点领域研发计划“精密仪器设备”重点专项(2022B0303040001)资助

声、时间依赖性等特性^[11]。TSC 深度学习研究方法包括多层感知机 (multi layer perceptron, MLP)、卷积神经网络 (convolutional neural network, CNN)、循环神经网络 (recurrent neural network, RNN) 等,通过端到端学习过程,自动提取时序数据中关键特征,直接应用或加以改进使用,已成为 TSC 任务的有效办法。

如文献[12]研究应用多尺度卷积神经网络 (multi-scale convolutional neural network, MCNN) 进行数据转换、局部卷积和全卷积步骤,从输入时间序列数据提取多尺度特征,并输出相应分类标签,其数据预处理过程复杂;文献[13]实验证明 MLP 在单变量 TSC 任务应用可行性,MLP 含三层全连接隐藏层架构,实现时序数据端到端学习,过于简单的 MLP 难以提取时间序列局部判别性特征;文献[14]研究多尺度注意力卷积神经网络 (multi-scale attention convolutional neural network, MACNN),通过多尺度卷积捕捉时间序列的多层次信息+注意力机制自动突出关键特征,降低数据预处理的复杂性,提升 TSC 准确率;Karim 等^[15]提出一种并行结构的长短时记忆全卷积网络 (long short-term memory fully convolutional network, LSTM-FCN),有效结合 FCN 高维特征提取能力与 LSTM 序列处理优势,在 TSC 实现中较单独 LSTM 性能显著提升;Elsayed 等^[16]对 LSTM-FCN 中 LSTM 部分替换为门控循环单元形成 (gated recurrent unit fully

convolutional network, GRU-FCN) 模型^[17],与 LSTM-FCN 相比,GRU-FCN 既提高分类精度又降低训练参数数量、缩短训练时间。LSTM-FCN、GRU-FCN 均在有效提取多尺度特征方面存在不足。

针对上述模型难以综合处理多尺度特征、时间依赖特征方面限制,本文设计多尺度条件卷积增强模块 (multi-scale conditional convolution and enhancement, MCCE) 和基于多头注意力机制的门控循环单元 (multi-head attention mechanism based GRU, MHAGRU),构建 MHAGRU-MCCE 混合模型,旨在增升 TSC 准确率。MHAGRU-MCCE 模型从时序数据中提取多尺度特征、时间依赖特征,再将这两种特征融合以构建更为丰富、综合特征表示,最后利用融合特征进行分类,提高时序分类的准确率。

1 MHAGRU-MCCE 模型结构与机理

本文所提出的面向单变量 TSC 的 MHAGRU-MCCE 混合模型结构如图 1 所示,模型包含基于多尺度条件卷积增强模块 (MCCE) 的残差结构支路 I、结合门控循环单元 (GRU)+多头注意力机制 (MHA)^[18] 序列处理支路 II 及特征融合与输出 III 等组成部分,基于 MCCE 的残差结构支路 I 由本文首次提出,结合 GRU+MHA 的序列处理支路 II 是对其进行改进。

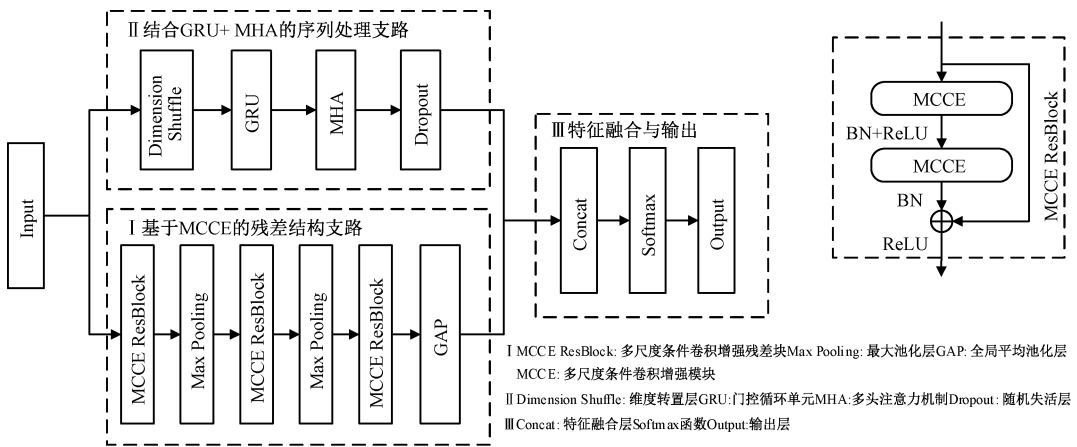


图 1 MHAGRU-MCCE 模型结构图

Fig. 1 MHAGRU-MCCE model structure diagram

其中,基于 MCCE 的残差结构支路 I 由多尺度条件卷积增强残差块 MCCE ResBlock、最大池化层 (Max Pooling) 和全局平均池化层 (global average pooling, GAP) 等组成,MCCE ResBlock (由 2 个 MCCE 堆叠残差连接而成) 提取处理不同尺度特征,Max Pooling 降低特征维度并减少模型参数数量,GAP 整合全局信息,捕捉整个序列数据的全局依赖关系,从而提升模型提取多尺度特征能力。结合 GRU+MHA 的序列处理支路 II 由维度转置层 (Dimension Shuffle)、GRU、MHA 和随机失活层 (Dropout) 等组成,Dimension Shuffle 通过将长度 N 单变

量时间序列转置为具有单个时间步长 N 维时间序列提升训练效率,GRU 对转置后数据处理从而捕捉时间序列长期依赖关系,MHA 增强序列特征捕捉能力,Dropout 特征正则化从而防止过拟合,有效提升模型提取时间依赖特征能力。特征融合与输出 III 由特征融合层 Concat、Softmax 函数和输出层 Output 等组成,Concat 融合 I、II 支路所提取多尺度特征、时间依赖特征,Softmax 函数计算类别概率、输出分类结果,Output 使用交叉熵损失函数对模型参数进行优化。下面重点分析 MCCE 与 MHAGRU 专用模块结构与原理,其中 MCCE 模块中的 MCCB①、EB②的原

创设计以及 GRU 的优化设计是难点。

1.1 多尺度条件卷积增强模块 MCCE

本文提出 MCCE 模块结构如图 2 所示,含多尺度条件卷积块 (multi-scale conditional convolution block, MCCB)、增强块 (enhancement block, EB)。其中, MCCB ①由条件卷积层 (CondConv)^[19]、特征融合层 Concat、批量归一化层 (batch normalization, BN) 和 ReLU 激活函数等组成; EB ②由 GAP、2 层全连接层 (fully connected, FC) 等组成。CondConv ③作为 MCCB 的核心块, 首先对输入数据 x 经过由 GAP、FC 和 Sigmoid 函数组成的路由函数 $r(x)$ 处理, 再生成动态调整权重参数 a_i 控制多组卷积核 W_i 的动态组合, 使得 CondConv 可以根据输入的特定条件适应不同的输入特征, 最终通过 Conv1D 层产生输出, 具体过程有:

$$\begin{cases} \alpha_i = r(x) = \text{Sigmoid} [\text{GAP}(x)\mathbf{R}] \\ \text{Output}(x) = \sigma [(\alpha_1\mathbf{W}_1 + \dots + \alpha_n\mathbf{W}_n) * x] \end{cases} \quad (1)$$

多尺度条件卷积块 MCCB ①中, 创新性采用 3 个独立并行 CondConv 结构, 每个 CondConv 不同卷积核根据输入样本动态调整其卷积核权重, 形成不同尺度特征图, 实现样本定制化特征提取^[20-21]; 不同尺度特征图通过 Concat

融合, 丰富特征图信息质量; 再经过 BN 处理, 提高网络训练过程稳定性; 最后 ReLU 激活函数, 有助于增加其非线性处理能力。因此, MCCB 显著提升模型的复杂、多变 TSC 分类性能。

增强块 EB ②中, MCCB 输出特征 Y 经 GAP 得到时序数据特征图全局信息 \bar{Y} , 其中特征图长度为 L_f 、系列特征元素 $y_n(i)$ 融合后特征元素表示为 \bar{y}_n ; 独特设计的 EB 两层结构中, 第一层 FC 利用降维系数 r 将特征图的通道数降低至原始输入的 $1/r$ 、用 ReLU 激活函数增加非线性处理, 第二层 FC 进行升维操作恢复至原始通道数、使用 Sigmoid 函数计算每个特征通道的权重 \mathbf{S} ; EB 输出权重 \mathbf{S} 与 MCCB 输出特征 \mathbf{Y} 进行逐通道加权 \otimes 得到增强后输出 \mathbf{Z} , 有效地选择、增强对分类结果有益特征, 提升模型特征表达质量。其计算式如下:

$$\begin{cases} \bar{Y} = \text{GAP}(\mathbf{Y}) \\ \bar{y}_n = \frac{1}{L_f} \sum_{i=1}^{L_f} y_n(i) \\ \mathbf{S} = \text{Sigmoid} [\mathbf{W}_2 \text{ReLU}(\mathbf{W}_1 \bar{Y})] \\ \mathbf{Z} = \mathbf{S} \otimes \mathbf{Y} \end{cases} \quad (2)$$

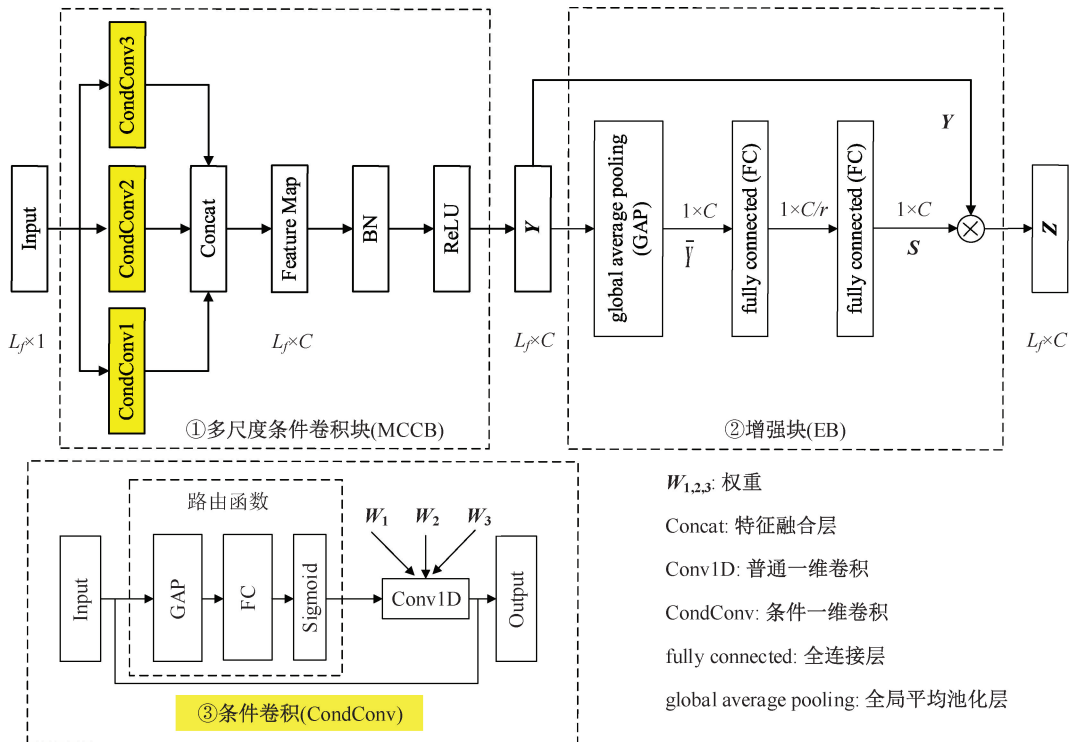


图 2 MCCE 模块结构图

Fig. 2 MCCE module structure diagram

1.2 多头注意力机制十门控循环单元 MHAGRU

在 TSC 任务中, GRU 可能不足以充分捕获时序数据在各个时间步之间复杂关系, 故本文综合应用基于多头注意力机制的门控循环单元 (MHAGRU), 以更精准地识别

时序数据的关键局部特征、筛选出重要信息, 从而增强模型的表达能力、泛化能力。MHAGRU 包括 GRU、MHA 两部分。

GRU 结构如图 3 所示, 它是为克服 RNN 在长序列依

赖、梯度消失问题而设计,与 LSTM 采用相似的门控机制,GRU 通过简化结构,仅包含重置门、更新门,实现与 LSTM 相当性能,又节省计算资源。

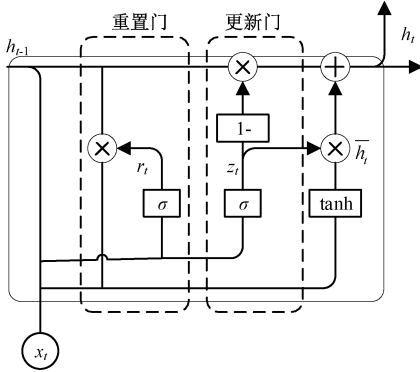


图 3 GRU 结构^[17]
Fig. 3 GRU structure

令当前时间步的输入信息为 x_t ;重置门 r_t 、更新门 z_t 分别决定前一时间步信息的忽略程度、保留程度; W_r 、 W_z 、 W_h 和 U_r 、 U_z 、 U_h 分别为 x_t 、 h_{t-1} 对应的权重矩阵, b_r 、 b_z 、 b_h 是偏置项; h_{t-1} 、 h_t 、 h_t 分别为前一时间步隐藏状态、当前时间步的候选隐藏状态及最终隐藏状态;Sigmoid 函数为 σ ,则 GRU 前向传播的计算过程为:

$$\begin{cases} r_t = \sigma(W_r x_t + U_r h_{t-1} + b_r) \\ z_t = \sigma(W_z x_t + U_z h_{t-1} + b_z) \\ \bar{h}_t = \tanh(W_h x_t + U_h (r_t \odot h_{t-1}) + b_h) \\ h_t = z_t \odot h_{t-1} + (1 - z_t) \odot \bar{h}_t \end{cases} \quad (3)$$

MHA 结构如图 4 所示,通过并行处理输入序列的各个部分,显著提升模型性能。

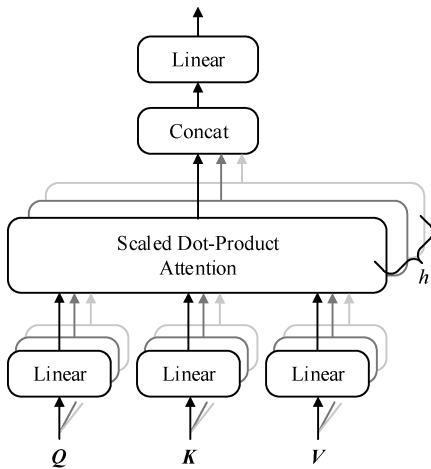


图 4 MHA 结构^[18]
Fig. 4 MHA structure

首先,MHA 设定 h 个独立的注意力“头”,每个注意力“头”分别对 GRU 输出向量执行三次不同线性映射,生成一组查询矩阵 Q 、键矩阵 K 和值矩阵 V ;再采用缩放点积

(Scaled Dot-Product)操作、Softmax 函数对查询矩阵 Q 与键矩阵 K 的相似度进行计算和归一化得到权重矩阵,与相应的值矩阵 V 相乘后,产生加权和输出,每个注意力“头”独立完成此过程,学习不同特征;最后将每个注意力“头”学到特征拼接并做线性变换得最终结果。令矩阵 K 的维度大小为 d_k 、 Q 、 K 、 V 权重矩阵分别为 WQ_i 、 WK_i 、 WV_i ,拼接操作为 Concat,第 i 个注意力“头”学到特征为 $head_i$,线性变换权重矩阵为 W^o ,则有:

$$\begin{cases} Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \\ head_i = Attention(QW_i^Q, KW_i^K, VW_i^V) \\ MultiHead(Q, K, V) = Concat(head_1, \dots, head_h)W^o \end{cases} \quad (4)$$

以上系统地分析 MHAGRU-MCCE 模型结构、专用模块结构及原理。

2 实验与结果分析

2.1 实验数据集与参数设置、对比模型

为系统比较 MHAGRU-MCCE 性能,本文选择行业著名的美国加利福尼亚大学河滨分校(university of california riverside, UCR)时间序列归档^[22]作为实验数据集,包括覆盖电气设备、传感器读数、运动捕捉、图像轮廓、心电图、光谱仪和模拟信号等领域的 85 个单变量时间序列数据集,均已预先划分为训练集、测试集,序列长度 24~2 709、训练集样本数 16~8 926、类别数量 2~60。UCR 数据集多样性、复杂性,可充分验证 MHAGRU-MCCE 模型分类性能、泛化能力。

MHAGRU-MCCE 模型采用 keras 框架编写;设置 Max Pooling = 2, Dropout 层失活率 0.8, 3 个 MCCE Resblock 卷积核数量分为 64、128、256;训练在 NVIDIA GeForce RTX 3090 GPU 进行,设置优化器为 Adam、初始学习率=0.001、批处理=128、训练轮数=2 000 等选项。MCCB 设置 3 个 CondConv 卷积核大小=3、6、12,卷积步长=1,不同卷积核组数量=2;EB 设置降维系数 $r=16$;MHAGRU 设置 GRU 记忆细胞数=8、64、128,MHA“头”数=2。

对比模型选择与 MHAGRU-MCCE 模型结构相似的现有深度学习模型,如采用多尺度卷积+注意力机制的 MACNN(2021)^[14]、AFFNet(2022)^[5];采用多尺度卷积的 OS-CNN(2022)^[6]、LITETime(2023)^[7]、及典型 MLP^[13]、LSTM-FCN^[15]等 6 种。

2.2 评价指标

选用平均准确率(mean accuracy, MA)、算术平均排名(arithmetic mean rank, AMR)、几何平均排名(geometric mean rank, GMR)以及 Nemenyi 非参数统计检验^[14]等评价模型性能。其中:

MA、AMR、GMR 指标分别表示模型在所有数据集上

的分类正确率算术平均值、准确率排名算术平均值、准确率排名几何平均值,MA 越大,AMR 及 GMR 越小,模型综合性能越好。令数据集总个数为 N ,在第 i 个数据集上的 Top-1 准确率、准确率排名分别 a_i, r_i ,则这些指标定义为:

$$\begin{cases} MA = \frac{1}{N} \sum_{i=1}^N a_i \\ AMR = \frac{1}{N} \sum_{i=1}^N r_i \\ GMR = \sqrt[N]{\prod_{i=1}^N r_i} \end{cases} \quad (5)$$

Nemenyi 非参数统计检验步骤如下:首先计算临界差值(critical difference, CD),模型个数为 k, q_α 由置信度、模型个数

决定(取置信度 $\alpha = 0.05$ 、查表 $q_\alpha = 2.949$),计算式为:

$$CD = q_\alpha \sqrt{\frac{k(k+1)}{6N}} \quad (6)$$

再显著性差异判断。若两个模型 AMR 差值超过 CD,则以相应置信度拒绝“两个模型性能相同”假设。

2.3 实验结果分析

MHAGRU-MCCE 与 6 种对比模型在 85 个 UCR 数据集上取得的 Top-1 准确率及 MA、AMR、GMR 指标结果表如表 1、2 所示,表 2 由表 1 计算得到;Nemenyi 非参数统计检验显著性差异图如图 5 所示,图 5 由表 2 及式(6)计算得到,本文研究模型 1 个加对比模型 6 个,即 $k=7$ 。表 1 粗体数值表示每个数据集上的最高准确率,获胜次数 Win 表示各模型在所有数据集上取得最高准确率次数。

表 1 7 种模型在 85 个 UCR 数据集上 Top-1 准确率结果对比

Table 1 Comparison of Top-1 accuracy results for 7 models on 85 UCR datasets

数据集名称	MACNN (2021)	AFFNet (2022)	OS-CNN (2022)	LITETime (2023)	经典的 MLP	经典的 LSTM-FCN	MHAGRU- MCCE
Adiac	0.812	0.831	0.835	0.834	0.752	0.862	0.880
ArrowHead	0.855	0.800	0.838	0.840	0.823	0.897	0.897
Beef	0.907	0.833	0.807	0.767	0.833	0.800	0.833
BeetleFly	0.995	0.850	0.815	0.900	0.850	0.900	1.000
BirdChicken	0.990	0.900	0.885	0.900	0.800	0.950	0.900
Car	0.910	0.900	0.933	0.933	0.833	0.950	0.933
CBF	0.999	0.994	1.000	0.999	0.860	0.996	0.997
Chlorine	0.882	0.862	0.839	0.848	0.872	0.803	0.871
CinCECGTorso	0.877	0.761	0.827	0.873	0.842	0.879	0.867
Coffee	1.000	1.000	1.000	1.000	1.000	1.000	1.000
Computers	0.814	0.840	0.707	0.812	0.540	0.836	0.836
CricketX	0.854	0.844	0.855	0.831	0.569	0.782	0.823
CricketY	0.863	0.869	0.867	0.877	0.595	0.777	0.826
CricketZ	0.870	0.869	0.863	0.862	0.592	0.818	0.849
DiatomSizeR	0.970	0.941	0.977	0.961	0.964	0.944	0.954
DisPhxAgeGp	0.762	0.741	0.738	0.712	0.827	0.833	0.784
DisPhxCorr	0.775	0.764	0.766	0.761	0.810	0.807	0.826
DisPhxTW	0.682	0.698	0.664	0.712	0.747	0.788	0.719
Earthquakes	0.748	0.748	0.670	0.734	0.792	0.804	0.784
ECG200	0.915	0.920	0.908	0.920	0.920	0.870	0.910
ECG5000	0.946	0.941	0.940	0.943	0.935	0.943	0.946
ECGFiveDays	0.997	1.000	1.000	0.998	0.970	0.992	0.929
ElectricDevices	0.691	0.743	0.724	0.703	0.580	0.747	0.764
FaceAll	0.869	0.918	0.845	0.793	0.885	0.922	0.951
FaceFour	0.956	0.943	0.951	0.955	0.830	0.943	0.932
FacesUCR	0.976	0.976	0.967	0.962	0.815	0.936	0.955
FiftyWords	0.862	0.815	0.816	0.807	0.712	0.802	0.815
Fish	0.989	0.971	0.987	0.977	0.874	0.954	0.983
FordA	0.950	0.951	0.955	0.958	0.769	0.932	0.965

续表 1
Continuation table 1

数据集名称	MACNN (2021)	AFFNet (2022)	OS-CNN (2022)	LITETime (2023)	经典的 MLP	经典的 LSTM-FCN	MHAGRU- MCCE
FordB	0.868	0.856	0.838	0.853	0.629	0.895	0.865
GunPoint	0.998	0.993	0.999	1.000	0.933	1.000	1.000
Ham	0.822	0.829	0.704	0.714	0.714	0.733	0.819
HandOutlines	0.938	0.897	0.929	0.954	0.807	0.865	0.957
Haptics	0.532	0.506	0.510	0.565	0.461	0.487	0.532
Herring	0.675	0.672	0.608	0.766	0.687	0.656	0.719
InlineSkate	0.483	0.336	0.429	0.551	0.351	0.489	0.515
InsectWing	0.643	0.652	0.635	0.633	0.631	0.657	0.653
ItalyPower	0.968	0.975	0.947	0.967	0.966	0.962	0.975
LrgKitApp	0.915	0.917	0.896	0.891	0.480	0.885	0.912
Lightning2	0.803	0.852	0.807	0.754	0.721	0.787	0.787
Lightning7	0.841	0.849	0.793	0.836	0.644	0.795	0.904
Mallat	0.975	0.958	0.964	0.959	0.936	0.973	0.974
Meat	0.978	0.950	0.947	0.933	0.933	0.983	1.000
MedicalImages	0.774	0.791	0.769	0.786	0.729	0.789	0.811
MidPhxAgeGp	0.609	0.545	0.536	0.513	0.735	0.753	0.656
MidPhxCorr	0.829	0.818	0.814	0.849	0.760	0.832	0.849
MidPhxTW	0.581	0.500	0.519	0.500	0.609	0.609	0.597
MoteStrain	0.910	0.899	0.926	0.883	0.869	0.930	0.915
NonInvThor1	0.945	0.952	0.963	0.960	0.942	0.962	0.972
NonInvThor2	0.953	0.952	0.960	0.961	0.943	0.957	0.968
OliveOil	0.850	0.767	0.787	0.800	0.400	0.733	0.833
OSULeaf	0.971	0.983	0.940	0.955	0.570	0.983	0.992
PhalCorr	0.826	0.815	0.830	0.840	0.830	0.828	0.841
Phoneme	0.336	0.359	0.305	0.337	0.098	0.325	0.361
Plane	1.000	1.000	1.000	1.000	0.981	1.000	1.000
ProxPhxAgeGp	0.847	0.844	0.844	0.854	0.824	0.868	0.883
ProxPhxCorr	0.916	0.918	0.908	0.918	0.887	0.911	0.921
ProxPhxTW	0.787	0.800	0.773	0.805	0.797	0.818	0.829
RefrigerationDevices	0.579	0.552	0.503	0.496	0.371	0.523	0.589
ScreenType	0.635	0.613	0.526	0.589	0.408	0.619	0.672
ShapeletSim	0.994	1.000	0.799	0.750	0.483	0.983	0.994
ShapesAll	0.933	0.937	0.920	0.913	0.775	0.897	0.920
SmlKitApp	0.796	0.795	0.721	0.771	0.389	0.747	0.779
SonyAIBORobot1	0.979	0.895	0.980	0.835	0.727	0.958	0.950
SonyAIBORobot2	0.959	0.967	0.954	0.941	0.839	0.973	0.979
StarLightCurves	0.976	0.978	0.975	0.977	0.957	0.966	0.981
Strawberry	0.973	0.976	0.982	0.984	0.967	0.962	0.978
SwedishLeaf	0.961	0.976	0.971	0.962	0.893	0.976	0.987
Symbols	0.975	0.975	0.961	0.983	0.853	0.983	0.980
SyntheticControl	0.999	1.000	0.999	1.000	0.950	0.987	1.000
ToeSegmentation1	0.966	0.974	0.954	0.978	0.601	0.982	0.965
ToeSegmentation2	0.939	0.923	0.946	0.938	0.746	0.908	0.931

续表 1
Continuation table 1

数据集名称	MACNN (2021)	AFFNet (2022)	OS-CNN (2022)	LITETime (2023)	经典的 MLP	经典的 LSTM-FCN	MHAGRU- MCCE
Trace	1.000	1.000	1.000	1.000	0.820	0.250	1.000
TwoLeadECG	0.999	0.998	0.999	1.000	0.886	0.995	1.000
TwoPatterns	1.000	1.000	1.000	0.998	0.853	0.999	1.000
UWaveGestAll	0.956	0.951	0.942	0.912	0.954	0.840	0.841
UWaveGestX	0.835	0.853	0.822	0.829	0.768	0.756	0.779
UWaveGestY	0.779	0.776	0.757	0.762	0.703	0.765	0.787
UWaveGestZ	0.784	0.791	0.764	0.781	0.705	0.958	0.965
Wafer	0.999	0.999	0.998	0.999	0.996	0.998	1.000
Wine	0.839	0.722	0.744	0.667	0.796	0.722	0.889
WordSynonyms	0.758	0.726	0.742	0.718	0.594	0.671	0.712
Worms	0.849	0.831	0.765	0.818	0.343	0.591	0.896
WormsTwoClass	0.859	0.857	0.657	0.766	0.597	0.779	0.844
Yoga	0.916	0.879	0.911	0.918	0.855	0.907	0.896
获胜次数 Win	18	16	10	15	3	15	38

表 2 7 种模型评价指标结果对比

Table 2 Comparison of results for evaluation metrics across 7 models

模型	MA	AMR	GMR
MHAGRU-MCCE	0.871 6	2.45	1.98
MACNN	0.865 0	3.07	2.62
AFFNet	0.851 2	3.62	3.05
OS-CNN	0.837 1	4.32	3.77
LITETime	0.844 6	3.80	3.24
MLP	0.751 6	5.94	5.46
LSTM-FCN	0.842 7	3.91	3.24

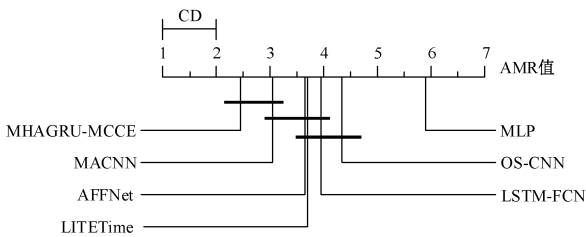


图 5 Nemenyi 非参数统计检验显著性差异图

Fig. 5 Nemenyi non-parametric statistical test significance difference diagram

由表 1、表 2、图 5 可以看出:

1) MHAGRU-MCCE 模型在 UCR 数据集中 85 个的 38 个取得最高 Top-1 准确率,即 $Win=38$,远优于其他对比模型的 18、16、10、15、3、15 个;MHAGRU-MCCE 在 MA、AMR 和 GMR 指标上均取得最优结果,即 $MA=87.16\%$ 、 $AMR=2.45$ 名、 $GMR=1.98$ 名,均优于其他对

比模型的评价指标。

2) 与串联模式多尺度卷积+注意力机制 MACNN、AFFNet 比较,MHAGRU-MCCE 模型 MA、AMR、GMR 分别提升 0.66%、2.04%、0.62 名、1.17 名及 0.64 名、1.07 名,表明基于 MCCE 残差结构支路、结合 GRU+MHA 序列处理支路并联结构的更好地捕捉多尺度特征、时间依赖特征。

3) 与多尺度卷积 OS-CNN、LITETime 比较,MHAGRU-MCCE 模型 MA、AMR、GMR 分别提升、3.45%、2.70%、1.87 名、1.35 名及 1.79 名、1.26 名,表明在多尺度卷积基础上结合 3 个并行 CondConv+EB 结构,能更有效地捕捉多尺度特征,各个尺度特征权重也得到动态调整。

4) 与典型 MLP、LSTM-FCN 比较,MHAGRU-MCCE 模型 MA、AMR、GMR 分别提升 12%、2.89%、3.49 名、1.46 名及 3.48 名、1.26 名,表明 MHAGRU-MCCE 通过堆叠多个 MCCE ResBlock,深层网络结构充分挖掘数据之间非线性关系,有效提取、利用对分类任务关键的潜在时间序列特征。

5) 在置信度 $\alpha=0.05$ 、 $k=7$ 的临界差值 $CD=0.98$,MHAGRU-MCCE 模型比其他 6 个模型的 AMR 指标更小,模型综合性能更好;MHAGRU-MCCE、MACNN 模型 $\Delta AMR=0.62 < CD$,可认为 MACNN 模型 MHAGRU-MCCE 模型性能相同。

3 结 论

针对现有 TSC 模型在多尺度特征、时间依赖特征方面综合利用不足,研究单变量时间序列的 MHAGRU-MCCE

分类方法,主要工作包括:创新提出结合多尺度条件卷积增强模块与基于多头注意力机制的门控循环单元的混合模型 MHAGRU-MCCE,通过堆叠多个 MCCE ResBlock 深层网络结构能更好地挖掘数据之间非线性关系;在多尺度卷积基础上结合 3 个并行 CondConv+EB 的独特结构,能有效地捕捉多尺度特征,使各个尺度特征权重得到动态调整。设计选择 UCR 的 85 个单变量时间序列数据集作为实验数据集;选用平均准确率 MA、算术平均排名 AMR、几何平均排名 GMR 等评价指标;应用 MACNN、AFFNet、LITETime、OS-CNN、MLP、LSTM-FCN 等 6 种模型作为对比模型的实验方案,设置模型参数。实验表明 MHAGRU-MCCE 模型在 UCR 数据集中 85 个的 38 个取得最高 Top-1 准确率,远优于其他对比模型的 18、16、10、15、3、15;MHAGRU-MCCE 在 MA、AMR 和 GMR 指标上均取得最优结果,即 MA = 87.16%、AMR = 2.45 名、GMR = 1.98 名,均优于其他对比模型的评价指标,证明 MHAGRU-MCCE 分类、泛化性能的优越性。后续将结合广东省重点领域研发计划专项(2022B0303040001)内容开展 MHAGRU-MCCE 相关应用研究。

参考文献

- [1] 李泓沛,刘桂雄,邓威. 动态工况模拟下动力锂电池建模与参数辨识方法[J]. 电子测量技术, 2022, 45(14): 103-108.
LI H P, LIU G X, DENG W. Modeling and parameter identification method for power lithium batteries under dynamic working condition simulation [J]. Electronic Measurement Technology, 2022, 45(14): 103-108.
- [2] 黄春荣,刘桂雄,戈燕红,等. 应用 GA-RBFNN 全光谱水体 COD 与 NO₃-N 检测方法研究[J]. 中国测试, 2024, 50(2): 66-72.
HUANG CH R, LIU G X, GE Y H, et al. Research on GA-RBFNN full spectrum water body COD and NO₃-N detection methods[J]. China Measurement & Test, 2024, 50(2): 66-72.
- [3] 陈韬文,宋家骏,彭湘安,等. 心电图 ECG 信号自动检测特征提取方法研究进展[J]. 电子测量技术, 2022, 45(19): 106-112.
CHEN T W, SONG J J, PENG X AN, et al. Overview of automatic detection method of ECG signal based on feature extraction [J]. Electronic Measurement Technology, 2022, 45(19): 106-112.
- [4] HOLDER C, MIDDLEHURST M, BAGNALL A. A review and evaluation of elastic distance functions for time series clustering[J]. Knowledge and Information Systems, 2024, 66(2): 765-809.
- [5] WANG T, LIU ZH Y, ZHANG T, et al. Adaptive feature fusion for time series classification [J]. Knowledge-Based Systems, 2022, 243: 108459.
- [6] TANG W S, LONG G D, LIU L, et al. Omni-scale cnns: A simple and effective kernel size configuration for time series classification[C]. 2022 10th International Conference on Learning Representations (ICLR), 2022:1-17.
- [7] ISMAIL-FAWAZ A, DEVANNE M, BERRETTI S, et al. Lite: Light inception with boosting techniques for time series classification [C]. 2023 IEEE 10th International Conference on Data Science and Advanced Analytics(DSAA), 2023: 1-10.
- [8] HAJIZADEH R. Unconstrained neighbor selection for minimum reconstruction error-based K-NN classifiers[J]. Complex & Intelligent Systems, 2023, 9 (5): 5715-5730.
- [9] MA F, WANG CH L, ZENG ZH. SVM-based subspace optimization domain transfer method for unsupervised cross-domain time series classification[J]. Knowledge and Information Systems, 2023, 65(2): 869-897.
- [10] CHENG W X, SUGANTHAN P N, KATUWAL R. Oblique random forests on residual network features[C]. Neural Information Processing: 27th International Conference, 2020: 306-317.
- [11] MIDDLEHURST M, SCHÄFER P, BAGNALL A. Bake off redux: A review and experimental evaluation of recent time series classification algorithms[J]. Data Mining and Knowledge Discovery, 2024, 38(4): 1-74.
- [12] CUI ZH CH, CHEN W L, CHEN Y X. Multi-scale convolutional neural networks for time series classification [J]. ArXiv preprint arXiv: 1603.06995, 2016.
- [13] WANG ZH G, YAN W ZH, OATES T. Time series classification from scratch with deep neural networks: A strong baseline [C]. 2017 International Joint Conference on Neural Networks (IJCNN), 2017: 1578-1585.
- [14] CHEN W, SHI K. Multi-scale attention convolutional neural network for time series classification [J]. Neural Networks, 2021, 136: 126-140.
- [15] KARIM F, MAJUMDAR S, DARABI H. Insights into LSTM fully convolutional networks for time series classification [J]. IEEE Access, 2019, 7: 67718-67725.
- [16] ELSAYED N, MAIDA A S, BAYOUMI M. Gated recurrent neural networks empirical utilization for time series classification[C]. 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and

- IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2019: 1207-1210.
- [17] CHUNG J Y, GULCEHRE C, CHO K, et al. Empirical evaluation of gated recurrent neural networks on sequence modeling[J]. ArXiv preprint arXiv:1412.3555, 2014.
- [18] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. ArXiv preprint arXiv:1706.03762, 2017.
- [19] YANG B, BENDER G, LE Q V, et al. Condconv: Conditionally parameterized convolutions for efficient inference[J]. ArXiv preprint arXiv:1904.04971, 2019.
- [20] BUZA K, ANTAL M. Convolutional neural networks with dynamic convolution for time series classification[C]. International Conference on Computational Collective Intelligence, 2021: 304-312.
- [21] WANG T, LIU ZH Y, ZHANG T, et al. Time series classification based on multi-scale dynamic convolutional features and distance features[C]. 2021 2nd Asia Symposium on Signal Processing (ASSP), 2021: 239-246.
- [22] DAU H A, BAGNALL A, KAMGAR K, et al. The UCR time series archive[J]. IEEE/CAA Journal of Automatica Sinica, 2019,6(6): 1293-1305.

作者简介

林泓, 硕士研究生, 主要研究方向为智能检测与仪器仪表。

E-mail: 202221002160@mail.scut.edu.cn

刘桂雄(通信作者), 博士, 教授, 主要研究方向为先进传感与仪器研究。

E-mail: megxliu@scut.edu.cn

戈燕红, 高级工程师, 主要研究方向为在线监测技术研究及产品开发。

E-mail: geyh@infore.com

崔怀丰, 博士研究生, 主要研究方向为智能检测与仪器仪表。

E-mail: chfsecond@163.com