

DOI:10.3969/j.issn.1003-5060.2023.12.009

# 基于样本间潜在关系的多变量时间序列分类

唐胜唐, 吴共庆, 台昌杨, 杨泽, 张赞

(合肥工业大学 计算机与信息学院, 安徽 合肥 230601)

**摘要:**多变量时间序列(multivariate time series, MTS)分类任务旨在确定多变量时间序列样本的标签。多变量时间序列数据存在时序关系和样本相似性关系等丰富的关系信息,然而现有的算法未能充分利用关系信息导致分类性能难以提升。基于此,文章提出一种基于图卷积网络(graph convolutional network, GCN)的多变量时间序列分类方法,通过挖掘样本间的潜在关系来提高分类性能。为了有效表示样本关系,设计基于样本相似度的构图规则,对样本数据进行建模从而将样本的时序特征和潜在关系信息映射到图空间中,提出基于图卷积的分类模型,通过聚合样本特征来捕获有利于分类的潜在样本关系,更新到样本自身特征向量以提升分类精度。在11个公共数据集上的大量实验结果表明,该文所提算法优于12种对比算法,可见通过挖掘时间序列数据之间潜在的关系用于分类对分类结果具有重要影响,从而为处理时间序列分类问题提供一种新的途径。

**关键词:**多变量时间序列分类;样本相似性;图卷积网络(GCN);潜在关系;特征聚合

中图分类号:TP181

文献标志码:A

文章编号:1003-5060(2023)12-1642-09

## Multivariate time series classification via potential relationship between samples

TANG Shengtang, WU Gongqing, TAI Changyang, YANG Ze, ZHANG Zan

(School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230601, China)

**Abstract:** The task of multivariate time series (MTS) classification aims to determine the label of MTS samples. MTS data have rich relationship information such as temporal relationship and sample similarity relationship. However, the existing methods fail to make full use of these relationship information, which makes it difficult to improve the classification performance. For this reason, this paper proposes an MTS classification method based on graph convolutional network (GCN), which improves the classification performance by mining the potential relationship between samples. Firstly, in order to effectively represent the sample relationship, rules of building graph based on sample similarity are designed to model the samples, which can map the potential relationship information of samples into a graph space. Then, a classification model based on graph convolution is proposed, which captures the potential sample relationship conducive to classification by aggregating sample features, and updates them to the sample's own feature vector to improve the classification accuracy. Extensive experiments on eleven public datasets show that the proposed method is superior to twelve comparison methods, which shows that the proposed method provides a new approach for dealing with the problem of time series classification. It really has an important influence on the classification results by mining the potential relationship between time series data for classification.

收稿日期:2022-12-07;修回日期:2023-03-27

基金项目:国家自然科学基金资助项目(61906059);国家自然科学基金国际(地区)合作与交流资助项目(62120106008)

作者简介:唐胜唐(1996—),男,安徽巢湖人,合肥工业大学硕士生;

吴共庆(1975—),男,安徽岳西人,博士,合肥工业大学教授,硕士生导师。

**Key words:** multivariate time series classification; sample similarity; graph convolutional network(GCN); potential relationship; feature aggregation

## 0 引言

多变量时间序列(multivariate time series, MTS)是通过监测多个指标(变量)收集的具有时间特性的一系列数据。多变量时间序列分类(multivariate time series classification, MTSC)是将时间序列分类为预定义类别的任务,它是数据挖掘领域中一个重要且具有挑战性的问题,吸引了众多研究者的关注,已广泛应用于医疗保健、运动识别和天气预测<sup>[1]</sup>等诸多领域。

早期的 MTS 分类方法大多是基于距离的方法<sup>[2]</sup>和基于特征的方法<sup>[3]</sup>,它们依赖于从原始 MTS 数据中提取特征并结合传统分类器进行分类,难以直接处理原始 MTS 数据,需要大量的领域专业知识和特征工程对数据进行处理。近年来,基于深度学习的方法<sup>[4]</sup>在时间序列分类中取得了优异的表现。其相比于传统的 MTS 分类方法能够更全面且自主地学习 MTS 数据中隐藏的丰富的有价值信息,取得更好的分类效果,然而在对 MTS 数据的关系利用上依旧存在不足。

MTS 样本间隐藏了丰富的关系信息。以医疗领域为例,利用心电图仪器可以对心脏活动进行实时监测并采样,从而获取一系列具有时序关系并且相互联系的 MTS 数据样本。医生可以根据这些样本的关系以及相关数据的分析结果对病人病情做出判断。然而,现有的 MTS 分类方法通常将 MTS 样本当作一个单独的个体进行处理,未考虑样本间隐藏的关系信息,难以有效提升分类性能。同时这些方法构建的分类模型只使用了标记样本的数据进行训练,未能有效利用未标记的样本,在标记的样本较少的数据集上分类性能不佳。

为了更好地捕获 MTS 数据中的关系信息,一些研究者将 MTS 数据映射到图空间,通过图神经网络(graph neural network, GNN)挖掘数据中潜在的关系<sup>[5-6]</sup>。近年来,图神经网络在许多任务上都取得了优异的表现<sup>[7]</sup>。以图卷积网络(graph convolutional network, GCN)为例,它通过图结构扩展信息,引入可以优化的卷积参数对节点的邻居进行卷积操作,使每个节点都能充分利用其邻接节点的特征信息,不但获得了节点更有效的特征表示,而且能将图中未标记节点的特征信息

充分利用起来,可以有效地处理节点分类任务。

为了充分挖掘 MTS 样本间的潜在关系,本文提出一种基于 GCN 的 MTS 分类框架,通过挖掘样本间的潜在关系和利用未标记数据来提高分类性能。为了对样本关系进行建模,本文设计了一种基于样本相似性的关系映射准则构造样本关系图,将时间序列数据映射到图空间来获取样本间潜在关系的特征表示。在图的构建过程中,图节点包括标记样本和未标记样本,因此,模型可以充分利用未标记样本包含的有价值的信息。为了获得样本关系图中的多阶邻接样本特征信息,提出使用由多个不同的图卷积层组成的 MTS 分类模型,通过图结构聚合和更新其  $l$  阶相邻样本信息学习图的深层结构表示。大量实验结果验证了本文提出的分类模型的有效性。本文的主要贡献概括如下:

1) 提出一种基于图卷积网络的 MTS 分类框架,能够获取 MTS 样本的关系信息,进而将其与样本特征信息进行融合生成语义更丰富的多变量时序数据的样本表示。

2) 设计了一种基于样本相似性的关系映射准则,将 MTS 样本映射到图空间,获取样本间潜在关系的特征表示。

3) 本文在 11 个数据集上进行了大量实验,并与 12 种多变量时间序列分类方法进行比较。实验结果表明,该方法在分类性能方面具有显著优越性。

## 1 相关工作

本文提出一种新的 MTS 分类模型,通过对 MTS 样本进行建模,将样本关系映射到图空间,使用 GCN 获取样本间潜在关系信息的特征表示用于分类。

### 1.1 多变量时间序列分类

MTS 分类方法可分为基于距离的方法、基于特征的方法、基于模型的方法和基于深度学习的方法 4 种。

基于距离的方法采用欧几里德距离<sup>[8]</sup>、短时间序列距离<sup>[9]</sup>、动态时间规整距离(dynamic time wrapping, DTW)<sup>[10-11]</sup>及其各种变体<sup>[12-13]</sup>等利用相似性度量准则计算时间序列之间的相似性,然后根据测试实例与训练实例之间的相似性对测试

实例进行分类。基于特征的方法采用时间序列 Shapelets 模型<sup>[14]</sup>、多变量时间序列的符号表示模型<sup>[15]</sup>和广义多变量 Shapelet 模型<sup>[16]</sup>等将时间序列转换为特征向量,从原始 MTS 数据中提取全局或局部的特征,提供给分类器进行分类。基于模型的方法采用高斯混合模型<sup>[17]</sup>、多变量高斯模型<sup>[18]</sup>和隐马尔可夫模型<sup>[19]</sup>等使用模型参数表示原始时间序列,根据假设模型对数据建模,通过衡量模型之间的相似度进行分类。

近年来,基于深度学习的方法在时间序列分类领域表现优异。文献<sup>[20-22]</sup>引入自动学习特征的理念,通过神经网络中逐层的特征变换,将 MTS 样本在原始空间上的特征表示映射到一个新特征空间,使用这些新特征能够更容易地实现分类的目的。与人工构造规则抽取特征、设计模型的方法相比,基于深度学习的方法可以自动地学习特征,提取数据蕴含的丰富信息以提升分类性能。

上述分类方法仅学习序列的特征信息,没有考虑序列间的复杂关系。鉴于图模型能方便地表示对象之间的关系,本文引入图模型,用于表示序列之间的关系,通过基于图模型的挖掘方法发现序列之间的关系辅助,提升 MTS 分类性能。

## 1.2 图卷积网络

随着图神经网络<sup>[23]</sup>的兴起,基于消息传递<sup>[24]</sup>、信息传播<sup>[25]</sup>和图卷积<sup>[26]</sup>设计的 GNN 模型已被应用于网络分析和自然语言处理等诸多领

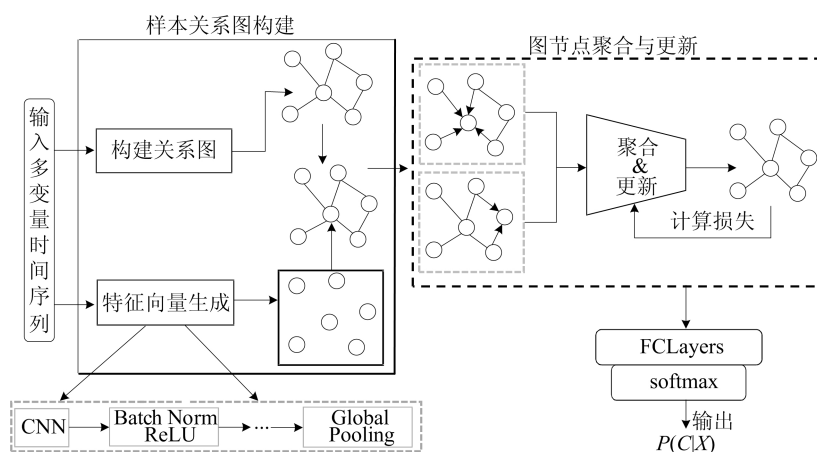
域。图卷积网络<sup>[26]</sup>和图注意网络<sup>[27]</sup>等学习有效的消息传递机制,对节点及其邻居节点进行加权求和,通过多层图卷积,聚合多阶节点之间的信息以表达节点的关系信息,利用最后一层卷积得到的节点特征向量执行分类或预测等任务。

已有研究工作探索将图神经网络应用于 MTS 处理任务。文献<sup>[28]</sup>设计一种新的池化层 MTPool,与 GCN 结合,捕获 MTS 变量之间的隐藏依赖关系和时序信息用于分类。StemGNN<sup>[29]</sup>是实现 MTS 预测的深度学习框架,集成了图傅里叶变换和离散傅里叶变换,以捕获序列间的相关性。MTGNN<sup>[30]</sup>结合了图卷积模块和时间卷积模块,捕获 MTS 变量之间的依赖关系用于 MTS 预测任务。这些基于 GNN 的方法能够有效获取 MTS 变量之间的依赖关系,在 MTS 分类和预测等领域取得了较好的效果。

然而,上述方法没有将 MTS 样本之间的关系用于提升 MTS 分类任务的精度。因此,本文探索将样本的关系信息映射到图空间,对 MTS 样本进行编码,利用 GCN 的关系获取能力和节点分类能力挖掘样本间潜在的关系信息以提升分类性能。

## 2 基于图卷积的 MTSC 模型

本节设计基于图卷积网络挖掘样本间潜在关系的多变量时间序列分类模型(MTSC based on GCN, GMTSC),详细介绍模型的组成以及基于模型的分类算法设计。MTSC 模型如图 1 所示。



### 2.1 GMTSC 模型

直推学习在训练时会同时使用训练集和测试集数据,在训练阶段不使用测试集样本的标签,只在测试阶段使用测试集样本的标签用于性能评

估,适合基于图卷积模型的学习。本文基于直推学习范式设计了一种半监督图卷积模型 GMTSC 用于多变量时序样本分类。由图 1 可知 GMTSC 包含 2 个模块:样本关系图构建模块以及图节点

聚合与更新模块。首先,输入原始时间序列样本数据,基于样本相似度量按照直推学习范式对训练集和测试集中的样本统一建图,将样本间的关系信息映射到图空间;然后,利用卷积神经网络对每个样本提取多个时间戳上的特征,映射为样本节点的特征向量,通过图卷积在图空间上挖掘样本间的潜在关系信息用于分类任务,此外通过迭代优化训练样本的预测值,提升模型对样本关系的获取能力,进而将关系信息有效地编码到样本的特征向量上;最后,通过一个线性分类器,将测试样本编码后的特征向量映射为该测试样本的概率向量以预测类别。

## 2.2 样本关系图构建

使用图卷积挖掘 MTS 样本之间的潜在关系信息需要预定义的图结构。本文使用样本建立图的节点,基于样本相似性的映射准则计算样本之间的相似度量,选择每个样本的前  $k$  个最相似样本作为邻接节点,构建无向边,为整个 MTS 数据集构建样本关系图。

在构建无向边时,由于 DTW 算法能够自动规整时间序列,通过时间轴上的局部缩放,可使 2 个序列的形态接近以计算相似度。因此,本文选择 DTW 算法进行样本的相似度量来寻找邻接节点构建无向边,生成图结构。对于多变量时间序列样本,基于多维变量计算维度相关的 DTW 距离  $D_{TW}$ ,而不是单独面向每个维度进行计算,计算公式为:

$$D_{TW}(t_1, t_2) = \{\text{dis}(t_1, t_2) + \min[D_{TW}(t_1 - 1, t_2), D_{TW}(t_1, t_2 - 1), D_{TW}(t_1 - 1, t_2 - 1)]\}^{1/2} \quad (1)$$

其中:  $\text{dis}(t_1, t_2) = (X_i(t_1) - X_j(t_2))^2$  为 MTS 数据集任意 2 个样本  $X_i$  的  $t_1$  时间戳和  $X_j$  的  $t_2$  时间戳相对应的所有变量观察值的局部距离;  $D_{TW}(t_1, t_2)$  为从  $X_i$  和  $X_j$  的第 1 个时间戳到  $X_i$  的  $t_1$  时间戳和  $X_j$  的  $t_2$  时间戳的最小累加距离。

关系图节点的数据表示形式为一维特征向量。然而,根据样本构建的节点对应的原始样本数据是一个二维矩阵。为了便于 GCN 对图的操作,需要重建节点的数据,将原始样本数据的二维矩阵映射为对应节点的一维特征向量。向量映射准则为:① 保留原始数据之间的特征和时序性;② 映射的特征向量符合图卷积网络的输入要求;③ 每个节点特征向量的维数是统一的。

根据上述向量映射准则,设计堆叠卷积神经网络(convolutional neural network, CCN),将节点的原始二维矩阵数据映射为一维特征向量,以

图中所有节点的原始样本数据作为输入,使用多个具有不同大小卷积核的卷积神经网络层提取样本中多变量和多时间戳下的特征,通过全局池化将采样的特征映射为一维向量,计算公式为:

$$\text{Global Pool}(\text{CNN})(X): \mathbf{R}^{N \times T \times M} \longrightarrow \mathbf{R}^{N \times D} \quad (2)$$

其中:  $X = \{X_1, X_2, \dots, X_N\}$  为原 MTS 始样本数据的集合;  $N$  为数据集中样本数;  $T, M$  分别为每个样本的时间戳数和变量数;  $D$  为映射生成的特征向量的维度。

## 2.3 图节点聚合与更新

上述过程通过对样本关系进行建模生成样本关系图,将关系信息映射到图空间。为了在分类时能充分利用样本间的潜在关系信息,本文基于图卷积模型对样本节点执行聚合和更新操作,将样本在图上多个邻接样本的特征信息结合起来作为样本特征的补足信息,并与样本自身特征融合,将潜在样本关系信息编码到节点特征向量上。通过不断融合节点的邻接样本信息(潜在样本关系信息)来更新节点自身样本特征,直到模型达到稳定均衡。

单个节点聚合与更新操作计算公式为:

$$\mathbf{v}_i' = \sigma\left(\frac{1}{|N_{\text{car}}(X_i)|} \sum_{u \in N_{\text{car}}(X_i)} \mathbf{v}_u \mathbf{W}\right) \quad (3)$$

其中:  $i \in [1, N]$ ,  $N_{\text{car}}(X_i)$  为样本  $X_i$  的邻居样本集合;  $\mathbf{v}_u \in \mathbf{R}^{1 \times d}$  为  $X_i$  的邻居样本的特征向量;  $\mathbf{v}_i' \in \mathbf{R}^{1 \times d'}$  为  $X_i$  经过邻域聚合与更新后的特征向量;  $d$  为样本原始特征长度;  $d'$  为模型预设的输出特征维度;  $\mathbf{W} \in \mathbf{R}^{d \times d'}$  为节点聚合其相邻节点特征后进行线性变换的权重矩阵;  $\sigma$  为激活函数。

对所有节点的聚合与更新操作计算公式为:

$$\mathbf{V}' = \sigma(\tilde{\mathbf{D}}^{-0.5} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-0.5} (\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_N) \mathbf{W}) \quad (4)$$

其中:  $\mathbf{V}' = (\mathbf{V}_1', \mathbf{V}_2', \dots, \mathbf{V}_N') \in \mathbf{R}^{N \times d'}$  为所有样本聚合与更新后得到的特征向量集合;  $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$  为样本关系图的邻接矩阵加上自连接,  $\mathbf{I}$  为单位矩阵;  $\tilde{\mathbf{D}}$  为  $\tilde{\mathbf{A}}$  的度矩阵;  $\tilde{\mathbf{A}}$  为一个没有标准化的矩阵,与特征矩阵相乘会改变特征的原始分布,导致网络训练时间增加或网络无法收敛等情况。为加快网络学习速度,对  $\tilde{\mathbf{A}}$  做了一个标准化处理  $\tilde{\mathbf{A}} = \tilde{\mathbf{D}}^{-0.5} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-0.5}$ 。

GCN 层数表示节点获取邻接节点聚合信息的最远距离。当图稀疏时,单层 GCN 的节点只能获取少量邻居节点的特征信息,难以有效聚合多跳节点间的关系信息,从而影响分类效果。本文通过加深 GCN 层数以获得更大的特征聚合范围,

使每个样本节点可以充分利用其多跳邻居样本的特征,提高模型对样本间潜在关系特征的捕获能力。第  $l$  层 GCN 中每个样本的输出特征  $\mathbf{Z}_i^l$  计算公式为:

$$\mathbf{Z}_i^l = \sigma \left( \frac{1}{|N_{\text{car}}(X_i)|} \sum_{u \in N_{\text{car}}(X_i)} \mathbf{Z}_u^{l-1} \mathbf{W}^l \right) \quad (5)$$

其中:  $\mathbf{W}^l$  为第  $l$  层 GCN 中线性变换的权重矩阵,它的第 2 维度决定了下一层的特征数;  $\mathbf{Z}_u^{l-1}$  为第  $l-1$  层 GCN 的输出特征,并作为第  $l$  层的输入,当  $l=1$  时,  $\mathbf{Z}_i^{l-1} = \mathbf{V}_i$ ,即初始输入是原始 MTS 数据集中的特征向量。

#### 2.4 GMTSC 算法

本节给出基于直推学习范式下的多变量时间序列分类算法 GMTSC,算法描述如下:

输入:时间序列训练集样本和标签、测试集样本和迭代次数 epoch。

输出:测试集样本的预测标签。

- 1) 使用训练集和测试集样本构建样本关系图;
- 2) 将原始 MTS 样本数据映射为图节点的一维特征向量;
- 3) 输入样本关系图到图卷积模块;
- 4) 迭代 epoch 次,每次迭代中在每层图卷积层都对样本节点的特征进行聚合和更新,得到更新后的特征向量;
- 5) 使用算法对训练集样本标签进行预测;
- 6) 根据训练集样本的预测标签与真实标签计算图卷积模块的损失;
- 7) 根据算法的损失优化算法的权重参数,返回步骤 4) 进入下一次迭代;
- 8) 迭代结束,根据训练好的 GMTSC 算法对测试集的标签进行预测;
- 9) 返回测试集的预测标签。

算法步骤 1) ~ 步骤 3) 根据式(1) ~ 式(2)构建样本关系图,输入到图卷积模块;步骤 4) ~ 步骤 7) 是算法的训练过程。首先算法步骤 4) 基于样本关系图,使用多层图卷积捕获图中样本节点的关系信息并编码到节点的特征向量上,然后通过一层全连接层对训练样本的标签进行预测,计算公式为:

$$P(C | X_i) = \text{softmax}(\mathbf{Z}_i^l \mathbf{W}' + b) \quad (6)$$

其中:  $\mathbf{Z}_i^l$  为样本  $X_i$  通过最后一层图卷积的输出特征;  $\mathbf{W}'$  为全连接层进行线性变换的权重矩阵;  $b$  为全连接层的偏置项;  $C$  为经过 softmax 函数计算后的属于每个标签的概率集合。

算法步骤 5) ~ 步骤 7) 每次迭代利用训练样本的预测标签与它们的真实标签对比计算算法的损失,将其作为反馈信息来指导算法中权重参数的更新。交叉熵可以用来计算学习算法分布与训练分布之间的差异,一般情况下可以收敛得到更好的局部极小值点,本文使用交叉熵函数计算算法的损失,计算公式为:

$$L_{\text{oss}} = -\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^k y_{\text{true}_i^{(j)}} \text{lb}(y_{\text{pred}_i^{(j)}}) \quad (7)$$

算法步骤 8) ~ 步骤 9) 是算法的分类过程,在算法训练结束后,通过式(6)对图中测试样本的标签进行预测以完成分类。

### 3 GMTSC 的实验设计及结果分析

本文使用分类精度、平均精度和每种算法能达到最佳精度的数据集数量作为实验的评价指标,迭代次数 epoch 为 100,实验环境为 Intel (R) Core(TM) i7-9800x CPU@3.80 GHz CPU、32 GiB RAM、Windows 10 Pro 64,开发平台为 Python 3.7。

#### 3.1 数据集

本文从 UEA&UCR 时间序列数据集网站中选择了 5 个类别共 11 个公开 MTS 数据集,见表 1 所列。

表 1 数据集所属的类别

数据集所属类别	数据集
人类活动识别	BasicMotions(BM)
	Epilepsy(EP)
	NATOPS(NATO)
运动分类	ArticularyWordRecognition(AWR)
	PenDigits(PD)
心电图分类	AtrialFibrillation(AF)
	StandWalkJump(SWJ)
脑电图 / 脑磁图分类	HandMovementDirection(HMD)
	MotorImagery(MI)
	SelfRegulationSCP2(SR2)
其他	PEMS-SF(PSF)

实验数据集的基本信息见表 2 所列。数据集由不同应用程序收集的实时多变量时间序列数据组成,包括人类活动识别、运动分类、音频频谱分类等,包含了广泛样本、维度和序列长度的真实多变量时序数据。

数据集中的变量数量从轨迹分类数据集集中的 2 维变量到交通流分类任务中的 963 维变量不等,时间序列长度在 8 ~ 3 000 之间,样本数量在 27 ~ 10 992 之间,样本数据的分布是随机的,具

有代表性。

表 2 实验数据集及其基本信息

数据集	训练集数	测试集数	变量数	长度	标签数
AWR	275	300	9	144	25
AF	15	15	2	640	3
BM	40	40	6	100	4
EP	137	138	3	206	4
HMD	160	74	10	400	4
MI	278	100	64	3 000	2
NATO	180	180	24	51	6
PSF	267	173	963	144	7
PD	7 494	3 498	2	8	10
SR2	200	180	7	1 152	2
SWJ	12	15	4	2 500	3

### 3.2 对比算法

本文对比算法选择在 MTSC 领域的 12 种代表性分类算法,包括 9 种监督分类算法和 3 种半监督分类算法。

9 种监督分类算法如下:

1) 1NN-ED<sup>[8]</sup>。使用欧氏距离作为 MTS 之间的距离度量的最近邻分类器。

2) 1NN-DTW-I<sup>[31]</sup>。最近邻分类器中分别计算 MTS 每个维度的 DTW 距离,并将其总和作为分类依据。

3) 1NN-DTW-D<sup>[31]</sup>。将每个时间戳对应的多个变量的观测值视为一个点,在最近邻分类器计算距离时,对齐过程与传统的 DTW 算法相同。

4) WEASEL-MUSE<sup>[32]</sup>。使用多个窗口长度来构建一个大的特征空间,然后使用卡方检验来确定

最相关的特征并将其输入到逻辑回归模型中。

5) MLSTM-FCN<sup>[4]</sup>。将全卷积网络模型与 LSTM 进行结合,同时向全卷积模块中添加了压缩模块和激励模块。

6) Tapnet<sup>[20]</sup>。设计了一个将时间序列的维度重构为组的随机组置换方法,从多变量时间序列数据中学习低维特征,利用标记样本和未标记样本设计了一个注意原型网络,有效解决了缺少标记样本训练模型的问题。

7) 1NN-ED(norm)。1NN-ED 在分类前对输入数据进行的归一化操作。

8) 1NN-DTW-I(norm)。1NN-DTW-I 在分类前对输入数据进行的归一化操作。

9) 1NN-DTW-D(norm)。1NN-DTW-D 在分类前对输入数据进行的归一化操作。

3 种半监督分类算法如下:

1) 1NN-DTW-D<sup>[33]</sup>。使用基于距离的最近邻分类器构建时间序列半监督学习算法。

2) MTL<sup>[34]</sup>。利用已建立的多任务学习算法,将模型预测作为次要任务与主分类任务一起优化。

3) Semi-Tapnet<sup>[20]</sup>。Tapnet 的注意原型网络利用未标记数据将模型扩展到半监督模式,应用于 MTS 半监督分类。

### 3.3 实验结果与分析

#### 3.3.1 监督分类算法的实验结果比较

GMTSC 与 9 种监督分类算法的分类精度、平均精度和能达到最佳精度的数据集数量(获胜数)见表 3 所列。

表 3 GMTSC 和 9 种 MTS 分类算法在 11 个数据集上的分类精度

数据集	GMTSC	Tapnet	MLSTM-FCN	WEASEL-MUSE	1NN-ED	1NN-DTW-I	1NN-DTW-D	1NN-ED (norm)	1NN-DTW-I (norm)	1NN-DTW-D (norm)
AWR	0.993	0.987	0.973	0.990	0.970	0.980	0.987	0.970	0.98	0.987
AF	0.533	0.333	0.267	0.333	0.267	0.267	0.200	0.267	0.267	0.220
BM	1.000	1.000	0.950	1.000	0.675	1.000	0.975	0.676	1.000	0.975
EP	0.957	0.971	0.761	1.000	0.667	0.978	0.964	0.667	0.978	0.964
HMD	0.527	0.378	0.365	0.365	0.279	0.306	0.231	0.278	0.306	0.231
MI	0.630	0.590	0.510	0.500	0.510	0.390	0.500	0.510	N/A	0.500
NATO	0.950	0.939	0.889	0.870	0.860	0.850	0.883	0.850	0.850	0.883
PSF	0.728	0.751	0.699	N/A	0.705	0.734	0.711	0.705	0.734	0.711
PD	0.985	0.980	0.978	0.948	0.973	0.939	0.977	0.973	0.939	0.977
SR2	0.617	0.550	0.472	0.460	0.483	0.533	0.539	0.483	0.533	0.539
SWJ	0.600	0.400	0.067	0.333	0.200	0.333	0.200	0.200	0.333	0.200
平均精度	0.775	0.716	0.630	0.680	0.590	0.665	0.652	0.598	0.692	0.653
获胜数	8	2	0	2	0	1	0	0	1	0

从表 3 可以看出,GMTSC 不仅在 8 个数据集上取得了最佳分类精度,而且平均分类精度也达到了最佳(0.775),远远优于分类精度次佳的 Tapnet(0.716)。这说明 GMTSC 通过将样本关系映射到图空间,利用图卷积挖掘样本间的潜在关系信息用于分类,为样本提供了更多的可区分特征。相比之下,几种传统的基于距离的方法平均分类精度只在 0.650 左右,这是由于它们只简单利用了与测试样本最相似的训练集样本进行分类,没有深入考虑样本间潜在的关系信息。

基于特征的分类方法和深度学习的分类方法虽然在 EP 和 PSF 等数据集上也获得了最佳的分类精度,但在 11 个数据集上的平均精度也只在 0.68 左右,远远不及 GMTSC。这是由于它们在处理每个 MTS 样本时,仅将其作为独立个体进行特征提取,很少考虑使用其他相似的样本特征信息来丰富自身特征表示。

此外,在 AF 和 SWJ 只有少量标记样本的数据集上, GMTSC 也取得了最佳分类精度,说明 GMTSC 充分利用了未标记样本包含的有价值的信息,有效提升了分类性能。

本文还给出了 GMTSC 与 9 种对比算法分类性能差异的临界差值图,如图 2 所示,具有最佳排名的分类算法位于图中右侧。从图 2 可以看出 GMTSC 的平均排名最高。根据 Friedman test,若算法性能相同,则它们的平均排名也相等,图 2 表明 GMTSC 以及其他 9 种对比方法的性能都不相同。使用 Nemenyi test 做进一步区分,设定显著性水平  $\alpha = 0.05$ ,计算出平均排名差别的临界值域(critical difference, CD),若 2 个算法的平均排名之差超出了该阈值说明 2 个算法性能有差异,图 2 表明 GMTSC 在性能上显著优于大部分方法。

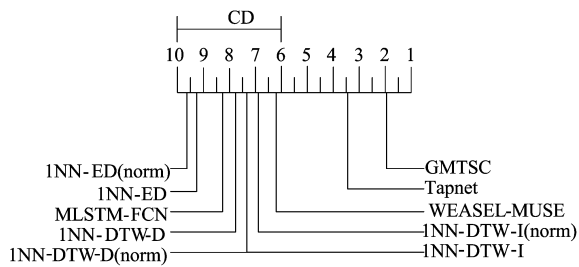


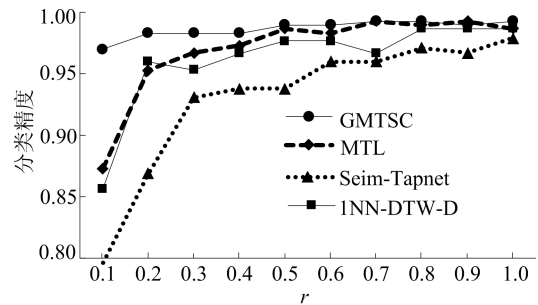
图 2 GMTSC 与 9 种对比算法分类性能差异的邻接差值图

### 3.3.2 半监督分类算法的实验结果比较

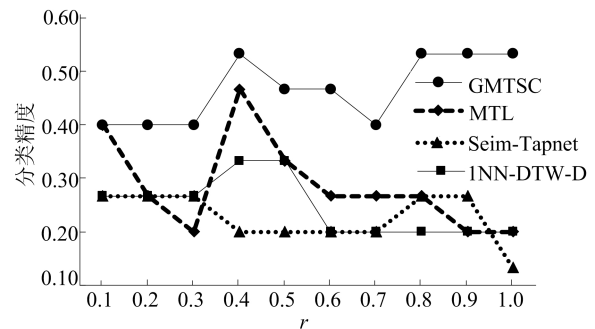
本文选择了 4 个属于不同领域的数据集来评估 GMTSC 在不同的监督水平下的分类效果,并

与 1NN-DTW-D、Semi-Tapnet 和 MTL 这 3 种代表性半监督分类算法进行对比。在划分数据集时,按照 0.1 ~ 1.0 递增比例  $r$ ,选取训练集中每个类别对应比例的样本数标记样本进行训练,剩下的样本作为无标记样本,保证选取的训练集样本包含所有的类别。

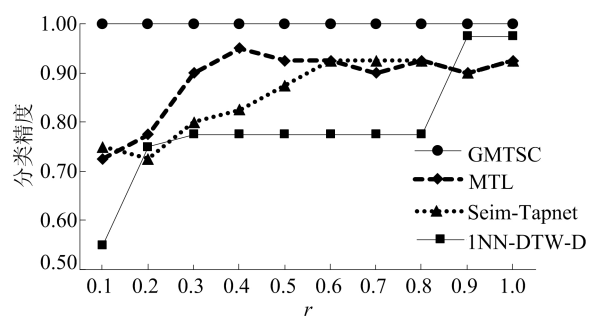
GMTSC 与其他 3 种算法在这 4 个数据集下的分类精度如图 3 所示。



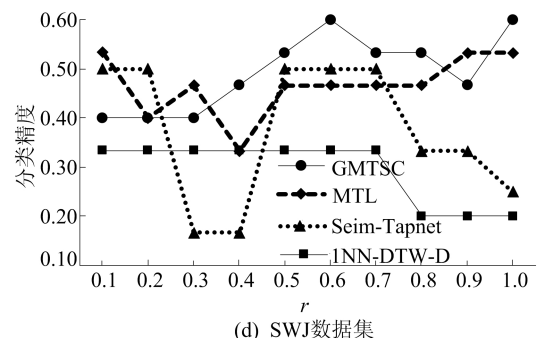
(a) AWR 数据集



(b) AF 数据集



(c) BM 数据集



(d) SWJ 数据集

图 3 GMTSC 与 3 种半监督分类算法的实验结果

从图 3 可以看出,GMTSC 在分类精度上明显优于对比算法,尤其在 AWR 数据集上,GMTSC 在只有 0.1 带标记样本集上训练时分类精度能达到 0.983,明显优于 Semi-Tapnet(0.790)、MTL (0.835) 和 1NN-DTW-D(0.400)。

由图 3 可知,随着划分比例的增加,GMTSC 的分类性能稳定提升,而 Semi-Tapnet、MTL 和 1NN-DTW-D 的分类性能整体变化趋势不稳定,尤其是基于距离的半监督算法 1NN-DTW-D 极易受到单 MTS 样本(噪声数据)的影响。

上述结果表明,即使只有较少的标记样本,GMTSC 通过对样本间关系信息的挖掘以及利用未标记样本,在 MTS 半监督分类任务中具有较好的分类性能。

#### 4 结 论

本文提出了 GMTSC 算法并用于 MTS 分类,设计了一种挖掘时间序列样本间潜在关系的图卷积模型,该模型能捕获映射到图空间的样本间关系信息用于 MTS 分类,通过对样本节点执行聚合和更新操作,将潜在样本关系信息编码到节点特征向量更新节点自身特征以提高分类精度。在 11 个数据集上与 12 种代表性算法进行对比,实验结果表明 GMTSC 在 MTS 分类任务上具有优越的性能。未来可进一步从以下 2 个方面开展研究:

1) 在样本关系图构建方面,探索设计一种新的构图规则,解决基于 DTW 算法构建样本关系图时间复杂度过高的问题,并能够很好地衡量时间序列之间的距离以处理时间序列不等长的问题,从而提高建图质量。

2) 在模型结构设计方面,改进图卷积网络,在对节点进行聚合和更新时,将与节点相连的边的特征信息附加到节点的特征向量上一起计算,以捕获更充分的样本间关系信息,以期进一步提高 MTS 分类性能。

#### [参 考 文 献]

- [1] MCGOVERN A,ROSENDAHL D H,BROWN R A,et al. Identifying predictive multi-dimensional time series motifs: an application to severe weather prediction[J]. *Data Mining and Knowledge Discovery*,2011,22(1):232-258.
- [2] SETO S,ZHANG W,ZHOU Y. Multivariate time series classification using dynamic time warping template selection for human activity recognition[C]//*Proceedings of the 2015 IEEE Symposium Series on Computational Intelligence*. [S. l. ]:IEEE,2015:1399-1406.
- [3] WU J,YAO L,LIU B. An overview on feature-based classification algorithms for multivariate time series[C]//*Proceedings of the 3rd International Conference on Cloud Computing and Big Data Analysis*. [S. l. ;s. n. ],2018:32-38.
- [4] KARIM F,MAJUMDAR S,DARABI H,et al. Multivariate LSTM-FCNs for time series classification[J]. *Neural Networks*,2019,116:237-245.
- [5] DENG A,HOOI B. Graph neural network-based anomaly detection in multivariate time series[C]//*Proceedings of the 2021 AAAI Conference on Artificial Intelligence*. [S. l. ]:AAAI Press,2021:4027-4035.
- [6] CHEN Y,SEGOVIA I,GEL Y R. Z-GCNETS:time zigzags at graph convolutional networks for time series forecasting[C]//*Proceedings of the 38th International Conference on Machine Learning*. [S. l. ;s. n. ],2021:1684-1694.
- [7] ZHOU J,CUI G,HU S,et al. Graph neural networks:a review of methods and applications[J]. *AI Open*,2020,1:57-81.
- [8] KIM S W,PARK D H,LEE H G. Efficient processing of subsequence matching with the Euclidean metric in time-series databases [J]. *Information Processing Letters*,2004,90(5):253-260.
- [9] MILLER-LEVET C S,KLAWONN F,CHO K H,et al. Fuzzy clustering of short time-series and unevenly distributed sampling points[C]//*Proceedings of the 2003 International Symposium on Intelligent Data Analysis*. [S. l. ;s. n. ],2003:330-340.
- [10] BAGNALL A,LINES J,BOSTROM A,et al. The great time series classification bake off:a review and experimental evaluation of recent algorithmic advances[J]. *Data Mining and Knowledge Discovery*,2017,31(3):606-660.
- [11] RUIZ A P,FLYNN M,LARGE J,et al. The great multivariate time series classification bake off:a review and experimental evaluation of recent algorithmic advances[J]. *Data Mining and Knowledge Discovery*,2021,35(2):401-449.
- [12] GRECKI T,LUCZAK M. Non-isometric transforms in time series classification using DTW[J]. *Knowledge-based Systems*,2014,61(1):98-108.
- [13] SHARABIANI A,DARABI H,REZAEI A,et al. Efficient classification of long time series by 3-D dynamic time warping[J]. *IEEE Transactions on Systems,Man, and Cybernetics:Systems*,2017,47(10):2688-2703.
- [14] YE L,KEOGH E J. Time series shapelets:a new primitive for data mining[C]//*Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. [S. l. ]:ACM,2009:947-956.
- [15] BAYDOGAN M G,RUNGER G C. Learning a symbolic representation for multivariate time series classification [J]. *Data Mining and Knowledge Discovery*,2015,29(2):400-422.
- [16] MEDICO R,RUYSSINCK J,DESCRIJVER D,et al. Learning multivariate shapelets with multi-layer neural networks for interpretable time-series classification[J].

- Advances in Data Analysis and Classification, 2021, 15(4): 911-936.
- [17] CAO H, TAN V Y F, PANG J Z F. A parsimonious mixture of gaussian trees model for oversampling in imbalanced and multimodal time-series classification[J]. IEEE Transactions on Neural Networks and Learning Systems, 2014, 25(12): 2226-2239.
- [18] TAI C, YANG Z, ZHANG H, et al. Gaussian model-based fully convolutional networks for multivariate time series classification[C]//Proceedings of the 2021 IEEE International Conference on Big Knowledge (ICBK). [S. l.]; IEEE, 2021: 141-148.
- [19] GUO H, PEDRYCZ W, LIU X. Hidden Markov models based approaches to long-term prediction for granular time series[J]. IEEE Transactions on Fuzzy Systems, 2018, 26(5): 2807-2817.
- [20] ZHANG X, GAO Y, LIN J, et al. Tapnet: multivariate time series classification with attentional prototypical network[C]//Proceedings of the 2020 AAAI Conference on Artificial Intelligence. [S. l.]; AAAI Press, 2020: 6845-6852.
- [21] LI G, CHOI B, XU J, et al. Shapenet: a shapelet-neural network approach for multivariate time series classification[C]//Proceedings of the 2021 AAAI Press Conference on Artificial Intelligence. [S. l.]; AAAI Press, 2021: 8375-8383.
- [22] ZERVEAS G, JAYARAMAN S, PATEL D, et al. A transformer-based framework for multivariate time series representation learning[C]//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. [S. l.]; ACM, 2021: 2114-2124.
- [23] SCARSELLI F, GORI M, TSOI A C, et al. The graph neural network model [J]. IEEE Transactions on Neural Networks, 2009, 20(1): 61-80.
- [24] GILMER J, SCHOENHOLZ S S, RILEY P F, et al. Neural message passing for quantum chemistry[C]//Proceedings of the 2017 International Conference on Machine Learning. [S. l.]; s. n., 2017: 1263-1272.
- [25] KLICPERA J, BOJCHEVSKI A, GNNEMANN S. Predict then propagate: Graph neural networks meet personalized pagerank[C]//Proceedings of the 7th International Conference on Learning Representations. [S. l.]; s. n., 2019: 1-15.
- [26] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[C]//Proceedings of the 5th International Conference on Learning Representations. [S. l.]; s. n., 2017: 1-14.
- [27] VELICKOVIC P, CUCURULL G, CASANOVA A, et al. Graph Attention Networks[C]//Proceedings of the 6th International Conference on Learning Representations. [S. l.]; s. n., 2018: 1-12.
- [28] DUAN Z, XU H, WANG Y, et al. Multivariate time-series classification with hierarchical variational graph pooling [J]. Neural Networks, 2022, 154: 481-490.
- [29] CAO D, WANG Y, DUAN J, et al. Spectral temporal graph neural network for multivariate time-series forecasting[J]. Advances in Neural Information Processing Systems, 2020, 33(1): 17766-17778.
- [30] WU Z, PAN S, LONG G, et al. Connecting the dots: multivariate time series forecasting with graph neural networks[C]//Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. [S. l.]; ACM, 2020: 753-763.
- [31] SHOKOOHI-YEKTA M, HU B, JIN H, et al. Generalizing DTW to the multi-dimensional case requires an adaptive approach [J]. Data Mining and Knowledge Discovery, 2017, 31(1): 1-31.
- [32] SCHFER P, LESER U. Fast and accurate time series classification with WEASEL[C]//Proceedings of the 2017 ACM on Conference on Information and Knowledge Management. [S. l.]; ACM, 2017: 637-646.
- [33] CHEN Y, HU B, KEOGH E, et al. DTW-D: time series semi-supervised learning from a single example[C]//Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. [S. l.]; ACM, 2013: 383-391.
- [34] JAWED S, GRABOCKA J, SCHMIDT-THIEME L. Self-supervised learning for semi-supervised time series classification[C]//Proceedings of the 2020 Pacific-Asia Conference on Knowledge Discovery and Data Mining. [S. l.]; s. n., 2020: 499-511.

(责任编辑 李 凯)