中图法分类号:TP183;TP391　　**文献标识码**:A　　**文章编号**:1006-8961(2026)11-3837-26

论文引用格式:Li S C, Sun J and Sun Z L. 2026. Application of graph neural networks in multivariate time-series classification: an overview. Journal of Image and Graphics, 31(11):3837-3862(李贯宸, 孙杰, 孙志林. 2026. 图神经网络在多变量时间序列分类中的应用综述. 中国图象图形学报, 31(11):3837-3862)[DOI:10.11834/jig.250529]

图神经网络在多变量时间序列分类中的应用综述

李贯宸[1,2],孙杰[1,2*],孙志林[3]

1. 南京邮电大学计算机学院/软件工程学院/网络空间安全学院,南京　210023; 2. 南京晓庄学院信息工程学院,南京　211171; 3. 南京杰辉智能科技有限公司,南京　210019

摘　要:多变量时间序列(multivariate time series,MTS)分类是数据科学中的一项关键任务,旨在从具有复杂时空依赖性的多维数据流中识别模式。图神经网络(graph neural network,GNN)凭借其强大的结构化数据建模能力为这一领域带来了范式转变。然而,现有研究呈现出方法多样化但缺乏统一理论指导的局面。本文系统梳理了GNN在MTS分类中的应用进展,并首次提出从"时空依赖耦合范式"(spatio-temporal dependency coupling paradigm)角度审视该领域的统一分析框架,将现有模型划分为解耦式(decoupled)、耦合式(coupled)和演化式(evolutionary)三大类别,揭示了不同设计在模型灵活性、计算效率和动态适应性之间的核心权衡。基于此框架,本文超越了对模型机制的简单描述,深入剖析了各类方法在处理现实世界数据中普遍存在的异质性(heterogeneity)、非平稳性(non-stationarity)和因果混淆(causal confounding)等根本性挑战时的内在假设与理论局限。此外,通过对交通、医疗、金融和工业等关键应用领域的跨域综合分析,提炼出驱动模型选择和图构建策略的普适性原则。最后,本文总结了现有挑战,并进一步勾勒出一份旨在构建更鲁棒、可解释和支持干预的下一代时空图智能模型的挑战性研究议程,强调了从关联建模向因果推理演进的必要性。研究表明,GNN在MTS分类中展现的巨大潜力正推动该领域迈向一个更深刻、更具挑战性的新阶段。为了促进相关研究的复现与发展,本文汇总了文中涉及的主流算法代码、数据集及评估指标,并在https://github.com/sunjieseu/GNN-MTSC-Survey上开源维护。

关键词:多变量时间序列分类(MTSC);图神经网络(GNN);时空数据挖掘;图表示学习;深度学习

Application of graph neural networks in multivariate time-series classification: an overview

Li Shichen[1,2], Sun Jie[1,2*], Sun Zhilin[3]

1. School of Computer Science/School of Software Engineering/School of Cyberspace Security, Nanjing University of Posts and Telecommunications, Nanjing 210023, China; 2. School of Information Engineering, Nanjing Xiaozhuang University, Nanjing 211171, China; 3. Nanjing Jiehui Intelligent Technology Co., Ltd., Nanjing 210019, China

Abstract: Multivariate time-series (MTS) classification is a foundational problem in modern data science and underpins critical applications ranging from intelligent transportation and industrial IoT monitoring to financial surveillance and medical diagnosis. Unlike univariate time series, MTS data exhibit two intertwined forms of dependency: temporal dynamics

收稿日期:2025-11-18;**修回日期**:2026-02-20;**预印本日期**:2026-02-27

*通信作者:孙杰　jie.sun@njxzc.edu.cn

基金项目:国家自然科学基金项目(62177028);江苏省教育科学"十四五"重点课题项目(B-b/2024/01/145,B-b/2024/01/169);江苏省高等教育学会重点课题项目(2024CXJG029)

Supported by: National Natural Science Foundation of China (62177028); Key Project of the 14th Five-Year Plan for Educational Science in Jiangsu Province (B-b/2024/01/145, B-b/2024/01/169); Key Project of Jiangsu Higher Education Society (2024CXJG029)

within each variable and complex spatial correlations among variables. Traditional approaches, such as dynamic time warping (DTW) with k-nearest neighbors or statistical feature-extraction methods, often suffer from poor scalability and strong dependence on manual feature engineering. Although deep-learning models such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs) automate feature extraction effectively, they are fundamentally designed for Euclidean data and therefore struggle to capture the complex non-Euclidean topological relationships—such as sensor networks, biological interactions, or stock-market linkages—that characterize many real-world systems. In recent years, graph neural networks (GNNs) have emerged as a transformative framework for structured data, representing variables as nodes and their interdependencies as edges. Through message passing, GNNs can explicitly model geometric structure and thereby capture spatial dependencies that conventional models often miss. However, despite the rapid growth of GNN-based methods for MTS tasks, the literature remains fragmented. Most existing surveys focus primarily on forecasting or imputation, with far less attention given specifically to classification. Moreover, the field lacks a unified theoretical framework for organizing the diverse strategies used to model coupled spatial and temporal features. To address this gap, this paper proposes a unified analytical framework, termed the "spatio-temporal dependency coupling paradigm". Under this framework, existing models are categorized into three philosophical approaches. 1) The decoupled paradigm: this divide-and-conquer strategy treats spatial and temporal modeling as separate stages. Typically, a GNN extracts spatial features at each time step, and those features are then passed to a temporal model such as an LSTM or TCN. Although this approach offers strong interpretability and modularity, it often fails to capture the intricate simultaneous interactions between space and time. 2) The coupled paradigm: represented by models such as STGCN and DCRNN, this paradigm treats spatial and temporal dependencies as inseparable. By integrating graph convolutions directly with temporal convolutions or recurrent units, these models capture local spatio-temporal correlations more effectively, though often at the cost of greater parameter complexity. 3) The evolutionary paradigm: this frontier direction addresses the non-stationarity of real-world systems. Unlike early models that assume a static graph structure, methods such as EvolveGCN and TodyNet explicitly model the dynamic evolution of the graph itself, reflecting the fact that relationships among variables can change over time—for example, when traffic patterns shift after an accident or stock correlations reverse during a financial crisis. A critical determinant of GNN performance in MTS classification is how the graph structure is defined, especially when no explicit topology is available. We systematically review four graph-construction strategies: 1) structure-driven methods, which rely on prior knowledge such as physical road networks in traffic tasks; 2) similarity-driven methods, which define edges using statistical criteria such as Pearson correlation or DTW; 3) attention-driven methods, which use mechanisms such as graph attention networks (GATs) to learn edge weights dynamically; and 4) latent graph learning, which infers hidden dependency structures end to end without manual intervention, as in MTSGNN. This survey goes beyond taxonomy to provide a comprehensive analysis of applications in four major domains: transportation, healthcare, industry, and finance. We synthesize general principles for model selection. For example, in transportation and industry, where physical sensors determine topology, structure-driven static graphs are often robust and efficient. By contrast, in finance and healthcare, where relationships are abstract, invisible, and highly non-stationary—such as information flow among stocks or physiological interactions in ICU monitoring—dynamic and latent graph-learning strategies are often essential. We also highlight representative case studies, including the use of the concurrency hypothesis in traffic analysis and knowledge-empowered dynamic graphs (KEDGN) in medical diagnosis, to illustrate how domain knowledge can improve GNN performance. Despite rapid progress, several open challenges continue to hinder the wider adoption of GNNs for MTS classification. 1) Scalability: the computational cost of graph construction remains a bottleneck for large-scale systems such as massive IoT networks. 2) Heterogeneity and non-stationarity: real-world data frequently involve irregular sampling and distribution shifts, or concept drift, which static models cannot handle well. 3) Interpretability versus performance: although GNNs offer richer topological insight than CNNs, deep GNNs and current explanation methods still struggle to provide causal interpretations of their predictions. Looking ahead, we propose a research agenda focused on causal spatio-temporal representation learning. Future models should move beyond capturing correlations and toward identifying causal structure, thereby improving robustness to spurious dependencies. We also discuss the promise of integrating large language models (LLMs) to infer semantic relationships among variables, self-supervised learning to address label scarcity, and privacy-

preserving federated graph learning for sensitive medical and financial data. This paper provides a holistic overview of the GNN-based MTS classification landscape, bridging the gap between theoretical modeling and practical application. By clarifying the trade-offs among model flexibility, computational efficiency, and dynamic adaptability, we aim to guide researchers in the selection and design of next-generation intelligent systems. To facilitate further research and reproducibility, the algorithms, datasets, and evaluation metrics discussed in this paper have been summarized and are available at https://github.com/sunjieseu/GNN-MTSC-Survey.

Key words: multivariate time series classification (MTSC); graph neural network (GNN); spatio-temporal data mining; graph representation learning; deep learning

0 引　言

多变量时间序列(multivariate time series, MTS)在金融市场监控、智能交通调度、环境气象预测、工业制造监测及医疗健康等领域具有广泛的应用(Wei, 2019)。例如,交通系统中多个传感器同时采集流量与车速数据,构成高度相关的时序数据流;金融领域中,多只股票的价格、成交量等共同演化构成典型的MTS数据。在这些领域中,数据往往由多个相互关联的变量构成,这些变量随时间演变且相互影响。对于多变量时间序列数据的挖掘主要包括3个方面的工作:分类(Ismail Fawaz等, 2019)、预测(Lim和Zohren, 2021)和异常检测(Blázquez-García等, 2022)。本文着重讨论多变量时间序列分类方面的相关研究进展。

多变量时间序列分类(multivariate time series classification, MTSC)(Liu等, 2019)作为MTS分析中的核心任务,其目标是:从有限长度的多变量时间序列样本中学习特征表示,从而将每条序列映射到一个预定义的类别标签。这一任务在众多领域中具有重要的实际意义,如交通系统(Gupta等, 2020)、医疗健康监测(Sun等, 2024);工业物联网与制造(Selim等, 2021)、金融市场(Sun等, 2026a)以及环境科学(Sun等, 2026b)等。

与单变量时间序列不同,MTS不仅表现出强烈的时序依赖性,同时存在复杂的变量间交互关系。这使得分类任务面临两方面挑战:1)如何有效捕捉时间维度上的短期与长期依赖关系;2)如何建模变量之间的非线性、多尺度交互关系。此外,实际数据往往存在噪声、采样不均和缺失值等问题,进一步加大了建模难度。

MTS分类最早依赖基于相似度度量的方法,如动态时间规整(dynamic time warping, DTW)(Berndt和Clifford, 1994)结合K近邻(K-NN)分类器,其对局部形变较为敏感,但计算复杂度高且难以扩展。随后,统计学习方法如Shapelets(Ye和Keogh, 2009)、Bag-of-Features(Baydogan等, 2013)等通过提取统计特征向量实现分类,但这些方法严重依赖人工特征设计与先验知识、合理特征选择策略和特征工程过程的计算效率。

随着深度学习方法,如卷积神经网络(convolutional neural network, CNN)(Gu等, 2018)和长短期记忆网络(long short-term memory network, LSTM)(Hochreiter和Schmidhuber, 1997)的出现,MTS分类迈向了机器学习和深度学习技术并取得显著成果。这些模型通过端到端训练方式自动提取时序特征,提高了特征表达能力。然而,它们通常基于欧几里得结构假设,难以捕捉变量之间的复杂拓扑依赖关系。

图神经网络(graph neural network, GNN)(Zhou等, 2020)近年来在社交网络(Fan等, 2019)、推荐系统(Gao等, 2022)和知识图谱(Ye等, 2022)等领域取得了广泛成功。Foumani等人(2024)综述了深度学习在时间序列分类中的最新进展,而Jin等人(2024b)与Corradini等人(2026)则分别从通用时序分析和时空建模视角,系统回顾了GNN在包括分类在内的多项任务中的应用,这些工作充分展示了该领域的蓬勃发展(孙杰 等, 2026)。相比于传统多变量时间序列分类方法基于欧几里得结构假设,通常难以捕捉变量之间的复杂拓扑依赖关系的缺陷,图神经网络作为近年来兴起的一类结构化数据建模方法,为MTS分类带来了新的突破。其核心思想是将多变量视为图中的节点,变量间的关系通过边进行建模,从而捕捉复杂的交互结构。通过消息传递机制,GNN可在保持拓扑结构信息的同时,实现高效

的时空特征提取。

尤其是最近热门的时空图神经网络(spatio-temporal GNN, ST-GNN)(Jin等, 2024a)框架,摆脱了传统的GNN模型在时间维度的不足,通过引入时间维度,融合图卷积与时间卷积(或循环)模块,能够同步建模变量间的空间依赖与时间维度的动态演化,显著提高了分类性能,解决了多变量时间序列分类任务中时序信息与空间关系无法同时捕捉的问题。已成功应用于交通预测(Bui等, 2022)、风功率预测(Khodayar和Wang, 2019)等任务。尽管已有部分综述涉及了时间序列分析或图神经网络,但它们在侧重点上与本文存在显著差异。例如,Ismail Fawaz等人(2019)主要回顾了通用深度学习模型在时间序列分类中的应用,未深入探讨GNN的图结构建模优势;Jin等人(2024b)和孙杰等人(2026)虽然系统总结了GNN在时空数据预测等任务中的进展,但其多涵盖预测、插补等广泛任务,且分类体系多基于模型架构(如循环神经网络(recurrent neural network, RNN)或CNN)。相比之下,本文聚焦多变量时间序列分类(MTSC)这一特定任务,并独创性地提出从"时空依赖耦合范式"角度将现有模型划分为解耦式、耦合式和演化式,填补了现有文献在MTSC细分领域缺乏统一理论分析框架的空白。为了更直观地展示本文与现有代表性综述的区别与联系,从任务范围、是否聚焦MTSC、图构建策略讨论、分类体系依据及应用挑战等维度进行了详细对比,如表1所示,本文是首篇专门聚焦GNN在MTSC任务中的应用,并从时空依赖耦合视角建立系统性分类的综述。

表1 本文与现有代表性时间序列综述的对比

Table 1 Comparison between this paper and existing representative time series reviews

代表性综述	核心任务范围	是否针对MTSC	是否讨论图构建	分类体系依据	跨领域案例与挑战
图结构(Ismail Fawaz等, 2019)	通用深度学习/分类	否*	否	模型架构(CNN、RNN等)	是(通用)
多尺度动态图(Foumani等, 2024)	深度学习/分类与回归	否	否	网络结构与训练机制	是
基于预测的图学习(Jin等, 2024b)	GNN/全时序任务(预测、插补等)	否	是(通用)	下游任务类型	是
时空图网络(Fang等, 2021)	时空GNN/预测与分类	否	是	模型组件类型	是
本文	GNN/多变量时间序列分类	是	是**	时空依赖耦合范式	是(聚焦MTSC)

注:*表示含单变量;**表示针对MTSC优化。

本文旨在系统梳理图神经网络在多变量时间序列分类中的研究进展,聚焦其核心建模机制、图构建策略、代表性模型架构以及应用案例。主要贡献如下:1)系统回顾传统方法与深度模型在MTS分类中的发展历程,指出其关键局限;2)梳理GNN基础知识并以不同视角提出其在MTSC中的建模范式,特别是静态图、动态图与时空图的构建方法;3)对比并分析代表性GNN模型的设计思路及性能表现;4)总结GNN在多个典型领域(如医疗、交通、金融)中的应用案例与评估方式;5)讨论当前GNN在MTSC任务中的挑战与未来研究方向,如动态图建模、自监督学习和可解释性等。

为了直观展示本文的逻辑结构与核心内容,在图1中给出了整体框架,涵盖了从图构建策略、建模范式到关键应用领域的全景视角。该框架以图构建策略为基础,通过3种时空依赖耦合范式进行特征建模,并服务于四大关键应用领域。

1 预备知识

本节从多变量时间序列、多变量时间序列分类、图神经网络基础几个方面介绍相关理论知识。

1.1 多变量时间序列的定义与特性

多变量时间序列(MTS)数据通常由多个传感器或数据源在不同时间点记录的多个变量组成。通常由多个变量组成的时间序列数据,每个变量在多个时间步长上都有观测值。与单变量时间序列(univariate time series, UTS)相比,MTS在建模时更加复

杂，因为每个变量不仅具有时间依赖性，还可能与其他变量之间存在复杂的空间依赖关系。

在 MTS 数据中，每个时间点的观测可以表示为一个向量，这些观测向量随着时间的推移而变化。具体而言，假设有一个包含 N 个变量和 T 个时间步的多变量时间序列数据集 $\boldsymbol{X} \in \mathbf{R}^{N \times T}$，其中 $\boldsymbol{X}_i = \left\{x_{i,1}, x_{i,2} \cdots, x_{i,T}\right\}$ 表示第 i 个变量在 T 个时间步的观测值 $(i = 1, 2, \cdots, N)$。每个变量的观测值是一个一维的时间序列，每个时刻的值 $x_{i,t}$ 表示在 t 时刻、第 i 个变量的观测结果。例如，在智能交通系统中，多个交通传感器每分钟采集的交通流量、车速等数据可以构成多变量时间序列数据（Qian 等，2025）。为了有效地处理这种数据，需要同时捕捉到时间步之间的依赖性以及不同变量之间的关系。

MTS 的主要特性包括高维性、非线性与动态性、时序依赖以及空间依赖性。高维性是指每个时间点可能包含多个变量的观测值，形成一个高维的特征向量；非线性与动态性是指各变量之间通常存在非线性关系，且随时间动态变化；时序依赖是指每个变量随时间演化，表现出短期与长期的时序依赖性；空间依赖性是指不同变量之间的关系可以通过空间拓扑结构表示，尤其在传感器网络或地理分布的情况下，变量间存在地理或物理上的联系（Yuan 和 Li，2021；Liu 等，2024）。

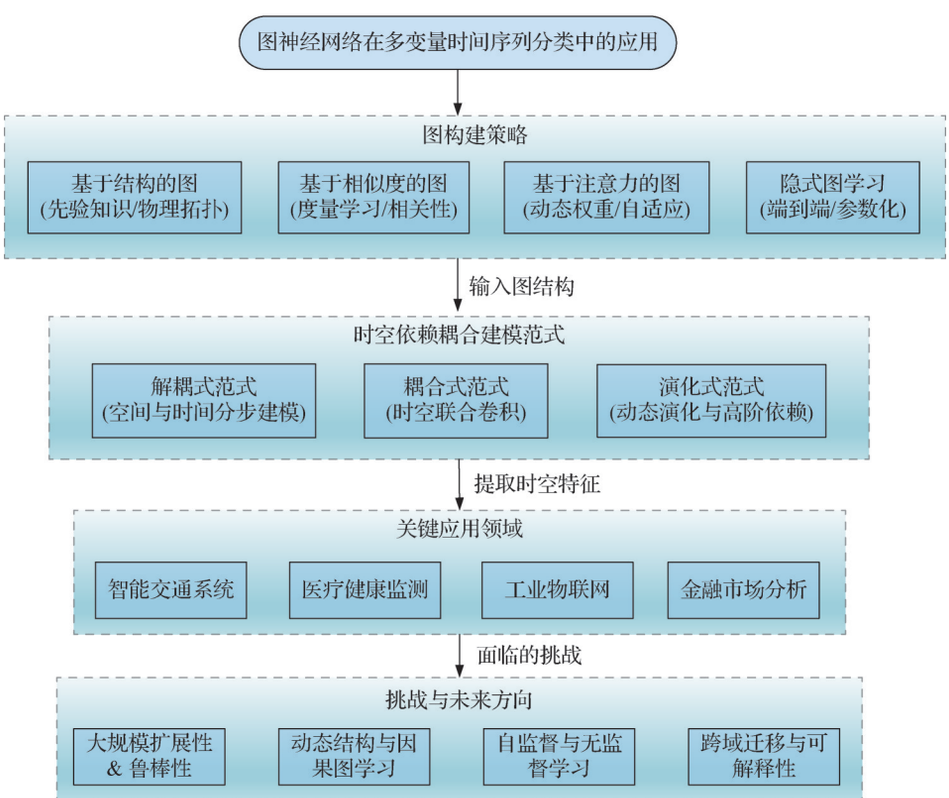


图 1　本文提出的 GNN 在 MTSC 任务中的统一分析框架

Fig. 1　The unified analysis framework of GNN proposed in this paper for the MTSC task

1.2　多变量时间序列的定义与特性

多变量时间序列分类（MTSC）任务是指将一段多变量时间序列数据，或一个多变量时间序列数据集分配到一个预定义的类别标签。具体而言，给定一个多变量时间序列 $\boldsymbol{X} \in \mathbf{R}^{N \times T}$，其中 N 表示变量数，T 表示时间步数，任务的目标是学习一个分类器 f：$\mathbf{R}^{N \times T} \rightarrow \boldsymbol{Y}$，将该序列映射到一个类别标签 $y \in \boldsymbol{Y}$。因其数据对时间步长依赖，数据存在很高的维度且不平衡等问题，MTSC 任务具有很大的困难性。MTSC 问题的难点主要体现在以下方面：1）时间依赖性建模。MTS 中的时间步长依赖性可能会呈现短期与长期的多层次特征，传统模型往往难以捕捉长时间跨度的依赖；2）高维数据建模。多变量时间序列通常具有高维的特征空间，如何从这些高维数据

中提取有效特征，是分类任务的一大挑战；3）空间依赖性。不同变量间存在复杂的相互关系，尤其是在传感器网络、地理信息系统等场景中，变量间的空间关系至关重要；4）数据不平衡性。在很多应用中，少数类事件（如设备故障、金融危机等）非常稀缺，导致分类模型在训练过程中容易偏向多数类，从而造成漏检或误报。

这些问题的存在，对MTSC任务带来极大的干扰，使得MTSC任务效果下降。

1.3 图神经网络

1.3.1 图术语与表示

图是一种描述对象（节点）及其关系（边）的数据结构，本质上是一种非欧几里得数据结构，可以用于表示很多系统。图可以是静态的，也可以是动态的，所以对图的分析专注于节点分类、链接预测方向。具体而言，在GNN中，图表示为$\boldsymbol{G}=(\boldsymbol{V},\boldsymbol{E})$，式中，$\boldsymbol{V}$是节点集，表示图中的对象，$\boldsymbol{E}$是边集，表示节点之间的关系。在MTS中，节点通常对应于时间序列中的各变量，而边则表示变量之间的依赖关系。邻接矩阵$\boldsymbol{A}$表示节点之间的连接关系，$\boldsymbol{A}_{i,j}=1$表示节点i和节点j之间存在边。特征矩阵$\boldsymbol{X}$表示每个节点的特征，$\boldsymbol{X}_i\in\mathbf{R}^d$表示节点$i$的特征向量。这些都是常见的图表示。

图神经网络通过对节点之间的信息传递与聚合，学习每个节点的嵌入表示。这些嵌入表示不仅包含节点本身的特征，还包含邻居节点的信息。

1.3.2 消息传递框架

图神经网络的核心操作是消息传递（message passing）（Wu等，2021），如图2所示，其基本思想是通过聚合邻居节点的信息更新节点的特征表示。通常，消息传递过程包括两个步骤：聚合函数和更新函数。聚合函数（AGGREGATE）是指通过聚合的方式从邻居节点收集信息，常见的聚合方式包括求和、平均、最大池化等。更新函数（UPDATE）是将聚合的信息与节点的当前特征结合，更新节点的特征表示。假设在第1层，对于图$\boldsymbol{G}=(\boldsymbol{V},\boldsymbol{E})$，节点$v\in\boldsymbol{V}$，定义节点$v$的邻居为$N(v)$，则有：

$$\boldsymbol{m}_v^{(l)}=AGGREGATE^{(l)}\left(\left\{\boldsymbol{h}_u^{(l-1)}:\boldsymbol{u}\in\boldsymbol{N}(\boldsymbol{v})\right\}\right) \tag{1}$$

$$\boldsymbol{h}_v^{(l)}=UPDATE^{(l)}\left(\boldsymbol{h}_v^{(l-1)},\boldsymbol{m}_v^{(l)}\right) \tag{2}$$

式中，$\boldsymbol{h}_v^{(l)}$是节点v在第l层的表示，*AGGREGATE*是邻居特征的聚合函数（如求和、平均、最大池化），

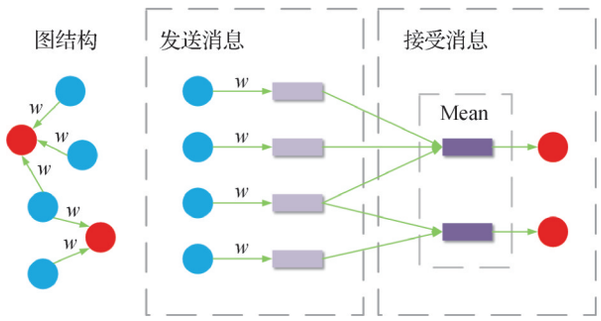


图2 消息传递机制

Fig. 2 Message passing mechanism

*UPDATE*是节点特征更新函数。

1.3.3 常见的GNN模型

本节详细列出常见的GNN模型：GCN（graph convolutional network）（Zhang等，2019）、GAT（graph attention network）（Veličković等，2018）、GraphSAGE（Hamilton等，2017）和ST-GNN（spatio-temporal GNN）（Yu等，2018）的原理及其公式以及它们的特点。

1）图卷积网络（GCN）。GCN将卷积操作从传统数据（图像或网格）推广到图数据。关键是学习一个函数f，通过聚合其自身的特征$\boldsymbol{X}_i$和邻居的特征$\boldsymbol{X}_j$生成节点$\boldsymbol{v}_i$的表示，其中$j\in N(\boldsymbol{v}_i)$。其思想是基于谱域卷积，将邻居节点特征按照归一化邻接矩阵进行聚合，并线性变换。GCN也是一个神经网络层，层与层之间的传播方式为

$$\boldsymbol{H}^{(l+1)}=\sigma\left(\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\boldsymbol{H}^{(l)}\boldsymbol{W}^{(l)}\right) \tag{3}$$

式中，$\tilde{\boldsymbol{A}}=\boldsymbol{A}+\boldsymbol{I}$表示在原邻接矩阵$\boldsymbol{A}$上加自环。$\tilde{\boldsymbol{D}}$对应$\tilde{\boldsymbol{A}}$的度矩阵，$\tilde{\boldsymbol{D}}_{ii}=\sum_j\tilde{\boldsymbol{A}}_{ij}$。$\boldsymbol{H}^{(l)}$、$\boldsymbol{W}^{(l)}$分别为第$l$层节点表示和第$l$层的可学习权重矩阵。$\sigma$是非线性激活函数，如ReLU（rectified linear unit）。通过对称归一化$\tilde{\boldsymbol{D}}^{-1/2}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{D}}^{-1/2}$平衡节点度数差异，聚合邻居信息并更新节点特征。GCN模型如图3所示。

在实际应用中，假设有两层GCN，激活函数为ReLU和softmax，则正向传播公式为

$$\boldsymbol{Z}=softmax\left(\tilde{\boldsymbol{A}}\cdot ReLU\left(\tilde{\boldsymbol{A}}\boldsymbol{X}\boldsymbol{W}^{0}\right)\cdot\boldsymbol{W}^{(1)}\right) \tag{4}$$

GCN在每一层进行一次局部（1-hop）邻居特征平均，这导致容易出现过平滑问题（deep GCN层数过多时所有节点表征趋同）。

2）图注意力网络（GAT）。GAT是一种基于注意力机制的图神经网络，专门用于处理图结构数据。GAT的思想是通过注意力机制学习节点对邻居节点

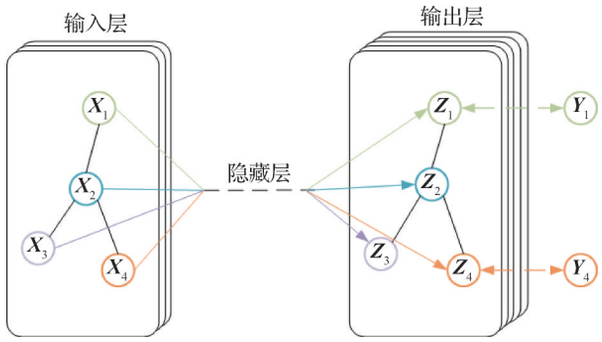


图3　经典GCN模型

Fig. 3　Classic GCN model

的重要性权重，自适应地加权聚合。GAT主要分为两步：注意力权重计算和特征聚合。

注意力权重计算式为

$$e_{ij} = LeakyReLU\left(\vec{a}^{\mathrm{T}}\left[\overrightarrow{\boldsymbol{W}\boldsymbol{h}_i} \,\|\, \overrightarrow{\boldsymbol{W}\boldsymbol{h}_j}\right]\right) \tag{5}$$

式中，‖ 表示特征连接，$\vec{\boldsymbol{a}}$ 表示可学习的注意力向量，$\boldsymbol{W}$ 表示特征变换矩阵。

归一化注意力权重（softmax）计算式为

$$\alpha_{ij} = \frac{\exp\left(e_{ij}\right)}{\sum_{k \in N(i)} \exp\left(e_{ik}\right)} \tag{6}$$

式中，$N(i)$表示节点i的一阶邻居集合，α_{ij}表示节点j对节点i的重要性权重。特征聚合（节点更新）如图4所示。

GAT的优势在于使用注意力机制进行动态权重分配，实现归纳式的学习，注意力权重的非对称性天然支持有向图处理。相比于GCN使用节点度预先定义权重，GAT注意力机制自动学习权重，提升模型

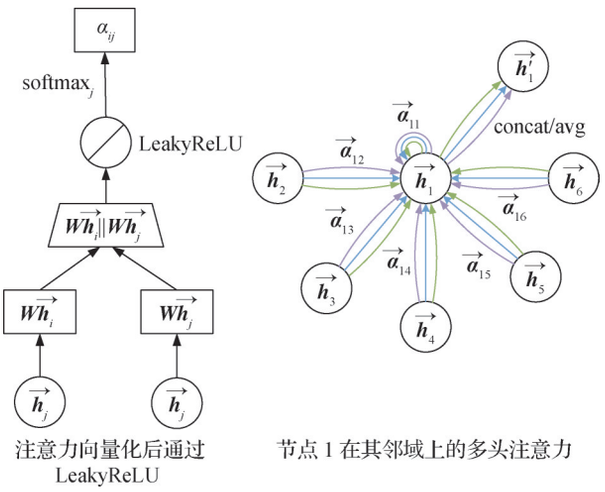


图4　特征聚合（节点更新）

Fig. 4　Feature aggregation (node update)

表达能力和任务适应性。

3）采样式图卷积（GraphSAGE）。GraphSAGE的思想是针对大规模图提出的，通过固定采样邻居节点进行高效聚合，避免内存爆炸。具体步骤是通过邻居采样和特征聚合生成节点嵌入，GraphSAGE的样本和聚合方法的可视化说明如图5所示。其分层传播计算为

$$\boldsymbol{h}_v^{(l+1)} = \sigma\left(\boldsymbol{W}^{(l)} \cdot AGGREGATE\left(\left\{\boldsymbol{h}_u^{(l)}: u \in N(v)\right\} \cup \left\{\boldsymbol{h}_v^{(l)}\right\}\right)\right) \tag{7}$$

式中，*AGGREGATE* 可取 Mean，Pooling，LSTM aggregator等。

GraphSAGE的优势在于通过固定数量邻居采样，支持归纳式学习，可泛化到新节点。在大规模图

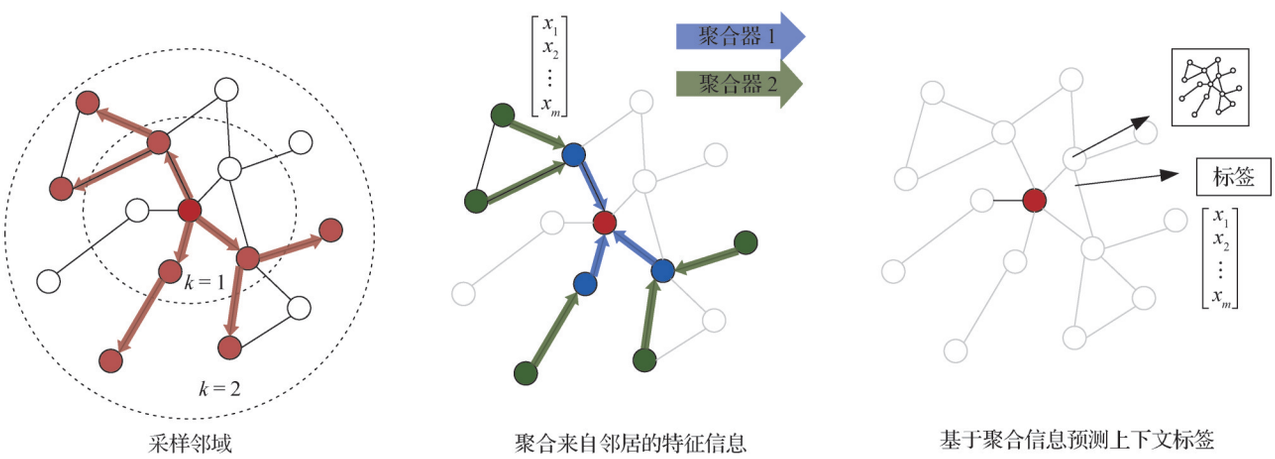


图5　GraphSAGE样本和聚合方法的可视化说明

Fig. 5　Visualization of GraphSAGE samples and aggregation methods

上聚合效率高。可以通过聚合函数动态调整权重，灵活性更强。

4)时空图神经网络(ST-GNN)。ST-GNN是一种专门用于处理兼具时间动态性和空间关联性复杂数据的深度学习模型。ST-GNN的思想是同时建模节点间的空间依赖与节点自身的时间演化特征，将在4.3小节进一步详细说明。ST-GNN主要分为两个模块时序卷积：时间卷积(temporal convolution，TCN)(Lea等，2016)和空间图卷积(GCN)(Zhang等，2019)。

时间卷积(TCN)和空间图卷积(GCN)计算为

$$\boldsymbol{H}' = TCH(\boldsymbol{X}) \tag{8}$$

$$\boldsymbol{H}^{(l+1)} = \sigma\left(\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\tilde{\boldsymbol{A}}\tilde{\boldsymbol{D}}^{-\frac{1}{2}}\boldsymbol{H}^{(l)}\boldsymbol{W}^{(l)}\right) \tag{9}$$

ST-GNN通过时空注意力(如ST-GAT(Song等，2022))动态调整节点权重，增强对关键时空关系的捕捉。传统时序模型(如ARIMA(autoregressive integrated moving average)(Ho和Xie，1998))忽略空间关联性，而ST-GNN通过图结构显式建模多变量间的空间交互(如相邻传感器数据的相互影响)。ST-GNN在交通流量预测(Bui等，2022)、气象与能源预测(Khodayar和Wang，2019)以及动态社交网络分析(Min等，2021)等方面应用广泛。传统方法和原始GNN方法的区别如表2所示。

表2　传统方法与原始GNN方法的对比

Table 2　Comparison between traditional methods and original GNN methods

对比维度	传统方法	原始GNN方法
依赖建模	依赖独立特征建模，无法捕捉多变量之间的复杂依赖	消息传递机制捕捉多变量复杂依赖，基于固定拓扑结构建模
空间建模	仅支持欧氏空间，例如基于距离和相关性建模	支持非欧氏空间的任意结构建模，如社交网络、知识图谱
时间建模	主要依赖RNN、LSTM、GRU，长时间依赖难捕捉	原始GNN无法建模时间依赖，需结合RNN或外部时序模型
异构数据处理	需要手动特征工程，多模型组合，难以对多模态数据统一建模	支持异构节点，但需要预先定义邻接关系
动态图处理	缺乏动态更新能力，通常需要重新训练	原始GNN不支持动态图，需要重新训练来更新结构
可解释性	部分方法具有解释性(例如决策树)，但深度模型难以解释	模型的可解释性较弱，难以跟踪多跳信息的来源
计算复杂度	对长时间序列和高维数据敏感，计算成本高	对大规模图计算开销大，复杂度为$O(N^2)$

2　面向MTSC的图构建策略

在多变量时间序列分类(MTSC)任务中，数据的复杂性源于时间序列中的多个变量及其相互之间的关系。传统的时间序列建模方法(如基于特征工程的统计学习(Nargesian等，2017)、RNN、LSTM等)往往忽视了时间序列数据中变量之间的空间关系和时序关系，导致这些方法在处理涉及多个变量之间复杂依赖关系时的局限性。为了有效捕捉这些关系，图结构作为一种天然的工具，能够在数据的每个时间点上对变量进行建模，进而增强模型对数据中复杂依赖结构的捕捉能力。

2.1　图构建策略

在面向多变量时间序列分类(MTSC)的任务中，图构建方法直接影响图神经网络(GNN)的表现。根据图构建时所依赖的先验信息和数据特性，图构建方法主要可分为4类：基于结构的图、基于相似度的图、基于注意力机制的图学习和基于预测或隐式优化的图学习(Jin等，2024b)。这些方法各有优缺点，根据不同的应用场景和任务需求，可以选择合适的图构建策略。

2.1.1　基于结构的图(structure-driven graphs)

在基于结构的图构建方法中，图的构建通常依据已知的领域知识或外部信息，如传感器的物理连接、设备的依赖关系等。这里的图结构是静态的，可

以通过已知的拓扑结构或物理位置定义节点和边。常见的应用场景包括智能交通（Yu 等，2018；Li 等，2018）、工业物联网（Selim 等，2021）和环境监测等。

假设有一个传感器网络，传感器节点 v_i 和 v_j 之间的边 e_{ij} 表示它们之间的物理或功能关系。邻接矩阵 $\boldsymbol{A}$ 的元素定义为

$$\boldsymbol{A}_{ij}=\begin{cases}1 & \text{传感器}v_i\text{与}v_j\text{之间存在连接}\\ 0 & \text{其他}\end{cases}$$

这个邻接矩阵 $\boldsymbol{A}$ 用于表示传感器之间的依赖关系，式中，$\boldsymbol{A}_{ij}=1$ 表示传感器节点 v_i 和 v_j 有物理连接或依赖关系，$\boldsymbol{A}_{ij}=0$ 表示没有直接的连接。

基于结构的图构建方法的优点是具有稳定性和高效性，图的结构明确且固定，容易理解和实现。在有明确物理连接或拓扑结构时，图结构的构建简单且高效，适合大规模应用。缺点是该方法高度依赖领域专家知识，对于没有明显结构信息或不具备领域知识的任务，难以使用。且在数据变化或领域迁移时，手工构建的图结构可能不再适用。

2.1.2　基于相似度的图（similarity-driven graphs）

基于相似度的图构建方法通常通过计算节点（即变量）之间的相似度来确定边的连接。常见的相似度度量包括 Pearson 相关系数、余弦相似度和动态时间规整（DTW）等。例如，Ahanger 等人（2025）在物联网入侵检测中利用欧几里得距离构建基于相似度的图结构；GraphCNNpred（Jin，2025）则通过可见度图将股票历史数据转化为图结构。

给定时间序列 $\boldsymbol{X}_i=\left\{\boldsymbol{x}_{i,1},\boldsymbol{x}_{i,2},\cdots,\boldsymbol{x}_{i,T}\right\}$ 和 $\boldsymbol{X}_j=\left\{\boldsymbol{x}_{j,1},\boldsymbol{x}_{j,2},\cdots,\boldsymbol{x}_{j,T}\right\}$，Pearson 相关系数 r_{ij} 可以表示为

$$r_{ij}=\frac{\sum_{t=1}^{T}\left(\boldsymbol{x}_{i,t}-\bar{\boldsymbol{x}}_i\right)\left(\boldsymbol{x}_{j,t}-\bar{\boldsymbol{x}}_j\right)}{\sqrt{\sum_{t=1}^{T}\left(\boldsymbol{x}_{i,t}-\bar{\boldsymbol{x}}_i\right)^2\sum_{t=1}^{T}\left(\boldsymbol{x}_{j,t}-\bar{\boldsymbol{x}}_j\right)^2}}\tag{10}$$

式中，$\bar{\boldsymbol{x}}_i$ 和 $\bar{\boldsymbol{x}}_j$ 分别是时间序列 $\boldsymbol{X}_i$ 和 $\boldsymbol{X}_j$ 的均值。式（10）用于衡量时间序列 $\boldsymbol{X}_i$ 和 $\boldsymbol{X}_j$ 在时间上的线性相关性。

对于两个向量 $\boldsymbol{x}_i$ 和 $\boldsymbol{x}_j$，余弦相似度 s_{ij} 的定义为

$$s_{ij}=\frac{\sum_{t=1}^{T}\boldsymbol{x}_{i,t}\boldsymbol{x}_{j,t}}{\sqrt{\sum_{t=1}^{T}\boldsymbol{x}_{i,t}^2}\sqrt{\sum_{t=1}^{T}\boldsymbol{x}_{j,t}^2}}\tag{11}$$

DTW 是衡量时间序列相似性的一种方法，特别适用于非线性对齐的情况。DTW 的核心思想是通过对齐两个序列的不同部分，最小化它们之间的距离，计算方式为

$$DTW\left(\boldsymbol{X}_i,\boldsymbol{X}_j\right)=\min_{\pi\in\Pi}\sum_{t=1}^{T}d\left(\boldsymbol{x}_{i,t},\boldsymbol{x}_{j,\pi(t)}\right)\tag{12}$$

式中，$d\left(\boldsymbol{x}_{i,t},\boldsymbol{x}_{j,\pi(t)}\right)$ 表示时间点 t 处两个时间序列之间的距离，Π 是所有可能的对齐路径。

根据相似度度量的不同，可以构建出不同的图邻接矩阵 $\boldsymbol{A}$，其中，$\boldsymbol{A}_{ij}$ 表示时间序列 $\boldsymbol{X}_i$ 和 $\boldsymbol{X}_j$ 的相似度，通常通过设定阈值 τ 来决定是否建立边

$$\boldsymbol{A}_{ij}=\begin{cases}1 & \boldsymbol{X}_i\text{与}\boldsymbol{X}_j\text{的相似度超过}\tau\\ 0 & \text{其他}\end{cases}$$

基于相似度的图构建方法的优点在于仅依赖于数据本身的相似性，可以适应各种不同的数据集，能够动态构建图，不需要固定的结构，特别适合数据分布不明确或无物理结构的任务。缺点是对噪声敏感且计算复杂度较高。

2.1.3　基于注意力机制的图学习（attention-driven graph learning）

基于注意力机制的图构建方法通过自适应学习节点之间的关系分配权重，这些权重代表了节点之间的相对重要性。

基于注意力机制的图构建方法由于注意力机制或者多头注意力机制（Jiang 等，2022）的引入能够根据数据的变化动态调整图结构，且能让模型聚焦于关键的依赖关系，从而减少不重要节点的干扰。如 ASTGCN（Zhu 等，2021）利用注意力机制在时空维度上自动选择重要节点以增强特征提取；Instance-Attention GNN（Han 等，2025）则引入实例级注意力机制以适应非规则时间序列的动态边权。

2.1.4　基于预测或隐式优化的图学习（latent graph learning）

基于预测或隐式优化的图学习方法通过端到端训练自动学习图结构。典型代表如 MTSGNN（graph neural network for multivariate time series）（Zhang，2024），该模型通过自适应图学习模块自动推断变量间的依赖关系，无需先验知识即可在训练中优化图结构。在这些方法中，邻接矩阵通常作为可学习的参数，与其他模型组件（如图卷积、全连接层等）一起优化。

假设邻接矩阵 $\boldsymbol{A}$ 是一个可学习的参数矩阵，图神经网络的损失函数可以定义为

$$L=\sum_i\left\|\boldsymbol{Y}_i-f\left(\boldsymbol{A},\boldsymbol{X}_i\right)\right\|^2\tag{13}$$

式中，$\boldsymbol{Y}_i$是标签，$\boldsymbol{X}_i$是节点特征，$f(\boldsymbol{A}, \boldsymbol{X}_i)$是基于图结构$\boldsymbol{A}$和节点特征$\boldsymbol{X}_i$的模型输出。通过最小化损失函数，可以学习出最优的图结构$\boldsymbol{A}$。

这种方法的好处是无需显式图结构，适应性强，而且可以让图学习与任务优化同时进行。但是缺点也很明显，由于图结构是可学习的，训练过程可能更加复杂，容易出现过拟合或收敛缓慢的问题。而且一旦学习到不合理的图，会让模型性能变得不稳定。

在MTSC任务中，图构建方法对模型性能影响显著。基于结构的静态图适用于领域场景中有可靠先验（如道路网络、传感器拓扑）的应用，计算开销低、解释性强，但对时变依赖难以适应；基于相似度的动态图通过Pearson相关、DTW等从数据中自动构建边，能捕获数据驱动的关联，但易受噪声干扰、计算复杂度高；基于注意力的图学习如GAT自适应分配边权，可抑制不重要关系，但需大量数据支撑注意力学习；隐式图学习是一种端到端学习，图结构最为灵活，但训练不稳定且易过拟合。

4种图构建方式各有利弊，若关注实时动态变化（如流式监测、股市交易），可优先考虑注意力或隐式图学习；若更注重可解释性与效率（如工业设备监控、路网管理），则结构驱动方法更合适。

2.2 静态与动态图表示

在多变量时间序列分类（MTSC）任务中，图的表示方式对于模型能够有效捕捉时空依赖关系至关重要。根据图结构是否随时间变化，可以将图分为静态图和动态图。静态图通常假设图的拓扑结构在整个时间序列的过程中保持不变，而动态图则认为随着时间的推进，图的结构会根据新的数据和时序信息进行动态调整。选择静态图还是动态图通常取决于数据的特点和应用的需求。

静态图是指在多变量时间序列分类任务中，整个时间序列过程的所有时间步使用相同的图结构。每个变量在不同时间步的依赖关系通过固定的图结构来表示。静态图假设数据中变量之间的关系在时间演化过程中保持不变，因此构建出的图是固定的，图中的节点和边在整个序列中都是一样的。静态图通常通过领域知识或外部数据（如传感器物理连接、设备依赖关系、网络拓扑等）构建图结构，确定节点间的边及其连接权重。例如，智能交通中的传感器网络可以通过路网图表示传感器之间的关系，而工业监控中的传感器依赖关系则可以通过设备拓扑来表示。静态图的核心特点是图的邻接矩阵（或类似的图结构表示方法）在整个序列中保持不变。因此，在图卷积操作中，图的拓扑结构不发生变化。静态图的优点是计算高效、稳定性强、简单且易实现。缺点是无法捕捉时变依赖且对动态系统的适应性差。

动态图则认为图结构随时间变化，每个时间步的图结构可能不同。动态图能够捕捉多变量时间序列中的时变依赖，即变量之间的关系在时间的推移下会发生变化。例如，某些设备在不同时间段内可能会产生不同的依赖关系，某些交通路段的拥堵情况会随时间发生变化。这些时间依赖关系需要通过动态图来建模。在动态图模型中，每个时间步的图结构都会根据当前时刻的观测数据进行动态调整。通常通过计算当前时刻节点之间的相似性或相关性更新图的邻接矩阵。例如，可以基于时间序列数据的相似度计算，动态地调整邻接矩阵的值。动态图的每个时间步都有一个单独的邻接矩阵，这意味着每个时间步都会进行图卷积操作。随着时间的推移，节点之间的连接关系和传播方式会随着数据的演变而不断更新（Liu等，2024）。动态图的特性决定了其灵活性强、建模高效、准确率较静态图高（Hamilton等，2017），但同样会带来计算复杂度高和不稳定的弊端。

尽管这类方法因其灵活性和自适应性而得到广泛应用，但其有效性建立在一个深刻且具有风险的核心假设之上："历史数据的统计相关性等同于未来交互的有效表征"。这一假设在面对真实世界的复杂系统时，暴露出其根本性的脆弱，主要体现在以下两个方面。其一，关联不等于因果（correlation is not causation）：在金融、医疗和社会科学等领域，虚假相关性（spurious correlation）是一种普遍现象。例如，在一个经济体中，冰淇淋销量和溺水人数可能呈现高度正相关，但两者并无直接因果联系，它们都是由一个未被观测到的混杂变量（confounding variable）"季节温度"驱动。如果基于这种虚假相关性构建图的边，GNN模型中进行的消息传递本质上是在放大和传播由混杂因素引起的噪声，而非系统内在的有效信息。这可能导致模型过拟合于历史的巧合，而在系统状态发生变化时（例如，宏观经济政策调整），其预测能力会急剧下降。其二，对非平稳性的敏感

性(sensitivity to non-stationarity)：许多真实世界的MTS,如金融时间序列,都具有显著的非平稳性,其统计特性(如均值、方差)会随时间变化。在一个固定的历史窗口上计算出的相似度矩阵,可能无法代表未来的变量关系。例如,在金融市场中,两家公司的股票在市场平稳期可能表现出低相关性,但在危机期间可能因为共同的风险暴露而呈现出高度的同步下跌。一个静态的、基于历史平均关系的图结构,将无法捕捉这种依赖关系的动态权重转换(regime shift)。

因此,数据驱动的图关联学习方法的理论有效性边界在于:系统变量间的内在关系必须足够稳定,且观测到的变量集应足以捕捉其主要驱动力,同时混杂效应较弱。这解释了为何此类方法在一些相对稳定的物理或生物系统中(例如,部分传感器网络应用)可以取得不错的效果,但在高度动态、充满噪声和潜在因果混淆的社会经济系统中,往往需要与更鲁棒的机制(如动态图更新、因果发现)相结合,才能构建出真正可靠的模型。

2.3 讨论:选择合适的图构建策略

在多变量时间序列分类(MTSC)任务中,图构建策略的选择直接影响图神经网络(GNN)模型的表现。选择合适的图构建方法需要综合考虑多方面因素,包括数据的特性、任务的需求、图结构的动态性以及计算资源等。以下讨论将详细分析这些因素,并为实际应用提供选择依据。

2.3.1 数据特性:时空依赖性与噪声水平

时空依赖性是多变量时间序列数据的一个核心特性。在一些应用场景中,变量之间的空间依赖性(例如传感器之间的物理连接)是非常明确且固定的,此时使用基于结构的图(如传感器布置图)构建方法较为合适。然而,在某些领域中,变量间的依赖关系随着时间的推移发生变化,尤其是在处理动态系统(如股市、交通流量等)时,基于结构的静态图可能无法有效捕捉时变依赖关系,此时动态图可能更加适用。另一方面,噪声水平也会影响图构建的选择。基于相似度的图方法(如相关性计算、DTW)通常依赖于数据的相似性或相关性度量,但这种方法对噪声较为敏感。如果数据中存在较高的噪声,可能导致计算出的相似度不准确,从而影响图的质量。在这种情况下,基于注意力机制的图学习方法能够通过自适应调整节点间的关系,减少噪声的干扰,可能更加适合。

2.3.2 任务需求:空间依赖与时间依赖的建模

在不同的应用任务中,对空间依赖性和时间依赖性的建模需求可能存在差异。例如,在医疗健康监测中,不同传感器或设备之间的空间依赖关系非常重要,而时间依赖性(如心率变化、体温波动)也是关键特征之一。因此,在此类任务中,同时捕捉空间和时间依赖显得尤为重要。为此,结合时空图神经网络(ST-GNN)和基于注意力的图学习方法可能是较为理想的选择,因为它能够在建图时灵活地捕捉空间和时间维度的依赖。对于传统的静态系统(如传感器网络、物联网设备监控等),空间依赖性较为固定,因此可以使用基于结构的图方法来建立节点间的连接关系。而对于时变系统(如金融市场分析、交通流量预测等),系统中变量之间的依赖关系随着时间不断变化,选择动态图或基于注意力机制的图学习方法可能会更加适合,因为这些方法能够动态调整图结构,更好地适应时序数据的变化。

2.3.3 计算资源:静态图与动态图的权衡

图的构建方法对计算资源的需求有着重要影响,尤其是在大规模数据处理时,图的更新和计算可能带来较大的计算开销。静态图的一个显著优点是图结构在整个时间序列中保持不变,因此只需计算一次图的邻接矩阵即可进行后续的图卷积操作。这使得静态图在计算上相对高效,尤其适用于较大规模、计算资源有限的场景。

然而,动态图需要在每个时间步重新构建图结构,并且随着时间的推移,节点之间的关系可能发生剧烈变化,因此需要动态更新图的邻接矩阵。这不仅增加了计算复杂度,而且也对存储和内存提出了较高要求。对于时序较长、时间步较多的任务,动态图可能会导致显著的性能瓶颈,尤其是在内存和存储资源有限的情况下。在此类情况下,如果能够有效地压缩动态图结构(例如使用图稀疏化技术或邻居采样),则可以减轻计算负担。

2.3.4 任务的可解释性:图结构与注意力机制的结合

在某些应用领域,可解释性是非常重要的,尤其是在医疗诊断、金融分析等高风险场景中,模型的决策过程需要透明且可追溯。基于结构的图构建方法通常能够提供较好的可解释性,因为图的结构是显式定义的,每个节点和边代表的数据关系可以通过

领域知识进行解释。然而，随着图的复杂性增加，尤其是动态图或基于注意力机制的图学习方法，模型的解释性可能变得较为困难。为了解决这一问题，一些研究提出将注意力机制与图神经网络结合（Huang等，2024；Veličković等，2018）。通过图注意力网络（GAT）等方法，模型能够根据不同节点的依赖关系自动调整权重，从而在一定程度上提高了可解释性。通过对图中的注意力权重进行分析，研究人员可以理解模型如何进行决策，以及各节点之间的相对重要性。

2.3.5　领域知识的融合：图构建的自动化与半自动化

尽管基于结构的图方法在一些领域中表现出色，但对于没有明确结构的任务，基于相似度的图构建方法和基于注意力的图学习方法能够更好地从数据中学习图结构。数据驱动的图学习方法不需要专家知识的干预，能够自动发现数据中的潜在关系，因此在复杂任务中更具灵活性和适应性。例如，在金融市场预测中，节点之间的关系是高度动态且难以明确预定义的（Wang等，2022a），基于相似度的图构建方法（如使用动态时间规整DTW计算相似性）能够自动捕捉股票之间的相关性变化。而在交通流量预测中，基于注意力机制的图学习方法可以通过自动分配不同交通路段间的权重，灵活调整模型的依赖结构（Wang等，2022b）。

3　面向多变量时间序列分类的GNN架构的分类体系

3.1　解耦式建模范式（the decoupled modeling paradigm）

解耦式建模范式的核心哲学思想是“分而治之”（divide and conquer）。它将复杂的时空依赖问题分解为两个相对独立且更易处理的子问题：首先，在每个时间步上捕捉纯粹的空间依赖；然后，将提取出的空间特征序列送入一个独立的时间模型来捕捉时间动态。这种范式在概念上将空间和时间维度视为可分离的。

时间GNN（temporal graph neural network）（Bai等，2020）是一种专门用于处理时序数据的图神经网络架构，尤其适用于那些图结构随时间变化的任务。传统的图神经网络（GNN）通常用于静态图建模，节点之间的依赖关系是固定的。然而，许多多变量时间序列数据的图结构在时间维度上是动态变化的，这使得时间GNN在建模动态时序依赖时表现出独特的优势。

时间GNN的一个关键特征是图结构随时间变化。每个时间步t中的图结构可能不同，节点之间的关系会根据当前时间步的数据而动态调整。这种方法非常适合处理那些依赖关系随时间变化的多变量时间序列数据。假设在每个时间步t都根据当前数据构建一个图结构。图的邻接矩阵$\boldsymbol{A}_t$可以根据当前时间步的变量之间的相似度、相关性或通过其他度量进行更新。具体而言，邻接矩阵$\boldsymbol{A}_t = f(\boldsymbol{X}_t, \boldsymbol{X}_{t-1}, \cdots, \boldsymbol{X}_1)$。$\boldsymbol{X}_t$表示在$t$时刻所有节点（变量）的特征，$f$是一个函数，表示如何根据时间序列数据动态构建邻接矩阵。例如，这可以是基于变量之间的相似度计算，或基于时序依赖的建模。节点的特征更新过程可表示为

$$\boldsymbol{H}_i^{(t)} = \sigma\left(\sum_{j \in \boldsymbol{N}(i)} \boldsymbol{A}_{ij}^{(t)} \boldsymbol{H}_j^{(t-1)} + \boldsymbol{W}_t \boldsymbol{X}_i^{(t)}\right) \tag{14}$$

式中，$\boldsymbol{H}_i^{(t)}$是节点i在时刻t的特征表示，$\boldsymbol{A}_{ij}^{(t)}$是$t$时刻邻接矩阵中的元素，表示节点$i$和节点$j$之间的连接强度，$\boldsymbol{N}(i)$表示节点$i$的邻居集合，$\boldsymbol{W}_t$表示时间步$t$的权重矩阵，$\boldsymbol{X}_i^{(t)}$表示节点$i$在时刻$t$的输入特征。

在时间GNN中，节点的表示不仅受到当前时间步的邻居节点的影响，还可能受到历史时间步的影响。在完成节点特征的动态更新后，针对MTSC任务，通常采用读出函数（readout function）（如全局平均池化）将所有时间步和节点的特征聚合为统一的图级向量，随后接入全连接层以输出整个时间序列的类别标签。这就要求在每个时间步中，图结构和节点特征都进行信息聚合和消息传递。消息传递的过程可以表示为

$$\boldsymbol{H}_i^{(t)} = \sum_{j \in N(i)} \boldsymbol{\alpha}_{ij}^{(t)} \boldsymbol{H}_j^{(t-1)} + \beta_t \boldsymbol{X}_i^{(t)} \tag{15}$$

式中，$\boldsymbol{\alpha}_{ij}^{(t)}$是由注意力机制计算的注意力权重，表示节点$i$和节点$j$之间的依赖强度，$\boldsymbol{H}_j^{(t-1)}$是节点$j$在上一时间步的特征表示（表示时间依赖性），$\beta_t$是时间步$t$的系数，用于调节时间特征的影响。

式（15）展示了信息聚合的过程，式中每个节点的表示不仅受到其邻居节点在当前时间步的影响，还受到邻居节点在上一时间步的影响。此外，注意

力机制使得模型能够自适应地为不同节点之间的关系分配不同的权重，进一步加强了模型在时序依赖建模中的灵活性。

时间 GNN 的优势在于同时建模时序数据的空间与时间依赖，能够通过动态图结构更新邻接关系，有效捕捉复杂的时空交互。然而，其应用仍面临挑战：一是每个时间步需更新图结构，计算复杂度较高；二是训练过程因结构动态变化而不稳定，且可解释性不足；三是复杂时空数据使节点关系与更新机制难以直观理解，信息传播变化可能导致收敛缓慢或过拟合。

在多变量时间序列分类（MTSC）任务中，图神经网络（GNN）通过建模节点之间的关系和依赖，展现了捕捉时空依赖的潜力。为了充分挖掘时间序列数据中的时序信息，GNN 模型可以与其他时间序列处理模型，如循环神经网络 RNN（Caterini 和 Chang，2018）和卷积神经网络 CNN 结合形成时间 GNN，从而提高对时序动态变化的捕捉能力。

在空间 GNN 层中，可以使用图卷积网络（GCN）或图注意力网络（GAT）等方法，捕捉不同变量之间的相互关系。通过节点间的信息传递，学习出每个时间步中节点（变量）之间的空间依赖。假设图的邻接矩阵为 $\boldsymbol{A}$，节点特征矩阵为 $\boldsymbol{H}$，在图卷积操作中，节点 i 的表示 $\boldsymbol{H}_i$ 会受到其邻居节点 j 信息的影响。随后，空间 GNN 的输出被分别输入到 RNN/CNN 中建模时间表示。

在这种组合架构中，空间 GNN 首先对每个时间步的多变量时序数据进行图卷积操作，学习出节点间的空间关系表示。然后，这些表示作为输入传递给 RNN 模型，进一步学习时间步之间的时序依赖。

获得空间 GNN 层的输出特征后，这些特征被传递给 RNN 层进行时序建模。RNN 通过递归地传递隐藏状态（Yu 等，2021）捕捉时间序列中的时序依赖。与序列预测不同，在分类任务中，通常提取 RNN 在最后一个时间步的隐藏状态作为整个序列的语义编码，直接输入 softmax 分类器进行判别。

空间 GNN 与 RNN 结合的架构能够同时处理多变量时间序列数据中的空间依赖和时序依赖。该模型能够有效提升对复杂时空依赖关系的建模能力。RNN 与 CNN 的结构区别如图 6 所示。

CNN 是一种擅长从数据中提取局部特征的模型，如图 6 所示。对于多变量时间序列数据，CNN 能够通过滑动窗口操作自动提取时间上的局部模式（Gu 等，2018）。然而，CNN 通常无法捕捉节点之间的空间依赖。将空间 GNN 与 CNN 结合，能够弥补这一缺点，同时有效捕捉时空依赖。

空间 GNN 层的输出特征将作为输入传递给 CNN 层。CNN 层通过卷积操作提取数据中的局部时间特征。通过滑动窗口，CNN 能够捕捉到时间序列中的局部模式和周期性变化。具体为

$$\boldsymbol{Z}_i = \sum_k \boldsymbol{W}_k * \boldsymbol{H}_i^{(t+k)} \tag{16}$$

式中，$\boldsymbol{Z}_i$ 是卷积操作的输出，表示节点 i 的局部时空特征，$\boldsymbol{W}_k$ 是卷积核，$\boldsymbol{H}_i^{(t+k)}$ 是时刻 $t+k$ 的节点特征。

CNN 提取到的时空特征将通过全连接层（full connection，FC）进行分类或回归预测。二者结合能够更有效地提取时空特征，能够同时考虑到空间依赖性和时间序列的局部相关性。例如，GraphCNNpred（Jin，2025）便采用了解耦策略，分别利用 CNN 和 GCN 提取时间与空间特征，成功应用于股票市场状态分类。

3.2 耦合式建模范式（the coupled modeling paradigm）

与解耦式相反，耦合式建模范式认为时空依赖是不可分割的整体（inseparable whole）。其设计哲学是必须在一个统一的操作中联合建模时空邻域信息，从而直接捕捉局部时空模式。

时空图神经网络（ST-GNN）结合了空间图卷积和时间卷积（或时间门控网络），能够同时建模多变量时间序列中的空间依赖和时间依赖，是典型的耦合式建模范式。ST-GNN 通过集成图卷积网络和时间卷积网络，能够有效捕捉到时间序列中多层次的时空依赖关系。在分类任务中，MTSGNN（Zhang，2024）沿用了这一耦合范式，通过联合建模时空依赖，在多个 MTSC 基准数据集上取得了优异性能。

3.2.1 基于卷积时间模块的架构

时空图卷积网络（spatio-temporal graph convolutional network）结合了图卷积（GCN）和时间卷积（如 1D-CNN），通过两级卷积操作分别提取时空特征。STGCN 首先在空间维度上对节点进行图卷积操作，然后对时间维度进行卷积，以捕捉时间序列数据的长期和短期依赖。

STGCN 结合了图卷积与时间卷积，能够有效捕捉时空依赖性，适用于动态时空数据。时间卷积操

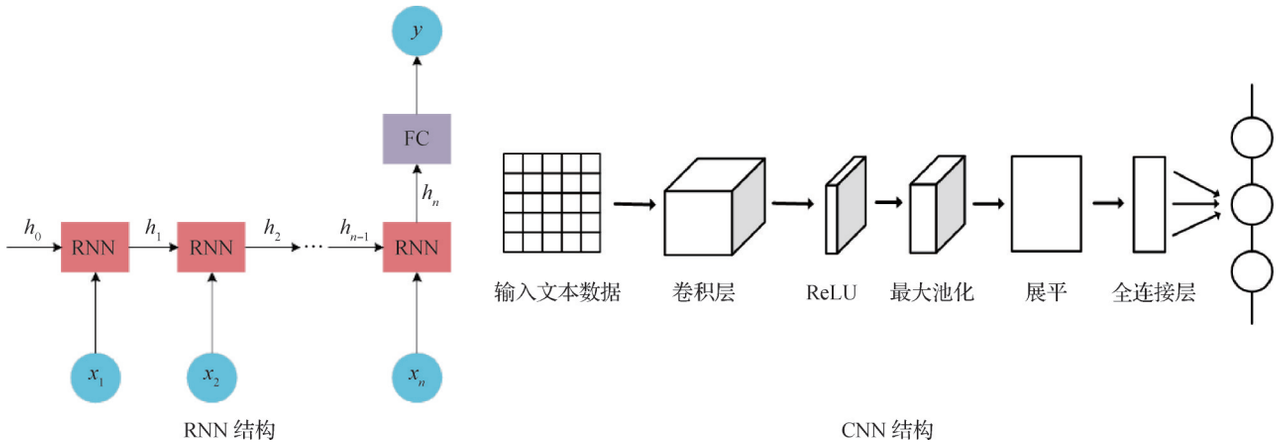


图6　RNN和CNN结构

Fig. 6　RNN and CNN structures

作适合处理长时间序列数据，尤其是对短期依赖关系的建模非常高效。同时，STGCN需要在每个时间步进行图卷积和时间卷积计算，计算复杂度较高。值得注意的是，原始 STGCN 多用于交通流预测，输出为未来时刻数值。在应用于 MTSC 分类（如动作识别）时，需要在时空卷积块之后引入时空全局池化层（global spatiotemporal pooling），将三维时空特征张量压缩为固定维度的特征向量，从而适配分类任务中单一类别标签的输出需求。虽然STGCN能够有效建模短期时序依赖，但对于更复杂的长期依赖关系，可能需要结合更强的时序建模能力（如RNN）。

3.2.2　基于循环时间模块的架构

基于循环时间模块的架构通过引入循环神经网络（RNN）或门控机制，进一步增强了模型在处理长期时序依赖方面的能力。与基于卷积的架构不同，这些方法通过结合图卷积与RNN（或GRU/LSTM）捕捉图结构的时空依赖。

扩散卷积递归神经网络（diffusion convolutional recurrent neural network，DCRNN）结合了图卷积与RNN，专门用于捕捉时序数据中的时空依赖。DCRNN（Li等，2018）采用扩散卷积建模空间依赖，通过RNN捕捉时间依赖。

DCRNN在空间卷积上使用扩散卷积，计算为

$$\boldsymbol{H}_i^{(t)} = \sum_{j \in N(i)} \boldsymbol{P}_{ij} \boldsymbol{H}_j^{(t-1)} + \boldsymbol{W}^{(t)} \boldsymbol{X}_i^{(t)} \tag{17}$$

式中，$\boldsymbol{P}_{ij}$是通过扩散操作生成的邻接矩阵，表示节点之间的空间依赖关系。

DCRNN在时间卷积上使用循环卷积，通过引入RNN能够有效捕捉长时间依赖，适合长时间序列数据的处理。DCRNN能够同时建模时空依赖，提升了对复杂动态系统的预测能力。但由于需要计算扩散卷积和RNN的反向传播，训练过程复杂且计算开销较大，需要同时优化图卷积和RNN参数，训练过程相对较为复杂。在处理 MTSC 任务（如故障诊断）时，DCRNN 的编码器—解码器结构通常会被简化：仅保留编码器部分，利用循环单元随时间演化的最终隐藏状态作为多变量序列的压缩表示，直接用于下游的故障类别判定，而非生成未来的序列预测值。

3.2.3　基于注意力/Transformer时间模块的架构

基于注意力/Transformer时间模块的架构，如ASTGCN（Zhu等，2021）、图Transformer（Lyu等，2019）引入了注意力机制和Transformer架构，以更好地捕捉多尺度时空依赖关系。与传统的卷积和循环神经网络不同，注意力机制能够根据每个节点在图中的重要性，自适应地调整节点间的信息传播。

注意力时空图卷积网络（attentive spatio-temporal graph convolutional network）结合了时空图卷积与自注意力机制，能够同时捕捉时空依赖，并通过注意力机制自适应调整节点之间的依赖强度。在时间卷积和注意力机制的共同加权下更新节点，这种设计确保了模型能够全面建模时空依赖，并通过自适应学习提升对不同数据模式的处理能力。

这种设计的好处是能够有效建模数据中的时空依赖，自适应调整传播的权重，高效提取时空特征。这种机制在分类任务中尤为关键，因为它允许分类器通过注意力权重自动聚焦于决定类别的关键时间片段或核心变量（例如，在医疗诊断中聚焦于心率异常波动的几秒钟），从而显著提高分类的准确性与可

解释性。缺点依旧是图神经网络的共有缺点，计算复杂度高和可解释性差。

不同ST-GNN在时间维度的处理上存在显著权衡，RNN可直接建模长期依赖，但序列化计算难以并行、训练时间成本高。1D-CNN并行性好、对短期依赖敏感，但对超长依赖建模能力有限。Transformer使用自注意力捕获多尺度时序依赖、并行效率高，但对小规模数据易过拟合，并且内存开销大。在长序列、小样本场景下，RNN或混合方案更稳健；在大规模、高频场景下，Transformer架构往往具备更好性能。

3.3 演化式建模范式（the evolutionary modeling paradigm）

演化式建模是迄今最高级、最具挑战性的范式。它的哲学思想是，在一个时变系统中，不仅节点状态在演化，描述节点间关系的依赖结构本身也在演化。因此，模型的目标不仅是建模时空依赖，更是建模依赖结构本身的动态演化过程。

随着多变量时间序列分类（MTSC）任务的复杂度不断增加，传统的图神经网络（GNN）方法逐渐难以满足高维、大规模数据的建模需求。为了解决这一问题，研究人员提出了层次化图神经网络（hierarchical GNN）（Xing等，2021）和自监督图神经网络（self-supervised GNN）（Xie等，2023）等高级GNN模型。这些方法能够在提高模型效率的同时，增强其对复杂时空依赖关系的建模能力。对于分类任务，演化式GNN通过聚合整个演化过程中的图快照特征，能够捕捉系统状态随时间的动态轨迹（如从正常运行模式逐渐演变为故障模式的过程），从而依据这种演化路径特征进行类别判定。如TodyNet（Liu等，2024）引入了动态图构建机制以适应MTSC中潜在的动态时空关联；Gao等人（2025）则利用自监督时空图互信息最大化，解决了网络安全分类中的类别不平衡问题。

上述3种范式在特定场景下仍存在局限性：解耦式范式假设时空特征可分，在处理时空耦合紧密的高度非平稳数据时往往表现欠佳；耦合式范式虽然能捕捉联合依赖，但其复杂的参数结构在面对高噪声数据时容易过拟合；演化式范式虽然旨在解决非平稳性，但动态图生成过程对数据质量极其敏感，输入数据中的噪声极易导致生成错误的拓扑结构，进而误导下游分类器。

3.3.1 层次化图神经网络

层次化图神经网络旨在通过引入层次化结构处理大规模图数据和多尺度依赖问题。传统的GNN通常处理单一层次的图结构，而层次化GNN则通过多层次图卷积来处理图的不同尺度，从而增强对复杂时空依赖的建模能力。

层次化GNN的基本思想是通过多层次图结构处理数据中的不同尺度特征。具体而言，层次化GNN通过图池化（graph pooling）操作将高维图数据压缩成低维的图表示，并将该过程递归地应用于多个层次，从而有效提取全局信息。

层次化GNN通过图卷积操作与邻居节点的特征进行信息传递。为了处理大规模图数据，层次化GNN通过图池化操作将图的尺寸逐层压缩，减少计算复杂度。图池化操作的目标是通过选择或聚合邻居节点的特征生成一个更小的图表示。常见的图池化方法包括平均池化、最大池化和加权池化等。在多层次图结构中，节点在每一层都会聚合来自邻居节点的信息，并通过池化操作逐步压缩图数据。通过这种方式，层次化GNN能够在不同尺度上捕捉到全局信息和局部信息。在经过图卷积和池化操作后，最后一层的节点特征会通过全连接层（FC）或其他分类头进一步处理，输出最终的分类结果或回归预测。

通过层次化结构，能够捕捉数据中不同尺度的依赖关系，提升了模型的表达能力。层次化GNN能够从多个层次提取全局信息，提升了模型对长时序依赖的建模能力。图池化操作可以有效减少计算复杂度，尤其是在大规模图数据上，能够加速训练过程。

3.3.2 自监督图神经网络

自监督任务通常设计为一个对比学习任务，目的是让模型在没有标签的情况下学习到节点的良好表示。一种常见的方法是基于对比学习的图嵌入（graph embedding）任务。例如，图节点掩蔽任务可以将图中的部分节点随机掩蔽，然后要求模型预测被掩蔽节点的特征。这种任务能够引导模型学习到节点间的潜在结构关系。

在自监督GNN中，首先，需要设计一个自监督任务，引导模型通过未标注的数据学习节点的表示。常见的自监督任务包括节点掩蔽任务（node masking）和对比学习任务（contrastive learning）。然后，进行与

传统GNN模型类似的图卷积操作。在自监督学习过程中，模型需要根据自监督任务的要求进行节点特征的预测。经过图卷积和自监督任务的训练，节点的特征表示逐步被更新。每一层的节点特征都会被传递到下一层，通过多层的图卷积操作，不断聚合信息，增强节点对图结构中时空依赖关系的理解。最后，模型通过全连接层（FC）或其他分类头对节点的表示进行处理，输出最终的分类结果或回归值。

自监督GNN的目标是最小化自监督任务的损失函数。在节点掩蔽任务中，模型需要预测被掩蔽节点的特征，损失函数通常是重建误差，具体为

$$L_{\text{self-supervised}} = \sum_{i} \left\| \boldsymbol{H}_i^{(t)} - \hat{\boldsymbol{H}}_i^{(t)} \right\|^2$$

式中，$\boldsymbol{H}_i^{(t)}$是节点i在时刻t的特征表示，$\hat{\boldsymbol{H}}_i^{(t)}$是通过自监督任务生成的节点特征的预测值，表示模型对被掩蔽节点的预测。通过这种方式，自监督GNN能够有效地从图数据中学习到节点的表示，从而提高模型在没有标签数据时的学习能力。自监督图神经网络的优点在于自监督学习不依赖于大量标注数据，且能够在未标注数据上进行预训练。自监督学习任务通常能够帮助模型在学习过程中学习到更加通用的表示，从而提升模型的泛化能力。

本节综述的主流GNN架构如表3所示。

4 关键应用领域与案例研究

图神经网络（GNN）近年来在多变量时间序列分类（MTSC）中取得了显著的进展，其在多种应用领域中的应用展现了强大的时空建模能力。通过结合时序数据的空间和时间依赖，GNN能够处理复杂的时空数据，广泛应用于智能交通、医疗监测、工业控制和金融分析等领域。下面详细讨论这些领域中的具体应用，并通过案例研究展示GNN在实际中的应用效果。

4.1 交通系统

4.1.1 应用背景

现代城市交通系统依赖于大量分布式传感器（如车载GPS、地磁传感器、摄像头、信号灯感应器等）实时收集数据（Massobrio和Nesmachnow，2020），形成高维、多变量的时间序列。这些数据反映了道路交通的速度、流量、密度以及车道占用率等特征，具有显著的时空相关性。对这类数据进行交通状态分类，即在给定某一时间段和区域的多变量观测数据下，判断当前道路属于“畅通”、“缓行”、“拥堵”等状态，对交通控制优化、信号灯调度以及应急响应等城市治理具有重要意义。

由于交通网络天然具有图结构特征（路段 = 节点，道路连接 = 边），并且节点间存在复杂的非欧氏关系，图神经网络在此类任务中表现出远优于传统机器学习模型的潜力。

4.1.2 数据建模、图构建策略与分类流程

每个交通节点（如传感器、路口）都可记录一系列连续的数值变量，例如平均车速$v_i^{(t)}$、流量$q_i^{(t)}$、车道占用率$o_i^{(t)}$、排队长度$l_i^{(t)}$，这些变量构成节点i在时间

表3 主流GNN架构

Table 3 Mainstream GNN architectures

模型	特点	复杂度	应用领域
GCN	基于拉普拉斯算子进行图卷积	O (N^2)	节点分类、链路预测
GAT	引入注意力机制，区分节点重要性	O (N^2)	社交网络、推荐系统
GraphSAGE	采用采样策略进行邻居特征聚合	O (kN)	大规模网络、工业监测
STGCN	空间—时间联合建模	O (N^2T)	交通预测、金融分析
DCRNN	基于扩散卷积捕捉时空依赖	O (N^2T)	智能交通、时序预测
ASTGCN	基于自注意力机制的时空建模	O (N^2T)	交通、工业、金融市场
MTSGNN	基于自适应图学习和GCN	O (N^2T)	多领域分类
DMGSTCN	基于多图注意力时空卷积	O (kNT^2)	交通、能耗
TodyNet	采用动态图、1D-CNN和自注意力	O (N^2T)	财务分类
Instance-Attention GNN	引入实例级注意力和动态边权	O (N^2T)	流式时序分类

t 的特征向量 $\boldsymbol{x}_i^{(t)} \in \mathbf{R}^d$，多个节点在多个时间步下构成输入张量 $\boldsymbol{X} \in \mathbf{R}^{N \times T \times d}$。

图 $\boldsymbol{G} = (\boldsymbol{V}, \boldsymbol{E})$ 通常基于路网拓扑、距离阈值、历史流量相关性和注意力机制学习的隐式结构图方式构建。分类的典型流程为首先输入表示，将 N 个节点的多变量时间序列输入图神经网络，构成图序列 $\left\{\boldsymbol{X}^{(t)}\right\}_{t=1}^{T}$。时空特征提取，使用ST-GNN模型进行联合建模。分类头设计，输出各个节点或整幅图的分类标签，例如“当前路段属于拥堵状态(类3)”，具体可设计为 softmax 分类或二分类。最后是损失函数，例如使用交叉熵损失：$\mathcal{L} = -\sum_{i=1}^{N}\sum_{c=1}^{C} y_{ic}\log\left(\hat{y}_{ic}\right)$，其中 y_{ic} 是真实标签，$\hat{y}_{ic}$ 是GNN输出的预测概率。

4.1.3　案例研究：交通状态分类

Concurrency Hypothesis（并发假设）(Chen 等，2024)基于图神经网络，通过构建并发先验(concurrency prior，CP)机制增强交通事故的预测能力。该方法通过引入交通流量、速度等多个变量构建图，并在此基础上动态调整图的边权重，学习交通事故的时空关联。CP 假设通过捕捉相邻路段的交通事故并发性，提高了图神经网络在时空数据中的表现。

如图 7 所示，该模型使用了美国多个州的交通事故数据以及路网的交通流量数据。具体数据集包括交通事故发生的时间、地点、天气状况、交通流量和车速等多个特征。模型将 GNN 提取到的路网拥堵传播特征映射到预定义的离散状态空间，利用 softmax 层输出当前时刻路网状态的类别概率，从而辅助交通管理部门进行分级管控。

在实际应用中，融合 CP 的 GNN 相较于传统的 STGCN、DCRNN 等方法在预测交通事故发生的 F1 分数上提升了 3%～13%，并且在 AUC（area under curve）指标上也有显著提高。尤其在处理大规模的路网数据时，CP 机制的引入能够显著提升分类准确度。

4.2　医疗健康监测

4.2.1　应用背景

随着可穿戴设备、生理监测系统（如ICU（intensive care unit）监控）及智能医疗平台的普及，患者心率、血压、体温和血氧饱和度等生理数据可被高频连续采集，形成多变量时间序列(Shi 等，2024)。这些数据揭示了神经、心血管和呼吸等系统的动态交互。通过分析与分类，可实现病人状态的实时评估，如正常/异常判别及疾病阶段识别。这类任务称为病人状态分类，是多变量时间序列分类(MTSC)在医疗监测中的核心应用。然而，由于不同生理信号间存在复杂耦合，传统 LSTM、CNN 模型难以刻画高阶交互结构，而图神经网络(GNN)在此方面具有天然优势。

4.2.2　数据建模、图构建策略与分类流程

每种生理信号（如心率、体温、呼吸频率）可视为一个节点，其时间序列可以表示为变量 i 在时间 t 的

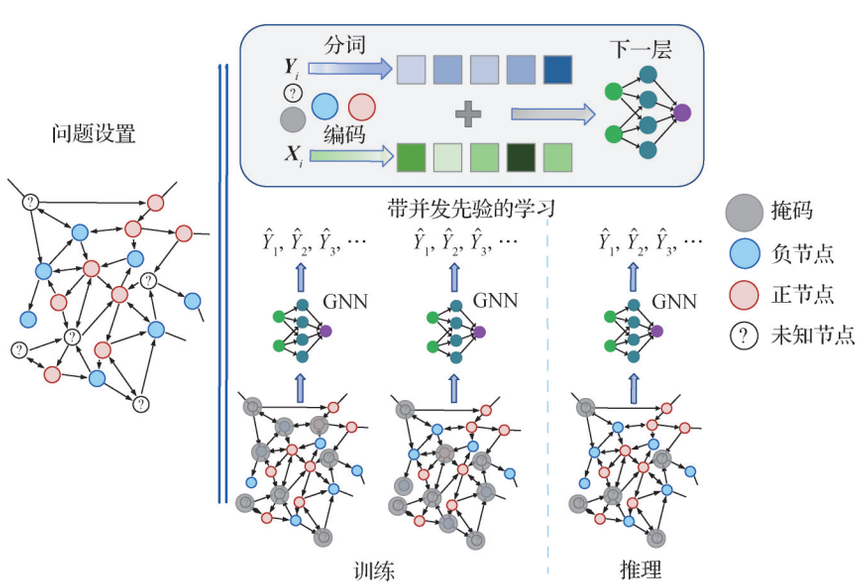


图 7　融合 CP 的 GNN 推理过程

Fig. 7　GNN inference process with CP fusion

观测值 $\boldsymbol{x}_i^{(t)} \in \mathbf{R}$，多变量数据可表示为张量 $\boldsymbol{X} \in \mathbf{R}^{N \times T}$。

由于生理系统之间存在明确的功能性和解剖学连接，可借助生理知识、动态相关性以及图结构学习等途径构建图结构。病人状态分类通常是一个图级分类任务，即模型需基于一段时间序列内所有变量的交互状态，判断当前病人处于哪种健康状态。基于这种关系，分类的典型流程可以表示为首先将 N 个节点的多变量时间序列输入图神经网络，构成图序列 $\left\{\boldsymbol{X}^{(t)}\right\}_{t=1}^{T}$，此时每帧图中节点为生理变量。使用图卷积（GCN、GAT）在每个时间步内建模变量之间的依赖关系，得到每一时间步下的嵌入表示 $\boldsymbol{H}^{(t)} \in \mathbf{R}^{N \times d}$。然后将所有时间步下的图嵌入串联，使用时间模块（如 GRU、1D-CNN、Transformer）提取动态演化特征。最后，对整段图序列特征进行池化或注意力加权后，通过多层感知机（multilayer perceptron，MLP）或 Softmax 分类器输出病人状态类别。

4.2.3　案例研究：病人状态分类

KEDGN（knowledge-empowered dynamic graph network）模型通过引入医学知识（如疾病进程、治疗反应等）增强图神经网络的表示能力（Luo 等，2024）。如图 8 所示，该模型在每个时间步根据采样密度动态调整图结构，并结合基于时序卷积的图卷积网络（GCN）进行多变量时间序列分类。在该应用中，GNN 的核心作用是将不同生理指标（节点）间的隐式关联显式化，分类器依据这些关联特征（如血压下降与心率上升的同步性）区分“高危”与“低危”病患，而非仅仅依赖单个指标的数值阈值。KEDGN 融合了知识图谱和动态图构建，使得模型能够处理不规则采样的医疗数据。

KEDGN 使用了 MIMIC-Ⅲ 和 MIMIC-Ⅳ 数据集，这些数据集包含了大量 ICU 病人的临床信息，包括血压、心率、血氧饱和度和呼吸频率等多个生理信号，并记录了病人各类病史、治疗方案等数据。

在多重耐药风险预测任务中，KEDGN 相比于传统的图神经网络模型（如 ST-GCN）和基于循环神经网络（RNN）的模型，ROC-AUC 提升了 7.1%，模型表现出较好的可解释性，能够有效识别病人健康状态与临床特征之间的关联。

KEDGN 在处理不规则采样的医疗数据时，能够动态适应不同时间点的采样密度，准确预测病人状态。其知识增强的机制使得模型能够结合领域知识，提高预测结果的可靠性和可解释性，对于疾病的早期预警和精准治疗具有重要意义。与之类似，ICU-TGNN 模型（Shi 等，2024）针对 ICU 患者的多模态生理时间序列，通过混合多任务 Transformer 和图神经网络，成功预测了患者的临床转归（如死亡率、住院时长），进一步拓宽了 GNN 在危重症监护中的应用边界。

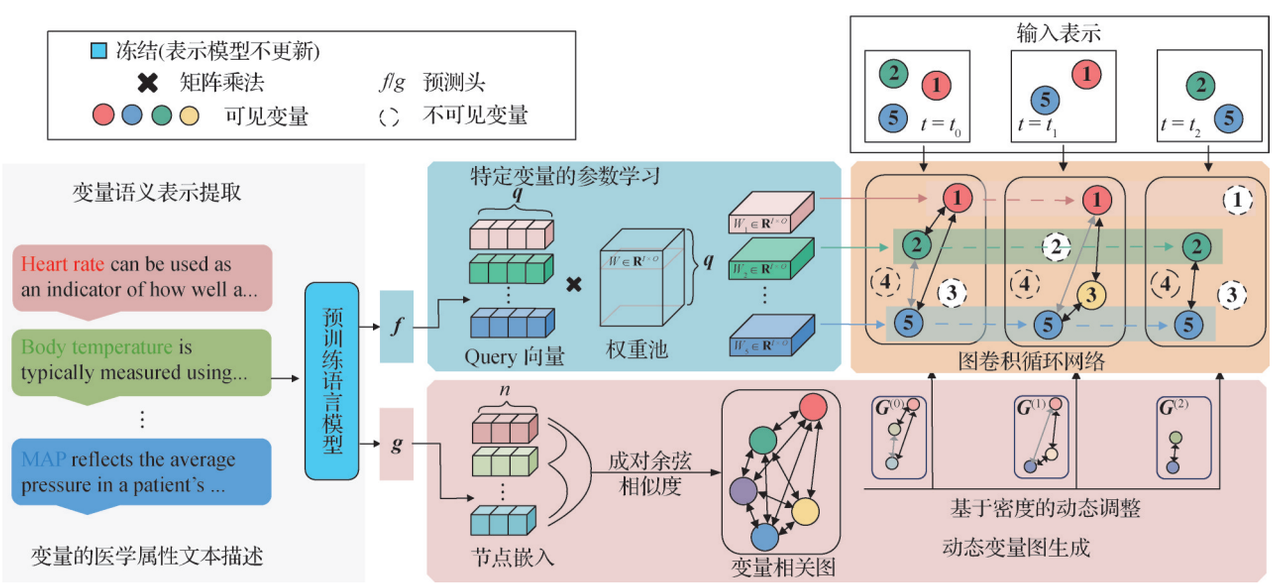


图 8　KEDGN 模型结构

Fig. 8　KEDGN model structure

4.3 工业物联网与制造

4.3.1 应用背景

在现代工业制造系统中，随着物联网（internet of things，IoT）与传感器技术的发展，工厂中大量关键设备（如轴承、电机、数控机床和压缩机等）被部署了高频监测装置（Adi 等，2020）。这些传感器可连续记录设备运行状态，如振动频率、电流、电压、转速和温度等，形成结构化的多变量时间序列数据。基于这些数据，工业系统可以实现故障检测与分类（fault detection and diagnosis，FDD），及时识别潜在问题并分类故障类型（如轴承磨损、电机失稳等），避免代价高昂的停机或事故（Abid 等，2021）。

工业设备中各个部件之间通常存在强耦合关系，故障具有时空扩散性和层级传染性，单一变量的分析无法反映系统行为的全貌。因此，将设备及其传感器网络建模为图结构，并使用图神经网络对其进行故障分类，成为近年来的研究热点。

4.3.2 数据建模、图构建策略与分类流程

设系统中有 N 个传感器，每个传感器记录 d 维时序变量（如振动三轴、电压和电流）。可构建张量表示 $\boldsymbol{X} \in \mathbf{R}^{N \times T \times d}$，故障类型以标签 $y \in \{1, \cdots, C\}$ 表示，C 是类别数。

工业系统中传感器间的关系通常较为稳定，可采用基于物理结构建图、基于信号相关性建图和基于注意力机制建图的图建模方式。典型工业故障分类流程如下，先进行多变量时间序列建模，将每个时间步的传感器网络视为一帧图 $\boldsymbol{G}_t = (\boldsymbol{V}, \boldsymbol{E}, \boldsymbol{X}_t)$，构成图序列或将整个时间窗口内构成单一图进行图序列建模。构建时空模块，使用 ST-GNN 架构同时学习时空联合表示。然后将整个图或图序列经过图池化、注意力加权后进行全连接分类，输出预测的故障类别标签。最后使用交叉熵或 focal loss 应对类别不均衡问题，结合置信度正则化或边注意力以增强可解释性。

4.3.3 案例研究：物联网入侵多层分类

Ahanger 等人（2025）提出了一种基于图注意力网络（GAT）的物联网入侵多层分类系统。该方法通过将网络流量数据转换为图结构（节点 = 流量记录，边 = 相似性关系），构建了一个专门用于二元分类（正常流量 vs 攻击流量）和多元分类（DoS/Probe/R2L/U2R 攻击类型）的图神经网络框架。该框架利用 GAT 聚合网络流量间的相似性模式，当攻击发生时，流量节点间的连接权重分布会呈现出与正常流量截然不同的聚集特征，分类头正是依据这种拓扑结构特征的差异识别具体的攻击类型。

实验采用双层注意力机制，通过动态权重分配聚焦关键邻居节点。二元分类用 2 神经元输出层，多元分类用 5 神经元输出层（对应 4 类攻击 + 正常）。最后使用稀疏分类交叉熵损失函数优化多类别判别能力。

在数据处理方面，对 NSL-KDD 数据集的分类特征（协议类型/服务/标志）进行 one-hot 编码，通过欧几里得距离构建 250 654 条加权边，建立分类所需的图结构关系。采用 80%/20% 的训练/测试划分，对分类模型的泛化能力进行验证。

在工业设备维护领域，Jiang 等人（2022）针对轴承故障诊断任务，提出了一种基于多头图注意力网络的方法。该方法将振动信号转换为图结构，通过注意力机制聚焦关键频段特征，有效识别了不同工况下的设备故障类型，展示了 GNN 在工业故障声学信号分析中的实用价值。

4.4 金融市场

4.4.1 应用背景

在金融市场中，市场状态的变化往往受到多种因素的共同影响，包括股票价格、交易量、利率以及宏观经济指标等。这些因素构成了一种多变量时间序列，其中变量之间的相互作用和依赖关系不仅限于时间维度，还表现出复杂的空间依赖性，例如跨市场板块的联动效应、全球市场的同步波动等。传统的时间序列分析方法，如 ARIMA、GARCH（generalized autoregressive conditional heteroskedasticity）（Kour 和 Arora，2020）以及基于 LSTM 和 RNN 的深度学习模型，通常关注于单一变量的时间演化，无法充分捕捉金融市场中跨变量的交互影响。然而，金融市场中的资产价格往往会受到多个因素的共同影响，例如宏观经济政策的变动、行业发展趋势以及国际局势等，这些交互影响在传统模型中难以建模。GNN 可以将不同的金融时间序列视为节点，通过边表示它们之间的关联关系，并通过消息传递机制建模节点之间的时空交互，进而提高市场状态分类的准确性。

4.4.2 数据建模、图构建策略与分类流程

在市场状态分类任务中，通常使用历史金融数

据构建多变量时间序列，包括股票价格（开盘价、收盘价、最高价和最低价）、交易量（日交易量、买卖盘数据）、宏观经济指标（GDP、失业率和消费者物价指数）、货币汇率（美元指数、日元汇率）、期货和商品价格（黄金、原油等）。这些数据以时间为步长构建多维矩阵 $\boldsymbol{X}=\{\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n\}$，假设 T 表示时间步长，F 表示特征数量，则 $\boldsymbol{X}_i\in\mathbf{R}^{T\times F}$。为了更好地捕捉金融市场各因素之间的关联性，常用基于相关性的图构建、基于动态图的构建以及基于注意力机制的图建模的策略进行图的构建。

4.4.3　案例研究：基于GraphCNNpred的股票市场状态分类

GraphCNNpred（Jin，2025）模型结合了图卷积网络（GCN）和一维卷积神经网络（CNN）来进行股票市场的状态分类。该方法通过可见度图（visibility graph）将多个资产的历史数据转化为图结构，并利用CNN提取局部时间特征，通过GCN捕捉资产之间的相互影响。最终，该模型对次日股市的涨跌进行二分类预测。

GraphCNNpred在多个股票指数数据集上的F1-Score提升了4%～15%，并且在模拟交易中，模型的Sharpe比率超过3，表现出了优异的投资收益率。这表明，GraphCNNpred能够有效捕捉资产间的联动关系，并提高市场状态预测的准确性。GraphCNNpred能够综合多个资产的数据，准确捕捉它们之间的关联性，提升市场状态分类的精度。相比于传统的基于单一资产的预测方法，该方法能更好地适应复杂的市场环境，具有较高的投资价值和风险控制能力。

本节综述的4类方法如表4所示，可以发现各方法在训练阶段具有显著的共性。首先是所述研究普遍选择领域公认的数据集，如交通领域广泛使用PEMS-SF和METR-LA；医疗领域多用PhysioNet与MIMIC系列；金融领域则普遍采用公开股票指数数据（如NASDAQ、S&P 500）或市场预测基准。这一共性便于各方法之间的公平比较与验证。其次，多数方法都采用标准化或归一化方式进行特征处理，同时广泛使用滑动窗口方式生成训练样本。研究人员普遍采用时间敏感的划分方法以防止数据泄露，这也是时间序列任务的通行惯例。还有大部分研究均采用基于Adam优化器的训练方式，常涉及学习率、节点嵌入维度、注意力头数以及dropout比例等超参数的调整。此外，大多仅展示最佳性能配置，建议未来研究可公开完整超参数搜索过程，以提升结果的复现性与可比较性。最后，所综述方法的性能评估基本上围绕准确率（Accuracy）、F1-score以及AUC-ROC等指标展开，尤其在类别不平衡的数据集中，多采用Macro-F1以避免指标偏倚。此外，建议未来研究增加实验多次重复的结果，以展现模型性能的稳定性。

综合来看，多变量时间序列分类（MTSC）在交通、医疗、金融和工业物联网等领域涌现出一系列基于图神经网络（GNN）的代表性模型，各具结构优势与任务适配性。交通预测中的DST-SGNN通过在Stiefel流形（Sun等，2024）上定义图谱卷积与图傅里叶变换，高效建模复杂时空相关，并以较低开销在多项基准数据集上超越当时最新模型性能。医疗监测领域的CST-GNN（causal spatiotemporal graph neural network）（Bi等，2025）将流行病学SIR（susceptible-infected-recorered）模型与GNN时空学习融合，利用静态与动态人群图叠加刻画疾病传播过程，在提升

表4　不同领域的案例研究汇总

Table 4　Summary of case studies from different fields

任务场景	数据集	模型	Accuracy	F1-Score	AUC
交通系统状态分类	美国49个州的全国道路网络与事故记录数据	GNN + 并发先验（CP）	–	+3%～13%提升	+1.3%～9%提升
ICU病人状态预测	MIMIC-III/IV数据集	KEDGN	0.82 ROC-AUC	+7.1%提升	–
物联网入侵多层分类	NSL-KDD数据集	GAT	98%（二分类）、99.2%（多分类）	99%	–
金融市场状态分类	S&P 500、NASDAQ、DJI、NYSE等股票指数数据	GraphCNNpred	–	+4%～15%提升	–

注：“–”表示原始参考文献中未报告。

预测精度的同时显著增强因果可解释性。金融分析方面，LED-GNN（LLM-enhanced dynamic graph neural network）（Xu 等，2025）借助大语言模型从财经新闻中抽取股票关系，结合历史交易数据构建动态图，对股票交易量预测效果优于现有方法。工业物联网中，多尺度动态图建模（Foumani 等，2024）通过滑动时间窗与自适应图结构捕捉传感器数据的时空演化，在设备异常检测任务上显著提升准确率。总体而言，交通与工业系统因物理拓扑稳定，更强调动态邻接更新与多尺度特征提取以平衡精度与部署代价；医疗与金融场景则因时空关联非平稳、异质性强，需更高的图结构自适应与演化能力，并结合领域知识确保可解释性与可靠性（如 CST-GNN 融入传染病动力学机理）。在 MTSC 任务中，评估协议与基准选择直接影响性能衡量与对比的可靠性，除传统分类指标外，还需关注图结构学习的鲁棒性与时间建模效果。

5 评估与基准

5.1 常用数据集

在基于 GNN 的多变量时间序列分类任务中，表 5 展示的公开数据集广泛用于评估和对比。这些数据集覆盖了工业、金融、医疗和交通等多个领域，并且能够反映时间序列的多尺度、多依赖特性，非常适合 GNN 在复杂系统中的建模与评估。

表 5 部分公开的多变量时间序列分类数据集

Table 5 Some publicly available multivariate time series classification datasets

数据集名称	描述	变量数量	样本数量	任务类别	来源
UEA Archive	多领域时间序列分类数据集	2～1 345	1 000+	分类	UEA
PhysioNet	ICU 监测数据，生理信号	30+	40 000+	分类/预测	MIT
PEMS-SF	交通流量监测	963	15 000+	分类/预测	Caltrans
Tennessee Eastman Process (TEP)	工业流程监测	52	5 000+	故障检测	Eastman Chemical
MIMIC-III	ICU 患者生理数据	40+	60 000+	分类/预测	MIT
HAR Dataset	人体动作识别数据	9	10 299	分类	UCI

5.2 评估指标

在基于 GNN 的时序分类与预测任务中，需同时评估图结构学习质量与预测性能。图结构方面，可采用节点分类准确率、边权重重构误差、聚类系数与模块化度。若节点存在类别标签，可直接计算分类准确率；边权重重构误差可通过比较预测邻接矩阵 $\boldsymbol{A}_{\text{pred}}$ 与真实依赖图 $\boldsymbol{A}_{\text{true}}$ 进行衡量，具体为

$$L_{\text{adj}} = \left| \boldsymbol{A}_{\text{pred}} - \boldsymbol{A}_{\text{true}} \right|^2 \tag{23}$$

聚类系数与模块化度则用于评估图是否能清晰刻画簇结构，对时序共变现象尤为重要。在预测性能方面，常用指标包括均方误差（mean square error，MSE）、平均绝对误差（mean absolute error，MAE）及动态时间规整（DTW），具体为

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i\right)^2 \tag{24}$$

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\boldsymbol{y}_i - \hat{\boldsymbol{y}}_i\right| \tag{25}$$

DTW 则通过对齐预测序列与真实序列，评估其形态相似性。上述指标共同构成了对图时序模型的结构与性能的综合评价体系。

6 开放性挑战与未来研究方向

尽管基于图神经网络的多变量时间序列分类在许多领域取得了显著进展，但仍存在许多挑战有待解决，同时也孕育着未来研究的巨大潜力。根据现有的研究现状与发展趋势，本节从 5 个方面详细讨论 GNN 在 MTSC 中的开放性挑战与未来研究方向。

6.1 面向大规模与真实场景 MTS 的可扩展与鲁棒方法

随着物联网 IoT（Fabrício 等，2020）、工业监测（Zhao 等，2024）和金融数据采集（Zhang 等，2023）的普及，MTS 数据规模呈指数级增长，伴随噪声、缺失与不规则采样，给 GNN 的计算与存储带来挑战。传统 GCN 复杂度为 $O(N^2)$，在成千上万节点的工业或

交通网络中易导致显存溢出与计算瓶颈。为提升可扩展性，研究提出邻居采样 GraphSAGE（Hamilton 等，2017）、重要性采样 FastGCN（Chen 等，2018）、分布式框架 DistDGL（Zheng 等，2020）、GraphX（Gonzalez 等，2014）及多尺度建模等策略。

真实数据中的缺失与噪声亦威胁模型鲁棒性。常见方法包括动态邻接修正、融合多模态信息，以及在预处理阶段利用 GRU-D（gated recurrent unit with decay）（Che 等，2018）、BRITS（bidirectional recurrent imputation for time series）（Cao 等，2018）进行插值与填补；同时设计鲁棒损失或图滤波以降低异常影响。然而，对图时序模型的理论理解仍不足，表达能力与收敛性缺乏严格分析。未来需在算法与理论两方面推进：发展稀疏高效的消息传递机制，增强对数据缺陷的容忍度，并建立严谨的理论框架，以保障性能与稳定性

6.2 动态结构与因果图学习

工业生产、金融市场与交通网络中的变量关系随时间动态演化，固定拓扑的 GNN 难以适应时变依赖。为此，动态图神经网络（D-GNN）成为重要方向：如 D-GNN（Skarding 等，2021）采用滑动窗口更新邻接矩阵，EvolveGCN（Pareja 等，2020）与 DySAT（Sankar 等，2020）自动学习结构变化；其他方法通过自适应边连接或事件检测动态调整信息传递路径。

除结构自适应外，提升可解释性与因果推理亦很关键。现有模型多为“黑箱”，研究者利用 GNN Explainer 与 PGExplainer 为边和节点分配重要性分数，揭示影响决策的子图结构（Luo 等，2020）。但相关性解释不足以揭示因果关系，因果图学习（Shanmugam 等，2015）正逐渐融入 GNN，通过扰动节点特征观察结果变化以识别因果依赖。未来方向是将因果发现与解释结合动态图建模，发展可边训练边解释的自适应 GNN，并设计因果性正则化约束，以提升透明度与泛化能力，推动其在高风险实时场景中的应用。

6.3 自监督、无监督与多视角表示学习

在多变量时间序列分类中，高质量标签获取成本高，因此充分利用无标签数据进行表征学习成为关键课题。自监督学习通过预文本身任务在无需人工标签的情况下学习特征，已在图领域展现潜力。典型方法包括对比学习、图重构和时序预测：如 Deep Graph Infomax（Gao 等，2025）与 GraphCL（You 等，2020）通过对比学习捕捉结构信息，GraphCL 利用数据增强生成不同视图并最大化相似度，策略可推广至时序数据以学习跨时间依赖。其他任务包括遮蔽预测与顺序恢复，以逼近真实时序关系。

多视角表示学习则从不同传感器子集或时间尺度构建多视图，通过跨视角对比或协同训练融合信息，提升对多模态、多尺度特征的刻画能力（Hassani 和 Khasahmadi，2020）。未来，自监督与无监督方法有望结合时序数据增强与跨模态信号设计更贴合的预训练任务，并开发表示质量评估指标，显著降低对人工标签的依赖，为跨领域 MTS 分类提供更通用稳健的表示基础。

6.4 隐私保护与联邦图学习

在医疗、金融和工业等领域，MTS 数据常涉及敏感信息，集中式训练需汇总数据至云端，存在隐私与合规风险。联邦学习（Wen 等，2023）提供“数据不出本地”的范式：各方在本地训练并上传参数，由服务器聚合形成全局模型，从而共享智能而无需共享原始数据，契合分散式多节点时序场景。然而，联邦图学习仍面临挑战：1）多源数据异构性显著，不同机构的特征空间与采样方式差异较大；2）异步更新导致参数延迟，影响收敛与性能；3）梯度更新可能泄露隐私（如差分攻击、梯度反演）。

针对这些问题，研究提出多种改进：在异构数据方面，可通过图对齐技术统一结构，或采用局部模型仅聚合共享层参数；在鲁棒性方面，可构建分布式 GNN 架构，利用 FedAvg 聚合中间表示以减少开销并保护分布细节；在隐私保护方面，可引入差分隐私（differential privacy，DP），在上传前添加噪声，或采用同态加密，在密文状态下完成聚合。未来，联邦图学习需结合跨域迁移与知识蒸馏，在保障隐私的同时提升泛化性能，并建立标准评测框架与安全协议，以在隐私、效率与精度间取得平衡，推动隐私保护的 GNN 在 MTS 分析中的应用。

6.5 领域知识融合与跨领域迁移

尽管数据驱动方法能自动提取特征，领域先验知识在多变量时间序列建模中仍不可或缺。工业物理定律、医疗生理机制、交通拓扑与规则等均蕴含宝贵信息，融合进 GNN 有助于提升可解释性与泛化性。知识注入面临三大挑战：1）难以形式化为图结构；2）先验权重设定困难；3）知识需随政策与工艺动态更新。现有探索主要通过构建独立先验图并

与数据驱动图并行消息传递，或在传递中加入规则约束，如交通预测中结合道路拓扑与传感器相关性，化工建模中依据守恒定律调整权重。这些方法在初步实验中展现了增强模型稳健性的潜力。

跨领域迁移学习是提升模型泛化能力的重要方向。实际应用中，不同领域或场景的 MTS 数据分布差异显著，源域模型直接应用于目标域往往效果不佳。迁移学习（transfer learning，TL）与领域自适应（domain adaptation，DA）通过跨域知识传递，使模型在小样本甚至无标签条件下仍具竞争力（Sun 等，2025a，b）。然而，源域与目标域的特征空间和图结构常不一致，且目标域缺乏标注限制了直接调优。为此，研究者提出多种策略：如在训练中引入对抗性损失以对齐分布，利用少量目标域数据微调源模型，或采用元学习初始化以快速适应新任务；在异常检测等特殊场景中，还尝试通过源域引入类似异常信号以弥补目标域缺失。未来，领域知识注入与迁移学习有望结合：以通用先验知识作为跨域桥梁，辅助模型理解新变量，并开发面向图结构的迁移算法，使模型在节点关系动态变化下保持性能。随着更多开放数据与基准的出现，跨领域评估将成为衡量模型鲁棒性的重要标准。

7 结　语

基于 GNN 的多变量时间序列分类方法，打破了传统时间序列单维度独立建模的局限，通过图结构显式表达时间序列的空间交互特性，在交通预测、工业监测、医疗诊断以及金融市场分析等多个领域展示出广泛的应用潜力。然而，现阶段的研究仍然存在诸多挑战，包括大规模数据的可扩展性与计算效率、缺失数据与不规则采样的鲁棒性、动态图的自适应学习与因果关系建模、多领域的迁移学习与自监督表示、隐私保护与跨机构联邦学习的推广。

未来的研究可以从以下几方面推进，更高效的稀疏化消息传递机制、引入联邦学习与安全计算、领域自适应与迁移学习以及可解释的动态图建模。

随着新技术的发展与数据资源的进一步开放，基于 GNN 的 MTSC 研究将有望解决更复杂的工业、交通、金融和医疗问题，成为时序数据智能预测的重要支撑。

参考文献（References）

Abid A，Khan M T and Iqbal J. 2021. A review on fault detection and diagnosis techniques：basics and beyond. Artificial Intelligence Review，54(5)：3639-3664［DOI：10.1007/s10462-020-09934-2］

Adi E，Anwar A，Baig Z and Zeadally S. 2020. Machine learning and data analytics for the IoT. Neural Computing and Applications，32(20)：16205-16233［DOI：10.1007/s00521-020-04874-y］

Ahanger A S，Khan S M，Masoodi F and Salau A O. 2025. Advanced intrusion detection in internet of things using graph attention networks. Scientific Reports，15(1)：#9831［DOI：10.1038/s41598-025-94624-8］

Bai T，Zhang Y J，Wu B and Nie J Y. 2020. Temporal graph neural networks for social recommendation//Proceedings of 2020 IEEE International Conference on Big Data. Atlanta，USA：IEEE：898-903［DOI：10.1109/BigData50022.2020.9378444］

Baydogan M G，Runger G and Tuv E. 2013. A bag-of-features framework to classify time series. IEEE Transactions on Pattern Analysis and Machine Intelligence，35(11)：2796-2802［DOI：10.1109/TPAMI.2013.72］

Berndt D J and Clifford J. 1994. Using dynamic time warping to find patterns in time series//Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining. Seattle，USA：AAAI Press：359-370

Bi X，Jin Q H，Song M L，Yao X，Zhao X G，Yuan Y，et al. 2025. Spatiotemporal learning with decoupled causal attention for multivariate time series. IEEE Transactions on Big Data，11(4)：1589-1599［DOI：10.1109/TBDATA.2024.3499312］

Blázquez-García A，Conde A，Mori U and Lozano J A. 2022. A review on outlier/anomaly detection in time series data. ACM Computing Surveys（CSUR），54(3)：#56［DOI：10.1145/3444690］

Bui K H N，Cho J and Yi H. 2022. Spatial-temporal graph neural network for traffic forecasting：an overview and open research issues. Applied Intelligence，52(3)：2763-2774［DOI：10.1007/s10489-021-02587-w］

Cao W，Wang D，Li J，Zhou H，Li L and Li Y. 2018. BRITS：bidirectional recurrent imputation for time series//Proceedings of the 31st International Conference on Neural Information Processing Systems. Montréal，Canada：Curran Associates，Inc.：6775-6785

Caterini A L and Chang D E. 2018. Recurrent neural networks//Caterini A L and Chang D E. Deep Neural Networks in a Mathematical Framework. Cham，Germany：Springer：59-79［DOI：10.1007/978-3-319-75304-1_5］

Che Z P，Purushotham S，Cho K，Sontag D and Liu Y. 2018. Recurrent neural networks for multivariate time series with missing values. Scientific Reports，8(1)：#6085［DOI：10.1038/s41598-018-24271-9］

Chen J，Ma T F and Xiao C. 2018. FastGCN：fast learning with graph convolutional networks via importance sampling//Proceedings of the 6th International Conference on Learning Representations. Vancou-

ver, Canada: OpenReview.net

Chen X W, Boroujeni S P H, Shu X, Li H Y and Razi A. 2024. Enhancing graph neural networks in large-scale traffic incident analysis with concurrency hypothesis//Proceedings of the 32nd ACM International Conference on Advances in Geographic Information Systems. Atlanta, USA: Association for Computing Machinery: 196-207 [DOI: 10.1145/3678717.3691256]

Corradini F, Gerosa F, Gori M, Lucheroni C, Piangerelli M and Zannotti M. 2026. A systematic literature review of spatio-temporal graph neural network models for time series forecasting and classification. Neural Networks, 195: #108269 [DOI: 10.1016/j.neunet.2025.108269]

Fabrício M A, Herman Behrens F and Bianchini D. 2020. Monitoring of industrial electrical equipment using IoT. IEEE Latin America Transactions, 18 (8): 1425-1432 [DOI: 10.1109/TLA.2020.9111678]

Fan W Q, Ma Y, Li Q, He Y, Zhao E, Tang J L, et al. 2019. Graph neural networks for social recommendation//Proceedings of 2019 World Wide Web Conference. San Francisco, USA: Association for Computing Machinery: 417-426 [DOI: 10.1145/3308558.3313488]

Fang Z, Long Q, Song G and Xie K. 2021. Spatial-temporal graph ODE networks for traffic flow forecasting//Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining (KDD'21). [s.l.]: Association for Computing Machinery: 364-373

Foumani N M, Miller L, Tan C W, Webb G I, Forestier G and Salehi M. 2024. Deep learning for time series classification and extrinsic regression: a current survey. ACM Computing Surveys, 56(9): #217 [DOI: 10.1145/3649448]

Gao C, Wang X, He X N and Li Y. 2022. Graph neural networks for recommender system//Proceedings of the 15th ACM International Conference on Web Search and Data Mining. [s.l.]: Association for Computing Machinery: 1623-1625 [DOI: 10.1145/3488560.3501396]

Gao P, Zhang H T, Wang M, Yang W Y, Wei X S, Lyu Z, et al. 2025. Deep temporal graph infomax for imbalanced insider threat detection. Journal of Computer Information Systems, 65(1): 108-118 [DOI: 10.1080/08874417.2023.2267510]

Gonzalez J E, Xin R S, Dave A, Crankshaw D, Franklin M J and Stoica I. 2014. GraphX: graph processing in a distributed dataflow framework//Proceedings of the 11th USENIX conference on Operating Systems Design and Implementation. Broomfield, USA: USENIX Association: 599-613

Gu J X, Wang Z H, Kuen J, Ma L Y, Shahroudy A, Shuai B, et al. 2018. Recent advances in convolutional neural networks. Pattern Recognition, 77: 354-377 [DOI: 10.1016/j.patcog.2017.10.013]

Gupta A, Gupta H P, Biswas B and Dutta T. 2020. An early classification approach for multivariate time series of on-vehicle sensors in transportation. IEEE Transactions on Intelligent Transportation Systems, 21(12): 5316-5327 [DOI: 10.1109/TITS.2019.2957325]

Hamilton W, Ying Z and Leskovec J. 2017. Inductive representation learning on large graphs//Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, USA: Curran Associates Inc.: 1024-1034

Han K, Koay A M Y, Ko R K L, Chen W T and Xu M. 2025. Adapting to the stream: an instance-attention GNN method for irregular multivariate time series data. Frontiers of Computer Science, 19(8): #198340 [DOI: 10.1007/s11704-024-40449-z]

Hassani K and Khasahmadi A H K. 2020. Contrastive multi-view representation learning on graphs//Proceedings of the 37th International Conference on Machine Learning. [s.l.]: PMLR: 4116-4126

Ho S L and Xie M. 1998. The use of ARIMA models for reliability forecasting and analysis. Computers and Industrial Engineering, 35(1/2): 213-216 [DOI: 10.1016/S0360-8352(98)00066-7]

Hochreiter S and Schmidhuber J. 1997. Long short-term memory. Neural Computation, 9(8): 1735-1780 [DOI: 10.1162/neco.1997.9.8.1735]

Huang X H, Chen N J, Deng Z Y and Huang S Q. 2024. Multivariate time series anomaly detection via dynamic graph attention network and Informer. Applied Intelligence, 54(17): 7636-7658 [DOI: 10.1007/s10489-024-05575-y]

Ismail Fawaz H, Forestier G, Weber J, Idoumghar L and Muller P A. 2019. Deep learning for time series classification: a review. Data Mining and Knowledge Discovery, 33(4): 917-963 [DOI: 10.1007/s10618-019-00619-1]

Jiang L, Li X J, Wu L and Li Y B. 2022. Bearing fault diagnosis method based on a multi-head graph attention network. Measurement Science and Technology, 33(7): #075012 [DOI: 10.1088/1361-6501/ac56f1]

Jin G Y, Liang Y X, Fang Y C, Shao Z Z, Huang J C, Zhang J B, et al. 2024a. Spatio-temporal graph neural networks for predictive learning in urban computing: a survey. IEEE Transactions on Knowledge and Data Engineering, 36(10): 5388-5408 [DOI: 10.1109/TKDE.2023.3333824]

Jin M, Koh H Y, Wen Q, Zambon D, Alippi C, Webb G I, et al. 2024b. A survey on graph neural networks for time series: forecasting, classification, imputation, and anomaly detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 46(12): 10466-10485 [DOI: 10.1109/TPAMI.2024.3443141]

Jin Y H. 2025. GraphCNNpred: a stock market indices prediction using a graph based deep learning system//Proceedings of the 2nd International Conference on Artificial Intelligence, Systems and Network Security. Mianyang, China: Association for Computing Machinery: 170-178 [DOI: 10.1145/3714334.3714364]

Khodayar M and Wang J H. 2019. Spatio-temporal graph deep neural network for short-term wind speed forecasting. IEEE Transactions on Sustainable Energy, 10(2): 670-681 [DOI: 10.1109/TSTE.2018.2844102]

Kour V P and Arora S. 2020. Recent developments of the internet of things in agriculture: a survey. IEEE Access, 8: 129924-129957 [DOI: 10.1109/ACCESS.2020.3009298]

Lea C, Vidal R, Reiter A and Hager G D. 2016. Temporal convolutional networks: a unified approach to action segmentation//Proceedings of 2016 Computer Vision — ECCV Workshops. Amsterdam, the Netherlands: Springer: 47-54 [DOI: 10.1007/978-3-319-49409-

8_7］

Li Y G, Yu R, Shahabi C and Liu Y. 2018. Diffusion convolutional recurrent neural network: data-driven traffic forecasting［EB/OL］.［2025-11-01］. https://arxiv.org/pdf/1707.01926.pdf

Lim B and Zohren S. 2021. Time-series forecasting with deep learning: a survey. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 379(2194): #20200209［DOI: 10.1098/rsta.2020.0209］

Liu C L, Hsaio W H and Tu Y C. 2019. Time series classification with multivariate convolutional neural network. IEEE Transactions on Industrial Electronics, 66(6): 4788-4797［DOI: 10.1109/TIE.2018.2864702］

Liu H Y, Yang D H, Liu X Z, Chen X L, Liang Z Y, Wang H Z, et al. 2024. TodyNet: temporal dynamic graph neural network for multivariate time series classification. Information Sciences, 677: #120914［DOI: 10.1016/j.ins.2024.120914］

Luo D S, Cheng W, Xu D K, Yu W C, Zong B, Chen H F, et al. 2020. Parameterized explainer for graph neural network//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: #1646

Luo Y C, Liu Z, Wang L H, Zheng J H, Wu B Q and Ma Q L. 2024. Knowledge-empowered dynamic graph network for irregularly sampled medical time series//Proceedings of the 38th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: #2144［DOI: 10.522021079017-2144］

Lyu H, Sha N, Qin S, Yan M, Xie Y and Wang R. 2019. Advances in neural information processing systems. Advances in neural information processing systems, #32

Massobrio R and Nesmachnow S. 2020. Urban mobility data analysis for public transportation systems: a case study in Montevideo, Uruguay. Applied Sciences, 10(16): #5400［DOI: 10.3390/app10165400］

Min S J, Gao Z, Peng J, Wang L, Qin K and Fang B. 2021. STGSN — A spatial-temporal graph neural network framework for time-evolving social networks. Knowledge-Based Systems, 214: #106746［DOI: 10.1016/j.knosys.2021.106746］

Nargesian F, Samulowitz H, Khurana U, Khalil E B and Turaga D S. 2017. Learning feature engineering for classification//Proceedings of the 26th International Joint Conference on Artificial Intelligence. Melbourne, Australia: ijcai.org: 2529-2535［DOI: 10.24963/ijcai.2017/352］

Pareja A, Domeniconi G, Chen J, Ma T F, Suzumura T, Kanezashi H, et al. 2020. EvolveGCN: evolving graph convolutional networks for dynamic graphs//Proceedings of the 34th AAAI Conference on Artificial Intelligence. New York, USA: AAAI Press: 5363-5370［DOI: 10.1609/aaai.v34i04.5984］

Qian L P, Zuo Q, Liu H G and Zhu H. 2025. Multivariate time series classification based on spatial-temporal attention dynamic graph neural network. Applied Intelligence, 55(2): #115［DOI: 10.1007/s10489-024-06014-8］

Sankar A, Wu Y H, Gou L, Zhang W and Yang H. 2020. DySAT: deep neural representation learning on dynamic graphs via self-attention networks//Proceedings of the 13th International Conference on Web Search and Data Mining. Houston, USA: Association for Computing Machinery: 519-527［DOI: 10.1145/3336191.3371845］

Selim G E I, El-Din Hemdan E Z Z, Shehata A M and El-Fishawy N A. 2021. Anomaly events classification and detection system in critical industrial internet of things infrastructure using machine learning algorithms. Multimedia Tools and Applications, 80(8): 12619-12640［DOI: 10.1007/s11042-020-10354-1］

Shanmugam K, Kocaoglu M, Dimakis A G and Vishwanath S. 2015. Learning causal graphs with small interventions//Proceedings of the 29th International Conference on Neural Information Processing Systems. Montreal, Canada: MIT Press: 3195-3203

Shi T Y, Xu H W, Ma J and Kong G L. 2024. ICU-TGNN: a hybrid multitask transformer and graph neural network model for predicting clinical outcomes of patients in the ICU//Proceedings of 2024 IEEE International Conference on Systems, Man, and Cybernetics. Kuching, Malaysia: IEEE: 1761-1766［DOI: 10.1109/SMC54092.2024.10831750］

Skarding J, Gabrys B and Musial K. 2021. Foundations and modeling of dynamic networks using dynamic graph neural networks: a survey. IEEE Access, 9: 79143-79168［DOI: 10.1109/ACCESS.2021.3082932］

Song J, Son J, Seo D H, Han K, Kim N and Kim S W. 2022. ST-GAT: a spatio-temporal graph attention network for accurate traffic speed prediction//Proceedings of the 31st ACM International Conference on Information and Knowledge Management. Atlanta, USA: Association for Computing Machinery: 4500-4504［DOI: 10.1145/3511808.3557705］

Sun J, Dong M Y, Lu F, Wang Z, Zhu J, Wang Y T, et al. 2025a. Transformer-driven multi-scale contextual alignment for robust E-nose drift compensation. Knowledge-Based Systems, 327: #114136［DOI: 10.1016/j.knosys.2025.114136］

Sun J, Li S C, Chen Z S, Wang X Z, Zheng H, Chen C W, et al. 2026. Temporal-spatial convolution and adversarial mutual learning for time series domain adaptation. Journal of Software, 37(3): 1340-1356（孙杰，李贯宸，陈仲珊，王小正，郑豪，陈长伟，等. 2026. 时序数据域自适应的时空卷积与对抗互学习. 软件学报，37(3): 1340-1356）［DOI: 10.13328/j.cnki.jos.007430］

Sun J, Sun Z L, Chen Z S, Dong M Y, Wang X Z, Chen C W, et al. 2026a. MSA-LR: enhancing multi-scale temporal dynamics in multivariate time series forecasting with low-rank self-attention. Neural Networks, 194: #108150［DOI: 10.1016/j.neunet.2025.108150］

Sun J, Sun Z L, Chen Z S, Xie Y L, Zhang Y K, Wang Z, et al. 2026b. A dynamic continuous-time attention framework for robust gas sensor drift compensation. Advanced Engineering Informatics, 69: #104043［DOI: 10.1016/j.aei.2025.104043］

Sun J, Sun Z L, Zheng H and Zhao X J. 2024. Stiefel-attentive transformer for refrigerator odor detection. IEEE Transactions on Automation Science and Engineering, 21(4): 7647-7661［DOI: 10.1109/TASE.2023.3346823］

Sun J, Zheng H, Diao W X, Sun Z L, Qi Z D and Wang X Z. 2025b.

Prototype-Optimized unsupervised domain adaptation via dynamic Transformer encoder for sensor drift compensation in electronic nose systems. Expert Systems with Applications, 260: #125444 [DOI: 10.1016/j.eswa.2024.125444]

Veličković P, Cucurull G, Casanova A, Romero A, Liò P and Bengio Y. 2018. Graph attention networks. arXiv: 1710.10903 [DOI: 10.48550/arXiv.1710.10903]

Wang J N, Zhang S, Xiao Y H and Song R. 2022a. A review on graph neural network methods in financial applications. Journal of Data Science, 20(2): 111-134 [DOI: 10.6339/22-JDS1047]

Wang Y, Jing C F, Xu S S and Guo T. 2022b. Attention based spatio-temporal graph attention networks for traffic flow forecasting. Information Sciences, 607: 869-883 [DOI: 10.1016/j.ins.2022.05.127]

Wei W W S. 2019. Multivariate Time Series Analysis and Applications. Hoboken, USA: John Wiley and Sons [DOI: 10.100219781119502951]

Wen J, Zhang Z X, Lan Y, Cui Z H, Cai J H and Zhang W S. 2023. A survey on federated learning: challenges and applications. International Journal of Machine Learning and Cybernetics, 14(2): 513-535 [DOI: 10.1007/s13042-022-01647-y]

Wu Z H, Pan S R, Chen F W, Long G D, Zhang C Q and Yu P S. 2021. A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems, 32(1): 4-24 [DOI: 10.1109/TNNLS.2020.2978386]

Xie Y C, Xu Z, Zhang J T, Wang Z Y and Ji S W. 2023. Self-supervised learning of graph neural networks: a unified review. IEEE Transactions on Pattern Analysis and Machine Intelligence, 45(2): 2412-2429 [DOI: 10.1109/TPAMI.2022.3170559]

Xing Y F, He T, Xiao T J, Wang Y X, Xiong Y J, Xia W, et al. 2021. Learning hierarchical graph neural networks for image clustering// Proceedings of 2021 IEEE/CVF International Conference on Computer Vision. Montreal, Canada: IEEE: 3467-3477 [DOI: 10.1109/ICCV48922.2021.00345]

Xu Z Y, Liu Y, Wang Y C, Bao R H, Harimoto K and Sun X. 2025. Modeling interactions between stocks using LLM-enhanced graphs for volume prediction//Proceedings of the Joint Workshop of the 9th Financial Technology and Natural Language Processing. Abu Dhabi, UAE: Association for Computational Linguistics: 153-163

Ye L X and Keogh E. 2009. Time series shapelets: a new primitive for data mining//Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Paris, France: Association for Computing Machinery: 947-956 [DOI: 10.1145/1557019.1557122]

Ye Z, Kumar Y J, Sing G O, Song F Y and Wang J S. 2022. A comprehensive survey of graph neural networks for knowledge graphs. IEEE Access, 10: 75729-75741 [DOI: 10.1109/ACCESS.2022.3191784]

You Y N, Chen T L, Sui Y D, Chen T, Wang Z Y and Shen Y. 2020. Graph contrastive learning with augmentations//Proceedings of the 34th International Conference on Neural Information Processing Systems. Vancouver, Canada: Curran Associates Inc.: #488

Yu B, Yin H T and Zhu Z X. 2018. Spatio-temporal graph convolutional networks: a deep learning framework for traffic forecasting//Proceedings of the 27th International Joint Conference on Artificial Intelligence. Stockholm, Sweden: ijcai.org: 3634-3640 [DOI: 10.24963/ijcai.2018/505]

Yu W N, Kim I Y and Mechefske C. 2021. Analysis of different RNN autoencoder variants for time series classification and machine prognostics. Mechanical Systems and Signal Processing, 149: #107322 [DOI: 10.1016/j.ymssp.2020.107322]

Yuan H T and Li G L. 2021. A survey of traffic prediction: from spatio-temporal data to intelligent transportation. Data Science and Engineering, 6(1): 63-85 [DOI: 10.1007/s41019-020-00151-z]

Zhang F, Ding Y and Liao Y H. 2023. Financial data collection based on big data intelligent processing. International Journal of Information Technologies and Systems Approach, 16(3): 1-13 [DOI: 10.4018/ijitsa.320514]

Zhang H. 2024. Multivariate time series classification with graph neural networks//Proceedings of the 4th International Conference on Artificial Intelligence, Big Data and Algorithms. Zhengzhou, China: ACM: 361-367 [DOI: 10.1145/3690407.3690469]

Zhang S, Tong H H, Xu J J and Maciejewski R. 2019. Graph convolutional networks: a comprehensive review. Computational Social Networks, 6(1): #11 [DOI: 10.1186/s40649-019-0069-y]

Zhao Z L, Xiao Z and Tao J. 2024. MSDG: multi-scale dynamic graph neural network for industrial time series anomaly detection. Sensors, 24(22): #7218 [DOI: 10.3390/s24227218]

Zheng D, Ma C, Wang M J, Zhou J J, Su Q D, Song X, et al. 2020. DistDGL: distributed graph neural network training for billion-scale graphs//Proceedings of the 10th IEEE/ACM Workshop on Irregular Applications: Architectures and Algorithms. Atlanta, USA: IEEE: 36-44 [DOI: 10.1109/IA351965.2020.00011]

Zhou J, Cui G Q, Hu S D, Zhang Z Y, Yang C, Liu Z Y, et al. 2020. Graph neural networks: a review of methods and applications. AI Open, 1: 57-81 [DOI: 10.1016/j.aiopen.2021.01.001]

Zhu J W, Wang Q J, Tao C, Deng H H, Zhao L and Li H F. 2021. AST-GCN: attribute-augmented spatiotemporal graph convolutional network for traffic forecasting. IEEE Access, 9: 35973-35983 [DOI: 10.1109/ACCESS.2021.3062114]

作者简介

李贯宸，男，硕士研究生，主要研究方向为机器学习、深度学习和时间序列分类。E-mail：1223045623@njupt.edu.cn

孙杰，通信作者，男，教授，硕士生导师，主要研究方向为机器学习、深度学习、时间序列分析和大语言模型。

E-mail：jie.sun@njxzc.edu.cn

孙志林，男，博士，高级工程师，主要研究方向为机器学习、信号处理、大数据分析与处理。

E-mail：sunzhilin@pgsensing.com