

# 融合多维过程视角：一种基于上下文感知图注意力的 业务流程预测框架

张希为<sup>1</sup>, 方贤文<sup>1,2\*</sup>, 毛古宝<sup>1</sup>

(1. 安徽理工大学数学与大数据学院, 安徽淮南 232001; 2. 同济大学嵌入式系统与服务计算教育部重点实验室, 上海 201804)

**摘要:** 随着数字化转型的深化, 以预测性流程监控为核心的数据驱动流程分析技术, 已成为企业提升运营效能与决策水平的关键。为提升预测性流程监控的精度与泛化能力, 现有研究致力于从海量的历史事件日志中挖掘流程的深层表征。然而, 真实业务流程的演化不仅遵循既定的时序逻辑, 也受到资源分配、数据依赖等潜在结构化因素的影响, 这对现有预测模型的表征能力构成了严峻挑战。具体而言, 主流预测方法的性能常受限于其单一的过程视角与静态的信息融合策略。多数方法, 即便是基于图神经网络, 也倾向于从单一的控制流视角建模, 忽略了资源交互、数据依赖等关键维度, 进而造成流程深层结构与多维关系表征的鸿沟。此外, 少数尝试融合多维信息的研究也大多采用静态融合策略, 缺乏对多维信息的上下文感知融合能力, 使得模型适应性不足。为应对上述挑战, 本文提出一种上下文感知多视角图融合预测框架 (Context-Aware Multi-view Graph Fusion, CAM-GF)。该框架首先突破控制流局限, 系统地构建了一个过程图谱, 该图谱不仅包含捕捉宏观规律的长期依赖图等基础控制流视角, 还涵盖了如揭示组织协作关系的资源交互图等扩展语义视角, 以捕获流程全局性、多层次的结构化知识。进而, 在时空信息融合层面, 设计了一种新颖的上下文感知图注意力机制, 它以案例的实时执行前缀为输入, 动态学习并分配各视角的融合权重。最后, 引入 Transformer 对动态融合后的特征序列进行深度时序建模, 以实现对下一活动的精准预测。为验证框架的有效性与实用价值, 本文在 6 个公开的真实业务流程数据集上进行了综合实验。结果表明, 相较于多种主流模型, CAM-GF 框架在下一活动预测任务上预测准确率平均提升 4.16 个百分点。此外, 框架动态生成的注意力权重为模型行为提供了高价值的可解释性, 揭示了模型如何根据预测反馈与实时上下文, 既能在局部上下文失效时回归对全局流程结构的依赖, 也能在特定情境下转而聚焦于资源分配等关键语义视角, 充分验证了所提框架在精度与透明度上的先进性与实用价值。

**关键词:** 预测性流程监控; 图注意力网络; 多视角表征; 上下文感知融合; 可解释

**基金项目:** 国家自然科学基金 (No.61572035, No.61402011); 嵌入式系统与服务计算教育部重点实验室开放课题 (No.ESSCKF2021-05); 安徽省重点研究与开发计划项目 (No.2022a05020005); 安徽省自然科学基金 (水科学联合基金) (No.2308085US11); 安徽省高校领军骨干人才项目 (No.2020-01-12); 安徽省学术和技术带头人资助项目 (No.2022D327)

中图分类号: TP391 文献标识码: A 文章编号: 0372-2112(2025)10-3705-13

电子学报 URL: <http://www.ejournal.org.cn> DOI: 10.12263/DZXB.20250599

## Fusing Multi-Dimensional Process Views: A Context-Aware Graph Attention Framework for Business Process Prediction

ZHANG Xi-wei<sup>1</sup>, FANG Xian-wen<sup>1,2\*</sup>, MAO Gu-bao<sup>1</sup>

(1. School of Mathematics and Big Data, Anhui University of Science and Technology, Huainan, Anhui 232001, China;

2. Key Laboratory of Embedded Systems and Service Computing, Ministry of Education, Tongji University, Shanghai 201804, China)

**Abstract:** Amidst deepening digital transformation, data-driven process analysis, with Predictive Process Monitoring (PPM) at its core, has become pivotal for enhancing enterprise operational efficiency and decision-making. To improve the accuracy and generalization of PPM, existing research focuses on mining deep representations from vast event logs. However, the evolution of real-world business processes is influenced not only by temporal logic but also by underlying structural factors such as resource allocation and data dependencies. This complexity poses a formidable challenge to the representa-

tional capabilities of existing predictive models. Specifically, the performance of mainstream predictive methods is often constrained by their reliance on a singular process view and static information fusion strategies. Most approaches, even those based on Graph Neural Networks (GNNs), tend to model processes from a single control-flow perspective. This overlooks critical dimensions such as resource interactions and data dependencies, creating a gap in the representation of deep process structures and multi-dimensional relationships. Furthermore, the few studies that attempt to integrate multi-dimensional information typically employ static fusion strategies, lacking a context-aware fusion capability and resulting in models with insufficient adaptability. To address these challenges, this paper proposes a context-aware multi-view graph fusion (CAM-GF) framework. The framework first transcends the limitations of the control-flow perspective by systematically constructing a process graph map. This graph map comprises not only basic control-flow views, such as a long-term dependency graph that captures macroscopic patterns, but also extended semantic views, like a resource interaction graph that reveals organizational collaboration, thereby capturing holistic and multi-level structural knowledge. Subsequently, a novel context-aware graph attention mechanism is designed for spatio-temporal information fusion. It takes the real-time prefix of a case as input to dynamically learn and assign fusion weights to each view. Finally, a Transformer is introduced to perform deep temporal modeling on the dynamically fused feature sequence to achieve precise next-activity prediction. To validate the framework's effectiveness and practical value, comprehensive experiments were conducted on six public, real-world business process datasets. The results demonstrate that, compared to various mainstream baseline models, the CAM-GF framework achieves an average accuracy improvement of 4.16 percentage points on the next-activity prediction task. Furthermore, the dynamically generated attention weights provide high-value interpretability for the model's behavior, revealing how the model, based on predictive feedback and real-time context, can both rely on global process structures when local context fails and pivot to focus on critical semantic views, such as resource allocation, in specific situations. This thoroughly validates the proposed framework's advancement in both accuracy and transparency.

**Key words:** predictive process monitoring; graph attention networks; multi-view representation; context-aware fusion; interpretability

**Foundation Item(s):** National Natural Science Foundation of China (No.61572035, No.61402011); Open Project Program of the Key Laboratory of Embedded System and Service Computing of Ministry of Education (No.ESSCKF2021-05); Key Research and Development Program of Anhui Province (No.2022a05020005); Anhui Provincial Natural Science Foundation (Water Science Joint Fund) (No.2308085US11); Leading Backbone Talent Project in Anhui Province of China (No.2020-01-12); Anhui Province Academic and Technical Leader Foundation (No.2022D327)

## 1 引言

随着数字化转型的深入,业务流程管理已成为企业提升运营效率与核心竞争力的关键.近年来,得益于事件日志的广泛可用性,数据驱动的流程分析,特别是预测性流程监控(Predictive Process Monitoring, PPM),受到了学术界和工业界的极大关注<sup>[1]</sup>. PPM旨在利用历史数据,在流程执行过程中对未来状态(如未来活动、剩余时间或最终结果)进行实时预测,从而为管理者提供前瞻性的决策支持,以规避潜在风险、优化资源配置.然而,尽管PPM前景广阔,但由于真实业务流程内在的复杂性和动态性,构建一个能够精确捕捉活动间复杂时空关系并适应动态环境的预测模型,仍然是一项艰巨的挑战.

早期的预测性流程监控研究主要采用经典的机器学习或基于深度学习的序列模型<sup>[2]</sup>.以长短期记忆网络(Long Short-Term Memory, LSTM)及其变体<sup>[3-6]</sup>Transformer模型<sup>[7]</sup>为代表的方法,在捕捉流程轨迹的时间依赖性方面取得了显著成功.这些方法将流程实例视为一个活动序列,通过学习历史序列模式来预测未来.然

而,它们的核心局限在于将流程过度简化为一维线性序列,忽视了流程中蕴含的丰富结构化信息.在复杂的业务场景中,活动间的关系远非简单的先后顺序,往往呈现出并发、循环、选择等复杂的拓扑结构,这些重要的空间关系被序列模型所忽略<sup>[8]</sup>,从而限制了其表达能力和预测精度.为了弥补这一不足,研究者们引入了图神经网络(Graph Neural Networks, GNN)<sup>[9,10]</sup>,它为直接在图结构数据上进行学习提供了强大的工具,能够有效捕捉流程中的空间拓扑关系.基于GNN的PPM方法<sup>[11-13]</sup>通过将流程建模为图,显著提升了对活动间结构关系的理解.

即便如此,现有基于GNN的方法仍存在两个尚未被充分解决的深层痛点:其一,是过程视角的片面性与信息鸿沟.绝大多数方法仅从单一的控制流视角(即活动间的直接跟随关系)来构建图,这本质上造成了巨大的信息损失.流程中同样关键的资源共享关系(如哪些活动由同一人处理)、数据依赖关系(如哪些活动处理相似的数据对象)等多个维度的信息被忽略,导致模型对流程的理解是片面的,预测能力也因此受限.其二,

是上下文建模的静态化与适应性缺失.即使少数研究尝试融合多源信息,也大多采用简单的静态融合策略.然而,不同信息维度的相对重要性并非一成不变,而是依赖于具体案例的实时上下文.例如,在一个常规案例中控制流可能起主导作用,但在一个涉及高级专家或大额订单的案例中,资源或数据视角的重要性可能瞬间凸显.现有模型普遍缺乏一种机制来动态、自适应地权衡不同视角的重要性,表现出明显的“僵化”和适应性不足.

综上所述,为了实现更精准、更具适应性的业务流程预测,必须解决上述两大核心挑战,即如何克服单一视角带来的信息偏差,以及如何突破静态建模导致的适应性瓶颈.为此,本文提出了一种全新的基于上下文感知多视角图融合的业务流程预测框架CAM-GF(Context-Aware Multi-view Graph Fusion).该框架的核心思想如下:首先从事件日志中构建一个包含控制流、资源、数据等多个维度的过程“图谱”,以获得对流程全面而立体的认知;然后,利用一个创新的上下文感知注意力机制,根据当前案例的实时前缀信息,动态地学习不同视角的重要性权重,并以此为指导进行自适应的信息融合;最后,将融合后的特征输入时序模型中进行预测.在多个真实事件日志上的大量实验表明,CAM-GF能够有效捕捉和融合多维度的时空信息,其预测性能显著优于现有的主流方法.

本文的主要贡献如下:

(1)提出了一种多视角过程图谱构建方法,该方法通过并行提取流程中的控制流、资源等多维结构关系并为各视角构建专属图表示,有效克服了传统单视角建模所带来的信息鸿沟问题.

(2)设计了一种上下文感知的注意力融合机制,它能够为每个流程实例的前缀,动态为不同的视角分配权重,解决了流程多视角静态融合策略缺乏适应性的难题.

(3)构建了一个可以学习复杂业务流程依赖关系的时空联合预测架构,该架构有机结合了视角特异性的图编码、上下文感知的动态融合与Transformer时序建模,以实现精准预测.

## 2 相关工作

在预测性流程监控的诸多任务中,下一活动预测任务因其能为流程参与者提供即时的操作指导和预警,从而有效避免错误操作或流程瓶颈,一直是研究的焦点.近年来,随着深度学习技术的飞速发展,利用其强大的非线性模式捕捉能力来解决此问题已成为主流范式.其中,以循环神经网络(Recurrent Neural Networks, RNN)为基础的序列模型被率先引入该领域.最具代表性

的是Tax等<sup>[3]</sup>提出的基于长短期记忆网络的方法,它将流程实例视为一个时间序列,并成功应用于下一活动预测任务.Pasquadibisceglie等提出使用卷积神经网络(Convolutional Neural Networks, CNN),把事件日志转为图像格式,进而做活动预测.Bukhsh等人<sup>[14]</sup>提出通过使用Transformer克服传统预测模型不能有效捕捉输入序列长期依赖关系的问题,实现对预测下一活动的准确预测.Smedt等人<sup>[15]</sup>提出了一种预测流程模型监控的方法,该方法通过建模流程模型中活动之间的约束关系,捕获活动间的详细信息,并预测整个流程模型的未来状态.除此之外,Taymouri等人<sup>[16]</sup>基于生成对抗网,通过LSTM体系结构和前馈神经网络组成的判别器和生成器的最小博弈来预测下一活动.然而,这些序列模型的共同范式是将流程轨迹视为一维的线性序列,这是一种仅包含时序控制流的单视角建模方式.这种范式天生无法捕捉业务流程中普遍存在的并发、选择、循环等复杂的拓扑结构,这些重要的空间信息被忽略,从而限制了模型对流程内在逻辑的理解深度,成为其性能瓶颈.

为了将流程的拓扑结构信息融入模型,研究社区开始全面拥抱图神经网络.例如,Weinzierl等人<sup>[12]</sup>通过显示的建模,将流程的直接跟随关系构建为图,并使用门控图神经网络来预测流程下一活动.这虽然引入了结构信息,但本质上仍是“单视角的控制流”思维,未能解决“信息鸿沟”问题.为了获得更全面的流程表示,研究者们开始探索融合多维度信息(即多视角)的方法.其中,最直观的策略是直接编码.Evermann等人<sup>[17]</sup>对活动相关的上下文数据进行编码,并与活动向量进行拼接,从而构建一个信息更丰富的向量作为LSTM的输入来预测下一活动.Camargo等人<sup>[4]</sup>同样也使用了直接编码与LSTM结合的方式来预测下一活动,除此之外,还引入IF-THEN Rule(条件规则)来解释预测结果.Rama-Maneiro等人<sup>[8]</sup>编码事件日志的上下文数据到属性特征矩阵中,并使用一种组合图卷积网络与循环神经网络的方法来预测下一活动.

这类方法在模型训练前,就将资源、数据等多源特征与活动自身特征进行拼接,形成一个“富特征”向量,然后在一个基于控制流的图上进行学习.这种“早期融合”策略虽然简单,但它模糊了不同类型信息的内在结构,例如,它无法区分“活动A后是活动B”和“活动A与活动B共享一个资源”这两种性质完全不同的关系,限制了模型从特定视角(如资源视角)中学习精细模式的能力.为了更精细地利用多视角信息,更为先进的方法采用了间接编码的策略.这类方法会为不同的视角构建独立的图或使用独立的编码路径.例如,Andrea等人<sup>[18]</sup>提出了利用实例图结合深度图卷积神经网络预测

下一活动的方法. 该方法不仅利用图结构编码了控制流信息, 还结合了多视角的数据特征来丰富预测的输入. Wang 等人<sup>[19]</sup>提出了一种历史信息驱动的图神经网络方法, 通过构建统一的历史图, 将事件日志中的控制流和多视角属性相结合并预测下一个活动. Pasquadi-bisceglie 等人<sup>[20]</sup>提出利用异构图神经网络进行业务流程监控的方法, 该方法结合多视角学习与异构图注意力网络, 构建包含不同节点和边类型的图结构, 以捕捉事件特征及其关系, 并在预测精度与可解释之间取得平衡.

### 3 背景知识

本节将介绍本文研究工作所涉及的基本概念和形式化定义, 包括事件日志的构成以及下一活动预测任务的数学描述.

#### 3.1 事件日志

**定义 1** 事件. 一个事件  $e$  是一个元组  $e = (c, a, t, \text{attrs})$ , 其中  $c$  是案例 ID, 且唯一标识一个流程实例;  $a$  是活动名称, 表示所执行的工作类型;  $t$  是事件的时间戳;  $\text{attrs}$  是一个属性集, 可以包含执行资源、事件成本等其他属性信息. 基于事件的定义, 可以进一步定义流程实例跟轨迹.

**定义 2** 流程实例. 一个流程实例, 也称案例 (Case), 代表了一次完整的业务流程执行过程 (例如, 一次客户订单的处理全过程). 在事件日志中, 所有具有相同案例 ID ( $c$ ) 的事件共同构成一个流程实例.

**定义 3** 轨迹. 一个轨迹  $\sigma$  是隶属于同一个流程实例, 并按时间戳非递减顺序排列的事件序列. 对于一个包含  $n$  个事件的案例, 其轨迹可以表示为  $\sigma = \langle e_1, e_2, \dots, e_n \rangle$ , 其中  $e_i. c = e_{i+1}. c$  且  $e_i. t \leq e_{i+1}. t$ . 因此, 整个事件日志  $L$  可以看作是一个由多条轨迹组成的集合, 即  $L = \{\sigma_1, \sigma_2, \dots, \sigma_m\}$ .

**定义 4** 前缀轨迹及其标签. 在预测任务中, 通常使用一个正在进行的案例的前缀部分来预测其下一活动. 对于一个长度为  $n$  的完整轨迹  $\sigma = \langle e_1, e_2, \dots, e_n \rangle$ , 其长度为  $k$  ( $1 \leq k \leq n$ ) 的前缀轨迹被定义为  $\sigma_k = \langle e_1, e_2, \dots, e_k \rangle$ . 对于下一活动预测任务, 该前缀轨迹  $\sigma_k$  的标签就是紧随其后的那个事件的活动名称, 即  $y_k = e_{k+1}. a$ . 历史事件日志中的每条轨迹都可以被切割成多个 (前缀轨迹, 标签) 的训练样本对.

#### 3.2 业务流程下一活动预测任务描述

基于上述定义, 业务流程下一活动预测可以被形式化地描述为一个监督学习问题.

设  $a \in A$  为事件日志  $L$  中所有可能活动的集合, 即活动空间. 预测模型的输入是一个来自某个正在运行

的案例的前缀轨迹  $\sigma_k = \langle e_1, e_2, \dots, e_k \rangle$ . 模型的目标是学习一个映射函数  $f: \sigma_k \rightarrow \hat{a}_{k+1}$ , 其中  $\hat{a}_{k+1}$  是模型预测出的最有可能发生的下一活动. 在应用阶段, 通常被建模为多分类问题. 模型  $f$  的实际输出是在所有活动空间  $A$  上的一个概率分布  $P(\cdot | \sigma_k)$ , 而最终的预测结果则是该分布中概率最高的活动:

$$\hat{a}_{k+1} = \arg \max_{a \in A} P(a | \sigma_k) \quad (1)$$

模型的优化目标是最小化预测活动  $\hat{a}_{k+1}$  与真实标签  $a_{k+1}$  之间的差异, 本文提出的模型正是旨在解决这个问题. 它不仅利用了前缀轨迹  $\sigma_k$  中的序列信息, 还从事件日志  $L$  中构建并融合了多维度的图结构信息, 以期获得更精准的预测性能.

## 4 上下文感知的多视角图融合预测框架

### 4.1 框架概览

上下文感知的多视角图融合预测框架如图 1 所示, 该框架可以分为三个阶段.

第一阶段: 多视角过程图谱构建. 该阶段旨在从事件日志中, 系统性地构建一个由两类核心视角组成的图谱: 一类是可从任何事件日志中提取的、作为流程“主干”的基础控制流视角; 另一类则是依赖于特定日志附加属性的扩展语义视角.

第二阶段: 自适应多视角图融合与时空表示学习. 该阶段通过并行的图编码器从多视角图谱中学习特定于视角的空间表示, 并利用一个创新的上下文感知注意力机制对这些表示进行动态加权融合. 随后, 融合后的特征序列被送入 Transformer 以捕捉其全局依赖, 最终生成一个深度融合、信息丰富的时空特征表示.

第三阶段: 下一活动预测与模型优化. 该阶段利用学习到的高级特征表示, 通过一个预测头生成下一活动预测结果, 并基于一个处理类别不平衡问题的损失函数对整个模型进行优化.

本框架的整体设计遵循空间优先的解耦式建模. 其核心思想是, 在进行时序依赖建模之前, 优先且充分地在全局事件日志中提取并编码多维度的空间结构信息, 并将这些信息动态地注入到当前流程实例的序列表示中. 这种职责分离的设计, 使各模块能专注于其擅长的学习任务, 从而实现了对流时空特性的更优表征与预测.

### 4.2 多视角过程图谱构建

单一的控制流图不足以全面地刻画业务流程的复杂性. 为了使模型能够从多个维度深度理解流程, 本节提出了一种灵活的图谱构建策略, 它从事件日志  $L$  系统性地构建一个过程图谱. 图谱中的每一个视角都由一个邻接矩阵  $A = \mathbb{R}^{N \times N}$  表示, 其中  $N = |A|$ .

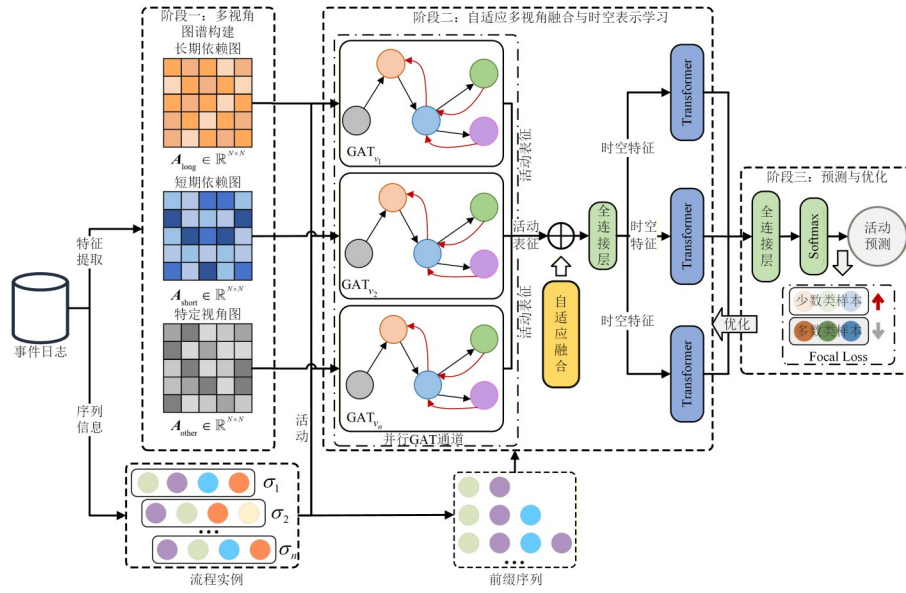


图1 上下文感知的多视角图融合预测框架

### 4.2.1 基础控制流视角

该类视角源自所有事件日志都必备的三元组信息(案例、活动名称、时间戳),它们共同构成了流程结构与动态的核心骨架。

(1)长期依赖图( $A_{long}$ ). 此图是流程最基础、稳固的结构模式. 它是一个有向加权图,其中邻接矩阵的元素  $A_{long}(i, j)$  定义为活动  $a_j$  紧跟在活动  $a_i$  之后出现的总次数.  $A_{long}$  捕获了流程中基于频率的、宏观的、确定性的长期依赖规律. 其形式化为

$$A_{long}(i, j) = \{ (e_k, e_{k+1}) \in \sigma, \forall \sigma \in L \mid e_k \cdot a = a_i, e_{k+1} \cdot a = a_j \} \quad (2)$$

(2)短期依赖图( $A_{short}$ ). 此图作为基础控制流的动态细化,着重强调了活动发生的即时性影响. 它同样是一个有向加权图,但其边权重是通过指数衰减函数进行累加的. 具体而言,  $A_{short}(i, j)$  的值由以下公式累加生成:

$$A_{short}(i, j) = \sum_{(e_k, e_{k+1}) \mid e_k \cdot a = a_i, e_{k+1} \cdot a = a_j} e^{-\lambda \Delta t} \quad (3)$$

其中,  $\Delta t_k = e_{k+1} \cdot t - e_k \cdot t$  是某次相继活动间的时间差;  $\lambda$  是控制衰减速率的超参数. 这种机制使得时间间隔越短的活动对拥有越高的连接权重,从而使模型能敏锐地捕捉到流程中的短期依赖模式。

### 4.2.2 扩展语义视角

为了进一步增强模型的表达能力,当事件日志提供除基础三元组之外的附加属性时,该框架能够灵活地构建相应的扩展语义视角图. 这些视角揭示了隐藏在控制流之下的深层关系. 例如:

(1)当日志中包含资源信息时,可以构建此图来刻画组织协同关系,即资源视角( $A_{res}$ ). 它通常被构建为

一个无向图,其中边权重  $A_{res}(i, j)$  表示活动  $a_i$  和  $a_j$  在同一案例中由同一资源执行的频率. 此图能揭示因人员分工、技能分组或部门协作而产生的隐性关联。

(2)其他特定视角图. 不同的业务场景包含其特定的数据属性,例如,贷款流程中,不同案例的贷款金额可以构建一个数据关联图,其边权重  $A_{data}(i, j)$  代表两个活动在处理同一业务案例中共同出现的频率. 例如,在特定场景下(如贷款审批流程),若事件带有“贷款金额”这一数值属性,则可以构建一个基于数据值相似性的图,该图可以揭示基于业务对象内在价值的流程模式. 同理,若日志包含地理位置、成本、产品型号等信息,亦可构建相应的地理邻近图、成本关联图等。

通过这种组合视角构建策略,CAM-GF框架在多视角构建上具备内在的灵活性与鲁棒性. 扩展语义视角的构建是数据驱动的,它严格依赖于事件日志中相应属性的存在性. 若日志缺乏特定属性(如资源、特定业务数据等),则对应的语义视角将不会被创建,模型会自动在剩余的可用视角集合上进行后续的动态融合. 除此之外,本框架存在一个清晰的退化路径. 即,在仅包含基础三元组信息的“纯控制流”场景下,CAM-GF将自动退化为一个仅利用长期和短期依赖图的预测模型. 这种设计不仅保证了其在所有标准事件日志上的适用性,也为其在富含属性的特定领域数据集上取得卓越性能提供了强大的可扩展性。

### 4.3 自适应多视角融合与时空表征学习

在获得多视角过程图谱后,需要整合异构的结构化信息与流程实例的时序信息,并学习一个综合的特征表示,以服务于下游的下一活动预测任务. 本节共包含三个核心操作:(1)通过多通道图编码器提取各视角

的空间特征;(2) 利用上下文感知的注意力机制对空间特征进行动态融合;(3) 使用时序编码器对融合后的序列进行处理以捕捉时间依赖. 经过这三个操作,该模块最终输出一个深度融合了多维空间信息与时间信息的最终序列表示.

#### 4.3.1 多通道图编码

此阶段的目标是从不同的视角图谱中,并行地提取特定于该视角的空间结构特征. 首先,为流程中的全部  $N$  个唯一活动  $a_i \in \mathcal{A}$  初始化一个可学习的嵌入矩阵  $\mathbf{E} \in \mathbb{R}^{N \times d_{\text{embed}}}$ , 其中  $d_{\text{embed}}$  为嵌入维度. 该矩阵作为所有活动的基础特征表示. 随后,为每个在第 4.2 节中定义的视角图  $\mathbf{A}_v$ , 配置一个独立的图注意力网络编码器(Graph Attention Network)<sup>[21]</sup>, 记为  $\text{GAT}_v$ . 每个编码器  $\text{GAT}_v$  将全局活动嵌入矩阵  $\mathbf{E}$  和其对应的邻接矩阵  $\mathbf{A}_v$  配作为输入, 输出一个视角特定的节点表示矩阵  $\mathbf{H}_v \in \mathbb{R}^{N \times d_{\text{embed}}}$ . 其形式化为

$$\mathbf{H}_v = \text{GAT}_v(\mathbf{E}, \mathbf{A}_v) \quad (4)$$

此阶段的最终输出是一组视角特定的活动表示矩阵集合  $\{\mathbf{H}_v | v \in V\}$ , 每一个矩阵  $\mathbf{H}_v$  都蕴含了从视角  $v$  中学到的结构知识.

#### 4.3.2 上下文感知自适应融合

此阶段的目标是根据当前流程实例的上下文,对上一阶段产出的多组并行空间特征进行动态加权融合,如图 2 所示. 具体来说,对于一个长度为  $k$  的前缀轨迹  $\sigma_k = \langle e_1, e_2, \dots, e_k \rangle$ , 首先使用上一阶段学习到的各视角表示矩阵  $\mathbf{H}_v$  为查找表,将输入前缀轨迹中的活动序列转换为各视角的序列表示  $S_v \in \mathbb{R}^{k \times d_{\text{embed}}}$ . 随后,为生成一个稳定且具有代表性的上下文向量,本文选取控制流视角(具体采用长期依赖图  $\mathbf{A}_{\text{long}}$ )作为主视角,这是因为控制流作为流程的通用骨架,在任何标准事件日志中都保证存在,因此是最稳定和最具代表性的信息源<sup>[22]</sup>. 利用该主视角的前缀序列表示  $S_v$ , 通过池化操作,得到一个能够概括当前前缀核心上下文向量  $\mathbf{c}$ , 然后将该向量输入到注意力网络  $f_{\text{attn}}$  中,并生成一组对应各个视角的动态注意力权重  $\alpha$ , 其计算式为

$$\alpha = \text{Softmax}(f_{\text{attn}}(\mathbf{c})) \quad (5)$$

最后使用权重  $\alpha$  对所有视角的序列表示  $\{S_v\}$  进行加权求和,得到一个经过动态融合的、统一的序列结构表示  $S_{\text{fused}}$ , 形式化为

$$S_{\text{fused}} = \sum_{v \in V} \alpha_v \cdot S_v \quad (6)$$

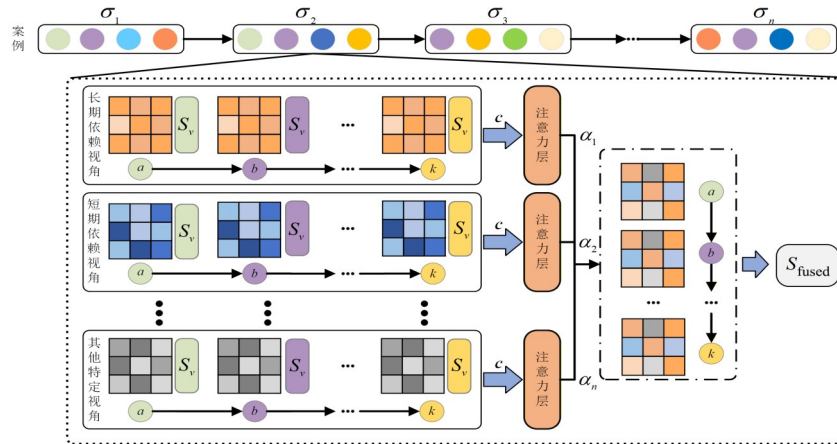


图2 自适应多视角图融合

#### 4.3.3 时序依赖建模

此阶段的目标是在融合了多维空间信息的基础上,进一步捕捉序列中事件之间的时间依赖关系. 首先,将上一阶段得到的融合结构表示  $S_{\text{fused}}$  与输入的原始时间特征序列  $x_{\text{time}}$  进行拼接,并通过一个线性投射层得到组合特征序列  $S_{\text{combined}} \in \mathbb{R}^{k \times d_{\text{embed}}}$ . 然后,为  $S_{\text{combined}}$  添加位置编码以注入序列的顺序信息. 处理后的序列被送入一个多层的 Transformer<sup>[23]</sup> 编码器,用于捕捉序列内事件间的长距离依赖关系. Transformer 编码器的输出  $H_{\text{final}} \in \mathbb{R}^{k \times d_{\text{embed}}}$  即为该模块最终产出的、融合了多维

时空信息的最终序列表示.

#### 4.4 下一活动预测及模型优化

在通过前述模块学习到融合了多维时空信息的最终序列  $H_{\text{final}}$  表示后,框架的最后一步是将其转化为具体的预测结果. 具体而言,首先提取序列表示  $H_{\text{final}}$  中对应于最后一个时间步的特征向量  $\mathbf{h}_{\text{last}}$ , 该向量凝聚了整个前缀轨迹的完整上下文信息. 然后该向量被送入一个预测头,即一个标准的全连接线性层. 该线性层负责将高维的特征向量  $\mathbf{h}_{\text{last}}$  投射到维度为  $N$  的输出空间,从而为每一个可能的下一活动生成一个原始的预测分

数. 最后,该分数向量会经过 Softmax 函数进行变换,生成模型最终的预测结果  $\hat{a}$ .

考虑到真实业务流程事件日志中普遍存在的类别不平衡问题,即少数核心活动的发生频率远高于大量其他活动,若采用标准的交叉熵损失进行优化,将导致模型过度关注多数类样本而忽略对稀有但关键的少数类活动的学习. 因此,为缓解此问题,本框架采用 Focal Loss 作为最终的优化目标. 其形式化如下:

$$FL(p_i) = -\alpha_i (1 - p_i)^\gamma \log(p_i) \quad (7)$$

其中,  $p_i$  是真实类别的预测概率;  $\gamma$  是动态缩放因子,该因子可以降低“简单”样本(多数类活动)的损失权重,从而使模型能够更加“聚焦”于学习那些难以分类的“困难”样本(少数类活动);  $\alpha_i$  是平衡系数,可以平衡不同类别的重要性.

## 5 实验

本节对提出的上下文感知的多视角图融合预测框架进行试验探究与分析.

表 1 事件日志统计信息

Dataset	案例总数	事件总数	活动种类	最短案例长度	最长案例长度	平均案例长度
Helpdesk	4 590	21 348	14	2	15	5
BPIC_2012_A	13 087	60 849	10	3	8	5
BPIC_2012_O	5 015	31 244	7	3	30	6
Hospital Billing	99 999	451 359	18	1	217	5
Prepaid Travel Costs	2 099	18 246	29	1	21	9
RTFM	150 370	561 470	11	2	20	4

### 5.2 实验设置

在实验开始前,首先对数据集进行划分. 为了防止同一案例的事件序列泄漏到不同的数据集中,本文严格遵循案例级别的划分原则,所有案例 ID 被划分为训练集(70%)、验证集(10%)和测试集(20%).

所有实验均基于一套统一的基线超参数配置. 采用 AdamW 优化器,初始学习率设置为  $5 \times 10^{-4}$ ,并配置了  $1 \times 10^{-4}$  的权重衰减. 模型的批处理大小(batch size)为 128. 在训练过程中,采用了早停机制,若验证集上的 Accuracy 分数连续 15 轮没有提升,则终止训练以防止过拟合. 对于模型架构,GAT 和 Transformer 的多头注意力头数均为 4,Transformer 编码器包含 4 个堆叠层. 此外,隐含层节点数,向量维数以及 Dropout 值的设置可根据不同事件日志的复杂性进行微调以达到最优性能. 最后,为确保评估严谨性,本次实验结果均为 5 次不同随机种子实验的均值与标准差,其硬件环境如下:酷睿(TM)i9-13900H CPU @2.60 GHz, NVIDIA GeForce GTX 4070 GPU, 32 GB RAM. 软件环境如下: Windows 11 操作系统,算法实现基于 Python 3.10 版本.

为了定量评估模型的预测性能,采用下一活动预测任务中常用的分类评估指标:准确率(Accuracy),加

### 5.1 实验数据集

为了全面评估模型的性能和泛化能力,实验选取了 6 个来自“4TU Center for Research Data”平台的真实业务流程事件日志. 这些数据集来源多样,涵盖了 IT 服务、医疗、金融等多个领域,具有不同的流程复杂度和数据特征. 选用的数据集如下:

Helpdesk<sup>①</sup>: 记录了 IT 服务台处理技术支持请求的流程. BPIC\_2012\_A<sup>②</sup>: 记录了某荷兰金融机构的贷款申请流程(受理部分). BPIC\_2012\_O<sup>③</sup>: 记录了同一金融机构贷款申请流程的后续(WGA 提交)部分. Hospital Billing<sup>④</sup>: 记录了某家医院的计费流程. Prepaid Travel Costs<sup>⑤</sup>: 记录了某公司预付差旅费用的申请与报销流程 Road Traffic Fine Management Process (RTFM)<sup>⑥</sup>: 记录了某地区交通罚款管理的流程.

表 1 展示了这些数据集的详细统计信息,包括案例总数、事件总数、活动种类、最短案例长度、最长案例长度和平均案例长度,以便了解其规模与特性.

权 F1 分数,精确率(Precision),召回率(Recall)以及精确率-召回率曲线下面积(Area Under the Precision-Recall Curve, PR-AUC). 其计算公式如下:

$$\text{Accuracy} = P_{\text{true}}/N \quad (8)$$

$$F1 = \sum_{n \in N} w_n \times F1_{\text{class}} \quad (9)$$

其中,  $w_n$  是类别  $n$  的样本在总样本中的比例;  $N$  代表样本总数;  $P_{\text{true}}$  代表正确预测的样本数;  $F1_{\text{class}}$  是精确率与召回率的调和平均值. 精确率与召回率计算公式为

$$\text{Precision} = P_t/P_m \quad (10)$$

$$\text{Recall} = P_t/P_a \quad (11)$$

其中,  $P_t$  是被正确预测为正类的样本数量;  $P_m$  与  $P_a$  分别是被模型预测为正类的总样本数量与数据集中实际为正类的总样本数量. 最后, PR-AUC 的计算公式

① <https://data.4tu.nl/datasets/94ee26c8-78f6-4387-b32b-f028f2103a2c>

② <https://data.4tu.nl/datasets/533f66a4-8911-4ac7-8612-1235d65d1f37/1>

③ <https://data.4tu.nl/datasets/533f66a4-8911-4ac7-8612-1235d65d1f37/1>

④ [https://data.4tu.nl/articles/\\_/12705113/1](https://data.4tu.nl/articles/_/12705113/1)

⑤ [https://data.4tu.nl/collections/\\_/5065541/1](https://data.4tu.nl/collections/_/5065541/1)

⑥ [https://data.4tu.nl/articles/\\_/12683249/1](https://data.4tu.nl/articles/_/12683249/1)

如下:

$$PR - AUC = \int_0^1 P(r) dr \quad (12)$$

其中,  $P(r)$  是将召回率视为自变量  $r$  时, 精确率随之变化的函数. 这个公式计算的是由一系列 (召回率, 精确率) 点对构成的 PR 曲线下方的总面积.

### 5.3 实验结果与分析

本节旨在从三个方面对模型的性能进行实证评估. 首先, 将提出的 CAM-GF 框架与最先进的方法进行严格的性能基准测试, 并分析所提框架的其他性能指标; 其次, 通过消融实验来量化分析模型关键组件的有效性; 然后, 通过具体案例深入剖析模型的可解释性; 最后分析框架的复杂度与效率.

#### 5.3.1 与基准方法性能对比

为全面评估 CAM-GF 框架的有效性和优越性, 本文选取了涵盖不同技术路径的多种方法进行比较. 这些方法分别是基于序列处理的方法 (如 LSTM<sup>[3]</sup>, Attn-LSTM<sup>[24]</sup>, Attn-BiLSTM<sup>[7]</sup>, Transformer<sup>[14]</sup>) 和基于图神经网络的方法 [如图卷积网络<sup>[11]</sup> (Graph Convolutional Network, GCN), 图卷积网络-有序神经元长短期记忆<sup>[25]</sup> (Graph Convolutional Network-Ordered Neurons Long Short-Term Memory, GCN-ONLSTM), 融合历史信息图的过程预测器<sup>[19]</sup> (History-informed Graph-based Process Predictor, HiGPP), 以及多层异构图<sup>[26]</sup> (Multi-layer Heterogeneous Graph-based, MHG)]. 各方法在 6 个公开数据集上的下一活动预测准确率对比详见表 2.

表 2 各模型在不同数据集上的下一活动预测准确率

单位: %

模型	Helpdesk	BPIC_2012_A	BPIC_2012_O	Hospital Billing	Prepaid Travel Costs	RTFM
LSTM	74.33	72.07	75.98	85.56	84.29	75.25
Attn-LSTM	77.81	73.78	78.66	85.01	76.58	86.72
Attn-BiLSTM	79.14	76.23	81.73	88.18	87.69	78.36
Transformer	81.26	76.88	80.46	86.15	85.21	81.73
GCN	73.05	73.65	72.86	70.68	69.67	70.15
GCN-ONLSTM	82.87	80.18	85.61	89.62	91.74	80.59
HiGPP	83.17	75.35	82.84	89.25	87.81	82.43
MHG	82.91	74.25	83.30	88.17	86.74	81.89
CAM-GF	86.78±0.12	78.58±0.08	84.16±0.05	93.20±0.01	88.50±0.09	83.45±0.21

对表 2 的分析揭示了几个关键洞察. 首先, 实验结果表明, 无论是经典的序列方法 (如 Attn-BiLSTM) 还是最近的 Transformer, 其性能均一致优于单纯基于图拓扑结构的 GCN 方法. 这一发现证实了在流程预测任务中, 有效捕捉由活动顺序构成的时序依赖是实现高性能的基础. 在此基础上, 一个关键的性能提升路径在于将流程的结构与上下文信息融入时序建模中. 第一种是通过增强图的表达能力来融入更丰富的上下文信息. 如 GCN-ONLSTM 这类混合方法, 通过结合图的拓扑信息与序列模型的时序捕捉能力, 其性能便超越了多数单一范式的模型, 初步验证了时空融合的有效性. 而最新的基于图的预测方法 (HiGPP 和 MHG) 将这一思想推向了新的高度, 它们通过构建更复杂的静态图结构来深度融合控制流与其他上下文维度, 进一步提升了预测精度. 因此, 这些方法的性能表现共同指向一个明确的结论: 在强大的时序建模基础上, 有效融入并利用流程的结构与上下文信息, 是提升预测性能的关键.

然而, 上述先进方法构建的图谱或融合策略在应用于具体案例时是相对固定的. 与之相比, 本文提出的 CAM-GF 框架在与所有基准方法的对比中取得了最佳的综合性能, 其核心优势在于引入了上下文感知的动态融合机制. CAM-GF 不仅通过并行 GAT 网络构建

了多维度的过程视角, 更关键的是, 它能够根据每个案例的实时演化状态, 动态地、自适应地权衡不同视角的重要性. 正是这种为每个具体情境“量身定制”最优信息融合策略的动态调整能力, 使得 CAM-GF 能够捕捉到其他先进方法所忽视的情境依赖性, 从而作出更为精准的活动预测.

为对各对比方法的总体性能进行严格的统计学评估, 本文采用了 Friedman 检验. 检验结果 ( $\chi^2 = 34.53$ ,  $p < 0.001$ ) 拒绝了“所有方法性能相同”的原假设, 表明不同方法间确实存在统计上的显著区别. 在此基础上, 采用了 Nemenyi 后续检验进行两两比较. 图 3 通过临界差异图可视化了 Nemenyi 后续检验的结果. 图 3 中展示了各方法的平均排名, 若排名差距超过临界差异 (Critical Difference, CD) 值, 则表明其性能存在统计显著性不同.

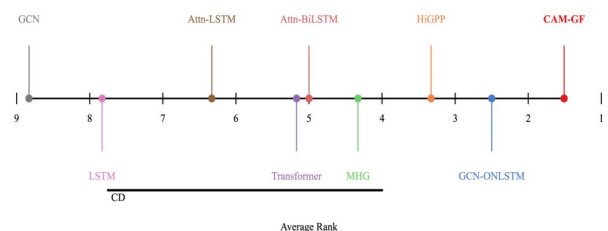


图 3 各方法性能排名的临界差异图

图 3 的检验结果表明 CAM-GF 的性能在统计上显著优于 Attn-LSTM、LSTM 及 GCN 等传统预测方法。其次, CAM-GF 在所有对比方法中获得了最低的平均排名,与 GCN-ONLSTM、HiGPP 等先进方法共同构成了性能的顶尖组。值得强调的是, CAM-GF 依赖的动态融合机制,是其区别于其他静态方法的根本优势。该机制在确保高精度的同时,赋予了模型根据实时上下文自适应调整决策的能力与可解释性,展现了更优越的综合性能。

### 5.3.2 CAM-GF 框架性能分析

本节旨在从类别不平衡鲁棒性与前缀长度敏感性

两个维度,对 CAM-GF 框架进行深度性能剖析。针对类别不平衡问题,表 3 展示了 CAM-GF 框架在 Precision、Recall、F1 及 PR-AUC 等多维度评价指标上的详细表现。表 3 的结果清晰地表明, CAM-GF 框架不仅在总体准确率上表现优异,其较高的 Precision 和 Recall 值也证明了模型在识别不同类活动时取得了良好的平衡。除此之外, CAM-GF 在所有数据集上均获得了较高的 PR-AUC 值,证实了所提框架能够有效识别频率较低但可能至关重要的业务路径,这在风险预警或机会发现等真实业务场景下具备很大的实用价值。

表 3 CAM-GF 框架的多指标性能评估

单位: %

数据集	Helpdesk	BPIC_2012_A	BPIC_2012_O	Hospital Billing	Prepaid Travel Costs	RTFM
Recall	86.78±0.12	78.58±0.08	84.16±0.05	93.20±0.01	88.50±0.09	83.45±0.21
Precision	82.62±0.03	74.82±0.06	83.20±0.02	91.29±0.00	85.29±0.02	80.62±0.00
F1	84.42±0.10	75.66±0.05	81.63±0.02	92.29±0.01	86.86±0.01	82.01±0.01
PR-AUC	80.02±0.03	74.57±0.04	81.62±0.03	92.27±0.01	85.30±0.01	81.63±0.00

其次,为探究模型在流程实例不同执行阶段的预测能力,图 4 分析了其性能随前缀长度的变化情况。图 4 的观察结果揭示了 CAM-GF 框架强大的阶段适应性与鲁棒性。对于高度结构化的流程(如 Prepaid Travel Costs),模型展现出快速学习能力,仅需较短的前缀便能收敛至高精度。这主要得益于框架的多视图图编码能力,使得模型能迅速从全局的、确定性的结构关系中定位当前实例。对于平均长度较短的流程(如 BPIC\_2012\_A),模型在整个执行周期内均保持了高水平的稳定预测。这表明模型有效平衡了即时前缀信息与全局结构知识,实现了在短序列上的稳定泛化。尤为值得注意的是,对于序列长度变化极大且包含超长序列的复杂流程(如 Hospital

Billing), CAM-GF 表现出卓越的鲁棒性。模型不仅在短前缀上保持高准确率,在长序列(长度>50)场景下其性能更是趋近于饱和,有效克服了长程依赖中的信息稀释问题。

本节的多指标评估的结果证实了其在处理类别不平衡数据时的性能鲁棒性,而前缀长度敏感性分析则证明了其能够稳定地适应流程在不同执行阶段的信息变化。这两种特性表明了该框架在复杂、动态的真实世界场景中具有高可靠性。

### 5.3.3 消融实验

为精准量化 CAM-GF 框架中各核心组件的边际贡献,本节设计了系统的消融实验,旨在隔离并评估框架内空间编码、时序建模以及动态融合机制的独立效能。

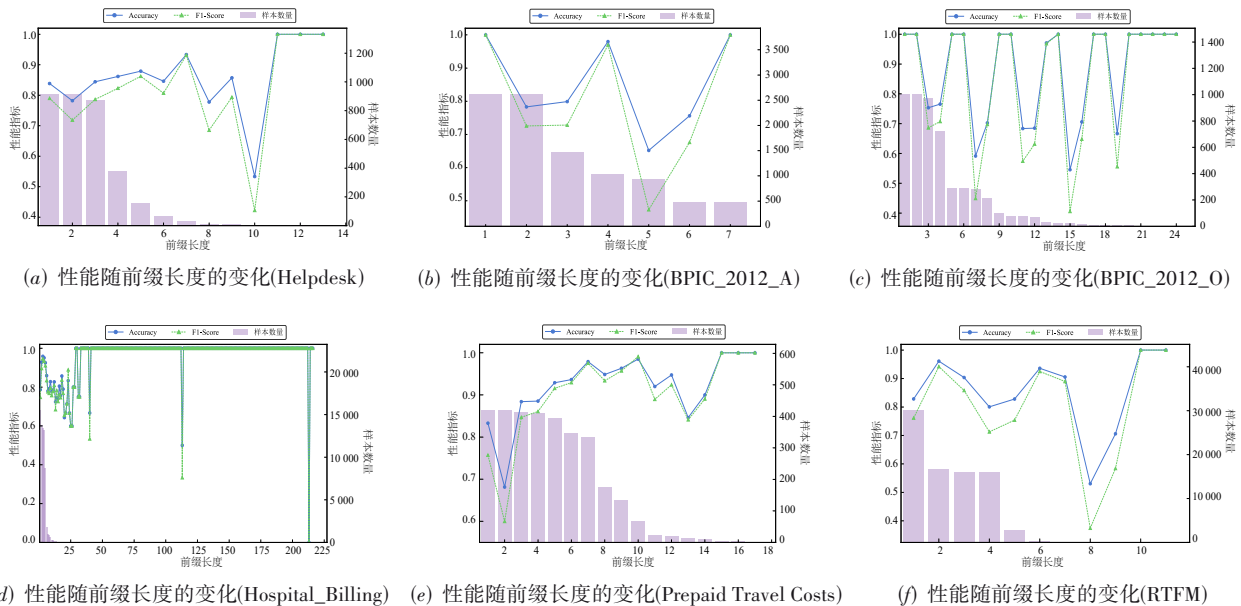


图 4 不同前缀长度下模型预测性能变化

实验涵盖了三个关键的框架变体：“CAM-GF（静态融合）”（采用简单的平均池化进行视角融合）、“仅GAT”（仅空间编码模块）和“仅Transformer”（仅时序建模模块）。所有模型的性能指标均为多次独立实验的均值，详细结果如表4所示。

通过对表4进行分析，可以得到以下结论。首先，通过比较“仅GAT”和“仅Transformer”这两个单一组件模型，可以发现二者各有侧重。在多数数据集中，“仅Transformer”凭借其强大的时序捕捉能力表现更优；但在如Prepaid Travel Costs这类结构化较强的流程中，“仅GAT”亦展现出极具竞争力的性能，这可能是因为在该类流程中，活动间的转移逻辑更多地由固定的结构化关系而非长距离时序模式所主导，而这正是GAT模型所擅长捕捉的。尽管各有侧重，但单一组件模型均无法企及完整框架的性能高度，这有力地证实了业务流程预测的复杂性，单纯依赖空间结构信息或时序依

赖信息均非最优解，二者的深度融合是必要的。

在此基础上，为证明信息融合的方式是决定模型性能上限的关键，实验进一步对比了完整CAM-GF与“CAM-GF（静态融合）”的性能差异。结果表明，在所有6个数据集上，采用上下文感知动态融合的CAM-GF均优于其静态融合版本。值得注意的是，这种性能增益在不同特性的数据集上表现各异，尤其是在Hospital Billing这一流程最长、变化最复杂的挑战性数据集上，动态融合带来了3.65个百分点的性能提升，远高于在其他数据集上的增益。这一现象暗示，当流程越复杂、上下文变化越丰富时，模型根据实时前缀自适应地调整视角权重的能力就越发关键。该消融实验的结果，从实证层面系统地验证了本框架设计的合理性。实验不仅证明了时空联合建模的必要性与多视角信息的有效性，更通过静态与动态融合的直接对比，将性能突破的根本原因指向了“上下文感知的动态融合机制”这一框架的核心理论。

表4 使用不同视角的预测性能对比

单位：%

数据集	CAM-GF(本文)	CAM-GF(静态融合)	仅GAT	仅Transformer
Helpdesk	86.78	83.01	78.66	81.94
BPIC_2012_A	78.58	76.34	74.66	75.91
BPIC_2012_O	84.16	82.38	81.55	81.85
Hospital Billing	93.20	89.55	82.81	88.19
Prepaid Travel Costs	88.50	85.36	84.89	81.85
RTFM	83.45	82.61	79.18	80.34

### 5.3.4 可解释性分析

除了卓越的定量性能，CAM-GF框架的核心优势在于其固有的可解释性。通过其上下文感知的注意力机制，能够洞察模型在进行预测时，是如何动态权衡不同过程视角的。本节以Helpdesk数据集为例，通过对4个代表性案例（Case 1995、Case 2301、Case 1097 和 Case 3031）的可视化分析（如图5所示），来揭示模型动态融合策略的内部工作机理。

首先，观察案例1995的注意力权重演化过程可以发现模型于活动“Take in charge ticket”之后进行预测时，发生了一次错误（预测为“Resolve ticket”，而实际应为“Wait”）。作为响应，模型权重分布发生了显著调整：它大幅提升了对长期依赖视角（Long View）的关注度，同时衰减了短期依赖（Short View）和严重程度（Seriousness View）等其他视角的权重。这种调整策略表明，当基于局部上下文的预测失败后，模型能自主地将信任度从直接的、短期的依赖关系，迁移到对流程整体宏观结构的理解上。正是这次精准的权重再分配，使得模型在后续步骤中恢复了准确性。案例2301也验证了类似的纠错模式，在一次错误预测后，模型同样提升了长期依赖视角的权重，并降低了先前占优的客户（Customer View）视角权重，展现了其在不同情境下回归全局视角的核心纠错能力。

然而，模型的自适应能力并非一成不变。在案例1097中，一次预测错误后，模型并未像之前那样优先提升长期依赖视角，而是逐步提高了资源视角（Resource View）的权重，并借此在后续取得了正确的预测。这深刻地表明，模型能够根据特定案例的上下文，判断出“由谁执行”（即资源）是比流程通用模式更关键的预测因子。更为复杂的适应性体现在案例3031中。该案例在预测中遭遇了两次连续错误。第一次错误后，模型尝试提升长期依赖视角但未能奏效。这次连续的失败触发了更深层次的策略调整：模型转而同时提升了短期依赖视角（Short View）与长期依赖视角的权重，并降低了其他数据属性的权重。这说明模型能够意识到单一的调整策略不足以应对高度模糊的场景，从而采取一种“远近兼顾”的复杂策略来稳定后续的预测。

这4个案例共同揭示了CAM-GF框架并非一个静态的“黑箱”。它能够执行一种实例级别、由预测反馈驱动的实时权重校准。当模型基于当前的权重策略做出错误判断时，其注意力机制会进行“反思”和“修正”，将焦点转移到可能更有价值的信息维度上。这种动态、透明的决策过程正是“上下文感知”的实践体现，它不仅是模型取得高精度的关键，更为流程分析人员深入理解模型决策提供了直观的洞察。

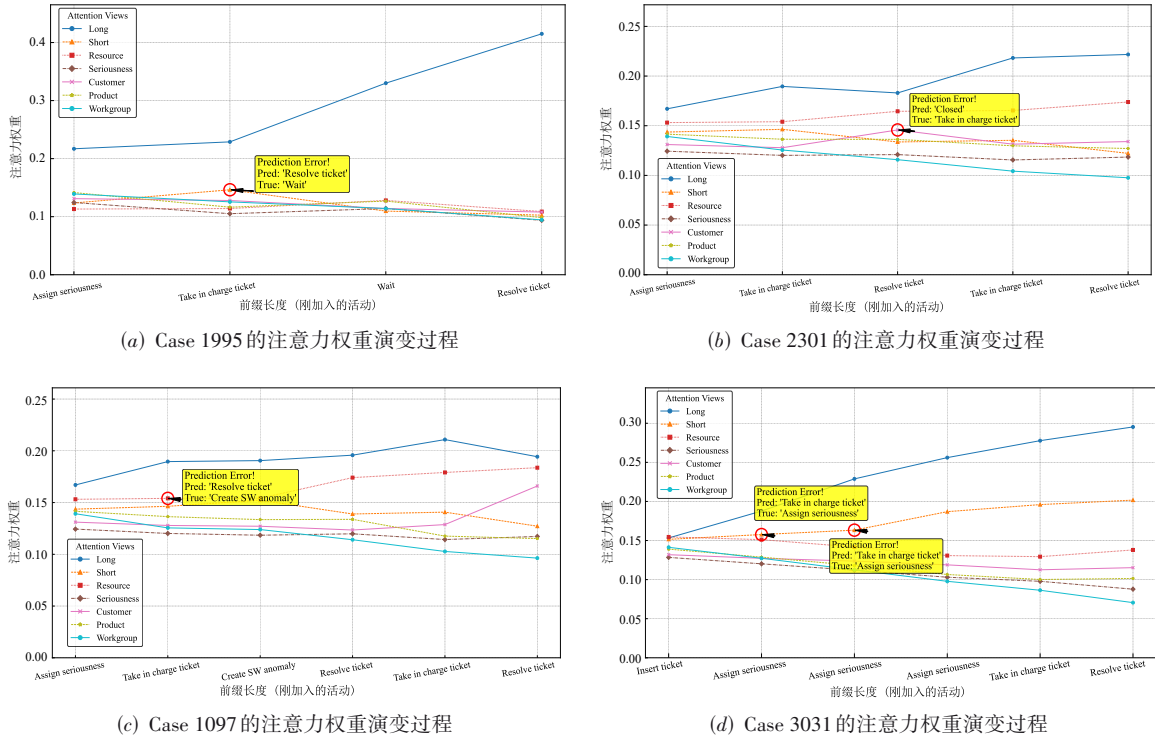


图5 不同视角关注度动态变化图

5.3.5 复杂度与效率分析

为全面评估CAM-GF框架的实用性,本节将从效率和理论复杂度两个方面,对模型的计算成本进行分析.

(1)效率评估

本节在与前述实验相同的硬件环境下,对CAM-GF(完整模型)及其两个核心消融变体(“仅GAT”和“仅Transformer”)进行效率测试.为提供具有代表性的综合性度量,表5展示了各模型在全部6个数据集上的平均性能指标.

从表5的平均结果可以看出,CAM-GF作为完整框架,其各项计算开销在三者中是最高的.这符合其融合了空间与时间双重建模能力的预期.值得注意的是,仅GAT模型的浮点运算次数(Floating-point Operations, FLOPs)远低于仅Transformer,但其实际推理时延反而更高.这揭示了一个有趣的现象:GAT所依赖的邻居聚合等操作,在当前GPU硬件上的优化效率可能不如Transformer所依赖的高度并行化的大规模矩阵运算.

表5 模型平均计算开销对比

模型	参数量/M	FLOPs/G	推理时延/ms	显存/MB
仅Transformer	4.87	0.038	2.91	156
仅GAT	0.94	0.004	6.11	82.5
CAM-GF	5.96	0.041	12.67	216

此外,CAM-GF的参数量低于仅Transformer和仅GAT的直接相加,这主要得益于框架内核心嵌入层的共享,体现了其设计的参数有效性.综上,尽管CAM-GF在计算资源上高于其他两个模型,但其展现出的显著预测精度优势,表明了这种为深度融合时空信息而增加的计算成本,在许多追求高精度的应用场景下是合理且有价值的.

(2)理论复杂度分析

CAM-GF的理论时间复杂度主要由其两个核心计算阶段决定:多视角图编码和时序依赖建模.首先,在多视角图编码阶段,该阶段包含 $V$ 个并行的GAT编码器.对于一个含有 $L_g$ 层的GAT,处理一个包含 $N$ 个节点和 $E$ 条边的视角图,其单层复杂度约为 $O(N \cdot d_{in} \cdot d_{out} + |E| \cdot d_{out})$ ,其中 $d$ 为特征维度.因此,此阶段的总复杂度与视角数 $V$ 、节点数 $N$ 、总边数 $E_{total}$ 呈大致线性关系.其次,在时序依赖建模阶段,该阶段由一个含有 $L_l$ 层的Transformer编码器主导.对于一个长度为 $k$ 的输入前缀序列,其核心自注意力机制的复杂度为 $O(k^2 \cdot d)$ .因此,此阶段的复杂度与前缀长度 $k$ 呈平方关系.

综合来看,CAM-GF的总体时间复杂度可表示为

$$O(L_g \cdot V \cdot (N + |E_{total}|) \cdot d + L_l \cdot k^2 \cdot d) \quad (13)$$

由此可见,模型的可扩展性主要受到两个因素的

制约:对于包含大量活动( $N$ 大)或关系极其复杂( $|E|$ 大)的流程,图编码阶段将成为需要突破的瓶颈;而对于需要处理极长执行序列( $k$ 大)的流程,Transformer的时序建模阶段将成为主要开销.这一分析为在不同资源约束下应用和优化本框架提供了理论指导.

## 6 结束语

现有预测性流程监控技术,特别是在使用基于图的深度学习模型预测未来状态时,通常因依赖单一的过程视角和采用静态的信息融合策略,而导致预测精度和泛化能力受限.为解决此问题,本文提出了一种新颖的、基于上下文感知多视角图融合的预测框架(CAM-GF).该框架首先从事件日志中并行构建一个涵盖控制流、资源和数据等维度的多视角过程图谱,以克服单视角带来的信息鸿沟;接着,引入一种创新的上下文感知注意力机制,该机制能根据每个流程实例的实时前缀信息,动态地学习并分配不同视角的融合权重,解决了静态融合缺乏适应性的难题;最后,将深度融合后的综合表示送入Transformer模型以捕捉长时间依赖关系,从而作出精准预测.在多个真实数据集上的大量实验证明,本文提出的CAM-GF框架在预测准确率上显著优于主流基线方法,并且其注意力权重为多视角融合过程提供了直观的可解释性.

尽管本文提出的CAM-GF框架取得了令人鼓舞的结果,但仍存在一些局限性与值得探索的未来方向.首先,本框架采用的图注意力网络虽然对数据噪声具有一定的鲁棒性,但如何量化和处理高噪声环境下的语义视角构建,仍是一个值得未来深入研究的方向.其次,对于包含海量活动和案例的超大规模事件日志,多视角图的构建和GNN的计算开销可能较大,未来可研究图采样或更高效的图表示学习技术以提升模型的可扩展性.然后,当前模型的超参数(如嵌入维度、注意力头数等)需要为不同数据集进行人工调优,未来可探索自动化机器学习(AutoML)技术,以实现超参数的自适应寻优.最后,本研究假设业务流程是静态的,而真实世界的流程可能随时间发生演变(即概念漂移),未来的工作可以将模型扩展为能够在线更新和适应流程变化的动态框架.此外,将此多视角动态融合思想应用到结果预测、剩余时间预测等其他流程挖掘任务中,也是一个极具价值的研究方向.

## 参考文献

- [1] NEU D A, LAHANN J, FETTKE P. A systematic literature review on state-of-the-art deep learning methods for process prediction[J]. *Artificial Intelligence Review*, 2022, 55(2): 801-827.
- [2] RAMA-MANEIRO E, VIDAL J C, LAMA M. Deep learning for predictive business process monitoring: Review and benchmark[J]. *IEEE Transactions on Services Computing*, 2023, 16(1): 739-756.
- [3] TAX N, VERENICH I, LA ROSA M, et al. Predictive business process monitoring with LSTM neural networks[M]// *Advanced Information Systems Engineering*. Cham: Springer, 2017: 477-492.
- [4] CAMARGO M, DUMAS M, GONZÁLEZ-ROJAS O. Learning accurate LSTM models of business processes[M]// *Business Process Management*. Cham: Springer, 2019: 286-302.
- [5] DI FRANCESCO MARINO C, GHIDINI C, MAGGI F M, et al. An eye into the future: Leveraging a-priori knowledge in predictive business process monitoring[M]// *Business Process Management*. Cham: Springer International Publishing, 2017: 252-268.
- [6] 徐兴荣, 刘聪, 李婷, 等. 基于双向准循环神经网络和注意力机制的业务流程剩余时间预测方法[J]. *电子学报*, 2022, 50(8): 1975-1984.  
XU X R, LIU C, LI T, et al. Business process remaining time prediction: An approach based on bidirectional quasi recurrent neural network with attention[J]. *Acta Electronica Sinica*, 2022, 50(8): 1975-1984. (in Chinese)
- [7] WICKRAMANAYAKE B, HE Z P, OUYANG C, et al. Building interpretable models for business process prediction using shared and specialised attention mechanisms[J]. *Knowledge-Based Systems*, 2022, 248: 108773.
- [8] RAMA-MANEIRO E, VIDAL J C, LAMA M. Embedding graph convolutional networks in recurrent neural networks for predictive monitoring[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2024, 36(1): 137-151.
- [9] WU Z H, PAN S R, CHEN F W, et al. A comprehensive survey on graph neural networks[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2021, 32(1): 4-24.
- [10] 万升, 杨健, 宫辰. 基于图神经网络的高光谱图像分类研究进展[J]. *电子学报*, 2023, 51(6): 1687-1709.  
WAN S, YANG J, GONG C. Advances of hyperspectral image classification based on graph neural networks[J]. *Acta Electronica Sinica*, 2023, 51(6): 1687-1709. (in Chinese)
- [11] VENUGOPAL I, TÖLLICH J, FAIRBANK M, et al. A comparison of deep-learning methods for analysing and predicting business processes[C]// *2021 International Joint Conference on Neural Networks*. Piscataway: IEEE, 2021: 1-8.
- [12] WEINZIERL S. Exploring gated graph sequence neural networks for predicting next process activities[M]// *Business Process Management Workshops*. Cham: Springer International Publishing, 2022: 30-42.
- [13] AMIRI ELYASI K, VAN DER AA H, STUCKENSCHMIDT H. PGTNet: A process graph Transformer network for remaining time prediction of business process in-

stances[M]//Advanced Information Systems Engineering. Cham: Springer Nature Switzerland, 2024: 124-140.

- [14] BUKHSH Z A, SAEED A, DIJKMAN R M. ProcessTransformer: Predictive business process monitoring with transformer network[EB/OL]. (2021-04-01)[2025-06-30]. <https://arXiv.org/abs/2104.00721>.
- [15] DE SMEDT J, DE WEERDT J. Predictive process model monitoring using long short-term memory networks[J]. Engineering Applications of Artificial Intelligence, 2024, 133: 108295.
- [16] TAYMOURI F, LA ROSA M, ERFANI S, et al. Predictive business process monitoring via generative adversarial nets: The case of next event prediction[EB/OL]. (2020-04-01)[2025-06-30]. <https://arXiv.org/abs/2003.11268>.
- [17] EVERMANN J, REHSE J R, FETTKE P. Predicting process behaviour using deep learning[J]. Decision Support Systems, 2017, 100: 129-140.
- [18] CHIORRINI A, DIAMANTINI C, GENGA L, et al. Multi-perspective enriched instance graphs for next activity prediction through graph neural network[J]. Journal of Intelligent Information Systems, 2023, 61(1): 5-25.
- [19] WANG J X, LU C L, YU Y F, et al. HiGPP: A history-informed graph-based process predictor for next activity[M]//Service-Oriented Computing. Singapore: Springer Nature Singapore, 2024: 337-353.
- [20] PASQUADIBISCEGLIE V, SCARINGI R, APPICE A, et al. PROPHET: Explainable predictive process monitor-

ing with heterogeneous graph neural networks[J]. IEEE Transactions on Services Computing, 2024, 17(6): 4111-4124.

- [21] 李鑫, 陆伟, 马召祎, 等. 基于图注意力和改进Transformer的节点分类方法[J]. 电子学报, 2024, 52(8): 2799-2810.
- LI X, LU W, MA Z Y, et al. A node classification method based on graph attention and improved Transformer[J]. Acta Electronica Sinica, 2024, 52(8): 2799-2810. (in Chinese)
- [22] VAN DER AALST W. Data science in action[M]//Process Mining. Berlin, Heidelberg: Springer, 2016: 3-23.
- [23] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook: Curran Associates Inc., 2017: 6000-6010.
- [24] SINDHGATTA R, MOREIRA C, OUYANG C, et al. Exploring interpretable predictive models for business processes[M]//Business Process Management. Cham: Springer, 2020: 257-272.
- [25] DONG L L, LIU C, REN C G. GCN-ONLSTM: Process next event prediction method based on spatio-temporal feature fusion[J]. International Journal of Science and Engineering Applications, 2023, 12(3): 1-4.
- [26] WANG J X, YU Y F, FANG N, et al. MHG-predictor: A multi-layer heterogeneous graph-based predictor for next activity in complex business processes[C]//Proceedings of the ACM on Web Conference 2025. New York: ACM, 2025: 500-509.

## 作者简介



**张希为** 男, 1999年4月出生于山东省泰安市. 现为安徽理工大学数学与大数据学院博士研究生. 主要研究方向为预测性流程监控、图神经网络等.

E-mail: zxw@aust.edu.cn



**方贤文** 男, 1975年10月出生于河南省信阳市. 现为安徽理工大学数学与大数据学院教授、博士生导师. 主要研究方向为流程挖掘、Petri网等.

E-mail: xwfang@aust.edu.cn



**毛古宝** 男, 1997年7月出生于安徽省安庆市. 现为安徽理工大学数学与大数据博士研究生. 主要研究方向为预测性流程监控、异常检测等.

E-mail: mgb@aust.edu.cn