计 算 机 研 究 与 发 展
Journal of Computer Research and Development
https://crad.ict.ac.cn
63(9): 2215−2230, 2026

语义对齐与自适应回放增强的动态多模态知识图谱持续补全

王 迪[1] 杜军平[1] 薛 哲[1] 叶冠华[1] 赵 桐[1] 孟晓龙[1] 李海生[2]
[1](北京邮电大学计算机学院(国家示范性软件学院) 北京 100876)
[2](北京工商大学计算机与人工智能学院 北京 100048)
(wdd123@bupt.edu.cn)

Continual Completion for Dynamic Multimodal Knowledge Graphs via Semantic Alignment and Adaptive Replay Enhancement

Wang Di[1], Du Junping[1], Xue Zhe[1], Ye Guanhua[1], Zhao Tong[1], Meng Xiaolong[1], and Li Haisheng[2]
[1](*School of Computer Science*(*National Pilot School of Software Engineering*), *Beijing University of Posts and Telecommunications*, *Beijing* 100876)
[2](*School of Computer and Artificial Intelligence*, *Beijing Technology and Business University*, *Beijing* 100048)

Abstract To meet the demand for dynamic knowledge updating and long-term knowledge retention in knowledge-enhanced artificial intelligence, dynamic multimodal knowledge graph continual completion infers missing facts incrementally as graph snapshots evolve while keeping historical knowledge stable. It thus provides fundamental support for dynamic knowledge updates. Existing methods predominantly focus on single-modal structural information, failing to fully exploit textual and visual semantics; directly incorporating multimodal information may introduce unexpected semantic shifts across snapshots, thereby exacerbating catastrophic forgetting. To this end, we propose SAREC, a semantic alignment and adaptive replay enhanced method for dynamic multimodal knowledge graph continual completion. SAREC leverages large language models to generate entity semantic descriptions and combines them with vision-language pretrained representations to construct a multimodal discrepancy-aware fusion mechanism, enhancing the completeness and robustness of entity representations. A multimodal temporal semantic alignment module is further introduced to constrain cross-snapshot representation evolution and maintain semantic consistency. Finally, a temporal-aware adaptive replay mechanism is proposed, which dynamically allocates the replay budget according to snapshot recency and selectively replays historical knowledge to alleviate catastrophic forgetting. Experiments demonstrate that SAREC achieves superior performance under various knowledge evolution modes, providing effective support for continual knowledge injection and adaptive knowledge enhancement in open dynamic environments.

Key words multimodal knowledge graph completion; continual learning; dynamic knowledge updating; large language model; catastrophic forgetting

摘 要 面向知识增强人工智能中动态知识更新与长期知识保持的需求,动态多模态知识图谱持续补全随图谱快照演化增量推断缺失事实、保持历史知识稳定,为动态知识更新提供基础支撑。现有持续补全方法多关注单一模态,难以充分利用文本与视觉语义;直接引入多模态信息又易导致跨快照语义非预期偏移,进而加剧灾难性遗忘。为此,提出语义对齐与自适应回放增强的动态多模态知识图谱持续补全方

收稿日期: 2026-06-06; **修回日期**: 2026-07-17
基金项目: 国家自然科学基金项目 (62192784, U22B2038, 62532002, 62422202)
This work was supported by the National Natural Science Foundation of China (62192784, U22B2038, 62532002, 62422202).
通信作者: 杜军平(junpingdu@126.com)

法 SAREC。利用大语言模型生成实体语义描述,并结合视觉-语言预训练表征构建多模态差异感知融合机制,提升实体表示的完整性与鲁棒性;进而通过多模态时序语义对齐,对实体表征的跨快照演化进行约束,保持内在语义一致性;最后,提出时序感知自适应回放机制,依据快照时序动态分配回放预算并选择性回放历史知识,缓解灾难性遗忘。实验结果表明,SAREC 在多种知识演化模式下取得优于基线方法的性能,为开放环境下的持续知识注入与自适应知识增强提供了有效支撑。

关键词 多模态知识图谱补全;持续学习;知识动态更新;大语言模型;灾难性遗忘

中图法分类号 TP391

DOI: 10.7544/issn1000-1239.202660445 **CSTR:** 32373.14.issn1000-1239.202660445

随着大模型在多任务场景中的广泛应用,外部知识资源的动态更新、持续注入与可靠调用[1]逐渐成为提升智能系统事实一致性和推理稳定性的关键环节。知识图谱以结构化方式存储大规模事实知识,是支撑大模型进行知识增强、提升推理可靠性的核心外部知识源。然而,真实世界的知识图谱往往是不完备的[2],并且,相较于静态知识库,真实环境中的知识图谱会随时间不断演化[3],持续产生新的实体、关系与事实,既有知识也可能发生扩展和重组。传统的静态知识图谱补全方法[4-7]在面对此类动态场景时,要么需要全量重训练、代价高昂,要么会因增量训练而引发灾难性遗忘[8],难以支撑大模型对最新事实的动态更新与持续注入。因此,需要进行持续知识图谱补全(continual knowledge graph completion, CKGC),以增量方式融入新知识,同时保留已学习的历史表征,这是实现知识动态演化与可持续注入的关键路径。

现有 CKGC 方法主要遵循 3 类技术路线:1)基于动态架构扩展的方法,通过扩充网络分支、插入轻量化模块或按需调整结构规模等手段,为持续涌入的新知识开辟独立的承载空间,使新旧知识在不同的结构区域中各司其职,从根本上规避参数共享所导致的相互覆盖,代表性技术有渐进式神经网络[9]、FastKGE[10]、SAGE[11]等;2)基于参数正则化的方法,该类方法不改变既有网络结构,而是从参数敏感性的角度量化每个参数对历史知识的贡献度,进而在模型更新时对关键参数施加方向性约束,以此在统一参数空间内实现新旧知识的协同保留,代表性技术有 EWC[12]、LKGE[13]、IncDE[14]、DebiasedKGE[15]等;3)基于知识回放的方法,从训练数据的角度缓解遗忘,其关键在于挑选并保存具有代表性的历史样本,或借助生成机制重构近似的历史数据分布,并在后续阶段将这些样本与新数据交替输入模型进行联合训练,使模型在持续接触历史信息的过程中巩固既有知识,代表性技术有 GEM[16]、EMR[17]、DGAR[18]、ERPP[19]等。这些方法虽然能够在一定程度上适应图谱快照(snapshot)间的结构变化,但大多聚焦于演化图谱中显式的结构变化,忽视了实体本身所承载的内在语义[20]。理想情况下,实体的局部连接结构可能随时间发生明显变化,而其内在语义应在其局部图结构动态变化时保持相对稳定。因此,若仅依赖结构信息进行持续建模,模型容易将图结构扰动误判为知识语义变化,进而引发表征非预期偏移、历史知识保持不足与跨阶段预测不稳定等问题。特别是在复杂知识持续注入场景中,缺乏对实体语义稳定性的显式建模,会削弱模型对图谱长期演化规律的刻画能力。

与此同时,现实中的知识图谱日益呈现出多模态属性和动态演化的双重特征。如图 1 所示,动态多模态知识图谱由连续演化的图谱快照构成,实体不仅通过结构化三元组进行刻画,还伴随着文本描述与视觉图像等多模态辅助信息[21],这些多模态信号为实体表示学习提供了重要的互补语义线索[22]。随着大语言模型[23]与视觉-语言预训练模型[24]的快速发展,自动获取并整合实体的文本描述与图像表征已变得切实可行,为知识图谱的补全更新注入更丰富的外部知识。然而,现有多模态知识图谱补全方法主要面向静态场景设计,如 SAFER[25],缺乏适应图谱持续演化的能力;MSCGL[26]虽率先将多模态信息引入持续知识图谱学习,但其任务设置仅限于实体分类,并非面向知识图谱补全。此外,面向多模态知识图谱持续补全的公开基准仍不充分,现有数据集难以同时覆盖多模态信息和知识动态接入等关键需求,这为该领域的方法评估与比较带来了额外挑战。更重要的是,将多模态信号直接用于持续学习,还会面临模态贡献不均衡[27]以及不同时间快照之间语义分布变化[28]等问题,这可能进一步放大跨快照表征非预期偏移,使模型难以在知识动态更新中同时兼顾新知识适应与旧知识保持。

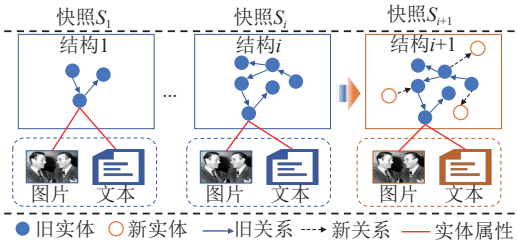


Fig. 1 Illustration of dynamic multimodal knowledge graph

图 1 动态多模态知识图谱示意图

基于上述讨论，本文认为有效的多模态知识图谱持续补全不仅需要保持结构上的连续性，更需在跨快照的动态演化中维护内在语义一致性，从而为大模型在高可靠性场景下的知识动态更新与注入提供稳定的外部知识支撑。因此，本文提出一种语义对齐与自适应回放增强的动态多模态知识图谱持续补全方法 SAREC（semantic alignment and adaptive replay enhancement for continual completion, SAREC），以应对开放环境知识演化场景下的灾难性遗忘与多模态语义非预期偏移两大挑战。具体而言，SAREC 首先利用大语言模型生成实体语义描述，并结合视觉-语言预训练表征，构建大语言模型引导的多模态差异感知融合机制，使模型能够根据实体特性自适应地融合各模态信息，在模态异构或缺失情形下仍可获得鲁棒表征；进一步设计多模态时序语义对齐策略，对实体与关系表征在相邻快照间的演化进行正则约束；最后，提出时序感知自适应回放机制，依据快照新近度动态分配回放预算，并选择性回放预测差异较大的历史三元组，从而重点保护最易被遗忘的知识。多模态时序语义对齐与时序感知自适应回放的协同设计，使 SAREC 既能将历史知识平稳传递至新引入的实体与关系，又能有效抑制语义非预期偏移，从而在持续演化的知识图谱中实现稳定有效的知识补全与持续注入。

本文主要贡献如下：

1）提出大语言模型引导的多模态差异感知融合机制，利用大语言模型生成实体语义描述，结合视觉-语言预训练表征实现差异感知融合，提升多模态实体表示的完整性与鲁棒性。

2）设计多模态时序语义对齐方法。对实体表征的跨快照演化进行时序约束，维护多模态实体内在语义一致性，避免跨快照语义非预期偏移，为开放场景下的知识动态更新提供稳定支撑。

3）提出时序感知自适应回放机制。依据快照新近度动态分配回放预算，以预测差异为依据选择性回放最易被遗忘的历史三元组，有效缓解持续演化场景下的灾难性遗忘问题，为知识增强人工智能在动态演化场景下的应用提供了可行路径。

1 相关工作

1.1 持续知识图谱补全

知识图谱补全旨在基于已知事实推理缺失三元组[29]，在开放环境知识演化场景中[30]，持续知识图谱补全进一步以增量方式更新实体与关系表征，并保留对历史知识的有效记忆，从而支撑大模型在演化环境下的知识动态更新与持续注入。根据缓解遗忘的策略不同，现有持续知识图谱补全方法大致可分为 3 类：基于动态架构扩展的方法、基于参数正则化的方法和基于知识回放的方法。

基于动态架构扩展的方法[9-11]通过扩展或调整模型结构来容纳新知识。早期工作如渐进式神经网络[9]为每个新任务引入独立的子网络分支，避免对已有参数的覆盖。近期面向知识图谱的工作进一步发展了这一思路：FastKGE[10]提出增量低秩适配器机制，将新知识隔离到特定层中以平衡训练效率与遗忘抑制；SAGE[11]则提出尺度感知的渐进式演化策略，根据知识规模动态调整模型容量。这类方法在避免全量重训练的同时提升了对动态环境的适应能力，但模型参数随快照数量增加而持续膨胀，难以适用于长期演化的大规模知识图谱场景。

基于参数正则化的方法[12-15]通过对模型参数的更新过程施加约束，限制其相对于已学习表征的过度偏移，从而保留历史知识。EWC[12]利用 Fisher 信息矩阵识别参数的重要性，并对重要参数施加更强的更新惩罚。LKGE[13]将其扩展至知识图谱领域，提出面向知识图谱终身学习的嵌入迁移框架；2024 年，IncDE[14]引入增量蒸馏机制，通过教师-学生范式约束新旧表征的演化；DebiasedKGE[15]进一步指出，传统正则化方法存在“伪遗忘”现象，模型在保留旧表征的同时反而抑制了对新模式的学习能力，并提出双视图正则化的知识对齐策略以缓解该问题。尽管正则化方法在参数效率上具有优势，但仅通过约束参数更新难以充分捕获新增知识中的复杂语义结构[31]。

基于知识回放的方法[16-19]通过在新任务训练中回放部分历史样本来缓解灾难性遗忘。GEM[16]以梯度情景记忆为代表，将历史样本作为约束保证梯度更新方向不损害旧任务性能；EMR[17]工作通过句子级表征对齐实现长期关系抽取中的知识保留；2025

年, DGAR[18] 将生成式自适应回放引入时序知识图谱推理, 通过合成历史样本替代真实存储以降低记忆开销; ERPP[19] 则通过演化关系路径传递机制, 将“知识遗忘”转化为“知识累积”, 利用路径级表征实现终身知识图谱嵌入。回放类方法在缓解遗忘方面效果显著, 但其性能高度依赖回放样本的选择策略与回放预算的分配方式, 简单的均匀回放难以重点保护最易被遗忘的关键知识。

总体而言, 现有持续知识图谱补全方法主要聚焦于演化图谱中显式的结构适应问题, 而对实体表征在跨快照演化过程中的语义一致性问题仍缺乏系统性的探索。理想情况下, 实体的内在语义应在其局部图结构动态变化时保持相对稳定, 但单一模态下的结构信号难以有效刻画这种内在语义不变性, 这限制了模型在复杂演化场景下对知识冲突与语义非预期偏移的应对能力。

1.2 多模态知识图谱嵌入

为缓解单一模态信号的表达局限, 多模态知识图谱嵌入(multimodal knowledge graph embedding, MKGE)通过引入文本描述[32]、视觉图像[33] 等辅助模态信息来增强实体表征。该方向的核心挑战在于如何在统一的语义空间中协同建模结构、文本与视觉信息, 并应对模态异构与模态缺失等问题。

现有 MKGE 方法大多面向静态知识图谱设计, 如 APKGC[34] 提出噪声增强的注意力惩罚机制以缓解多模态融合中的特征污染。这类方法通常默认训练数据一次性给定, 并在固定实体、关系和模态集合上学习统一表示。随着视觉-语言预训练模型的快速发展, 以 CLIP[35] 为代表的对比学习框架在大规模图文对上学习到了对齐良好的跨模态表征, 为多模态知识图谱嵌入提供了高质量的预训练特征来源。近期工作如 VL-KGE[36] 进一步将视觉-语言模型表征系统性地融入知识图谱嵌入框架, 显著提升了模型对模态异构和模态错位的鲁棒性。此外, 伴随大语言模型的兴起[37-38], 越来越多的研究开始借助大语言模型为知识图谱生成丰富的文本描述[39], 以缓解知识图谱中实体文本信息稀疏的问题, 并增强多模态融合的语义基础。上述静态多模态方法充分验证了多模态信息对增强实体表征的价值, 但它们缺乏对图谱持续演化的适应能力, 难以直接迁移至持续学习场景。当知识库不断演化时, 直接在新快照上微调静态 MKGE 模型容易破坏已有实体的结构表示和多模态对齐关系, 进而导致旧知识遗忘和跨阶段表示不一致。

近年来, 一些研究开始尝试将 MKGE 与持续学习相结合: MSPT[40] 面向多模态知识图谱持续构建任务, 通过平衡稳定性与塑性来缓解多媒体数据处理中的遗忘问题, 但其目标更侧重于新知识抽取与图谱构建, 而非基于已有事实的缺失三元组推理; MSCGL[26] 结合神经架构搜索, 将多模态信息引入持续图学习框架以支持实体分类任务, 验证了多模态信息在持续图表示学习中的有效性, 但其任务设置与知识图谱补全中的链接预测仍存在明显差异; MMRotatH[41] 模型则面向未见模态的增量嵌入, 强调新模态知识向已有嵌入空间的迁移能力。

尽管上述工作已在多模态知识图谱持续学习方向上做出了有益探索, 但对快照演化的时序特性关注尚显不足, 且多模态信号与时序演化之间的联合建模仍有待深入。因此, 如何协同建模模态差异与时序演化, 并实现对历史知识的差异化保护, 仍是该领域亟待深入研究的问题。

2 任务说明

本节主要定义动态多模态知识图谱持续补全任务相关的概念, 以及多模态实体表征的形式化定义。

定义 1. 动态多模态知识图谱。动态多模态知识图谱被定义为一个按时间顺序排列的快照序列 $\mathcal{G}=\{\mathcal{S}_1,\mathcal{S}_2,\cdots,\mathcal{S}_I\}$, 其中每个快照 $\mathcal{S}_i=(\mathcal{E}_i,\mathcal{R}_i,\mathcal{T}_i)$ 表示知识图谱在第 i 步的演化状态, $\mathcal{E}_i$ 代表实体集, $\mathcal{R}_i$ 代表关系集, $\mathcal{T}_i$ 代表事实集。事实集中的每条事实被表示为一个三元组 $(h,r,t)\in\mathcal{T}_i$, 其中 $h,t\in\mathcal{E}_i$ 分别为头实体和尾实体, $r\in\mathcal{R}_i$ 为 2 个实体之间的关系。

与单模态知识图谱不同, 多模态知识图谱中的每个实体除了与图结构相关联, 还伴随着多种辅助模态信息。具体来说, 对于实体 $e\in\mathcal{E}_i$, 除了从图拓扑结构中学习到的结构嵌入 z_e^{struct} 之外, 还考虑其文本特征 x_e^{text} 和视觉特征 x_e^{img} 作为辅助语义信息。基于这 3 类模态信息, 实体 e 的最终多模态表征通过融合函数 f_{fusion} 得到, 即:

$$z_e=f_{\text{fusion}}(z_e^{\text{struct}},x_e^{\text{text}},x_e^{\text{img}}), \tag{1}$$

其中 f_{fusion} 为多模态融合函数, 用于将结构、文本与视觉信息整合为统一的实体表征。

动态多模态知识图谱随时间不断演化, 新的实体、关系与事实持续涌现, 对于任意相邻的 2 个快照 $\mathcal{E}_{i-1}$ 与 $\mathcal{E}_i$, 在第 i 个快照中新引入的实体、关系与事实分别被定义为: $\Delta\mathcal{E}_i=\mathcal{E}_i-\mathcal{E}_{i-1}$、$\Delta\mathcal{R}_i=\mathcal{R}_i-\mathcal{R}_{i-1}$ 和 $\Delta\mathcal{T}_i=\mathcal{T}_i-\mathcal{T}_{i-1}$。

基于上述动态多模态知识图谱的形式化定义，多模态知识图谱持续补全任务定义如下。

定义 2. 多模态知识图谱持续补全。给定一个动态多模态知识图谱的快照序列$\mathcal{G}=\{\mathcal{S}_1,\mathcal{S}_2,\cdots,\mathcal{S}_I\}$，多模态知识图谱持续补全任务在每个快照$\mathcal{S}_i$上以链接预测的形式进行，即对于形如$(h,r,?)$或$(?,r,t)$的待推理三元组，基于截至第$i$步所观测到的全部多模态知识，预测其缺失的目标实体。在快照$\mathcal{S}_i$上，模型可访问的新增信息包括新引入的事实$\Delta\mathcal{T}_i$、实体$\Delta\mathcal{E}_i$与关系$\Delta\mathcal{R}_i$；模型基于$\Delta\mathcal{T}_i$进行增量式更新，同时需要保留对历史实体集合$\mathcal{E}_{i-1}$与关系集合$\mathcal{R}_{i-1}$已学习到的表征。通过持续学习，实体与关系的表征被逐步精化，以与当前快照$\mathcal{S}_i$保持一致，并支撑动态演化场景下的可靠链接预测。

3 SAREC 方法介绍

本节详细介绍语义对齐与自适应回放增强的动态多模态知识图谱持续补全框架 SAREC，其整体框架如图 2 所示，主要由 3 个模块组成：大语言模型引导的多模态差异感知融合、多模态时序语义对齐、时序感知自适应回放。其中，大语言模型引导的多模态差异感知融合模块通过自适应整合结构嵌入与文本、视觉模态信息来构建鲁棒的实体表征；在此基础上，多模态时序语义对齐模块对多模态实体表征在快照间的演化进行约束以缓解跨快照的语义非预期偏移；为进一步稳定持续学习过程，提出时序感知自适应回放模块，依据快照新近度与预测差异动态回放历史三元组，有效保留历史知识并缓解灾难性遗忘。

3.1 大语言模型引导的多模态差异感知融合

在多模态知识图谱持续补全场景下，新引入的实体往往结构连接稀疏、邻域信息匮乏，仅依赖图结构难以学习到鲁棒的实体表征。为此，本文提出一种大语言模型引导的多模态差异感知融合范式：通过预训练的大语言模型与视觉-语言模型注入文本与视觉语义知识，并经由模态感知的注意力机制与实体感知的门控融合机制，将多模态语义与图结构信息在实体粒度上自适应整合，构建语义丰富、结构稳健的实体表征，增强模型在持续演化的开放知识环境中对复杂实体语义的理解与推理能力。

3.1.1 多模态语义注入

受面向持续学习的大语言模型语义增强方法[42]启发，本文采用预训练大语言模型为实体注入辅助文本知识，同时借助外部实体图像资源（如 MMKG[21]提供的实体图像集）为实体集合补充视觉模态信息。

具体来说，对于每个实体$e\in\mathcal{E}_i$，将大语言模型生成的简洁文本描述与其对应的实体图像分别送入预训练的视觉-语言模型编码器（如 CLIP[35]）进行编码。得益于视觉-语言模型在大规模图文对上的联合预训练，文本与图像 2 种模态被映射到同一共享语义空间，从而在同一向量空间中实现跨模态对齐。对于关联多张图像的实体，取其所有图像向量的均值作为该实体的视觉表征。对于缺失模态，则将缺失模态的输入向量置为零向量，仅作为批处理占位，并不代表“零语义”。本文在融合过程中引入模态可用性掩码，在注意力归一化前屏蔽缺失模态，并在加权求和时将其贡献置零。因此，缺失模态不会参与权重竞争，也不会由投影层偏置引入额外语义偏差。该处理既保证输入维度一致，也增强了模型对模态缺

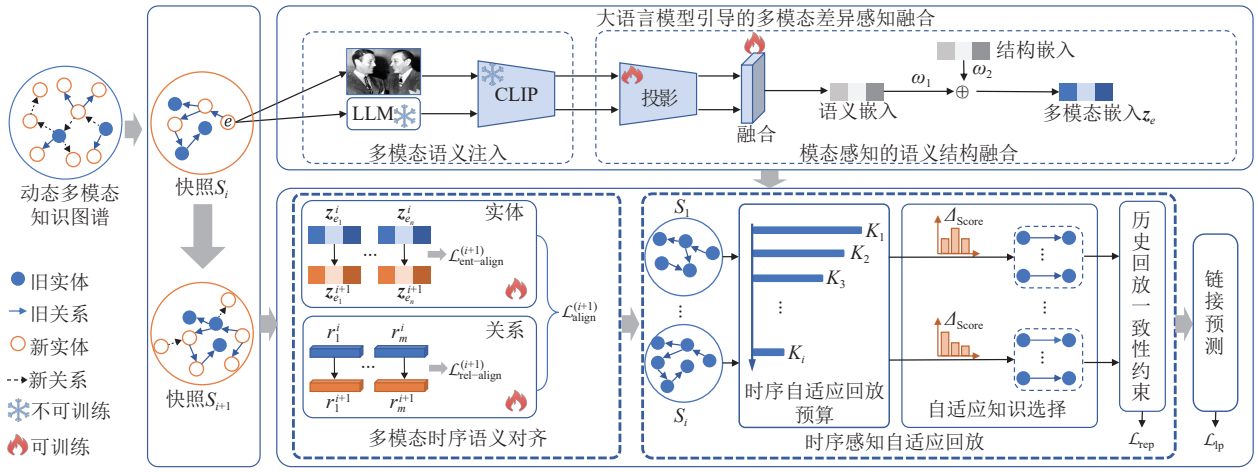


Fig. 2 SAREC framework for dynamic multimodal knowledge graph continual completion

图 2 动态多模态知识图谱持续补全的 SAREC 框架

失场景的鲁棒性。最终得到的文本与视觉表征分别记为$\boldsymbol{x}_e^{\text{text}} \in \mathbb{R}^{d_c}$和$\boldsymbol{x}_e^{\text{img}} \in \mathbb{R}^{d_c}$，其中$d_c$表示视觉-语言模型编码器输出向量的维度。

3.1.2 模态感知的语义-结构融合

在完成多模态语义注入后，需要进一步将文本与视觉模态的表征与结构嵌入空间对齐，并自适应地融合3类模态信息以获得鲁棒的实体表征。具体过程包括模态感知的语义融合与实体感知的语义-结构融合2个步骤。

1)模态感知的语义融合。为将文本与视觉特征进一步对齐到多模态知识图谱持续补全所使用的嵌入空间，设计一个共享的线性投影层，即：$\tilde{\boldsymbol{x}}_e^{\text{text}} = \boldsymbol{W}_{\text{s}}\boldsymbol{x}_e^{\text{text}} + \boldsymbol{b}_{\text{s}}$，$\tilde{\boldsymbol{x}}_e^{\text{img}} = \boldsymbol{W}_{\text{s}}\boldsymbol{x}_e^{\text{img}} + \boldsymbol{b}_{\text{s}}$。其中，$\boldsymbol{W}_{\text{s}} \in \mathbb{R}^{d\times d_c}$与$\boldsymbol{b}_{\text{s}} \in \mathbb{R}^d$是共享的可学习参数，$d$表示目标嵌入维度。考虑到不同实体在文本与视觉模态上的信息丰度差异显著，为自适应地捕获并利用2种模态的互补语义信息，本文采用一种轻量级的基于注意力的融合机制。具体来说，文本与视觉模态的注意力权重计算为：

$$(\alpha_{e,1},\alpha_{e,2})^{\text{T}} = softmax\left(\boldsymbol{W}_{\text{a}}[\tilde{\boldsymbol{x}}_e^{\text{text}} \| \tilde{\boldsymbol{x}}_e^{\text{img}}] + \boldsymbol{b}_{\text{a}}\right) \tag{2}$$

其中$\boldsymbol{W}_{\text{a}} \in \mathbb{R}^{2\times 2d}$和$\boldsymbol{b}_{\text{a}} \in \mathbb{R}^2$是可学习参数，$[\cdot\|\cdot]$表示向量拼接操作，$\alpha_{e,1}$与$\alpha_{e,2}$分别为文本与视觉模态对应的注意力权重。实体$e$的语义表征通过对2种模态的加权组合得到：

$$\boldsymbol{z}_e^{\text{sem}} = \alpha_{e,1}\tilde{\boldsymbol{x}}_e^{\text{text}} + \alpha_{e,2}\tilde{\boldsymbol{x}}_e^{\text{img}}, \tag{3}$$

其中$\boldsymbol{z}_e^{\text{sem}} \in \mathbb{R}^d$表示实体$e$经多模态融合后的语义嵌入。该机制使模型能够根据实体特性自适应地加权各模态信息，在某一模态信号较弱或缺失时，仍可获得鲁棒的语义表征。

2)实体感知的语义-结构融合。对于多模态知识图谱持续补全任务，实体表征不仅需要反映其在图谱中的结构信息，也需要刻画其内在的多模态语义。因此，本文进一步将语义表征与可训练的结构嵌入相融合，以获得最终的实体表征。将实体e的结构嵌入记为$\boldsymbol{z}_e^{\text{struct}} \in \mathbb{R}^d$，为在实体粒度上自适应地平衡结构信息与语义信息的贡献，本文采用一种实体感知的门控融合机制，最终实体表征定义为：

$$\boldsymbol{z}_e = g_e \cdot \boldsymbol{z}_e^{\text{struct}} + (1-g_e)\cdot \boldsymbol{z}_e^{\text{sem}}, \tag{4}$$

其中g_e为融合门控，由以下方式计算：

$$g_e = \sigma\left(\boldsymbol{w}_g^{\text{T}}[\boldsymbol{z}_e^{\text{struct}} \| \boldsymbol{z}_e^{\text{sem}}]\right), \tag{5}$$

其中$\boldsymbol{w}_g \in \mathbb{R}^{2d}$是可学习的参数向量，$\sigma(\cdot)$表示sigmoid激活函数。门控值$g_e \in (0,1)$由实体$e$的结构与语义信息共同确定：当$g_e$趋近于1时，模型更倚重结构嵌入；当$g_e$趋近于0时，模型则更多依赖多模态语义嵌入。这种实体感知的门控设计使结构与语义信息的贡献能够在实体粒度上动态调整，从而在持续演化的多模态知识图谱中获得稳定可靠的实体表征。

3.2 多模态时序语义对齐

在多模态知识图谱持续补全任务中，每个新快照所引入的实体、关系与事实会在增量训练过程中持续更新模型参数，直接引入多模态信息会导致历史实体与关系的表征发生非预期偏移。一方面，预训练文本与视觉特征由外部图文数据学习得到，具有相对稳定但与当前图谱关系模式不完全一致的语义分布；另一方面，图谱拓扑的局部剧烈变化会改变实体在链接预测中的上下文角色，若缺乏跨快照一致性约束，模型可能将这种结构角色变化误解释为实体内在语义变化，并在多模态投影与融合过程中放大弱相关模态或噪声模态的影响。因此，本文提出一种多模态时序语义对齐机制，对历史实体与关系在相邻快照间的表征施加一致性约束，在允许语义合理演化的同时，抑制无关扰动导致的表征退化。

时序对齐机制同时作用于实体与关系2个层面。$\mathcal{E}_i$和$\mathcal{R}_i$分别表示截至快照$\mathcal{S}_i$已观测到的实体集合与关系集合。记$\boldsymbol{z}_e^{(i)}$和$\boldsymbol{z}_e^{(i+1)}$分别为实体e在快照$\mathcal{S}_i$更新前后由3.1节得到的实体融合表征，$\boldsymbol{r}^{(i)}$与$\boldsymbol{r}^{(i+1)}$分别为关系r在相邻快照中的表征。

为约束历史实体表征在跨快照演化中的偏移，定义实体级知识对齐损失为：

$$\mathcal{L}_{\text{ent-align}}^{(i+1)} = \frac{1}{|\mathcal{E}_i|}\sum_{e\in\mathcal{E}_i}\left\|\boldsymbol{z}_e^{(i+1)} - \boldsymbol{z}_e^{(i)}\right\|_2^2。 \tag{6}$$

该损失通过最小化已观测实体在相邻快照间表征的ℓ_2距离，鼓励历史实体表征在快照更新后保持稳定。

为进一步保留关系层面的语义一致性，定义关系级知识对齐损失为：

$$\mathcal{L}_{\text{rel-align}}^{(i+1)} = \frac{1}{|\mathcal{R}_i|}\sum_{r\in\mathcal{R}_i}\left\|\boldsymbol{r}^{(i+1)} - \boldsymbol{r}^{(i)}\right\|_2^2。 \tag{7}$$

最终，多模态时序语义对齐的总体损失由实体级与关系级2项知识对齐损失加和得到：

$$\mathcal{L}_{\text{align}}^{(i+1)} = \mathcal{L}_{\text{ent-align}}^{(i+1)} + \mathcal{L}_{\text{rel-align}}^{(i+1)}。 \tag{8}$$

通过多模态时序语义对齐机制，以前一快照学习到的多模态表征为语义锚点，对当前快照的训练过程施加一致性约束，从而在增量学习中有效抑制历史表征的非预期偏移，保留已习得的语义知识。

3.3 时序感知自适应回放

不同于经验回放方法中常用的随机或均匀采样

策略[43]，本文提出一种时序感知与差异引导相结合的知识自适应回放机制：依据快照新近度自适应地分配回放预算，并基于预测分数差异选择最具回放价值的关键三元组，从而在有限的回放开销下高效缓解灾难性遗忘。

3.3.1 时序感知的回放预算分配

给定三元组的全局回放比例$\rho\in(0,1)$，首先确定回放的总预算$K=\lfloor\rho|\mathcal{T}_i|\rfloor$，$K$即在增量训练中用于回放的三元组总数，$|\mathcal{T}_i|$表示截至快照$\mathcal{S}_i$已观测到的三元组总数。考虑到越早的历史快照经历了更多次增量更新，其表征更易发生非预期偏移、更易被遗忘，因而需要分配更多的回放预算加以保护；而越晚的快照所学习到的知识尚处于较稳定的阶段，所需回放强度相对较低。为此，采用线性加权方案将K分配至各历史快照，快照$\mathcal{S}_j$的回放预算定义为：

$$K_j=\left\lfloor\frac{\frac{i-j+1}{i}K}{\sum_{k=0}^{i}(k+1)}\right\rfloor,j=0,1,\cdots,i。\tag{9}$$

3.3.2 自适应关键三元组选择

在确定每个快照的回放预算后，需要进一步从该快照的历史三元组中筛选出最易被遗忘、最具回放价值的样本。本文以预测分数差异作为筛选依据：预测分数差异越大，说明模型在增量更新后对该三元组的评估发生了显著变化，其原有知识更可能受到干扰，因而具有更高的回放优先级。

对于来自历史快照$\mathcal{S}_j(j\leqslant i)$的三元组$(h,r,t)$，分别使用快照$\mathcal{S}_i$训练完成后保存的冻结模型参数$\theta^{(i)}_{\text{frozen}}$与当前增量训练阶段的实时模型参数$\theta^{(i+1)}_{\text{current}}$计算其预测分数，并定义预测分数差异为：

$$\Delta(h,r,t)=\left|f_{\theta^{(i)}_{\text{frozen}}}(h,r,t)-f_{\theta^{(i+1)}_{\text{current}}}(h,r,t)\right|,\tag{10}$$

其中$f_\theta(\cdot)$表示由模型参数θ参数化的链接预测评分函数。在每个历史快照内部，按预测分数差异$\Delta(h,r,t)$由大到小对候选三元组排序，选取差异最大的前K_j个三元组作为该快照$\mathcal{S}_j$的回放样本；最终将所有历史快照的回放样本取并集，构成回放集合：

$$\mathcal{T}_{\text{rep}}=\bigcup_{j=0}^{i}TopK_{(h,r,t)\in\mathcal{T}_j}(\Delta(h,r,t),K_j)。\tag{11}$$

时序感知自适应回放的流程如算法1所示。

算法 1. 时序感知自适应回放算法。

输入：历史快照集合$\{\mathcal{S}_0,\cdots,\mathcal{S}_i\}$，各快照三元组集合$\{\mathcal{T}_0,\cdots,\mathcal{T}_i\}$，全局回放比例$\rho$，冻结模型参数$\theta^{(i)}_{\text{frozen}}$，当前模型参数$\theta^{(i+1)}_{\text{current}}$；

输出：回放三元组集合$\mathcal{T}_{\text{rep}}$。

① 由ρ与$|\mathcal{T}_i|$计算总回放预算K；

② 初始化回放集合$\mathcal{T}_{\text{rep}}=\varnothing$；

③ for $j=0,\cdots,i$ do

④ 按式(9)计算快照$\mathcal{S}_j$的回放预算K_j；

⑤ for each $(h,r,t)\in\mathcal{T}_j$ do

⑥ 用$\theta^{(i)}_{\text{frozen}}$与$\theta^{(i+1)}_{\text{current}}$计算预测分数；

⑦ 按式(10)计算预测分数差异$\Delta(h,r,t)$；

⑧ end for

⑨ 按$\Delta(h,r,t)$由大到小对$\mathcal{T}_j$中三元组排序；

⑩ 选取前K_j个三元组加入$\mathcal{T}_{\text{rep}}$；

⑪ end for

⑫ 按式(11)取并集得到最终回放集合$\mathcal{T}_{\text{rep}}$；

⑬ return $\mathcal{T}_{\text{rep}}$。

3.3.3 历史回放一致性约束

对于选中的回放三元组集合$\mathcal{T}_{\text{rep}}$，本文在分数层面对冻结历史模型与当前模型的预测施加一致性约束，以保护历史知识在增量训练中不被遗忘。具体来说，对每个$(h,r,t)\in\mathcal{T}_{\text{rep}}$，分别基于$\theta^{(i)}_{\text{frozen}}$与$\theta^{(i+1)}_{\text{current}}$计算其预测分数，回放损失定义为：

$$\mathcal{L}_{\text{rep}}=\sum_{(h,r,t)\in\mathcal{T}_{\text{rep}}}H(f_{\theta^{(i)}_{\text{frozen}}}(h,r,t),f_{\theta^{(i+1)}_{\text{current}}}(h,r,t)),\tag{12}$$

其中$H(\cdot,\cdot)$表示Huber损失函数。相较于ℓ_2损失，Huber损失对高差异样本的惩罚更为平缓，避免少数极端样本主导梯度。

回放集合在快照切换阶段更新一次，而非每个epoch重新计算。模型缓存上一快照的实体与关系表征作为冻结历史表征，并以当前模型和冻结历史表征对候选历史三元组的评分差异作为遗忘风险度量，按*TopK*选择高风险三元组加入回放集合。设历史三元组的候选池规模为C，嵌入维度为d，回放集合规模为K，则一次筛选的评分计算复杂度为$O(Cd)$，*TopK*选择复杂度约为$O(C\log K)$，训练阶段每个epoch仅对缓存的K个回放三元组计算一致性损失，额外复杂度为$O(Kd)$。该实现不保存全部候选三元组分数，仅缓存历史实体/关系表征和最终回放集合，因此内存开销可控。

3.4 链接预测与优化目标

在前述多模态融合、时序对齐与自适应回放机制的基础上，本节给出SAREC在每个快照上的链接预测目标及总体优化目标。

在快照$\mathcal{S}_i$上，对于给定的训练三元组(h,r,t)，本

文采用经典的 TransE[5] 评分函数，基于 3.1 节得到的多模态实体融合表征评估三元组的合理性，评分函数定义为：

$$f(h,r,t)=-\|z_h+z_r-z_t\|_2, \tag{13}$$

其中 z_h 与 z_t 分别表示头实体 h 与尾实体 t 经多模态融合得到的最终表征，z_r 表示关系 r 的嵌入。沿用知识图谱补全任务中的通用负采样策略，通过替换头实体或尾实体生成对应的负样本三元组，并将正、负样本的评分分别记为 s^+ 与 s^-。

链接预测目标通过带边际的排序损失进行优化：

$$\mathcal{L}_{\text{lp}}=\frac{1}{|\mathcal{B}|}\sum_{(h,r,t)\in\mathcal{B}} w(h,t)\cdot\ell_{\text{MR}}(s^+,s^-), \tag{14}$$

其中 $\mathcal{B}$ 表示训练批次，$\ell_{\text{MR}}(s^+,s^-)=\max(0,\gamma+s^--s^+)$ 表示边际为 γ 的排序损失，考虑到新引入的实体往往结构连接稀疏、参与训练的三元组数量有限，难以与历史实体在同等条件下被充分优化，本文进一步引入新实体感知的样本权重 $w(h,t)$：若头实体 h 或尾实体 t 为快照 $\mathcal{S}_i$ 中新引入的实体，则 $w(h,t)=\beta$，否则 $w(h,t)=1$，其中 $\beta>1$ 为放大因子。通过将新实体感知直接嵌入到学习目标中，模型能够在统一的优化框架下优先关注稀疏的新增实体，加强其表征学习。

最终，SAREC 在快照 $\mathcal{S}_i$ 上的总体训练目标由链接预测损失、回放一致性损失与时序对齐损失 3 部分加权组合而成：

$$\mathcal{L}^{(i)}=\mathcal{L}^{(i)}_{\text{lp}}+\lambda_{\text{rep}}\mathcal{L}^{(i)}_{\text{rep}}+\lambda_{\text{align}}\mathcal{L}^{(i)}_{\text{align}}, \tag{15}$$

其中 λ_{rep} 与 λ_{align} 为 2 项约束损失的权衡超参数。链接预测损失 $\mathcal{L}^{(i)}_{\text{lp}}$ 保证模型对当前快照新增知识的有效学习，回放一致性损失 $\mathcal{L}^{(i)}_{\text{rep}}$ 通过有针对性地复现易遗忘的历史三元组来缓解灾难性遗忘，时序对齐损失 $\mathcal{L}^{(i)}_{\text{align}}$ 则约束历史实体与关系表征在快照间的稳定演化。三者协同作用，使 SAREC 能够在持续演化的多模态知识图谱中实现历史知识的有效保持与新增知识的稳定学习，增强模型在开放动态环境下的持续链接预测能力。

4 实　　验

4.1 实验设置

1)数据集。为验证所提模型在多模态知识图谱持续补全任务上的有效性，本文在 7 个公开数据集上开展实验，包括 Cui 等人[13] 构建的 ENTITY、RELATION、FACT 与 HYBRID，以及 Liu 等人[14] 构建的 GraphEqual、GraphHigher 与 GraphLower，覆盖了知识动态演化的多种典型场景。为验证多模态信息在持续知识图谱补全中的作用，本文在保留原始三元组结构与数据划分方式不变的前提下，为每个实体补充文本描述与图像信息，以保证与现有持续知识图谱补全工作的公平对比。所有数据集均被划分为 5 个快照，每个快照内的训练集、验证集与测试集按 3∶1∶1 的比例划分，详细的数据统计信息可参见表 1。

2)基线模型。为系统验证 SAREC 相对于现有持续知识图谱补全方法的性能优势，本文选取了 3 种类别的 11 种代表性基线方法进行对比：基于动态架构扩展的方法包括 PNN[9]、CWR[44]、FastKGE[10] 与 SAGE[11]；基于参数正则化的方法包括 SI[45]、EWC[12]、LKGE[13]、IncDE[14] 与 DebiasedKGE[15]；基于知识回放的方法包括 GEM[16]、EMR[17]。所选基线覆盖了持续学习领域三大主流技术路线，能够充分反映 SAREC 在不同对比维度下的性能表现。为保证公平性，本文在相同文本与视觉特征输入条件下构建了 SAGE-MM、LKGE-MM 和 IncDE-MM 多模态增强基线，将文本和图像特征拼接并投影到对应模型的结构嵌入空间，通过可学习标量门控与结构实体嵌入融合，不

Table 1　Statistics for the Variation of Dataset with Time

表 1　数据集随时间变化的统计信息

数据集	快照 1			快照 2			快照 3			快照 4			快照 5		
	N_{E}	N_{R}	N_{T}	N_{E}	N_{R}	N_{T}	N_{E}	N_{R}	N_{T}	N_{E}	N_{R}	N_{T}	N_{E}	N_{R}	N_{T}
ENTITY	2 909	233	46 388	5 817	236	72 111	8 275	236	73 785	11 633	237	70 506	14 541	237	47 326
RELATION	11 560	48	98 819	13 343	96	93 535	13 754	143	66 136	14 387	190	30 032	14 541	237	21 594
FACT	10 513	237	62 024	12 779	237	62 023	13 586	237	62 023	13 894	237	62 023	14 541	237	62 023
HYBRID	8 628	86	57 561	10 040	102	20 873	12 779	151	88 017	14 393	209	103 339	14 541	237	40 326
GraphEqual	2 908	226	57 636	5 816	235	62 023	8 724	237	62 023	11 632	237	62 023	14 541	237	66 411
GraphHigher	900	197	10 000	1 838	221	20 000	3 714	234	40 000	7 467	237	80 000	14 541	237	160 116
GraphLower	7 505	237	160 000	11 258	237	80 000	13 134	237	40 000	14 072	237	20 000	14 541	237	10 116

注：N_{E}、N_{R} 分别表示截至该快照的累计实体数与关系数，N_{T} 表示该快照内新增三元组数。

改变原始关系表示、打分函数和持续学习目标，保证性能差异主要反映多模态实体表示接入后的影响。

3）评价指标。本文在链接预测任务上采用通用的排序类指标作为性能评价标准，包括平均倒数排名（mean reciprocal rank, MRR）以及预测实体在前 1 与前 10 的命中率（Hits@1 和 Hits@10，简记为 H@1 和 H@10），指标值越高表示模型性能越好。此外，为衡量模型在开放动态知识环境中的前向适应与历史知识保持能力，引入前向迁移（forward transfer, FWT[16]）与后向迁移（backward transfer, BWT[16]）两项指标。具体来说，对于快照i，在第 1 个至第i个快照的测试集上分别计算各项指标，并取其平均值作为该快照上的整体性能；最终的实验结果来自最后时刻生成的模型。

4）实现细节。所有实验基于 PyTorch 框架实现，并在一台配有 48G NVIDIA RTX A6000 GPU 的设备上完成。模型优化采用 Adam 优化器，学习率范围设为{1E−5, 1E−4, 1E−3}，若验证集性能连续 5 个 epoch 未提升则触发早停。实体嵌入与关系嵌入的维度均为 500。训练的批次大小在{512, 1 024, 2 048}范围内进行调整，早停步数设置为 5。回放比例ρ在{0.1,0.2, 0.3,0.4,0.5}范围内搜索，超参数λ_{rep}和λ_{align}均在{0.1, 0.2,⋯,1.0}范围内进行网格搜索以选取最优配置。新实体感知的放大因子β在(1.0,3.0]区间内搜索。

5）大语言模型与 Prompt 设置。本文采用 Llama-3.1-8B-Instruct 生成实体语义描述，基于固定 Prompt 模板为每个实体生成语义描述。给定实体名称，输入大语言模型的提示模板如下：

<|begin_of_text|>

<|system|>

You are a knowledgeable assistant.

<|user|>

Provide a concise encyclopedic description of the entity "[*ENTITY_NAME*]" in one sentence.

<|assistant|>

其中，[*ENTITY_NAME*] 表示当前实体的可读名称。模型输出为该实体的一句简洁百科式描述，并作为后续文本模态特征编码的输入。为保证可复现性，生成过程采用确定性解码，即 *do_sample*=False。

4.2 链接预测任务结果

表 2 和表 3 展示了 SAREC 与基线方法在 7 个演化数据集上的链接预测结果。总体来看，SAREC 在绝大多数数据集上均取得最优性能，验证了所提方法的有效性。以下从不同的知识演化场景展开分析。

1）实体快速增长场景。以 ENTITY、GraphEqual 2 个数据集以实体的增量变化为核心，呈现实体空间

Table 2 Experimental Results of Different Methods on ENTITY, RELATION, FACT, HYBRID and GraphEqual Datasets

表 2 不同方法在 ENTITY、RELATION、FACT、HYBRID 及 GraphEqual 数据集上的实验结果

方法	ENTITY			RELATION			FACT			HYBRID			GraphEqual		
	MRR	H@1	H@10	MRR	H@1	H@10	MRR	H@1	H@10	MRR	H@1	H@10	MRR	H@1	H@10
CWR*	0.084	0.028	0.191	0.022	0.010	0.043	0.082	0.030	0.189	0.040	0.016	0.081	0.121	0.041	0.274
PNN*	0.232	0.132	0.428	0.166	0.095	0.305	0.157	0.086	0.289	0.188	0.104	0.352	0.213	0.118	0.408
FastKGE	0.223	0.137	0.390	0.168	0.100	0.294	0.171	0.104	0.292	0.194	0.115	0.344	0.133	0.074	0.243
SAGE	0.255	0.152	0.441	0.188	0.100	0.350	0.215	**0.132**	0.391	0.210	0.123	0.375	0.231	0.135	0.420
SAGE-MM	0.208	0.133	0.353	0.038	0.017	0.074	0.161	0.104	0.273	0.118	0.067	0.211	0.179	0.111	0.311
SI*	0.121	0.048	0.257	0.119	0.061	0.229	0.154	0.073	0.319	0.103	0.044	0.214	0.170	0.083	0.342
EWC*	0.232	0.133	0.429	0.163	0.091	0.304	0.201	0.114	0.380	0.189	0.106	0.355	0.208	0.114	0.400
LKGE*	0.241	0.142	0.436	0.179	0.095	0.348	0.209	0.122	0.386	0.204	0.116	0.377	0.223	0.126	0.421
LKGE-MM	0.225	0.137	0.393	0.095	0.052	0.176	0.177	0.111	0.304	0.145	0.089	0.249	0.198	0.115	0.361
IncDE	0.254	0.150	0.448	0.198	0.109	0.368	0.213	0.127	0.387	0.222	0.128	0.398	0.231	0.131	0.428
IncDE-MM	0.205	0.120	0.370	0.078	0.042	0.142	0.143	0.086	0.251	0.122	0.075	0.208	0.184	0.104	0.338
DebiasedKGE	0.238	0.141	0.433	0.164	0.089	0.318	0.181	0.096	0.343	0.191	0.112	0.351	0.219	0.125	0.411
EMR*	0.169	0.089	0.327	0.112	0.054	0.221	0.159	0.082	0.317	0.140	0.074	0.267	0.182	0.097	0.354
GEM*	0.157	0.079	0.310	0.089	0.041	0.182	0.164	0.084	0.326	0.133	0.069	0.255	0.181	0.095	0.355
SAREC（本文）	**0.274**	**0.163**	**0.484**	**0.243**	**0.144**	**0.433**	**0.219**	0.128	**0.415**	**0.241**	**0.138**	**0.436**	**0.251**	**0.141**	**0.467**

注：黑体数值显示的结果代表最佳性能，下划线标注的数值对应次优结果；标有*的方法基于 LKGE 提供的代码框架复现，其余方法则在我们的实验环境中参照各自原始论文独立实现。

Table 3 Experimental Results of Different Methods on GraphHigher and GraphLower Datasets

表 3 不同方法在 GraphHigher 和 GraphLower 数据集上的实验结果

方法	GraphHigher			GraphLower		
	MRR	H@1	H@10	MRR	H@1	H@10
CWR*	0.190	0.098	0.372	0.032	0.005	0.079
PNN*	0.186	0.097	0.364	0.215	0.122	0.405
FastKGE	0.097	0.061	0.159	0.204	0.117	0.376
SAGE	0.238	0.143	0.425	0.212	0.120	0.397
SAGE-MM	0.180	0.113	0.311	0.168	0.107	0.286
SI*	0.182	0.093	0.358	0.177	0.093	0.346
EWC*	0.197	0.106	0.383	0.205	0.114	0.389
LKGE*	0.218	0.125	0.402	0.215	0.122	0.406
LKGE-MM	0.199	0.120	0.354	0.189	0.110	0.345
IncDE	0.229	0.135	0.414	0.227	0.128	0.422
IncDE-MM	0.185	0.107	0.336	0.175	0.102	0.321
DebiasedKGE	0.205	0.119	0.380	0.213	0.120	0.404
EMR*	0.200	0.111	0.377	0.172	0.091	0.336
GEM*	0.194	0.106	0.370	0.178	0.094	0.346
SAREC（本文）	**0.250**	**0.145**	**0.454**	**0.234**	**0.132**	**0.440**

注：黑体数值显示的结果代表最佳性能，下划线标注的数值对应次优结果；标有*的方法基于 LKGE 提供的代码框架复现，其余方法则在我们的实验环境中参照各自原始论文独立实现。

持续快速扩张的特征。SAREC 在 ENTITY 上的 MRR、H@1、H@10 较最强基线分别提升 1.9、1.1、3.6 个百分点；在 GraphEqual 上较最强基线（SAGE/IncDE）分别提升 2.0、0.6、3.9 个百分点。这表明在新实体持续涌入且空间持续快速扩张的场景下，多模态语义与新实体感知加权机制能够有效弥补结构信息不足，显著提升新增实体的表征质量。

2）关系快速增长场景。SAREC 在 RELATION 数据集上优势最为突出，MRR 达 0.243，相较最强基线 IncDE（MRR 为 0.198）提升 22.7%，相较 H@1 提升 32.1%。这表明在关系语义持续扩展的场景下，多模态时序语义对齐机制能够有效约束关系表征的非预期漂移，维持跨快照的语义一致性。

3）多维度混合演化场景。HYBRID 的实体、关系与三元组数量同步变化，是演化模式最为复杂的场景。SAREC 在 MRR、H@1、H@10 三项指标上均优于所有基线，较最强基线 IncDE 分别提升 1.9、1.0、3.8 个百分点，表明 3 个模块的协同设计在复杂多变的演化过程中具有较强的鲁棒性。

4）三元组增量模式主导场景。FACT、GraphHigher、GraphLower 以新增三元组的增量轨迹为核心变量，分别对应均匀、指数递增、指数递减 3 种典型模式。FACT 结构信息充足，SAREC 的 MRR 与 H@10 最优，H@1（0.128）略低于结构型基线 SAGE（0.132），表明结构充足时多模态增益相对有限。GraphHigher 增量指数递增，SAREC 三项指标虽居首但优势收窄，反映知识爆发式增长加剧了各方法的遗忘压力。GraphLower 增量持续递减，SAREC 较 IncDE 三项指标分别提升 0.7、0.4、1.8 个百分点，说明时序感知回放与多模态语义锚点有助于维持知识稀疏演化下的长期稳定性。

4.3 消融实验

为验证 SAREC 中各核心模块以及多模态信息对模型性能的影响与贡献，本文在 ENTITY 与 RELATION 数据集上设计了系统的消融实验。具体来说，构造了 7 种变体模型与完整模型作对比：1）w/o T：移除实体文本模态信息，仅保留结构与视觉信息；2）w/o I：移除实体视觉模态信息，仅保留结构与文本信息；3）w/o T&I：同时移除文本与视觉模态信息，仅保留结构嵌入；4）w/o MTA：移除多模态时序语义对齐模块，不再对跨快照表征演化施加一致性约束；5）w/o TAAR：移除时序感知自适应回放模块，仅依赖当前快照数据进行增量训练；6）w/ SimpleF：将模态感知的语义-结构融合模块替换为简化版本，将投影后的文本与图像嵌入以等权相加得到多模态表示，再与结构嵌入拼接后经单层 MLP 映射至目标维度，以验证本文自适应门控融合设计的必要性；7）TAAR-Uni：将时序感知自适应回放模块中的快照预算分配策略替换为均匀分配，以验证时序感知的线性加权分配策略的必要性。表 4 展示了完整的 SAREC 模型与 7 种变体的对比结果。

从回放机制来看，TAAR 模块对模型性能的贡献

Table 4 Ablation Experiment Results

表 4 消融实验结果

方法	ENTITY			RELATION		
	MRR	H@1	H@10	MRR	H@1	H@10
w/o T	0.255	0.145	0.471	0.216	0.118	0.412
w/o I	0.259	0.149	0.473	0.222	0.124	0.417
w/o T&I	0.256	0.143	0.472	0.215	0.117	0.413
w/o MTA	0.257	0.146	0.470	0.223	0.125	0.416
w/o TAAR	0.172	0.081	0.347	0.114	0.032	0.271
w/ SimpleF	0.192	0.114	0.344	0.175	0.109	0.304
TAAR-Uni	0.167	0.075	0.342	0.114	0.033	0.269
SAREC（本文）	**0.274**	**0.163**	**0.484**	**0.243**	**0.144**	**0.433**

注：黑体数值显示的结果代表最佳性能。

最为显著，移除 TAAR 后 ENTITY 与 RELATION 数据集上的 MRR 分别由 0.274 降至 0.172、由 0.243 降至 0.114，绝对下降幅度远超其他模块；将时序感知预算分配替换为均匀分配后性能进一步下滑，表明低价值历史样本的引入会干扰当前快照的训练，验证了时序感知线性加权分配的必要性。从多模态融合来看，移除单一模态或双模态均导致性能稳定下降，而 w/o T 与 w/o T&I 性能相近，说明文本与视觉模态在共享语义空间中可能存在一定重叠；将自适应门控融合替换为简单等权求和加 MLP（w/ SimpleF）后性能明显下滑，验证了自适应融合机制的优越性；移除 MTA 模块后 2 个数据集性能均出现稳定下降，说明跨快照时序对齐约束能够有效抑制历史表征的非预期偏移。

4.4 知识迁移能力分析

为评估不同方法在持续学习场景下的知识迁移能力，图 3 与图 4 分别展示了 SAREC 与不同基线方法在 FWT 与 BWT 指标上的对比结果。

1）前向迁移能力分析。如图 3 所示，SAREC 在全部 7 个数据集上均取得最高 FWT 分数，表明 SAREC 能够在多种图谱演化模式下稳定地将历史知识迁移至新快照的学习过程中。这一优势主要源于多模态语义建模机制：即使新引入实体在结构层面连接稀疏，文本与视觉模态所提供的语义信息仍能为实体构建具有可迁移性的稳定表征基础。

2）后向迁移能力分析。如图 4 所示，大多数基线方法的 BWT 指标呈负值，反映出持续训练过程中不同程度的灾难性遗忘。与之对比，SAREC 在所有数据集上基本取得了正的或最接近 0 的负 BWT 分数，表明 SAREC 不仅有效抑制了历史知识的遗忘，还能在新快照训练过程中进一步巩固对历史知识的理解。这一优势说明 SAREC 能够将有限的回放资源精准投放至最需保护的历史知识，在动态演化的知识环境中实现稳定的知识保持与链接预测能力持续增强。

4.5 跨快照知识保持能力分析

图 5 展示了 SAREC 在 ENTITY、RELATION 与 GraphEqual 数据集上跨快照的 MRR 演化情况。总体而言，SAREC 在持续学习过程中保持了稳定的性能表现，展现出良好的抗遗忘能力。

在 ENTITY 数据集上，对于早期快照测试集（如 D1），随着模型的持续训练（如快照 1~2），模型 MRR 呈上升趋势，表明新快照引入的知识能够反向强化模型对早期实体的表征理解，虽然 D1 在第 4 和第 5 快照中出现了轻微的性能下降，但 MRR 的最大降幅保持在 1 个百分点以内，表明 SAREC 对灾难性遗忘

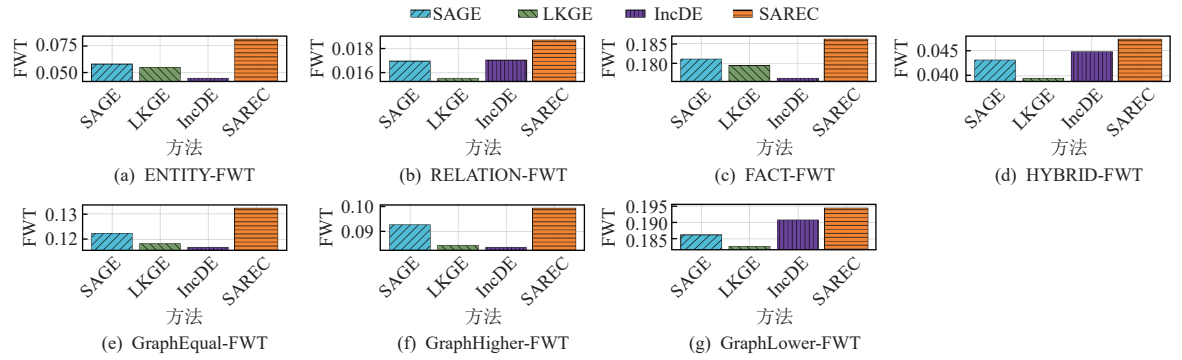


Fig. 3 Comparison of FWT metrics using different methods

图 3 不同方法的 FWT 指标对比

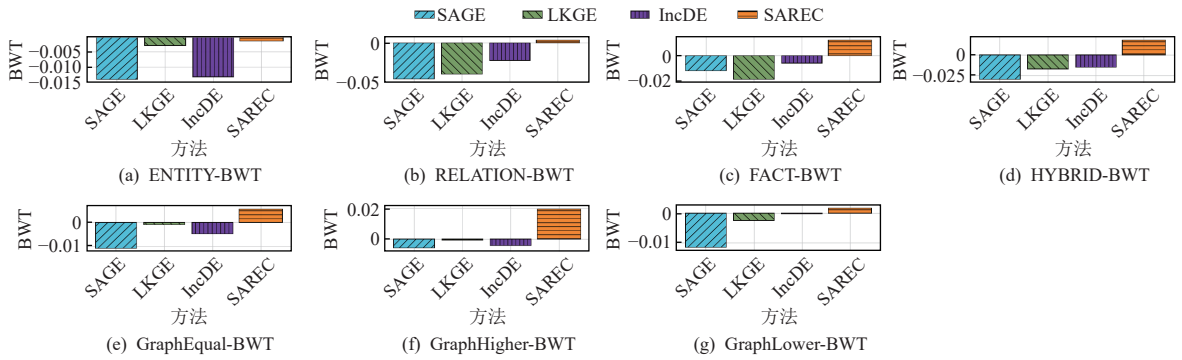
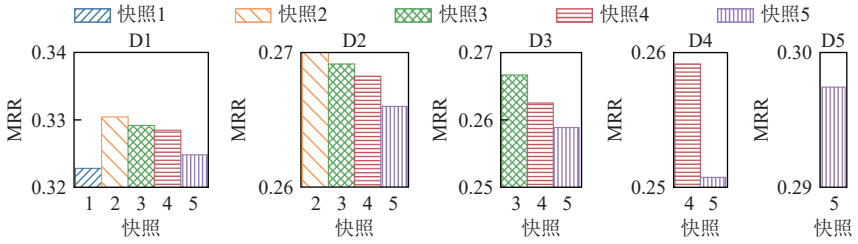
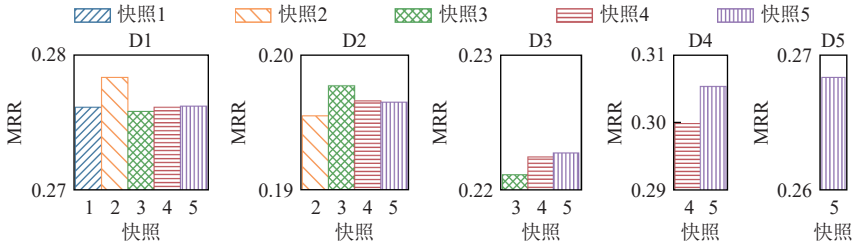


Fig. 4 Comparison of BWT metrics using different methods

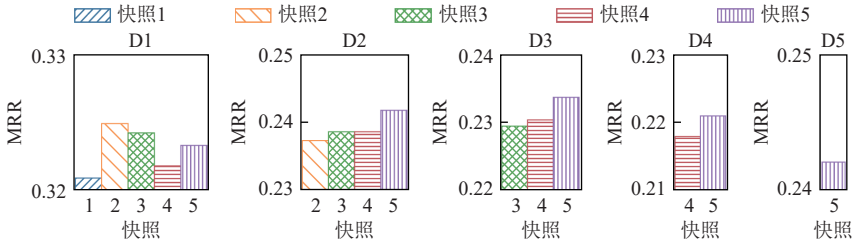
图 4 不同方法的 BWT 指标对比



(a) ENTITY数据集



(b) RELATION数据集



(c) GraphEqual数据集

Fig. 5 The evolutionary trends of SAREC across snapshots

图 5 SAREC 的跨快照演化趋势

的抑制效果良好。在 RELATION 的 D3 和 D4 以及 GraphEqual 的 D2~D4 测试集上，SAREC 的 MRR 指标呈现渐进式提升，验证了上述观察的普遍性。

上述结果说明，SAREC 不仅有效抑制了持续学习过程中的灾难性遗忘，还能将新增知识转化为对历史知识的进一步强化，在动态演化的开放知识环境中实现了历史知识保持与新知识学习的协同增益。

4.6 超参数分析

为深入理解时序感知自适应回放策略中关键超参数对模型性能的影响，本文在 ENTITY 数据集上对回放比例ρ与回放损失权重λ_{rep}进行了敏感性分析，结果如图 6 所示。

1)回放比例。如图 6 左侧所示，当ρ从 0.1 增大至 0.3 时，各项评价指标均快速上升，表明适当增加回放预算能够显著改善历史知识的保持能力；超过 0.3 后性能提升趋于平缓，并在$\rho=0.7$附近达到最优。这一趋势表明，得益于自适应关键三元组选择机制将有限回放资源精准投放至最易遗忘的历史知识，所提自适应回放策略在较低回放开销下即可取得接近全量回放的效果。

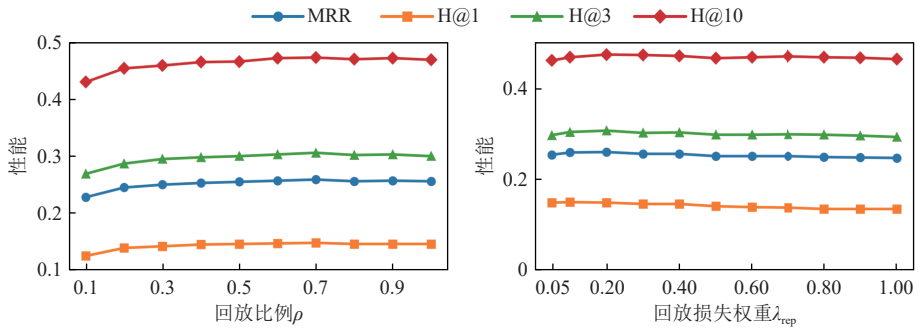


Fig. 6 Sensitivity analysis of hyperparameters for replay on ENTITY dataset

图 6 回放超参数在 ENTITY 数据集上的敏感性分析

2）回放损失权重。如图 6 右侧所示，λ_{rep}在 0.2 时取得最优性能，随后随权重增大各项指标性能逐渐下降。这表明，λ_{rep}过小时历史知识保持力度不足，过大时则过度约束模型对新增知识的学习能力，合理的λ_{rep}取值能够在历史知识保持与新知识吸收之间取得有效平衡。

4.7 案例分析

为更直观地展示 SAREC 在跨快照演化场景下的工作机制，本文选取一个具有代表性的三元组查询 (Hulk Hogan, profession, ?) 作为案例，其正确答案为 Film_producer，该三元组出现于第 3 个快照，使用在第 4 个快照上训练得到的模型对其进行预测，本文选取了 Singer-songwriter、Film_producer 和 Author 这 3 个候选实体进行分析，结果如表 5 所示。其中，Δ_{esem} = 10.918表示该三元组所涉及实体的融合表示在相邻快照间的偏移幅度，Δ_{score}=0.288 9 表示快照更新前后该三元组的预测分数变化量（绝对值），两者均处于较高水平，说明若不加以约束，该三元组的表征与评分在持续训练中将出现明显的不稳定性，具有较高的遗忘风险。

Table 5 Case Study

表 5 案例分析

方法	前 3 名候选	回放
IncDE	Singer-songwriter, Film_producer, Author	
TAAR-Uni	Singer-songwriter,Author, Film_producer	No
w/o T&I	Singer-songwriter, Film_producer, Author	Yes
SAREC	Film_producer, Singer-songwriter, Author	Yes

从预测结果来看，IncDE 未能将正确尾实体 Film_producer 排在候选列表首位；时序感知自适应回放模块的变体（TAAR-Uni）同样未能正确排序，且该三元组未被选入回放集合，表明在缺乏针对性回放保护的情况下，该高风险三元组在增量训练中未受到任何针对性保护，导致排序退化；不使用文本和图像模态信息的模型变体（w/o T&I）虽将 Film_producer 排至第 2 位，但仍未达到最优；完整的 SAREC 模型则成功将 Film_producer 排在第 1 位，且该三元组被选入回放集合。上述结果表明，SAREC 通过计算预测分数的差异，可以识别高遗忘风险的三元组并优先进行回放，结合多模态语义增强，能够在持续演化的知识环境中实现对关键历史知识的精准保护。

4.8 效率分析

4.8.1 整体时间开销对比

表 6 对比了 SAREC 与 4 个基线方法在 ENTITY 数据集上的训练时间开销，SAREC 的回放比例设置为 0.5。相较于轻量级纯结构方法，SAREC 引入多模态融合与自适应回放机制会带来一定的额外计算开销，这在多模态持续学习框架中是合理且可接受的。值得注意的是，大语言模型的文本生成与 CLIP 特征提取均为离线一次性预处理操作，在实际训练与推理阶段无需重复加载，不构成持续开销。从每轮训练的平均耗时来看，SAREC（22.733 s/epoch）与基于知识回放的 EMR（27.034 s/epoch）和 GEM（21.333 s/epoch）处于同一量级，表明所提框架在引入多模态语义增强的同时保持了合理的训练效率。此外，由于 SAREC 的回放比例是可灵活调节的，通过控制每轮回放的历史三元组数量，可在补全性能与训练时间开销之间取得合理平衡，使框架在资源受限场景下具备良好的适应性。

Table 6 Time Overhead Comparison of Different Methods on ENTITY Dataset

表 6 不同方法在 ENTITY 数据集上的时间开销对比

指标	SAREC	EMR	GEM	FastKGE	DebiasedKGE
总时长/s	9 821	6 326	7 296	1 003	2 459
epochs	432	234	342	73	217
平均耗时（s/epoch）	22.733	27.034	21.333	13.740	11.332

4.8.2 模块时间开销分析

表 7 展示了 SAREC 各模块在 RELATION 数据集上的时间开销分解。从每轮平均耗时来看，移除文本与视觉模态（w/o T&I）与移除自适应回放（w/o TAAR）均带来显著的开销降低，分别节省约 18.5% 与 15.4%，说明多模态语义注入与自适应回放是框架计算开销的 2 个主要来源；而移除多模态时序语义对齐模块（w/o MTA）带来的开销降低相对有限，表明跨快照时序对齐约束以较小的计算代价提供了互补性的表征稳定效果。综合各模块的性能贡献与计算开销，SAREC 在效率与性能之间取得了合理的平衡，适用于动态开放环境下的多模态知识图谱持续补全任务。

Table 7 Time Overhead Analysis of Modules on RELATION Dataset

表 7 RELATION 数据集上模块的时间开销分析

指标	SAREC	w/o T	w/o I	w/o T&I	w/o TAAR	w/o MTA
总时长/s	7 921	9 098	6 401	7 037	6 781	6 190
epochs	354	452	322	386	358	299
平均耗时/（s/epoch）	22.376	20.128	19.879	18.231	18.941	20.702

5 总结与展望

本文针对动态多模态知识图谱持续补全任务中灾难性遗忘与跨快照语义非预期偏移的核心挑战，提出了语义对齐与自适应回放增强的持续补全框架SAREC，该框架由3个模块构成：大语言模型引导的多模态差异感知融合、多模态时序语义对齐和时序感知自适应回放。在7个具有不同演化特点的数据集上的实验表明，SAREC在多种知识演化模式下均取得了优越的链接预测性能，并在前向迁移与后向迁移指标上验证了其持续知识积累与历史知识保持能力。

在未来工作中，将从以下2个方向进一步拓展本文工作。其一，在多模态融合层面引入更细粒度的模态可靠性建模机制，以缓解噪声或缺失模态对融合表征质量的影响，提升框架在真实开放环境下的鲁棒性。其二，在应用层面，探索将SAREC持续补全的多模态知识图谱作为外部知识源，嵌入到知识增强的大模型问答、智慧医疗与智能教育等高可靠性下游场景中，推动动态知识图谱在知识增强人工智能领域的实际落地与应用。

作者贡献声明：王迪负责论文撰写与修改、方法设计、实验验证及结果分析；杜军平负责论文审阅、修订与课题指导；薛哲负责论文选题与研究思路指导；叶冠华负责实验方案设计与结果核实；赵桐、孟晓龙负责文献调研、数据整理与分析；李海生负责论文讨论与修改建议。

参 考 文 献

[1] Ma Jie, Sun Wangchun, Wang Pinghui, et al. Knowledge graph question answering based on relevance prompts[J]. Journal of Software, 2025, 36(9): 4056−4071 (in Chinese)
(马杰, 孙望淳, 王平辉, 等. 基于相关性提示的知识图谱问答[J]. 软件学报, 2025, 36(9): 4056−4071)

[2] Cai Borui, Xiang Yong, Gao Longxiang, et al. Temporal knowledge graph completion: A survey[C]//Proc of the 32nd Int Joint Conf on Artificial Intelligence. Palo Alto, CA: IJCAI, 2023: 6545−6553

[3] Li Linyu, Jin Zhi, He Yuanpeng, et al. Learning to evolve: Bayesian-guided continual knowledge graph embedding[C]//Proc of the ACM Web Conf 2026. New York: ACM, 2026: 4372−4383

[4] Zhang Yichi, Chen Zhuo, Guo Lingbing, et al. Tokenization, fusion, and augmentation: Towards fine-grained multi-modal entity representation[C]//Proc of the 39th AAAI Conf on Artificial Intelligence. Palo Alto, CA : AAAI, 2025, 39(12): 13322−13330

[5] Bordes A, Usunier N, García-Durán A, et al. Translating embeddings for modeling multi-relational data[C]//Proc of the 27th Int Conf on Neural Information Processing Systems. Red Hook, NY: Curran Associates, 2013: 2787−2795

[6] Sun Zhiqing, Deng Zhihong, Nie Jianyun, et al. RotatE: Knowledge graph embedding by relational rotation in complex space[C/OL]// Proc of the 7th Int Conf on Learning Representations. Amherst, MA: OpenReview. net, 2019[2026-05-20]. https://openreview.net/forum?id=HkgEQnRqYQ

[7] Yang Bishan, Yih W, He Xiaodong, et al. Embedding entities and relations for learning and inference in knowledge bases[C/OL]// Proc of the 3rd Int Conf on Learning Representations. Amherst, MA: OpenReview. net, 2015[2026-05-20]. https://arxiv.org/abs/1412.6575

[8] Zhao Tianzhe, Chen Jiaoyan, Ru Yanchi, et al. Rethinking continual knowledge graph embedding: Benchmarks and analysis[C]//Proc of the 48th Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2025: 138−147

[9] Rusu A A, Rabinowitz N C, Desjardins G, et al. Progressive neural networks[J]. arXiv preprint, arXiv: 1606.04671, 2016

[10] Liu Jiajun, Ke Wenjun, Wang Peng, et al. Fast and continual knowledge graph embedding via incremental LoRA[C]//Proc of the 33rd Int Joint Conf on Artificial Intelligence. Palo Alto, CA: IJCAI, 2024: 2198−2206

[11] Li Yifei, Zhang Lingling, Yan Hang, et al. SAGE: Scale-aware gradual evolution for continual knowledge graph embedding[C]//Proc of the 31st ACM SIGKDD Conf on Knowledge Discovery and Data Mining. New York: ACM, 2025: 1600−1611

[12] Kirkpatrick J, Pascanu R, Rabinowitz N, et al. Overcoming catastrophic forgetting in neural networks[J]. Proceedings of the National Academy of Sciences of the United States of America, 2017, 114(13): 3521−3526

[13] Cui Yuanning, Wang Yuxin, Sun Zequn, et al. Lifelong embedding learning and transfer for growing knowledge graphs[C]//Proc of the 37th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2023: 4217−4224

[14] Liu Jiajun, Ke Wenjun, Wang Peng, et al. Towards continual knowledge graph embedding via incremental distillation[C]//Proc of the 38th AAAI Conf on Artificial Intelligence. Palo Alto, CA : AAAI, 2024: 8759−8768

[15] Zhu Junlin, Fu Bo, Duan Guiduo. DebiasedKGE: Towards mitigating spurious forgetting in continual knowledge graph embedding[C]// Proc of the 34th ACM Int Conf on Information and Knowledge Management. New York: ACM, 2025: 4519−4528

[16] Lopez-Paz D, Ranzato M. Gradient episodic memory for continual learning[C]//Proc of the 31st Int Conf on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc., 2017: 6470−6479

[17] Wang Hong, Xiong Wenhan, Yu Mo, et al. Sentence embedding alignment for lifelong relation extraction[C]//Proc of the 2019 Conf

of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. Stroudsburg, PA: ACL, 2019: 796−806

[18] Zhang Zhiyu, Chen Wei, Lin Youfang, et al. A generative adaptive replay continual learning model for temporal knowledge graph reasoning[C]//Proc of the 63rd Annual Meeting of the Association for Computational Linguistics. Stroudsburg, PA: ACL, 2025: 10964−10977

[19] Yang Jing, Jiang Xinfa, Jiang Xiaowen, et al. From knowledge forgetting to accumulation: Evolutionary relation path passing for lifelong knowledge graph embedding[C]//Proc of the 48th Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2025: 1197−1206

[20] Xie Ruobing, Liu Zhiyuan, Jia Jia, et al. Representation learning of knowledge graphs with entity descriptions[C]//Proc of the 30th AAAI Conf on Artificial Intelligence. Palo Alto, CA: AAAI, 2016: 2659−2665

[21] Liu Ye, Li Hui, García-Durán A, et al. MMKG: Multi-modal knowledge graphs[C]//Proc of the 16th Extended Semantic Web Conf. Berlin: Springer, 2019: 459−474

[22] Dong Sibo, Shaheen I, Shen M, et al. ViSTA: Visual storytelling using multi-modal adapters for text-to-image diffusion models[C]//Proc of the IEEE/CVF Winter Conf on Applications of Computer Vision. Los Alamitos, CA: IEEE Computer Society, 2026: 12−21

[23] Zhang Wei, Zhou Dong'ao, Gong Yongshun, et al. Iterative two-stage schema-enhanced SQL generation with large language models[J]. Journal of Computer Research and Development, 2026, 63(5): 1261−1272 (in Chinese)
(张伟, 周东傲, 宫永顺, 等. 基于大语言模型的 2 阶段迭代模式增强 SQL 生成[J]. 计算机研究与发展, 2026, 63(5): 1261−1272)

[24] Wang Xiao, Alabdulmohsin I, Salz D, et al. Scaling pre-training to one hundred billion data for vision language models[C]//Proc of the IEEE/CVF Conf on Computer Vision and Pattern Recognition Findings. Piscataway, NJ: IEEE, 2026: 6185−6196

[25] Li Peining, Liang Meiyu, Huang Wei, et al. Multi-granularity multi-modal knowledge graph representation learning via subgraph-aware adaptive fusion and hierarchical relation modeling[C]//Proc of the ACM Web Conf 2026. New York: ACM, 2026: 3677−3685

[26] Cai Jie, Wang Xin, Guan Chaoyu, et al. Multimodal continual graph learning with neural architecture search[C]//Proc of the ACM Web Conf 2022. New York: ACM, 2022: 1292−1300

[27] Zhang Yichi, Chen Zhuo, Liang Lei, et al. Unleashing the power of imbalanced modality information for multi-modal knowledge graph completion[C]//Proc of the 2024 Joint Int Conf on Computational Linguistics, Language Resources and Evaluation. Torino, Italia: ELRA and ICCL, 2024: 17120−17130

[28] Semedo D, Magalhaes J. Diachronic cross-modal embeddings[C]//Proc of the 27th ACM Int Conf on Multimedia. New York: ACM, 2019: 2061−2069

[29] Guo Wenbin, Wang Xin, Li Zhao, et al. Attention-enhanced semantic fusion knowledge graph representation learning framework[J]. Journal of Computer Research and Development, 2025, 62(11): 2687−2696 (in Chinese)
(郭文斌, 王鑫, 李钊, 等. 注意力增强的语义融合知识图谱表示学习框架[J]. 计算机研究与发展, 2025, 62(11): 2687−2696)

[30] Liu Zhuang, Dong Zichen, Dong Yilin, et al. Lifelong graph learning: A comprehensive review[J]. Journal of Computer Research and Development, 2024, 61(8): 2067−2096 (in Chinese)
(刘壮, 董子宸, 董宜琳, 等. 图终身学习: 综述[J]. 计算机研究与发展, 2024, 61(8): 2067−2096)

[31] Wang Jia, Li Yawen, Xue Zhe, et al. Horizontal federated heterogeneous graph learning: A multi-scale adaptive solution to data distribution challenges[C]//Proc of the ACM on Web Conf 2025. New York: ACM, 2025: 4582−4591

[32] Fang Yi, Fan Dongzhe, Zha Daochen, et al. GaugLLM: Improving graph contrastive learning for text-attributed graphs with large language models[C]//Proc of the 30th ACM SIGKDD Conf on Knowledge Discovery and Data Mining. New York: ACM, 2024: 747−758

[33] Wang Meng, Wang Sen, Yang Han, et al. Is visual context really helpful for knowledge graph? A representation learning perspective [C]//Proc of the 29th ACM Int Conf on Multimedia. New York: ACM, 2021: 2735−2743

[34] Jian Yue, Luo Xiangyu, Li Zhifei, et al. APKGC: Noise-enhanced multi-modal knowledge graph completion with attention penalty[C]// Proc of the 39th AAAI Conf on Artificial Intelligence. Palo Alto, CA : AAAI, 2025: 15005−15013

[35] Radford A, Kim J W, Hallacy C, et al. Learning transferable visual models from natural language supervision[C]//Proc of the 38th Int Conf on Machine Learning. Cambridge, MA: PMLR, 2021: 8748−8763

[36] Efthymiou A, Rudinac S, Kackovic M, et al. VL-KGE: Vision-language models meet knowledge graph embeddings[C]//Proc of the ACM Web Conf 2026. New York: ACM, 2026: 7552−7563

[37] Cui Yuanning, Sun Zequn, Hu Wei. A pre-trained universal knowledge graph reasoning model based on rule prompts[J]. Journal of Computer Research and Development, 2024, 61(8): 2030−2044 (in Chinese)
(崔员宁, 孙泽群, 胡伟. 基于规则提示的知识图谱通用推理预训练模型[J]. 计算机研究与发展, 2024, 61(8): 2030−2044)

[38] Liu Jingping, Cen Zhedong, Zhang Weiyan, et al. Open domain relation reasoning method based on generative model[J]. Journal of Computer Research and Development, 2026, 63(2): 378−392 (in Chinese)
(刘井平, 岑哲栋, 张维彦, 等. 基于生成模型的开放域关系推理方法[J]. 计算机研究与发展, 2026, 63(2): 378−392)

[39] Wang Liqin, Gan Yifan, Wang Xu, et al. Textual and structural dual enhancement for knowledge graph completion with large language models[J]. Journal of Intelligent Information Systems, 2025, 63(5): 1625−1643

[40] Chen Xiang, Zhang Jingtian, Wang Xiaohan, et al. Continual multimodal knowledge graph construction[C]//Proc of the 33rd Int

Joint Conf on Artificial Intelligence. Palo Alto, CA: IJCAI, 2024: 6225−6233

[41] Wei Yuyang, Chen Wei, Wen Shiting, et al. Knowledge graph incremental embedding for unseen modalities[J]. Knowledge and Information Systems, 2023, 65(9): 3611−3631

[42] Zhou Dawei, Li Kaiwen, Ning Jingyi, et al. External knowledge injection for CLIP-based class-incremental learning[C]//Proc of the IEEE/CVF Int Conf on Computer Vision. Piscataway, NJ: IEEE, 2025: 3314−3325

[43] Rolnick D, Ahuja A, Schwarz J, et al. Experience replay for continual learning[C]//Proc of the 33rd Int Conf on Neural Information Processing Systems. Red Hook, NY: Curran Associates Inc., 2019: 350−360

[44] Lomonaco V, Maltoni D. CORe50: A new dataset and benchmark for continuous object recognition[C]//Proc of the 1st Annual Conf on Robot Learning. Cambridge, MA: PMLR, 2017: 17−26

[45] Zenke F, Poole B, Ganguli S. Continual learning through synaptic intelligence[C]//Proc of the 34th Int Conf on Machine Learning. Cambridge, MA: PMLR, 2017: 3987−3995



Wang Di, born in 1999. PhD candidate. Student member of CCF. Her main research interests include knowledge representation learning and artificial intelligence.

王　迪, 1999 年生。博士研究生。CCF 学生会员。主要研究方向为知识表示学习、人工智能。



Du Junping, born in 1963. Professor. Fellow of CCF. Her main research interests include artificial intelligence and pattern recognition.

杜军平, 1963 年生。教授。CCF 会士。主要研究方向为人工智能、模式识别。



Xue Zhe, born in 1987. Professor. Member of CCF. His main research interests include machine learning and data mining.

薛　哲, 1987 年生。教授。CCF 会员。主要研究方向为机器学习、数据挖掘。



Ye Guanhua, born in 1997. Associate professor. Member of CCF. His main research interests include machine learning and distributed collaboration.

叶冠华, 1997 年生。副教授。CCF 会员。主要研究方向为机器学习、分布式协作。



Zhao Tong, born in 1993. PhD candidate. Her main research interests include data mining and artificial intelligence.

赵　桐, 1993 年生。博士研究生。主要研究方向为数据挖掘、人工智能。



Meng Xiaolong, born in 2000. PhD candidate. His main research interests include data mining and large language model.

孟晓龙, 2000 年生。博士研究生。主要研究方向为数据挖掘、大语言模型。



Li Haisheng, born in 1974. Professor. Distinguished member of CCF. His main research interests include computer graphics and data mining.

李海生, 1974 年生。教授。CCF 杰出会员。主要研究方向为计算机图形学、数据挖掘。