

LBMKGC

Large Model-Driven Balanced Multimodal Knowledge Graph Completion

大模型驱动的均衡多模态知识图谱补全

组会汇报 · Paper Reading

NeurIPS 2025 | Yuan Guo, Qian Ma, Hui Li et al.

大连海事大学 · 江南大学 · 东北大学

汇报大纲

Outline

- 1. 研究背景与三大核心挑战
- 2. LBMKGC 整体框架概览
- 3. CMoA
- 4. CGuAF
- 5. 主实验结果 (RQ1)
- 6. 消融实验分析 (RQ2)
- 7. 模态缺失鲁棒性 (RQ3)
- 8. 案例研究：自适应模态权重 (RQ4)

研究背景与三大挑战

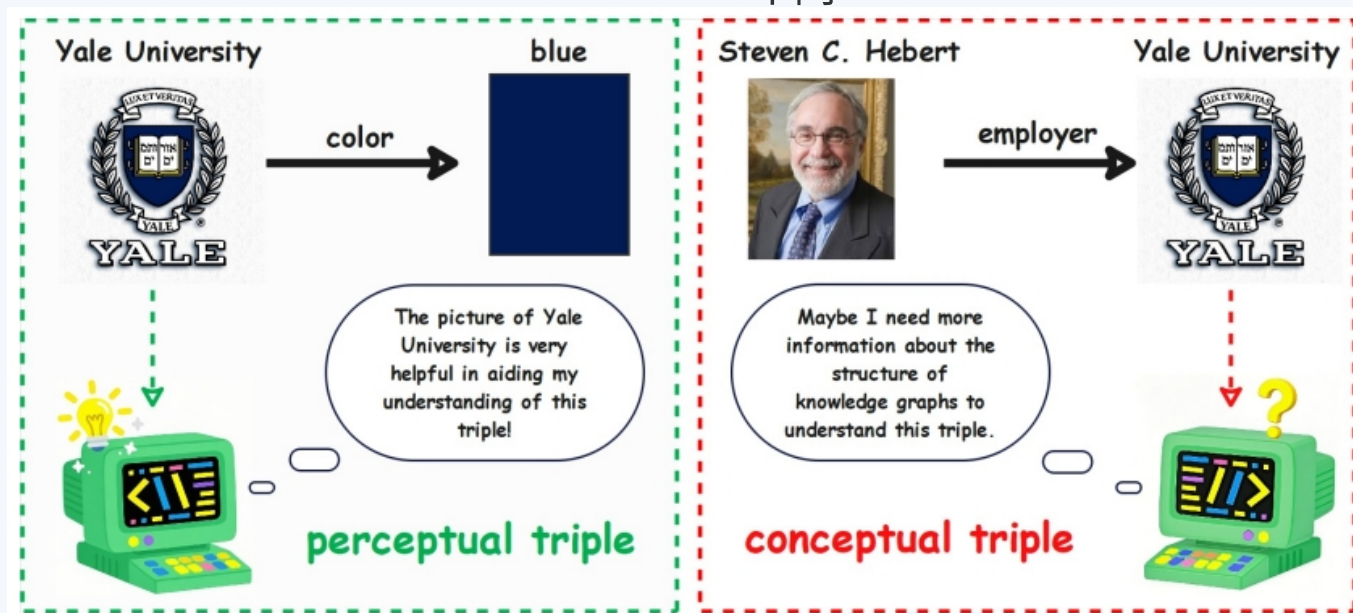
Motivation & Three Key Challenges

背景 / Background

- **MMKGC**: 融合文本、图像、视频等多模态信息以补全缺失三元组
- **应用**: 推荐系统、视觉问答、自然语言处理等
- **主流方法**: 多模态融合机制 + 负采样优化

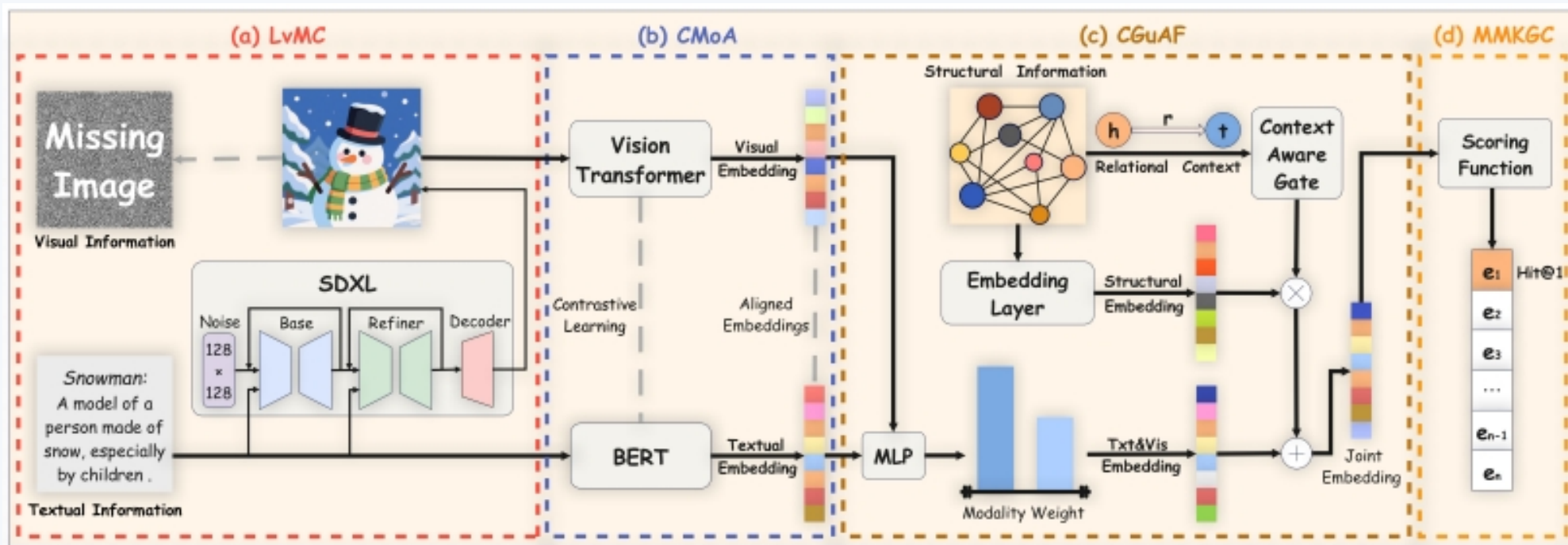
现有方法的三大瓶颈

- **跨实体模态信息不均衡**: 部分实体缺少图像，依赖纯文本
- **实体内部多模态异构性**: 图像2D像素 vs 文本符号语义鸿沟
- **模态贡献度不一致**: 感知型三元组 vs 概念型三元组需求不同



LBMKGC 整体框架

Overall Framework



(a) LvMC

大模型视觉补全
从实体的文本描述生成图
像。

(b) CMoA

跨模态语义对齐
CLIP 对比学习

(c) CGuAF

上下文自适应融合
感知/概念动态权重

(d) MMKGC

三元组建模
RotatE 评分

跨模态语义对齐

CMoA

将图像和文本拉到同一语义空间，解决模态异构问题

step ①

特征提取

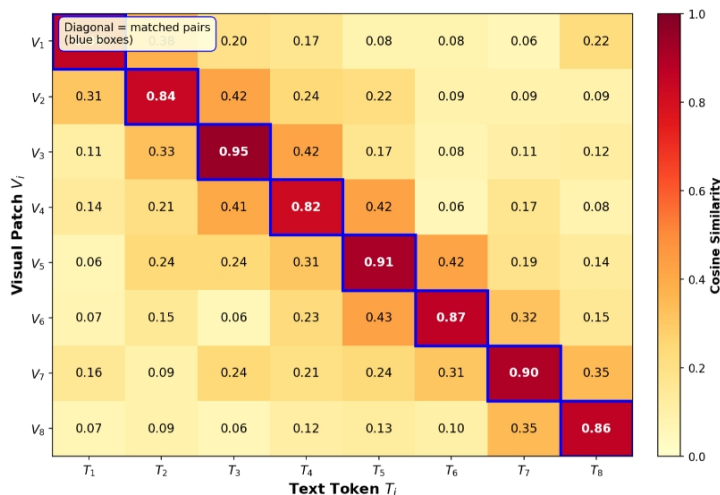
文本: $ft = \text{BERT}(e_text)$

图像: $fv = \text{ViT}(e_image)$

得到两个N维向量

step ②

计算相似
度矩阵



$$\mathcal{L}_{CLIP} = -\frac{1}{2N} \sum_{i=1}^N \left[\log \frac{e^{S_{i,i}/\tau}}{\sum_{j=1}^N e^{S_{i,j}/\tau}} + \log \frac{e^{S_{i,i}/\tau}}{\sum_{j=1}^N e^{S_{j,i}/\tau}} \right]$$

利用CLIP原版的损失函数，让对角线上对应的文本token和图像块特征相似度更高。

step ③

投影到同
一向量空
间

对齐后的特征通过MLP+ReLU投影到同一向量空间

对齐后的特征 $\xrightarrow{MLP} e_t^{emb}, e_v^{emb}$

上下文自适应融合

CGuAF

核心问题

同一个实体在不同三元组里，模态贡献不一样。

解决方法

两阶段动态调整

先按实体本身调模态权重

再按关系类型调结构-多模态占比

完整流程

1. 模态加权

$$\omega_m(e) = \frac{\exp(V \cdot \tanh(e_m^{emb}))}{\sum_{n \in M} \exp(V \cdot \tanh(e_n^{emb}))}$$

V可学习，让模型自己发现哪些实体偏视觉、哪些实体偏文本。

"蓝色"跟V点积大→视觉权重大
"法律"跟V点积大→文本权重大

2. 模态融合

$$h_{mm}^{emb} = \sum_{m \in M} \omega_m \cdot e_m^{emb}$$

加权求和，得到实体的多模态融合表示。

3. 关系温度调制

$$h_{joint}^{emb} = \zeta_r \cdot e_s^{emb} + (1 - \zeta_r) \cdot h_{mm}^{emb}$$

ζ_r 是关系r专属的温度系数，取值 (0, 1)，由关系类型决定。

ζ_r 大 → 头/尾实体的联合嵌入更偏向**结构信息**
 ζ_r 小 → 更偏向**多模态融**

主实验结果：超越 21 个 SOTA 基线

Main Results (RQ1)

关键结论

3 个数据集上 Hit@3 / Hit@10 全面领先

表 1: KGC 模型在三个基准上的链接预测对比

Table 1: Knowledge Graph Completion Models Comparison												
Model	DB15K				MKG-W				MKG-Y			
	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10
TransE	24.86	12.78	31.48	47.07	29.19	21.06	33.20	44.23	30.73	23.45	35.18	43.37
DistMult	23.03	14.78	26.28	39.59	20.99	15.93	22.28	30.86	25.04	19.33	27.80	35.95
ComplEx	27.48	18.37	31.57	45.37	24.93	19.06	26.69	36.73	28.71	22.26	32.12	40.93
RotatE	29.28	17.87	36.12	49.66	33.67	26.80	36.68	46.73	34.95	29.10	38.35	45.30
PairRE	31.13	21.62	35.91	49.30	34.40	28.24	36.71	46.04	32.01	25.53	35.84	43.89
GC-OTE	31.85	22.11	36.52	51.18	33.92	26.55	39.65	46.90	32.95	26.77	38.41	44.08
IKRL	26.82	14.09	34.93	49.09	32.36	26.11	34.75	44.07	33.22	30.37	34.28	38.26
TBKGC	28.40	15.61	37.03	49.86	31.48	25.31	33.98	43.24	33.99	30.47	35.27	40.07
TransAE	28.09	21.25	31.17	41.17	30.00	21.23	34.91	44.72	28.10	25.31	29.10	33.03
MMKRL	26.81	13.85	35.07	49.39	30.10	22.16	34.09	44.69	36.81	31.66	39.79	45.31
RSME	29.76	24.15	32.12	40.29	29.23	23.36	31.97	40.43	34.44	31.78	36.07	39.09
OTKGE	23.86	18.45	25.89	34.23	34.36	28.85	36.25	44.88	35.51	31.97	37.18	41.38
IMF	32.25	24.20	36.00	48.19	34.50	28.77	36.62	45.44	35.79	32.95	37.14	40.63
QEB	28.18	14.82	36.67	51.55	32.38	25.47	35.06	45.32	34.37	29.49	36.95	42.32
VISTA	30.42	22.49	33.56	45.94	32.91	26.12	35.38	45.61	30.45	24.87	32.39	41.53
AdaMF	32.51	21.31	39.67	51.68	34.27	27.21	37.86	47.12	38.06	33.49	40.44	45.48
KBGAN	25.73	9.91	36.95	51.93	29.47	22.21	34.87	40.64	29.71	22.81	34.88	40.21
MANS	28.82	16.87	36.58	49.26	30.88	24.89	33.63	41.78	29.03	25.25	31.35	34.49
MMRNS	32.68	23.01	37.86	51.01	35.03	28.59	37.49	47.47	35.93	30.53	39.07	45.47
AdaMF-MAT	35.14	25.30	41.11	<u>52.92</u>	35.77	28.87	<u>38.85</u>	<u>48.71</u>	38.57	<u>34.34</u>	<u>40.59</u>	<u>45.76</u>
MyGO	<u>37.17</u>	29.62	<u>41.66</u>	51.87	<u>36.09</u>	<u>30.02</u>	38.26	46.89	<u>38.66</u>	34.85	39.96	44.37
Ours	37.23	<u>27.78</u>	42.75	54.71	38.46	31.20	41.78	51.46	40.03	33.89	43.11	50.81
	+0.16%	-	+2.6%	+3.38%	+6.6%	+3.9%	+7.5%	+5.6%	+3.5%	-	+6.2%	+11.0%

消融实验：每个模块都不可或缺

Ablation Study (RQ2)

表 2：消融结果 (DB15K / MKG-W / MKG-Y)

Table 2: Ablation Study												
Settings	DB15K				MKG-W				MKG-Y			
	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10
LBMKGC	37.23	27.78	42.75	54.71	38.46	31.20	41.78	51.46	40.03	33.89	43.11	50.81
w/o LvMC	36.73	27.19	42.01	53.76	37.12	29.43	40.28	49.16	39.40	33.42	42.47	49.59
w/o CMoA	36.81	27.69	41.56	53.37	38.03	29.88	40.86	49.34	39.34	33.58	42.28	49.37
w/o CGuAF	36.36	26.55	42.39	54.48	37.53	30.51	41.76	50.77	39.29	33.01	42.92	50.86
w/o e_s^{emb}	31.21	16.96	41.12	54.60	34.76	23.84	41.86	51.97	37.79	29.33	42.43	51.02
w/o e_t^{emb}	36.51	27.09	41.45	53.19	36.72	30.01	40.02	48.79	38.33	32.72	40.95	47.67
w/o e_v^{emb}	37.08	27.34	42.01	53.92	37.89	30.63	41.05	50.89	39.24	33.13	41.76	49.73

消融设置

- w/o LvMC：移除 SDXL 视觉补全
- w/o CMoA：用 VGG+BERT 替代 CLIP 对齐
- w/o CGuAF：替换为均值融合
- w/o $e_s^{emb}/e_t^{emb}/e_v^{emb}$ ：去除结构/文本/视觉

每个模块、每种模态都带来正向贡献,且 CGuAF 模块提升最显著。

模态缺失下的鲁棒性

Robustness under Missing Modality (RQ3)

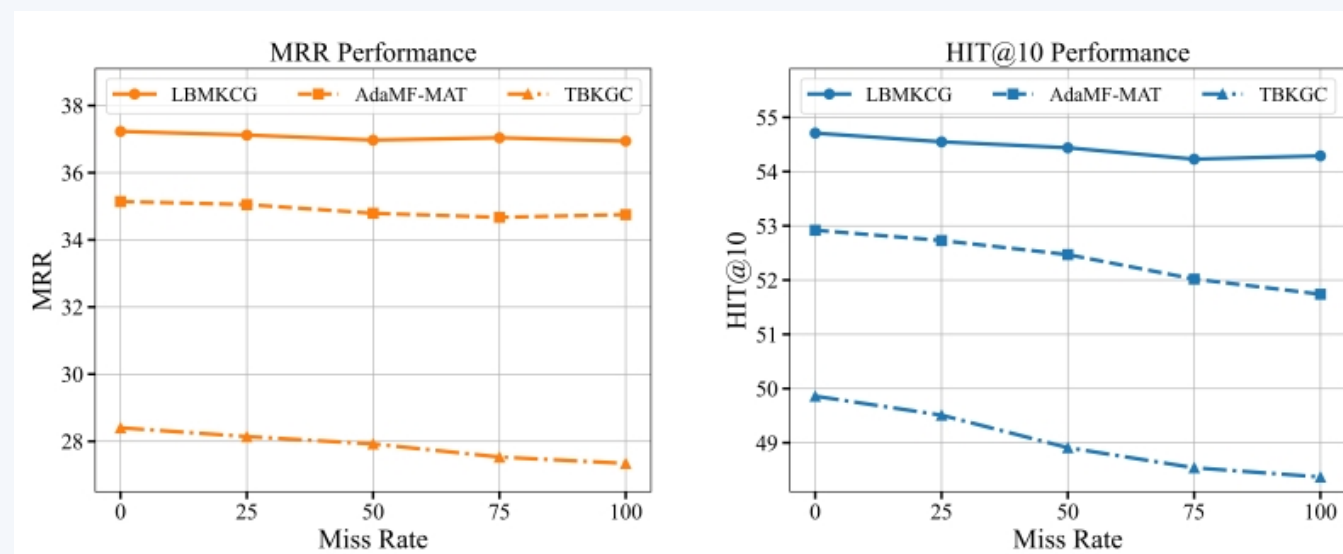
实验设置

- 数据集：DB15K
- 图像丢弃比例 $\eta \in \{0, 0.25, 0.5, 0.75, 1.0\}$
- LBMKGC：LvMC显式补全缺失图像

关键观察

- 随缺失率上升，TBKGC / AdaMF-MAT 显著下降
- LBMKGC 性能基本平稳，曲线接近水平

图 3：不同缺失率下的 MRR 与 Hit@10



案例研究：自适应模态权重

Case Study - Adaptive Modality Weights (RQ4)

CGuAF 自适应行为

- 实体模态权重由可学习向量调节
 - 关系温度让同一实体在不同关系下动态调整
 - 概念型实体 → 文本权重更高
 - 感知型实体 → 视觉权重更高
- 案例1: Power_是一部电影，属于概念型实体，文本权重远高于视觉权重。
- 案例2: Australian_National_University是一所大学，属于感知型，视觉模态占比会更高一些。

图 4: MKG-W 视觉模态权重分布 & 文本-视觉对比

