# Unleashing the Power of Imbalanced Modality Information for Multi-modal Knowledge Graph Completion

Yichi Zhang, Zhuo Chen, Lei Liang, Huajun Chen, Wen Zhang

COLING 2024

Presenter: Yichi Zhang

Date: 04/30/2024

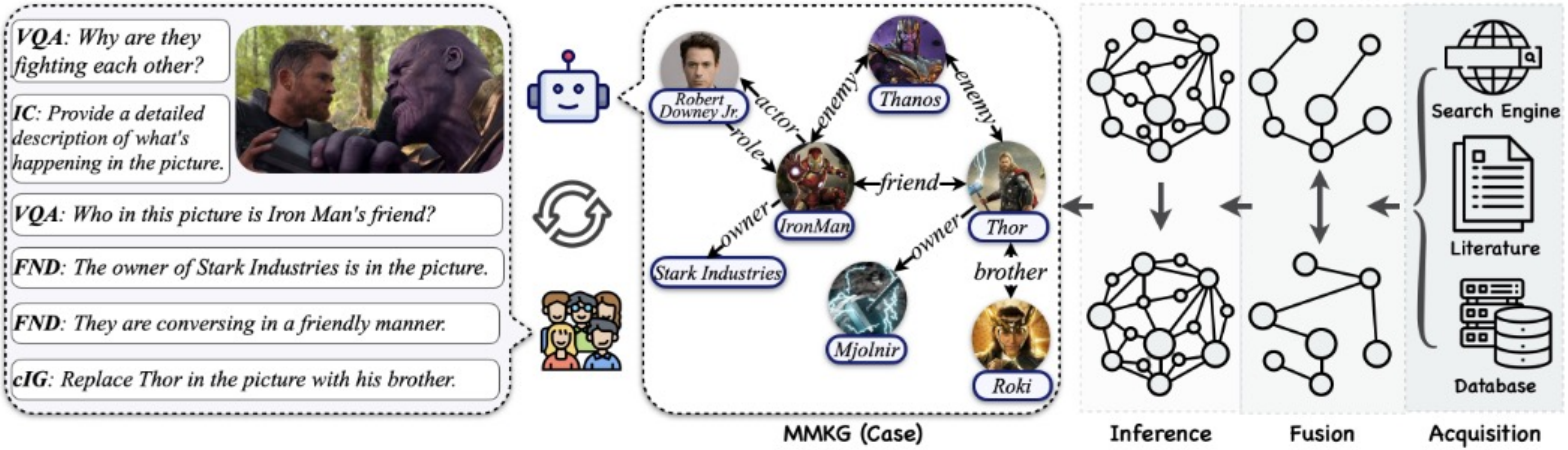
蚂蚁集团 ANT GROUP

ZHEJIANG UNIVERSITY 1897

# 01. Research Background

- Multi-modal Knowledge Graph
- Imbalance of Multi-modal Knowledge Graphs

# Multi-modal Knowledge Graphs

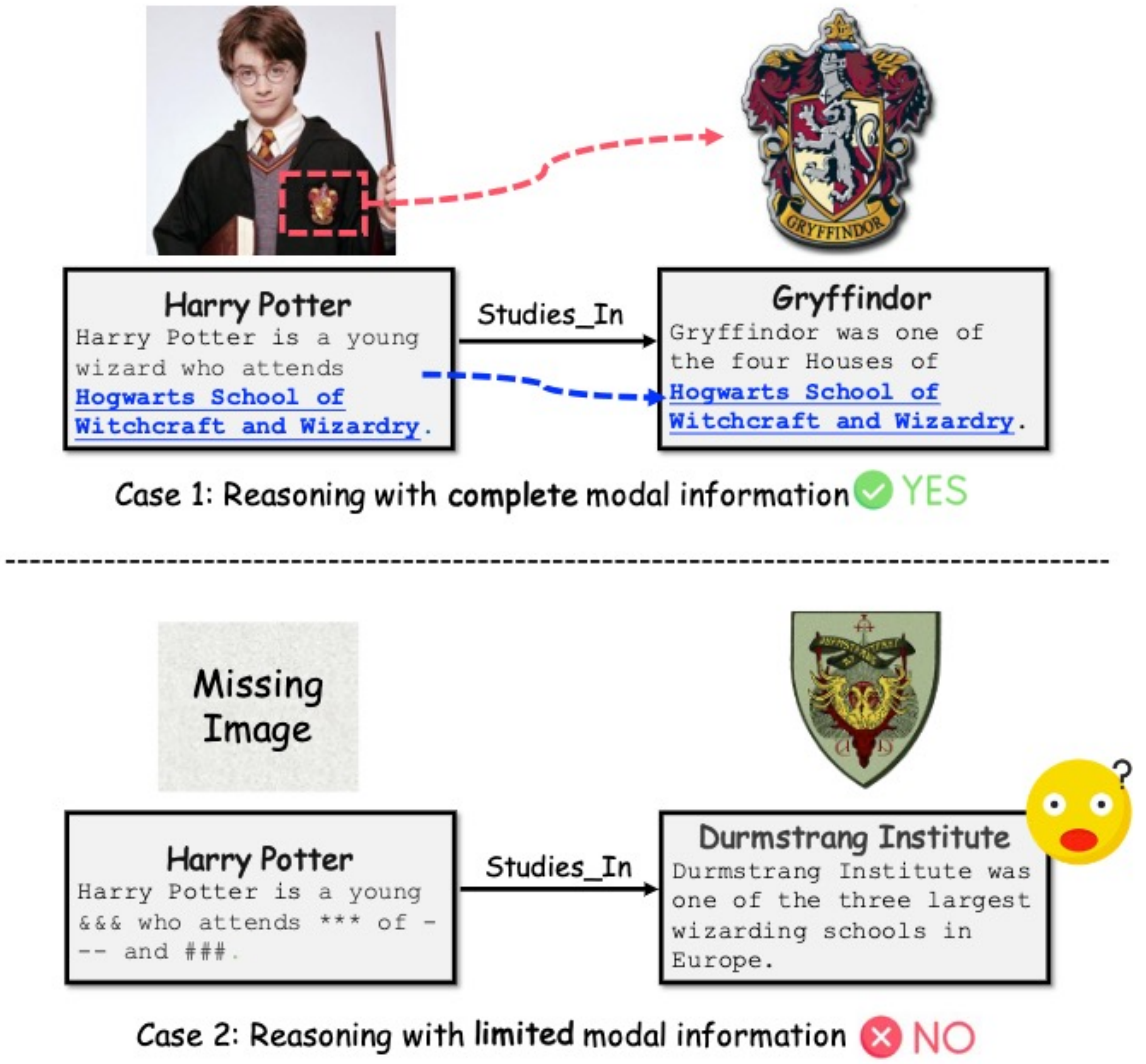



Multi-modal Knowledge Graphs

[Preprint] Knowledge Graphs Meet Multi-Modal Learning: A Comprehensive Survey

# Knowledge Graph Completion


Multi-modal Knowledge Graph Completion

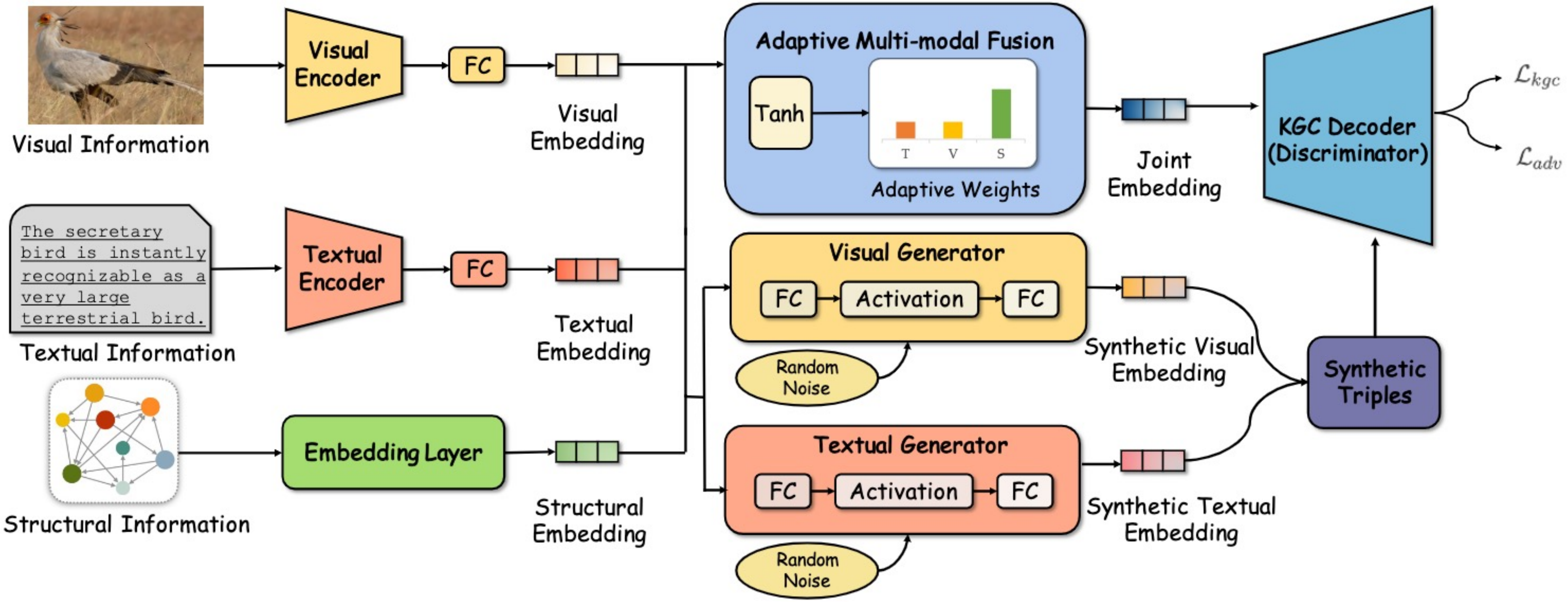



Imbalance in Multi-modal Knowledge Graph

# 02. Our Method

- Overview
- Adaptive Multi-modal Fusion
- Modality Adversarial Training

# Overview of Our Method




Overview of Our AdaMF-MAT framework

# Multi-modal Feature Encoding

Encoding Visual and Textual Information for each entity

$$f_v = \frac{1}{|\mathcal{V}_e|} \sum_{img_i \in \mathcal{V}_e} \mathrm{PVE}(img_i)$$

$$e_v = \mathbf{W}_v \cdot f_v + \mathbf{b}_v$$

# Adaptive Multi-modal Fusion

- Get the joint embedding of each entity:

$$\alpha_m = \frac{\exp(w_m \oplus \tanh(e_m))}{\sum_{n\in\mathcal{M}} \exp(w_n \oplus \tanh(e_n))} \qquad e_{joint} = \sum_{m\in\mathcal{M}} \alpha_m e_m$$

- Triple plausibility

$$\mathcal{F}(h, r, t) = ||h_{joint} \circ r - t_{joint}||$$

- Training MMKGC model:

$$\mathcal{L}_{kgc} = \frac{1}{|\mathcal{T}|} \sum_{(h,r,t)\in\mathcal{T}} \Big( -\log \sigma(\gamma - \mathcal{F}(h, r, t)) - \sum_{i=1}^{K} p_i \log \sigma(\mathcal{F}(h'_i, r'_i, t'_i) - \gamma) \Big)$$

# Modality Adversarial Training

- Multi-modal Embedding Generator

$$\mathbf{G}_m(e_s, z) = \mathbf{W}_2 \cdot \delta(\mathbf{W}_1 \cdot [e_s; z] + \mathbf{b}_1) + \mathbf{b}_2$$

- Synthetic Embedding → Synthetic Entity → Synthetic Triples
- Adversarial Training Loss:

$$\mathcal{L}_{adv} = \frac{1}{|\mathcal{T}|} \sum_{(h,r,t)\in\mathcal{T}} \Big( -\log\sigma(\gamma - \mathcal{F}(h,r,t)) - \frac{1}{|\mathcal{S}|} \sum_{\substack{(h^*,r^*,t^*) \\ \in \mathcal{S}(h,r,t)}} \log\sigma(\mathcal{F}(h^*,r^*,t^*) - \gamma) \Big)$$

# Training Objective

The pseudo-code of our framework:

- Iterative training for G and D
- Multi-task learning

$$\min_{\mathbf{D}} \mathcal{L}_{kgc} + \min_{\mathbf{D}} \max_{\mathbf{G}} \lambda \mathcal{L}_{adv}$$

**Algorithm 1:** Pseudo-code for training AdaMF-MAT

**Input:** A batch of training triple $\mathcal{B}$ sampled from $\mathcal{T}$, the multi-modal information of the entities', the AdaMF model $\mathbf{D}$, the generator $\mathbf{G}$.

**Output:** The MMKGC model $\mathbf{D}$ trained with MAT.

**for** *each triple* $(h, r, t) \in \mathcal{B}$ **do**
- // Training $\mathbf{D}$
- Get the joint embeddings $\boldsymbol{h}_{joint}, \boldsymbol{t}_{joint}$.
- Calculate the triple score $\mathcal{F}(h, r, t)$ and the nagative triple scores $\mathcal{F}(h'_i, r'_i, t'_i)$.
- Calculate the kgc loss $\mathcal{L}_{kgc}$.
- Generate the adversarial example set $\mathcal{S}$ with $\mathbf{G}$.
- Calculate the adversarial loss $\mathcal{L}_{adv}$.
- Calculate the overall loss $\mathcal{L}_{kgc} + \lambda \mathcal{L}_{adv}$.
- Back propagation and optimize $\mathbf{D}$.
- // Training $\mathbf{G}$.
- Get the joint embeddings $\boldsymbol{h}_{joint}, \boldsymbol{t}_{joint}$.
- Generate the adversarial example set $\mathcal{S}$ with $\mathbf{G}$.
- Calculate the adversarial loss $\mathcal{L}_{adv}$.
- Back propagation and optimize $\mathbf{G}$.

**end**

# 03. Experiments and Evaluation

- Experiment Settings
- Main Results
- Further Exploration

# Experiment Settings

- **Dataset:** DB15K, MKG-W, MKG-Y
- **Task**: Link Prediction (Knowledge Graph Completion)
- **Evaluation Metrics**: MRR, Hit@K (K=1,3,10)

Table 1: Statistical information of the benchmarks.

| Dataset | $\lvert\mathcal{E}\rvert$ | $\lvert\mathcal{R}\rvert$ | #Train | #Valid | #Test |
|---|---|---|---|---|---|
| DB15K | 12842 | 279 | 79222 | 9902 | 9904 |
| MKG-W | 15000 | 169 | 34196 | 4276 | 4274 |
| MKG-Y | 15000 | 28 | 21310 | 2665 | 2663 |

$$\mathbf{MRR} = \frac{1}{|\mathcal{T}_{test}|}\sum_{i=1}^{|\mathcal{T}_{test}|}\left(\frac{1}{r_{h,i}} + \frac{1}{r_{t,i}}\right)$$

$$\mathbf{Hit@K} = \frac{1}{|\mathcal{T}_{test}|}\sum_{i=1}^{|\mathcal{T}_{test}|}\left(\mathbf{1}(r_{h,i} \leq K) + \mathbf{1}(r_{t,i} \leq K)\right)$$

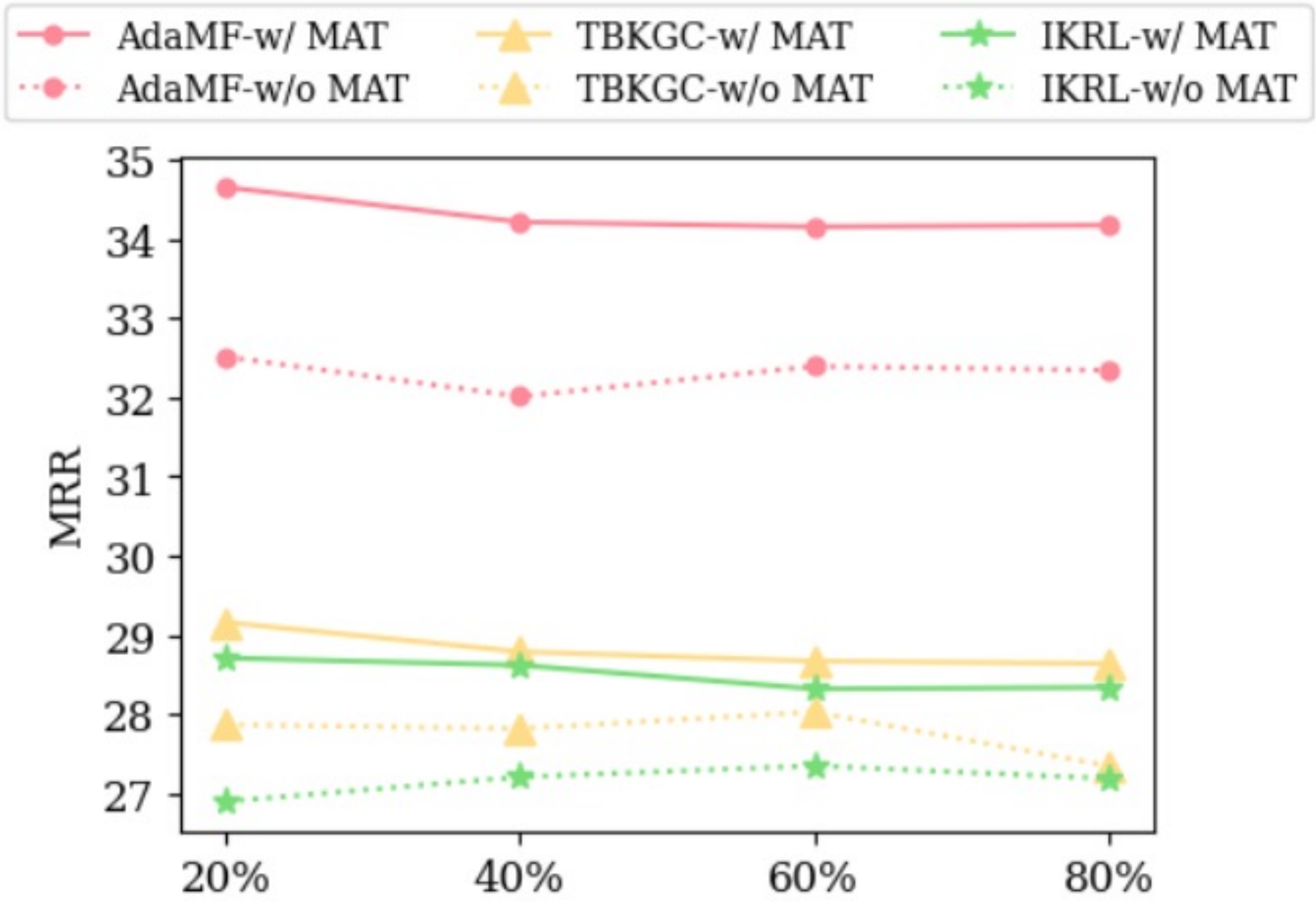
# Main Experiment Results

- **Link Prediction Results on Three MMKGs**

| 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 |
| Unimodal KGC | 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 |
| | TransD | 21.52 | 8.34 | 29.93 | 44.24 | 25.56 | 15.88 | 32.99 | 40.18 | 26.39 | 17.01 | 33.60 | 40.31 |
| | 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 | 19.09 | 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 | 35.96 | 46.05 | 32.95 | 26.77 | 36.44 | 44.08 |
| Multi-modal KGC | 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 | <u>24.15</u> | 32.12 | 40.29 | 29.23 | 23.36 | 31.97 | 40.43 | 34.44 | 31.78 | 36.07 | 39.09 |
| | VBKGC | 30.61 | 19.75 | 37.18 | 49.44 | 30.61 | 24.91 | 33.01 | 40.88 | 37.04 | <u>33.76</u> | 38.75 | 42.30 |
| | OTKGE | 23.86 | 18.45 | 25.89 | 34.23 | 34.36 | <u>28.85</u> | 36.25 | 44.88 | 35.51 | 31.97 | 37.18 | 41.38 |
| Negative Sampling | KBGAN(TransE) | 25.73 | 9.91 | 36.95 | 51.93 | 29.47 | 22.21 | 34.87 | 40.64 | 29.71 | 22.81 | 34.88 | 40.21 |
| | KBGAN(TransD) | 23.74 | 9.34 | 33.51 | 47.94 | 29.67 | 22.38 | 35.24 | 40.80 | 28.73 | 20.99 | 34.64 | 40.76 |
| | 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(RotatE) | 29.67 | 17.89 | 36.66 | 51.01 | 34.13 | 27.37 | 37.48 | 46.82 | 35.93 | 30.53 | 39.07 | 45.47 |
| | MMRNS(SOTA) | <u>32.68</u> | 23.01 | 37.86 | 51.01 | <u>35.03</u> | 28.59 | 37.49 | <u>47.47</u> | 35.93 | 30.53 | 39.07 | 45.47 |
| Ours | AdaMF | 32.51 | 21.31 | <u>39.67</u> | <u>51.68</u> | 34.27 | 27.21 | <u>37.86</u> | 47.21 | <u>38.06</u> | 33.49 | <u>40.44</u> | <u>45.48</u> |
| | AdaMF-MAT | **35.14** | **25.30** | **41.11** | **52.92** | **35.85** | **29.04** | **39.01** | **48.42** | **38.57** | **34.34** | **40.59** | **45.76** |

# Further Exploration

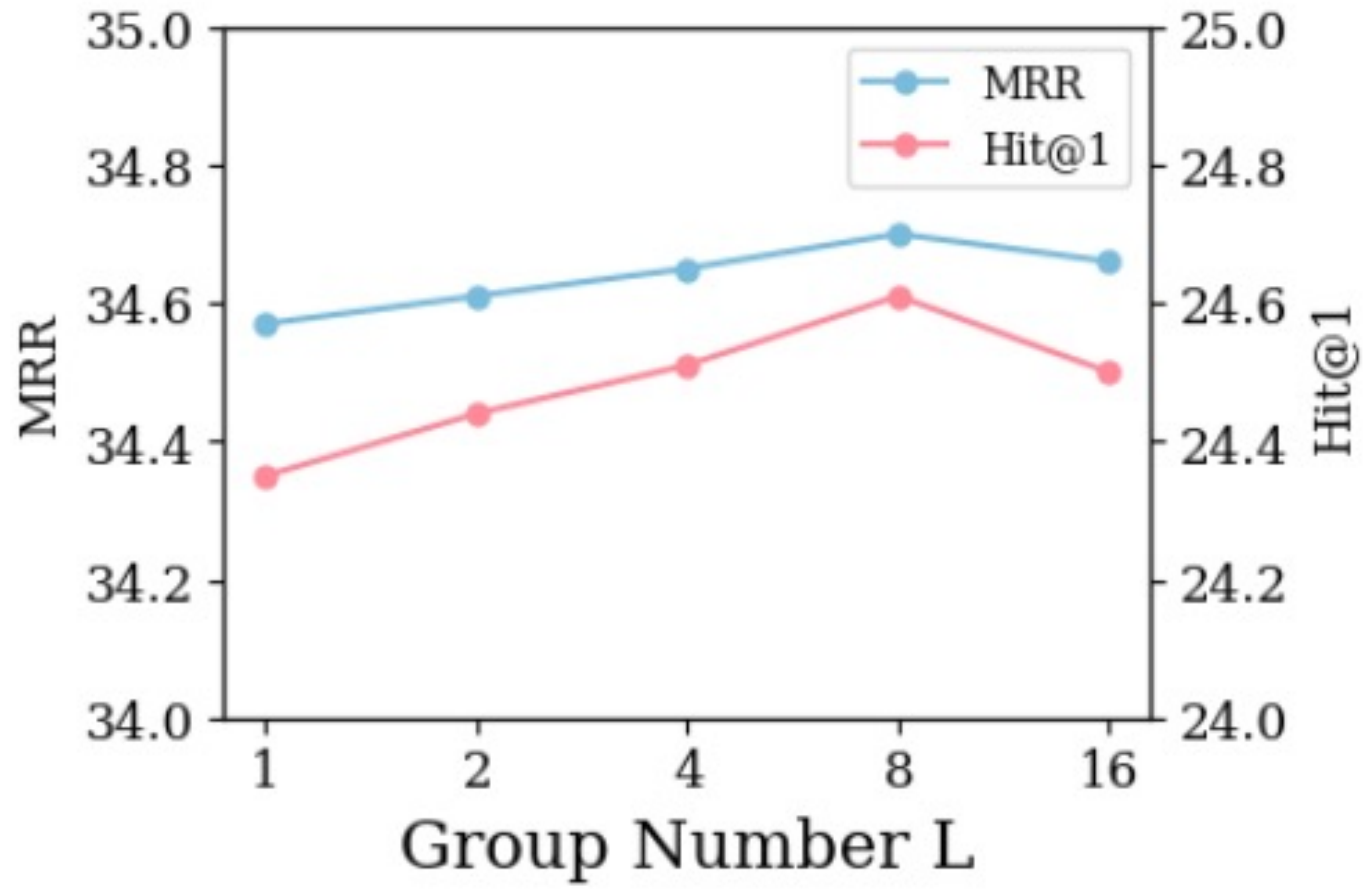
- **Imbalanced Link Prediction Results on DB15K**



# Further Exploration

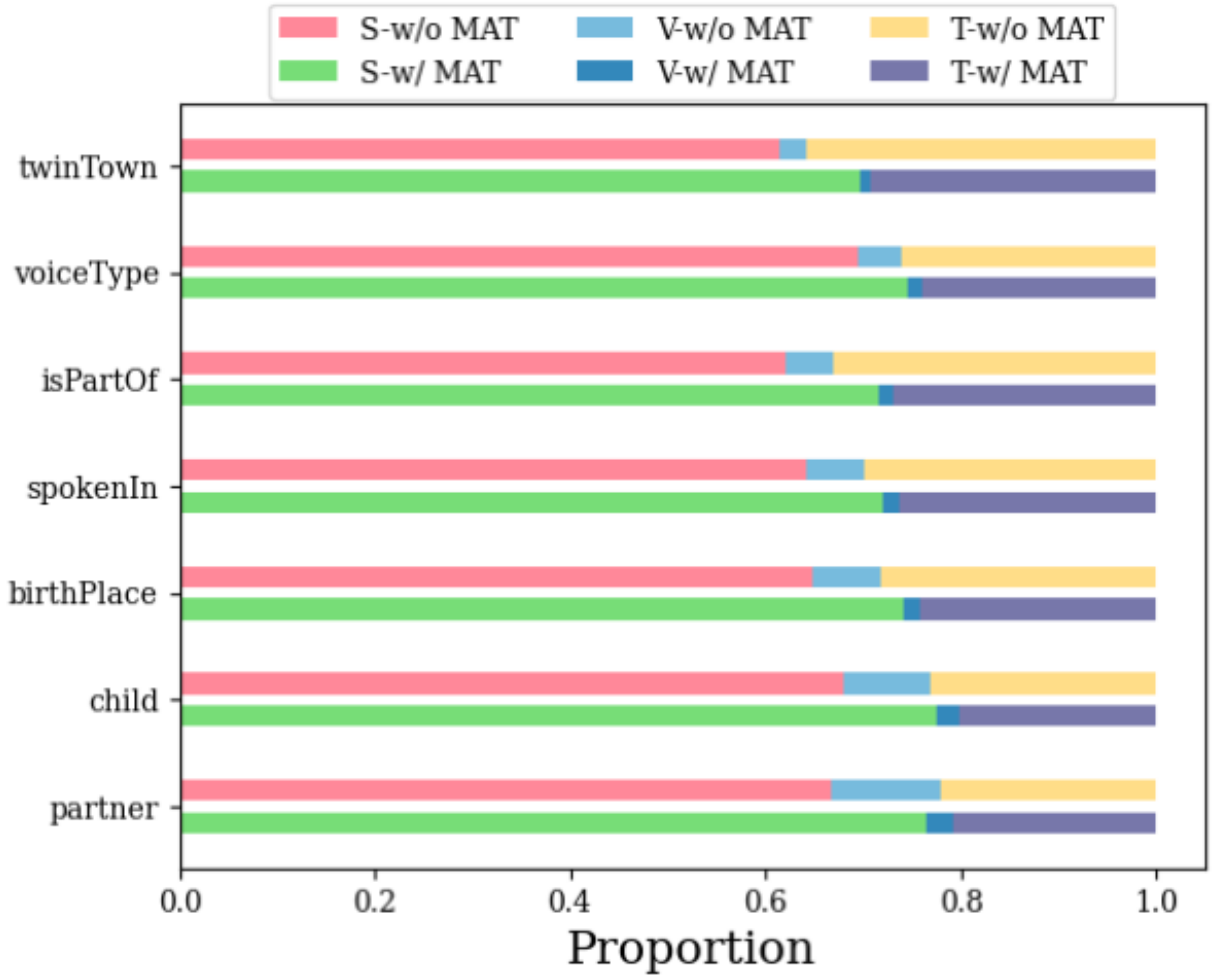
- **Ablation Study Results**

| Model | | MRR | Hit@1 | Hit@3 | Hit@10 |
|---|---|---|---|---|---|
| AdaMF-MAT | | 35.14 | 25.30 | 41.11 | 52.92 |
| AdaMF (w/o MAT) | S+V+T(Adaptive) | 33.19 | 23.08 | 40.34 | 52.47 |
| | S+V+T(Mean) | 32.57 | 21.45 | 39.71 | 51.68 |
| | S+V(w/o T) | 32.34 | 21.84 | 38.90 | 50.76 |
| | S+T(w/o V) | 31.82 | 19.63 | 39.69 | 52.51 |
| | V+T(w/o S) | 31.01 | 18.45 | 39.38 | 52.27 |
| MAT | w/o $(h^*, r, t)$ | 34.64 | 24.52 | 40.98 | 52.49 |
| | w/o $(h, r, t^*)$ | 34.65 | 24.49 | 41.13 | 52.61 |
| | w/o $(h^*, r, t^*)$ | 34.61 | 24.36 | 40.98 | 52.65 |



# Further Exploration

- **Adaptive Weight Visualization**



# 04. Conclusion

- Conclusion
- Our Future Work

# Conclusion

- In this paper, we mainly discuss the problem of utilizing modal information in MMKGC and propose a novel MMKGC framework called AdaMF-MAT to address the limitations of the existing methods.
- Our method AdaMF-MAT employs adaptive modal fusion to utilize the multi-modal information diversely and augment the multi-modal embeddings through modality-adversarial training. Experiments demonstrate that AdaMF-MAT can outperform all the existing baseline methods and achieve SOTA results in MMKGC tasks.

# Future Work

- [SIGIR 2024] Multi-modal Knowledge Graph Completion in the Wild
- [Preprint] MyGO: Discrete Modality Information as Fine-Grained Tokens for Multi-modal Knowledge Graph Completion
- [Preprint] The Power of Noise: Toward a Unified Multi-modal Knowledge Graph Representation Framework
- [Preprint] Knowledge Graphs Meet Multi-Modal Learning: A Comprehensive Survey

# 谢谢！

蚂蚁集团
ANT GROUP