ACM Multimedia 2023
Ottawa, Canada

MEAformer: Multi-modal Entity Alignment Transformer for Meta Modality Hybrid

Zhuo Chen[1], Jiaoyan Chen[2], Wen Zhang[1‡], Lingbing Guo[1], Yin Fang[1], Yufeng Huang[1], Yichi Zhang[1],

Yuxia Geng[1], Jeff Z. Pan[3], Wenting Song[4], Huajun Chen[1,5]

https://github.com/zjukg/MEAformer

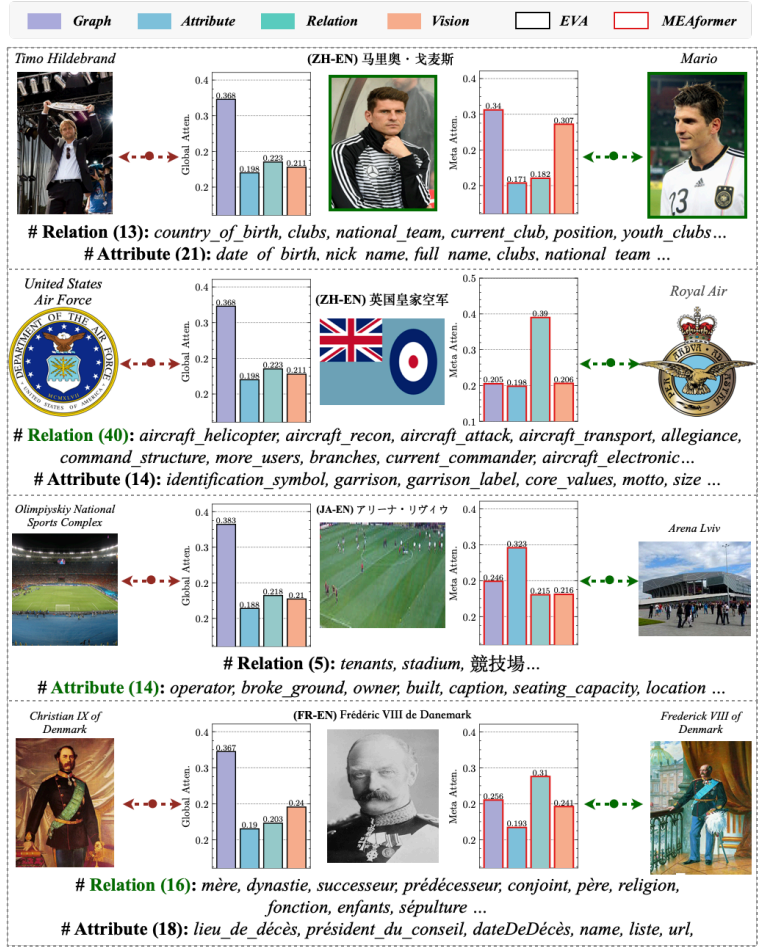
1 Zhejiang University 2 The University of Manchester and The University of Oxford
3 The University of Edinburgh 4 Huawei Technologies Ltd 5 Donghai laboratory

ZHEJIANG UNIVERSITY · MANCHESTER 1824 The University of Manchester · UNIVERSITY OF OXFORD · THE UNIVERSITY OF EDINBURGH · HUAWEI

Background:

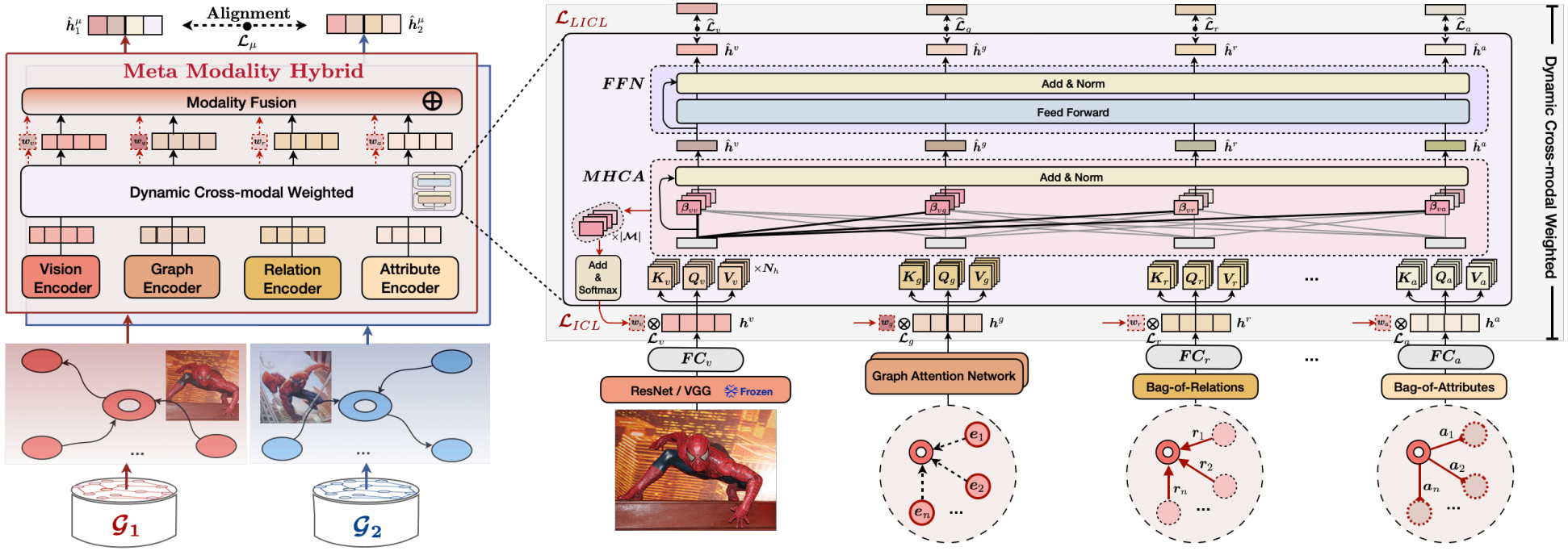
◆**Static modality weights in EVA (left) and dynamic meta modality weights in our MEAformer (right).**

- **Previous MMEA Methods:**
  - ***KG-level weights for modality fusion***:
    - *Static modality weights*
  - ***Disregard intra-modal discrepancies*** :
    - *E.g., node degrees or relation numbers*
  - ***Disregard inter-modal preferences*** :
    - *E.g., modality absence or ambiguity*



Our MEAformer

- ***An alternative MMEA paradigm that generates reasonable multi-modal entity hybrid features to foster modality preferences adaptable to entities***



(a) MEAformer. (b) Meta modality hybrid.

- ***Entity-level weights for modality fusion***
- ***Meta-learning alike paradigm***
- ***Modal-aware hard entity replay***

Main Experiments

◎ Dataset

(a) ***Bilingual:***

- *DBP15K$_{ZH\text{-}EN}$, DBP15K$_{JA\text{-}EN}$ and DBP15K$_{FR\text{-}EN}$*
- *Seed alignments (R_{sa} = 0.3)*

(b) ***Monolingual:***

- *FBDB15K and FBYG15K from MMKG*
- *$R_{sa} \in \{0.2, 0.5, 0.8\}$*

Dataset	KG	# Ent.	# Rel.	# Attr.	# Rel. Triples	# Attr. Triples	# Image	# EA pairs
DBP15K$_{ZH-EN}$	ZH (Chinese)	19,388	1,701	8,111	70,414	248,035	15,912	15,000
	EN (English)	19,572	1,323	7,173	95,142	343,218	14,125	
DBP15K$_{JA-EN}$	JA (Japanese)	19,814	1,299	5,882	77,214	248,991	12,739	15,000
	EN (English)	19,780	1,153	6,066	93,484	320,616	13,741	
DBP15K$_{FR-EN}$	FR (French)	19,661	903	4,547	105,998	273,825	14,174	15,000
	EN (English)	19,993	1,208	6,422	115,722	351,094	13,858	
FBDB15K	FB15K	14,951	1,345	116	592,213	29,395	13,444	12,846
	DB15K	12,842	279	225	89,197	48,080	12,837	
FBYG15K	FB15K	14,951	1,345	116	592,213	29,395	13,444	11,199
	YAGO15K	15,404	32	7	122,886	23,532	11,194	

◎ Iterative Training

- *Concretely, every K_e (where K_e = 5) epochs, we propose cross-KG entity pairs that are **mutual nearest neighbors** in the vector space and add them to a **candidate list** N_{cd}*
- *An entity pair in N_{cd} will be added into the training set if it **remains a mutual nearest neighbor** for K_s (= 10) consecutive rounds.*

◎ Unsupervised Training

- *Utilizing the visual similarity of entities to construct a **pseudo seed dictionary** (S_{dic}) to eliminate the reliance on gold labels.*
- ***Surface** and **visual** modality information serve as entity reference to obtain unsupervised results.*

Main Experiments

- **Non-iter. Results on Bilingual Datasets**

	Models	Para.	DBP15K$_{ZH-EN}$			DBP15K$_{JA-EN}$			DBP15K$_{FR-EN}$		
			H@1	H@10	MRR	H@1	H@10	MRR	H@1	H@10	MRR
w/o SF	MUGNN [3]	-	.494	.844	.611	.501	.857	.621	.495	.870	.621
	AliNet [32]	-	.539	.826	.628	.549	.831	.645	.552	.852	.657
	EVA* [21]	13.3M	.680	.910	.762	.673	.908	.757	.683	.923	.767
	MSNEA* [5]	14.1M	.601	.830	.684	.535	.775	.617	.543	.801	.630
	MCLEA* [20]	13.2M	.715	.923	.788	.715	.909	.785	.711	.909	.782
	MEAformer (Ours)	13.7M	**.771**	**.951**	**.835**	**.764**	**.959**	**.834**	**.770**	**.961**	**.841**
	w/ MERP	13.7M	**.772**	**.951**	**.835**	**.769**	**.961**	**.840**	**.771**	**.962**	**.841**
w/ SF	AttrGNN† [23]	-	.777	.920	.829	.763	.909	.816	.942	.987	.959
	RNM [53]	-	.840	.919	.870	.872	.944	.899	.938	.981	.954
	CLEM† [37]	-	.854	.935	.879	.885	.958	.904	.936	.977	.952
	RPR-RHGT [2]	-	.693	-	.754	.886	-	.912	.889	-	.919
	ERMC† [45]	-	.903	.946	.899	.942	.944	.925	.962	.982	.973
	EVA* [21]	13.8M	.929	.986	.951	.964	.997	.976	.990	.999	.994
	MSNEA* [5]	14.7M	.887	.961	.913	.938	.983	.955	.969	.997	.980
	MCLEA* [20]	13.7M	.926	.983	.946	.961	.994	.973	.987	.999	.992
	MEAformer (Ours)	14.2M	**.948**	**.993**	**.965**	**.977**	**.999**	**.986**	**.991**	**1.00**	**.995**
	w/ MERP	14.2M	**.949**	**.993**	**.965**	**.978**	**.999**	**.986**	**.991**	**1.00**	**.995**
(S) Unsup.	EVA* [21]	13.8M	.883	.967	.913	.930	.985	.951	.968	.995	.978
	MSNEA* [5]	14.7M	.858	.935	.886	.921	.973	.939	.953	.990	.967
	MCLEA* [20]	13.7M	.879	.963	.909	.931	.983	.951	.959	.993	.972
	MEAformer (Ours)	14.2M	**.917**	**.980**	**.941**	**.958**	**.992**	**.972**	**.973**	**.998**	**.982**
(V) Unsup.	EVA* [21]	13.8M	.891	.961	.917	.941	.986	.958	.970	996	.982
	MSNEA* [5]	14.7M	.859	.936	.887	.921	.970	.939	.954	.989	.968
	MCLEA* [20]	13.7M	.860	.950	.893	.914	.975	.938	.953	.990	.967
	MEAformer (Ours)	14.2M	**.909**	**.974**	**.933**	**.950**	**.990**	**.965**	**.972**	**.997**	**.983**

- **Iter. Results on Bilingual Datasets**

	Models	Para.	DBP15K$_{ZH-EN}$			DBP15K$_{JA-EN}$			DBP15K$_{FR-EN}$		
			H@1	H@10	MRR	H@1	H@10	MRR	H@1	H@10	MRR
w/o SF	BootEA [30]	-	.629	.847	.703	.622	.854	.701	.653	.874	.731
	NAEA [52]	-	.650	.867	.720	.641	.873	.718	.673	.894	.752
	EVA* [21]	13.3M	.746	.910	.807	.741	.918	.805	.767	.939	.831
	MSNEA* [5]	14.1M	.643	.865	.719	.572	.832	.660	.584	.841	.671
	MCLEA* [20]	13.2M	.811	.954	.865	.806	.953	.861	.811	.954	.865
	MEAformer (Ours)	13.7M	**.847**	**.970**	**.892**	**.842**	**.974**	**.892**	**.845**	**.976**	**.894**
w/ SF	EVA* [21]	13.8M	.956	.993	.969	.979	.998	.987	.995	.999	.997
	MSNEA* [5]	14.7M	.896	.969	.922	.942	.986	.958	.971	.998	.982
	MCLEA* [20]	13.7M	.964	.996	.977	.986	.999	.992	.995	1.00	.997
	MEAformer (Ours)	14.2M	**.973**	**.998**	**.983**	**.991**	**1.00**	**.995**	**.996**	**1.00**	**.998**
(S) Unsup.	EASY† [14]	-	.898	.979	.930	.943	.990	.960	.980	.998	.990
	EVA* [21]	13.8M	.937	.991	.957	.974	.998	.983	.992	1.00	.996
	MSNEA* [5]	14.7M	.870	.946	.897	.933	.980	.950	.961	.992	.973
	MCLEA* [20]	13.7M	.947	.995	.966	.977	.999	.986	.990	1.00	.994
	MEAformer (Ours)	14.2M	**.962**	**.998**	**.976**	**.987**	**.999**	**.992**	**.993**	**1.00**	**.996**
(V) Unsup.	EVA* [21]	13.8M	.948	.992	.964	.977	.997	.985	.990	1.00	.995
	MSNEA* [5]	14.7M	.871	.943	.898	.927	.976	.945	.959	.993	.979
	MCLEA* [20]	13.7M	.942	.994	.963	.974	.999	.984	.986	1.00	.992
	MEAformer (Ours)	14.2M	**.964**	**.998**	**.975**	**.985**	**1.00**	**.991**	**.992**	**1.00**	**.996**

- **Non-iter. Monolingual Results**

	Models	FB15K-DB15K			FB15K-YAGO15K		
		H@1	H@10	MRR	H@1	H@10	MRR
20%	MMEA	.265	.541	.357	.234	.480	.317
	EVA∗	.199	.448	.283	.153	.361	.224
	MSNEA∗	.114	.296	.175	.103	.249	.153
	MCLEA∗	.295	.582	.393	.254	.484	.332
	MEAformer	**.417**	**.715**	**.518**	**.327**	**.595**	**.417**
	w/ MERP	**.434**	**.728**	**.534**	**.325**	**.598**	**.416**
50%	MMEA	.417	.703	.512	.403	.645	.486
	EVA∗	.334	.589	.422	.311	.534	.388
	MSNEA∗	.288	.590	.388	.320	.589	.413
	MCLEA∗	.555	.784	.637	.501	.705	.574
	MEAformer	**.619**	**.843**	**.698**	**.560**	**.778**	**.639**
	w/ MERP	**.625**	**.847**	**.704**	**.560**	**.780**	**.640**
80%	MMEA	.590	.869	.685	.598	.839	.682
	EVA∗	.484	.696	.563	.491	.692	.565
	MSNEA∗	.518	.779	.613	.531	.778	.620
	MCLEA∗	.735	.890	.790	.667	.824	.722
	MEAformer	**.765**	**.916**	**.820**	**.703**	**.873**	**.766**
	w/ MERP	**.773**	**.918**	**.825**	**.705**	**.874**	**.768**

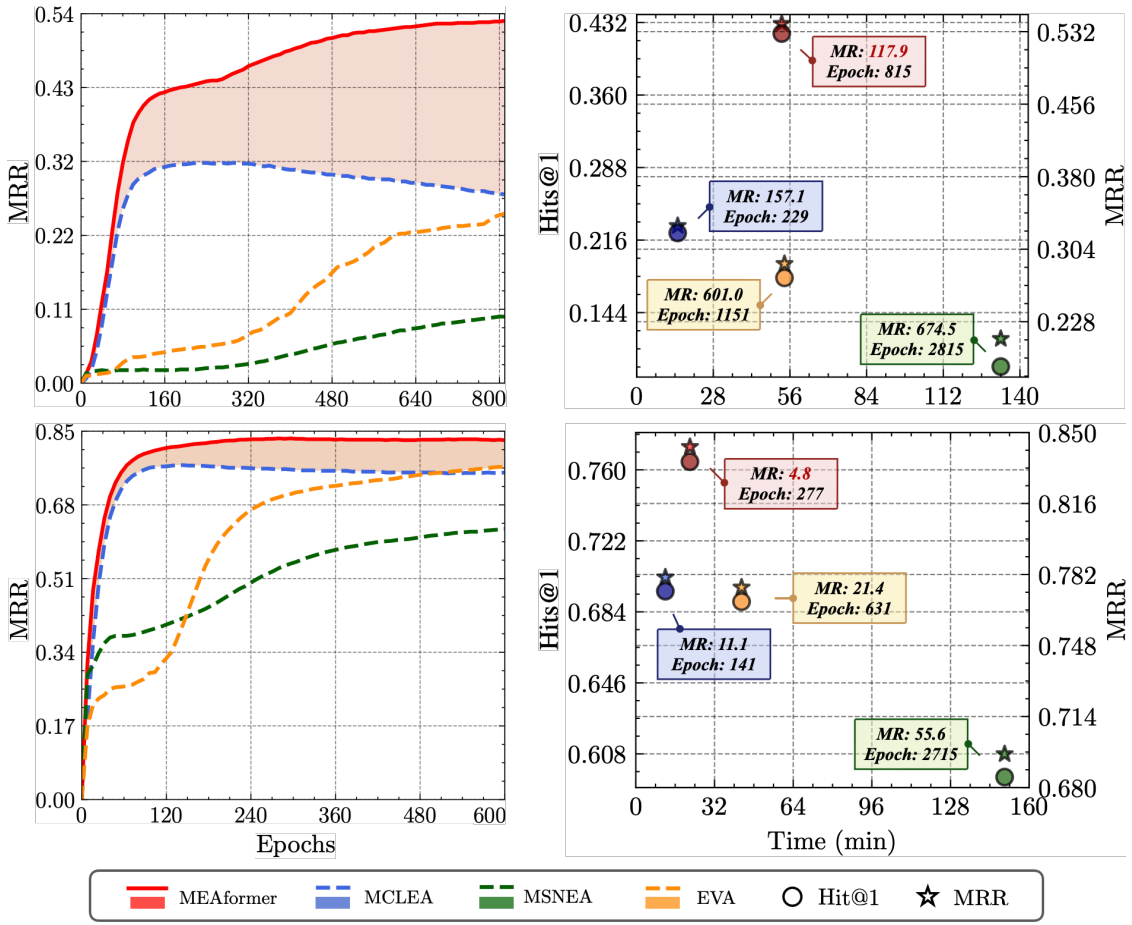
- **Iter. Monolingual Results**

	Models	FB15K-DB15K			FB15K-YAGO15K		
		H@1	H@10	MRR	H@1	H@10	MRR
20%	EVA∗	.231	.488	.318	.188	.403	.260
	MSNEA∗	.149	.392	.232	.138	.346	.21
	MCLEA∗	.395	.656	.487	.322	.546	.400
	MEAformer	**.578**	**.812**	**.661**	**.444**	**.692**	**.529**
50%	EVA∗	.364	.606	.449	.325	.560	.404
	MSNEA∗	.358	.656	.459	.376	.646	.472
	MCLEA∗	.620	.832	.696	.563	.751	.631
	MEAformer	**.690**	**.871**	**.755**	**.612**	**.808**	**.682**
80%	EVA∗	.491	.711	.573	.493	.695	.572
	MSNEA∗	.565	.810	.651	.593	.806	.668
	MCLEA∗	.741	.900	.802	.681	.837	.737
	MEAformer	**.784**	**.921**	**.834**	**.724**	**.880**	**.783**

➢ *Experiment results without (w/o) and with (w/) surface forms (SF), where † denotes that the PLMs were applied for surface/attribute embedding generation. "Para." refers to the number of learnable parameters. "Unsup." refers to the unsupervised learning depending on entities' surface (S) or visual (V) information. The symbol ∗ denotes our reproduced results on the same experiment setting.*

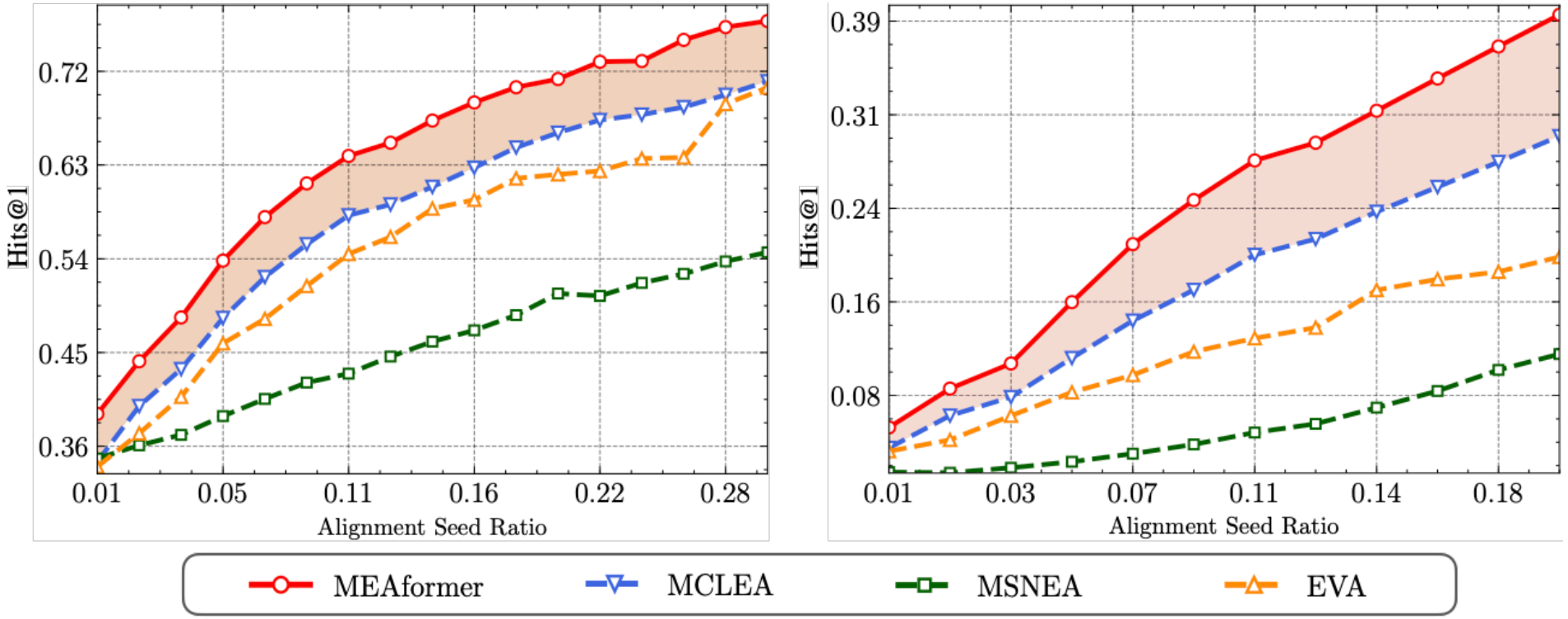
Efficiency Analysis

◎ **Performance *vs.* training epoch (left) & training time (right) with R_{sa} = 0.2. FBDB15K (up) & DBP15K$_{ZH-EN}$ (down)**



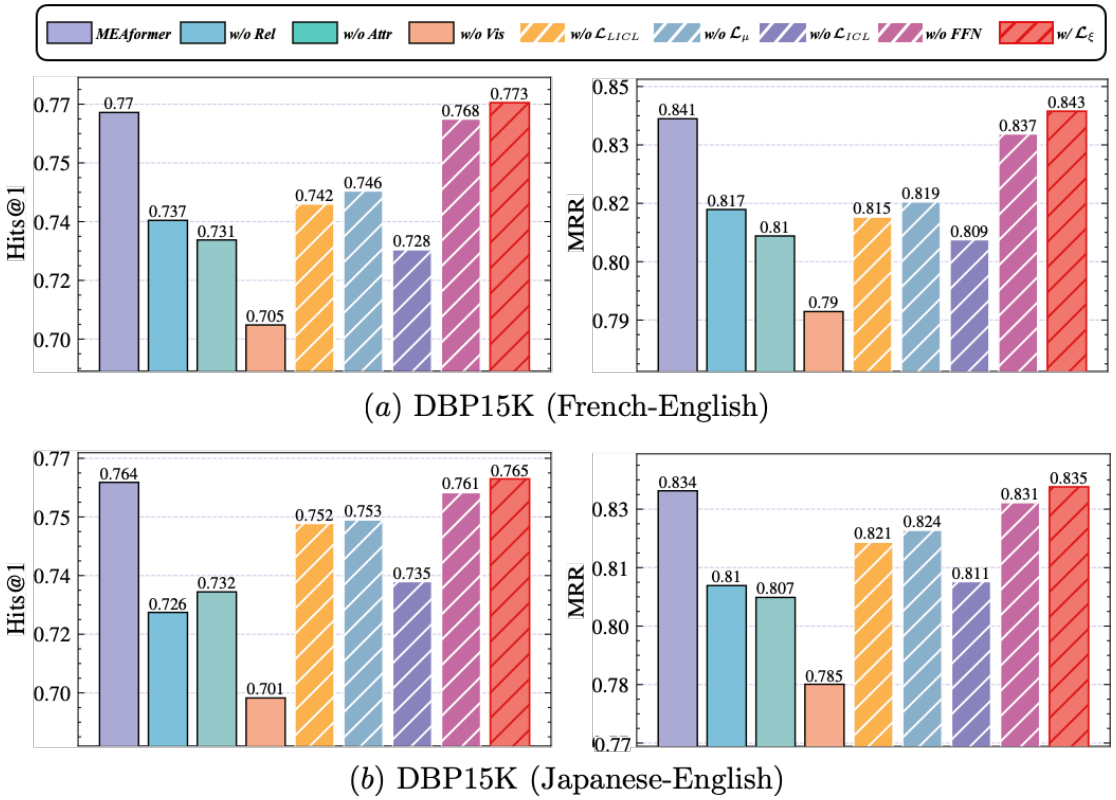
Low Resource (Fewer Seed Alignments)

◎ Models' Hits@1 performance with fewer seed alignments on DPB15K *FR-EN* (left) and FBDB15K (right).

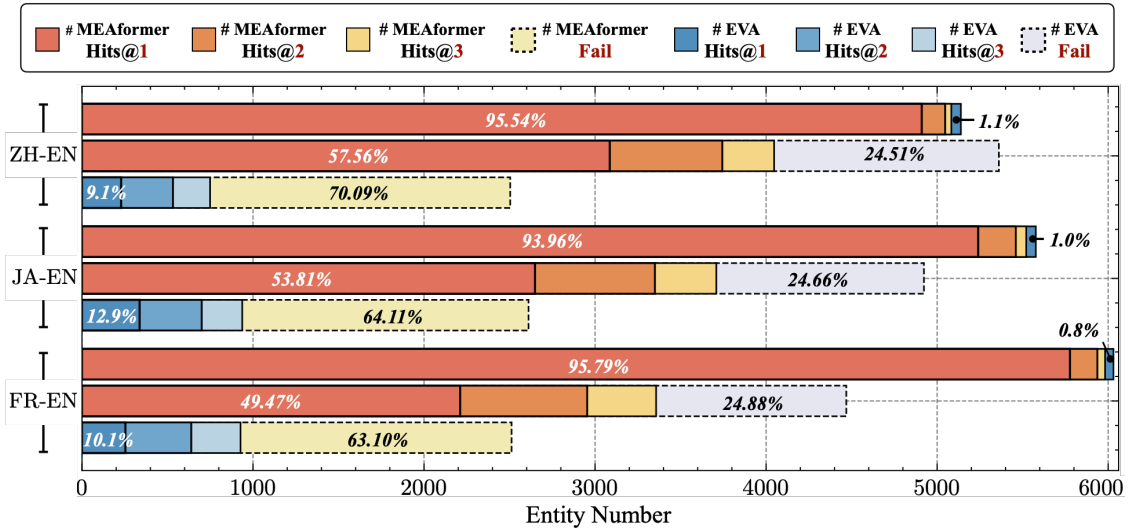


Component & Error Analysis

◎ **evaluate stripped-down versions of MEAformer**



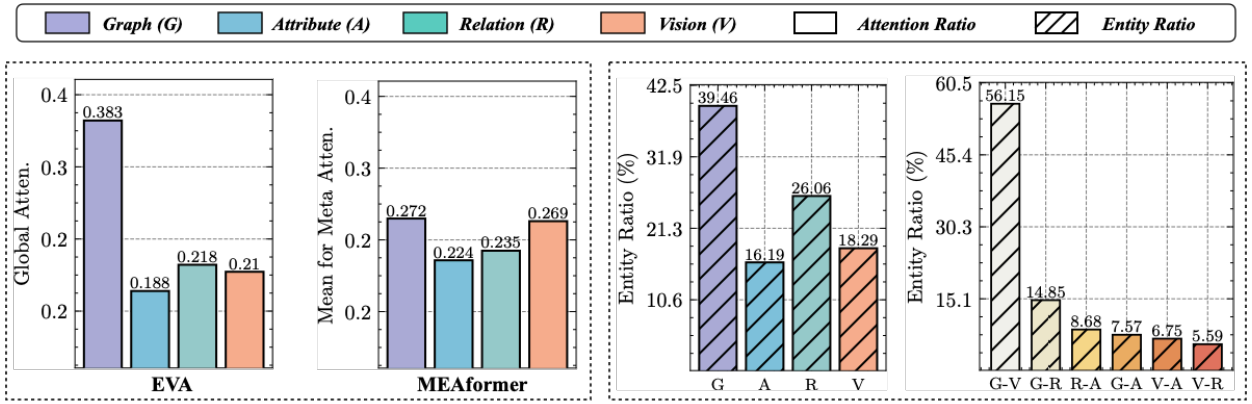
◎ **"# Hits@N" refers to the number of entities e_{uc} ranked in model's top N list, and "#Fail" refers to the number of e_{uc} that model failed in.**



➢ *MEAformer demonstrates a higher degree of stability than EVA.*

Case Analysis & Modality Distribution

◎ Although our average weights are evenly distributed, a distinct modality preference distribution for different entities still exists, resembling EVA's modality weight.



◎ Part of the representative examples whose meta weights of corresponding modalities rank relatively high

Modality	Examples
Vision	Rivaldo, Reinaldo_da_Cruz_Oliveira, Gilberto_Silva, Didier_Zokora, Kate_Bush, Billie_Jean, Take_My_Breath_Away, Léon_Blum, Celine_Dion, Ophiuchus, Lacerta, Costinha
Attribute	Kingdom_of_England, Roman Republic, Timurid_Empire, 2014_AFC_Cup, 2015–16_Football_League_Cup, Kuwait_City, Calabria, Hong_Kong, Amsterdam, Hamburg, Kiev
Relation	United_States, France, Michael_Jackson, The_Beatles, Sony_Mobile, Microsoft, Apple_Inc, Queen_Victoria, Akihito, European_People's_Party,English_language, Aramaic_alphabet
Graph	Vágner_Love, David_Villa, Jimmy_Bullard, Torsten_Frings, Hydra_(constellation), Ursa_Major, Montricher, Delphinus, Scorpius, Volans, Locarno, Crissier, Schmiedrued, Tramelan

- Football players (e.g., *David_Villa*) – DBP15K$_{ZH\text{-}EN}$.
- Constellations (e.g., *Ursa_Major*) – DBP15K$_{JA\text{-}EN}$.
- Switzerland's areas (e.g., *Crissier*) – DBP15K$_{FR\text{-}EN}$.

ACM Multimedia 2023
Ottawa, Canada

Thank you!

https://github.com/zjukg/MEAformer

MEAformer: Multi-modal Entity Alignment Transformer for Meta Modality Hybrid

ZHEJIANG UNIVERSITY · The University of Manchester · UNIVERSITY OF OXFORD · THE UNIVERSITY OF EDINBURGH · HUAWEI