HyConvE - A Novel Embedding Model for Knowledge Hypergraph Link Prediction with Convolutional Neural Networks

Chenxu Wang, Xin Wang, Zhao Li, Zirui Chen Jianxin Li

Published in WWW, 2023

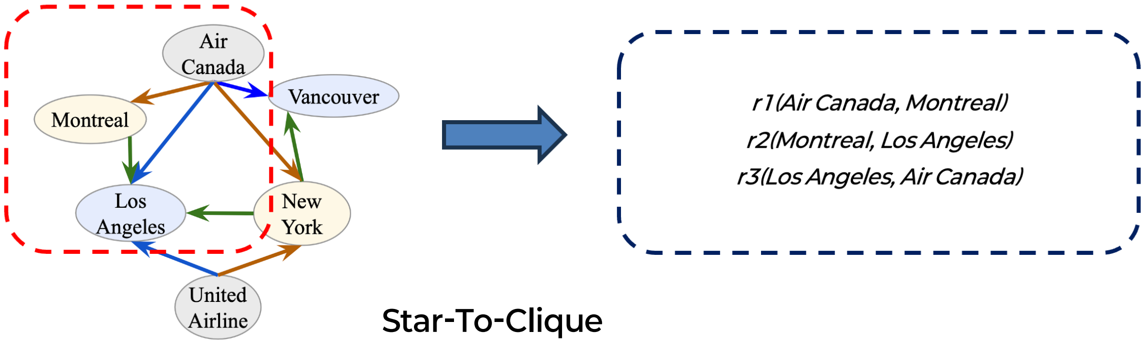
2025-03-13
HoonUi Lee

Contents

Previous Work

❖ Limitation in modeling restrictions

- ❑ Some approach that does **not decompose** the original tuple
  - ▸ **HSimple, HypE** in last presentation, etc..
  - ▸ Better preserved the **positional information** and the **intrinsic pattern**



Star-To-Clique
(Tuple to Triple)

Previous Work

- **Limitation in modeling restrictions**
  - However, **shallow model** is not sufficient to fully learn the **latent and implicit knowledge** inherent in n-ary data
    - mapping representation of entities to a low-dimensional space during the learning process
    - not sufficient for learning complex interactions in multi-ary relations

➔ **To learn rich and complex expressions, Convolutional Neural Network is introduced**

Previous Work

- CNNs

❖ Some approaches using CNNs

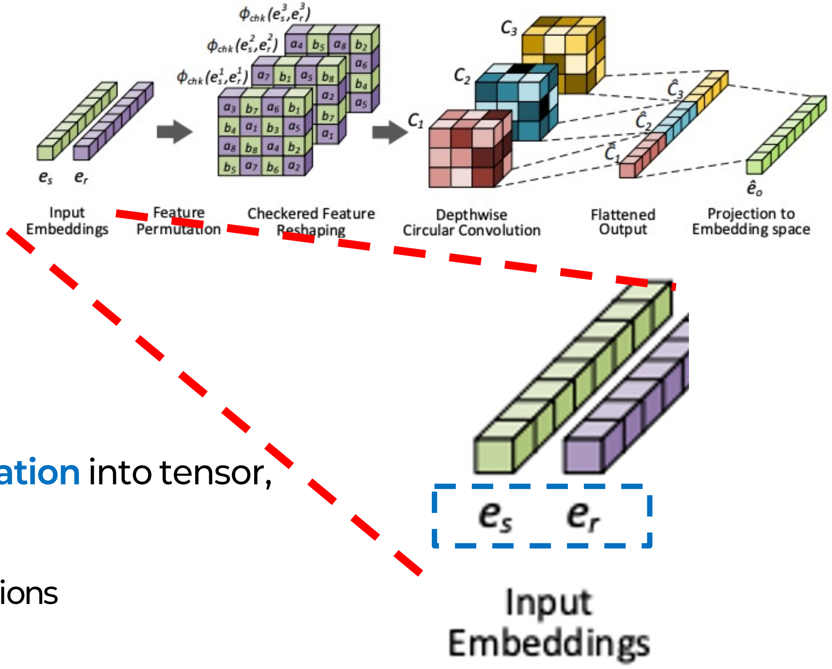
- ❑ Starting with **ConvE**, subsequent studies have attempted to characterize the interactions between entities and relations well.
  - HypER, ConvR, AcrE, InteractE

- ❑ BUT, There is tendency to **use complex convolutional layers**
  - leading to a **denser** model structure and **increased complexity**

Previous Work

- CNNs

InteractE



❖ Loss of translation property

❑ $h + r \approx t$

- intuitive relationship representation between entities

❑ Existing approach transforms **head entity and relation** into tensor, applies filters to drive the final vector

- linear flow is interpreted through nonlinear computations
- **weakens** the ability to capture **surface knowledge**

➔ Using 3D Convolution

HyConvE

- framework

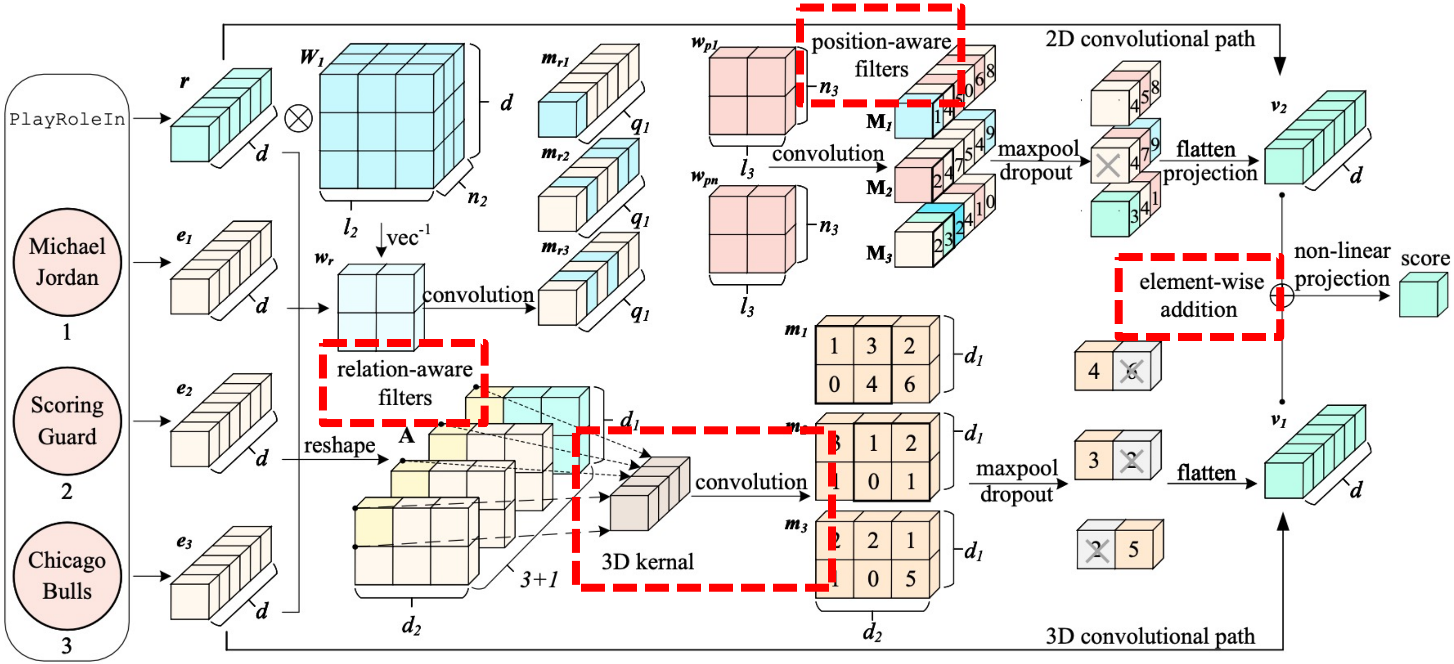


Figure 2: The framework of the HyConvE model.

HyConvE

- framework

❖ Two feed-forward paths

❑ Expanding ConvE's 2D convolution to **3D convolution**

- learning by incorporating information from the **depth** dimension
- in KHG, this represents the interaction between relations and entities

❑ Two 2D convolution filters

- **relation-aware, position-aware**
- To effectively capture the positional information of each relation and the patterns between entities

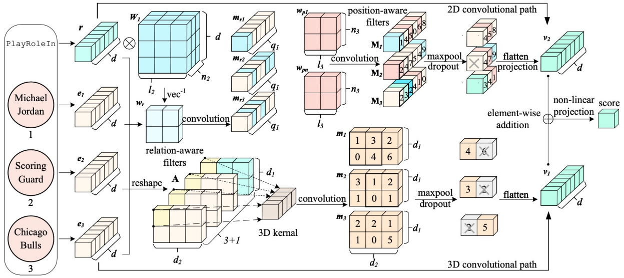


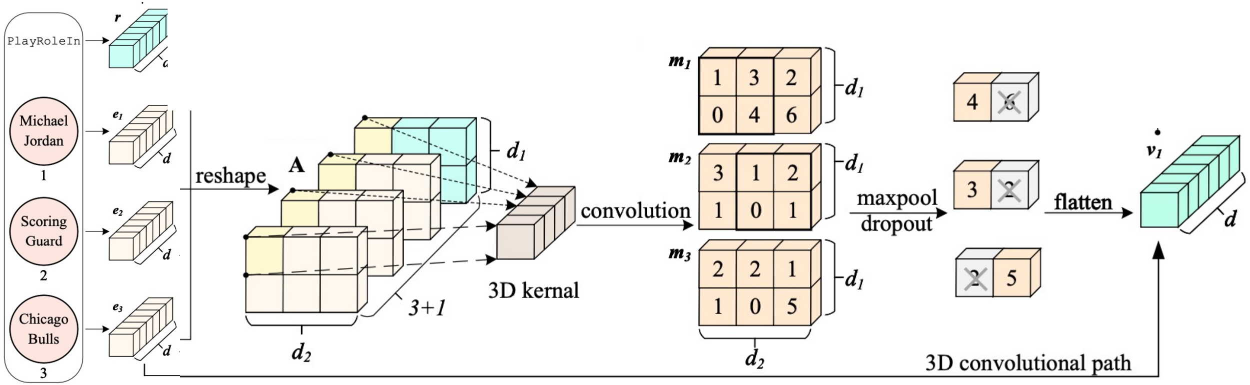
Figure 2: The framework of the HyConvE model.

HyConvE

- framework

❖ Latent and Surface Knowledge Extraction

- ❑ Reshaping and stacking entity and relation vectors to form a 3D structure
- ❑ Performing 3D convolution operations on the constructed cube with n1 filters
- ❑ Get output feature v1 from $m_1, m_2, \ldots, m_k \in \mathbb{R}^{d_1 \times d_2}$ with some operation



HyConvE

- framework

❖ Intrinsic Semantic Information Capture

❑ An **entity's semantic information** is associated with the **relation**

❑ Also depends on its **position** within the n-ary relation

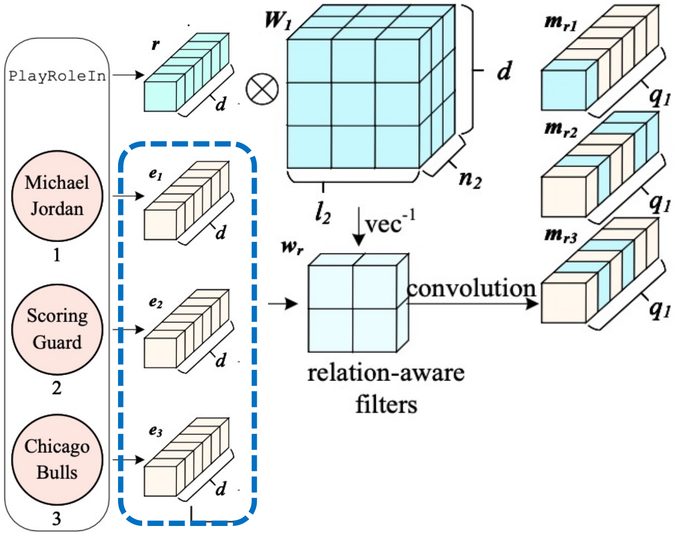
❑ Introducing 2D relation-aware filter and position-aware filter

HyConvE

- framework

❖ Relation-aware filters

- ❑ Transform it into 2D filter matrix through a vectorization inverse operation
- ❑ Relation filters are convolved with the associated entities

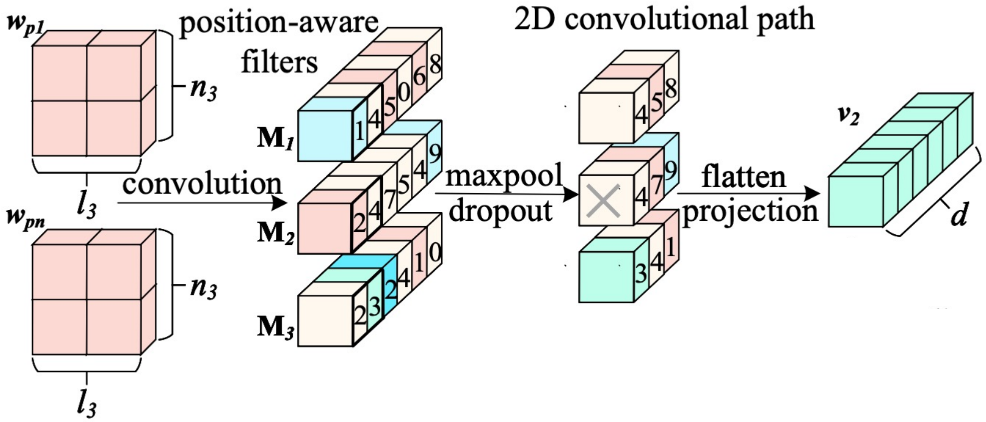


HyConvE

- framework

❖ Position-aware filters

- ❑ After through the relation filter,
  entity passes through the position filter corresponding to its position
- ❑ Get output feature v2 with some operation



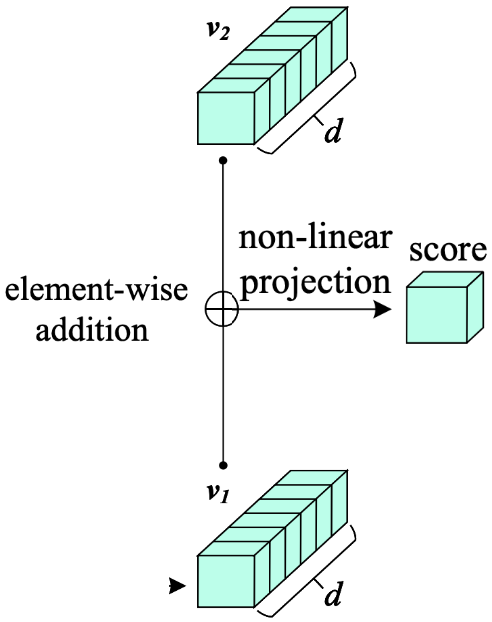
HyConvE

- framework

❖ Adding the two pieces of information

- ❑ The vector obtained through 3D convolution and through 2D convolution are combined using element-wise addition

- ❑ $score = g(v_1 + v_2)W_3$
  - introducing activation function
  - randomly discarding some neurons
  - get final score by projection layer W3



HyConvE

- VS. Previous Work

- **Limitations in learning deep and complex representations with shallow model**
  - Using Convolution Neural Networks
  - **Relation-aware filter, Position-aware filter**

- **Limitations in capturing the translation property**
  - Using **3D Convolution**
  - Providing both relations and entities tensors
    - Learning the interactions between entities and relations within a tuple more effectively

HyConvE

- Training

❖ Negative Sampling

❑ $$\bigcup_{i=1}^{k} \mathcal{N}_x^{(i)} \equiv \bigcup_{i=1}^{k} \{e_1, \cdots, \bar{e}_i, \cdots, e_k \notin \mathcal{F} \,|\, \bar{e}_i \in \mathcal{E}, \bar{e}_i \neq e_i\}$$

- a set of negative samples of size N|r|
- by replacing each of the entities with N random entities in the tuple
- $\mathcal{N}_x^{(i)}$ replaces the entity in the i-th position

HyConvE

- Training

$$l_{(r,e_1,\ldots,e_k)} = \begin{cases} 1 & \text{for } r(e_1,\ldots,e_k) \in \mathcal{H} \\ -1 & \text{for } r(e_1,\ldots,e_k) \in \mathcal{H}' \end{cases}$$

❖ Loss Function

❑ $$\mathcal{L} = \sum_{r(e_1,\ldots,e_k)\in\{\mathcal{H}\cup\mathcal{H}'\}} \log(1+\exp(l_{r(e_1,\ldots,e_k)} \cdot f(r(e_1,\ldots,e_k)))) + \boxed{\phi}$$

- based on logistic loss function with L2 regularization

❑ $$\phi = \lambda(\|\boldsymbol{c}\|_2^2 + \|\boldsymbol{w}\|_2^2 + \|\boldsymbol{p}\|_2^2 + \sum_{i=1}^{k} \|\boldsymbol{e}_i\|_2^2 + \|\boldsymbol{r}\|_2^2$$

- for convolution layers, max-pool layers, fully connected layers in two path and entity and relation embeddings

Experiment

- Dataset

Table 1: Dataset Statistics. The size of the train, valid, and test columns represent the number of triples or tuples, respectively. "Arity" denotes the involved arities of relations.

Dataset	$\lvert\mathcal{E}\rvert$	$\lvert\mathcal{R}\rvert$	Arity	# train	# valid	# test	# arity=2	# arity=3	# arity=4	# arity $\geq$ 5
FB15k-237	14,541	237	2	272,115	17,535	20,466	310,116	–	–	–
WN18RR	40,943	11	2	86,835	3,034	3,134	93,003	–	–	–
JF17K	29,177	327	2-6	61,104	15,275	24,568	56,332	34,550	9,509	2,267
WikiPeople	47,765	707	2-9	305,725	38,223	38,281	337,914	25,820	15,188	3,307
FB-AUTO	3,388	8	2, 4, 5	6,778	2,255	2,180	3,786	–	125	7,212
JF17K-3	11,541	104	3	27,645	3,454	3,455	–	34,544	–	–
JF17K-4	6,536	23	4	7,607	951	951	–	–	9509	–
WikiPeople-3	12,270	66	3	20,656	2,582	2,582	–	25,820	–	–
WikiPeople-4	9,528	50	4	12,150	1,519	1,519	–	–	15188	–

❑ **WN18RR, FB15K-237** : with binary relation

❑ **JF17K, WikiPeople, FB-AUTO** : with non-binary relation

Experiment

- Knowledge Hypergraph Completion

Table 2: Results of Link Prediction on Knowledge Hypergraph Datasets. The best results are in boldface and the second best are underlined. Experimental results with "-" are those results that were not presented in the original paper. All experimental results are obtained locally.

Model	JF17K				WikiPeople				FB-AUTO			
	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10
RAE [41]	0.392	0.312	0.433	0.561	0.253	0.118	0.343	0.463	0.703	0.614	0.764	0.854
NaLP [15]	0.310	0.239	0.334	0.450	0.338	0.272	0.362	0.466	0.672	0.611	0.712	0.774
HINGE[28]	0.473	0.397	0.490	0.618	0.333	0.259	0.361	0.477	0.678	0.630	0.706	0.765
NeuInfer [14]	0.451	0.373	0.484	0.604	0.351	0.274	0.381	0.467	0.737	0.700	0.755	0.805
HypE [10]	0.494	0.399	0.532	0.650	0.263	0.127	0.355	0.486	0.804	0.774	0.824	0.856
tNaLP+ [13]	0.449	0.370	0.484	0.598	0.339	0.269	0.369	0.473	0.729	0.645	0.748	0.826
S2S [8]	0.528	0.457	0.570	0.690	**0.364**	0.273	0.402	**0.503**	-	-	-	-
RAM [22]	0.539	0.463	0.572	0.690	0.363	0.271	**0.405**	0.500	0.830	0.803	0.851	0.876
HyConvE (ours)	**0.590**	**0.478**	**0.610**	**0.729**	0.362	**0.275**	0.388	0.501	**0.847**	**0.820**	**0.872**	**0.901**

Complex Network Design

Using Decomposition

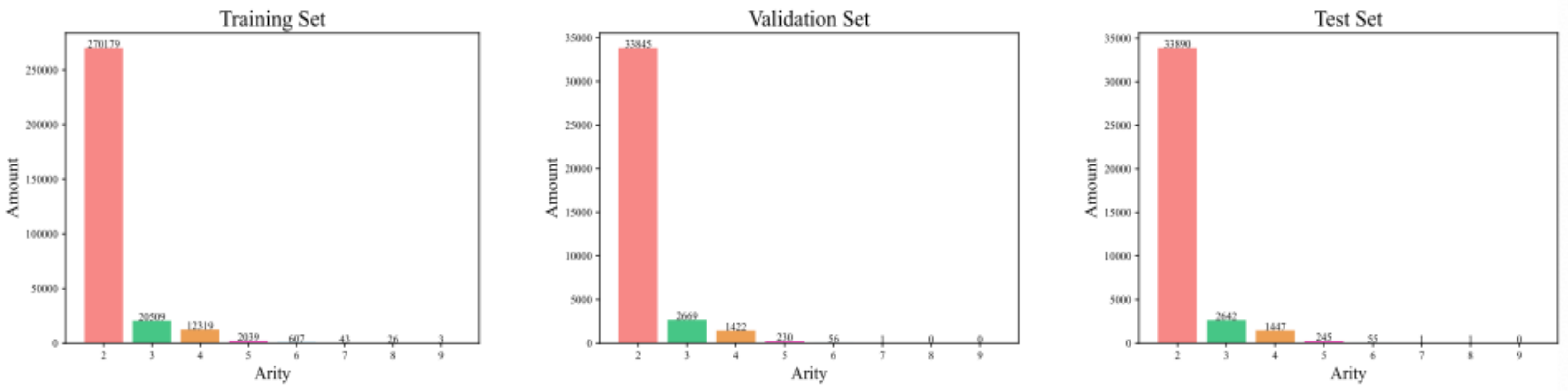
Only with Position Not with Relation

❑ Low performance in WikiPeople because of high-arity relation

- sparsity problem occur

Experiment

- Knowledge Hypergraph Completion



(a) Training Set (b) Validation Set (c) Test Set

Figure 7: Distribution of relational data of different arities in the WikiPeople dataset.

Experiment

- Result on fixed arity set

Table 3: Results on fixed arity datasets. The best results are in boldface and the second best are underlined.

Model	JF17K-3			JF17K-4			WikiPeople-3			WikiPeople-4		
	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10
RAE [41]	0.505	0.430	0.644	0.707	0.636	0.835	0.239	0.168	0.379	0.150	0.080	0.273
NaLP [15]	0.515	0.431	0.679	0.719	0.673	0.805	0.301	0.226	0.445	0.342	0.237	0.540
n-CP [21]	0.669	0.613	0.801	0.754	0.701	0.855	0.313	0.237	0.476	0.253	0.163	0.432
n-tucker [21]	0.727	0.664	0.852	0.786	0.723	0.851	0.315	0.236	0.478	0.335	0.225	0.536
GETD [21]	<u>0.725</u>	<u>0.660</u>	<u>0.858</u>	<u>0.822</u>	<u>0.761</u>	<u>0.924</u>	**0.363**	**0.272**	**0.545**	<u>0.346</u>	<u>0.229</u>	<u>0.542</u>
RAM [22]	0.578	0.505	0.722	0.743	0.701	0.845	0.254	0.190	0.383	0.226	0.161	0.367
HyConvE (ours)	**0.729**	**0.670**	**0.861**	**0.827**	**0.770**	**0.931**	<u>0.318</u>	<u>0.240</u>	<u>0.482</u>	**0.386**	**0.271**	**0.607**

- ❑ Low performance in WikiPeople-3, because of noise introduced by other relational data
  - Since the data distribution among relations in the dataset is not uniform

Experiment

- Knowledge Graph Completion

Table 4: Results of Link Prediction on Knowledge Graph Datasets. The best results are in boldface and the second best are underlined. Experimental results with "-" are those results that were not presented in the original paper. All experimental results are obtained locally.

Model	FB15k-237			WN18RR			JF17K			WikiPeople			FB-AUTO		
	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10	MRR	Hit@1	Hit@10
TransE [5]	0.294	-	0.561	0.226	-	0.501	0.276	0.167	0.495	0.312	0.146	0.574	0.313	0.132	0.562
DistMult[28]	0.241	0.155	0.419	0.431	0.390	0.490	0.228	0.144	0.411	0.275	0.193	0.388	0.494	0.444	0.566
ComplEx [14]	0.253	0.158	0.428	0.440	0.411	0.512	0.308	0.219	0.498	0.326	0.232	0.461	0.487	0.442	0.568
HypE [10]	0.240	0.160	0.400	0.363	0.332	0.473	-	-	-	-	-	-	-	-	-
S2S [8]	**0.348**	**0.256**	**0.540**	**0.498**	**0.455**	**0.577**	-	-	-	-	-	-	-	-	-
RAM [22]	-	-	-	-	-	-	0.324	0.232	0.515	**0.408**	**0.313**	**0.568**	0.489	0.444	**0.576**
HyConvE (ours)	0.339	0.212	0.458	0.461	0.432	0.534	**0.338**	**0.246**	**0.525**	0.388	0.281	0.556	**0.493**	**0.445**	0.572

❑ Binary relational facts are analyzed to lack position and semantic information

Experiment

- Ablation Study on Path

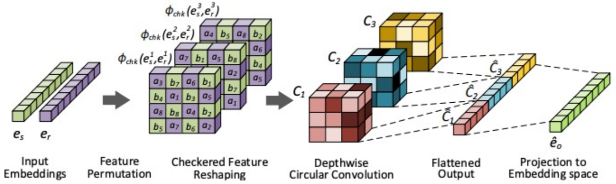
Table 5: Results of ablation study. The best results are in boldface. HyConvE-path1-only means to use only the 3D path of HyConvE when conducting experiments and HyConvE-path2-only means the other.

Model	JF17K				WikiPeople				FB-AUTO			
	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10	MRR	Hit@1	Hit@3	Hit@10
HyConvE-path1-only	0.528	0.457	0.570	0.690	0.323	0.227	0.344	0.478	0.831	0.796	0.851	0.899
HyConvE-path2-only	0.102	0.054	0.094	0.168	0.072	0.048	0.094	0.172	0.145	0.082	0.164	0.212
HyConvE (ours)	**0.590**	**0.478**	**0.610**	**0.729**	**0.352**	**0.275**	**0.388**	**0.501**	**0.847**	**0.820**	**0.872**	**0.901**

❑ Compare with ”Only 2D convolution” and “Only 3D convolution”

Conclusion

❖ Previous Work



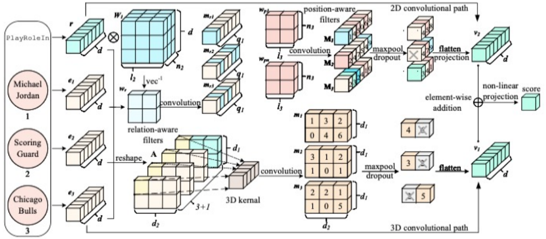
- ❑ Shallow models is unable to learn latent and implicit knowledge
- ❑ Some Attempts exist to learn such representations using CNN
  - while they lose translation property, weaken in capturing surface knowledge

❖ HyConvE

- ❑ Using 3D convolution to capture deep interactions of entities and relations without compromising translation property
- ❑ Using relation-aware, position-aware filters to capture intrinsic patterns and position information

❖ Experiment

- ❑ Performance evaluation in KHG completion and KG completion
- ❑ Ablation study about two feed-forward paths



HyCubE: Efficient Knowledge Hypergraph 3D Circular Convolutional Embedding

Zhao Li, Xin Wang, Jun Zhao, Wenbin Guo, Jianxin Li

Published in TKDE, 2025

2025-03-13
HoonUi Lee

Contents

- **Previous Work**
- **HyCubE**
  - Framework
  - Alternative Mask Stack
  - 3D Convolution
  - Model Training
- **Experiment**
  - Knowledge Hypergraph Completion
  - Model Efficiency
  - Ablation Study
- **Conclusion**

Previous Work

❖ Effectiveness vs Efficiency trade-off

- ❑ KHG emb. methods have focused on **"Effectiveness"**
  - significantly higher computational costs and memory usage
  - handle more complex inherent semantic information

- ❑ Argued that the **efficiency** of the model must be considered

Model	Parameters (Millions)			GPU Memory Usage (MB)			Time Usage (s)		
	JF17K	WikiPeople	FB-AUTO	JF17K	WikiPeople	FB-AUTO	JF17K	WikiPeople	FB-AUTO
RAM [5]	≈ 14.24	≈ 27.34	≈ 1.63	15,174	20,742	2,860	74.2	215.8	4.0
PosKHG [6]	≈ 14.34	≈ 27.53	≈ 1.65	15,388	21,147	2,874	87.4	229.7	4.3
HyConvE [7]	≈ 12.80	≈ 21.44	≈ 4.80	7,718	15,430	3,032	98.7	247.3	4.9
ReAlE [3]	≈ 14.88	≈ 29.61	≈ 1.64	16,316	17,970	14,488	333.7	507.9	13.9

HyCubE

- Effectiveness / Efficiency

Effectiveness

Alternate Mask Stack

3D Convolution

Circular Padding

Efficiency

3D MaxPooling

1-N Multilinear Scoring

HyCubE

- Effectiveness / Efficiency

Effectiveness

Alternate Mask Stack

3D Convolution

Circular Padding

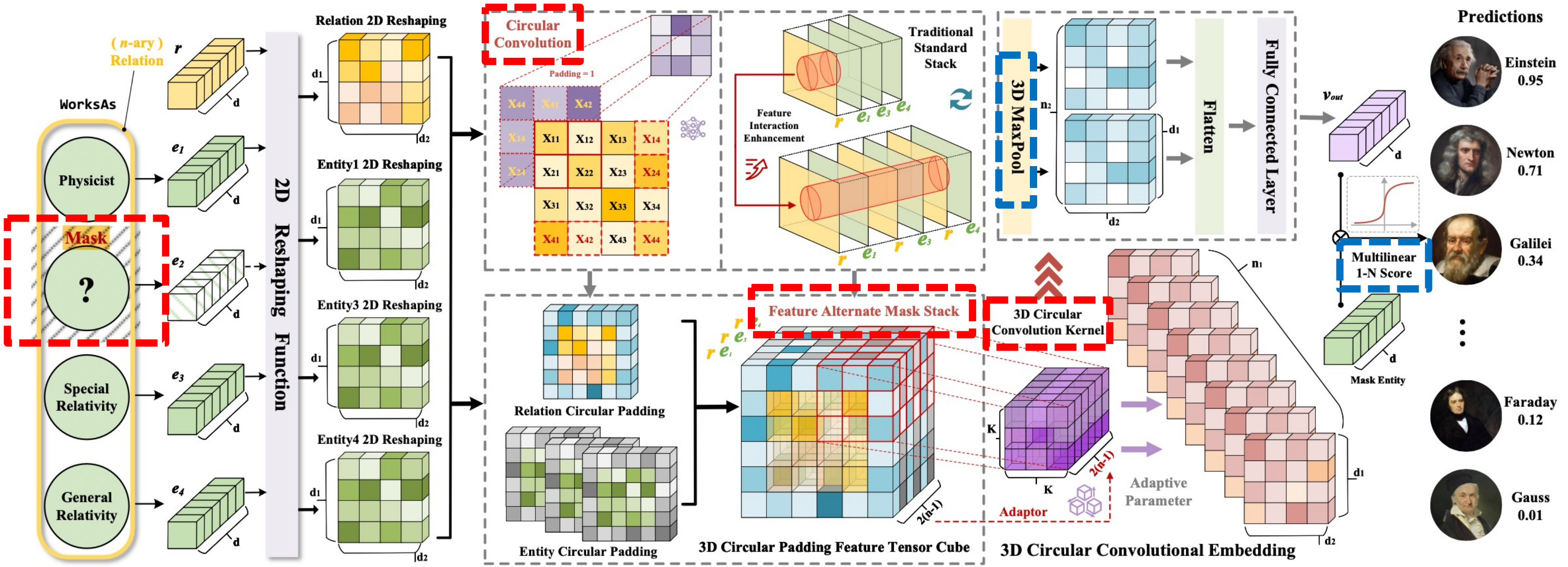
Efficiency

3D MaxPooling

1-N Multilinear Scoring

HyCubE

- framework



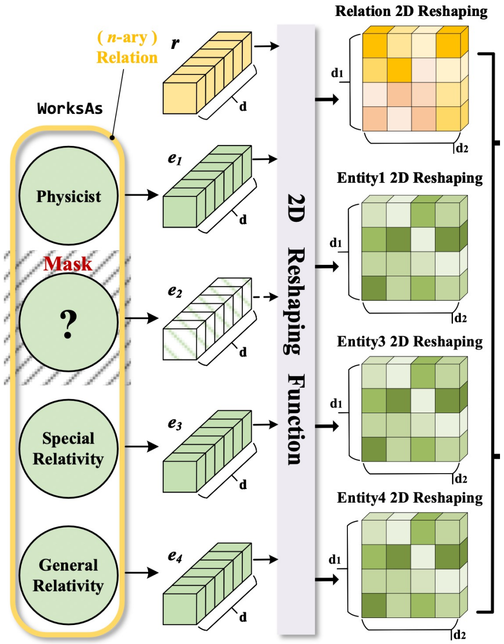
HyCubE

- Alternative Mask Stack

❖ Masked entity prediction

- ❑ Masks the predicted entity (missing entity) e_m
  - ‣ Generating new inputs by masking one entity at a time within the tuple
  - ‣ Inspired by the mask mechanism of the large language model

- ❑ Enhance the interaction and extraction of feature information
  - ‣ just as LLMs can better understand the structure and context of language



HyCubE

- Alternative Mask Stack

❖ Strategies to improve feature interaction

❑ Alternately stacking relations and entities

- for feature interaction enhancement
- $\mathscr{C}_n = [(\overline{\boldsymbol{r}}, \overline{\boldsymbol{e}}_1)||\cdots||(\overline{\boldsymbol{r}}, \overline{\boldsymbol{e}}_{m-1})||(\overline{\boldsymbol{r}}, \overline{\boldsymbol{e}}_{m+1})||\cdots||(\overline{\boldsymbol{r}}, \overline{\boldsymbol{e}}_n)]$

❑ Use 3D Convolution with **Circular Padding**

- improve the interaction area between relations and entities
- enhancing the effectiveness of the KHG embedding

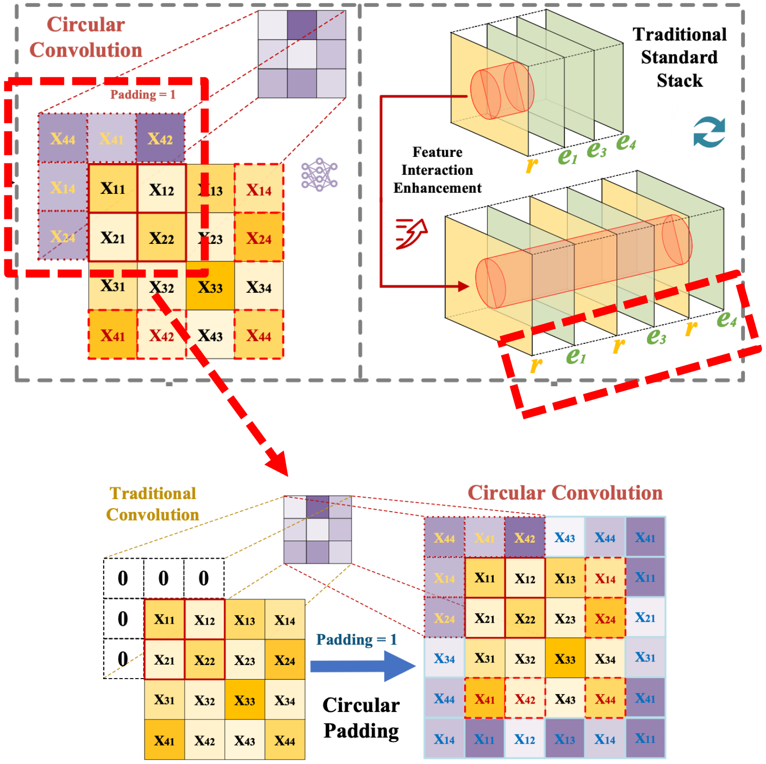


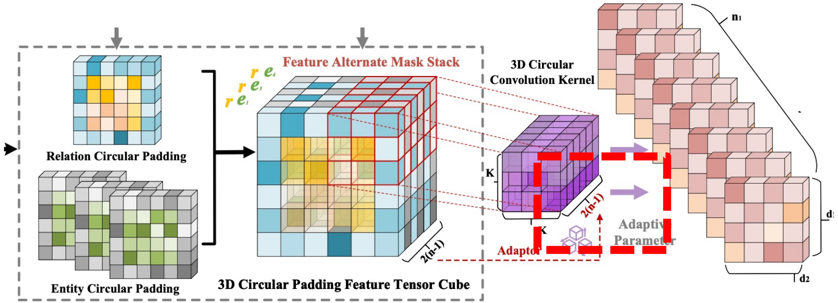
Fig. 4. A front view of the 3D circular padding.

HyCubE

- 3D Convolution

❖ Adaptive kernel about tuple arity

- ❑ Depth dimension of the 3D CNN is adaptively matched to **depth of Cube**
- ❑ $\mathcal{F}_i = \mathcal{C}_n \circledast w_c\,(k_h = k, k_w = k, k_d = \boxed{\theta_{adp}})$ parameter adjusts adaptively
  - actually, filter exists at each depth
- ❑ Embed different n-ary tuple simultaneously without redundant operations
  - tuple decomposition and summation

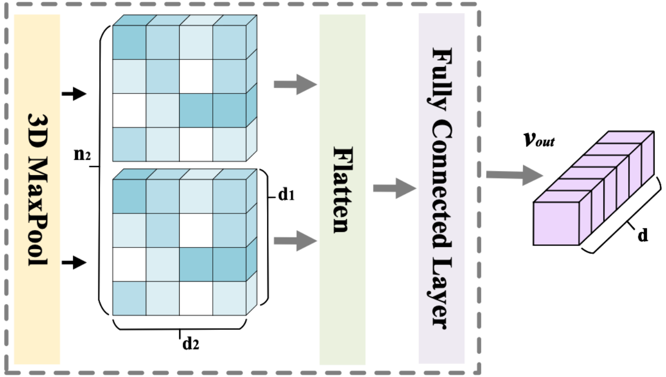


HyCubE

- 3D Convolution

❖ 3D Maxpooling

- ❑ Extract salient features, reduce model parameters, and mitigate overfitting
  - ensure model training efficiency
- ❑ $\mathcal{F}_{i}^{MP} = \mathbf{MaxPool3D}\left(\mathcal{F}_{i}\right)$
  - $\mathcal{F}_{i}^{MP} \in \mathbb{R}^{d_1 \times d_2}$, $i = 1, 2, \ldots, n_2$, and $n_2 = n_1/4$



HyCubE

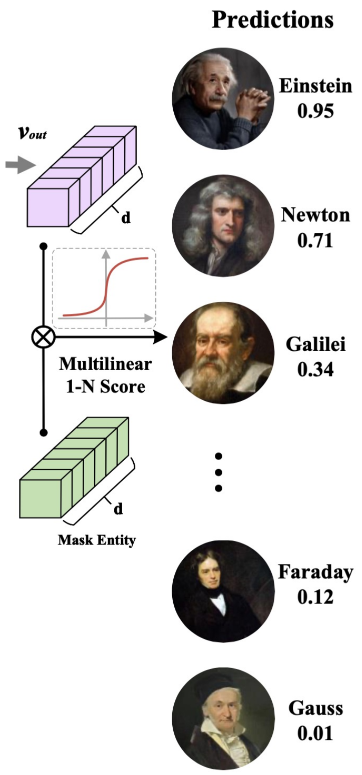
- Model Training

❖ 1-N multilinear scoring way

❑ $\phi(x) = \textbf{softmax}\left(\boldsymbol{v}_{\boldsymbol{out}} \cdot \boldsymbol{e}_{\boldsymbol{m}}^{\mathbf{T}} + \boldsymbol{b}\right)$

❑ Treat each entity in KHG as a candidate prediction entity set

❑ Use the output feature vector v_{out} with each entity embedding



❖ Loss Function

❑ $$\mathcal{L} = \sum_{x \in \mathcal{T}} \sum_{i=1}^{n} -\log\left[e^{\phi(x)} / \left(e^{\phi(x)} + \sum\nolimits_{y \in \mathcal{N}_x^{(i)}} e^{\phi(y)}\right)\right]$$

❑ Softmax-based multi-class cross-entropy

Experiment

- Dataset

TABLE I
DATASET STATISTICS

Dataset		$\lvert\mathcal{E}\rvert$	$\lvert\mathcal{R}\rvert$	Arity	#Train	#Valid	#Test	#Arity=2	#Arity=3	#Arity=4	#Arity$\geq$ 5
Mixed Arity	JF17K	28,645	322	2-6	61,104	15,275	24,568	54,627	34,544	9,509	2,267
	WikiPeople	47,765	707	2-9	305,725	38,223	38,281	337,914	25,820	15,188	3,307
	FB-AUTO	3,388	8	2, 4, 5	6,778	2,255	2,180	3,786	-	215	7,212
Fixed Arity	JF17K-3	11,541	104	3	18,910	4,904	10,730	-	34,544	-	-
	JF17K-4	6,536	23	4	5,641	1,296	2,572	-	-	9,509	-
	WikiPeople-3	12,270	205	3	20,509	2,669	2,642	-	25,820	-	-
	WikiPeople-4	9,528	177	4	12,319	1,422	1,447	-	-	15,188	-

❑ Classified into mixed and fixed

Experiment

- KHG Completion : Mixed Arity

TABLE II
RESULTS OF LINK PREDICTION ON MIXED ARITY KNOWLEDGE HYPERGRAPH DATASETS

Model	JF17K				WikiPeople				FB-AUTO			
	MRR	Hits@1	Hits@3	Hits@10	MRR	Hits@1	Hits@3	Hits@10	MRR	Hits@1	Hits@3	Hits@10
RAE [19]	0.392	0.312	0.433	0.561	0.253	0.118	0.343	0.463	0.703	0.614	0.764	0.854
NaLP [20]	0.310	0.239	0.334	0.450	0.338	0.272	0.362	0.466	0.672	0.611	0.712	0.774
HypE [2]	0.494	0.399	0.532	0.650	0.263	0.127	0.355	0.486	0.804	0.774	0.824	0.856
RAM [5]	0.539	0.463	0.573	0.690	(0.363)	0.271	0.405	0.500	0.830	0.803	0.851	0.876
HyperMLN [22]	0.556	(0.482)	0.597	0.717	$0.351^{\dagger}$	$0.270^{\dagger}$	$0.394^{\dagger}$	$0.497^{\dagger}$	0.831	0.803	0.851	0.877
tNaLP+ [23]	0.449	0.370	0.484	0.598	0.339	0.269	0.369	0.473	0.729	0.645	0.748	0.826
PosKHG [6]	0.545	0.469	0.582	0.706	$0.315^{\dagger}$	$0.214^{\dagger}$	$0.377^{\dagger}$	$0.475^{\dagger}$	0.856	0.821	0.876	0.895
ReAlE [3]	0.530	0.454	0.563	0.677	$0.332^{\dagger}$	$0.207^{\dagger}$	$(0.417^{\dagger})$	$(0.514^{\dagger})$	(0.861)	(0.836)	0.877	(0.908)
RD-MPNN [11]	0.512	0.445	0.573	0.685	-	-	-	-	0.810	0.714	(0.880)	0.888
HyConvE [7]	(0.580)	0.478	(0.610)	(0.729)	0.362	(0.275)	0.388	(0.501)	0.847	0.820	0.872	0.901
HyCubE (Ours)	**0.584**	<u>0.508</u>	**0.616**	**0.730**	**0.448**	**0.368**	**0.490**	**0.592**	<u>0.881</u>	<u>0.860</u>	<u>0.894</u>	<u>0.918</u>
HyCubE+ (Ours)	<u>0.582</u>	**0.511**	<u>0.611</u>	0.720	<u>0.433</u>	<u>0.347</u>	<u>0.478</u>	<u>0.591</u>	**0.891**	**0.872**	**0.901**	**0.923**
HyPlanE (Ours)	0.569	0.496	0.600	0.708	0.402	0.323	0.443	0.549	0.866	0.843	0.880	0.909

- HyPlanE: **2D convolution** + {alternate mask stack + circular padding + 1-N multilinear scoring}
- HyCubE+: **Residual module** + HyCubE

Experiment

- KHG Completion : Mixed Arity

TABLE II
RESULTS OF LINK PREDICTION ON MIXED ARITY KNOWLEDGE HYPERGRAPH DATASETS

Model	JF17K				WikiPeople				FB-AUTO			
	MRR	Hits@1	Hits@3	Hits@10	MRR	Hits@1	Hits@					
RAE [19]	0.392	0.312	0.433	0.561	0.253	0.118	0.343					
NaLP [20]	0.310	0.239	0.334	0.450	0.338	0.272	0.362					
HypE [2]	0.494	0.399	0.532	0.650	0.263	0.127	0.355					
RAM [5]	0.539	0.463	0.573	0.690	(0.363)	0.271	0.405					
HyperMLN [22]	0.556	(0.482)	0.597	0.717	0.351†	0.270†	0.394†					
tNaLP+ [23]	0.449	0.370	0.484	0.598	0.339	0.269	0.369					
PosKHG [6]	0.545	0.469	0.582	0.706	0.315†	0.214†	0.377†					
ReAlE [3]	0.530	0.454	0.563	0.677	0.332†	0.207†	(0.417†					
RD-MPNN [11]	0.512	0.445	0.573	0.685	-	-	-					
HyConvE [7]	(0.580)	0.478	(0.610)	(0.729)	0.362	(0.275)	0.388					
HyCubE (Ours)	**0.584**	0.508	**0.616**	**0.730**	**0.448**	**0.368**	**0.490**	**0.592**	0.881	0.860	0.894	0.918
HyCubE+ (Ours)	0.582	**0.511**	0.611	0.720	0.433	0.347	0.478	0.591	**0.891**	**0.872**	**0.901**	**0.923**
HyPlanE (Ours)	0.569	0.496	0.600	0.708	0.402	0.323	0.443	0.549	0.866	0.843	0.880	0.909

	$\|\mathcal{E}\|$	$\|\mathcal{R}\|$	Arity
JF17K	28,645	322	2-6
WikiPeople	47,765	707	2-9
FB-AUTO	3,388	8	2, 4, 5

❑ HyCubE+ can alleviate problem of training on datasets with low relation-specific

- low relation arity, less relation => less n-ary implicit semantic information
- prone to gradient vanishing

Experiment

- KHG Completion : Fixed Arity

TABLE III
RESULTS OF LINK PREDICTION ON FIXED ARITY KNOWLEDGE HYPERGRAPH DATASETS

Model	JF17K-3			JF17K-4			WikiPeople-3			WikiPeople-4		
	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10
GETD [21]	(0.602)	0.525	0.726	0.750	(0.700)	0.842	0.303	(0.226)	(0.463)	0.340	(0.233)	0.531
HypE [2]	0.364	0.255	0.573	0.408	0.300	0.627	0.266	0.183	0.443	0.304	0.191	0.527
HSimplE [2]	0.429	0.326	0.612	0.575	0.508	0.703	0.233	0.181	0.342	0.177	0.151	0.218
RAM [5]	0.591	0.516	0.725	0.717	0.661	0.813	0.270	0.205	0.401	0.223	0.150	0.378
HyperMLN [22]	0.574	0.501	0.711	0.734	0.687	0.831	0.252	0.193	0.385	0.224	0.167	0.370
tNaLP+ [23]	0.411	0.325	0.617	0.630	0.531	0.722	0.270	0.185	0.444	(0.344)	0.223	(0.578)
PosKHG [6]	0.597	(0.532)	(0.727)	0.749	0.692	(0.855)	0.283	0.207	0.435	0.284	0.192	0.468
ReAlE [3]	0.587	0.511	0.724	0.702	0.642	0.819	0.304	0.218	0.461	0.303	0.190	0.540
RD-MPNN [11]	0.581	0.497	0.716	0.727	0.661	0.822	0.247	0.195	0.367	0.234	0.187	0.401
HyConvE [7]	0.573	0.490	0.709	(0.751)	0.670	0.831	(0.309)	0.217	0.457	0.336	0.227	0.507
HyCubE (Ours)	0.599	0.534	0.723	0.793	0.742	**0.887**	0.336	0.256	0.499	0.367	0.258	0.578
HyCubE+ (Ours)	**0.603**	**0.537**	**0.728**	**0.794**	**0.743**	0.886	**0.345**	**0.260**	**0.515**	**0.383**	**0.269**	**0.602**
HyPlanE (Ours)	0.574	0.510	0.700	0.757	0.699	0.862	0.318	0.237	0.479	0.324	0.210	0.559

❑ HyCubE+ shows best performance in fixed arity datasets

- fixed arity makes more challenging to learn patterns from diverse relations

Experiment

- Model Efficiency: Mixed Arity

TABLE IV
RESULTS OF MODEL EFFICIENCY COMPARISON ON MIXED ARITY KNOWLEDGE HYPERGRAPH DATASETS

Model	Parameters (Millions)			GPU Memory Usage (MB)			Time Usage (s)		
	JF17K	WikiPeople	FB-AUTO	JF17K	WikiPeople	FB-AUTO	JF17K	WikiPeople	FB-AUTO
RAM [5]	≈ 14.24	≈ 27.34	≈ 1.63	15,174	20,742	2,860	74.2	215.8	4.0
PosKHG [6]	≈ 14.34	≈ 27.53	≈ 1.65	15,388	21,147	2,874	87.4	229.7	4.3
HyConvE [7]	≈ 12.80	≈ 21.44	≈ 4.80	7,718	15,430	3,032	98.7	247.3	4.9
ReAlE [3]	≈ 14.88	≈ 29.61	≈ 1.64	16,316	17,970	14,488	333.7	507.9	13.9
HyCubE (Ours)	≈ **1.28**	≈ **2.24**	≈ **0.96**	2,536	**3,108**	**1,704**	**9.7**	40.1	2.9
HyCubE+ (Ours)	≈ 5.77	≈ 13.46	≈ 1.28	**2,290**	3,424	1,914	14.9	**33.7**	**1.9**
HyPlanE (Ours)	≈ 11.52	≈ 19.20	≈ 3.84	2,628	3,482	1,972	17.4	42.1	4.1

❑ Time usage: time required for each epoch iteration of KHG embedding models

- calculated from the average of 10 epochs iteration time

Experiment

- Model Efficiency: Fixed Arity

TABLE V
RESULTS OF MODEL EFFICIENCY COMPARISON ON FIXED ARITY KNOWLEDGE HYPERGRAPH DATASETS

Model	Parameters (Millions)				GPU Memory Usage (MB)				Time Usage (s)			
	JF-3	JF-4	WP-3	WP-4	JF-3	JF-4	WP-3	WP-4	JF-3	JF-4	WP-3	WP-4
RAM [5]	$\approx$ 4.62	$\approx$ 2.61	$\approx$ 4.91	$\approx$ 3.81	3,114	2,370	1,982	1,962	14.9	5.4	11.9	10.9
PosKHG [6]	$\approx$ 4.67	$\approx$ 2.64	$\approx$ 4.95	$\approx$ 3.85	3,214	2,375	2,180	2,007	14.8	5.3	11.7	10.5
HyConvE [7]	$\approx$ 6.72	$\approx$ 9.28	$\approx$ 6.72	$\approx$ 9.28	2,774	3,354	2,580	2,248	22.4	9.5	19.1	17.7
ReAlE [3]	$\approx$ 4.66	$\approx$ 2.62	$\approx$ 4.99	$\approx$ 3.88	4,346	5,562	4,348	5,564	46.7	14.6	41.5	30.6
HyCubE (Ours)	$\approx$ **0.32**	$\approx$ **0.32**	$\approx$ **0.32**	$\approx$ **0.32**	1,972	**1,760**	**1,906**	**1,790**	**4.9**	**1.9**	**5.4**	8.9
HyCubE+ (Ours)	$\approx$ 1.60	$\approx$ 1.93	$\approx$ 1.60	$\approx$ 1.93	**1,928**	1,890	1,940	1,906	7.2	2.2	6.4	**7.4**
HyPlanE (Ours)	$\approx$ 6.40	$\approx$ 8.96	$\approx$ 6.40	$\approx$ 8.96	2,052	1,914	1,968	1,948	9.4	4.4	9.7	10.2

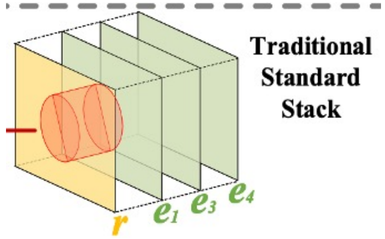
Experiment

- Ablation Study: Mixed Arity

TABLE VI
RESULTS OF ABLATION STUDY ON MIXED ARITY KNOWLEDGE HYPERGRAPH DATASETS

Model	JF17K				WikiPeople				FB-AUTO			
	MRR	Hits@1	Hits@3	Hits@10	MRR	Hits@1	Hits@3	Hits@10	MRR	Hits@1	Hits@3	Hits@10
HyCubE	**0.584**	**0.508**	**0.616**	**0.730**	**0.448**	**0.368**	**0.490**	**0.592**	**0.881**	**0.860**	**0.894**	**0.918**
w/o Alternate	0.579	0.506	0.612	0.723	0.447	0.367	0.487	0.591	0.875	0.853	0.888	0.911
w/o Circular	0.583	0.507	0.614	0.727	0.437	0.356	0.479	0.584	0.877	0.855	0.890	0.911
HyCubE+	**0.582**	**0.511**	**0.611**	**0.720**	**0.433**	**0.347**	**0.478**	**0.591**	**0.891**	**0.872**	**0.901**	**0.923**
w/o Alternate	0.577	0.505	0.607	0.719	0.432	0.345	0.477	0.590	0.887	0.869	0.897	0.919
w/o Circular	0.532	0.451	0.568	0.690	0.432	0.346	0.477	0.589	0.888	0.871	0.898	0.919
HyPlanE	0.569	0.496	0.600	0.708	0.402	0.323	0.443	0.549	0.866	0.843	0.880	0.909

- ❑ w/o Alternate: using Standard Stack
- ❑ w/o Circular: using Zero-padding



Experiment

- Ablation Study: Fixed Arity

TABLE VII
RESULTS OF ABLATION STUDY ON FIXED ARITY KNOWLEDGE HYPERGRAPH DATASETS

Model	JF17K-3			JF17K-4			WikiPeople-3			WikiPeople-4		
	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10	MRR	Hits@1	Hits@10
HyCubE	**0.599**	**0.534**	**0.723**	**0.793**	**0.742**	**0.887**	**0.336**	**0.256**	**0.499**	**0.367**	**0.258**	**0.578**
w/o Alternate	0.596	0.532	0.718	0.789	0.738	0.884	0.332	0.252	0.492	0.363	0.248	0.577
w/o Circular	0.590	0.525	0.714	0.790	0.738	0.886	0.330	0.252	0.490	0.361	0.249	0.576
HyCubE+	**0.603**	**0.537**	**0.728**	**0.794**	**0.743**	**0.886**	**0.345**	**0.260**	**0.515**	**0.383**	**0.269**	**0.602**
w/o Alternate	0.601	0.534	0.727	0.791	0.740	0.882	0.343	0.255	0.514	0.377	0.260	0.597
w/o Circular	0.601	0.535	0.726	0.792	0.741	0.885	0.340	0.254	0.512	0.373	0.253	0.597
HyPlanE	0.574	0.510	0.700	0.757	0.699	0.862	0.318	0.237	0.479	0.324	0.210	0.559

Conclusion

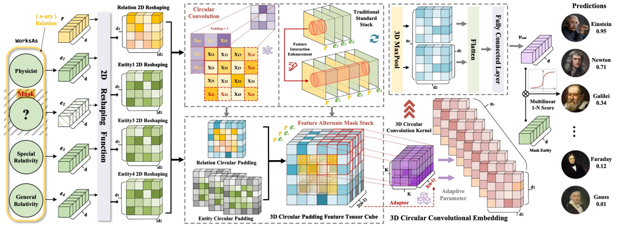
❖ Previous Work

- ❑ Existing methods focused on “effectiveness”
  - ▸ higher computational costs and memory usage

Model	Parameters (Millions)			GPU Memory Usage (MB)			Time Usage (s)		
	JF17K	WikiPeople	FB-AUTO	JF17K	WikiPeople	FB-AUTO	JF17K	WikiPeople	FB-AUTO
RAM [5]	≈ 14.24	≈ 27.34	≈ 1.63	15,174	20,742	2,860	74.2	215.8	4.0
PosKHG [6]	≈ 14.34	≈ 27.53	≈ 1.65	15,388	21,147	2,874	87.4	229.7	4.3
HyConvE [7]	≈ 12.80	≈ 21.44	≈ 4.80	7,718	15,430	3,032	98.7	247.3	4.9
ReAlE [3]	≈ 14.88	≈ 29.61	≈ 1.64	16,316	17,970	14,488	333.7	507.9	13.9

❖ HyCubE

- ❑ Alternative Mask Stacking, 3D convolution w/ circular padding improve the interaction area between relations, entities
- ❑ 3D maxpooling, 1-N Multilinear Scoring enhanced “efficiency”



❖ Experiment

- ❑ HyCubE shows best performance both of ”effectiveness” and “efficiency”
- ❑ Alternative Mask Stacking, Circular padding have shown the impact in ablation study