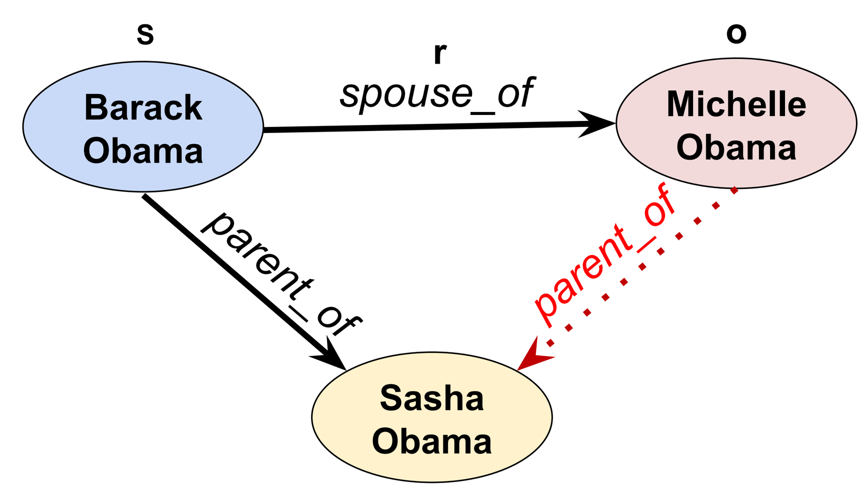
InteractE: Improving Convolution-based Knowledge Graph Embeddings by Increasing Feature Interactions

Shikhar Vashishth[*1,2], Soumya Sanyal[*1], Vikram Nitin[3], Nilesh Agrawal[1], Partha Talukdar[1]
[1]Indian Institute of Science, [2]Carnegie Mellon University, [3]Columbia University

Knowledge Graph Link Prediction

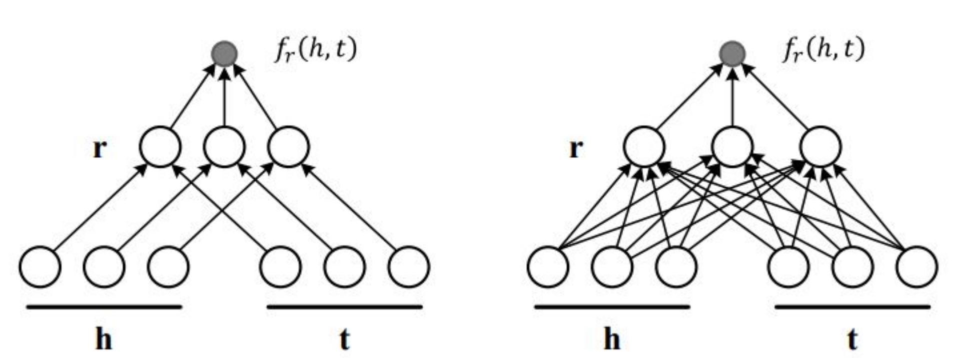
The task of inferring new facts based on the existing facts in the Knowledge Graph (KG).



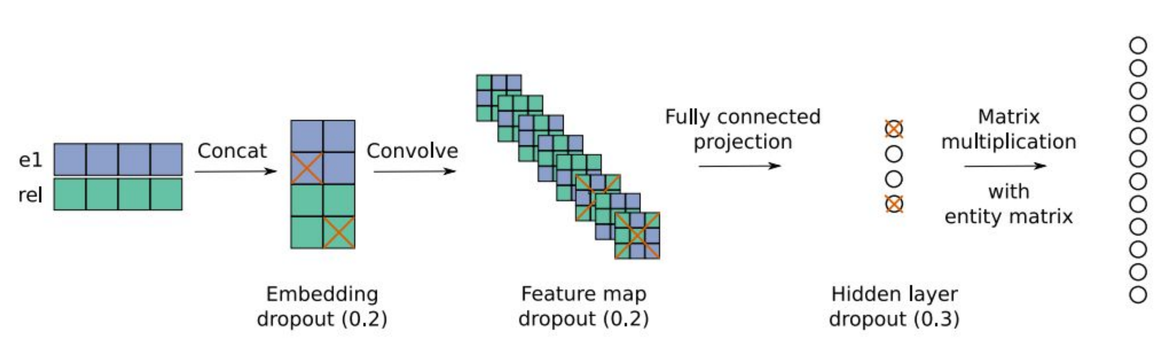
General technique involves learning a representation for all entities and relations in KG.

Increasing Interactions Helps!

Prior works have demonstrated that increasing interaction between features improves link prediction performance.



DistMult (Product) HolE (Circular Correlation)

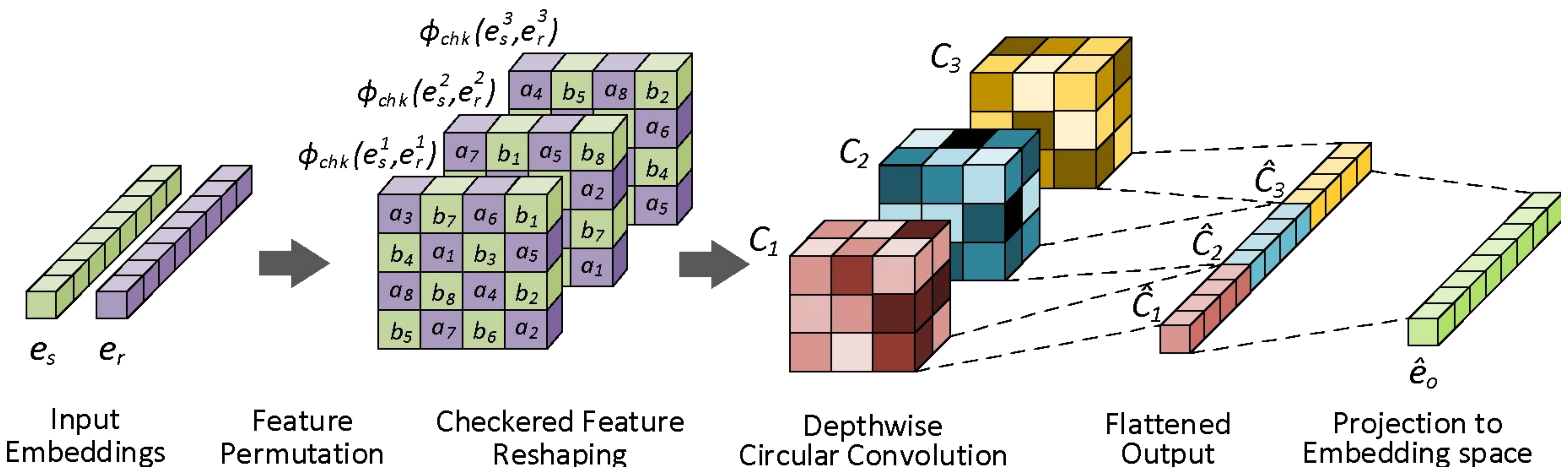


ConvE (2D-Convolution)

Contributions

1. We propose InteractE, which augments the expressive power of ConvE through feature permutation, "checkered" reshaping, and circular convolution.
2. Establish correlation between the number of interactions and link prediction performance. Theoretically, we demonstrate that InteractE increases interactions compared to ConvE.

InteractE Overview

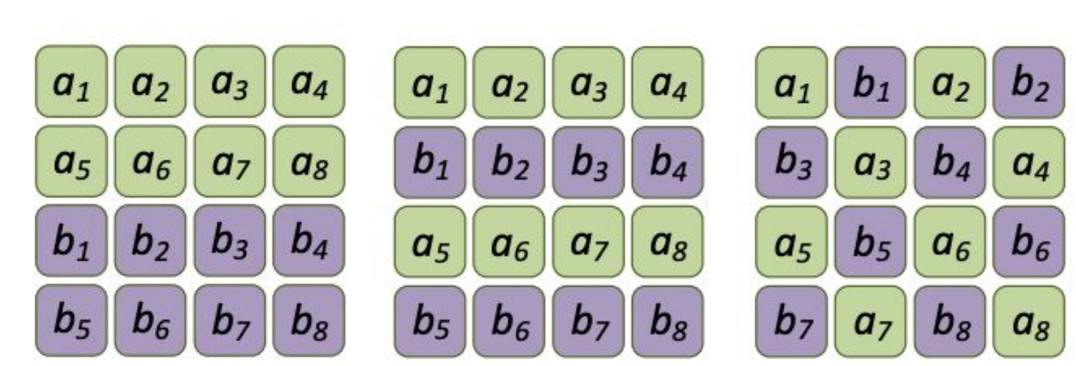


Given entity and relation embeddings, InteractE generates multiple permutations ($\mathcal{P}_k$) of these embeddings and reshapes them using a "Checkered" reshaping function (ϕ). Depth-wise circular convolution ($\circledast$) is employed to convolve each of the reshaped permutations, which are then flattened (vec) and fed to a fully-connected layer to generate the predicted object embedding.

Score Function: $\psi(s,r,o) = g(\mathrm{vec}(f(\phi(\mathcal{P}_k) \circledast \boldsymbol{w}))\boldsymbol{W})e_o$

Components of InteractE

Feature Reshaping:

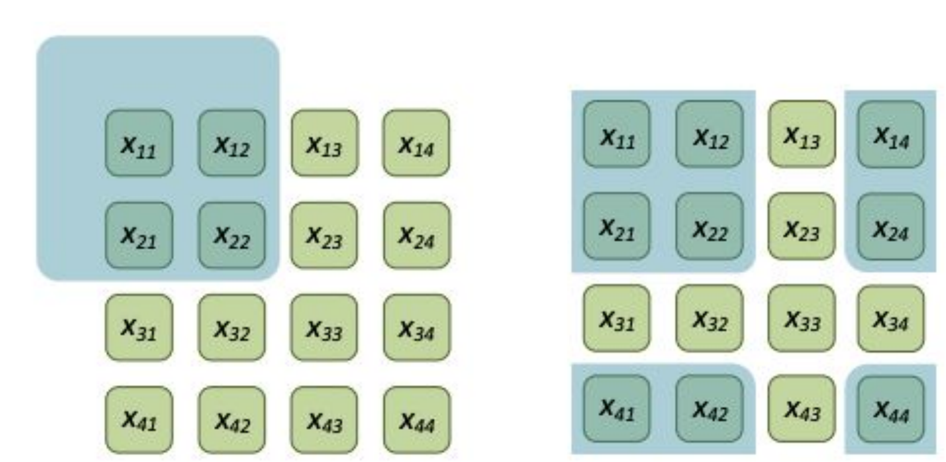


(a) Stack (b) Alternate (c) Chequer

Proposition 7.3. *For any kernel w of size k and for all reshaping functions $\phi : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^{n \times n}$, the following statement holds:*

$$\mathcal{N}_{het}(\phi_{chk}, k) \geq \mathcal{N}_{het}(\phi, k)$$

Circular Convolution:



(a) Standard Convolution (b) Circular Convolution

Proposition 7.4. *Let Ω_0, $\Omega_c : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{(n+p)\times(n+p)}$ denote zero padding and circular padding functions respectively, for some $p > 0$. Then for any reshaping function ϕ,*

$$\mathcal{N}_{het}(\Omega_c(\phi), k) \geq \mathcal{N}_{het}(\Omega_0(\phi), k)$$

Acknowledgement

This work is supported in part by the Ministry of Human Resource Development (Government of India) and Google PhD Fellowship.

Source Code

Source code is available at:
github.com/malllabiisc/InteractE

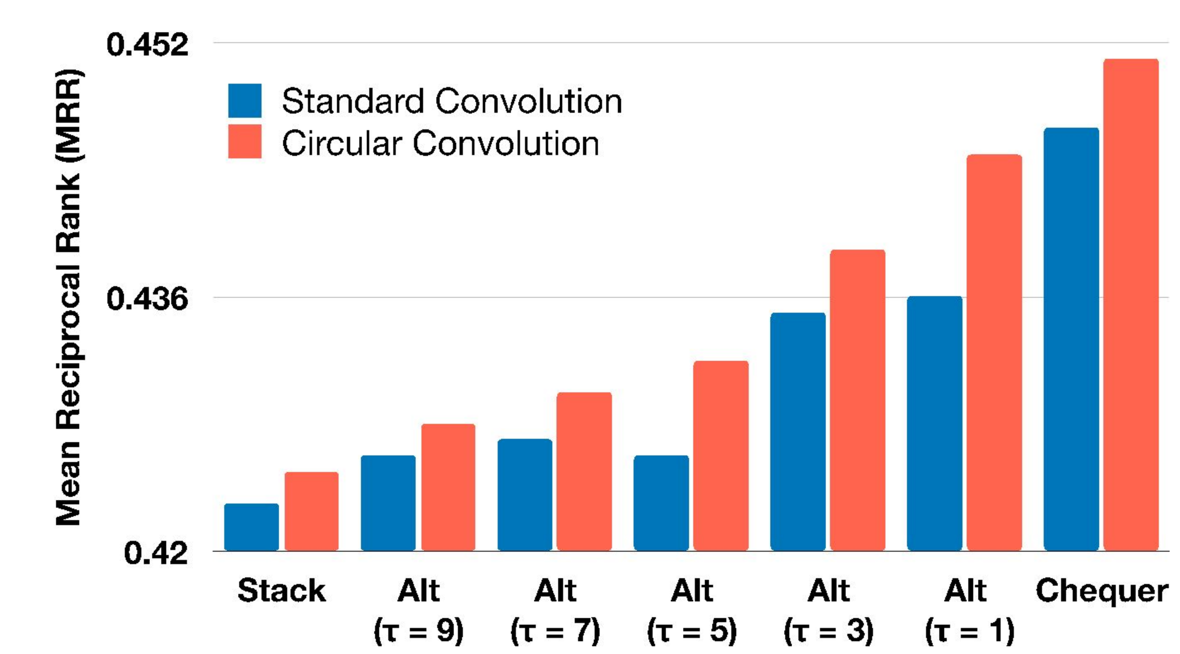
Contact
soumyasanyal@iisc.ac.in
svashish@cs.cmu.edu

Results

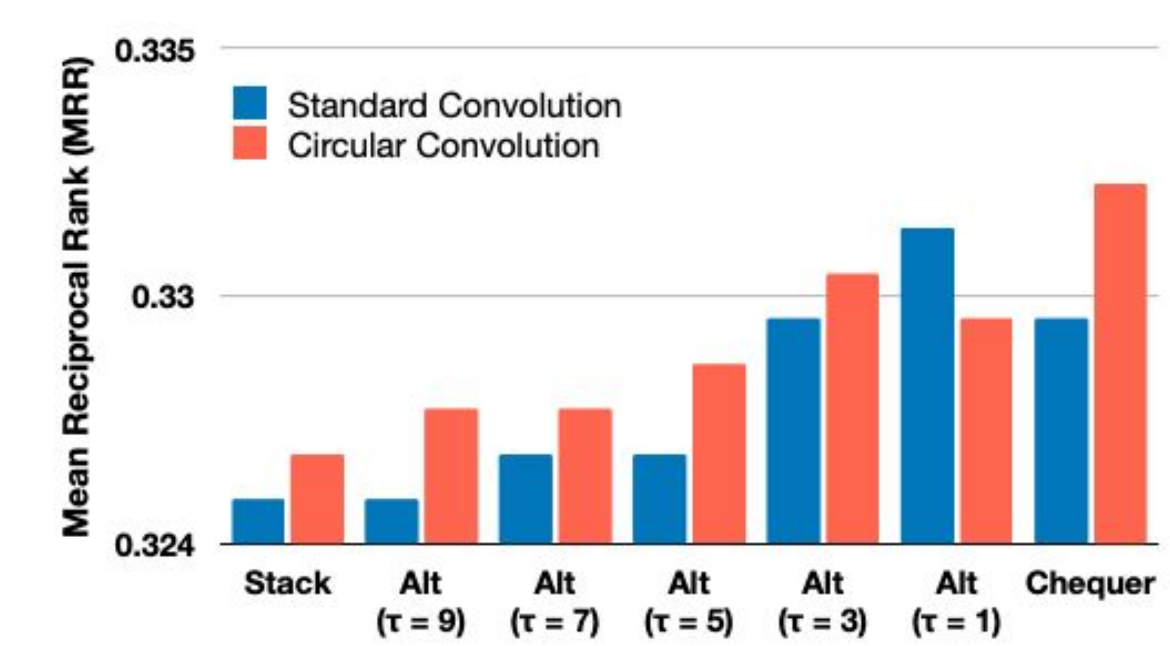
	FB15k-237			WN18RR		
	MRR	H@10	H@1	MRR	H@10	H@1
SACN	.35	.54	.26	.47	.54	.43
RotatE	.338	.533	.241	**.476**	**.571**	.428
ConvE	.325	.501	.237	.43	.52	.40
InteractE	**.354**	**.535**	**.263**	.463	.528	**.430**

Link prediction results on FB15k-237 and WN18RR. We find that InteractE outperforms or gives comparable performance across all the datasets.

Effect of Feature Reshaping Function



Performance on WN18RR



Performance on FB15k-237

As we decrease τ, the number of heterogeneous interactions increases. The results empirically validate that increasing heterogeneous interactions improve link prediction.