

# Zero-shot Recognition via Semantic Embeddings and Knowledge Graphs (CVPR 2018)

Jonghyun Lee

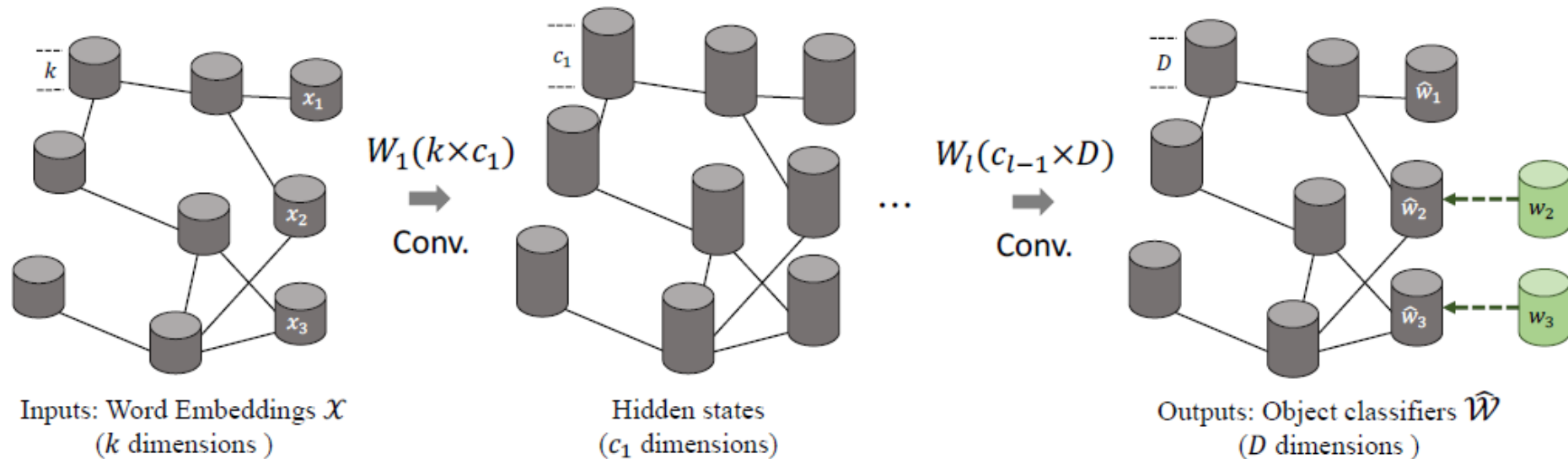
Seoul National University

# Zero-shot Recognition

- Training 중 한 번도 보지 못한 category를 맞추는 image classifier를 얻는 것이 목표
- 당시 기준 최근 work들은 image의 feature와 category의 word embedding을 matching 시키는 방법을 고안
- 여전히 잘 아는 class의 정보를 어떻게 처음 보는 class에 distill시킬지가 큰 문제

# GCN for Zero-shot Learning

- 총  $n$ 개의 category
  - Image와 label이 같이 존재하는  $m$ 개의 category를 포함
- Class의 word embedding과 그 relationship을 이용해 relationship graph 생성 가능
- 기존 방법들보다 deep하게 convolution할 수 있다.(2->6 layers)



# GCN for Zero-shot Learning

- 각 class를 word embedding한다. (ex. GloVe)
- 우리가 알고있는 m개의 class에 대해서 logistic regression classifier weight를 학습시킨다.
- Class word embedding을 node로 relationship graph를 만들고 이를 GCN을 통과시켜 output을 내보낸다.
- 그 output과 우리가 알고있는 m개의 class의 logistic regression weight간의 MSE를 줄여 둘의 space를 인접하도록 만든다.
- Test할 때 각 class의 logistic regression weight를 이용해 classification을 진행한다.

# Experiments

- ImageNet에서 한 번도 보지 않은 class에 대한 classification 결과
- Training set : ImageNet2012 1K dataset
- WordNet의 sub-graph를 knowledge graph로 사용
- 2-hops가 가장 training set과 비슷함
- All이 가장 training set과 차이 큼
- 2-hops에서는 약 20% point차이로 앞지름

Test Set	Model	ConvNets	Hit@k (%)				
			1	2	5	10	20
2-hops	ConSE [4]	Inception-v1	8.3	12.9	21.8	30.9	41.7
	ConSE(us)	Inception-v1	12.4	18.4	25.3	28.5	31.8
	SYNC [4]	Inception-v1	10.5	17.7	28.6	40.1	52.0
	EXEM [5]	Inception-v1	12.5	19.5	32.3	43.7	55.2
	Ours	Inception-v1	18.5	31.3	50.1	62.4	72.0
	Ours	ResNet-50	<b>19.8</b>	<b>33.3</b>	<b>53.2</b>	<b>65.4</b>	<b>74.6</b>
3-hops	ConSE [4]	Inception-v1	2.6	4.1	7.3	11.1	16.4
	ConSE(us)	Inception-v1	3.2	4.9	7.6	9.7	11.4
	SYNC [4]	Inception-v1	2.9	4.9	9.2	14.2	20.9
	EXEM [5]	Inception-v1	3.6	5.9	10.7	16.1	23.1
	Ours	Inception-v1	3.8	6.9	13.1	18.8	26.0
	Ours	ResNet-50	<b>4.1</b>	<b>7.5</b>	<b>14.2</b>	<b>20.2</b>	<b>27.7</b>
All	ConSE [4]	Inception-v1	1.3	2.1	3.8	5.8	8.7
	ConSE(us)	Inception-v1	1.5	2.2	3.6	4.6	5.7
	SYNC [4]	Inception-v1	1.4	2.4	4.5	7.1	10.9
	EXEM [5]	Inception-v1	<b>1.8</b>	2.9	5.3	8.2	12.2
	Ours	Inception-v1	1.7	3.0	5.8	8.4	11.8
	Ours	ResNet-50	<b>1.8</b>	<b>3.3</b>	<b>6.3</b>	<b>9.1</b>	<b>12.7</b>

# Conclusion

- Zero-shot recognition을 word embedding과 GCN with knowledge graph를 연관시켜 novel한 방법론을 제시
- Zero-shot recognition task에서 state-of-the-art 성능을 큰 차이로 얻어냄