Rethinking Knowledge Graph Propagation for Zero-Shot Learning
Supplementary Materials

Michael Kampffmeyer*[1], Yinbo Chen*[2], Xiaodan Liang[3], Hao Wang[4], Yujia Zhang[5], and Eric P. Xing[6]

[1]UiT The Arctic University of Norway, [2]Tsinghua University, [3]Sun Yat-sen University, [4]Massachusetts Institute of Technology, [5]Institute of Automation, Chinese Academy of Sciences, [6]Carnegie Mellon University

1. Additional Qualitative Examples

Figure 1 and 2 provide further qualitative results of our single-hidden-layer GCN (SGCN) and Dense Graph Propagation Module (DGP) compared to a standard ResNet and GCNZ, our reimplementation of [6].

2. Performance improvements between GCNZ and SGCN

Table 1 explains the performance difference between our SGCN, our reimplementation of GCNZ and the reported results in [6]. Note, unless otherwise stated training is performed for 3000 epochs. Non-symmetric normalization ($D^{-1}A$) is denoted as *non-sym* in the normalization column, while a symmetric normalization ($D^{-1/2}AD^{-1/2}$) is denoted as *sym*. No finetuning has been performed for SGCN in these results.

Table 1: Illustration of the improvements between the original results of GCNZ in [6], our reimplementation of GCNZ and our SGCN. GCNZ† corresponds to updated results from [6] (taken from https://github.com/JudyYe/zero-shot-gcn). GCNZ‡ is our reimplementation of [6].

Model	Norm	**Hit@k (%)**				
		1	2	5	10	20
GCNZ (300 epochs) [6]	sym	19.8	33.3	53.2	65.4	74.6
GCNZ† (300 epochs) [6]	sym	21.0	33.7	52.7	64.8	74.3
GCNZ‡ (300 epochs)	sym	21.4	34.7	54.3	67.5	77.6
GCNZ‡	sym	23.5	36.9	56.5	68.8	78.0
SGCN (ours)	sym	24.6	38.1	57.6	70.0	79.7
SGCN (ours)	non-sym	24.8	38.3	57.5	69.9	79.6

*Indicates equal contribution.

3. Performance on AWA2

AWA2 is a replacement for the original AWA dataset and represents more traditional zero-shot learning datasets, where most approaches rely on class-attribute information. It consists of 50 animal classes, with a total of 37,322 images and an average of 746 per class. The dataset further consists of 85-attribute features per class. We report results on the proposed split in [7] to ensure that there is no overlap between the test classes and the ImageNet 2012 dataset. In the proposed split, 40 classes are used for training and 10 for testing. AWA2 test classes are contained in the 21K ImageNet classes and several of the training classes (24 out of 40) that are in the proposed split overlap with the ImageNet 2012 dataset. We, therefore, use a unified approach for both datasets.

Results for the AWA2 dataset are presented in Table 2. Note that our model differs considerably from the baselines as it does not make use of the attributes provided in the dataset. To illustrate the merits of our approach, we reimplement [6], as it represents the method which is closest related to our approach and also makes use of word embeddings and a knowledge graph. We observe that our methods also outperforms [6], however, the improvement is lower than on the ImageNet dataset, which we believe is due to the arguably simpler task with the number of classes being considerably lower. Note, all methods, except SYNC, use a pretrained network trained on the 1K ImageNet classes. GCNZ and our DGP do not make use of the attribute information supplied for AWA2, however, both methods use the ImageNet knowledge graph.

4. Comparison to Graph Attention Networks

Table 3 illustrates the results for a 1-hidden-layer and 2-hidden-layer GCN with the attention mechanism proposed in GAT [5]. Note, performance degrades compared to a 1-hidden-layer GCN (i.e. SGCN(-f)). The hidden dimension

Table 2: Top-1 accuracy results for unseen classes on AWA2. Results for ConSE, Devise and SYNC obtained from [7].

Model	**ACC (%)**
ConSE [4]	44.5
Devise [2]	59.7
SYNC [1]	46.6
SE-GZSL [3]	69.2
Gaussian-Ort [8]	70.5
GCNZ [6]	70.7
DGP (ours)	**77.3**

is 2048 and training settings are the same as in the paper.

Table 3: Accuracy on ImageNet for a 1- and 2-hidden-layer GAT [5] compared to a 1-hidden-layer GCN without attention.

Test set	**Model**	**Hit@k (%)**				
		1	2	5	10	20
2-hops	GAT-1	24.1	37.5	57.2	69.7	79.4
	GAT-2	23.3	36.9	56.8	68.7	77.9
	GCN-1 (ours)	**24.8**	**38.3**	**57.5**	**69.9**	**79.6**

References

[1] S. Changpinyo, W.-L. Chao, B. Gong, and F. Sha. Synthesized classifiers for zero-shot learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5327–5336, 2016. 2

[2] A. Frome, G. S. Corrado, J. Shlens, S. Bengio, J. Dean, T. Mikolov, et al. Devise: A deep visual-semantic embedding model. In *Advances in Neural Information Processing Systems*, pages 2121–2129, 2013. 2

[3] V. Kumar Verma, G. Arora, A. Mishra, and P. Rai. Generalized zero-shot learning via synthesized examples. In *CVPR*, 2018. 2

[4] M. Norouzi, T. Mikolov, S. Bengio, Y. Singer, J. Shlens, A. Frome, G. S. Corrado, and J. Dean. Zero-shot learning by convex combination of semantic embeddings. *International Conference for Learning Representation*, 2014. 2

[5] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio. Graph attention networks. *ICLR*, 2018. 1, 2

[6] X. Wang, Y. Ye, and A. Gupta. Zero-shot recognition via semantic embeddings and knowledge graphs. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018. 1, 2

[7] Y. Xian, C. H. Lampert, B. Schiele, and Z. Akata. Zero-shot learning-a comprehensive evaluation of the good, the bad and the ugly. *IEEE transactions on pattern analysis and machine intelligence*, 2018. 1, 2

[8] H. Zhang and P. Koniusz. Zero-shot kernel learning. In *CVPR*, 2018. 2

ResNet: upright, grand piano, organ, accordion, barbershop
GCNZ: piano, spinet, keyboard instrument, concert grand, **baby grand**
SGCN: piano, spinet, concert grand, **baby grand**, keyboard instrument
DGP: piano, **baby grand**, concert grand, spinet, keyboard instrument

ResNet: breakwater, aircraft carrier, seashore, wing, sandbar
GCNZ: barrier, **bar**, shore, grate, geological formation
SGCN: littoral, **bar**, seaside, barrier, landfall,
DGP: **bar**, littoral, shore, seaside, landfall

ResNet: lemon, orange, banana, spaghetti squash, fig
GCNZ: **bitter orange**, temple orange, citrus, sweet orange, edible fruit
SGCN: citrus, **bitter orange**, temple orange, sweet orange, edible fruit,
DGP: citrus, **bitter orange**, sweet orange, temple orange, edible fruit

ResNet: lycaenid, cabbage butterfly, ringlet, sulphur butterfly, damselfly
GCNZ: pierid, small white, large white, hairstreak, southern cabbage butterfly
SGCN: **blue**, hairstreak, copper, pierid, butterfly,
DGP: **blue**, hairstreak, copper, pierid, butterfly

ResNet: candle, altar, lighter, lipstick, perfume
GCNZ: vigil light, rushlight, **chandlery**, dip, lamp
SGCN: vigil light, rushlight, **chandlery**, dip, high altar
DGP: vigil light, **chandlery**, rushlight, dip, flambeau

ResNet: bagel, french loaf, cheeseburger, dough, hotdog
GCNZ: onion bagel, bun, loaf of bread, **cracker**, bread dough
SGCN: onion bagel, bun, bread dough, pastry, sandwich
DGP: bun, onion bagel, bread dough, **cracker**, pastry

ResNet: walking stick, jacamar, hip, house finch, chainlink fence
GCNZ: **diapheromera**, phasmid, finch, oscine, praying mantis
SGCN: **diapheromera**, phasmid, neuropteron, thrush, finch
DGP: **diapheromera**, phasmid, thrush, titmouse, oscine

ResNet: desktop computer, monitor, screen, computer keyboard, mouse
GCNZ: personal computer, portable computer, planner, computer, computer screen
SGCN: personal computer, computer, computer screen, **display**, television monitor
DGP: personal computer, background, computer screen, portable computer, **display**

ResNet: bittern, partridge, coucal, ruffed grouse, kite
GCNZ: least bittern, american bittern, **european bittern**, phasianid, crow pheasant
SGCN: american bittern, **european bittern**, least bittern, plain turkey, great bustard
DGP: american bittern, least bittern, **european bittern**, heron, egret

ResNet: damselfly, dragonfly, lacewing, walking stick, grasshopper
GCNZ: **odonate**, neuropteron, hymenopterous insect, phasmid, brown lacewing
SGCN: **odonate**, neuropteron, brown lacewing, green lacewing, phasmid
DGP: **odonate**, brown lacewing, green lacewing, neuropteron, phasmid

ResNet: macaw, lorikeet, bee eater, sulphur-crested cockatoo, house finch
GCNZ: lory, **parrot**, rainbow lorikeet, varied lorikeet, cockatoo
SGCN: **parrot**, lory, rainbow lorikeet, varied lorikeet, cockatoo
DGP: lory, **parrot**, cockatoo, rainbow lorikeet, varied lorikeet

ResNet: grocery store, confectionery, tobacco shop, restaurant, butcher shop
GCNZ: marketplace, greengrocery, **supermarket**, shop, tuck shop
SGCN: **supermarket**, marketplace, greengrocery, tuck shop, shop
DGP: **supermarket**, greengrocery, marketplace, tuck shop, shop

ResNet: cliff, valley, lakeside, alp, promontory
GCNZ: geological formation, natural elevation, natural depression, mountain, ravine
SGCN: **precipice**, crag, natural depression, ravine, natural elevation
DGP: natural depression, geological formation, natural elevation, crag, **precipice**

ResNet: church, monastery, dome, bell cote, mosque
GCNZ: kirk, cathedral, abbey, basilica, cathedral
SGCN: abbey, cathedral, friary, basilica, cathedral
DGP: cathedral, abbey, cathedral, basilica, kirk

true label: **place_of_worship**