Heterogeneous Graph-based Knowledge Transfer for Generalized Zero-shot Learning

Junjie Wang jasonwang.ecnu@gmail.com
Software Engineering Dept, ECNU, CHINA

Summary

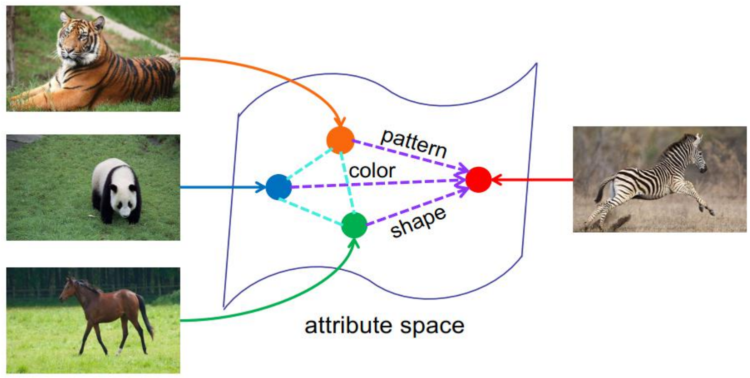
Problem: **(1)** how to capture relationship between all seen and unseen classes? **(2)** how to transfer knowledge based on this relationship?

Contributions:

1. We capture **inter-class** and **intra-class** relationship jointly by constructing a heterogeneous structured graph.
2. Instead of averaging instances directly, we utilize **Wasserstein** metric to extract more representative node of each class.
3. Our approach is the novel **inductive** GNN-based GZSL method that is **agnostic** to unseen information during training. Knowledge is transferred from seen classes to new unseen classes based on the learned aggregation and embedding functions.

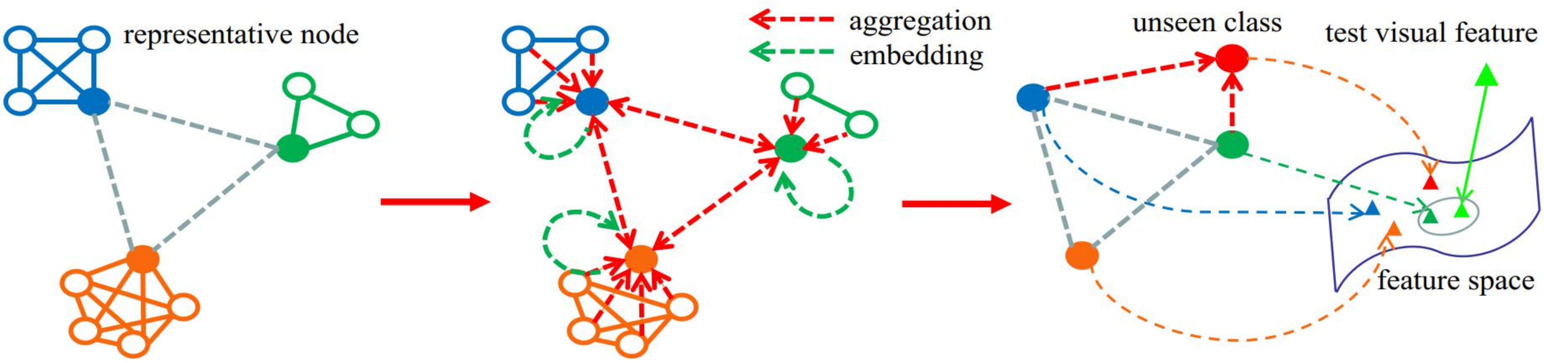
Problem Setting

Instances can be described by some common high-level **semantic information**.
Transfer knowledge learned from seen classes exploring those common semantic information.



Algorithm

Left: each class corresponds to a **complete graph** and all complete subgraphs are connected based on their representative nodes. Middle: the embedding vector of representative nodes are produced by both **aggregation** and **embedding** functions. Right: connect the new unseen class with **k-nearest** seen classes in visual feature space; select the nearest class by visual feature distance as the prediction for each test sample.



Main Result

Comparison of GZSL methods on public datasets.

Dataset		SUN			CUB			AWA1		
Method	Type	ts	tr	H	ts	tr	H	ts	tr	H
DAP [11]	Inductive	4.2	25.1	7.2	1.7	67.9	3.3	0.0	**88.7**	0.0
IAP [11]	Inductive	1.0	37.8	1.8	0.2	72.8	0.4	2.1	78.2	4.1
CONSE [29]	Inductive	6.8	39.9	11.6	1.6	72.2	3.1	0.4	88.6	0.8
CMT [30]	Inductive	8.1	21.8	11.8	7.2	49.8	12.6	0.9	87.6	1.8
CMT* [30]	Inductive	8.7	28.0	13.3	4.7	60.1	8.7	8.4	86.9	15.3
SSE [12]	Inductive	2.1	36.4	4.0	8.5	46.9	14.4	7.0	80.5	12.9
LATEM [31]	Inductive	14.7	28.8	19.5	15.2	57.3	24.0	7.3	71.7	13.3
ALE [32]	Inductive	21.8	33.1	26.3	23.7	62.8	34.4	16.8	76.1	27.5
DEVISE [18]	Inductive	16.9	27.4	20.9	23.8	53.0	32.8	13.4	68.7	22.4
SJE [33]	Inductive	14.7	30.5	19.8	23.5	59.2	33.6	11.3	74.6	19.6
ESZSL [13]	Inductive	11.0	27.9	15.8	12.6	63.8	21.0	6.6	75.6	12.1
SYNC [34]	Inductive	7.9	**43.4**	13.4	11.5	**70.9**	19.8	8.9	87.3	16.2
SAE [35]	Inductive	8.8	18.0	11.8	7.8	54.0	13.6	1.8	77.1	3.5
GFZSL [36]	Inductive	0.0	39.6	0.0	0.0	45.7	0.0	1.8	80.3	3.5
DEM [15]	Inductive	20.5	34.3	25.6	19.6	57.9	29.2	32.8	84.7	47.3
PSRZSL [19]	Inductive	20.8	37.2	26.7	24.6	54.3	33.9	-	-	-
GAFE [21]	Inductive	19.6	31.9	24.3	22.5	52.1	31.4	25.5	76.6	38.2
HGKT (Ours)	Inductive	**22.3**	36.5	**27.7**	**25.2**	56.9	**34.9**	**39.4**	83.5	**53.6**

Ablation Study

Effectiveness of **intra-class** relationship. our model can transfer knowledge from seen class efficiently.

Dataset	Intra-class	ts	tr	H
SUN	None	22.2	33.4	26.7
	CG	22.3	36.5	**27.7**
CUB	None	24.4	56.7	34.1
	CG	25.2	56.9	**34.9**
AwA1	None	34.9	82.6	49.1
	CG	39.4	83.5	**53.6**
AwA2	None	36.1	85.5	50.7
	CG	37.9	86.5	**52.7**
aPY	None	14.3	56.7	22.8
	CG	18.3	79.0	**29.7**

Selecting methods:

(1) Wasserstein metric

(2) Euclidean metric

(3) Random selection