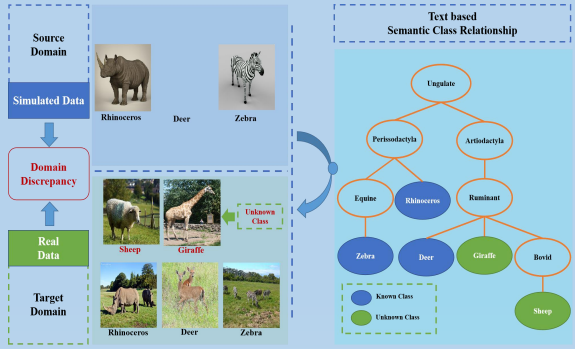
Open Set Domain Recognition via Attention-Based GCN and Semantic Matching Optimization

Xinxing He, Yuan Yuan, Zhiyu Jiang*
School of Computer Science and Center for OPTical IMagery Analysis and Learning(OPTIMAL),
Northwestern Polytechnical University, Xi'an, Shaanxi, P.R. China

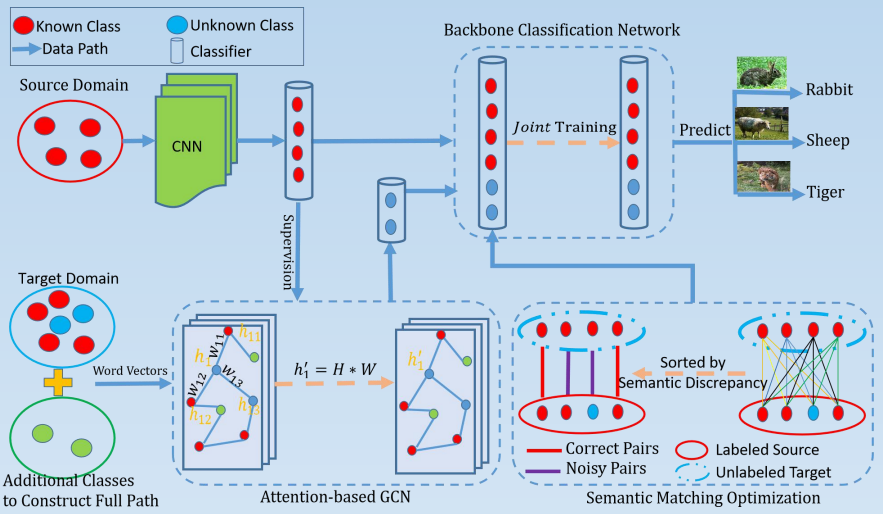
ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

INTRODUCTION



- **Task**: The goal of open set domain recognition is to specifically classify each sample in the practical unlabeled target domain, which consists of all known classes in the simulative labeled source domain and target-specific unknown categories.
- **Motivation**: Considering that the domain discrepancy between the target domain and the source domain has restricted the ability to generalize of current approaches which may transfer biased classification rules from known to unknown categories.
- **Contributions**: First, Attention-based GCN is designed to make full use of category similarity on knowledge transfer, so that the unknown classes can learn more discriminating feature representations and further obtain more accurate visual classifier. Besides, semantic matching optimization method is proposed to perform domain invariant feature learning by minishing the domain discrepancy measured on the optimal matching pair.

OUR APPROACH



Attention-based GCN for Knowledge Propagation

We first build a graph with n nodes where each node represents a distinct class. Self-attention mechanism is performed on these nodes. The attention coefficient is computed as

$$\partial_{i,j} = f_a(W \times h_i, W \times h_j).$$

After obtaining attention coefficient, the final output features of every node is expressed as

$$z_i = \sum_{j \in n_i} \partial_{ij} \times W \times h_j.$$

By minimizing the loss from the m classes with abundant supervision to estimate the network parameter for the attention-based GCN. The loss is

$$L_{tran} = \frac{1}{m} \sum_{i=1}^{m} (z_i - w_i)^2.$$

Semantic Matching Optimization

For each sample in the target domain, we calculate the feature distance between it and each sample in the source domain. Then the source sample with the minimum distance is selected as its optimal pair. Semantic consistency is utilized to filter such noisy matching. The filter function is

$$I = \begin{cases} 1, & d(p_i^s, p_i^t) < \tau, \\ 0, & else. \end{cases}$$

After filtering out the noisy matching, the domain discrepancy loss can be measured as

$$L_d = \sum_{I} d(f_i^s, f_i^t) \times I.$$

EXPERIMENTS

Table1: The Top1 classification accuracy on I2AwA dataset.

Method	Known	Unknown	All
zGCN	77.2	21.0	65.0
dGCN	78.2	11.6	64.0
adGCN	77.3	15.0	64.1
bGCN	84.6	28.0	72.6
pmd-bGCN	84.7	27.1	72.5
UODTN	84.7	31.7	73.5
AGCN-SMO(ours)	**85.1**	**34.3**	**74.3**

Table2: The Top1 classification accuracy on I2CIFAR dataset.

Method	Known	Unknown	All
zGCN	50.1	5.6	21.7
dGCN	59.9	3.1	22.1
adGCN	59.3	4.2	22.6
bGCN	65.3	9.3	27.9
pmd-bGCN	63.9	11.8	29.2
UODTN	64.2	13.6	30.4
AGCN-SMO(ours)	**65.1**	**15.7**	**32.1**

Figure1: Visualization of top 3 prediction results for various image inputs on I2AwA dataset.

CONCLUSIONS

01

Attention-based GCN can effectively improve the unknown categories to learn more discriminating feature representations and further obtain more accurate visual classifier.

Make full use of category similarity on knowledge transfer.

Semantic matching optimization can achieve good performance on domain invariant feature learning in the presence of distractions from various unknown categories.

Minish the domain gap measured on the optimal matching pair.

02