Learning The Set Graphs: Image-Set Classification Using Sparse Graph Convolutional Networks

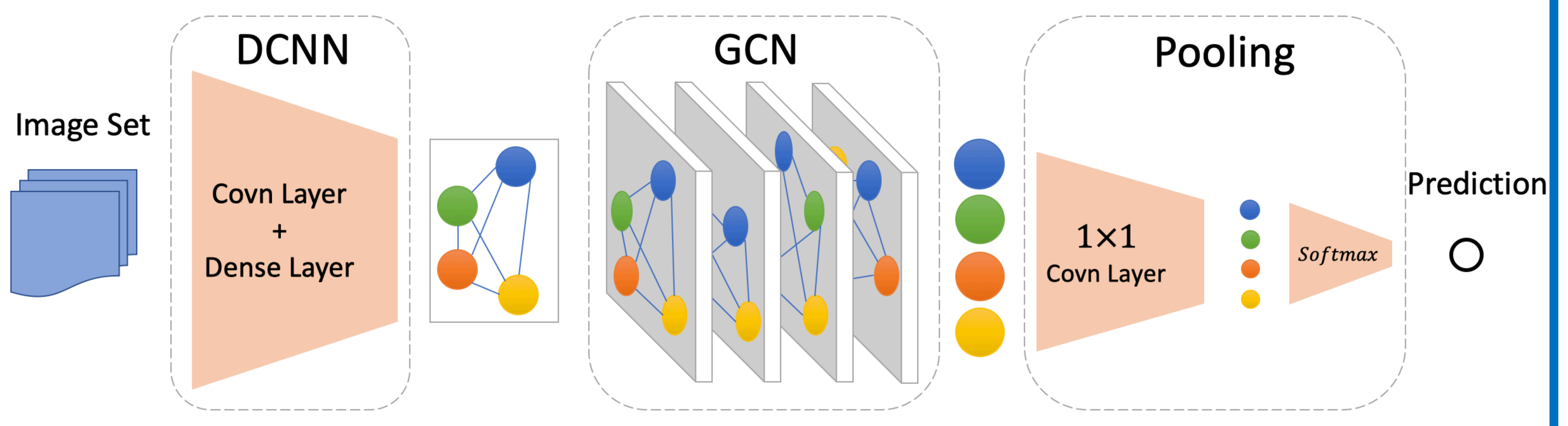
Haoliang Sun[1], Xiantong Zhen[2], and Yilong Yin[1]

[1] Shandong University, China [2] Inception Institute of Artificial Intelligence, UAE

Introduction

- We model the image set as a graph and formulate image set classification as the graph matching task.
- We build the first end-to-end graph convolutional network, the Deep SetNet, to learn the graph structure of an image set.
- We impose the $\ell_{1,2}$-norm based sparsity constraint to select vertex features in the set graph to improve the model generalization capability.

Deep SetNet



Deep CNN

- Convolutional layers and fully-connected layers.
- Input a set of images with different cardinality into the DCNN each time with variant batch size

Deep GCN

- The graph convolutional operation for the $G = (A, Z)$

$$\hat{Z} = g(D^{-1}AZW)$$

- Stacking to the deep GCN

$$Z^{t+1} = g(D^{-1}AZ^tW^t)$$

Graph Pooling Layers

- Vertex features are aggregated to the graph feature

$$v = \frac{1}{m}\sum_k Z'_k$$

Joint Sparsity for Vertex Selection

- Impose the $\ell_{1,2}$-norm constraint on the vertex matrix Z

The Objective

$$\frac{1}{n}\sum_{\forall X} y\log(\hat{y}) + \lambda \sum_{k=1}^{l} ||Z^k||_{1,2}$$

Results

Table 1. The performance on the UCSD dataset.

Discriminant Canonical Correlations (DCC) [14]	91.5 ± 3.4
Manifold Discriminant Analysis (MDA) [25]	92.7 ± 3.6
Grassmann Discriminant analysis (GDA) [17]	92.5 ± 2.6
Covariance Discriminative learning (CDL) [11]	91.7 ± 0.9
SPD Manifold Learning (SPD-ML) [15]	92.1 ± 1.5
Log-Euclidean metric learning (LEML) [16]	92.5 ± 2.9
Discriminant Grassmann kernels (DGK) [2]	96.5 ± 1.7
Deep SetNet (ours)	**98.3 ± 0.8**

Table 2. The performance on the ETH-80 dataset.

Manifold Discriminant Analysis (MDA) [25]	89.0 ± 2.0
Graph-embeded Discriminant Analysis (GEDA) [18]	92.1 ± 2.0
Analysis on Riemannian of Gaussian (DARG) [13]	92.3 ± 2.4
kFDA with the Hellinger Distance (kFDA-HL) [7]	93.7 ± 1.4
Deep Reconstrction Model (DRM) [5]	94.1 ± 1.9
NN Classifier with the J-divergence (NN-J-DR) [7]	93.8 ± 2.8
Deep Match Kernel (DMK)[3]	96.8 ± 1.5
Deep SetNet (ours)	**97.0 ± 2.7**

Conclusion

- We propose the first end-to-end graph convolutional networks, the Deep SetNet, for image-set classification.
- Our model can match the image set in an efficient way.
- Extensive experiments and analysis show the great effectiveness of our model for image-set classification.