PaperID: 3170

Adaptive Attention Graph Capsule Network

Xiangping Zheng, Xun Liang, Bo Wu, Yuhui Guo, Hui Tang

School of Information, Renmin University of China, Beijing, China 100872

{wubochn, xliang, xpzheng, yhguo, huitang}@ruc.edu.cn

Backgraound

From the perspective of the spatial domain, Graph Convolutional Network (GCN) is essentially a process of iteratively aggregating neighbor nodes.

Problems:Current methods using simple average or sum aggregation may neglect the characteristics of each node and topology between nodes, resulting in a large amount of early-stage information lost during the graph convolution step.

Approach

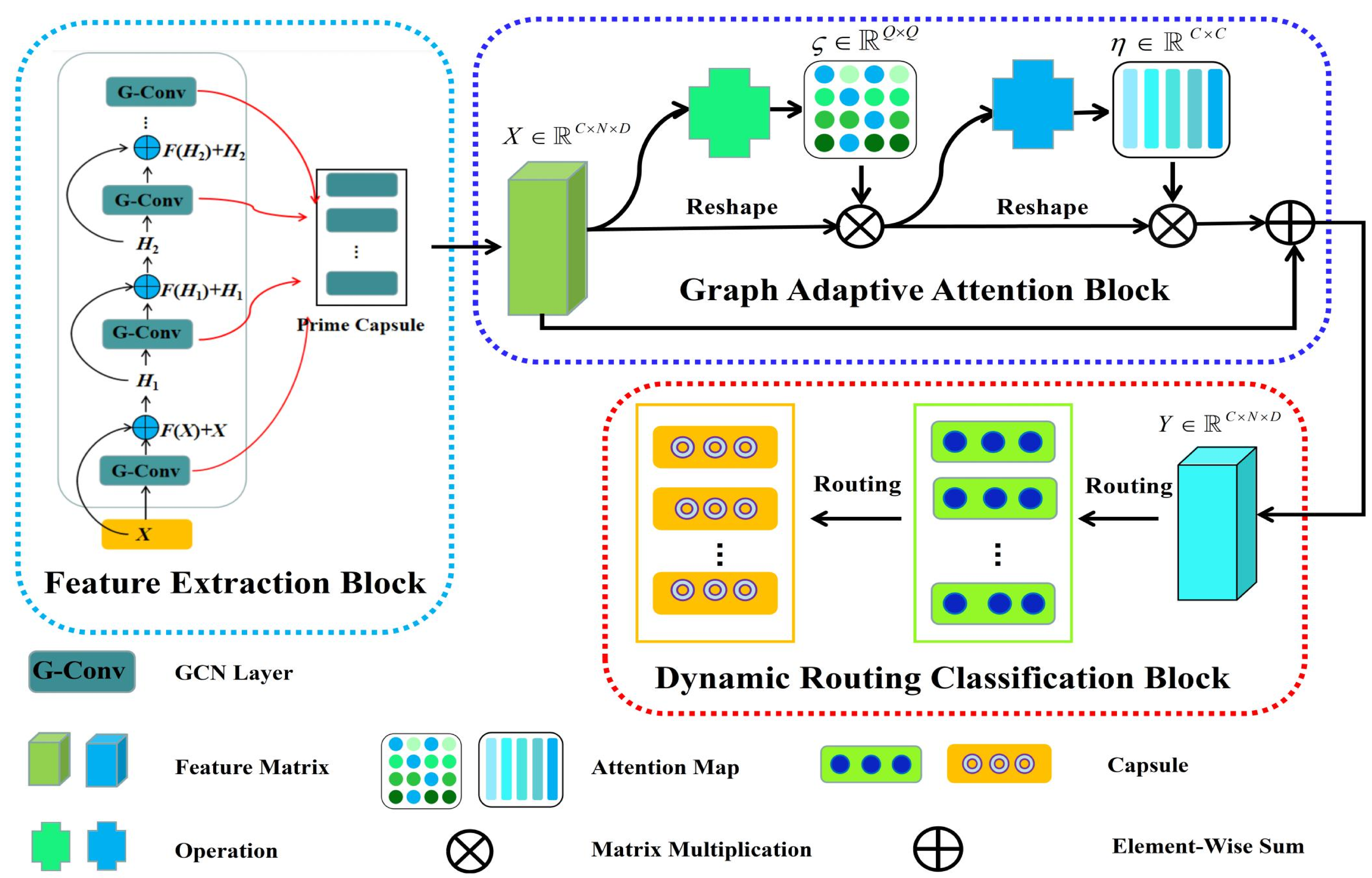
Our AA-GCN model comprise three blocks: feature extraction, graph adaptive attention mechanism, and capsule dynamic routing classification.

Feature Extraction We present an attention mechanism combined with graph propagation and capsules to generate capsule nodes to capture moregraph attribute features.

Adaptive Attention We present a graph adaptive attention mechanism to make the model pay more attention to the related nodes and spatial domain information of some supernodes.

Dynamic Routing We designed this block for graph classification.The primary capsule outputs to the capsule layer through the graph adaptive attention mechanism module

Framework



Main Result

Table 2. Graph classification accuracy (%) of our method compared with state-of-the-art graph classification methods on Biological Datasets. The top-2 performers on each dataset are shown in **bold**.

Methods	Biological MUTAG	PTC	ENZYMES	PROTEINS	Social Network COLLAB	IMDB-B	IMDB-M
GK [13]	81.58 ± 0.36	57.26 ± 1.41	26.61 ± 0.99	71.67 ± 0.55	72.84 ± 0.28	65.87 ± 0.98	43.89 ± 0.38
WL [14]	82.05 ± 0.36	57.97 ± 0.49	52.22 ± 1.26	74.68 ± 0.49	79.02 ± 1.77	73.40 ± 4.63	49.33 ± 4.75
DGK [15]	87.44 ± 2.72	60.08 ± 2.55	53.43 ± 0.91	75.68 ± 0.54	73.09 ± 0.25	66.96 ± 0.56	44.55 ± 0.52
AWE [16]	-	-	35.77 ± 0.36	-	73.90 ± 1.90	74.50 ± 5.80	51.50 ± 3.60
WEGL [17]	-	-	60.00 ± 6.30	76.50 ± 4.25	80.60 ± 2.00	75.40 ± 5.00	52.30 ± 2.90
GMT [18]	83.44 ± 1.33	-	-	75.09 ± 0.59	80.74 ± 0.54	73.48 ± 0.76	50.66 ± 0.82
GraphNorm [19]	91.60 ± 6.50	64.90 ± 7.50	-	$\mathbf{77.40 \pm 4.90}$	80.20 ± 1.00	76.00 ± 3.70	-
ASAP [20]	77.83 ± 1.49	-	-	73.92 ± 0.63	78.64 ± 0.50	72.81 ± 0.50	50.78 ± 0.75
DGCNN [21]	85.83 ± 1.66	58.59 ± 2.47	51.00 ± 7.29	75.54 ± 0.94	73.76 ± 0.49	70.03 ± 0.86	47.83 ± 0.85
GCN [9]	87.20 ± 5.11	-	66.50 ± 6.91	75.65 ± 3.24	81.72 ± 1.64	73.30 ± 5.29	51.20 ± 5.13
GIN [22]	89.40 ± 5.60	64.60 ± 7.00	-	76.20 ± 2.80	80.20 ± 1.90	75.10 ± 5.10	52.30 ± 2.80
GFN [23]	90.84 ± 7.22	-	70.17 ± 5.58	76.46 ± 4.06	81.50 ± 2.42	73.00 ± 4.35	51.80 ± 5.16
GCAPS-CNN [24]	-	66.01 ± 5.91	61.83 ± 5.39	76.40 ± 4.17	77.71 ± 2.51	71.69 ± 3.40	48.50 ± 4.10
CapsGNN [25]	86.67 ± 6.88	-	54.67 ± 5.67	76.28 ± 3.63	79.62 ± 0.91	73.10 ± 4.83	50.27 ± 2.65
Basic-AAGCN	92.49 ± 3.36	$\mathbf{66.63 \pm 4.23}$	69.25 ± 4.23	77.14 ± 2.68	81.82 ± 1.25	$\mathbf{76.84 \pm 4.12}$	51.68 ± 2.36
Plain-AAGCN	93.24 ± 4.10	65.52 ± 3.54	71.14 ± 5.16	77.28 ± 4.85	80.75 ± 2.21	75.61 ± 2.76	51.96 ± 4.10
Dense-AAGCN	$\mathbf{93.85 \pm 4.23}$	66.54 ± 0.36	$\mathbf{72.12 \pm 5.12}$	77.21 ± 3.36	$\mathbf{82.92 \pm 2.62}$	75.82 ± 3.25	$\mathbf{52.63 \pm 3.28}$
Res-AAGCN	$\mathbf{94.01 \pm 3.16}$	$\mathbf{67.02 \pm 6.31}$	$\mathbf{72.35 \pm 5.85}$	$\mathbf{78.20 \pm 4.12}$	$\mathbf{83.17 \pm 1.85}$	$\mathbf{77.23 \pm 3.57}$	$\mathbf{52.94 \pm 4.10}$

GCN Propagation

Table 1. Four different GCN propagation formulas

Methods	Propagation formula
Basic-GCN	$\mathbf{H}^{(l+1)} = \sigma\left(\hat{\mathbf{A}}\mathbf{H}^{(l)}\mathbf{W}^{(l)}\right)$
Plain-GCN	$\mathbf{H}^{(l+1)} = \sigma\left(\hat{\mathbf{A}}\mathbf{H}^{(l)}\mathbf{W}^{(l)}\right) + \mathbf{X}$
Res-GCN	$\mathbf{H}^{(l+l)} = \sigma\left(\hat{\mathbf{A}}\mathbf{H}^{(l)}\mathbf{W}^{(l)}\right) + \mathbf{H}^{(l)}$
Dense-GCN	$\mathbf{H}^{(l+1)} = \sigma\left(\hat{\mathbf{A}}\mathbf{H}^{(l)}\mathbf{W}^{(l)}\right) + \sum_{i=1}^{l}\mathbf{H}^{i} + \mathbf{X}$

Ablation

Table 3. Ablation study on **COLLAB** dataset.

Basic-AAGCN			Plain-AAGCN		
NAM	**LAM**	**Accuracy**	**NAM**	**LAM**	**Accuracy**
		77.72%			77.15%
✓		81.53%	✓		80.28%
	✓	80.62%		✓	79.45%
✓	✓	81.82%	✓	✓	80.75%
Res-AAGCN			**Dense-AAGCN**		
NAM	**LAM**	**Accuracy**	**NAM**	**LAM**	**Accuracy**
		79.66%			79.43%
✓		82.28%	✓		81.88%
	✓	81.56%		✓	81.28%
✓	✓	83.17%	✓	✓	82.92%

Conclusions

In this work, we propose acapsule graph neural network based on an adaptive attention mechanism to tackle the potential loss of relevant node information caused by graph convolution operations. Experiments show that our model can learn better feature representation for graph classification .

Acknowledgements

This work was supported by the National Natural ScienceFoundation of China (62072463, 71531012), and the National Social Science Foundation of China (18ZDA309).