Beijing Normal University
School of Artificial Intelligence

MultiNet: Adaptive Multi-Viewed Subgraph Convolutional Networks for Graph Classification

Xinya Qin, Lu Bai, Lixin Cui, Ming Li,

Hangyuan Du, Edwin R. Hancock

XinyaQin@mail.bnu.edu.cn, bailu@bnu.edu.cn

Speaker: **Xinya Qin** Affiliation: **School of Artificial Intelligence, Beijing Normal University**

北京師範大學
BEIJING NORMAL UNIVERSITY

NEURAL INFORMATION PROCESSING SYSTEMS

MultiNet: Adaptive Multi-Viewed Subgraph Convolutional Networks for Graph Classification

- Motivation
- Contributions
- Methodology
- Experimental Results

Motivation

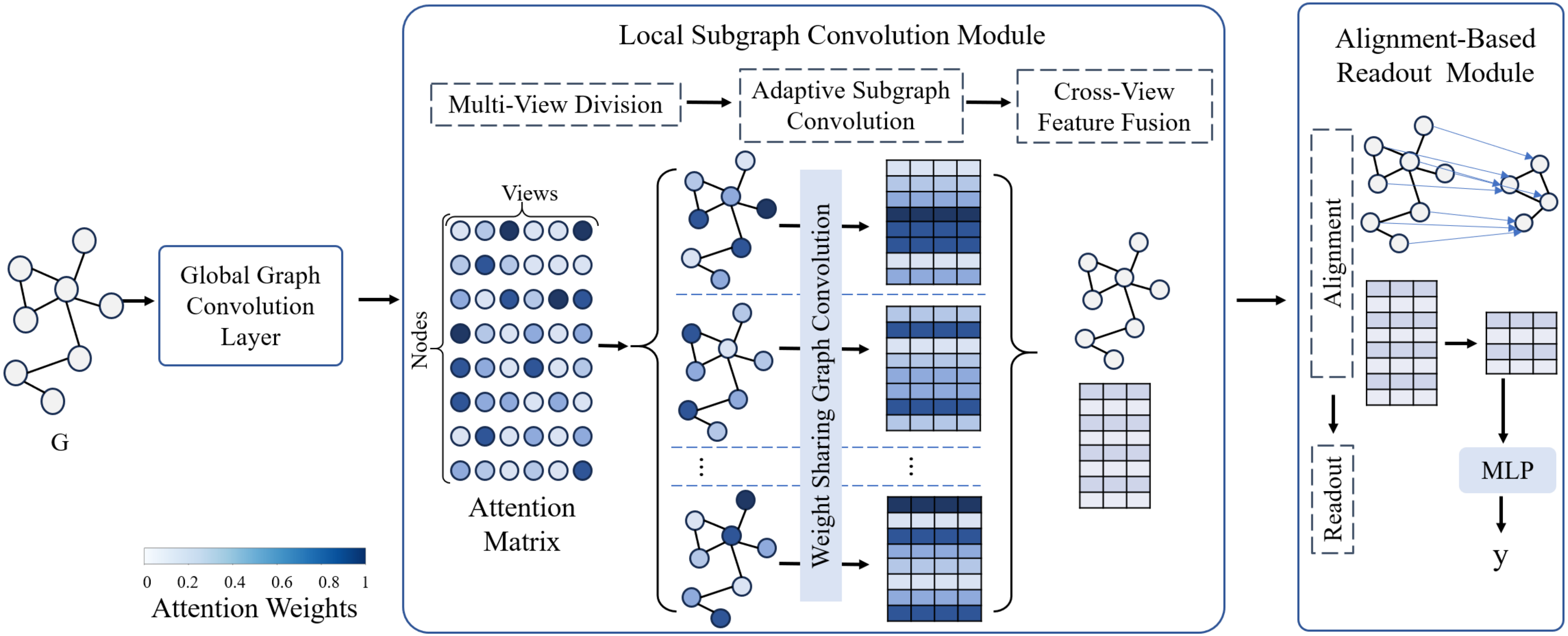
- Over-smoothing in GCNs → node representations become indistinguishable.
- New observation: graph-level over-smoothing.
  - Node representations become indistinguishable.
  - Global readouts like sum or mean ignore local structure differences.
- Preserve local discriminative information and generate more distinct graph-level representations.

Contributions

- Adaptive Multi-Viewed Subgraph Convolutional Networks (MultiNet)
  - Local Subgraph Convolution:
    - Adaptively divides graphs into multiple views to capture discriminative local structures.
  - Alignment-Based Readout:
    - Establishes consistent cluster-level correspondences to preserve structural information.
  - Theoretical & Empirical Validation:
    - Formally defines and validates the mitigation of graph-level over-smoothing.

Methodology

- Adaptive Multi-Viewed Subgraph Convolutional Networks (MultiNet)



Methodology

- Local Subgraph Convolution Module
  - Multi-Viewed Subgraph Division：

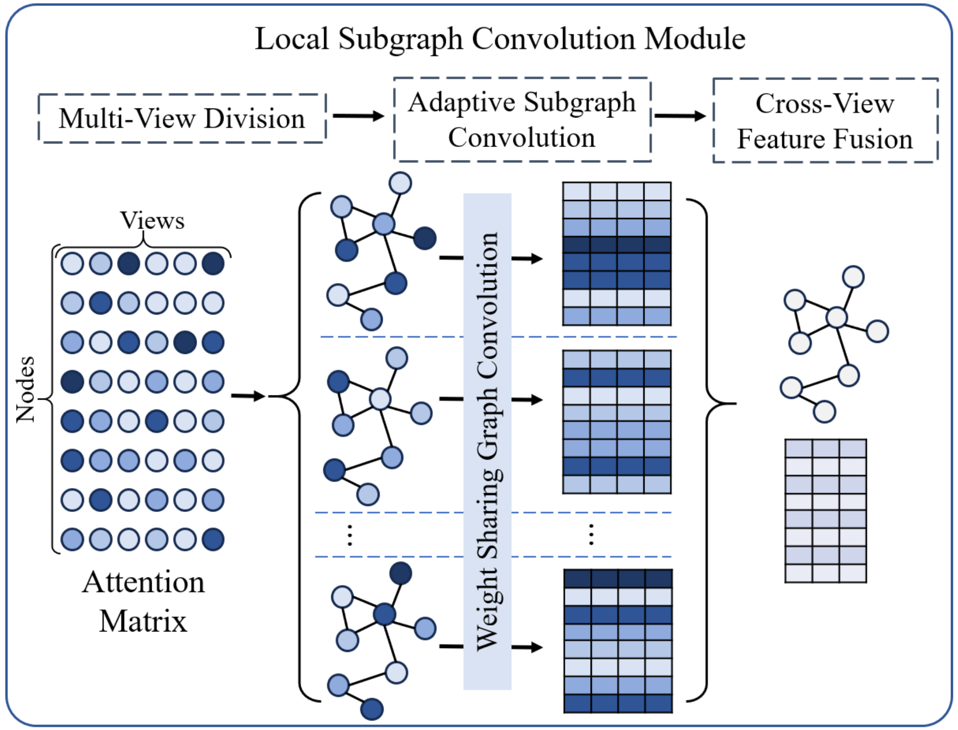
$$P = \mathrm{softmax}(\tilde{A}X^{(0)}W_P),$$

  - Adaptive Subgraph Convolution Operation:

$$X_j^{(l+1)} = \sigma\left((\tilde{A}X_j^{(l)}) \odot P_j W^{(l)}\right)$$

  - Cross-Viewed Feature Fusion

$$X^{(L)} = \mathrm{MLP}([X_1^{(L)} \parallel X_2^{(L)} \parallel \cdots \parallel X_m^{(L)}])$$



Methodology

- Alignment-based Readout
  - Learn a shared cluster assignment:

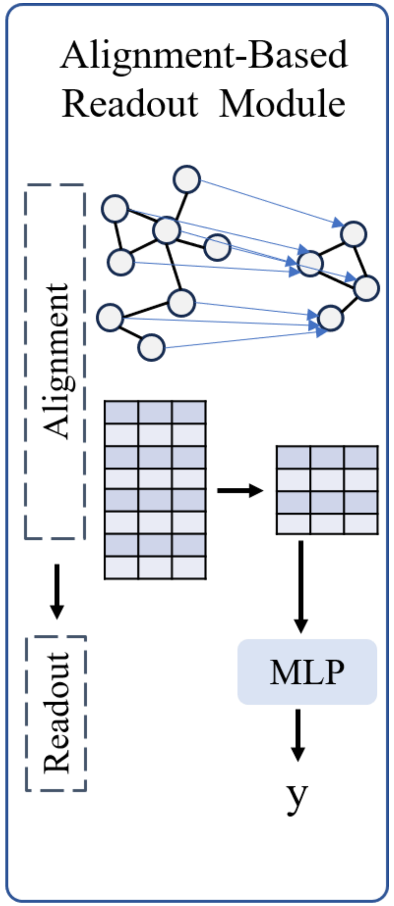
$$S = softmax(\tilde{A}X^{(L)}W_S)$$

  - Align the node features:

$$\tilde{X} = S^\top X^{(L)}$$

  - Apply an MLP for expressive readout:

$$y = \mathrm{MLP}(\tilde{X})$$



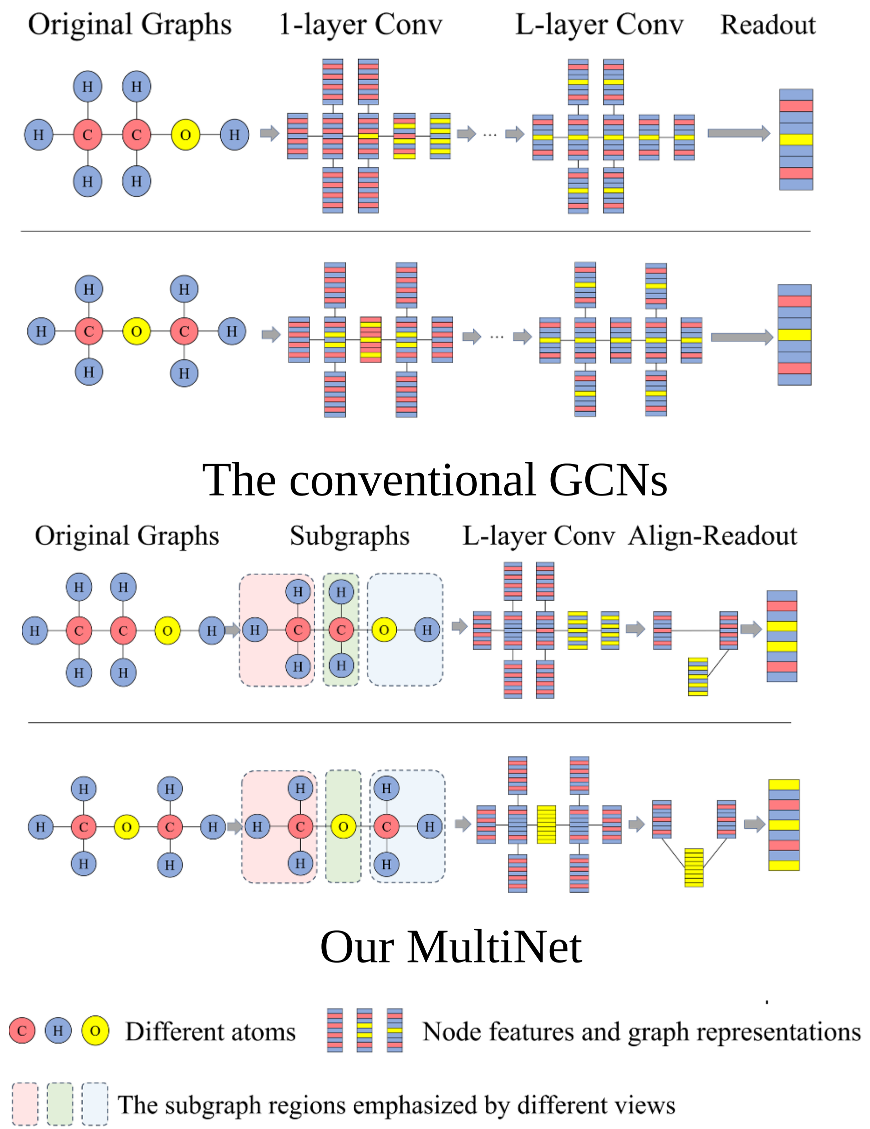
Methodology

- Theoretical Analysis
  - Constrained propagation
    - view-specific propagation kernel

$$\tilde{A}_j = \mathrm{diag}(P_j)\,\tilde{A}\,\mathrm{diag}(P_j)$$

  - Spectral radius $\rho(\tilde{A}_j) < \rho(\tilde{A}) = 1$
    - limits diffusion, prevents convergence to the dominant eigenvector

$\rightarrow$ mitigates over-smoothing.



Experimental Results

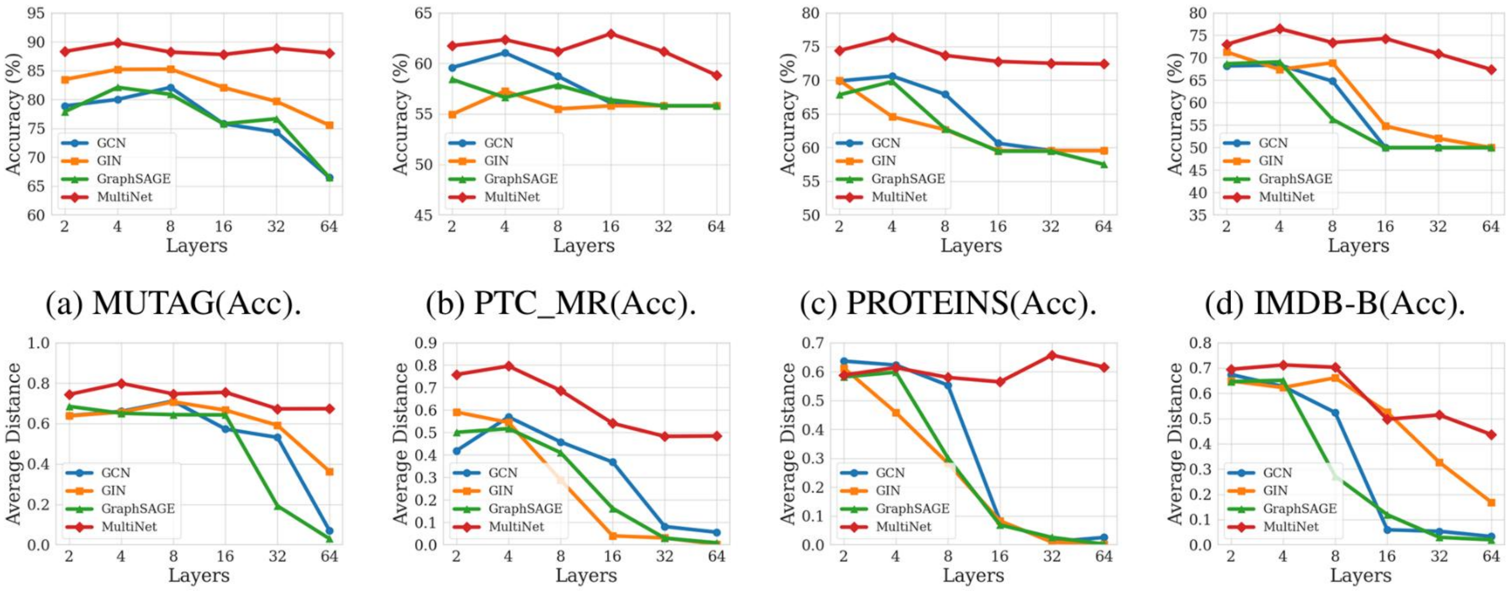
- Experiments on Graph Classification

Table 1: Classification accuracy (In % ± standard error) on benchmark datasets.

Method	MUTAG	PTC_MR	ENZYMES	PROTEINS	DD	IMDB-B	IMDB-M	Rank
DGCNN	84.0±6.7	58.3±7.0	38.9±5.7	72.9±3.5	76.6±4.3	69.2±3.0	45.6±3.4	12.40
DiffPool	79.8±7.1	60.8±7.0	59.5±5.6	73.7±3.5	75.0±3.5	68.4±3.3	45.6±3.4	10.80
ECC	75.4±6.2	55.7±3.3	29.5±8.2	72.3±3.4	72.69±4.1	67.7±2.8	43.5±3.1	15.60
GIN	84.7±6.7	58.8±5.5	**59.6±4.5**	73.3±4.0	75.3±2.9	71.2±3.9	48.5±3.3	7.40
GraphSAGE	83.6±9.6	60.1±4.7	58.2±6.0	73.0±4.5	72.9±2.0	68.8±4.5	47.6±3.5	10.80
DGK	82.66±1.45	57.32±1.13	53.4±0.9	71.68±0.50	78.50±0.22	66.96±0.56	44.55±0.52	10.57
1-RWNN	89.2±4.3	—	56.7±5.2	74.7±3.3	77.6±4.7	70.8±4.8	47.8±3.8	6.33
2-RWNN	88.1±4.8	—	57.4±4.9	74.1±2.8	76.9±4.6	70.6±4.4	48.8±2.9	7.16
3-RWNN	88.6±4.1	—	57.6±6.3	74.3±3.3	77.4±4.9	70.7±3.9	47.8±3.5	6.50
GKNN-WL	85.73±2.70	59.29±2.54	—	74.94±1.10	—	69.70±2.20	47.87±1.78	7.00
GKNN-GL	85.24±2.28	60.13 ±1.94	—	75.36±1.12	—	69.90±2.20	45.67±1.22	7.40
RWGK	80.77±0.72	55.91±0.37	22.37±0.35	74.20±0.40	71.70±0.47	67.94±0.77	46.72±0.30	13.14
SPGK	83.38±0.31	56.55±0.53	29.00±0.48	75.10±0.50	78.45±0.26	71.26±1.04	51.33±0.57	6.71
GK	81.66±0.11	—	24.87±0.22	71.67±0.55	78.45±0.26	65.87±0.98	45.42±0.87	14.17
WLSK	82.88±0.57	56.05±0.51	52.75±0.44	73.52±0.43	**79.78±0.36**	71.88±0.77	49.50±0.49	7.14
JTQK	85.50±0.55	57.39±0.46	56.41±0.42	72.86±0.41	79.49±0.32	72.45±0.81	50.33±0.49	6.00
ASK	87.50±0.65	—	—	—	70.38±0.22	—	50.12±0.51	9.33
EDBMK	86.35	56.75	36.85	—	78.19	—	—	8.25
QBMK	88.55±0.43	59.38±0.36	—	—	77.60±0.47	—	—	5.00
MultiNet	**89.81±1.46**	**62.65±0.88**	54.83±1.55	**76.40±0.87**	78.90±0.51	**76.49±0.60**	**51.93±0.25**	**2.28**

Experimental Results

- Evaluations on Mitigating the Over-Smoothing
  - Average Cosine Distance (AD): $\mu(H) = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(1 - \frac{H_i^\top H_j}{\|H_i\|\|H_j\|}\right)$



(a) MUTAG(Acc). (b) PTC_MR(Acc). (c) PROTEINS(Acc). (d) IMDB-B(Acc).

(e) MUTAG(AD). (f) PTC_MR(AD). (g) PROTEINS(AD). (h) IMDB-B(AD).

Figure 4: Classification accuracies (%) and the AD values on the four datasets.

Thanks For Your Listening

Contact us at: XinyaQin@mail.bnu.edu.cn, bailu@bnu.edu.cn,