M.Sc Defence Thesis Presentation

# Empowering Graph Wavelet Convolution for Node Classification: A Novel Approach with Local Lifting Scheme

RME 5401: Thesis

Swakshar Deb
Exam Roll: 120706
Registration No. 2016-614-703

Supervisor
Dr. Sejuti Rahman
Associate Professor
Robotics and Mechatronics Engineering
University of Dhaka

# Overview

## Part 1

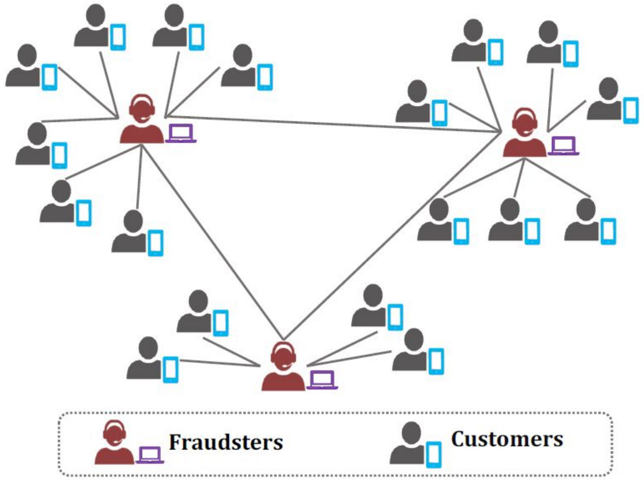
- Motivation
- Homophilic and Heterophilic graphs
- Problem Statement
- Related Works
- Proposed Generalized Adaptive Graph Wavelet Neural Network (GA-GWNN)
- Limitations and our Contributions
- Experimental Analysis

## Part 2

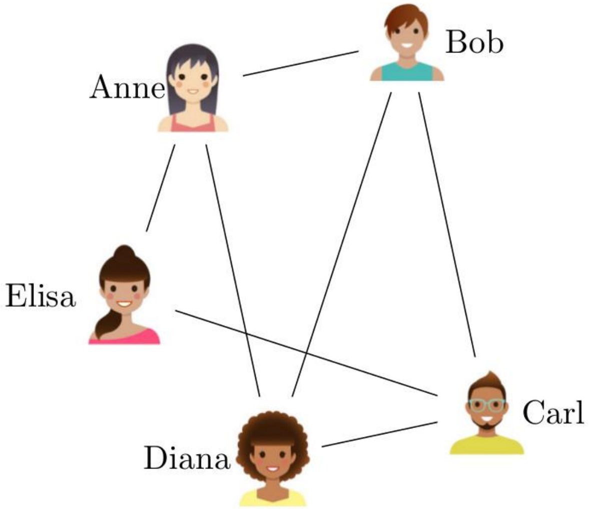
- Simple and Effective Graph Wavelet Neural Network (SEA-GWNN)
- Experimental Analysis
- Conclusion

# Motivation for Node Classification

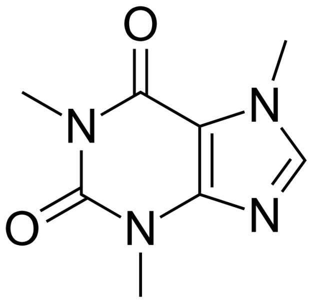
Fraudster Detection




Friend Suggestion

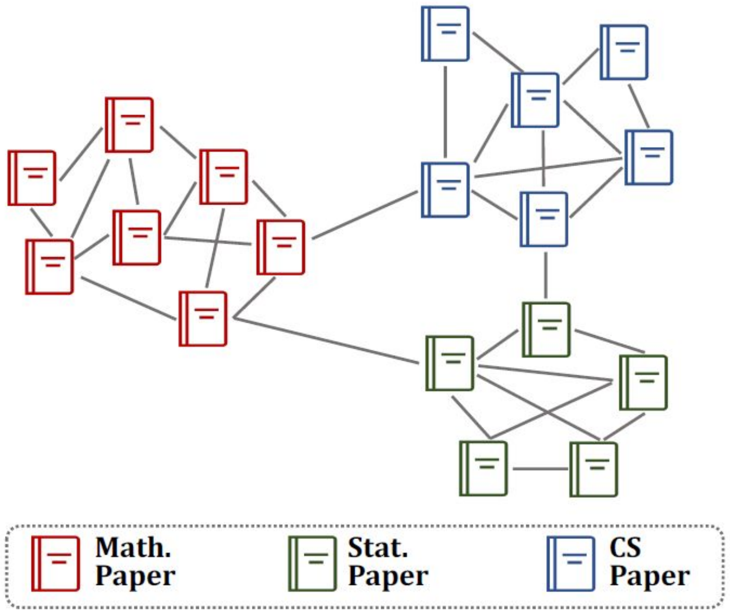



Drug Discovery

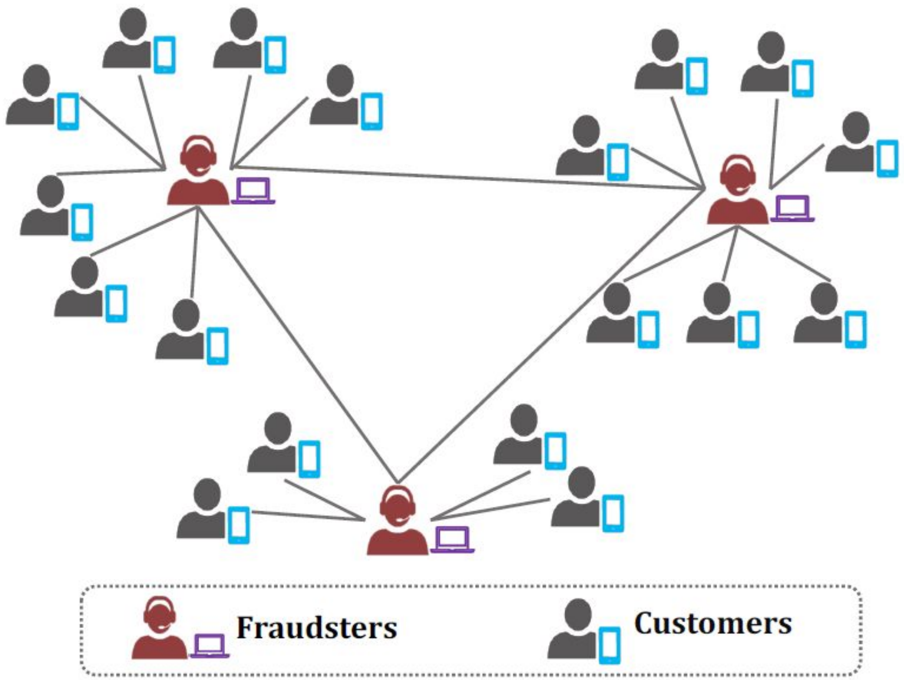


# Homophilic and Heterophilic Graphs

Homophilic Graph



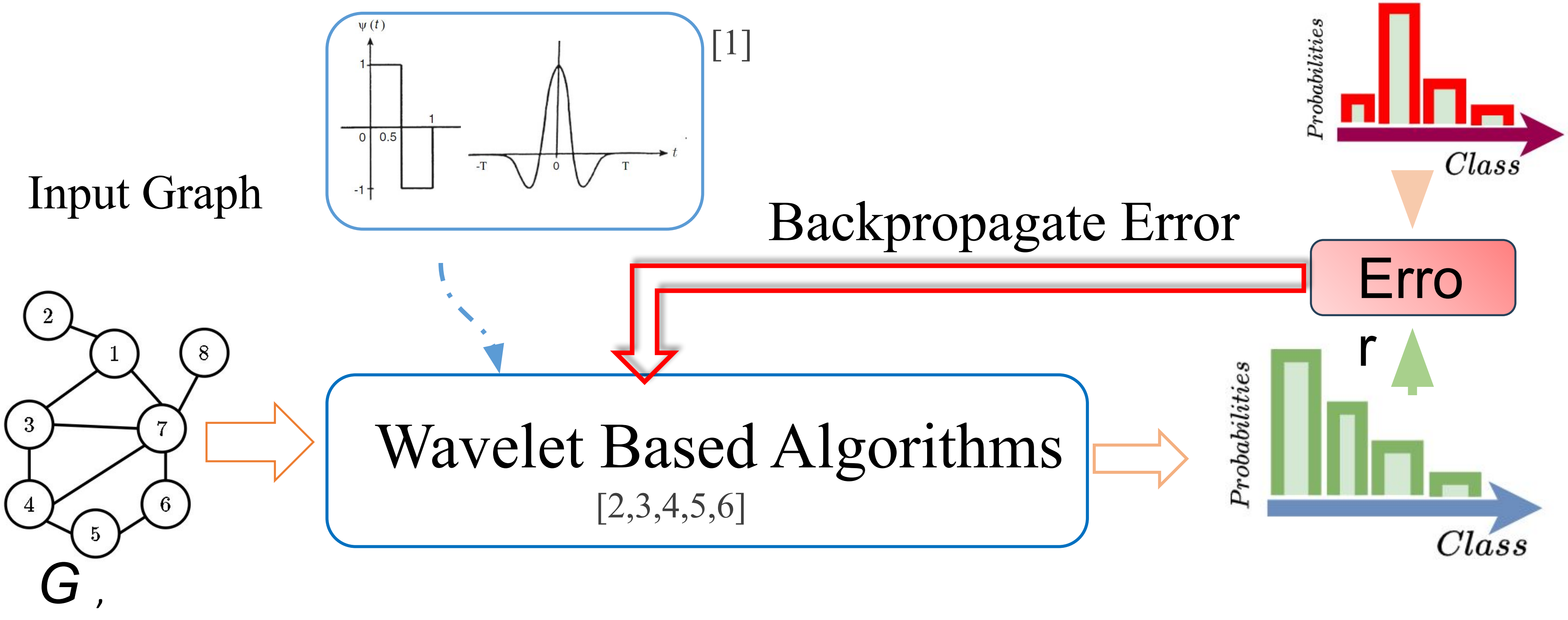

Heterophilic Graph




Due to similar neighbours, homophilic graphs are suitable for low pass filtering

Due to dissimilar neighbour, it is necessary to recognise high frequency pattern via high pass filter

# Problem Statement





[1]Mallat et. al., “A wavelet Tour of Signal Processing”,Elsevier,1999

[2]Xu et. al., “Graph wavelet neural network,” ICLR, 2018

[3]Zheng et. al., “How framelets enhance graph neural networks,“ ICLR, 2021

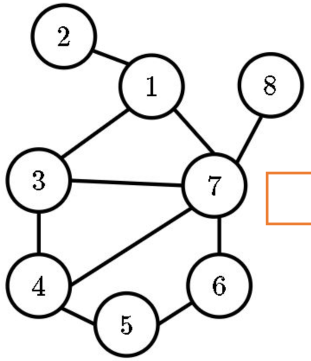
[4] Zheng et. al. , “Mathnet: Haarlike wavelet multiresolution analysis for graph representation learning,“ Knowledge-Based Systems, 2023.

[5]Li et. al., “Fast haar transforms for graph neural networks,” Neural Networks, 2020.

[6]Xu et. al., “Graph neural networks with lifting-based adaptive graph wavelets,” IEEE Transactions on Signal and Information Processing over Networks, 2022

# Related Work

Input Graph




$G$ , $\mathbf{X}$

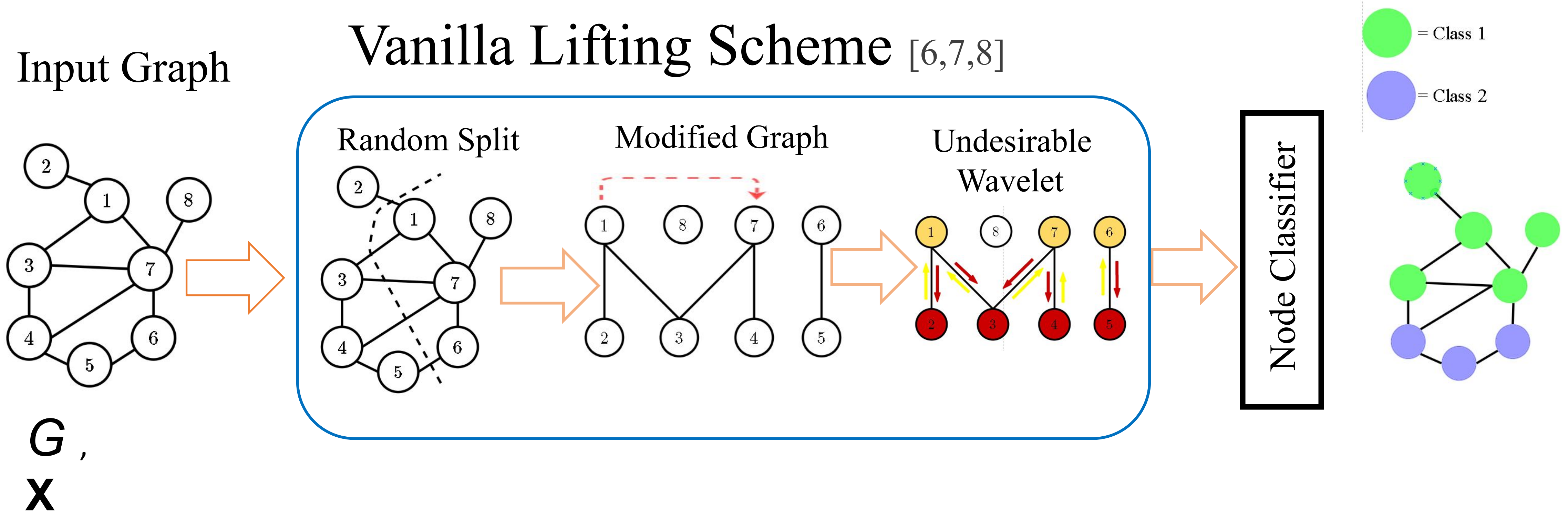
Wavelet Based Algorithms
[2,3,4,5,6,7,8]

Node Classification

- Predefined Wavelet based Approaches [2,3,4,5]
  - Utilize fix wavelet filters
  - Require domain specific knowledge
  - Data preprocessing step is mandatory
- Adaptive Wavelet based Approaches [6,7,8]
  - Lead to undesirable wavelet filters

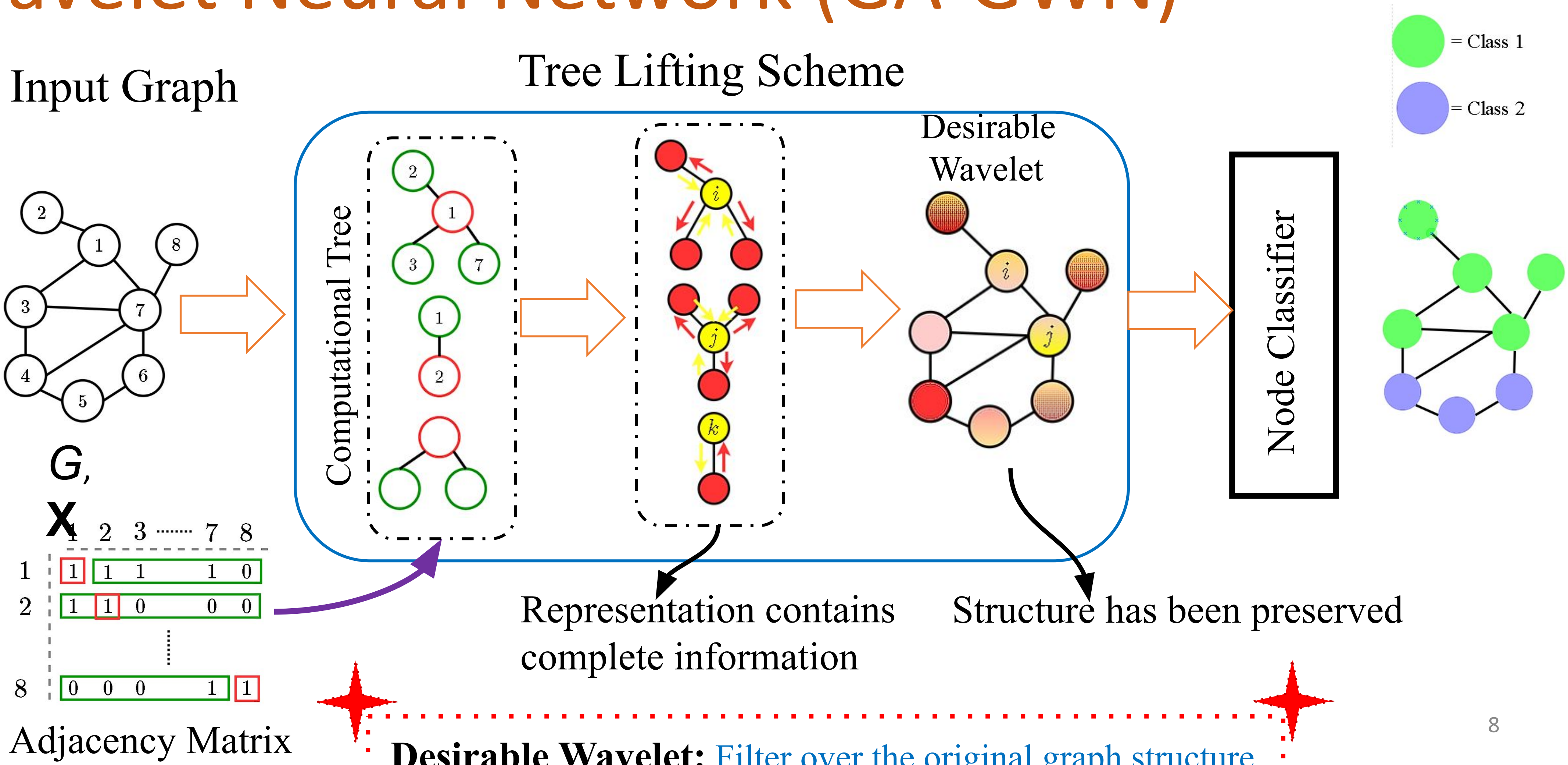

[7]Shen et. al., “Optimized distributed 2d transforms for irregularly sampled sensor network grids using wavelet lifting," in IEEE ICASS, 2008
[8]Narang et. al., Lifting based wavelet transforms on graphs,“ APSIPA, 2009.

# Related Works: Adaptive Wavelet based Approach




**Undesirable Wavelet:** A Filter produced over a disrupted graph structure

# Proposed Generalized Adaptive Graph Wavelet Neural Network (GA-GWN)




**Desirable Wavelet:** Filter over the original graph structure

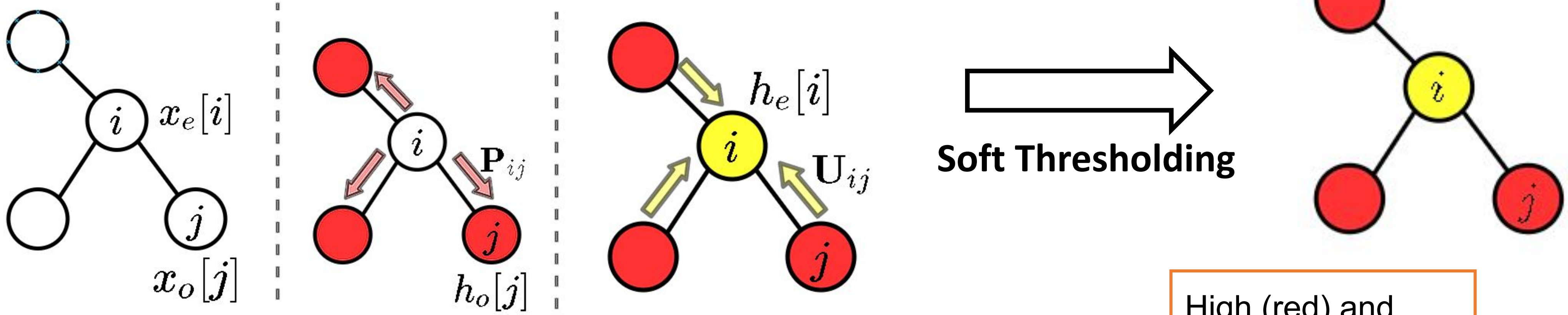
# Our Contributions

- Limitations of existing wavelet based methods
  - Predefined wavelet filter based methods consider only homophily assumption.
  - Predefined wavelet filter based methods require domain specific knowledge.
  - Adaptive wavelet filter based methods produces undesirable filters.

- Contributions
  - Our Proposed GA-GWNN can generalize to both homophilic and heterophilic graphs.
  - Since proposed GA-GWNN is adaptive wavelet based approach thus does not require domain specific knowledge.
  - GA-GWNN is able to produce desirable wavelet filters.
  - Also proposed further simple and effective version SEA-GWNN.
    - No need of inverse transform.
    - Attention detachment.
    - Multiscale information.

# Message Propagation of GA-GWNN

Theorem 4.1 $h_e[j] = x_e[j] + \mathbf{U}(h_o)[j]$




Tree, T_i

High Frequency

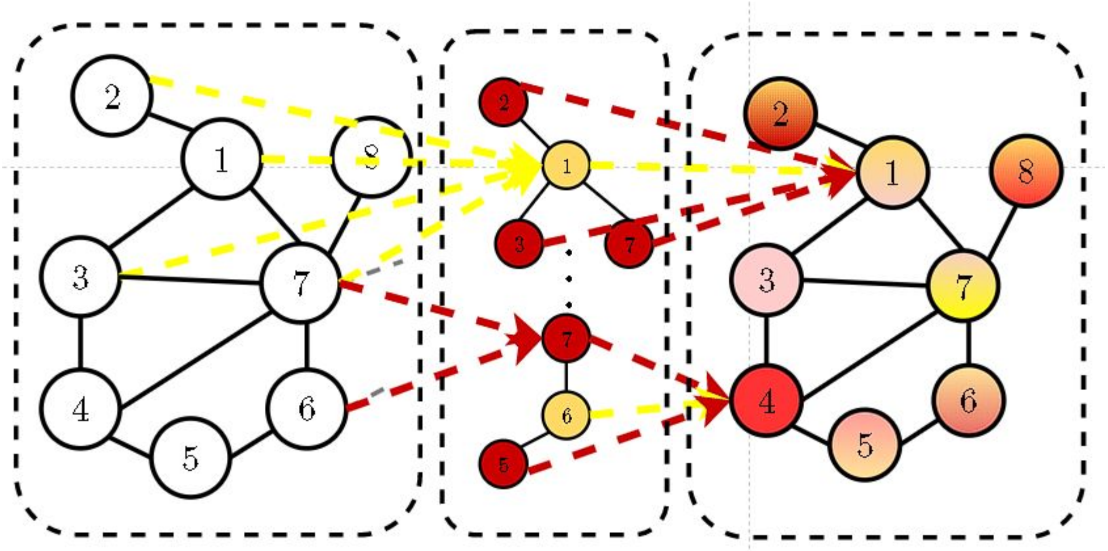
Low Frequency

High (red) and Low (yellow) pass filter.

$h_o[i] = x_o[i] - \mathbf{P}(x_e)[i]$ Theorem 4.1

# Fusion Module/Operation

- Fuses both the low and high frequency information to enhance representation




Input Graph

Learned wavelet through lifting

Fusion module

# Dataset and Evaluation Metrics

- Homophilic dataset
  - Citation graph[24] (Cora, Citeseer, PubMed)
- Heterophilic dataset
  - Webpage graph[13] (Cornell, Texas, Wisconsin)
  - Film industry graph[13] (Film)
- Large scale graphs
  - Ogbn-Arxiv[19]
  - Ogbn-Products[19]

- Evaluation Metrics
  - Accuracy
    - Percentage correct prediction
  - Precision
  - Recall
  - F1-Score
    - Harmonic mean of precision and recall

Dataset Statistics

| Dataset | Classes | Nodes | Edges | Features | Hom. Ratio |
|---|---|---|---|---|---|
| Cora | 7 | 2,708 | 5,429 | 1,433 | 0.81 |
| Citeseer | 6 | 3,327 | 4,732 | 3,703 | 0.74 |
| Pubmed | 3 | 19,717 | 44,338 | 500 | 0.80 |
| Film | 5 | 7,600 | 33,544 | 931 | 0.22 |
| Cornell | 5 | 183 | 295 | 1,703 | 0.30 |
| Texas | 5 | 183 | 309 | 1,703 | 0.11 |
| Wisconsin | 5 | 251 | 499 | 1,703 | 0.21 |
| Ogbn-Arxiv | 40 | 169,343 | 1,166,243 | 128 | 0.66 |
| Ogbn-Products | 47 | 2,449,029 | 61,859,140 | 100 | 0.81 |

# Overall Results

Mean accuracy on semi supervised node classification. Best results are heighted in bold

| Methods | Cora | Citeseer | PubMed |
|---|---|---|---|
| GraphSAGE[9] | 74.5 ± 0.8 | 67.2 ± 1.0 | 76.8 ± 0.6 |
| GAT[10] | 83.0 ± 0.7 | 72.5 ± 0.7 | 79.0 ± 0.3 |
| HANet[5] | 81.9 | 70.1 | 79.3 |
| GWNN[2] | 81.6 ± 0.7 | 70.5 ± 0.6 | 78.6 ± 0.3 |
| UFGConvS[3] | 83.0 ± 0.5 | 71.0 ± 0.6 | 79.4 ±0.4 |
| UFGConvR[3] | 83.6 ± 0.6 | 72.7 ± 0.6 | 79.9 ± 0.1 |
| LGWNN[6] | 83.4 ± 0.6 | 71.1 ± 0.4 | 79.5 ± 0.5 |
| GA-GWNN | **84.4** ± **0.2** | **73.1** ± **0.3** | **80.1** ±**0.2** |

Predefined Wavelet Filter

Adaptive Wavelet method but undesirable filter

Adaptive wavelet method with desirable filters

[9]Hamilton et. al., "Inductive representation learning on large graphs,"Neurips, 2017
[10] Velickovic et. al., " Graph attention networks," ICLR, 2018.

# Overall Results(3)

Mean Precision, Recall, F1-score on Semi supervised node classification

| Method | Precision | | | Recall | | | F1 Score | | |
|---|---|---|---|---|---|---|---|---|---|
| | Cora | Citeseer | Pubmed | Cora | Citeseer | PubMed | Cora | Citeseer | PubMed |
| GRAPHSAGE[9] | 75.77±0.1 | 60.68±0.2 | 77.13±0.6 | 79.78±1.5 | 62.70±0.5 | 75.16±0.6 | 78.88±0.7 | 61.44 ±0.6 | 76.44 ±0.6 |
| GAT[10] | 75.79±0.1 | 63.68±0.2 | 79.13±0.6 | 79.78±1.5 | 64.70±0.5 | 75.16±0.6 | 79.88±0.7 | 63.44 ±0.6 | 74.44 ±0.6 |
| HANET[5] | 75.77±0.1 | 60.68±0.2 | 77.13±0.6 | 79.78±1.5 | 62.70±0.5 | 75.16±0.6 | 78.88±0.7 | 61.44 ±0.6 | 76.44 ±0.6 |
| GWNN[2] | 77.31 ±0.2 | 66.93 ±0.1 | 76.14 ±0.1 | 79.62 ±1.4 | 61.89 ±0.6 | 79.00 ±0.5 | 78.04 ±0.7 | 63.41 ±0.5 | 77.41 ±0.5 |
| UFGCONVS[3] | 79.08±1.8 | 67.59±2.6 | 80.04±0.5 | 78.74±0.6 | 68.12±9.1 | 78.19±3.4 | 66.51±5.1 | 66.99±5.1 | 79.75±5.1 |
| UFGCONVR[3] | 80.01±1.1 | 67.16±1.9 | 79.41±0.6 | 78.39±1.5 | 68.30±8.4 | 77.22±6.7 | 80.55±5.2 | 66.47±3.6 | 79.47±3.6 |
| LGWNN[6] | 80.21±1.3 | 67.01±1.9 | 78.78±0.6 | 78.60±1.1 | 68.38±4.9 | 77.00±7.2 | 80.45±4.9 | 66.51±3.1 | 79.40±3.1 |
| GA-GWNN | **80.32**±0.2 | **68.30**±0.1 | **80.25**±0.1 | **83.00**±1.8 | **68.16**±4.6 | 79.45±3.5 | **81.43**±3.3 | **67.08**±1.9 | **80.08**±1.9 |

# Overall Results(2)

Mean accuracy on full supervised node classification. Best results are in bold

| Method/ Acc.(%) | homophilic graph datasets | | | heterophilic graph datasets | | | | |
|---|---|---|---|---|---|---|---|---|
| | Cora | Cite. | Pubm. | Film | Corn. | Texa. | Wisc. | Avg. Acc. |
| GCN[11] | 85.77±0.1 | 73.68±0.2 | 88.13±0.6 | 28.78±1.5 | 52.70±0.5 | 52.16±0.6 | 45.88±0.7 | 61.44 ±0.6 |
| GAT[10] | 86.37±0.2 | 74.32±0.2 | 87.62±0.4 | 28.99±1.4 | 54.32±0.3 | 58.38±0.5 | 49.41±0.9 | 62.77 ±0.5 |
| APPNP[11] | 87.87±0.2 | 76.53±0.2 | 89.40±1.4 | 34.86 ±1.1 | 73.51±1.1 | 65.41±0.1 | 69.02±1.6 | 70.94 ±0.8 |
| GWNN[2] | 85.31 ±0.2 | 73.93 ±0.1 | 88.14 ±0.1 | 26.62 ±1.4 | 61.89 ±0.6 | 60.00 ±0.5 | 48.04 ±0.7 | 63.41 ±0.5 |
| DeepWalk[12] | 80.08±1.8 | 53.59±2.6 | 81.14±0.5 | 23.74±0.6 | 44.12±9.1 | 49.19±3.4 | 53.51±5.1 | 55.05±5.1 |
| G-GCN[13] | 84.91±1.1 | 75.16±1.9 | 88.41±0.6 | 32.39±1.5 | 55.68±8.4 | 66.22±6.7 | 62.55±5.2 | 66.47±3.6 |
| MixHop[14] | 87.61±0.8 | 76.26±1.3 | 85.31±0.6 | 32.22±2.3 | 73.51±6.3 | 77.84±7.7 | 75.88±4.9 | 72.66±3.4 |
| H2GCN[15] | 87.69±1.3 | 75.95±2.1 | 88.78±0.5 | 36.71±1.4 | 78.92±5.2 | 82.16±8.2 | 82.57±3.2 | 76.11±3.1 |
| CPGNN[16] | 87.18±1.1 | 75.52±1.8 | 89.08±0.6 | 35.51±1.8 | 63.51±5.8 | 74.32±7.3 | 81.76±6.7 | 72.44±3.6 |
| GPR-GNN[17] | 86.70±1.1 | 75.12±1.9 | 87.38±0.6 | 36.47±1.3 | 82.97±5.6 | **84.59**±4.3 | 83.92±3.1 | 76.73±2.5 |
| AM-GCN[18] | 86.66±1.3 | 76.01±1.9 | 86.78±0.6 | 33.60±1.1 | 78.38±4.9 | 78.38±7.2 | 81.76±4.9 | 74.51±3.1 |
| GA-GWNN | **88.32**±0.1 | **76.61**±0.2 | **89.70**±0.1 | **37.16**±1.8 | **85.16**±4.6 | 84.45±3.5 | **85.43**±3.3 | **78.08**±1.9 |

Homophilic Graph Neural Network

Heterophilic Graph Neural Network

Adaptive wavelet filters

# Overall Results(3)

Mean accuracy on large scale graphs (full supervised). Best results are highlighted in bold. OOM denotes *Out of Memory*

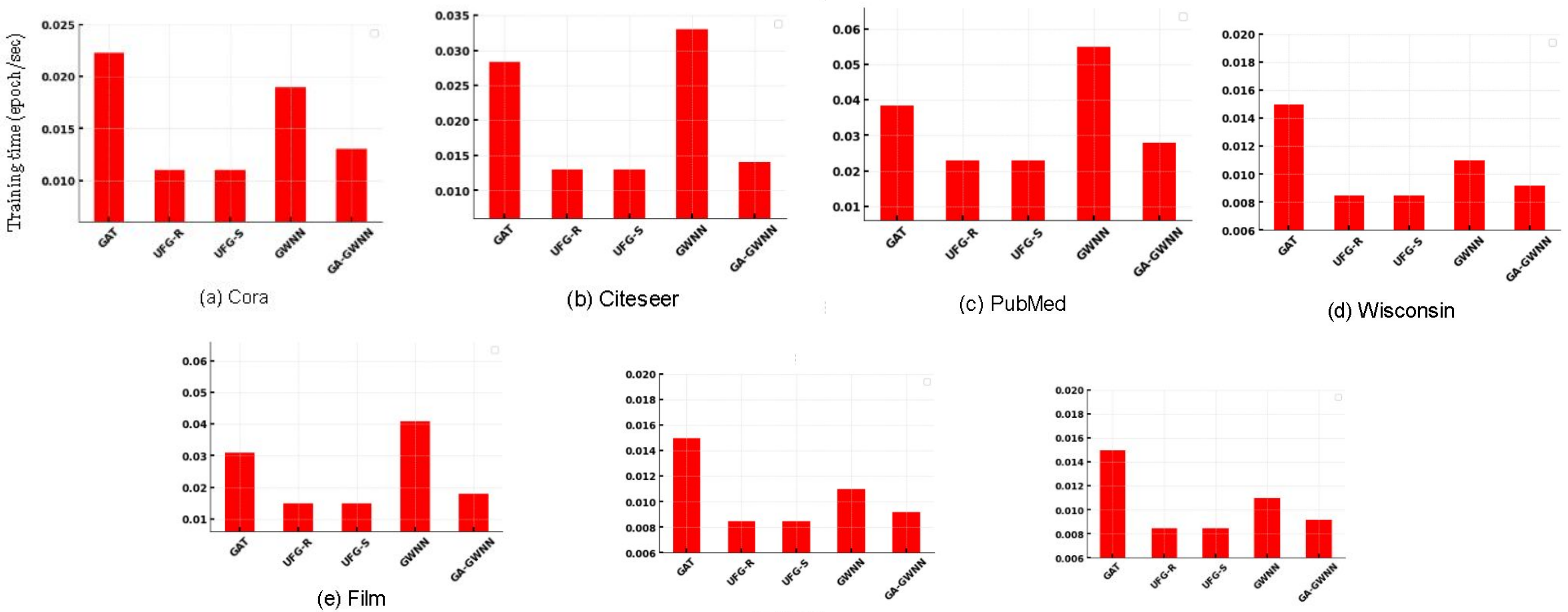
| Method/ Acc.(%) | Ogbn-Arxiv | | Ogbn-Products | |
|---|---|---|---|---|
| | Test Acc. | Val. Acc. | Test Acc. | Val. Acc. |
| MLP[19] | 55.50±0.23 | 57.65±0.12 | 61.06±0.08 | 75.54±0.02 |
| NODE2VEC[20] | 70.07±0.13 | 71.29±0.13 | 72.49±0.10 | 90.32±0.06 |
| GRAPHZOOM[21] | 71.18±0.18 | 72.20±0.07 | 74.06±0.26 | 90.66±0.11 |
| GRAPHSAGE[9] | 71.49±0.27 | 72.77±0.17 | 78.29±0.16 | 92.24±0.07 |
| GCN[11] | 71.74±0.29 | 73.00±0.17 | 75.64±0.21 | 92.00±0.03 |
| DEEPERGCN[12] | 71.92±0.17 | 72.62±0.14 | **80.98**±0.20 | 92.38±0.09 |
| SIGN[13] | 71.95±0.11 | 73.23±0.06 | 80.52±0.16 | **92.99**±0.04 |
| UFGCONV-S[3] | 70.04±0.22 | 71.04±0.11 | OOM | OOM |
| UFGCONV-R[3] | 71.97±0.12 | 73.21±0.05 | OOM | OOM |
| GA-GWNN | **72.27**±0.21 | **73.64**±0.35 | 80.91 ±0.18 | 92.30 ±0.15 |

Only utilize graph structure

Only utilize graph signal

More depth
More the cost

# Computational Cost




(a) Cora (b) Citeseer (c) PubMed (d) Wisconsin

(e) Film (f) Texas (g) Cornell

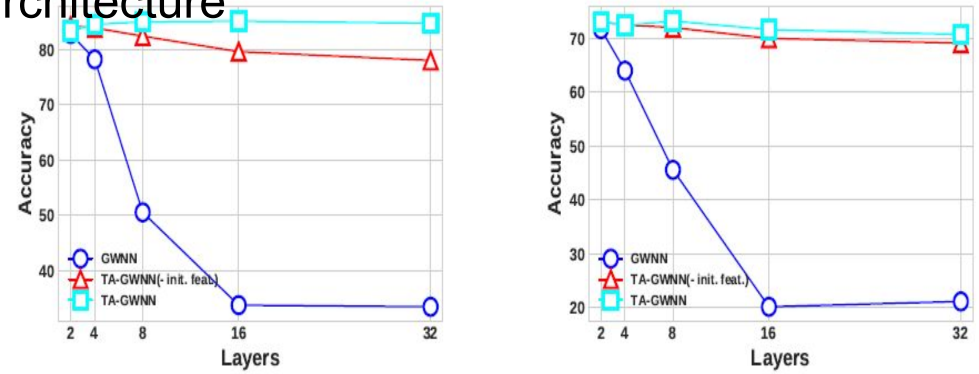
**GA-GWNN Time Complexity:** O(L|N|d_in d_out + L |E| d_out)

# Ablation Studies

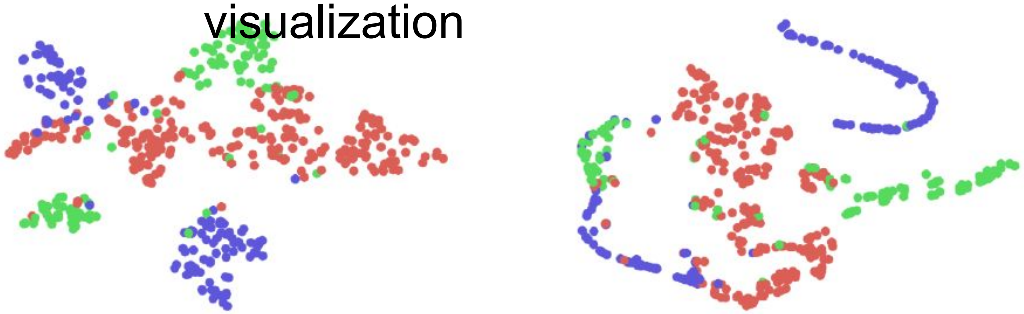
Performance of GA-GWNN with deeper layer in supervised task

| Dataset | 2 | 4 | 6 | 8 | 12 | 16 | 20 | 26 |
|---|---|---|---|---|---|---|---|---|
| | #Layers | | | | | | | |
| Cora | 88.32 | 88.49 | 88.61 | **88.63** | 88.29 | 88.43 | 88.40 | 88.53 |
| Citeseer | 77.08 | 76.92 | 76.87 | **77.10** | 76.37 | 76.76 | 77.02 | 76.87 |
| Film | 37.16 | 36.87 | 37.11 | 36.94 | **37.25** | 36.92 | 36.75 | 36.88 |
| Cornell | **85.16** | 84.60 | 81.35 | 81.76 | 80.94 | 79.86 | 79.80 | 79.70 |
| Texas | 84.15 | **84.83** | 84.16 | 84.30 | 82.38 | 83.65 | 82.23 | 83.00 |
| Wisconsin | 85.43 | 85.32 | 85.29 | **85.49** | 85.29 | 84.71 | 83.53 | 83.33 |

Semi supervised performance with deeper architecture




Feature map visualization



(a) TA-GWNN

(b) GWNN

# Simple and Effective GWNN (SEA-GWNN)

GA-GWNN

**Algorithm 1:** Implementation of the proposed GA-GWNN model

**Input:** Graph ($\mathcal{G}$) adjacency matrix $\tilde{\mathcal{A}}$, node feature matrix $\mathbf{X} \in \mathbb{R}^{N\times d}$, set of ground truth $\mathcal{Y}$, max layer $L$, number of sample node $\eta_v$

**Output:** Predicted class label: $\mathbf{Y}$.

Initialize model parameters
**for** *epoch* $\leftarrow 1 \dots m$ **do**
  $h^0 \leftarrow \mathbf{X}$;
  **for** *layers*, $\ell \leftarrow 1 \dots L$ **do**
    /* Structure aware lifting operators */
    $\boxed{\mathcal{U}^\ell \leftarrow \mathrm{ATT}_{gat}(\mathcal{G}, h^\ell)}$
    $\boxed{\mathcal{P}^\ell \leftarrow \frac{1}{2}\mathcal{U}^\ell}$
    /* Filtered signal in the wavelet domain */
    $\bar{h}^\ell = \mathrm{T}_k\left(\tilde{\mathcal{A}}\cdot h^\ell + \mathscr{U}^\ell_{+\mathrm{I}}\cdot h^\ell - h^\ell \circledast \tilde{\xi}\right)$
    /* Reconstruction in the spatial domain */
    $\cancel{\tilde{h}^\ell = \tilde{\mathcal{A}}\cdot\bar{h}^\ell + \mathscr{U}^\ell_{-\mathrm{I}}\cdot\bar{h}^\ell \circledast \xi + h^0}$
  **end**
  $\mathbf{Y} \leftarrow \mathrm{Softmax}(\tilde{h}^\ell)$ ; /* Outputs the softmax probability */
  $\mathcal{L} = \mathrm{Loss}(\mathbf{Y}, \mathcal{Y})$ ; /* Compute loss */
  Backpropagation and update parameters
**end**

SEA-GWNN

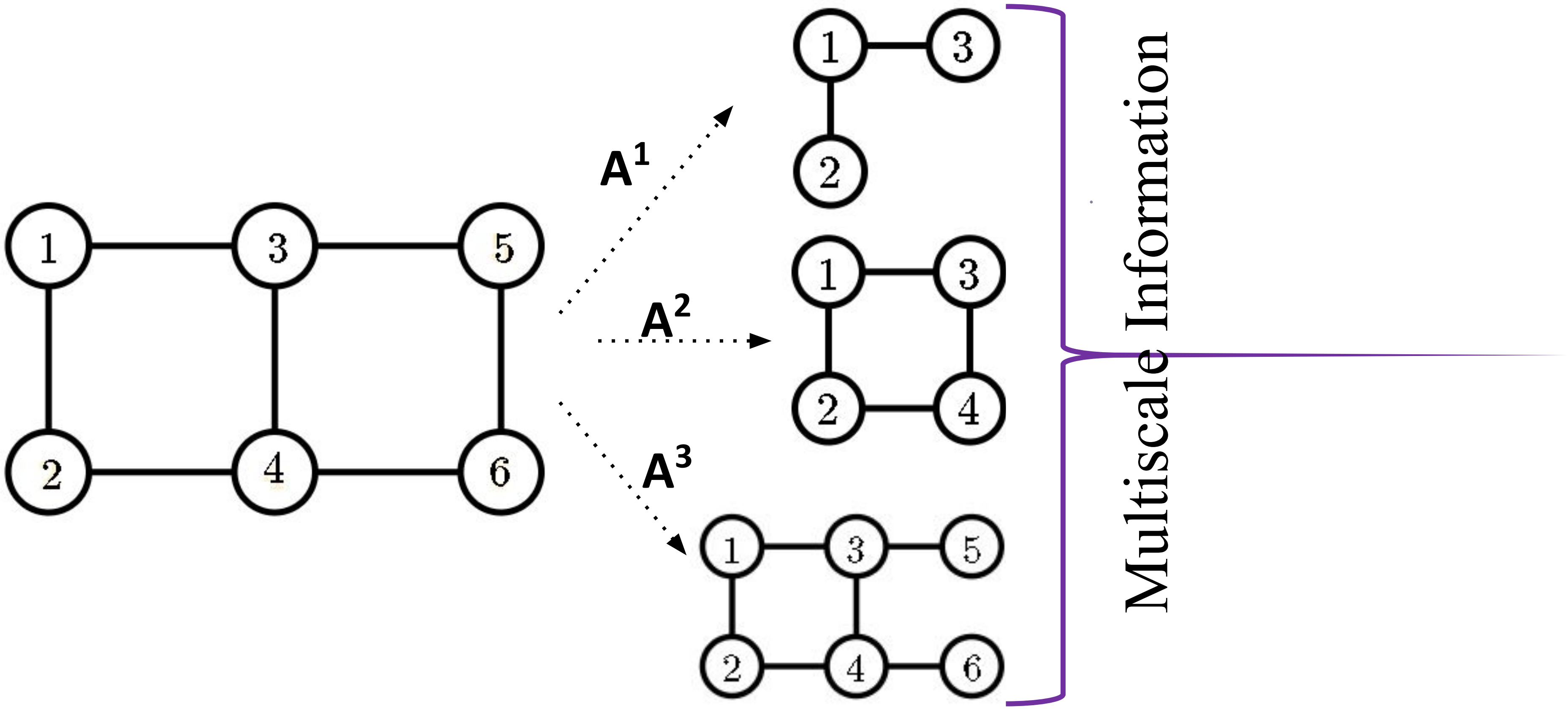
**Algorithm 2:** Implementation of the proposed SEA-GWNN model

**Input:** Graph ($\mathcal{G}$) adjacency matrix $\mathcal{A}$, node feature matrix $\mathbf{X} \in \mathbb{R}^{N\times d}$, set of ground truth $\mathcal{Y}$, max layer $L$

**Output:** Predicted class label: $\mathbf{Y}$.

Initialize model parameters
**for** *epoch* $\leftarrow 1 \dots m$ **do**
  $\mathbf{H}_0, \tilde{\mathbf{H}}_0 \leftarrow \mathbf{X}$;
  /* Structure aware lifting operators */
  $\mathcal{U} \leftarrow \mathrm{ATT}_{gat}(\mathcal{G}, \mathbf{X})$
  $\mathcal{P} \leftarrow \frac{1}{2}\mathcal{U}$
  **for** *layers*, $\ell \leftarrow 1 \dots L$ **do**
    $\mathbf{H}_\ell \leftarrow \mathcal{A}\mathbf{H}_{\ell-1}$
    $\tilde{\mathbf{H}}_\ell \leftarrow \tilde{\mathcal{U}}\tilde{\mathbf{H}}_{\ell-1}$
    /* Decompose signal in the wavelet domain */
    $\mathbf{Z}_\ell \leftarrow \alpha\mathbf{H}_\ell + \tilde{\mathbf{H}}_\ell \circledast \xi$
    $\mathbf{Z} \mathrel{+}= \gamma_\ell \mathbf{Z}_\ell$
  **end**
  $\mathbf{Y} \leftarrow \mathrm{Softmax}(\mathbf{Z})$ ; /* Outputs the softmax probability */
  $\mathcal{L} = \mathrm{Loss}(\mathbf{Y}, \mathcal{Y})$ ; /* Compute loss */
  Backpropagation and update parameters
**end**

# SEA-GWNN(2)



# Overall Results

Mean accuracy on full supervised node classification. Best results are highlighted in bold

| Method/ Acc.(%) | homophilic graph datasets | | | heterophilic graph datasets | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Cora | Cite. | Pubm. | Film | Cham. | Squi. | Corn. | Texa. | Wisc. |
| GCN[11] | 85.77±0.1 | 73.68±0.2 | 88.13±0.6 | 28.78±1.5 | 28.18±0.78 | 36.89±1.34 | 52.70±0.5 | 52.16±0.6 | 45.88±0.7 |
| GAT[10] | 86.37±0.2 | 74.32±0.2 | 87.62±0.4 | 28.99±1.4 | 42.93±0.46 | 30.62±2.11 | 54.32±0.3 | 58.38±0.5 | 49.41±0.9 |
| APPNP[11] | 87.87±0.2 | 76.53±0.2 | 89.40±1.4 | 34.86 ±1.1 | 54.30 ±0.56 | 34.77 ±0.34 | 73.51±1.1 | 65.41±0.1 | 69.02±1.6 |
| GWNN[2] | 85.31 ±0.2 | 73.93 ±0.1 | 88.14 ±0.1 | 26.62 ±1.4 | - | - | 61.89 ±0.6 | 60.00 ±0.5 | 48.04 ±0.7 |
| G-GCN[13] | 84.91±1.1 | 75.16±1.9 | 88.41±0.6 | 32.39±1.5 | 61.06±0.49 | 38.28±0.27 | 55.68±8.4 | 66.22±6.7 | 62.55±5.2 |
| MixHop[14] | 87.61±0.8 | 76.26±1.3 | 85.31±0.6 | 32.22±2.3 | 60.50±2.53 | 43.80±1.48 | 73.51±6.3 | 77.84±7.7 | 75.88±4.9 |
| H2GCN[15] | 87.69±1.3 | 75.95±2.1 | 88.78±0.5 | 36.71±1.4 | 58.38±1.76 | 37.90±2.02 | 78.92±5.2 | 82.16±8.2 | 82.57±3.2 |
| CPGNN[16] | 87.18±1.1 | 75.52±1.8 | 89.08±0.6 | 35.51±1.8 | 65.24±0.87 | 45.00±1.40 | 63.51±5.8 | 74.32±7.3 | 81.76±6.7 |
| GPR-GNN[17] | 86.70±1.1 | 75.12±1.9 | 87.38±0.6 | 36.47±1.3 | 65.42±2.04 | 49.93±0.53 | 82.97±5.6 | **84.59** ±4.3 | 83.92±3.1 |
| AM-GCN[18] | 86.66±1.3 | 76.01±1.9 | 86.78±0.6 | 33.60±1.1 | 68.46±1.70 | 40.02±0.96 | 78.38±4.9 | 78.38±7.2 | 81.76±4.9 |
| SEA-GWNN | **88.72** ±0.1 | **76.08** ±0.2 | **89.70**±0.1 | **37.16**±1.8 | **71.67**±0.18 | **54.02**±0.17 | **85.16**±4.6 | **84.45**±3.5 | **85.16**±3.3 |

Mean accuracy on semi supervised node classification with best results are highlighted in bold

| Methods | Cora | Citeseer | PubMed |
|---|---|---|---|
| GraphSAGE[9] | 74.5 ±0.8 | 67.2 ±1.0 | 76.8 ±0.6 |
| GAT[10] | 83.0 ±0.7 | 72.5 ±0.7 | 79.0 ±0.3 |
| HANet[5] | 81.9 | 70.1 | 79.3 |
| GWNN[2] | 81.6 ±0.7 | 70.5 ±0.6 | 78.6 ±0.3 |
| UFGConvS[3] | 83.0 ±0.5 | 71.0 ±0.6 | 79.4 ±0.4 |
| UFGConvR[3] | 83.6 ±0.6 | 72.7 ±0.6 | 79.9 ±0.1 |
| LGWNN[6] | 83.4 ±0.6 | 71.1 ±0.4 | 79.5 ±0.5 |
| SEA-GWNN | **84.4** ±0.3 | **72.8** ±0.3 | **80.7** ±0.2 |

# Overall Results(2)

Mean accuracy on large scale graphs node classification. Best results are highlighted in bold

| Methods | Penn94 | Arxiv-Year | Genius | Arxiv |
|---|---|---|---|---|
| MLP[19] | 73.61 ±0.40 | 36.70 ±0.21 | 86.68 ±0.09 | 55.00 ±0.025 |
| Label Prop.[19] | 74.13 ±0.46 | 46.07 ±0.15 | 67.04 ±0.20 | 68.32 ±0.00 |
| GCN[11] | 82.47 ±0.27 | 46.02 ±0.26 | 87.42 ±0.37 | 71.74 ±0.29 |
| ChebNet[20] | 82.51 ±0.31 | 46.76 ±0.24 | 89.36 ±0.31 | 71.72 ±0.22 |
| GAT[11] | 81.53 ±0.55 | 46.05 ±0.51 | 55.80 ±0.87 | 71.95 ±0.11 |
| GCNJK[20] | 81.63 ±0.54 | 46.28 ±0.29 | 89.30 ±0.19 | 72.19 ±0.21 |
| GCNII[21] | 82.92 ±0.59 | 47.21 ±0.28 | 90.24 ±0.09 | 72.74 ±0.16 |
| H2GCN[15] | 81.31 ±0.60 | OOM | OOM | OOM |
| GPRGNN[17] | 81.38 ±0.16 | 45.97 ±0.26 | 90.05 ±0.31 | 71.78 ±0.18 |
| UFGConv[3] | OOM | OOM | OOM | 71.97 ±0.12 |
| SEA-GWNN | **84.35** ±0.35 | **49.70** ±0.41 | **90.53** ±0.14 | **72.85** ±0.51 |

# Overall Results(3)

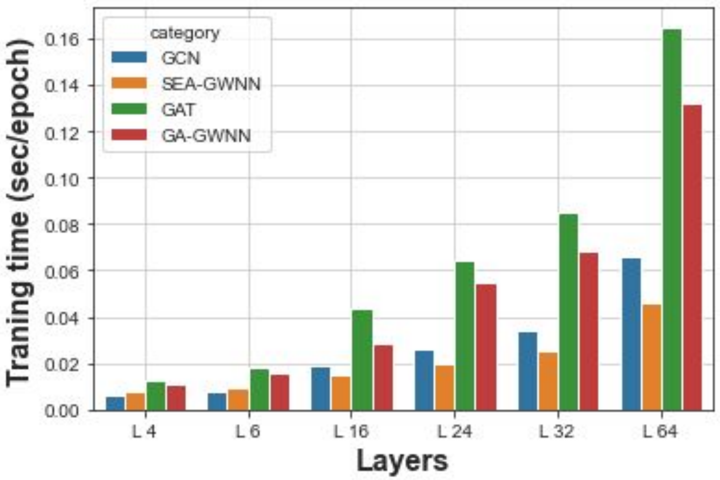
Graph level classification and prediction

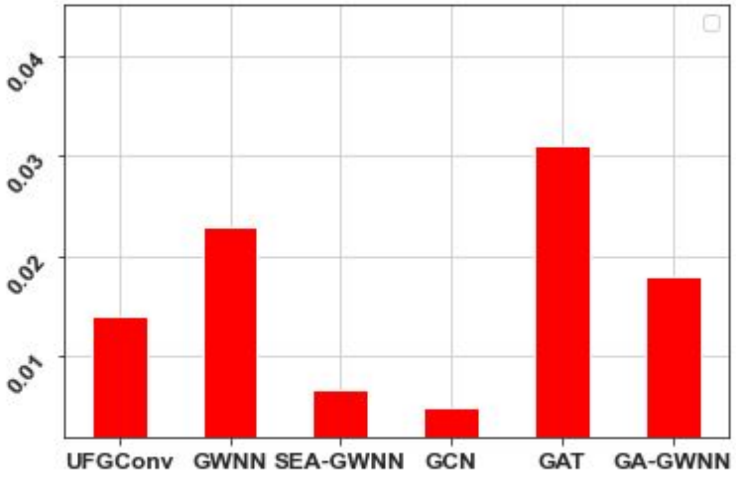
| Datasets | PROTEINS | Mutagenicity | D&D | NCI1 | QM7 |
|---|---|---|---|---|---|
| TopKPool | 73.48 ±3.57 | 79.84 ±2.46 | 74.87 ±4.12 | 75.11 ±3.45 | 175.41 ±3.16 |
| Attention | 73.93 ±5.37 | 80.25±2.22 | 77.48 ±2.65 | 74.04±1.27 | 177.99 ±2.22 |
| SAGPool | 75.89 ±2.91 | 79.86 ±2.36 | 74.96±3.60 | 76.30±1.53 | 41.93 ±1.14 |
| SUM | 74.91 ±4.08 | 80.69 ±3.26 | 78.91 ±3.37 | 76.96 ±1.70 | 42.09±0.91 |
| MAX | 73.57 ±3.94 | 78.83 ±1.70 | 75.80 ±4.11 | 75.96 ±1.82 | 177.48 ±4.70 |
| MEAN | 73.13 ±3.18 | 80.37 ±2.44 | 76.89 ±2.23 | 73.70 ±2.55 | 177.49 ±4.69 |
| UFGPool | 77.77±2.60 | 81.59 ±1.40 | 80.92±1.68 | 77.88±1.24 | 41.74±0.84 |
| GWNN [31] | 73.35 ±3.71 | 74.26 ±2.29 | 75.04 ±4.55 | 69.79±1.67 | - |
| LGWNN [33] | 74.02 ±5.23 | **82.47** ±1.90 | 78.72 ±4.33 | 78.97 ±2.07 | - |
| Ours | **80.23**±0.51 | 80.29 ±1.62 | **80.39**±0.49 | - | **40.95**±0.83 |

[5]H. Pei, B. Wei, K. C.-C. Chang, Y. Lei, and B. Yang, Geom-gcn: Geometric graph convolutional networks," in International Conference on Learning Representations, 2019.

# Model Analysis

**Computational Cost** O(|N|d_in d_out + L|E|d_out)




Model performance on deeper architecture

| Datasets/ Accuracy (%) | The number of graph convolutional layers $k$ | | | | | | |
|---|---|---|---|---|---|---|---|
| | 2 | 4 | 6 | 8 | 16 | 28 | 32 |
| Cora | 82.46 | 84.04 | **84.38** | 84.34 | 82.52 | 82.86 | 80.86 |
| PubMed | 79.90 | 79.44 | 79.84 | 80.34 | **80.74** | 79.94 | 79.92 |

# Conclusion

- Proposed a novel class of algorithm namely GA-GWNN that produces desirable wavelet filters.
- Proposed a novel lifting scheme namely tree lifting scheme that preserve the original graph structure.
- Our proposed GA-GWNN can learn wavelet filters on arbitrary graphs.
- Experimental results demonstrate the superiority of our proposed algorithm.
- Further proposed a simple and more scalable version of GA-GWNN namely SEA-GWNN

# References

[1]Mallat et. al., “A wavelet Tour of Signal Processing”,Elsevier,1999
[2]Xu et. al., “Graph wavelet neural network,” ICLR, 2018
[3]Zheng et. al., “How framelets enhance graph neural networks,“ ICLR, 2021
[4] Zheng et. al. , “Mathnet: Haarlike wavelet multiresolution analysis for graph representation learning,“ Knowledge-Based Systems, 2023.
[5]Li et. al., “Fast haar transforms for graph neural networks,” Neural Networks, 2020.
[6]Xu et. al., “Graph neural networks with lifting-based adaptive graph wavelets,” IEEE Transactions on Signal and Information Processing over Networks, 2022
[7]Shen et. al., “Optimized distributed 2d transforms for irregularly sampled sensor network grids using wavelet lifting," in IEEE ICASS, 2008
[8]Narang et. al., Lifting based wavelet transforms on graphs,“ APSIPA, 2009.
[9]Hamilton et. al., “Inductive representation learning on large graphs,”Neurips, 2017
[10] Velickovic et. al., “ Graph attention networks,” ICLR, 2018.
[11] Kipf et. al., “Semi-supervised classication with graph convolutional networks," ICLR, 2017
[11] Gasteiger et. al., “Predict then propagate: Graph neural networks meet personalized pagerank,“ICLR, 2018
[12] Perozzi et al, Deepwalk: Online learning of social representations,“ ICDM, 2014
[13] Pei et. al., Geom-gcn: Geometric graph convolutional networks," ICLR, 2019.
[14] Abu et. al. Mixhop: Higher-order graph convolutional architectures via sparsied neighborhood mixing," PMLR, 2019
[15] Zhu et. al., Beyond homophily in graph neural networks: Current limitations and effective designs, Neurips,, 2020.
[16] Zhu et. al., Graph neural networks with heterophily, AAAI, 2021.
[17] Chien et. al., Adaptive universal generalized pagerank graph neural network, ICLR, 2021.
[18] Wang et. al., Am-gcn: Adaptive multi-channel graph convolutional networks, ICDM, 2020.
[19] Hu et. al., Open graph benchmark for machine learning on graphs, Neurips, 2020.
[20]Grover et. al., Node2vec: Scalable feature learning for networks, ICDM, 2016
[21 ]Derg et. al., Graphzoom: A multi level spectral approach for accurate and scalable graph embedding, ICLR, 2019
[22]Li et. al., DeeperGCN: All you need to train deeper gcn, ICLR, 2020
[23]Fabrizzio et. al., Scalable inception graph neural network, ICML, 2020.
[24]Sen et. al., Collective classication in network data, AI magazine, 2008