



# 面向零样本节点分类的对偶双向图卷积网络

岳琴, 梁吉业\*, 崔军彪, 白亮  
山西大学, 计算机与信息技术学院

## 背景

图数据是一种重要的数据表现形式, 在开放世界中, 新类是不断涌现的, 因此标记样本不能完全覆盖所有类别, 即零样本节点分类。例如: 在引文网络中, 新的研究主题会不断出现。

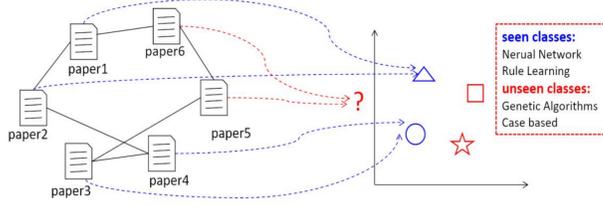


图1 引文网络中的零样本节点分类问题

## 相关工作: 图卷积网络 (GCN)

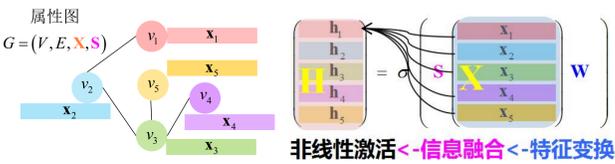


图2 传统图卷积网络的示意图

- 每个节点的新表示 (每行) 是所有节点表示 (行) 的线性组合;
- 组合系数由图的拓扑结构决定;
- 是拓扑结构指导的信息融合的过程

## 双向图卷积网络 (BiGCN)

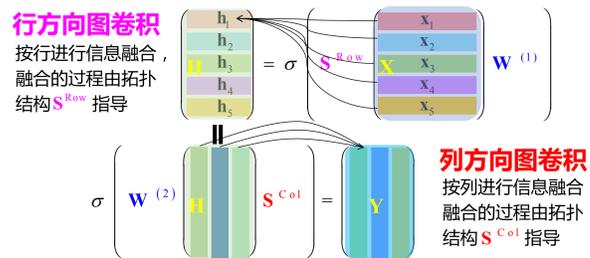


图3 双向图卷积网络的示意图

## 对偶双向图卷积网络 (DBiGCN)

### 问题定义

令  $G = (V, E, X, S^V)$  表示一个节点集为  $V$ 、边集为  $E$  的属性图,  $S^V \in \mathbb{R}^{n \times n}$  表示节点的邻接矩阵,  $X \in \mathbb{R}^{n \times d}$  表示节点的属性矩阵。

对于零样本节点分类问题, 类空间由两个不交集的部分组成, 即  $\mathcal{Y} = \mathcal{Y}_s \cup \mathcal{Y}_u$  and  $\mathcal{Y}_s \cap \mathcal{Y}_u = \emptyset$ 。为了方便, 假设有  $c_s$  个可见类和  $c_u$  个未见类 ( $c_s + c_u = c$ )。且每个类通过一个语义描述向量进行描述,  $A \in \mathbb{R}^{c \times d_c}$  表示对应的类语义描述矩阵。

### 方法

#### (1) 节点角度的双向图卷积网络

基于节点角度学习节点和类别的联合表示, 可以同时融合节点信息、节点之间的关系以及类别之间的关系, 即

$$Y^V = \text{softmax}(\text{relu}(\hat{S}^V X W^{(1)}) W^{(2)} \hat{S}^A)$$

然后, 对于所有的标记样本, 我们应用交叉熵损失, 有

$$\mathcal{L}_{\text{nodes}} = - \sum_{i=1}^L \sum_{j=1}^c y_{Lij}^{\text{true}} \ln y_{ij}^V$$

#### 框架

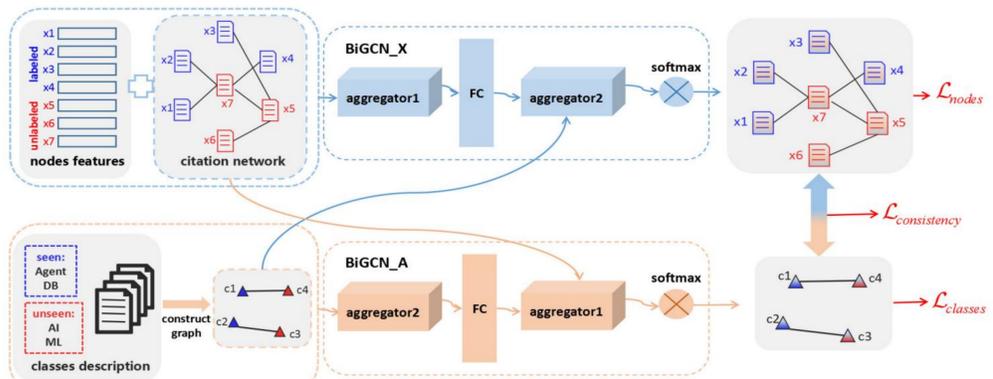


图4 对偶双向图卷积网络的示意图

#### (2) 类别角度的双向图卷积网络

类似地, 基于类别角度学习节点和类别的联合表示, 可以同时融合类别语义描述信息、节点之间的关系以及类别之间的关系, 即

$$Y^A = \text{softmax}(\hat{S}^A A W^{(3)} \hat{S}^V)$$

然后, 对于所有的标记样本, 我们应用交叉熵损失, 有

$$\mathcal{L}_{\text{classes}} = - \sum_{i=1}^L \sum_{j=1}^c y_{Lij}^{\text{true}} \ln y_{ji}^A$$

#### (3) 标签一致性损失

基于节点和类别的角度分别学到了节点和类别的联合表示, 很自然地我们希望两个双向图卷积网络可以协作以实现互相指导、互相提升。因此, 我们设计了标签一致性损失使得两个双向图卷积网络的表示保持一致, 同时标记样本对的输出要与真实标签保持一致, 即

$$\mathcal{L}_{\text{consistency}} = \left\| Y_L^V Y_L^A - Y_L^{\text{true}} (Y_L^{\text{true}})^T \right\|_F^2$$

综上, 对偶双向图卷积网络的最终的目标函数为

$$\mathcal{L}_{\text{overall}} = \mathcal{L}_{\text{nodes}} + \alpha \mathcal{L}_{\text{classes}} + \beta \mathcal{L}_{\text{consistency}}$$

## 实验

### 数据集

表1 数据集划分信息

	Class split I		Class split II	
	Train	Test	Train	Test
Cora	{1, 2, 3}	{4, 5, 6, 7}	{1, 2}	{3, 4}, {5, 6, 7}
Citeseer	{1, 2}	{3, 4, 5, 6}	{1, 2}	{3, 4}, {5, 6}
C-M10M	{1, 2, 3}	{4, 5, 6}	{1, 2}	{3, 4}, {5, 6}

### 实验1—不同数据划分下的实验结果

表2 使用TEXT-CSDs下的零样本节点分类精度

	Cora			Citeseer			C-M10M		
	Acc	Std	Impr	Acc	Std	Impr	Acc	Std	Impr
RandomGuess	25.35	0.00	0.00	24.86	0.00	0.00	33.21	0.00	0.00
DAP	26.56	0.00	0.00	34.01 <sup>3</sup>	0.00	0.00	38.71 <sup>3</sup>	0.00	0.00
DAP(CNN)	27.80	0.00	0.00	30.45	0.00	0.00	32.97	0.00	0.00
ESZSL	27.35	0.00	0.00	30.32	0.00	0.00	37.00	0.00	0.00
ZS-GCN	25.73	0.00	0.00	28.62	0.00	0.00	37.89	0.00	0.00
ZS-GCN(CNN)	16.01	0.00	0.00	21.18	0.00	0.00	36.44	0.00	0.00
WDVSc	30.62 <sup>3</sup>	0.00	0.00	23.46	0.00	0.00	38.12	0.00	0.00
Hyperbolic-ZSL	26.36	0.00	0.00	34.18	0.00	0.00	35.80	0.00	0.00
DGPN	33.78 <sup>2</sup>	0.00	0.00	38.02 <sup>2</sup>	0.00	0.00	41.98 <sup>2</sup>	0.00	0.00
DBiGCN	45.14 <sup>1</sup>	0.00	0.00	40.97 <sup>1</sup>	0.00	0.00	45.45 <sup>1</sup>	0.00	0.00
Improve rate	33.63%	0.00	0.00	7.76%	0.00	0.00	8.27%	0.00	0.00

结论:

a) 所提方法 DBiGCN 所有情况下都优于 RandomGuess (随机猜) 以及经典的零样本学习方法, 说明所提方法的对于零样本节点分类的有效性。

b) 所提方法 DBiGCN 在大多数情况下优于零样本节点分类的方法, 说明所提方法的优越性和适用性。

### 实验2—不同类语义描述下的实验结果

表3 使用不同类语义描述下的零样本节点分类精度比较

Class Split I	Cora			Citeseer			C-M10M		
	TEXT	LABEL	Decline rate	TEXT	LABEL	Decline rate	TEXT	LABEL	Decline rate
DAP	26.56	25.34	-4.59%	34.01	30.01	-11.76%	38.71	32.67	-15.60%
ESZSL	27.35	25.79	-5.70%	30.32	28.52	-5.94%	37.00	35.02	-5.35%
ZS-GCN	25.73	23.73	-7.77%	28.62	26.11	-8.77%	37.89	33.32	-12.06%
WDVSc	30.62	18.73	-38.83%	23.46	19.70	-16.02%	38.12	30.82	-19.15%
Hyperbolic-ZSL	26.36	25.47	-3.38%	34.18	21.04	-38.44%	35.80	34.49	-3.66%
DGPN	33.78	32.55	-3.64%	38.02	31.83	-16.28%	41.98	35.05	-16.51%
DBiGCN	45.14	39.05	-13.49%	40.97	39.10	-3.10%	45.45	43.71	-3.83%

结论:

类别语义 TEXT-CSDs 比 LABEL-CSDs 所包含的信息多, 所以基于 TEXT-CSDs 的零样本节点分类性能更好。与其他方法相比, 所提方法 DBiGCN 对于类别语义的质量的下降有更稳定的性能。

### 实验3—相对优势分析

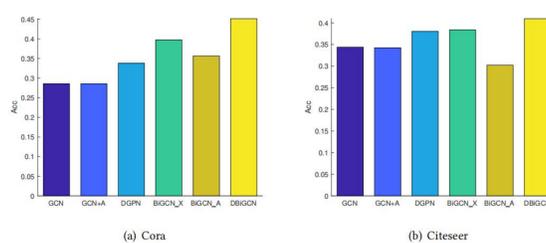


图5 不同图卷积网络在零样本节点分类中上的性能比较

结论:



### 实验4—消融实验

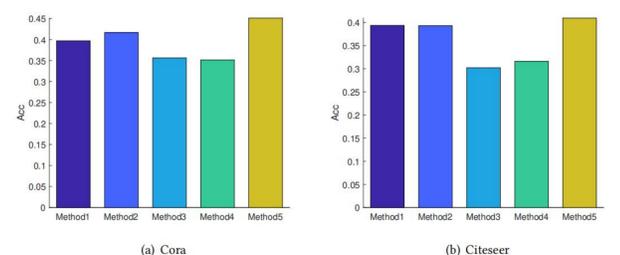


图6 不同消融模型的零样本节点分类精度

- Method1: BiGCN\_X + 损失  $\mathcal{L}_{\text{nodes}}$
- Method2: BiGCN\_X + 损失  $\mathcal{L}_{\text{consistency}}$
- Method3: BiGCN\_A + 损失  $\mathcal{L}_{\text{classes}}$
- Method4: BiGCN\_A + 损失  $\mathcal{L}_{\text{consistency}}$
- Method5: 全模型

## 总结

- 提出了一种双向图卷积网络, 能有效融合特征信息、拓扑信息和类别信息
- 提出了一种对偶双向图卷积网络, 能有效实现二者的相互指导
- 为零样本学习提供了一个有效的解决方案

## 致谢

本工作受到国家重点研发计划 (2020AAA0106100) 以及国家自然科学基金 (U21A20473, 62022052) 的支持。



论文链接: <https://dl.acm.org/doi/10.1145/3534678.3539316>

代码链接: <https://github.com/warmerspring/DBiGCN>

