

联合图随机游走和跳跃连接的动态超图神经网络

牛雪琼¹ 农丽萍^{2,3} 梁海^{4*} 王俊义¹ 林基明¹

¹(桂林电子科技大学信息与通信学院 广西 桂林 541004)

²(西安电子科技大学通信工程学院 陕西 西安 710071)

³(广西师范大学物理科学与技术学院 广西 桂林 541004)

⁴(桂林电子科技大学计算机与信息安全学院 广西 桂林 541004)

摘要 针对传统超图神经网络难以提取节点直接邻域外关联度高的节点特征,导致全局特征信息不完整的问题,对动态超图神经网络(DHGNN)进行改进,提出联合图随机游走和跳跃连接的动态超图神经网络(RWS-DHGNN),用于非欧几里得数据的分类。该网络在DHGNN的基础上,引入了图随机游走,从而有效地获取直接邻域外关联度高的节点特征。同时,引入残差网络的思想在超图的顶点卷积处增加跳跃连接构成残差结构。所提网络模型充分发挥图结构和超图结构的优势。在Cora数据集的标准分割和随机分割上将所提网络与GCN、HGNN、GAT和DHGNN进行对比实验,实验结果表明,该模型可以有效提高分类准确率。

关键词 超图神经网络 随机游走 跳跃连接 节点分类

中图分类号 TP3

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.03.028

DYNAMIC HYPERGRAPH NEURAL NETWORK WITH JOINT GRAPH RANDOM WALK AND SKIP CONNECTION

Niu Xueqiong¹ Nong Liping^{2,3} Liang Hai^{4*} Wang Junyi¹ Lin Jiming¹

¹(School of Information and Communication, Guilin University of Electronic Technology, Guilin 541004, Guangxi, China)

²(School of Telecommunication Engineering, Xidian University, Xi'an 710071, Shaanxi, China)

³(College of Physics and Technology, Guangxi Normal University, Guilin 541004, Guangxi, China)

⁴(School of Computer and Information Security, Guilin University of Electronic Technology, Guilin 541004, Guangxi, China)

Abstract Traditional hypergraph neural network is difficult to extract node features with high degree of correlation outside the direct neighborhood of nodes, which leads to incomplete global feature information. Dynamic hypergraph neural network (DHGNN) is improved, and a dynamic hypergraph neural network with joint graph random walk and skip connection (RWS-DHGNN) is proposed. RWS-DHGNN was used to classify non Euclidean data. Based on the DHGNN, graph random walk was introduced into the network to effectively obtain the node features with high degree of correlation outside the direct neighborhood of nodes. Meanwhile, the idea of residual network was introduced, and the residual structure was formed by adding skip connections at the vertex convolution of hypergraph. RWS-DHGNN gave full play to the advantages of graph structure and hypergraph structure. RWS-DHGNN was compared with GCN, HGNN, GAT and DHGNN on Cora dataset. The experimental results show that RWS-DHGNN can effectively improve the classification accuracy.

Keywords Hypergraph neural network Random walk Skip connection Node classification

收稿日期:2021-01-18。国家自然科学基金项目(61966007);“认知无线电与信息处理”教育部重点实验室开放基金项目(CRKL180201);广西无线宽带通信与信号处理重点实验室开放基金项目(GXKL06190204);广西无线宽带通信与信号处理重点实验室主任基金项目(GXKL06190117,GXKL06200116)。牛雪琼,硕士生,主研领域:图信号处理,超图神经网络。农丽萍,讲师。梁海,讲师。王俊义,研究员。林基明,教授。

0 引言

随着大量数据的产生和人工智能发展的需要,如何处理非欧几里得数据具有重要的研究意义。图是表达非欧几里得数据的一种重要形式,描述的是一组对象的二元关系,现实中的很多问题大都可以抽象为图去解决,比如社交网络中人际关系的推断、天气网络中温度变化的预测、引文网络中文章的分类等。通过挖掘图中的节点信息和节点之间连边潜在的关系,获取节点的属性信息与结构信息,从而更好地利用图的表征能力实现各种任务,比如节点分类^[1]、图分类^[2]和链接预测^[3]等。

近年来图卷积网络获得了极大的研究关注,在交通预测、网络分析、生物化学、计算机视觉、自然语言处理、推荐系统等领域表现出了优越的性能。在图卷积神经网络节点分类的研究工作中,Kipf等^[4]提出了图卷积网络半监督分类(GCN)方法。该方法将谱图卷积的局部一阶近似作为卷积结构,这为完成图结构数据节点分类任务提供了一种有效的思路。Velickovic等^[5]提出图注意力网络(GAT)。该模型将注意机制引入到图卷积模型中,明显提高了图结构数据节点分类准确率。然而,图卷积网络是在表征成对关系的图结构上挖掘信息,往往忽略了对对象之间的多元关系。在实际应用中对象之间复杂、多元的关系普遍存在。比如文章的引用关系就是多元关系,一篇文章可以被多篇文章引用,也可以引用多篇文章。当多元关系被强制转换为二元关系时,虽然可以降低解决问题的难度,但这将导致很多有用信息的丢失。相对图结构而言,超图可以更加准确地描述对象之间的多元关系,同时在网络分析、计算机视觉等领域中应用广泛,并取得了不错的效果。近年来超图神经网络逐渐成为机器学习领域研究的热点。例如,Feng等^[6]提出超图神经网络(HGNN),利用K-NN算法(K-Nearest Neighbor algorithm)进行超图构造。该模型编码了高阶和复杂的数据关系,该网络在利用高阶、复杂的相关性学习数据表征上取得了较好的效果。为了解决HGNN忽略了训练过程中通过调整特征嵌入动态修改超图结构的问题,Jiang等^[7]提出了动态超图神经网络(DHGNN)。该网络较好地获取了固有结构没有直接表示出来的隐藏的数据之间的高阶相关性信息。DHGNN有利于提取对象间高阶复杂的关系,增加高阶有效信息。然而其没有很好地提取节点直接邻域外关联度高的节点特征。鉴于DHGNN所采用的原始数据是Cora数据集的

图结构数据,为了充分地捕获节点邻域外的关联度高的节点特征信息,本文拟在DHGNN结构的基础上,首先增加图数据的预处理工作,即采用图上的随机游走过程^[8-9]生成从中心节点走 r 步转移到某个节点的概率,得到转移概率矩阵来计算节点间的相似度,获取中心节点直接邻域外关联度高的节点特征,有效地增加全局特征信息。其次,通过引入残差网络的思想^[10]在超图的顶点卷积处增加跳跃连接构成残差结构来提高节点分类准确率。

本文所提的网络充分发挥了图结构和超图结构的优势。图结构在表征对象之间二元关系方面具有显著优势,而超图具有能够很好地表示数据间高阶相关性的优势。将两者相结合,我们能够更准确地获取节点间的二元关系和多元关系,进而获得更出色的特征表示,从而得到更好的预测性能。在公开数据集Cora上进行实验,实验结果表明本文模型可以更好地提取直接邻域外节点的有效信息,得到较好的分类效果。

1 模型设计

本节对DHGNN和顶点卷积的构成、图上的随机游走过程和残差结构的相关理论依次进行阐述,最后详细讲解RWS-DHGNN的构成。

1.1 DHGNN和顶点卷积

DHGNN模型结构示意图如图1所示。该模型主要由超图构造和超图卷积两个模块构成,用于捕获固有结构没有直接表示出来的隐藏的数据之间的高阶相关性信息。它联合K-NN算法和K-means算法进行动态超图构造。超图卷积层是由顶点卷积和超边卷积构成,这可以更好地细化特征。在将节点特征送入超图卷积层后,将执行节点特征到超边特征,再到节点特征的转换,然后得到中心节点新的特征表示,接着通过迭代训练后得到分类模型。

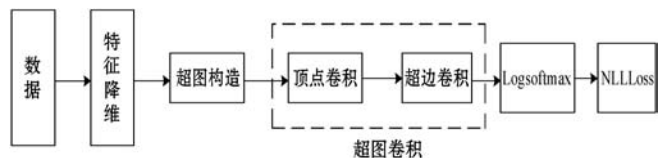


图1 DHGNN节点分类模型

DHGNN中的顶点卷积是将超边中所连接的顶点信息聚合到超边上,其由中间带有Softmax函数的两个一维卷积构成。其先由顶点特征来生成变换矩阵 B ,如式所示:

$$B = \text{softmax}(\text{conv1d}(X)) \quad (1)$$

式中:conv1d表示一维卷积。

然后通过生成的变换矩阵聚合顶点特征 X 到包含这些顶点的超边上,如式所示:

$$y(X) = \text{conv1d}(BX) \quad (2)$$

式中: $y(X)$ 表示超边特征。

如上所述,DHGNN 可以很好地学习数据之间的有效信息。但是存在有效信息不完整的问题,因此本文通过引入图随机游走来获取中心节点直接邻域外关联度高的节点特征,增加全局特征信息的获取。并且在超图的顶点卷积处增加跳跃连接构成残差结构来提高网络模型的节点分类准确率。

1.2 图上的随机游走过程

增加有效信息是提高节点分类准确率的一种方法,本文倾向于通过图数据预处理增加节点的全局特征信息,因此引入了图上的随机游走过程来捕获中心节点直接邻域外关联度高的节点特征。图上的随机游走过程是指从当前节点出发以一定概率随机游走 r 步到达某个节点。当图上的随机游走过程中的转移概率越大时,节点间越趋向于有着相同的标签。

用 R 表示图单步随机游走过程中的转移概率矩阵, R 中的每一个元素定义为 $r(u,v)$ 如式所示:

$$r(u,v) = \begin{cases} \frac{w(u,v)}{d(u)} & \{u,v\} \in E \\ 0 & \text{其他} \end{cases} \quad (3)$$

式中: $r(u,v)$ 表示从顶点 u 转移到顶点 v 的转移概率; $w(u,v)$ 表示顶点 u 与顶点 v 的权重函数,顶点度 $d(u) = \sum_{\{u,v\} \in E} w(u,v)$ 。 R 用矩阵的形式表示如式所示:

$$R = D^{-1}W \quad (4)$$

式中: D 为度矩阵; W 为权重矩阵; $D = \text{diag}(d_1, d_2, \dots, d_n)$; $d_i = \sum_j W_{ij}$; R 为图单步随机游走过程中的转移概率矩阵。图 r 步随机游走过程中的转移概率矩阵 R^r 表示如式所示:

$$R^r = (D^{-1}W)^r \quad (5)$$

通过在图上进行 r 步的随机游走,节点可以到达其直接邻域外的节点。因此本文通过式(6)针对图信号 $X \in \mathbf{R}^{N \times D_1}$,其中 N 表示节点个数和 D_1 表示节点的特征维数,利用该公式进行 r 步游走过程,可以获得中心节点直接邻域外关联度高的节点特征,增加了全局特征信息,进而提高节点分类准确率,其中新的特征表示 S 为信号 X 的输出。

$$S = R^r X \quad (6)$$

为了更好地说明式(6)的原理,设 r 为 2 即节点之间进行 2 步随机游走,以从节点 A 出发随机游走两步

为例,图上的随机游走过程如图 2 中的箭头所示。根据图 2 可以构成 2 步转移概率矩阵,如表 1 所示。当以节点 A 作为起点,游走两步可以到达距离较远的节点 E 处,节点 E 是节点 A 的直接邻域外节点。综上所述,一个节点经过多步随机游走可以跳出直接邻域区域,捕获到距离中心节点较远的节点的信息,从而得到该节点直接邻域外关联度高的节点特征。

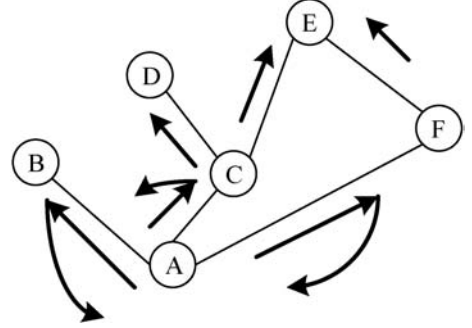


图 2 图上的随机游走过程

表 1 2 步转移概率矩阵

节点	A	B	C	D	E	F
A	0.611 1	0	0	0.111 1	0.277 8	0
B	0	0.333 3	0.333 3	0	0	0.333 3
C	0	0.111 1	0.611 1	0	0	0.277 8
D	0.333 3	0	0	0.333 3	0.333 3	0
E	0.416 7	0	0	0.166 7	0.416 7	0
F	0	0.166 7	0.416 7	0	0	0.416 7

1.3 残差结构

本文引入残差网络的思想,分别在 K-means 和图随机游走之后的顶点卷积处增加跳跃连接构成残差结构,提高了网络模型的节点分类准确率。残差是指实际观测值与预测值之间的差。残差网络^[10]是由来自 Microsoft Research 的 4 位学者提出的卷积神经网络,其特点是容易优化,并且其内部的残差块使用跳跃连接可以缓解网络层数多时带来的网络退化问题。

DHGNN 中的顶点卷积由中间带有 Softmax 函数的两个一维卷积构成,其将顶点特征聚合到超边上,有利于对顶点特征之间有鉴别力的信息进行建模。特征送入顶点卷积后的具体操作如式(1)和式(2)所示。本文利用残差网络的思想,在顶点卷积上进行跳跃连接构成残差块,如图 3 所示,其中,⊕表示以张量的形式进行相加。

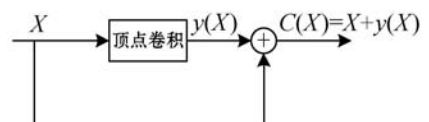


图 3 顶点卷积构成的残差结构

1.4 RWS-DHGNN

为了获得更好的节点分类准确率,需要捕获更多的有效信息和节点之间具有鉴别力的特征信息,因此本文在 DHGNN 的基础上引入了图上的随机游走过程和残差结构,整体框架如图 4 所示,其中虚线框部分是本文所增加的模块。同时在超图构造中 K-means 算法之后的顶点卷积处也增加了跳跃连接构成残差结构。RWS-DHGNN 的详细描述如下。

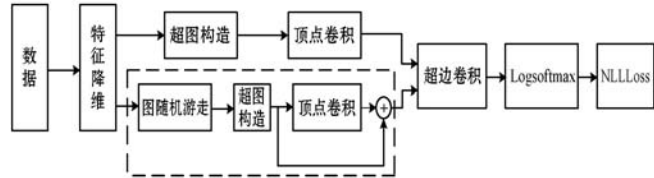


图4 RWS-DHGNN 图结构数据节点分类模型

(1) 特征降维:采用全连接层将以张量形式输入的 $2\,708 \times 1\,433$ 维度的特征信息进行降维操作,将特征矩阵由 1 433 维度降为 256 维度,降维的同时也保留了有效信息。

(2) 超图构造:采用 K-NN 算法来生成基本超边,通过 K-means 算法扩展相邻超边集,将生成的顶点特征送入顶点卷积操作中。

(3) 图随机游走:首先通过图上的随机游走过程提取中心节点直接邻域外关联度高的节点特征,然后利用 K-NN 算法进行超图构造。

(4) 顶点卷积:顶点卷积是由中间带有 Softmax 函数的两个一维卷积构成。将顶点特征送入顶点卷积中,顶点特征被聚合到超边上从而得到超边特征。

(5) 超边卷积:将超边特征送入超边卷积中,超图中的超边特征被聚合到顶点上从而得到顶点特征,获得中心顶点新的特征表示用于节点分类。

(6) 残差结构:通过引入残差网络的思想,在顶点卷积处增加跳跃连接构成残差结构,提高节点分类准确率。

(7) 损失函数:采用 NLLoss 作为损失函数,将信息输入 NLLoss 前需要进行 Logsoftmax 操作。

2 实验与结果分析

为了更好地说明本文网络的有效性,本文通过数值实验将 RWS-DHGNN 与其他方法进行比较。实验分为三个模块:图结构数据节点分类实验分析、随机游走步数对分类准确率的影响分析和对比实验有效性分析。

2.1 图结构数据节点分类

本文采用 Cora 基准数据集作为实验数据集,常用于图网络的有效性验证,此数据集是引文网络数据集,一共有 2 708 篇文章,每篇文章至少引用一篇文章或被一篇文章引用,每篇文章有 1 433 个特征维度。文章作为顶点,顶点之间的连边表示文章之间有引用关系。

Yang 等^[11]描述的 Cora 数据集标准分割的标签率是 5.2%,7 类文章都包含了,每类文章各有 20 篇,共 140 个节点有标签。Cora 数据集随机分割选择 2%、5.2%、10%、20%、30% 和 44% 比例作为训练集中带标签的节点。

实验中学习率参数为 0.01,epoch 参数为 25,权重衰减参数为 0.000 5,使用 Adam 优化器,K-NN 算法中的 K 为 64,K-means 算法中的簇数为 400。

实验的硬件环境为 PC:英特尔 Core i5-9400 处理器、4 GB 显存和 8 GB 内存;软件环境为 Ubuntu 18.04.5 LTS 64 位操作系统、CUDA 9.2、PyTorch 0.4.1 和 Python 3.6。

为了验证本文模型的有效性,在 Cora 基准数据集标准分割上进行验证,140 个带标签节点作为训练集,随机游走步数 r 取为 30。所提网络 RWS-DHGNN 的测试集分类准确率与 GCN、HGNN、GAT 和 DHGNN 进行比较,如表 2 所示。

表2 Cora 数据集标准分割上的分类准确率对比(%)

方法	准确率
GCN	81.5
HGNN	81.6
GAT	83.0
DHGNN	82.5
本文模型	84.7

RWS-DHGNN 在 Cora 数据集标准分割上的分类准确率是最优的。RWS-DHGNN 的分类准确率比 GCN 方法提高 3.2 个百分点,比 GAT 方法提高了 1.7 个百分点,这充分说明了 RWS-DHGNN 可以更好地提取有效信息,更好地完成节点分类任务;与 HGNN 和 DHGNN 这两种方法相比较,RWS-DHGNN 在节点分类准确率上分别高出 3.1 个百分点和 2.2 个百分点,表明所提网络可以更好地提取中心节点直接邻域外关联度高的节点特征,网络模型的表征能力更强,有利于有效信息的提取,从而提高节点分类的准确率。

在 Cora 数据集随机分割中,在随机选取标签率为 2%、5.2%、10%、20%、30% 和 44% 情况下分别取随机游走步数 r 为 5、10、15、20、25、30 和 35,在相同标签率不同随机游走步数下运行 10 次计算出测试集节点分类准确率的平均值,取平均值中的最大值与最小值的平均值作为中间值写入表 3 中,然后将最大值与中间值的差值写入表 3 中作为节点分类准确率,将 RWS-DHGNN 的节点分类准确率与 GCN、HGNN、GAT 和 DHGNN 进行比较,结果如表 3 所示。

表 3 Cora 数据集随机分割上的分类准确率对比

标签率	节点数	模型准确率/%				
		GCN	HGNN	GAT	DHGNN	本文
2%	54	69.6	75.4	74.8	76.9	73.81 ± 3.19
5.2%	140	77.8	79.7	79.4	80.2	81.55 ± 0.94
10%	270	79.9	80.0	81.5	81.6	83.35 ± 0.98
20%	540	81.4	80.1	83.5	83.6	85.69 ± 0.74
30%	812	81.9	82.0	84.5	85.0	85.76 ± 1.09
44%	1 200	82.0	81.9	85.2	85.6	86.83 ± 0.25

由表 3 可以看出本文模型除了标签率为 2% 相对 HGNN、GAT 和 DHGNN 较差外,其他情况均占据了优势。例如在标签率为 10% 时,RWS-DHGNN 的节点分类准确率中间值相比 GCN、HGNN、GAT 和 DHGNN 分别高出 3.45、3.35、1.85 和 1.75 百分点。在标签率为 20% 时,RWS-DHGNN 的节点分类准确率中间值相比 GCN、HGNN、GAT 和 DHGNN 分别提高 4.29、5.59、2.19 和 2.09 百分点。在大多数情况下 RWS-DHGNN 都可以得到较好的节点分类准确率,这表明所提网络可以更好地提取较为丰富的有效信息,特征学习更加准确。

2.2 随机游走步数对分类准确率的影响分析

本文通过图上随机游走过程来获取中心节点直接邻域外关联度高的节点特征,因此随机游走步数 r 的选择对准确获取有效信息至关重要。在 Cora 数据集标准分割上,分别选取游走步数 r 为 5、10、15、20、25、30 和 35 进行实验,结果如图 5 所示。当 r 为 30、35 的时候,分类效果最好,表明此时可以更准确地获取直接邻域外关联度高的节点特征,但考虑到随机游走步数越多,运算量越大,因此选择 r 为 30 作为最合适的游走步数,此时节点分类的准确率最高,为 84.7%。当随机游走 15 步时分类准确率下降,这是因为随机游走 15 步时,刚好此时的节点和目标节点的属性比较不相似,具有较大差别;当随机游走的远时提取了较多直接邻域外关联度高的节点特征,故准确率有所

回升。

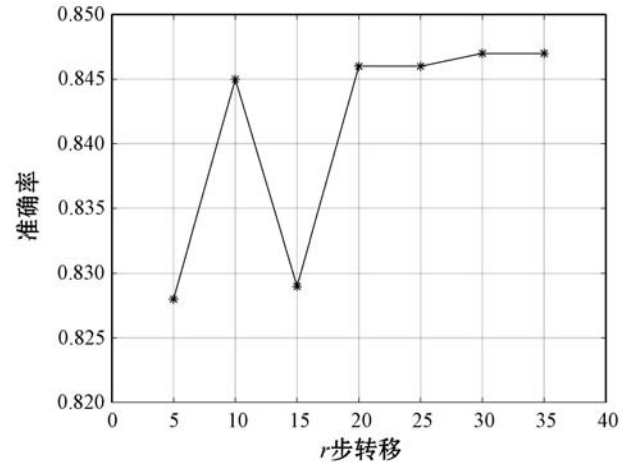


图 5 RWS-DHGNN 中随机游走步数 r 对分类准确率的影响

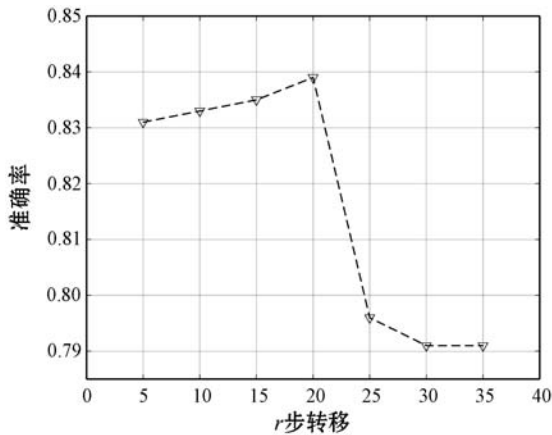
2.3 对比实验有效性分析

在 Cora 数据集标准分割上,分别在 DHGNN 加超图构造加顶点卷积加图随机游走 ($r=30$)、DHGNN 加超图构造加顶点卷积加跳跃连接和 RWS-DHGNN ($r=30$) 上进行实验,结果如表 4 所示,DHGNN 加超图构造加顶点卷积加跳跃连接上得到的分类准确率比 DHGNN 分类准确率高 1.7 百分点,通过引入残差网络的思想有利于更好地提取有效信息,提高模型的节点分类准确率,由实验可证,残差结构是有效的。

表 4 方法的有效性实验

方法	测试集节点分类准确率/%
DHGNN 加超图构造加顶点卷积加图随机游走 ($r=30$)	79.1
DHGNN 加超图构造加顶点卷积加跳跃连接	84.2
本文 ($r=30$)	84.7

在 DHGNN 加超图构造加顶点卷积加图随机游走上,本文分别取随机游走步数 r 为 5、10、15、20、25、30、35 进行实验,结果如图 6 所示。当随机游走 30 步时,测试集节点分类准确率不高,为 79.1%。然而,当随机游走步数 r 为 20 时,节点分类准确率最高,为 83.9%。此时 DHGNN 加超图构造加顶点卷积加图随机游走 ($r=20$) 的分类准确率比 DHGNN 的分类准确率高 1.4 百分点,这是因为通过采用图上的随机游走过程可以游走到距离中心节点较远的节点处,可以更好地提取到直接邻域外关联度高的节点特征,提高特征表述的准确性。由实验可证图随机游走在此模型中是有效的。

图6 随机游走步数 r 对分类准确率的影响

在标准分割上,随机游走步数 r 依次为 5、10、15、20、25、30 和 35 的情况下,将所提网络 RWS-DHGNN、DHGNN 加超图构造加顶点卷积加图随机游走、DHGNN 加超图构造加顶点卷积加跳跃连接与 DHGNN 的测试集节点分类准确率进行比较,如图 7 所示。可以看出所提网络模型在大多数情况下均占据优势。例如在随机游走步数 r 为 10 时,RWS-DHGNN 的节点分类准确率相比 DHGNN 加超图构造加顶点卷积加图随机游走、DHGNN 加超图构造加顶点卷积加跳跃连接和 DHGNN 分别高出 1.2、0.3 和 2 个百分点。在随机游走步数 r 为 30 时,RWS-DHGNN 的节点分类准确率相比 DHGNN 加超图构造加顶点卷积加图随机游走、DHGNN 加超图构造加顶点卷积加跳跃连接和 DHGNN 分别提高 5.6、0.5 和 2.2 个百分点。这充分揭示了本文网络可以更好地捕获中心节点直接邻域外关联度高的节点特征,并且具有较强的表征能力。

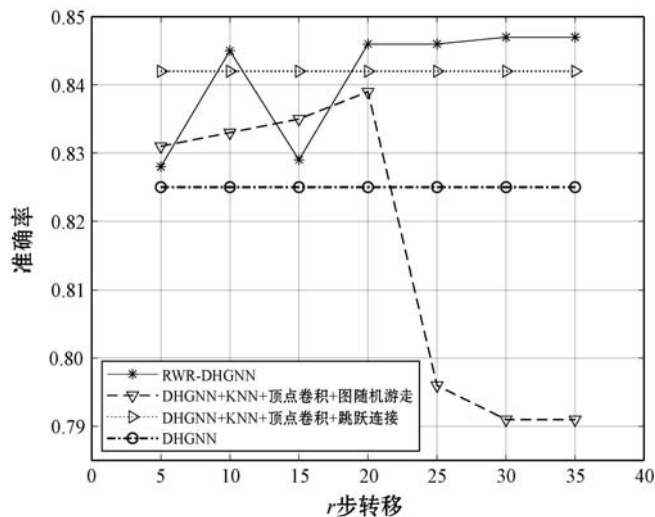


图7 四种网络分类准确率的对比

3 结语

本文针对 DHGNN 深入研究对其进行改进,通过

引入图随机游走以及在顶点卷积处增加跳跃连接构成残差结构,提出能够获取直接邻域外关联度高的节点特征和获得鉴别力更强的特征表述的 RWS-DHGNN 用于非欧几里得数据的分类。在 Cora 数据集的标准分割和随机分割上的实验表明,RWS-DHGNN 表现出较好的分类效果。下一步研究重心将探索解决全局特征信息不完整问题的其他方法,寻找更优的特征降维方法,进一步提高节点分类准确率。

参考文献

- [1] Zhou C, Shi B, Qiu H, et al. Graph convolutional architectures via arbitrary order of information aggregation[J]. IEEE Access, 2020 (8):92802-92813.
- [2] Ma T H, Wang H M, Zhang L J, et al. Graph classification based on structural features of significant nodes and spatial convolutional neural networks[J]. Neurocomputing, 2020, 423:639-650.
- [3] Wang X, Chai Y B, Li H, et al. Link prediction in heterogeneous information networks: An improved deep graph convolution approach [J]. Decision Support Systems, 2020, 141:1-12.
- [4] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks[EB]. arXiv:1609.02907,2016.
- [5] Velickovic P, Cucurull G, Casanova A, et al. Graph attention networks[EB]. arXiv:1710.10903,2017.
- [6] Feng Y, You H, Zhang Z, et al. Hypergraph neural networks[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2019, 33:3558-3565.
- [7] Jiang J, Wei Y, Feng Y, et al. Dynamic hypergraph neural networks[C]//28th International Joint Conference on Artificial Intelligence, 2019.
- [8] Cooper C, Radzik T, Siantos Y. Estimating network parameters using random walks [C]//2012 Fourth International Conference on Computational Aspects of Social Networks (CASoN), 2012:33-40.
- [9] Xia F, Liu J, Nie H, et al. Random Walks: A review of algorithms and applications[J]. IEEE Transactions on Emerging Topics in Computational Intelligence, 2019(99):1-13.
- [10] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]//IEEE Conference on Computer Vision & Pattern Recognition. IEEE, 2016.
- [11] Yang Z, Cohen W W, Salakhutdinov R. Revisiting semi-supervised learning with graph embeddings [C]//Proceedings of the 33rd International Conference on Machine Learning, 2016: 40-48.