

面向多中心数据的超图卷积神经网络及应用

周海榆 张道强

1 南京航空航天大学计算机科学与技术学院 南京 210016

2 南京航空航天大学模式分析与机器智能工信部重点实验室 南京 210016

(fishzzhou@163.com)

摘要 近年来,图神经网络在神经性脑疾病诊断中的应用引起了广泛关注。然而,现有研究中使用的图通常只是基于简单的点对点连接,无法反映3个或更多受试者之间的复杂关联,尤其是在多中心数据集中,即由不同医疗机构所使用的不同采集设备和不同受试人群而集成的具有异质性的数据集。为解决医疗影像数据中存在的多中心异质性问题,提出了一种多中心超图数据结构来描述多中心数据之间的关系。这种超图由两种不同的超边构成,一种是描述单个中心内部关系的中心内超边,另一种是描述不同中心之间关系的跨中心超边。另外,还提出了一种超图卷积神经网络来学习节点的特征表示,这种超图卷积由两部分构成,第一部分是超图节点卷积,第二部分是超边卷积。在两个多中心数据集上的实验结果证明了所提方法的有效性。

关键词: 多中心数据;数据异质性;脑疾病诊断;图卷积网络;超图卷积网络

中图法分类号 TP181

Multi-site Hyper-graph Convolutional Neural Networks and Application

ZHOU Hai-yu and ZHANG Dao-qiang

1 College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

2 MIIT Key Laboratory of Pattern Analysis and Machine Intelligence, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

Abstract Recently, the exploitation of graph neural networks for neurological brain disorder diagnosis has attracted much attention. However, the graphs used in the existing studies are usually based on the pairwise connections of different nodes, and thus cannot reflect the complex correlation of three or more subjects, especially in the multi-site dataset, i. e., the dataset collected from different medical institutions with the problem of data heterogeneity resulted from various scanning parameters or subject population. To address this issue, a multi-site hypergraph data structure is proposed to describe the relationship between multi-site data. This hypergraph consists of two types hyper-edge, one is intra-site hyper-edge that describes the relationship within the site, and the other is inter-site hyper-edge that describes relationship between different sites. Also, a hypergraph convolutional network is proposed to learn the feature representation of each node. The hypergraph convolution consists of two parts: the first part is the hypergraph node convolution, the second part is the super edge convolution. Experimental results on two multi-site datasets can also validate the effectiveness of the proposed method.

Keywords Multi-site dataset, Data heterogeneity, Brain diseases diagnosis, Graph convolutional networks, Hyper-graph convolutional networks

1 引言

图(graph)是一种用于描述不同样本间关联的数据表示形式,而图神经网络(Graph Neural Network, GNN)是目前研究的热点。Scarselli等最先提出图神经网络模型^[1],通过不断聚合一个节点和它邻域的节点特征来学习该节点的特征表示。随后,Bruna等提出基于图的拉普拉斯的谱域图卷积神经网络(Graph Convolutional Network, GCN)模型^[2]。Defferrard等提出用Chebyshev多项式来优化卷积计算过程^[3],

Kipf等使用一阶多项式代替Chebyshev多项式,以进一步优化计算复杂度^[4]。另外,非谱域的图卷积^[5-7]也被提出。

基于GCN的脑疾病诊断方法也是目前研究的热点之一。根据构建图的不同策略,基于GCN的脑疾病诊断研究大致可分为以下两类:第一类研究在大脑脑区层面上构建图,其中图的节点代表大脑脑区,图的边表示大脑脑区之间的连接^[8]。第二类研究将每个受试者视为图上的一个节点,而受试者之间的关联或相似性则用图的边来表示^[9]。第一类研究侧重于利用大脑脑区之间的相关性,第二类研究则是考虑

收稿日期:2020-11-23 返修日期:2021-12-08

基金项目:国家自然科学基金(61876082,61861130366,61732006)

This work was supported by the National Natural Science Foundation of China(61876082,61861130366,61732006).

通信作者:张道强(dqzhang@nuaa.edu.cn)

受试者之间关系。

尽管基于图的研究方法已在脑疾病诊断方面取得了不错的结果,但这些研究大多假设多中心数据来自同一分布^[8],没有考虑到多中心数据中存在的特性——由于扫描设备及参数或受试者群体的不同,数据存在很强的异质性。为了解决这一问题,Parisot 等利用被试之间的成对连接来寻找多中心数据之间的关联,提升了最终的疾病诊断性能^[9]。由于被试者之间的关系非常复杂,这种简单的成对关系很难反映出实际应用中不同被试者之间的复杂关联。因此,如何表示和利用这些关系是提升脑疾病诊断性能的关键。

超图(hyper-graph)是简单图的一种扩展形式,其在描述多个样本之间的关系上有着比图更优秀的性能。近年来,对超图的研究也取得了不错的进展。对超图的研究的第一步是构建超图,Huang 等使用 kNN 方法为每个节点选择最近的 k 个邻居节点来构造超边,然后用半径搜索法选择 k 的最优值^[10];Zhang 等提出了一种动态超图结构学习方案,其超图结构是不断更新的^[11]。Feng 等基于超图拉普拉斯提出了超图神经网络^[12]。Bai 等将注意力机制引入超图神经网络中,提出了基于注意力机制的超图神经网络^[13]。

基于以上考虑,我们利用超图来描述被试者之间的关系,并提出了一种多中心超图卷积神经网络(Multi-site Hyper-Graph Convolutional Networks, MH-GCN),利用多中心数据进行脑疾病诊断。所提 MH-GCN 由 3 个主要模块组成,分别是多中心数据对齐(Multi-site Data Alignment,MDA)、动态多中心超图构造(Dynamic Multi-site Hyper-Graph Construction,DMHG)和超图卷积(Hyper-Graph Convolution,HGC)。具体来说,MDA 模块通过将多中心数据投影到公共空间来解决数据的异质性问题;而 DMHG 模块用于构造两种不同类型的超边,即中心内超边(intra-site hyper-edges)和中心间超边(inter-site hyper-edges),分别表示同一个中心内受试者的关系和不同中心间受试者之间的关系。最后,利用 HGC 模块对超图的结构和特征信息进行编码,获得每一个节点的特征嵌入。另外,值得注意的是,所提模型可以根据每一层的特征嵌入动态更新超图结构,以获取更精确的超图。本文的主要贡献如下:

- (1)现有的很多研究忽视了多中心数据中存在的异质数据特性,而所提方法考虑到了这种特性,并且提出了一个 MDA 模块来进行数据对齐。
- (2)提出的 DMHG 模块通过构建中心内超边和中心间超边来表示不同受试群体之间的关系,符合数据的特性。
- (3)提出的 MH-GCN 可以根据每一次卷积学到的信息来更新超图结构,以更好地表示数据之间的关系。
- (4)分别在自闭症多中心数据集和精神分裂症多中心数据集上进行实验,实验结果证明了本文方法的有效性。

本文第 2 节介绍了所提多中心超图卷积神经网络框架;第 3 节介绍实验所选取的数据及实验结果和分析;最后总结全文并展望未来。

2 多中心超图卷积神经网络

本节将对所提出的多中心超图卷积神经网络框架进行

详细介绍。MH-GCN 由 3 个主要模块构成:多中心数据对齐(MDA)、动态多中心超图构造(DMHG)和超图卷积(HGC)。下面我们将依次介绍 MDA,DMHG 和 HGC 模块。

2.1 多中心数据对齐

一般来说,数据之间的相似度可以使用数据特征之间的欧氏距离或余弦距离等来描述。在来自单个中心的数据中,这种方法是可行的。然而,多中心数据具有很强的数据异质性,直接使用不同中心的受试者特征的距离来描述他们之间的相似性是不合理的,因为其无法正确地描述数据之间的关系。

为解决这个问题,提出了一个多中心数据对齐(MDA)模块来解决多中心数据之间的数据异质性问题。采用这种方法的直觉是,所有样本原本都应来自同一个分布,只不过由于不同的扫描设备或参数等使得数据产生了异质性,如果能够将所有数据投影到一个公共空间,则问题就迎刃而解了。

令 $\mathbf{X}_i = [\mathbf{x}_{i1}; \mathbf{x}_{i2}; \cdots; \mathbf{x}_{im}] \in R^{m \times d}$ 表示第 i 个中心的 m 个受试者的特征构成的矩阵,其中 $\mathbf{x}_{ij} (j = 1, \cdots, m)$ 表示第 i 个中心中第 j 个受试者的特征向量, d 代表特征维度。通过最小化不同中心数据之间的平均特征距离,可以训练得到一个变换矩阵 $\mathbf{T}$,即:

$$L_1 = \arg \min_{\mathbf{T}} \sum_{i=1}^n \sum_{j=1}^{m_i} \left\| \frac{1}{m_i} \mathbf{1}^T \mathbf{X}_i \mathbf{T} - \frac{1}{m_j} \mathbf{1}^T \mathbf{X}_j \mathbf{T} \right\| \tag{1}$$

其中, m_i 是第 i 个中心内的受试者人数。转换矩阵 $\mathbf{T}$ 可以将数据转换到一个公共空间,在这个公共空间中,数据之间的相似性可以通过计算特征距离来度量。

2.2 动态多中心超图构建

为了描述受试者之间的复杂关联,本文在受试者水平上构建超图,超图中的节点代表受试者,超边表示受试者之间的关联性。与已有的研究不同,本文提出了一种全新的多中心超图结构。具体来说,多中心超图由两种类型的超边构成,一种是每个中心内构建的超边,另一种是不同中心之间构建的超边。通过构造这种多中心超图,可以描述中心内和中心间受试者的关系。

令 $Con(e)$ 为超边 e 内节点组成的集合, $Adj(u)$ 表示所有包含节点 u 的超边组成的集合,即:

$$\begin{aligned} Con(e) &= \{u_1, u_2, \cdots, u_{k_e}\} \\ Adj(u) &= \{e_1, e_2, \cdots, e_{k_u}\} \end{aligned} \tag{2}$$

其中, k_e 和 k_u 分别表示超边 e 中的节点数和含有节点 u 的超边数,节点 u 表示超边集 $Adj(u)$ 的中心节点。

本文利用 kNN 方法和 k -means 方法构造多中心超图,这样就可以同时获取局部结构信息和全局结构信息。对于 k -NN 方法,我们选择距离节点 u 最近的 $k-1$ 个节点以及节点 u 本身,构造一条超边并加入 $Adj(u)$ 中。对于 k -means 方法,我们将节点聚类为 k' 个簇。最近的 $S-1$ 个簇以及包含 u 的簇将被选为 $Adj(u)$ 中的超边。具体的超边构建过程见算法 1。

算法 1 多中心超图构建算法

输入:中心数量 n ,每个中心的特征向量矩阵 $\mathbf{X}_i, i = 1, \cdots, n$,超边内采样节点数量 k
输出:超边集合 Adj
1. $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_n\}$

```
2. for i in range(n) do
3.   对 $\mathbf{X}_i$ 进行 k-means 聚类得到 m 个中心内聚类;
4.   对 $\mathbf{C}_X\mathbf{X}_i$ 进行 k-means 聚类得到 m' 个中心间聚类簇;
5.   for u in len( $\mathbf{X}_i$ ) do
6.     在 $\mathbf{X}_i$ 内计算离 u 最近的 k-1 个节点构造一条 kNN 超边 e 加入 Adj(u);
7.     选取含有 u 的簇以及 S-1 个最近的中心内聚类作为超边加入 Adj(u);
8.     在 $\mathbf{C}_X\mathbf{X}_i$ 内计算离 u 最近的 k-1 个节点构造一条 kNN 超边 e 加入 Adj(u);
9.     选取 S-1 个最近的中心间聚类作为超边加入 Adj(u);
10.  end for
11. end for
```

由于最初构造的超图可能不能准确地描述多中心数据,因此本文对每一层卷积获得的特征嵌入重复执行上述操作,随着卷积深度和训练轮数的增加,超图结构会得到动态更新,使其对数据之间关系的描述更加精确。

2.3 超图卷积

受到 Jiang 等的启发^[14],本文提出了一种多中心超图卷积,其由一个节点卷积模块和一个超边卷积模块构成。

2.3.1 超图节点卷积

节点卷积模块的作用是聚合超边内部节点的特征。聚合特征的方法有很多,如平均池化、最大池化或使用变换矩阵。但是这些简单的非学习方法无法捕捉到超边中具有判别性的节点特征。为了解决这个问题,本文从节点特性中学习一个矩阵 $\mathbf{P}$,用于节点特征的排序和加权。具体来说,我们采用多层感知机(Multi-Layer Perception,MLP)来学习变换矩阵 $\mathbf{P}$ 。变换矩阵 $\mathbf{P}$ 同时学习节点间的结构信息和节点上的特征信息。最后,利用一维卷积来生成超边特征。上述操作可表述为:

$$\mathbf{P} = \text{MLP}(\mathbf{X}_u)$$
$$\mathbf{x}_e = \text{conv}(\mathbf{P} \times \mathbf{X}_u)$$

(3)

其中, $\mathbf{X}_u \in R^{k \times d}$ 表示超边 e 中 k 个节点特征向量组成的矩阵, d 是特征维度, $\mathbf{x}_e \in R^d$ 表示超边 e 的特征向量。

2.3.2 超边卷积

超边卷积的功能是将超边集合内的超边特征聚集到中心节点上。利用多层感知(MLP)计算每个超边的权重得分。然后计算这些超边特征的加权和作为中心节点特征,上述过程可以表示为:

$$\mathbf{w} = \text{softmax}(\mathbf{x}_e \mathbf{W} + \mathbf{b})$$
$$\mathbf{x}_u = \sum_{i=1}^{|\text{Adj}(u)|} \mathbf{w}^i \mathbf{x}_e^i$$

(4)

其中, $\mathbf{x}_u$ 表示中心节点 u 的特征, $\mathbf{x}_e$ 表示含有中心节点 u 的超边的特征, $\mathbf{W}$ 和 $\mathbf{b}$ 是可学习的参数。

2.3.3 超图卷积

超图卷积由节点卷积和超边卷积组成。对于中心节点 u 的超边集合 $\text{Adj}(u)$ 中的每条超边 e ,采样 k 个节点,从而得到 $\mathbf{X}_u$ 。然后,节点卷积将节点特征 $\mathbf{X}_u$ 转换为超边特征 $\mathbf{x}_e$ 。在得到所有的 $|\text{Adj}(u)|$ 条超边的特征后,使用超边卷积将这些超边特征聚合为中心节点特征 $\mathbf{x}_u$ 。最后,利用非线性激活函数 σ 将 $\mathbf{x}_u$ 更新为 $\mathbf{y}_u$,即:

$$\mathbf{y}_u = \sigma(\mathbf{w} \mathbf{x}_u + \mathbf{b})$$

(5)

2.4 多中心超图神经网络

接下来介绍整个模型架构。首先将数据输入多中心数据对齐模块,以解决数据的异质性。其次使用多层动态超图构造模块和超图卷积模块得到每一个节点的特征表示。然后 $\mathbf{x}_u$ 通过非线性激活函数 σ 更新为 $\mathbf{y}_u$ 。最后加入交叉熵损失函数来训练模型,即:

$$L_2 = - \sum_i y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i)$$

(6)

其中, y_i 和 $\hat{y}_i$ 表示第 i 个样本的标签和模型的输出结果。最终损失函数可以描述为:

$$L = \gamma L_1 + L_2$$

(7)

其中, L_1 和 L_2 分别表示多中心数据对齐损失和交叉熵损失, γ 是调节这两个损失函数占比的系数。

3 实验结果与分析

本节首先介绍数据集和实验的相关设置。随后进行消融实验,分别研究 MDA 模块、DMHG 模块和使用不同种类超边的影响。最后,在两个多中心数据集上测试所提方法,并将实验结果与几种最新的方法进行比较。

3.1 数据集与预处理

我们在两个多中心数据集上进行实验,包括一个自闭症(Autistic Spectrum Disorder,ASD)数据集和一个精神分裂症(Schizophrenia,SZ)数据集。ASD 数据集来自(Autism Brain Imaging Data Exchange,ABIDE)^[15],我们选取了 5 个受试者人数多于 50 的中心数据。SZ 数据集则收集于 5 家不同的医疗机构。详细的数据统计信息如表 1 和表 2 所列。

表 1 ASD 数据集上的详细统计信息

Table 1 Detailed statistics information of ASD dataset

数据中心	患者	正常人
Leuven	25	32
NYU	71	93
UCLA	36	38
UM	48	65
USM	38	22

表 2 SZ 数据集上的详细统计信息

Table 2 Detailed statistics information of SZ dataset

数据中心	患者	正常人
COBRE	53	67
Huaxi	161	150
Nottingham	32	36
Taiwan	69	62
Xiangya	83	60

所有 rs-fMRI 图像的预处理都是通过 DPARSF 工具箱 2.0 版本中的软件包 SPM8 完成的。具体来说,预处理内容主要包括切片定时校正、运动校正和强度归一化。首先,采用干扰回归方法去除由呼吸、心脏搏动、头部运动和扫描仪漂移等引起的信号波动。然后,将图像与解剖自动标记(AAL)图谱对齐,提取每个受试者脑区 ROI 的平均时间序列。最后,基于皮尔逊相关系数,计算两两脑区之间相关的功能连接矩阵。由于连接矩阵是对称的,信息具有冗余性,因此去除矩阵的上三角和对角线值(即每个脑区与自身的相关性),会使其余部分被转换成向量作为每一个被试者的特征。

3.2 实验设置

本文所有的实验都是在联想 P320 图形工作站上实现,实验硬件环境为:英特尔 I7-7700K 型号 cpu 和单张 NVIDIA GTX1080TI 显卡,32 GB 内存;软件环境为:Ubuntu18. 04. 2, python3. 6,pytorch0. 4. 1。实验采取 10 折分层交叉验证,并报告平均结果。我们将 MDA 中转换矩阵 T 的维数设置为 6670×128 。整个模型采用两层的 DMHG 和 HGC 模块。在 DMHG 构建超边的过程中,对于 k -means 聚类方法,将聚类中心数设置为 10 并选取 2 个最近的簇作为超边;在 kNN 中,选取 8 个最近的节点构成一条超边。在节点卷积中,采样的节点数 k 设置为 8。其他相关参数为:丢弃率 0. 5,学习率 0. 01,权重衰减 0. 0005,衰减步长 200,衰减率 0. 7,训练轮数 100,激活函数 Relu。

3.3 消融实验

本节介绍对所提 MH-GCN 进行的消融实验,包括对 MDA 模块和动态构造超图模块,以及使用不同种类超边进行的实验以及结果分析。

3.3.1 MDA 模块的消融实验

为了研究所提 MDA 模块对模型的影响,对 MDA 模块进行了消融实验,消融模型中 MDA 模块被从完整模型中移除。从表 3 可以看到,使用 MDA 模块的完整 MH-GCN 模型在处理多中心数据方面表现更好,这是由于 MDA 对数据进行了对齐,在一定程度上减轻了数据异质性造成的影响。

表 3 消融实验结果
Table 3 Result of ablation experiments
(单位: %)

Method	Accuracy	
	ASD	SZ
移除 MDA	71. 40±3. 15	73. 16±4. 07
移除 DMHG	70. 37±2. 76	72. 47±3. 52
仅使用中心内超边	71. 88±2. 32	72. 29±3. 68
仅使用中心间超边	71. 32±3. 12	72. 35±4. 15
完整模型	72. 65±2. 41	74. 38±3. 98

3.3.2 动态构建超图的影响

为了研究动态超图是否有效,我们将所有的 DMHG 模块移除并使用固定的超图结构。具体来说,先用 kNN 和 k -means 构建一张初始的超图,之后始终使用这张超图。表 3 中的结果表明,移除了 DMHG 模块的模型性能比使用 DMHG 模块的完整模型差。与静态的超图结构相比,动态超图结构能更好地反映数据之间的关系,因此使用动态超图结构的模型具有更好的性能。

3.3.3 不同种类超边对结果的影响

在 DMHG 模块中,我们同时构建并使用了两种不同的超边,即中心内超边和中心间超边。为了研究不同种类的超边对模型的影响,进行了以下实验:1)仅使用中心内超边的 MH-GCN;2)仅使用中心间超边的 MH-GCN;3)同时使用两种超边的 MH-GCN 模型。从表 3 中可以发现,将两种类型的超边结合起来使用的方法具有更好的性能。这是因为两种超边分别关注了中心内样本关系和中心间样本关系,同时使用会使模型更完整。

3.4 完整模型的实验结果

我们将所提 MH-GCN 方法与几种方法进行了比较,

包括 NRF-CNN^[16],sGCN^[17],DP-GCN^[9],HGNN^[12],DHG-NN^[14]。NRF-CNN 是一种非 GNN 方法,它采用基于节点的行滤波卷积和结构保留池化的方法来诊断脑疾病。sGCN 和 DP-GCN 是目前脑疾病诊断中两种较新的 GCN 方法。HG-NN 和 DHGNN 是目前比较流行的两种超图神经网络方法,但尚未用于脑疾病分析。特别地,DP-GCN 利用一些额外的表型信息(如性别、年龄、基因等)来辅助建立模型,而本文方法并没有使用这些信息。为了公平起见,本文只比较了没有使用额外的表型信息 DP-GCN 的结果。实验结果以分类准确度(ACC)、敏感度(SEN)、特异度(SPE)和 ROC 曲线下面积(AUC) 4 个标准展示。

如表 4 所列,我们可以发现一些比较直观的信息。首先,与基于图的模型 NRF-CNN,sGCN 和 DP-GCN 相比,基于超图的方法获得了更好的结果,证明超图在表示多个样本间关系上优于简单图。相比简单图,超图中包含的超边能更好地表示多个样本之间的关系,可以促进节点间信息的传播。其次,在基于超图的方法中,DHGNN 和 MH-GCN 都采用了动态学习超图的策略,因此比只使用静态超图结构的 HGNN 具有更好的性能。初始静态超图结构可能并不能准确地描述节点之间的关系,而动态的超图结构则可以根据模型的训练进行调整,以更加适应数据。另外,与不考虑数据异质性的 DHGNN 方法相比,所提 MH-GCN 方法的性能更好。与 DHGNN 相比,MH-GCN 模型采用了 MDA 模块来解决数据异质性,并构造了两种不同种类的超边,在处理多中心数据时更加有效。

表 4 对比实验结果
Table 4 Result of the compared methods
(单位: %)

Dataset	Method	ACC	SEN	SPE	AUC
ASD	NRF-CNN	69. 67±4. 00	72. 41±5. 03	67. 01±6. 59	68. 22±6. 10
	sGCN	62. 83±3. 40	66. 82±2. 16	58. 89±7. 11	61. 06±5. 88
	DP-GCN	67. 32±3. 75	69. 43±5. 16	66. 30±5. 48	68. 17±6. 32
	HGNN	69. 88±3. 53	73. 37±4. 65	66. 13±5. 53	68. 78±5. 80
	DHGNN	70. 10±3. 15	74. 49±4. 49	65. 41±4. 35	69. 32±4. 74
	MH-GCN (ours)	72. 65±2. 41	79. 60±3. 96	65. 72±6. 01	71. 44±5. 72
SZ	NRF-CNN	70. 38±4. 16	73. 33±3. 74	67. 44±5. 77	68. 25±4. 52
	sGCN	68. 83±3. 91	72. 06±4. 37	65. 65±7. 73	65. 08±4. 87
	DP-GCN	69. 17±3. 25	73. 02±4. 72	66. 13±5. 85	67. 33±5. 96
	HGNN	69. 73±4. 21	72. 41±5. 07	67. 01±6. 34	68. 17±6. 35
	DHGNN	70. 76±4. 07	74. 59±3. 48	67. 96±6. 31	70. 01±5. 04
	MH-GCN (ours)	74. 38±3. 98	79. 28±3. 99	69. 54±5. 32	70. 75±4. 67

3.5 超边质量分析

为了进一步探索模型分类的情况,我们对 DHGNN 和 MH-GCN 构建的超边内的节点进行数据统计。本文提出了一种评价超边质量的指标,称为有效百分比(Effective Percentage,EP),即正负样本差占总样本的比率,可表示为:

$$EP = \frac{|\sum_{x \in S^+} x - \sum_{x \in S^-} x|}{|Con(e)|}$$

(8)

其中, $x \in S^+$ 表示正样本, $x \in S^-$ 表示负样本。表 5 列出了两个数据集上不同类型超边的平均 EP 值。

表 5 不同类型超边的 EP 值

Table 5 EP value of each type of hyper-edge

(单位: %)

超边类型	EP	
	ASD	SZ
DHGNN 中的超边	53.75	52.68
MH-GCN 中的中心内超边	57.42	59.37
MH-GCN 中的中心间超边	56.36	56.84

从表中可以看出,由于不考虑数据的异质性,DHGNN 中而超边 EP 值相对较低,无法准确地表示多中心之间的相关性。所提 MH-GCN 中的中心内超边和中心间超边的 EP 都相对较高,因为它们可以有效地表示中心内和中心间数据的关系。

结束语 本文提出了一种多中心超图卷积神经网络(MH-GCN)模型,针对多中心数据进行脑疾病诊断。该模型包括 3 个主要模块:多中心数据对齐(MDA)模块、动态多中心超图构造(DMHG)模块和超图卷积(HGC)模块。具体来说,为解决多中心数据之间的异质性问题,MDA 模块学习一个转换矩阵将来自不同中心的数据转换到公共空间。在此基础上,DMHG 通过构建每个中心内的超边和不同中心之间的超边来表示被试者之间的关系。由于初始的超图并不能准确地表示数据间的关系,因此 DMHG 模块在每一层上动态更新超图结构。HGC 模块通过节点卷积和超边卷积来学习节点的特征表示。在两个多中心数据集上的实验结果也证明了所提方法的有效性。未来的工作将考虑利用数据集中额外包含的一些表型信息来更好地建立超图,提高模型的分类精度。

参 考 文 献

[1] SCARSELLI F,GORI M,TSOI A C,et al. The Graph Neural Network Model[J]. IEEE Transactions on Neural Networks, 2009,20(1):61-80.

[2] BRUNA J,ZAREMBA W,SZLAM A. Yann LeCun:Spectral Networks and Locally Connected Networks on Graphs[C]//International Conference on Learning Representations. 2014.

[3] DEFFERRARD M,BRESSON X,VANDERGHEYNST P. Convolutional neural networks on graphs with fast localized spectral filtering [C]//Advances in Neural Information Processing Systems. 2016:3844-3852.

[4] KIPF T N,WELLING M. Semi-supervised classification with graph convolutional networks[C]//International Conference on Learning Representations. 2017.

[5] ATWOOD J,TOWSLEY D. Diffusion convolutional neural networks[C]//Advances in Neural Information Processing Systems. 2016:1993-2001.

[6] DUVENAUD D K,MACLAURIN D,IPARRAGUIRRE J, et al. Convolutional networks on graphs for learning molecular fingerprints [C]//Advances in Neural Information Processing Systems. 2015:2224-2232.

[7] HAMILTON W,YING Z T,LESKOVEC J. Inductive representation learning on large graphs [C]//Advances in Neural Information Processing Systems. 2017:1024-1034.

[8] K TENA S I,PARISOT S,FERRANTE E,et al. Distance metric learning using graph convolutional networks: Application to functional brain networks [C] // International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer,2017:469-477.

[9] PARISOT S,K TENA S I,FERRANTE E,et al. Disease prediction using graph convolutional networks: application to autism spectrum disorder and Alzheimer’s disease[J]. Medical Image Analysis,2018,48:117-130.

[10] HUANG Y C,LIU Q S,METAXAS D. Video object segmentation by hypergraph cut [C]//IEEE Conference on Computer Vision and Pattern Recognition. 2009:1738-1745.

[11] ZHANG Z Z,LIN H J,GAO Y,et al. Dynamic Hypergraph Structure Learning [C]//Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence. 2018: 3162-3169.

[12] FENG Y F,YOU H X,ZHANG Z Z,et al. Hypergraph neural networks [C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2019:3558-3565.

[13] BAI S,ZHANG F H,TORR P H S. Hypergraph convolution and hypergraph attention[J]. arXiv:1901.08150,2019.

[14] JIANG J W,WEI Y X,FENG Y F,et al. Dynamic hypergraph neural networks[C]//Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. 2019:2635-2641.

[15] MARTINO A D,YAN C G,LI Q,et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism[J]. Molecular Psychiatry, 2014,19(6):659-667.

[16] MAO B C,HUANG J S,ZHANG D Q. Node based row filter convolutional neural network for brain network classification [C]//Pacific Rim International Conference on Artificial Intelligence. 2018:1069-1080.

[17] K TENA S I,PARISOT S,FERRANTE E,et al. Metric learning with spectral graph convolutions on brain connectivity networks [J]. NeuroImage,2018,169:431-442.



ZHOU Hai-yu, born in 1994, postgraduate. His main research interests include computer vision and medical image analysis.



ZHANG Dao-qiang, born in 1978, professor, Ph.D supervisor, is a member of China Computer Federation. His main research interests include machine learning, pattern recognition, data mining and medical image analysis.