

www.pibb.ac.cn



基于TI-GNN的青少年吸烟成瘾诊断*

王旭雯1) 喻大华2)** 薛 婷3)** 李晓娇1) 麦珍珍1) 董 芳1)

马宇欣1) 王 娟1) 袁 凯1,2,4)

(¹⁾ 内蒙古科技大学数智产业学院,包头 014010;²⁾ 内蒙古科技大学自动化与电气工程学院,包头 014010;
 ³⁾ 内蒙古科技大学理学院,包头 014010;⁴⁾ 西安电子科技大学生命科学与技术学院,西安 710071)

摘要 目的 烟草相关疾病是全球可预防的主要健康问题之一,且它们是导致过早死亡的主要原因之一。而且吸烟成瘾作 为一种慢性大脑疾病,已广泛被认定为影响大脑结构和功能的关键因素。然而,目前有效的诊断方法仍存在挑战。为了更 好地理解吸烟成瘾的神经机制,并提高诊断的准确性,本研究提出了一种新型的图神经网络框架——TI-GNN,旨在通过功 能磁共振成像(fMRI)数据揭示吸烟成瘾与大脑连接异常之间的关系。方法 本研究基于fMRI数据,利用图神经网络 (GNN)对吸烟成瘾的功能连接模式进行建模。TI-GNN通过Transformer提取全局交互信息和空间注意机制有效获取脑区之 间的联系,以提高模型的诊断性能。此外,模型内置因果解释模块,以深入挖掘大脑不同区域的因果关系,从而增强模型 的可解释性。结果 实验结果表明,TI-GNN模型在吸烟成瘾数据集上的分类效果显著优于现有的最佳基线方法。特别地, TI-GNN在提高区分效果、准确识别吸烟成瘾与健康对照之间的差异方面表现出色,准确率、F1分数和马修斯系数分别达 到0.91、0.91和0.83。同时揭示了杏仁核、前扣带皮层等关键脑区的异常连接模式,与临床研究结果一致。结论 TI-GNN 框架为吸烟成瘾的客观诊断提供了高效工具,其揭示的脑网络异常与因果关联机制,深化了对成瘾病理机制的理解,为靶 向干预策略和个性化治疗奠定了重要理论基础。

关键词 空间注意力, Transformer, 图神经网络, 功能磁共振成像, 分类, 青少年, 吸烟成瘾
 中图分类号 TP391, R338 DOI: 10.3724/j.pibb.20250107 CSTR: 32369.14.pibb.20250107

全球约有1/5的人口处于吸烟状态¹¹,全球吸 烟人数从 1980 年的 4.96 万亿^[2] 上升至 2016 年的 5.5万亿^[3]。随着卷烟消费的进一步增加,预计到 2025年这一数字将达到9万亿支^[4]。吸烟不仅是导 致一系列严重健康问题的主因,还与全球范围内的 过早死亡密切相关。每6秒就有一人死于烟草相关 疾病,烟草致死占全球死亡人数的1/5^[5]。更为严 重的是,吸烟者的平均寿命比不吸烟者短10年以 上^[6]。特别是18~25岁的年轻人群体,其吸烟率较 高,且从青少年时期开始吸烟的人群更容易发展为 成年后的烟草依赖^[7]。过去研究表明^[8],青春期 大脑中会发生一系列重要而敏感的神经生理学变 化,吸烟引起的尼古丁摄入对大脑的影响显著,甚 至会造成脑网络可控性的异常改变^[9]。这一群体 的吸烟行为不仅增加了他们日后成为烟草依赖者的 风险,还可能提高他们对其他成瘾药物的使用

概率。

成瘾是一种影响大脑奖赏系统、记忆、动机和 行为控制的慢性疾病,其特征是大脑结构和功能的 深刻改变^[10-11]。因此,探索吸烟成瘾的神经机制 对于更好地预防和治疗烟草依赖具有重要意义。

功 能 磁 共 振 成 像 (functional magnetic resonance imaging, fMRI)^[12] 作为一种非侵入性

^{*} 国家脑科学和类脑智能技术计划(2022ZD0214500),科技创新 2030项目重大项目(2022ZD0207100),国家自然科学基金 (82260359,82371500,U22A20303,61971451),内蒙古自然科 学基金(2023QN08007,2025MS08027,2025MS08098),内蒙古 高校基本科研业务金(2023QNJS204,2023QNJS206,2024QN-JS119)和内蒙古高校青年科技人才发展计划(NJYT24030)资助。 **通讯联系人。

喻大华 Tel: 0472-5950060, E-mail: fmydh@imust.edu.cn 薛婷 Tel: 0472-5950060, E-mail: xueting41@imust.edu.cn 收稿日期: 2025-3-14, 接受日期: 2025-07-09

的神经成像技术,近年来已成为研究大脑功能神经 模式的主流方法之一。fMRI通过血氧水平依赖 (blood oxygen level dependent, BOLD)信号来量 化大脑区域之间的功能连接^[13],为揭示吸烟成瘾 相关的大脑生物标志物提供了重要依据。随着人工 智能技术的飞速发展,机器学习方法在fMRI信号 的分析和生物标志物提取中得到了广泛应用。通过 对fMRI时间序列中的两两相关性进行量化,功能 连接矩阵可以作为区分吸烟成瘾者和健康对照组的 重要特征。

然而,目前基于功能连接矩阵的成瘾诊断模型 大多采用传统的两阶段训练策略(图1a),其中包 括特征选择和浅层分类模型,如支持向量机 (support vector machines, SVM)^[14]、LASSO 回 归^[15]和随机森林(random forests, RF)^[16]等。这 些方法虽然可以有效地处理某些类型的特征数据, 但它们在捕捉复杂的大脑网络拓扑结构和非线性信 息方面存在局限性,且过于依赖特征选择的准确性 和分类器的性能,因此在实际应用中可能出现较高 的误差和不稳定性。此外,传统的深度学习框架如 卷 积 神 经 网 络 (convolutional neural network, CNN)和循环神经网络(recurrent neural network, RNN)也未能充分考虑大脑网络的非欧几里得几 何特性,这限制了它们在大脑网络分析中的表现。



(b) Modern addiction diagnosis model with built-in interpretable graph neural networks

Fig. 1 Flowchart of traditional model and emerging model framework

The fMRI of the resting state is segmented by automated anatomical labeling (AAL) atlas to generate a functional connection matrix. The traditional addiction diagnosis model is to extract features from the data through a two-stage training strategy, and then classify the features through a shallow classification model. For modern addiction models with built-in interpretable networks (*e.g.* TI-GNN), the functional connection matrix is transformed into a graph that is input to TI-GNN for diagnosis. In addition, TI-GNN can detect potential biomarkers that influence addiction.

近年来,基于图神经网络的统一的端到端吸烟 成瘾诊断模型成为一种新兴的研究趋势。通过将大 脑建模为图结构,其中节点代表大脑的感兴趣区 (region of interest, ROI),边缘则表示这些区域 之间的功能连接, 图神经网络能够有效地捕捉大脑 网络中的复杂模式。尽管这种方法在性能上已有显 著提升,但目前大多数图神经网络模型仍然存在 "黑盒"问题,缺乏足够的可解释性[17-19],这使得 这些模型的决策过程难以理解和信任。此外,图神 经网络是分析结构化数据的重要模型,在社交网络 和生物信息学等多个领域得到了广泛的应用。其中 图注意力神经网络 (graph attention networks, GAT)通过注意力机制赋予不同邻居节点不同的权 重,从而实现更加精确的特征聚合。但GAT依赖 邻域聚合,仅能捕捉局部信息,难以学习远程节点 之间的联系,且主要聚焦于局部子图,难以提取全 局结构信息,从而影响整体的分类效果。

为了解决上述问题,本文提出了基于 Transformer 的可解释图神经网络(transformerbased and interpretable graph neural network, TI-GNN),旨在用于吸烟成瘾的诊断。与传统的两阶 段训练策略不同,TI-GNN能够直接从原始fMRI 数据中提取深层次特征,区分吸烟成瘾者与健康对 照组,并进一步识别出导致吸烟成瘾的生物标志物 (图1b)。为了提高模型的准确性和鲁棒性,本框 架引入了Transformer^[20]和空间注意力机制,以便 更好地捕捉数据中节点间的关系和全局信息。通过 在获取全局信息和空间维度上对大脑连接模式进行 建模,TI-GNN显著提升了对吸烟成瘾的分类性 能。此外,TI-GNN还内置了因果解释模块,通过 对神经网络的决策过程进行可解释性分析,帮助识 别影响吸烟成瘾的关键大脑区域和异常连接模式, 从而为临床诊断和干预提供更为直观的依据。

1 吸烟成瘾分类框架

本文提出的吸烟成瘾分类框架如图2所示,旨 在通过一系列步骤将fMRI数据转化为有用的分类 特征。首先,对原始fMRI数据进行一系列预处理 操作,包括去噪、标准化等。接着,使用自动解剖 标记图谱提取感兴趣的脑区,并计算这些脑区的平 均时间序列。基于此,进一步生成功能连接矩阵, 这些矩阵反映了各脑区之间的功能连接关系。然 后,图生成模块将这些矩阵转化为图结构,图神经 网络通过这一结构进行学习。

模型的主要流程包括: a. 将图结构输入到图 CNN模块; b. 通过Transformer模块捕获数据的全 局信息,增强节点间的关系; c. 输出通过因果解释 模块进行因果关系分析并且通过空间注意力机制进 行聚合; d. 最终分类结果通过输出模块给出。



Fig. 2 The overall framework of the TI–GNN model

1.1 图的构建

图神经网络因具有出色的图表示学习能力,被 广泛应用于各种图任务,包括图分类^[21]、边的预 测^[22]和对节点分类^[23]。本实验中,任务是基于大 脑网络进行图分类。我们将大脑分割为若干个 ROI,每个脑区对应一个图中的节点,节点之间的 边权则由脑区之间的功能连接来定义。

具体地,将大脑划分为N个ROI,定义无向加 权图为 $G_n = (V, E)$,图的大小为 $N \times N$ 。其中 $V = \{v_1, \dots, v_N\}$ 是一组表示大脑区域的节点,E = $\{(V_i, V_j)|V_i, V_j \in V\}$ 是一个邻接矩阵,其中每个边 $E_{i,j} \in E$ 表示每对大脑区域 $V_i n V_j$ 之间的连接强度, 该结构通常表示为邻接矩阵的形式 $A \in R^{N \times N}$ 。网络 利用聚合策略A来学习G中节点v的表示形式并进 一步使用READOUT策略R来学习图G的表示。它 们的信息传递过程定义为:

$$h_{v}^{(k)} = A^{(k)} \Big(\Big\{ h_{v}^{(k-1)} : u \in N(v) \Big\} \Big)$$
(1)

$$h_{G} = \mathbf{R}\left(\left\{h_{v}^{(k)}|v \in G\right\}\right) \tag{2}$$

其中 $h_v^{(k)}$ 表示节点v在第K层的嵌入向量, h_c 表示

整个图, N(v)表示的是相邻节点的集合, $A \in \mathbb{R}^{N \times N}$ 为图的邻接矩阵。

1.2 基于Transformer的特征增强

在经过节点嵌入后,通过Transformer进一步 强化节点的特征表示,使其能够综合考虑远程节点 信息(图3)。

Transformer 采用多头自注意力机制,通过自适应加权的方式,使得节点不仅能获取邻居节点的影响,还能从远程节点中获取信息。将节点特征 $X \in R^{n \times d}$ 映射到Q、K 和 V。

$$Q = XW_0 \tag{3}$$

$$K = XW_{\kappa} \tag{4}$$

$$V = XW_{V} \tag{5}$$

其中 W_{q} , W_{κ} , $W_{v} \in R^{d \times d}$ 。通过计算Q和K的相似性, 生成注意力权重:

Attention
$$(Q,K,V) = softmax \left(\frac{QK^{T}}{\sqrt{d_{k}}}\right) V$$
 (6)

这里*d*_{*k}</sub>是缩放因子,来保证数值稳定性。该注 意力机制能够捕捉远程节点的相关性,使得节点表 示更加全面。此外,使用残差连接来增强流动梯度 <i>X*'。随后经过前馈神经网络进行特征变换*X*["],最终 得到增强后的节点嵌入*X*_{new}。</sub>

$$X' = LayerNorm(X + Dropout(MSA(X))) (7)$$

$$X'' = ReLu(X'W_1 + b_1)W_2 + b_2$$
(8)

$$X_{new} = LayerNorm\left(X' + Dropout\left(X''\right)\right) \quad (9)$$

在计算完节点嵌入后,对所有的节点进行全局 池化,来生成整个图的表示:

 $\widetilde{H}_{v} = Concat(Readout_{1}(X), Readout_{2}(X), \cdots)(10)$ 1.3 基于空间注意机制的聚合函数

模型提出一种新的空间注意机制,来获取空间 注意力向量*Z*∈*R*^{*},仅限于图中相邻节点对之间, 图的拓扑结构依赖于节点之间的邻接关系,体现了 局部的空间感知能力,能有效避免全局注意力导致的冗余与过度平滑问题,增强节点在图中的语义表达能力。该向量旨在聚合节点表示有效地融合了空间信息,以产生最终的全局图表示 $H_c \in R^d$ 。本文采用了一种类似自注意机制的方法,涉及查询和关键向量的初始生成:

$$q = \varphi_{mean} \tilde{H} W_q \tag{11}$$

$$K = \tilde{H}W_{K}$$
(12)
其中 $W_{K} \in \mathbb{R}^{d \times d}$, $W_{q} \in \mathbb{R}^{d \times d}$ 以及 $\varphi_{mean} =$

 $\left[\frac{1}{N}, \cdots, \frac{1}{N}\right]$ 。注意力向量由公式(13)所示:

$$Z = \sigma(\frac{qK}{\sqrt{d}}) \tag{13}$$

此处*σ*(.)表示 sigmoid 激活函数,基于获得的 注意力向量,可以通过以下过程确定为最终的全局 图表示:

$$H_{c} = \frac{1}{N} \sum_{i=1}^{N} z_{i} h_{i}$$
(14)

1.4 图形编码器和因果效应估计量

本模块旨在促进模型学习更具有表现力和适用 性的整体大脑表征。图形分自编码器模块旨在捕捉 图的结构信息和节点间的关系。因果效应估计量模 块来识别和量化变量之间的因果关系,在图结构中 通过分析不同节点之间的因果关系,估计干预对特 定结果的影响。通过上述方式,该模块有助于降低 过拟合的风险,并提高模型泛化的能力。详细细节 如下。

1.4.1 图形编码器

给定一个输入图G = (X, A),其中 $A \in R^{n \times n}$ 是 图的邻接矩阵, $X \in R^{n \times n}$ 为节点特征矩阵,使用图 形编码器得到重构图 $G = (X_c, A_c)$ 。此模块的编码 器是一个基本的图卷积神经网络。具体地,第k层 中输出的Z可以定义为:

$$Z^{k} = \sigma(\tilde{A}Z^{k-1}W^{k-1}) \tag{15}$$



其中, $\tilde{A} = D^{-\frac{1}{2}} \hat{A} D^{-\frac{1}{2}}$, $\hat{A} = A + I$, $D = \sum_{i} \widehat{A}_{ij}$ 是一个 对角矩阵。此外, W是一个可训练矩阵, $\sigma(.)$ 为激 活函数。

此模块改进解码器具有独立的头部:多层感知器用于重构*X*,线性内积解码器用于恢复变量。重建的公式为:

$$A_c = \sigma(ZZ^T) \tag{16}$$

$$X_c = MLP(Z) \tag{17}$$

图形编码器的损失函数为:

$$L_{VAE} = E\left[\left|\left|X - X_{c}\right|\right|_{F}\right] + E\left[\left|\left|A - A_{c}\right|\right|_{F}\right] - E\left[D_{KL}\left[q\left(Z|A,X\right)||p\left(Z\right)\right]\right]$$
(18)

其中q(ZlA, X)表示编码器模型, II.II_F表示费罗贝 尼乌斯范数, p(Z)表示先验分布, 假设它遵循各 向同性高斯分布。

1.4.2 因果效应估计量

图编码器学习到的隐藏表示 $H \in R^{n \times d}$ 可以分解 为 $\alpha \in R^{n \times k}$ 和 $\beta \in R^{n \times l}$ 两部分,其中 α 表示影响分 类结果的因果因素, β 表示非因果因素(图4)。这 里k和l是约束k + l = d的超参数。



Fig. 4 Causality diagram

The graph shows a directed acyclic graph, reflecting a causal relationship between latent factor α and label *Y*, and a spurious correlation between β and label *Y*.

为了使模型学习到 α 对分类预测Y的因果影响 最大化,采用了信息流 $I(\alpha \rightarrow Y)$ 作为定量评估 α 对Y因果影响的特定度量,根据现有研究^[24],从 α 到Y的信息流可以看作是 α 与Y之间的互信息,表 示为 $I(\alpha \rightarrow Y) = I(\alpha; Y)$ 。因此,可以通过显式 最大化 $I(\alpha; Y)$ 。对于非因果因子 β ,我们通过最 小化 $I(\alpha; \beta)$ 。综上所述,该模块的损失函数为: $L_{caugal} = -I(\alpha; Y) + I(\alpha; \beta)$ (19)

1.5 分类器

在模型的分类阶段,本文设计了一种基于多层 感知机(multi-layer perceptron, MLP)的图级分 类器。该分类器以图神经网络模块输出的全局图嵌 入作为输入,经过多层线性变换、非线性激活函数 和Dropout 正则化,有效挖掘高阶特征表示并抑制 过拟合风险。最终,输出层使用 Softmax 函数生成 每个图样本所属类别的预测概率。模型训练过程 中,采用交叉熵损失函数作为目标函数,以优化预 测结果与真实标签之间的匹配程度。

$$y = MLP(H_G) \tag{20}$$

将第*i*个个体的大脑网络表示为*H*⁽ⁱ⁾,分类的损 失函数可以用如下方式表示:

$$L_{ce} = -\sum_{i=1}^{C} y_i \log(y'_i)$$
(21)

其中 y'_i 是模型输出的softmax概率, C是类的数量。

1.6 总体损失函数

综合上述,最终的损失函数可以表示为:

$$L = L_{ce} + L_{VAE} + \lambda L_{causal}$$
(22)

其中λ是超参数。

2 实验及结果分析

2.1 数据集与预处理

2.1.1 参与者

本研究在内蒙古科技大学筛选了55 例男性吸烟者,其筛选符合《精神障碍诊断与统计手册》第五版尼古丁依赖的诊断标准^[25]。尼古丁依赖测试(Fagerstrom test for nicotine dependence, FTND)用于测量尼古丁依赖的严重程度,评估青少年吸烟成瘾者的尼古丁依赖水平。同时筛选年龄、教育程度匹配的55 名健康男性非吸烟者。所有被试均为右利手且民族为汉族。相关信息见表1。青少年吸烟者排除标准为: a. 根据临床评估的任何身体疾病,如脑瘤、阻塞性肺病、肝炎或癫痫; b. 目前可能影响认知功能的任何药物; c. 使用酒精、咖啡或任何药物滥用; d. 存在神经系统疾病; e. 患有幽闭恐惧症等。

在充分解释实验流程后,所有被试均签署知情同意书。本研究经过内蒙古科技大学包头医学院第一附属医院医学伦理委员会批准(批件号:2020032601)。

Table 1 Basic statistics for the dataset

Dataset	Smoking	Control	P
Dataset	group	group	1
Number	55	55	-
Age range	16-23	16-23	-
Age (mean±std)	19.69±1.99	$20.11{\pm}1.97$	0.56
Education years (mean±std)	14.53 ± 0.75	14.21 ± 0.81	0.18

2.1.2 数据采集

在被试招募工作以及对问卷计算完成后,将所 有符合标准的被试分为吸烟组和健康对照组,分别 对其进行fMRI扫描,每个参与者被要求保持放松, 闭上眼睛,不去想任何特别的事情。实验在包头医 学院第一附属医院的3.0 T Philips 磁共振扫描仪进 行。磁共振扫描仪带有八通道相位阵列头线圈。通 过以下参数获得实验者的T1 加权像:重复时间 (*TR*) =8.5 ms,回波时间(*TE*) =3.4 ms,翻转角 (*FA*) =12°,矩阵=240×240,视野(*FOV*) =256 mm×256mm,切片厚层=1.0 mm,层数=140。

2.1.3 数据处理

静息态功能磁共振数据采用统计参数映射 (statistical parametric mapping, SPM)软件进行预 处理,预处理过程包括丢弃前10卷、切片定时矫 正、头动矫正、空间归一化、平滑(6 mm)、时间 带通滤波(0.01~0.1 HZ)、回归协变量、白质和脑 脊液信号的影响。

在进行预处理后,根据自动解剖标记图谱提取 ROI的平均时间序列。在此基础上,利用皮尔逊相 关系数计算ROI之间的功能连通性,如式(23)所 示,生成116×116的功能连接矩阵。

$$P_{uv} = \frac{Cov(R_u, R_v)}{\sigma_{R_u} \sigma_{R_v}}$$
(23)

其中,时间序列 $R \in R^{N \times T}$, $P \in R^{N \times N}$

2.2 实验细节

本研究使用 PyTorch 搭建网络,在8 GB的 NVIDIA GeForce RTX 4060的计算设备训练。对于

$$TP \times TN - FP \times FN$$

$$\sqrt{(\text{TP} + \text{FP}) \times (\text{TP} + \text{FN}) \times (\text{TN} + \text{FP}) \times (\text{TN} + \text{FN})}$$

2.4 结果与分析

MCC = -

表2展示了不同模型分类性能方面的ACC、F1 分数和MCC。随机划分数据集,80%用于训练, 10%用于验证,剩下的10%用于测试。本文对比的 模型方法有 SVM^[14]、NB^[27]、RF^[16]、GBDT^[28]、 LASSO^[15]、GCN^[29]、GAT^[30]、GIN^[31]、CI-GNN^[32]、DIR-GNN^[33]和BrainGNN^[34]。 模型训练,最大训练周期设置为300,数据输入的 batch_size设置为32,利用Adam优化器^[26]对训练 过程进行优化,初始学习率为0.001,随机失活率 设置为0.1,权重衰减设置为0.0005。图卷积网络 (graph convolutional networks, GCN)、GAT和图 同构神经网络(graph isomorphism network,GIN) 设置为3层。对于TI-GNN, λ 的值被设置为0.001, 预定义特征维度k和l分别设置为56和8。对于基 线模型,实验采用网格搜索或推荐设置以确定最终 的设置。对于GCN、GAT和GIN,实验使用相关 工作中推荐的超参数来训练模型。

2.3 评价指标

采用准确率(accuracy, ACC)、F1分数(F1 score)和马修斯相关系数(Matthews correlation coefficient, MCC)作为网络性能评价指标。准确 率作为分类问题中最常见的评价指标,它是正确预 测的样本与样本总数的比率。对于二元分类问题, 准确率定义如下:

$$ACC = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}}$$
(24)

其中TP代表真阳性(true positive), TN代表真阴 性(true negative), FP代表假阳性(false positive), FN代表假阴性(false negative)。

在类别平衡的情况下,ACC可以准确衡量数 据集的预测效果。但在数据严重失衡的情况下则会 影响结果。为此,本文引入了F1分数和MCC。其 中F1分数是分类器准确率和召回率的调和平均值, F1分数定义如下:

F1 score =
$$\frac{\text{TP}}{\text{TP} + \frac{1}{2}(\text{FP} + \text{FN})}$$
 (25)

MCC 是一个综合性能指标,它同时考虑了4 个分类结果,包括真反例和真正例,可以在不平衡 的数据集中提供更准确的性能评估,MCC 定义 如下:

(26)

大量实验表明,在吸烟数据集上,本实验中的 模型与所有对比模型上的所有评价指标都有显著的 改进。与经典的非深度学习方法 SVM-rbf 相比, TI-GNN 在准确率上提升了 0.27, F1 分数显著提高 了 0.41, MCC 提升了 0.62。在所有深度学习方法 中,TI-GNN 也具有突出的性能,与近几年的模型 相比,TI-GNN 具有明显的优势,即使跟 2024 年的 最先进的方法 CI-GNN 相比,在 ACC、F1 分数和 MCC 上分别高出了 0.18、0.13 和 0.27。从表 2 可以 看出在 ACC 达到最高的同时,F1 分数和 MCC 更是 优于当前最先进方法,说明本模型在面对数据不平 衡的情况,也能够表现出比其他方法更加卓越且稳 定的性能表现。

Table 2	Comparison	of	classification	results	of	different
	models in the	sm	oking addicti	on datas	set	

Method	ACC	F1 score	MCC
SVM-rbf	0.64	0.50	0.21
SVM-linear	0.55	0.44	0.10
NB	0.82	0.67	0.62
RF	0.73	0.75	0.61
GBDT	0.64	0.64	0.28
LASSO	0.73	0.67	0.48
GCN	0.64	0.50	0.26
GAT	0.73	0.73	0.46
GIN	0.73	0.67	0.45
BrainGNN	0.82	0.83	0.69
DIR-GNN	0.82	0.80	0.63
CI-GNN	0.73	0.78	0.56
TI-GNN	0.91	0.91	0.83

Notes: The best results are shown in bold.

2.5 消融实验与分析

2.5.1 超参数分析

本实验使用青少年吸烟成瘾数据集研究了预定 义特征维度k和l对结果的影响。实验通过采用不 同比率k/(k+l)和不同的特征维数z对TI-GNN模 型进行训练。可以观察到,当k/(k+l)设置为0.8 (图5)和z = k + l = 64(图6)时,TI-GNN在青 少年吸烟成瘾的脑部数据集上准确率最高。

参数λ影响着因果效应估计量对模型学习目标



Fig. 5 Effect of k/(k + l) on classification performance



的贡献程度,其取值对于实现最优诊断性能具有重要作用。为评估λ的影响,实验对其进行了系统性的调整,来测量其对模型诊断性能的影响。将λ设置为0.001时,模型在各项性指标上表现最佳(图7)。



Fig. 7 Impact of λ on model performance



2.5.2 不同模块有效性分析

为了验证模型中模块的有效性,在保证基本网络结构完整的前提下,在吸烟数据集上进行了消融试验。实验方案如下:设置4组消融试验, Network_1代表基准模型,Network_2代表移除空间注意力机制模块后的模型,Network_3代表了移除因果解释模块后的模型,Network_4表示移除Transformer部分后的模型。



Fig. 8 Histogram of ablation experiment

由图8可知,实验采用的3个模块对整体网络性能均有提升,当这几个模块组合使用对整体性能提升最为显著。移除空间注意力机制模块后,模型的准确率下降了0.18,表明此模块能有效捕捉数据的空间信息;移除因果解释模块后,模型的准确率下降约0.18,表明因果解释模块能够有效增强网络的识别和分类能力;在移除Transformer部分后,模型的准确率下降了0.09,表明此部分在获取全局信息和特征识别上起到了增强的作用。实验结果表明,本实验所采用的三个模块均有助于提升模型的分类性能。

本实验还对功能连接矩阵的构建方式进行了深 入讨论,采用了4种不同的计算策略,分别是基于 皮尔逊相关系数 (Pearson's correlation, PC)的 方 法^[35]、基于稀 疏 矩 阵 表 示 (sparse representation, SR)^[36]的方法、相干系数^[37]和锁 相值 (phase locking value, PLV)^[38]。如表 3 所 示,基于PC构建的功能连接矩阵在分类任务中取 得了最佳的性能,无论是在ACC还是F1分数等指 标上均优于其他方法。这表明在本实验中,脑区间 的线性相关性可能更能有效捕捉到与青少年吸烟成 瘾行为相关的特征。相比之下,相干系数与PLV 所构建的连接矩阵在MCC指标上表现相对较差。 SR方法虽未达到皮尔逊方法的整体性能,但在保 持连接矩阵稀疏性和降噪方面仍具有一定优势。

此外,本文对本模型的READOUT策略和不同 的基本分类器进行了进一步的研究。通过分别采用 求和、平均和最大三种不同的池化方式探讨对分类

 Table 3 Comparison with different brain functional connectivity network construction methods

Method	ACC	F1 score	MCC	
Pearson' s correlation (PC)	0.91	0.91	0.83	
Sparse representation (SR)	0.91	0.89	0.83	
Coherence	0.82	0.80	0.63	
Phase locking value (PLV)	0.82	0.83	0.69	

Notes: The best results are shown in bold.

性能的影响。实验结果表明,吸烟成瘾数据集在最 大池化策略(表4)和基本分类器为GAT(表5) 时得到的分类效果最好。

 Table 4
 Comparison of classification performance using different pooling methods

Method	ACC	F1 score	MCC
TI-GNN (sum)	0.73	0.77	0.56
TI-GNN (ave)	0.64	0.60	0.27
TI-GNN (max)	0.91	0.91	0.83

Notes: The best results are shown in bold.

 Table 5
 The effect of different basic classifiers on classification performance

	-			
Method	ACC	F1 score	MCC	
TI-GNN (GCN)	0.82	0.83	0.69	
TI-GNN (GIN)	0.82	0.80	0.63	
TI-GNN (GAT)	0.91	0.91	0.83	

Notes: The best results are shown in bold.

2.6 可解释性分析

为了评估TI-GNN模型的解释,实	:验进一步	₩
-------------------	-------	---

•9•

究了吸烟成瘾数据集对图分类任务的解释,图9中 展示了吸烟成瘾数据集上的连接体情况。这个可视 化中,每个节点的颜色代表一个不同的大脑网络, 而每条边则反映了脑区之间存在联系。数据集中定 义的ROI节点被映射到9个常用的大脑网络上,包 括视觉网络 (visual network, VN)、躯体运动网 络(somatomotor network, SMN)、背侧注意网络 (dorsal attention network, DAN)、腹侧注意网络 (ventral attention network, VAN)、边缘网络 (limbic network, LIN)、额顶叶网络 (frontoparietal network, FPN)、默认模式网络 (default mode network, DMN) 、小脑 (cerebellum, CBL) 和皮层下网络 (subcortical network, SBN)。由图9可以看出右侧背外侧额上 回以及左侧内扣带回存在连接,青少年吸烟成瘾可 能会对情绪调控产生影响,这一结果与之前的研究 结果^[39]相一致。此外,基于获得的空间注意力向。 量Z,实验对不同大脑区域的注意力得分进行了比 较,来确定青少年吸烟成瘾诊断的关键因素。在图

10中,展示了影响诊断的前30个大脑区域,其相 应的功能子网络如表6所示。根据获得的注意力得 分可以看出,其中前额叶皮层涉及到的区域,如背 外侧额上回、额中回、内侧额上回和眶内额上回都 有较高得分,表明青少年吸烟成瘾的情况与认知发 展、执行功能和抑制能力密切相关,有可能会加剧 对尼古丁的依赖, 甚至会造成焦虑和类似抑郁的行 为,此发现与过去的研究发现相一致^[40]。左岛叶 也具有较高的得分,说明对吸烟成瘾对青少年的认 知和情感功能存在一定程度的影响,这与近期的研 究发现结果相一致^[41]。此外,杏仁核也受到一定 程度的影响,这表明青少年吸烟成瘾会对执行能力 和感知能力造成影响,这与近期的研究相一致^[42]。 从功能子网络的角度出发,可以看出异常主要集中 在皮层下网络,这与近期的相关研究吻合[43]。上 述研究表明,青少年吸烟成瘾的病理机制可能与认 知/情感加工区和执行/感知加工区受损有关。在未 来需要更多的样本来验证这一发现。





The brain network can be divided into: visual network (VN), somatomotor network (SMN), dorsal attention network (DAN), ventral attention network (VAN), limbic network (LIN), frontoparietal network (FPN), default mode network (DMN), cerebellum (CBL) and subcortical network (SBN).





The brain network can be divided into: visual network (VN), somatomotor network (SMN), dorsal attention network (DAN), ventral attention network (VAN), limbic network (LIN), frontoparietal network (FPN), default mode network (DMN), cerebellum (CBL) and subcortical network (SBN).

Region No.	Brain Region	Affiliated Functional	Region No.	Brain Region	Affiliated Functional
		Sublictwork			Sublictwork
11	Frontal_Inf_Oper_L	FPN	2	Precentral_R	SMN
3	Frontal_Sup_L	DMN	7	Frontal_Mid_L	DMN
4	Frontal_Sup_R	FPN	8	Frontal_Mid_R	FPN
12	Frontal_Inf_Oper_R	DAN	21	Olfactory_L	LIN
29	Insula_L	VAN	22	Olfactory_R	LIN
30	Insula_R	VAN	23	Frontal_Sup_Medial_L	DMN
72	Caudate_R	SBN	24	Frontal_Sup_Medial_R	FPN
71	Caudate_L	SBN	26	Frontal_Mid_Orb_R	LIN
77	Thalamus_L	SBN	25	Frontal_Mid_Orb_L	LIN
41	Amygdala_L	SBN	38	Hippocampus_R	SBN
60	Parietal_Sup_R	DAN	37	Hippocampus_L	SBN
59	Parietal_Sup_L	DAN	31	Cingulum_Ant_L	DMN
5	Frontal_Sup_Orb_L	LIN	32	Cingulum_Ant_R	SBN
6	Frontal_Sup_Orb_R	FPN	42	Amygdala_R	SBN
1	Precentral_L	SMN	91	Cerebelum_Crus1_L	CBL

Table 6 TOP 30 Brain regions with the highest attention scores

3 结论

总之,针对现有图神经网络在影像分类领域的 局限性,本文介绍了TI-GNN,它有效地模拟了静态的脑功能网络,充分考虑大脑区域级的功能连 接,将空间注意力与表征增强相结合,旨在实现对 吸烟成瘾的精确分类。此外,空间注意力机制增强 了模型的有效性和空间维度的可解释性。在吸烟数 据集上的实验验证通过性能比较和消融试验证明了 该模型的有效性,而且使用学习到的空间注意力机 制进行可解释性分析来证明它的优越性。此模型在 吸烟成瘾数据集中始终实现了最高的分类准确性, 且识别的生物标志物与临床观察结果一致,突出 TI-GNN 对吸烟成瘾研究脑网络分析的贡献。尽管 本文提出的模型在性能上展现了显著的优势,但仍 需指出其存在的局限性,第一,没有相应临床专家 的相关知识纳入TI-GNN 的设计中;第二,涉及到 的青少年样本量数量相对有限,且被试均为男性, 可能限制了模型的泛化能力导致模型未能学习到女 性受试者相关的判别特征,在一定程度上降低了其 临床应用的广度和鲁棒性。未来工作应当注重性别 多样性,在样本招募中适当平衡男女比例,并结合 更多临床信息进行联合建模,或可进一步提高模型 的泛化性能和诊断的精确性,从而更加贴合临床实 际应用的需求。

参考文献

- Ritchie H, Roser M. Smoking [M]. Our World in Data. 2019[2023-11]. https://ourworldindate.org/smoking
- [2] Ng M, Freeman M K, Fleming T D, *et al.* Smoking prevalence and cigarette consumption in 187 countries, 1980-2012. Jama, 2014, 311(2): 183
- [3] Drope J, Schluger N. The tobacco atlas: Sixth Edition. American Cancer Societ, USA, 2018. https://tobaccoatlas.org/wp-content/ uploads/2018/03/TobaccoAtlas_6thEdition_LoRes_Rev0318.pdf
- [4] Mackay J, Eriksen M, Shafey O. The Tobacco Atlas 2nd edition. Atlanta: American Cancer Society, 2006,1106-1107
- [5] Jha P, Ramasundarahettige C, Landsman V, et al. 21st-century hazards of smoking and benefits of cessation in the United States. N Engl J Med, 2013, 368(4): 341-350
- [6] Organization W H. WHO report on the global tobacco epidemic, 2013, WHO, Geneva, Switzerland, 2018, http://www. who. int/ tobacco/global_report/2013/en/
- [7] White H R, Bray B C, Fleming C B, et al. Transitions into and out of light and intermittent smoking during emerging adulthood. Nicotine Tob Res, 2009, 11(2): 211-219
- [8] 隋文禹,薛婷,曹景超,等.吸烟成瘾青少年大脑自发活动变化.中国医学影像技术,2023,39(7):978-981
 Sui W Y, Xue T, Cao J C, et al. Chin J Med Imag Technol, 2023, 39 (7):978-981
- [9] 丁静静, 董芳, 王宏德, 等. 青少年吸烟者结构脑网络的可控性分析. 生物化学与生物物理进展, 2025, 52(1): 182-193
 Ding J J, Dong F, Wang H D, *et al.* Prog Biochem Biophys, 2025, 52(1): 182-193
- [10] Lüscher C, Robbins T W, Everitt B J. The transition to compulsion in addiction. Nat Rev Neurosci, 2020, 21(5): 247-263
- [11] Roland J L, Snyder A Z, Hacker C D, et al. On the role of the corpus callosum in interhemispheric functional connectivity in humans. Proc Natl Acad Sci USA, 2017, 114(50): 13278-13283
- [12] Weisskoff R, Rosen B R. Functional magnetic resonance imaging. Acad Radiol, 1995, 2: S133-S135
- [13] Kambeitz J, Cabral C, Sacchet M D, *et al*. Detecting neuroimaging biomarkers for depression: a meta-analysis of multivariate pattern recognition studies. Biol Psychiatry, 2017, 82(5): 330-338
- [14] Pan X, Xu Y. A novel and safe two-stage screening method for support vector machine. IEEE Trans Neural Netw Learn Syst, 2019, 30(8): 2263-2274
- [15] Ranstam J, Cook JA. LASSO regression. Br J Surg, 2018, 105(10):

1348

王旭雯,等:基于TI-GNN的青少年吸烟成瘾诊断

- [16] Rigatti S J. Random forest. J Insur Med, 2017, 47(1): 31-39
- [17] Yuan H, Yu H, Wang J, et al. On explainability of graph neural networks via subgraph explorations// Meila M, Zhang T. Proceedings of the 38th International Conference on Machine Learning (ICML 2021). PMLR: Proceedings of Machine Learning Research, 2021: 12241 – 12252.
- [18] Zeng Y, Wang X, Liu J, et al. SHADEWATCHER: Recommendation-guided cyber threat analysis using system audit records// IEEE. Proceedings of the 2022 IEEE Symposium on Security and Privacy (SP). Los Alamitos, CA: IEEE Press, 2022: 884 – 901.
- [19] Yuan H, Tang J, Hu X, et al. XGNN: Towards Model-Level Explanations of Graph Neural Networks. ACM, Association for Computing Machinery, NY, USA, 2020: 430-438
- [20] Vaswani A, Shazeer N, Parmar N, et al. Attention Is All You Need// Proceedings of the 31st International Conference on Neural Information Processing Systems. Red Hook, NY, USA : Curran Associates Inc, 2017: 6000-6010
- [21] Zhang M, Cui Z, Neumann M, et al. An end-to-end deep learning architecture for graph classification// McIlraith S, Weinberger K. Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence: Volume 6. New Orleans, Louisiana, USA: AAAI Press, 2018: 4438 – 4445
- [22] Schlichtkrull M, Kipf T N, Bloem P, et al. Modeling Relational Data with Graph Convolutional Networks. Springer, Cham, Springer-Verlag, Berlin, Heidelberg, 2018, 593-607
- [23] Bo D, Wang X, Shi C, et al. Beyond Low-frequency Information in Graph Convolutional Networks // AAAI. Proceedings of the AAAI Conference on Artificial Intelligence: Vol. 35, No. 5. Palo Alto (CA): AAAI Press, 2021: 3950 – 3957
- [24] O'shaughnessy M, Canal G, Connor M, et al. Generative causal explanations of black-box classifiers // Larochelle H, Ranzato M, Hadsell R, et al. Proceedings of the Annual Conference on Neural Information Processing Systems, Red Hook, NY, USA, Curran Associates Inc, 2020: 5453-5467
- [25] Vahia V N. Diagnostic and statistical manual of mental disorders 5: a quick glance. Indian J Psychiatry, 2013, **55**(3): 220-223
- [26] Kingma D, Ba J. Adam: A Method for Stochastic Optimization. arXiv, 2014. DOI:10.48550/arXiv.1412.6980
- [27] Chien Y. Pattern classification and scene analysis. IEEE Trans Automat Contr, 1974, 19(4): 462-463
- [28] Friedman J H. Greedy function approximation: a gradient boosting machine. Ann Statist, 2001, 29(5): 1189-1232
- [29] Kipf T N, Welling M. Semi-supervised classification with graph convolutional networks // Bengio Y, LeCun Y. Proceedings of the 5th International Conference on Learning Representations (ICLR 2017. Toulon, France: OpenReview, 2017: 1 – 14
- [30] Veličković P, Cucurull G, Casanova A, et al. Graph Attention Networks. arXiv, 2017. DOI: 10.48550/arXiv.1710.10903
- [31] Xu K, Hu W, Leskovec J, et al. How powerful are graph neural networks?// Bengio Y, LeCun Y. Proceedings of the 7th

International Conference on Learning Representations (ICLR 2019. OrleansNew, Louisiana, USA: OpenReview, 2019: 1 - 16

- [32] Zheng K, Yu S, Chen B. CI-GNN: a Granger causality-inspired graph neural network for interpretable brain network-based psychiatric diagnosis. Neural Netw, 2024, 172: 106147
- [33] Wu Y, Wang X, Zhang A, et al. Discovering invariant rationales for graph neural networks. arXiv, 2022. DOI: 10.48550/ arXiv.2201.12872
- [34] Li X, Zhou Y, Dvornek N, *et al.* BrainGNN: interpretable brain graph neural network for fMRI analysis. Med Image Anal, 2021, 74: 102233
- [35] Sedgwick P. Pearson's correlation coefficient. Bmj, 2012, 345 (jul041): e4483
- [36] Lee H, Lee D S, Kang H, et al. Sparse brain network recovery under compressed sensing. IEEE Trans Med Imaging, 2011, 30(5): 1154-1165
- [37] Sun F T, Miller L M, D'Esposito M. Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data. Neuroimage, 2004, 21(2): 647-658
- [38] Bruña R, Maestú F, Pereda E. Phase locking value revisited:

teaching new tricks to an old dog. J Neural Eng, 2018, 15(5): 056011

- [39] Zhang X, Salmeron B J, Ross T J, et al. Factors underlying prefrontal and insula structural alterations in smokers. Neuroimage, 2011, 54(1): 42-48
- [40] DeBry S C, Tiffany S T. Tobacco-induced neurotoxicity of adolescent cognitive development (TINACD): a proposed model for the development of impulsivity in nicotine dependence. Nicotine Tob Res, 2008, 10(1): 11-25
- [41] Ghahremani D G, Pochon J F, Diaz M P, et al. Nicotine dependence and insula subregions: functional connectivity and cue-induced activation. Neuropsychopharmacology, 2023, 48(6): 936-945
- [42] Mihov Y, Hurlemann R. Altered amygdala function in nicotine addiction: Insights from human neuroimaging studies. Neuropsychologia, 2012, 50(8): 1719-1729
- [43] Gazula H, Rootes-Murdy K, Holla B, et al. Federated Analysis in COINSTAC Reveals Functional Network Connectivity and Spectral Links to Smoking and Alcohol Consumption in Nearly 2, 000 Adolescent Brains. Neuroinformatics, 2023, 21(2):287-301

Adolescent Smoking Addiction Diagnosis Based on TI-GNN*

WANG Xu-Wen¹, YU Da-Hua^{2)**}, XUE Ting^{3)**}, LI Xiao-Jiao¹, MAI Zhen-Zhen¹, DONG Fang¹, MA Yu-Xin¹, WANG Juan¹, YUAN Kai^{1,2,4})

(¹⁾School of Digital and Intelligent Industry, Inner Mongolia University of Science and Technology, Baotou 014010, China;

²⁾School of Automation and Electrical Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China;

³School of Science, Inner Mongolia University of Science and Technology, Baotou 014010, China;

⁴⁾School of Life Science and Technology, Xidian University, Xi'an 710071, China)

Graphical abstract



Abstract Objective Tobacco-related diseases remain one of the leading preventable public health challenges worldwide and are among the primary causes of premature death. In recent years, accumulating evidence has supported the classification of nicotine addiction as a chronic brain disease, profoundly affecting both brain structure and function. Despite the urgency, effective diagnostic methods for smoking addiction remain lacking, posing significant challenges for early intervention and treatment. To address this issue and gain deeper insights into the neural mechanisms underlying nicotine dependence, this study proposes a novel graph neural network framework, termed TI-GNN. This model leverages functional magnetic resonance imaging (fMRI) data to identify complex and subtle abnormalities in brain connectivity patterns associated with smoking addiction. **Methods** The study utilizes fMRI data to construct functional connectivity matrices that represent interaction

patterns among brain regions. These matrices are interpreted as graphs, where brain regions are nodes and the strength of functional connectivity between them serves as edges. The proposed TI-GNN model integrates a Transformer module to effectively capture global interactions across the entire brain network, enabling a comprehensive understanding of high-level connectivity patterns. Additionally, a spatial attention mechanism is employed to selectively focus on informative inter-regional connections while filtering out irrelevant or noisy features. This design enhances the model's ability to learn meaningful neural representations crucial for classification tasks. A key innovation of TI-GNN lies in its built-in causal interpretation module, which aims to infer directional and potentially causal relationships among brain regions. This not only improves predictive performance but also enhances model interpretability—an essential attribute for clinical applications. The identification of causal links provides valuable insights into the neuropathological basis of addiction and contributes to the development of biologically plausible and trustworthy diagnostic tools. **Results** Experimental results demonstrate that the TI-GNN model achieves superior classification performance on the smoking addiction dataset, outperforming several state-of-the-art baseline models. Specifically, TI-GNN attains an accuracy of 0.91, an F1-score of 0.91, and a Matthews correlation coefficient (MCC) of 0.83, indicating strong robustness and reliability. Beyond performance metrics, TI-GNN identifies critical abnormal connectivity patterns in several brain regions implicated in addiction. Notably, it highlights dysregulations in the amygdala and the anterior cingulate cortex, consistent with prior clinical and neuroimaging findings. These regions are well known for their roles in emotional regulation, reward processing, and impulse control-functions that are frequently disrupted in nicotine dependence. Conclusion The TI-GNN framework offers a powerful and interpretable tool for the objective diagnosis of smoking addiction. By integrating advanced graph learning techniques with causal inference capabilities, the model not only achieves high diagnostic accuracy but also elucidates the neurobiological underpinnings of addiction. The identification of specific abnormal brain networks and their causal interactions deepens our understanding of addiction pathophysiology and lays the groundwork for developing targeted intervention strategies and personalized treatment approaches in the future.

Key words spatial attention, transformer, graph neural network, functional magnetic resonance imaging, classification, adolescent, smoking addiction

DOI: 10.3724/j.pibb.2025.0107 **CSTR:** 32369.14.pibb.20250107

XUE Ting. Tel: 86-472-5950060, E-mail: xueting41@imust.edu.cn

^{*} This work was supported by grants from Chinese National Programs for Brain Science and Brain-like Intelligence Technology (2022ZD0214500), the Scientific and Technological Innovation (STI)2030-Major Project (2022ZD0207100), The National Natural Science Foundation of China (82260359, 82371500, U22A20303, 61971451), Natural Science Foundation of Inner Mongolia (2023QN08007, 2025MS08027, 2025MS08098), the Fundamental Research Fonds for the Universities of Inner Mongolia (2023QNJS204, 2023QNJS206, 2024QNJS119) and the Development Program for Young Talents of Science and Technology in Universities of Inner Mongolia (NJYT24030).

^{**} Corresponding author.

YU Da-Hua. Tel: 86-472-5950060, E-mail: fmydh@imust.edu.cn

Received: March 14, 2025 Accepted: July 9, 2025