arXiv:2512.20059v1 [cs.MM] 23 Dec 2025

DS-HGCN：一种通过社会传染预测学生参与度的双流超图卷积网络

范子扬[1], 陶丽[1](✉)，王毅[1],屈静薇[1],王颖[1],和姜飞[2]

[1] 中国重庆市西南大学，邮编 400715，tli@swu.edu.cn[2] 中国重庆市重庆科学技术研究院

摘要。 学生参与度是影响学业成功和学习成果的关键因素。准确预测学生参与度对于优化教学策略和提供个性化干预至关重要。然而，大多数方法侧重于单维特征分析，并根据学生个体因素评估参与度。本文提出一种基于超图卷积网络的双流多特征融合模型（DS-HGCN），并纳入学生参与度的**社会传染**。DS-HGCN通过对多维特征及其在学生之间的传播机制进行建模，实现对学生参与状态的准确预测。该框架构建超图结构，以编码学生之间参与度的传染，并通过多频信号捕捉情绪和行为上的差异与共性。此外，我们引入超图注意力机制，动态加权每位学生的影响力，从而考虑传播过程中的个体差异。在公开基准数据集上进行的大量实验表明，我们提出的方法性能更优，显著优于现有最先进的方法。

关键词： 学生参与度 · 超图卷积网络 · 社会传染 · 多特征融合 · 超图注意力

1 引言

学生参与度是影响学业成功和学习成果的关键因素 [16]。它是一个多维构念，涵盖行为、情绪和认知方面，这些方面共同提供了对学生参与学习过程的全面视角 [10]。在现代数字化课堂中，准确预测学生参与度对于改进教学策略和提供个性化干预至关重要。

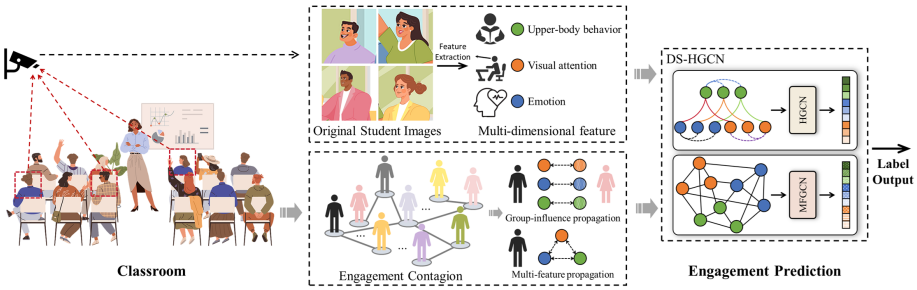


图1：该图展示了教室学生参与度预测的示意图。DS-HGCN 利用学生的多维特征和参与度传染来推断其参与状态。

鉴于参与度的复杂性，相关研究已从调查问卷和观察法等主观方法 [2,28,23]，发展到数据驱动技术，即利用机器学习分析多种数据来源，包括面部表情、肢体语言和生理信号 [2]。然而，这些方法往往聚焦于个别学生或单模态特征，只能提供片面的理解。这忽视了已有充分记录的**参与度传染**现象，即情绪和行为会在群体中传播，并显著影响个体的学习和动机 [19,14]。这造成了显著的研究空白，因为当前模型缺乏明确建模教室参与度传染中固有的群体层面高阶依赖关系的能力。

为弥补这一不足，我们提出了**双流超图卷积网络（DS-HGCN）**，这是一个专为参与度预测中的社会传染建模而设计的框架。我们的方法利用超图来捕捉学生群体内部复杂的多对多关系。该方法由两个并行组件构成：用于建模参与度如何在学生之间传播的多变量传播模块，以及用于捕捉个体行为细粒度差异的多频传播模块。为考虑影响程度的差异，我们进一步整合了超图注意力机制，动态地为这些交互分配权重。我们在 RoomReader 数据集上开展的大量实验验证了该方法的优越性。DS-HGCN 在二分类和三分类任务中分别达到了 94.02% 和 81.37% 的最先进准确率，较以往方法分别高出 2.5% 和 5.95%。总而言之，本文的主要贡献如下：

- 我们将参与度传染的概念引入预测模型，并使用超图捕捉教室中学生之间复杂的群体层面影响。
- 我们提出了一种名为 DS-HGCN 的双流架构，协同建模高阶社会依赖关系和细粒度特征细节。
- 我们设计了一种超图注意力机制，能够动态学习参与度传播中的个体差异。

2 相关工作

2.1 学生参与度分析

学生参与度分析是教育数据挖掘中的一个重要研究领域。传统的参与度分析依赖于问卷调查和直接观察等主观方法，缺乏实时适用性 [2,28,23]。因此，现代方法利用机器学习分析注意力等可见行为 [15,11,13,22],，但其通常依赖单一数据模态，无法捕捉参与度的多面性。

为克服单模态方法的局限，近期研究探索了多模态数据融合。例如，研究人员结合视觉和时序数据（如卷积神经网络和门控循环单元）来捕捉情绪动态 [8],整合来自可穿戴设备的各种生物信号，以评估生理状态 [25],并在虚拟现实环境中利用眼动及其他信号 [1]。另有研究从行为、情绪和认知视角综合分析参与度[29]。最近，研究人员提出了基于图的模型，用于融合文本、音频和视觉数据，为多模态分析提供了新的视角 [17]。在此基础上，我们的工作不仅纳入了与情绪参与度和行为参与度相关的特征，还采用一种新颖的基于图的架构对二者之间复杂的依赖关系进行建模。

2.2 图神经网络

图神经网络（GNN）已被证明能够有效地对不同领域中的关系数据进行建模。在教育场景中，图神经网络已被用于对学生的行为模式进行分类 [21] 并通过融合多模态数据源来评估参与度 [17]。

标准图神经网络仅限于成对关系，而教室互动通常以群体为基础。超图神经网络（HGNN） [9] 更适合处理这些高阶关系，相关发展包括动态构建 [12] 和注意力机制 [3]。然而，一个关键的研究空白仍然存在：现有方法忽视了< style id='4'>参与度传染。我们的工作首次使用超图明确建模群体层面的同伴影响，弥补了以往研究中的一个关键局限。

3 方法

所提出的 DS-HGCN 的完整流程如图 2 所示，由四个部分组成。首先，多特征编码器旨在提取学生的多维特征（第 3.1 节）。随后，多变量传播模块对学生之间参与度的传染过程进行建模（第 3.2 节）。与此同时，多频传播模块从提取的特征中捕获多频信息（第 3.3 节）。最后，参与度分类器根据融合表示预测参与度等级（第 3.4 节）。

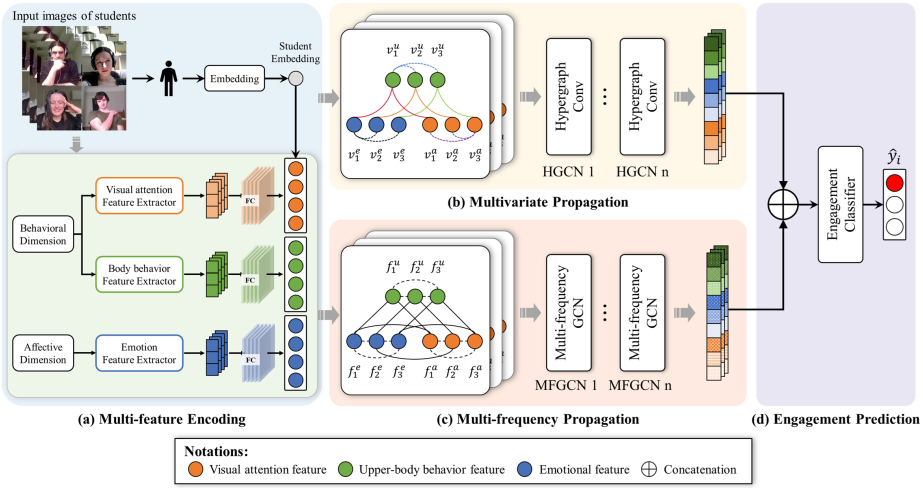


图 2：所提出的 DS-HGCN 的流程。

3.1 多特征编码器

学生参与度包括行为参与度、情感参与度和认知参与度。多特征编码器处理一组图像 $\{x_1, x_2, \ldots, x_N\}$，每张图像都捕捉到某一时刻教室中的 N 名学生。对于每位学生 S_i，编码器都会提取与参与度的情感和行为维度相对应的特征，这也是本视觉研究的主要关注点。

具体而言，我们提取三种单模态表示。对于情感特征（x_i^e），我们使用预训练的 DDAMFN++ [32] 模型，并省略其最终分类层。对于视觉注意特征（x_i^a），我们遵循先前的工作 [24,17]，使用 OpenFace [4] 获取视线方向、头部姿态和面部动作单元。对于上半身行为特征（x_i^u），我们采用 HRNet [18] 提取人体骨骼运动。

随后，我们使用三个独立的多层感知机（MLP），将这些表示投影到统一的 D_h 维空间中：

$$c_i^t = W_t x_i^t + b_i^t, \quad t \in \{e, a, u\} \tag{1}$$

其中 $c_i^t \in \mathbb{R}^{D_h}$。为考虑个体差异，我们引入一个学生专属嵌入 $\hat{s}_i$，并将其添加到每个特征表示中。该嵌入由独热标识符 s_i 计算得出：

$$\hat{s}_i = W_s * s_i, \tag{2}$$

其中 $\hat{s}_i \in \mathbb{R}^{D_h}$ 和 W_s 是一个可训练权重矩阵。因此，每个特征的最终编码表示为：

$$\hat{c}_i^t = c_i^t + \hat{s}_i, \quad t \in \{e, a, u\}. \tag{3}$$

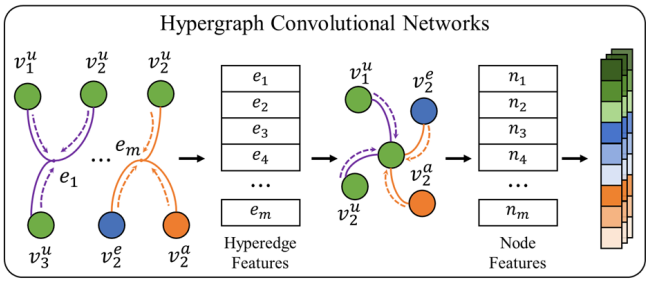


图3：超边卷积层示意图，该层可执行节点-超边-节点变换，并利用超图结构进一步细化特征。

3.2 多变量传播

为对参与度传染进行建模，我们在超图 $\mathcal{H}$上采用多变量传播模块。该结构能够捕捉每个学生内部的异质特征影响，并实现不同学生及维度之间的信息融合。

超图构建 对于一个包含 N 名学生的教室，我们构建超图 $\mathcal{H} = (\mathcal{V}_{\mathcal{H}},\ \mathcal{E}_{\mathcal{H}})$。其结构由节点、超边和关联矩阵定义。

节点。 节点集 $\mathcal{V}_{\mathcal{H}}$ 包含 $3N$ 个节点，其中每名学生 S_i 由三个节点的三元组 $\{v_i^e, v_i^a, v_i^u\}$表示，分别对应其情感、注意力和上半身行为特征。每个节点的初始表示是由公式3推导出的相应编码特征 $\{\hat{c}_i^e、\hat{c}_i^a、\hat{c}_i^u\}$。**超边**。 为捕捉内部特征交互和外部群体影响，我们在 $\mathcal{E}_{\mathcal{H}}$中定义两种类型的超边：

– **多维超边**：为建模单个学生的情感与行为特征之间的依赖关系，对于每名学生，超边连接其三个特征节点 $\{v_i^e, v_i^a, v_i^u\}$。– **群体影响超边**：为建模社会传染，对于每种特征类型 $t \in \{e, a, u\}$，一条超边连接所有学生中对应的节点 $\{v_1^t, \ldots, v_N^t\}$。

这种设计包含 $e \in \mathcal{E}_{\mathcal{H}}(|\mathcal{E}_{\mathcal{H}}| = 3 + N)$ 条超边，使模型能够捕捉超越简单成对交互的高阶相关性。
关联矩阵。超图的拓扑结构由关联矩阵$\mathbf{H}$ $\in \mathbb{R}^{|\mathcal{V}_{\mathcal{H}}|\times|\mathcal{E}_{\mathcal{H}}|}$编码，其中每个条目表示一个节点是否属于某条超边。

$$H_{ve} = \begin{cases} 1, & \text{if node } v \in \text{hyperedge } e \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

该二值矩阵构成基础，并将在**超图注意力**中由我们的注意力机制动态加权。

超图卷积 超图中的信息传播通过超图卷积操作实现，该操作通过共享超边聚合相连节点的信息，从而更新节点表示。我们将超图卷积层表述如下：

$$\mathbf{Q}^{(l+1)} = \sigma(\mathbf{D}_{\mathcal{H}}^{-1}\mathbf{H}\mathbf{W}_e\mathbf{B}^{-1}\mathbf{H}^T\mathbf{Q}^{(l)}\mathbf{P}^{(l)}) \tag{5}$$

其中，$\mathbf{Q}^{(l)} \in \mathbb{R}^{|\mathcal{V}_{\mathcal{H}}| \times D_h}$ 是第 l 层的节点特征矩阵，σ 是非线性激活函数。$\mathbf{D}_{\mathcal{H}}$ 和 $\mathbf{B}$ 分别是顶点和超边的对角度矩阵。$\mathbf{W}_e$ 是一个对角矩阵，其中包含每条超边对应的可学习权重，而 $\mathbf{P}^{(l)}$ 是第 l 层线性变换的可训练权重矩阵。

该操作完全可微，因此模型可以通过梯度下降进行端到端训练。如图 3 所示，该模块执行节点-边-节点变换，根据超图中编码的高阶关系有效地细化特征表示。经过 L 层传播后，我们从最后一层的输出 $v_{i,(L)}^{x}$中获得最终的多变量表示：

$$\overline{v_i^e} = v_{i,(L)}^e, \overline{v_i^a} = v_{i,(L)}^a, \overline{v_i^u} = v_{i,(L)}^u \tag{6}$$

超图注意力为克服二元关联矩阵 $\mathbf{H}$的静态特性（这限制了模型适应不同交互强度的能力），我们引入超图注意力机制。受 [3], 启发，该机制通过计算注意力分数 $\gamma_e(v)$，动态学习超边e 内每个节点 v 的重要性。

$$\gamma_{e_j}(v_i) = \frac{\exp(\sigma(\text{sim}(v_i\mathbf{P}, e_j\mathbf{P})))}{\sum_{k\in\mathcal{N}_v}\exp(\sigma(\text{sim}(v_i\mathbf{P}, e_k\mathbf{P})))} \tag{7}$$

其中，$\sigma(\cdot)$ 是非线性激活函数，$\mathcal{N}_v$ 表示节点 v_i 的邻域集合。相似度函数 $\text{sim}(\cdot)$ 定义为可学习投影：

$$\text{sim}(v_i, e_j) = a^T[v_i][e_j] \tag{8}$$

其中，$[\cdot][\cdot]$ 表示拼接，$\mathbf{a}$ 是一个可训练的权重向量。通过这种方式，超图的初始关联矩阵得到了丰富。我们将注意力分数表示为边依赖节点权重 $\gamma_e(v)$，从而形成动态加权关联矩阵：

$$H_{ve} = \begin{cases} \gamma_e(v), & \text{if node } v \in \text{hyperedge } e \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

通过使式 5 中的$\mathbf{H}$ 依赖于这些可学习的注意力分数，模型能够动态调节信息流，从而有效捕捉传播过程中学生特征与群体影响之间细微且不断变化的依赖关系。

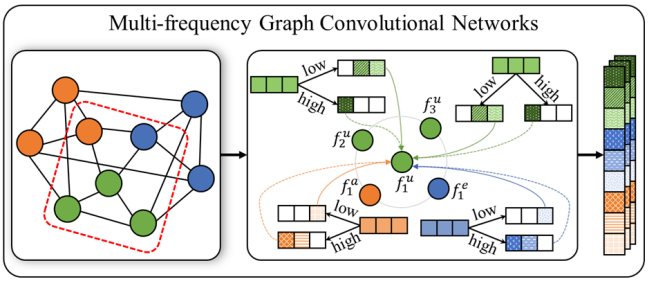


图4：展示了 MFGCN 的示意图。它聚合来自邻近节点的低频和高频信号。

3.3 多频传播

虽然多变量模块的聚合起到低通滤波器的作用，可用于捕捉平滑依赖关系，但也有丢失高频细节的风险。这些细节并非噪声，往往承载着对预测至关重要的特征差异。为保留这些信号，我们引入了一个并行的多频传播模块，该模块受 MFGCN [5],启发，旨在整合学生特征中的低频与高频信息。

图构建 我们构建了一个与超图 $\mathcal{H}$并行运行的无向图 $\mathcal{G} = (\mathcal{V}_\mathcal{G}, \mathcal{E}_\mathcal{G})$。节点集 $\mathcal{V}_\mathcal{G}$ 与 $\mathcal{V}_\mathcal{H}$相同，其中每个节点 $f_i^t (t \in \{e, a, u\})$均使用编码表示$\{\hat{c}_i^e, \hat{c}_i^a, \hat{c}_i^u\}$进行初始化。

不同于使用超边表示群组依赖关系的超图，$\mathcal{G}$ 基于成对连接，以促进直接的频率感知传播。超图结构中相互连接的任意两个节点之间都会建立一条边。这意味着每个节点 f_i^t都与以下节点相连：(1) 来自其他学生、表示相同特征的所有节点 $\{f_j^t | j \neq i\}$；(2) 来自同一学生的另外两个特征节点 $\{f_i^z | z \neq t\}$。该结构由邻接矩阵 $\mathbf{A} \in \mathbb{R}^{|\mathcal{V}_\mathcal{G}| \times |\mathcal{V}_\mathcal{G}|}$表示，并据此推导出归一化图拉普拉斯矩阵：

$$\mathbf{L} = \mathbf{I} - \mathbf{D}_\mathcal{G}^{-1/2} \mathbf{A} \mathbf{D}_\mathcal{G}^{-1/2} \tag{10}$$

其中，$\mathbf{D}_\mathcal{G}$ 是 $\mathcal{G}$ 的对角度矩阵，$\mathbf{I}$ 是单位矩阵。该拉普拉斯矩阵是定义我们的基于频率的滤波器的基础。

多频图卷积为明确捕捉频谱中的信号，我们采用两种不同的图滤波器：低通滤波器$\mathcal{F}_l$ 和高通滤波器 $\mathcal{F}_h$。按照 [5],中的设计，这些滤波器根据图结构定义如下：

$$\begin{aligned} \mathcal{F}_l &= \mathbf{I} + \mathbf{D}_\mathcal{G}^{-1/2} \mathbf{A} \mathbf{D}_\mathcal{G}^{-1/2} = 2\mathbf{I} - \mathbf{L} \\ \mathcal{F}_h &= \mathbf{I} - \mathbf{D}_\mathcal{G}^{-1/2} \mathbf{A} \mathbf{D}_\mathcal{G}^{-1/2} = \mathbf{L} \end{aligned} \tag{11}$$

低通滤波器 $\mathcal{F}_l$ 对邻居节点的特征取平均，以促进平滑；高通滤波器 $\mathcal{F}_h$则是图拉普拉斯算子本身，它会放大节点与其邻居节点之间的差异，从而捕捉高频差异。

然后，我们将这些滤波器整合到一个统一的传播层中，使模型能够自适应地组合低频和高频分量。第 k 层的特征更新规则以特征矩阵 $\mathbf{F}^{(k)}$ 作为输入，表示如下：

$$\mathbf{F}^{(k+1)} = \sigma\left((\mathcal{F}_l\mathbf{F}^{(k)})\mathbf{W}_l^{(k)} + (\mathcal{F}_h\mathbf{F}^{(k)})\mathbf{W}_h^{(k)}\right) \tag{12}$$

其中，$\mathbf{F}^{(k)}$ 是第 k 层的节点特征矩阵$f^t_{i,\,(k)}$。$\mathbf{W}_l^{(k)}$ 和$\mathbf{W}_h^{(k)}$ 是可学习的权重矩阵，用于自适应地平衡低频和高频信号的影响，σ 是一个非线性激活函数。

通过堆叠 K 个这样的层，该模块迭代地传播并细化多频信息。经过 K 层消息传递后，最终得到频率感知型参与度表示：

$$\overline{f_i^e} = f^e_{i,(K)}, \overline{f_i^a} = f^a_{i,(K)}, \overline{f_i^u} = f^u_{i,(K)} \tag{13}$$

3.4 参与度预测

在获得多变量模块和多频模块的表示后，最后一步是预测每位学生的参与状态，并将其划分为 C 个预定义类别之一。为此，我们首先将两个传播流的输出拼接起来，为每位学生 S_i 构建一个综合表示 μ_i ：

$$\mu_i = \overline{v_i^e} \oplus \overline{f_i^e} \oplus \overline{v_i^a} \oplus \overline{f_i^a} \oplus \overline{v_i^u} \oplus \overline{f_i^u} \tag{14}$$

其中，μ_i 作为全面的特征表示，捕捉与参与度预测相关的参与度传染和多频交互。

随后，将 μ_i 输入一个变换层，再通过 Softmax 函数进行参与度分类：

$$\begin{aligned} \mathcal{P}_i &= \mathrm{softmax}(W_4 \cdot \mathrm{ReLU}(\mu_i) + b_4), \\ \hat{y}_i &= \arg\max_{\tau}(\mathcal{P}_i[\tau]), \end{aligned} \tag{15}$$

其中，W_4 是一个可训练权重矩阵，$\mathcal{P}_i \in \mathbb{R}^C$ 表示参与度类别上的概率分布，$\hat{}y_i$ 是对学生 S_i的参与度标签的预测。

模型通过最小化训练集 $\mathcal{D}$中所有学生样本的标准分类交叉熵损失的平均值进行端到端训练，并采用 L2 正则化防止过拟合。损失函数定义如下：

$$\mathcal{L} = -\frac{1}{|\mathcal{D}|}\sum_{i\in\mathcal{D}} \log \mathcal{P}_i[y_i] + \lambda\|\theta\|_2^2 \tag{16}$$

其中，y_i 是学生 S_i的真实标签，$\mathcal{P}_i[y_i]$ 是模型对该真实类别的预测概率，θ 表示模型的所有可训练参数，λ 是控制 L2 正则化强度的超参数。

4 实验结果

本节将开展一系列全面的实验，严格评估所提出的 DS-HGCN 的有效性和鲁棒性。首先，我们在第 4.1 节详细介绍实验设置。随后，在第 4.2 节通过与当前最先进的方法进行比较，展示我们模型的优越性。为验证设计选择，我们在第 4.3 节开展深入的消融研究，分析各个核心组件的作用。最后，在第 4.4 节进一步分析超参数敏感性和数据效率，从而全面评估模型。

4.1 实验设置

数据准备。 我们在 RoomReader 数据集 [24],（一个用于学生参与度分析的基准数据集）上评估我们的方法。经过数据筛选和帧提取后，我们处理了此前研究 [17] 指出的数据集显著的类别不平衡问题。按照既定方案 [17,26],，我们根据连续标注构建了两项任务：一项二**分类** 任务（“参与”与“未参与”），以及一项三**分类** 任务（“低”“中”“高”参与度），最终评估集包含 1,252 张独特的教室快照。

实现细节 我们的模型使用 PyTorch 实现，并在单块 NVIDIA 3090 GPU 上训练。我们使用 Adam 优化器，初始学习率为 1×10^{-5}。网络采用 0.5 的丢弃率以减轻过拟合。为处理剩余的类别不平衡问题，我们在训练过程中将类别权重纳入交叉熵损失函数。超图卷积层数（L）和图卷积层数（K）分别设为 3 和 3。模型性能通过准确率、F_1-分数和曲线下面积（AUC）进行评估。

4.2 基线方法比较

为验证我们提出的 DS-HGCN 的有效性，我们将其与多种具有代表性且达到当前最先进水平的参与度预测模型进行了全面比较。为确保评估公平，我们使用 RoomReader 数据集重新实现了这些基线模型，并遵循指定的数据格式。基线模型包括：

- **单特征模型：** 此类方法对传统的非图方法进行基准评估，这些方法基于单一特征对参与度建模。主要示例包括：**TCCT-NET** [30], 通过双流融合网络分析行为信号；**EnsModel** [27], 利用深度网络根据面部行为预测参与度；**ConvLSTM** [7], 使用集成卷积回归模型处理视频帧；**ED-MTT** [6], 通过多任务训练框架提高准确率；以及**EG-NET** [20], 采用标准卷积神经网络进行面部表情分析。

表1：在 RoomReader 数据集上与此前最先进方法的比较。标有 ◇ 的条目取自论文 [17]，而标有 † 的条目则来自我们使用开源代码重新实现的结果。

方法	二分类 准确率（%）	三分类 特征维度 准确率（%）	Type
TCCT-NET† [30]	73.99 ± 1.43	60.48 ± 1.18	单一
EnsModel† [27]	75.30 ± 3.50	69.53 ± 3.54	单一
ConvLSTM† [7]	76.50 ± 1.85	74.53 ± 1.50	单一
EG-NET† [20]	77.38 ± 1.27	72.76 ± 2.24	单一
ED-MTT† [6]	73.80 ± 3.35	71.20 ± 1.35	单次
Bootstrap† [31]	82.42 ± 2.15	75.42 ± 2.43	多次
Haar-MGL◇ [17]	90.18 ± 1.34	-	多重
DS-HGCN（我们的方法）	**94.02 ± 0.49**	**81.37 ± 0.34**	多个

- **多特征模型：** 为展示纳入多个特征的优势，我们比较了**Bootstrap** [31]，这是一种同时结合头部和身体特征的方法，代表了在单特征分析基础上的进一步发展。
- **基于图的模型（当前最先进水平）：** 最后，为了将我们的工作与直接的当前最先进方法进行比较，我们以**Haar-MGL** [17]作为基准。该模型采用多模态图学习框架来融合文本、音频和图像数据，为基于图的参与度预测提供了最相关且最有力的比较对象。

表1展示了五次独立试验的平均定量结果，采用80%/20%的训练/测试划分。如表所示，DS-HGCN在二分类和三分类任务中均显著优于所有基线方法。比较结果揭示了两点重要认识。首先，多特征方法始终优于单特征模型，证实了整体行为分析的价值。其次，DS-HGCN确立了新的当前最先进水平，优于强大的基于图的Haar-MGL。这一优势凸显了我们的双流超图网络在显式建模高阶社会传染方面的独特优势——这是以往方法所忽略的一项关键动态。

4.3 消融研究

我们的消融研究（表2）验证了 DS-HGCN 各组成部分的作用。移除多变量传播流（变体1）会导致准确率下降幅度最大（二分类/三分类任务中分别下降4.18%/5.67%），证实通过超图建模社会传染是模型取得成功的主要驱动力。多频流（变体2）也至关重要，缺少该流会显著降低性能。对输入特征（变体3至5）的分析表明，情绪线索不可或缺；移除这些线索会导致性能下降最为明显（二分类准确率下降12.46%）。最后，禁用超图注意力（变体6）会损害性能，验证了我们的动态

表2：DS-HGCN 的消融研究。我们将所提方法与第4.3节中介绍的一系列变体进行比较。

候选项	二分类任务 准确率（%）	F_1	AUC	三分类任务 准确率（%）	F_1	AUC
DS-HGCN（完整模型）	**94.02**	**0.939**	**0.893**	**81.37**	**0.816**	**0.921**
1 不含多变量信息	89.84	0.891	0.845	75.70	0.755	0.885
2 不含多频信息	92.13	0.917	0.857	77.29	0.776	0.904
3 不含视觉注意力特征	84.50	0.832	0.805	68.45	0.685	0.843
4 不含身体行为特征	86.37	0.859	0.823	73.37	0.742	0.862
5 不含情绪特征	81.56	0.803	0.784	62.77	0.631	0.784
6 不使用超图注意力。	92.73	0.925	0.872	79.23	0.793	0.867

加权方案。这些结果证实，每项设计选择都是模型达到最先进性能不可或缺的一环。

4.4 进一步分析

关于图层的讨论。 我们通过将超图（L）和图（K）层的数量从1调整到6，分析了模型深度的影响。如图5所示，在两项分类任务中，性能分别在 $L = 3$ 和 $K = 3$达到峰值。这表明，深度为3时最适合捕捉参与度传染，同时不会导致过平滑，因为更深的模型带来的增益可以忽略不计。

训练数据规模的影响。 为评估数据效率，我们在训练数据的子集上训练模型，子集规模从20%到80%不等。图6显示，即使在低数据量条件下，DS-HGCN也始终优于基线模型。值得注意的是，在所有数据规模下，其标准差更小且更稳定，证实了其更强的鲁棒性和泛化能力。

计算复杂度分析。 我们评估了超图注意力（HA）机制带来的开销。表 3 显示，与不使用注意力的版本相比，完整的 DS-HGCN 模型仅使参数量略有增加（从 3.076M 增至 3.079M），训练时间也略有上升。更重要的是，推理时间仍然非常高效（每批次 8.52 毫秒），因此我们的方法适用于生成定期参与度报告等应用场景。

表3：计算成本与效率比较。

模型	参数 (M)	训练时间 （秒/轮）	推理时间 （毫秒/批次）
DS-HGCN（不含 HA）	3.076	3.69	7.31
DS-HGCN（完整模型）	3.079	4.39	8.52

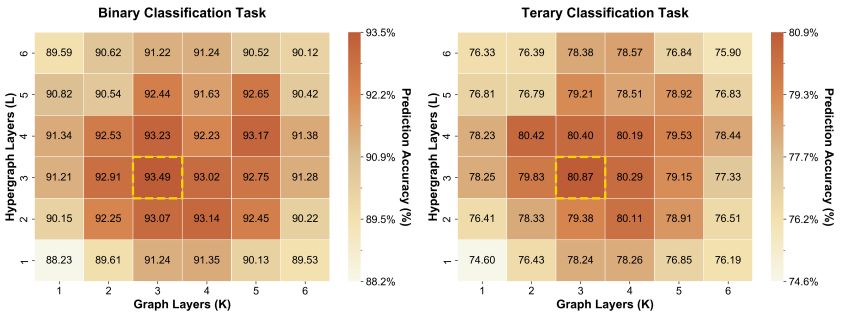


图 5：热力图展示了在超图层数（L）和图层数（K）不同的情况下，二分类和三分类任务的性能。

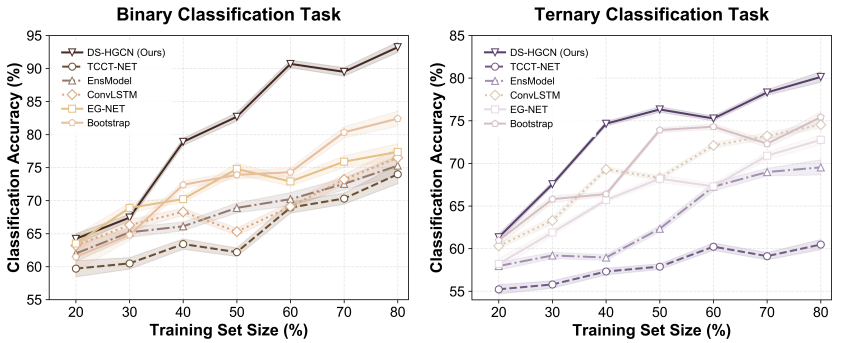


图 6：比较了两种分类任务中模型在不同训练数据比例下的性能。阴影区域表示标准差。

5 结论

本文介绍了 DS-HGCN，这是一种通过对社会传染进行建模来预测学生参与度的双流超图卷积网络。我们的方法采用并行的超图卷积网络和图卷积网络，探索不同学生特征之间的高阶复杂关系，并充分利用不同频率信息中特征之间的相关性和依赖关系。在 RoomReader 数据集上进行的大量实验表明，我们的方法表现出色，在二分类和三分类任务中的准确率分别达到 94.02% 和 81.37%。虽然基于图像的方法效果显著，但未来的工作可以纳入音频和文本等其他模态，或探索认知参与度的代理指标，以进一步提升该方法。总之，DS-HGCN 凸显了社会情境在学习分析中的重要性，并为自动化参与度监测提供了有力工具。

参考文献

1. Asish, S.M., Kulshreshth, A.K., Borst, C.W.：利用眼动数据上的机器学习检测虚拟现实教育环境中的分心学生。《计算机与图形学》**109**，75–87（2022） 2. Azevedo, R.：界定和衡量科学学习中的参与度与学习：概念、理论、方法和分析问题。《教育心理学家》**50**(1)，84–94（2015） 3. Bai, S., Zhang, F., Torr, P.H.：超图卷积与超图注意力。《模式识别》**110**，107637（2021） 4. Baltrušaitis, T., Robinson, P., Morency, L.P.：OpenFace：一款开源面部行为分析工具包。载于：2016年IEEE冬季计算机视觉应用会议（WACV）。第1–10页。IEEE（2016） 5. Bo, D., Wang, X., Shi, C., Shen, H.：超越图卷积网络中的低频信息。载于：AAAI人工智能会议论文集。第35卷，第3950–3957页（2021） 6. Copur, O., Nakıp, M., Scardapane, S., Slowack, J.：在电子学习环境中通过多任务训练检测参与度。载于：国际图像分析与处理会议。第411–422页。Springer（2022） 7. Del Duchetto, F., Baxter, P., Hanheide, M.：你还在听我说吗？从机器人视角进行持续投入度评估。《机器人学与人工智能前沿》**7**，116（2020） 8. Ding, L.：基于GRU和非线性Transformer算法的在线教学情绪分析。《PeerJ计算机科学》**9**，e1696（2023） 9. Feng, Y., You, H., Zhang, Z., Ji, R., Gao, Y.：超图神经网络。载于：AAAI人工智能会议论文集。第33卷，第3558–3565页（2019） 10. Fredricks, J.A., Blumenfeld, P.C., Paris, A.H.：学校投入度：这一概念的潜力与证据现状。《教育研究评论》**74**(1)，59–109（2004） 11. Gupta, S., Kumar, P., Tekchandani, R.K.：在在线学习情境中，基于深度学习模型的实时学习者面部情绪识别与参与度检测系统。《多媒体工具与应用》**82**(8)，11365–11394（2023） 12. Jiang, J., Wei, Y., Feng, Y., Cao, J., Gao, Y.：动态超图神经网络。载于：国际人工智能联合会议论文集（IJCAI）。第2635–2641页（2019） 13. Khenkar, S., Jarraya, S.K.：利用三维卷积神经网络分析微小身体手势进行参与度检测。《计算机、材料与连续体》**70**(2)（2022） 14. King, R.B., Mendoza, N.B.：学生社会目标的社会传染及其对学校投入度的影响。《学习与个体差异》**88**，102004（2021） 15. Kodithuwakku, J., Arachchi, D.D., Rajasekera, J.：一种情绪与注意力识别系统，利用机器学习模型和神经加速器芯片，实时识别视频对话参与者的参与度水平。《算法》**15**(5)，150（2022） 16. Kuh, G.D.：通过全国学生投入度调查评估真正影响学生学习的因素。《变革：高等教育杂志》**33**(3)，10–17（2001） 17. Li, M., Zhuang, X., Bai, L., Ding, W.：基于三维Haar半紧框架小波的多模态图学习，用于学生参与度预测。《信息融合》**105**，102224（2024）

18. Lu, B., Sun, Y., Yang, Z., Song, R., Jiang, H., Liu, Y.：HRNet：具有层次化细化功能的点云三维目标检测网络。《模式识别》**149**，110254（2024） 19. Mendoza, N.B., King, R.B.：学校中学生参与度的社会传染。《国际学校心理学》**41**(5)，454–474（2020） 20. Mohamad Nezami, O., Dras, M., Hamey, L., Richards, D., Wan, S., Paris, C.：利用深度学习和面部表情自动识别学生参与度。载于：欧洲机器学习与数据库知识发现联合会议。第273–289页。Springer（2019） 21. Mubarak, A.A., Cao, H., Hezam, I.M., Hao, F.：利用图卷积网络建模学生表现。《复杂与智能系统》**8**(3)，2183–2201（2022） 22. Ngoc Anh, B., Tung Son, N., Truong Lam, P., Phuong Chi, L., Huu Tuan, N., Cong Dat, N., Huu Trung, N., Umar Aftab, M., Van Dinh, T.：一种基于计算机视觉的教室学生行为监测应用。《应用科学》**9**(22)，4729（2019） 23. Pekrun, R., Goetz, T., Frenzel, A.C., Barchfeld, P., Perry, R.P.：衡量学生学习与表现中的情绪：成就情绪问卷（AEQ）。《当代教育心理学》**36**(1)，36–48（2011） 24. Reverdy, J., Russell, S.O., Duquenne, L., Garaialde, D., Cowan, B.R., Harte, N.：RoomReader：在线多人会话互动的多模态语料库。载于：第十三届语言资源与评测会议论文集。第2517–2527页（2022） 25. Södergård, C., Laakko, T.：利用可穿戴设备的生物信号数据推断日常生活中的学生自评专注水平。《IEEE Access》**11**，30308–30323（2023） 26. Sümer, Ö., Goldberg, P., D'Mello, S., Gerjets, P., Trautwein, U., Kasneci, E.：基于教室面部视频的多模态参与度分析。《IEEE情感计算汇刊》**14**(2)，1012–1027（2021） 27. Thong Huynh, V., Kim, S.H., Lee, G.S., Yang, H.J.：基于面部行为特征的投入强度预测。载于：2019年国际多模态交互会议。第567–571页（2019） 28. Vallerand, R.J., Pelletier, L.G., Blais, M.R., Briere, N.M., Senecal, C., Vallieres, E.F.：学业动机量表：衡量教育中的内在动机、外在动机和无动机。《教育与心理测量》**52**(4)，1003–1017（1992） 29. Vanneste, P., Oramas, J., Verelst, T., Tuytelaars, T., Raes, A., Depaepe, F., Van den Noortgate, W.：计算机视觉与人类行为、情绪和认知检测：一项关于学生参与度的案例研究。《数学》**9**(3)，287（2021） 30. Vedernikov, A., Kumar, P., Chen, H., Seppänen, T., Li, X.：TCCT-Net：一种双流网络架构，通过行为特征信号快速、高效地估计参与度。载于：IEEE/CVF计算机视觉与模式识别会议论文集。第4723–4732页（2024） 31. Wang, K., Yang, J., Guo, D., Zhang, K., Peng, X., Qiao, Y.：用于投入强度回归的自助法模型集成与排序损失。载于：2019年国际多模态交互会议。第551–556页（2019） 32. Zhang, S., Zhang, Y., Zhang, Y., Wang, Y., Song, Z.：用于面部表情识别的双向注意力混合特征网络。《电子学》**12**(17)，3595（2023）