

【电子与信息科学 / Electronics and Information Science】

基于拓扑数据分析与卷积神经网络的特征融合方法

杨含¹, 秦广军¹, 刘子源², 胡永庆¹, 刘光南¹, 戴庆龙¹

1) 北京联合大学智慧城市学院, 北京 100101; 2) 国家电力投资集团数字科技有限公司, 北京 102209

摘要: 针对卷积神经网络(convolutional neural networks, CNN)难以捕获和利用复杂高维数据的多维结构信息, 限制了其特征学习能力的问题, 提出一种融合了拓扑数据分析(topological data analysis, TDA)与CNN的特征融合方法——TDA-CNN. 该方法将CNN捕获的数值分布特征与TDA提取的拓扑结构特征相融合, CNN通道负责提取数值分布特征, TDA通道专注于提取拓扑结构特征, 然后, 将这两类特征融合形成组合特征表示, 并利用注意力机制自适应地学习每种特征的重要性权重, 为后续全连接网络提供更全面的决策依据. 在Intel Image、Gender Images和Chinese Calligraphy Styles by Calligraphers等数据集上的实验表明, TDA-CNN在改进特征聚类与识别关键特征方面表现出色, 分别将基线模型VGG16、EfficientNet V2和DenseNet121的性能提升了21.89%、22.66%和8.26%, 有效增强了模型的判别能力.

关键词: 人工智能; 模式识别; 计算机神经网络; 拓扑数据分析; 卷积神经网络; 注意力机制; 计算机图像处理

中图分类号: TP391.4; TP3-05

文献标志码: A

DOI: 10.3724/SP.J.1249.2025.05624

A feature fusion method based on topological data analysis and convolutional neural network

YANG Han¹, QIN Guangjun¹, LIU Ziyuan², HU Yongqing¹, LIU Guangnan¹, and DAI Qinglong¹

1) School of Smart Cities, Beijing Union University, Beijing 100101, P. R. China

2) State Power Investment Group Digital Technology Co., Ltd., Beijing 102209, P. R. China

Abstract: A feature fusion method, TDA-CNN, is proposed to address the problem of convolutional neural networks (CNNs) being difficult to capture and utilize the multidimensional structural information of complex high-dimensional data, which limits their feature learning ability. The numerical distribution features captured by CNNs with the topological structure features extracted by TDA, and the CNN channel is responsible for extracting numerical distribution features, while the TDA channel focuses on extracting topological structure features. The numerical distribution features and topological structure features are then fused to form a combined feature representation. An attention mechanism is introduced to adaptively learn the importance weight of each feature, thus providing a more robust foundation for decision-making in the fully connected layers. Experiments on datasets such as Intel Image, Gender Images, and Chinese Calligraphy Styles by Calligraphers show that TDA-CNN performs well in improving feature clustering and identifying key features, improving the performance of baseline models VGG16, EfficientNet V2, and DenseNet121 by 21.89%, 22.66%, and 8.26%, effectively enhancing the discriminative ability of the

Received: 2024-11-09; **Accepted:** 2025-03-16; **Online (CNKI):** 2025-05-16

Foundation: National Natural Science Foundation of China (62172045); Beijing Union University Research Project Funding (ZK10202403)

Corresponding author: Associate professor QIN Guangjun (zhtguangjun@bnu.edu.cn)

Citation: YANG Han, QIN Guangjun, LIU Ziyuan, et al. A feature fusion method based on topological data analysis and convolutional neural network [J]. Journal of Shenzhen University Science and Engineering, 2025, 42(5): 624-630. (in Chinese)

Open access: This article is licensed under the CC BY 4.0 License (<https://creativecommons.org/licenses/by/4.0/>).

<http://journal.szu.edu.cn>



model.

Key words: artificial intelligence; pattern recognition; computer neural networks; topological data analysis; convolutional neural network; attention mechanism; computer image processing

卷积神经网络(convolutional neural networks, CNN)是深度学习的一种重要实现方式,能够有效捕捉图像中的边缘、角和纹理等局部和全局特征^[1-3],其核心是通过多级非线性变换和特征提取,逐步学习数据中的抽象特征表示.近年来,图神经网络(graph neural networks, GNN)、几何深度学习(geometric deep learning, GDL)等模型试图将神经网络推广到非欧空间^[4](如图、流形、网格或字符串表示等),以期提升神经网络的特征表示和决策能力.拓扑数据分析(topological data analysis, TDA)^[5]作为一种基于代数拓扑的数据分析方法,能有效提取数据的深层次拓扑结构信息,已被引入神经网络领域,且表现出优异性能.

TDA可被视为基于拓扑学原理的无监督学习方法^[6],用于描述数据的形状和连接性.通过将数据点云化,采用持久同源性(persistent homology, PH)^[7-8]和Mapper^[9-11]等方法, TDA将高维数据映射到拓扑空间,分析点云数据的拓扑结构及其相互关系,从而有效揭示了数据中的潜在相关性并提取结构特征.具体而言,持久同源通过构建数据点之间的相似度或距离,逐步增加数据点并跟踪拓扑特征的变化,从而提取拓扑结构特征;Mapper将数据空间分割为多个簇,通过连接重叠或关联的簇,最终形成拓扑图来展示数据的全局结构.

为适配机器学习和数据分析任务,多种基于PH的特征表示方法被陆续提出,包括持久条形码(persistence barcodes)、持久性图(persistence diagram, PD)^[12]、持久曲线(persistence curves, PC)^[13]、持久统计(persistence statistics, PS)、持久景观(persistence landscapes, PL)^[14]和持久性映像(persistent image, PI)^[15]等,这些方法将高维数据的拓扑特征转换为固定维度的向量表示.

TDA擅长提取全局结构信息,但在区分细微视觉差异时存在局限,常需结合其他视觉特征以提升适用性.近年来, TDA在机器学习和深度学习领域得到了广泛应用,显著提升了机器学习和深度学习模型的决策能力和泛化能力. TDA与其他方法结合的核心目标是通过将拓扑特征与其他特征的融合,丰富数据的特征空间,提升分类与识别性能.

SWENSON等^[16]将图表示学习与TDA方法结

合,构建了PersGNN模型,通过处理蛋白质结构的1维和2维PDs来捕获数据的局部和全局拓扑结构特征,实现了高准确性的蛋白质功能预测. CANG等^[17]提出蛋白质特异性持久同源(element-specific persistent homology, ESPH)方法,用于提取蛋白质结构的1维拓扑特征,揭示了生物分子中隐藏的结构-功能关系,并进一步结合深度卷积神经网络,构建了多任务多通道拓扑神经网络(multi-task multi-channel topological convolutional neural network, MM-TCNN),用于预测蛋白质-配体结合的亲和力和突变后的蛋白质稳定性变化. AHMED等^[18]将TDA用于眼底成像领域,提出Topo-ML模型和Topo-Net模型,将拓扑特征分别与机器学习技术和深度学习模型相结合,显著提高了模型的分类性能. HU等^[19]提出TopoResNet-101深度学习架构,将PC和PS等拓扑特征融入残差神经网络(residual neural network, ResNet),稳健性较ResNet-101更好.

然而,现有方法主要依赖于PD、PC、PS和PL等拓扑特征的表征方式,特征的表现能力存在不足;此外,这些拓扑特征通常无法直接输入到CNN中进行端到端的训练,限制了其在深度学习模型中的应用.而且,这些方法在融合拓扑特征与数值特征时多采用固定的权重或简单拼接方式,未能有效自适应地分配不同特征的权重.本研究采用更为稳健和鲁棒的持久性映像特征表征方法^[20],并通过注意力机制(attention mechanism, AM)自动学习并分配拓扑特征和数值特征之间的权重,有效提升了特征融合的能力和效果.

1 TDA-CNN模型设计与实现

本研究采用双通道特征流设计,结合TDA与CNN构造TDA-CNN网络结构,将持久性映像作为拓扑特征表示,融合CNN提取的数值分布特征,再采用AM调整特征权重,提升特征学习和表示能力.

1.1 TDA-CNN模型结构

TDA-CNN模型结构如图1.设输入图像为 D ($D \in \mathbb{R}^{H \times W \times C}$,其中, H 、 W 和 C 分别为图像的高度、宽度和通道数),对 D 进行TDA生成持久性映像 I_{pers} .同时,选用任意CNN模型提取图像的数值

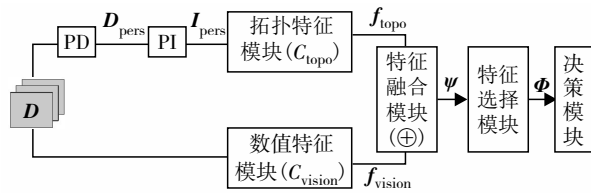


图1 TDA-CNN模型架构

Fig. 1 TDA-CNN model architecture diagram.

分布特征.

数值特征模块采用CNN模型 C_{vision} 提取输入图像的视觉数值特征^[21], 得到视觉特征向量为

$$f_{\text{vision}} = C_{\text{vision}}(D) \quad (1)$$

拓扑特征模块将高维拓扑特征转化为适用于深度学习的图像数据, 实现对数据的全局拓扑结构和局部细节的有效提取, 同时对齐特征融合前的数据特征, 是TDA发挥作用的关键步骤.

先计算图像数据的拓扑特征(I_{pers}), 再用对应的CNN模型(C_{topo})提取 I_{pers} 的数值特征, 即

$$f_{\text{topo}} = C_{\text{topo}}(I_{\text{pers}}) \quad (2)$$

特征融合模块将视觉特征与拓扑特征按通道拼接, 融合后可得

$$\Psi = f_{\text{vision}} \oplus f_{\text{topo}} \quad (3)$$

基于代数拓扑的TDA擅长发现数据内在的结构特征, 但对数值特征的敏感度不如CNN, 二者融合有助于从不同角度丰富数据的特征表示.

为自适应适配数据的结构特征和数值特征, 增强模型对关键特征的关注, 确保在特征融合过程中更精准地捕捉数据中的重要模式, 本研究在组合特征上引入注意力机制压缩与激励(squeeze and excitation, SE)模块, 使TDA-CNN能够动态地调整不同特征的权重, 更有效地提取和整合拓扑结构与数值信息, 从而提升模型在复杂数据分析中的表征能力和分类精度. 经过AM处理后的特征为

$$\Phi = f_{\text{SE}}(\Psi) \quad (4)$$

其中, f_{SE} 为SE模块.

决策模块采用全连接神经网络(fully connected neural networks, FCNN), 将引入AM后的组合特征输入至全连接层, 用于执行分类任务. 该步骤通过全连接层的处理, 将高度抽象的特征表示映射至类别空间, 从而实现对图像的精确分类.

$$O = N_{\text{FC}}(\Phi) \quad (5)$$

其中, O 为分类结果; N_{FC} 为全连接层映射函数.

在模型训练过程中, 采用交叉熵损失函数用于

优化 C_{topo} 和 C_{vision} 两个子网络. 通过计算每个子网络的预测输出与真实标签之间的交叉熵损失, 指导这两个网络的参数更新, 以提升模型的整体分类性能. 最终的输出经过SE模块融合两个子网络的特征, 使模型能够更准确地完成分类任务.

1.2 计算持久性映像

在将持久性图向量化为持久性映像时, 一方面由于0维特征的出生值和死亡值相等, 导致拓扑特征持续时间为0, 这些特征点聚集在持久性图的 y 轴上, 无法有效呈现持久性映像图的宽度; 另一方面, 持久性图中存在持续时间为无穷大的点(通常表示为数据中的孤立点或噪声点), 无法在有限空间内绘制为持久性映像. 上述问题阻碍了持久性图向持久性映像图的有效转换. 本研究对转换过程进行改进, 具体流程为:

1) 将图像数据转化为点云, 在其上构建复形结构, 计算持久性并提取拓扑特征. 令提取的第 K 维拓扑特征集合为 $T_K = \{(x_n, y_n) | n = 1, 2, \dots, N\}$, 其中, x_n 和 y_n 分别为第 N 个拓扑特征的出生时间和死亡时间; N 为第 K 维拓扑特征的特征总数. 使用距离阈值(ε)筛选出符合条件的拓扑特征, 形成筛选后的持久性图, 记作 $P_K^{(\varepsilon)} = \{T_K(x_m, y_m) | m = 1, 2, \dots, M\}$. 其中, 每个点代表1个拓扑特征及其生命周期; M 为符合筛选条件的特征总数.

提取0维和1维的拓扑特征, 这既能有效捕捉数据的基本结构信息, 又能保持计算效率和可行性. 按维度解耦同调计算, 独立记录生命周期并输出0维和1维持久性图 $P^{(0)}$ 和 $P^{(1)}$, 并以图形化方式直观展示数据的拓扑特征及其持久性.

2) 对于 $P^{(0)}$ 和 $P^{(1)}$, 通过计算特征点的出生时间和持续时间之间的关联来构造相应的出生持久(birth-persistence, BP)图, 将拓扑特征 i 的持久性表示为 $\pi_i = y_i - x_i$, 从而将特征转换为其出生时间和持久时间对 (x_i, π_i) .

在0维出生持久图($P_B^{(0)}$)中, 针对持续时间为无穷大的特征 i , 将持续时间替换为最大持续时间 $\pi_{\max}$; 引入辅助坐标点 (τ_x, τ_y) , 其中, $\tau_y > \tau_x > 0$. 转换示例如图2所示. 经过上述操作, 得到0维和1维的出生持久图 $P_B^{(0)}$ 和 $P_B^{(1)}$.

3) 通过将0维和1维出生持久图上的特征点信息及其对应的持续时间转化为像素值的分布信息,

生成持久性映像. 转换过程在2维平面上运用高斯核函数对特征点进行加权, 加权后的像素值分布作为持久性映像的具象表征.

以出生持久图上每个特征点 (x_i, π_i) 为中心, 通过2维高斯函数映射将BP图转换为 I_{pers} 图,

$$\phi(x, y; x_i, \pi_i) = \exp\left[-\frac{(x - x_i)^2 + (y - \pi_i)^2}{2\xi^2}\right] \quad (6)$$

其中, ξ 为高斯核的标准差; x 和 y 为图像网格的坐标. 出生持久图上的特征点 (x_i, π_i) 通过坐标映射到

图像网格空间 (x, y) , 再采用二维高斯核函数进行加权计算, 使拓扑特征的抽象信息“扩散”至图像空间. 通过叠加所有特征点的高斯响应, 抽象的拓扑特征信息被转化为以 (x, y) 为空间坐标的持久性映像, 实现从拓扑特征到图像表征的转换. 因此, 在图像网格上对式(6)求和, 可生成持久性图像

$$I_{\text{pers}}(x, y) = \sum_i \phi(x, y; x_i, \pi_i) \quad (7)$$

将0维和1维的持久性映像分别表示为 $I_{\text{pers}}^{(0)}$ 和 $I_{\text{pers}}^{(1)}$.

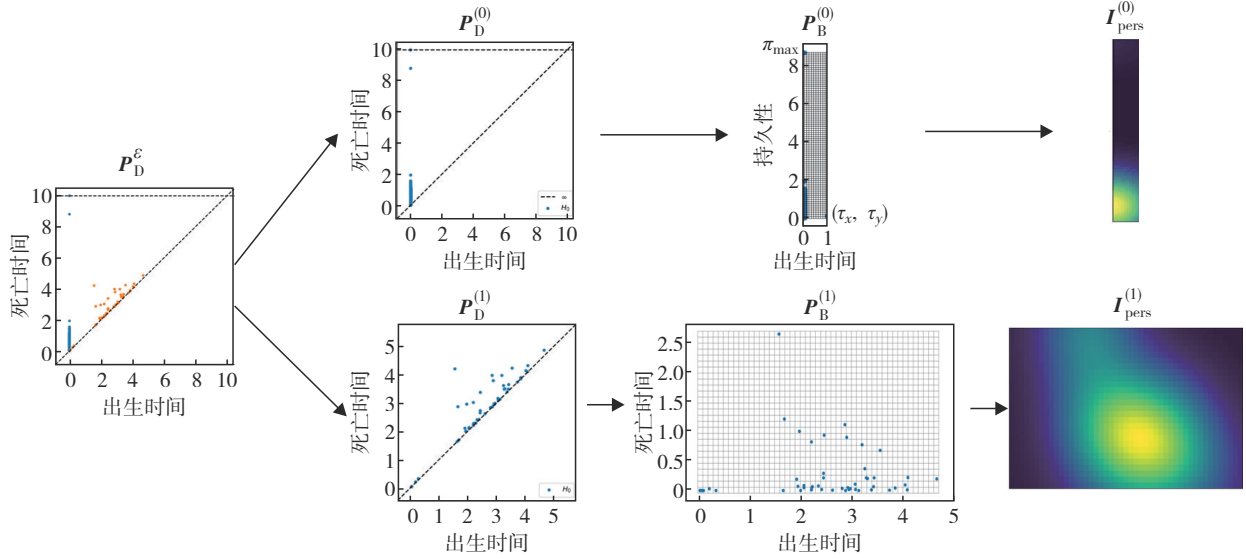


图2 持久性图像转换到持久性映像的过程

Fig. 2 Illustrates the process of converting a persistence diagram to a persistent image.

4) $I_{\text{pers}}^{(0)}$ 和 $I_{\text{pers}}^{(1)}$ 具有不同维度, 可采用零填充方法将各维度的持久性映像对齐到相同大小, 然后, 将 $I_{\text{pers}}^{(0)}$ 和 $I_{\text{pers}}^{(1)}$ 通过最大值堆叠方法进行堆叠, 得

$$I_{\text{stk}} = \max(I_{\text{pers}}^{(0)}, I_{\text{pers}}^{(1)}) \quad (8)$$

此外, 为实现对持久性映像的图像化处理, 以增强拓扑特征表示能力, 捕获更丰富和全面的信息内容, 并确保其通道数与原始图像保持一致, 将 $I_{\text{pers}}^{(0)}$ 、 $I_{\text{pers}}^{(1)}$ 和 I_{stk} 三幅持久性映像图堆叠形成一个三通道图像 $I_{\text{pers}}^{\text{final}}$, 作为最终所需的持久性映像, 如图3所示. Channel₁为从 $I_{\text{pers}}^{(0)}$ 和 $I_{\text{pers}}^{(1)}$ 经过堆叠得到的更具优势的特征信息. Channel₂保留了 $I_{\text{pers}}^{(0)}$ 原始的特征维度信息, Channel₃保留 $I_{\text{pers}}^{(1)}$ 原始特征情况.

1.3 多通道特征融合

为有效融合拓扑特征和视觉特征, 图1中 C_{vision} 提取图像的视觉数值特征, C_{topo} 提取数据的拓扑结构特征, 则由式(3)可将得到的两类特征按通道拼接, 形成内容更丰富的组合特征

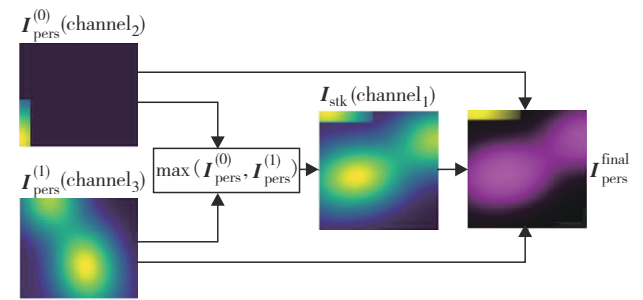


图3 持久性映像的生成过程

Fig. 3 Generation process of persistent images.

$$\Psi = C_{\text{vision}}(D) \oplus C_{\text{topo}}(I_{\text{pers}}^{\text{final}}) \quad (9)$$

1.4 AM

为提高模型对关键特征的关注度, 并增强特定特征的表达能力, 在拼接后的组合特征 Ψ 上应用SE模块.

首先, 对拼接后的特征图 Ψ 进行全局平均池化, 得到每个通道的重要性为

$$z = \frac{1}{H \times W} \sum_{a=1}^H \sum_{b=1}^W \Psi(a, b) \quad (10)$$

其中, $z \in \mathbf{R}^c$; a 和 b 分别为特征图在高度方向和宽度方向上的坐标索引, $a, b \in [1, H]$.

然后, 利用全连接网络学习得到对应的通道权重为

$$s = \sigma[W_n \cdot \delta(W_1 \cdot Z)] \quad (11)$$

其中, $W_1 \in \mathbf{R}^{\frac{c}{g} \times c}$ 是全连接层的权重矩阵; σ 和 δ 分别为 sigmoid 激活函数和 ReLU 激活函数; g 为缩放系数; W_n 为除 W_1 外的其他全连接层的权重矩阵.

最后, 经 SE 模块调整后每个通道的特征值调整为

$$\Phi = \Psi(a, b, c) \cdot s(c) \quad (12)$$

其中, c 为通道索引. 将 Φ 输入到全连接层, 可得到分类任务的输出.

2 实验与结果分析

2.1 实验配置

实验所用高性能工作站的中央处理器(central processing unit, CPU)采用 2 个 Hygon 3350, 图形处理器(graphic process unit, GPU)为 2 个 NVIDIA GeForce RTX 4090, 24 Gbyte 显存, 内存为 128 Gbyte DDR4-3200 ECC, 存储为 256 Gbyte SSD (SATA 6G)+ 8 Tbyte HDD, 操作系统采用 Ubuntu 22.04, 使用 PyTorch 1.12.1 作为深度学习框架. 数据集与评价指标

实验使用 3 个不同领域的数据集, 数据集 1 为 Intel Image Classification 数据集, 包含来自世界各地自然场景的图像数据 17 046 幅图像, 为 6 分类任务; 数据集 2 为性别分类的 Gender Classification 200K Images/CelebA 数据集, 包含 20 2381 幅图像, 是一个 2 分类任务; 数据集 3 是由 20 位中国著名书法家所书写的 105 080 幅图像组成的汉字数据集 Chinese Calligraphy Styles by Calligraphers, 是 20 分类任务.

为全面评估模型性能, 选用准确率(A)、精确率(P)、召回率(R)以及 F_1 值作为主要的评价指标.

2.3 实验设计及参数设置

在拓扑数据分析阶段, 用 Ripser 库处理点云或数据点集合. 使用 Ripser 库处理灰度矩阵的点云数据时, 需先计算点云中各点之间的距离, 并基于此距离分析并选择合适的距离阈值, 构建复形并获取

持久同调类的结构信息进而生成 0 维和 1 维的持久性图, 并得到持久性映像, 用于揭示数据集的拓扑结构特征.

选择典型 CNN 模型包括 VGG16(2014)及其改进版本 VGG16_BN(2015)、Inception v1(2014)、DenseNet121(2017)、ResNet34(2015), 并引入 EfficientNet v2(2021)和 ConvNeXt(2022)模型作为基线模型, 按照前述方法结合 TDA 构建了 TDA-VGG16、TDA-Inception v1、TDA-DenseNet121、TDA-ResNet34、TDA-EfficientNet 和 TDA-ConvNeXt 模型. 设置批量大小为 32, 训练周期为 50, 使用随机梯度下降(stochastic gradient descent, SGD)优化器更新参数.

2.4 结果分析

表 1 给出了不同模型在 3 个数据集上的性能表现. 由表 1 可见, TDA-CNN 模型在不同数据集上均表现更优, 尤其是在复杂性较高的数据集 3 上, TDA-CNN 方法分别将基线模型 VGG16、EfficientNet v2、DenseNet121 的图像分类准确率提升了 21.89%、22.16% 和 8.26%. TDA-VGG16 模型的性能指标显著高于其他对比模型, TDA-VGG16 和 TDA-DenseNet121 的 F_1 分数分别达到 0.974 4 和 0.931 6, 显著优于基线模型和其他改进模型. TDA 通过捕获数据的拓扑特征, 有效提升了 CNN 对复杂数据结构的理解和表征能力, 使其在处理具有复杂模式的数据集时表现出更强的鲁棒性和泛化能力.

2.5 方法分析

选择采用 CNN 模型的 GoogleNet 分类模型作为基线模型, 并构建融合了 TDA-CNN 的 TDA-GoogleNet 模型, 对两模型进行可视化分析(图 4), 以验证 CNN 模型和 TDA-CNN 模型特征对于分类准确率的重要性. 图 4 中的纵坐标表示特征指标数量, 横坐标表示特征重要性.

由图 4 可见, TDA-GoogleNet 模型提取到的特征数量显著超过 GoogleNet 模型, 且更多特征具有较高的重要性, 最高超过了 1.75, 表明 TDA-CNN 模型能够从数据中提取出更多有用且重要的特征, 增强了特征的表示能力和学习能力, 充分证明了 TDA-CNN 模型的有效性和优越性.

3 消融实验

为明确各组件在图像分类任务中的具体作用和

表1 不同模型在3个数据集上的性能指标

Table 1 Comparison among the different models on three datasets

数据集	模型	A	P	R	F_1
1	VGG16	0.897 6	0.897 7	0.897 6	0.897 4
	VGG16_BN	0.902 6	0.902 4	0.902 6	0.902 1
	TDA-VGG16	0.906 6	0.906 3	0.906 6	0.906 2
	Inception v1	0.886 9	0.886 5	0.886 9	0.886 6
	TDA-Inception v1	0.912 3	0.911 8	0.912 3	0.911 8
	DenseNet121	0.905 9	0.905 6	0.905 9	0.905 7
	TDA-DenseNet121	0.915 7	0.915 3	0.915 7	0.915 4
	ResNet34	0.902 6	0.902 5	0.902 6	0.902 4
	TDA-ResNet34	0.910 6	0.910 0	0.910 6	0.910 2
	EfficientNet V2	0.862 2	0.861 9	0.862 2	0.861 7
	TDA-EfficientNet V2	0.919 4	0.919 1	0.919 4	0.919 1
	ConvNeXt	0.915 2	0.915 4	0.915 2	0.915 0
2	TDA-ConvNeXt	0.920 7	0.920 6	0.920 7	0.920 5
	VGG16	0.913 6	0.908 1	0.885 2	0.896 5
	VGG16_BN	0.932 9	0.933 2	0.933 2	0.933 1
	TDA-VGG16	0.966 8	0.966 0	0.966 1	0.9660
	Inception v1	0.931 7	0.917 7	0.921 0	0.919 4
	TDA-Inception v1	0.940 2	0.927 2	0.931 6	0.929 4
	DenseNet121	0.938 6	0.940 1	0.912 8	0.926 2
	TDA-DenseNet121	0.947 1	0.933 9	0.941 4	0.937 6
	ResNet34	0.936 8	0.925 9	0.924 8	0.925 4
	TDA-ResNet34	0.937 2	0.925 5	0.925 9	0.925 7
	EfficientNet V2	0.925 6	0.925 5	0.925 6	0.925 5
	TDA-EfficientNet V2	0.958 3	0.958 3	0.958 3	0.958 3
3	ConvNeXt	0.956 3	0.956 4	0.956 3	0.956 4
	TDA-ConvNeXt	0.958 7	0.958 7	0.958 7	0.958 7
	VGG16	0.799 5	0.800 6	0.799 5	0.794 3
	VGG16_BN	0.849 2	0.848 0	0.849 2	0.845 8
	TDA-VGG16	0.974 5	0.974 4	0.974 5	0.974 4
	Inception v1	0.824 5	0.820 3	0.824 5	0.814 5
	TDA-Inception v1	0.869 0	0.866 5	0.869 0	0.865 9
	DenseNet121	0.861 2	0.860 2	0.861 2	0.857 9
	TDA-DenseNet121	0.932 3	0.931 6	0.932 3	0.931 6
	ResNet34	0.841 4	0.838 4	0.841 4	0.837 4
	TDA-ResNet34	0.868 5	0.866 9	0.868 5	0.866 1
	EfficientNet V2	0.711 1	0.703 3	0.711 1	0.694 8
	TDA-EfficientNet V2	0.872 2	0.869 8	0.872 2	0.869 7
	ConvNeXt	0.907 9	0.907 3	0.907 9	0.903 2
	TDA-ConvNeXt	0.917 4	0.916 7	0.917 4	0.914 4

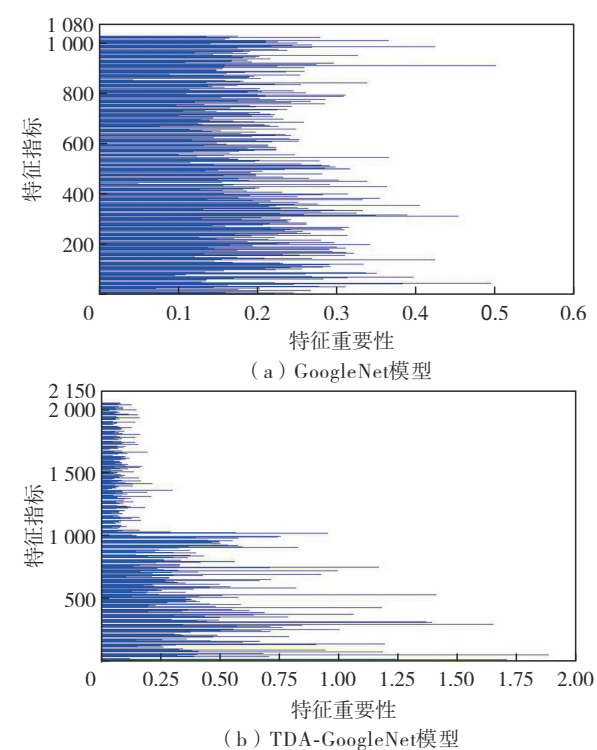


图4 GoogleNet模型和TDA-GoogleNet模型的特征对于分类精度的重要性

Fig. 4 The importance of features for classification accuracy in (a) GoogleNet model and (b) TDA GoogleNet model.

影响, 选择DenseNet121模型作为基线模型进行消融实验. 采用TDA方法生成PI, 并引入SE模块, 在数据集1上对各组成部分进行分析, 以评估各组件对模型性能的贡献和影响, 结果见表2.

表2 DenseNet121模型在数据集1上的消融实验

Table 2 DenseNet121 ablation experiments on dataset 1

实验	SE	PI	A	P	R	F_1
1	×	×	0.905 9	0.905 6	0.905 9	0.905 7
2	×	√	0.911 0	0.910 7	0.911 0	0.910 8
3	√	√	0.915 7	0.915 3	0.915 7	0.915 4

注: ×表示未添加相应组件; √表示添加了相应组件.

表2可见, 在未使用额外特征提取模块的实验组1基础上, 引入PI的实验组2在准确率、精确率、召回率和 F_1 值等指标上均有显著提升, 彰显了PI在捕捉数据复杂结构和关系方面的有效性. 进一步地, 实验组3在引入PI的基础上又加入了SE模块, 展现了更优异的性能, 准确率比实验组1提高了近1.0%, 比实验组2提高了近0.5%, 其他指标亦有显著提升. 这说明, SE与PI的加入能够更全面地理解和表示数据特征, 增强了模型的特征学习能力和分类效果.

结 语

提出一种拓扑结构特征与数值特征的融合方法, 通过 CNN 通道提取数值特征, 通过 TDA 通道提取拓扑特征, 然后融合这些特征来增强模型在特征学习和表示方面的能力, 并采用 SE 模块让模型自适应学习两种特征产生特征点的权重分配, 以进一步优化模型的特征提取和表示能力. 通过引入辅助坐标点, 解决了特征表示转换中存在持续时间为零或无穷大导致无法正常转换的问题. 最后, 在 Intel Image Classification、Gender Classification 200K Images/CelebA、Chinese Calligraphy Styles by Calligraphers 等数据集上和 VGG16、EfficientNet v2、DenseNet121 网络上验证了本文方法的有效性.

未来也可将 TDA 方法与更多机器学习和深度学习方法的相融合, 同时也需研究如何降低引入 TDA 带来的额外计算量.

基金项目: 国家自然科学基金资助项目(62172045); 北京联合大学校级科研资助项目(ZK10202403)

作者简介: 杨含(monica_yh2024@163.com), 北京联合大学硕士研究生. 研究方向: 拓扑数据分析与机器学习.

引 文: 杨含, 秦广军, 刘子源, 等. 基于拓扑数据分析与卷积神经网络的特征融合方法[J]. 深圳大学学报理工版, 2025, 42(5): 624-630.

参考文献 / References:

- [1] SMYS S, CHEN J I Z, SHAKYA S. Survey on neural network architectures with deep learning [J]. Journal of Soft Computing Paradigm, 2020, 2(3): 186-194.
- [2] GU Jiuxiang, WANG Zhenhua, KUEN J, et al. Recent advances in convolutional neural networks [J]. Pattern Recognition, 2018, 77: 354-377.
- [3] XIE Lingxi, YUILLE A. Genetic CNN [C]// 2017 IEEE International Conference on Computer Vision. Piscataway, USA: IEEE, 2017: 1388-1397.
- [4] BRONSTEIN M M, BRUNA J, LECUN Y, et al. Geometric deep learning: going beyond Euclidean data [J]. IEEE Signal Processing Magazine, 2017, 34(4): 18-42.
- [5] CARLSSON G. Topology and data [J]. Bulletin of the American Mathematical Society, 2009, 46(2): 255-308.
- [6] LI Ou. TDA 3D data model and feature simplification analysis [J]. Journal of Physics: Conference Series, 2021, 1883(1): 012023.
- [7] DEY T K, WANG Yusu. Computational topology for data analysis [M]. Cambridge: Cambridge University Press, 2022.
- [8] DEY T K, WANG Yusu. Computational topology for data analysis [M]. Cambridge: Cambridge University Press, 2022.
- [9] SARDIU M E, GILMORE J M, GROPE B, et al. Identification of topological network modules in perturbed protein interaction networks [J]. Scientific Reports, 2017, 7(1): 43845.
- [10] LEE H, KANG H, CHUNG M K, et al. Persistent brain network homology from the perspective of dendrogram [J]. IEEE Transactions on Medical Imaging, 2012, 31(12): 2267-2277.
- [11] NIELSON J L, PAQUETTE J, LIU A W, et al. Topological data analysis for discovery in preclinical spinal cord injury and traumatic brain injury [J]. Nature Communications, 2015, 6(1): 8581.
- [12] BERRY E, CHEN Y C, CISEWSKI-KEHE J, et al. Functional summaries of persistence diagrams [J]. Journal of Applied and Computational Topology, 2020, 4(2): 211-262.
- [13] CHUNG Y M, LAWSON A. Persistence curves: a canonical framework for summarizing persistence diagrams [J]. Advances in Computational Mathematics, 2022, 48(1): 6.
- [14] BUBENIK P. The persistence landscape and some of its properties [C]// Topological Data Analysis. Cham: Springer International Publishing, 2020: 97-117.
- [15] ADAMS H, EMERSON T, KIRBY M, et al. Persistence images: a stable vector representation of persistent homology [J]. The Journal of Machine Learning Research, 2017, 18(1): 218-252.
- [16] SWENSON N, KRISHNAPRIYAN A S, BULUC A, et al. PersGNN: applying topological data analysis and geometric deep learning to structure-based protein function prediction [EB/OL] [2024-10-10]. <https://arxiv.org/abs/2010.16027>.
- [17] CANG Zixuan, WEI Guowei. TopologyNet: topology based deep convolutional and multi-task neural networks for biomolecular property predictions [J]. PLoS Computational Biology, 2017, 13(7): e1005690.
- [18] AHMED F, COSKUNUZER B. Topo-Net: retinal image analysis with topological deep learning [EB/OL] [2024-10-10]. <https://doi.org/10.1101/2024.02.03.24302291>.
- [19] HU Chuanshen, LAWSON A, CHEN J S, et al. TopoResNet: a hybrid deep learning architecture and its application to skin lesion classification [J]. Mathematics, 2021, 9(22): 2924.
- [20] ATHIWARATKUN B, KANG K. Feature representation in convolutional neural networks [EB/OL] [2024-10-10]. <https://arxiv.org/abs/1507.02313>.
- [21] VER HOEF L, ADAMS H, KING E J, et al. A primer on topological data analysis to support image analysis tasks in environmental science [J]. Artificial Intelligence for the Earth Systems, 2023, 2(1): e220039.

【中文责编: 英子; 英文责编: 木柯】