arXiv:2303.04231v1 [cs.LG] 7 Mar 2023

一种表征脑状态的拓扑分类器：形状何以比方差更重要

Aina Ferr`a [1], Gloria Cecchini [1], Fritz-Pere Nobbe Fisas [1], Carles Casacuberta [1,2], Ignasi Cos [1,2,3]

摘要

尽管机器学习分类器在监督方式下区分复杂数据集时取得了令人瞩目的准确率，但其大多数运作方式仍不足以帮助人们深入直观地理解数据结构，更重要的是，无法深入理解给定数据集所表征的现象。相比之下，拓扑数据分析（TDA）致力于借助持久性描述符研究数据云的形状，并对所考察数据集的特定拓扑特征进行定量表征。

本文介绍一种新颖的基于拓扑数据分析的分类器，其原理是评估向数据子集中加入新输入后，拓扑度量发生的可量化变化。我们将该分类器用于一个高维脑电（EEG）数据集，该数据集记录了11名参与者在一项决策实验中的脑电信号；实验通过操纵社会压力诱发了三种动机状态。对经过带通滤波器处理的脑电信号进行处理后，我们计算了与每种动机状态相关的持久图的轮廓，并根据未标记信号对各参考轮廓的影响对其进行分类。结果表明，除准确率达到最近邻分类器的水平外，拓扑数据分析分类器还提供了对数据集结构的形式化直观认识，并估计其内在维数。为此，我们将降维方法纳入流程，发现拓扑数据分析分类器的准确率通常并不受解释方差影响，而是取决于形状；这与大多数机器学习分类器的情况相反。

引言

深度学习能够捕捉大型数据集的详细高维统计特征，因此被誉为解决各类复杂问题的通用方法 [1, 2, 3]：从图像分析[4, 5] 、目标识别 [6], 、气候预测 [7, 8], 、基因组分析和行为预测[9] ，到脑动力学中的隐变量识别 [3]。然而，深度学习与任何黑箱人工智能技术一样，也有其缺点：结果和数据集结构的可解释性较差，这必然限制了其可解释性 [10, 11, 12]。更令人困惑的是，深度学习偶尔会出现难以预测的聚类错误，例如把狮子错认成图书馆，或把台灯错认成交通信号灯 [13]。尽管这些错误看似荒唐，但正是缺乏合理解释，导致此类不可预测的失误，也使深度学习面临严重的信任问题，尤其是在涉及危及生命的决策的人工智能应用场景中。

[1]巴塞罗那大学（UB）数学与计算机科学系，Gran Via de les Corts Catalanes 585号，08007 巴塞罗那，加泰罗尼亚，西班牙[2]巴塞罗那大学数学研究所（IMUB）[3]加泰罗尼亚 Serra-Hunter 研究员计划，西班牙 本研究得到加泰罗尼亚政府CERCA计划的支持（I. Cos）；并得到欧盟地平线2020框架计划研究与创新项目的支持，具体资助协议编号为945539，即人类脑计划SGA3（I. Cos）；此外，还得到MCIN/AEI/10.13039/501100011033的资助，资助编号为PRE2020-094372（A. Ferr`a），以及项目PID2019-105093GB-I00（I. Cos、A. Ferr`a）和PID2020-117971GB-C22（C. Casacuberta）。

Keywords: 拓扑数据分析；机器学习；分类；神经科学；脑状态；可解释性*Mathematics Subject Classification:* 55N31、92C20、62R40、68T09

相比之下，拓扑是数学的一个分支，致力于通过形式化方法刻画高维数据集的结构 [14, 15]。简言之，如果一个数据集可以表示为超空间中的点云，拓扑数据分析（TDA）便可以借助涵盖所有维数的一组度量（或摘要），刻画其连通性、点云中的环，或其中包含的簇的数量。因此，尽管拓扑数据分析并非为解决问题或捕捉数据变化本身而设计，但若将其应用于分类问题中不同类别的数据集，就能得到各类别特有的摘要。与此相符，一些研究提出将拓扑数据分析度量用作初步阶段，为机器学习分类器提取特征，例如用于医学图像分析 [16, 17, 18], 、化学成分 [19, 20], 或计算机科学问题 [21]。因此，这些分类问题并非在多维数据集或后续的降维空间中进行 [22, 23], ，而是在具有明确含义的拓扑特征向量领域中进行 [24, 25, 26], ，从而提升可解释性。

尽管这是一条颇具前景的研究路径，能够帮助我们追溯类别特性与数据结构之间的联系，但我们面临的难题是如何将拓扑数据分析用作分类器。简言之，如果不同的数据子集（属于不同类别）产生不同的拓扑摘要，那么这些摘要之间的差异本身就可能直接用于分类。为此，我们在此介绍一种基于这一原理的分类器。与传统机器学习分类器相比，这种分类器的优势在于，除了给出准确率和混淆矩阵外，还应能让人深入了解数据集中哪些具体方面导致了类别之间的可分性。已有研究采用拓扑持久性开展分类或半监督学习，见 [27]和 [28]。

我们以一个亟需可解释性的棘手分类问题作为分类器的测试平台：执行特定任务时的皮层脑网络状态 [29, 30, 31]。具体而言，我们重点刻画动机脑网络；该网络是在一项决策任务中，根据不同程度的社会压力下进行精确伸手动作来定义的 [32] 。我们分析了从十一位参与者身上记录的一组脑电图（脑电图），这些记录涵盖了决策逐步形成的过程。我们将问题转化为三分类问题，旨在借助拓扑数据分析分类器解释这些状态之间的差异。我们得出两项主要结果：第一，拓扑数据分析分类器的准确率与最近邻分类器相当；第二，准确率很大程度上取决于数据的形状，而非降维投影所达到的解释方差。总之，特定的拓扑描述符确实能够可靠地整体刻画高维神经状态，并为数据可解释性提供一条有别于机器学习的途径。

1 材料与方法

1.1 持久性摘要

拓扑数据分析是数学的一个分支，基于代数拓扑，旨在检测并表征数据集的结构特征，例如稀疏性、尖峰或环。其主要工具是持久同调 [33, 34, 35, 36], ，这是配备实值过滤函数的单纯复形的一种代数特征。持久同调非常适合描述点云在一系列分辨率尺度下的形状。

形式上，点云是一个有限子集 $X \subset \mathbb{R}^d$ ，其中 $d \geq 2$ 通过欧氏距离赋予度量空间结构。与 X 相关联的*Vietoris–Rips*过滤 [37, 38] 是一个嵌套的抽象单纯复形族 $V_t(X)$，其中 $t \geq 0$，而 $V_t(X)$为 X 中的每个点设有一个顶点，并且对于 $v_0, \ldots, v_k$ 中的每组点，若 $\|v_i - v_j\| \leq 2t$ 对所有 i, j均成立，则设有一个 k-单纯形 $k \geq 1$。

包含映射 $V_s(X) \subseteq V_t(X)$ 对 $s \leq t$ 诱导出态射 $H_n(V_s(X)) \to H_n(V_t(X))$，其中 $n \geq 0$， H_n 表示 n维单纯同调 [39]。我们使用有限域上的系数计算同调——采用 GUDHI Python库 [40]，默认选择域 $\mathbb{F}_{11}$。同调生成元的出生参数 b 是使 $H_n(V_t(X))$ 包含给定生成元的最小 n 维值 t；而死亡参数 d 是该生成元映射为零时的最小值 t。同调生成元的持久性 或生存期 为二者之差 $d - b$。

与 X 的 Vietoris–Rips过滤相关联的持久性图 由同调维数为 n 时的一组 n维同调生成元基的所有出生-死亡对（b, d）构成，这些点绘制在 $\mathbb{R}^2$中第一象限的对角线 $y = x$ 上方。靠近对角线的点（即生存期较短的点）可能对应无关紧要的现象，而生存期较长的点则反映给定数据集的持久形状特征。不过，在某些情况下，对角线附近的点的分布也包含不应忽视的相关信息。

0 维持久性图包含关于 Vietoris– Rips复形连通分支的信息，具体而言，描述了参数 t增大时连通分支如何合并。1 维持久性图描绘了 1-循环的出现与消失，而 $n \geq 2$ 维的持久性图则表示 Vietoris–Rips复形中 n维空腔的演变。

两项基本结果赋予持久同调所需的稳健性，使其成为一套严谨且适用于现实应用的数学理论。第一项是持久性图定义良好这一事实 [34], ，也就是说，它们不依赖于同调生成元基底的选择。第二项是稳定性 [41, 42], ，即数据中的微小扰动只会导致相应持久性图发生微小变化。

持久性图之间的差异度可以用瓶颈距离度量；稳定性定理在声称两个图彼此接近时，所指的正是这种距离。两个图 D_1 与 D_2 之间的瓶颈距离定义为点 $x \in D_1$ 与 φ之间按坐标计算的最大距离 $x \in D_2$，其中 φ 是在所有可能的、从 D_1 到 D_2的双射中使结果最小的匹配，并允许将各自对角线上的点作为匹配对象 [33]。然而，配备瓶颈距离的持久性图并不适合进行统计分析。一方面，无法计算持久性图的平均值 [43]；另一方面，许多用途都要求使用希尔伯特空间中的持久性图表示。为此，持久性图的一种便捷概括是其景观 [44, 45]。持久性景观是由一系列分段线性函数构成的序列：将图顺时针旋转 45 度（图 1），然后在所得图形中，为每个 $k \geq 1$ 选择第 k高的点 [44]。

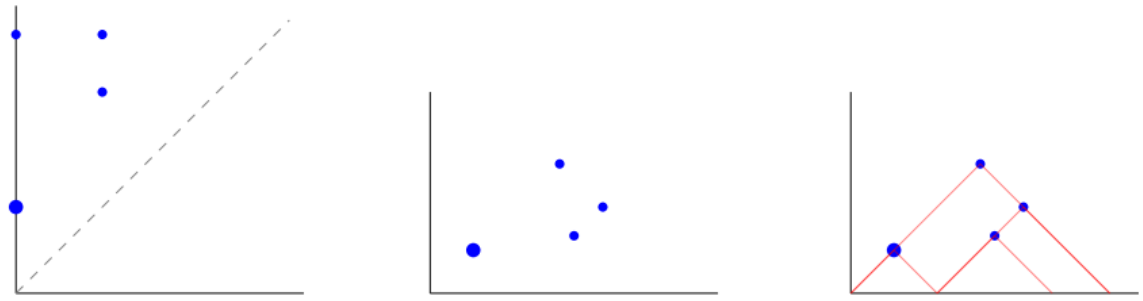


图 1：通过旋转 45°并缩放（中图），由持久性图（左图）得到的持久性景观（右图）。较大的圆点表示重数较大。

更准确地说，对于给定持久性图中的每个点（b, d），都考虑一个帐篷函数$\Lambda_{(b,d)}(t) = \max\{0, \min\{t-b, d-t\}\}$并定义 $\lambda_k \colon \mathbb{R} \to \mathbb{R}$对每个 $k \geq 1$，令其为 $\lambda_k(t) = \text{kmax}\{\Lambda_{(b_i,d_i)}(t)\}$，其中 $\{(b_i, d_i)\}$ 是给定持久性图中所有点的集合，而 kmax 返回给定数字集合中第 k 大的值；若不存在第 k 大的值，则返回零。因此，当 k 足够大时， $\lambda_k = 0$。最初的景观层级 λ_1, $\lambda_2 \ldots$表示持久性图中最显著的特征，而最后几个层级则对应于靠近对角线的点，因而代表短暂现象。

尽管持久性景观很有用，但处理大型数据集时会占用大量内存。为解决这一问题，轮廓于 [46] 被提出，其做法是对构建景观所用的帐篷函数进行加权平均：

$$\phi_w(t) = \frac{\sum_{i=1}^{m} w_i \, \Lambda_{(b_i,d_i)}(t)}{\sum_{i=1}^{m} w_i}.$$

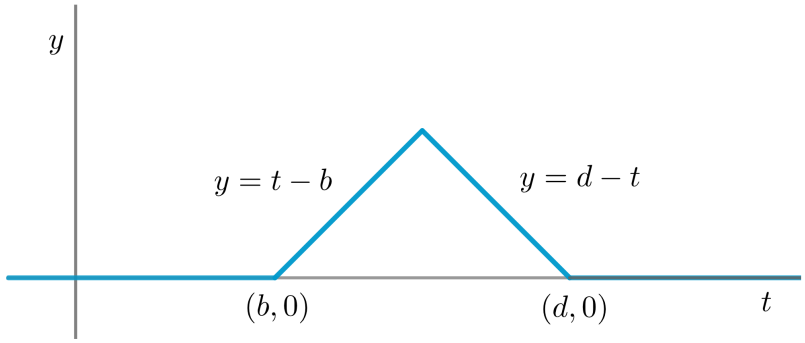


图 2：出生参数 b 和死亡参数 d 对应的帐篷函数。

本文选择寿命 $w_i = d_i - b_i$ 作为权重，并将由此得到的轮廓用作分析中的持久性摘要。由于寿命必须是有限的，我们会舍弃零维中具有无限持久性的点。

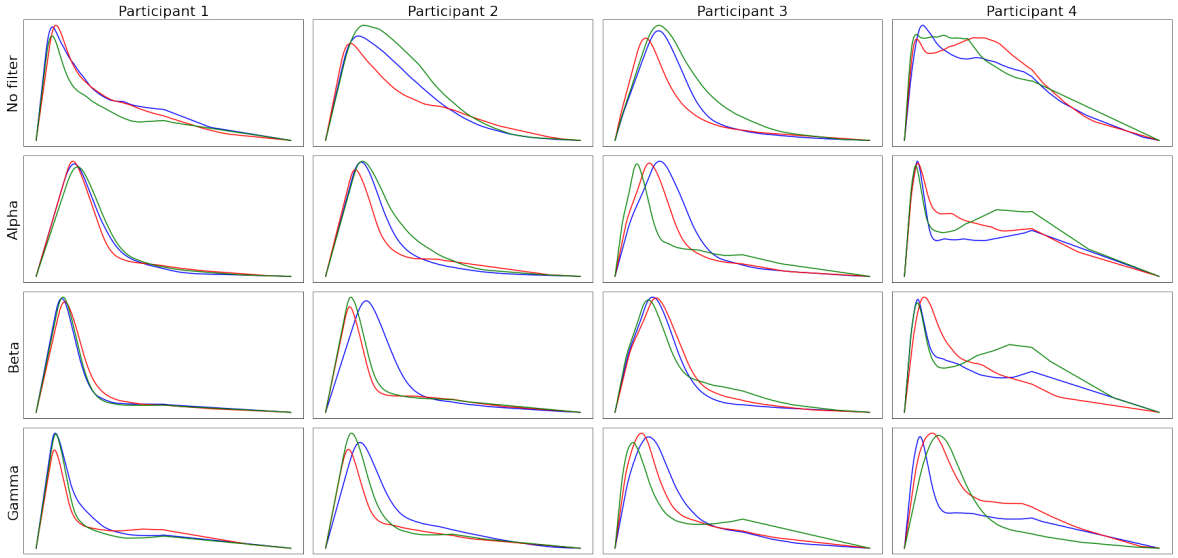


图 3：未进行降维的源空间中，参与者 1–4 在每种动机状态下（M_0：蓝色， M_1：红色， M_2：绿色），各频带（阿尔法、贝塔、伽马）以及未过滤数据集（未使用滤波器）对应的零维持久性图轮廓。数据集将在接下来的小节中介绍。

我们将轮廓离散化为包含 1000 个分量的向量，并计算这些向量之间的欧几里得距离，以衡量持久性图之间的差异度。

分类时仅使用对应于零阶同调维数的持久性轮廓。本研究未考虑更高维数，因为通常情况下，向训练数据集添加单个点不会使持久性摘要发生显著变化，零维的情况除外。

1.2 非拓扑分类器

典型的分类过程首先将带标签数据集 D 划分为训练集和测试集，并假设已知训练集中每个输入 x 所属的标签类别 c 。

准确率 准确率定义为测试集中被正确分类的数据点所占的百分比。

因此，若我们有一个数据集 D，其中包含 m 个训练数据点，如 $D = \{(x_i,\ c_i)\}$ 所示，且$i = 1,$ ……，m，那么分类的目的是：对于任意给定的点 x，返回其最佳预测类别 $c(x)$。因此，对于分类器算法而言，一个关键问题是定义数据点之间适当的相似性度量，使相似的点属于同一类别。通常，给定两个点 x 和y，它们之间的相似程度由距离函数 $d(x,\ y)$ 量化。例如，最近邻算法 使用欧氏距离。给定测试数据集中的一个数据点 x，以及训练数据 $\{(x_i,\ c_i)\}$，最近邻算法按如下方式对 x 进行分类：

1. 计算 点 x 到每个训练数据点的欧氏距离 $d_i = d(x,\ x_i)$ $x_i.$
2. 在训练数据集中找到一个点 x_{i_*}，使得 $d_{i_*} =$ 最小。
3. 指定类别标签 $c(x) = c_{i_*}$。如果存在标签不同且距离相等的点，算法会选择包含点数最多的类别。

而我们的基于持久性的拓扑分类器采用以下步骤：

1. 根据给定的标签将训练集划分为不同类别。
2. 对于训练集中的每个类别标签 c，计算其在同调维数 0 下对应的轮廓 S_c，并以寿命作为权重。
3. 要对来自测试集的输入 x 进行分类，将 x 加入每个类别标签 c对应的训练数据点云 X_c。然后，重新计算数据集 $X_c \cup \{x\}$的持久性轮廓 $S_{c,\ x}$，最后计算新得到的轮廓与原轮廓之间的欧氏距离 $d(S_c,\ S_{c,\ x})$。
4. 指定类别标签 $c(x) = c_*$，其中 $c_* = \operatorname{argmin}_c \{d(S_c,\ S_{c,\ x})\}$。

其基本假设是，不同类别的点云呈现出可辨识的不同形状。更准确地说，如果向正确类别的点云中添加一个点，所得拓扑不应发生根本性变化，也就是说，所有距离都应保持较小。相比之下，如果向其他类别的点云中添加一个点，拓扑则应出现更明显的变化。

1.3 动机脑状态时间序列

出于实际考虑，我们选择在先前采集的数据集上测试拓扑数据分析分类器，而没有为此临时采集新数据。下文将介绍该数据集及其采集背景。除了可以利用一组有利的基线结果进行比较、无疑有助于验证拓扑数据分析分类器之外，我们之所以特意选择该数据集，还因为我们希望拓扑数据分析分类器能够比机器学习算法更深入地刻画该数据集的底层结构及其生成过程。

奖励和动机是人类行为的两种基本驱动力。与此相符，神经科学领域的大量研究旨在仔细识别并刻画奖励处理脑中枢，这些研究大多基于对人类功能性磁共振成像（fMRI）记录的分析。然而，仍有一个问题尚未得到充分解答：动机的不同表现形式在脑网络中如何分布？换言之，社会动机脑网络的分布是怎样的?

为回答这一问题，我们参考了一项在巴塞罗那开展的研究，该研究招募了十一名参与者[32]，，在一项通过社会压力调节动机的决策任务中，记录了参与者的高密度脑电图（EEG）。这项任务旨在正式刻画社会压力对运动决策和精确伸手动作所选运动参数的影响，以及社会动机相关脑网络 [32]。动机操控通过社会压力实现，具体依据参与者相对于虚拟伙伴的瞄准准确度进行调节。简而言之，

我们的设想是，伙伴的存在会引入一种隐性偏差，促使参与者提高准确率。我们告知参与者，他们将在一个玩家群体中完成任务，并且每个区块都会有一位来自该群体的不同伙伴与其一同完成任务。为强化参与者对伙伴存在的信念，每个试次结束时，我们都会显示一条水平绿色条，表示参与者达到的准确率（标准化至 0 到 100%），其下方是一条显示伙伴准确率的红色条。引入模拟伙伴的目的是造成隐性偏差，从而调节参与者更准确地伸手的动机。因此，我们指示每位参与者不要与伙伴竞争，只需专注于自己的表现，并忽略伙伴。为了量化虚拟伙伴存在所引入的偏差，我们设置了两类伙伴：准确率低于或高于参与者（图 1D）。我们还让伙伴与参与者的性别一致，以控制跨性别效应。

每个区块中的社会压力水平始终保持不变，虚拟伙伴也保持相同。不同区块中的伙伴则以平衡顺序进行更换。每位参与者完成两个实验阶段，每个阶段包含六个区块，每个区块由一百零八个试次组成。六个区块分为两组，每组三个区块。每组包含一个单人条件区块，以及两个与瞄准技能较低或较高的伙伴共同进行的区块。此操控旨在诱发三种不同的动机状态，具体取决于所施加的社会压力水平：

- 单人条件：参与者独自完成任务。在此条件下没有社会压力。
- 简单条件：参与者与一位技能低于自己的虚拟伙伴共同完成任务。预计此条件下会产生一定的社会压力。
- 困难条件：参与者与一位技能高于自己的虚拟伙伴共同完成任务。预计此条件下会产生较大的压力。

用于拓扑数据分析（TDA）的数据集记录了参与者执行此任务时的数据。数据集的结构由每位参与者在各个诱发社会压力水平下的脑电图（EEG）记录片段组成。由于我们特别关注动机基线的变化，因此在每个试次中选取了一个 1200 毫秒的感兴趣区间：从第一个刺激开始（即初始提示）前 800 毫秒开始，到其后 400 毫秒结束（见图 4A）；此时该试次中尚未呈现任何动作或刺激。为消除电源线造成的电气干扰，对每个通道应用了 50、100 和 150 Hz 的四阶陷波滤波器，并应用了 0.1 至 100 Hz 的四阶巴特沃斯带通滤波器，以限定感兴趣频带。在 200 毫秒区间内，脑电图电压超过 200 V 或电压变化步长/采样超过 50 V 的电极均从后续分析中剔除。数据经过基线校正，并使用每个实验阶段中试次区块类型 1 和 2（参与者单独进行任务时）的记录作为动机基线参考，对数据集进行 z 标准化。通过独立成分分析（ICA）去除眼部伪迹；该分析使用定制的 FieldTrip 开源工具箱（www.fieldtrip.com）和 EEGLAB 脚本（http://sccn.ucsd.edu/eeglab，加州大学圣迭戈分校，美国加利福尼亚州）实现。识别与眼动相关的源的过程是半自动化的：首先，将每个源与记录眼动的电极信号进行相关性分析，以获得初步的相关性指标；其次，对所有源进行目视检查，以确认其形状和空间位置与眼部伪迹相符。随后移除与眼部相关的源，并通过逆转 ICA 过程获得清理后的信号。最后，再次应用 ICA 得到源空间，但将所得空间的维度限制为小于电极空间的维度，以捕获若干个将信号传递至电极的大脑独立区域。因此，每位参与者的源空间维度可能不同，通常少于 50。

数据在源空间中的投影是通过基于 EEGLAB 库并结合电极空间位置图的自定义 MATLAB 脚本计算的。研究使用 Brain Products Unicap 64（Brain Products GmbH，德国吉尔欣）配置，在电极布局之间建立空间参照，并进行源定位。我们假设头部为球形模型。请注意，ICA投影具有双重用途：首先，将电极信号中包含的信息沿着各维度间独立性最大的方向分布；其次，为大脑皮层表面上产生电极所记录信号的脑源位置提供粗略的解剖学估计。

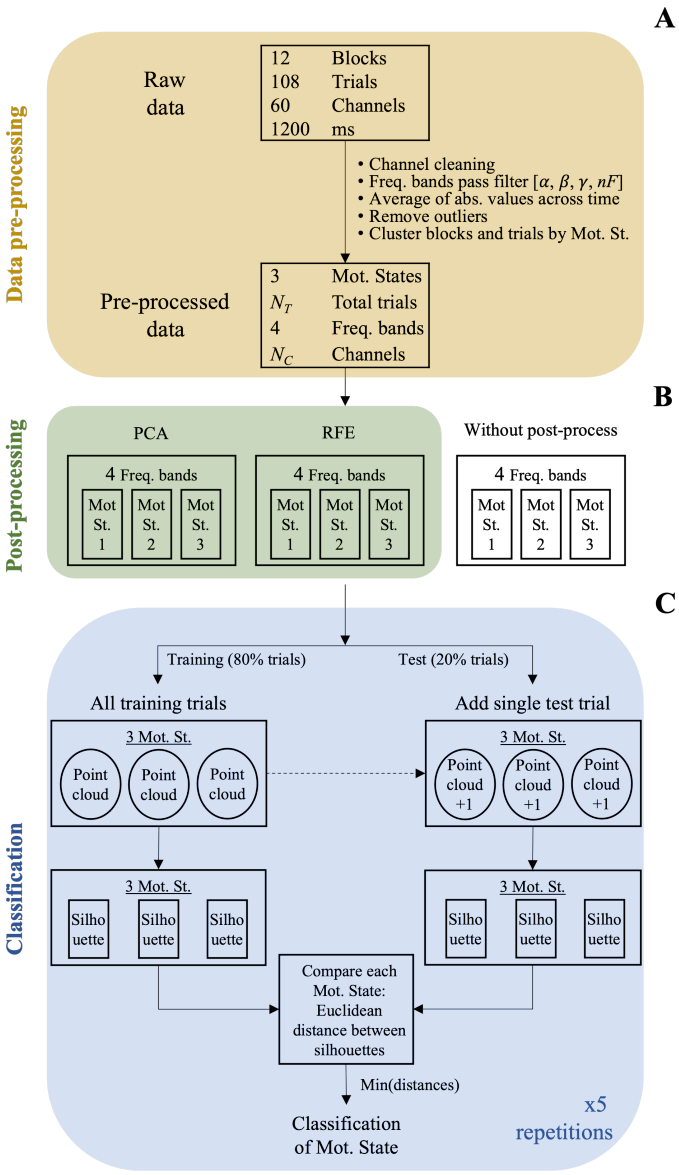


图 4：**A. 预处理示意图**。通道可以是电极，也可以是源。对每个通道的原始数据进行带通滤波，得到三个典型的脑电图频段（α：低；β：中；γ：高）。移除含有伪迹或噪声信息的通道，最终得到 N_C个通道。对每个通道的时间序列取绝对值并求平均。从所得数据集中移除离群值。**B. 后处理示意图**。使用主成分分析（PCA）或递归特征消除（RFE）。不对原始数据进行后处理。**C. 分类示意图**。使用 80% 的数据进行训练，20% 的数据用于测试。使用持久性轮廓进行分类。

总而言之，该数据集中的每个试次包含数量可变的通道（电极或源），记为 N_C，每个试次持续 1200 毫秒。每位参与者完成了 12 个区块，每个区块包含 108 个试次。请注意，这 12 个区块是在两个不同的实验阶段中记录的，每个阶段包含 6 个区块，以平衡动机区块的数量。

1.4 分析流程

研究中的 11 名参与者（4 名男性和 7 名女性，年龄为 55 ± 5.8）的每个数据集均包含 1200 ms 的 $\times\ 60$ 电极脑电图片段，这些片段分布在 12 个区块中（每个实验阶段 6 个区块），每个区块包含 108 个试次（图 4A）。每个区块内，促成特定动机状态的社会压力水平保持不变；每种动机状态各有 4 个区块（每个实验阶段各 2 个）。

预处理

- 首先，通过目视检查数据，识别并剔除噪声通道或伪迹通道，以免其参与后续分析。我们将经过此步骤后保留下来的通道数量称为 N_C。
- 其次，脑电信号的频谱范围为 0.01–100 Hz，我们将其划分为三个典型频段： α (8–15 Hz)、β (15–32 Hz)、 γ (32–80 Hz)，由此得到原始时间序列的三个带通版本。对每个信号均采用相同方法进行分析，彼此独立，并与原始信号未经滤波的基线版本一同分析。这种频段划分依据的是大脑功能与特定频段及电极位置上的功率波动之间已确立的关联。带通滤波使用 Python 编写的自定义脚本完成，调用 scipy.signal 库中的 iirfilter() 函数 [47]。以往研究的分析表明，动机相关调制可能属于高伽马频段，因此动机相关偏差可能更好地编码在较高频段中 [32]。

- 第三，我们在 1200 ms 的观测窗口内，对每个电极或源的绝对值取平均，得到一个按 N_T 试次 $\times$ N_C 通道组织的矩阵数据集。对于数据点范数的平均值，超出其两倍标准差的离群值不纳入后续分析。

我们分别对每个带通滤波后的信号以及未经滤波的原始信号进行分类。

后处理

我们使用 sklearn库 [48],对预处理后的数据集应用了两种不同的降维方法，得到的维数介于 2 至 10 之间。

- 主成分分析（PCA）是向由主成分构成的低维空间进行线性投影。其中，点云的第一个主成分能够解释最多的方差；在去除先前各主成分的影响后，每个后续主成分则解释剩余数据中最多的方差 [22]。

- 递归特征消除（RFE）通过依次移除对分类器准确率影响最小的坐标来实现 [49]。在本研究中，我们将 RFE 应用于预处理后的数据集，并使用逻辑回归模型为特征分配权重。

分类

每个数据集（针对每个频带）均经过适当格式化后，输入拓扑数据分析分类器（图 4B）。从此开始，各试次均被视为待分类的点云。分类操作按照“拓扑分类器”小节所述进行。每次分类重复 5 次，并对 5 次重复所得的准确率取平均值。每次重复时，后处理后的数据集均划分为 80% 的训练数据和 20% 的测试数据。对于原始数据（未进行降维），则使用预处理后的数据集。

分析既使用电极空间中的原始信号进行，也使用投影到脑源空间的信号进行。

2 结果

首先，为了建立一个准确率基线，以便评估降维技术对分类过程的影响，我们使用投影到源空间的原始脑电图信号测试了拓扑数据分析分类器；参见分析流程。其次，我们测试了两种降维方法（主成分分析和递归特征消除）对分类过程的影响。第三，为了进一步评估独立成分分析投影对分类器的影响，我们还测试了在电极空间中采集的原始数据集。上述每项测试均分别使用脑电信号的各个带通版本（α, β, γ）进行；参见分析流程。

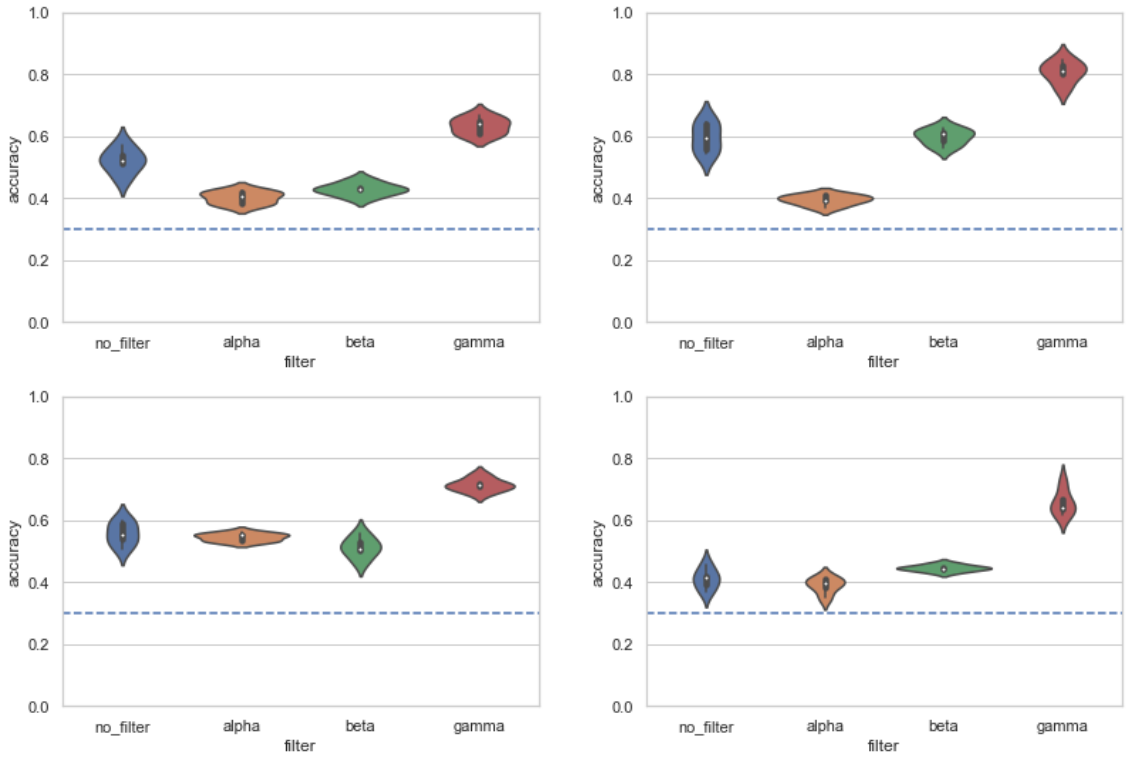


图5：采用源空间且未进行降维时，参与者1、3（右上）、7和11在各频带上的源空间拓扑分类器准确率。所有参与者的完整结果见附录。

源空间中的分类

我们针对每位参与者的数据集，在每个频带内进行分类，使用的是降维前投影到源空间的数据。图5展示了拓扑数据分析分类器对四位典型参与者的分类准确率。尽管不同参与者的分类结果存在一些差异，但主要结果显示，除一位参与者外，所有参与者在伽马频带的准确率均稳定居首，平均介于60%至80%之间。图5中的小提琴图呈现了数据的中位数、四分位距和核平滑概率密度。小提琴图使用 seaborn 库绘制 [50]。

以此为基线，我们评估了降维对分类的影响。其基本假设是，如果拓扑数据分析分类器关注点云形状的特征来判定可分性，那么在特定空间中重新排列的数据集应能产生更有利于分类的点云形状。为验证这一点，我们采用了两种互补的降维技术。对于每位参与者的数据集，我们首先进行了主成分分析

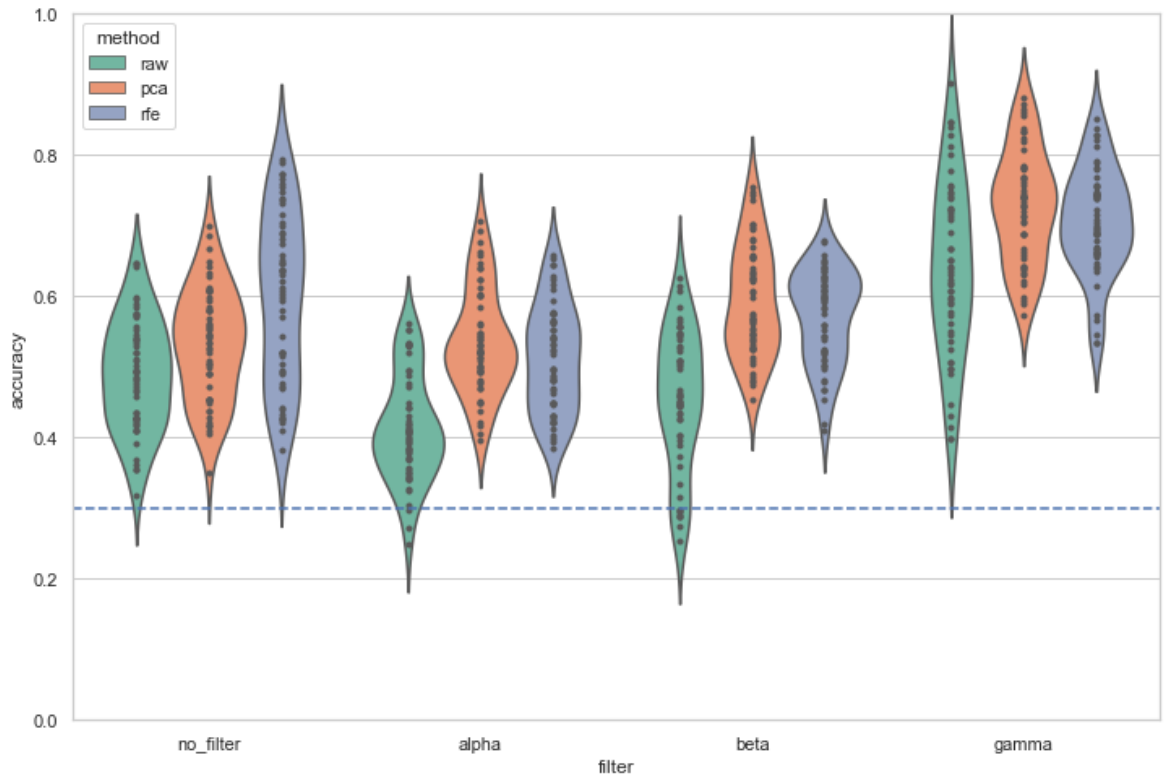


图6：对于参与者1至11，比较拓扑分类器在源空间中各频带下的基线准确率（原始）与采用主成分分析（主成分分析）和递归特征消除（递归特征消除）降维后得到的准确率。每个小提琴图包含每位参与者五次重复的准确率百分比。

（PCA）分解，选择能够解释所有参与者数据方差至少 95% 的维数，最终保留 2 至 10 个维数。

作为补充，我们在相同的数据集上采用了第二种降维技术。与主成分分析不同，递归特征消除（RFE）通过评估原始数据集中各特定成分对分类过程的贡献来进行。在源空间中执行时，这会产生源排序。我们还对主成分分析和递归特征消除得到的数据集进行了分类，涵盖每位参与者和每个频带。图13列出了这些分类的汇总结果，显示了拓扑分类器对11位参与者在各频带下采用每种降维方法时的准确率。图13中的准确率是对每个维数下五次重复所得的准确率取平均后得出的。由于有三个不同的类别，机会水平为0.33。

这些分析得出了两个主要结果。首先，与拓扑数据分析分类器的基线分类结果以及以往机器学习分类器的结果一致，除一位参与者外，所有参与者在伽马频带中都取得了最高的准确率 [32]。其次，与降维会将点云重新排列成有利于分类的形状这一假设一致，我们的结果一贯证实，降维空间中的点云比基线高维数据集具有更高的准确率。

降维的影响

由于在低维空间中可以获得更高的分类准确率，因此需要探讨准确率与维度之间的关系。对于递归特征消除，我们旨在寻找最优的源数量

，具体依据先前的结果 [32]。为此，我们对投影到源空间的脑电数据应用递归特征消除，并使用所得空间测试拓扑分类器，同时将算法选出的源数量从 2 个逐步增加到 10 个。我们对每位参与者和每个频带并行进行分类。

PCA 的情况对维数效应进行了更深入的研究，这主要是由于准确率与解释方差之间存在关联。在大多数机器学习分类器中，二者之间存在直接关系：解释方差越高，准确率通常也越高，并且在达到 95% 的解释方差后往往趋于收敛。我们的假设是，TDA 技术能够从数据中捕捉到的形状概念与其变异性无关。相反，我们预计会存在一个最优的主成分数量，使所得的投影空间呈现出更有利于分类的形状。我们使用将主成分分析应用于源空间后得到的空间来测试拓扑分类器，并在每一步将选取的主成分数量从 2 增加到 10。我们记录了每次分类的准确率，以及在该步骤中由相应数量的成分所解释的方差。分类过程针对每位参与者和每个频带并行进行。为了进一步评估独立成分分析对数据形状的影响，我们还使用电极空间进行了相同的测试，即对脑电数据不应用 ICA 算法。

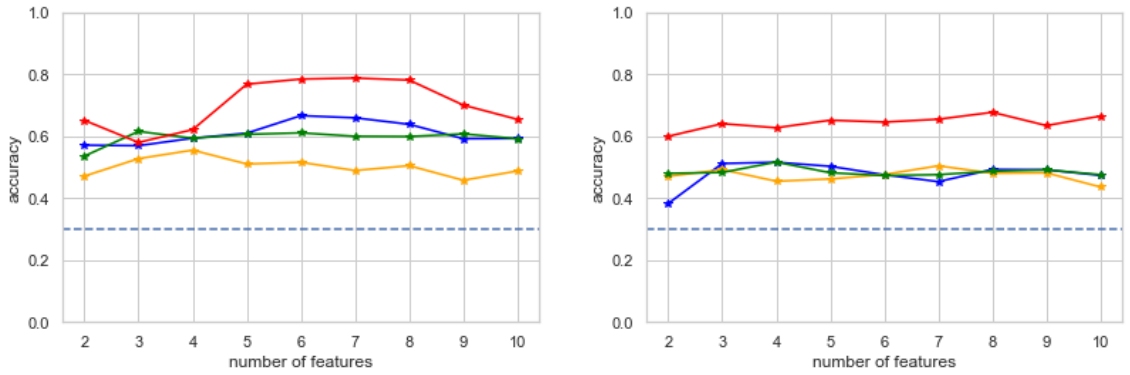


图7：参与者1和2中，拓扑分类器的准确率随递归特征选择所选源数量的变化情况，按频段区分（蓝色：未使用滤波器；黄色：阿尔法；绿色：贝塔；红色：伽马）。蓝色虚线表示机会水平。

图7展示了两名典型参与者的准确率随RFE所选源数量（从2到10个）的变化情况，并按频段区分。在伽马频带内，准确率始终较高。准确率变化图表明，通常只需少量源即可获得最高准确率。更具体地说，对于所有参与者，5个源就足以完成分类任务，这与 [32]中得到的结果一致。

图8上排展示了在电极空间应用主成分分析时，准确率随维度的变化情况。在维度为3和4时观察到峰值。相比之下，图8下排比较了在源空间应用主成分分析时，准确率随维度的变化情况。图中显示，准确率的提升（红色）基本上独立于主成分分析所解释的方差（蓝色）；在电极的情况下，这一点更为明显。

3 讨论

本研究聚焦于设计并表征一种利用拓扑数据分析进行数据分类的新方法。与基于评估不同类别点之间相对距离度量的机器学习分类器不同，本文所述的拓扑分类器通过利用各类别点云之间的形状差异进行分类。在实际操作中，我们通过计算持久性轮廓之间的差异来量化此类差异 [46]。

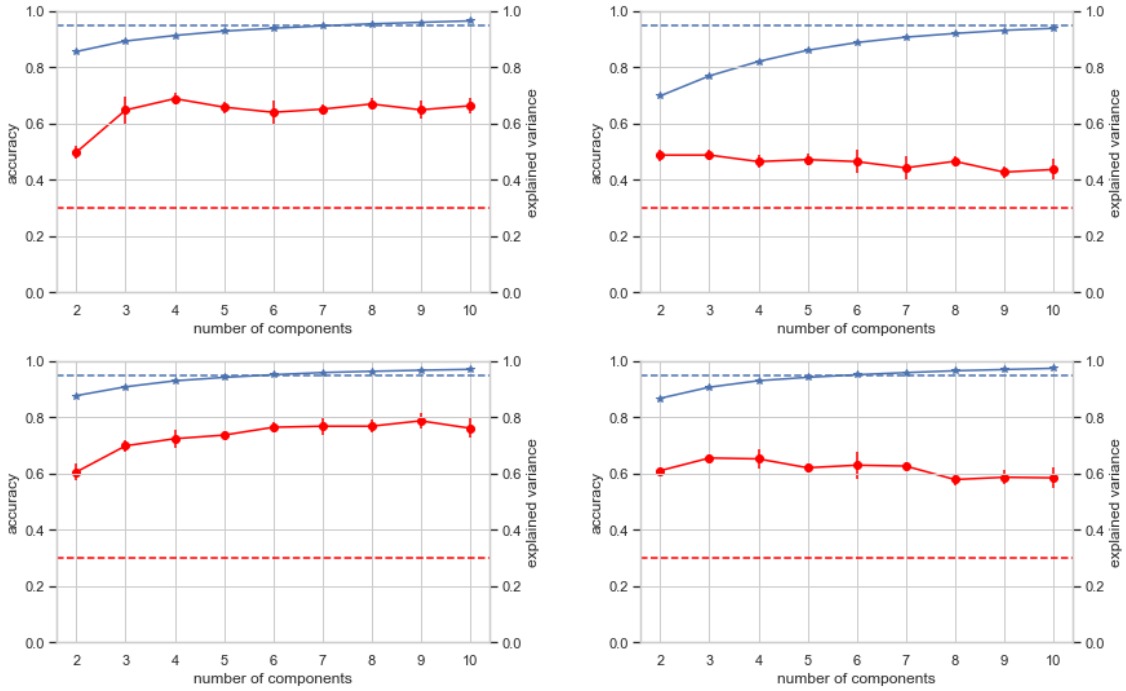


图8：随着维度增加，参与者1（左）和参与者2（右）的准确率变化（红色）与主成分分析解释方差（蓝色）的比较。上排对应电极空间，下排对应源空间。蓝色虚线表示95%的解释方差，红色虚线表示机会水平。准确率（红色）的标准差是在分类器重复五次后计算得到的。

持久条形码的稳定性定理 [46, 35] 保证，点云的微小变化只会使其持久性描述符发生微小变化。因此，如果一个点云被划分为不同类别——在我们的数据集中为三类——那么向该点所属的类别中添加该点，几乎不会影响最终的拓扑描述符；而将同一个点纳入其他类别，则可能引起更明显的形状变化，我们会对这些变化进行量化，并将其用于分类。我们利用脑电图数据集检验了这一原理。这些数据集是在一项针对人类参与者的动机行为研究中记录的 [32], ，该研究旨在量化三个层级的社会压力对人类动机系统神经动力学特征的影响。与大多数脑电图分析一样，我们将数据时间序列分解为三个关注频段：α、β 和 γ （见方法），从而使分析量增加了三倍；此外，我们还使用了未经滤波的数据集信号，以供比较。

数据集的选择

数据集的选择需要满足以下几项要求：

1. 需要拥有丰富的高维数据集，其中包含元素数量可观的数据云。2. 为了对测试结果进行双重验证，数据集此前已通过可靠的方法进行分析，这一点很有帮助，因为这样便能为拓扑数据分析分类器的潜在准确率提供明确的参照。在我们的研究中，数据来自 11 位参与者，他们各自完成了 1296 个试次，并分为三类。3. 先前的机器学习分析 [32] 表明，对于所有参与者，γ 频段的表现始终最佳，准确率介于 75% 至 90% 之间。

维数

我们的研究首先在电极空间中进行，随后在脑源空间中进行；脑源空间是通过独立成分分析（ICA）从电极空间中获得的。从统计学角度看，这种技术很有优势，因为源空间中的各个维度在统计上彼此独立；从神经科学角度看，它也很有用，因为它能够粗略估计负责所记录数据的脑源定位。

我们通过主成分分析（PCA）逐步增加维数，并相应地逐渐提高解释方差，以评估维数对拓扑数据分析分类器性能的影响。PCA分析最多涵盖十个维数，以确保解释超过95%的数据方差。结果令人瞩目：解释方差在十个维数范围内渐近增加，而拓扑数据分析分类器的准确率在达到某一维数后便趋于平稳——在电极空间中，这一维数通常为四。这有力地表明，分类器的运行对数据云的潜在维数比对解释方差的大小更为敏感。与1-近邻分类器的性能对比分析进一步支持了这一结论：随着解释方差增加，该分类器的准确率持续提高，直至维数达到十。

作为对维数测试的补充验证分析，我们还进行了递归特征消除分解（RFE），以确定源空间中对分类贡献最大的信号维度。该分析也得出了与主成分分析类似的上限效应，不过这一效应出现在维数约为五时。

我们最初的预测是，通过 ICA 将原始数据集投影到独立成分空间（或源空间），会形成定义更精细的点云，最终获得高于原始数据集的准确率。我们对 ICA 数据集维数影响准确率的分析支持了这一假设：若事先对数据集应用 ICA，准确率对维数的敏感性会降低。同样，我们使用递归特征消除进行的验证测试也表明，要达到渐近准确率，所需的源数量是有限的（通常为五个）。

最终，能够有效促进拓扑数据分析分类器进行分类的维数，与原研究中足以表示动机脑网络的源数量相符 [32]。此外，效果最佳的频带为 γ 频带，这与先前的结果一致。

解释方差与解释形状

我们基于拓扑数据分析的分类器所获得的准确率与最近邻分类器（1-近邻分类器）相当，也与先前的结果一致 [32]。然而，性能上的一个主要差异是，随着用于分类的维数增加，1-近邻分类器的准确率稳步提高。换言之，分类器训练所用信号的解释方差越高，准确率就越高。尽管如此，拓扑数据分析分类器的准确率通常会在达到某个维数后停滞，这一现象在不同受试者中一致。我们认为，1-近邻分类器具有更高敏感性，是因为它具有局部性，与拓扑摘要的全局视角相对。总之，我们可以推断，拓扑数据分析提供了定量方法，用以评估模型所“解释的形状”而非解释方差。我们的结果表明，环境维数对点云形状的影响与其对数据方差的影响性质不同，因为形状往往会在某个由给定数据本身决定的维数处趋于稳定。

补充信息

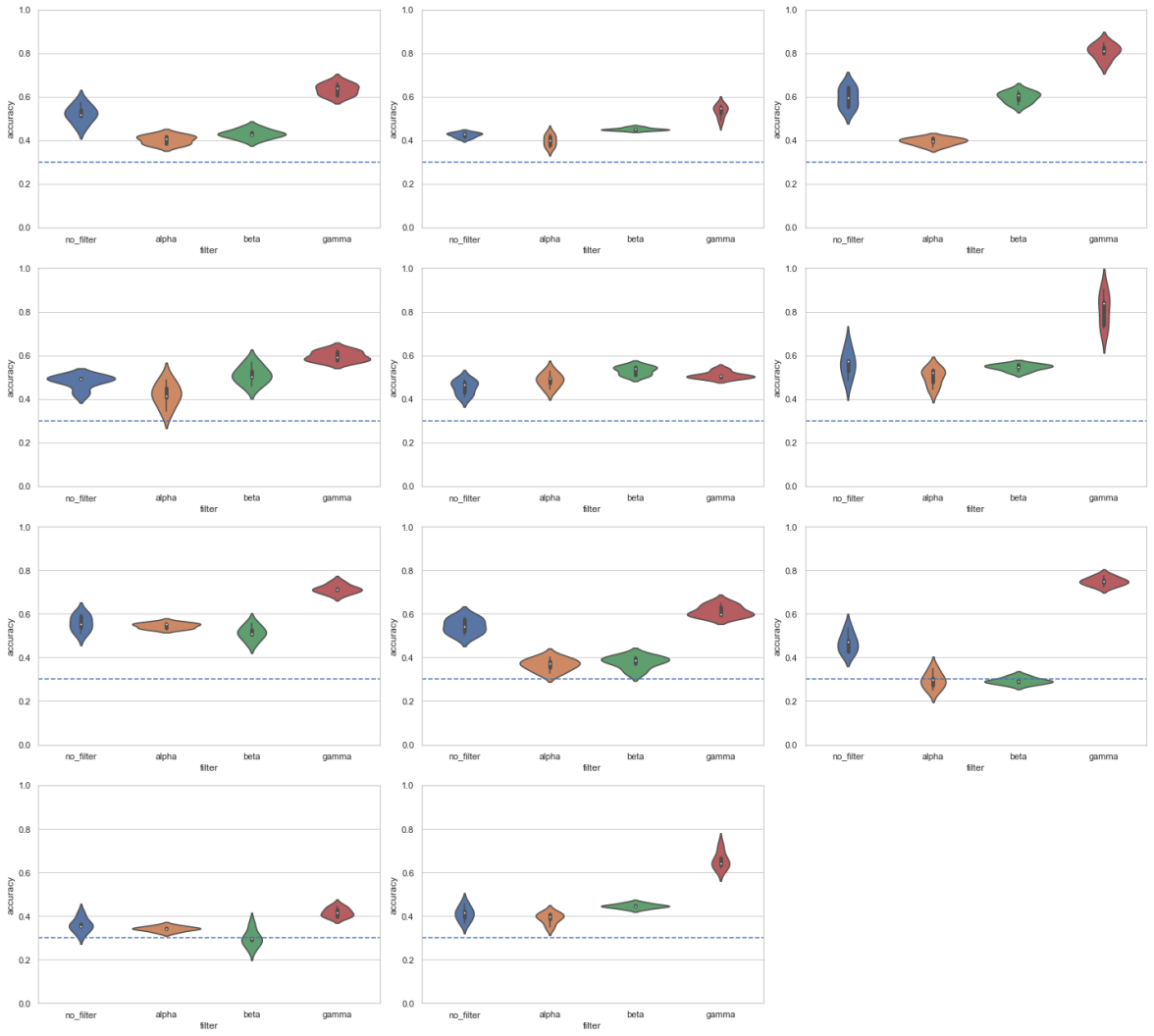


图9：在未进行任何降维的源空间中，参与者1至11的拓扑分类器在各频带（无滤波器、阿尔法、贝塔和伽马）下的准确率

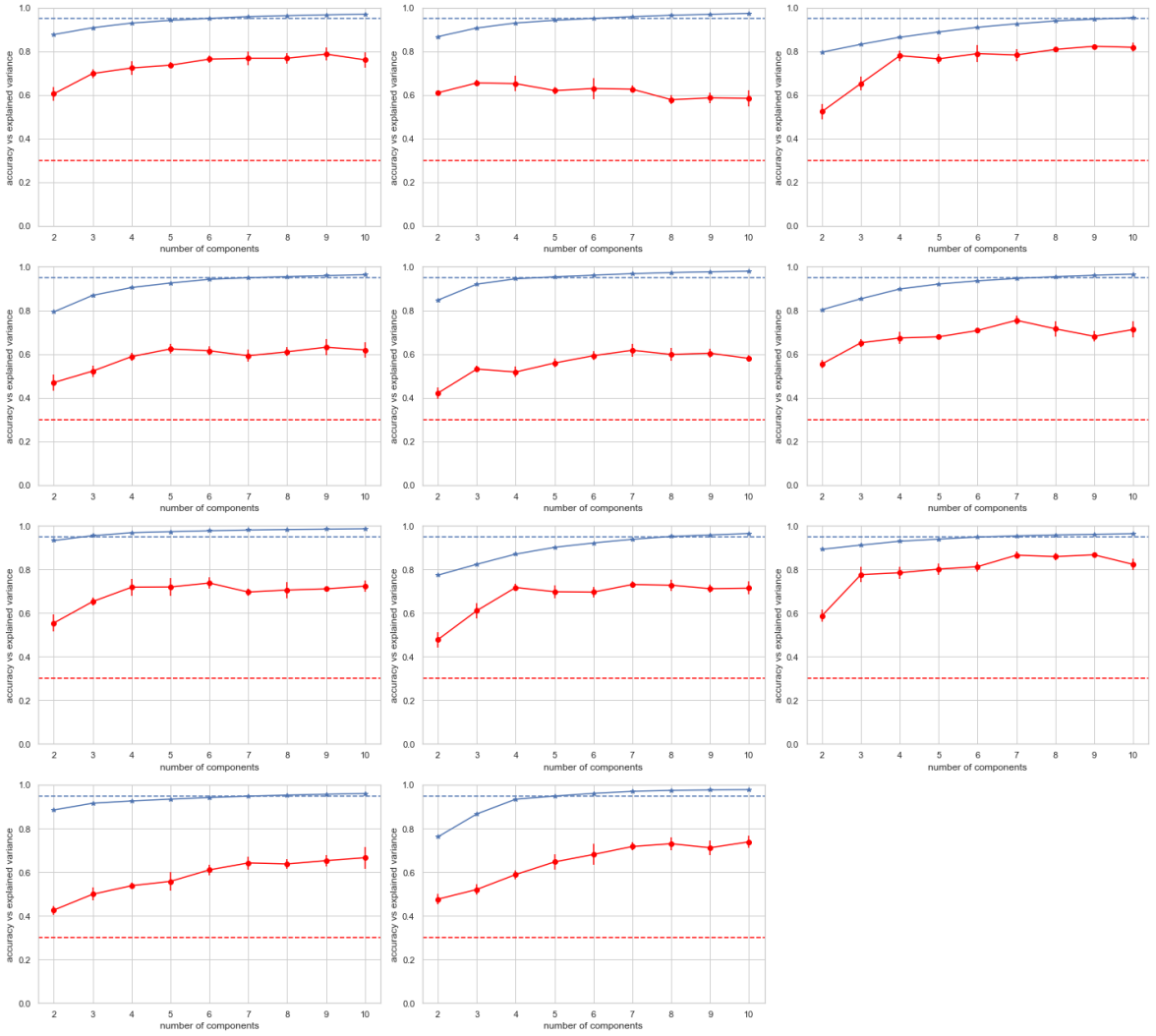


图10：对于伽马频段内的**源空间**，随着维度增加，比较所有参与者的准确率（红色）与**主成分分析**解释方差（蓝色）的变化。蓝色虚线表示95%的解释方差，红色虚线表示机会水平。准确率（红色）的标准差是在分类器重复五次后计算得到的。

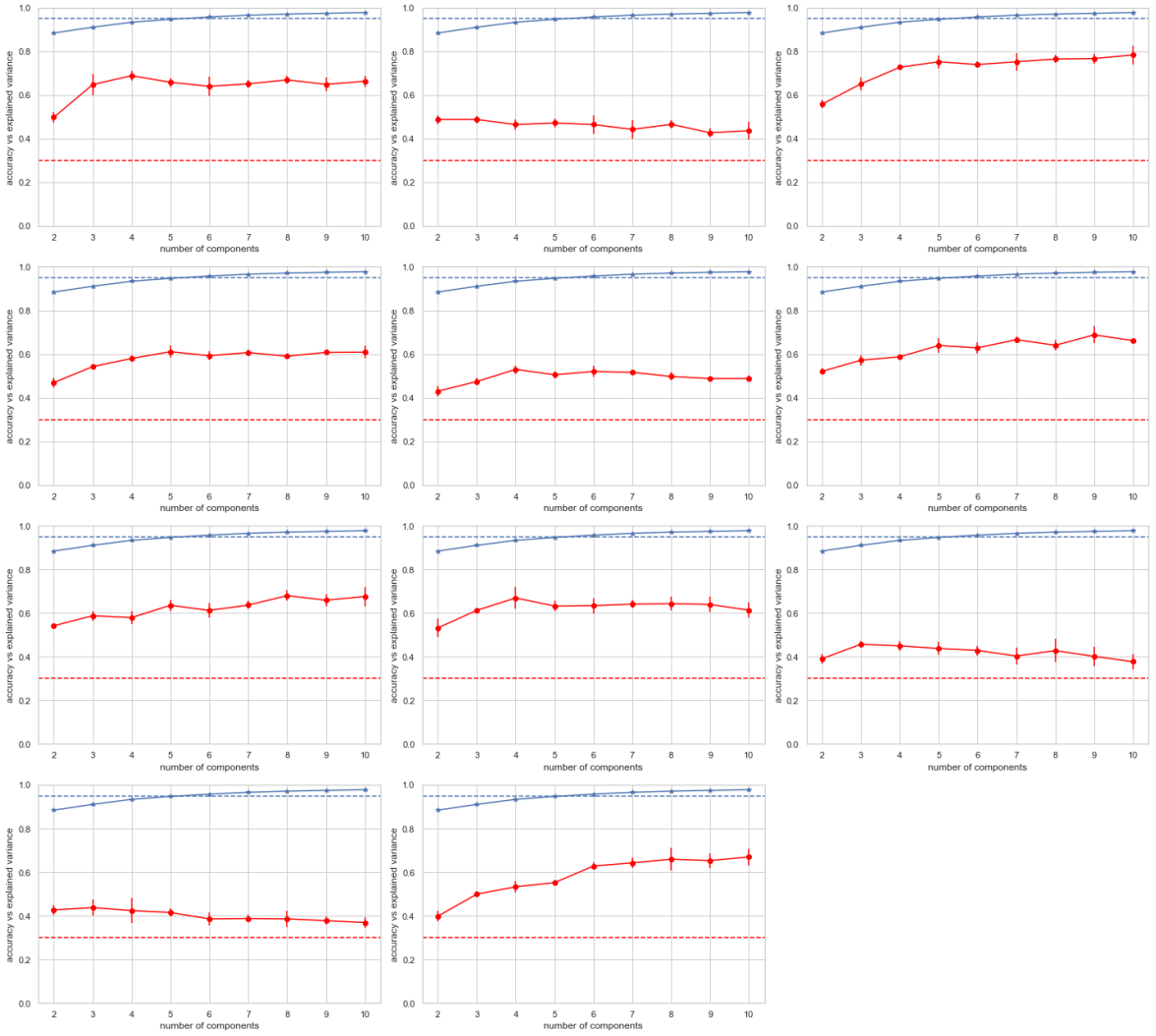


图11：随着维度增加，比较所有参与者在伽马频段电极空间内的准确率（红色）与**主成分分析**解释方差（蓝色）的变化。

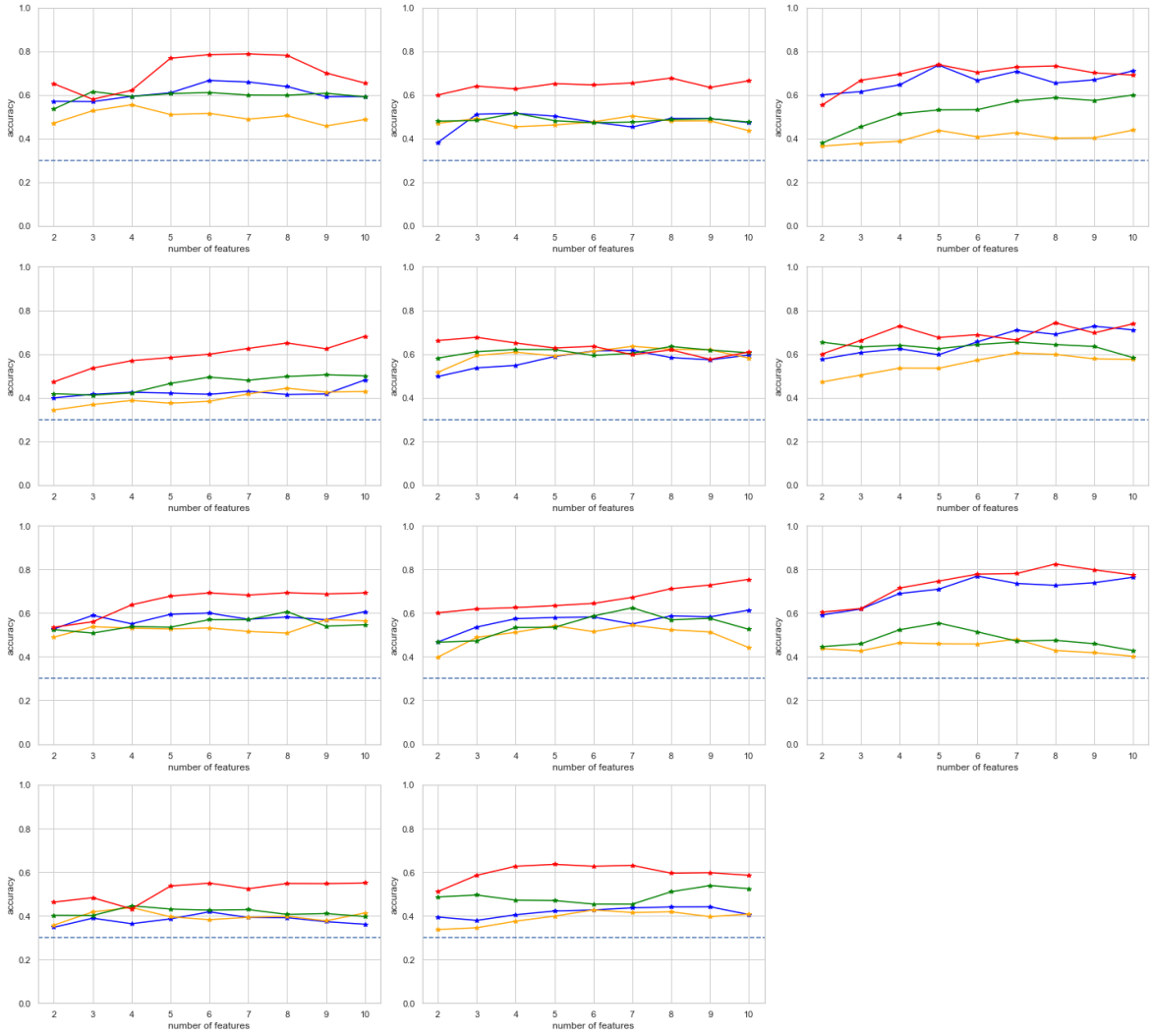


图12：在源空间中，对所有参与者而言，随着通过**递归特征消除**算法选取的源数量从2增加到10，各频带的准确率变化比较（蓝色：无滤波器；黄色：阿尔法；绿色：贝塔；红色：伽马）。

	no filter			alpha			beta			gamma		
	raw	pca	rfe	raw	pca	rfe	raw	pca	rfe	raw	pca	rfe
1	0.57	0.56	0.69	0.42	0.60	0.58	0.46	0.64	0.65	0.67	0.79	**0.83**
2	0.43	0.47	0.54	0.41	0.52	0.58	0.46	0.56	0.54	0.56	0.66	**0.71**
3	0.65	0.56	0.77	0.49	0.48	0.50	0.63	0.63	0.65	**0.85**	0.82	0.76
4	0.50	0.48	0.50	0.41	0.52	0.48	0.57	0.56	0.53	0.62	0.63	**0.70**
5	0.49	0.58	0.65	0.53	0.67	0.68	0.55	0.66	0.68	0.54	0.62	**0.70**
6	0.65	0.63	0.77	0.53	0.64	0.66	0.56	0.74	0.70	**0.90**	0.75	0.78
7	0.60	0.57	0.65	0.56	0.54	0.59	0.56	0.55	0.63	**0.75**	0.74	**0.75**
8	0.58	0.54	0.66	0.40	0.53	0.59	0.40	0.67	0.64	0.65	0.73	**0.79**
9	0.51	0.64	0.79	0.35	0.47	0.53	0.32	0.49	0.60	0.78	**0.87**	0.85
10	0.41	0.42	0.44	0.36	0.51	0.49	0.36	0.51	0.48	0.45	**0.67**	0.59
11	0.46	0.43	0.50	0.41	0.42	0.47	0.46	0.53	0.59	0.72	**0.74**	0.66

图13：比较参与者1至11在各频带下，拓扑分类器在源空间中的基线准确率（原始）与经过主成分分析（主成分分析）和递归特征消除（递归特征消除）降维后获得的准确率。最高准确率以粗体标示。

无滤波器			阿尔法		
raw	pca	rfe	raw	pca	rfe
0.57 ± 0.04	0.56 ± 0.06	0.69 ± 0.04	0.42 ± 0.02	0.60 ± 0.01	0.58 ± 0.04
0.43 ± 0.01	0.47 ± 0.02	0.54 ± 0.05	0.43 ± 0.03	0.52 ± 0.03	0.58 ± 0.04
0.65 ± 0.05	0.56 ± 0.03	0.77 ± 0.06	0.41 ± 0.02	0.48 ± 0.02	0.50 ± 0.04
0.50 ± 0.03	0.48 ± 0.04	0.50 ± 0.03	0.49 ± 0.05	0.52 ± 0.02	0.48 ± 0.04
0.49 ± 0.03	0.56 ± 0.03	0.65 ± 0.04	0.53 ± 0.03	0.67 ± 0.03	0.68 ± 0.04
0.65 ± 0.06	0.63 ± 0.05	0.77 ± 0.06	0.53 ± 0.04	0.64 ± 0.02	0.66 ± 0.05
0.60 ± 0.04	0.57 ± 0.06	0.65 ± 0.04	0.56 ± 0.01	0.54 ± 0.02	0.59 ± 0.03
0.58 ± 0.03	0.54 ± 0.03	0.66 ± 0.05	0.40 ± 0.03	0.53 ± 0.02	0.59 ± 0.06
0.54 ± 0.05	0.64 ± 0.04	0.79 ± 0.06	0.35 ± 0.04	0.47 ± 0.03	0.53 ± 0.03
0.41 ± 0.03	0.42 ± 0.06	0.44 ± 0.03	0.36 ± 0.01	0.51 ± 0.03	0.49 ± 0.03
0.46 ± 0.03	0.43 ± 0.01	0.50 ± 0.04	0.41 ± 0.03	0.42 ± 0.02	0.47 ± 0.04

表1：比较患者1至11在无频率滤波器和阿尔法频带下，拓扑分类器在源空间中针对各频率滤波器的基线准确率（原始）与经过主成分分析（主成分分析）和递归特征消除（递归特征消除）降维后获得的准确率，并列出重复5次后的标准差。

beta			伽马		
raw	pca	rfe	raw	pca	rfe
0.46 ± 0.02	0.64 ± 0.03	0.65 ± 0.03	0.67 ± 0.03	0.79 ± 0.03	0.83 ± 0.08
0.46 ± 0.01	0.56 ± 0.03	0.54 ± 0.03	0.56 ± 0.03	0.66 ± 0.01	0.71 ± 0.04
0.63 ± 0.03	0.63 ± 0.02	0.65 ± 0.07	0.85 ± 0.03	0.82 ± 0.01	0.76 ± 0.06
0.57 ± 0.04	0.56 ± 0.02	0.53 ± 0.04	0.62 ± 0.02	0.63 ± 0.04	0.70 ± 0.07
0.55 ± 0.02	0.66 ± 0.04	0.68 ± 0.02	0.54 ± 0.02	0.62 ± 0.03	0.70 ± 0.04
0.56 ± 0.01	0.74 ± 0.02	0.70 ± 0.03	0.90 ± 0.08	0.75 ± 0.02	0.78 ± 0.05
0.56 ± 0.03	0.55 ± 0.01	0.63 ± 0.04	0.75 ± 0.02	0.74 ± 0.02	0.75 ± 0.06
0.40 ± 0.03	0.67 ± 0.02	0.64 ± 0.05	0.65 ± 0.03	0.73 ± 0.02	0.79 ± 0.06
0.32 ± 0.01	0.49 ± 0.03	0.60 ± 0.05	0.78 ± 0.02	0.87 ± 0.01	0.85 ± 0.08
0.36 ± 0.04	0.51 ± 0.03	0.48 ± 0.03	0.45 ± 0.02	0.67 ± 0.05	0.59 ± 0.05
0.46 ± 0.01	0.53 ± 0.02	0.59 ± 0.04	0.72 ± 0.04	0.74 ± 0.03	0.66 ± 0.04

表 2：对于患者 1 至 11，在贝塔和伽马频段中，拓扑分类器在源空间上针对每个频率滤波器的基线准确率（原始）与采用主成分分析（主成分分析）和递归特征消除（递归特征消除）进行降维后获得的准确率的比较；结果为重复五次后的标准差。

参考文献

[1] Grossberg S. 通往可解释人工智能与自主自适应智能之路：深度学习、自适应共振，以及感知、情绪和行动模型。《神经机器人学前沿》2020；14。[2] Sumathi A, Yasotha K, Nandhinidevi S. 面向演化概念的高维深度数据聚类架构。NVEO - Natural Volatiles & Essential Oils Journal 2021；1695–1703。[3] Yazdani A, Lu L, Raissi M, Karniadakis GE. 利用系统生物学信息的深度学习推断参数与隐藏动力学。《PLOS计算生物学》2020；16:e1007575。[4] Moen E, Bannon D, Kudo T, Graf W, Covert M, Van Valen D. 用于细胞图像分析的深度学习。《自然·方法》2019；16:1233–1246。[5] Wang L, Wang H, Huang Y, Yan B, Chang Z, Liu Z, Zhao M, Cui L, Song J, Li F. 深度学习网络在医学图像分析中的应用趋势：2012至2020年的演进。《欧洲放射学杂志》2022；146:110069。[6] Pathak AR, Pandey M, Rautaray S. 深度学习在目标检测中的应用。《普罗赛迪亚计算机科学》132: 1706–1717, 2018。[7] Gibson PB, Chapman WE, Altinok A, Delle Monache L, DeFlorio MJ, Waliser DE. 在气候模型输出上训练机器学习模型，可产生技巧高超且可解释的季节性降水预报。《通讯·地球与环境》2021；2:1–13。[8] Rasp S, Pritchard MS, Gentine P. 利用深度学习表示气候模型中的亚网格过程。《美国国家科学院院刊》2018；115:9684–9689。[9] Zou J, Huss M, Abid A, Mohammadi P, Torkamani A, Telenti A. 基因组学深度学习入门。《自然·遗传学》2019；51:12–18。[10] Ghassemi M, Oakden-Rayner L, Beam AL. 当前医疗保健领域可解释人工智能方法所带来的虚假希望。《柳叶刀·数字健康》2021；3:e745–e750。[11] Petch J, Di S, Nelson W. 打开黑箱：可解释机器学习在心脏病学中的前景与局限。《加拿大心脏病学杂志》2021。

[12] Razavi S. 深度学习解析：基础、可解释性及与过程模型的衔接。《Environmental Modelling & Software》2021；144:105159。[13] Heaven D. 为什么深度学习人工智能如此容易受骗。《自然》2019；574:163–166。[14] LeCun Y, Bengio Y, Hinton G. 深度学习。《自然》2015年5月；521:436–444。[15] Nguyen A, Yosinski J, Clune J. 深度神经网络很容易受骗：对不可识别图像给出高置信度预测。2015年IEEE计算机视觉与模式识别会议（CVPR），2015；427–436。[16] Belchi F, Pirashvili M, Conway J, Bennett M, Djukanovic R, Brodzki, J. 慢性阻塞性肺疾病患者的肺部拓扑特征。《科学报告》2018；8(1):1–12。[17] Sizemore AE, Phillips-Cremins JE, Ghrist R, Baset DS. 整体的重要性：面向网络神经科学家的拓扑数据分析。《Network Neuroscience》2018；3(3):656–673。[18] Lawson P, Sholl AB, Brown, JQ, Fasy BT, Wenk C. 利用持久同调定量评估前列腺癌组织学中的结构特征。《科学报告》2019；9(1):1–15。[19] Xia K, Zhao Z, Wei G-W. 面向超大型生物分子数据集的多分辨率持久同调。《化学物理学杂志》2015；143(13): 10B603 1。[20] Hiraoka Y, Nakamura T, Hirata A, Escolar EG, Matsue K, Nishiura Y. 利用持久同调表征非晶固体的层级结构。《美国国家科学院院刊》2016；113(26):7035–7040。[21]Gebhart T, Schrater P. 通过持久同调检测神经网络中的对抗样本。arXiv:1711.10056, 2017。[22] Pearson K. LIII. 关于空间点系中最佳拟合的直线与平面。《伦敦、爱丁堡和都柏林哲学杂志与科学杂志》1901；2(11):559–572。[23] Comon, P. 独立成分分析。《高阶统计》1992；29–38。[24] Carri`ere M, Blumberg A. 用于拓扑机器学习的多参数持久图像。神经信息处理系统大会，2020。[25] Hofer C, Kwitt R, Niethammer M, Uhl A. 利用拓扑特征进行深度学习。arXiv: 1707.04041, 2017。[26] Yan Y, Ivanov K, Cen J, Liu Q-H, Wang, L. 基于持久性景观的拓扑数据分析，用于个性化心律失常分类。Preprints 2019, 2019080320。[27] Kindelan R, Fr´ıas J, Cerda M, Hitschfeld N. 基于拓扑数据分析的分类。arXiv:2102.03709, 2021。[28] In´es A, Dom´ınguez C, Heras J, Mata G, Rubio J. 半监督机器学习：一种同调方法。arXiv:2301.11658, 2023。[29] Allen EA, Damaraju E, Eichele T, Wu L, Calhoun VD. 动态功能网络连接状态的脑电图特征。《脑地形学》2017；1–16。[30] Cattai T, Colonnese S, Corsi M-C, Bassett DS, Scarano G, De Vico Fallani F. 通过脑信号之间的节点连接表征心理状态。第26届欧洲信号处理会议（EUSIPCO）：1377–1381。[31] Gilson M, Zamora-L´opez G, Pallar´es V, Adhikari MH, Senden M, Campo AT, Mantini D, Cor- betta M, Deco G, Insabato A. 基于模型的全脑有效连接，用于研究健康与疾病状态下的分布式认知。《Network Neuroscience》2019；4:338–373。

[32] Cos I、Deco G、Gilson M。决策过程中精确伸手动作的社会压力行为与神经相关性。DOI:10.21203/rs.3.rs-1974463/v1[33] Edelsbrunner H、Letscher D、Zomorodian A。拓扑持久性与简化。《离散与计算几何》2002；28：511–533。[34] Zomorodian A、Carlsson G。持久同调的计算。《离散与计算几何》2005；33(2)：249–274。[35] Edelsbrunner H、Harer J。持久同调——综述。《离散与计算几何综述：二十年后》。《当代数学》2008；453：257–282；普罗维登斯：美国数学会。[36] Carlsson G。拓扑与数据。《美国数学会通报》（新系列）2009；46(2)：255–308。[37] De Silva V、Ghrist R。通过持久同调实现传感器网络覆盖。《代数与几何拓扑》2007；7(1)：339–358。[38] Ghrist R。条形码：数据的持久拓扑。《美国数学会通报》2008；45(1)：61–75。[39] Hatcher A。《代数拓扑》。剑桥：剑桥大学出版社，2002。[40] GUDHI 项目。《GUDHI 用户与参考手册》，第3.4.1版。GUDHI 编辑委员会，2021。[41] Cohen-Steiner D、Edelsbrunner H、Harer J。持久图的稳定性。《离散与计算几何》2007；37(1)：103–120。[42] Chazal F、De Silva V、Oudot S。几何复形的持久性稳定性。《Geometriae Dedicata》2014；173：193–214。[43] Mileyko Y、Mukherhee S、Harer J。持久图空间上的概率测度。《反问题》2011年11月；27(12)：124007。[44] Bubenik, P。使用持久性景观进行统计拓扑数据分析。《机器学习研究杂志》2015；16(1)：77–102。[45] Bubenik P、D-lotko P。用于拓扑统计的持久性景观工具箱。《符号计算杂志》2017；78：91–114。[46] Chazal F、Fasy BT、Lecci F、Rinaldo A、Wasserman L。持久性景观与轮廓的随机收敛。第30届计算几何年度研讨会论文集，2014：474–483。[47] Virtanen P 等。SciPy 1.0：用于 Python 科学计算的基础算法。《自然·方法》2020；17：261–272。[48] Pedregosa F 等。Scikit-learn：Python 机器学习。《机器学习研究杂志》2011；12：2825–2830。[49] Guyon I、Weston J、Barnhill S、Vapnik V。使用支持向量机进行癌症分类的基因选择。《机器学习》2002；46(1-3)：389—422。[50] Waskom ML。Seaborn：统计数据可视化。《开放源代码软件杂志》2021；6(60)：3021。