arXiv:2305.01807v2 [cs.LG] 4 May 2023

协方差神经网络的可迁移性及其在使用解剖特征进行可解释脑龄预测中的应用

Saurabh Sihag[1], Gonzalo Mateos[2], Corey T. McMillan[1],和 Alejandro Ribeiro[1]

[1]宾夕法尼亚大学，费城，宾夕法尼亚州。[2]
罗切斯特大学，罗切斯特，纽约州。

摘要

图卷积网络（GCN）利用由拓扑结构驱动的图卷积运算，整合图中的信息以执行推理任务。在我们近期的工作中，我们研究了以协方差矩阵为图的GCN，即协方差神经网络（VNN）；这类网络与传统的主成分分析驱动的数据分析方法有相似之处，同时又具备显著优势。本文首先聚焦于从理论上刻画VNN的可迁移性。可迁移性的概念源于一种直观预期：学习模型经过最少的调整，便能泛化到“兼容”的数据集（其维度可能不同）。VNN继承了GCN的无尺度数据处理架构；本文表明，对于协方差矩阵收敛到某一极限对象的数据集，VNN能够实现性能可迁移性（无需重新训练）。多尺度神经影像数据集能够支持在多个尺度上研究脑，因此为验证VNN可迁移性的理论结果提供了理想场景。为了评估VNN在神经影像数据分析中的优势，我们聚焦于利用皮层厚度特征进行“脑龄”预测。在临床神经科学领域，人们对基于磁共振成像皮层厚度特征的机器学习算法日益关注；这类算法能够估计“脑龄”，且估计结果与实际年龄存在偏差。重要的是，脑龄与实际年龄之间的不一致（“脑龄差”）可能反映个体对神经系统疾病或认知障碍的易感性或韧性增强。我们利用VNN的架构，将研究从阿尔茨海默病（AD）中的粗略脑龄差指标拓展至更深入的分析，并得出两项重要发现：（i）VNN能够通过识别相关脑区，为AD患者脑龄差升高提供解剖学解释；（ii）VNN所提供的可解释性取决于其利用解剖协方差矩阵特定主成分的能力。我们进一步利用VNN的可迁移性，在维度不同的数据集间交叉验证了上述发现。

审稿中。本稿部分内容已发表于 [1] 和 [2]。有关数据与代码可用性的信息，请参阅第7节。

1 引言

在各种现代应用中，数据集的特征数量（记为 m）是数据采集的基本组成部分，通常由目标应用及相关后勤因素决定 [3,4]。大多数机器学习算法和统计推断方法都是针对预先定义的特征集设计的，因此其计算复杂度和样本复杂度本质上取决于维度 m [5,6]。本文研究一种名为协方差神经网络（VNN）的深度学习框架 [1]，它基于在给定数据集的样本协方差矩阵上运行的图神经网络，并且与尺度无关，即VNN中的可学习参数数量不受数据集维度 m 影响（图 1）。VNN的尺度无关特性使其能够在维度不同的数据集之间实现可迁移性，而无需改变其架构；也就是说，在维度为 $m = m_1$ 的数据集上训练的VNN，可以使用同一组已学习参数处理维度为 $m = m_2$ 的另一数据集。

虽然数据集中包含更多特征 m 可能意味着所收集数据具有更高的分辨率或质量，但特征过多也可能给存储、计算复杂度以及统计模型的可解释性带来挑战，从而影响有效的数据分析。另一方面，特征过少的数据集可能缺乏足够的相关信息，难以保证准确性和推理能力。尽管如此，如果两组特征描述的是不同尺度或分辨率下的相似现象，那么直观上可以预期，由 $m = m_1$ 个特征组成的数据集与由$m = m_2$ 个特征组成的另一个数据集之间存在某种对应关系。当推理任务所需的可学习参数数量由数据集的维度 m 决定时，传统的数据分析方法（例如主成分分析）和机器学习算法无法利用或适应数据集之间这一相似性。因此，一旦数据维度发生变化，就需要从头重新设计此类统计模型。在此背景下，我们给出了关于协方差矩阵的理论条件，在这些条件下，针对一个数据集上的推理任务所取得的 VNN 模型性能，可以迁移到维度不同的另一个数据集上，而无需重新训练或更改 VNN 模型（图3）。除了相较于传统数据分析方法所具有的方法学优势外，VNN 框架在管理计算复杂度方面也具有优势。事实上，在适当的条件下，基于特征数量较少的数据集训练的模型可以直接应用于从特征数量较多的数据集中进行推理。

神经影像学是一个现代应用的例子，在该应用中，不同数据集的特征数量高度可变，但不同数据集包含相似的信息 [7, 8]。具体而言，MRI数据可以用多种尺度表示，从单个体素（通常为 $\sim 1\text{mm}^3$）到源自多尺度脑图谱的感兴趣区域（感兴趣区域），后者涵盖从数十个到数千个不等的脑区划分（例如，在多尺度脑图谱中，脑区划分数量可从100到1000不等 [9,10]）。多尺度脑图谱提供了将大脑皮层划分为不同数量脑区划分的映射，因此，相关神经影像数据集以不同尺度描述大脑皮层上的相似信息。因此，我们通过实验证实

了根据一个常用脑图谱的不同尺度整理的一组神经影像数据集上VNN可迁移性的理论结果。

VNN 框架的简洁性使我们能够分析每个特征对 VNN 模型最终统计结果的贡献。对于使用神经影像数据集的应用而言，这一点意义重大，因为其中的特征与不同的脑区相关。因此，在使用包含解剖特征的数据集进行训练时，VNN 使我们能够为学习结果赋予解剖学或区域层面的 可解释性 。受这些观察结果的启发，我们将基于磁共振成像（MRI）获得的皮层厚度特征进行“脑龄”预测，作为 VNN 的一项应用（图 4）。重要的是，个体可能以不同的速率经历与年龄相关的影响，这种现象被称为“生物学衰老”。因此，加速衰老（例如，生物学年龄高于实际年龄时）可能预示与年龄相关的脆弱性，例如认知衰退风险，或阿尔茨海默病及相关痴呆症（ADRD）等神经系统疾病 [11]。在这一领域，所关注的指标是 脑龄差，即生物学年龄与实际年龄之差。我们使用记号Δ-Age 表示脑龄差。由于脑龄没有真实值， Δ-Age 本质上是一项定性指标；与健康人群相比，预计潜在患有神经退行性疾病的个体该指标会更高。大量基于广泛机器学习方法的现有研究报告称，在包括阿尔茨海默病 [12] 和精神分裂症 [13]在内的神经退行性疾病中， Δ-Age 会升高。然而，使用机器学习进行脑龄评估的方法也受到了一些批评。主要批评包括： Δ-Age 过于粗略，导致无法明确哪些脑区促成了脑龄升高；以及这些机器学习模型的设计对实际年龄预测准确率存在缺乏解释的依赖 [14]。VNN 的架构使我们能够提出一种有原则的脑龄预测框架，通过分离出促成神经退行性变中 Δ-Age 升高的脑区来实现可解释性。通过分析神经退行性疾病组与健康对照组之间的组间差异，可以识别出这些相关脑区，具体而言，是比较不同特征对 VNN 所作相应预测的贡献（图 4 ）。附录 A 收录了面向普通读者的概述，介绍 VNN 在为脑龄预测增添可解释性方面所提供的优势。

VNN 的可解释性与可迁移性共同为脑龄预测框架中训练模型以预测实际年龄的作用提供了新见解。现有大量利用深度学习模型研究脑龄的文献，将模型准确预测健康对照组实际年龄（出生后经过的时间）的能力 [15–17] 视为评估其方法质量的一项相关指标。与此同时，在拟合健康对照组实际年龄方面相对适中的深度学习模型，也可能比拟合得更紧密的模型更能帮助深入理解脑龄 [16]。因此，对于在预测有意义的脑龄时，健康对照组实际年龄的拟合质量究竟发挥什么作用，目前在概念上尚不明晰 [18], ，而且

如果缺乏透明度，任何用于评估脑龄的统计方法都不太可能解决这一问题（附录 A）。

可以合理地推测，在健康对照组中准确或近乎完美地预测实际年龄，并不一定能使深度学习模型适应神经退行性变所引发的脑部变化。对 VNN 可解释性与可迁移性方面的研究表明，将 VNN 从一个数据集迁移到另一个数据集后，仍可能定性地识别出神经退行性变中导致 Δ-Age 升高的脑区。此外，我们还发现，VNN 利用解剖协方差矩阵中特定特征向量的能力，是 VNN 在 Δ-Age 情境下实现解剖学可解释性的根本因素。因此，VNN 有助于有原则地解耦推断脑龄这一目标与在健康对照组中实现实际年龄预测近乎完美表现这一目标。

贡献： 本文的贡献总结如下：

- **VNN 的可迁移性：** 我们从理论上刻画了 VNN 在不同维度的数据集之间的可迁移性。对于一个具有 m_1 个特征且协方差矩阵为 $\mathbf{C}^{m_1}$ 的数据集，以及另一个具有 m_2 个特征且协方差矩阵为 $\mathbf{C}^{m_2}$ 的数据集，我们证明，在协方差矩阵 $\mathbf{C}^{m_1}$ 和 $\mathbf{C}^{m_2}$ 满足适当条件时，分别以 $\mathbf{C}^{m_1}$ 和 $\mathbf{C}^{m_2}$ 初始化的 VNN 输出在某种意义上是接近的（见定理3）。我们在一项回归任务中验证了关于 VNN 可迁移性的理论结果；该任务基于一组皮层厚度数据集，这些数据集按照常用多尺度脑图谱的不同尺度整理而成（图 5 和表 1 和表 2）。

- **使用 VNN 进行脑龄预测：**我们将 VNN 应用于脑龄预测任务，并比较了健康对照组与确诊阿尔茨海默病的个体之间的 Δ-脑龄。以下总结了这组实验得出的见解：

  a) **VNN 可提供具有解剖学可解释性的 Δ-脑龄，以反映神经退行性变：**确诊阿尔茨海默病的个体，其 Δ-脑龄高于健康对照组。VNN架构简单，因此我们能够刻画导致 Δ-脑龄升高的区域贡献因素，从而赋予 Δ-脑龄以解剖学可解释性（图 7）。在多尺度数据集上，VNN 的可迁移性还有助于交叉验证神经退行性变中 Δ-脑龄所观察到区域特征的空间稳健性（附录K中的图 20）。

  b) **解剖学可解释性与解剖协方差矩阵的特征向量相关：**我们的实验表明，解剖协方差矩阵的特定特征向量与促进 Δ-Age 解剖学可解释性的特征之间存在相关性（图 8）。因此， Δ-Age 与 VNN 利用解剖协方差矩阵中特定特征向量的能力相关。

c) **明确了训练 VNN 以预测实际年龄的作用：**我们的实验确凿地表明，训练 VNN 预测实际年龄，使其能够利用与神经退行性变中 Δ-年龄升高相关的解剖协方差矩阵的特征向量（图 9、图 10 和图 19）。在此背景下所观察到的现象，有助于将脑龄预测任务与在健康对照组中实现近乎完美的实际年龄预测性能这一目标区分开来。

接下来，我们将回顾与本文贡献相关的文献。

相关文献。 图神经网络（GNNs）是卷积神经网络针对图结构数据的一种广受欢迎的改编形式 [19–21]。图是复杂、空间分布现象的自然描述符，因此，图结构数据集广泛存在于各种应用领域 [22],, 包括物理基础设施 [23], 、社交网络分析 [24], 、生物学 [25], 、Network Neuroscience [26]和自然科学 [27]。图结构数据的处理面临各种实际挑战（如泛化能力、可复现性和可扩展性） [22], 因此，人们提出了多种 GNN 变体来应对这些挑战。我们建议读者参阅近期的综述文章 [21] 和 [28] ，这些文章依据多种标准对 GNN 架构进行分类，包括数学公式、算法以及硬件／软件实现。卷积 GNN [29], 、图自编码器 [30], 、循环 GNN 和门控 GNN [31] 是目前研究和应用较为广泛的几类 GNN。

与本文相关的分类体系是图卷积网络 [29]。GCN 通常依赖于在图结构上进行信息聚合（称为图卷积）的过程来处理数据。文献中已提出多种图卷积运算的实现策略，包括谱卷积 [32], 切比雪夫多项式 [33], 普通多项式 [34],以及基于扩散的表示 [35]。在各种设置下，GCN 具备对拓扑扰动的稳定性，以及在不同规模图之间的可迁移性 [36–39], ，这使其成为处理图结构数据的合理数据分析工具。

在我们最近的工作 [1], 中，我们研究了协方差神经网络（VNN），这是一类以样本协方差矩阵作为图、以多项式图滤波器作为卷积运算的 GCN。协方差矩阵和主成分分析（PCA）构成了现实世界应用中非参数分析的两大基石；这些应用具有空间分布式的多变量数据采集协议，包括神经影像学 [40], 、计算机视觉 [41, 42], 、天气建模 [43], 、交通流分析 [44], 以及云计算 [45]。我们在 [1] 中的结果确立了以下重要观察：i）协方差矩阵上图卷积的谱分析与标准 PCA 变换之间存在相似性；ii）VNN 对用于估计样本协方差矩阵的样本数量具有鲁棒性，从而克服了 PCA 统计推断中潜在的不稳定性和不可复现性来源 [46,47]。

图神经网络从训练图迁移到某些兼容的测试图族的可迁移性，先前已从不同的理论视角进行研究 [39, 48–50]。图神经网络可迁移性的概念

总体而言，该概念体现了这样一种直觉：当图神经网络应用于描述与训练图相同现象的测试图时（无论其规模如何），可能仍能在某些推理任务中保持性能。在此背景下， [48] 中的研究考察了图神经网络在表示底层拓扑空间离散化的图上的可迁移性。另有若干研究考察了图神经网络在收敛序列中的图上的可迁移性；随着节点数量趋于无穷，该序列会逼近一个极限对象 [39, 50]。在 [49], 中，基于图拓扑构建的自我图分布之间的相似性，是评估图神经网络可迁移性的主要依据。图神经网络的可迁移性还具有计算复杂度方面的优势：对于含有 m 个节点的稠密图，其计算复杂度按 $\mathcal{O}(m^2)$ 的规模增长。本文将可迁移性的概念扩展至 VNN，并确立了不同尺寸的协方差矩阵之间的迁移性，这些矩阵以某种意义收敛。与此相对，对于传统的、由主成分分析驱动的统计模型，可迁移性并不可行，因为主成分受限于原始数据集的特征空间；一旦特征数量发生变化，就需要重新评估。具体而言，主成分分析无法衡量从大小为 $m_1 \times m_1$ 的协方差矩阵中提取的主成分与从另一个大小为 $m_2 \times m_2$的协方差矩阵中提取的主成分之间的相似性。因此，VNN的可迁移性与多元统计领域密切相关。

神经科学领域对多尺度数据集的兴趣日益浓厚 [8, 51–54]。这类数据集依托脑图谱或模板，可对大脑表面进行多尺度分区（例如 Schae- fer 图谱 [9] 和 Lausanne 图谱 [10]）。多尺度脑图谱会在不同尺度下将大脑皮层划分为数量不等的区域。然而，目前仍缺乏能够以统计学上可靠的方式，充分利用包含多尺度特征的数据集中的冗余信息的方法。本文利用依据多尺度脑图谱整理的皮层厚度数据集，验证 VNN 的可迁移性。在依据不同脑图谱整理的数据集之间不满足 VNN 可迁移性的理论保证时，可能需要采用适当的映射来考虑不同脑图谱之间的差异 [7, 8] ，从而在将 VNN 从一个数据集迁移到另一个数据集时，对 VNN 输出进行缩放，以便在实际年龄预测任务中取得良好的定量表现。然而，这些映射不太可能包含有关神经退行性变或脑龄的信息，因此本文不予研究。

人类衰老过程的特征是大脑在解剖结构和功能方面发生渐进性变化 [55]。病理性《脑》-年龄可视为对大脑中由病理驱动的纵向非典型变化的标量表征 [12]。包括结构性磁共振成像（MRI）、功能性磁共振成像和正电子发射断层扫描在内的神经影像学模态数据，能够捕捉神经退行性变和健康衰老导致的大脑变化 [56, 57]。因此，利用不同神经影像学模态推断脑龄一直是一个活跃的研究领域 [58–63]。本文利用 OASIS-3 数据集中从结构性MRI图像得出的皮层厚度测量值数据集，研究脑龄 [64]。皮层厚度会随正常衰老而变化 [65] ，并受阿尔茨海默病（AD）中的神经退行性变影响 [66, 67]。

因此，从人群中皮层厚度测量值之间的相关性计算得出的解剖协方差矩阵，也呈现出与年龄和疾病严重程度相关的变化 [40]。

由于 VNN 能够对其统计结果提供内在的可解释性，因此可以借助 VNN 解释神经退行性变中升高的 Δ-脑龄的区域特征。在这一方面，探究脑龄差与区域特征之间关联的类似研究仍然有限 [15,68]。 [15] 中的研究采用卷积神经网络方法，直接根据磁共振成像（MRI）图像推断脑龄，并在评估脑龄时为各脑区分配重要性。从原理上说，VNN 在脑龄背景下提供的可解释性与此类似：我们通过识别导致神经退行性变中 Δ-脑龄升高的脑区，推断出 Δ-脑龄的区域特征。此外，VNN 识别出的区域特征与解剖协方差矩阵的特定特征向量或主成分相关。因此，我们的框架推断出的 Δ-脑龄取决于 VNN 根据解剖协方差矩阵的特定主成分操控输入数据的能力。此外，与 [15] 中研究的深度学习模型相比，VNN 的复杂度显著更低。我们的结果表明，在阿尔茨海默病中，使用少于 2000 个可学习参数训练的 VNN 在脑龄背景下表现出（空间上稳健的）区域可解释性。总体而言，VNN 提供的区域表达能力与大量现有相关研究形成鲜明对比；这些研究依赖透明度较低的统计方法，并进一步采用事后分析（例如消融分析 [69–71] ，或探究其与区域特异性标志物 [58] 及精神症状 [72,73]的相关性），以赋予标量升高的 Δ-脑龄效应可解释性。

2 协方差神经网络

VNN 作用于协方差矩阵，其架构与图神经网络 [1]类似。我们首先给出与架构相关的基本定义，随后讨论 VNN 的理论性质。

考虑一个 $m-$ 维随机向量 $\mathbf{x} \in \mathbb{R}^{m \times 1}$，其总体协方差矩阵定义为

$$\mathbf{C} \triangleq \mathbb{E}[(\mathbf{x} - \mathbb{E}[\mathbf{x}])(\mathbf{x} - \mathbb{E}[\mathbf{x}])^{\mathsf{T}}] \, , \tag{1}$$

其中，$\cdot^{\mathsf{T}}$ 是转置算子，$\mathbb{E}[\cdot]$ 是关于$\mathbf{x}$的概率分布的期望。在实践中，我们通常可以获得一个数据集，其中包含$\mathbf{x}$的统计信息。因此，我们还考虑由$\mathbf{x}$的 n 个随机、相互独立且同分布（i.i.d）的样本组成的数据集，表示为$\mathbf{x}_i \in \mathbb{R}^{m \times 1}, \forall i \in \{1, \ldots, n\}$，其中该数据集可以用矩阵形式表示为$\mathbf{X}_n = [\mathbf{x}_1, \ldots, \mathbf{x}_n]$。利用 $\mathbf{X}_n$，我们估计总体协方差矩阵，通常称为样本协方差矩阵，如下所示

$$\hat{\mathbf{C}} \triangleq \frac{1}{n-1} \sum_{i=1}^{n} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^{\mathsf{T}} \, , \tag{2}$$

[1]在本文其余部分中，GCNs 和图神经网络可互换使用。

其中， $\bar{\mathbf{x}}$ 是$\mathbf{X}_n$中样本的样本均值。接下来，我们将讨论为何要将VNN与图神经网络分开研究。

2.1 动机

协方差矩阵广泛存在于采用空间分布式多变量数据采集协议的现实世界应用中 [40, 43–45]。协方差矩阵的特征向量被称为数据集的主成分，并构成广为人知的主成分分析变换 [74]。本节的讨论表明，将以协方差矩阵为图实例化的图卷积算子的谱域表示应用于该图，可得到主成分分析变换。这一观察表明，使用以协方差矩阵为图的图神经网络进行学习，至少部分是通过根据协方差矩阵的特征基操纵数据来实现的。因此，本文为以协方差矩阵为基础、以 VNN 形式实例化的图神经网络研究提供了概念性贡献。我们还给出了理论保证，表明与使用主成分分析进行统计推断的模型相比，该方法具有显著优势。

协方差矩阵 $\mathbf{C}$ 可以被视为表示向量 $\mathbf{x}$随机结构的图的邻接矩阵，其中，$\mathbf{x}$ 的 m 个维度可以看作一个具有 m 个节点的无向图中的节点，其边表示$\mathbf{x}$中元素之间的成对协方差。此外，$\mathbf{C}$ 的特征值编码了数据集沿正交空间中不同方向的变异性，而该正交空间由相应的特征向量或主成分决定。

在图信号处理中，定义在图的节点上的向量被视为图信号，而将图信号投影到图的特征基上即可得到图傅里叶变换 [75]。图傅里叶变换为分析图上的卷积滤波器提供了一种系统的数学工具 [29, 76]。有趣的是，对于表示在有向循环图上的离散周期时间序列，经典傅里叶变换与图傅里叶变换会趋于一致 [77]。与图傅里叶变换类似，我们可以将协方差傅里叶变换定义为随机实例$\mathbf{x}$[2] 在协方差矩阵$\mathbf{C}$ [1, 的特征向量上的投影。定义 1]。 [1] 中的协方差傅里叶变换定义如下。为此，我们利用$\mathbf{C}$ 的特征分解:

$$\mathbf{C} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{\mathsf{T}} , \tag{3}$$

其中，$\mathbf{V} = [\mathbf{v}_1, \ldots, \mathbf{v}_m]$ 是一个大小为 $m \times m$ 的矩阵，其列为特征向量； $\mathbf{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_m)$ 是一个对角矩阵，其对角元素表示 $\mathbf{C}$ 的特征值。

定义 1（协方差傅里叶变换）。随机样本$\mathbf{x}$ 的协方差傅里叶变换（*VFT*）定义为其在$\mathbf{C}$ 的特征子空间上的投影，表示为

$$\tilde{\mathbf{x}} \triangleq \mathbf{V}^{\mathsf{T}}\mathbf{x} . \tag{4}$$

<escape>[2]</escape>为便于记述，后续我们将使用符号 $\mathbf{x}$ 来表示协方差矩阵为$\mathbf{C}$的随机向量的一个随机实例。

i-th entry of $\tilde{\mathbf{x}}$，即 $[\tilde{\mathbf{x}}]_i$ ，表示$\mathbf{x}$ 在特征向量$\mathbf{v}_i$ 上的投影，因此与特征值 λ_i相关联。因此，（4）中的主成分分析变换与 VFT 之间的相似性表明，特征值 λ_i 编码了数据集$\mathbf{X}_n$ 沿主成分$\mathbf{v}_i$方向的变异性。在此背景下，协方差矩阵的特征值在数学上等价于图信号处理中的图频率概念 [75]。

采用依赖线性移位求和运算的卷积滤波器的图神经网络，在根本上表现出对图拓扑变化的ˆ稳定性 [38]。由于样本协方差矩阵$\mathbf{C}$可能相对于$\mathbf{C}$ [78], 受到扰动，因此，稳定性有助于减轻样本数量对统计推断的影响。受此观察启发，我们定义了协方差滤波器（VF）的概念，即协方差矩阵的多项式，并据此刻画 VNN 中的卷积运算。

定义2 (协方差滤波器**)**。 给定一组实数标量参数 $\{h_k\}_{k=0}^{K}$，协方差矩阵$\mathbf{C}$ 上的协方差滤波器定义为

$$\mathbf{H}(\mathbf{C}) \triangleq \sum_{k=0}^{K} h_k \mathbf{C}^k \ . \tag{5}$$

此外，对于输入$\mathbf{x}$ ，协方差滤波器$\mathbf{H}(\mathbf{C})$ 的输出为

$$\mathbf{z} = \mathbf{H}(\mathbf{C})\mathbf{x} \ . \tag{6}$$

将协方差滤波器**H(C)** 应用于输入$\mathbf{x}$ ，相当于整合不同大小邻域中的信息。为阐明这一点，考虑一个参数为 $K = 1$ 和 $h_0 = 0$的协方差滤波器。在此情形下$\mathbf{z}$ 的第 i个元素计算如下

$$[\mathbf{z}]_i = h_1 \sum_{j=1}^{m} [\mathbf{C}]_{ij} [\mathbf{x}]_j \ . \tag{7}$$

因此， $[\mathbf{z}]_i$ 表示按照$\mathbf{C}$的第 i行对所有元素进行线性组合。这一观察表明，$\mathbf{x}$ 中第 i维的邻域（由图表示$\mathbf{C}$推导而来）决定了式（7）中卷积运算的结果 $[\mathbf{z}]_i$ 。对于 $K > 1$，卷积运算会根据权重 h_k 组合多跳邻域（最多 K跳）中的信息。

通过对定义2中的协方差滤波操作的滤波器输出$\mathbf{z}$进行 VFT 谱分析，可得到协方差滤波器的频率响应，并揭示协方差滤波与主成分分析之间的相似性。对$\mathbf{z}$进行 VFT 后，我们有

$$\tilde{\mathbf{z}} = \mathbf{V}^{\mathsf{T}} \mathbf{H}(\mathbf{C})\mathbf{x} \ , \tag{8}$$

$$= \sum_{k=0}^{K} h_k \mathbf{\Lambda}^k \mathbf{V}^{\mathsf{T}} \mathbf{x} = \sum_{k=0}^{K} h_k \mathbf{\Lambda}^k \tilde{\mathbf{x}} \ , \tag{9}$$

其中，$\tilde{\mathbf{x}} = \mathbf{V}^{\mathsf{T}}\mathbf{x}$ 是$\mathbf{x}$ 的协方差傅里叶变换，式 (9) 由式 (8) 根据$\mathbf{C}$的特征向量的正交归一性得出。协方差滤波器的频率响应取决于滤波器抽头 $\{h_k\}$ 和$\mathbf{C}$ 的特征值，其表达式为

$$h(\lambda) = \sum_{k=0}^{K} h_k \lambda^k \ . \tag{10}$$

此外，由于 $\tilde{\mathbf{x}}$ 是$\mathbf{V}$ 中特征向量空间上的$\mathbf{x}$ 的投影，且 $[\tilde{\mathbf{z}}]_i = h(\lambda_i)[\mathbf{V}^{\mathsf{T}}\mathbf{x}]_i$，因此$\mathbf{z}$ 的第 i 个元素与标准主成分分析变换具有相似性。这一观察结果在引理 1 中得到了形式化表述。

引理 1（协方差滤波器与主成分分析的谱）。给定一个协方差矩阵 $\mathbf{C}$，其特征分解见(3)。若输入$\mathbf{x}$ 的主成分分析变换为$\mathbf{y} = \mathbf{V}^{\mathsf{T}}\mathbf{x}$，则存在一个协方差滤波器组 $\{\mathbf{H}_i(\mathbf{C}) : i \in \{1, \ldots, m\}\}$，使得输入$\mathbf{x}$ 在特征向量$\mathbf{v}_i$ 上的投影分数可通过应用协方差滤波器$\mathbf{H}_i(\mathbf{C})$ 恢复，如下所示：

$$[\mathbf{y}]_i = \mathbf{v}_i^{\mathsf{T}} \mathbf{H}_i(\mathbf{C})\mathbf{x} \ , \tag{11}$$

其中，滤波器 $\mathbf{H}_i(\mathbf{C})$ 的频率响应 $h_i(\lambda)$ 由下式给出

$$h_i(\lambda) = \begin{cases} \omega_i, & \textit{if} \quad \lambda = \lambda_i \ , \\ 0, & \textit{otherwise} \end{cases} \quad . \tag{12}$$

引理 1 确立了使用主成分分析处理数据样本与使用特定多项式处理协方差矩阵之间的等价性。

我们之前在 [1] 中的工作表明，与涉及样本协方差矩阵特征向量的主成分分析（PCA）相比，使用样本协方差矩阵的多项式进行信息处理能够对有限样本引起的扰动保持稳定。实际上，在实践中，我们可能只能获取样本协方差矩阵 $\hat{\mathbf{C}}$，它是$\mathbf{C}$的估计量。由于$\mathbf{C}$ 是$\mathbf{C}$的相合估计量，$\hat{\mathbf{C}}$ 的特征值和特征向量会在样本数量趋于无穷时趋近于$\mathbf{C}$ 的特征值和特征向量，即 $n \to \infty$。然而，对于有限的样本数量 n，$\mathbf{C}$ 的特征向量和特征值相对于$\mathbf{C}$ 的特征向量和特征值会发生扰动 [78]。原则上，使用主成分分析进行统计推断可能容易受到不稳定性的影响，这是由于总体协方差矩阵中对应于彼此接近的特征值的特征向量 [46]，从而导致无法复现的统计结论 [47]。在此背景下，我们非正式地陈述 [1]中的定理2。

定理 1（协方差滤波器的稳定性）**.** 考虑一个数据集，其样本协方差矩阵 $\mathbf{C}$ 由n 个样本构成，而对应的总体协方差矩阵为 $\mathbf{C}$。在 *[1],* 的假设 *1* 下，以下结论以高概率成立：

$$\left\| \mathbf{H}(\hat{\mathbf{C}}) - \mathbf{H}(\mathbf{C}) \right\| = \mathcal{O}\left(\frac{\nu}{n^{\frac{1}{2}-\varepsilon}} \right) \ , \tag{13}$$

对于某些 $\nu > 0$ 和 $\varepsilon \in (0, 1/2)$。

定理 1 证明，使用协方差矩阵多项式进行信息处理，可确保对样本协方差矩阵 $\mathbf{C}$ 与 $\mathbf{C}$ [1]之间扰动的稳定性。此外，作为定理 1 的推论，我们可以得出，在不同样本协方差矩阵上实例化的协方差滤波器，其输出差异有界。这些观察结果表明，基于协方差滤波器的统计推断具有稳健性，能够抵御有限样本量的影响，因此可在高置信度下获得一致的统计结果。PCA 不具备此类保证。接下来，我们将讨论基于协方差滤波器的 VNN 架构；因此，VNN 能够继承 coVariance滤波器所提供的稳定性。

2.2 架构

我们从构成VNN架构一层的协方差感知器描述开始。为此，我们借助逐点非线性激活函数 $\sigma(\cdot)$ 的定义，使得对于 $\mathbf{x} = [x_1, \ldots, x_m]$，有 $\sigma(\mathbf{x}) = [\sigma(x_1) \ldots, \sigma(x_m)]$。逐点非线性激活函数的示例包括 ReLU 和双曲正切。

定义 3(协方差感知机**)**。 给定逐点非线性激活函数 $\sigma(\cdot)$、输入 $\mathbf{x}$、协方差滤波器 $\mathbf{H}(\mathbf{C})$ 及其对应的滤波器抽头集合 $\mathcal{H}$，协方差感知机定义为

$$\Phi(\mathbf{x}; \mathbf{C}, \mathcal{H}) \triangleq \sigma(\mathbf{H}(\mathbf{C})\mathbf{x}) \; . \tag{14}$$

VNN 可以通过堆叠感知机来构建，从而形成多层信息处理架构。接下来将对这一观察进行形式化表述。

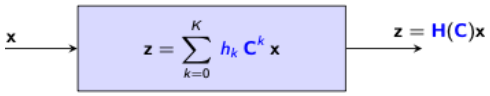
注 1 (多层 VNN**)**。 考虑一个由堆叠 L 个定义 *3* 中定义的协方差感知机组成的 L 层架构。在此情形下，我们用 $\mathbf{H}_\ell(\mathbf{C})$ 表示 *VNN* 的第 ℓ 层中的协方差滤波器，其对应的滤波器抽头集合为 $\mathcal{H}_\ell$。对于给定的逐点非线性激活函数 $\sigma(\cdot)$，第 ℓ 层中协方差感知机的输入 $\mathbf{x}_{\ell-1}$ 与输出 $\mathbf{x}_\ell$ 之间的关系为

$$\mathbf{x}_\ell = \sigma(\mathbf{H}_\ell(\mathbf{C})\mathbf{x}_{\ell-1}) \quad \textit{for} \quad \ell \in \{1, \ldots, L\}, \tag{15}$$

其中 $\mathbf{x}_0$ 是输入$\mathbf{x}$。我们将这种 L 层架构称为 L 层 *VNN*。

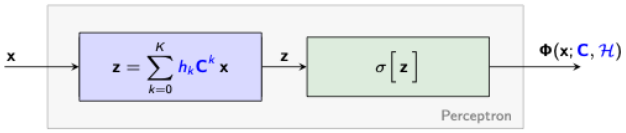
图 1 展示了多层 VNN 的示意图。此外，与其他深度学习模型类似，通过在每一层采用多输入多输出（MIMO）处理，可以增强 VNN架构的表达能力。形式上，考虑一个 VNN 层 ℓ，它可处理 $F_{\ell-1}$ 个 m 维输入，并通过 $F_{\ell-1} \times F_\ell$ 个滤波器组 [38] 输出 F_ℓ 个 m 维输出。在此情形下，输入表示为 $\mathbf{X}_{\mathsf{in}} = [\mathbf{x}_{\mathsf{in}}[1], \ldots\ldots, \mathbf{x}_{\mathsf{in}}[F_{\mathsf{in}}]]$,

a



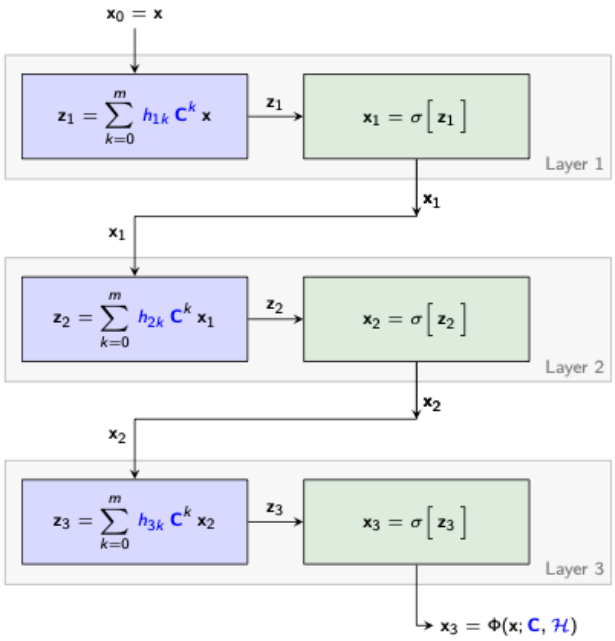
协方差滤波器

b



协方差感知器

c



三层VNN

图 1：**VNN架构基础**。面板**a** 展示了协方差滤波器**H(C)**是**C** 的多项式，以及它作用于输入**x**的情况。面板**b** 展示了基于协方差滤波器**H(C)**和逐点非线性函数 σ构建协方差感知器的过程。协方差感知器定义了VNN的一层。面板**c** 展示了由三个协方差感知器堆叠而成的基本多层VNN架构。

输出定义为 $\mathbf{X}_{\mathsf{out}} = [\mathbf{x}_{\mathsf{out}}[1], \ldots, \mathbf{x}_{\mathsf{out}}[F_{\mathsf{out}}]]$。第 f 个输出$\mathbf{x}_{\mathsf{out}}[f]$ 与输入 $\mathbf{x}_{\mathsf{in}}$ 之间的关系由下式给出

$$\mathbf{x}_{\mathsf{out}}[f] = \sigma\left(\sum_{g=1}^{F_{\mathsf{in}}} \mathbf{H}_{fg}(\mathbf{C})\mathbf{x}_{\mathsf{in}}[g]\right) , \tag{16}$$

其中，$\mathbf{H}_{fg}(\mathbf{C})$ 是用于处理$\mathbf{x}_{\mathsf{in}}[g]$的协方差滤波器。一般性不失，后续讨论聚焦于 $F_\ell = F$, $\forall \ell \in \{1, \ldots, L\}$ 的情形。在这种情况下，所有滤波器抽头的集合为 $\mathcal{H} = \{\mathcal{H}_{fg}^{\ell}\}, \forall f$, $g \in \{1, \ldots, F\}, \ell \in \{1, \ldots, L\}$，其中 $\mathcal{H}_{fg} = \{h_{fg}^{\ell}[k]\}_{k=0}^{K}$和 $h_{fg}^{\ell}[k]$分别是滤波器$\mathbf{H}_{fg}(\mathbf{C})$ 的第 k 个滤波器抽头。因此，我们可以用记号 $\Phi(\mathbf{x};\mathbf{C}\ \mathcal{H})$ 简洁地表示能够执行 MIMO处理的多层 VNN 架构，因为滤波器抽头集合 $\mathcal{H}$ 涵盖了其架构的全部范围。我们还使用记号 $\Phi(\mathbf{x};\mathbf{C}\ \mathcal{H})$ 表示 VNN 的输出。对于最终层有 F 个 m 维输出的 VNN，体素神经网络输出 $\Phi(\mathbf{x};\ \mathbf{C}\ \mathcal{H})$ 的大小为 $m \times F$。

输出 $\Phi(\mathbf{x};\ \mathbf{C}\ \mathcal{H})$ 后接一个读出函数，该函数将 $\Phi(\mathbf{x};\mathbf{C}\ \mathcal{H})$ 映射为所需的输出。本文假设采用非自适应或不可学习的读出函数（例如均值、最大值或最小值函数），以保持 VNN 模型的置换不变性。此外，非自适应读出函数对于 VNN 的可迁移性至关重要（详见第 2.3 节）。

研究 VNN 输出对样本数量 n 的稳健性至关重要，这样才能确保 VNN 统计结果的可复现性。具体而言，当使用从 n_1 或 n_2 个样本估计得到的样本协方差矩阵训练架构时，若 $n_1 \neq n_2$, VNN 输出的变化应受到控制或限制。定理2通过分析 $\|\Phi(\mathbf{x};\mathbf{C}\ \mathcal{H}) - \Phi(\mathbf{x};\ \mathbf{C}\ \mathcal{H})\|$，非正式地陈述了 VNN 稳定性的结果；也就是说，该结果描述了样本协方差矩阵 $\mathbf{C}$ 与总体协方差矩阵 $\mathbf{C}$对应的 VNN 输出之间的差异。该定理此前也已在 [1] 中得到证明。

ˆ**定理2 (**VNN 稳定性**)**。 考虑总体协方差矩阵 $\mathbf{C}$ 及其基于 n个样本得到的估计值 $\hat{\mathbf{C}}$。给定一组滤波器抽头为 $\mathcal{H} = \{\mathcal{H}_{fg}^{\ell}\colon\ f \in \{1, \ldots, F\}, \ell \in \{1, \ldots, L\}\}$ 的协方差滤波器，这些协方差滤波器具有稳定性，并满足

$$\|\mathbf{H}_{fg}^{\ell}(\hat{\mathbf{C}}) - \mathbf{H}_{fg}^{\ell}(\mathbf{C})\| \leq \alpha_n , \tag{17}$$

对于某个 $\alpha_n > 0$，该结论以高概率成立（定理 *1*）。此外，对于逐点非线性函数 $\sigma(\cdot)$，若满足 $|\sigma(a) - \sigma(b)| \leq |a - b|$，则有

$$\|\Phi(\mathbf{x}; \hat{\mathbf{C}}, \mathcal{H}) - \Phi(\mathbf{x}; \mathbf{C}, \mathcal{H})\| \leq LF^{L}\alpha_n . \tag{18}$$

证明。 参见补充材料附录 B。 □

ˆ (17) 中的参数 α_n 表示扰动中相对于$\mathbf{C}$ 和$\mathbf{C}$的有限样本效应，其借鉴自定理 1。通过利用协方差矩阵的扰动理论分析

coVariance滤波器的稳定性，我们还在定理 2 的证明中表明，对于某个 $\varepsilon \in (0, 1/2)$， α_n 随样本数量 n 按 $1/n^{\frac{1}{2}-\varepsilon}$ 的尺度变化。我们注意到，随着 F 或 L 增大，(18) 中的界会变宽，这与图神经网络的相应结果 [38] 一致。然而，如果不对 $\|\Phi(\mathbf{x};\mathbf{C}\ \mathcal{H}) - \Phi(\mathbf{x};\mathbf{C}\ \mathcal{H})\|$ 的下界进行 ˆ 分析，我们就无法断言 VNN 的鲁棒性确实会随着 F 或 L 增大而变差。此外，我们指出，VNN 通过牺牲与相近特征值相关的特征向量之间的判别能力来实现稳定性 [1]。作为推论，我们还指出，利用 (18) 并应用三角不等式，可以直接扩展定理 2，以刻画由 n_1 个样本和 n_2 个样本估计得到的样本协方差矩阵所对应的 VNN 输出之间的差异。

2.3 VNN的可迁移性

VNN的可迁移性这一概念之所以可行，是因为 coVariance 滤波器（定义2）具有可在任意维度的协方差矩阵上实例化的特性，因此 VNN 可以轻松迁移，以处理维度不同的数据集。从实现角度来看，将 VNN 迁移到特征数量不同的数据集，等同于将 VNN 模型 $\Phi(\cdot;\mathbf{C}\ \mathcal{H})$ 中的协方差矩阵替换为另一种尺寸的协方差矩阵，同时保持参数$\mathcal{H}$ 不变。由于我们考虑不同维度的协方差矩阵，因此将尺寸为 $m \times m$ 的协方差矩阵$\mathbf{C}$ 记为 $\mathbf{C}_m$。非正式地说，我们可以将评估可迁移性的目标表述如下。

变分神经网络可迁移性的非正式问题陈述。给定一个数据点 $\mathbf{x}_{m_1}$，它来自一个包含 m_1 个特征且具有相应协方差矩阵 $\mathbf{C}_{m_1}$ 的数据集；另一个数据点 $\mathbf{x}_{m_2}$，它来自一个包含 m_2 个特征且具有相应协方差矩阵 $\mathbf{C}_{m_2}$ 的数据集，我们旨在刻画 VNN输出$\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$ 和 $\Phi(\mathbf{x}_{m_2}; \mathbf{C}_{m_2}, \mathcal{H})$ 收敛的条件。当 $\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$ 和 $\Phi(\mathbf{x}_{m_2}; \mathbf{C}_{m_2}, \mathcal{H})$ 收敛时，我们可以得出结论：参数 $\mathcal{H}$ 可在由 m_1 个特征和 m_2 个特征构成的两个数据集之间迁移。

请注意，当 $m_1 \neq m_2$ 时，VNN 输出 $\Phi(\cdot; \mathbf{C}_{m_1}, \mathcal{H})$ 和 $\Phi(\cdot; \mathbf{C}_{m_2}, \mathcal{H})$ 具有不同的维度，因此直接比较二者并不自然。从根本上说，为了能够分析 VNN 的可迁移性，必须提供一个数学框架，以比较大小不同的向量和协方差矩阵。为了便于比较大小不同的向量，我们考虑一种简单的映射，将向量表示在连续区间 $[0, 1]$ 上。具体而言，给定一个 m 维向量 $\mathbf{x} = [x_1, \ldots, x_m]$，我们可以将 $\mathbf{x}$ 定义为函数$y_{\mathbf{x}}$ 的连续表示：$[0, 1] \mapsto \mathbb{R}$，使得对于 $u \in \mathcal{U}_i$，$y_{\mathbf{x}}(u) = x_i$，其中 $\mathcal{U}_i$ 是与 $\mathbf{x}$ 的第 i 个元素相关联的预定义区间。类似地，我们可以使用映射 $\mathbf{W}_{\mathbf{C}_m}$ 将协方差矩阵 $\mathbf{C}_m$ 映射到由 $[0, 1]^2$ 定义的紧空间：$[0, 1]^2 \mapsto \mathbb{R}$，其中，对于 $u \in \mathcal{U}_i$ 和 $v \in \mathcal{U}_j$，有 $\mathbf{W}_{\mathbf{C}_m}(u, v) = [\mathbf{C}_m]_{ij}$。图 2 中给出了协方差矩阵 $\mathbf{C}$ 的 $y_{\mathbf{x}}$ 和 $\mathbf{W}_{\mathbf{C}}$ 的示意图。其中，区间 $\mathcal{U}_i$由变量 ρ_i 参数化，后文将在式（20）中讨论。请注意，我们可以从 $y_{\mathbf{x}}$ 恢复 $\mathbf{x}$，反之亦然（对于 $\mathbf{C}_m$ 和 $\mathbf{W}_{\mathbf{C}_m}$ 也同样如此）。因此，对于由.....组成的数据点 $\mathbf{x}_{m_1}$ 和 $\mathbf{x}_{m_2}$

m_1 和 m_2 元素，分别而言，连续表示 $y_{\mathbf{x}_{m_1}}$ 和 $y_{\mathbf{x}_{m_2}}$ 之间的接近程度可用作衡量多尺度数据集中数据点相似性的指标。这一观察结果也适用于协方差矩阵$\mathbf{C}_{m_1}$ 和 $\mathbf{C}_{m_2}$之间的比较。

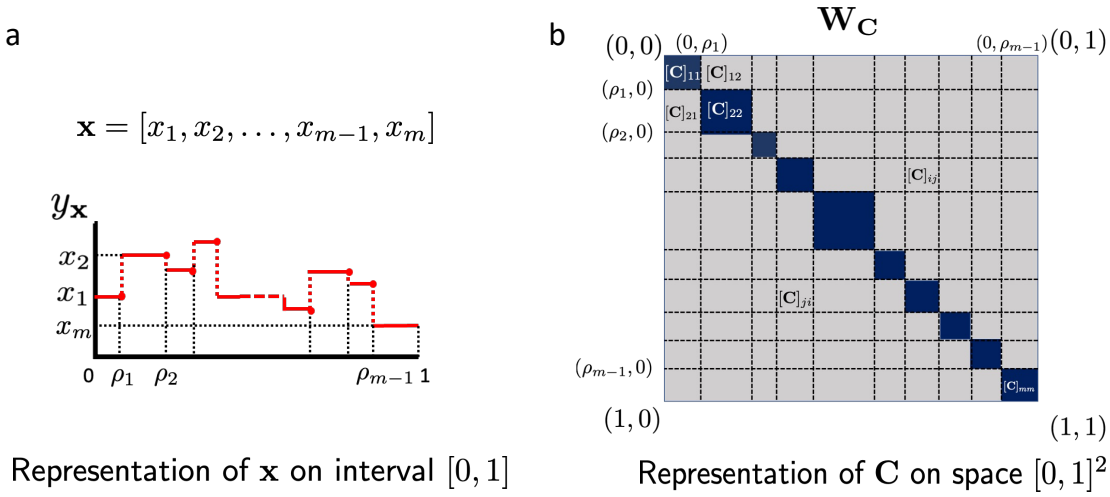


图2：**连续域中 m维向量$\mathbf{x}$ 及其对应协方差矩阵$\mathbf{C}$ 的表示。a.** 通过离散化空间 $[0,1]$获得$\mathbf{x}$ 的表示。**b.** 根据（20）离散化空间 $[0,1]^2$，获得$\mathbf{C}$ 的表示$\mathbf{W}_{\mathbf{C}}$。因此，$\mathbf{W}_{\mathbf{C}}$ 保留了$\mathbf{C}$的对称性。$\mathbf{C}$ 的对角元素所覆盖的区域以蓝色标记。分配给对角元素的方形区域大小与其值成正比。网格的其他部分容纳$\mathbf{C}$的非对角元素。

对于最终层具有 F 个输出的VNN架构，VNN输出 $\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$的维度为 $m_1 \times F$，而 $\Phi(\mathbf{x}_{m_2}; \mathbf{C}_{m_2}, \mathcal{H})$的维度为 $m_2 \times F$。因此，我们可以通过比较输出 $\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$和 $\Phi(\mathbf{x}_{m_2}; \mathbf{C}_{m_2}, \mathcal{H})$中每一列的连续表示，来比较$\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$和 $\Phi(\mathbf{x}_{m_2}; \mathbf{C}_{m_2}, \mathcal{H})$，其中连续表示的定义方式与上述 $y_{\mathbf{x}}$ 相同。对于VNN $\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$，我们使用记号 $y_{m_1}[f]$ 表示 $\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$中 f号输出的连续表示，即$y_{[\Phi(\mathbf{x}_{m_1};\ \mathbf{C}_{m_1},\mathcal{H})]_f}$。类似于 $y_{\mathbf{x}}$与$\mathbf{x}$之间的关系，$[\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$的 f号VNN输出$]_f$及其连续表示 $y_{m_1}[f]$ 在操作上可以互换（详见附录C）。利用上述连续表示，我们现在可以更加具体地描述VNN的可迁移性评估。

VNN 可迁移性的形式化问题陈述：考虑定义在分别具有 m_1 和 m_2 个特征的数据上的两个 VNN $\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$ 和 $\Phi(\mathbf{x}_{m_2}; \mathbf{C}_{m_2}, \mathcal{H})$。若满足以下条件：(a) 输入 $\mathbf{x}_{m_1}$ 和 $\mathbf{x}_{m_2}$ 的连续近似彼此接近，即 $\|y_{\mathbf{x}_{m_1}} - y_{\mathbf{x}_{m_2}}\|_2$ 有界；以及 (b) 协方差矩阵 $\mathbf{C}_{m_1}$ 和 $\mathbf{C}_{m_2}$ 的连续近似彼此接近，即 $\|\mathbf{W}^{\mathbf{C}_{m_1}} - \mathbf{W}_{\mathbf{C}_{m_2}}\|_2$ 有界；我们的目标是证明 VNN 输出 $\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$ 和 $\Phi(\mathbf{x}_{m_2}; \mathbf{C}_{m_2}, \mathcal{H})$ 的连续表示彼此接近，即对于所有

$f \in \{1, \ldots, F\}$。

接下来，我们非正式地陈述本节的主要结果，该结果确立了处理由 m_1 和 m_2 特征组成的数据集的VNN之间的可迁移性。

定理3（VNN的可迁移性）。考虑两个*VNN* $\Phi(\mathbf{x}_{m_1};\ \mathbf{C}_{m_1}, \mathcal{H})$ 和 $\Phi(\mathbf{x}_{m_2};\ \mathbf{C}_{m_2}, \mathcal{H})$，它们均由 L 层组成，且每层有 F 个输出。如果输入的连续表示与协方差矩阵彼此接近，即 $\|y_{\mathbf{x}_{m_1}} - y_{\mathbf{x}_{m_2}}\|_2$ 且 $\|\mathbf{W}^{\mathbf{C}_{m_1}} - \mathbf{W}_{\mathbf{C}_{m_2}}\|_2$ 有界，并且假设*A1*和*A2*成立（见（20）和（24）），则有

$$\|y_{m_1}[f] - y_{m_2}[f]\|_2 \leq LF^L\beta\Big(\frac{1}{m_1^{3\zeta/2-1}} + \frac{1}{m_2^{3\zeta/2-1}}\Big)\,, \forall f \in \{1, \ldots, F\}\,, \tag{19}$$

对于某些 $\beta > 0$ 和 $\zeta \in (2/3, 1]$。

证明。见附录 E。 □

定理3意味着，随着 m_1 和 m_2的增加，VNN各自最终层的全部 F 输出的连续表示$\Phi(\mathbf{x}_{m_1};\ \mathbf{C}_{m_1}, \mathcal{H})$和 $\Phi(\mathbf{x}_{m_2};\ \mathbf{C}_{m_2}, \mathcal{H})$将趋于一致。由于连续表示 $y_{m_1}[f]$ 与VNN输出 $[\Phi(\mathbf{x}_{m_1};\ \mathbf{C}_{m_1}, \mathcal{H})]_f$在操作上可以互换，因此我们也预计输出 $[\Phi(\mathbf{x}_{m_1};\ \mathbf{C}_{m_1}, \mathcal{H})]_f$和 $[\Phi(\mathbf{x}_{m_2};\ \mathbf{C}_{m_2}, \mathcal{H})]_f$的集中趋势度量（例如均值、中位数）将趋于一致。进一步地，如果定理3成立，我们还预计 $\Phi(\mathbf{x}_{m_1};\ \mathbf{C}_{m_1}, \mathcal{H})$和 $\Phi(\mathbf{x}_{m_2};\ \mathbf{C}_{m_2}, \mathcal{H})$的集中趋势度量也将趋于一致。在此背景下，如果VNN的读出函数是非加权平均，我们预计VNN $\Phi(\mathbf{x}_{m_1};\ \mathbf{C}_{m_1}, \mathcal{H})$和 $\Phi(\mathbf{x}_{m_2};\ \mathbf{C}_{m_2}, \mathcal{H})$的统计结果将较为接近，并且这种收敛在 m_1 和 m_2较大时会更加明显。定理3的影响十分广泛，因为我们已经证明，在保持推理任务性能的同时，参数 $\mathcal{H}$ 可以与尺度无关。具体而言，只要训练数据集与新数据集的数据样本和协方差矩阵在连续表示意义下接近，VNN就可以实例化于维度不同于训练数据集的数据集上，并且对于两个数据集使用相同的参数 $\mathcal{H}$ 时，VNN能够恢复出接近的统计结果。因此，相较于基于PCA的分析方法，VNN还具有显著优势，因为主成分受限于某个数据集的特征空间，无法对另一个维度不同的数据集的结构提供任何数学洞见，即使这些数据集可能彼此相关。图3展示了VNN可迁移性的概览。虽然定理3证明背后的严格细节留在附录 E中，但我们将在第19页的高亮文本中简要讨论使我们能够在定理3中确立（19）的主要数学概念和假设。

迄今为止，我们讨论了 VNN 在实现相近统计结果（例如回归模型的结果）方面的可迁移性，前提是不同维度的数据集以某种方式对齐。如前所述，多尺度神经影像数据集为在这一情境下验证定理3的理论保证提供了理想条件。不过，请注意，

VNN 还能够在最终层实现特征层面的表达能力。例如，如果将 VNN 用于回归任务，且读出层是简单的均值函数，则可利用 VNN 最终层输出来刻画数据集中各特征对最终回归结果的贡献。在基于神经影像学的应用中，这一发现尤为值得关注，因为神经影像数据集中的每个特征通常都对应一个特定脑区。因此，VNN 自然而然地提供了一种可行的方法来解释 最终统计结果。此外，如果借助 VNN 实现了这种可解释性，我们还可以进一步将分析与 VNN 的可迁移性结合起来，以检验其在不同维度的数据集之间的空间稳健性。这里的发现促使我们进一步研究 VNN 在刻画因年龄相关神经退行性变而导致 Δ-Age 升高时所涉及脑区方面的效用。

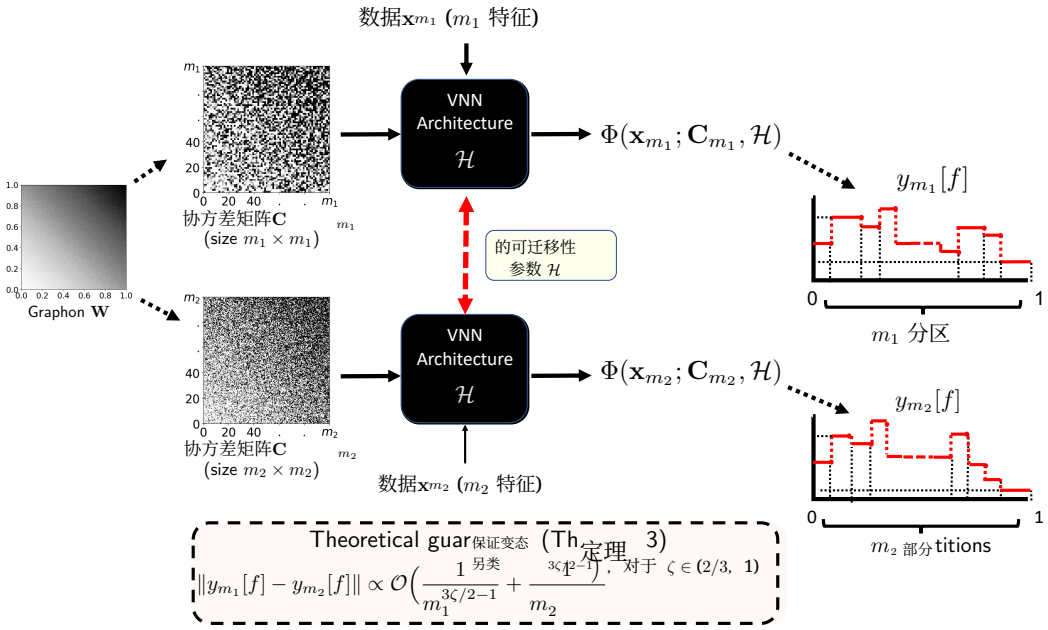


图3：**VNN的可迁移性概述**。$y_{m_1}[f]$ 是 VNN $\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$ 在具有 m_1 个特征的数据和协方差矩阵上实例化后，其第 f 个输出的连续表示。类似地，$y_{m_2}[f]$ 表示 VNN $\Phi(\mathbf{x}_{m_2}; \mathbf{C}_{m_2}, \mathcal{H})$ 在具有 m_2 个特征的数据集上实例化后，其第 f 个输出。如果协方差矩阵 $\mathbf{C}_{m_1}$ 和 $\mathbf{C}_{m_2}$ 的连续对应物，即 $\mathbf{W}^{\mathbf{C}_{m_1}}$ 和 $\mathbf{W}^{\mathbf{C}_{m_2}}$，属于一个收敛到图子 $\mathbf{W}$ （定义4）的序列，并且输入 $\mathbf{x}_{m_1}$ 和 $\mathbf{x}_{m_1}$ 的连续表示彼此接近，那么 $y_{m_1}[f]$ 和 $y_{m_2}[f]$ 之间的收敛性可通过定理3中的 m_1 和 m_2 来刻画。

可迁移性的数学基础

图信号和图的连续表示此前已被用于研究图子信息处理领域中图神经网络的可迁移性 [79]。具体而言，如果一个收敛序列中的图收敛于一种称为图子的极限对象，并且节点数趋于无穷大，那么图神经网络可以在该序列中的图之间实现可迁移 [80]。类似地，我们利用图子理论 [80]和图子信号处理 [79] 建立了定理3。图子是稠密 图的极限（即边数为 $\Theta(m^2)$数量级的图） [81] ，因此适合用于研究通常为稠密矩阵的协方差矩阵的极限。图子的定义见定义4。

定义4 （图子）。图子是一个有界、对称且可测的函数$\mathbf{W}$： $[0,1]^2 \mapsto [-1,1]$。

在将协方差矩阵视为加权图并进行归一化，使其最大特征值为 1 的设定下，我们声称，协方差矩阵的收敛序列$\{\mathbf{C}_m\}$ 收敛，意味着其对应连续表示的序列，即 $\{\mathbf{W}_{\mathbf{C}_m}\}$，收敛到某个图子$\mathbf{W}$ ，前提是$\mathbf{W}_{\mathbf{C}_m}$ 由$\mathbf{C}_m$适当构造而成。该论断基于将 [80,推论 3.9] 推广到我们的设定，而相关的正式表述包含在附录C中的备注 2中。定理3适用于收敛序列 $\{\mathbf{C}_m\}$ 中的任意一对协方差矩阵（在某些稍后将讨论的假设下），因此，参数 $\mathcal{H}$ 可以在基于该序列中不同协方差矩阵实例化的任意两个 VNN 之间进行传递。$\mathbf{W}_{\mathbf{C}_m}$ 的构造依赖于对区间 $\mathcal{U}_i$ 进行适当定义，具体描述如下。

a. 将区间 $[0,1]$ 划分为 m 个互不相交的区间 $[\mathcal{U}_1,\ldots,\mathcal{U}_m]$，使得，

$$\mathcal{U}_i = \begin{cases} [0,\rho_1] & \text{if} \quad i=1 \\ (\rho_{i-1},\rho_i] & \text{if} \quad i\in\{2,\ldots,m\} \end{cases}, \tag{20}$$

其中

$$\rho_i \triangleq \frac{1}{\mathsf{tr}(\mathbf{C}_m)}\sum_{j=1}^{i}[\mathbf{C}_m]_{jj}\,, \tag{21}$$

且 $\mathsf{tr}(\mathbf{C}_m)$ 是$\mathbf{C}_m$的迹。显然， $\rho_m = 1$。

b. 特征 i 与特征 j 之间的关系由 $\mathbf{W}_{\mathbf{C}_m}(u_i,\ u_j) = [\mathbf{C}_m]_{ij}$ 给出，其中 $u_i \in \mathcal{U}_i,\ u_j \in \mathcal{U}_j$。

如果按照上述步骤构造$\mathbf{W}_{\mathbf{C}_m}$ 的连续表示$\mathbf{C}_m$ ，我们将$\mathbf{W}_{\mathbf{C}_m}$ 称为$\mathbf{C}_m$的图子近似。因此，图子极限$\mathbf{W}$ 构成了这样一种模式：协方差矩阵$\mathbf{C}_m$表示其分辨率为 m 时的协方差实现。接下来，我们指出，结果 (19) 取决于与协方差矩阵$\mathbf{C}_m$ 和图子极限$\mathbf{W}$相关的两个主要假设。以下将介绍这些假设。

A1. （协方差矩阵的($(\Omega\ \zeta)$-支配性质）对于序列 $\{\mathbf{C}_m\}$，存在正常数 Ω 和 ζ，使得我们有

$$\frac{1}{\mathsf{tr}(\mathbf{C}_m)}\max_{j\in\{1,\ldots,m\}}[\mathbf{C}_m]_{jj} \le \frac{\Omega}{m^{\zeta}}\,, \tag{22}$$

对于所有有限的 m。我们在附录E中的分析表明，定理3对于任意 $\Omega > 0$ 和 $\zeta \in (2/3,\ 1]$ 均成立。式22中的性质意味着，

$$\frac{1}{\mathsf{tr}(\mathbf{C}_m)}\max_{j\in\{1,\ldots,m\}}[\mathbf{C}_m]_{jj} \to 0 \tag{23}$$

如 $m \to \infty$。我们将满足 (22) 中性质的协方差矩阵$\mathbf{C}_m$ 称为 $(\Omega\ \zeta)$-占优。我们还注意到，(23) 表明，数据集中各个特征的方差分布（由协方差矩阵中相应的对角元素刻画）不能集中在少数特征上。

A2. （图子的利普希茨连续性。）若$\mathbf{W}$ 是序列 $\{\mathbf{W}_{\mathbf{C}_m}\}$ 在 $m\to\infty$ 时的极限，则要使定理 3 成立，$\mathbf{W}$ 必须满足

$$|\mathbf{W}(u_1,v_1) - \mathbf{W}(u_2,v_2)| \le \alpha(|u_1-u_2| + |v_1-v_2|)\,, \tag{24}$$

对于任意 u_1、v_1、u_2、$v_2 \in [0,1]$。满足 (24) 的任意图子称为 α-利普希茨图子。图子$\mathbf{W}$ 的利普希茨连续性决定了任意两个坐标（u_1、v_1）和（u_2、v_2）之间信息的平滑程度。因此，直观上可以预期，对于给定的 m，利普希茨常数 α 较小的图子$\mathbf{W}$ 相比利普希茨常数较大的图子，将能被$\mathbf{W}_{\mathbf{C}_m}$ 更好地近似。

3 应用：脑龄预测

Δ脑龄是认知衰退和神经退行性变的已知生物标志物 [14, 82]。然而，由于缺乏真实值， Δ脑龄这一概念较为抽象；若不识别神经退行性变导致脑龄升高的主要因素，其临床效用便十分有限。本文以阿尔茨海默病（AD）为例，研究病理性年龄相关神经退行性变。年龄是 AD 的主要风险因素，因此，AD 具有表明加速衰老的生物学特征 [83]。我们利用 VNN 的架构，从区域角度进行脑龄预测；结果表明，AD 患者脑龄升高 Δ，同时伴有 AD 特征性多个感兴趣区域的异常。数据集的描述见方法部分第 4.1 小节。接下来，我们将讨论对 VNN 最终层输出的分析如何捕捉年龄相关神经退行性变对大脑各区域的影响。

3.1 识别　　使用 VNN架构识别与神经退行性变相关的脑区

首先，我们使用 VNN 作为回归模型，将健康对照组（HC）的多变量皮层厚度数据拟合到实际年龄。用于脑龄分析的 VNN模型的训练和所选架构的详细信息，见方法部分第 4.2 节。因此，在我们的实验中，VNN通过多层架构处理皮层厚度数据，该架构由根据 HC组皮层厚度推导出的解剖协方差矩阵构成。VNN 模型的最终预测由最终层输出的无权重聚合确定，可以将其理解为各个脑区年龄预测值的非加权平均。因此，VNN架构使我们能够计算每个脑区的“区域残差”（给定脑区的标量输出，即由 VNN 最终层输出得出，减去聚合后的 VNN 输出，或由 VNN 得出的年龄估计），以评估各脑区对 VNN 最终输出的贡献。有关区域残差计算的具体细节，请参阅方法部分第 4.4.1 节。

根据第4.2节中的流程在HC组数据上训练的VNN 模型，从皮层厚度及其相关的解剖协方差矩阵中捕捉到了有关健康衰老的信息。需要注意的是，这些信息可能不足以准确预测实际年龄，即我们可能无法对HC组的实际年龄实现完美拟合。然而，我们在定理2中关于VNN稳定性的理论结果，以及在定理3中关于VNN可迁移性的理论结果表明，在此背景下，VNN所达到的回归性能预计能够对由样本量变化导致的解剖协方差矩阵扰动保持稳定；如果不同数据集各自的协方差矩阵满足特定的理论假设，该回归性能还能够在具有不同维度的数据集之间迁移。

接下来，我们通过将 HC组的解剖协方差矩阵替换为合并组的解剖协方差矩阵，在 HC 和阿尔茨海默病合并组上测试了这些训练好的 VNN模型。

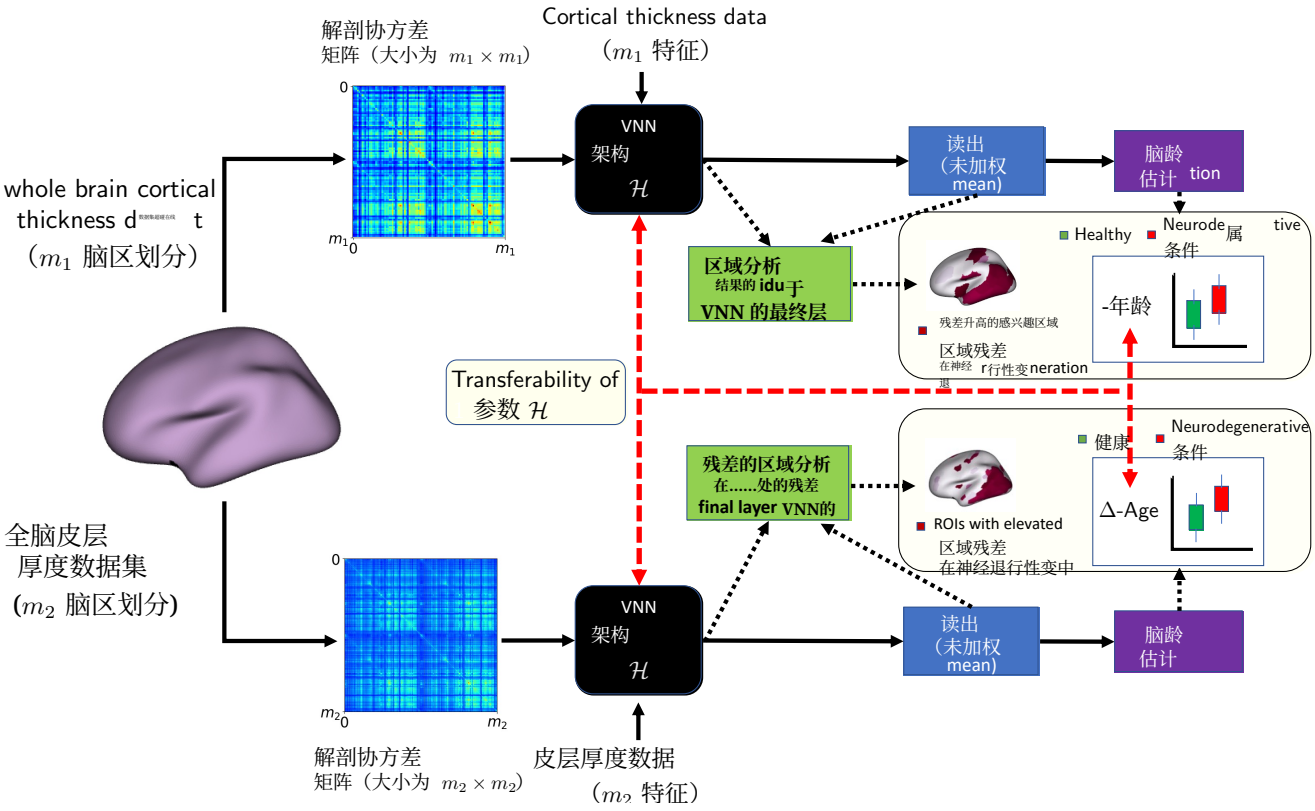


图4：**使用 VNN 进行脑龄预测的框架概述**。通过优化滤波器抽头 $\mathcal{H}$，训练 VNN 作为回归模型，利用 HC组的皮层厚度特征预测实际年龄。VNN 的预测结果通过读出函数获得，该函数对 VNN 最终层的输出进行平均。对于脑龄预测，我们将训练好的 VNN 模型中的协方差矩阵替换为来自 HC组与 AD 个体联合人群的解剖协方差矩阵。对于每个个体，VNN 首先通过读出函数整合 VNN 最终层的输出，形成标量输出，然后应用偏差校正，最终评估脑龄。在联合人群中，对 VNN 最终层的输出进行统计分析，以识别神经退行性变中局部升高的残余效应。这些局部升高的残余效应最终会促成升高的 Δ-年龄效应，因此能够提供因年龄相关神经退行性变而导致的升高 Δ-年龄的区域特征。如果在不同数据集上，实际年龄预测任务的性能可迁移性成立（根据定理3），那么我们预期无需重新训练，就能在这些数据集中观察到相似的 Δ-年龄分布。然而，区域特征并不依赖于 VNN 对 HC组实际年龄的预测性能，并且是在应用年龄偏差校正以计算个体的 Δ-年龄之前得出的。因此，如果这些区域特征能够反映加速衰老，那么即使定理3可能不成立（例如，将 VNN 模型从 FTDC数据集迁移到 OASIS-3 数据集时），我们仍预期在将 VNN 迁移到不同数据集后观察到相似的区域特征。

健康对照组与阿尔茨海默病患者组成的群体。鉴于我们关注的是捕捉加速衰老，我们假设，与HC组相比，具有阿尔茨海默病特征的脑区（例如皮层萎缩程度更高的脑区）在阿尔茨海默病患者中可能会有更高的区域残差。此外，由于VNN的输出是通过依据协方差矩阵的主成分或特征向量对数据进行操作而获得的（第2.1节），我们假设区域残差可能与合并群体的解剖协方差矩阵的主成分相关。此外，定理2所确立的VNN稳定性意味着，区域残差中观察到的任何组间差异都预期能抵御协方差矩阵中的扰动，因此也能抵御用于估计解剖协方差矩阵的健康对照组与阿尔茨海默病患者合并群体构成的变化。

VNN的无尺度架构使我们能够评估在年龄相关神经退行性变中观察到的区域残差升高在空间和统计上的稳健性，无论是在性能可迁移性有保证（定理3和图3）的情形下，还是在无法保证的情形下。附录F中的讨论表明，根据Schaefer图谱的不同脑区划分整理的皮层厚度数据集（称为FTDC数据集，详见第4.1节）属于图3所述的可迁移性理论保证范围。由此可推知，为HC组实际年龄预测训练的VNN模型，其性能可迁移至依据Schaefer图谱不同分辨率组织的数据集。因此，在Schaefer图谱的不同尺度下评估区域残差升高所涉及的脑区，可以为脑龄预测中的可迁移性提供概念验证。附录K中包含了与此相关的实验。

我们重点考察这样一种情形：性能可迁移性的理论保证（定理3）可能不成立，以评估将在依据不同脑图谱整理的数据集之间迁移 VNN，是否仍能保留与 AD相关神经退行性变有关的区域特征。在此背景下，我们研究了将 VNN 从依据 Schaefer图谱整理的数据集迁移至依据 Destrieux（DKT） [84]图谱整理的数据集时的区域特征。在这一背景下，两个数据集来源于不同的群体。依据 DKT图谱整理的数据集不属于针对依据不同尺度的 Schaefer图谱整理的数据集训练的 VNN 可迁移性理论保证的适用范围（见附录F中的图 13及相关讨论）。对于依据 DKT图谱整理的数据集与依据 Schaefer图谱整理的数据集之间的 VNN 可迁移性，缺乏理论保证，这意味着，为了在实际年龄预测任务上取得可比的性能，可能需要通过一种映射对 VNN 进行增强，以考虑DKT图谱与 Schaefer图谱之间的差异。因此，尽管无法保证在依据 Schaefer图谱和 DKT图谱整理的数据集之间，实际年龄预测任务的性能能够直接迁移，但我们假设，如果观测到的区域残差升高效应是由解剖协方差矩阵中的皮层萎缩或神经退行性变相关特征驱动的，那么即使在迁移 VNN 后，也可能观察到这些效应。因此，在数据集之间迁移 VNN 后评估区域残差

在依据不同脑图谱整理的数据集之间进行评估，可能有助于将脑龄升高的潜在影响因素，与在HC组中实现接近完美的实际年龄预测性能这一目标解耦 [3]。

我们进一步扩展了上述实验，考察区域残差与临床痴呆评定（CDR）指标之间的相关性。CDR各分项评分总和常用于临床和研究场景，以评定痴呆严重程度 [85]。较高的CDR评分与更严重的认知和功能状态相关。如果我们关于区域残差升高是由年龄相关的神经退行性变所致这一假设成立，那么我们预计，区域残差升高所涉及的脑区范围，与其区域残差和CDR指标相关的脑区范围之间会显著吻合。

3.2 个体层面的脑龄预测

接下来，我们采用与文献中其他研究一致的流程，根据VNN输出预测个体的 Δ-年龄。对于HC组中的个体，VNN所预测年龄与实际年龄之间的差值所计算出的残差容易产生偏倚。具体而言，年轻个体的年龄往往会被高估，而年长个体的年龄则往往会被低估 [86, 87]。例如，当实际年龄与预测年龄之间的相关性显著小于1时，就可能出现这一现象。为校正残差中与年龄相关的偏倚，我们采用了一种基于线性回归的方法 [88]，并评估了 Δ-年龄。我们假设，AD个体组的 Δ-年龄显著高于HC组。年龄偏差校正的具体方法以及评估 Δ-年龄的流程，详见方法部分第4.4.2小节。此外，基于VNN的可迁移性，我们此前已报告：按照多尺度脑图谱的不同尺度整理的数据集，其 Δ-年龄的分布相似 [2]。

4 方法

4.1 数据

We 本文考虑来自两个独立群体的数据集，具体如下所述。

多尺度 FTDC 数据集。 这些数据集包含从健康对照组（HC； $n = 105$，年龄 $= 62.6 \pm 7.62$ 岁，女性 57 名）中提取的不同分辨率的皮层厚度数据。对于每位受试者，均依据多分辨率Schaefer图谱 [9], 整理皮层厚度数据，采用 100、300 和 500 个脑区的分辨率；随着脑区划分数量增加，皮层厚度估算的分辨率也随之提高

[3]我们指出，此处的讨论并不局限于 Schaefer图谱和DKT图谱，也可以推广到根据不同脑图谱整理的任意一对数据集；在这种情况下，VNN可能无法在它们之间直接迁移。

。使用 ANTs 皮层厚度处理流程 [89, 90]，根据 3T T1 加权磁共振成像（1 mm 各向同性分辨率）计算各图谱脑区内的平均皮层厚度。我们报告了三个数据集的结果：FTDC100、FTDC300 和 FTDC500，它们分别是包含 100、300 和 500 个皮层厚度特征的皮层厚度数据集。此外，FTDC100、FTDC300 和 FTDC500 数据集统称为 FTDC数据集。

OASIS-3 数据集。该数据集源自公开可用的 FreeSurfer 皮层厚度估计数据（托管于 central.xnat.org），如先前报告所述 [64],，其包含认知功能正常的个体（健康对照组；$n = 652$, 年龄为 $= 67.76 \pm 7.88$ 岁，女性 382 人），以及确诊为阿尔茨海默病痴呆且处于不同认知衰退阶段的个体（$n = 209$，年龄为 $= 74.61 \pm 7.13$ 岁，女性 102 人）。根据 Destrieux（DKT）图谱 [84]（包含 148 个皮层区域）整理了皮层厚度特征。在可迁移性研究中，OASIS-3 提供了一个按照不同于 FTDC 数据集中多尺度 Schaefer 图谱的独立脑图谱整理的数据集，因此使我们能够研究 VNN 在多分辨率数据集以外的场景中的可迁移性。为了清晰阐述脑龄预测方法，我们未对确诊为阿尔茨海默病痴呆的个体进行痴呆分期，以进一步划分为轻度认知障碍（MCI）或阿尔茨海默病，并使用标签 AD+ 指代这一群体。AD+ 组中的个体年龄显著高于健康对照组（t检验：p值 $= 4.15 \times 10^{-27}$）。健康对照组和 AD+ 组实际年龄分布的箱线图见图 16。对于 AD+ 组中的 206 名个体，我们获得了在磁共振成像扫描后一年（365 天）内评估的 CDR盒子总分（CDR盒子总分 $= 3.38 \pm 1.73$）。该人群的 CDR盒子总分根据 [91]进行评估。

利用FTDC和OASIS-3中HC组个体的数据，根据第2节所提供的理论框架，研究VNN的可迁移性。此外，我们主要使用OASIS-3 数据集，得出关于 Δ-年龄作为AD+组中加速衰老标志物的结论。在FTDC与OASIS-3数据集之间迁移VNN后所得的观察结果，被用于深入理解在脑龄预测框架中训练VNN以预测脑部实际年龄所具有的方法学意义。除上述数据集外，我们还将VNN的可迁移性扩展至一个小型多尺度数据集上的脑龄预测框架。该数据集由确诊阿尔茨海默病诊断的个体组成（独立于OASIS-3），并依据Schaefer图谱的不同尺度（附录K）进行整理，旨在展示我们的脑龄预测框架所获得的区域特征在多个尺度下的空间稳健性，并验证在OASIS-3 数据集上的研究发现。

4.2 VNN 学习

我们将 VNN 模型用于回归任务，其中多变量特征集被回归为一个标量。请注意，对于一个 m 维输入，由 $\Phi(\mathbf{x}; \hat{\mathbf{C}}\ \mathcal{H})$[4] 表示的架构的体素神经网络输出维度为 $m \times F$，前提是 VNN 架构的最终层具有 F m 维输出。回归输出由读出层确定，该层计算 VNN 最终层所有输出的非加权平均。因此，对于输入$\mathbf{x}$，回归输出为

$$\hat{y} = \frac{1}{Fm} \sum_{j=1}^{m} \sum_{f=1}^{F} [\Phi(\mathbf{x}; \hat{\mathbf{C}}, \mathcal{H})]_{jf} \ . \tag{25}$$

在输出端使用非加权平均进行预测意味着，VNN 模型具有置换不变性（即，最终输出与输入特征和协方差矩阵的排列无关）以及可迁移性。尽管通过添加可学习的或自适应读出函数（例如加权平均或单层感知机），将 VNN 的最终层输出通过可学习映射映射为标量 $\hat{}y$，可能会提升 VNN 模型的性能，但我们后续的实验表明，在脑龄预测的背景下，这会对 VNN 模型的可解释性产生负面影响。对于回归 $\hat{}$ 任务，训练数据集 $\{\mathbf{x}_i, y_i\}_{i=1}^{n}$ 被用于学习 $\mathcal{H}$ 中 VNN $\Phi(\cdot;\mathbf{C}\ \mathcal{H})$ 的滤波器抽头，使其最小化训练损失，即，

$$\mathcal{H}_{\mathsf{opt}} = \min_{\mathcal{H}} \frac{1}{n} \sum_{i=1}^{n} \ell(\hat{y}_i, y_i) \ , \tag{26}$$

其中

$$\hat{y}_i = \frac{1}{Fm} \sum_{j=1}^{m} \sum_{f=1}^{F} [\Phi(\mathbf{x}_i; \hat{\mathbf{C}}, \mathcal{H})]_{jf} \ , \tag{27}$$

且 $\ell(\cdot)$为均方误差（MSE）损失函数。

在我们的实验中，我们训练了四组 VNN 模型，分别针对 FTDC100、FTDC300、FTDC500 和 OASIS-3 数据集中的 HC 组。所有 VNN 的训练过程均相似。我们将数据集随机划分为约 $90/10$ 的训练集/测试集。因此，FTDC数据集的测试集包含 10 名个体，而 OASIS-3 数据集的测试集包含 65 名个体。使用训练集中的所有样本（FTDC数据集为 $n = 95$，OASIS-3 数据集为 $n = 587$ ）计算样本协方差矩阵，并得到大小为 $m \times m$ 的样本协方差矩阵 $\mathbf{C}$ （其中，FTDC100 为 $m = 100$，FTDC300 为$m = 300$，FTDC500 为 $m = 500$，OASIS-3 为 $m = 148$ ）。此外，对于所有数据集，$\mathbf{C}$ 均经过归一化处理，使其最大特征值为 1。接下来，我们在训练集内部进行随机划分，使得 VNN 根据以下两者之间的均方误差损失进行训练。

[4]在方法部分中，我们使用符号 $\hat{}\mathbf{C}$ 表示协方差矩阵，因为在实际实现中，VNN架构是基于样本协方差矩阵构建的。

对于 FTDC 数据集中的 $n = 84$ 个样本和 OASIS-3 数据集中的 $n = 513$ 个样本，预测年龄与真实年龄之间的差异。我们使用 PyTorch 库 [92] 中提供的 Adam 优化器，通过小批量随机梯度下降来优化损失，最多训练 100 轮。FTDC100 数据集的批次大小为 34，FTDC300 数据集为 8，FTDC500 数据集为 12，OASIS-3 数据集为 78。对于训练集的这次排列，将在训练集剩余样本（作为验证集）上均方误差表现最佳的 VNN 模型纳入名义模型集合。对于每个数据集，我们针对 FTDC 各数据集的完整训练集 $n = 95$ 个样本以及 OASIS-3 数据集的 $n = 587$ 个样本，对 VNN 模型进行 100 次排列的训练和验证，因此，每个数据集最终得到 100 个训练好的 VNN 模型（也称为名义模型）。

VNN架构的超参数和优化器的学习率是通过使用 Optuna [93] 套件进行超参数搜索后选定的。对于 FTDC100，VNN采用 $L = 2$层架构，滤波器组的设置为 $F = 26$，每层有 2 个滤波器抽头。优化器的学习率为 0.058。该 VNN 的可学习参数数量为 1456。对于 FTDC300，VNN采用 $L = 2$层架构，滤波器组的设置为 $F = 39$，第一层有 3 个滤波器抽头，第二层有 2 个滤波器抽头。优化器的学习率为 0.0241。该 VNN 的可学习参数数量为 3237。对于 FTDC500，VNN 模型采用 $L = 2$ 层架构，滤波器组的设置为 $F = 27$，第一层有 4 个滤波器抽头，第二层有 2 个滤波器抽头。该 VNN 的可学习参数数量为 1620。Adam优化器的学习率设为 0.0631。对于 OASIS-3，VNN 模型采用 $L = 2$层架构，滤波器组的设置为 $F = 5$，第一层有 6 个滤波器抽头，第二层有 10 个滤波器抽头。Adam优化器的学习率设为 0.059。该 VNN 的可学习参数数量为 290。

请注意，所有VNN的参数均在各自数据集中的健康对照组上学习，之后未针对脑龄预测进行训练。上述将数据集划分为训练集/验证集/测试集来训练VNN的过程，旨在确保VNN未对训练集过拟合。不同数据集中VNN在测试集上的表现见附录G。不过，我们的结果主要关注可迁移性和脑龄预测这两种设置，结果部分报告的两者结果均基于完整数据集。

4.3 VNN的可迁移性

在我们的实验中，我们通过实证研究了第 4.1 节所述不同数据集之间 VNN 模型的可迁移性。一般而言，将 VNN 模型从数据集 A 迁移到数据集 B，意味着该 VNN 在数据集 A 上针对某项推理任务进行了训练，并在数据集 B 上接受同一推理任务的测试。VNN架构的尺度无关特性使 VNN 能够在维度不同的数据集之间迁移。

我们首先描述 评估 VNN 在多尺度皮层厚度 s

ˆ 数据集上的可迁移性，这些数据集源自同一人群。令$\mathbf{C}_{m_1}$ 表示用于训练 VNN 的数据集中的大小为 $m_1 \times \hat{} m_1$ 的样本协方差矩阵。此处，将$\mathbf{C}_{m_1}$ 归一化，使其最大特征值为 1，从而与定义4中的图子定义保持一致。考虑一名实际年龄为 y 且具有皮层厚度数据的个体，其数据表示为大小为 $m_1 \times 1$ 的$\mathbf{x}_{m_1}$ 和大小为$m_2 \times 1$的$\mathbf{x}_{m_2}$。因此，由训练得到的 VNN 模型对该个体的预测为

$$\hat{y}_1 = \frac{1}{Fm_1} \sum_{j=1}^{m_1} \sum_{f=1}^{F} [\Phi(\mathbf{x}_{m_1}; \hat{\mathbf{C}}_{m_1}, \mathcal{H})]_{jf} \ , \tag{28}$$

其中，滤波器抽头集合 $\mathcal{H}$ 由使用包含 m_1 个特征的皮层厚度数据集进行训练确定；我们用记号 $\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$ 表示 VNN 最终层的输出。为了通过实证验证关于 VNN 可迁移性的理论结果，我们旨在评估将协方差矩阵 $\mathbf{C}_{m_1}$ 替换为$\mathbf{C}_{m_2}$时性能的变化，其中$\mathbf{C}_{m_2}$ 由包含 m_2个皮层厚度特征的数据集生成。因此，同一个体基于皮层厚度特征$\mathbf{x}_{m_2}$ 得出的预测年龄为

$$\hat{y}_2 = \frac{1}{Fm_2} \sum_{j=1}^{m_2} \sum_{f=1}^{F} [\Phi(\mathbf{x}_{m_2}; \hat{\mathbf{C}}_{m_2}, \mathcal{H})]_{jf} \ , \tag{29}$$

其中，滤波器抽头 $\mathcal{H}$ 与式（28）中的相同。如果 VNN 模型能够从包含m_1 个皮层厚度特征的数据集迁移到包含 m_2 个特征的数据集，我们预期预测值 ˆy_1 和 ˆy_2 会比较接近。

研究不同维度、不同群体的数据集之间 VNN 可迁移性的流程与上述类似，但有以下扩展。原则上，为了评估从数据集 B 到数据集 A 的可迁移性，我们比较在数据集 A 上训练的 VNN 与从数据集 B 迁移到数据集 A 的 VNN 的性能。显然，如果从数据集 B 迁移到数据集 A 的 VNN 达到了与在数据集 A 上训练的 VNN 相当的性能，我们就可以得出结论：VNN 实现了从数据集 B 到数据集 A 的迁移。

4.4 脑龄预测

使用健康对照组的皮层厚度数据训练 VNN 模型，使其作为回归模型预测实际年龄，预期能够捕捉健康衰老对皮层厚度特征的影响。接下来，我们将介绍在 AD 背景下以区域可解释性刻画脑龄的策略。我们重点关注 OASIS-3 得出的脑龄结果。附录K 收录了在脑龄背景下对多尺度数据集的其他观察结果。

4.4.1 年龄相关神经退行性变中的 VNN残差区域分析

VNN架构允许我们为最终层中 m 个维度中的每一个关联一个标量输出。具体而言，我们有

$$\mathbf{p}_i = \frac{1}{F}\sum_{f=1}^{F}[\Phi(\mathbf{x}_i; \hat{\mathbf{C}}_m, \mathcal{H})]_f \ , \tag{30}$$

其中，$\mathbf{p}_i$ 是一个向量，表示最终层滤波器组中各滤波器对应的所有最终层输出的均值。请注意，$\mathbf{p}_i$ 所有元素的均值即为式（28）中得到的预测值 $\hat{}y_i$ 。对于皮层厚度数据集，我们可以将$\mathbf{p}_i$ 的每个元素与一个独立脑区对应起来。因此，向量$\mathbf{p}_i$ 表示 VNN 对输出 $\hat{}y_i$ 的“区域贡献”。如前所述，参数$\mathcal{H}$ 是在 HC组上学习得到的，并在后续分析中保持不变。本节后续内容表明，我们主要聚焦于 OASIS-3 数据集，以研究 $\hat{}$ 脑龄。对于 OASIS-3 数据集，我们用 $\mathbf{C}_{148}$ 表示由 HC 组的皮层厚度特征构成的协方差矩阵。

接下来，我们利用式（30）研究神经退行性变对脑区的影响。为此，在 OASIS-3 数据集中，我们根据 HC 组和 AD+ 组的合并皮层厚度数据，评估协方差矩阵 $\hat{}\mathbf{C}_{148}^{\mathsf{AD}+}$。根据定理3的稳定性，我们预期 VNN 得出的推断对协方差矩阵中的扰动具有稳定性。因此，后续使用 VNN 进行的评估预计不会受到用于估计解剖协方差矩阵 $\hat{}\mathbf{C}_{148}^{\mathsf{AD}+}$的 HC 组和 AD+ 组合并构成的影响。

对于合并数据集中 HC 和阿尔茨海默病+ 组的每位个体，我们将其皮层厚度数据$\mathbf{x}$ 输入模型 $\Phi(\mathbf{x}; \hat{}\mathbf{C}_{148}^{\mathsf{AD}+}, \mathcal{H})$进行处理，其中参数 $\mathcal{H}$ 是在此前所述的回归任务中，基于 HC 组的数据学习得到的。因此，对于皮层厚度输入$\mathbf{x}$ ，所有最终层输出的均值向量为

$$\mathbf{p} = \frac{1}{F}\sum_{f=1}^{F}[\Phi(\mathbf{x}; \hat{\mathbf{C}}_{148}^{\mathsf{AD}+}, \mathcal{H})]_f \ , \tag{31}$$

而体素神经网络输出为

$$\hat{y} = \frac{1}{148}\sum_{j=1}^{148}[\mathbf{p}]_j \ . \tag{32}$$

此外，我们将特征 a （或本例中由特征 a 表示的脑区）的残差定义为

$$[\mathbf{r}]_a \triangleq [\mathbf{p}]_a - \hat{y} \ . \tag{33}$$

因此，方程 33 使我们能够在区域层面，针对具有皮层厚度数据 $\mathbf{x}$ 的个体，刻画各个脑区相对于 VNN 输出 $\hat{}y$ 的残差。下文中，我们将

根据 (33) 计算的残差称为“区域残差”。因此，当预测某个个体的脑龄高于其实际年龄，且原因是神经退行性疾病时，我们假设这一观察结果是若干生物学上合理的脑区共同作用的综合效应。通过分析式 (33) 中定义的区域残差，可以在区域层面识别导致观察到的脑龄偏高的脑区。因此，残差向量 $\mathbf{r}$ 的各个元素有望作为一种生物标志物，用于识别受年龄相关神经退行性变影响的脑区。

在我们的实验中，对于给定的 VNN 模型，我们评估了数据集中每个个体的残差向量 $\mathbf{r}$ 。此外，在针对 OASIS-3 数据集开展的实验中，我们将健康对照组的残差向量总体表示为 $\mathbf{r}_{\mathrm{HC}}$，将 AD+ 组个体的残差向量总体表示为 $\mathbf{r}_{\mathrm{AD+}}$。残差向量的长度与数据集中皮层厚度特征的数量相同。由于残差向量中的每个元素都与一个不同的脑区相关联，因此我们采用方差分析（ANOVA）检验 HC 组与 AD+ 组个体之间的组间差异。此外，由于 Δ-Age 的升高是本分析关注的生物标志物，我们假设，与 HC 组相比，AD+ 组区域残差均值较高的脑区与捕捉加速衰老最为相关。因此，我们仅报告相对于 HC 组而言，AD+ 组区域残差分布呈升高趋势的脑区结果。此外，如果某一脑区对应的残差向量元素在 AD+ 组与 HC 组之间的组间差异满足以下标准，则认为该差异具有统计学显著性：i）在方差分析模型中，临床诊断标签对应的校正后 p-值（Bonferroni校正）小于 0.05；ii）在以年龄和性别为协变量的协方差分析模型中，临床诊断标签对应的未校正 p-值小于 0.05。附录 H 中提供了该区域分析的一个示例。

请回顾一下，我们将来自HC组的皮层厚度特征训练集进行不同排列，并据此训练了100个不同的VNN模型，将其作为回归模型。我们利用这些训练好的模型，评估观察到的区域残差分布组间差异的稳健性。下文将介绍这一流程。

区域分析结果的稳健性。我们针对每个训练好的VNN模型，分别执行上述区域分析，并统计在上述区域分析中被判定为显著的脑区对应的VNN模型数量。若更多VNN模型将某个脑区判定为显著，则说明该脑区观察到的效应具有更高的稳健性。例如，在几乎所有100个训练好的VNN模型中，AD+组相较于HC组均表现出某脑区区域残差升高，这表明该脑区很可能是导致AD+ 组Δ-年龄升高的重要且稳健的因素。我们使用R中的fsbrain软件包，将区域残差显著升高的稳健性投射到脑模板 [94]上。

VNN 的无尺度架构由协方差滤波器（定义2）实现。此外，非自适应读出函数使我们能够将 VNN 迁移到不同维度的数据集上进行处理。

架构没有任何变化。我们利用所提出框架的这些特性，研究不同数据集之间“可解释性的可迁移性”，而不论这些数据集是否属于 VNN 性能可迁移性理论保证的适用范围。接下来将对此进行说明。

可解释性的可迁移性。请注意，区域残差分析独立于 VNN 模型在回归任务中的表现。因此，我们利用 VNN 的尺度无关特性，评估在上述分析中使某些脑区被判定为显著的统计效应，在迁移到维度不同的数据集后是否仍然保留。为此，我们将一个在 FTDC数据集上训练的 VNN 模型迁移到 OASIS-3 数据集，并将 OASIS-3 中 AD+ 组和健康对照组的区域残差分析判定为显著的脑区标注在脑模板上。如果 AD 的加速衰老特征是 AD+组与健康对照组区域残差中观察到的统计差异背后的驱动因素，那么我们预计，在 OASIS-3 数据集上，无论是使用 OASIS-3 数据集训练的 VNN，还是从 FTDC数据集迁移到 OASIS-3 数据集的 VNN，都会得到定性一致的发现。这一观察结果将为脑龄 VNN 在不同数据集间的“可解释性可迁移性”提供证据，因为在 FTDC数据集上训练的 VNN 模型能够将其构建可解释区域残差的能力从 FTDC 迁移到 OASIS-3。这里需要说明，定性一致是指：对于由在 OASIS-3 和 FTDC数据集上训练的 VNN 得到的区域残差，AD+ 组与健康对照组之间观察到方向相同的显著组间差异（但效应量不一定相同）。通过研究可解释性的可迁移性，我们还可能将 AD+ 组中观察到的加速衰老效应背后的影响因素，与既有文献中对健康对照组实际年龄预测达到近乎完美表现的预期区分开来 [15, 16]。我们假设，可解释性的可迁移性可能由皮层萎缩与 VNN 利用协方差矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$中特征向量（主成分）的能力共同驱动；这些特征向量（主成分）与脑龄预测最为相关。

4.4.2 个体层面的脑龄预测

现有研究主要关注估计脑龄与实际年龄之间的差值，将其作为病理生物标志物 [12], ，其中脑龄是一个标量。正如第 4.4.1 节所述，VNN 可通过区域分析，以一种可解释且系统的方法评估脑龄。不过，借助 VNN，我们也可以通过与该领域现有研究一致的流程，获得脑龄的标量估计值。为了根据 VNN回归输出评估脑龄，我们首先处理了 $\hat{y}$ 与 y 之间差值的系统性偏差；对于年龄较大者，年龄可能被低估，而对于年龄较小者，年龄可能被高估 [86]。这种偏差可能会干扰对脑龄的解读。因此，为了校正这种年龄驱动的偏差，我们采用基于线性回归模型的方法，校正 VNN 年龄估计值中的年龄偏差 [86, 88]。根据这种方法，我们对

VNN 估计年龄 $\hat{y}$ 采用以下偏差校正步骤，以根据实际年龄为 y 且具有皮层厚度数据$\mathbf{x}$ 的个体，获得脑龄 $\hat{y}_{\mathsf{B}}$：

步骤 1. 在完整数据集上拟合一个线性回归模型，以确定以下线性模型中的系数 α 和 β：

$$\hat{y} - y = \alpha y + \beta \, . \tag{34}$$

步骤 2. 按如下方式评估脑龄:

$$\hat{y}_{\mathsf{B}} = \hat{y} - (\alpha y + \beta) \, . \tag{35}$$

$\hat{y}_{\mathsf{B}}$ 和 y 之间的差值通常是关注的生物标志物，定义如下。对于皮层厚度$\mathbf{x}$ 为 y 且实际年龄为 y 的个体，脑龄差 Δ-Age 定义为

$$\Delta\text{-Age} \triangleq \hat{y}_{\mathsf{B}} - y \, , \tag{36}$$

其中，$\hat{y}_{\mathsf{B}}$ 根据（34）和（35）中的步骤，由 Φ（$\mathbf{x}$；$\mathbf{C}$ $\mathcal{H}$）和 y 中的VNN年龄估计值$\hat{y}$ 确定。（34）和（35）中的年龄偏差校正仅针对HC组进行，以校正健康衰老导致的VNN估计值偏差，随后将其应用于AD+ 组。此外，还获得了HC组和AD+组所有个体的 Δ-年龄分布。由于VNN模型得出的区域残差升高，预期AD+组的 Δ-年龄会高于HC组。为阐明这一点，我们考虑一个简单示例:两名个体具有相同的实际年龄 y，其中一名属于AD+ 组，另一名属于HC组。公式（35）表明，使用相同的项 $\alpha y + \beta$对两者相应的VNN输出（分别记为AD+ 组个体的$\hat{y}_{\mathsf{AD+}}$ 和HC组个体的$\hat{y}_{\mathsf{HC}}$）进行年龄偏差校正。因此，只有当AD+ 组个体的VNN预测值$\hat{y}_{\mathsf{AD+}}$ 高于HC组个体的$\hat{y}_{\mathsf{HC}}$时，该AD+ 组个体的 Δ-年龄才会高于HC组个体的相应值。由于VNN预测值$\hat{y}_{\mathsf{AD+}}$ 和$\hat{y}_{\mathsf{HC}}$ 可视为区域层面估计值的无权重聚合（见（32）），因此，AD+ 组中$\hat{y}_{\mathsf{AD+}}$ 高于HC组中$\hat{y}_{\mathsf{HC}}$，直接源于AD+ 组的区域残差（见（33））高于HC组。当此示例中的个体具有不同的实际年龄时，预期年龄偏差校正能够消除 Δ-年龄中由实际年龄引起的任何方差。由于AD+组和HC组的年龄分布不同，我们还通过以年龄和性别为协变量的协方差分析，验证了AD+ 组和HC组在 Δ-年龄方面的差异并非由年龄或性别差异所致。

5 结果

5.1 VNN在回归任务中的可迁移性

我们评估了在 FTDC数据集的 HC组皮层厚度数据上训练为回归模型的 VNN模型，在 Schaefer图谱不同分辨率之间的可迁移性。首先，我们

指出，根据$100-500$分区提取的 FTDC数据集中 HC组皮层厚度特征所形成的协方差矩阵序列是收敛的（附录F）。这一评估很重要，因为我们的定理3理论结果适用于协方差矩阵的收敛序列。此外，我们还评估了 FTDC数据集中的协方差矩阵与 OASIS-3 数据集中健康对照组的协方差矩阵之间的距离（$\mathbf{C}_{148}$，针对 DKT图谱）。这里，OASIS-3 数据集得出的协方差矩阵与 FTDC数据集得出的协方差矩阵之间的距离，显著大于 Schaefer图谱不同分辨率所对应协方差矩阵之间的成对距离（附录F中的图 13）。这一现象可能是由 Schaefer图谱和 DKT图谱构建方式的差异所致。具体而言，DKT图谱是一种解剖图谱，可同时映射脑回和脑沟区域 [84]。相比之下，Schaefer图谱是利用功能性磁共振成像数据 [9] 通过梯度加权马尔可夫随机场模型构建的，本身并不具有解剖学边界。鉴于构建方式存在这些显著差异，我们并不认为根据 DKT图谱整理的数据所得到的解剖协方差矩阵会属于 FTDC数据集协方差矩阵的收敛序列。因此，基于以上讨论，我们预期 VNN的可迁移性适用于 FTDC数据集上的统计推断任务。我们还针对在 FTDC数据集上训练的 VNN，开展了从 FTDC迁移至 OASIS-3 数据集以及反向迁移的可迁移性探索性分析，以评估回归性能的下降情况。

为了研究VNN的可迁移性，我们按照第4.2节中描述的流程，针对HC组个体的皮层厚度与实际年龄之间的回归任务，分别在FTDC100、FTDC300、FTDC500和OASIS-3数据集上训练了VNN模型。每个数据集最终得到100个名义VNN模型（每个模型均基于训练集的一种不同排列进行训练）。VNN中的读出层是非自适应的，它通过计算VNN末层输出的非加权平均值来生成实际年龄的估计值。因此，训练后的VNN无需重新训练或更改架构，即可直接处理具有不同特征数量的数据集。模型性能通过平均绝对误差（MAE）以及VNN输出与真实值之间的皮尔逊相关系数进行量化。

我们在表 1 中列出平均绝对误差，在表 2 中列出真实值与体素神经网络输出之间的皮尔逊相关系数。对于这两个表，行 ID 表示训练 VNN 模型所用的数据集，列 ID 表示报告 VNN 性能所用的数据集（如果训练数据集和测试数据集不同，则表示迁移 VNN 后所用的数据集）。例如，表 1 中行 ID 为“FTDC100”、列 ID 为“FTDC300”的单元格，表示在 FTDC300 数据集 ($m = 300$) 上评估的、针对在 FTDC100 数据集 ($m = 100$) 上训练的 100 个名义VNN模型的平均绝对误差均值和标准差。表 1 和表 2 中行 ID 与列 ID 相同的单元格提供了基线性能，可用于衡量迁移 VNN 后的性能。

Ta 中的结果　表 1 和表 2 表明，VNN 在平均绝对误差和相关性方面的表现

表 1：迁移能力。 不同数据集之间的可迁移性（VNN回归输出相对于真实值的平均绝对误差） .

训练 \ 测试	FTDC100（健康对照组）组）	FTDC300（健康对照组）	FTDC500（健康对照	OASIS-3（健康对照组）
FTDC100（健康对照组）	5.39 ± 0.084	5.5 ± 0.101	5.61 ± 0.132	7.55 ± 0.356
FTDC300（健康对照组）	5.39 ± 0.193	5.41 ± 0.167	5.47 ± 0.169	7.22 ± 0.4
FTDC500（健康对照组）	5.43 ± 0.2	5.38 ± 0.15	5.4 ± 0.169	7.14 ± 0.6
OASIS-3（健康对照组）	6.82 ± 0.365	6.98 ± 0.42	6.47 ± 0.34	5.72 ± 0.076

表 2：跨数据集的可迁移性（VNN输出与真实值之间的皮尔逊相关系数）。

训练 \ 测试	FTDC100（健康对照组）组）	FTDC300（健康对照组）	FTDC500（健康对照	OASIS-3（健康对照组）
FTDC100（健康对照组）	0.49 ± 0.017	0.47 ± 0.018	0.468 ± 0.018	0.387 ± 0.021
FTDC300（健康对照组）	0.498 ± 0.05	0.49 ± 0.042	0.486 ± 0.04	0.374 ± 0.06
FTDC500（健康对照组）	0.51 ± 0.021	0.509 ± 0.02	0.51 ± 0.021	0.374 ± 0.031
OASIS-3（健康对照组）	0.454 ± 0.006	0.437 ± 0.004	0.432 ± 0.003	0.43 ± 0.012

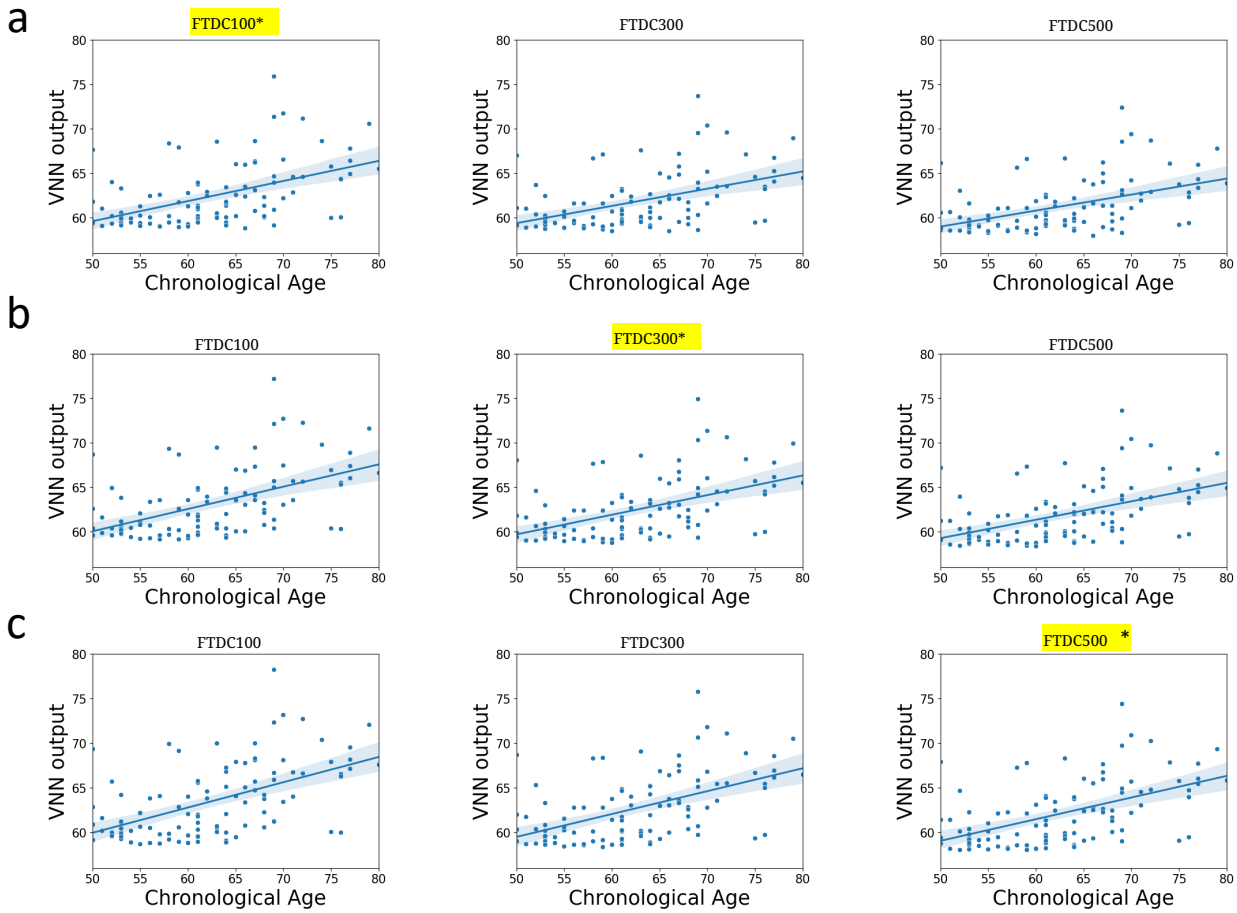


图5：**跨 FTDC 数据集迁移 VNN 后，VNN 回归输出与实际年龄的关系**。面板**a** 展示了在 FTDC100 数据集上训练的全部 100 个 VNN 模型的输出均值；这些模型在依据 Schaefer 图谱不同分辨率整理的 FTDC 数据集上进行测试。图中展示了所有 100 个 VNN 输出的回归输出均值与真实值（实际年龄）的关系；VNN 处理的数据集分别为 FTDC100 数据集（$m = 100$；带星号标记，因为 VNN 在该数据集上训练）、FTDC300 数据集（$m = 300$）和 FTDC500 数据集（$m = 500$）。面板**b** 和面板**c** 分别展示了在 FTDC300 和 FTDC500 数据集上训练的 VNN 所得到的、与面板**a** 类似的结果。

VNN输出与真实值之间的关系在将VNN迁移到根据不同Schaefer图谱分辨率整理的FTDC数据集后仍得以保持。图 5进一步证实了VNN在不同FTDC数据集之间的可迁移性。该图展示了100个VNN模型输出均值的曲线，这些VNN模型在FTDC100 数据集（或FTDC300或FTDC500 数据集）上训练，并用于处理FTDC100、FTDC300和FTDC500数据集。我们还指出，对于PCA-回归模型而言，该实验不可行，因为一个数据集的主成分和回归模型不能直接迁移，用于处理另一个具有不同特征数量的数据集。因此，表 1、表 2和图 5中关于FTDC数据集的观察结果验证了我们关于VNN可迁移性的理论结果。OASIS-3 数据集中HC组的VNN输出与实际年龄之间的曲线见附录 I中的图 17。

此外，表 1 和表 2 还报告了在 FTDC 数据集上训练的 VNN 模型迁移至 OASIS-3 数据集，以及反向迁移时的表现。正如图 13b 中的观察结果所示，在处理根据 DKT图谱整理的数据集时，VNN 的表现显著下降。然而，在 FTDC100、FTDC300 或 FTDC500 数据集上训练的 VNN 模型，在 OASIS-3 数据集上仍保留了预测年龄与真实年龄之间显著的相关性（表 2），且平均绝对误差为 < 8年（表 1）。在 OASIS-3 数据集上训练的 VNN 模型迁移至 FTDC 数据集时，也有类似的观察结果。这些观察结果表明，在 FTDC 数据集上训练的 VNN 模型的输出可能未经过适当缩放，因而无法迁移至 OASIS-3 数据集，反之亦然。因此，要实现回归模型的最佳迁移，Schaefer图谱与 DKT图谱之间需要进行映射或转换。尽管如此，我们的实验表明，在 FTDC 数据集（或 OASIS-3 数据集）上训练的 VNN 模型至少包含了一些有关实际年龄的信息，在处理 OASIS-3 数据集（或 FTDC 数据集）时，这些信息可以迁移。我们将在第 5.2 节进一步探讨这一点。虽然上述观察结果并非本文的重点，但它们为以下问题提供了一些背景：为何针对特定应用对预训练深度学习模型进行微调，可能有助于从神经影像数据中进行统计推断 [7, 95, 96]。

5.2 OASIS-3中HC组的VNN回归模型输出与C_{148}的第一$\hat{}$特征向量相关

引理 1 中的结果表明，基于 VNN 的统计推断在概念上与 PCA 驱动的分析有相似之处。因此，我们进一步探究了表 1 中 VNN 的回归性能是否可由解剖协方差矩阵各主成分的贡献来表征。这里我们重点讨论在 OASIS-3 数据集上训练的 VNN，以及从 FTDC100 数据集迁移到 OASIS-3 数据集的 VNN （在 FTDC 数据集中，FTDC100 数据集的样本数与特征数之比最大）。然而，通过在 FTDC300 或 FTDC500 数据集上训练 VNN 所得的实验观察结果，与在 FTDC100 数据集上训练的结果相似，并且

详见附录 J。

回顾一下，VNN 的最终回归输出是通过无权平均函数作为读出函数形成的。因此，我们可以将读出的功能等价地表示为对不同特征或脑区对 VNN 最终估计值所作贡献的简单汇总（见（30））。因此，对于每位受试者，我们计算了所有 100 个 VNN 模型中每个脑区的贡献向量与协方差矩阵 $\mathbf{C}_{148}$ 的主成分ˆ之间内积（也可等价地称为向量间的点积）的均值。请注意，区域贡献向量的长度与皮层厚度特征的数量相同（即 OASIS-3 中为 148），因此，该向量的每个元素都对应一个不同的脑区。我们使用记号 $\mathbf{p}_{\mathsf{HC}}$ 表示从 OASIS-3 的健康对照组中获得的向量总体。为了计算内积，我们使用了$\bar{}\mathbf{p}_{\mathsf{HC}}$，它是对$\mathbf{p}_{\mathsf{HC}}$ 进行归一化后得到的（范数 = 1）。我们用 $|<\bar{}\mathbf{p}_{\mathsf{HC}},\mathbf{v}_i>|$ 表示 OASIS-3 健康对照组中，与$\mathbf{C}_{148}$的特征向量$\mathbf{v}_i$ 对应的内积总体。请注意，由于 $\bar{}\mathbf{p}_{\mathsf{HC}}$ 中的所有向量都已归一化为范数 1，且$\mathbf{C}_{148}$ 的特征向量默认长度为 1，因此，$|<\bar{}\mathbf{p}_{\mathsf{HC}},\mathbf{v}_i>|$ 表示 OASIS-3 健康对照组中，$\mathbf{p}_{\mathsf{HC}}$ˆ中的向量与$\mathbf{C}_{148}$ 的特征向量$\mathbf{v}_i$ 之间夹角余弦值的总体。

图 6a 绘制了在 OASIS-3 数据集上训练的 VNN 在 HC组中观察到的前 30 个特征向量 ˆ of $\mathbf{C}$ $_{148}$ 的内积均值。第一特征向量 ˆ$\mathbf{C}_{148}$ 与体素神经网络输出的区域贡献向量之间的对齐程度显著强于其他特征向量（0.991 ± 0.0003 在 HC组中），而特征向量 $\mathbf{v}_2$（0.04 ± 0.0029）、$\mathbf{v}_3$（0.0746 ± 0.0023）和 $\mathbf{v}_4$（0.0716 ± 0.0026）的相关性相对较小。因此，可以得出结论：VNN 主要利用了 $\mathbf{C}_{148}$ 的第一主成分（即 $\mathbf{v}_1$）中的信息，从而实现了表 1 和表 2 所示的性能。图 6b 展示了 $\mathbf{v}_1$ 在脑模板上的投影。在图 6b 中，$\mathbf{v}_1$ 主要包含位于海马旁回、楔前叶、颞中回下部和中央前回的双侧脑解剖区域。

ˆ 接下来，我们将讨论 $\mathbf{C}_{148}$ 的特征向量与从 FTDC100 迁移至 OASIS-3 的 VNN 所得区域贡献之间的相关性。请回顾表 1 和表 2：从 FTDC100 迁移至 OASIS-3 的 VNN 在 OASIS-3 数据集上执行实际年龄预测任务时，性能迁移性有所减弱。然而，与图 6a 类似，在此设置下，第一主成分 $\mathbf{v}_1$ 与 VNN 所得区域贡献的相关性最高（0.985 ± 0.0015），而 $\mathbf{v}_3$（0.072 ± 0.004）和 $\mathbf{v}_4$（0.111 ± 0.005）的相关性则相对较低。图 6c 展示了从 FTDC100 迁移至 OASIS-3 数据集的 VNN 所得 $\mathbf{C}_{148}$ 的前 30 个特征向量对应的内积在健康对照组中的均值。因此，尽管 FTDC 与 OASIS-3 数据集之间的定量迁移能力有所减弱（表 1 和表 2），从 FTDC100 数据集迁移至 OASIS-3 数据集的 VNN 仍利用 $\mathbf{C}_{148}$ 的第一特征向量，为 OASIS-3 数据集中的健康对照组生成实际年龄预测值。

这一观察表明，在某一数据集上训练以预测实际年龄的 VNN，迁移到另一个数据集进行处理时，能够获得利用解剖协方差矩阵相关特征向量的能力，而这一能力不受实际年龄预测任务性能可迁移性的影响。

5.3 使用 VNN 在 OASIS-3 中进行脑龄预测

接下来，我们按照第 4.4.1 节所述的流程，对 VNN 模型针对 HC 组和 AD+ 组个体得出的残差进行了区域分析。需要指出的是，已有多项脑龄预测研究所采用的模型，在 HC 组上的平均绝对误差优于我们在表 1 和表 2 [15–17,73] 中报告的结果。然而，我们在此主要旨在证明，使用 HC 组数据训练的 VNN 最终层的输出，使我们能够将特征性地与 AD 神经退行性变相关的贡献归因于相应脑区。因此，本节总结了 VNN 所揭示的洞见；基于深度学习模型黑箱实现的现有脑龄评估框架，大多既未探索这些洞见，也无法获得这些洞见。

首先，我们利用第 5.1 节中使用 OASIS-3 健康对照组数据训练的 VNN 模型，对包含健康对照组和阿尔茨海默病+ 组的合并皮层厚度数据集进行处理。此外，在这一情形下，协方差矩阵由合并数据集构建而成（基于健康对照组和阿尔茨海默病+ 组的皮层厚度数据得到的 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$）。这一选择旨在将神经退行性变引起的解剖结构协方差变化纳入模型。接下来，对于每个 VNN 模型，我们按照第 4.4.1 节所述，评估阿尔茨海默病+ 组和 HC 组中区域残差的分布。附录 H 中给出了一个 VNN 模型区域分析输出结果的示例。

我们注意到，VNN模型得出的区域残差与其预测HC组个体实际年龄时的平均绝对误差表现无关。因此，除了分析OASIS-3 数据集中的脑龄外，我们还进行了探索性区域分析，以评估在FTDC数据集上训练并迁移用于处理OASIS-3 数据集的VNN模型所具有的可解释性。我们旨在通过分析AD+ 组与HC组之间区域残差的组间差异，以及AD+ 组中区域残差与CDR盒子总分之间的关联，来探究VNN所具有的可解释性。

5.3.1 VNN 输出的区域分析揭示了阿尔茨海默病的特征性脑区

OASIS-3 数据集的研究结果表明，与健康对照组相比，阿尔茨海默病+ 组中具有疾病特征的脑区在区域残差分布（定义见（33））方面存在显著差异。请注意，本节所呈现的结果已使用一组包含100个名义模型的模型族进行了稳健性检验，这些模型均经过训练，用于预测 OASIS-3 中健康对照组个体的实际年龄（详见第5.1节）。

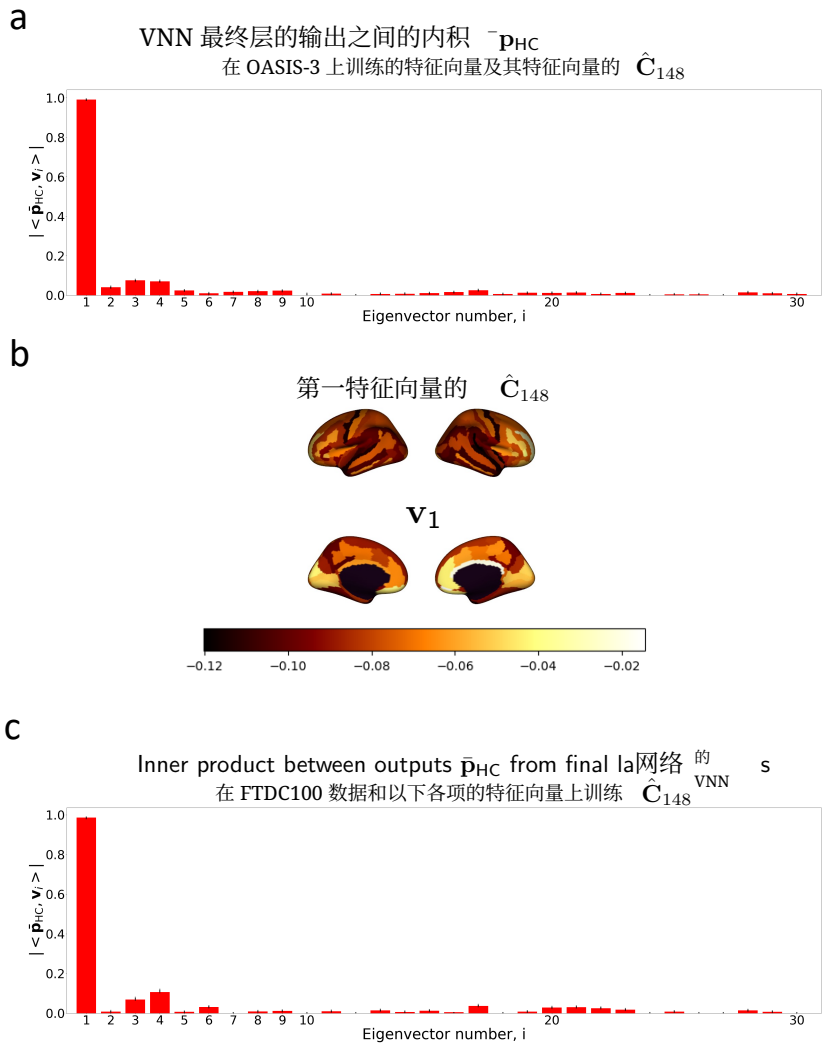


图6：VNN输出的区域贡献归一化向量 $\bar{\mathbf{p}}_{HC}$ 与 $\mathbf{C}_{148}$ （OASIS-3 中 HC组的解剖协方差矩阵）的特征向量之间的内积。图a展示了 $|<\bar{\mathbf{p}}_{HC}$ 与 $\mathbf{v}_i>|$ 的条形图，其中 $\mathbf{v}_i$ 是协方差矩阵 $\mathbf{C}_{148}$ 的第 i 个特征向量（主成分），其对应的特征值按绝对值大小排序后位列第 i，而区域贡献向量 $\bar{\mathbf{p}}_{HC}$ 则由在 OASIS-3 数据集上训练的 VNN 得到。排除了在 OASIS-3 的 HC组中变异系数大于 30% 的特征向量的内积结果（因此，其相应条目设为 0）。对于 HC组中的每位个体，均针对 100 个名义 VNN模型，评估其对应的区域贡献向量 $\bar{\mathbf{p}}_{HC}$ 与 $\mathbf{C}_{148}$ 的特征向量之间的关联。第一特征向量（$\mathbf{v}_1$）与 $\bar{\mathbf{p}}_{HC}$ 的关联最大（0.991 ± 0.0003，在 HC组中）。图b将特征向量 $\mathbf{v}_1$ 绘制在脑模板上。图c展示了在 VNN 从 FTDC100 迁移至 OASIS-3 数据集的情形下，$|<\bar{\mathbf{p}}_{HC}$ 与 $\mathbf{v}_i>|$ 的条形图，涉及 $\mathbf{C}_{148}$ 的前 30 个特征向量以及区域贡献向量 $\bar{\mathbf{p}}_{HC}$。

图7a展示了阿尔茨海默病组与健康对照组在皮层厚度方面显著的组间差异（t 检验；将经 Bonferroni 校正后 p 值达到 < 0.05 的脑区 t统计量投影到脑模板上）。由于阿尔茨海默病组的特征是多个脑区的皮层厚度加速减小，我们预计区域残差分析的结果会与图 7a 中的皮层萎缩模式有显著重叠。图7b展示了在阿尔茨海默病组相对于健康对照组中，各脑区对应残差元素的区域效应升高这一结果的稳健性（通过分析 100 个 VNN 模型确定）。每个 VNN 模型中脑区的显著性，是通过比较健康对照组与阿尔茨海默病组，并依据 VNN最终层输出确定的；VNN 的参数则通过训练确定，训练目标是预测 OASIS-3 中健康对照组的实际年龄。阿尔茨海默病组相对于健康对照组中，区域残差升高最显著的脑区集中在双侧顶下叶、颞叶、内嗅区、海马旁回、楔前叶、胼胝体下区和中央前回。除中央前回和胼胝体下区外，这些脑区均呈现出图 7a 中的皮层萎缩特征；并且已知这些脑区与海马高度相互连接 [97]。

我们的研究结果也与现有关于AD解剖学变化的研究一致。顶下小叶和内侧颞叶区域是阿尔茨海默病典型变异型中众所周知的萎缩区域 [98, 99]。这些区域的皮层萎缩还伴随着tau蛋白沉积 [100]。楔前叶和海马旁区与AD早期阶段有关 [101, 102] ，而楔前叶也是早发型AD患者的一个重要研究区域 [103]。上述脑区在文献 [15]的研究中也被赋予了可解释性。

尽管图 7b 中的结果为 AD+ 组提供了有意义的区域特征，但我们进一步进行了探索性分析，以检验区域残差是否具有任何临床意义。为此，我们针对所有 100 个 VNN 模型，评估了 AD+ 组的 CDR盒子总分与源自 VNN 最终层输出的区域残差之间的相关性。图7c在脑模板上绘制了 AD+ 组中 100 个 VNN 模型所观察到的区域残差与 CDR盒子总分之间的 Pearson 相关系数的均值。有趣的是，相关性最显著的脑区与图 7b 中 AD+ 组区域残差升高所对应的区域特征中最稳健的脑区相一致。我们澄清，在这一背景下，我们的分析并未局限于图 7b 中突出显示的脑区。因此，从图 7b 中 VNN 最终层输出提取的区域特征也捕捉到了有关痴呆严重程度的信息。接下来，我们检验了图 7b 和图7c 中的区域性发现是否转化为 Δ-Age 的升高，以及 AD+ 组中 Δ-Age 与 CDR盒子总分之间的相关性。

5.3.2 Δ-年龄在阿尔茨海默病+ 组中升高，并与临床痴呆评定量表相关

我们按照第 4.4.2 节所述的流程，评估了 OASIS-3 中健康对照组和阿尔茨海默病+ 组的 Δ-年龄。我们还研究了 Δ-年龄与临床痴呆评定量表 CDR各分项评分总和之间的皮尔逊相关系数

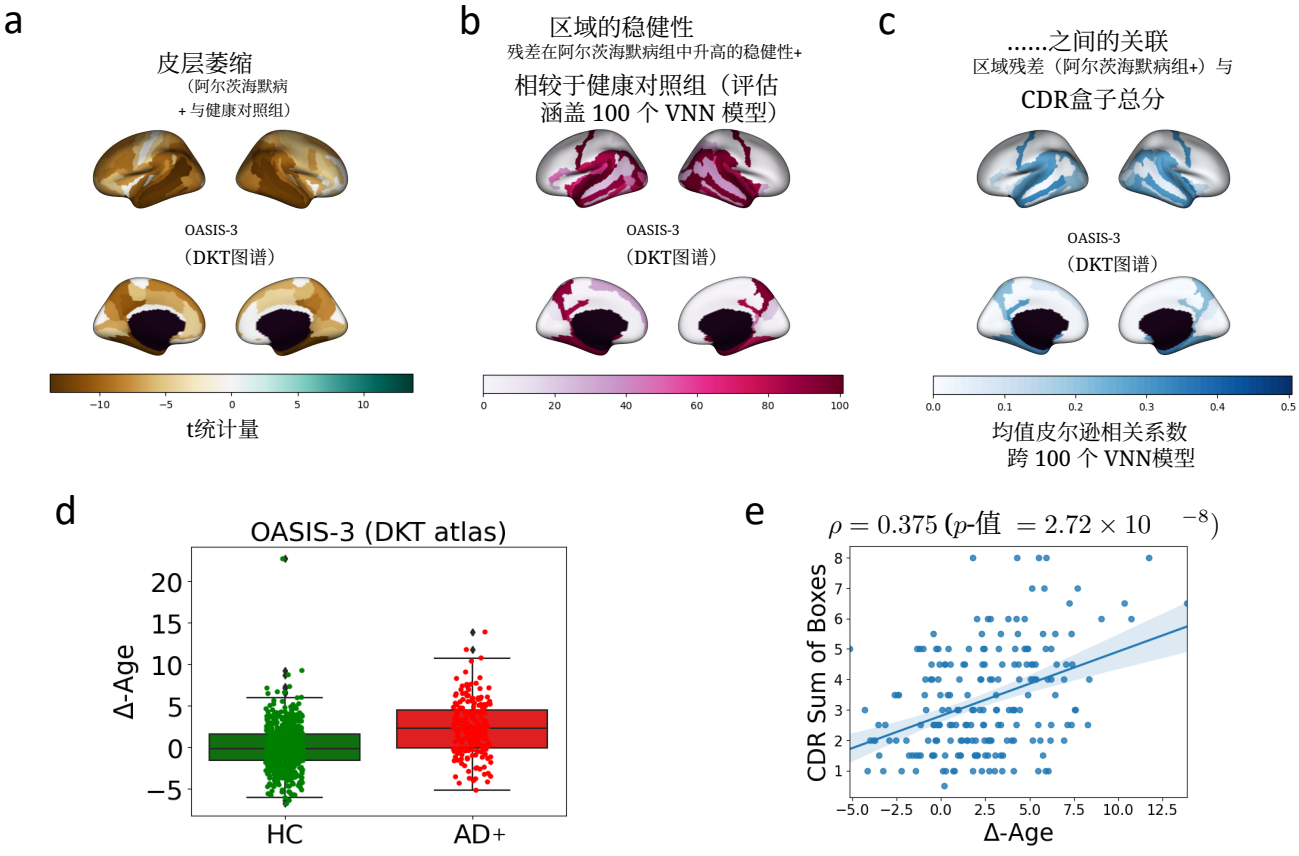

a
皮层萎缩
（阿尔茨海默病
+ 与健康对照组）
OASIS-3
（DKT图谱）
−10
−5
0
5
10
t统计量
b
区域的稳健性
残差在阿尔茨海默病组中升高的稳健性+
相较于健康对照组（评估
涵盖 100 个 VNN 模型）
OASIS-3
（DKT图谱）
0
20
40
60
80
100
c
……之间的关联
区域残差（阿尔茨海默病组+）与
CDR盒子总分
OASIS-3
（DKT图谱）
0.0
0.1
0.2
0.3
0.4
0.5
均值皮尔逊相关系数
跨 100 个 VNN模型
d
OASIS-3 (DKT atlas)
Δ-Age
20
15
10
5
0
−5
HC
AD+
e
$\rho = 0.375$ (p-值 $= 2.72 \times 10^{-8}$)
CDR Sum of Boxes
8
7
6
5
4
3
2
1
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
12.5
Δ-Age

图 7 （续前页）：**可解释的 Δ-Age 在 OASIS-3 数据集中的评估**。面板**a** 显示了阿尔茨海默病+ 组与 HC 组之间皮层厚度组间差异的结果。识别出存在显著差异的脑区（双侧 t 检验，经 Bonferroni 校正的 p-值 < 0.05），并将相应的 t-统计量投射到脑模板上。某脑区的负 t-统计量表明，与 HC 组相比，阿尔茨海默病+ 组在该脑区存在显著的皮层萎缩。面板**b** 显示了在不同脑区中，阿尔茨海默病+ 组相对于 HC 组显著升高的区域残差的稳健性。对于在 OASIS-3 中使用 HC 组训练的 100 个名义 VNN 模型，每个模型的区域残差分析均有助于识别阿尔茨海默病+组相对于 HC 组区域残差显著升高的脑区（方差分析：Bonferroni 校正后的 p-值< 0.05；以年龄和性别为协变量的协方差分析： p-值 < 0.05）。对 100 个 VNN 模型分别进行此实验后，通过统计某脑区被判定为阿尔茨海默病+ 组相对于 HC 组区域残差显著升高的次数，评估该脑区中观察到的显著效应的稳健性。某脑区与阿尔茨海默病+ 组相对于 HC 组区域残差显著升高相关联的次数被投射到脑模板上，可视为该脑区作为阿尔茨海默病+ 组加速衰老的影响因素的稳健性指标。面板**c**将阿尔茨海默病+ 组的区域残差（由 100 个名义 VNN 模型中的每个模型得出）与 CDR盒子总分之间的皮尔逊相关系数均值投射到脑模板上。面板**d**显示了 HC 组和阿尔茨海默病+ 组的 Δ-Age 分布。阿尔茨海默病+组与 HC 组的 Δ-Age 均值相差 2.36 年，且差异显著（Cohen's $d = 0.797$）。此处脑龄升高的效应由面板**b**中的区域特征体现。面板**e** 显示了阿尔茨海默病+ 组中 CDR盒子总分与 Δ-Age 的散点图。 Δ-Age 与 CDR盒子总分之间的相关性，可能归因于面板**c**中观察到的区域残差与 CDR盒子总分之间的相关性。

在阿尔茨海默病+ 组中。图 7d 展示了 HC组和阿尔茨海默病+ 组的 Δ-年龄分布（HC组的Δ-年龄：0 ± 2.66 岁，阿尔茨海默病+组的 Δ-年龄： 2.36 ± 3.22 岁）。阿尔茨海默病+ 组与 HC组的 Δ-年龄差异显著（Cohen's $d = 0.797$，以年龄和性别为协变量进行协方差分析： p-值 $= 3.98 \times 10^{-22}$，偏$\eta^2 = 0.103$）。此外，协方差分析中与年龄相关的 p-值为 0。61，因此，HC组与阿尔茨海默病+ 组实际年龄分布的差异并未导致两组间Δ-年龄的差异。图 7e 绘制了阿尔茨海默病+ 组的 CDR盒子总分与 Δ-年龄之间的散点图。CDR盒子总分与 Δ-年龄之间的皮尔逊相关系数为 0。375（p-值 $= 2.72 \times 10^{-8}$），这表明，在阿尔茨海默病+ 组中评估的 Δ-年龄捕捉到了有关痴呆严重程度的信息。因此，正如预期，随着 CDR盒子总分增加，阿尔茨海默病+ 组的 Δ-年龄可能也会增大。例如，CDR盒子总分大于 4 的个体，其 Δ-年龄均值为 3。76 岁；而 CDR盒子总分小于或等于 4 的个体，其均值为 1。79 岁。

鉴于该 a ge-偏差校正过程是对VNN输出进行的线性变换，因此可以轻松地】【。

由此可以得出结论，图 7b 和图 7c 中区域残差的统计模式导致了图 7d 中升高的 Δ-Age，以及图 7e 中 Δ-Age 与 CDR各分项评分总和之间的相关性。因此，图 7 中研究结果的一个重要意义在于，与健康对照组相比，在 AD相关神经退行性变的背景下，某些脑区对 VNN模型所形成的年龄估计产生了异常贡献。因此，对 VNN 模型最终层输出所推导出的区域残差进行分析，为通过区域特征描述 AD+ 组中的加速生物衰老提供了一种可行的方法。附录 I 中包含了有关 VNN输出以及年龄偏差校正应用前后脑龄的更多图表和详细信息，这些内容展示了 OASIS-3 中健康对照组和 AD+ 组的 Δ-Age 与脑龄的分布情况。

我们此前也曾在 [2] 中报告，经过 VNN模型迁移后，AD、MCI 和 HC 组之间的 Δ-Age 差异在 Schaefer图谱的不同分辨率下均得以保留。我们在 [2] 中的先前研究结果直接源于 VNN 在 Schaefer图谱不同尺度之间的可迁移性。为确保完整性，我们进一步在附录K中展示了，对于一个不同的数据集，在 Schaefer图谱的不同分辨率下评估的 AD 组区域特征之间具有空间一致性。

5.4 AD+ 组中升高的 Δ-年龄区域特征的稳健性取决于区域残差与解剖学协方差矩阵 $\hat{\mathbf{C}}_{148}^{AD+}$ 的主成分之间的相关性

迄今为止，我们已分别证明，作为回归模型训练的 VNN 可在多尺度数据集之间迁移，无需任何重新训练；同时，VNN 为神经退行性变中的加速脑龄提供了可解释视角。然而，尽管 FTDC 与 OASIS-3 数据集之间缺乏定量迁移能力，我们在图 6 中的结果显示，从 FTDC100 迁移至 OASIS-3 的 VNN，以及在 OASIS-3 上训练的 VNN，利用了解剖协方差矩阵的相似特征向量，为 HC组的实际年龄形成预测。在此，我们探究 VNN 为 Δ-年龄 提供的区域可解释性是否取决于解剖协方差矩阵 $\hat{\mathbf{C}}_{148}^{AD+}$ 的主成分。

5.4.1 基于 OASIS-3 训练的 VNN 得到的区域残差与解剖协方差矩阵的主成分相关

我们首先研究了在 OASIS-3 上训练的 VNN 得出的区域残差与 $\hat{\mathbf{C}}_{148}^{AD+}$ 的主成分之间的关系。为此，我们评估了在 OASIS-3 上训练的 VNN 所得归一化残差向量（范数 = 1）与协方差矩阵 $\hat{\mathbf{C}}_{148}^{AD+}$ 的特征向量在阿尔茨海默病+ 组个体中的内积，以确定 $\hat{\mathbf{C}}_{148}^{AD+}$ 的某些特定特征向量（主成分）是否有助于重现图 7b 中的发现。归一化残差向量的.....

标准化残差向量记为 $\bar{\mathbf{r}}^{\mathrm{AD+}}$。对于每个个体，针对用于得出图 7 结果的 100 个 VNN 模型，计算了绝对值内积 $|<\bar{\mathbf{r}}^{\mathrm{AD+}}, \mathbf{v}_i>|$（其中，$\mathbf{v}_i$ 是 $\hat{\mathbf{C}}_{148}^{\mathrm{AD+}}$ 的第 i 个特征向量）的均值。图8a绘制了与 $\hat{\mathbf{C}}_{148}^{\mathrm{AD+}}$ 的 50 个最大特征值相关联的特征向量的平均内积。AD+ 组中与区域残差相关性最大的三个主成分分别是 $\hat{\mathbf{C}}_{148}^{\mathrm{AD+}}$ 的第四个主成分（$|<\bar{\mathbf{r}}^{\mathrm{AD+}}, \mathbf{v}_4>| = 0.6 \pm 0.006$）、第三个主成分（$|<\bar{\mathbf{r}}^{\mathrm{AD+}}, \mathbf{v}_3>| = 0.384 \pm 0.007$）和第一个主成分（$|<\bar{\mathbf{r}}^{\mathrm{AD+}}, \mathbf{v}_1>| = 0.294 \pm 0.0003$）。图 8b 将这些主成分绘制在脑模板上。有趣的是，最大的关联出现在 $\hat{\mathbf{C}}_{148}^{\mathrm{AD+}}$ 的第四个特征向量上，该向量横跨右侧大脑半球的胼胝体下区和内侧额叶区域。第三个特征向量横跨右半球的胼胝体下区，以及两个大脑半球的颞极和海马旁回区域。$\hat{\mathbf{C}}_{148}^{\mathrm{AD+}}$ 的第一个主成分横跨的脑区与 $\hat{\mathbf{C}}_{148}$ 的第一个主成分所横跨的脑区相似（图 6b）。这些观察结果表明，胼胝体下区、海马旁回和颞极区域中观察到的区域效应，是图 7 中观察到的升高 Δ-年龄效应最显著的 AD相关贡献因素。在附录K所述的独立数据集中，这些脑区也被认为是 AD 患者脑龄升高的贡献因素。

5.4.2 从 FTDC100 迁移至 OASIS-3 的 VNN 所得区域残差与 $\hat{\mathbf{C}}_{148}^{\mathrm{AD+}}$ 的主成分相关

在这组实验中，我们使用在 FTDC100 数据集上训练、并迁移至 OASIS-3 进行处理的 VNN，评估了 OASIS-3 数据集中所有个体的区域残差。图9展示了阿尔茨海默病+ 组区域残差与$\hat{\mathbf{C}}_{148}^{\mathrm{AD+}}$的前 30 个特征向量之间内积均值的条形图。与图 8a 类似，由从 FTDC100 迁移至 OASIS-3 的 VNN 得出的区域残差，与解剖协方差矩阵$\hat{\mathbf{C}}_{148}^{\mathrm{AD+}}$的第一、第三和第四个特征向量相关性最大（$|<\bar{\mathbf{r}}^{\mathrm{AD+}}, \mathbf{v}_4>| = 0.56 \pm 0.013$，$|<\bar{\mathbf{r}}^{\mathrm{AD+}}, \mathbf{v}_3>| = 0.362 \pm 0.014$，$|<\bar{\mathbf{r}}^{\mathrm{AD+}}, \mathbf{v}_1>| = 0.29 \pm 0.004$）。这些观察结果表明，在 OASIS-3 数据集中进行脑龄预测时，基于 FTDC100 数据集训练的 VNN 模型能够利用相关特征向量来形成区域残差。根据图 8a 和图 9 的发现，我们可以得出结论：在 OASIS-3 和 FTDC100 上训练、用于预测实际年龄的 VNN 模型，能够利用$\hat{\mathbf{C}}_{148}^{\mathrm{AD+}}$的主成分；这些主成分与图 7 中升高的 Δ-Age 的区域特征相关。$\hat{\mathbf{C}}_{148}^{\mathrm{AD+}}$的第一、第三和第四个特征向量，也是与从 FTDC300 或 FTDC500 数据集迁移至 OASIS-3 数据集的 VNN 得出的区域残差关联最显著的三个特征向量。此处的结果见附录 J 中的图 19。

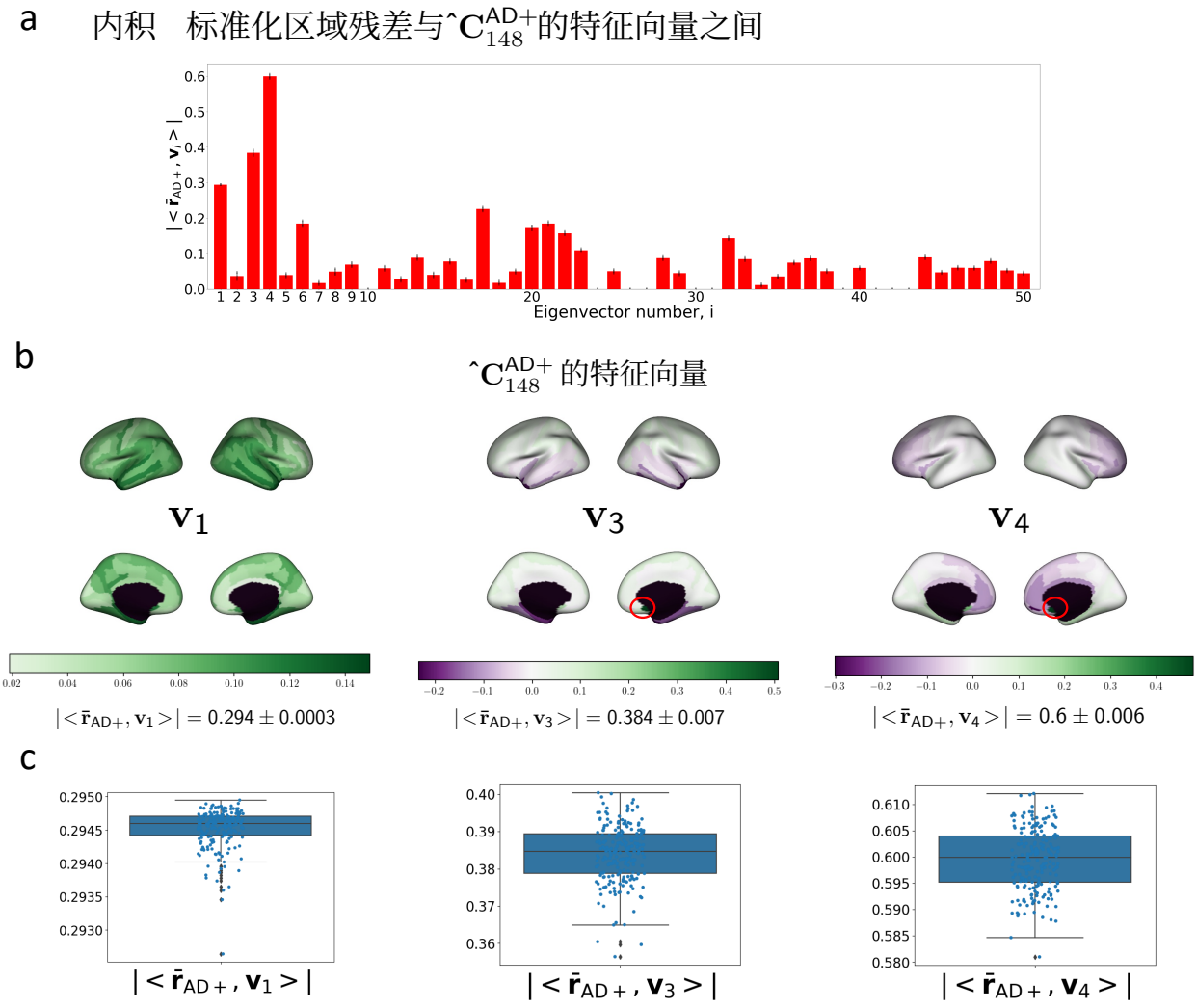


图 8：阿尔茨海默病+ 组中获得的 VNN 输出（由 OASIS-3 上训练的 VNN）之区域残差归一化向量（范数 = 1）与$\hat{C}_{148}^{AD+}$的特征向量之间内积的均值（$\hat{C}_{148}^{AD+}$为健康对照组与阿尔茨海默病+ 组的合并协方差矩阵），以及特征向量图。图a展示了 $i \in \{1, \ldots, 50\}$对应的 $|<\bar{r}^{AD+}, v_i>|$条形图，其中 v_i 是协方差矩阵$\hat{C}_{148}^{AD+}$中与其第 i大特征值对应的第 i个特征向量。条形值根据阿尔茨海默病+ 组个体的 $|<\bar{r}^{AD+}, v_i>|$均值计算得出（阿尔茨海默病+ 组中 $|<\bar{r}^{AD+}, v_i>|$ 的变异系数大于 30% 的特征向量所对应的结果不予纳入）。对于阿尔茨海默病+ 组中的每位个体，其区域残差与$\hat{C}_{148}^{AD+}$的特征向量之间的关联度，是通过对 100 个名义VNN模型（在 OASIS-3 数据集上训练）所得关联度取均值计算得出的。图b在脑模板上绘出了 $|<\bar{r}^{AD+}, v_i>|$取值最大的前三个对应特征向量。右半球的胼胝体下区与 v_3 和 v_4 中绝对值最大的元素相对应，并在相应图中以红圈标出。图c展示了阿尔茨海默病+ 组中 v_1, v_3, 和 v_4 对应的 $|<\bar{r}^{AD+}, v_i>|$分布箱线图。

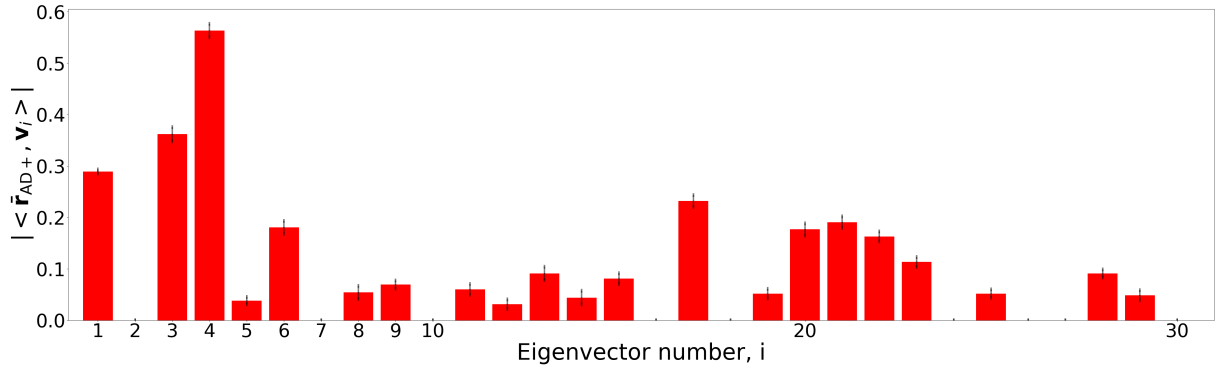


图 9：**阿尔茨海默病 = 1组中个体的体素神经网络输出（从 FTDC100 转移至 OASIS-3）的区域残差归一化向量（范数 = 1）与 $\hat{\mathbf{C}}_{148}^{AD+}$（健康对照组和阿尔茨海默病+ 组合协方差矩阵）的特征向量之间的均值内积**。 此图展示了 $|<\bar{\mathbf{r}}^{AD+}, \mathbf{v}_i>|$关于 $i \in \{1, \ldots, 30\}$ 的条形图，其中 $\mathbf{v}_i$是与协方差矩阵 $\hat{\mathbf{C}}_{148}^{AD+}$的第 i 大特征值相对应的第 i 个特征向量。条形图的数值根据阿尔茨海默病+组中个体的 $|<\bar{\mathbf{r}}^{AD+}, \mathbf{v}_i>|$均值计算得出（排除了变异系数大于 30% 的 $|<\bar{\mathbf{r}}^{AD+}, \mathbf{v}_i>|$对应的特征向量结果）。对于阿尔茨海默病+ 组中的每个个体，均使用 100 个名义VNN模型（在 FTDC100 数据集上训练）评估其区域残差与 $\hat{\mathbf{C}}_{148}^{AD+}$特征向量之间的关联。

5.4.3 将 VNN 从 FTDC100 迁移至 OASIS-3 后，AD+ 组中区域残差升高的脑区的区域特征在定性上得以保留

接下来，我们检查了图 8a 和图 9 中与主成分的相关性是否一致，以及基于在 FTDC100 数据集上训练的 VNN 得出的区域特征，在 OASIS-3 中相对于健康对照组的 AD+ 组区域残差升高相关脑区中，是否也呈现定性一致性。图 10b 展示了将 AD+ 组相对于健康对照组观察到的显著较高区域残差的稳健性投射到脑模板上的结果；与图 7b 的结果类似，稳健性是使用 100 个名义VNN模型评估的。为便于比较，图 10a 纳入了由在 OASIS-3 数据集上训练的 VNN 得出的升高的 Δ-年龄区域特征。

图 10b 中使用在 FTDC100 上训练的 VNN 获得的区域特征与图 10a 中的基本一致。这一观察结果表明，在 FTDC100 数据集上训练的 VNN 能够提取 OASIS-3 数据集中加速衰老的特征性区域特征。这可能是因为在 FTDC100 上训练的 VNN 能够利用与在 OASIS-3 数据集上训练的 VNN 相似的 $\hat{\mathbf{C}}_{148}^{AD+}$ 主成分。

通过训练 VNN 来预测健康个体的实际年龄，我们旨在获取有关健康衰老的信息，这些信息有望有助于检测加速衰老的模式。迄今为止，我们的结果表明，以这种方式训练的 VNN 所识别的加速衰老标志物

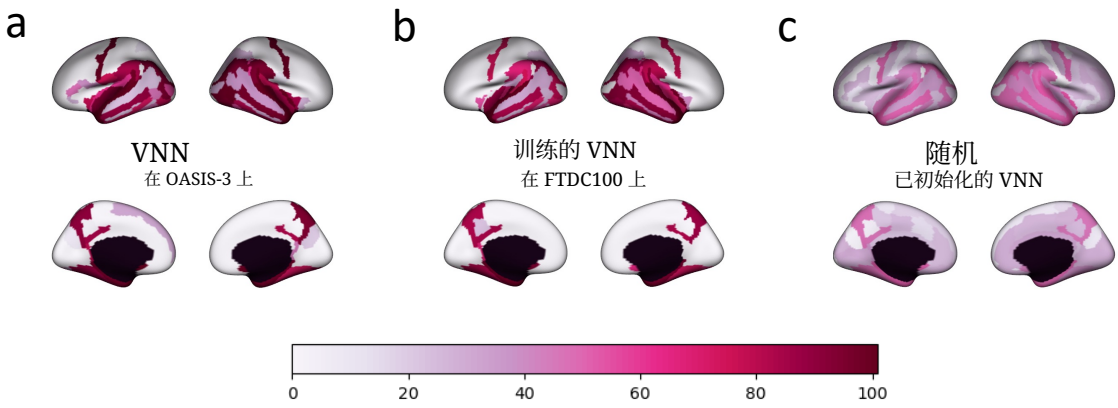


图 10：**基于 AD 区域残差升高稳健性得出的区域特征+在使用不同 VNN 集合的 OASIS-3 中相对于 HC组的区域特征**。面板 **a** 将 OASIS-3 数据集上训练的 VNN 在脑模板上对 AD+ 相对于 HC组的区域残差升高稳健性进行投影，其中 100 个 VNN模型中的各个模型均参与了分析。对于某个脑区，如果有更多 VNN模型判定该脑区在 AD+ 组中相对于 HC组具有显著更高的区域残差，则该脑区作为 AD+ 相对于 HC 观察到的升高的 Δ-年龄的贡献者，其稳健性也越高。面板 **b** 展示了针对 AD+ 和 HC组的区域残差分析所得到的区域特征，其中区域残差来自在 FTDC100 中训练、用于预测 HC组实际年龄的 100 个 VNN。面板 **c** 展示了由 100 个随机初始化的 VNN 得出的相应结果。

无论定量迁移能力如何，都可以将其与解剖协方差矩阵的主成分联系起来。因此，我们推测，训练 VNN 预测实际年龄，有助于微调其利用与阿尔茨海默病+ 组中升高的 Δ-年龄效应相关的主成分来处理输入数据的能力。为进一步支持这一推测，我们使用 100 个随机初始化（即完全未经训练）且架构与在 OASIS-3 上训练的 VNN 相同的 VNN，评估了 OASIS-3 中相对于健康对照组，阿尔茨海默病+ 组区域残差升高的脑区的区域特征。图 10c 展示了在 OASIS-3 数据集上由随机初始化的 VNN 得出的区域特征。图 10c 表明，相对于健康对照组，阿尔茨海默病+ 组区域残差升高的稳健性显著减弱，与图 10a 和图 10b 中相应的结果相比尤为明显。

我们的实验表明，VNN从皮层厚度数据中提取有关加速衰老信息的能力，依赖于其利用$\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$的特征向量或主成分的能力，而不一定取决于其预测健康个体实际年龄的表现。因此，在HC组中训练VNN以预测实际年龄，有助于微调其参数，使其能够利用$\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$中与 AD+ 组的 Δ-年龄相关的特征向量或主成分。上述观察结果有助于将脑龄预测任务与在HC组中实现实际年龄预测近乎完美表现这一目标区分开来。

其他实验：附录 J 的图 19 展示了区域残差与 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$ 特征向量之间的关联，以及从 FTDC300 或 FTDC500 迁移 VNN 后，AD+ 组中区域残差较高的脑区的区域特征。附录 J 中的结果进一步表明，对于在 OASIS-3 上训练的 VNN，在图 6a 中观察到的最显著相关性，在将 VNN 从 FTDC300 或 FTDC500 数据集迁移至 OASIS-3 时大多得以保留。附录 L 中的其他实验有助于确认图 7b 中区域特征对于协方差矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$扰动的稳定性，并表明此处使用的 VNN 模型并未对用于估计解剖协方差矩阵的人群构成过拟合。在附录 M 中，我们报告称，使用自适应读出提升 VNN 在实际年龄预测任务上的表现，可能会削弱 VNN 在脑龄研究中所具有的区域可解释性（同时牺牲 VNN 与尺度无关的特性及其可迁移性保证）。

6 讨论

协方差矩阵上的图卷积算子称为协方差滤波器，它构成了 VNN 架构的基础。协方差滤波器的系数刻画了它根据协方差矩阵的特征谱操控数据以实现学习目标的能力。因此，使用 VNN 进行统计推断与 PCA 驱动的统计方法有相似之处。然而，PCA 通常在给定数据集的特征空间内运作，因此无法提供任何关于……之间相似性的概念。

不同特征数量的数据集的主成分。在本文中，我们研究了 VNN 模型可迁移性的关键属性，该属性使 VNN 能够在具有相似特征但特征数量不同的数据集之间进行迁移。对于由不同数量特征组成的数据集之间的相似性概念，我们借鉴了现有的研究稠密图极限的图子理论 [81]。具体而言，我们的理论结果表明，如果存在一个协方差矩阵序列，该序列在特征数量趋于无穷大的渐近情况下收敛于一个连续极限对象，那么对于统计推断，该序列中的任意两个协方差矩阵之间都可以进行 VNN 迁移。其背后的理论结果依赖于协方差矩阵连续近似的特征谱收敛性，这使得属于收敛序列的协方差矩阵的协方差滤波器输出收敛，进而使 VNN 输出收敛。我们的实验涉及稠密解剖协方差矩阵，因此，基于图子模型的分析无疑适用于研究 VNN 的可迁移性。我们还指出，文献 [79] 中关于图神经网络可迁移性的平行研究结果仅适用于不含自环的图或无标记图，而当我们满足 $\zeta = 1$（即数据集中的所有特征均具有等方差）时，这些结果可以从定理 3 中推导出来。此外，我们指出，稀疏协方差矩阵也具有实际研究价值，因为它们有助于控制计算复杂度 [104]。因此，研究稀疏协方差矩阵上的 VNN 可迁移性是一个值得关注的未来研究方向。

我们还观察到，VNN输出的函数与底层协方差矩阵的若干特征向量相关（图 8），从而验证了我们基于引理 1 提出的假设：VNN会根据协方差矩阵的特征谱处理数据。因此，使用VNN进行信息处理与基于主成分分析的分析有相似之处。与主成分分析不同的是，VNN能够保证在底层协方差矩阵发生扰动时保持稳定（定理2）。此前，我们已通过经验验证了在协方差矩阵受到扰动时，VNN回归性能的稳定性（见 [1]）。图21表明，在脑龄背景下，协方差矩阵受到扰动时VNN的稳定性也延伸到了与评估VNN输出可解释性相关的观察结果。

从数据分析的角度来看，我们通过实证表明，VNN能够从健康对照组的皮层厚度特征中提取有关实际年龄的信息，而且在根据 Schaefer图谱的不同尺度整理的皮层厚度数据集之间，这些信息具有可迁移性，差异极小。我们的理论结果预测了这些观察结果，因为根据 Schaefer图谱不同尺度的皮层厚度特征得出的协方差矩阵构成了一个收敛序列。尽管我们使用 Schaefer图谱，在理论结果的范围内展示了 VNN的可迁移性，但我们的发现并不局限于 Schaefer图谱，还可轻松扩展到任何支持多尺度分区的其他脑图谱。我们还注意到，这种性能可迁移性似乎是以牺牲性能为代价的。具体而言，在附录 M 中，我们报告称，使用自适应读出可以显著提升 VNN的回归性能。使用自

适应读出会牺牲 VNN的端到端可迁移性和尺度无关特性，前提是读出函数中的可学习特征数量取决于数据的维度。因此，对于注重性能的应用，可能有必要进一步研究采用自适应读出的 VNN。

尽管人们对神经科学中类似的多尺度数据集日益感兴趣 [51–53], ，但目前仍缺乏能够在不同尺度上有效识别或利用此类数据集冗余信息的分析策略。在此背景下，VNN 还提供了一种新颖的分析框架，可轻松迁移至多个尺度。对依据不同脑图谱整理的数据集进行分析，也揭示了 VNN 的迁移能力局限（表 1 和表 2）。例如，在 FTDC100 数据集（Schaefer图谱）的皮层厚度特征上训练的 VNN，在 OASIS-3 数据集上未能保持回归性能。这些观察结果表明，为了在多尺度皮层厚度数据集之外，在实际年龄预测中实现最佳泛化，可能需要为 VNN 加入脑图谱之间的映射。

在我们的实验中，VNN是在从老年人群体（年龄> 50 岁）中提取的数据集上训练的。然而，不同年龄组的大脑变化并非在所有解剖区域中都一致——与其他区域相比，有些区域在发育期间增长更多 [105, 106] ，而在衰老期间萎缩也更明显 [107, 108]。因此，与发育和神经退行性变对大脑的影响相关的因素，可能会使本文使用的VNN无法从皮层厚度数据中捕捉到衰老相关信息，尤其是当目标人群的年龄特征比本文所研究的数据集涵盖的范围更广或有所不同时。从技术角度来看，在评估由不同年龄组健康对照组构成的神经影像数据集之间的迁移性时，这些生物学因素可能导致违反VNN可迁移性理论保证所需的假设。此外，在纵向分析中也可能需要考虑类似因素，因为在特定年龄组数据上训练的VNN，可能无法在个体生命周期的所有阶段进行准确推断。

尽管本文严格聚焦于源自结构性MRI的皮层厚度数据集，但神经影像学还包括多种成像模态，例如弥散磁共振成像、功能性磁共振成像和脑电图（EEG），它们能够为脑活动和结构提供不同的见解。与结构性MRI相比，这些模态的信噪比（SNR）通常较低 [109, 110]。因此，必须研究VNN在不同神经影像学模态间的可迁移性，以评估其可迁移性在多尺度数据集上的广泛适用性。

脑龄预测任务面临独特的统计学挑战，因为脑龄没有真实值，且所用的机器学习模型必须能够捕捉年龄相关神经退行性变所驱动的变化，而这些变化会导致加速衰老。在此背景下，我们对 VNN 的研究为脑龄预测方法学作出了奠基性贡献。结果表明，非自适应读出函数不仅有助于提升可迁移性，也对刻画 VNN 模型在脑龄预测任务中的区域可解释性至关重要。具体而言，最终回归输出可以表示为若干个.....的均值

m维向量的元素（如式（32）所示），因此我们能够评估各脑区对最终回归输出的贡献。因此，当发病状态下观察到脑龄升高时，VNN 模型还能够提供区域特征。此外，从 VNN 模型中提取的区域特征与解剖协方差矩阵的某些主成分相关。重要的是，训练 VNN 以预测实际年龄，有助于微调其参数，使其能够利用解剖协方差矩阵中相关的主成分。因此，年龄偏差校正步骤的作用仅限于将 VNN 输出投影到一个空间中，从外行人的角度可以在其中观察相对于实际年龄的生物学衰老。

有趣的是，从 FTDC数据集迁移到 OASIS-3数据集的 VNN，能够利用 OASIS-3 数据集解剖协方差矩阵中与脑龄相关的同一组主成分，正如在 OASIS-3 数据集上训练的 VNN 一样。这使得两组 VNN 在 AD+ 组区域残差升高的脑区中，得出了相一致的区域特征。这一观察进一步表明，VNN 所识别的区域特征能够抵御表征 FTDC 与 OASIS-3数据集间异质性的各种因素（例如神经影像数据质量的差异、MRI采集过程中扫描仪之间的对比差异，以及不同的队列构成）。因此，我们的研究结果表明，VNN 的协方差滤波器所建模的卷积运算，为从皮层厚度数据集中提取具有可解释性、空间稳健性和可重复性，且与脑龄预测相关的信息，提供了一种有用的分析工具。

附录 M 中的结果表明，提高实际年龄预测任务的表现未必能改善脑龄预测，无论是提高 Δ-年龄，还是更好地刻画被认为会导致 Δ-年龄升高的脑区，都是如此。脑龄是一种粗略指标，预计在各种神经退行性疾病中会高于实际年龄，但其判别能力可能不足以区分这些疾病。因此，通过将 Δ-年龄与区域特征相关联，VNN 也提供了一种可行的工具：即使不同病理类型的 Δ-年龄分布相互重叠，也能够对它们进行区分。

目前，对于机器学习模型在实际年龄预测任务中达到何种性能才算合格，尚无明确基准；而该领域现有的大多数方法都缺乏可解释性，因此难以据此预测脑龄。例如， [18] 中的研究指出，使用对实际年龄“拟合较松”的模型进行脑龄预测，是分析中一个潜在的陷阱；而另一项研究则报告称，使用对实际年龄“拟合适中”的深度学习模型能够获得更好的脑龄预测效果 [16]。 [15] 中的研究将体素级可解释性与脑龄联系起来，但并未讨论其方法得出的可解释性模式与实际年龄预测准确度之间的关联。采用对实际年龄拟合质量较高的模型进行脑龄预测，一个动机或许是减轻年龄偏差校正步骤对最终结果的影响 [18]。当脑龄估计方法无法分离出导致其预测出偏高 Δ-年龄的异常时，这种批评似乎是有道理的。然而，其简单性

使 VNN 模型能够在应用年龄偏差校正之前，分析脑龄估计中间步骤的偏差。在此背景下，我们注意到，以非加权平均作为读出函数的 VNN 无法对健康对照组实现完美的实际年龄预测，但它们对阿尔茨海默病患者的脑龄预测却与稳健且可迁移的区域特征相关。因此，VNN 似乎在方法上善于发现促成Δ-年龄升高的神经退行性变驱动因素。尽管拥有数百万个可学习参数，但如果脑龄预测方法依赖复杂且不透明的深度学习模型，那么上述关于 VNN 和脑龄的见解便根本无法获得。

基于上述讨论，我们得出结论：学习如何预测健康对照组的实际年龄，对于 VNN 提供升高的 Δ-Age 的可解释性至关重要。然而，仅凭对健康对照组实际年龄的近乎完美的预测，并不能决定神经退行性变中脑龄预测的质量。从更广泛的角度来看，即使是对健康对照组进行脑龄预测，也是一项复杂的任务，因为在不存在不良健康状况的情况下，多种因素仍可能导致加速衰老 [111–113]。尽管我们并未声称 VNN 在任何评价指标上都能提供“最佳”的脑龄预测，但我们的实验有力地证明，VNN 能够从皮层厚度数据集中提取足够且稳健的信息，从而在神经退行性变中实现具有解剖学可解释性和合理依据的脑龄预测。基于 VNN 的脑龄预测及其相关的区域性解释，仍有可能通过某种方式进一步优化或改进。例如，在 VNN 的训练过程中纳入 DNA甲基化衰老 [114] 的衰老指标，是一个很有前景的未来研究方向，有助于拓展我们对衰老的理解。

包括本文在内的现有研究，都未能明确界定最优脑龄预测的概念。然而，我们注意到，为了在神经退行性变中获得理想的 Δ-年龄，或使 Δ-年龄与辅助指标呈现相关性，而对机器学习模型进行微调，使其更贴合实际年龄，可能会导致这类模型对 Δ-年龄本身过拟合。因此，需要更加重视具有原则性的脑龄预测统计方法，以捕捉导致加速衰老的因素。VNN 等局部可解释且具有理论依据的深度学习模型，为在更广泛的背景下构建统计上和概念上均合理的脑龄预测模型提供了可行且前景广阔的未来方向。

7 数据与代码可用性

FTDC数据集中的个体的磁共振成像和临床数据可通过 https://www.pennbindlab.com/data-sharing 申请获取；经宾夕法尼亚大学神经退行性疾病数据共享委员会审核后，将在合理请求的情况下授予访问权限。OASIS-3 数据集可公开获取，并托管于 central.xnat.org。用于演示VNN的可迁移性和脑龄评估的代码可在 https://github.com/pennbindlab/VNN_Brain_Age 获取。

有关OASIS-3中个体ID以及所有图表源数据的详细信息，可发送邮件至 sihags@pennmedicine.upenn.edu 进行申请。

8 致谢

FTDC数据集的磁共振成像数据由宾夕法尼亚额颞叶变性中心（NIH AG066597）和宾夕法尼亚老年研究所提供。皮层厚度数据由宾夕法尼亚大学图像计算与《科学》实验室提供。OASIS-3数据由纵向多模态神经影像学项目提供：首席研究员：T. Benzinger、D. Marcus、J. Morris；NIH P30 AG066444、P50 AG00561、P30 NS09857781、P01 AG026276、P01 AG003991、R01 AG043434、UL1 TR000448、R01 EB009352。

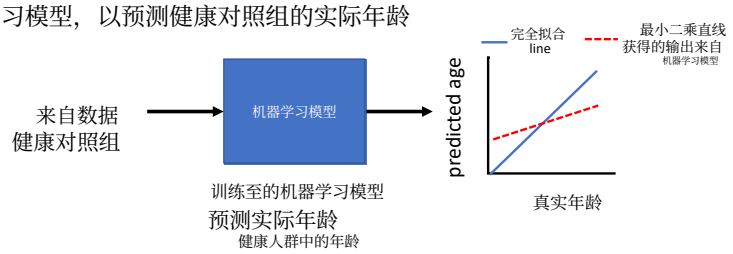
附录

附录 A：基于 VNN 的脑龄预测摘要概述

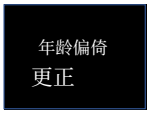
图 11 概述了使用机器学习（ML）模型评估脑龄的一般流程。从图 11 中，我们注意到，如果 ML 模型是一个黑箱，那么在步骤 3 中可能无法确定导致年龄差距升高的因素。此外，在此背景下，年龄偏差校正步骤是否会通过某种统计伪影 [18] 影响最终的 Δ-年龄预测，也尚不明确。因此，可以通过选择一种在步骤 1 中能够对健康对照组的实际年龄实现近乎完美拟合的 ML 模型，尽可能降低年龄偏差校正在 Δ-年龄评估中的作用。然而，无法保证在健康对照组的真实年龄上实现“完美拟合”，就能使 ML 模型捕捉到神经退行性变个体中神经退行性变的影响。

VNN 能够分析每个特征（脑区）对最终输出的贡献。因此，通过研究区域残差中的组间差异，分析不同脑区贡献的增加情况，我们能够确定在患有年龄相关神经退行性变的个体中，促成加速衰老（或更大的 Δ- 年龄）的脑区（图 12）。由此，我们可以验证 VNN 捕获了由神经退行性变驱动的效应，这些效应最终导致个体的 Δ-年龄升高。我们的实验表明，VNN 无法对健康个体的实际年龄实现完美拟合。因此，年龄偏差校正对于通过线性模型将 VNN 模型的输出适当地映射到合适的空间至关重要，从而使临床医生能够观察到神经退行性疾病患者（本文中为阿尔茨海默病患者）的 Δ-年龄升高效应。基于这些观察结果，我们认为 VNN 提供了一个用于脑龄预测的可解释框架。

步骤 1. 训练机器学习模型，以预测健康对照组的实际年龄



步骤 2。 基于线性回归的年龄偏倚校正，用于机器学习模型的输出



步骤 3。获取健康对照组和神经退行性疾病患者群体的年龄差距

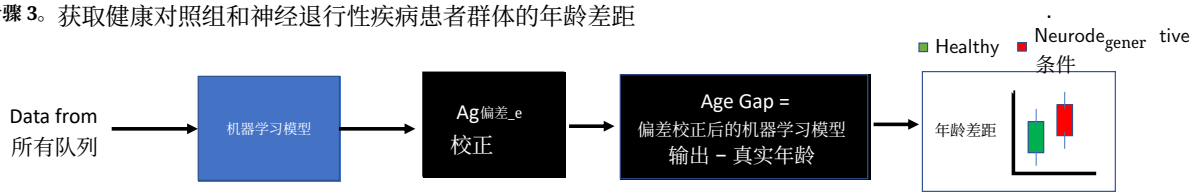


图 11：**现有文献中使用机器学习算法进行脑龄评估的一般概述。步骤 1**包括训练机器学习（ML）模型，以预测健康对照组的实际年龄（真实年龄）。如果预测年龄与真实年龄之间的相关性小于 1，则 ML 模型的输出存在年龄偏倚：模型往往会低估年长者的年龄，而高估年轻者的年龄。为纠正这种偏倚，**步骤 2**中会对 ML 模型的输出应用基于线性回归的模型。根据 ML 模型能够捕捉年龄相关神经退行性变中的加速衰老这一假设，预计神经退行性变患者的 Δ-年龄将显著高于健康对照组（**步骤 3**）。

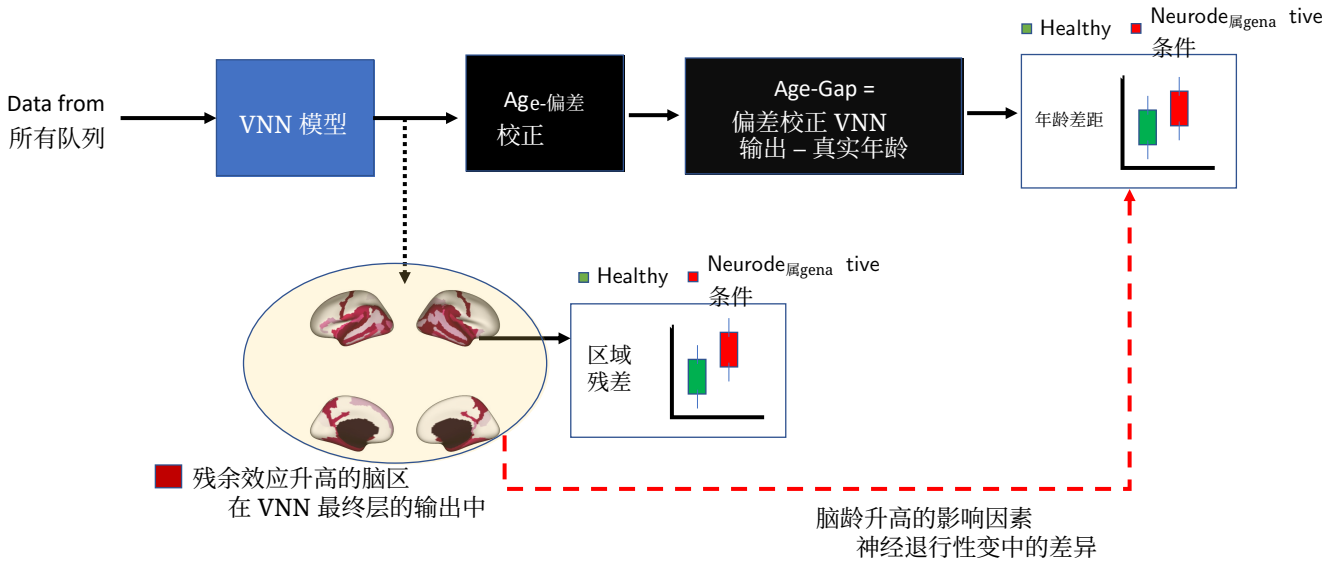


图 12：**VNN 在脑龄预测中提供的可解释性**。通过分析 VNN 的最终层输出，我们可以识别出相较于健康对照组，AD 个体中区域残差较大的脑区。此外，经过年龄偏差校正后，这些脑区中升高的区域残差最终会导致升高的 Δ-年龄。

附录 B 定理 2 的证明

ˆ 我们重述 [1]中的证明。首先，我们注意到，相对于$\mathbf{C}$ 和$\mathbf{C}$ 的协方差滤波器定义为

$$\mathbf{H}(\hat{\mathbf{C}}) = \sum_{k=0}^{m} h_k \hat{\mathbf{C}}^k \quad \text{and} \quad \mathbf{H}(\mathbf{C}) = \sum_{k=0}^{m} h_k \mathbf{C}^k \ . \tag{37}$$

ˆ 我们首先在引理 1 中刻画样本协方差矩阵$\mathbf{C}$ 相对于$\mathbf{C}$ 的扰动。为此，我们定义

$$\mathbf{E} \triangleq \hat{\mathbf{C}} - \mathbf{C} \ , \tag{38}$$

ˆ 并将$\mathbf{I}_m$ 定义为一个 $m \times m$ 单位矩阵。此外，样本协方差矩阵$\mathbf{C}$ 的特征值分解为

$$\hat{\mathbf{C}} = \hat{\mathbf{V}} \hat{\Lambda} \hat{\mathbf{V}}^{\mathsf{T}} \ , \tag{39}$$

ˆ ˆ ˆ 其中$\mathbf{V} =$ $[\hat{\mathbf{v}}_1, \ldots, \hat{\mathbf{v}}_m]$ 是由$\mathbf{C}$ 的正交归一特征向量构成的矩阵， Λ =ˆ ˆ ˆ ˆ ˆ ˆ diag($\lambda_1, \ldots, \lambda_m$) 是 $\mathbf{C}$ 的特征值对角矩阵，因此， $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_m$。显然，$\mathbf{C}$ 的特征值和特征向量是对真实协方差矩阵$\mathbf{C}$的特征值和特征向量的估计。

引理 2。 设总体协方差矩阵$\mathbf{C}$ 的特征分解如(3) 所示，样本ˆ协方差矩阵$\mathbf{C}$ 的特征分解如(39)所示。对于$\mathbf{C}$的任意特征值 $\lambda_i > 0$ ，扰动$\mathbf{E}$ 满足

$$\mathbf{E}\mathbf{v}_i = \beta_i \delta \mathbf{v}_i + \delta\lambda_i \mathbf{v}_i + (\delta\lambda_i \mathbf{I}_m - \mathbf{E})\delta\mathbf{v}_i \tag{40}$$

其中

$$\beta_i \triangleq (\lambda_i \mathbf{I}_m - \mathbf{C}), \quad \delta\mathbf{v}_i \triangleq \hat{\mathbf{v}}_i - \mathbf{v}_i, \quad \delta\lambda_i \triangleq \hat{\lambda}_i - \lambda_i \ . \tag{41}$$

证明。 注意，根据特征向量和特征值的定义，我们有

$$\hat{\mathbf{C}}\hat{\mathbf{v}}_i = \hat{\lambda}_i \hat{\mathbf{v}}_i \ . \tag{42}$$

我们可以将 (42) 重写为关于总体协方差矩阵$\mathbf{C}$ 的扰动及其特征分解输出的形式，如下所示:

$$(\hat{\mathbf{C}} - \mathbf{C})(\mathbf{v}_i + \delta\mathbf{v}_i) + \mathbf{C}(\mathbf{v}_i + \delta\mathbf{v}_i) = (\lambda_i + \delta\lambda_i)(\mathbf{v}_i + \delta\mathbf{v}_i) \ , \tag{43}$$

其中我们使用了 $\lambda_i = \lambda_i + \delta\lambda_i$ 和 $\hat{}\mathbf{v}_i = \mathbf{v}_i + \delta\mathbf{v}_i$。利用$\mathbf{C}\mathbf{v}_i = \lambda_i\mathbf{v}_i$ 这一事实，并重新排列式（43）中的各项，我们得到

$$(\hat{\mathbf{C}} - \mathbf{C})\mathbf{v}_i = (\lambda_i\mathbf{I}_m - \mathbf{C})\delta\mathbf{v}_i + \delta\lambda_i(\mathbf{v}_i + \delta\mathbf{v}_i) - (\hat{\mathbf{C}} - \mathbf{C})\delta\mathbf{v}_i \ . \tag{44}$$

$\hat{}$通过设定 $\mathbf{E} = \mathbf{C} - \mathbf{C}$ 以及 $\beta_i = \lambda_i\mathbf{I}_m - \mathbf{C}$，我们可以将 (44) 重写为

$$\mathbf{E}\mathbf{v}_i = \beta_i\delta\mathbf{v}_i + \delta\lambda_i\mathbf{v}_i + (\delta\lambda_i\mathbf{I}_m - \mathbf{E})\delta\mathbf{v}_i \ . \tag{45}$$

□

接下来，我们建立 $\hat{}\mathbf{C}^k$ 关于 $\mathbf{C}$ 和 $\mathbf{E}$的一阶近似。$\hat{}\mathbf{C}^k$ 的一阶近似为

$$(\mathbf{C} + \mathbf{E})^k = \mathbf{C}^k + \sum_{r=0}^{m} \mathbf{C}^r\mathbf{E}\mathbf{C}^{k-r-1} + \tilde{\mathbf{E}} \ , \tag{46}$$

其中 $\tilde{}\mathbf{E}$ 满足 $\| \tilde{}\mathbf{E}\| \leq \sum^{k} \binom{k}{r}\|\mathbf{E}\|^r\|\mathbf{C}\|^{k-r}$。利用 (46)，我们有$_{r=2}$

$$\mathbf{H}(\hat{\mathbf{C}}) - \mathbf{H}(\mathbf{C}) = \sum_{k=0}^{m} h_k[(\mathbf{C} + \mathbf{E})^k - \mathbf{C}^k] \ , \tag{47}$$

$$= \sum_{k=0}^{m} h_k \sum_{r=0}^{k-1} \mathbf{C}^r\mathbf{E}\mathbf{C}^{k-r-1} + \tilde{\mathbf{E}} \ , \tag{48}$$

其中 $\tilde{}\mathbf{E}$ 满足 $\| \tilde{}\mathbf{E}\|^2 = \mathcal{O}(\|\mathbf{E}\|^2)$ [38]。我们后续分析将聚焦于 (48) 中的第一项。对于随机数据样本 $\mathbf{x} = [x_1, \ldots, x_m]^\mathsf{T}$，满足 $\|\mathbf{x}\|_2 \leq 1$和 $\mathbf{x} \in \mathbb{R}^{m\times 1}$，其关于 $\mathbf{C}$的傅里叶变换为 $\tilde{}\mathbf{x} = \mathbf{V}^\mathsf{T}\mathbf{x}$，其中 $\tilde{}\mathbf{x} = [\tilde{x}_1, \ldots, \tilde{x}_m]^\mathsf{T}$ [1]。 $\tilde{}\mathbf{x}$与 $\mathbf{x}$之间的关系可表示为

$$\mathbf{x} = \sum_{i=1}^{m} \tilde{x}_i\mathbf{v}_i \ . \tag{49}$$

将（48）式两边同时乘以$\mathbf{x}$，并利用（49）式，可得

$$[\mathbf{H}(\hat{\mathbf{C}}) - \mathbf{H}(\mathbf{C})]\mathbf{x} = \sum_{k=0}^{m} h_k \sum_{r=0}^{k-1} \mathbf{C}^r\mathbf{E}\mathbf{C}^{k-r-1}\mathbf{x} + \tilde{\mathbf{E}}\mathbf{x} \ , \tag{50}$$

$$= \sum_{i=1}^{m} \tilde{x}_i \sum_{k=0}^{m} h_k \sum_{r=0}^{k=1} \mathbf{C}^r\mathbf{E}\mathbf{C}^{k-r-1}\mathbf{v}_i + \tilde{\mathbf{E}}\mathbf{x} \ , \tag{51}$$

$$= \sum_{i=1}^{m} \tilde{x}_i \sum_{k=0}^{m} h_k \sum_{r=0}^{k-1} \mathbf{C}^r\lambda_i^{k-r-1}\mathbf{E}\mathbf{v}_i + \tilde{\mathbf{E}}\mathbf{x} \ , \tag{52}$$

其中，我们在从 (51) 到 (52) 的推导中使用了 **协方差**$_i = \lambda_i \mathbf{v}_i$。我们只关注 (52) 中的第一项，并利用引理 1 的结果展开$\mathbf{E}\mathbf{v}_i$，得到

$$\begin{aligned}\sum_{i=1}^{m}\tilde{x}_i\sum_{k=0}^{m}h_k\sum_{r=0}^{k-1}\mathbf{C}^r\lambda_i^{k-r-1}\mathbf{E}\mathbf{v}_i = &\underbrace{\sum_{i=1}^{m}\tilde{x}_i\sum_{k=0}^{m}h_k\sum_{r=0}^{k-1}\mathbf{C}^r\lambda_i^{k-r-1}\beta_i\delta\mathbf{v}_i}_{\text{Term 1}}\\ &+\underbrace{\sum_{i=1}^{m}\tilde{x}_i\sum_{k=0}^{m}h_k\sum_{r=0}^{k-1}\mathbf{C}^r\lambda_i^{k-r-1}\delta\lambda_i\mathbf{v}_i}_{\text{Term 2}}\\ &+\underbrace{\sum_{i=1}^{m}\tilde{x}_i\sum_{k=0}^{m}h_k\sum_{r=0}^{k-1}\mathbf{C}^r\lambda_i^{k-r-1}(\delta\lambda_i\mathbf{I}_m-\mathbf{E})\delta\mathbf{v}_i}_{\text{Term 3}}\ . \end{aligned}\tag{53}$$

接下来，我们分别分析（53）中的项 1、项 2 和项 3。

项 1 的分析（53）。在分析项 1 时，我们首先注意到

$$\beta_i = \lambda_i\mathbf{I}_m - \mathbf{C}\ , \tag{54}$$

$$= \sum_{j=1}^{m}(\lambda_i-\lambda_j)\mathbf{v}_j\mathbf{v}_j^{\mathsf{T}}\ , \tag{55}$$

$$= \mathbf{V}(\lambda_i\mathbf{I}_m-\Lambda)\mathbf{V}^{\mathsf{T}}\ . \tag{56}$$

利用（56）以及项 1 中的 $\delta\mathbf{v}_i = \mathbf{u}_i - \mathbf{v}_i$ （53），可得

$$\sum_{i=1}^{m}\tilde{x}_i\sum_{k=0}^{m}h_k\sum_{r=0}^{k-1}\mathbf{C}^r\lambda_i^{k-r-1}\mathbf{V}(\lambda_i\mathbf{I}_m-\Lambda)\mathbf{V}^{\mathsf{T}}(\hat{\mathbf{v}}_i-\mathbf{v}_i)\ . \tag{57}$$

利用（57）中的$\mathbf{C}^r = \mathbf{V}\Lambda^r\mathbf{V}^{\mathsf{T}}$，项 1（53）等价于

$$\sum_{i=1}^{m}\tilde{x}_i\sum_{k=0}^{m}h_k\sum_{r=0}^{k-1}\lambda_i^{k-r-1}\mathbf{V}\Lambda^r(\lambda_i\mathbf{I}_m-\Lambda)\mathbf{V}^{\mathsf{T}}(\hat{\mathbf{v}}_i-\mathbf{v}_i)\ , \tag{58}$$

$$= \sum_{i=1}^{m}\tilde{x}_i\mathbf{V}\mathbf{L}_i\mathbf{V}^{\mathsf{T}}(\hat{\mathbf{v}}_i-\mathbf{v}_i)\ , \tag{59}$$

其中$\mathbf{L}_i$ 是一个对角矩阵，其 j项为

$$[\mathbf{L}_i]_j = \sum_{k=0}^{m}h_k\sum_{r=0}^{k-1}(\lambda_i-\lambda_j)\lambda_i^{k-r-1}\lambda_j^r\ , \tag{60}$$

$$= \sum_{k=0}^{m}h_k(\lambda_i-\lambda_j)\frac{\lambda_i^k-\lambda_j^k}{\lambda_i-\lambda_j}\ , \tag{61}$$

$$= \sum_{k=0}^{m}h_k\lambda_i^k - \sum_{k=0}^{m}h_k\lambda_j^k\ , \tag{62}$$

$$= h(\lambda_i) - h(\lambda_j)\ , \tag{63}$$

其中， $h(\lambda_i)$ 表示协方差滤波器的频率响应，其定义见式 (10)。因此，我们有$\mathbf{L}_i = \mathsf{diag}([h(\lambda_i) - h(\lambda_j)]_j)$。接下来，在式 (59) 中，我们注意到

$$\mathbf{V}^{\mathsf{T}}(\hat{\mathbf{v}}_i - \mathbf{v}_i) = [\mathbf{v}_1^{\mathsf{T}}(\hat{\mathbf{v}}_i - \mathbf{v}_i), \cdots, \mathbf{v}_m^{\mathsf{T}}(\hat{\mathbf{v}}_i - \mathbf{v}_i)]^{\mathsf{T}} . \tag{64}$$

将 (64) 和 (63) 代入 (59) 以及 $\mathbf{v}_j^{\mathsf{T}}\mathbf{v}_i = 0, \ \forall j \neq i$，我们推得 (53) 中的项 1 等价于

$$\sum_{i=1}^{m} \tilde{x}_i \mathbf{V}\mathbf{L}_i\mathbf{V}^{\mathsf{T}}(\hat{\mathbf{v}}_i - \mathbf{v}_i) = \sum_{i=1}^{m} \tilde{x}_i \mathbf{V}\mathbf{J}_i , \tag{65}$$

其中$\mathbf{J}_i$ 的第 j个元素为

$$[\mathbf{J}_i]_j = \begin{cases} 0 , \quad \text{if} \quad j = i , \\ (h(\lambda_i) - h(\lambda_j))\mathbf{v}_j^{\mathsf{T}}\hat{\mathbf{v}}_i , \ \text{otherwise} \end{cases} . \tag{66}$$

对于稳定性分析，我们关注项 1 的范数。因此，注意到式 (53) 和式 (65) 中的项 1 等价，在取范数后，我们得到

$$\left\| \sum_{i=1}^{m} \tilde{x}_i \sum_{k=0}^{m} h_k \sum_{r=0}^{k-1} \mathbf{C}^r \lambda_i^{k-r-1} \beta_i \delta \mathbf{v}_i \right\|_2 = \left\| \sum_{i=1}^{m} \tilde{x}_i \mathbf{V}\mathbf{J}_i \right\|_2 , \tag{67}$$

$$\leq \sqrt{m} \sum_{i=1}^{m} |\tilde{x}_i| \max_{j, i\neq j} |h(\lambda_i) - h(\lambda_j)| |\mathbf{v}_j^{\mathsf{T}}\hat{\mathbf{v}}_i| . \tag{68}$$

注意，$\mathbf{v}_j^{\mathsf{T}}\hat{}\mathbf{v}_i$ 是总体协方差矩阵的特征向量$\mathbf{v}j$ 与$\mathbf{C}$ 样本协方差矩阵的特征向量 $\hat{}\mathbf{v}_i$ 之间的内积 $\hat{}\mathbf{C}$。关于$\mathbf{v}_j^{\mathsf{T}}\hat{}\mathbf{v}_i$ 的界（以数据样本数量 n 为变量）已在现有文献中得到研究。在此，我们利用 [78, 定理 4.1]$\hat{}$ 的结果来得出：如果 $\mathsf{sgn}(\lambda_j - \lambda_i)2\lambda_j > \mathsf{sgn}(\lambda_j - \lambda_i)(\lambda_j - \lambda_i)$ 对于 $\lambda_i \neq \lambda_j$，则条件

$$\left\| \sum_{i=1}^{m} \tilde{x}_i \sum_{k=0}^{m} h_k \sum_{r=0}^{k-1} \mathbf{C}^r \lambda_i^{k-r-1} \beta_i \delta \mathbf{v}_i \right\|_2 \leq \sqrt{m} \sum_{i=1}^{m} |\tilde{x}_i| \max_{j, i\neq j} |h(\lambda_i) - h(\lambda_j)| \frac{2k_i}{n^{1/2-\varepsilon}|\lambda_i - \lambda_j|} , \tag{69}$$

对于某个 $\varepsilon \in$，当 0， $1/2]$ 时，该结论以至少 $\left(1 - \frac{1}{n^{2\varepsilon}}\right)$ 的概率成立，其中 $k_i = \left(\mathbb{E}[\|\mathbf{X}\mathbf{X}^{\mathsf{T}}\mathbf{v}_i\|_2^2] - \lambda_i^2\right)^{\frac{1}{2}}$。此外，我们注意到，条件 $\mathsf{sgn}(\lambda_j - \lambda_i)2\lambda_j > \mathsf{sgn}(\lambda_j - \lambda_i)(\lambda_j - \lambda_i)$ 以至少 $1 - \frac{2k_i^2}{|\lambda_i - \lambda_j|}$ [78, Corollary 4.2], 的概率成立。通过并集界和泰勒级数的一阶近似，这意味着 (69) 以至少 $1 - \frac{1}{n^{2\varepsilon}} - \frac{2\kappa m}{n}$ 的概率成立，其中 κ 定义为

$$\kappa \triangleq \max_{i,j:\lambda_i\neq\lambda_j} \frac{k_i^2}{|\lambda_i - \lambda_j|} \quad \text{where} \quad k_{\mathsf{min}} \triangleq \min_{i\in\{1,\ldots,m\}, \lambda_i>0} k_i . \tag{70}$$

因此，对于具有以下性质的协方差滤波器

$$\frac{|h(\lambda_i) - h(\lambda_j)|}{|\lambda_i - \lambda_j|} \leq \frac{M}{k_i} , \tag{71}$$

对于某个实常数 $M > 0$，式（69）中的条件等价于

$$\left\| \sum_{i=1}^{m} \tilde{x}_i \sum_{k=0}^{m} h_k \sum_{r=0}^{k-1} \mathbf{C}^r \lambda_i^{k-r-1} \beta_i \delta \mathbf{v}_i \right\|_2 \leq \frac{2\sqrt{m}M}{n^{\frac{1}{2}-\varepsilon}} \sum_{i=1}^{m} |\tilde{x}_i| \,, \tag{72}$$

m该结论以至少 $1 - \frac{1}{n^{2\varepsilon}} - \frac{2\kappa m}{n}$ 的概率成立。此外，注意到 $\sum_{i=1} |\tilde{x}_i| \leq \sqrt{m}\|\mathbf{x}\|_{2}$。当随机样本 $\mathbf{x}$ 满足 $\|\mathbf{x}\|_2 \leq 1$时，我们有

$$\mathbb{P}\left(\left\| \sum_{i=1}^{m} \tilde{x}_i \sum_{k=0}^{m} h_k \sum_{r=0}^{k-1} \mathbf{C}^r \lambda_i^{k-r-1} \beta_i \delta \mathbf{v}_i \right\|_2 \leq \frac{2}{n^{\frac{1}{2}-\varepsilon}} mM \right) \geq 1 - \frac{1}{n^{2\varepsilon}} - \frac{2\kappa m}{n} \,, \tag{73}$$

对于任意 $\varepsilon \in (0,\ 1/2]$。

An对式 (53) 中项 2 的分析。利用 $\mathbf{C}\mathbf{v}_i = \lambda_i \mathbf{v}_i$，注意到式 (53) 中的项 2 等价于　to

$$\sum_{i=1}^{m} \tilde{x}_i \sum_{k=0}^{m} h_k \sum_{r=0}^{k-1} \mathbf{C}^r \lambda_i^{k-r-1} \delta\lambda_i \mathbf{v}_i = \sum_{i=1}^{m} \tilde{x}_i \sum_{k=0}^{m} h_k \sum_{r=0}^{k-1} \lambda_i^{k-1} \delta\lambda_i \mathbf{v}_i \,, \tag{74}$$

$$= \sum_{i=1}^{m} \tilde{x}_i \sum_{k=0}^{m} k h_k \lambda_i^{k-1} \delta\lambda_i \mathbf{v}_i \,, \tag{75}$$

$$= \sum_{i=1}^{m} \tilde{x}_i h'(\lambda_i) \delta\lambda_i \mathbf{v}_i \,. \tag{76}$$

接下来，利用 Weyl 定理 [115, 定理 8.1。6], 我们注意到 $\|\mathbf{E}\| \leq \alpha$ 意味着 $|\delta\lambda_i| \leq \alpha$ 对任意$\alpha > 0$ 均成立。对于随机向量 $\mathbf{X}$ 的一个随机实例 $\mathbf{x}$，其概率分布支撑在半径为 1 的球内（不失一般性），且满足 $\|\mathbf{x}\|_2 \leq 1$，我们有

$$\mathbb{P}\left(\mathbf{E} \leq B \left(\frac{\|\mathbf{C}\|\sqrt{\log m + u}}{\sqrt{n}} + \frac{(1 + \|\mathbf{C}\|)(\log m + u)}{n} \right) \right) \geq 1 - 2^{-u} \,, \tag{77}$$

对于某些常数 $B > 0$ 和 $u > 0$。式 (77) 中的结果可直接由 [116, 定理 5.6.1] 得出。因此，利用式 (76)，我们有

$$\left\| \sum_{i=1}^{m} \tilde{x}_i \sum_{k=0}^{m} h_k \sum_{r=0}^{k-1} \mathbf{C}^r \lambda_i^{k-r-1} \delta\lambda_i \mathbf{v}_i \right\|_2 \leq \sum_{i=1}^{m} |\tilde{x}_i| |h'(\lambda_i)| |\delta\lambda_i| \|\mathbf{v}_i\|_2 \,. \tag{78}$$

使用（77），$|h'(\lambda_i)| \leq M/k_{\min}$ (其中 $k_{\min} = \min_{i \in \{1,\ldots,m\},\ \lambda_i > 0} k_i$) 来自 (71)，并且 $\|\mathbf{v}_i\|_2 = 1$，我们有　e

$$\mathbb{P}\left(\left\| \sum_{i=1}^{m} \tilde{x}_i \sum_{k=0}^{m} h_k \sum_{r=0}^{k-1} \mathbf{C}^r \lambda_i^{k-r-1} \delta\lambda_i \mathbf{v}_i \right\|_2 \right.$$
$$\left. \leq \frac{A}{k_{\min}} \sqrt{m} M \left(\frac{\|\mathbf{C}\|\sqrt{\log m + u}}{\sqrt{n}} + \frac{(1 + \|\mathbf{C}\|)(\log m + u)}{n} \right) \right) \geq 1 - 2^{-u} \,, \tag{79}$$

对于某些常数 $A > 0$ 和 $u > 0$。

对（53）中项 3 的分析（53）。我们指出，（53）中的项 3 由二阶误差项组成，其随样本数 n 增加而衰减的速度快于项 1 和项 2。这可以得出

以下观察。首先，根据（77）和 Weyl 定理，我们注意到，以高概率有 $\|\delta\lambda_i\mathbf{I}_m - \mathbf{E}\| \le 2\|\mathbf{E}\|$以及 $\|\mathbf{E}\| \simeq \mathcal{O}(1/\sqrt{n})$。其次，项 $\delta\mathbf{v}_i$ 的上界也具有与$\mathbf{E}$ [117]类似的缩放行为。

因此，二阶误差项 $(\delta\lambda_i\mathbf{I}_m - \mathbf{E})\delta\mathbf{v}_i$ 的衰减速度快于 $\mathcal{O}(1/\sqrt{n})$，其速度也快于项 1 和项 2；后两者分别按 $\mathcal{O}(1/n^{1/2-\varepsilon})$ 的速率缩放，其中 $\varepsilon \in (0, 1/2]$ 和 $\mathcal{O}(1/\sqrt{n})$ 分别对应这两项。最后，注意到关于 $\|[\mathbf{H}(\mathbf{C}) - \mathbf{H}(\mathbf{C})]\mathbf{x}\|_2$ 的条件可归结为搜索空间 $\max_{\mathbf{x}\in\mathbb{R}^{m\times 1}, \|\mathbf{x}\|_2\le 1}\{\|[\mathbf{H}(\mathbf{C}) - \mathbf{H}(\mathbf{C})]\mathbf{x}\|_2\}$ 上的算子范数条件 $\|[\mathbf{H}(\mathbf{C}) - \mathbf{H}(\mathbf{C})]\|$，并且 (73) 和 (79) 中按 $1/\sqrt{n}$ 或更慢速率缩放的项主导了 $\|\mathbf{H}(\mathbf{C}) - \mathbf{H}(\mathbf{C})\|$ 上界的缩放行为，由此得到以下定理。

定理4（协方差滤波器的稳定性）。考虑一个随机向量 $\mathbf{X} \in \mathbb{R}^{m\times 1}$，其对应的协方差矩阵为 $\mathbf{C} = \mathbb{E}[(\mathbf{X} - \mathbb{E}[\mathbf{X}])(\mathbf{X} - \mathbb{E}[\mathbf{X}])^{\mathsf{T}}]$。对于使用 n 个独立同分布的$\mathbf{X}$ 实例构造的样本协方差 $\hat{}$ 矩阵 $\mathbf{C}$，以及$\mathbf{X}$的一个随机实例$\mathbf{x}$，满足 $\|\mathbf{x}\|_2 \le 1$，并且在假设(71)下，对任意$\varepsilon \in (0, 1/2]$，以下结论以至少$(1 - n^{-2\varepsilon} - 2\kappa m/n)(1 - 1/n)$ 的概率成立

$$\left\|\mathbf{H}(\hat{\mathbf{C}}) - \mathbf{H}(\mathbf{C})\right\| = \frac{M}{n^{\frac{1}{2}-\varepsilon}} \cdot \mathcal{O}\left(m + \frac{\sqrt{m}\|\mathbf{C}\|\sqrt{\log mn}}{k_{\min} n^{\varepsilon}}\right). \tag{80}$$

因此，(80) 的右侧给出了定理2中 α_n 的结构。根据定理4，我们注意到$\|\mathbf{H}(\hat{\mathbf{C}}) - \mathbf{H}(\mathbf{C})\|$ 会随着样本数量 n 的增加而衰减，且衰减速率至少为 $1/n^{\frac{1}{2}-\varepsilon}$。因此，我们得出结论：随着样本数量 n 增加，协方差滤波器的稳定性会提高。这一观察符合预期，因为根据大数定律，估计值$\mathbf{C}$ 会逐渐接近总体协方差矩阵$\mathbf{C}$ 。关于假设 (71) 对滤波器设计的影响，读者可参见 [1] 以了解更多讨论。证明的其余部分可直接由 [1, 定理 3] 得出。

附录C 图子信息处理

此前，图子理论已被用于研究同一图子族中图之间图神经网络的可迁移性 [39]。定理3的证明依赖于这样一个性质：当不同数据集所对应的协方差矩阵属于由图子刻画的收敛序列时，VNN在这些数据集之间具有可迁移性。本节的主要目标是说明，协方差滤波器上的数据处理可以等价地通过其图子近似在连续域中表示。确立这一性质最终将使我们能够比较基于来自特征数量不同的数据集的协方差矩阵实例化的VNN。我们先介绍一些关于图子的预备知识。

C.1 预备知识。

下面重新陈述图子的定义。

定义 5 (图子**)**。 图子是一个有界、对称且可测的函数 $\mathbf{W}:\ [0,1]^2 \mapsto [-1,1]$。

利用图子的收敛理论，并将协方差矩阵解释为数据的加权图表示，如果协方差矩阵序列 $\{\mathbf{C}_m\}$ 在割距离 [80]下收敛，那么存在一个图子$\mathbf{W}$ 作为图子近似序列$\{\mathbf{W}_{\mathbf{C}_m}\}$ 的极限对象。在备注 2 中，我们形式化地给出了对于给定协方差矩阵序列，极限对象存在的充分条件。备注 2 中的陈述将 [80, 推论 3.9] 扩展到了我们的设定，即将协方差矩阵视为加权图。

备注 2 （图子作为极限对象 [80]）。 **</style}若协方差矩阵序列** $\mathbf{C}_m\}$ 相对于割距离 $\delta_{\square}$构成柯西序列，则称该序列收敛，其中协方差矩阵$\mathbf{C}m_1$ 与$\mathbf{C}m_2$ 之间的割距离（$\mathbf{C}m_1$，$\mathbf{C}m_2$） $\delta_{\square}$在附录 *D* 的（93）中定义。此外，对于任意收敛的协方差矩阵序列 $\{\mathbf{C}_m\}$，相应的图子近似序列$\{\mathbf{W}_{\mathbf{C}_m}\}$ 收敛到一个图子。

割距离的一个显著特征是，它允许比较不同大小的协方差矩阵。因此，所有图子近似收敛到某个图子的协方差矩阵，都可以视为该图子族的一部分。此外，图子$\mathbf{W}$ 可以解释为一种模式，而协方差矩阵$\mathbf{C}m$ 则表示该模式在分辨率 m下的协方差实现。

C.2 使用图子进行信息处理

接下来，我们将证明，协方差滤波器 $\mathbf{H(C}m\mathbf{)}$ 可以等价地表示为连续域中基于图子表示 $\mathbf{W}_{\mathbf{C}m}$ 的卷积运算。给定协方差滤波器的输出$\mathbf{z}\ =\mathbf{H(C}m\mathbf{)x}$，$\mathbf{x}$ 的连续表示为 $y_{\mathbf{x}}$，而 $\mathbf{C}m$ 的连续表示为 $\mathbf{W}_{\mathbf{C}m}$。运算 $\mathbf{Cx}$ 是 $\mathbf{H(C}m\mathbf{)x}$ 中卷积运算的基础，因此，我们首先给出其对应的连续形式。对于 $\mathbf{s}\ =\mathbf{Cx}$，$\mathbf{s}$ 的第 i 个元素定义为

$$[\mathbf{s}]_i = \sum_{j=0}^{m} [\mathbf{C}_m]_{ij}[\mathbf{x}]_j \ . \tag{81}$$

因此，根据$\mathbf{C}m$的第 i行，$[\mathbf{s}]_i$ 是$\mathbf{x}$ 中元素的线性组合。在连续空间中，我们也可以将（81）等价地写为

$$y_{\mathbf{s}}(u) = \int_0^1 \mathbf{W}_{\mathbf{C}_m}(u,v) y_{\mathbf{x}}(v) \mathrm{d}v \ , \tag{82}$$

其中，$y_{\mathbf{x}}$ 是根据（20）中定义的区间获得的$\mathbf{x}$ 的连续表示。注意，$y_{\mathbf{s}}$ 是$\mathbf{s}$的连续表示，即它们满足 $y_{\mathbf{s}}(u)\ =[\mathbf{s}]$，其中 $u \in \mathcal{U}_i$。因此，$y_{\mathbf{s}}$ 和$\mathbf{s}$ 可以相互恢复。这一观察可以进一步推广，以定义.....的连续等价形式。

一个协方差滤波器。这是可行的，因为我们可以将实体$\mathbf{C}_m^k\mathbf{x}$ 在 $\mathbf{H}(\mathbf{C})$ 中写成递归形式。具体而言，如果我们有$\mathbf{s}_k = \mathbf{C}_m^k\mathbf{x}$，那么我们可以将$\mathbf{s}_k$ 重写为

$$\mathbf{s}_k = \mathbf{C}_m\mathbf{s}_{k-1} \ , \tag{83}$$

其中 $\mathbf{s}_0 = \mathbf{x}$。因此，沿用证明 (81) 与 (82) 等价时的相同推理，我们得出，$\mathbf{s}_k$ 的连续表示 $y_{\mathbf{s}_k}$ 可通过以下运算恢复

$$y_{\mathbf{s}_k}(u) = \int_0^1 \mathbf{W}_{\mathbf{C}_m}(u,v)y_{\mathbf{s}_{k-1}}(v)\mathsf{d}v \ . \tag{84}$$

由于协方差滤波器的输出 $\mathbf{z}$ 是各项 $\mathbf{s}_k$的加权聚合，我们可以将其连续表示 $y_{\mathbf{z}}$ 写为

$$y_{\mathbf{z}}(u) = \sum_{k=0}^{K} h_k y_{\mathbf{s}_k}(u) \ . \tag{85}$$

通过推导 (85) 所涉及的数学步骤，我们已经证明，协方差滤波器输出 $\mathbf{z}$ 的连续表示可以通过对 (82) 和 (84) 中的图子表示$\mathbf{W}_{\mathbf{C}_m}$ 执行卷积运算来恢复。此外，$\mathbf{z}$ 与 $y_{\mathbf{z}}$ 在操作上可以互换。此外，我们还可以将这种对应关系推广到具有多层架构和 MIMO信息处理的协方差感知器及 VNN。将这一观察推广到协方差感知器和基础 VNN 是显而易见的，因为协方差输出是在对$\mathbf{z}$ 应用逐点非线性 σ 后计算得出的，而基础 VNN 则由多个协方差感知器堆叠而成，并且每层的输入和输出数量（即 F ）均设为 $\mathbf{1}$。

我们使用记号 $\mathbf{x}_m$ 表示具有 m 个特征的输入向量。因此，若体素神经网络输出 $\Phi(\mathbf{x}_m; \mathbf{C}_m, \mathcal{H})$ 的大小为 $m \times 1$，且我们有 $F = 1$ 和层数 L，则其连续近似 $y_{\Phi(\mathbf{x}_m;\mathbf{C}_m,\mathcal{H})}$ 可通过在 $\mathbf{W}_{\mathbf{C}_m}$ 上实例化、以 $y_{\mathbf{x}_m}$为输入的卷积架构恢复。对于采用 MIMO处理的 VNN，每个 VNN 层都有多个 m维输入和多个 m维输出。因此，我们也可以将能够执行 MIMO处理、在 $\mathbf{W}_{\mathbf{C}_m}$ 和 $\mathbf{x}_m$ 上实例化并以多个连续表示作为输出的架构等价地定义出来。这种架构此前已以图子神经网络的形式进行过研究 [79]。在此背景下，我们定义模型$\tilde{\Phi}(y_{\mathbf{x}_m}; \mathbf{W}_{\mathbf{C}_m}, \mathcal{H})$，该模型通过式 (84) 中对 $\mathbf{W}_{\mathbf{C}_m}$ 执行卷积运算来建模，并具有与 VNN $\Phi(\mathbf{x}_m; \mathbf{C}_m, \mathcal{H})$ 相同的架构。请注意，$\tilde{\Phi}(y_{\mathbf{x}_m}; \mathbf{W}_{\mathbf{C}_m}, \mathcal{H})$ 的输出是 VNN $\Phi(\mathbf{x}_m; \mathbf{C}_m, \mathcal{H})$ 输出的连续表示（另见图 3 中的示意图）。因此，我们可以通过分析 $\tilde{\Phi}(y_{\mathbf{x}_{m_1}}; \mathbf{W}_{\mathbf{C}_{m_1}}, \mathcal{H})$ 与 $\tilde{\Phi}(y_{\mathbf{x}_{m_2}}; \mathbf{W}_{\mathbf{C}_{m_2}}, \mathcal{H})$ 之间的差异，来研究实例化于协方差矩阵 $\mathbf{C}_{m_1}$ 和 $\mathbf{C}_{m_2}$ 上的 VNN 之间参数 $\mathcal{H}$ 的可迁移性。

在此背景下，我们的分析依赖于这样一种设定：图子近似$\mathbf{W}_{\mathbf{C}_{m_1}}$ 和$\mathbf{W}_{\mathbf{C}_{m_2}}$ 属于图子近似序列 $\{\mathbf{W}_{\mathbf{C}_m}\}$，该序列收敛于图子$\mathbf{W}$。因此，~我们还考虑在图子$\mathbf{W}$上实例化的信息处理架构 $\Phi(y;\mathbf{W} \ \mathcal{H})$，使得

$\tilde{y}$ 且连续表示 $y_{\mathbf{x}_m}$ 始终满足 $y_{\mathbf{x}_m}(\rho_i) = \tilde{y}(\rho_i)\ \forall i \in \{1,\dots,m\}$。在这里，我们还可以将 $\tilde{\Phi}(y; \mathbf{W}\ \mathcal{H})$ 解释为生成模型，其中 $\Phi(y_{\mathbf{x}_m}; \mathbf{W}_{\mathbf{C}_m},\ \mathcal{H})$ 是 $\Phi(y; \mathbf{W}\ \mathcal{H})$ 在分辨率 m 下的一个实例。因此，我们对 VNN 可迁移性的分析也包括研究 $\Phi(y_{\mathbf{x}_m}; \mathbf{W}_{\mathbf{C}_m},\ \mathcal{H})$ 的输出与 $\Phi(y;\mathbf{W}\ \mathcal{H})$ 的输出之间的收敛性。

为此，我们现在正式定义图子上的卷积滤波器，并刻画其频率响应。我们将对$\mathbf{W}_{\mathbf{C}_m}$ 和连续表示$y_{\mathbf{x}_m}$进行 k跳聚合（类似于$\mathbf{C}^k\mathbf{x}$）的算子记为 $T^k{}_{\mathbf{W}_{\mathbf{C}_m}} y_{\mathbf{x}_m}$，其定义如下：

$$(T^k_{\mathbf{W}_{\mathbf{C}_m}} y_{\mathbf{x}_m})(u) \triangleq \int_0^1 \mathbf{W}_{\mathbf{C}_m}(u,v)(T^{k-1}_{\mathbf{W}_{\mathbf{C}_m}} y_{\mathbf{x}_m})(v)\mathsf{d}v\ , \tag{86}$$

对任意 $k > 1$，其中

$$(T_{\mathbf{W}_{\mathbf{C}_m}} y_{\mathbf{x}_m})(u) \triangleq \int_0^1 \mathbf{W}_{\mathbf{C}_m}(u,v) y_{\mathbf{x}_m}(v)\mathsf{d}v\ . \tag{87}$$

因此，根据上述讨论，$T^k{}_{\mathbf{W}_{\mathbf{C}_m}} y_{\mathbf{x}_m}$和$\mathbf{C}^k_m\mathbf{x}_m$在操作上可以互换。当 y 与 $y_{\mathbf{x}_m}$ 满足 $y_{\mathbf{x}_m}(\rho_i) = y(\rho_i)$ 的关系时，我们也可以使用算子 $T_{\mathbf{W}}y$ 定义$\mathbf{W}$ 上的 k跳聚合，其中 ρ_i 在式（20）中定义。因此，图子$\mathbf{W}$ 和连续表示 y 可以视为协方差矩阵$\mathbf{C}_m$和数据点$\mathbf{x}_m$的生成模型。这一观察与图和图子背景下的相关观察 [79]相呼应。我们将一组滤波器抽头 $\mathcal{H} = \{h_k\}_{k=0}^K$ 对应的图子滤波器记为$\Psi(y; \mathbf{W}\ \mathcal{H})$: $[0,1] \to \mathbb{R}$，其定义如下：

$$\Psi(y; \mathbf{W}, \mathcal{H})(u) \triangleq \sum_{k=0}^K h_k (T^k_{\mathbf{W}} y)(u)\ . \tag{88}$$

与协方差滤波器类似，我们可以通过对 (88) 中的$\mathbf{W}$ 进行特征分解，来刻画图子滤波器的频率响应。由于$\mathbf{W}$ 有界且对称，我们可以将$\mathbf{W}$ 的谱分解表示为

$$\mathbf{W}(u,v) = \sum_{i\in\mathbb{Z}\backslash\{0\}} \eta_i \Gamma_i(u)\Gamma_i(v)\ , \tag{89}$$

其中 $\eta_i, \forall i \in \mathbb{Z}\backslash\{0\}$ 是特征值，Γ_i 是$\mathbf{W}$的特征信号。因此，我们可以将 (88) 重写为

$$\Psi(y; \mathbf{W}, \mathcal{H})(u) = \sum_{i\in\mathbb{Z}\backslash\{0\}} \sum_{k=0}^K h_k \eta_i^k \Gamma_i(u) \int_0^1 \Gamma_i(v) y(v)\mathsf{d}v\ , \tag{90}$$

$$= \sum_{i\in\mathbb{Z}\backslash\{0\}} \tilde{h}(\eta_i)\Gamma_i(u) \int_0^1 \Gamma_i(v) y(v)\mathsf{d}v\ , \tag{91}$$

对于 $u \in [0,1]$。请注意，(90) 是利用 (89) 和 (86) 从 (88) 推导得出的，并且我们在 (91) 中使用了定义 $\tilde{h}(\eta) \triangleq \sum_{k=0}^K h_k \eta^k$。项 $\tilde{h}(\eta_i)$ 表征了图子滤波器的频率响应，并取决于滤波器抽头 $\{h_k\}$ 和图子特征值。附录 E 中对 $\|\Psi(y_{\mathbf{x}_m}; \mathbf{W}_{\mathbf{C}_m}, \mathcal{H}) - \Psi(y; \mathbf{W}\ \mathcal{H})\|_2$ 的分析表明，图子滤波器近似 $\Psi(y_{\mathbf{x}}; \mathbf{W}_{\mathbf{C}_m}, \mathcal{H})$ 与图子滤波器 $\Psi(y;\mathbf{W}\ \mathcal{H})$ 之间的接近程度与维度 m 成反比。

附录 D 割距离

此处，我们借鉴文献 [80] 中加权图的定义，来定义两个协方差矩阵 $\mathbf{C}_{m_1}$ 和 $\mathbf{C}_{m_2}$ 之间的割距离，其中 $m_1 \neq m_2$。首先，对 $\mathbf{C}_{m_1}$ 和 $\mathbf{C}_{m_2}$ 进行归一化，使得 $\mathsf{tr}(\mathbf{C}_{m_1}) = 1$且 $\mathsf{tr}(\mathbf{C}_{m_2}) = 1$。接下来，我们定义一族大小为 $m_1 \times m_2$ 的矩阵 $\mathcal{K}(\mathbf{C}_{m_1},\mathbf{C}_{m_2})$，使得对于其中的每个元素 $\mathbf{K} \in \mathcal{K}(\mathbf{C}_{m_1},\mathbf{C}_{m_2})$，都有

$$\sum_{j=1}^{m_2}[\mathbf{K}]_{ij} = [\mathbf{C}_{m_1}]_{ii} \quad \text{and} \quad \sum_{i=1}^{m_1}[\mathbf{K}]_{ij} = [\mathbf{C}_{m_2}]_{jj} \; . \tag{92}$$

族 $\mathcal{K}(\mathsf{C}_{m_1}, \mathsf{C}_{m_2})$ 称为“分数叠加”，它在概念上定义了 C_{m_1} 与 C_{m_2} 各个特征之间的映射。由于协方差矩阵提供了一种以特征为节点的加权图表示，分数叠加描述了两个节点数量不同的加权图中节点之间的重叠。接下来，利用分数叠加 $\mathcal{K}(\mathsf{C}_{m_1}, \mathsf{C}_{m_2})$ 中的一个成员 K，我们在 $[m_1] \times [m_2]$ 中的成对元素集合上定义矩阵 $\mathsf{C}_{m_1}[\mathsf{K}]$ 和 $\mathsf{C}_{m_2}[\mathsf{K}^\mathsf{T}]$，其中$[m_1] = [1, \ldots, m_1]$。因此，矩阵 $\mathsf{C}_{m_1}[\mathsf{K}]$ 和 $\mathsf{C}_{m_2}[\mathsf{K}]$ 的大小为 $m_1 m_2 \times m_1 m_2$。在 $\mathsf{C}_{m_1}[\mathsf{K}]$和 $\mathsf{C}_{m_2}[\mathsf{K}^\mathsf{T}]$中，与 $[m_1] \times [m_2]$ 中元素 (i, j) 对应的对角元素的权重均为 $[\mathsf{K}]_{ij}$。此外，对于 $i, j \in [m_1]$ 和 $a, b \in [m_2]$，与配对元素 $((i, a), (j, b))$ 对应的非对角元素，在 $\mathsf{C}_{m_1}[\mathsf{K}]$ 中为 $[\mathsf{C}_{m_1}]_{ij}$，在 $\mathsf{C}_{m_2}[\mathsf{K}^\mathsf{T}]$中为 $[\mathsf{C}_{m_2}]_{ab}$。由于由 C_{m_1} 和 C_{m_2} 得到的 $\mathsf{C}_{m_1}[\mathsf{K}]$ 和 $\mathsf{C}_{m_2}[\mathsf{K}^\mathsf{T}]$ 大小相同，因此可以直接定义它们之间的距离。C_{m_1} 与 C_{m_2} 之间的割距离等价于 $\mathsf{C}_{m_1}[\mathsf{K}]$ 与 $\mathsf{C}_{m_2}[\mathsf{K}^\mathsf{T}]$ 之间的割距离，其定义为

$$\delta_\square(\mathbf{C}_{m_1}, \mathbf{C}_{m_2}) \triangleq \min_{\mathbf{K} \in \mathcal{K}(\mathbf{C}_{m_1},\mathbf{C}_{m_2})} d_\square(\mathbf{C}_{m_1}[\mathbf{K}], \mathbf{C}_{m_2}[\mathbf{K}^\mathsf{T}]]) \; , \tag{93}$$

其中，$d_\square$ 是相同大小矩阵之间的一种距离度量，定义如下。对于具有相同对角元素的两个矩阵 $\mathbf{C}_m$ 和$\mathbf{D}_m$，距离 $d_\square(\mathbf{C}_m,\mathbf{D}_m)$ 定义为

$$d_\square(\mathbf{C}_m, \mathbf{D}_m) \triangleq \max_{S,T \in [m]} \frac{1}{\mathsf{tr}(\mathbf{C}_m)^2} |e_{\mathbf{C}_m}(S, T) - e_{\mathbf{D}_m}(S, T)| \; , \tag{94}$$

其中

$$e_{\mathbf{C}_m}(S, T) \triangleq \sum_{i \in S, j \in T} [\mathbf{C}_m]_{ii}[\mathbf{C}_m]_{jj}[\mathbf{C}_m]_{ij} \; . \tag{95}$$

附录 E 定理 3 的证明

在定理3中，我们比较 VNN $\Phi(\mathbf{x}_{m_1}; \mathbf{C}_{m_1}, \mathcal{H})$ 和 $\Phi(\mathbf{x}_{m_2};\mathbf{C}_{m_2}, \mathcal{H})$ 的第 f 个输出的连续表示。我们在附录C中的讨论表明，这些连续表示自然地作为架构 $\Phi(y_{\mathbf{x}_{m_1}}; \mathbf{W}^{\mathbf{C}_{m_1}}, \mathcal{H})$ 和 $\Phi(y_{\mathbf{x}_{m_1}}; \mathbf{W}^{\mathbf{C}_{m_1}}, \mathcal{H})$ 分别作用于图子近似 $\mathbf{W}^{\mathbf{C}_{m_1}}$ 和 $\mathbf{W}^{\mathbf{C}_{m_2}}$ 时的输出出现。因此，我们后续的分析聚焦于

其组成图子滤波器之间的比较（定义见附录C）最终使我们能够确立 f-th outputs of $\Phi(y_{\mathbf{x}_{m_1}};\mathbf{W}^{\mathbf{C}_{m_1}},\mathcal{H})$ and $\Phi(y_{\mathbf{x}_{m_1}};\mathbf{W}^{\mathbf{C}_{m_1}},\mathcal{H})$ 之间的收敛性。有关定义在图子近似上的信息处理架构的详细信息，以及理解 $\Phi(y_{\mathbf{x}_{m_1}};\mathbf{W}^{\mathbf{C}_{m_1}},\mathcal{H})$ 与相应 VNN $\Phi(\mathbf{x}_{m_1};\mathbf{C}_{m_1},\mathcal{H})$ 之间关系的内容，请参阅附录C。

我们首先建立若干结果，以研究$\mathbf{W}$ 与$\mathbf{W}^{\mathbf{C}_m}$, y 与$y_{\mathbf{x}_m}$之间的比较，以及两个不同图子之间特征值的差异。我们利用式（22）中协方差矩阵序列 $\{\mathbf{C}_m\}$ 的（Ω ζ）-占优性质，以及式（24）中图子的利普希茨条件，建立以下结果。

引理 3. 给定一个 α-利普希茨图子 $\mathbf{W}$ ，且 $\mathbf{W}^{\mathbf{C}_m}$ 是一个 $(\Omega$ $\zeta)$-主导协方差矩阵$\mathbf{C}_m$ 的图子表示，我们有

$$\|\mathbf{W}-\mathbf{W}_{\mathbf{C}_m}\|_2 \leq \frac{\alpha\Omega^{3/2}}{m^{3\zeta/2-1}}\,. \tag{96}$$

证明。 根据$\mathbf{W}^{\mathbf{C}_m}$的构造，我们有

$$\|\mathbf{W}-\mathbf{W}_{\mathbf{C}_m}\|_2 = \Big(\int_0^1\int_0^1 \|\mathbf{W}(u,v)-\mathbf{W}_{\mathbf{C}_m}(u,v)\|^2 \mathsf{d}u\mathsf{d}v\Big)^{\frac{1}{2}}\,, \tag{97}$$

$$= \Big(\sum_{i,j}\int_{\mathcal{U}_i}\int_{\mathcal{U}_j} \|\mathbf{W}(u,v)-\mathbf{W}_{\mathbf{C}_m}(u,v)\|^2 \mathsf{d}u\mathsf{d}v\Big)^{\frac{1}{2}}\,. \tag{98}$$

不失一般性，假设 $\mathcal{U}_1 = [0,\rho_1]$ 是最大的区间。利用图子 $\mathbf{W}^{\mathbf{C}_m}$ 的 α-利普希茨连续性，并注意到 $\mathbf{W}^{\mathbf{C}_m}(\rho_i,\rho_j) = \mathbf{W}(\rho_i,\rho_j)$，我们有

$$\|\mathbf{W}-\mathbf{W}_{\mathbf{C}_m}\|_2 \leq \Big(\sum_{i,j}\int_{I_i}\int_{I_j} \alpha^2(|u|+|v|)^2 \mathsf{d}u\mathsf{d}v\Big)^{\frac{1}{2}}\,, \tag{99}$$

$$\leq \Big(m^2\int_0^{\rho_1}\int_0^{\rho_1} \alpha^2(|u|+|v|)^2 \mathsf{d}u\mathsf{d}v\Big)^{\frac{1}{2}}\,, \tag{100}$$

$$\leq \Big(m^2\int_0^{\rho_1}\int_0^{\rho_1} \alpha^2(|u|+|v|) \mathsf{d}u\mathsf{d}v\Big)^{\frac{1}{2}}\,, \tag{101}$$

$$\leq \alpha m\rho_1^{3/2}\,. \tag{102}$$

利用 $\mathbf{C}_m$ 是 $(\Omega$ $\zeta)$-占优这一假设，我们有

$$\|\mathbf{W}-\mathbf{W}_{\mathbf{C}_m}\|_2 \leq \frac{\alpha\Omega^{3/2}}{m^{3\zeta/2-1}}\,. \tag{103}$$

□

接下来，我们在引理 4 中刻画图子信号 $y \in L_2([0,1])$ 与从随机样本 $\mathbf{x}$ 中获得的近似值 $y_{\mathbf{x}_m}$ ob- 之间的差异。为此，我们作出如下假设：图子信号 y 满足 $|y(a)-y(b)| \leq \alpha_2|a-b|,\ \ \forall a,\ \ b \in [0,1]$。我们将满足此性质的图子信号称为 α_2-利普希茨图子信号。

引理 4。 给定一个 α_2-利普希茨图子信号 y 以及从$\mathbf{x}_m \in \mathbb{R}^{m\times 1}$中获得的图子信号近似值 $y_{\mathbf{x}_m}$，有

$$\|y - y_{\mathbf{x}_m}\|_2 \leq \frac{\alpha_2 \Omega^{3/2}}{m^{3\zeta/2-1}} \,. \tag{104}$$

证明。 注意

$$\|y - y_{\mathbf{x}_m}\|_2 = \sum_{\mathcal{U}_i} \|y - y_{\mathbf{x}_m}\|_{L_2[I_i]} \,, \tag{105}$$

$$= \sum_{i=1}^{m} \Big(\int_{\rho_{i-1}}^{\rho_i} (y(u) - y_{\mathbf{x}_m}(u))^2 \mathsf{d}u \Big)^{\frac{1}{2}} \,, \tag{106}$$

其中我们有 $\rho_0 = 0$。利用图子信号的 Lipschitz 性质以及 $\mathbf{C}_m$ 的 $(\Omega\ \zeta)$-性质，我们得到

$$\|y - y_{\mathbf{x}_m}\|_2 \leq m\Big(\alpha_2^2 \int_0^{\rho_1} u^2 \mathsf{d}u\Big)^{\frac{1}{2}} \leq \frac{\alpha_2 \Omega^{3/2}}{m^{3\zeta/2-1}} \,. \tag{107}$$

□

接下来，我们陈述 [39] 中的命题 4，该命题刻画了两个图子之间特征值差异的一个界。

引理 5 （ [39]中的命题 4）。 考虑两个图子 $\mathbf{W}$ 和 $\mathbf{W}'$，其特征值集合分别为 $\{\eta_i\}_{i=1}^{\infty}$和 $\{\beta_i\}_{i=1}^{\infty}$。那么，对于所有 $i \in \mathbb{Z}^+$，我们有

$$|\eta_i - \beta_i| \leq \|T_{\mathbf{W}-\mathbf{W}'}\|_2 \leq \|\mathbf{W} - \mathbf{W}'\|_2 \,. \tag{108}$$

接下来，我们利用引理 3、4 和 5，对图子卷积 $\Psi(y;\ \mathbf{W}\ \mathcal{H})$ 与由 $\mathbf{C}_m$ 上的图滤波器实现的近似卷积 $\Psi(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})$ 之间的差异进行界定。接下来的分析与 [79] 中的分析类似。为此，我们作出以下假设。

A1 图子$\mathbf{W}$ 是 α-利普希茨的。

A2 图子信号 y 是 α_2-利普希茨的。

A3 协方差矩阵$\mathbf{C}_m$ 属于一个收敛的 $(\Omega\ \zeta)$-占优协方差矩阵序列。此外，对应的图子近似序列属于图子$\mathbf{W}$的族。

˜A4 频率响应是带限的，即 $|h(\eta)| = 0$ 对于 $\eta \leq \eta_{\mathsf{c}}$。此外，我们假设图子$\mathbf{W}$ 的按幅值排序的 m_{c} 个最大特征值满足 $|\eta| > \eta_{\mathsf{c}}$，这些˜特征值的集合记为 $\mathcal{C}$。此外，图子滤波器是非扩张的，并满足 $|h(\eta)| \leq 1,\ \forall\tilde{\eta}$ ˜以及 $|h(\eta_i) - h(\hat{\eta}_i)| \leq \alpha_3|\eta_i - \hat{\eta}_i|$。

ˆ 在下面的引理中，我们分别用记号 $\{\hat{\eta}_i\}$ 和 $\{\Gamma_i\}$ 表示$\mathbf{W}_{\mathbf{C}_m}$的特征值集合和特征函数集合。

引理 6 (图子滤波器的可迁移性**)**。 对于卷积 $\Psi(y;\mathbf{W}\ \mathcal{H})$ 及其近似 $\Psi(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})$，在假设 *A1-A4* 下且对于 $\|y\|_2 \leq 1$，我们有

$$\|\Psi(y;\mathbf{W},\mathcal{H}) - \Psi(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2 \leq \frac{\Omega^{3/2}}{m^{3\zeta/2-1}}\left(\alpha_2 + \alpha\Big[\alpha_3 + \frac{\pi m_{\mathsf{c}}}{2\Delta_c}\Big]\right)\ , \tag{109}$$

其中 $\Delta_c =$ 最小值$_{i\neq j;i,\ j\in\mathcal{C}}\{|\eta_i - \hat{}\eta_j|\}$。

证明。 注意，我们可以将 $\|\Psi(y;\mathbf{W}\ \mathcal{H}) - \Psi(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\ \mathcal{H})\|_2$ 改写为

$$\begin{aligned}\|\Psi(y;\mathbf{W},\mathcal{H}) - \Psi(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2 = \|&\Psi(y;\mathbf{W},\mathcal{H}) - \Psi(y;\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\\ &+ \Psi(y;\mathbf{W}_{\mathbf{C}_m},\mathcal{H}) - \Psi(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2\ ,\end{aligned} \tag{110}$$

利用三角不等式可得

$$\begin{aligned}\|\Psi(y;\mathbf{W},\mathcal{H}) - \Psi(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2 \leq\ &\underbrace{\|\Psi(y;\mathbf{W},\mathcal{H}) - \Psi(y;\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2}_{\text{Term 1}}\\ &+ \underbrace{\|\Psi(y;\mathbf{W}_{\mathbf{C}_m},\mathcal{H}) - \Psi(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2}_{\text{Term 2}}\ .\end{aligned} \tag{111}$$

接下来，我们分别分析式 (111) 中的项 1 和项 2。

项 1 的分析。利用 $\Psi(y;\mathbf{W}\ \mathcal{H})$ 和 $\Psi(y;\mathbf{W}_{\mathbf{C}_m},\ \mathcal{H})$ 的展开式，可得

$$\|\Psi(y;\mathbf{W},\mathcal{H}) - \Psi(y;\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2 = \Big(\int_0^1 f^2(v)\mathsf{d}v\Big)^{1/2}\ , \tag{112}$$

其中

$$f(v) = \sum_{i=\in\mathbb{Z}\setminus\{0\}}\Big[\tilde{h}(\eta_i)\Gamma_i(v)\int_0^1 y(u)\Gamma_i(u)\mathsf{d}u - \tilde{h}(\hat{\eta}_i)\hat{\Gamma}_i(v)\int_0^1 y(u)\hat{\Gamma}_i(u)\mathsf{d}u\Big]\ . \tag{113}$$

通过添加并 su 在（113）中减去 $\tilde{h}(\hat{}\eta_i)\Gamma_i(v)\int_0^1 y(u)\Gamma_i(u)\mathsf{d}u$，并使用三角不等式，我们有 e

$$\|\Psi(y;\mathbf{W},\mathcal{H}) - \Psi(y;\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2 \leq \Big(\int_0^1 f_1^2(v)\mathsf{d}v\Big)^{1/2} + \Big(\int_0^1 f_2^2(v)\mathsf{d}v\Big)^{1/2} = \|f_1\|_2 + \|f_2\|_2\ , \tag{114}$$

其中

$$f_1(v) = \sum_{i\in\mathbb{Z}\setminus\{0\}}\Big[(\tilde{h}(\eta_i) - \tilde{h}(\hat{\eta}_i))\Gamma_i(v)\int_0^1 y(u)\Gamma_i(u)\mathsf{d}u\Big]\ , \tag{115}$$

and

$$f_2(v) = \sum_{i\in\mathbb{Z}\setminus\{0\}}\Big[\tilde{h}(\hat{\eta}_i)\Gamma_i(v)\int_0^1 y(u)(\Gamma_i(u) - \hat{\Gamma}_i(u))\mathsf{d}u\Big]\ . \tag{116}$$

˜ ˜利用图子滤波器的 Lipschitz 性质，我们有 $|h(\eta_i) - h(\hat{}\eta_i)| \leq \alpha_3|\eta_i - \hat{}\eta_i|$。因此，利用引理 5 和引理 3，我们有

$$\|f_1\|_2 \leq \frac{\alpha_3\alpha\Omega^{3/2}}{m^{3\zeta/2-1}}\ . \tag{117}$$

对于满足 $\|y\|_2 \leq 1$的任意 y。为了分析 $\|f_2\|_2$，注意到根据柯西-施瓦茨不等式，我们有

$$\|f_2\|_2 \leq \sum_{i\in\mathbb{Z}\setminus\{0\}} |\tilde{h}(\hat{\eta}_i)|\|\Gamma_i\|_2\|y(\Gamma_i - \hat{\Gamma}_i)\|_2 \,, \tag{118}$$

$$\leq \sum_{i\in\mathbb{Z}\setminus\{0\}} |\tilde{h}(\hat{\eta}_i)|\|\Gamma_i - \hat{\Gamma}_i\|_2 \,, \tag{119}$$

其中，式 (119) 由式 (118) 得出；不失一般性地取 $\|y\|_2 = 1$、$\|\Gamma_i\|_2 = 1$，并再次应用柯西-施瓦茨不等式。接下来注意到，积分算子 $T_{\mathbf{W}}$ 满足 $(T_{\mathbf{W}}y)(v) = \int_0^1 \mathbf{W}(u,\ v)y(u)\mathrm{d}u$，是一个自伴希尔伯特–施密特算子，且 $\mathbf{W}$ 存在谱分解 $\hat{}$，其中特征值为 $\{\eta_i\}$，特征信号为 $\{\Gamma_i\}$。因此，为了分析 $\|\Gamma_i - \Gamma_i\|_2$，注意到 Γ_i 是与特征值 η_i 相关的算子 $T_{\mathbf{W}}$ 的$\hat{}$投影，而 Γ_i 是与特征值 $\hat{}\eta_i$相关的算子 $T_{\mathbf{W}_{\mathbf{C}_m}}$ 的投影。将 $T_{\mathbf{W}}$ 的谱划分为 $\mathrm{spec}(T_{\mathbf{W}}) = \{\eta_i\} \cup \{\eta_j\}_{j\neq i}$，并将 $T_{\mathbf{W}_{\mathbf{C}_m}}$ 的谱划分为 $\mathrm{spec}(T_{\mathbf{W}_{\mathbf{C}_m}}) = \{\hat{}\eta_i\} \cup \{\hat{}\eta_j\}_{j\neq i}$，我们应用 [118] 中的命题 2.3，得到

$$\|\Gamma_i - \hat{\Gamma}_i\|_2 \leq \frac{\pi}{2}\frac{\|T_{\mathbf{W}} - T_{\mathbf{W}_{\mathbf{C}_m}}\|_2}{d_i} \,, \tag{120}$$

其中，$d_i > 0$ 是一个满足 $|\eta_i - \hat{}\eta_{i+1}| \geq d_i$、$|\eta_i - \hat{}\eta_{i-1}| \geq d_i$、$|\eta_{i+1} - \hat{}\eta_i| \geq d_i$和$|\eta_{i-1} - \hat{}\eta_i| \geq d_i$的常数。根据（120）、引理 5 和引理 3，可得（119）中的结果。

$$\|f_2\|_2 \leq \frac{\pi\alpha\Omega^{3/2}}{2\Delta_c m^{3\zeta/2-1}} \sum_{i\in\mathbb{Z}\setminus 0} |\tilde{h}(\hat{\eta}_i)| \,, \tag{121}$$

其中 $\Delta_c = \min_i d_i$。接下来，我们注意到，在假设 A4 下，图子滤波器$\tilde{}$的频率响应是带限的，即当频率响应根据 A4 受到限制时，我们有 $\sum_{i\in\mathbb{Z}\setminus 0} |h(\hat{}\eta_i)| \leq m_{\mathsf{c}}$。我们将$\mathbf{W}$ 的 m_{c} 个最大特征值（按绝对值计）组成的集合记为 $\mathcal{C}$。在这种情况下，我们可以将 (121) 重写为

$$\|f_2\|_2 \leq \frac{\pi\alpha\Omega^{3/2}m_{\mathsf{c}}}{2\Delta_c m^{3\zeta/2-1}} \,. \tag{122}$$

显然，m_c 与 ζ 之间存在权衡，因为为了使（122）关于 m呈现递减趋势，必须满足 $m_c < m^{3\zeta/2-1}$ 和 $\zeta > 2/3$。方程117和122给出了项 1的上界。

项 2 的分析。我们可以将项 2 展开为

$$\|\Psi(y;\mathbf{W}_{\mathbf{C}_m},\mathcal{H}) - \Psi(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2 = \Big(\int_0^1 g^2(v)\mathrm{d}v\Big)^{1/2} \,, \tag{123}$$

其中

$$g(v) = \sum_{i=1}^{\infty}\Big[\tilde{h}(\eta_i)\hat{\Gamma}_i(v)\int_0^1 (y(u) - y_{\mathbf{x}_m}(u))\hat{\Gamma}_i(u)\mathrm{d}u\Big] \,. \tag{124}$$

因此，利用 (124)，我们有

$$\|\Psi(y;\mathbf{W}_{\mathbf{C}_m},\mathcal{H})-\Psi(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2=\|\Psi(y-y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2\,. \tag{125}$$

注意，对于满足 $h(\eta)<1$的频率响应，图子滤波器不会扩张，因此我们有

$$\|\Psi(y;\mathbf{W}_{\mathbf{C}_m},\mathcal{H})-\Psi(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2\leq\|y-y_{\mathbf{x}_m}\|_2\,. \tag{126}$$

根据引理 4，我们有

$$\|\Psi(y;\mathbf{W}_{\mathbf{C}_m},\mathcal{H})-\Psi(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})\|_2\leq\frac{\alpha_2\Omega^{3/2}}{m^{3\zeta/2-1}}\,. \tag{127}$$

因此，结合式 (117)、(122) 和 (127) 中对项 1 和项 2 的上界，引理 6 的证明得证。

□

引理 6 建立了图子$\mathbf{W}$ 与由协方差矩阵$\mathbf{C}_m$得到的图子近似$\mathbf{W}_{\mathbf{C}_m}$ 之间的迁移性。我们利用引理 6 的结果，在类似的设定下建立图子神经网络的迁移性。我们将图子神经网络 $\tilde{\Psi}(y;\mathbf{W}_{\mathbf{C}_m},\mathcal{H})$ 的第 f 个输出（其最终层有 F 个输出）记为 $[\tilde{\Psi}(y;\mathbf{W}_{\mathbf{C}_m},\mathcal{H})]_f$。

引理 7（图子神经网络的可迁移性）。考虑一个有 L 层且每层有 F 个输出的图子神经网络 $\tilde{\Phi}(\cdot;\mathbf{W},\mathcal{H})$，以及一个 VNN $\Phi(\cdot;\mathbf{C}_m,\mathcal{H})$，其图子神经网络表示为 $\tilde{\Phi}(\cdot;\mathbf{W}_{\mathbf{C}_m},\mathcal{H})$。若协方差矩阵 $\mathbf{C}_m$ 属于一个 (Ω,ζ)-占优协方差矩阵序列，且其图子近似 $\mathbf{W}_{\mathbf{C}_m}$ 属于一个由 α-利普希茨图子 $\mathbf{W}$ 构成的图子族，则在假设 A1-A4 下，对于 $\|y\|_2\leq 1$ 和 $2/3<\beta\leq 1$，我们有

$$\|[\tilde{\Phi}(y;\mathbf{W},\mathcal{H})]_f-[\tilde{\Phi}(y_{\mathbf{x}_m};\mathbf{W}_{\mathbf{C}_m},\mathcal{H})]_f\|_2\leq LF^L\left(\frac{\Omega^{3/2}}{m^{3\zeta/2-1}}\left[\alpha_2+\alpha\left[\alpha_3+\frac{\pi m_{\mathbf{c}}}{2\Delta_c}\right]\right]\right)\,. \tag{128}$$

引理 7 的证明利用了引理 6，并考虑了多层 VNN 架构的影响。关于确切的分析步骤，请参阅 [39] 中的公式 (23)-(28)。最后，通过对 (128) 应用三角不等式，我们分别建立了图子神经网络近似 $\tilde{\Phi}(\cdot;\mathbf{W}^{\mathbf{C}_{m_1}},\mathcal{H})$ 与 $\tilde{\Phi}(\cdot;\mathbf{W}^{\mathbf{C}_{m_2}},\mathcal{H})$，以及 VNN $\Phi(\cdot;\mathbf{C}_{m_1},\mathcal{H})$ 与 $\Phi(\cdot;\mathbf{C}_{m_2},\mathcal{H})$ 之间的迁移性，并完成了定理3的证明。

附录F FTDC数据集协方差矩阵的收敛性

割度量可以比较不同规模图的矩阵表示 [81]。此处，使用 Python 中的 cutnorm 软件包估计了一对协方差矩阵之间的割范数（实现基于 [119]）。在此，我们研究按照 Schaefer图谱的不同尺度整理的数据集所形成的协方差矩阵的收敛性。为此，除 FTDC数据集外，我们还考虑了按照 Schaefer图谱的 200 个和 400 个脑区划分整理的数据集。图 13a 绘制了针对序列 $\{\mathbf{C}_{100}, \mathbf{C}_{200}, \mathbf{C}_{300}, \mathbf{C}_{400}, \mathbf{C}_{500}\}$ 评估的割范数，其中 $\mathbf{C}_m$ 是根据 Schaefer图谱的 m 个脑区划分整理的皮层厚度特征构建的解剖协方差矩阵。随着脑区数量增加，该序列中连续成员之间的割范数减小，这与柯西序列在该度量下的性质一致。我们还注意到，由 OASIS-3 得到的协方差矩阵之间的距离

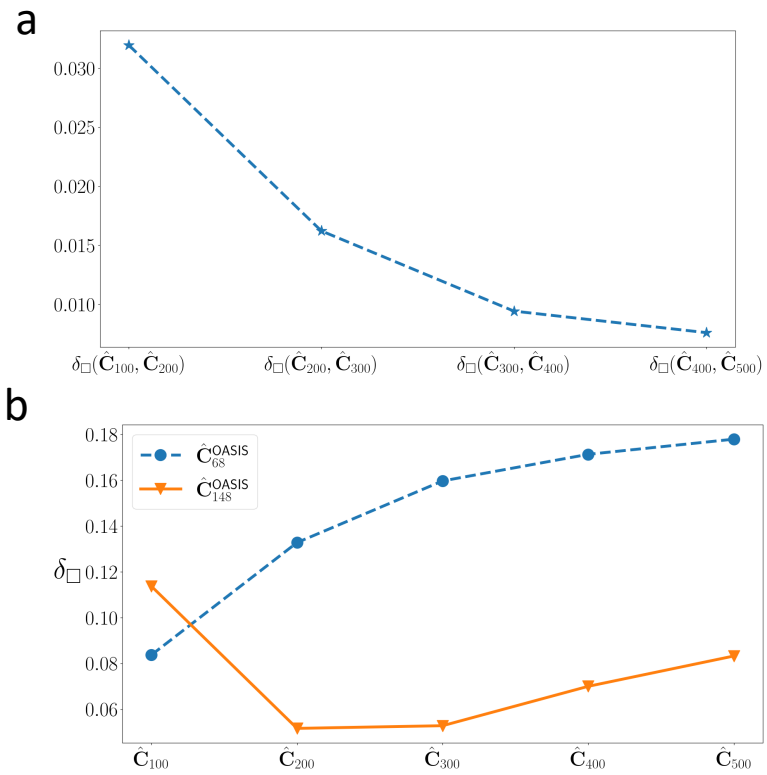


图 13：**协方差矩阵之间的割范数**。 面板 **a**绘制了协方差矩阵序列 $\{\mathbf{C}_{100}, \mathbf{C}_{200}, \mathbf{C}_{300}, \mathbf{C}_{400}, \mathbf{C}_{500}\}$ 中连续元素之间的估计割范数。面板 **b** 绘制了根据 DK 和 DKT图谱从 OASIS-3 数据集得到的协方差矩阵与来自 FTDC数据集的协方差矩阵之间的估计割范数。

数据集与 FTDC数据集中的数据之间的距离，显著大于与 Schaefer图谱不同分辨率相关的协方差矩阵之间的距离（图 13b）。在图 13b中，协方差

矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{OASIS}}$是根据健康对照组按照 DKT图谱整理的皮层厚度特征估计得到的。$\hat{\mathbf{C}}_{68}^{\mathrm{OASIS}}$是根据同一组按照 Desikan-Killiany图谱整理的皮层厚度特征估计得到的。根据定理3，VNN的性能可在其协方差矩阵属于收敛序列的数据集之间迁移。因此，我们预期，按照 Schaefer图谱不同尺度整理的皮层厚度数据集之间具有 VNN 可迁移性。然而，按照 Schaefer图谱整理的数据集与按照 DKT图谱整理的数据集之间，并不能保证 VNN 可迁移性。

附录 G：VNN 在测试集上的性能总结

此处，我们提供了分别在 FTDC100、FTDC500 和 OASIS-3 数据集上训练的不同 VNN 集合中，标称模型在测试集上的性能。训练过程中完全未使用该测试集。

OASIS-3：测试集上 100 个名义模型的平均绝对误差为 5.87 ± 0.177 年，皮尔逊相关系数为 0.425 ± 0.006。

FTDC100：测试集上 100 个名义模型的平均绝对误差为 3.72 ± 0.22 年，皮尔逊相关系数为 0.78 ± 0.01。

FTDC300：测试集上 100 个名义模型的平均绝对误差为 3.93 ± 0.34 年，皮尔逊相关系数为 0.75 ± 0.008。

FTDC500：在测试集上，100 个名义模型的平均绝对误差为 4.07 ± 0.36 年，皮尔逊相关系数为 0.72 ± 0.007。

附录 H VNN 模型输出的区域残差分析示意图

在本节中，我们展示第 4.4.1 节所述的区域分析，该分析针对一个 VNN 模型进行。该模型经过训练，用于预测 OASIS-3 数据集中 HC 组的实际年龄。本节中提及的所有数学符号均沿用第 4.4.1 节中的定义。请注意，为了评估脑龄或区域残差，我们没有对该 VNN 模型进行进一步训练。

在此 VNN 模型中，协方差矩阵被替换为 $\mathbf{C}_{148}^{\mathsf{AD}+}$，并对 HC 和阿尔茨海默病+ 个体组成的合并组进行处理。因此，对于该合并数据集中的每位个体 i，我们获得了一个年龄预测值 $\hat{}y_i$ 以及一个残差向量 $\mathbf{r}_i$。残差向量 $\mathbf{r}_i$的大小为 148×1，因此，$\mathbf{r}_i$ 中的每个元素都对应 DKT 脑图谱所定义的一个独立脑区，该图谱包含 148 个分区。通过评估合并数据集中每位个体的残差向量$\mathbf{r}_i$，我们得以构建一个由 HC 组残差向量组成的群体（记作$\mathbf{r}_{\mathsf{HC}}$）和一个由阿尔茨海默病+组残差向量组成的群体（记作$\mathbf{r}_{\mathsf{AD}+}$）。本文将这些残差向量的各个元素称为区域残差。

如第 4.4.1 节所述，我们通过方差分析考察这些残差向量的每个维度，以检验 HC 组和阿尔茨海默病+组之间的组间差异。因此，对于每个 VNN 模型，我们最终进行了$m = 148$ 次方差分析检验，并评估各脑区的残差是否存在显著组间差异。针对某一脑区，HC 组与阿尔茨海默病+ 组区域残差分布之间组间差异的显著性，是通过对方差分析检验的 p值进行 Bonferroni 多重比较校正后确定的（Bonferroni 校正后的 p值< 0.05）。此外，还使用以年龄和性别为协变量的协方差分析，在未经校正的显著性水平下检验组间差异。

图 14 展示了在此背景下通过方差分析（ANOVA）获得的结果。根据方差分析的 F 值大小对脑区进行着色，所有着色区域均符合第 4.4.1 节中给出的标准，具有统计学显著性。不同脑区的箱线图显示，与健康对照组（HC）相比，阿尔茨海默病（AD）组的残差显著升高。尽管识别出了许多具有显著性的脑区，但我们注意到显著性存在明显的变异性，即某些脑区的组间差异比其他脑区更为显著。我们针对 OASIS-3 数据集训练了 100 个 VNN 模型，并分别对每个模型进行了类似的分析。此外，我们统计了上述分析结果中将某一脑区判定为具有显著性的模型数量。我们预期，在 HC 与 AD 之间的脑区残差分布中具有稳健组间差异的脑区，会更频繁地被 VNN 模型标记为具有显著性。针对 OASIS-3 数据集进行的这项稳健性分析结果如图 7b 所示。

图 15 展示了将 VNN 模型从 FTDC100 数据集迁移到 OASIS-3 数据集后，对所得区域残差进行分析的结果。在这里，我们观察到，与图 14 中相似的脑区在 AD+ 组中表现出显著升高的区域残差（尽管在方差分析检验的 + -值方面显著性较低）。

因此，图 15 证明，尽管定量性能缺乏可迁移性，但在 FTDC100 数据集上训练的 VNN 仍可将可解释性从 FTDC100 数据集迁移到 OASIS-3 数据集。

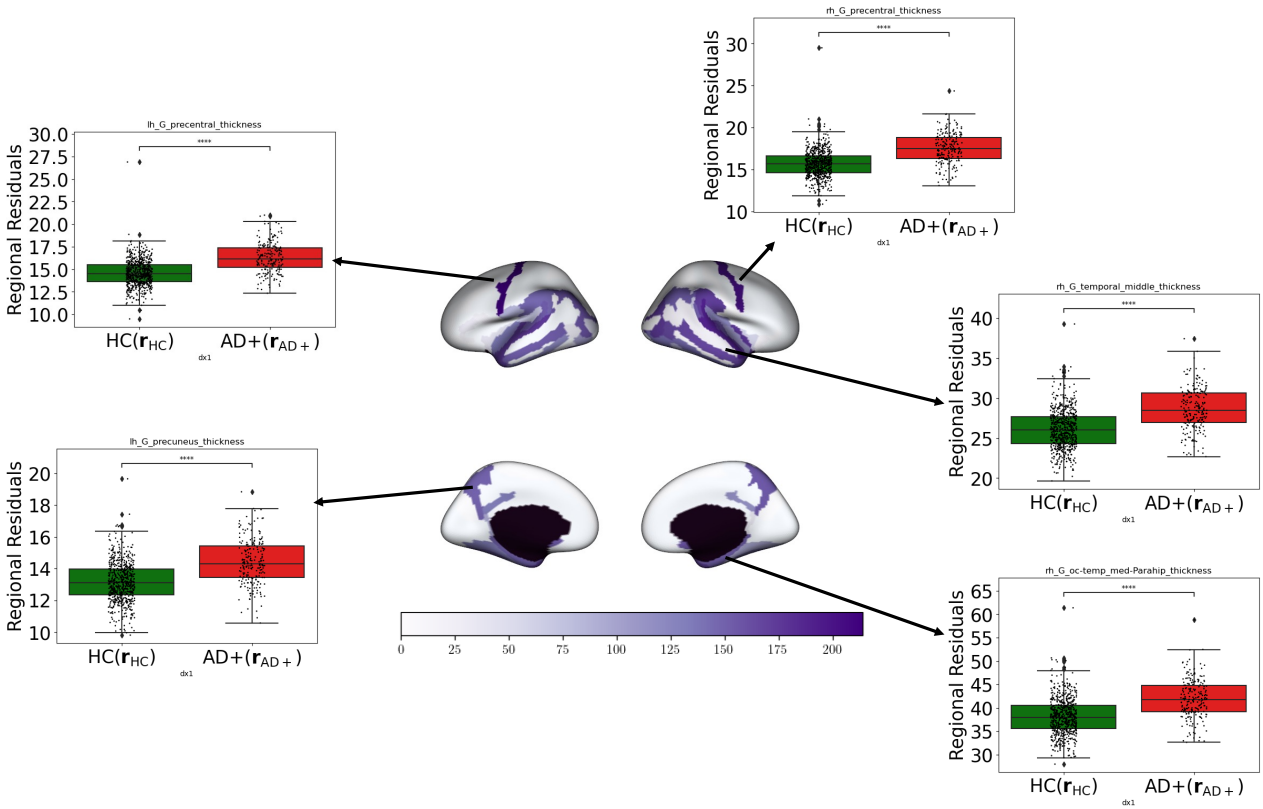


图 14：展示 OASIS-3 中 AD+ 组相较于 HC组区域残差显著升高的脑区。此处结果来自一个 VNN 模型，该模型以 OASIS-3 中 HC组的皮层厚度数据为基础进行训练，作为回归模型预测实际年龄。阴影区域的对比度与通过方差分析得到的 F 值成正比，且所有阴影区域在 AD+中相较于 HC 均表现出升高的区域残差。

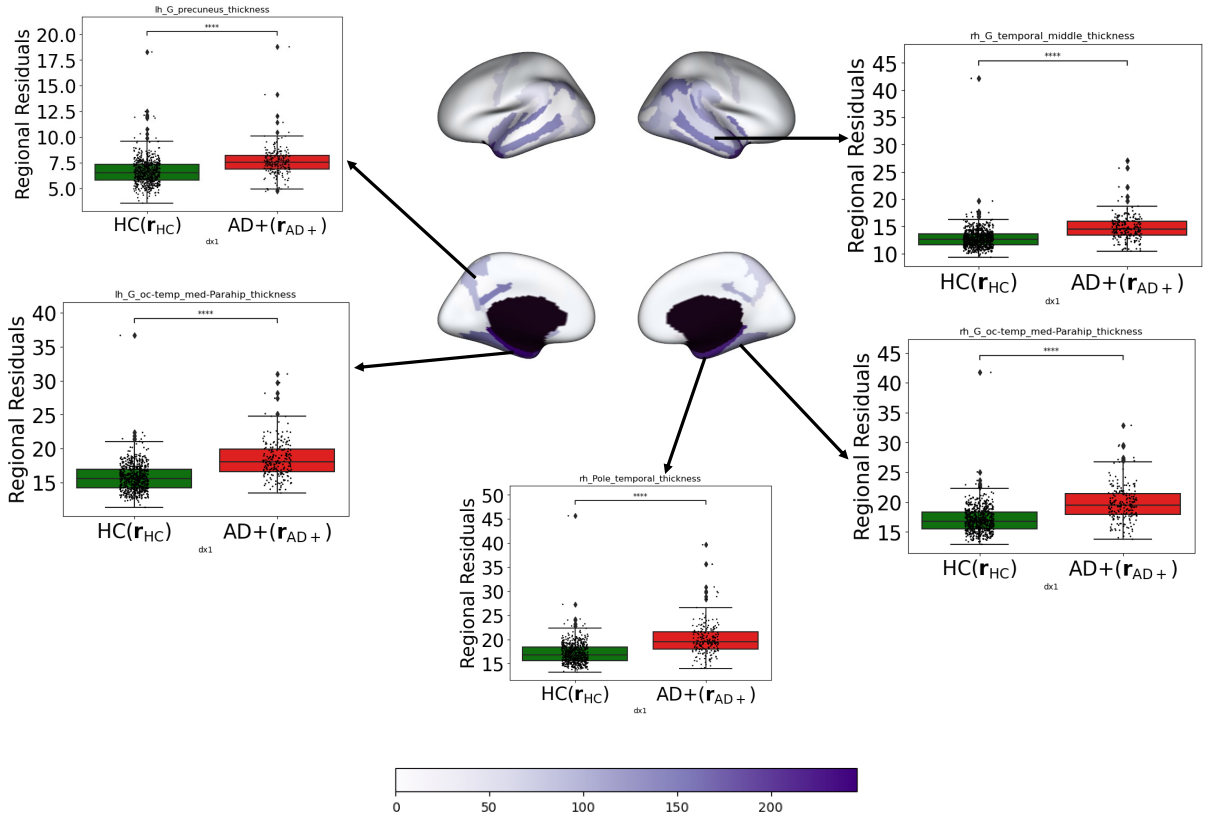


图 15：展示了在 OASIS-3 中 AD+ 组相对于 HC 组具有显著升高的区域残差的脑区。此处的结果由一个从 FTDC100 迁移而来的 VNN 模型得出，该模型用于处理 OASIS-3，并已作为回归模型进行训练，以根据 FTDC100 数据集中健康对照组的皮层厚度数据预测实际年龄。阴影区域的对比程度与通过方差分析获得的 F 值成正比，且所有阴影区域均表现出 AD+ 相对于 HC 显著升高的区域残差（Bonferroni 校正后的 p值小于 0.05）。

附录 I：OASIS-3 中脑龄预测的更多细节

本节提供了关于图 7 中可解释脑龄预测结果的补充图表和讨论。图 16 展示了 OASIS-3 数据集中阿尔茨海默病+组和健康对照组的实际年龄分布。

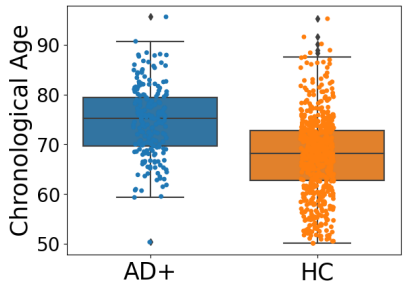


图 16：阿尔茨海默病+组和健康对照组的实际年龄分布。

由于 VNN 是针对回归任务进行训练的，因此 VNN 会处理皮层厚度数据，并为每位个体提供实际年龄估计。我们在 OASIS-3 的不同训练集排列上训练了 100 个 VNN 模型，因此将所有 VNN 估计值的均值作为每位个体的 VNN 预测值。随后利用该 VNN 预测值形成脑龄估计值和 Δ-Age 指标。图 17a 展示了完整健康对照组的 VNN 预测值与实际年龄（真实值）的关系图。健康对照组中，VNN 预测值与实际年龄（真实值）之间的皮尔逊相关系数为 0。43，与表 1 报告的结果相似。这一观察结果表明，尽管将解剖协方差矩阵从仅基于健康对照组替换为基于健康对照组和阿尔茨海默病+组合组（$\hat{\mathbf{C}}_{148}^{AD+}$），VNN 在健康对照组实际年龄预测任务上的性能并未下降。然而，对于健康对照组中年龄较大的个体，VNN 输出明显低估了实际年龄；对于年龄分布中较年轻一端的个体，则明显高估了实际年龄。

图 17b 显示了完整的 AD+ 组中 VNN 预测值与实际年龄（真实值）的散点图。AD+ 组中 VNN 预测值与实际年龄（真实值）的皮尔逊相关系数为 0.28。我们进一步指出，VNN架构以及我们对区域残差的分析，有助于量化各脑区对图 17a 和图 17b 中数据点的贡献。因此，与HC组相比，AD+ 组中某些脑区的贡献较大，这可能会影响图 17b 中的散点图。

图 17c 展示了通过 VNN 预测值与实际年龄之差评估的 HC 和阿尔茨海默病+ 组的残差箱线图。图 17c 表明，与 HC 组相比，阿尔茨海默病+组的实际年龄被低估了。由于阿尔茨海默病+ 组的年龄显著高于 HC 组，这一观察结果也是预期之中的。然而，我们预计，稳健升高的区域

图 7b 中脑区的残差在一定程度上缓解了 AD+ 组因年龄较高而导致的低估效应。

图17d-f展示了对VNN输出应用年龄偏差校正后的结果。正如预期，图17d中健康对照组的脑龄主要集中在相等线（$x = y$ line）附近。相比之下，图17e中AD+ 组的脑龄则集中在相等线上方。这些效应体现在图17f中 Δ-Age的箱线图中，我们观察到，与健康对照组相比，AD+ 组的 Δ-Age 较高。

VNN架构有助于在应用年龄偏差校正之前，分离加速衰老的影响。因此，图 17a-c中的VNN输出转换为图17d-f中的脑龄并不令人意外。然而，对于缺乏透明度的机器学习方法而言，这类洞见可能难以获得，因此无法解读或分离神经退行性变导致的偏离健康对照人群的影响。在此背景下，如果学习模型是一个黑箱，图 17a-c可能看起来与检测AD+ 组加速衰老的目标背道而驰，而年龄偏差校正的影响也可能不明确，进而招致诸多批评 [18]。

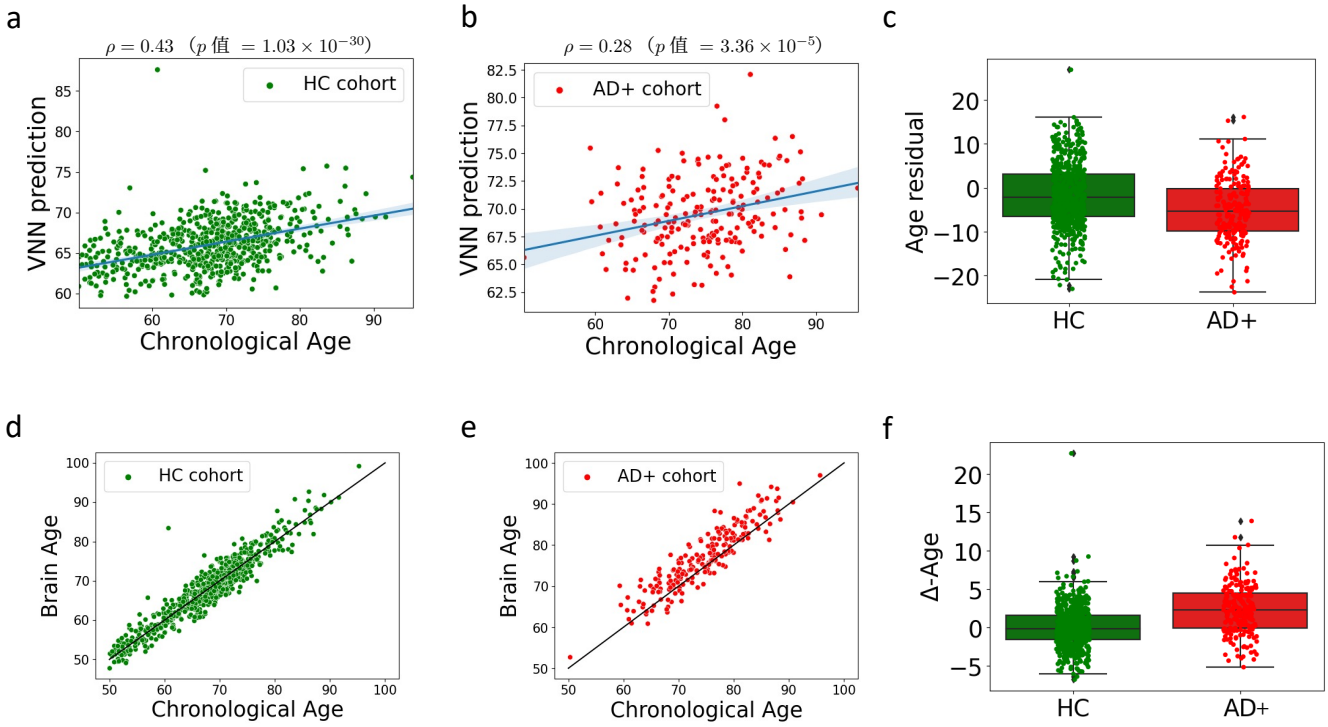


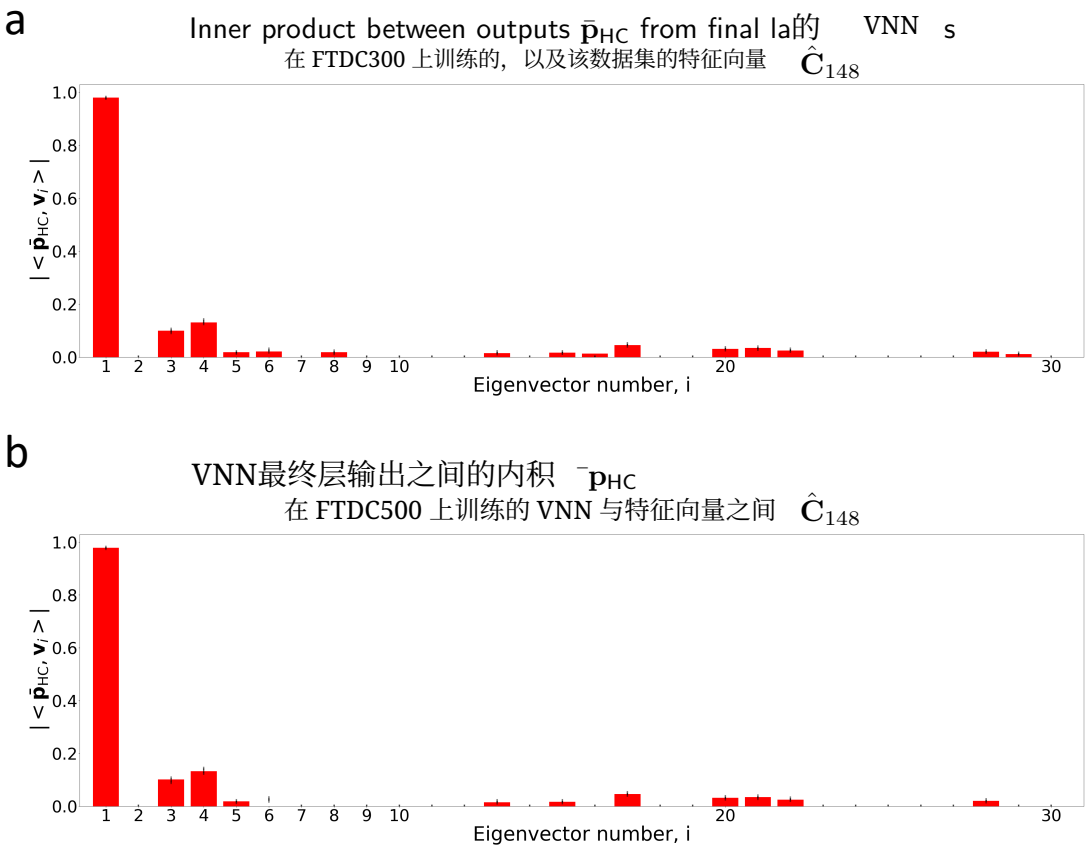
图 17：**图 7 的补充图**。 面板 **a** 显示了 HC组中 VNN 预测值与实际年龄的关系图。VNN 预测值是 100 个在 OASIS-3 上训练并作用于解剖协方差矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$ 的标称变分神经网络的输出平均值。面板 **b** 显示了 AD+ 组中与面板 **a** 类似的结果。面板**a** 和 **b** 中的实线为最小二乘拟合线。面板**c** 包含 HC组和 AD+ 组中由 VNN 预测值与实际年龄之差得出的残差箱线图。面板**d** 和 **e** 分别显示了 HC 组和 AD+组中脑龄与实际年龄的关系图。面板 **d** 和 **e** 中的实线为恒等线。面板 **f**显示了 HC组和 AD+ 组中的 Δ-年龄箱线图。

附录 J VNN 得出的区域残差与 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$ 特征向量之间关联的其他结果

图18展示了在 FTDC300 或 FTDC500 数据集上训练的 VNN，并将其迁移用于处理 OASIS-3 数据集中 HC组数据时，VNN 输出 $\hat{}(\bar{\mathbf{p}}_{\mathrm{HC}})$ 的区域贡献与解剖协方差矩阵 $\mathbf{C}_{148}$ 的特征向量之间的关联结果。图 18 中的结果表明，$\mathbf{C}_{148}$ 的第一、第三和第四个特征向量与基于 FTDC300 或 FTDC500 数据集训练的 VNN 所得的区域贡献具有最高的三项关联，这与图 6 中基于 OASIS-3 或 FTDC100 数据集训练的 VNN 所得的结果一致。除这三个特征向量外，与图 6a 中基于 OASIS-3 数据集训练的 VNN 所得结果相比，图18中基于 FTDC300 或 FTDC500 数据集训练的 VNN 所得结果中，满足 $|<\bar{\mathbf{p}}_{\mathrm{HC}},\mathbf{v}_i>|$ 在 HC组中的变异系数小于 30% 的特征向量数量较少。

图19a展示了使用在不同数据集上训练的 VNN 得出的阿尔茨海默病+组区域残差与 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$的特征向量之间的关联。对于在 OASIS-3 上训练的 VNN，区域残差与 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$的前三大关联分别出现在第一、第三和第四个特征向量上。有趣的是，使用从 FTDC 数据集迁移至 OASIS-3 数据集的 VNN 的最终层输出评估区域残差时，这一现象依然存在。不过，我们还发现，与在 OASIS-3 数据集上训练的 VNN 相比，在 FTDC 数据集上训练的 VNN 的内积方差更大。因此，在图 19a 中，对于在 FTDC300 或 FTDC500 数据集上训练的 VNN，与区域残差相关联且变异系数小于 30% 的特征向量数量显著减少。

图19b展示了由不同的100个VNN集合得出的区域特征，这些特征对应于OASIS-3中阿尔茨海默病+组相对于健康对照组的区域残差升高的稳健性。以基于OASIS-3数据集训练的VNN得出的区域特征为基准，基于FTDC100训练的VNN所获得的区域残差分析得出的区域特征最为一致。基于FTDC300或FTDC500数据集训练的VNN得出的区域特征，在海马旁区和胼胝体下区保留了最高的稳健性。根据图19a中的结果，这些观察结果是意料之中的，因为$\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$的第三和第四特征向量涵盖了海马旁区和胼胝体下区，且与基于FTDC数据集训练的VNN得出的区域残差高度相关。图19中的结果进一步证实了我们的结论：导致 Δ-年龄升高的因素所具有的特征性区域特征，取决于VNN利用解剖协方差矩阵中特定特征向量的能力。



ˆ 图18：**特征向量$\mathbf{C}_{148}$ 与 VNN 输出中的区域贡献（$\bar{\mathbf{p}}_{HC}$）之间的关联，这些 VNN 从 FTDC300 或 FTDC500 数据集迁移至 OASIS-3 数据集**。面板**a**展示了 $|<\bar{\mathbf{p}}_{HC}$的条形图，$\mathbf{v}_i>|$ 对应于 $i \in \{1,\ldots,30\}$，其中$\mathbf{v}_i$是协方差矩阵$\mathbf{C}_{148}$ 的第 i 个特征向量（主ˆ成分），并且与按幅值计第 i 大的特征值相关联；区域贡献向量 $\bar{\mathbf{p}}_{HC}$ 由在 FTDC300 数据集上训练、并迁移用于处理 OASIS-3 数据集中 HC 组数据的 VNN 获得。在 OASIS-3 的 HC 组中，变异系数大于 30% 的特征向量的内积结果被排除（因此，其贡献设为 0）。对于 HC 组中的每名个体，在 100 个名义VNN模型上评估了其相应的区域贡献向量 $\bar{\mathbf{p}}_{HC}$ 与$\mathbf{C}_{148}$ 的特征向量之间的关联。面板**b** 展示了与面板**a** 类似的结果，对象为在 FTDC500 数据集上训练、并迁移用于处理 OASIS-3 数据集中 HC 组数据的 VNN。

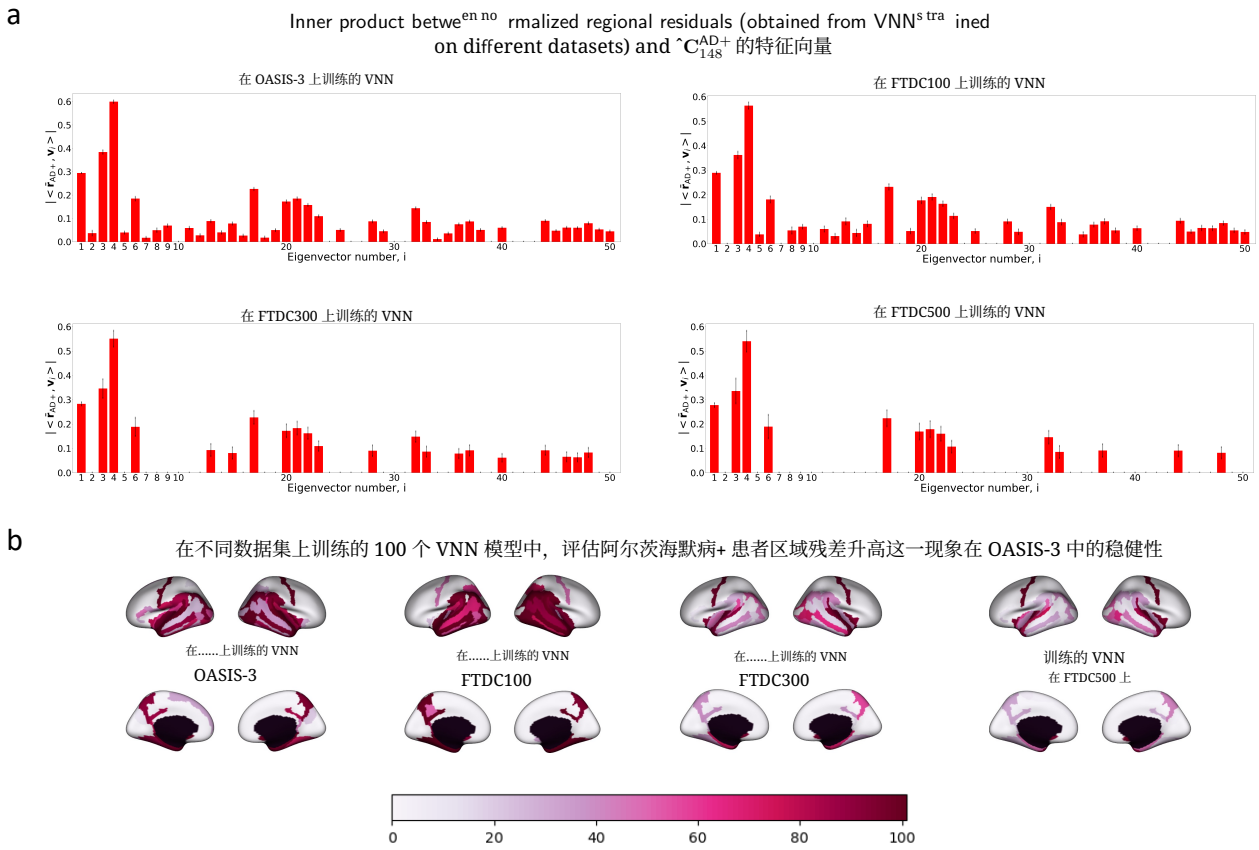


图19：**解剖协方差矩阵 $\hat{\mathbf{C}}_{148}^{AD+}$的特征向量与阿尔茨海默病+ 组的区域残差之间的关联；以及阿尔茨海默病+ 组中与升高的区域残差相对应的区域特征**。面板**a** 展示了内积 $|<\bar{\mathbf{r}}^{AD+}, \mathbf{v}_i>|$的条形图，该内积表示解剖协方差矩阵 $\hat{\mathbf{C}}_{148}^{AD+}$的前 50 个特征向量与在阿尔茨海默病+ 组中评估的标准化区域残差之间的内积，其中区域残差来自不同 VNN 集合的最终层输出。对于 $|<\bar{\mathbf{r}}^{AD+}, \mathbf{v}_i>|$ 在阿尔茨海默病+ 组中的变异系数大于 30% 的特征向量，其结果被排除（因此在条形图中对应的数值为 0）。面板**b** 展示了在 OASIS-3 中，与健康对照组相比，阿尔茨海默病+ 组中具有稳健且显著升高的区域残差所关联脑区的区域特征。其中包括对来自在 OASIS-3 和 FTDC数据集上训练的 VNN 所生成区域残差进行分析后获得的区域特征。

附录K VNN的可迁移性：用于评估多尺度FTDC数据集中的区域特征和Δ-年龄

本节提供一项概念验证：当 VNN 在对应于不同 Schaefer图谱分辨率的皮层厚度数据集之间迁移时，不同 Schaefer图谱尺度之间用于实际年龄预测的可迁移性，会使由区域残差确定的 Δ-年龄呈现一致的区域特征。下文将介绍本实验所用的数据集。

数据集： 该数据集包含来自 FTDC 数据集的 105 名健康对照组，以及 67 名轻度认知障碍或阿尔茨海默病患者（AD+；年龄 $= 68.52 \pm 9.29$ 岁，女性 28 名）。皮层厚度数据采用 Schaefer图谱的 100、300 和 500 分辨率。

以上数据集的区域残差由一个 VNN 模型计算得出，该模型曾在 FTDC100 数据集上训练，用于预测实际年龄。我们使用方差分析评估了 Schaefer图谱 100、300 和 500 分辨率下健康对照组与 AD+ 组（合并）的皮层厚度区域残差组间差异。图 20 将 AD+ 组相较于健康对照组呈现稳健且显著升高的区域残差所对应的脑区投射到 100、300 和 500 分辨率的脑模板上。图 20 显示，在所有 Schaefer图谱分辨率下，分离出的脑区均集中于相似区域，且与图 7b 中的脑区一致（但双侧楔前叶和顶上小叶区域除外）。

随后，对所有个体的 Δ年龄进行了评估。使用包含 100 个特征的皮层厚度数据集评估得出的阿尔茨海默病组 Δ年龄为 3.67 ± 3.73 岁，健康对照组为 0 ± 2.06 岁。图20b展示了FTDC100 数据集中健康对照组和阿尔茨海默病+ 组的 Δ年龄箱线图，以及将 VNN 从 FTDC100 迁移至 FTDC300 和 FTDC500 数据集后得到的箱线图。由于 VNN 在 Schaefer图谱不同尺度间具有可迁移性，即使 VNN 仅在 FTDC100 数据集上训练，100、300 和 500 分辨率数据中的 Δ年龄分布及脑龄与实际年龄的关系图仍保持一致。

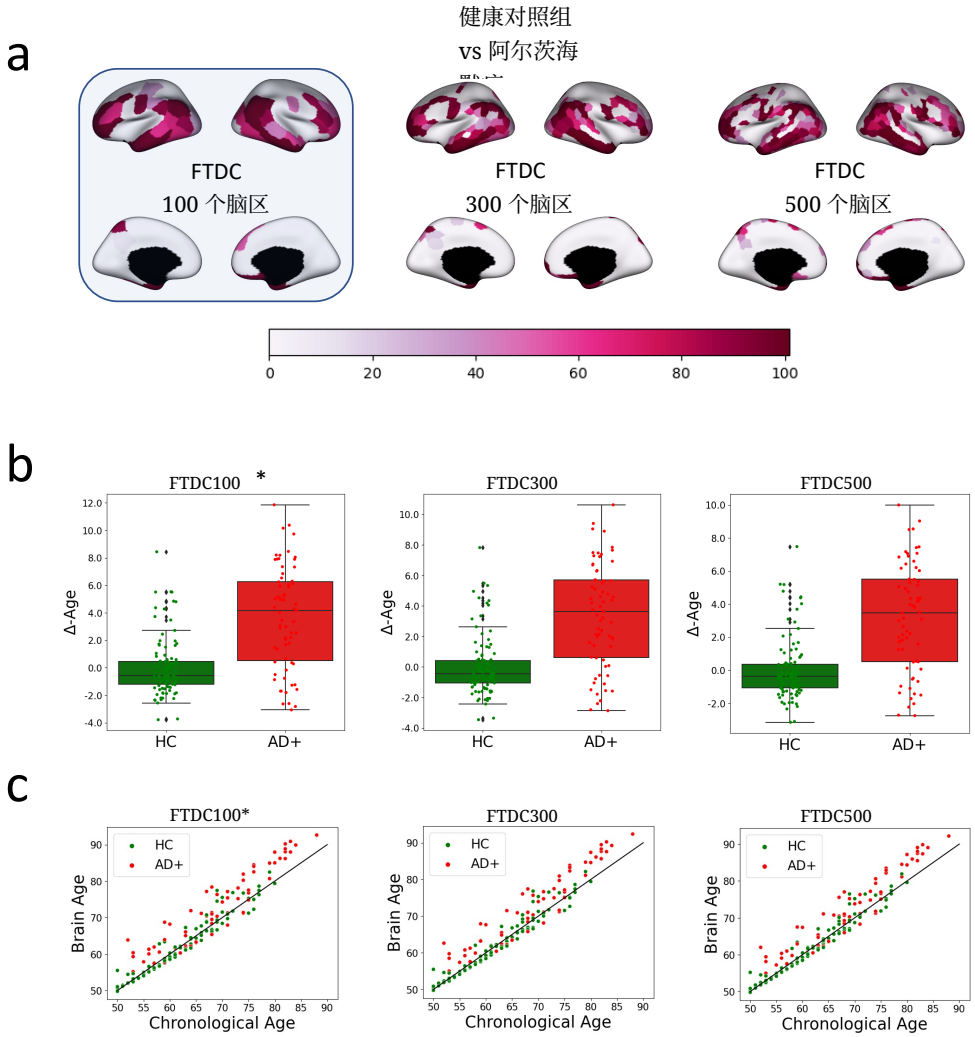


图20：**根据 Schaefer图谱的多个尺度构建的数据集进行脑龄预测**。面板**a** 展示了区域特征，其中包含合并AD+ 组相对于HC组具有稳健且升高的区域残差的脑区。VNN在FTDC100上训练，以预测实际年龄，并通过100个名义VNN模型评估其稳健性。将VNN从FTDC100迁移至FTDC300和FTDC500后，获得了具有300个特征和500个特征的数据集的区域特征。面板**b** 展示了与面板**a**中的区域特征相对应的 Δ-年龄箱线图。面板**c** 绘制了根据Schaefer图谱不同尺度构建的、分别具有100、300和500个皮层厚度特征的数据集的脑龄与实际年龄之间的关系图。

附录 L AD+ 组中与较高区域残差对应的区域特征，对用于估计解剖协方差矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$的数据构成具有稳定性

回顾一下， Δ-年龄及相关区域特征是通过作用于复合解剖协方差矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$的 VNN 进行评估的。接下来，我们检验了 VNN 得出的结果是否会因用于估计解剖协方差矩阵的 HC 组和 AD+ 组的合并数据构成变化而发生变化。为此，我们进行了两组实验。第一组实验中，我们纳入整个 HC 组，并逐步改变纳入 AD+组的人数，以构建 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$的估计值。图 21a 展示了随机选取的一个 VNN 模型所得结果，该模型对应的解剖协方差矩阵由 HC 组和 AD+ 组个体的不同组合构建而成。图 21a 中的结果显示，当 AD+ 组的区域残差高于 HC 组时，方差分析的 F 值（经 Bonferroni 校正的 p-值 < 0.05）。VNN 使用由全部 652 名 HC 个体和 209 名 AD+ 个体估计得到的 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$时，在图 7b 所示的 AD+ 组与 HC 组间 Δ-年龄差异的区域特征所包含的脑区中，呈现出最稳健的组间差异。通过减少用于估计协方差矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$的 AD+ 组个体数量来扰动该矩阵时，我们发现，排除约 134 名 AD+ 个体之前，所得结果仍然保持稳定。VNN 使用由 50 名或更少 AD+ 个体以及全部 652 名 HC 个体估计得到的 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$时，整体而言，所得结果的 F-值明显降低；其中，双侧胼胝体下区、海马旁回、颞极和内侧颞叶脑区的 F-值受影响尤为明显。因此，从解剖协方差矩阵中移除与 AD 相关的信息，会使 AD 特征性脑区中的结果显著性降低。

图21b展示了与上述类似实验的结果，不同之处在于：当通过减少用于估计解剖协方差矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$ 的HC组个体数量对其进行扰动时，评估了VNN的区域残差。以图21a中根据652名HC个体和209名阿尔茨海默病+ 个体估计得到的 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$ 的结果为基线，只要纳入至少100名HC组个体来构建 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$，阿尔茨海默病+ 组与HC组区域残差之间的方差分析结果（阿尔茨海默病+ 组的残差高于HC组）就保持一致。纳入 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$ 的HC个体少于100名时，左半球的楔前叶和缘上回区的结果显著性明显降低。

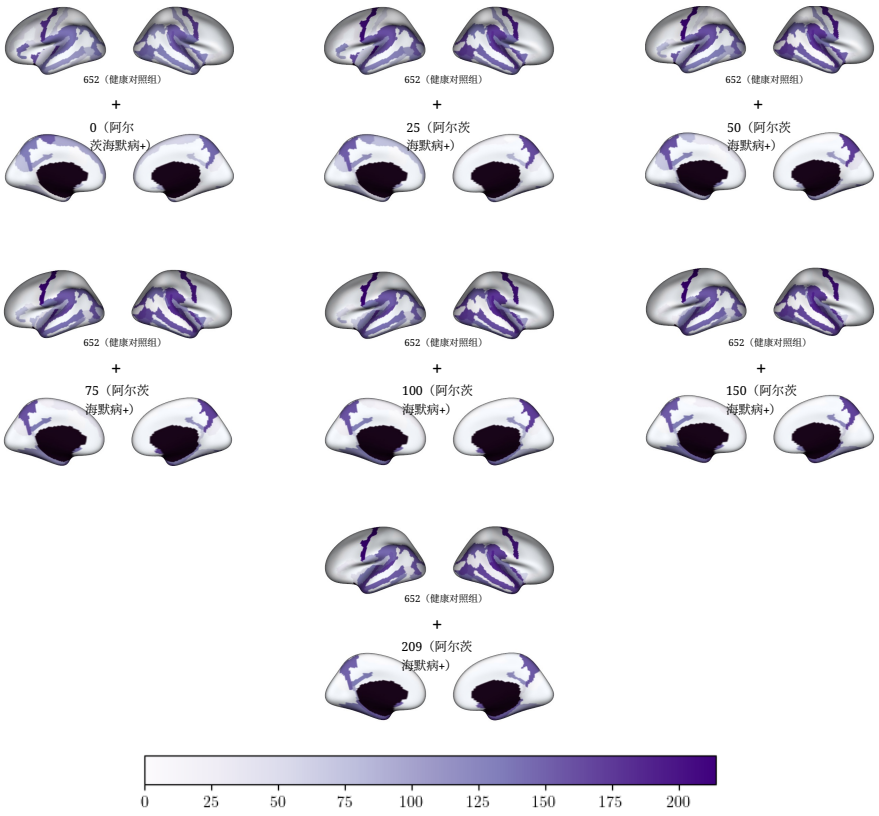
总之，在 OASIS-3 数据集中，阿尔茨海默病组+与健康对照组在区域残差方面观察到的组间差异，对协方差矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$的扰动具有稳健性；这些扰动是通过采用不同的健康对照组和阿尔茨海默病+ 个体组合来估计该矩阵，从基线引入的。然而，从 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$中（几乎）完全排除健康对照组或阿尔茨海默病+ 组，导致升高不再具有统计显著性。

AD+ 在多个区域的区域残差，包括双侧海马旁回和颞极区域，以及左半球的楔前叶和缘上回区。因此，健康对照组和 AD+ 组均与解剖协方差矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$相关，该矩阵得出了图 7 中的区域特征；但图 7 中的区域特征并未针对用于估算 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$ 的健康对照组和 AD+ 个体组合进行过拟合。

附录 M 自适应读出可能会损害区域残差和 Δ-年龄的可解释性

到目前为止，我们一直关注采用非自适应读出（未加权平均）函数的 VNN。然而，预计借助自适应读出函数，可以显著提升 VNN 在其原始任务——实际年龄预测——上的性能。我们的实验确实表明了这一点。如果在 VNN 中添加一个由 10 个神经元组成的单层全连接感知器，并使其架构与在 OASIS-3 数据集上训练的 VNN 相同，那么我们就能够提升其实际年龄预测任务的性能。对于在训练数据的随机排列上训练的 100 个采用自适应读出的 VNN，HC组的平均绝对误差中位数为 4.64 岁，显著小于采用非自适应读出的 VNN 所达到的平均绝对误差（表 1）。在这 100 个采用自适应读出的 VNN 模型中，我们分析了其中一个模型的区域残差；该模型在 HC组的实际年龄预测中表现最佳（测试集：平均绝对误差 $= 4.17$ 岁，预测值与真实值之间的皮尔逊相关系数为 $= 0.73$；完整 HC组：平均绝对误差 $= 4.26$ 岁，预测值与真实值之间的皮尔逊相关系数为 $= 0.725$）。我们的区域残差显示，AD+ 组与 HC组之间的区域残差没有显著差异。这一观察结果表明，由于加入了自适应读出函数，VNN 失去了其可解释性。此外，我们还观察到，使用该 VNN 模型确定的 AD+ 组与 HC组之间的 Δ-年龄差距有所缩小（AD组的Δ-年龄：1.58 ± 4.67 岁；HC组的 Δ-年龄：0 ± 3.45 岁；Cohen's $d = 0.384$）。这里讨论的研究结果表明，使用自适应读出函数提升实际年龄预测任务的性能，可能会损害采用非自适应读出的 VNN 所提供的可解释性，同时也会缩小 AD+ 组与 HC 组之间的 Δ- 年龄差距。

a



方差分析检验的 F 值

b

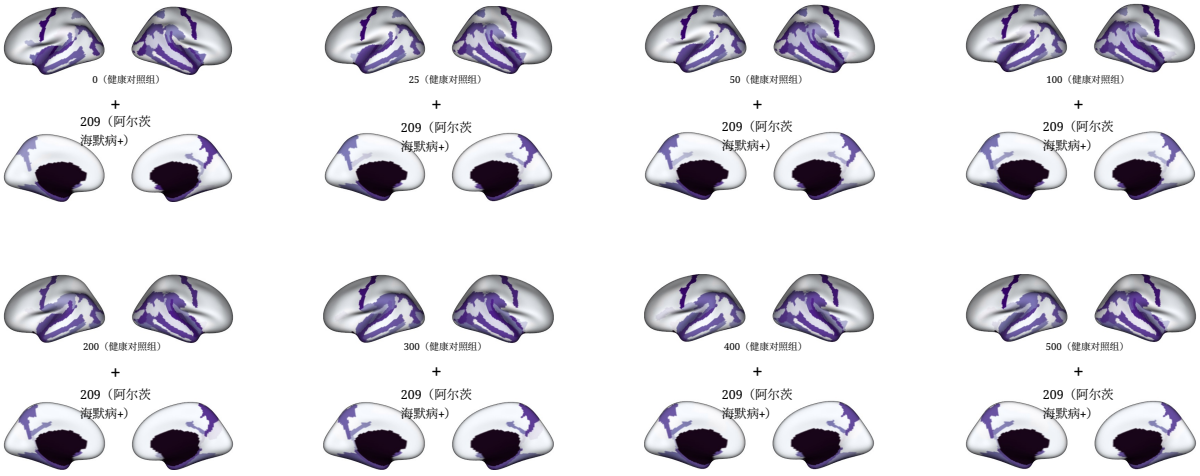


图21 （前页）：**AD+ 组与健康对照组在区域残差中观察到的组间差异，其解剖协方差矩阵对扰动的稳定性**。对于在 OASIS-3 数据集中针对健康对照组训练、用于预测实际年龄的 VNN 模型，首先使用由完整 OASIS-3 数据集（即 652 名健康对照组个体和 209 名 AD+ 组个体）的皮层厚度数据构成的解剖协方差矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$确定区域残差。根据《方法部分》第 4.4.1 小节所述流程，考察 AD+ 组与健康对照组之间区域残差的组间差异。在该流程中，我们评估了 AD+ 组与健康对照组区域残差之间方差分析检验的 F 值。将 AD+ 组相对于健康对照组升高的区域残差 F 值投影到图a中的脑模板上。我们还通过改变用于估计 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$的 AD+ 组和健康对照组皮层厚度数据的组成，进一步考察组间差异对扰动的稳定性。图**a** 展示了通过分析 VNN 对完整 OASIS-3 数据集皮层厚度数据进行处理所得的区域残差而获得的结果；该 VNN 使用基于 652 名健康对照组个体及数量不等的 AD+组个体估计的解剖协方差矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$。其中，使用 652 名健康对照组个体、但不包含 AD+ 组个体的皮层厚度数据估计 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$的情形构成了这些实验的一个极端。图**b** 展示了与图**a** 类似的实验结果，区别在于解剖协方差矩阵 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$使用 AD+ 组全部 209 名个体以及数量不等的健康对照组个体估计。使用 209 名 AD+ 组个体和 652 名健康对照组个体估计的 $\hat{\mathbf{C}}_{148}^{\mathrm{AD}+}$所对应的结果，作为图**a**和**b**中所有情形的比较基准。

参考文献

[1] S. Sihag、G. Mateos、C. McMillan 和 A. Ribeiro，“协方差神经网络”，载于神经信息处理系统会议论文集，2022年11月。[2] ——，“使用可迁移协方差神经网络预测脑龄”，载于*IEEE* 声学、语音与信号处理国际会议论文集，2023年6月。[3] P. Liu，“遥感大数据综述”，环境科学前沿，第3卷，第45页，2015年。[4] B. H. Brinkmann、M. R. Bower、K. A. Stengel、G. A. Worrell 和 M. Stead，“大规模电生理学：大数据的采集、压缩、加密与存储”，神经科学方法杂志，第180卷，第1期，第185–192页，2009年。[5] J. Cai、J. Luo、S. Wang 和 S. Yang，“机器学习中的特征选择：一种新视角”，*Neurocomputing*，第300卷，第70–79页，2018年。[6] A. Y. Ng，“特征选择、l 1 与 l 2 正则化及旋转不变性”，载于第二十一届国际机器学习会议论文集，2004年，第78页。[7] Y. Yang、Y. Zhu、H. Cui、X. Kan、L. He、Y. Guo 和 C. Yang，“通过多任务元学习实现数据高效的脑连接组分析”，*arXiv* 预印本 *arXiv:2206.04486*，2022年。[8] R. D. Markello、J. Y. Hansen、Z.-Q. Liu、V. Bazinet、G. Shafiei、L. E. Su´árez、N. Blostein、J. Seidlitz、S. Baillet、T. D. Satterthwaite 等，“Neuromaps：脑图谱的结构与功能解释”，自然·方法，第19卷，第11期，第1472–1479页，2022年。[9] A. Schaefer、R. Kong、E. M. Gordon、T. O. Laumann、X.-N. Zuo、A. J. Holmes、S. B. Eickhoff 和 B. T. Yeo，“基于内在功能连接 MRI 对人类大脑皮层进行局部—整体分区”，《大脑皮层》，第28卷，第9期，第3095–3114页，2018年。[10] P. Hagmann、L. Cammoun、X. Gigandet、R. Meuli、C. J. Honey、V. J. Wedeen 和 O. Sporns，“绘制人类大脑皮层的结构核心”，*PLoS Biology*，第6卷，第7期，文章 e159，2008年。[11] M. Habes、D. Janowitz、G. Erus、J. Toledo、S. Resnick、J. Doshi、S. Van der Auwera、K. Wittfeld、K. Hegenscheid、N. Hosten等，“脑衰老加速：与流行病学和遗传风险因素的关系，以及与阿尔茨海默病萎缩模式的重叠”，转化精神病学，第6卷，第4期，第 e775–e775 页，2016年。[12] K. Franke 和 C. Gaser，“健康衰老、轻度认知障碍和阿尔茨海默病中个体脑龄的纵向变化。”老年心理学与老年精神病学杂志 *GeroPsych*，第25卷，第4期，第235页，2012年。[13] T. Hajek、K. Franke、M. Kolenic、J. Capkova、M. Matejka、L. Propper、R. Uher、P. Stopkova、T. Novak、T. Paus等，“双相障碍或精神分裂症早期阶段的脑龄”，《精神分裂症通报》，第45卷，第1期，第190–198页，2019年。[14] J. H. Cole 和 K. Franke，“基于神经影像学预测年龄：创新型脑衰老生物标志物”，神经科学趋势，第40卷，第12期，第681–690页，2017年。[15] C. Yin、P. I mms、M. Cheng、A. Amgalan、N. F. Chowdhury、R. J. Massett、N. N. Chaudhari、X. Chen、P. M. Thompson、P. Bogdan 等，“具有解剖学可解释性的脑龄深度学习捕捉到领域特异性认知障碍”，美国国家科学院院刊，第120卷，第2期，文章 e2214634120，2023年。[16] V. M. Bashyam、G. Erus、J. Doshi、M. Habes、I. M. Nasrallah、M. Truelove-Hill、D. Srinivasan、L. Mamourian、R. Pomponio、Y. Fan 等，“基于深度脑网络及全球 14 468 名个体的全生命周期脑龄与疾病磁共振成像特征”，《脑》，第143卷，第7期，第2312–2324页，2020年。[17] B. Couvy-Duchesne、J. Faouzi、B. Martin、E. Thibeau-Sutre、A. Wild、M. Ansart、S. Durrleman、D. Dormont、N. Burgos 和 O. Colliot，“卷积神经网络、支持向量机和最佳线性无偏

脑龄预测器：Aramis 对 2019 年预测分析竞赛挑战的贡献，《Frontiers in Psychiatry》，第 11 卷，第 593336 页，2020年。[18] E. R. Butler、A. Chen、R. Ramadan、T. T. Le、K. Ruparel、T. M. Moore、T. D. Satterthwaite、F. Zhang、H. Shou、R. C. Gur 等，“脑龄分析中的陷阱”，Wiley Online Library，技术报告，2021年。[19] J. Zhou、G. Cui、S. Hu、Z. Zhang、C. Yang、Z. Liu、L. Wang、C. Li 和 M. Sun，“图神经网络：方法与应用综述”，AI Open，第 1 卷，第 57–81 页，2020年。[20] F. Scarselli、M. Gori、A. C. Tsoi、M. Hagenbuchner 和 G. Monfardini，“图神经网络模型”，IEEE Transactions on Neural Networks，第 20 卷，第 1 期，第 61–80 页，2008年。[21] Z. Wu、S. Pan、F. Chen、G. Long、C. Zhang 和 P. S Yu，“图神经网络综合综述”，IEEE Transactions on Neural Networks and Learning Systems，第 32 卷，第 1 期，第 4–24 页，2020年。[22] S. Sahu、A. Mhedhbi、S. Salihoglu、J. Lin 和 M. T. ¨Ozsu，“大型图无处不在及图处理面临的意外挑战”，Proceedings of the VLDB Endowment，第 11 卷，第 4 期，第 420–431 页，2017年。[23] V. Quintana、A. Simoes-Costa 和 A. Mandel，“采用直接图论方法实现电力系统拓扑可观测性”，IEEE Transactions on Power Apparatus and Systems，第 3 期，第 617–626 页，1982年。[24] S. A. Myers、A. Sharma、P. Gupta 和 J. Lin，“信息网络还是社交网络？Twitter关注图的结构”，载于《第23届万维网国际会议论文集》，2014年，第 493–498 页。[25] A. Fout、J. Byrd、B. Shariat 和 A. Ben-Hur，“利用图卷积网络预测蛋白质界面”，《神经信息处理系统进展》，第 30 卷，2017年。[26] O. Sporns，“图论方法：在脑网络中的应用”，Dialogues in Clinical Neuroscience，2022 年。[27] A. Sanchez-Gonzalez、N. Heess、J. T. Springenberg、J. Merel、M. Riedmiller、R. Hadsell 和 P. Battaglia，“将图网络作为可学习的物理引擎，用于推断和控制”，载于《国际机器学习会议》。PMLR，2018年，第 4470–4479 页。[28] S. Abadal、A. Jain、R. Guirado、J. L´opez-Alonso 和 E. Alarc´on，“图神经网络计算：从算法到加速器的综述”，ACM Computing Surveys (CSUR)，第 54 卷，第 9 期，第 1–38 页，2021年。[29] S. Zhang、H. Tong、J. Xu 和 R. Maciejewski，“图卷积网络：综合综述”，Computational Social Networks，第 6 卷，第 1 期，第 1–23 页，2019年。[30] T. N. Kipf 和 M. Welling，“变分图自编码器”，arXiv预印本 arXiv:1611.07308，2016年。[31] Y. Li、R. Zemel、M. Brockschmidt 和 D. Tarlow，“门控图序列神经网络”，载于《ICLR’16论文集》，2016年。[32] J. Bruna、W. Zaremba、A. Szlam 和 Y. LeCun，“图上的谱网络与局部连接网络”，arXiv预印本 arXiv:1312.6203，2013年。[33] D. K. Hammond、P. Vandergheynst 和 R. Gribonval，“基于谱图理论的图小波”，《应用与计算调和分析》，第 30 卷，第 2 期，第 129–150 页，2011年。[34] F. Gama、E. Isufi、G. Leus 和 A. Ribeiro，“图、卷积与神经网络：从图滤波器到图神经网络”，IEEE Signal Processing Magazine，第 37 卷，第 6 期，第 128–138 页，2020年。[35] J. Atwood 和 D. Towsley，“扩散卷积神经网络”，《神经信息处理系统进展》，第 29 卷，2016 年。[36] S. Verma 和 Z.-L. Zhang，“ 图卷积神经网络的稳定性与泛化 ”，载于《第 25 届 ACM SIGKDD知识发现与数据挖掘国际会议论文集》，2019年，第 1539–1548 页。

[37] N. Keriven、A. Bietti 和 S. Vaiter，“大型随机图上图卷积网络的收敛性与稳定性”，《神经信息处理系统进展》，第 33 卷，第 21 512–21 523 页，2020年。[38] F. Gama、J. Bruna 和 A. Ribeiro，“图神经网络的稳定性性质”，IEEE Transactions on Signal Processing，第 68 卷，第 5680–5695 页，2020年。[39] L. Ruiz、L. Chamon 和 A. Ribeiro，“图子神经网络与图神经网络的可迁移性”，《神经信息处理系统进展》，第 33 卷，第 1702–1712 页，2020年。[40] A. C. Evans，“解剖协方差网络”，NeuroImage，第 80 卷，第 489–504 页，2013年。[41] H. Murase 和 S. K. Nayar，“利用参数特征空间进行目标识别的照明规划”，IEEE模式分析与机器智能汇刊，第 16 卷，第 12 期，第 1219–1227 页，1994年。[42] F. De la Torre 和 M. J. Black，“用于计算机视觉的鲁棒主成分分析”，载于《第八届IEEE国际计算机视觉会议论文集。ICCV 2001》，第 1 卷。IEEE，2001年，第 362–369 页。[43] D. Stephenson，“空间气候/天气图的相关性及在预测中使用马氏度量的优势”，Tellus A，第 49 卷，第 5 期，第 513–527 页，1997年。[44] H. Shao、W. H. Lam、A. Sumalee、A. Chen 和 M. L. Hazelton，“根据逐日交通计数估计高峰时段起讫点需求的均值和协方差”，交通研究B辑：方法论，第 68 卷，第 52–75 页，2014年。[45] M. N. Ismail、A. Aborujilah、S. Musa 和 A. Shahzad，“利用协方差矩阵方法检测云计算环境中的基于洪泛的拒绝服务攻击”，载于《第七届普适信息管理与通信国际会议论文集》，2013年，第 1–6 页。[46] I. T. Joliffe 和 B. Morgan，“主成分分析与探索性因子分析”，医学研究统计方法，第 1 卷，第 1 期，第 69–95 页，1992年。[47] E. Elhaik，“基于主成分分析（PCA）的群体遗传学研究结果存在严重偏差，必须重新评估”，Scientific Reports，第 12 卷，第 1 期，第 1–35 页，2022年。[48] R. Levie、W. Huang、L. Bucci、M. M. Bronstein 和 G. Kutyniok，“谱图卷积神经网络的可迁移性”。J. Mach. Learn. Res.，第 22 卷，第 272–1 页，2021年。[49] Q. Zhu、C. Yang、Y. Xu、H. Wang、C. Zhang 和 J. Han，“利用自我图信息最大化实现图神经网络迁移学习”，《神经信息处理系统进展》，第 34 卷，第 1766–1779 页，2021年。[50] S. Maskey、R. Levie 和 G. Kutyniok，“图神经网络的可迁移性：一种扩展Graphon方法”，《应用与计算调和分析》，第 63 卷，第 48–83 页，2023年。[51] Y. Zeighami 和 A. C. Evans，“关联与预测：皮层表面平滑和分区对脑龄的影响”，Frontiers in Big Data，第 4 卷，第 637724 页，2021年。[52] A. I. Luppi 和 E. A. Stamatakis，“结合网络拓扑与信息论构建代表性脑网络”，Network Neuroscience，第 5 卷，第 1 期，第 96–124 页，2021年。[53] J. Royer、R. Rodr´ıguez-Cruces、S. Tavakol、S. Larivi`ere、P. Herholz、Q. Li、R. Vos de Wael、C. Paquola、O. Benkarim、B.-y. Park 等，“用于多尺度神经科学的开放MRI数据集”，Scientific Data，第 9 卷，第 1 期，第 1–12 页，2022年。[54] B. S. Khundrakpam、J. Tohka、A. C. Evans、B. D. C. Group 等，“基于不同空间分辨率下皮层厚度预测脑成熟度”，NeuroImage，第 111 卷，第 350–359 页，2015年。[55] C. L´opez-Ot´ın 等，“衰老标志”，Cell，第 153 卷，第 6 期，第 1194–1217 页，2013年。[56] G. B. Frisoni、N. C. Fox、C. R. Jack、P. Scheltens 和 P. M. Thompson，“结构性MRI在阿尔茨海默病临床应用中的作用”，Nature Reviews Neurology，第 6 卷，第 2 期，第 67–77 页，2010年。

[57] J. J. Chen，“衰老和神经退行性疾病中的脑生理学功能性磁共振成像”，*NeuroImage*，第187卷，第209–225页，2019年。[58] J. Lee、B. J. Burkett、H.-K. Min、M. L. Senjem、E. S. Lundt、H. Botha、J. Graff-Radford、L. R. Barnard、J. L. Gunter、C. G. Schwarz 等，“正常衰老和痴呆中的基于深度学习的脑龄预测”，*Nature Aging*，第2卷，第5期，第412–424页，2022年。[59] W. Yin、L. Li 和 F.-X. Wu，“基于功能性磁共振成像图像的脑疾病诊断深度学习”，*Neurocomputing*，第469卷，第332–345页，2022年。[60] P. R. Millar、B. A. Gordon、P. H. Luckett、T. Benzinger、C. Cruchaga、A. M. Fagan、J. Hassenstab、R. J. Perrin、S. E. Schindler、R. F. Allegri 等，“多模态脑龄估计值与早期阿尔茨海默病生物标志物及认知的关联：一项横断面观察性研究”，*bioRxiv*，2022年。[61] J. Dafflon、W. H. Pinaya、F. Turkheimer、J. H. Cole、R. Leech、M. A. Harris、S. R. Cox、H. C. Whalley、A. M. McIntosh 和 P. J. Hellyer，“一种基于自动化机器学习的方法：根据皮层解剖测量预测脑龄”，*HumanBrain Mapping*，第41卷，第13期，第3555–3566页，2020年。[62] L. Baecker、J. Dafflon、P. F. Da Costa、R. Garcia-Dias、S. Vieira、C. Scarpazza、V. D. Calhoun、J. R. Sato、A. Mechelli 和 W. H. Pinaya，“脑龄预测：使用基于区域和基于体素的形态测量数据对机器学习模型进行比较”，*Human Brain Mapping*，第42卷，第8期，第2332–2346页，2021年。[63] H. M. Aycheh、J.-K. Seong、J.-H. Shin、D. L. Na、B. Kang、S. W. Seo 和 K.-A. Sohn，“利用皮层厚度数据进行生物学脑龄预测：一项大规模队列研究”，*Frontiers in Aging Neuroscience*，第10卷，第252页，2018年。[64] P. J. LaMontagne、T. L. Benzinger、J. C. Morris、S. Keefe、R. Hornbeck、C. Xiong、E. Grant、J. Hassenstab、K. Moulder、A. G. Vlassenko等，“OASIS-3：用于正常衰老和阿尔茨海默病的纵向神经影像学、临床及认知数据集”，*MedRxiv*，2019年。[65] S. M. McGinnis、M. Brickhouse、B. Pascual 和 B. C. Dickerson，“人类皮层区域厚度的年龄相关变化”，*Brain Topography*，第24卷，第3期，第279–291页，2011年。[66] M. Lehmann、S. J. Crutch、G. R. Ridgway、B. H. Ridha、J. Barnes、E. K. Warrington、M. N. Rossor 和 N. C. Fox，“后部皮层萎缩和典型阿尔茨海默病中的皮层厚度与基于体素的形态测量”，《衰老神经生物学》，第32卷，第8期，第1466–1476页，2011年。[67] M. R. Sabuncu、R. S. Desikan、J. Sepulcre、B. T. T. Yeo、H. Liu、N. J. Schmansky、M. Reuter、M. W. Weiner、R. L. Buckner、R. A. Sperling 等，“阿尔茨海默病中皮层和海马萎缩的动态变化”，*Archives of Neurology*，第68卷，第8期，第1040–1048页，2011年。[68] S. G. Popescu、B. Glocker、D. J. Sharp 和 J. H. Cole，“局部脑龄：一种 U-Net 模型”，*Frontiers in Aging Neuroscience*，第13卷，第761954页，2021年。[69] X. Feng、Z. C. Lipton、J. Yang、S. A. Small、F. A. Provenzano、阿尔茨海默病神经影像学倡议、F. L. D. N. Initiative 等，“利用深度学习和结构性磁共振成像，根据统一健康人群估计脑龄”，《衰老神经生物学》，第91卷，第15–25页，2020年。[70] X. Feng、J. Yang、Z. C. Lipton、S. A. Small、F. A. Provenzano、阿尔茨海默病神经影像学倡议等，“基于磁共振成像的深度学习证实海马结构在阿尔茨海默病分类中的突出作用”，*bioRxiv*，第456277页，2018年。[71] A. Essemlali、E. St-Onge、M. Descoteaux 和 P.-M. Jodoin，“通过可解释人工智能理解阿尔茨海默病的结构连接”，载于*Medical Imaging with Deep Learning*。PMLR，2020年，第217–229页。[72] H. T. Karim、M. Ly、G. Yu、R. Krafty、D. L. Tudorascu、H. J. Aizenstein 和 C. Andreescu，“加速衰老：晚年的担忧和反刍思维与脑龄偏高有关”，*Neurobiology of Aging*，第101卷，第13–21页，2021年。

[73] F. Liem、G. Varoquaux、J. Kynast、F. Beyer、S. K. Masouleh、J. M. Huntenburg、L. Lampe、M. Rahim、A. Abraham、R. C. Craddock 等，“从多模态成像数据预测脑龄可捕捉认知障碍”，NeuroImage，第148卷，第179–188页，2017年。[74] J. Shlens，“主成分分析教程”，arXiv 预印本 arXiv:1404.1100，2014年。[75] A. Ortega、P. Frossard、J. Kovaˇcevi´c、J. M. Moura 和 P. Vandergheynst，“图信号处理：概述、挑战与应用”，Proceedings of the IEEE，第106卷，第5期，第808–828页，2018年。[76] D. I. Shuman、S. K. Narang、P. Frossard、A. Ortega 和 P. Vandergheynst，“图上信号处理的新兴领域：将高维数据分析扩展到网络和其他非规则域”，IEEE signal processing maga- zine，第30卷，第3期，第83–98页，2013年。[77] A. Sandryhaila 和 J. M. Moura，“图上的离散信号处理”，IEEE transactions on signal processing，第61卷，第7期，第1644–1656页，2013年。[78] A. Loukas，“样本协方差矩阵的特征向量与实际协方差矩阵的特征向量有多接近？”，载于《国际机器学习会议》。PMLR，2017年，第2228–2237页。[79] L. Ruiz、L. F. Chamon 和 A. Ribeiro，“图子信号处理”，IEEE Transactions on Signal Processing，第69卷，第4961–4976页，2021年。[80] C. Borgs、J. T. Chayes、L. Lov´asz、V. T. S´os 和 K. Vesztergombi，“稠密图的收敛序列 I：子图频率、度量性质与测试”，Advances in Mathematics，第219卷，第6期，第1801–1851页，2008年。[81] L. Lov´asz，《大型网络与图极限》。A merican Mathematical Soc.，2012年，第60卷。[82] K. Franke 和 C. Gaser，“脑龄作为脑衰老神经影像生物标志物的十年发展：我们获得了哪些认识？”，Frontiers in Neurology，第789页，2019年。[83] M. Jov´e、M. Portero-Ot´ın、A. Naud´ı、I. Ferrer 和 R. Pamplona，“人类脑衰老及年龄相关神经退行性疾病的代谢组学”，Journal of Neuropathology & Experimental Neurology，第73卷，第7期，第640–657页，2014年。[84] C. Destrieux、B. Fischl、A. Dale 和 E. Halgren，“使用标准解剖学命名法自动划分人类大脑皮层脑回和脑沟”，NeuroImage，第53卷，第1期，第1–15页，2010年。[85] S. E. O’Bryant、S. C. Waring、C. M. Cullum、J. Hall、L. Lacritz、P. J. Massman、P. J. Lupo、J. S. Reisch、R. Doody、T. Alzheimer’s Research Consortium 等，“使用临床痴呆评定量表方框总分对痴呆进行分期：得克萨斯州阿尔茨海默病研究联盟的一项研究”，Archives of Neurology，第65卷，第8期，第1091–1095页，2008年。[86] I. Beheshti、S. Nugent、O. Potvin 和 S. Duchesne，“神经影像脑龄框架中的偏差校正：一种稳健方案”，NeuroImage: Clinical，第24卷，第102063页，2019年。[87] T. T. Le、R. T. Kuplicki、B. A. McKinney、H.-W. Yeh、W. K. Thompson、M. P. Paulus 和 T. . Investigators，“非线性模拟框架支持在分析脑龄时进行年龄校正”，Frontiers in Aging Neuroscience，第10卷，第317页，2018年。[88] A.-M. G. de Lange 和 J. H. Cole，“评论：脑龄预测中的校正程序”，NeuroImage: Clinical，第26卷，2020年。[89] N. J. Tustison、P. A. Cook、A. Klein、G. Song、S. R. Das、J. T. Duda、B. M. Kandel、N. van Strien、J. R. Stone、J. C. Gee 等，“对 ANTs 和 FreeSurfer 皮层厚度测量进行大规模评估”，NeuroImage，第99卷，第166–179页，2014年。[90] S. R. Das、B. B. Avants、M. Grossman 和 J. C. Gee，“基于配准的皮层厚度测量”，NeuroImage，第45卷，第3期，第867–879页，2009年。

[91] J. C. Morris，“临床痴呆评定量表（CDR）：当前版本与评分规则。”《神经学》，1993年。[92] A. Paszke、S. Gross、F. Massa、A. Lerer、J. Bradbury、G. Chanan、T. Killeen、Z. Lin、N. Gimelshein、L. Antiga 等，“PyTorch：一种命令式风格的高性能深度学习库”，《神经信息处理系统进展》，第32卷，2019年。[93] T. Akiba、S. Sano、T. Yanase、T. Ohta 和 M. Koyama，“Optuna：新一代超参数优化框架”，载于2019年ACM SIGKDD知识发现与数据挖掘国际会议第25届会议论文集，第2623–2631页。[94] T. Sch¨afer 和 C. Ecker，“fsbrain：用于结构神经影像数据可视化的R语言软件包”，bioRxiv，第2020–09页，2020年。[95] S. Wein、W. M. Malloni、A. M. Tom´e、S. M. Frank、G.-I. Henze、S. W¨ust、M. W. Greenlee 和 E. W. Lang，“用于脑网络因果推断的图神经网络框架”，Scientific Reports，第11卷，第1期，第1–18页，2021年。[96] Z. Ardalan 和 V. Subbian，“神经影像分析中的迁移学习方法：范围综述”，《人工智能前沿》，第5卷，2022年。[97] O. Lindberg、E. Westman、S. Karlsson、P. ¨Ostberg、L. A. Svensson、A. Simmons 和 L.-O. Wahlund，“胼胝体下内侧前额叶皮层是阿尔茨海默病和额颞叶变性的共同萎缩部位吗？”Frontiers in Aging Neuroscience，第4卷，第32页，2012年。[98] R. Ossenkoppele、Y. A. Pijnenburg、D. C. Perry、B. I. Cohn-Sheehy、N. M. Scheltens、J. W. Vogel、J. H. Kramer、A. E. Van Der Vlies、R. L. Joie、H. J. Rosen 等，“阿尔茨海默病的行为型/执行功能障碍型变异：临床、神经影像学和病理特征”，《脑》，第138卷，第9期，第2732–2749页，2015年。[99] D. Ferreira、A. Nordberg 和 E. Westman，“阿尔茨海默病的生物学亚型：系统综述与荟萃分析”，《神经学》，第94卷，第10期，第436–448页，2020年。[100] O. Hansson、M. J. Grothe、T. O. Strandberg、T. Ohlsson、D. H¨agerstr¨om、J. J¨ogi、R. Smith 和 M. Sch¨oll，“阿尔茨海默病中tau蛋白病理的分布与认知相关功能性脑网络呈差异性对应”，《神经科学前沿》，第11卷，第167页，2017年。[101] S.-Y. Ryu、M. J. Kwon、S.-B. Lee、D. W. Yang、T.-W. Kim、I.-U. Song、P. S. Yang、H. J. Kim 和 A. Y. Lee，“使用磁共振容积测量阿尔茨海默病患者楔前叶和海马体积”，《临床神经病学杂志》，第6卷，第4期，第196–203页，2010年。[102] C. Ech´avarri、P. Aalten、H. B. Uylings、H. Jacobs、P. J. Visser、E. Gronenschild、F. Verhey 和 S. Burgmans，“海马旁回萎缩作为阿尔茨海默病的早期生物标志物”，《脑结构与功能》，第215卷，第3期，第265–271页，2011年。[103] G. Karas、P. Scheltens、S. Rombouts、R. Van Schijndel、M. Klein、B. Jones、W. Van Der Flier、H. Vrenken 和 F. Barkhof，“早发性阿尔茨海默病中的楔前叶萎缩：一项形态测量结构磁共振成像研究”，《神经放射学》，第49卷，第12期，第967–976页，2007年。[104] J. Bien 和 R. J. Tibshirani，“协方差矩阵的稀疏估计”，《生物统计学》，第98卷，第4期，第807–820页，2011年。[105] R. K. Lenroot 和 J. N. Giedd，“儿童和青少年的脑发育：来自解剖磁共振成像的见解”，《神经科学与生物行为评论》，第30卷，第6期，第718–729页，2006年。[106] P. Reardon、J. Seidlitz、S. Vandekar、S. Liu、R. Patel、M. T. M. Park、A. Alexander-Bloch、L. S. Clasen、J. D. Blumenthal、F. M. Lalonde 等，“人类正常脑大小变异与脑形态多样性”，《科学》，第360卷，第6394期，第1222–1227页，2018年。

[107] C. Coffey、W. Wilkinson、L. Parashos、S. Soady、R. Sullivan、L. Patterson、G. Figiel、M. Webb、C. Spritzer 和 W. Djang，“衰老的人脑的定量脑解剖学：一项使用磁共振成像的横断面研究”，《神经学》，第42卷，第3期，第527–527页，1992年。[108] N. Raz、F. M. Gunning、D. Head、J. H. Dupuis、J. McQuain、S. D. Briggs、W. J. Loken、A. E. Thornton 和 J. D. Acker，“体内观察到的人类大脑皮层选择性衰老：前额叶灰质的差异性易损性。”《大脑皮层》（纽约，纽约州：*1991*），第7卷，第3期，第268–282页，1997年。[109] J. Polzehl 和 K. Tabelow，“弥散磁共振成像模型中的低信噪比”，《美国统计学会杂志》，第111卷，第516期，第1480–1490页，2016年。[110] C. M. Bennett 和 M. B. Miller，“功能性磁共振成像的结果有多可靠？”《纽约科学院年刊》，第1191卷，第1期，第133–155页，2010年。[111] L. Ronan、A. F. Alexander-Bloch、K. Wagstyl、S. Farooqi、C. Brayne、L. K. Tyler、P. C. Fletcher 等，“肥胖与中年期脑龄增加相关”，《衰老神经生物学》，第47卷，第63–70页，2016年。[112] K. Mareckova、R. Marecek、L. Andryskova、M. Brazdil 和 Y. S. Nikolova，“妊娠期间的母亲抑郁症状与年轻成年子代的脑龄：产前出生队列研究结果”，《大脑皮层》，第30卷，第7期，第3991–3999页，2020年。[113] L. T. Westlye、I. Reinvang、H. Rootwelt 和 T. Espeseth，“APOE对健康成年人脑白质微观结构的影响”，《神经学》，第79卷，第19期，第1961–1969页，2012年。[114] S. Horvath，“人体组织和细胞类型的DNA甲基化年龄”，《基因组生物学》，第14卷，第10期，第1–20页，2013年。[115] G. H. Golub 和 C. F. Van Loan，矩阵计算。约翰·霍普金斯大学出版社，2013年。[116] R. Vershynin，高维概率：数据科学中的导论与应用。剑桥大学出版社，2018年，第47卷。[117] Y. Yu、T. Wang 和 R. J. Samworth，“统计学家适用的Davis–Kahan定理变体”，《生物统计学》，第102卷，第2期，第315–323页，2015年。[118] A. Seelmann，“关于正弦 2Θ 定理的札记”，《积分方程与算子理论》，第79卷，第4期，第579–597页，2014年。[119] N. Alon 和 A. Naor，“通过Grothendieck不等式逼近割范数”，载于第三十六届年度*ACM*计算理论研讨会论文集，2004年，第72–80页。[120] R. S. Desikan、F. S´egonne、B. Fischl、B. T. Quinn、B. C. Dickerson、D. Blacker、R. L. Buckner、A. M. Dale、R. P. Maguire、B. T. Hyman 等，“一种自动标记系统，用于将MRI扫描图像中的人类大脑皮层划分为基于脑回的感兴趣区域”，*NeuroImage*，第31卷，第3期，第968–980页，2006年。