arXiv:2601.01229v1 [eess.SP] 3 Jan 2026

NeuroSSM：面向上下文感知功能磁共振成像分析的多尺度差分状态空间建模

Furkan Genç[1,2]，Boran İsmet Macun[1,2]，Sait Sarper Özaslan[1,2]，EmineU.Saritas[1,2,3]，和 Tolga Çukur[1,2,3,*]

[1]比尔肯特大学电气与电子工程系，土耳其安卡拉 06800[2]比尔肯特大学国家磁共振研究中心（UMRAM），土耳其安卡拉 06800[3]比尔肯特大学神经科学系，土耳其安卡拉 06800[*]通讯作者：cukur@ee.bilkent.edu.tr

摘要

准确的功能磁共振成像分析需要关注多个尺度上的时间结构，因为血氧水平依赖信号编码了认知过程，而这些过程从快速瞬态动力学到较慢的大尺度波动均有体现。现有用于时间建模的深度学习（DL）方法，在联合捕捉长功能性磁共振成像时间序列中的这些动力学方面面临挑战。在当前的深度学习模型中，Transformer 通过注意力机制显式建模成对交互，以处理长程依赖关系；但其相关的二次计算成本限制了在长功能磁共振成像序列中有效整合时间依赖关系的能力。相比之下，选择性状态空间模型（SSM）通过动力系统中的潜在状态演化，隐式建模长程时间依赖关系，从而能够高效地随时间传播依赖信息。然而，近期基于状态空间模型的功能磁共振成像方法通常基于衍生的功能连接表征进行处理，并采用单尺度时间处理。这些设计选择限制了单个模型联合表征快速瞬态动力学与较慢全局趋势的能力。我们提出 NeuroSSM，这是一种选择性状态空间架构，旨在对功能性磁共振成像时间序列中的原始血氧水平依赖信号进行端到端分析。NeuroSSM 通过两个互补的设计组件克服上述局限：一个能够并行捕捉快速与缓慢动力学的多尺度状态空间骨干网络，以及一个能够提高对瞬态状态变化敏感度的并行差分分支。在临床和非临床数据集上的实验表明，与最先进的功能磁共振成像分析方法相比，NeuroSSM 取得了具有竞争力的性能和效率。

关键词：功能磁共振成像、时间序列、深度学习、状态空间模型、认知、帕金森病

1. 引言

功能磁共振成像（fMRI）可无创测量血氧水平依赖（BOLD）信号，这些信号间接反映分布于整个大脑的神经活动 [1]。因此，人们广泛分析任务态 [2, 3, 4]或静息态 [5, 6] 功能性磁共振成像扫描记录的脑活动，以刻画与正常脑功能和神经系统疾病相关的认知过程 [7, 8]。这些认知过程涉及多个时间尺度上的神经活动，从快速瞬态动力学到较慢、全局协调的变化趋势 [9]。然而，功能磁共振成像无法直接观测这些活动：神经活动会经过缓慢且具有空间异质性的血流动力学反应滤波，并以相对较粗的时间分辨率进行采样 [10, 11]。因此，观测到的血氧水平依赖信号呈现长程时间依赖关系，且快速与缓慢成分之间存在显著分离 [9]。因此，有效的功能磁共振成像

分析需要采用能够整合较长BOLD时间序列信息，同时对短暂的状态变化保持敏感，并能在神经影像学研究常见的有限样本量下保持稳健的模型。

许多现有的功能磁共振成像分析深度学习（DL）方法，都隐含地依赖于与这些要求相冲突的建模假设。卷积、基于图的和循环模型已广泛应用于功能磁共振成像 [12,13, 14], , 但它们通常基于血氧水平依赖信号提取的功能连接（FC）特征进行建模 [15, 16, 17, 18]。虽然功能连接表征能够降低维度和噪声，但它们通常通过时间平均相关性度量计算得出，而这类度量假设数据具有平稳性和线性交互，从而抹去了时间顺序 [19]。这可能会掩盖原始血氧水平依赖信号中存在的非线性和时变动力学 [20, 21]。因此，这类方法以牺牲时间分辨率为代价换取计算上的可行性，限制了其捕捉长时间序列中时间结构的能力。

近年来，人们采用基于注意力的变换器架构，通过显式建模时间维度上的交互，来处理功能磁共振成像中的长程时间依赖关系 [22, 23, 24, 25]。在这些模型中，每个时间点都被视为一个离散词元，自注意力则根据词元相似性捕捉成对依赖关系。然而，这种建模方式存在一些局限，在功能磁共振成像分析中尤为突出。自注意力的二次方计算和内存开销，限制了对长血氧水平依赖序列进行有效整合的能力 [26]。更根本的是，高阶和多尺度时间结构必须从成对交互中间接推断，而非显式表示；即使通过稀疏化或近似提高了注意力机制的计算效率，这一局限依然存在 [25, 27]。

在当前的功能磁共振成像分析模型中，如何在捕捉潜在时间动态的同时，对长时间序列进行稳定且计算上切实可行的处理，仍然颇具挑战。状态空间模型（SSM）通过以潜在状态动力学描述信号演化，为建模时间结构提供了一个有原则的框架 [28]。然而，将状态空间模型应用于功能磁共振成像，还带来了额外的设计问题：如何在多个时间尺度上表示时间结构。近期基于状态空间模型的功能磁共振成像研究探索了一些特定的设计选择 [29, 30, 31], , 通常侧重于基于连接的表征和单尺度时间建模。尽管这些方法证明了状态空间模型适用于功能磁共振成像分析，但并未明确解决如何在单一模型中联合表示快速瞬态动力学与较慢的全局趋势。

在此，我们介绍 NeuroSSM，这是一种选择性状态空间模型，通过两项互补的设计组件实现高保真fMRI分析。首先，NeuroSSM采用多尺度状态空间骨干网络，在完整时间分辨率和较粗尺度上处理血氧水平依赖信号，从而能够同时建模快速瞬态动力学与较慢的全局趋势。其次，并行差分分支将血氧水平依赖信号的一阶时间导数纳入状态空间模型，从而在保留长程时间整合能力的同时，提高对瞬态状态变化的敏感度。这些组件相结合，使 NeuroSSM 能够在多个尺度上编码时间结构，同时使计算成本与序列长度保持线性关系。

我们在两个公开数据集上评估了 NeuroSSM：人类连接组计划（HCP） [32], , 包含静息态和任务态功能磁共振成像数据；以及帕金森病进展标志物计划（PPMI） [33], , 旨在区分帕金森病患者与对照组。在这两个数据集上，与经典深度学习、Transformer 和状态空间模型基线方法相比，NeuroSSM 均展现出有竞争力的预测性能。消融研究进一步证实，多尺度骨干网络和差分分支对于模型整体性能都至关重要。

贡献

- 我们提出 NeuroSSM，这是一种选择性状态空间架构，可用于对原始功能性磁共振成像血氧水平依赖时间序列进行端到端分析，其计算成本随序列长度线性增长。
- 我们提出了一种多尺度状态空间骨干网络，可在单个模型中联合建模血氧水平依赖信号中的快速瞬态动力学和较慢的全局趋势。

- 我们在状态空间公式中加入了并行差分分支，以提高对大脑状态快速变化的敏感度。

- 我们在涵盖临床和非临床场景的大规模公共功能性磁共振成像数据集上评估了 NeuroSSM，结果表明其性能可与最先进的功能磁共振成像分析方法相媲美。

2. 相关工作

2.1. 早期深度学习架构

在早期将深度学习应用于功能性磁共振成像时，模型通常通过降维来应对原始血氧水平依赖信号的高维和噪声特性。一个常见策略是将 BOLD时间序列预处理为功能连接矩阵，以捕获脑区之间的成对线性关系 [17, 18]。随后，这些矩阵被用作深度学习模型的输入，例如卷积神经网络 [12] 或图神经网络 [13, 34, 35], , 以从连接结构中提取空间信息。虽然功能连接特征降低了计算负担，但其主要编码静态的一阶交互，因此忽略了血氧水平依赖信号中固有的复杂非线性动力学 [19], [20]。随着该领域转向显式的时间序列建模，循环神经网络被采用为一种自然的选择 [36, 37]。然而，在实践中，循环神经网络的性能往往受到梯度消失或爆炸问题的制约，阻碍了其在长距离范围内有效传播信息 [38]。此外，其顺序执行且无法并行化的特性增加了长时功能磁共振成像扫描的训练延迟[39]。

2.2. Transformer架构

近年来，自注意力机制被视为一种强大的机制，可用于捕捉跨时间的上下文交互 [26, 40, 41]。Transformer模型将血氧水平依赖信号视为一系列词元，其中每个词元代表单个时间点（重复时间）上所有脑区的多变量活动。然而，基础自注意力机制依据词元之间的成对相似度提取依赖关系，因此限制了其整合长序列上下文交互以及捕捉神经活动特有高阶交互的能力 [27]。此外，基础自注意力机制的复杂度随序列长度呈二次增长，这给较长的功能性磁共振成像扫描带来了显著的计算瓶颈 [26]。

随后出现了高效的Transformer变体，包括将完整扫描划分为重叠窗口以实现局部注意力 [22]；采用并行混合网络，通过交互式融合提取局部和全局表征 [42]；或利用低秩分解近似完整的注意力矩阵 [23]。虽然这些方法在运行时间和内存方面实现了近线性扩展，但也带来了上下文效率权衡：窗口化限制了信息在不同分区之间的传播，下采样牺牲了局部时间敏感性，而低秩近似则可能因舍弃细微的高秩时间模式而导致欠拟合 [25, 43]。这些近似往往会削弱全局上下文整合这一初衷，尤其容易遗漏功能性磁共振成像数据中固有的超长程交互 [9], , 并且需要针对特定功能性磁共振成像数据集，谨慎地重新调整定制的近似策略。

2.3. 状态空间模型（状态空间模型）

状态空间模型通过将序列建模为离散线性动力系统，为解决序列建模难题提供了一种很有前景的方案 [28]。它们天生适合建模极长的时间依赖关系，因此近期有多项研究将状态空间模型应用于功能磁共振成像分析。然而，现有用于功能磁共振成像的状态空间模型架构存在一些共同局限，使其无法充分利用血氧水平依赖信号的层次结构。现有方法的一个特点是依赖派生的功能连接特征，这可以避免直接对原始血氧水平依赖信号进行时间建模

信号 [29, 30]。另一方面，处理血氧水平依赖信号的方法通常只在单一时间尺度上进行处理 [31]。值得注意的是，状态空间模型的效率取决于在序列聚合过程中压缩上下文表征，但这种单尺度压缩会造成次优折衷，无法同时关注不同的时间频率。因此，单尺度状态空间模型难以同时解析快速波动和全局趋势。

NeuroSSM 与现有状态空间模型的不同之处在于，它摒弃了功能连接特征和单尺度处理。它引入了一个多尺度骨干网络，并结合时间差分路径，在多个时间分辨率上共同重塑输入序列，以捕捉快速瞬变和长程趋势。这种统一架构能够端到端地建模大脑的层级动态，同时保持严格的线性复杂度。

3. 理论

对于四维功能性磁共振成像数据的多变量分析，NeuroSSM 首先使用解剖图谱将大脑划分为 N 个感兴趣区域。在每个感兴趣区域内，对体素级BOLD信号取平均，生成单条时间序列，再进行 z 分数标准化，使其均值为零、方差为单位方差。我们采用 K 个时间尺度，步长为 $\boldsymbol{\tau} = (\tau_1、\ldots\ldots、\ \tau_K)$。对于每个尺度 k，时间重塑得到的序列长度为 $T_k = T/\tau_k$，特征维度为 $d_k = \tau_k N$。我们的模型学习将这些区域BOLD响应映射到目标标签 y（例如受试者诊断或任务条件），并将每个按时间索引的向量视为时间序列中的一个词元。

令 $\mathbf{x} = \left(x^{(0)}, \ldots, x^{(T-1)}\right) \in \mathbb{R}^{T\times N}$表示从持续时间为$T$的功能性磁共振成像扫描中提取的 N维血氧水平依赖时间序列。我们将每个向量 $x^{(t)}$ 视为长度为N 的词元，并将它们堆叠成序列

$$\mathbf{B}_0 = \left(x^{(0)}, x^{(1)}, \ldots, x^{(T-1)}\right) \in \mathbb{R}^{T\times N}.$$

该原始序列由一系列 L *NeuroSSM* 模块处理，随后接入一个任务专用头部。每个*NeuroSSM*模块首先对其输入进行归一化，然后应用一组多尺度差分状态空间块（MSD-SSB）。图1展示了该处理流程的示意图。

3.1. NeuroSSM 模块

每个*NeuroSSM* 模块接收序列$\mathbf{B}_{\ell-1} \in \mathbb{R}^{T_{\ell-1}\times N}$并输出$\mathbf{B}_\ell \in \mathbb{R}^{T_\ell \times N}$（$l = 1 \ldots L$）。为简便起见，此处省略层索引，仅分析单个块。模块首先通过输入层归一化来平衡通道统计量。对于序列张量$\mathbf{B} \in \mathbb{R}^{T\times d}$（长度为 T，特征维度为 d），可通过以下方式实现：

$$\widetilde{\mathbf{B}} = \mathrm{LN}_{\mathrm{in}}(\mathbf{B}) = \gamma_{\mathrm{in}} \odot \frac{\mathbf{B} - \mu(\mathbf{B})}{\sqrt{\sigma^2(\mathbf{B}) + \varepsilon}} + \beta_{\mathrm{in}}, \tag{1}$$

其中

$$\mu(\mathbf{B}) = \frac{1}{d}\sum_{c=1}^{d} \mathbf{B}_c, \quad \sigma^2(\mathbf{B}) = \frac{1}{d}\sum_{c=1}^{d} \left(\mathbf{B}_c - \mu(\mathbf{B})\right)^2,$$

以及 γ_{in}，$\beta_{\mathrm{in}} \in \mathbb{R}^d$ 是可学习的增益向量和偏置向量，其中 $\varepsilon > 0$ 用于保证数值稳定性。随后将得到的 $\widetilde{\mathbf{B}}$ 传递至多尺度处理阶段。

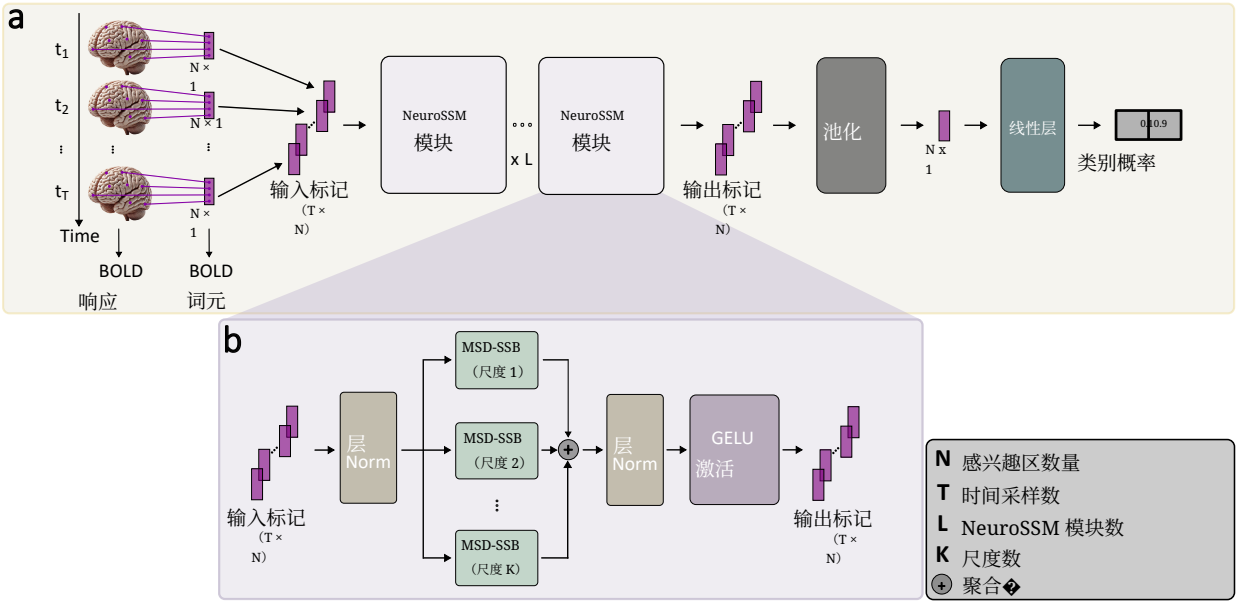


图1：**NeuroSSM 框架概览。(a)** 高层处理流程。从四维功能性磁共振成像扫描中提取感兴趣区域（ROI）的血氧水平依赖信号，构成维度为 $T \times N$的多变量时间序列，其中 T 为序列长度， N 为感兴趣区数量。该序列经过由 L 个 NeuroSSM 模块组成的堆栈处理，在保留序列长度的同时提取层次化时空特征。最终输出序列通过时间池化聚合为单个 N维表示，然后由线性分类器映射为特定任务的类别概率（例如疾病存在或不存在的概率）。**(b)** NeuroSSM 模块的内部结构。输入序列经过归一化后，分发至一组并行运行的多尺度微分状态空间块（MSD-SSB），这些模块在 K 个时间尺度上进行处理。每个尺度 k 以不同的步长 τ_k处理该序列，从而能够同时捕捉细粒度的瞬态变化和缓慢变化的全局趋势。多尺度输出经求和、归一化后，通过 GeLU 激活，再传递至后续层。

3.1.1 多尺度差分状态空间块（MSD-SSB）

MSD-SSB 旨在捕获跨多个时间层级的动态。对于每个尺度 $k = 1, \ldots, K$，我们首先以步长 τ_k对输入 $\widetilde{\mathbf{B}} \in \mathbb{R}^{T \times N}$ 进行时间重缩放，将每 τ_k 个连续词元分组为一个重缩放后的词元：

$$\widetilde{\mathbf{B}}_t^{\langle k \rangle} = \mathrm{vec}\left(\widetilde{\mathbf{B}}_{t\,\tau_k : t\,\tau_k + \tau_k - 1}\right), \quad t = 0, 1, \ldots, \frac{T}{\tau_k} - 1, \tag{2}$$

因此，重缩放后的序列张量定义为

$$\widetilde{\mathbf{B}}^{\langle k \rangle} \in \mathbb{R}^{\frac{T}{\tau_k} \times (\tau_k N)}.$$

其中， τ_k 表示尺度为 k时的时间重缩放步长；例如，选择 $\tau_k = k$ 时，当 $K = 3$，步长集合为 $\boldsymbol{\tau} = (1, 2, 3)$。

同时，在每个重缩放级别上，我们计算一阶时间差分，以提高对瞬态活动的敏感度，明确注入信号变化率：

$$\Delta\widetilde{\mathbf{B}}_0^{\langle k \rangle} = \mathbf{0}, \quad \Delta\widetilde{\mathbf{B}}_t^{\langle k \rangle} = \widetilde{\mathbf{B}}_t^{\langle k \rangle} - \widetilde{\mathbf{B}}_{t-1}^{\langle k \rangle}, \quad t = 1, \ldots, \frac{T}{\tau_k} - 1. \tag{3}$$

在 $t = 0$ 处初始化的零向量可防止虚假启动伪影，而后向差分则突出帧间变化，同时不改变序列长度或维度。随后，时间重缩放后的信号 $\widetilde{\mathbf{B}}^{\langle k \rangle}$及其时间微分 $\Delta\widetilde{\mathbf{B}}^{\langle k \rangle}$ 由我们的双流状态空间块（Dual-SSM）并行处理，生成 $\mathbf{U}^{\langle k \rangle}$ 和$\mathbf{V}^{\langle k \rangle}$，其输出在后续融合阶段进行融合，并重塑回原始时间分辨率

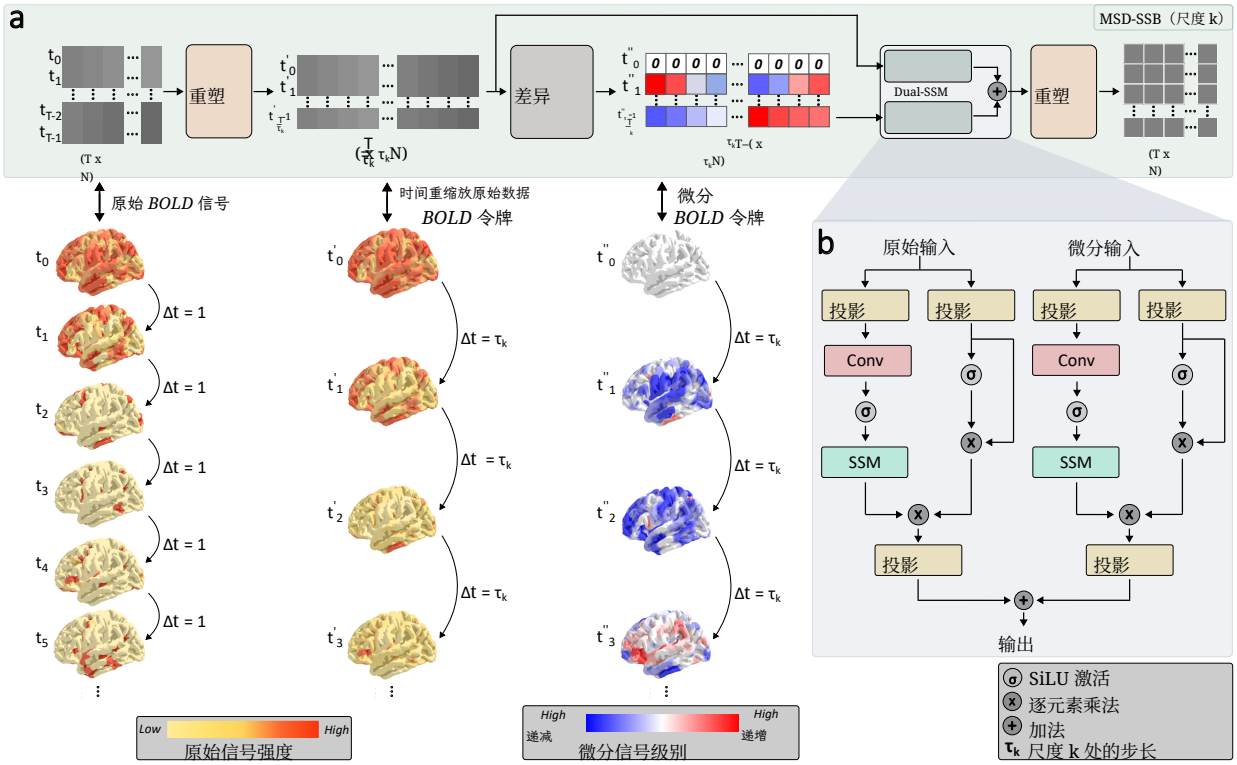


图 2：**MSD-SSB（多尺度微分状态空间块）的详细结构。(a)** 输入序列在时间维度上以尺度 k、步长 τ_k 进行重塑，得到大小为 $\lfloor T/\tau_k \rfloor \times (\tau_k \cdot N)$ 的表示。随后生成两个并行输入流：(i) 经 τ_k 重缩放的 BOLD序列；(ii) 其一阶时间差分，以突出快速状态转换。这两个序列被输入双流状态空间模块。输出通过残差连接求和，并重塑回原始的 $T \times N$ 维度，以便在层间传播。**(b)** 双流状态空间模块使用参数共享的选择性状态空间核，并行处理原始流和差分流。每个流的输入均经过主路径处理，包括逐通道卷积层、SiLU 激活和状态空间模型，以捕捉长程依赖；同时经过第二条路径，包括非线性投影和 SiLU 激活，实现乘法门控机制。最后，将原始流和差分流的输出相加融合。

sub。图 2 展示了配备 Dual-SSM 的 MSD-SSB 结构 .

3.1.2 双流状态空间模块 (Dual-SSM)

Dual-SSM 使用两个专用的状态空间核 $\mathrm{SSM}_k^{(\mathrm{rescaled})}$ 和 $\mathrm{SSM}_k^{(\mathrm{diff})}$，并行处理时间重缩放流 及其时间差分 。对于每个尺度 k，我们首先分别使用可学习映射 $s^{(k)}$ rescaled 和 $s^{(k)}$ diff 对重缩放输入和差分输入进行线性投影：

$$X^{\langle k \rangle} = s^{(k)}_{\mathrm{rescaled}}\big(\widetilde{\mathbf{B}}^{\langle k \rangle}\big), \quad D^{\langle k \rangle} = s^{(k)}_{\mathrm{diff}}\big(\Delta\widetilde{\mathbf{B}}^{\langle k \rangle}\big), \tag{4}$$

其中，$s^{(k)}$ rescaled 和 $s^{(k)}$ diff 将 $d_k = \tau_k N$ 通道映射至内部宽度 $d_{\mathrm{inner},\ k} = \mathrm{expand} \cdot d_k$（通过门控线性投影实现）。随后，我们使用卷积核宽度为 d_{conv} 且具有每尺度卷积核 $C_{(k)}$rescaled 和 $C^{(k)}$diff 的深度一维卷积（记为 $*$），

$$\overline{X}^{\langle k \rangle} = X^{\langle k \rangle} * C^{(k)}_{\mathrm{rescaled}}, \quad \overline{D}^{\langle k \rangle} = D^{\langle k \rangle} * C^{(k)}_{\mathrm{diff}}. \tag{5}$$

这些卷积后的输出随后被输入两个专用的选择性状态空间核：$\mathrm{SSM}_k^{(\mathrm{rescaled})}$ 用于重缩放流，$\mathrm{SSM}_k^{(\mathrm{diff})}$ 用于差分流：

$$\mathbf{U}^{\langle k \rangle} = \mathrm{SSM}_k^{(\mathrm{rescaled})}\big(\overline{X}^{\langle k \rangle}\big), \quad \mathbf{V}^{\langle k \rangle} = \mathrm{SSM}_k^{(\mathrm{diff})}\big(\overline{D}^{\langle k \rangle}\big). \tag{6}$$

每个核 SSM_k 都会维护学习得到的对角状态系数 $\Lambda \in \mathbb{R}^{d_{\text{inner},\,k} \times d_{\text{state}}}$，其中 d_{state} 表示每个通道的状态维度；此外还有三个紧凑的多层感知机 s_δ，s_β，s_γ，用于将时间 t 的卷积后词元 $r_t^{\langle k \rangle}$（即，对于 $\mathrm{SSM}_k^{(\text{rescaled})}$ 为 $X_{\langle k \rangle t}$，对于 $\mathrm{SSM}_k^{(\text{diff})}$ 则为 $D_t^{\langle k \rangle}$）映射为内容感知选择参数：

$$\delta^{(t)} = s_\delta(\bar{r}_t^{\langle k \rangle}), \quad \beta^{(t)} = s_\beta(\bar{r}_t^{\langle k \rangle}), \quad \gamma^{(t)} = s_\gamma(\bar{r}_t^{\langle k \rangle}), \tag{7}$$

并结合 $\delta^{(t)} > 0$。它们确定线性递归的离散更新系数：

$$A^{(t)} = \exp\big(\delta^{(t)}\Lambda\big), \quad P^{(t)} = \big(\delta^{(t)}\Lambda\big)^{-1}\big(e^{\delta^{(t)}\Lambda} - I\big)\,\delta^{(t)}\beta^{(t)}, \tag{8}$$

由此驱动线性递归关系：

$$h^{(t)} = A^{(t)}\,h^{(t-1)} + P^{(t)}\,\bar{r}_t^{\langle k \rangle}, \quad u^{(t)} = \gamma^{(t)} \odot h^{(t)}, \tag{9}$$

其中，$\odot$表示 逐元素乘法。同时，另一个轻量级 MLP s_z 生成一个门控向量：

$$z^{(t)} = s_z(\bar{r}_t^{\langle k \rangle}). \tag{10}$$

并通过最终的线性投影 s_{out}应用三门非线性函数：

$$y^{(t)} = u^{(t)} \odot \big(z^{(t)} \odot \sigma(z^{(t)})\big), \quad o^{(t)} = s_{\text{out}}\big(y^{(t)}\big). \tag{11}$$

由于 Λ 是对角矩阵（以逐通道状态系数的形式存储），式（8）和（9）中的循环机制在 $d_{\text{inner},k}$ 个通道和 d_{state} 个状态维度上均为逐元素运算，因此可以实现融合的 *selective-scan*，每个时间步的计算成本为 $O\big(d_{\text{inner},k}\,d_{\text{state}}\big)$（即当 d_{state}固定时，与 $d_{\text{inner},k}$ 呈线性关系）。扫描完成后，再应用式（11）中的最终投影。最后，将 Dual-SSM 的两个流相加，并重塑回原始序列长度：

$$\mathbf{Z}^{\langle k \rangle} = \big(\mathbf{U}^{\langle k \rangle} + \mathbf{V}^{\langle k \rangle}\big) \xrightarrow{\text{reshape}} \mathbb{R}^{T \times N}. \tag{12}$$

3.1.3 多尺度融合与预测

经过 Dual-SSM 处理后，将各尺度的输出对齐至公共 $T \times N$ 形状，并通过无参数残差求和在不同尺度间进行融合：

$$\mathbf{Z} = \sum_{k=1}^{K} \mathbf{Z}^{\langle k \rangle}. \tag{13}$$

这种简单的聚合保留了每个尺度的信息，同时避免了学习式融合方案（例如基于注意力机制的方案）带来的额外参数和计算开销。

融合后的序列仍包含特定于尺度的统计信息；因此，我们应用第二次层归一化 $\mathrm{LN}_{\text{out}}(\cdot)$，其函数形式与式 (1) 相同，但使用独立的可学习参数。随后逐元素应用高斯误差线性单元（GELU），

$$\mathbf{B}_{\text{out}} = \phi\big(\mathrm{LN}_{\text{out}}(\mathbf{Z})\big), \quad \phi(z) = \tfrac{1}{2}z\big[1 + \mathrm{erf}(z/\sqrt{2})\big], \tag{14}$$

提供平滑的激活并改善梯度流。堆叠 L 个此类*NeuroSSM* 模块可得到深度表征$\mathbf{B}_L$，其同时保留局部动态与长程依赖。为获得固定大小的表征，我们沿时间进行全局均值池化，

$$\mathbf{g} = \frac{1}{T}\sum_{t=0}^{T-1} \mathbf{B}_{L,t}, \tag{15}$$

从而在不增加额外参数的情况下捕获整体时间上下文。池化表征$\mathbf{g}$随后通过线性分类器进行最终预测。

3.2. 模型复杂度

序列模型的计算成本，根本上取决于一个 N 维词元与序列中其余 $T-1$ 个词元交互的频率。在经典深度学习架构中，循环神经网络会在每个时间步更新一个 $N \times N$ 转移矩阵，因此时间复杂度为 $O(N^2T)$，内存复杂度为 $O(N^2)$ [44, 45, 46]。核宽度为 k_{cnn} 的一维卷积神经网络会执行 k_{cnn} 次此类矩阵-向量乘法，时间复杂度为 $O(k_{\text{cnn}}N^2T)$，内存复杂度为 $O(k_{\text{cnn}}N^2)$。原始 Transformer 会计算 T 查询与键之间每一对的点积；因此，多头自注意力的时间和内存复杂度均为 $O(NT^2)$ [26, 40, 42]。与此同时，高效 Transformer 将注意力交互限制在局部窗口或稀疏模式内。例如，采用时间窗口化或局部注意力的方法会将序列划分为长度为 W 的重叠窗口，并将自注意力交互限制在该窗口内的词元之间 [25]。由此，整个序列的总计算成本降至 $O(NTW)$，相对于序列长度 T 实现线性扩展，同时引入了W 这一新的、对性能至关重要的超参数，用于决定最大上下文长度 [25]。因此，随着W 增大以更有效地捕捉长程上下文，其效率优势会随之减弱，复杂度也会回到二次级。

相比之下，NeuroSSM 通过使用 K 并行选择性状态空间模型流，完全省去了注意力机制。在尺度 k下，时间重塑会生成长度为 $T_k = T/\tau_k$ 、词元宽度为 $d_k = \tau_k N$的序列。Dual-SSM 将该宽度扩展至 $d_{\text{inner},k} = \text{expand} \cdot d_k$ ，并维持大小为 d_{state}的逐通道潜在状态。由于状态转移是对角的，选择性扫描更新会在各通道上逐元素计算，每个时间步的计算量为 $O(d_{\text{inner},k} d_{\text{state}})$，每个尺度的计算量为 $O(T_k d_{\text{inner},k} d_{\text{state}})$。代入 $T_k = T/\tau_k$和 $d_{\text{inner},k} = \text{expand} \cdot \tau_k N$后，每个尺度的时间复杂度为 $O(\text{expand} \cdot N \cdot d_{\text{state}} \cdot T)$，因此总体时间复杂度为 $O(K \cdot \text{expand} \cdot N \cdot d_{\text{state}} \cdot T)$。这对应于每个尺度 $O(d_{\text{inner},k} d_{\text{state}})$ 的内存开销（当 d_{state} 固定时，随 N 线性增长）。门控和最终投影在扫描之后执行，不会改变随 T线性增长的复杂度。在我们的实现中， $K = 3$、$\text{expand} = 3$和 $d_{\text{state}} = 2$ 均为较小的常数；因此，主导时间开销按 $O(NT)$ 增长，内存开销按 $O(N)$ 增长，另有固定因子 $K \cdot \text{expand} \cdot d_{\text{state}}$。

4. 方法

4.1. 数据集

演示基于 HCP S1200 [1] [32] 和 PPMI 发布版本 [33]进行。在 HCP S1200 中，使用静息态功能磁共振成像数据（HCP-Rest）进行性别检测，使用任务态功能磁共振成像数据（HCP-Task）进行认知任务检测。在 PPMI 中，使用静息态功能磁共振成像数据检测帕金森病（PD）。建模前，所有数据均经过预处理，包括与结构扫描进行刚性配准、归一化至 MNI空间、感兴趣区域分区，以及对每个感兴趣区域内的时间序列进行 Z分数标准化。

HCP-Rest： 使用了预处理后的静息态功能磁共振成像数据的首个扫描时段 [47]。剔除了时间点少于 1200 个的扫描。分析的数据集共包含 871 次扫描，其中女性受试者 484 名，男性受试者 387 名。使用 Schaefer图谱进行分区，得到 400 个感兴趣区 [48]。

HCP-Task: 使用了预处理后的任务态功能磁共振成像数据的第一场次 [47]。剔除了长度短于七种认知任务（情绪、关系推理、赌博、语言、社交、运动和工作记忆）各自标准时长的扫描数据。分析的数据集包含来自 61 名受试者的 427 次扫描。通过 Schaefer图谱进行脑区分区，得到 400 个感兴趣区 [48]。

PPMI： 使用了预处理后的静息态功能磁共振成像数据 [33]。采用二分类疾病检测任务，阳性类为帕金森病受试者，阴性类为神经系统健康的

[1] https://db.humanconnectome.org

对照组。排除了前驱期、SWEDD 和所有遗传队列组。对于拥有多次静息态扫描的受试者，仅保留首次可用扫描。分析数据集共包含 114 名受试者，其中 99 例为帕金森病，15 例为对照病例。使用 AAL3v2图谱进行脑区分区，共得到 166 个感兴趣区 [49]。

4.2. 实验设置

所有模型均使用 PyTorch 框架在单张 Nvidia RTX 5090 GPU 上实现。除非另有说明，模型采用 Adam优化器，通过最小化分类交叉熵损失进行训练，共训练 20 个周期，迷你批次大小为 32，权重衰减为 4×10^{-5}。

对于所有数据集，我们采用分层交叉验证程序评估模型性能。首先将数据划分为开发集（80%）和留出测试集（20%），以保持不同类别之间的平衡。在开发集中，使用 5 折分组交叉验证调整超参数。在每个内层折中，模型使用约 64% 的样本进行训练，并在剩余的 16% 样本上进行验证。根据平均验证性能选定超参数后，使用完整开发集训练最终模型，并在留出测试集上评估一次。在所有实验中，训练集、验证集和测试集之间的受试者身份严格保持互斥。

为分析学习效率，我们在每个数据集上构建了学习曲线。对于给定数据集，我们从开发集中对受试者进行子采样，仅使用可用训练受试者的 $S \in \{5、10、20、50、100\}\%$，同时固定留出测试集。对于每个 S，我们在开发集的相应子集上从头重新训练模型，并在测试集上进行评估。

为提高训练效率并引入随机性，随机时间裁剪仅应用于训练阶段：HCP-Rest裁剪为600个时间点，HCP-Task为150个，PPMI为100个。测试序列保留其完整的原始长度。为确保公平比较，所有模型均采用相同的按受试者划分方式和训练样本。实验使用五个随机种子进行，并对结果取平均以提高可靠性，同时在不同模型之间保持种子一致。模型性能通过准确率、F1分数和曲线下面积进行量化。NeuroSSM与基线方法之间性能差异的统计显著性通过非参数Wilcoxon符号秩检验进行评估。

4.3. 实现细节

核心 NeuroSSM 模型在所有实验中均采用单个块层（$L = 1$）。每个时间步的输入都是维度为 N 的分区化感兴趣区域（ROI）向量（例如，HCP 为 $N = 400$，PPMI 为 $N = 166$）。该输入向量经过 NeuroSSM 模块处理，该模块由以下部分组成：初始层归一化、一个在 $K = 3$ 个时间尺度上运行且步长为 $\boldsymbol{\tau} =$（1，2，3）的多尺度差分状态空间块（MSD-SSB）组、第二个层归一化，以及 GeLU 激活函数。

对于每个尺度 τ_k，MSD-SSB 首先进行 τ_k 倍时间重缩放，并计算一阶时间差分流。随后，两个流由双流状态空间模块（Dual-SSM）并行处理。Dual-SSM 包含两个相同的单输入单输出状态空间核（一个用于重缩放后的输入，另一个用于差分输入），二者共享参数化以提高效率。这些核的潜在状态大小为 $d_{\text{state}} = 2$，深度卷积窗口大小为 $d_{\text{conv}} = 1$，MLP 扩展因子为 $\text{expand} = 3$。

4.4. 竞争方法

NeuroSSM 与八种基线方法进行了基准比较，涵盖传统机器学习、经典深度学习，以及近期的 Transformer 和状态空间模型方法。

支持向量机： 一种基于静态皮尔逊相关功能连接特征 [50] 运行的线性l_2–正则化支持向量机。交叉验证得到的正则化权重为 $C = 1$。

表1：竞争方法在 HCP-Rest、HCP-Task 和 PPMI 的功能性磁共振成像扫描上的表现。准确率（准确率）、F1 分数和 AUC 指标报告为所有评估训练集比例 $S \in \{5, 10, 20, 50, 100\}\%$ 下的均值±标准差。粗体表示每个数据集各项指标的最佳方法，下划线表示次佳方法。

	HCP-Rest			HCP-Task			PPMI		
方法	准确率（%）	F1（%）	曲线下面积 (%)	准确率(%)	F1分数 (%)	曲线下面积 (%)	准确率(%)	F1分数 (%)	曲线下面积 (%)
SVM	75.20±4.78	69.20±7.88	83.23±5.97	53.85±6.11	52.50±5.57	69.68±6.36	83.22±4.68	89.96±4.03	60.28±18.36
BrainNetCNN	72.05±3.20	64.93±8.71	77.13±3.49	29.23±6.12	24.45±6.62	65.88±4.26	82.26±5.59	89.63±4.01	52.67±18.23
LSTM	73.35±2.97	62.85±8.62	83.03±2.86	73.67±4.49	72.43±5.15	95.39±1.05	82.78±6.66	89.86±5.02	61.70±17.59
STAGIN	71.75±2.88	65.10±5.40	74.52±3.55	40.53±2.93	36.98±3.50	74.12±2.06	79.37±14.17	85.41±14.28	62.26±14.70
BolT	78.33±3.69	70.18±10.58	87.51±2.31	83.47±3.06	82.32±3.30	98.95±0.44	84.00±5.49	90.43±4.94	49.33±16.43
FST-Mamba	67.89±3.09	58.37±13.75	76.03±3.41	58.20±7.55	55.87±8.63	88.06±2.93	77.22±11.88	84.75±10.46	58.87±13.12
Mamba	68.87±5.41	62.92±8.54	75.34±7.53	73.63±5.17	72.52±5.46	94.13±2.08	84.87±4.47	91.50±2.95	59.87±9.03
NeuroSSM	**81.76±4.02**	**74.68±8.63**	**89.81±2.02**	**87.82±2.69**	**87.24±2.48**	**99.10±0.27**	**86.43±3.56**	**92.55±2.24**	**67.47±17.75**

BrainNetCNN：一种基于功能连接特征运行的卷积模型 [12]。交叉验证得到的超参数为1×10^{-4} 学习率。

LSTM： 一种处理原始血氧水平依赖信号的循环模型，通过对时间维度上的隐藏状态取平均进行分类 [44]。交叉验证得到的超参数包括：1 层、32 个隐藏单元、0.5 的 dropout，以及 0.05 的学习率。

STAGIN： 一种处理滑动窗口血氧水平依赖信号的图变换器模型 [24]。交叉验证得到的超参 数为：4 个模型层、128 个隐藏单元、窗口大小为 50、步长为 3，以及 1×10^{-4} 学习率。

BolT： 一种采用融合局部窗口注意力机制处理原始血氧水平依赖信号的 Transformer 模型 [25]。其配置包括 4 个 Transformer 层、400 个隐藏单元、窗口大小为 20，以及学习率 2×10^{-4}。

FST-Mamba： 一种处理动态功能连接特征的分层时空状态空间模型 [30]。该模型采用 $d_{\text{state}} = 16$、$d_{\text{conv}} = 4$，扩展系数为 = 1，学习率为 1×10^{-3}。

Mamba： 一种专为处理原始血氧水平依赖信号而构建的状态空间模型 [51]。交叉验证得到的超参数为$d_{\text{state}} = 2, d_{\text{conv}} = 2$，扩展因子为 = 3，学习率为 5×10^{-4}。

5. 结果

5.1. 比较实验

我们在三个功能性磁共振成像分析任务上评估了 NeuroSSM：在 *HCP-Rest* 上进行性别检测，在 *HCP-Task* 上进行认知任务检测，以及在 *PPMI* 上进行疾病检测。我们将其与已有的基线方法进行比较，包括传统方法（支持向量机）、经典深度学习方法（BrainNetCNN、长短期记忆网络）、Transformer（STAGIN、BolT）和状态空间模型（FST-Mamba、Mamba）。为评估竞争方法的学习效果，我们系统地调整了各数据集中可用训练受试者的比例，并考察模型性能随训练集规模变化的情况 $S \in \{5, 10, 20, 50, 100\}\%$。

准确率、F1 分数和曲线下面积随训练集大小变化的学习曲线如图3所示。在 HCP-Rest 和 HCP-Task 上，NeuroSSM 在 S 上始终取得最优表现，尤其在低数据量情形下（即 $S \leq 20\%$）优势显著，表明其相较于 Transformer 和状态空间模型基线方法具有更高的样本效率。在 PPMI 上，大多数方法的准确率和 F1 分数已接近饱和，而曲线下面积在 S 间的波动则显著更大。总体而言，NeuroSSM 的曲线下面积表现具有竞争力，并且在训练集比例较小至中等时排序表现更佳，体现了其在数据受限情形下的稳健性。

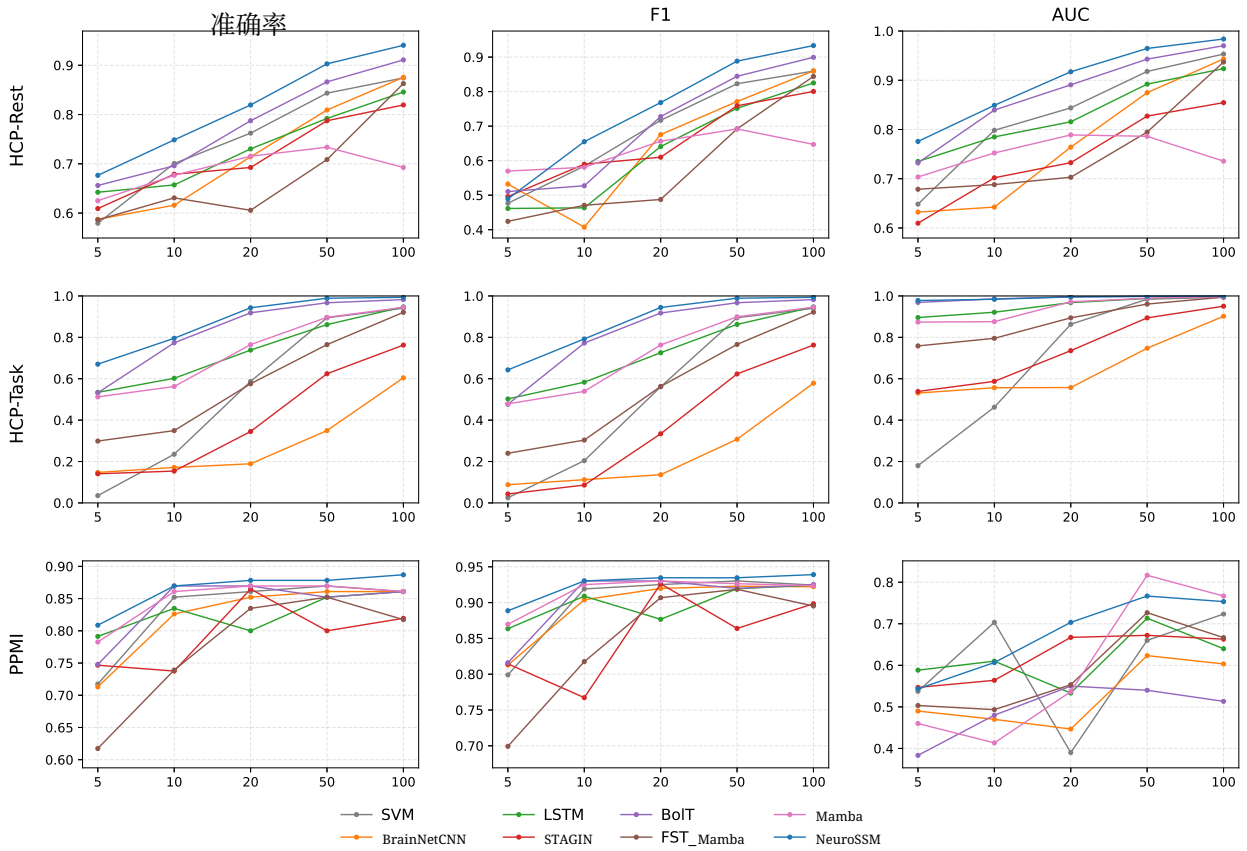


图 3：HCP-Rest、HCP-Task 和 PPMI 上的学习效率曲线。每个面板展示了在留出测试集上测得的准确率、F1分数或曲线下面积；模型使用开发集中不同占比 $S \in \{5,10,20,50,100\}\%$ 的受试者进行训练。x 轴表示 S（训练集占比），y 轴显示相应的指标值。

不同训练集比例下的综合表现汇总于表1。在 HCP-Rest 上，NeuroSSM 的静息态性别检测指标优于所有基线方法（$p < 0.05$，符号秩检验）。与传统基线方法相比，其表现分别提升（+6.6%、 +5.5%、 +6.6%）；与经典深度学习相比，分别提升（+9.1%、 +10.8%、 +9.7%）；与 Transformer 基线方法相比，分别提升（+6.7%、 +7.0%、 +8.8%）；与状态空间基线方法相比，分别提升（+13.4%、 +14.0%、 +14.1%）（依次对应准确率、F1分数和曲线下面积）。在 HCP-Task 上，NeuroSSM 也取得了最高的各项指标，且在训练集较小至中等时优势尤为明显（$p < 0.05$）。与传统基线方法相比，NeuroSSM 的表现分别提升（+34.0%、 +34.7%、 +29.4%）；与经典深度学习相比，分别提升（+36.4%、 +38.8%、 +18.5%）；与 Transformer 相比，分别提升（+25.8%、 +27.6%、 +12.6%）；与状态空间基线方法相比，分别提升（+21.9%、 +23.0%、 +8.0%）。尽管所有方法在训练集规模较大时都接近性能上限S，NeuroSSM 在 S 上始终保持最高的曲线下面积。最后，在 PPMI 上，NeuroSSM 的疾病检测指标最高（$p < 0.05$）。与传统基线方法相比，其表现分别提升（+3.2%、 +2.6%、 +7.2%）；与经典深度学习相比，分别提升（+3.9%、 +2.8%、 +10.3%）；与 Transformer 相比，分别提升（+4.8%、 +4.6%、 +11.7%）；与状态空间基线方法相比，分别提升（+5.4%、 +4.4%、 +8.1%）。尽管支持向量机等较简单的线性基线方法取得了有竞争力的 F1 分数，但其曲线下面积较低，反映出可靠性较差，这可能是由于 PPMI 具有较高的数据异质性。综合来看，这些结果表明，NeuroSSM 能更有效地从功能性磁共振成像时间序列中提取受试者层面的潜在结构，既能捕捉内在认知变量（静息态性别和任务诱发认知），也能捕捉疾病相关特征，尤其是在监督信号有限时。

表 2：NeuroSSM 变体在来自 HCP-Rest、PPMI 及其平均值的静息态功能磁共振成像扫描上的性能。指标列说明该变体是否包含多尺度流和差分流。粗体表示每个数据集上各项指标的最佳方法，下划线表示次佳方法。

多尺度	差分	HCP-Rest			PPMI			平均值		
		准确率（%）	F1分数 (%)	曲线下面积 (%)	准确率(%)	F1分数 (%)	曲线下面积 (%)	准确率(%)	F1分数 (%)	曲线下面积 (%)
✗	✗	79.43±10.36	71.36±16.93	87.68±9.18	86.09±1.74	92.46±1.13	49.87±7.87	82.76±6.05	81.91±9.03	68.77±8.53
✓	✗	80.59±9.59	73.07±16.12	88.96±7.87	84.52±5.32	91.14±3.99	64.93±11.50	82.56±7.45	82.11±10.06	76.95±9.69
✗	✓	80.05±10.61	72.78±16.69	88.11±9.32	86.26±1.39	**92.58±0.88**	52.53±7.67	83.15±6.00	82.68±8.78	70.32±8.49
✓	✓	**81.76±9.70**	**74.68±16.13**	**89.81±7.68**	**86.43±2.84**	92.55±1.86	**67.47±8.64**	**84.10±6.27**	**83.61±9.00**	**78.64±8.16**

表3：NeuroSSM 变体在来自 HCP-Rest、PPMI 及其平均数据集的静息态功能磁共振成像扫描上的表现。变体模型针对不同的分辨率配置 τ 构建。粗体表示每个数据集各项指标的最佳方法，下划线表示次佳方法。

τ 集	HCP-Rest			PPMI			平均值		
	准确率(%)	F1分数 (%)	曲线下面积 (%)	准确率(%)	F1分数 (%)	曲线下面积 (%)	准确率(%)	F1分数 (%)	曲线下面积 (%)
{1}	80.05±10.61	72.78±16.69	88.11±9.32	86.26±1.39	**92.58±0.88**	52.53±7.67	83.15±6.00	82.68±8.78	70.32±8.49
{1, 2}	81.14±10.17	74.48±15.77	89.31±8.37	86.26±2.36	92.51±1.52	64.73±10.37	83.70±6.27	83.49±8.64	77.02±9.37
{1, 2,3}	81.76±9.70	74.68±16.13	89.81±7.68	**86.43±2.84**	92.55±1.86	**67.47±8.64**	**84.10±6.27**	83.61±9.00	**78.64±8.16**
{1, 2,3, 4}	**81.90±9.11**	**75.65±15.12**	**89.98±7.36**	85.91±3.45	92.14±2.42	66.33±10.28	83.91±6.28	**83.90±8.77**	78.15±8.82
{1, 2,3,4, 5}	81.30±9.96	74.86±16.40	89.81±7.42	86.26±2.29	92.48±1.48	67.47±8.92	83.78±6.13	83.67±8.94	78.64±8.17

5.2. 消融研究

我们开展了消融研究，以分别评估多尺度处理和 NeuroSSM 中差分输入流的贡献。我们分析了四种变体：(i) 不进行多尺度处理或差分的单尺度基线，(ii) 仅多尺度变体，(iii) 仅差分变体，以及 (iv) 结合这两个组件的完整模型。表 2 列出了静息态 HCP-Rest 和 PPMI 数据集上的结果，结果为不同训练集比例下的平均值。与单尺度基线相比，多尺度处理和差分流均能提升性能，而完整模型的总体表现最佳，表明这两个组件具有互补优势。

接下来，我们评估了多尺度骨干网络中时间分辨率配置 τ 的影响（表3）。总体而言，$\tau = \{1,2,3\}$ 和 $\tau = \{1,2,3,4\}$ 在各数据集上的表现十分接近。这些发现进一步证实，在选定的 $\tau = \{1,2,3\}$ 配置中增加更多尺度，会带来额外的模型复杂度，却没有明显的性能收益。

6. 讨论

我们提出了 NeuroSSM，这是一种新颖的选择性状态空间架构，旨在捕捉功能性磁共振成像时间序列中的层次化、上下文感知表征，以用于下游分析。NeuroSSM 在多个时间尺度上处理血氧水平依赖信号，同时注入一阶导数信息，从而整合长程上下文，且计算成本与序列长度（即扫描时长）呈线性关系，避免了 Transformer 中注意力机制的二次复杂度。随后，基于这些丰富的时空特征的池化表征进行分类。在静息态和任务态功能性磁共振成像数据集上的实验表明，与最先进的基线方法相比，NeuroSSM 性能更优；这些基线方法包括卷积、循环、基于图、Transformer 以及近期的状态空间模型。

本研究主要关注针对性别、认知任务和疾病检测的类别输出分类模型。对最终隐藏特征进行全局池化，可得到整个功能性磁共振成像扫描的精简高层表征。然而，脑活动还编码了有关连续变量的信息 [52]。NeuroSSM 非常适合分析这些特征的皮层表征。最终的池化表征向量可连接线性输出层或采用线性激活函数的多层感知机，使 NeuroSSM 能够从功能磁共振成像扫描中解码连续变量。

计算神经影像学中一种常见做法是构建解码模型，根据测得的脑活动预测外部刺激或任务条件 [53, 54], , 我们在此采用的正是这一框架。另一种方法是构建编码模型，根据刺激或任务特征预测脑激活 [55, 56, 57, 58, 59]。对于认知神经影像学研究，可以将任务的实验时间进程输入基于 NeuroSSM 的架构，再按脑区训练该架构以预测测得的血氧水平依赖响应。在这一神经编码框架下评估 NeuroSSM 的有效性，尤其是通过多尺度连接，使更细尺度特征以更粗尺度表征为条件，从而确保上下文传播一致 [60], , 仍是未来研究的一个重要方向。

我们对 PPMI 数据集中的帕金森病进行的分析仅依赖静息态功能磁共振成像扫描的信息。尽管我们的结果令人鼓舞，但研究表明，纳入患者人口统计信息或临床评估等辅助信息，有助于提升疾病检测效果 [61]。此外，许多神经系统疾病（包括帕金森病）在其他成像模态（如结构磁共振成像或弥散加权磁共振成像）中也呈现出独特的特征 [62]。整合此类多模态数据有望进一步提升 NeuroSSM 的性能。可在最终分类层之前，将辅助信息与潜在表征融合；对于其他成像模态，则可通过并行编码器流进行处理，并将其输出与功能性磁共振成像特征拼接。

按照神经影像学中的常见做法，我们使用标准图谱在模板空间中对每位受试者的大脑进行分区，并将这些感兴趣区域的平均血氧水平依赖响应输入模型。虽然这确保了不同受试者之间感兴趣区域定义的一致性，但向共同模板进行空间配准可能会造成信息损失。通过在每位受试者的原生脑空间中定义感兴趣区域，或可减轻此类损失。具体而言，可使用配准变换的逆变换对模板感兴趣区进行反投影。另一种方法是将三维卷积神经网络作为空间前端集成到 NeuroSSM 中，在状态空间主干网络对时间演化进行建模之前，先从体积功能性磁共振成像数据中学习提取特征 [22, 40, 63]。在此背景下，采用专门的状态空间模型模块，分别捕捉空间依赖和通道级上下文依赖，可能会进一步提升处理高维体积数据时的表征能力 [64]。

我们使用包含数百名受试者的数据集，从头开始训练了所有参比模型。尽管所有参比模型均从头开始训练，但 HCP-Task 上的学习曲线结果表明，即使训练受试者数量有限，NeuroSSM 仍能稳健运行。对于仅有小规模数据集可用的应用场景，可通过迁移学习进一步提升其性能 [65, 66]。例如，可以先在 HCP 这类大规模公共数据集上对 NeuroSSM 进行预训练，再针对规模较小的特定队列进行微调。基于先进流程进行图像合成，以可靠的方式扩充数据，也有助于缓解数据稀缺问题 [67, 68, 69, 70]。联邦学习等其他策略则可在不损害患者隐私的前提下，促进模型在多样化的多机构数据集上进行训练，从而进一步增强模型的泛化能力 [71, 72]。系统探索 NeuroSSM 的数据效率仍是未来研究的重要课题。

7. 结论

在本研究中，我们提出了 NeuroSSM，这是一种用于功能性磁共振成像分类的新型选择性状态空间架构，能够高效处理原始血氧水平依赖序列。该架构采用多尺度骨干网络，以捕捉多样的时间动态，并行

差分流以增强对瞬态活动的敏感性，同时保持线性计算成本。在大规模 HCP 和 PPMI 数据集上进行验证后，NeuroSSM 在性别、疾病（帕金森病）和认知任务检测方面取得了最先进的准确率，超越了现有基线方法。NeuroSSM 为功能磁共振成像分析提供了高效且敏感的框架，非常适合未来在临床和认知领域的扩展，包括连续变量的解码以及多模态数据的整合。

致谢

本研究部分经费来自授予 F. Genc¸ 的 TUBITAK BIDEB 奖学金，以及授予 T. C¸ukur 的 TUBA GEBIP 2015奖学金、BAGEP 2017奖学金和 TUBITAK 121N029资助。Neuripides项目获得欧洲神经退行性疾病研究联合计划（JPND）2020年征集项目“与神经退行性疾病相关的新型成像和脑刺激方法与技术”的资助，项目名称为“神经反馈：通过自我刺激大脑治疗帕金森病”。Neuripides项目还获得了以下机构在 JPND 框架下提供的资助：荷兰——荷兰健康研究与发展组织（ZonMw）；德国——联邦教育与研究部（BMBF）；捷克共和国——教育、青年与体育部（MEYS）；法国——法国国家研究署（ANR）；加拿大——加拿大卫生研究院（CIHR）；土耳其——土耳其科学技术研究委员会（TUBITAK）。

参考文献

[1] E. M. Hillman，“血氧水平依赖信号的耦合机制及其意义：现状报告”，*Annu. Rev. Neurosci.*，第37卷，第161–181页，2014年。[2] K. Li、L. Guo、J. Nie、G. Li和T. Liu，“使用功能磁共振成像检测功能性脑连接的方法综述”，*Comput. Med. Imaging Graph.*，第33卷，第2期，第131–139页，2009年。[3] A. Venkataraman、K. R. Van Dijk、R. L. Buckner和P. Golland，“通过聚类探索功能磁共振成像中的功能连接”，载于*ICASSP*，IEEE，2009年，第441–444页。[4] S. Nishimoto、A. T. Vu、T. Naselaris、Y. Benjamini、B. Yu和J. L. Gallant，“从自然影片诱发的脑活动中重建视觉体验”，*Curr.Biol.*，第21卷，第19期，第1641–1646页，2011年。[5] C. Niu等，“使用机器学习从个体受试者的静息态功能磁共振成像中建模运动任务激活”，*Brain Imaging Behav.*，第15卷，第1期，第122–132页，2021年。[6] B. T. Thomas Yeo等，“通过内在功能连接估计人类大脑皮层的组织结构”，*J. Neurophysiol.*，第106卷，第3期，第1125–1165页，2011年。[7] R. Kong等，“个体特异性皮层网络的空间拓扑结构可预测人类认知、人格与情绪”，*Cereb. Cortex*，第29卷，第6期，第2533–2551页，2019年。[8] K. Rajpoot、A. Riaz、W. Majeed和N. Rajpoot，“静息态功能磁共振成像中癫痫的功能连接改变”，*PloS One*，第10卷，第8期，e0134944，2015年。[9] O. L. Campbell和A. M. Weber，“功能磁共振成像的单分形分析：入门综述”，*Human Brain Mapping*，第43卷，第8期，第2693–2706页，2022年。[10] N. Logothetis、J. Pauls、M. Augath、T. Trinath和A. Oeltermann，“功能磁共振成像信号基础的神经生理学研究”，*Nature*，第412卷，第150–157页，2001年。[11] D. A. Handwerker、J. M. Ollinger和M. D’Esposito，“不同受试者和脑区之间血氧水平依赖血流动力学反应的变异及其对统计分析的影响”，*NeuroImage*，第21卷，第4期，第1639–1651页，2004年。

[12] J. Kawahara等，“BrainNetCNN：用于脑网络的卷积神经网络；迈向神经发育预测”，NeuroImage，第146卷，第1038–1049页，2017年。[13] S. Parisot等，“使用图卷积网络进行疾病预测：孤独症谱系障碍和阿尔茨海默病的应用”，Med. Image Anal.，第48卷，第117–130页，2018年。[14] L. Fan、J. Su、J. Qin、D. Hu和H. Shen，“一种用于动态功能连接的深度网络模型及其在性别分类和智力预测中的应用”，Front. Neurosci.，第14卷，第881页，2020年。[15] M. J. McKeown和T. J. Sejnowski，“功能性磁共振成像数据的独立成分分析：考察相关假设”，Hum. Brain Mapp.，第6卷，第5–6期，第368–372页，1998年。[16] M. Svens´en、F. Kruggel和H. Benali，“功能性磁共振成像组研究数据的独立成分分析”，NeuroImage，第16卷，第3期，第551–563页，2002年。[17] X. Zhang、B. Hu、X. Ma和L. Xu，“使用L2正则化逻辑回归进行轻度认知障碍分类的静息态全脑功能连接网络”，IEEE Tran. NanoBio.，第14卷，第2期，第237–247页，2015年。[18] C. Wang、Z. Xiao和J. Wu，“使用SVM-RFECV基于静息态功能磁共振成像数据对孤独症和对照组进行基于功能连接的分类”，Phys. Med.，第65卷，第99–105页，2019年。[19] P.-J. Lahaye、J.-B. Poline、G. Flandin、S. Dodel和L. Garnero，“功能连接：研究血氧水平依赖信号之间的非线性延迟相互作用”，NeuroImage，第20卷，第2期，第962–974页，2003年。[20] Z. Hu和P. Shi，“血氧水平依赖信号的非线性分析：生物物理建模、生理状态与功能激活”，载于MICCAI，Springer，2007年，第734–741页。[21] P. Ciuciu、G. Varoquaux、P. Abry、S. Sadaghiani和A. Kleinschmidt，“静息和任务状态下功能性磁共振成像信号的无标度多重分形时间动力学”，Frontiers in Physiology，第3卷，2012年。[22] I. Malkiel、G. Rosenman、L. Wolf和T. Hendler，“用于功能性磁共振成像预测任务的Transformer预训练与微调”，arXiv:2112.05761，2021年。[23] Z. Peng、Z. He、Y. Jiang、P. Wang和Y. Yuan，“GBT：用于孤独症诊断的几何导向脑Transformer”，载于MICCAI，Springer，2024年，第142–152页。[24] B.-H. Kim、J. C. Ye和J.-J. Kim，“利用时空注意力学习脑连接组的动态图表示”，载于NIPS，Curran Associates Inc.，2021年，第4314–4317页。[25] H. A. Bedel、I. Sivgin、O. Dalmaz、S. U. Dar和T. C¸ukur，“BolT：用于功能性磁共振成像时间序列分析的融合窗口Transformer”，Med. Image Anal.，第88卷，第102841页，2023年。[26] A. Vaswani等，“Attention Is All You Need”，载于NeurIPS，第30卷，Curran Associates, Inc.，2017年。[27] J. Zhang、L. Zhou、L. Wang、M. Liu和D. Shen，“用于脑部疾病分类的扩散核注意力网络”，IEEE Tran. Med. Imaging，第44卷，第10期，第2814–2827页，2022年。[28] A. Gu、K. Goel和C. Re，“利用结构化状态空间高效建模长序列”，载于ICLR，2022年。[29] A. Behrouz和F. Hashemi，“Brain-Mamba：通过选择性状态空间模型编码脑活动”，载于PCHIL，丛书PMLR，第248卷，PMLR，2024年，第233–250页。[30] Y. Wei、A. Abrol和V. D. Calhoun，“用于功能磁共振成像分析的分层时空状态空间建模”，载于RECOMB，Springer，2025年，第86–98页。[31] W. Deng、F. Han、Q. Ling、Q. Liu和H. Han，“Causal fMRI-Mamba：用于神经解码和脑任务状态识别的因果状态空间模型”，载于ICASSP，2025年，第1–5页。[32] D. C. Van Essen等，“WU-Minn人类连接组计划：概述”，NeuroImage，第80卷，第62–79页，2013年。

[33] K. Marek 等，“帕金森病进展标志物计划（PPMI）——建立帕金森病生物标志物队列”，《临床与转化神经病学年鉴》，第 5 卷，第 12 期，第 1460–1477 页，2018年。[34] B.-H. Kim 和 J. C. Ye，“理解用于静息态功能磁共振成像功能连接分析的图同构网络”，*Front. Neurosci.*，第 14 卷，第 630 页，2020年。[35] X. Kan、H. Cui、J. Lukemire、Y. Guo 和 C. Yang，“FBNetGen：通过功能性脑网络生成进行任务感知的基于图神经网络的功能磁共振成像分析”，载于 *PMLR*，第 172 卷，2022年，第 618–637 页。[36] Z. Yang、X. Zhuang、K. Sreenivasan、V. Mishra、T. Curran 和 D. Cordes，“一种用于任务态功能性磁共振成像数据去噪的稳健深度神经网络：在工作记忆和情景记忆中的应用”，医学图像分析，第 60 卷，第 101 622 页，2019年。[37] H. Li 和 Y. Fan，“使用长短期记忆循环神经网络从功能磁共振成像中进行脑解码”，载于*MICCAI*，Springer，2018年，第 320–328 页。[38] G. Kerg、B. Kanuparthi、A. G. ALIAS PARTH GOYAL、K. Goyette、Y. Bengio 和 G. Lajoie，“梳理人工神经网络中循环机制与自注意力之间的权衡”，*NeurIPS*，第 33 卷，第 19443–19 454 页，2020年。[39] R. Pascanu、T. Mikolov 和 Y. Bengio，“论循环神经网络训练的困难”，载于 *ICML*，PMLR 系列，第 28 卷，2013年，第 1310–1318 页。[40] S. Nguyen、B. Ng、A. D. Kaplan 和 P. Ray，“注意并解码：使用注意力模型进行 4D 功能磁共振成像任务状态解码”，载于*ML4H*，PMLR，第 136 卷，2020年，第 267–279 页。[41] P. He、Z. Shi、Y. Cui、R. Wang 和 D. Wu，“用于基于静息态功能磁共振成像的阿尔茨海默病诊断的时空图 Transformer 方法”，《生物学与医学中的计算机》，第 178 卷，第 108 762 页，2024。[42] C. Zhao 等，“IFT-Net：用于儿科超声心动图定量分析的交互式融合 Transformer 网络”，医学图像分析，第 82 卷，第 102 648 页，2022年。[43] X. Tian、H. Ma、Y. Guan、L. Xu、J. Liu 和 L. Tian，“Transformer3：用于基于功能磁共振成像的人脑功能表征的纯 Transformer 框架”，《*IEEE* 生物医学与健康杂志》，第 29 卷，第 1 期，第 468–481 页，2025年。[44] N. Dvornek、P. Ventola、K. Pelphrey 和 J. Duncan，“使用长短期记忆网络从静息态功能磁共振成像中识别孤独症”，载于 *MLMI*，第 10541 卷，2017年，第 362–370 页。[45] C. Zhao、H. Li、Z. Jiao、T. Du 和 Y. Fan，“用于功能磁共振成像数据去噪的三维卷积封装长短期记忆网络（3DConv-LSTM）模型”，载于*MICCAI*，Springer，第 12267 卷，2020年，第 479–488 页。[46] X. Xing 等，“采用辅助任务训练的动态谱图卷积网络用于早期轻度认知障碍诊断”，载于*MICCAI*，Springer，2019年，第 639–646 页。[47] M. F. Glasser 等，“人类连接组计划的最简预处理流程”，*NeuroImage*，第 80 卷，第 105–124 页，2013年。[48] A. Schaefer 等，“基于内在功能连接磁共振成像的人类大脑皮层局部-整体分区”，《大脑皮层》，第 28 卷，第 9 期，第 3095–3114 页，2017年。[49] E. T. Rolls、C.-C. Huang、C.-P. Lin、J. Feng 和 M. Joliot，“自动解剖标记图谱 3”，*NeuroImage*，第 206 卷，第 116 189 页，2020年。[50] A. Abraham 等，“从多中心静息态数据中提取可重复的生物标志物：一个基于自闭症的示例”，*NeuroImage*，第 147 卷，第 736–745 页，2017年。[51] A. Gu 和 T. Dao，“Mamba：采用选择性状态空间的线性时间序列建模”，载于*COLM*，2024。

[52] T. C¸ukur、S. Nishimoto、A. G. Huth 和 J. L. Gallant，“自然视觉过程中的注意力会扭曲人脑中的语义表征”，Nat. Neurosci.，第 16 卷，第 6 期，第 763–770 页，2013年。[53] S. M. LaConte，“实时解码功能磁共振成像脑状态”，NeuroImage，第 56 卷，第 2 期，第 440–454 页，2011年。[54] P. Andersson、J. P. Pluim、J. C. Siero、S. Klein、M. A. Viergever 和 N. F. Ramsey，“使用 7T功能磁共振成像实时解码对视空间注意的脑反应”，PloS One，第 6 卷，第 11 期，e27638，2011年。[55] S. Nishimoto，“使用运动能量及时空视觉变换器特征对电影诱发的人脑活动进行建模”，bioRxiv，2021年。. [56] E. Celik、U. Keles、I. Kiremitc¸i、J. L. Gallant 和 T. C¸ukur，“人脑中动态场景类别表征的皮层网络”，Cortex，第 143 卷，第 127–147 页，2021年。[57] M. Shahdloo、E. C¸ elik、B. A. Urgen、J. L. Gallant 和 T. C¸ukur，“视觉动作类别搜索过程中语义表征的任务依赖性扭曲”，J. Neurosci.，第 42 卷，第 35 期，第 6782–6799 页，2022年。[58] A. J. Anderson、B. D. Zinszer 和 R. D. Raizada，“功能磁共振成像的表征相似性编码：基于模式的合成，使用刺激-模型相似性预测脑活动”，NeuroImage，第 128 卷，第 44–53 页，2016年。[59] G. H. Ngo、M. Nguyen、N. F. Chen 和 M. R. Sabuncu，“一种基于变换器的神经语言模型，可根据自由形式文本查询合成脑激活图”，医学图像分析，第 81 卷，第 102 540 页，2022年。[60] B. Kabas、F. Arslan、V. A. Nezhad、S. Ozturk、E. U. Saritas 和 T. C¸ukur，“用于医学图像重建的物理驱动自回归状态空间模型”，arXiv:2412.09331v3，2025年。[61] N. C. Dvornek、P. Ventola 和 J. S. Duncan，“使用循环神经网络结合表型数据与静息态功能磁共振成像数据进行孤独症分类”，载于 ISBI，IEEE，2018年，第 725–728 页。[62] N. Pyatigorskaya、C. Gallea、D. Garcia-Lorenzo、M. Vidailhet 和 S. Lehericy，“磁共振成像在帕金森病中的应用综述”，Thera. Adv. Neurol. Disord.，第 7 卷，第 4 期，第 206–220 页，2014年。[63] J. Choi、D. K. Park、J. Kwon、S. Yoo 和 J. Cha，“NeuroMamba：用于功能磁共振成像的状态空间基础模型”，载于 NeurIPS，2025年。.[64] S.. Ozt¨urk、O. Duran 和 T. Cukur，“Denomamba：用于低剂量CT去噪的融合状态空间模型”，IEEE Journal of Biomedical and Health Informatics，第 PP 卷，第 1–14 页，2025年11月。[65] J. Devlin、M.-W. Chang、K. Lee 和 K. Toutanova，“BERT：用于语言理解的深度双向变换器的预训练”，载于 NAACL，2019年。[66] O. Dalmaz、M. Yurt 和 T. C¸ukur，“ResViT：用于多模态医学图像合成的残差视觉变换器”，IEEE Tran. Med. Imaging，第 41 卷，第 10 期，第 2598–2614 页，2022年。[67] A. G¨ung¨or 等，“用于加速磁共振成像重建的自适应扩散先验”，医学图像分析，第 88 卷，第 102872 页，2023年。[68] M. Ozbey 等，“使用对抗扩散模型进行无监督医学图像转换”，IEEE Tran. Med. Imaging，第 42 卷，第 10 期，第 3524–3539 页，2022年。[69] S. Chen、R. Zhang、H. Liang、Y. Qian 和 X. Zhou，“状态空间模块与注意力机制的耦合：一种输入感知的多对比度磁共振成像合成方法”，Medical Physics，第 52 卷，第 4 期，第 2269–2278 页，2025年。[70] O. F. Atli 等，“I2i-mamba：通过选择性状态空间建模实现多模态医学图像合成”，arXiv:2405.14022v6，2025年。[71] G. Elmas 等，“用于磁共振成像重建的生成式图像先验联邦学习”，IEEE Tran. Med. Imaging，第 42 卷，第 7 期，第 1996–2009 页，2022年11月。

[72] O. Dalmaz 等人，“一个模型统一所有模型：多对比度磁共振成像合成的个性化联邦学习”，医学图像分析，第 94 卷，页码 103121，2024 年。