受物理启发的神经声学计算：基于可调非线性多重散射

Ali Momeni[1]、**Xinxin Guo**[2]、**Herv´e Lissek**[2]和 **Romain Fleury**[1,*]

[1]波工程实验室，洛桑联邦理工学院（EPFL），瑞士洛桑 CH-1015。[2]LTS2信号处理实验室，洛桑联邦理工学院（EPFL），瑞士洛桑 CH-1015*电子邮件：romain.fleury@epfl.ch

arXiv:2304.08380v1 [cs.SD] 17 Apr 2023

摘要

光和声音等波会因周围复杂材料物体引发的多重散射而自然反弹并相互混合。这种散射过程会严重扰乱波所携带的信息，对传统通信系统、传感范式和基于波的计算方案构成挑战。在这里，我们展示了多重散射非但不是阻碍，反而有助于实现并增强模拟非线性信息映射，从而能够直接以物理方式实现储备池计算和极限学习机等计算范式。我们提出了此类计算架构的一种受物理启发的实现方案，用于语音和元音识别；该方案直接在输入信号的原生域中运行，即处理真实声音，无需任何数字预处理、编码转换或反向传播训练计算。我们首先在一个概念验证原型中实现了该方案：一个非线性混沌声学腔，其中包含多个可调谐且高能效的非线性超散射体。我们证明，这种基于声学的计算系统在元音识别任务中效率很高，测试分类准确率达到 91.4%。最后，我们展示了该系统在混响室这一自然环境中进行元音识别的高性能。我们的研究成果为高效声学学习机铺平了道路，使其能够直接处理输入声音，并利用物理原理实现自然语言处理（NLP）。

1 引言

波在复杂的非均匀线性介质中的传播，可视为一组输入模式与输出模式之间的线性混合过程。能够承载信息的空间自由度数量与系统尺寸和波长之比有关；在光学中，这一数量可能极其庞大[1–4]。利用波进行模拟计算和信号处理任务，例如矩阵乘法、卷积、图像微分和随机投影，因其能够以极低功耗在大规模并行架构中运行而备受关注；与传统数字计算机不同，这种方式无需使用记忆[5–18]。事实上，从传统计算机到用于机器学习的专用架构（如图形处理器（GPU）或谷歌的张量处理单元（TPU）），当前所有计算机架构都面临所谓的冯·诺依曼瓶颈：记忆与处理单元分离会带来显著的速度和能耗代价[19]。另一方面，利用波进行模拟计算属于一种独特的计算范式，即神经形态计算，不受这些缺陷的影响[20,21]。

非均匀和/或非线性介质中的波传播与单层神经网络之间的类比，可用于直接处理波所携带的信息。波物理学与计算机科学之间的此类类比，催生了两个领域中的众多概念，例如非确定性多项式时间、NP难问题与光子伊辛机之间的类比[22,23]。复杂和/或非线性介质中的波传播，使人们能够以不受神经元数量影响的速度，为不同的学习任务执行随机核或非线性核；借此，简单的线性支持向量机（SVM）[24] 或线性回归[25] 即可对经复杂介质变换后获得的数据进行分类（例如基于波的储备池计算（RC）和极限学习机（ELM））[26–34]。然而，实现此类基于波的RC和ELM系统所需的非线性，通常是在基于波的系统的数字前端实现，或通过材料固有非线性引入（例如克尔介质），这会损害能效、灵活性和速度[28,31,35,36]。

另一方面，非线性强散射介质在基于波的

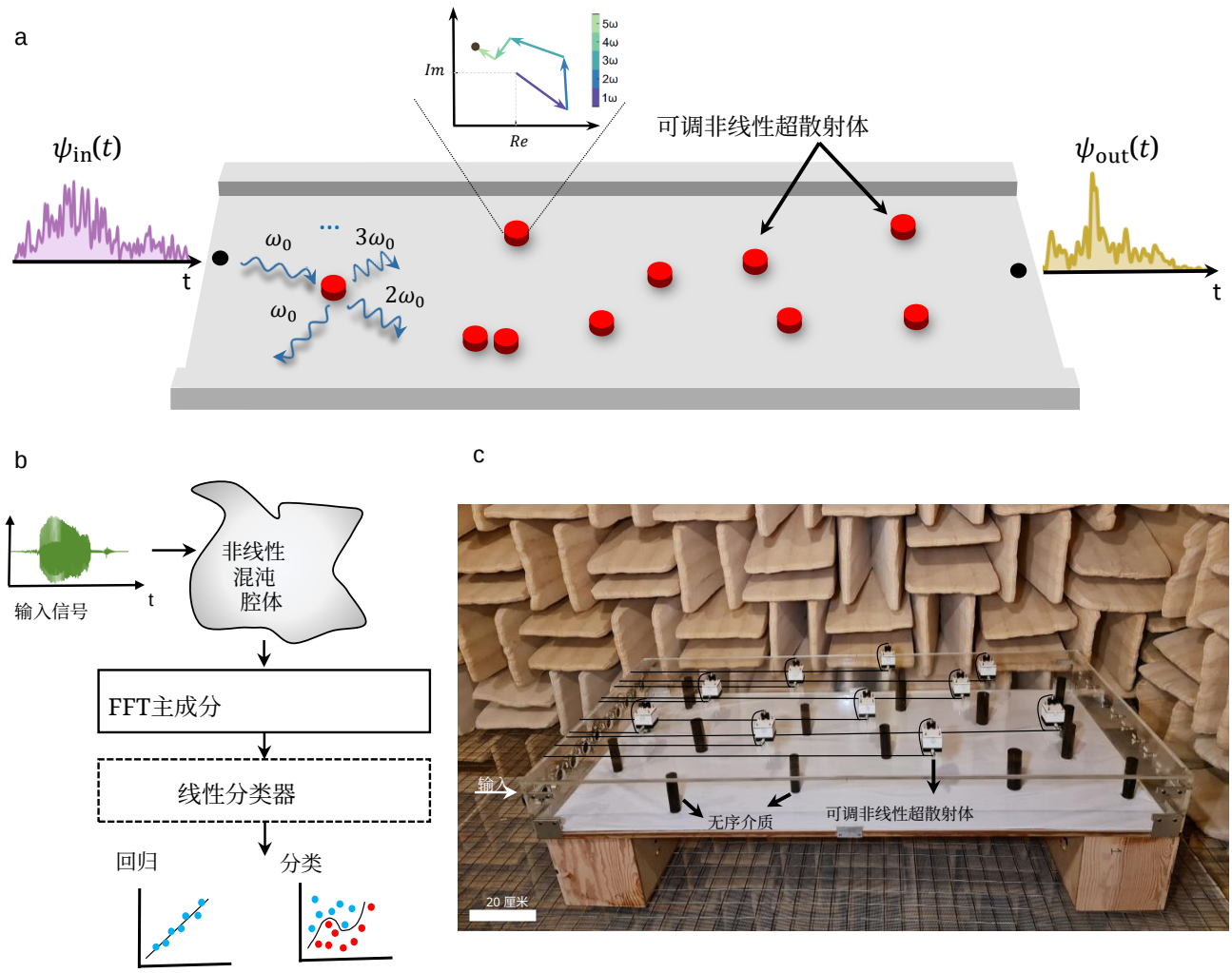


图 1. 受物理启发的储备池计算。a，物理非线性混沌计算系统的示意图，以及**b**，对应的框图。原始波形 Ψin(*t*) 从左侧输入端注入，并在非线性混沌腔体中传播。波在腔体内传播时，会遇到可调非线性超散射体，将波散射为高次谐波。输出ψout(*t*)被馈入一个具有重新缩放系数的可调密集层（线性分类器），用于回归和分类。**c**，所提出计算系统的实验实现：系统使用声波和一个声学腔，其中包含十五根亚波长杆和十个非线性超散射体；这些超散射体呈非线性活性膜的形式，随机安装在腔壁上。

神经形态计算中可能具备的优势，至今仍鲜少有人探索。多重散射通常与强空间色散相关，即介质中相距较远的点之间的响应相互交织。这可能增强信号流的互连性、加剧非线性，并使系统超越单层模型。这为在真实环境中处理信息提供了可能：只需在已有的无序介质中加入一些非线性超散射体即可。这一点在射频或声波领域尤为重要，因为在这些领域中，腔体

并且非线性超散射体很容易获得。

在此，我们展示了一种基于波的神经形态储层计算机，它能直接对待感知的波信号进行快速、低功耗的非线性数据变换，无需进行转换或预处理，也无需执行繁重的反向传播训练计算[4,37]。我们展示了该方案对于一项颇具挑战性的元音识别问题的适用性：让声音在掺入非线性超散射体的多重散射腔体中传播，并用简单的线性分类器处理少数几个麦克风的输出，即可解决这一问题。我们通过实验研究了随机分布于均匀介质中的非线性超散射体如何提升首个原型——非线性声学腔体——的计算性能。我们证明了基于声学的计算系统在具有挑战性的元音识别任务中的效率，对测试声音取得了较高的分类准确率（91.4%）。随后，我们通过展示该方案在真实且强散射的混响室这一现实场景中的高性能元音识别能力，证明了其灵活性。这类基于波的计算系统不仅速度快、易于训练、能效高且用途广泛，还具有独特的准确率表现，可与最精密的数字方案相媲美。

2 结果

2.1 受物理启发的神经形态储备池计算

我们研究一种特殊的循环神经网络方案，称为储备池计算。在储备池计算中，固定随机权重的循环连接神经元构成储备池，并通过一个可训练层与少数输出神经元相连[38]。固定随机权重的连接神经元可大幅简化训练过程，这是储备池计算的一项关键优势。此外，储备池计算的循环特性使其适用于时间序列预测和分类[39–44]。

现在简要介绍传统 RC 的概念。输入向量 $\psi\mathrm{in}(t)$ 被注入一个称为“储备池”的高维动态系统。向量 $\mathbf{h}_t$ 表示储备池状态，其初始状态是随机定义的。储备池神经元的内部连接，以及输入与储备池神经元之间的连接，分别由 **Win** 和 **Wr** 定义。这两个矩阵均以随机方式初始化，并在整个 RC 训练过程中保持固定。神经元的演化

储层 储备池计算机可以表示为如下形式的递归非线性映射[45]:

$$\mathbf{h}_{t+1} = \mathrm{F_{NL}}\left(\mathbf{W}_{\mathrm{in}}\psi_t^{\mathrm{in}} + \mathbf{W}_{\mathrm{r}}\mathbf{h}_t\right) \tag{1}$$

其中 FNL 是一个非线性函数。在训练流程中，输入 ψin 被馈入储层，并递归计算相应的储层状态。训练过程的最后一步是执行简单的线性支持向量机或线性回归，以最小化用于调整**W**out权重的损失函数。输出可通过$\mathbf{O}_t$ =**W**out$\mathbf{h}_t$计算。根据式（1），储层被定义为一种具有独特记忆和非线性特性的动力系统。在这类系统中，储层状态会在每个时间步根据其先前的值进行更新，使系统能够保留过去信息的记忆，并学习数据中的长程依赖。

该方案可采用图 1a所示的非线性多重散射神经声学系统实现，其框图见图 1b。该系统由多个随机分布在均匀介质（例如腔体）中的非线性超散射体组成。我们将输入声音 ψin(t)从左侧注入，使其在腔体中传播。线性和非线性超散射体产生大量多重散射事件。在我们的实验中，我们设计了适用于声波的高能效有源非线性超散射体，其特性将在后文详细介绍。

声压$p(x, y, z)$的动力学由声波方程 $\partial^2 p \partial t^2 - c^2 \cdot \nabla^2 p = \psi \mathrm{in}_t$控制，其中 $\nabla^2 = \frac{\partial^2}{\partial} x^2 + \frac{\partial^2}{\partial} y^2 + \frac{\partial^2}{\partial} z^2$ 是拉普拉斯算子，$c = c(x, y, z)$是波速的空间分布，而 $\psi \mathrm{in}_t = \psi \mathrm{in}_t(x, y, z, t)$是源项。对该方程进行有限差分时间离散化后，可得到如下形式的递推方程

$$\mathbf{h}_t = \mathbf{Q}^{(\mathrm{c})}(\mathbf{h}_{t-1}) \cdot \mathbf{h}_{t-1} + \mathbf{Q}^{(\mathrm{i})} \cdot \mathbf{x}_t \tag{2}$$

其中$\mathbf{h}_t = \begin{pmatrix} p_{t+1} \\ p_t \end{pmatrix}$ 以及稀疏矩阵$\mathbf{Q}^{(\mathrm{c})}$描述了波场 p_t 和p_{t-1}在无声源时的更新。矩阵

$\mathbf{Q}^{(\mathrm{c})}$ 和 $\mathbf{Q}^{(\mathrm{i})}$的构造，以及式（2）的推导，见补充材料第S1节。 $\mathbf{Q}^{(\mathrm{c})}$ 对p_{t-1} 的依赖通常可能是非线性的，从而使波动系统产生非线性响应。比较式（1）和式（2）可以发现二者有明显的相似之处

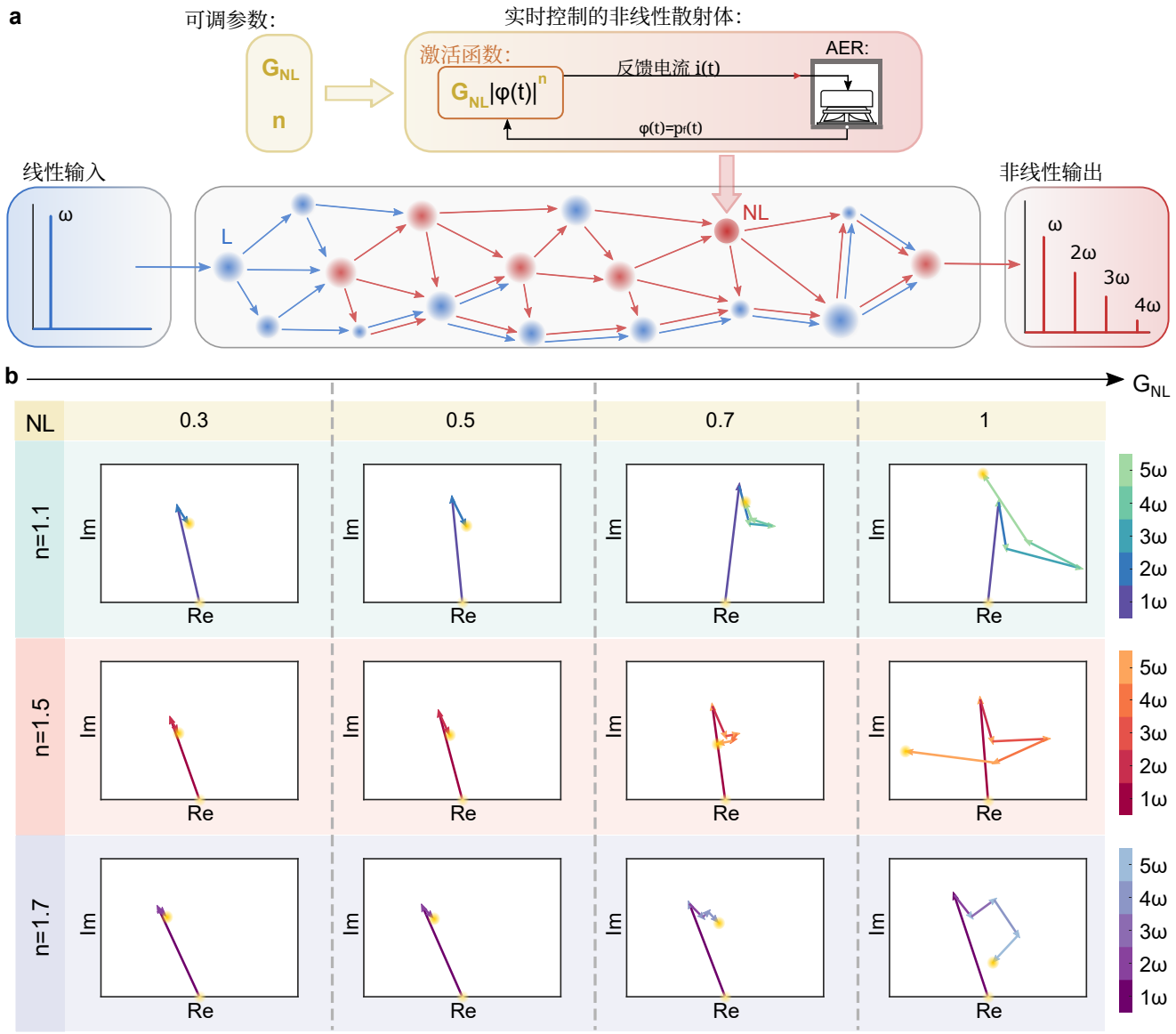


图 2. 完全可调的非线性多重散射。a非线性散射体（红色）由有源电动力学谐振器（AER）构成，其形式为扬声器振膜，通过反馈电流 $i(t)$ 实时控制，非线性根据前方声压 pf 指定。**b** 非线性散射过程可调性的演示：改变非线性的类型和强度（由增益 GNL 定义）时，在非线性散射体前方测得的前五个声压谐波的频闪图（复相量的演变）。非线性分别由幂n 和增益 GNL 定义。整个系统由 500 Hz 的正弦波激励。在每种非线性情形下，所有谐波的实部（横坐标）和虚部（纵坐标）均按对应基波分量的绝对值缩放 ω。浅黄色和亮黄色点分别表示初始位置 (0,0) 和最终位置。

在动力学方面，传统 RC 与波动系统有所不同。该递推关系表达了时间步 t 的波场对时间步 $t-1$ 状态的依赖性，揭示了基于波的系统固有的短期记忆。除了波在介质中传播所产生的内部记忆，多重散射还将长期记忆显式融入复杂结构中，并进一步增强它。为此，我们利用并设计共振散射元件，通过固有的波反馈实现增强的长期记忆，从而构建出复杂的可能信息传播路径“迷宫”。当基于波的系统受到狄拉克δ函数 $\psi\mathrm{in}(t) = \delta(t - t0)$激励时，可通过脉冲响应在实验中直观呈现记忆效应。时间步 $t0$ 的输入影响会延续至特定时间范围，这一事实体现了记忆效应。我们在补充材料第S3节中，分别针对非线性混沌腔和混响室展示了这一结果。

2.2 完全可调的非线性多重散射过程

我们构建了一种用于神经形态计算的基于物理波的系统原型，如图1c所示。所采用的声学腔由多个随机放置在腔体顶壁上的非线性超散射体组成。时间输入信号由腔体右侧的十个扬声器注入，发出的波形穿过包含十个非线性超散射体的无序腔体传播。波形由位于非线性超散射体处的十个麦克风接收。

为实现所需的非线性超散射体，我们对亚波长电动扬声器采用了一种主动控制策略，该策略受一种已知方法启发，这种方法最初用于吸声应用[46,47]。通过实时反馈电流$i(t)$，该策略能够调整扬声器的声学特性。

这种方法能够在任意低的输入功率下产生显著的非线性效应[48]。这意味着，中心频率为 ω_0 的波在传播过程中会产生频率为 ω_0整数倍的频谱分量。为实现向高次谐波的最佳相干能量传递，本文设计的非线性控制律形式为$G\mathrm{NL}p\mathrm{nf}$，如图2a所示，其中$p\mathrm{f}(t)$是受控扬声器前方测得的声压。$G\mathrm{NL}$和n 是两个控制参数，分别用于调节非线性的程度（由$G\mathrm{NL}$控制）和形式（由指数n控制）。

通过主动方式在超散射体中引入非线性后，所诱导的非线性

多重散射过程变得完全可调。通过测量给定位置处声压的时间信号，对其进行单边傅里叶变换，并将各次谐波的傅里叶分量绘制为复平面上的箭头，即可展示这一点。将这些分量相加，便能直观呈现离散时刻的声压值，从而得到每个基频谐波周期内多频声音的频闪图，如图2所示。本文考虑了图2b中的三种非线性类型，其特征是幂指数n = 1分别为 1、1.5 和 1.7。对于每种非线性类型，均通过非线性增益G_{NL} 来调节非线性程度。当其增大到稳定性所允许的最大值时，高次谐波的振幅可达到与某个基波分量相当的水平，在G_{NL} = 1的情况下尤为明显。这凸显了混沌腔中多重散射的影响：它显著增强了非线性，即提高了高次谐波的强度，而这在自由空间情形下无法实现。此外，幂指数n从n = 1.1 变为n = 1.7，即使只是微小变化，也会导致所有谐波的复域轮廓截然不同（见图2b）。这种由非线性多重散射带来的、对非线性参数的增强敏感性，使我们能够针对特定计算任务优化合适的非线性类型和程度。

通过一个简单的回归问题，可以展示所提出的物理启发计算系统的性能。该问题基于一个由非线性关系生成的数据集。对非线性函数进行线性回归，是评估神经网络模型学习能力的一种标准且广为人知的分析方法；若没有非线性变换[30]，则无法进行这种分析。输入信息（ζ_{in}）是一组在 $-\pi$ 到 π 之间随机生成的数值，相应的输出标签（y_{i}）根据非线性函数 $y = \mathrm{sinc}(\zeta_{\mathrm{in}})$生成。每个输入值（$\zeta_{\mathrm{in}}$）通过与固定的随机二维掩模相乘，映射为一个矩阵（$\mathbf{S}_{\mathrm{in}}$）。此处掩模的维度为 M×N（N=M=10），其中 M 和 N 分别对应扬声器数量和输入数量。我们假设$\mathbf{S}_{\mathrm{in}}$的每一行都以彼此相近但各不相同的频率ω_{n}进行调制，因此：

$$\psi_m^{\mathrm{in}}(t) = \sum_{n=1}^{\mathrm{N}} \mathbf{S}_{n,m}^{\mathrm{in}} \sin(\omega_n t) \tag{3}$$

归一化后的 $\Psi_{\mathrm{in}m}(t)$ 由第 m 个扬声器注入腔体（见 图 1c）。非线性

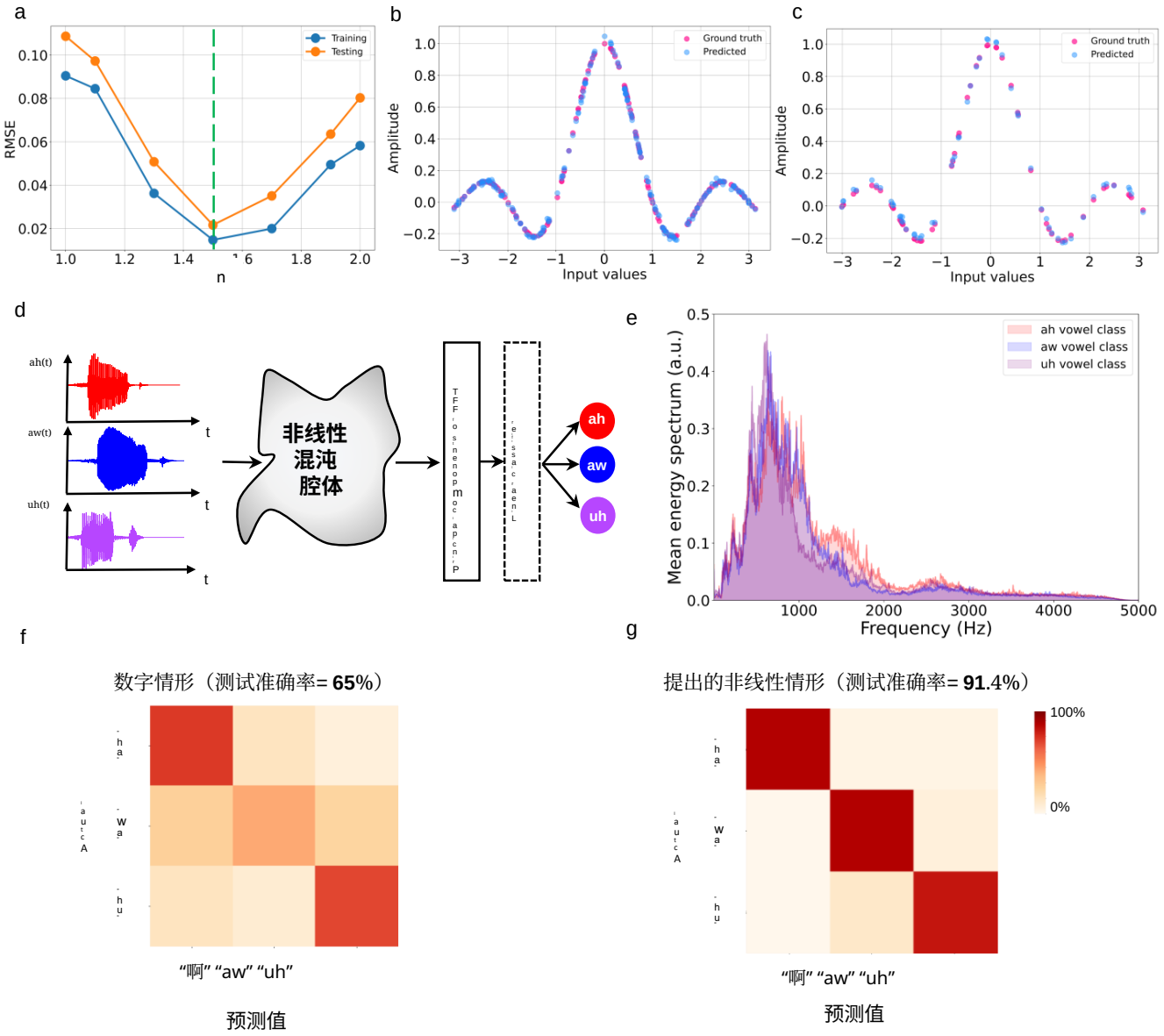


图3. 非线性回归任务和元音识别的学习后性能。a调节非线性超散射体的非线性指数n，可以使标准回归学习任务（即插值 sinc函数）的均方根误差达到最小值。**b** 和**c** 分别表示n = 1.5 时的回归训练和测试性能。**d** 元音识别装置示意图。来自三个类别的口语元音样本的原始音频波形分别通过声源扬声器注入。信号经过非线性混沌腔后，将输出信号快速傅里叶变换的主成分输入线性分类器。**e** “ah”、“aw”和“uh”元音类别的频率成分。**f** 和**g** 分别为数字情形和非线性情形的混淆矩阵及对应的测试准确率。

变换后的波形由十个麦克风接收。对输出进行线性回归，输出定义为针对每个训练输入 ζ_{in} 测得的二次谐波强度。学习性能通过均方根误差（RMSE）衡量。为了使某个

对于回归学习任务，我们调节非线性超散射体的非线性指数 n。回归任务的均方根误差随超参数 n 变化的情况（训练和测试数据集）如图 3a所示。当 n = 1.5 时，学习性能显著，均方根误差极低。将n 设为这一最优值时，训练和测试数据集的回归结果分别如图 3b 和c所示。这些结果表明，所提出的计算系统中实现的非线性数据映射足以执行标准的非线性回归任务。值得注意的是，为了计算决策层的二次谐波强度，我们只需重新缩放线性回归权重，无需付出任何额外的大量计算成本（更多详情请参阅方法部分）。

2.3 元音识别

所提出的基于波的系统可直接用真实声音进行训练，无需对输入进行任何预处理或编码。因此，我们转而研究一项更具挑战性的任务：对口语元音进行分类。为此，我们使用标准元音识别数据集进行模型训练[49]、[50]，并在训练时未使用的新未见数据上测量预测准确率。我们使用的数据集包含来自45名不同男性说话者和48名不同女性说话者的10类元音共930段原始音频录音[49]。在本学习任务中，我们选取了其中279段录音，涉及线性分类器无法分类的三类元音，对应的元音为“ah”“aw”和“uh”（见图 3d）。为作说明，我们绘制了所有样本（男性和女性说话者）中这三类元音的平均能量谱，见图 3e。在整个频率范围内，这三类元音的谱能量密度有很大重叠。因此，所选的元音识别任务极具挑战性。

每个元音的音频波形以 10 毫秒的时间间隔依次输入腔体左侧的十个扬声器，发出的波形穿过无序腔体传播（更多详情请参见方法）。位于非线性超散射体处的十个麦克风接收这些波形。通过麦克风记录非线性传递波形后，对变换后的信号进行快速傅里叶变换（FFT），并提取其主成分。最后，我们使用简单的线性支持向量机对元音类别进行分类（更多详情请参见方法）。

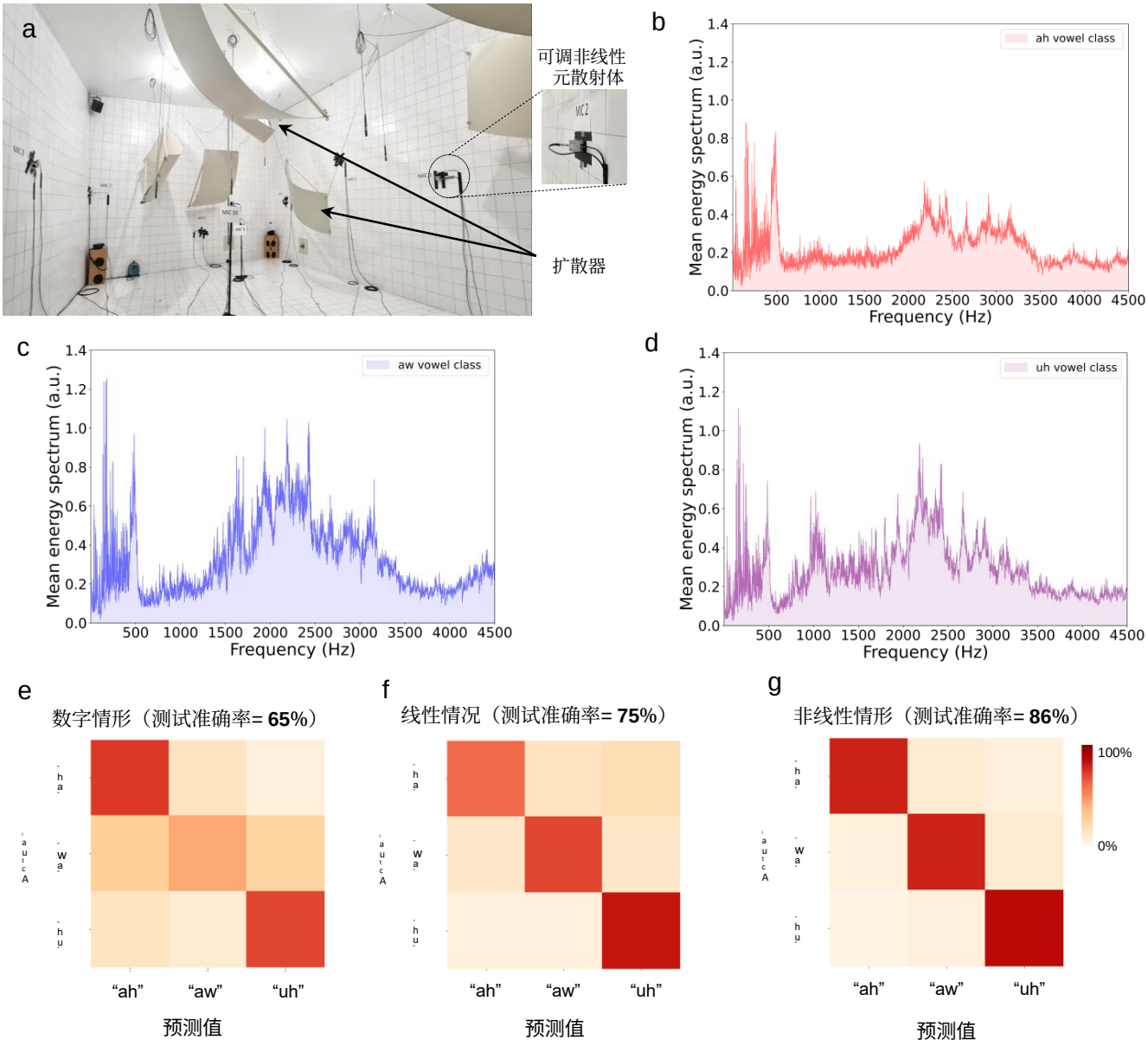


图4. 复杂混响室中的元音识别。a混响室及其安装设备的照片。**b-d** 三组变换后元音数据（男性和女性说话人）的平均能量谱，均以原始数据的平均能量谱归一化。**e-g** 分别为数字、线性（控制关闭）和非线性（控制开启）情况下的测试准确率。

测试集的训练结果及对应的混淆矩阵如 图 3f-g 所示。我们使用混淆矩阵比较测试数据集上的元音识别质量。在此测试场景中，混淆矩阵（本例中为一个 3×3 矩阵）体现了网络的泛化能力。我们观察到，所提出的基于声音的系统在未见数据上取得了出色的准确率。其测试准确率为 91.4%，与最先进的研究相当，例如物理

循环神经网络[35,51]。然而，后者需要使用成对输入-输出数据进行成本高昂的训练，而所提出的方法无需如此；此外，后者仅在三个元音类别上进行了测试，而这三个类别在频域中已可通过线性支持向量机实现线性可分（更多详情请参阅补充信息）。我们将研究结果与最先进的研究以及数字神经网络（如非线性支持向量机）进行了比较（参见补充信息中的补充表 1 和 2）。其中一些先前研究也利用了材料固有非线性，这会降低能效。另一方面，所提出的基于波的计算系统利用非线性超散射体，因而能效较高。此外，所提出的基于波的计算系统不仅用途广泛、易于训练，还能在声学数据的原生域中进行处理，直接处理真实声音，无需进行剪辑或滤波等预处理。

为了研究腔体中声波非线性混频的影响，我们从图 3d的框图中移除非线性混沌腔体，并计算测试准确率。在这种情况下，测试准确率降至 65%，这表明非线性波混频发挥着关键作用。

2.4 真实场景中的元音识别

最后，我们在混响室中重复上述实验，以展示所提方法在真实声学环境中进行元音识别的能力。混响室通常会配备额外的声学扩散体，用于产生入射方向在统计意义上随机的声波。在这种环境中，声音传播会产生大量多重散射事件，使声场变得杂乱并形成散斑。我们在房间内随机布置了十个非线性超散射体和麦克风（见图4a）。每个元音的音频波形由放置在房间角落的单个扬声器播放。麦克风记录变换后的信号，并按照图3d 所示的相同训练流程进行处理。图4b-d 展示了三组变换后元音数据（男性和女性说话人）的平均能量谱，并以图3e中数据的初始平均能量谱为基准进行了归一化。我们发现，非线性多重散射能够非常有效地将这些起初相互重叠的频谱转换成彼此不同的声音，尤其是在两个频率范围内（50至600 Hz和1000至3500 Hz）。与较小的腔体原型一样，这使得线性分类器能够显著提升分类效果。不同情形（包括数字情形、线性情况和

非线性情形）的训练结果分别见图3e-g。数字情形的混淆矩阵结果采用与上一节类似的方法计算，即从图3d的框图中移除非线性混沌腔；结果表明，数字支持向量机的性能较差。为进行公平比较，我们报告了线性和非线性情形的混淆矩阵结果，并指出两者的唯一区别在于：线性情况下禁用了非线性超散射体。线性和非线性情形的测试准确率分别为75%和86%，证实了非线性对该器件性能的重要性。

总之，我们表明，非线性多重散射介质中的波传播动力学在概念上等同于储备池计算系统。这种概念上的联系为基于波的机器学习开辟了新途径，使信息能够在其原生域中得到高效处理。我们通过实验表明，在真实声学系统中，非线性多重散射可用于实现基于波的语音识别：它能将真实声音强有力地映射至高维空间，并可针对性地调整这种非线性映射，无需任何预处理，且可在真实环境中运行。

3 方法

3.1 实验装置

用于进行储备池计算的腔体图 1 至图 3 中所示的腔体长 2 米、宽 1 米、高 0.2 米。在 500 Hz 以下，该腔体有 11 个模态（其记忆特性表征见补充材料 S3）。对于元音识别任务，每个元音的音频波形由腔体左侧的十个扬声器注入（每个扬声器之间的延迟为 10 毫秒）。图 4 中最终元音识别实验是在真实的混响室中进行的。在这种情况下，每个元音的波形由一个位于混响室角落的扬声器注入。两种配置中使用的非线性散射体均随机放置。这些散射体通过控制电动扬声器（Visaton FRWS 5 SC）实现。主动控制由基于FPGA的Speedgoat性能实时目标机执行，并通过 MATLAB 的 xPC目标环境（经由 SIMULINK）进行控制。输出电压经 Howland电流泵（HCP）转换器转换为电流[46]，转换增益为 10 mA/V，最后反馈至每个受控扬声器。测量（声源生成和信号记录）也由实时目标机完成。

3.2 读出训练

这里，我们展示如何利用读出节点接收到的时间信号数据训练决策层，而无需使用额外的滤波操作。设 ψoutm (t) 为输出麦克风 m 接收到的时间信号，其离散化信号的离散傅里叶变换为 $\bar{\psi}$ outm $=\left(\psi\text{outm}(0), \psi\text{outm}(t0), \ldots, \psi\text{outm}((n-1)t0)\right)$T，定义为 yf $=\Sigma$(n–1) (k=0) ψout m (tk)exp(-2πin jk)，其中 j $=0, \ldots, n-1$，(.)T 表示转置运算，t0 为采样时间。傅里叶系数 yfj 与离散化信号之间的关系可表示为乘积 ymf $=$F $\bar{\psi}$outm，其中 F 为傅里叶矩阵：

$$\mathbf{F}=\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & \kappa^{2} & \cdots & \kappa^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \kappa^{n-1} & \cdots & \kappa^{(n-1)^{2}} \end{pmatrix}$$

其中$\kappa = e^{-2\pi i/n}$。通常，我们只关注某些频率谐波，其分量m可通过$\mathbf{y}_m\text{f} = \mathbf{F_s}\,\bar{\psi}\text{out}_m$ 计算，其中$\mathbf{F_s}$ 是对应于所需频率谐波的傅里叶矩阵。为了通过线性回归训练读出函数（其通常定义为形如$\mathbf{Y} = \mathbf{WX}$T的线性矩阵运算，其中$\mathbf{W}$ 为权重矩阵），我们必须将这两个运算组合起来，将 $\mathbf{F_s}$ 矩阵与权重矩阵相乘：

$$\mathbf{Y}=\mathbf{W}\left(\text{vec}\left(\mathbf{F_s X}\right)\right)^{\mathrm{T}} \tag{4}$$

其中，$\mathbf{X} = \left(\bar{\psi}\text{out1}、\ \bar{\psi}\text{out2}、\ \ldots\ldots、\ \bar{\psi}\text{out}_m\right)$。vec(**X**) 是通过线性变换将矩阵 **X** 向量化为列向量，其定义为 vec(**X**) $=\Sigma_{ni=1}\,\mathbf{e}_i \otimes \mathbf{X}\mathbf{e}_i$，其中 $\mathbf{e}_i$ 是 n 维空间的第i 个标准基向量，即 $\mathbf{e}_i$ = [0, ...、0、1、0、.....、0]T，$\otimes$表示克罗内克积。与常规线性回归（**Y** =**WX**T）类似，所提出的基于波的计算方案只需进行简单的矩阵乘法，无需使用敏感或复杂的滤波器。也可以将其视为对数字层权重矩阵进行简单的重新缩放。因此，该方案不会带来额外的计算成本。

为学习非线性函数（sinc函数），我们采用线性回归这一监督学习算法来训练读出函数。将预测输出与真实值进行比较，计算误差，并按照标准线性回归学习规则更新读出网络中的权重。

对于元音识别任务的读出网络训练，我们使用了 PyTorch，它提供了一个高级应用程序接口，可用于访问 TensorFlow。如图 3d 所示，我们首先对归一化后的 $\mathbf{y}_m\mathrm{f}$ 进行了主成分分析（我们对数据在 10-1kHz 和 1-3.5kHz 两个不同范围内的傅里叶分量进行了归一化（最小-最大归一化））。主成分分析（PCA）是一种线性降维方法，利用数据的奇异值分解将其投影到更低维的空间中。随后，我们使用线性支持向量机对元音类别进行分类。

致谢：A. Momeni 和 R.Fleury 感谢瑞士国家科学基金会通过编号为 181232 的 Eccellenza资助项目提供资助。X. Guo 和 H. Lissek 感谢瑞士国家科学基金会（SNSF）通过编号为 200020_200498 的资助项目提供资助。

贡献 A.M 和 R.F 共同构思了这一想法。A.M 和 X.G. 的贡献相同。A.M. 设计了计算引擎，并进行了理论研究和数值模拟。X.G. 设计了非线性主动控制并开展了实验。H.L. 监督了实验工作。R.F 发起并监督了该项目。所有作者都参与了解释结果和撰写稿件。

参考文献

1. Gigan, S. 利用无序进行成像与计算。《自然·物理》**18**, 980–985 (2022).**2.** Cao, H., Mosk, A. P. & Rotter, S. 调控复杂介质中的光传播。《自然·物理》**18**,994–1007 (2022).**3.** Zangeneh-Nejad, F., Sounas, D. L., Alù & Fleury. 利用超材料进行模拟计算。《自然综述：材料》**6**, 207–225 (2021).**4.** Wright, L. G. 等 利用反向传播训练的深度物理神经网络。《自然》**601**,549–555(2022).

5. Feldmann, J. 等 利用集成光子张量核心进行并行卷积处理。 《自然》**589**, 52–58 (2021).**6.** Silva, A.等 利用超材料执行数学运算。 《科学》 **343**, 160–163 (2014).**7.** Wang, T. 等 一种每次乘法所需光子数少于 1 个的光学神经网络。 《自然·通讯》**13**,123 (2022).**8.** Momeni, A. 等 用于轴上反射式光学计算的互易超表面。 *IEEE* 天线与传播汇刊**69**, 7709–7719 (2021).

9. Sol, J.、Smith, D. R. & Del Hougne, P. 元可编程模拟微分器。 《自然·通讯》**13**,1–10 (2022).

10. Babaee, A.、Momeni, A.、Abdolali, A. & Fleury, R. 基于 2× 2具有非对称响应的多输入多输出超表面处理器的并行模拟计算。 《物理评论应用》**15**,044015 (2021).

11. del Hougne, P. & Lerosey, G. 利用混沌实现基于波的模拟计算：以室内无线通信信号为例进行演示。 《物理评论*X*》**8**,041037(2018).

12. Matthès, M. W.、Del Hougne, P.、De Rosny, J.、Lerosey, G. & Popoff, S. M. 光学复杂介质作为通用可重构线性算子。 《光学》**6**, 465–472 (2019).

13. Momeni, A., Rajabalipanah, H., Abdolali, A. & Achouri, K. 基于极化率张量合成的多算子超表面的广义光信号处理。 《物理评论应用》**11**,064042 (2019).

14. Momeni, A., Rouhi, K. & Fleury, R. 利用计算型石墨烯基多层结构实现可切换同步时空模拟计算。 《碳》**186**, 599–611 (2022).

15. Anderson, M. G., Ma, S.-Y., Wang, T., Wright, L. G. & McMahon, P. L. 光学变换器。 *arXiv* 预印本 *arXiv:2302.10360*(2023).

16. Wetzstein, G.等 利用深度光学与光子学实现人工智能推理。 《自然》 **588**, 39–47（2020）。

17. Wu, Z., Zhou, M., Khoram, E., Liu, B. & Yu, Z. 神经形态超表面。《光子学研究》8, 46–50（2020）。18. Lin, X. 等。利用衍射深度神经网络实现全光机器学习。《科学》361, 1004–1008（2018）。19. Sebastian, A., Le Gallo, M., Khaddam-Aljameh, R. & Eleftheriou, E. 用于存内计算的存储器件及应用。《自然·纳米技术》15, 529–544（2020）。20. Markovi´c, D., Mizrahi, A., Querlioz, D. & Grollier, J. 神经形态计算物理学。《自然·物理评论》2, 499–510（2020）。21. Grollier, J. 等。神经形态自旋电子学。《自然·电子学》3, 360–370（2020）。22. Pierangeli, D., Marcucci, G. & Conti, C. 通过空间光调制实现大规模光子伊辛机。《物理评论快报》122, 213902（2019）。23. Huang, J., Fang, Y. & Ruan, Z. 通过光电相关计算实现反铁磁空间光子伊辛机。《通信物理》4, 242（2021）。24. Suykens, J. A. & Vandewalle, J. 最小二乘支持向量机分类器。《神经处理快报》9, 293–300（1999）。25. Seber, G. A. & Lee, A. J. 线性回归分析，第330卷（约翰·威立父子出版公司，2003）。26. Rafayelyan, M., Dong, J., Tan, Y., Krzakala, F. & Gigan, S. 用于预测时空混沌系统的大规模光学储备池计算。《物理评论X》10, 041037（2020）。27. Pierangeli, D., Marcucci, G. & Conti, C. 通过自由空间光传播实现光子极限学习机。《光子学研究》9, 1446–1454（2021）。˙28. Teˇgin, U., Yıldırım, M., Oˇguz, I., Moser, C. & Psaltis, D. 可扩展光学学习算子。《自然·计算科学》1, 542–549（2021）。29. Pathak, J., Hunt, B., Girvan, M., Lu, Z. & Ott, E. 从数据中对大规模时空混沌系统进行无模型预测：一种储备池计算方法。《物理评论快报》120, 024102（2018）。30. Momeni, A. & Fleury, R. 通过非线性时间-弗洛凯纠缠实现基于电磁波的极限深度学习。《自然·通讯》13, 1–11（2022）。

31. Yildirim, M. 等 用于机器学习的非线性光学数据变换器。 *arXiv* 预印本 *arXiv:2208.09398* (2022)。

32. Oguz, I. 等 编程非线性传播，以实现高效的光学学习机器。 *arXiv* 预印本 *arXiv:2208.04951*(2022)。

33. Vandoorne, K. 等 在硅光子芯片上演示储备池计算。《自然》通讯**5**，3541（2014）。

34. Zhou, T. 等 利用可重构衍射处理单元实现大规模神经形态光电计算。《自然》光子学**15**, 367–373（2021）。

35. Hughes, T. W., Williamson, I. A., Minkov, M. & Fan, S. 将波动物理学作为模拟循环神经网络。《科学》进展**5**，eaay6946（2019）。

36. Skinner, S. R., Steck, J. E. & Behrman, E. C. 使用克尔型非线性材料的光学神经网络。载于第四届神经网络与模糊系统微电子学国际会议论文集，12–15 （IEEE，1994）。

37. Rumelhart, D. E., Hinton, G. E. & Williams, R. J. 通过反向传播误差学习表征。《自然》**323**，533–536（1986）。

38. Zhong, Y.等。 基于动态忆阻器的储备池计算，用于高效时间信号处理。《自然·通讯》**12**, 408（2021）。

39. Sun, L. 等。 通过二维忆阻器实现用于语言学习的传感器内储备池计算。《科学·进展》**7**, eabg1455（2021）。

40. Marinella, M. J. & Agarwal, S. 利用忆阻器实现高效储备池计算。《自然·电子学》 **2**, 437–438（2019）。

41. Liu, K. 等 一种基于 α-in2se3 的光电突触，具有可控的时间动态特性，可用于多模态和多尺度储备池计算。《自然·电子学》**5**, 761–773 (2022).

42. Du, C.等。 利用动态忆阻器进行储备池计算，以处理时间信息。《自然·通讯》**8**, 2204（2017）。

43. Milano, G. 等。基于自组织纳米线网络的全忆阻架构实现 In materia 储备池计算。《自然·材料》21, 195–202（2022）。44. Moon, J. 等。使用基于忆阻器的储备池计算系统进行时间数据分类与预测。《自然·电子学》2, 480–487（2019）。45. Tanaka, G. 等。物理储备池计算的最新进展：综述。《神经网络》115, 100–123（2019）。46. Rivet, E., Karkar, S. & Lissek, H. 通过混合传感器／并联支路阻抗控制实现宽带低频电声吸收器。《IEEE 控制系统技术汇刊》25, 63–72（2017）。DOI 10.1109/TCST.2016.2547981。47. Guo, X., Volery, M. & Lissek, H. 用于设计可调单自由度声吸收器的电声谐振器类 PID 主动阻抗控制。《声学与振动杂志》525, 116784（2022）。DOI https://doi.org/10.1016/j.jsv.2022.116784。48. Guo, X., Lissek, H. & Fleury, R. 通过非线性主动电声谐振器提升吸声效果。《物理评论应用》13, 014018（2020）。DOI 10.1103/PhysRevApplied.13.014018。49. Hillenbrand, J., Getty, L. A., Clark, M. J. & Wheeler, K. 美式英语元音的声学特征。《美国声学学会杂志》97, 3099–3111（1995）。50. Romera, M. 等。利用四个耦合的自旋转矩纳米振荡器识别元音。《自然》563, 230–234（2018）。51. Qu, Y., Zhou, M., Khoram, E., Yu, N. & Yu, Z. 用于模拟循环神经网络的共振。《ACS 光子学》9, 1647–1654（2022）。

补充信息：

基于可调非线性多重散射的受物理启发的神经声学计算

Ali Momeni、Xinxin Guo、Hervé Lissek 和 Romain Fleury

此 PDF 文件包括：

S1 节。将声波方程推导为储备池计算更新方程 S2节。非线性有源散射体 S3节。记忆：所用腔体和混响室的脉冲响应 S4节。将所提出的计算系统与最先进的研究及数字神经网络进行比较 参考文献

S1 节。推导　　将声波方程表述为储备池计算更新方程　　**s**

正如我们在正文中提到的，标量波场分布 $p(x,y,z)$ 的动力学由波动方程支配：

$$\frac{\partial^2 p}{\partial t^2} - c^2 \cdot \nabla^2 p = \psi_t^{\text{in}} \tag{S1}$$

我们可以 用时间步长为 Δt的中心有限差分对 (S1) 进行离散化：

$$\frac{p_{t+1} - 2\mathrm{p}_t + p_{\mathrm{t}-1}}{\Delta_{\mathrm{t}}^2} - \mathrm{c}^2 \cdot \nabla^2 p_{\mathrm{t}} = \psi_t^{\text{in}} \tag{S2}$$

下标 t 用于表示标量场在给定时间步的取值。我们将 (S2) 改写为矩阵形式：

$$\mathbf{h}_t = \mathbf{Q}^{(\mathrm{c})}(\mathbf{h}_{\mathrm{t}-1}) \cdot \mathbf{h}_{\mathrm{t}-1} + \mathbf{Q}^{(\mathrm{i})} \cdot \mathbf{x}_t \tag{5}$$

where $\mathbf{h}_t = \begin{pmatrix} p_{t+1} \\ p_t \end{pmatrix}$. Also $\mathbf{Q}^{(\mathrm{c})}$ and $\mathbf{Q}^{(\mathrm{i})}$ are defined as

$$\mathbf{Q}^{(\mathrm{c})} = \begin{pmatrix} 2 + \Delta_{\mathrm{t}}^2 \cdot c^2 \cdot \nabla^2 & -1 \\ 1 & 0 \end{pmatrix} \tag{S4}$$

$$\mathbf{Q}^{(\mathrm{i})} = \Delta_{\mathrm{t}}^2 \begin{pmatrix} \mathbf{G}^{(\mathrm{i})} \\ \mathbf{0} \end{pmatrix}$$

其中，**0**是一个全零矩阵。系统的输出可表示为

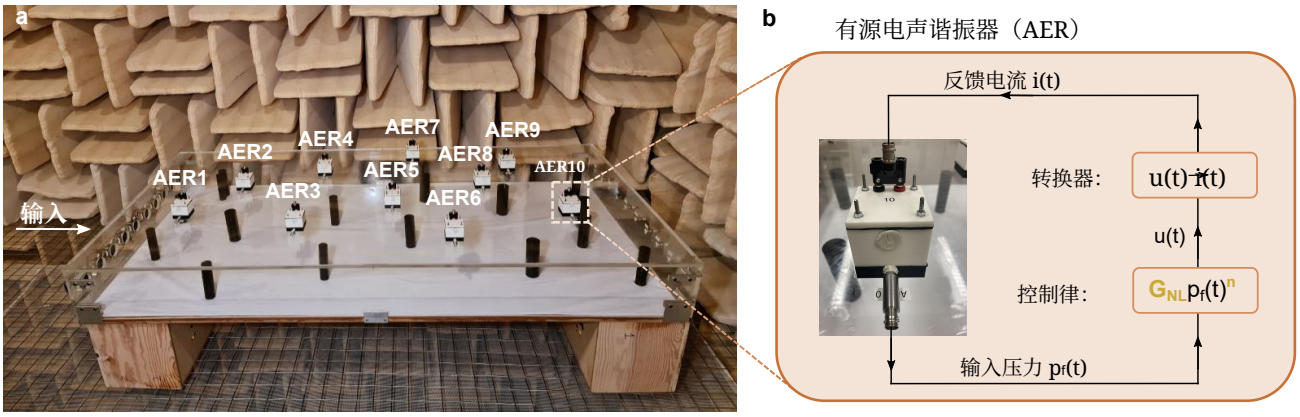
$$\psi^{\text{out}} = \left| [\mathbf{G}^{(\mathrm{o})\mathrm{T}}, 0] \cdot \mathbf{h}_t \right| \tag{S5}$$

其中，**G(i)** 和 **G(o)** 分别表示注入点和测量点在该

波传播域。

S2节。非线性有源散射体

图中所示的反馈控制回路补充图1b 用于实现所需的非线性散射体。实时感测每个扬声器前方的压力（pf），并将其用作控制输入。控制输出（以电压表示）定义为指定形式的 $(t) = G\mathrm{NL}p\mathrm{f}(t)n$，并转换为反馈电流。指数 n 设为 1.5，这是针对所考虑的两个元音识别任务的最佳值。调节常数控制增益GNL，使其在低输入功率下触发非线性效应。激励源的pf被限制在数帕的范围内，从而使扬声器在被动情况下（控制关闭）呈线性工作。因此，只有在对扬声器施加主动控制时，才会产生非线性。

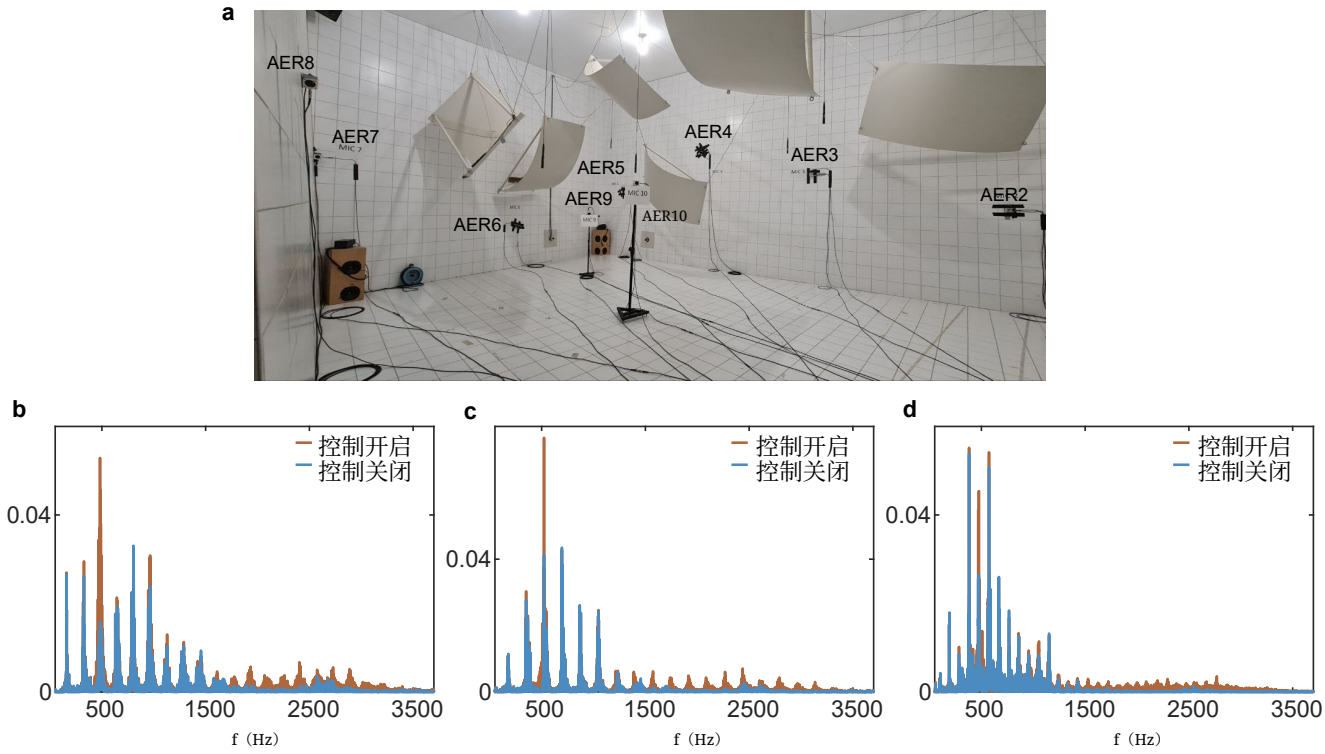


补充图1。 腔体顶侧非线性有源散射体的排列。a十个受控散射体的位置。**b** 施加的非线性主动控制。

控制增益 GNL 的设置分别汇总于表1，适用于两个元音识别任务。扬声器的相应编号分别见补充图1a和补充图2a-b。

补充表1：非线性主动控制的参数设置		
任务	**散射体**	**设置详情**
空腔中的 V 识别-ity	AER1 至 AER10（指数如图1a所示）	AER1 的 GNL = 7.0 AER2 的 GNL = 5.0 AER3 的 GNL = 6.0 AER4 的 GNL = 6.8 AER5 的 GNL = 7.0 AER6 的 GNL = 5.0 AER7 的 GNL = 6.0 AER8 的 GNL = 6.3 AER9 的 GNL = 6.0 AER10 的 GNL = 5.8
元音识别 **真实混响室**	AER1 至 AER10（指数如图2a所示）	AER1 的 GNL = 9.0 AER2 的 GNL = 6.4 AER3 的 GNL = 5.4 AER4 的 GNL = 7.9 AER5 的 GNL = 7.5 AER6 的 GNL = 6.4 AER7 的 GNL = 6.9 AER8 的 GNL = 7.1 AER9 的 GNL = 7.1 用于 AER10 的 GNL = 9.8

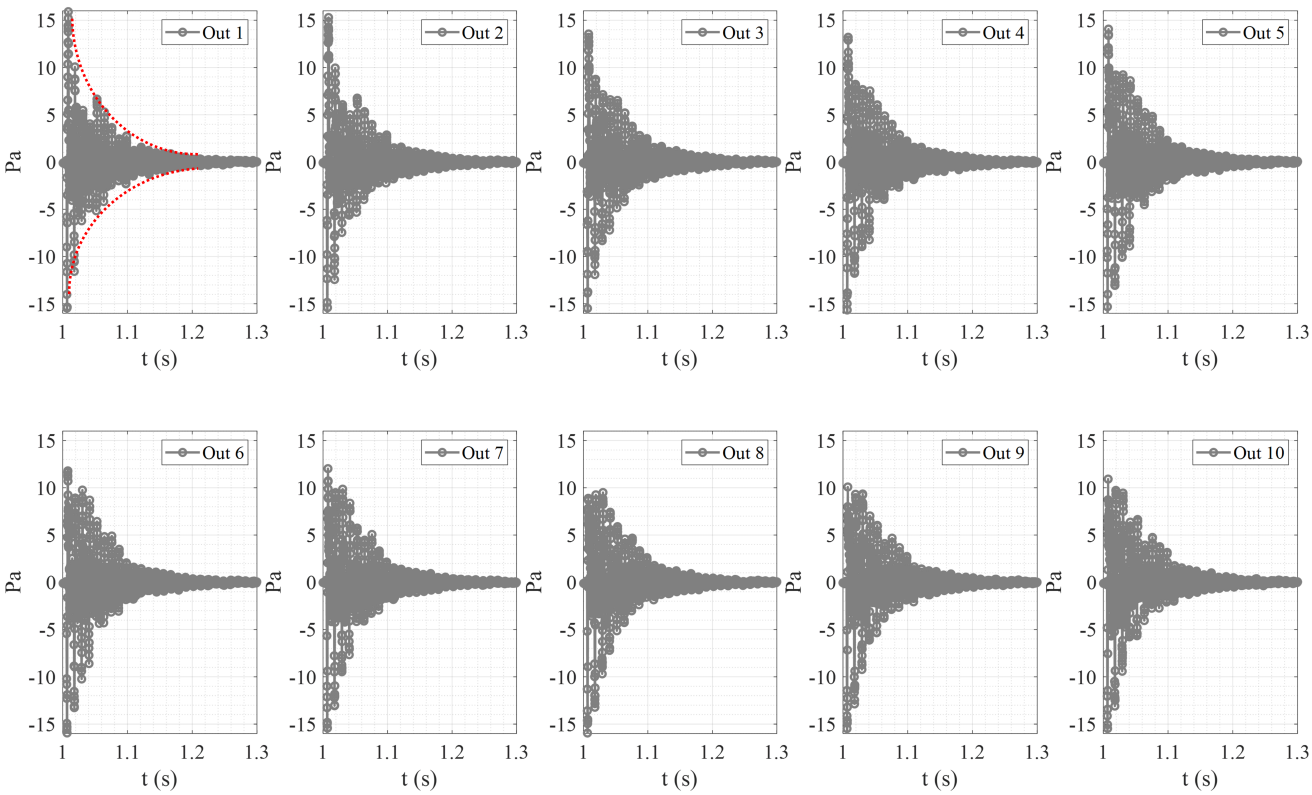
为了更好地展示触发的非线性效应，我们在补充图2c-e中展示了各类别元音信号的传输示例。线性（控制关闭）与非线性（控制开启）两种情况下存在明显差异。只有引入非线性后，所考察的三类元音才变得可区分，这一点在主文稿的图4b-d中有所体现。



补充图2。混响室中非线性有源散射体的排列。a混响室中受控制的散射体的位置。**b c d** 控制关闭（线性）与控制开启（非线性）情况下的比较。所示结果为当每类元音中的一个元音（**b**：ah类，**c**：aw类，**d**：uh类）作为声源时，在AER1附近由麦克风测得的传输压力。

S3节。记忆：所用腔体和混响室的脉冲响应

记忆效应可以通过脉冲响应直接直观呈现，即所提出的计算系统受到狄拉克δ函数 $\psi in(t) = \delta(t - t0)$ 激励时的输出。脉冲位于时间步 t0。记忆效应体现在时间步 t0 的输入影响会延续到一定的时间范围内。这里，所提出的腔体和房间受到一个 0.1 毫秒的脉冲信号激励，以测量不同读出节点处的时间响应（见补充图1和图2）。显而易见

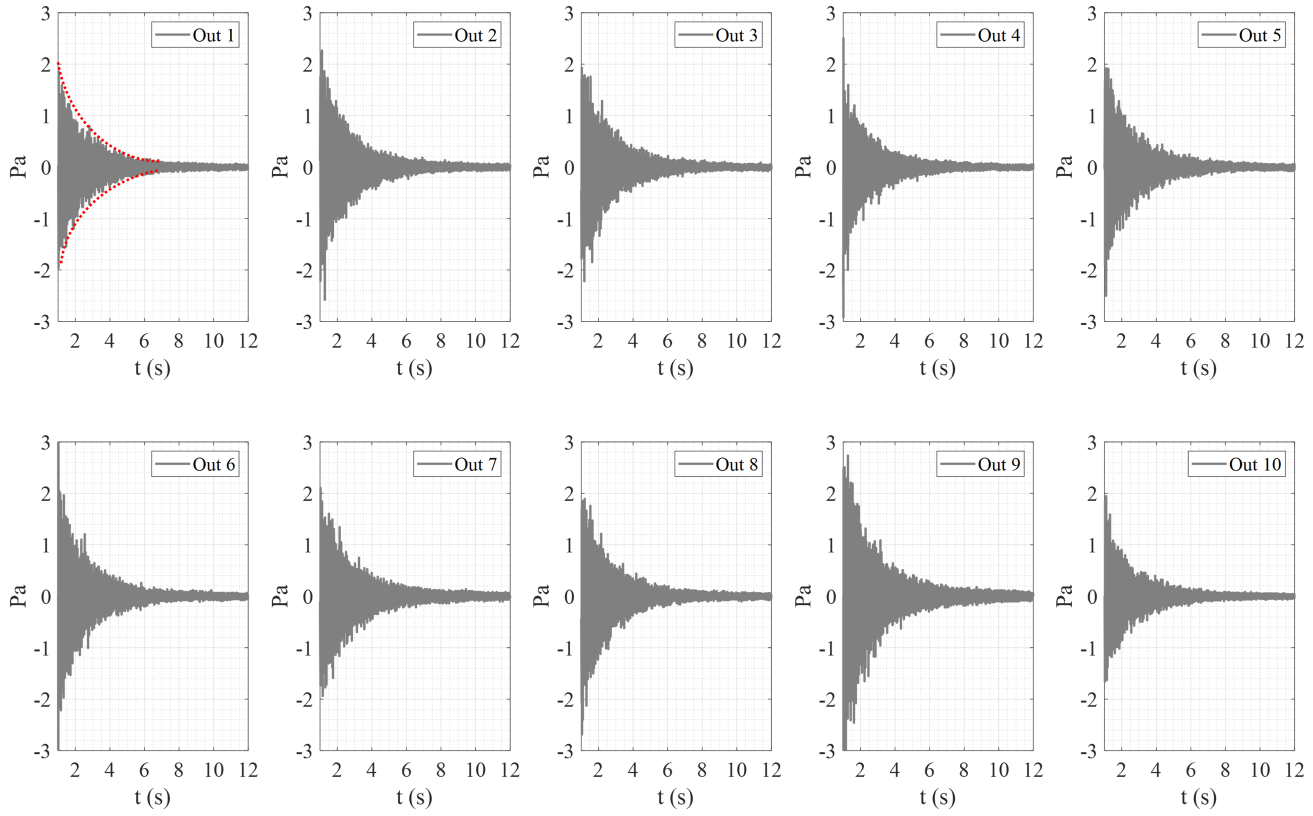


补充图3。 非线性混沌腔的脉冲响应。非线性混沌腔在 t = 1s 时受到1毫秒脉冲激励时，不同读出节点处的脉冲响应。

从补充图1和图2可以看出，每个读出节点的时间脉冲响应在振幅和时间演化方面各不相同，呈现出具有不同寿命的独特时间映射。

S4节。所提出的计算系统与最先进的研究和数字神经网络的比较

在此，我们将所提出的计算系统的性能与先前研究以及数字神经网络进行比较。补充表2和表3展示了不同计算架构之间的准确率和均方根误差对比。



补充图4。 混响室的脉冲响应。。混响室在 t = ls 时受到 1毫秒脉冲激励时，不同读出节点处的脉冲响应。

补充表2：所提出的计算系统与先前研究的比较		
任务	**基准测试**	**详情**
学习非线性函数（sinc(*x*))	参考 [28]：均方根误差：0.0671 我们的工作：均方根误差：0.017	• 参考文献 [28] 使用了 GRIN 50/125多模光纤（多模光纤），支持 240 种空间模式。在本研究中，作者在 or- 中使用了高功率光脉冲以激发渐变折射率多模光纤的非线性（输入光学峰值功率脉冲为 3.43 kW）。
元音识别	参考 [51]：准确率：83.3% 参考 [35]：准确率：86.3% 我们的工作：准确率：91.4%	• 参考文献 [51]和 [35]采用了基于波的循环神经网络。他们使用了 ae、ei 和 iy 元音类别，这些类别在频域中线性可分（使用线性支持向量机时测试准确率为 90.3%，且在频域中有 200 个可训练参数）。 Ref[35] 使用了 4200 个可训练参数。 • 我们使用了 ah、aw 和 uh 元音类别（测试准确率为 65.0%，使用线性支持向量机且仅有 200 个可训练参数）。相应类别的元音识别并非易事。

补充表3：数字神经网络与所提出的基于波的系统的比较		
网络结构	**元音数据集上的准确率 (%)**	**详情**
线性支持向量机	测试准确率= 65 ± 0.5	具有 200 个可训练参数的线性支持向量机。
非线性支持向量机（Sigmoid）	测试准确率= 68± 1	采用 Sigmoid 非线性 核且具有 200 个可训练参数的非线性支持向量机。
非线性支持向量机（径向基函数）	测试准确率= 74± 1	采用径向基函数 非线性核和 200 个可训练参数的非线性支持向量机 量 ·
循环神经网络	测试准确率= 80.2	采用 Leaky ReLU 非线性且包含 5250 个可训练参数的数字循环神经 ters[35].
所提出的系统	测试准确率= 91.4	仅有 200 个可训练参数，并在决策层采用线性 支持向量机