

基于手写的高性能大脑-文本的交流系统^{*}

XXX^{1*}, 某某某²

¹College of Electrical and Control Engineering, Xi'an University of Science and Technology, Xi'an 710054, P. R. China

²西安科技大学电气与控制工程学院, 西安710054

Abstract

脑机接口 (BCI) 可为失去行动或说话能力的人恢复交流的能力。截至目前, BCI研究的一个主要焦点是恢复粗大运动技能 (Gross Motor Skills), 例如, 伸手、抓握^[1-5] 或使用计算机光标点击打字^[6,7]。然而, 一系列快速和高度灵巧的动作 (例如手写或触摸打字) (Rapid Sequences of Highly Dexterous Behaviours) 可能会达到更快的交流速度。本文开发了一种植入式BCI系统。当受试者尝试手写动作时, 该系统采用循环神经网络解码运动皮层神经活动, 并将神经活动实时转换为文本。利用该系统, 受试者 (因脊髓损伤导致手部瘫痪) 实现了每分钟90个字符的打字速度, 其打字的在线准确率达到94.1%, 离线自动校正后的准确率超过99%。经调研, 这一打字速度超过了有记录的所有BCI系统, 且与受试者同龄人群使用智能手机打字的速度 (每分钟115个字符) 相当^[8]。最后, 我们在理论上解释了为什么**时间上复杂的动作** (如手写) 比点对点的动作更容易解码。我们的研究结果为BCI系统开辟了一条新的途径, 并证明了在患者瘫痪数年后准确解码快速灵巧的动作仍具有可行性。

Key words:

研究表明, 人体在瘫痪后, 关于粗大运动技能 (如伸手、抓握或移动计算机光标) 的动作意图仍会在运动皮层中进行神经编码^[1-7]。然而, 对于快速和高度灵巧的动作 (如手写), 其神经表征是否完整还需进一步研究确认。因此, 当T5 (“BrainGate” 研究的参与者) 尝试手写单个字母和符号时, 我们通过记录中央前回 (precentral gyrus, 一个运动神经前部的区域) ^[9,10] 手部“旋钮” (hand ‘knob’) 区域中两个微电极阵列采集到的神经活动 (图1a) 来验证这一点。受试者T5受到过严重的脊髓损伤, 颈部以下全部瘫痪, 其手部仅能抽搐和微动。我们指导T5 “尝试” 想象自己正拿着一支笔, 像正常人一样在横格纸上进行书写。

1 手写动作的神经表征

为了将受试者在尝试手写过程中产生的神经活动 (即多单元阈值跨越率 (multiunit threshold crossing rates)) 可视化, 用主成分分析法 (principal components analysis) 展示出 (神经活动中) **方差最大 (most variance)** 的前三种神经维度 (图1b)。由于 (受试者) 在书写的速度上有波动, 导致了不同的试验中神经活动达到峰值和低谷的时间有所不同, 但仍能从图中看到 (尝试手写时产生的) 神经活动似乎是强烈且可重复的。我们采用时间对齐技术 (time-alignment technique) 来消除时间上的差异 (temporal variability) 这个过程揭示了: 每一个字符都具有明显一致的潜在的神经活动模式, 这些神经活动的模式具备独特性 (图1c)。为了验证尝试书写过程中, 神经活动是否编码了每个字符形状所需的笔迹, 我们从试验平均 (trial-averaged) 过的神经活动中线性解码笔尖速度, 以重建每个字符 (图1d)。由图1d可知, 重建的字母形状易于辨识, 这说明笔尖速度是能够被稳健编码的。代表笔尖速度的神经维度占总神经活方差 (neural variance) 的30%。

接下来, 在给出 “开始” 的提示后, 我们使用非线性降维法 (nonlinear dimensionality reduction method) (即t-分布随机邻域嵌入, t-distributed stochastic neighbour embedding; t-SNE), 为每一次独立实验中记录到的神经活动生成二维 (2D) 可视图 (two-dimensional (2D) visualization) (图1e)。T-SNE可视化法揭示了每个字符对应

^{*} This work is supported by the National Science Foundation of China(61603295, 51905416), Outstanding Youth Science Fund of Xi'an University of Science and Technology (2018YQ2-07), Shaanxi Postdoctoral Science Foundation (2018BSHEDZZ124), and China Postdoctoral Science Foundation (2017M623207).

^{*} Corresponding author: Hongguang Pan, E-mail: hongguangpan@163.com

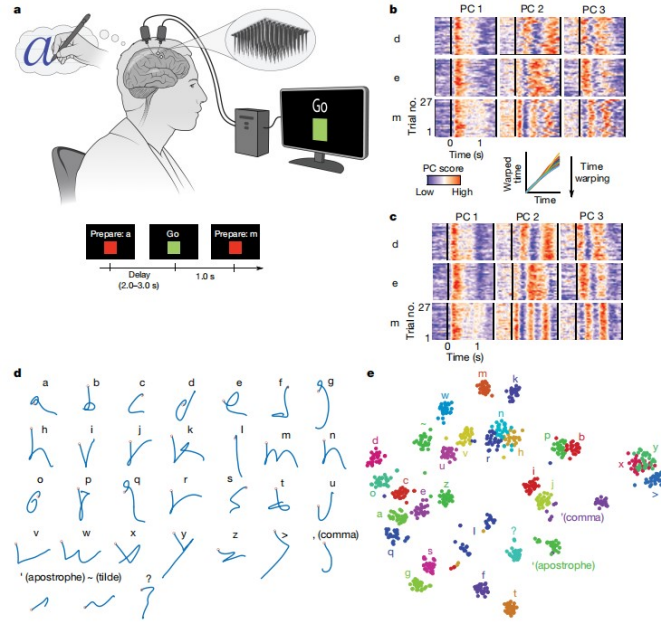


Fig. 1. a) 为了评估尝试手写时产生的神经表征，受试者T5跟随电脑屏幕上给出的指令（下方的图展示了屏幕上随着时间而给出的指令内容），每次尝试书写一个字符。来源：E. Woodrum 创作的人体剪影图。b) 展示了三个样例字母（d, e 和m）的神经活动的前三项主要成分（PCs），每个字母都重复做了27次（实验）。分别规范了每个板块里的颜色范围，以便于可视化。c) 时间扭曲神经活动，消除实验与实验间因书写速度造成的差异，这种方法揭示出：活动的模式具有一致性，且对于每个字符来说都是独一无二的。上方插图c，展示了字母“m”的时间扭曲函数，扭曲函数紧密地贴近恒等线（用不同颜色的线绘制每次试验的扭曲函数）。d) 展示了实验用到的31个解码后的字符，使用交叉验证（cross-validation）法线性解码意图产生二维笔尖速度的神经活动（每个字符都被保留），在多次实验中为译码器的输出进行平均去噪。橙色的圆圈表示起笔位置。e) 使用t-SNE对神经活动进行二维可视化处理。每个圆圈都表示一次独立的试验（31个字符进行的27次试验都展示在图中）。

的密集神经活动簇和以肌肉运动为主的编码方式，在该编码方式中，书写方式相似的字符靠得更近。将k-近邻分类算法（k-nearest-neighbour classifier）应用于离线的神经活动，可进行字符分类，分类准确率可达94.1%（95%的置信区间（confidence interval; CI）=[92.6, 95.8]）。上述结果表明，即使在瘫痪数年后，运动皮层中的手写神经表征也能强大到用于BCI系统中。

2 解码手写语句

接下来，我们测试了是否可以实时解码完整的手写语句，从而使四肢瘫痪的患者能够通过尝试手写他们想要表达的信息进行交流。为此，我们训练了一个循环神经网络（Recurrent Neural Network, RNN），将神经活动转换成概率来描述每个字符在每个时刻被实时写入的可能性（图2a，扩展图1）。在线地，我们采用简单方式设置阈值以输出离散字符（“原始在线输出”，图2a）；或者离线地，采用大词汇量语言模型更广泛地处理这些概率以实现自动校正功能（“语言模型的离线输出”，图2a）。我们使用了图1d所示31个有限字符集，包括26个小写字母，以及逗号、撇号、问号、句号（T5将其书写为波浪符号，‘~’）和空格（T5将其书写为大于号，‘>’）组成。选择‘~’和‘>’符号是为了使句号和空格更容易被检测。T5采用印刷体（不是草书）书写每个字符，**每个字符都紧接着前一个字符输出**。

为了收集RNN的训练数据，我们记录了T5按照电脑指示，尝试手写完整句子时的神经活动。在实时评估前，我们在3个试验日内总共收集了242个语句，这些语句被组合起来训练RNN。在后续的实时测试中，我们继续收集额外的训练数据以重新校准RNN，在此过程中共产生572个训练语句（包括7.6小时和31472个字符）。为训练RNN，我们采用自动语音识别^[12,14]中的神经网络方法以克服两个关键挑战：（1）训练数据中每个字母的书写时间未知（因为T5的手是瘫痪的），这导致使用监督学习技术具有挑战性；（2）与典型的RNN数据集相比，我们数据集的规模有限，很难避免训练时出现过拟合。（见补充方法，扩展图2，3）。

在后续实验的5天内（图2c，5天不连续），我们对RNN的性能进行了评估。评估共有4个评估模块，其中包含7-10个RNN从未接受过的语句（确保RNN不会出现过拟合）。根据屏幕语句提示，T5逐一尝试书写语句中的字符，同时RNN检测到被解码的字符，并将其实时展示在屏幕上（补充视频1、2、扩展数据表1）。尝试书写后，经过短暂的延时（大约为0.4-0.7 s），T5尝试书写的字符就会显示在屏幕上。解码后的语句非常清晰（“原始输出”，图2b）。值得注意的是，该系统输出速度很高，稳定在每分钟90个字符，且字符的平均错误率为5.4%（最后一天四

个模块的平均错误率) (图2c)。由于未设置“Backspace”功能,即便系统出现解码错误的情况,我们不会中止T5书写。

当使用离线语言模型自动校正错误时,其错误率明显下降(图2c,表1):字符错误率下降到0.89%,单词错误率下降到3.4%。这与最先进的语音识别系统(单词错误率为4~5%)相当^[14,15]。因此,该系统具有较好的实用性。进一步地,为探究解码性能的极限,我们采用所有可用数据训练新的RNN,并以非因果方式处理整个语句(与其他的BCI研究^[16,17]类似)。在这种情况下,该系统准确率极高(0.17%的字符错误率)。这一结果表明,尽管该解码器不能向用户提供逐字反馈,但其性能的潜在上限很高。

为在较少约束下评估解码器性能,我们让T5在两天内使用BCI自由地输入开放式问题的答案,并收集这两天的数据(补充视频3,扩展数据表2)。结果表明,与复现屏幕提示语句相比,当T5自由书写相关语句时,解码器仍具有较高性能(输出速度每分钟73.8个字符,实时字符错误率为8.54%,校正后字符错误率为2.25%)。据我们所知,植入式BCI自由打字记录是每分钟24.4个正确字符。

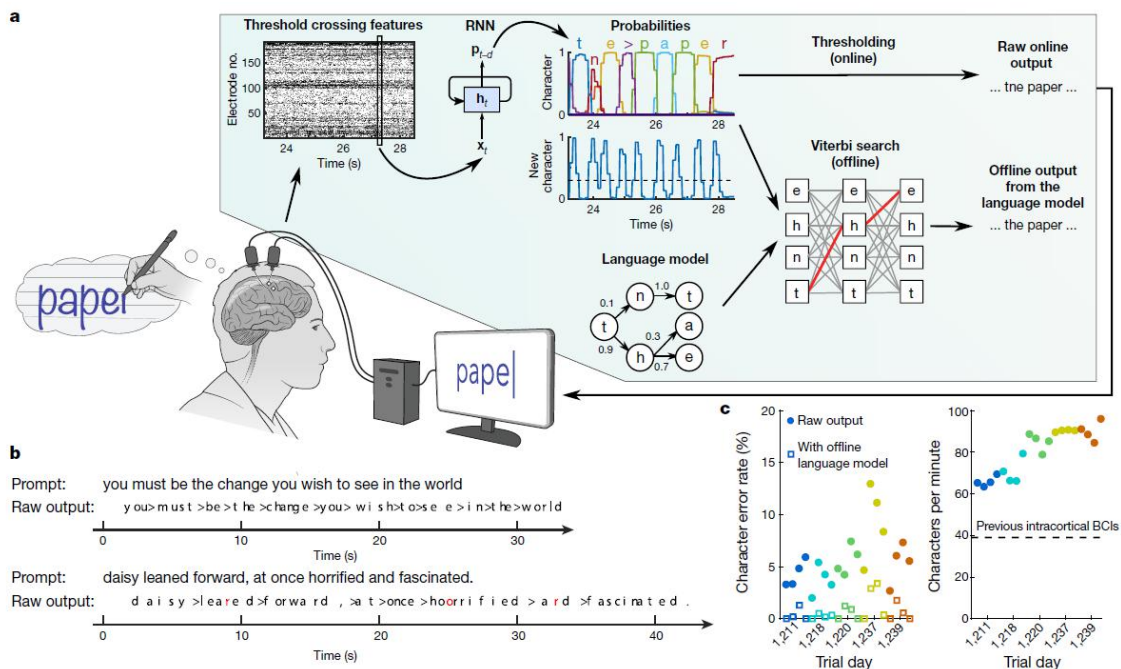


Fig. 2. 尝试手写的实时神经解码。a.解码算法示意图。首先,将神经活动(多单元阈值跨越率)在每个电极(20毫秒bins)上暂时分块并平滑处理。然后,RNN将该神经簇时间序列(X_t)转换为概率时间序列(p_{t-d}),描述每个字符出现的可能性和任何新字符开始的概率。RNN有一个1秒的输出延迟(d),让它有时间在确定每个字符的身份之前充分观察每个字符。最后,对字符概率进行阈值化,以产生可实时使用的“原始在线输出”(当“新字符”概率在时间 t 超过阈值时,在时间 $t + 0.3s$ 生成最可能的字符并在屏幕上显示)。在一项离线回顾性分析中,将字符概率与大词汇量语言模型相结合,以解码参与者最有可能手写的文本(使用自定义的50000字二元模型)。来源: E.Woodrum创作的人体轮廓图。b.显示了两个实时训练的例子,展示了RNN在从未接受过训练的语句上具有解码易于理解文本的能力。错误用红色突出标出,空格由‘>’符号表示。c.显示5天,4个模块,每个模块7-10个语句的错误率和打字速度(每个模块用一个圆圈表示,并根据试验日上色)。速度是次快植入式BCI速度的两倍以上,如虚线所示。

3 解码器日常再训练

根据标准惯例^[1,2,4,5,18],借助每天开始时收集到的“校准”数据,在每天评估之前再训练手写解码器。再训练有助于解释神经记录随时间产生的变化,该变化可能是由神经可塑性或电极阵列的微运动引起的。理想情况下,为了减少使用者的负担,只需极少或不需要校准数据。在针对图2b数据的回顾性分析中,我们对每天使用少于50个校准语句是否能够实现高性能进行评估(图3a)。评估发现,需要30个校正语句将在线错误率降低到5.9%。但是10个语句(8.5分钟)足够将错误率降低到8.5%(使用语言模型时为1.7%)。

Table 1

5天内手写BCI的平均字符和单词错误率（具有95%的置信区间）。“原始在线输出”是在线（实时）解码的内容。“在线输出+离线语言模型”是通过将语言模型回顾性地应用于在线解码的内容而获得的（以模拟自动校正功能）。除了应用语言模型外，还通过使用所有可用数据离线再训练双向（非因果）解码器获得“离线双向RNN+语言模型”。单词错误率可能远高于字符错误率，因为如果一个单词中的任何字符错误，则该单词被认为是错误的。

	字符错误率	单词错误率
原始在线输出	5.9%[5.3,6.5]	25.1%[22.5,27.4]
在线输出+离线语言模型	0.89%[0.61,1.2]	3.4%[2.5,4.4]
离线双向RNN+语言模型	0.17%[0,0.36]	1.5%[0,3.2]

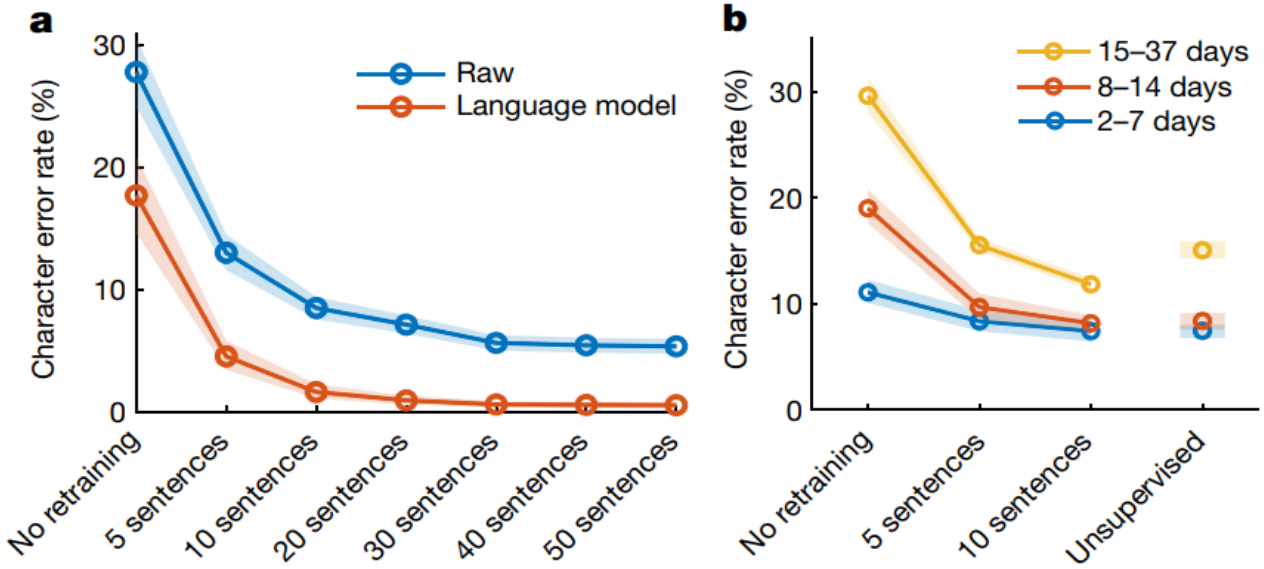


Fig. 3. 当每日解码器再训练时间缩短（或无监督）时，性能仍然很高。a, 为了说明随着时间推移神经活动的变化，我们每天在评估之前再训练手写解码器。在这里，我们离线模拟了如果使用少于50个校准语句，解码性能将如何变化。折线显示所有数据的平均错误率，阴影区域表示95%的置信区间（CIs）。b, 八次区间的复制打字数据用于评估如果区间发生的时间更近，是否会减少对校准数据的需要。所有区间（X, Y）都被考虑。解码器首先使用来自区间X和更早的训练数据进行初始化，然后在不同的再训练方法下对区间Y进行评估（无再训练，在有限的校准数据上再训练，或无监督再训练）。折线表示平均错误率，阴影区域表示95%的置信区间（CIs）。

然而，我们的数据是跨越28天收集的，这些数据可能包含了神经活动更大的变化。我们采用离线分析评估了间隔长短的区间能否减少对校准数据的需求（图3b）。我们发现，当两次区间仅间隔2-7天时，没有再训练的情况下（11.1%的错误率，使用语言模型时为1.5%），解码器的性能是可以接受的。这与之前神经记录短期稳定性的研究工作一致^[19–21]。最后，我们通过使用语言模型对解码器进行错误纠正和再训练，并测试解码器能否以无监督的方式进行再训练，从而避免因校正而中断用户使用（通过在正常使用期间启用自动重新校准）。令人振奋的是，当区间为7天或更短时，无监督再训练将错误率降低到7.3%（使用语言模型时为0.84%）。

最终，能否用最少的重新校准数据成功地再训练解码器取决于神经活动随时间变化的快慢。我们评估了与每个字符相关的神经模式的稳定性，发现了很高的短期稳定性（间隔7天或更短时的平均相关系数为0.85）和神经活动的变化。这些神经活动的变化似乎以稳定和可预测的速度积累（扩展数据图4）。这些结果表明，临床实施是有希望的；在无监督情况下再训练解码器，同时结合有限的监督情况下的再训练，可使长时间未工作的解码器保持较高性能。尽管如此，将来必须在线实施对解码的再训练，因为离线模拟的性能与在线实施的性能并不总能保持一致。

4 时间多样性改进解码

据调研，目前报道的BCI系统类型中最高打字速度是每分钟90个字符（见“讨论”）。对于植入式BCI系统，表现最好的方法是二维计算机光标点击打字，每分钟最多能正确输入40个字符（直接比较见补充视频4）^[7]。点击式BCI系统的速度主要受到解码精度的限制。在参数优化过程中，系统通过增加光标增益提高打字速度，直到

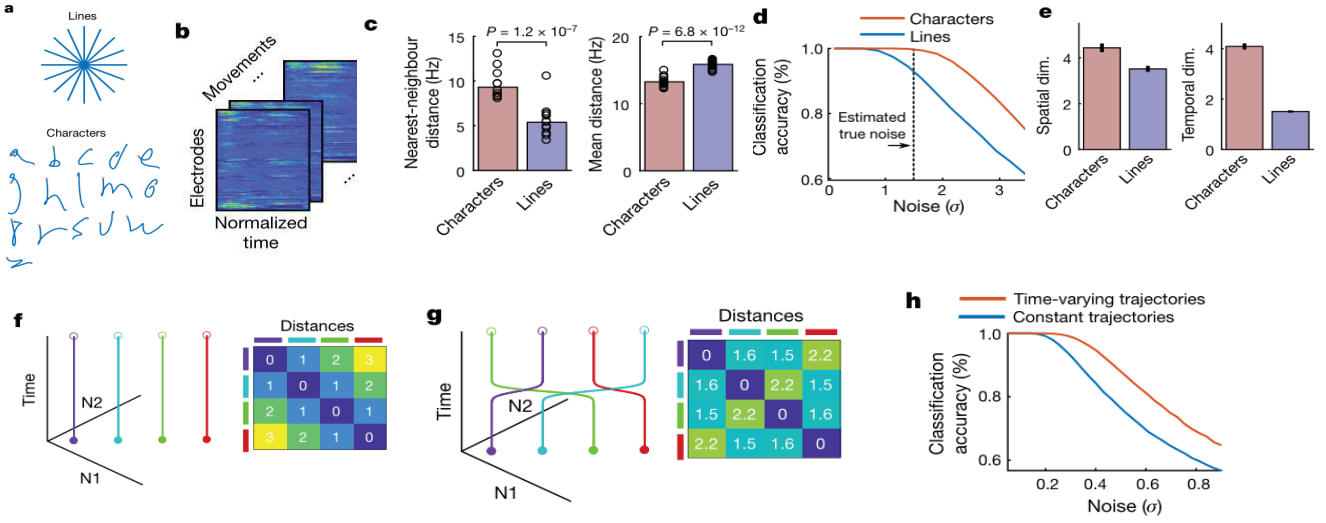


Fig. 4. 增加时间多样性可以使动作更容易解码。a.我们分析了16个手写字符（持续时间为1s）与16个手写直线动作（持续时间为0.6s）对应的神经活动的时空模式。b.通过对给定动作的所有试验进行平均处理来发现神经活动的时空模式（经过时间扭曲以使试验时间对齐）。对神经活动进行重采样，使每组动作的持续时间相等，得出每个动作的矩阵为 192×100 （192个电极和100个时间步长）。c.计算每组神经活动样本之间的欧氏距离，显示出字符较大的最近邻距离（但不是平均距离）。每个圆代表一个单独的动作，条形高度表示平均值。d.较大的最近邻距离使字符比直线更容易分类。噪声以标准差为单位，并与图c中的距离比例相匹配。e.字符和直线的空间维度相似，但字符的时间维度是直线的两倍多，这表明较多的时间维度是使最近邻距离增加和分类性能更好的基础。错误栏显示95%的置信区间。维度通过参与率进行量化。f-h.这是一个toy example，直观地说明了时间维度的增加如何使神经轨迹更易区分。描述了三条神经轨迹（N1和N2是两个假设的神经元，其活动受限于单一空间维度，即单位对角线）。通过添加一个弯曲，使神经轨迹在时间维度上变化，从而将时间维度从1（f）增加到2（g），从而能增加最近邻距离以及更好的分类（h）。

光标移动速度过快，导致解码错误而失控^[22]。因此，我们提出这样一个问题：如何能以两倍以上的时间和相当的准确率解码手写字母。

我们的想法是，手写字母运动可能比点对点的运动更容易区分，因为尝试手写字母的神经活动的时空模式比直线运动的更具有多样性。为了验证这一想法，我们分析了16个直线运动和16个字母运动（不需要将笔从页面上抬离）的神经活动样本，这两个动作均由T5尝试手写完成（图4a, b）。

首先，我们分析了每个神经活动样本之间的欧氏距离。我们发现，手写字母运动的最近邻距离要比直线运动大72%（95% CI = [60%, 86%]），这使得解码器混淆两个相邻字符的概率较低（图4c）。为证实这一点，我们将每组动作的分类精度模拟为神经噪声函数（图4d），这表明字符比直线更容易分类。

为了深入了解字符的最近邻距离相对增加的原因，我们检验了神经样本的空间和时间维度。使用主成分分析（PCA）特征值谱的“参与率”来估计空间和时间维度，该特征值谱大概量化了需要多少空间或时间维度来解释神经活动模式中80%的方差^[23]。研究发现，字符的空间维度略大（1.24倍；95%CI=[1.19,1.30]），但时间维度则大得多（2.65倍；95%CI=[2.58, 2.72]），这表明字母书写活动中时间模式的多样性增加，提高了每个动作的可区分性（图4e）。

为了说明增加的时间维度如何使运动更容易区分，我们构建了一个包含四个运动轨迹和两个神经元的toy model，将神经活动限制在单个维度上（图4f,g）。只需让神经轨迹随时间变化（图4g），就可以增加神经轨迹之间的最近邻距离，从而在噪声水平足够大时提高分类精度（图4h）。虽然toy model的神经噪声被假定为独立的白噪声，但我们发现这些结果也适用于与时间和神经元相关的噪声（扩展数据图5，补充说明1）。

这些结果表明，运动的时变模式（比如手写字母）从根本上来说比点对点的运动更容易解码。我们认为这是使手写BCI系统比连续动作的点击式BCI系统更快的一个重要因素，但不一定是唯一的。其他离散的（基于分类的）BCI系统通常使用时间变化较小的定向运动，这可能会限制其准确性和（或）运动集的大小^[24,25]。

一般来说，利用手写动作之间最近邻距离最大化的原则，应该可以优化一组动作以便于分类^[26]。我们研究了这种可能性，并设计了一种理论上比拉丁字母更容易分类的字母表（扩展数据图6）。优化后的字母表避免了大量书写相似的冗余字母（大多数拉丁字母的开头不是向下就是逆时针卷曲）。

5 讨论

闭锁综合征（几乎所有主动肌麻痹）严重损害或阻碍交流，最常见的原因是脑中风或晚期肌萎缩侧索硬化（闭锁综合征的估计患病率：100000人中有1例^[27]）。恢复交流常用的BCIs不是闪烁EEG拼写^[18,28–32]，就是植入点击式（point-and-click）BCIs^[6,7,33]。基于P300电位（oddball potentials）或运动想象的脑电拼写器通常每分钟能达到1–5个字符^[28–32]。使用视觉诱发电位的EEG拼写器已达到每分钟60个字符的速度^[18]，但在使用上有明显的局限性，主要表现在使用过程需要遮住拼写者眼睛，通常不能自定步速，而且需要在屏幕上安装闪光灯。基于二维光标移动的植入BCIs，拼写者可自由地环顾四周，并可以设定自己的交流速度，但每分钟还没超过40个正确字符^[7]。最近，语音解码BCI在恢复快速交流方面显示出令人振奋的前景，但其准确度和词汇量需要进一步提高以支持正常交流^[16,17,34]。

在本文中，我们介绍了一种新的交流BCI系统——解码四肢瘫痪患者快速、灵巧的动作——为交流速率设定了一个基准，即每分钟90个字符。该实时系统具有通用性（用户可以表达任何句子）、易用性（完全自定步速，眼睛可以自由移动）和准确性（94.1%的原始准确率，以及在具有大词汇表语言模型中大于99%的离线准确率），可应用于实际。为获得更高性能，我们开发了数据有限情况下有效处理未标记神经序列的解码方法。这种方法可以更普遍地应用于无法直接观察到的任何连续行为（例如，解码不能说话人的语音）的情况。

值得注意的是，该系统证明了高性能手写BCI的可行性（在单个受试者中）；但它还不是完整的、可应用于临床的系统。因此，后续需要更多的工作来证明在其他人中也能实现高性能手写，包括扩展字符集（例如，大写字母），启用文本编辑和删除，并在不中断用户再训练的情况下保持对神经活动变化的鲁棒性。推而广之，植入微电极阵列技术仍在趋向成熟阶段，其寿命、安全性和有效性需进一步证明后才能广泛应用于临床^[35,36]。拼写者表现的差异性也是一个潜在的问题（在之前的一项研究中，T5在三名受试者中取得了最好的表现^[7]）。

尽管如此，我们仍相信植入式BCIs的前景是光明的。目前，微电极阵列技术（包括此处；参见扩展数据图11）可保证植入后具有1000天以上的效能^[37,38]，与其他已有技术（例如EEG或皮层电图）相比，其在恢复交流^[7]、手臂控制^[2,5]和通用计算机使用^[39]等方面实现了最高BCI性能。可植入设备设计方面已有新进展，这些设计将电极数量增加一个数量级，将进一步提高可植入设备的性能和寿命^[35,36,40,41]。最后，我们计划结合算法创新^[42–44]和设备稳定性改进，将减少解码器日常训练的需求。在本文中，**离线分析结果显示了有限监督解码器设计、在线无监督解码器训练具有潜在前景（图3）**。

6 在线内容

任何方法、附加参考、自然研究报告总结、源数据、扩展数据、补充信息、致谢、同行评审信息；作者贡献和竞争利益的详细信息；有关数据和代码可用性的声明，请访问<https://doi.org/10.1038/s41586-021-03506-2>。

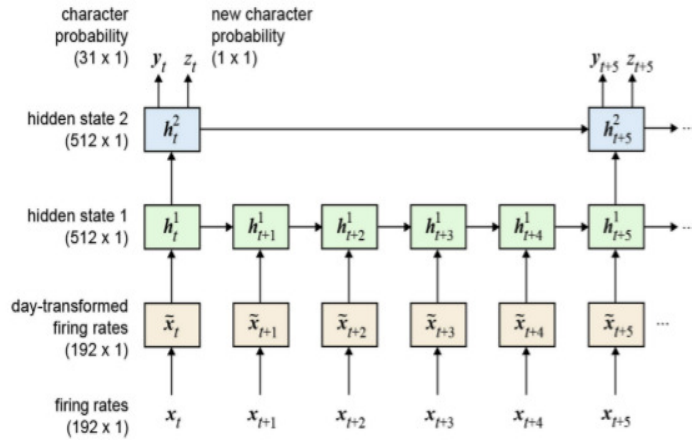


Fig. 5. RNN结构图。我们使用两层门控递归单元（GRU）结构的递归神经网络将神经放电频率向量 X_t 序列（在20ms时进行时间平滑和装箱）转换为字符概率向量 y_t 和新字符概率标量 z_t 序列。 y_t 向量描述了每个字符在该时刻被写入的概率，并且每当RNN检测到T5开始写入任何新字符时， z_t 标量变高。请注意，我们发现RNN上一层的运行频率低于下一层更容易在内存中长时间保存信息，从而提高了训练速度。因此，RNN的输出每100毫秒更新一次。此外，请注意，我们使用了一种特定于一天的仿射变换来解释神经活动的日常变化（底部一行）——这有助于RNN解释在多天合并训练数据时，由电极阵列微动或大脑可塑性引起的神经调节变化。

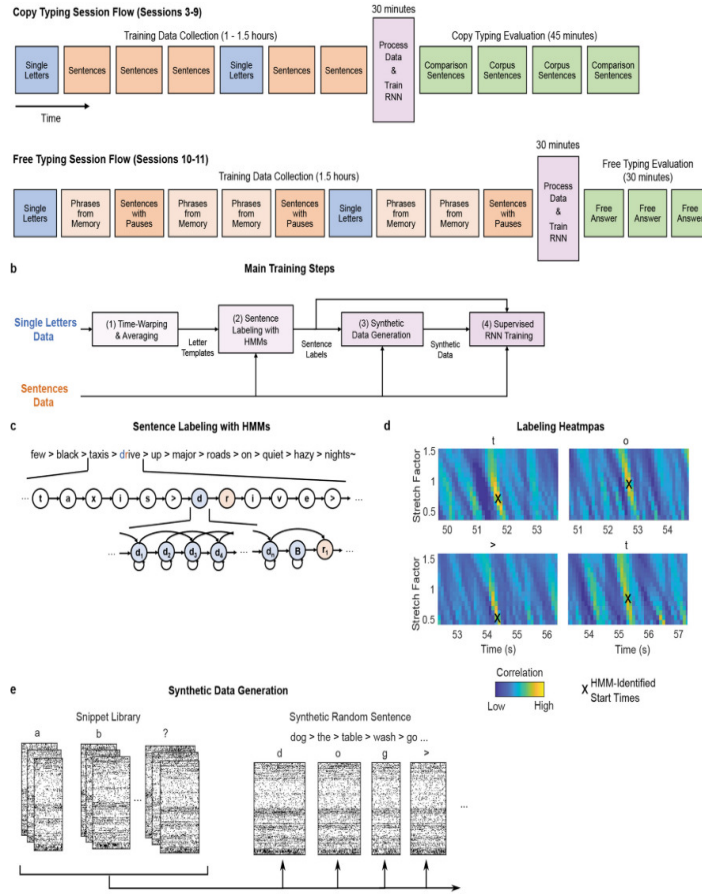


Fig. 6. RNN训练方法概述。a，复制键入和自由键入会话的会话流程图（每个矩形对应一个数据块）。首先，收集单个字母和语句的训练数据（蓝色和红色块）。接下来，使用新收集的数据加上前几天的所有数据（紫色块）对RNN进行训练。最后，RNN保持固定并进行评估（绿色块）。b，数据处理和RNN训练流程图（a中的紫色块）。首先，对单个字母数据进行时间扭曲和平均处理，为每个字符创建神经活动的时空模板。这些模板用作语句标签去初始化隐马尔可夫模型（HMM）。标记后，将观察到的数据分割并重新排列成新的字符序列，以生成合成语句。最后，将合成句与真实句相结合训练RNN。c，用于标记“少数黑色出租车在安静朦胧的夜晚行驶在主要道路上”的强制对齐HMM图。HMM状态对应于上述语句的字符序列。d，标签质量可通过将单字符神经模板与真实数据关联成的互相关热图进行验证。HMM识别的字符开始时间在热图上形成清晰的热点。请注意，这些热图仅用于定性显示标签质量，不用于训练（仅需要字符开始时间来生成RNN训练的目标）。e，为了生成新的合成语句，将真实数据中每个标记字符对应的神经数据从数据流中剪切出来并放入片段库。然后从库中随机抽取这些片段，在时间上拉伸或压缩高达30%（以增加更多人为的时间变化），并组合成新的语句。

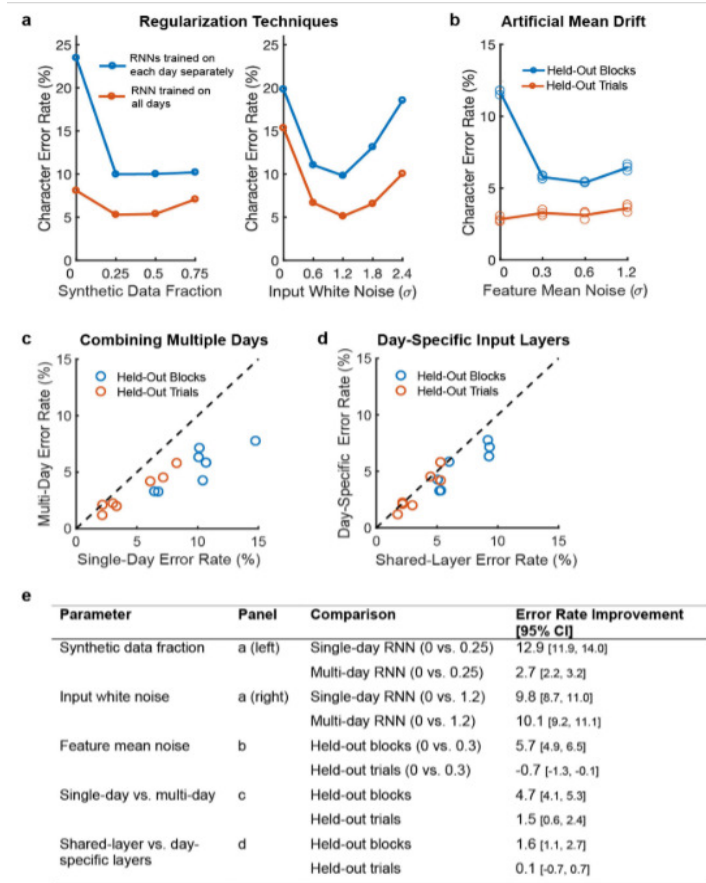


Fig. 7. 关键RNN参数对性能的影响。a，使用合成数据（左）和人工白噪声输入（右）进行训练对于高性能都是必不可少的。数据显示在两个参数上的网格搜索中，行显示另一个参数最佳值的性能。结果表明，这两个参数都是高性能所必需的，即使另一个参数处于最佳值。在数据集的大小非常有限的情况下，使用合成数据更为重要，就像仅在一天的数据上进行训练（蓝线）。注意，给RNN的输入是z分数，因此输入白噪声以输入特征的标准偏差为单位。b，添加到特征段中的人工噪声（随机偏移和基线发射率的缓慢变化）极大地提高了RNN的能力，使其能够泛化到以后出现的新数据块中。但是，这并不能帮助RNN在已经训练过的数据块中推广新的试验，这是因为特征意义随着时间的推移缓慢变化。对于每个参数设置，训练三个单独的RNN（圆圈）；结果表明，RNN训练运行的可变性较低。c，与每天单独进行的训练相比，将所有数据集结合起来对RNN进行训练可以提高性能。每个圆圈显示七天中的一天的性能。d，每天使用单独的输入层比每天使用同一个输入层要好。e、总结了每个参数在字符错误率方面的改进。通过单次试验（ $n=10000$ ）的bootstrap重采样计算95%的CIs。如表所示，所有参数至少在一种情况下（CI不与零相交）显示出统计上的显著改善。

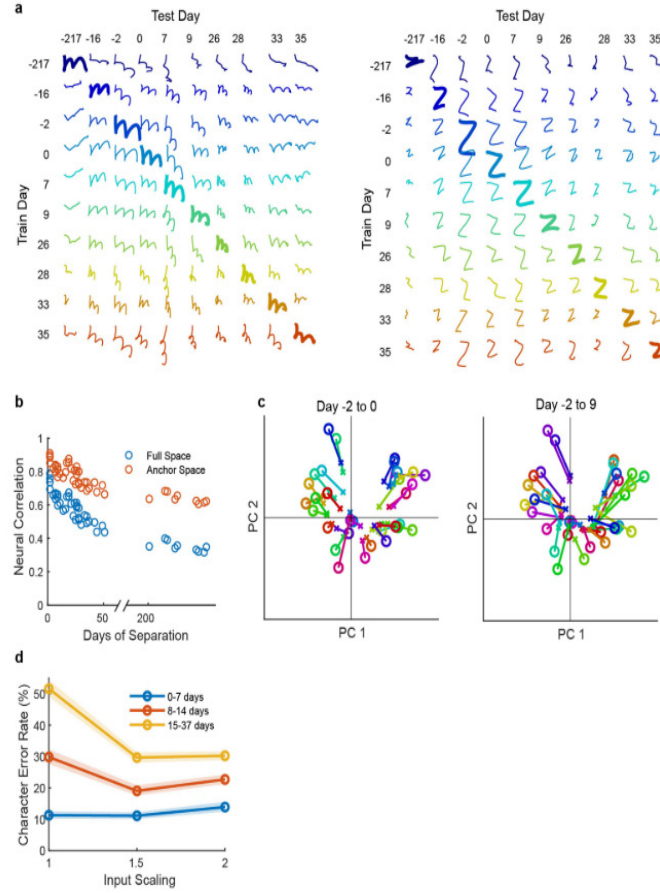


Fig. 8. 数天内神经记录的变化。a, 为了可视化神经记录随时间变化的程度, 使用在所有其他日期 (行) 训练的解码器, 为所有10天的数据 (列) 绘制了两个示例字母 (m和z) 的解码笔尖轨迹。根据2019年12月9日 (第4天) 经过的天数对每一次会话进行标记。结果表明, 尽管神经活动的模式随着时间的推移而明显改变, 但它们的基本结构在很大程度上是保守的 (因为过去训练的解码器很容易转移到未来)。b, 计算每对会话的神经活动模式之间的相关性 (Pearson's r), 并绘制为每对会话间隔天数的函数。蓝色圆圈显示在整个神经空间 (所有192个电极) 中计算的相关性, 而红色圆圈显示在“锚定”空间中的相关性 (前一阶段的前10个主要成分)。高值表示字符在几天内的神经编码方式具有很高的相似性。锚定空间中相关性较高的事实表明, 神经模式的结构在缓慢转移到新空间时基本保持不变, 导致原始空间收缩, 但结构变化不大。c, 通过原始“锚定”空间顶部的两台电脑查看每个字符的神经表征如何随时间变化的可视化。每个圆圈代表单个字符的神经活动模式, 每个x符号在以后的一天显示相同的字符 (直线连接匹配的字符)。左边是一对只有两天的会话 (2到0天); 右边是11天之间的一对会话 (2到9天)。神经模式的相对位置在几天内保持相似, 但大多数情况下明显向原点收缩。这与神经表征从原始空间缓慢转移到新空间是一致的, 并表明放大输入特征可能有助于解码器更准确地传输到未来会话 (通过抵消这种收缩效应)。d, 类似于图3b, 在未来会话上评估解码器时, 来自八个会话的复制数据被用于离线评估 (当不使用解码器重新训练时) 放大解码器输入是否可以提高性能。考虑了所有会话对 (X, Y)。首先使用会话X和更早的所有数据初始化解码器, 然后在会话Y上以不同的缩放因子进行评估 (例如, 输入比例为1.5意味着输入特征按50%的比例放大)。线条表示平均原始字符错误率, 阴影区域显示95% CIs。结果表明, 当会话之间经过长时间时, 输入缩放可以提高性能。因此, 在图3的“无再训练”条件下评估解码器性能时, 我们使用了1.5的缩放因子。

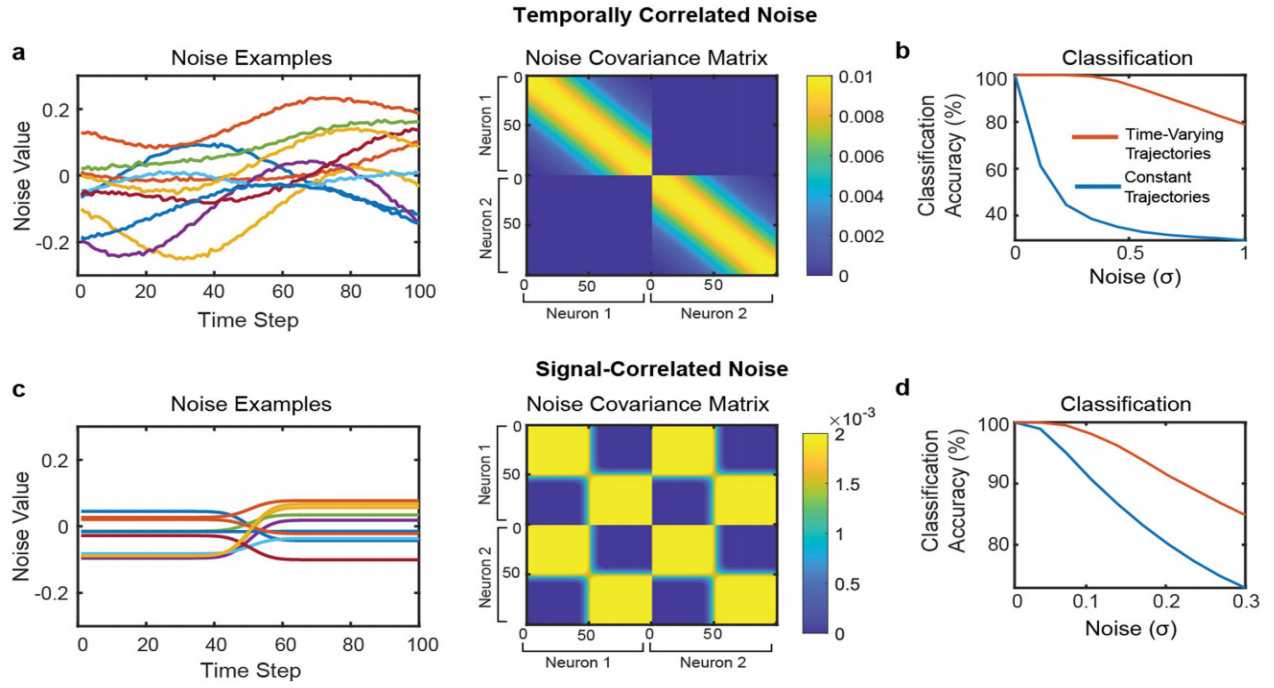


Fig. 9. 扩展数据图5|相关噪声对时间维度玩具模型 (toy model) 的影响a, 时间相关噪声的噪声向量和协方差矩阵示例。在左边, 噪声向量示例被绘制出来 (每条线描述了一个示例)。给出了神经元1的所有100个时间步长的噪声向量。在右边, 用于生成时间相关噪声的协方差矩阵被绘制出来 (维度= 200×200)。前100个时间步长描述神经元1的噪声, 后100个时间步描述神经元2的噪声。对角线带在每个模拟神经元内产生时间相关的噪声 (但这两个神经元彼此不相关)。b, 在存在时间相关噪声的情况下, 使用最大似然分类器对所有四种可能的轨迹进行分类时的分类精度。即使存在时间相关的噪声, 时变轨迹仍然更容易分类。c, 与信号相关的噪声向量和噪声协方差矩阵的例子 (即噪声只集中在跨越类均值的时空维度上)。与时间相关噪声不同, 该协方差矩阵生成时间步长与神经元之间具有相关性的时空噪声, d, 信号相关噪声下的分类精度。时变轨迹比常数轨迹更容易分类。图片的详细解释见补充说明1。

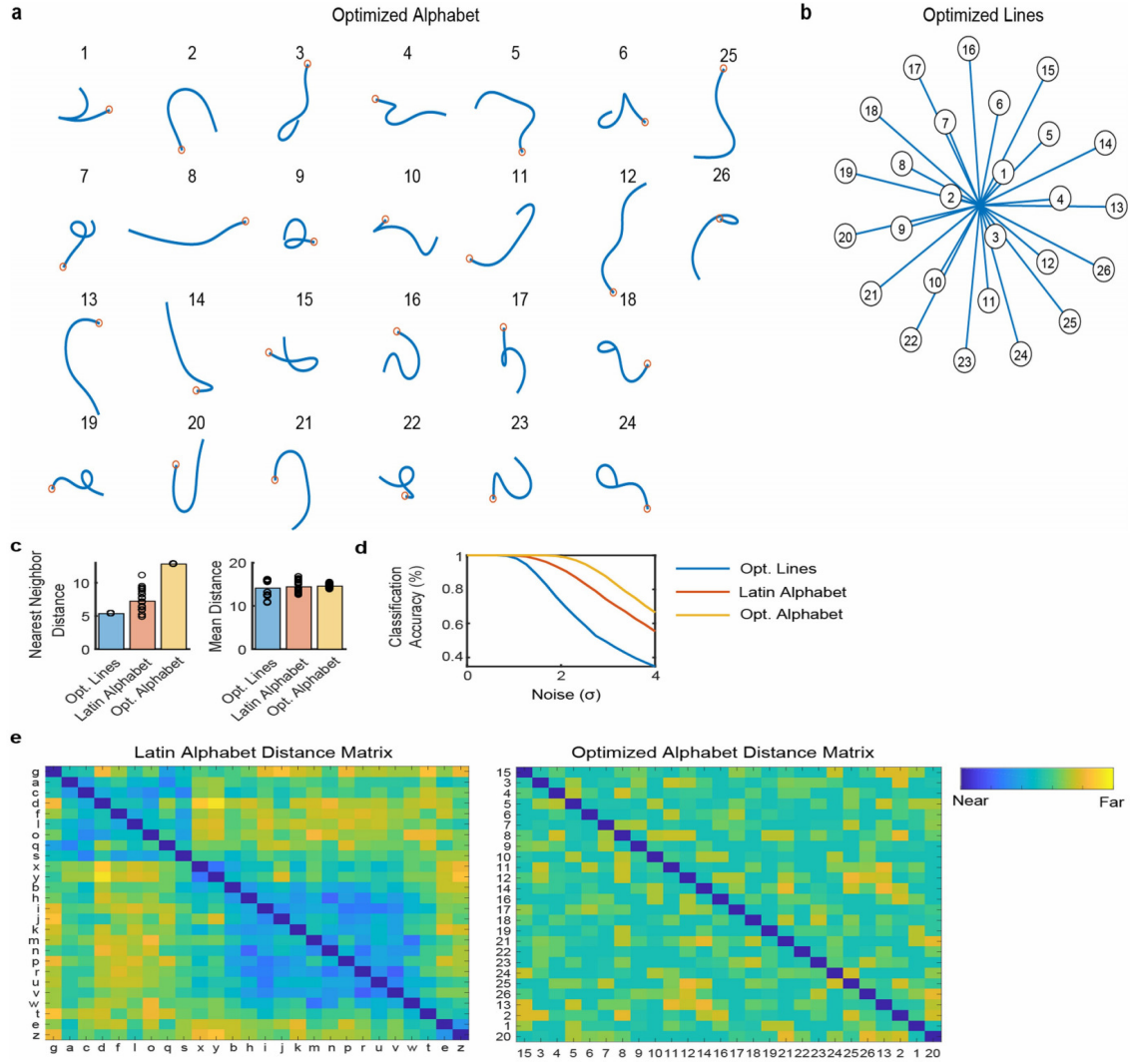


Fig. 10. 扩展数据图6|最大限度地提高神经解码能力的人工字母优化。a，基于最大化最近邻距离的原则，我们优化了一组笔迹，理论上它比拉丁字母更容易分类（使用线性神经调节笔尖速度的标准假设）。b，为了比较，我们还优化了一组直线（26条），使最近邻距离最大化。c，计算每组笔尖轨迹之间的成对欧几里得距离，揭示了优化后的字母比拉丁字母显示出更大的最近邻居距离（不是平均距离）。每个圆圈代表一次移动，柱状高度代表平均值。d，模拟分类精度作为人工噪声增加量的函数。结果证实，优化后的字母确实比拉丁字母更容易分类，而且拉丁字母比直线更容易分类，即使在直线经过优化后。e，拉丁字母和优化字母的距离矩阵显示笔迹之间的成对欧几里得距离。采用单链层次聚类方法将距离矩阵分为7个聚类。优化后的字母表的距离矩阵没有明显的结构；相比之下，拉丁字母表有两大组相似的字母（以逆时针旋度开头的字母和以下描线开头的字母）。

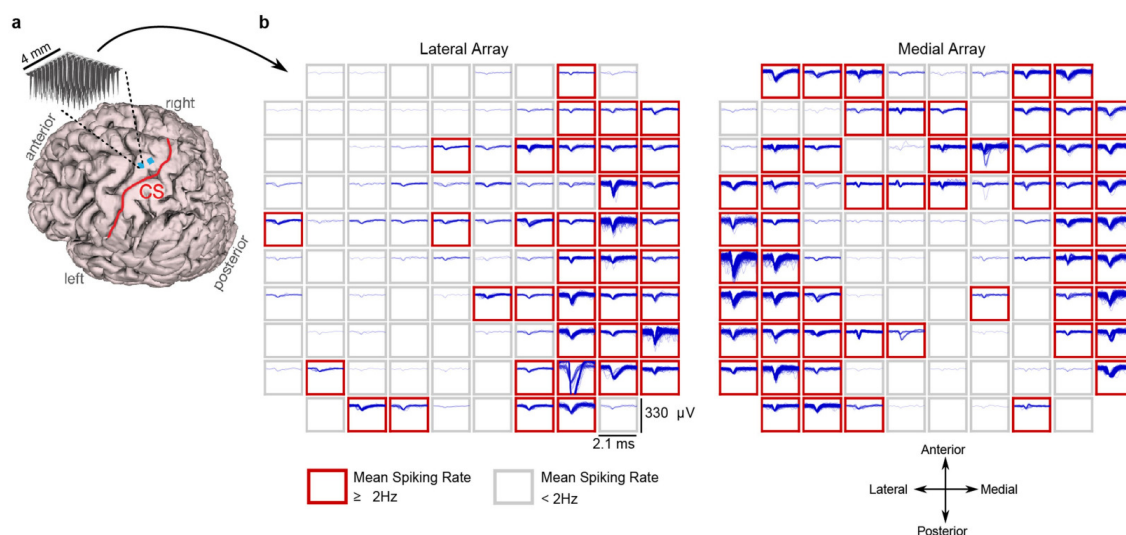


Fig. 11. 扩展数据图7|每个微电极阵列记录的峰值活动示例。a, 受试者T5大脑解剖的磁共振成像（MRI: Magnetic resonance imaging）。通过术后CT图像与术前MRI图像共同配准确定微电极阵列位置（蓝色方块）。b, 绘制每个电极在10秒时间窗内检测到尖峰波形的示例（在植入后1218天内数据被记录）。每个矩形面板相当于一个单一的电极，且每个蓝色轨迹是一个单一的尖峰波形（2.1-ms持续时间）。峰值事件使用4.5均方根（RMS）阈值检测，因此排除了几乎所有的背景活动。平均阈值越过率至少2Hz的电极被认为具有峰值活动，并且用红色表示（注意，这是一个保守估计，意味着只包括可能来自单个神经元的峰值活动，而不是多单元 `textcolor[rgb]1.00,0.00,0.00 'hash'` ）。结果表明，许多电极仍能记录远高于噪声底限的尖峰波形（每个面板的y轴跨度为330 μV ，而背景活动的平均均方根值仅为6.4 μV ）。在这一天，192个电极中有92个的阈值越过率至少为2Hz。

Prompt	Real-Time Output	Real-Time Output (Annotated)	Offline Language Model (Annotated)
infected adults develop a cough and their skin and ears turn blue.	infected adults dercep a cough and thhir skin and ears tunn blue	infected adults dercep a cough and thhir skin and ears tunn blue.	infected adults develop a cough and their skin and ears turn blue.
i interrupted, unable to keep silent.	i interrupted, unabee to keep silent.	i interrupted, unabee to keep silent.	i interrupted, unable to keep silent.
i dumped the tools in the hut.	i dumped the tools in the hut.	i dumped the tools in the hut.	i dumped the tools in the hut.
within thirty seconds the army had landed.	within thirty seconds the ammy had landed	within thirty seconds the ammy had landed.	within thirty seconds the army had landed.
that's when i threw up on the carpet.	that's when i threw up on the carpet.	that's when i threw up on the carpet.	that's when i threw up on the carpet.
he didn't want to rub salt into her wounds.	he didn't want to rub salt into her wounds.	he didn't want to rub salt into her wounds.	he didn't want to rub salt into her wounds.
shouting and swearing, i yelled for an epidural.	shouting and sweerin, i yeled for an epidural.	shouting and sweering, i yeled for an epidural.	shouting and swearing, i yelled for a an epidermal .
you wish to purchase something?	you wish to purchase something?	you wish to purchase somathing?	you wish to purchase something?
lowell felt like a soldier on a battlefield, stripped of ammunition.	lowel felt like a soldier on a battlefield, stripped ef ammunition.	lowell felt like a soldier on a battlefield, stripped ef ammunition.	lowell felt like a soldier on a battlefield, stripped of ammunition.
there are only one or two minor casualties.	thee ane only on or two minor cafualties.	there ane only one or two minor cafualties.	there are only one or two minor casualties.

Fig. 12. 扩展数据表1|一段通过拷贝打字而解码语句的例子。在最右边的列中，用红色突出显示错误（用红色方框表示额外的空格，用划线表示遗漏的字母）。需要注意到，由于在词汇表中没有“硬脑膜外的”（‘epidural’），我们在语言模型用“表皮的”（‘epidermal’）代替了“硬脑膜外的”（‘epidural’）。该区块的平均打字速度为每分钟86.47个字符，字符错误率为4.18%（实时输出）和1.22%（语言模型）。语句提示是从英国国家语料库中随机抽取的（详细信息请参见补充方法）。

Prompt	Real-Time Output	Real-Time Output (Annotated)	Offline Language Model (Annotated)
what made you first get interested in machining?	it was the famil business	it was the family business	it was the family business
what is the hardest part of machining?	forming the tooling	forming the tooling	forming the tooling
how much spice do you like in your food?	lots and lots	lots and lots	lots and lots
what type of music do you enjoy most?	ii like loud rock musica	ii like loud rock musica	i like loud rock music
what are some of your favorite games to play?	i like to play cards	i like to play cards	i like to play cards
what has taken you the longest to get good or decent at?	i worked for years to perfecet my photography	i worked for years to perfecet my photography	i worked for years to perfect my photography
what food do you love that a lot of people might find a little odd?	very few poope can eat smekek sailfish	very few poople can eat smekek sailfish	very few people can eat smoked salt fish
if you could start a charity, what would it be for?	i woud proovide furdng to smal families	i woud proovide furdng to small families	i would provide funding to small families
what advice would you give to your younger self?	be patientt it wil got better	be patientt it will got better	be patient, it will get better

Fig. 13. 扩展数据表2]一段通过自由打字而解码语句的例子。在最右边的列中，用红色突出显示错误（用划线表示遗漏的字母）。需要注意到，由于在词汇表中没有旗鱼（‘sailfish’），我们在语言模型用咸鱼（‘salt fish’）代替了旗鱼（‘sailfish’）。该区块的平均打字速度为每分钟73.8 个字符，字符错误率为6.82%（实时输出）和1.14%（语言模型）。

7 报告总结

《自然研究》希望提高我们出版的作品的可再现性。这个表格提供了报告的一致性和透明度的结构。有关《自然研究》政策的进一步信息，请参阅我们的编辑政策和编辑政策检查表。

7.1 统计数据

对于所有的统计分析，确认图图例、表图例、主文本或方法部分中存在以下项。

n/a	存在
✓	每个实验组/条件的确切样本量（n），以离散数字和计量单位给出。
✓	是否从不同的样本中进行测量，或是否重复测量相同的样本。
✓	所使用的统计检验以及它们是单边的还是双向的。
✓	只有常见的测试才应该仅通过名称来描述;在“方法”一节中描述更复杂的技术。
✓	所有被测试协变量的描述。
✓	对任何假设或修正的描述，如常态检验和多次比较的调整。
✓	统计参数的完整描述，包括集中趋势（例如平均值）、其他基本估计（例如回归系数）、变异（例如标准差）或不确定性的相关估计（例如置信区间）。
✓	对于无效假设检验，检验统计量（如F, t, r）具有置信区间、效应大小、自由度和P值。
✓	在适当的时候给出P值作为精确值。
✓	对于贝叶斯分析，信息的选择先验和马尔可夫链蒙特卡罗设置。
✓	对于分层和复杂的设计，确定适当的测试级别，并全面报告结果。
✓	效应大小的估计（e.g.Cohen’s d, Pearson’s r），表明它们是如何计算的。

7.2 软件和代码

关于计算机代码可用性的政策信息。

数据收集	使用MATLAB和Simulink real（MathWorks, Natick, MA）定制开发了运行实验任务、记录数据和实现实时解码系统的软件。
数据分析	数据分析使用自定义MATLAB和python代码。自定义代码在接受后将在GitHub上公开。

这些自定义算法或软件是研究的核心，但其手稿尚未在已发表的文献中描述，软件必须提供编辑和审稿人。我们强烈鼓励代码存储在社区存储库中（例如github）。关于提交代码和软件的进一步信息，请参阅Nature Resrarch指南。

7.3 数据

关于数据可用性的策略信息。
所有手稿必须包括一个数据可用性声明。在适用的情况下，该声明应提供以下信息:

- 接入码、唯一标识符或公开数据集的网络链接。
- 与原始数据相关的数据列表。
- 对数据可用性的任何限制的描述。

接受后，数据将在Dryad上公开。相关链接和标识符将包括在最终手稿中。

8 领域的报道

请选择一个最适合您的研究下面。如果你不确定，在做出选择之前，先阅读相应的部分。有关包含所有章节的文档参考副本，
✓ 生命科学 行为与社会科学 生态、进化和环境科学
请参见nature.com/documents/nr-reporting-summary-flat.odf

9 生命科学研究设计

所有的研究必须公开这些点，即使公开是负面的。

样本大小	没有进行样本量计算。研究人员收集了单个参与者的数据，以描述脑机接口的性能。 用置信区间对性能估计中的不确定性进行了量化，并给出了稳定的结果。
数据筛选	这项研究是基于连续几天收集的脑-机接口性能评估数据。研究报告了所有的天数，并包括了所有相关数据。
重复实验	这项研究评估了单个参与者的脑-机接口性能。结果由重复了多天的绩效评估得出。
随机化实验	随机化实验与本研究无关——本研究只纳入了一名参与者。
混合实验	混合实验与这项研究无关，该研究评估了单个个体的脑-机接口性能

10 汇报具体的材料、系统和方法

我们需要从作者的一些类型的材料，实验系统和方法在许多研究中使用的信息。在这里，注明所列的材料、系统或方法是否与你的研究相关。如果你不确定所列的项目是否适用于你的研究，在选择回答之前，先阅读相应的部分。

材料与实验系统		
n/a	参与了这项研究	
×		抗体
×		真核细胞系
×		古生物学和考古学
×		动物和其他生物
	×	人类研究参与者
×		临床数据
×		关注的双重用途研究
方法		
n/a	参与了这项研究	
×		ChIP-seq
×		流式细胞仪
×		MIR-based神经成像

11 人类研究参与者

涉及人类研究参与者的研究的政策信息。注意，研究方案批准的完整信息也必须在手稿中提供。

人口特征	该研究包括一名接受此实验情况并参加了BrainGate2神经接口系统临床试验的参与者（确认为T5）的数据（ClinicalTrials.gov identifier: NCT00912041，注册于2009年6月3日），但请注意，该研究没有报告临床试验结果。T5为右利手男性，收集数据时为65岁，在纳入研究前约9年发生C4 AIS - C脊髓损伤。
招募	在满足了部分基于疾病特征的纳入标准后，参与者T5在设计和执行本研究之前被纳入了BrainGate2试点临床试验。纳入和排除标准可在网上查询（ClinicalTrials.gov）。
道德监督	BrainGate2神经接口系统临床试验根据美国食品和药物管理局的研究器械豁免（IDE）获得批准（研究器械豁免#G090003）。斯坦福大学的机构审查委员会也给予了许可（协议#20804）。所有的研究都是按照相关的指南/规定进行的。

References

- [1] Hochberg, L. R. et al. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature* 485, 372375 (2012).
- [2] Collinger, J. L. et al. High-performance neuroprosthetic control by an individual with tetraplegia. *Lancet* 381, 557564 (2013).
- [3] Aflalo, T. et al. Neurophysiology. Decoding motor imagery from the posterior parietal cortex of a tetraplegic human. *Science* 348, 906910 (2015).
- [4] Bouton, C. E. et al. Restoring cortical control of functional movement in a human with quadriplegia. *Nature* 533, 247250 (2016).
- [5] Ajiboye, A. B. et al. Restoration of reaching and grasping movements through brain-controlled muscle stimulation in a person with tetraplegia: a proof-of-concept demonstration. *Lancet* 389, 18211830 (2017).
- [6] Jarosiewicz, B. et al. Virtual typing by people with tetraplegia using a self-calibrating intracortical braincomputer interface. *Sci. Transl. Med.* 7, 313ra179 (2015).
- [7] Pandarinath, C. et al. High performance communication by people with paralysis using an intracortical braincomputer interface. *eLife* 6, e18554 (2017).
- [8] Palin, K., Feit, A. M., Kim, S., Kristensson, P. O. & Oulasvirta, A. How do people type on mobile devices? Observations from a study with 37,000 volunteers. In *Proc. 21st International Conference on HumanComputer Interaction with Mobile Devices and Services* 112 (Association for Computing Machinery, 2019).
- [9] Yousry, T. A. et al. Localization of the motor hand area to a knob on the precentral gyrus. A new landmark. *Brain* 120, 141157 (1997).
- [10] Willett, F. R. et al. Hand knob area of premotor cortex represents the whole body in a compositional way. *Cell* 181, 396409 (2020).
- [11] Williams, A. H. et al. Discovering precise temporal patterns in large-scale neural recordings through robust and interpretable time warping. *Neuron* 105, 246259 (2020).
- [12] Hinton, G. et al. Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups. *IEEE Signal Process. Mag.* 29, 8297 (2012).
- [13] Graves, A., Mohamed, A. & Hinton, G. Speech recognition with deep recurrent neural networks. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing* 66456649 (2013).
- [14] Xiong, W. et al. The Microsoft 2017 Conversational Speech Recognition System. Preprint at <https://arxiv.org/abs/1708.06073> (2017).
- [15] He, Y. et al. Streaming end-to-end speech recognition for mobile devices. In *2019 IEEE International Conference on Acoustics, Speech and Signal Processing* 63816385 (2019).
- [16] Anumanchipalli, G. K., Chartier, J. & Chang, E. F. Speech synthesis from neural decoding of spoken sentences. *Nature* 568, 493498 (2019).
- [17] Makin, J. G., Moses, D. A. & Chang, E. F. Machine translation of cortical activity to text with an encoder-decoder framework. *Nat. Neurosci.* 23, 575582 (2020).
- [18] Chen, X. et al. High-speed spelling with a noninvasive braincomputer interface. *Proc. Natl Acad. Sci. USA* 112, E6058E6067 (2015).
- [19] Dickey, A. S., Suminski, A., Amit, Y. & Hatsopoulos, N. G. Single-unit stability using chronically implanted multielectrode arrays. *J. Neurophysiol.* 102, 13311339 (2009).
- [20] Eleryan, A. et al. Tracking single units in chronic, large scale, neural recordings for brain machine interface applications. *Front. Neuroeng.* 7, 23 (2014).

- [21] Downey, J. E., Schwed, N., Chase, S. M., Schwartz, A. B. & Collinger, J. L. Intracortical recording stability in human braincomputer interface users. *J. Neural Eng.* 15, 046016 (2018).
- [22] Willett, F. R. et al. Signal-independent noise in intracortical braincomputer interfaces causes movement time properties inconsistent with Fitts' law. *J. Neural Eng.* 14, 026010 (2017).
- [23] Gao, P. et al. A theory of multineuronal dimensionality, dynamics and measurement. Preprint at <https://doi.org/10.1101/214262> (2017).
- [24] Musallam, S., Corneil, B. D., Greger, B., Scherberger, H. & Andersen, R. A. Cognitive control signals for neural prosthetics. *Science* 305, 258262 (2004).
- [25] Santhanam, G., Ryu, S. I., Yu, B. M., Afshar, A. & Shenoy, K. V. A high-performance braincomputer interface. *Nature* 442, 195198 (2006).
- [26] Cunningham, J. P., Yu, B. M., Gilja, V., Ryu, S. I. & Shenoy, K. V. Toward optimal target placement for neural prosthetic devices. *J. Neurophysiol.* 100, 34453457 (2008).
- [27] Pels, E. G. M., Aarnoutse, E. J., Ramsey, N. F. & Vansteensel, M. J. Estimated prevalence of the target population for braincomputer interface neurotechnology in the Netherlands. *Neurorehabil. Neural Repair* 31, 677685 (2017).
- [28] Vansteensel, M. J. et al. Fully implanted braincomputer interface in a locked-in patient with ALS. *N. Engl. J. Med.* 375, 20602066 (2016).
- [29] Nijboer, F. et al. A P300-based braincomputer interface for people with amyotrophic lateral sclerosis. *Clin. Neurophysiol.* 119, 19091916 (2008).
- [30] Townsend, G. et al. A novel P300-based braincomputer interface stimulus presentation paradigm: moving beyond rows and columns. *Clin. Neurophysiol.* 121, 11091120 (2010).
- [31] McCane, L. M. et al. P300-based braincomputer interface (BCI) event-related potentials (ERPs): people with amyotrophic lateral sclerosis (ALS) vs. age-matched controls. *Clin. Neurophysiol.* 126, 21242131 (2015).
- [32] Wolpaw, J. R. et al. Independent home use of a braincomputer interface by people with amyotrophic lateral sclerosis. *Neurology* 91, e258e267 (2018).
- [33] Bacher, D. et al. Neural point-and-click communication by a person with incomplete locked-in syndrome. *Neurorehabil. Neural Repair* 29, 462471 (2015).
- [34] Mugler, E. M. et al. Direct classification of all American English phonemes using signals from functional speech motor cortex. *J. Neural Eng.* 11, 035015 (2014).
- [35] Nurmikko, A. Challenges for large-scale cortical interfaces. *Neuron* 108, 259269 (2020).
- [36] Vázquez-Guardado, A., Yang, Y., Bandodkar, A. J. & Rogers, J. A. Recent advances in neurotechnologies with broad potential for neuroscience research. *Nat. Neurosci.* 23, 15221536 (2020).
- [37] Simeral, J. D., Kim, S.-P., Black, M. J., Donoghue, J. P. & Hochberg, L. R. Neural control of cursor trajectory and click by a human with tetraplegia 1000 days after implant of an intracortical microelectrode array. *J. Neural Eng.* 8, 025027 (2011).
- [38] Bullard, A. J., Hutchison, B. C., Lee, J., Chestek, C. A. & Patil, P. G. Estimating risk for future intracranial, fully implanted, modular neuroprosthetic systems: a systematic review of hardware complications in clinical deep brain stimulation and experimental human intracortical arrays. *Neuromodulation* 23, 411426 (2020).
- [39] Nuyujukian, P. et al. Cortical control of a tablet computer by people with paralysis. *PLoS One* 13, e0204566 (2018).
- [40] Musk, E. An integrated brainmachine interface platform with thousands of channels. *J. Med. Internet Res.* 21, e16194 (2019).
- [41] Sahasrabudhe, K. et al. The Argo: a high channel count recording system for neural recording in vivo. *J. Neural Eng.* 18, 015002 (2021).
- [42] Sussillo, D., Stavisky, S. D., Kao, J. C., Ryu, S. I. & Shenoy, K. V. Making brainmachine interfaces robust to future neural variability. *Nat. Commun.* 7, 13749 (2016).
- [43] Dyer, E. L. et al. A cryptography-based approach for movement decoding. *Nat. Biomed. Eng.* 1, 967976 (2017).
- [44] Degenhart, A. D. et al. Stabilization of a braincomputer interface via the alignment of low-dimensional spaces of neural activity. *Nat. Biomed. Eng.* 4, 672685 (2020).