

非侵入信号神经语音解码

谢旭荣、秦颖、王昊铭、张高媛、周超、
陈辉、田丰

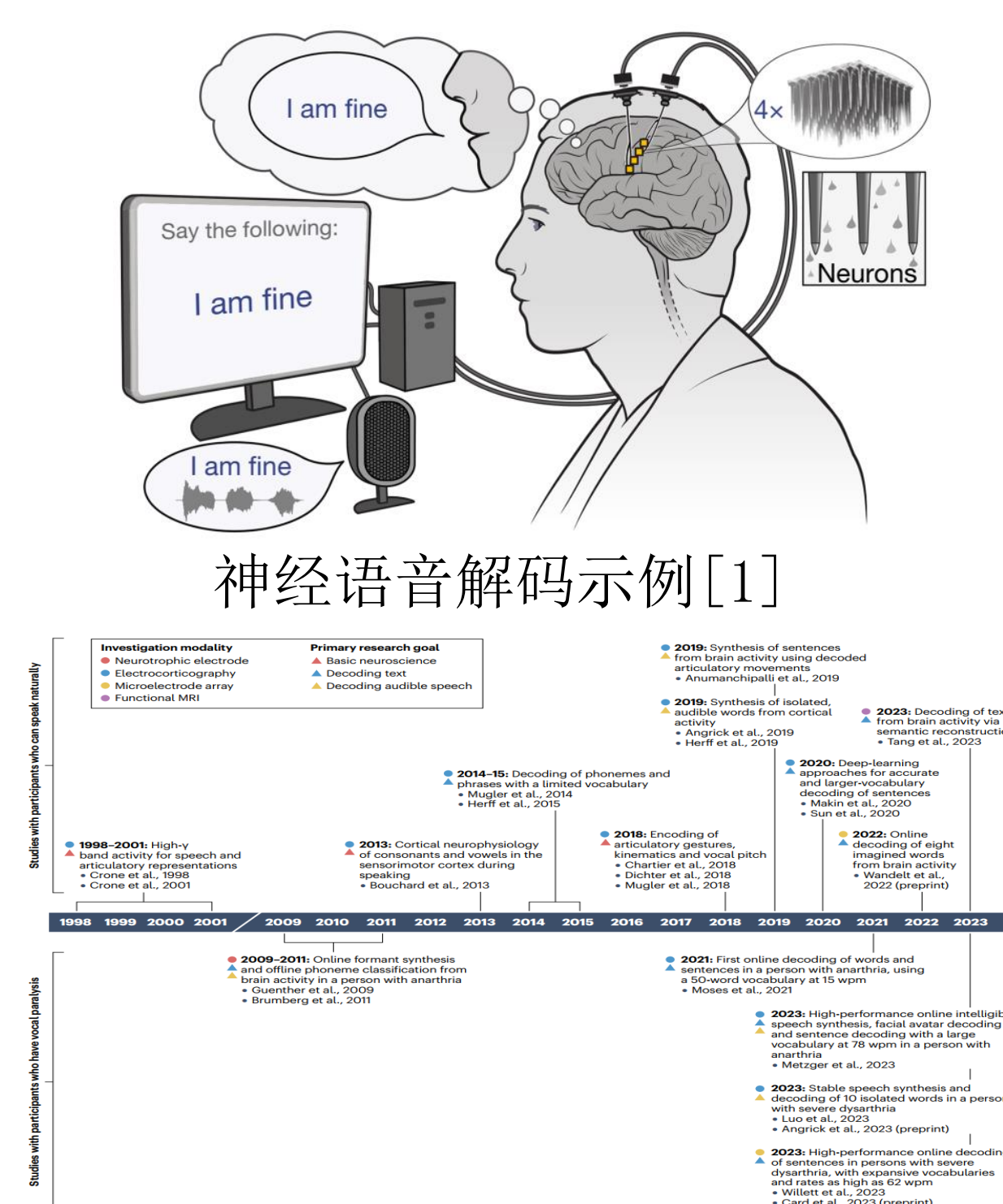
联系方式（谢旭荣 xurong@iscas.ac.cn）

研究背景

神经系统疾病如脑卒中等患者可能失去口头言语能力。神经语音解码技术从言语交互过程神经信号提取患者语音信息，并解码为可理解文本或重建为清晰语音，**能有助于患者康复及辅助交流，显著提高其生活质量。**

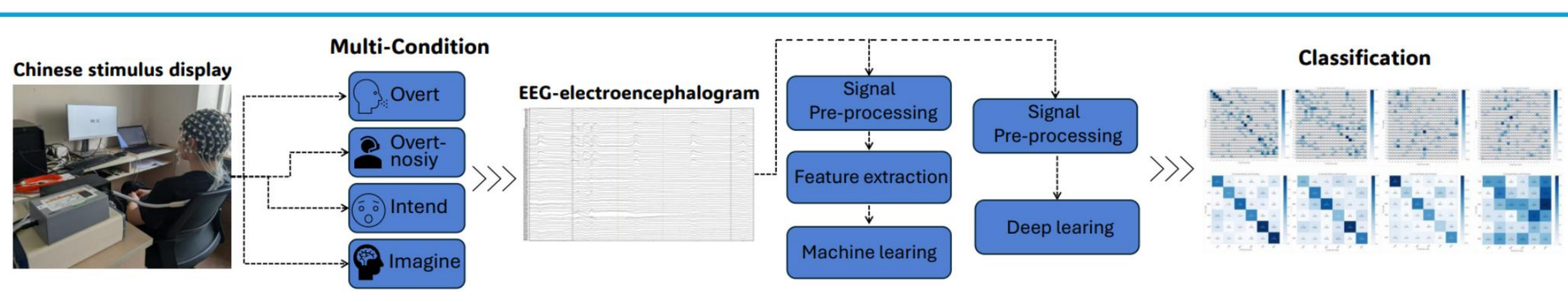
神经语音解码（也称语音神经假体）是当前人机交互及人工智能的前沿研究领域，主要基于皮层电极等侵入式神经信号。**对于非侵入信号，尤其是成本较低、使用较方便的脑电（EEG）信号研究较少，且解码目标主要为非中文内容。**

本工作面向基于**EEG信号的中文神经语音解码开展系统性研究**，包括交互任务内容、条件、范式设计，数据处理与建模，结果分析等方面，并提出了一套行之有效的关键技术路线。



关键技术进展[2]

关键技术



本工作在科技创新2030—新一代人工智能重大项目“面向神经系统疾病预警的智能人机交互关键技术”支持下开展，关键技术如下：

任务内容

主要面向中文言语任务，包括**带声调的中文元音**和覆盖**日常场景及医院场景的短语**，如“吃饭”、“厕所”、“头疼”、“换药”等。

实验范式

单次流程为：**刺激呈现 -> 提示准备 -> 任务执行 -> 休息**。其中刺激以屏幕显示文本方式呈现。

实验设置

为了模拟不同条件患者，设置**4种条件**：“**正常发音**”；“**噪声发音**”使用噪声剥夺听觉反馈；“**意图发音**”有发音动作但不发声；“**想象发音**”无发音动作。

数据处理与建模

通过降采样、坏导插值、滤波、ICA去干扰、傅里叶变换等方法提取数据特征，并建立多种机器学习模型和深度学习模型，对元音和短语进行分类。

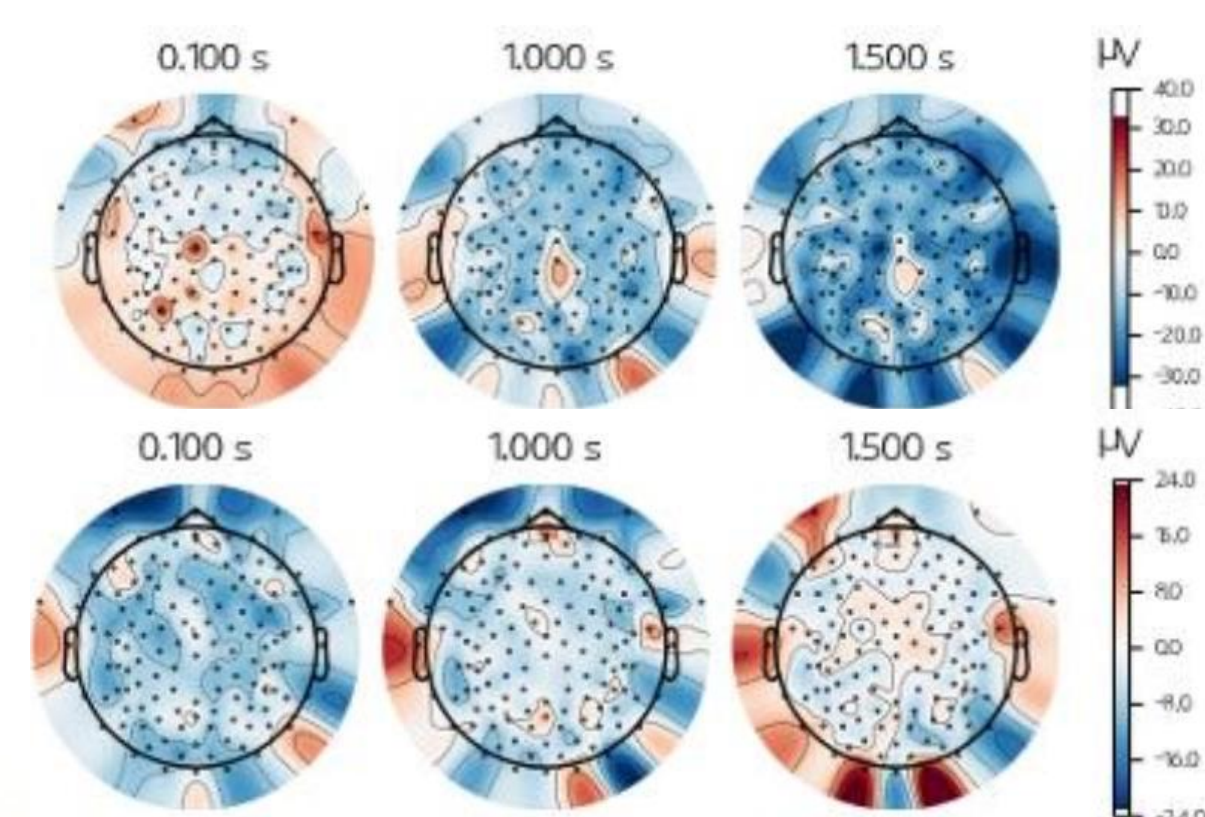
数据与结果

基于EGI 128导脑电设备采集1名被试约**11h有效数据**，如下

Dataset structure	Healthy population	Patient
vowel with tone	Overt(1.64h)	Overt-noisy(2.1h), Intend(1.87h), Imagine(1.67h)
word	Overt(1.15h)	Overt-noisy(1.15h), Intend(1.17h), Imagine(1.23h)

以CNN模型短语分类结果为例，所有结果均**显著优于随机**，短语分类**基本达到可应用水平**，其中“噪声发音”和“意图发音”效果相较“正常发音”差异较小，“想象发音”效果显著下降，提示其困难性。

Task type	Accuracy	Precision	Recall	F1
Word-Overt	75.6%	71.2%	69.3%	70.24%
Word-Overt-noisy	60.6%	58.7%	56.2%	57.4%
Word-intend	69.5%	66.5%	65.6%	66%
Word-imagine	28.1%	29.3%	26.6%	27.9%



不同实验条件下神经激活

引用

- Wairagkar, M., Card, N. S., Singer-Clark, T., Hou, X., Iacobacci, C., Miller, L. M., ... & Stavisky, S. D. (2025). An instantaneous voice-synthesis neuroprosthesis. *Nature*, 1-8.
- Silva, A. B., Littlejohn, K. T., Liu, J. R., Moses, D. A., & Chang, E. F. (2024). The speech neuroprosthesis. *Nature Reviews Neuroscience*, 25(7), 473-492.