arXiv:2601.03098v1 [cs.LG] 6 Jan 2026

借助 MyoText 从肌肉到文本：通过手指分类和基于 Transformer 的解码，将表面肌电转化为文本

MEGHNA ROY CHOWDHURY, 普渡大学，美国
SHREYAS SEN, 普渡大学，美国YI DING, 普渡大学，美国

表面肌电图（表面肌电）可作为一种直接的神经接口，用于解码肌肉活动，并为可穿戴和混合现实系统中的无键盘文本输入提供了有前景的基础。以往的肌电转文本研究主要致力于直接从表面肌电信号中识别字母，这是将肌肉活动转化为文本的重要第一步。在此基础上，我们提出了 MyoText，这是一种分层框架，通过基于生理机制的中间阶段，将表面肌电信号解码为文本。MyoText 首先使用 CNN–BiLSTM–Attention模型从多通道表面肌电中分类手指激活，再结合符合人体工学的打字先验推断字母，并使用经过微调的 T5 Transformer 重建完整句子。这种模块化设计体现了打字的自然层次结构，将肌肉意图与语言输出联系起来，并缩小了解码的搜索空间。在 emg2qwerty 数据集的 30 名用户上进行评估时，MyoText 的表现优于基线，手指分类准确率达到 85.4%，字符错误率（CER）为 5.4%，词错误率（WER）为 6.5%。除准确率提升外，这一方法还为从神经肌肉信号到文本的转换建立了有原则的途径，为完全无需实体键盘的虚拟现实和增强现实打字界面提供了蓝图。通过将符合人体工学的结构与基于 Transformer 的语言推理相结合，MyoText 推进了无缝可穿戴神经输入在未来普适计算环境中的应用可行性。

CCS 概念： • **以人为本的计算 → 普适计算与移动计算中的实证研究**； • **计算方法 →可演化硬件**；**神经网络**。

美国计算机协会参考文献格式：
Meghna Roy Chowdhury、Shreyas Sen 和 Yi Ding。2026年。从肌肉到文本：借助 MyoText 通过手指分类和基于 Transformer 的解码将表面肌电转换为文本。 美国计算机协会交互、移动、可穿戴与普适技术论文集。 0，0，文章 0（2026年），25页。 https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 引言

可穿戴计算和普适计算的发展正日益模糊人与机器之间的界限，使人们能够在物理世界和虚拟世界之间进行持续的免手交互 [87]。随着用于测量生物信号的生物传感设备不断小型化，并进一步融入日常生活，用户需要可靠、自然的输入方式，而不依赖特定表面 [11]。然而，触摸屏、摄像机和运动传感器等常见输入方式仍依赖视觉、光照或大幅度手势，从而限制了移动设备和增强现实的性能 [78]。

在众多新型输入技术中，肌电图（EMG）通过测量肌肉运动时产生的电信号，直接连接人体的运动系统 [86]。表面肌电（sEMG）可捕捉这些信号

作者联系方式：Meghna Roy Chowdhury，mroycho@purdue.edu，美国印第安纳州西拉斐特普渡大学；Shreyas Sen，shreyas@purdue.edu，美国印第安纳州西拉斐特普渡大学；Yi Ding，yiding@pudue.edu，美国印第安纳州西拉斐特普渡大学。

允许出于个人或课堂用途免费制作本作品全部或部分内容的数字或纸质副本，但不得为营利或商业利益而制作或分发副本，且副本首页必须附有本声明和完整引文。本作品中归其他作者所有的组成部分，其版权必须得到尊重。注明出处后允许摘录。若要以其他方式复制、再版、发布到服务器或分发至列表，须事先获得特定许可和／或支付费用。请发送邮件至 permissions@acm.org 申请许可。©2026年版权归所有者／作者所有。出版权已授权给美国计算机协会。美国计算机协会
2474-9567/2026/0-ART0https://doi.org/10.1145/nnnnnnn.nnnnnnn

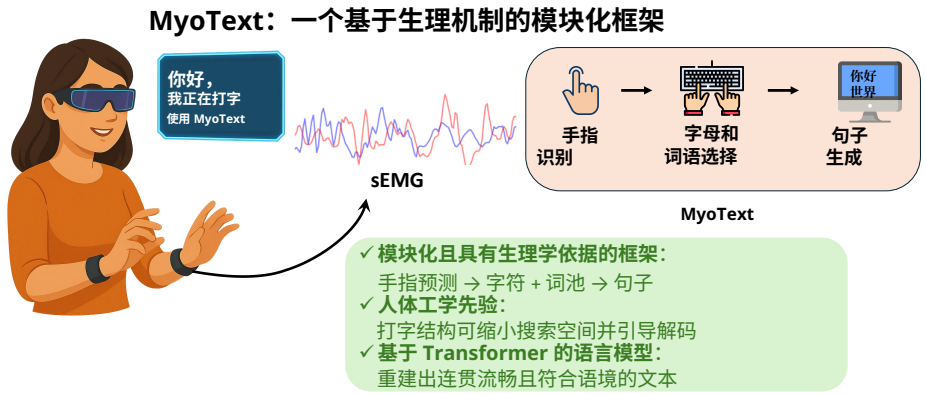


图 1。 **MyoText**。 一种基于生理机制的模块化表面肌电到文本解码框架，结合手指分类 n, 符合人体工学的字母输入与句子重构。

通过皮肤贴附式电极进行非侵入式采集 [17, 21], 支持安静、灵活的交互。这些特性使表面肌电适用于智能手表、AR眼镜和神经腕带等可穿戴设备与混合现实设备 [80, 86]。Meta Neural Band 等近期系统表明，基于腕部的表面肌电传感器能够识别细微的手指动作，从而实现轻松、隐蔽的交互 [48, 54, 72]。这些能力对于 AR 和虚拟现实中的文本输入至关重要，使用户无需实体键盘即可在空中打字。在此类环境中，通过表面肌电解码肌肉活动，是实现扩展现实（XR）无缝文本输入的重要一步。

尽管前景广阔，要实现稳健的表面肌电转文本系统，仍需应对三个关键挑战。首先，由于电极放置、肌肉疲劳和个体解剖结构等因素，表面肌电信号在不同用户和会话之间会自然发生变化 [58, 60, 89, 93, 111]。研究人员已经开发出重要的稳定化技术，包括信号滤波、标准化电极定位和用户个体校准，但在不进行校准的情况下实现可靠的跨用户泛化仍是一个悬而未决的问题。其次，为应对这种变异性，该领域通过分层解码流程取得了进展：先检测动作，再对手势或字符进行分类 [14, 56], ；同时，循环神经网络等序列模型 [43] 和 Transformer [98] 能够捕捉时间依赖关系。这些架构方面的进展提高了会话内的稳定性，而将其扩展到大规模、具有语言结构的数据集则代表着下一阶段的前沿方向。第三，尽管 NinaPro（27 名用户） [68], 、MyoGym（12 名用户） [52], 、CSL-HDEMG（5 名用户） [3], 和 CapgMyo（18 名用户） [24], 等基础数据集已广泛用于手势识别，但连续的文本级解码需要规模更大、具有自然语言结构的数据。Meta Reality Labs 的 *emg2qwerty* 数据集 [92]代表了一个重要里程碑：该数据集在打字任务期间采集了来自 108 名用户的标注的表面肌电信号。参与者经过筛选，以确保其在标准 QWERTY 键盘布局下具有至少 90% 的正确手指到按键映射 [8, 105, 114], ，从而大规模捕捉符合人体工学的打字行为。利用该数据集，作者训练了一个卷积神经网络，将 EMG活动直接映射为输入字符；在加入 6元语言模型后，字符错误率（CER）取得了显著改善。这一结果凸显了一个重要认识：系统对语言先验的强烈依赖表明，直接从表面肌电信号预测字符可能与信号的生理结构并不一致。表面肌电信号自然编码了手指级肌肉激活，因此促使研究人员采用先解码手指、再通过符合人体工学的约束和语言约束消解字符歧义的方法。

基于这一洞见，我们提出 MyoText，这是一个基于生理机制、模块化的表面肌电到文本框架，将打字建模为运动—语言层级。如图 1所示，MyoText 分三个阶段运行：(1) 根据表面肌电信号进行手指分类；(2) 进行符合人体工学约束的字母与单词汇集；(3) 进行基于 Transformer 的句子生成。这一设计基于三项核心原则。首先，表面肌电信号

自然编码的是手指级肌肉激活，而非抽象的字符身份。通过预测手指而非直接预测字母，MyoText 使解码任务与信号的生理结构相契合，从而减少歧义并提升跨用户泛化能力。其次，打字人体工学提供了可缩小解码搜索空间的结构性约束。标准化的 QWERTY 手指到按键映射定义了一种自然先验，可引导字母和单词池化，同时借助下游 Transformer 对非典型打字模式保持鲁棒性。第三，将运动解码与语言推理分离，既能提高准确率，也有助于部署。受脑机接口中的层级解码启发 [14, 56], ，这种模块化设计支持分布式执行：轻量级手指分类器可在资源受限的可穿戴设备上运行，而 Transformer 解码器则可在智能手机等资源充足的配对设备上运行 [16]。

我们在包含30名参与者的 *emg2qwerty* 数据集上评估了 MyoText，这些参与者经过筛选，兼顾了多样性与计算可行性。MyoText đạt到85.4%的手指分类准确率，并以5.4%的字符错误率（CER）和6.5%的词错误率（WER）重构句子，显著优于以往直接进行字符预测的方法。我们还在非规范手指到按键映射和语言变体条件下进一步测试了鲁棒性，观察到一致的解码结果。

总而言之，本研究作出以下贡献：

- 我们提出了 MyoText，这是一种基于生理机制的分层表面肌电到文本解码框架，通过手指分类、人体工学约束和基于 Transformer 的句子生成方法，将运动意图与语言解码分离开来。
- 我们设计了一个用于手指激活识别的 CNN–BiLSTM–Attention模型，以及一个受约束的 T5解码器；在 *emg2qwerty* 数据集上，该模型实现了85.4%的手指预测准确率和5.4%的字符错误率。
- 我们将打字人体工学编码为一种结构先验，用于约束基于 Transformer 的解码的搜索空间，并提升对非规范手指—按键映射的鲁棒性。
- 我们分析了 *emg2qwerty* 数据集，揭示了一致的手指依赖的表面肌电结构，并展现出——通过留一用户评估、在英语语料库上进行跨数据集测试，以及在非典型打字模式下测试鲁棒性，证明了 MyoText 的泛化能力。

这些贡献共同推动了基于表面肌电的文本系统的发展，表明实现可泛化的解码是可能的。通过结合人体工学设计、模块化架构和鲁棒语言模型，这项工作为表面肌电驱动的通信奠定了实用且可扩展的基础，为人体互联网的未来铺平了道路 [88]。

2 背景与相关工作

本节首先介绍肌电图（EMG）生物信号及其背后的生理学原理。随后，我们重点介绍肌电在可穿戴设备中的关键应用，尤其关注表面肌电到文本解码系统。最后，我们讨论当前表面肌电到文本研究中的空白。

肌电图（肌电）。肌电图（肌电）是一种生理信号，可捕捉骨骼肌激活产生的电活动，为人体神经肌肉系统提供直接接口 [67]。肌电信号由运动单位动作电位产生，通常范围为 1–10 mV，主要频率介于 50–150 Hz [21]。这些生物电位信号既可通过针电极以侵入方式测量，也可通过皮肤贴附式电极以非侵入方式测量表面肌电（sEMG） [40]。虽然针式肌电具有临床

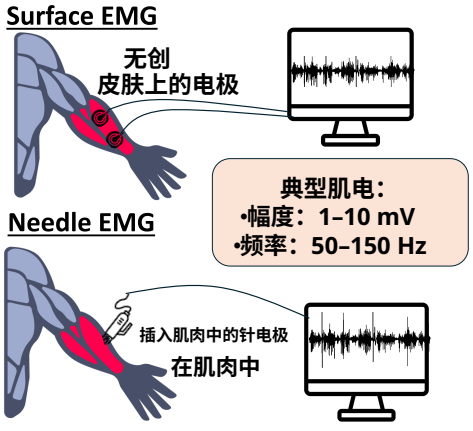


图2. **肌电感知模态**。 针电极和表面肌电信号采集方法。

精确度较高，但表面肌电由于舒适、易于使用，且有望集成到日常接口中，因此在可穿戴和移动场景中得到更广泛的应用 [82]。肌肉组织在静息时保持电静息状态，但在收缩过程中会产生独特的时空模式，以编码运动意图 [29]。这些模式构成了解读用户动作并实现响应式、意图感知系统的基础。

可穿戴设备上的表面肌电。 可穿戴设备设计的最新进展显著提升了基于表面肌电的系统的易用性 [86]。如今，这些设备已集成到臂带、纺织品、智能鞋和手套中，覆盖了广泛的平台。表 1列出了具有表面肌电功能的代表性可穿戴设备及其应用领域 [22, 32, 34, 49, 56, 62, 81, 113]。其中许多设备支持手势识别、运动能力分析或康复，展现了表面肌电信号在人机交互中的多功能性。表面肌电可穿戴设备的最新进展是 Meta Neural Band [72], ，这是一种神经接口，能够通过基于腕部的表面肌电，以低延迟响应解码细粒度的运动意图。这些进展体现了表面肌电从临床和生物力学工具向普及型输入接口的转变，为新一代打字和控制范式提供了支持。

表面肌电的应用。表面肌电最初被用作神经肌肉疾病的诊断工具，例如肌营养不良症、重症肌无力和腕管综合征 [95, 96]。随后，它推动了肌电假肢的发展，使残余肌肉信号能够直观地控制假肢 [18, 97, 118]。随着技术日趋成熟，表面肌电的应用扩展至运动医学、康复和人体工学领域，通过实时肌肉反馈改善姿势、降低受伤风险并提升表现 [2, 19, 53, 107]。其应用范围涵盖可穿戴机器人技术、步态分析、跌倒风险监测以及年龄相关性肌少症研究 [36, 42, 57, 70, 101]。除了身体层面的应用，表面肌电也日益用于将人的意图转化为交流内容：相关系统能够识别手语手势，为聋人群体提供更自然的交互界面；还可以通过解码预期动作并将其转换为文本或控制指令，帮助部分瘫痪或存在运动障碍的人士 [1, 9, 15, 28, 75, 94]。表面肌电到文本的发展路径，通过采集前臂、面部和喉部的肌肉活动，并将其映射为书面语言或控制信号 [46, 47, 108]。表面肌电也已融入沉浸式虚拟环境，将肌肉意图直接转化为数字控制指令 [13, 25, 104]。最新进展是 Meta 的表面肌电神经接口腕带，它能够检测细微的肌肉信号，并将其转换为文本和指令，推动面向日常交流与互动的无缝非侵入式接口发展 [48, 54]。

表 1. **表面肌电可穿戴设备概览**。

Ref.	形态	应用
[49]	可穿戴腕带	手势识别（太阳能供电）
[62]	智能鞋类	足部手势识别
[34]	纺织短裤	表面肌电监测
[56]	智能袜	腿部肌肉评估
[32]	机器人手套	手部康复
[113]	Msleeve 和 Mbody 服	运动与人体工学
[81]	Myo 臂带	用于手势控制的表面肌电传感
[72]	Meta 神经腕带	神经接口与输入解码

sEMG 信号面临的挑战。 尽管 sEMG 信号的应用日益受到关注，但其仍存在诸多技术挑战，尤其是在动态的现实世界人机交互环境中。sEMG 信号天然具有非平稳性，且通常表现出较低的信噪比（SNR） [76]。表面肌电记录会受到多种生理和环境因素的影响，包括电干扰、运动伪迹、出汗、电极接触变化以及用户疲劳 [27, 102]。此外，sEMG 捕获的是多个重叠肌纤维的总体活动，这可能因解剖结构变异和肌间串扰而降低信号的特异性 [5]。即使在不同会话之间电极放置位置仅发生细微变化，也可能导致信号模式发生显著改变，从而使个性化以及跨用户的稳健泛化更加复杂。

sEMG 到文本输入的相关研究。 在本研究中，我们聚焦于表面肌电到文本解码。既有研究表明，该领域经历了三个主要阶段：从离散按键分类，到连续手势映射，最终发展到从流式信号中进行连续文本解码。早期研究证明，前臂表面肌电信号能够清晰地区分单个按键动作。例如，Sharma 等人 [90]提取统计特征，并使用由强化信号引导的 k 近邻分类器，可靠地区分不同的手指动作。Fu 等人 [37] 在此基础上进一步将协调的三指手势映射到 T9 键盘布局；其 CNN–LSTM 模型捕获了多按键序列中的频谱和时序结构，从而超越了逐事件识别。随后，Crouch 等人 [20] 将基于肌电的打字重新定义为流式序列建模问题，其方式类似于语音识别，能够直接根据重叠的肌肉活动生成文本，而无需进行显式分段。在此思路的基础上，Sivakumar 等人 [92] 将自动语音识别风格的模型应用于 *emg2qwerty* 数据集，采用时间–深度可分离卷积神经网络（具有 1 秒感受野）以及特定频带的旋转不变滤波器。他们使用连接主义时序分类（CTC）损失训练模型，并通过无词典的 6 元字符级语言模型对输出进行解码，同时采用改进的 Kneser–Ney 平滑和束搜索。尽管这些系统专注于复现实体键盘打字，但下一步是实现虚拟键盘文本交互，这正是本研究方法的动机所在。

当前最先进技术面临的挑战。 尽管近期取得了进展，表面肌电到文本解码仍面临诸多挑战。表面肌电信号本身存在噪声、因用户而异，且往往表达含糊的信息。Sivakumar 等人 [92] 近期的研究结果表明，模型在同一用户上的准确率很高，但跨用户迁移能力有限。经过微调的个性化模型搭配 6元语言模型时，字符错误率（CER）达到 6.9%；而通用模型（跨用户）则达到 51.78%。在不使用语言模型时，字符错误率分别升至 11.28% 和 55.38%，这凸显了当前系统对语言先验的高度依赖。这些结果表明，亟需开发能够更有效地跨用户泛化，并更高效地利用语言模型的方法。未来研究应着重开发只需极少校准便能学习可迁移肌肉-语言表征的模型，直接从表面肌电动态中捕捉局部上下文，并主要利用语言模型提供长程支持。最终目标是打造舒适、快速且稳健的可穿戴表面肌电到文本系统，以支持连续无键盘打字。

3 字母到手指的人体工学映射及其变异性

为应对上述挑战并指导我们的设计，我们采用了一种符合人体工学的字母到手指的映射方式，以捕捉共同的运动模式，并使不同用户和会话中的解码更加稳定。我们首先根据表面肌电推断当前活动的手指，然后在该手指对应的按键集合中解码字母，从而缩小有效搜索空间，并支持虚拟／无键盘操作，同时将语言模型作为轻量级的长程先验。QWERTY键盘布局天然适用于这一方法。作为英语打字的标准布局，它通过尽量减少手指运动，在速度与舒适度之间取得平衡 [71]。如图所示，3（a），盲打遵循一项人体工学规则：每个字母按键都对应八根手指中的一根（不包括拇指） [8, 105, 114]。打字员将手指放在基准行上，ASDF对应左手，HJKL 对应右手，并在每次击键后将手指放回基准行 [103]。这种习惯限制了动作幅度，使手指负担分配均匀；同时，受训打字员无需再判断该使用哪根手指，因此能够以更少的脑力更快地打字。这种映射体现了打字训练中长期沿用的惯例，并广泛应用于人机交互和运动控制研究 [84, 115, 116]。研究表明，盲打者依靠稳定的双手手指摆放和动觉反馈，而非视觉线索 [23], ；受训打字员则会形成牢固的肌肉记忆模式，使他们无需看键盘也能打字 [39]。

我们对 *emg2qwerty* 数据集应用相同的规则，将每个字母映射到这八根手指之一（编号为 0–9，不包括拇指）。在图3(b) 中，我们发现，e、t、n和 o 等常见字母主要由食指和中指（2、3、6）输入，而 q、z和 p 等罕见字母则由小指（0、9）输入。这种模式

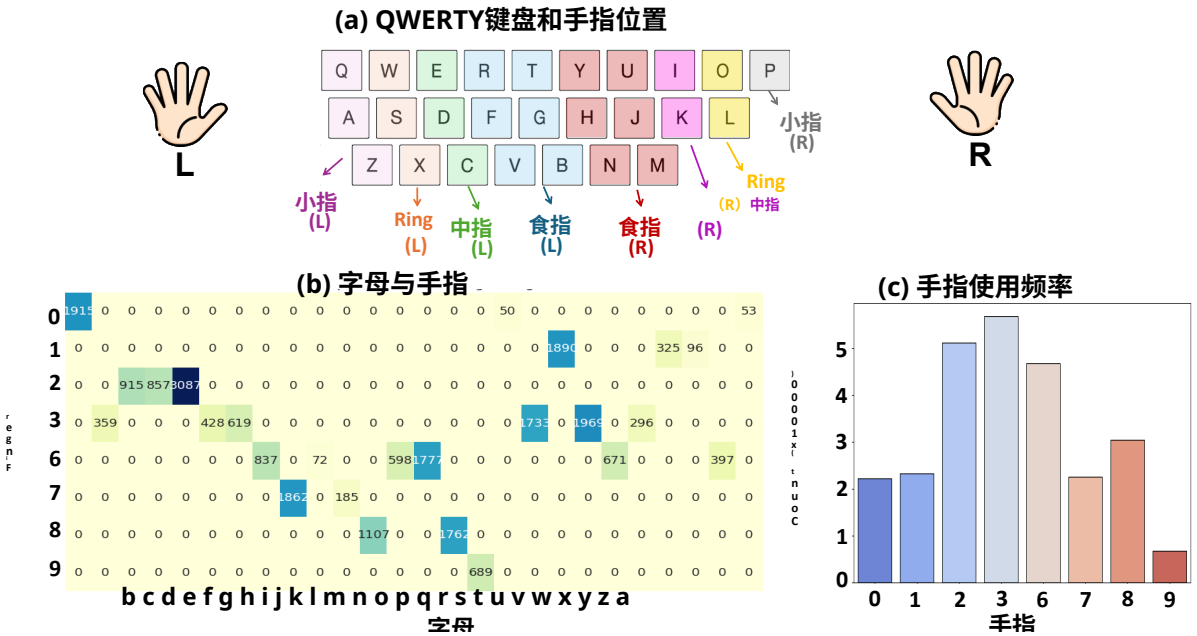


图 3. **QWERTY 人体工学与手指使用情况**。(a) 展示了标准 QWERTY键盘布局及其常规的手指与按键分配。字母分配给八根活动手指；为完整起见，也纳入了拇指（4、5）。(b) 基于手指的字母映射。(c) 手指使用频率，突出显示了在 *emg2qwerty* 数据集中食指和中指占主导地位。

使表面肌电信号呈现出有规律的变化：中间的手指活动更频繁，产生的信号也更稳定；而边缘手指激活较少，变异性更大。图 3(c) 中的手指使用分布证实了这种不对称性，反映了人体工学设计和有意义的生理结构。我们也承认，许多用户的盲打方式并不符合标准惯例（例如采用其他指法，尤其是在键盘中间区域的按键上）。为考虑这一点，我们的评估明确纳入了 *emg2qwerty* 中的打字变异性（§5.2.2），以评估模型在不同打字风格和非标准手指与字母分配情况下的表现。解码模型可以利用这种手指与字母层级关系作为内置结构，以降低噪声，并更好地泛化到不同用户和会话。简而言之，以基准行惯例为基础的手指组织方式，既是有用的约束，也是从表面肌电信号构建稳健解码器的重要先验。

4 提出的方法：MyoText框架

我们提出了**MyoText**，这是一种模块化的表面肌电到文本解码框架，其设计仿照了人类的打字行为。如图 4所示，MyoText框架包含四个步骤：(i) 数据准备，我们使用 *emg2qwerty* 数据集中的表面肌电信号，并将其与 QWERTY键盘上键入的字母对齐，随后根据打字人体工学，将每个字母映射到对应的手指；(ii) 手指级分类，使用带有注意力机制的双向长短期记忆网络，根据表面肌电信号对手指活动进行分类；(iii)字母汇聚与候选词构建，将预测手指映射为候选字母，并据此生成合理的候选词；(iv) 通过基于 *Transformer* 的解码进行句子预测，使用经过微调的、基于 T5 的 Transformer语言模型，将候选词解码为语法和语义上均合理的句子。接下来的小节将详细介绍 MyoText 各步骤的设计。

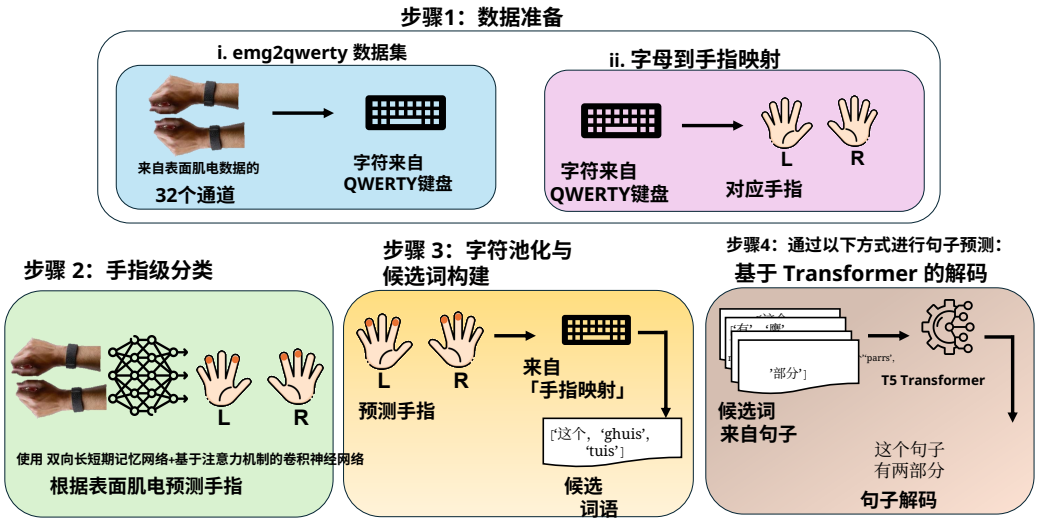


图 4. **MyoText：模块化表面肌电转文本框架**。从表面肌电信号到句子解码的四阶段示意图：(i) 使用符合人体工学的手指映射进行数据准备，(ii) 手指分类，(iii) 字母与候选词汇集，以及 (iv) 基于 T5 的解码。

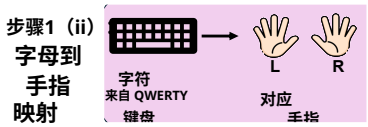
4.1 数据准备

首先，我们从 *emg2qwerty*数据集的原始 HDF5 日志中，为每次按键提取与时间对齐的表面肌电片段 [30]。尽管此流程是专为 *emg2qwerty*设计的，但我们的模块化流水线也能轻松适配其他数据集或预处理方案。每条日志都记录了来自 32 个电极（每侧前臂各 16 个）的同步表面肌电信号，以及每次击键的时间戳。对于每次按键，我们根据记录的按键及其时间戳定位对应的表面肌电时间区间，然后提取事件前后各 1 秒的短片段，以捕捉肌肉的预备活动和残余活动。我们保留全部 32 个通道的原始分辨率，并将处理后的片段保存为压缩文件，以加快后续分析。

接下来，我们遵循标准打字人体工学 [8, 105,114] 和 QWERTY键盘布局 [71]，为每个字母标注通常按下该字母的手指。我们将字母映射到八根活动手指（左手索引为 0–3，右手索引为 6–9），并将拇指留作按空格键使用。表2 展示了完整的按键到手指的映射。尽管并非所有用户都严格遵循 QWERTY 规范，但此前的打字生物力学和表面肌电研究均采用了这一映射作为一致的基线。经典的人体工学研究测量了十名专业打字员的细线表面肌电信号 [33], ，之后的研究也继续招募经验丰富的打字员进行受控实验 [27, 91, 112]。*emg2qwerty* 数据集 [92] 还对参与者进行了筛选，以确保他们能够不看键盘打字，并且至少 90% 的时间都能遵循正确的手指到按键映射。这样的连贯性支持在基于表面肌电的打字研究中继续采用标准化映射。为验证我们的方法处理用户差异的能力，我们还在评估期间对手指到按键映射施加了受控扰动，以测试其鲁棒性（见 §5.2.2）。结果表明，即使用户偏离标准映射，我们的方法仍能保持出色表现，证实其对个人打字风格具有鲁棒性。

表 2. **字母到手指映射**

左手			右手		
Key	手指	#	Key	手指	#
a	左手小指	0	n	右手食指	6
b	左手中指	3	o	右手无名指	8
c	中指（左）	2	p	小指（右）	9
d	中指（左）	2	h	食指（右）	6
e	中指（左）	2	u	食指（右）	7
f	左手食指	3	i	右手中指	7
g	左手食指	3	j	右手食指	6
q	左手小指	0	k	右手中指	7
r	左手食指	3	l	右手无名指	8
s	左手食指	1	y	右手食指	6
t	左手食指	3	m	右手食指	6
v	左手食指	3			
w	左手无名指	1			
x	左手无名指	1			
z	左手小指	0			



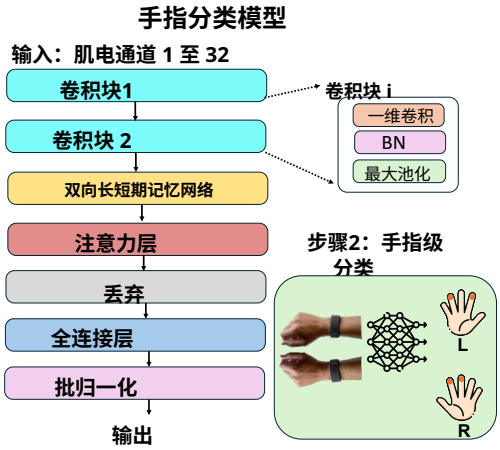


图 5.**手指分类网络**。用于根据多通道表面肌电预测手指标签的 CNN–BiLSTM–Attention 架构。

4.2 手指级分类

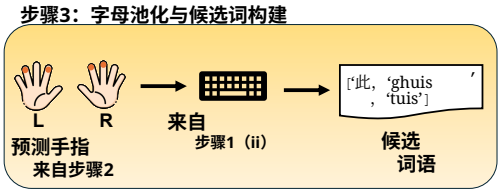
给定提取出的（表面肌电片段、手指标签）对，我们训练一个模型，使其能够直接根据原始多通道表面肌电输入，预测每个肌肉信号由哪根手指产生。通过将任务设定在手指级而非字母级，我们将分类空间从 26 个字母缩减为 8 根手指，从而更贴合打字时肌肉激活的实际情况。这种设计有助于模型应对字母级类别不平衡以及字母之间重叠的肌肉活动等问题，使其更稳健，也更容易泛化到新用户。

手指级分类是一个监督学习问题，包含两个目标：（i）捕捉多个表面肌电通道之间的短期相关性；（ii）建模双向时间依赖关系，因为肌肉激活通常同时包含预期信号和残余信号。我们使用卷积神经网络学习局部时空特征，并使用 LSTM 跟踪两个时间方向上的较长期依赖关系 [31, 59, 83]。基于这些优势，我们的模型将卷积块与双向长短期记忆层依次堆叠。我们还引入注意力机制，以突出最具信息量的时间步，并通过全连接层优化表示，从而预测正确的手指类别。简而言之，我们的架构结合了卷积神经网络、双向长短期记忆网络和注意力模块，并通过全连接层完成最终分类（图5）。

模型接收输入序列 $\mathbf{X} \in \mathbb{R}^{C\times T}$，其中 C 为电极通道数，T 为时间步数，并输出预测手指标签 $y \in \{0, \ldots\ldots, 7\}$。该架构依次经过四个阶段：

- **卷积特征提取**。卷积层通过 $\mathbf{H}^l = \sigma(\mathbf{W}^l * \mathbf{H}^{l-1} + \mathbf{b}^l)$ 对 $\mathbf{H}^0 = \mathbf{X}$ 执行卷积，其中 $*$ 表示卷积运算，σ 表示 ReLU 激活，而 $\mathbf{W}^l, \mathbf{b}^l$ 是可训练卷积核和偏置。每个卷积滤波器都会学习检测表面肌电波形中的局部时序特征。最大池化将时间维度从 T 降至 T'。因此，该阶段编码了与运动单位募集相关的短期依赖和频率特异性能量变化。

表 3. **手指到字母的映射和词语候选池化示例**。



字母	预测手指	字母池	候选词
t	3	b f g r t v	
h	6	h j m n u y	['ghis'（这个）、'this'（这个）、'tuis']
i	7	i k	
s	1	s w x	

• **使用双向长短期记忆网络进行时序建模**。 卷积特征 $\mathbf{H}^L \in \mathbb{R}^{F\times T}$，（其中 F 是学习到的特征图数量）被输入双向长短期记忆网络，以建模长程时间依赖关系。前向更新 $\overrightarrow{\mathbf{h}}_t = \mathrm{LSTM}_f(\overrightarrow{\mathbf{h}}_{t-1}, \mathbf{H}_t^L)$ 捕捉过去的上下文，而后向更新 $\overleftarrow{\mathbf{h}}_t = \mathrm{LSTM}_b(\overleftarrow{\mathbf{h}}_{t+1}, \mathbf{H}_t^L)$ $\overleftarrow{}$捕捉未来的上下文。拼接后的隐藏状态 $\mathbf{h}_t = [\overrightarrow{\mathbf{h}}_t; \overleftarrow{\mathbf{h}}_t]$编码了双向时序信息，同时保留前瞻性活动和残余活动。这对于表面肌电至关重要，因为肌肉激活和放松阶段往往会重叠，而且可以根据信号动态的前后变化推断运动意图。

• **基于注意力的聚焦**。 为了强调具有生理信息的片段，模型会学习随时间变化的注意力分布。每个双向长短期记忆网络的输出 $\mathbf{h}_t \in \mathbb{R}^{d_h}$ 表示一个 d_h维的隐藏特征向量，用于捕捉时间 t处的肌电上下文，包括收缩强度和转换动态。模型计算相关性分数 $e_t = \mathbf{v}^\top \tanh(\mathbf{W}_a\mathbf{h}_t)$，其中 $\mathbf{W}_a \in \mathbb{R}^{d_a\times d_h}$ 是一个可训练投影矩阵，用于将隐藏状态映射到注意力空间；$\mathbf{v} \in \mathbb{R}^{d_a}$ 是一个衡量其生理相关性的上下文向量。归一化权重 $\alpha_t = \exp(e_t)/\sum_{t'}\exp(e_{t'})$量化了各时间点的贡献，而上下文向量 $\mathbf{c} = \sum_t \alpha_t\mathbf{h}_t$ 则根据特征所学得的重要性对其进行加权汇总。

• 全连接层。上下文向量 $\mathbf{c} \in \mathbb{R}^{d_c}$ 经过带有 ReLU 的稠密层，计算得到 $\mathbf{z} = \phi(\mathbf{W}_f\mathbf{c} + \mathbf{b}_f)$，随后进行丢弃和批归一化。分类器输出为 $\hat{\mathbf{y}} = \mathrm{softmax}(\mathbf{W}_o\mathbf{z} + \mathbf{b}_o)$，其中 $y \in \{0, \ldots, 7\}$、$\hat{y}_k \in [0,1]$和 $\sum_{k=0}^{7}\hat{y}_k = 1$；对于 $K = 8$个类别，交叉熵损失为 $\mathrm{L} = -\sum_{k=0}^{K-1}\mathbb{1}[y=k]\log\hat{y}_k = -\log\hat{y}_y$。其中：$\phi$表示 ReLU；$\mathbf{z} \in \mathbb{R}^{d_z}$；$\mathbf{W}_f \in \mathbb{R}^{d_z\times d_c}$、$\mathbf{b}_f \in \mathbb{R}^{d_z}$；$\hat{\mathbf{y}} \in \mathbb{R}^K$及其 $K = 8$；$\mathbf{W}_o \in \mathbb{R}^{K\times d_z}$、$\mathbf{b}_o \in \mathbb{R}^K$；$y$表示真实类别索引；$\hat{y}_k$表示类别 k的预测概率；$\mathbb{1}[y=k]$ 表示指示函数。注意：如果直接使用 $\mathbf{c} = \sum_t \alpha_t\mathbf{h}_t$，则 $d_c = d_h$。

该模型结合卷积块、双向长短期记忆网络单元、注意力机制和全连接层，将局部空间特征与时间信息及上下文信息融合起来。它使分类任务与人们的打字方式相一致，使整个架构建立在真实的肌肉和手指运动之上。因此，该模型能够学习多通道表面肌电信号，并预测当前处于活动状态的手指。随后，这些手指预测结果会输入 MyoText 流水线的下一阶段，在该阶段，系统将手指映射为可能的字母，并开始重构词语。

4.3 字母池化与候选词构建

预测出每次击键由哪根手指完成后，我们会根据标准 QWERTY 盲打布局中手指与字母子集之间的固定映射，缩小字母搜索范围。表3 列出了一些手指到字母映射的示例。例如，手指编号为三的手指对应的字母包括 b、f、g、r、t和 v，人们通常用左手食指敲击这些字母。此步骤既缩小了解码搜索空间，也将生物力学知识融入预测过程，使模型更快速、更贴近实际。对于某个词的每个预测手指序列（例如 *this* $\rightarrow [3, 6, 7, 1]$），我们将每根手指映射到其对应的可能字母集合，从而生成一个字母池序列，列出该词每个位置上可能出现的字母。

表4. 基于手指的字母池所构建的候选词池示例。

句子	字母池序列	候选词
this	[b f g r t v ; h j m n u y ; i k ; s w x]	ghis、this、tuis
has	[h j m n u y ; a q z ; s w x]	has、haw、jaw、mas、max、naw、yas、yaw
two	[b f g r t v ; s w x ; l o]	two
部分	[p ; a q z ; b f g r t v ; b f g r t v ; s w x]	parrs，部分
合法的	[l o ; c d e ; b f g r t v ; i k ; b f g r t v ; i k ; h j m n u y ; a q z ; b f g r t v ; c d e]	合法的
load	[l o ; l o ; a q z ; c d e]	负载，渗出
经济的	[c d e ; c d e ; l o ; h j m n u y ; l o ; h j m n u y ; i k ; c d e]	经济的
女性	[s w x ; l o ; h j m n u y ; c d e ; h j m n u y]	soncy、女性、wouch

根据这个字母池序列，下一步是构建候选词池。其思路是先对所有字母池取笛卡尔积，生成所有可能的词语组合，然后使用大型英语语料库（如布朗语料库 [69], WordNet [73], 以及公开的英语词典 [26]）筛选这些词语。此筛选步骤会剔除无效或无意义的组合（例如 *“bbs”*、*“jnd”*），同时保留真实存在的英语词语。表 4 展示了最终候选词池的示例。

4.4 基于 Transformer 的解码进行句子预测

在我们框架的最后阶段，我们从候选词池中构建完整句子。第一阶段的手指分类会引入一定程度的不确定性，这种不确定性可能会传播到词语池中。为了解决这一问题，句子解码器应检查整个句子，而不是只关注相邻词语。因此，我们采用基于 Transformer 的语言模型，同时处理所有候选集合，并从每个集合中选择一个词语，以形成最终连贯的句子。

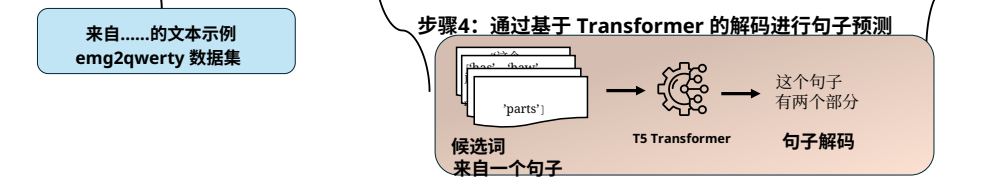
T5 Transformer。 我们使用谷歌在 2019 [77] 中提出的文本到文本迁移 Transformer（T5）。T5 在大规模清洁抓取语料库（C4）上进行了预训练，词汇量为 $|V| \approx 32,000$。T5模型包含一个编码器和一个解码器。编码器包含多层（例如，T5-Base 中有 12 层），每层都具有双向自注意力、带 ReLU 激活的两层前馈网络、层归一化和用于稳定训练的残差连接。解码器的结构与此类似，但增加了掩码自注意力，以避免查看未来词元，并包含编码器–解码器注意力，以实现基于上下文的生成 [77]。T5 是一种通用的序列到序列模型，可以针对摘要生成、问答和文本分类等任务进行微调 [6, 7, 61, 64, 106, 117]。它具备良好的可扩展性，并在 GLUE 和 SuperGLUE 等基准测试中表现出色 [109, 110]，因此被广泛应用于下游任务。

我们选择 T5 来执行句子解码任务，因为其编码器—解码器结构能够对整个句子进行全局推理。编码器从输入序列中捕获双向上下文，而自回归解码器则根据编码器的输出和此前生成的词元生成文本。这种设计使模型能够利用后续上下文来消除先前的不确定性。例如，在处理短语“two parts”时，区分“has”和“haw”就至关重要，这对于 MyoText 的准确解码不可或缺。

输入表示适配。 与输入本身已是词元序列的典型序列到序列任务不同，MyoText 的句子解码器接收结构化的候选词池（步骤 3 之后）。标准 T5 无法直接处理此类结构化输入，因为它需要扁平的文本序列。为解决这一问题，我们将候选词池序列 G_1，……，G_N（其中 N 是句子长度）转换为基于指令的提示

表5. **利用候选词池生成句子的示例。**

句子	候选池	数据表示	解码后的句子
这有两个部分	ghis this tuis; has haw jaw mas maw max naw yas yaw; two；parrs 部分；	请选择恰好 4 个词语：ghis this tuis <SEP> has haw jaw mas maw max naw yas yaw <SEP> two <SEP> parrs 部分 <SEP>	这有两个部分
合法负载 经济女性	合法的；负载渗出； 经济的；soncy女性 wouch；	请选择恰好 4 个词语：合法的 <SEP> 负载渗出 <SEP> 经济的 <SEP> soncy女性 wouch <SEP>	合法负载 经济女性
卡片解决问题 选民	咖啡馆 笼子 卡片 照料 分类 洞穴 达德 达雷 日期 戴夫 檐口；解决；问题； 钻孔工人 捡拾工人 罗杰斯 流浪者 搬运工 选民；	准确选择 4 个词语：咖啡馆 笼子 卡片 护理 凯特 山洞 达德 达尔 日期 戴夫 屋檐 <SEP> 解决 <SEP> 问题 <SEP> 蛀虫 加油工 罗杰斯 漫游者 搬运工 选民 <SEP>	笼子 解决问题 选民



字符串，并使用分隔符标记 <SEP>进行分隔，即：“恰好选择 N 个词语：G_1 <SEP> G_2 <SEP>……G_N <SEP>”。这会限制位置 n 只能从相应的候选词池中选择。例如，在表5第 3 列中，句子“this has two parts”及其候选词池 $G_1 = \{ghis, this, tuis\}$、$G_2 = \{has, haw\}$、$G_3 = \{two\}$、$G_4 = \{parrs, parts\}$会生成序列化输入 $S = \phi(G1, \ldots, G_4)$：“恰好选择 4 个词语：ghis this tuis <SEP> has haw <SEP> two <SEP> parrs parts <SEP>”。

通过约束束搜索生成句子。 我们将束搜索与微调后的 T5模型相结合，从候选词池中生成句子。T5 的编码器读取一次经过改编的序列化输入字符串，并生成上下文嵌入，以联合表示所有可能的词语。在解码过程中，T5 逐词生成句子。在每个位置 n，T5 只考虑相应候选池中的词语 $\mathcal{G}_n$作为有效选项。对于每个候选词，我们计算其在 T5 解码器下的概率，该概率以编码器上下文和已选词语为条件。随后，束搜索保留得分最高的B 个部分句子（其中 B 为束宽度），并用 $\mathcal{G}_n$中的所有候选词扩展每个部分句子，同时保留得分最高的后续序列。填入所有位置后，MyoText 选择总得分最高的句子。此过程使 T5 能够运用预训练中获得的全局语言知识，同时由束搜索系统地探索并评估多种词语组合，最终生成既流畅又符合生理解码约束的句子。表5 第4列列出了由候选词池生成的解码后句子示例。

5 评估

在本节中，我们将评估MyoText。我们将报告手指级分类准确率、句子级错误率（CER/WER）、对打字变异性和跨数据集偏移的鲁棒性，以及与现有表面肌电转文本系统的端到端比较。我们还将分析*emg2qwerty* 数据集，以评估信号质量。除非另有说明，所有实验均使用*emg2qwerty* 数据集。我们将回答以下研究问题（RQ）：

- **研究问题1：** MyoText 的模块化解码框架，从手指级分类到句子生成，性能如何?
- **研究问题2：** MyoText 对更长的自然文本和非典型打字模式的泛化能力如何?
- **研究问题3：** *emg2qwerty* 数据集是否具备足够的规模和击键级对齐能力，能否作为可靠的表面肌电转文本解码基准?
- **研究问题4：** MyoText 在表面肌电转文本系统的发展历程中处于什么位置？它为部署提供了哪些见解?

5.1 研究问题1：MyoText 的性能

我们使用*emg2qwerty* 数据集评估了 MyoText，因为它是目前发布的最多样化的表面肌电转文本数据集，发布于 2024 [92]。该数据集包含来自双侧前臂的同步表面肌电记录（每侧 16 个通道），以及带注释的击键事件。我们提取了结构化样例，其中包含用户ID、击键时序和多通道表面肌电片段。这确保每次击键都与相应的表面肌电窗口配对。数据预处理流程在 Google Colab 上使用 CPU资源运行，耗时约 15 小时。我们总共使用了 30名参与者的数据进行评估。对于手指分类，我们使用 Adam优化器训练神经网络模型 [50]（学习率为 10^{-3}，批量大小为 32），训练 100 个周期。为防止过拟合，我们结合使用了多种正则化策略，包括自适应学习率调度器（当验证损失趋于平稳时，ReduceLROnPlateau [100] 以 0.01 的衰减因子降低学习率）以及耐心值为 10 个周期的提前停止。对于句子解码，我们使用教师强制对预训练T5变换器 [77]进行微调。微调时批量大小为 32，学习率为 5×10^{-5}，最多训练 30 个周期，使用 Hugging Face Trainer API。所有模型均使用 `TensorFlow`实现，预训练模型从 `Hugging Face` [99]导入。训练和微调在 NVIDIA T4 或 NVIDIA L40 上进行，具体取决于 Google Colab 上的资源可用情况。

5.1.1 手指级分类。 MyoText 的一个关键设计选择是，先对手指激活（8 根活动手指）进行分类，还是直接预测字母（26 个类别）作为最终输出。基于 RNN 的方法很适合这个问题，因为 LSTM 等循环模型广泛用于时间序列解码，并能有效捕捉表面肌电信号中的时间依赖关系。我们的初步实验表明，将左臂和右臂信号分别取平均并合并为两个输入通道，会因空间细节丢失而显著降低性能（准确率为 40.3%）。相比之下，在相同的 CNN–LSTM架构中使用完整的 32 通道输入，将准确率提高到了 71.2%，凸显了高分辨率表面肌电表示的重要性。因此，我们训练了两个采用相同 CNN–BiLSTM–Attention架构（§4.2）的变体，两者仅在输出目标上有所不同：

- 模型 1（手指分类）：预测 8 个手指标签；平均准确率为 85.4%。
- 模型 2（字母分类）：直接预测 26 个字母；平均准确率达到 51.2%。

这一性能差距凸显了手指级解码的可行性。表面肌电信号会因手指活动而被激活，因此往往按照手指激活情况聚类，而不是按照单个字母聚类，因为同一手指敲击多个按键时会产生相似的肌肉模式。这种重叠使字母级分类变得更加模糊。图6展示了这种对比。模型 1 的混淆矩阵（图 6a）显示出明显的对角线占优现象，表明手指分离清晰

表6. **手指与字母变体准确率（双向长短期记忆网络+注意力）。**

Fold	模型 1（%）		模型 2（%）	
	CV	LOUO	CV	LOUO
1	82.71	85.52	43.85	29.76
2	82.01	85.41	47.52	28.46
3	88.66	80.98	51.2	30.84
4	89.30	78.84	54.87	25.52
5	84.61	77.27	58.55	26.73
均值 ± 标准差	**85.45 ± 3.36**	**81.4 ± 3.76**	**51.2 ± 5.81**	**28.5 ± 2.25**

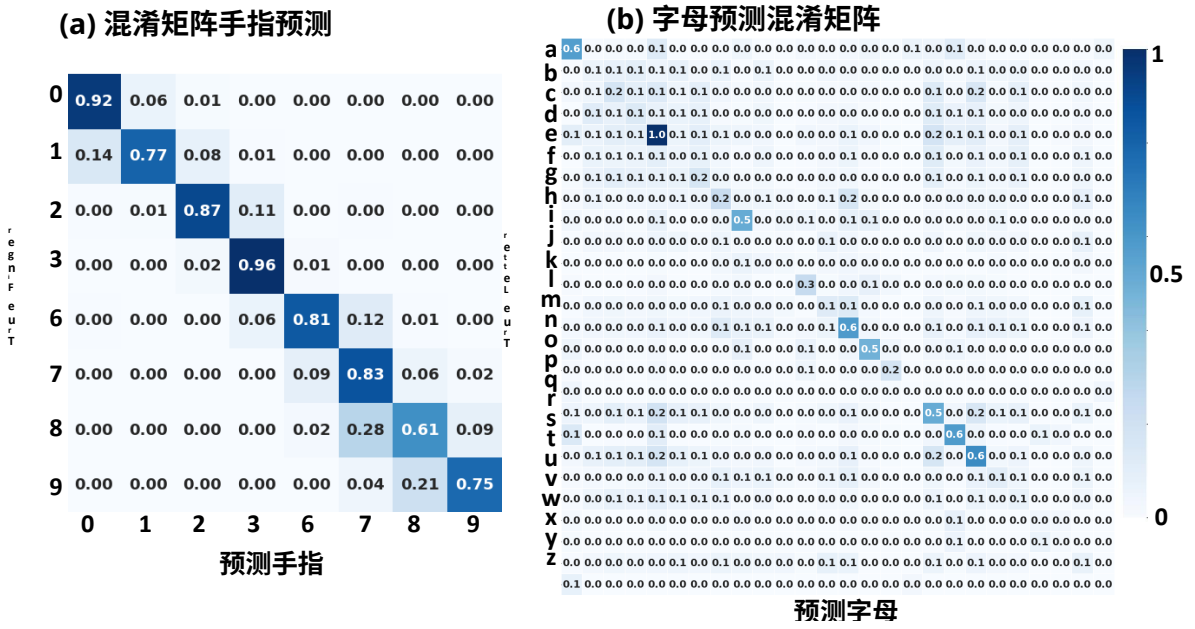


图 6. **手指与字母预测的混淆矩阵** (a) 模型 1 (b) 模型 2

。相比之下，模型 2 的混淆矩阵（图 6b）显示出大量误分类，尤其是在空间相邻的按键之间（或涉及相似肌肉活动的按键之间）。我们进一步使用两种验证方案评估模型的鲁棒性。如表 6所示，5 折交叉验证的平均准确率为 85.45% ± 3.36%，而在五名留出用户上进行留一用户交叉验证（LOUO）时，模型 1 的准确率达到 81.4% ± 3.76%。模型 2 的结果也显示在同一表格中。 尽管我们的模型遵循标准的 *QWERTY* 手指到按键映射（表 *2*），但用户之间可能存在特定偏差，例如用右手食指而非左手食指输入*“B”*。我们在 *§5.2.2*中考虑了这类情况，并且我们的下游 *Transformer*解码器能够利用句子级上下文有效解决这些不一致问题，从而确保模型能够适应手指预测错误和现实世界中的打字变异性。

5.1.2 句子级预测. 在手指分类之后，我们使用经过微调的 T5 Transformer，将候选词池（§4.3）解码为完整句子 [77]。句子重建性能采用词错误率（WER）和字符错误率（CER）进行评估，其定义见公式 1 和 2，其中 S， D，以及 I分别表示替换、删除和插入，而 N 表示参考词或字符的总数。

$$\mathrm{WER} = \frac{S_W + D_W + I_W}{N_W} \tag{1}$$

$$\mathrm{CER} = \frac{S_C + D_C + I_C}{N_C} \tag{2}$$

模型变体与结果. 我们评估了三个 T5-Base 变体，每个变体均针对候选项到句子的生成任务进行训练:

- **T5 模型 1：** 采用标准交叉熵损失的基线（$\mathrm{LM}_1 = \mathrm{L}_{CE}$）。
- **T5 模型 2：** 加入替换感知辅助损失（$\mathrm{LM}_2 = \mathrm{L}_{CE} + \lambda \mathrm{L}_{aux}$）。

- **T5 模型 3：** 采用模型 2 的损失，并加入权重衰减、标签平滑和提前停止。

图 7 显示，T5 模型 3 的表现最佳，词错误率为 6.5%，字符错误率为 5.4%。这些结果表明，解码器能够有效利用句子级上下文，克服手指级预测噪声和候选歧义。

错误分析。表7 列举了具有代表性的解码错误。大多数错误分为三类：（i）语义或语音相似的替换（例如，“卡片” → “笼子”）；（ii）遗漏专有名词（例如，“乌兹别克斯坦”“摩尔多瓦”）；（iii）过早终止或截断。

尽管平均词错误率和字符错误率较低，我们仍观察到约 7% 的标准差。这种差异主要归因于*emg2qwerty* 数据集的特性，其中包含较短且句法稀疏的短语。在这种情况下，即使只错一个词，也可能对词错误率产生不成比例的影响（例如，一个包含 4 个词的句子错了 1 个词，词错误率就是 25%）。尽管词汇准确率有所波动，模型通常仍能保留句子的长度和流畅度。尽管手指误分类或变异性带来了上游错误，

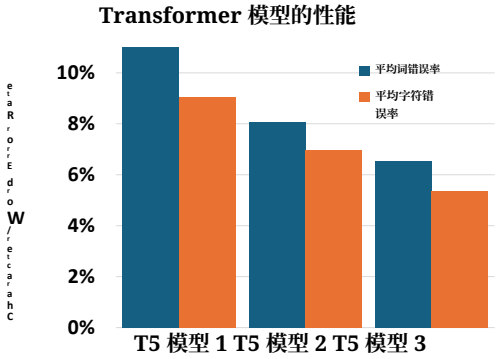


图 7。**T5 模型性能**。T5变体的字符错误率和词错误率。数值越低越好。

变异或变化 手指到按键的映射中，基于 Transformer 的解码器表现始终如一，因此支持了我们的 r
模块化框架 ork.

表 7. **emg2qwerty 数据中的参考句与预测句对比** **set.**

序号 编号	参考文本 #	预测*
1	敏捷的棕色狐狸**跳过**一只懒狗	敏捷的棕色狐狸 一只懒狗 ▲
2	法院 融资 **艾伦**姐妹	法院 融资 **所有** 姐妹 ★+
3	**卡片** 解决问题 选民	**笼子** 解决问题 选民 +
4	交易提取**某人的** 采购	交易提取 采购 ▲
5	奥林匹克**摩尔多瓦** 岩石 语料库	奥林匹克 岩石 语料库 ★ ▲
6	脊柱 保留**脱落的** 发烧	脊柱 保留**谢** 发烧 +
7	**乌兹别克斯坦** 雇用 犹他州 英国的	雇用 犹他州 英国的 ★ ▲
8	疑似 可能的 便士 **手持设备**	疑似 可能的 便士 ▲
9	横跨 **变更日志** 请愿书 上升	提出的请愿书 ▲
10	充电器 **丹佛** 反映 宽屏	充电器反映宽屏 ★▲
11	**星期二**立陶宛 独立的 包裹	**Thesean**立陶宛 独立的 包裹 ★+

+词语替换 ▲缺失词语 ★专有名词 / 语音混淆 错误类型：

***预测基于候选词构建（步骤 3），而候选词构建又由手指级分类（步骤2）生成；#句子的真实标签来自 emg2qwerty**

表8. **使用自然英语句子对参考句与预测句进行比较。**

参考文本	**预测**
女孩住在一个小镇	一个女孩住在小镇上 **缺失词"a"**
她每天都去农场	她每天都去农场
农场离她家很近	农场离她家很近
她喂奶牛	她喂奶牛
她给它们喂草和水	她给它们喂草和水
奶牛喜欢她	奶牛喜欢她
有一天，一头奶牛病了	有一天，一头奶牛病了
女孩告诉她的父亲	女孩告诉她的父亲
他打电话叫医生。	他打电话叫医生。
医生帮助了这头牛。	医生帮助了这头牛。
女孩很开心。	女孩很开心。

注：预测仅基于 T5 模型 3；此处没有手指级分类

研究问题1的主要结论

在 *emg2qwerty* 上，MyoText 通过与运动→人体工学→语言结构相契合的三阶段模块化框架，实现了 85.4 % 的手指分类准确率和 5.4% 的句子 CER：(1) 手指：表面肌电识别手指的效果优于识别字母（直接识别字母的准确率为 51.2%）。(2) 人体工学：QWERTY 手指—按键约束根据标准手指到字母的映射，缩小了每个位置的搜索范围。(3)句子生成：T5解码器消除了残余歧义，无需用户个体校准即可生成连贯的句子。

5.2 研究问题2：真实场景打字与句子变异性下的泛化

实际应用中的表面肌电到文本解码的一项关键要求，是能够突破训练数据集的限制实现泛化。尽管 *emg2qwerty* 是表面肌电到文本研究的重要首个基准，但其在语言多样性和打字变异性方面仍然有限。为评估更广泛的泛化能力，我们在三种设置下评估所提出的方法：(i) 在自然、段落规模的文本上进行句子级解码；(ii) 评估对偏离规范手指到按键映射的鲁棒性；(iii) 使用具有模拟打字变化的开源文本数据集进行跨域评估，因为目前不存在开源的表面肌电到句子数据集。

5.2.1 自然语言泛化。 为了评估句子解码器是否对 *emg2qwerty* 中简短、句法结构稀疏的句子过拟合（平均句长为 ≈4 词），我们使用 ChatGPT-4 生成的一段英语陈述句，对表现最佳的 Transformer 模型（T5 模型 3）进行了测试。每个句子的长度都相近。此项评估仅使用候选词池来考察语言模型的能力，不涉及表面肌电信号或手指分类。如表 8所示，句子解码器能够以近乎完美的准确度重建长篇句子。唯一的偏差是轻微的冠词省略（"a girl lives in a small town" → "a girl lives in small town"）。

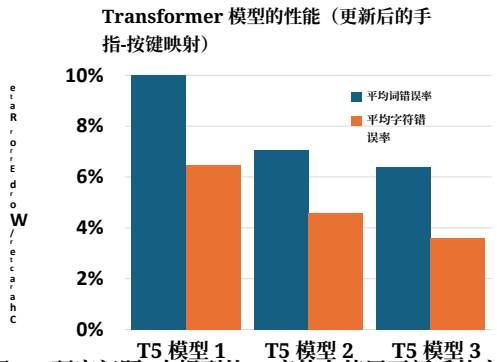


图8。 **研究问题1中提到的T5变体在使用更新后的字母集合时的解码性能**。CER和WER越低，表示性能越好。

表9。**扩展的手指到字母映射**。 标准 QWERTY 字母池通过添加额外字母进行扩展，以反映合理的打字偏差、模拟用户差异性并评估鲁棒性。

手指（左手）	更新后的字母集合	手指（右手）	更新后的字母集合
小指 (0)	[q、a、z]	小指 (9)	[p]
无名指 (1)	[w,s,x]	无名指 (8)	[o，l]
中指 (2)	[e、d、c、w、s、x]	中指 (7)	[i,k,o,l,p]
食指 (3)	[r,f,v,t,g,b]	食指 (6)	[y,h,n,u,j,m,t,g,b]

最终词错误率为 1.43%，字符错误率为 0.41%。这些结果证实，该模型具有良好的泛化能力，不会记忆固定的短语结构。

5.2.2 对打字变异性的鲁棒性。 虽然 QWERTY 键盘布局规定了标准的手指到按键映射，但现实中的打字往往与之不同，尤其是位置居中的按键。例如，字母b通常分配给左手食指（手指 3），但经常由右手食指（手指 6）敲击；同样，x、s和w 可能会从左手无名指（手指 1）转由左手中指（手指 2）敲击；而o、l和p 则常由右手中指（手指 7）而非右手无名指或小指敲击。为捕捉此类差异，我们根据观察到的打字行为，为候选字母池增补了合理的替代项（表 9）。这为解码引入了符合现实的模糊性，以模拟行为噪声。我们使用这些扩展后的字母池，在*emg2qwerty*上重新评估了所有 T5模型（模型1–3），其他流水线组件均保持不变。如图 8所示，性能基本相近，字符错误率和词错误率仅略有上升。这些发现证实，MyoText 不仅能在语言层面实现泛化，也能在行为层面实现泛化，在真实的手指到按键变异情况下保持解码保真度。这种鲁棒性使其能够部署于多种用户打字风格和输入条件下。

5.2.3 跨数据集泛化。 我们方法的一个关键应用场景是增强现实/虚拟现实系统，例如 Meta 的 Neural BandandRay- Ban Display [45], ，其中打字主要用于短篇通信（例如发短信）。为评估超越*emg2qwerty*的泛化能力，我们使用 Kaggle [10, 79], 上的两个公开语料库评估了句子解码器，分别记为 D1（聊天情感；741 个句子）和 D2（随机英语句子；584 个句子）。

表10。**跨数据集泛化**。

训练	Test	字符错误率 (%)
D1（741）	D2（584）	1.7
D2（584）	D1（300）	0.9

虽然这些数据集不包含表面肌电记录，但它们用于评估句子解码器在模拟输入变异性下的语言泛化能力。由表 9 中增强映射推导出的手指级扰动，使解码鲁棒性能够得到一致的基准评估。我们在这些条件下应用了表现最佳的模型（T5 模型 3），并进行了跨领域测试：在 D1 上训练、在 D2 上评估，随后反转这一设置，并从 D1 中随机保留一个子集用于测试（表 10）。此次评估有意排除 *emg2qwerty*，因为该数据集主要用于验证硬件层面的交互，因此大多包含较短的（四词）序列，句法结构有限，适用于简单的统计基线（例如 *n*元组）。相比之下，D1 和 D2 包含更长且语法连贯的句子，为跨领域解码提供了更具挑战性和代表性的测试环境。在这些条件下，尽管存在语言变异性，句子解码器仍保持较低的字符错误率（$<2\%$），表明基于T5的解码器具有良好的鲁棒性，适合部署于受控实验室环境之外。

研究问题2的主要结论

MyoText框架能够有效泛化到 *emg2qwerty* 数据集之外。它能够高保真地重建更长、更自然的句子，在非规范的手指到按键映射引入的打字变异性下仍保持鲁棒性，并且在未见过的开放领域数据集上进行测试时仍能维持性能。这些结果凸显了句子解码器对输入噪声和领域偏移的适应能力，彰显了其在多样化现实场景中部署的潜力。

5.3 研究问题 3：分析 *emg2qwerty* 数据集

我们对评估中使用的 *emg2qwerty* 数据集进行深入分析，以验证其质量和相关性。这是一个近期发布的大规模数据集，目前尚未得到广泛研究。该数据集包含参与者在 QWERTY键盘上打字时，通过腕部佩戴的阵列采集的表面肌电记录。语料库包含 108 名参与者的表面肌电数据及对应的击键数据，提示语取自词典和维基百科，以确保涵盖充分的词汇和句法内容。为评估该数据集是否适用于我们的研究，我们分析了数据集中带有高质量标签的用户的信号一致性和手指特定模式。

信号质量和手指层面的一致性。 为验证信号稳定性，我们计算了以每次击键为中心、时长为 2 秒的窗口内各手指的均方根（RMS）振幅分布（击键前 1 秒和击键后 1 秒）。该窗口涵盖预激活和放松阶段，因此其中位振幅低于峰值收缩振幅。如图 9所示，RMS 振幅中位数介于 10.6 至 11.0 μV 之间，标准差为 4.6 至 5.1 μV。四分位距保持较小（约为 5 至 15 μV），而较高幅值的点（最高可达 50 至 60 μV）对应的是按键按压，而非统计离群值。变异系数（CV = $\sigma_{\text{RMS}}/\mu_{\text{RMS}}$）介于 42% 至 47% 之间，表明各手指的尺度一致，且变异具有系统性而非随机性。

为量化信号一致性，我们计算了每根手指的信噪比（SNR），其定义为变异系数的倒数（SNR = $\mu_{\text{RMS}}/\sigma_{\text{RMS}}$）。如图10所示，所有手指的信号都稳定且可重复，SNR 值介于 2.12 和 2.37 之间（均值=2.3）。手指 3 的 SNR 最高（2.37），表明其肌肉激活具有很高的可重复性，这可能归因于其独立的运动控制 [38]。Finger9 的信噪比最低（2.12），这可能是由于其解剖结构复杂，且与邻近肌肉的信号重叠严重 [55]。

这些稳定的幅度分布和信噪比数值证实，该数据集捕捉到了稳定且符合生理特征的肌肉激活模式，信号质量足以支持可靠的击键级解码。

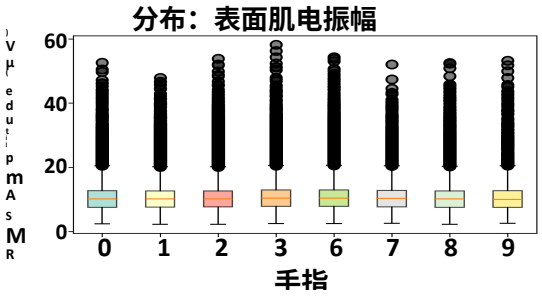


图 9。**按手指类别划分的表面肌电振幅**。 根据以每次按键为中心的 2 秒窗口计算。高于 40–50 μV 的点对应按键按压，而非离群值。

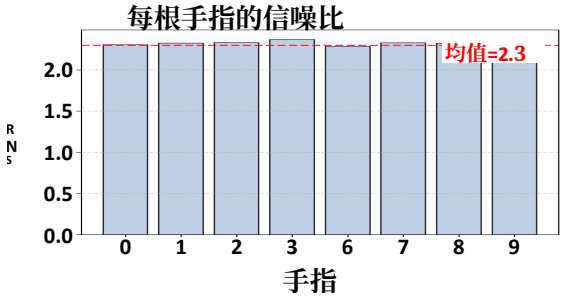


图 10。**类别内均方根稳定性**（μ/σ）。按手指计算 2 秒窗口内的均值与标准差之比；数值集中在 2.3 附近，表明信号质量稳定。

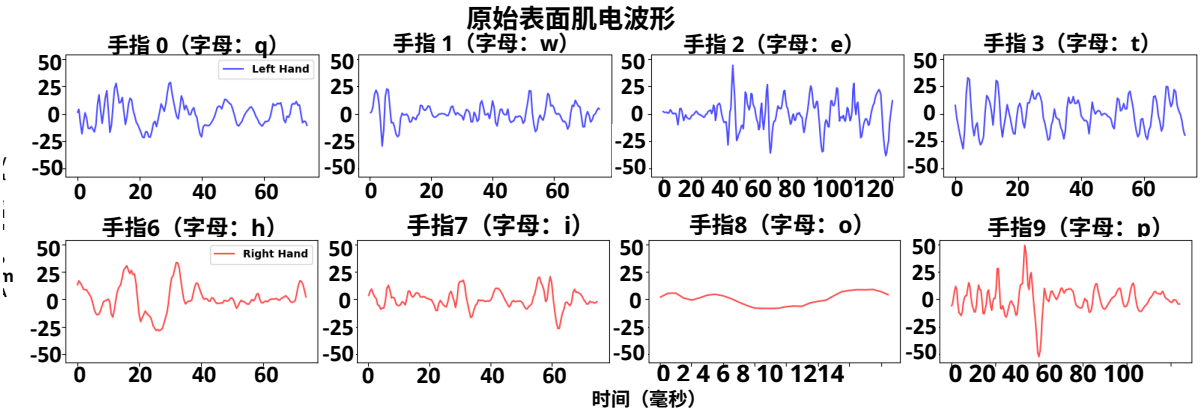


图 11. **手指特异性表面肌电波形**。平均单次按压波形显示，不同手指具有不同的时间形态，这支持采用考虑时间信息的解码方法。

手指特异性波形形态。 除了总体幅度指标之外，我们还考察不同手指是否会产生可区分的时间模式。图 11 展示了手指 0–9 单次按键的原始表面肌电。波形在时长（15–110 ms）、峰值幅度（±20-50 μV）和频率成分方面呈现出明显的手指特异性差异。手指 8（'o'）表现为短暂、平滑的激活（~15 ms， ±20 μV），而手指 2（'e'）则呈现持续的高频活动（~110ms）。手指 9（'p'）的幅度最高（±50 μV），手指 6（'h'）则表现出密集的瞬变。这些独特的时间特征表明手指级可区分性，并促使我们采用能够捕捉波形形态而非仅关注幅度的时序建模方法。这些独特的时间特征直观地体现了手指级可区分性，并促使我们采用能够捕捉波形形态而非仅关注幅度的时序建模方法。

研究问题3的关键结论

emg2qwerty 数据集呈现出稳定的信号质量，并且在幅度和时间域上均具有明显的手指特定模式。这种稳定性和可区分性在不同用户间均有所体现，使其适合用于评估有监督的击键解码方法。

5.4 研究问题4：MyoText 在表面肌电到文本系统及现实场景部署中的定位

为评估 MyoText 的实际效用，我们将其置于既有表面肌电转文本文献的背景下，并考察其模块化设计如何支持其在当代可穿戴平台上的部署。

现有的表面肌电转文本系统方法。 表11 总结了文献中报告的具有代表性的表面肌电转文本系统。虽然这些研究在数据集、词汇量和任务定义方面有所不同，但总体上展现了基于肌电的文本系统的发展，从手势驱动界面到紧凑型可穿戴设备上的句子级解码。在这一不断发展的领域中，MyoText 扩大了评估规模：MyoText 和*emg2qwerty* 基线均在 ≥30名参与者中进行了测试。与早期研究中常见的小规模评估相比，更大的队列有助于深入了解泛化能力和用户适应。在*emg2qwerty* 数据集上，使用腕戴式表面肌电设备和相同的30名用户配置时，MyoText 的字符错误率（CER）为5.4%，而 [92]所用方法的 CER 为 64%（即字符准确率为36%，因为CER = 1 – 准确率）——两者条件相同（表11）。这一

表 11. **表面肌电转文本领域**。 文献中报告的代表性系统。指标按原文报告：准确率（Acc）表示正确预测字符所占的比例，而字符错误率（CER）表示错误预测字符所占的比例（CER = 1 – Acc）。

系统	用户	通道	设备	Task	指标
Sharma 等人 [90]	10	2	BIOPAC	7 个按键	92.7% 准确率
Fu 等人 [37]	–	8	Myo 腕带	T9 手势（9 个按键）	85.9% 准确率
Crouch 等人 [20]‡	14.8 万个按键†	32	双袖套	播客转录（完整词汇表）	91% 准确率
Wang 等人 [112]	–	32	16×16 频带	30 个符号（26 个字母 + 空格/退格/回车/句点)	3.8% 字符错误率
Sivakumar 等人 [92]†	108	32	腕带	完整句子（26 个按键，~4 每句单词数）	52% 字符错误率
Sivakumar 等人 [92]†	30*	32	腕带	整句解码（26 按键， ~4 词语/句子）	64% 字符错误率（36% 准确率）
MyoText（我们的方案） †	**30**	**32**	**腕带**	**完整句子（26 个按键，~4 每句单词数）**	**5.4% 字符错误率（94.6% 准确率）**

注： 用户队列按原文报告（— = 未说明）。 † 表示在相同数据集和硬件上进行评估（*emg2qwerty*，32 通道腕带）。 * 表示在 MyoText 使用的 30名用户子集上重新实现并重新评估。‡ 表示报告的是击键数，而非参与者人数。

准确率水平与我们在研究问题1（§5.1）中的发现一致，字母分类达到了相近的准确率范围。更广泛文献中报告的结果涵盖了多种系统设计和评估设置。例如，Wang 等人 [112] 在30符号集上报告了3.8%的 CER，Crouch 等人 [20] 则使用双前臂袖套进行播客转录，达到91%的准确率。这些研究共同勾勒出基于表面肌电的解码系统在不同词汇量规模、数据集和设备配置下的当前性能范围。MyoText 的 CER 降低体现了其模块化框架的有效性；该框架将生理手指运动分类器与基于Transformer的句子解码器相结合。通过将生理信号解码与语言推理分离，该方法能够推广至不同用户。

部署可行性。 除了准确率之外，表面肌电转文本系统的实际可行性还取决于计算效率，以及能否在可穿戴设备严格的功耗和内存限制下实时运行。将手指动作解码与文本生成整合在一起的单体式端到端模型，所需的计算资源和功耗将超出可穿戴设备通常的预算。相反，纯云端方案会带来隐私风险、网络依赖和延迟 [35, 51]。为应对这些挑战，MyoText采用分离式执行设计：手指分类器可在可穿戴设备本地运行，而基于Transformer的句子解码器则可在配对的手机或边缘处理器上运行。对MyoText两个主要组件的性能分析表明，步骤2（手指分类器）每个样本的执行时间为 ~60毫秒，仅需 ≈12百万次操作和 < 12MB内存，完全满足当代可穿戴设备的交互阈值 [63, 85, 112]。现代边缘神经处理单元可提供38–80 TOPS（Apple M4、Snapdragon X2 Elite） [4, 44, 65, 74], , 足以支持该模块在设备端运行。步骤4（Transformer解码器）的计算量更大，但移动级硬件能够高效处理。在NVIDIA L4 GPU上，推理速度为每个Token 18.851毫秒。量化、内核融合和注意力缓存方面的进展，如今已能让十亿参数级Transformer在移动神经处理单元上运行 [12, 41, 66]。商业部署（如Apple ANE、谷歌端侧大语言模型）已证明，无需依赖云端即可实现低延迟文本生成。

综合来看，这些结果表明，MyoText能够部署在当前硬件上：手指分类器在资源受限的可穿戴设备上本地运行，而语言模型则在配对的高资源设备（如智能手机或边缘处理器）上运行。

研究问题4的关键结论

MyoText 在实现较高准确率（5.4%的字符错误率）的同时，还具备跨用户泛化和基于生理机制的架构等关键优势。手指分类器能够在腕上计算资源预算内运行，Transformer解码器则可在配对的智能手机上高效运行，从而使当前硬件能够实现实用且保护隐私的sEMG文本输入。

6 讨论与未来工作

MyoText 首次通过明确建模肌肉激活与打字行为之间的关系，提出了基于生理机制的基于sEMG的文本解码方法。不同于以往将肌电信号直接映射到字母的端到端系统，MyoText 首先预测活跃的手指——这一特征与EMG活动直接相关——然后再调用基于 Transformer 的语言模型进行句子解码。这种分层设计显著缩小了可能字母的有效搜索空间，使语言模型能够在定义明确且具有上下文感知能力的约束下运行。最终，该解码框架将生理学可解释性与语言流畅性相结合，在保持适用于可穿戴部署的计算效率的同时，实现了较低的字符错误率。

然而，我们的研究结果仍有进一步加强的空间。首先，尽管评估仅使用了一个数据集，但该数据集是目前规模最大的 sEMG 打字资源，包含 108 名参与者。受计算资源限制，我们在 30 名用户的数据上评估了模型；这一样本规模较为充足，也大于许多既往肌电打字研究所采用的规模。模块化框架为这一样本规模提供了合理依据，而 LOUO 验证也证实了模型具有稳定的跨用户泛化能力。我们对 30 名用户进行的队列评估采用了明确的 LOUO 评估，其表现与先前的表面肌电到文本研究相当或更优（见表 11）。未来，我们将对完整的参与者群体进行评估，以进一步验证泛化能力。其次，本研究侧重于验证 MyoText 框架的核心创新，而非端到端部署。未来，若能结合实时 sEMG 采集和设备端处理进行实现，将有助于深入了解实际可穿戴部署中的系统延迟和功耗。第三，尽管我们采用了标准化的 QWERTY 手指-按键映射，但个人打字策略存在显著差异。我们的实验纳入了噪声模型以模拟这些偏差，而 *emg2qwerty* 数据集中的受训打字员比普通用户表现出更一致的模式。全面调查群体层面的手指-按键偏好，将有助于制定更加个性化的解码策略。MyoText 为免手持文本输入奠定了可行的基础；未来的挑战在于扩大数据规模、优化模型，并在受控环境之外验证系统。MyoText 的灵活性意味着它具有广泛的应用前景，包括手势界面、面向运动障碍人士的辅助技术，以及噪声敏感环境中的无声通信，可为新一代基于生物信号的人机交互提供蓝图。

7 结论

我们介绍了 MyoText，这是一种基于生理机制的模块化表面肌电到文本解码框架，融合了手指级分类、符合人体工学约束的字母映射和基于 Transformer 的句子生成模型。与以往的单体式系统或基于手势的系统不同，MyoText 建立了一种结构化的运动—语言层级，反映了打字的神经肌肉过程。借助 *emg2qwerty* 数据集，MyoText 的手指分类准确率达到 85.4%，字符错误率（CER）为 5.4%，展现出跨用户的可靠解码性能。这项工作为开发可扩展的无键盘通信接口迈出了一步，

推动神经输入在 AR/VR 环境、辅助技术和泛在计算平台中的实际应用。

参考文本

[1] Bashaer A Al Abdullah、Ghada A Amoudi 和 Hanan S Alghamdi。2024年。《手语识别的进展：全面综述与未来展望》。IEEE Access 12（2024年），128871–128895。[2] Muhammad Al-Ayyad、Hamza Abu Owida、Roberto De Fazio、Bassam Al-Naami 和 Paolo Visconti。2023年。《康复中的肌电监测系统：临床应用、可穿戴设备和信号采集方法综述》。《电子学》12，7（2023年），1520。[3] Christoph Amma、Thomas Krings、Jonas Böer 和 Tanja Schultz。2015年。《利用高密度肌电图推进肌肉-计算机接口》。载于第33届美国计算机协会人机交互系统年度会议论文集。929–938。[4] 苹果公司。2024年。《Apple推出M4芯片》。Apple Newsroom。https://www.apple.com/newsroom/2024/05/apple-introduces-m4-chip/ [5] P Artemiadis。2012年。《基于肌电的机器人控制接口：过去、现在与未来》。《机器人与自动化进展》1，2（2012年），1–3。[6] Stefan Costin Atitienei 和 Anca Marginean。2024年。《探索T5的极限：面向推理、摘要生成、情感分析和比喻性语言理解的多任务微调》。载于2024年IEEE第20届国际智能计算机通信与处理会议（ICCP）。IEEE，01–08。[7] Halim Wildan Awalurahman 和 Indra Budi。2024年。《多项选择题中的自动干扰项生成：系统性文献综述》。PeerJ Computer Science 10（2024年），e2441。[8] Cord Bolte。2025年。《十指打字快速入门 - TypingAcademy》。https://www.typing.academy/10-finger-typing [在线；访问日期：2025-04-14]。[9] Dextina Alana Booker。2020年。《时尚与人体手势控制的未来：面向手语使用者的可穿戴通信设备探索》。博士学位论文。麻省理工学院。[10] ChatSent12：在线，2025年。《Chat 情感数据集》。https://www.kaggle.com/datasets/nursyahrina/chat-sentiment-dataset [在线；访问日期：2025-09-26]。[11] Baibhab Chatterjee、Pedram Mohseni 和 Shreyas Sen。2023年。《人体互联网的生物电子传感器节点》。《生物医学工程年度评论》25，1（2023年），101–129。[12] Kevin Chen。2025年。《在手机上运行70亿参数 Transformer：量化如何突破极限》。AI in Plain English。https://ai.plainenglish.io/running-a-7b-transformer-on-your-phone-how-quantization-is-pushing-the-limits-e70b3307595d [13] Karen B. Chen、Kevin Ponto、Ross D. Tredinnick 和 Robert G. Radwin。2015年。《虚拟用力：利用手势和肌肉活动在虚拟现实中唤起施力感》。《人因》57，4（2015年），658–673。数字对象标识符：10.1177/0018720814562231 arXiv：https://doi.org/10.1177/0018720814562231 PMID：25977324。[14] Yongming Chen、Haihong Zhang、Chuanchu Wang、Kai Keng Ang、Soon Huat Ng、Huiwen Jin 和 Zhiping Lin。2023年。《一种用于基于肌电图预测自主运动意图的分层动态贝叶斯学习网络》。《科学报告》13，1（2023年），4730。[15] Manisha Choudhary、Monika Lokhande、Rushikesh Borse 和 Avinash Bhute。2022年。《一种利用肌电信号辅助瘫痪患者的机器学习方法》。载于《社会计算高级数据挖掘工具与方法》。爱思唯尔，107–125。[16] Meghna Roy Chowdhury、Archisman Ghosh、Md Faizul Bari 和 Shreyas Sen。2024年。《利用分布式神经网络发挥超低功耗可穿戴设备的优势》。载于2024年IEEE第20届国际人体传感器网络会议（BSN）。IEEE，1–4。[17] Rubana H Chowdhury、Mamun BI Reaz、Mohd Alauddin Bin Mohd Ali、Ashrif AA Bakar、Kalaivani Chellappan 和 Tae G Chang。2013年。《表面肌电信号处理与分类技术》。《传感器》13，9（2013年），12431–12466。[18] Andrea Cimolato、Josephus JM Driessen、Leonardo S Mattos、Elena De Momi、Matteo Laffranchi 和 Lorenzo De Michieli。2022年。《下肢假肢中的肌电驱动控制：基于主题的系统性综述》。《神经工程与康复杂志》19，1（2022年），43。[19] Jan Pieter Clarys 和 Jan Cabri。1993年。《肌电图与体育运动研究：综述》。《体育科学杂志》11，5（1993年），379–448。[20] Michael S Crouch、Mingde Zheng 和 Michael S Eggleston。2021年。《通过表面肌电图识别自然打字》。arXiv预印本 arXiv:2109.10743（2021年）。[21] Carlo J De Luca。2002年。《表面肌电图：检测与记录》。DelSys Incorporated 10，2（2002年），1–10。[22] Riccardo Di Giminiani、Marco Cardinale、Marco Ferrari 和 Valentina Quaresima。2020年。《验证织物式大腿可穿戴肌电传感器和血氧测量在力量与耐力训练期间监测股四头肌活动的效果》。《传感器》20，17（2020年），4664。[23] Denise K Donica、Peter Giroux 和 Amber Faust。2018年。《键盘输入教学：提高小学生键盘输入技能的技巧比较》。《职业治疗、学校与早期干预杂志》11，4（2018年），396–410。[24] Yu Du、Wenguang Jin、Wentao Wei、Yu Hu 和 Weidong Geng。2017年。《通过深度领域自适应增强基于表面肌电的跨会话手势识别》。《传感器》17，3（2017年），458。

[25] Anany Dwivedi、Yongje Kwon 和 Minas Liarokapis。2020年。《虚拟现实中操作动作的基于肌电图解码：迈向沉浸式界面》。载于2020年IEEE系统、人和控制论国际会议（SMC）。IEEE，3296–3303。[26] dwyl。2025年。english-words/words_alpha.txt, 位于master · dwyl/english-words · GitHub。https://github.com/dwyl/english-words/blob/ master/words_alpha.txt [在线；访问日期：2025-04-28]。[27] Jonathan Eby、Moshe Beutel、David Koivisto、Idan Achituve、Ethan Fetaya 和 José Zariffa。2025年。《面向神经技术人机界面的肌电打字手势分类数据集》。《Scientific Data》12，1（2025年），440。[28] Oussama EL GHOUL、Mohamed JEMNI 等。2024年。《基于肌电图的通信界面的设计与实现：增强与聋人的互动》。《信息科学杂志》23，2（2024年），1–20。[29] Electrom90：在线，2023年。《肌电图（EMG）》。|约翰·霍普金斯医学。https://www.hopkinsmedicine.org/health/treatment- tests-and-therapies/electromyography-emg [在线；访问日期：2025-04-10]。[30] facebookresearch。2025年。GitHub - facebookresearch/emg2qwerty：一个带有基线方法的表面肌电盲打数据集。https://arxiv.org/abs/2410.20081。https://github.com/facebookresearch/emg2qwerty [在线；访问日期：2025-04-29]。[31] Yongxin Fan、Qian Tang、Yangming Guo 和 Yifei Wei。2024年。《BiLSTM-MLAM：一种基于双向长短期记忆网络和局部注意力机制的传感器数据多尺度时间序列预测模型》。《传感器》24，12（2024年），3962。[32] Yongfei Feng、Mingwei Zhong、Xusheng Wang、Hao Lu、Hongbo Wang、Pengcheng Liu 和 Luige Vladareanu。2021年。《基于表面肌电传感器的气动康复手套主动触发控制》。PeerJ Computer Science 7（2021年），e448。[33] ELISABETH FERNSTRÖM、Mats O Ericson 和 HANS MALKER。1994年。《使用打字机和键盘时的肌电活动》。《人体工学》37，3（1994年），477–484。[34] T Finni、Min Hu、P Kettunen、T Vilavuo 和 S Cheng。2007年。《测量嵌入服装中的纺织电极所采集的肌电活动》。《生理测量》28，11（2007年），1405。[35] Fiveable。2024年。《边缘设备的功耗挑战》。Fiveable 学习指南。https://fiveable.me/edge-ai-and-computing/unit- 8/power-consumption-challenges-edge-devices/study-guide/kU7ZtkGCmjmI4oOW [36] Adam Freed、Adrian DC Chan、Edward D Lemaire 和 Avi Parush。2011年。《用于康复的可穿戴肌电分析（WEAR）——临床步态分析中的表面肌电图》。载于2011年IEEE医学测量与应用国际研讨会。IEEE，601–604。[37] Zongkai Fu、Huiyong Li、Zhenchao Ouyang、Xuefeng Liu 和 Jianwei Niu。2020年。《借助肌电键盘随处打字：一种基于MYO臂带的新型人机交互工具》。载于国际算法与并行处理架构会议。Springer，247–261。[38] Andrew J Fuglevand、Vaughan G Macefield 和 Brenda Bigland-Ritchie。1999年。《控制人手手指的肌肉中运动单位的力-频率与疲劳特性》。《神经生理学杂志》81，4（1999年），1718–1729。[39] Donald R Gentner。1983年。《誊写打字中的击键时序》。载于《熟练打字的认知方面》。Springer，95–120。[40] Vidhi Gohel 和 Ninad Mehendale。2020年。《肌电信号采集与处理综述》。《生物物理学评论》12，6（2020年），1361–1367。[41] Dibakar Gope、David Mansell 等。2025年。《面向Arm CPU上大语言模型推理的高度优化内核与细粒度码本》。arXiv预印本 arXiv:2501.00032（2025年）。[42] Reed D Gurchiek、Nicole Donahue、Niccolo M Fiorentino 和 Ryan S McGinnis。2021年。《仅使用可穿戴设备分析肌肉和关节力学：一种肌电驱动方法》。《IEEE生物医学工程汇刊》69，2（2021年），580–589。[43] Awni Hannun、Ann Lee、Qiantong Xu 和 Ronan Collobert。2019年。《利用时间深度可分离卷积进行序列到序列语音识别》。arXiv预印本 arXiv:1904.02619（2019年）。[44] Adam Hartley。2024年。《TOPS详解——Apple全新M4 iPad芯片究竟有多强大？》TechRadar。https://www.techradar.com/computing/artificial-intelligence/tops-explained-exactly-how-powerful-is-apples-new-m4-ipad-chip [45] heathera。2025年。《Meta Ray-Ban Display：配备肌电腕带的AI眼镜》。https://about.fb.com/news/2025/09/meta-ray-ban-display-ai-glasses-emg-wristband/ [在线；访问日期：2025-09-29]。[46] Matthias Janke 和 Lorenz Diener。2017年。《肌电转语音：直接从面部肌电信号生成语音》。《IEEE/ACM音频、语音与语言处理汇刊》25，12（2017年），2375–2385。[47] Szu-Chen Stan Jou、Tanja Schultz、Matthias Walliczek、Florian Kraft 和 Alex Waibel。2006年。《迈向使用表面肌电图进行连续语音识别》。载于Interspeech。573–576。[48] P Kaifosh 和 T Reardon。2024年。《一种用于人机交互的通用非侵入式神经运动界面》。bioRxiv（2024年）。[49] Victor Kartsch、Simone Benatti、Mattia Mancini、Michele Magno 和 Luca Benini。2018年。《由太阳能采集器供电的智能可穿戴腕带，用于基于肌电的手势识别》。载于2018年IEEE国际电路与系统研讨会（ISCAS）。IEEE，1–5。[50] Diederik P Kingma 和 Jimmy Ba。2014年。《Adam：一种随机优化方法》。arXiv预印本 arXiv:1412.6980（2014年）。[51] Kisaco Research。2024年。《边缘人工智能能效实用指南》。Kisaco Research。https://www.kisacoresearch.com/sites/default/ files/documents/halio_whitepaper_edge_ai_power_efficiency.pdf [52] Heli Koskimäki、Pekka Siirtola 和 Juha Röning。2017年。《Myogym：介绍一个使用Myo臂带采集、用于活动识别的开放式健身数据集》。载于2017年美国计算机协会普适与泛在计算国际联合会议论文集以及会议论文集

2017年美国计算机协会可穿戴计算机国际研讨会论文集。537–546。[53] Yi-Liang Kuo、Pei-San Wang、Po-Yen Ko、Kuo-Yuan Huang 和 Yi-Ju Tsai。2019年。实时姿势生物反馈对颈痛成人脊柱姿势、肌肉活动和感知疼痛严重程度的即时影响。《步态与姿势》67（2019年），187–193。[54] Meta Reality Labs。2021年。推出一种用于神经输入的腕戴式可穿戴设备。https://tech.fb.com/inside-facebook-reality-labs-wrist- based-interaction。访问于2025年4月。[55] JNAL Leijnse、PM Quesada 和 CW Spoor。2010年。手指指间关节耦合机制的运动学评估——变异性、屈伸差异、触发、锁定、天鹅颈畸形及人体测量学相关性。《生物力学杂志》43，12（2010年），2381–2393。[56] Alessandro Leone、Gabriele Rescio、Lucia Giampetruzzi 和 Pietro Siciliano。2019年。用于评估腿部肌肉收缩的智能肌电袜。载于《2019年IEEE国际测量与网络研讨会（M&N）》。IEEE，1–6。[57] Alessandro Leone、Gabriele Rescio 和 Pietro Siciliano。2017年。利用表面肌电图技术评估跌倒风险。载于《2017年国际工程、技术与创新会议（ICE/ITMC）》。IEEE，1092–1095。[58] Dongming Li、Linnan Huang、Yangrong Wen、Yueming Gao、Željka Lučev Vasić、Mario Cifrek 和 Min Du。2020年。基于有限元法分析用于局部肌肉疲劳评估的电阻抗肌描记术电极配置。IEEE Access 8（2020年），172233–172243。[59] Mengfang Li、Yuanyuan Jiang、Yanzhou Zhang 和 Haisheng Zhu。2023年。利用深度学习算法进行医学图像分析。Frontiers in Public Health 11（2023年），1273253。[60] Na Li、Rui Zhou、Bharath Krishna、Ashirbad Pradhan、Hyowon Lee、Jiayuan He 和 Ning Jiang。2024年。肌肉疲劳监测的非侵入式技术：全面综述。Comput. Surveys 56，9（2024年），1–40。[61] Jintao Ling 和 Muhammad Afzaal。2024年。在高等教育中使用预训练大语言模型自动生成问答对。Computers and Education: Artificial Intelligence 6（2024年），100252。[62] Simone Maragliulo、Pedro Filipe Alhais Lopes、Luís Bica Osório、Aníbal T De Almeida 和 Mahmoud Tavakoli。2019年。通过双通道可穿戴肌电系统实现足部手势识别。IEEE Sensors Journal 19，22（2019年），10187–10197。[63] Jean-Matthieu Maro 等。2024年。Helios：一种用于常开智能眼镜的超低功耗事件式手势识别技术。arXiv 预印本 arXiv:2407.05206（2024年）。[64] Alaa Marshan、Anwar Nais Almutairi、Athina Ioannou、David Bell、Asmat Monaghan 和 Mahir Arzoky。2024年。MedT5SQL：用于医疗领域文本到 SQL 转换的基于 Transformer 的大语言模型。Frontiers in Big Data 7（2024年），1371680。[65] Jon Martindale。2025年。Snapdragon X2 Elite 与 Snapdragon X Elite 对比：有哪些新变化。Tom’s Guide。https://www.tomsguide.com/computing/cpus/snapdragon-x2-elite-vs-snapdragon-x-elite-heres-whats-new [66] Meta AI。2024年。推出速度更快、内存占用更低的量化 Llama 模型。Meta AI Blog。https: //ai.meta.com/blog/meta-llama-quantized-lightweight-models/ [67] Kerry R Mills。2005年。肌电图基础。《Journal of Neurology, Neurosurgery & Psychiatry》76，增刊2（2005年），ii32–ii35。[68] Ninapro25:online 2025年。Ninapro。https://ninapro.hevs.ch/ [在线；访问于2025-04-14]。[69] NLTKSamp10:online 2025年。NLTK：语料库示例用法。https://www.nltk.org/howto/corpus.html [在线；访问于2025-04-28]。[70] Mina Nouredanesh、Sunil L Kukreja 和 James Tung。2016年。使用可穿戴肌电传感器检测代偿性平衡反应以评估跌倒风险。载于《2016年第38届IEEE医学与生物工程学会国际年会（EMBC）》。IEEE，1680–1683。[71] Jan Noyes。1983年。QWERTY键盘：综述。《International Journal of Man-Machine Studies》18，3（1983年），265–281。[72] OrionAIG74:online 2025年。Orion AI眼镜：AR眼镜技术的未来 | Meta。https://about.meta.com/realitylabs/orion/ [在线；访问于2025-04-08]。[73] Ted Pedersen、Siddharth Patwardhan、Jason Michelizzi 等。2004年。WordNet：概念相关性的相似度测量。载于 AAAI，第4卷。25–29。[74] Tom Petri。2025年。Snapdragon X2 Elite Extreme 在新基准测试中击败 Apple M4、Intel 和 AMD。Windows Central。https://www.w i n dowscentral.com/hardware/qualcomm/snapdragon-x2-elite-extreme-crushes-apple-m4-intel-and-amd-in-new-benchmarks [75] Carlos G Pinheiro Jr、Eduardo LM Naves、Pierre Pino、Etienne Losson、Adriano O Andrade 和 Guy Bourhis。2011年。面向严重运动障碍人士的替代性交流系统：综述。Biomedical Engineering Online 10，1（2011年），31。[76] M Raez、M Hussain 和 F Mohd-Yasin。2006年。肌电信号分析技术：检测、处理、分类与应用。检索于2018年10月16日。[77] Colin Raffel、Noam Shazeer、Adam Roberts、Katherine Lee、Sharan Narang、Michael Matena、Yanqi Zhou、Wei Li 和 Peter J Liu。2020年。探索统一文本到文本 Transformer 的迁移学习极限。Journal of Machine Learning Research 21，140（2020年），1–67。[78] Ismo Rakkolainen、Ahmed Farooq、Jari Kangas、Jaakko Hakulinen、Jussi Rantala、Markku Turunen 和 Roope Raisamo。2021年。扩展现实中多模态交互技术——范围综述。Multimodal Technologies and Interaction 5，12（2021年），81。

[79] Randomen9:online 2025年。随机英语句子。 https://www.kaggle.com/datasets/nikitricky/random-english-sentences [在线；访问于2025-09-26]。[80] Gundala Jhansi Rani、Mohammad Farukh Hashmi 和 Aditya Gupta。2023年。表面肌电图与人工智能在人体活动识别中的应用——方法、新兴趋势、应用、挑战及未来实施的系统综述。IEEE Access 11（2023年），105140–105169。[81] Seema Rawat、Somya Vats 和 Praveen Kumar。2016年。评估与探索 MYO 臂带。载于《2016年国际系统建模与研究趋势发展会议（SMART）》。IEEE，115–120。[82] Devon I Rubin。2019年。针极肌电图：基本概念。《临床神经病学手册》160（2019年），243–256。[83] Tara N Sainath、Oriol Vinyals、Andrew Senior 和 Haşim Sak。2015年。卷积、长短期记忆、全连接深度神经网络。载于《2015年IEEE声学、语音与信号处理国际会议（ICASSP）》。IEEE，4580–4584。[84] Christian Sax、Hannes Lau 和 Elaine Lawrence。2011年。Liquid Keyboard：适用于触摸屏和表面的符合人体工学、自适应 QWERTY键盘。feedback 64（2011年），19。[85] Md Abu Sayeed 和 Fatahia Nasrin。2023年。面向癫痫可穿戴医疗设备的低延迟、低功耗边缘计算平台。载于《2023年IEEE德克萨斯无线与微波电路及系统研讨会（WMCS）》。IEEE，1–4。[86] Ovishake Sen、Anna M Sheehan、Pranay R Raman、Kabir S Khara、Adam Khalifa 和 Baibhab Chatterjee。2023年。利用神经信号进行语音与手写检测的机器学习方法：综述。《传感器》23，12（2023年），5575。[87] Shreyas Sen 和 Arunashish Datta。2024年。以人为本的分布式可穿戴人工智能。载于《第61届美国计算机协会/IEEE设计自动化会议论文集》。1–4。[88] Shreyas Sen 和 Arunashish Datta。2024年。特邀论文：以人为本的分布式可穿戴人工智能。载于《第61届美国计算机协会/IEEE设计自动化会议论文集》（美国加利福尼亚州旧金山）（DAC '24 364，4 10 1145 3649329）。计算机协会，纽约，美国，文章页码。数字对象标识符：./3605573.[89] Mehran Shabanpour 、 Kasra Rad 、 Sadaf Khademi 和 Arash Mohammadi。2025年。MoEMba：一种基于 Mamba 的专家混合模型，用于高密度肌电的手势识别。arXiv 预印本 arXiv:2502.17457（2025年）。[90] Rajneesh Sharma 和 Amit Kukker。2017年。基于神经强化学习、利用前臂肌电信号识别打字按键。载于《第9届信号处理系统国际会议论文集》。225–229。[91] Guy G Simoneau、Richard W Marklin 和 Joseph E Berman。2003年。电脑键盘倾斜度对无肌肉骨骼疾病打字者腕部位置和前臂肌电图的影响。《物理治疗》83，9（2003年），816–830。[92] Viswanath Sivakumar、Jeffrey Seely、Alan Du、Sean Bittner、Adam Berenzweig、Anuoluwapo Bolarinwa、Alex Gramfort 和 Michael Mandel。2024年。emg2qwerty：一个用于表面肌电图盲打的大型数据集及其基线。《神经信息处理系统进展》37（2024年），91373–91389。[93] Andreia SP Sousa、Andreia Noites、Rui Vilarinho 和 Rubim Santos。2023年。肌电测量中电极-皮肤阻抗的长期变化。《传感器》23，20（2023年），8582。[94] LV Srinivas、R Shiva Shankar、N Deshai、K Sravani 和 V Maheswararao。2024年。促进包容性交流：可穿戴技术与 GloSign 的进展——实现无缝手语交互的基于手套的解决方案。载于《高级人工智能算法》。CRC出版社，492–501。[95] Erik Stålberg 和 Peter Dioszeghy。1991年。正常肌肉和神经肌肉疾病中的扫描肌电图。《脑电图与临床神经生理学/诱发电位分册》81，6（1991年），403–416。[96] Abdulhamit Subasi。2013年。使用粒子群优化的支持向量机对肌电信号进行分类，以诊断神经肌肉疾病。《Computers in Biology and Medicine》43，5（2013年），576–586。[97] S Sudarsan、IEEE学生会员和 E Chandra Sekaran。2012年。肌电控制假肢肢体的设计与开发。《Procedia Engineering》38（2012年），3547–3551。[98] Gabriel Synnaeve、Qiantong Xu、Jacob Kahn、Tatiana Likhomanenko、Edouard Grave、Vineel Pratap、Anuroop Sriram、Vitaliy Liptchinsky 和 Ronan Collobert。2019年。端到端自动语音识别：从监督学习到基于现代架构的半监督学习。arXiv 预印本 arXiv:1911.08460（2019年）。[99] T511:online 2025年。T5。https://huggingface.co/docs/transformers/en/model_doc/t5 [在线；访问于2025-04-29]。[100] Keras 团队。2025年。ReduceLROnPlateau。https://keras.io/api/callbacks/reduce_lr_on_plateau/ [在线；访问于2025-04-29]。[101] Shi-Liu Tian、Yu Liu、Li Li、Wei-Jie Fu 和 Chien-Hua Peng。2010年。机械肌动图在检测年龄相关性肌少症方面比肌电图更敏感。《生物力学杂志》43，3（2010年），551–556。[102] Dennis Tkach、He Huang 和 Todd A Kuiken。2010年。肌电模式识别时域特征稳定性的研究。《神经工程与康复杂志》7（2010年），1–13。[103] 维基媒体项目贡献者。2004年。盲打 - 维基百科。https://en.wikipedia.org/wiki/Touch_typing?utm_source=chatgpt.com [在线；访问于2025-09-10]。[104] Cinthya Lourdes Toledo-Peral、Gabriel Vega-Martinez、Jorge Airy Mercado-Gutiérrez、Gerardo Rodriguez-Reyes、Arturo Vera-Hernandez、Lorenzo Leija-Salas 和 Josefina Gutierrez-Martinez。2022年。将肌电图用作控制/生物反馈的康复应用中的虚拟/增强现实：系统性文献综述。《电子学》11，14（2022年），2271。

[105] TypingLe73:online 2025年。《打字课程》——《打字基础》第 2 部分：键盘上的手指分配。https://www.learn2type.com/ typingtest/lessons/typingkeyboards.cfm [在线；访问日期：2025-04-14]。[106] Hisao Usui 和 Kanako Komiya。2023年。使用日语T5将历史日语翻译为现代日语。载于《第三届自然语言处理与数字人文联合国际会议暨第八届乌拉尔语族计算语言学国际研讨会论文集》。27–35。[107] A Prasad Vinod 和 Chai Yee Da。2013 2013 7 98 102 。一种用于运动医学应用的集成式表面肌电数据采集系统。载于第届国际医学信息与通信技术研讨会（ISMICT）。IEEE，–。[108] Michael Wand 和 Tanja Schultz。2011年。基于肌电的语音识别：不依赖会话的语音识别。载于《生物信号》。295–300。[109] Alex Wang、Yada Pruksachatkun、Nikita Nangia、Amanpreet Singh、Julian Michael、Felix Hill、Omer Levy 和 Samuel Bowman。2019年。SuperGLUE：一个更具挑战性的通用语言理解系统基准。《神经信息处理系统进展》32（2019年）。[110] Alex Wang、Amanpreet Singh、Julian Michael、Felix Hill、Omer Levy 和 Samuel R Bowman。2018年。GLUE：一个用于自然语言理解的多任务基准与分析平台。arXiv 预印本 arXiv:1804.07461（2018年）。[111] Jinqiang Wang、Dianguo Cao、Yang Li、Jiashuai Wang 和 Yuqiang Wu。2022年。通过判别典型相关分析和自适应降维，利用表面肌电进行多用户动作识别。《神经机器人学前沿》16（2022年），997134。[112] Yi Wang、Youhao Wang、Ruilin Zhao、Yue Shi 和 Yingnan Bian。2024年。高效的基于肌电图的打字系统：迈向一种新颖的人机交互文本输入方法。载于2024年IEEE医学与生物工程学会第46届年度国际会议（EMBC）。IEEE，1–5。[113] Wearable26:online 2025年。《可穿戴技术》| Myontec。https://www.myontec.com/ [在线；访问日期：2025-04-08]。[114] WhereSho97:online 2025年。键盘上手指应该放在哪里？https://www.computerhope.com/issues/ch001346.htm [在线；访问日期：2025-04-14]。[115] Shumin Zhai 和 Per Ola Kristensson。2012年。单词手势键盘：重新构想键盘交互。《美国计算机协会通讯》55，9（2012年），91–101。[116] Mingrui Ray Zhang、Shumin Zhai 和 Jacob O Wobbrock。2022年。TypeAnywhere：一种面向普适计算的基于QWERTY的文本输入解决方案。载于《2022年CHI人因计算系统会议论文集》。1–16。[117] Qingqing Zhao、Yuhan Xia、Yunfei Long、Ge Xu 和 Jia Wang。2025年。将感知知识融入文本到文本迁移 Transformer，以增强情感分析。《信息处理与管理》62，1（2025年），103876。[118] Ping Zhou、Madeleine M Lowery、Kevin B Englehart、He Huang、Guanglin Li、Levi Hargrove、Julius PA Dewald 和 Todd A Kuiken。2007年。解码一种用于控制假肢的新型神经机器接口。《神经生理学杂志》98，5（2007年），2974–2982。