arXiv:2601.01658v1 [cs.NE] 4 Jan 2026

STEMNIST：脉冲触觉扩展 MNIST 神经形态数据集

AnubhabTripathi[1], **Li Gaishan**[1], **Zhengnan Fu**[1], **ChiaraBartolozzi**[2], **Bert E. Shi**[3] **和 ArindamBasu**[1, 4, *]

[1] 香港城市大学电气工程系，香港九龙达之路 83 号[2] 机器人事件驱动感知，意大利理工学院，Via Morego 30，16163 热那亚，意大利[3] 香港科技大学电子及计算机工程学系，香港九龙清水湾[4]香港城市大学太赫兹及毫米波国家重点实验室，香港九龙达之路 83 号

* 请将任何通信联系发送给本文作者。

电子邮件：arinbasu@cityu.edu.hk

摘要。 触觉传感对于机器人操作、假肢和辅助技术至关重要，但与视觉对应数据集相比，神经形态触觉数据集仍然有限。我们推出 STEMNIST，这是一个大规模神经形态触觉数据集，将 ST-MNIST 从 10 个数字扩展到 35 个字母数字类别（大写字母 A–Z 和数字 1–9），为基于事件的触觉识别提供了具有挑战性的基准测试。该数据集包含从 34 位参与者处采集的 7,700 个样本，使用定制的 16×16 触觉传感器阵列以 120 赫兹运行，并通过自适应时间微分编码为 1,005,592 个脉冲事件。遵循 EMNIST 的视觉字符识别协议，STEMNIST 填补了简化数字分类与需要字母数字辨别的现实触觉交互场景之间的关键空白。采用传统卷积神经网络（测试准确率为 90.91%）和脉冲神经网络（89.16%）进行的基线实验确立了性能基准。该数据集基于事件的格式、不受限制的空间变化和丰富的时间结构，使其适用于测试神经形态硬件和仿生学习算法。STEMNIST 支持对触觉识别系统进行可复现的评估，并为推动机器人学、生物医学工程和人机界面领域的高能效神经形态感知奠定基础。该数据集、文档和代码均已公开，以加速神经形态触觉计算研究。

1. 引言

触觉感知 [1] 与传感长期以来令科学家着迷，因为当视觉和听觉等其他感官途径受阻或无法使用时，它是帮助我们了解周围环境的最重要感知模态之一。其中一个主要组成部分是“触觉” [2]——人类最重要的感官之一，使我们能够

通过直接接触和形变感知，与环境进行自然互动。触觉传感对于开发人工智能机器至关重要，使其不仅能够安全地灵巧操作物体，还能通过实时提供接触动力学和材料属性信息，高效地参与人机交互。电阻式触觉阵列 [3] 凭借可扩展性和低成本，在当前机器人学研究中的传感器技术中脱颖而出，使大规模部署成为可能。施加在传感器元件阵列上的压力会引起电阻显著变化，这些变化可以被扫描、数字化并进一步处理，以用于后续任务。

诸如 [5] 的开创性工作展示了神经形态方法在以高时间分辨率捕捉触觉数据方面的优势；神经科学家认为，这一特性是生物体能够快速产生感觉反应的原因 [6]。然而，事件驱动神经形态领域缺少大规模数据集，无法帮助定量评估高时间精度触觉数据所承载的信息差异。鉴于近期机器人学和生物医学工程领域的人工触觉系统迅速涌现，这一问题如今尤为重要。具有开创意义的 MNIST [7] 数据集推动了手写数字识别机器学习的变革性进展，其扩展版本 EMNIST [8] 则通过标准化的数据划分和任务，将基准测试扩展至完整的英文字母表。然而，神经形态触觉类比数据集一直十分有限。ST-MNIST [9] 推出了首个公开可用的神经形态触觉数据集，捕捉在一组 10×10上传感器进行数字手写时产生的异步事件特征，并报告了卷积神经网络（CNN）和脉冲神经网络（SNN）的基线结果，证实了时空触觉模式适用于高效的事件驱动分类。尽管如此，ST-MNIST仅涵盖数字（0至9），且空间分辨率低于现代触觉皮肤，限制了它在学习字母符号中更丰富多样的时空形式方面的用途。相比之下，STAG [10]数据集提供了大规模触觉视频，记录了配备548传感器手套的人手与多种物体交互的过程；该数据集空间覆盖丰富，适用于全手抓握分析，但帧率显著较低（∼70.3 赫兹），这会模糊神经形态编码所需的精细事件时序信息，并妨碍时间敏感型脉冲推理。近期，盲文字符阅读数据集 [11]使用 iCub机器人的指尖，以12个采样频率为40赫兹的电容式触觉传感器记录了27种盲文类别（字母A至Z及空格）的时空模式，并展示了在 Intel Loihi 上进行能效较高的神经形态硬件推理，部分满足了这一需求。然而，其空间分辨率极低，时间采样也有限，仍不足以捕捉更复杂触觉模式的全部复杂性。表1比较了这些数据集的详细信息。

为弥补这些互补性不足，我们提出 STEMNIST，这是一种神经形态触觉数据集，将触觉事件记录扩展至手写英文字母大写字母（A 至 Z）和数字（1 至 9），使用 16×16 触觉传感器阵列（TSA）采集，具有较高的时间分辨率（120 赫兹）和准确的

表1. 基于事件的触觉数据集比较

数据集	模态	类别	规模	显著特征
ST-MNIST [9]	事件驱动触觉阵列	10 个数字	6,953 样本	10×10触觉像素，类似快速适应型响应，2 秒记录，人工神经网络／脉冲神经网络基线
STAG [10]	同步触觉手套	26 个物体	135,187 帧	548 个传感器，~7.3 赫兹，物体编号和重量估计
盲文 [11]	电容式触觉传感器	27 个（A-Z 和空格）	5,400 样本	12 个触觉像素，40 赫兹，iCub 指尖，Σ-Δ 编码，Loihi
STEMNIST（本研究）	事件驱动触觉阵列	35 个字符（26 个字母，A-Z 和 9 个数字，1-9）	7,700 样本 1,848,000 帧	16×16 触觉像素，120 赫兹，2 秒录制，差分帧预处理，神经网络基线

类事件预处理。为确保人工神经网络和脉冲神经网络方法的可复现性及公平基准测试，STEMNIST 采用适用于神经形态的差分编码，并遵循受 EMNIST 启发的标准化任务划分。这一资源旨在为机器人学、生物医学工程和神经形态计算领域提供基准测试先进学习算法所需的丰富类别和时空保真度，推动触觉感知的发展，并促进触觉接口、辅助设备等领域的研究。

本文其余部分安排如下：第2 节深入探讨相关研究。第3.1节介绍触觉硬件系统，并在第3.2节详细阐述数据采集过程。第4节全面讨论该数据集。第5节报告分类结果。第6 节概述未来研究的潜在方向，最后在第7节总结。

2. 相关工作

MNIST 数据集为手写数字分类确立了基准标准，催生了众多深度学习方法，并成为测试神经网络架构的试验平台。随后，EMNIST 以此为基础，通过纳入取自 NIST 特殊数据库的字母，系统地扩展了类别词汇表， 19 [12] 并规范了任务划分，以便进行公平评估和比较研究。然而，这些基于帧的数据集将手写内容表示为静态空间模式，从根本上丢弃了时间信息，因而限制了

它们在评估利用脉冲时序依赖处理的神经形态计算系统方面的效用。为弥补这一不足，研究者将静态图像转换为事件流，借助仿生视觉传感器对 MNIST 进行神经形态改造。MNIST-DVS [13] 使用动态视觉传感器（DVS），在图像以受控运动方式显示于液晶显示器上时记录像素级亮度变化，从而生成异步地址事件表示（AER）脉冲序列。类似地，N-MNIST [14]采用异步时间型图像传感器（ATIS） [15] 通过模拟扫视眼动对 MNIST 数字进行编码，从而能够基于时间事件数据评估脉冲神经网络。

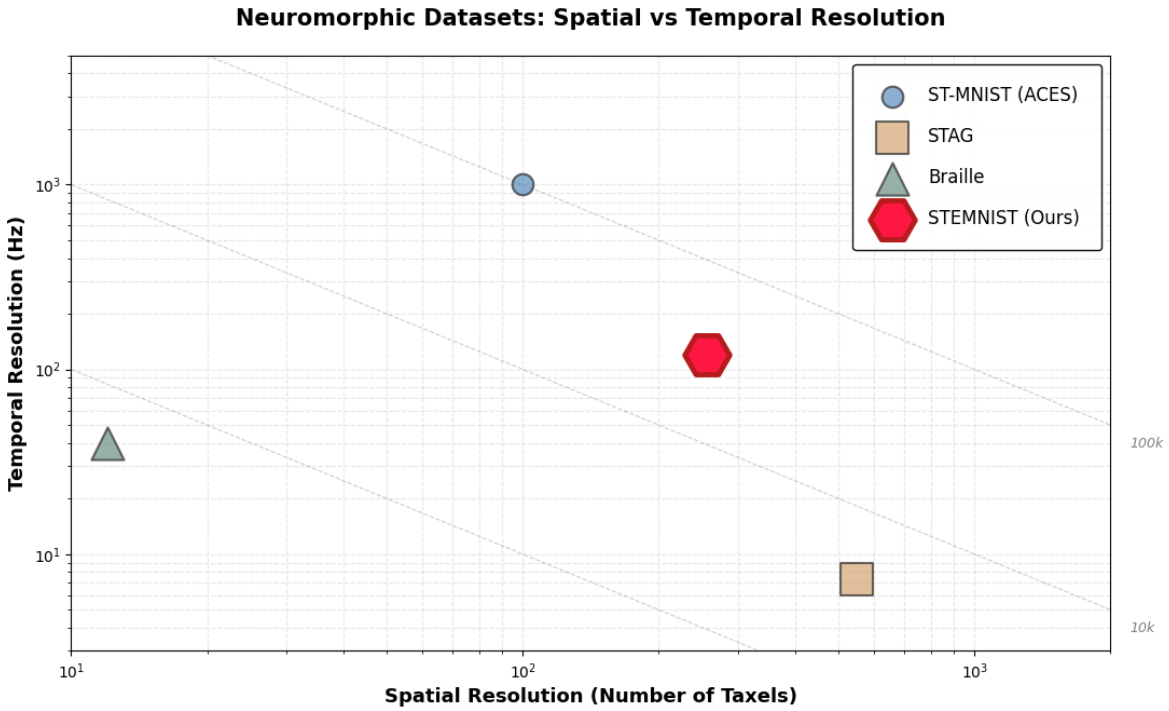

图1. 神经形态时空数据集的比较。

然而，这些合成生成的神经形态数据集的时间丰富性一直受到质疑。Iyer等人 [16] 证明，对时间折叠后的 N-MNIST 和 N-Caltech101 进行训练的基于帧的人工神经网络，准确率分别达到 99.23% 和 78.01%，与在相同数据集上最先进的脉冲网络性能相当，甚至更高。通过比较脉冲时序依赖可塑性算法（用于纯空间处理的 RD-STDP 和用于时空处理的 STDP-tempotron），他们表明，仅凭空间特征就足以对 N-MNIST 进行分类，而 DvsGesture [17]这类真正动态的数据集则需要时间算法。这一发现揭示了一个根本局限：通过人为的相机运动将静态图像转换为事件流，并不能保证其中包含足以证明在脉冲神经网络中采用依赖脉冲时序的处理所需计算开销合理的判别性时间信息。因此，这类合成数据集可能会低估以下方面的潜在优势：

神经形态架构旨在利用毫秒级精度的时序动态，而非分布在经过人为拉长的时间窗口中的空间模式。

这些观察结果促使研究者开发基于本质上具有动态特性的感觉模态的神经形态基准测试。在这些模态中，时间结构源自真实的物理交互，而非合成方法。手写过程中的触觉传感是理想的研究对象：手指在传感器阵列上描绘字符笔画时，压力变化持续发生，从而编码与书写速度、笔画顺序和接触动力学相关的真实时间依赖关系；这些特征在静态视觉图像中并不存在。与事后将时间信息附加到空间信息完整的图像上的视觉数据集不同，触觉手写天然会产生稀疏、异步且分布于时空中的事件，与事件驱动神经形态处理范式天然契合。然而，现有触觉数据集尚未充分发挥这一潜力。用于机器人操作的高密度触觉阵列 [18] 以及配备先进模拟前端的传感器架构，推动了空间分辨率和材料科学的发展，但同时具备基于事件的编码、高时空分辨率以及足以支持字符级识别的丰富类别的公开基准测试仍然十分稀缺。ST-MNIST [9] 首次推出神经形态触觉字符数据集，包含 10 个手写数字，采用 10×10阵列，并报告了卷积神经网络和脉冲神经网络基线，以验证事件驱动触觉分类的可行性。然而，该数据集仅包含数字，且空间分辨率不高，因此难以用于评估神经形态算法处理完整字母数字识别所需的多样时空模式的能力。尽管盲文字符阅读数据集[11]将类别数扩展至 27 个盲文字符（A–Z 和空格），但其空间分辨率极低（12 个触觉像素），时间采样也有限（40 赫兹），因而难以应用于需要更精细时空辨别能力的复杂触觉交互场景。神经形态计算通过对依赖活动的时空数据流进行异步、稀疏的事件驱动编码，为可扩展、低功耗的嵌入式感知带来了变革性优势。要实现这一潜力，需要能够保留自然触觉动态、具备丰富类别并支持标准化评估的数据集。现有神经形态视觉数据集缺乏这种感觉运动丰富性，其时间结构依赖合成方式构建。STEMNIST 通过高空间分辨率（256 个触觉像素）和高时间分辨率（120 赫兹）下的直接触觉测量，捕捉人类手写固有的时空复杂性，从而解决了这一问题。该数据集通过自适应时间微分，将 35 个字母数字类别（A–Z、1–9）编码为异步脉冲事件，在延续 ST-MNIST 挑战的同时，也支持在自然动态领域中检验时间编码的优势。

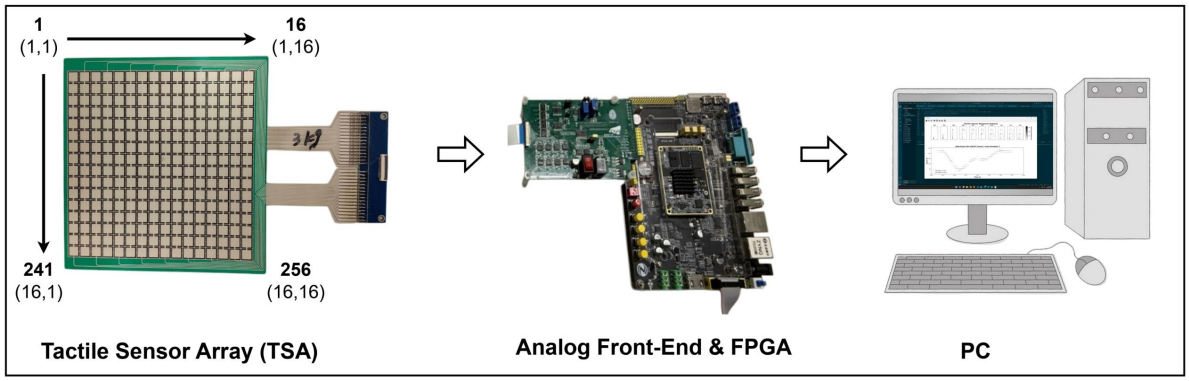

图 2. 用于创建数据集的触觉传感硬件系统，由16 × 16 触觉传感器阵列（TSA）、读出电子电路以及用于向个人计算机传输数据的现场可编程门阵列（FPGA）组成。

3. 材料与方法

3.1. 神经形态触觉传感硬件系统

图 2 展示了分布式柔性压力传感器阵列。该阵列通过在薄膜基底上精确印制纳米材料力敏层和银互连线制成。传感器由 256 个单元组成，每个单元的尺寸为 7.5 毫米× 7.5 毫米，并以 16×16阵列的形式排列在面积为 150 毫米 × 150 毫米的区域内。施加压力增大时，由于压阻效应，电阻会降低；机械臂校准证实了这一现象。事实上，在测试范围内，电阻从数 GΩ 单调下降至 3 kΩ，表明电阻与施加压力呈反比关系。这一明确特性确保了适当的动态范围和读出分辨率，为前端放大器的设计提供了重要指导。

图 3 展示了一个以现场可编程门阵列（FPGA）为核心构建的多通道数据采集系统，以及 16×16 电阻式触觉传感器阵列。每个传感器单元都集成在并联负反馈电路中。系统采用由寄存器（74HC16D）控制的开关芯片（ADG734），将阵列各行连接到精密基准电压 V_{ref} 或地（GND）。将某一行连接到地时，会启用采用 OPA2991 运算放大器的反馈放大器配置。这种可切换架构支持多种监测模式，并提升系统的数据压缩能力。输出电压定义为 $V_o = V_{ref}(1 + R_f/R_s)$，其中反馈电阻 $R_f = 100\,\text{k}\Omega$，传感器电阻 R_s 范围从数 kΩ 到数百 GΩ。为实现米勒补偿，在 C_f 两端并联 R_f 补偿电容，以增强电路稳定性，抵消高阻抗源产生的寄生电容效应。

为充分利用模数转换器的动态范围，信号链采用了八个低噪声、轨到轨 OPA2991 运算放大器。使用温度系数为 3 ppm/°C 的高稳定性 REF5025 电压基准来生成参考 V_{ref}。经放大的

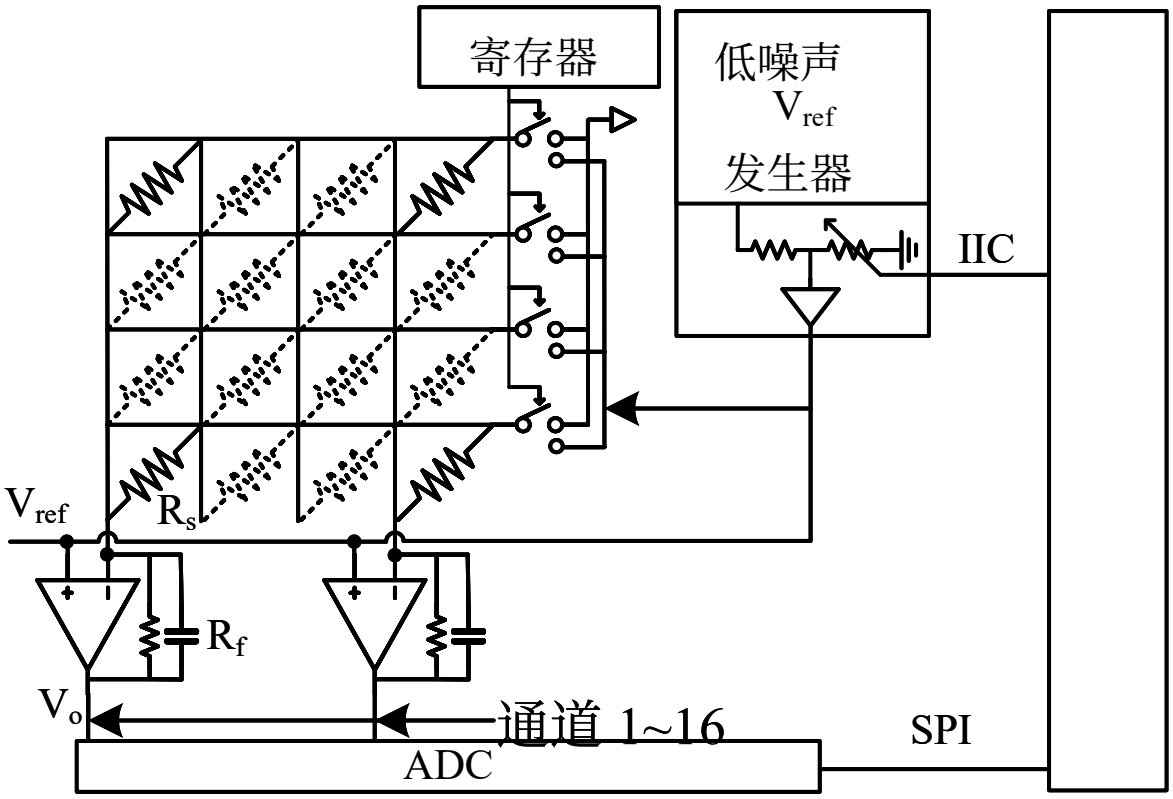

图 3. 读出硬件示意图。

模拟信号由 16 通道 ADS7961模数转换器进行数字化处理，其可编程输入范围为 0 至 V_{ref} 或 0 至 $V_{ref}/2$，采样率最高可达 1 兆赫兹。随后，数字数据通过串行外设接口（SPI）通信协议传输至现场可编程门阵列，以进行进一步处理。

3.2. 数据采集

共有 34 名参与者参加试验，每人生成了 35 个不同的字符类别，包括手写的大写英文字母（A 到 Z）和数字（1 到 9）。参与者被要求用食指直接在 16×16 TSA 表面书写字符。采集协议将这 35 个字符分为连续的七组，每组五个字符，以便系统、均衡地进行数据采集。七组分别为集合 1（A 到 E）、集合 2（F 到 J）、集合 3（K 到 O）、集合 4（P 到 T）、集合 5（U 到 Y）、集合 6（Z、1 到 4）和集合 7（5 到 9）。每位参与者依次完成全部七组；最初有 10 名参与者各重复这一完整周期 10 次。之后，重复次数减少到 5 次，从而每位参与者的每个字符获得 5 个样本。每个周期结束后都安排了适当的休息时间，以防参与者感到疲劳。这确保了整个记录阶段所采集数据的质量始终一致。

为确保时间间隔精确，并为参与者提供清晰易懂的视觉反馈，由现场可编程门阵列控制的发光二极管（LED）指示灯经过编程，每个字符恰好亮起 2 秒，以标示 TSA 正在扫描的数据采集窗口。每个 2 秒录制时段结束后，会安排 1 秒的试次间隔，提示参与者准备书写下一个字符；此间隔期间 LED 熄灭，随后再次亮起 2 秒以记录下一个字符。集合 1 中 5 个字符的记录流程如图 4 所示。LED 提示确保所有参与者大体上同时开始作答，并清楚掌握确切的记录时长，从而有助于保持时间安排一致，并减少因时间不确定性导致的参与者反应时间差异。

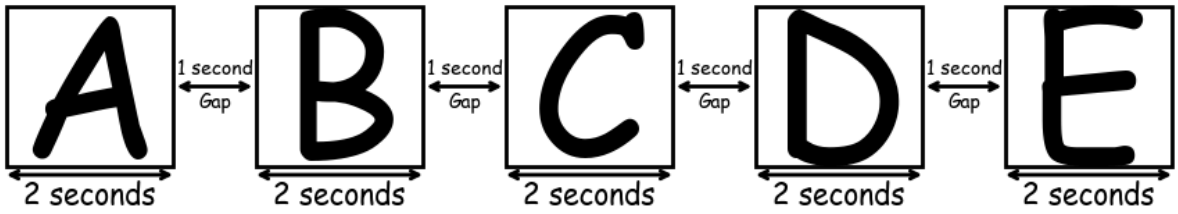

图 4. 数据采集方案。

在每个 2 秒记录窗口内，TSA 均以 120 赫兹的采样率持续扫描，从所有 256 个触觉像素（16×16 阵列）采集高时间分辨率的压力分布数据。在数据采集期间，参与者在每组中依次书写五个字符，现场可编程门阵列控制的数据采集系统则通过 SPI 以 115200 的波特率将原始传感器读数流式传输到主机。ADS7961 的原始模数转换器输出被数字化为 8 位无符号整数。由于压阻式传感器的特性，放大器增益配置和基准参考电压 V_{ref} 针对经过测试的压力范围进行了优化，因此在主动手写期间观察到的输出值有效动态范围主要介于 75 和 255 之间；较低的基准值（$\sim$75）对应未施加压力的状态，较高的值（接近 255）则对应较大的接触压力。

为确保数据完整性，采集流程采用了基于 Python 特别设计的图形用户界面（GUI），用于同步 LED 时序与串行数据采集，并实时验证帧标记（0x0F 十六进制分隔符）。每组中的五字符记录最初存储为单个 HDF5 文件（每组共 1,200 帧），并附带采集时间戳、采样率和阵列尺寸等元数据。随后，使用自动化的后处理 Python 脚本，将帧序列均匀拆分，以平衡不同字符的时间窗口；每个包含五个字符的 HDF5 文件由此被分割为独立的字符样本，生成五个单独的文件，每个文件包含一个字符的帧。最终数据集共包含 7,700 个样本和 1,848,000 帧，均匀分布于全部 35 个字符类别中，以确保各类别的样本数量相同。每个样本均以独立的层次数据格式第 5 版

（HDF5）文件存储在集中式原始字符存储库中，并采用标准化命名约定 `<ParticipantID>_<Character>_<RepetitionNumber>.h5`，以便后续进行基于事件的预处理和神经形态编码。原始压力数据文件以完整时间分辨率（120 赫兹）存储，以便灵活地进行基于阈值的脉冲生成和差分帧分析，详情见第4节。

4. 脉冲触觉扩展MNIST数据集

4.1. 脉冲生成与编码

以120 赫兹记录的原始压力帧经过后处理，生成适用于神经形态计算的类事件脉冲表示。脉冲生成采用连续帧之间的时间微分：当压力增量超过阈值时，生成ON脉冲（极性 $+1$）；当压力减小并低于相应的负阈值时，生成OFF脉冲（极性 -1）。

每个样本在帧 f 处的时间微分计算如下：

$$\Delta P(f, x, y) = P(f, x, y) - P(f-1, x, y) \tag{1}$$

其中， $P(f$、 x、 $y)$表示 16×16阵列中帧 f和触觉像素位置（x、 y）处的压力值。为适应参与者间变异性带来的书写压力差异，我们采用了自适应逐样本阈值化方案。对于每个样本，阈值 θ_{sample} 计算为其时间微分分布的第90百分位数，该分布基于非零绝对值计算：

$$\theta_{\text{sample}} = \text{percentile}_{90}\big(\{\, |\Delta P(f, x, y)| : |\Delta P(f, x, y)| \neq 0 \,\}\big) \tag{2}$$

确保将观测到的非零绝对压力变化幅值中最高的十分位（前 10%）归类为脉冲事件。由于触觉手写在空间和时间上都具有稀疏性，通常在任意时刻仅激活少数触觉像素，因此计算百分位数时排除了零值（表示压力没有变化的触觉像素）。由于绝大多数时间差分值为零，若基于完整分布计算百分位数，得到的阈值将主要受背景静止状态影响，而非实际触觉事件的幅值分布。排除这些零值可确保阈值反映活动轨迹上的笔画动态。为避免噪声影响，计算出的阈值被限制为不低于 4：

$$\theta_{\text{sample}}^{\text{clipped}} = \max(\theta_{\text{sample}}, 4) \tag{3}$$

该下界大致对应于传感器的噪声底；低于此值时，电子噪声产生的伪事件会占据主导，掩盖真实的触觉信号。我们未设置上界，因此自适应方法可以自然适应参与者从轻触到重压的各种书写压力范围。这种基于百分位数的裁剪策略类似于事件驱动神经网络的稳健归一化方法， [19] 能够维持足够的脉冲率，从而保留时间

动态特性，适用于大多数样本。保守的基于离群值的归一化（例如，不设界限而使用每个样本的最大值）可能导致大多数样本的发放率不足，而基于百分位数的界限则能确保不同书写风格下都能稳定地产生事件。在全部 7,700 个样本中，自适应阈值的均值为 $\mu = 7.77$，标准差为 $\sigma = 4.44$，证实了该方法能够有效适应不同的书写风格，同时维持抗噪声的下界。

随后根据以下准则生成脉冲：

$$\text{Spike}(f,x,y) = \begin{cases} +1 & \text{if } \Delta P(f,x,y) > \theta^{\text{clipped}}_{\text{sample}} \quad \text{(ON spike)} \\ -1 & \text{if } \Delta P(f,x,y) < -\theta^{\text{clipped}}_{\text{sample}} \quad \text{(OFF spike)} \\ 0 & \text{otherwise} \quad \text{(NO spike)} \end{cases} \tag{4}$$

每个脉冲事件都按照神经形态系统的地址事件表示（AER）格式标准编码为一个元组（t_s,$\text{ID}_{\text{taxel}}, p$）。其中，$t_s$ 表示以秒为单位的脉冲时间戳，计算方式为 $t_s = f/f_{\text{sample}}$，其中 f为帧索引，$f_{\text{sample}} = 120$ 赫兹为采样率。触觉像素标识符 $\text{ID}_{\text{taxel}} \in [1, 256]$ 的计算方式如下：

$$\text{ID}_{\text{taxel}} = (X-1) \times 16 + Y \tag{5}$$

对于行 X 和列 Y（均从 1 开始索引），$p \in \{-1, +1\}$ 表示脉冲极性。这种稀疏的事件驱动表征保留了触觉手写的异步时间结构，同时支持在神经形态硬件上进行高效存储和处理。

4.2. 文件格式与组织

STEMNIST数据集采用分层目录结构，以支持基于脉冲的神经形态处理和基于帧的深度学习方法。该数据集由三个主要部分组成：

原始压力数据：与每个手写字符相关的原始 8 位压力数据最初集中存储在一个包含五个字符的 HDF5 文件中，随后通过基于 Python 的后处理进行拆分，并分别保存为独立的 HDF5 文件。每个样本包含一个压力读数三维数组（240 帧 $\times$ 16 $\times$ 16），以及完整的元数据，包括参与者标识符、字符标签、重复索引、采样频率（120 赫兹）和时间采集时间戳。文件命名约定遵循第 3.2 节中详细说明的方案，结构如下 `<ParticipantID> <Character> <RepetitionNumber>.h5`。

处理后的脉冲数据： 这些是基于事件的数据表示，以结构化 NumPy 数组的形式存储在 HDF5 文件中。每个脉冲文件都包含三个字段：时间戳（float32）、触觉像素 ID（int32）和极性（int16）。此外，每个脉冲文件还包含属性，用于显示脉冲总数和自适应阈值

值（θ^{clipped} 样本），以及对原始文件的引用。脉冲文件采用 `<ParticipantID> <Character> <RepetitionNumber> spikes.h5` 的命名规则，并根据各自的字符类别整理到 35 个子目录中。

训练-测试划分： 分类实验采用分层随机 80-20 训练-测试划分，并使用 5 个随机种子（42、123、456、789、1011）进行五次独立试验，确保各次划分中的类别分布均衡。为评估模型泛化一致性和鲁棒性，结果以第五节中五次试验的平均准确率 ± 标准差形式呈现5。脉冲生成后的全部样本均包含在 `ProcessedSpikes` 目录中，其中所有脉冲文件进一步按字符类别整理，以供参考并确保可复现性。

4.3. 数据集属性

STEMNIST数据集包含 7,700 个样本，均匀分布于 35 个字符类别（26 个大写字母 A–Z 和 9 个数字 1–9）中，每个类别包含 220 个样本。每个样本包含来自 16×16TSA的 2 秒录制（240 帧，采样频率为 120 赫兹），每个样本共计 61,440 个原始压力测量值。经过自适应脉冲生成后，数据集总计包含 1,005,592 个脉冲事件，其中有 498,230 个 ON脉冲和 507,362 个 OFF脉冲，平均每个样本包含 130.60 个脉冲（64.7 个 ON脉冲和 65.9 个 OFF脉冲）。

表2。 A 至 E 字母的描述性统计

字符	**A**	**B**	**C**	**D**	**E**
ON事件	64.76 (17.26)	78.05 (23.49)	59.55 (20.28)	67.85 (22.27)	69.96 (22.10)
OFF事件	65.21 (16.88)	78.56 (22.91)	58.54 (19.22)	68.05（21.43）	72.46（23.94）
X mean	8.56（0.90）	8.89（0.98）	8.01（0.87）	8.80（1.02）	8.43 (1.00)
Y mean	8.75 (0.81)	8.95 (0.77)	9.05 (0.93)	8.92 (0.78)	8.98 (0.87)
X 范围	7.71 (1.74)	7.73 (1.81)	8.23 (1.78)	8.05 (1.81)	7.97 (1.62)
Y 范围	10.16（1.62）	10.81（1.65）	9.68（1.68）	10.18（1.72）	9.96（1.65）
样本量	220	220	220	220	220

注：所有统计表中的数值均以均值（标准差）表示，样本量除外；样本量表示每个类别的样本数。

表 3. 字母 F 至 J 的描述性统计

字符	**F**	**G**	**H**	**I**	**J**
ON事件	60.83 (15.00)	68.83 (21.15)	64.93 (18.35)	61.69 (19.98)	60.70 (17.94)
OFF事件	63.00 (16.24)	67.61 (19.81)	66.22 (18.80)	63.74 (21.82)	61.68 (17.77)
X mean	8.03 (1.10)	8.73 (0.93)	9.07 (0.97)	8.78 (1.05)	8.45 (1.11)
Y mean	7.67 (0.74)	9.20 (0.74)	8.98 (0.75)	9.27 (0.90)	9.32 (1.02)
X 范围	7.27 (1.88)	7.54 (1.71)	6.97 (1.66)	7.31 (2.23)	7.51 (1.76)
Y 范围	9.80 (1.53)	10.07 (1.60)	10.14 (1.60)	9.59 (1.44)	10.40 (1.55)
样本量	220	220	220	220	220

表 2-8 按照 ST-MNIST 确立的格式，详细呈现了每个字符类别的统计数据。这些统计数据中的高标准差表明

表4. 字母K至O的描述性统计

字符	K	L	M	N	O
ON事件	65.79 (18.17)	50.76 (15.42)	74.66 (18.56)	66.45 (18.95)	66.66 (21.63)
OFF事件	66.82 (17.64)	52.71 (15.76)	75.99 (18.38)	69.12 (21.22)	67.09 (21.75)
X mean	8.39 (0.96)	8.07 (0.92)	9.18 (0.93)	8.93 (0.87)	8.81 (0.79)
Y mean	9.40 (0.87)	10.45 (0.82)	8.79 (1.03)	9.02 (0.92)	8.59 (0.94)
X 范围	7.00 (1.70)	7.65 (1.93)	9.15 (1.83)	8.14 (1.99)	8.36 (1.74)
Y 范围	9.93 (1.62)	9.37 (1.45)	9.66 (1.73)	9.61 (1.70)	9.63 (1.58)
样本量	220	220	220	220	220

表5. 字母P至T的描述性统计

字符	P	Q	R	S	T
ON事件	63.60 (17.55)	69.76 (18.22)	72.33 (20.86)	63.67 (20.20)	55.49 (18.73)
OFF事件	65.44 (17.71)	70.16 (18.19)	72.78 (21.04)	64.02 (19.98)	57.82 (19.36)
X mean	7.97 (1.17)	8.82 (0.96)	8.84（1.17）	8.69（0.91）	8.97（0.90）
Y mean	7.67（0.84）	9.14（0.90）	8.81（0.74）	9.14 (0.87)	7.73 (1.04)
X 范围	6.91 (1.65)	8.08 (1.64)	7.18 (1.72)	7.20 (1.50)	8.70 (2.14)
Y 范围	10.28 (1.40)	9.62 (1.52)	10.14 (1.64)	9.97 (1.56)	9.87 (1.61)
样本量	220	220	220	220	220

表6. 字母 U 至 Y 的描述性统计

字符	U	V	W	X	Y
ON事件	61.19 (14.99)	57.98 (17.92)	76.28 (20.76)	62.93 (19.81)	63.75 (20.42)
OFF事件	63.35 (15.18)	60.65 (18.06)	78.45 (22.20)	64.47 (20.52)	63.63 (19.88)
X mean	8.78 (1.05)	9.03 (0.83)	8.93 (0.86)	8.58 (0.89)	8.95 (0.82)
Y mean	9.20 (0.76)	8.83 (0.81)	9.20 (0.73)	9.37 (0.85)	8.91 (0.89)
X 范围	7.36 (1.59)	7.81 (1.93)	9.60（1.76）	7.80（1.66）	7.48（1.77）
Y 范围	10.06（1.40）	9.90（1.49）	9.31 (1.74)	9.58 (1.57)	11.39 (1.57)
样本量	220	220	220	220	220

表7. 字母Z和数字1至4的描述性统计

字符	Z	1	2	3	4
ON事件	71.97 (22.50)	46.60 (19.41)	63.65 (18.45)	65.21 (18.69)	68.74 (18.75)
OFF事件	75.00 (24.49)	47.69 (19.24)	65.52 (18.89)	66.11 (17.92)	71.69 (20.80)
X mean	8.58 (0.92)	8.55 (0.99)	8.71 (0.74)	9.03 (0.81)	8.86 (0.97)
Y mean	9.08 (0.85)	9.47 (1.06)	9.69 (0.99)	9.26 (1.01)	9.23 (0.83)
X 范围	8.31 (1.64)	4.02 (3.10)	8.20 (1.56)	7.22 (1.68)	8.45 (1.82)
Y 范围	8.75 (1.59)	10.06 (1.67)	8.95 (1.58)	10.33 (1.65)	10.54 (1.69)
样本量	220	220	220	220	220

表8. 数字5至9的描述性统计

字符	**5**	**6**	**7**	**8**	**9**
ON事件	65.26 (19.26)	59.52 (17.75)	53.53 (17.38)	72.05 (19.49)	69.71 (18.69)
OFF事件	67.21 (19.52)	59.31 (17.23)	53.91 (17.14)	72.77 (19.62)	69.42 (17.46)
X mean	8.55 (0.87)	8.40 (0.94)	9.21 (0.87)	8.71 (0.88)	9.26 (1.03)
Y mean	8.61 (0.78)	9.37 (0.91)	7.99 (1.02)	8.75 (0.94)	7.93 (1.09)
X 范围	7.85 (1.69)	6.45 (1.67)	6.91 (1.68)	6.23 (1.57)	6.34 (1.68)
Y 范围	9.81 (1.39)	9.80 (1.41)	10.09 (1.39)	10.16 (1.47)	11.15（1.66）
样本量	220	220	220	220	220

在某一类别内的所有样本中，脉冲计数存在显著差异。这种差异源于人们书写方式、书写速度以及施加压力大小的不同。通过该示例，可以清楚地反映这些字符之间的相对复杂度和笔画数量差异。类别“B”平均有 78.1±23.5 个 ON事件和 78.6±22.9 个 OFF事件，而类别“L”的数量要少得多，平均只有 50.8±15.4 个 ON事件和 52.7±15.8 个 OFF事件。空间统计量（X 均值、Y 均值、X 范围和 Y 范围）也表明，同一字符类别内以及不同字符类别之间存在很大的变异性。STEMNIST 的设计没有对记录的起点和终点施加任何空间约束，因此能够保留人手手写中的大量变异性；相比之下，N-MNIST 源自经过预居中的静态 MNIST 图像，其空间范围是固定的。因此，不受限制的触觉手写中书写位置、字符大小和方向的变化，便通过空间范围统计量较大的标准差得以体现。

图5展示了字母“B”这一代表性样本的触觉表征随时间演变的过程。第一行显示的是基于时间积分原始压力数据生成的归一化触觉图像，呈现了字符笔画轨迹逐渐形成的过程。第二行展示了随时间在空间上累积的脉冲事件图像。ON事件和OFF事件的累积情况表明，事件驱动表征如何保留书写手势模式的底层时空结构。底部的脉冲栅格图显示了在 2 秒记录窗口内，256 个触觉像素产生的 185 个 ON事件和 178 个 OFF事件的实际时间分布。字母“B”的竖干、上环和下环分别对应不同的时间簇，而特定触觉像素范围内脉冲活动的集中情况则与构成该字符的连续笔画相对应。基于事件的表征能够有效捕捉触觉手写笔画随时间和空间展开的过程，提供丰富的时空特征，适用于神经形态脉冲神经网络处理。

根据所书写的具体字符，这些统计数据可用作分类任务的判别特征。然而，先前针对神经形态数据集的研究 [14] 表明，如果将汇总统计量直接用作简单分类器的特征

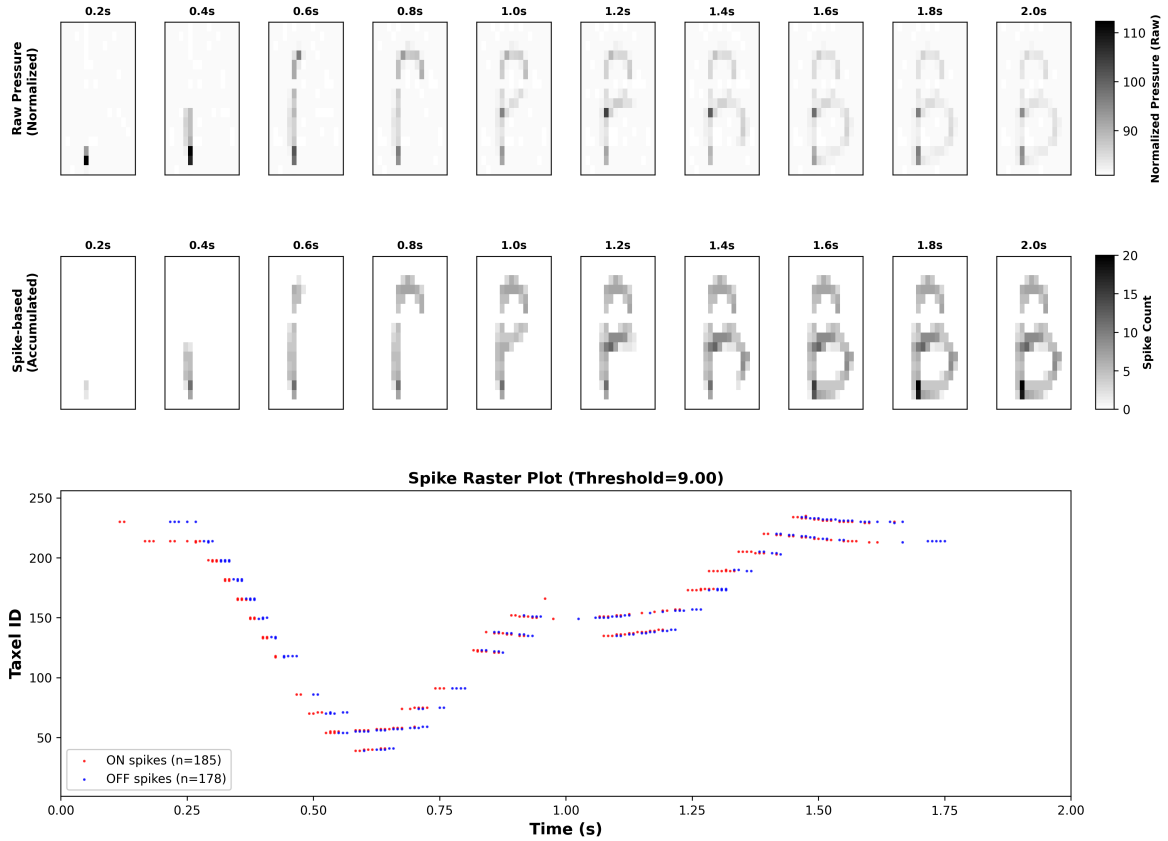

图 5. 顶部行：字母“B”的归一化“触觉图像”随时间的变化，由累积的原始压力帧生成。**中间行：**基于脉冲的触觉表示随时间的变化，通过对 ON 和 OFF 脉冲事件的数量随时间求和得到。**底部行：**脉冲栅格图显示了 2 秒录制期间，沿 256 个触觉像素分布的 185 个 ON 事件（红色）和 178 个 OFF 事件（蓝色）。对于此样本，自适应阈值为 $\theta_{\text{sample}} = 9.0$。

（例如 k近邻），其性能会逊于利用脉冲序列完整时空结构的分类器。因此，第5节讨论的分类实验侧重于利用完整的脉冲序列，而非仅使用汇总统计量。

5. 分类

本研究探讨的核心触觉识别任务是对由 35 个类别构成的集合进行多类分类，包括大写字母**A-Z**和数字**1-9**。这一直接类比 EMNIST 方案的任务，相较于 ST-MNIST 的 10 类数字识别任务，为评估提供了更具挑战性且更具代表性的基准测试，要求模型区分共享重叠笔画基元的拓扑相似字符（例如 B/P/R/K、C/O/Q、I/J/L、U/V/W）。为建立基线性能指标并验证我们的基于事件的触觉数据集的可判别性，我们实现了两种分类方法：一种是传统的卷积

神经网络（CNN），用于处理双通道脉冲累积图；另一种是利用时间脉冲动态、与神经形态计算兼容的脉冲神经网络（SNN）。我们的评估方案分为三个阶段：（1）在训练集上通过五折交叉验证进行超参数优化；（2）通过采用不同随机数据划分进行多次训练，评估鲁棒性；（3）遵循严格的机器学习最佳实践，在最终留出测试集上进行评估，以防止数据泄漏。

5.1. 模型架构

卷积神经网络（*CNN*）：该 CNN 采用紧凑的前馈设计，针对触觉传感器阵列的 16×16空间分辨率进行了优化。输入为双通道脉冲累积图，其中通道 0 编码 ON事件（正向压力变化），通道 1 编码 OFF事件（负向压力变化），并在 2 秒记录窗口内进行汇总。这种双极性表示保留了双向触觉动态，同时能够通过标准卷积运算进行高效处理。该架构由两个卷积块和一个全连接分类器组成：

- **卷积块 1：** 8 个滤波器，使用 4×4卷积核，步幅为 1，不填充（输出： $8\times13\times13$），采用修正线性单元（ReLU）激活， 2×2 最大池化，步幅为 2（输出：$8\times6\times6$）。

- **卷积块 2：** 16 个滤波器，使用 3×3卷积核，步幅为 1，填充为 1（输出： $16\times6\times6$），采用 ReLU 激活， 2×2 最大池化，步幅为 2（输出： $16\times3\times3$）。
- **随机失活：** 应用于 144 维展平特征向量的正则化层，丢弃概率为 0.3。

- **分类器：** 将 144 个特征映射到 35 个输出类别的全连接层（26 个字母 A-Z $+$ 9 数字 1-9）。

参数总数为 6,507（Conv1：264，Conv2：1,168，FC：5,075），确保计算效率，适合在资源受限系统上进行边缘部署。

脉冲神经网络 *(SNN)*：该脉冲神经网络架构与卷积神经网络的空间拓扑相同，但以漏积分发放 (LIF) 神经元取代静态修正线性单元激活，直接处理时间脉冲序列。输入脉冲序列在 2 秒录制窗口内划分为 20 个离散时间步（每个时间步的时间分辨率为 100 毫秒），从而保留触觉手写笔画的异步动态。每个卷积层和全连接层之后都接有一个漏积分发放神经元群，其膜衰减常数为 $\beta = 0.9$，用于模拟符合生物特性的脉冲时序动态。前向传播按顺序遍历全部 20 个时间步，膜电位对脉冲输入进行积分，并在超过发放阈值时产生输出脉冲。最终通过频率编码，对时间维度上的输出脉冲计数求和，以完成分类。架构如下：

- **卷积块 1：** 8 个滤波器，使用 4×4 卷积核；漏积分发放神经元，使用 $\beta = 0.9$，2×2 最大池化。
- **卷积块 2：** 16 个滤波器，使用 3×3 卷积核；采用带有 $\beta = 0.9$、2×2 最大池化的泄漏积分发放神经元。
- **随机失活：** 在分类之前对池化特征应用丢弃概率为 0.3 的正则化。

- **分类器：** 通过全连接层映射到 35 个类别，并使用输出泄漏积分发放神经元；在时间步上汇总脉冲计数用于分类。

训练采用 snnTorch [20] 进行基于替代梯度的时间反向传播，使用反正切（ATan）替代梯度函数，这是 snnTorch 针对不可微 Heaviside 脉冲函数的默认梯度近似。ATan 替代梯度提供平滑的钟形梯度，在广泛的膜电位范围内平衡训练稳定性与梯度幅度，从而尽管脉冲生成机制不可微，仍能实现端到端的梯度下降。两种架构具有相同的参数数量（6,507），以确保表征能力比较的公平性。所有实验均在配备 CUDA 加速的 NVIDIA 图形处理器上，使用 PyTorch 2.0 完成。未采用任何数据增强（例如旋转、缩放、噪声注入），以保留自然的触觉手写变异性。

5.2. 超参数优化

为确定最佳训练配置并防止测试集污染，我们仅在训练集划分上（共 6,160 个样本，占完整数据集的 80%）进行了五折分层交叉验证。每一折均保持类别分布均衡，每折包含 4,928 个训练样本和 1,232 个验证样本。我们针对两种架构评估了多种超参数组合，探索了学习率 $\{0.001, 0.005\}$、批次大小 $\{32, 64\}$；对于脉冲神经网络，则还探索了时间分辨率 $\{10, 20\}$ 时间步。所有配置均采用随机失活正则化（概率为 0.3）和 L2 权重衰减（10^{-4}），以防止过拟合。

CNN 每折训练 20 个轮次，而 SNN 由于通过替代梯度学习时间脉冲模式的复杂性更高，需要训练 50 个轮次。训练采用交叉熵损失和 Adam 优化器。CNN 的最佳超参数配置根据 5 折验证准确率均值最高确定，具体为：学习率 0.005、批量大小 64，交叉验证准确率达到 $\mathbf{89.72\%} \pm \mathbf{0.61\%}$。SNN 的最优设置为：学习率 0.005、批量大小 64、20 个时间步，交叉验证准确率达到 $\mathbf{88.25\%} \pm \mathbf{0.98\%}$。这一阶段仅 1.47% 的微小性能差距表明，双通道脉冲编码和时间分箱策略有效保留了适用于神经形态处理的判别性触觉特征。

5.3. 多重划分鲁棒性评估

为评估模型稳定性和泛化一致性，避免结论仅依赖单次训练-测试划分，我们针对每种架构使用训练集的不同随机 80-20 划分，训练了五个独立模型，所用随机种子为 $\{42, 123, 456, 789, 1011\}$。每次划分均保持分层类别平衡，并在每次划分时分别重新初始化模型权重和随机数生成器。该方案确保报告的性能指标反映模型对数据采样变异性的鲁棒性，而非偶然的划分结果。使用 5.2 中确定的最优超参数，我们从头开始训练每个模型，并逐轮次跟踪训练损失、验证损失、训练准确率和验证准确率。随后汇总五次划分的这些指标，计算带标准差带的平均学习曲线，以展示收敛行为及其方差（图 6）。CNN 在各次划分中的平均验证准确率为 $\mathbf{90.36\%} \pm \mathbf{0.70\%}$（范围：89.04 - 91.07%），而 SNN 达到 $\mathbf{87.81\%} \pm \mathbf{1.20\%}$（范围：85.80 - 89.29%）。SNN 性能的较大方差反映了替代梯度训练对初始化敏感这一已知挑战；不过，两种架构都表现出稳定的收敛趋势，验证损失持续下降或趋于平稳，证实正则化有效。

5.4. 最终测试集评估

完成超参数选择和鲁棒性验证后，我们使用最优配置在完整的 6,160 个样本训练集（占完整数据集的 80%）上训练最终模型，并仅在此前所有阶段均未使用过的 1,540 个样本留出测试集上评估一次。这一严格流程可防止信息泄漏，并提供能够代表真实部署场景的无偏性能估计。

量化性能：最终的卷积神经网络达到**90.91% 的测试准确率**，而脉冲神经网络达到 89.16%，两者的**性能差距仅为 1.75%**。两个模型的表现都远胜于随机猜测（35 个等概率类别时为 2.86%），甚至超过了 See 等人报告的准确率 [9]（针对 10 类 ST-MNIST）。值得注意的是，我们的 35 类卷积神经网络和脉冲神经网络的准确率均超过了针对 10 类 ST-MNIST 的人工神经网络 89% 的表现，尽管我们面对的是一个$5\times$难度更高的分类问题，这展现了系统化正则化和双通道编码的有效性。

混淆分析：图7 展示了五个子划分验证集的平均归一化混淆矩阵，揭示了各类别的性能模式。两种架构的混淆矩阵都呈现明显的对角线占优，表明逐类别准确率较高；误分类主要集中在拓扑相似字符对之间：I/J/L（竖笔画）、M/N/W（多峰结构）、U/V（弧形笔画）以及数字 5/6/8（共享环形结构）。这些错误模式与不受限制的触觉手写中预期的人类感知混淆相符，也凸显了在现实环境中部署仍面临的挑战。

特征空间可视化：为定性评估学习到的表征，我们

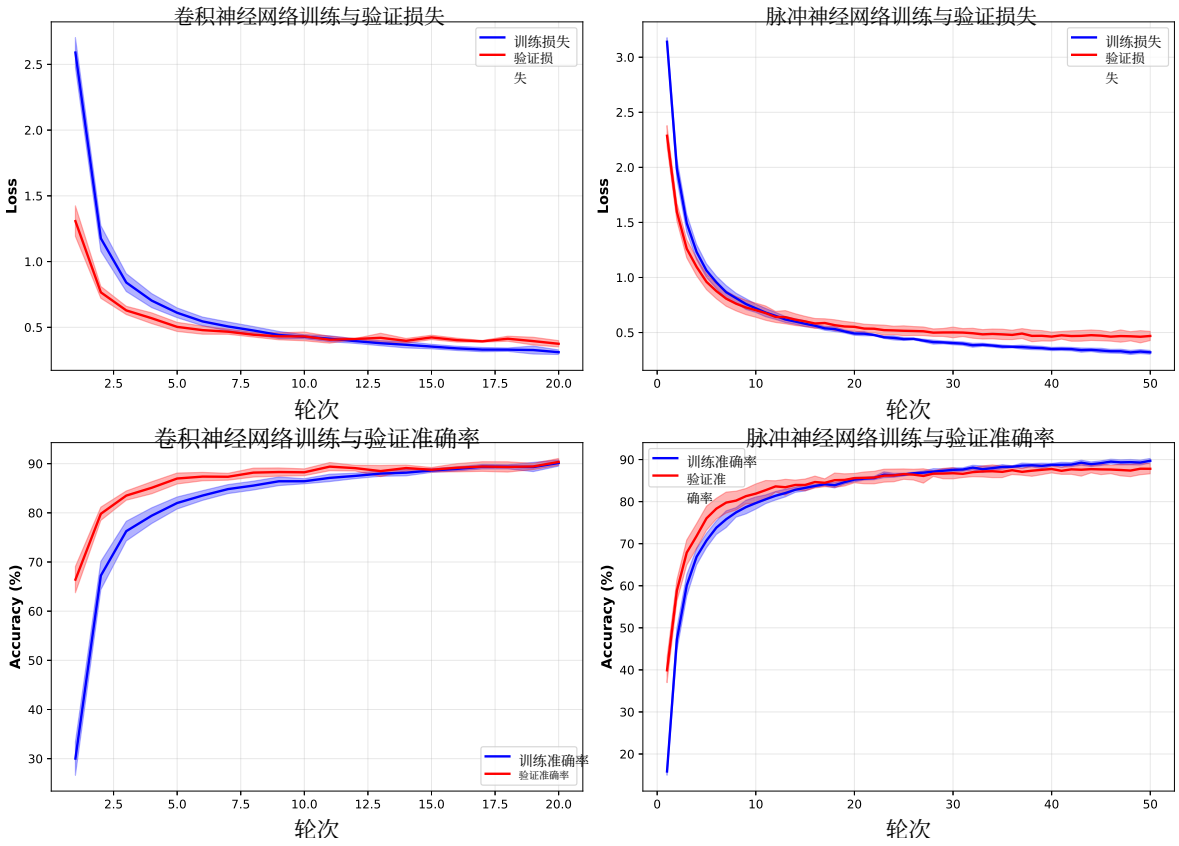

图 6. 在五个随机子集划分上取平均得到的学习曲线。**顶部：**训练与验证损失。**底部：**训练与验证准确率。阴影区域表示 ± 1 标准差。卷积神经网络（左）在 20 个训练轮次内收敛；脉冲神经网络（右）需要 50 个训练轮次才能达到相当的性能。随机失活和权重衰减可防止过拟合，这一点从稳定的验证损失中可以看出。

采用 t 分布随机邻域嵌入（t-SNE [22]）对从一个验证集划分中提取的 144 维倒数第二层特征进行分析（图8）。对于卷积神经网络，特征直接取自展平后的池化后激活（$16\times3\times3 = 144$ 维）。对于脉冲神经网络，我们按照速率编码范式，对倒数第二层每个神经元在全部 20 个时间步上的膜电位求和，以此提取特征；在该范式中，时间脉冲计数（与积分后的膜电位成正比）编码类别信息。这种时间积分将 20×144时空特征张量压缩为每个样本一个 144 维向量，从而能够在相同嵌入空间中与卷积神经网络特征直接比较。对于大多数字符类别，卷积神经网络和脉冲神经网络都形成了分离良好的簇，而相似字符之间存在一些预期的重叠（例如 U/V/W、K/R/Y）。与卷积神经网络相比，脉冲神经网络的特征空间呈现出略明显的簇扩散，这可能反映了脉冲时序动力学的随机性，以及 2 秒记录窗口内时间积分的变化。尽管如此，清晰的类别边界依然存在，证实经时间积分的脉冲模式能够编码足够的判别信息，从而实现高准确率的神经形态分类。

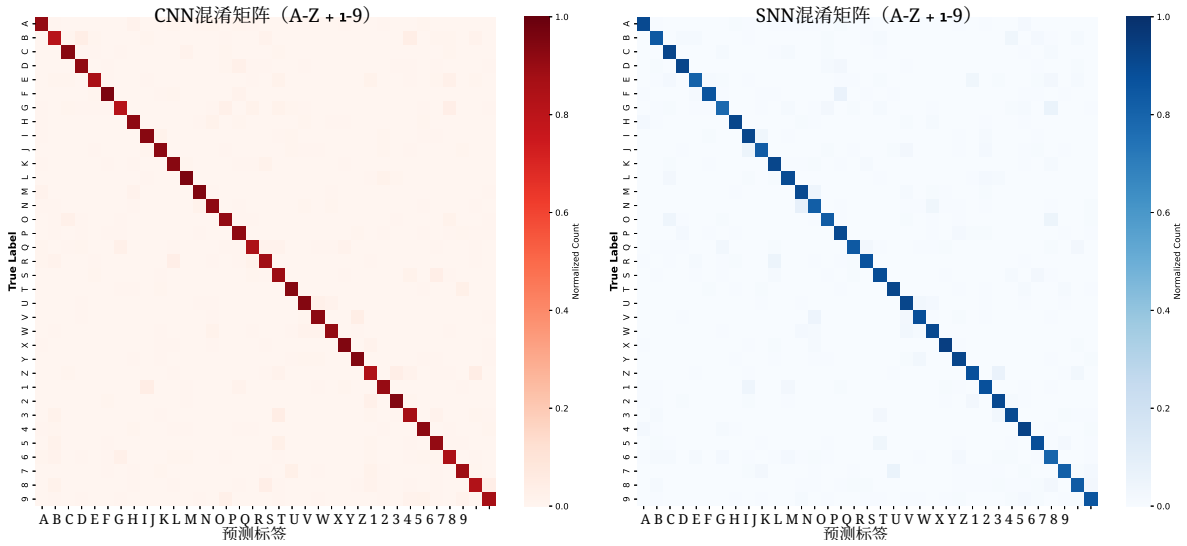

图 7. 五个子划分模型的平均归一化混淆矩阵。 **左：** 卷积神经网络。**右：** 脉冲神经网络。两者均呈现明显的对角线优势，常见错误发生在视觉上相似的字符之间（例如字母中的 I/J/L，数字中的 5/6/8），这与人类的感知模糊性一致 [21] 这种模糊性也见于触觉字母识别任务：共享笔画基元（垂直笔画、弯曲笔段）的拓扑相似字符会产生不对称的混淆模式，即使在经过训练的参与者中也是如此。

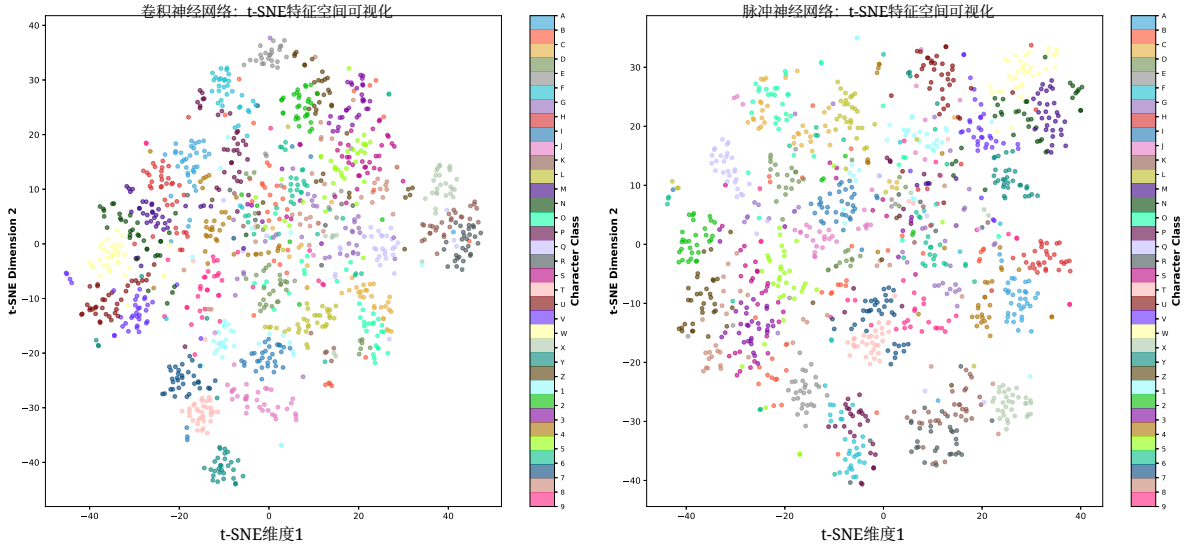

图 8. 使用 t-SNE 对从一个验证集划分中提取的倒数第二层特征（144 维）进行特征空间可视化。**左：** 卷积神经网络特征。**右：** 脉冲神经网络特征。

6. 未来方向

这些基线结果确立了 STEMNIST 作为触觉识别领域具有挑战性的神经形态基准，并表明其与 ST-MNIST 相比仅存在少数性能差距，因此为算法创新和数据集扩展提供了充足的机会。未来

的发展方向涵盖三个互补领域：采用基于事件的触觉传感的现实世界应用；扩大覆盖范围和模态的数据集扩展；以及利用时空脉冲动力学的算法开发。

6.1. 潜在应用

该数据集具有事件驱动特性，适用于多种神经形态计算场景。该数据集可助力面向视障用户的辅助技术（如触觉盲文阅读器或触觉文本界面），通过包含 35 个类别的字母数字字符集，弥合传统基于视觉的光学字符识别（OCR）与触摸识别之间的差距。人机界面和假肢可以利用神经形态触觉反馈，以更低的延迟和更少的能耗实现闭环感觉恢复。利用该数据集在神经形态硬件上训练高能效的事件识别流水线，还可为非结构化环境中的灵巧操作任务提供实时触觉反馈，助力机器人触觉感知系统。

6.2. 数据集扩展

未来可通过增加类别多样性来扩展这项工作，从而提升泛化能力。纳入小写字母“a”至“z”后，类别数可增至 61；再加入零（0），即可与 EMNIST 完全对齐。这将支持大小写混合手写识别。加入天城文、汉字笔画或孟加拉数字等多语言字符，可将应用拓展至全球范围。通过摄像机采集工具轨迹，将这些触觉记录与视觉信息同步，可实现多模态数据融合，用于跨模态学习和传感器融合基准测试。

6.3. 算法发展

CNN 与 SNN 之间有限的性能差异促使我们探索多种算法创新。采用残差连接、批归一化和自适应漏积分发放阈值的更深层架构，能够更充分地利用空间特征和时序动态。循环神经网络（如 LSTM-SNN 混合模型）可以对笔画序列中的长程时间依赖关系进行建模，从而捕捉笔画顺序和笔画间时序。针对脉冲序列的注意力机制，例如在适当数量的时间分箱上应用时间注意力，或在 16×16 阵列上应用空间注意力，可以动态调整判别特征的权重。在无标签触觉数据上进行自监督预训练，例如对笔画嵌入进行对比学习，可以在微调之前改善特征表示。图神经网络可以将节点视为触觉像素、边视为时间连通性，通过建模笔画拓扑来捕捉结构先验。借助片上学习规则（例如脉冲时序依赖可塑性、资格迹）在神经形态硬件上部署，或可实现实时且高能效的

训练。混合人工神经网络-脉冲神经网络通路结合了卷积神经网络的空间特征提取能力与脉冲神经网络的时序处理能力，从而能够发挥各自的互补优势。最后，少样本学习和元学习可用于适应参与者特有的手写变化，同时只需少量用户校准数据。

7. 结论

我们推出了 STEMNIST，这是首个大规模神经形态触觉数据集。它将原始 ST-MNIST 框架从 10 个数字扩展至 35 个字母数字类别（A-Z 和 1-9），填补了神经形态计算领域基于事件的触觉基准测试中的关键空白。该数据集包含从 34 名参与者处采集的 7,700 个样本，使用定制的 16×16触觉传感器阵列以 120 赫兹运行采集，并通过自适应阈值处理编码生成了 1,005,592 个脉冲事件。与早期仅包含数字的数据集相比，这一扩展显著提高了分类难度，并在触觉领域带来了 EMNIST 视觉字符识别挑战的复杂性。采用传统卷积神经网络和脉冲神经网络进行的基准分类实验，既展现了该数据集的判别能力，也证明了双通道脉冲编码在仿生触觉处理中的有效性。通过五折交叉验证进行系统的超参数优化（卷积神经网络达到 89.72%$\pm$ 0.61%，脉冲神经网络达到 88.25% $\pm$ 0.98%），并在五个独立的数据划分上进行鲁棒性评估（卷积神经网络为 90.36% $\pm$ 0.70%，脉冲神经网络为 87.81% $\pm$ 1.20%），我们的最终模型分别取得了 90.91% 和 89.16% 的测试准确率，卷积神经网络与脉冲神经网络的性能差距仅为 1.75%；相比之下，先前的神经形态字符识别研究报告的典型差距为 8%+。两种神经网络的测试准确率均显著高于随机猜测，甚至超过了 ST-MNIST 分类性能的卷积神经网络和脉冲神经网络基线。STEMNIST 弥合了简化的触觉数字识别与更复杂的现实触觉交互场景之间的差距，后者需要区分字母数字字符。基于事件的格式、不受限制的空间变化以及丰富的时序动态，使该数据集成为神经形态硬件部署以及脉冲时序依赖可塑性、循环脉冲神经网络和注意力机制等仿生学习算法的理想测试平台。本文提供的标准化训练-测试划分、丰富的统计数据和基准测试，有助于确保在传统深度学习、神经形态计算和混合架构中开展触觉识别研究评估的可复现性。此外，未来还将扩展该数据集的范围和适用性，纳入小写字母、多语言字符和多模态融合。我们希望 STEMNIST 能推动机器人学、辅助技术和人机界面领域节能型触觉感知的发展，助力实现神经形态人工智能这一更广阔的愿景，使其效率和适应能力可与生物感觉处理相媲美。

8. 数据集可用性

STEM NIST 数据集及相关文档可在此处获取：

9. 致谢

作者谨向所有参与 STEMNIST数据集收集工作的人员，以及在项目构思阶段于2025年班加罗尔神经形态工程研讨会上参与讨论的各位致以谢意。本研究获中国香港特别行政区研究资助局资助（项目编号：CityU 11212823）。

参考文献

[1]S. J. Lederman 和 R. L. Klatzky，《触觉感知：教程》，《注意、知觉与心理物理学》，第 71 卷，第 7 期，第 1439–1459 页，2009 年。 [在线]。可访问：https://pubmed.ncbi.nlm.nih.gov/19801605/[2] J. J. Gibson，《关于主动触觉的观察》，《心理学评论》，第 69 卷，第 6 期，第 477 页，1962 年。[3] J. Jin、S. Wang、Z. Zhang、D. Mei 和 Y. Wang，《机器人应用中用于物体属性识别、操作和人机交互的柔性触觉传感器研究进展》，《软科学》，第 3 卷，第 1 期，第 N–A 页，2023 年。[4] P. Roberts、M. Zadan 和 C. Majidi，《用于机器人学的柔性触觉传感皮肤》，《当代机器人学报告》，第 2 卷，第 3 期，第 343–354 页，2021 年。[5] W. W. Lee、S. Kukreja 和 N. Thakor，《用于研究神经形态触觉中时空特征的千赫兹千触元触觉传感器阵列》，载于 2015 年 IEEE生物医学电路与系统会议（BioCAS）。IEEE，2015 年，第 1–4 页。[6] J. A. Pruszynski 和 R. S. Johansson，《一级触觉神经元中的边缘方向处理》，《自然·神经科学》，第 17 卷，第 1404–1409 页，2014 年。[7] Y. LeCun、L. Bottou、Y. Bengio 和 P. Haffner，《将基于梯度的学习应用于文档识别》，《IEEE会刊》，第 86 卷，第 11 期，第 2278–2324 页，2002 年。[8] G. Cohen、S. Afshar、J. Tapson 和 A. Van Schaik，《EMNIST：将 MNIST 扩展至手写字母》，载于 2017 年国际神经网络联合会议（IJCNN）。IEEE，2017 年，第 2921–2926 页。[9] H. H. See、B. Lim、S. Li、H. Yao、W. Cheng、H. Soh 和 B. C. Tee，《ST-MNIST——脉冲触觉MNIST神经形态数据集》，arXiv 预印本 arXiv:2005.04319，2020 年。[10]S. Sundaram、P. Kellnhofer、Y. Li、J.-Y. Zhu、A. Torralba 和 W. Matusik，《使用可扩展触觉手套学习人类抓握特征》，《自然》，第 569 卷，第 7758 期，第 698–702 页，2019 年。[11] S. F. M¨uller-Cleve、V. Fra、L. Khacef、A. Pequeˆno-Zurro、D. Klepatsch、E. Forno、D. G. Ivanovich、S. Rastogi、G. Urgese、F. Zenke 等，《盲文字符阅读：神经形态硬件上时空模式识别的基准测试》，《神经科学前沿》，第 16 卷，第 951164 页，2022 年。[12] P. J. Grother，《NIST 特殊数据库 19》，《手写表格与字符数据库》，美国国家标准与技术研究院，第 10 卷，第 69 页，1995 年。[13] T. Serrano-Gotarredona 和 B. Linares-Barranco，《Poker-DVS 和 MNIST-DVS：它们的历史、制作方法及其他细节》，《神经科学前沿》，第 9 卷，第 481 页，2015 年。[14] G. Orchard、A. Jayawant、G. K. Cohen 和 N. Thakor，《使用扫视将静态图像数据集转换为脉冲神经形态数据集》，《神经科学前沿》，第 9 卷，第 437 页，2015 年。[15] C. Posch、D. Matolin 和 R. Wohlgenannt，《一种异步时间型图像传感器》，载于 2008 年 IEEE国际电路与系统研讨会（ISCAS）。IEEE，2008 年，第 2130–2133 页。

[16] L. R. Iyer、Y. Chua 和 H. Li，《神经形态 MNIST 是否真的具有神经形态特性？分析神经形态数据集在时域中的判别能力》，《神经科学前沿》，第 15 卷，第 608567 页，2021 年。[17] A. Amir、B. Taba、D. Berg、T. Melano、J. McKinstry、C. Di Nolfo、T. Nayak、A. Andreopoulos、G. Garreau、M. Mendoza 等，《一种低功耗、完全基于事件的手势识别系统》，载于*IEEE*计算机视觉与模式识别会议论文集，2017 年，第 7243–7252 页。[18] D. Kong、Y. Lu、S. Zhou、M. Wang、G. Pang、B. Wang、L. Chen、X. Huang、H. Lyu、K. Xu等，《深度学习赋能的稀疏单元超分辨率触觉传感器阵列》，《科学进展》，第 11 卷，第 27 期，第 eadv2124 页，2025 年。[19]B. Rueckauer、I.-A. Lungu、Y. Hu、M. Pfeiffer 和 S.-C. Liu，《将连续值深度网络转换为用于图像分类的高效事件驱动网络》，《神经科学前沿》，第 11 卷，第 682 页，2017 年。[20] J. K. Eshraghian、M. Ward、E. O. Neftci、X. Wang、G. Lenz、G. Dwivedi、M. Bennamoun、D. S. Jeong 和 W. D. Lu，《借鉴深度学习经验训练脉冲神经网络》，《*IEEE*会刊》，第 111 卷，第 9 期，第 1016–1054 页，2023 年。[21] F. Vega-Bermudez、K. O. Johnson 和 S. S. Hsiao，《人体触觉模式识别：主动触觉与被动触觉、速度效应及混淆模式》，《神经生理学杂志》，第 65 卷，第 3 期，第 531–546 页，1991 年。[22] L. v. d. Maaten 和 G. Hinton，《使用 t-SNE 可视化数据》，《机器学习研究杂志》，第 9 卷，11 月，第 2579–2605 页，2008 年。