arXiv:2601.08447v1 [cs.NE] 13 Jan 2026

基于睡眠的稳态正则化：稳定循环脉冲神经网络中的脉冲时序依赖可塑性

Andreas Massey Aliaksandr Hubin StefanoNichele Solve Sæbø

2026年1月14日

摘要

脉冲时序依赖可塑性（STDP）为脉冲神经网络（SNNs）提供了一种符合生物学原理的学习机制；然而，具有循环连接的架构中的赫布式权重更新会引发病理性权重动态：权重无界增长、灾难性遗忘以及表征多样性丧失。我们提出一种受突触稳态假说启发的神经形态正则化方案：定期进入离线阶段，在此期间抑制外部输入，使突触权重随机衰减至稳态基线，并通过自发活动促进记忆巩固。我们证明，这种睡眠-觉醒周期能够防止权重饱和，同时保留已学习的结构。实证结果表明，在我们的STDP-SNN 模型中，较短至中等时长的睡眠（占训练时长的 10–20%）能够提升类似MNIST的基准测试上的稳定性，且无需针对特定数据进行超参数调优。相比之下，同样的睡眠干预对替代梯度脉冲神经网络（SG-SNN）没有带来可测量的益处。总而言之，这些结果表明，定期进行基于睡眠的重整化或许是稳定神经形态系统中局部赫布学习的一种基本机制，同时也说明，将此类协议与现有的基于梯度的优化方法相结合时需要格外谨慎。

关键词： 脉冲神经网络、脉冲时序依赖可塑性、稳态、在线学习、神经形态计算

1 引言

脉冲神经网络（脉冲神经网络）通过离散脉冲事件的时空分布编码信息，从而能够在神经形态器件上实现高效的事件驱动计算[Su等人， 2023]。与稠密的连续值架构相比，这种时间稀疏性可使功耗降低数个数量级 [Isik等人， 2024]。然而，脉冲神经网络的学习景观仍然效率低下且缺乏深入理解 [Akil等人，2021], 尤其是在生物学上合理的局部学习规则的稳定性和收敛性方面 [Stratton等人， 2022]。

脉冲时序依赖可塑性（STDP）自然成为脉冲神经网络的典型局部学习规则 [Song等人， 2000]。权重的变化完全取决于突触前与突触后脉冲之间毫秒级的时序差异：当突触前活动在时间上先于突触后放电时，权重增强；当顺序相反时，权重减弱。这种遵循因果关系的更新规则不仅符合生物学规律，在理论上也有其依据，因为它通过时间邻接实现信用分配，并源自脉冲巧合检测原理 [Jayabal 等，2024年，Chauhan等人， 2018]。然而，在对时间处理和内部状态维持至关重要的循环架构中，脉冲时序依赖可塑性会表现出严重的问题。循环连接性通过正反馈回路放大初始权重扰动，从而导致三种相互交织的失效模式：（i）**权重饱和**，共同激活的突触发生不受限制的赫布式增强，使权重趋向无界值或截断上限 [Gilson等人， 2010]。前者会造成灾难性不稳定，而后者则会消除进一步学习的能力；（ii）**表征坍塌**，权重耦合会诱发虚假相关，侵蚀判别性结构 [Akil等人， 2021]；以及（iii）**灾难性遗忘**，序列任务学习会覆盖先前学到的表征 [Tadros等人， 2022]。

该领域主要采用了两种稳定脉冲时序依赖可塑性的方法，但各自都伴随着显著代价。第一种方法是，稳态机制 采用全局归一化方案：突触缩放会重新归一化权重分布，使其趋向放电率目标；内在可塑性会调整神经元阈值；结构可塑性则会修改连接性 [Carlson 等， 2013]。尽管这些方法有效，但它们牺牲了严格局部性——需要全局放电率统计——并引发了对生物学不合理性的担忧——突触缩放通常被建模为乘性全局因子，而非局部机制 [Qiao 等， 2016]。第二种方法是，替代梯度 （SG）方法通过在反向传播过程中使用平滑的膜电位代理函数替代不连续性，从而规避脉冲不可微性，并实现基于误差的权重更新 [Huh 和 Sejnowski， 2017]。因此，SG 方法可以与标准的 ℓ_1 或 ℓ_2正则化相结合，并使脉冲神经网络在 MNIST 上取得较高准确率（$>$97%），但其代价高昂，违背了采用脉冲神经网络的初衷：它们需要误差反向传播（牺牲局部性）、需要大量内存进行梯度存储、使用损失函数，并且消除了事件驱动稀疏性——后者可以说是神经形态计算最重要的优势之一 [Li 等，2024年，Wang 等， 2023]。

我们提出第三种方法，其理论基础是突触稳态假说（SHy）。这一神经科学理论认为，睡眠通过活动加权突触重新归一化来维持稳定性并巩固学习 [Tononi 和 Cirelli， 2014]。在生物睡眠期间，高活动突触的下调幅度大于低活动突触，从而既实现权重重新归一化，也能选择性修剪与任务无关的连接；此前的睡眠启发神经形态模型（如 Tadros 等、[2022], Kinoshita 等、[2025], Nere 等， [2013]）已对此进行探索。我们将这一原理移植到神经形态算法中，采用类似 Thiele 等人提出的周期性睡眠/清醒周期 [2017]：在睡眠阶段，（1）抑制外部输入；（2）突触权重受控地按照幂律衰减，趋向稳态基线；（3）受 Ma 等人 [2023] 以及 Deperrois 等人 [2022], 启发，由内在膜噪声单独驱动的自发神经活动，使脉冲时序依赖可塑性得以持续应用，并通过隐式回放实现巩固。已有研究表明，后者能够提升类内一致性，并降低数据过拟合风险 [Deperrois 等，2022，Hoel，2021，Tadros 等， 2022]。重要的是，该机制保留了完全的局部性（只需突触前和

突触后统计量），几乎不增加计算开销，并且仍与神经形态硬件兼容。此外，我们强调基于睡眠的稳态调节与不同学习范式相互作用时的一个关键区别。睡眠阶段在机制上与赫布学习兼容，因为二者都基于局部活动统计量运作。相比之下，在脉冲神经网络中，经过评估的脉冲时序依赖可塑性/反向传播（BP）混合方法寥寥无几，而现有方法通常试图通过引入全局奖励信号或概率学习规则来绕过局部性约束。这反映出一种根本性的不兼容：脉冲时序依赖可塑性从其构造上就无法直接支持反向传播所依赖的全局误差梯度。我们的实验结果与这一观点一致。尽管睡眠期间的自发活动和噪声驱动的权重降低原则上可能帮助替代梯度训练的脉冲神经网络摆脱较差的局部最优解，但我们没有发现证据表明这能转化为学习性能的系统性提升[1]。

我们的贡献主要有三点。首先，我们提出了一种有原则的睡眠-觉醒机制，将稳态权重衰减与自发重放相结合，并推导出稳定性条件，表明周期性重新归一化能够限制权重增长并维持表征结构，这与 [Capone et al., 2019] 的结果一致。其次，我们提出了防止性能饱和的驱动因素：通过重放进行巩固，以及维持稳态平衡，并给出了相应的机制性解释。第三，我们在 MNIST 系列基准上进行了实证验证，结果表明，在我们的基于 STDP 的模型中，中等睡眠时长（基线训练时长的 10%–20%）能够稳定提升性能；而对于替代梯度脉冲神经网络模型，无睡眠机制并未带来性能增益。

本文其余部分的组织如下：第 2 节介绍本研究所采用的方法，包括所研究的两种脉冲神经网络实现、它们的架构、学习规则以及睡眠-觉醒正则化协议。作为初步验证步骤，我们首先在一个受控的玩具分类任务上评估 STDP-SNN，以验证无监督 STDP 下稳定的学习动态，随后使用两个模型扩展到 MNIST 系列基准。第 3 节介绍结果。第 4 节讨论其对神经形态硬件的更广泛影响及未来研究方向。

2 方法

2.1 概述

我们比较了两种脉冲神经网络实现，它们在架构和学习范式上有所不同，但都采用了相同的睡眠启发式稳态调节机制。第一个模型——脉冲时序依赖可塑性脉冲神经网络（STDP-SNN）——采用循环兴奋-抑制架构，并通过局部脉冲时序依赖可塑性和反向脉冲时序依赖可塑性学习规则进行训练。该模型从头开始在 `numpy` 中实现，并借助 `numba`加速，明确模拟漏积分发放神经元动力学、突触可塑性以及睡眠—觉醒周期。在睡眠阶段，突触权重会经历噪声驱动的重新归一化，同时脉冲时序依赖可塑性仍保持活跃，从而实现类似重放的巩固。

[1]已有一些尝试将替代方法与脉冲时序依赖可塑性相结合，例如 Ma 等人开发的噪声脉冲神经网络（NSNN）模型 [2023]。该模型利用噪声膜活动来确保网络具有更强的鲁棒性，并将损失输入膜电位，其方式与其他替代方法类似。

第二个模型——替代梯度脉冲神经网络（SG-SNN）——使用 `snntorch`框架实现，由前馈线性层与漏积分发放单元构成，并通过替代梯度反向传播进行训练。在这一设置中，睡眠通过定期中断任务驱动训练来引入，以应用相同的权重衰减和噪声机制，但不会持续进行脉冲时序依赖可塑性。

因此，睡眠在这两个模型中发挥着截然不同的功能作用：在 STDP-SNN 中促进基于重放的巩固，而在替代梯度机制中则仅作为随机正则化项。

STDP-SNN 模型的全部代码已在 Massey [2025]公开发布 .

2.2 STDP模型

STDP-SNN 是我们主要的生物学启发模型，它将局部脉冲时序依赖可塑性与周期性的、受睡眠启发的稳态正则化机制相结合。网络完全采用局部学习规则进行无监督训练，并在交替进行的睡眠阶段通过带噪权重重新归一化来维持稳定性。这种设计使模型能够学习类别结构化表征，同时避免基于循环STDP的网络通常会出现的病理性权重动态。

2.2.1 架构

构建网络时，我们借鉴了此前关于脉冲神经网络无监督聚类的研究 [Zenke et al., 2015], ，该研究表明，采用 STDP 训练的循环兴奋–抑制架构能够自组织形成类别选择性表征。这类模型特别适合用于研究赫布学习中的稳定性，因为循环兴奋会放大微小的权重扰动，因此若缺乏明确的正则化，就容易出现病理性动态。通过采用类似架构，我们有意将模型置于一个具有挑战性的机制中；在这种机制下，众所周知，不受控制的 STDP 会失效，因此我们可以直接评估受睡眠启发的稳态机制能否在不牺牲局部性的前提下稳定学习。

为确保结果具有可比性，我们选择了以下网络参数：将泊松过程输入编码为脉冲层（$N_{\text{in}} = 225$ 个神经元）、循环兴奋性读出层（$N_{\text{exc}} = 200$）和侧向抑制层（$N_{\text{inh}} = 50$）。请注意，Zenke 等人 [2015]还在侧向抑制层中加入了循环权重。由于这会增加超参数估计的复杂度，我们没有采用这一做法。连接是随机且稀疏的： $P_{\text{in}\to\text{exc}} = 10\%$、 $P_{\text{exc}\to\text{exc}} = 15\%$、 $P_{\text{exc}\to\text{inh}} = 20\%$、 $P_{\text{inh}\to\text{exc}} = 25\%$，并且没有自连接。该架构体现了感觉皮层的关键特性：前馈丘脑输入、循环兴奋性动力学和局部抑制控制。

2.2.2 神经元模型

每个神经元通过漏积分发放（LIF）动力学整合突触输入：

$$\tau_m \frac{dU_i(t)}{dt} = -(U_i(t) - U_{\text{rest}}) + R_m I_i^{\text{syn}}(t) + \xi_i(t), \tag{1}$$

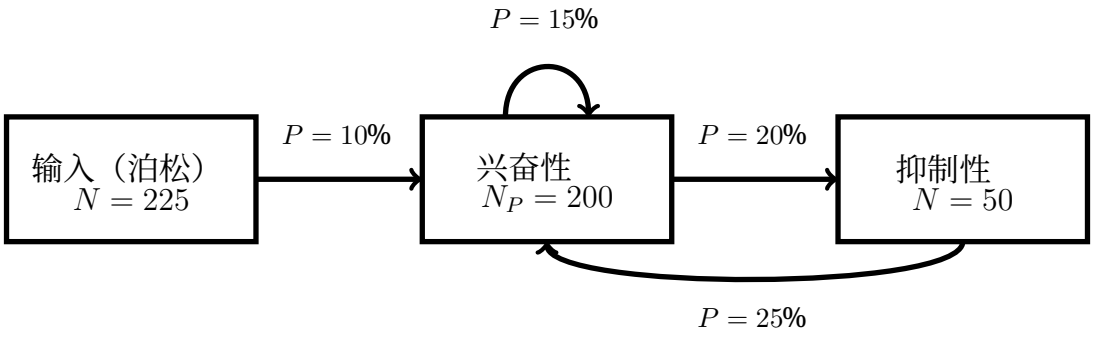


图1：网络架构：包含前馈输入（$P = 10\%$）、循环兴奋（$P = 15\%$）和侧向抑制（$P = 25\%$）的三层脉冲神经网络。稀疏连接支持事件驱动计算。

其中，$U_i(t)$ 表示膜电位，τ_m 表示时间常数，R_m 表示膜电阻，U_{rest} 表示静息电位，$I_i^{\text{syn}}(t) = \sum_j w_{ij} S_j(t)$ 表示突触电流，$\xi_i(t) \sim \mathcal{N}(0, \sigma^2)$ 表示内在噪声。当 $U_i(t) \geq U_{\text{th}}^i(t)$ 时，神经元 i 会发放一个脉冲（$S_i(t) = 1$），并重置至 U_{reset}。脉冲阈值会自适应调整（以防止失控放电）：

$$\frac{dU_{\text{th}}^i(t)}{dt} = U_{\text{th}}(0) + \alpha_i,$$
$$\frac{d\alpha_i}{dt} = -\frac{\alpha_i}{\tau_{th}} + S_i(t)\delta,$$

基线脉冲阈值为 $U_{\text{th}}(0)$，时间常数为 τ_{th}，跳变幅度为 δ。例如，采用 $\delta = 3$ 时，我们会降低持续脉冲序列出现的概率；也就是说，每当神经元 i 发放脉冲时，脉冲阈值就会升高 3 mV（每个突触后脉冲增加 $\sim 6.5\%$）。

2.2.3 学习规则

兴奋性突触通过脉冲时序依赖可塑性（STDP）进行更新：

$$\Delta w_{ij}^{\text{exc}}(\Delta t) = \eta_{\text{exc}} \begin{cases} A_+ \exp(-\Delta t/\tau_+) & \text{if } \Delta t \geq 0 \\ -A_- \exp(\Delta t/\tau_-) & \text{otherwise,} \end{cases} \tag{2}$$

其中，$\Delta t = t_{\text{post}} - t_{\text{pre}}$ 表示脉冲时序，η_{exc} 表示学习率，$A_\pm$、$\tau_\pm$ 分别表示学习幅度和时间常数。抑制性突触采用反向脉冲时序依赖可塑性（iSTDP）：

$$\Delta w_{ij}^{\text{inh}}(\Delta t) = \eta_{\text{inh}} \begin{cases} -A_+ \exp(-\Delta t/\tau_+) & \text{if } \Delta t \geq 0 \\ A_- \exp(\Delta t/\tau_-) & \text{otherwise,} \end{cases} \tag{3}$$

其中 η_{inh} 是学习率（或反学习率）。反向脉冲时序依赖可塑性确保抑制性神经元专门针对高活跃度的兴奋性神经元，从而实现稳态负载均衡。

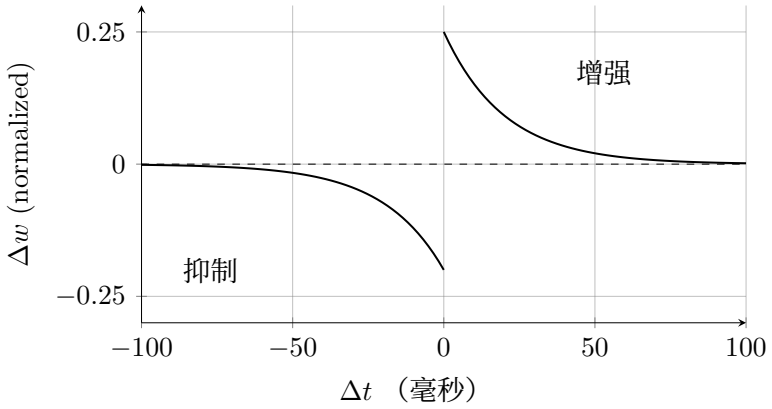


图 2：STDP 学习窗口：当突触前脉冲先于突触后脉冲时，权重增强（ Δt 为正）；否则权重减弱（ Δt 为负）。

2.2.4 睡眠协议

本文的核心创新在于实现了由睡眠比例控制的周期性睡眠–清醒时间表。时间被划分为固定长度的睡眠间隔。每个间隔开始时，模拟器进入硬暂停 睡眠阶段，并执行最多

$$T_{\text{sleep}} = \max(1, \lfloor \texttt{sleep-ratio} \cdot \texttt{sleep interval} \rceil)$$

虚拟睡眠迭代，而不推进实际模拟时间。

睡眠期间，权重通过幂律动力学衰减，同时保持相对排序：

$$w_i(t+1) = w_{\text{tgt}} \left(\frac{w_i(t)}{w_{\text{tgt}}} \right)^{\lambda}, \tag{4}$$

其中， w_{tgt} 是函数趋近的目标， λ 表示衰减率。睡眠期间，感觉输入受到抑制（所有刺激脉冲均设为零），膜动力学由内在噪声驱动，产生自发脉冲活动，并可持续引发 STDP 更新。此外，通过睡眠算子将突触权重重新归一化至目标幅值。

当权重之和 $\mathcal{W}_x(t)$ 达到 $w_x^{\min}$ 时，睡眠结束

$$\mathcal{W}_x(t) \le w_x^{\min} = \alpha_{\text{base}} \cdot \mathcal{W}_x(0), \quad \alpha_{\text{base}} < \alpha_{\text{trig}} \tag{5}$$

或者，当睡眠间隔 超出且尚未达到 $w_x^{\min}$ 时。

2.2.5 评估

为计算准确率，我们首先汇总每个训练样本的平均脉冲率，并将特征向量标准化为 $\mu = 0$ 且方差为 1。然后，使用主成分分析（PCA）对标准化后的特征向量进行降维，同时保留 95% 的方差。最后，我们使用变换后的特征训练多项逻辑回归（MLR）。训练好的分类器随后可对网络处理过的未见测试数据进行类别预测。

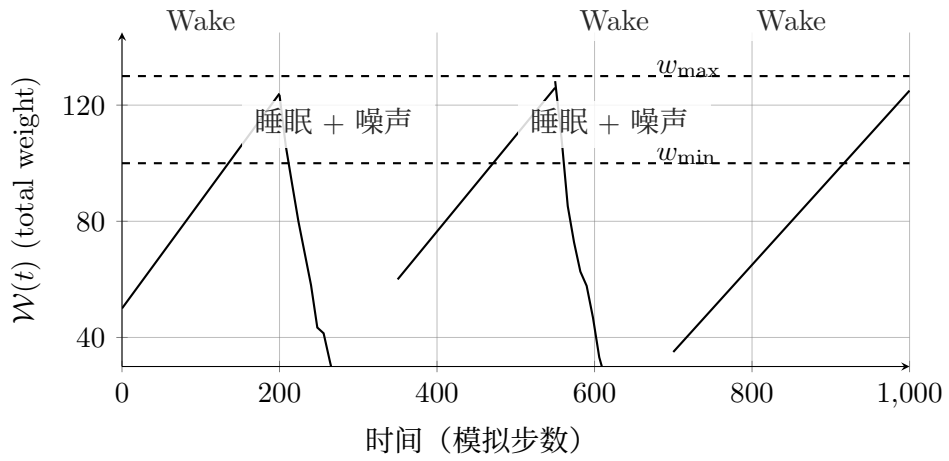


图 3：权重动力学：清醒阶段表现为由脉冲时序依赖可塑性驱动的确定性线性增长；睡眠阶段则表现为幂律正则化（公式 4），并叠加类似于高斯随机变量实现值的随机噪声，以表示巩固期间的自发神经活动。系统在清醒（输入驱动）与睡眠（稳态噪声衰减）状态间稳定循环，使权重保持在合理范围内。

2.3 玩具几何实验

为了评估无监督脉冲时序依赖可塑性脉冲神经网络的聚类行为，并将基于睡眠的正则化效果与数据集特定混杂因素区分开来，我们设计了一项受控的玩具分类任务。该实验可作为诊断性验证，使我们能够在评估大规模图像基准测试上的性能之前，验证网络在局部赫布学习下能否形成稳定且具有区分性的内部表征。

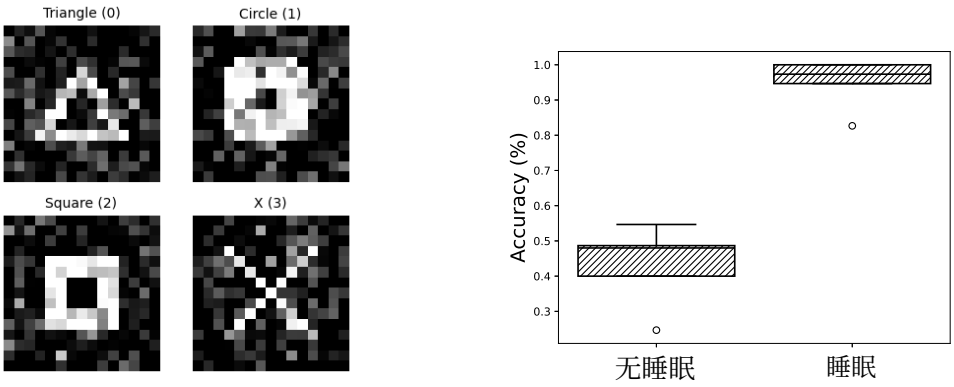
2.3.1 数据

我们使用了一个合成几何数据集，其中包含 7,100 张分辨率为 15×15 的灰度图像（图 4a）。每张图像属于四种几何类别之一，并受到加性高斯噪声 $\zeta_i \sim \mathcal{N}(0, 0.02)$ 的干扰，以引入可控的变化，同时保留类别身份。每个刺激呈现给网络的时间为 100 毫秒。

数据集被分别划分为 6,000/100/1,000 个用于训练、验证和测试的样本。构建的小批次包含均衡的类别表征，确保学习动态反映的是聚类行为，而非类别频率效应。

2.3.2 设计

为评估睡眠的效果，我们比较了两种训练机制下的模型性能：睡眠和无睡眠。在睡眠条件下，突触权重会定期进行正则化；而在无睡眠条件下，训练则持续进行，不包含这些阶段。在两种机制下，突触权重都通过符号钳制被限制在各自的功能组（兴奋性或抑制性）内。符号钳制通过阻止权重跨越零边界、从而改变其突触后效应，来强制保持固定极性。



(a) (b)

图4：（a）由四个几何类别（三角形、圆形、正方形和十字形）组成的训练集，每个类别均加入高斯噪声 $\mathcal{N}$，0.02。所有训练样本均使用相同的基本形状，噪声作为引入逐样本方差的随机成分。（b）显示了采用无睡眠和睡眠协议时达到的准确率。

2.3.3 结果

我们观察到，睡眠协议与符号钳制相结合，成功稳定了突触权重，而单独使用符号钳制则未能做到这一点（见图5）。在15个批次内，睡眠协议实现了较高的准确率（均值$= 94.93$、最小值$= 82.67$ 和最大值$= 100.00$）；相比之下，无睡眠 配置的性能明显较差：（均值 $=$ 43.20，最小值 $= 24.67$，最大值 $= 54.67$）。在图5a中，训练过程中抑制性组和兴奋性组的权重均值及其最小值/最大值范围保持一致，而在无睡眠 条件下则并非如此（见图5b）。在批次内，这一趋势也同样明显：图5c展示了批次内的动态变化；启用睡眠协议后，权重保持在有界范围内，同时仍呈现出良好的异质性。与此同时，图5d表明，如果没有适当的正则化方法，权重会迅速爆炸。

相比之下，移除睡眠阶段会导致快速失稳。即使在单个批次内，权重也开始发散；请注意睡眠和无睡眠条件之间 y 轴刻度的差异，最终导致学习过程受阻。这种性能下降体现在图6中，其中无睡眠模型6b显示出受损的聚类结构。由于睡眠协议在保持变异性的同时抑制了权重增长，因此能够带来更稳定的性能，并在本例中实现清晰的类别分离。

2.4 替代梯度模型

虽然前文各节重点讨论了如何稳定循环兴奋性/抑制性脉冲神经网络中的完全局部脉冲时序依赖可塑性，但目前仍不清楚观察到的性能提升究竟是赫布学习所特有的，还是体现了更普遍的正则化效应。为区分这两个因素，我们引入了第二种采用全局误差驱动优化进行训练的模型。该替代梯度模型为以下方面提供了重要参照：（i）时间反向传播下的预测性能；（ii）相同的睡眠干预是否也适用于基于梯度的脉冲神经网络训练。

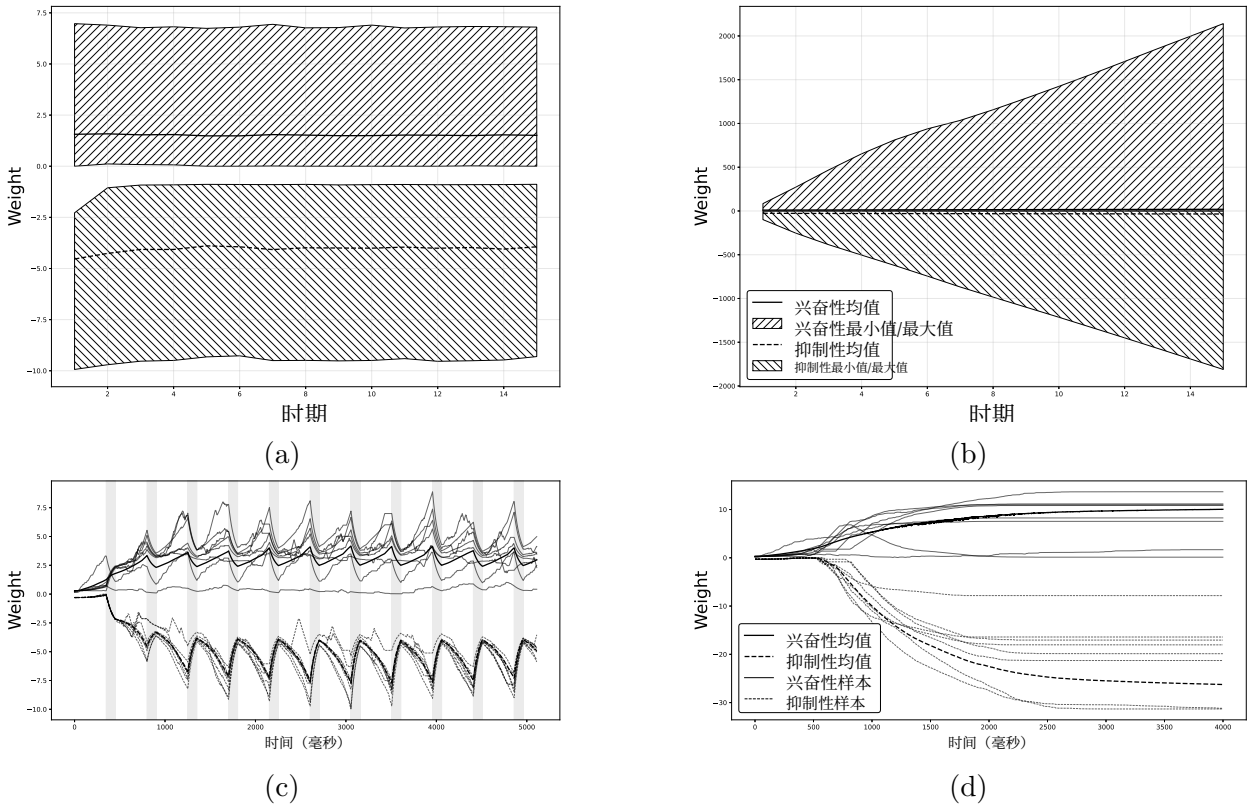
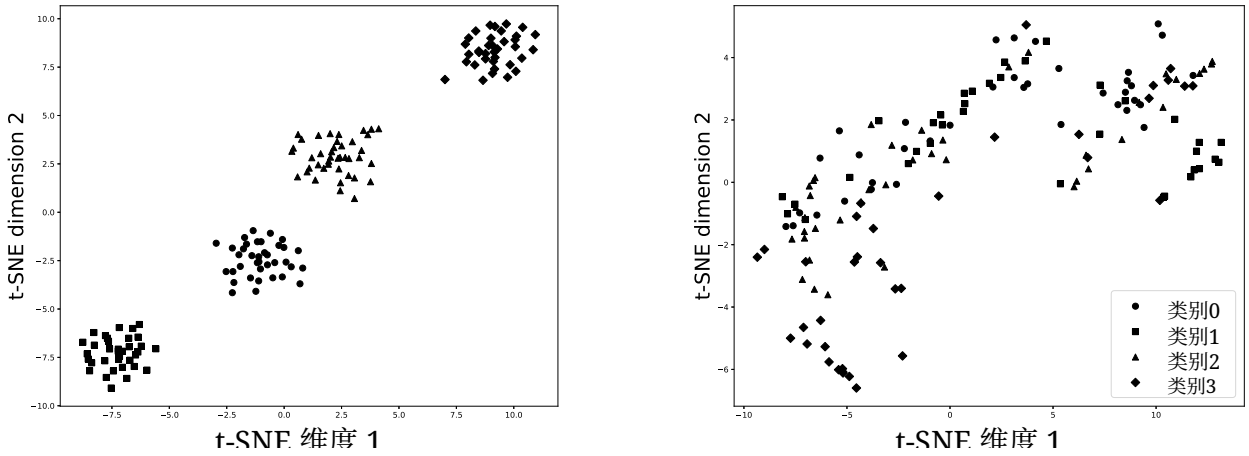


图5：（a）表示采用睡眠正则化进行训练时，各批次中的兴奋性和抑制性权重平均值。灰色区域表示每个批次权重的最小值和最大值。（b）表示相同的数据，但未采用睡眠来稳定学习。（c）显示启用睡眠协议时，单个批次训练会话中的权重变化。横轴表示以毫秒为单位的时间，纵轴表示权重变化。与（d）类似，点线表示抑制性权重，实线表示兴奋性权重的最小值、最大值和平均值。



(a) (b)

图6：睡眠 (a) 与无睡眠 (b) 训练范式下的 t-SNE 对比图。两幅图均展示了按训练样本分箱的脉冲率向量的 t-SNE 可视化结果。每种几何形状对应一个不同的类别，每个点代表嵌入到 t-SNE 降维空间中的一个测试观测。

2.4.1 架构

该 n etwork 是一个具有脉冲动力学的两层多层感知机（MLP） cs:

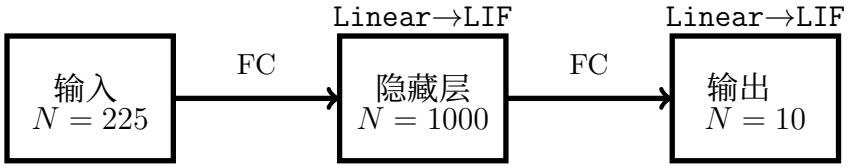


图 7：替代梯度基线：前馈脉冲多层感知机（225–1000–10），由全连接突触和漏积分发放动力学构成（`nn.Linear`→`snn.Leaky`）。

由两个全连接层（`Linear`）以及漏积分发放（LIF）神经元（`snn.Leaky`）构成。该架构纯粹采用前馈结构（即没有像 STDP 模型那样的循环连接）。每张静态输入图像呈现 $T = 100$ 离散仿真步；时间维度来自 LIF 状态动力学，而非随时间变化的基于脉冲的输入。

2.4.2 神经元模型

每个脉冲层都使用膜衰减参数为 $\beta = 0.95$ 的离散时间漏积分发放神经元。令 $U(t)$ 表示膜电位，并令 $S(t) \in \{0, 1\}$ 表示时间步 t 时发放的脉冲。在每一层中，膜电位按照标准的漏积分发放递推进行更新（参见公式 1），并通过阈值穿越生成脉冲：

$$S(t) = H(U(t) - U_{\text{thr}}),$$

其中， $H(\cdot)$ 是海维赛德阶跃函数。

2.4.3 学习规则

作为比较，我们还实现了替代梯度（SG）训练，在该训练过程中，反向传播期间使用平滑代理函数替代不可微的脉冲函数。前向传播：神经元通过确定性 LIF 阈值穿越规则生成脉冲（公式 1，其中使用了平移后的海维赛德阶跃函数）：

$$S_i(t) = \begin{cases} 1 & \text{if } U_i(t) \geq U_{thr} \\ 0 & \text{if } U_i(t) < U_{thr}. \end{cases} \tag{6}$$

反向传播过程中，脉冲非线性由反正切形替代函数$\sim S_i(t)$取代，从而平滑近似阈值函数并使梯度流得以传播。

$$\tilde{S}_i(t) = \frac{1}{\pi}\arctan\left(\pi U_i(t)\frac{\alpha}{2}\right) \tag{7}$$

基于 Fang 等人 [2021], ，其中 α 是决定替代梯度形状的常数因子。随后，通过对任务损失（$\mathcal{L}$）执行标准梯度下降来更新权重：

$$w_{ij}(t+1) \leftarrow w_{ij}(t) - \alpha_{\mathrm{SG}} \frac{\partial \mathcal{L}}{\partial w_{ij}(t)}, \tag{8}$$

其中 α_{SG} 是 SG 训练的学习率。

2.4.4 睡眠协议

为探究基于睡眠的正则化是否也能使经 SG 训练的脉冲神经网络受益，我们采用上述相同的睡眠-清醒方案。在清醒阶段，网络处理任务驱动输入，并禁用脉冲时序依赖可塑性（所有更新均由梯度控制）。在睡眠阶段，我们抑制外部输入（$I^{\mathrm{task}} = 0$），维持膜电位中的内在噪声，并采用与（4）中相同的随机权重衰减。

与赫布训练不同，赫布训练中脉冲时序依赖可塑性在睡眠期间仍会持续进行，并实现基于重放的巩固；而在 SG 训练中，睡眠机制会提供随机扰动，或有助于摆脱误差景观中的次优局部极小值。不过，乘法衰减也会干扰基于梯度的权重更新。因此，净效应预计会在很大程度上取决于睡眠强度：睡眠率较低时，噪声驱动的探索或许能通过帮助摆脱损失函数的局部极值带来适度益处；睡眠率较高时，衰减很可能占据主导，并彻底破坏 SG 收敛性。

2.4.5 评估

训练通过最小化随时间累积的分类目标函数来完成。对于每个输入样本，网络在每个仿真步都会生成一个输出脉冲向量 $\mathbf{s}^{(2)}(t) \in \{0,1\}^C$ ，其中 $C = 10$ 类别。我们在每个仿真步计算交叉熵，并在刺激窗口内求和：$t \in \{1, \ldots, T\}$

$$\mathcal{L} = \sum_{t=1}^{T} \mathrm{CE}\big(\mathbf{s}^{(2)}(t),\, y\big)\,, \tag{9}$$

并使用 Adam 更新线性层的突触参数。

为从脉冲输出中获得单一类别预测，我们在最终层应用脉冲计数解码。令**类别** $= \sum_{t=1}^{T}$**脉冲计数**$^{(2)}(t)$ 表示各类别的脉冲计数；预测标签为

$$\hat{y} = \arg \max_{k \in \{1,\ldots,C\}} c_k. \tag{10}$$

准确率以验证集和测试集中 $\hat{} y = y$ 的样本比例报告。

2.5 MNIST 系列实验

为正式评估我们的模型，我们在 MNIST 系列数据集上对其进行了训练：KMNIST、MNIST、FMNIST 和 NotMNIST。

2.5.1 数据

每个数据集均包含 60，000 张用于训练的图像和 10，000 张用于测试的图像。由于 STDP-SNN 模型的计算资源受限，每张图像均为 15×15，且每次运行仅包含一个时期运行，使用 6，000 张随机抽取的图像，直到两个模型的类别分布均衡。随后，使用 100 个样本验证每个批次，最后，按照玩具实验的方式，使用 1,000 个测试样本进行测试。为比较睡眠的影响，我们使用 11 种不同的睡眠率（0% 至 100%，每次递增 10%）训练模型。此外，为提高统计置信度并确保不同睡眠率之间的可比性，我们为每种睡眠率使用一组固定的五个随机种子训练模型五次，并在所有睡眠条件下重复使用这组种子。运行总次数可表示为

$$\text{total runs} = \text{seeds} \times \text{sleep rates} \times \text{datasets} \tag{11}$$

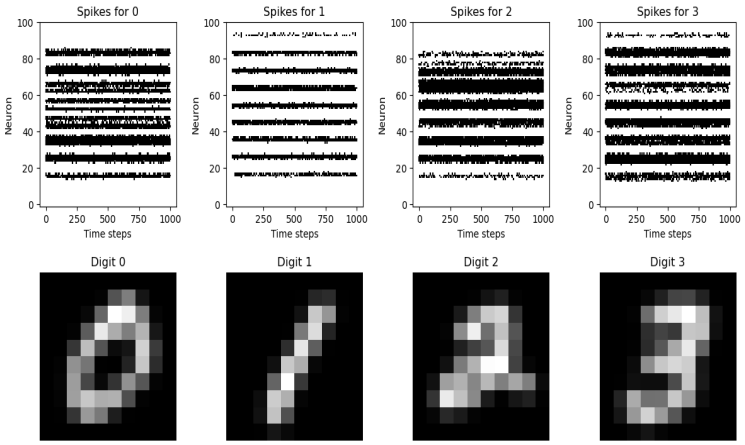


图 8：图中展示了 MNIST 数据集中的四个训练数据示例。顶行展示了利用 snntorch 的 *spikegen.rate()* 函数生成的脉冲编码训练数据。图中以时间为横轴绘制神经元活动，每个点代表一个脉冲。

2.5.2 GLM 模型

为量化基于睡眠的正则化的效果，并给出性能的 95% 置信区间，我们拟合了一个分层贝塔回归，其中包含

$$y_{ijkl} \sim \mathrm{Beta}(\mu_{ijkl}, \phi), \tag{12}$$

$$\begin{aligned}\mathrm{logit}(\mu_{ijkl}) = \beta_0 + \beta_j^{(\mathrm{sleep})} + \beta_k^{(\mathrm{model})} + \beta_{jk}^{(\mathrm{sleep}\times\mathrm{model})} \\ + b_l^{(\mathrm{dataset})} + b_i^{(\mathrm{seed})} + b_{li}^{(\mathrm{dataset}\times\mathrm{seed})}.\end{aligned} \tag{13}$$

其中， y_{ijkl} 是模型 k在数据集 l 和种子 i下、睡眠率为 j 时的测试准确率。每个 β_j量化了相对于 0% 睡眠率，采用睡眠率 j's% 的固定效应；所有睡眠水平均作为因子变量纳入，以便直接解释。

3 结果

固定效应贝塔回归估计（表 1）从统计学上证实了观察到的趋势。截距表示 STDP-SNN 模型在 0% 睡眠率下的逻辑斯蒂准确率，而系数 $\beta_{j=10}$–$\beta_{j=100}$ 则反映相对于该基线，各睡眠水平下逻辑斯蒂准确率的变化。最强效应出现在 $\beta_{j=10}$ (1.56, $p < .001$)和 $\beta_{j=20}$ (1.59, $p < .001$)，而睡眠水平更高时，效应逐渐减弱——尽管仍具有显著性。这一模式表明，睡眠达到中等水平后，收益递减；这与其他正则化方案（例如 ℓ_1 和 ℓ_2 正则化 [Tarr`es 和 Yao， 2013]）中观察到的偏差-方差权衡一致。

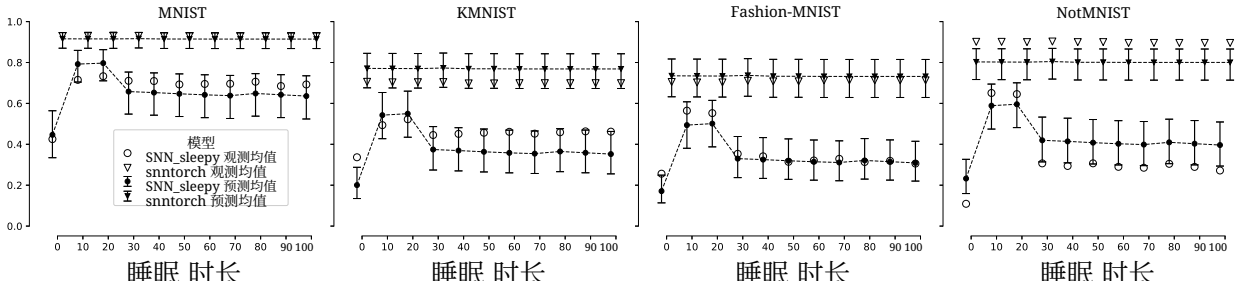


图 9：STDP-SNN 模型和 SG-SNN model-SG 在 MNIST 系列基准测试上的平均预测准确率（±95% 置信区间）与睡眠时长的关系。每个点表示基于 5 个种子的平均 `glmmTMB` 估计值。空心点表示观测准确率。

相比之下，SG-SNN 模型呈现出截然不同的行为。负向且显著的交互项与睡眠的正向主效应几乎相互抵消，使得在所有睡眠水平下的净效应都接近于零。因此，模型性能基本不受睡眠时长影响。

图 9 体现了这种差异：SG-SNN 模型保持稳定且接近最优的准确率，并在睡眠比例为 0% 时达到最优；相比之下，STDP-SNN 模型呈现非单调响应，在中等睡眠比例（10%–20%）时达到峰值，并在更高比例下下降。这一模式在全部四个数据集上均保持一致。尽管睡眠正则化带来了提升，SG-STDP 模型总体上仍保持更高的绝对准确率。

表 1 下半部分的方差估计表明了数据集层面的异质性，同时也证实了正则化效应的稳健性。

表 1：采用贝塔回归的混合效应模型

效应	估计	标准误	z 值	p值	
截距	-1.098	.295	-3.715	.192	
睡眠时长的主效应					
$\beta_{j=10}$	1.563	.156	10.026	$< .001$	***
$\beta_{j=20}$	1.592	.156	10.212	$< .001$	***
$\beta_{j=30}$	.875	.152	5.764	$< .001$	***
$\beta_{j=40}$	.853	.152	5.619	$< .001$	***
$\beta_{j=50}$	.827	.152	5.444	$< .001$	***
$\beta_{j=60}$	.804	.152	5.295	$< .001$	***
$\beta_{j=70}$	.787	.152	5.190	$< .001$	***
$\beta_{j=80}$	.833	.152	5.497	$< .001$	***
$\beta_{j=90}$	.806	.152	5.298	$< .001$	***
$\beta_{j=100}$	.778	.152	5.118	$< .001$	***
模型的主效应（参照 =*STDP-SNN* 模型）					
模型：SG-SNN	2.614	.167	15.657	$< .001$	***
$\beta_{j=10}^{(\text{sleep}\times\text{model})}$	-1.566	.230	-6.797	$< .001$	***
$\beta_{j=20}^{(\text{sleep}\times\text{model})}$	-1.599	.230	-6.939	$< .001$	***
$\beta_{j=30}^{(\text{sleep}\times\text{model})}$	-.863	.228	-3.789	$< .001$	***
$\beta_{j=40}^{(\text{sleep}\times\text{model})}$	-.864	.228	-3.797	$< .001$	***
$\beta_{j=50}^{(\text{sleep}\times\text{model})}$	-.838	.228	-3.682	$< .001$	***
$\beta_{j=60}^{(\text{sleep}\times\text{model})}$	-.819	.228	-3.600	$< .001$	***
$\beta_{j=70}^{(\text{sleep}\times\text{model})}$	-.802	.228	-3.528	$< .001$	***
$\beta_{j=80}^{(\text{sleep}\times\text{model})}$	-.849	.227	-3.731	$< .001$	***
$\beta_{j=90}^{(\text{sleep}\times\text{model})}$	-.821	.228	-3.605	$< .001$	***
$\beta_{j=100}^{(\text{sleep}\times\text{model})}$	-.793	.228	-3.483	$< .001$	***
随机效应方差					
组	Name	方差	标准差		
数据集	(截距)	.289	.538		
Seed	(截距)	$< .001$	$< .001$		
数据集：种子	(截距)	.037	.192		

4 讨论

本研究表明，稳态类睡眠正则化可以显著提升局部训练的脉冲神经网络的学习稳定性和泛化能力，但前提是将其应用于狭窄且明确界定的范围内。在完整的 MNIST 基准测试系列中，我们发现了一致的

证据表明，适度的睡眠时长能够带来可测量的性能提升，而过长的睡眠则会削弱学习效果。总体而言，这些结果从机制和统计两个层面，稳健地解释了类睡眠过程如何与局部突触可塑性相互作用。

交错安排睡眠阶段的主要作用是防止权重病态增长和灾难性干扰。在采用脉冲时序依赖可塑性训练的循环脉冲神经网络中，突触更新会不断累积，且没有内在的归一化机制，因此即使在兴奋性和抑制性群体均衡的网络中，权重幅值也会失控增长。在睡眠期间定期施加幂律权重衰减，可将突触限制在有效范围内，从而抵消这种不稳定性。这恢复了有效的可塑性，并在较长的训练过程中稳定了学习动态。

睡眠的作用不止于控制权重。在睡眠阶段，外部驱动输入会被移除，取而代之的是低振幅高斯噪声，从而诱发网络自发活动。这种背景活动会使神经元随机地重新激活先前学到的神经元集群。由于在这些静息时段脉冲时序依赖可塑性仍然活跃，突触结构能够得到有选择的强化，而不受新感觉模式的干扰。这种隐式回放机制构成了一种巩固形式，或有助于提升泛化能力。

至关重要的是，我们的结果表明，睡眠带来的益处有着明确的界限。对于脉冲时序依赖可塑性-脉冲神经网络模型，只有轻度正则化——约 10% – 20%——才能提升性能。超过这一范围后，所有数据集上的准确率都会单调下降。表 1 中的贝塔回归分析量化了这一效应：尽管 $\beta_{j=10}$ 和 $\beta_{j=20}$ 呈强正相关，但与更长睡眠时间相关的系数会逐渐减小，最终相较于最佳范围产生负面影响。这一模式与经典正则化方法（例如 ℓ_1 或 ℓ_2 惩罚项）中的表现相似：正则化过强会增加偏差并损害性能。

最后，我们的研究结果揭示了学习规则与睡眠效能之间的重要相互作用。在赫布式脉冲神经网络中，睡眠衰减与脉冲时序依赖可塑性作用于相互兼容的局部活动统计特征，共同实现权重稳态和记忆稳定。相比之下，替代梯度模型表现出本质上不同的响应。在极低睡眠率下，睡眠期间的随机扰动偶尔有助于探索损失景观，从而带来小幅性能提升。然而，随着睡眠时长增加，全局重新归一化与基于梯度的优化之间的冲突日益加剧，可能会扰乱误差信号与权重更新之间的协调。不过，从实证结果来看，这些相反的效应在很大程度上相互抵消，因此在 MNIST 系列基准测试中，性能基本不受睡眠时长影响。

我们的实现适用于神经形态部署。所有正则化和可塑性操作本质上都是局部的，不会增加推理时的开销，并且在睡眠期间每个权重仅带来 $O(1)$的计算负担，因此与事件驱动硬件完全兼容。

4.1 局限性

需要指出几项局限。首先，不同任务中每个样本的模拟步数有所不同（几何数据集为 1000 毫秒，而 MNIST- 系列数据集为 100 毫秒）。这一缩减是出于计算资源限制，因为 MNIST 系列实验需要针对每个睡眠率进行多次评估（在十一个睡眠水平上各运行五次），

因此，延长每个样本的模拟时间会造成难以承受的计算开销。

其次，SG-SNN 在学习动态和数据表征方面都与 STDP-SNN 存在根本差异。STDP-SNN 在清醒和睡眠期间都完全依赖局部脉冲时序依赖可塑性，而 SG-SNN 则在清醒阶段通过替代梯度反向传播进行训练，睡眠阶段仅引入随机权重衰减和辅助的基于迹的更新。此外，SG-SNN 处理以连续电流形式呈现的静态像素值输入，而 STDP-SNN 则处理脉冲编码的感觉输入。两种模型还采用不同的解码和准确率估算方法。因此，观察到的性能差异不能仅归因于是否采用基于睡眠的正则化，还必须结合各模型所依据的不同学习信号和表征假设来理解。

最重要的是，所提出模型的预测性能远未达到最先进水平；我们的工作应被视为在脉冲神经网络中对赫布学习进行小幅渐进改进的基础，而这类网络目前仍难以达到常规人工神经网络的性能。就这一点而言，我们的工作也不例外。

4.2 未来工作

本研究展现出若干值得探索的方向。首先，将所提出的睡眠正则化融入更深层的脉冲神经网络架构，并在更大规模的数据集上进行评估，有助于考察其可扩展性、鲁棒性和实际局限性。将这些实验扩展到受生物学启发的多模态数据，尤其契合该方法以生物学为基础的特性。

由于睡眠协议依赖噪声膜活动来保持突触强度的相对关系，下一步一项重要工作是厘清噪声如何促成这种稳定作用。具体而言，在不同睡眠率下系统地改变噪声水平，或许能够揭示其交互效应，并阐明睡眠在何种条件下最能促进学习。

除了噪声驱动的机制之外，研究睡眠如何与其他形式的稳态调节相互作用也很有价值，例如内在可塑性、注意调制或基于奖励的学习。研究这些相互作用或许有助于揭示睡眠是与其他稳定策略相辅相成，还是与之相互竞争。

最后，将所提出的模型部署在神经形态硬件上，将有力检验其在现实场景中的适用性，同时也能与其他在本地运行的连续时间学习算法进行直接比较。

4.3 结论

我们表明，受睡眠启发的稳态正则化能够稳定并提升局部训练的基于STDP的脉冲神经网络，但其作用仅限于狭窄的工作区间。在 MNIST 系列基准测试中，适度的睡眠比例（10–20%）始终能提升性能，而同样的干预在替代梯度训练中却未产生系统性收益。结合机制层面的动机、系统性参数扫描和混合效应贝塔回归，我们刻画了局部睡眠式重归一化的益处与局限。

尽管我们的模型未达到最先进的准确率，所提出的协议仍保留了严格局部性，且计算开销低，因此非常适用于神经形态计算和片上学习。更广泛地说，结果表明，类睡眠重归一化是控制赫布式脉冲神经网络中病理性权重动态的一种实用途径，也是构建具有生物学基础的鲁棒学习系统的一个有前景的方向。

A 超参数

表2：所有采用 *STDP-SNN* 模型 的实验的关键调优参数

组件	参数	描述符	值
网络			
	N_{in}	输入神经元	225
	N_{exc}	兴奋性神经元	200
	N_{inh}	抑制性神经元	50
	$P_{\text{in}\to\text{exc}}$	连线概率	15%
	$P_{\text{exc}\to\text{exc}}$	连线概率	10%
	$P_{\text{exc}\to\text{inh}}$	连线概率	20%
	$P_{\text{inh}\to\text{exc}}$	连线概率	25%
	$W_{\text{in}\to\text{exc}}$	突触权重	0.10
	$W_{\text{exc}\to\text{exc}}$	突触权重	0.15
	$W_{\text{inh}\to\text{exc}}$	突触权重	0.30
	$W_{\text{exc}\to\text{inh}}$	突触权重	-0.30
STDP			
	η_{exc}	尖峰时序依赖可塑性速率	5×10^{-4}
	η_{inh}	反向尖峰时序依赖可塑性速率	5×10^{-4}
	A_+	增强	0.5
	A_-	抑制	0.3
	τ_+	LTP时间常数	10 毫秒
	τ_-	LTD时间常数	7.5 毫秒
神经动力学			
	Δt	时间常数	1 ms
	τ_m	膜时间常数	30 毫秒
	U_{rest}	静息电位	-70 毫伏
	U_{reset}	复位电位	-80 毫伏
	U_{max}	最大电位	40 毫伏
	U_{min}	最小电位	-100 毫伏
	R_m	膜电阻	30 毫伏
	λ_μ	膜噪声均值 0毫伏	
	λ_σ	膜噪声方差	3mV
	α	LIF阈值基线	-55 毫伏
	τ_{th}	阈值 tc	100 毫秒
	δ	阈值跳变	3
睡眠/稳态			
	β	睡眠百分比	10%
	α_{base}	睡眠下界	1.0
	λ	睡眠衰减幂	0.9997
	w_{tgt}	目标权重	| 0.2 |
数据处理			
	N_{train}	训练样本	6000

组件	参数	描述符	值
	N_{test}	测试样本	1000
	N_{val}	验证样本	100
	S_{image}	图像缩放尺寸	(15，15)
	$f_{\max}$	最大脉冲频率	1000 赫兹
	$T_{\text{图像}}$	刺激持续时间	100 毫秒
	B	批次大小	400
	$N_{\text{批次}}$	批次数量	15
早停			
	p_{patience}	训练耐心值	0.2

表 3：使用 *SG-SNN* 模型进行所有实验的关键调优参数

组件	参数	描述符	值
网络			
	N_{in}	输入神经元	225
	N_{exc}	隐藏神经元	1000
	N_{exc}	输出神经元	10
	$P_{\text{in}\to\text{hid}}$	连线概率	100%
	$P_{\text{hid}\to\text{out}}$	连线概率	100%
	$W_{\text{in}\to\text{hid}}$	突触权重	1.0
	$W_{\text{hid}\to\text{out}}$	突触权重	1.0
STDP			
	η_{exc}	尖峰时序依赖可塑性速率	1×10^{-3}
	η_{inh}	反向尖峰时序依赖可塑性速率	1×10^{-3}
	A_+	增强	0.01
	A_-	抑制	0.012
	τ_{pre}	突触前痕迹时间常数	20 毫秒
	τ_{post}	突触后痕迹时间常数	20 毫秒
神经动力学			
	Δt	时间常数	1 ms
	τ_m	膜衰减	0.95 毫秒
	U_{reset}	复位电位	0
	λ_μ	膜噪声均值 0	
	λ_σ	膜噪声方差	0.5
	α	LIF阈值基线	1 mV
睡眠/稳态			
	β	睡眠百分比	[0:100%]
	α_{base}	睡眠下界	1.0
	λ	睡眠衰减幂	0.999
	w_{tgt}	目标权重	0.2
输入编码			
	$f_{\max}$	最大脉冲频率	1000 赫兹

组件	参数	描述符	值
Data pre-处理	$T_{\text{图像}}$	刺激持续时间	100 毫秒
	N_{train}	训练样本	6000
	N_{test}	测试样本	1000
	N_{val}	验证样本	100
	S_{image}	图像缩放尺寸	(15，15)
	$f_{\max}$	最大脉冲频率	1000 赫兹
	$T_{\text{图像}}$	刺激持续时间	100 毫秒
	B	批次大小	400
	$N_{\text{批次}}$	批次数量	15
优化			
	η_{Adam}	Adam 学习率	5×10^{-4}
	β_1	Adam 第一衰减率	0.9
	β_2	Adam二阶矩衰减率	0.999
早停			
	p_{patience}	耐心比例	0.2
	$\Delta_{\min}$	最小改进阈值	0.001

参考文献

Alan Eric Akil、Robert Rosenbaum 和 Kre˘simir Josi´c。脉冲时间依赖可塑性下的平衡网络。*PLOS Computational Biology*，17(5):e1008958，2021年5月。ISSN 1553-7358。doi: 10.1371/journal.pcbi.1008958。URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1008958。出版方：公共科学图书馆。

Cristiano Capone、Elena Pastorelli、Bruno Golosio 和 Pier Stanislao Paolucci。类睡眠慢振荡通过丘脑-皮层模型中的突触稳态和记忆联结改善视觉分类。*Scientific Reports*，9(1):8990，2019年6月。ISSN 2045-2322。doi: 10.1038/s41598-019-45525-0。URL https://www.nature.com/articles/s41598-019-45525-0。出版商：自然出版集团。

Kristofor D. Carlson、Micah Richert、Nikil Dutt 和 Jeffrey L. Krichmar。稳态和脉冲时序依赖可塑性的生物学合理模型：脉冲神经网络中的稳定性与学习。发表于*2013*年国际神经网络联合会议 *(IJCNN)*，第1–8页，2013年8月。doi: 10.1109/IJCNN.2013.6706961。URL https://ieeexplore.ieee.org/document/6706961。ISSN：2161-4407。

Tushar Chauhan、Timoth´ee Masquelier、Alexandre Montlibert 和 Benoit R. Cottereau。《通过赫布学习产生双目视差选择性》。*Journal ofNeuroscience*，38(44):9563–9578，2018年10月。ISSN 0270-6474、1529-2401。doi: 10.1523/

JNEUROSCI.1259-18.2018。URL https://www.jneurosci.org/content/38/44/9563。出版方：神经科学学会。栏目：研究论文。

Nicolas Deperrois、Mihai A Petrovici、Walter Senn 和 Jakob Jordan。通过扰动和对抗性梦境学习皮层表征。*eLife*，11:e76384，2022年4月。ISSN 2050-084X。doi：10.7554/eLife.76384。URL https://doi.org/10.7554/eLife.76384。出版方：eLife Sciences Publications, Ltd.

Wei Fang、Zhaofei Yu、Yanqi Chen、Timothee Masquelier、Tiejun Huang 和 Yonghong Tian。引入可学习膜时间常数以增强脉冲神经网络的学习能力，2021年8月。URL http://arxiv.org/abs/2007.05785。arXiv:2007.05785 [cs]。

Matthieu Gilson、Anthony Burkitt 和 J. Leo van Hemmen。循环神经元网络中的脉冲时序依赖可塑性。计算神经科学前沿，4:23，2010年9月。ISSN 1662- 5188。doi：10.3389/fncom.2010.00023。URL https://pmc.ncbi.nlm.nih.gov/articles/PMC2947928/。

Erik Hoel。《过拟合的大脑：梦境的演化旨在辅助泛化》。*Patterns*，2(5)：100244，2021年5月。ISSN 2666-3899。doi：10.1016/j.patter.2021.100244。URL https://www.sciencedirect.com/science/article/pii/S2666389921000647。

Dongsung Huh 和 Terrence J. Sejnowski。《脉冲神经网络的梯度下降》，2017年6月。URL http://arxiv.org/abs/1706.04698。arXiv:1706.04698 [q-bio]。

Murat Isik、Karn Tiwari、Muhammed Burak Eryilmaz 和 I. Can Dikmen。《加速神经形态计算中的传感器融合：Loihi-2 案例研究》，2024年8月。URLhttp://arxiv.org/abs/2408.16096。arXiv:2408.16096 [cs]。

Sriram Jayabal、Brandon J. Bhasin、Maxwell Kounga、Jennifer DiSanto、Aparna Suvrathan、Mark S. Goldman 和 Jennifer L.Raymond。《经验会改变支配突触可塑性与学习的时序规则》，2024年12月。URL https://www.biorxiv.org/content/10.1101/2022.11.28.518128v2。页码：2022.11.28.518128。栏目：新结果。

Fukuaki L. Kinoshita、Rikuhiro G. Yamada、Koji L. Ode 和 Hiroki R. Ueda。《睡眠—觉醒周期中突触动力学建模的统一框架》。*PLOS Biology*，23(6)：e3003198，2025年6月。ISSN 1545-7885。doi：10.1371/journal.pbio.3003198。URL https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3003198。出版商：公共科学图书馆。

Yang Li、Feifei Zhao、Dongcheng Zhao 和 Yi Zeng。使用稀疏替代梯度直接训练时序脉冲神经网络，2024年6月。网址 http://arxiv.org/abs/2406.19645。arXiv：2406.19645 [cs]。

Gehua Ma、Rui Yan 和 Huajin Tang。将噪声作为脉冲神经网络中计算和学习的资源。*Patterns*，4(10):100831，2023年10月。ISSN 26663899。doi: 10.1016/j.patter.2023.100831。URL http://arxiv.org/abs/2305.16044。arXiv:2305.16044 [cs]。

Andreas Massey。SpikingNeuralNetwork，2025年1月。URL https://github.com/AndrlmMass/SpikingNeuralNetwork。

Andrew Thomas Nere、Atif Hashmi、Chiara Cirelli 和 Giulio Tononi。《睡眠依赖性突触下调（I）：睡眠对记忆巩固与整合益处的建模》。《神经学前沿》，第4卷，2013年9月。ISSN 1664-2295。doi: 10.3389/fneur.2013.00143。URL https://www.frontiersin.org/journals/neurology/articles/10.3389/fneur.2013.00143/full。出版商：Frontiers。

Ning Qiao、Giacomo Indiveri 和 Chiara Bartolozzi。超低漏电突触缩放稳态可塑性电路的自动增益控制。载于*2016*年*IEEE*生物医学电路与系统会议（*BioCAS*），第156–159页，2016年10月。doi：10.1109/BioCAS。2016.7833755。URL https://ieeexplore.ieee.org/document/7833755。

Sen Song、Kenneth D. Miller 和 L. F. Abbott。通过脉冲时序依赖性突触可塑性实现竞争性赫布学习。《自然·神经科学》，3(9):919–926，2000年9月。ISSN 1546-1726。doi: 10.1038/78829。URL https://www.nature.com/articles/nn0900_919。出版商：自然出版集团。

Peter G. Stratton、Andrew Wabnitz、Chip Essam、Allen Cheung 和 Tara J. Hamilton。让脉冲网络发挥作用：稳健的类脑无监督机器学习，2022年9月。URL http://arxiv.org/abs/2208.01204。arXiv:2208.01204 [cs]。

Qiaoyi Su、Yuhong Chou、Yifan Hu、Jianing Li、Shijie Mei、Ziyang Zhang 和 Guoqi Li。用于目标检测的深度直接训练脉冲神经网络。发表于*2023*年*IEEE/CVF*国际计算机视觉会议（*ICCV*），第6532–6542页，2023年10月。doi: 10.1109/ICCV51070.2023.00603。URL https://ieeexplore.ieee.org/document/10376840。ISSN：2380-7504。

Timothy Tadros、Giri P. Krishnan、Ramyaa Ramyaa 和 Maxim Bazhenov。类似睡眠的无监督重放可减少人工神经网络中的灾难性遗忘。自然通讯，13(1):7742，2022年12月。ISSN 2041-1723。doi: 10.1038/s41467-022-34938-7。URL https://www.nature.com/articles/s41467-022-34938-7。出版方：自然出版集团。

Pierre Tarr`es 和 Yuan Yao。《在线学习》作为正则化路径的随机逼近，2013年1月。URL http://arxiv.org/abs/1103.5538。arXiv:1103.5538 [math]。

Johannes Thiele、Peter Diehl 和 Matthew Cook。用于循环脉冲神经网络的唤醒-睡眠算法，2017年3月。URL http://arxiv.org/abs/1703.06290。arXiv:1703.06290 [cs]。

Giulio Tononi 和 Chiara Cirelli。《睡眠与可塑性的代价：从突触和细胞稳态到记忆巩固与整合》。《神经元》，81(1):12–34，2014年1月。ISSN 0896-6273。doi: 10.1016/j.neuron.2013.12.025。URL https://pmc.ncbi.nlm.nih.gov/articles/PMC3921176/。

Guan Wang、Yuhao Sun、Sijie Cheng 和 Sen Song。演化连接性的循环脉冲神经网络，2023年5月。URL http://arxiv.org/abs/2305.17650。arXiv:2305.17650 [cs]。

Friedemann Zenke、Everton J. Agnes 和 Wulfram Gerstner。多种突触可塑性机制协同作用，在脉冲神经网络中形成和提取记忆。《自然·通讯》，6(1):6922，2015年4月。ISSN 2041-1723。doi: 10.1038/ncomms7922。URLhttps://www.nature.com/articles/ncomms7922。出版商：自然出版集团。