引入突触缩放对脉冲神经网络学习的影响

***Shinnosuke Touda*[1]**
大阪工业大学信息科学与技术研究生院，日本大阪

***Hirotsugu Okuno* [2]**
大阪工业大学信息科学与技术学院，日本大阪 hirotsugu.okuno@oit.ac.jp

摘要— **采用受神经可塑性启发的无监督学习方法的脉冲神经网络（SNN）有望成为人工智能的新框架。本研究考察了多种神经可塑性机制（如脉冲时序依赖可塑性（STDP）和突触缩放）对由脉冲神经元组成的胜者全得（WTA）网络学习的影响。我们使用 Python 实现了具有多种神经可塑性机制的胜者全得网络。训练和测试使用了 MNIST 和 Fashion-MNIST 数据集。我们改变了神经元数量、STDP 的时间常数以及突触缩放采用的归一化方法，以比较分类准确率。结果表明，基于 L2 范数的突触缩放在提升分类性能方面最为有效。采用基于 L2 范数的突触缩放，并将兴奋层和抑制层中的神经元数量均设为 400 后，网络经过一个训练轮次，在 MNIST 数据集和 Fashion-MNIST 数据集上的分类准确率分别达到 88.84% 和 68.01%**。

关键词——脉冲神经网络；突触缩放；脉冲时序依赖可塑性

I. 引言

脉冲神经网络（SNN）由能够产生脉冲式电位变化（即脉冲）的神经元模型组成，与大脑中的神经元类似，有望成为一种新型人工智能框架，能够像大脑一样进行灵活识别。与构成许多传统神经网络（NN）的形式神经元不同，脉冲神经元具有类似大脑神经元的复杂时序特性，因此，脉冲神经网络通常需要采用不同于由形式神经元组成的神经网络的学习规则。

在大脑中发生的学习规则中，脉冲时序依赖可塑性（STDP）[1][2] 和稳态可塑性（HP）[3] 长期以来一直是研究的主题。STDP 是一种学习规则，它根据直接连接的神经元之间脉冲发生时间的差异更新突触权重。例如，在兴奋性突触中，如果突触后细胞在突触前细胞放电后的一段特定时间窗口内发放脉冲，连接强度就会增强。由于这种变化本质上属于正反馈，因此需要调节机制来稳定网络，而稳态可塑性正是发挥这一作用的机制。稳态可塑性包括突触缩放 [4]，（它会调整单个神经元输入突触的整体权重），以及根据神经元的活动调节兴奋性的机制（综述见[5][6]）。在结构不同于本研究所用结构的脉冲神经网络中，STDP 与稳态可塑性的协同效应也得到了研究[7][8]。

上述学习规则在单个神经元内部或直接连接的神经元之间局部发挥作用。这一点不同于反向传播（BP）[9]，，后者是一种广泛应用于形式神经元网络的全局学习规则。由于反向传播需要将信息反向传播至整个网络，因此学习的计算负担很重，而且在作为分类器运行的网络中，反向传播不易与其共存。相比之下，上述受大脑启发的规则仅在脉冲发放的位置和时序上发挥作用，这既减轻了计算开销，又使系统能够在执行分类的同时继续学习。此外，这些受大脑启发的规则以无监督方式运行，训练时不需要带标签数据。

赢家通吃网络已用于实现脑启发学习规则。胜者全得网络由兴奋层和抑制层组成，在通过上述规则进行适当训练后，兴奋层中的神经元会针对特定输入模式选择性放电。先前研究已证明，配备某些脑启发学习规则的胜者全得网络可在多个图像数据集上进行无监督学习（[10][11] 例如）。此外，研究人员还致力于降低在硬件中实现胜者全得网络所需的存储器和能耗 [11]。先前研究还将缺乏生理合理性的条件处理应用于胜者全得网络，以提高学习效率和分类准确率。这些方法包括：当脉冲输出不足时，提高输入脉冲频率后重新输入数据[10]，；以及针对每张训练图像，仅对放电次数最多的神经元应用突触可塑性 [11]。

本研究采用了与前一项研究相同的网络配置 [10]，并考察了引入多种脑启发学习规则的影响。本研究的主要贡献有以下两点。首先，我们证明，胜者全得网络无需采用参数变化时重新输入等生理上不可行的方法，也能得到恰当训练。消除这类条件处理也能加快训练。其次，我们

© 2025年 IEEE。允许个人使用本材料。对于任何其他用途，包括在任何当前或未来媒体中重印或再版本材料以用于广告或宣传目的、创作新的汇编作品、转售或重新分发至服务器或列表，或在其他作品中重复使用本作品的任何受版权保护部分，均须获得 IEEE 的许可。

证明了用于突触缩放的函数对分类准确率的影响。如上所述，本质上属于正反馈的脉冲时序依赖可塑性需要配合一些机制，才能有效调节权重的增加。本文提出了适合与脉冲时序依赖可塑性配合使用的互补机制。

II. 相关工作

适用于分类器的脉冲神经网络构建方法大致可分为以下两类。

第一种方法利用基于形式神经元模型的深度神经网络研究成果。这些方法包括将经过训练的深度神经网络中的形式神经元转换为脉冲神经元 [12][13], , 以及调整反向传播 [9] 以训练脉冲神经网络 [14][15][16]。与转换前的原始深度神经网络相比，这类方法通常会导致一定程度的性能下降。

第二种方法是使用脑启发学习规则（如脉冲时序依赖可塑性和稳态可塑性）构建脉冲神经网络。目前，采用这种方法开发的神经网络，其准确率低于采用基于深度神经网络的方法开发的神经网络；但这种方法能够利用第 I 节所述脑启发学习规则的优势。Diehl 等人证明，基于脑启发学习规则的胜者全得网络能够以足够高的准确率对手写数字进行分类 [10]。此外，也有研究报道，通过加入卷积层以及自适应突触计算、阈值和侧向抑制，实现了准确率更高的脉冲神经网络 [17][18]。

本研究采用后一种方法。本研究的范围仅限于神经网络架构中的分类部分，因此，本研究使用的网络不包含任何特征提取组件，例如卷积神经网络。

III. 神经网络与神经元-突触模型

A. 神经网络架构

图1展示了本研究使用的胜者全得网络，该网络由输入层、兴奋层和抑制层组成。输入层由784个神经元组成，每个神经元对应输入图像中的一个像素。输入层中的神经元与兴奋层中的神经元全连接。兴奋层和抑制层的神经元数量相同，为100个或400个。兴奋层中的每个神经元都与抑制层中的一个神经元相连，并通过抑制层中的神经元抑制兴奋层中除自身以外的其他神经元。

在胜者全得网络中，获胜者按以下方式选出。兴奋层中的每个神经元都通过抑制性神经元相互抑制。因此，除脉冲率高于某一阈值的“获胜神经元”外，其他神经元的脉冲发放都会受到抑制。在这种情况下，脉冲时序依赖可塑性（第 I V 节对此进行了说明）会增加获胜神经元的输入权重，使获胜神经元与其他神经元之间的脉冲数量差异逐渐增大。

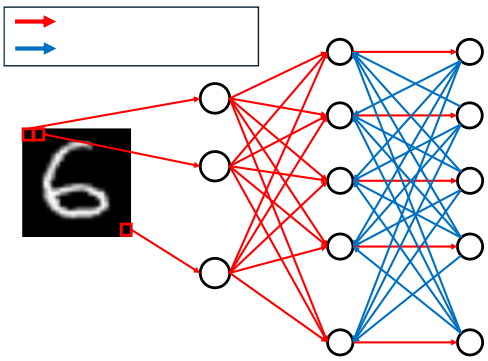

图1. 胜者全得网络架构。

B. 神经元模型

本节介绍本研究设计的神经网络中使用的两种神经元模型。对于输入层神经元，采用了脉冲数量服从泊松分布的模型。当某事件发生的概率服从期望值为 λ 的泊松分布时，该事件在短时间间隔 Δt 内发生一次的概率可由下式表示：

$$P[X(\Delta t) = 1] = \frac{(\lambda \Delta t)^1 e^{-\lambda \Delta t}}{1!} \simeq \lambda \Delta t. \qquad (1)$$

在所实现的程序中，会生成一个服从均匀分布 $U(0,1)$ 的随机数；如果其值小于 $\lambda \Delta t$，则生成一个脉冲。

对于兴奋层和抑制层，采用漏积分发放（LIF）模型作为神经元模型。图2(a)展示了LIF神经元的电路模型。v_i 表示神经元i的膜电位。C_m 和 g_m 分别表示细胞膜的电容和电导。I_S 表示化学突触的突触电流，E_L 表示静息膜电位。膜电位根据以下微分方程变化：

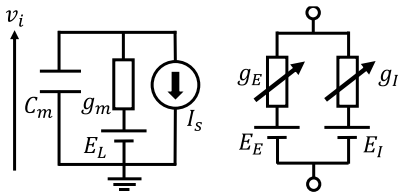

图2。神经元和突触的电路模型。(a) 兴奋层和抑制层中神经元的电路模型。(b) 化学突触的电路模型。

$$C_m \frac{dv_i}{dt} = -g_m(v_i - E_L) + I_S. \quad (2)$$

当膜电位达到某个阈值v_θ时，神经元会产生脉冲。下一次放电时，膜电位会下降至复位电位，并进入一段膜电位不发生变化的时期，即不应期；该时期的长度为t_r。兴奋性层神经元的阈值会随着学习而变化。详情见第 IV 节。抑制性神经元的阈值为常数$v_\theta = v_{\theta 0}$。

表 I 展示了兴奋层和抑制层中神经元的参数。表中的 τ_m 表示神经元的时间常数。

C. 突触模型

图 2(b)展示了化学突触的电路模型 I_S。$g_{E,\ k}$ 和 $g_{I,\ k}$ 分别表示其突触前神经元为 k的兴奋性和抑制性突触电导，而 E_E 和 E_I 分别表示兴奋性和抑制性平衡电位。兴奋性和抑制性化学突触的总电流 I_S 表示为

$$I_S = \sum_{k=1}^{N_E} g_{E,k}(v_i - E_E) + \sum_{k=1}^{N_I} g_{I,k}(v_i - E_I)\ , \quad (3)$$

其中，N_E 和 N_I 分别表示兴奋性和抑制性突触前神经元的总数。这里使用的突触电流模型是单指数突触模型，其中突触前神经元发生脉冲发放时，突触后神经元的突触电导（$g_{E,\ k}$ 和 $g_{I,\ k}$）会增加，随后呈指数衰减。设 t_{pre} 为突触前神经元脉冲发放的时刻，则突触后神经元的兴奋性突触电导 $g_{E,\ k}$ 表示为

$$g_{E,k}(t) = \frac{w_{E,k}}{\tau_s} \exp\left(-\frac{t - t_{pre}}{\tau_s}\right), \quad (4)$$

其中，$w_{E,\ k}$ 是突触前神经元 k对应的兴奋性突触权重。抑制性电导 $g_{I,\ k}$ 以相同方式计算。

表 II 展示了连接到各层神经元的突触参数。

IV. 学习规则

A. 脉冲时序依赖可塑性

输入层到兴奋层的突触权重通过脉冲时序依赖可塑性和稳态可塑性的组合进行更新，其中包括突触缩放和兴奋性调节。

表 I.　兴奋层和抑制层神经元的参数

	兴奋层	抑制层
$v_{\theta 0}$	-40 mV	-40 mV
E_L	-60 mV	-60 mV
t_r	2 ms	5 ms
$\tau_m\ (= C_m/g_m)$	30 毫秒	5 ms
τ_θ	2×10^5 毫秒	-
α	0.11 毫伏	-

表 II.　各层神经元所连接的突触参数

	兴奋层	抑制层
E_E	0 mV	0 mV
E_I	-80 毫伏	-
τ_s	20 毫秒	20 毫秒

脉冲时序依赖可塑性是一种存在于生物神经网络中的学习规则，它根据突触前神经元与突触后神经元之间脉冲发放时间的差异 $\Delta t_f (= t_{post} - t_{pre})$ 来更新权重 [1][2]。该学习规则中的权重变化 Δw 由以下方程表示：

$$\Delta w = \begin{cases} A_+ \exp\left(\frac{\Delta t_f}{\tau_p}\right) & \text{if } \Delta t > 0 \\ -A_- \exp\left(-\frac{\Delta t_f}{\tau_p}\right) & \text{if } \Delta t < 0 \end{cases}. \quad (5)$$

此处，A_+ 和 A_- 对应学习率。本研究采用的数值如下：

$A_+ = 0.005$　$A_- =$。，. . 图展示了Δt_f 0.0025　3　Δw 与 的关系。

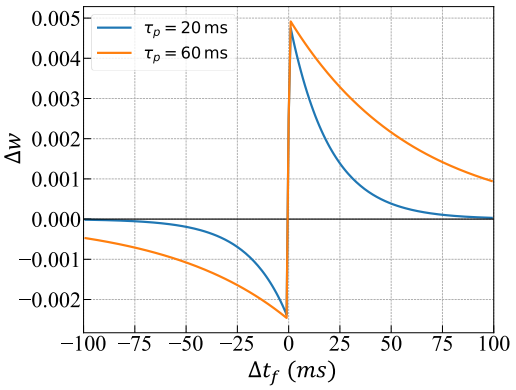

图 3. 脉冲发放时序差异 Δt_f 与权重变化 Δw之间的关系。

B. 突触缩放

除了脉冲时序依赖可塑性之外，突触缩放也会更新突触权重。脉冲时序依赖可塑性会更新单个突触权重，而突触缩放会更新输入到单个神经元的所有突触权重。本研究所实现的突触缩放根据以下公式对权重向量 $\boldsymbol{w}$ 进行归一化，该权重向量表示从输入层到兴奋层中某个神经元的权重：

$$\boldsymbol{w}' = \frac{W_t}{\|\boldsymbol{w}\|_p}\boldsymbol{w}, \tag{6}$$

其中 $\|\boldsymbol{w}\|_p$ 表示L_p 范数，使用了 L1 范数和 L2 范数。系数 W_t 表示目标范数，对应于权重保持不变时的稳态范数。

C. 兴奋性调节

由于基于脉冲时序依赖可塑性的权重更新本质上是一种正反馈，若不加以适当控制，只有某些神经元的权重会增加，放电也会被少数神经元垄断。为防止垄断，除了脉冲时序依赖可塑性之外，还需要调节神经元的兴奋性。调节兴奋性的方法多种多样；为求简便，本研究通过改变放电阈值来调节兴奋性。

兴奋层中神经元的放电阈值定义为 $v_\theta = v_{\theta 0} + \theta$，其中 $v_{\theta 0}$ 是表示初始阈值的常数，而 θ 则根据以下微分方程动态变化：

$$\frac{d\theta}{dt} = -\frac{\theta}{\tau_\theta} + \alpha\delta\left(t - t_f\right), \tag{7}$$

其中 δ 是狄拉克 δ 函数，t_f 是神经元的放电时间。兴奋层中神经元的放电阈值每次放电都会增加 α mV，并以时间常数 τ_θ衰减。

V. 仿真方法

A. 仿真设置

我们使用 Python 3.12 实现了上述胜者全得网络和学习规则。仿真的时间步长（Δt）设为 1 ms。用于训练和测试的数据集为 MNIST [19], , 这是一个由 0 到 9 的手写数字组成的图像数据集；以及 Fashion-MNIST [20], , 这也是一个图像数据集，包含 10 类服装和配饰。

两个数据集中的每张图像都包含 784 个像素（28×28）。我们使用不同的初始权重进行了三次仿真，并以三次仿真结果的平均值作为评估对象。

训练脉冲神经网络时，从 MNIST 的 60,000 张训练图像中随机选取了 40,000 张，每类 4,000 张或 Fashion-MNIST。每张输入图像的呈现时间为 150 ms。输入层到兴奋层的突触权重初始值为服从均匀分布 U(0,1) 的随机数。测试时使用了 10,000 张测试图像。每张输入图像的呈现时间为 500 ms。训练和测试期间，每次切换输入图像时，都会将每个神经元的膜电位重置为其静息膜电位。

B. 分类准确率

为评估分类准确率，应将每种脉冲输出模式与网络预测的类别关联起来。训练数据集的类别标签仅用于此关联过程。首先，针对每个类别 $l \in \{0, 1, \cdots, 9\}$的所有训练图像，计算兴奋层中每个神经元 i 的平均发放率 m_{il}。根据以下规则确定分配给神经元 i 的类别 C_i：

$$C_i = \underset{l}{\operatorname{argmax}}\, m_{il}\,. \tag{8}$$

接下来，输入图像 j 的预测类别由对图像 j响应时脉冲计数最高的神经元所对应的标签确定。这种仅依据放电率最高的神经元进行预测的方法，其准确率高于以往研究 [10], 采用的方法；后者根据分配给各类别的神经元的平均发放率确定预测类别。尤其是在网络包含大量神经元时，这种方法的准确率更高。

VI. 结果

A. 突触缩放方法的影响

本研究在 MNIST 和 Fashion-MNIST 数据集上考察了采用 L1 范数和 L2 范数进行突触缩放的方法的分类准确率。兴奋层和抑制层中的神经元数量均设为 100 或 400，STDP 时间常数 τ_p 设为 20 ms 或 60 ms。对不同 W_t取值下的分类准确率进行了比较。表 III 和表 IV 分别展示了在 MNIST 数据集上采用 L1 范数和 L2 范数的方法的分类准确率，而表 V 和表 VI 分别展示了在 Fashion-MNIST 数据集上采用 L1 范数和 L2 范数的方法的分类准确率。表中的红色数字表示每一行中的最高值。

这些结果表明，无论使用何种数据集，采用 L2 范数的方法的分类准确率始终高于采用 L1 范数的方法。造成这一现象的潜在原因可能在于下一节图 4 所示的训练后权重。

研究发现，分类准确率高度依赖于参数 W_t，而能获得最高分类准确率的 W_t 值会随神经元数量和 τ_p的值而变化。由于 W_t 表示突触权重更新后稳定状态下权重向量的 L1 或 L2 范数，

表 III.　采用 L1 范数突触缩放的 MNIST 数据集分类准确率（%）。

神经元数量	τ_p	W_t				
		60	80	120	160	200
100	20	53.17	72.70	77.02	78.57	68.94
	60	61.08	73.57	81.18	83.61	81.73
400	20	57.90	74.02	79.10	78.51	69.68
	60	63.38	74.88	85.63	87.05	85.02

表 IV.　采用 L2 范数突触缩放的 MNIST 数据集分类准确率（%）。

神经元数量	τ_p	W_t				
		3	5	7	9	11
100	20	83.46	82.74	81.76	80.78	73.41
	60	85.21	85.08	85.31	86.30	84.66
400	20	84.25	84.01	82.62	81.90	76.47
	60	85.80	88.06	88.58	88.84	86.96

表 V.　使用 L1 范数突触缩放的 Fashion-MNIST 数据集分类准确率（%）。

神经元数量	τ_p	W_t				
		60	80	120	160	200
100	20	12.50	50.40	45.42	44.69	41.20
	60	17.83	50.50	45.73	43.62	40.05
400	20	19.64	50.21	43.79	40.55	38.59
	60	22.42	48.47	43.14	42.85	38.74

表 VI.　使用 L2 范数突触缩放的 Fashion-MNIST 数据集分类准确率（%）。

神经元数量	τ_p	W_t				
		3	5	7	9	11
100	20	64.14	52.86	54.40	48.91	47.90
	60	51.92	56.67	53.89	53.28	54.68
400	20	68.01	57.32	51.05	48.55	48.04
	60	65.02	54.67	52.45	50.93	50.18

这意味着稳定状态下范数的最优值取决于各种参数。此外，当 W_t 设置得过小时，权重不足以推动学习继续进行，也无法形成下一节所示的类别特定的权重模式。表 V 中 $W_t = 60$ 处分类准确率极低，正是这一情况所致。这些事实表明，有必要采用能够让 W_t 也通过学习发生变化的方法。

关于神经元的数量，当基于 L2 范数的方法中的 W_t 设置恰当时，由 400 个神经元组成的网络在各个数据集上的表现始终优于由 100 个神经元组成的网络。神经元数量与分类准确率之间的这种正相关关系，与先前研究报告的发现一致。

另一方面，能够获得较高分类准确率的 τ_p 值因数据集而异。对于 MNIST 数据集，将 τ_p 设为 60 ms 可获得更高的准确率；而对于 Fashion-MNIST 数据集，τ_p 设为 20 ms 时表现更佳。由于最优的 τ_p 预计取决于输入脉冲的统计特性，因此有必要从这一角度开展进一步实验。

在以下条件下，采用 L2 范数的方法在 MNIST 数据集上的分类准确率最高，为 88.84 %：神经元数量 = 400、τ_p = 60 ms 和 W_t = 9。同样，在以下条件下，L2 范数方法在 Fashion-MNIST 数据集上的分类准确率最高，为 68.01 %：神经元数量 = 400、τ_p = 20 ms 和 W_t = 3。Fashion-MNIST 上的这一准确率超过了此前一项采用脉冲时序依赖可塑性进行无监督学习的研究所取得的准确率 [9]。

B. 训练过程中的突触权重自组织

图 4 展示了训练过程中突触权重的组织方式。图 4(a) 和图 4(b) 分别展示了在 MNIST 数据集上使用 L1 范数和 L2 范数的方法所获得的权重。图 4(c) 展示了在 MNIST 数据集上使用 L2 范数且 W_t 值较低（W_t = 2）的方法所获得的权重；该方法会抑制学习进程。图 4(d) 展示了在 Fashion-MNIST 数据集上使用 L2 范数的方法所获得的权重。图 4(a)、(b) 和 (d) 中的权重是使用以下参数获得的：

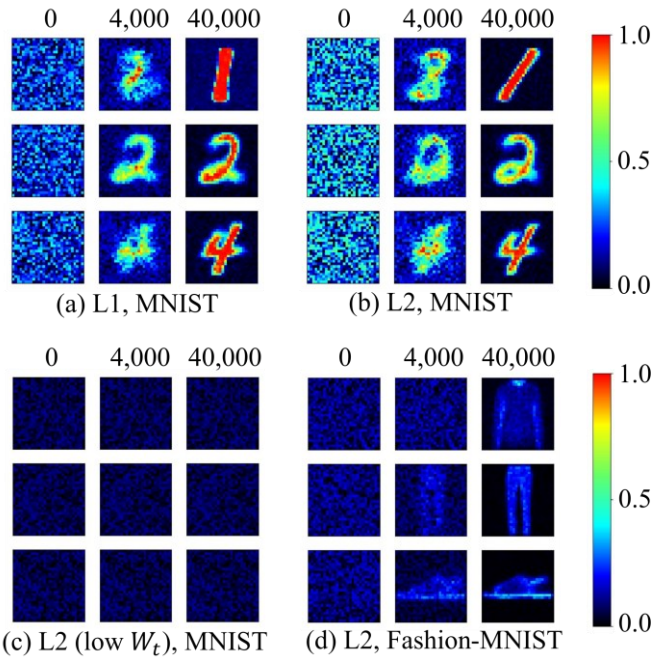

图 4. 展示训练过程中突触权重变化的延时图像。图中顶部的数字表示用于训练的输入图像数量。(a) 和 (b) 分别表示在 MNIST 数据集上使用基于 L1 范数和 L2 范数的方法得到的权重。(c) 表示在 MNIST 数据集上使用 L2 范数方法得到的权重，其中 W_t 值较低（W_t = 2），会抑制学习过程。(d) 表示在 Fashion-MNIST 数据集上使用 L2 范数方法得到的权重。

这些参数获得了最高的分类准确率。如这些图所示，权重从随机数开始，逐渐接近某个输入模式的形式。

比较图 4(a) 和 4(b) 可见，使用基于 L1 范数的方法得到的权重通常高于使用基于 L2 范数的方法得到的权重。在本文未展示的其他示例中，所有神经元权重的统计数据也体现了这一关系。尽管这种差异可能会影响分类准确率，但仍需进一步研究以确定其根本原因。

VII. 结论

本研究考察了在胜者全得网络的训练中引入突触缩放的影响，以及突触缩放的归一化方法对分类准确率的影响。

归一化方法的比较表明，使用 L2 范数的方法比使用 L1 范数的方法具有更高的分类准确率。在不依赖诸如参数变化后重新输入等生理上不可行的操作的情况下，采用基于 L2 范数的方法的网络在 MNIST 数据集和 Fashion-MNIST 数据集上的分类准确率分别达到 88.84% 和 68.01%。形成有助于分类的权重需要使 W_t 值高于某个阈值，而 STDP 时间常数对这一过程影响甚微。仍需开展更多研究，以阐明基于 L2 范数的方法为何比其他方法表现更优。

致谢

本研究获得日本学术振兴会科学研究费助成事业（JSPS KAKENHI）项目 JP24H02338 的资助。 我们谨此感谢森田明人博士在基于 Python 实现脉冲神经网络方面给予的宝贵建议。

参考文献

[1] G. Q. Bi and M. M. Poo, 《培养的海马神经元中的突触修饰：对脉冲时序、突触强度和突触后细胞类型的依赖》, 《神经科学杂志》, 第18卷, 第24期, 第10464–10472页, 1998年12月。[2] H. Markram、J. Lübke、M. Frotscher 和 B. Sakmann, 《突触后动作电位与兴奋性突触后电位的同时发生对突触效能的调节》, 《科学》, 第275卷, 第5297期, 第213–215页, 1997年1月。[3] G. G. Turrigiano, 《神经元网络中的稳态可塑性：万物变迁，万物依旧》, 《神经科学趋势》, 第22卷, 第5期, 第221–227页, 1999年5月。[4] G. G. Turrigiano、K. R. Leslie、N. S. Desai、L. C. Rutherford 和 S. B. Nelson, 《新皮层神经元中依赖活动的量子振幅缩放》, 《自然》, 第391卷, 第6670期, 第892–896页, 1998年2月。[5] G. G. Turrigiano 和 S. B. Nelson, 《神经元可塑性中的赫布机制与稳态》, 《神经生物学观点》, 第10卷, 第3期, 第358–364页, 2000年6月。[6] A. J. Watt 和 N. S. Desai, 《稳态可塑性与脉冲时序依赖可塑性：在变幻的世界中让神经元保持“冷静”》, 《突触神经科学前沿》, 第2卷, 第5页, 2010年。[7] A. Lazar、G. Pipa 和 J. Triesch, 《SORN：一种自组织循环神经网络》, 《计算神经科学前沿》, 第3卷, 第23页, 2009年10月。

[8] K. D. Carlson、M. Richert、N. Dutt 和 J. L. Krichmar, 《稳态与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》, 载于：2013年国际联合神经网络会议（IJCNN）, 第1–8页, 2013年8月。[9] D. E. Rumelhart、G. E. Hinton 和 R. J. Williams, 《通过误差反向传播学习表征》, 《自然》, 第323卷, 第6088期, 第533–536页, 1986年10月。[10] P. U. Diehl 和 M. Cook, 《利用脉冲时序依赖可塑性进行数字识别的无监督学习》, 《计算神经科学前沿》, 第9卷, 2015年8月。[11] R. V. W. Putra 和 M. Shafique, 《FSpiNN：面向存储器与能效的脉冲神经网络优化框架》, 《IEEE集成电路与系统计算机辅助设计汇刊》, 第39卷, 第11期, 第3601–3613页, 2020年11月。[12] P. U. Diehl、D. Neil、J. Binas、M. Cook、S. C. Liu 和 M. Pfeiffer, 《通过权重与阈值平衡实现快速分类、高精度的脉冲深度网络》, 载于：2015年国际联合神经网络会议（IJCNN）, 第1–8页, 2015年7月。[13] A. Sengupta、Y. Ye、R. Wang、C. Liu 和 K. Roy, 《深入探索脉冲神经网络：VGG与残差架构》, 《神经科学前沿》, 第13卷, 2019年3月。[14] J. H. Lee、T. Delbruck 和 M. Pfeiffer, 《使用反向传播训练深度脉冲神经网络》, 《神经科学前沿》, 第10卷, 2016年11月。[15] Y. Wu、L. Deng、G. Li、J. Zhu 和 L. Shi, 《用于训练高性能脉冲神经网络的时空反向传播》, 《神经科学前沿》, 第12卷, 2018年5月。[16] W. Zhang 和 P. Li, 《通过反向传播进行深度脉冲神经网络的时间脉冲序列学习》, 载于：第34届神经信息处理系统国际会议论文集, 第12022–12033页, 2020年12月。[17] C. Lee、P. Panda、G. Srinivasan 和 K. Roy, 《通过基于脉冲时序依赖可塑性的无监督预训练和监督微调训练深度脉冲卷积神经网络》, 《神经科学前沿》, 第12卷, 第435页, 2018年。[18] Y. Dong、D. Zhao、Y. Li 和 Y. Zeng, 《受生物学上合理的学习规则与连接启发的基于脉冲时序依赖可塑性的无监督脉冲神经网络》, 《神经网络》, 第165卷, 第799–808页, 2023年8月。[19] 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, 1998年1月。[20] H. Xiao、K. Rasul 和 R. Vollgraf, 《Fashion-MNIST：用于机器学习算法基准测试的新型图像数据集》, arXiv, 2017年9月。