

学校代码: 10284

分类号: TP181

密级: 公开

U D C: 004.8

学号: 502022370027



南京大學

# 硕士学位论文

论文题目 基于生物神经元非线性  
整合机制的脉冲神经元  
模型研究

作者姓名 刘佩涵

专业名称 计算机科学与技术

研究方向 脉冲神经网络

导师姓名 申富饶 教授

2025年5月16日

答辩委员会主席 武港山 教授

评 阅 人 戴新宇 教授

徐明华 教授

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

论文答辩日期 2025年5月16日

研究生签名:

导师签名:

# **Research on Spiking Neuron Models Based on the Nonlinear Integration Mechanisms of Biological Neurons**

by  
**Liu Peihan**

Supervised by  
Professor Shen Furao

A dissertation submitted to  
the graduate school of Nanjing University  
in partial fulfilment of the requirements for the degree of

MASTER

in

Computer Science and Technology



School of Artificial Intelligence  
Nanjing University

May 16, 2025



# 南京大学学位论文原创性声明

本人郑重声明，所提交的学位论文是本人在导师指导下独立进行科学研究工作所取得的成果。除本论文中已经注明引用的内容外，本论文不包含其他个人或集体已经发表或撰写过的研究成果，也不包含为获得南京大学或其他教育机构的学位证书而使用过的材料。对本文的研究做出重要贡献的个人和集体，均已在论文的致谢部分明确标明。本人郑重声明愿承担本声明的法律责任。

研究生签名：

日期：



# 南京大学研究生毕业论文中文摘要首页用纸

毕业论文题目：基于生物神经元非线性整合机制的脉冲神经元模型研究

计算机科学与技术 专业 2022 级硕士生姓名：刘佩涵

指导教师（姓名、职称）：申富饶 教授

## 摘 要

人工神经网络作为一种广泛应用的机器学习模型，对学术界与产业界均具有显著影响。然而，人工神经网络的推理过程通常需要大量的计算资源，限制了其在算力、能耗受限场景下的应用。脉冲神经网络作为一种脑启发的事件驱动计算模型，不仅具备更强的生物学可解释性，还在低能耗、低延迟推理和神经形态数据处理等方面展现出独特优势，因而逐渐成为研究热点。当前脉冲神经网络研究主要可分为三类：一是利用生物神经网络的部分性质构建新型脉冲神经网络；二是将人工神经网络的模型和训练方法迁移到脉冲神经网络；三是脉冲神经网络相关硬件设计以及计算加速。其中，基于生物神经元特性设计新的脉冲神经元模型的研究较为稀缺。

脉冲神经元是脉冲神经网络的基本计算单元，其计算能力直接影响网络的整体表现。当前大多数脉冲神经网络基于 LIF (Leaky Integrate-and-Fire) 模型构建，然而 LIF 模型沿用了人工神经元的线性整合方式来处理同一时刻的脉冲输入，未能充分模拟生物神经元中更为复杂的树突整合机制，这在一定程度上限制了其表达能力。为此，本文充分借鉴生物神经元的整合特性，提出了两种改进的脉冲神经元模型，使其具备更强的计算表达能力和更好的生物学可解释性。将这些模型应用于常见脉冲神经网络结构中，不仅能够提升主流脉冲神经网络的精度表现，同时也能够加速网络的收敛过程。本文的主要贡献包括：

1. **基于双线性整合的脉冲神经元模型。** 现有的脉冲神经元模型通常采用线性整合方式处理当前时刻的脉冲输入，但生物神经元的膜电位响应展现出明显的双线性特性。尽管双线性整合模式能够有效捕捉输入之间的二阶特征交互，但其计算复杂度较高且难以直接实现。为此，本文基于因子分解机

的矩阵分解技术对双线性整合 (Bilinear Integration) 机制进行了创新性的实现, 在保持较低计算复杂度的同时, 增强了脉冲神经元对非线性特征的学习能力。

2. **基于分流抑制的脉冲神经元模型。** 脉冲神经元的输入通常包括兴奋性输入和抑制性输入, 分别对应生物神经元去极化和超极化过程。然而, 生物神经元的分流抑制 (Shunting Inhibition) 效应常常被忽视。为此, 本文设计了一种具有分流抑制特性的脉冲神经元模型, 能够更准确地模拟生物神经元的抑制机制, 从而提高神经元对复杂输入模式的响应能力。
3. **应用: 基于脉冲神经网络的推荐系统。** 本文提出的双线性整合脉冲神经元在保持脉冲神经网络高能效优势的同时, 具备学习二阶特征交互的能力。为验证其实际应用价值, 本文构建了一个基于脉冲神经网络的推荐系统, 为推荐场景提供了一个低能耗、低延迟的创新解决方案。

**关键词:** 脉冲神经元; 脉冲神经网络; 树突整合; 双线性整合; 分流抑制

# 南京大学研究生毕业论文英文摘要首页用纸

THESIS: Research on Spiking Neuron Models Based on the Nonlinear  
Integration Mechanisms of Biological Neurons

SPECIALIZATION: Computer Science and Technology

POSTGRADUATE: Liu Peihan

MENTOR: Professor Shen Furao

## ABSTRACT

As a widely used machine learning model, artificial neural networks (ANNs) have a significant impact on both academia and industry. However, the inference process of ANNs typically requires substantial computational resources, limiting their application in scenarios where computing power and energy consumption are constrained. In contrast, spiking neural networks (SNNs), as a brain-inspired event-driven computing model, not only offer enhanced biological interpretability but also demonstrate unique advantages in low-energy, low-latency inference and neuromorphic data processing. Consequently, they have gradually become a research hotspot. Current research on SNNs can be broadly categorized into three groups: first, constructing novel SNNs by incorporating certain properties of biological neural networks; second, adapting the models and training methods of artificial neural networks to SNNs; and third, designing specialized hardware and computational accelerators for SNNs. Among these, studies focusing on the design of new spiking neuron models based on the properties of biological neurons remain relatively scarce.

Spiking neurons are the fundamental computational units in SNNs, and their computational power directly influences the overall performance of the network. Most existing SNNs are based on the Leaky Integrate-and-Fire (LIF) model, which employs a linear integration method to process spike inputs at a given moment. This approach fails to fully emulate the complex dendritic integration mechanisms observed in biological neurons, thereby limiting the expressive power of the model. To address this limitation, this paper draws inspiration from the integration characteristics of biological neurons

and proposes two improved spiking neuron models with enhanced computational expressiveness and superior biological interpretability. When integrated into common SNNs, these models not only improve the accuracy of mainstream SNNs but also accelerate network convergence. The primary contributions of this paper are as follows:

1. **A spiking neuron model based on bilinear integration.** Existing spiking neuron models typically employ a linear integration scheme for processing current spike inputs, yet the membrane potential responses of biological neurons exhibit distinct bilinear characteristics. Although bilinear models can effectively capture second-order feature interactions between inputs, their high computational complexity makes direct implementation challenging. To overcome this, the proposed approach innovatively realizes the bilinear integration mechanism based on the matrix factorization techniques of Factorization Machines, thereby enhancing the spiking neuron's ability to fit nonlinear distributions while retaining linear computational efficiency.
2. **A spiking neuron model based on shunting inhibition.** Spiking neuron inputs generally include excitatory and inhibitory components, corresponding to the depolarization and hyperpolarization processes in biological neurons, respectively. However, the shunting inhibition effect is often neglected. In response, this paper designs a spiking neuron model incorporating shunting inhibition, which more accurately simulates the inhibitory mechanisms of biological neurons and improves the neuron's responsiveness to complex input patterns.
3. **Application: A Recommendation System Based on SNNs.** The proposed spiking neuron model based on bilinear integration not only preserves the high energy efficiency inherent to SNNs but also possesses the ability to learn second-order feature interactions. To validate its practical value, we develop a recommendation system based on SNNs, offering an innovative solution characterized by low energy consumption and low latency for recommendation scenarios.

**KEYWORDS:** Spiking Neurons; Spiking Neural Networks; Dendritic Integration; Bilinear Integration; Shunting Inhibition

# 目 录

中文摘要 .....	I
ABSTRACT .....	III
目 录 .....	V
插图目录 .....	IX
表格目录 .....	XI
<b>第一章 绪论</b> .....	<b>1</b>
1.1 脉冲神经网络研究背景 .....	1
1.2 脉冲神经元建模研究现状与面临的问题 .....	4
1.3 本文主要工作 .....	6
<b>第二章 相关工作</b> .....	<b>9</b>
2.1 脉冲神经网络预备知识 .....	9
2.1.1 神经元 .....	9
2.1.2 脉冲编码 .....	13
2.1.3 神经网络结构 .....	14
2.1.4 脉冲神经网络的训练 .....	16
2.2 生物学启发的脉冲神经元研究 .....	19
2.2.1 脉冲延迟 .....	19
2.2.2 二阶脉冲神经元 .....	20
2.2.3 脉冲神经元不应期 .....	21
2.3 本章小结 .....	22

<b>第三章 基于双线性整合的脉冲神经元模型</b> . . . . .	23
3.1 研究动机 . . . . .	23
3.2 生物学依据与关键技术 . . . . .	24
3.2.1 生物神经元的非线性整合 . . . . .	24
3.2.2 二阶特征交互的高效建模方法 . . . . .	25
3.3 非线性整合脉冲神经元模型实现与分析 . . . . .	27
3.3.1 神经元实现 . . . . .	27
3.3.2 计算优势分析 . . . . .	29
3.3.3 二阶特征交互与脉冲编码的适应性分析 . . . . .	30
3.3.4 全连接层与卷积层实现 . . . . .	31
3.4 实验与分析 . . . . .	33
3.4.1 实验设置 . . . . .	33
3.4.2 全连接网络对比实验结果与分析 . . . . .	38
3.4.3 卷积网络对比实验结果与分析 . . . . .	40
3.4.4 消融实验 . . . . .	42
3.5 本章小结 . . . . .	44
<b>第四章 基于分流抑制的脉冲神经元模型</b> . . . . .	45
4.1 研究动机 . . . . .	45
4.2 生物学依据与关键技术 . . . . .	46
4.2.1 生物神经元的去极化与超极化 . . . . .	46
4.2.2 生物神经元的分流抑制 . . . . .	47
4.2.3 人工神经元的门控机制 . . . . .	48
4.3 分流抑制脉冲神经元模型实现与分析 . . . . .	50
4.3.1 脉冲神经元正负输入拆解 . . . . .	50
4.3.2 神经元实现 . . . . .	51
4.3.3 计算优势分析 . . . . .	52
4.3.4 全连接层与卷积层实现 . . . . .	53
4.4 实验与分析 . . . . .	54
4.4.1 实验设置 . . . . .	54

---

4.4.2	全连接网络对比实验结果与分析 . . . . .	55
4.4.3	卷积网络对比实验结果与分析 . . . . .	57
4.4.4	消融实验 . . . . .	58
4.5	本章小结 . . . . .	61
<b>第五章</b>	<b>应用：基于脉冲神经网络的推荐系统 . . . . .</b>	<b>63</b>
5.1	研发背景 . . . . .	63
5.2	系统设计 . . . . .	64
5.2.1	需求分析 . . . . .	64
5.2.2	系统架构 . . . . .	65
5.3	系统实现 . . . . .	65
5.3.1	数据模块 . . . . .	66
5.3.2	后端模块 . . . . .	66
5.3.3	前端模块 . . . . .	68
5.4	效果展示 . . . . .	68
5.4.1	界面展示 . . . . .	68
5.4.2	推荐结果量化对比 . . . . .	69
5.5	本章小结 . . . . .	70
<b>第六章</b>	<b>总结与展望 . . . . .</b>	<b>71</b>
	<b>参考文献 . . . . .</b>	<b>73</b>
	<b>致    谢 . . . . .</b>	<b>81</b>
	<b>简历与科研成果 . . . . .</b>	<b>83</b>
	<b>学位论文出版授权书 . . . . .</b>	<b>85</b>



# 插图目录

1-1	三代神经网络的神经元对比 . . . . .	2
1-2	HH 模型示意图 . . . . .	4
1-3	异或问题示意图 . . . . .	6
1-4	本文研究动机 . . . . .	8
2-1	LIF 神经元模型示意图 . . . . .	11
2-2	频率编码和时间编码的对比 . . . . .	14
3-1	生物神经元膜电位响应示意图 . . . . .	23
3-2	双线性整合神经元的异或计算 . . . . .	25
3-3	二阶特征交互学习方式对比 . . . . .	30
3-4	MNIST 和 Fashion-MNIST 数据集 . . . . .	34
3-5	CIFAR100 和 CIFAR10-DVS 数据集 . . . . .	34
3-6	MNIST 数据集上全连接脉冲神经网络使用线性与双线性整合的 训练曲线对比 . . . . .	39
3-7	基于 PLIF 的深度脉冲神经网络在 CIFAR10-DVS 数据集上的训练 曲线 . . . . .	42
4-1	分流抑制示意图 . . . . .	45
4-2	传统脉冲神经元与分流抑制脉冲神经元模型对比 . . . . .	52
4-3	分流抑制神经元的异或计算 . . . . .	52
4-4	MNIST 数据集上全连接脉冲神经网络使用线性与分流抑制整合 的训练曲线对比 . . . . .	56
4-5	分流抑制脉冲神经网络消融实验训练曲线 . . . . .	60
5-1	系统架构 . . . . .	65

5-2	系统登录/注册界面 . . . . .	68
5-3	系统主页电影推荐板块 . . . . .	69

## 表格目录

3-1	对图像数据的预处理 . . . . .	35
3-2	实验网络参数设置 . . . . .	37
3-3	基于 PLIF 搭建的脉冲神经网络结构 . . . . .	38
3-4	双线性整合在全连接网络中的实验结果 . . . . .	38
3-5	双线性整合在卷积网络中的实验结果 . . . . .	41
3-6	双线性整合消融实验的模型设置 . . . . .	43
3-7	双线性整合消融实验结果 . . . . .	43
4-1	实验网络参数设置 . . . . .	54
4-2	分流抑制在全连接网络中的实验结果 . . . . .	55
4-3	分流抑制在卷积网络中的实验结果 . . . . .	57
4-4	分流抑制消融实验的模型设置 . . . . .	59
4-5	分流抑制消融实验结果 . . . . .	59
5-1	MovieLens 数据集实验结果对比 . . . . .	70



# 第一章 绪论

## 1.1 脉冲神经网络研究背景

人工神经网络（Artificial Neural Networks, ANNs）是人工智能联结主义学派的核心组成部分，联结主义的基本观点是人类的智能来源于神经元之间的连接与互动，通过人工神经网络模型能够解构大脑的结构和功能，从而模拟人类的感知、学习、记忆等功能。人工神经网络也是机器学习领域的重要模型，相较于传统的机器学习模型具备更强的自动特征学习能力、高维数据处理能力、泛化能力和可扩展性。人工神经网络的历史可追溯至 1943 年，当时 Warren Sturgis McCulloch 和 Walter Harry Pitts 基于数学和阈值逻辑提出了 McCulloch-Pitts (MP) 神经元模型<sup>[1]</sup>，奠定了神经网络理论的基础。随后，人工神经网络的研究经历了几次兴盛和低谷，曾因硬件计算能力和学习理论瓶颈而陷入停滞。随着反向传播（Backpropagation）算法<sup>[2]</sup>的提出和计算硬件（尤其是 GPU 等高性能并行计算平台）的进步，人工神经网络逐渐在各领域发挥重要作用。近三十年以来，网络结构的改进，如卷积神经网络（Convolutional Neural Networks, CNNs）<sup>[3]</sup>、长短期记忆（Long Short-Term Memory, LSTM）<sup>[4]</sup>、残差连接（Residual Connections）<sup>[5]</sup>、Transformer<sup>[6]</sup>等，以及训练方法的优化，如归一化（Normalization）<sup>[7-9]</sup>、正则化（Regularization）、dropout<sup>[10]</sup>等，拓展了神经网络的应用范围，有效提升了深度神经网络的性能，推动了其在复杂任务中的应用，从而使人工神经网络在多个领域取得了显著成果。2024 年，John Joseph Hopfield 和 Geoffrey Everest Hinton 因在人工神经网络和机器学习领域的开创性贡献而获得诺贝尔物理学奖。这也体现了人工神经网络在科研界的深远影响和重要地位。

然而，人工神经网络在实际应用中仍面临一些挑战与局限。其中，高推理能耗问题尤为突出，特别是在移动设备和边缘计算环境中，深度神经网络的高计算需求往往导致能耗过大，同时伴随较高的延时。此外，人工神经网络在处理带有噪声、干扰或不完整的数据时，鲁棒性较差。例如，当图像中加入人类无法察觉

的某些特定的微小噪声时，神经网络可能会出现意外的误判<sup>[11-12]</sup>，这限制了其在一些实际应用中的可靠性。与之相对，生物神经网络具有低推理能耗、高稳定性以及较强的噪声抑制能力等优势。因此，部分研究者希望通过更精准地模拟生物神经网络的结构和功能，来实现这些特性，从而克服人工神经网络在能效、鲁棒性等方面的局限，脉冲神经网络（Spiking Neural Networks, SNNs）作为一种更接近生物学现实的模型，逐渐成为研究的热点。目前，脉冲神经网络已在语音处理<sup>[13]</sup>、视频分析<sup>[14-15]</sup>、机器人控制<sup>[16]</sup>以及自动驾驶<sup>[17]</sup>等领域取得实际应用。

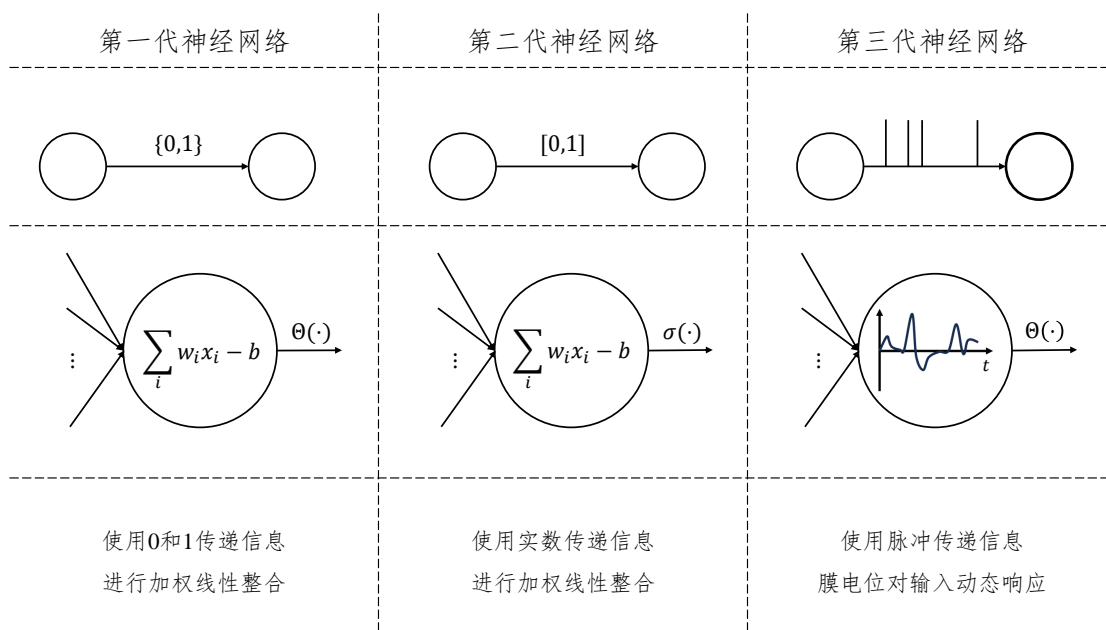


图 1-1 三代神经网络的神经元对比。不同代际神经网络的区别主要体现在神经元输入类型以及输入整合模式的不同。图中  $\Theta(\cdot)$  表示阶跃函数， $\sigma(\cdot)$  表示连续激活函数。

脉冲神经网络被称为“第三代神经网络”，这种代际划分的依据主要是神经网络所使用的神经元类型。第一代神经网络基于 MP 神经元构建，例如多层感知机（Multilayer Perceptron, MLP）<sup>[2]</sup>、霍普菲尔德网络（Hopfield Network）<sup>[18]</sup>和玻尔兹曼机（Boltzmann Machine）<sup>[19]</sup>等。这些网络的主要局限在于其输出是离散值，且缺乏生物神经元输出信号的时序特性，导致其对模拟信号的表达能力受限。第二代神经网络在感知器模型的基础上，改用了连续的激活函数，如 Sigmoid、Tanh、ReLU 等。这些网络具有更强的模拟信号表达能力，被广泛应用于现代人工智能中，包括全连接神经网络、卷积神经网络、循环神经网络、Transformer 等。目前学术界中“人工神经网络”一词通常指第二代神经网络，本文也遵循这一惯例。作为第三代神经网络，脉冲神经网络则通过引入脉冲神经元，更准确地模仿

生物神经元的离散脉冲传递机制、时序整合特性以及激活机制。脉冲神经元通过接收、发放离散脉冲来传递信息，能够更自然地处理时间依赖型数据，同时在计算效率上具有显著优势。与人工神经网络类似，脉冲神经网络也满足通用近似定理<sup>[20]</sup>，即在具有足够层数和神经元数量的情况下，脉冲神经网络可以表示任意目标函数。此外，脉冲神经网络在拟合某些函数时所需要的神经元数量远小于人工神经网络<sup>[20]</sup>。

脉冲神经网络作为一种生物学启发的计算模型，在神经元建模、网络架构设计以及学习机制等方面有很多对生物神经网络的借鉴。脉冲神经元是本文的核心研究内容，其发展历程和具体实现会在后续章节中详细阐述。在网络结构方面，传统的人工神经网络通常采用全连接结构，而生物神经网络中的神经元连接则通常呈现高度稀疏性，这种稀疏连接结构反映了生物大脑神经元的实际布局，部分脉冲神经网络的研究<sup>[21-22]</sup>也开始借鉴这一稀疏连接模式。在脉冲神经网络中，稀疏连接的主要优势在于能够减少计算复杂度、降低模型的计算成本。此外，稀疏连接使得一个神经元的异常输出只会影响下一层的部分神经元，从而提高网络对异常情况的鲁棒性。生物神经元的放电不仅是信息传递的过程，也是生物神经网络学习（主要指连接权重调节）的基础。赫布学习（Hebbian Learning）<sup>[23]</sup>法则是神经学习的经典理论之一，提出当两个神经元的活动在时间上紧密关联时，它们之间的连接应当变得更强。基于赫布法则，脉冲时间依赖性可塑性（Spike-Timing-Dependent Plasticity, STDP）理论提供了该过程的量化描述，STDP 及其改进的生物学习算法也被广泛应用于脉冲神经网络的训练任务中<sup>[24-25]</sup>。

尽管脉冲神经网络的提出已有较长时间，其发展仍然相较于人工神经网络滞后，主要原因之一在于，脉冲神经网络不连续的激活机制导致其缺乏有效的有监督学习算法。因此，目前脉冲神经网络的研究方向之一集中在将人工神经网络中经过验证的训练方法与优化技巧迁移到脉冲神经网络中。由于脉冲神经网络具备类似于循环神经网络的递归结构，沿时间反向传播（Backpropagation Through Time, BPTT）算法<sup>[26-28]</sup>被引入到脉冲神经网络训练中。此外，由于脉冲神经网络的激活函数具有不连续和不可导的特性，研究者提出在反向传播过程中使用连续函数替代原激活函数来完成梯度计算。因此这类直接基于反向传播训练脉冲神经网络的算法被称作替代梯度下降法（Surrogate Gradient Descent），其中一种典型方法是 STBP（Spatio-Temporal Backpropagation）算法<sup>[29]</sup>。然而，递归结构

和替代梯度的性质导致替代梯度下降法在训练脉冲神经网络时面临梯度消失的问题。为了缓解这一问题，批归一化技术（Batch Normalization, BN）也被应用于脉冲神经网络训练，并基于网络特性发展出 tdBN<sup>[30]</sup>、BNTT<sup>[31]</sup>以及 MPBN<sup>[32]</sup>等改进方法。这些优化显著提升了替代梯度下降法的训练效果，并加速了脉冲神经网络的收敛。另一方面，由于脉冲神经网络输入、输出的 0/1 脉冲序列可以被视为对实数的编码，因此另一种训练脉冲神经网络的方法就是 ANN-SNN 转化。该方法通过先训练表现良好的人工神经网络模型，并将其权重矩阵迁移到结构相似的脉冲神经网络中，从而使脉冲神经网络能够模拟人工神经网络的行为。这种方法常用于大型网络的训练<sup>[33-34]</sup>，且有研究者针对带有残差链接的脉冲神经网络提出了进一步的优化方法<sup>[35-36]</sup>。

本节分析了人工神经网络的优势与局限性，突出脉冲神经网络的独特优势及其研究意义，并总结了生物学研究成果与人工神经网络优化经验对脉冲神经网络研究的推动作用。

## 1.2 脉冲神经元建模研究现状与面临的问题

上一节介绍了脉冲神经网络的研究背景，本节则聚焦于脉冲神经元模型，总结其研究现状，并分析现有研究的局限性，最后引出本文所要解决的关键问题。

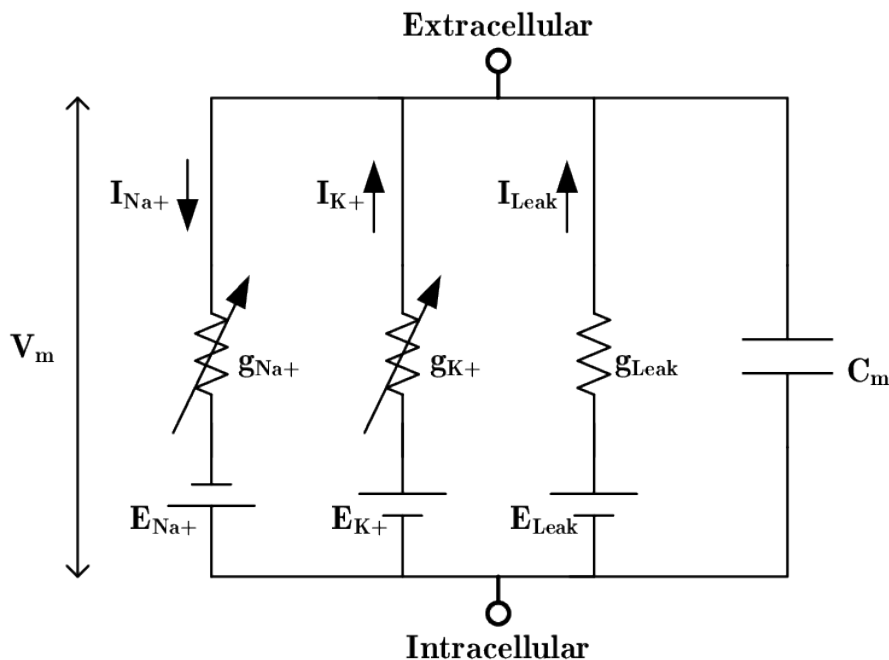


图 1-2 HH 模型示意图<sup>[37]</sup>

在生物神经元建模领域，早期的经典探索之一是 HH (Hodgkin-Huxley) 模型<sup>[38]</sup>，该模型考虑了多个离子通道的动力学及其对膜电位的影响，将神经元的细胞膜视为一个电容器，能够储存电荷，并通过一系列离子通道导电。这些通道分别允许钠离子电流 ( $I_{Na+}$ )、钾离子电流 ( $I_{K+}$ ) 和泄漏电流 ( $I_{Leak+}$ ) 通过，从而导致膜电位的变化。HH 模型可由如下微分方程表示：

$$C_m \frac{dV(t)}{dt} = -g_{Na}m^3(V(t)-E_{Na})-g_Kn^4(V(t)-E_K)-g_{Leak}(V(t)-E_{Leak})+I(t). \quad (1-1)$$

其中  $V(t)$  表示  $t$  时刻的膜电位， $I(t)$  为外部输入电流， $m, n$  为电压门控变量， $g_{Na}, g_K, g_{Leak}$  为电导， $E_{Na}, E_K, E_{Leak}$  为各通道的反极化电位， $C_m$  为膜电容常量。HH 模型不仅成功解释了神经元在不同条件下的电生理活动，还为后来的脉冲神经元模型奠定了基础。

然而，HH 模型较为复杂，对于大规模神经网络仿真有较大的计算开销。相比之下，LIF 模型<sup>[39]</sup>忽略了离子通道的具体机制，通过积分和泄漏的电流动态来描述神经元膜电位的变化，该模型也能够有效地模拟生物神经元的脉冲发放行为，并且计算效率更高，因此成为了脉冲神经网络中最常用的神经元模型。在 LIF 模型的基础上，研究者们提出了若干新的脉冲神经元模型。例如，借鉴 LSTM 和 GRU 等改进，有研究提出了脉冲 LSTM (Spiking LSTM)<sup>[40]</sup>和脉冲 GRU (Spiking GRU)<sup>[41]</sup>等模型。尽管 LSTM 和 GRU 已在循环神经网络中基本取代了原有模型，但它们在脉冲神经网络中的应用仍然很受限。

LIF 模型仍然是现有脉冲神经网络中的主流模型。目前，脉冲神经元模型的改进很少从生物神经元的特性出发，而是更多地借鉴了人工神经元的设计理念，这也导致脉冲神经元模型保留了许多人工神经元的特性。例如，尽管 LIF 神经元相比于人工神经元具备了时序整合特性，但它们对于同一时刻输入信号的整合机制仍保持一致，两者均采用线性整合。这导致单个脉冲神经元与人工神经元在表达能力上存在不足，无法完成生物神经元的某些非线性计算。1969 年，Marvin Lee Minsky 和 Seymour Aubrey Papert 在《感知机》一书中证明了单层人工神经网络无法解决异或问题<sup>[42]</sup>，此后人工神经网络的研究经历了长时间的低谷期。虽然通过增加神经元数量和网络层数可以解决这一问题，但现实中网络规模是有限的。在有限资源的约束下，单个神经元缺乏某些基本计算能力会

影响整个网络的表现。导致单一人工神经元无法解决异或问题的原因主要有两点：线性整合方式和单调激活函数。假设人工神经元接受两个输入  $x_1, x_2$ ，输出为  $y = \sigma(w_1x_1 + w_2x_2)$ ，如果可以进行异或计算，那么有：

$$\sigma(0) = 0, \tag{1-2}$$

$$\sigma(w_1) = 1, \tag{1-3}$$

$$\sigma(w_2) = 1, \tag{1-4}$$

$$\sigma(w_1 + w_2) = 0. \tag{1-5}$$

显然，如果  $\sigma(\cdot)$  是单调函数，那么根据其单调性，公式 (1-2)、(1-3) 和公式 (1-4)、(1-5) 必然会推导出矛盾的结论，因此单个人工神经元无法完成异或计算。生物学研究表明，单个大脑神经元就能够完成异或计算<sup>[43]</sup>，而生物神经元的激活函数通常是单调递增的。这表明生物神经元采用了非线性的整合方式，一些研究也证实了这一点<sup>[44-46]</sup>。

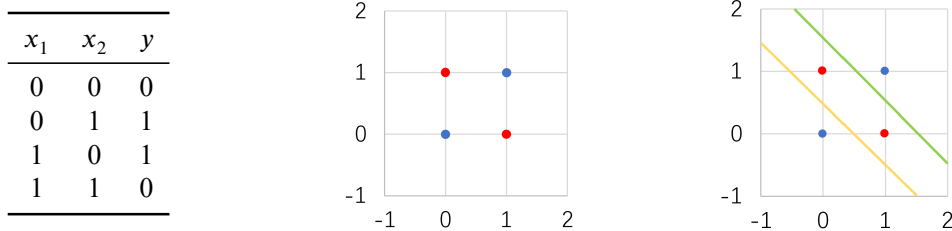


图 1-3 异或问题示意图。在二维平面上，一条直线无法将异或数据区分。

与生物神经元相比，当前的脉冲神经元模型缺乏某些基本的计算能力，这可能导致脉冲神经网络无法充分发挥生物神经网络的优势。线性的输入整合方式是限制其计算能力的主要原因之一。现有脉冲神经元未能采用生物神经元复杂的整合机制，关键原因在于后者的计算复杂度较高。如何在借鉴生物神经元整合模式的同时有效控制计算复杂度，成为亟待解决的一大问题。

### 1.3 本文主要工作

脉冲神经元是脉冲神经网络的基本计算单元，通过模拟生物神经元的整合方式，可以提升计算能力，从而增强脉冲神经网络的表现。然而，由于计算复杂度的限制，生物神经元的整合方式尚无法完全实现。目前，基于生物神经元整合

机制改进脉冲神经元的研究较为稀少。本文将聚焦于脉冲神经元的设计，分析生物神经元整合方式的优缺点，并结合机器学习技术，在有限的计算复杂度下对生物神经元的整合特性进行高效的实现。本文主要研究内容如下：

**双线性整合脉冲神经元的设计。**生物神经元的膜电位对输入的响应表现出双线性特性。然而，由于计算复杂度的限制，现有脉冲神经元仍沿用人工神经元的线性整合模式，这使得脉冲神经元在计算能力上相较于生物神经元有所欠缺。本文基于生物神经元的非线性整合模式，并结合因子分解机中的低秩矩阵分解技术，设计了一种新的脉冲神经元模型。在有限的计算复杂度下，该模型能够捕捉额外的二阶特征交互，从而使单个神经元也能够解决异或问题。随后，我们将该模型应用于不同的网络结构，并观察到网络表现出更强的学习能力，展现了其广泛的适应性。

**分流抑制脉冲神经元的设计。**生物神经元的抑制机制大致可分为超极化和分流抑制两种模式。传统的脉冲神经元沿用了人工神经元的线性整合模式，仅保留了超极化这一抑制机制。尽管分流抑制效应长期未被广泛重视，现代细胞内记录技术已证实其在大脑皮层细胞中的重要作用<sup>[47]</sup>。分流抑制通过改变细胞膜的电导率，引入更多的非线性因素，从而影响输入的整合模式，使单个脉冲神经元可以完成异或计算。本文基于生物神经元的分流抑制效应，并结合人工神经元中常见的门控机制，设计了一种新的脉冲神经元模型，并将其应用于脉冲神经网络中，以提高同规模网络对复杂输入模式的学习能力。

虽然这两个研究都是从非线性整合机制入手改进脉冲神经元模型，在研究思路和实验过程上有相似之处，但它们的研究出发点和优势侧重点存在明显差异。双线性整合脉冲神经元依托生物学数值拟合的研究结论，侧重于通过数据驱动的方法提升模型的计算能力，从而改善神经元在特征交互学习方面的表现；而分流抑制脉冲神经元则依据电生理学证据，注重模拟生物神经元真实的抑制机制，使模型在生物学合理性上更为贴近真实神经系统。两者各有侧重，共同为非线性整合脉冲神经元的设计提供了富有意义的两种思路。

本文除应用系统外的研究思路总结如图1-4所示，后续章节组织方式如下：

第二章简要介绍了生物神经元、人工神经元和脉冲神经元的基本模型，讨论了脉冲信号编码方案、神经网络结构及脉冲神经网络的训练方法。随后，总结了基于生物神经元对脉冲神经元进行改进的相关研究，并分析了这些改进的优势

与局限性，从而体现本文研究价值。

第三章介绍了生物神经元树突整合的双线性特性，分析了其异或计算能力，并指出该整合模式在计算复杂度上的不足。随后，此章基于矩阵分解思想，对双线性整合的实现进行了改进，提出了一种新的脉冲神经元模型，并将其应用于不同的网络结构中。最后，通过大量实验证明了该改进的有效性，并通过消融实验证明各模块在整体性能提升中的作用。

第四章探讨了生物神经元中抑制性输入的电生理学机制，并通过门控机制融合了超极化和分流抑制两种机制，提出了一种新的脉冲神经元模型。接着，本章将该模型应用于不同的网络结构中，并通过实验证实了此改进在提升网络性能方面的实际意义，同时通过消融实验证明了各组件设计的重要性。

第五章首先讨论了推荐系统的重要性以及脉冲神经网络在其中的意义，随后详细介绍了基于脉冲神经网络的推荐系统的设计与实现方案，最后展示了系统运行效果。

第六章对全文进行总结，并对未来的脉冲神经元相关研究进行展望。

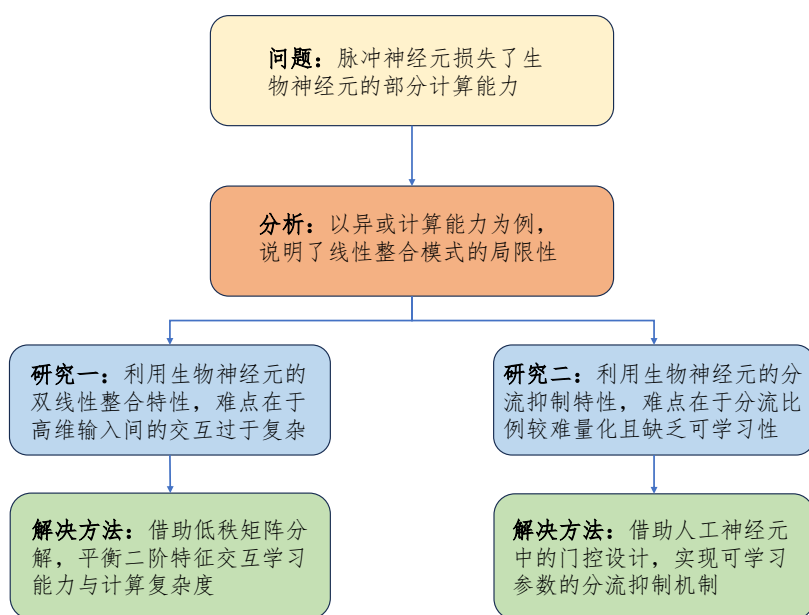


图 1-4 本文研究动机。研究一、研究二分别对应第三章和第四章所提出的模型。

## 第二章 相关工作

本章首先介绍脉冲神经网络相关的基础知识，以便读者更好地理解本文研究内容。随后，本章介绍了其他基于生物神经元特性设计新型脉冲神经元的工作，并对其各自的优势与局限性进行了分析，从而凸显本文研究意义。

### 2.1 脉冲神经网络预备知识

#### 2.1.1 神经元

##### 生物神经元

生物神经元是生物神经系统的基本功能单元，负责信息的接收、处理和传递，其基本组成包括细胞体、树突、轴突和突触等。细胞体是神经元的代谢中心，包含细胞核及遗传物质，负责能量的产生和神经递质的制造等。树突是从细胞体伸出的分支结构，主要功能是接收来自其他神经元或感受器的信号，通常较短。轴突是细长的突起，负责将神经元的输出信号传递给其他神经元。

神经元的细胞内膜与外膜之间存在电势差，称为膜电位（Membrane Potential）。在无外部刺激条件下，膜电位通常维持在-60至-70毫伏（mV），称为静息电位（Resting Potential），此时神经元处于极化（Polarization）状态，细胞内电位低于细胞外电位。生物神经元的细胞膜上分布有多种离子通道，这些通道对神经元的电生理活动具有重要意义。离子通道不仅参与了静息电位的形成与维持，还在神经元的信息处理过程中发挥关键作用。当神经元树突接收到兴奋性输入时，钠离子通道开放，钠离子迅速内流，导致膜电位向零电位或正电位方向变化，这一过程称为去极化（Depolarization）。相反，当树突接收到抑制性输入时，钾离子通道延迟关闭或氯离子通道开放，引起钾离子外流或氯离子内流，使膜电位进一步向负值偏移，称为超极化（Hyperpolarization）。

当神经元的兴奋性输入达到一定强度时，去极化使膜电位迅速升高，形成短

暂的电位波动，称为动作电位（Action Potential）。动作电位传导至轴突末梢时，促使神经递质释放，神经递质通过突触传递至下游神经元的树突，从而实现神经元之间的信息传递。这个过程称为神经元的激发。动作电位发生后的某段时间内，神经元无法再次激发，这段时间称为不应期，在此期间膜电位逐渐恢复到静息电位。

神经元间的信息传递主要通过突触实现，突触是神经元之间的连接点，分为电突触和化学突触两种类型。电突触信号传递速度快，几乎没有延迟，且能够双向传递，但不具备学习能力。相比之下，化学突触是神经元之间最常见的连接形式，其信号传递依赖于神经递质的释放，因此具有单向传递的特性，即信号只能从突触前神经元传递至突触后神经元。不同突触接收到的神经递质对膜电位的影响程度不同，这一影响程度被称为连接强度。突触前后神经元的连接强度是可变的，这一变化过程称为突触可塑性（Synaptic Plasticity）。突触可塑性的生理机制包括神经递质受体的改变和突触形态的重塑等，主要表现为长期增强（Long-Term Potentiation, LTP）和长期抑制（Long-Term Depression, LTD）。突触可塑性是大脑学习和记忆的基础。

## MP 神经元

生物神经元模型过于复杂，且其中许多机制至今仍未被完全理解，因此难以直接应用于计算系统中。MP 神经元是最早的人工神经元模型之一，采用了许多简化假设。例如：使用实数输入代替脉冲信号输入；忽略时间整合过程，当前输入直接决定膜电位，并在达到阈值时立即激发；每个神经元被视为一个多输入单输出的信息处理单元。尽管 MP 模型非常简单，但它揭示了神经元在信息处理中的基本作用，为神经网络的发展提供了理论基础。其数学表示如下：

$$\begin{aligned}z &= \mathbf{w}^\top \mathbf{x} + b, \\y &= \sigma(z).\end{aligned}\tag{2-1}$$

其中  $\mathbf{x}$ ,  $\mathbf{w}$  分别表示输入和权重， $b$  为神经元偏置项，对应生物神经元的静息电位， $z$  为输入的整合值，类似于生物神经元的膜电位， $y$  为神经元的输出。当输入整合值超过阈值时，MP 神经元会被激发，输出值为 1；否则，输出值为 0。因

此激活函数  $\sigma(\cdot)$  采用阶跃函数，其表达式如下：

$$\sigma(z) = \begin{cases} 1 & \text{if } z > \text{threshold,} \\ 0 & \text{otherwise.} \end{cases} \quad (2-2)$$

由于输入采用实数表示而非脉冲信号，MP 神经元仅需在单一时刻进行计算，无需像生物神经元那样在一段时间内持续产生输出。

当前大多数人工神经元是基于 MP 神经元改进的。虽然 MP 神经元能够执行逻辑计算，但在需要获取连续输出值时，阶跃函数表现不佳。且由于阶跃函数的梯度性质较差，使用反向传播训练基于 MP 神经元构成的神经网络很困难。为了适应不同数据集的需求并支持反向传播训练，人工神经元通常采用连续的激活函数，如 Sigmoid 函数和 Tanh 函数。为了缓解深度神经网络中梯度消失的问题，ReLU 激活函数被广泛应用于神经网络中，同时 Leaky-ReLU 被用来解决 ReLU 引起的神经元坏死问题。此外，门控机制也被引入激活函数设计中，如门控线性单元（GLU）<sup>[48]</sup>和 Swish<sup>[49]</sup>函数，它们在加速网络收敛速度方面表现出色。

## 脉冲神经元

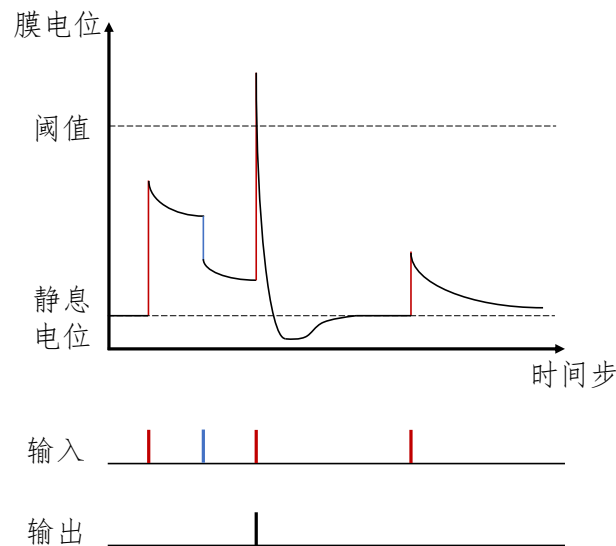


图 2-1 LIF 神经元模型示意图

LIF 模型是脉冲神经网络中最常用的神经元模型。与 MP 神经元相比，LIF 神经元设计更贴近生物神经元，其输入和输出均为脉冲信号，并具有膜电位的时

序整合作用。LIF 模型具有以下几个主要特点：

1. **泄漏 (Leaky)**。在不接受外部刺激的情况下，神经元的膜电位会逐渐向静息电位衰减。
2. **整合 (Integrate)**。神经元不仅会整合当前的输入，还会通过膜电位将历史输入与当前输入进行整合。
3. **激发 (Fire)**。当膜电位超过阈值时，神经元会激发，产生输出 1，并重置膜电位至静息电位。

在 LIF 模型中，输入电流与膜电位之间的关系可以通过以下微分方程表述：

$$\tau_m \frac{dV(t)}{dt} = -(V(t) - V_{\text{rest}}) + R_m I(t). \quad (2-3)$$

其中  $V(t)$  表示  $t$  时刻的膜电位， $V_{\text{rest}}$  为静息电位， $I(t)$  为电流输入， $\tau_m, R_m$  为常量。上述表达式中未考虑神经元在激发后膜电位的重置过程。在脉冲神经网络的实现中，通常会对时间进行离散化处理，LIF 模型可以表示为以下离散形式：

$$u_t = \alpha(u_{t-1} - b) + \mathbf{w}^\top \mathbf{x}_t + b. \quad (2-4)$$

其中  $u_t, u_{t-1}$  表示  $t$  和  $t-1$  时间步的膜电位， $\alpha$  为衰减系数，由公式 (2-3) 中的  $\tau_m$  决定， $\mathbf{w}, \mathbf{x}_t$  分别代表权重和  $t$  时间步的输入， $\mathbf{w}^\top \mathbf{x}_t$  对应  $R_m I(t)$ ，偏置  $b$  则代表了神经元的静息电位  $V_{\text{rest}}$ 。神经元的输出  $o_t$  由膜电位  $u_t$  决定：

$$o_t = \sigma(u_t) = \begin{cases} 1 & \text{if } u_t > u_{\text{thr}}, \\ 0 & \text{otherwise.} \end{cases} \quad (2-5)$$

其中  $u_{\text{thr}}$  表示膜电位的激发阈值。如果  $o_t = 1$ ，即神经元激发，则神经元会重置膜电位至静息电位。因此公式 (2-4) 可以改写为如下形式：

$$u_t = \alpha(u_{t-1} - b)(1 - o_{t-1}) + \mathbf{w}^\top \mathbf{x}_t + b. \quad (2-6)$$

由于 LIF 模型在保留生物神经元许多关键特征的同时，仍然保持足够的简单性和计算效率，因此在脉冲神经网络中得到了广泛应用。此外，LIF 模型具有

较好的可扩展性，因为它与人工神经元共享一些相似的结构，因此许多针对人工神经元的改进也可以被有效地应用于脉冲神经元。

### 2.1.2 脉冲编码

动态视觉传感器 (Dynamic Vision Sensor, DVS)<sup>[50]</sup>, 也称为事件相机 (Event Camera), 是一种事件驱动的视觉传感器。与传统相机逐帧记录数据不同, DVS 通过捕捉像素点的亮度变化事件 (包括时间戳、像素坐标和极性, 其中极性指亮度的升高或降低) 来记录数据。DVS 具有低能耗、低延迟和高动态范围等优势, 在机器视觉和自动驾驶等领域展现出良好的应用前景。DVS 采集的数据是一种神经形态数据, 即与生物神经系统使用的数据形态类似, 因此可以通过脉冲神经网络进行处理和计算。尽管目前已有一些公开的神经形态数据集, 但大部分数据集采集的仍为实数数据, 无法直接用于脉冲神经网络的计算。因此, 在使用脉冲神经网络进行实数数据的学习和推理时, 需要将实数数据转化为脉冲数据, 这一过程称为脉冲编码 (Spike Encoding)。与之相对, 将脉冲数据转化为实数数据的过程叫做脉冲解码 (Spike Decoding)。目前有多种脉冲编码方案, 这些方案源自于对生物脉冲的不同角度的研究。不同的脉冲编码方案适用于不同的脉冲神经网络模型和数据类型。

频率编码 (Rate Coding) 源自以下生物学研究观察: 生物神经元所受的刺激越强, 脉冲发放频率越高。频率编码认为, 神经元之间信息传递强度的度量方式是脉冲的频率。泊松编码 (Poisson Coding) 是频率编码的一种实现方式, 也是最常用的脉冲编码方案。其实现过程非常简单, 首先通过归一化等方法将原始输入变换到  $[0, 1]$  区间, 对于一个实数  $p \in [0, 1]$ , 泊松编码的脉冲序列  $\mathbf{x} = (x_0, x_1, \dots, x_{T-1})$  满足:

$$\text{Prob}(x_t = 1) = p, \text{ where } 0 \leq t \leq T - 1. \quad (2-7)$$

频率编码仅将脉冲发放频率作为信号强度的度量方式, 忽略了脉冲发放时间所蕴含的信息。近年越来越多的研究表明, 这种编码方式与生物神经元的实际工作机制存在差异。例如, 部分生物神经元仅需利用脉冲序列中的前几个脉冲即可完成计算任务。尽管如此, 由于频率编码实现简单且具有较强的抗噪声能力, 其仍然是当前大多数脉冲神经网络所采用的主要编码方案。

与忽略脉冲到达时间的频率编码不同，时间编码（Temporal Coding）将脉冲发放时间作为信号强度的衡量依据。延迟编码（Latency Coding）是时间编码的一种简单实现，其基于以下生物学观察：生物神经元受到的刺激越强，最早的脉冲到达时间越短。延迟编码生成的脉冲序列中仅包含一次脉冲发放，其发放时间代表了脉冲信号的强度，即发放时间越早，信号强度越高。全时间编码（Fully Temporal Coding）是一种更为复杂的时间编码方式，它将每一次脉冲的到达时间综合起来，共同作为脉冲强度的度量。相较于频率编码，时间编码具有更高的编码效率，但其抗噪声能力较弱，且与 LIF 神经元的匹配度较低，因此应用范围相对较窄。

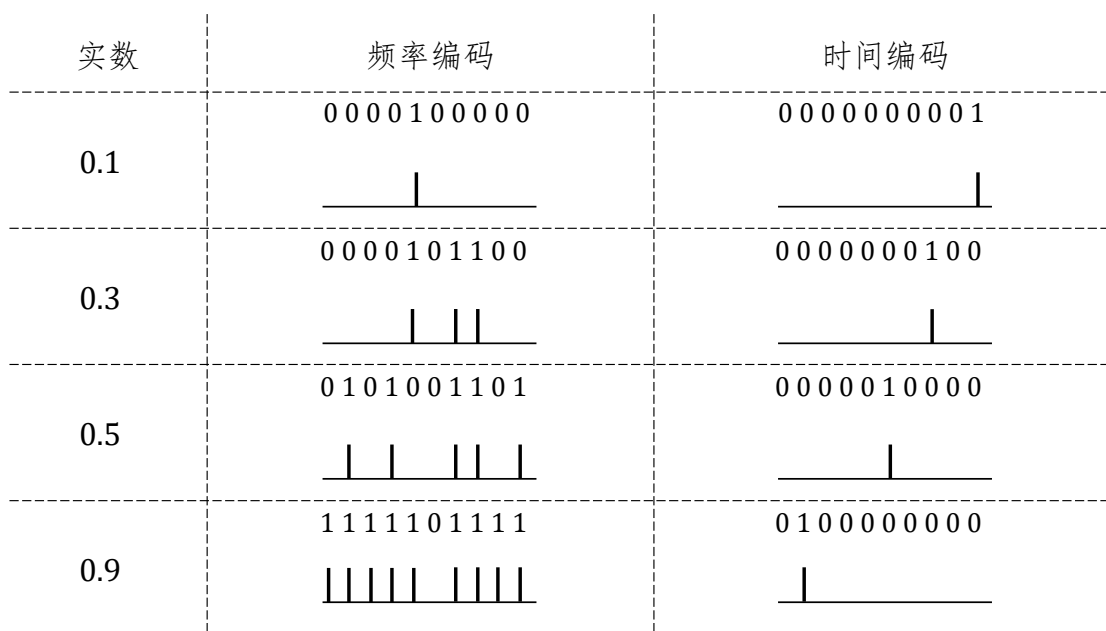


图 2-2 频率编码和时间编码的对比。频率编码采用泊松编码，时间编码采用延迟编码。

除了频率编码和时间编码，脉冲编码还有一些其他的实现方案，例如二元编码（Binary Coding）、相位编码（Phase Coding）等。此外，还有结合了多种编码方式的混合编码<sup>[51]</sup>，在保留良好抗噪声特性的同时，优化了编码效率和推理能效，能够实现比传统脉冲编码方案更优的效果。

### 2.1.3 神经网络结构

目前，大多数人工神经网络的信息传递是单向的，即从突触前神经元传递至突触后神经元。尽管存在一些具有特殊结构的网络，如霍普菲尔德网络和玻尔兹曼机，当前应用最广泛的仍是前馈神经网络。循环神经网络虽然能够整合当前输

入和历史状态，但其信息传递依然是单向的，不涉及反向信息传递。对于脉冲神经网络而言，全连接结构和卷积结构仍然是其主要架构。

全连接结构是神经网络中最常用的结构之一。除了仅由全连接层组成的全连接神经网络外，大多数神经网络即使包含其他结构，最后几层也通常是全连接层。在全连接结构中，网络层中的每个神经元与上一层的所有神经元都有连接，因此每个神经元通过不同的权重对上一层的所有神经元的信息进行汇总，从中提取出不同模式的信息。对于上一层的输入  $\mathbf{x}$ ，用  $\mathbf{W}$  表示权重矩阵，本层的输出可以表示为以下形式：

$$\mathbf{y} = \sigma(\mathbf{W}\mathbf{x}). \quad (2-8)$$

其中  $\sigma(\cdot)$  为某种激活函数。尽管全连接层实现简单、表达能力强，但其面临着计算开销大、容易过拟合等问题，尤其在处理高维数据或具有大量局部特征的数据时，往往表现不佳。

大脑皮层的神经元通常只与有限区域内的神经元相连接，也就是说，生物神经元的连接一般是稀疏且局部化的。受到这一现象的启发，卷积结构被引入到人工神经网络中。卷积结构使用共享权重的卷积核，利用滑动窗口对输入数据进行卷积操作，从而计算局部区域内神经元的加权和。通过共享权重和局部计算，卷积结构显著减少了所需的参数量。尽管如此，卷积操作对数据的整合模式仍然是线性的。以常用于图像数据的二维卷积为例，对于输入数据  $\mathbf{X} \in \mathbb{R}^{C \times H \times W}$ ，卷积核  $\mathbf{W} \in \mathbb{R}^{C \times S \times S}$ ，其中  $C$  表示通道数， $H, W$  为图像的高和宽， $S$  表示卷积核的大小，单个通道的输出特征图  $\mathbf{Y} \in \mathbb{R}^{1 \times H \times W}$  可表示为：

$$\mathbf{Y} = \sigma(\mathbf{W} * \mathbf{X}). \quad (2-9)$$

以上计算中使用了填充（Padding），即在输入特征图的边缘添加 0 元素，以保证卷积操作后特征图的大小不变。为了提取更多不同模式的特征，可以使用多个卷积核，每个卷积核的输出作为整体输出的一个通道。此外，还有一些特殊的卷积类型，如空洞卷积（Dilated Convolution）<sup>[52]</sup>、分组卷积（Group Convolution）<sup>[53]</sup>和深度可分离卷积（Depthwise Separable Convolution）<sup>[54]</sup>等。这些改进方法可以优化网络表现或减少计算开销，从而提高卷积神经网络的计算效率。

人工神经网络通常在 CPU、GPU 等通用计算硬件上进行同步、密集且连续

的计算，而脉冲神经网络则能够利用专用的神经计算硬件，进行异步、事件驱动且稀疏的计算。然而，由于学习算法和硬件的限制，目前脉冲神经网络的训练大多仍在通用计算硬件上进行。与人工神经网络通常使用 PyTorch、TensorFlow 框架实现不同，脉冲神经网络尚缺乏统一的框架。目前已有一些在通用计算硬件上实现脉冲神经网络的框架，如 SpikingJelly<sup>[55]</sup>、snnTorch<sup>[56]</sup>和 BindsNET<sup>[57]</sup>等。在这些框架下实现的脉冲神经网络通常具有类似于循环神经网络的结构，网络的输出通常表示为最后一层神经元的激发率，即脉冲输出在时间窗口内的累积值除以时间窗口长度。对于网络最后一层神经元的输出  $X \in \{0, 1\}^{T \times N}$ ，网络最终输出可表示为如下形式：

$$\mathbf{y} = \frac{\sum_{t=0}^{T-1} \mathbf{x}_t}{T}. \quad (2-10)$$

网络输出  $\mathbf{y} \in [0, 1]^N$  的每个元素  $y_i \in [0, 1]$  可以视为对每个类别的概率估计。

## 2.1.4 脉冲神经网络的训练

脉冲神经网络的学习方法主要可以分为三种类型：以 STDP 为代表的生物学启发式学习法则、ANN-SNN 转化以及替代梯度下降。每类方法都有其应用场景和优势，同时也存在一定的局限性。

### 生物启发的学习法则

赫布学习法则是神经科学领域早期提出的生物学习理论之一，为神经系统的无监督学习提供了生物学基础。该法则认为，如果两个神经元的激发时间接近，那么它们之间的连接强度会得到增强。然而，赫布法则仅描述了增强连接的机制，并未提供减弱连接的途径，而且也没有考虑神经元激发时间先后的影响。在赫布法则的基础上，STDP 考虑了时序依赖性，即神经元激发顺序的影响，并提出了基于时序依赖性的神经元连接强度调节机制，包括连接的增强和减弱。STDP 源自生物神经元的突触可塑性理论，包括长期增强作用和长期抑制作用。具体而言，若突触前神经元的激发早于突触后神经元，连接会增强；反之，若突触前神经元的激发晚于突触后神经元，则连接会减弱。例如，突触前神经元和突触后神经元分别在  $t_{\text{pre}}$  和  $t_{\text{post}}$  时刻激发，那么两个神经元之间的连接权重的改变

量可以表示为如下形式:

$$\Delta w = \begin{cases} A_+ \exp\left(\frac{t_{\text{pre}} - t_{\text{post}}}{\tau_+}\right) & \text{if } t_{\text{pre}} < t_{\text{post}}, \\ A_- \exp\left(\frac{t_{\text{pre}} - t_{\text{post}}}{\tau_-}\right) & \text{if } t_{\text{pre}} > t_{\text{post}}. \end{cases} \quad (2-11)$$

其中  $A_+ > 0$  和  $A_- < 0$  表示权重的最大改变量,  $\tau_+ > 0$  和  $\tau_- < 0$  表示时间常数。从上述公式中可见, 突触前后神经元的激发时刻间隔越短, 对连接权重的影响越显著, 表明两个神经元之间的关系越为密切。权重的正负则反映了神经元间关系的性质, 正权重表示兴奋性连接, 负权重则表示抑制性连接。

STDP 是脉冲神经网络中最常见且基础的生物学习法则, 已在无监督学习任务中取得了较多应用。公式 (2-11) 只是其实现方式之一, 还有其他多种具体的实现。STDP 具有比较充分的生物学可解释性, 并且是一种局部学习规则, 不需要反向传播, 计算负担小。然而, STDP 仅适用于浅层网络结构和无监督学习任务。部分研究将 STDP 扩展, 或将其与反向传播算法结合, 以应用到有监督任务和深层网络中<sup>[58-61]</sup>, 以拓宽其应用场景。尽管生物启发的学习法则在某些场景下具有计算开销和可解释性的优势, 但在提高网络准确性等性能指标方面, 仍难以与另外两种训练算法相竞争。

## ANN-SNN 转化

由于脉冲神经网络具有递归结构和非连续激活函数, 通过直接训练来优化其参数较为困难。作为一种替代方案, ANN-SNN 转化方法通过先训练一个人工神经网络, 然后利用该网络的参数来构造一个结构相同的脉冲神经网络, 从而实现脉冲神经网络的学习。从人工神经网络到脉冲神经网络的转换过程有多种实现方式, 其中一种常见的方法是权重映射, 并随后微调阈值。权重的映射通常以如下方法实现:

$$w_{\text{SNN}} = \alpha w_{\text{ANN}} \quad (2-12)$$

其中  $\alpha$  表示比例系数。阈值微调的目的是使脉冲神经元的输出与人工神经元的输出相对应, 即将连续值映射到脉冲的发放频率或发放时间。以频率编码为例, 人工神经元的输出经过归一化等变换之后, 得到实数  $y \in [0, 1]$ , 通过阈值微调,

使得脉冲神经元的输出  $\mathbf{y} \in \{0, 1\}^T$  满足以下关系:

$$y = \frac{\sum_{t=0}^{T-1} y_t}{T} \quad (2-13)$$

即人工神经元的输出与脉冲神经元的脉冲发放频率成正比。然而, 由于脉冲的离散性以及脉冲神经元的复杂时序特性, 以上转化过程通常会引入一定的误差。此外, 一些网络结构不支持这种操作, 例如最大池化层, 因为其运算是非线性的, 无法保持脉冲频率与实数之间的对应关系。因此, 一些研究致力于降低转化误差, 提出了对转化算法、损失函数等方面的优化方法<sup>[62-63]</sup>。

人工神经网络拥有成熟且高效的有监督学习算法, 反向传播算法在深度网络的训练中被广泛应用且有很多优化。ANN-SNN 转化方法基于人工神经网络的训练, 能够利用更加成熟的训练框架, 在大规模脉冲神经网络的训练中具有显著优势。处理实数数据时, ANN-SNN 转化是目前精度最高的脉冲神经网络训练方式。然而, 更高的转化精度的代价是更长的时间窗口, 这显著增加了脉冲神经网络推理阶段的计算开销。此外, 由于传统人工神经网络缺乏对时序特征的建模和学习能力, 在处理神经形态数据时, 通过 ANN-SNN 转化训练的脉冲神经网络无法充分利用数据中的时序特征, 导致性能表现欠佳。

## 替代梯度下降

替代梯度下降法在部分研究中也称为直接训练 (Direct Training)。ANN-SNN 转化法将预训练人工神经网络转化为脉冲神经网络, 与之不同, 替代梯度下降法直接在脉冲神经网络上进行端到端的训练。在前向传播过程中, 脉冲神经元的激活以公式 (2-5) 的形式进行; 对应地, 在反向传播过程中  $o_t$  对  $u_t$  的梯度计算如下:

$$\frac{do_t}{du_t} = \begin{cases} 0 & \text{if } u_t \neq u_{\text{thr}}, \\ +\infty & \text{if } u_t = u_{\text{thr}}. \end{cases} \quad (2-14)$$

由于脉冲神经元的激活函数总是为 0 或无穷大, 这导致传统梯度下降法无法直接应用于脉冲神经网络的训练。为了解决这一问题, 替代梯度下降法使用连续可微的函数来代替阶跃函数进行梯度计算, 即使用  $f(u_t)$  来代替  $\frac{do_t}{du_t}$ 。替代梯度  $f(\cdot)$

通常满足以下性质：

$$\begin{aligned} f(x) &\geq 0, \\ f(x) &\leq f(u_{\text{thr}}), \\ \int_{-\infty}^{\infty} f(x)dx &= 1. \end{aligned} \tag{2-15}$$

以确保参数优化方向的正确性以及梯度计算的稳定性。尽管替代梯度下降法使得脉冲神经网络的反向传播训练成为可能，但其仍面临梯度消失的挑战。具体而言，当膜电位远离阈值  $u_{\text{thr}}$  时，替代梯度  $f(u_{\text{thr}})$  趋近于零，导致梯度消失。为了缓解这一问题，研究者提出了多种方法，例如通过归一化技术优化膜电位分布<sup>[30-32]</sup>，或利用预训练对网络进行合适的参数初始化<sup>[64]</sup>。

在使用替代梯度下降法训练脉冲神经网络时，网络需要沿时间维度展开，这一过程与循环神经网络的训练类似，会带来较大的计算负担。尽管如此，替代梯度下降法能够有效利用数据中的时序特征，在处理神经形态数据时表现出显著优势。此外，该方法能够更充分地利用数据中的时序特性，因此可以使用更短的时间窗口，从而提升推理阶段的计算效率。值得注意的是，在反向传播过程中，替代梯度下降法仅在神经元的激活部分引入替代梯度，这使其与采用不同整合机制的新型脉冲神经元模型具有良好的兼容性。

## 2.2 生物学启发的脉冲神经元研究

尽管脉冲神经元相较于人工神经元更贴近生物学现实，但与真实的生物神经元相比，其模型仍然过于简化。在生物神经元的众多复杂特性中，许多为了计算简便性而被舍弃的机制可能对神经元的计算能力和表达能力具有重要影响。因此，脉冲神经网络研究的一个重要方向是基于生物神经元的特性设计新型脉冲神经元模型，以增强其表达能力或提升其在特定任务中的表现。

### 2.2.1 脉冲延迟

传统的人工神经元通常缺乏时序特性，其激发过程仅基于当前输入进行一次计算，并输出相应的激活值。相比之下，LIF 神经元具有时序特性，其膜电位会随时间动态变化。在每个时间步，LIF 神经元通过综合当前输入信号和历史膜电位状态计算当前膜电位，并根据阈值决定是否进行脉冲发放。值得注意的是，

在 LIF 神经元模型中，膜电位超过阈值和脉冲发放的过程是在同一时间步内完成的。然而，这种激发过程相较于生物神经元而言过于简化。实际上，当生物神经元的膜电位超过阈值后，神经元会进入激发模式（Active Mode），在此时膜电位会继续升高一段时间，随后神经元才会释放神经递质。这一现象中的时间间隔被称作脉冲延迟（Spike Latency）或突触时延（Synaptic Delay）。脉冲延迟的大小并非固定值，而是与膜电位的具体数值密切相关。具体而言，一种简化的描述脉冲延迟的模型<sup>[65]</sup>可以表示如下：

$$t_f = \frac{1}{u_t - 1}, \text{ if } u_t > u_{\text{thr}}. \quad (2-16)$$

其中  $t_f$  表示脉冲延迟， $u_{\text{thr}} = 1 + \epsilon$  为膜电位的阈值， $\epsilon > 0$ 。

在 LIF 模型中，由于使用阶跃函数作为神经元的激活函数，只要膜电位超过了阈值，神经元的响应模式便完全相同，而与膜电位的具体数值无关。在引入脉冲延迟之后，膜电位的数值差异会直接影响神经元的脉冲发放时间。这种机制不仅更贴近生物神经元的动态特性，还显著提升了脉冲神经元的表达能力，使脉冲能够编码更多的信息。然而，引入脉冲延迟会给脉冲神经网络的训练带来很大的挑战。在 LIF 模型中，神经元的输出  $o_t$  仅由当前时间步的膜电位  $u_t$  决定。而脉冲延迟的引入带来了复杂的时序依赖关系，使得脉冲输出  $o_t$  不仅可能由当前膜电位决定，还可能由历史膜电位  $u_{t-t_f}$  决定，其中  $t_f$  为动态变化的脉冲延迟时间。由于  $t_f$  并非固定值，基于时间展开的训练方法难以直接应用。此外，由于缺少对应的人工神经网络，带有脉冲延迟的脉冲神经网络也难以基于人工神经元进行模型转换。因此，常用的高性能训练方法，如 ANN-SNN 转换法和替代梯度下降法，均难以适用于此类模型。目前，针对带有脉冲延迟的脉冲神经网络，研究者<sup>[66]</sup>主要依赖于基于脉冲时间依赖可塑性的无监督学习方法，但这些方法通常仅能处理简单数据集上的任务，其性能仍有待进一步提升。

### 2.2.2 二阶脉冲神经元

在传统的 LIF 神经元中，受体接收到输入  $\mathbf{x}_t$  后，会产生一个瞬时的突触后电流  $i_t = \mathbf{w}^\top \mathbf{x}_t$ ，从而对膜电位  $u_t$  产生一次即时影响。然而，在生物神经元中，受体接收的刺激所引发的影响并非总是短暂且瞬时的。相反，这种影响可能具有

明显的持续性，导致突触后电流在较长时间内逐渐变化，从而对膜电位产生长期的动态调节。部分研究基于神经递质的持久影响特性，构建了可以利用二阶动力学特征的脉冲神经元模型<sup>[56]</sup>。其中一种实现方式如下：

$$\begin{aligned} i_t &= \beta i_{t-1} + \mathbf{w}^\top \mathbf{x}_t, \\ u_t &= \alpha u_{t-1} + i_t. \end{aligned} \quad (2-17)$$

其中， $\beta$  和  $\alpha$  分别代表电流和膜电位的衰减系数。为了表达式的简洁性，上述公式未考虑脉冲发放后的膜电位重置，且假设静息膜电位为 0。

二阶脉冲神经元在稀疏数据处理和捕捉长时间时序依赖关系等场景中表现更好。当脉冲编码方案中包含时序特性时，二阶脉冲神经元能够更有效地捕捉这些特性，从而提升模型的表现。然而，二阶脉冲神经元引入了更多的递归结构，这显著增加了模型的训练难度。此外，在使用频率编码时，二阶神经元试图学习长时间依赖模式往往是无效的，反而可能导致过拟合，从而对模型性能产生负面影响。因此，尽管二阶脉冲神经元在某些特定场景中表现优异，但在大多数实际应用中，其性能仍不及传统的 LIF 模型。

### 2.2.3 脉冲神经元不应期

生物神经元在激发后，会进入一段时间的不应期。在此期间，离子通道的通透性发生变化，神经元暂时无法接收新的输入，同时膜电位逐渐向静息电位衰减。这一现象具有一定的生物学合理性，因为神经元的神经递质资源并非无限的，其合成、运输和释放过程均需要消耗能量和时间，因此神经元难以持续保持激发状态。然而，目前绝大多数脉冲神经元模型为了简化实现，忽略了不应期的特性。部分研究设计并实现了具有不应期的脉冲神经元模型<sup>[67-68]</sup>，其主要实现方式包括但不限于以下几种：

1. 直接设置不应期：在神经元激活后，设置一段固定的时间窗口，在此期间神经元无法接收输入。
2. 动态阈值调整：在神经元激发后，提高其激发阈值，并使其随时间逐渐恢复到初始水平。
3. 超极化重置：在神经元激发后，将膜电位重置到一个比静息电位更低的值。

上述方法通过限制神经元的连续激发，一定程度上可以达到增强脉冲神经网络稳定性的效果。

然而，当前应用了不应期的脉冲神经元模型普遍存在实现复杂度高的问题。部分引入不应期的方式会形成复杂的时序依赖关系，导致参数学习过程面临梯度传播不稳定、时序对齐困难等挑战，显著增加了训练的计算负担。此外，对神经元连续激发的限制会直接降低其脉冲发放率，这种稀疏的脉冲发放模式可能削弱脉冲神经网络的时序信息传递效率，从而对模型的学习能力产生负面影响。值得注意的是，神经元连续激发引发的网络不稳定的问题，通常可通过膜电位归一化或梯度裁剪等方法有效缓解。因此，尽管不应期的生物合理性得到广泛认可，但其在现有脉冲神经网络中的实际应用价值有限，尚未在高性能任务中展现出明显优势。

### 2.3 本章小结

本章首先概述了脉冲神经元与脉冲神经网络的基础知识，为理解本文研究工作奠定理论基础。随后，从生物可解释性角度总结了现有的针对脉冲神经元模型的一些改进。然而，这些研究鲜有涉及生物神经元的—个关键特性，即非线性整合机制。此外，上述生物学特性在复杂任务中普遍面临两大挑战：其一，复杂的时序依赖关系导致训练稳定性下降；其二，较多引入了额外的计算开销。这些局限性表明，现有的基于生物学特性改进的脉冲神经元模型在生物合理性、表达能力与计算效率之间尚未达成有效平衡。

## 第三章 基于双线性整合的脉冲神经元模型

### 3.1 研究动机

本文第一章以异或计算为例，指出人工神经元在处理某些非线性计算时因采用线性整合机制而存在局限性，脉冲神经元亦有相同问题。为此，本章通过引入生物神经元的非线性整合机制解决脉冲神经元的非线性计算问题。本章模型基于以下生物学研究：双线性的整合模式能更准确地反映生物神经元的实际膜电位响应。随后，本章进一步以异或计算为例，证明了引入双线性整合机制可显著增强神经元的非线性计算能力。

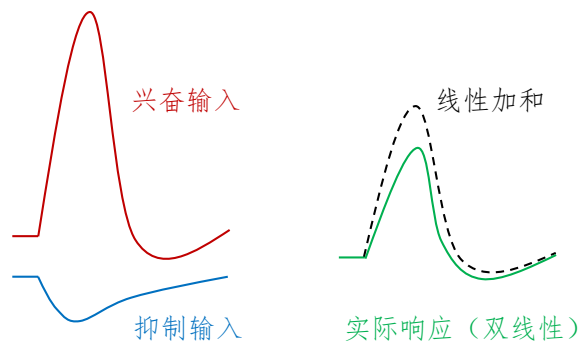


图 3-1 生物神经元膜电位响应示意图。线性整合不能很好地拟合实际膜电位响应。

然而，双线性的整合方式带来了计算复杂度大幅增加的问题，主要开销在于二阶特征交互的计算。为此，本文参考了因子分解机模型的低秩矩阵分解方法，实现了对双线性整合机制中二阶特征交互项的有效近似，从而在增强脉冲神经元非线性计算能力的同时，有效地控制了计算资源的需求。此外，考虑到双线性整合可能引发的数据分布问题及其他潜在挑战，进一步优化了模型设计。随后，通过分析脉冲编码对二阶特征交互的保持能力，论证了在脉冲神经元层面实施双线性改进的合理性。值得注意的是，针对整合机制的改进保持了脉冲神经元自身良好的泛化能力，使其能够无缝集成于各种网络架构中，而无需对现有系统进行重大调整或重新配置，这大幅拓宽了其应用范围。

## 3.2 生物学依据与关键技术

### 3.2.1 生物神经元的双线性整合

MP 神经元模型假设神经元对输入信号的整合是线性的，大多数人工神经元模型沿用了这种方式。在此基础上，以 LIF 模型为代表的脉冲神经元引入了时间维度的整合机制，即考虑了不同时刻输入的时序整合，从而使得脉冲神经元能够处理时序数据。然而，LIF 神经元在单一时刻对输入的整合仍然是线性的，这种线性的空间整合模式限制了其对脉冲时空特征的学习能力。生物学理论推导和仿真实验研究表明，使用双线性函数可以更好地拟合生物神经元的树突整合机制<sup>[45]</sup>。为了简化讨论，以没有膜电位以及时序整合机制的神经元为例，假设输入信号为  $x_1$  和  $x_2$ ，它们单独作用引起的电流分别为  $i_1 = w_1 x_1$  和  $i_2 = w_2 x_2$ ，其中  $w_1$  和  $w_2$  为输入对应的权重。根据双线性整合法则，这两个输入共同作用所产生的总电流可以表示为：

$$i = i_1 + i_2 + \kappa i_1 i_2. \quad (3-1)$$

其中， $\kappa$  常数表示分流系数，用于控制线性项与二阶特征交互项的比例。由于  $\kappa i_1 i_2 = \kappa w_1 w_2 x_1 x_2$ ，为了简化表达式，并且使其易于学习和优化，可以进一步将此系数合并为一个可学习参数，从而得到：

$$i = w_1 x_1 + w_2 x_2 + \kappa x_1 x_2. \quad (3-2)$$

对于多输入的情况，可以将总电流表示为各对输入之间双线性整合的和。使用向量  $\mathbf{x}$  表示神经元接收的输入信号，神经元的双线性整合电流可以表示为：

$$\begin{aligned} i &= \mathbf{w}^\top \mathbf{x} + \mathbf{x}^\top \mathbf{K} \mathbf{x}, \\ \mathbf{K}_{ij} &= \mathbf{K}_{ji} = \kappa_{ij}/2. \end{aligned} \quad (3-3)$$

其中对称矩阵  $\mathbf{K} \in \mathbb{R}^{N \times N}$  表示分流系数，且满足  $\mathbf{K}_{ii} = 0 (\forall i)$  以消除自交互项。

双线性整合模式不仅使得神经元更加贴近生物学现实，还显著增强了单个神经元的非线性表达能力。以异或计算为例，采用双线性整合的神经元可以仅通过一个神经元实现异或运算。具体而言，只需设置  $\mathbf{w} = (1, 1)$ ， $\mathbf{K} = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$ ，阈

值  $u_{\text{thr}} = 0.5$ ，即可完成异或计算。此外，双线性整合的神经元保留了传统神经元的线性计算部分，因此不会损失其原有的线性表达能力。

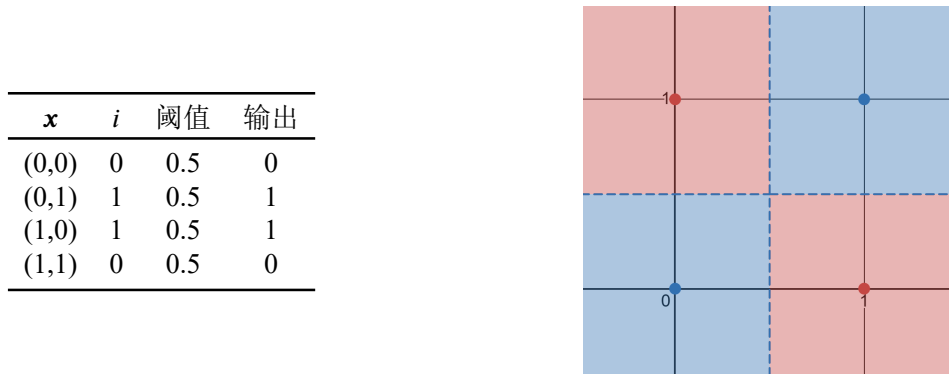


图 3-2 双线性整合神经元的异或计算。右图为神经元划分边界的可视化。

尽管双线性整合法则在神经元中的应用可以带来诸多优势，如增强非线性表达能力和生物学可解释性，但它在实际应用中却很少被采用，主要原因是其带来了过高的计算负担。以  $N$  输入  $M$  输出的全连接层为例，线性整合的参数数量和计算开销是  $\mathcal{O}(MN)$  的，而采用以上双线性整合方式时，参数数量和计算开销则是  $\mathcal{O}(MN^2)$  的。由于这一量级差距，双线性整合的实现和应用面临巨大的挑战，这也是其难以广泛应用的主要原因。

### 3.2.2 二阶特征交互的高效建模方法

在双线性整合模式中，计算瓶颈主要集中在二阶特征交互项上，因此高效地学习二阶特征交互成为本章研究的关键问题。实际上，许多实际场景中的数据普遍存在大量的二阶特征交互，而此类数据通常具有高维和稀疏的特点。例如，在推荐系统中，用户和商品的数量往往十分庞大，但单个用户通常只与商品集中的极少数项发生交互，因此直接对所有二阶特征交互进行建模通常因计算开销过高而不适用；而实现简单、计算效率高的线性模型又难以捕捉更复杂的高阶交互关系。在此背景下，诸如因子分解机<sup>[69]</sup>等模型通过采用高效的二阶特征交互建模方案，在模型表达能力和计算效率之间达到了较好的平衡，并因此得到了广泛应用。基于对现有机器学习模型的分析与借鉴，我们提出了对双线性整合模式中二阶特征交互项的高效建模方案。

接下来，本文将基于因子分解机的矩阵分解技术，对公式 (3-3) 的二阶特征交互项进行高效近似表示。根据矩阵分解理论，当矩阵  $\mathbf{K}$  正定时，存在 Cholesky

分解  $K = V^T V$ ，其中  $V = (\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{N-1}) \in \mathbb{R}^{k \times N}$ ，且满足  $k \geq \text{Rank}(K)$ 。因此，对于矩阵  $K$  的每个元素有  $K_{ij} = \mathbf{v}_i^T \mathbf{v}_j$ 。对于实际稀疏数据， $K$  通常具有低秩特性，因此  $k \ll N$ 。实际上，即使矩阵  $K$  不是正定的，也可以通过  $K_{ij} \approx \mathbf{v}_i^T \mathbf{v}_j$  来逼近。公式 (3-3) 的二次项部分可以重构为：

$$\begin{aligned}
 \mathbf{x}^T K \mathbf{x} &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} K_{ij} x_i x_j \\
 &\approx \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (\mathbf{v}_i^T \mathbf{v}_j) x_i x_j \\
 &= \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \left( \sum_{f=0}^{k-1} v_{if} v_{jf} \right) x_i x_j \\
 &= \sum_{f=0}^{k-1} \left( \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} v_{if} v_{jf} x_i x_j \right) \\
 &= \sum_{f=0}^{k-1} \left( \sum_{i=0}^{N-1} v_{if} x_i \right)^2 \\
 &= \sum_{f=0}^{k-1} (\mathbf{v}_f^T \mathbf{x})^2.
 \end{aligned} \tag{3-4}$$

由于  $K$  的对角线元素为 0，为了满足这一约束条件，需要去除所有元素与自身乘积的项。因此，二阶特征交互项可以表示为：

$$\mathbf{x}^T K \mathbf{x} \approx \sum_{f=0}^{k-1} ((\mathbf{v}_f^T \mathbf{x})^2 - (\mathbf{v}_f^2)^T \mathbf{x}^2). \tag{3-5}$$

其中  $\mathbf{v}^2$  表示  $\mathbf{v} \odot \mathbf{v}$ ， $\mathbf{x}^2$  表示  $\mathbf{x} \odot \mathbf{x}$ 。为了公式的简洁性，本文中向量的平方、矩阵的平方均指代其与自身的 Hadamard 积。通过这样的流程，原本  $\mathcal{O}(N^2)$  的计算开销被压缩到  $\mathcal{O}(kN)$ ，且对于稀疏矩阵有  $k \ll N$ 。

尽管传统的二阶特征交互建模方法也会与神经网络进行结合，但结合方法与本章在神经元层面深度融合的做法有较大差异。以因子分解机为例，其对于高维稀疏数据的情况有良好的表现，而神经网络对于稀疏数据表现往往不佳；相对地，因子分解机的表征能力和学习能力相较于深度神经网络而言仍显不足。为了解决这一问题，研究者们探索了将因子分解机与深度神经网络结合的方法，以充

分利用两者的优势。一个典型的例子是 DeepFM 模型<sup>[70]</sup>，它整合了因子分解机与神经网络模型。对于给定的输入向量  $\mathbf{x}$ ，DeepFM 的最终输出可以表示为：

$$\begin{aligned} y &= \sigma(y_{\text{DNN}} + y_{\text{FM}}), \\ y_{\text{DNN}} &= \text{DNN}(\text{Emb}(\mathbf{x})), \\ y_{\text{FM}} &= \text{FM}(\mathbf{x}). \end{aligned} \quad (3-6)$$

其中 Emb 是嵌入层，作用是将高维稀疏的输入转换为低维稠密的嵌入向量；DNN 为神经网络模型，负责学习嵌入向量中所蕴含的表征；FM 为因子分解机，用于提取输入变量间的二阶特征交互。上述传统方法仅停留在结果融合层面，并未深入神经网络内部结构。

### 3.3 双线性整合脉冲神经元模型实现与分析

#### 3.3.1 神经元实现

针对生物神经元双线性整合计算复杂度过高的问题，将 3.2.2 节所述方案应用到脉冲神经元中，这可以在增强神经元对二阶特征交互学习能力、增强非线性表达能力的同时，控制计算的开销。具体而言，通过矩阵低秩分解方法，神经元的双线性整合的公式 (3-3) 可以被重构为：

$$i = \mathbf{w}^T \mathbf{x} + \sum_{f=0}^{k-1} ((\mathbf{v}_f^T \mathbf{x})^2 - (\mathbf{v}_f^2)^T \mathbf{x}^2). \quad (3-7)$$

为了平衡计算能力和算法复杂度，并确保公式的简洁性，在本文后续的研究过程中，将设置隐含向量维度  $k = 1$ 。尽管如此，该模型仍能有效捕捉重要的二阶特征交互信息，从而提升模型的整体性能。

公式 (2-6) 给出了 LIF 神经元膜电位对输入的响应模式，其中  $\mathbf{w}^T \mathbf{x}_t$  这一项可以解释为  $t$  时刻输入引起的电流。当引入上述高效的双线性电流整合模式时，脉冲神经元的膜电位响应可以表示为：

$$u_t = \alpha(u_{t-1} - b)(1 - o_{t-1}) + \mathbf{w}^T \mathbf{x}_t + (\mathbf{v}^T \mathbf{x}_t)^2 - (\mathbf{v}^2)^T \mathbf{x}_t^2 + b. \quad (3-8)$$

通过这种方式，单个脉冲神经元能够以  $\mathcal{O}(N)$  的参数量和计算开销完成对二阶特征交互的建模，与一般的 LIF 神经元模型具有相同的渐进复杂度。这意味着增强了脉冲神经元对二阶特征交互的学习能力，使其更好地处理脉冲这种稀疏数据，同时模型的计算效率却并未显著下降。

以上线性项和二阶特征交互项的结合是通过简单加和实现的。然而，当这种机制应用于神经元时，会面临以下几个主要问题：

1. **数据尺度差异**。线性项和二阶特征交互项的计算方式不同，因此具有不同的数据分布，如果以相同的方法初始化权重矩阵，可能会导致网络出现无法训练的情况。例如，在某些情况下，二阶特征交互项的值可能远大于线性项，使得膜电位分布震荡明显，大部分膜电位大幅偏离阈值，从而导致梯度消失。
2. **特征贡献差异**。数据中线性项和二阶特征交互项的贡献不同，例如某些数据集可能以线性的特征分布为主，单纯地将两种特征加和可能导致模型过度拟合二阶特征交互部分，从而捕捉到噪声而非有效的特征，使网络训练不稳定，在处理脉冲这类稀疏数据时尤为突出。
3. **梯度计算与学习速率的差异**。二阶特征交互项的梯度和线性项的梯度计算方式不同，导致它们的学习速率也不同。具体来说，二阶特征交互项的权重参数的梯度涉及更复杂的非线性计算，与线性项权重的梯度分布有明显差异，可能会导致两个权重更新的速度不一致。
4. **二阶项权重矩阵调节能力受限**。因子分解机将表征二阶特征交互的矩阵分解为更小矩阵的乘积，导致其的每个维度的参数之间是耦合的，参数的学习会比较不灵活。这种耦合限制了模型的表达能力和优化效率。

为了解决以上这些问题，可以考虑给二阶特征交互项加入比例系数  $\lambda$ ，用以平衡线性特征和二阶特征交互的作用比例。公式 (3-8) 可以被重构为：

$$u_t = \alpha(u_{t-1} - b)(1 - o_{t-1}) + \mathbf{w}^\top \mathbf{x}_t + \lambda \cdot ((\mathbf{v}^\top \mathbf{x}_t)^2 - (\mathbf{v}^2)^\top \mathbf{x}_t^2) + b. \quad (3-9)$$

由于不同神经元可能对应的两种特征的比例都是不同的，因此比例参数  $\lambda$  被设置为可学习的。通过设置自适应的特征比例参数，可以灵活调整两种特征的贡献比例，增强神经元的表达能力，并且加速网络的收敛。

### 3.3.2 计算优势分析

如小节 3.2.1 所述，采用原始的双线性整合机制的神经元只需要单个神经元就可以完成异或计算。实际上，3.3.1 所述采用了高效双线性整合模式的脉冲神经元同样保留了部分异或计算能力。为了简化分析，仅对一个时刻的输入  $\mathbf{x}_t$  引发的电流进行分析：

$$i_t = \mathbf{w}^\top \mathbf{x}_t + \lambda \cdot ((\mathbf{v}^\top \mathbf{x}_t)^2 - (\mathbf{v}^2)^\top \mathbf{x}_t^2). \quad (3-10)$$

对于二维输入  $\mathbf{x}_t \in \{0, 1\}^2$  的情况，可以选择特定的参数值来实现异或计算。具体而言，只需令  $\mathbf{w} = (1, 1)$ ， $\mathbf{v} = (1, 1)$ ， $\lambda = -1$  即可。对于多维输入  $\mathbf{x} \in \mathbb{R}^N$  的情况，单个神经元也可以同时捕捉  $\Omega(N)$  数量的异或特征。虽然与原始的双线性整合机制相比，此方法能够表达的异或特征数量更少，但是对于脉冲这种稀疏数据而言，此方法通常能够捕捉足够多的特征，满足单个神经元的计算需求。

根据通用近似定理，当多层神经网络具备足够数量的神经元时，理论上能够逼近任意函数。然而，在实际应用中，神经网络的规模往往受到计算资源与时间的限制。因此，为了在有限资源条件下实现高效学习，需要根据数据特性选择合适的神经元模型和网络结构。脉冲数据作为一种稀疏且富含非线性特征的数据，其处理难度较大。本章所提出的模型在神经元层面支持二阶特征交互的学习，通过引入两个独立的权重矩阵分别捕捉脉冲输入的线性特征与二阶特征交互，不仅提升了脉冲神经网络对数据的拟合精度，同时优化了特征分解过程，有更快的网络收敛速度。

除了相对于传统脉冲神经元模型有所提升外，本章所述模型也提供了一种传统机器学习方法与神经网络结合的新模式。在以往的研究中，因子分解机和深度神经网络通常被视为并行存在的两个独立模型，它们各自的输出被结合在一起，两个模型共同参与训练过程。例如，在 DeepFM 模型中，因子分解机负责捕捉输入数据之间的二阶特征交互，而深度神经网络则用于挖掘复杂的非线性关系。这种组合方式虽然在一定程度上有效，但其局限性在于因子分解机仅学习了输入数据之间的二阶特征交互，并未充分利用其在捕捉这些二阶特征交互方面的强大能力来增强神经网络每一层的表达能力。具体来说，因子分解机并未被应用于神经网络内部的改进，尤其是神经元层面的改进。这意味着，尽管因子分解

机能够有效地处理高维稀疏数据中的特征交互问题，但它并没有直接参与到神经网络模型本身的学习过程中。

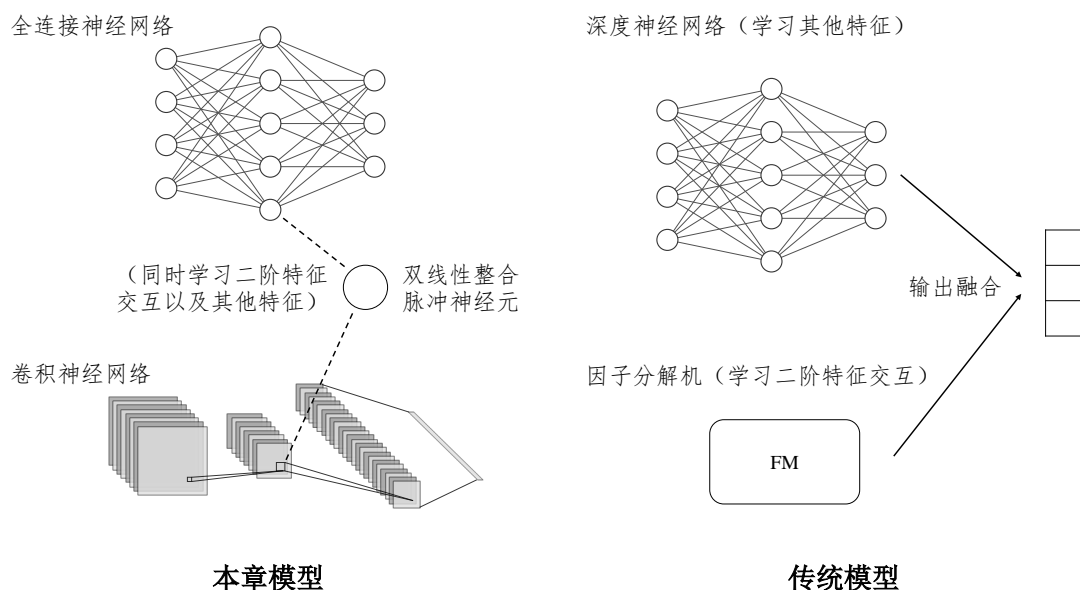


图 3-3 二阶特征交互学习方式对比。以往的工作将二阶特征交互的提取过程与深度神经网络并行，而本文工作在神经元层面实现二阶特征交互建模，拥有更好的拓展性。

与以往研究不同，本章工作在神经网络的每一个神经元中引入类似于因子分解机的机制。以此为基础，因子分解机和神经网络的结合不仅仅局限在两者的输出的融合，而是可以从根本上改变神经元的工作原理，使单个神经元也能够更精确地捕捉其输入特征间的复杂交互特征。如图3-3所示，将因子分解机在神经元级别实现，可以增强模型的适应性，使其能够处理不同类型的数据。例如，因子分解机通常难以直接学习图像数据中的特征，但基于因子分解机的神经元可以轻松集成到卷积神经网络中。

### 3.3.3 二阶特征交互与脉冲编码的适应性分析

双线性整合增强了神经元的非线性表达能力，例如异或计算任务。除此之外，公式 (3-3) 所表示的整合方式直接综合了线性项与二阶特征交互项。因此，采用双线性整合的人工神经元在处理包含较多二阶特征交互的数据时，通常能取得更好的表现。然而，在脉冲神经元中，输入脉冲的空间整合是在每个时间步进行的，随后通过时序整合来综合不同时间步的输入。对于包含二阶特征交互的实数数据，将其转化为脉冲后是否能保留原有的二阶特征交互，仍然是一个值得讨论的问题。

以实现最简单的频率编码——泊松编码为例，接下来将探讨其对线性特征和二阶特征交互的保留能力。假设输入为实数  $p, q \in [0, 1]$ ，有连接权重  $w_1, w_2$ ，通过泊松编码得到的脉冲序列为  $\mathbf{p}, \mathbf{q} \in \{0, 1\}^T$ 。根据泊松编码的定义，对每个时间步  $t$  ( $0 \leq t \leq T - 1$ )，有：

$$\begin{aligned} \text{Prob}(\mathbf{p}_t = 1) &= p, \\ \text{Prob}(\mathbf{q}_t = 1) &= q. \end{aligned} \quad (3-11)$$

这一过程可以视为  $T$  重伯努利实验。根据泊松小数定律，当  $T$  足够大，且  $p$  和  $q$  足够小，并且  $Tp$  和  $Tq$  大小适中时， $\mathbf{p}$  和  $\mathbf{q}$  的概率分布可以使用泊松分布来逼近。根据泊松分布变量的数学期望性质，可得：

$$\begin{aligned} E\left[\frac{\sum_{t=0}^{T-1} p_t}{T}\right] &= p, \\ E\left[\frac{\sum_{t=0}^{T-1} q_t}{T}\right] &= q. \end{aligned} \quad (3-12)$$

由于不同时间步生成的脉冲是独立的，进一步得出：

$$E\left[\frac{\sum_{t=0}^{T-1} w_1 p_t + w_2 q_t}{T}\right] \propto w_1 p + w_2 q. \quad (3-13)$$

这表明，泊松编码能够保留原有的线性特征。由于  $\mathbf{p}$  和  $\mathbf{q}$  的生成过程是互相独立的，因此两个随机变量的分布也彼此独立。假设  $\kappa$  为分流系数，则有：

$$E\left[\frac{\sum_{t=1}^{T-1} \kappa p_t q_t}{T}\right] \propto \kappa p q. \quad (3-14)$$

因此，泊松编码也能够保留原有的二阶特征交互。

由于泊松编码实现简单且能够同时保留线性特征和二阶特征交互，本章选择泊松编码作为编码方案，以适应提出的新型脉冲神经元模型。

### 3.3.4 全连接层与卷积层实现

在探索神经元层面的改进时，使用高效的双线性整合可以增强单个脉冲神经元对稀疏脉冲输入的复杂特征学习能力。这种改进策略不仅限于单个神经元，

而是能够广泛适用于脉冲神经网络的多种常见的结构。由于几乎所有神经元本质上都需要执行对输入信号的整合，因此这种从神经元整合机制入手的改进展现出广泛的适应性，而不受限于具体的网络拓扑结构。对于  $N \times M$  的全连接层，其中每个神经元都会接收到来自前一层  $N$  个神经元的相同输出，只不过后层  $M$  个神经元均拥有各自的权重，用于学习前一层神经元输出中所包含的不同的特征。对于卷积层，每个神经元仅关注前一层部分通道临近区域内的神经元的输出，而且同一个通道内的神经元会共享权重，但是单个神经元的运作机制仍然保持不变。因此，基于高效的双线性整合机制改进的脉冲神经元能够无缝集成到这些不同的网络结构中。

将对神经元的改进应用到脉冲神经网络的全连接结构中，对于  $t$  时间步的膜电位  $u_t \in \mathbb{R}^M$ ，其对输入  $x_t \in \{0, 1\}^N$  的响应可由以下表达式描述：

$$u_t = \alpha(u_{t-1} - b) \odot (1 - o_{t-1}) + Wx_t + \Lambda \odot ((Vx_t)^2 - V^2x_t^2) + b. \quad (3-15)$$

其中， $W \in \mathbb{R}^{M \times N}$  代表线性项权重矩阵， $V \in \mathbb{R}^{M \times N}$  为二阶特征交互项权重矩阵， $b \in \mathbb{R}^M$  为偏置项，可视为静息膜电位， $\Lambda \in \mathbb{R}^N$  为特征比例系数，允许每个神经元单独调节各自的线性项和二阶特征交互项的利用比例， $o_{t-1} \in \{0, 1\}^M$  为上一个时间步的神经元输出情况，决定是否将历史膜电位重置。这种改进只影响到了脉冲神经元的整合过程，而不涉及从膜电位到脉冲发放的作用机制，因此激发过程和传统的脉冲神经网络一致，在此不再赘述。

将对神经元的改进应用到脉冲神经网络的卷积结构中，对于  $t$  时间步的输入数据  $X_t \in \{0, 1\}^{C \times H \times W}$ ，神经元膜电位为  $U_t \in \mathbb{R}^{C' \times H \times W}$ ，其中  $C'$  为输出的通道数，即输出特征图的数量，膜电位对输入的响应方程可以表示如下：

$$U_t = \alpha(U_{t-1} - b) \odot (1 - O_{t-1}) + W * X_t + \Lambda \odot ((V * X_t)^2 - V^2 * X_t^2) + b. \quad (3-16)$$

其中， $W \in \mathbb{R}^{C' \times C \times S \times S}$  表示提取线性特征的卷积核， $V \in \mathbb{R}^{C' \times C \times S \times S}$  为提取二阶特征交互的卷积核， $b \in \mathbb{R}^{C'}$  为 channel-wise 的偏置项， $\Lambda \in \mathbb{R}^{C'}$  是 channel-wise 的特征比例系数，同一个通道内的神经元会共享特征比例， $O_{t-1} \in \{0, 1\}^{C' \times H \times W}$  为上一个时间步神经元的激发值。

## 3.4 实验与分析

本节通过对比实验来评估线性整合方式与双线性整合方式在脉冲神经网络中的表现差异。为了确保实验结果具有广泛的适用性和可靠性，本文选择了多样化的数据集进行测试，包括 IMDB、GTZAN、MNIST、N-MNIST、Fashion-MNIST、CIFAR-10、CIFAR10-DVS 以及 CIFAR-100 等。此外，实验采用了多种网络模型，具体包括基于 STBP<sup>[29]</sup>、PLIF<sup>[71]</sup>和 MPBN<sup>[32]</sup>的网络架构。其中，STBP 是广泛使用的脉冲神经网络训练方式（属于替代梯度下降法的一种实现），PLIF 和 MPBN 分别是 2021 年和 2023 年基于替代梯度下降法训练 SNN 的最佳模型代表。为进一步验证所提出方法的有效性，后续开展了消融实验，探讨不同组件和结构对整体性能的影响。

### 3.4.1 实验设置

#### 数据集与数据预处理

实验所用的数据集根据数据格式主要可以分为以下三类：向量数据、一般图像数据、神经形态数据。下面介绍这些数据集的基本信息和处理方式。

IMDB 数据集包含 50,000 条电影评论，每条评论标注为正面或负面情感，适用于二分类任务。该数据集分为训练集和测试集，各包含 25,000 条数据。为了便于模型处理，每条评论通过提取 TF-IDF 特征转化为 5000 维的向量表示。GTZAN 数据集是一个音乐数据集，包含 1,000 个音乐片段，覆盖 10 种不同的音乐类别。在实验中，选取 900 条数据作为训练集，剩余 100 条作为测试集。为了支持模型学习，原始音频被转换为梅尔频谱图及其他特征，转换后的每条数据包含 169 维特征。考虑到脉冲神经网络的神经编码过程，以上数据会在每个维度上进行归一化处理，以确保所有特征的取值位于  $[0, 1]$  范围内。

MNIST 数据集包含手写数字的灰度图像，共有 60,000 张训练图像和 10,000 张测试图像。每张图像的尺寸为  $28 \times 28$ ，每个像素值在 0 到 255 之间。尽管 MNIST 数据集应用广泛，但由于其相对简单，大多数图像仅需一个像素点即可区分，导致算法改进效果较为有限。因此，Fashion-MNIST 作为 MNIST 的替代品而被提出。Fashion-MNIST 数据集的规模与 MNIST 相同，包含 70,000 张  $28 \times 28$  的灰度图像。不同于 MNIST 中的数字，Fashion-MNIST 中的每张图像代表一件

服装，数据集共包含 10 个服装类别。由于 Fashion-MNIST 在规格上与 MNIST 一致，它成为了在更复杂模型中进行实验的理想替代。

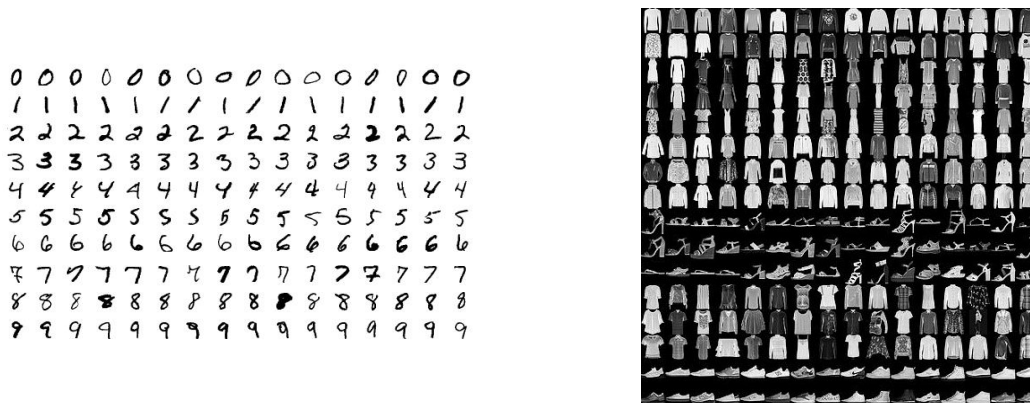


图 3-4 MNIST 和 Fashion-MNIST 数据集

CIFAR 系列数据集对脉冲神经网络而言具有较高的难度。CIFAR-10 数据集包含 60,000 张  $32 \times 32$  的三通道彩色图像，分为 10 个类别，其中 50,000 张为训练样本，10,000 张为测试样本。CIFAR-100 数据集与 CIFAR-10 在规模上相同，也包含 60,000 张  $32 \times 32$  的三通道彩色图像，但其包含 100 个类别，因此学习难度相较 CIFAR-10 更大。

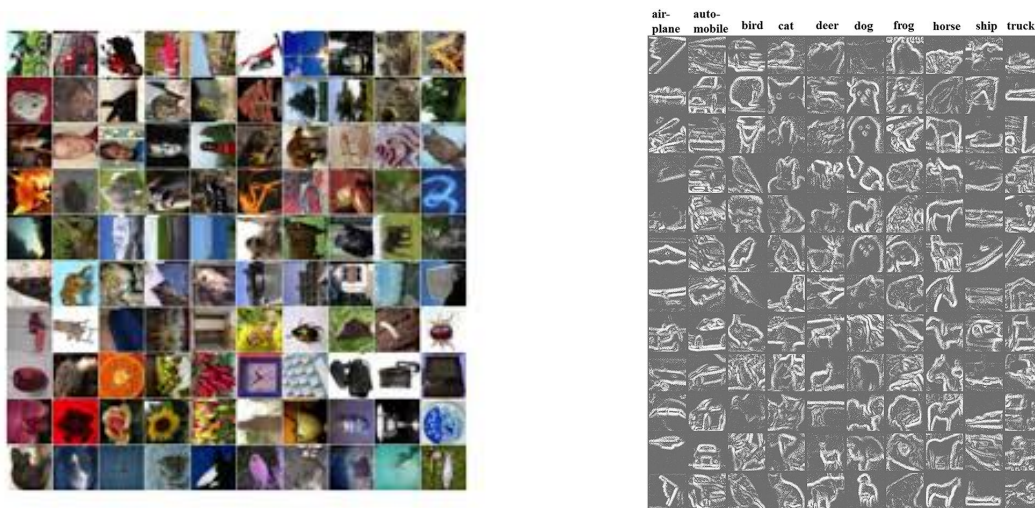


图 3-5 CIFAR100 和 CIFAR10-DVS 数据集

除了向量数据和一般图像数据，本章也在 N-MNIST 与 CIFAR10-DVS 这两个神经形态数据集上进行了实验。N-MNIST 是通过 DVS 传感器捕捉手写数字创建过程而形成的数据集，而 CIFAR10-DVS 则是通过 DVS 记录 CIFAR-10 重复闭环平滑（Repeated Closed-loop Smooth, RCLS）移动而形成的。神经形态数据集通常是基于事件的（Event-based）数据，即时间戳-像素坐标-极性的格式，其中极

性指的是亮度的变化方向（上升或下降），在输入给脉冲神经网络之前需要将其转化为基于框架的（Frame-based）数据，即  $T \times 2 \times W \times H$  的形式，这里  $T$  代表时间窗口的长度，两个通道对应两种极性。实验中神经形态数据的加载和转化是通过 SpikingJelly 框架实现的。N-MNIST 数据集的数据量以及类别数量和 MNIST 相同，每个样本可以转化成  $T \times 2 \times 34 \times 34$  的张量。CIFAR10-DVS 包含 10000 条数据，每条数据可以转化为  $T \times 2 \times 128 \times 128$  的张量。由于 CIFAR10-DVS 本身未进行训练集和测试集的划分，实验中抽取 10% 作为测试集。

表 3-1 对图像数据的预处理。\* 标记的操作只在训练集中进行。

数据集	预处理
CIFAR-10	RandomCrop(32, padding=4),* RandomHorizontalFlip(*),* ToTensor(*), Normalize((0.4914, 0.4822, 0.4465), (0.2023, 0.1994, 0.2010))
CIFAR-100	RandomCrop(32, padding=4),* RandomHorizontalFlip(*),* ToTensor(*), Normalize((0.5071, 0.4867, 0.4408), (0.2675, 0.2565, 0.2761))
其他图像数据集	ToTensor(*)

神经形态数据经过格式转化可以直接作为脉冲神经网络的输入，但是实数数据还需要经过其他预处理步骤。向量数据的预处理已在数据集介绍中说明，对于一般图像数据，还需要进行图像预处理，具体方式如表3-1所示，包括填充、裁剪、归一化等操作。之后，实数数据还需要经过脉冲编码处理，本实验中采用泊松编码方式，将连续的实数值转换为脉冲序列，作为脉冲神经网络的输入。

在进行实验时，对于使用 PLIF 搭建的网络，从原始训练集中随机抽取 15% 作为验证集，其余数据用于训练模型。训练完成后，根据验证集准确率确定最佳模型，并报告其在测试集上的准确率。对于其他网络，则不设置验证集，报告的是测试集上的最高准确率。以上准确率数据会取三次运行的平均值。

## 网络设置

本实验采用多种网络架构作为基线，通过替换脉冲神经元的整合机制来突出双线性整合的优势。需要指出的是，由于实验资源和时间开销的限制，本文的实验配置与原始基线文献中的设置有所差异，导致准确率存在一定差异。例

如，PLIF 模型原工作采用了带动量的随机梯度下降（SGD with Momentum）进行优化。然而，在现有实验环境下（NVIDIA GeForce RTX 2080 Ti），使用该方法训练一个完整的网络耗时超过 20 天。为提高效率，本文选择了收敛速度更快的 Adam 优化器，并适当减少了训练轮数。此外，由于显存的限制，部分实验调整了时间窗口大小并降低了批大小。为了保证实验对比的公平性，若无特殊说明，基线和双线性整合神经元的实验采用相同的超参数设置。实验中作为基线的神经网络包含以下三类：

1. **基本的脉冲神经网络**。使用 STBP 方法训练基本的脉冲神经网络，实验中全连接网络和浅层的卷积网络使用此方法搭建。
2. **基于 PLIF 的脉冲神经网络**。此方法用于搭建处理 CIFAR-10 和 CIFAR10-DVS 数据的深度神经网络。
3. **基于 MPBN 的脉冲神经网络**。此方法用于搭建处理 CIFAR-10 和 CIFAR-100 的深度神经网络。

在实验中，为了兼顾不同的数据集和模型，所有的脉冲神经网络均使用替代梯度下降法进行训练，即在正向传播阶段使用阶跃函数作为激活函数，而在反向传播时，则使用替代梯度来进行梯度近似计算。实验中第一类和第三类的脉冲神经网络采用如下替代梯度函数：

$$f_1(u_t) = \begin{cases} \frac{1}{2a} & \text{if } |u_t - u_{\text{thr}}| \leq a, \\ 0 & \text{otherwise.} \end{cases} \quad (3-17)$$

在实验中取  $a = 0.5$ 。第二类脉冲神经网络采用的替代梯度函数为  $\arctan$  函数的导数，具体表示如下：

$$f_2(u_t) = \frac{1}{1 + (u_t - u_{\text{thr}})^2}. \quad (3-18)$$

两种替代梯度函数均满足公式 (2-15) 的约束条件。

在基础的脉冲神经网络中，实验使用了 LIF 模型。在 LIF 模型构建的神经网络中，所有神经元共享一个衰减系数，该系数通常是一个预先设定且不可调节的超参数。这导致所有神经元对输入的时间整合采用完全相同的模式，从而限制了脉冲神经网络在挖掘脉冲输入间复杂时空特征方面的能力。已有研究<sup>[72]</sup>从理论上证明了可调节衰减系数对脉冲神经网络时序学习能力的显著影响，并提

出了一种根据特征值设置衰减系数的方法。相比之下，PLIF 模型则是将衰减系数设为可学习参数，通常衰减系数的取值在  $(0, 1)$  之间，因此 PLIF 使用 Sigmoid 函数对衰减系数进行约束。PLIF 神经元模型可以表示如下：

$$u_t = (1 - \sigma(a))(u_{t-1} - b)(1 - o_t) + \sigma(a)i_t + b. \quad (3-19)$$

其中， $\sigma(\cdot)$  为 Sigmoid 函数， $a$  为可学习参数，用于控制衰减系数， $u_t$  和  $i_t$  分别为  $t$  时间步的膜电位和电流。需要注意的是，由于本章对神经元的改进仅涉及神经元输入的整合机制，因此通过修改输入项  $i_t$ ，即可实现双线性整合与 PLIF 模型的结合。这体现了本章提出的双线性整合神经元与其他模型的良好兼容性。

PLIF 是对脉冲神经元的改进，而 MPBN 则是针对脉冲神经网络训练过程的改进。批归一化常用于深度神经网络的训练，通过规范化数据输入，减少协变量偏移现象，从而加速网络的收敛并提高网络的稳定性。然而，由于脉冲神经元具有复杂的时空整合机制，尽管批归一化可以使输入整合得到电流分布更加合理，但膜电位的时间整合、膜电位重置重置等非线性操作会破坏这种分布。因此，MPBN 方法提出了一种直接对膜电位进行批归一化的做法，以解决这一问题。实验中使用 MPBN 搭建了深度脉冲神经网络，体现了本章提出的神经元改进方法在不同网络中的适应性。

**表 3-2 实验网络参数设置。有括号标注的设置仅在指定数据集或模型上生效。**

参数名称	设置
优化器	Adam
初始学习率	0.001, 0.0001 (IMDB)
训练轮数	100, 20 (IMDB), 200 (CIFAR-100)
批大小	100, 128 (MPBN), 16/8 (PLIF, CIFAR-10/CIFAR10-DVS)
阈值	0.5, 1.0 (PLIF)
衰减系数 $\alpha$	0.25, 自适应 (PLIF)
偏置	有, 无 (PLIF 和 MPBN)

本实验中，网络的参数设置详见表3-2。由于实验中使用的脉冲神经网络结构较多，具体的结构将在实验结果章节中详细说明。使用 PLIF 构建的脉冲神经网络结构如表3-3所示，其中投票层的含义是每 10 个神经元的输出之和作为最终输出的一个维度。100 个神经元为 10 个类别提供概率估计，这样的设计有助于提高网络的稳定性。为了便于表示，后文将使用 PLIF-Net1 和 PLIF-Net2

分别表示在 CIFAR-10 和 CIFAR10-DVS 数据集上使用 PLIF 构建的网络。使用 MPBN 构建的脉冲神经网络则以 ResNet20 作为基础结构。

**表 3-3 基于 PLIF 搭建的脉冲神经网络结构。其中 P2 指 MaxPool2d, Drop 为 dropout 层, 丢失概率设置为 0.5, Vote 为投票层。**

数据集	结构
CIFAR-10	256C3-256C3-256C3-P2-256C3-256C3-256C3-P2-Drop-2048-100-Vote
CIFAR10-DVS	128C3-P2-128C3-P2-128C3-P2-128C3-P2-Drop-2048-100-Vote

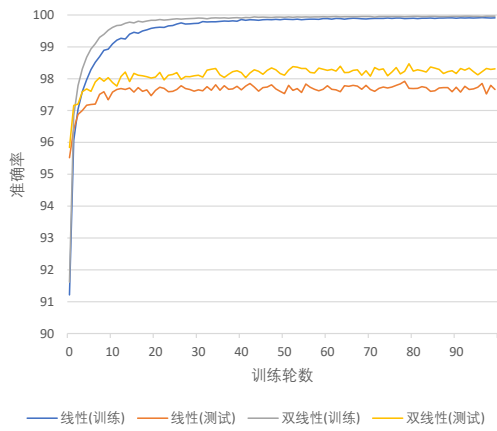
需要注意的是, 在脉冲神经网络的实现中, 为了便于进行替代梯度的计算, 通常会设置一个统一且固定的阈值电位, 并通过一个可学习的偏置参数来调整静息电位和阈值电位之间的差值。这种方式与设置可学习的阈值具有相同的表达能力, 同时简化了模型实现的复杂性。但在某些研究的具体实现中, 由于代码设计的原因, 膜电位会向其他值而非向静息电位(偏置)衰减, 这一现象可能导致模型行为偏离理论预期, 影响实验结果的有效性和一致性。因此, 本实验会对脉冲神经元的实现进行统一调整, 使其符合公式 (2-6) 所表示的规范的脉冲神经元模型。

### 3.4.2 全连接网络对比实验结果与分析

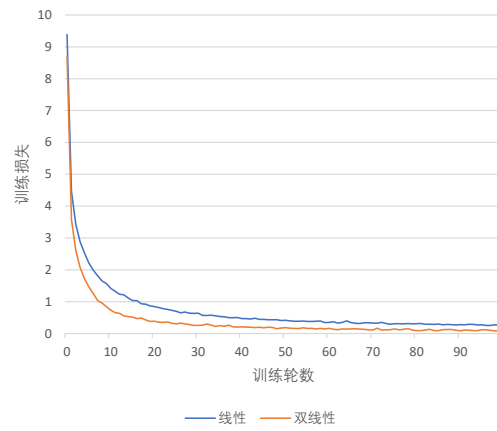
**表 3-4 双线性整合在全连接网络中的实验结果**

数据集	时间窗口	网络结构	整合方式	$\lambda$ 初值	准确率 (%)
IMDB	20	5000-100-2	线性	-	85.91
			双线性	0.01	<b>86.36</b>
GTZAN	20	169-50-10	线性	-	65.00
			双线性	0.01	<b>68.33</b>
MNIST	20	784-100-10	线性	-	97.96
			双线性	0.1	<b>98.36</b>
		784-200-10	线性	-	98.42
			双线性	0.1	<b>98.66</b>
784-400-10	线性	-	98.62		
	双线性	0.1	<b>98.74</b>		

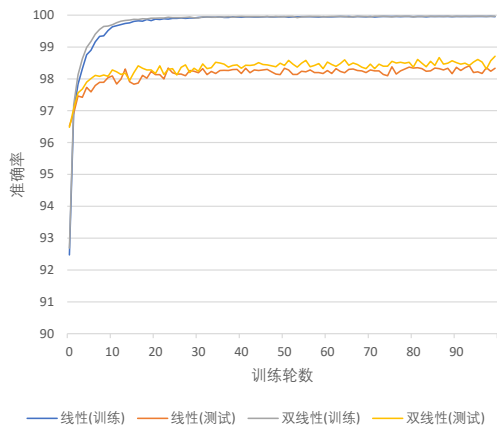
实验首先在全连接脉冲神经网络上进行, 因为全连接结构是脉冲神经网络的最基础形式。考虑到全连接网络在学习能力上的局限性, 本节在一些简单数据集上使用浅层网络进行了测试, 结果记录在表3-4中。从数据中可以看出, 双线性整合方式在不同结构的全连接网络中, 对不同数据集均有一定程度的提升。



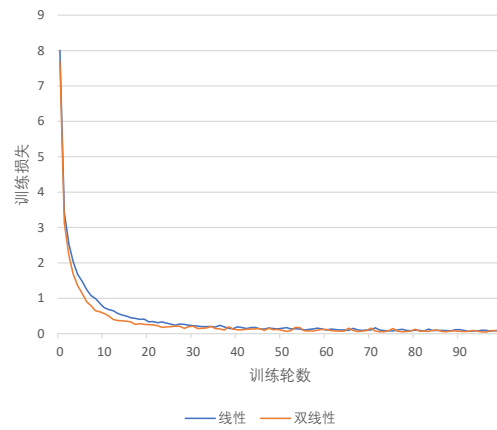
准确率曲线, 100 个隐层神经元



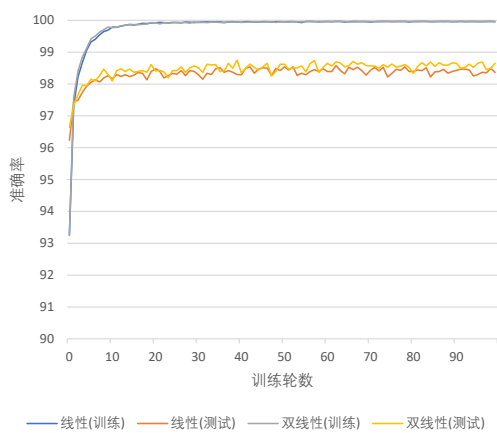
损失曲线, 100 个隐层神经元



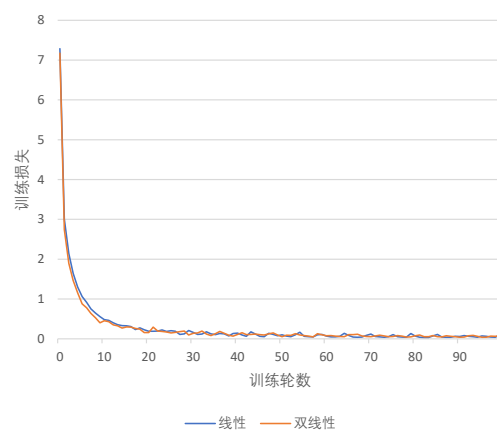
准确率曲线, 200 个隐层神经元



损失曲线, 200 个隐层神经元



准确率曲线, 400 个隐层神经元



损失曲线, 400 个隐层神经元

图 3-6 MNIST 数据集上全连接脉冲神经网络使用线性与双线性整合的训练曲线对比

在处理 IMDB 和 GTZAN 等向量数据集时，双线性整合机制相较于传统线性整合方法展现出更显著的性能优势。这主要是由于向量数据的不同特征具有较为明确的含义，不同特征之间更具有交互性。传统的神经元通常依赖于深层网络结构来学习复杂的非线性特征交互，但这在神经元数量有限或网络层数较浅的情况下可能导致学习效果不佳。而双线性模型能够更高效地捕捉特征之间的二阶交互关系，从而增强模型的非线性表达能力，这在神经元数量较少的浅层网络结构中表现尤为突出。

对于 MNIST 这类图像数据，实验将其展平为 784 维的向量作为输入。实验结果表明，双线性整合仍然具有一定的优势。随着神经元数量的增加，两种整合方式的差距有所缩小。然而，双线性整合神经元凭借更强的非线性表达能力，仍保持更高的计算效率。例如，使用 200 个隐层神经元的非线性整合脉冲神经网络，其表现超过了拥有 400 个隐层神经元的传统脉冲神经网络。

除了测试集准确率这一指标外，模型训练过程中的训练集损失和训练集准确率变化也能有效反映模型的学习能力和学习速度。全连接脉冲神经网络在 MNIST 数据集上的训练过程中的准确率和损失曲线如图3-6所示。从图中可以得知，使用双线性整合的脉冲神经网络最终的损失值更小，表明该模型在拟合数据分布方面具有更强的能力。同时，训练损失曲线的变化趋势表明，双线性整合脉冲神经网络具有更快的收敛速度。以上的这些优势在神经元数量更少时会更加凸显，契合脉冲神经网络轻量化、低能耗应用场景的需求。

#### 3.4.3 卷积网络对比实验结果与分析

为了展示神经元对不同网络结构的适应能力，实验也在卷积网络上进行了测试。卷积网络通常用于处理图像数据，因此实验分别在 MNIST、Fashion-MNIST、CIFAR-10 和 CIFAR-100 数据集上进行。此外，本章也在两个神经形态数据集 N-MNIST 和 CIFAR10-DVS 上进行了实验。实验使用的网络包括基于 STBP 构建的浅层脉冲神经网络，以及基于 PLIF 或 MPBN 构建的深度脉冲神经网络。实验结果见表3-5，数据表明，在大多数场景下，脉冲神经元使用双线性整合方式整体优于线性整合方式。

首先，对于基于 STBP 构建的一些基础的浅层卷积脉冲神经网络，我们在 MNIST 和 N-MNIST 数据集上进行了实验。实验结果表明，双线性整合能够提升

表 3-5 双线性整合在卷积网络中的实验结果。表中的卷积网络结构省略了输入部分。

数据集	时间窗口	网络结构	整合方式	$\lambda$ 初值	准确率 (%)
MNIST	20	32C3-P2-32C3-P2-128-10	线性	-	99.41
			双线性	0.01	<b>99.48</b>
N-MNIST	10	32C3-P2-32C3-P2-128-10	线性	-	99.42
			双线性	0.01	<b>99.52</b>
Fashion-MNIST	20	32C3-P2-32C3-P2-128-10	线性	-	91.12
			双线性	0.01	<b>91.41</b>
CIFAR10	8	PLIF-Net1	线性	-	89.95
			双线性	0.1	<b>91.08</b>
	4	MPBN-ResNet20	线性	-	89.82
			双线性	0.1	<b>90.07</b>
CIFAR10-DVS	10	PLIF-Net2	线性	-	<b>67.67</b>
			双线性	0.1	67.53
CIFAR100	4	MPBN-ResNet20	线性	-	66.01
			双线性	0.1	<b>66.42</b>

脉冲神经网络的表现，但提升幅度相对有限。这是因为 MNIST 和 N-MNIST 数据集对于卷积网络而言较为简单，网络的精度已经接近饱和，因此不太需要额外的表达能力增强。为进一步验证双线性整合的效果，在难度更大的 Fashion-MNIST 数据集上进行了实验，结果表明，双线性整合在该数据集上能够带来更为明显的精度提升。

CIFAR-10 和 CIFAR-100 数据集相对较为复杂，因此在这些数据集上的实验使用了基于 PLIF 和 MPBN 的深层脉冲神经网络。对于 CIFAR-10 数据集，使用双线性整合能够较大提高基于 PLIF 的网络的精度，但在 MPBN 网络上提升较小。这可能是由于双线性整合优化了膜电位的分布，而膜电位的批归一化操作在一定程度上抹平了这一优势。尽管如此，双线性整合仍保持了非线性计算能力上的优势，因此在 CIFAR-10 和 CIFAR-100 数据集上，整体精度仍有所提升。

对于 CIFAR10-DVS 数据集，由于 MPBN 与神经形态数据的兼容性较差（原因在于事件流中不同物体的出现时间可能不同，不适合批归一化，这与大语言模型不使用批归一化的原理相似），因此实验中仅使用 PLIF 构建深度脉冲神经网络。实验结果表明，双线性整合在 CIFAR10-DVS 数据集上并未表现出明显的优势，甚至略低于线性整合。为了探究这一现象，我们分析了训练过程中的训练集准确率和损失曲线，结果如图3-7所示。双线性整合使得网络在训练集上的准确率更高，损失更低，表明双线性整合的脉冲神经网络在训练集的拟合效果更好。

基于此事实，我们推测 CIFAR10-DVS 数据集中的二阶特征交互较少，导致双线性整合模型在训练数据上发生了过拟合，从而影响了测试集的准确率。通常，图像数据经过泊松编码后可以保留更多的二阶特征交互，但 CIFAR10-DVS 的产生原理不同，因此其特征分布与一般图像数据集有所不同。

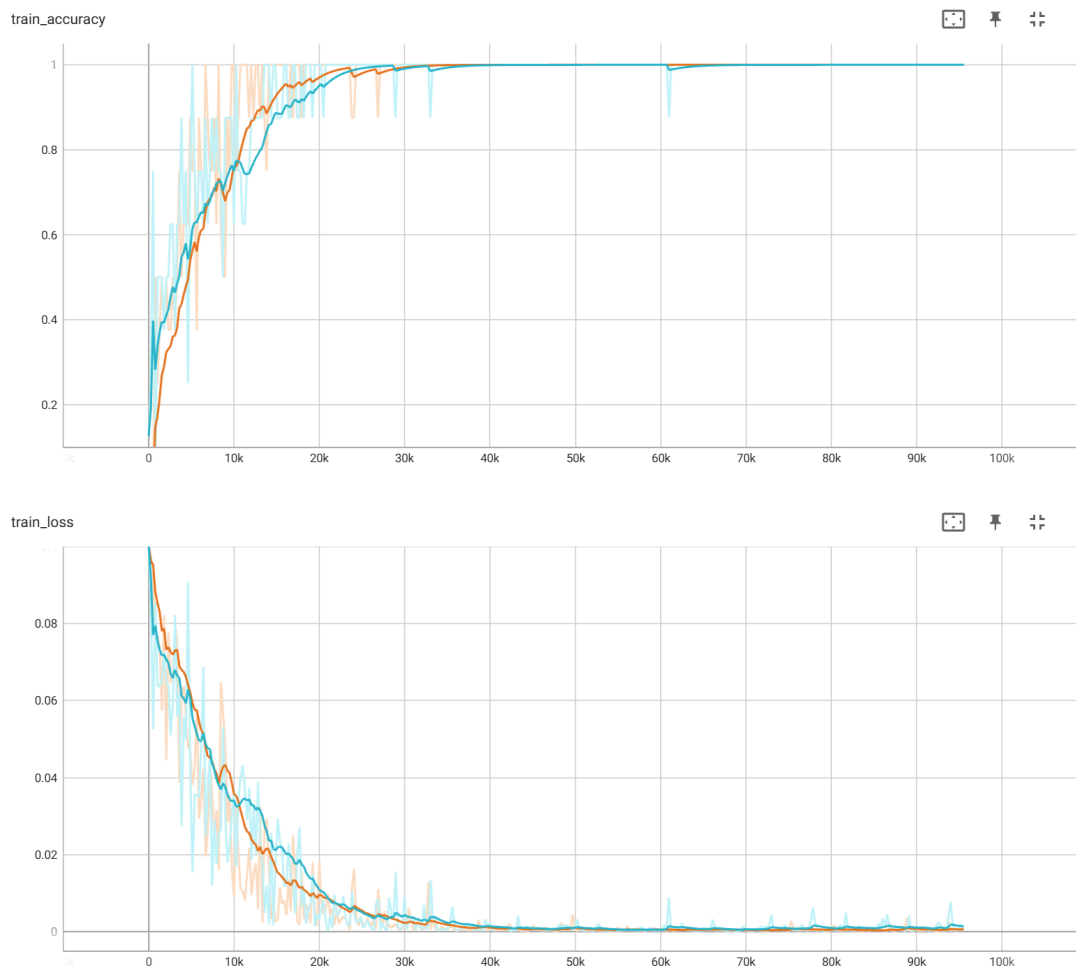


图 3-7 基于 PLIF 的深度脉冲神经网络在 CIFAR10-DVS 数据集上的训练曲线。蓝色为线性整合，橙色为双线性整合，横坐标为训练的步数，TensorBoard 展示的曲线经过了平滑来达到更好的展示效果，浅色曲线为未经平滑的原始数据。

#### 3.4.4 消融实验

本小节通过消融实验分析不同模块对双线性整合脉冲神经网络的影响。对比模型的配置如表3-6所示，其中 LIF-A、LIF-B 和 LIF-C 所代表的模型各包含双线性整合神经元的一部分组件。由于双线性整合脉冲神经元包含两个权重矩阵，这可能是导致网络更快拟合速度和更强拟合能力的原因，因此 LIF-A 模型使用了两个权重矩阵来捕捉线性特征，其中可学习参数  $a$  用于调整两个特征的比例。

LIF-B 模型仅包含二阶特征交互部分，而不包括线性特征，用于评估二阶特征交互的价值。LIF-C 模型则使用固定的比例参数，以验证可学习比例参数的作用。

表 3-6 双线性整合消融实验的模型设置

模型	线性项	二阶特征交互项	公式
LIF-A	✓		$i_t = a \cdot \mathbf{w}_1^\top \mathbf{x}_t + (1 - a) \cdot \mathbf{w}_2^\top \mathbf{x}_t$ , 两个权重矩阵
LIF-B		✓	$i_t = (\mathbf{v}^\top \mathbf{x}_t)^2 - (\mathbf{v}^2)^\top \mathbf{x}_t^2$
LIF-C	✓	✓	$i_t = \mathbf{w}^\top \mathbf{x}_t + \lambda \cdot ((\mathbf{v}^\top \mathbf{x}_t)^2 - (\mathbf{v}^2)^\top \mathbf{x}_t^2)$ , $\lambda$ 为定值

从卷积网络对比实验的结果数据中可以得知，使用 PLIF 构建的脉冲神经网络在处理 CIFAR-10 数据集时，两种整合方式的表现存在较大差异。为了提升消融实验结果的显著性，我们以此数据集和网络设置为标准开展实验，消融实验结果如表3-7所示。

表 3-7 双线性整合消融实验结果

模型	$\lambda$ 设置	准确率 (%)
LIF	-	89.95
双线性 LIF	初始值 0.1	<b>91.08</b>
LIF-A	-	90.55
LIF-B	-	88.86
LIF-C	0.1	91.01
	0.2	90.94
	0.5	88.58

首先，LIF-A 模型的结果表明，使用两个权重矩阵进行线性特征计算确实能比使用单个权重矩阵取得更好的效果。尽管在正向传播中，双权重矩阵的线性整合  $a \cdot \mathbf{w}_1^\top \mathbf{x}_t + (1 - a) \cdot \mathbf{w}_2^\top \mathbf{x}_t$  与单个权重矩阵的  $\mathbf{w}^\top \mathbf{x}_t$  在表达能力上是相同的，但  $\mathbf{w}$  与  $\mathbf{w}_1 + \mathbf{w}_2$  具有不同的初始化参数，并且在反向传播过程中，两个权重矩阵的参数优化可能会更快一些。LIF-B 模型仅使用了双线性整合模式，尽管实验结果不如线性整合表现优异，说明数据中的特征仍以线性特征为主导，但其准确率仍然较高，表明输入和网络之间传播的数据中存在大量的二阶特征交互。LIF-C 模型则设置了不同的比例参数  $\lambda$ ，其中  $\lambda = 0.1$  和  $\lambda = 0.2$  的实验表现接近可学习比例参数的实验结果，但仍存在差距。不适当的参数设置，如  $\lambda = 0.5$ ，会导致性能大幅下降。因此，设置可学习的比例参数对于平衡线性特征与二阶特征交互的权重是更好的选择。

### 3.5 本章小结

本章基于生物神经元的非线性整合机制，分析了其在特定非线性计算任务上的优势，并说明了其计算复杂度过高的局限性。随后，通过借助传统机器学习技术对二阶特征交互的高效处理方式，对双线性整合脉冲神经元进行了创新性的实现，做到了神经元计算能力和计算开销的平衡，还使得神经元具有广泛的兼容性，能够轻松集成到常见的脉冲神经网络结构中。本章还深入探讨了脉冲编码在保留二阶特征交互方面的特性，展示了在脉冲神经元中引入二阶特征交互建模的可行性。为了验证该模型的有效性，本章在多个数据集和多种网络结构上进行了实验。实验结果表明，改进后的神经元模型可以较大提升脉冲神经网络的表现。此外，还通过消融实验进一步分析了不同组件对模型性能的影响。由于此模型内在机制支持二阶特征交互的学习而未显著增加计算负担，因此在保持广泛适应性的同时，还适合于包含大量高阶交互关系以及数据稀疏的应用场景，例如推荐系统。鉴于脉冲神经网络固有的高能效和低时延特性，此改进尤其适合资源受限和高时效性应用，为低功耗、快速响应的智能推荐系统提供了新的可能。

## 第四章 基于分流抑制的脉冲神经元模型

### 4.1 研究动机

本文第三章研究的出发点为生物神经元的非线性整合，这是基于数值拟合的发现。而本章则是从神经元电生理机制的本质层面出发，基于细胞膜的分流抑制机制，提出了一种新型的非线性整合脉冲神经元模型。分流抑制也被称为除法抑制（Divisive Inhibition），因为其对兴奋性电流的影响是按比例进行的。分流抑制的本质在于膜电导的升高导致旁路电流的形成，进而对兴奋电流产生分流作用，从而降低兴奋电流产生去极化的幅度。

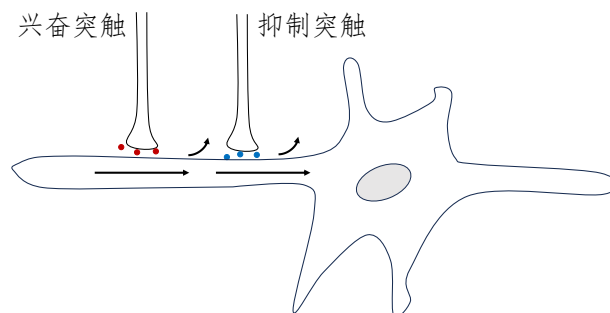


图 4-1 分流抑制示意图。膜电导率改变，兴奋电流被分流。

尽管分流抑制效应具有明确的生物学依据，其计算建模仍面临两个主要挑战：一是离子通道的动态时变特性难以量化建模；二是传统电导参数体系缺乏自适应学习能力。这限制了分流抑制在脉冲神经网络中的应用。因此，本章通过引入人工神经元中常用的门控机制来量化分流比例，并通过可学习参数来实现可调的分流机制。同时，为了平衡正负两种电流的数值，确保网络训练的稳定性，本章在兴奋输入和抑制输入上都引入了门控，并提出了一种根据兴奋和抑制特性对输入进行拆解的方案，最终形成了基于分流抑制的脉冲神经元模型。由于同样是对脉冲神经元整合机制的改进，本章的模型同样可以嵌入到不同的网络结构中，具有广泛的结构兼容性。

## 4.2 生物学依据与关键技术

### 4.2.1 生物神经元的去极化与超极化

从计算主义的角度来看，去极化与超极化是生物神经元处理信息的关键机制。在简化的神经元模型中，对于单个时刻的输入  $x \in \{0, 1\}$ ，神经元的响应电流  $i = wx$ ，这反映了突触权重  $w$  对输入信号整合的重要影响。由于神经元的输入  $x$  只有 0 和 1 两种取值，因此  $w$  的正负决定了电流的正负。当  $w > 0$  时，连接是兴奋性的；当  $w < 0$  时连接是抑制性的。在人工神经元中，无论  $w$  的正负，都没有改变加权线性整合的本质，并且  $w$  的正负性在神经网络训练的过程中有可能会改变，从而改变连接的性质。然而，在生物神经元中，正负权重的连接对应着两种不同的机制。当探讨神经元的去极化与超极化时，需要从神经元的基本结构和机制入手，深入研究这些过程背后的复杂机制。

在没有外部刺激的情况下，神经元能够维持一个稳定的静息膜电位而不发生离子顺电势移动导致膜电位归零的现象，这主要归功于细胞膜的选择透过性。具体来说，细胞膜通过不同的机制控制离子等物质的进出，主要包括主动运输和被动运输两种途径。主动运输允许离子逆浓度梯度或逆电势梯度移动，但这一过程需要消耗能量，通常以 ATP 的形式提供，同时这一过程也需要离子泵等结构的支持，例如钠钾泵。被动运输则是物质顺着浓度梯度或电势梯度自然扩散的过程，不需要额外的能量输入。氧气等小分子可以直接穿透磷脂双分子层进入细胞，而像钠离子、钾离子这样的带电粒子则需要借助特定的通道蛋白才能穿过细胞膜。离子通道不仅协助神经元维持静息膜电位，还在去极化与超极化过程中起到了关键作用，直接影响生物神经元的信息处理模式。因此，深入研究离子通道有助于开发具有更强计算能力的脉冲神经元模型。

当神经元树突接收到足够多的兴奋性神经递质时，这通常会导致局部区域的细胞膜开放部分钠离子通道。由于细胞外的钠离子浓度远高于细胞内，且膜内电势低于膜外，钠离子通道的开放会使钠离子迅速进入细胞，从而使膜电位快速升高，这就是去极化过程的开始。一旦神经元细胞膜的局部区域发生去极化，它会引起一个局部电流，这个电流可以通过扩散传导来影响相邻区域的钠离子通道，从而影响整个膜电位。虽然钠离子是去极化过程中导致膜电位变化的主要离子，但钙离子有时也会参与这一过程。当去极化的膜电位超过某一阈值后，会产

生一个较高的动作电位，大量钠离子通道打开，导致膜电位急剧上升，并沿轴突传导到突触末端。此时钙离子通道也会开放，钙离子进入细胞，触发神经递质的释放，通过神经递质将信号传递给后续神经元。

去极化是生物神经元触发动作电位的核心机制，而超极化则通过降低膜电位来抑制动作电位的发生。超极化主要由氯离子和钾离子通道的活动调控。具体而言，当神经元树突接收到抑制性神经递质时，氯离子通道开放，允许氯离子顺着浓度梯度流入细胞内，导致膜电位进一步下降，从而实现超极化。此外，另一种超极化现象发生在神经元激发之后，钾离子通道打开，大量钾离子流出细胞，使膜电位迅速恢复至静息水平。而此时钾离子通道往往是延迟关闭的，钾离子的外流会继续进行，使膜电位继续降低，短时间难以再次产生动作电位，这就是后超极化（Afterhyperpolarization）现象。

总而言之，生物神经元的去极化与超极化为人工神经元的输入整合机制提供了生物学基础，同时生物神经元膜电位调整机制的新研究也为设计新的人工神经元模型提供了更多可能性。

#### 4.2.2 生物神经元的分流抑制

本文第三章中所述的生物神经元双线性整合是基于数值模拟的结果，并未深入探索导致这种特性的本质。由于生物神经元本身十分复杂，其对输入的整合机制至今也没有被完全研究透彻，实际上，生物神经元的非线性整合现象是由诸多具体机制共同作用导致的，其中分流抑制机制就是一种典型的非线性因素，在大脑皮层细胞的信息处理过程中起到了重要作用。去极化与超极化是神经元输入整合的主要机制，因此目前大多数神经元对输入的整合是线性的。而分流抑制在生物神经元中的输入整合过程中发挥了重要作用，这可以为设计新的神经元模型提供启示。

在超极化和去极化过程中，神经递质引起对应离子通道的开放，这导致离子的流动并会带来电流，从而引起电压的改变。在此过程中，影响膜电位变化幅度的还有一个关键因素——电阻。根据公式 (2-3) 所示的 LIF 模型微分方程表述，由于膜电位向静息电位的自然衰减通常是持续、缓慢的，而去极化、超极化的作用往往是迅速的，因此在短时间  $\delta t$  内，膜电位的自然衰减可以忽略，膜电位的

变化量可描述如下：

$$\frac{R_m}{\tau_m} I(t) \delta t. \quad (4-1)$$

其中  $I(t)$  表示神经元细胞膜的电流， $\tau_m$  为膜时间常数。假设神经元的膜电阻  $R_m$  是固定的，那么神经元电流和膜电位的变化量会呈线性关系。然而，当神经元接收到某些抑制性神经递质（例如  $\gamma$ -氨基丁酸）时，氯离子通道的开放不仅会引发超极化，还会增加细胞膜的总电导率，从而降低膜电阻，产生旁路电流。此时，兴奋性神经递质引起去极化现象所产生的电流被分流，对膜电位的影响会减小，这种现象称为分流抑制。分流抑制机制引入了非线性因素，增强了单个神经元的非线性特征计算能力。

神经元在产生动作电位并释放神经递质后，钠离子通道会进入失活状态，导致神经元进入不应期。在此期间，无论兴奋性输入有多强，膜电位都难以再次达到阈值水平，这意味着兴奋输入信号都被绝对抑制了。这一机制为神经元提供了必要的时间来准备下一次的激发，包括恢复离子浓度梯度，以及准备激发所需物质和能量等。然而，这一机制仅在神经元激发之后才会发生。而在神经元达到阈值之前，分流抑制也起到了调节神经元的兴奋水平的作用，阻止过高频率的动作电位产生，避免过度激活。分流抑制与不应期的绝对抑制两者共同作用，对于维持神经元的功能和整个神经系统的稳定性有重要意义。

### 4.2.3 人工神经元的门控机制

LSTM 是为了解决传统循环神经网络中存在的梯度消失问题及处理长时间依赖关系而设计的一种模型。LSTM 通过引入遗忘门、输入门和输出门等门控机制来控制信息的流动，是较早将门控机制引入人工神经元的尝试之一，其具体实现如下：

$$\begin{aligned} c_t &= f_t c_{t-1} + i_t \cdot \tanh(\mathbf{w}_c^\top [h_{t-1}, \mathbf{x}_t] + b_c), \\ f_t &= \sigma(\mathbf{w}_f^\top [h_{t-1}, \mathbf{x}_t] + b_f), \\ i_t &= \sigma(\mathbf{w}_i^\top [h_{t-1}, \mathbf{x}_t] + b_i), \\ o_t &= \sigma(\mathbf{w}_o^\top [h_{t-1}, \mathbf{x}_t] + b_o), \\ h_t &= o_t \cdot \tanh(c_t). \end{aligned} \quad (4-2)$$

其中  $c_t$  和  $h_t$  分别代表  $t$  时刻神经元的状态和隐状态,  $f_t, i_t, o_t$  分别是遗忘门、输入门和输出门,  $[h_{t-1}, \mathbf{x}_t]$  表示前一时刻隐状态和当前输入的拼接向量,  $\mathbf{w}$  和  $b$  等为不同门对应的权重和偏置,  $\sigma(\cdot)$  一般使用 Sigmoid 函数,  $o_t$  为神经元输出。LSTM 模型实现较为复杂, 计算开销也较高。为了简化这一过程, 实现更为简单的 GRU 应运而生。GRU 仅使用了一个状态, 并将遗忘门和输入门合并为更新门, 其数学表示如下:

$$\begin{aligned} h_t &= (1 - z_t)h_t + z_t \cdot \tanh(\mathbf{w}_h^\top [r_t h_{t-1}, \mathbf{x}_t] + b_h), \\ z_t &= \sigma(\mathbf{w}_z^\top [h_{t-1}, \mathbf{x}_t] + b_z), \\ r_t &= \sigma(\mathbf{w}_r^\top [h_{t-1}, \mathbf{x}_t] + b_r). \end{aligned} \quad (4-3)$$

其中  $h_t$  为  $t$  时刻神经元状态,  $z_t$  和  $r_t$  分别为更新门和重置门。GRU 可以用更简单的结构和更小的计算开销达到和 LSTM 类似的效果。

在神经网络设计中, 门控机制不仅限于特定结构的神经元, 它也被应用于激活函数的设计。GLU 最早在语言建模任务中被提出, 其在梯度计算时会留下线性通道<sup>[48]</sup>, 能够抑制梯度消失, 加速网络收敛, 因此在大语言模型中应用较多<sup>[73-74]</sup>。GLU 的数学表示如下:

$$\text{GLU}(\mathbf{x}) = (\mathbf{w}^\top \mathbf{x} + b) \cdot \sigma(\mathbf{v}^\top \mathbf{x} + c) \quad (4-4)$$

其中  $\mathbf{w}$  和  $\mathbf{v}$  表示权重,  $b$  和  $c$  表示偏置,  $\sigma(\cdot)$  为 Sigmoid 函数。相比于使用了两个权重的 GLU, Swish 激活函数则使用了自门控机制:

$$\text{Swish}(z) = z \cdot \sigma(z). \quad (4-5)$$

其中  $z = \mathbf{w}^\top \mathbf{x} + b$ 。Swish 函数连续、处处可微, 并且上无界、下有界, 这些良好特性使其相较于传统的激活函数有更好的表现<sup>[49]</sup>。

尽管上述基于门控机制的设计并未直接涉及生物神经元的具体工作原理, 如离子通道或分流抑制等复杂机制, 但这些工作为优化人工神经元的表现提供了基础, 也为如何利用门控机制改进脉冲神经元模型带来了启示。

### 4.3 分流抑制脉冲神经元模型实现与分析

#### 4.3.1 脉冲神经元正负输入拆解

在生物神经元中，分流抑制可以使去极化电流产生分流，从而减少膜电位上升的幅度。为了在人工神经元中模拟这一机制，首先需要首先将兴奋性输入和抑制性输入分离。具体而言，在人工神经元模型中，如果有输入  $x \in \{0, 1\}$ ，那么权重  $w$  的正负决定了连接的性质， $w > 0$  表示兴奋性连接， $w < 0$  表示抑制性连接。对于给定输入向量  $\mathbf{x} \in \{0, 1\}^N$ ，其引起的电流可以按以下方式进行分解：

$$\begin{aligned} i &= \mathbf{w}^\top \mathbf{x} = i_+ + i_-, \\ i_+ &= \sum_{w_i > 0} w_i x_i, \\ i_- &= \sum_{w_i < 0} w_i x_i. \end{aligned} \quad (4-6)$$

其中， $i_+$  为所有兴奋性输入的电流，对应去极化程度， $i_-$  为所有抑制性输入的电流，对应超极化程度。

基于现有的深度学习框架（如 PyTorch），为了在人工神经元中依据以上方式对输入进行拆解，可以利用 ReLU 函数来实现。ReLU 函数可以表示如下：

$$\text{ReLU}(z) = \max(0, z). \quad (4-7)$$

ReLU 函数实现简单，并且梯度性质良好，可以有效缓解梯度消失和梯度爆炸问题，因此 ReLU 及其变体经常用作神经网络的激活函数。但对于正负拆解的场景而言，直接对电流应用 ReLU 并不能满足需求，因为神经元输入一旦被整合成电流  $i$ ，就无法再进行拆分了。因此，我们需要在整合电流之前，将 ReLU 作用在连接权重上：

$$\begin{aligned} i_+ &= \mathbf{w}_+^\top \mathbf{x}, \\ i_- &= \mathbf{w}_-^\top \mathbf{x}, \\ \mathbf{w}_+ &= \text{ReLU}(\mathbf{w}), \\ \mathbf{w}_- &= -\text{ReLU}(-\text{ReLU}(\mathbf{w})). \end{aligned} \quad (4-8)$$

这样一来， $\mathbf{w}_+$  和  $\mathbf{w}_-$  分别提取了原有权重  $\mathbf{w}$  里所有的正值和负值部分。由于输入是 0/1 的，因此以上拆解方式保证了电流  $i_+$  总是兴奋电流， $i_-$  总是抑制电流。为了表达式的简洁性，本文后续所有  $(\cdot)_+$  和  $(\cdot)_-$  形式的变量均表示按此方法进行拆解的结果。

### 4.3.2 神经元实现

分流抑制在生物神经元中可以通过分流来削弱兴奋性输入对膜电位的影响，然而，如果直接将此机制引入人工神经元中会带来一些问题。例如，在处理  $t$  时间步的脉冲输入  $\mathbf{x}_t$  时，正负电流可以表示如下：

$$\begin{aligned} i_t^+ &= \mathbf{w}_+^\top \mathbf{x}_t \cdot \sigma(\mathbf{v}^\top \mathbf{x}_t), \\ i_t^- &= \mathbf{w}_-^\top \mathbf{x}_t. \end{aligned} \quad (4-9)$$

其中  $\sigma(\cdot)$  表示 Sigmoid 函数。目前，线性层、卷积层的权重矩阵初始化通常采用的是 Kaiming 均匀分布 (Kaiming Uniform Distribution)<sup>[75]</sup>，权重的期望均值为 0。这使得正负权重的绝对值分布相同，从而保证线性整合值的稳定性。但如果按照上述公式仅对正电流施加门控，会使得  $i_t^+$  与  $i_t^-$  的数值分布存在较大偏差，从而导致极端值的出现，使网络训练变得不稳定。

分流抑制现象源于膜电导率的改变，对生物神经元而言，这一改变对抑制性输入引起的膜电位变化幅度同样有影响。因此，无论是从生物学合理性出发，还是就神经网络的训练稳定性而言，为负的电导也引入门控机制是合理的做法。尽管给正负电流施加相同的分流程度（即使用相同的门控）在技术实现上是可行的，但考虑到膜电阻是一个描述细胞膜对电流总体阻碍程度的参数，它并不意味着细胞膜对所有正负离子的运动具有相同的阻碍程度。因此，更为合理的做法是对神经元中的正负电流施加两个不同的分流抑制比例。对于脉冲输入  $\mathbf{x}_t$ ，基于分流抑制调控的脉冲神经元膜电位响应可表述如下：

$$u_t = \alpha(u_{t-1} - b) + \mathbf{w}_+^\top \mathbf{x}_t \cdot \sigma(\mathbf{v}_p^\top \mathbf{x}_t) + \mathbf{w}_-^\top \mathbf{x}_t \cdot \sigma(\mathbf{v}_n^\top \mathbf{x}_t) + b. \quad (4-10)$$

其中， $\mathbf{v}_p$  和  $\mathbf{v}_n$  分别表示正负电流对应门控的权重。通过这种方式，脉冲神经元可以更精确地模拟生物神经元的膜电位响应机制，同时能够提高脉冲神经网络

的表达能力和训练稳定性。图4-2展示了分流抑制神经元的实现方式。

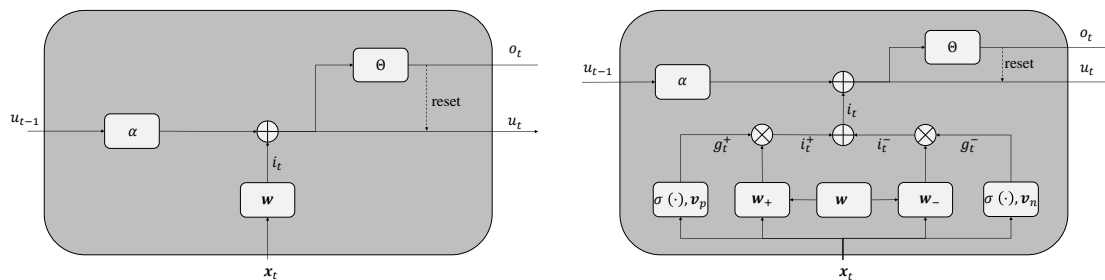


图 4-2 传统脉冲神经元与分流抑制脉冲神经元模型对比

### 4.3.3 计算优势分析

$\mathbf{x}$	$i$	阈值	输出
(0,0)	0	0.25	0
(0,1)	0.269	0.25	1
(1,0)	0.269	0.25	1
(1,1)	0.238	0.25	0

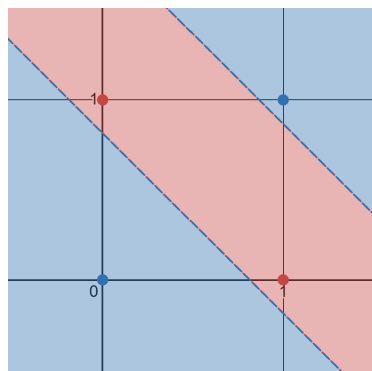


图 4-3 分流抑制神经元的异或计算。右图为神经元划分边界的可视化。

在脉冲神经元中，通过对电流施加门控实现对输入的分流抑制机制，可以增强膜电位的非线性表达能力。以经典的异或问题为例，本章提出的基于分流抑制的脉冲神经元同样只需要一个神经元就可以完成异或计算任务。假设有  $t$  时间步的输入  $\mathbf{x}_t$ ，其引发的电流可以表示为：

$$i_t = \mathbf{w}_+^T \mathbf{x}_t \cdot \sigma(\mathbf{v}_p^T \mathbf{x}_t) + \mathbf{w}_-^T \mathbf{x}_t \cdot \sigma(\mathbf{v}_n^T \mathbf{x}_t). \quad (4-11)$$

对于二维输入  $\mathbf{x}_t \in \{0, 1\}^2$  的情况，通过选择适当的参数值，例如令  $\mathbf{w} = (1, 1)$ ， $\mathbf{v}_p = \mathbf{v}_n = (-1, -1)$ ，阈值  $u_{\text{thr}} = 0.25$ ，即可有效地解决异或问题。

单个神经元的非线性表达能力提升可以增强整个网络的表达能力。此外，门控机制通过引入三个独立的可学习参数  $\mathbf{w}, \mathbf{v}_p, \mathbf{v}_n$ ，实现了对神经元活动更精准的调控，增加了模型的灵活性，并促进了脉冲神经网络更快的收敛速度。

从生物学角度而言，分流抑制可以防止神经元的过度激活，增强生物神经系统的稳定性。对应到脉冲神经网络中，这种门控机制限定了输入对膜电位的影响程度，减少了极端值的出现。当采用替代梯度下降法对脉冲神经网络进行训练时，由于输入整合值的稳定性得到了保证，膜电位偏离阈值的程度得以减小，这进一步缓解了梯度消失的问题。

#### 4.3.4 全连接层与卷积层实现

本章所提出的基于分流抑制对脉冲神经元整合机制的改进，可以更充分的利用生物神经元计算特性，从而增强脉冲神经元对非线性特征的学习能力。和第三章所提模型类似，由于同样是仅限于空间整合机制的改进，因此本章所提的改进也能够无缝嵌入到多种脉冲神经网络结构中，而不受限于网络具体拓扑结构，与多种模型具有广泛的兼容性。

将对神经元的改进应用到脉冲神经网络的全连接结构中，对于  $t$  时间步的膜电位  $\mathbf{u}_t \in \mathbb{R}^M$  对输入  $\mathbf{x}_t \in \{0, 1\}^N$  的响应可由以下表达式描述：

$$\mathbf{u}_t = \alpha(\mathbf{u}_{t-1} - \mathbf{b}) \odot (\mathbf{1} - \mathbf{o}_{t-1}) + \mathbf{W}_+ \mathbf{x}_t \odot \sigma(\mathbf{V}_p \mathbf{x}_t) + \mathbf{W}_- \mathbf{x}_t \odot \sigma(\mathbf{V}_n \mathbf{x}_t) + \mathbf{b}. \quad (4-12)$$

其中， $\mathbf{W} \in \mathbb{R}^{M \times N}$  代表权重矩阵， $\mathbf{V}_p, \mathbf{V}_n \in \mathbb{R}^{M \times N}$  分别代表正输入和负输入门控的权重矩阵， $\mathbf{b} \in \mathbb{R}^M$  为偏置项， $\mathbf{o}_{t-1} \in \{0, 1\}^M$  为上一个时间步的神经元输出情况，决定是否将历史膜电位重置。

将对神经元的改进应用到脉冲神经网络的卷积结构中，对于  $t$  时间步的输入数据  $\mathbf{X}_t \in \{0, 1\}^{C \times H \times W}$ ，神经元膜电位为  $\mathbf{U}_t \in \mathbb{R}^{C' \times H \times W}$ ，其中  $C'$  为输出的通道数，即输出特征图的数量，膜电位对输入的响应方程可以表示如下：

$$\begin{aligned} \mathbf{U}_t = & \alpha(\mathbf{U}_{t-1} - \mathbf{b}) \odot (\mathbf{1} - \mathbf{O}_{t-1}) + (\mathbf{W}_+ * \mathbf{X}_t) \odot \sigma(\mathbf{V}_p * \mathbf{X}_t) \\ & + (\mathbf{W}_- * \mathbf{X}_t) \odot \sigma(\mathbf{V}_n * \mathbf{X}_t) + \mathbf{b}. \end{aligned} \quad (4-13)$$

其中， $\mathbf{W} \in \mathbb{R}^{C' \times C \times S \times S}$  表示卷积核， $\mathbf{V}_p, \mathbf{V}_n \in \mathbb{R}^{C' \times C \times S \times S}$  分别代表正负输入对应门控的卷积核， $\mathbf{b} \in \mathbb{R}^{C'}$  为 channel-wise 的偏置项。 $\mathbf{O}_{t-1} \in \{0, 1\}^{C' \times H \times W}$  为上一个时间步此通道的神经元的激发值。

值得注意的是，PyTorch 默认采用 Kaiming 均匀分布来对矩阵进行初始化，这种初始化方法适用于 ReLU 及其变种激活函数的网络，但在某些特殊结构中，例如将其应用在门控内的权重矩阵的初始化时，默认的初始化策略可能表现不佳。如果门控值的初始分布过于极端（即接近 0 或 1），可能会导致某些门控几乎完全开启或关闭，这会使得网络前期学习困难。为了应对这个问题，可以在门控内的权重矩阵上乘以一个缩放因子  $k$ ，使门控在网络训练初期更加倾向于中间值，有助于网络对更均等地考虑所有信息的价值，这种方法本质上与改变初始分布的方差是等价的，实验部分缩放比例指的就是缩放因子  $k$  的取值。

## 4.4 实验与分析

本节通过对比实验来评估线性整合方式与基于分流抑制的非线性整合方式在脉冲神经网络中的表现差异。为进一步验证所提出方法的有效性，后续开展了消融实验，探讨不同组件和结构对整体性能的影响。

### 4.4.1 实验设置

表 4-1 实验网络参数设置。有括号标注的设置仅在指定数据集或模型上生效。

参数名称	设置
优化器	Adam
初始学习率	0.001, 0.0001 (IMDB)
训练轮数	100, 20 (IMDB), 200 (CIFAR-100)
批大小	100, 128 (MPBN), 16/4 (PLIF, CIFAR-10/CIFAR10-DVS)
阈值	0.5, 1.0 (PLIF)
衰减系数 $\alpha$	0.25, 自适应 (PLIF)
偏置	有, 无 (PLIF 和 MPBN)

为了验证模型在不同数据上的学习能力增强效果，本章选择了多样化的数据集进行测试，包括 IMDB、GTZAN、MNIST、N-MNIST、Fashion-MNIST、CIFAR-10、CIFAR10-DVS 以及 CIFAR-100 等。各数据集的选取和预处理方式均与第三章实验保持一致。为了体现本章提出的整合方式对脉冲神经网络不同结构的适应性，实验采用了基于 STBP<sup>[29]</sup>、PLIF<sup>[71]</sup> 和 MPBN<sup>[32]</sup> 的多种网络架构。

实验的网络参数设置详见表4-1。由于本章所提出的改进同样基于非线性整合机制设计新的脉冲神经元，因此实验设置基本延续了第三章的配置。与第三章

不同的是，在使用分流抑制的 PLIF 网络学习 CIFAR10-DVS 数据时，出于显存限制考虑，批大小被设置为较小的值。

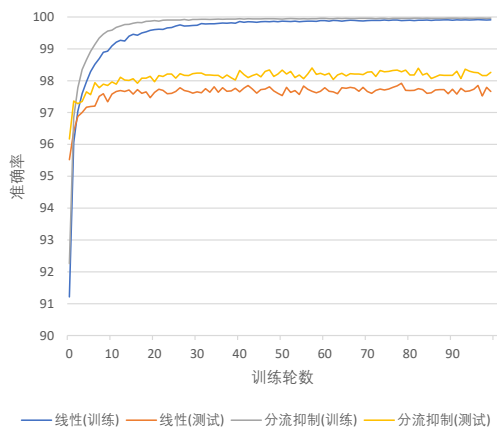
#### 4.4.2 全连接网络对比实验结果与分析

表 4-2 分流抑制在全连接网络中的实验结果

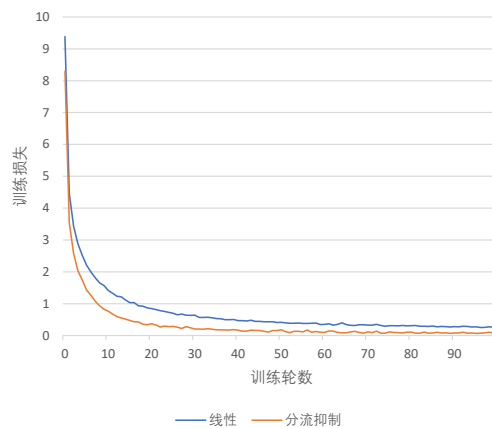
数据集	时间窗口	网络结构	整合方式	缩放比例	准确率 (%)
IMDB	20	5000-100-2	线性	-	85.91
			分流抑制	0.01	<b>86.31</b>
GTZAN	20	169-50-10	线性	-	65.00
			分流抑制	0.01	<b>68.00</b>
MNIST	20	784-100-10	线性	-	97.96
			分流抑制	0.1	<b>98.20</b>
		784-200-10	线性	-	98.42
		分流抑制	0.1	<b>98.56</b>	
		784-400-10	线性	-	98.62
			分流抑制	0.1	<b>98.66</b>

实验首先在最基础的全连接脉冲神经网络上进行，以展示模型在简单任务上的学习能力。表4-2展示了分流抑制机制在全连接脉冲神经网络中的实验结果。从数据可以看出，分流抑制在大多数简单数据集（如 IMDB、GTZAN 和 MNIST）上均表现出一定的性能优势，但提升幅度在不同数据集上存在一定差异。例如，对于 IMDB 和 GTZAN 这类向量形式的数据，分流抑制的提升幅度分别为 0.4% 和 3.0%，而对于 MNIST 的提升幅度相对较小，为 0.04% 到 0.24%，这种差异可能源于数据集本身的特征分布特性。IMDB 和 GTZAN 的数据通常包含较多的非线性特征交互关系，而分流抑制机制通过引入门控对输入进行动态调节，能够更有效地捕捉这些复杂关系，从而在表现上的优势更加明显。相比之下，MNIST 数据的线性可分性较强，分流抑制的非线性优势在神经元数量较多时逐渐被稀释，导致提升效果减弱。

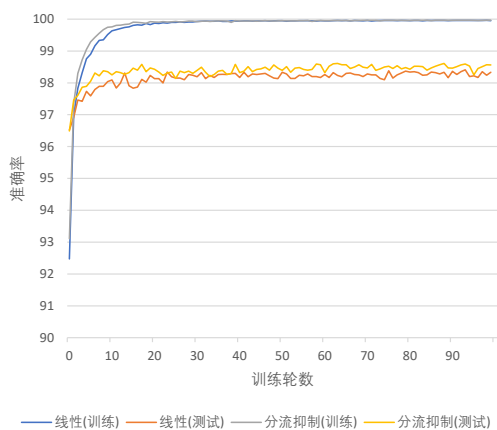
此外，MNIST 数据集上的实验结果显示，神经元的数量也对分流抑制机制带来的性能提升幅度有一定影响，即神经元数量越少提升越明显。这表明采用传统线性整合机制的神经元需要借助多层网络的非线性表达能力才能够学习到 MNIST 数据集中的非线性特征，而这需要更多的神经元数量支持。分流抑制的改进效果在资源受限的场景下更为显著。



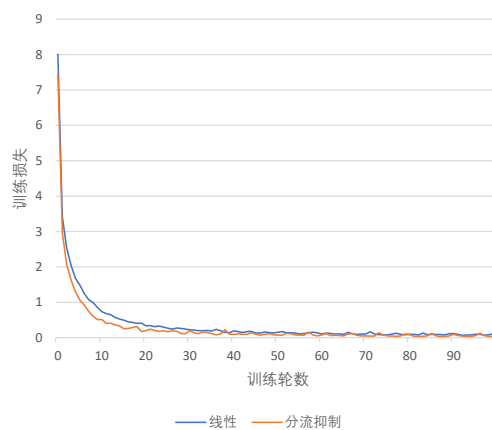
准确率曲线, 100 个隐层神经元



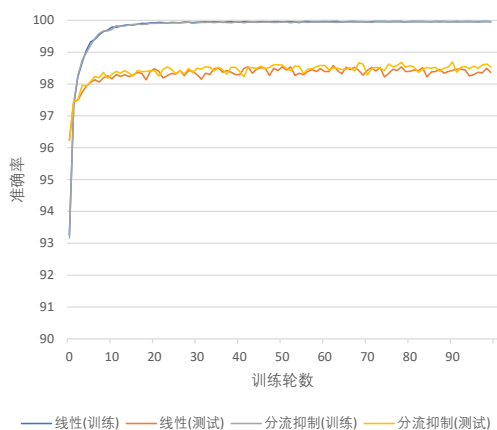
损失曲线, 100 个隐层神经元



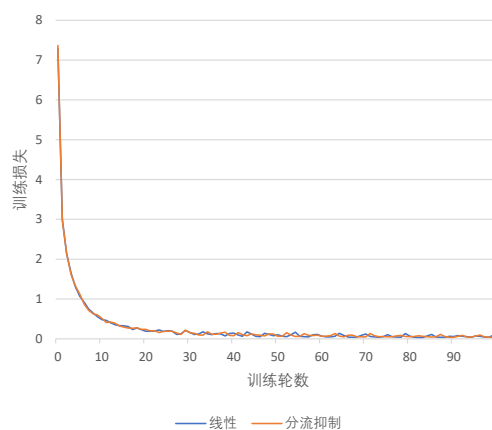
准确率曲线, 200 个隐层神经元



损失曲线, 200 个隐层神经元



准确率曲线, 400 个隐层神经元



损失曲线, 400 个隐层神经元

图 4-4 MNIST 数据集上全连接脉冲神经网络使用线性与分流抑制整合的训练曲线对比

为了更直观地展示分流抑制机制对模型学习能力的提升，我们记录了模型在训练过程中的准确率和损失变化的情况。如图4-4所示，采用分流抑制的脉冲神经网络整体上具有更高的训练集准确率和更低的损失，体现了其对训练数据更强的拟合能力。例如，使用 100 个隐层神经元时，分流抑制模型的训练集准确率始终高于线性整合模型，且损失值下降速度更快。这一现象可能与分流抑制对膜电位动态的调控有关，传统的线性整合方式可能导致膜电位分布过于集中或发散，而分流抑制通过门控机制动态调整输入对膜电位的影响，使得膜电位更接近阈值区域，从而优化了梯度传播的稳定性。这表明分流抑制能够更快地适应数据分布，减少训练初期的探索成本。随着隐层神经元数量的增加，两种整合机制的差距逐渐缩小，训练集最终准确率均接近 100%，说明在此任务上的精度已接近饱和，需要在更复杂的任务上进行实验。

#### 4.4.3 卷积网络对比实验结果与分析

表 4-3 分流抑制在卷积网络中的实验结果。表中的卷积网络结构省略了输入部分。

数据集	时间窗口	网络结构	整合方式	缩放比例	准确率 (%)
MNIST	20	32C3-P2-32C3-P2-128-10	线性	-	99.41
			分流抑制	0.1	<b>99.46</b>
N-MNIST	10	32C3-P2-32C3-P2-128-10	线性	-	99.42
			分流抑制	0.1	<b>99.47</b>
Fashion-MNIST	20	32C3-P2-32C3-P2-128-10	线性	-	91.12
			分流抑制	0.01	<b>91.26</b>
CIFAR10	8	PLIF-Net1	线性	-	89.95
			分流抑制	0.01	<b>90.65</b>
CIFAR10-DVS	10	PLIF-Net2	线性	-	<b>67.67</b>
			分流抑制	0.01	67.50
CIFAR100	4	MPBN-ResNet20	线性	-	66.01
			分流抑制	0.01	<b>66.31</b>

为了测试分流抑制机制在更复杂脉冲神经网络结构上的适应性及其对复杂数据集学习能力的提升效果，后续实验采用卷积网络在 MNIST、Fashion-MNIST、CIFAR-10 和 CIFAR-100 等图像数据集上进行，同时也在神经形态数据集 N-MNIST 和 CIFAR10-DVS 上也进行了实验。所采用的网络结构包括基于 STBP 构建的浅层卷积脉冲神经网络以及基于 PLIF 或 MPBN 构建的深度脉冲神经网络。

络，实验结果详见表4-3。数据表明，分流抑制机制在脉冲神经网络中整体上能够提升网络性能。

对于基于 STBP 构建的浅层卷积脉冲神经网络，在 MNIST、N-MNIST 和 Fashion-MNIST 数据集上的实验结果表明，分流抑制能够提升网络对图像数据的学习能力。尽管提升幅度相对有限，但该机制至少能够保持线性整合的性能，不会对一般图像数据上的任务产生负面影响，从而初步验证了分流抑制在卷积脉冲神经网络中的有效性。

接下来，实验在更复杂、区分度更高的 CIFAR-10 和 CIFAR-100 数据集上进行。为验证本章改进对深层网络性能的提升效果，分别采用 PLIF 和 MPBN 构建了不同的深度脉冲神经网络。实验结果表明，分流抑制机制能够为这两种网络结构在不同数据集上带来不同程度的性能提升。在 CIFAR-10 数据集中，分流抑制在 MPBN 网络上的提升幅度小于在 PLIF 网络上的提升，这可能是由于分流抑制减少了过高或过低的膜电位，使膜电位趋于集中于阈值附近，从而缓解了替代梯度计算中梯度消失的问题；而 MPBN 本身也具有类似的稳定作用，因此两者之间的性能差异较小。

对于 CIFAR10-DVS 数据集，实验结果显示分流抑制的表现反而比线性整合低 0.17%，这与第三章中双线性整合面临的问题类似。综合其他数据集和网络的实验结果，我们认为这是由于 CIFAR10-DVS 数据集的特性所致，其更适合采用线性整合模式的脉冲神经网络。但在多数应用场景中，分流抑制机制可为脉冲神经网络带来普遍的性能提升。

#### 4.4.4 消融实验

本小节通过系统性消融研究评估分流抑制脉冲神经元中各模块的贡献，重点关注以下三个核心问题：

1. 分流抑制与超极化抑制机制的功能优先级比较，验证单独使用分流抑制作为抑制机制是否可行。
2. 给负电流也添加门控机制对网络学习能力是否有提升。
3. 正负电流分别设置门控相较于对整体添加门控的计算能力是否有区别。

为此，实验构建了三个消融模型 (LIF-A/B/C)，具体实现方式如表4-4所示。

表 4-4 分流抑制消融实验的模型设置

模型	超极化	负电流分流	整合方式
LIF-A			$i_t = \mathbf{w}_+^\top \mathbf{x}_t \cdot \sigma(\mathbf{v}_p^\top \mathbf{x}_t)$
LIF-B	✓		$i_t = \mathbf{w}_+^\top \mathbf{x}_t \cdot \sigma(\mathbf{v}_p^\top \mathbf{x}_t) + \mathbf{w}_-^\top \mathbf{x}_t$
LIF-C	✓	✓	$i_t = \mathbf{w}^\top \mathbf{x}_t \cdot \sigma(\mathbf{v}^\top \mathbf{x}_t)$

LIF-A 模型仅保留对正电流的分流作为抑制机制，不包含超极化抑制，以验证分流抑制单独作用的效果。LIF-B 模型仅实施对正电流的门控，不改变负电流。LIF-C 模型对正负电流采用统一的门控机制，即对电流整体施加门控。鉴于初步实验结果表明不适宜的缩放比例会导致权重初始化不良，进而引发训练过程中准确率和损失的停滞现象，因此未将不同的缩放比例设置纳入消融实验范畴。

从 4.4.3 小节卷积网络对比实验中的数据可知，使用 PLIF 构建的脉冲神经网络在处理 CIFAR-10 数据集时，分流抑制机制带来的性能提升较为明显。为了增强消融实验结果的显著性，我们选择在 CIFAR-10 数据集上使用基于 PLIF 构建的脉冲神经网络进行消融实验。该实验的网络结构和参数设置与对比实验保持一致，消融实验结果见表4-5。

表 4-5 分流抑制消融实验结果

模型	放缩比例	准确率 (%)
线性	-	89.95
分流抑制	0.01	<b>90.65</b>
LIF-A	1.0	43.75
LIF-B	0.01	90.31
LIF-C	0.01	89.89

消融实验结果表明，LIF-A 模型仅使用分流抑制机制而忽略超极化抑制会导致网络性能显著下降，超极化抑制作为神经元的核心机制，能够提供稳定的基本整合能力，而 LIF-A 模型缺乏这一基础，导致模型难以捕捉输入数据中的基本特征，会严重削弱模型的表达能力。LIF-B 模型表现略低于对正负电流均施加门控的模型，但高于纯线性整合模型，表明只对正电流分流可以增强神经元表达能力，但是会面临 4.3.2 中所述的数据分布问题。LIF-C 模型与线性整合模型的差异不大，可能原因在于 PLIF 本身对电流与膜电位的调控的使得单一的门控作用被部分抵消，而对正负电流采用不同门控的方法受到的影响相对较小。

此外，图4-5中记录了消融实验的训练曲线，由于 LIF-A 的表现明显落后于其他模型，因此未参与比较。训练曲线显示，对正负电流均应用分流抑制机制的模型（橙色曲线）在训练集准确率和损失上相对于其他模型具有较为明显的优势，体现了模型具有更强的数据拟合能力。而且根据损失曲线，分流抑制模型在 150k 步附近就达到了收敛，相较于其他模型具有更快的收敛速度，这一现象可能源于分流抑制通过动态调节膜电位分布，优化了梯度传播的稳定性，从而加速了模型收敛。总体而言，训练曲线的差异验证了分流抑制机制在动态调节膜电位分布中的关键作用，同时凸显了正负电流独立门控设计的必要性。

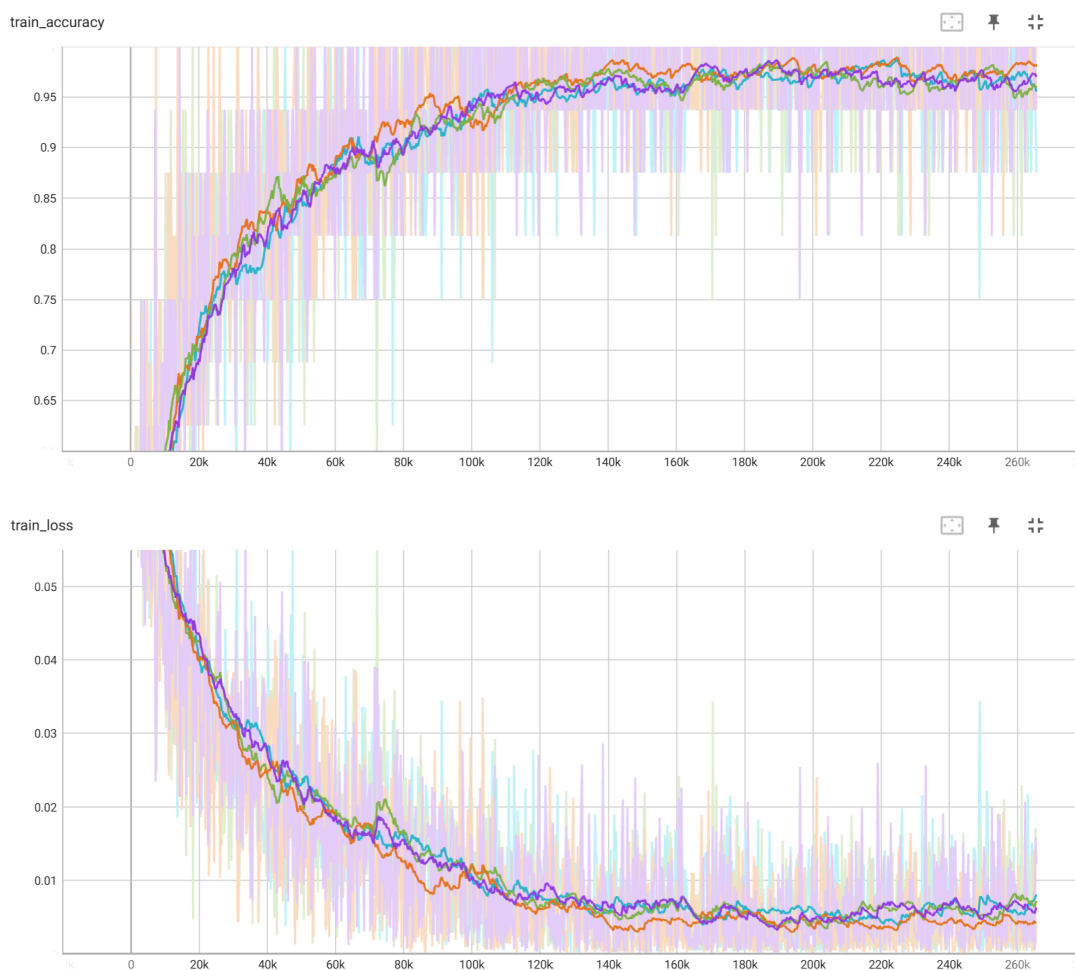


图 4-5 分流抑制脉冲神经网络消融实验训练曲线。蓝色为线性整合，橙色为分流抑制整合，绿色为 LIF-B，紫色为 LIF-C，横坐标为训练的步数，TensorBoard 展示的曲线经过了平滑来达到更好的展示效果，浅色曲线为未经平滑的原始数据。

## 4.5 本章小结

本章基于生物神经元的分流抑制机制的研究，从神经元去极化、超极化等生理过程中的离子通过透过性和膜电导率变化等基础机制出发，对传统的 LIF 神经元模型的线性整合方式进行了改进，提出了一种新型脉冲神经元模型。该模型通过引入基于分流抑制的非线性整合机制，增强了单个脉冲神经元的非线性表达能力，并能便捷地集成到常见的脉冲神经网络中，从而优化整体网络性能。为了验证该模型的有效性，在多个数据集和多种网络结构上进行了系统实验。实验结果表明，改进后的神经元模型能够普遍提升脉冲神经网络在各类任务上的表现。同时，通过消融实验，本章进一步分析了模型中不同组件对性能的影响，结果显示对神经元正负电流分别施加单独门控具有重要意义。尽管改进后的神经元模型整体上的表现上略低于第三章提出的基于双线性整合的脉冲神经元模型，但本章的工作侧重于从生物神经元电生理机制的角度探索神经元特性，为后续脉冲神经元模型设计相关研究提供了更为丰富的生物学依据和启示。



## 第五章 应用：基于脉冲神经网络的推荐系统

为了验证本文第三章提出的双线性整合脉冲神经元在实际应用场景中的有效性，本章搭建了一个基于脉冲神经网络的推荐系统。本章将从系统研发背景、设计实现以及结果展示等方面对该系统进行详细介绍。

### 5.1 研发背景

随着全球经济的持续增长及社会生产力和文化产业的不断革新，人们的物质和精神生活日益丰富，实体商品和文化产品的选择呈现出前所未有的多样性。同时，生活水平的提升使用户需求变得更加个性化和差异化，导致商品供给与需求之间的关系愈发复杂。在这一背景下，信息技术的迅猛发展推动了数据量的爆炸式增长，使用户在海量信息中迅速准确地找到所需内容成为一大难题。这不仅增加了用户获取信息的时间成本，也对商家精准信息投放和营销策略提出了更高要求。传统的信息推送方式在处理海量、多元用户需求时，往往难以实现精准匹配，从而导致信息投放的效率和有效性大幅降低。例如，传统电视广告由于传播方式较为单一和粗放，其转化率普遍较低，难以满足现代商业环境对精准营销的需求。在此背景下，智能推荐系统应运而生，成为缓解用户选择困难、提升商家转化率的重要技术手段。借助机器学习、数据挖掘等先进技术，推荐系统能够自动筛选并推送个性化信息，不仅显著降低了用户在海量信息中寻找所需内容的时间成本，而且提升了用户体验和满意度，从而增强了用户黏性并推动了平台商业价值的实现。因此，智能推荐系统在现代信息化社会中具有广泛的应用前景和重要的研究价值。

面对实际的应用场景，推荐系统不仅需要关注推荐结果的准确性，还必须保证计算的实时性，以使用户能够及时获取有效信息。因此，设计推荐系统时必须在提升推荐效果与优化计算效率之间取得恰当平衡。近年来，脉冲神经网络作为一种低功耗、低延迟的计算模型逐渐受到学界和工业界的关注。尤其是在应对

稀疏数据场景时，其事件驱动的特性带来了明显的低功耗推理优势，为构建高效推荐系统提供了新的解决方案。

## 5.2 系统设计

本章结合影视推荐场景，搭建了一个基于脉冲神经网络的电影推荐系统，并将本文第三章提出的双线性整合脉冲神经元应用其中。系统利用脉冲神经网络模型计算候选电影与用户之间的推荐值，并以此对电影进行排序。为验证系统表现，本章使用公开的 MovieLens 电影评分数据集<sup>[76]</sup>进行测试与效果展示。

### 5.2.1 需求分析

本系统旨在通过综合分析用户个人特征、电影属性、评分统计数据以及用户与电影之间的历史交互记录，为用户提供精准且个性化的电影推荐服务，从而帮助用户发现潜在感兴趣的新电影。基于上述目标，本系统需要满足以下需求：

1. **用户信息管理**。为了在海量电影资源中迅速筛选出候选影片，以及后续根据用户具体情况对电影进行推荐排序，系统需收集并管理用户基本信息。
2. **电影资料管理**。系统应构建一个涵盖电影上映时间、类型、大众评分等信息的电影信息库，并定期更新数据库，确保数据的准确性和时效性。
3. **交互行为收集与分析**。为深入了解用户偏好，并不断优化推荐算法，系统必须具备完善的用户行为追踪机制，能够及时记录和更新用户对电影的评分及其他交互数据。
4. **精准的个性化图书推荐**。在基于召回算法确定推荐范围的基础上，还需要采用细致的排序算法对候选电影进行推荐值的计算与排序，以提升推荐结果的精准性。
5. **推荐系统冷启动**。对于新注册用户或新上线电影，系统应设计有效的冷启动策略，以应对初始数据稀缺问题。
6. **隐私管理**。为保护用户个人信息，系统需允许用户选择是否上传个人信息；对于未上传信息的用户，则应制定空缺值处理策略，确保在信息缺失条件下仍能提供合理的推荐结果。

## 5.2.2 系统架构

根据以上需求分析的结果，我们设计了如图5-1所示的系统架构。本系统主要分为以下三个模块：（1）数据模块：负责存储必要的的数据资源，包括用户数据、电影数据以及交互数据等。为支持实时推荐功能，除表格型数据外，数据模块还需存储预训练模型相关数据，并确保这些模型能够快速加载以供推理过程使用。（2）后端模块：负责数据管理及推荐系统核心算法逻辑的实现。该模块不仅需响应前端请求，对数据进行更新和读取，还需要依据数据完成特征提取、候选项召回、推荐值计算及排序等功能，以生成最终的推荐结果并返回给前端。（3）前端模块：负责用户交互，提供直观、友好的界面供用户浏览推荐结果并进行互动。同时，前端模块应具备向后端发送查询请求以及接收、可视化展示返回推荐结果的能力。以上三个模块通过定义明确的 API 进行通信，实现模块间的解耦，简化了系统维护工作，并为未来功能扩展提供了良好的支撑。

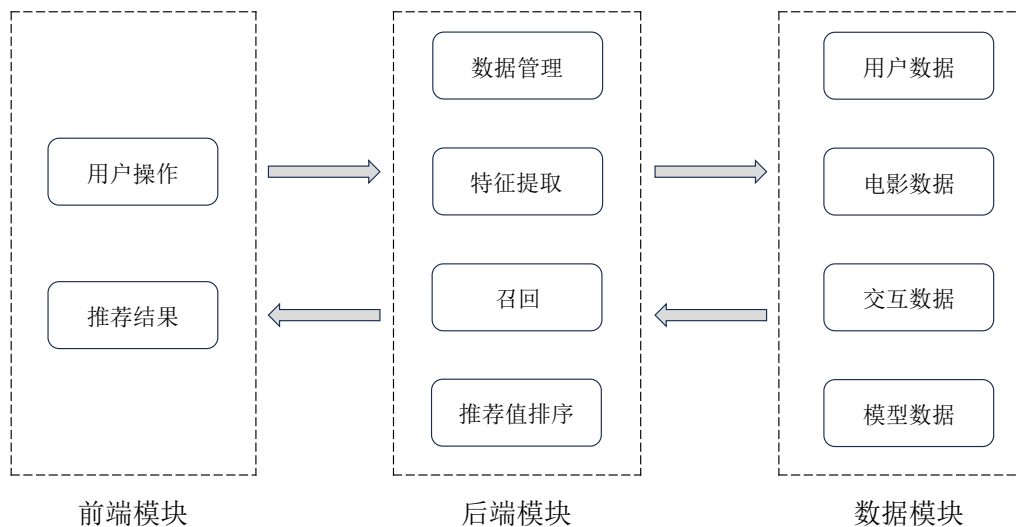


图 5-1 系统架构

## 5.3 系统实现

基于上述系统架构设计，本文对基于脉冲神经网络的推荐系统进行了具体实现。本节将从框架选取、技术选型以及算法方案等角度，介绍数据模块、后端模块和前端模块的具体实现细节。

### 5.3.1 数据模块

本系统涉及多种数据类型，不同数据会有各自的采集方式、结构及存储需求，具体描述如下：

1. **用户数据**。包括用户编号、性别、年龄、职业和邮编，其中用户编号为唯一标识符，确保数据的唯一性。用户数据在用户注册或更新个人信息时采集和更新。
2. **电影数据**。包括电影编号、名称、上映时间和类型，其中电影编号为唯一标识符。电影数据既可以由系统运营方录入，也支持用户上传，上传后经审核后可纳入数据库。
3. **交互数据**。记录用户与电影之间的行为，包括用户编号、电影编号、评分及时间戳，其中用户编号与电影编号共同构成单次评分行为的标识。对于同一用户对同一电影的评分，后续更新将覆盖先前记录。
4. **模型数据**。包括召回与排序阶段使用的预训练模型。

对于表格型数据的存储，系统使用开源的关系型数据库 MySQL 进行管理，使用 SQL 语言完成表和记录的增删查改。所有表均使用自增主键作为主键生成策略。对于预训练模型，系统通过直接存储其 PyTorch 导出文件。

### 5.3.2 后端模块

本系统的后端采用 Python 的 Flask 框架实现，通过该框架连接数据库并执行 SQL 语句，从而完成与数据模块的交互。同时，系统利用 PyTorch 实现脉冲神经网络模型的训练、保存及加载。

## 数据预处理与特征构建

在推荐系统中，特征工程是至关重要的一步，其数据质量和预处理方法往往对系统的最终表现有着比模型本身更显著的影响。因此，本系统也采用了一整套完善的数据预处理与特征构建流程。对于用户的编号、性别、职业以及电影编号等单类别型变量，系统将其编码为由 0 开始的整数，作为后续模型嵌入层的索引。由于一个电影可能同时涵盖多个类型，因此电影类别采用了 Multi-Hot 编码

策略。针对用户年龄和电影上映时间等连续型变量，系统首先对其进行分段，再根据其所属的区间赋予整数索引。此方法不仅简化了数据的表示，同时有助于抑制过拟合和减少异常值的影响。

在实际应用中，部分用户可能选择不上传某些个人信息，导致部分数据出现缺失。对于用户性别、职业及电影类别等离散变量的缺失值，系统采用单独归类的方式以保留缺失状态的信息。而对于年龄和电影上映时间等数值型变量，则采用均一填充方法，即在编码的每个维度上使用统一数值填补缺失数据。

## 召回模型

由于电影数据量通常较大，对每个电影进行精准排序的计算开销较高，因此需要设计高效的召回策略，以便快速从海量电影中筛选出候选集。系统采用了多路并行的召回策略，主要包括以下几类：基于用户基本信息计算与预设典型用户的相似度，从而召回相似用户高评分的电影；通过电影偏好类型召回，结合用户在注册时主动提供的偏好或历史评价记录，筛选出用户可能感兴趣的电影类型；热门及新电影召回，筛选出近期评分较高、评价数量较多的电影以及新上映电影。在召回过程中，为避免重复推荐，系统会过滤掉用户已评分过的电影。

多种召回策略的结合丰富了候选样本的多样性，为后续的精确定序提供了更为全面的候选数据集。此外，不同的召回策略可以应对更多的应用场景，在有效利用用户的历史交互数据的同时，兼顾解决新用户推荐内容的冷启动问题。

## 推荐值排序模型

为了从召回的候选电影中筛选出用户最可能感兴趣的电影，系统进一步采用脉冲神经网络对候选电影进行推荐值计算与排序。尽管工业级推荐系统通常将排序过程细分为粗排、精排、重排等多个阶段，但考虑到本系统所处理的数据量较为有限，无法与实际工业环境相比，因此仅采用单一模型实现推荐值计算。

在推荐值排序模型中，每种类别型变量会通过嵌入层转化为嵌入向量，其中用户编号、性别、年龄、职业以及电影编号、上映时间等单类别变量通过查表（Lookup）转化，电影类别这种多类别变量则通过矩阵乘法（Matmul）得到嵌入向量。系统将这些嵌入向量拼接，作为一条评分的特征。由于脉冲神经网络的输

入、输出是 0/1 的，因此嵌入向量需要经过 Sigmoid 函数，以将数据变换到 [0, 1] 区间，以对应脉冲发放频率。由于排序模型使用的特征不包含历史交互信息，因此无需考虑新用户的冷启动问题。接下来，通过泊松编码将拼接后的特征转化为脉冲编码，最后由脉冲神经网络输出对应的推荐值。

### 5.3.3 前端模块

推荐系统的前端作为用户直接交互的模块，承担着展示推荐结果和支持用户操作的关键功能。为确保用户体验的友好性与直观性，本系统的前端基于开源的 Bootstrap 框架实现。前端模块将用户操作请求发送至后端模块，并以可视化方式呈现返回的推荐结果和相关电影数据。我们设计了用户登录/注册、电影数据上传、电影评分、电影推荐等界面，来对用户操作进行支持。

## 5.4 效果展示

### 5.4.1 界面展示

用户访问系统后，首先进入登录界面。在首次使用时，用户需完成注册，填写基本信息并提交注册请求后，系统将分配一个唯一账号。用户可使用该账号及注册时设置的密码在登录界面进行系统登录。需要注意的是，用户在注册时，用户名与密码为必填项，而其他信息则可选择性填写。图5-2展示了系统的登录和注册界面。

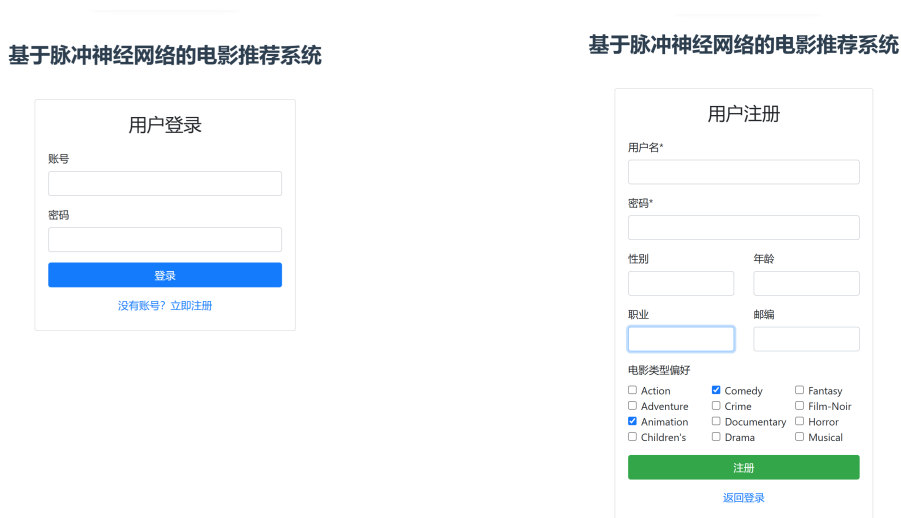


图 5-2 系统登录/注册界面

用户登录后，即可进入系统主页。主页包含三个主要板块：电影推荐、电影评分以及电影上传。页面顶部导航栏右侧展示用户头像与用户名，点击后可跳转至修改基本信息的页面。进入主页时系统默认展示电影推荐板块。电影推荐页展示了电影的基本信息，如名称、初始上映时间、类型以及评分等，并提供“查看详情”按钮以使用户查询电影的详细信息。用户在注册阶段时，电影类型偏好选择了喜剧，因此召回的内容中包括了较多的喜剧，从而影响最终推荐结果。由图5-3可见，系统推荐电影以喜剧为主，也包含其他类型的高评分电影。如果用户对已观看的电影进行评分，可切换至电影评分板块，并通过关键词搜索功能查找目标电影。在搜索结果页面中，用户不仅可以对电影进行评分，还可以撰写短评（此项为可选操作）。此外，为满足系统中未涵盖但用户希望评价的电影需求，用户可进入电影上传板块，自行填写电影相关信息，创建新的电影条目，从而丰富电影数据集。

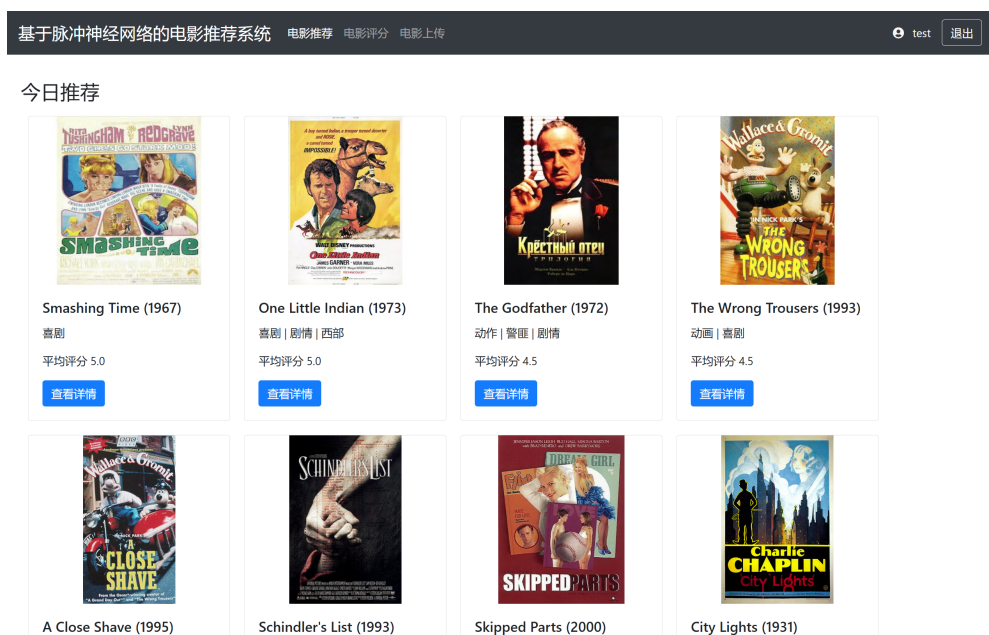


图 5-3 系统主页电影推荐板块

### 5.4.2 推荐结果量化对比

为验证本文双线性整合脉冲神经元在二阶特征交互学习上的优势，并探究其在推荐系统中的应用潜力，本小节选取公开的 MovieLens 数据集开展对比实验。鉴于系统实际运行环境中推荐效果难以通过指标精确量化，因此使用公开数据集上的客观结果评估模型性能。MovieLens 数据集包含 6040 名用户、3952 部

电影与 1000209 条评分记录，本使用以 8:2 的比例随机划分训练集和测试集。由于实验环境与实际运行的系统存在差异，因此使用流程中省略了召回阶段，直接在全量数据上训练排序模型；同时，数据集中无缺失或异常值，所以无需进行缺失值填充或异常处理。

在数据预处理阶段，年龄特征的划分规则为：18 岁以下、18-22 岁、22-25 岁各为一类，之后每 10 年为一类别，以捕捉不同年龄段的偏好差异。随后，通过嵌入层，将用户编号映射为 32 维向量，用户性别、年龄、职业以及电影编号、类型、上映年份各映射为 16 维向量，并在拼接后得到 128 维联合特征表示。考虑到嵌入层并非脉冲结构，因此实验中首先将嵌入层接入一般的全连接人工神经网络进行预训练，之后将嵌入层单独作为静态特征提取器。嵌入层输出经 Sigmoid 激活，再通过泊松编码转换为脉冲序列，作为后续脉冲神经网络的输入，参与脉冲神经网络的训练与推理过程。

表 5-1 展示了脉冲神经网络使用传统神经元与双线性整合脉冲神经元的实验结果，均取训练过程中测试集上最低的 MSE 结果。可见，采用双线性整合的脉冲神经元相比线性整合在相同网络结构及训练轮数下取得了更低的误差，说明其在二阶特征交互建模方面具有更强的学习能力，更契合推荐系统的应用需求。

表 5-1 MovieLens 数据集实验结果对比。表中网络结构省略了特征嵌入层。

时间窗口	网络结构	训练轮数	整合方式	$\lambda$ 初值	MSE
20	128-100-1	10	线性	-	0.857
20	128-100-1	10	双线性	0.1	<b>0.845</b>

## 5.5 本章小结

本章介绍了基于脉冲神经网络构建的电影推荐系统。该系统为每位用户提供个性化的电影推荐，帮助用户高效找到感兴趣的电影，并支持用户对观看过的电影进行评分与评价。系统在推荐值计算过程中采用了脉冲神经网络模型，并应用了本文第三章提出的双线性整合脉冲神经元，以增强对稀疏数据的处理能力。该方法在推荐任务中表现良好，并且能够优化推荐系统的功耗和延迟表现，具备实际应用价值。

## 第六章 总结与展望

脉冲神经网络因其事件驱动特性以及高能效、低时延的固有优势，正逐渐成为新的研究热点。然而，目前基于生物神经元特性构建新型脉冲神经元模型的基础性研究仍相对匮乏。本文针对传统 LIF 脉冲神经元模型在非线性表达能力上的局限性，基于生物神经元整合特性，提出了两种具有生物可解释性的改进方案。通过系统分析生物神经元的双线性膜电位响应特性和分流抑制机制，本研究突破了传统 LIF 模型的线性整合范式，在神经元层面增强了脉冲网络的非线性计算能力。

基于生物神经元膜电位响应呈现的双线性特性，本文创新性地设计了双线性整合脉冲神经元模型。该模型不仅在数学建模层面能够更精确地模拟生物神经元膜电位的动态响应特性，还特别增强了对输入信号中二阶特征交互的建模能力，从而提升了网络对复杂非线性关系的计算性能。通过借鉴因子分解机中的矩阵分解技术，本研究实现了对双线性整合模式的高效近似，在保持与线性整合相同渐进计算复杂度的同时，提升了神经元对异或关系等非线性特征的学习能力，在计算效率与模型性能之间实现了更优平衡。大量对比实验与消融实验表明，该模型具有良好的多架构适应性和多数据集泛化能力。为验证其应用价值，我们搭建了基于脉冲神经网络的推荐系统，将双线性整合脉冲神经元应用其中，验证了其工程应用价值。

基于神经元膜电位非线性响应背后的电生理学机制，本文在脉冲神经元模型中引入具有生物物理基础的分流抑制特性，以更精准地模拟生物神经元的信息处理过程。基于参数可学习动态门控机制，该模型能够自适应调节分流比例，从而模拟生物神经元的非线性抑制过程，并进一步增强网络对复杂时空模式的表征能力。异或计算能力的理论分析表明，分流抑制机制同样能够赋予脉冲神经元更强的非线性表达能力。实验结果显示，该模型在多个主流脉冲神经网络架构和数据集上实现了性能提升。

本文的两项工作虽然都是针对脉冲神经元的非线性计算能力不足问题而展开，但其具体机制和适用场景存在一定区别。双线性整合脉冲神经元模型通过低秩矩阵分解直接建模二阶特征交互，在稀疏数据（如推荐系统场景）中表现出色，适合处理存在高阶特征交互的数据。而分流抑制脉冲神经元模型更倾向于模拟真实的生物神经元响应机制，从而实现膜电位的非线性响应模式，具有更丰富的神经生理学意义。

当然，本文的研究工作尚存在一些局限性。双线性整合脉冲神经元在部分二阶特征交互不明显的数据集上可能存在过拟合问题，未来的研究工作可进一步探索更合理的线性项与二阶特征交互项整合范式，以适应不同特征分布的数据；分流抑制脉冲神经元模型设计较为复杂，计算成本在传统脉冲神经元的三倍以上，后续工作应着力于开发更轻量化的实现方案。值得注意的是，脉冲神经网络的硬件协同设计也是一个重要的研究方向，例如对二阶特征交互项或门控计算的硬件级加速可以提高新型脉冲神经网络的计算效率，这对新模型的实际部署具有重要意义。

## 参考文献

- [1] MCCULLOCH W S, PITTS W. A logical calculus of the ideas immanent in nervous activity[J]. Bulletin of Mathematical Biophysics, 1943, 5: 115-133.
- [2] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning representations by back-propagating errors[J]. Nature, 1986, 323(6088): 533-536.
- [3] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [4] HOCHREITER S. Long Short-term Memory[J]. Neural Computation MIT-Press, 1997.
- [5] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (CVPR). 2016: 770-778.
- [6] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[C] // Advances in Neural Information Processing Systems.(NeurIPS). 2017: 5998-6008.
- [7] IOFFE S, SZEGEDY C. Batch Normalization: accelerating deep network training by reducing internal covariate shift[C] // Proceedings of the International Conference on Machine Learning. (ICML): vol. 37. 2015: 448-456.
- [8] BA L J, KIROS J R, HINTON G E. Layer Normalization[J]. CoRR, 2016, abs/1607.06450.
- [9] WU Y, HE K. Group Normalization[C] // European Conference on Computer Vision.(ECCV): vol. 11217. 2018: 3-19.
- [10] SRIVASTAVA N, HINTON G E, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. Journal of Machine Learning Research, 2014, 15(1): 1929-1958.

- [11] NGUYEN A M, YOSINSKI J, CLUNE J. Deep neural networks are easily fooled: high confidence predictions for unrecognizable images[C]//IEEE Conference on Computer Vision and Pattern Recognition.(CVPR). 2015: 427-436.
- [12] MADRY A, MAKELOV A, SCHMIDT L, et al. Towards deep learning models resistant to adversarial attacks[C]//International Conference on Learning Representations.(ICLR). 2018.
- [13] WU J, CHUA Y, LI H. A biologically plausible speech recognition framework based on spiking neural networks[C]//International Joint Conference on Neural Networks.(IJCNN). 2018: 1-8.
- [14] ZHUL, WANG X, CHANG Y, et al. Event-based video reconstruction via potential-assisted spiking neural network[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition. (CVPR). 2022: 3584-3594.
- [15] LIU D, BELLOTTO N, YUE S. Deep spiking neural network for video-based disguise face recognition based on dynamic facial movements[J]. IEEE Transactions on Neural Networks and Learning Systems, 2020, 31(6): 1843-1855.
- [16] BING Z, MESCHEDE C, RÖHRBEIN F, et al. A survey of robotics control based on learning-inspired spiking neural networks[J]. Frontiers in Neurobotics, 2018, 12: 35.
- [17] MAQUEDA A I, LOQUERCIO A, GALLEGO G, et al. Event-based vision meets deep learning on steering prediction for self-driving cars[C]//IEEE Conference on Computer Vision and Pattern Recognition. (CVPR). 2018: 5419-5427.
- [18] HOPFIELD J J. Neural networks and physical systems with emergent collective computational abilities[J]. Proceedings of the National Academy of Sciences, 1982, 79(8): 2554-2558.
- [19] ACKLEY D H, HINTON G E, SEJNOWSKI T J. A learning algorithm for Boltzmann machines[J]. Cognitive Science, 1985, 9(1): 147-169.
- [20] MAASS W. Networks of spiking neurons: the third generation of neural network models[J]. Neural Networks, 1997, 10(9): 1659-1671.

- 
- [21] SAUNDERS D J, PATEL D, HAZAN H, et al. Locally connected spiking neural networks for unsupervised feature learning[J]. *Neural Networks*, 2019, 119: 332-340.
- [22] GHAEMI H, MIRZAEI E, NOURI M, et al. BioLCNet: reward-modulated locally connected spiking neural networks[C]// *International Conference on Machine Learning (ICML)*: vol. 13811. 2022: 564-578.
- [23] HEBB D O. *The organization of behavior: a neuropsychological theory*[M]. Psychology Press, 2005.
- [24] DIEHL P U, COOK M. Unsupervised learning of digit recognition using spike-timing-dependent plasticity[J]. *Frontiers in Computational Neuroscience*, 2015, 9: 99.
- [25] DONG Y, ZHAO D, LI Y, et al. An unsupervised STDP-based spiking neural network inspired by biologically plausible learning rules and connections[J]. *Neural Networks*, 2023, 165: 799-808.
- [26] MOZER M C. A focused backpropagation algorithm for temporal pattern recognition[J]. *Complex Systems*, 1989, 3(4).
- [27] ROBINSON A J, FALLSIDE F. *The utility driven dynamic error propagation network*[M]. 1987.
- [28] WERBOS P J. Generalization of backpropagation with application to a recurrent gas market model[J]. *Neural Networks*, 1988, 1(4): 339-356.
- [29] WU Y, DENG L, LI G, et al. Spatio-temporal backpropagation for training high-performance spiking neural networks[J]. *Frontiers in Neuroscience*, 2018, 12: 331.
- [30] ZHENG H, WU Y, DENG L, et al. Going deeper with directly-trained larger spiking neural networks[C]// *AAAI Conference on Artificial Intelligence*. 2021: 11062-11070.
- [31] KIM Y, PANDA P. Revisiting batch normalization for training low-latency deep spiking neural networks from scratch[J]. *CoRR*, 2020, abs/2010.01729.

- [32] GUO Y, ZHANG Y, CHEN Y, et al. Membrane potential batch normalization for spiking neural networks[C]//IEEE/CVF International Conference on Computer Vision.(ICCV). 2023: 19363-19373.
- [33] HUNSBERGER E, ELIASMITH C. Training spiking deep networks for neuro-morphic hardware[J]. CoRR, 2016, abs/1611.05141.
- [34] RUECKAUER B, LUNGU I A, HU Y, et al. Conversion of continuous-valued deep networks to efficient event-driven networks for image classification[J]. *Frontiers in Neuroscience*, 2017, 11: 682.
- [35] SENGUPTA A, YE Y, WANG R, et al. Going deeper in spiking neural networks: VGG and residual architectures[J]. CoRR, 2018, abs/1802.02627.
- [36] HAN B, SRINIVASAN G, ROY K. RMP-SNN: residual membrane potential neuron for enabling deeper high-accuracy and low-latency spiking neural network[C]//IEEE/CVF Conference on Computer Vision and Pattern Recognition.(CVPR). 2020: 13555-13564.
- [37] THANAPITAK S, TOUMAZOU C. A bionics chemical synapse[J]. *IEEE Transactions on Biomedical Circuits and Systems.(CASS)*, 2012, 7(3): 296-306.
- [38] HODGKIN A L, HUXLEY A F. A quantitative description of membrane current and its application to conduction and excitation in nerve[J]. *Journal of Physiology*, 1952, 117(4): 500.
- [39] ABBOTT L F. Lapique's introduction of the integrate-and-fire model neuron (1907)[J]. *Brain Research Bulletin*, 1999, 50(5-6): 303-304.
- [40] LOTFI-REZAABAD A, VISHWANATH S. Long short-term memory spiking networks and their applications[C]//Proceedings of the International Conference on Neuromorphic Systems.(ICONS). 2020: 3:1-3:9.
- [41] DAMPFHOFFER M, MESQUIDA T, VALENTIAN A, et al. Investigating current-based and gating approaches for accurate and energy-efficient spiking recurrent neural networks[C]//International Conference on Artificial Neural Networks.(ICANN): vol. 13531. 2022: 359-370.

- [42] MINSKY M, PAPER S A. Perceptrons: An Introduction to Computational Geometry[M]. The MIT Press, 2017.
- [43] GIDON A, ZOLNIK T A, FIDZINSKI P, et al. Dendritic action potentials and computation in human layer 2/3 cortical neurons[J]. Science, 2020, 367(6473): 83-87.
- [44] KOCH C, SEGEV I. The role of single neurons in information processing[J]. Nature Neuroscience, 2000, 3(11): 1171-1177.
- [45] LI S, LIU N, ZHANG X, et al. Bilinearity in spatiotemporal integration of synaptic inputs[J]. PLOS Computational Biology, 2014, 10(12).
- [46] STUART G J, SPRUSTON N. Dendritic integration: 60 years of progress[J]. Nature Neuroscience, 2015, 18(12): 1713-1721.
- [47] BORG-GRAHAM L J, MONIER C, FREGNAC Y. Visual input evokes transient and strong shunting inhibition in visual cortical neurons[J]. Nature, 1998, 393(6683): 369-373.
- [48] DAUPHIN Y N, FAN A, AULI M, et al. Language modeling with gated convolutional networks[C] // PRECUP D, TEH Y W. Proceedings of the International Conference on Machine Learning.(ICML): vol. 70. 2017: 933-941.
- [49] RAMACHANDRAN P, ZOPH B, LE Q V. Searching for activation functions[C] // International Conference on Learning Representations.(ICLR). 2018.
- [50] LIU M, DELBRÜCK T. Block-matching optical flow for dynamic vision sensors: algorithm and FPGA implementation[C] // IEEE International Symposium on Circuits and Systems(ISCAS). 2017: 1-4.
- [51] CHEN X, YANG Q, WU J, et al. A hybrid neural coding approach for pattern recognition with spiking neural networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(5): 3064-3078.
- [52] YU F, KOLTUN V. Multi-scale context aggregation by dilated convolutions[C] // International Conference on Learning Representations.(ICLR). 2016.

- [53] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [54] HOWARD A G, ZHU M, CHEN B, et al. MobileNets: efficient convolutional neural networks for mobile vision applications[J]. CoRR, 2017, abs/1704.04861.
- [55] FANG W, CHEN Y, DING J, et al. SpikingJelly: an open-source machine learning infrastructure platform for spike-based intelligence[J]. CoRR, 2023, abs/2310.16620.
- [56] ESHRAGHIAN J K, WARD M, NEFTCI E, et al. Training spiking neural networks using lessons from deep learning[J]. Proceedings of the IEEE, 2023, 111(9): 1016-1054.
- [57] HAZAN H, SAUNDERS D J, KHAN H, et al. BindsNET: a machine learning-oriented spiking neural networks library in Python[J]. Frontiers Neuroinformatics, 2018, 12: 89.
- [58] HAO Y, HUANG X, DONG M, et al. A biologically plausible supervised learning method for spiking neural networks using the symmetric STDP rule[J]. Neural Networks, 2020, 121: 387-395.
- [59] HU Z, WANG T, HU X. An STDP-based supervised learning algorithm for spiking neural networks[J]. CoRR, 2022, abs/2203.03379.
- [60] LU S, SENGUPTA A. Deep unsupervised learning using spike-timing-dependent plasticity[J]. Neuromorphic Computing and Engineering, 2024, 4(2): 24004.
- [61] ZHU Y, YU Z, FANG W, et al. Training spiking neural networks with event-driven backpropagation[C]//Advances in Neural Information Processing Systems.(NeurIPS). 2022.
- [62] HAO Z, BU T, DING J, et al. Reducing ANN-SNN conversion error through residual membrane potential[C]//AAAI Conference on Artificial Intelligence. 2023: 11-21.

- [63] DENG S, GU S. Optimal conversion of conventional artificial neural networks to spiking neural networks[C]//International Conference on Learning Representations.(ICLR). 2021.
- [64] SAMADZADEH A, FAR F S T, JAVADI A, et al. Convolutional spiking neural networks for spatio-temporal feature extraction[J]. Neural Processing Letters, 2023, 55(6): 6979-6995.
- [65] SALERNO M, SUSI G, CRISTINI A. Accurate latency characterization for very large asynchronous spiking neural networks[C]//. 2011: 116-124.
- [66] SUSI G, ANTÓN TORO L, CANUET L, et al. A neuro-inspired system for on-line learning and recognition of parallel spike trains, based on spike latency, and heterosynaptic STDP[J]. Frontiers in neuroscience, 2018, 12: 780.
- [67] LEHMANN H M, HILLE J, GRASSMANN C, et al. Leaky Integrate-and-Fire neuron with a refractory period mechanism for invariant spikes[C]//Conference on Ph.D Research in Microelectronics and Electronics.(PRIME). 2022: 365-368.
- [68] ZHAO Y, LIU Y, LIU X, et al. Self-adapting spiking neural P systems with refractory period and propagation delay[J]. Information Sciences, 2022, 589: 80-93.
- [69] RENDLE S. Factorization Machines[C]//IEEE International Conference on Data Mining.(ICDM). 2010: 995-1000.
- [70] GUO H, TANG R, YE Y, et al. DeepFM: a Factorization-Machine based neural network for CTR prediction[C]//International Joint Conference on Artificial Intelligence.(IJCAI). 2017: 1725-1731.
- [71] FANG W, YU Z, CHEN Y, et al. Incorporating learnable membrane time constant to enhance learning of spiking neural networks[C]//IEEE/CVF International Conference on Computer Vision.(ICCV). 2021: 2641-2651.
- [72] ZHANG S, ZHANG Z, ZHOU Z. Bifurcation spiking neural network[J]. Journal of Machine Learning Research, 2021, 22: 253:1-253:21.

- [73] WANG Q, ZHANG T, HAN M, et al. Complex dynamic neurons improved spiking Transformer network for efficient automatic speech recognition[C]// AAAI Conference on Artificial Intelligence. 2023: 102-109.
- [74] ZHU R, ZHAO Q, LI G, et al. SpikeGPT: generative pre-trained language model with spiking neural networks[J]. Transactions on Machine Learning Research, 2024, 2024.
- [75] HE K, ZHANG X, REN S, et al. Delving deep into rectifiers: surpassing human-Level performance on imageNet classification[C]// IEEE International Conference on Computer Vision.(ICCV). 2015: 1026-1034.
- [76] HARPER F M, KONSTAN J A. The MovieLens datasets: history and context[J]. ACM Transactions on Interactive Intelligent Systems, 2015, 5(4).

## 致 谢

在南京大学本硕七年的学习生活即将结束，站在校园到社会的转折点，我心中充满了对未来的憧憬，以及对过去点滴时光的怀念与不舍。作为人工智能学院的第一届本科生，我很幸运能够和学院一起成长，度过一段充实而愉快的时光。学院完善的课程设置与培养体系让我受益良多，对人工智能技术有了更加全面深刻的理解。这也为我未来的工作打下了坚实基础，在此我深表感谢。

首先，衷心感谢申老师。您为我指明了科研方向，每次与您的交流都让我在思路和方法上获益良多；在项目实践中，您提出的建议也让我更快地解决了遇到的问题。这些经历对我的研究和求职都帮助很大。

其次，非常感谢徐百乐师兄。您在我的毕业论文写作过程中耐心指导，从论文结构编排，到模型优化方法，再到实验设计，都给了我许多实用的建议，让我能够顺利完成论文。

还要特别感谢我的家人和朋友。父母始终支持我的选择，帮我解决生活中的难题；朋友们也会与我互相交流信息，并在我遇到困难时给予帮助。正是这些支持让我能够专注于学业，顺利完成本硕阶段的学习。

最后，感谢所有帮助过我的人。希望人工智能学院不断发展壮大，培养出更多优秀的人才，在如今 AI 时代大放异彩。



# 简历与科研成果

## 基本信息

刘佩涵，男，2000年8月9日出生，河北省邢台人。

## 教育背景

南京大学人工智能学院	硕士	2022年9月—2025年6月
南京大学人工智能学院	本科	2018年9月—2022年6月

## 攻读硕士学位期间参与的科研课题

1. 科技创新 2030 — “脑科学与类脑研究” 重大项目 “基于神经可塑性的脉冲网络高效学习机制与类脑智能系统”（参与课题年限 2022 年 9 月—2025 年 6 月），负责脉冲神经网络模型相关研究。

## 攻读硕士学位期间的发明专利

1. 徐百乐，刘佩涵，申富饶. 《一种基于新型脉冲神经网络的稀疏数据高效处理方法》。专利申请号：202510436326.7。



# 学位论文出版授权书

本人完全同意《中国优秀博硕士学位论文全文数据库出版章程》（以下简称“章程”），愿意将本人的学位论文提交“中国学术期刊（光盘版）电子杂志社”在《中国博士学位论文全文数据库》、《中国优秀硕士学位论文全文数据库》中全文发表。《中国博士学位论文全文数据库》、《中国优秀硕士学位论文全文数据库》可以以电子、网络及其他数字媒体形式公开出版，并同意编入《中国知识资源总库》，在《中国博硕士学位论文评价数据库》中使用和在互联网上传播，同意按“章程”规定享受相关权益。

作者签名：\_\_\_\_\_

\_\_\_\_\_年\_\_月\_\_日

论文题名	基于生物神经元非线性整合机制的脉冲神经元模型研究				
研究生学号	502022370027	所在院系	人工智能学 院	学位年度	2025
论文级别	<input checked="" type="checkbox"/> 学术学位硕士 <input type="checkbox"/> 专业学位硕士 <input type="checkbox"/> 学术学位博士 <input type="checkbox"/> 专业学位博士				
作者 Email	peihanliu@smail.nju.edu.cn				
导师姓名	申富饶				

论文涉密情况：

不保密

保密，保密期（\_\_\_\_\_年\_\_月\_\_日至\_\_\_\_\_年\_\_月\_\_日）

