

基于多子网络预训练的脉冲神经网络分类模型

卓明松, 莫凌飞

引用本文

卓明松, 莫凌飞. [基于多子网络预训练的脉冲神经网络分类模型](#)[J]. 计算机科学, 2024, 51(11A): 240300191-6.

ZHUO Mingsong, MO Lingfei. [Spiking Neural Network Classification Model Based on Multi-subnetwork Pre-training](#) [J]. Computer Science, 2024, 51(11A): 240300191-6.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于策略融合及Spiking DRL的移动机器人路径规划方法](#)

Mobile Robots' Path Planning Method Based on Policy Fusion and Spiking Deep Reinforcement Learning

计算机科学, 2024, 51(11A): 240100211-11. <https://doi.org/10.11896/jsjcx.240100211>

[自校准首脉冲时间编码神经元模型](#)

Self-calibrating First Spike Temporal Encoding Neuron Model

计算机科学, 2024, 51(3): 244-250. <https://doi.org/10.11896/jsjcx.221200003>

[类脑计算脉冲神经网络模型及其学习算法研究进展](#)

Spiking Neural Network Model for Brain-like Computing and Progress of Its Learning Algorithm

计算机科学, 2023, 50(1): 229-242. <https://doi.org/10.11896/jsjcx.220100058>

[一种基于脉冲神经网络的星体表面岩石检测算法](#)

Onboard Rock Detection Algorithm Based on Spiking Neural Network

计算机科学, 2023, 50(1): 98-104. <https://doi.org/10.11896/jsjcx.211100149>

[类脑计算新发展——“TrueNorth”神经元芯片](#)

New Development of Artificial Cognitive Computation: TrueNorth Neuron Chip

计算机科学, 2016, 43(Z6): 17-20. <https://doi.org/10.11896/j.issn.1002-137X.2016.6A.002>

基于多子网络预训练的脉冲神经网络分类模型

卓明松 莫凌飞

东南大学仪器科学与工程学院 南京 210096

(220213666@seu.edu.cn)

摘要 脉冲神经网络(Spiking Neural Network, SNN)被认为是最符合生物大脑机制的类脑计算模型,凭借其事件驱动、高效、可解释等特点吸引了越来越多的研究关注。然而,由于脉冲的二值输出与不可微分性,SNN的训练方法仍存在一定空缺。于是借鉴皮层记忆单元通过局部网络存储记忆信息的方式,提出一种基于多子网络预训练的脉冲神经网络分类方法。该方法使用样本标签信息优化了脉冲序列特征提取过程,采用改进的脉冲时间依赖可塑性学习规则预训练多个单类别特征提取子网络,并将预训练后的子网络进行无监督特征融合,有效提高了网络的特征分类能力。此外,在权重可视化与 t-SNE 可视化工具的帮助下,分析了方法的有效性。所提方法在 MNIST 与 Fashion-MNIST 数据集上分别取得了 97.40% 与 88.81% 的分类准确度。

关键词: 脉冲神经网络;脉冲时间依赖可塑性;单类别特征提取子网络;无监督特征融合;类脑计算

中图分类号 TP389.1

Spiking Neural Network Classification Model Based on Multi-subnetwork Pre-training

ZHUO Mingsong and MO Lingfei

School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China

Abstract Spiking neural network(SNN) is widely regarded as the most biologically plausible model, aligning closely with the mechanisms of the biological brain. It has garnered increasing research attention due to its event-driven nature, high energy efficiency, and interpretability. However, the training methods of SNN still have some limitations due to the binary output and non-differentiability of the spike. This paper proposes a SNN classification method based on multi-subnetwork pre-training, which draws inspiration from the way cortical memory units store memory information through local networks. This approach leverages sample label information to optimize the feature extraction process, employs enhanced spike-timing-dependent-plasticity learning rules for pre-training multiple single-class feature extraction subnetworks, and conducts unsupervised feature fusion on these pre-trained subnetworks to effectively enhance the network's feature classification capability. Furthermore, the effectiveness of this method is analyzed through weight visualization and t-SNE visualization tools. Finally, the classification accuracy of 97.40% and 88.81% is achieved on MNIST and Fashion MNIST datasets, respectively.

Keywords Spiking neural network, Spike-timing-dependent-plasticity, Single-class feature extraction subnetworks, Unsupervised feature fusion, Brain-like computing

1 引言

近年来,以卷积神经网络(Convolutional Neural Network, CNN)^[1]为代表的人工神经网络技术(Artificial Neural Network, ANN)快速发展,在诸多领域得到了广泛应用,极大地推动了信号处理、模式识别等领域的发展。然而,ANN目前仍存在着一定的问题。首先,大多数 ANN 模型高度依赖于样本标签,这一过程需要消耗大量的人力。同时,当前深度学习算法需要耗费许多算力与资源,以复现 AlphaGo Zero^[2]论文中的实验为例,经过估算和折合,大约需要花费 3 500 万美元。与此形成鲜明对比的是,人脑中大约有 1 00 万亿个神经突触和 1 000 亿个神经元相互联结组成复杂的脑神经网络,其每天处理着各种复杂的事务,但大脑的等效功耗仅

有 24 瓦特^[4]。此外,ANN 中常用的误差反向传播机制在生物神经网络中缺乏广泛的证据。尽管 ANN 采用仿生结构,但其学习算法从原理上来说并不具有生物可解释性。综上所述,为了创造具备更低功耗且具有人类智能特征的人工智能,越来越多的研究人员把目光投向了脑科学和类脑智能技术。

脉冲神经网络^[5]作为一种重要的类脑计算模型,在原理上更加接近生物神经网络。SNN 神经元之间使用脉冲信号传递信息,相对于 ANN 使用数值信号传递信息,神经元之间使用脉冲信号通讯可以大大降低通讯开销。其无监督训练过程也是基于生理上合理的脉冲时序依赖可塑性(Spike Timing-Dependence Plasticity, STDP)^[6],而不是误差反向传播(Back Propagation, BP)。得益于 SNN 事件驱动计算的特点,未被激活的神经元不会参与实际计算。因此,SNN 非常适用

基金项目:江苏省高校青蓝工程

This work was supported by the Blue Project of Jiangsu Province.

通信作者:莫凌飞(lfmo@seu.edu.cn)

于部署在神经拟态芯片上以实现低功耗计算。国内外学者也开发了许多 SNN 低功耗计算芯片,例如 SpiNNaker^[7], Tianjic^[8], Loihi^[9], Darwin^[10]等。利用这些芯片, SNN 在计算功耗上可以低于 ANN 计算功耗的百分之一^[11]。SNN 凭借其高生物可解释性、低标注依赖以及低功耗等优点吸引着越来越多的关注,被称为“第三代人工神经网络”。

目前,制约 SNN 发展的主要原因是缺乏有效的训练算法。在 SNN 的无监督学习算法中,其核心是 Hebb 规则,即一段时间共同发放脉冲的两个神经元之间的连接会增强。这种模式下,为了避免 SNN 对输入模式的重复学习,许多学者引入了竞争机制。Diehl 等^[12]通过“赢家通吃”与“全局侧向抑制”的方式保证特征层每个神经元所学习的时空模式不同。也有学者借鉴神经网络连接的稀疏性,提出自组织映射(Self-Organizing Mapping, SOM)^[13]方法。该方法认为距离较远的神经元之间倾向于彼此抑制,距离较近的神经元倾向于相互激发。越来越多的证据表明^[14],神经元之间的竞争在塑造学习中起着重要作用,这种竞争迫使神经元表现出不同的特征。

在 SNN 的监督学习算法中,由于脉冲信号的不可微性,在 ANN 中广泛使用的误差反向传播不能用于训练 SNN。因此,很难找到具有生物可解释性的 SNN 监督训练算法。部分学者采取替代梯度解决脉冲不可微的问题,对应的 STBP (Spatio-temporal Backpropagation)算法能够实现精确和高效的梯度计算,但这一类算法往往计算复杂度较高,且可能遇到梯度消失或梯度爆炸等问题。鉴于此,一些研究人员开始研究结合 STDP 规则的监督学习。远程监督方法^[15]使用 STDP 规则,并通过远程教师信号使输出神经元在期望的时间点上达到峰值。虽然网络输出层的突触权值可以学习,但隐藏层的突触在初始化后是不可修改的,不利于 SNN 提取时空特征。中科院 Zhang 等^[16]将神经元膜电位设置为可训练对象,并提出对应的四步训练法。通过教师信号监督网络最后一层的学习,通过膜电位驱动该层对前三层的 STDP 学习进行弱有监督。上述方法都是将监督信号作用于网络的最后一层,优化了网络特征分类过程。若监督信号能够参与网络特征提取,将会进一步提高分类精度,这也是本文改进的方向之一。

针对上述问题,本文提出一种基于 STDP 的 SNN 分类方法。该方法可以分为两个步骤:单类别的特征提取与无监督的特征融合。首先,在特征提取阶段引入教师信号,通过单类别特征提取器学习不同类别样本的精细特征。接着,进行无监督特征融合,将多个子网络的特征进行融合,利用赢家通吃与全局侧向抑制规则,保证每个神经元学习到不同的特征,提高网络分类能力。借鉴脑区中通过局部网络存储记忆信息的模式,本文建立了一种具有生物可解释性的 SNN 两步训练方法,利用教师信号进一步提高算法的特征提取能力,并在 MNIST 与 Fashion-MNIST 数据集上验证算法的有效性。

2 方法

本文提出一种基于多子网络预训练的 SNN 分类方法,其整体结构如图 1 所示。该方法主要包括脉冲编码、单类别特征提取与无监督特征融合 3 个部分。具体来说,图像通过时间编码方法将像素灰度值编码为脉冲序列信号。特征提取过程分为两个步骤:第一步,根据输入样本的标签激活对应单类别特征提取器的神经元;第二步,根据神经元间竞争的方式提

取脉冲信号的时空特征,通过改进的 STDP 突触学习规则更新输入层与特征层对应神经元之间的连接权重。最后,对特征层进行低学习率的无监督特征融合学习,提高网络的特征分类能力。

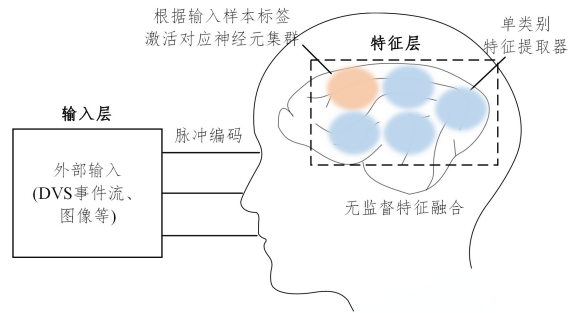


图 1 方法整体结构

Fig. 1 Overall structure of the proposed method

2.1 脉冲编码

在生物神经系统中,需要将外界输入信号编码为脉冲序列,神经元之间通过脉冲信号进行信息传递。常见的编码方式包括频率编码以及时间编码。频率编码模型认为大部分信息包含在神经元的脉冲发放频率中,随着外部输入刺激的增强,神经元脉冲发放频率也会相应增加。该脉冲编码方式更注重脉冲频率,具有不错的抗噪能力,但编码效率较低。时间编码有多种形式,本文采用的是首脉冲编码方式,该方式认为从开始接受刺激到输出首个脉冲的时延包含一定的信息。该编码方式信息编码效率更高,可以用较少的脉冲表达相同的信息。

对于图像数据集而言,每一张图像需要编码为脉冲序列,才能输入到脉冲神经网络中。脉冲通常用地址事件法(Address Event Representation, AER)进行表示,其主要包括发放脉冲的神经元序号(Address)以及脉冲事件的时间戳(Event)。每一张图像需要展平为一维向量,再根据像素灰度值进行时间编码,灰度值越大,对应的脉冲发放时间越靠前。

2.2 单类别特征提取子网络

2.2.1 神经元模型

本文中使用的是 LIF (Leaky Integrated and Fire)^[17]模型,这是一种在 SNN 计算与计算神经科学仿真领域广泛使用的神经元模型。该模型通过 HH (Hodgkin Huxley)模型的简化获得^[18],在保留 HH 模型主要特性的同时去除了那些相对不重要的特性,进而在获得接近 HH 模型计算效果的同时极大地降低了模型复杂度与计算量。其模型计算如式(1)所示:

$$C_m \frac{dV}{dt} = -g_L(V - V_L) + I_{syn} \quad (1)$$

其中, C_m 为膜电容, V 为膜电位, g_L 为漏电导, V_L 为漏电势, I_{syn} 为突触前神经元的输入电流。假设总电导值为 g_E , 并定义常数 $\tau_m = \frac{C_m}{g_L}$, 则式(1)可转化为式(2)。

$$\tau_m \frac{dV}{dt} = -(V - V_L) - \frac{g_E}{g_L}(V - V_E) \quad (2)$$

$$\tau_E \frac{dg_E}{dt} = -g_E + \sum_{j \in N_E} \omega_{j,i} \delta_i \quad (3)$$

式(2)中 g_E 在突触前脉冲的影响下会发生动态变化,具体变化如式(3)所示。即突触前神经元一旦产生脉冲, g_E 会呈非线性增加。 V_E 是兴奋性神经元的反转电位, τ_E 是兴奋性

神经元的电导衰减时间常数, N_E 为突触前神经元的数目, δ_i 为突触前神经元产生脉冲的具体时刻, $\omega_{j,i}$ 为突触前神经元 j 与突触后神经元 i 的连接权值。

$$\text{if}(V > V_{\text{thr}}) \begin{cases} V = V_L \\ T_{\text{ref}} = T_0 \\ V_{\text{thr}} = V_{\text{thr}} + V_{\text{thrDelta}} \end{cases} \quad (4)$$

如式(4)所示,当膜电位 V 增大到超过膜电位阈值 V_{thr} 时,膜电位将被复位,不应期 T_{ref} 设置为 T_0 。在不应期内,神经元不会响应突触前神经元脉冲。同时,为了保证神经元的脉冲放电频率稳定在某一特定范围内,避免某些神经元几乎不发放脉冲,而其他神经元却一直发放脉冲,本文引入了神经元动态阈值机制^[12]。如式(4)所示,每当神经元产生一个脉冲时,神经元的阈值就会相应提高,从而提高下一个脉冲发放的难度。 V_{thrDelta} 是一个超参数,用于控制神经元脉冲生成的难易程度。

$$\tau_{\text{thr}} \frac{dV_{\text{thr}}}{dt} = -(V_{\text{thr}} - V_{\text{thrBase}}) \quad (5)$$

如式(5)所示, V_{thr} 会逐渐衰减为 V_{thrBase} , 暂时降低不产生脉冲的神经元的膜电位阈值。结合式(4),将神经元脉冲放电难度控制在合理范围内。 τ_{thr} 是动态阈值衰减时间常数。

2.2.2 突触模型

本文使用的突触模型主要基于 STDP 规则,并对经典 STDP 规则进行适当的调整,使之更加符合本文模型的需要。

$$\Delta\omega = \begin{cases} \eta(\alpha + \beta \cdot e^{-\frac{ISI}{\tau}}), & \text{if } (ISI > 0) \\ 0, & \text{else} \end{cases} \quad (6)$$

$$ISI = t_{\text{post}} - t_{\text{pre}} \quad (7)$$

式(6)是本文使用突触可塑性公式,其中 $\Delta\omega$ 是每次脉冲后突触权值的更新量; ISI (Inter-Spike Interval) 是突触前后神经元最近脉冲时间的差值,如式(7)所示; η 是学习率; α 是常数偏置项,通常小于 0,用于模拟异突触长时程抑制(Long-Term Depression, LTD)^[19] 现象; β 用于调整权值变化强度,通常大于 0; τ 是 LTP 变化部分的时间常数,用于控制 LTP 部分变化曲线。

2.2.3 单类别特征提取器

在基于突触可塑性学习规则的方法中,神经元之间的竞争方式影响着网络特征提取能力。Diehl 采用赢家通吃与全局侧向抑制的方式训练网络,取得了不错的效果。但该方法训练效率较低,每次训练仅有一个神经元获胜,在大规模网络仿真中需要耗费较多的计算资源。在不同输入样本的影响下,特征层神经元可能会学习到不同类别样本的混乱特征,导致在测试中错误发放脉冲。SOM 聚类算法引入了区域学习的模式,距离较近的神经元可以习得相似的特征,一定程度上提高了学习效率。但该方法缺乏具体的目标函数,常会出现一个自然簇分割到多个 SOM 簇的情况;且该方法受初始网络参数影响较大,每次训练并不能保证收敛。

为了提高网络的特征提取能力,本文提出了单类别特征提取器,训练过程如图 2 所示。受到生物学上的启发^[14],皮层的记忆单元主要通过局部网络的方式存储记忆信息。故本文结合突触可塑性,利用标签信息将输入样本进行分区训练,每个子网络称为单类别特征提取器,仅输入对应类别的样本。子网络中的神经元以改进的 STDP 突触学习规则提取该类别样本的特征。

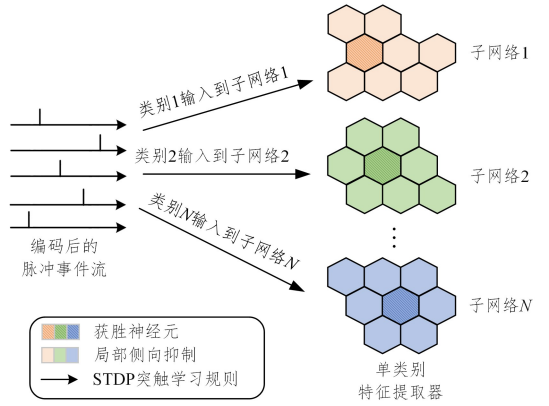


图 2 单类别特征提取器的训练过程

Fig. 2 Training process of single-class feature extractor

训练算法流程如算法 1 所示,标签信息仅用于样本的分区输入这一过程,单类别特征提取还是基于突触可塑性的规则进行的。基于突触可塑性的方法可以很好地提取输入中反复出现的脉冲模式,在侧向抑制的规则下避免同一模式被重复记忆。本文通过简单的监督信号,设定了神经元竞争的范围,影响了网络学习的方向。借助于具象化的目标,即每个子网络仅学习该类别样本下的特征,提高网络特征提取能力。相比于 Diehl 的方法,由于每个输入样本仅会激活单个子网络,在特征层规模同样大的情况下,激活参与计算的突触连接仅为其 1/10,网络整体学习效率更高。

单类别特征提取预训练算法流程如算法 1 所示。

算法 1 单类别特征提取预训练算法

1. 初始化输入层与 10 个特征子网络的连接权重 w_0 , 学习率 $\eta=0.15$
2. for $i=0$ to 59999 do
3. 输入样本脉冲序列 $I_i(t)$
4. 根据输入样本的标签,将脉冲序列输入对应类别的特征子网络
5. 根据特征子网络的输出脉冲事件,触发 STDP 学习规则更新输入层与输出神经元之间的连接权重
6. 输出神经元在子网络内触发局部侧向抑制
7. $i=i+1$

需要注意的是,单类别特征提取子网络仅属于一种预处理的操作,后续还需要将多个子网络的特征进行融合。

2.3 无监督特征融合与分类

在经过单类别特征提取后,每个子网络能够提取到该类别下较为精细的特征,但由于缺少全局侧向抑制,各个子网络间的特征可能存在重复学习。因此,需要一种方法融合不同特征子网络,避免神经元学习重复的特征,以提高网络的特征分类能力。

将预训练后的各个单类别特征提取器组合为特征层,进行低学习率的无监督特征融合。训练过程如图 3 所示。此时,输入层与特征层神经元之间为全连接,每个输入样本只会激活一个特征层神经元,获胜神经元与该输入模式的突触连接几乎不变。而除获胜神经元之外的所有特征层神经元将会抑制对应输入模式的突触连接。在赢家通吃与全局侧向抑制的作用下,连接权重进一步微调。若将每个特征层神经元与输入层之间的突触连接视为一个向量,那么特征融合的作用是进一步扩大各个向量间的距离,抑制重复输入模式的学习,提高网络的特征分类能力。无监督特征融合的算法流程如算法 2 所示。

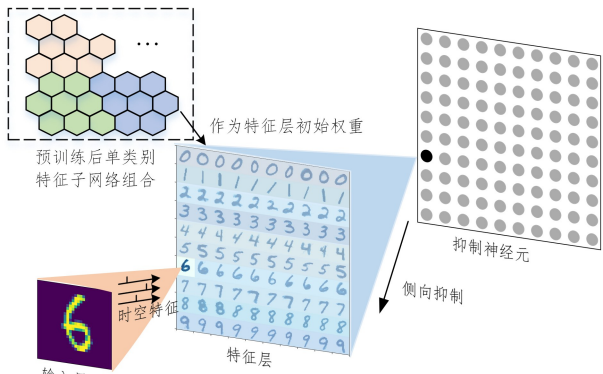


图3 基于预训练的单类别特征提取子网络的特征融合

Fig. 3 Feature fusion of single-class feature extraction sub-network based on pre-training

算法2 无监督特征融合算法

1. 初始化输入层与特征层连接权重为预训练权重 w_1 , 学习率 $\eta=0.01$
2. for $i=0$ to 59999 do
3. 输入样本脉冲序列 $I_i(t)$
4. 根据特征层的输出脉冲事件, 触发异突触 STDP 学习规则更新输入层与输出神经元之间的连接权重
5. 输出神经元在特征层触发全局侧向抑制
6. $i=i+1$

3 实验与结果

下面将通过 MNIST 数据集以及 Fashion-MNIST 数据集的分类任务来证明该方法的性能。

MNIST 数据集广泛应用在各种神经网络的性能测试中, 它包含从 0 到 9 的 10 类手写数字的灰度图像, 其中训练集共 60000 个样本, 测试集共 10000 个样本。Fashion-MNIST 中的 10 类物品为衣服、鞋子等, 贴近实际应用场景, 具有更加丰富的特征。

实验网络的结构包括 3 部分: 输入编码、单类别特征提取以及无监督特征融合。输入层在对输入数据进行编码后, 与各个单类别特征提取器完全连接。编码方式采用时间编码, 即对应像素值越大, 神经元脉冲信号发射越早, 如果该值低于超参数编码阈值, 则不会发射脉冲信号。每张图片在网络中保持 200 ms, 两张图片之间有 50 ms 的间隔, 在这段时间内输入层不会产生脉冲, 用于重置网络状态。

单类别特征提取器的数量与输入样本类别数一致。在特征提取过程中, 标签神经元首先激活对应类别的特征层提取子网络。此后, 输入层与特征子网络之间通过异突触 STDP 学习规则进行权重更新。在无监督特征融合中, 特征层的每个神经元与除自身外的所有神经元都有抑制性突触连接, 用于实现侧向抑制, 以防止模式的重复学习。具体的网络参数如下: 输入层 28×28 , 与数据集中的样本分辨率一致。实验分别在 3 个特征层规模下进行, 分别为 20×20 , 40×40 以及 80×80 , 样本类别数为 10, 所以单类别特征提取器的神经元数量对应为 40, 160 与 640。编码阈值为 0.3。输入层到特征层的突触权重在 0.45 ~ 0.55 之间均匀分布, $\eta=0.015$, $\alpha=-0.3$, $\beta=1.3$, $\tau=20$ 。

参照 Diehl 的两层网络的无监督训练方法, 在异突触 STDP 学习规则下, 20×20 特征层可视化后的权重图如图 4(a)所示。在赢家通吃与全局侧向抑制的权重更新规则下,

每个神经元所学习的模式各不相同, 部分神经元在不同标签样本的影响下, 权重分布较为模糊。图 4(b)在异突触 STDP 学习规则的基础上, 引入了类别监督信号。监督信号将特征层划分为有序的 10 个单类别特征提取器, 各个子网络在改进的 STDP 突触规则下能够更加有效地提取该类别样本的特征, 可视化后的权重图也更加清晰。

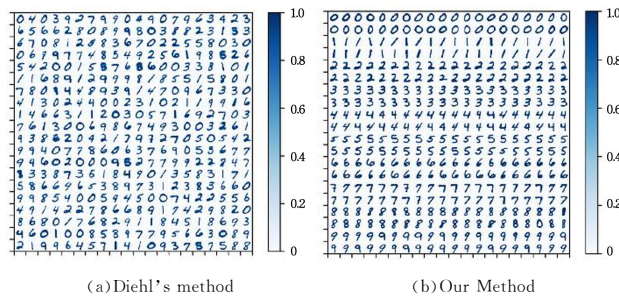


图4 两种方法下特征层权重可视化结果对比

Fig. 4 Visualization results comparison of feature layer weights in two methods

为了进一步可视化方法的特征提取能力, 本文采用高维数据可视化降维工具 t-SNE 比较两种方法所提取特征在二维空间的分布情况, 如图 5 所示。其中, 每张图具有 6400 个点, 对应 6400 个特征层神经元所记忆的特征模式, 两个点之间的距离表示两种特征的相似性, 不同标签的特征通过颜色进行区分。在图 5(a)中, 输入层仅经过全局竞争学习, 特征的聚类效果较为清晰, 但部分类之间仍存在交叉混杂情况, 例如数字 4 与数字 9。经过本文方法训练后的特征层神经元的聚类效果如图 5(b)所示。与图 5(a)相比, 图 5(b)中各类特征聚类效果更加清晰, 类之间距离更大, 类内距离更小, 说明本文方法最终提取的特征有助于网络实现更好的聚类与分类性能。

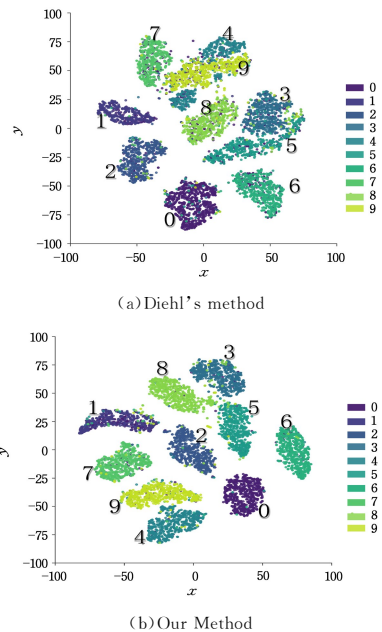


图5 不同方法提取的特征经过 t-SNE 降维后的可视化结果
Fig. 5 Visualization results of features extracted by different methods after t-SNE dimensionality reduction

图 6 为提出方法在 MNIST 数据集上的性能表现, 特征层分别设置为 20×20 , 40×40 以及 80×80 。其中, 精度计算采用的是统计法^[12]。

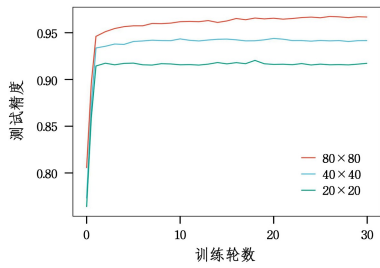


图6 在MNIST数据集不同特征层规模下训练轮数对测试精度的影响

Fig. 6 Impact of training rounds on testing accuracy under different feature layer scales on MNIST dataset

由于MNIST数据集较为简单,网络在经过几轮训练后就能够达到不错的识别性能,随着训练集迭代次数的增加,网络识别精度逐步收敛。当特征层规模扩大时,特征层能够记住更多的输入模式,网络分类精度也会有有一定的提升。当特征层神经元个数为400与1600时,最高识别精度为91.7%与94.27%;当特征层神经元为6400个时,网络可以达到的最佳分类准确率为97.4%,此时的分类混淆矩阵如图7所示。可以看出大部分测试样本的实际类别与预测类别是一致的。

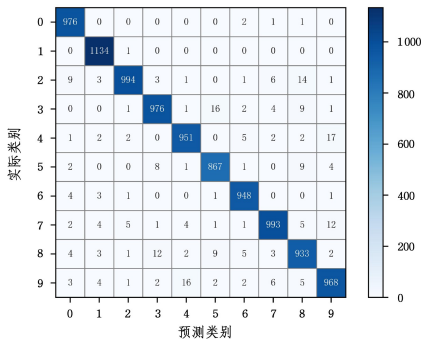


图7 在MNIST数据集的分类混淆矩阵

Fig. 7 Classification confusion matrix on MNIST dataset

为了进一步验证方法的有效性,本文在Fashion-MNIST也进行了实验,网络的具体性能表现如图8所示。

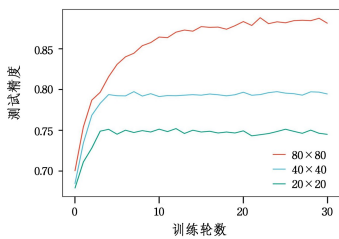


图8 在Fashion-MNIST数据集不同特征层规模下训练轮数对测试精度的影响

Fig. 8 Impact of training rounds on testing accuracy under different feature layer scales on Fashion-MNIST dataset

由于Fashion-MNIST数据集中的服装具有更丰富的特征,网络需要更多的训练轮数才能收敛。最终在20x20,40x40特征层规模下的最高精度分别为75.12%,79.74%。当特征层规模为80x80时,可以达到最佳分类精度88.81%,此时对应的混淆矩阵如图9所示。可以看出,大部分测试样本的预测类别与实际类别一致,标签为4(大衣)与标签为6(衬衫)的测试样本的预测准确度还需进一步提升。

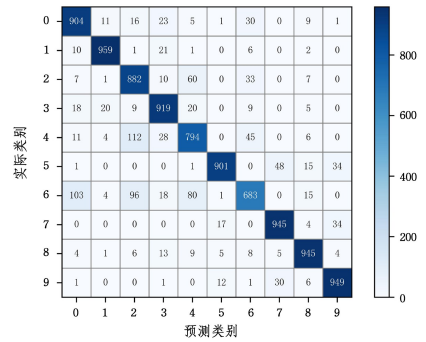


图9 在Fashion-MNIST数据集的分类混淆矩阵

Fig. 9 Classification confusion matrix on Fashion-MNIST

表1列出了部分方法在MNIST数据集的分类准确度。其中,Diehl等利用无监督STDP学习规则,在特征层神经元数量为6400时取得了95.0%的准确度,为后续工作奠定了一定基础。ALSA为我们团队的另一项工作,该方法的教师信号参与特征决策过程,当特征层神经元数量为1600时网络分类精度为91.6%。对比之下,本文将教师信号引入至特征提取阶段,有效提高网络的特征提取能力,在同样规模的特征层神经元数量下,精度可以达到94.27%。当特征层规模扩大为6400时,特征层规模与Diehl方法一致,而本文方法的精度能够提升2.4%,证明了本文提出方法的优越性。也有学者将STDP与BP机制相结合,通过替代梯度解决脉冲不可微的问题,取得的精度略高于本文提出方法的精度。但BP机制在生物神经网络中缺乏广泛的证据,生物可解释性较低。

表1 部分方法在MNIST数据集上的分类表现

Table 1 Classification performance of some methods on MNIST dataset

模型	学习规则	准确度 (%)
STDP ^[20]	无监督	93.5
STDP ^[12]	无监督	95.0
对称STDP+SVM ^[21]	有监督	96.7
ALSA ^[22]	有监督	91.6
STDP+BP ^[23]	有监督	98.1
BP-STDP ^[24]	有监督	97.2
本文方法	有监督+无监督	97.4

表2列出了部分方法在Fashion-MNIST数据集的分类表现。本文方法优于大多数基于STDP规则的脉冲神经网络模型。GLSNN通过多层网络模型以及BP机制,分类精度稍高于本文方法。从表中可以看出,基于突触可塑性的训练方法在满足生物可解释性的前提下,具有不错的分类表现。

表2 部分方法在Fashion-MNIST数据集上的分类表现

Table 2 Classification performance of some methods on Fashion-MNIST dataset

模型	学习规则	准确度 (%)
FSpiSNN ^[25]	无监督	68.8
A-STDP ^[26]	无监督	75.9
对称STDP+SVM ^[21]	有监督	85.3
VPSNN ^[16]	有监督	83.0
CBSNN ^[27]	有监督	85.7
GLSNN ^[28]	有监督	89.1
本文方法	有监督+无监督	88.8

结束语 本文提出了一种基于多子网络预训练的脉冲神经网络分类方法,其主要包含脉冲编码、单类别特征提取器以及无监督特征融合 3 个部分。首先,图像通过时间编码方法转化为脉冲序列,输入至 SNN 中。接着,教师神经元利用样本标签信息激活对应的单类别特征提取器,由激活子网络通过异触 STDP 学习规则完成脉冲序列的特征提取,最终经过无监督特征融合后,采用统计法计算实验精度。不同于其他基于 STDP 的 SNN 监督学习方法,本文通过调整竞争的范围与模式,有效提高了网络对脉冲序列的特征提取与分类能力,并通过 t-SNE 特征降维方法可视化其效果。本方法在 MNIST 与 Fashion-MNIST 数据集上分别取得了 97.40% 与 88.81% 的准确度,优于大多数同类型方法。后续工作可以通过扩大特征层规模,或加深网络模型层数,进一步提高实验精度。

参考文献

- [1] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
- [2] KALCHBRENNER N, GREFFENSTETTE E, BLUNSON P. A Convolutional Neural Network for Modelling Sentences[C] // Annual meeting of the Association for Computational Linguistics, 2014.
- [3] SILVER D, SCHRITTWIESER J, SIMONYAN K, et al. Mastering the game of go without human knowledge[J]. Nature, 2017, 550(7676): 354-359.
- [4] WANG X Q, ZENG Y H, HAN D M, et al. Brain like computing based on spiking neural networks [J]. Journal of Beijing Institute of Technology, 2019, 45(12): 1277-1286.
- [5] MAASS W. Networks of spiking neurons: the third generation of neural network models[J]. Neural Networks, 1997, 10(9): 1659-1671.
- [6] CAPORALE N, DAN Y. Spike timing-dependent plasticity: a Hebbian learning rule [J]. Annu. Rev. Neurosci., 2008, 31: 25-46.
- [7] LESTER D, FURBER S. Spinnaker: Distributed computer engineering for neuromorphics [M] // Neural Nets WIRN11. IOS Press, 2011: 324-331.
- [8] PEI J, DENG L, SONG S, et al. Towards artificial general intelligence with hybrid Tianjic chip architecture [J]. Nature, 2019, 572(7767): 106-111.
- [9] DAVIES M, SRINIVASA N, LIN T H, et al. Loihi: A neuromorphic manycore processor with on-chip learning [J]. IEEE Micro, 2018, 38(1): 82-99.
- [10] MA D, SHEN J, GU Z, et al. Darwin: A neuromorphic hardware co-processor based on spiking neural networks [J]. Journal of Systems Architecture, 2017, 77: 43-51.
- [11] BOUVIER M, VALENTIAN A, MESQUIDA T, et al. Spiking neural networks hardware implementations and challenges: A survey [J]. ACM Journal on Emerging Technologies in Computing Systems (JETC), 2019, 15(2): 1-35.
- [12] DIEHL P U, COOK M. Unsupervised learning of digit recognition using spike-timing-dependent plasticity [J]. Frontiers in Computational Neuroscience, 2015, 9: 99.
- [13] KOHONEN T. The self-organizing map [J]. Proceedings of the IEEE, 1990, 78(9): 1464-1480.
- [14] TEYLER T J, RUDY J W. The hippocampal indexing theory and episodic memory: updating the index [J]. Hippocampus, 2007, 17(12): 1158-1169.
- [15] PONULAK F. ReSuMe-new supervised learning method for spiking neural networks. Institute of Control and Information Engineering, Poznań University of Technology [R]. Tech. rep, 2005.
- [16] ZHANG T, ZENG Y, ZHAO D, et al. A plasticity-centric approach to train the non-differential spiking neural networks [C] // Proceedings of the AAAI Conference on Artificial Intelligence, 2018.
- [17] KOCH C, SEGEV I. Methods in neuronal modeling: from ions to networks [M]. MIT press, 1998.
- [18] HODGKIN A L, HUXLEY F. A quantitative description of membrane current and its application to conduction and excitation in nerve [J]. The Journal of Physiology, 1952, 117(4): 500.
- [19] KRUG M, MÜLLER-WELDE P, WAGNER M, et al. Functional plasticity in two afferent systems of the granule cells in the rat dentate area: frequency-related changes, long-term potentiation and heterosynaptic depression [J]. Brain Research, 1985, 360(1/2): 264-272.
- [20] QUERLIOZ D, BICHLER O, DOLLFUS P, et al. Immunity to device variations in a spiking neural network with memristive nanodevices [J]. IEEE Transactions on Nanotechnology, 2013, 12(3): 288-295.
- [21] 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.
- [22] MO L, WANG G, LONG E, et al. ALSA: Associative Learning Based Supervised Learning Algorithm for SNN [J]. Frontiers in Neuroscience, 2022, 16: 838832.
- [23] LIU F, ZHAO W, CHEN Y, et al. SSTDP: Supervised spike timing dependent plasticity for efficient spiking neural network training [J]. Frontiers in Neuroscience, 2021, 15: 756876.
- [24] TAVANAIE A, MAIDA A. BP-STDP: Approximating back-propagation using spike timing dependent plasticity [J]. Neurocomputing, 2019, 330: 39-47.
- [25] PUTRA R V W, SHAFIQUE M. Fspinn: An optimization framework for memory-efficient and energy-efficient spiking neural networks [J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2020, 39(11): 3601-3613.
- [26] RASTOGI M, LU S, ISLAM N, et al. On the self-repair role of astrocytes in STDP enabled unsupervised SNNs [J]. Frontiers in Neuroscience, 2021, 14: 603796.
- [27] SHI M, ZHANG T, ZENG Y. A curiosity-based learning method for spiking neural networks [J]. Frontiers in Computational Neuroscience, 2020, 14: 7.
- [28] ZHAO D, ZENG Y, ZHANG T, et al. GLSNN: A multi-layer spiking neural network based on global feedback alignment and local STDP plasticity [J]. Frontiers in Computational Neuroscience, 2020, 14: 576841.



ZHUO Mingsong, born in 1999, post-graduate. His main research interest is spiking neural network learning algorithms and their applications.



MO Lingfei, born in 1981, Ph.D, associate professor, Ph. D supervisor. His main research interests include neuromorphic intelligence and neuromorphic perception, and intelligent perception of the Internet of Things.