

基于DT-LIF神经元与SSD的脉冲神经网络目标检测方法

周雅 栗心怡 武喜艳 赵宇飞 宋勇*

(北京理工大学光电学院 北京 100081)

摘要: 相对于传统人工神经网络(ANN), 脉冲神经网络(SNN)具有生物可解释性、计算效率高等优势。然而, 对于目标检测任务, SNN存在训练难度大、精度低等问题。针对上述问题, 该文提出一种基于动态阈值LIF神经元(DT-LIF)与单镜头多盒检测器(SSD)的SNN目标检测方法。首先, 设计了一种DT-LIF神经元模型, 该模型可根据累积的膜电位动态调整神经元的阈值, 以驱动深层网络的脉冲活动, 提高推理速度。同时, 以DT-LIF神经元为基元, 构建了一种基于SSD的混合SNN。该网络以脉冲视觉几何群网络(Spiking VGG)和脉冲密集连接卷积网络(Spiking DenseNet)为主干(Backbone), 具有由批处理归一化(BN)层、脉冲卷积(SC)层与DT-LIF神经元构成的3个额外层和SSD预测框头(Head)。实验结果表明, 相对于LIF神经网络, DT-LIF神经网络在Prophesee GEN1数据集上的目标检测精度提高了25.2%。对比AsyNet算法, 所提方法的目标检测精度提高了17.9%。

关键词: 计算机视觉; 目标检测; 脉冲神经网络; 神经元

中图分类号: TN911.73; TP391.41

文献标识码: A

文章编号: 1009-5896(2023)08-2722-09

DOI: 10.11999/JEIT221367

Object Detection Method with Spiking Neural Network Based on DT-LIF Neuron and SSD

ZHOU Ya LI Xinyi WU Xiyang ZHAO Yufei SONG Yong

(School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China)

Abstract: Compared with traditional Artificial Neural Network (ANN), the Spiking Neural Network (SNN) has advantages of biological reliability and high computational efficiency. However, for object detection task, SNN has problems such as high training difficulty and low accuracy. In response to the above problems, an object detection method with SNN based on Dynamic Threshold Leaky Integrate-and-Fire (DT-LIF) neuron and Single Shot multibox Detector (SSD) is proposed. First, a DT-LIF neuron is designed, which can dynamically adjust the threshold of neuron according to the cumulative membrane potential to drive spike activity of the deep network and improve the inference speed. Meanwhile, using DT-LIF neuron as primitive, a hybrid SNN based on SSD is constructed. The network uses Spiking Visual Geometry Group (Spiking VGG) and Spiking Densely Connected Convolutional Network (Spiking DenseNet) as the backbone, and combines with SSD prediction head and three additional layers composed of Batch Normalization (BN) layer, Spiking Convolution (SC) layer, and DT-LIF neuron. Experimental results show that compared with LIF neuron network, the object detection accuracy of DT-LIF neuron network on the Prophesee GEN1 dataset is improved by 25.2%. Compared with the AsyNet algorithm, the object detection accuracy of the proposed method is improved by 17.9%.

Key words: Computer vision; Object detection; Spiking Neural Network (SNN); Neuron

1 引言

目标检测是一种以图像分类为基础的视觉任务,

是计算机视觉领域的重要研究方向之一。一般情况下, 目标检测主要解决两个问题: 一是通过绘制边界框精准定位单个或多个目标所在位置(定位), 二是判断目标属于哪种类别(识别)。近年来, 基于深度学习的目标检测算法得到了快速发展, 并逐渐成为主流。此类算法可分为两类: 一类是以区域卷积神经网络(Region-CNN, R-CNN)^[1]、空间金字塔池化网络(Spatial Pyramid Pooling-Net, SPP-Net)^[2]

收稿日期: 2022-11-01; 改回日期: 2023-05-11; 网络出版: 2023-05-20

*通信作者: 宋勇 yongsong@bit.edu.cn

基金项目: 国家自然科学基金(82272130, U22A20103)

Foundation Items: The National Natural Science Foundation of China (82272130, U22A20103)

为代表Two-stage算法,即在特征提取后首先生成一个可能包含待检测目标的候选框,再通过卷积神经网络进行分类,进一步实现目标定位和识别;另一类是以“你只看1次”算法(You Only Look Once, YOLO)^[3]、单镜头多盒检测器(Single Shot multibox Detector, SSD)^[4]为代表的one-stage算法,它利用端对端的方法,仅通过一步即可同时实现目标位置与类别的预测,在推理速度方面具有优势。

另一方面,虽然更深、更复杂的人工神经网络(Artificial Neural Network, ANN)可以实现更高的精度,但其对计算成本的需求也大大增加^[5]。相对而言,脉冲神经网络(Spiking Neural Network, SNN)的神经元之间通过尖峰脉冲的独立、离散电信号相互通信,只有当膜电位达到放电阈值^[6]时,才通过脉冲传递信息。这种推理方法更接近生物神经元的特性,使得SNN比常规ANN具有更小的功耗和更快的推理速度。

目前,已有的SNN主要应用于图像分类等相对简单的任务,对于同时实现定位与分类的目标检测等相对复杂的任务,其精度较低。例如:Kim等人^[7]利用间接有监督学习,同时通过信道归一化为深度SNN传输信息,将YOLO转换成SNN版本的目标检测架构,但这种转换方法一定程度上丢失了SNN低功耗的优势,其性能仍低于标准YOLO。Chakraborty等人^[8]提出了一种基于RetinaNet^[9]的SNN目标检测器,采用混合学习方式,在主干网络中添加了由ANN对应网络转换而来的Spiking网络块,但同样由于这种转换方式会带来性能上的损失,其检测结果较RetinaNet模型仍有差距。Kugele等人^[10]提出了一种混合SNN-ANN的架构用以解决目标检测任务,通过结合反向传播和替代梯度规则进行端到端训练,与类似的ANN网络架构相比精度损失很小,然而该方法使用常规脉冲神经元,存在推理时间较长等问题。

针对上述问题,本文提出一种基于动态阈值LIF神经元(Dynamic Threshold Leaky Integrate-and-Fire, DT-LIF)与SSD的脉冲神经网络目标检测方法。首先,设计了一种动态阈值LIF神经元DT-LIF,该神经元模型可以在每个时间步长上动态调整阈值来增加脉冲活动,以减少延迟;然后,以DT-LIF神经元为基元,构建了一种基于SSD的混合SNN。该网络以脉冲视觉几何群网络(Spiking Visual Geometry Group, Spiking VGG)和脉冲密集连接卷积网络(Spiking Densely connected convolutional Network, Spiking DenseNet)为主干

(Backbone),具有由批处理归一化(Batch Normalization, BN)层、脉冲卷积(Spiking Convolution, SC)层与DT-LIF神经元构成的3个额外层和SSD预测框头(Head)。最终,在无需额外转换(ANN-to-SNN)的前提下,实现基于SNN的高精度目标检测。

2 相关工作

2.1 脉冲神经网络

与生物神经系统类似,SNN采用脉冲序列表示信息,通过发射脉冲在空间域和时间域两个维度传递信息,可模拟信息在人类大脑中的编码和处理过程。在脉冲事件稀疏性的驱动下,SNN兼具计算高效性和生物可解释性,较ANN具有更高的计算效率和更快的推理速度。目前,基于SNN的学习方法主要可分为3类^[11]:

(1)基于脉冲时间依赖可塑性(Spike-Timing Dependent Plasticity, STDP)^[12]的无监督学习。STDP是Hebb^[13]无监督学习算法的一种变体,连接突触前、后神经元的权重根据二者之间的放电延迟而修改。其中,Kheradpisheh等人^[14]提出了一种基于STDP的脉冲深度卷积神经网络,用于目标识别。由于STDP是利用突触可塑性进行局部调整,不足以实现突触个体与神经系统整体目标的协调,因而难以应用于大规模深度网络。

(2)ANN-to-SNN间接有监督学习。该方法先训练一个ANN模型,再利用脉冲神经元平均脉冲发放率和ReLU激活函数之间的关系,在完成原始ANN训练后,再将学习完成的权重迁移到具有相同结构的SNN。其中,Diehl等人^[15]实现了快速ANN-to-SNN转换,采用参数归一化在MNIST数据集上实现了几乎无损的转换。然而,该方法实质上是利用SNN逼近ANN,损失了SNN的时间特性。

(3)直接监督学习。近年来,基于反向传播(Back Propagation, BP)的SNN监督学习在提高网络性能方面取得了较大进展。Neftci等人^[16]提出使用替代梯度学习训练SNN:在正向传播中,使用Heaviside阶跃函数产生脉冲;在反向传播中,采用替代梯度近似不可微函数的梯度(例如s型函数)。这项工作表明SNN能够在通用深度学习框架中通过时间的反向传播进行学习。Wu等人^[17]提出了一种时空反向传播(Spatio-Temporal BackPropagation, STBP)学习算法,同时为了解决SNN中不可微分的问题,提出了适合于梯度下降训练的脉冲活动近似导数。基于该学习算法,Zheng等人^[18]提出了阈值依赖批处理归一化(threshold-dependent Batch Normalization, tdBN)训练方法,用于直接训练高性能的深层SNN。其中,使用替代梯度学习规则,

产生了基于Pytorch的SNN训练框架,例如SpikingJelly^[19]。借助于自动微分和GPU加速,新的框架已经使得SNN获得了更深层的训练,在分类问题上取得了重要进展^[19]。

2.2 脉冲神经元

SNN的基本计算单元是脉冲神经元,LIF神经元模型^[20]目前是SNN中最常见的脉冲神经元之一。LIF神经元可以等效为一个阻容(RC)电路,如图1所示,电路内部并联了一个膜电容 C_m 和膜电阻 R_m 。没有电流输入时,膜电位 $V(t)$ 处于静息电压 V_{reset} 。如果突触前神经元给突触后神经元传来脉冲,则在连接的突触上产生相应的电流 $I(t)$ 。输入电流 $I(t)$ 一部分给 C_m 充电,此时 $V(t)$ 开始累积电压;另一部分则由 R_m 进行电子转移,即泄漏电流。当 C_m 两端的 $V(t)$ 超过神经元阈值电压 V_{th} 时,神经元会触发一个脉冲传给通过突触连接的下一个神经元,随后 C_m 产生放电现象, $V(t)$ 复位到静息电压。

在LIF神经元中引入膜时间参数 $\tau_m = RC_m$,膜电位 $V(t)$ 的变化过程可由式(1)表示

$$\tau_m \frac{dV(t)}{dt} = -(V(t) - V_{\text{reset}}) + R_m I(t) \quad (1)$$

令 $X(t) = R_m I(t)$ 作为外界对神经元的输入,将式(1)改写为

$$\tau_m \frac{dV(t)}{dt} = -(V(t) - V_{\text{reset}}) + X(t) \quad (2)$$

将LIF神经元模型的连续时间离散化到一个时间步长中,可以用式(3)—式(5)分别描述神经元的充电、放电、复位动力学过程

$$H_t = f(V_{t-1}, X_t) = V_{t-1} + \frac{1}{\tau_m} (-(V_{t-1} - V_{\text{reset}}) + X_t) \quad (3)$$

$$S_t = \Theta(H_t - V_{\text{th}}) \quad (4)$$

$$V_t = H_t (1 - S_t) + V_{\text{reset}} \cdot S_t \quad (5)$$

其中, H_t 为 t 时刻神经元充电后的膜电位, V_t 为触发脉冲后神经元的膜电位, X_t 为外部对神经元的输入, V_{th} 为发射脉冲阈值, τ_m 表示膜时间参数。 $\Theta(x)$ 是阶跃函数, S_t 表示 t 时刻的输出脉冲,如果有脉冲产生则等于1,否则为0。

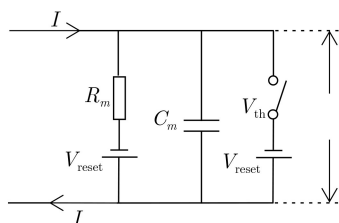


图1 LIF神经元模型等效电路

2.3 SSD框架

SSD结合了YOLO和Faster-RCNN的优点^[21],是近年来广泛使用的one-stage算法之一,在检测精度与速度上都具有一定优势。一方面,SSD一般由一个主干特征提取器和多个预测头组成,它采用多尺度方法提取特征,根据与anchor相似的不同尺度先验框进行初步检测,以提高检测精度。另一方面,SSD采用端到端的训练,借鉴回归模式完成目标的检测与定位,可在保证精度的同时,更快速地完成实时检测^[22]。

SSD的网络结构包括主干网络、额外层(extra layer)、预测头3个部分。其中,主干网络是一个改进的VGG16^[23]网络,利用卷积层替换VGG16最后两个全连接层及输出层。额外层是在基础主干网络之后添加的4个卷积层,以生成更多不同尺度的特征,提高网络的目标特征提取能力。此外,SSD网络还采用了6个卷积特征图,包括Conv4_3(38, 38), Conv7(19, 19), Conv8_2(10, 10), Conv9_2(5, 5), Conv10_2(3, 3)和Conv11_2(1, 1)。不同尺度特征图设置的先验框数目不同,分别为30, 60, 111, 162, 213, 264。

SSD算法流程主要分3部分。首先,将图片输入到预训练好的网络中获得不同大小的特征映射。然后,根据上述6个卷积特征图,从主干网络和额外层中提取特征并构建不同尺度的检测分支,输入到检测层中进行边框回归和分类。最后,采用非最大值抑制(Non-Maximum Suppression, NMS)消除高重叠预测,从而获得最优解。

3 基于DT-LIF与SSD的目标检测算法

3.1 算法结构

基于DT-LIF神经元与SSD的目标检测算法的结构主要由主干网络、3个额外层和SSD预测头等构成。其中,主干网络通过将SSD算法的CNN替换为Spiking VGG/Spiking DenseNet形成;3个额外层均由BN层、 1×1 SC层、 3×3 SC层和所设计的DT-LIF神经元构成(1×1 SC层用以减少通道数量, 3×3 SC层的步长为2);SSD头部主要利用所接收3个额外层输出的特征图进行预测。图2为基于DT-LIF神经元与SSD的目标检测算法的结构。

3.2 DT-LIF神经元模型

已有研究表明生物神经元具有可变的阈值^[24]。Azouz等人^[24]发现,体内皮质神经元的电压阈值不是固定的,脉冲阈值与脉冲发生前的膜去极化速率成反比。除此之外,Fontaine等人^[25]验证阈值可变性是神经元的内在特征之一。然而,目前常规LIF

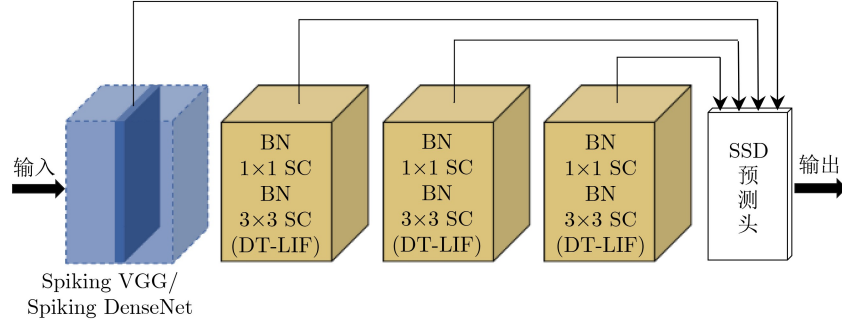


图2 基于DT-LIF神经元与SSD的目标检测算法的结构

神经元模型一般采用固定阈值，需要积累足够的电压才能达到阈值并触发脉冲，从而限制了脉冲神经元的放电速率。当网络层次加深时，处于深层的神经元将很难被驱动，导致收敛速度下降。

针对上述问题，本文设计了一种DT-LIF神经元模型(图3)，以增加脉冲活动，加快收敛速度。

首先，脉冲阈值与膜电位变化率之间的反比例关系的连续时间形式可以用式(6)表示

$$\theta = p + qe^{-\frac{\psi}{\tau}} \quad (6)$$

其中， θ 为脉冲阈值， $\psi = dV_m/dt$ 为膜去极化速率， V_m 为膜电位。同时，由于模拟SNN是一个离散的过程，因此对膜去极化速率的连续时间描述可离散化为

$$\psi = \frac{\Delta V_m}{\Delta t} = \frac{H_t - V_{t-1}}{\Delta t} \quad (7)$$

其中， ΔV_m 为膜电位变化量，相当于式(3)中的 $1/\tau_m \cdot -(V_{t-1} - V_{\text{reset}}) + X_t$ 。

将式(7)代入式(6)中，可得到动态脉冲阈值 θ_t 的离散化描述，如式(8)所示

$$\theta_t = p + qe^{-\frac{\frac{1}{\tau_m} \cdot -(V_{t-1} - V_{\text{reset}}) + X_t}{c}} \quad (8)$$

其中， $p \in [0,1]$ 是电压门控阈值， p, q 满足 $p+q=1$ ， c 为常数。需要说明的是，为了避免过饱和和发射脉冲，应满足 $\theta_t \geq p$ 。

式(8)描述了DT-LIF神经元模型的动态阈值，将其与式(4)中的固定阈值进行比较可以发现，DT-LIF神经元的动态阈值更接近生物神经元内在特征。因此，将式(4)中的 V_{th} 替换为式(8)中的 θ_t ，以实现DT-LIF神经元模型的脉冲产生机制(如图3所示)

$$S_t = \Theta(H_t - \theta_t) \quad (9)$$

从式(8)可以看出，脉冲阈值与膜电位变化量 ΔV_m 成反比。如果 ΔV_m 为负，则增大阈值以防止神经元放电，这与已有研究中膜超极化会降低去极化电位的上升速率^[24]的结论一致；如果 ΔV_m 为正，则降低阈值以增加神经元激发脉冲的数量，这与已有

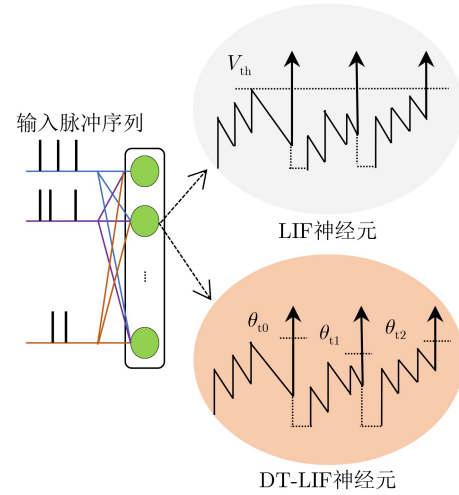


图3 DT-LIF神经元模型示意图

研究中膜去极化会导致较低的脉冲阈值的结论一致^[24]。同时，当输入 $X_t = V_{t-1} - V_{\text{reset}}$ 时，式(8)中的 θ_t 等于式(4)中的 V_{th} ，式(9)等价于式(4)。在这种情况下，DT-LIF模型退化为LIF模型。

算法1给出了DT-LIF神经元发射脉冲的动力学过程。

3.3 主干网络

本文分别构建了基于VGG, DenseNet的Spiking VGG, Spiking DenseNet, 作为所设计的目标检测算法结构的主干网络。同时，设计了一种脉冲分类器作为主干网络的最终层，并先将空间域的所有输出脉冲累加，然后在时间域上得到预测结果。

网络中所使用的神经元基元均为本文所设计的DT-LIF神经元，并采用脉冲卷积(Spiking Convolution, SC)^[26]、最大池化、批处理归一化(Batch Normalization, BN)操作。其中，BN层的参数可以在推理时与前后的卷积层参数融合，使得无论是否接收到脉冲，通过脉冲卷积学习到的权值会通过反向传播来更新，以加快收敛速度，本文将其放置在网络中DT-LIF神经元之前。此外，本文设计的脉冲分类器替代了VGG, DenseNet主干网络中用于分类的最终层，以实现与SNN的兼容性。脉冲分类器

算法1 DT-LIF发射脉冲过程

参数: $\theta, p, q, V_{th}, \tau_m$

- (1) $\theta = V_{th} = 1; V = 0; V_{reset} = 0$ // 初始化
- (2) **for** $t = 1$ **to** timesteps **do**
- (3) **for** $l = 2$ **to** L **do**
- (4) **for** $i = 1$ **to** neurons **do**
- (5) $H_{i,t}^l = V_{i,t-1}^l + (X_{i,t}^l - (V_{i,t-1}^l - V_{reset})) * \tau_m$ // $X_{i,t}^l$ 是正向传递的输入
- (6) $\delta = H_{i,t}^l - V_{i,t-1}^l$
- (7) $\theta_{i,t}^l = p + q \exp(-\delta / c)$
- (8) **if** $H_{i,t}^l \geq \theta_{i,t}^l$ **then**
- (9) $S_{i,t}^l = 1$
- (10) $V_{i,t}^l = V_{reset}$
- (11) **end for**
- (12) **end for**
- (13) **end for**

由BN层、DT-LIF神经元、输出 1×1 的num_classes通道组成。其中，BN层将脉冲的高度稀疏特征映射转换为密集的十进制表示。 1×1 卷积输出可以保证该分类器能处理所有大小的特征图，而无需使用平均池化层。同时，DT-LIF神经元可以取代ReLU激活功能。

本文构建的两种主干网络具体描述如下：

(1)Spiking VGG: VGG^[23]是一个卷积神经网络。本文利用脉冲分类器取代VGG最后的分类层，其余架构与原网络保持一致，图4为Spiking VGG网络结构图(以VGG11为例)。该网络包含5个Spiking块与1个脉冲分类器，每个Spiking块均由BN层、SC层、DT-LIF神经元组成，SC层全部使用 3×3 的卷积核。每一块的SC层特征图大小分别为：SC1(224, 224), SC2(112, 112), SC3(56, 56), SC4(28, 28), SC5(14, 14)。

(2)Spiking DenseNet: 为了促进训练中梯度的反向传播，ResNet^[27]模型建立了前后层之间的“短路连接”，连接机制为元素级相加。DenseNet^[28]

的连接机制相较ResNet更密集，是一种通过使用通道级连接来促进梯度传播的架构。Spiking DenseNet的网络结构主要由Spiking Dense块和Transition层组成。Spiking Dense块中每层的特征图大小一致，每个Spiking Dense块均包含数个由 $[1 \times 1$ SC, 3×3 SC]组合的Bottleneck，每个SC层前均添加了BN层与DT-LIF神经元。Transition层主要是连接两个相邻的Spiking Dense块，并降低特征图尺寸，结构为BN层、 1×1 SC层、DT-LIF神经元和最大池化层的组合。图5为Spiking DenseNet网络结构图(以DenseNet121为例)，该网络包含4个Spiking Dense块，每一块的特征图大小分别为：(56, 56), (28, 28), (14, 14), (7, 7)。

4 实验及结果分析

4.1 实验设置

本文的实验在CentOS系统下进行，GPU为NVIDIA RTX 2080，软件环境为基于PyTorch 1.11的SpikingJelly深度学习框架。训练时初始学习率设置为 $3e-4$ ，batchsize设为8，在Prophesee GEN1数据集上训练了50个epoch，使用余弦退火的学习率下降方式(cosine annealing learning rate scheduler)将学习率逐渐降低到0。实验中所有SC操作均采用Kaiming统一方法初始化，所有BN层的初始化权重为1，偏差为0。DT-LIF神经元的膜时间参数 τ_m 均为2，固定阈值 V_{th} 为1， $p=q=0.5$ ， $c=20$ ，采用atan作为替代函数。

本文在Prophesee GEN1检测数据集^[29]上评估了所提出的脉冲神经网络目标检测方法。该数据集由一个空间分辨率为 304×240 的Prophesee GEN1传感器记录，共有39 h的城市、高速公路等场景，并将其剪切为数个60 s录制文件，包含超过255 000个手动标注的汽车、行人边界框。数据包含 (x, y, t, p) 信息，分别为事件的2维像素坐标、时间戳与极性，图6为部分典型数据集示例。本文将数据集按照3:1的比例分为训练集和验证集。

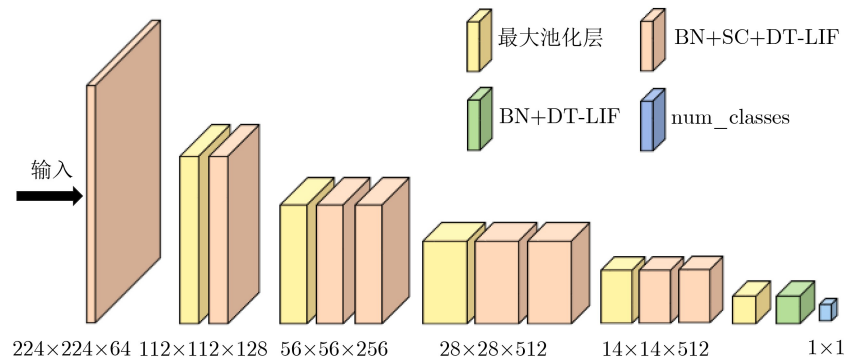


图4 Spiking VGG网络结构图(以VGG11为例)

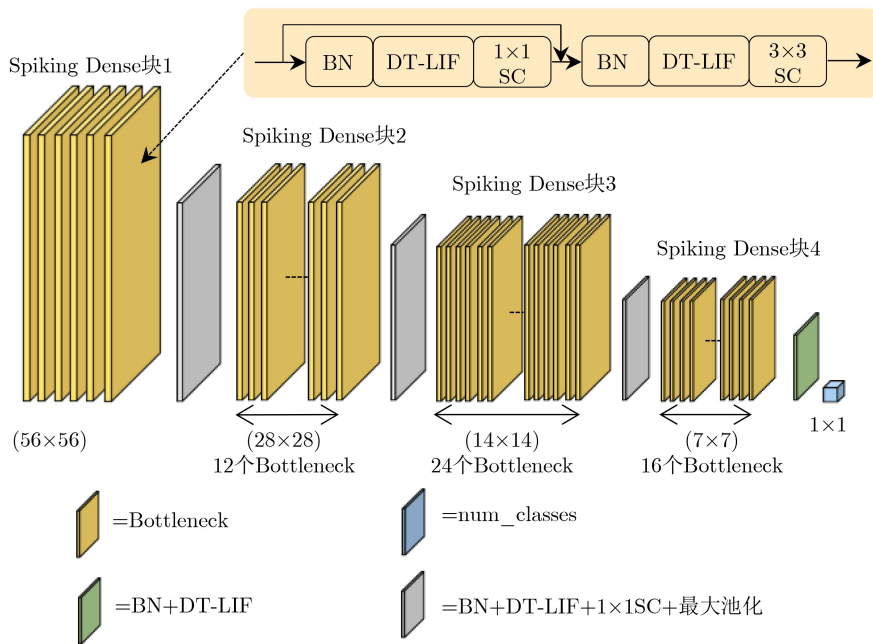
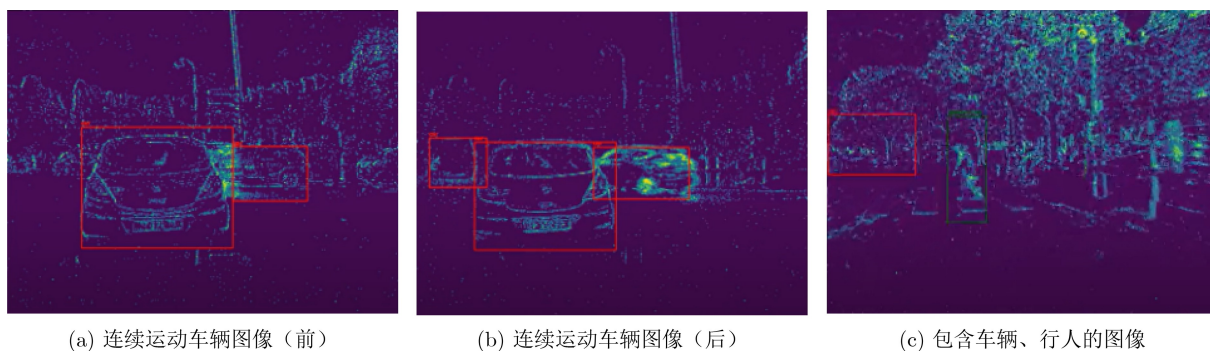


图 5 Spiking DenseNet网络结构图(以DenseNet121为例)



(a) 连续运动车辆图像 (前) (b) 连续运动车辆图像 (后) (c) 包含车辆、行人的图像

图 6 Prophesee GEN1数据集示例

4.2 评价指标

本文采用平均精度均值(mean Average Precision, mAP)作为性能评价指标。mAP是所有类别P-R曲线下面积的平均值，即计算mAP前需先得到准确率(Precision, P)和召回率(Recall, R)。

以预测框与真实框的交并比(Intersection of Union, IoU)的阈值作为判断是否检测成功的指标，若IoU大于0.5则认为是正确检测出的正样本，反之则认为是检测错误的负样本^[30]。本文使用的具体指标为mAP(0.5:0.95)，即IoU阈值等于0.5时求出的mAP与IoU设为0.95时求出的mAP之比。

4.3 实验结果

首先，为了评价本文设计的DT-LIF神经元的性能，分别以Spiking VGG11和Spiking DenseNet121为主干网络，进行了LIF神经元和DT-LIF神经元的对比实验。其次，为了验证本文提出的基于DT-LIF与SSD的目标检测算法的性能，开展了

其与AsyNet^[31]目标检测算法的对比实验。实验结果如表1所示。

如表1所示，以Spiking VGG11为主干网络时，相对LIF神经元模型，采用DT-LIF神经元模型的目标检测算法mAP提高了25.2%。类似地，以Spiking DenseNet121为主干网络时，采用DT-LIF神经元模型的目标检测算法mAP相对提高了11.5%，表明所设计的DT-LIF神经元模型可有效提高目标检测网络的检测精度。

表 1 Prophesee GEN1数据集上的对比实验结果

方法	mAP(0.5:0.95)
Spiking VGG11+LIF	0.127
Spiking VGG11+DT-LIF	0.159
Spiking DenseNet+LIF	0.148
Spiking DenseNet+DT-LIF	0.165
AsyNet ^[31]	0.140

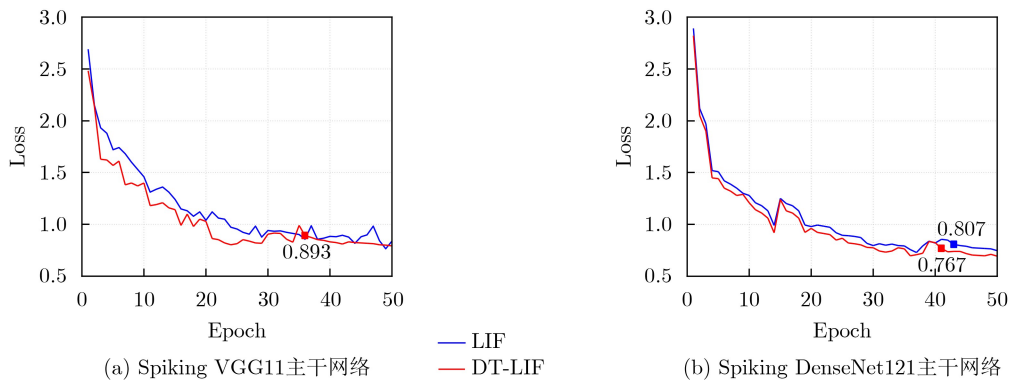


图7 训练损失(Loss)曲线图

同时, 本文对比了训练时每一轮的损失(Loss), 训练损失曲线图如图7所示。可以看出, 以Spiking VGG11为主干网络时, 第3轮起采用不同神经元模型的训练损失差距变得明显。整个训练过程中LIF神经网络的损失一直呈波动趋势, 而第36轮时DT-LIF神经网络的训练损失已下降至0.893, 之后保持稳定。类似地, 以Spiking DenseNet121为主干网络时, 第4轮起采用不同神经元模型的训练损失差距开始变得明显。第43轮时LIF神经网络的损失下降至0.807, 第41轮时DT-LIF神经网络的训练损失已下降至0.767, 之后均趋于平稳。上述结果表明, 相较于基于LIF神经元的目标检测网络, 由于DT-LIF神经元具有阈值可变性, 基于DT-LIF神经元的网络训练损失更早收敛到稳定的数值, 推理速度更快。

另一方面, 由表1可以看出, 相对AsyNet目标检测算法, 本文提出的基于DT-LIF与SSD的目标检测算法mAP提高了17.9%, 表明本文提出的算法具有更强的处理真实事件数据特征的能力, 有效提高了目标检测精度。

5 结论

本文提出一种基于DT-LIF神经元与SSD框架的SNN目标检测方法。首先, 设计了一种DT-LIF神经元模型。同时, 以DT-LIF神经元为基元, 构建了以Spiking VGG和Spiking DenseNet为主干(Backbone), 包括3个额外层和SSD预测框头(Head)的混合SNN。在Prophesee GEN1检测数据集上的实验结果表明, 相对于采用LIF神经元模型, 所设计的DT-LIF神经元模型的mAP提高了25.2%, 且损失收敛速度更快, 与AsyNet算法对比, 所提出的目标检测算法的mAP提高了17.9%, 从而为解决目标检测任务中SNN存在的训练难度大、精度低等问题, 实现基于SNN的高精度目标检测奠定了基础。

参考文献

- [1] GIRSHICK R, DONAHUE J, DARRELL T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation[C]. 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, USA, 2014: 580–587. doi: [10.1109/CVPR.2014.81](https://doi.org/10.1109/CVPR.2014.81).
- [2] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2015, 37(9): 1904–1916. doi: [10.1109/TPAMI.2015.2389824](https://doi.org/10.1109/TPAMI.2015.2389824).
- [3] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: Unified, real-time object detection[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, USA, 2016: 779–788. doi: [10.1109/CVPR.2016.91](https://doi.org/10.1109/CVPR.2016.91).
- [4] LIU Wei, ANGUELOV D, ERHAN D, *et al.* SSD: Single shot MultiBox detector[C]. The 14th European Conference on Computer Vision, Amsterdam, The Netherlands, 2016: 21–37. doi: [10.1007/978-3-319-46448-0_2](https://doi.org/10.1007/978-3-319-46448-0_2).
- [5] TAN Mingxing and LE Q. EfficientNet: Rethinking model scaling for convolutional neural networks[C]. The 36th International Conference on Machine Learning, Long Beach, USA, 2019: 6105–6114.
- [6] GERSTNER W and KISTLER W M. Spiking Neuron Models: Single Neurons, Populations, Plasticity[M]. Cambridge: Cambridge University Press, 2002: 421–454. doi: [10.1017/CBO9780511815706](https://doi.org/10.1017/CBO9780511815706).
- [7] KIM S, PARK S, NA B, *et al.* Spiking-YOLO: Spiking neural network for energy-efficient object detection[C]. The 34th AAAI Conference on Artificial Intelligence, New York, USA, 2020: 11270–11277. doi: [10.1609/AAAI.v34i07.6787](https://doi.org/10.1609/AAAI.v34i07.6787).
- [8] CHAKRABORTY B, SHE Xueyuan, and MUKHOPADHYAY S. A fully spiking hybrid neural network for energy-efficient object detection[J]. *IEEE Transactions on Image Processing*, 2021, 30: 9014–9029. doi: [10.1109/TIP.2021.3122092](https://doi.org/10.1109/TIP.2021.3122092).

- [9] LIN T Y, GOYAL P, GIRSHICK R, *et al.* Focal loss for dense object detection[C]. 2017 IEEE International Conference on Computer Vision, Venice, Italy, 2017: 2999–3007. doi: [10.1109/ICCV.2017.324](https://doi.org/10.1109/ICCV.2017.324).
- [10] KUGELE A, PFEIL T, PFEIFFER M, *et al.* Hybrid SNN-ANN: Energy-efficient classification and object detection for event-based vision[C]. 43rd DAGM German Conference on Pattern Recognition, Bonn, Germany, 2022: 297–312. doi: [10.1007/978-3-030-92659-5_19](https://doi.org/10.1007/978-3-030-92659-5_19).
- [11] 胡一凡, 李国齐, 吴郁杰, 等. 脉冲神经网络研究进展综述[J]. 控制与决策, 2021, 36(1): 1–26. doi: [10.13195/j.kzyjc.2020.1006](https://doi.org/10.13195/j.kzyjc.2020.1006).
- HU Yifan, LI Guoqi, WU Yujie, *et al.* Spiking neural networks: A survey on recent advances and new directions[J]. *Control and Decision*, 2021, 36(1): 1–26. doi: [10.13195/j.kzyjc.2020.1006](https://doi.org/10.13195/j.kzyjc.2020.1006).
- [12] TOYOIZUMI T, PFISTER J P, AIHARA K, *et al.* Spike-timing dependent plasticity and mutual information maximization for a spiking neuron model[C]. The 17th International Conference on Neural Information Processing Systems, Vancouver, Canada, 2004: 1409–1416.
- [13] HEBB D O. The Organization of Behavior: A Neuropsychological Theory[M]. New York: Psychology Press, 2002. doi: [10.4324/9781410612403](https://doi.org/10.4324/9781410612403).
- [14] KHERADPISHEH S R, GANJTABESH M, THORPE S J, *et al.* STDP-based spiking deep convolutional neural networks for object recognition[J]. *Neural Networks*, 2018, 99: 56–67. doi: [10.1016/j.neunet.2017.12.005](https://doi.org/10.1016/j.neunet.2017.12.005).
- [15] DIEHL P U, NEIL D, BINAS J, *et al.* Fast-classifying, high-accuracy spiking deep networks through weight and threshold balancing[C]. 2015 International Joint Conference on Neural Networks (IJCNN), Killarney, Ireland, 2015: 1–8. doi: [10.1109/IJCNN.2015.7280696](https://doi.org/10.1109/IJCNN.2015.7280696).
- [16] NEFTCI E O, MOSTAFA H, and ZENKE F. Surrogate gradient learning in spiking neural networks: Bringing the power of gradient-based optimization to spiking neural networks[J]. *IEEE Signal Processing Magazine*, 2019, 36(6): 51–63. doi: [10.1109/msp.2019.2931595](https://doi.org/10.1109/msp.2019.2931595).
- [17] WU Yujie, DENG Lei, LI Guoqi, *et al.* Spatio-temporal backpropagation for training high-performance spiking neural networks[J]. *Frontiers in Neuroscience*, 2018, 12: 331. doi: [10.3389/fnins.2018.00331](https://doi.org/10.3389/fnins.2018.00331).
- [18] ZHENG Hanle, WU Yujie, DENG Lei, *et al.* Going deeper with directly-trained larger spiking neural networks[C]. The 35th AAAI Conference on Artificial Intelligence, Palo Alto, USA, 2021: 11062–11070. doi: [10.1609/aaai.v35i12.17320](https://doi.org/10.1609/aaai.v35i12.17320).
- [19] FANG Wei, YU Zhaofei, CHEN Yanqi, *et al.* Incorporating learnable membrane time constant to enhance learning of spiking neural networks[C]. 2021 IEEE/CVF International Conference on Computer Vision, Montreal, Canada, 2021: 2641–2651. doi: [10.1109/ICCV48922.2021.00266](https://doi.org/10.1109/ICCV48922.2021.00266).
- [20] GERSTNER W, KISTLER W M, NAUD R, *et al.* Neuronal Dynamics: From Single Neurons to Networks and Models of Cognition[M]. Cambridge: Cambridge University Press, 2014. doi: [10.1017/CBO9781107447615](https://doi.org/10.1017/CBO9781107447615).
- [21] 贺丰收, 何友, 刘准钊, 等. 卷积神经网络在雷达自动目标识别中的研究进展[J]. 电子与信息学报, 2020, 42(1): 119–131. doi: [10.11999/JEIT180899](https://doi.org/10.11999/JEIT180899).
- HE Fengshou, HE You, LIU Zhunga, *et al.* Research and development on applications of convolutional neural networks of radar automatic target recognition[J]. *Journal of Electronics & Information Technology*, 2020, 42(1): 119–131. doi: [10.11999/JEIT180899](https://doi.org/10.11999/JEIT180899).
- [22] 董小伟, 韩悦, 张正, 等. 基于多尺度加权特征融合网络的地铁行人目标检测算法[J]. 电子与信息学报, 2021, 43(7): 2113–2120. doi: [10.11999/JEIT200450](https://doi.org/10.11999/JEIT200450).
- DONG Xiaowei, HAN Yue, ZHANG Zheng, *et al.* Metro pedestrian detection algorithm based on multi-scale weighted feature fusion network[J]. *Journal of Electronics & Information Technology*, 2021, 43(7): 2113–2120. doi: [10.11999/JEIT200450](https://doi.org/10.11999/JEIT200450).
- [23] SIMONYAN K and ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C]. 3rd International Conference on Learning Representations, San Diego, USA, 2015. doi: [10.48550/arXiv.1409.1556](https://doi.org/10.48550/arXiv.1409.1556).
- [24] AZOUZ R and GRAY C M. Dynamic spike threshold reveals a mechanism for synaptic coincidence detection in cortical neurons *in vivo*[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2000, 97(14): 8110–8115. doi: [10.1073/PNAS.130200797](https://doi.org/10.1073/PNAS.130200797).
- [25] FONTAINE B, PEÑA J L, and BRETTE R. Spike-threshold adaptation predicted by membrane potential dynamics *in vivo*[J]. *PLoS Computational Biology*, 2014, 10(4): e1003560. doi: [10.1371/journal.PCBI.1003560](https://doi.org/10.1371/journal.PCBI.1003560).
- [26] XIAO Rong, TANG Huajin, MA Yuhao, *et al.* An event-driven categorization model for AER image sensors using multispikes encoding and learning[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31(9): 3649–3657. doi: [10.1109/tnnls.2019.2945630](https://doi.org/10.1109/tnnls.2019.2945630).
- [27] FANG Wei, YU Zhaofei, CHEN Yanqi, *et al.* Deep residual learning in spiking neural networks[C/OL]. The 34th International Conference on Neural Information Processing Systems, 2021: 21056–21069.
- [28] HUANG Gao, LIU Zhuang, VAN DER MAATEN L, *et al.* Densely connected convolutional networks[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA, 2017: 2261–2269. doi: [10.1109/CVPR.2017.243](https://doi.org/10.1109/CVPR.2017.243).

- [29] DE TOURNEMIRE P, NITTI D, PEROT E, *et al.* A large scale event-based detection dataset for automotive[EB/OL]. <https://doi.org/10.48550/arXiv.2001.08499>, 2020.
- [30] 张德祥, 王俊, 袁培成. 基于注意力机制的多尺度全场景监控目标检测方法[J]. 电子与信息学报, 2022, 44(9): 3249–3257. doi: [10.11999/JEIT210664](https://doi.org/10.11999/JEIT210664).
ZHANG Dexiang, WANG Jun, and YUAN Peicheng. Object detection method for multi-scale full-scene surveillance based on attention mechanism[J]. *Journal of Electronics & Information Technology*, 2022, 44(9): 3249–3257. doi: [10.11999/JEIT210664](https://doi.org/10.11999/JEIT210664).
- [31] MESSIKOMMER N, GEHRIG D, LOQUERCIO A, *et al.* Event-based asynchronous sparse convolutional networks[C]. 16th European Conference on Computer Vision, Glasgow, UK, 2020: 415–431. doi: [10.1007/978-3-030-58598-3_25](https://doi.org/10.1007/978-3-030-58598-3_25).
- 周 雅: 女, 副教授, 研究方向为智能光电信息处理.
栗心怡: 女, 硕士生, 研究方向为类脑计算.
武喜艳: 女, 博士生, 研究方向为脉冲神经网络及其应用.
赵宇飞: 男, 博士后, 研究方向为面向计算机视觉的类脑计算.
宋 勇: 男, 教授, 研究方向为类脑计算、智能交互等.
- 责任编辑: 余 蓉