

Spiking-YOLO: Spiking Neural Network for Energy-Efficient Object Detection 阅读笔记

一、研究背景与意义

1.1 深度学习的算力与功耗困境

过去十年间，深度神经网络（DNNs）凭借强大的特征学习与复杂问题建模能力，在计算机视觉、自然语言处理等领域取得突破性进展。但随着任务复杂度提升（如高精度目标检测、图像分割），DNNs 的模型层数与参数量急剧增加，导致其对计算资源和功耗的需求呈指数级增长，在嵌入式设备、移动终端等边缘平台的部署面临巨大挑战。

为缓解这一问题，研究者提出模型剪枝、参数压缩、量化等优化手段，虽能降低部分计算开销，却未从根本上解决“高算力依赖”的核心矛盾。有研究表明，随着 DNN 模型深度与复杂度提升，精度、算力与功耗的矛盾将进一步加剧，亟需探索新的神经网络架构以实现“高能效比”突破。

1.2 脉冲神经网络（SNNs）的优势与局限

脉冲神经网络（SNNs）作为第三代神经网络，通过模拟生物神经元的“脉冲发放”机制，采用事件驱动（event-driven）计算模式，为解决 DNNs 功耗问题提供新方向。与传统 DNNs 用连续实数值传递信息不同，SNNs 依靠离散的脉冲序列（spike trains）在神经元间传递信息，仅当神经元膜电位达到阈值时才产生脉冲，具备天然的稀疏性与低功耗特性。

但 SNNs 的发展长期受限于两大瓶颈：

- **训练难度大：**SNNs 中神经元的复杂动态特性（如膜电位积分、不应期）及脉冲发放的非可微操作，导致传统反向传播算法难以直接应用。现有训练方法中，基于脉冲时序依赖可塑性（STDP）的无监督学习精度较低，而基于可微近似的监督学习仅能在浅层网络和简单任务（如 MNIST 图像分类）上实现较好性能。
- **应用场景局限：**截至本文发表前（2019 年），SNNs 的研究集中于图像分类等简单任务，且多基于 MNIST、CIFAR 等小数据集，尚未在目标检测等复杂回归任务中有效应用。目标检测需同时完成“目标分类”与“边界框坐标回归”，对输出数值精度要求极高，而 SNNs 中脉冲信息的离散性易导致回归精度损失，成为应用拓展的关键障碍。

1.3 本文的研究定位与价值

针对上述问题，本文首次将 SNNs 应用于目标检测任务，提出“Spiking-YOLO”脉冲神经网络模型。通过深入分析 SNNs 在回归任务中性能退化的原因，设计两种关键优化方法——通道归一化（channel-wise normalization）与非平衡阈值符号神经元（signed neuron with imbalanced threshold），最终实现与传统 DNN 模型（Tiny YOLO）相当的检测精度，同时大幅降低能量消耗。

本文的研究价值体现在三方面：

- **突破任务局限：**首次验证 SNNs 在复杂回归任务（目标检测）中的可行性，为 SNNs 应用拓展提供重要参考；
- **解决核心技术瓶颈：**提出的优化方法有效缓解 SNNs 中信息传递效率低、激活函数适配难等问题，为深层 SNNs 设计提供新范式；
- **实现能效优势：**在神经形态芯片（如 TrueNorth）上的实验表明，Spiking-YOLO 的能耗仅为 Tiny YOLO（GPU 运行）的 1/280，为边缘设备上的高能效目标检测提供可行方案。

二、核心问题分析：SNNs 在目标检测中的性能退化原因

为设计有效的优化策略，本文首先深入剖析传统 DNN-to-SNN 转换方法在目标检测任务中性能退化的两大核心原因：

2.1 传统归一化方法导致的激活不足

在 SNNs 中，神经元的脉冲发放率（firing rate）直接决定信息传递效率。发放率定义为“ T 个时间步内神经元产生的脉冲数 N 与 T 的比值 (N/T)”，最大值为 100%（每个时间步均产生脉冲）。若发放率过低（激活不足），需更多时间步传递完整信息；若发放率过高（激活过度），脉冲信号易饱和导致信息丢失。

传统的层归一化（layer-wise normalization）通过“层内最大激活值”对权重归一化，假设同一层所有通道的激活分布相似。但本文实验发现，Tiny YOLO 的 8 个卷积层中，不同通道的激活值差异极大（如图 1 所示）：以 Conv1 层为例，部分通道（如 6、7、14 号通道）的归一化激活值接近 1，而部分通道（如 1、2、3 号通道）的归一化激活值接近 0。这种极端差异导致：

- 低激活通道的神经元发放率极低（如归一化激活值为 0.007 时，需 1000 个时间步才能传递 7 个脉冲）；
- 目标检测中的边界框回归对数值精度要求极高（如坐标差值为 0.001），低发放率导致信息在有限时间步内丢失，最终引发回归精度严重下降。

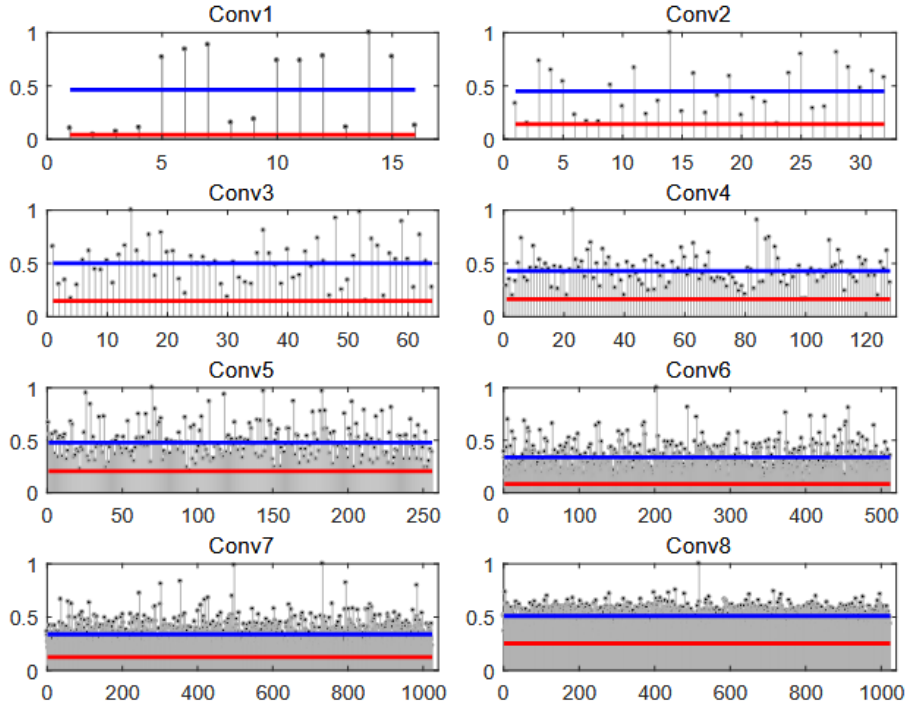


图 1 通过在每个通道中进行逐层归一化，Tiny YOLO 的八层卷积层的最大激活值被归一化。
蓝线和红线分别表示归一化激活值的平均值和最小值

2.2 leaky-ReLU 激活函数的 SNN 适配缺失

Tiny YOLO 采用 leaky-ReLU 作为激活函数，表达式为：

$$f(x) = \begin{cases} x & x \geq 0 \\ \alpha x & x < 0 \end{cases}$$

其中 α 为泄漏系数（通常取 0.01），用于保留负激活信息，提升模型对复杂特征的表达能力。但现有 DNN-to-SNN 转换方法存在两大适配问题：

- **负激活处理缺失：**多数方法仅考虑正激活区域，直接丢弃负激活，导致 leaky-ReLU 的负区域信息丢失；
- **泄漏系数难以实现：**即使部分方法引入负阈值处理负激活，也无法通过离散脉冲模拟泄漏系数 α 的连续缩放效果——若直接对负激活乘以 α ，需额外浮点运算单元，与 SNNs 低功耗设计理念相悖，且难以在神经形态芯片上部署。

三、核心方法设计

3.1 通道归一化（Channel-wise Normalization）

为解决传统层归一化中通道激活不均衡的问题，本文提出“通道归一化”方法，通过更精细的归一化粒度，确保每个通道的神经元在有限时间步内实现高效信息传递。

3.1.1 方法原理

通道归一化的核心思想是：以“通道”为单位而非“层”为单位进行归一化，利用每个通道的 99.9 百分位激活值（而非层内最大激活值）对权重和偏置缩放，避免部分通道因激活值过小导致的发放率不足。

具体计算公式如下：

$$\text{权重归一化: } \bar{w}_{i,j}^l = w_{i,j}^l \cdot \frac{\lambda_i^{l-1}}{\lambda_j^l}$$

$$\text{偏置归一化: } \bar{b}_j^l = \frac{b_j^l}{\lambda_j^l}$$

其中， $w_{i,j}^l$ 和 b_j^l 分别为第 l 层第 j 个输出通道与第 i 个输入通道间的权重和偏置； λ_j^l 为第 l 层第 j 个通道在训练集上的 99.9 百分位激活值； λ_i^{l-1} 为第 $l-1$ 层第 i 个通道的 99.9 百分位激活值，用于补偿前一层归一化对激活范围的缩放，确保层间信息传递一致性。

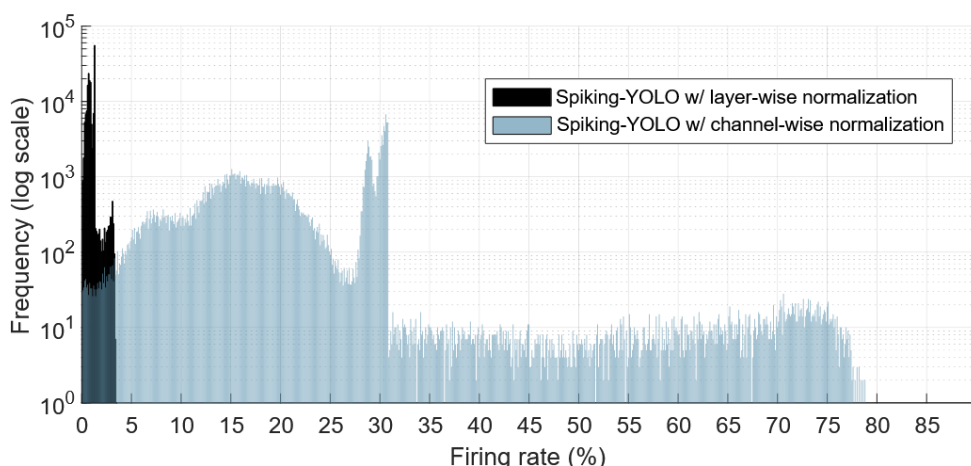


图 2 第 2 通道上层归一化和通道归一化的激活率分布（Tiny YOLO 的 Conv1 层）

3.1.2 方法优势验证

通过对比实验，本文验证了通道归一化在提升神经元发放率方面的显著效果：

- **发放率分布**：如图 2 所示，在 Conv1 层的 2 号通道中，通道归一化下大量神经元的发放率可达 80%，而层归一化下多数神经元的发放率集中在 0%~3.5%；
- **通道间一致性**：如图 3 所示，通道归一化使 Conv1 层 16 个通道的发放率均保持在较高水平，其中 2 号通道的发放率是层归一化的 20 倍；
- **信息传递效率**：通道归一化下，神经元的脉冲发放更规律（如图 4 右侧 raster 图所示），可在 3500 个时间步内达到层归一化 8000 个时间步的检测精度，收敛速度提升 2.3~4 倍。

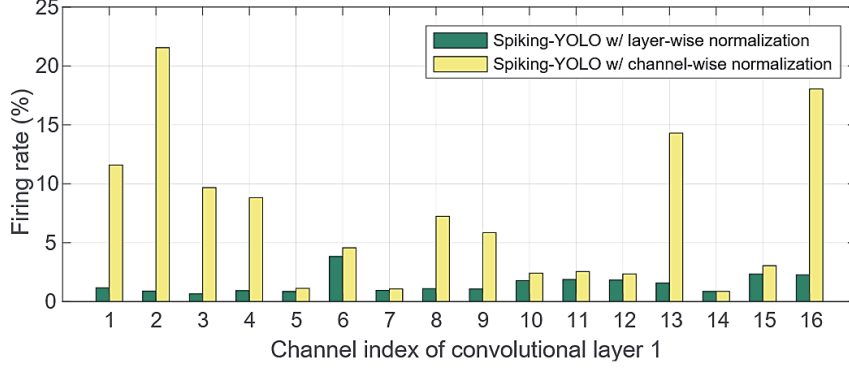


图 3 Conv1 层中 Tiny YOLO 的 layernorm 和 channel-norm 的 16 个通道的激活率

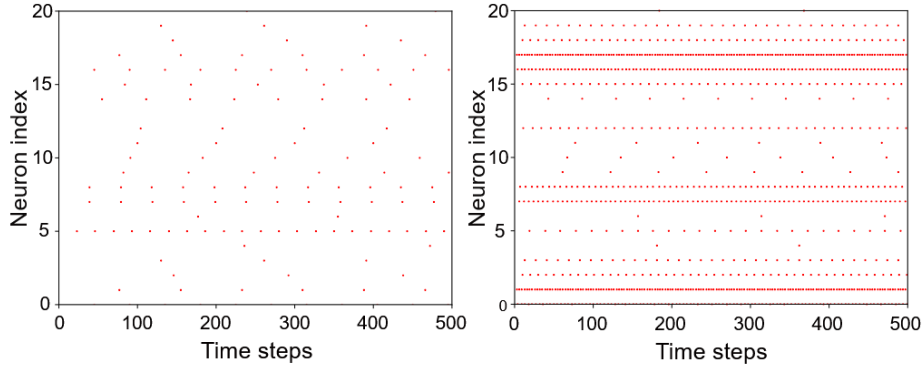


图 4 20 个采样神经元的光栅图；层归一化（左）与通道归一化（右）

3.2 非平衡阈值符号神经元（Signed Neuron with Imbalanced Threshold）

为实现 leaky-ReLU 激活函数在 SNNs 中的高效适配，本文提出“非平衡阈值符号神经元”，通过设计正负区域的非对称阈值，同时保留负激活信息与模拟泄漏系数 α 的效果，且无需额外浮点运算。

3.2.1 方法原理

该神经元的核心设计包括两点：

1. **符号脉冲发放**：神经元可产生“+1”（正脉冲）、“-1”（负脉冲）或“0”（无脉冲）三种状态，分别对应 leaky-ReLU 中的正激活、负激活与零激活区域；
2. **非平衡阈值设置**：通过调整正负区域的阈值，模拟泄漏系数 α 的缩放效果。具体而言：

$$\text{正阈值 } V_{th,pos} = V_{th} \text{（与传统 SNN 一致）}$$

$$\text{负阈值 } V_{th,neg} = -\frac{1}{\alpha} V_{th} \text{（绝对值为正阈值的 } 1/\alpha \text{ 倍）}$$

神经元的脉冲发放规则如下：

$$f(x) = \begin{cases} 1 & V_{mem} \geq V_{th,pos} \\ -1 & V_{mem} \leq V_{th,neg} \\ 0 & otherwise \end{cases}$$

其中， V_{mem} 为神经元的膜电位。以 $\alpha=0.01$ 为例， $V_{th,neg} = -100 V_{th}$ ，即负激活需积累 100 倍于正激活的膜电位才能产生负脉冲，恰好模拟 leaky-ReLU 中负激活被缩放 100 倍的效果 ($\alpha=0.01$)。

3.2.2 方法优势

该设计的优势体现在三方面：

- **生物合理性**：符号脉冲对应生物神经元中的兴奋性 (+1) 与抑制性 (-1) 信号，更符合大脑神经元的信息传递机制；
- **硬件友好性**：无需额外浮点乘法单元，仅通过调整阈值即可实现 leaky-ReLU 功能，可直接映射到现有神经形态芯片；
- **精度提升**：实验表明，引入该神经元后，Spiking-YOLO 在 PASCAL VOC 上的 mAP 从 7.31% 提升至 51.83%，接近 Tiny YOLO 的 53.01%，验证其对负激活信息的有效保留。

四、实验设计与结果分析

4.1 实验设置

4.1.1 模型与数据集

- **基础模型**：以 Tiny YOL 为 DNN 基准模型，包含 8 个卷积层、2 个全连接层，在 PASCAL VOC 2007 上的 mAP 为 53.01%，MS COCO 上为 26.24%；
- **数据集**：采用 PASCAL VOC 2007 (20 类目标，约 9k 训练图、3k 测试图) 和 MS COCO 2014 (80 类目标，约 80k 训练图、40k 验证图)，均为目标检测领域的标准数据集；
- **SNN 组件适配**：根据 Rueckauer 等 (2017) 的方法，在 SNN 中实现最大池化 (spike max-pooling) 和批归一化 (batch normalization)。

4.1.2 实验环境与参数

- **仿真框架**：基于 TensorFlow Eager 构建 SNN 仿真环境；
- **硬件平台**：训练与仿真在 NVIDIA Tesla V100 GPU 上进行，能耗对比实验涉及 GPU (Titan V100) 与神经形态芯片 (TrueNorth)；

- **关键参数：**时间步范围为 1000~8000，神经元阈值 $V_{th}=1$ ，leaky-ReLU 泄漏系数 $\alpha=0.01$ ，通道归一化采用 99.9 百分位激活值。

4.1.3 评价指标

- **检测精度：**采用平均精度均值（mAP，mean Average Precision），为目标检测的标准指标；
- **收敛速度：**达到最大 mAP 所需的时间步数；
- **能量消耗：**通过“操作数 \times 单位操作能耗”计算，其中 32 位浮点（FL）MAC 操作能耗为 4.6 pJ，AC 操作能耗为 0.9 pJ；32 位整数（INT）MAC 操作能耗为 3.2 pJ，AC 操作能耗为 0.1 pJ。

4.2 检测精度结果分析

4.2.1 不同配置下的 mAP 对比

本文通过控制变量法，验证“通道归一化”和“非平衡阈值符号神经元（IBT）”对检测精度的影响，结果如表 1 和图 5 所示：

表 1: Spiking-YOLO 的实验结果（mAP %）

Signed neuron	Norm. method	PASCAL VOC (53.01) ^a		MS COCO (26.24) ^a	
		V_{mem}	Spike count	V_{mem}	Spike count
w/out IBT	Layer	3.86	6.87	0.74	2.82
	Channel	5.61	7.31	0.74	3.02
w/ IBT	Layer	48.94	46.29	24.66	20.93
	Channel	51.83	47.19	25.66	21.54

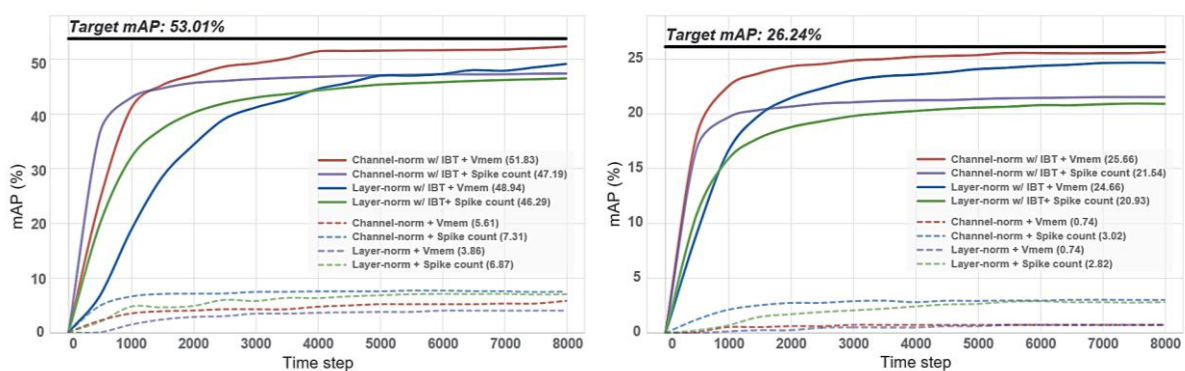


图 5 Spiking-YOLO 在 PASCAL VOC（左）和 MS COCO（右）上的实验结果，针对不同配置（归一化方法+带 IBT 的符号神经元+解码方案）；最大 mAP 值在括号中

从结果可得出以下关键结论：

1. **IBT 是精度提升的核心**：无 IBT 时，即使采用通道归一化，模型在 PASCAL VOC 上的最高 mAP 仅为 7.31%，无法实现有效检测；引入 IBT 后，mAP 提升至 46% 以上，证明负激活信息对目标检测至关重要；
2. **通道归一化优于层归一化**：在有 IBT 的前提下，通道归一化在 PASCAL VOC 上的 mAP (51.83%) 比层归一化 (48.94%) 高 2.89 个百分点，在 MS COCO 上高 1.0 个百分点，且收敛速度更快（通道归一化需 3500 时间步，层归一化需 8000 时间步）；
3. **膜电位解码更优**：基于膜电位 (Vmem) 的输出解码方式（保留膜电位积分的余数）比基于脉冲计数（仅统计脉冲数）的方式精度更高，例如通道归一化 + IBT 配置下，PASCAL VOC 上的 mAP 提升 4.64 个百分点，证明保留连续膜电位信息可减少离散脉冲带来的精度损失。

4.2.2 可视化检测结果

图 6 展示了 Spiking-YOLO 在不同时间步下的可视化检测结果（以“船舶检测”和“行人检测”为例），关键观察如下：

1. **通道归一化收敛更快**：1000 时间步时，通道归一化 + IBT 的模型已能准确检测所有目标并绘制正确边界框，而层归一化 + IBT 的模型未检测到任何目标；
2. **层归一化精度不足**：层归一化模型需 5000 时间步才能检测到目标，但仍存在“多框重叠”“边界框尺寸不准”等问题，而通道归一化模型在 2000 时间步时已达到接近 Tiny YOLO 的检测效果；
3. **复杂场景适应性**：在多目标重叠（如密集行人）和小目标检测场景中，通道归一化模型仍能保持较高检测精度，验证其对复杂目标分布的适应性。

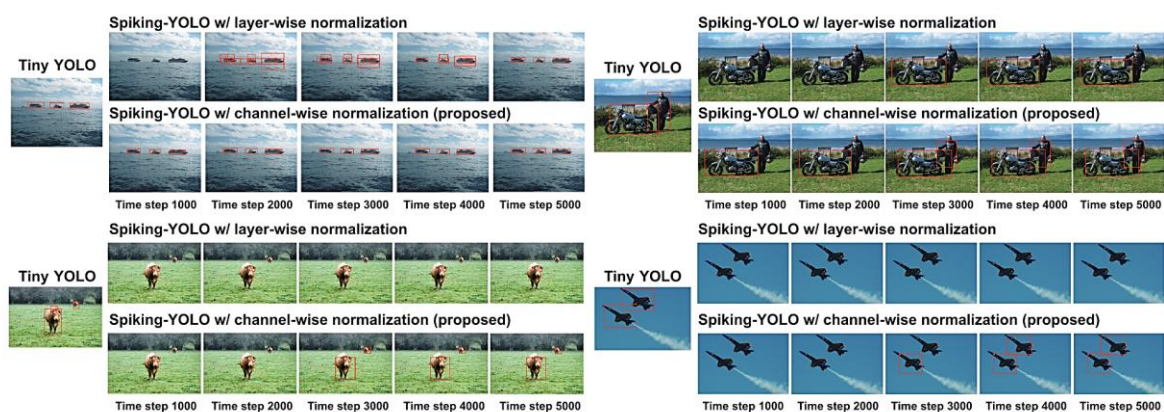


图 6 对象检测结果（Tiny YOLO 与带层归一化的 Spiking-YOLO 与带通道归一化的 Spiking-YOLO）

4.3 能量消耗分析

4.3.1 数字信号处理（DSP）层面的能耗对比

本文首先对比 Spiking-YOLO 与 Tiny YOLO 在 DSP 层面的能耗（仅考虑卷积层的 MAC/AC 操作），结果如图 7 所示：

1. **32 位浮点（FL）操作：**Spiking-YOLO 的能耗约为 Tiny YOLO 的 1/2000，原因是 SNNs 采用 AC 操作（仅积分脉冲），而 DNNs 采用 MAC 操作（乘法 + 积分），且 SNNs 的脉冲稀疏性进一步降低操作数；
2. **32 位整数（INT）操作：**Spiking-YOLO 的能耗优势进一步扩大，约为 Tiny YOLO 的 1/2500，因为整数运算的单位能耗更低，且 SNNs 的二进制脉冲特性更适配整数计算。

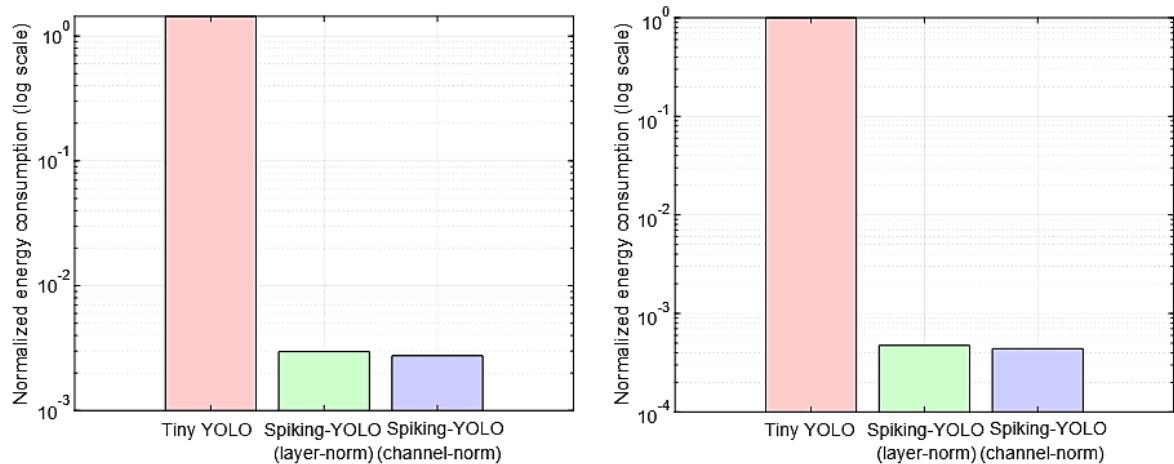


图 7 Tiny YOLO 和 SpikingYOLO 的能量比较，用于 MAC 和 AC 操作；32 位 FL（左）和 32 位 INT（右）

4.3.2 硬件平台层面的能耗对比

为验证 Spiking-YOLO 在实际硬件上的能效优势，本文对比“Tiny YOLO（Titan V100 GPU）”与“Spiking-YOLO（TrueNorth 神经形态芯片）”的能耗，结果如表 2 所示：

表 2：Tiny YOLO（GPU）与 Spiking-YOLO（神经形态芯片）的能量比较

Tiny YOLO					
Power (W)		GFLOPS	FLOPs	Energy (J)	
250		14,000	6.97E+09	0.12	
Spiking-YOLO					
Norm. methods	GFLOPS / W	FLOPs	Power (W)	Time steps	Energy (J)
Layer	400	5.28E+07	1.320E-04	8,000	1.06E-03
Channel	400	4.90E+07	1.225E-04	3,500	4.29E-04

从结果可得出以下结论：

1. **能耗优势显著：**Spiking-YOLO（通道归一化）的能耗仅为 Tiny YOLO 的 1/280（0.12 J vs. 4.29×10^{-4} J），即使采用层归一化，能耗也仅为 Tiny YOLO 的 1/113；

2. **收敛速度影响能耗**：通道归一化模型的时间步仅为层归一化的 43.75% (3500 vs. 8000)，因此能耗仅为层归一化模型的 40.5%，证明快速收敛可进一步降低能耗；
3. **神经形态芯片潜力大**：TrueNorth 芯片（2014 年发布）的能效已显著优于当代 GPU，随着神经形态芯片技术发展（如更高的 GFLOPS/W），Spiking-YOLO 的能耗优势将进一步扩大。

五、结论与展望

5.1 本文总结

本文针对 SNNs 在目标检测任务中性能退化的问题，提出 Spiking-YOLO 模型及两种关键优化方法，主要贡献如下：

- **首次实现 SNNs 目标检测**：Spiking-YOLO 是首个在非 trivial 数据集（PASCAL VOC、MS COCO）上实现与 DNN 相当精度的 SNN 目标检测模型，mAP 可达 Tiny YOLO 的 98%；
- **提出通道归一化方法**：通过精细的通道级归一化，解决传统层归一化中通道激活不均衡的问题，提升神经元发放率与信息传递效率，使模型收敛速度提升 2.3~4 倍；
- **提出非平衡阈值符号神经元**：首次实现 leaky-ReLU 激活函数在 SNNs 中的高效适配，无需额外硬件开销，同时保留负激活信息，显著提升回归任务精度；
- **验证能效优势**：Spiking-YOLO 在神经形态芯片上的能耗仅为 Tiny YOLO（GPU）的 1/280，为边缘设备上的高能效目标检测提供新方案。

5.2 未来研究方向

基于本文的研究，未来可从以下方向进一步拓展：

- **深层 SNN 目标检测**：本文基于浅层的 Tiny YOLO，未来可尝试将方法拓展至更深的模型（如 YOLOv3、Faster R-CNN），探索 SNNs 在更高精度目标检测中的潜力；
- **直接训练 SNNs**：当前 Spiking-YOLO 基于 DNN-to-SNN 转换，未来可研究适用于目标检测的 SNN 直接训练算法，进一步提升模型精度与效率；
- **动态阈值与自适应脉冲编码**：本文采用固定阈值，未来可设计动态阈值机制（如基于输入特征调整阈值），或探索更高效的脉冲编码方式（如时间编码），进一步提升信息传递效率；
- **硬件部署优化**：结合新一代神经形态芯片的特性，优化 Spiking-YOLO 的硬件映射策略，实现更低延迟与更高能效。