NUS
National University
of Singapore

基于脉冲神经网络的编码和算法

新加坡国立大学
张马路
2022年1月

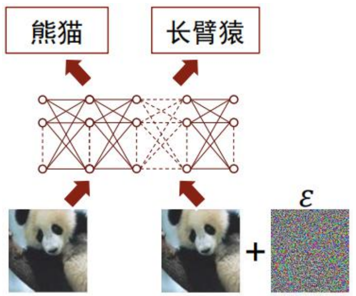
01 研究背景和意义

背景需求 Background

➢ 现有神经网络模型的局限性制约了神经网络的进一步应用

安全性脆弱

现有神经网络模型容易受到对抗样本的攻击，网络安全性脆弱

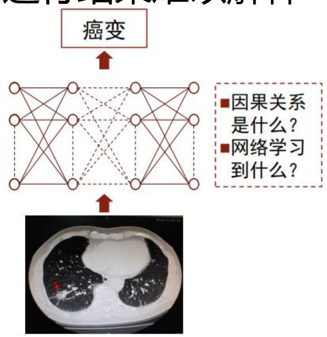


过分依赖大数据

现有神经网络模型的训练过分依赖大数据

可解释性差

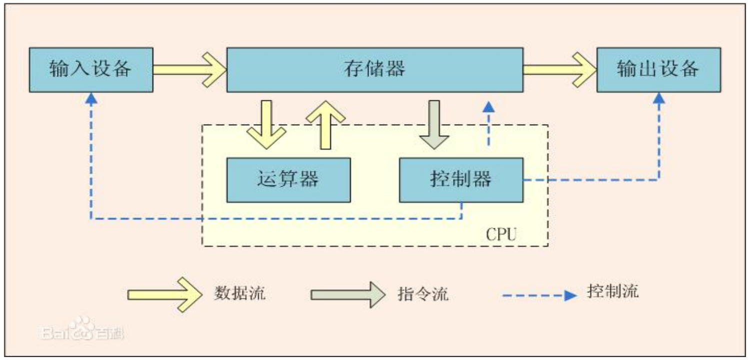
现有神经网络模型被认为是黑箱，网络学习和运行结果难以解释



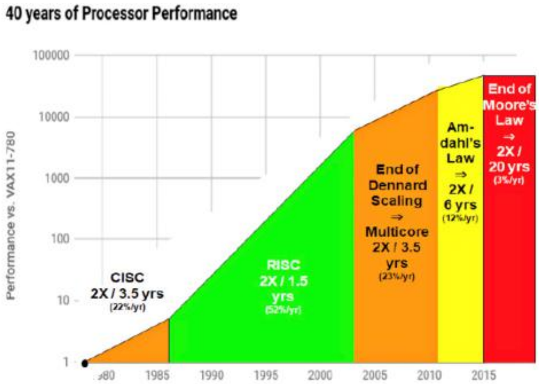
功耗高

现有深度学习功耗高，实际部署困难

➢ 现有计算体系瓶颈：冯诺依曼计算体系效能低下和摩尔定律逐渐失效



存-算分离的冯诺依曼计算体系

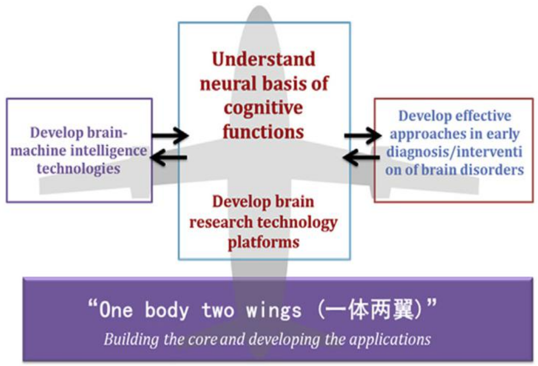


摩尔定律逐渐失效

背景需求 Background

- 类脑计算深入借鉴大脑信息处理的规律，是提升智能模型性能和打破现有计算体系瓶颈的重要方案，蕴含着重大创新的可能性。

中国脑计划（2021-）

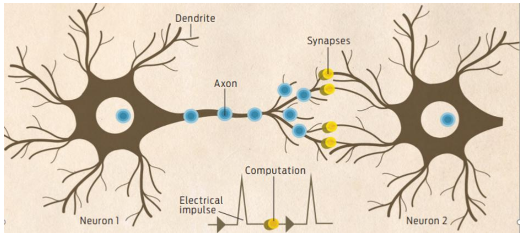


中华人民共和国科学技术部
Ministry of Science and Technology of the People's Republic of China

科技创新 2030— “脑科学与类脑研究”
重大项目 2021 年度项目申报指南

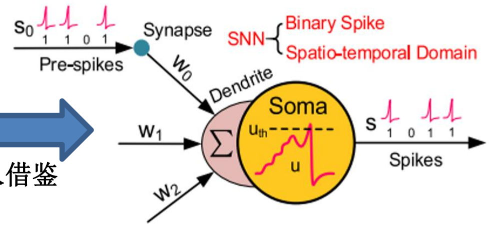
中国脑计划，首批拨款31.48亿，
未来五年拨款190亿。

- 类脑计算包含芯片和软件算法，其中脉冲神经网络（Spiking Neural Networks, SNNs）是软件算法的核心。



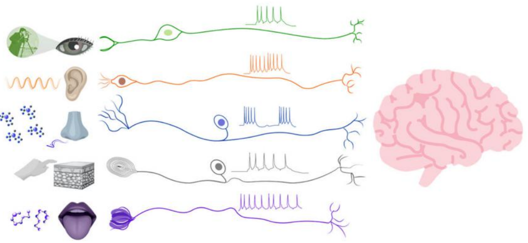
生物神经元

深入借鉴

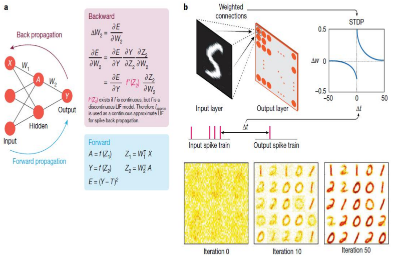


脉冲神经元

编码?

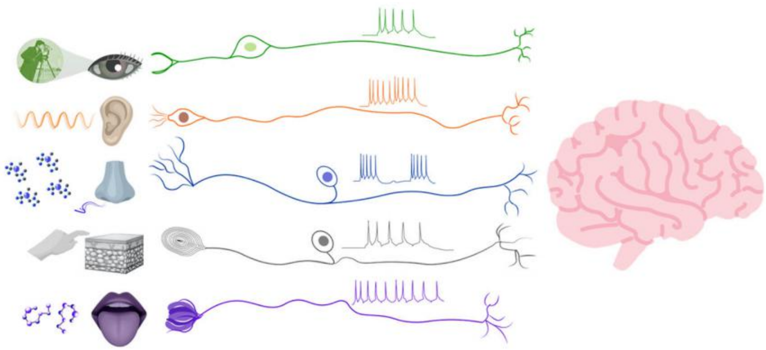


算法?

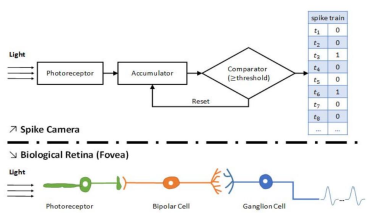


类脑脉冲时空编码

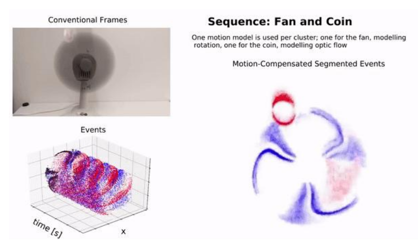
- **在现实生活环境中，大脑通过多种感觉器官（眼睛、耳朵、皮肤等）接收复杂且快速时变的多模态信息。如何对这些异质的多模态信息进行有效的脉冲时空编码，对后续脉冲神经网络模型的学习、推理和认知具有重大影响。**



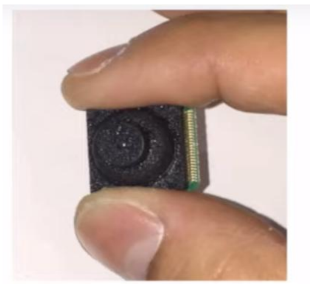
- **视觉**



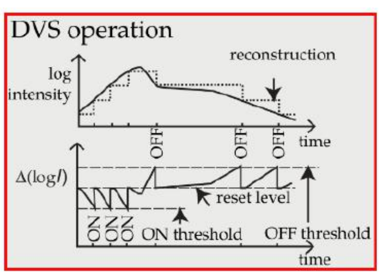
积分型



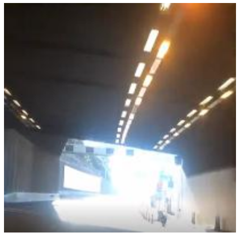
高时域分辨率



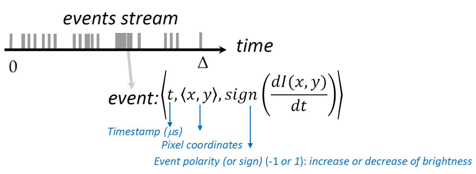
低功耗



差分型



高动态范围

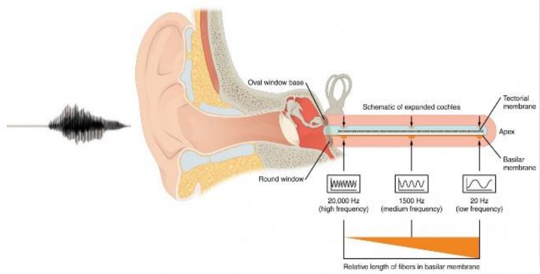


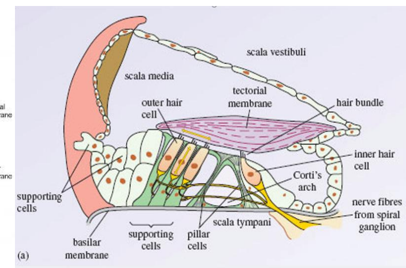
异步输出

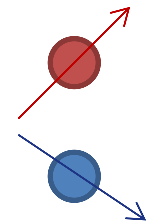
编码1: 基于语音识别任务的类脑脉冲时空编码

➤ **基于语音识别任务的类脑脉冲时空编码方案。**

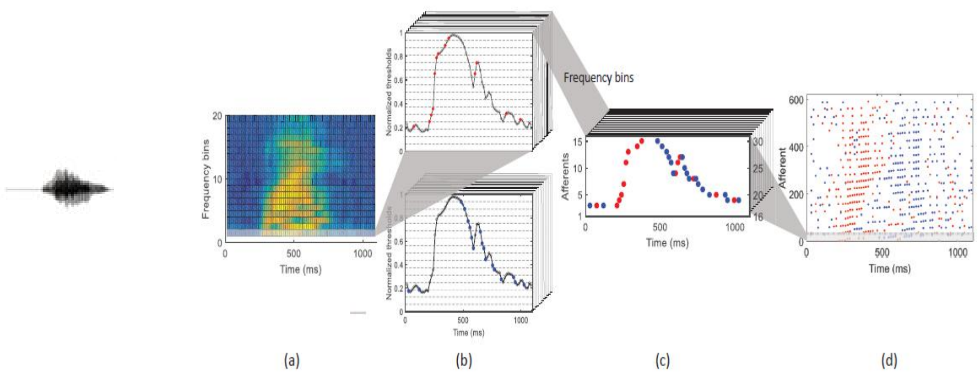
- 大脑听觉前端







- 类脑语音编码



基于识别任务的脉冲时空编码方案

编码1: 基于语音识别任务的类脑脉冲时空编码

- **基于语音识别任务的类脑脉冲时空编码方案。**

在语音数据集**TIDIGITS**上的测试结果

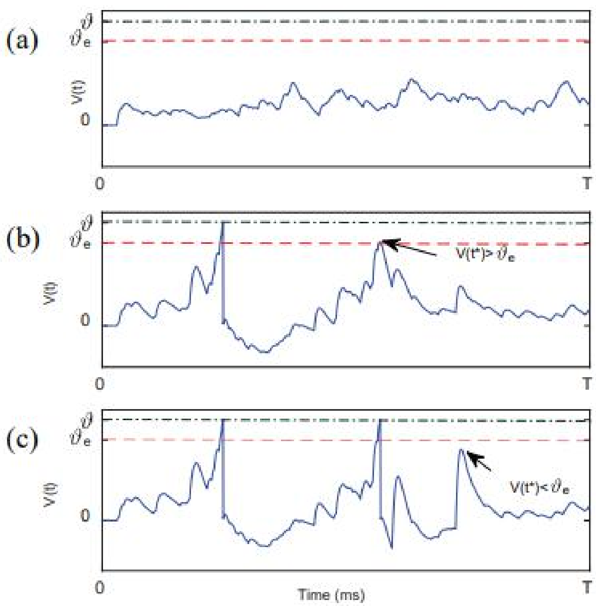


Table 1: Comparison of the proposed framework against other baseline frameworks.

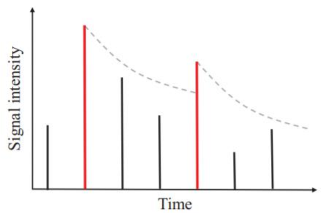
Model	Accuracy
Single-layer SNN and SVM (Tavanaei and Maida 2017b)	91.00%
Spiking CNN and HMM (Tavanaei and Maida 2017a)	96.00%
AER Silicon Cochlea and SVM (Abdollahi and Liu 2011)	95.58%
Auditory Spectrogram and SVM (Abdollahi and Liu 2011)	78.73%
AER Silicon Cochlea and Deep RNN (Neil and Liu 2016)	96.10%
Liquid State Machine (Zhang et al. 2015)	92.30%
MPD-AL with $N_d = 3$	95.35%
MPD-AL with Dynamic Decoding	97.52%

"MPD-AL: an efficient membrane potential driven learning algorithm for SN." ***AAAI-2019.***

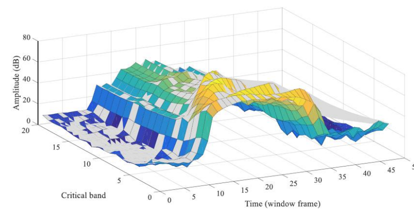
编码1: 基于语音识别任务的类脑脉冲时空编码

➢ **基于语音识别任务的类脑脉冲时空编码方案。**

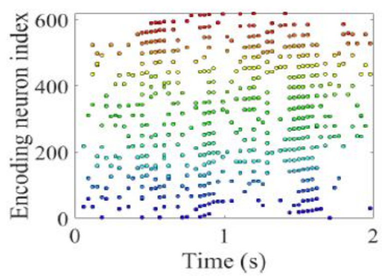
- 听觉掩码机制

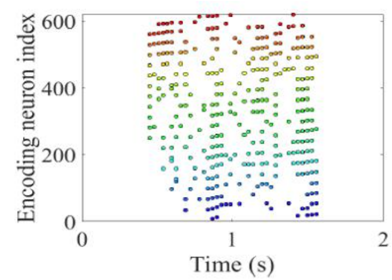


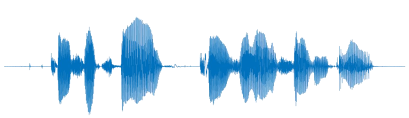
时间掩码



频率掩码



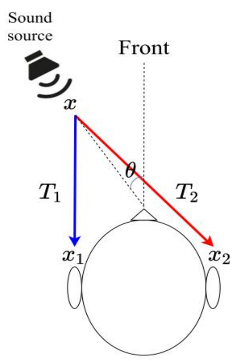




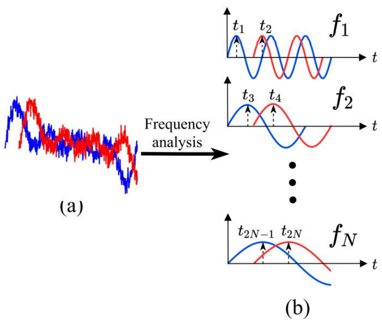
[2] "An efficient and perceptually motivated auditory neural encoding and decoding algorithm for spiking neural networks." Frontiers in neuroscience 13 (2020): 1420.

编码2: 基于声源定位任务的类脑脉冲时空编码

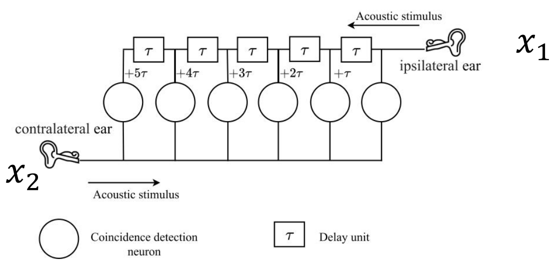
➢ **基于声源定位任务的类脑脉冲时空编码方案。**



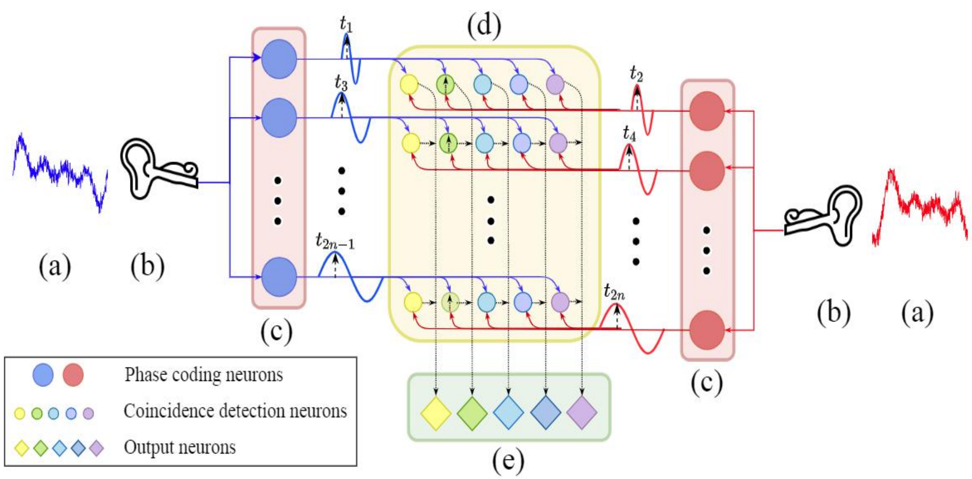
定位任务



相位分析



Jeffress model



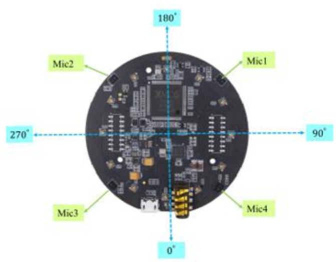
基于定位任务的多频带相位编码

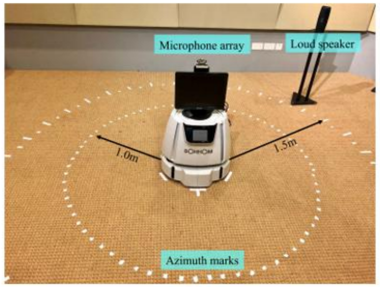
Multi-Tone Phase Coding of Interaural Time Difference for Sound Source Localization With Spiking Neural Networks. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 29, 2656-2670.

编码2: 基于声源定位任务的类脑脉冲时空编码

➢ 基于声源定位任务的类脑脉冲时空编码方案。

- 实验环境

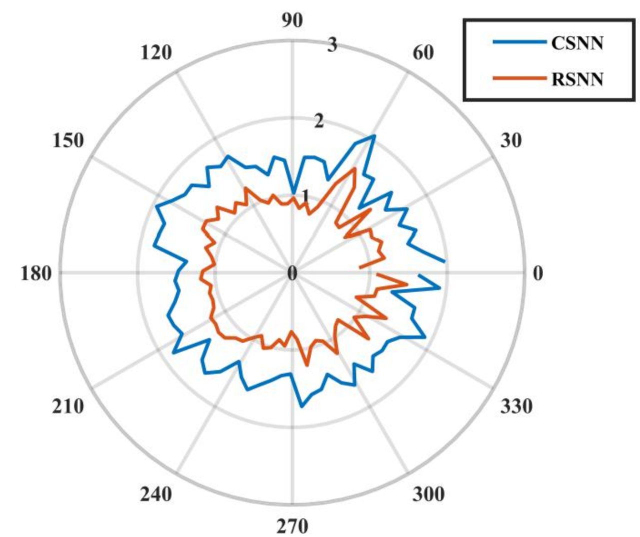




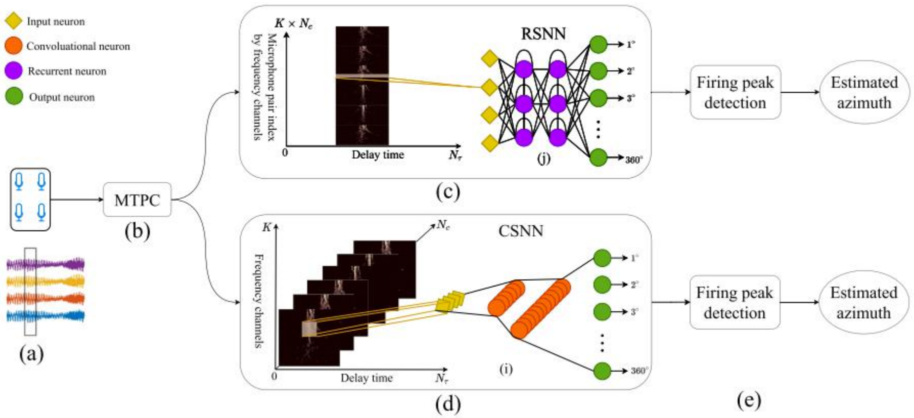
- 实验结果

TABLE VII
A COMPARISON OF MAE BETWEEN MTPC-SNN AND OTHER COMPETITIVE SOUND SOURCE LOCALIZATION TECHNIQUES

Computational model	MAE (1.5m)	MAE (1.0m)
MTPC-CSNN [55]	1.61	4.84
MTPC-RSNN [44]	1.02	4.09
MTPC-CNN [56]	1.20	4.06
MTPC-LSTM [57]	0.41	3.89
GCC-Phat + CNN [52]	1.57	4.38
MUSIC [54]	2.35	3.79



- 计算模型

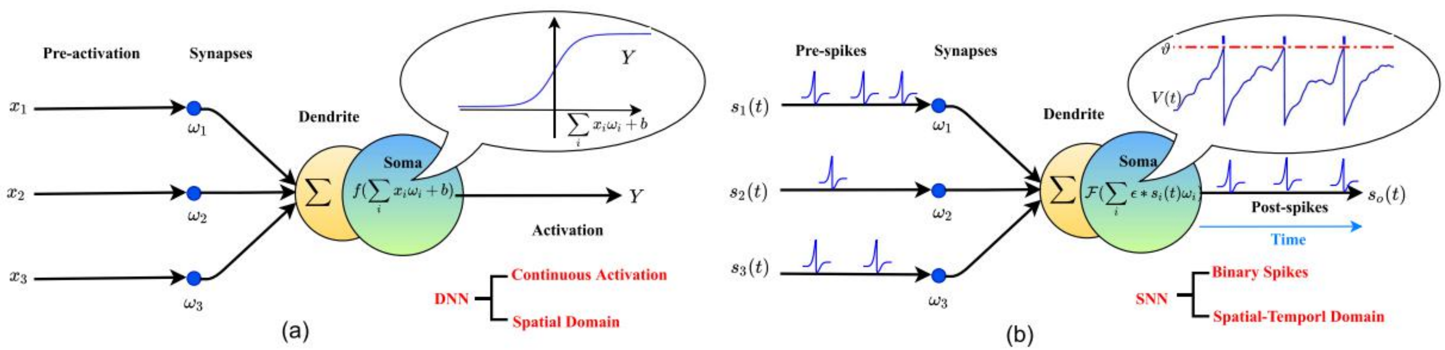


编码2: 基于声源定

- 基于声源定位任务的类

基于脉冲神经网络的算法

- **脉冲神经网络具有复杂的时序依赖关系且脉冲激发函数不可导，因此如何高效训练脉冲神经网络一直是一个研究热点。**



单个神经元 → 三层神经网络 → 深度神经网络

ANN-SNN

代理梯度

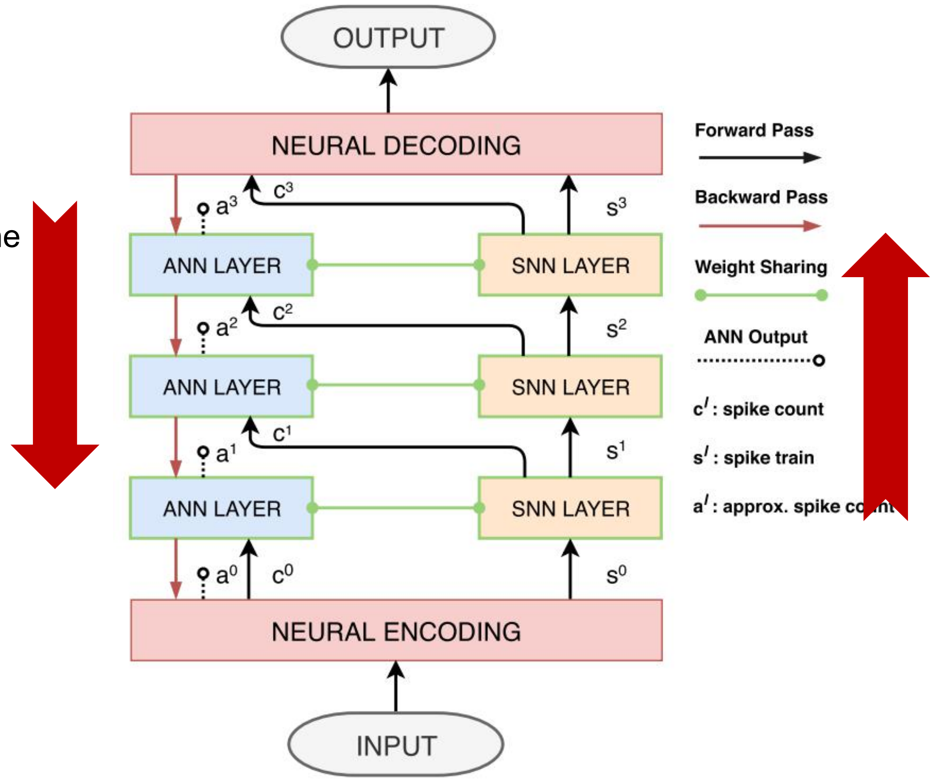
脉冲驱动BP

等

算法1: 基于深度脉冲神经网络的学习算法-Tandem Learning

➢ 异构融合的学习策略-Tandem Learning

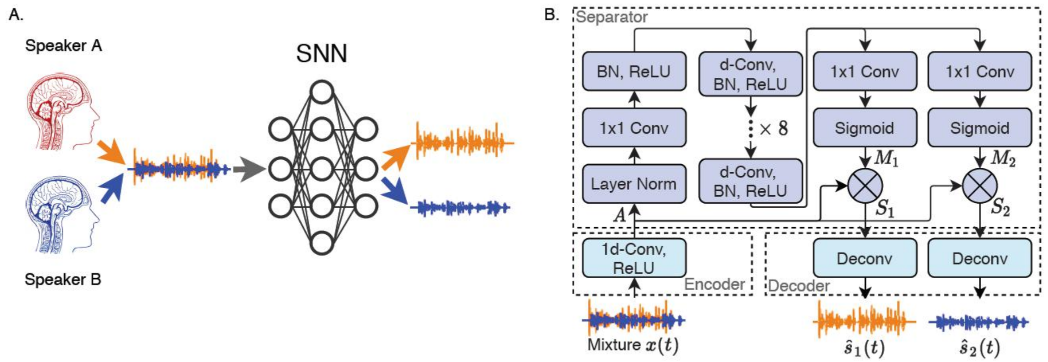
Backward Pass
Approximate the gradient of the coupled SNN layers with the weight-sharing ANN.

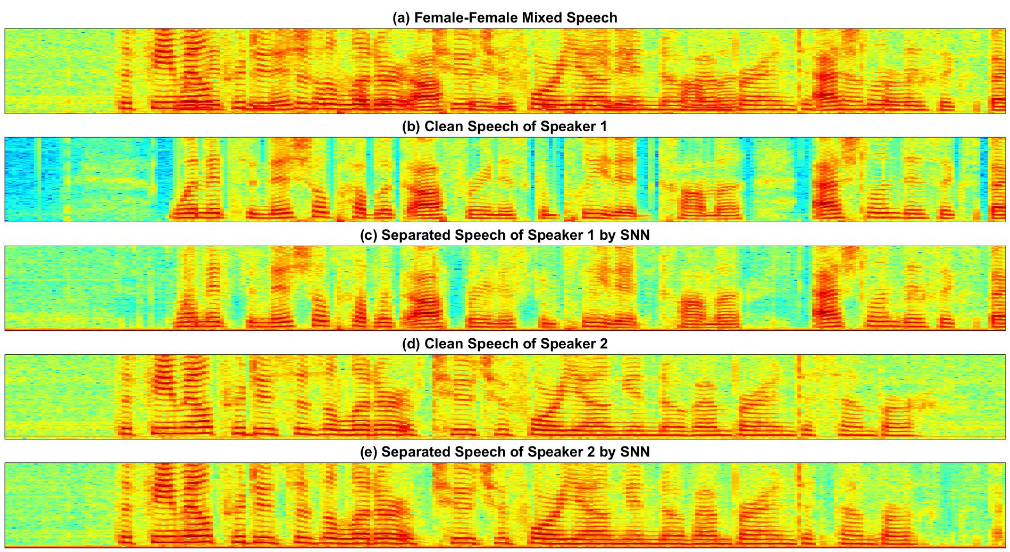


Forward Pass
Exact spike representation determined from the SNN layers

算法1: 基于深度脉冲神经网络的学习算法-Tandem Learning

➢ **实验结果。**





"Progressive tandem learning for pattern recognition with deep spiking neural networks." IEEE Transactions on Pattern Analysis and Machine Intelligence (2021).

算法2: 基于脉冲时间信息的误差反传算法 STDBP

- **基于脉冲时间信息的误差反传算法 STDBP （Spike-Timing-Dependent BP）**

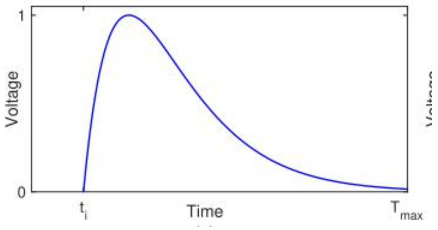
问题分析：

以脉冲驱动的方式进行权重更新，具有更高的学习效率；适用于时间编码，能充分利用脉冲神经网络的时间信息。

- 脉冲发放函数不可微分

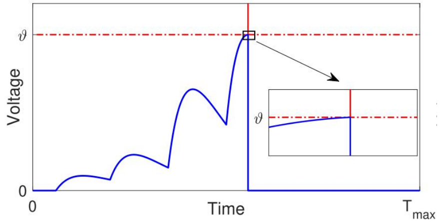
$$t_j^l = \mathcal{F}\{t | V_j^l(t) = \vartheta, t \geq 0\}$$

$$\boxed{\frac{\partial t_j^l}{\partial V_j^l(t_j^l)}}$$

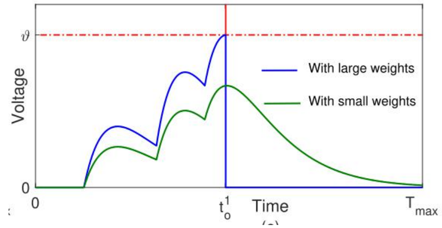


- 梯度爆炸

$$\frac{\partial t_j^l}{\partial V_j^l(t_j^l)} = \frac{-1}{\partial V_j^l(t_j^l)/\partial t_j^l}$$



- ‘静默’神经元



算法2: 基于脉冲驱动的误差反传算法 STDBP

- **基于脉冲时间信息的误差反传算法 STDBP （Spike-Timing-Dependent BP）**

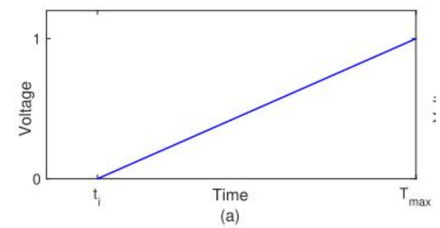
问题分析：

以脉冲驱动的方式进行权重更新，具有更高的学习效率；适用于时间编码，能充分利用脉冲神经网络的时间信息。

- 脉冲发放函数不可微分

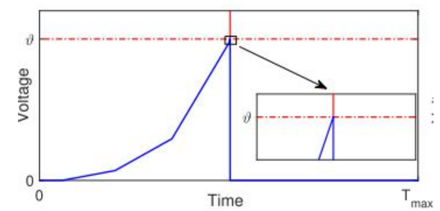
$$t_j^l = \mathcal{F}\{t|V_j^l(t) = \vartheta, t \geq 0\}$$

$$\boxed{\frac{\partial t_j^l}{\partial V_j^l(t_j^l)}}$$

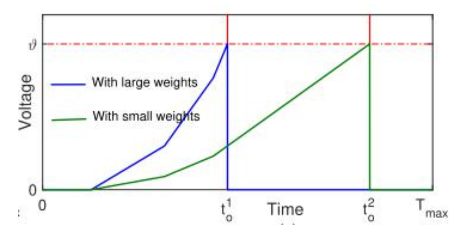


- 梯度爆炸

$$\frac{\partial t_j^l}{\partial V_j^l(t_j^l)} = \frac{-1}{\partial V_j^l(t_j^l)/\partial t_j^l}$$



- ‘静默’神经元



算法2: 基于脉冲时间信息的误差反传算法 STDBP

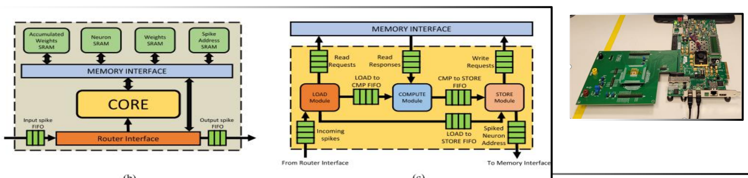
➢ 基于脉冲时间信息的误差反传算法 STDBP （Spike-Timing-Dependent BP）

- 实验结果

TABLE I

CLASSIFICATION ACCURACIES OF EXISTING SPIKE-DRIVEN LEARNING ALGORITHMS ON THE MNIST DATASET. WE USE THE FOLLOWING NOTATION TO INDICATE THE SNN ARCHITECTURE. LAYERS ARE SEPARATED BY—AND SPATIAL DIMENSIONS ARE SEPARATED BY ×. THE CONVOLUTION LAYER AND POOLING LAYER ARE REPRESENTED BY C AND P, RESPECTIVELY

Model	Coding	Network Architecture	Additional Strategy	Acc. (%)
Mostafa [36]	Temporal	784-800-10	Weight and Gradient Constraint	97.5
Tavanaei et al [67]	Rate	784-1000-10	None	96.6
Comsa et al [49]	Temporal	784-340-10	Weight and Gradient Constraint	97.9
Kheradpisheh et al [48]	Temporal	784-400-10	Weight Constraint	97.4
ANN	Rate	784-800-10	None	98.6
STDBP (This work)	Temporal	784-340-10	None	**98.0**
STDBP (This work)	Temporal	784-400-10	None	**98.1**
STDBP (This work)	Temporal	784-800-10	None	**98.5**
STDBP (This work)	Temporal	784-1000-10	None	**98.5**
CNN	Rate	28×28-16C5-P2-32C5 -P2-800-128-10	None	99.5
STDBP (This work)	Temporal	28×28-16C5-P2-32C5 -P2-800-128-10	None	**99.4**



Accelerator	Coding	Acc. (%)	fps	Tech	Power	uJ/frame
SNNwt [76]	Rate	91.82	-	65	-	214.700
TrueNorth-a [24]	Rate	92.70	1000	28	0.268	0.268
TrueNorth-b [24]	Rate	99.42	1000	28	108.000	108.000
Loihi-a [77]	Rate	98.79	120	14	-	-
Loihi-b [78]	Rate	99.21	150	14	99.248	660
Spinnaker [56]	Rate	95.01	77	130	300.000	3896.000
Tianji [57]	Rate	96.59	-	120	120.000	-
Shenjing [55]	Rate	96.11	40	28	1.260	38.000
STDBP+YOSO (This work)	Temp.	98.45	21	22	**(0.878*) 0.751**	**(41.93*) 35.839**

**Scaled for 28 nm process (×1.17 for half a generation)

"Rectified linear postsynaptic potential function for backpropagation in deep spiking neural networks." IEEE Transactions on Neural Networks and Learning Systems (2021).

谢谢大家！