FTBC: Forward Temporal Bias Correction for Optimizing ANN-SNN Conversion

Xiaofeng Wu[*1] Velibor Bojkovic[*2] Bin Gu[23] Kun Suo[4] Kai Zou[5]

Faculty of Data Science, City University of Macau, Macau, China

Department of Machine Learning, Mohamed bin Zayed University of Artificial Intelligence, Abu Dhabi, United Arab Emirates

School of Artificial Intelligence, Jilin University, China

Department of Computer Science, Kennesaw State University, Georgia, USA

ProtagoLabs Inc

Abstract

Spiking Neural Networks (SNNs) offer a promising avenue for energy-efficient computing compared with Artificial Neural Networks (ANNs), closely mirroring biological neural processes. However, this potential comes with inherent challenges in directly training SNNs through spatio-temporal backpropagation — stemming from the temporal dynamics of spiking neurons and their discrete signal processing — which necessitates alternative ways of training, most notably through ANN-SNN conversion. In this work, we introduce a lightweight Forward Temporal Bias Correction (FTBC) technique, aimed at enhancing conversion accuracy without the computational overhead. We ground our method on provided theoretical findings that through proper temporal bias calibration the expected error of ANN-SNN conversion can be reduced to be zero after each time step. We further propose a heuristic algorithm for finding the temporal bias only in the forward pass, thus eliminating the computational burden of backpropagation and we evaluate our method on CIFAR-10/100 and ImageNet datasets, achieving a notable increase in accuracy on all datasets.

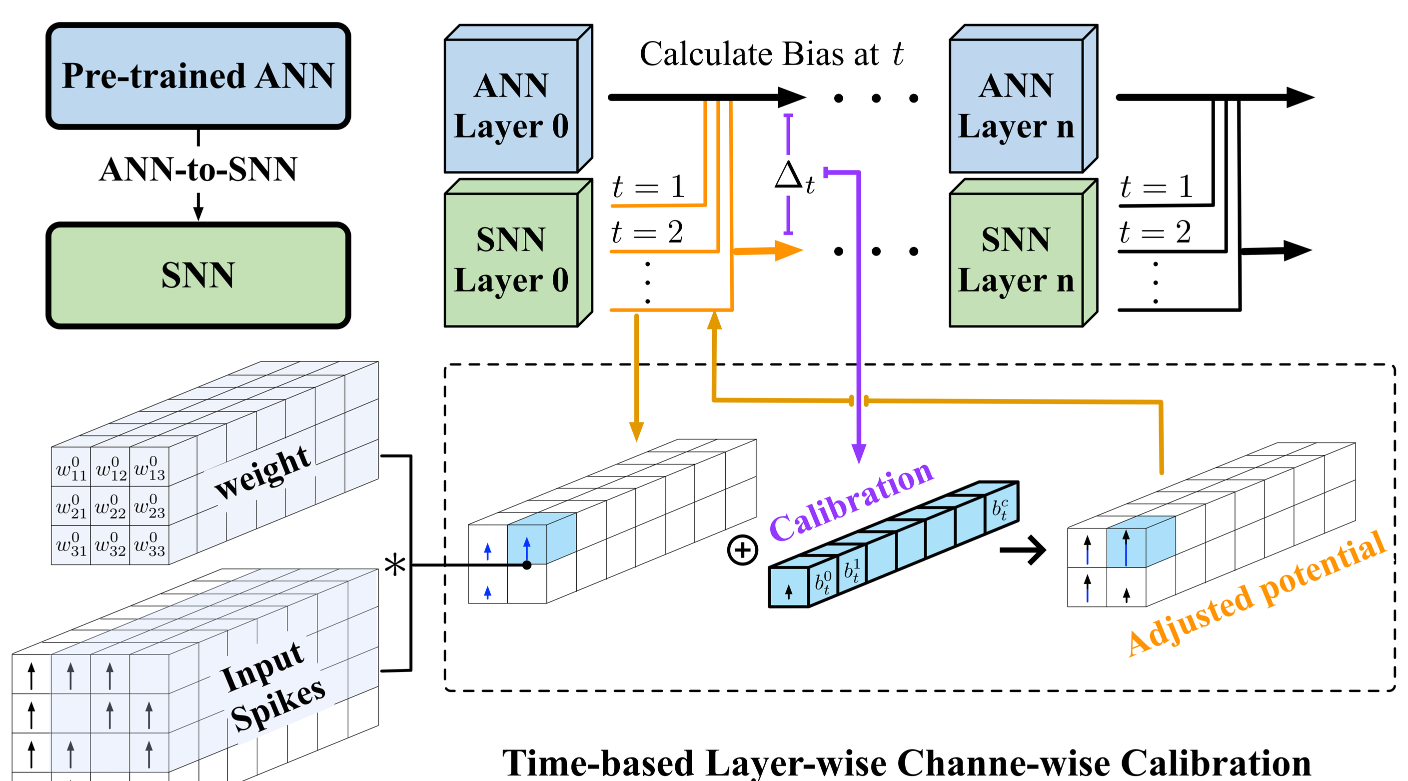


Figure 1. Overview of our proposed Forward Temporal Bias Correction method for ANN-SNN conversion. This approach calibrates time-based channel-wise bias terms (b_t) by dynamically adjusting membrane potential based on the temporal activation patterns observed in the pre-trained ANN. These adjustments are distributed across timesteps from $t = 1$ to $t = T$, ensuring the temporal precision of spike dynamics in SNNs is maintained.

Introduction

SNNs provide an energy-efficient alternative to ANNs, but direct training of SNNs remains challenging due to the discrete and temporal nature of spiking neurons. ANN-to-SNN conversion offers a solution by leveraging pre-trained ANNs, converting them to SNNs while preserving performance. However, the process introduces errors due to differences in neuronal dynamics. Our work addresses these challenges with Forward Temporal Bias Correction (FTBC), enhancing conversion accuracy by dynamically adjusting membrane potential over time. This approach eliminates the need for backpropagation during training, reducing computational complexity.

SNNs often employ integrate-and-fire (IF) neurons, which accumulate membrane potential from incoming weighted spike inputs. Once the membrane potential surpasses a threshold, the neuron emits a spike and resets its potential by subtracting the threshold value. This dynamic occurs across discrete time steps. The membrane potential of a spiking neuron is represented as:

$$v^{(\ell)}[t+1] = v^{(\ell)}[t] + W^{(\ell)}\theta^{(\ell-1)}s^{(\ell-1)}[t+1] - \theta^{(\ell)}s^{(\ell)}[t],$$

where $v[t]$ is the membrane potential at time t, $s[t]$ indicates whether a spike occurs, W is the synaptic weight, and θ is the spiking threshold. The function $H(v^{(\ell)}[t+1] - \theta^{(\ell)})$ governs spike firing.

The goal of ANN-to-SNN conversion is to approximate ANN outputs by summing SNN neuron outputs over time, reducing errors introduced by temporal dynamics. Summing over time and scaling the output yields:

$$\frac{1}{T}\left(\theta^{(\ell)}\sum_{t=1}^{T} s^{(\ell)}[t]\right) = W^{(\ell)}\frac{1}{T}\left(\theta^{(\ell-1)}\sum_{t=1}^{T} s^{(\ell-1)}[t]\right) - \frac{1}{T}v^{(\ell)}[T],$$

where T is the total simulation time. This equation forms the foundation of our method to ensure accurate conversion between ANNs and SNNs.

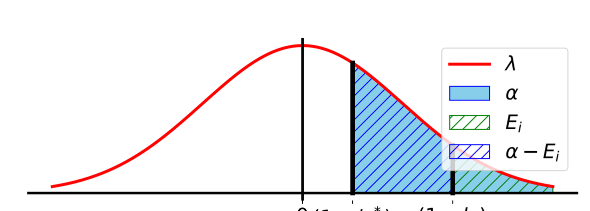
Methodology

We propose the Forward Temporal Bias Correction (FTBC) method to address the conversion errors between ANNs and SNNs. FTBC introduces a time-dependent bias term to correct the membrane potential of spiking neurons during conversion.

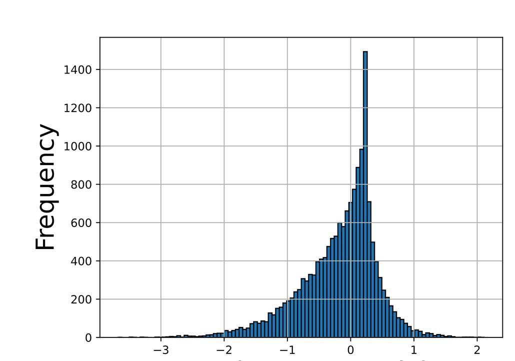
1. **Bias Calculation:** At each time step, the membrane potential update equation for spiking neurons is modified to include a bias term:

   $$v^{(\ell)}[t+1] = v^{(\ell)}[t] + W^{(\ell)}\theta^{(\ell-1)}s^{(\ell-1)}[t+1] - \theta^{(\ell)}s^{(\ell)}[t] + b^{(\ell)}[t]$$

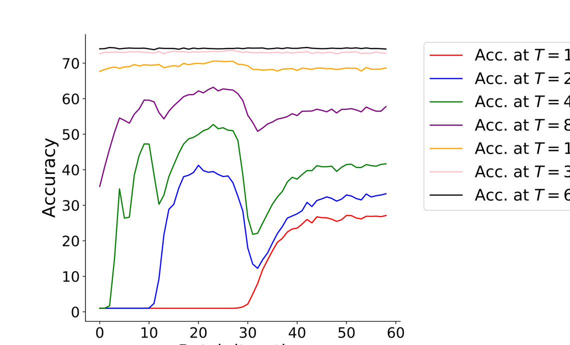
   The bias $b^{(\ell)}[t]$ is calculated in a forward pass, ensuring the expected output of the SNN matches the corresponding ANN at each time step.
2. **Heuristic Algorithm:** We employ a heuristic approach to iteratively adjust the bias, minimizing the error between the ANN and SNN outputs. The bias is computed channel-wise and time-step-wise, aligning the SNN activations with the pre-trained ANN by comparing output differences at each timestep.



(a) Visualization of the FTBC process where the bias term b_t is adjusted based on the temporal differences between ANN and SNN activations.



(b) Layer $l = 6$, $t = 1$. Depiction of membrane potential distributions for a VGG16 model pretrained on the CIFAR100 dataset.



(c) Change of accuracy at various time steps with batch iteration for a VGG16 on CIFAR100 and $\alpha = 0.5$ as a hyperparameter in calibration.

Theoretical Analysis

The foundation of Forward Temporal Bias Correction (FTBC) relies on correcting the expected error between ANN and SNN outputs by introducing a time-dependent bias term.

Proposition for Bias Calculation: Let λ be a continuous distribution with compact support $[a, c]$, and $\alpha \in [0, 1]$. There exists a unique bias b^* such that:

$$\mathbb{E}_{x\sim\lambda}[H(x + b^* - 1)] = \alpha$$

where $H(\cdot)$ is the Heaviside function. This proposition ensures that the expected outputs of the SNN neurons match those of the corresponding ANN neurons by finding the optimal bias.

Experiment

Table 1. Comparison between our proposed method and other ANN-SNN conversion methods on ImageNet.

Architecture	Method	ANN	T=4	T=8	T=16	T=32	T=64	T=128
ResNet-34	SNNC-AP*	75.66	–	–	–	64.54	71.12	73.45
	QCFS	74.32	–	–	59.35	69.37	72.35	73.15
	Ours*	75.66	13.70	38.55	60.68	70.88	74.29	75.28
	Ours (+QCFS)	74.32	49.94	65.28	71.66	73.57	74.07	74.23
VGG-16	SNNC-AP*	75.36	–	–	–	63.64	70.69	73.32
	SNM*	73.18	–	–	–	64.78	71.50	72.86
	RTS	72.16	–	–	55.80	67.73	70.97	71.89
	QCFS	74.29	–	–	50.97	68.47	72.85	73.97
	SlipReLU	71.99	–	–	51.54	67.48	71.25	72.02
	OPI*	74.85	–	6.25	36.02	64.70	72.47	74.24
	Ours*	75.36	51.13	64.20	71.19	73.89	74.86	75.29
	Ours (+RTS)	72.16	29.61	55.22	67.14	70.74	71.86	72.13
	Ours (+QCFS)	73.91	58.83	69.31	72.98	74.05	74.16	74.21

* Without modification to ReLU of ANNs.

References

[1] Tong Bu, Wei Fang, Jianhao Ding, Penglin Dai, Zhaofei Yu, and Tiejun Huang.
Optimal ANN-SNN conversion for high-accuracy and ultra-low-latency spiking neural networks.
In *The Tenth International Conference on Learning Representations, ICLR 2022, Virtual Event, April 25-29, 2022*, 2022.

[2] Shikuang Deng and Shi Gu.
Optimal conversion of conventional artificial neural networks to spiking neural networks.
In *ICLR*, 2021.

[3] Yuhang Li, Shikuang Deng, Xin Dong, Ruihao Gong, and Shi Gu.
A free lunch from ann: Towards efficient, accurate spiking neural networks calibration.
In *International conference on machine learning*. PMLR, 2021.