



HAL
open science

DNM-SNN: Spiking Neural Network Based on Dual Network Model

Zhen Cao, Hongwei Zhang, Qian Wang, Chuanfeng Ma

► **To cite this version:**

Zhen Cao, Hongwei Zhang, Qian Wang, Chuanfeng Ma. DNM-SNN: Spiking Neural Network Based on Dual Network Model. 5th International Conference on Intelligence Science (ICIS), Oct 2022, Xi'an, China. pp.11-19, 10.1007/978-3-031-14903-0_2 . hal-04666439

HAL Id: hal-04666439

<https://hal.science/hal-04666439v1>

Submitted on 1 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

DNM-SNN: Spiking Neural Network Based on Dual Network Model

Zhen Cao, Hongwei Zhang, Qian Wang, Chuanfeng Ma

School of Artificial Intelligence
Xidian University, Xi'an, Shaanxi, P.R. China
icaozhen@163.com

Abstract. In recent years, deep neural network (DNN) has shown excellent performance in many applications. However, the huge energy consumption leads to many problems. In order to solve this problem, spiking neural network (SNN) has attracted extensive research attention. SNN is the third-generation neural network that is used to process complex spatiotemporal data, and it has become a hot research topic due to its event-driven and low-power characteristics. However, the propagation function of spiking neurons is usually non-differentiable, which prevents back propagation and makes the training of SNN difficult. This paper proposes an efficient supervised learning algorithm framework based on dual-network-model spiking neural network (DNM-SNN), which is universal to various supervised learning algorithms of spiking neural networks and can effectively improve the prediction accuracy. DNM-SNN includes two key methods. Firstly, a dual model training method in training stage is proposed, which requires an additional auxiliary network same as the network used. Single model training is easy to fall into local optimal problems. By maintaining two networks, the same problem can be viewed from different perspectives, which solves the problem and improves the training effect. Second, we propose a multi-channel mix module inference method in the prediction stage. The prediction accuracy of the model is improved and the performance of the spiking neural network is optimized by multi-channel optimization of mix module. Experimental results show that the DNM-SNN outperforms the single-model algorithm on classification tasks, with a slight improvement on the MNIST dataset and a 3% improvement on the CIFAE-10 dataset.

Keywords: Spiking Neural Network, Dual Model, Mix Module, DNM-SNN.

1. Introduction

The emergence of back-propagation algorithm [1] has significantly promoted the development of traditional artificial neural network (ANN), which can effectively solve many real-life problems, such as image recognition, control and prediction. With the advent of training algorithms for deep convolutional neural networks [2], deep neural networks were only understood in the fields of computer vision and

machine learning and became the mainstream method for solving tasks such as image recognition and speech recognition. Research on spiking neural networks dating back to the 1950s has not received as much attention as ANNs, one of the main reasons being the lack of efficient deep spiking due to its replicated dynamics and non-differentiable operations Learning algorithms for neural networks. Many researchers have proposed some algorithm frameworks, such as the deep convolutional neural network based on STDP rules proposed by Kheradpisheh [3], which achieved an accuracy level of 98.4% on the MNIST dataset; Rueckauer et al. VGG-16 and GoogLeNet models; Hu et al. [4] used the deep structure of ResNet-50 to achieve 72.75% accuracy classification results. However, their algorithms are either only used for classification tasks on simple datasets, or the accuracy of the algorithms is not guaranteed and cannot achieve the same performance as ANN. In recent years, Cao [5] and Sengupta [6], etc. proposed a new simple and effective algorithm, which imported the pre-trained parameters from DNN to SNN, and realized the conversion from DNN to SNN. In MNIST and CIFAR datasets Competitive results have been achieved. However, due to the existence of the temporal dimension, transforming SNNs on deep networks often suffers from the problem of vanishing gradients, and simple transplant transforms can also cause the problem of insufficient burst firing rate.

It can be seen that the spiking neural network lacks an efficient and stable supervised training algorithm, and it is a difficult task to train a spiking neural network well. There are three main reasons for the poor performance of the algorithm: a) The neural information is expressed in the form of a pulse sequence, which is essentially a discrete signal, and the state variables and error propagation functions inside the spiking neuron do not satisfy the continuous differentiability property, which makes the gradient descent learning method based on backpropagation cannot be directly applied to the spiking neural network. b) In the prediction stage, due to the low firing rate of spiking neurons, the output neurons of the last layer of the deep spiking neural network are prone to not firing pulses, resulting in serious deviations from the target pulse sequence. c) During the training process, due to the low firing rate of spiking neurons, the model cannot converge, and the single-network model is prone to fall into the local optimal solution.

To sum up, the main contribution of this paper is as follows:

Dual-model spiking network training method: In order to enhance the reliability of the algorithm, the dual networks jointly learn a problem, improve the convergence speed of training, and prevent the trouble of falling into local optimum.

Multi-channel mix module prediction method: multi-channel prediction, find the most suitable prediction channel, and effectively improve the prediction accuracy.

Dual model achieves better performance than single model: On the CIFAR-10 dataset, the dual-model Resnet11 is 3% higher than the single-model, and the dual-model Resnet20 is 1% to 3% higher than the state-of-the-art single-model Resnet20.

2. Methods

The traditional supervised learning algorithm of spiking neural network is to find a suitable synaptic weight matrix for the network for a given set of multiple input pulse sequences and a set of multiple target pulse sequences through a network model. The resulting matrix minimizes the error between the pulse sequence output by the network and the corresponding target pulse sequence. However, in classification tasks, traditional supervised learning algorithms often predict wrong classification results, and the performance of the algorithm is seriously degraded. In-depth analysis of its reasons, the possible explanations for this are:

(a) The single-network model is often biased, and it is easy to fall into the local optimal solution, resulting in the degradation of the performance of the algorithm.

(b) Unstable pulse firing rates lead to very severe biases in predictions.

In order to overcome these two problems, this paper proposes two methods: the dual-model spiking network supervised learning training method and the mix module prediction method. The overall framework is shown in Figure 1.

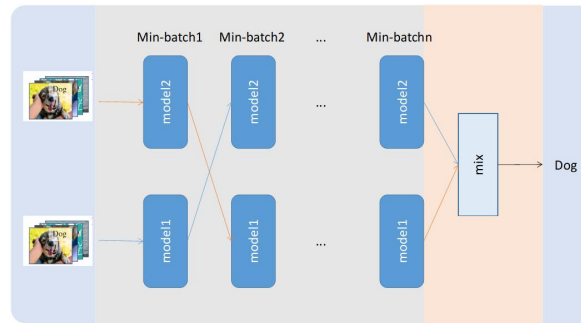


Fig. 1. The framework of the dual-model supervised learning algorithm is divided into four parts. The first part is the encoding operation of the data set, which encodes the data set into a pulse sequence. The second part is the training of dual models, where model1 and model2 are two identical models, and the output pulse sequences of each batch of the two models are exchanged with each other. The third part is the multi-channel mix prediction module, which can select the channel with the best accuracy to complete the classification task. The last part is the output of the classification result, and the selected mix channel outputs the classification result.

1.1 Traditional SNN supervised learning algorithm framework and its limitations.

The training of the traditional spiking neural network supervised learning algorithm [7] is divided into three stages: first, the sample data is encoded into a spike sequence through a specific encoding method. Second, the pulse sequence is input into the neural network, and a certain simulation strategy is applied to run the same neural network to obtain the actual output pulse sequence. Then, according to the target pulse sequence, the synaptic weights of the current neural network are adjusted

through the loss function and its learning rules. Finally, when the prediction error reaches a pre-specified value or the number of iterations is completed, the training is complete. Obviously, single-model training is very risky. Once there is a deviation, it is difficult to correct, because the data obtained is always the result predicted by its own network, so it is easy to fall into the trouble of local optimization. The traditional prediction method is to directly apply the trained weights to the current network, and the output result is the prediction result. However, due to the problem of low firing rate of the pulse, the final result has a large deviation from the actual target result. There may be cases where the category cannot be predicted or the classification effect is poor.

1.2 Proposed Dual-Model Spike Network Supervised Learning Algorithm.

This paper proposes a more efficient SNN training algorithm, which we call a dual-model spiking network supervised learning training method. Our method is to establish two network models with the same structure, instead of one network model in the traditional method, as shown in Figure 1, establish two identical networks model1 and model2, and input the encoded data sets into the two neural networks respectively. In the network, after the prediction results of each batch are inferred, the prediction results of model1 are respectively passed to model2, and the prediction results of model2 are passed to model1. The two networks constrain each other to prevent a large deviation in one network. Then repeat the above steps to complete the training.

This is like two experts solving the same problem, the final answer is the same, but the ideas of solving the problem are different. Even if one of them is in trouble, as long as the other person is right, he will quickly get out of trouble with his help. It is difficult for the two networks to fall into local optima at the same time. Even if they fall into local optima at the same time, they will iterate again because of the difference between the two local optima. The worst case scenario is that both networks fall into the same local optimum at the same time. This paper has not encountered this situation in the experiment. If this situation occurs, it is likely that the two selected network models are inappropriate.

1.3 Proposed Multi-Channel Mix Module Prediction Method.

The output of a spiking neural network is a binary sequence of spikes, and the differences between categories are enormous. Therefore, when the pulse firing rate is insufficient, all the classification results will be 0. In order to solve this problem, a multi-channel mix module prediction method is proposed, which can cooperate with the dual-model supervised learning algorithm to optimize the prediction results and improve the prediction accuracy. The specific algorithm is as follows:

Algorithm 1: Multi-Channel mix Module Prediction Method

```

1 class Mix:
  # Get the output of the dual model
2   def __init__(self, model1output, model2output):
3     self.model1output = model1output

```

```

4     self.model2output = model2output
# maximum channel
5     def maxChannel(self):
6         newoutput = getMax(self.model1output, self.model2output)
7         return newoutput
# Average channel
8     def averageChannel(self):
9         newoutput = getAverage(self.model1output, self.model2output)
10        return newoutput
# Choose the channel best suited for the current task
11    def betterOutput(self):
12        flag = whichBetter(self.maxChannel(), self.averageChannel())
13        if flag:
14            return 1
15        else:
16            return 2

```

1.4 The chosen network model.

We choose Resnet[8] model as the dual model, which is of great significance to the study of deep spike neural network[9]. Here is our snn-resnet11 model:

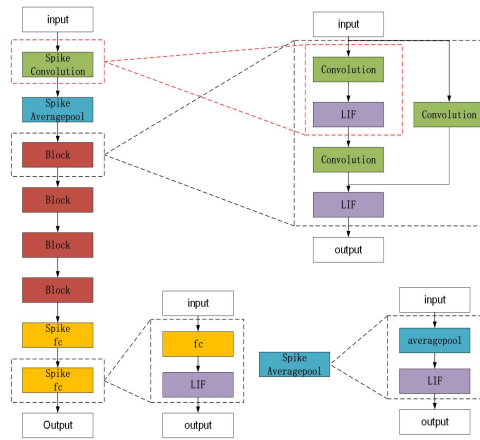


Fig. 2. SNN-Resnet11 single network architecture

1.5 Selection of spiking neurons.

Figure 3 shows the difference in electrical properties between spiking and non-spiking neurons. Non-spiking neurons are memoryless, their current output is completely dependent on the current input, while spiking neurons have inherent memory, they all have a memory element[11].



Fig. 3. Spiking neuron (right) and non-spiking neuron (left) circuit primitives

As shown in Figure 2, the Leaky Integrate-and-Fire (LIF) spiking neuron [12] is selected in this paper. Its working process is shown in Figure 4.

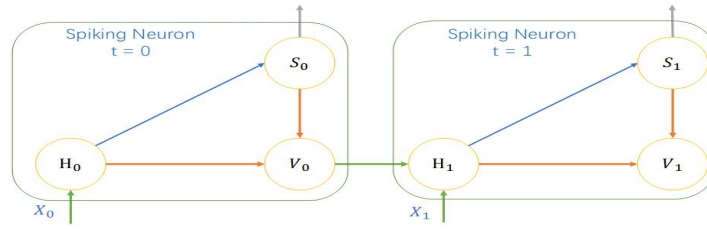


Fig. 4. LIF model working process

Figure 4 shows the working process of LIF neurons [13] and its dynamic model can be described by three discrete equations of charging, discharging and resetting.

The charging equation is:

$$H_t = f(V_{t-1}, X_t) \quad (1)$$

$$f(V_{t-1}, X_t) = V_t + \frac{1}{\tau_m} ((-V_{t-1} - V_{reset}) + X_t) \quad (2)$$

Among them, X_t is the external input, which represents the input of the pulse sequence at time t , such as the increment of the voltage, H_t is the hidden state of the neuron, and is the membrane voltage of the neuron before the pulse at time t . τ_m is the time constant, V_t is the membrane voltage of the neuron at time t , and V_{reset} is the reset voltage.

The discharge equation is as follows:

$$S_t = \theta(H_t - V_{th}) = \begin{cases} 1, & H_t \geq V_{th} \\ 0, & H_t < V_{th} \end{cases} \quad (3)$$

where V_{th} is the threshold voltage, is an $\theta(H_t - V_{th})$ step function, and S_t represents the neuron firing the pulse.

Its reset equation is as follows:

$$V_t = r(H_t, S_t) \quad (4)$$

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

3. Experimental results

3.1 Single- and dual-model Resnet11 performance on the CIFAR-10 dataset.

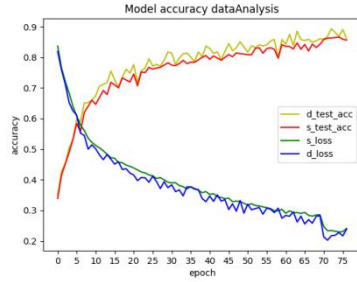


Fig. 5. Performance of the two models on the CIFAR-10 dataset

Table 1. Results for CIFAR-10 and MNIST dataset.

Model	Dataset	Test Acc
Resnet11(single)	MNIST	98.27%
Resnet11(Dual)	MNIST	99.11%
Resnet11(single)	CIFAR-10	86.62%
Resnet11(Dual)	CIFAR-10	89.93%

3.2 Related work comparison.

Table 2. Comparison with other conversion methods on CIFAR-10.

Model	Author	Dataset	Test Acc
Resnet11(single)	This work	CIFAR-10	86.62%
Resnet20	Abhronil Sengupta [14]	CIFAR-10	89.1%
Resnet20	Bing Han [15]	CIFAR-10	91.36%
Resnet20(single)	This work	CIFAR-10	88.74%
Resnet20(Dual)	This work	CIFAR-10	92.33%

4. Conclusion

In this paper, we introduce a dual-model spiking neural network supervised learning algorithm framework, and propose a dual-model training method and a multi-channel mix module prediction method corresponding to the algorithm. We successfully applied it to the SNN-Resnet11 single-network model and obtained a dual-model SNN-Resnet11. Experiments show that the dual-model structure achieves an accuracy improvement of about 3% over the single-network model on the CIFAR-10 dataset.

We believe that this algorithm can be applied to a variety of network architectures to better facilitate the exploration of learning algorithms for spiking neural networks.

5. Acknowledgements

This work was supported in part by the China Postdoctoral Science Foundation under Grant 2019M663637, in part by the Natural Science Basic Research Program of Shaanxi under Program 2021JQ-201, and in part by the National Natural Science Foundation of China under Grant 62104176.

References

1. Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors[J]. *nature*, 1986, 323(6088): 533-536.
2. Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. *Advances in neural information processing systems*, 2012, 25.
3. 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.
4. Rueckauer B, Lungu I A, Hu Y, et al. Conversion of continuous-valued deep networks to efficient event-driven networks for image classification[J]. *Frontiers in neuroscience*, 2017, 11: 682.
5. Cao Y, Chen Y, Khosla D. Spiking deep convolutional neural networks for energy-efficient object recognition[J]. *International Journal of Computer Vision*, 2015, 113(1): 54-66.
6. Sengupta A, Ye Y, Wang R, et al. Going deeper in spiking neural networks: VGG and residual architectures[J]. *Frontiers in neuroscience*, 2019, 13: 95.
7. LIN X, WANG X, ZHANG N, et al. Supervised learning algorithms for spiking neural networks: a review[J]. *ACTA ELECTONICA SINICA*, 2015, 43(3): 577.
8. He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016: 770-778.
9. Hu Y, Tang H, Pan G. Spiking Deep Residual Networks[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2018.
10. Zambrano D, Nusselder R, Scholte H S, et al. Sparse computation in adaptive spiking neural networks[J]. *Frontiers in neuroscience*, 2019, 12: 987.
11. Kim S, Park S, Na B, et al. Spiking-yolo: spiking neural network for energy-efficient object detection[C]//*Proceedings of the AAAI Conference on Artificial Intelligence*. 2020, 34(07): 11270-11277.
12. Lapique L. Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarisation[J]. *Journal of Physiology and Pathology*, 1907, 9: 620-635.
13. Dayan P, Abbott L F. *Computational and mathematical modeling of neural systems*[M]//*Theoretical neuroscience*. MIT Press, 2001.
14. Sengupta A, Ye Y, Wang R, et al. Going deeper in spiking neural networks: VGG and residual architectures[J]. *Frontiers in neuroscience*, 2019, 13: 95.
15. Han B, Srinivasan G, Roy K. Rmp-snn: Residual membrane potential neuron for enabling deeper high-accuracy and low-latency spiking neural network[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020: 13558-13567.