Proceedings of the 5th International Conference on Computing and Data Science
DOI: 10.54254/2755-2721/20/20231096

Comparing performance of neural network frameworks for image recognition: A review

Luting Wang

School of computing science, Georgia Institute of technology, Georgia, United States, GA 30332

lwang797@gatech.edu

Abstract. Spiking Neural Networks (SNNs) are increasingly recognized as a promising approach to simulating the biological behavior of neurons. This study conducted an in-depth performance comparison of different SNN models with supervised learning algorithms, particularly in the field of image recognition. Using datasets like MNIST, CIFAR-10, and ImageNet, the research analyzed the performance and accuracy of SNNs compared to traditional Artificial Neural Networks (ANNs). This research contributes to understanding SNNs' performance in supervised learning environments and offers insights into the optimization of SNN architectures, thus influencing future research and the development of next-generation neural network designs. Based on this comparative study, it was found that while SNNs perform admirably with smaller, less complex datasets such as MNIST and CIFAR, they encounter difficulties with larger, more complex datasets like ImageNet. Despite these challenges, advances in conversion techniques have led to models that more closely simulate the behavior of biological neurons. The study identified significant future work areas, including addressing issues of local learning methods, the limited scale of SNNs, hardware implementation difficulties, a lack of substantial benchmark datasets, and limited support for complex computations.

Keywords: spiking neural networks, supervised learning, image recognition, artificial neural networks, performance comparison.

1. Introduction

Spiking Neural Networks (SNNs), simulating biological neuron behavior [1] and providing advanced information encoding and processing, are an emerging area with key advances in hardware implementations and learning algorithm [2-6]. Despite challenges in directly implementing traditional supervised learning algorithms like backpropagation, promising strategies, including Spike-Timing-Dependent Plasticity (SDTP) and surrogate gradient methods, have emerged, and the integration of SNNs into image recognition tasks has shown notable progress [7-10]. Supervised Learning for SNNs is a promising research topic. Trying to improve the SNNs' performance on the dataset for image recognitions is a significant field.

This paper aims to compare the performance of different SNN models for supervised learning algorithms on the image recognition task and, at the end, address future directions. The comparison of different SNN models provides insights into the strengths and weaknesses of different SNN architectures and learning algorithms, which can inform future research on SNNs. The rest of the article will be

DOI: 10.54254/2755-2721/20/20231096

structured as follows. The background in section 2 will be split into two sections based on the two separate datasets, MNIST and CIFAR-10. A performance comparison of spike train learning using a particular representative convolution model will be presented for each dataset. The comparison evaluates the effectiveness of SNNs with different learning algorithms in improving the performance of Convolutional Neural Networks (CNNs) on image recognition tasks. Section 4 includes the conclusion and the future directions for SNNs.

This paper lays the groundwork for shaping and directing future research in the domain of SNNs. More specifically, this article contributes by revealing the effectiveness of SNN architectures for image recognition tasks, thereby contributing to the optimization of current methodologies and informing the development of next-generation neural network designs. This study fosters the understanding of how varying learning algorithms can impact SNNs' performance, which can have profound implications for both academic research and practical applications in machine learning.

2. Analysis and discussion

2.1. Dataset

The comparison focuses on the performances of supervised learning on ANN-oriented datasets, like MNIST, CIFAR-10, and ImageNet [11-13].

MNIST: The MNIST dataset is a collection of 70,000 grayscale images of handwritten digits. Each image is 28x28 pixels in size and contains a single digit from 0 to 9[11].

CIFAR: The CIFAR dataset is a collection of color images. CIFAR-10 contains 60,000 32x32 color images in 10 classes, with 6,000 images per class. CIFAR-100 is a dataset of 60,000 32x32 color images in 100 classes, with 600 images per class [12].

ImageNet: The ImageNet dataset contains over 14 million images in more than 21,000 categories [13]. Because of the larger size and more diverse range of images, the ImageNet dataset is much more challenging than MNIST and CIFAR.

2.2. Performance comparison

Most of the work on SNNs has used image classification datasets as benchmarks. To show how the different conversion methods and neuron models influence the performance of SNN learning rules, Table 1 compares the existing methods tested on static datasets.

Table 1. The result on static database, including MNIST, CIFAR-10/100, and ImageNet[14-18].

Dataset	Model	Learning method	Neuron model	Network model	Depth	ANN	Converted SNN	Loss
MNIST	Diehl et al model (2015)	ANN pre-training	IF	CovNet	4	99.12%	99.14%	-0.02%
	Hu et al. model (2021)	ANN pre-training	IF	S-ResNet	8	99.59%	99.59%	0.00%
CIFAR – 10	Sengupta et al. model (2019)	ANN pre-training	IF	VGG-16	16	91.70%	91.55%	0.15%
	Sengupta et al. model (2019)	ANN pre-training	IF	ResNet	20	89.10%	87.46%	1.64%
	Hu et al. model (2021)	ANN pre-training	IF	S-ResNet	44	92.85%	92.37%	0.48%
	Han et al. model (2020)	ANN Pre-training	RMP	S-ResNet	20	91.47%	91.36%	0.11%
	Han et al. model (2020)	ANN Pre-training	RMP	VGG-16	16	93.63%	93.63%	0.00%

DOI: 10.54254/2755-2721/20/20231096

Table 1. (Continued)

	Deng et al. model (2022)	Activation gradient	LIF	ResNet	19	94.97%	94.50%	0.47%
CIFRA – 100	Sengupta et al. model (2019)	ANN Pre-training	IF	VGG-16	16	71.22%	70.77%	0.45%
	Sengupta et al. model (2019)	ANN Pre-training	IF	ResNet	20	68.72%	64.09%	4.63%
	Han et al. model (2020)	ANN Pre-training	RMP	S-ResNet	34	70.64%	69.89%	0.75%
	Han et al. model (2020)	ANN Pre-training	RMP	VGG-16	16	71.22%	70.93%	0.29%
	Deng et al. model (2022)	Activation gradient	LIF	ResNet	19	75.35%	74.72%	0.63%
ImgaeNet	Sengupta et al. model (2019)	ANN Pre-training	IF	VGG-16	16	70.52%	69.96%	0.56%
	Sengupta et al. model (2020)	ANN Pre-training	IF	ResNet	20	70.69%	65.47%	5.22%
	Hu et al. model (2021)	ANN pre-training	IF	S-ResNet	34	27.12%	28.39%	-1.27%
	Hu et al. model (2021)	ANN pre-training	IF	S-ResNet	50	24.55%	27.25%	-2.70%
	Han et al. model (2020)	ANN Pre-training	RMP	S-ResNet	34	70.64%	69.89%	0.75%
	Han et al. model (2020)	ANN Pre-training	RMP	VGG-16	16	73.49%	73.09%	0.40%
	Deng et al. model (2022)	Activation gradient	LIF	ResNet	34		64.79%	

Table 1. The result on static database, including MNIST, CIFAR-10/100, and ImageNet. The table included the accuracy between ANN and ANN after being converted to SNN. Dataset: This column lists the specific datasets that have been used to train and test the models. Model specifies the particular model being used. Learning method describes the learning algorithm used for training the model, such as backpropagation, gradient descent, etc. Neuron model indicates the specific neuron model employed in the network. Network model denotes the type of architecture used for the network like, feedforward, convolutional, recurrent, etc.

Due to the simplicity of MNIST, the recent accuracy reports on MNIST are pretty high. The accuracy between the converted SNN and ANN is almost equal, with an overall accuracy > 99%. In recent studies, because of its complexity compared to MNIST, CIFAR-10/100 and ImageNet have been used to evaluate deep SNNs. More Recent results focused on ANN pre-training and activation gradient approaches. By improving and modifying the neuron model, the performance of SNN becomes more accurate. Even though using ANN pre-training as the learning method is better than those gradient-based methods. ANN pre-training methods are used to implement more steps. Though ANN pre-training methods are better than activation gradient-based methods in terms of accuracy, ANN pre-training methods need more time steps for inference. The different regularizations also influenced the performance. In this situation, regularization, like dropout regularization, prevents a limited subset of neurons from dominating the others, reducing overfitting [6].

DOI: 10.54254/2755-2721/20/20231096

3. Future work

Supervised learning of SNNs is an important research field with a decent amount of research results. Because of SNNs' biological structure, there are many potential learning mechanisms and algorithms that need to be explored. By analyzing the current performance of SNNs, it can be predicted the major problems that will need to be solved in the future.

3.1. Large–scale SNNs

Since there is a need for real-time simulation and accurate modeling of neuronal dynamics , simulating large-scale SNNs can be challenging. It is necessary to convert the formal datasets to event-driven datasets [19]. Although plenty of studies have tried to improve the performance of deep SNN architecture, it is still very challenging to build an SNN with a large-scale deep architecture for obtaining high accuracy [15, 20]. In the future, the development of large-scale or deep spiking neural networks could focus on how to improve the computation cost for hardware and accuracy.

3.2. Recurrence connection in SNNs

SNNs just partially support recurrent and feedback interactions, which are essential to many applications like speech recognition and video processing [21]. The absence of appropriate spiking neuron models that can accurately represent the dynamics of recurrent networks is the cause of this problem. In addition to the benefits of feed-forward SNNs and recurrent SNNs, it is crucial to build effective hybrid SNNs [22]. In most cases, the recurrent architecture in hybrid SNNs is liquid state machines (LSM) [23]. So far, there have been few attempts at supervised learning algorithms for hybrid SNNs with dynamic recurrent architecture [5]. It is necessary to study the dynamic characteristics of recurrent architecture in hybrid SNNs and develop algorithms for large-scale SNNs.

4. Large-scale benchmark datasets

The lack of large-scale benchmark datasets limits the development and evaluation of SNN models. Most studies still evaluate the performances of SNNs on static-image datasets, like CIFAR and ImageNet [24-25]. Inspired by the biological visual system, several neuromorphic datasets for neuromorphic study have been gathered. Although several datasets, such as the Neuromorphic MNIST (N-MNIST) and event-based vision that is derived from the CIFAR-10 dataset (CIFAR10-DVS), have been developed for SNNs, they are relatively small compared to traditional datasets such as ImageNet [25]. In the future, not only the deep spiking neural network architecture should be developed, but more suitable evaluation datasets should also be collected.

5. Conclusion

For relatively small and simple datasets like MNIST and CIFAR data sets shown in Table 1, the performance of SNNs is approximated to have equal high accuracy with ANNs. The datasets are much less complex than those used in deep ANN applications. However, there has been recent success in using SNNs to train ImageNet, as shown in Table 1.

The biological plausibility of the model can be increased by the performance triumphs of more current conversion techniques, which will result in a model that behaves more like genuine brain neurons. However, for deep SNNs, using local learning methods, such as STDP, which is bioplausible, and taking advantage of the inherent sparsity of spiking neural networks to train large-scale networks remains a deep challenge.

Future deep spiking neural network architecture must solve a number of SNN's drawbacks and difficulties. They include the difficulty in utilizing the intrinsic features of SNNs, the constrained scale of SNNS, difficulties with hardware implementation, the lack of substantial benchmark datasets, and the restricted support for complicated computations.

References

[1] Maass, W. (1997). Networks of spiking neurons: The third generation of neural network models. Neural Networks, 10(9), 1659–1671. https://doi.org/10.1016/s0893-6080(97)00011-7
[2] Zhang, X., Xu, Z., Henriquez, C., and Ferrari, S. (2013). Spike-based indirect training of a spiking neural network-controlled virtual insect, in Decision and Control (CDC), 2013 IEEE 52nd Annual Conference on (Florence: IEEE), 6798–6805.
[3] Bohte, S. M. (2004). The evidence for neural information processing with precise spike- times: A survey. Natural Computing, 3, 195-206.
[4] Lee, J. H., Delbruck, T., & Pfeiffer, M. (2016). Training deep spiking neural networks using backpropagation. Frontiers in neuroscience, 10, 508.
[5] Wu, J., Yılmaz, E., Zhang, M., Li, H., & Tan, K. C. (2020). Deep spiking neural networks for large vocabulary automatic speech recognition. Frontiers in neuroscience, 14, 199.
[6] Nunes, J. D., Carvalho, M., Carneiro, D., & Cardoso, J. S. (2022). Spiking neural networks: A survey. IEEE Access, 10, 60738-60764.
[7] Gerstner, W., Kistler, W. M., Naud, R., & Paninski, L. (2014). Neuronal dynamics: From single neurons to networks and models of cognition. Cambridge University Press.
[8] Lee, C., Sarwar, S. S., Panda, P., Srinivasan, G., & Roy, K. (2020). Enabling Spike-Based Backpropagation for Training Deep Neural Network Architectures. Frontiers in neuroscience, 14, 119. https://doi.org/10.3389/fnins.2020.00119
[9] Rueckauer, B., Lungu, I., Hu, Y., Pfeiffer, M., & Liu, S. (2017). Conversion of Continuous-Valued Deep Networks to Efficient Event-Driven Networks for Image Classification. Frontiers in Neuroscience, 11. https://doi.org/10.3389/fnins.2017.00682
[10] Pfeiffer, M., & Pfeil, T. (2018). Deep Learning With Spiking Neurons: Opportunities and Challenges. Frontiers in Neuroscience, 12. https://doi.org/10.3389/fnins.2018.00774
[11] LeCun, Y., Cortes, C., & Burges, C. (1998). The MNIST database of handwritten digits.
[12] Krizhevsky, A., & Hinton, G. (2009). Learning multiple layers of features from tiny images. Technical report, University of Toronto.
[13] Deng, L., Wang, G., Li, G., Li, S., Liang, L., Zhu, M., ... & Shi, L. (2020). Tianjic: A unified and scalable chip bridging spike-based and continuous neural computation. IEEE Journal of Solid-State Circuits, 55(8), 2228-2246.
[14] Diehl, P. U., Neil, D., Binas, J., Cook, M., Liu, S. C., & Pfeiffer, M. (2015, July). Fast-classifying, high-accuracy spiking deep networks through weight and threshold balancing. In 2015 International joint conference on neural networks (IJCNN)(pp. 1-8). ieee.
[15] Hu, Y., Tang, H., & Pan, G. (2021). Spiking deep residual networks. IEEE Transactions on Neural Networks and Learning Systems.
[16] Sengupta, A., Ye, Y., Wang, R., Liu, C., & Roy, K. (2019). Going deeper in spiking neural networks: VGG and residual architectures. Frontiers in neuroscience, 13, 95.
[17] Han, B., Srinivasan, G., & Roy, K. (2020). Rmp-snn: Residual membrane potential neuron for enabling deeper high-accuracy and low-latency spiking neural network. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 13558-13567).
[18] Deng, S., Li, Y., Zhang, S., & Gu, S. (2022). Temporal efficient training of spiking neural network via gradient re-weighting. arXiv preprint arXiv:2202.11946.
[19] Furber, S. B., Lester, D. R., Plana, L. A., Garside, J. D., Painkras, E., Temple, S., & Brown, A. D. (2012). Overview of the SpiNNaker system architecture. IEEE transactions on computers, 62(12), 2454-2467.
[20] Bouvier, M., Valentian, A., Mesquida, T., Rummens, F., Reyboz, M., Vianello, E., & Beigne, E.(2019). Spiking neural networks hardware implementations and challenges: A survey. ACM Journal on Emerging Technologies in Computing Systems (JETC), 15(2), 1-35.
[21] Wu, Y., Deng, L., Li, G., Zhu, J., & Shi, L. (2018). Spatio-temporal backpropagation for training high-performance spiking neural networks. Frontiers in neuroscience, 12, 331.

DOI: 10.54254/2755-2721/20/20231096

[22] Esser, S. K., Appuswamy, R., Merolla, P., Arthur, J. V., & Modha, D. S. (2015). Backpropagation for energy-efficient neuromorphic computing. Advances in neural information processing systems, 28.
[23] Bellec, G., Salaj, D., Subramoney, A., Legenstein, R., & Maass, W. (2018). Long short-term memory and learning-to-learn in networks of spiking neurons. Advances in neural information processing systems, 31.
[24] Yi, Z., Lian, J., Liu, Q., Zhu, H., Liang, D., & Liu, J. (2023). Learning rules in spiking neural networks: A survey. Neurocomputing, 531, 163-179.
[25] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee.