4-17-2019

Deep Spiking Neural Networks: Study on the MNIST and N-MNIST Data Sets

Ruthvik Vaila
Boise State University

John Chiasson
Boise State University

Vishal Saxena
University of Idaho

Deep Spiking Neural Networks

Study on the MNIST and N-MNIST data sets

Ruthvik Vaila & John Chiasson
Boise State University, ECE dept.
Vishal Saxena
University of Idaho, ECE dept.

Contact Information: Ruthvik Vaila
ECE Department
Boise State University
1910 University drive, Idaho, USA
Phone: +1 209-689-6867
Email: ruthvikvaila@boisestate.edu

BOISE STATE UNIVERSITY

Introduction

Deep learning, i.e., the use of deep convolutional neural networks (DCNN), is a powerful tool for pattern recognition (image classification) and natural language (speech) processing. Deep convolutional networks use multiple convolution layers to learn the input data. They have been used to classify the large data set Imagenet with an accuracy of 96.6%. Spiking neural networks are biologically inspired in that the communication and learning algorithms are biologically plausible. In this work deep spiking networks are considered.

Spike Timing Dependant Plasticity (STDP)

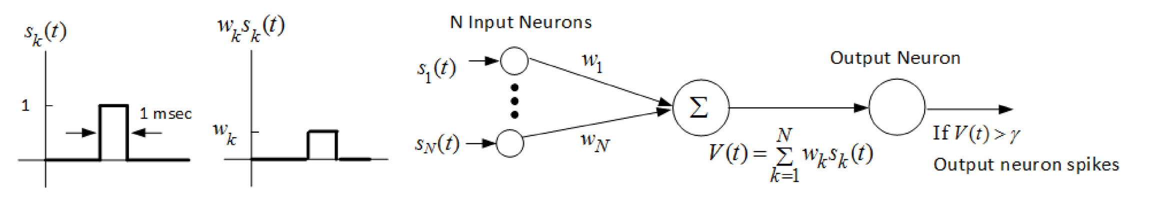

Figure 1: A simple fully connected spiking network.

- Spike timing dependant plasticity (STDP) has been shown to be able to detect hidden (in noise) patterns in spiking data [3]. Figure 1 shows a simple 2 layer fully connected network with N input (pre-synaptic) neurons and 1 output neuron.
- The spike signals $s_i(t)$ are modelled as being either 0 or 1 in one millisecond increments. That is, 1 msec pulse of unit amplitude represents a spike while a value of 0 represents no spike present. See the left side of the Figure 1.

The potentials are then summed as

$$V(t) = \sum_{i=1}^{N} w_k s_k(t). \tag{1}$$

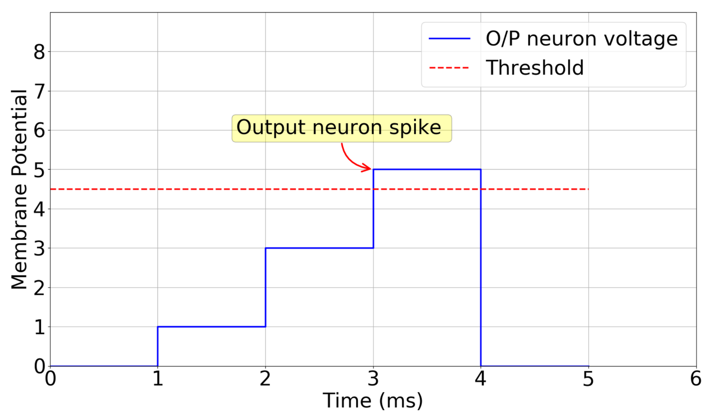

Figure 2: Spike generation by the output neuron.

- $V(t)$ is called the *membrane potential* of the output neuron. At any time t if the membrane potential $V(t)$ is greater than a specified threshold γ, then the output neuron spikes. By this we mean that the output neuron produces a 1 msec pulse of unit amplitude.
- The idea here is that the weights can be updated according to an unsupervised learning rule that results in the output spiking if and only if the fixed pattern is present. This weight update is called STDP. [1]

$$w_i \leftarrow w_i + \Delta w_i,\ \ \Delta w_i = \begin{cases} +a^+ w_i(1-w_i), & \text{if } t_{out} - t_i \le 0 \\ -a^- w_i(1-w_i), & \text{if } t_{out} - t_i > 0. \end{cases} \tag{2}$$

Here t_i and t_{out} are the spike times of the pre-synaptic (input) and the post-synaptic (output) neuron, respectively. That is, if the i^{th} input neuron spikes before the output neuron spikes then the weight w_i is increased otherwise the weight is decreased.[1]

Network and Features extracted

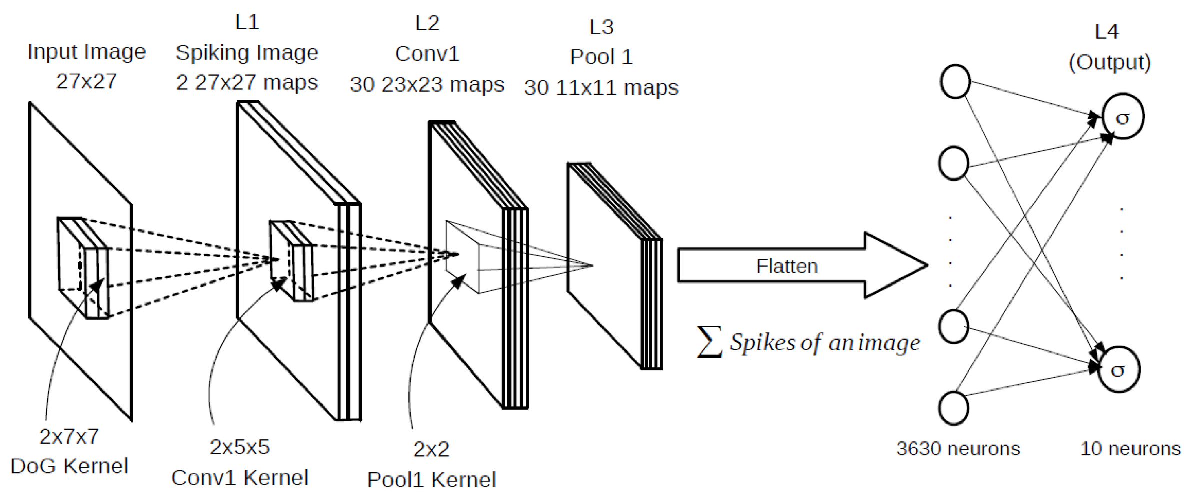

Figure 3: Deep spiking convolutional network architecture for classification of the MNIST data set.

Images in the MNIST are converted to spatio temporal spikes using rank order coding (ROC). N-MNIST data set is a recorded set images in the MNIST data set using ATIS, a silicon retina that detects changes in the pixel intensity.

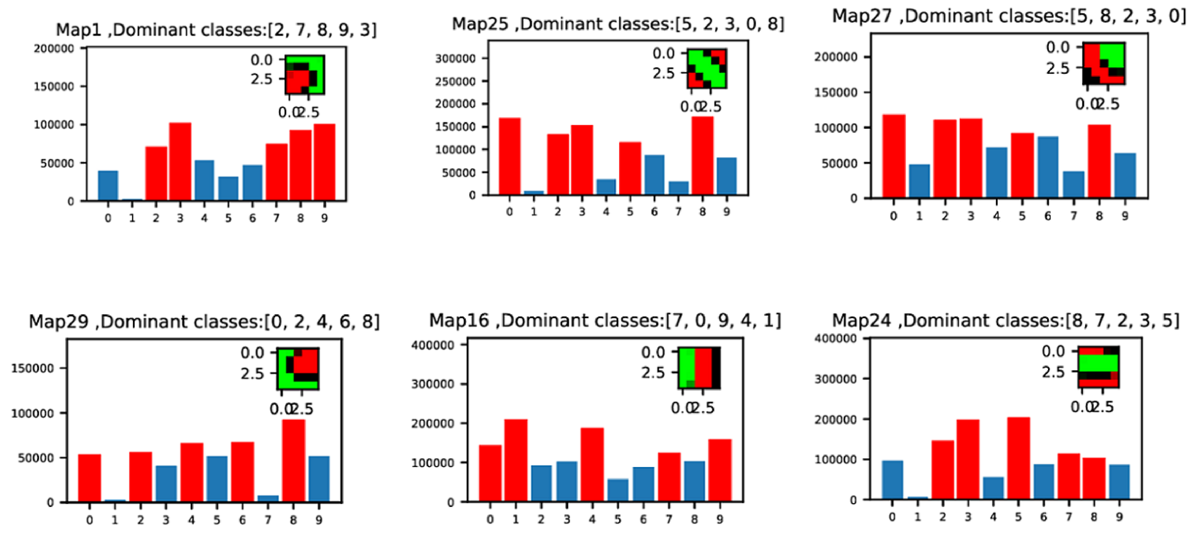

Figure 4: Spikes per map per digit. Headings for each of the sub-plots indicate the dominant (most spiking) digit for respective features.

Training algorithms for L4 layer

We used a simple two layer back propagation algorithm to perform classification of the spike vectors collected in layer L3. The gradient of a quadratic cost $C = \sum_{i=1}^{n_{0ut}}(y - a^{L4})^2$ gives the error from the last layer as

$$\delta^{L4} = \frac{\partial C}{\partial a^{L4}} \sigma'(z^{L4}) \tag{3}$$

a^L is the activation of the neurons in the output layer, σ is the activation function and z is the net input to the output layer. The weights and biases of the last layer (L4) are updated as follows:

$$\frac{\partial C}{\partial b_j^L} = \delta_j^{L4} \tag{4}$$

$$\frac{\partial C}{\partial W_{jk}^{L4}} = a_k^{L3} \delta_j^{L4} \tag{5}$$

A simple two layer backprop is a linear classifier and it achieved an accuracy of 88% [2] on the MNIST data set. We show in the later sections that a spiking convolutional network combined with a two layer backprop can achieve a classification accuracy of 98.4% on the MNIST data set.

Catastrophic forgetting

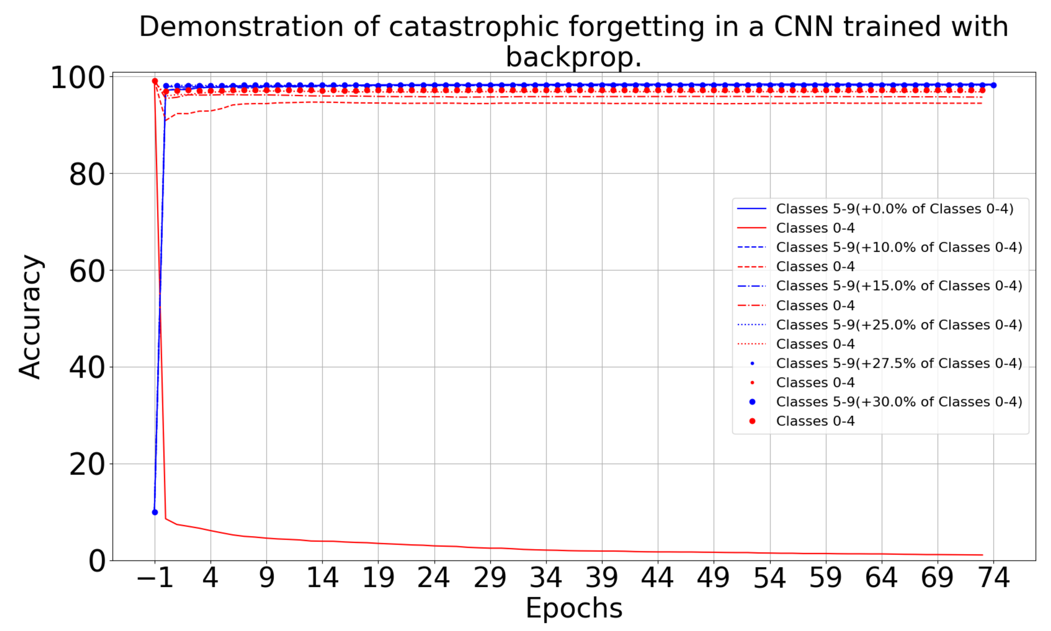

Figure 5: Catastrophic forgetting in a convolutional network while revising a fraction of the previously trained classes. Note that epoch -1 indicates that the network was tested for validation accuracy before training of the classes 5-9 started. Brackets in the legend shows the fraction of previously trained classes that were used to revise the weights from the previous classes.

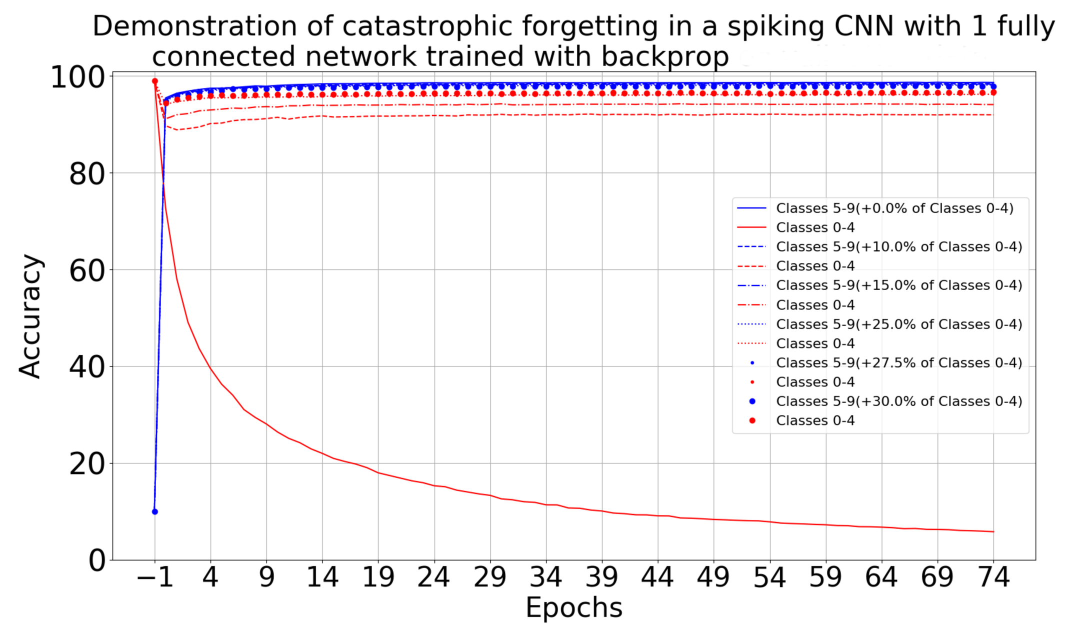

Figure 6: Catastrophic forgetting in a spiking convolutional neural networks. Note that the solid red line in this plot indicates that he forgetting in spiking networks is not catastrophic.

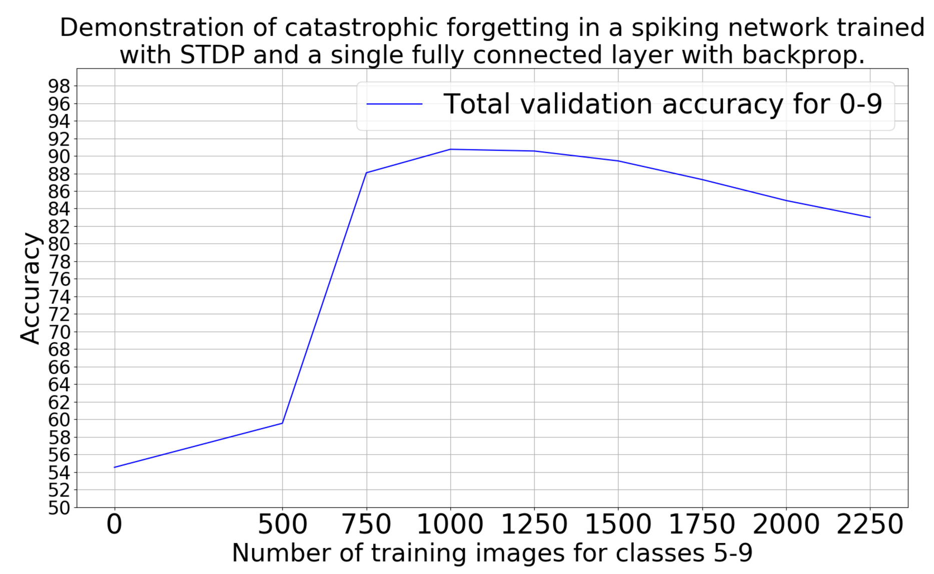

Figure 7: Note that as the number of training images for the classes 5-9 increases the total accuracy drops.

Results

Saeed et al [1] used a linear SVM and an additional convolution layer in the Figure 3 and achieved an accuracy of 98.3%. Our research indicates that using a simpler two layer back propagation and a single convolution/pool layer is enough to achieve an accuracy of 98.4% on the MNIST data set.

Classifier	Test Acc.	Val Acc.	Data set
2 layer FCN	98.4%	98.5%	MNIST
SVM (RBF)	98.8%	98.87%	MNIST
SVM (linear)	98.41%	98.31%	MNIST
2 layer FCN	97.45%	97.62%	N-MNIST
SVM (RBF)	98.32%	98.40%	N-MNIST
SVM (linear)	97.64%	97.71%	N-MNIST

Table 1: Classification accuracy on the MNIST data set

Stromatias et al reported an accuracy of 97.23% accuracy by using artificially generated features for the kernels of the first convolutional layer and training a 3 layer fully connected neural network classifier on spikes collected at the first pooling layer [4]. Results for the MNIST and N-MNIST data sets are presented in the Table 1.

Conclusions

- We have shown that combining feature extraction in spiking networks when combined with a simple two layer backprop can result in 98.4% accuracy and we have also shown that training the features of the L2 layer instead of artificially generating them results in an accuracy of 97.45%.
- We have shown that spiking convolutional networks can retain up to 91% test accuracy when trained with disjoint sets.

Forthcoming Research

We plan to test our network using bigger data sets like EMNIST, Caltech 101 etc.

References

[1] Saeed Reza Kheradpisheh, Mohammad Ganjtabesh, Simon J. Thorpe, and Timothée Masquelier. STDP-based spiking deep convolutional neural networks for object recognition. *Neural Networks*, 99:56 – 67, 2018.

[2] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, Nov 1998.

[3] Timothée Masquelier, Rudy Guyonneau, and Simon J. Thorpe. Spike timing dependent plasticity finds the start of repeating patterns in continuous spike trains. *PLOS ONE*, 3(1):1–9, 01 2008.

[4] Evangelos Stromatias, Miguel Soto, Teresa Serrano-Gotarredona, and Bernab Linares-Barranco. An event-driven classifier for spiking neural networks fed with synthetic or dynamic vision sensor data. *Frontiers in Neuroscience*, 11:350, 2017.

Acknowledgements

We would like to express our deep gratitude to Professor Timothe Masquelier and Dr. Saeed Reza Kheradpisheh for answering our many questions about their work [1].

[1] The input neuron is assumed to have spiked *after* the output neuron spiked.