STDP-based spiking deep convolutional neural networks for object recognition

Neural Networks

Edoardo Ghini
Gianluca Cerilli

Sapienza Università di Roma

Table of contents

Introduction

Introduction

Spiking Neural Networks communicate through *spikes*, that are discrete events that occur at certain points in time.

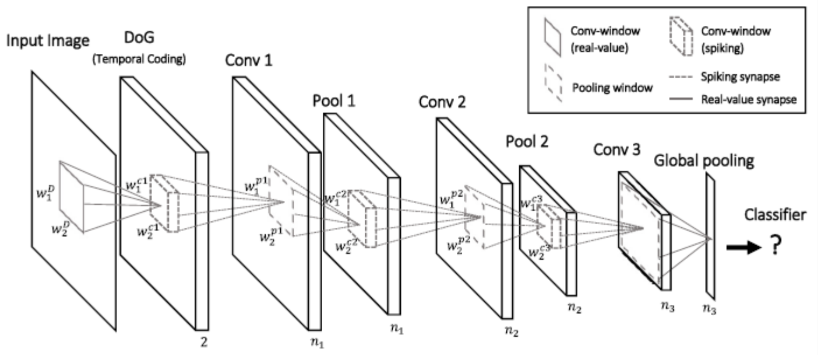
When a neuron reaches a certain potential value, it spikes, so it resets its value.

This modelling allows us to handle spatio-temporal data, that are real-world sensory data.

Network architecture

Architecture

STDP-based spiking deep neural network (SDNN) with a spike-time neural coding:



Temporal coding

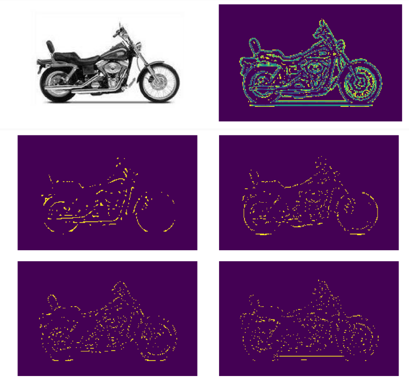
Input signal is encoded into temporal-discrete spike events.

- A Difference of Gaussians (DoG) filter detects the contrasts in the input image and emit a spike, accordingly.
- The higher the contrast in a cell, the more strongly this one is activated in order to fire.
- The firing time $\tau = 1/r$ of a cell is inversely proportional to its activation value r.

Temporal coding

DoG temporal coding (human face)

DoG temporal coding (motorbike)



Convolutional layers

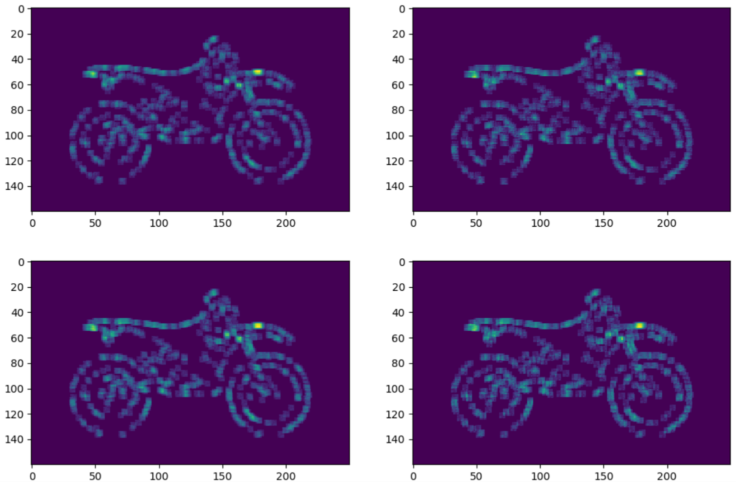
Several neuronal maps detect visual features at different locations.

- Same synaptic weights assigned to the same neuronal maps
- Neurons generate spikes $S(t)$ according to the output of the pre-synaptic neurons and when their potential $V(t)$ exceeds a certain threshold

$$V_i(t) = V_i(t-1) + \sum_j W_{j,i} S_j(t-1)$$

- After it fires, the neuron can not fire again and it also inhibits other neurons in its same location

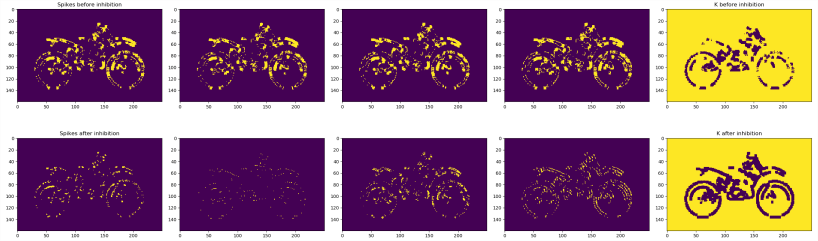
Convolutional layers



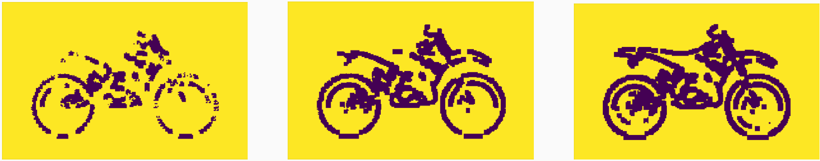
Output of the first convolutional layer

Convolutional layers

Spiking neurons before (top) and after (bottom) inhibition



Inhibition matrix for the first three outputs of the DoG filter

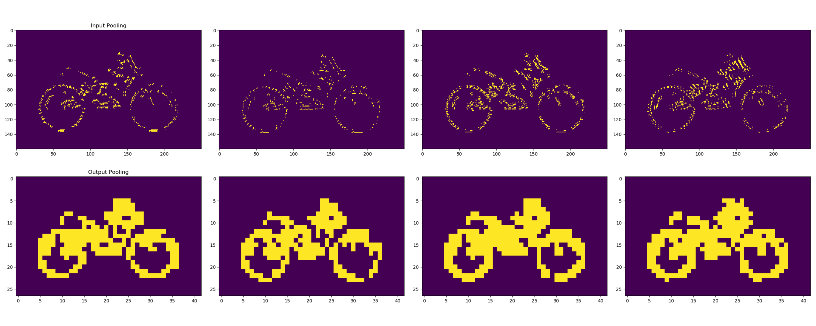


Pooling layers

Neurons perform a max pooling operation over a window in the corresponding neuronal map of the previous layer.

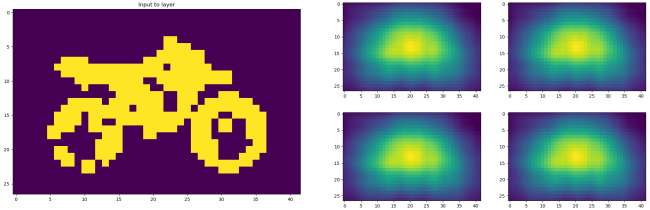
- No learning occurs
- As before, each neuron can emit only one spike.
- Visual information is compressed

Pooling layers



Input (top) and output (bottom) of the first pooling layer.

2nd convolutional layer



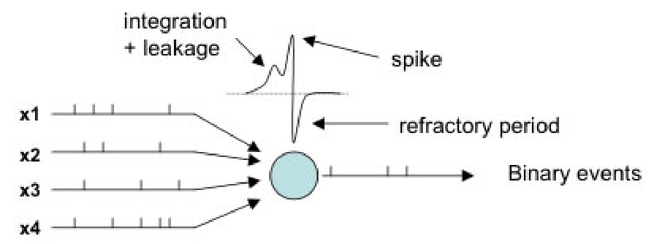
Output of the first pooling layer (left) and of the second convolutional layer (four images)

The image is not recognizable any more after the second convolution.

STDP-based learning

Spike-timing-dependent plasticity (STDP) is a biological process used by brain to modify its own synapses.

Here, it allows to strengthen (weaken) synapses if they contribute (or not) to the firing of a post-synaptic neuron.



STDP-based learning

Learning occurs only in convolutional layers, so that neurons compete to fire earlier than others and use STDP to learn input patterns.

STDP is represented by:

$$\Delta w_{ij} = \begin{cases} a^{+} w_{ij}(1 - w_{ij}), & if \;\; t_j - t_i \leq 0, \\ a^{-} w_{ij}(1 - w_{ij}), & if \;\; t_j - t_i > 0, \end{cases}$$

where Δw_{ij} is the synaptic weight modification and a^{+} and a^{-} are the learning rate parameters.

STDP-based learning

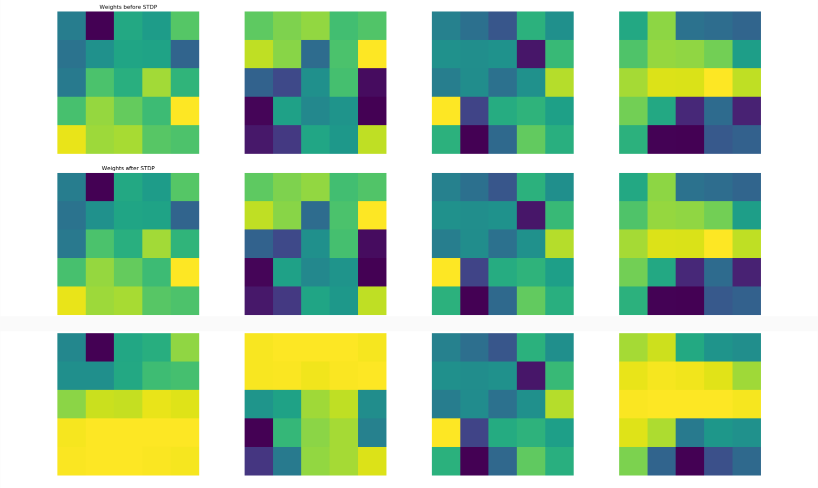
Weights are randomly initiated with $\mu = 0.8$ and $\sigma = 0.05$.

The learning convergence of the *l-th* convolutional layer is represented by:

$$C_l = \sum_f \sum_i w_{f,i}(1 - w_{f,i})/n_w$$

where n_w is the total number of synaptic weights in that layer and $w_{f,i}$ is the *i-th* synaptic weight of the *f-th* feature.

STDP-based learning



Weights randomly initialised (top) and updated after one (middle) and three images processed (bottom).

Classification

The output of the global max pooling layer is used to build a dataset of labelled features which are split in a training and testing set for the classifier.

Features are the greatest potentials reached by the neuronal maps in the third convolutional layer.

The resulting dataset is standardised and used to feed a linear SVM.

Conclusion

Summary

- New *spiking* neuron models used to handle a bio-inspired neural network.
- Spatio-temporal information acquired in a better way and good results obtained with few input data
- 88% of accuracy with some existing functions
  62% - 75% (depending on the parameters and the classifier) with every function implemented from scratch

Possible improvements

- Modify the STDP, in order to apply penalty on more neurons not involved in the spiking process
- Extend the encoding time of the DoG filter to have a better differentiation between useful and useless information (side-effect of computational effort)

References

- S. R. Kheradpisheh, M. Ganjtabesh, S. J. Thorpe, and T. Masquelier, "Stdp-based spiking deep convolutional neural networks for object recognition," Neural Networks, vol. 99, pp. 56–67, 2017.
- Spiking Neural Networks, the Next Generation of Machine Learning, https://towardsdatascience.com/spiking-neural-networks-the-next-generation-of-machine-learning-84e167f4eb2b
- Spiking-Neural-Network, https://github.com/Shikhargupta/Spiking-Neural-Network