IEEE WCCI 2020
IEEE World Congress on Computational Intelligence
Virtual Conference – July 19-24, 2020

Implementing a foveal-pit inspired filter in a Spiking Convolutional Neural Network: a preliminary study

Shriya T.P. Gupta [1] and Basabdatta S. Bhattacharya [2]

Department of Computer Science and Information Systems,
BITS Pilani Goa Campus, Goa, India

email: [1] shriyatp99@gmail.com , [2] basabdattab@goa.bits-pilani.ac.in

SPONSORS:

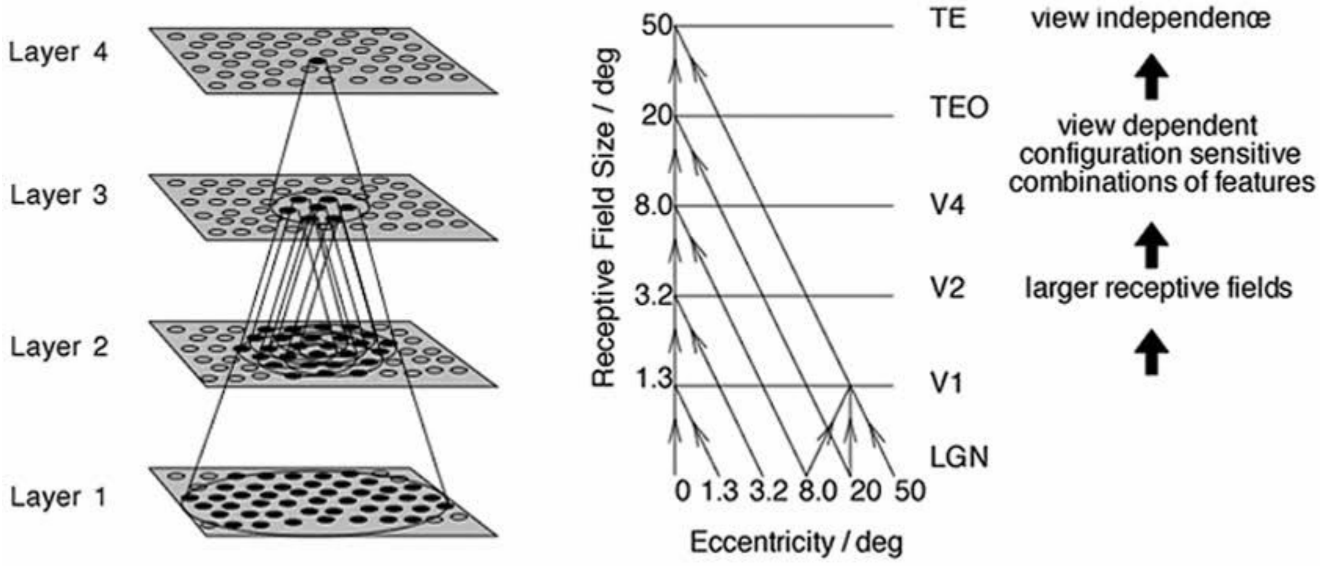
IEEE Advancing Technology for Humanity
IEEE Computational Intelligence Society
IET
The International Neural Network Society (INNS)
EPS Evolutionary Programming Society

Outline

- Motivation
- Related Work
- Datasets
- Proposed Approach
- Results and Discussion
- Conclusions

Motivation

The primate visual pathway and its hierarchical organization



Source: Wikipedia images

Related Work

SCNN with DoG filters

- Diehl et al [1] - two-layer fully connected SCNN, trained using STDP
- Kheradphisheh et al [2] - SCNN with multiple convolutional and pooling layers composed of LiF neurons.

Modifications in backpropagation for SCNNs

- Lee et al [3] have developed a semi-supervised training framework by using a unsupervised STDP rule for pre-training the network, and then using the supervised gradient descent algorithm for fine tuning the model.
- Hunsberger et al [4] have proposed a backpropagation training of spiking neurons, which is implemented in the Nengo library.

[1] Diehl et al.: "Unsupervised learning of digit recognition using spike timing dependent plasticity." Frontiers in computational neuroscience.
[2] Kheradpisheh et al.: " Stdp - based spiking deep convolutional neural networks for object recognition.", Neural Networks.
[3] Lee et al.: "Training deep SCNNs with stdp based unsupervised pre-training followed by supervised fine-tuning", Frontiers in neuroscience.
[4] Hunsberger et al.: "Training spiking deep networks for neuromorphic hardware", arXiv preprint.

Datasets

Digit Recognition

- The MNIST digit recognition database [6] - benchmark for image recognition tasks
- For the noise-free scenario - as they only contain digits without any associated background
- Each image is of size 28×28 and represents a digit between 0 to 9
- Has 55,000 training images and 10,000 testing images

Vehicle Recognition

- Sampled a subset of Caltech 101 image dataset [7] belonging to the motorbike and airplane classes
- For the noisy case - corresponding background details of real world images
- This subset has roughly 800 images from each class with varying dimensions
- For purposes that suit the filter dimension in this work, images are resized to 150×200 pixels

[6] LeCun et al., "Gradient-based learning applied to document recognition," Proceedings of the IEEE,vol. 86, no. 11, pp. 2278–2324, 1998.
[7] Fei-Fei et al., "Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories," in Conference on Computer Vision and Pattern Recognition Workshop. IEEE, 2004, pp. 178–178.

Proposed Approach

Foveal-pit inspired neural filters

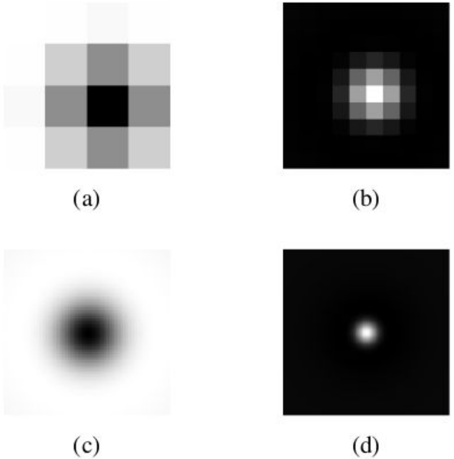


Fig. 2. The ganglion cells modelled using DoG functions representing the (a) off-center midget cell (b) on-center midget cell (c) off-center parasol cell and (d) on-center parasol cell.

Ganglion Cell Types		Receptive Field Simulation Parameters: Matrix Size (n)	Std. Dev. Center (σ_c)	Center Width in Pixels (w_c)	Std. Dev. Surround (σ_s)	Sampling Resolution
Midget	OFF-center	5×5	0.8	3	$6.7 \times \sigma_c$	$\frac{1}{\sqrt{2}}$
	On-center	11×11	1.04 $\simeq (1.3 \times 0.8)$	5		
Parasol	Off-center	61×61	8 $\simeq (10 \times 0.8)$	33	$4.8 \times \sigma_c$	$\frac{5}{\sqrt{2}}$
	On-center	243×243	10.4 $\simeq (10 \times 1.04)$	53		

Source: Bhattacharya et al.. "Biologically inspired means for rank-order encoding images: A quantitative analysis." IEEE transactions on neural networks.

Rank ordering

$$Contrast_{\mathrm{Im}}(x, y, s) = \sum_i \sum_j (\mathrm{Im}(i + x, j + y) \cdot DoG_s(i, j))$$

Highest pixel (coefficient) value corresponds to the first ganglion cell to have fired - carries the most weight (importance)

Reconstruction

$$\mathrm{Im}_{\mathrm{Rec}}(i, j) = \sum_x \sum_y \sum_s Contrast_{\mathrm{Im}}(x, y, s) \cdot DoG_s(x - i, y - j)$$

Source: R. V. Rullen and S. J. Thorpe, “Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex,”Neural computation, vol. 13, no. 6, pp. 1255–1283, 2001.

Overview of the proposed approach

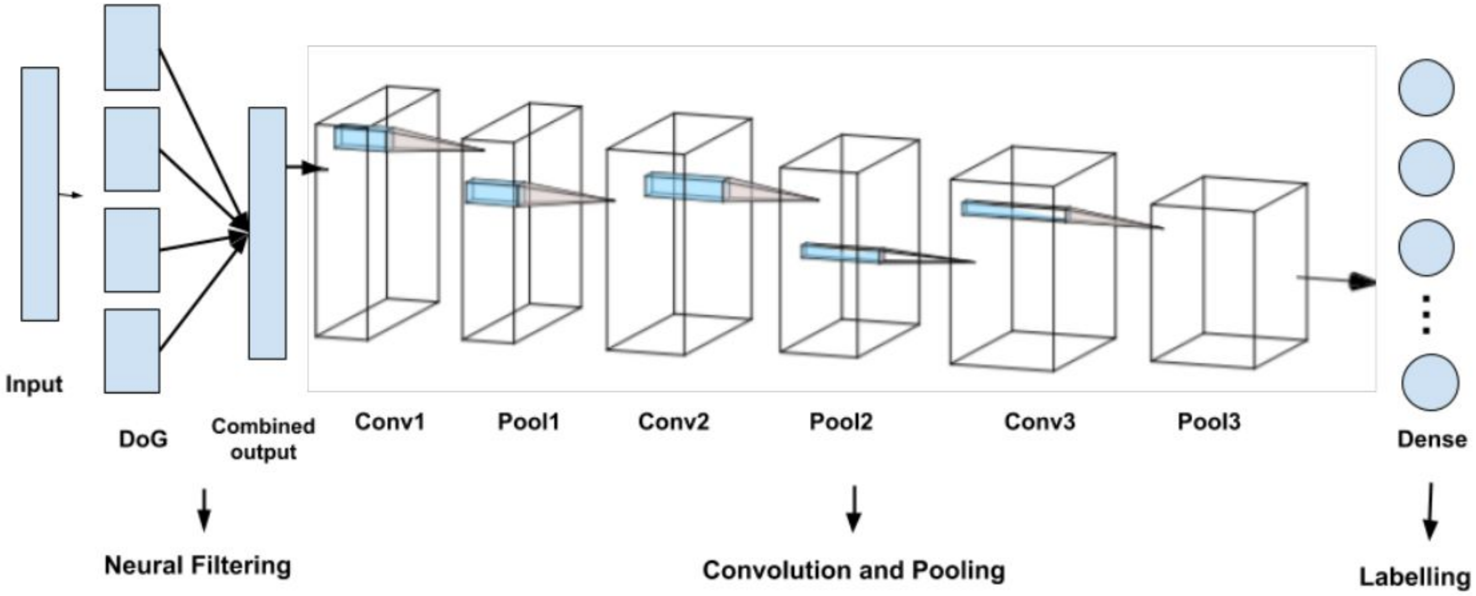
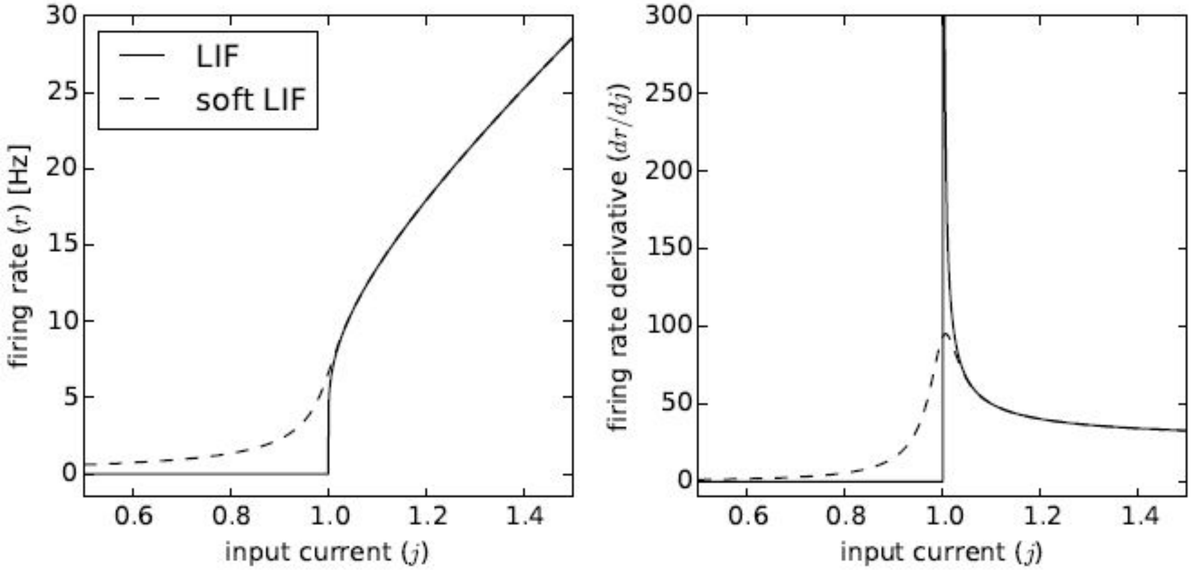


Fig. 1. The overall architecture of the proposed Spiking Convolutional Neural Network.

Spiking approximation of backpropagation for training



Source: Hunsberger et al. "Training spiking deep networks for neuromorphic hardware."

Results and Discussion

Quantitative Results

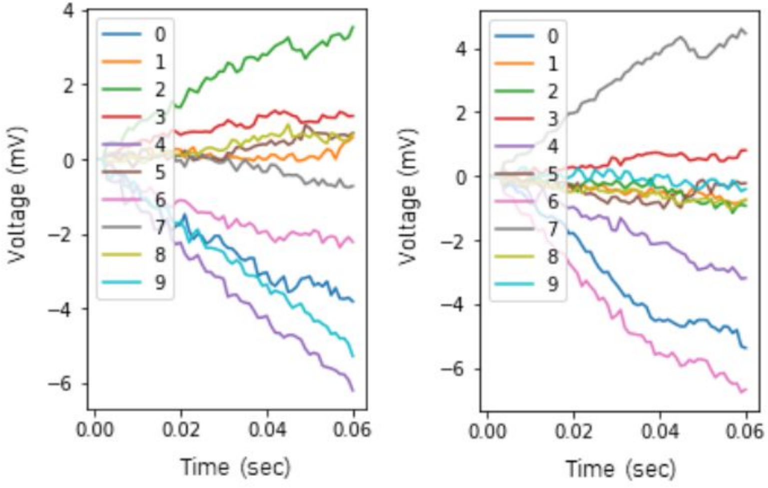
ACCURACIES (%) FOR THE DIGIT RECOGNITION TASK

Neural filter	3 epochs	6 epochs	10 epochs
off-center midget cell	98.25	99.00	96.50
on-center midget cel	97.75	98.25	96.25

ACCURACIES (%) FOR THE VEHICLE RECOGNITION TASK

Neural filter	3 epochs	5 epochs	7 epochs
without any filtering	57.50	52.50	47.50
only off-center midget cells	90.00	52.50	67.50
midget-off, midget-on, parasol-off	60.00	52.50	52.50

Detailed analysis on the MNIST dataset



(a) (b)

Fig. 4. Outputs of the final layer neurons for (a) digit 2 and (b) digit 7. The timestep of simulation is 1 ms and total time of simulation is 60ms.

COMPARISON OF RESULTS ON THE DIGIT RECOGNITION TASK

Model	Learning Rule	Accuracy (%)
Two layer SNN [1]	STDP	95.00
Convolutional SNN [18]	Backpropagation	**99.10**
Spiking Deep CNN [2]	STDP	98.40
Proposed SCNN	Backpropagation	99.00

Qualitative outputs of the progressive reconstruction

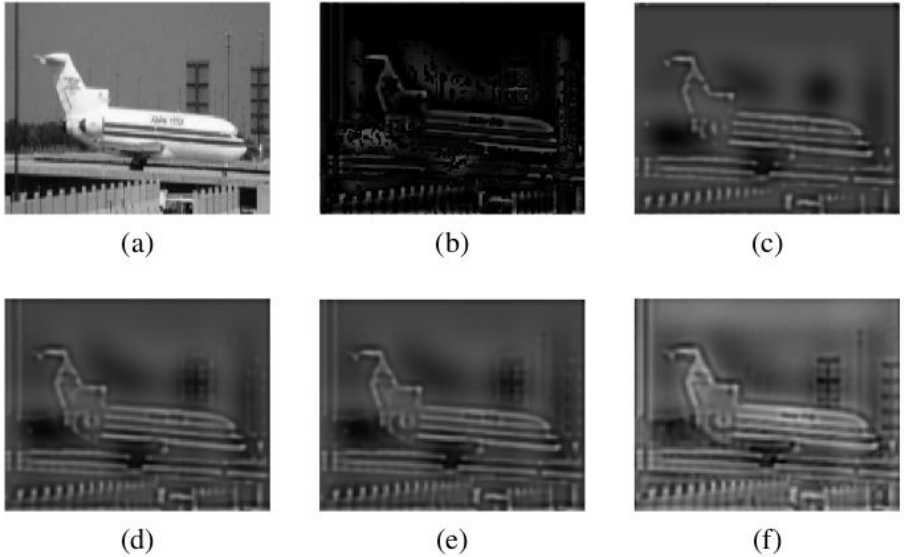


Fig. 7. The output images after the qualitative analysis. (a) The original image of the airplane (b) the rank ordered output of the airplane (c) the reconstructed image using only 10% of the coefficients (d) using only 20% of the coefficients (e) using only 30% of the coefficients and (f) using all the coefficients for the reconstruction.

Conclusions

Conclusions

- SCNN with retinal foveal-pit inspired DoG filters - improves classification in both a noise-free and a noisy environment
- Achieves an accuracy of up to 99.00% on MNIST while upto 90% on the Caltech dataset - an improvement on around 57% over the sans-filtering approach
- Importance of the bio-inspired neural filters in redundancy reduction of input images + in eliminating irrelevant background details

Future Directions

- Explored the qualitative effect of the filtering by performing a reconstruction of the rank ordered images
- Implementing lateral-inhibition based Filter-overlap Correction[8] for refining the rank ordered outputs

[8] B. Sen Bhattacharya and S. Furber, "A biologically inspired algorithm to deal with filter-overlap in retinal models," in Supplementary BMC Neuroscience. Springer Nature, 2009, p. P126

Thank you