

Multi-column Deep Neural Networks for Image Classification*

Bekir Bakar

Bursa Technical University

b.bakar@outlook.com

*Ciregan, D., Meier, U., & Schmidhuber, J. (2012, June). Multi-column deep neural networks for image classification. In *2012 IEEE conference on computer vision and pattern recognition*(pp. 3642-3649). IEEE.

Problem

- Traditional methods of computer vision and machine learning cannot match human performance tasks such as the recognition of handwritten digits or traffic signs
- Biologically plausible deep artificial neural network architectures can

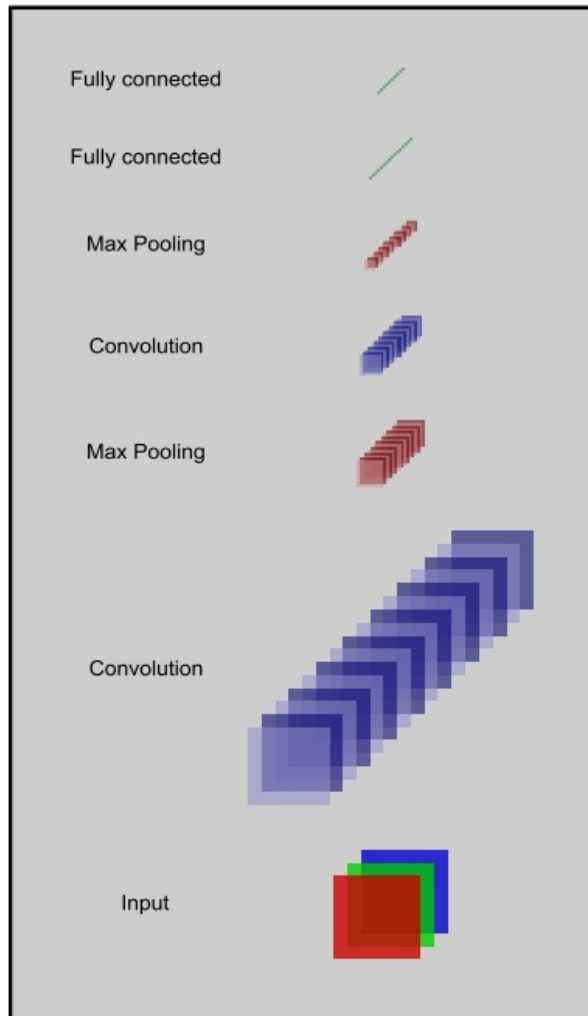
Propose

- Only winner neurons are trained. Several deep neural columns become experts on inputs preprocessed in different ways; their predictions are averaged
- On the very competitive MNIST handwriting benchmark, our method is the first to achieve near-human performance

Architecture

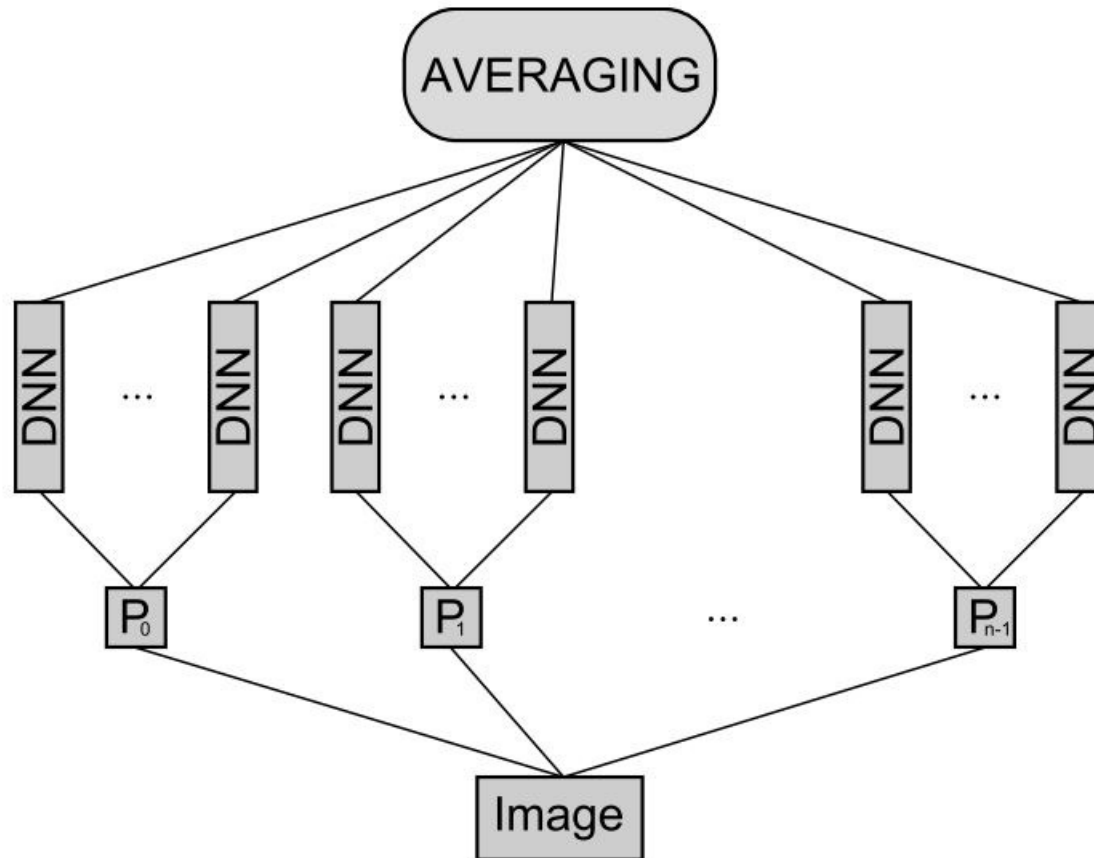
- Inspired by micro-columns of neurons in the cerebral cortex
- Given some input pattern, the predictions of all columns are democratically averaged
- Before training, the weights (synapses) of all columns are randomly initialized
- Various columns can be trained on the same inputs, or on inputs preprocessed in different ways

Architecture: DNN

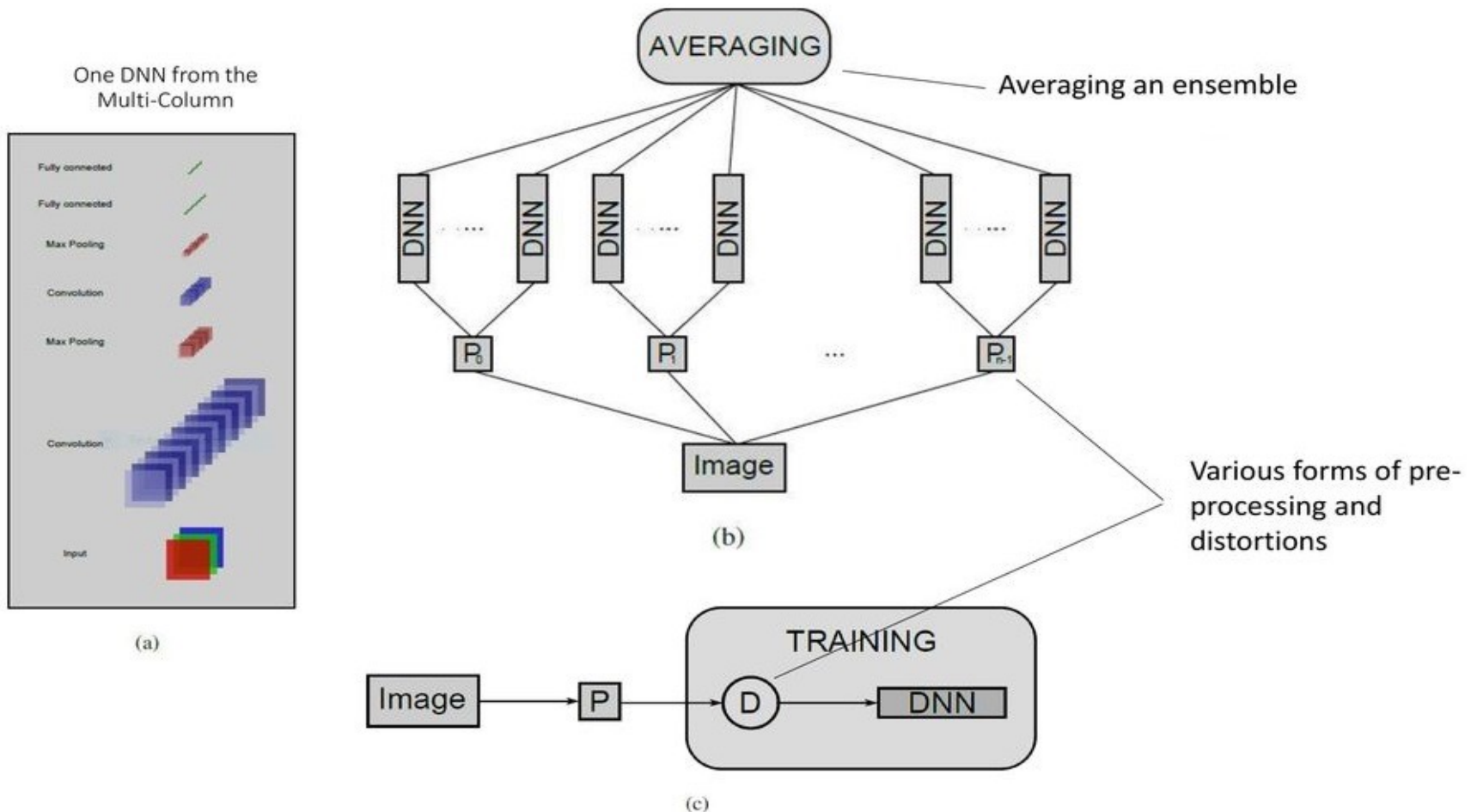


- 2x48x48-100C5-MP2-100C5-MP2-100C4-MP2-300N-100N-6N represents a net with 2 input images of size 48x48
- a convolutional layer with 100 maps and 5x5 filter
- a max-pooling layer over non overlapping regions of size 2x2
- a convolutional layer with 100 maps and 4x4 filters
- a max-pooling layer over non overlapping regions of size 2x2
- a fully connected layer with 300 hidden units
- a fully connected layer with 100 hidden units and
- a fully connected output layer with 6 neurons (one per class)

Architecture: Averaging



Architecture: MCDNN



Datasets

- MNIST
- NIST SD 19
- Chinese characters
- Traffic signs
- CIFAR 10
- NORB

MNIST

- The MNIST database is a large database of handwritten digits that is commonly used for training various image processing systems
- The original MNIST digits [20] are normalized such that the width or height of the bounding box equals 20 pixels.

Keynotes

- Proposed method is fully supervised and does not use any additional unlabeled data source
- On many other image classification datasets MCDNN improves the state-of-the-art by 30-80%
- Single DNN already are sufficient to obtain new state-of-the-art results; combining them into MCDNNs yields further dramatic performance boosts

Results

Dataset	Best result of others [%]	MCDNN[%]	Relative improvement [%]
MNIST	0.39	0.23	41
HWDB1.0	7.61	5.61	26
CIFAR10	18.50	11.21	39
Traffic Signs	1.69	0.54	72
NORB	5.00	2.70	46