On-chip Training of Deep Neural Networks Using Unsupervised and Supervised Learning Methods

Raqibul Hasan and Tarek M. Taha
Department of Electrical and Computer Engineering
University of Dayton, Ohio

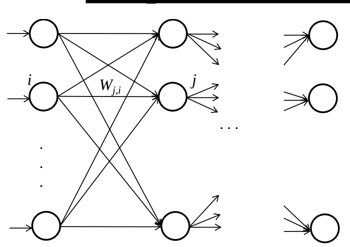
Abstract

This poster presents on-chip training circuits for memristor based deep neural networks utilizing unsupervised and supervised learning methods. Memristor crossbars are prone to device variations and faults. On chip training circuits would allow the training algorithm to account for device variability and faults in these circuits. We utilize autoencoders for layer-wise pre-training of the deep networks and utilize the back-propagation algorithm for supervised fine tuning. We demonstrate successful training of memristor based deep networks for the MNIST digit classification and the KDD intrusion detection datasets.

Objectives

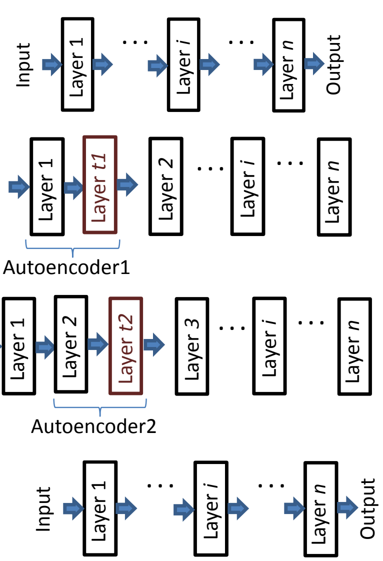
- Memristor based deep network training.
- Autoencoders for layer-wise pre-training.
- Circuit for unsupervised and supervised training.

Deep Neural Network



$$DP_j = \sum_{i=1}^{n} x_i w_{ij} \; y_j = f(DP_j)$$

Fig. 1. Deep network.



Memristor Crossbar Neural Network

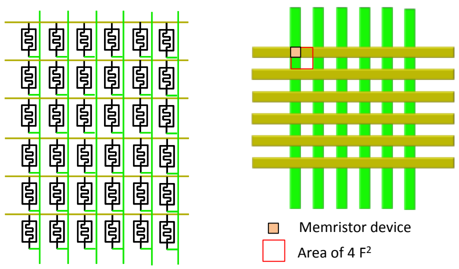
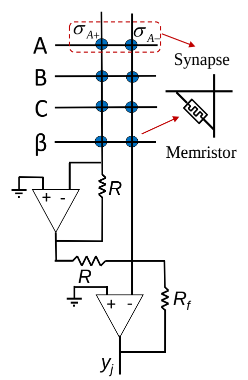


Fig. 2. Memristor crossbar.



$$y_j = R_f\,[A(\sigma_{A+} - \sigma_{A-}) + \cdots + \beta(\sigma_{\beta+} - \sigma_{\beta-})]$$

Fig. 3. Memristor based neuron circuit performing parallel analog operation.

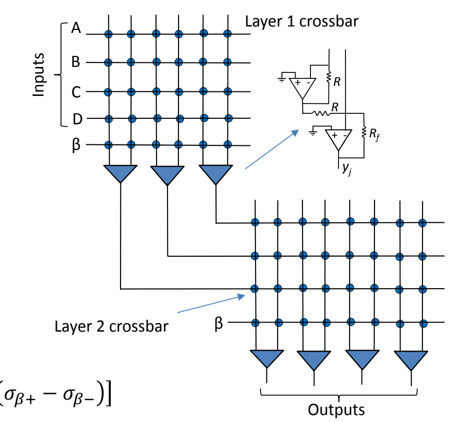


Fig. 4. Memristor based 2 layer neural network.

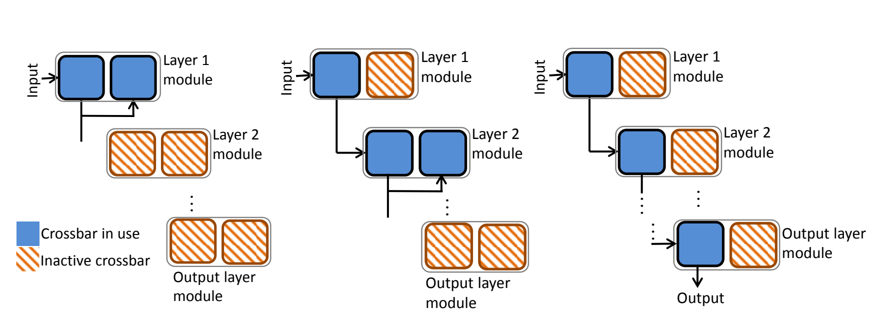
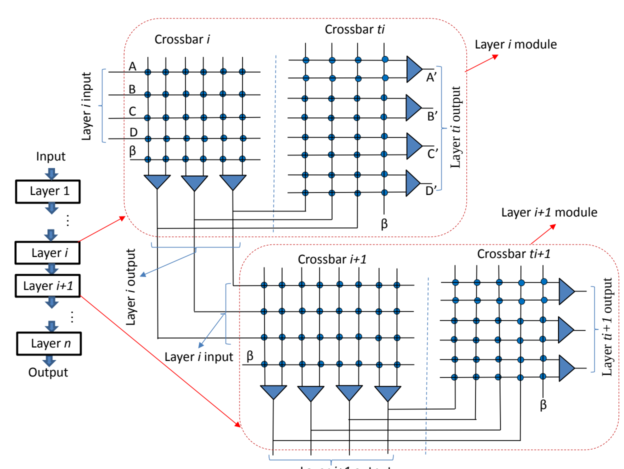


Fig. 5. Training memristor based deep network.

Memristor Based Deep Network



Experimental Evaluations

Table I: Neural network configurations.

MNIST	784→200→100→10
KDD	41→100→50→20→1

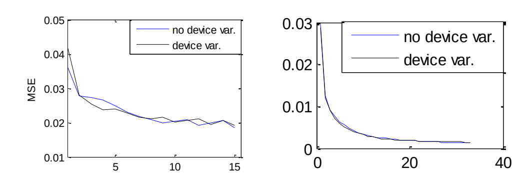


Fig. 6. Training graph for KDD (left) and MNIST (right) datasets.

Table II: Recognition error on test data.

	Recognition error (%)		
	no device var.	device var.	s/w impl.
KDD (5000 test data)	3.36	3.2	3.32
MNIST (5000 test data)	8.06	8.04	6.84

Conclusion

In this work we have designed on-chip training system for memristor based deep neural networks utilizing two memristors per synapse. We have utilized autoencoders for layer-wise pre-training of the networks and utilized the back-propagation algorithm for supervised fine tuning.