Scalable Training of Artificial Neural Networks with Adaptive Sparse Connectivity

Pronob Sarker

Gettysburg College

April 16, 2025

Introduction

Paper Details

- Title: “Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science”
- Authors: Decebal Constantin Mocanu et al.
- Journal: Nature Communications (2018)
- Intended audience: Computer scientists and machine learning researchers

The Problem

Current Limitation of Neural Networks

- **Key Question:** Can we make artificial neural networks more efficient by using sparse connectivity rather than fully-connected layers?
- **Problem:** Current ANNs use fully-connected layers requiring quadratic number of parameters, limiting scalability
- **Relevance:** As neural networks grow larger, computational and memory requirements become prohibitive
- **Observation:** Biological neural networks have sparse, scale-free topologies - not fully-connected

The SET Solution

What is SET (Sparse Evolutionary Training)?

- Replace fully-connected layers with **sparse, evolving connections**
- Start with randomly sparse layers (instead of fully-connected)
- During training, after each epoch:
  - Remove weakest connections (smallest weights)
  - Add new random connections in equal number

How SET Networks Adapt to Data

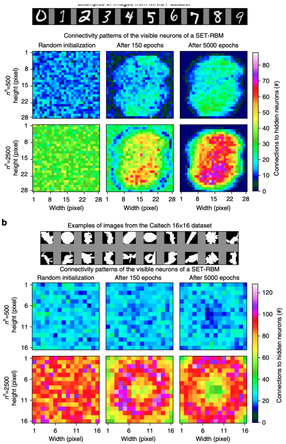

Fig. 4 SET-RBMs connectivity patterns for the visible neurons. **a** On the MNIST dataset. **b** On the Caltech 101 16 × 16 dataset. For each dataset, we have analyzed two SET-RBM architectures, i.e. 500 and 2500 hidden neurons. The heat-map matrices are obtained by reshaping the visible neurons vector to match the size of the original input images. In all cases, it can be observed that the connectivity starts from an initial Erdös–Rényi distribution. Then, during the training process, it evolves towards organized patterns which depend on the input images

Figure: Evolution of connectivity patterns in SET-RBM networks

Experiment Results & Impact

Tested across diverse datasets

- Tested on 15 datasets across various domains (biology, computer vision, physics, etc.)
- With three neural network architectures: RBMs, MLPs, and CNNs

Major findings

- SET models achieve **similar or better accuracy with only 1-2% of parameters**
- Example: SET-MLP outperforms standard MLP on CIFAR10 with $\sim 100\times$ fewer parameters

Limitations Critiques

Challenges:

- Current deep learning hardware/software optimized for dense matrix multiplication
- Implementation requires specialized sparse matrix operations that are less optimized
- Paper doesn't fully explore how different connection removal strategies might affect performance

Benefits & Future Possibilities

Benefits

- **Efficiency:** Dramatically fewer parameters
- **Performance:** Equal or better accuracy
- **Scalability:** Could enable billion-neuron networks with current hardware

Future opportunities

- Specialized hardware for sparse neural networks
- On-device training for edge computing
- Dramatically larger neural networks

Conclusion

SET offers a promising approach to building more efficient neural networks

- By mimicking biological neural network properties (sparsity, scale-freeness), we can create models that:
- Use dramatically fewer parameters
- Maintain or improve accuracy
- Future work could address hardware/software implementation challenges and enable billion-node scale neural networks

Questions?