

Fast Deep Neural Network (DNN) Training/Execution on Hardware Platforms

Tech ID: 30511 / UC Case 2019-065-0

BACKGROUND

With the growing range of applications for Deep Neural Networks (DNNs), the demand for higher accuracy has directly impacted the depth of the state-of-the-art models. Although deeper networks are shown to have higher accuracy, they suffer from drastically long training time and slow convergence speed with high computational complexity.

TECHNOLOGY DESCRIPTION

Researchers at UC San Diego have invented a network modification algorithm that takes as input a conventional CNN architecture and enforces a small-world property on its topology to generate a new network, called SWNet. The approach leverages a quantitative metric for small-worldness and devises a customized rewiring algorithm. The algorithm restructures the inter-layer connections in the in-put CNN to find a topology that balances regularity and randomness, which is the key characteristic of SWNs. Small-world properties in CNNs translates to an architecture where all layers are interlinked via sparse connections. Such networks have similar quality of prediction and number of trainable parameters as their baseline feed-forward architectures, but due to the added sparse links and the optimal SWN connectivity, they warrant better data flow. As a result, the architecture modification has three main benefits:

- ▶ it removes non-critical connections and reduces computational implications;
- ▶ it increases the degrees of freedom during training, allowing faster convergence; and
- ▶ it provides customized data paths in the model for better cross-layer information propagation.

APPLICATIONS

The disclosed technology can enhance the efficiency of any DNN, as used in image recognition, autonomous vehicles and robotics, etc.

ADVANTAGES

During benchmarking experiments on various network architectures, SWNets' performance on popular image classification benchmarks including CI-FAR10, CIFAR100, and ImageNet achieve an average of 2.1-fold improvement in training iterations required to achieve comparable classification accuracy as the baseline models. Comparisons of SWNet with the state-of-the-art DenseNet model shows that with 10×fewer parameters, SWNets demonstrate identical performance during training.

STATE OF DEVELOPMENT

Working software has been developed and is available for evaluation.

INTELLECTUAL PROPERTY INFO

The idea is patent pending and available for licensing.

RELATED MATERIALS

- ▶ Mojan Javaheripi, Bitra Darvish Rouhani, Farinaz Koushanfar. SWNet: Small-World Neural Networks and Rapid Convergence. 4/9/2019 eprint arXiv:1904.04862 - 04/09/2019

PATENT STATUS

Country	Type	Number	Dated	Case
United States Of America	Published Application	0125960-A1	04/23/2020	2019-065

CONTACT

University of California, San Diego
Office of Innovation and Commercialization
innovation@ucsd.edu
tel: 858.534.5815.



OTHER INFORMATION

KEYWORDS

Deep Neural Networks, DNN, Machine Learning, AI, Small world networks

CATEGORIZED AS

- ▶ Computer
 - ▶ Hardware
 - ▶ Software

RELATED CASES

2019-065-0

University of California, San Diego

Office of Innovation and Commercialization

9500 Gilman Drive, MC 0910, ,

La Jolla, CA 92093-0910

Tel: 858.534.5815

innovation@ucsd.edu

<https://innovation.ucsd.edu>

Fax: 858.534.7345

© 2019 - 2020, The

Regents of the University of
California

[Terms of use](#)

[Privacy Notice](#)