Cannistraci-Hebb Training of Convolutional NeuralNetworks

Hanming Li, Yusong Wang,
Yingtao Zhang, Carlo Vittorio Cannistraci

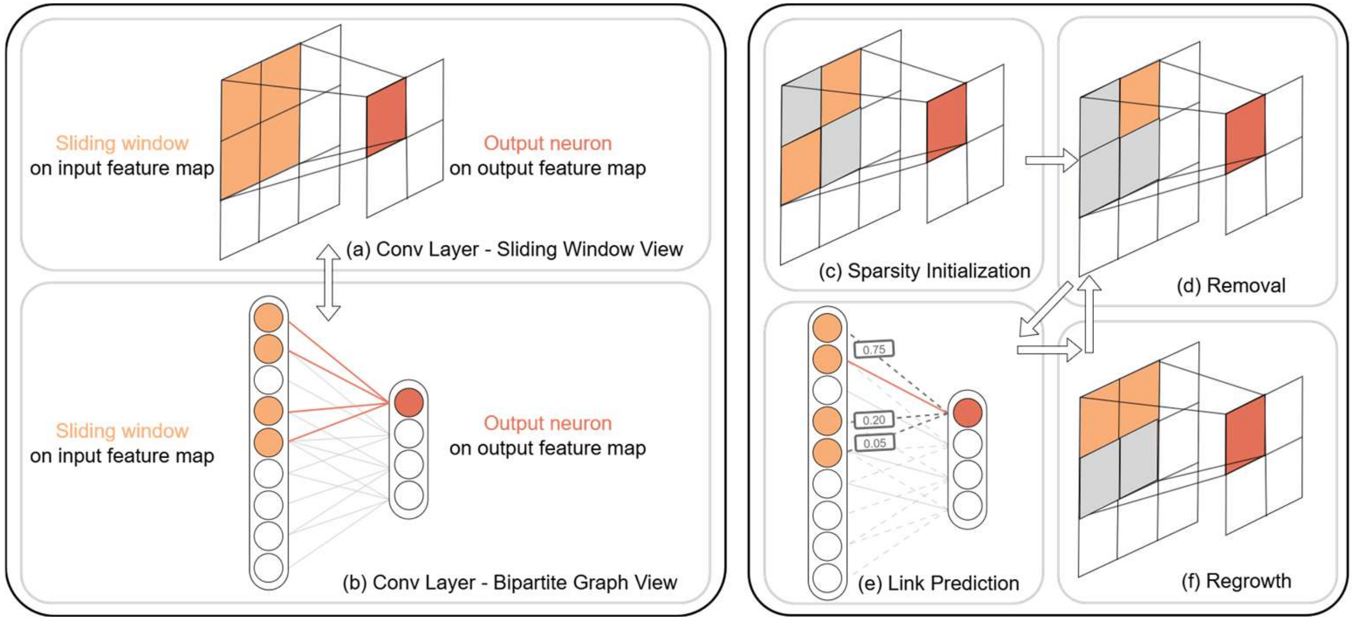
NEURAL INFORMATION PROCESSING SYSTEMS

Introduction

Dynamic sparse training (DST) enables neural networks to evolve their topology during training, which reduces computational overhead while maintaining performance. Cannistraci-Hebb Training (CHT), a brain-inspired method based on epitopological learning principles, has demonstrated significant advantages in building ultra-sparse fully connected networks.

We propose CHT-Conv, extending CHT to convolutional layers while adhering to the inherent constraints of convolutional layers. Experiments on CIFAR-10 and CIFAR-100 using VGG16 architectures show CHT-Conv achieves competitive or superior performance compared to SET baseline at 50\% and 70\% sparsity levels.

$$CH3\_L3(u,v) = \sum_{z_1,z_2 \in L3} \frac{1}{\sqrt{de^*_{z_1} * de^*_{z_2}}}$$



Proposed Method

Initialization, Fig (c)

For each convolutional layer, we initialize a random Boolean tensor mask of the same shape as each kernel to govern which positions are removed at this phase.

Removal, Fig (d)

On each kernel, a fixed fraction positions with the smallest weight magnitudes are removed.

Link Prediction and Regrowth, Fig (e)(f)

With CH link predictors, each inactive positions is assigned a likelihood scores, by which we choose the positions to regrow.

Experimental Results

We evaluate the performance of CHT-Conv on CIFAR10 and CIFAR100 with VGG16.

Preliminary experimental results indicate that:
(1) As sparsity progressively increases, the performance of the network declines;
(2) Compared to SET, the CHT method either outperforms SET (on CIFAR-10) or performs at least comparably to SET (on CIFAR-100).

Sparsity	Method	CIFAR-10	CIFAR-100
0%	Dense	92.14 ± 0.06	$\mathbf{72.58 \pm 0.11}$
50%	SET	92.08 ± 0.10	71.86 ± 0.12
	CHT-CH2	92.04 ± 0.21	72.08 ± 0.06
	CHT-CH3	$\mathbf{92.32 \pm 0.06}$	$\mathbf{72.14 \pm 0.06}$
70%	SET	91.63 ± 0.07	$\mathbf{71.03 \pm 0.13}$
	CHT-CH2	91.73 ± 0.15	70.89 ± 0.23
	CHT-CH3	$\mathbf{92.04 \pm 0.05}$	70.75 ± 0.00

References

Carlo Vittorio Cannistraci et al. From link-prediction in brain connectomes and protein interactomes to the local-community-paradigm in complex networks. Scientific reports, 3(1):1613, 2013.

Decebal Constantin Mocanu et al. Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science. Nature communications, 9(1):2383, 2018.

Yingtao Zhang et al. Epitopological learning and cannistraci-Hebb network shape intelligence brain-inspired theory for ultra-sparse advantage in deep learning. In The Twelfth International Conference on Learning Representations, 2024.

Yingtao Zhang et al. Brain network science modelling of sparse neural networks enables transformers and llms to perform as fully connected. arXiv preprintarXiv:2501.19107, 2025.