Introduction

- One-shot Network Pruning at Initialization (OPaI) is an effective method to decrease network pruning costs. Recently, there is a growing belief that data is unnecessary in OPaI.
- However, extensive experiments reveal that OPaI is data-dependent in two representative OPaI methods.
- We propose two novel methods, Discriminative One-shot Network Pruning (DOP) and Super Stitching, to prune the network by high-level visual discriminative image patches.

Related Work

- One-cut Network Pruning at Initialization (OPaI)
  - Samples a randomize mini-batch in training data.
  - Mask unimportant parameters with 0, fine-tuning pruned network.
- SNIP computes the connection sensitivity
  \[
  s(\theta_i) = \frac{\partial \mathcal{L}(T^b, \theta_i; m)}{\partial m}
  \]
  where \(\mathcal{L} = \mathcal{L}(T^b, \theta_i; m)\) is the loss function.
- GraSP uses the Hessian \(H\) to preserve the gradient flow
  \[
  s(\theta_i) = -\theta_i \frac{\partial^2 \mathcal{L}(T^b, \theta_i)}{\partial \theta_i^2}.
  \]
- OPaI data-independent?
  - A much-debated question is whether data-independent in OPaI. Recent research demonstrate in sanity check approaches, that the data using in pruning step is unnecessary.

Proposed Method

- Discriminative One-shot Network Pruning (DOP) and Super Stitching.
  - (1) Cluster segments in trained network’s activation space, extract Discriminative Image Patches. The green is meaningful in network prediction, and the red is meaningless. (2) Using Discriminative Image Patches or Super Stitching to prune unimportant parameters by a specific OPaI algorithm.

Experiments and Results

- Pruning ResNet-50 with Varying Levels of Sparsity with Discriminative Image Patches.
- Sparsity percentage: 60%, 80%, 90%, 95% (Baseline: 76.47%)
- SNIP: 74.7% 70.94% 61.06% 38.43%
- SNIP with DOP (Ours): 74.29% 71.15% 64.12% 48.14%
- GraSP: 73.87% 71.14% 67.07% 61.76%
- GraSP with DOP (Ours): 74.19% 71.76% 67.65% 60.02%

Table: DOP: Top-1 Test Accuracy of ResNet-50 on ImageNet.

- We design three ablation experiments to investigate the function of data in OPaI by gradually changing the content of input data.
- We create Super Stitching to further improve the discriminative image patches in OPaI. We aim to use fewer but more informative samples to enhance the gradient flow.

Conclusion

- Our research reveals that informative data is helpful in OPaI, and provides a new route for OPaI advancement.
- Our novel proposed methods DOP and Super Stitching can significantly improve pruning performance.
- Our work refreshes our typical views of the OPaI methods.

Acknowledgments

This work is supported by the Initiative for Realizing Diversity in the Research Environment (Advanced Type).