# Supplymentary for EagleEye: Fast Sub-net Evaluation for Efficient Neural Network Pruning

Bailin Li[1], Bowen Wu[2][0000−0001−6624−8744], Jiang Su[1], and Guangrun Wang[2]

[1]Dark Matter AI Inc. [2]Sun Yat-sen University
bl-zorro@163.com, {wubw6,wanggrun}@mail2.sysu.edu.cn, sujiang@dm-ai.cn

## 1 Quantitative Analysis of Correlation for More Datasets and Architectures

In Section 4.1, we have already shown the quantitative analysis results for MobileNetV1 [2] on ImageNet [1]. In this section, more exhaustive experiments are presented to prove the generalizability of the adaptive-BN-based evaluation method. Experiments results for MobileNetV1, MobileNetV2 and ResNet-50 on CIFAR-10 [3] and ImageNet are presented, see Figure 1, 2, 3, 4, 5 and Table 1.

**Table 1.** Correlation analysis for more datasets(CIFAR-10 and ImageNet) and architectures(MobileNetV2, ResNet-50 and MobileNetV1).

| Model | Dataset | FLOPs constraints | $\rho_{X_1,Y}$ | $\rho_{X_2,Y}$ | $\phi_{X_1,Y}$ | $\phi_{X_2,Y}$ | $\tau_{X_1,Y}$ | $\tau_{X_2,Y}$ |
|---|---|---|---|---|---|---|---|---|
| MobileNetV1 | CIFAR-10 | Not Fixed | 0.613 | 0.933 | 0.689 | 0.937 | 0.500 | 0.790 |
| | | 50% FLOPs | 0.638 | 0.854 | 0.511 | 0.829 | 0.369 | 0.656 |
| | | 62.5% FLOPs | 0.655 | 0.857 | 0.632 | 0.785 | 0.457 | 0.610 |
| ResNet-50 | CIFAR-10 | Not Fixed | 0.279 | 0.681 | 0.450 | 0.519 | 0.324 | 0.365 |
| MobileNetV2 | ImageNet | Not Fixed | -0.09 | 0.634 | -0.216 | 0.564 | -0.162 | 0.404 |

All the above tables and figures prove that the adaptive-BN-based evaluation shows stronger correlation, and hence a more robust prediction, between the evaluated and fine-tuned accuracy for the pruning candidates.

## 2 Quantitative Analysis of Correlation for Pruning Method Other than $L1$-norm-based

To demonstrate the generalizability of the adaptive-BN-based evaluation method, we perform correlation analysis for $L2$-norm-based Filter Pruning method. See Figure 6 and Table 2, the adaptive-BN-based evaluation still shows stronger correlation comparing to vanilla evaluation.

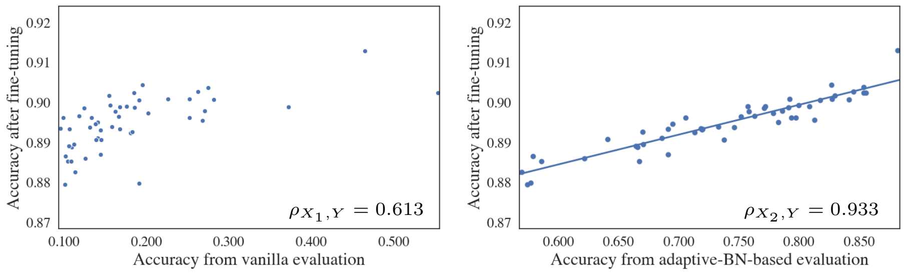


**Fig. 1.** Vanilla vs. adaptive-BN evaluation: MobileNetV1 on CIFAR-10 without fixing FLOPs constraint.

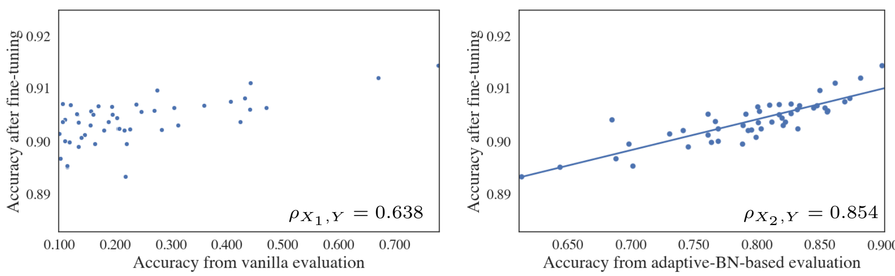


**Fig. 2.** Vanilla vs. adaptive-BN evaluation: MobileNetV1 on CIFAR-10 under the constraint of 50% FLOPs remaining.

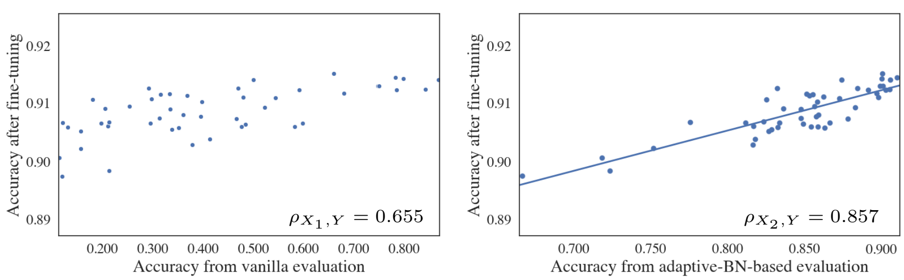


**Fig. 3.** Vanilla vs. adaptive-BN evaluation: MobileNetV1 on CIFAR-10 under the constraint of 62.5% FLOPs remaining.

**Table 2.** Correlation analysis for $L2$-norm-based Filter Pruning method.

| Model | Dataset | FLOPs constraints | $\rho_{X_1,Y}$ | $\rho_{X_2,Y}$ | $\phi_{X_1,Y}$ | $\phi_{X_2,Y}$ | $\tau_{X_1,Y}$ | $\tau_{X_2,Y}$ |
|---|---|---|---|---|---|---|---|---|
| MobileNetV1 | ImageNet | Not Fixed | 0.074 | 0.535 | 0.021 | 0.516 | 0.014 | 0.367 |

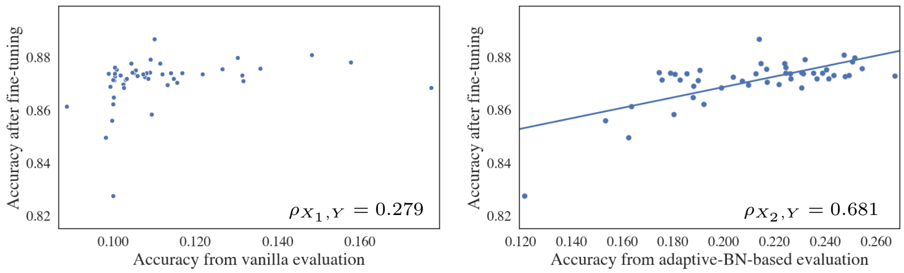


**Fig. 4.** Vanilla vs. adaptive-BN evaluation: ResNet-50 on CIFAR-10 without fixing FLOPs constraint.

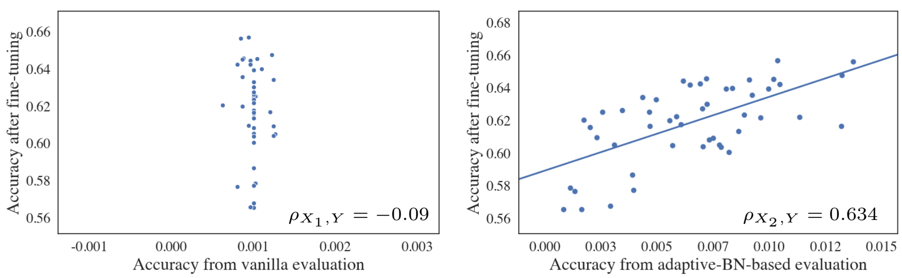


**Fig. 5.** Vanilla vs. adaptive-BN evaluation: MobileNetV2 on ImageNet without fixing FLOPs constraint.

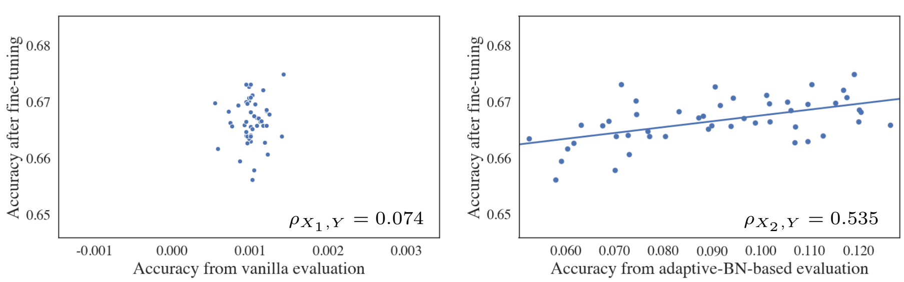


**Fig. 6.** Vanilla vs. adaptive-BN evaluation: MobileNetV1 on ImageNet without fixing FLOPs constraint, pruning by ranking filters by their $L2$-norm.

# 3 Implementation Details for Pruning MobileNetV1 on CIFAR-10

As mentioned in Section 4.4, we compare the our EagleEye method with Filter Pruning and the directly-scaled models. In this section, we show the detail hyperparameters for the experiments.

For the directly-scaled models(0.75, 0.5 and 0.25 $\times$ MobileNetV1), we train for 350 epochs with the base learning rate of 1e-1 and decrease the it by 10 at 150, 250 epochs. The batch size is 1024 and the weight-decay is set to 5e-4.

For Filter Pruning [4] method, we uniformly prune each layer with the same pruning rate according to their $L1$-norm magnitude. After pruning out filters, both Filter Pruning and our EagleEye method fine-tune the sub-nets for 30 epochs with the learning rate of 1e-3.

# References

1. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 248–255. Ieee (2009)
2. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861 (2017)
3. Krizhevsky, A.: Learning multiple layers of features from tiny images (2009)
4. Li, H., Kadav, A., Durdanovic, I., Samet, H., Graf, H.P.: Pruning filters for efficient convnets. arXiv preprint arXiv:1608.08710 (2016)