Lightweight Alpha Matting Network Using Distillation-Based Channel Pruning -SUPPLEMENTARY MATERIAL-

Donggeun Yoon[1][0000-0001-6153-8056], Jinsun Park[2][0000-0002-2296-819X], and Donghyeon Cho*[1][0000-0002-2184-921X]

[1] Chungnam National University, Daejeon, South Korea
202250187@o.cnu.ac.kr, cdh12242@cnu.ac.kr
[2] Pusan National University, Pusan, South Korea
jspark@pnu.ac.kr

1 Introduction

The following supplementary materials provide additional experiments and details of the proposed distillation-based channel pruning method. First, we show additional experimental results of our methods in other pruning methods and the various distillation balancing factors. Second, we plot histograms of the remaining channels after channel pruning. Third, we show additional alpha matting results on the Adobe-1k [6], Distinctions-646 [4], and alphamatting.com [5] testsets. Forth, we present implementation details for semantic segmentation. Lastly, we provide additional semantic segmentation results on the PASCAL VOC2012 validation set.

2 Comparisons with Recent Channel Pruning.

We analyze pruned models by recent channel pruning method: context-aware channel pruning (CAP) [1]. CAP has been proposed for the semantic segmentation task and shows good performance. As shown in Table 1, the performance of CAP is inferior to our method and the NS method when the pruned model is trained using knowledge distillation rather than from scratch in the training stage. Note that the pruned model by CAP degrades its performance when knowledge distillation is applied during the training stage.

3 Knowldege Distillation Parameter Tuning.

To analyze the sensitivity about distillation loss, we experiment by changing the balance factors. The experiment is conducted with SPKD, and the balance factors are set to 0.5, 1, and 2 in training and pruning stages. As reported in Table 2, performance is good when (λ_4, w_3)=0.5, but the number of parameters is more than the case of (λ_4, w_3)=1. It means that the performance changes according to the tuning hyperparameter, and there is a trade-off according to

Table 1. Performance analysis by applying knowledge distillation to the model pruned by context-aware channel pruning (CAP). All evaluations are conducted by GCA-50% model on the Adobe-1k.

Methods		MSE	SAD	Grad	Conn
KD	Prune				
-	UNI	0.017	52.61	35.27	46.24
-	NS	0.017	51.57	28.70	45.66
-	CAP	0.014	48.55	26.01	42.76
NST	CAP	0.019	54.02	30.67	47.42
OFD	CAP	0.016	50.67	27.76	44.35
SPKD	CAP	0.015	47.69	25.87	41.66
SPKD	NS	0.012	42.69	21.88	37.54
SPKD	Ours	0.011	41.26	21.42	35.87

Table 2. Results according to knowledge distillation balancing factor.

λ_4	w_3	MSE	SAD	Grad	Conn	#Param
0.5	0.5	0.010	39.96	20.00	35.07	4.93M
1	1	0.011	41.26	21.42	35.87	4.66M
2	2	0.012	43.15	22.42	37.92	4.78M

the performance and the number of parameters. However, when (λ_4, w_3)=2, the performance is lower than when (λ_4, w_3)=1, and the number of parameters is large.

4 Channel Histogram of Pruned Model

To analyze which channels are pruned, we visualized the channel histogram of our pruned model. Figure 1 and Figure 2 show the channel histograms of pruned GCA model using our method on Adobe-1k and Distinctions-646 datasets. In addition, Figure 3 and Figure 4 show the channel histograms of the pruned DIM and IndexNet models on Adobe-1k dataset. Each histogram shows the number of channels in the encoder and decoder layers sequentially. As shown in those histograms, it can be seen that the number of channels in the high-level layers of encoder is significantly reduced, compared to the unpruned teacher model.

5 Additional Qualitative Matting Results

Furthermore, as shown in Figure 6 and Figure 7, more qualitative alpha matting results on Adobe-1k and Distinctions-646 are provided. As competitors, we choose NS [3], CAP [1] channel pruning methods. We also show qualitative results on the alphamatting.com benchmark in Figure 5. It is easy to see that our channel pruning method is better by zooming in on the result or looking at the inserted snapshot.

6 Implementation Details for Semantic Segmentation

In semantic segmentation experiments, we train PSPNet-50 network for 50 epochs with batch size of 8 and the learning rate 0.01 to prune network. We employ Adam optimizer [2] with momentum coefficient 0.9 and weight decay coefficient 0.0001. We randomly crop the input images into the size 473 $\times$ 473. After pruning, the pruned network is trained the same setting from pruning step. Distillation loss is used only the encoder part except first layer like image matting.

7 Additional Qualitative Segmentation Results

Figure 8 shows additional quantitative semantic segmentation results on PASCAL VOC2012 validation set. Again, we compare our best performed method to NS and CAP. As shown from the comparison, our proposed method is more effective than the other methods.

Upon acceptance of this paper, we will release the source codes and pretrained weights for all models used in the experiments.

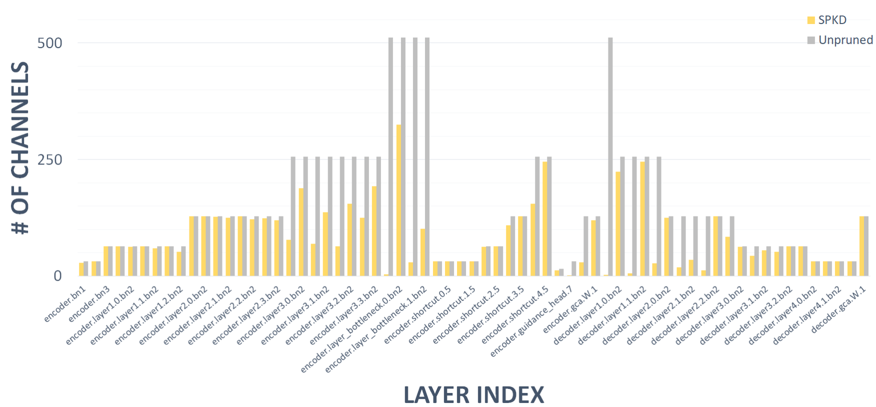

Fig. 1. Channel histogram of our pruned GCA model (trained by Adobe-1k dataset).

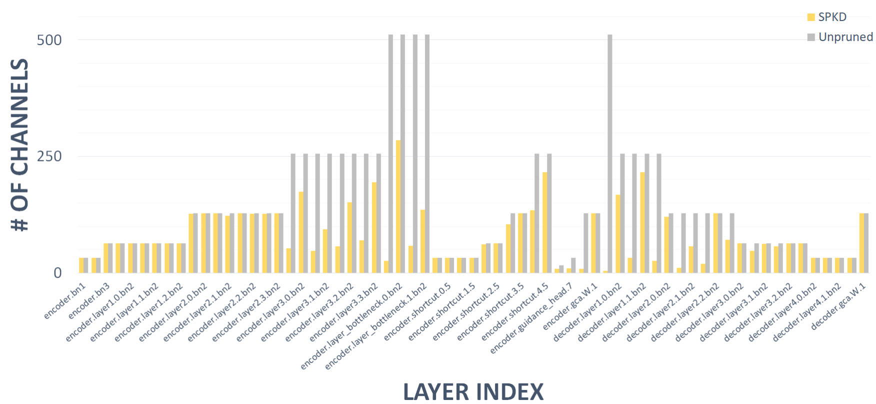

Fig. 2. Channel histogram of our pruned GCA model (trained by Distinctions-646 dataset).

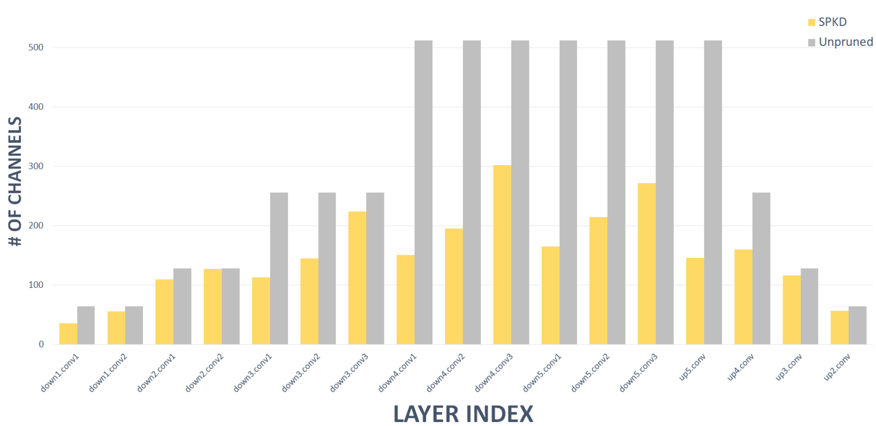

Fig. 3. Channel histogram of our pruned DIM model (trained by Adobe-1k dataset).

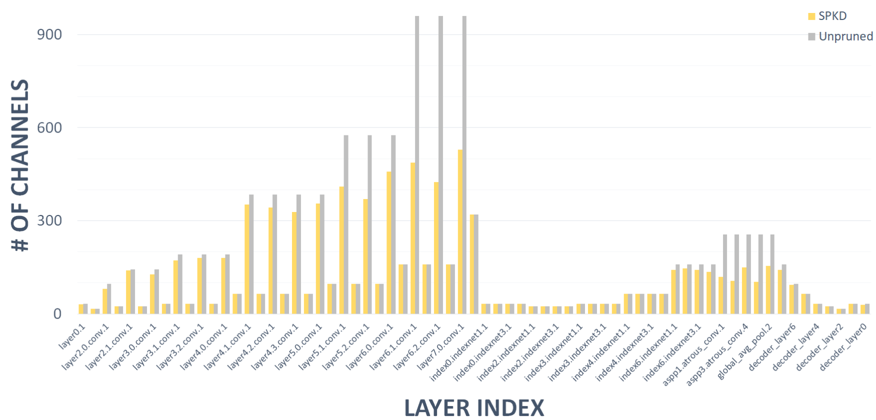

Fig. 4. Channel histogram of our pruned IndexNet model (trained by Adobe-1k dataset).

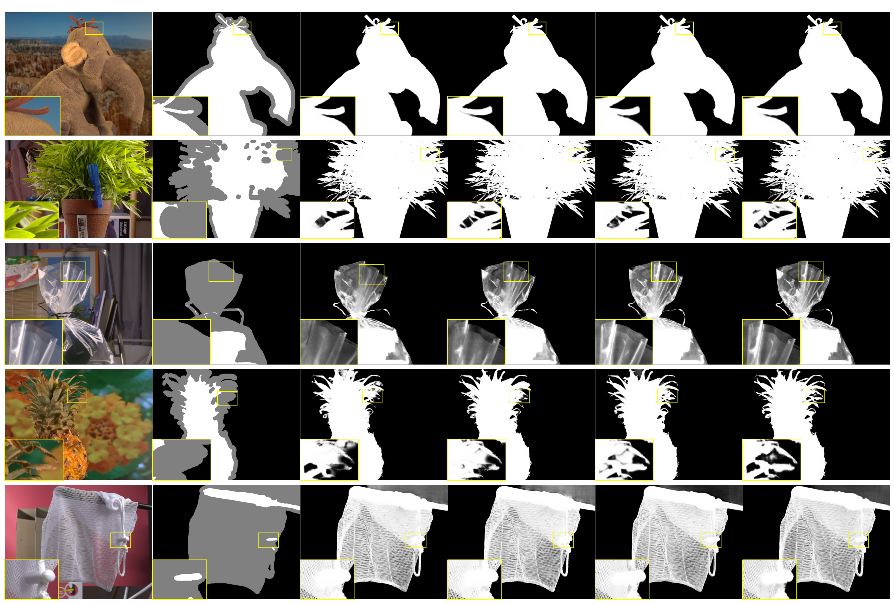

Fig. 5. Qualitative image matting results by GCA on alphamatting.com testset. (a) Input images. (b) trimaps. (c) NS [3]. (d) CAP [1]. (e) NS-SPKD. (f) Ours-SPKD.

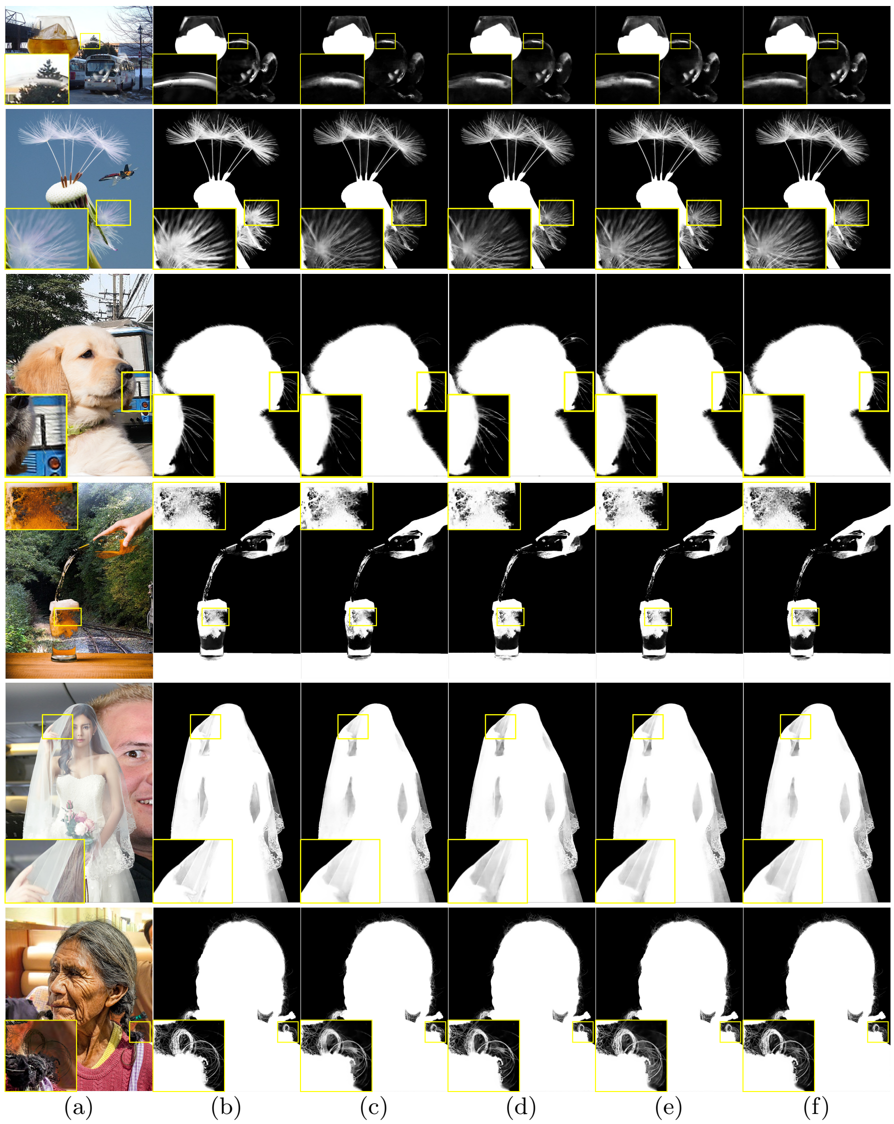

Fig. 6. Qualitative image matting results by GCA on the Adobe-1k. (a) Input images. (b) Ground truths. (c) NS [3]. (d) CAP [1]. (e) NS-SPKD. (f) Ours-SPKD.

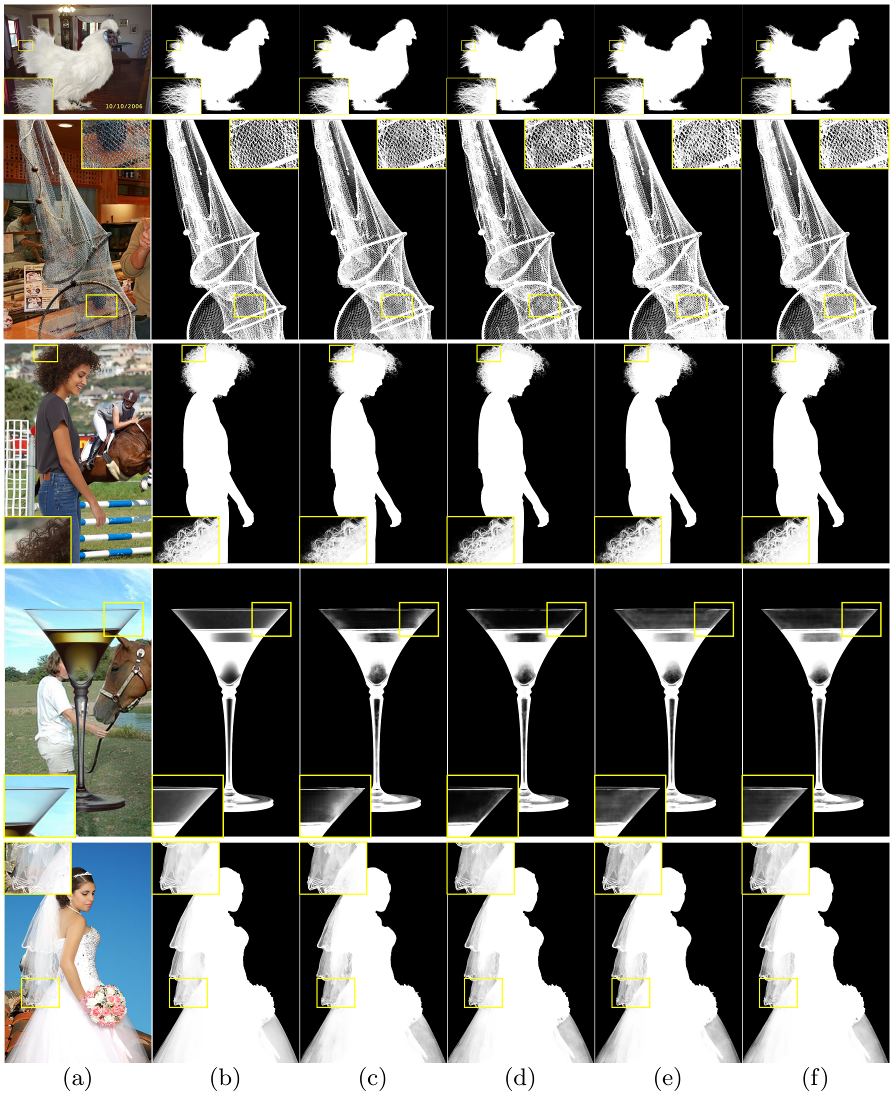

Fig. 7. Qualitative image matting results by GCA on the Distinctions-646. (a) Input images. (b) Ground truths. (c) NS [3]. (d) CAP [1]. (e) NS-SPKD. (f) Ours-SPKD.

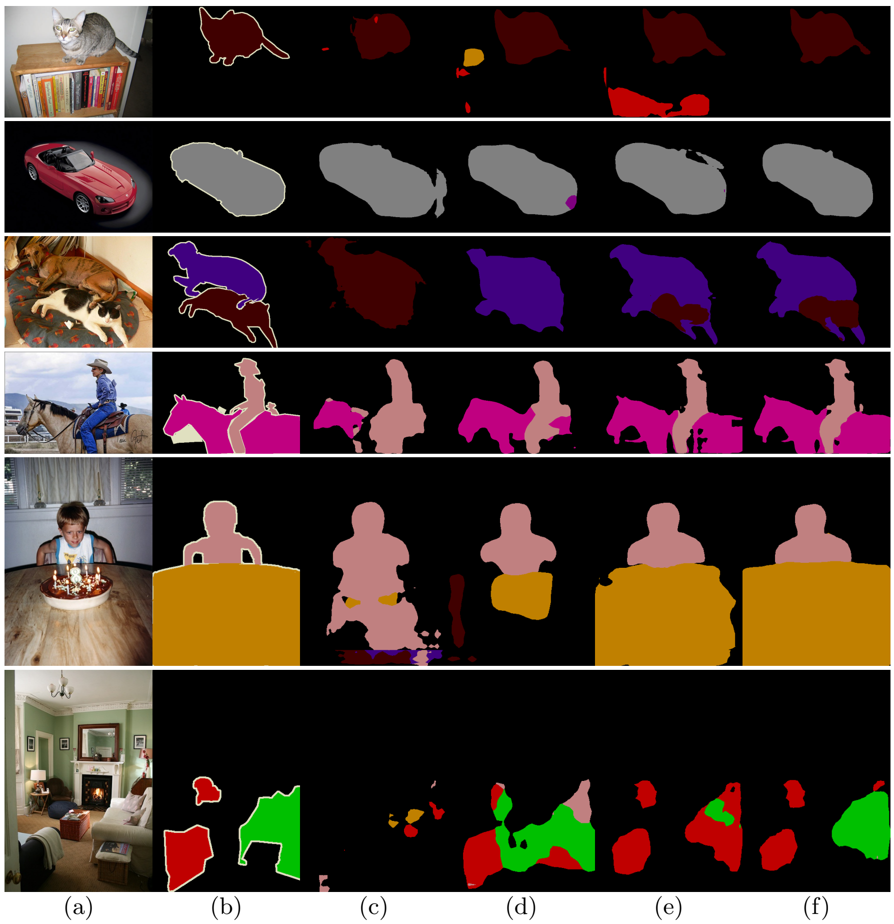

Fig. 8. Qualitative semantic segmentation results by PSPNet-50 on PASCAL VOC2012 validation set. (a) Input images. (b) Ground truths. (c) NS [3]. (d) CAP [1]. (e) NS-OFD. (f) Ours-OFD.

References

1. He, W., Wu, M., Liang, M., Lam, S.K.: Cap: Context-aware pruning for semantic segmentation. In: Proc. of Winter Conference on Applications of Computer Vision (WACV). pp. 960–969 (2021)
2. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. International Conference on Learning Representation (ICLR) (2014)
3. Liu, Z., Li, J., Shen, Z., Huang, G., Yan, S., Zhang, C.: Learning efficient convolutional networks through network slimming. In: Proc. of Int'l Conf. on Computer Vision (ICCV). pp. 2736–2744 (2017)
4. Qiao, Y., Liu, Y., Yang, X., Zhou, D., Xu, M., Zhang, Q., Wei, X.: Attention-guided hierarchical structure aggregation for image matting. In: Proc. of Computer Vision and Pattern Recognition (CVPR) (2020)
5. Rhemann, C., Rother, C., Wang, J., Gelautz, M., Kohli, P., Rott, P.: A perceptually motivated online benchmark for image matting. In: Proc. of Computer Vision and Pattern Recognition (CVPR) (2009)
6. Xu, N., Price, B.L., Cohen, S., Huang, T.S.: Deep image matting. In: Proc. of Computer Vision and Pattern Recognition (CVPR) (2017)