

Supplementary Material for 'Learning to ignore: rethinking attention in CNNs'

Firas Laakom^{*1}

firas.laakom@tuni.fi

Kateryna Chumachenko^{*1}

kateryna.chumachenko@tuni.fi

Jenni Raitoharju²

jenni.raitoharju@syke.fi

Alexandros Iosifidis³

ai@ece.au.dk

Moncef Gabbouj¹

moncef.gabbouj@tuni.fi

¹ Department of Computing Sciences,
Tampere University, Tampere, Finland

² Finnish Environment Institute,
Jyväskylä, Finland

³ Department of Electrical and Computer
Engineering, Aarhus University, Aarhus,
Denmark

1 Experimental results

Here, we report the results of additional experiments of different architectures with CBAM baseline with identical settings, i.e., initialization from identical weights, identical random augmentation, and identical data loading order to the model. We initialize from three different weight sets. The average and the standard deviation of the test errors on CIFAR10 and CIFAR100 are reported in Table 2. Additionally, we report ImageNet results initialized from identical weights on ResNet18 architecture in Table 1.

	Top-1 Accuracy%	Top-5 Accuracy%
Standard	69.29 ± 0.01	88.71 ± 0.05
CBAM	69.72 ± 0.07	89.06 ± 0.05
CBAM-Ign ₁ ($\alpha=1$)	69.79 ± 0.04	88.98 ± 0.04
CBAM-Ign ₁ ($\alpha=0.5$)	69.90 ± 0.08	89.13 ± 0.04
CBAM-Ign ₁ ($\alpha=0.8$)	69.71 ± 0.17	89.03 ± 0.06
CBAM-Ign ₂	69.81 ± 0.04	89.10 ± 0.01
CBAM-Ign ₃	69.93 ± 0.05	89.04 ± 0.09

Table 1: Accuracy of ResNet18 with different CBAM variants on ImageNet dataset

2 Validation loss curves

Here, we present additional validation loss curves of different ignoring approaches and corresponding baselines with ResNet50 on both CIFAR10 and CIFAR100 datasets.

		CIFAR 10	CIFAR 100	
		Top-1 Error%	Top-1 Error%	Top-5 Error%
ResNet50	Standard	8.22 ± 0.65	33.30 ± 0.83	10.78 ± 0.33
	CBAM	7.47 ± 0.30	31.20 ± 0.18	9.41 ± 0.25
	CBAM-Ign ₁ ($\alpha=1$)	7.63 ± 0.05	31.02 ± 0.33	9.52 ± 0.12
	CBAM-Ign ₁ ($\alpha=0.5$)	7.31 ± 0.29	31.11 ± 0.31	9.03 ± 0.28
	CBAM-Ign ₁ ($\alpha=0.8$)	7.30 ± 0.11	30.44 ± 0.04	9.23 ± 0.14
	CBAM-Ign ₂	7.44 ± 0.05	31.86 ± 0.22	9.64 ± 0.26
	CBAM-Ign ₃	7.47 ± 0.17	30.92 ± 0.48	9.41 ± 0.42
DenseNet	Standard	7.28 ± 0.20	29.60 ± 0.20	8.29 ± 0.16
	CBAM	6.64 ± 0.09	27.94 ± 0.25	7.74 ± 0.21
	CBAM-Ign ₁ ($\alpha=1$)	6.54 ± 0.16	27.60 ± 0.25	7.81 ± 0.19
	CBAM-Ign ₁ ($\alpha=0.5$)	6.58 ± 0.17	27.36 ± 0.33	7.36 ± 0.15
	CBAM-Ign ₁ ($\alpha=0.8$)	6.51 ± 0.14	27.34 ± 0.51	7.45 ± 0.11
	CBAM-Ign ₂	6.69 ± 0.05	28.09 ± 0.12	7.67 ± 0.22
	CBAM-Ign ₃	6.67 ± 0.04	27.75 ± 0.11	7.69 ± 0.06
Resnet18	Standard	9.05 ± 0.04	34.98 ± 0.37	10.27 ± 0.21
	CBAM	8.58 ± 0.35	33.62 ± 0.17	9.97 ± 0.17
	CBAM-Ign ₁ ($\alpha=1$)	8.52 ± 0.22	33.65 ± 0.09	9.72 ± 0.29
	CBAM-Ign ₁ ($\alpha=0.5$)	8.45 ± 0.13	33.10 ± 0.32	9.83 ± 0.12
	CBAM-Ign ₁ ($\alpha=0.8$)	8.68 ± 0.27	33.14 ± 0.05	9.75 ± 0.11
	CBAM-Ign ₂	8.33 ± 0.17	33.37 ± 0.39	9.76 ± 0.22
	CBAM-Ign ₃	8.74 ± 0.19	33.49 ± 0.39	9.88 ± 0.25

Table 2: Average and standard deviation of error rates for different CBAM variants on CIFAR10 and CIFAR100 datasets.

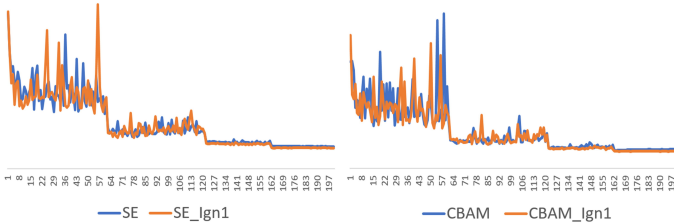


Figure 1: Validation loss curves of ResNet50 on CIFAR10 using the Ign1 approach.

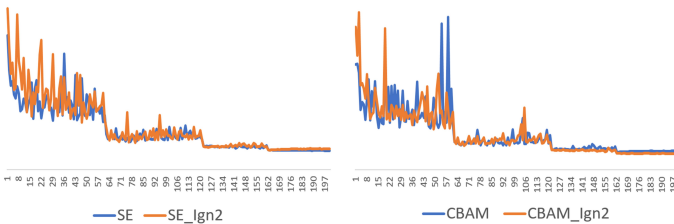


Figure 2: Validation loss curves of ResNet50 on CIFAR10 using the Ign2 approach.

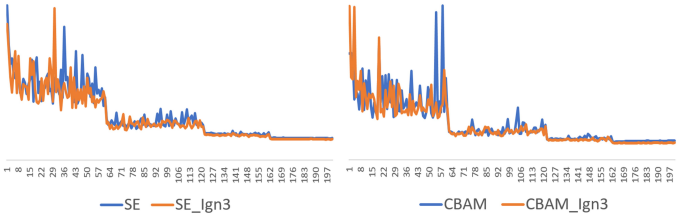


Figure 3: Validation loss curves of ResNet50 on CIFAR10 using the Ign3 approach.

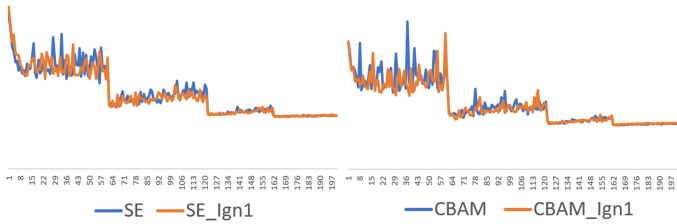


Figure 4: Validation loss curves of ResNet50 on CIFAR100 using the Ign1 approach.

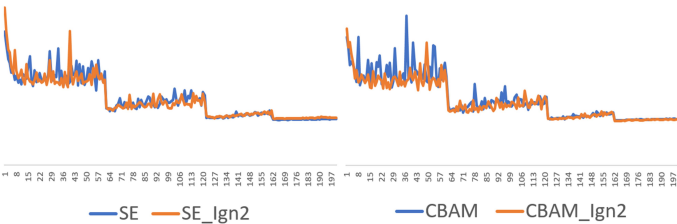


Figure 5: Validation loss curves of ResNet50 on CIFAR100 using the Ign2 approach.

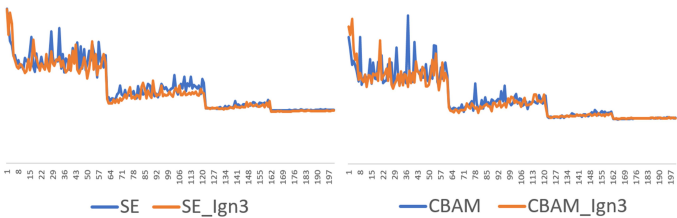


Figure 6: Validation loss curves of ResNet50 on CIFAR100 using the Ign3 approach.

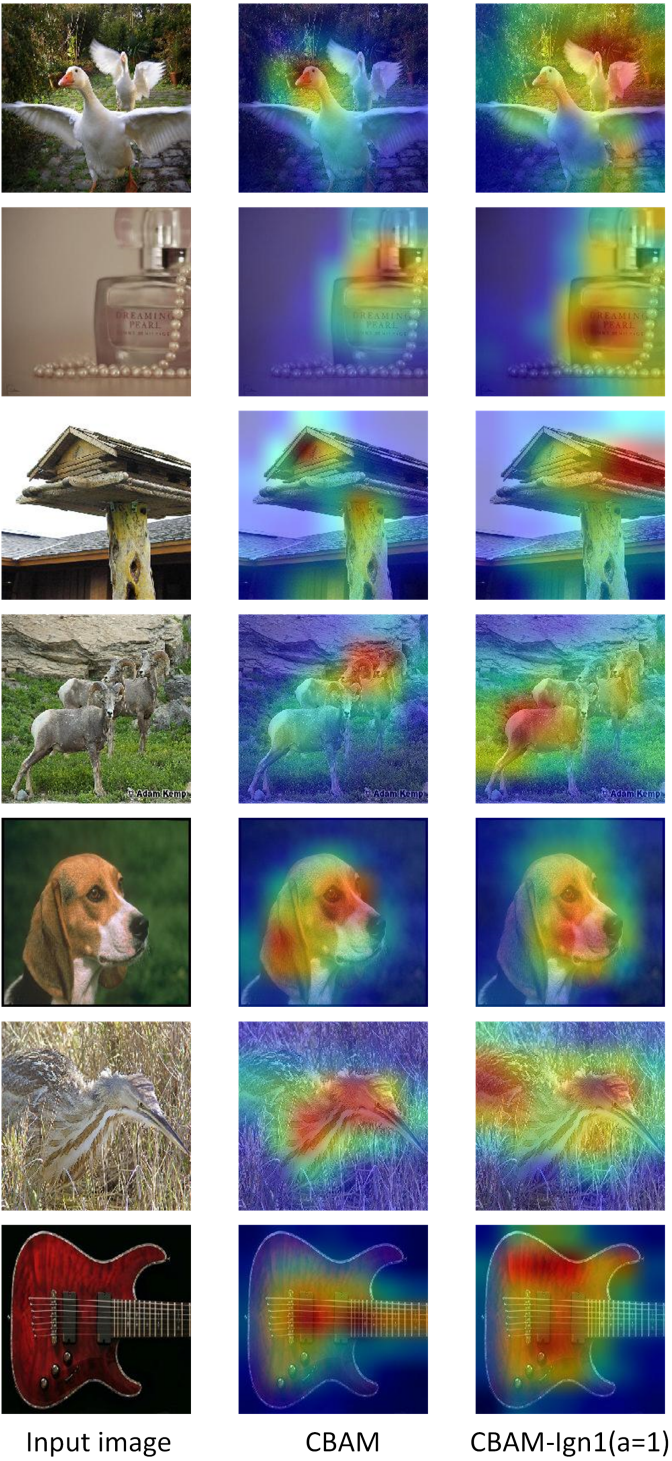


Figure 7: Attention masks for CBAM, CBAM-Ign1($\alpha = 1$)