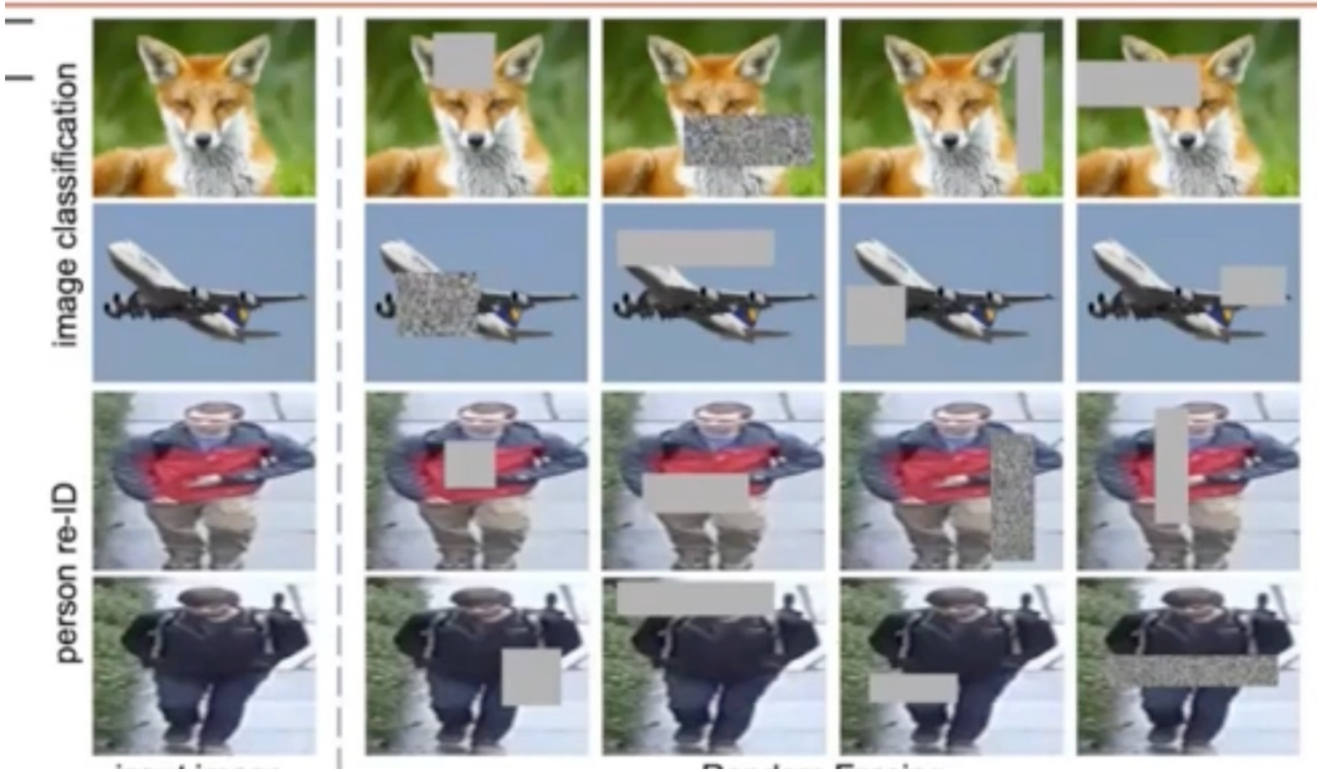
Random Erasing Data Augmentation

Zhun Zhong, Liang Zheng, Guoliang Kang, Shaozi Li, Yi Yang

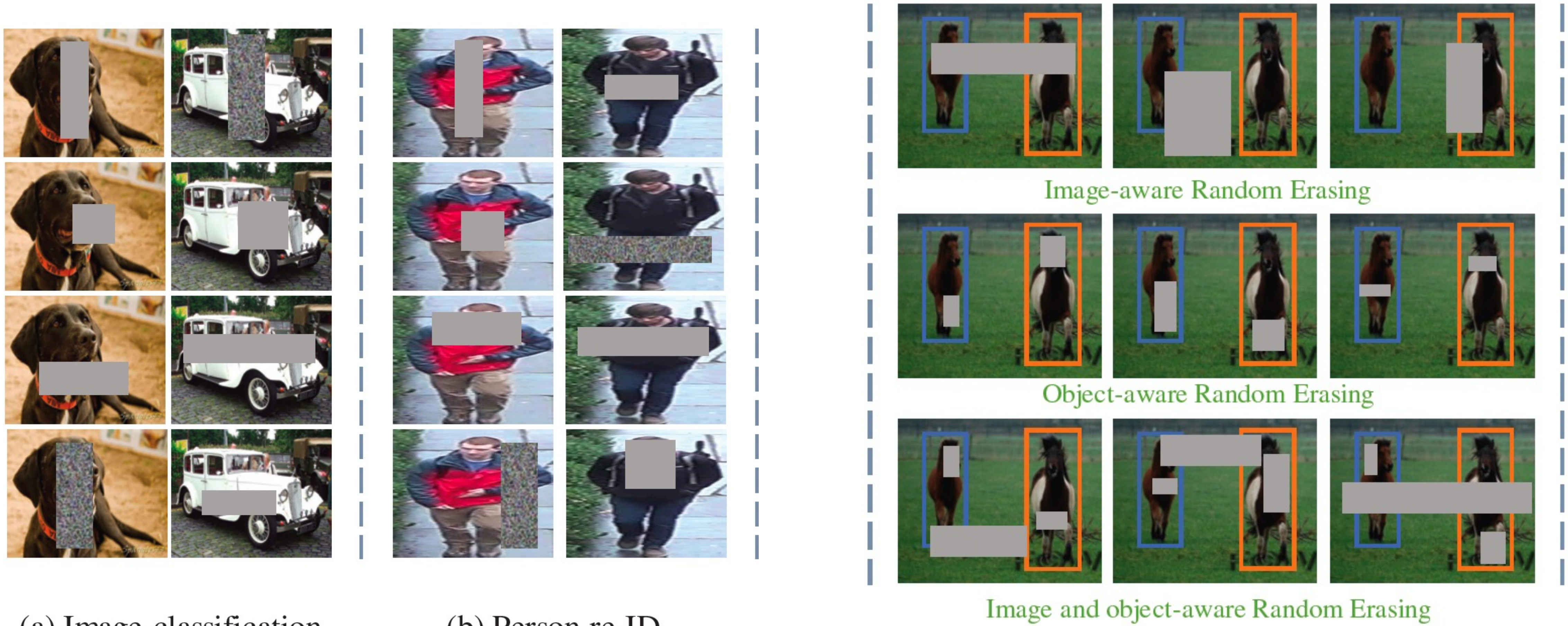
Published in 2020 in AAAI
3919 citations as of this week



Presenter: Abdullah Mamun

Date: August 7, 2024

Applications



(a) Image classification

(b) Person re-ID

(c) Object detection

Algorithm

Algorithm 1: Random Erasing Procedure

Input : Input image I; Image size W and H; Area of image S; Erasing probability p; Erasing area ratio range s_l and s_h; Erasing aspect ratio range r_1 and r_2.

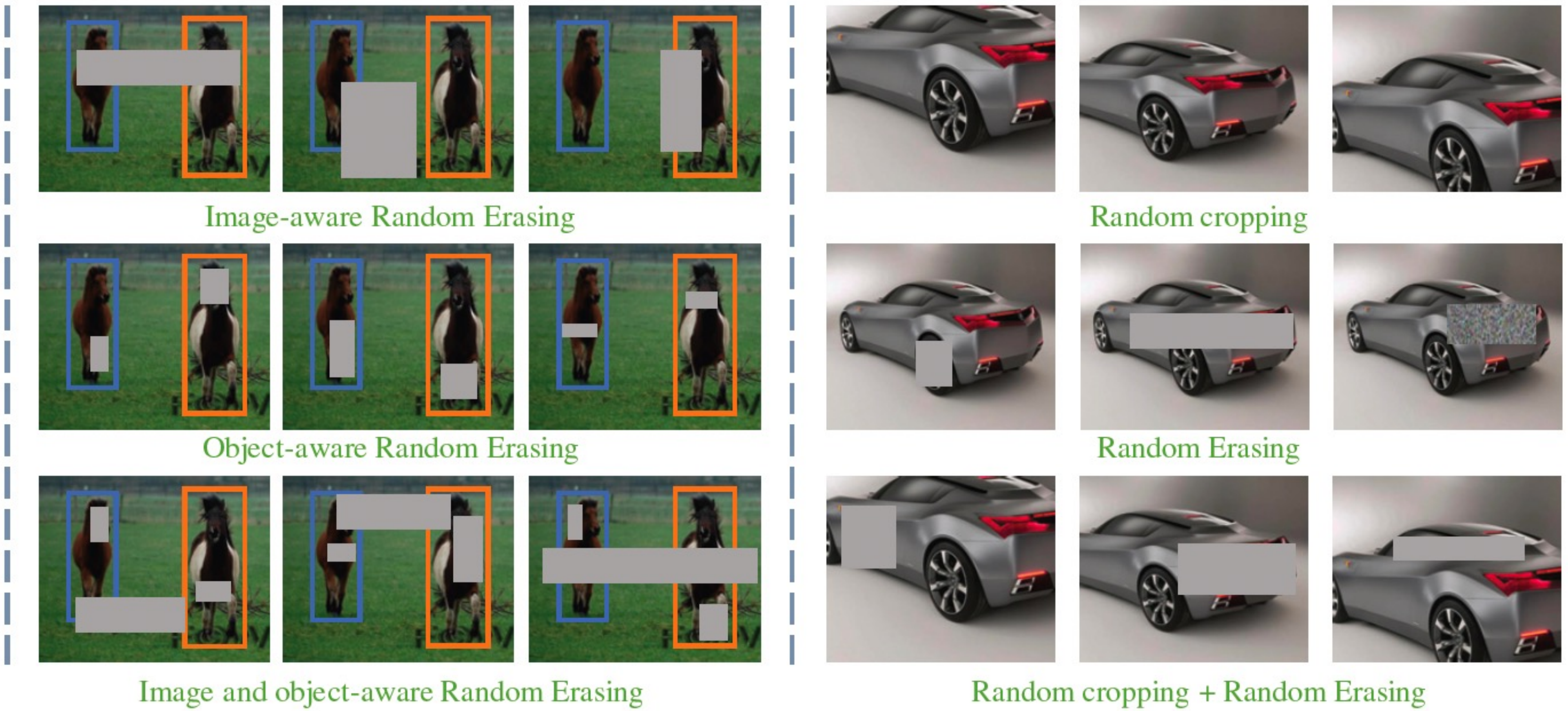
Output: Erased image I^*.

Initialization: $p_1 \leftarrow$ Rand (0, 1).

if $p_1 \geq p$ **then**
    $I^* \leftarrow I$;
    return I^*.
else
    while *True* **do**
        $S_e \leftarrow \text{Rand}\ (s_l, s_h) \times S$;
        $r_e \leftarrow \text{Rand}\ (r_1, r_2)$;
        $H_e \leftarrow \sqrt{S_e \times r_e},\ W_e \leftarrow \sqrt{\frac{S_e}{r_e}}$;
        $x_e \leftarrow \text{Rand}\ (0, W),\ y_e \leftarrow \text{Rand}\ (0, H)$;
        if $x_e + W_e \leq W$ *and* $y_e + H_e \leq H$ **then**
            $I_e \leftarrow (x_e, y_e, x_e + W_e, y_e + H_e)$;
            $I(I_e) \leftarrow \text{Rand}\ (0, 255)$;
            $I^* \leftarrow I$;
            return I^*.
        end
    end
end

(a) Image classification (b) Person re-ID

Applications and Combining with Others



(c) Object detection

(d) Different augmentation methods

Datasets

3 Datasets

For **image classification**, we evaluate on four image classification datasets, including two well-known datasets, CIFAR-10 and CIFAR-100 (Krizhevsky and Hinton 2009), a new dataset Fashion-MNIST (Xiao, Rasul, and Vollgraf 2017), and a large-scale dataset ImageNet2012 (Deng et al. 2009). **CIFAR-10** and **CIFAR-100** contain 50,000 training and 10,000 testing 32×32 color images drawn from 10 and 100 classes, respectively. **Fashion-MNIST** consists of 60,000 training and 10,000 testing 28x28 gray-scale images. Each image is associated with a label from 10 classes. **ImageNet2012** consists of 1,000 classes, including 1.28 million training images and 50k validation images. For CIFAR-10, CIFAR-100 and Fashion-MNIST, we evaluate top-1 error rates in the format "mean $\pm$ std" based on 5 runs. For ImageNet2012, we evaluate the top-1 and top-5 error rates on the validation set.

For **object detection**, we use the **PASCAL VOC 2007** (Everingham et al. 2010) dataset which contains 9,963 images of 24,640 annotated objects in training/validation and testing sets. We use the "trainval" set for training and "test" set for testing. We evaluate the performance using mean average precision (mAP).

For **person re-identification (re-ID)**, the **Market-1501** dataset (Zheng et al. 2015) contains 12,936 images with 751 identities for training, 19,732 images with 750 identities and 3,368 query images for testing. **DukeMTMC-reID** (Zheng, Zheng, and Yang 2017; Ristani et al. 2016) includes 16,522 training images of 702 identities, 2,228 query images of the other 702 identities and 17,661 gallery images. For **CUHK03** (Li et al. 2014), we use the **new training/testing protocol** proposed in (Zhong et al. 2017). There are 767 identities in the training set and 700 identities in the testing set. We conduct experiment on both "detected" and "labeled" sets. Rank-1 accuracy and mean average precision (mAP) are evaluated on these three datasets.

Variations in object detection

4.3 Random Erasing for Object Detection

Object detection aims at detecting instances of semantic objects of a certain class in images. Since the location of each object in the training image is known, we implement Random Erasing with three schemes: 1) Image-aware Random Erasing (**IRE**): selecting erasing region on the whole image, the same as image classification and person re-identification; 2) Object-aware Random Erasing (**ORE**): selecting erasing regions in the bounding box of each object. In the latter, if there are multiple objects in the image, Random Erasing is applied on each object separately. 3) Image and object-aware Random Erasing (**I+ORE**): selecting erasing regions in both the whole image and each object bounding box. Examples of Random Erasing for object detection with the three schemes are shown in Fig. 1(c).

(c) Object detection

Results – Classification

Table 1: Test errors (%) with different architectures on CIFAR-10, CIFAR-100 and Fashion-MNIST. **RE:** Random Erasing.

Model	CIFAR-10		CIFAR-100		Fashion-MNIST	
	Baseline	RE	Baseline	RE	Baseline	RE
ResNet-20	7.21 ± 0.17	6.73 ± 0.09	30.84 ± 0.19	29.97 ± 0.11	4.39 ± 0.08	4.02 ± 0.07
ResNet-32	6.41 ± 0.06	5.66 ± 0.10	28.50 ± 0.37	27.18 ± 0.32	4.16 ± 0.13	3.80 ± 0.05
ResNet-44	5.53 ± 0.08	5.13 ± 0.09	25.27 ± 0.21	24.29 ± 0.16	4.41 ± 0.09	4.01 ± 0.14
ResNet-56	5.31 ± 0.07	4.89 ± 0.07	24.82 ± 0.27	23.69 ± 0.33	4.39 ± 0.10	4.13 ± 0.42
ResNet-110	5.10 ± 0.07	4.61 ± 0.06	23.73 ± 0.37	22.10 ± 0.41	4.40 ± 0.10	4.01 ± 0.13
ResNet-18-PreAct	5.17 ± 0.18	4.31 ± 0.07	24.50 ± 0.29	24.03 ± 0.19	4.31 ± 0.06	3.90 ± 0.06
WRN-28-10	3.80 ± 0.07	$\mathbf{3.08 \pm 0.05}$	$\mathbf{18.49 \pm 0.11}$	$\mathbf{17.73 \pm 0.15}$	$\mathbf{4.01 \pm 0.10}$	$\mathbf{3.65 \pm 0.03}$
ResNeXt-8-64	$\mathbf{3.54 \pm 0.04}$	3.24 ± 0.03	19.27 ± 0.30	18.84 ± 0.18	4.02 ± 0.05	3.79 ± 0.06

Results – Object detection

Table 6: **VOC 2007 test** detection average precision (%). ⋆ refers to training schedule in (Wang, Shrivastava, and Gupta 2017).

Method	train set	mAP	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	persn	plant	sheep	sofa	train	tv
FRCN	07	66.9	74.5	78.3	69.2	53.2	36.6	77.3	78.2	82.0	40.7	72.7	67.9	79.6	79.2	73.0	69.0	30.1	65.4	70.2	75.8	65.8
FRCN⋆	07	69.1	75.4	80.8	67.3	59.9	37.6	81.9	80.0	84.5	50.0	77.1	68.2	81.0	82.5	74.3	69.9	28.4	71.1	70.2	75.8	66.6
A-Fast-RCNN	07	71.0	74.4	81.3	67.6	57.0	46.6	81.0	79.3	86.0	52.9	75.9	73.7	82.6	83.2	77.7	72.7	37.4	66.3	71.2	78.2	74.3
Ours (IRE)	07	70.5	75.9	78.9	69.0	57.7	46.4	81.7	79.5	82.9	49.3	76.9	67.9	81.5	83.3	76.7	73.2	40.7	72.8	66.9	75.4	74.2
Ours (ORE)	07	71.0	75.1	79.8	69.7	60.8	46.0	80.4	79.0	83.8	51.6	76.2	67.8	81.2	83.7	76.8	73.8	43.1	70.8	67.4	78.3	75.6
Ours (I+ORE)	07	**71.5**	76.1	81.6	69.5	60.1	45.6	82.2	79.2	84.5	52.5	78.7	71.6	80.4	83.3	76.7	73.9	39.4	68.9	69.8	79.2	77.4
FRCN	07+12	70.0	77.0	78.1	69.3	59.4	38.3	81.6	78.6	86.7	42.8	78.8	68.9	84.7	82.0	76.6	69.9	31.8	70.1	74.8	80.4	70.4
FRCN⋆	07+12	74.8	78.5	81.0	74.7	67.9	53.4	85.6	84.4	86.2	57.4	80.1	72.2	85.2	84.2	77.6	76.1	45.3	75.7	72.3	81.8	77.3
Ours (IRE)	07+12	75.6	79.0	84.1	76.3	66.9	52.7	84.5	84.4	88.7	58.0	82.9	71.1	84.8	84.4	78.6	76.7	45.5	77.1	76.3	82.5	76.8
Ours (ORE)	07+12	75.8	79.4	81.6	75.6	66.5	52.7	85.5	84.7	88.3	58.7	82.9	72.8	85.0	84.3	79.3	76.3	46.3	76.3	74.9	86.0	78.2
Ours (I+ORE)	07+12	**76.2**	79.6	82.5	75.7	70.5	55.1	85.2	84.4	88.4	58.6	82.6	73.9	84.2	84.7	78.8	76.3	46.7	77.9	75.9	83.3	79.3

Results – Re-identification

Table 7: Person re-identification performance with Random Erasing (RE) on Market-1501, DukeMTMC-reID, and CUHK03 based on different models. We evaluate CUHK03 under the new evaluation protocol in (Zhong et al. 2017).

Method	Model	RE	Market		Duke		CUHK03 (L)		CUHK03 (D)	
			Rank-1	mAP	Rank-1	mAP	Rank-1	mAP	Rank-1	mAP
IDE	ResNet-18	No	79.87	57.37	67.73	46.87	28.36	25.65	26.86	25.04
		Yes	82.36	62.06	70.60	51.41	36.07	32.58	34.21	31.20
	ResNet-34	No	82.93	62.34	71.63	49.71	31.57	28.66	30.14	27.55
		Yes	84.80	65.68	73.56	54.46	40.29	35.50	36.36	33.46
	ResNet-50	No	83.14	63.56	71.99	51.29	30.29	27.37	28.36	26.74
		Yes	85.24	68.28	74.24	56.17	41.46	36.77	38.50	34.75
TriNet	ResNet-18	No	77.32	58.43	67.50	46.27	43.00	39.16	40.50	37.36
		Yes	79.84	61.68	71.81	51.84	48.29	43.80	46.57	43.20
	ResNet-34	No	80.73	62.65	72.04	51.56	46.00	43.79	45.07	42.58
		Yes	83.11	65.98	72.89	55.38	53.07	48.80	53.21	48.03
	ResNet-50	No	82.60	65.79	72.44	53.50	49.86	46.74	50.50	46.47
		Yes	83.94	68.67	72.98	56.60	**58.14**	**53.83**	**55.50**	**50.74**
SVDNet	ResNet-50	No	84.41	65.60	76.82	57.70	42.21	38.73	41.85	38.24
		Yes	**87.08**	**71.31**	**79.31**	**62.44**	49.43	45.07	48.71	43.50

Results – Performance vs level of occlusion

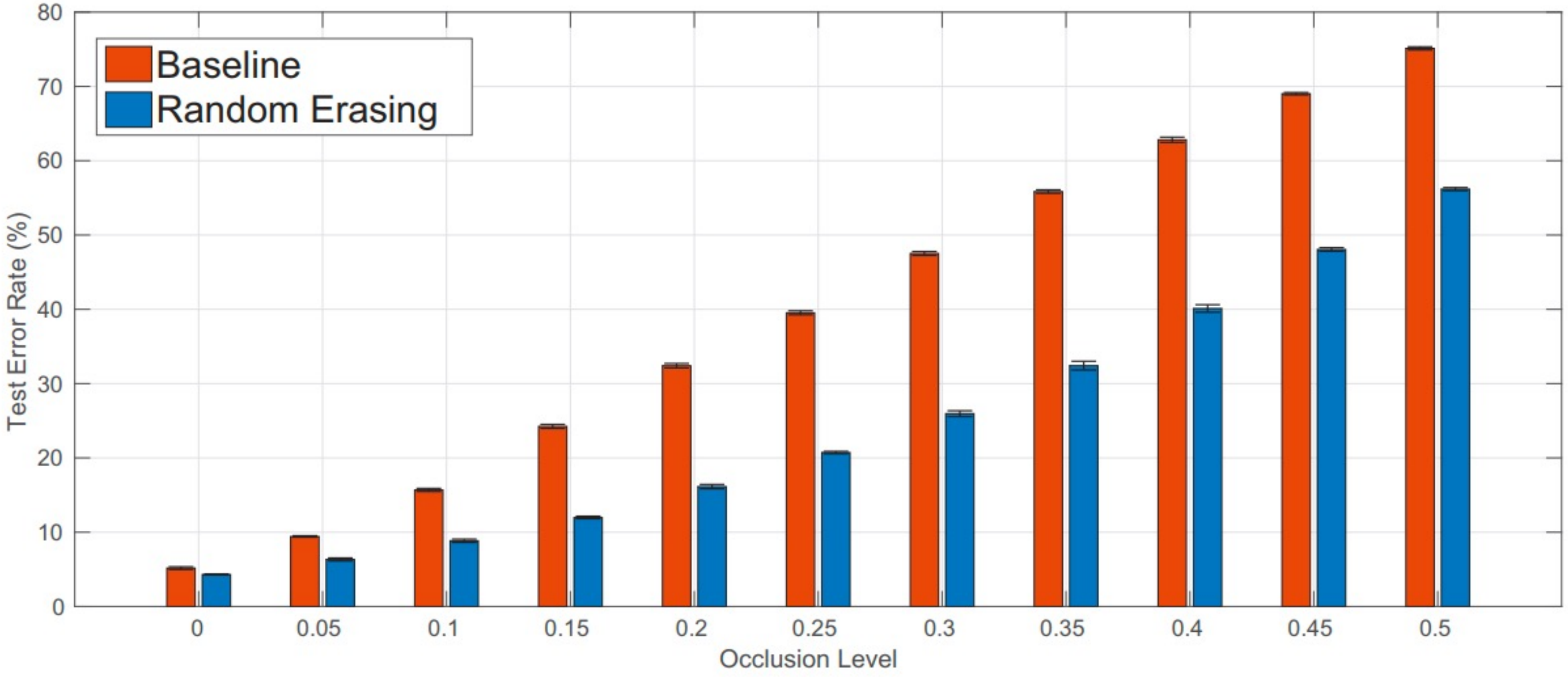


Figure 3: Test errors (%) under different levels of occlusion on CIFAR-10 based on ResNet18 (pre-act). The model trained with Random Erasing is more robust to occlusion.

https://abdullah-mamun.com
a.mamun@asu.edu