Random Erasing Data Augmentation

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

†Cognitive Science Department, Xiamen University, China

§University of Technology Sydney, Australia

August 28, 2019

Presented by Fatir Qureshi

Overview

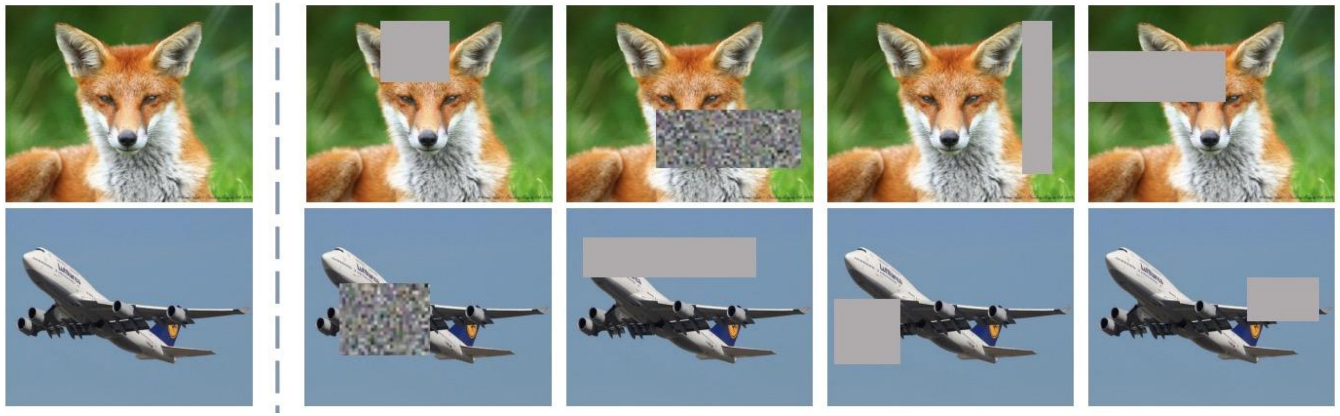
- Introduction
- Random Erasing Overview
- Related Work
- Experiments
- Conclusions

Introduction

- Over fitting is one of the primary challenges which impacts the utility of CNNs
- Many techniques have been introduced to improve the generalization ability of CNNs
  - Random cropping
  - Flipping
  - Dropout
- Occlusion is a significant factor that influences the generalizability of CNNs

Introduction

- Collected training samples usually exhibit limited variance in occlusion
- Manually added occluded images are limited and costly
- These issues necessitate the need for a novel data augmentation approach enhancing robustness to occlusion



Random Erasing

- Random erasing is the proposed approach to address the issue of the occlusion problem in CNNs
- Principle advantages of this method are:
  - Does not require any extra parameter learning or memory consumption
  - Improving the robustness of CNNs to partially occluded samples
  - Consistently improving the performance of recent state-of-the-art deep models (as of 2017) on image classification, object detection, and person re-ID

Random Erasing Overview

- During training an image within a mini-batch randomly undergoes either of the two operations:
  - 1) Kept unchanged
  - 2) randomly choose a rectangle region of an arbitrary size, and assign the pixels within the selected region with random values
- Thus, in Operation 2), an image is partially occluded in a random position with a random sized mask

2

1

Method overview

- Random Erasing randomly selects a rectangle region in an image I denoted by I_e
- The area of the image is S = W × H. and the erasing region is Se
- Randomly initialize the area of erasing rectangle region to Se, where Se/S is in range specified by minimum sl and maximum sh
- The aspect ratio of erasing rectangle region is randomly initialized between r1 and r2, we set it to re
- The size of Ie is thus governed by $H_e = \sqrt{S_e \times r_e}$ and $W_e = \sqrt{\frac{S_e}{r_e}}$

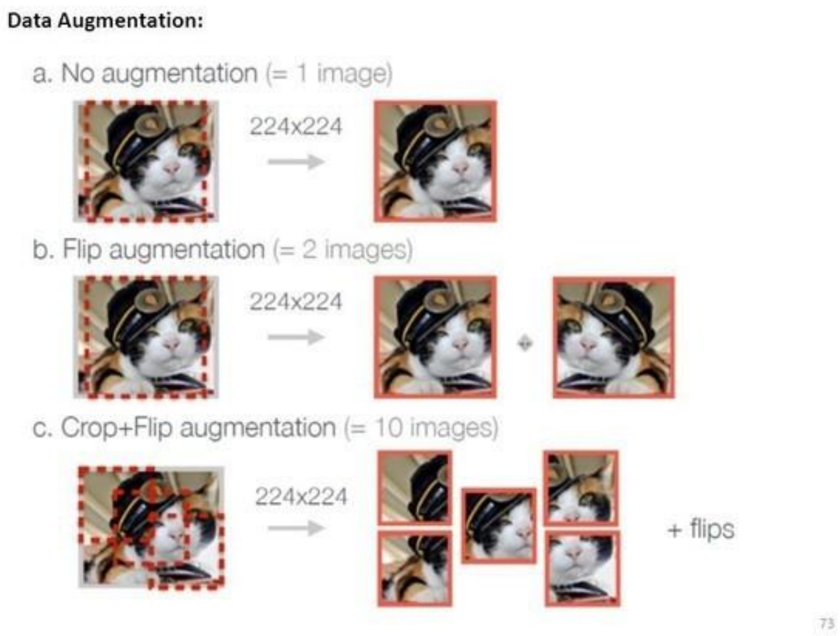
Algorithm 1: Random Erasing Procedure

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

Similarity to other techniques

- Random flipping, random cropping and Dropout function in several of the same ways as random erasing
- Random flipping does not lead to information loss during augmentation
- Random erasing is distinct from random cropping as it only results in partial occluding
- Random Erasing can be described as similar to performing Dropout on the image level



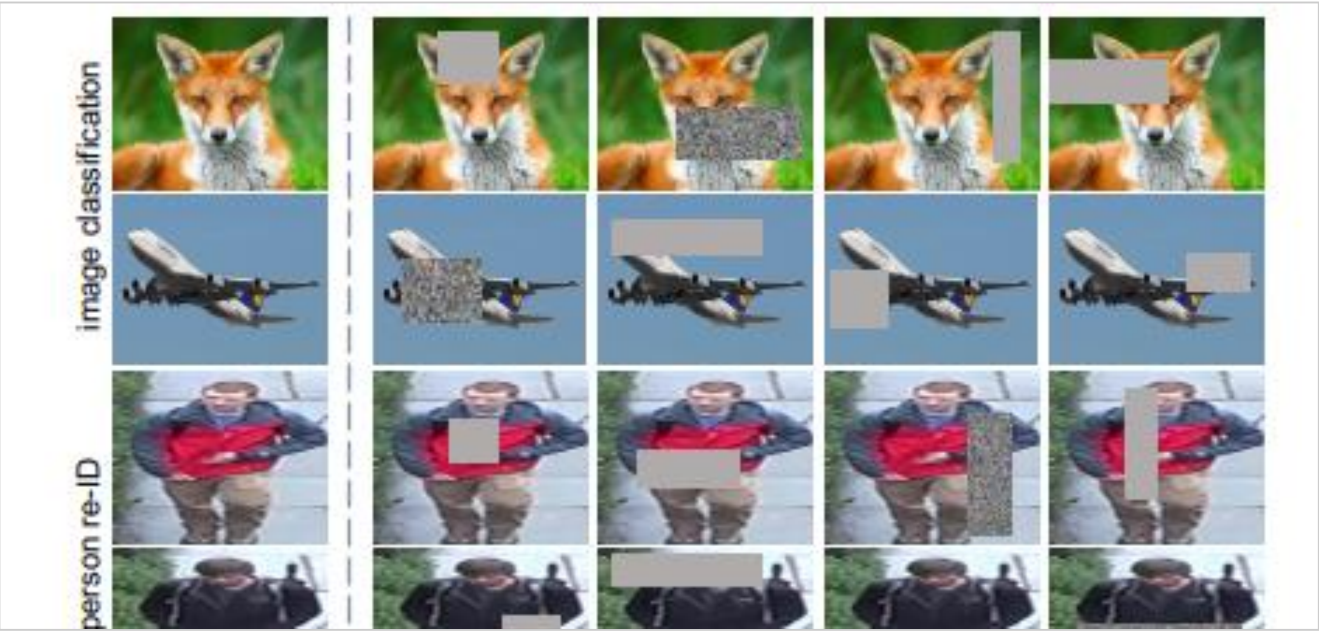
Random Erasing

- Regularization is a key component in preventing overfitting in the training of CNN models
- Data augmentation is an explicit form of regularization that is also widely used in the training of deep CNN
- Random flipping and random cropping are the two most widely used data augmentation techniques
- Random erasing serves as an analogous choice for data augmentation
- Random erasing can be combined with other techniques to create more versatile augmented training data

Experiments

- Three tasks were evaluated with the introduction of random erasing
  - Image Classification
    - Utilization of CIFAR-10 and CIFAR-100
    - 50,000 training and 10,000 testing 32×32 color images drawn from 10 and 100 classes, respectively
  - Object Detection
    - PASCAL VOC 2007 was utilized
    - 9,963 images of 24,640 annotated objects in training/validation and testing sets
  - Person re-identification
    - Market-1501 was utilized
    - 32,668 labeled bounding boxes of 1,501 identities captured from 6 different cameras

Random Erasing for Image Classification



- In image classification, an image is classified according to its visual content
- Training data does not provide the location of the object for this task
- For this reason, random erasing is carried out over the whole of the image
- Four architectures were examined using data erasing augmented datasets: ResNet, RESNeXt, pre-activation ResNet and Wide Residual Networks

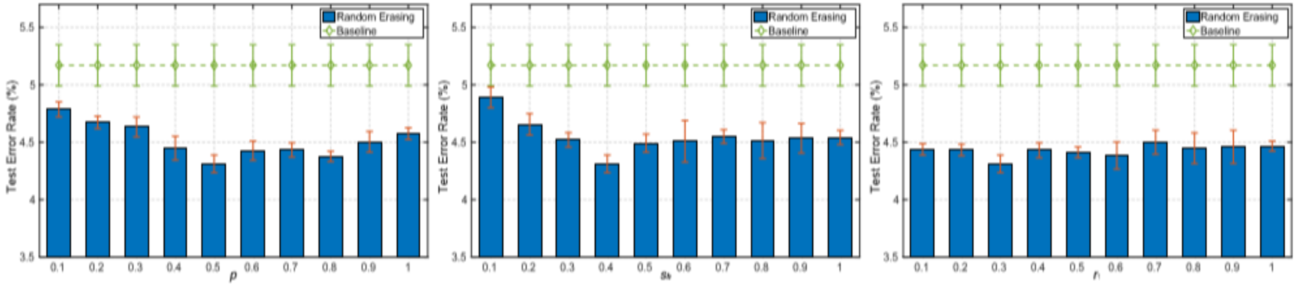
Model	CIFAR-10		CIFAR-100		Fashion-MNIST	
	Baseline	Random Erasing	Baseline	Random Erasing	Baseline	Random Erasing
ResNet-20	7.21 ± 0.17	6.73 ± 0.09	30.84 ± 0.19	29.97 ± 0.11	4.39 ± 0.08	4.02 ± 0.07
ResNet-32	6.41 ± 0.06	5.66 ± 0.10	28.50 ± 0.37	27.18 ± 0.32	4.16 ± 0.13	3.80 ± 0.05
ResNet-44	5.53 ± 0.08	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-20-PreAct	7.36 ± 0.11	6.78 ± 0.06	30.58 ± 0.16	30.18 ± 0.13	4.43 ± 0.19	4.02 ± 0.09
ResNet-32-PreAct	6.42 ± 0.11	5.79 ± 0.10	29.04 ± 0.25	27.82 ± 0.28	4.36 ± 0.02	4.00 ± 0.05
ResNet-44-PreAct	5.54 ± 0.16	5.09 ± 0.10	25.22 ± 0.19	24.10 ± 0.26	4.92 ± 0.30	4.23 ± 0.15
ResNet-56-PreAct	5.28 ± 0.12	4.84 ± 0.09	24.14 ± 0.25	22.93 ± 0.27	4.55 ± 0.30	3.99 ± 0.08
ResNet-110-PreAct	4.80 ± 0.09	4.47 ± 0.11	22.11 ± 0.20	20.99 ± 0.11	5.11 ± 0.55	4.19 ± 0.15
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

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

Random Erasing for Image Classification and Person Re-identification

- The results of applying Random Erasing on CIFAR-10 ,CIFAR-100 and Fashion-MNIST with different architectures are shown in the table above
- Results indicate that models trained with Random Erasing have significant improvement, demonstrating that the authors method is applicable to various CNN architectures

Impact of hyper-parameters



(a) probability p (b) area ratio s_h (c) aspect ratio r_1

Figure 4. Test errors (%) under different hyper-parameters on CIFAR-10 with using ResNet18 (pre-act).

- To demonstrate the importance of hyper parameters in this approach, different permutations of parameters were evaluated
- There are three hyperparameters that govern random erasing
  - Area ratio range of erasing region
  - Aspect ratio range of erasing region
  - The erasing probability p
- Erasing consistently outperforms the ResNet18 (pre-act) baseline under all parameter settings

Four types of random values for erasing

- The authors examined four distinct ways in which pixels in erased regions would be replaced
  - each pixel is assigned with a random value ranging in [0, 255]
  - All pixels are assigned with the mean ImageNet pixel value
  - All pixels are assigned with 0, denoted as RE-0
  - All pixels are assigned with 255, denoted as RE-255
- The authors observed that all erasing schemes outperform the baseline
- The first and second methods listed above performed the best

Types of Erasing Value	Baseline	RE-R	RE-M	RE-0	RE-255
Test error rate (%)	5.17 ± 0.18	$\mathbf{4.31 \pm 0.07}$	4.35 ± 0.12	4.62 ± 0.09	4.85 ± 0.13

Comparison to variant methods

Method	Test error (%)	Method	Test error (%)
Baseline	5.17 ± 0.18	Baseline	5.17 ± 0.18
Random Erasing	$\mathbf{4.31 \pm 0.07}$	Random Erasing	$\mathbf{4.31 \pm 0.07}$
Dropout	Test error (%)	Random Noise	Test error (%)
$\lambda_1 = 0.001$	5.37 ± 0.12	$\lambda_2 = 0.01$	5.38 ± 0.07
$\lambda_1 = 0.005$	5.48 ± 0.15	$\lambda_2 = 0.05$	5.79 ± 0.14
$\lambda_1 = 0.01$	5.89 ± 0.14	$\lambda_2 = 0.1$	6.13 ± 0.12
$\lambda_1 = 0.05$	6.23 ± 0.11	$\lambda_2 = 0.2$	6.25 ± 0.09
$\lambda_1 = 0.1$	6.38 ± 0.18	$\lambda_2 = 0.4$	6.52 ± 0.12

Table 3. Comparing Random Erasing with dropout and random noise on CIFAR-10 with using ResNet18 (pre-act).

- Comparing this method to Dropout and random noise demonstrates its advantages as a variant method
- Applying dropout or adding random noise at the image layer fails to improve the accuracy

Comparing with data augmentation method

Method	RF	RC	RE	Test errors (%)
Baseline				11.31 ± 0.18
	✓			8.30 ± 0.17
		✓		6.33 ± 0.15
			✓	10.13 ± 0.14
	✓	✓		5.17 ± 0.18
	✓		✓	7.19 ± 0.10
		✓	✓	5.21 ± 0.14
	✓	✓	✓	**4.31 ± 0.07**

Table 4. Test errors (%) with different data augmentation methods on CIFAR-10 based on ResNet18 (pre-act). **RF**: Random flipping, **RC**: Random cropping, **RE**: Random Erasing.

- When applied alone, random cropping (6.33%) outperforms the other two methods
- Random Erasing and the two competing techniques are complementary
- Combining these three methods achieves 4.31% error rate

Random Erasing for Object Detection

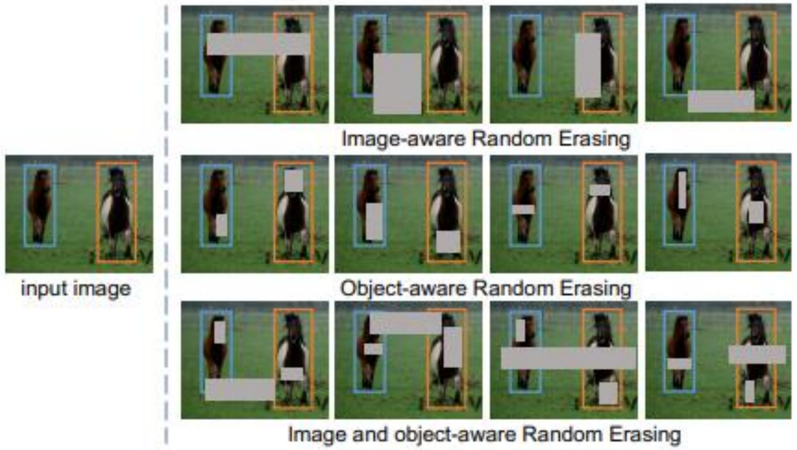


Figure 2. Examples of Random Erasing for object detection with Image-aware Random Erasing (**IRE**), Object-aware Random Erasing (**ORE**) and Image and object-aware Random Erasing (**I+ORE**).

- The goal of object detection is to detect instances of semantic objects of a certain image class
- Three main schemes of random erasing
  - Image-aware Random Erasing- selecting erasing regions on the whole image, which is the same as with image classification
  - Object-aware Random Erasing - selecting erasing regions in the bounding box of each object
  - Image and object-aware Random Erasing- selecting erasing regions in both the whole image and each object bounding box

Random Erasing for Object Detection

- Experiment is conducted based on the Fast-RCNN detector
- The authors report results with using IRE, ORE and I+ORE during training Fast-RCNN
- Two training sets are utilized
  - VOC07
  - The union of VOC07 and VOC12

Detection Evaluation

Method	train set	mAP	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	r
FRCN [7]	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	
FRCN⋆ [27]	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	
ASDN [27]	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	
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	
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	
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	
FRCN [7]	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	
FRCN⋆ [27]	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	
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	
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	
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	

Table 5. **VOC 2007 test** detection average precision (%). FRCN⋆ refers to FRCN with

- It was observed that Object-aware Random Erasing (ORE) performs slightly better than Image-aware Random Erasing (IRE)
- The authors approach (I+ORE) outperforms A-FastRCNN
- Parameter learning is not utilized in this approach to image detection in contrast to other methods

Person re-identification

- Person re-identification was another application that was evaluated using random erasing
- This method serves as a special case of image classification and used the same Image-aware Random Erasing approach
- Random erasing was compared against three baselines
  - IDdiscriminative Embedding (IDE)
  - TriNet
  - SVDNet

Method	Model	RE	Market-1501		DukeMTMC-reID		CUHK03 (labeled)		CUHK03 (detected)	
			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

Table 6. 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 [41].

Person Re-identification

- The results of Random Erasing on Market-1501, DukeMTMC-reID, and CUHK03 with different baselines and architectures are shown
- Random erasing improves baseline models

Conclusions

- Random data erasing is proposed as a novel data augmentation technique with significant promise
- Experiment conducted on CIFAR10, CIFAR100, and Fashion-MNIST with various architectures validate the effectiveness of this method
- Reasonable improvement on object detection and person re-identification are observed through this approach

Questions?