Exploring and Extending Random Erasing Data Augmentation

Anonymous Authors[1]

Abstract

This study delves into the enhancement of Random Erasing, an innovative data augmentation technique for Convolutional Neural Networks (CNNs). It begins with the re-implementation of the original Random Erasing methodology, applied to the Fashion-MNIST dataset using a custom-designed 1-channel input ResNet-20 model. Subsequently, novel adaptations are introduced, including the use of random shapes for erasure and the implementation of edge blurring. These advancements are aimed at bolstering the robustness of CNNs against various occlusions and image variations. The effectiveness of these refined techniques is assessed through classification accuracy and test error metrics, revealing notable performance improvements over the standard Random Erasing method. The findings of this research highlight the efficacy of advanced Random Erasing methods in enhancing the generalization capabilities of CNNs, thereby offering significant contributions to their applicability in real-world scenarios.

1. Introduction

The quest for improved generalization in Convolutional Neural Networks (CNNs) continues to be a pivotal area of research focus, particularly in addressing the challenge of overfitting in complex models. Overfitting typically manifests in models with a high parameter-to-sample ratio, leading to scenarios where the model captures noise instead of discerning the underlying data patterns (Zhong et al., 2020). This issue is notably apparent when models exhibit high performance on training data but underperform on new, unseen data. To enhance the generalization abilities of CNNs, various data augmentation and regularization techniques have been proposed, such as random cropping (Krizhevsky et al., 2012), flipping (Simonyan & Zisserman, 2015), dropout (Srivastava et al., 2014), and batch normalization (Szegedy et al., 2016).

Among these techniques, Random Erasing (RE) has emerged as an effective method, particularly in addressing challenges related to occlusions. Occlusions significantly affect the generalization capacity of CNNs, necessitating the ability to recognize objects even when partially obscured. However, standard training samples often lack sufficient variability in occlusion, leading to diminished model performance on partially occluded images. While the manual addition of occluded images to the training set is an option, it tends to be resource-intensive and limited in scope.

In response to the potential of RE in surmounting these challenges, this study aims to re-implement and expand upon the standard RE technique. The initial phase of the research involves replicating the RE method using a custom-designed 1-channel input ResNet-20 model applied to the Fashion-MNIST dataset. This foundational stage paves the way for the exploration of innovative variations and extensions to the RE approach. These extensions include the implementation of random shapes for erasure and the strategic blurring of erased region edges to minimize the introduction of artifacts typically associated with standard RE. The goal of these enhancements is to evaluate whether such modifications can bolster the robustness of CNN models, particularly in the context of diverse occlusions and image variations.

The research presented in this paper comprises a comprehensive study that not only replicates but also extends the existing RE methodology. It endeavors to demonstrate that the judicious customization and augmentation of Random Erasing can significantly elevate the generalization capabilities of CNNs. The evaluation of these enhancements is conducted using classification accuracy and test error metrics, with a focus on their effects across various test datasets. The findings of this research are intended to contribute to the evolving landscape of data augmentation techniques, offering insights with potential significance for real-world applications, especially in scenarios characterized by data variability and frequent occlusions.

2. Review of paper to implement or extend

2.1. Storyline

High-level motivation/problem The challenge of improving the generalization of Convolutional Neural Networks (CNNs) in computer vision tasks, emphasizing how data augmentation plays a crucial role in enhancing model robustness (Zhong et al., 2020).

Prior work on this problem Traditional data augmentation techniques like flipping, cropping, and rotation have been employed in deep learning models to boost their performance. There's also Dropout, which can be seen as a kind of structural regularization method in neural networks (Zhong et al., 2020).

Research gap Existing augmentation techniques may not comprehensively cater to various object occlusion scenarios. An augmentation method that randomly erases parts of the input image might better simulate real-world scenarios (Zhong et al., 2020).

Contributions Introduction of a novel "Random Erasing" data augmentation technique. Demonstrated its effectiveness on several datasets, including CIFAR-10, CIFAR-100, and Fashion-MNIST. Outperformed other methods in certain tasks, e.g., person re-identification (Zhong et al., 2020).

2.2. Proposed solution

The paper introduces a novel data augmentation technique termed "Random Erasing." At the heart of this approach is a simple yet effective strategy: within an image, a rectangular region is randomly chosen and then filled with random values. This procedure enhances the robustness of models by presenting them with altered visual information. What makes Random Erasing particularly appealing is its ease of implementation. Furthermore, it can be seamlessly integrated into existing Convolutional Neural Network (CNN) architectures. Algorithm 1 provides the pseudocode detailing how to implement the Random Erasing data augmentation technique (Zhong et al., 2020).

2.3. Claims-Evidence

Claim 1: Random Erasing consistently outperforms the baseline across various datasets.

Evidence 1: Table 1 shows that Random Erasing has lower error rates on CIFAR-10, CIFAR-100, and Fashion-MNIST.

Claim 2: Random Erasing is robust against variations in hyper-parameters.

Evidence 2: Figure 1 shows that Random Erasing consistently outperforms the ResNet18 (pre-act) baseline under all parameter settings in a range of hyper-parameters values.

Claim 3: Random Erasing improves the performance under different levels of occlusion.

Evidence 3: Figure 2 shows that Random Erasing consistently outperforms the ResNet18 (pre-act) baseline under different levels of occlusion on CIFAR-10.

Algorithm 1 Random Erasing Procedure (Zhong et al., 2020)

Input: Input image I; Image size W and H; Area of image S; Erasing probability p; Erasing area ratio range s_1 and s_2; Erasing aspect ratio range r_1 and r_2.
Output: Erased image I^*.
Initialization: $p_1 \leftarrow \text{Rand}(0, 1)$
if $p_1 \geq p$ **then**
  $I^* \leftarrow I$;
  return I^*;
else
  while true **do**
    for $i = 1$ **to** $m - 1$ **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 if
    end for
  end while
end if

2.4. Critique and Discussion

"Random Erasing Data Augmentation" introduces a simple yet effective technique that randomly selects a rectangle region in an image and erases its pixels. At its core, the strategy is designed to make models more robust and less sensitive to the specific details of an image, driving them to learn more diverse and representative features.

One of the main strengths of Random Erasing is its simplicity. Unlike some data augmentation techniques that require complex computations or transformations, Random Erasing can be implemented easily and integrated into any training pipeline with minimal effort. This simplicity also means it can be combined with other augmentation techniques, potentially offering a compounded regularization effect.

However, there are some potential concerns and areas for improvement. The random nature of the erasing process might sometimes lead to the removal of critical image features. For instance, in the context of facial recognition, if the erasing process consistently removes the eyes, the model might struggle with recognizing certain facial features that are essential for accurate identification.

It's also worth noting that the method involves various parameters, such as the aspect ratio and proportion of the area to be erased. An in-depth exploration of the sensitivity of these parameters would provide users with a clearer guide on the optimal settings for various datasets and tasks.

Lastly, while Random Erasing offers a fresh perspective on data augmentation, it's crucial to evaluate its utility against the backdrop of evolving techniques in the data augmentation landscape. As newer methods emerge, it would be beneficial to understand where Random Erasing stands in comparison and in combination with these methods.

Furthermore, while the paper has shown effectiveness across a variety of tasks, it would be interesting to explore its impact in more specialized domains, such as medical imaging. In domains where certain regions of the image hold significant importance, random erasing might not always be the best choice.

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	$\mathbf{3.08} \pm \mathbf{0.05}$	$\mathbf{18.49} \pm \mathbf{0.11}$	$\mathbf{17.73} \pm \mathbf{0.15}$	$\mathbf{4.01} \pm \mathbf{0.10}$	$\mathbf{3.65} \pm \mathbf{0.03}$
ResNeXt-8-64	$\mathbf{3.54} \pm \mathbf{0.04}$	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. RE: Random Erasing. Source: (Zhong et al., 2020)

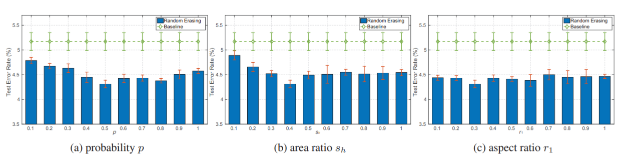

Figure 1. Test errors (%) under different hyper-parameters on CIFAR-10 with using ResNet18 (pre-act). Source: (Zhong et al., 2020)

3. Implementation

3.1. Implementation motivation

The primary objective of this study is to explore the versatility of the Random Erasing technique further by trialing a range of modifications, such as employing varied shapes for erasing, and blurring the edge of the erased region. In real-world contexts, objects often face multiple occlusions that come in shapes beyond mere rectangle which is used as the erased region in the standard RE technique. Introducing varied shapes aims to mirror these varied shape occlusion scenarios, potentially boosting the model's generalization and resilience.

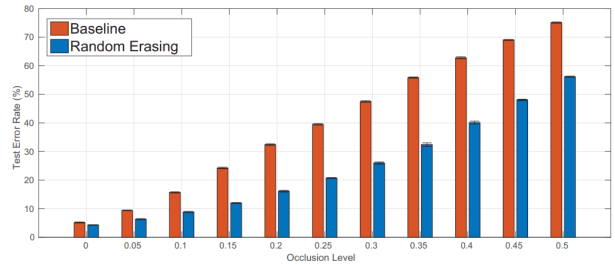

Figure 2. Test errors (%) under different levels of occlusion on CIFAR-10 based on ResNet18 (pre-act). The model trained with Random Erasing is more robust to occlusion. Source: (Zhong et al., 2020)

A prevalent issue with several data augmentation techniques is the unintentional introduction of artifacts or patterns that do not align with real-world data characteristics. While these anomalies can sometimes aid models during the training phase, they might adversely affect their generalization capabilities in practical settings. A particular concern is the sharp edges between erased regions and other pixels in a image data. The approach of blurring these edges seeks to counteract the unintended effects of the Random Erasing method, making the augmented data more representative of natural images.

The overarching goal of these experiments is to assess the improvements in CNN models' ability to handle a spectrum of image disturbances, aiming for heightened real-world relevance.

3.2. Implementation setup and plan

Code Base: PyTorch

Datasets: Fashion-MNIST

Methods: Random Erasing as primary, extend to erasing with various shapes rather than rectangle, and softening the boundaries of the erased regions.

Architectures: 1-channel input ResNet-20.

Epochs: 300 as primary, 100 to experiment the relation between complex RE approaches and small epochs.

Learning rate decay schedule: Starts from 0.1, decays to 0.01 after 50% of epochs, and to 0.001 after 75% of epochs.

Transforms: Extended Random Erasing, Random Horizontal Flip, Random Cropping with 4px padding.

Evaluation metrics:
Classification accuracy is used to evaluate the model's performance. Defined as the percentage of correct predictions made by the model, the formula is:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \times 100\%$$

A higher classification accuracy suggests that the model is more effective in classifying unseen data. Note that it can also be inversely represented as test error in this paper, where:

$$\text{Test Error} = 100\% - \text{Accuracy}$$

Code reuse: Existing ResNets architectures and train functions. (Idelbayev)

Code to write: 1-channel input ResNet-20, circular and random shapes Random Erasing functions, edge blurring functions.

3.3. Details

Following are the addtional features compare to the original Random Erasing data augmentation method.

shape: Determines the shape of the erased region.

blend_edges: A boolean that decides if the edges of the erased region should be blended or not.

blend_type: Specifies the type of blending applied to the edges. Options include linear, sigmoid, quadratic, cubic, gaussian, and cosine.

blend_factor: An integer that impacts the extent of blurring on the edges. A smaller value results in more pronounced blurring.

Visual representation of updated implementation Figure 4 showcases a visual depiction of datasets processed using the enhanced data augmentation techniques. For a comparative perspective, Figure 3 illustrates the outcomes of the original Random Erasing data augmentation method.

3.4. Results and Interpretation

The empirical investigation into various Random Erasing (RE) data augmentation strategies on ResNet models, specif-

Figure 3. Dataset processed using the original Random Erasing data augmentation technique

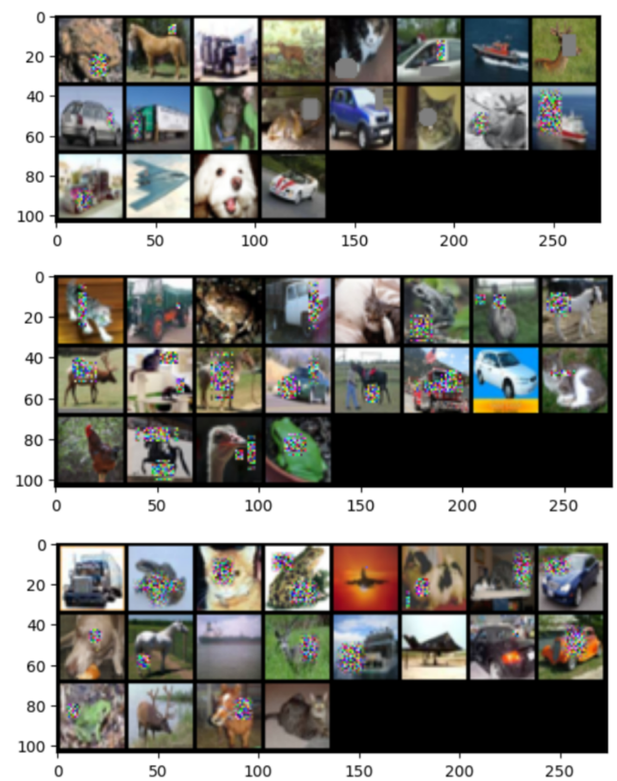

Figure 4. Figures showcasing datasets post-application of the refined data augmentation techniques. The images exhibit multiple erased regions, including circular shapes. Additionally, the edges of these erased regions demonstrate varying degrees of blurring.

Table 2. Test error rates (%) for various ResNet models on Fashion-MNIST using various Random Erasing (RE) augmentations (cs: circular shape, rs: random shape, be: blurred edge).

Model	Fasion-MNIST						
	Baseline	RE	RE cs	RE rs	RE be	RE cs-be	RE rs-be
Baseline	5.26 ± 0	15.43 ± 0.46	10.08 ± 0.17	12.18 ± 0.17	9.27 ± 0.30	8.78 ± 0.09	8.73 ± 0.20
RE	4.68 ± 0	5.23 ± 0.09	5.73 ± 0.09	5.40 ± 0.14	5.05 ± 0.13	5.21 ± 0.11	5.31 ± 0.14
RE cs	4.81 ± 0	7.45 ± 0.17	5.07 ± 0.12	6.31 ± 0.16	5.84 ± 0.18	5.03 ± 0.14	5.07 ± 0.13
RE rs	4.79 ± 0	5.48 ± 0.11	4.90 ± 0.05	5.18 ± 0.13	5.26 ± 0.11	5.01 ± 0.05	5.03 ± 0.06
RE be	4.60 ± 0	5.31 ± 0.18	5.51 ± 0.22	5.32 ± 0.12	5.02 ± 0.08	5.07 ± 0.12	5.05 ± 0.14
RE cs-be	4.50 ± 0	6.90 ± 0.18	4.77 ± 0.05	5.76 ± 0.15	5.24 ± 0.04	4.82 ± 0.09	4.75 ± 0.04
RE rs-be	4.51 ± 0	5.89 ± 0.12	4.84 ± 0.08	5.25 ± 0.04	4.97 ± 0.12	4.82 ± 0.10	4.76 ± 0.08

ically applied to the Fashion-MNIST dataset, is encapsulated in the test error rates shown in Table 2. These rates, presented as percentages along with standard deviations, serve as a measure of the models' robustness and their ability to generalize in the presence of occluded or altered image data. The findings from this study offer valuable benchmarks for assessing the effectiveness of different RE techniques in enhancing model performance under varied conditions.

One of the most striking observations is the baseline model's performance degradation when confronted with RE augmentation (15.43 ± 0.46%), as opposed to its training conditions (5.26 ± 0). This notable discrepancy underscores the model's limited adaptability to unfamiliar occlusion patterns, an aspect that resonates with findings from similar studies in the field (Zhong et al., 2020). In contrast, models trained with RE augmentation demonstrated enhanced capability in handling occlusions, as evidenced by significantly reduced test error rates against RE-augmented data (5.23 ± 0.09%). This improvement not only reflects the model's increased resilience, a trait crucial for real-world applications where occlusion variability is common, but also aligns with the broader literature on data augmentation's impact on model robustness.

The introduction of shape-specific occlusions, namely circular shape (RE cs) and random shape (RE rs), initially hypothesized to improve generalization, paradoxically resulted in higher error rates on baseline data. This counterintuitive outcome, suggesting potential overfitting to specific occlusion patterns, was effectively mitigated by integrating blurring with these shape alterations (RE cs-be and RE rs-be). The models trained with these combined augmentations (RE cs-be and RE rs-be) not only performed notably better on corresponding augmented data but also on baseline data (4.50 ± 0% and 4.51 ± 0%, respectively), outperforming models trained with standard RE and solely with blurred edges (RE be) (4.68 ± 0% and 4.60 ± 0%, respectively). Moreover, these models maintained comparable or superior performance on datasets processed with circular and random shape erasings compared to RE cs and RE rs, indicating a more nuanced understanding of occlusion handling.

The role of the blurring technique in enhancing the models' generalization capabilities beyond occluded training data emerges as a pivotal finding. Blurring effectively softens the abrupt transitions caused by occlusions, reducing shape-specific artifacts, thus preventing overfitting to pronounced occlusion patterns. This adaptability enables the models to maintain robustness and generalize effectively when exposed to clean, non-augmented baseline images, while being equipped to handle a variety of occlusion patterns encountered in real-world scenarios. When combined with RE that incorporates different shapes, blurring contributes to creating models that are versatile and applicable in diverse practical applications.

These insights suggest a refined approach in designing RE augmentation strategies, emphasizing the importance of combining shape alterations with edge blurring. This balanced approach introduces the necessary variability while preserving essential image characteristics, significantly enhancing CNN robustness and generalization across varied scenarios. The findings advocate for strategic use of advanced Random Erasing methods, especially those incorporating edge blurring, in developing robust and generalizable CNNs. This contribution is novel in its demonstration of how specific augmentation techniques can be optimized for better model performance, presenting a potential avenue for future research in exploring further augmentation combinations and their effects on different neural network architectures and application domains.

References

Idelbayev, Y. Proper ResNet implementation for CIFAR10/CIFAR100 in PyTorch. https://github.com/akamaster/pytorch_resnet_cifar10. Accessed: 2023-10-13.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. In Pereira, F., Burges, C., Bottou, L., and

Weinberger, K. (eds.), *Advances in Neural Information Processing Systems*, volume 25. Curran Associates, Inc., 2012. URL https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf.

Simonyan, K. and Zisserman, A. Very deep convolutional networks for large-scale image recognition. In *International Conference on Learning Representations*, 2015.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958, 2014. URL http://jmlr.org/papers/v15/srivastava14a.html.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. Rethinking the inception architecture for computer vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, 2016. doi: 10.1109/CVPR.2016.308.

Zhong, Z., Zheng, L., Kang, G., Li, S., and Yang, Y. Random erasing data augmentation. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(07): 13001–13008, Apr. 2020. doi: 10.1609/aaai.v34i07.7000. URL https://ojs.aaai.org/index.php/AAAI/article/view/7000.