MoireDB: A Formula-driven Image Dataset for Robustness Enhancement

Anonymous CVPR submission

Paper ID 15

Abstract

Image recognition models have struggled to achieve robustness against real-world degradations and adversarial attacks. In this context, data augmentation methods like PixMix have been shown to enhance robustness. The PixMix framework utilizes generative Fractal arts and Feature Visualizations of CNNs (FVis) as mixing images, which are combined with images from the original training dataset. However, these mixing images suffer from copyright restrictions and high construction costs. To address these challenges, we propose Moire DataBase (MoireDB), a formula-driven Moiré image dataset. MoireDB eliminates copyright concerns, reduces dataset construction costs compared to previous mixing images, and effectively diversifies the perturbations applied to the original images during training. Since each Moiré image is generated from simple mathematical formulas, MoireDB is computationally efficient, eliminating the need for advanced image generation AI and minimizing resource consumption. Experiments on CIFAR-C and CIFAR-based adversarial robustness demonstrate that MoireDB-augmented images in the CIFAR training dataset partially outperform traditional augmentations based on Fractal arts and FVis.

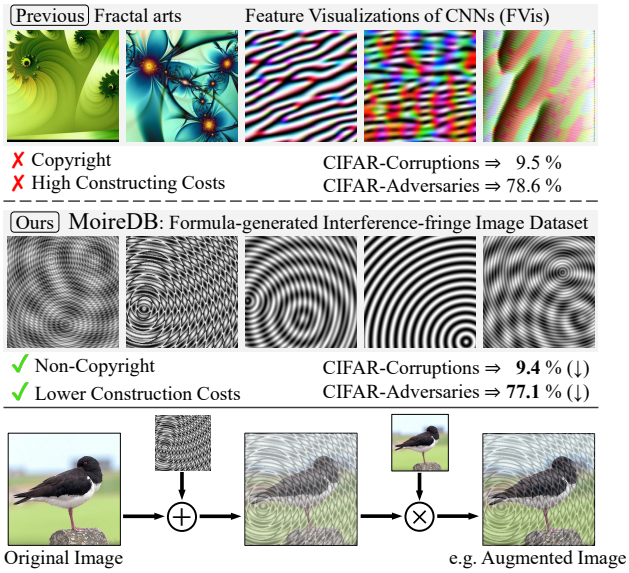

Figure 1. Example of data augmentation using MoireDB. MoireDB eliminates copyright issues and can be constructed with lower computational costs than Fractal arts and Feature Visualizations (FVis) [4]. Moreover, MoireDB achieved the best performance in CIFAR corruption benchmarks [3] and adversarial robustness metrics under certain experimental conditions.

1. Introduction

Image recognition techniques, particularly those based on deep learning, are promising for real-world applications; however, image classification using deep learning models is known to be less robust than human visual perception when faced with diverse real-world degradations, such as noise and blurring, as well as adversarial attacks [3, 8].

One promising technique for improving the robustness of image recognition models for classification task is *data augmentation* such as Mixup [16] and CutMix [14]. Using these data augmentation methods, we can increase image counts while reducing overfitting, thus potentially improving robustness. A notable data augmentation method known for enhancing robustness is PixMix [4], which extends training datasets by blending training images with synthetic images from predefined mixing sets. The mixing set images used by PixMix include mathematically generated Fractal arts and feature visualizations (FVis). Fractal arts are collected on DeviantArt[1]. The images are visually diverse. FVis are collected using OpenAI Microscope[2]. The images are created from convolutional neural networks (CNNs) such as AlexNet [7], VGGNet [12], and ResNet [2], which are basically pre-trained on ImageNet [10, 17].

However, the use of these images entail at least two practical disadvantages: i) Some of the human-designed digital patterns and generative arts are protected by copyright, and thus commercial use of PixMix data augmentation remains questionable. ii) Generating FVis requires multiple CNNs

[1]https://www.deviantart.com/
[2]https://openai.com/index/microscope/

trained on large image datasets, and is thus a high-cost operation for assembling images into datasets.

To address these issues, we propose Moiré DataBase (MoireDB) as a new mixing set for PixMix. All Moiré images in MoireDB are generated by formulas, eliminating copyright concerns and reducing construction costs. Moiré images, similar to FVis, exhibit structured texture patterns. However, unlike FVis, they possess circular structures, which eliminates directional biases. This characteristic is expected to allow any Moiré image to effectively enrich the information of an original image during augmentation. Moreover, despite being generated with a small number of parameters in a simple manner, MoireDB produces textures that induce visual illusions in humans. This dataset may provide insights into how such visually illusion-inducing texture images affect the robustness of image recognition models.

Experimental results confirm that MoireDB partially outperforms Fractal arts and FVis, and it also surpasses other mathematically generated datasets, such as Fractal DataBase (FractalDB) [5] and VisualAtom [13], in both corruption and adversarial robustness.

Our proposal of MoireDB offers several key advantages, including the following.

- The use of MoireDB for data augmentation improves robustness with respect to real-world degradations and adversarial attacks.
- MoireDB contains only formula-generated images, eliminating copyright problems.
- The images constituting MoireDB are auto-generated, which reduces the cost of assembling datasets and eliminates the need for large-scale image generation AI models.

2. Related Work

2.1. Robustness in Image Classification

Digital images are susceptible to noise, compression, and other sources of corruption caused by a broad range of mechanisms. Although such corruption does not prevent human visual perception from identifying images with high accuracy, it *does* significantly reduce the image identification accuracy of image recognition models, and improving the robustness of image recognition models is a central challenge for image recognition research.

The robustness of image recognition models may be quantified by testing on specialized datasets such as ImageNet-C [3] and CIFAR-C, which consist of images that have been corrupted in various ways—such as by adding noise, blurring, weathering, or applying digital transformations—to reflect 15 types of corruption commonly experienced by digital images.

To quantify robustness using ImageNet-C or CIFAR-C, the image classification accuracy is measured for each of the 15 categories of image corruption, and an average is performed over all categories to yield a *mean corruption error* (mCE); smaller mCE values indicate greater robustness.

In addition to quantifying robustness against corruption, robustness may also be quantified against adversaries, i.e., adversarial attacks, by measuring image classification accuracy for special test images in the ImageNet [1] and CIFAR [6] datasets to which adversarial attacks have been applied; again, lower values of the image classification accuracy indicate greater robustness.

2.2. Data Augmentation Using Additional Images

One of the most promising data augmentation methods for enhancing robustness is PixMix [4]. In the PixMix approach to data augmentation, training images from databases such as ImageNet or CIFAR are combined additively or multiplicatively with an auxiliary set of structurally complex images to yield an augmented dataset; deep learning models trained on the augmented dataset then exhibit improved image identification accuracy and robustness compared to models trained on the non-augmented dataset. In the original PixMix proposal, the auxiliary set of structurally complex images included two types of images: Fractal arts and FVis.

Fractal arts (note that this is different from FractalDB [5]) are manually designed images downloaded from DeviantArt; these images contain shapes and color schemes designed to pique the curiosity of human visual perception, and are thus expected to be structurally complex.

FVis are machine-generated images that may be downloaded from OpenAI Microscope. This database allows visualization results for image features [11]—as extracted by various pre-trained CNN models operating on a large image dataset—to be downloaded in the form of image files.

Given an input image dataset, PixMix produces an augmented dataset by performing repeated mixing operations. Specifically, each input image is subjected to a randomly chosen number (at most 4) of mixing steps and in each step, the image is mixed either with an input image or with an image chosen from the auxiliary image set, and the mixing is performed either additively or multiplicatively (chosen at random). Deep learning models trained on PixMix augmented datasets are known to exhibit improved image identification accuracy and robustness compared to other data augmentation methods such as Mixup [16] or CutMix [14].

However, some Fractal arts are protected by copyright, and thus commercial use of PixMix remains questionable. Moreover, both Fractal arts and FVis are enormously costly to generate, and the number of images that may be feasibly assembled into a dataset is limited in practice.

To address these issues, we propose a synthetic image dataset as a new mixing set.

CVPR #15

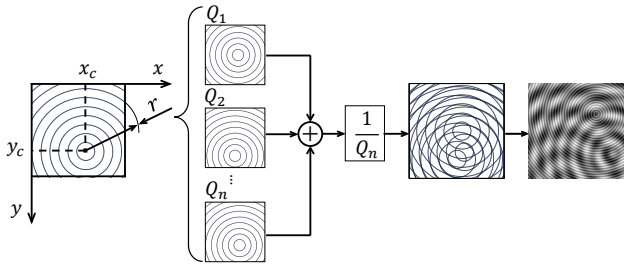

Figure 2. Algorithm for generating Moiré images

3. Proposed Method Details

In the present study, we propose MoireDB, a formula-generated image dataset for data augmentation.

Section 3.1 describes our procedure for generating the Moiré images comprising MoireDB, while Section 3.2 discusses our strategy for data augmentation using generated Moiré images.

This idea of generating images through formulas was inspired by formula-driven supervised learning (FDSL) datasets such as FractalDB [5] and VisualAtom [13]. MoireDB belongs to the same family of FDSL datasets. A key distinction is that MoireDB requires only three parameters to generate each Moiré image, which is significantly fewer than the number used in the rendering processes of FractalDB and VisualAtom. This minimal parameterization allows for more controllable and interpretable image generation while eliminating the need for large-scale image generation AI models.

3.1. Generation of Moiré Images

Our algorithm for generating the Moiré images constituting MoireDB is depicted schematically in Fig. 2. The starting point is a simple procedure (Fig. 2, far left) for generating a concentric-circle pattern; this procedure is described by a formula, discussed below, containing multiple adjustable parameters such as the coordinates (x_c, y_c) of the common center point. To generate a single Moiré image, we invoke this formula multiple times—with randomly chosen values for the adjustable parameters—to yield a set of multiple distinct concentric-circle patterns (Fig. 2, center), then simply *superpose* these to yield the Moiré image (Fig. 2, far right). The superposition of randomly generated concentric-circle patterns gives rise to the characteristic interference fringes of Moiré images, and varying the adjustable parameters defining the concentric-circle patterns allows a wide range of distinct fringe patterns to be realized.

The image representing each concentric-circle pattern is generated by a formula that computes a brightness value for each pixel in the image. Each Moiré image depends on several adjustable parameters: the number Q_n of concentric-circle patterns superposed, and, for each of these patterns, the center-point coordinates (x_c, y_c) and an interval frequency parameter ν described below. Values for all of these parameters are chosen randomly within the ranges listed in Table 1.

Each concentric-circle pattern may be described as a superposition of circles of the form

$$f_{Q_n} = \frac{1}{Q_n} \sum_{k=1}^{m} \eta_k \in \mathbb{R}^2 \tag{1}$$

where m is the number of circles drawn in the pattern and η_k represents the k-th circle. Denoting the radius of this circle by r_k, and recalling that the circle is centered at (x_c, y_c), we may express η_k in the form

$$\eta_k = \begin{cases} x = (r_k \cos\theta + x_c) \times g \\ y = (r_k \sin\theta + y_c) \times g \end{cases} \quad (0 \le \theta < 2\pi) \tag{2}$$

The center-point coordinates (x_c,y_c) are chosen at random from a uniform distribution. The quantity g in this expression, representing the brightness at point (x, y), is a sinusoidally varying function of the radial distance r:

$$g = (V_M(\cos\nu \times \pi \times r) + 1) \times 255, \tag{3}$$

where V_M is the amplitude of the sinusoidal brightness variation. Using the brightness g to define a grayscale value for each pixel yields an image representing the concentric-circle pattern. Choosing the number of concentric-circle patterns $Q_n > 1$ then ensures interference between the patterns, yielding the desired Moiré image.

We set the size of generated images to be 512×512 px; the number of circles m drawn for each concentric-circle pattern is determined as appropriate based on the image size and the interval frequency ν.

3.2. Data Augmentation Using Moiré Images

Our strategy for data augmentation using Moiré images is outlined schematically in the lower part of Fig. 1 shows a detailed diagram of the operational pipeline of our PixMix implementation with Moiré images, in this case for an example involving 1 time additive mixing operation and 1 time multiplicative mixing operations. Other settings in our data augmentation procedure are the same as that used in PixMix.

The number of Moiré images we generate for data augmentation is 14,230, chosen to match the number of Fractal arts used in [4]. For each image, the parameter values in the image-generation formulas are chosen at random from the ranges listed in Table 1. These parameters were determined through several experiments. For each mixing step, we choose an image at random from the set of generated images and mix it either additively or multiplicatively with the selected Moiré image or with the input image.

Table 1. Adjustable parameters for auto-generated Moiré images

Parameter	Symbol	Range
Interval frequency	ν	$0.03 \leq \nu < 0.05$
Center-point coordinates	x_c, y_c	$0 \leq (x_c, y_c) < 600$
Number of superposed concentric-circle patterns	Q_n	$Q_n = \{1, 2, 3\}$

Table 2. Robustness values measured for various data augmentation image datasets [4]. The experiments have been conducted on CIFAR-10-C and -100-C by using MoireDB within the framework of PixMix. Lower is better for the listed scores.

Dataset		Baseline	Fractal arts	FVis	FractalDB	VisualAtom	MoireDB
CIFAR-10-C	Clean	4.4	4.2	4.8	4.0	4.4	4.6
	Corruptions	26.4	10.8	9.5	11.9	10.8	**9.4**
	Adversaries	91.3	82.0	78.6	92.2	93.9	**77.1**
CIFAR-100-C	Clean	21.3	20.3	21.0	20.0	21.4	21.1
	Corruptions	50.0	33.3	**30.3**	35.0	33.4	30.9
	Adversaries	96.8	93.2	**92.3**	98.5	98.5	93.5

4. Experimental Evaluation

4.1. Robustness Tests Procedure

We conducted experimental tests to assess the effectiveness of data augmentation using MoireDB, comparing the results against robustness values obtained via several alternative datasets: data augmentation using Fractal arts and FVis, as originally proposed for PixMix, and PixMix with augmentation images taken from FractalDB and VisualAtom. The data augmentation settings follow the default configurations provided in PixMix's GitHub repository [4]. The number of images in Fractal arts and FVis is the same as in [4], with Fractal arts containing 14,230 images and FVis consisting of 4,700 images. FractalDB contains 1,000 images, while VisualAtom consists of 14,230 images. As mentioned earlier in this paper, MoireDB also contains 14,230 images.

The training model used in our experiments is WideResNet [15]. We use CIFAR as the training image dataset. For each data augmentation strategy, we generate an augmented version of the CIFAR training dataset while keeping the total number of training images fixed at 50,000. We then train WideResNet on the augmented dataset for 100 epochs and evaluate the robustness of the trained model.

Robustness is measured on the CIFAR-C dataset of test images. The Corruptions task involves using CIFAR-C to measure robustness against image corruption [3]. The metric for this assessment is the previously mentioned mCE, which is smaller for greater robustness. mCE is computed as the mean image identification accuracy for the 15 types of image corruption represented by CIFAR-C.

The Adversaries task involves measuring robustness against adversarial attack [9]. The metric for this assessment is the image identification accuracy, with lower values indicating better performance. Adversarial attacks are applied to CIFAR test images.

4.2. Results of Robustness Tests

Table 2 shows the results of tests to assess the impact of MoireDB-based data augmentation on the robustness of image classification. The column labeled "Baseline" presents the results obtained using the original (non-augmented) CIFAR training data, with the values sourced from [4].

From Table 2 we see that data augmentation using MoireDB achieves better image identification robustness than any other method for CIFAR-10-C. When comparing the results for the FDSL datasets—FractalDB, VisualAtom, and MoireDB—we observe that in *every* robustness test for both CIFAR-10-C and CIFAR-100-C, the greatest improvement in robustness is achieved with data augmentation using MoireDB.

These results demonstrate that data augmentation with MoireDB can achieve robustness improvements comparable to or even surpassing those of Fractal arts and FVis-based augmentations.

5. Conclusion

In the present study, we proposed MoireDB, a formula-generated dataset of interference-fringe images for use with the PixMix method of data augmentation, and conducted experiments to assess its impact on robustness. Our results showed that, for several test categories, data augmentation using MoireDB achieved a greater improvement in robustness than data augmentation with Fractal arts or FVis. Furthermore, an important point to mention is that every single image in MoireDB eliminates copyright issues and also reduces construction costs.

References

[1] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *CVPR*, 2009. 2

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *CVPR*, 2016. 1

[3] Dan Hendrycks and Thomas Dietterich. Benchmarking neural network robustness to common corruptions and perturbations. In *ICLR*, 2019. 1, 2, 4

[4] Dan Hendrycks, Andy Zou, Mantas Mazeika, Leonard Tang, Bo Li, Dawn Song, and Jacob Steinhardt. Pixmix: Dreamlike pictures comprehensively improve safety measures. In *CVPR*, 2022. 1, 2, 3, 4

[5] Hirokatsu Kataoka, Kazushige Okayasu, Asato Matsumoto, Eisuke Yamagata, Ryosuke Yamada, Nakamasa Inoue, Akio Nakamura, and Yutaka Satoh. Pre-training without natural images. In *ACCV*, 2020. 2, 3

[6] Alex Krizhevsky and Geoffrey Hinton. Learning multiple layers of features from tiny images. 2009. 2

[7] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In *NeurIPS*, 2012. 1

[8] Raphael Gontijo Lopes, Dong YinandBen Poole, Justin Gilmer, and Ekin Dogus Cubuk. Improving robustness without sacrificing accuracy with patch gaussian augmentation. *arXiv preprint arXiv:1906.02611*, 2019. 1

[9] Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks. In *ICLR*, 2018. 4

[10] Chris Olah, Alexander Mordvintsev, and Ludwig Schubert. Feature visualization, 2017. *Distill*, https://distill.pub/2017/feature-visualization. 1

[11] Chris Olah, Alexander Mordvintsev, and Ludwig Schubert. Feature visualization. *Distill*, 2017. https://distill.pub/2017/feature-visualization. 2

[12] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. In *NeurIPS*, 2014. 1

[13] Sora Takashima, Ryo Hayamizu, Nakamasa Inoue, Hirokatsu Kataoka, and Rio Yokota. Visual atoms: Pre-training vision transformers with sinusoidal waves. In *CVPR*, 2023. 2, 3

[14] Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, and Youngjoon Yoo. Cutmix: Regularization strategy to train strong classifiers with localizable features. In *ICCV*, 2019. 1, 2

[15] Sergey Zagoruyko and Nikos Komodakis. Wide residual networks. In *BMVC*, 2016. 4

[16] Hongyi Zhang, Moustapha Cisse, Yann Dauphin, and David Lopez-Paz. Mixup: Beyond empirical risk minimization. In *ICLR*, 2017. 1, 2

[17] Roland S. Zimmermann, Judy Borowski, Robert Geirhos, Matthias Bethge, Thomas S. A. Wallis, and Wieland Brendel. How well do feature visualizations support causal understanding of cnn activations? In *NeurIPS*, 2021. 1