Mitigating motion blur in images using Encoding Generative Adversarial Network

Antonio Fernández-Rodríguez[1], Jose D. Fernández-Rodríguez[1,2][0000−0003−3702−2230], Ezequiel López-Rubio[1,2][0000−0001−8231−5687], and Miguel A. Molina-Cabello[1,2][0000−0002−8929−6017]

[1] ITIS Software, University of Malaga, Málaga, Spain,
[2] Instituto de Investigación Biomédica de Málaga y Plataforma en Nanomedicina-IBIMA Plataforma BIONAND, Málaga, Spain.
afr012240@uma.es; {josedavid,ezeqlr,miguelangel}@lcc.uma.es

Abstract. Motion blur is an undesired effect in images where objects in motion appear blurred due to the relative motion between the camera and the scene. An effective way against noisy images involves the use of GANs (Generative Adversarial Networks). In particular, the Encoding Generative Adversarial Network (EGAN) enhances the traditional GAN architecture by incorporating an encoder, which identifies the essential features of an image, contributing to noise reduction, and thereby simplifying the image classification task. This paper studies the effectiveness of EGAN in the detection and correction of motion blur noise in images so that it can improve the accuracy of a classifying model.

Keywords: Generative Adversarial Network · Encoder · Image Classification · Motion blur noise

1 Introduction

Motion blur is characterized by residual traces in images caused by the movement of objects while the image is captured [5]. It can be characterized with two parameters: the disturbance angle θ and the disturbance length l.

An effective deblurring strategy involves the use of GANs (Generative Adversarial Networks) [3], while the Encoding Generative Adversarial Network (EGAN) [4] enhances the traditional GAN architecture by incorporating an encoder [1]. The encoder plays a crucial role in identifying the essential features of an image, aiding the generator network in generating an equivalent image. By discerning and capturing key characteristics, the encoder contributes to noise reduction, thereby simplifying the image classification task.

This paper aims to study the effectiveness of the EGAN model for the detection and correction of motion blur noise in images such that it can improve the accuracy of a classifying model. The proposed model is trained on the Modified National Institute of Standards and Technology (MNIST) dataset [2] which contains a large number of handwritten digit images. To assess how well the

EGAN [4] model defends against these noises, we will compare the pre-trained classification model's accuracy with and without the EGAN defense mechanism.

2 Methodology

First, our motion blur model for vision sensors is presented. The linear motion can be defined as a pair (l, θ), where l signifies the magnitude of the displacement and θ denotes the direction of displacement. Starting with a clear, original image, we can create blurred versions of it by first taking its Fourier transform and then followed by applying a motion blur degradation calculated from the parameter vector (l, θ) [5]. Regarding the EGAN architecture, it contains a GAN that produces an image from a latent vector, and an encoder that produces a latent vector from an image [4].

3 Experiments and results

The dataset used for our experiments is the MNIST database [2]. It consists of a set of 28x28 grayscale images of handwritten digits from 0 to 9. The dataset comprises 70,000 images divided into two sets: the training set (consisting of 60,000 images), and the validation set (with 10,000 images). Initially, we train the EGAN model through two stages: GAN training and encoder training. The GAN training takes 300 epochs with a batch size of 128 images, resulting in trained generator and discriminator models. We selectively choose only the generator from these models. For the encoder training, a 5-fold cross-validation is performed. In each fold, the training process involves 20 epochs, with model validation at the end of each epoch. The goal is to iteratively minimize the encoder error until it can effectively encode an input image, leading to the generator producing an output image with minimal error concerning the input. Then, we validate the classifier previously trained on the dataset. We measure its performance on the original validation set without any data alterations, using the accuracy metric as our measure of success. The result of the classifier validation is an accuracy rate of 98.42%. Finally, we validate the EGAN model creating a set of images generated from the validation set using the EGAN. This process is carried out with the aim of observing how effective our encoder and generator are. We evaluate the EGAN model by the accuracy of the classification model over the validation set previously modified by the EGAN model. The result of the validation of the EGAN model is an accuracy rate of 95.99%.

3.1 Results of the noise add process

We add noise at validation time, applying motion blur with various combinations of its two parameters: the angle θ and the length l. We study all combinations of lengths $l \in \{0.01, 0.02, 0.03, 0.04, 0.05\}$ and angles $\theta \in \{0°, 15°, 30°, 45°\}$. Table 1 shows the accuracy of the classifier after the noisy add process for the images in the validation dataset, for different combinations of the motion blur parameters. The results obtained in the performance of the noise only take into account those images that had been correctly classified in the validation phase of the classifier. We assume an image that has not been correctly classified without having been altered will not be correctly classified if noise is applied to it; therefore, in order to reduce the execution time of the validations of the noise, when the classifier fails its prediction for the original image, it moves on to the next image, without applying noise to it. The results obtained from the noise add process are surprising. Initially, we expected that for any value of disturbance length, increasing the disturbance angle would result in decreased accuracy rates. However, it is observed the opposite situation: increased disturbance angles results in better accuracy rates. We attribute this counterintuitive result to the relatively small size of the images (28x28). With such size, the effects of noise produced by the θ parameter are not as strong as those of the length parameter, being more prevalent as the value of the θ parameter gets closer to zero.

3.2 Results of the denoising process

Now, the noisy images created in the previous step are fed to the EGAN model, and it generates images with similar characteristics to those of the training set, reducing the noise. This image is then fed to the classification model to check how this pre-processing affects predictions. Table 2 shows the accuracy rate of each noise parameters combination in the denoising process with the EGAN model. An analysis of the results shows that they can be divided into two different categories: lower intensity noise ($l \leq 0.02$), where the accuracy rate obtained without the EGAN model is higher than the accuracy rate obtained using the EGAN model to thwart the noise add; and higher intensity noise ($l > 0.02$), where the accuracy rate obtained using the EGAN model to thwart the noise add is higher than the accuracy rate obtained without the EGAN model.

4 Conclusions

Our validations and results analysis underscore the impact of noise add on the studied model. While we observed a clear inverse correlation between the intensity of the disturbance length and the accuracy rates, we also observed a correlation between the intensity of the disturbance angle and the accuracy rates. Comparing results from noise add processes and our proposed denoising strategy using the EGAN model reveals both successes and failures in achieving our initial research goals. In cases with low disturbance length, we encountered less favorable outcomes, resulting in decreased accuracy rates. This can be attributed to the operation of the generating network of the GAN, which deviated from expected results in specific scenarios. Conversely, in situations with more intense blur, the EGAN model was more successful in mitigating the noise. This dual outcome highlights the intricate relationship between model robustness and the intensity of adversarial perturbations. In essence, our study emphasizes the

Length	Epsilon			
	0°	15°	30°	45°
0	98.42%			
0.01	97.31%	97.35%	97.50%	97.70%
0.02	95.19%	95.31%	95.77%	96.47%
0.03	93.59%	93.72%	94.15%	94.98%
0.04	90.54%	91.14%	92.82%	93.85%
0.05	87.04%	87.60%	89.55%	92.69%

Table 1. Accuracy after noisy process.

Length	Epsilon			
	0°	15°	30°	45°
0	95.99%			
0.01	95.29%	95.35%	95.41%	95.46%
0.02	94.84%	94.91%	95.04%	95.18%
0.03	93.89%	94.05%	94.35%	94.74%
0.04	91.89%	92.25%	93.12%	94.13%
0.05	89.17%	89.83%	91.14%	93.05%

Table 2. Accuracy after denoising.

importance of considering varying noise intensities when assessing a model's resilience against noise. A potential improvement to enhance the work is to fine-tune the model for lower-intensity noise, where accuracy rates dropped. A more detailed exploration of the generating network's operations and fine-tuning responses to specific scenarios might enhance overall performance.

Acknowledgments

This work was partially supported by the Autonomous Government of Andalusia (Spain) under grant UMA20-FEDERJA-108, the Ministry of Science and Innovation of Spain under grant PID2022-136764OA-I00, the University of Málaga (Spain) under grants B4-2022, B1-2021_20 and B1-2022_14, and the Fundación Unicaja under grant PUNI-003_2023. The authors thankfully acknowledge the computer resources and assistance provided by the SCBI (Supercomputing and Bioinformatics) center of the University of Malaga, the support of NVIDIA Corporation with the donation of a RTX A6000 GPU with 48Gb, the grant of the Universidad de Málaga, and the Instituto de Investigación Biomédica de Málaga y Plataforma en Nanomedicina-IBIMA Plataforma BIONAND.

References

1. Bank, D., Koenigstein, N., Giryes, R.: Autoencoders (2021)
2. Deng, L.: The mnist database of handwritten digit images for machine learning research [best of the web]. IEEE Signal Processing Magazine **29**(6), 141–142 (2012). https://doi.org/10.1109/MSP.2012.2211477
3. Goodfellow, I.J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative adversarial networks (2014)
4. Pérez-Bravo, J.M., Rodríguez-Rodríguez, J.A., García-González, J., Molina-Cabello, M.A., Thurnhofer-Hemsi, K., López-Rubio, E.: Encoding generative adversarial networks for defense against image classification attacks. In: International Work-Conference on the Interplay Between Natural and Artificial Computation. pp. 163–172. Springer (2022)
5. Rodríguez-Rodríguez, J.A., Molina-Cabello, M.A., Benítez-Rochel, R., López-Rubio, E.: The impact of linear motion blur on the object recognition efficiency of deep convolutional neural networks. In: International Conference on Pattern Recognition. pp. 611–622. Springer (2021)