# A Clean to Noisy Image Generation Scheme Using Generative Adversarial Network

Sadat Hossain and Bumshik Lee



November 18, 2022

# GAN을 이용한 노이즈 영상 생성 기법

사다트 호세인, 이범식*
e-mail: hossain@chosun.kr, bslee@chosun.ac.kr

# A clean to noisy image generation scheme using generative adversarial network

Sadat Hossain, Bumshik Lee*
Department of Information and Communications Engineering,
Chosun University

## Abstract

For training, current deep learning real denoising algorithms need a lot of noisy, clean image pairings. Nevertheless, it is an extremely expensive and time-consuming process to capture a true noisy-clean dataset. To address this issue we looks towards creating realistic noisy visuals. We propose a generative adversarial network (GAN) based noise generation model which utilizes a pre-trained image denoiser to construct the fake and real noisy images into a nearly noise-free solution space. Utilizing this denoiser we have developed a network to generate realistic looking noisy images.

## 1. Introduction

Image denoising is a significant yet difficult issue in low-level vision. It attempts to transform the noisy counterpart into a clean image. Image denoising has made major advancements with the rise of deep learning. The performance of deep convolutional neural network (CNN) proposed in the papers [1-4], which uses a strong learning model to remove noise, is impressive. The aforementioned deep CNN based denoisers depend on a vast dataset of noisy-clean pairs. However, gathering even small datasets requires a lot of time and effort. The procedure of acquiring real-world noisy-clean pairs is to average hundreds of noisy images captured of the same scene. Researchers attempt to synthesize noisy data in order to obtain more image pairings. Because it is easy to produce a pair of noisy and noise-free pictures by adding artificial noise to noise-free images, the bulk of prior learning-based solutions focus on the classic Gaussian denoising issue and give careful consideration to the architecture design of networks, particularly CNNs. As an alternative, few research [5, 6] have gathered well-aligned noisy and clean image pairings, to allow the denoisers to learn in a supervised manner. Although such a method is effective in reducing real-world noise, achieving large-scale pairings is still challenging because of a number of practical issues. When no corresponding clean photos are provided to the target noisy images, no generation-based approach that accurately mimics real-world noise has been presented. In this, paper we utilize generative adversarial network with a pre-trained denoiser in order to generate noisy images which matches the noise distribution of real world noisy images.

## 2. Related Works

GAN was firstly proposed by Ian Goodfellow *et al.* [7], later it has been successfully used for image synthesis and translation [8, 9]. GAN has also been used to image restoration, style transfer, deraining, dehazing and super resolution. The use of data-driven deep learning techniques has substantially enhanced the effectiveness of image enhancement. Although GAN is often used for low-level vision applications, the realistic noise generation problem has received very little attention. The SIDD dataset [5] uses five different smartphone cameras to collect noisy photos in a range of lighting situations. Denoising models with different attention modules are able to eliminate certain complicated real noise from these datasets, however, it is challenging to build a large real-world dataset for a particular purpose. GAN establishes a min-max game between generator and discriminator. Generating convincing samples that deceive the discriminator is the generator's main objective in order to separate generated samples from the real world data. CycleGAN

* 교신저자

[10], an universal GAN based unpaired image translator, was the first and most well-known example. It translates images from two distinct domains. To overcome the problems mention above we propose a method which is an extension of CycleGAN where we have utilized a single generator, discriminator and a pre-trained denoiser to generate realistic noisy images in order to help the denoisers learn better about the complex distribution of real world noise.

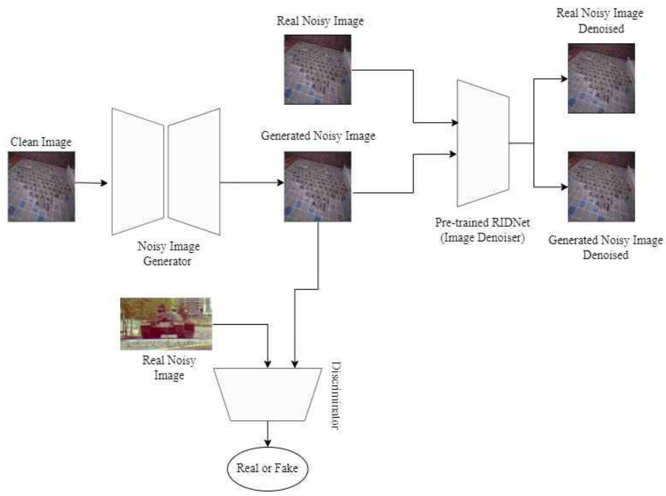


Figure 1. Overall architecture of the proposed noise generator network

## 3. Proposed Method

Figure 1 show the architecture of the GAN network that generates noisy images when a clean image is passed though it. Once the generator generates a noisy image, it is passed to the discriminator as well as the pre-trained RIDnet denoiser [11], the denoiser denoises the generated noisy image and the ground truth noisy image then we calculate the $L_1$ loss between the two, the generator learns from the $L_1$ loss, discriminator's loss and the perceptual quality loss between the generated and the ground truth image which is calculated using VGG-19 network. Following the vanilla GAN we define a discriminator $D$ that is optimized in an oscillating manner with $G$ for the adversarial mix-max problem to be solved. The losses utilized to train the network are expressed in (1).

$$\min\max E_{D^n \sim p_{data(D^n)}}[\log D(D^n)] + E_{D^g \sim p_{data(D^c)}}[\log(1-D^c)] \tag{1}$$

The overall loss of the generator is given as (2), (3) and (4)

$$L_1 = \mid D^n - D^g \mid \tag{2}$$

$$L_{VGG} = VGG-19(D^g - D^n) \tag{3}$$

$$L_g = \sum -\log D(G(D^c)) \tag{4}$$

where $L_{VGG}$ is the perceptual quality loss between the ground truth denoised image and the generated noisy denoised image, $L_g$ is the generators loss based on the output of the discriminator. We add up all these three losses to the generator while backpropagating the weights.

### 3.1 Generator Architecture

To generate realisitc noisy image dataset, we have used the generator architecture of recurrent residual U-Net [12], this design has a number of benefits for activities involving image synthesis. A residual unit first aids deep architecture training. Second, superior feature representation is ensured by feature accumulation using recurrent residual convolutional layers. Third, it enables us to create a more effective U-Net architecture with the same amount of network parameters. The generator consists of with convolutional encoder and decoder part. ReLU activation occurs in both sections of the network once the fundamental convolution processes are completed. To perform down sampling in the encoding unit, 2×2 max-pooling operations are done. The convolution transpose (representing up-convolution, or de-convolution) operations are carried out during the decoding phase to up-sample the feature maps. The Recurrent Convolutional Layers (RCL) operations are carried out in accordance with the discrete time steps specified by the RCNN. The recurrent convolutional operation in this case denoted by $t=2$ consists of a single convolution layer followed by two more recurrent convolution layers. Concatenation has been used in this implementation to combine the feature maps from the encoding unit and decoding unit for the R2U-Net model.


Figure 2. Images generated by the proposed noise generator network

Figure 2 shows the noisy images generated by our

* 교신저자

proposed model, these noisy images are similar to the ground truth images from SIDD dataset. The leftmost images shows the clean image which is used as an input to the trained model, the center images are the ground truth images from SIDD dataset and the leftmost images are the ones generated from our proposed noisy image generator.

## Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) funded by the Korean Government under Grant 2022R1I1A3065473

## References

[1] K. Zhang, W. Zuo, S. Gu, and L. Zhang, “Learning Deep CNN Denoiser Prior for Image Restoration.” arXiv, 2017. doi: 10.48550/ARXIV.1704.03264.

[2] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, “Residual Dense Network for Image Restoration.” arXiv, 2018. doi: 10.48550/ARXIV.1812.10477.

[3] Cai, Y., Wang, Z., Luo, Z., Yin, B., Du, A., Wang, H., Zhang, X., Zhou, X., Zhou, E., & Sun, J. “Learning Delicate Local Representations for Multi-Person Pose Estimation.” arXiv, 2020. doi: 10.48550/ARXIV.2003.04030.

[4] Luo, Z., Wang, Z., Cai, Y., Wang, G., Huang, Y., Wang, L., Zhou, E., Tan, T., & Sun, J. “Efficient Human Pose Estimation by Learning Deeply Aggregated Representations.” arXiv, 2020. doi: 10.48550/ARXIV.2012.07033.

[5] Abdelrahman Abdelhamed, Stephen Lin, and Michael S Brown. A high-quality denoising dataset for smartphone cameras. CVPR, 2018.

[6] Tobias Plotz and Stefan Roth. Benchmarking denoising algorithms with real photographs. CVPR, 2017.

[7] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. “Generative Adversarial Networks.” arXiv, 2014. doi: 10.48550/ARXIV.1406.2661.

[8] X. Gong, S. Chang, Y. Jiang, and Z. Wang, “AutoGAN: Neural Architecture Search for Generative Adversarial Networks.” arXiv, 2019. doi: 10.48550/ARXIV.1908.03835.

[9] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, “Image-to-Image Translation with Conditional Adversarial Networks.” arXiv, 2016. doi: 10.48550/ARXIV.1611.07004.

[10] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks.” arXiv, 2017. doi: 10.48550/ARXIV.1703.10593.

[11] S. Anwar and N. Barnes, “Real Image Denoising with Feature Attention.” arXiv, 2019. doi: 10.48550/ARXIV.1904.07396.

[12] M. Z. Alom, M. Hasan, C. Yakopcic, T. M. Taha, and V. K. Asari, “Recurrent Residual Convolutional Neural Network based on U-Net (R2U-Net) for Medical Image Segmentation.” arXiv, 2018. doi: 10.48550/ARXIV.1802.06955.