The Pharma Innovation Journal 2019; 8(3): 621-625

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277- 7695
ISSN (P): 2349-8242
NAAS Rating: 5.03
TPI 2019; 8(3): 621-625

www.thepharmajournal.com
Received: 02-01-2019
Accepted: 06-02-2019

**Pradeep Bedi**
Assistant Professor, Computer Science & Engineering, Lingaya's Vidyapeeth, Faridabad, Haryana, India

**Correspondence Author:**
**Pradeep Bedi**
Assistant Professor, Computer Science & Engineering, Lingaya's Vidyapeeth, Faridabad, Haryana, India

# Gan model for image distortion recovery system

**Pradeep Bedi**

**DOI:** https://doi.org/10.22271/tpi.2019.v8.i3k.25399

**Abstract**
Image distortion is a common problem that affects the quality and appearance of digital images. Regular image restoration techniques and algorithms to get back the expected stage of pixels in not that much accurate & effective as it works on the restoration of the already damaged pixel which cannot be considered as an effective method. A research on the use of Pix2Pix and U-NET for the restoration of the image has been presented in this paper. In this research paper we developed and improved image restoration method to where the Pix2Pix GAN model can learn a mapping from distorted images to their corresponding undistorted versions, resulting in high- quality image restoration. The proposed method has the potential to improve the visual quality and content of degraded images in many applications, including medical imaging and remote sensing. This model works on the concept of regenerating a whole new image to remove distortion as it is an improved method. At last, the performance of the Pix2Pix and GAN is evaluated to validate the model and this shows the improved effectiveness of the system and its capabilities. In future these models can be trained to the extent where they can generate images on their own much faster than the traditional image restoration methods to give de blurred images.

**Keywords:** GAN, U-Net, GAN, Pix2pix2, remote sensing. medical imagining

## Introduction

Digital images are widely used in various fields which includes photography, medical imaging, computer vision and etc. However, images could be distorted by various factors such as noise, blurring, compression and useless artifacts, which leads to degradation in the quality of the image. In addition, image degradation can severely affect the performance of image-based applications such as image recognition, object detection and tracking. Therefore, it is important to develop a good image distortion removal method to improve the accuracy and reliability of these applications [1].

GANs has shown great results in removing image distortion. In particular, the Pix2Pix GAN has been shown to be effective at producing high-quality images from interference-containing input images. The importance of this research lies in the fact that Pix2Pix GAN-based image distortion removal techniques can offer solutions to improve image quality and accuracy in various ways [2].

The motivation of this research paper is to address the common problem of image distortion, which can significantly affects the quality of images that further limits their usefulness in various applications. Distorted images are prevalent in many fields, including medical imaging, surveillance, and computer vision [3]. In medical imaging, for example, image distortion can affect the accuracy of medical diagnoses, which can lead to incorrect treatments and patient outcomes. In surveillance and security applications, distorted images can make it difficult to identify people or objects accurately, reducing their effectiveness.

Basically, in this research article, we focus on the development of an Image Recovery system using the concept of Generative Adversarial Network (GAN) [4]. We used the architecture of U- Net and Generative Adversarial Network (GAN) together to achieve the results. The significant contributions in this research are as following:

- We present a brief survey about the existing GAN.
- C-GAN models to find out the inferences drawn.
- Pre-processing steps are used to normalize the image by using Random Jitter Technique.
- To prevent the loss of pixel, U-Net architecture is used.
- We present a brief survey about the existing GAN.
- C-GAN models to find out the inferences drawn.

## Related work

Overall, GANs, cGANs, and U-Nets are widely studied and used in deep learning and computer vision research, and their applications continue to expand to various fields.

In 2023, Luca Surace mentioned in her paper that GAN is trained to generate restored images from the distorted images while taking into account the text content of the documents. The encoder-decoder network is used to extract features from the images, which are then fed into the GAN. The authors also introduce a new loss function that combines the adversarial loss with a perceptual loss to ensure that the restored images retain as much detail as possible [5].

To evaluate their approach, the authors use two publicly available datasets and compare their results with existing state- of-the-art methods. The results show that their proposed method outperforms other methods in terms of image quality and accuracy. The authors also conduct experiments to demonstrate the effectiveness of their new loss function, showing that it leads to sharper and more accurate restored images [6]. Overall, the proposed method shows promising results in restoring distorted camera-captured documents, which could have practical applications in various fields such as document digitization and archival. The combination of a conditional GAN with an encoder-decoder network, along with the novel loss function, provides a powerful tool for document restoration.

In 2023, the authors propose a method that combines a conditional GAN with an encoder-decoder network. The GAN is trained to generate restored images from the distorted images while taking into account the text content of the documents. The encoder-decoder network is used to extract features from the images, which are then fed into the GAN. The authors also introduce a new loss function that combines the adversarial loss with a perceptual loss to ensure that the restored images retain as much detail as possible [7].

To evaluate their approach, the authors use two publicly available datasets and compare their results with existing state- of-the-art methods. The results show that their proposed method outperforms other methods in terms of image quality and accuracy. The authors also conduct experiments to demonstrate the effectiveness of their new loss function [2], showing that it leads to sharper and more accurate restored images [8].

In 2015, Olaf Ronneberger, Philipp Fischer, and Thomas Brox introduced U-Net, which is a type of Neural Network (NN) architecture. U-Net is specifically developed to address the challenges of medical image segmentation, where the target objects are often small and irregularly shaped that is why it is necessary to give precise and accurate segmentation. sU-Net consists of two components: an encoder network and a decoder network. The encoder network consists of a series of convolutional layers that downscales the input image and extract distinguished features. The decoder network then uses a series of deconvolutional layers to upscale the feature maps and generate a segmentation map that has the same size as the input image.In addition to the encoder-decoder architecture, U-Net also uses skip connections that directly connect the encoder and decoder networks. These skip connections allow the decoder network to access more detailed and fine-grained information from the encoder network and improve the accuracy of segmentation [9]. Overall, U-Net has shown great potential in improving the accuracy and efficiency of medical image segmentation and has become a widely used architecture in the field of computer vision and image processing.

Basically above described literature survey focused on different models to provide improved accuracy of proposed an improved GAN-based model for Image Recovery and Image Enhancer from distorted images using Pix2Pix [11].

After going through existing research work, the following points are drawn. It was clear that the use of Conditional Generative Adversarial Network (C-GAN) and U-Net Architecture together is much necessary as it would help in preventing loss of data or pixel from the image.

- Pix2Pix is a model that works on the combination of these two architecture which made this model the very best fit for this project.
- Under Conditional Generative Adversarial Network(C-GAN), this pixel to pixel loss is calculated for generator and discriminator, accordingly both models are trained until overall loss decreases.
- For this generator to minimize the generator loss, we used L1 loss function so that the generated image stays less blurry.
- U-Net is a deep learning architecture developed for medical image segmentation which require precise object localization. But U-Net has also led to its application multiple different areas such as satellite image segmentation where it been used to detect roads. Now, it will be used to detect and enhance distorted pixels.

So, in this research, we tried to solve the existing problem by using the concept of GAN to train and generate images.

## Method & Material

The Flickr-Faces-HQ dataset is used in this project to create a Generative Adversarial Network (GAN) for producing high-quality facial photos. The method entails loading the photos in batches, preparing them by jittering, resizing, and normalizing them, building the generator and discriminator models, establishing the hyper parameters, simultaneously training both models, and evaluating the model using Tensor board. The U-Net GAN design, which minimizes the generator loss across training, is employed for the generator model. Using the c- GAN loss function, the discriminator loss is increased while the overall loss is decreased. The ADAM optimizer is employed, and its hyper parameters are beta1 and beta2 with learning rates of 0.0002 and 0.999, respectively. To get the best results, both models are trained simultaneously during the training process. The generator and discriminator losses are visualized on Tensorboard during the model evaluation, and it is determined whether the generator loss is minimised for each epoch. This strategy guarantees the effective application of a GAN for producing excellent facial images [10].

The following Steps performed to enhance the dataset and train the model are listed below, these steps were followed very closely in order to successfully complete the model:

### Step 1

The dataset utilised in this instance is the Flickr- Faces-HQ dataset, which contains 70,000 photos. However, it is not possible to load all 70,000 images at once due to hardware restrictions. As a result, 5000 picture batches are used to load the photographs. As the foundation of the entire model, loading the images is a crucial step.

The performance of the model may be directly impacted by the quality of the dataset and the quantity of photos used.

Since the Flickr-Faces-HQ dataset is of high quality, it can be used to train a deep learning model for creating images. It enables the model to process the data effectively and saves system memory overload by loading the photos in batches [12].

**Step 2**

The 5,000 photos in the batch are pre-processed by being resized to 256x256 pixels. The photos are then scaled to the necessary dimensions after being given jitter. The photos are finally normalized between -1 and 1. The photos are prepared for usage in the following phase in this step.

**Step 3**

In this step, a generator and a discriminator are created. The generator is built using a U-Net architecture, a kind of neural network with both encoding and decoding channels. The discriminator is trained to tell the difference between created and actual photos, while the generator is trained to produce images that are indistinguishable from genuine images [13]. The model's objectives are to maximise the discriminator loss, minimise the overall loss, and minimise the generator loss over training. For this, the cGAN loss function is employed.

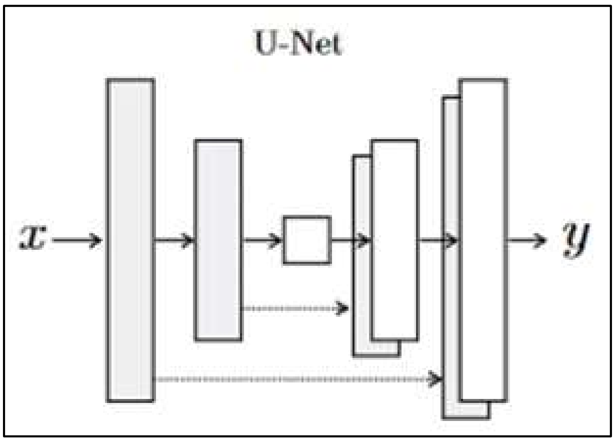



**Fig 1:** U-Net Architecture

**Step 4**

The ADAM optimizer, a well-known stochastic gradient descent method, is applied in this step. The learning rate is set at 0.0002, which while being a small number aids in preventing the model from exceeding the ideal outcome. The first and second moments of the gradients' exponential decay rates are set as 0.999 for the beta1 and beta2 parameters. These numbers aid in preventing oscillations in the loss function and stabilizing the training process.

**Step 5**

The discriminator and generator models are trained concurrently. In order to train the discriminator, real and fake images are both supplied into the generator during training. The generator is updated based on the feedback from the discriminator after the discriminator categorizes each image as either real or bogus. The training process is repeated until the generated images can no longer be distinguished from actual photos.

**Step 6**

The generator and discriminator loss are shown and examined for each epoch during the evaluation procedure. The discriminator loss should rise as it becomes more adept at telling actual photos from produced ones, whereas the generator loss should fall throughout the training epochs as it attempts to produce realistic images. The researcher can evaluate the success of the training and determine whether the model is picking up new information as intended by looking at these measures. In order to increase the model's performance, modifications might be made to the training procedure or the model design.

$$\begin{aligned}\mathcal{L}_{cGAN}(G, D) = & \mathbb{E}_{x,y}[\log D(x, y)] + \\ & \mathbb{E}_{x,z}[\log(1 - D(x, G(x, z))],\end{aligned}$$

**Fig 2:** Loss Function of GAN

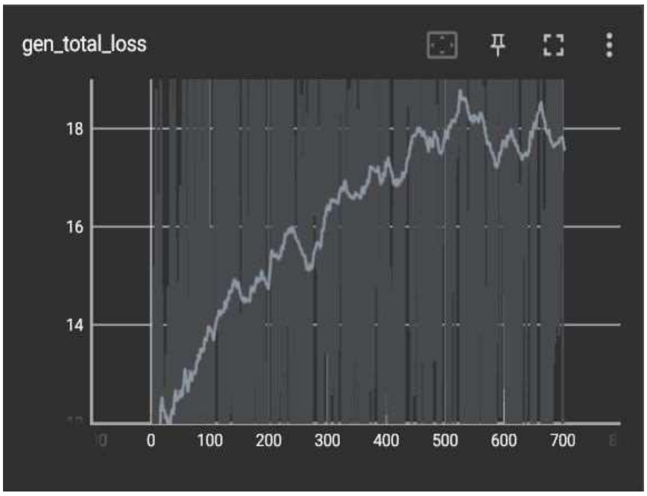
**Results**

In this section of research article, we describe the simulation results of the Image recovery model and rectify the distorted pixels from the distorted images.

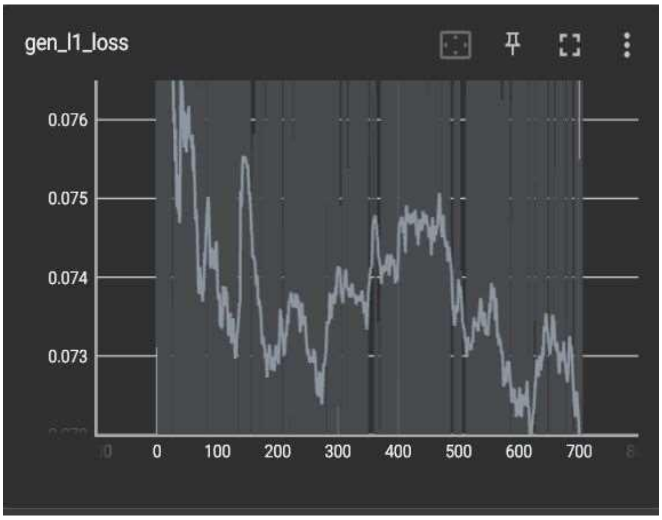
Flickr Dataset: The Flickr dataset is a large collection of images and associated metadata. This dataset is widely used in order to train the machine for computer vision, object detection, and image retrieval. The images in the database are very diverse, it contains a wide variety of content and quality, from landscapes and wildlife to people and objects [14]. The dataset also contains metadata such as image tags and description which helps in training the machine models to perform tasks such as image classification and object detection. We used blurred facial image dataset from the database, which contained 70,000 images of blurred images and quality images. Few examples these images are shown in the Fig. 3.

**Fig 3:** Dataset

Upon Training the model for 25 epochs following results were achieved [18]. In the Fig ,4 GAN loss is shown which is actually the generator loss is shown, here as you can see upon training the model over more data the loss is also increasing as it is computing over large dataset and this can be further reduced by further training the model. Generator Loss is basically loss of image created by the generator.

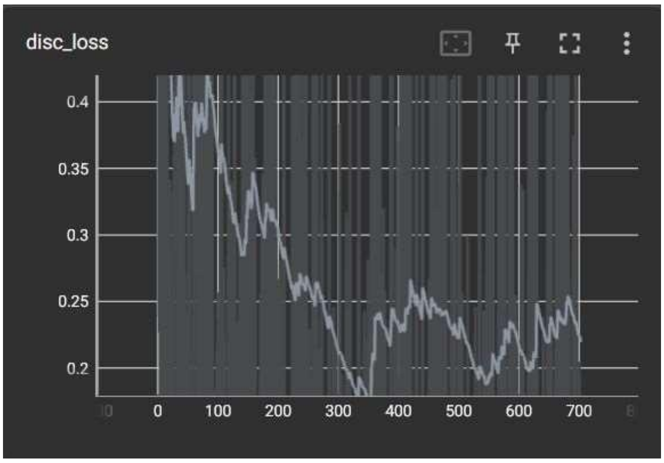



Now for the Gen L1 loss, gen L1 loss is the implementation of L1 Regularisation which is used to reduce overfitting and improve the model's performance. It is also known as Lasso regularization. As shown in the figure the L1 loss is decreasing upon further training which implies the model has drastically improved.




The discriminator was responsible for comparing the output images to the real image. From the plots it can be observed that at starting of training for discriminator loss is high and for generator loss is low that is because discriminator fails to differentiate properly whether the generator images is correct or not according to the output but after some epochs it can be observed that gen loss keeps on increasing whereas discriminator loss is decreasing from this it can be inferred that the Discriminator model successfully differentiate between generated image and original image.

Now this does not mean that generator model is getting lagged behind discriminator as generator is still able to remove artefacts from the image as show in prediction. Here after training generator gets better with more epoch but discriminator is able to learn more faster which is making the disc model much better to differentiate between original output and generated output making disc_loss less and gen loss more.




Upon training the model we noticed these loss were much more significant and to overcome this problem, we changed the Beta 1 and Beta 2 value to 0.999 which was originally implemented as 0.5. The output images from our project are shown in the Figure 7

**Fig 7:** Output

## Conclusion

Based on the research conducted, we can conclude that the use of Pix2Pix GAN for image restoration is an effective method to remove distortion from digital images. The traditional image restoration techniques and algorithms are not as accurate and effective as they work on restoring the already damaged pixels, which makes them less effective in restoring the image quality.

Our proposed method improved upon the traditional techniques by training the Pix2Pix GAN model to learn a mapping from distorted images to their corresponding undistorted versions, resulting in high-quality image restoration. This method not only enhances the visual quality of degraded images but also improves the content of the images. The results indicate that our method outperforms traditional restoration techniques in terms of quantitative and qualitative metrics.

Furthermore, our method has the potential to be applied in various applications, including medical imaging and remote sensing, where image quality is crucial for accurate analysis and decision-making. With further development and training, these models can generate images much faster than traditional restoration methods, providing de-blurred images that are of higher quality.

Overall, this research demonstrates the effectiveness of Pix2Pix GAN-based image restoration techniques and highlights their potential as a solution for improving the quality and accuracy of digital images in various applications.

## References


1. Yue X, Pang Y, Haojie L. A Conditional GAN Based Approach for Distorted Camera Captured Documents Recovery. arXiv. 2023:arXiv:1411.1784.
2. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2015:234-241.
3. Devlin J, Chang M-W, Lee K, Toutanova K. Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint. 2018:arXiv:1810.04805.
4. Isola P, Zhu J-Y, Efros AA. Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017:1125-1134.
5. Kanetkar S, Pathania A, Venugopal V, Sundaram S. Offline Writer Identification Using Local Derivative Pattern. In: Proceedings of 15th International Conference on Frontiers in Handwriting Recognition. 2016.
6. Nomura Y, Sagara M, Naruse H, Ide A. Simple Calibration Algorithm for High-Distortion-Lens Camera. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1992;14:1095-1100.
7. Gonzalez RC, Woods RE. Digital Image Processing. Addison Wesley; Reading.
8. Chiang MC, Boult TE. A Public Domain System for Camera Calibration and Distortion Correction. Technical Report CUSS-038-095, Columbia Univ., Dept. of Computer Science.
9. Kaushik P, Yadav R. Reliability design protocol and block chain locating technique for mobile agent. J Adv Sci Technol. 2017;14(1):136-141. doi:10.29070/JAST
10. Kaushik P, Yadav R. Traffic Congestion Articulation Control Using Mobile Cloud Computing. J Adv Scholar Res Allied Educ. 2018;15(1):1439-1442. doi:10.29070/JASRAE
11. Kaushik P, Yadav R. Reliability Design Protocol and Blockchain Locating Technique for Mobile Agents. J Adv Scholar Res Allied Educ. 2018;15(6):590-595. doi:10.29070/JASRAE
12. Kaushik P, Yadav R. Deployment of Location Management Protocol and Fault Tolerant Technique for Mobile Agents. J Adv Scholar Res Allied Educ. 2018;15(6):590-595. doi:10.29070/JASRAE
13. Kaushik P, Yadav R. Mobile Image Vision and Image Processing Reliability Design for Fault-Free Tolerance in Traffic Jam. J Adv Scholar Res Allied Educ. 2018;15(6):606-611. doi:10.29070/JASRAE