FACE COMPLETION AND SUPER RESOLUTION OF OCCLUDED IMAGES USING GAN

A PROJECT REPORT

Submitted by

AKASHRAM J 2019115011

THARANYAA R 2019115113

VIJAY S 2019115120

submitted to the Faculty of

INFORMATION AND COMMUNICATION ENGINEERING

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY

in

INFORMATION TECHNOLOGY

DEPARTMENT OF INFORMATION SCIENCE AND TECHNOLOGY

COLLEGE OF ENGINEERING, GUINDY

ANNA UNIVERSITY

CHENNAI 600 025

DECEMBER 2022

ANNA UNIVERSITY

CHENNAI - 600 025

BONA FIDE CERTIFICATE

Certified that this project report titled **FACE COMPLETION AND SUPER RESOLUTION OF OCCLUDED IMAGES USING GAN** is the bonafide work of **AKASHRAM J, THARANYAA R, VIJAY S** who carried out project work under my supervision. Certified further that to the best of my knowledge and belief, the work reported herein does not form part of any other thesis or dissertation on the basis of which a degree or an award was conferred on an earlier occasion on this or any other candidate.

PLACE:CHENNAI
DATE:

Ms.S.KANIMOZHI
TEACHING FELLOW
PROJECT GUIDE
DEPARTMENT OF IST, CEG
ANNA UNIVERSITY
CHENNAI 600025

COUNTERSIGNED

Dr.S SRIDHAR
PROFESSOR AND HEAD OF THE DEPARTMENT
DEPARTMENT OF INFORMATION SCIENCE AND TECHNOLOGY
COLLEGE OF ENGINEERING, GUINDY
ANNA UNIVERSITY
CHENNAI 600025

ABSTRACT

Low-resolution and occlusion often present in face images affect the accuracy of face recognition. e.g. under the scenario of video surveillance. While most of the existing face image recovery approaches can handle only one type of variation per model, in this work, we propose a Deep Convolutional Generative Adversarial Network (DCGAN) for performing face completion and super-resolution of image. The generator model aims to recover a high-resolution face image without occlusion, given an input low-resolution face image with occlusion.The Discriminator model is trained on real data for n batches and see if it can correctly predict them as real or fake . This model can be trained end-to-end using two submodels. Finally, evaluation for proposed model done over the available datasets CelebA. Our model is useful for improving face identification performance when there are low-resolution and occlusion in face images.

ABSTRACT

Low-resolution and occlusion often present in face images affect the accuracy of face recognition. e.g. under the scenario of video surveillance. While most of the existing face image recovery approaches can handle only one type of variation per model, in this work, we propose a Deep Convolutional Generative Adversarial Network (DCGAN) for performing face completion and super-resolution of image. The generator model aims to recover a high-resolution face image without occlusion, given an input low-resolution face image with occlusion.The Discriminator model is trained on real data for n batches and see if it can correctly predict them as real or fake . This model can be trained end-to-end using two submodels. Finally, evaluation for proposed model done over the available datasets CelebA. Our model is useful for improving face identification performance when there are low-resolution and occlusion in face images.

ACKNOWLEDGEMENT

It is our privilege to express our sincere thanks to our project guide Ms. S. Kanimozhi, Teaching Fellow, Department of Information Science and Technology, College of Engineering, Guindy, Anna University, Chennai for her keen interest, inspiring guidance, constant encouragement and support with our work during all the stages, to bring this report into fruition.

We deeply express our sincere thanks to Dr. S SRIDHAR, Professor and Head of the Department and the project coordinator DR. S. SWAMYNATHAN, Professor, Department of Information Science and Technology, Anna University, Chennai for their constant support.

We would like to express our sincere thanks to the project committee members, Dr. K. VANI, Professor, Dr. S. BAMA, Assistant Professor, Dr. P. GEETHA , Assistant Professor, Dr. M. Deivamani, Teaching Fellow, Department of Information Science and Technology, Anna University, Chennai for giving their valuable suggestions, encouragement and constant motivation throughout the duration of our project.

J AKASHRAM
R THARANYAA
S VIJAY

TABLE OF CONTENTS

	ABSTRACT	iv
	ABSTRACT	iv
	LIST OF FIGURES	viii
	LIST OF SYMBOLS AND ABBREVIATIONS	ix
1	**INTRODUCTION**	**1**
	1.1 BACKGROUND	1
	1.2 OBJECTIVE	2
	1.3 PROBLEM STATEMENT	2
	1.4 SOLUTION OVERVIEW	2
	1.5 ORGANIZATON OF THE REPORT	2
2	**LITERATURE SURVEY**	**4**
	2.1 DE-GAN: DOMAIN EMBEDDED GAN FOR HIGH QUALITY FACE IMAGE INPAINTING	4
	2.2 SIMULTANEOUS FACE COMPLETION AND FRONTALIZATION VIA MASK GUIDED TWO-STAGE GAN	4
	2.3 FACE COMPLETION WITH HYBRID DILATED CONVOLUTION	5
	2.4 EDGECONNECT: STRUCTURE GUIDED IMAGE INPAINTING USING EDGE PREDICTION.	6
	2.5 RECURRENT GENERATIVE ADVERSARIAL NETWORK FOR FACE COMPLETION	6
	2.6 SUMMARY	7
3	**SYSTEM DESIGN**	**8**
	3.1 SYSTEM ARCHITECTURE	8
	3.2 GENERATOR NETWORK	9
	3.3 DISCRIMINATOR NETWORK	10
4	**IMPLEMENTATION OF YOUR WORK**	**12**
	4.1 ALGORITHM	12
	4.1.1 Mode Collapse	13
	4.1.2 Minibatch Discrimination	13
	4.1.3 One-sided label smoothing	13

4.1.4 CelebA dataset 14

5 RESULTS AND PERFORMANCE ANALYSIS 15

5.1 RESULTS OF 2000 EPOCHS 15

5.2 TESTING ON REAL TIME FACE IMAGE 16

5.3 GRAPH ANALYSIS 18

5.3.1 Generator Loss 18

5.3.2 Discriminator Loss 18

5.3.3 Gan Loss 19

5.4 EVALUATION OF GAN 19

5.4.1 Qualitative GAN Generator Evaluation 20

5.4.2 Quantitative GAN Generator Evaluation 20

5.4.3 Fréchet Inception Distance and Inception Score 20

5.4.4 Structural Similarity Index Measure 21

6 CONCLUSION AND FUTURE WORK 22

REFERENCES 23

LIST OF FIGURES

3.1	System Architecture.	9
3.2	Generator Network	10
3.3	Discriminator Network.	10
5.1	Occluded Face	15
5.2	Ground Face	15
5.3	Generated Face	15
5.4	Person-1 Occluded Face	16
5.5	Person-1 Ground Face	16
5.6	Person-1 Generated Face	16
5.7	Person-2 Occluded Face	17
5.8	Person-2 Ground Face	17
5.9	Person-2 Generated Face	17
5.10	2000 Epochs Generator Loss	18
5.11	2000 Epochs Discriminator Loss	19
5.12	2000 Epochs GAN Loss	20

LIST OF SYMBOLS AND ABBREVIATIONS

CNN	Convolutional Neural Networks
DCGAN	Deep Convolutional Generative Adversarial Network
DE-GAN	Domain Embedded Generative Adversarial Network
GAN	Generative Adversarial Network
FID	Frechet Inception Distance
SSIM	Structural Similarity Index Measure

CHAPTER 1

INTRODUCTION

Image completion method aim to repair the defects of digital images with plausibly synthesized content to make images look more natural. Image completion is one of the technologies used to fill in missing parts of images. The generated contents can either be as accurate as the original, or simply fit well within the context such that the completed image appears to be visually realistic. It is important to find ways to repair the defects. For example, a face image captured by a camera may be occluded by other objects, resulting in loss of face information. This undoubtedly increases the difficulty of face recognition. While face completion can repair the missing face information and generate a complete face to improve the accuracy of face recognition. To solve this problem, we propose a Deep Convolutional Generative Adversarial Network (DCGAN) which is a deep learning-based model that automatically completes the missing region so that the completed face not only looks natural and realistic but also has consistency with the rest of the image.

1.1 BACKGROUND

DCGAN is an approach for generative modeling using deep learning methods such as CNN (Convolutional Neural Network). Generative modeling is an unsupervised learning approach that involves automatically discovering and learning patterns in input data such that the model can be used to generate new examples from the original dataset. There are two components of GANs: Generator: It is trained to generate new dataset, for example in computer vision it generates new images from existing real-world images. Discriminator: It compares those images with some real-world examples and classify real and fake images.

1.2 OBJECTIVE

To provide a technical solution to repair the defects of digital images. Our aim is to reconstruct missing or damaged regions of an incomplete face image given the context information using a deep generative completion model to synthesize the missing contents from random noise.

1.3 PROBLEM STATEMENT

General inpainting methods focus mainly on the resolution of generated images without considering the particular structure of human faces and generally produce inharmonious facial parts. Existing face completion methods incorporate only one type of facial feature for face completion, and their results are still undesirable.

1.4 SOLUTION OVERVIEW

The Deep Convolutional Generative Adversarial Network(DCGAN) is proposed, which is a deep learning-based model that automatically completes the missing region so that the completed face not only looks natural and realistic but also has consistency with the rest of the image.

1.5 ORGANIZATON OF THE REPORT

This paper is organized into 6 chapters, describing each part of the project with detailed illustrations and system design diagrams. The chapters are as follows: Chapter 2 explains the literature survey details of the existing systems with their methodologies, advantages, disadvantages, etc. Chapter 3 consists of the system design of the project with the overall architecture and the modules of the architecture and the description of the modules used in the project. Chapter 4 consists of Implementation part of the project. Chapter 5 consists of Final Result. Chapter 6 consists of conclusion and future scope.

CHAPTER 2

LITERATURE SURVEY

This chapter explains the literature survey done in the existing system and the analyses of the problem statement and issues in the existing system.

2.1 DE-GAN: DOMAIN EMBEDDED GAN FOR HIGH QUALITY FACE IMAGE INPAINTING

Zhang et al.[12] Domain knowledge of face shapes and structures plays an important role in face inpainting. However, general inpainting methods focus mainly on the resolution of generated images without considering the particular structure of human faces and generally produce inharmonious facial parts. Existing face-inpainting methods incorporate only one type of facial feature for face completion, and their results are still undesirable. To improve face inpainting quality, they proposed a Domain Embedded Generative Adversarial Network (DE-GAN) for face inpainting. DE-GAN embeds three types of face domain knowledge (i.e., face mask, face part, and landmark image) via a hierarchical variational auto-encoder (HVAE) into a latent variable space to guide face completion. Two adversarial discriminators, a global discriminator and a patch discriminator, are used to judge whether the generated distribution is close to the real distribution or not. Experimental results demonstrated that method achieves a better performance than the existing state-of-the-art face-inpainting methods. Moreover, DE-GAN provides great facilitation for face recognition face alignment. However, DE-GAN has some drawbacks such as limited application scenarios, and, similar to other inpainting methods, DE-GAN cannot complete face inpainting when there are two or more faces in an image.

2.2 SIMULTANEOUS FACE COMPLETION AND FRONTALIZATION VIA MASK GUIDED TWO-STAGE GAN

Duan et al. [13] Pose variation and occlusion are two key factors that affect the accuracy of face recognition. Most of the previous work alleviate the impacts of pose and occlusion by performing the tasks of face frontalization and face completion, respectively. Specially, generative adversarial networks (GANs) based methods have made a great progress on both of these two tasks. However, the two tasks are rarely paid attention simultaneously. Hence, the synthesis and recognition from the profile but occluded facial image is still an understudied and challenging problem in two aspects. 1) Occlusion mask, as a kind of noise, can be some very important prior information in the corrupted image. In particular, the occlusion mask is often used to fit this noise and help face restoration of the occluded region. However, a prior work, such as BoostGAN, failed to utilize the mask guided noise prior information. 2) The two tasks, de-occlusion and frontalization, are collaborative, so the identity discriminative information is easily lost if the two tasks are not organically unified in the training phase. In order to overcome these challenges, they proposed a novel mask guided two-stage generative adversarial network TSGAN. The objective of this work is de-occlusion and frontalization.The aim of face deocclusion and frontalization is to improve the recognition accuracy. This paper addresses an occluded but profile face recognition task. Failed to implement Cross-age face recognition.

2.3 FACE COMPLETION WITH HYBRID DILATED CONVOLUTION

Fang et al.[6] In this paper, they focused on face image inpainting tasks, aiming at reconstructing missing or damaged regions of an incomplete face image given the context information. We specially design the U-Net architecture to tackle the problem. The proposed U-Net based method combines Hybrid Dilated Convolution (HDC) and spectral normalization to fill in missing regions of any shape with sharp structures and fine-detailed textures. They performed both qualitative and quantitative evaluation on two challenging face datasets.The proposed method can generate realistic and semantically plausible images. Failed to produce high-resolution inpainting image.

2.4 EDGECONNECT: STRUCTURE GUIDED IMAGE INPAINTING USING EDGE PREDICTION.

Nazeri et al. [8] Many deep learning techniques have been applied to the image inpainting problem, the task of filling incomplete regions of an image. However,these models struggle to recover or preserve image structure especially when significant portions of the image are missing. In this paper they proposed a two-stage model that separates the inpainting problem into structure prediction and image completion. Their model first predicts the image structure of the missing region in the form of edge maps. Predicted edge maps are passed to the second stage to guide the inpainting process.Their method achieves state of art results on standard benchmarks, and is able to deal with images with multiple, irregularly shaped missing regions. Edge generating model sometimes struggles to accurately depict the edges in highly textured areas, or when a large portion of the image is missing especially for higher resolution images.

2.5 RECURRENT GENERATIVE ADVERSARIAL NETWORK FOR FACE COMPLETION

Most recently-proposed face completion algorithms use high-level features extracted from convolutional neural networks (CNNs) to recover semantic texture content. Although the completed face is natural-looking, the synthesized content still lacks lots of high-frequency details, since the high-level features cannot supply sufficient spatial information for details recovery. To tackle this limitation, in this paper, they proposed a Recurrent Generative Adversarial Network (RGAN) for face completion. Unlike previous algorithms, RGAN can take full advantage of multi-level features, and further provide advanced representations from multiple perspectives, which can well restore spatial information and details in face completion. Specifically, our RGAN model is composed of a CompletionNet and a DisctiminationNet, where the CompletionNet consists of two deep CNNs and a recurrent neural network (RNN). The first deep CNN is presented to learn the internal regulations of a masked image and represent it with multi-level features. The RNN model then exploits the relationships among the multi-level features and transfers these features in another domain, which can be used to complete the face image. Benefiting from bidirectional short links, another CNN is used to fuse multi-level features transferred from RNN and reconstruct the face image in different scales. Meanwhile, two context discrimination networks in the DisctiminationNet are adopted to ensure the completed image consistency globally and locally. Experimental results on benchmark datasets demonstrate qualitatively and quantitatively that this model performs better than the state-of-the-art face completion models, and simultaneously generates realistic image content and high-frequency details.

2.6 SUMMARY

From the literary survey,we can conclude that there have been various attempts to solve face completion. Most work has been focused on converting low resolution image to high resolution. There is a need for image completion. While face completion can repair the missing face information and generate a complete face to improve the accuracy of face recognition.

CHAPTER 3

SYSTEM DESIGN

This chapter consists of the system design of the project with the technical architecture and the modules of the architecture and the description of the modules used in the project.

3.1 SYSTEM ARCHITECTURE

A generative adversarial network (GAN) has two parts: The generator learns to generate plausible data. The generated instances become negative training examples for the discriminator. The discriminator learns to distinguish the generator's fake data from real data. The discriminator penalizes the generator for producing implausible results. When training begins, the generator produces obviously fake data, and the discriminator quickly learns to tell that it's fake. As training progresses, the generator gets closer to producing output that can fool the discriminator. Finally, if generator training goes well, the discriminator gets worse at telling the difference between real and fake. It starts to classify fake data as real, and its accuracy decreases. Both the generator and the discriminator are neural networks. The generator output is connected directly to the discriminator input. Through back propagation, the discriminator's classification provides a signal that the generator uses to update its weights. The technical architecture diagram of the proposed system is shown in Figure 3.1.

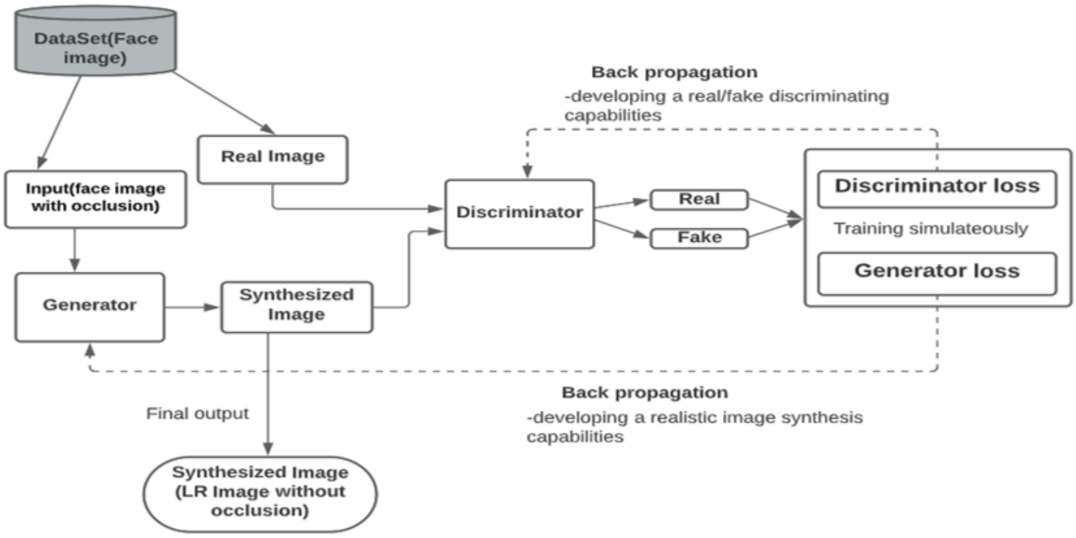

Figure 3.1: System Architecture.

3.2 GENERATOR NETWORK

Generator takes 64 * 64 occluded image as input and then this image goes through convolution layers. The purpose of convolution layers is to learn meaningful features from the image. With help of convolution layers, we get a feature abstract from the image. Batch Normalization occurs at each layer in CNN. The purpose of batch normalization is to convert all weights of each into the same scale to stabilize the Neural Network. Activation function takes the normal weight value and converts this value between 0 to 1. Transpose convolution is to up sample the feature abstract to get the desired output.The Generator network is shown in Figure 3.2

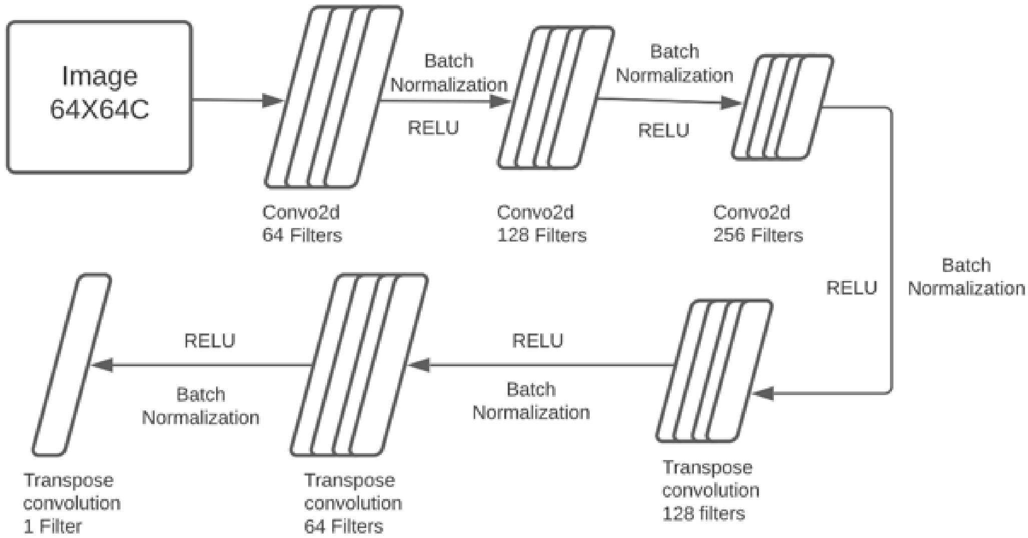

Figure 3.2: Generator Network

3.3 DISCRIMINATOR NETWORK

Discriminator takes a real or synthesized image as input and convolution layers learn meaningful features from the image. With help of convolution layers we get a feature abstract from the image. At each convolution layer, the input image gets boiled down to reduced size of the feature map and eventually ends up with one output neuron. Output value infers the probability of an image that is real. Figure 3.3

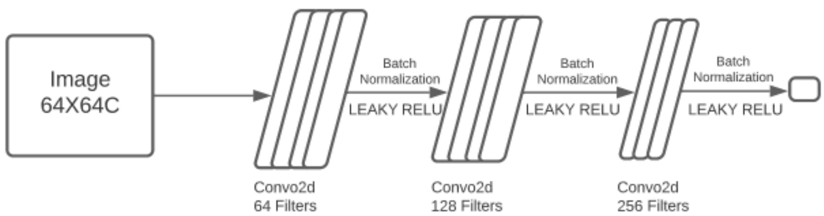

Figure 3.3: Discriminator Network.

Equation1: $$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

Here, p(x) – real image

D(x) – Discriminator prediction on real data

P(z) – fake image

D(G(x)) - Discriminator prediction on fake data

Above equation infers the cost function of the GAN model.

$\mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})]$	This term infers the probability of an image that is real. Discriminator wants D(x) to be a large number because D(x) represents high confidence that the real sample is actually real.
$\mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]$	This term infers the probability of an image that is fake. Discriminator wants D(G(x)) to be as small as possible because this term represents high confidence that the fake sample is actually fake. Since the generator wants to fool the discriminator, Generator wants this value to be large as possible or discriminator to think the fake sample is real with confidence.

CHAPTER 4

IMPLEMENTATION OF YOUR WORK

This chapter describes the algorithm for module that has been implemented for training a deep convolution generative adversarial network.

4.1 ALGORITHM

Input: Occluded face image	
Output: Clear face image	
Variables **Gan_label:** GAN label, **Gen_label:** Generator label, **Disc_label:** Discriminator Label, **Gen_pred:** Generator predict, **Occ_img[]:** Array occluded image, **Disc_inp []:** Array of input images at discriminator, **Real_img[]:** Array of real image **Tgan:** Test gan **Tdisc**: Test discriminator	
1	Initialize the Gan_label = 1, Disc_label = 0, batch = 64, nepoch = 500, spilt = 125, Disc_inp[]
2	For epoch in range (nepoch)
3	{
4	For n in range (spilt)
5	{

6	#Generator makes a prediction.
	Gen_pred = generator.predict. (Occ_img[batch*n : batch*(n+1)])

7	#Minibatch isolation and label smoothing is done here: If epoch % 2 == 0{
8	Disc_inp = Gen_pred
9	Smooth the discriminator label to identify fake image
10	}
11	Else
12	{
13	Disc_inp = real_img[n * batch : (n + 1) * batch]
14	Smooth the discriminator label to identify real image
15	}
16	Gen_label = real_img[n * batch : (n + 1) * batch]
17	Calculate loss a) Discriminator Loss $= \frac{\text{real image}}{\textit{discriminator label}}$ b) Generator Loss $= \frac{\textit{Occlusion image}}{\textit{Generator label}}$ c) Gan Loss $= \frac{\textit{Occlusion image}}{\textit{Gan Label}}$
18	}

```
If epochs = epochs – 1
    #Save the model
#Test the model
(Tgan, Tdisc) = test (Occ_img [8000:9000], real_img [8000:9000])
}
```

4.1.1 Mode Collapse

Usually GAN to produce a wide variety of outputs. However, if a generator produces an especially plausible output, the generator may learn to produce only that output. In fact, the generator is always trying to find the one output that seems most plausible to the discriminator. If the generator starts producing the same output (or a small set of outputs) over and over again, the discriminator's best strategy is to learn to always reject that output. But if the next generation of discriminator gets stuck in a local minimum and doesn't find the best strategy, then it's too easy for the next generator iteration to find the most plausible output for the current discriminator. Each iteration of generator over-optimizes for a particular discriminator, and the discriminator never manages to learn its way out of the trap. As a result, the generators rotate through a small set of output types. This form of GAN failure is called mode collapse.

4.1.2 Minibatch Discrimination

It is a discriminative technique for generative adversarial networks where we discriminate between whole minibatches of samples rather than between individual samples. This is intended to avoid mode collapse of the generator.

4.1.3 One-sided label smoothing

Earlier, label/target values for a classifier were 0 or 1; 0 for fake images and 1 for real images. Because of this, GANs were prone to adversarial , examples which are inputs to a neural network that result in an incorrect output from the network. Label smoothing is an approach to provide smoothed labels to the discriminator network. This means we can have decimal values such as 0.9 (true), 0.8 (true), 0.1 (fake), or 0.2 (fake), instead of labelling every example as either 1 (true) or 0 (fake). We smooth the target values (label values) of the real images as well as of the fake images. Label smoothing can reduce the risk of adversarial examples in GANs.

4.1.4 CelebA dataset

The dataset contains 10000 real images and 10000 occluded images and train, test, validation split is 8000, 1000, 1000 respectively for both real and occluded dataset. Feed the 64 images to‘ the generator model and discriminator model at time. For 8000 images ,split the images like 64 * 125.

CHAPTER 5

RESULTS AND PERFORMANCE ANALYSIS

This chapter illustrates the experimental results of the proposed system and the evaluation part.

5.1 RESULTS OF 2000 EPOCHS

CelebA dataset is spilt into training, test, validation set. Below image is output of 10 images that randomly pick from validation set.

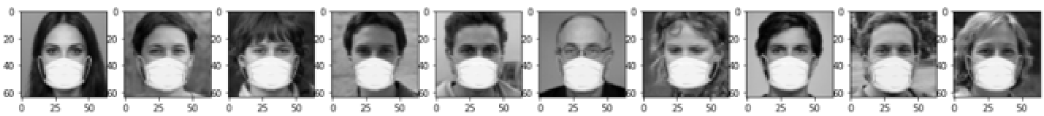

Figure 5.1: Occluded Face

Figure 5.2: Ground Face

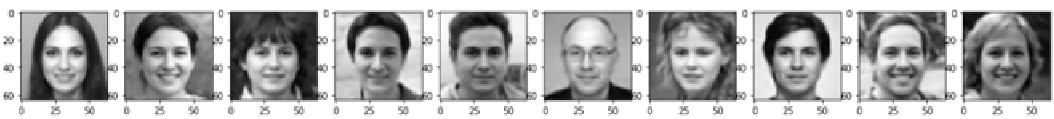

Figure 5.3: Generated Face

5.2 TESTING ON REAL TIME FACE IMAGE

Below images are randomly picked real face images with occlusion and generated their respected face image.

Figure 5.4: Person-1 Occluded Face

Figure 5.5: Person-1 Ground Face

Figure 5.6: Person-1 Generated Face

Figure 5.7: Person-2 Occluded Face

Figure 5.8: Person-2 Ground Face

Figure 5.9: Person-2 Generated Face

5.3 GRAPH ANALYSIS

5.3.1 Generator Loss

While the generator is trained, it samples random noise and produces an output from that noise. The output then goes through the discriminator and gets classified as either “Real” or “Fake” based on the ability of the discriminator to tell one from the other. The generator loss is then calculated from the discriminator’s classification – it gets rewarded if it successfully fools the discriminator, and gets penalized otherwise. Here, the mean squared error loss function as generator loss function. It is the evaluation measure to check the performance of the generator model. It measures the amount of divergence of predicted image with the actual image. So lesser the loss value, more the perfectness of model. For a perfect model, loss value = 0. For instance, as accuracy is the count of correct predictions i.e. the predicted image that matches the actual image, Loss value is the measure of uncertainty of our predicted image based on how it varies from the actual image. Generator loss is plotted as graph and x axis as Number of epochs, y axis as Generator loss. From graph, Generator loss curve tend to reach zero so, this infers that generator model is more perfect enough to generate fake.The generator loss graph is shown below in the figure 5.10.

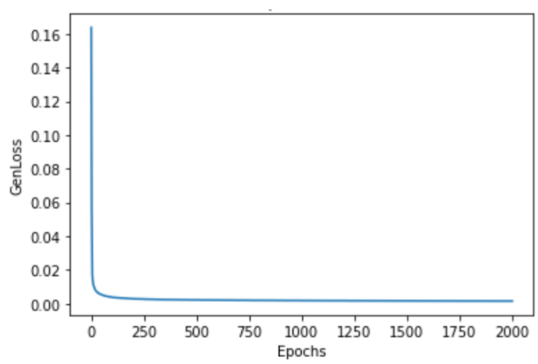

Figure 5.10: 2000 Epochs Generator Loss

5.3.2 Discriminator Loss

While the discriminator is trained, it classifies both the real data and the fake data from the generator. It penalizes itself for misclassifying a real instance as fake, or a fake instance (created by the generator) as real. For a binary classification like classifying real and fake images, the typical loss function is the binary cross-entropy. So, discriminator uses a binary cross entropy loss function. In figure 5.11 , test discriminator loss value is constant and discriminator loss value lies in the range of (0, 16). Now Discriminator able to classify the real and fake images in more accurate because of test discriminator loss is constant.

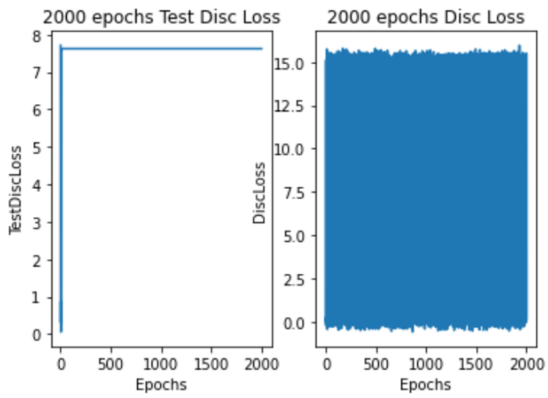

Figure 5.11: 2000 Epochs Discriminator Loss

5.3.3 Gan Loss

GANs try to replicate a probability distribution. They should therefore use loss functions that reflect the distance between the distribution of the data generated by the GAN and the distribution of the real data. To capture the difference between two distributions in GAN loss functions, minimax loss is the loss function commonly used in GAN. From figure 5.12, Gan loss value is constant so, this infers that discriminator model and generator model are more stable.

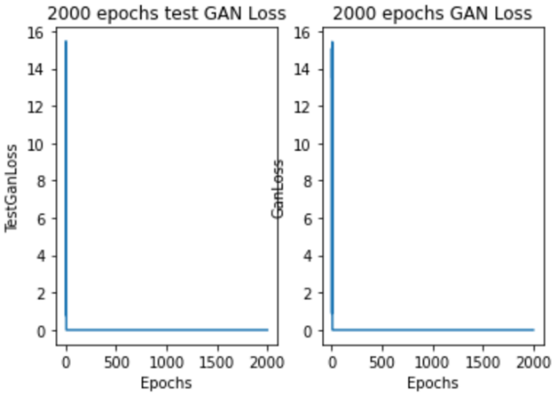

Figure 5.12: 2000 Epochs GAN Loss

5.4 EVALUATION OF GAN

Evaluation of GAN can be done in two ways: 1) Qualitative GAN Generator Evaluation and 2) Quantitative GAN Generator Evaluation

5.4.1 Qualitative GAN Generator Evaluation

Qualitative measures are those measures that are not numerical and often involve human subjective evaluation or evaluation via comparison

5.4.2 Quantitative GAN Generator Evaluation

Quantitative GAN generator evaluation refers to the calculation of specific numerical scores used to summarize the quality of generated images. Three popular methods for evaluating performance of GAN are 1) Inception score (IS) and 2) Fréchet inception distance (FID) 3) Structural similarity index measure (SSIM)

5.4.3 Fréchet Inception Distance and Inception Score

Usually, GAN performance measured in inception score (IS) and Fréchet inception distance (FID) score. The Fréchet inception distance (FID) is a metric used to assess the quality of images created by a generative model, like a generative adversarial network. Unlike the earlier inception score (IS), which evaluates only the distribution of generated images, the FID compares the distribution of generated images with the distribution of a set of real images. The Fréchet score for this face competition using DCGAN is 0.690.

5.4.4 Structural Similarity Index Measure

SSIM is a metric of comparison to check the similarity between the two images. It measures the perceptual difference between the two images and SSIM score range between 0 and 1. SSIM score value between two images nearer to 1 infers that two images are more similar or value tend to 0 infers that two images are dissimilar. For DCGAN, SSIM score is 0.93. Pluralistic Face Image Completion Based on Style GAN is an IEEE paper published in 2021, their SSIM score is 0.925

CHAPTER 6

CONCLUSION AND FUTURE WORK

In this paper, we proposed a Deep Convolutional Generative Adversarial Network (DCGAN) for face image completion. DCGAN improves the face completion performance and retains the identity information of the original face. Experimental results demonstrate that our method achieves a better performance than the existing methods. Future work consists of super resolution of the image.

REFERENCES

[1] Ali. Borji. Pros and cons of gan evaluation measures. *Computer Vision and Image Understanding*, 179:41–65, 2019.

[2] D. K. Vishwakarma. Comparative analysis of deep convolutional generative adversarial network and conditional generative adversarial network using hand written digits. *In 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 1072–1075, 2020.

[3] Jason Brownlee. *A Gentle Introduction to Generative Adversarial Networks (GANs).* June 17, 2019.

[4] Kurita T. Ide, H. Improvement of learning for cnn with relu activation by sparse regularization. *International Joint Conference on Neural Networks (IJCNN)*, pages 2684–2691, 2017.

[5] Li G. Deng L. Liu L. Wu D. Xie Y. Shi L. Wu, S. L1-norm batch normalization for efficient training of deep neural networks. *IEEE transactions on neural networks and learning systems*, pages 2043–2051, 2018.

[6] Li Y. Tu X. Tan T. Wang X Fang, Y. Face completion with hybrid dilated convolution. *Signal Processing: Image Communication*, 80, 2020.

[7] Li Z. Du B. Zhang M. Liu J. Xu, J. Reluplex made more practical: Leaky relu. *In 2020 IEEE Symposium on Computers and communications (ISCC)*, pages 1–7, 2020.

[8] Ng E. Joseph T. Qureshi F. Ebrahimi M Nazeri, K. Edgeconnect: Structure guided image inpainting using edge prediction. *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 30:0–0, 2019.

[9] Slama R. Tabia H. Ouni T. Abid M Kammoun, A. Generative adversarial networks for face generation: A survey. *ACM Computing Surveys*, 55:1–37, 2022.

[10] Syed Abdul Gaffar Shakhadri. *Deep Convolutional Generative Adversarial Network (DCGAN) for Beginners*. Published On July 7, 2021.

[11] White T. Dumoulin V. Arulkumaran K. Sengupta B. Bharath A. Creswell, A. Generative adversarial networks: An overview. *IEEE signal processing magazine*, 35:53–65, 2018.

[12] Zhang et al. De-gan: Domain embedded gan for high quality face image inpainting. *Pattern Recognition*, 124, 2022.

[13] Zhang L. Gao X Duan, Q. Simultaneous face completion and frontalization via mask guided two-stage gan. *IEEE Transactions on Circuits and Systems for Video Technology*, 32:3761 – 3773, 2021.

[14] Zhang Y. Zhang L. Wang, N. Dynamic selection network for image inpainting. *IEEE Transactions on Image Processing*, 30:1784–1798, 2021.

[15] Zijun. Zhang. Improved adam optimizer for deep neural networks. *In 2018 IEEE/ACM 26th International Symposium on Quality of Service (IWQoS)*, pages 1–2, 2018.