Project Synopsis

on

Blind Image Restoration and Data Augmentation

Submitted as a part of course curriculum for

Bachelor of Technology

in

Computer Science

Submitted by

Nandini Tyagi (2000290120099)

Piyush Gupta (2000290120105)

Avishi Tayal (2000290120051)

Under the Supervision of

Dr. Harsh Khatter

Assistant Professor in Computer Science Department

KIET Group of Institutions, Ghaziabad

Department of Computer Science

Dr. A.P.J. Abdul Kalam Technical University

2022-2023

TABLE OF CONTENTS

DECLARATION

We hereby declare that this submission is our work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgement has been made in the text.

Signature of Students
Name: Nandini Tyagi, Piyush Gupta, Avishi Tayal
Roll No.: 2000290120099, 2000290120105, 2000290120051
Date: 11/11/2022

CERTIFICATE

This is to certify that Project Report entitled "Blind Image Restoration and Data Augmentation" which is submitted by **Nandini Tyagi, Piyush Gupta, Avishi Tayal** in partial fulfilment of the requirement for the award of degree B. Tech. in Department of Computer Science of Dr A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

Date:

Dr. Harsh Khatter
Assistant Professor
Department of Computer Science

ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the synopsis of the B.Tech Mini Project undertaken during B.Tech. Third Year. We owe a special debt of gratitude to Dr. Harsh Khatter, Assistant Professor Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his constant support and guidance throughout the course of our work.

We also take the opportunity to acknowledge the contribution of Dr. Ajay Kumar Shrivastava, Head of the Department of Computer Science, KIET Group of Institutions, Delhi- NCR, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

Last but not the least, we acknowledge our friends for their contribution to the completion of the project.

Signature:
Date :
Name :
Roll No:

ABSTRACT

Some of the most amazing discoveries in this decade have been made thanks to technology and technological advancements. The world seems to get better every day thanks to the rapidly evolving technology, systems to support them, and back-end processing power.

Image restoration is currently one of the most difficult problems. Image restoration is the process of estimating the clean, original image from a corrupted or noisy image. Motion blur, noise, and camera focus issues are just a few examples of corruption. The main objective of the restoration process, which enhances the appearance of the image, is to return it to the way it appeared when it was first synthesized.

Images of various cancerous cells can be restored and studied using this technique. Astronomical aspects of satellite images can also be restored. Face image restoration can be used to identify criminal activity, among other real-world applications. To accomplish the mentioned goals, we are using the Generative Adversarial Network (GAN) models in this situation. Our use of the GAN prior embedded network enables us to produce a smooth and finely tuned image from which we can further deduce some conclusions. The biggest challenge facing the computer vision industry today is obtaining an image of a high quality without background noise. Because in this scenario, every pixel of the image is crucial to the training of a complete model. Therefore, image restoration is the most difficult task to complete, which will also help to solve many real-world and industrial problems. For example, in the real world, because of the degradation of some cctv images, it is difficult to easily record criminal activity. In a similar vein, a model in the medical field allows for the quick and accurate visualization of cancerous cells. The training of numerous economic projects, including text to image conversion and clothing translation, will undoubtedly be helpful to industry.

CHAPTER 1: Introduction

1.1 INTRODUCTION:

Technology and Technological developments in this decade have led to some of the most awe-inspiring discoveries. With rapidly changing technology and systems to support them and provide back-end processing power, the world seems to be becoming a better place to live day by day.

One of the biggest challenge from the times is Image restoration. Image restoration is **the operation of taking a corrupt/noisy image and estimating the clean, original image**. Corruption may come in many forms such as motion blur, noise and camera misfocus. The restoration process improves the image's appearance, and the main goal is to restore it to how it looked when it was first synthesized originally.

Image restoration plays an important role in the real life- medical field- different cancerous cells images can be restored and studied by the process, astronomical aspects obtained satellite image can be restored, face image restoration- detects the criminal activities etc.

Here, we are using the Generative Adversarial Network(GAN) models to achieve the mention goals. GAN prior embedded network helps us to generate a fine-tuned and smooth image which can be further used to draw some insights.

1.2 PROBLEM STATEMENT:

To achieve a high quality image without background noise, proper colorization and inpainted image is biggest challenge in today's computer vision industry. Because the training of a complete model in such case depends on each pixel of the image. Hence image restoration is the biggest task to achieve which further help to solve various real world and industrial problems such as in real world – various cctv image are degraded due to which criminal activities can't be record easily, similarly in medical field the cancerous cell can visualize easily and effectively by the model. It will certainly help in industries to train various economical projects like- clothes translation, text to image conversion and many more.

1.3 OBJECTIVE:

Objective is to build a website using Streamlit where the process of blind image restoration can be carried out and different poses for the same images can be generated that provides a better view to the user. A low quality real world image will be uploaded to user and in return it will provide a highly resolved image with different painting.

The model is efficient enough to perform the image colorization, image inpainting and synthesize the image efficiently.

1.4 SCOPE:

Now a day's industries are crazy about high resolution human face datasets which can be utilized by them to train their models for different purposes like virtual try-on, clothes translation and many more but limit is a dataset here GPEN along with STYLEGAN 3 can be used to generate the different poses of a image. The generated data is realistic and this produced dataset can be used in many ways like as local brands are unable to afford the superstars for advertisement in such case Stylegan can be used to generate the synthesized image with style mixing. And actors can't even do any case on them for the deed. Many other industrial applications also sustain for the same.

CHAPTER 2: Literature Review

Ian J. Goodfellow et al.[1] have implemented Generative Adversarial Network. Author employed this model with the help of two functions Generator and Discriminator. Here both functions have different meanings. According to the research Generator is responsible for capturing the data whereas Discriminator is responsible for estimating the probability that the feature comes from training data or captured image. Author has trained model on various famous datasets that are MNIST, CIFAR-10, TFD etc. Author had not used any Markov chains or unrolled approximate inference files during the training of model. The study completes with a note that it can be improved by taking better devising methods for Generator G and Discriminator D. This is how the whole study takes place.

Tao Yang et al.[2] have worked upon "Blind Face Restoration" by using GAN prior embedded network (GPEN). The author has researched that GAN models are used to improve the quality of degraded image but they make the image over smooth which is not a better result. That's why concept of GPEN came in which GAN model is mapped with U shaped Deep neural network to achieve better results in terms of quality and quantity as well. In the research, U-shaped Deep neural network is trained by using U-net model to make the results better. This model basically works on converting low quality face image to higher quality face image. According to their research they employed that GAN prior embedded network is able to generate photo realistic results. The model overall work on following a U-shaped encoder decoder architecture for being capable to generating high quality face images. Stylegan convolutions and U-net model is used for making of GPEN model for real world blind face restoration.

Feida Zhu et al.[3] aims to reconstruct high quality images from low quality images. It was founded that real world is suffering from degradation of low quality face images during data acquisition and Internet transmission. So, Author has employed SGPN Shape and Generated prior integrated network to solve this problem. Approach of the whole methodology starts with the shape restoration module to balance the face geometry. After that by following D3DFR author regress 3DMM coefficients with Resnet50. Then, shape S and

coloured texture C formed which is used to make 2D image plane to obtain 3D image. Author has utilized stylegan2 model as the generative prior to improve the model capability as well as results. The main focus of this study is to restore low quality face images in the wild. The model overlooks only the facial part and little about background part.

Yang Zhao et al.[4] have observed various limitations of previously done GAN models. The main issue they find out is that by going through a GAN model image lost it's so many originals features like shape of any face feature or texture or colour dissimilarity. Author's team has figure out a methodology to improve and overcome the limitations of GAN model by injecting a noise layer in the image. They employed an adaptive conditional noise layer to the model which makes model capable to consider non-deterministic nature of face image restoration problem. The two benchmarks of the study were Blind face restoration (BFR) and Super resolution (SR). Author has included the method of image feature quantization and linear gated feature fusion to make best results. The main aim of this whole study is to make an authenticated blind face restoration model which may leads to many beneficial applications. At last they have compared all the related models like Bicubic, GFP-GAN, GPEN with their model and there model gave extremely great results with respect to all.

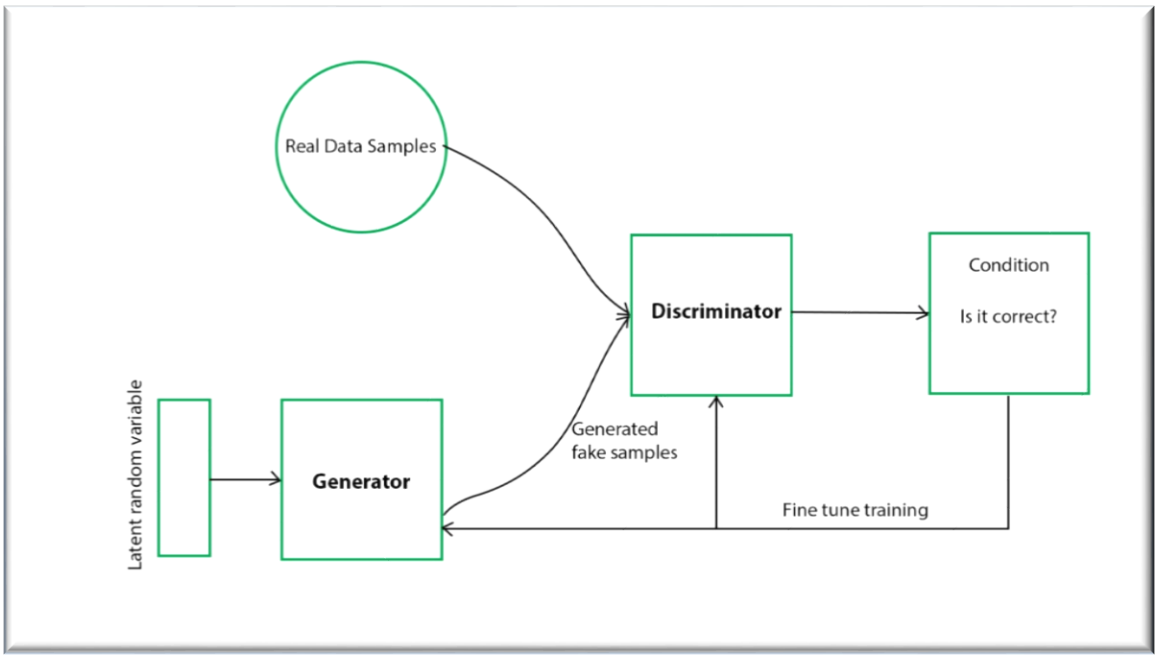
Tero Karras et al.[5] has proposed a different generator architecture for Generative Adversarial Network that is style based generator. Author has introduced new dataset of human faces as well like Flickr-Faces-HQ, FFHQ that offers a great quality content. Study says that most of the work is to focus on improving discriminator by using multiple types of discriminator. Author has considered stochastic variation of the image to get the best outcomes. The study mainly focuses on making great content in the field of Stylegan network by adding some latent space in between the layers to improve. Firstly, author has compared various methods of generator function like traditional, style based, noise addition according to their separability and path length to analyse their dependencies for enhancing the performance in a better way. The conclusion of the study is that Traditional GAN generator architectures are subordinate in every regard to style-based network architecture.

Tero Karras et al.[6] have find out a generative adversarial network model that is better suited for video and animation category. Author has matched the FID of stayle-gan2 model with their resulting model, then they know that their model performs better. In the proposed methodology author has converted the stylegan2 model generator to be fully equivariant to translation and rotation. During the whole research author has implemented alias free generator model that contains implicit assumptions about the behaviour of training dataset. In StyleGAN3-R, the emergent positional encoding patterns appear to be somewhat more well-defined. Author believe that the existence of a coordinate system that allows precise localization on the surfaces of objects will prove useful in various applications, including advanced image and video editing.

CHAPTER 3: Proposed Methodology

3.1 Flowchart:

Our complete study and deployment is based upon the embedded GAN models. GAN models basically helps in the generation of images based on the concept of generator and discriminator. Generator basically generates the images on the basis of the random vector from the Gaussian distribution. And at the same time discriminator takes this fake generated image and real images from the training data as input and tests that generated fake images. If these generated images fool the discriminator predicting that its true, then only the image is allowed to pass through discriminator.



3.2 Algorithm Proposed:

Dataset: Here, for the study and training purpose we had made a custom dataset that comprises of approximately 100 images of actors. Each image in the dataset is resized and converted to particular dimensions of 256 X 256.
GPEN- Severely degraded face images is the biggest challenging problem faced by the industries for the training purpose of their different models such as virtual try on-clothes, and moreover in the real world the degraded images don't allow us to interpret the incidents or individuals. Here, the main aim of using the GPEN model is to restore the blind portion of images. Steps involved in building the model:

1) Firstly, here we will train a GAN network based on the high quality images dataset with the same settings as that of Stylegan. The GAN model generator and discriminator both are trained on the basis of the high quality images.
2) Then the GPEN model is embedded with this trained GAN model to further perform the fine-tuning of images. Corresponding to each low quality image and a finely grained high quality is generated. The degraded images are generated /synthesized based on the training of GAN network provided by the high quality images and degradation model:

$$Id = ((I \otimes k) \downarrow s + n\sigma)JPEGq$$

3) Here, we have used different gradients of GPEN model apart from the face restoration that includes face colorization, face inpainting, selfie restoration and conditional image synthesis. Here we have used the pretrained weights of each gradient to train the embedded GAN both with different properties.
4) Further the trained model is deployed using streamlit to provide a user interface.

CHAPTER 4: Technology Used

Technologies used for the implementation of the model are:

1) Deep Learning
2) Open cv
3) Generative Adversarial Network
4) GAN Prior Embedded Network
5) Streamlit

Platform needed:

1) Google Colaboratory
2) Vs-code

Requirements:

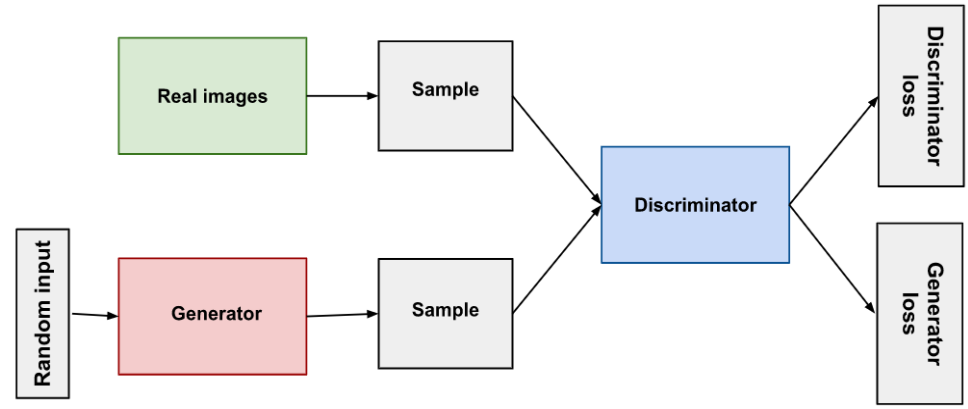
1) GPU for training the model

CHAPTER 5: Conclusion

Currently, the industry is wound up with with high-resolution human face datasets, which can be used to train models for various purposes such as virtual fitting, clothing translation, etc. There are limits to the datasets available. Can be used to create different poses of the image. The data generated is realistic and this generated dataset can be used in many ways as local brands cannot afford advertising superstars. In this case, Stylegan can be used to generate composite images using style mixing. And the actor can't even charge her for the act. Many other industrial applications also contribute.

CHAPTER 6: Diagrams

GAN architecture



Face Restoration

References:

[1] I. Goodfellow *et al.*, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020, doi: 10.1145/3422622.

[2] T. Yang, P. Ren, X. Xie, and L. Zhang, "GaN Prior Embedded Network for Blind Face Restoration in the Wild," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 672–681, 2021, doi: 10.1109/CVPR46437.2021.00073.

[3] F. Zhu *et al.*, "Blind Face Restoration via Integrating Face Shape and Generative Priors," pp. 7652–7661, 2022, doi: 10.1109/cvpr52688.2022.00751.

[4] Y. Zhao *et al.*, "Rethinking Deep Face Restoration," pp. 7642–7651, 2022, doi: 10.1109/cvpr52688.2022.00750.

[5] T. Karras, S. Laine, and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 12, pp. 4217–4228, 2021, doi: 10.1109/TPAMI.2020.2970919.

[6] T. Karras *et al.*, "StyleGAN3," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, no. NeurIPS, pp. 438–448, 2021, [Online]. Available: https://blog.faradars.org/generative-adversarial-networks/.