**International Islamic University Chittagong**

**Department of Computer Science and Engineering**



**A Thesis Proposal on**

## **Fake Face Generator: Generating Fake Human Faces Using GAN**

**Supervised by**

Md. Khaliluzzaman

Assistant Professor

**Submitted by**

Name: Md. Shahariar Younus Ashik

ID: C171017

Name: Md. Al Asad Nur Riyad

ID: C171021

**Approval of the Supervisor**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**[Md. Khaliluzzaman]**

[AssistantProfessor]

**1. Introduction**

Machine Learning has become one of the most influential technologies of the 21st century. It has many applications extending from precise diagnosis of skin diseases, detection faults in credit lending systems to recommendations on streaming channels and gaming, this technology is omnipresent. However, there are darker, malignant sides too. One of the most concerning challenges is fooling the existing algorithms of neural networks by adding a small amount of noise into the original data. And after many iteration cycles and feedback, the same model can now produce counterfeit data or wrong output. This happens because, after adding noise, the model gets higher confidence in doing wrong predictions compared to then when it is predicted correctly.

We would like to do our research in the area of Deep Learning for Computer Vision which is a subfield of Machine Learning. There are many Deep Learning methods. Generative Adversarial Network(GAN) is one of them. GAN is an approach to generative modeling using Deep Learning methods, such as Convolutional Neural Networks. GAN will actively be used in our proposed work to train a generative model for generating fake human faces.

The main aim of our thesis is to generate fake human faces for tackling misclassification, using the Generative Adversarial Network (GAN).

Challenges in generating fake human faces are:

* Vanishing Gradients
* Mode Collapse
* Failure to Converge
* Stabilization

**2. Literature Review**

We are greatly influenced by some of the previous works related to GAN. Some of them are listed below:

* Zhang et al. [1] in 2020 proposed a method to improve the quality of generated images that is high-quality face image generation. For this, they replaced MLP with convolutional neural networks (CNN) and removed pooling layers. They replaced the pooling layer with a convolution layer. They tested their method on LFW and CelebA datasets.
* Wang et al. [2] in 2016 proposed a method that simplified the overall generative process, led to more realistic high resolution images, highly stable and robust learning procedure, modified the underlying 3D structure of an input image and rendered a completely new image. They tested their method on the NYUv2 dataset.
* Ghatas et al. [3] in 2020 proposed a pipeline to build a complex modular pipeline using pre-trained models to generate the kin image. The modular nature of the pipeline changes the way the GAN problems are approached and introduces a systematic way that can be easily generalized to more generative problems and it widens the door of possible solutions. They tested their method on Family101 and FIW dataset.
* Hamdi et al. [4] in 2019 proposed a new regularizer for GAN that uses K-nearest neighbor (K-NN) selective feature matching to target set Y in high-level feature space during the adversarial training of any GAN, they named it as K-GAN. They also presented a cascade framework for GAN to push its output away from the base domain X to target Y. They called this setup the Imaginative Adversarial Network (IAN). They tested their method on CelebA and ImageNet dataset.
* Tolosana et al. [5] in 2020 proposed a method to detect fake news, fake image, fake media, face manipulation. They specially covered four types of manipulations: i) entire face synthesis, ii) identity swap, iii) attribute manipulation, and iv) expression swap. They used a bunch of public databases for their work.
* Karras et al. [6] in 2019 proposed a method to re-design the generator architecture in a way that exposes novel ways to control the image synthesis process. They proposed two new automated metrics: i) perceptual path length and ii) linear separability — for quantifying these aspects of the generator. They used Flickr-Faces-HQ and FFHQ datasets for their work.

**3. Objectives**

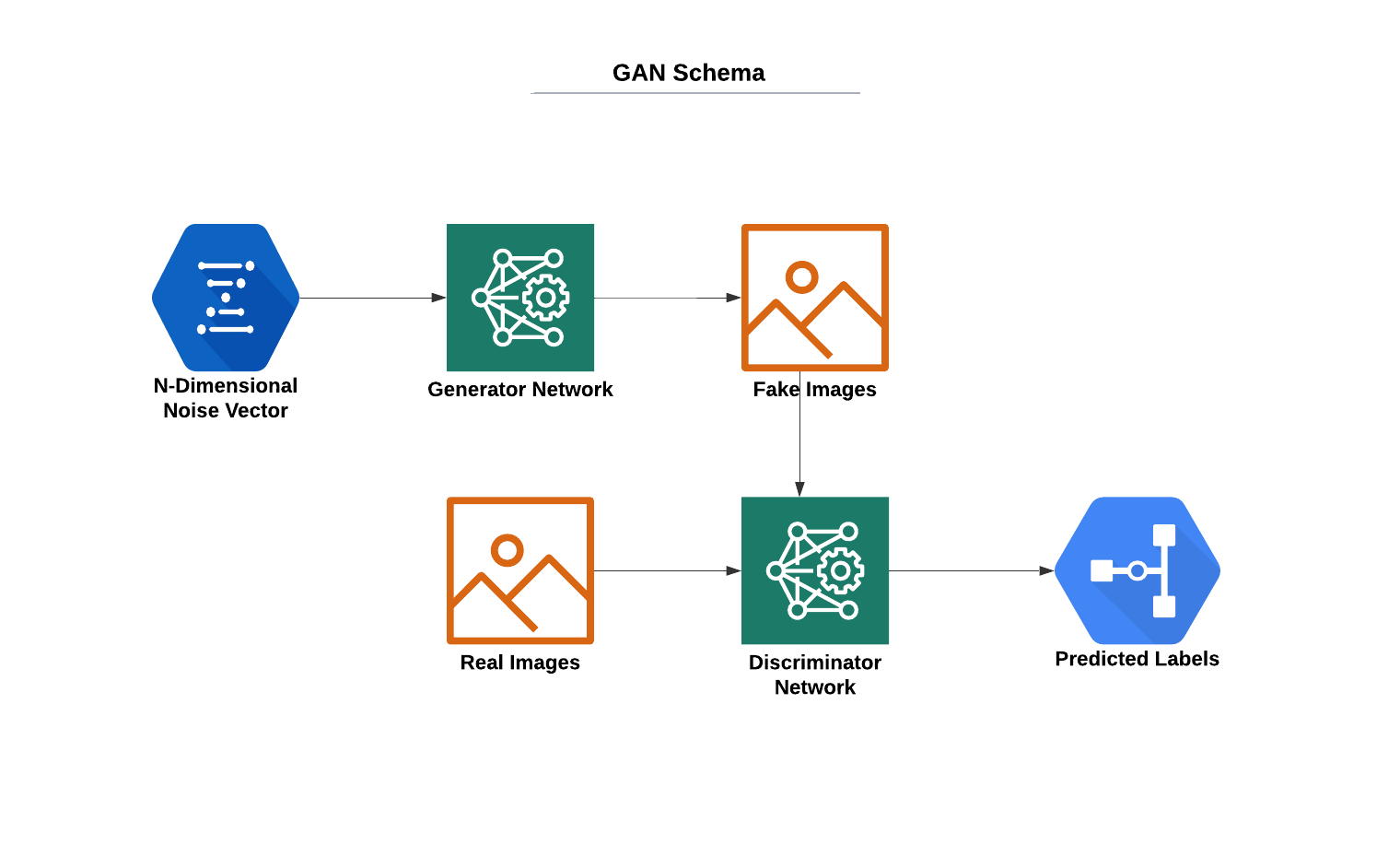
The aim of our work is to generate comparatively better real-life fake human faces, rather than removing all the noise and maximising the stabilisation which was the main challenge of the previous related works.

The main features of our proposed system will be:

* To generate comparatively better real-life fake human faces
* To make a pipeline based on GANs that reduce maximum noise.
* To improve runtime optimization

**4. Methodology**

In our work, we will propose a method for generating fake human faces based on Generative Adversarial Network which can be applied in face classification, face recognition, and age prediction. We will train our network with the CelebA dataset. Working with the CelebA dataset will be a great challenge as it is a large-scale face attributes dataset with more than 200K celebrity images, each with 40 attributes.



For this, we have to train the generation model G and the discriminant model D very well. At first, the generator G will take in random numbers z and return an image G(z). This generated image G(z) will be fed into the discriminator model D alongside the images x taken from the actual dataset. Then the discriminator will take in both real and fake images and will return probability, a number between 0 and 1, where 1 will be representing as real and 0 as fake. When the discriminator will not be able to distinguish between the real data x and the generated data G(z), the generator G will be considered to be optimal. The goal of the discriminator D will be to discriminate the two parts, which will make D as large as possible, while keeping D(G(z)) as small as possible, and the difference between the two parts as large as possible. While the goal of the generation model G will be to make the performance of D(G(z)) consistent with the performance of D(x), so that the discriminator could not be able to identify the difference between the real and the fake data. The performance of the both generator and discriminator module will improve during the process until the performance of the D(G(z)) would be consistent with the D(x). Finally, we will be able to generate some fake photographs of human faces.

**5. Work Plan and Timeline:**

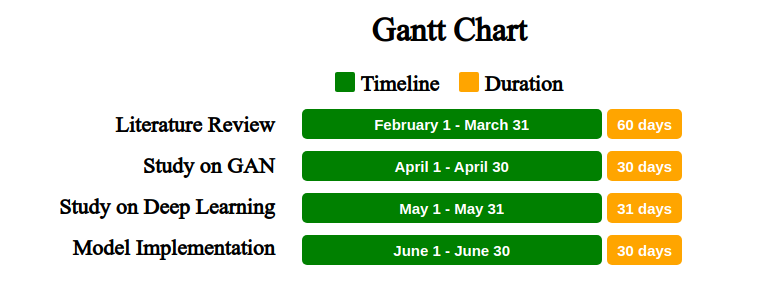


Figure: Gantt Chart of our work

**6. Conclusion:**

As there are some limitations in previous works while generating fake human faces using GAN, there are many places to improve the overall work.

We hope that our method will work out, remove the major limitations, and give high-quality fake human faces.

**7. References Cited and/or Bibliography of Planned Reading:**

1. **Zhang, Z., Pan, X., Jiang, S., & Zhao, P. (2020). High-quality face image generation based on generative adversarial networks. *Journal of Visual Communication and Image Representation*, *71*, 102719.**
2. **Wang, X., & Gupta, A. (2016, October). Generative image modeling using style and structure adversarial networks. In *European conference on computer vision* (pp. 318-335). Springer, Cham.**
3. **Ghatas, F. S., & Hemayed, E. E. (2020). Gankin: generating kin faces using disentangled gan. *SN Applied Sciences*, *2*(2), 1-10.**
4. **Hamdi, A., & Ghanem, B. (2019). IAN: Combining Generative Adversarial Networks for Imaginative Face Generation. *arXiv preprint arXiv:1904.07916*.**
5. **Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., & Ortega-Garcia, J. (2020). Deepfakes and beyond: A survey of face manipulation and fake detection. *Information Fusion*, *64*, 131-148.**
6. **Karras, T., Laine, S., & Aila, T. (2019). A style-based generator architecture for generative adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4401-4410).**

**8. List of Tools/Software required**

1. Jupyter Notebook
2. Google Colaboratory
3. Anaconda
4. Tensorflow
5. Keras