Be Comp a Project

Statement of Purpose

# Title:

# Design of system to generate photo realistic face images from textual description using GAN.

# Class:

BE COMP A

## Group no:

A17

# Name of group leader:

Nikita Mahajan 56

## Name of group members:

### Nikita Mahajan 56

Vikas Kushwaha 55

Rajeshwari Mishra 62

# Domain:

Intelligent System Design and Development.

# Category of project:

Core

# Type of project:

Inhouse

# Application:

Useful for Crime cases, engineers, scientist, designers. Designed for government and industry usage

# Problem definition:

Synthesizing high-quality images from text descriptions is a challenging problem in computer vision and has many practical applications. Samples generated by existing text-to-image approaches can roughly reflect the meaning of the given descriptions, but they fail to contain necessary details and vivid object parts. We are motivated by the potential of automated face generation to impact and assist critical tasks such as criminal face reconstruction.

# Literature Survey:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Sr.No** | **Title of the paper** | **Journal and year of publication** | **Summary of the paper** | **Gap identified** |
| 1. | Text2FaceGAN: Face Generation from Fine Grained Textual Descriptions | 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM) | In this work they presented captions for the CelebA dataset to facilitate face synthesis from text. Then they used Generative Adversarial Network to learn the conditional multimodality in synthesis of face from captions | Improve the selection of the wrong image for the GAN-CLS algorithm. Currently, they randomly select images from the dataset as the wrong image .Explore better language models such as BERT, analyze and compare performance of other GAN architectures with their model for face generation from captions |
| 6. | AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks | 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition | Attentional Generative Adversarial Network is proposed for fine-grained text-to-image synthesis which gives effective results for complex scenes. | To work on global coherent structure to eliminate the failure which could not be resolved by the current method.. |
| 5. | StackGAN: Text to Photo-realistic Image Synthesis  with Stacked Generative Adversarial Networks | 2017 IEEE International Conference on Computer Vision | It proposed the Stacked Generative Adversarial Networks (StackGAN) to generate 256×256 photo realistic images conditioned on text description & the outcome generated has higher resolution as compared to text-to-image generative models. | To improve the diversity of synthesized images and to stabilize the training of Conditional GAN introduced Conditioning Augmentation Technique. |
| 3. | A style-based generator architecture for generative adversarial networks | 2019 IEEE Conference on Computer Vision and Pattern Recognition | They proposed an alternative generator architecture for generative adversarial networks, borrowing from style transfer literature. The new architecture leads to an automatically learned, unsupervised separation of high-level attributes and stochastic variation in the generated images and it enables intuitive, scale-specific control of the synthesis. | We should also focus on text irrelevant contents in the original image. |
| 4. | Photographic Text-to-Image Synthesis with a Hierarchically-nested Adversarial Network | 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition | This paper presents a novel method to deal with the challenging task of generating photographic images conditioned on semantic image descriptions. The method introduces accompanying hierarchical-nested adversarial objectives inside the network hierarchies, which regularize mid-level representations and assist generator training to capture the complex image statistics. | Focus on increasing the visual quality and resolution of generated images. |

# References:

[1] Osaid Rehman Nasir, Shailesh Kumar Jha, Manraj Singh Grover, Yi Yu†, Ajit Kumar and Rajiv Ratn Shah, “Text2FaceGAN: Face Generation from Fine Grained Textual Descriptions” 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM).

[2] Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaogang Wang, Xiaolei Huang, Dimitris Metaxas in “StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks” 2017 IEEE International Conference on Computer Vision, pp. 5907–5915.

[3] Tao Xu, Pengchuan Zhang, Qiuyuan Huang, Han Zhang, Zhe Gan, Xiaolei Huang, Xiaodong He, “AttnGAN: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks” 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition.

[4] ] Zizhao Zhang, Yuanpu Xie, and Lin Yang, “Photographic text-to-image synthesis with a hierarchically-nested adversarial network,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 6199– 6208.

[5] Tero Karras, Samuli Laine, and Timo Aila, “A style-based generator architecture for generative adversarial networks,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 4401–4410.

[6] Guojun Yin, Bin Liu , Lu Sheng, Nenghai Yu , Xiaogang Wang , Jing Shao, “Semantics Disentangling for Text-to-Image Generation” arXiv:1904.01480v1 [cs.CV] 2 Apr 2019.

[7] Micah Hodosh, Peter Young, and Julia Hockenmaier, “Framing image description as a ranking task: Data, models and evaluation metrics,” Journal of Artificial Intelligence Research, vol. 47, pp. 853–899, 2013.

[8] Zizhao Zhang, Yuanpu Xie, and Lin Yang, “Photographic text-to-image synthesis with a hierarchically-nested adversarial network,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 6199– 6208.

# Name of the guide:

Mrs. Shiwani Gupta