|  |
| --- |
| BeCompAProject |

Day-2

# 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 |
| 2. | Controllable Text-to-Image Generation | arXiv:1909.07083v2  19 Dec 2019 | They have proposed a controllable generative adversarial network (ControlGAN), which can generate and manipulate the generation of images based on natural language descriptions. Our ControlGAN can successfully disentangle different visual attributes and allow parts of the synthetic image to be manipulated accurately. | They experimentally find that the channel-wise attention correlates closely with semantic parts of objects, while the spatial attention focuses mainly on colour descriptions. Without channel-wise attention, their model fails to generate controllable results when we modify the text related to parts of a bird. In contrast, their model with channel-wise attention can generate better controllable results. |
| 3. | Text-Adaptive Generative Adversarial Networks | arXiv:1904.01480v1  2 April  2019 | In this paper, they have proposed the text-adaptive generative adversarial network(TAGAN) to generate semantically manipulated images while preserving the irrelevant contents. | We should also focus on text irrelevant contents in the original image. |
| 4. | Semantics Disentangling for Text-to-Image Generation | arXiv:1810.11919v2  28 Nov 2018 | They design a Siamese machine in the discriminator to learn consistent high level semantics, and a visual semantic embedding strategy by semantic conditioned batch normalization to find diverse low level Semantics. | Focus on increasing the visual quality and resolution of generated images. |
| 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. |
| 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.. |

**Objectives:**

1.To generate images that would produce stable appearance patterns and learn the diverse expressions of humans.

2.To focus on text irrelevant contents in the original image and global coherent structure.

## 3. To improve the diversity of synthesized images and increasing the visual quality and resolution.