LITERATURE SURVEY

1) The research paper “Effective Removal of User-Selected Foreground Object From Facial Image Using a Novel GAN-Based Network” provides a novel method of removing a user select foreground object and generate the face with the help of GAN[ Generative Adversarial Network] architecture. This paper was published in the year 2020 by the authors NIZAM UD DIN, KAMRAN JAVED, SEHO BAE, and JUNEHO YI.

The authors of this paper has used CelebA-HQ dataset which is a collection of facial images of the celebrity across the world. They have prepared the occluded dataset using Photoshop and added occlusion such as mask, mic, sunglasses, hand etc…

The proposed method consist of two stages, a segmentation network which detects the foreground object and segment the foreground object and a GAN based network that fills the missing regions with the pixels in the background. GAN consist of two parts generator and discriminator, generator consist of encoder and decoder which is connected using U-Net architecture. They have two encoders that is vanilla conv encoder and partial conv encoder. Discriminator is divided into two parts global discriminator and local discriminator, global discriminator looks after the entire generated image while local discriminator focuses on the generated region. By using this method face generated will be accurate to the ground truth. The proposed method is evaluated on various dataset and results demonstrate its effectiveness in removing foreground objects and generating occluded region of the face.

The evaluation metric used in this paper are as follows: Structural Similarity(SSIM), Peak Signal to Noise Ratio(PSNR), Blind/Referenceless Image Spatial Quality Evaluator(BRISQUE). The performance of their model is high for removing microphone and sunglasses are high compared to mask and hands occlusion because of large areas of the face is covered using these object and reconstructing this is complex.

The paper provide a promising result for removing the foreground object and generating face. The author has compared their model both qualitatively and quantitatively to five different methods: GICA, EdgeConnect, PConv, GConv, and FaceD. The results shows that their model gives better output face image which is similar to the ground truth image compared to other method.

2) The research paper “Face Inpainting via Nested Generative Adversarial Networks” proposes a method of face inpainting using Nested GAN based architecture. This paper was published on October 25 2019 by the authors ZHIJIANG LI, HAONAH ZHU, LIQIN CAO, LEI JIAO, YANFEI, and AILONG MA. The dataset used is celebA dataset and occlusion to these images were added. 10% of dataset was used as test set and remaining 90% were used for training set. The proposed approach uses a nested GANs, where each level of the network is responsible for filling in fine details to the region of the face after the occlusion is removed. The network consist of coarse-level GAN and multiple fine level GAN. Masked image is provided as input to the coarse-level GAN which removes the masked object and give low resolution de-occluded image. Then this low resolution image is passed through multiple fine-level GAN and generates higher resolution image. The authors has also mentioned about the loss functions they have used, that is perceptual and contextual loss, which measures the similarity between the generated and ground truth images at inpainting benchmarks. The evaluation metric they have user are mean square error(MSE), peak signal-to-noise ratio(PSNR), structural similarity index(SSIM) and FACENET distance of repaired images are calculated. This results are compared with other two method, that is RMSProp optimizer and dualistic game theory combined with convolution network among which nested GAN perfomed well and gave better results.

3) The research paper “Object Removal and Inpainting from Image using Combined GAN” provides an efficiend method of removing any object from an image and inpainting it using combined GAN method. This paper was published in the year 2020 by the authors Jeongwon Pyo, Yuri Goncalves Rocha, Arpan Ghosh, Kwangheelee, Gungyo In and Taeyoung Kuc.

The dataset that they have used here is the cityscapes dataset which contains about 20,000 coarse annotation out of which 5000 are used for this model, of this 2975 for testing and 500 for testing purpose. The proposed architecture contains of 2 generators and 2 discriminators. The first generator( object removal image generator) generates fake object removal images and these images along with original object removal image are provided to first discriminator( Object removal image discriminator). Losses are updated. Once this is done the second generator ( inpainted image generator) generates fake inpainted images and these along with the original inpainted image are passed to the second discriminator. Once again losses are updated. This is done twice once for discriminator network updates and once for generative network updates. For better training purpose they have used batch normalization and leaky RELU after each convolution layer. The average accuracy of the object removal image discriminator obtained is 72.28% and the inpainted image discriminator is 72.63.

4) The research paper “Occluded Face Recognition in the Wild by identity-Diversity Inpainting” amins to aid occluded face recognition by proposing an identity-diversity inpainting method. The central idea is to combine GAN with a pre-trained CJJ recognizer that has been optimised to act as a third player and complete with the generator by identifying diversity within the same identification class. This paper was published in the year 2019 by the authors Shiming GE, Chenyu Li, Shengwei Zhao and Dan Zeng. The authors have used CelebA dataset for training and LFW dataset for testing. For training the model occluded images has been generated by taking a mXm mask to randomly cover on the set, following which these occluded face data is provided into each model. The model is divided into phases, first of which is the removal of occlusion and reconstructing of the missing region using GAN and using a pre-trained CNN model that recognize the reconstructed image. A pre-trained recognizer has been included into the GAN, acting as a third player to compete with the generator by spotting players, including the generator, the discriminator and the recognizer. An encoder-decoder architecture with dilated convolution is used by the inpainting network. The evaluation metric used in this paper are Peak Signal-to-Noise Ratio( PSNR) and Structural Similarity Index( SSIM). Speaking of the quantitative performance on inpainting results the proposed model gives a PSNR value as 31.5588 and the SSIM to be 0.9598. The model’s effectiveness in identifying obscured faces in the wild has been demonstrated by comparison with the four state-of-the-art systems by the author.

5) The research paper “Free-From Image In-painting with Gated Convolution” provides a new approach for image in-painting that is capable of generating high quality, consistent results for arbitrary missing regions. The authors aim to overcome the limitation of existing in-painting methods, which often require fixed in-painting masks and may produce unnatural or unconvincing results. The paper was published in the year 2019 by the authors Jiahui Yu1, Zhe Lin2. The authors of this paper has used celebA-hq dataset which is a collection of facial image of the celebrity across the world. They have prepared the occluded dataset using Photoshop and added occlusion such as mask, mic , sunglasses, hand etc… The proposed method is a Gated convolution encoder-decoder(GCE) that has been used for the task of image inpainting, it consist of an encoder that extracts features from the input image, a decoder that generates the output image, and a gating mechanism that controls the flow of information between the encoder and decoder. The gating mechanism allows the model to selectively use the features extracted by the encoder, resulting in more accurate and detailed in-painting results. By using this method face generated will be accurate to the ground truth. The Prosed method is evaluated on various dataset and results demonstrate its effectiveness in removing foreground objects and generating occluded region of the face. The evaluation metric used in this paper are as follows: Structural Similarity(SSIM), Peak Signal to Noise Ratio(PSNR), Blind/Referenceless Image Spatial Quality Evaluator(BRISQUE). Along with these metrics, the authors also carried out a user study in which participants were asked to rate the inpainted images' visual quality. The evaluation's findings demonstrated that the suggested strategy delivered cutting-edge outcomes in terms of both objective visual quality and quantitative measures.The efficiency of various parts of the authors' suggested method's various components was examined by ablation studies. They discovered that the gated convolutional neural network was an essential element that considerably enhanced the model's performance. Additionally, they discovered that the use of adversarial training and feature propagation helped their method perform better.Overall, the study's findings show that the suggested technique works well for free-form picture inpainting and has the potential to be applied in a number of situations, including image editing, image restoration, and the repair of damaged images.