

Review

Effective Removal of User-Selected Foreground Object From Facial Images Using a Novel GAN-Based Network

date of publication June 11, 2020.

This paper proposes a unique and easy-to-use method for face de-occlusion in facial images, where the user can choose the object to be removed focusing on commonly occurring occluding objects including (hands, a medical mask, microphone, sunglasses, and eyeglasses).

In this review we will focus on the main points in this method in order to support the removal of the mask from the face. This paper, conducted many researches on previous papers related to the same topic, as they found that most of the previous methods were only use vanilla convolution as the backbone of their deep-learning-based networks, this method helps in achieving well-incorporated predictions but leads to severe visual artifacts, especially at the boundaries of the valid and affected regions of the image and then they improved its schema to partial convolution, gated convolution. These methods perform better. However, they focus more on valid regions and do not effectively consider affected regions.

So to address the limitations they propose a novel GAN-based with two discriminators. The model consists of two stages: an object detection module, and an image completion module. And they integrate both vanilla and partial convolution. By using these two encoders in parallel it generates well-incorporated and sharp content with fine details. implemented in tensorflow.

Because there is no available dataset that contains facial image with and without occlusion objects they constructed dataset using CelebFaces Attributes Dataset (CelebA) and CelebA-HQ dataset.

The quantitative performance using four metrics: 1) Structural Similarity (SSIM) the result on mask was **0.908**, 2) Peak Signal to Noise Ratio (PSNR) with **28.727**, 3) Naturalness Image Quality Evaluator (NIQE) with **4.425**, and 4) Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) with **40.883**. The accuracy is not mentioned, they were comparing the performance of their model with different models.

The advantage on this paper is about how they use vanilla and partial convolution in parallel to give results that are visually pleasing and removed complex objects like mask which covers a large part of the face which perform better compared with other models. Disadvantage that the Failure cases occur when more than 70% of the face is occluded, especially on hand.