CSC 528 – FINAL PROJECT REPORT

Acne Detection and Removal in Digital Images

Name – Goutham Selvakumar

TABLE OF CONTENTS

I. ABSTRACT:

This project aims to develop an automatic pipeline for detecting and removing acne from facial images, saving valuable time spent on manual image processing. The proposed system will leverage advanced face detection and key point detection techniques, coupled with multiple acne detection methods, to create a reliable and accurate solution for users seeking to enhance their images for social media platforms.

II. INTRODUCTION:

In today's world, we are able to see a lot of people affected by various skin diseases. These skin diseases causes them to lose the self confidence in many specific fields. Here, Acne is considered to be the most spread skin diseases in comparison to the other skin diseases. In the society that we live right now and with this generation of people living with us, all of them look upon our skin and judge a person's characteristics, so, having a flawless and healthy skin is considered to be the most among everyone.

As we can see, the people on the social media nowadays are obsessed with their skin being perfect, so they are spending lots of money on the plastic surgery, with the various apps on the store that has both the unpaid and paid versions. Furthermore, by implementing this project, we are able to create an automatic solution for acne detection and removal, which can be helpful for users who are seeking to improve their facial images for social media platforms. The comparison of different acne detection methods will further help in identifying the most effective and accurate approach for this specific application.

III. BACKGROUND:

The **"How far are we from solving the 2D & 3D Face Alignment problem?"** [1], Paper illustrated that that the deep neural network is attained using an improved performance on the existing 2D and 3D face alignment datasets. They combined the localization with state-of-the-art residual block for a strong baseline and trained it on a 2D facial landmark datasets. Next, they processed the 2D landmarks network into the largest and difficult 3D facial landmark dataset with more than 230,000 images. Final step was to train the network for 3D face alignment and the LS3D-W was introduced. They were able to show that the accuracy was very much improved with both the 2D and 3D face alignment.

[2] The second paper "**One millisecond face alignment with an ensemble of regression trees**" explains about the problem of face alignment for a single image. They tried to show the regression trees that were used in order to estimate the face's landmark positions directly from the pixel intensities, were able to get an enormous performance results for predictions to be easier. The regression trees here optimizes the sum of square error loss and the missing labels are also handled. The researchers also analysed the quantity of training data with the accuracy and the effect of data augmentation using the synthesized data.

"**Deep Image Prior**" [3] illustrates that the deep convolutional networks are suitable for image restoration and generation. They were able to represent the architecture of a network that is capable of capturing the low level image to any learning. It was achieved by taking the randomly selected network that was used for standard inverse problems like denoising, super resolution. With the same method they were able to restore the images completely, they highlighted the inductive bias from the network. They were able to achieve it successfully by combing both the method of deep convolutional networks and handcrafted image known as the self-similarity.

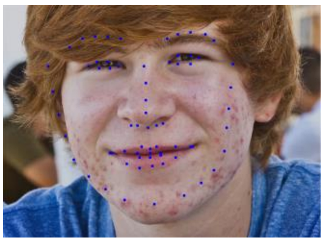
IV. METHODS:

For this project, I'm using the dataset called as 'ACNE04' that consists of 1,457 images that represents different impacts of acne that covers the face. This file consists of XML file that has the corresponding bounding boxes for pimples. Firstly, I tried to calculate and find the co-ordinates of the forehead points based on the facial landmarks with the image from the dataset.

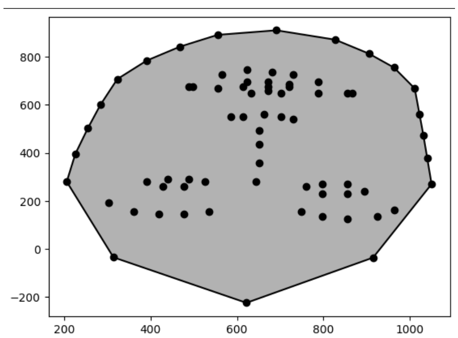
I loaded the image and the corresponding mask, resized the image if it is exceeding above 2500 pixels in which the image is converted to the RGB and grayscale method that reads and executes the ground truth mask.

Next step, was to perform the facial landmark detection with the help of FaceAlignment Library on the RGB image shown below. From the output, we can see that it draws a circle around the detected landmarks and image is displayed afterwards with the marked landmarks.

The extra landmarks around the forehead is also calculated and the resulting image is to be found with the landmarks and the forehead points. Using the script library from I was able to compute the convex hull of a set of points. What it does is, the facial landmarks and an additional forehead points are combined into an array which calculates the convex hull. Thus, we might see the circles drawn on the images specifically at the vertex places by iterating over the vertices on the convex hull.

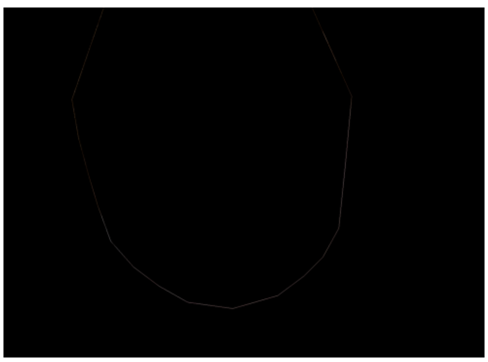
With the help of Scipy library and Matplotlib a set of points was plotted and the respective convex hull was visualized. The procedure is that the facial landmarks and forehead points are combined in order to create a singly numpy array and then the convex hull is computed. From the below output, I can illustrate that those black points connects the vertices and the areas the enclosed by the convex hull.



After computing the convex hull for the facial landmarks in the given image, a binary acne mask was created that initializes an image array and its ROI [Region of Interest] was defined for various parts in the image like the eyebrows, eyes, nose, and mouth.

From that the regions where these parts are present is dilated especially for eyes and nose regions. The later photo represents the acne present in the facial regions of interest. In order to show the acne present, I used the LAB colour space that highlights that specific region where a binary mask was created by the thresholding technique and the morphological operations were done in order.

A set of points were implemented for the convex hull with a mask based on the hull is also created. This mask initializes an image mask with zeroes and draws a triangle on the mask. Moreover, we can see from the image below that the mask wash applied with the mask being set to black by the pixels outside. The masked image, the only visible thing is the pixels within the convex hull.



V. RESULTS:

I created a function known as the find_mask for every index from 1 to 30 together in two different modes. The returned scores are stored in two lists and the index for the respective scores from both sides are printed.

0.72	0.613
0.548	0.6
0.458	0.531
0.634	0.772
0.13	0.142
0.353	0.597
0.075	0.07
0.354	0.791
0.27	0.667
0.693	0.465
0.771	0.681
0.373	0.314
0.685	0.567
0.611	0.715
0.553	0.544
0.39	0.448
0.333	0.539
0.439	0.594
0.53	0.602
0.239	0.299
0.358	0.3
0.479	0.581
0.517	0.493
0.097	0.052
0.084	0.109
0.143	0.317
0.332	0.492
0.668	0.798
0.371	0.372
0.33	0.507

From the above output, we can see that the scores are printed from that function for each of the index from 1 to 30 that ran in two different modes. The scores for first and second mode is showed, that corresponds to the index and the output is also rounded in three decimal places.

A U-Net architecture for a convolutional neural network was developed in which the network is often used in the image segmentation. It basically consists of encoder and decoder where the encoder uses normalization, dropout, and ReLU activation function whereas the decoder does upsampling.

I trained a neural network model that was done for a specific number of epochs on an image with the respective mask that uses an optimizer and loss criterion. The model's parameters are updated using the backpropagation and gradient descent, the model is evaluated and a prediction gets generated that is later appended to a list of images which will be returned for final prediction.

VI. CONCLUSION:

From the above output, it's shown that the facial landmarks were able to remove the acne only in some of the areas. I tried initializing the model with the MSE (Mean Squared Error), and Adam optimizer. I tried implementing the techniques like the segmentation mask, U-Net architecture, and deep learning tasks in order to remove an acne from the facial landmark regions using the ACNE04 dataset. From the above output, it's shown that the facial landmarks were able to remove the acne only in some of the areas.

VII. RESEARCH REFERENCES:

1. Bulat, G. Tzimiropoulos. (2017). *How far are we from solving the 2D \& 3D Face Alignment problem?* (And a dataset of 230,000 3D facial landmarks). IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION. https://openaccess.thecvf.com/content_ICCV_2017/papers/Bulat_How_Far_Are_ICCV_2017_paper.pdf

2. Kazemi, V., & Sullivan, J. (2014*). One millisecond face alignment with an ensemble of regression trees*. IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION. https://www.cv-foundation.org/openaccess/content_cvpr_2014/papers/Kazemi_One_Millisecond_Face_2014_CVPR_paper.pdf

3. Ulyanov Dmitry, Vedaldi Andrea, Lempitsky Victor. (2017). *Deep Image Prior*. INTERNATIONAL CONFERENCE ON COMPUTER AND INFORMATION TECHNOLOGY (ICCIT). *https://openaccess.thecvf.com/content_cvpr_2018/papers/Ulyanov_Deep_Image_Prior_CVPR_2018_paper.pdf*