COL783: Assignment 2
Site: https://knight94.github.io/

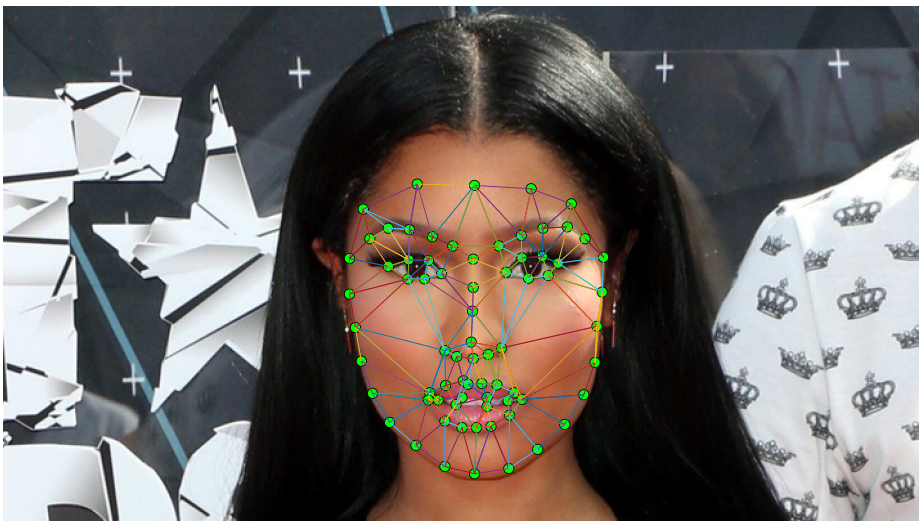
Naman Jain (2020anz8848) and Siddharth Shrivastava (2020anz8849)

1 Objective

To transfer the facial makeup feature from reference image to target image.

2 Get the facial landmarks using TCDCN

Using TCDCN, initial face landmarks are detected. Image is display with the points, user can input the final points. Total of 73 points are collected which include marking of eyes, nose, lips and eyebrows.



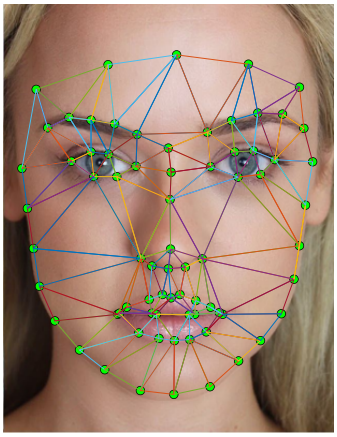


Figure 1: Reference images

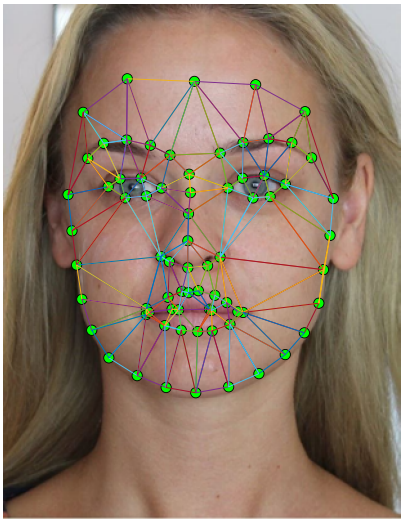
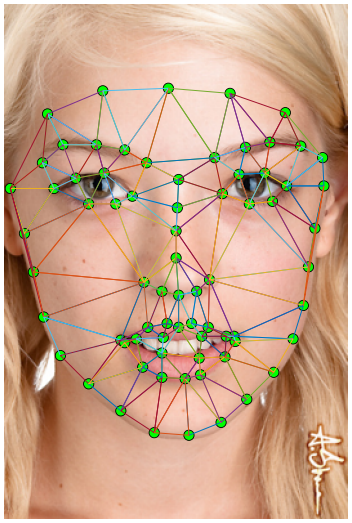



Figure 2: Target images

2.1 Image warping

- Now, using the 73 facial points number of triangles are determined using delaunay method. This will give 125 number of triangle in form of triplets that is coordinates of (x,y) points. Index matrix will be same for target and reference images.
- Determine the affine transformation function for each triangle in the both images.
  - This is done via following operation:

$$\begin{matrix} Rx \\ Ry \\ 1 \end{matrix} = \begin{matrix} a & b & c \\ f & g & h \\ 0 & 0 & 1 \end{matrix} * \begin{matrix} x \\ y \\ 1 \end{matrix}$$

$$X = \begin{matrix} x1 & y1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x1 & y1 & 1 \\ x2 & y2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x2 & y2 & 1 \\ x3 & y3 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x3 & y3 & 1 \end{matrix} \quad b = \begin{matrix} Rx1 \\ Ry2 \\ Rx2 \\ Ry2 \\ Rx3 \\ Ry3 \end{matrix} \text{ then } \begin{matrix} a \\ b \\ c \\ f \\ g \\ h \end{matrix} = b * X^{-1}$$

- Using the calculated matrix, determine the transformed spatial location for each pixel in Target image. This is done using baycentric method to locate the triangle and transformation matrix. This will give some (x,y) in reference image then the pixel intensity is determined using interpolation.

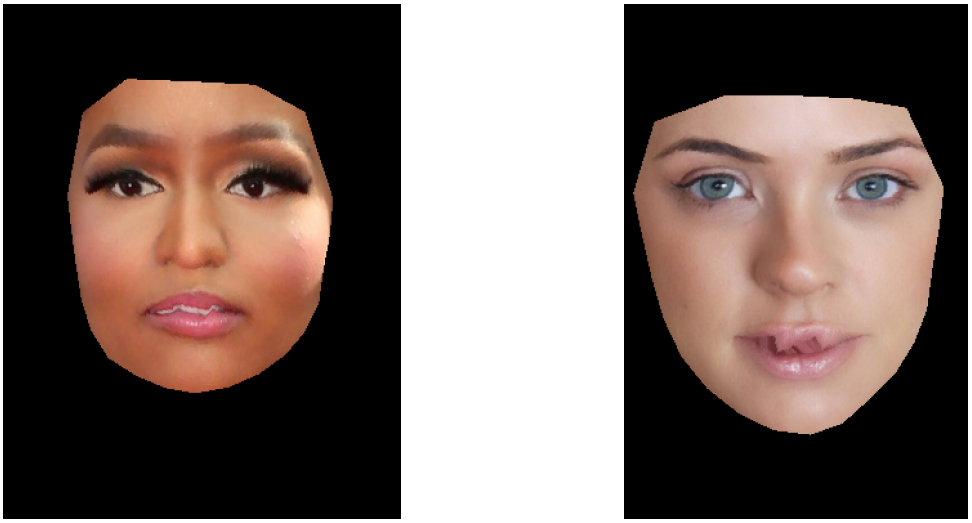


Figure 3: Warped images

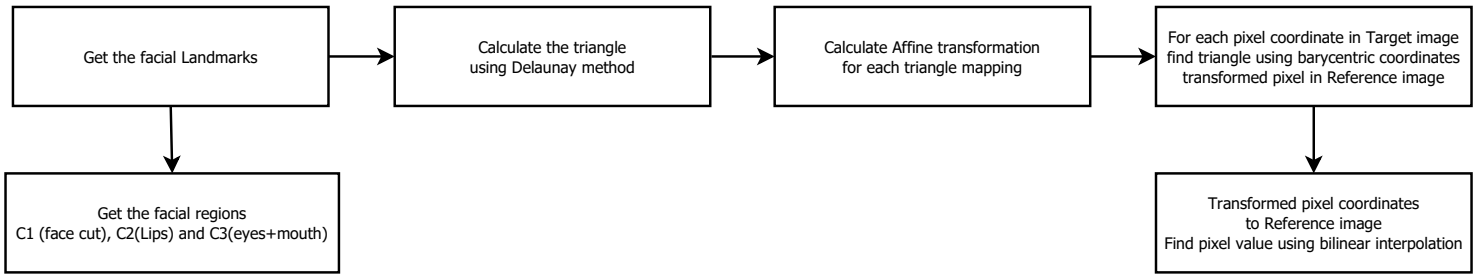


Figure 4: Warping flowchart

- Further, face region is partitioned into eyes, mouth, lips and rest of the face.

2.2 Decomposition of the images

- RGB images are converted to LAB space to separate color information.
- Next images are separated into detail and large layer using bilateral filter
  - Large scale layer is output of bilateral filter on L image
  - Subtract/divide previous from L to get skin/detail layer
- Decomposed images are shown in Figure 5.

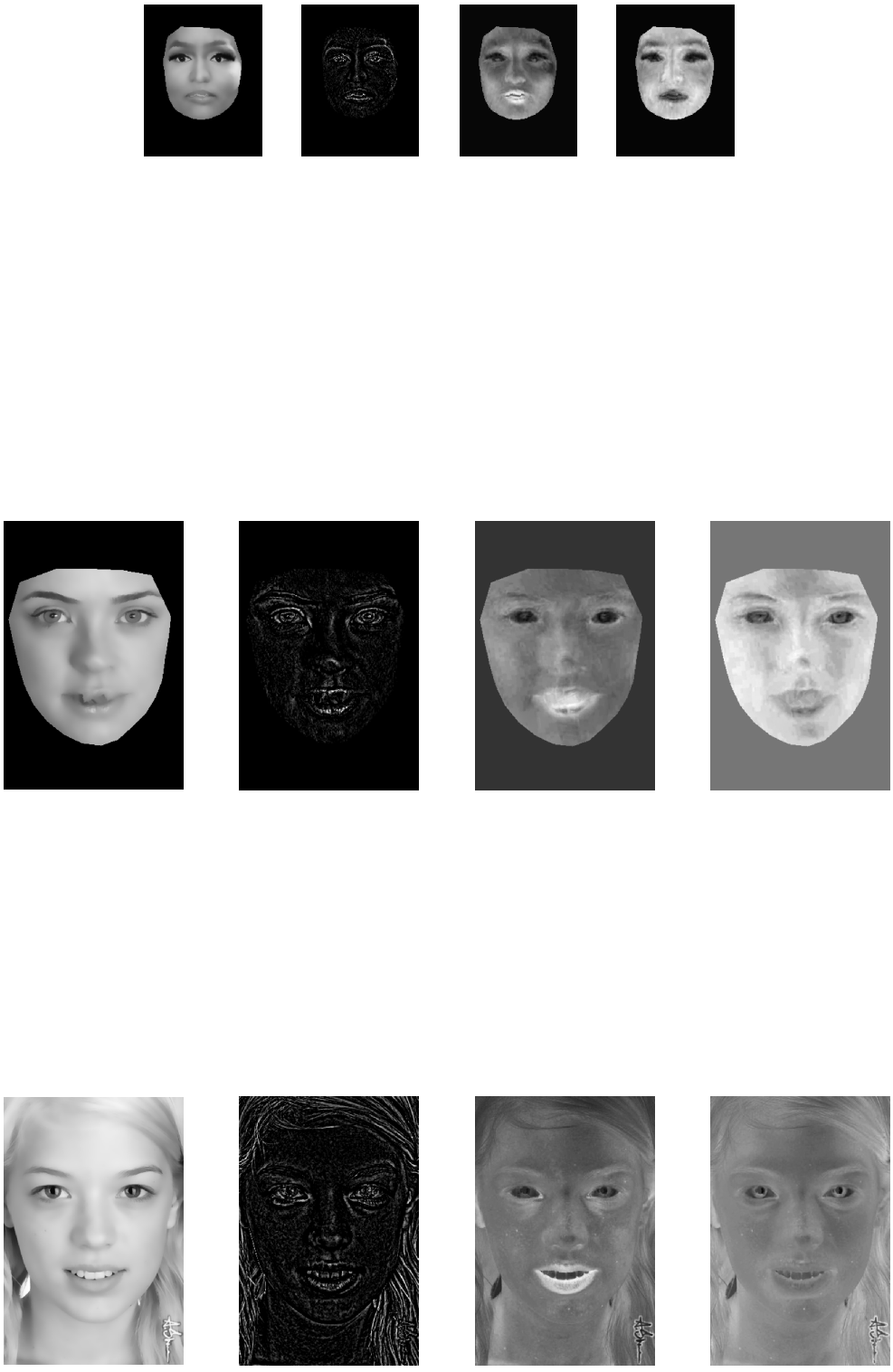


Figure 5: Decomposed images

3 Image transfer

- Different components of the images are merged separately-
  - Large scale layer is transfer using gradient editing method
  - Skin detail is transfer using weighted method in which target image contribution is set to 0.5
  - Color transfer is done via alpha blending
  - Lip transfer flow is depicted in the flowchart

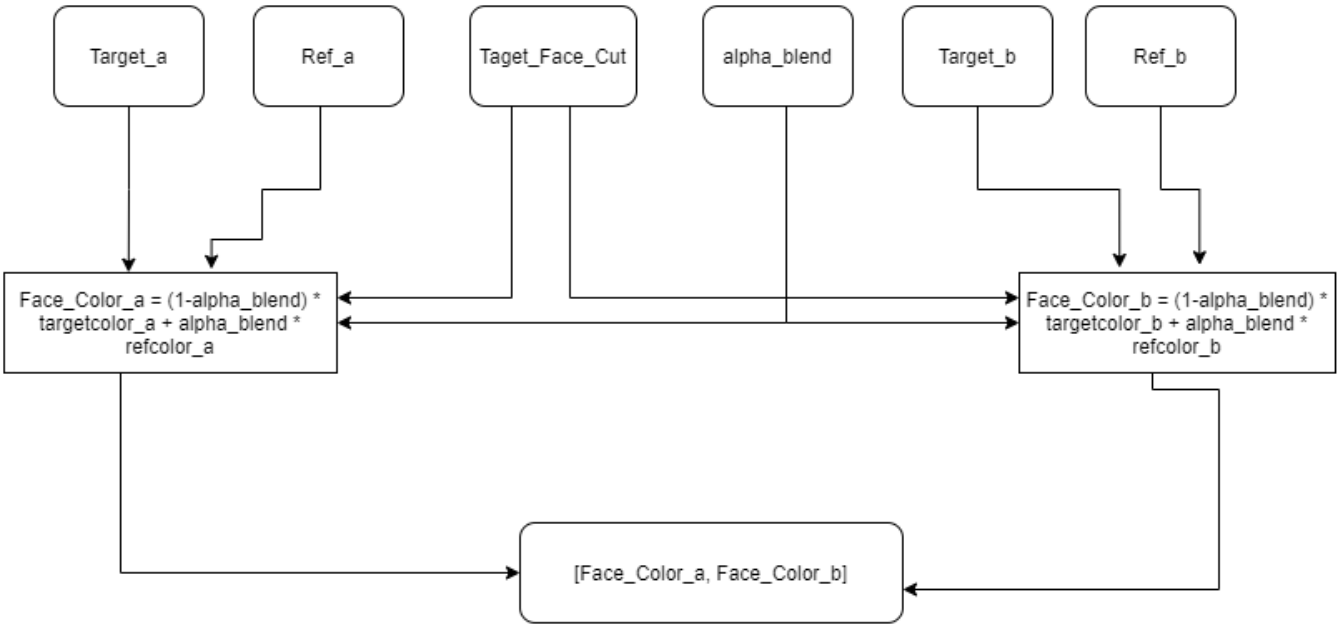


Figure 6: Color transfer

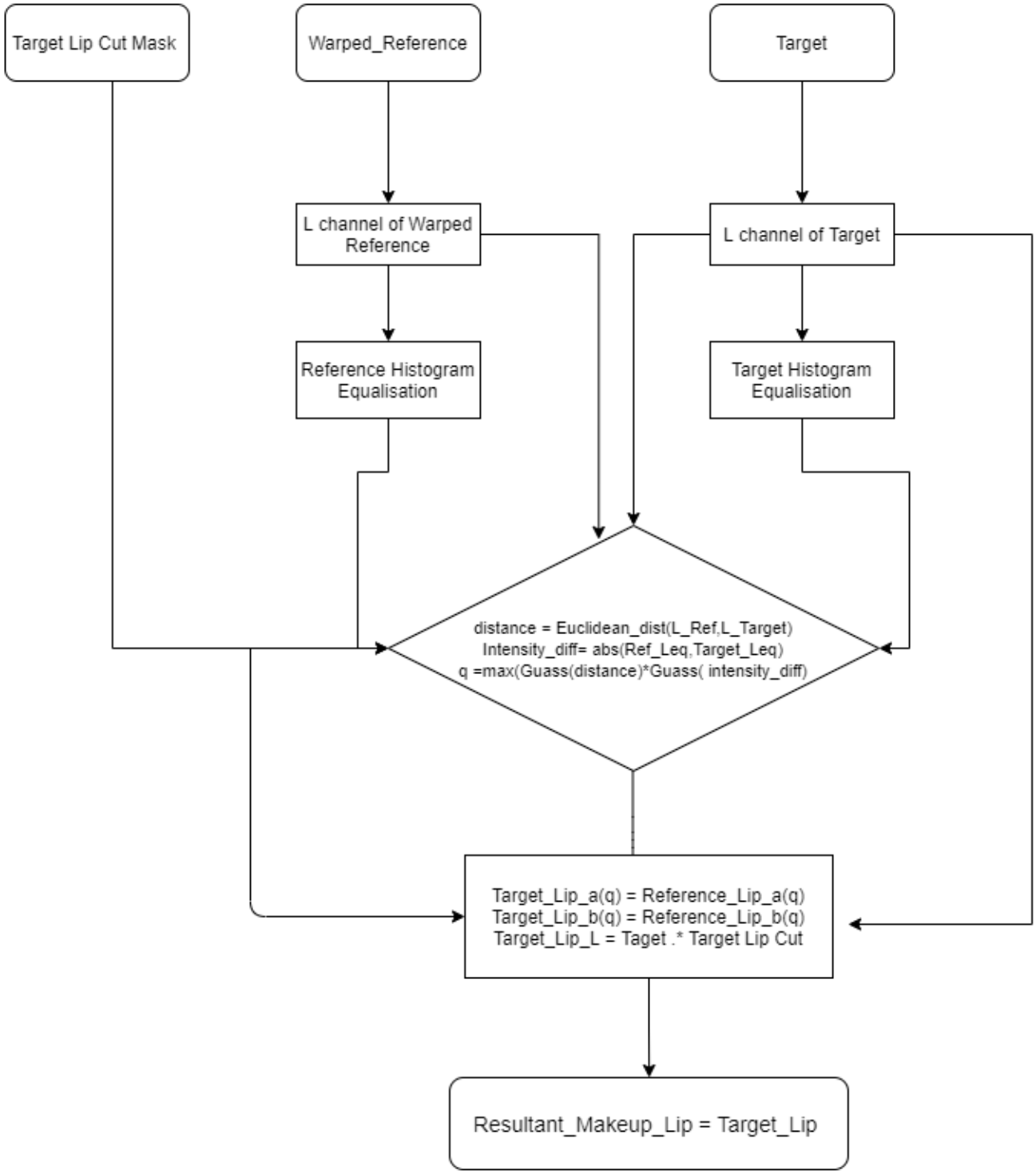


Figure 7: Lip transfer

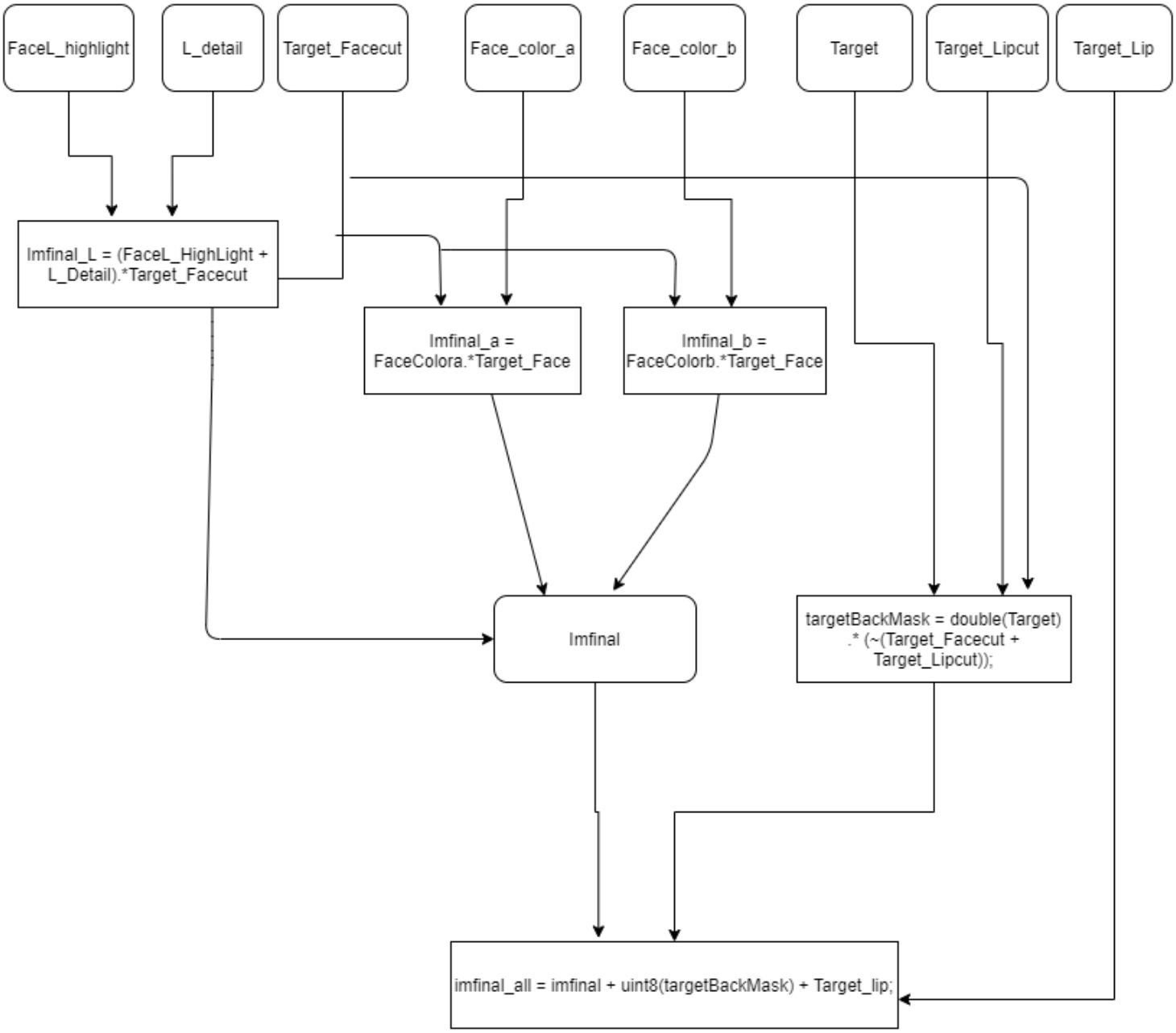
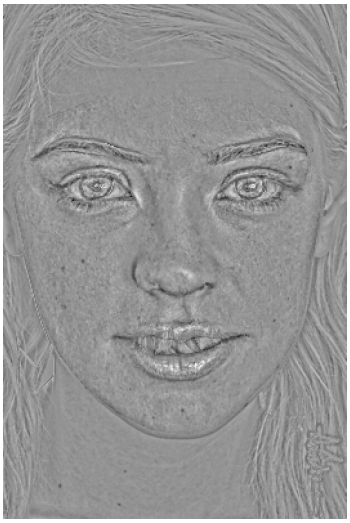
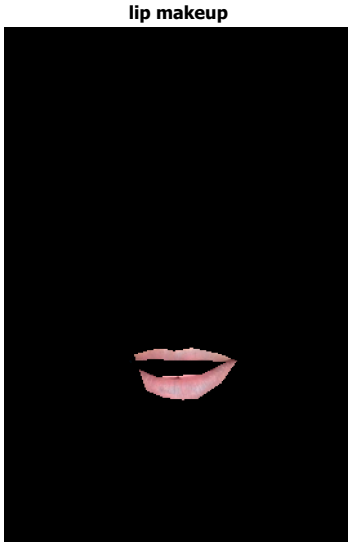


Figure 8: Merge images





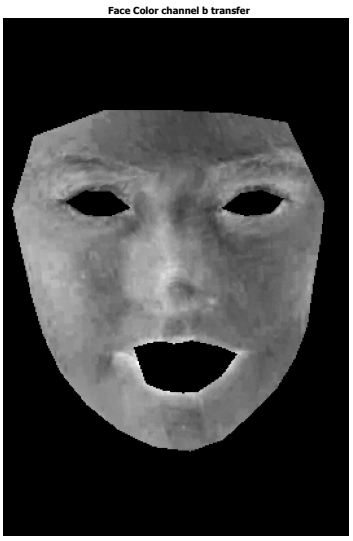


Figure 9: Transfer result for test case 1

4 Results

Final images are shown -

In the gradient editing, the process took the gradient from the reference image thus reconstructed image features are not matching in the final image. In the time frame, we were not able to improve the implementation of gradient editing. Need to sanitize the code for the beta generation. The main parts of the assignments (i) warping (ii) decomposition of the image into components are implemented correctly.

Figure 10: Final image

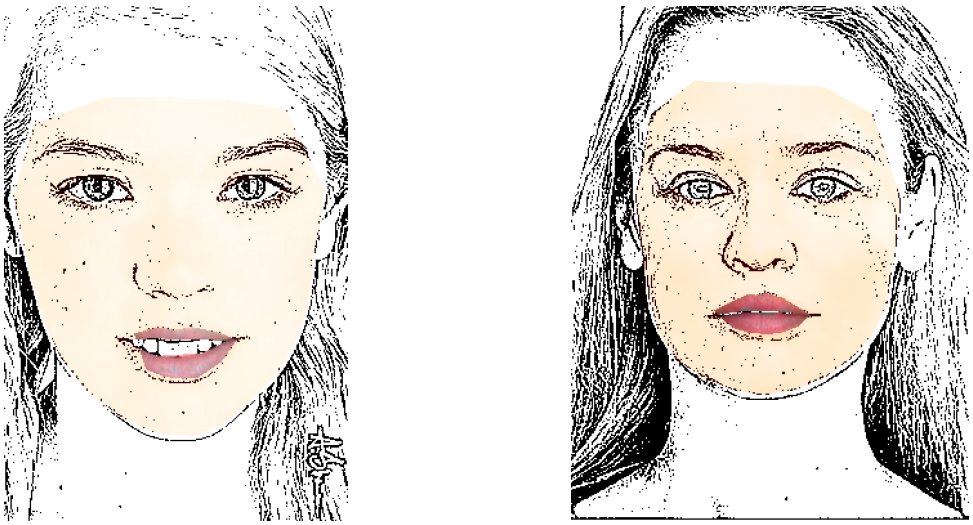


Figure 11: X_DoG image

References

[1] Guo D, Sim T. Digital face makeup by example. In2009 IEEE Conference on Computer Vision and Pattern Recognition 2009 Jun 20 (pp. 73-79). IEEE.

[2] Winnemöller H, Kyprianidis JE, Olsen SC. Xdog: an extended difference-of-gaussians compendium including advanced image stylization. Computers & Graphics. 2012 Oct 1;36(6):740-53.

[3] Zhanpeng Zhang, Ping Luo, Chen Change Loy, Xiaoou Tang. Facial Landmark Detection by Deep Multi-task Learning, in Proceedings of European Conference on Computer Vision (ECCV), 2014