Makeup-Go: Reversion of Portrait Edit

Team: Yehu Chen, Mark Jin, Zixuan Li, Yi Wen, Haoxiang Wu

The University of Michigan – College of Engineering EECS442 Computer Vision

Executive Summary

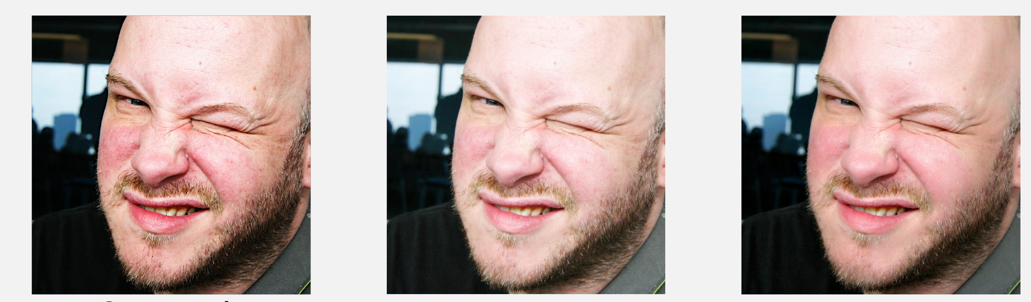
Advance of camera and image processing APPs gives rise to an overwhelming trend of artificial beautification. Our project aims to reverse the portrait beautification despite lack of information of the exact process applied. To precisely capture the subtle features of facial images, we apply PCA to the residual images of the original and beautified images, and then train several convolutional neural networks to regress each normalized components, and finally restore the unbeautified faces.

Background & Impact

Virtual Face beautification or markup becomes common operations in camera or image processing APPs, which is actually deceiving and is leading an unhealthy trend in aesthetics. Our project serves to divert people's overwhelming attention from external beauty to inner beauty and make an appeal for enhancing one's beauty by building self-recognition, self-confidence, and self-esteem and thus puts the general idea of beauty back onto the right track.

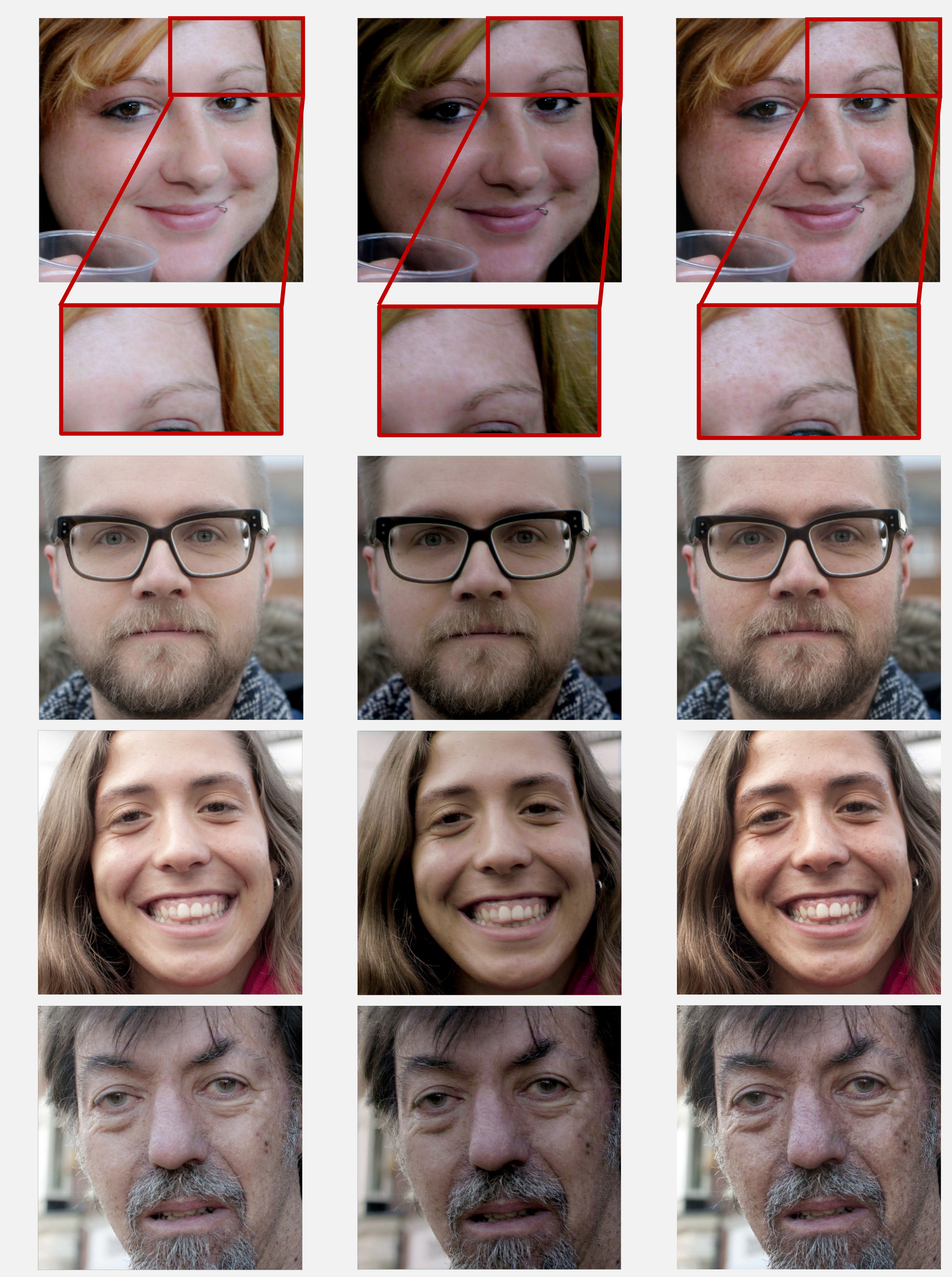
Scope of Application

Virtual beautifications are performed quite differently in software. To make the problem trackable, we limit the operations to whitening and skin smoothing. Also, we assume subtle cues still exist after virtual beautification.



Original Whitening Skin Smoothing

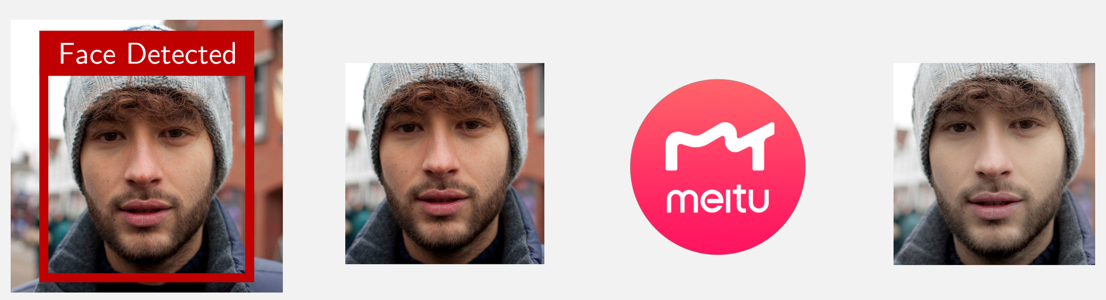
Results



The validation result indicates that

- Our model can capture and reverse the whitening effect, but still variant to the different extent.
- Our model can capture and reverse the skin smoothing basically and is subject to the amount of remaining clues on the beautified image.

Dataset Preparation



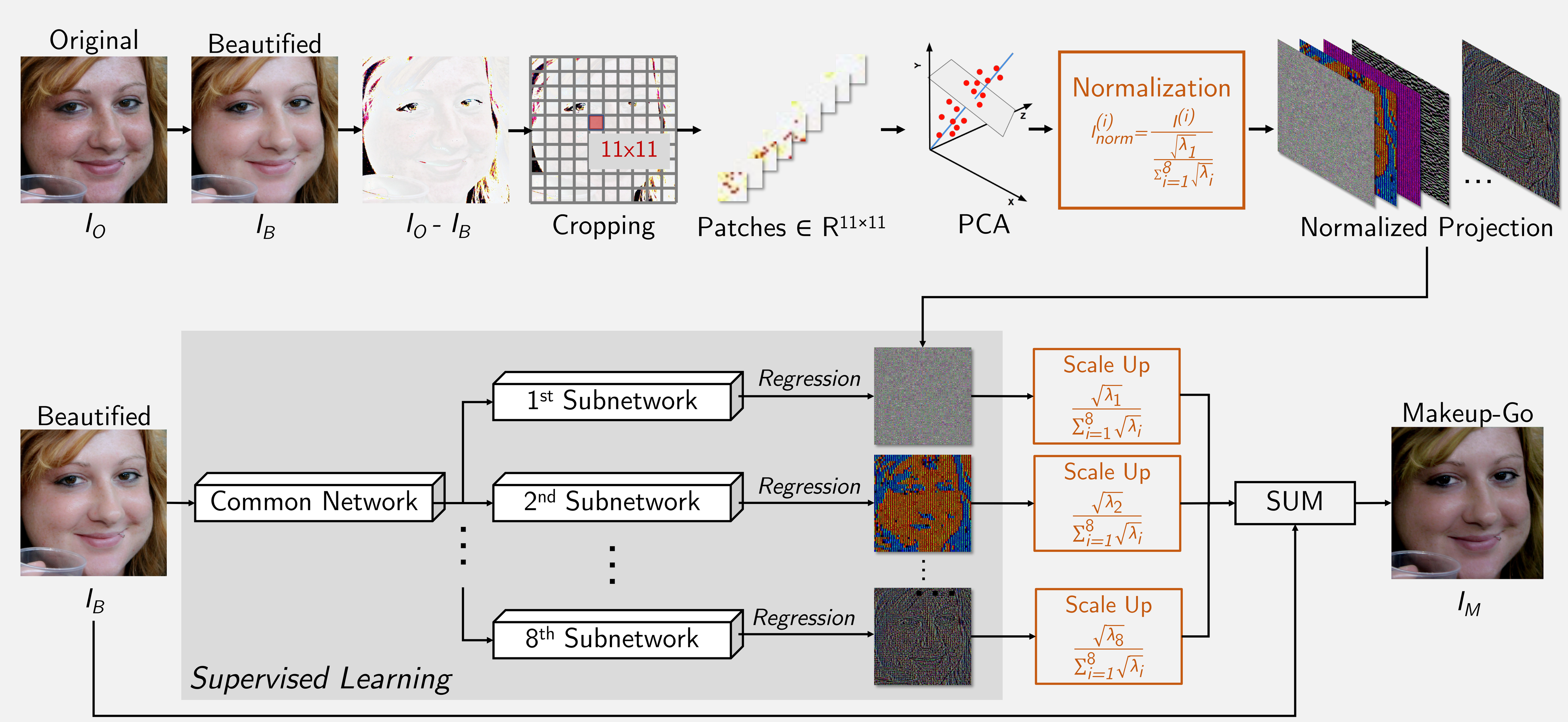
Crawled Image Cropped Image Beautification Beautified Image

Due to the lack of dataset available to approach the target, a dataset of 1200 photos has been manually constructed. 2000 portraits were crawled from Flickr. Face detection is performed on the portraits to crop the face so that the face lies in the center and scale invariance is eliminated. 1200 images are selected after deliberate filtering and outlier removal. MeituPic, a portrait beautifying software is used to perform beautification with different levels of whitening and skin smoothing.

Component Regression Network

Component Regression Network is built to reverse the beautification of a portrait without knowing the previous editing details. The input is the beautified image I_B and the output is the recovered image I_M.

The residual image is cropped into patches with size 11×11 each to learn subtle modifications. Principle Component Analysis is performed on the cropped patches on the whole training set. After projecting the cropped patches onto the first 8 principle components, we perform the normalization to make all components share the similar variance scale. The Dominance Effect is thus eliminated. The projected patches are rearranged into the original order and are fed into the model to supervise regression.



The regression network consists of two parts: a common network and a group of subnetworks. The common network works on extracting features from the beautified image I_B. The output of the common network is passed to every subnetwork. Every subnetwork performs convolutional regression for a scaled component. Except taking the common features as input, these subnetworks work individually.

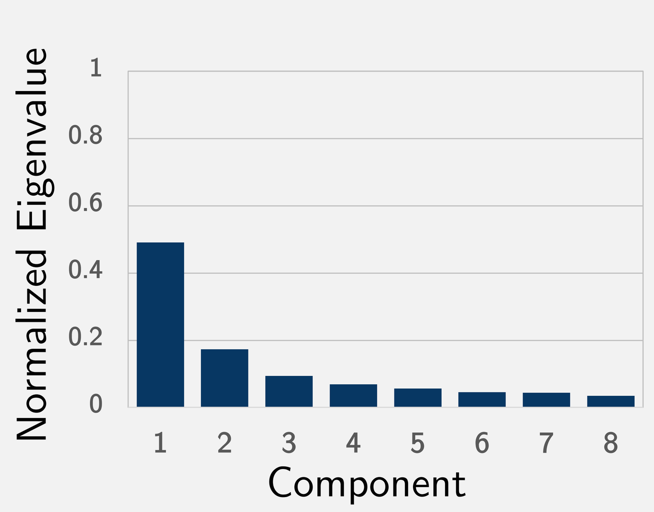
The loss is calculated as the sum of the MSE of each regression. Adaptive learning rate scheduler and gradient clipping is used to fasten learning. The model is trained with 4 images per batch and 15 epochs.

Dominance Effect

The reason traditional networks fail to regress low-ranking components is the deployment of the Euclidean loss, which is mainly controlled by the largest principle components.

$$J = \frac{1}{2m^2}\|F(v_X) - v_Y\|^2 = \frac{1}{2m^2}\|U^T F(v_X) - U^T v_Y\|^2$$
$$= \frac{1}{2m^2}\sum_{i=1}^{m^2}\|u_i^T F(v_X) - u_i v_Y\|^2$$

Normalization is performed to make all components share the same scale. Low-ranking components are now learnable.



Discussion & Future Work

Difference from Original Paper

During normalization, we divide the image by $\frac{\sqrt{\lambda_j}}{\sum_{i=1}^{8}\sqrt{\lambda_i}}$ rather than $\sqrt{\lambda_j}$, which makes the loss fall in a reasonable range and facilitates backpropagation.

Potential Ways to Improve Performance

Our current model is not powerful enough due to constrained time and resources. We can further

- Increase the number of images in the dataset, the batch size and the number of iterations.
- Use stronger computational resources to perform hyperparameter tuning.
- Research on how portrait beautifying software performs beautification operations and improve the model structure.

Reference: Y. Chen, X. Shen and J. Jia, "Makeup-Go: Blind Reversion of Portrait Edit," *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2018, pp. 4511-4519.