Beyond Colour Matching for Makeup Transfer

Chinmayi C. R.
Dept. of Information Technology
NITK Surathkal
Mangalore, India
chinmayicr27@gmail.com

Utkarsh Meshram
Dept. of Information Technology
NITK Surathkal
Mangalore, India
utkarsh.181it250@nitk.edu.in

***Abstract*—Transferring makeup from a reference image to a supply face is referred to as makeup transfer. The makeup that we are used to within the real world are a lot of varied and wild, including both colors and patterns, like stickers, blushes, and jewelry. Throughout history, people have admired beautiful faces. It is still very difficult due to the variety of makeup combinations, but makeup recognition can be very useful in a variety of commercial and recreational activities. The generated models based on deep learning do not give satisfactory results. Hence, we aim to modify the colour distribution while simultaneously maintaining the form and look of the makeup pattern. We also show excellent results for heavy makeup styles which are not accounted for in many previous works.**

***Index Terms*—makeup transfer, patterns, stickers**

I. Introduction

There has been a makeup application for a long time. In the past few decades, makeup as an industry has grown rapidly. With the advent of technology, facial beautification with makeup has been popularly used. Many applications that makeup features like eyeshadows, blushes, eyecolor, lipsticks etc. exist. Makeup transfer technology is used to try new styles by users and companies. Makeup transfer is the process of transferring makeup style from a reference image to the target image. This is particularly challenging because makeup components has to be extracted from the face considering the facial structure and orientation, occulusions, different lighting and head poses.

A. Motivation

- It is important to handle transfer of arbitrary makeup styles without any particular interaction.
- Most previous works do not support different makeup styles including light and extreme with accurate performance
- A model to accomodate tattoos or facial patterns along with vivid makeup styles.
- People want to understand what kind of makeup style best suits their facial structure.
- Technology can be provided as a tool to automatically facilitate this process so that even a novice user can interact with the makeup transfer system.

B. Contribution

- Transfer makeup style while retaining the natural face features.
- Account for different scenarios of makeup transfer
- Develop a comprehensive makeup transfer method that works for both light and extreme styles.
- Retain the original shape and quality of the facial image
- Introduce new makeup transfer data sets
- Successfully transfer the face blushes to capture the glowing skin foundation.
- Obtain state-of-the-art quantitative and qualitative performance.

II. Literature Survey

Beautyglow: On demand makeup transfer framework[1] can be easily extended to other applications in which two latent vectors need to be decomposed. It is possible to manipulate latent vectors to produce realistic images, even when the makeup is light. It focuses only on the conventional regions of the face i.e. lips and eyes and is a naive approach to transform from light to heavy makeup.

The paper [2] uses Local adversarial disentangling network to transfer and remove dramatic makeup styles seamlessly. It is possible to generate photorealistic results that preserve facial identity. It struggles to remove extreme makeup styles where colors are highly consistent in local regions but vary sharply across local patches. Extreme makeup styles fail to produce satisfactory results. Suffers when the head pose of the source and the reference faces are different, producing noticeable artifacts.

Adaptive Makeup Transfer via Bat Algorithm[3] achieves to have the most suitable makeup but also the most appropriate lightness for a certain makeup. They combine a beauty algorithm with an intelligent optimization algorithm to achieve good results. Their model however, only takes the frontal and upright face as target. It also does not work for extreme makeup styles.

PSGAN[4] can modify the shade of makeup transfer by virtue of its spatial-aware makeup matrices, substantially extending the application range of makeup transfer. It is capable of handling various expressions and poses of the head. However, they do not achieve results on retaining facial features following makeup transfer. Their methods fail to perform transfer on faces with different sizes.

PairedCycleGAN[5] use their network for makeup styles that can be transferred between people with different skin tones while preserving the original skin tone and other important identity details. An attempt is made to predict a natural face using a makeup face as an input. The network will not work well when extreme makeup styles are used that have not been learned during training. Also, some details may not be transferred. Since their training data set consists mostly of clear skin, their model cannot reconstruct blemishes.

CA-GAN[6] can be trained on unlabeled images by passing a noisy proxy for the relevant attribute through a weakly supervised learning approach. Their architecture demonstrates that this architecture can be used to perform makeup synthesis with a continuous color space and transfer makeup style. The model learns to modify color only, without affecting the other image attributes such as shine and specularities. The use of segmentation masks would require that the images be annotated when trained in a weakly supervised manner.

Disentangled Makeup Transfer with Generative Adversarial Network[7] can faithfully maintain the one-to-one mappings between face images and identity / makeup codes. Depending on how much the foundation color changes, their model produces results indicating a visible boundary between the upper and lower body.

The authors Anpei Chen et al.[8] provide a model with high-fidelity proxy geometry and fine geometric details based on emotional cues, expressions, appearance, and lighting. Their technique can produce, from a single image, ultra high quality 3D faces with fine geometric details under various expressions and lighting conditions. Their model generates realistic facial features like wrinkles. There was no consideration of occlusions (such as hair or glasses), hard shadows, which could lead to incorrect displacement estimations. Because it's reliant on accurate pixel appearance distributions, geometric detail prediction can't handle low resolution images.

III. Datasets

A. CPM-Real

This dataset consists of facial images including various makeup styles from light to heavy. It contains makeup styles with both makeup and patterns. It consists of 3895 RGB pictures. It has been chosen as it provides big selection of makeup designs, ages, head poses, etc. Multiple size.

B. CPM-Synt-1

This dataset consists of faces with synthetically created makeup patterns. The dataset is split into disjoint training and testing subsets of size 4182 and 1373, respectively. There are 5555 RGB texture pictures.

C. CPM-Synt-2

The requirement of CPM Synt 2 dataset is due to the fact that CPM Synt 1 is not suitable for evaluation purposes. It does not follow the transfer setup and hence the need for CPM Synt 2 dataset. This dataset contains three images: source image, reference image and the ground truth. It is specifically designed for pattern based evaluation. The dataset consists of 1625 triplets of images as mentioned. The use of this dataset for evaluation needs to satisfy one condition i.e. the source and the reference image of the same type of colour style.

D. Stickers

577 RGBA pictures were used for pattern transfer. The patterns are of varied sizes. It contains several patterns such as flowers, crystals, gems, henna, daisy, leaf, tattoo etc.

IV. Methodology

The input images I_{ns} , I_{mr} are converted to UV texture maps T_{ns} , T_{mr}.

This is done to remove misalignment i.e. head poses, 3D shape etc, between input and targeted image. We use PRNet, which has extended this idea and introduced a UV position map representation to encode any 3D face shape. PRNet is a network that process 2D image with three channels encoding the XYZ coordinates of the 3D face with respect to the camera coordinates. Alongside the UV position map, we do texture mapping to get the paired texture map.

The makeup-transferred texture T_{ms} is formed by combining the outputs of those branches, and this UV texture map is converted to the image space to obtain the final output I_{ms}. The texture maps are passed to two parallel branches for makeup transfer.

A. Colour Transfer Branch

A generator G is designed with two inputs and two outputs. To be specific, the network contains two separate input branches with convolutions, respectively. The two branches are then concatenated together and fed them into several residual blocks. After that, the output feature maps will be upsampled by two individual branches of transposed convolutions to generate two result images. Train a color-based makeup swapping network C that swaps makeup color on cosmetic regions between the source and the reference image with a loss function The loss function is a weighted sum of adversarial loss (L_{adv}), cycle consistency loss (L_{cyc}), perceptual loss (L_{per}) and histogram matching loss (L_{his}).

$$L = L_{adv} + L_{cyc} + L_{per} + L_{his} \tag{1}$$

B. Pattern Transfer Branch

Other than the Color Transfer Branch, we propose a clever example Transfer Branch expecting to recognize and move the example based cosmetics parts like stickers, facial drawings, and enhancing adornments.

The input texture map $T^{m}{}_{r}$ is considered to extract a binary segmentation mask for its makeup patterns. We used Pytorch pre trained segmentation models to detect stickers and patterns on the face.

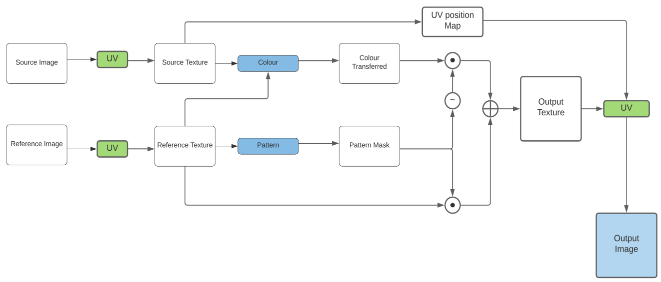

Fig. 1. Architecture of the model

We transformed both the branches into one to create a complete makeup transfer model. To do that we took the texture mapping from colour transfer branch and on those mapping we implemented pattern transfer branch.

V. Results and Analysis

First Progress: We implemented the colour transfer branch.

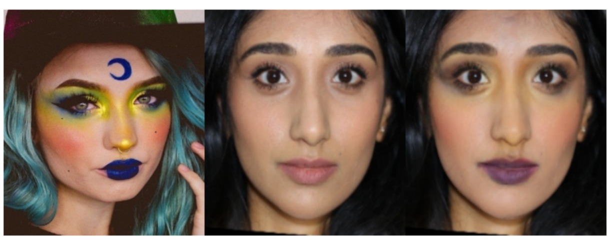

Fig. 2. Makeup Colour Transfer

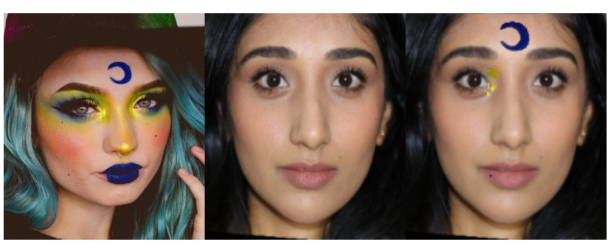

Fig. 3. Pattern Transfer

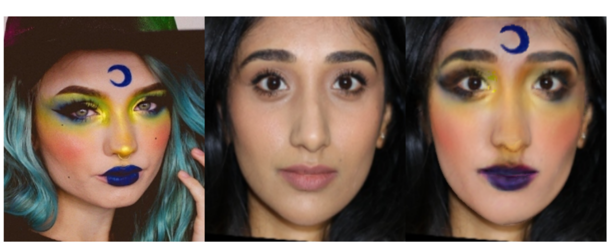

Fig. 4. Combined Branch: Colour + Pattern


```
{
  "location": {
    "left": 48.53,
    "top": 10.25,
    "width": 320,
    "height": 315,
    "rotation": 0
  },
  "face-probability": 1,
  "angle": {
    "yaw": -3.15,
    "pitch": 9.46,
    "roll": -1.07
  },
  "age": 22,
  "beauty": 0.8754,
  "expression": {
    "type": "none",
    "probability": 1
  },
  "face-shape": {
    "type": "heart",
    "probability": 0.3
  },
  "face-type": {
    "type": "human",
    "probability": 0.71
  }
}
```


Fig. 5. Beauty Score of the resultant image

TABLE I
Comparision of Beauty Scores on different datasets and models

Dataset	DMT	BGAN	LADN	PSGAN	Ours
CPM Real	.853	.847	.837	.859	.861
CPM Synt 1	.874	.876	.871	.880	.878
CPM Synt 2	.836	.838	.834	.839	.841

The result analysis shows that the model we present has little edge over rest of the models. For CPM Synt 1 dataset the PSGAN model performs slightly better but overall beauty score for our model is better.

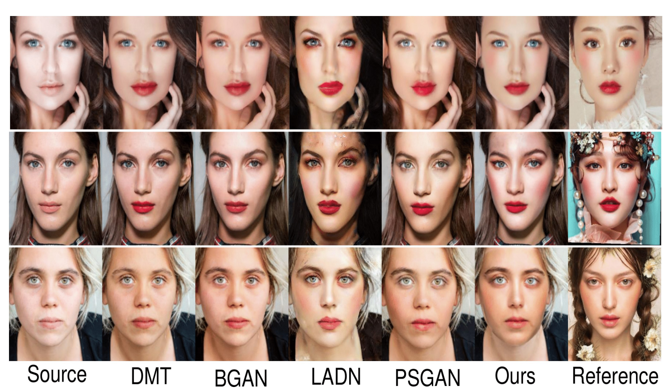

Fig. 6. Images from different models

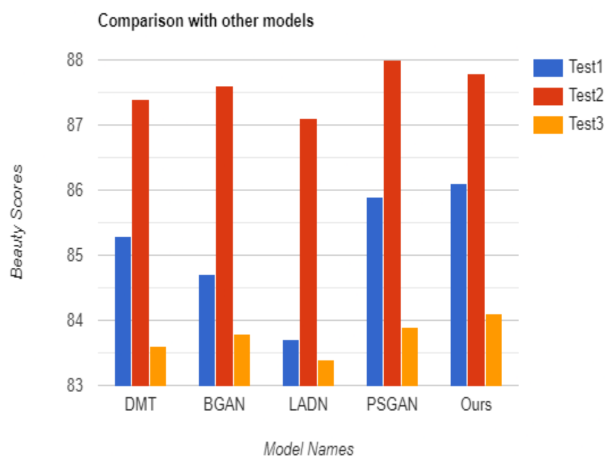

Fig. 7. Beauty Scores for different models

VI. Conclusion

Our work implemented two branches for makeup transfer: colour style transfer and pattern transfer. Most previous works have worked on colour transfer and pattern transfer separately. The images have incorporated UV representation to achieve better results in both branches. The colour transfer branch extends to different features of the face like eyes, lips, cheeks and chin. The model was implemented on new datasets. Most previous works on makeup transfer do not perform well for extreme makeup styles and do not combine patterns. We provide a combination of both the branches with state-of-the-art qualitative and quantitative results. There is no perfect evaluation criteria except for the human eye. So as a part of future work, feedback results from sample of people for each image can be incorporated. This would provide better beauty score rating for the images. Hence, concluding the best model.

References

[1] H. Chen, K. Hui, S. Wang, L. Tsao, H. Shuai, and W. Cheng, "Beautyglow: On-demand makeup transfer framework with reversible generative network," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2019, pp. 10 034–10 042.

[2] Qiao Gu, Guanzhi Wang, Mang Tik Chiu, Yu-Wing Tai, and Chi-Keung Tang. Ladn: Local adversarial disentangling network for facial makeup and de-makeup. In Proceedings of the International Conference on Computer Vision, 2019.

[3] Ren, Yeqing Sun, Youqiang Jing, Xuechun Cui, Zhihua Shi, Zhentao. (2019). Adaptive Makeup Transfer via Bat Algorithm. Mathematics. 7. 273. 10.3390/math7030273.

[4] Wentao Jiang, Si Liu, Chen Gao, Jie Cao, Ran He, Jiashi Feng, and Shuicheng Yan. Psgan: Pose and expression robust spatial-aware gan for customizable makeup transfer. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019.

[5] H. Chang, J. Lu, F. Yu and A. Finkelstein, "PairedCycleGAN: Asymmetric Style Transfer for Applying and Removing Makeup," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 40-48, doi: 10.1109/CVPR.2018.00012.

[6] Kips, Robin et al. "CA-GAN: Weakly Supervised Color Aware GAN for Controllable Makeup Transfer." ArXiv abs/2008.10298 (2020): n. pag.

[7] Zhang, Honglun et al. "Disentangled Makeup Transfer with Generative Adversarial Network." ArXiv abs/1907.01144 (2019): n. pag.

[8] Chen, Anpei et al. "Photo-Realistic Facial Details Synthesis From Single Image." 2019 IEEE/CVF International Conference on Computer Vision (ICCV) (2019): 9428-9438.