CS543 Computer Vision Final Project Proposal

Team members

Sally Rong(chenyan7),Loria Sun(jiajies3),Chaobo Cai(chaoboc2),Weixian Yi(weixian3)

Informative Project Title

Face makeup with CycleGAN and CV methods

Project Description

Image fusion is a classic topic in the CV field. Both the traditional computer vision work and modern deep learning frameworks have been explored it, In our project, we will perform the whole pipeline of face makeup, which is a popular application of image fusion.

Considering the faces dataset as our main resource, we will start with two faces – one without makeup and one with makeup. We address multiple techniques to perform a great fusion, which will be listed below.

1. Preprocessing
   a. Face Alignment
      Standardize the face orientation and scale to ensure that input images are aligned similarly, making the makeup application consistent.
   b. Face Detection
      Considering not all the face input will be a single headshot, we will apply face detection techniques to extract the main face features. We will try out the active shape model and also methods provided in OpenCV(Build Face Detection with Python using OpenCV (With link to the code) (youtube.com))
   c. Face Segmentation
      We use facial landmarks to segment areas such as eyes, lips, and cheeks.
2. Intrinsic Image Decomposition
   Intrinsic Image Decomposition is the process of separating an image into its formation components such as reflectance (albedo) and shading (illumination).
   We apply intrinsic image decomposition to isolate facial texture from this invariant characteristic, which is important when makeup is applied later.
   We will try the intrinsic method proposed by Bell(http://opensurfaces.cs.cornell.edu/intrinsic/).

3. Make up transfer

The GAN(Generative Adversarial Networks) model has been quite powerful in the image-to-image translation field. Among them, we learn CycleGAN (Cycle-Consistent Generative Adversarial Networks), which stands out when paired training data is unavailable. It is one of the most common frameworks in today's fusion field.

We perform a comparison between classic methods and deep learning-based methods(CycleGAN). In the classic method, we will apply Poisson Laplace editing, weight transfer, and alpha blending based on image layers. In CycleGAN, we will apply the pre-train model with the same input. Afterward, we will evaluate the quality based on predictability, realism, and faithfulness.

4. Post-Processing
    a. Shape from shading
    With shading information extracted from intrinsic image decomposition, we will try out shapes from shading in the face geometry reconstruction phase, particularly in enhancing the detail map. We will apply it in light after makeup transfer, trying to simulate the result as realistic.
    b. Blend/Smoothing/Adjust lighting
    We will apply these basic techniques to make the result more natural.

Member roles

Data Preprocessing

Assigned to: Sally, Weixian

- Main Task: Prepare the YMU datasets. Align and segment face images for consistent inputs across all stages. Use traditional CV techniques for face detection such as the Haar Cascade Classifier. Ensure accurate face recognition, alignment, and segmentation to standardize the input images for the CycleGAN model. Perform intrinsic image decomposition to separate reflectance (R) and shading (S) for each image.

CycleGAN implementation and optimization

Assigned to: Sally, Loria, Chaobo, Weixian

- Main task: Implement the CycleGAN model based on the official code repository, focusing on makeup style transfer between the no-makeup and makeup domains. Refine it for virtual makeup transfer on 2D reflectance images. Ensure

CycleGAN does not interfere with any 3D information. Focus on optimizing the model for realistic makeup styles.

Shape from Shading(SfS) implementation and 3D shape estimation
Assigned to: Chaobo, Weixian, Sally

- Implement shape from shading to extract 3D shape and surface normals from the shading (S) component. After makeup transfer via CycleGAN, use the 3D shape data to adjust and fuse makeup on the face, ensuring it respects the facial contours and lighting. Ensure 3D information is accurately captured.

Makeup fusion and postprocessing
Assigned to: Loria,

- Makeup fusion:
    - Blend Reflectance and Shading: Merge the reflectance map with transferred makeup (R') and the shading map (S), ensuring makeup naturally follows the face's contours using 3D shape data from Shape from Shading (SfS).
    - Lighting Adjustments: Ensure that the makeup blends with the existing lighting and shadows, adjusting highlights and shadows on the face to reflect the 3D structure.

Model testing and result visualization
Assigned to: Chaobo, Loria, Weixian

- Main Task: Test the trained model, demonstrate the effects of makeup styles and face fusion, and generate visual results, including images and videos, showcasing style transfer outcomes.

Resources

As a deep learning model, CycleGAN consumes many GPU resources. As evaluated, we are planning to use Google Colab as our working platform. If more resources are required, we will rent virtual clouds or using school resources(https://engrit.illinois.edu/services/research-services/campus-research-computing-options). While we will try to refrain from fine-tuning tasks and focus more on applications with pre-trained models.

We will use code from the official repository(https://github.com/junyanz/CycleGAN). We plan on reimplementing the code and refine upon that.

We are comparing traditional face makeup transfer techniques, as presented in the paper (https://web.stanford.edu/class/ee368/Project_Autumn_1516/Reports/Oo.pdf), with the CycleGAN approach.

For the data perspective, these are the datasets we so far find useful for our tasks and may be applied:

Large-scale CelebFaces Attributes(CelebA) Dataset*(Liu, Z., Luo, P., Wang, X., & Tang, X. (2015). Deep learning face attributes in the wild. CelebA Dataset. MMLab, Chinese University of Hong Kong. http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html).* This dataset contains 202,559 celebrity faces with 40 facial attributes, which contains adequate make-up attributes.

Youtube Makeup(YMU) Dataset*(Guo, Y., Lei, Z., Wan, J., Wang, S. Z., & Li, S. Z. (2014). Face authentication with makeup changes. YouTube Makeup (YMU) Dataset.https://iprobe.cse.msu.edu/dataset_detail.php?id=3&?title=YouTube_Makeup_Dataset_(YMU)).*

Reservations

During the implementation of our project on face makeup through image fusion, several potential challenges and considerations need to be addressed:

1. Accuracy of Preprocessing Steps: Precise face detection, alignment, and segmentation are critical. Errors in these steps, especially with non-standard headshots or images containing multiple faces, can significantly impact the quality and realism of the makeup application.

2. Complexity of Intrinsic Image Decomposition: Separating images into reflectance and shading components is complex and may not always produce accurate results. Inaccuracies here can hinder the isolation of facial textures, affecting subsequent makeup application stages.

3. Challenges with Makeup Transfer Methods: Both classic methods (like Poisson Laplace editing and alpha blending) and deep learning approaches (such as CycleGAN) require careful parameter tuning and substantial computational resources. Pre-trained models may not perform optimally without fine-tuning to our specific dataset.

With these considerations, our goal is to implement a comprehensive face makeup pipeline that applies basic makeup styles to a diverse set of face images using both classic and deep learning techniques. We will continuously refine each component to improve the overall quality, realism, and robustness of the results

Relationship to our background

Our project is closely related to the class. While none of the four members of the team have experience in computer vision, this project provides us with an opportunity to build foundational skills in the field. We are familiarizing ourselves with core concepts such as image processing, Generative Adversarial Networks, and unsupervised learning. We have experience with Python, deep learning frameworks (e.g. PyTorch, TensorFlow), and related libraries, but applying these skills to CV tasks like CycleGAN is a new challenge that fits well with our course goals. Our goal is to deepen our understanding of CV methodology through hands-on exploration while moving beyond our comfort zone by incorporating advanced techniques like StyleGAN-based image restoration.