Replication research regarding the HairMapper Algorithm

Maud Groen and **Jeroen Wasser** and **Mathieu Janssen**

Abstract

The *HairMapper: Removing Hair from Portraits Using GANs* paper proposes a solution that would easy the manipulation of hairstyles on 3d modelled faces, by removing existing hair completely. Creating a blank canvas to operate on for modellers, the HairMapper algorithm attempts to remove existing hair and hair-artifacts. We attempted to replicate their method using the literal replication method, and compared their results with other models. The code provided on their GitHub did not function as explained in the ReadMe provided by the authors. As such, we could not effectively replicate their results as provided due to the inadequate explanation and limited resources available to us. In spite of these limitations, we concluded that the model does not generalize to sub-optimal portrait conditions.

1 Introduction

Generative Adversarial Networks, or GANs for short, have been used in multiple fields of research and have been gaining popularity over the years. GANs consist of two neural networks that work together, namely the generator and the discriminator. Because of this approach, there are many applications in image manipulation, such as in image-to-image translation.

In their paper on HairMapping, (Wu, Yang, & Jin, 2022) develop an algorithm using these GANs to remove hair from portraits. Removing and changing hair from portraits has been researched before, however the research has not been successful yet. This is because there are many challenges that could easily cause the algorithm to not work properly. For example, a direct overlay of the new hair can cause problems due to it being mixed up with the old hair (Wu et al., 2022).

Many different approaches have already been tried using StyleGAN (Tero Karras, 2019), however none of these methods can remove hair while also preserving facial structure. This is why (Wu et al., 2022) took a new approach, where they design two pipelines ”to generate paired latent codes with/without hair for males and females while keeping their facial identities”. With this data they train the HairMapper network after which the final result is obtained by exploiting Poisson editing to blend the mapped portrait with the original portrait.

In this replication paper, we will attempt to reproduce the results found in (Wu et al., 2022). More specifically, we will test their training code as well as their final HairMapper model to see whether the code is reproducible. We find that the results in their paper are not reproducible. Because of lacking computing power, it is not possible to replicate the training process. The final HairMapper model does work, however there are many limitations that will be discussed further.

2 Background

The development of Generative Adversarial Networks (GANs) started in (Ian Goodfellow, 2014). In essence, it works by training two neural networks against each other in a zero-sum game. In most cases, one network is a generator and the other is the discriminator. The images created by the generator are compared to images of actual objects by the discriminator, to see if it can tell which is real and which is not. Therefore, the generator is not necessarily being trained to replicate specific images as closely as possible, but rather it is trained to generate images that ”fool” the discriminator.

As with many applications of GANs, the input images are encoded to a latent space vector. In general terms, a latent space is a space of vectors in which object that appear ”similar” are close together (Toan Pham Van, 2021). The images used as inputs for their HairMapper algorithm are first mapped to their latent space vector representation. One of the main advantages of this mapping is that is compresses the information in the images. This allows the neural network to learn the distribution of this information, such that it can sample this learned distribution to create new images.

In their paper, (Wu et al., 2022) refer to having to train two separation boundaries in this latent space; one for distinguishing between bald men and men with hair, and one for distinguishing between men and women. The male-female separator is an SVM-based algorithm the researchers trained themselves. The hair-bald separator for men is based on previous research into this area (Yujun Shen, 2019).

Due to the fact that their implementation relies on the StyleGAN (Tero Karras, 2019), there are certain combinations that cannot be sampled in the latent space, one of which is *bald-female*. At first glance this would imply that the developed algorithm can only work on males, but the researchers found a workaround. Namely, the latent vector corresponding to the original image of the female is edited using the aforementioned gender separation boundary. Once a male with similar skin characteristics and facial posture has been found, this male is turned bald and the original female face is diffused back on.

3 Target Study Overview

3.1 Technique Foundations

The HairMapper: Removing Hair From Portraits Using GANs paper employs several techniques as the basis for their findings. These are:

1. The implementation of classifiers
2. The implementation of unique pipelines to generate more viable data.
3. The mapping of an image to the latent space
4. The diffusion of latent space encoded images

3.2 Technique Internals

1. Two separate distinct pipelines are used for male and female faces due to the lack of paired bald and non bald female faces available as training data.
2. Hair separation boundaries are used to remove hair and generate bald faces for training purposes.
3. Paired latent codes are used to train HairMappers.
4. Female faces are first transformed to male faces with similar skin characteristics and face posture. These faces are then made bald using diffusion and afterwards transformed back to female faces but this time without hair.
5. The original image with hair and the generated bald image are diffused together to create a bald face with the same characteristics as the face in the original image.
6. Parts of the images are dilated and blurred to better hide the mask applied to the original image (with hair)

The CelebAHQ-mask dataset is leveraged by the authors to train their hair classifier. They made use of 602 bald images found in this dataset, they supplemented this dataset with 184 bald images found on the internet. ResNet50 was leveraged to train the hair classifier.

For the diffusion the Adam solver was used as the optimizer with a learning rate of 0.01.

The authors only made changes to the first 8 layers of the latent code to make sure that the fine features of a face are unchanged.

3.3 Study context and targeted domain

The results of this study are applicable to the 3D face reconstruction domain. Replacing old hair on reconstructed faces is error prone due to the mixing up with the old hair. By removing the old hair before applying new hair this problem is removed.

3.4 Validation Approach

The method proposed in the original paper was tested on various images collected on the internet. These images varied in age, ethnicity, gender group and lightning conditions. Pseudo ground-truth experiments where held as well where hair was manually added to already bald portraits. The method was then applied to these images to see if the original bald portrait was generated afterwards.

Additionally a user study was held where 40 participants were shown 20 images which were randomly chosen from 40 real bald portraits and 40 randomly sampled results which had their hair removed. They were then asked to judge whether images were untouched or not to see if they could distinguish real bald portraits from generated ones.

3.5 Study limitations and replication problem statement

Some limitations were mentioned within the original paper. Their method relies on the e4e encode to embed real images. Images that cannot be precisely encoded by e4e may cause artifacts to change. Additionally, hair that is not covered by the hair mask is noticeable after diffusion. Finally, extreme lighting conditions can also make the hair mask more noticeable.
Out of these findings we propose the following problem statement:

The effectiveness of the proposed solution on real-life images is questionable, as the generalizability of the data used for training is limited.

4 Replication Study Definition

4.1 Decision Tree Diagram

As this is a replication study, it's important to get an overview of what is being replicated from the original paper. The overall setup of this replication study can be found in Figure 1. The corresponding data flow diagram is given in Figure 2. The replication method chosen is that of the *Literal Replication*. That is, we intended to follow the original authors' process as closely as possible. We wanted to test the true effectiveness of their proposed solution, and felt this could be best achieved by following their own process. However, due to limitations in computing resources, we did not manage to do the data preparation and training phases.

4.2 Results

Since the original papers evaluates their network by comparing it to different networks, we intended to do the same for this replication study. However, the code provided in their GitHub did not work well, and the process of downloading

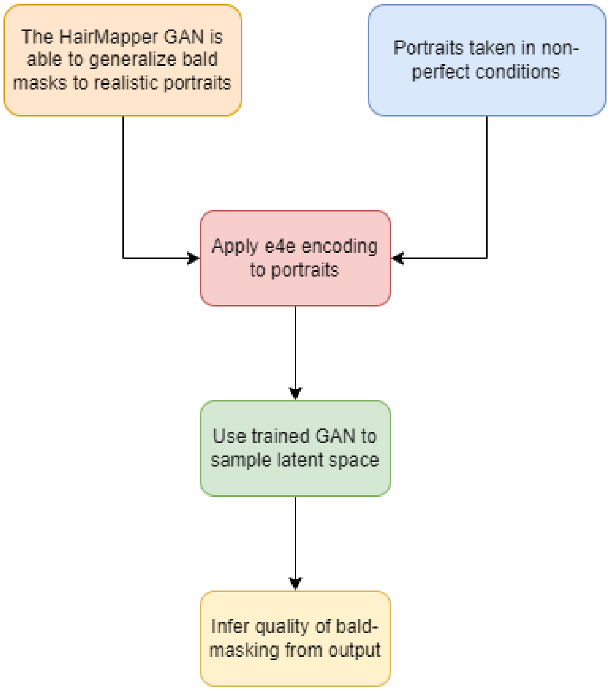

Figure 1: Decision Tree Diagram

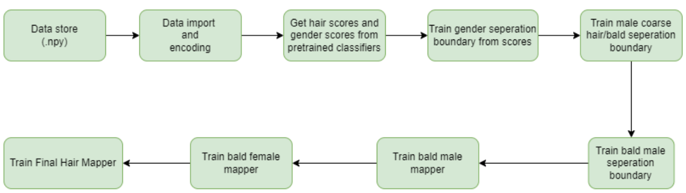

Figure 2: Data Flow diagram

and evaluating the other networks did not work as the authors describes. However, we still managed to evaluate the HairMapper GAN on images taken by ourselves, under various different conditions that one might experience in everyday life.

In Figures 3 & 4 one can see an example of a type of portrait that HairMapper cannot handle well. The conditions for this portrait are sub-optimal, especially compared to the quality of the portraits used to train the HairMapper. The face is not centered enough in the framing of the portrait, and the lighting conditions are too bright. However, it is a setting that one could expect to find in day-to-day life, and the assumption that every portrait fed to HairMapper is of excellent quality is not realistic.

4.3 Comparison to original paper

In their own paper, the portraits fed to HairMapper are ”baldified” well. They mention that, for a certain dataset, participants were not able to tell wether the portrait shown had been altered or not. They found that in 48.37% of cases, people labeled portraits altered by HairMapper as

Figure 3: Portrait of one of the authors

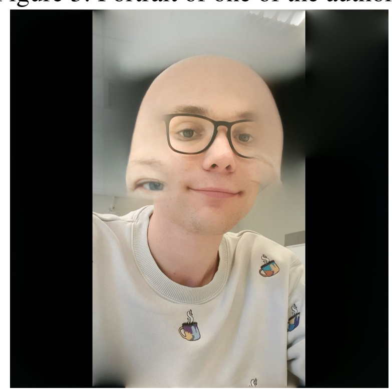

Figure 4: Portrait after having been passed through HairMapper

untouched, while the number for actual bald portraits was 59.74%. This implies that, for data of appropriate quality, the HairMapper model actually works well. However, if these conditions are not met strictly, the portraits generated by HairMapper are exposed quite readily.

Figures 3 & 4 show the lack of generalizability of HairMapper. These are not the only test portraits that were used, but they show the overall shortcomings of the HairMapper model. The model seems to have a hard time with sub-optimal lighting conditions (too bright or too dark), the face not being centered and prominent in the image and it also seems to struggle with glasses.

In their paper, they mention that the bald mask might not always work as intended. However, we also notice the introduction of certain artifacts in the generated portraits. For instance, in Figure 4 one can see that a new eye is placed on the right side of the head. This is of course not correct, and similar artifacts are found in other portraits put through the model.

5 Conclusions

The HairMapper technique discussed in the previous sections, has two applications, namely in digital hair design and 3D face reconstruction. The removing of hair will make it easier to show new hair designs. When the old hair is removed from an image, the new hair templates can be added directly onto that image. The same principle holds for 3D face reconstruction, when the hair is removed from the portrait, a "well-reconstructed face with limited geometry change and no hair artifacts" (Wu et al., 2022) is left.

A big addition made by this research, is the new dataset they created containing 6,000 non-hair portraits. One of the occuring problems before this paper, was the fact that there is a very limited amount of images of bald women. With this new dataset, these images are more available, which could inspire and facilitate more research in the future. This future research could focus on many different contexts, such as the digital hair designs and 3D face reconstructions mentioned before. However, other contexts to be considered could be beauty and fashion, medical imaging and forensics. For example, in forensic investigations the facial recognition algorithms might work better when there is no hair present in the portrait. With this, the algorithm could focus only on the structure of the face. The same could be true for medical imaging, it could be possible that removing hair from a head scan or portrait can improve accuracy.

In this paper, the goal was to replicate the findings of the HairMapper paper (Wu et al., 2022). The source code of the algorithm consisted roughly of two parts, the training of the model, and a demonstration of the created model. Unfortunately, due to limited computational powers, it was not possible to recreate the training process of the model. Even using the GPU computational power of Google Colab was not enough to run the algorithm. Because of that, the focus was on recreating the results using the demonstration. After uploading multiple portrait images, the conclusion was drawn that the algorithm only works for portraits with a limited background, or where the camera zoomed in so far that there was almost no background to detect. When there was some background and lighting in the picture, the model was not able to correctly process long hair and glasses. It caused the bald portrait (the portrait without hair but a different facial identity) to show through the original image. For example, the glasses on the original image were replaced with the glasses visible on this bald portrait. Overall, our analysis suggests that the generalizability of the model may not be as extensive as reported by the authors.

Due to limited computational resources, our attempts to replicate the results reported in the paper were limited. By conducting the demonstration of the HairMapper technique, we were able to produce a literal replication. The constraints posed by limited computational resources restrict our ability to modify either the dataset or the tooling, thereby rendering operational and instrumental replications infeasible. A constructive replication might be achievable, since the original procedure and research design should be disregarded. The new, alternative procedure should be chosen in such a way that it can be executed within the bounds of our computational resources.

Finally, we believe that our findings of limited generalizability might be due to the dataset used. The data used for training this model, came from the CelebAHQ-mask dataset, which contained 602 bald portraits. The complete dataset contains many images with different hairstyles, facial structures and skin colors. To these bald images, the researchers added 184 images of bald people manually, by collecting them from the internet. However, the amount of images of bald women is very limited, which is why the researchers create a female-to-male pipeline to be able to use the HairMapper technique on women. Moreover, these images were all selected based on their high quality. Because of this, we believe that the data might still be too limited and the model may not generalize well to other datasets or real-world scenarios.

References

Ian Goodfellow, M. M. B. X., Jean Pouget-Abadi. (2014, June). Generative adversarial nets.

Tero Karras, T. A., Samuli Laine. (2019, March). A style-based generator architecture for generative adversarial networks.

Toan Pham Van, N. N. T. H. V. N. L. B. D. H. Q. D. T. T. M., Tam Minh Nguyen. (2021, February). Interpreting the latent space of generative adversarial networks using supervised learning.

Wu, Y., Yang, Y.-L., & Jin, X. (2022, June). Hairmapper: Removing hair from portraits using gans. , 4227-4236.

Yujun Shen, X. T. B. Z., Jinjin Gu. (2019, July). Interpreting the latent space of gans for semantic face editing.