Applications of Generative Adversarial Networks in Hairstyle Transfer

Student
Ta Dang Khoa
Pham Cao Bang

Supervisor
Dr.Phan Duy Hung

Bachelor Of Computer Science
Fpt University - Hoa Lac Campus

Contents

1.Introduction

Problem and motivation

Contribution

Problem and motivation

- The success of StyleGAN, where stochastic variation is incorporated in the realistic looking synthesized images

- The notion brought up about the existence of a hyperplane in the latent space serving as the separation boundary for any binary semantic

Contribution

Semantic editing latent code

- We take advantages of InterFaceGan assumption about the existence of hyperplane in the latent space serving as the separation boundary for any binary semantic to develop system that can change single hair attribute

Contribution

Overwrite hair image over face image:

- We propose a method, which is possible to change the hairstyle according to a given hair photo, without altering other facial properties.

2. Background

Generative Adversarial Networks (GAN)

StyleGan

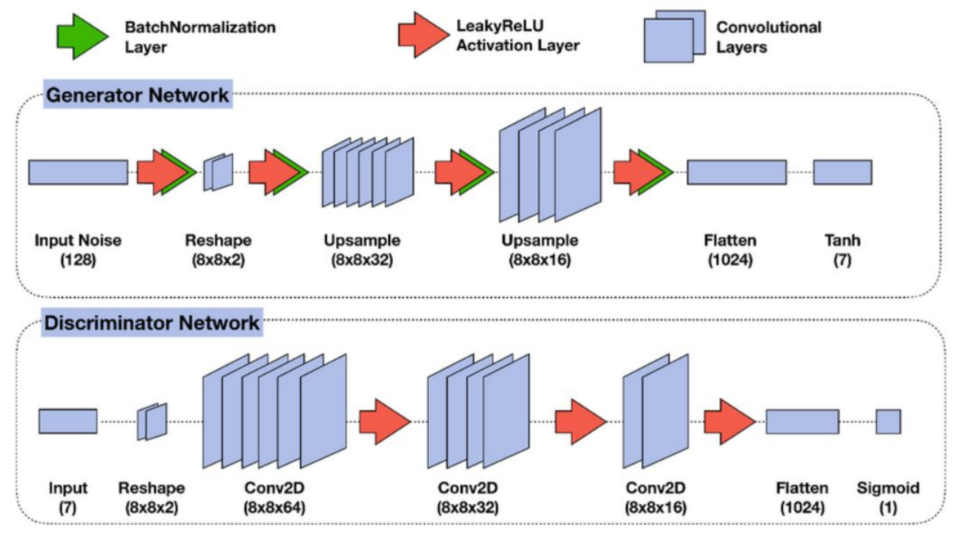
Perceptual loss

Resnet

StyleGan Encoder

InterFaceGan

Generative Adversarial Networks (GAN)

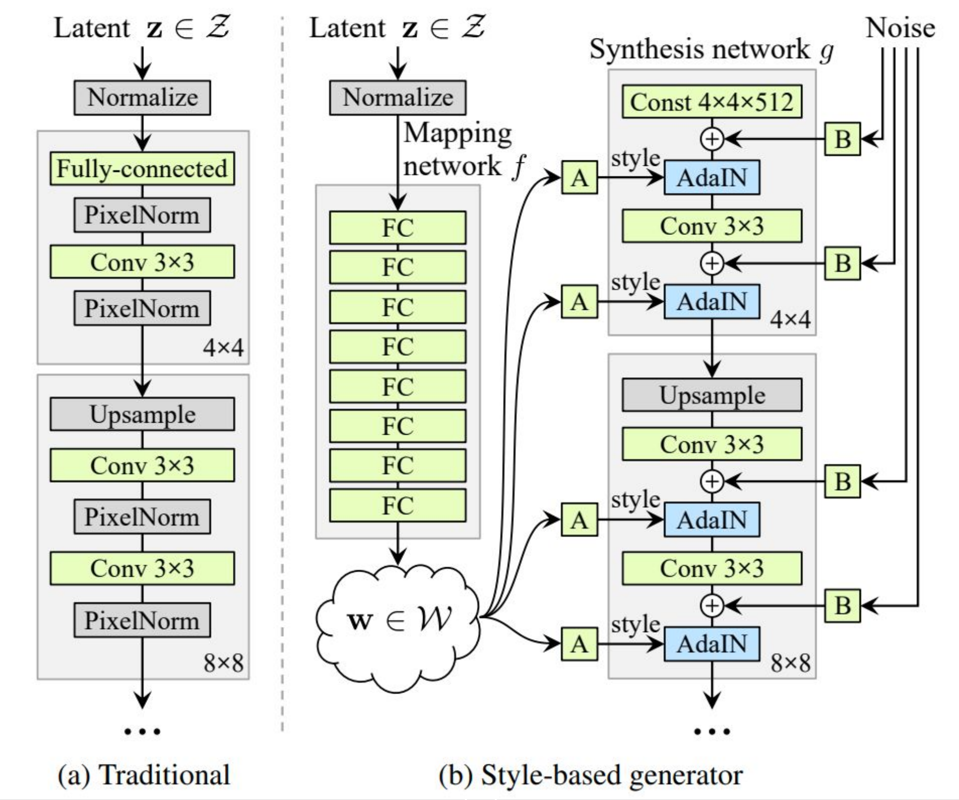


Generative Adversarial Networks (GAN)

Generator (G) and Discriminator (D) play the following two-player minimax game with value function V (G, D)

$$\min_{G} \max_{D} V(D, G) = \mathbb{E}_{x \sim p_{data}(x)}[log D(x)] + \mathbb{E}_{x \sim p_{z}(z)}[log(1 - D(G(z)))]$$

A Style-Based Generator Architecture for Generative Adversarial Networks (StyleGAN)



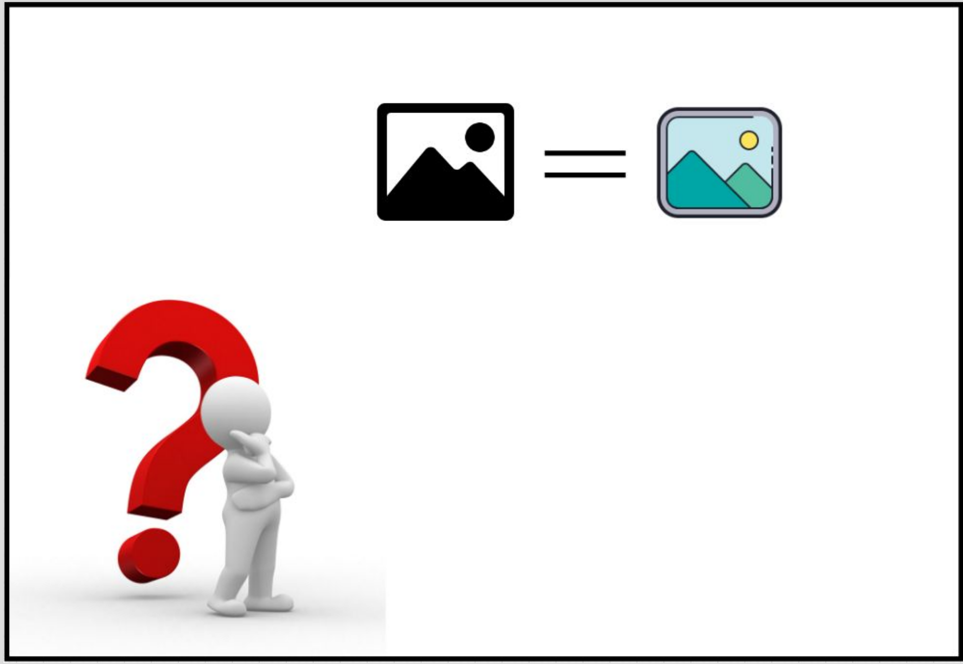
Style mixing

Stochastic variation

FID on 2 Datasets

	Method	CelebA-HQ	FFHQ
A	Baseline Progressive GAN [26]	7.79	8.04
B	+ Tuning (incl. bilinear up/down)	6.11	5.25
C	+ Add mapping and styles	5.34	4.85
D	+ Remove traditional input	5.07	4.88
E	+ Add noise inputs	**5.06**	4.42
F	+ Mixing regularization	5.17	**4.40**

Perceptual loss

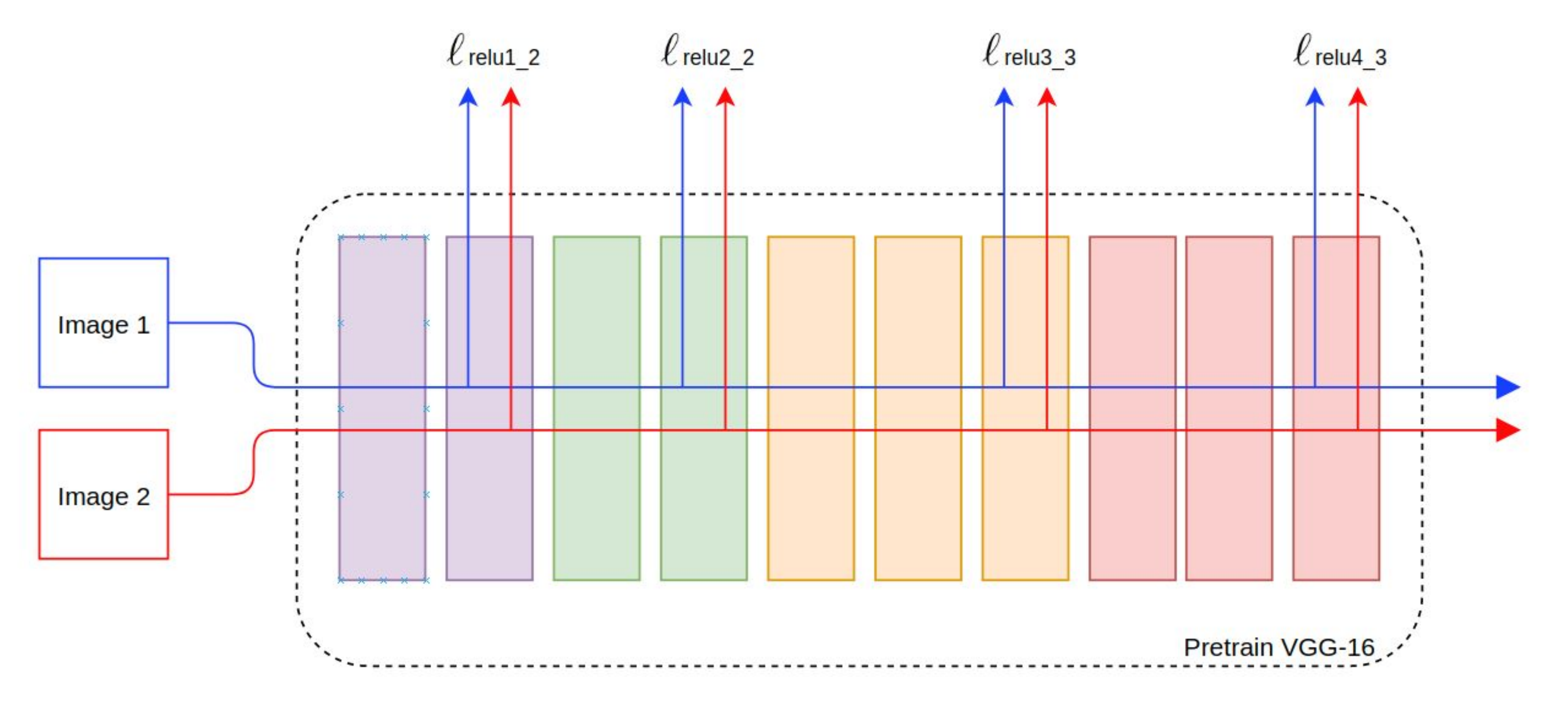


Per-pixel measures

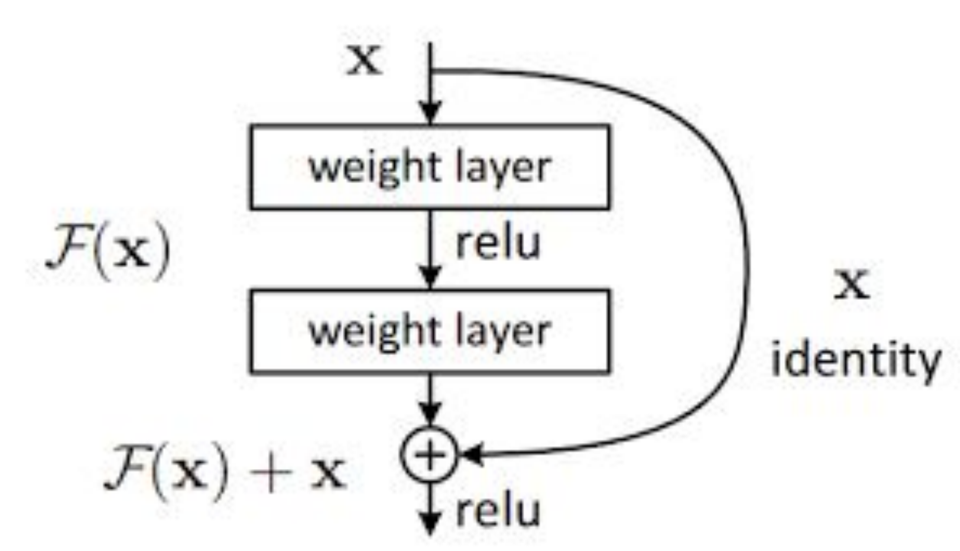
	Origin image	Blur 3x3	Blur 5x5	Blur 7x7
Image				
L2 loss	0	6000.650881362788	7041.8693540848935	7692.398455618378

Blurring images causes small perceptual but large L2 change

VGG16 weighted combination loss

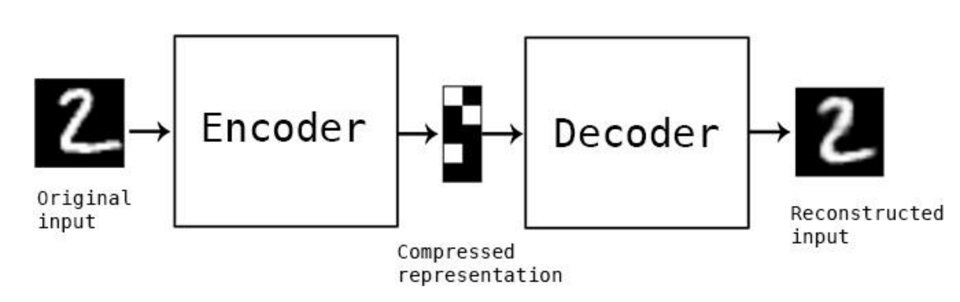


Resnet

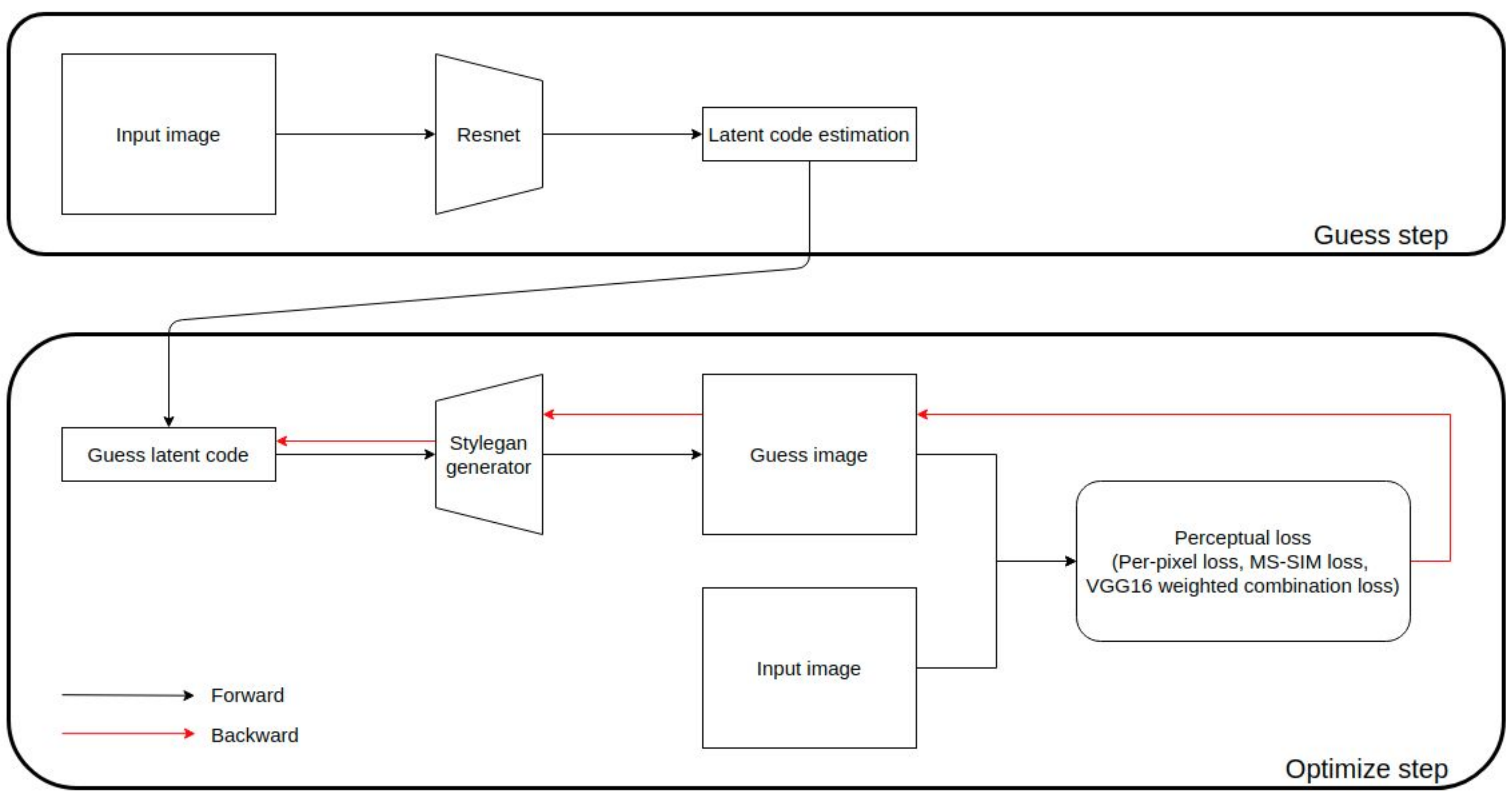


Residual block

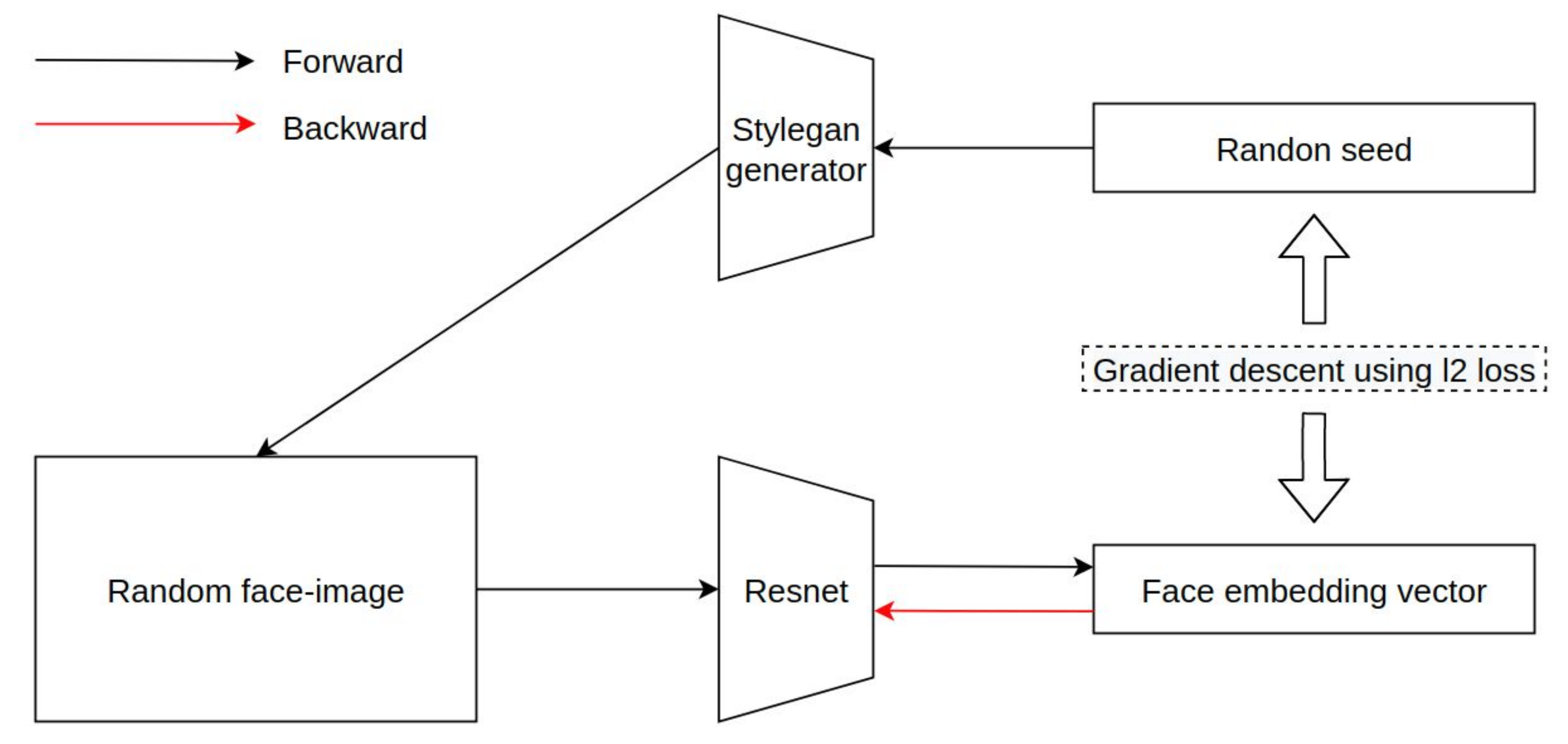
StyleGan Encoder



Encoder process



Training resnet process

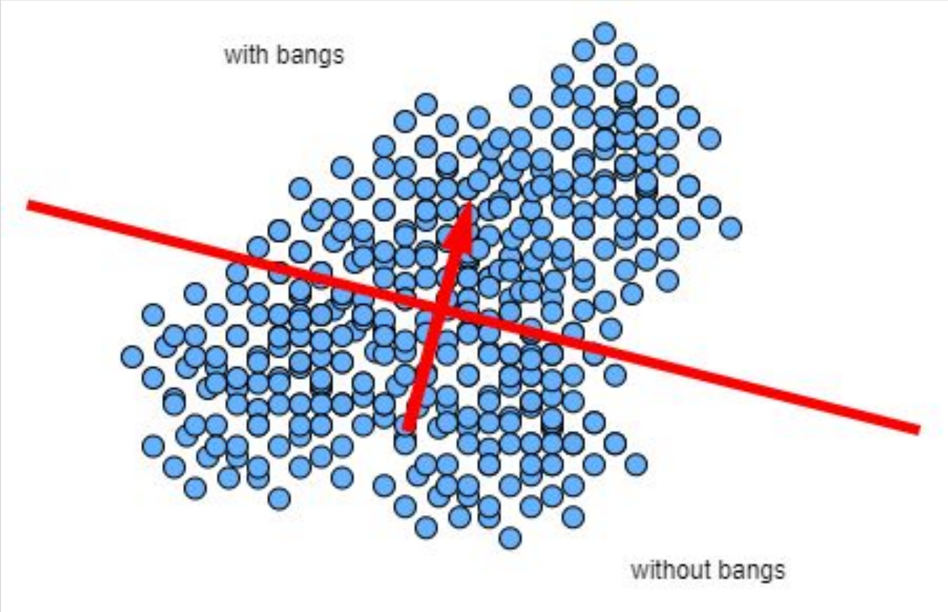


InterFaceGan

- Single attribute manipulation
- Multiple attribute manipulation
- Real image manipulation

Single attribute manipulation

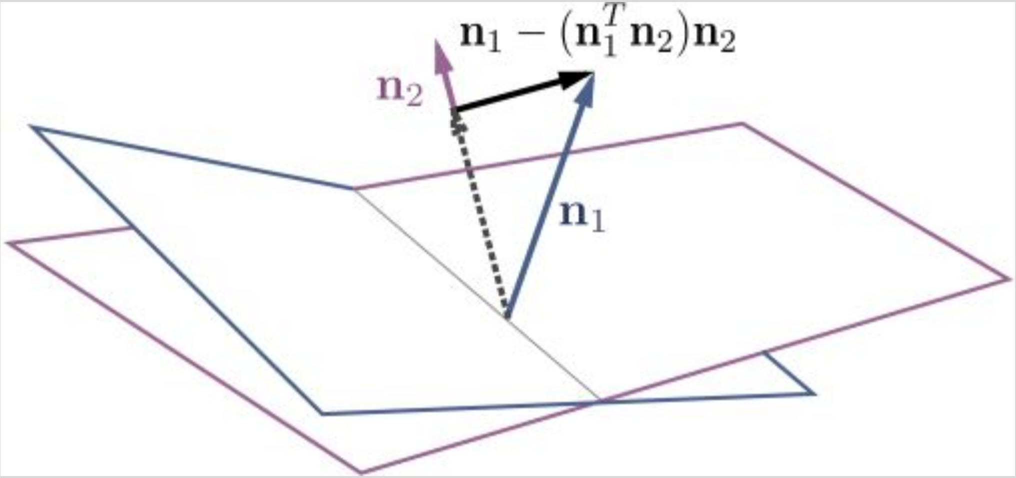
For any binary semantic such as male and female there exists a hyperplane in the latent space serving as the separation boundary



We can increase or reduce the target attribute by moving the latent vector of the image along the normal vector of separate hyperplane

Multiple attribute manipulation

When there are more than one attribute editing one may affect another. Conditional manipulation can be used to solve this problem.



➔ Instead of moving along n_1 we moving along the projected direction $\mathrm{n}_1 - (\mathrm{n}_1^T \mathrm{n}_2)\mathrm{n}_2$ which are orthogonal with n_2 , so attribute of this normal vector does not change.

Real image manipulation

The method of InterFaceGan will edit latent vector in fixed GANs latent space.

➔ We need to map our real image into GAN latent space.

For this purpose, existing methods have proposed to directly optimize the latent code to minimize the reconstruction loss, or to learn an extra encoder to invert the target image back to latent space. There are also some models that have already involved an encoder along with the training process of GANs.

3. Methodology

Semantic editing latent code

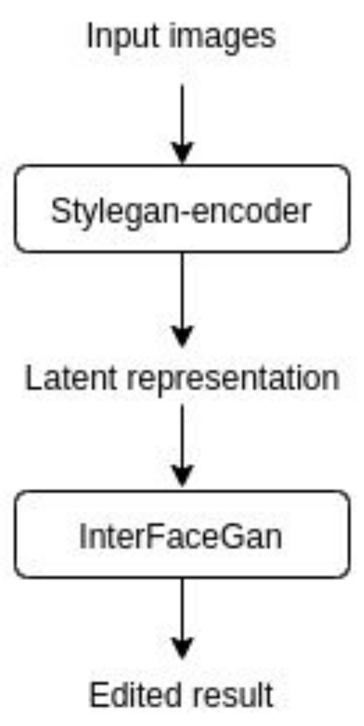
Overwrite hair image over face image

Semantic editing latent code

- Detail Implementation
- Experiment Result
- Analysis

Detail Implementation

Workflow



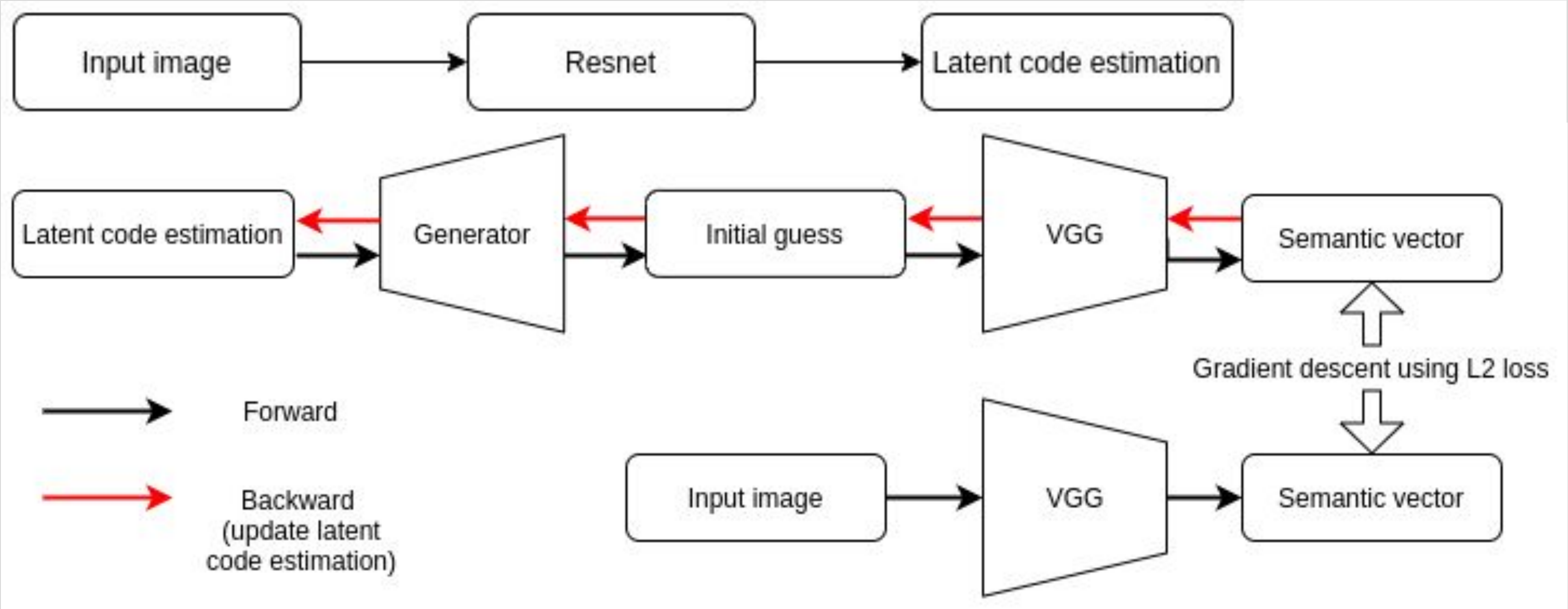
We take advantage of single attribute manipulation, which proposed in InterFaceGan, for semantic edit hair style of given image.

- Step 1: Encode image using StyleGan Encoder network, this network allow we to find latent representation of given image in GAN latent space.

- Step 2: Edit latent representation of image by a boundary (hyperplane) for target attribute

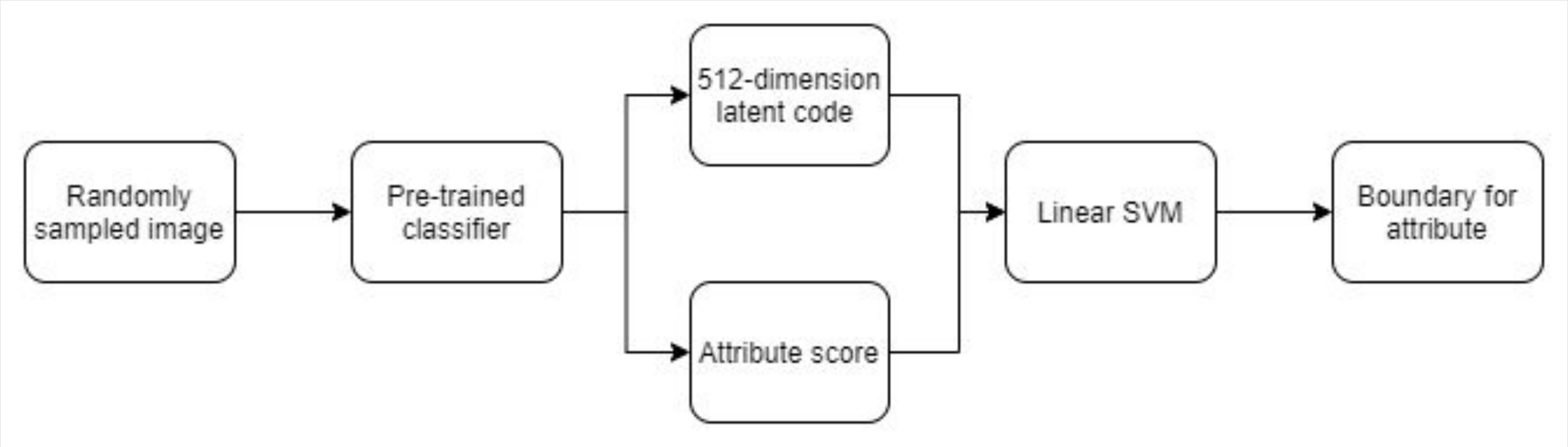
Detail Implementation

StyleGan Encoder Process



Detail Implementation

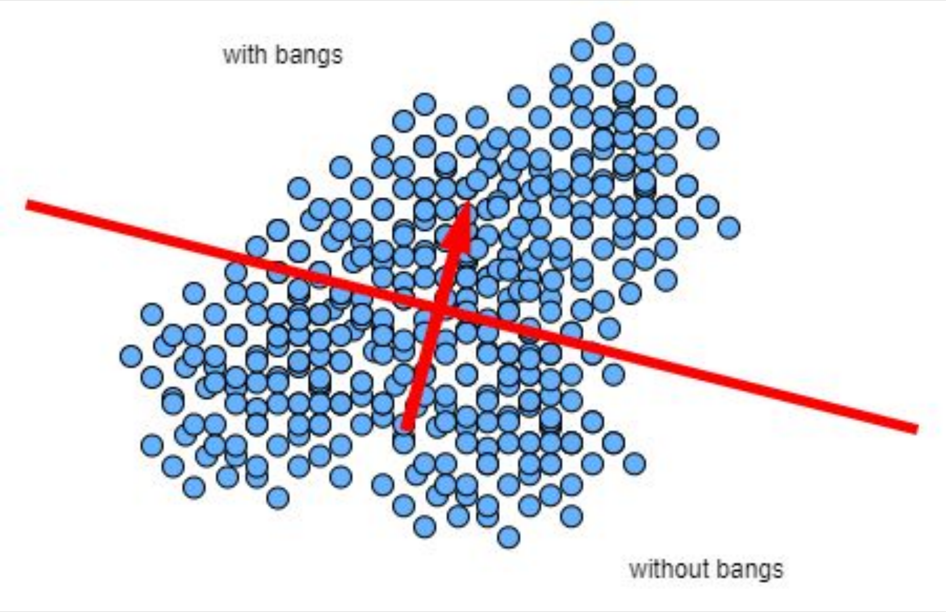
Training boundary (hyperplane) process



We random 100000 image every time we train boundary for an attribute. The pre-trained classifiers are provided by StyleGan, that are trained with a large dataset for single hair attribute classification

Detail Implementation

Editing image with InterFaceGan



➔ With latent representation of image and target attribute boundary received from 2 last steps, we modify latent code by moving along direction of normal vector of boundary.

Experiment Result

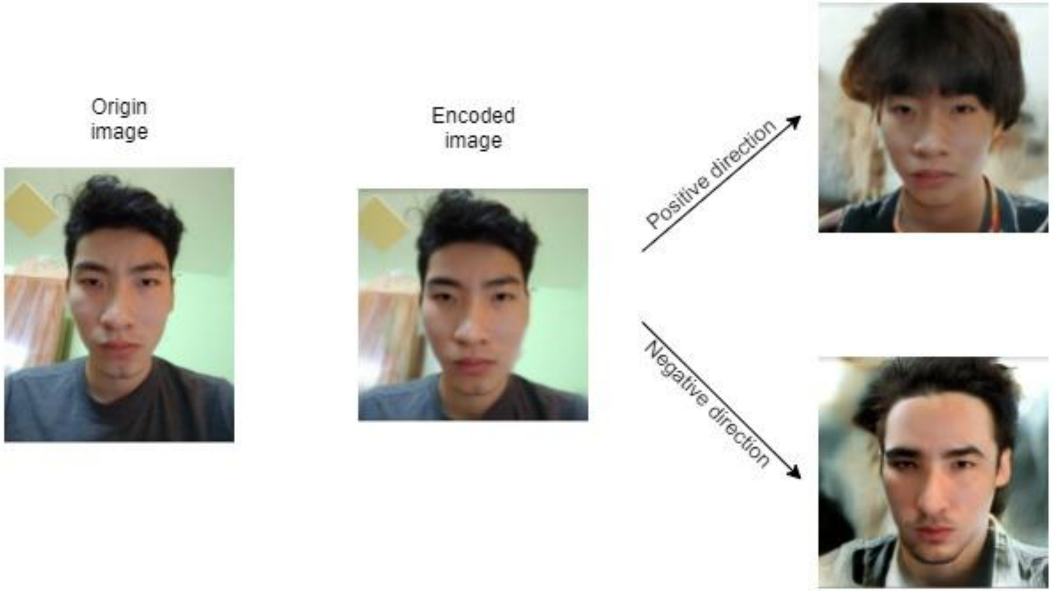
Attribute \ Direction	negative	origin	positive
bangs			
straight			
wavy			

Attribute \ Direction	negative	origin	positive
black hair			
blond hair			
bald			

Experiment Result

Changing latent vector too much regard to the target attribute (wavy hair)

Problem with InterFaceGan (bangs attribute)



Analysis

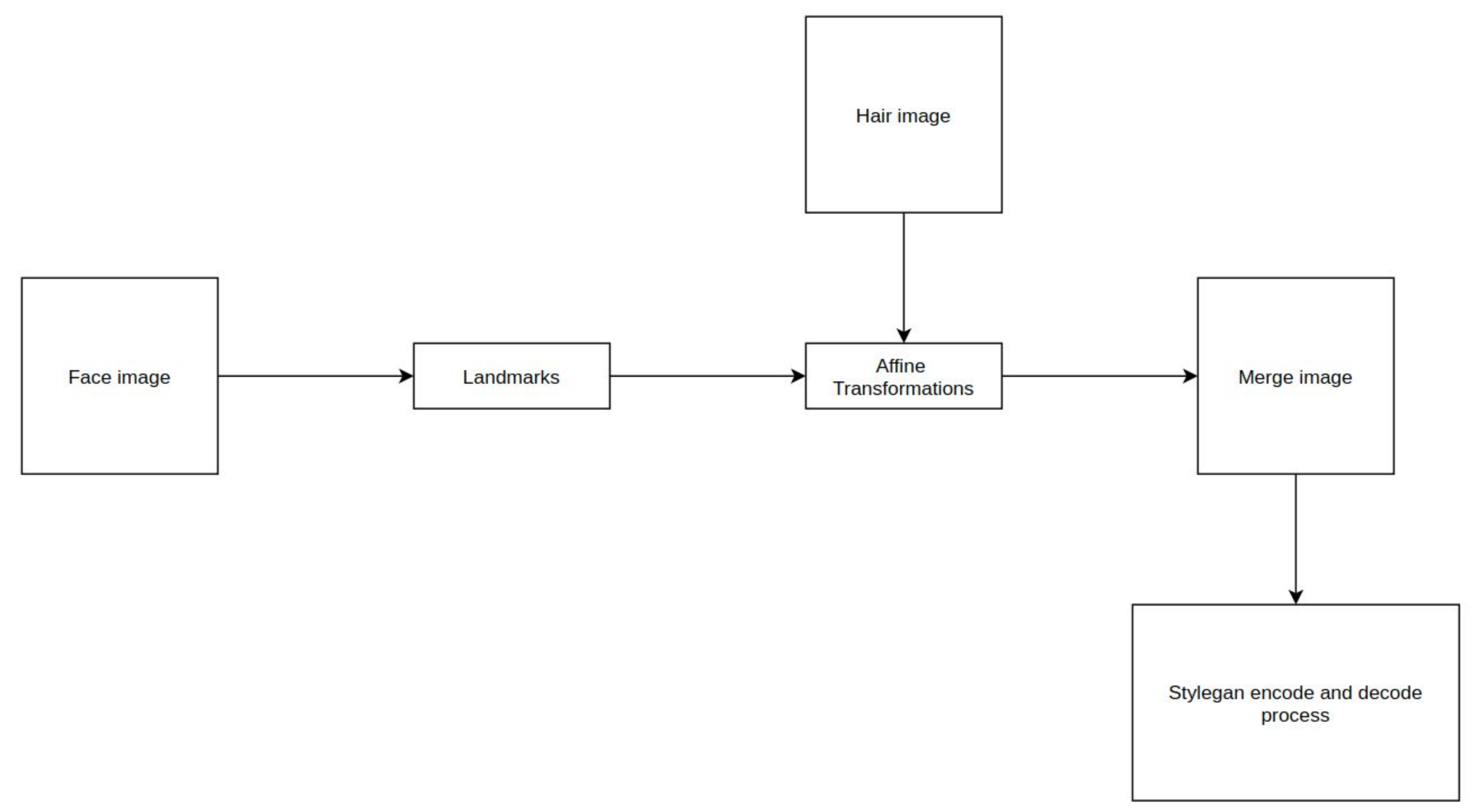
- When there are two or more attribute that correlated with each other changing one attribute may affect others. This property will be fixed with condition manipulation which proposed in InterFaceGan.

- InterFaceGan have problem when input image is Asia people. InterFaceGan change the face slightly like European people, that is problem of training dataset.

Overwrite hair image over face image

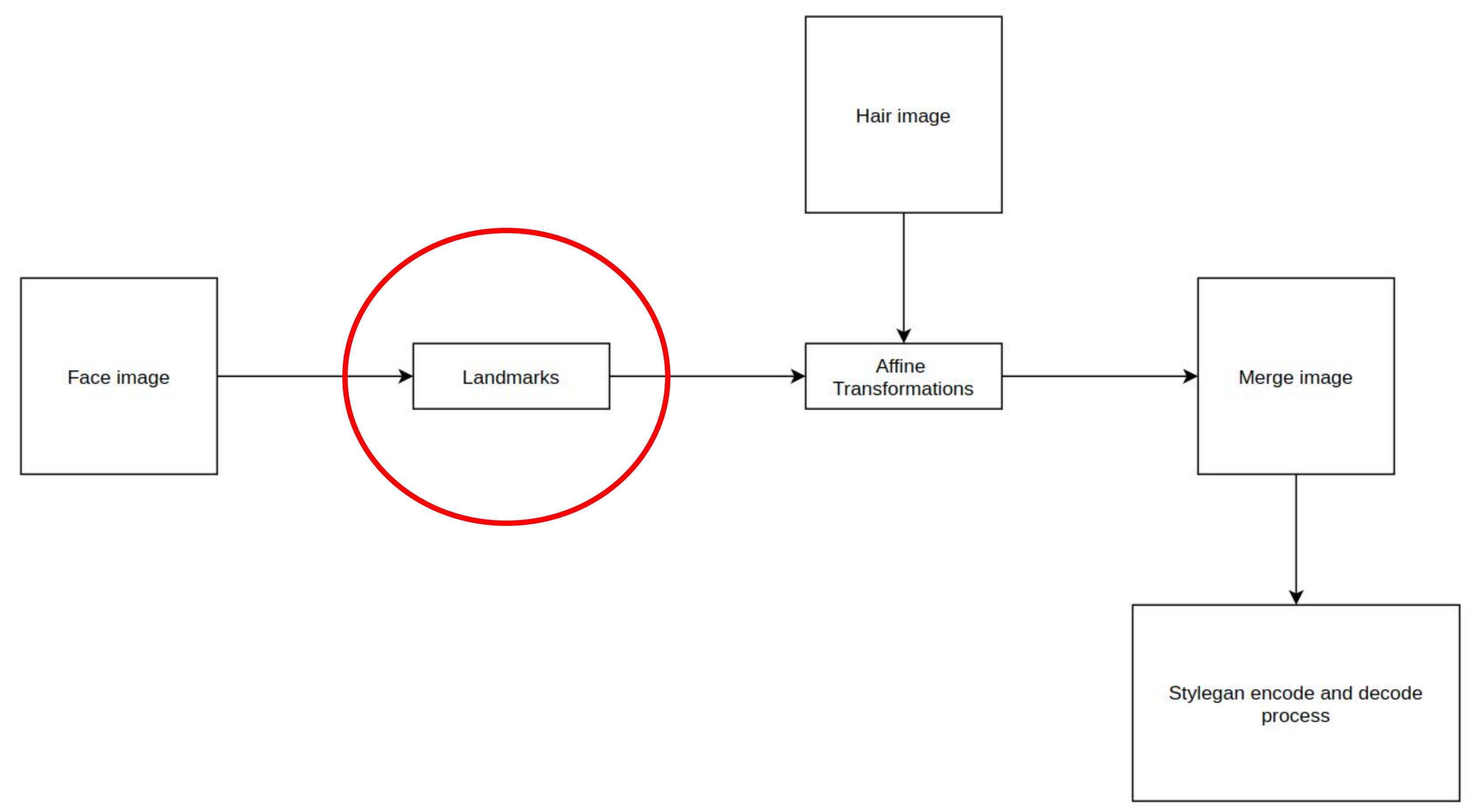
- Detail Implementation
- Experiment Result
- Analysis

Process

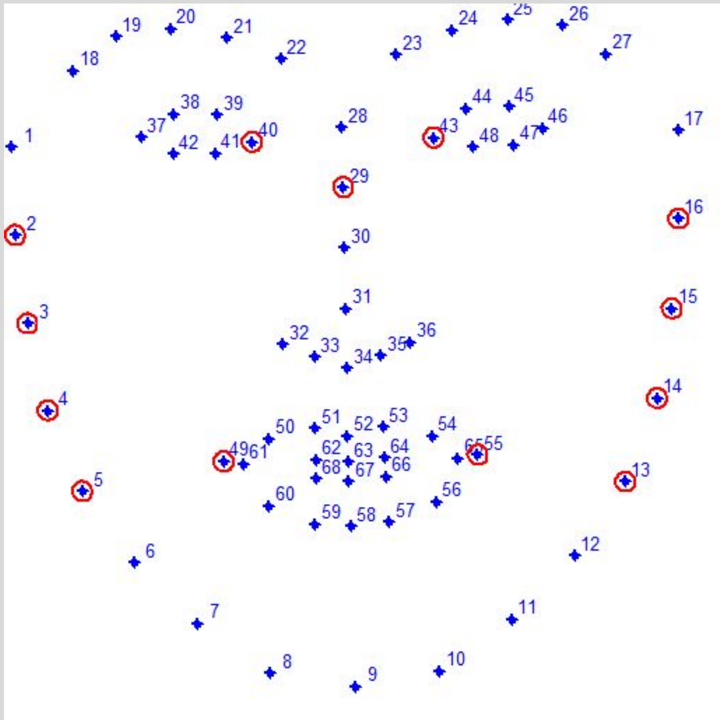


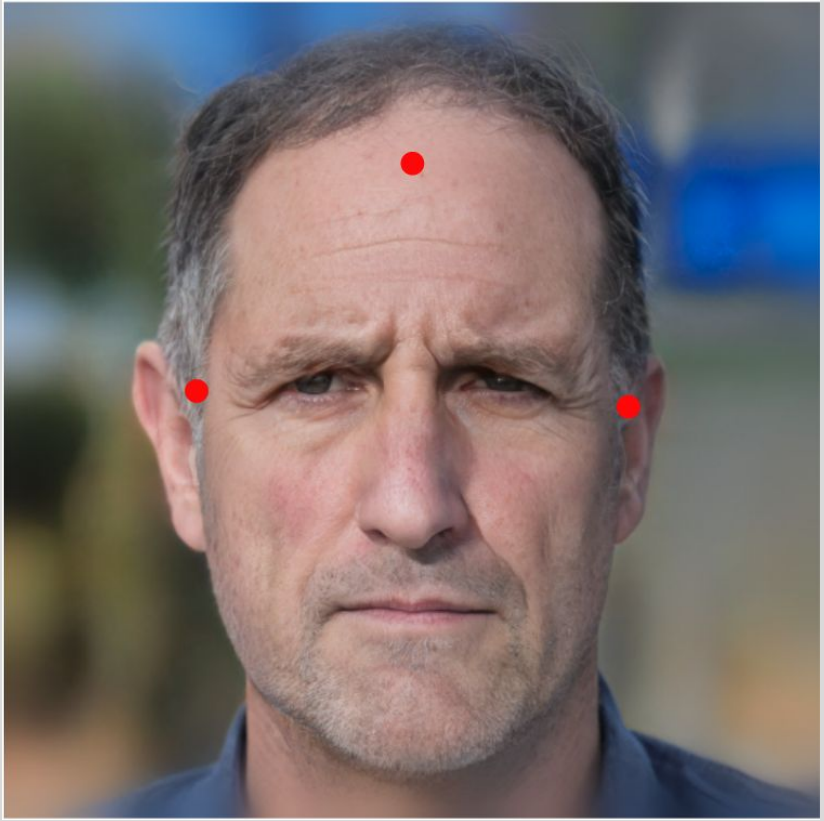
Hair image
Face image
Landmarks
Affine Transformations
Merge image
Stylegan encode and decode process

Process


Hair image
Face image
Landmarks
Affine Transformations
Merge image
Stylegan encode and decode process

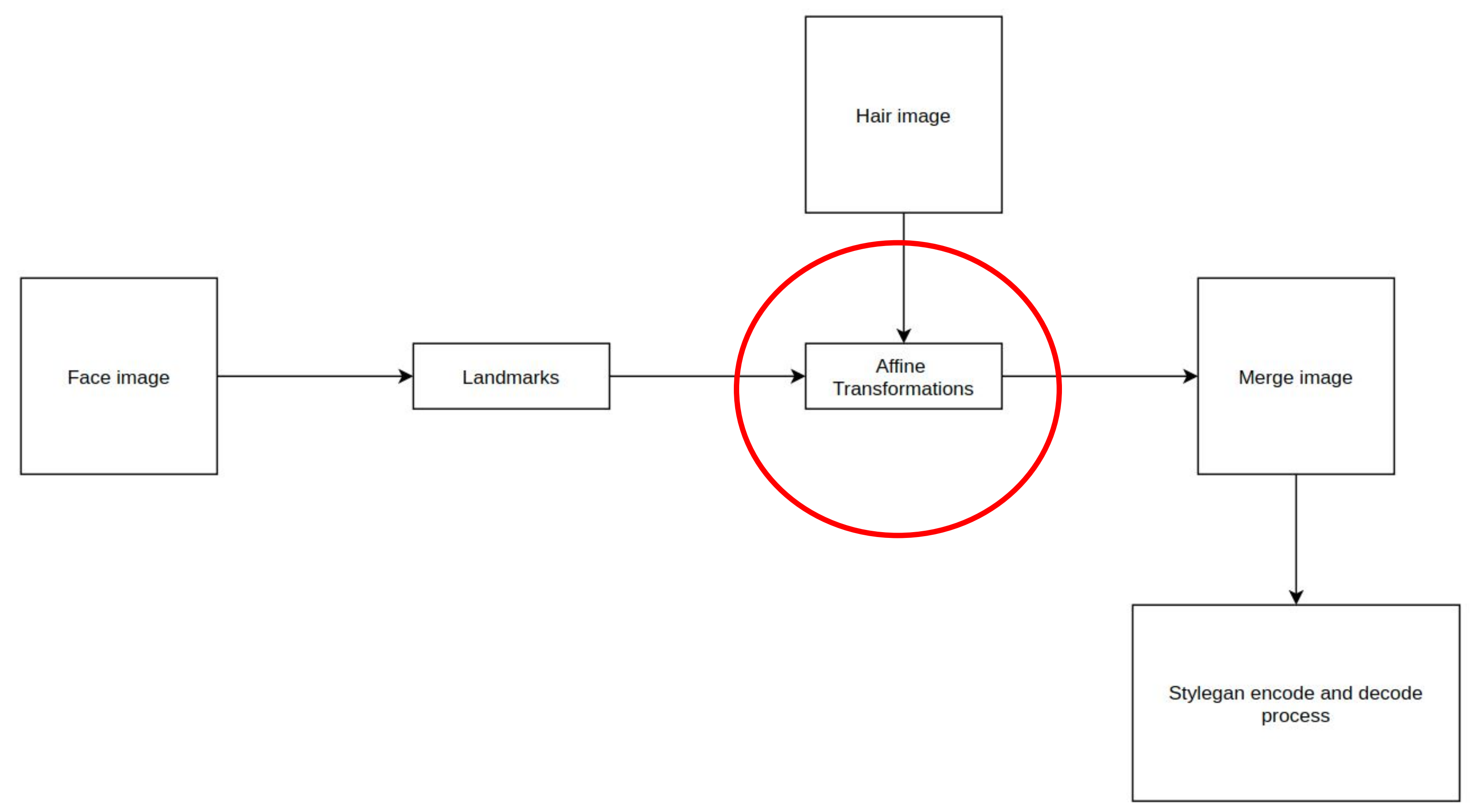
Landmarks



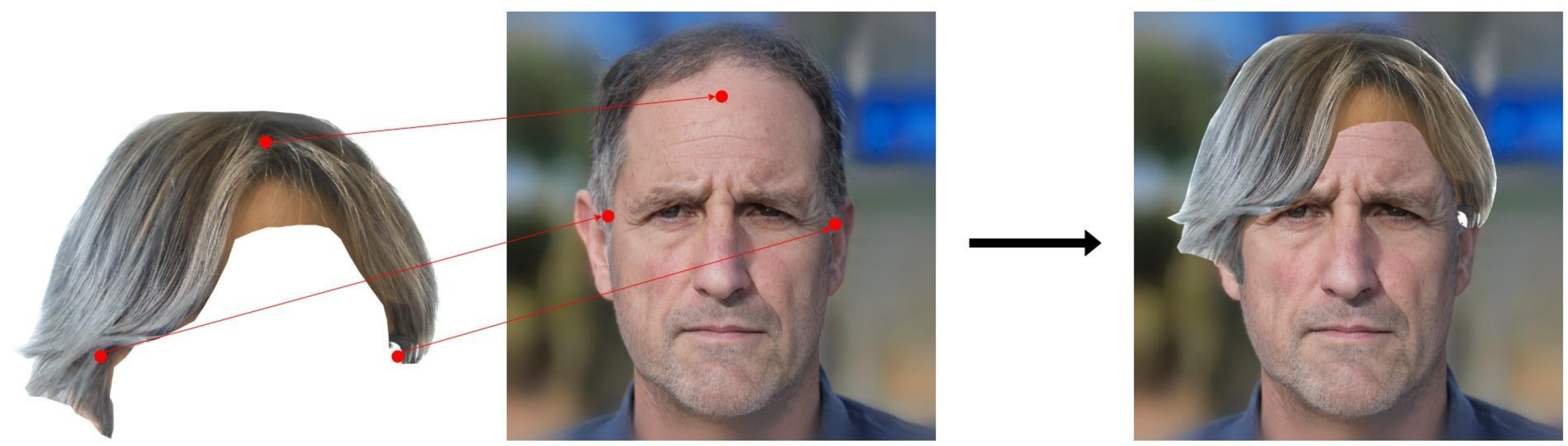


68 facial landmark detection

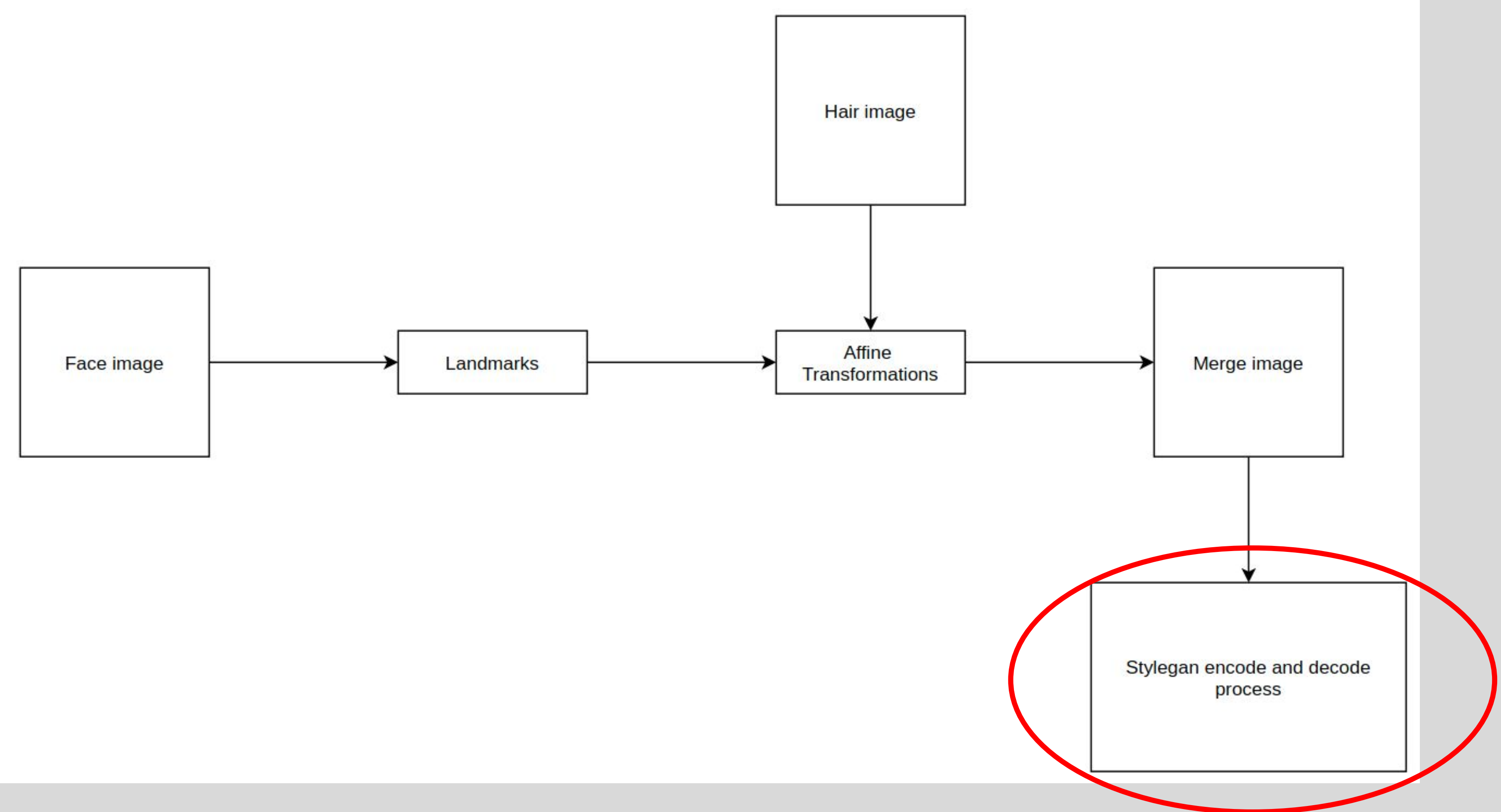
Process


Hair image
Face image
Landmarks
Affine Transformations
Merge image
Stylegan encode and decode process

Affine transformations

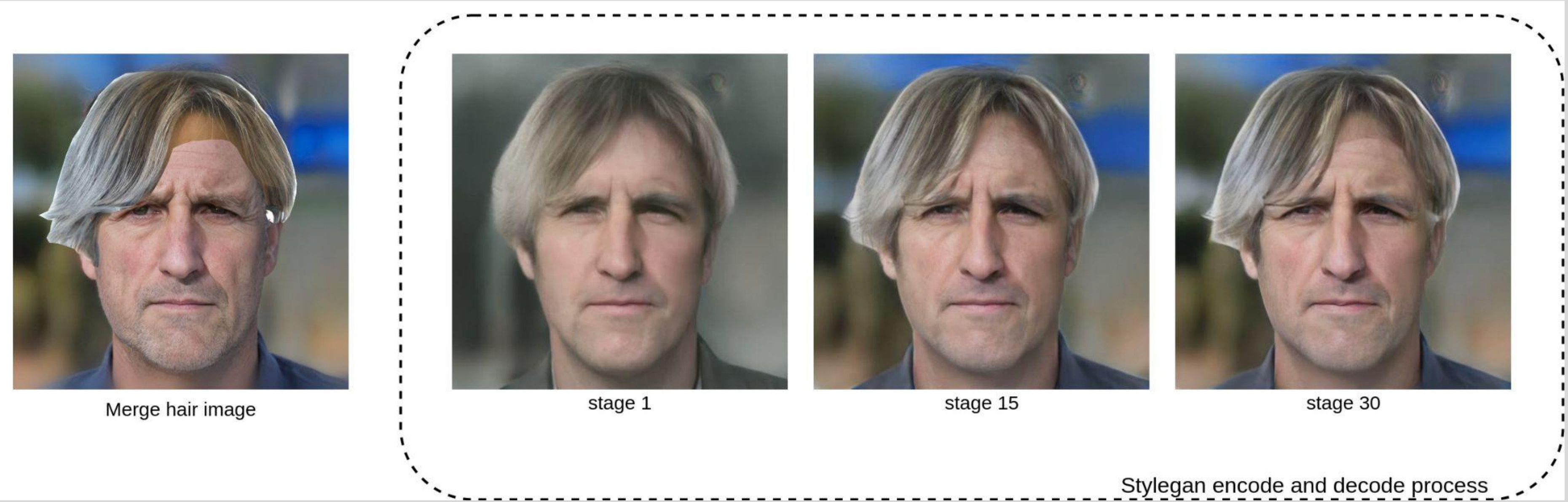


Process

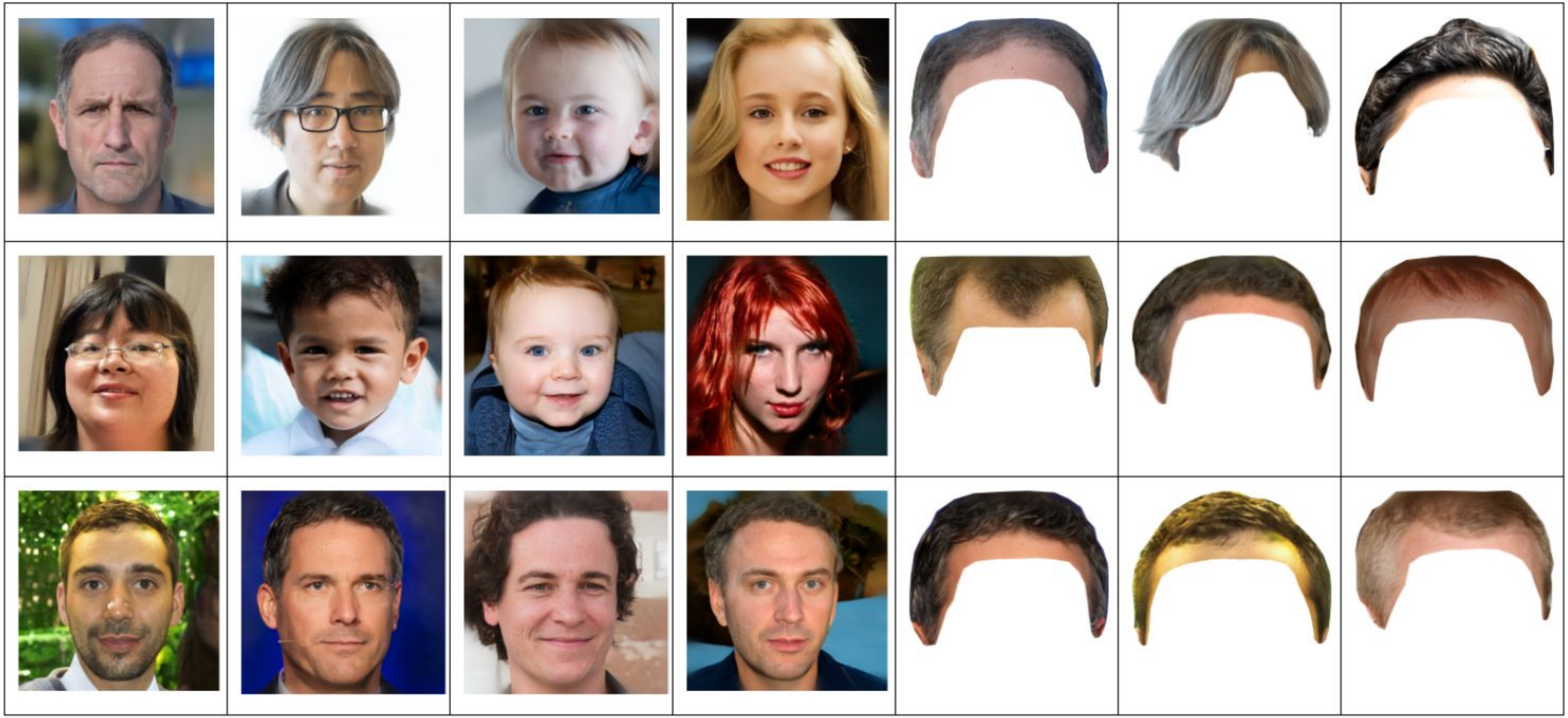


Hair image
Face image
Landmarks
Affine Transformations
Merge image
Stylegan encode and decode process

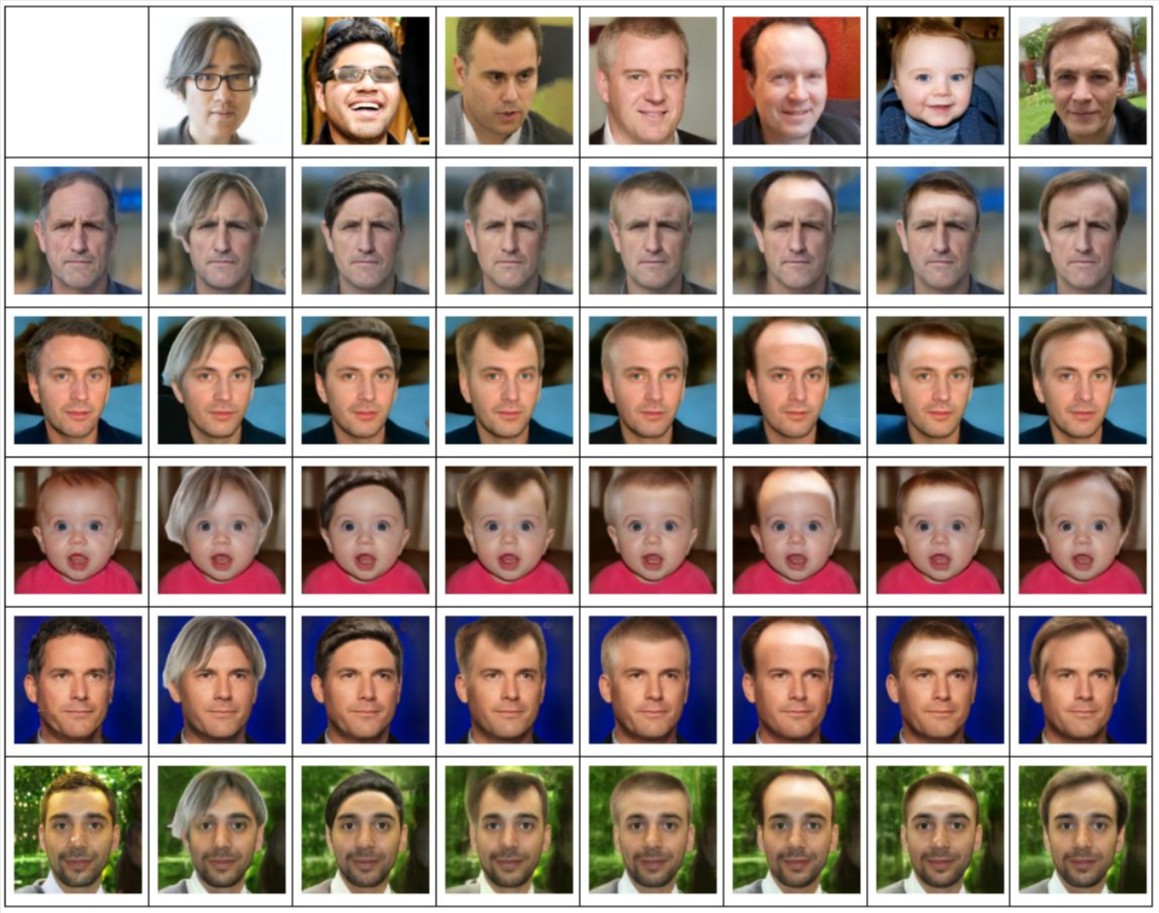
Improve quality process



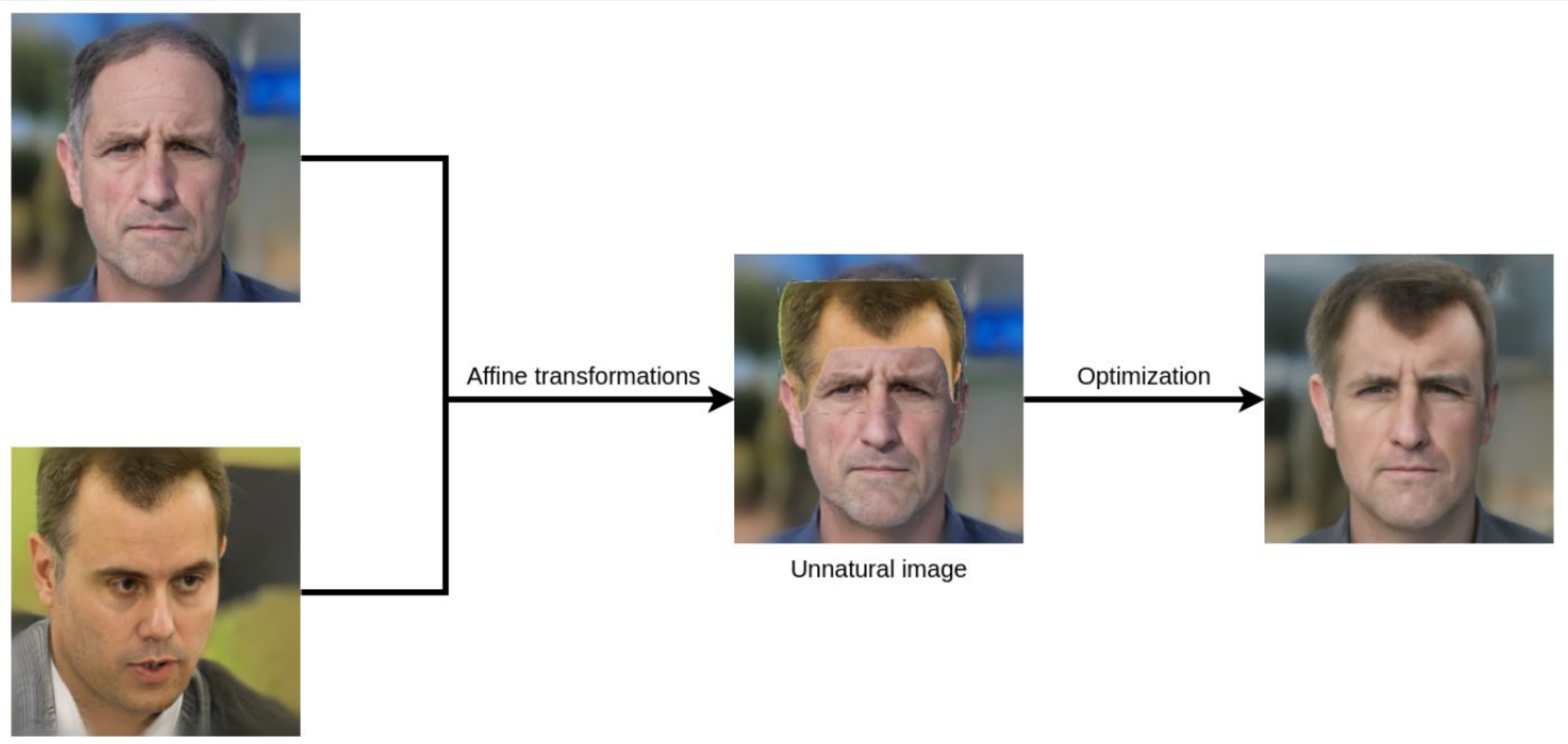
Generated dataset and Evaluation



Result



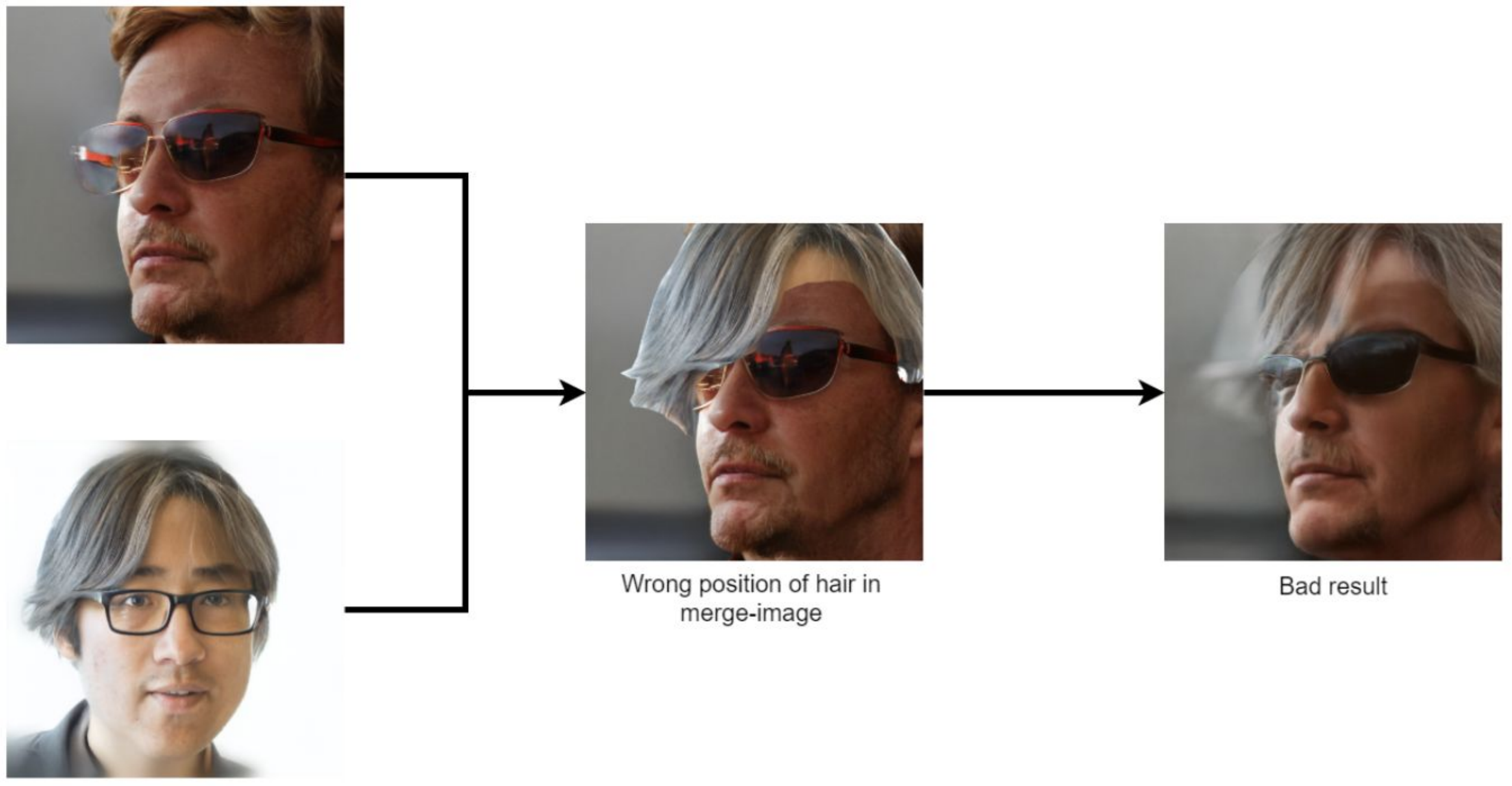
Advantages - Optimize model could remove unnatural features



Advantages - Properties in other areas are not affected

Baby's age is not affected

Disadvantages - Different directions images could generate bad result



Disadvantages - Effect of brightness to the result

4. Conclusion and Future works

Conclusion

- Provide semantic editing latent code method, which based on the assumption of hyperplane in the latent space serving as the separation boundary for any binary semantic, to edit single hair attribute.

- Provide a method, which uses Affine transformation to merge a hair image to a given image and then optimize the image using StyleGan. This method can completely transform hairstyles from one photo to another.

Future works

- Improve problem about more than one attribute correlate with others by using conditional manipulation, which is also mentioned in InterFaceGan.

- Training InterFaceGan again with diverse data sets.

- Training Stylegan encoder with more diverse hairstyles dataset.

Thanks for your attention

Any question please contact our via email
bangpche130602@fpt.edu.vn
khoatdhe130813@fpt.edu.vn

Q & A