

Anime Faces Generation Corresponding to Human Faces

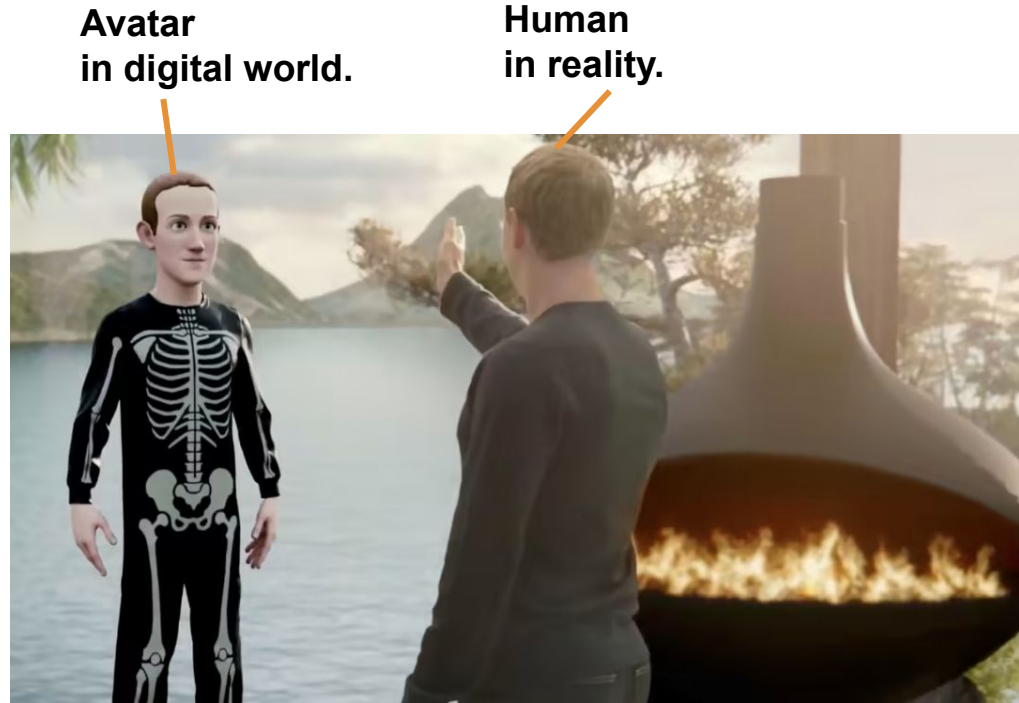
Advisor: Professor Hiroshi Ishikawa
Research on Computer Vision and Pattern Analysis
Wang Yiwen 5120FG52
July 18th, 202

Content

1. Background
2. Target
3. Literature Review
4. Method
5. Result
6. Summary

Background

People want to project themselves into the digital world.



Facebook CEO Mark Zuckerberg speaks to an avatar of himself during a live-streamed virtual and augmented reality conference to announce the rebrand of Facebook as Meta.
© Facebook/Reuters

Objective

Input



An image of real
human face



Output

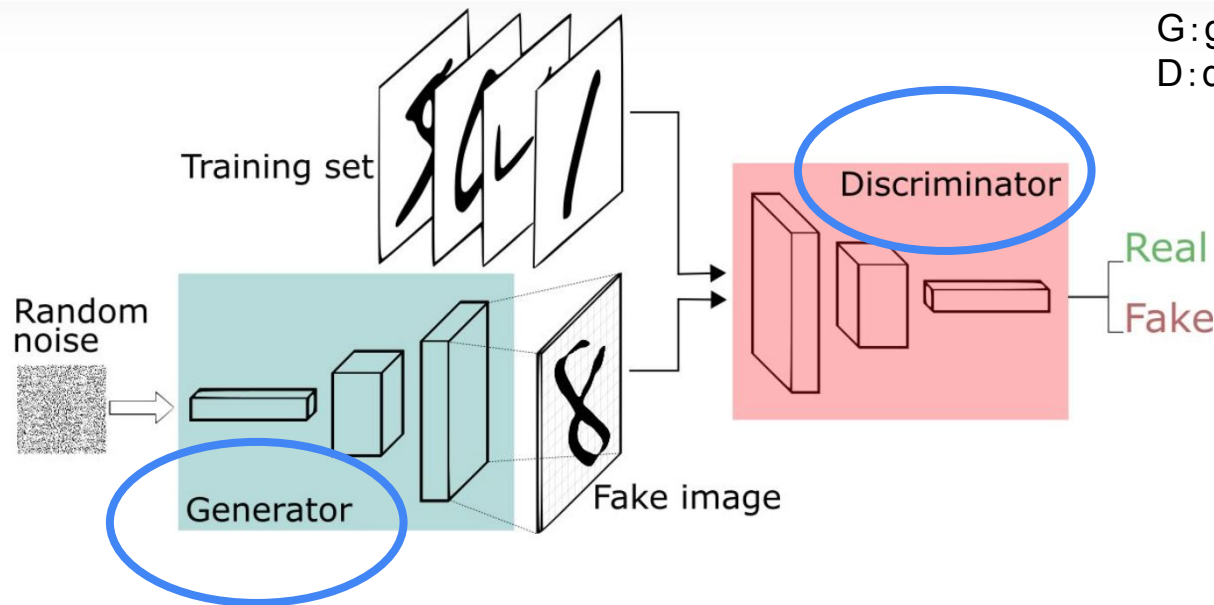


An image of a
similar anime face

Literature Review GAN

Generator+Discriminator = GAN

G: generate images from noise
D: distinguish between genuine/ fake

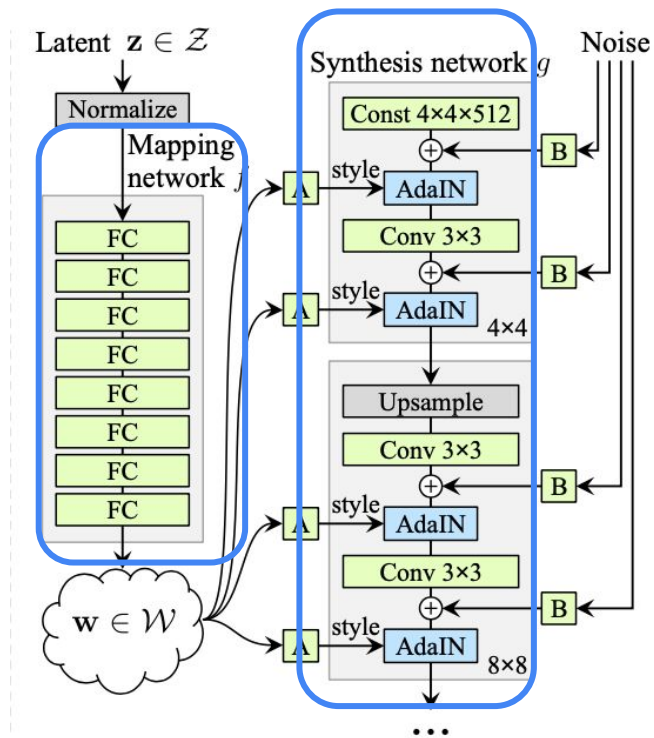


An example of a Generative Adversarial Networks (GAN).

Literature Review StyleGAN

StyleGAN is GAN with a new generator consists of

- a mapping network to provide style
- a synthesis network to generate images



(b) Style-based generator



$$\text{AdaIN}(x, y) = \sigma(y) \left(\frac{x - \mu(x)}{\sigma(x)} \right) + \mu(y)$$

Literature Review Previous works doing this task



cons of previous works:

1. reflecting almost no facial appearances from the input
2. too close to the input and few anime artifact

GANs&Roses

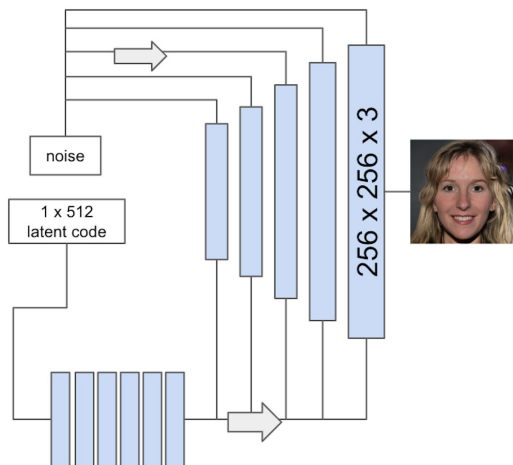
An image-to-image framework generating images by extracting and combining content and style

Swapping AutoEncoder

An image-to-image framework generating images by converting the input images into latent code and dividing it into structure code and texture code

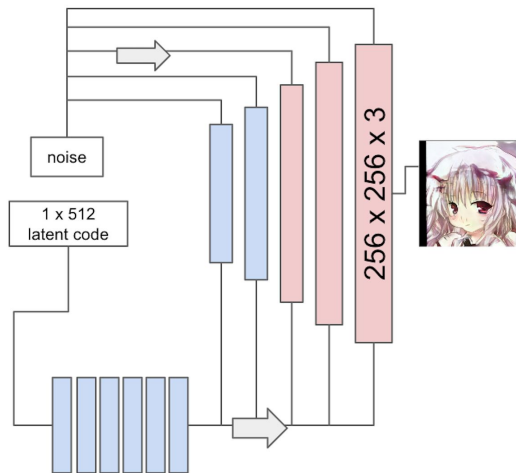
Proposed Method Generation

We use styleGAN to train



pre-train generator on FFHQ dataset

Flickr-Faces-HQ consists of 70,000 high-quality PNG images and contains considerable variation in terms of age, ethnicity and image background.

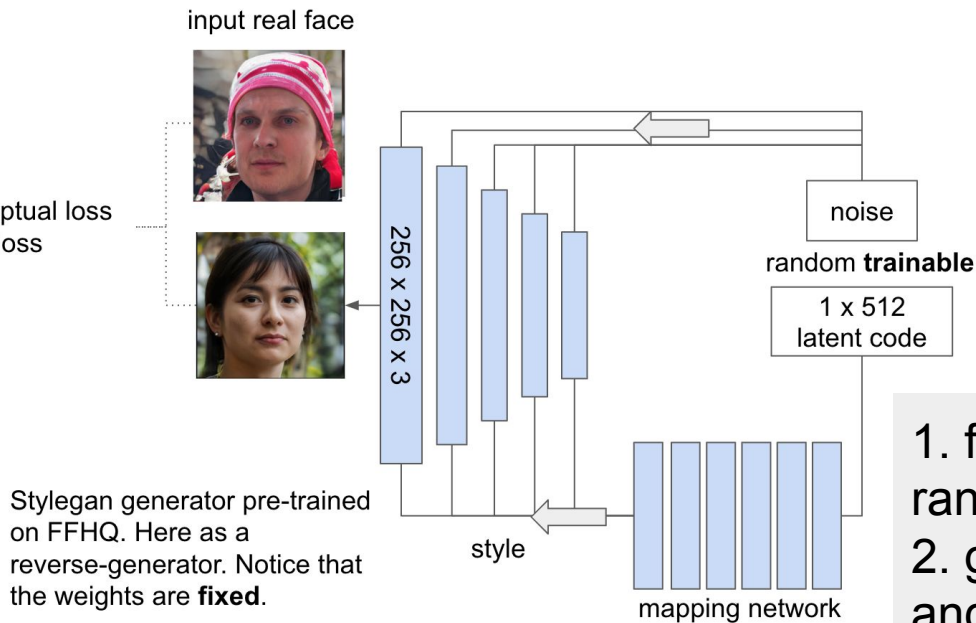


fine-tune generator on a different dataset (Anime)

We use an anime dataset consisting of 60k+ images from Kaggle, which is an online community platform for data scientists and machine learning enthusiasts.

Proposed Method Reverse Generation

- Perceptual loss
- MSE loss



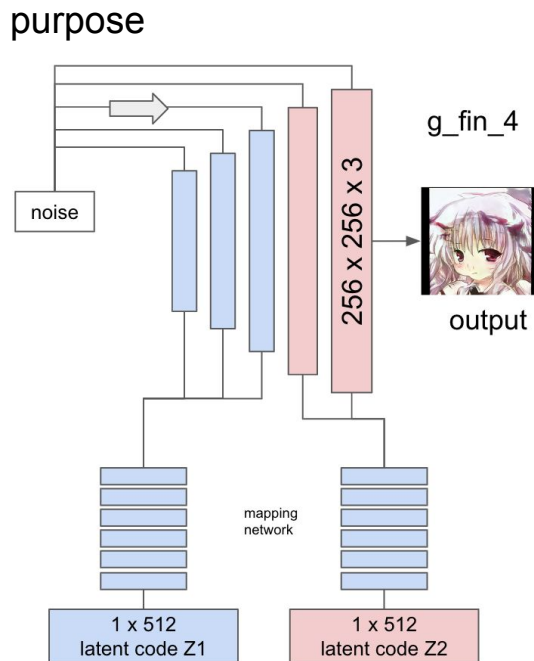
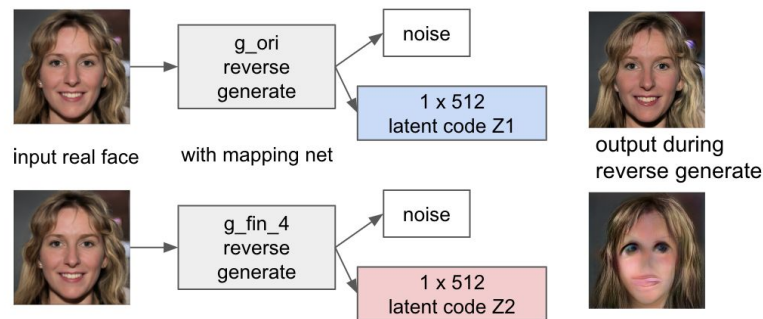
To find the latent code that realizes the image, using optimization

1. fix the generator and initiate a random&learnable latent vector
2. generate from this latent vector and compare output&input
3. edit the latent code until output is almost the same as input

Proposed Method Cross Domains Reverse Generation

g_ori: stylegan2 generator trained on FFHQ

g_fin_4: g_ori finetuned with 4 layers on anime dataset

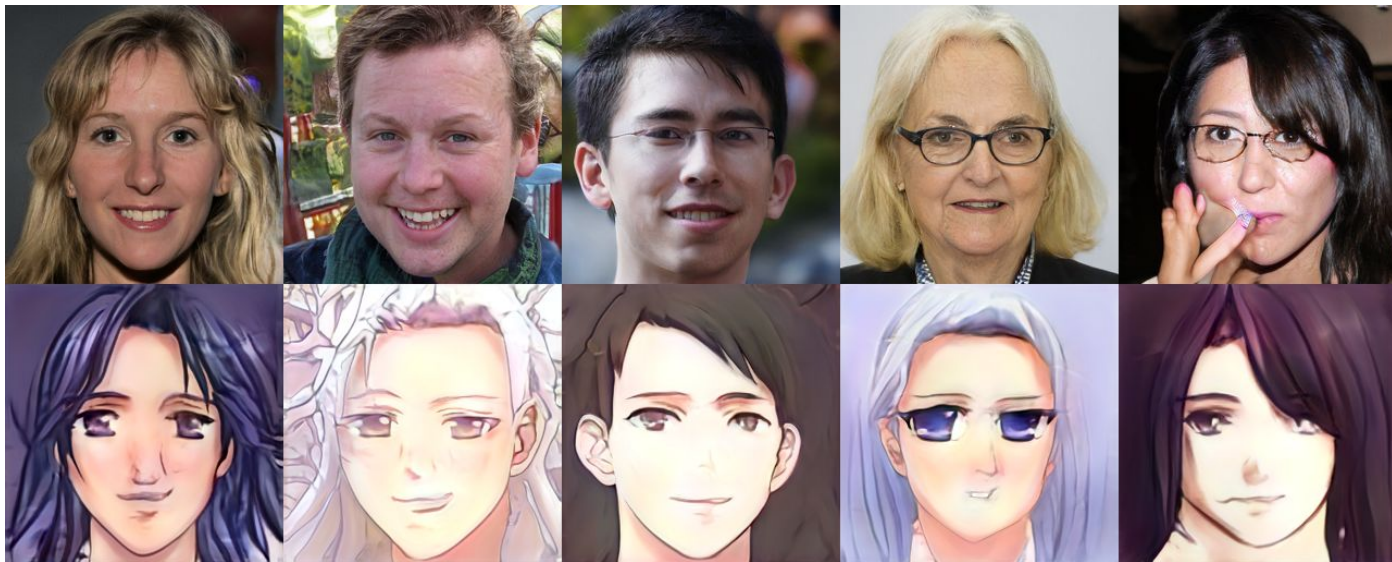


Not require
correspondence
between two domains
when training.

But suppose two
different domains
share latent at a higher
semantic level.

Result Final Examples

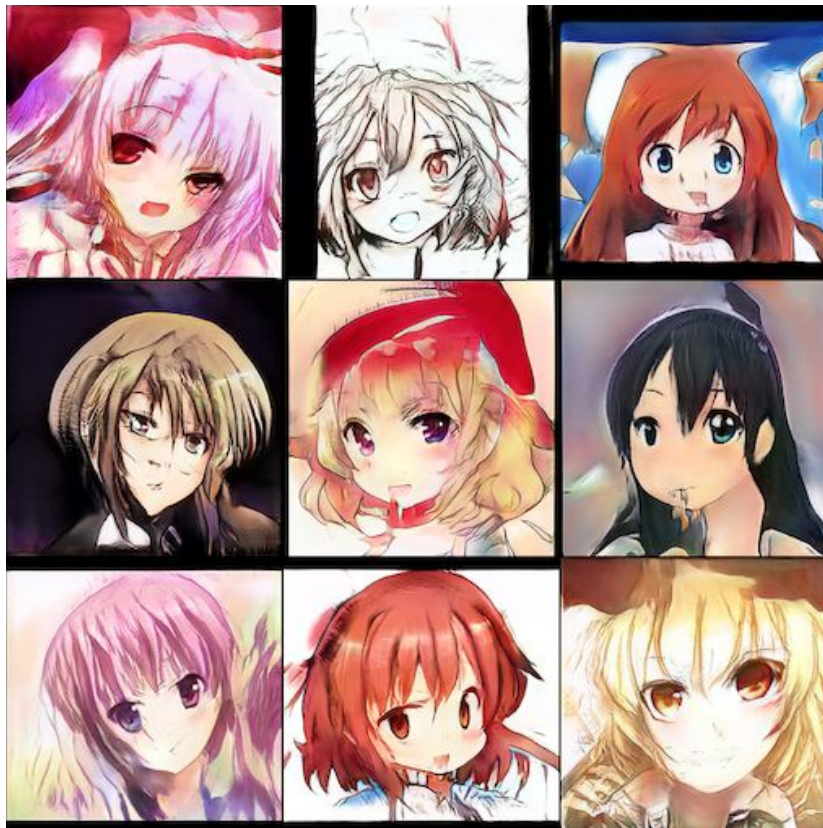
1. reverse generation does well for women with long hair & big eyes
2. salient colors in the input affect the whole of output.
3. the contours of the anime faces' hair are not clear



input

output

Result Anime Images from Fine-tuned Generator



All of the images here are generated from random latent vectors.

The fine-tuned generator can generate nice anime images with little training cost

Result Reverse Generaion

When finding the similar latent code, we can generate almost the same images.
For a different domain, we can find a most similar image.



Reverse generate with the pre-trained generator
left: input image
right: reverse generated image



Reverse generate with the fine-tuned generator
left: input image
right: reverse generated image

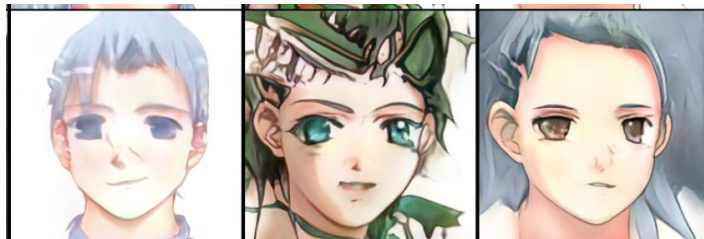
Summary

This thesis

- creates anime similar to input real-life human faces
- based on forward and reverse generation
- requires the image correspondence of two domains in latent space



input



output