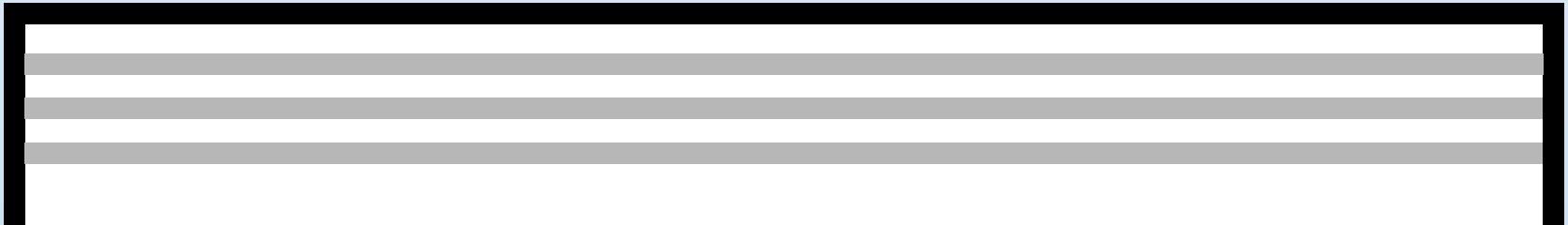
Graduation Project Proposal

4분반 5조 - 김다혜, 선다혜, 오찬희, 심우석

Table of Contents

01

Motivation



Motivation

그림 1

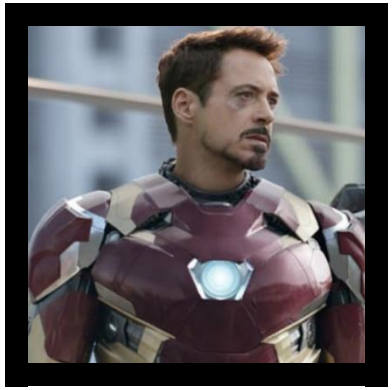


그림 2

그림 3

IF...

Motivation

그림 4

그림 5

Motivation

그림 6

Fun

그림 7

그림 3

Reminiscence

Motivation

그림 8

- Not a smartphone app
- Paid service

Motivation

그림 9

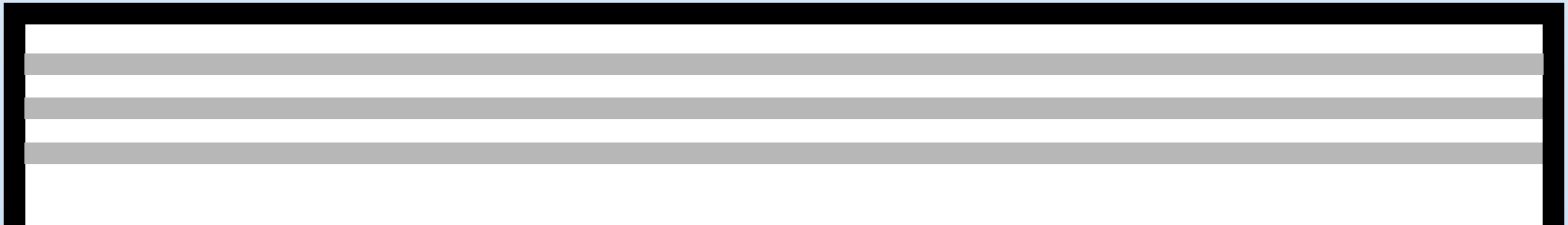
- Vulnerable to security
- Just face swapping
- Inconvenient use

Motivation

- Smartphone app
- Remind person in memory
- Make my own deep fake video

02

Key features



Key features

그림 10

그림 11

DeepFake(Face Swap) + Facial Expression(Lip sync)

Key features

그림 12

Quality ?
Or
Time ?

Key features

- Easy to Use
- DeepFake + Facial Expression(Lip Sync)
- High Quality + Done Quickly Or

  Low Quality + Done Slowly

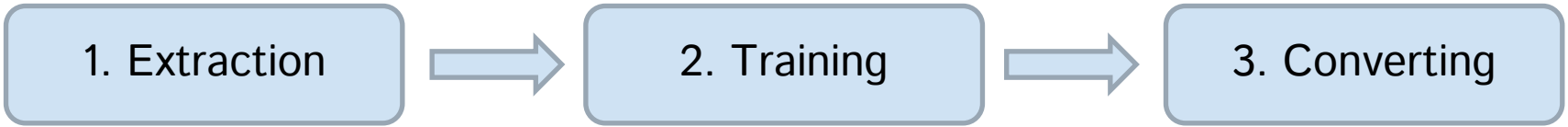
03

Technologies used

Technologies used – Deep Fake(Face swap)

[Deep Fake]

- A deep learning – based technique able to create fake images/videos.
- Swapping the face of a person in an image or video by the face of another person.
- To create an image or video, it has to go through **3 processes**.



Technologies used – Deep Fake(Face swap)

1. Extraction

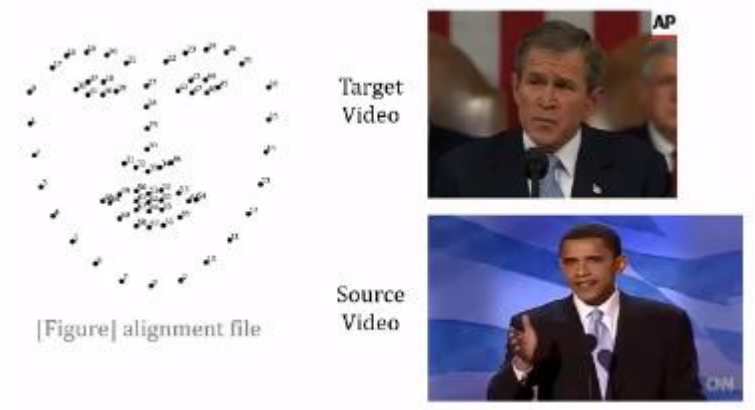


그림 13

Aligned photo

Semantic map

그림 14

To generate a set of faces, and optionally on alignments file and mask, for training.

To generate an alignments file and mask for converting your final frames.

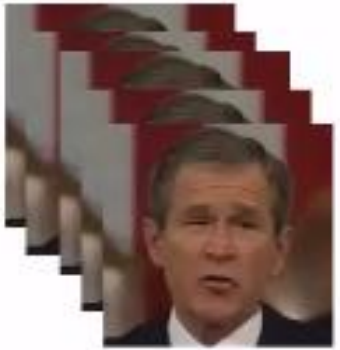
Technologies used – Deep Fake(Face swap)

2. Training

그림 13

Few-shot

3 ~ 5 images



Multi-shot

2000 > images

Technologies used – Deep Fake(Face swap)

2. Training – Few-Shot

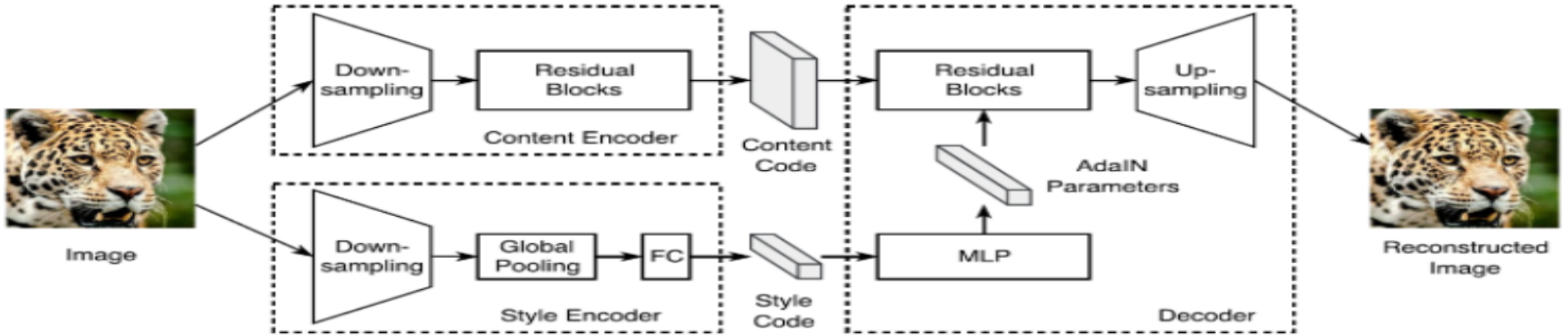


그림 15

FUNIT(Few-Shot Unsupervised Image-to-Image Translation)

Technologies used – Deep Fake(Face swap)

2. Training – Few-Shot

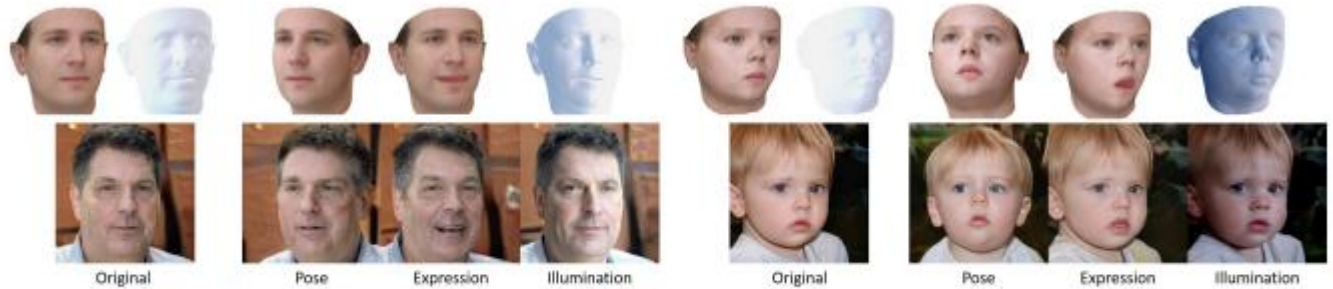


Figure 1: StyleRig allows for face rig-like control over StyleGAN generated portrait images, by translating semantic edits on 3D face meshes to the input space of StyleGAN.

그림 16

StyleGAN(A Style-Based Generator Architecture for GANs

Technologies used – Deep Fake(Face swap)

2. Training – Few-Shot

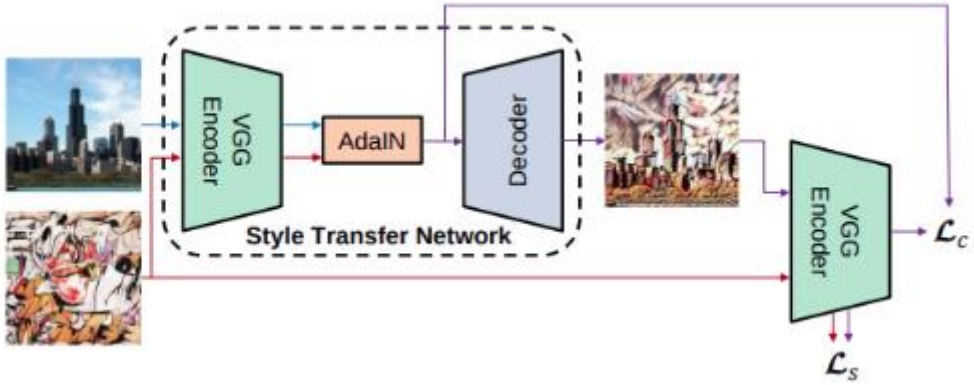


그림 17

AdaIN (Adaptive Instance Normalization)

Technologies used – Deep Fake(Face swap)

2. Training – Few-Shot

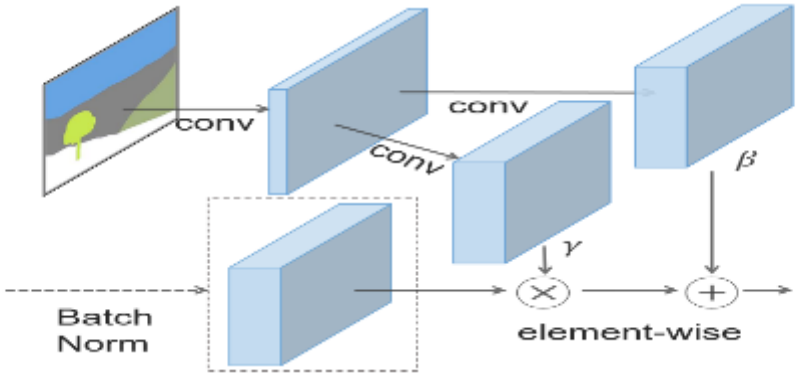


그림 18

SPADE (Spatially-Adaptive Normalization)

Technologies used – Deep Fake(Face swap)

2. Training (TensorFlow) – Few-Shot

FUNIT or StylerGAN + ADAIN + SPADE

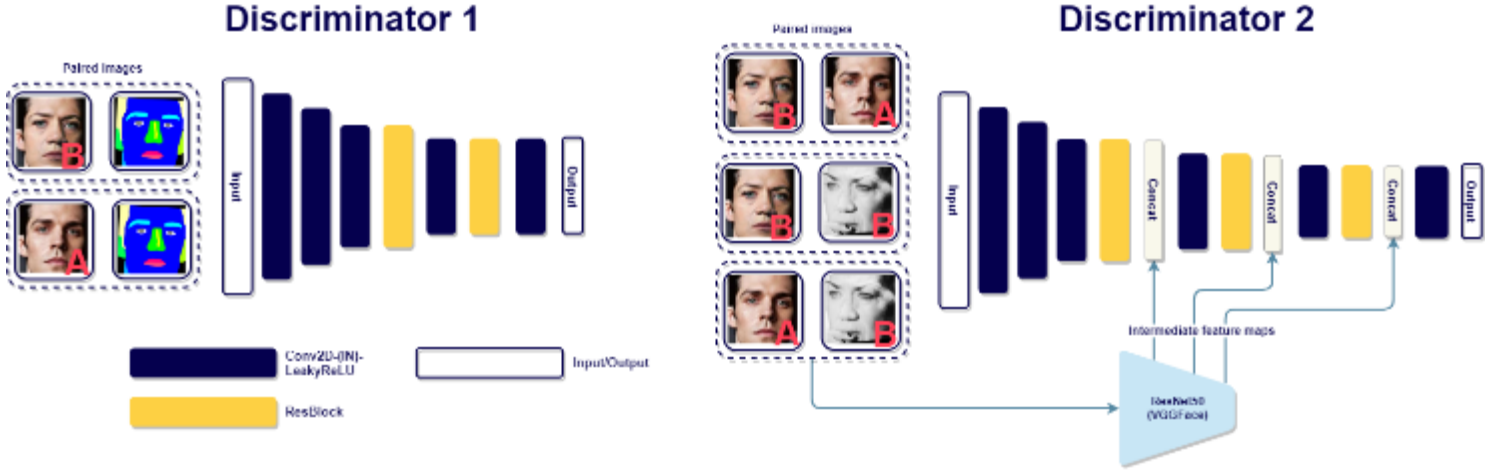


그림 19

Technologies used – Deep Fake(Face swap)

2. Training – Multi-shot

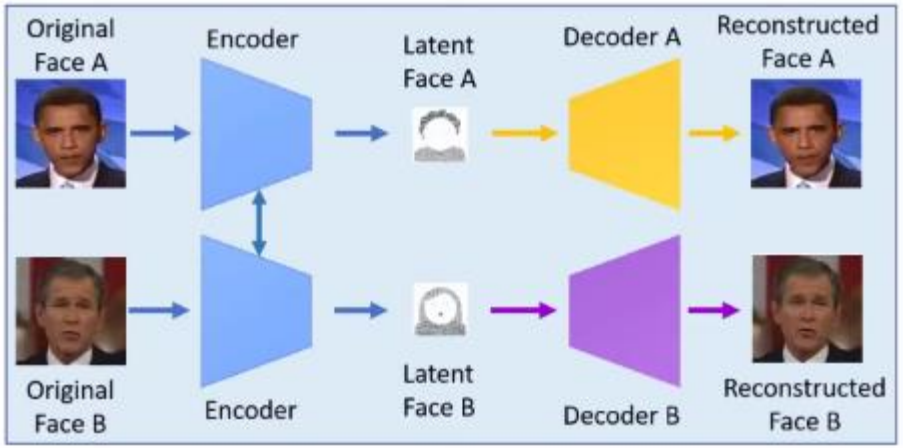


그림 13

AutoEncoder – For troubleshooting

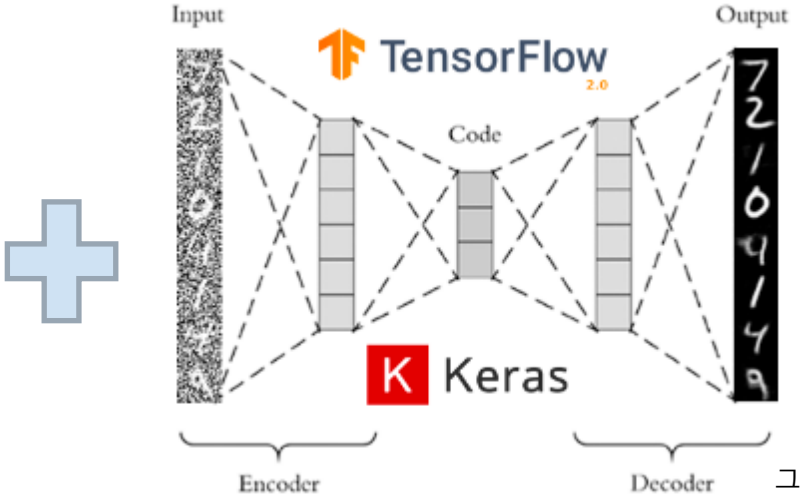


그림 20

Technologies used – Deep Fake(Face swap)

2. Training – Multi-shot

GAN (Generative adversarial networks)

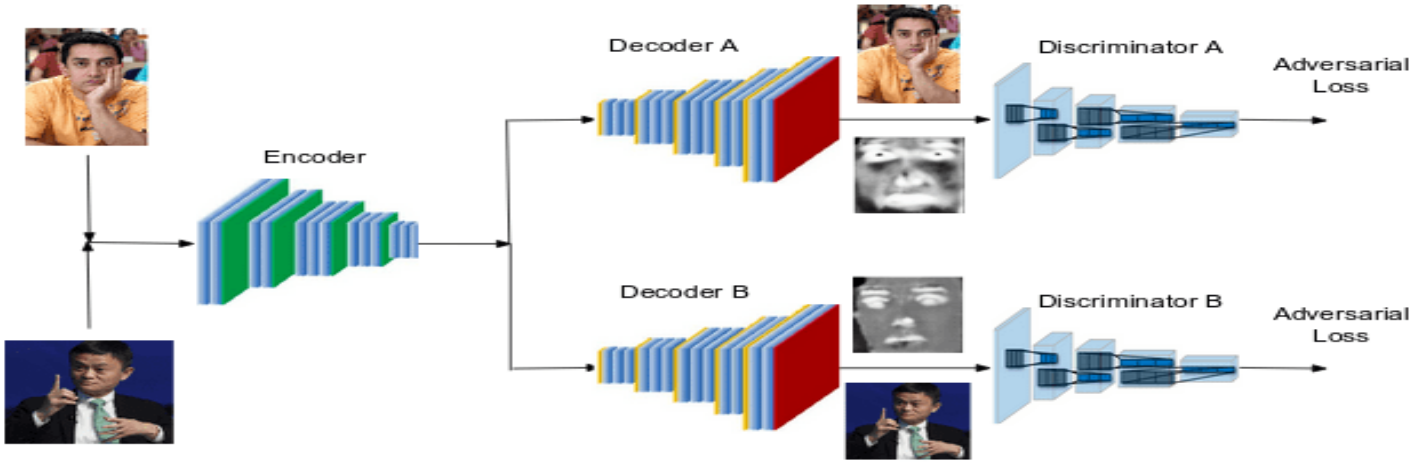


그림 21

Technologies used – Deep Fake(Face swap)

3. Converting

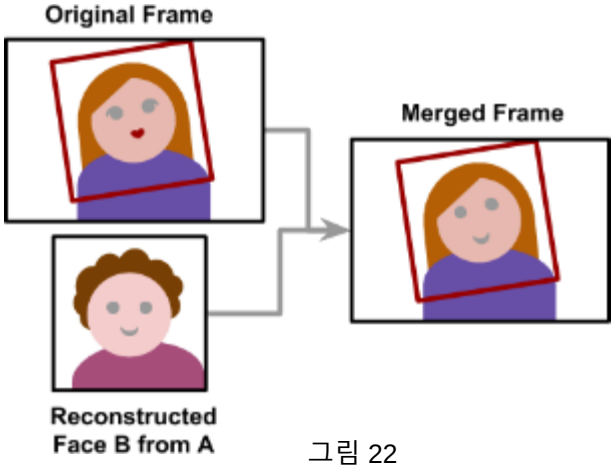


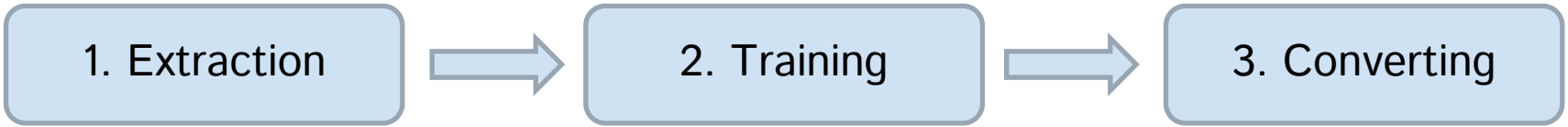
그림 22

그림 10

Technologies used – Speech-To-Lip sync

[Speech-to-Lip Sync]

- A deep learning – based technique able to change the sound to lib sync.
- Technology that converts acoustic speech signals obtained through sound sensors such as voice recordings and microphones into lib sync.
- In order to achieve recognized results, it has to go through **3 processes**.



Technologies used – Speech-To-Lip sync

1. Extraction

그림 23

- The extraction of the intonation, voice size, etc. of a voice signal in numerical form is called a characteristic vector.
- SST use characteristic vectors to generate criteria to determine the meaning of voice signals.

Technologies used – Speech-To-Lip sync

2. Training

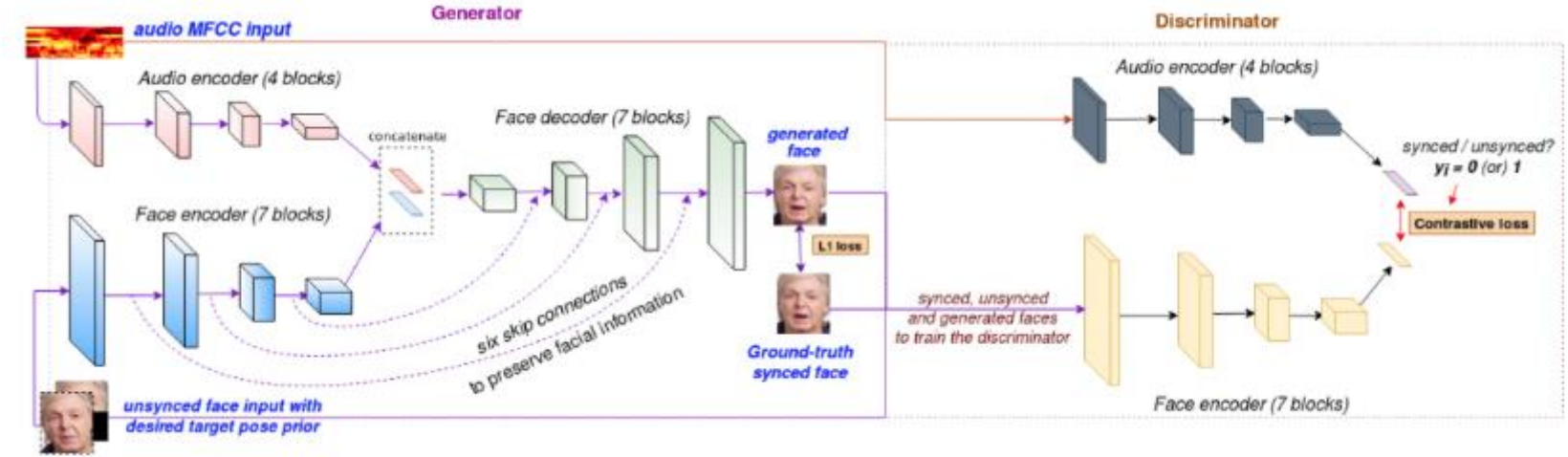


그림 24

Technologies used – Speech-To-Lip sync

3. Converting

Audio

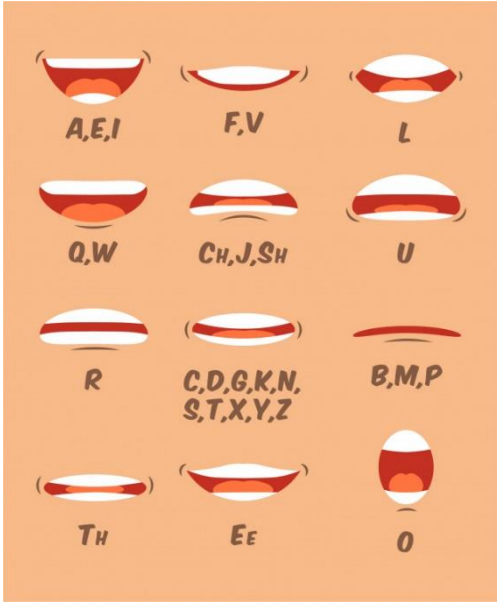


그림 25

Technologies used – Speech-To-Lip sync

3. Converting

그림 26

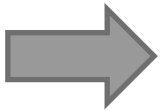
Technologies used – Framework

[Modeling]

그림 27

그림 28

그림 29



[Mobile implementation + Server]

그림 30

그림 32

그림 31

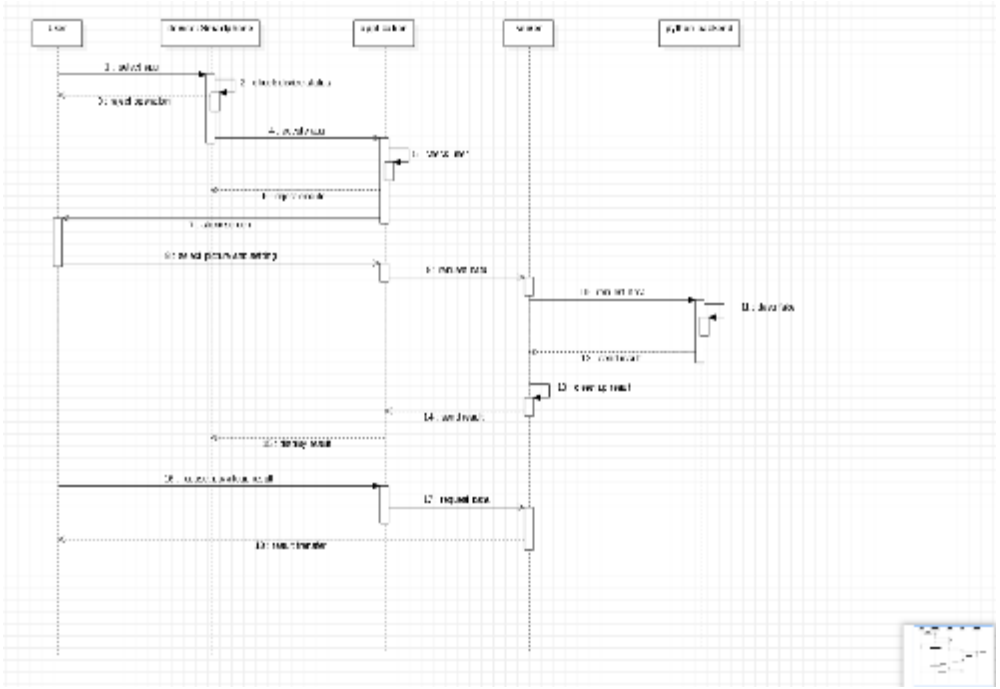
[Version control]

그림 33

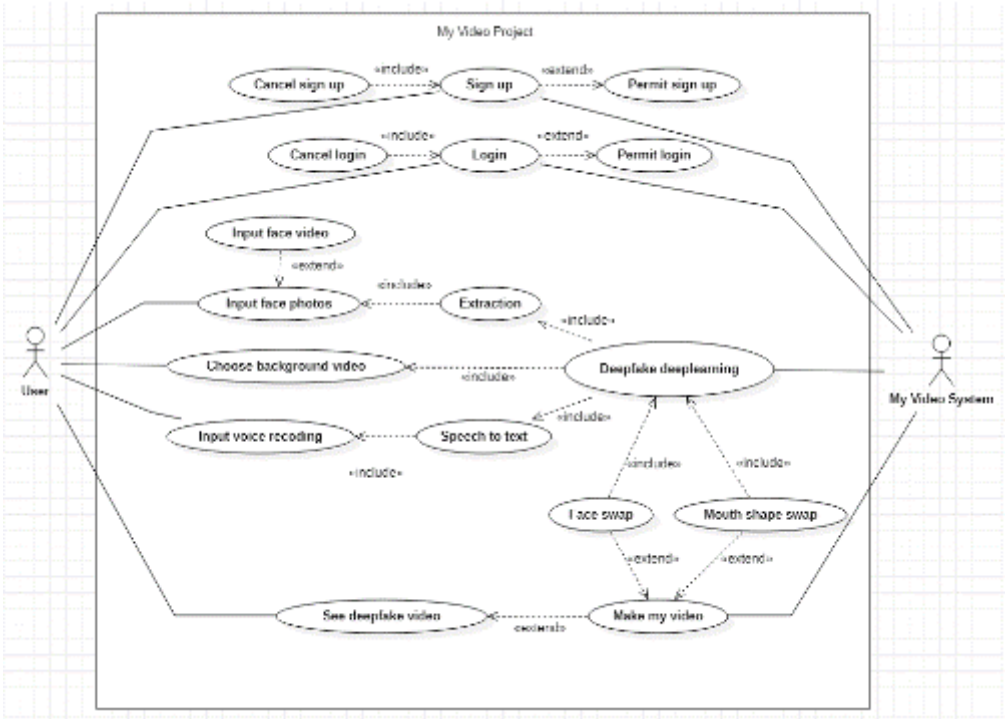
04

Progress

Sequence diagram



Usecase diagram



Agile Development

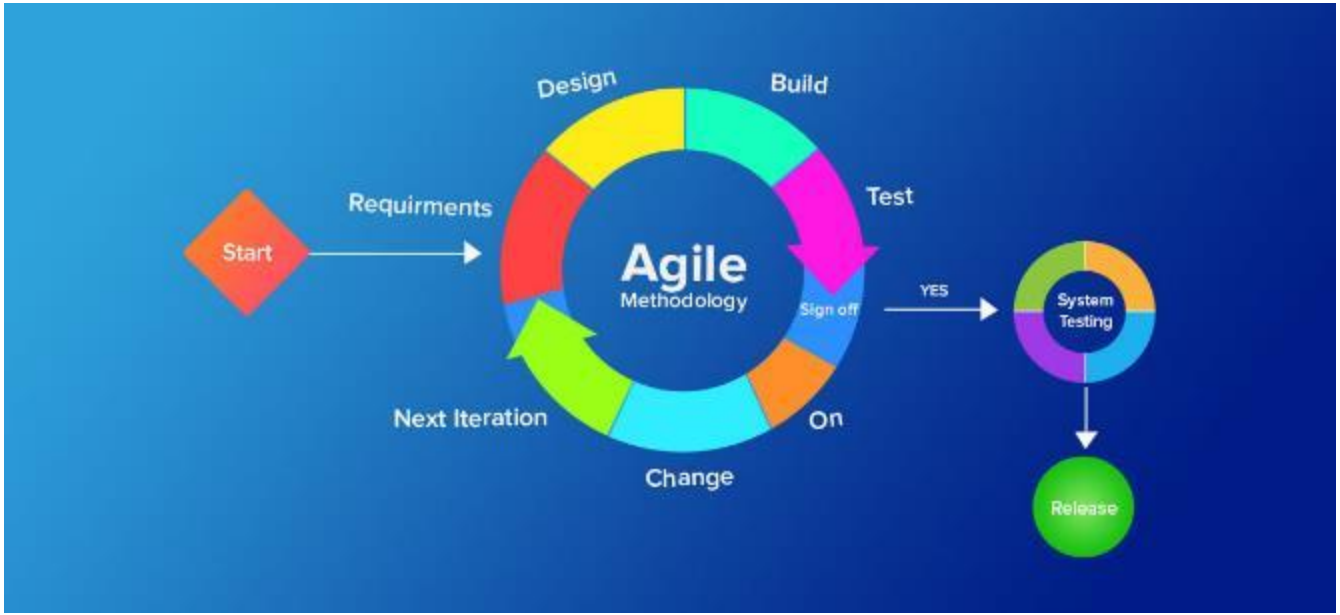


그림 33

05

Development Plan

Focus Direction

[Distinction]

- DeepFake + Lip-synching

- Existing

-> cover only human face

-> My image or base image requires mouth shape to move

[Effective]

- If a person who wants to remember has only a picture and a voice left, the prospect is that it can be provided as if he or she is speaking the voice.

- User can synthesize your own face and say your own lines in the best scenes of the movie.

Application Screen

[Home Screen]

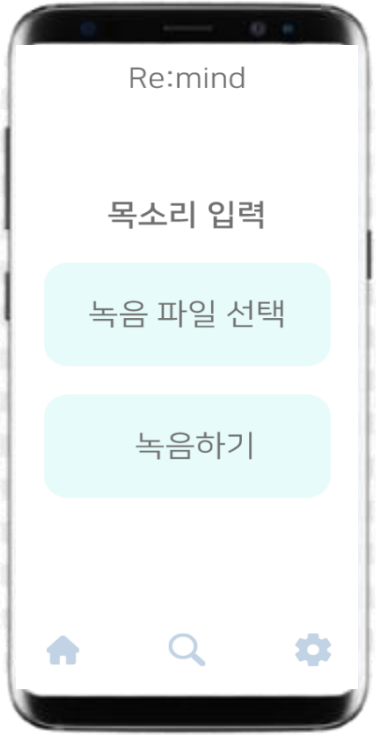
- Provides a list of application self-recommended base images.
- Home / Search / Settings
  Search can use base images and search functions divided by tags

Application Screen

[Choice Fabrication Mode]

- First screen available when base image is selected.
- Quick: fast time, low quality
  Quality: slow time, high quality

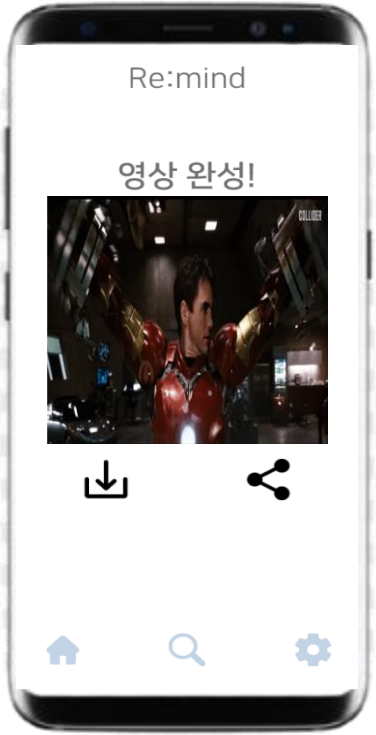
Application Screen



[Voice Input]

- Enter a voice for the lip-syncing function.
- Select files that have already been recorded,
  Proceeding to a live recording

Application Screen



[Result]

- Download and share
- Using voice function -> mp4
  If not -> gif or jpg

06

Role & Plan

Plan

2021 Summer

Dataset collection, preprocessing

Train Deep Fake, Speech-to-text

2021 Fall

2021 Winter

Deep Fake Lip Syncing

Testing Documentation

2022 Spring

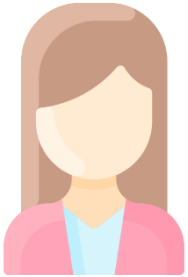
Roles

심우석
201636417

- DeepFake Research
- GAN, SPADE, AdaIN, Tensorflow Keras study

오찬희
201735855

- DeepFake Research
- GAN, SPADE, AdaIN, Tensorflow Keras study



김다혜
201835414

- Speech To Rib sync Research
- Server Study

선다혜
201835466

- Speech To Rib sync Research
- Development Environment
- Research

Wiki & YouTube

https://github.com/dntjr41/Graduation_project/wiki

https://www.youtube.com/watch?v=OO7-k4CzkLs

References

- 그림1: https://www.pexels.com/photo/woman-in-red-tank-top-holding-surfboard-3722185/
- 그림2: http://m.sports.khan.co.kr/view.html?art_id=202009021555013&sec_id=540401
- 그림3: https://www.pexels.com/photo/smiling-man-and-woman-wearing-jackets-1642883/
- 그림4: https://www.youtube.com/watch?v=uflTK8c4w0c
- 그림5: https://www.bbc.com/Korean/news-51498614
- 그림6: https://www.techadvisor.com/how-to/social-networks/how-face-swap-in-snapchat-3636570/
- 그림7: https://smartkoh.tistory.com/326
- 그림8: https://www.myheritage.co.kr/deep-nostalgia
- 그림9: https://play.google.com/store/apps/details?id=com.deepworkings.dfstudio&hl=ko&gl=US
- 그림10: https://www.digitalmusicnews.com/2020/01/07/facebook-bans-deepfakes/
- 그림11: https://github.com/astorfi/lip-reading-deeplearning
- 그림12: https://www.amst.com/resources/looking-for-a-fast-cheap-high-quality-website-30339
- 그림13: https://www.youtube.com/watch?v=O9UmjWZL3KQ
- 그림14: https://openaccess.thecvf.com/content_ICCV_2019/papers/Shi_Face-to-Parameter_Translation_for_Game_Character_Auto-Creation_ICCV_2019_paper.pdf
- 그림15: https://velog.io/@tobigs-gm1/UNIT
- 그림16: https://www.youtube.com/watch?v=eaW_P85wQ9k
- 그림17: https://openreview.net/pdf/ea6c88006e848ddecc005e35f1b7e2b70c30c7a6.pdf

References

- 그림18: https://towardsdatascience.com/implementing-spade-using-fastai-6ad86b94030a
- 그림19: https://github.com/shaoanlu/fewshot-face-translation-GAN
- 그림20: https://www.pyimagesearch.com/2020/02/24/denoising-autoencoders-with-keras-tensorflow-and-deep-learning/
- 그림21: https://imrahulr.github.io/deepfakes/gans/autoencoder/coe-cnds/2018/01/09/deepfakes.html
- 그림22: https://www.alanzucconi.com/2018/03/14/introduction-to-deepfakes/
- 그림23: http://www.hellot.net/new_hellot/magazine/magazine_read.html?code=202&sub=002&idx=42081&page=&list=
- 그림24: http://cvit.iiit.ac.in/research/projects/cvit-projects/facetoface-translation
- 그림25: https://www.adobe.com/th_th/creativecloud/video/discover/animation-lip-sync.html
- 그림26: https://github.com/Rudrabha/LipGAN
- 그림27: https://dora-guide.com/pycharm-install/
- 그림28: https://github.com/alvarobartt/serving-tensorflow-models
- 그림29: https://keras.io/ko/
- 그림30: https://blog.just4fun.kr/2017/10/android-studio-30.html
- 그림31: https://firebase.google.com/brand-guidelines?hl=ko
- 그림32: https://aws.amazon.com/ko/
- 그림33: https://www.kindpng.com/imgv/woooRm_github-logo-png-github-transparent-png/
- 그림33: https://medium.com/@preston.elliott/what-is-agile-methodology-20ca93aa7100

Thank you!!!