KHUDA / Computer Vision

Voice-Conversion을 이용한 내 가수 커버곡 듣기

KHUDA / 구태형, 임정우, 유혜지, 백지원

심화 프로젝트

목 차

0. 프로젝트 배경

프 로 젝 트 배 경

[바라던 바다▶모아듣기] 석양도 로제(ROSÉ) 앞에선 조명일 뿐„ 월클 로제 모아듣기 (무대.ver) 〈바라던 바다 (sea of hope)〉 | JTBC 210803 방송

조회수 1093만회 • 1년 전

JTBC Entertainment

로제 Playlist 00:00:00 The Only Exception 00:04:05 사랑하기 때문에 00:06:51 Lucky(with 온유) 00:10:05 If I Ain't Got You(With ...

The Only Exception | 사랑하기 때문에 | Lucky(with 온유) | If I Ain't Got You(With 온유&이수현) | Read my... 챕터 7

어쿠스틱 연주와 음색의 paradise💛 권진아(KWON JIN-AH)의 'Attention' | 비긴어게인 오픈마이크

조회수 75만회 • 7개월 전

Beginagain 비긴어게인

어쿠스틱 연주와 음색의 paradise 권진아(KWON JIN-AH)의 'Attention' #BeginAgain #OpenMic #오픈마이크 #Attention #권진아 ...

다양한 커버곡들이 인기

내가 좋아하는 가수의 버전으로 듣자

심 화 프 로 젝 트

1. Voice-Conversion

심화 프로젝트

Voice-Conversion

음성에서 언어적 내용(linguistic contents)는 변하지 않고, 화자의 음성 특징(리듬, 음역, 음색…)만을 변환하는 것.

- 발화 내용
- Source 화자의 특징
  - 리듬, 음역, 음색, …

“범인은 바로 당신이야!”

- 발화 내용
- Target 화자의 특징
  - 리듬, 음역, 음색, …

“범인은 바로 당신이야!”

2. 모델 선택 과정

대표적인 Voice conversion model 3가지

GAN 계열

Generator(생성자)는 discriminator(판별자)를 속이기 위해 실제 음성과 비슷한 음성을 만들어내고, discriminator(판별자)는 generator(생성자)가 생성해낸 음성과 실제 음성을 구별하는 것을 목적으로 학습하는 모델. 결국 생성자는 원본 음성과 구별할 수 없는 고품질의 음성을 생성해낼 수 있음.

Autoencoder 계열

Encoder를 거치며 노드 수를 줄여 특징적 요소만을 남게 하고, 이를 다시 업샘플링하는 decoder에 넣어 원본 음성과 유사한 음성을 만들어내는 모델.

Diff-svc -> pretrained model 사용

Forward process를 통해 복잡한 데이터에 점진적으로 noise를 더하고, reverse process를 통해 noise에서 시작하여 데이터를 복구하는 과정을 반복하며 유사한 목소리를 만들어내는 알고리즘.

GAN은 가장 현실적인 음성을 생성할 수 있지만, 학습 시간이 오래 걸림.

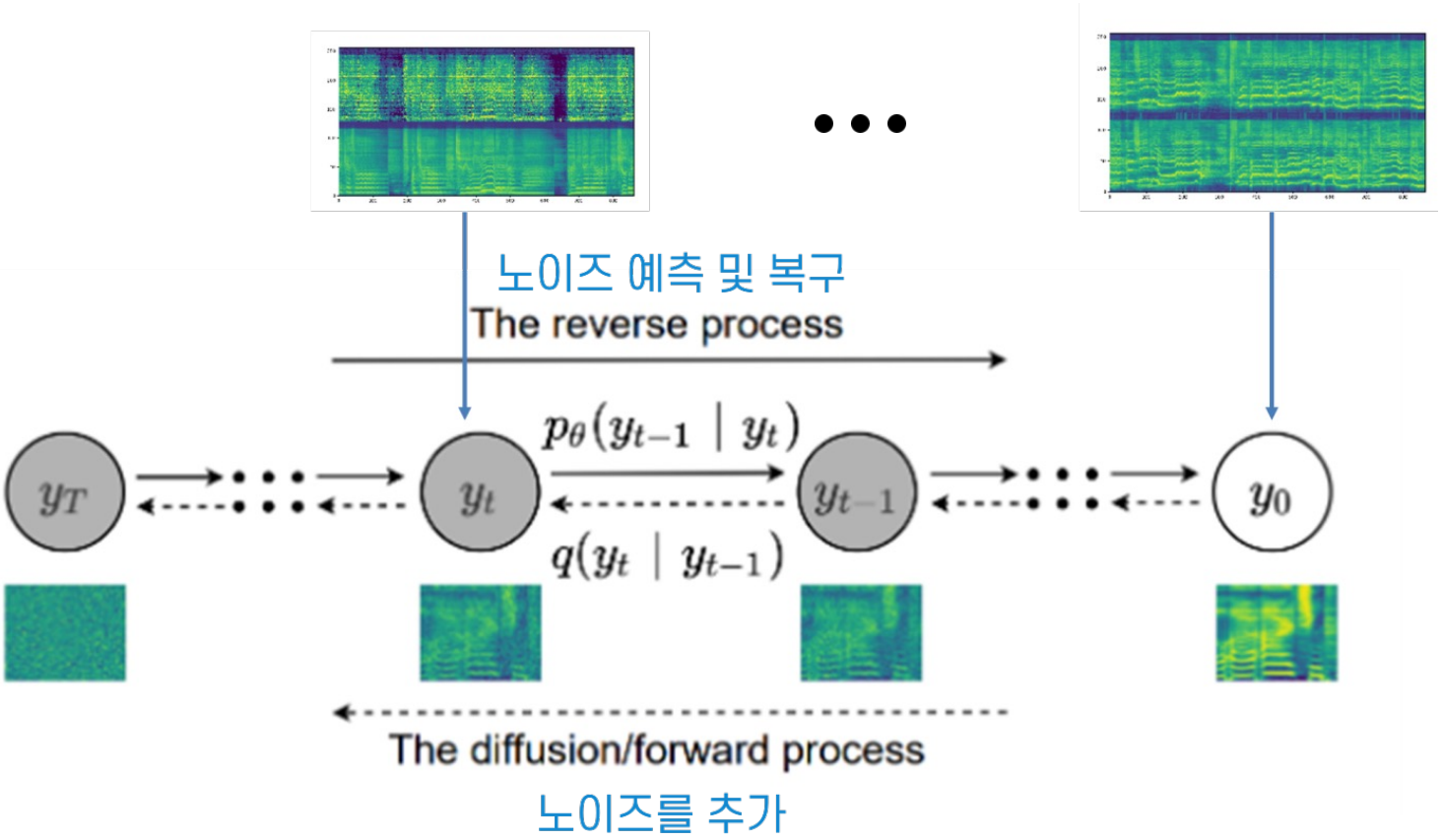
오토인코더는 GAN보다 학습 시간이 짧지만, GAN만큼 현실적인 음성을 생성하지는 못함.

Diff-SVC는 가장 빠르고 고품질의 음성 변환 알고리즘이지만, GAN 또는 오토인코더만큼 현실적인 음성을 생성하지는 못함.

3. Diffusion-SVC

심화 프로젝트

Diffusion-SVC



Data processing

4. 데이터 전처리

심화 프로젝트

Data processing

언어

한국어와 영어는 발음의 특성이 다르기 때문에 한국어 발화자의 음성을 영어 발화자의 음성으로 변환하기 어렵다.

한국어는 자음과 모음이 모두 소리나는 반면, 영어는 자음만 소리

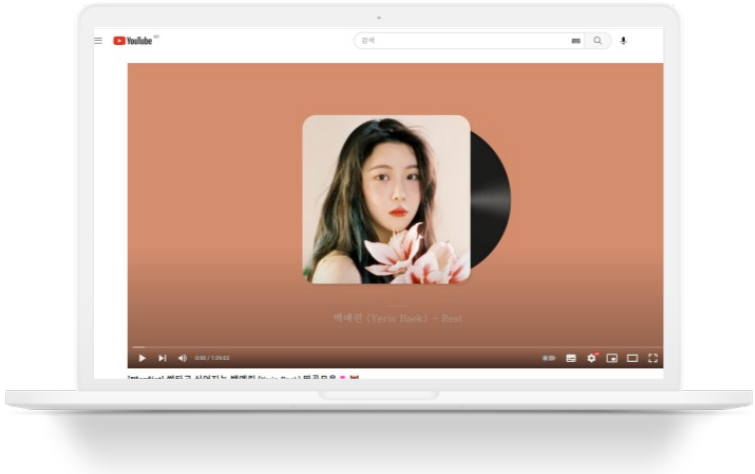
한국어가 영어보다 억양이 다양하다.

음색

남성의 목소리가 여성의 목소리보다 깊고 굵다.

때문에 성별이 다른 발화자의 음성으로 변환하는 것은 더 어렵다.

Data processing

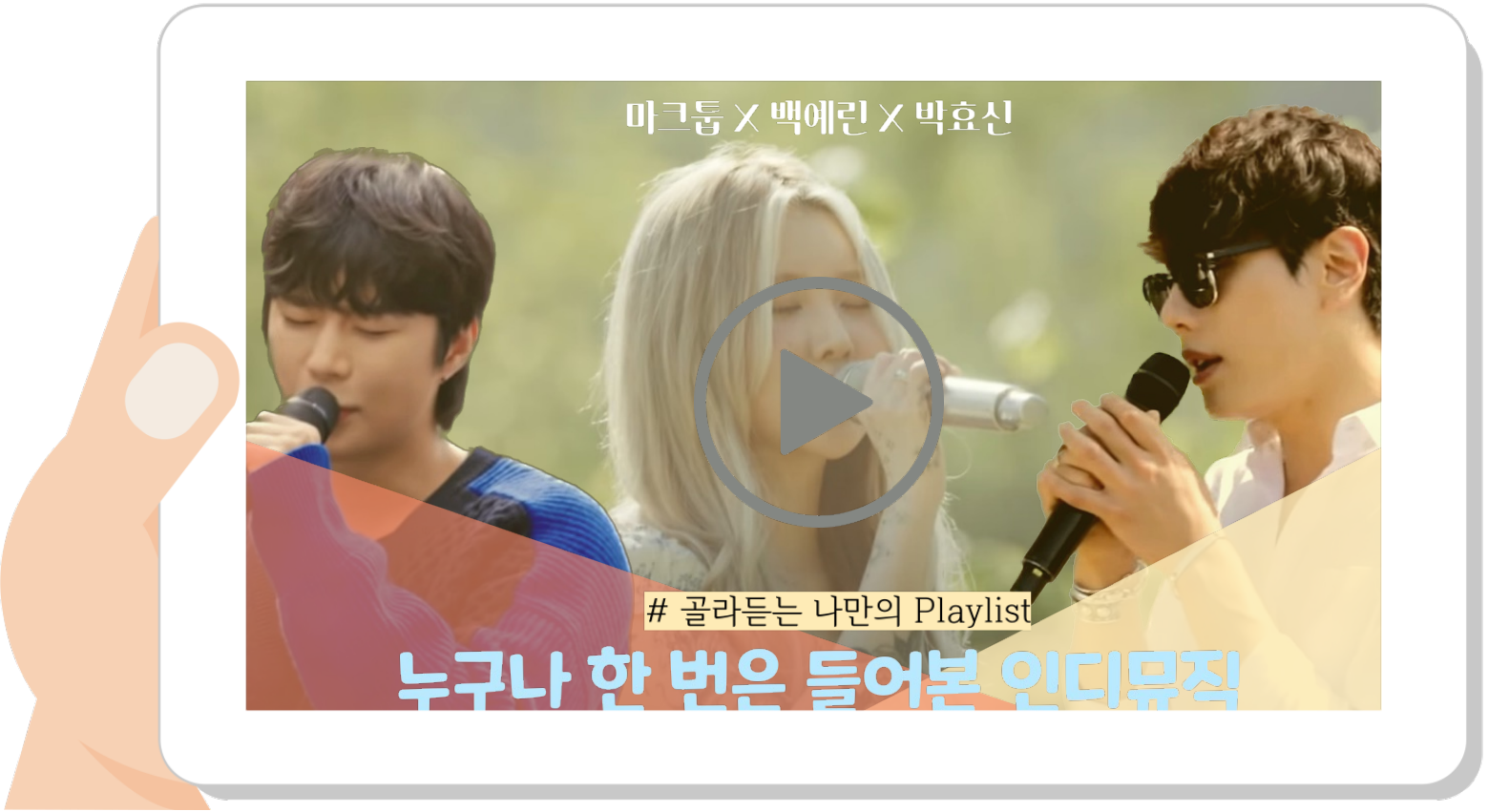


유튜브 영상 추출 및 mp3 변환

-> mr제거 -> wav 파일 변환 -> 공백 구간 제거 -> 15초 길이로 cut

result6.wav
result7.wav
result8.wav
result9.wav
result10.wav
result11.wav
result12.wav
result13.wav
result14.wav

5. 결과


마크툽 X 백예린 X 박효신
골라듣는 나만의 Playlist
누구나 한 번은 들어본 인디뮤직

6. 한계점

limitation

2023

THANK YOU

구태형, 임정우, 유혜지, 백지원

쿠다 심화 프로젝트