



Smootify

Phone, email, or username

Password

Log in

Forgot password? [Sign up for Smootify](#)



Smootify

TobigTonix1516@tobigs.com

Log in

Forgot password? [Sign up for Smootify](#)



음악을 환승하다,
**Infinite AI Music
Streaming**

PLAY

Overview Topics

Background

Motivation
Goal



Related Work

Mash-up
Transition



Dataset



Me

Mas
Tran

Background

Related Work

Dataset

Methods

Results

Demo

Reference

Team

Results

MuseGAN
Smootify



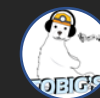
Demo



Reference



Tea



TobigTonix

프로필 보기



Overview Topics

Background

Motivation
Goal



Related Work

Mash-up
Transition



Dataset



Me

Mas
Tran

Results

MuseGAN
Smootify



Demo



Reference



Tea



TobigTonix

프로필 보기



Background



Gapless playback



Related Work



Dataset



Methods



Results



Demo



Reference



Team



Motivation

Gapless playback 란?

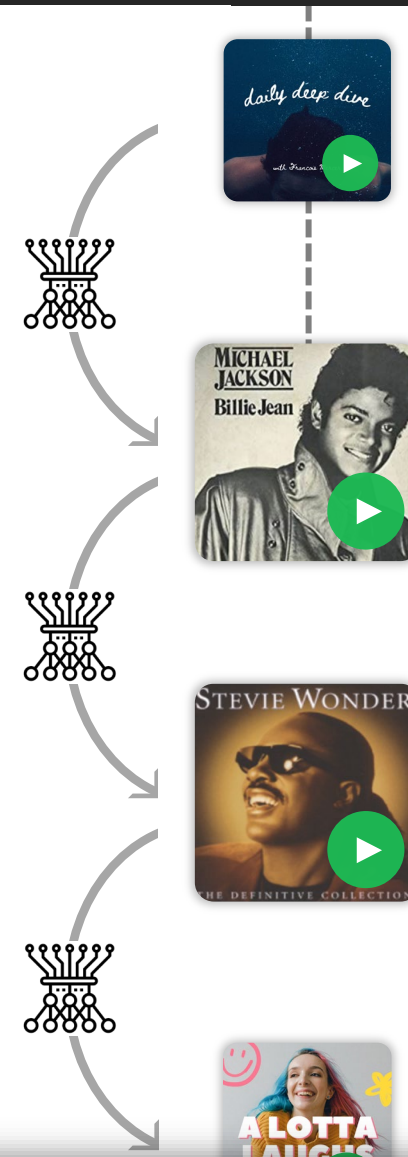
노래와 노래 사이에 공백을 없애 플레이리스트를 끊임없이 재생하는 기술

1 무한 재생, 새로운 음악 재생 방식 제안

클래식, 라이브, 믹스 앨범 등을 감상할 때 노래 간의 공백은 몰입도를 떨어뜨리고 괴리감을 느끼게 할 수 있다.

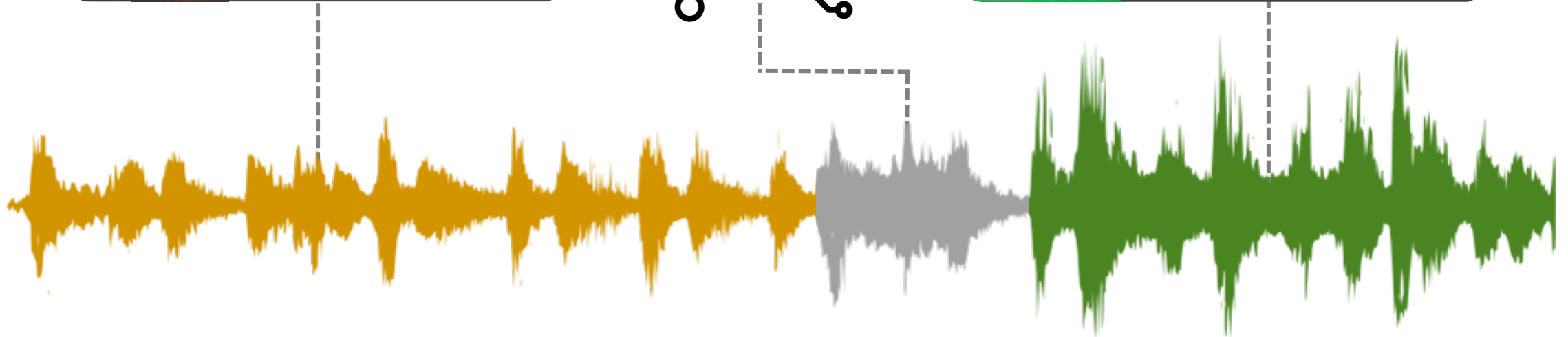
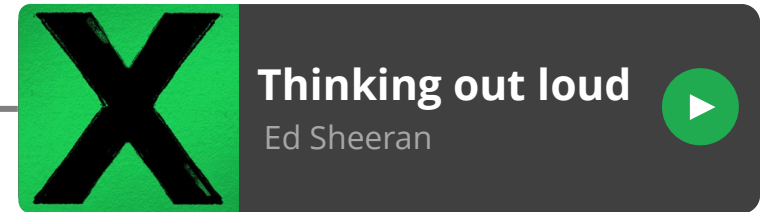
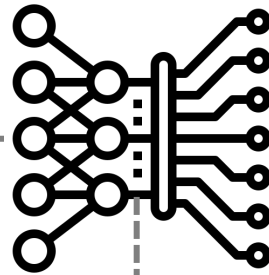
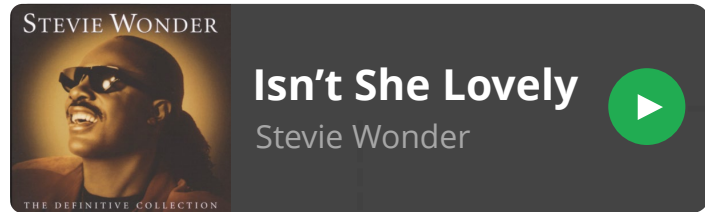
2 카페 등 배경음악이 계속 재생되어야 하는 장소에 활용 가능

음악 사이 공백으로 인한 거슬림을 제거함으로써 흐름을 끊지 않아 집중력을 요하는 상황에 효과적이다.



Goal

두 음악을 자연스럽게 잇는 연결 부위 작곡



Mash-up

Related-Work

Mash Up

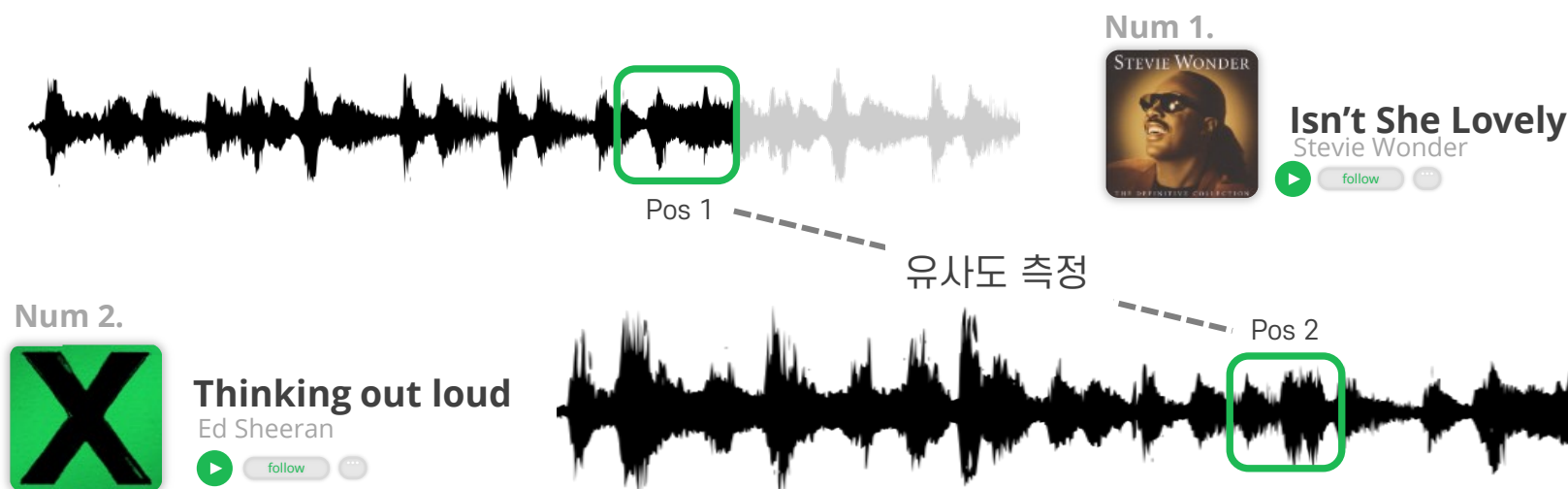
서로 다른 음악을 하나로 믹스하는 방법

Mix Point

Mashup을 구성하는 음악에서 믹스할 지점

Mix Point를 구하는 기존의 방법

1. 노래의 마디 시작 부분을 찾아서 곡 분할 포인트를 구함
2. peak highlight와 beat의 유사성을 분석하여 Mashup Point를 결정
3. novelty function으로 새로운 부분 가중치



Transition

Related-Work

음악의 특징

- 시간과 연관된 동적인 예술
- 다양한 악기와 트랙으로 구성됨
각 트랙은 동적인 고유의 시간적 특징을 가지고 있으면서도 상호의존적
- 이산적인 Musical event의 Sequence
- 미세한 변화에도 민감함

Composition Model

- Waveform 생성 / MIDI 생성
- Neural Net을 이용한 기존의 방법론
VAE, GAN, LSTM, Transformer 등 ...
- MuseGAN
binary multi-track sequence를 생성하기 위해
convolutional GAN을 사용함

→ 음악 도메인 고유의 문법이 존재

U-Net

U-Net 이란?

Biomedical Image Segmentation을 목적으로 제안된
Fully-Connection Network 기반 End-to-End 모델

U-Net의 구조

Contracting path : Context 추출

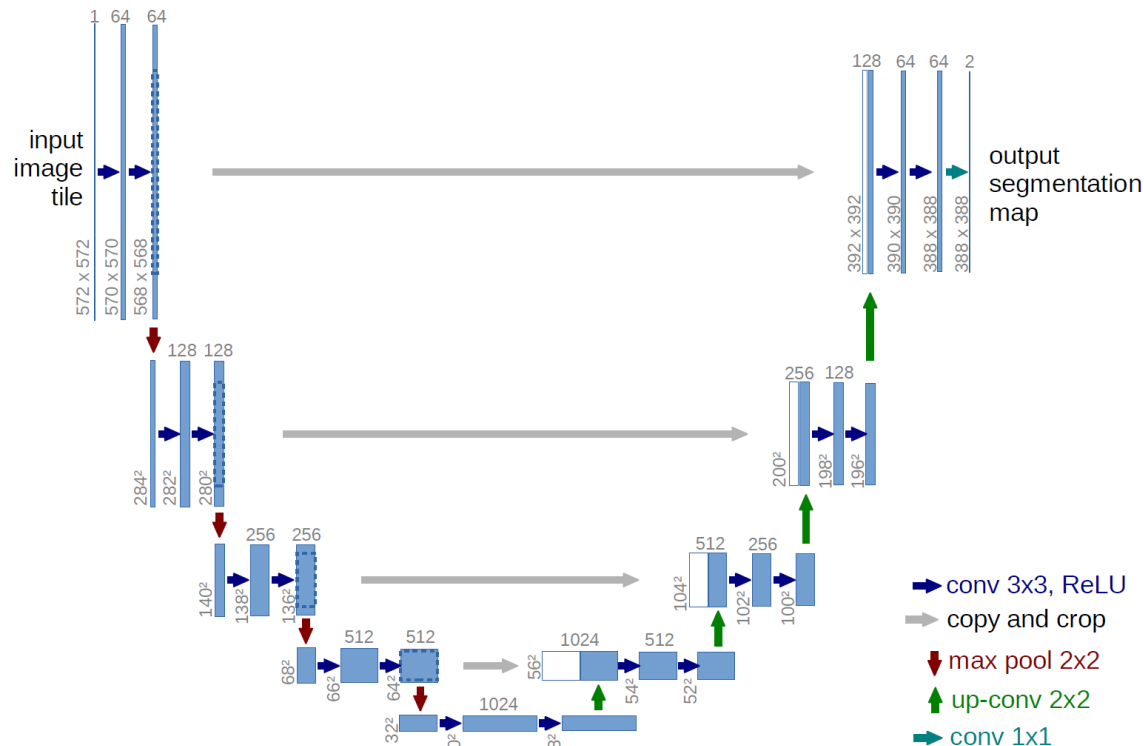
Expanding path : 세밀하고 정확한 Localization

Skip-Connection

각 Expanding Step의 feature map과 대응되는

Contracting path의 feature map을 Concatenation

→ Low-level의 정보를 잘 보존할 수 있음



Dataset

MIDI Data란?

전자 악기끼리 디지털 신호를 주고받기 위해 각 신호를 규칙화한 일종의 규약.

각 미디 파일마다 악기별 연주 방법에 대한 정보가 들어있으며, 재생 시 각 악기에 해당하는 가상악기를 적용시킨다.

Dataset

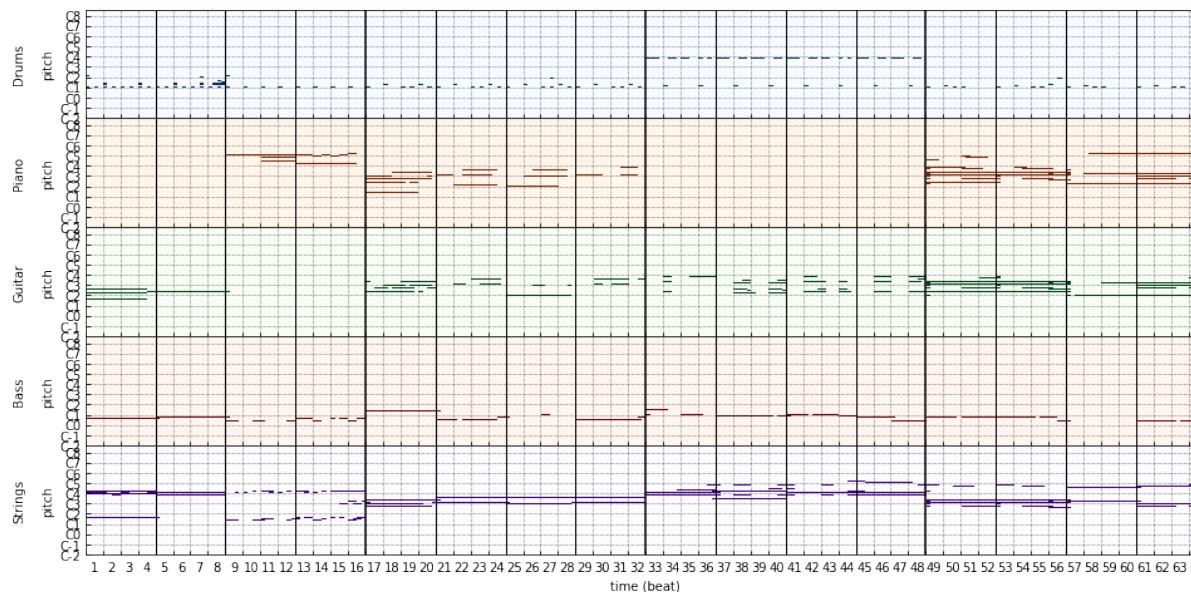
Lakh Pianoroll Dataset

Train Dataset

팝 음악으로 이루어진 Million Song Dataset에 기반한
멀티트랙 미디 데이터

POP Dataset

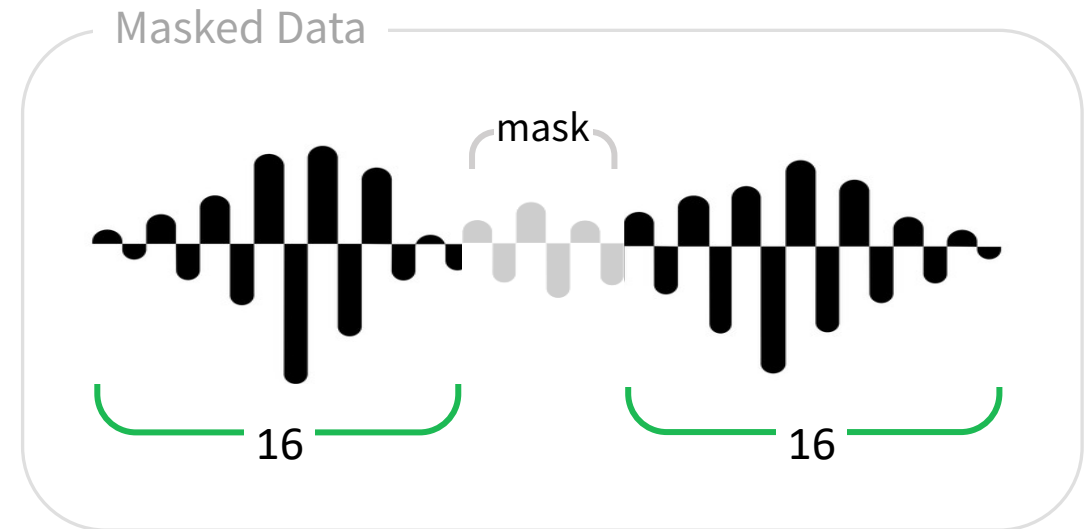
Inference Dataset



Dataset

Data Processing

- 5개의 트랙으로 구성
Drums, Piano, Guitar, Bass, Strings
구성 악기에 해당하는 트랙으로 전처리
- 각 데이터의 가운데 두 마디를 Masking
- Tempo : 100
- Input shape : 5 (track) * 72 (pitch) * 32 (16+16)



Mash-up

Waveform ↔ MIDI

Waveform

많은 정보들이 신호 형태로 축약
→ Downbeat tracking 등 추론 작업이 필요함

MIDI

추론 작업 없이 미디 파일의 정보를 모두 사용 가능

Methods

1. Dynamic Time Warping (DTW)
2. 마디의 분포 간 유사도 Jensen-Shannon Divergence

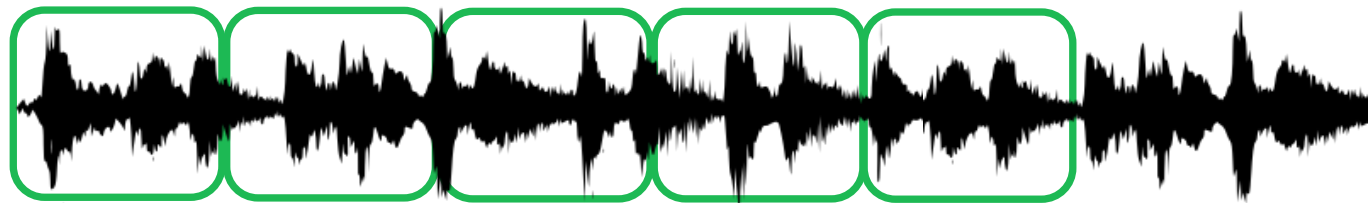
키 별 & 마디 별 분할



Thinking out loud Ed Sheeran



follow



Mash-up

키 별 & 마디 별 분할

- MIDI 파일의 QPM를 사용해 downbeat 위치를 구함
- 4/4 박자로 통일하였으므로 4개의 quarter note 를 한 마디로 추정하여 분할

유사도 측정을 위한 설정

- 미디 파일에 포함된 음의 분포 정보 이용
- 멜로디 분포 사용

자연스러운 Mashup을 위해서는 비교하는 두 마디의 화음이 같아야 한다.
화음은 게이름의 분포로 정의된다.
- 앞 노래는 초반일수록 뒤 노래는 후반일수록 가중치를 부여

비어 있는 마디와 음 분포가 sparse한 곡 초기 마디의 영향을 줄인다.

Dynamic Time Warping

DTW



- 두 개의 시계열 데이터 간의 유사도를 측정하기 위한 알고리즘 중 하나
- two time series에서의 temporal distortions을 추출하여 global alignment를 optimal하기 위한 방법
- $\min(D[i-1, j-1], [i-1, j], [i, j-1]) + |Reference_i - Query_j|$
위의 matrix를 통해 최소값을 갖는 Warping Path를 구하여 Mix Point를 추출함
- 문제점 : mixed MIDI와 original MIDI를 추출하여 Warping Path 구해서 일일이 확인해야 함

분포 기반 유사도

KL-Divergence & Jensen-Shannon Divergence

1. KL-Divergence

$$D_{\text{KL}}(P \parallel Q) = \sum_{x \in \mathcal{X}} P(x) \log \left(\frac{P(x)}{Q(x)} \right)$$

유사성 척도 : 분포 간 얼마나 가까운지에 대한 정보 손실량의 기댓값

Symmetry 하지 않아 분포 간 거리를 척도로 사용할 수 없음 → 잘못된 계산

2. Jensen-Shannon Divergence

$$JSD(P, Q) = \frac{1}{2} D(P \parallel M) + \frac{1}{2} D(Q \parallel M)$$

where $M = \frac{1}{2} (P + Q)$

P와 Q의 평균값 M과 KL-Divergence를 계산 → Symmetry

∴ 두 확률 분포 간 거리를 척도로 계산 가능

Smootify

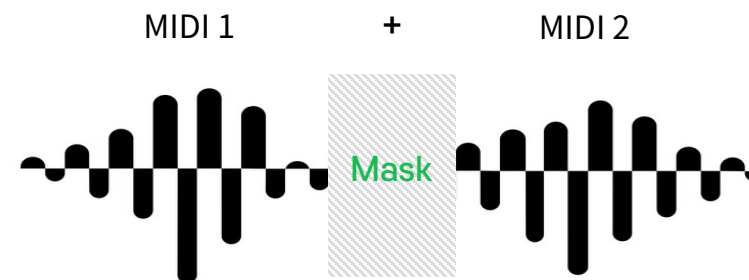
U-Net을 사용한 이유

- 픽셀의 변화에 둔감한 이미지와 달리 음악 및 음성은 미세한 변화에 거슬리는 artifact가 생성됨
- U-Net은 Low-level을 잘 보존하므로 본 음악 생성 프로젝트에 적합함

Model

- MuseGAN의 convolution 구조 차용
- 6 층의 Convolution Block으로 구성된 U-Net 구조
- 각 노래의 가운데 두 마디를 Masking 한 후, 이를 예측하도록 학습
- Loss : MSE

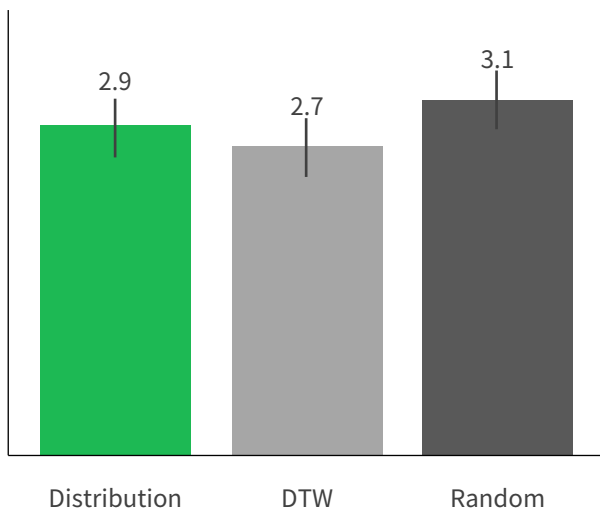
데이터 입력



Results

Mash-up

MOS Mean Opinion Score



5곡의 순열로 생성된 30곡 중 무작위 10곡에 대해
5점 척도로 평가

표준편차

	Distribution	DTW	Random
표준편차	0.477	0.447	0.680

결과

미미한 차이로 Random method의 MOS가 가장 높지만 표준편차도 가장 크다.

→ 3가지 Methods의 성능은 비슷하다.

원인

- 표본의 크기가 작음
- Mix Point를 찾기 위한 방법론은 Rule-based 이므로 한계점이 존재

Results

Transition

		MuseGAN		Smootify
		Generator	Discriminator	
Memory	Total params	359,823	204,321	619,696
	Params size	1.44 (MB)	0.82 (MB)	2.48 (MB)
	Total mult-adds	1.70 (G)	39.65 (M)	258.22 (M)
	Estimated Total Size	108.74 (MB)	13.73 (MB)	29.71 (MB)
Computation	Forward/backward pass size	106.56 (MB)	11.44 (MB)	26.49 (MB)

→ MuseGAN 보다 파라미터 수는 많으나 계산 속도가 빠르다

Conclusion

의의

음악의 고유한 특성으로 인한 어려움을 극복하고 자연스러운 곡 전환을 이루어 냄

결과

Mashup Point 찾기

DTW / Distribution

→ Random을 포함한 3가지 방법론의 성능이 비슷함

Bridge 생성

U-Net 구조의 Transition 모델

→ 계산 속도가 빠름

향후 연구

Mix Point를 찾는 방법 개선

- 장르를 구분하여 간주 생성 고도화
- 화성학 이론을 도입하여 음악적 유사도 찾기

다양한 형식의 음악을 MashUp

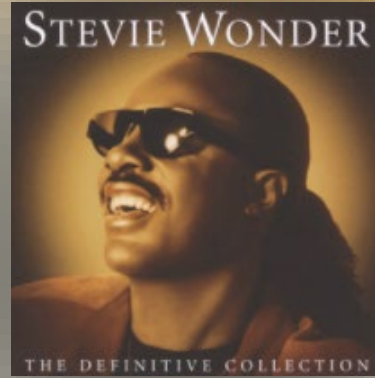
- 더 다양한 트랙으로 구성된 MIDI 파일
- Wave 파일

Demo

- Background
- Related Work
- Dataset
- Methods
- Results
- Conclusion
- Demo
- Reference



+



Billie Jean + Isn't She Lovely



FOLLOWING



제목

- 1 **Billie Jean**
Michael Jackson
- 2 **Isn't She Lovely**
Stevie Wonder

Reference

- MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment, 2017 (Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, Yi-Hsuan Yang)
- U-Net: Convolutional Networks for Biomedical Image Segmentation, 2015 (Olaf Ronneberger, Philipp Fischer, Thomas Brox)
- MIDI-VAE: Modeling Dynamics and Instrumentation of Music with Applications to Style Transfer, 2018 (Gino Brunner, Andres Konrad, Yuyi Wang, Roger Wattenhofer)
- AutoMashUpper: Automatic Creation of Multi-Song Music Mashups, 2014 (Matthew E. P. Davies, Philippe Hamel, Kazuyoshi Yoshii, Masataka Goto)
- [salu133445/musegan](https://salu133445.github.io/musegan)
- [salu133445/lakh-pianoroll-dataset](https://salu133445.github.io/lakh-pianoroll-dataset)
- [brunnerfino/MIDI-VAE](https://brunnerfino.github.io/MIDI-VAE)



TobigTonix



15기

권오현

음악만이 나라에서 허락하는
유일한 마약..

FOLLOW



15기

김태희

악기 연주를 넘어
컴퓨터로 음악 생성하기 ♪

FOLLOW



15기

안민준

음악과 함께 투빅스를 졸업해서
너무 행복합니다.. 엉엉

FOLLOW



15기

황보진경

스무스~

FOLLOW



16기

김윤희

음~ 악!

FOLLOW

SHUFFLE PLAY



Thanks for tuning in
Sound Conference