

음을 그리다

1. 작품 정의

- 1) 사용자에게 작곡의 편의성을 제공하는 Android 애플리케이션 서비스로, 사용자로부터 'Humming' 데이터를 입력받아 다음과 같은 서비스를 제공한다.

■ 악보 생성

추출한 음성 데이터를 일정한 박자 단위로 변형하여 실제 오선지 상에 기보

■ 음악 검색

Deep Learning을 통해 음성 데이터를 기존의 곡들과 비교하여 목소리만으로 음악 검색

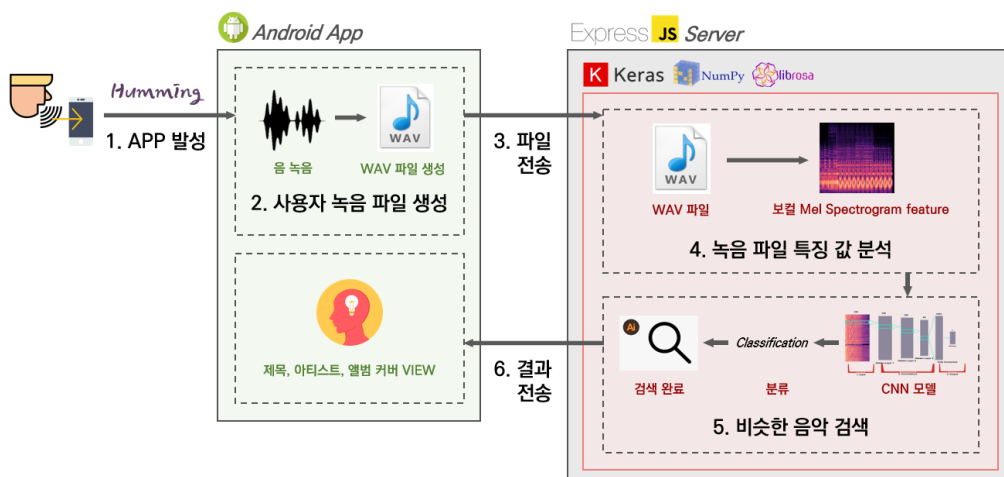
- 2) 음악적 초보자로 하여금 작곡의 진입장벽을 낮추어 자신의 음악을 만들 수 있도록 한다.
- 3) 만든 음악과 유사한 곡이 있는지 검색함으로써 저작권을 확보하고, 이로 인해 사용자가 수익을 창출할 수 있도록 함을 궁극적인 목표로 한다.

2. 핵심 기능 흐름도

2-1. 악보 생성



2-2. 음악 검색



3. 핵심 기술

3-1. 박자 알고리즘

- 1) 짧은 시간 내에 급격하게 한 옥타브 위, 아래로 **튀는음**을 결합한다.
- 2) 같은 코드가 반복되는 수를 세어 **박자값**을 생성한다.
- 3) 음이 시작되기 전과 끝날 때 함께 저장된 **공백음 'Z'**를 제거한다.
- 4) **떨림음**을 정형화시킨다.
 - a. 반음 차이나고 박자값이 1로 매우 짧게 들어온 음을 이웃하는 음으로 변경한다.
 - b. 같은 음이 연이어서 나타난 경우, 각 음의 박자값을 더해 하나의 박자값으로 결합한다.
- 5) **음표를 생성**하기 위해 박자값을 각 음표에 맞는 수치로 변경한다.
 - a. 박자값 1은 매우 짧은 음이므로 앞 단계에서 제거되지 못한 음이라고 판단하고 제거한다.
 - b. 박자값 2는 8분음표, 3-4는 4분음표, 5-6은 2분음표, 7-8은 온음표를 위한 박자로 사용한다.
- 6) **악보 기호**를 추가한다.
 - a. 박자가 8이 되면 마디를 나누는 기호 '|' 를 추가하여 마디를 나눈다.
 - b. 같은 음이 마디로 나뉘게 된다면, 기호 '-' 를 추가하여 이음줄을 나타낸다.

3-2. 딥러닝 모델 (CNN)

■ Train Data Set

음원 보컬 추출 데이터, 허밍 데이터에 대한 Mel Spectrogram

■ 모델 및 하이퍼 파라미터 구성

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 215, 503, 16)	208
max_pooling2d_1 (MaxPooling2D)	(None, 107, 251, 16)	0
conv2d_2 (Conv2D)	(None, 107, 251, 32)	2048
batch_normalization_1 (Batch Normalization)	(None, 107, 251, 32)	128
activation_1 (Activation)	(None, 107, 251, 32)	0
max_pooling2d_2 (MaxPooling2D)	(None, 53, 125, 32)	0
conv2d_3 (Conv2D)	(None, 53, 125, 64)	8192
batch_normalization_2 (Batch Normalization)	(None, 53, 125, 64)	256
activation_2 (Activation)	(None, 53, 125, 64)	0
max_pooling2d_3 (MaxPooling2D)	(None, 27, 63, 64)	0
conv2d_4 (Conv2D)	(None, 27, 63, 128)	32768
batch_normalization_3 (Batch Normalization)	(None, 27, 63, 128)	512
activation_3 (Activation)	(None, 27, 63, 128)	0
max_pooling2d_4 (MaxPooling2D)	(None, 14, 32, 128)	0
conv2d_5 (Conv2D)	(None, 14, 32, 256)	131072
batch_normalization_4 (Batch Normalization)	(None, 14, 32, 256)	1024
activation_4 (Activation)	(None, 14, 32, 256)	0
max_pooling2d_5 (MaxPooling2D)	(None, 7, 16, 256)	0
flatten_1 (Flatten)	(None, 28672)	0
dense_1 (Dense)	(None, 10)	286730
Total params: 462,938		
Trainable params: 491,978		
Non-trainable params: 960		

Input : 216 X 504 X 3 (Mel Spectro Graph Image)

Hidden Layer : 13

Drop Out : 0.2

Kernel Size : 2 X 2

Optimizer : Adam

Learning Rate : 0.00005

Loss Function : Categorical Cross-Entropy

Epochs : 30

Batch Size : 20

4. 활용 분야

작곡 지원	청음이 되지 않거나 음악적 지식이 없더라도, 쉽게 악보 기보를 가능하게 함으로써 음악 창작의 열의를 충족시킨다.
음원 소스 다양화	다수 사용자의 작곡을 통해 음원 소스 다양화 및 저작권 확보와 이로 인해 발생하는 사용자의 수익 창출에 기여한다.
음원 비교 고도화	음원으로만 음악 검색이 가능한 현재의 서비스에서 진보한 서비스를 제공한다.