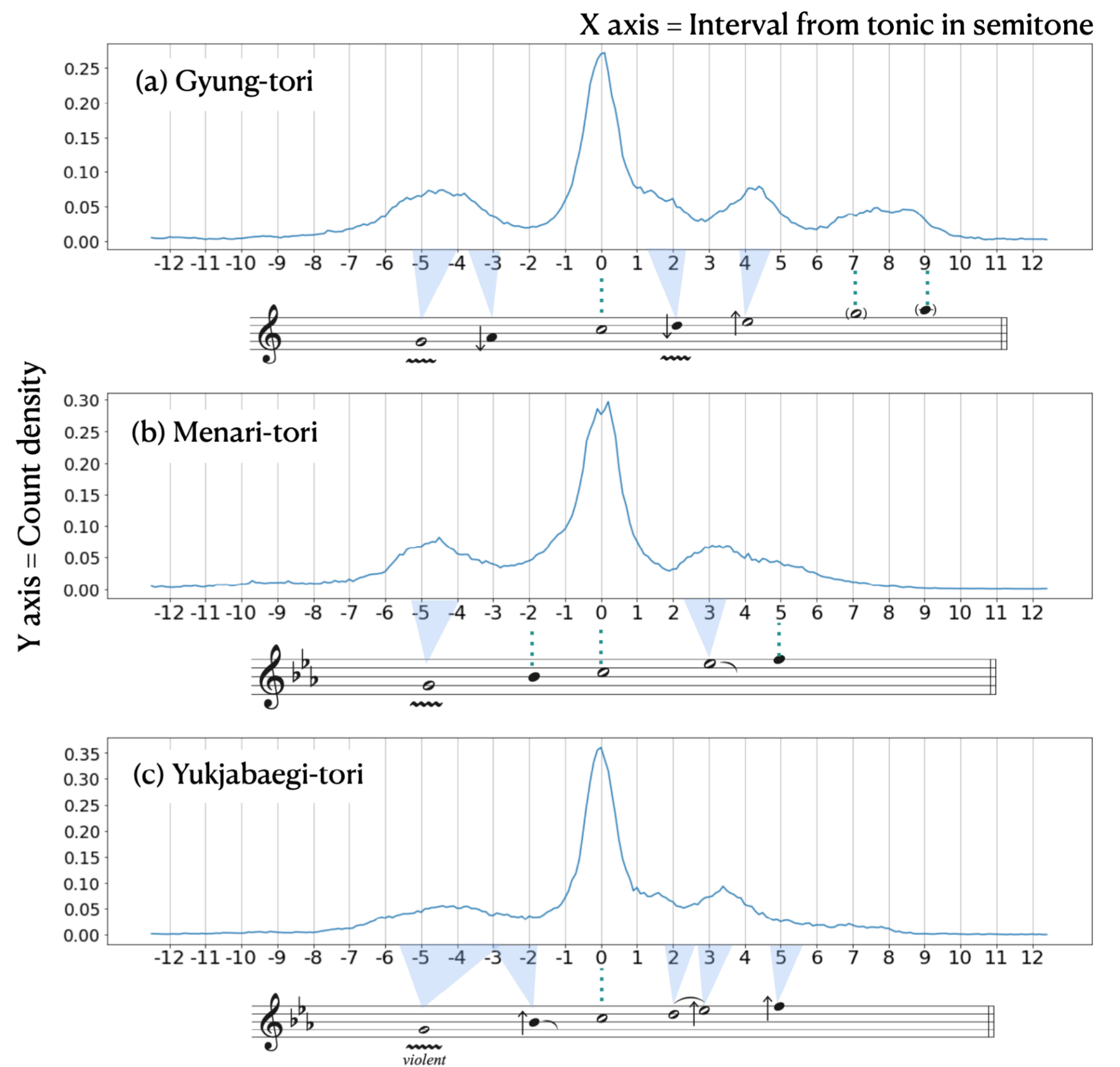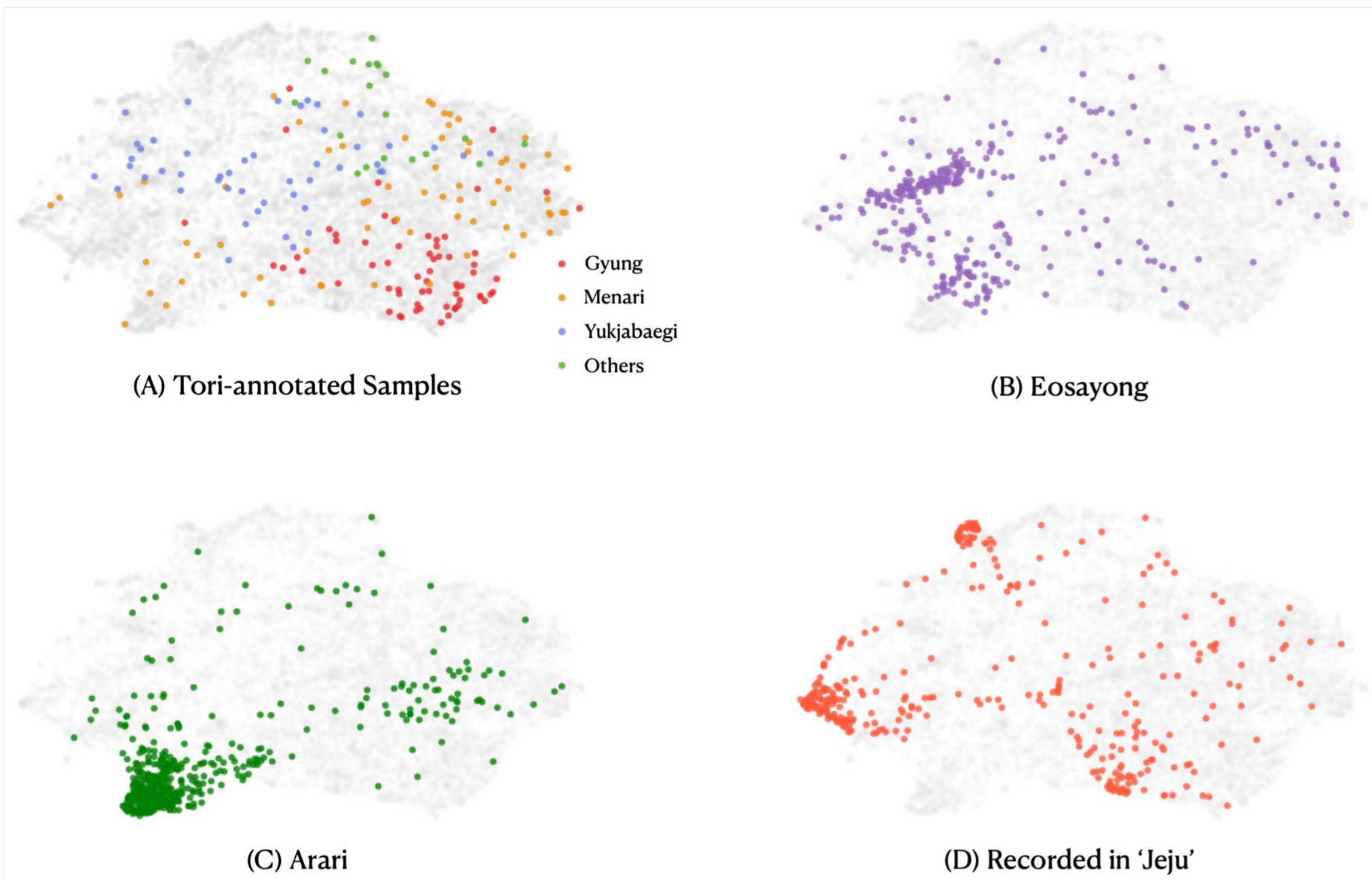# Finding Tori:
Self-supervised Learning for Analyzing Korean Folk Songs

Danbinaerin Han[1], Rafael Caro Repetto[2], Dasaem Jeong[1]

1) Music & Art Learning Lab, Department of Art & Technology in Sogang University, 2) Institute of Ethnomusicology, Kunstuniversität Graz

Our goal is to validate Korean traditional musicology theories through the lens of MIR, with vast amount of audio field recording of folk songs. The results show inherent link between 'tori' (토리) and the distinctive characteristics of each folk song.

(A) Tori-annotated Samples
- Gyung
- Menari
- Yukjabaegi
- Others

(B) Eosayong

(C) Arari

(D) Recorded in 'Jeju'

X axis = Interval from tonic in semitone

(a) Gyung-tori

(b) Menari-tori

(c) Yukjabaegi-tori

Y axis = Count density

violent

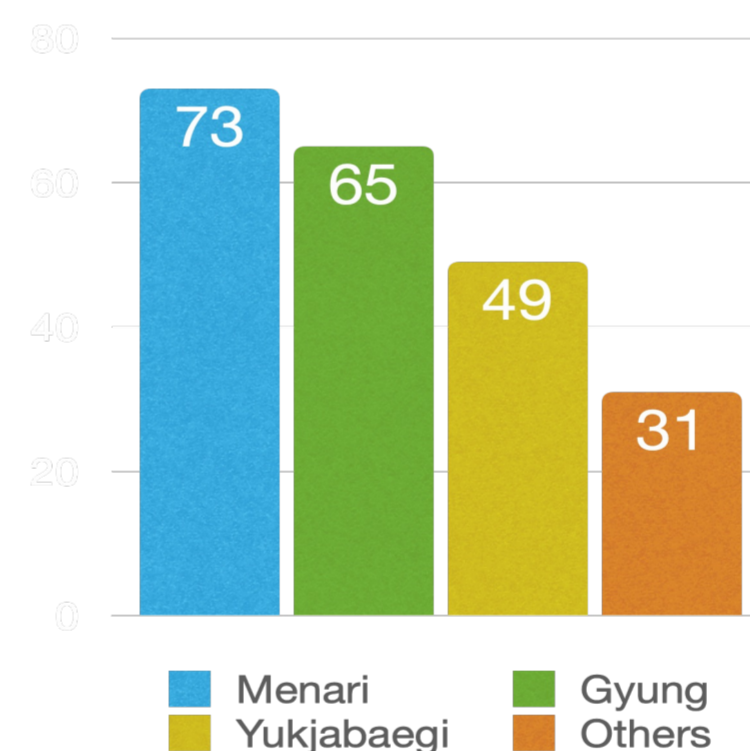## Dataset

<Anthology of Korean Folk Songs>
- A Large-scale audio collection comprising 15,861 Korean folk song in South Korea
- Approximately 700 hours
- Munhwa Broadcasting Corporation (MBC)
- 9 regions, 153 cities, and 1,010 villages
- Detailed information : recorded location, singer, recorded date, lyrics, etc.

We present our research for aligning lyrics at DLfM! (Friday)

Methodology

0. Data preprocessing

- Filtered by Sound event Detection(Choir/Percussion)
- Pitch Contour (CREPE)
- Masking under 0.8 confidence score
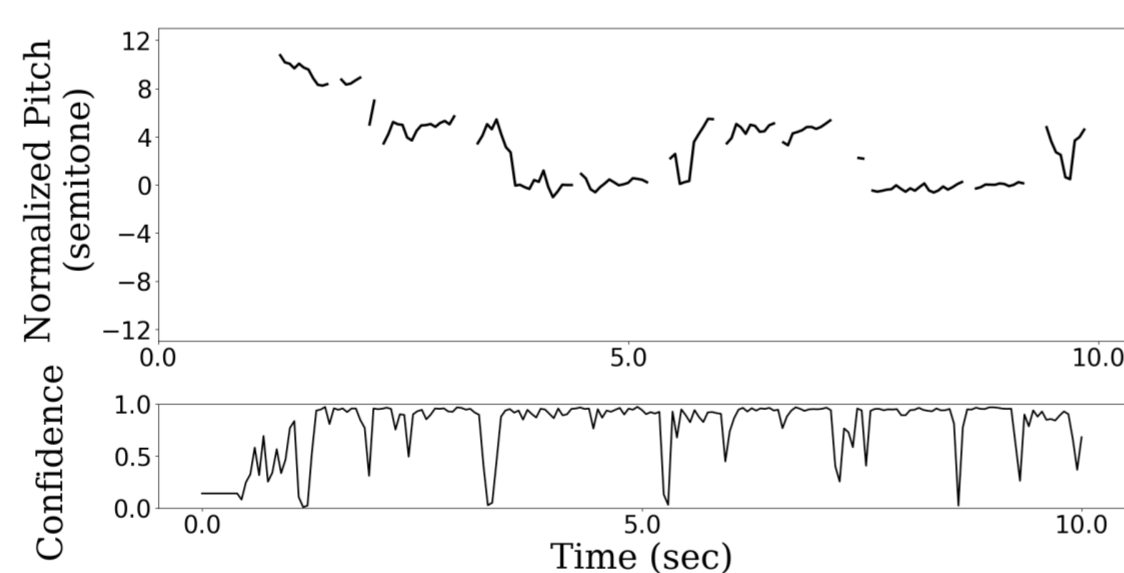
1. Pitch Histogram

- Normalize by adjusting the pitch center

2. Self-supervised CNN

We assumed that pitch contours from the same song share the same musical characteristics, while those from different songs do not.

We applied contrastive learning to cluster similar embeddings together and separate dissimilar ones.

Widen
Narrow

Context attention
Max pooling (size 5 or 4)
ReLU
Batch normalization
1D CNN (kernel size 3)
× 4 layers

Sampled from a single song   Sampled from other songs

Audio 1 (**Tori A**)
30 sec   30 sec
Tori 토리 A   Tori 토리 A

Audio 2 (**Tori B**)
Sample 3 **(Negative)**
30 sec
Tori 토리 B

Tori

The system for classifying the musical characteristics of regional folk songs. It contains primary notes, ornamentation, ending notes and idiomatic melodic patterns.
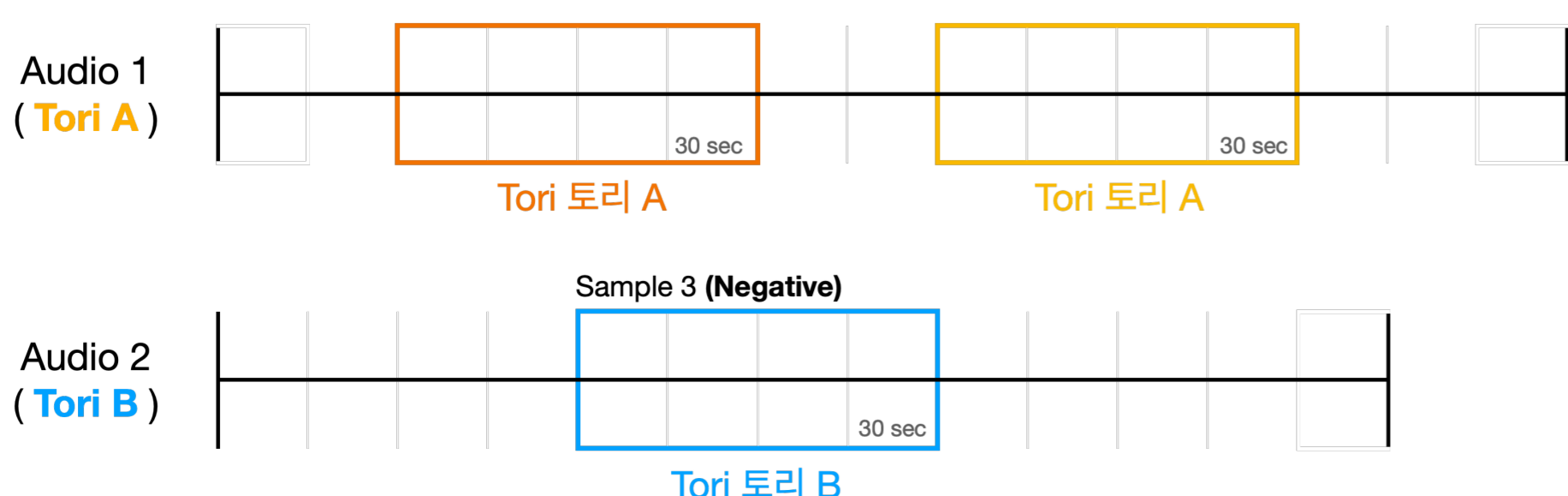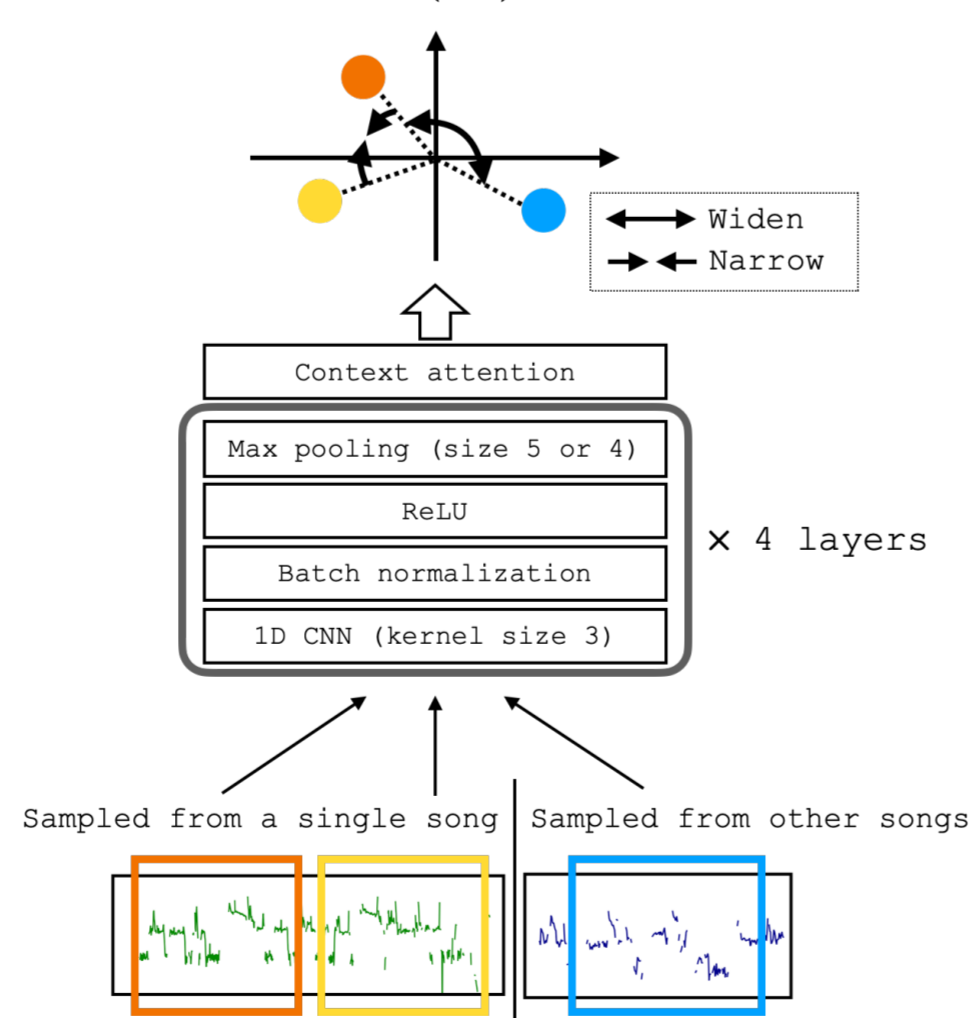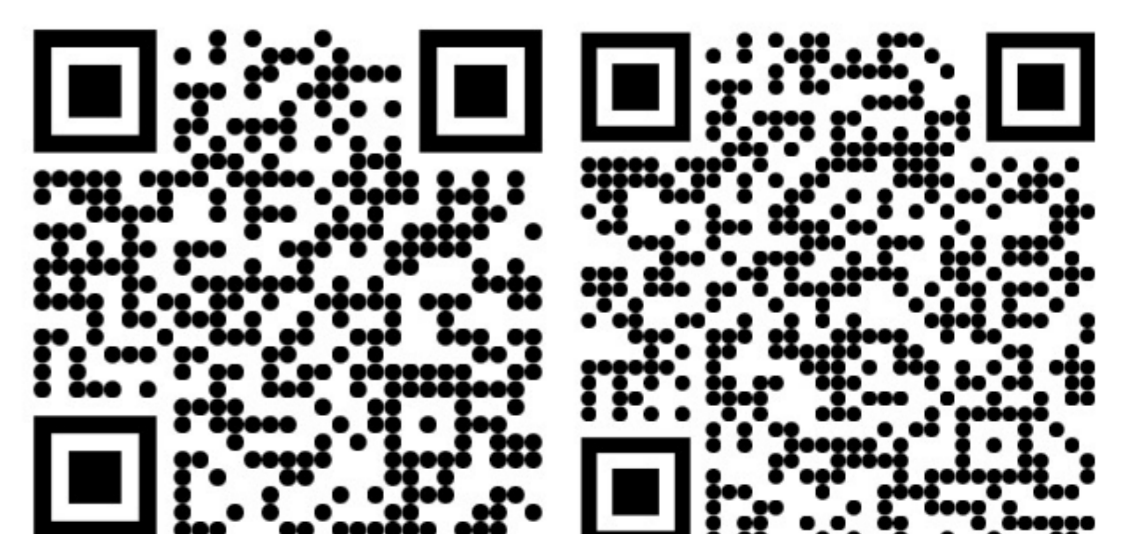

Susimga-tori
Menari-tori
Gyung-tori
Yukjabaegi-tori
Korean Peninsula

73   65   49   31
- Menari
- Yukjabaegi
- Gyung
- Others

Since there were **no** tori annotations in the dataset, we manually labeled using **4** distinct categories.

Evaluation

Embedding	Similarity Ranking nDCG	Random Forest Classifier Accuracy
Hist. (25bin)	0.783	0.744 ±0.058
Hist. (124bin)	0.777	0.722 ±0.054
CNN (region)	0.792	0.634 ±0.055
CNN (SSL)	0.853	0.848 ±0.039

Table 1. Experiment results on the tori subset.

Acknowledgements & Links

Git hub   **Visualization Demo**