SongNet: Real-time Music Classification

Chen Chen, Chi Zhang, Yue Zhang

{chenc2, czhang94, yzhang16}@stanford.edu

Overview

❑ Motivation

✓ Helps online music companies such as Spotify or Apple Music to mange their music base.

✓ Saves time and effort from manual classification.

❑ Objectives

✓ Build deep learning (C-RNN) models to automatically classify music genres for real time.

✓ Improve baseline models accuracy by C-RNN.

Dataset

❑ Free Music Archive

✓ The dataset contains 8000 tracks of 30 seconds clips, with 8 balanced genres[2] listed below.
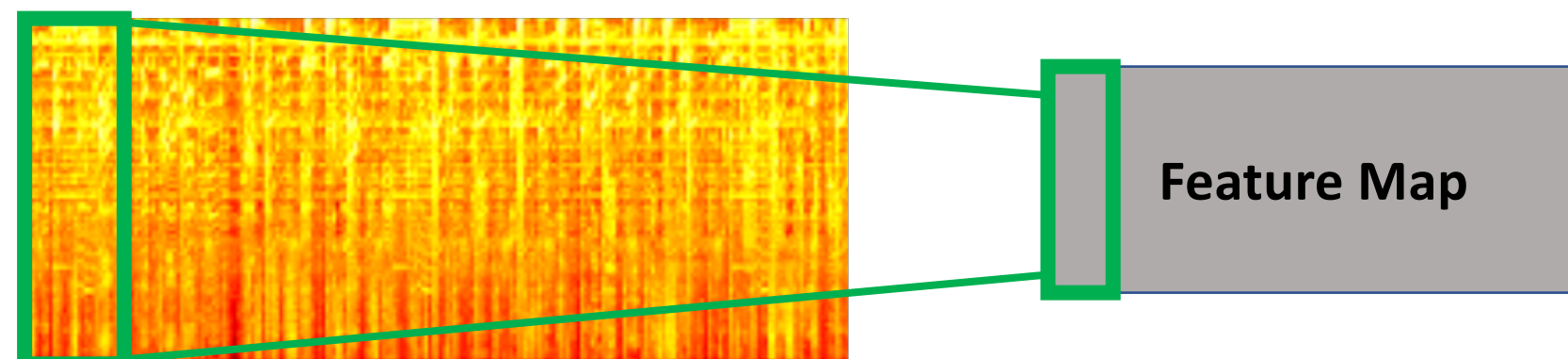
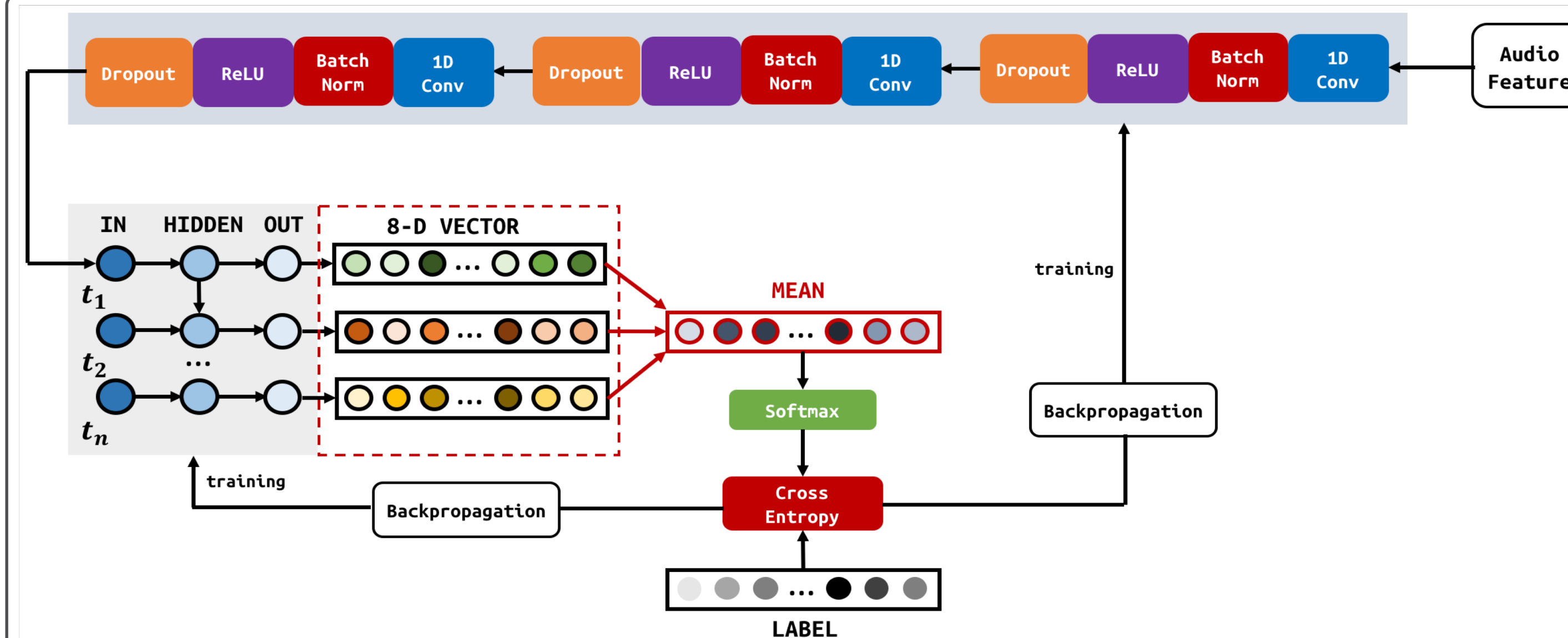| Instrumental	Rock	Electronic
International	Pop	Experimental
Hip-pop	Folk	

✓ 70% training / 20% Validation / 10% Test

Features

❑ Mel-spectrogram

✓ Used librosa[3] package to calculate mel-spectrogram.

✓ Designed CONV layers to extract features (DEMO)

Feature Map

C-RNN Model Architecture

❑ Architecture

✓ Input: mel-spectrogram

✓ 3 Convolutional layers
  ▪ Batch Normalization
  ▪ ReLU activation
  ▪ Dropout Regularization

✓ Recurrent layers

✓ Output: probability of each genre.

Results & Discussion

❑ Results

Models	Accuracy
Random guess	12.5%
K nearest neighbors	36.38%
Logistic regression	42.25%
Multilayer perceptron	44.88%
Support vector machine	46.38%
C-RNN	**65.32%**

❑ Discussion

✓ Genre "Experimental" is hard to be classified correctly.

✓ Classifications of other genres perform well.

✓ CONV layers extracted useful genre clips, listen to our demos.

✓ Recurrent models enable us to do real-time classification.

✓ C-RNN does not include music metadata, while baseline model does.
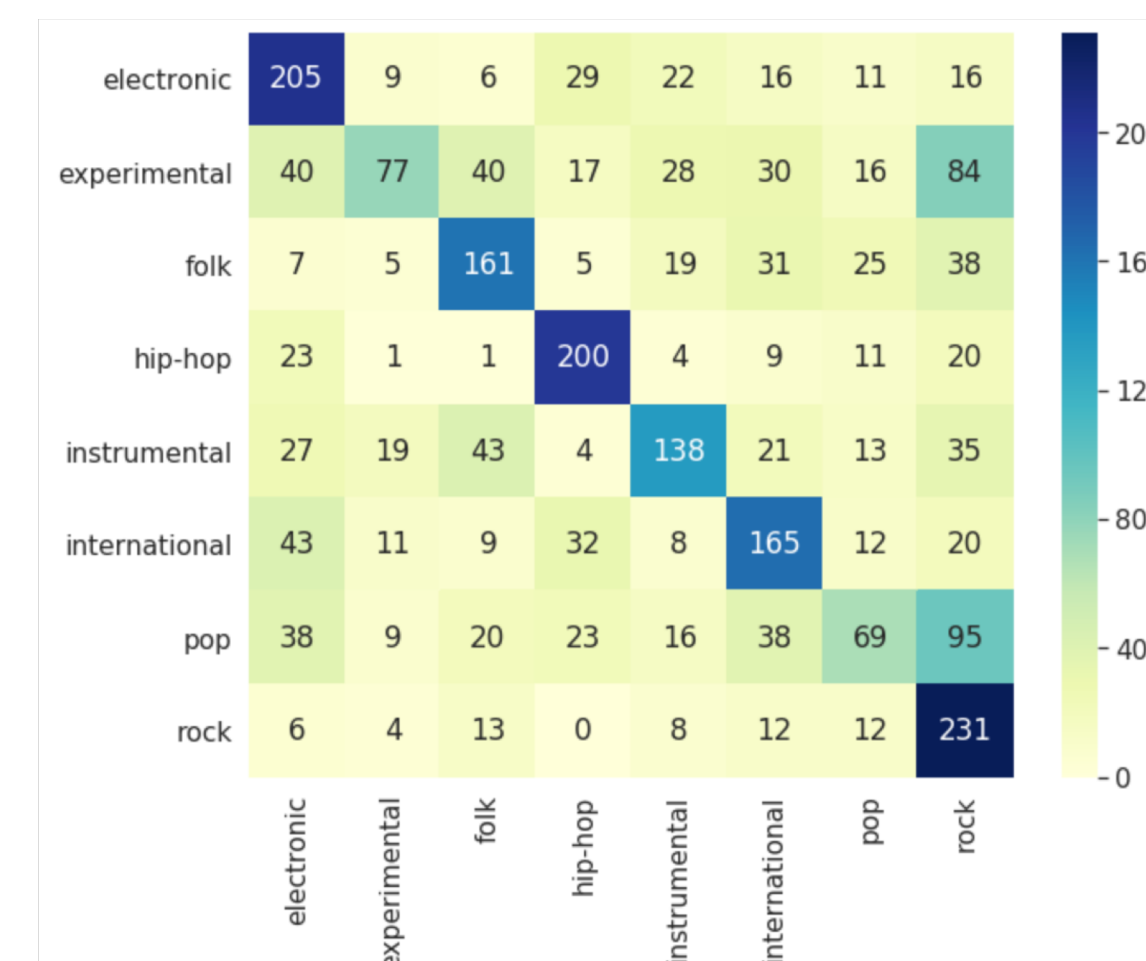
Future

✓ Understand why C-RNN cannot perform well on "Experimental" genre and improve the accuracy of that genre.

✓ Consider adding music metadata to C-RNN models and further improve the accuracy.

✓ Implement an user interface to allow users input a music clip and visualize the real-time music classification online.

Reference

[1] Kozakowski, P., & Michalak, B. (2016, October 26). Music Genre Recognition.

[2] K. Benzi, M. Defferrard, P. Vandergheynst, and X. Bresson. (2016). FMA: A dataset for music analysis.

[3] LibROSA (n.d.). http://librosa.github.io/librosa/