

DEEP LEARNING AND ITS APPLICATIONS

PROJECT PRESENTATION ON DEEPSYMPHONY

GROUP-18

Akhilesh Devrari, Chirag Vashist, Prabhakar Prasad



March 20, 2019

1 Introduction

- Motivation
- Problem Statement
- Work Done

2 Data-sets

3 Methodology

- Methodology explored
- Proposed Methodology

4 Results and Discussion

- Results and Discussion

Motivation

Deep learning is disrupting every field of technology. A particular area of interest is fusion of mathematics and arts. DeepSymphony is our experiment to see if deep learning can be used to produce music that can accompany a given vocal track given by user, and produce a song out of it.

Applications:

- Fun activity for amateur singers and every music lover.
- Can be even used by professionals as creative suggestions to produce new melodies and songs.

Problem Statement

- To train a deep learning model which can produce a complete song taking only song vocals as input.
- We expect DeepSymphony to be able to:
 - Take input audio file which will contain vocals of a singer.
 - Produce background music to accompany input vocal track.
 - Mix vocals with produced music and output a complete song.

Work Done

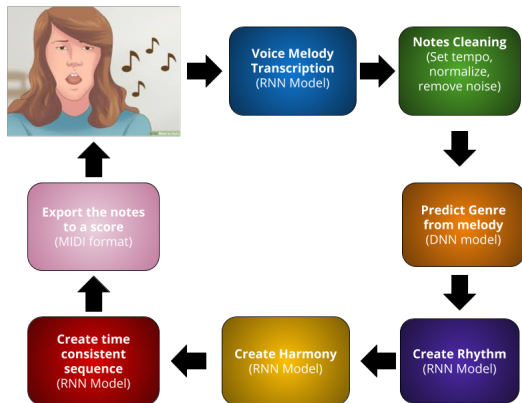
- In order to understand the problem, we have read on how music is artificially produced using Conditional Generative Adversarial Networks.
- Next, we looked into how music files are stored.
- We also looked into vocals to wordless melodies, which would later be used to find common sequences to align chord progressions.
- Furthermore, a basic understanding of keys, beats, scales, notes and chords was acquired. This would be crucial for training the network to create music that matches with the vocals.

Data-sets

- For data sets, both only vocals and vocals+music data are needed.
- We also need to distinguish different type of instruments used in music.
- For that, we need data sets of individual instrument's instrumental music for a particular vocal song.
- One very good source we found is <https://sisec.inria.fr/sisec-2016/2016-professionally-produced-music-recordings/>
- We came across this [research paper](#) on MuseGAN which has used [MuseGAN Data](#) that is actually derived from [Lakh Midi Dataset](#)

Methodology explored

- We came across different open-source projects and research papers.
<https://github.com/ybayle/awesome-deep-learning-music>
- Another open-source web application we come across is **Hum2Song**
<https://github.com/toxtli/hum2song>



Proposed Methodology

The following pipeline has been proposed:

- ① **Extract Melody** : Remove words from vocals to get underlying melody.
- ② **Sequence Matching** : Obtain repetitive bits within vocal sequence for placing riffs and chord progressions.
- ③ **Pass data to C-GAN with keys**
- ④ **Post-process music** : Tune sound characteristics to add desired effects to the music.

Challenges:

The principal challenges faced would be producing music that actually sounds good. Most of the research on producing music has been producing instrumental music. In our scenario, we have the additional constraints of matching the music with the input vocals.

Results and Discussion

We have explored a variety of models and architectures.

- ① **Extracting Melody from Voice** : Using a tool called **Audio-to-Melodia**, we are able to extract the melody from a song. This was one of our main objectives.
- ② **Exploring Models** : We have explored many state-of-the art models for the project. This includes **MuseGAN** and **CocoNet**.