Music Composer Classification Using Deep Learning

Final Project

Paul Parks

University of San Diego

AAI511 - Neural Networks and Deep Learning

Professor: Kahila Mokhtari

August 11, 2024

Table of contents

Introduction

This report details the methodology, data preprocessing steps, feature extraction techniques, model architecture, and training process used to classify musical compositions by 4 famous composers: Bach, Beethoven, Chopin, and Mozart. I utilized LSTM and CNN models to achieve this classification, with the CNN model outperforming the LSTM model. The report aims to provide a comprehensive and reproducible guide for future reference.

Classifying musical compositions by composer is challenging and requires understanding and analyzing complex musical features. This project aims to classify compositions by Bach, Beethoven, Chopin, and Mozart using deep learning techniques. The dataset consists of MIDI files, and the methodology includes data preprocessing, feature extraction, model building, and training.

Data Collection

The dataset used in this project is sourced from Kaggle. It contains MIDI files of compositions by Bach, Beethoven, Chopin, and Mozart, organized into directories named after the composers. Although the dataset contains MIDI files for several other composers, only Bach, Beethoven, Chopin, and Mozart were used in this project.

Data Preprocessing

MIDI files were loaded and labeled based on the composer. The dataset is structured with subdirectories for each composer, containing the respective MIDI files. All compositions were transposed to the key of C major to allow the model to predict the composer based on note selection within a key. Several features were extracted from the MIDI files, including key signatures, note histograms, chord histograms, tempos, pitch classes, and durations. Note histograms capture the distribution of notes within the key, while chord histograms represent the distribution of chords. The most common chords were added to the detection algorithm, and all non-common chords were labeled ‘other.’ Tempo changes were extracted to include the speed of compositions. However, many of these classical pieces contained tempo changes, which may be improved in further iterations.

Model

To classify the musical compositions, both LSTM and CNN models were created. The models were trained using the extracted features, with early stopping to prevent overfitting. Hyperparameter tuning was performed using the Keras Tuner. The models were evaluated using accuracy and classification reports.

Results

The CNN model outperformed the LSTM model, achieving an accuracy of 76%. Hyperparameter tuning improved the LSTM model slightly, but the CNN model remained superior. The models struggled with class imbalance, particularly with underrepresented composers.

Key Metrics

- LSTM Accuracy: 64%
    - Precision, Recall, F1-Score for class 0 (Bach): 0.70, 0.93, 0.80
    - Precision, Recall, F1-Score for class 1 (Beethoven): 0.65, 0.19, 0.30
    - Precision, Recall, F1-Score for class 2 (Chopin): 0.41, 0.43, 0.42
    - Precision, Recall, F1-Score for class 3 (Mozart): 0.35, 0.20, 0.25

- CNN Accuracy: 76%
    - Precision, Recall, F1-Score for class 0 (Bach): 0.81, 0.97, 0.88
    - Precision, Recall, F1-Score for class 1 (Beethoven): 0.70, 0.49, 0.58
    - Precision, Recall, F1-Score for class 2 (Chopin): 0.65, 0.61, 0.63
    - Precision, Recall, F1-Score for class 3 (Mozart): 0.59, 0.41, 0.49
- Tuned LSTM Accuracy: 66%
    - Precision, Recall, F1-Score for class 0 (Bach): 0.71, 0.94, 0.81
    - Precision, Recall, F1-Score for class 1 (Beethoven): 0.55, 0.21, 0.30
    - Precision, Recall, F1-Score for class 2 (Chopin): 0.52, 0.46, 0.49
    - Precision, Recall, F1-Score for class 3 (Mozart): 0.41, 0.24, 0.30
- Tuned CNN Accuracy: 75%
    - Precision, Recall, F1-Score for class 0 (Bach): 0.88, 0.90, 0.89
    - Precision, Recall, F1-Score for class 1 (Beethoven): 0.82, 0.40, 0.54
    - Precision, Recall, F1-Score for class 2 (Chopin): 0.54, 0.75, 0.63
    - Precision, Recall, F1-Score for class 3 (Mozart): 0.46, 0.59, 0.51

Conclusion

This project successfully demonstrated the feasibility of classifying musical compositions by famous composers using deep learning models. The CNN model showed superior performance compared to the LSTM model. However, there is room for further improvement in feature extraction, model architecture, and handling of class imbalance. Future work can build on these insights to develop more robust and accurate composer classification systems.

In conclusion, this project provides a comprehensive methodology for classifying musical compositions using deep learning. The CNN model achieved the highest accuracy, highlighting its effectiveness. The findings and methodology described in this report offer a foundation for future research and improvements in composer classification.

References

Dataset retrieved from:

https://www.kaggle.com/datasets/blanderbuss/midi-classic-music

Appendix

Project GitHub:

https://github.com/p-parks/AAI-511-FinalProject