Non-chord Tone Identification Using Deep Neural Networks

Yaolong Ju, Nathaniel Condit-Schultz, Claire Authur and Ichiro Fujinaga

McGill University and CIRMMT

SIMSSA | Single Interface for Music Score Searching and Analysis

Overview

This demo addresses the problem of harmonic analysis by proposing a non-chord tone identification model using deep neural network (DNN).

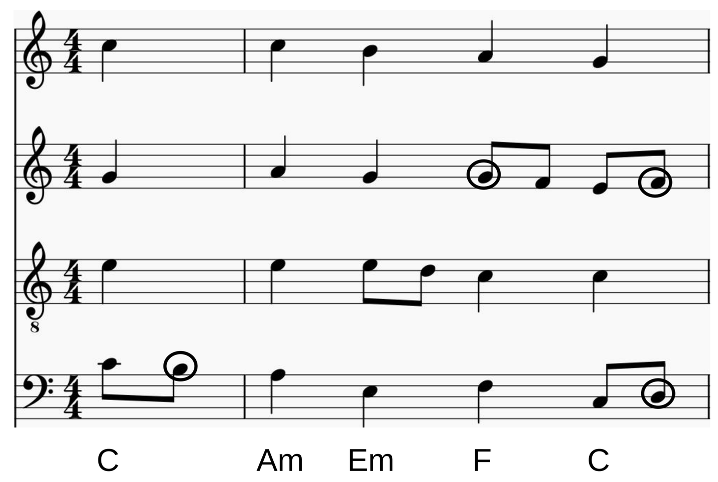
By identifying non-chord tones, the task of harmonic analysis is much simplified.

Trained and tested on a dataset of 140 Bach chorales, the DNN model was able to identify non-chord tones with F1-measure of 72.19% using pitch-class, metric information, and a small size contextual window as input features.

These results suggest that DNNs offer an innovative and promising approach to tackling the problem of non-chord tone identification, as well as harmonic analysis.

Introduction

- Non-chord tones are elaborative notes, usually marked by particular step-wise melodic contours, which do not belong to the local structural harmony (notes in circle).



- Non-chord tone can be used in:
  - ➢ Melodic analysis
  - ➢ Polyphonic music retrieval
  - ➢ Harmonization
  - ➢ **Harmonic analysis**

Motivation

- Machine learning of harmonic analysis is difficult due to the large number of chord classes, which require large amounts of training data to learn (Fig.1 left).
- In contrast, the relatively simple task of non-chord tone identification requires much less training data (Fig.1 right above).
- Once non-chord tones are identified, harmonic analysis becomes a relatively simple task, which can be accomplished by a rule-based algorithm (Fig.1 right below).

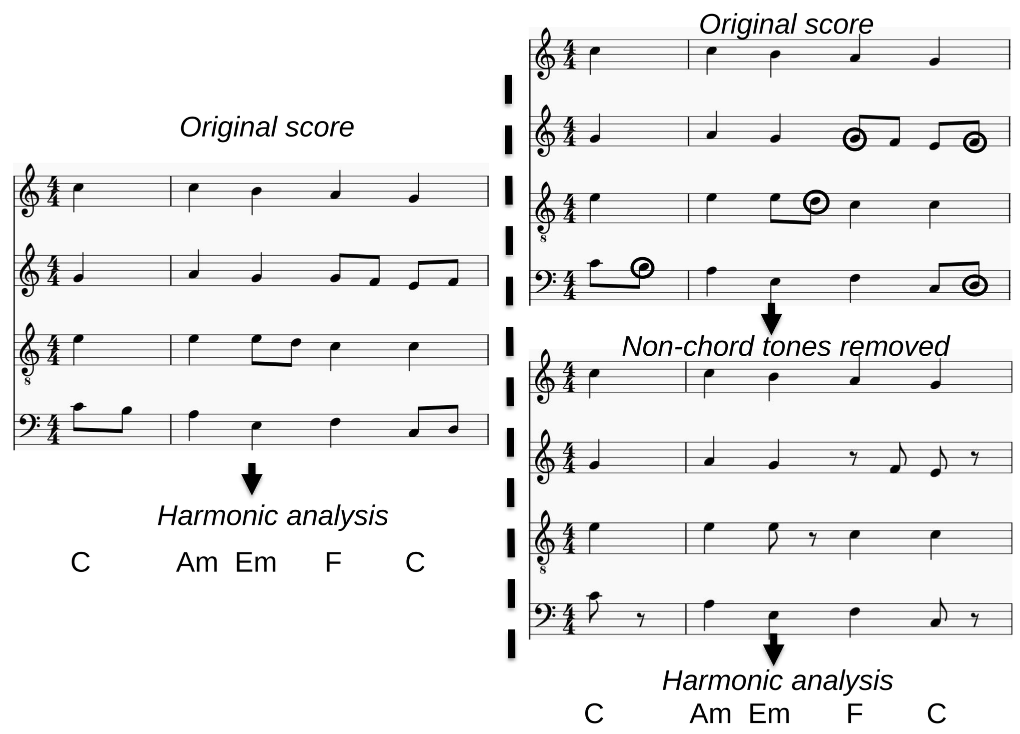


Fig. 1: Direct harmonic analysis (left) and the non-chord-tone-first approach (right).

Dataset

- This project draws on a convenient dataset, Rameau, consisting of 140 Bach chorales with expert harmonic annotations.
- Harmonic labels in the Rameau dataset are aligned with the music as *salami-slice*:
  - ➢ Formed whenever a new note onset occurs in *any* musical voice (Fig.2).
- To make the tonal relationships between pitch-classes consistent across the dataset, we also transposed all the chorales into the same key.
- Non-chord tones can be identified and labeled from each chord label associated with the slice.

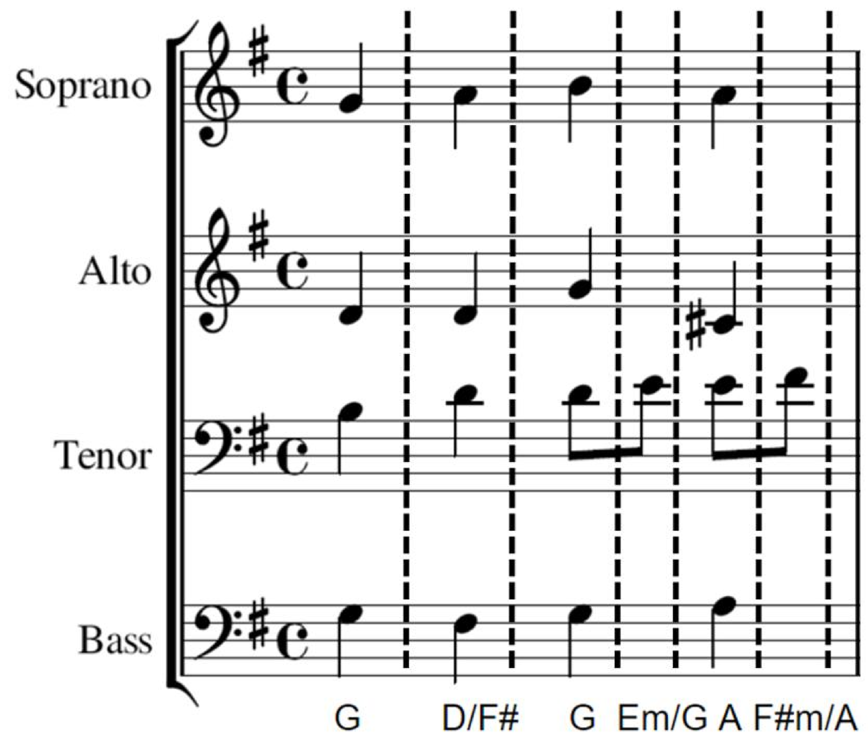


Fig. 2 Illustration of an excerpt from Rameau dataset and *salami slices* (Dotted lines indicate boundaries between slices).

Method

Table 1: The DNN settings. The number in parenthesis indicates the dimension of the features.

Network structure	2 hidden layers, 200 nodes each
Optimizer	ADAM
Loss function	Binary cross-entropy
Data division	8:1:1 (training : validation : test)
Input features	Pitch-class (12), on/off beat (2)
Output	Specify which present pitch-class (in an order of C, D, …, B) is non-chord tone
Window size	Three slices (42)

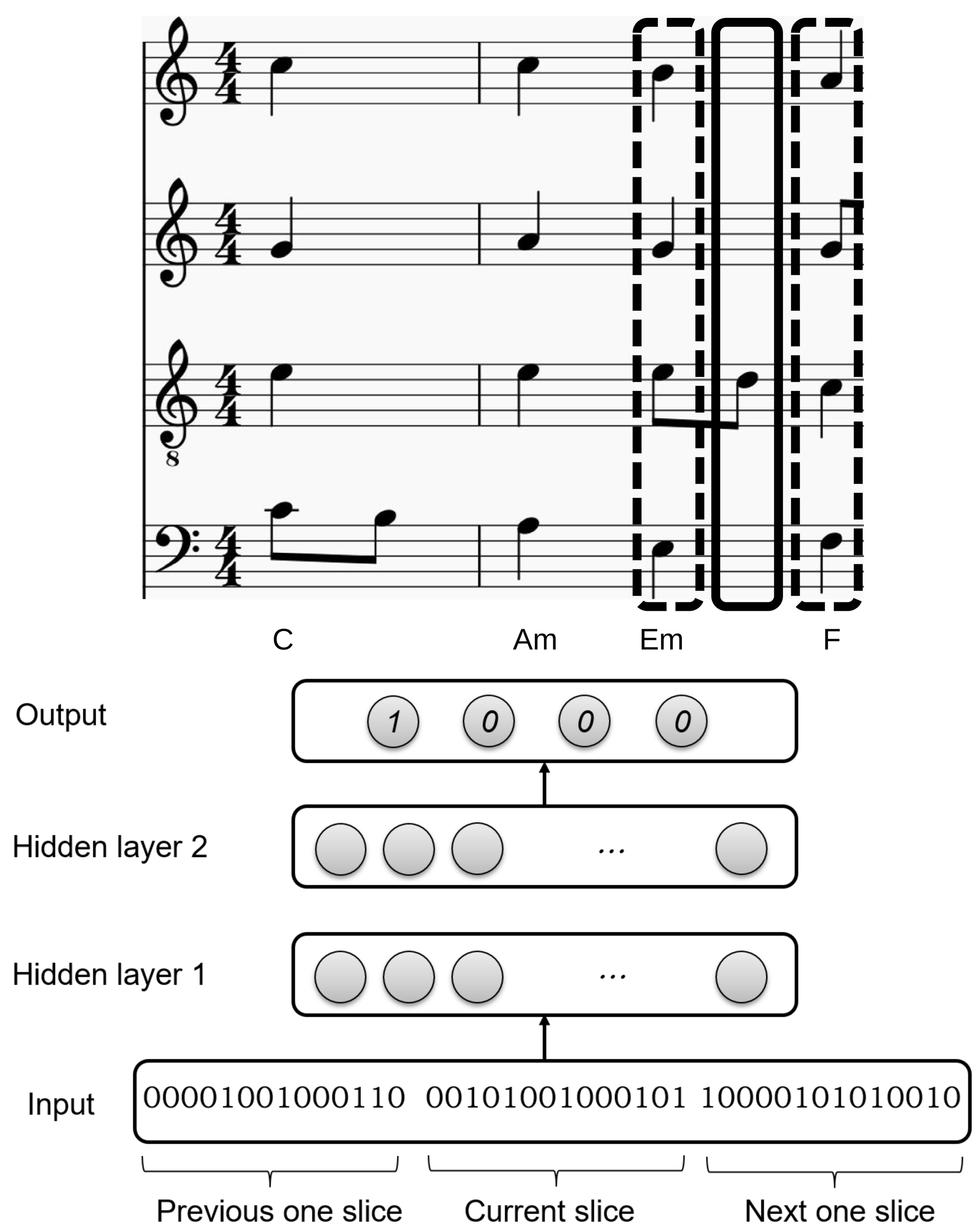


Fig. 3: The training process of DNN (below). Samples are outlined within the black rectangles (above).

Evaluation

- There is a significant imbalance between the numbers of chord tones and non-chord tones (92% and 8%).
  - ➢ If the model guesses every note as chord tone, it still can achieve an accuracy above 90%!
- To better measure the model's performance on identifying non-chord tones, **precision, recall** and **F1-measure** are reported.

Table 2: The experimental results.

Evaluation metric	Precision, recall, F1-measure
Evaluation method	10-fold cross validation
Precision	**86.02±3.35%**
Recall	**63.14±10.81%**
F1-measure	**72.19±7.68%**

Fig. 4: The first 18 measures of BWV 389 "Nun lob, mein Seel, den Herren". The first line of text underneath the score is the original chord labels. The second line is the non-chord-tone ground truth (note names). The third line is the model's predicted non-chord tones.

- As we can see, the model is correct for the first six measures, with some errors in the rest of the chorale. Experienced music analysts will see that many of the "errors" in fact represent plausible analytical choices.

Conclusion

- An innovative and promising approach to tackling the problem of non-chord tone identification, as well as harmonic analysis.
- Given a very limited amount of data, where the proportion of non-chord tones is less than 10%, the model still achieved an F1-measure above 72%.
- Including on/off metric information and a contextual window of size three slices achieves the optimal performance.

Future Work

- We plan to complete the whole Bach chorale dataset, with 371 chorales fully annotated with harmonic labels.
  - ➢ Providing more training data for our model.
  - ➢ The dataset can be used as a digital library for pedagogical, music theoretic, and musicological usage.

McGill | Schulich School of Music / École de musique Schulich | CIRMMT Centre for Interdisciplinary Research in Music Media and Technology | Social Sciences and Humanities Research Council of Canada / Conseil de recherches en sciences humaines du Canada | Canada | Fonds de recherche sur la société et la culture Québec | DDMAL Distributed Digital Music Archives & Libraries Lab

Acknowledgements

We would like to thank the Social Sciences and Humanities Research Council of Canada (SSHRC) and the Fonds de recherche du Québec - Société et culture (FRQSC) for their generous funding. We would also like to acknowledge the formidable contributions of our many collaborators on SIMSSA projects.

Presented in Proceedings of the 18th International Society for Music Information Retrieval Conference, Suzhou, China. October 27th, 2017