NON-CHORD TONE IDENTIFICATION USING DEEP NEURAL NETWORKS

Yaolong Ju **Nathaniel Condit-Schultz** **Claire Arthur** **Ichiro Fujinaga**
Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT)
Schulich School of Music, McGill University
Montréal, Canada
{yaolong.ju, nathaniel.condit-schultz, claire.arthur, ichiro.fujinaga}@mcgill.ca

ABSTRACT

This demo addresses the problem of harmonic analysis by proposing a non-chord tone identification model using deep neural networks (DNNs). 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 contextual window around each input sonority 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.

1. INTRODUCTION

Non-chord tones are elaborative notes, created by idiomatic step-wise melodic contours, which do not belong to the local structural harmony. Identifying non-chord tones is essential to many music analytic tasks, including polyphonic music retrieval [4], harmonization [1], and harmonic analysis [3]. Although the theoretical difficulty and importance of non-chord tone identification are addressed, few scholars have proposed complete, dedicated non-chord tone identification models.

In this demo, we construct a non-chord tone identification model based on deep neural networks (DNNs), trained on chorale music by J.S. Bach. Machine learning of harmonic analysis is difficult due to the large number of chord classes, which require large amounts of training data to learn. In contrast, the relatively simple task of non-chord tone identification requires much less training data. Once non-chord tones are identified, harmonic analysis becomes a relatively simple task, which can be accomplished by a rule-based algorithm. Deep learning has achieved substantial success in numerous complex tasks, and sometimes even surpassing human performance [2]. DNNs are nonetheless relatively simple, suitable for our preliminary research.

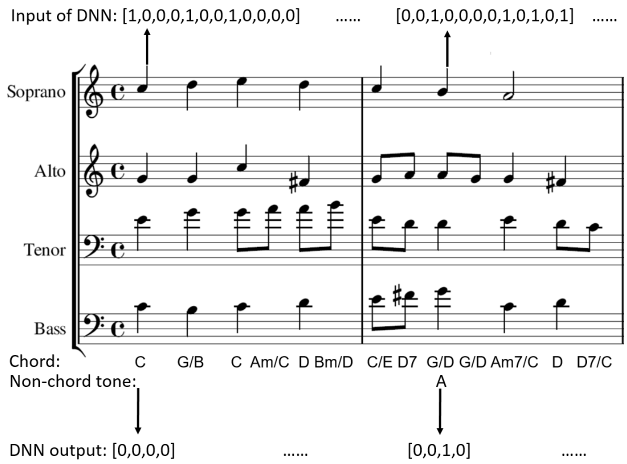

Figure 1. Non-chord tones in the first two measures of BWV 30/6 (transposed), and the corresponding DNN inputs and outputs.

2. DATASET

This project draws on a convenient dataset, Rameau [3], consisting of 140 Bach chorales with expert harmonic annotations. [1] Harmonic labels in the Rameau dataset are aligned with the music as *salami-slices*: a "salami-slice" is formed whenever a new note onset occurs in *any* musical voice. To make the tonal relationships between picth-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.

3. METHOD

As input to the DNN, each slice is represented by a vector of twelve ones or zeros, representing which pitch classes (C, C#/Db, D, D#/Eb, etc.) are present (1) or absent (0) in the slice. Also, metric information about each slice was added to each input vector, specifically whether the current

[1] Available at https://github.com/kroger/rameau/tree/master/rameau-deps/genos-corpus.

Network structure	2 hidden layers, 200 nodes each
Optimizer	ADAM (Adaptive Moment Estimation)
Loss function	Binary cross-entropy
Data division	8:1:1 (training:validation:test)
Evaluation metric	Precision, recall, F1-measure
Evaluation method	10-fold cross validation

Table 1. The DNN settings.

Input features	PC + B + WS1 (D:42)
Precision	86.02±3.35%
Recall	63.14±10.81%
F1-measure	72.19±7.68%

Table 2. Model performances with a combination of input features. (PC: pitch-class; B: on/off beat feature; WS: window size; D: dimension of the feature vector).

slice is on beat (1) or off (0). For the DNN's output we use a similar vector of length four, indicating which, if any, of the four voices contains a non-chord tone. These input and output vectors are illustrated in Fig. 1: The third slice in the second measure contains the pitch-classes D, G, A, and B, represented by the input vector [0,0,1,0,0,0,0,1,0,1,0,1]. The corresponding output vector for this slice [0,0,1,0] indicates that the pitch (A), which is the third "1" from the left in the input vector, is a non-chord tone. The experimental settings were determined empirically (shown in Table 1).

4. EVALUATION

Because of the significant imbalance between the number of chord tones and non-chord tones (92% and 8%), we report the metrics of precision, recall, and F1-measure. The results are shown in Table 2. Note that one slice adjacent (before and after) to the current one were added to the input vector, creating a windowed input that allows the model to consider context. As we can see, the model achieved F1-measure of 72.19%.

Fig. 2 illustrates the output of the model on a Bach chorale: Chord labels are placed immediately below the staves, while the next two lines of text indicate non-chord tones in the ground-truth data and the model's output respectively. 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.

5. CONCLUSION

In this demo, a non-chord tone identification model for Bach chorales using feedforward DNN is proposed. This model is trained and tested using the Rameau dataset. With an F1-measure of 72.19%, we hope that better performances will be achieved with more higher-quality data. Thus, we intend to complete the whole Bach chorale dataset with 371 chorales fully annotated with harmonic/

Figure 2. The first nine 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.

contrapuntal labels. [2] Not only will this dataset help to train and test our model further, it will be useful for many other music analytical tasks.

6. REFERENCES

[1] Ching-Hua Chuan and Elaine Chew. Generating and evaluating musical harmonizations that emulate style. *Computer Music Journal*, 35(4):64–82, 2011.

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: surpassing human-level performance on ImageNet classification. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1026–1034, 2015.

[3] Pedro Kröger, Alexandre Passos, Marcos Sampaio, and Givaldo De Cidra. Rameau: a system for automatic harmonic analysis. In *Proceedings of International Computer Music Conference*, Belfast, N. Ireland, 2008.

[4] Jeremy Pickens. *Harmonic modeling for polyphonic music retrieval*. PhD thesis, University of Massachusetts at Amherst, 2004.

[2] We would like to thank SSHRC for partially funding this research.