

TOWARD HARMONIC CONVOLUTIONS IN NEURAL NETWORKS

First Author

Affiliation1

author1@ismir.edu

Second Author

Retain these fake authors in

submission to preserve the formatting

Third Author

Affiliation3

author3@ismir.edu

ABSTRACT

Convolutional neural networks (CNNs), originally pioneered in computer vision, have shown great performance in content-based MIR across a variety of challenges. These models typically operate on time-frequency representations of audio, however, which are statistically distinct from natural images. Whereas images exhibit spatial correlations between pixels, sound is harmonic, and thus fundamental assumptions about the convolutional architecture are potentially invalid. Therefore, in an effort to better adapt CNNs to audio, we craft two input representations that aim to exploit this known behavior directly, and determine what effect this has on the learned model. We study these approaches in a large, solo instrument classification task, taking the union of three datasets. Experimental results show that the use of harmonic convolutions leads to better generalization, faster computation time, and more compact models per unit of performance. Finally, we provide an open source framework to reproduce both our results and the data used herein.

1. INTRODUCTION

This template includes all the information about formatting manuscripts for the ISMIR 2016 Conference. Please follow these guidelines to give the final proceedings a uniform look. Most of the required formatting is achieved automatically by using the supplied style file (L^AT_EX) or template (Word). If you have any questions, please contact the Program Committee (ismir2016-papers@ismir.net). This template can be downloaded from the ISMIR 2016 web site (<http://ismir2016.ismir.net>).

2. PAPER LENGTH & FILE SIZE

Instead of the strict limit of six pages (last used in ISMIR 2014), we adopt a “(6+1)-page policy” from ISMIR 2015 for ISMIR 2016. This means, the paper may have a maximum of 6 pages for technical content including figures and possible references with one additional optional 7th page containing only references. Note that this is a strict

requirement. **The seventh page (if used at all) must not contain any other material except for references.**

Paper should be submitted as PDFs and the **file size is limited to 2MB**. Please compress images and figures as necessary before submitting.

3. PAGE SIZE

The proceedings will be printed on portrait A4-size paper (21.0cm x 29.7cm). All material on each page should fit within a rectangle of 17.2cm x 25.2cm, centered on the page, beginning 2.0cm from the top of the page and ending with 2.5cm from the bottom. The left and right margins should be 1.9cm. The text should be in two 8.2cm columns with a 0.8cm gutter. All text must be in a two-column format. Text must be fully justified.

4. TYPESET TEXT

4.1 Normal or Body Text

Please use a 10pt (point) Times font. Sans-serif or non-proportional fonts can be used only for special purposes, such as distinguishing source code text.

The first paragraph in each section should not be indented, but all other paragraphs should be.

4.2 Title and Authors

The title is 14pt Times, bold, caps, upper case, centered. Authors’ names are omitted when submitting for double-blind reviewing. The following is for making a camera-ready version. Authors’ names are centered. The lead author’s name is to be listed first (left-most), and the co-authors’ names after. If the addresses for all authors are the same, include the address only once, centered. If the authors have different addresses, put the addresses, evenly spaced, under each authors’ name.

4.3 First Page Copyright Notice

Please include the copyright notice exactly as it appears here in the lower left-hand corner of the page. It is set in 8pt Times.

4.4 Page Numbering, Headers and Footers

Do not include headers, footers or page numbers in your submission. These will be added when the publications are assembled.



© First Author, Second Author, Third Author. Licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0). **Attribution:** First Author, Second Author, Third Author. “Toward Harmonic Convolutions in Neural Networks”, 17th International Society for Music Information Retrieval Conference, 2016.

String value	Numeric value
Hello ISMIR	2016

Table 1. Table captions should be placed below the table.

5. FIRST LEVEL HEADINGS

First level headings are in Times 10pt bold, centered with 1 line of space above the section head, and 1/2 space below it. For a section header immediately followed by a subsection header, the space should be merged.

5.1 Second Level Headings

Second level headings are in Times 10pt bold, flush left, with 1 line of space above the section head, and 1/2 space below it. The first letter of each significant word is capitalized.

5.1.1 Third and Further Level Headings

Third level headings are in Times 10pt italic, flush left, with 1/2 line of space above the section head, and 1/2 space below it. The first letter of each significant word is capitalized.

Using more than three levels of headings is highly discouraged.

6. FOOTNOTES AND FIGURES

6.1 Footnotes

Indicate footnotes with a number in the text.¹ Use 8pt type for footnotes. Place the footnotes at the bottom of the page on which they appear. Precede the footnote with a 0.5pt horizontal rule.

6.2 Figures, Tables and Captions

All artwork must be centered, neat, clean, and legible. All lines should be very dark for purposes of reproduction and art work should not be hand-drawn. The proceedings are not in color, and therefore all figures must make sense in black-and-white form. Figure and table numbers and captions always appear below the figure. Leave 1 line space between the figure or table and the caption. Each figure or table is numbered consecutively. Captions should be Times 10pt. Place tables/figures in text as close to the reference as possible. References to tables and figures should be capitalized, for example: see Figure ?? and Table ?. Figures and tables may extend across both columns to a maximum width of 17.2cm.

7. EQUATIONS

Equations should be placed on separate lines and numbered. The number should be on the right side, in parentheses, as in Eqn (??).

$$E = mc^2 \quad (1)$$

¹ This is a footnote.

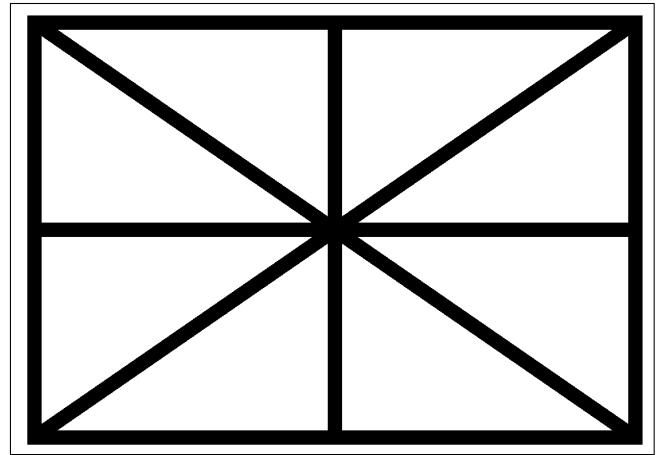


Figure 1. Figure captions should be placed below the figure.

8. CITATIONS

All bibliographical references should be listed at the end, inside a section named “REFERENCES,” numbered and in alphabetical order. All references listed should be cited in the text. When referring to a document, type the number in square brackets [?], or for a range [?, ?, ?].