



12th IAPR International Workshop on Graphics REcognition
Kyoto, 9-10 Nov. 2017



Session 3: Optical Music Recognition

Chairs:

Nina Hirata (University of São Paulo)

Josep Lladós (CVC, Universitat Autònoma de Barcelona)

Session outline (each paper: 10 min presentation)

- **On the Potential of Fully Convolutional Neural Networks for Musical Symbol Detection.** Matthias Dorfer, Jan Hajič, Gerhard Widmer
- **Towards a Universal Music Symbol Classifier.** Alexander Pacha, Horst Eidenberger
- **Bootstrapping Samples of Accidentals in Dense Piano Scores for CNN-Based Detection.** Kwon-Young Choi, Bertrand Coüasnon, Richard Zanibbi, Yann Ricquebourg
- **Optical Music Recognition by Recurrent Neural Networks.** Arnau Baró, Pau Riba, Jorge Calvo Zaragoza, Alicia Fornés
- **Pen-based Music Document Transcription.** Javier Sober-Mira, Jorge Calvo Zaragoza, David Rizo, Jose Manuel Inesta.

Introduction to the topic

- Optical music recognition (OMR) can be seen as the application of optical character recognition to interpret sheet music or printed scores into editable or playable form.
- Traditionally, given a musical score image, it follows the traditional pipeline of layout analysis (staff removal, and symbol segmentation), and classification.
- The use of syntactic models is highly important.
- One of the major challenges is the recognition of handwritten scores, both off line (e.g. historical musical documents) or on line (creation and edition).



12th IAPR International Workshop on Graphics REcognition
Kyoto, 9-10 Nov. 2017



Paper 1:

On the Potential of Fully Convolutional Neural Networks for Musical Symbol Detection. Matthias Dorfer, Jan Hajič, Gerhard Widmer

- Notehead detection on handwritten scores
- CNN that produces a pixel level probability map, followed by searching of local peaks in the map (considered centroids of the noteheads)
- high F-score both for writer dependent and independent tests
- adaptable to the detection of other types of symbols



12th IAPR International Workshop on Graphics REcognition
Kyoto, 9-10 Nov. 2017



Paper 2:

Towards a Universal Music Symbol Classifier. Alexander Pacha, Horst Eidenberger

- building of a large musical symbol dataset, by unifying existing datasets
- unification of class names and treatment of ambiguous symbols
- preliminary classification tests give error rates below 2%

Paper 3:

Bootstrapping Samples of Accidentals in Dense Piano Scores for CNN-Based Detection. Kwon-Young Choi, Bertrand Coüasnon, Richard Zanibbi, Yann Ricquebourg

- A CNN based detector (visual attention) to localize and classify three accidental symbols associated with a note head, or the note head if there is no accidental.
- To cope with “few data samples”, a data augmentation bootstrapping method is used.
- Complex and damaged piano scores.



12th IAPR International Workshop on Graphics REcognition
Kyoto, 9-10 Nov. 2017



Paper 4:

Optical Music Recognition by Recurrent Neural Networks. Arnau Baró, Pau Riba, Jorge Calvo Zaragoza, Alicia Fornés

- Recognition of musical scores as a sequence using BLSTM Recurrent Neural Networks.
- Context can be introduced.
- Use of a training synthetic dataset of more than 40.000 images labeled at primitive level.



12th IAPR International Workshop on Graphics REcognition
Kyoto, 9-10 Nov. 2017



Paper 5:

Pen-based Music Document Transcription. Javier Sober-Mira, Jorge Calvo Zaragoza, David Rizo, Jose Manuel Inesta.

- Human-machine interaction task for music notation creation.
- Combination of two modalities for the recognition: on-line and off-line data.
- Convolutional Neural Network approach.

Discussion

Open questions:

- Overall summary: The papers address problems from symbol detection and classification, to transcription of music scores.
- Deep learning models are the current choice (which model works best?)
- Many works are proposing the introduction of deep neural network approaches, but still for a particular step in the pipeline. Can we think in end-to-end systems?
- What about the “personal” component of music interpretation? Each composer has his/her own style. Can we use this type of contextual knowledge? How?
- Shall we approach other communities (e.g. NLP)?
- Availability and imbalance of training data is perceived as an important concern. What about databases? What is required? How to obtain it?
- Which is the “big challenge”?