International Journal on Document Analysis and Recognition (IJDAR) (2023) 26:347–362
https://doi.org/10.1007/s10032-023-00432-z

SPECIAL ISSUE PAPER

End-to-end optical music recognition for pianoform sheet music

Antonio Ríos-Vila[1] · **David Rizo**[1,2] · **José M. Iñesta**[1] · **Jorge Calvo-Zaragoza**[1]

Received: 14 November 2022 / Revised: 2 April 2023 / Accepted: 8 April 2023 / Published online: 12 May 2023

Abstract
End-to-end solutions have brought about significant advances in the field of Optical Music Recognition. These approaches directly provide the symbolic representation of a given image of a musical score. Despite this, several documents, such as pianoform musical scores, cannot yet benefit from these solutions since their structural complexity does not allow their effective transcription. This paper presents a neural method whose objective is to transcribe these musical scores in an end-to-end fashion. We also introduce the GRANDSTAFF dataset, which contains 53,882 single-system piano scores in common western modern notation. The sources are encoded in both a standard digital music representation and its adaptation for current transcription technologies. The method proposed in this paper is trained and evaluated using this dataset. The results show that the approach presented is, for the first time, able to effectively transcribe pianoform notation in an end-to-end manner.

Keywords Optical music recognition · Polyphonic music scores · GrandStaff · Neural networks

1 Introduction

Transcribing the content of musical documents to structured formats brings benefits to digital humanities and musicology, as it enables the application of algorithms that rely on symbolic music data and makes musical score libraries more browsable. Given the price of manual transcription, it is unaffordable to transcribe large historical archives manually. In this scenario, the reading of music notation invites automation, much in the same way as modern technology in the fields of Optical Character Recognition (OCR) or Handwritten Text Recognition (HTR) has enabled the automatic processing of written texts. The field of Optical Music Recognition (OMR) covers the automation of this computational *reading* in the context of music [1].

Holistic approaches, also referred to as *end-to-end* approaches, have begun to dominate the fields of sequential labeling, with notable examples such as HTR or Automatic Speech Recognition. In OMR, these approaches have proved successful in those contexts in which music notation retrieval can be easily expressed as a sequence. This applies to monophonic scores, or legacy music-notation languages in which different voices were written individually. However, the scores of many compositions are written using *grand staves*, i.e., a combination of two staves put together, such as those used for the piano (see Fig. 1). In the related literature, this kind of scores is also referred to as *pianoform* [1–3]. However, no end-to-end system that has attempted to recognize the content of this type of scores is known to date.

This work proposes the first end-to-end recognition approach for pianoform scores. This constitutes a first step in the application of holistic models to the full spectrum of OMR applications. We consider a neural approach inspired by state-of-the-art full-paragraph HTR research, with which the OMR problem shares some of its challenges. This approach provides a serialization of the scores based on textual encodings of music notation. Likewise, since it is the first attempt to address this problem, this work also introduces the GRANDSTAFF dataset, a large corpus of isolated grand staves

✉ Antonio Ríos-Vila
arios@dlsi.ua.es
URL: https://praig.ua.es/author/rios-vila-antonio/

David Rizo
drizo@dlsi.ua.es
URL: https://praig.ua.es/author/drizo/

José M. Iñesta
inesta@dlsi.ua.es
URL: https://praig.ua.es/author/inesta/

Jorge Calvo-Zaragoza
jcalvo@dlsi.ua.es
URL: https://praig.ua.es/author/jcalvo/

[1] U.I for Computing Research, University of Alicante, Alicante, Spain

[2] Instituto Superior de Enseñanzas Artísticas de la Comunidad Valenciana (ISEA. CV), Alicante, Spain