Combining image processing techniques, OCR, and OMR for the digitization of musical books

DAS 2022
May 22-25, La Rochelle

Gonzalo Santamaría, César Domínguez, Jónathan Heras, Eloy Mata, Vico Pascual
Department of Mathematics and Computer Science, University of La Rioja, Spain
{gonzalo.santamaria, cesar.dominguez, jonathan.heras, eloy.mata, vico.pascual}@unirioja.es

Abstract

Digitizing historical music books can be challenging since staves are usually mixed with typewritten text explaining some characteristics of them. In this work, we propose a new methodology to undertake such a digitization task. After scanning the pages of the book, the different blocks of text and staves can be detected and organized into music pieces using image processing techniques. Then, optical character recognition (OCR) and optical music recognition (OMR) methods can be applied to text and stave blocks, respectively, and the information conveniently stored using the *MusicXML* format. In addition, we explain how this methodology was successfully applied in the digitization of a book entitled "The Music in the Santo Domingo's Cathedral". In particular, we provide a new annotated database of musical symbols from the staves included in that book. This database was used to develop two new OMR deep learning models for the detection and classification of music scores. The detection model obtained a F1-score of 90 % on symbol detection; and the classification model a note pitch accuracy of 98.4 %. The method allows us to conduct text searches, obtain clean PDF files of music pieces, or reproduce the sound represented by the pieces.

Methodology

The methodology proposed in this work consists of the following steps: (1) Digital scanning of each page of the book. (2) Staves, text blocks, and musical pieces detection. (3) OCR or OMR techniques are used to extract the information from these blocks. (4) Finally, the obtained information is properly recomposed and stored in a database. With this information it is possible to create, for instance, a web page to access the digital copy of the book through different types of searches.
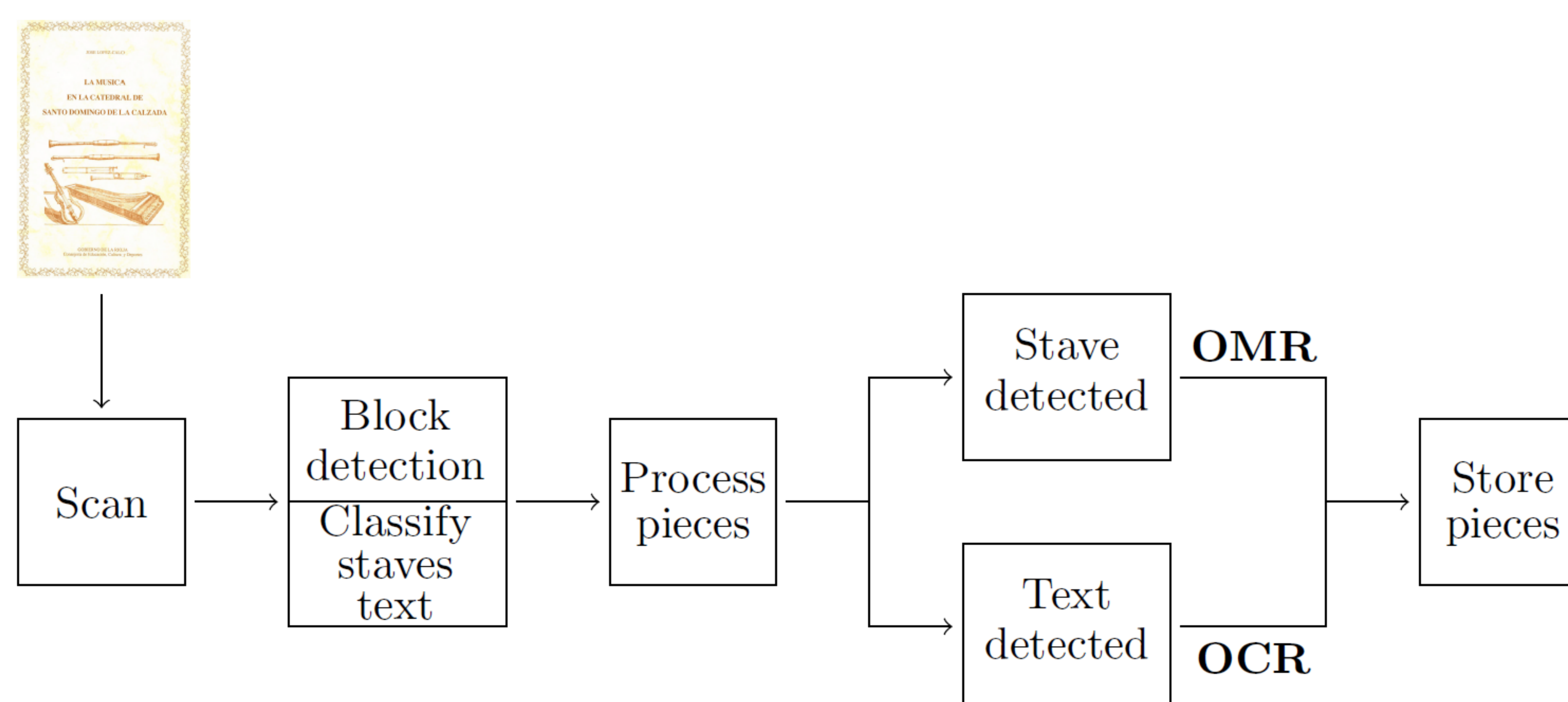
Implementation

The previous methodology has been applied to the digitization of "The Music in the Santo Domingo's Cathedral" book [1].

(1) Scanned image   (2) Staves and text blocks detection   (3) **OCR** and **OMR** combination   (4) Input image reconstruction

Results

Results achieved detecting musical symbols by OMR models are included in the following table. In bold face the best results. In the case of pitch notes classification, ResNet-18 model obtained an accuracy of 98.4 % (the training of this model took 56 seconds). A combination of YOLO and ResNet-18 models were choosen to process all the staves of the book.

	Precision	Recall	F1-score	mAP	COCO	Training time
EfficientDet	0.28	0.14	0.19	13.86 %	0.16	0.37 h
Faster R-CNN	0.86	0.77	0.82	**76.73 %**	**0.59**	**0.21 h**
RetinaNet	0.73	0.15	0.25	14.73 %	0.19	5.49 h
YOLO v4	**0.89**	**0.90**	**0.90**	68.25 %	0.48	72 h

Conclusions

- We have proposed a new methodology for the digitization of music books which merge text with staves.
- We have produced a new annotated database of musical symbols coming from the historical music produced for the Cathedral of Santo Domingo.
- We have developed a detection and a classification deep learning OMR models for structured representation of historical music scores and conducted a thorough analysis of the results of the models.
- We have publicly released all the code, annotated images, and models developed in this work on the project website `https://github.com/joheras/MusicaCatedralStoDomingoIER`.

Bibliography

1. López-Caro, José. La Música en la Catedral de Santo Domingo de la Calzada, Vol. I: Catálogo del Archivo de Música, Gobierno de La Rioja, 1988.