First Approach for a Music Sheet Automatic Player - Project Proposal

Houdaifa Ahmidan
Ibrahim Kaddouri
Nicolás Acuña Reyes
Camilo Carvajal Reyes

1. Motivation and Problem Definition

Music theory can go from simple concepts to deep and complex systems, then, due to its variety and complexity, one of the main challenges for musicians is to imagine sounds when reading music sheets. As not all musicians possess the ability to sight-read music scores, the goal of this work is to build a first approach of an algorithm capable of transforming images of music sheets into actual sounds.

This could help people who can't use an instrument to create music and play it using the program. It could also help those who are learning new instruments by showing them how the music they are trying to play sounds, so that they can try to imitate it. Furthermore, with more time and effort, the algorithm could be converted into one capable of drawing with music or even creating music with paintings. This would be a great achievement to help deaf people "listen" to music by just looking at images or to help blind people to see paintings by listening to the music that it generates.

2. State of the Art

Optical Music Recognition (OMR) is a challenge that has called the attention of the scientific and musical community. Some of its difficulties and up-to-date progress can be found on the work of Calvo-Zaragosa et al. [1]. Some applications such as SmartScore [2] or PhotoScore [3] provide a music score scanning service but at a very high price and usually relying on complicated applications and interfaces that are usually used by experts, which are not really in the scope of this project. An interesting work is the one of Rios-Vila et al., which proposes *ReadSco*, an open-source optical music recognition tool that is currently under early development [4]. The OMR part of this work, which focuses mainly in the system's architecture, is based on the approach by Alfaro-Contreras et al. [5]. It uses *Convolutional Recurrent Neural Network* so that the layers interpret image features in terms of sequences of musical symbols and it applies *Back-Propagation* for training. Similarly, Huang et al. also make use of deep learning for music object recognition, although with previously computerised images [6].

3. Methodology

Our purpose is to train an algorithm that would be able to recognise musical patterns. To do so, we will need to start by using existing algorithms of pattern recognition and adapting them to our dataset. The use of the dataset *Muscima* seems appropriated for this task, especially because it is endowed with a tutorial and a python library, making the dataset easier for us to handle [7].

As done in the works of Huang et al. and of Alfaro-Contreras et al., the problem will be dealt with using a deep learning approach [6][5]. The idea of our project is to use pattern recognition based algorithms (maybe some additional ones) and to compare the results with those of the articles.

For evaluating the project, the dataset is going to be split into a training set and a test set. Obviously, an acceptable performance on the testing set is expected. The determination of the threshold between acceptable and poor accomplishment will depend on the behaviour over the training set. Moreover, the work is going to be qualitatively tested with musicians in order to asses its real impact.

References

[1] Calvo-Zaragoza, J., Hajic Jr, J., Pacha, A. (2019). Understanding optical music recognition. arXiv preprint arXiv:1908.03608.
https://arxiv.org/abs/1908.03608

[2] Musitek. Smartscore
http://www.musitek.com/smartscore-pro.html

[3] Neuratron. Photoscore 2018
http://www.neuratron.com/photoscore.html

[4] Rıos-Vila, A., Calvo-Zaragoza, J., Rizo, D., Inesta, J. M. ReadSco: An Open-Source Web-Based Optical Music Recognition Tool.
https://sao.dlsi.ua.es/repositori/grfia/pubs/439/ReadSco_Worms2019.pdf

[5] Alfaro-Contreras, M., Calvo-Zaragoza, J., Iñesta, J. M. (2019, July). Approaching end-to-end optical music recognition for homophonic scores. In Iberian Conference on Pattern Recognition and Image Analysis (pp. 147-158). Springer, Cham.
https://link.springer.com/chapter/10.1007/978-3-030-31321-0_13

[6] Huang, Z., Jia, X., Guo, Y. (2019). State-of-the-Art Model for Music Object Recognition with Deep Learning. Applied Sciences, 9(13), 2645.
https://www.mdpi.com/2076-3417/9/13/2645

[7] Hajič, J., Pecina, P. (2017, November). The MUSCIMA++ dataset for handwritten optical music recognition. In 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR) (Vol. 1, pp. 39-46). IEEE.
https://ufal.mff.cuni.cz/muscima