International Workshop on Machine Learning and Music
European Conference on Machine Learning
2024, Vilnius, Lithuania

University of Alicante, Spain
Pattern Recognition and Artificial Intelligence Group

Can patch selection heuristics enhance layout analysis of music scores?

Francisco J. Castellanos, Juan P. Martinez-Esteso, Alejandro Galán-Cuenca, Antonio Javier Gallego
fcastellanos@dlsi.ua.es - juan.martinez11@ua.es - a.galan@ua.es - jgallego@dlsi.ua.es

- **Preservation** and **accessibility** of music scores are possible by means of **Optical Music Recognition (OMR) systems**.
- This paper focuses on **Layout Analysis (LA)**, a process in OMR that **classifies each pixel** in different layers of information.
- The trend is to use neural networks at the expense of the **high cost of manual labeling**.
- **Objective**: reducing the amount of labeled data.
- **Proposal**: labeling a portion of an image, selected according to a **specific criterion**.

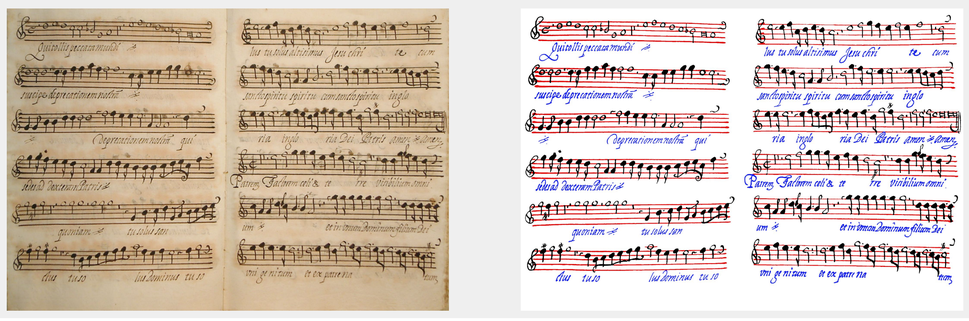


Fig. 1: Example of LA for an input image to detect staff lines, text, music symbols and background.

Methodology

We use an existing few-shot method for LA as our basis, called **Few-shot Selectional Auto-encoder (FSAE)**:[1]

1. Manual partial annotations.
2. Extracting random patch samples.
3. Training the model with the random samples.

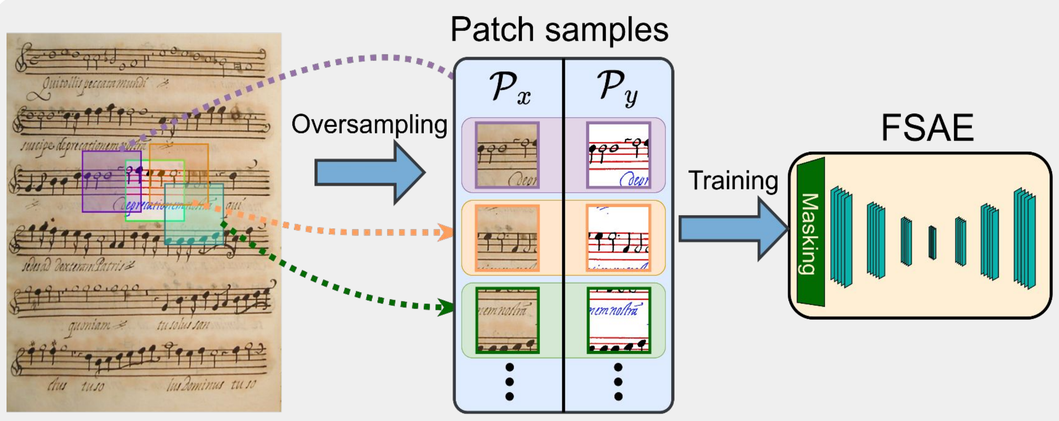


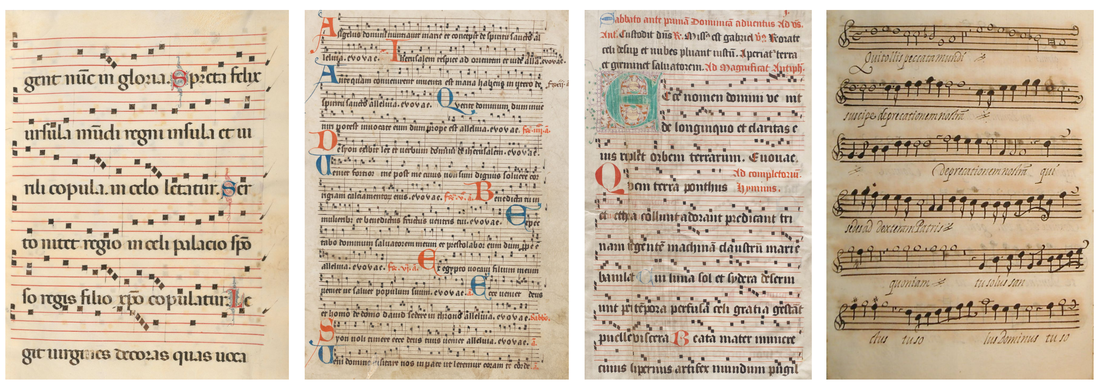
Fig. 2: Scheme of the few-shot method in which we based.

We study how the selection of an image portion (**step 1**) affects the FSAE's training with different selection policies:

1. Random selection.
2. Sequential selection.
3. Ink-rate selection.
4. Entropy-based selection.

Experiments

Corpora



a) ***MS73*** b) ***Einsiedeln*** c) ***Salzinnes*** d) ***Capitan***

Fig. 3: Examples of the corpora considered for experiments.

[1] *Castellanos, Gallego, Fujinaga: A few-shot neural approach for layout analysis of music score images. In ISMIR Conference (2023)*

Results and analysis

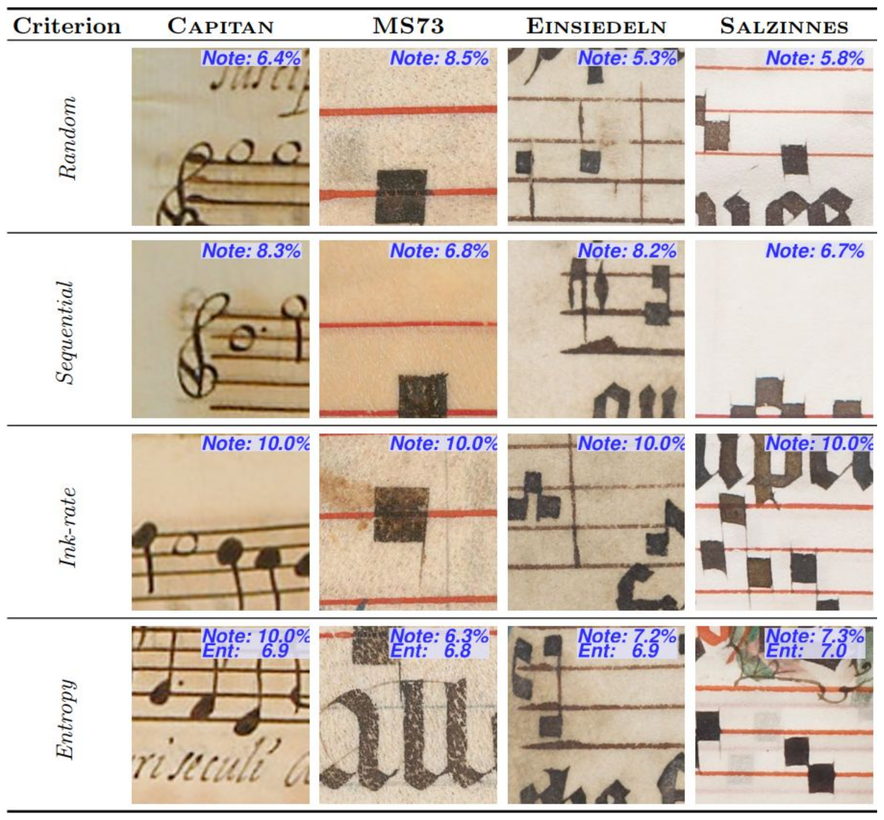


*Fig. 4: Examples of selected patch samples according to a specific criterion for the **Notes** layer.*

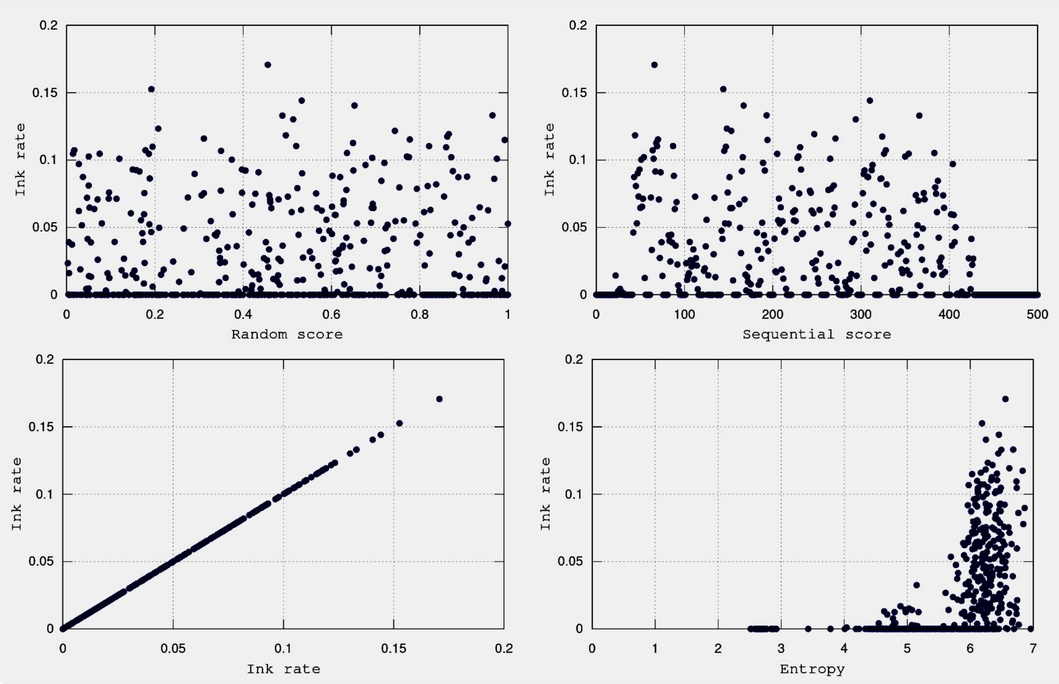


*Fig. 5: Analysis of the ink rate levels for the **Notes** layer according to the selected policy.*

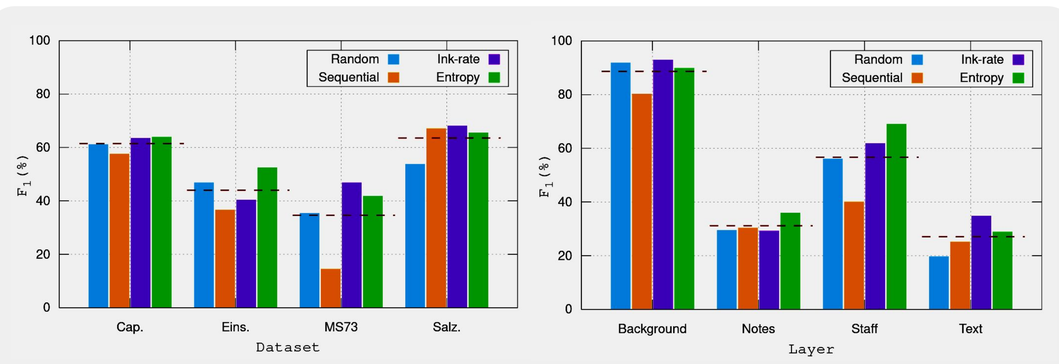


Fig. 6: F_1 results for the four selection policies. Dashed lines denote the average result.

Metric	Random	Sequential	Ink-rate	Entropy
F_1 (%)	49.2	43.9 $^{-5.3}$	54.6 $^{+5.4}$	**55.9** $^{+6.7}$

Tab. 1: Average F_1 (%) results for each patch selection method. Superscript values denote the percentage improvement with respect to the random method.

Conclusions

- Strategies based on entropy and ink rate are the most effective.
- Sequential and random do not consider the amount of information, resulting in lower performance.
- On average, entropy seems to be the best option with a 6.7% improvement in F_1 score, but not always.
- Other strategies should be studied.

Universitat d'Alacant
Universidad de Alicante
GOBIERNO DE ESPAÑA MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES AGENCIA ESTATAL DE INVESTIGACIÓN
Funded by the European Union
NextGenerationEU
Plan de Recuperación, Transformación y Resiliencia