

The DeepScoresV2 Dataset and Benchmark for Music Object Detection

Lukas Tuggener

Yvan Putra Satyawan

Alexander Pacha

Jürgen Schmidhuber

Thilo Stadelmann

25th International Conference on Pattern Recognition

January 2021



Università
della
Svizzera
italiana



datalab

www.zhaw.ch/datalab

What is DeepScores (and DeepScoresV2)

Large scale dataset for optical music recognition.

- Collection of over 300.000 sheets
- High-quality renderings using different musical fonts
- Annotated for object classification, detection, and semantic segmentation

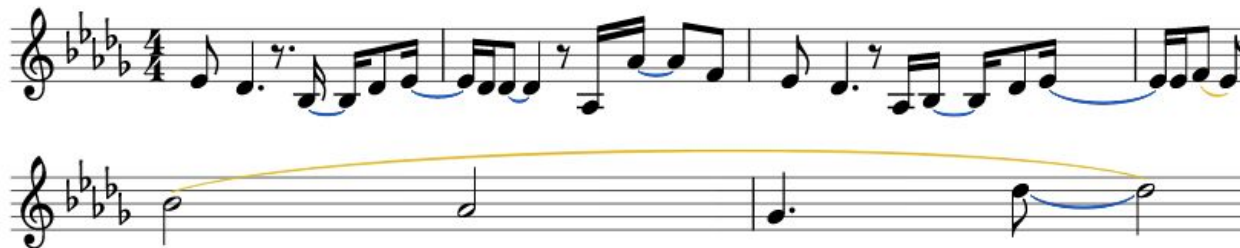


Bounding box annotation



Pixelwise annotations

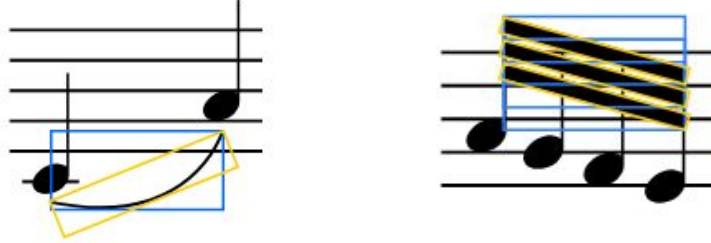
Novelties in DeepScoresV2 - additional classes



Wider variety of classes, totaling 135.

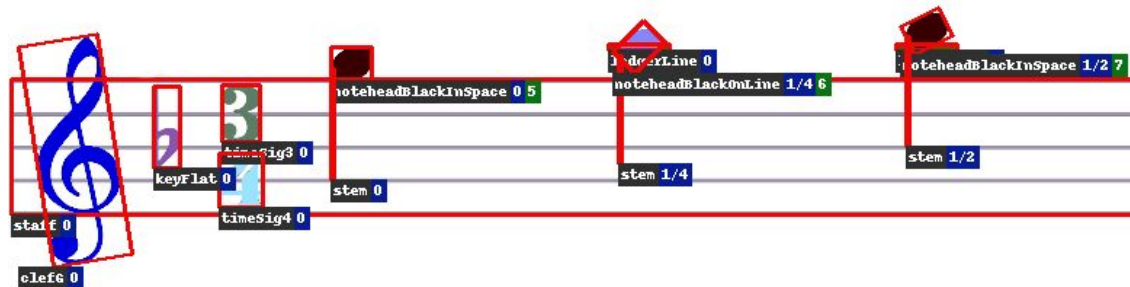
Introducing non-fixed shape symbols such as ties (blue) and slurs (yellow).

Novelties in DeepScoresV2 - oriented bounding boxes



- Aligned bounding boxes can be very imprecise and cause a lot of overlap
- Oriented decrease enclosed white pixels by 13%

Novelties in DeepScoresV2 - higher level information & instance segmentation



- per object onset
- relative notehead position

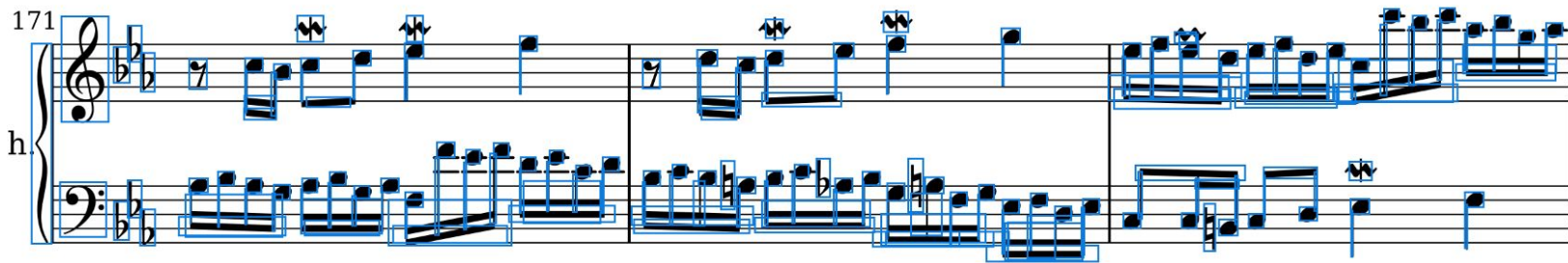


Benchmarks on DeepScoresV2

The Deep Watershed Detector

- Trained on full resolution images (trained on crops, tested using stitched crops)
- Ledgers and staves disabled due to overlaps
 - Low accuracy bounding boxes
 - + Finds also very slim objects such as staves

AP@0.5 of 0.503 but a mAP of only 0.203

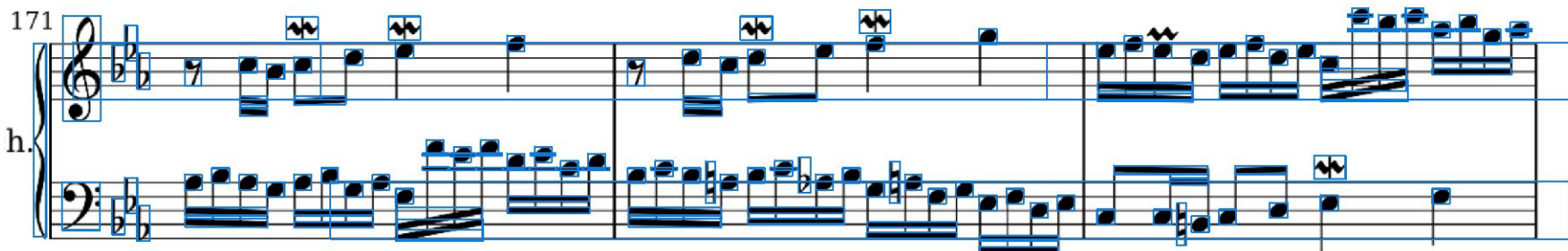


Benchmarks on DeepScoresV2

Faster R-CNN

- Trained on half resolution images (trained on crops, tested on full images)
 - Does not find the very slim objects such as stems
 - + Very accurate bounding boxes

AP@0.5 of 0.799 and a mAP of 0.700



Thanks for your attention!

Data and code available at:

<https://zenodo.org/record/4012193>



On me:

- Doctoral Student ZHAW / USI
- lukas.tuggener@zhaw.ch
- <https://tuggeluk.github.io/>

Happy to answer questions & requests.