

AI for Digitization of historic architectural plans

An AI-pipeline using OCR and NER models for calligraphic text

Student



Kevin Löffler

Introduction: The Swiss Archive for Landscaping Architecture, located at OST in Rapperswil, administers more than 100'000 historic plans. The collection contains estates of renowned Swiss and European architects who worked in the 20th century. Besides using the plans for teachings at OST, the archive is an important point of reference for today's architects. The archive is in the process of manually digitizing its documents which is a very time-consuming process. This SA is an exploration of different approaches and deep learning models to find an image processing pipeline that automates as much of the manual digitization process as possible.

Approach: This paper proposes a three-model architecture consisting of a layout model to find text on the plans, an optical character recognition model to extract the found words, and finally, a named entity recognition model to label the relevant words like the client, location, or date. K-means clustering is used to group the text blocks from the layout model into related blocks for OCR. Different deep-learning models are compared and evaluated. The most suitable models are then retrained on the NVIDIA DGX-2 system in a custom-built aptainer image with different training strategies to improve their accuracy.

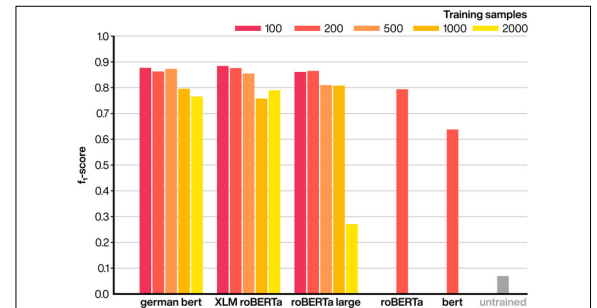
Result: The final image pipeline acts as a proof of concept for the given task. Testing the entire pipeline on various test images resulted in an F1 score of 48% with 35% precision and 77% recall. The chosen NER model "German BERT" scored an F1-score of 86% after re-training and the OCR pipeline extracted 54% of words perfectly, 18% were close, and 28% were bad matches. All models could be improved with further training, subject to training data being available. Other potential improvements are pre-processing techniques like denoising, thresholding,

and distance transforms, among others.

The insights from this SA can be applied to future projects to build an application usable by the archive, enabling it to catalog its documents and make them accessible to the world.

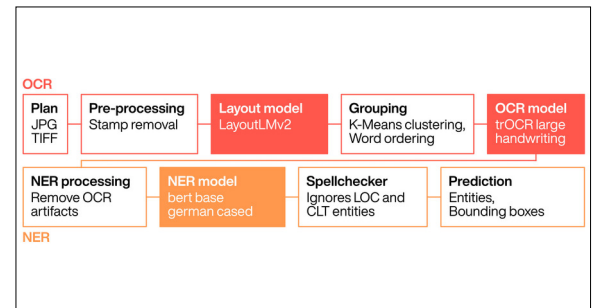
NER model comparison with different amounts of generated training data

Own presentation

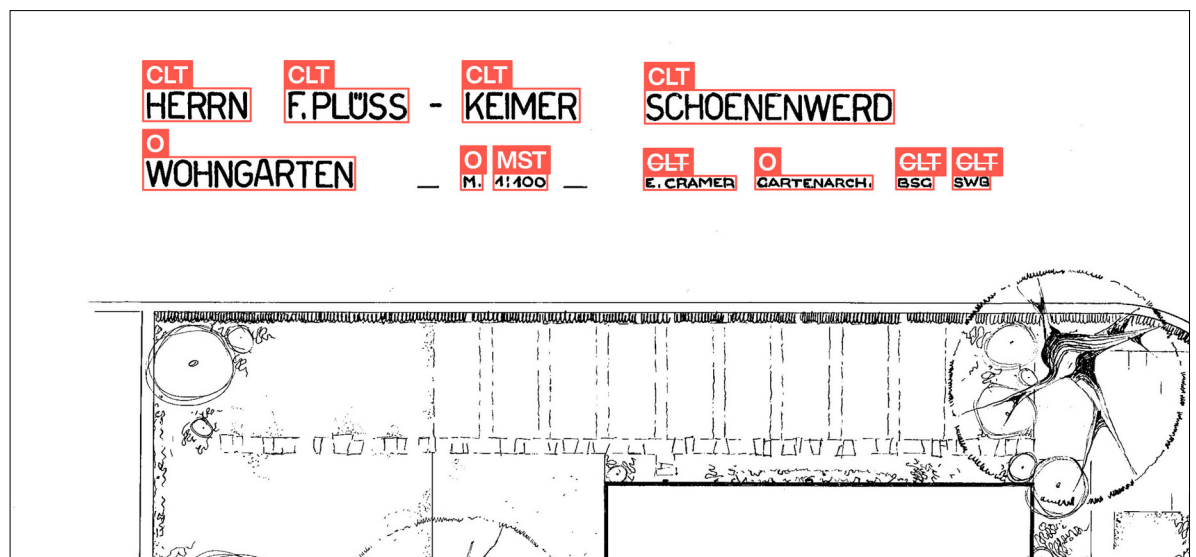


Pipeline architecture

Own presentation



Output of the pipeline with OCR bounding boxes and NER label annotations
Plan from ASLA



Advisor

Prof. Dr. Mitra Purandare

Subject Area

Software, Miscellaneous

Project Partner

OST Rapperswil:
Institute of Landscape Architecture,
Rapperswil-Jona, St. Gallen