

# Document Layout Analysis System

Author: BKZA team

Duc Nguyen ([nguyentanduc96@gmail.com](mailto:nguyentanduc96@gmail.com))

Cuong Ha ([nhatcuong94@gmail.com](mailto:nhatcuong94@gmail.com))

## Overview

Our method uses deep learning for segmentation and some heuristic algorithms for post-processing.

## Method

Our method has 3 phases

Phase 1: Segmentation with Deep learning

- Using U-net with Residual block
- Using 1 network for each class
- Data set: Layout Analysis Dataset

Phase 2: Post Processing

- Image processing algorithms to expand regions, remove noise regions and re-correct region type.
- OCR with tesseract.

Phase 3: Export to XML format with java

## Run

Run faster with GPU and Cuda driver

Read the README to know how to run

## References

[1] Ronneberger, Olaf, Philipp Fischer, and Thomas Brox. "U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015.

[2] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.