

What is TextBite?

A deep learning tool for extraction of articles, news, dictionary entries, book segments, etc. from historical documents.

TextBite enhances search capabilities in digitized documents used by librarians and scientists.

01 / Data

A custom dataset was created and labeled based upon data from Czech Digital Library.

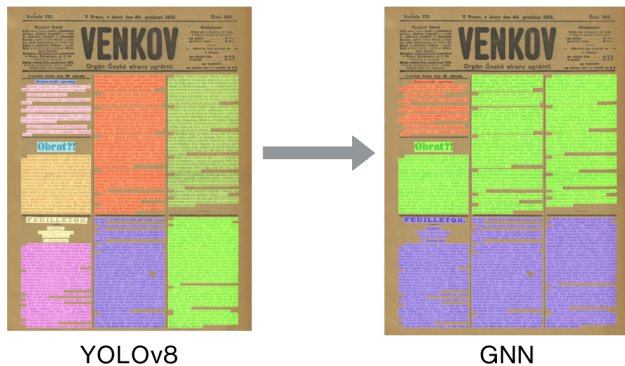
In total, 4044 pages were manually selected and labeled with the help of librarians and other students.

02 / Methodology

1. Regions of interest are detected using the YOLOv8 detector.
2. A complete graph is constructed from the detected regions. Each node is represented by both geometric and language features.

3. The graph is fed into a graph neural network, which works as a binary edge classifier. Regions connected by a positive edge are merged.

03 / Graph Neural Network Effect



04 / Evaluation and Results

The problem is defined as a line clustering problem. Each cluster represents a coherent text segment.

The baseline solution is based on imposing geometric constraints on the consecution of text lines, with additional distance based splitting or language model splitting.

The performance is evaluated using the V-Measure clustering metric (higher is better).

Method	V-Measure			
	Books	Dictionaries	Periodical	Time [s]
Baseline	19.72	40.11	52.20	0.16
Baseline+Dist	49.77	64.67	77.40	0.16
Baseline+LM	44.32	77.19	70.66	1.50
YOLOv8	65.33	93.36	83.93	0.14
YOLOv8+GNN	77.93	95.79	90.23	0.66

On NVIDIA RTX A6000

05 / Showcase

