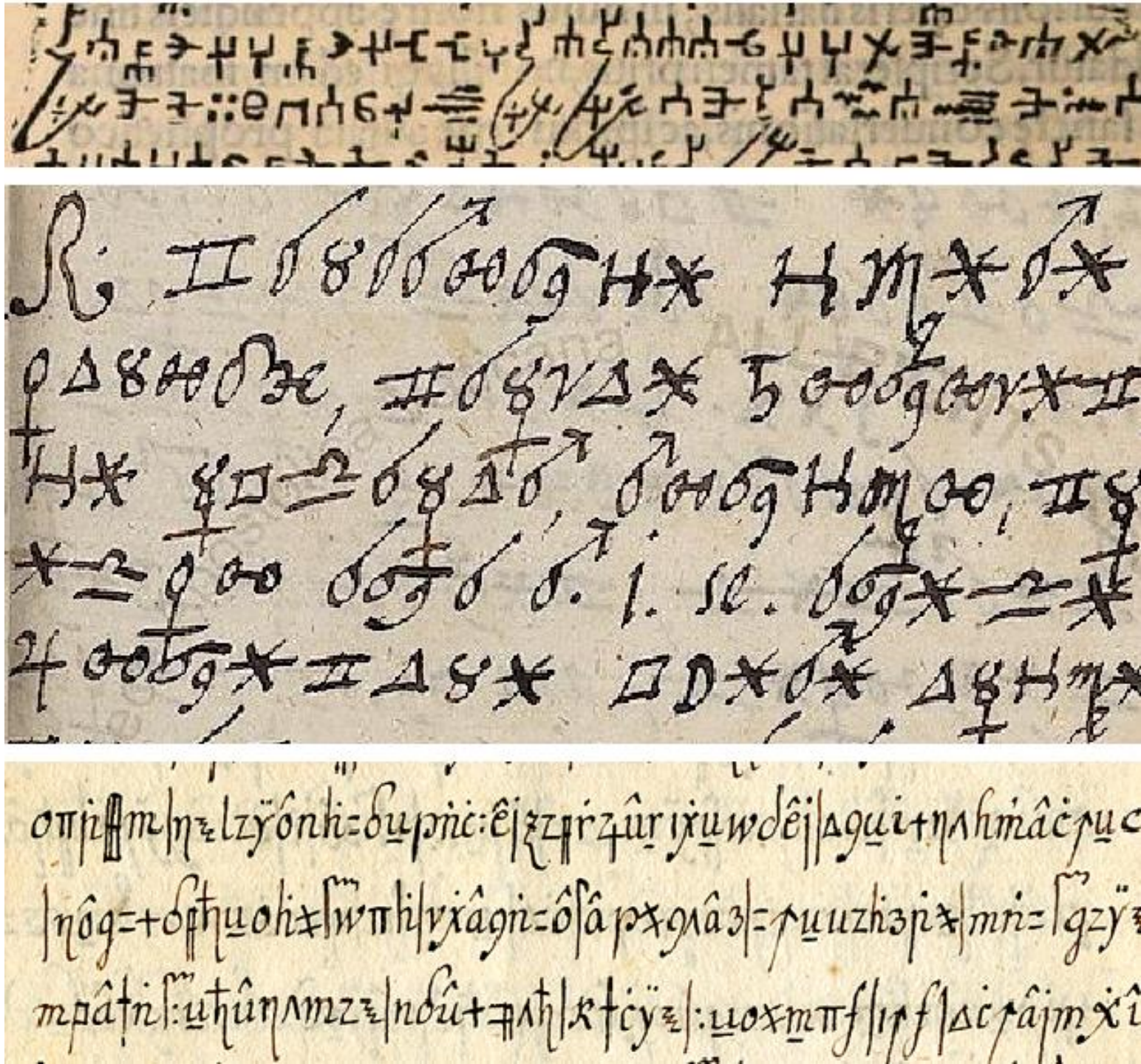
Handwritten Text (Character) Recognition

30.04.2024

Maria Levchenko / DHDK



References

- Ströbel, Phillip Benjamin, Simon Clematide, Martin Volk, and Tobias Hodel. "Transformer-based htr for historical documents." arXiv preprint arXiv:2203.11008 (2022).
- Muehlberger, G., Seaward, L., et all (2019), "Transforming scholarship in the archives through handwritten text recognition: Transkribus as a case study", Journal of Documentation, Vol. 75 No. 5, pp. 954-976. https://doi.org/10.1108/JD-07-2018-0114
- Li, Minghao, Tengchao Lv, Jingye Chen, Lei Cui, Yijuan Lu, Dinei Florencio, Cha Zhang, Zhoujun Li, and Furu Wei. 2023. "TrOCR: Transformer-Based Optical Character Recognition With Pre-Trained Models". Proceedings of the AAAI Conference on Artificial Intelligence 37 (11):13094-102. https://doi.org/10.1609/aaai.v37i11.26538.
- Najem-Meyer, Sven and Matteo Romanello. "Page Layout Analysis of Text-heavy Historical Documents: a Comparison of Textual and Visual Approaches." Workshop on Computational Humanities Research (2022).
- Matteo Romanello, Sven Najem-Meyer, and Bruce Robertson. 2021. Optical Character Recognition of 19th Century Classical Commentaries: the Current State of Affairs. In Proceedings of the 6th International Workshop on Historical Document Imaging and Processing (HIP '21). Association for Computing Machinery, New York, NY, USA, 1–6. https://doi.org/10.1145/3476887.3476911

Ayush Purohit et al, A Literature Survey on Handwritten Character Recognition,(IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (1) , 2016, 1-5

Survey on Image Preprocessing Techniques to Improve OCR Accuracy https://medium.com/technovators/survey-on-image-preprocessing-techniques-to-improve-ocr-accuracy-616ddb931b76

F. Simistira et al., "ICDAR2017 Competition on Layout Analysis for Challenging Medieval Manuscripts," 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 2017, pp. 1361-1370, doi: 10.1109/ICDAR.2017.223.

Clérice, Thibault. (2022). You Actually Look Twice At it (YALTAi): using an object detection approach instead of region segmentation within the Kraken engine. 10.48550/arXiv.2207.11230.

Fizaine FC, Bard P, Paindavoine M, Robin C, Bouyé E, Lefèvre R, Vinter A. Historical Text Line Segmentation Using Deep Learning Algorithms: Mask-RCNN against U-Net Networks. Journal of Imaging. 2024; 10(3):65. https://doi.org/10.3390/jimaging10030065

Leifert, Gundram, Christel Annemieke Romein, Achim Rabus, Phillip Benjamin Ströbel, Benjamin Kiessling, & Tobias Hödel. Evaluating State-of-the-art Handwritten Text Recognition (HTR) Engines; with Large Language Models (llms) for Historical Document Digitisation. Zenodo, 7 December 2023 r. https://doi.org/10.5281/zenodo.8102666

Weidemann, M., Michael, J., Gruning, T., and Labahn, R. ¨ (2018). HTR Engine Based on NNs P2 Building Deep Architectures with TensorFlow. Technical report.

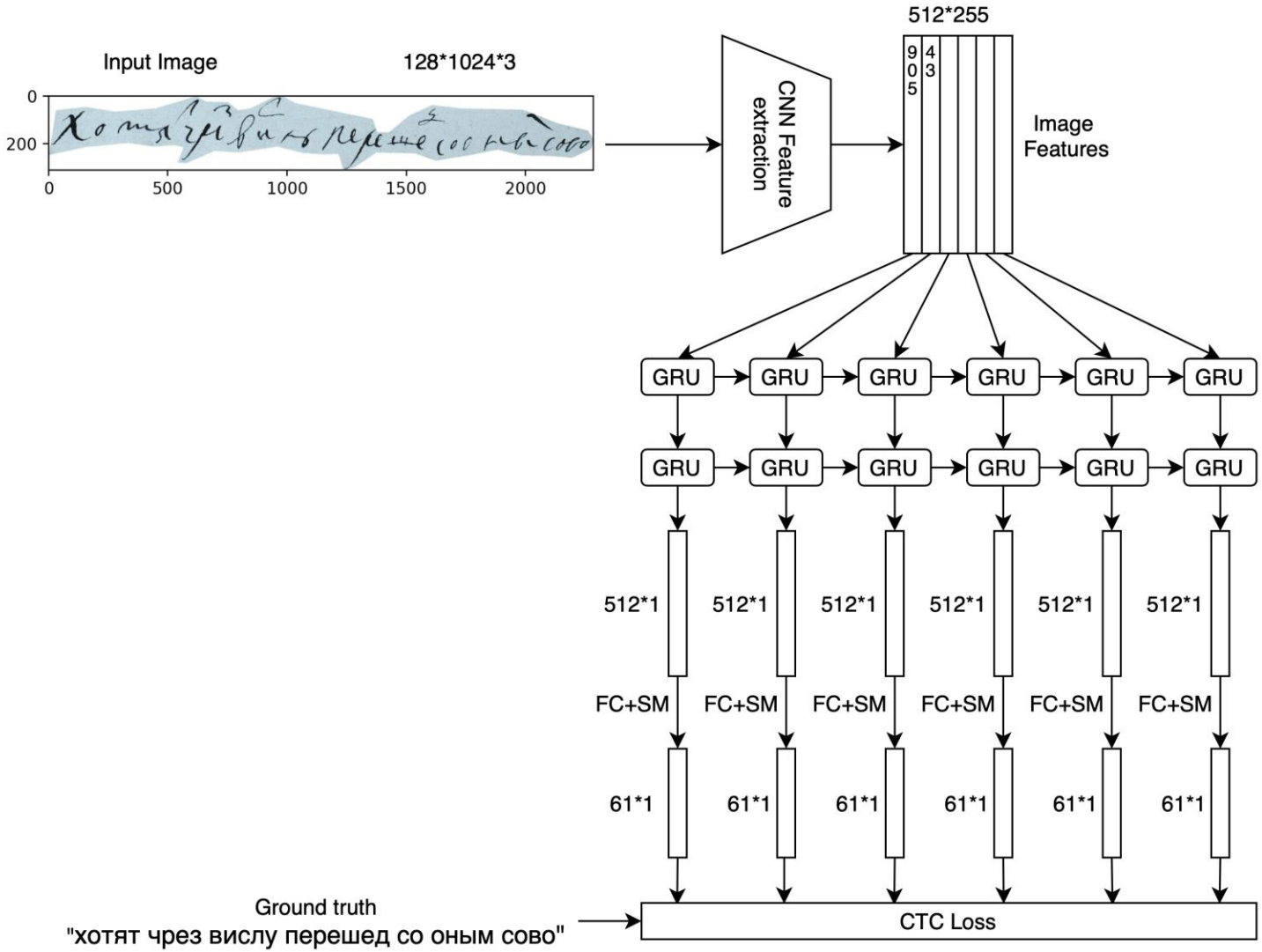
Peter Stokes, Benjamin Kiessling. Sharing Data for Handwritten Text Recognition (HTR). Digital Humanities in Practice, In press. ffhal-04444641f

Unread references (for the future)

- Petitpierre, R., Kramer, M. & Rappo, L. An end-to-end pipeline for historical censuses processing. *IJDAR* **26**, 419–432 (2023). https://doi.org/10.1007/s10032-023-00428-9
- Kass, Dmitrijs, and Ekta Vats. "AttentionHTR: Handwritten text recognition based on attention encoder-decoder networks." In *International Workshop on Document Analysis Systems*, pp. 507-522. Cham: Springer International Publishing, 2022.
- Ströbel, Phillip & Clematide, Simon & Volk, Martin & Schwitter, Raphael & Hodel, Tobias & Schoch, David. (2022). Evaluation of HTR models without Ground Truth Material.
- de Sousa Neto, A.F., Bezerra, B.L.D., de Moura, G.C.D. *et al.* Data Augmentation for Offline Handwritten Text Recognition: A Systematic Literature Review. *SN COMPUT. SCI.* **5**, 258 (2024). https://doi.org/10.1007/s42979-023-02583-6
- Ströbel, Phillip & Hodel, Tobias & Boente, Walter & Volk, Martin. (2023). The Adaptability of a Transformer-Based OCR Model for Historical Documents. 10.1007/978-3-031-41498-5_3.
- Ströbel, Phillip & Clematide, Simon & Volk, Martin. (2020). How Much Data Do You Need? About the Creation of a Ground Truth for Black Letter and the Effectiveness of Neural OCR.
- AlKendi, Wissam, Franck Gechter, Laurent Heyberger, and Christophe Guyeux. 2024. "Advancements and Challenges in Handwritten Text Recognition: A Comprehensive Survey" *Journal of Imaging* 10, no. 1: 18. https://doi.org/10.3390/jimaging10010018

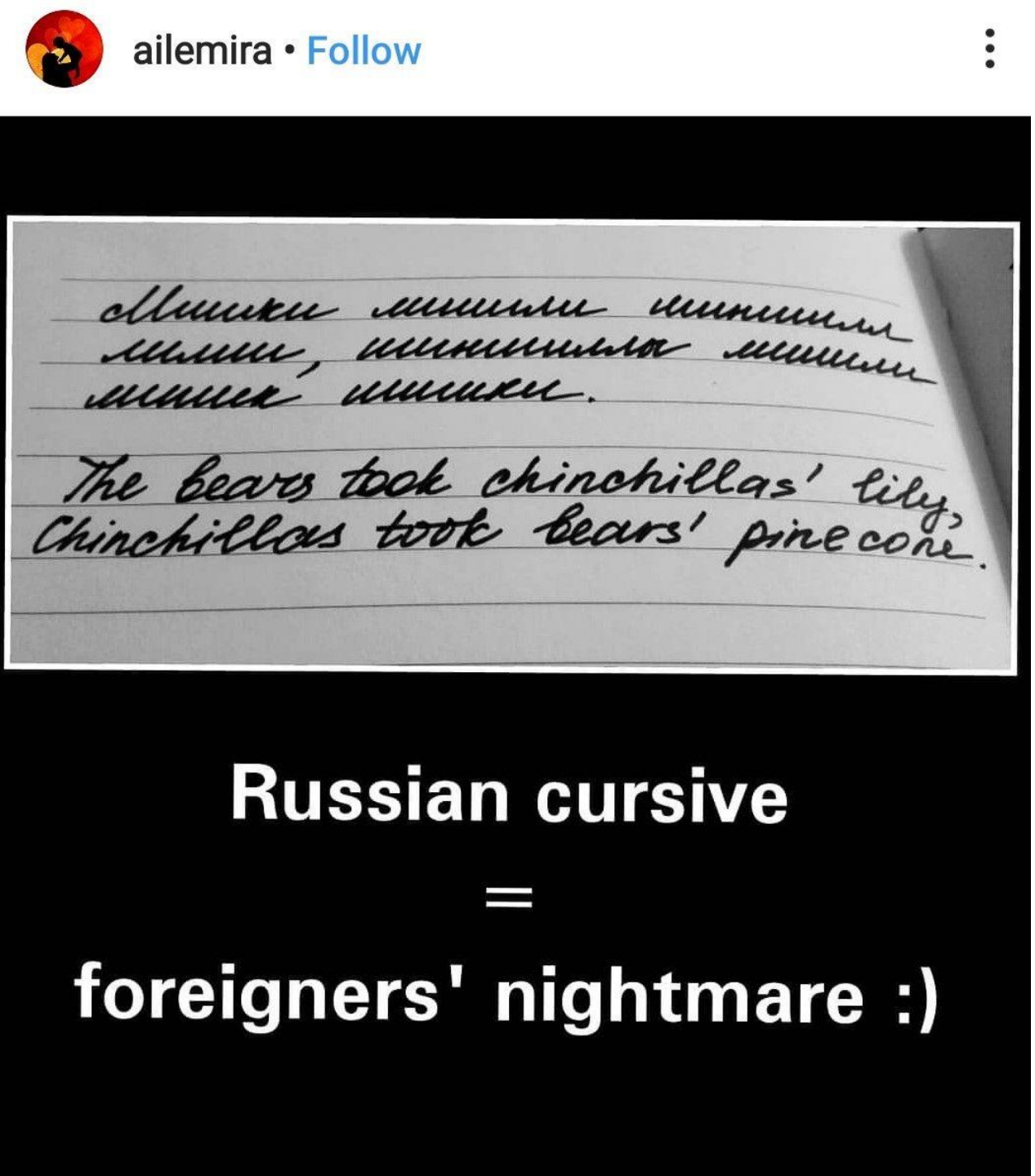
HTR:
automatic transcription of handwritten texts in manuscripts
(and early printed books)

- postal address interpretation,
- bank-cheque processing,
- signature verification,
- biometric writer identification,
- **manuscripts, documents, archives**

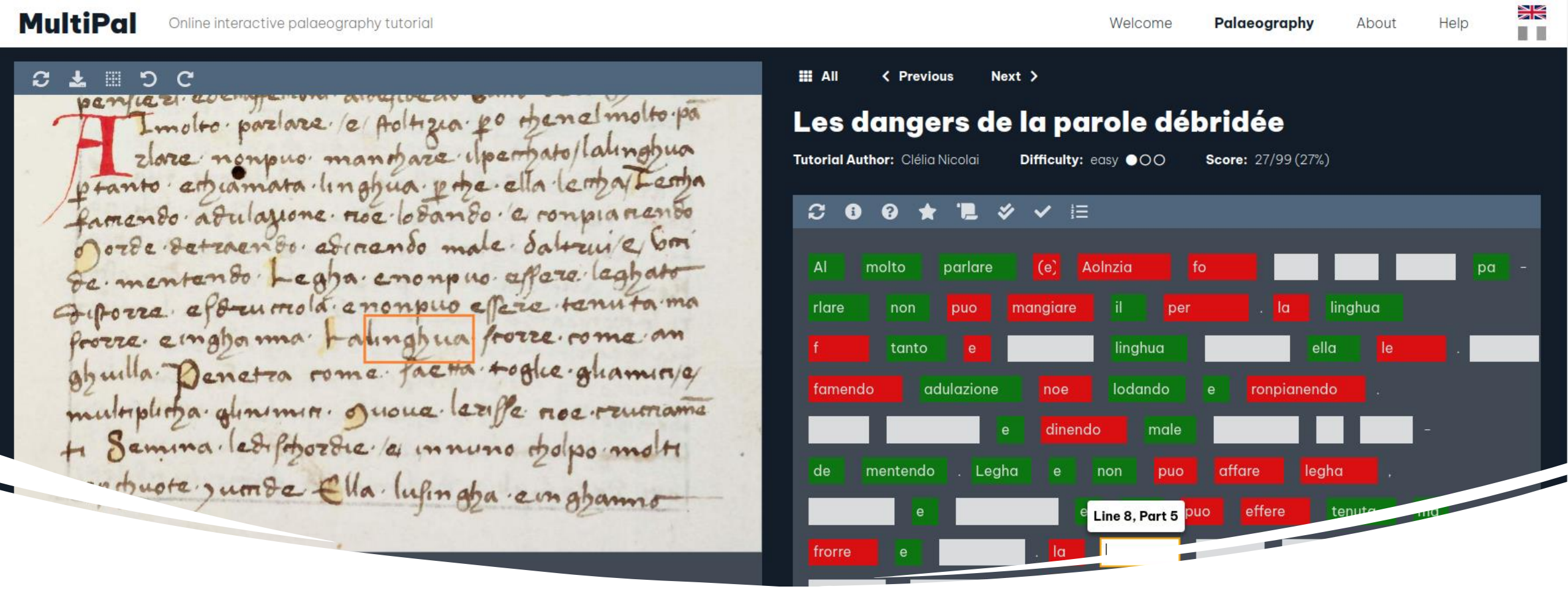


https://github.com/ai-forever/digital_peter_aij2020

Still tough, but even for humans



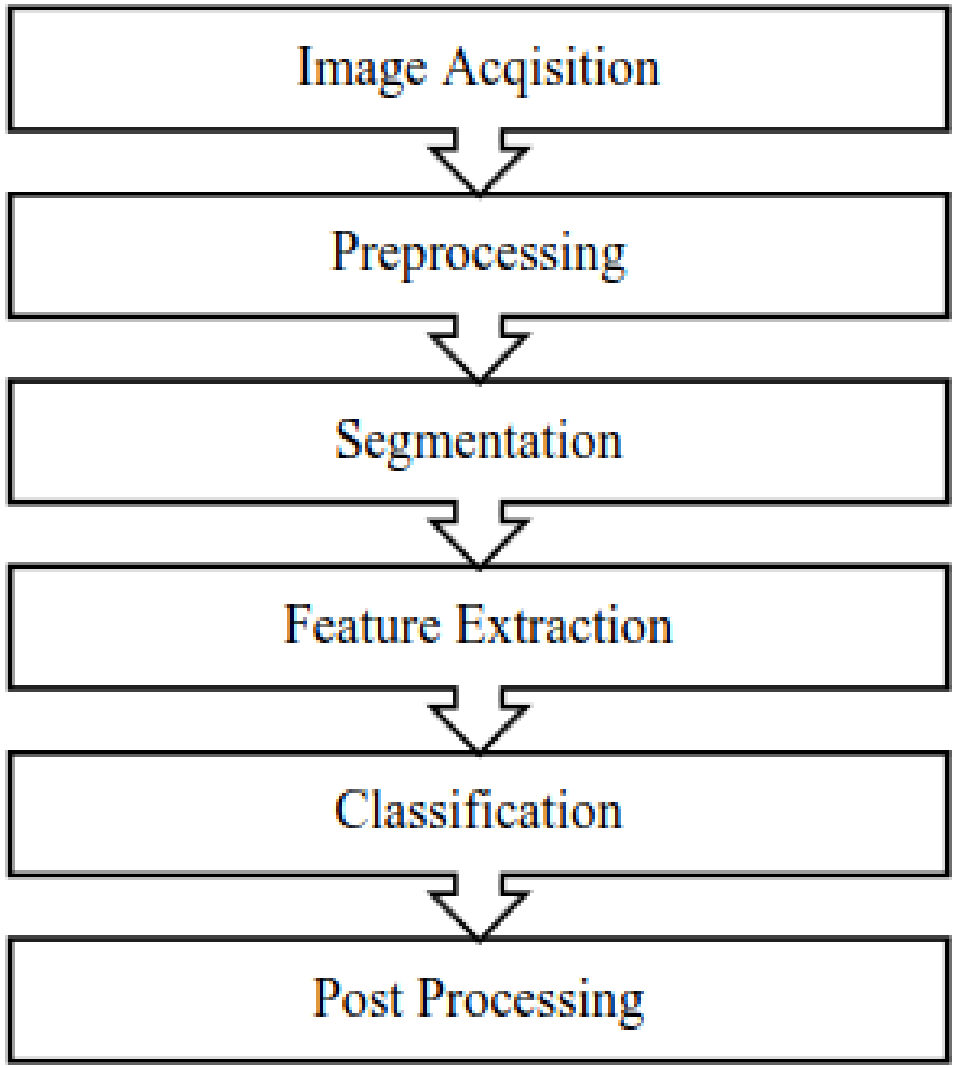
Unreadable
by non-natives



Unreadable by non-experts

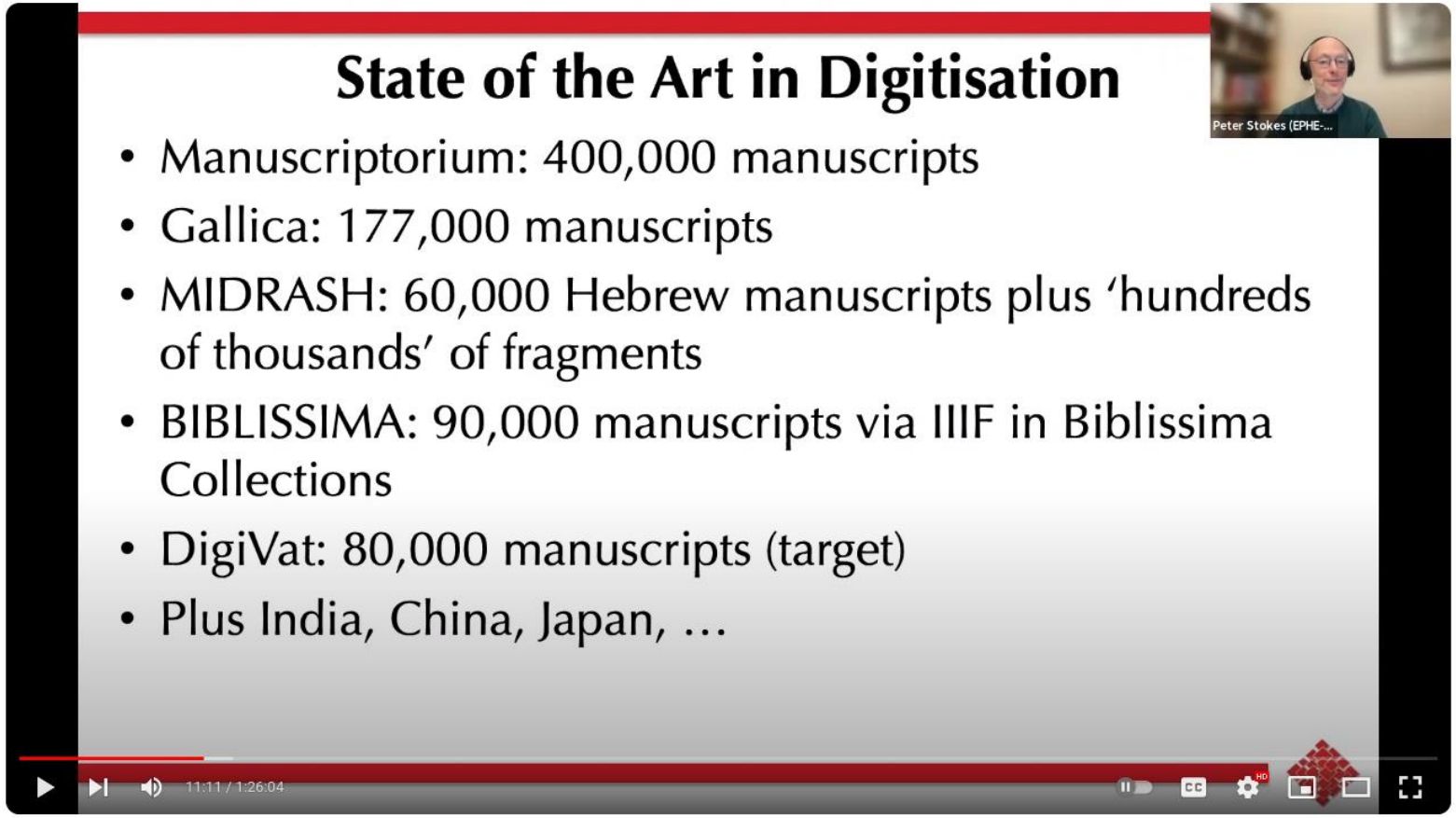
Test your paleographer skills: https://www.multipal.fr

HTR Pipeline



Ayush Purohit et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (1) , 2016, 1-5

Image Acquisition



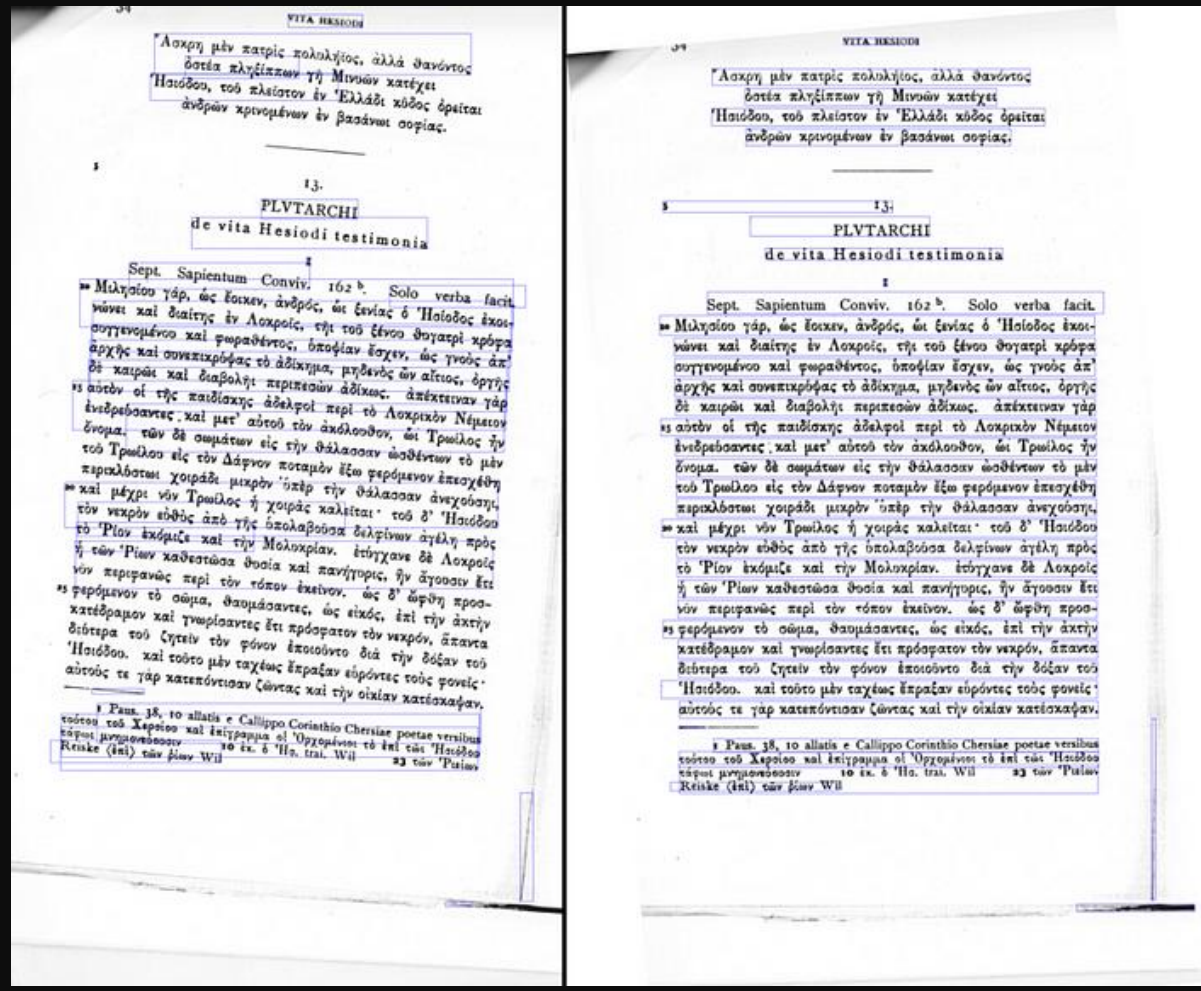
How to Transcribe a Million Manuscripts with eScriptorium

[How to Transcribe a Million Manuscripts with eScriptorium]

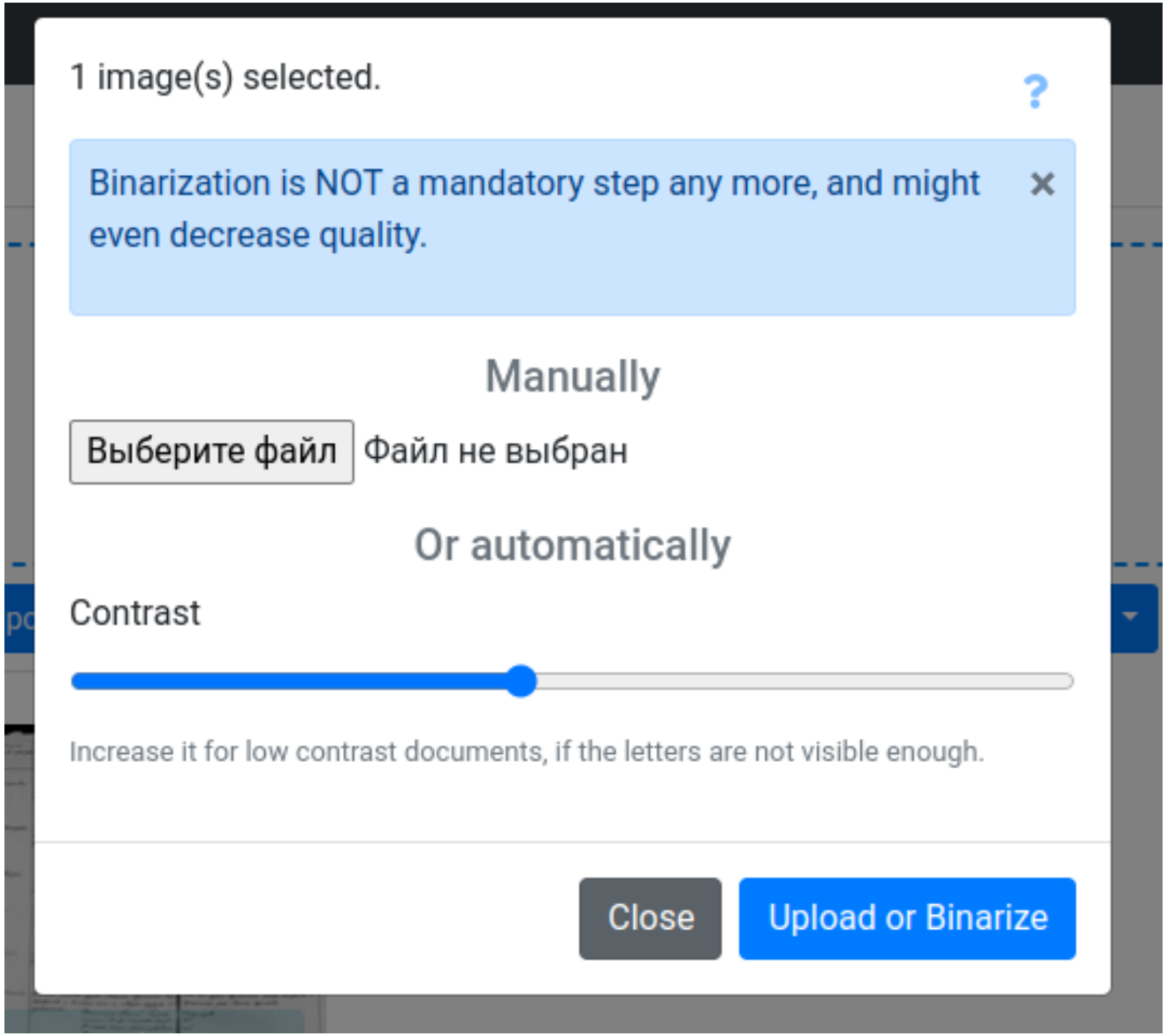
Preprocessing: Image Quality

- Resolution
- Contrast & Sharpness
- Geometric transformations and so on

Survey on Image Preprocessing Techniques to Improve OCR Accuracy https://medium.com/technovators/survey-on-image-preprocessing-techniques-to-improve-ocr-accuracy-616ddb931b76



Preprocessing: Binarisation?



Segmentation

Why is it important?

- Reduces document complexity,
- Improves the accuracy of recognition algorithms,
- Is required for further analysis and automatic markup at the post-processing and publishing stages

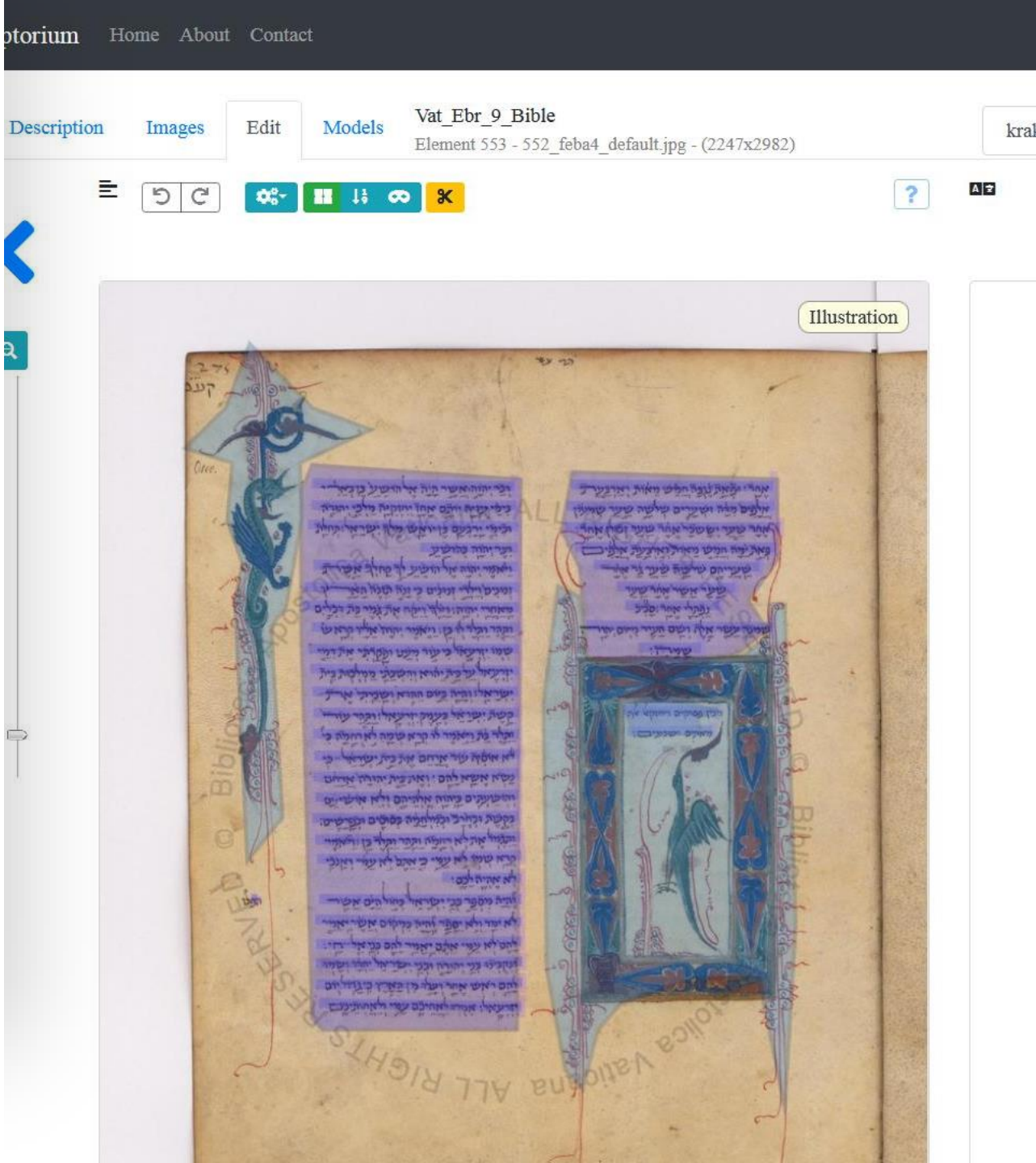
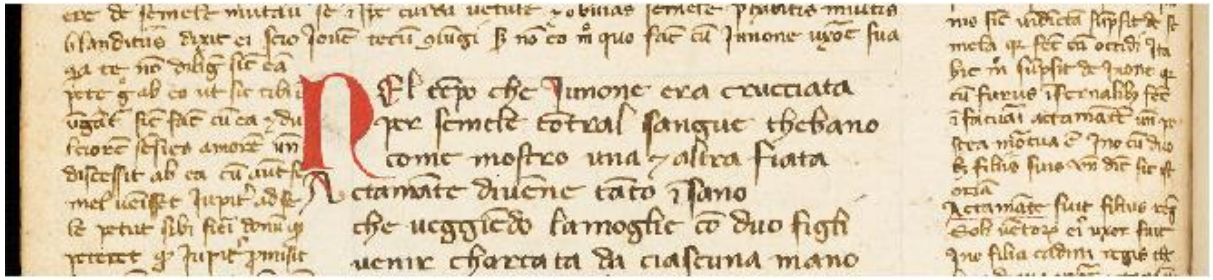


Figure 2: The main layout elements of a scholarly commentary page.

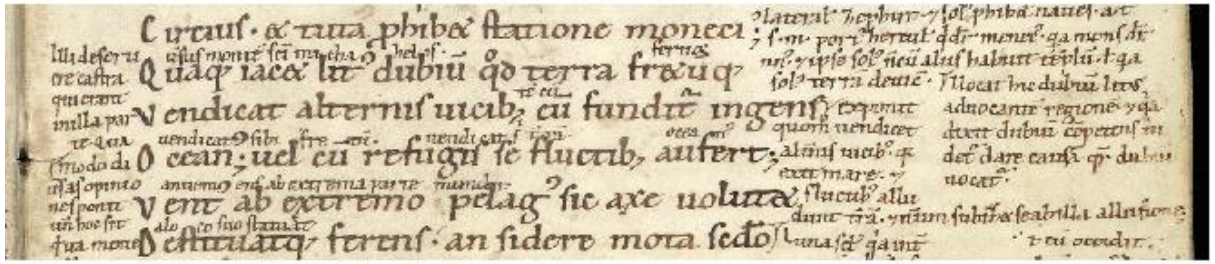
Najem-Meyer, Sven and Matteo Romanello. “Page Layout Analysis of Text-heavy Historical Documents: a Comparison of Textual and Visual Approaches.” *Workshop on Computational Humanities Research* (2022).

Most segmenters focus on pixel classification



(a) CB55, p.67v

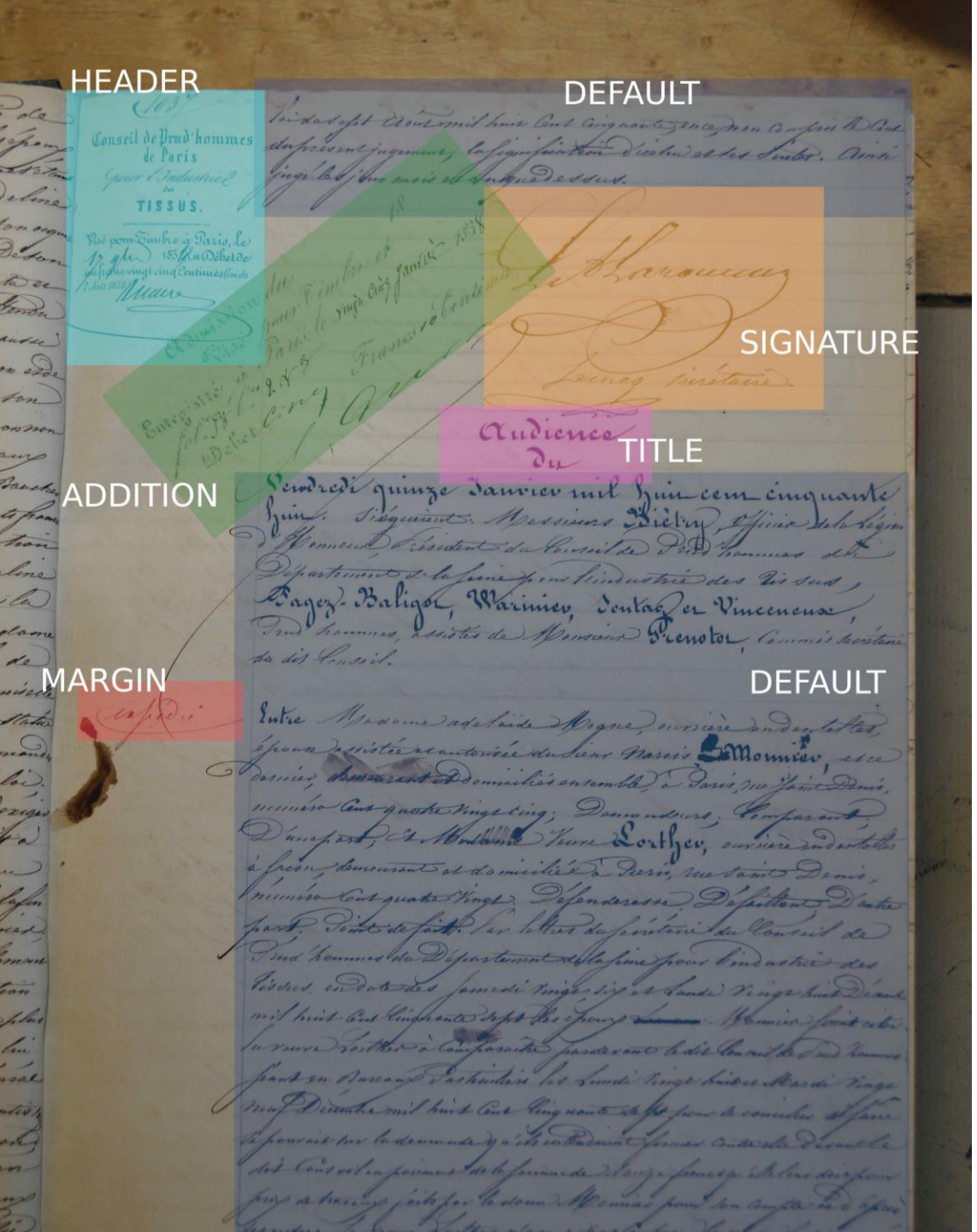
(b) CSG18, p.116



(c) CSG863, p.17

Figure 1: Samples of pages of the three medieval manuscripts

F. Simistira *et al.*, "ICDAR2017 Competition on Layout Analysis for Challenging Medieval Manuscripts," *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*, Kyoto, Japan, 2017, pp. 1361-1370, doi: 10.1109/ICDAR.2017.223.



Segmentation: SegmOnto

https://segmonto.github.io/

A Controlled Vocabulary to Describe the Layout of Pages

SegmOnto Region classes

... to make the published dataset as widely reusable as possible, we mapped our classes to the SegmOnto controlled vocabulary (Romanello, Najem-Meyer 2022)

Table 1
Complete list of fine- and coarse-grained page region classes used for layout annotation, with their corresponding mapping to SegmOnto's controlled vocabulary.

Fine	Coarse	SegmOnto Type:Subtype
commentary	commentary	MainZone:commentary
critical apparatus	critical apparatus	MarginTextZone:criticalApparatus
footnotes	footnotes	MarginTextZone:footnotes
page number	number	NumberingZone:pageNumber
text number	number	NumberingZone:textNumber
bibliography	others	MainZone:bibliography
handwritten marginalia	others	MarginTextZone:handwrittenNote
index	others	MainZone:index
others	others	CustomZone
printed marginalia	others	MarginTextZone:printedNote
table of contents	others	MainZone:ToC
title	others	TitlePageZone
translation	others	MainZone:translation
appendix	paratext	MainZone:appendix
introduction	paratext	MainZone:introduction
preface	paratext	MainZone:preface
primary text	primary text	MainZone:primaryText
running header	running header	RunningTitleZone

Zone & Lines types

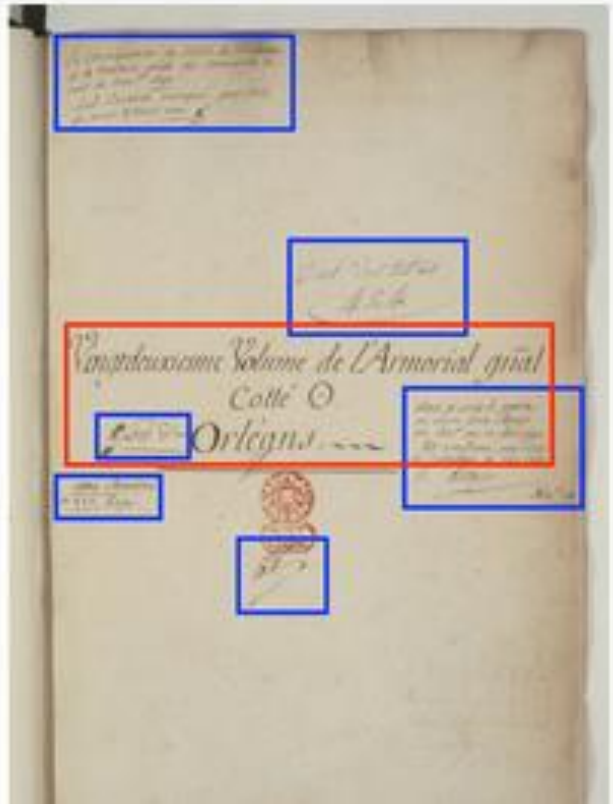
Why it is important?

MainZone:column#1

MainZone:column#2

TitlePageZone

and MarginTextZone

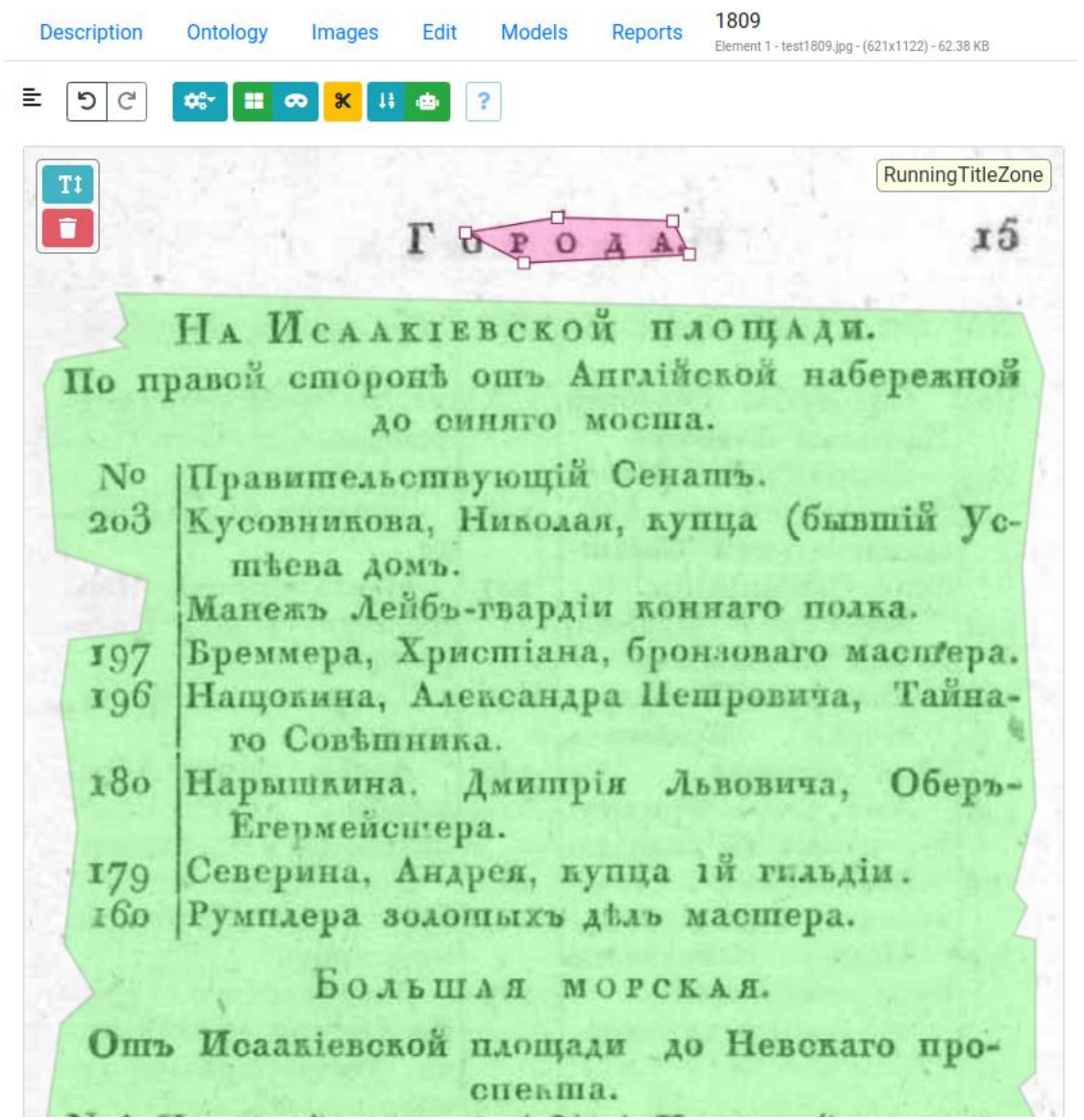


SegmOnto model

Layout analysis model trained with YALTAi, relying on YOLO models, and Kraken. Data are annotated with the SegmOnto controlled vocabulary.

https://zenodo.org/records/10972956

Published April 15, 2024



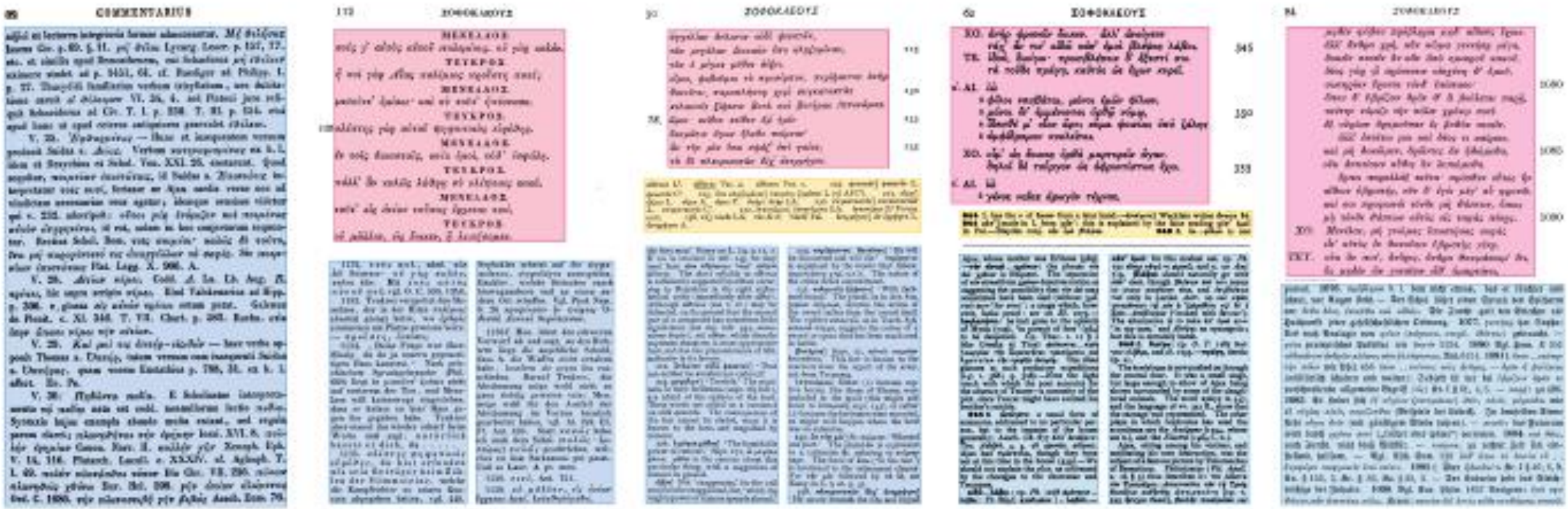


Figure 1: Example pages from 19th century commentaries on Sophocles' *Ajax*. The commentaries are by (from left to right): Lobeck (1835), Schneidewin (1853), Campbell (1881), Jebb (1896) and Wecklein (1894). Commentary sections are highlighted in blue, primary texts in pink and critical apparatus in yellow.

Image annotation tools

- **VGG Image Annotator (VIA)** https://www.robots.ox.ac.uk/~vgg/software/via/
- **CVAT (Computer Vision Annotation Tool)** https://www.cvat.ai/
- For manuscripts / books:
  - eScriptorium
  - Transkribus

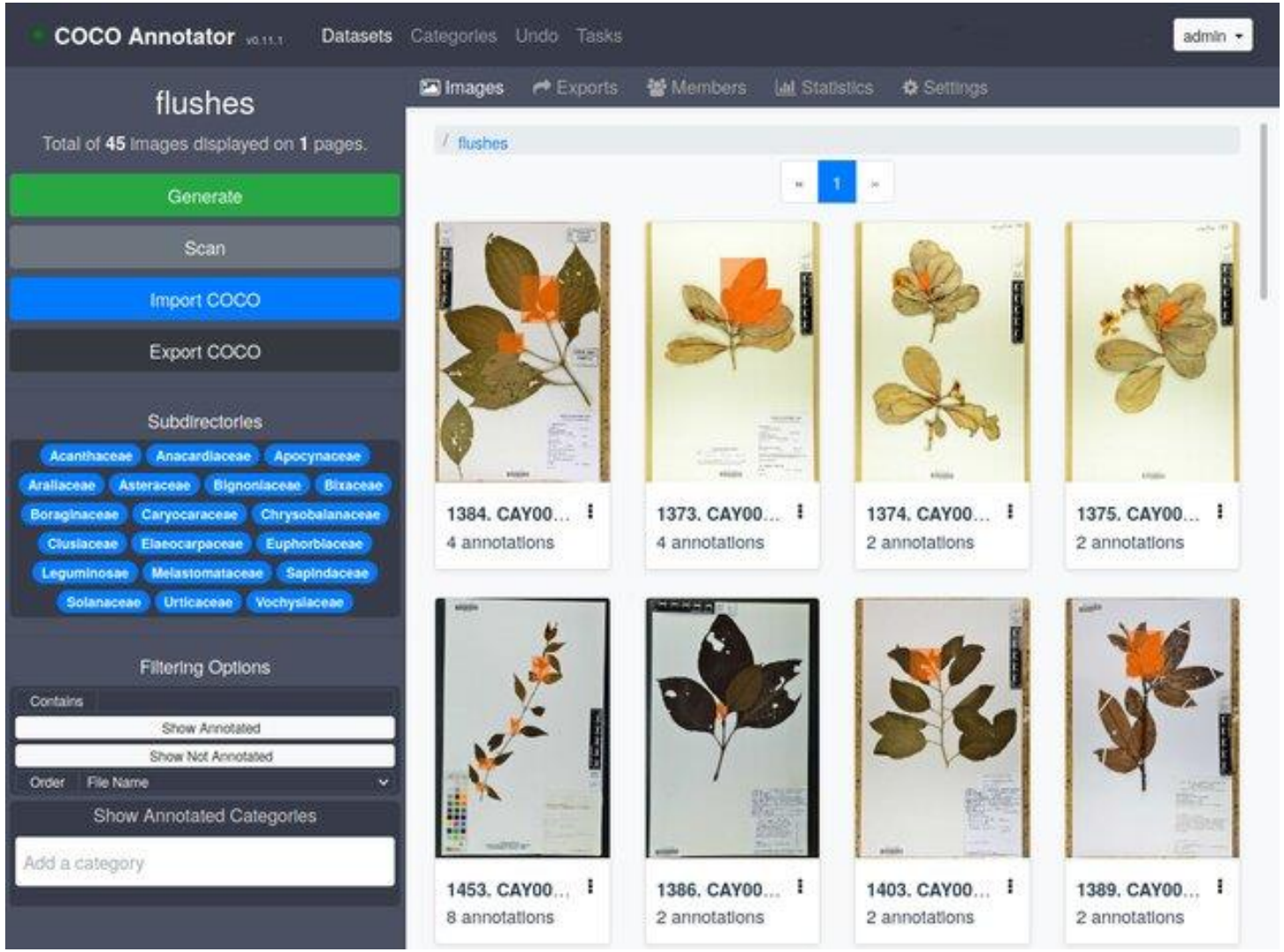
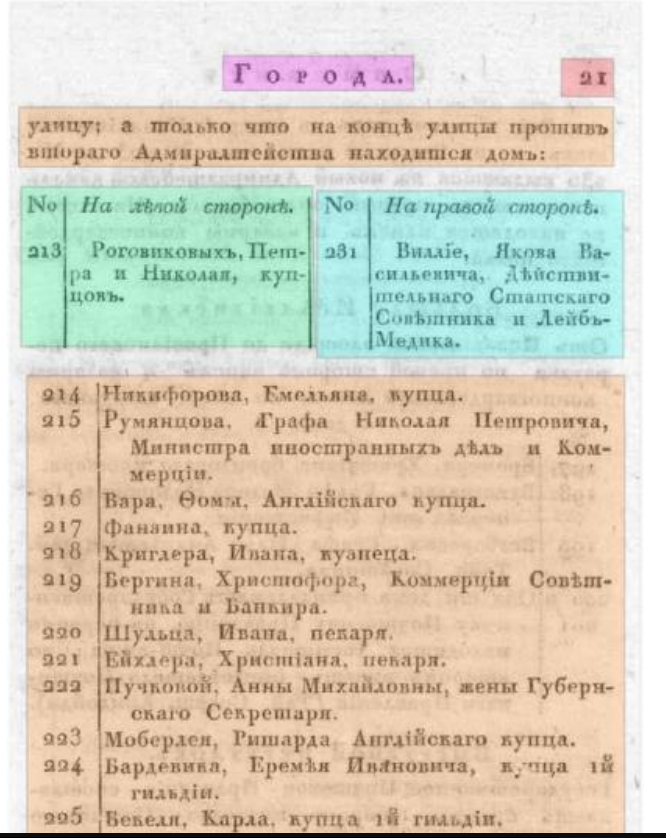


Image annotation

- Define an ontology of your segments and lines regarding your project goals
- Manually annotate (or use a model and then correct the result manually) a sample
- Train a model
- Check the output
- Repeat



Города. 21

улицу; а только что на концѣ улицы прошивъ второго Адмиралтейства находится домъ:

No	На лѣвой сторонѣ.	No	На правой сторонѣ.
213	Роговиковыхъ, Петра и Николая, купцовъ.	231	Виліе, Якова Васильевича, Дѣйствительнаго Статскаго Совѣтника и Лейбъ-Медика.

214 Никифорова, Емельяна, купца.
215 Румянцова, Графа Николая Петровича, Министра иностранныхъ дѣлъ и Коммерціи.
216 Вара, Ѳомы, Англійскаго купца.
217 Фанзина, купца.
218 Кригера, Ивана, кузнеца.
219 Бергина, Христофора, Коммерціи Совѣтника и Банкира.
220 Шульца, Ивана, пекаря.
221 Ейхлера, Христіана, пекаря.
222 Пучковой, Анны Михайловны, жены Губернскаго Секретаря.
223 Моберлея, Ришарда, Англійскаго купца.
224 Бардевика, Еремѣя Ивановича, купца 1й гильдіи.
225 Бекеля, Карла, купца 1й гильдіи.

- 0 0.929952 0.064840 0.079710 0.025401
- 2 0.248792 0.233957 0.479469 0.144385
- 2 0.729469 0.235294 0.478261 0.145722
- 2 0.507246 0.643048 0.948068 0.617647
- 2 0.489130 0.117647 0.955314 0.056818
- 1 0.489130 0.062166 0.299517 0.036096

```
<Tags>
  <OtherTag ID="BT45" LABEL="NumberingZone:page" DESCRIPTION="block type NumberingZone:page"/>
  <OtherTag ID="BT46" LABEL="MainZone" DESCRIPTION="block type MainZone"/>
  <OtherTag ID="BT47" LABEL="RunningTitleZone" DESCRIPTION="block type RunningTitleZone"/>
  <OtherTag ID="BT48" LABEL="MainZone:column#1" DESCRIPTION="block type MainZone:column#1"/>
  <OtherTag ID="BT49" LABEL="MainZone:column#2" DESCRIPTION="block type MainZone:column#2"/>
</Tags>
<Layout>
  <Page WIDTH="621"
        HEIGHT="1122"
        PHYSICAL_IMG_NR="14"
        ID="eSc_dummypage_">
    <PrintSpace HPOS="0"
                VPOS="0"
                WIDTH="621"
                HEIGHT="1122">
      <TextBlock HPOS="304"
                 VPOS="187"
                 WIDTH="303"
                 HEIGHT="165"
                 ID="eSc_textblock_28141c0c"
                 TAGREFS="BT49">
        <Shape><Polygon POINTS="304 187 304 352 607 352 607 187"/></Shape>
      </TextBlock>
      <TextBlock HPOS="211"
                 VPOS="53"
                 WIDTH="188"
                 HEIGHT="34"
                 ID="eSc_textblock_33338d66"
                 TAGREFS="BT47">
        <Shape><Polygon POINTS="211 53 211 87 399 87 399 53"/></Shape>
      </TextBlock>
```

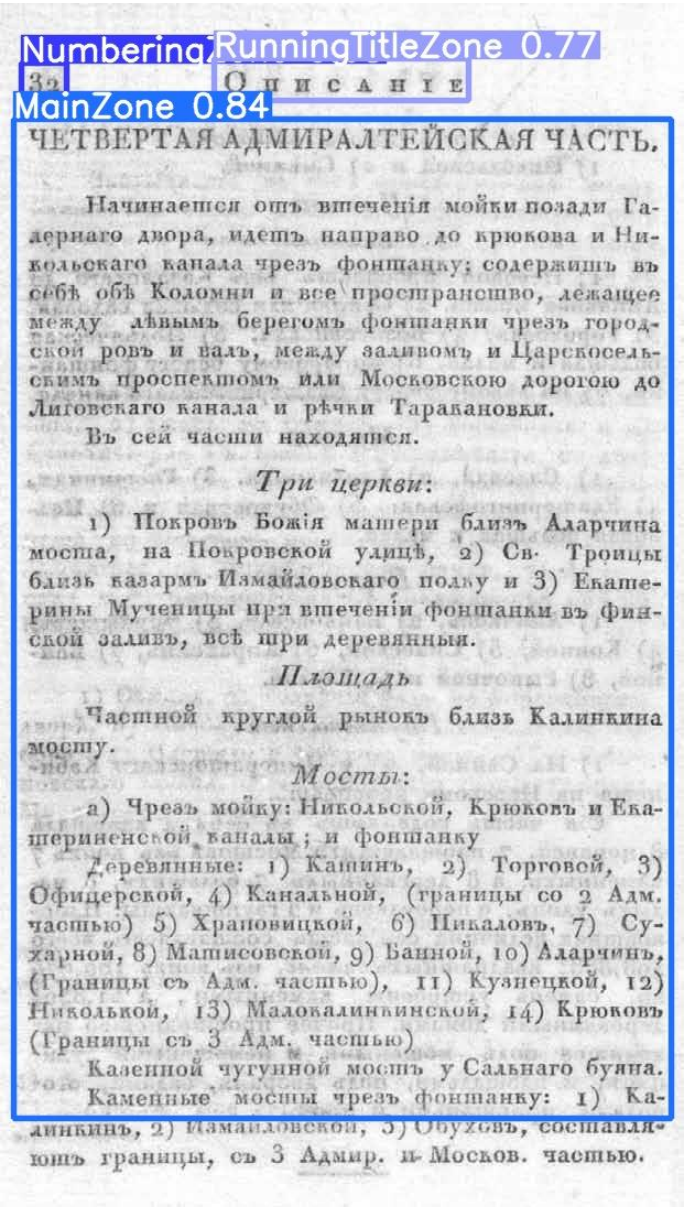
Training a model: YOLOv8

```
(venv) maria@maria-H410M-S2H:~/projects/spb$ yaltai alto-to-yolo export/*.xml my-dataset --shuffle .1 --segmonto region
Using list of inputs.
Found 20 to convert.
3/20 image for validation.
Shuffling data with a ratio of 0.1 for validation.
20it [00:00, 1246.01it/s]
20 ground truth XML files converted.
Configuration available at my-dataset/config.yml.
Label Map available at my-dataset/labelmap.txt.
Regions count:
        - 00040 MainZone
        - 00020 RunningTitleZone
        - 00020 NumberingZone
```

You only look once (*YOLO*)

Train and Predict

```
      Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
     97/100         0G      1.159       1.81      1.233          6        960: 100%|██████████| 3/3 [00:06<00:00,  2.01s/it]
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 1/1 [00:00<00:00,  2.18it/s]
                   all          6         28      0.639      0.683      0.597      0.434

      Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
     98/100         0G     0.9454       1.63      1.281          7        960: 100%|██████████| 3/3 [00:05<00:00,  1.79s/it]
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 1/1 [00:00<00:00,  1.98it/s]
                   all          6         28      0.639      0.683      0.597      0.434

      Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
     99/100         0G      1.012      1.812       1.21          3        960: 100%|██████████| 3/3 [00:05<00:00,  1.99s/it]
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 1/1 [00:00<00:00,  1.72it/s]
                   all          6         28      0.639      0.683      0.597      0.434

      Epoch    GPU_mem   box_loss   cls_loss   dfl_loss  Instances       Size
    100/100         0G     0.9336      1.666       1.17          3        960: 100%|██████████| 3/3 [00:05<00:00,  1.97s/it]
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 1/1 [00:00<00:00,  2.15it/s]
                   all          6         28      0.632      0.683      0.596       0.43

100 epochs completed in 0.186 hours.
Optimizer stripped from runs/detect/train/weights/last.pt, 6.3MB
Optimizer stripped from runs/detect/train/weights/best.pt, 6.3MB

Validating runs/detect/train/weights/best.pt...
Ultralytics YOLOv8.0.209 🚀 Python-3.10.12 torch-2.0.1+cu117 CPU (Intel Core(TM) i9-10900 2.80GHz)
Model summary (fused): 168 layers, 3006233 parameters, 0 gradients, 8.1 GFLOPs
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95): 100%|██████████| 1/1 [00:00<00:00,  2.44it/s]
                   all          6         28      0.642      0.683      0.599      0.458
              MainZone          6         12      0.502       0.75      0.533      0.469
      RunningTitleZone          6          6      0.943          1      0.995       0.78
         NumberingZone          6         10       0.48        0.3      0.268      0.124
Speed: 1.3ms preprocess, 49.6ms inference, 0.0ms loss, 7.1ms postprocess per image
Results saved to runs/detect/train
```



x1: tensor(12.9008, device='cuda:0') y1: tensor(105.7754, device='cuda:0') x2: tensor(609.1998, device='cuda:0') y2: tensor(1012.1533, device='cuda:0')
x1: tensor(194.3830, device='cuda:0') y1: tensor(50.6112, device='cuda:0') x2: tensor(425.4306, device='cuda:0') y2: tensor(88.9174, device='cuda:0')
x1: tensor(19.5257, device='cuda:0') y1: tensor(54.7710, device='cuda:0') x2: tensor(60.7622, device='cuda:0') y2: tensor(87.9373, device='cuda:0')

Model architecture for page layout

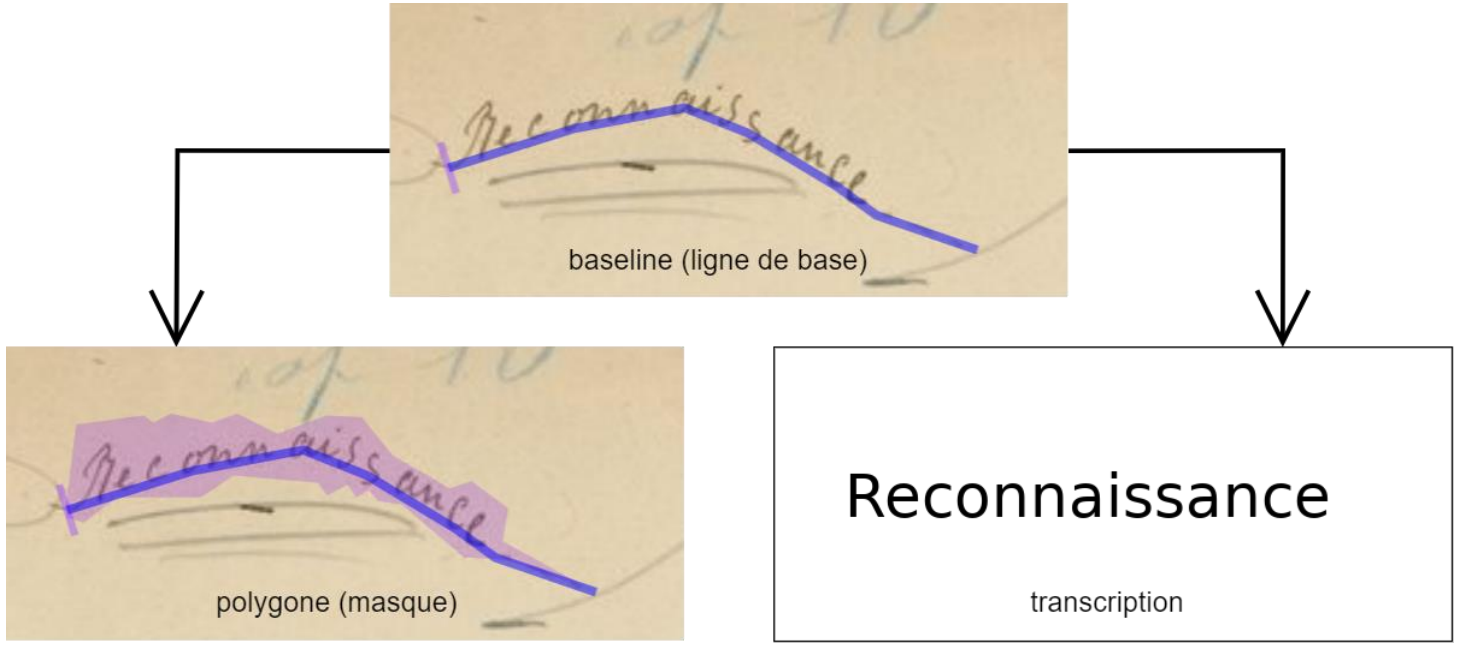
- Convolutional Neural Networks (CNNs)

From polygon (or bounding box) detection to pixel classification-based polygonization and now to object detection using isothetic rectangles

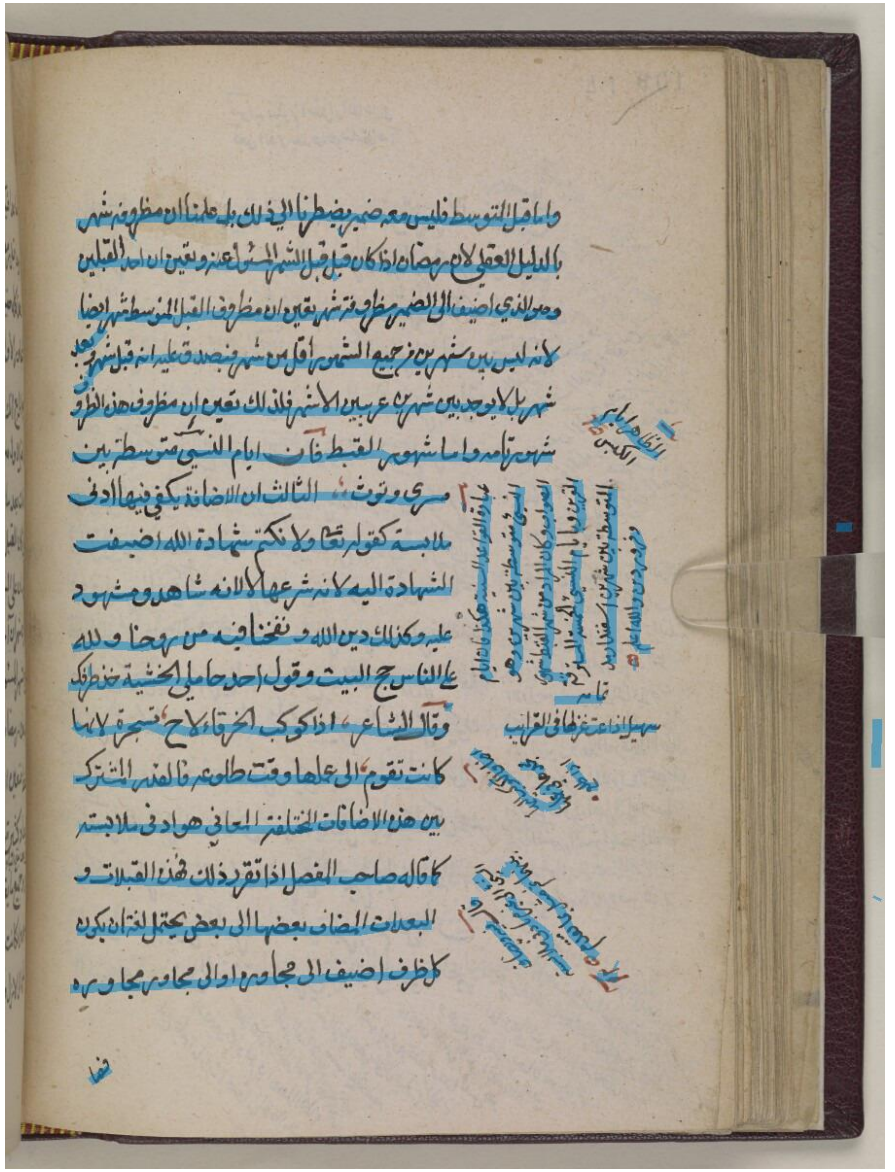
Clérice, Thibault. (2022). You Actually Look Twice At it (YALTAi): using an object detection approach instead of region segmentation within the Kraken engine. 10.48550/arXiv.2207.11230.

Baselines and Line Masks

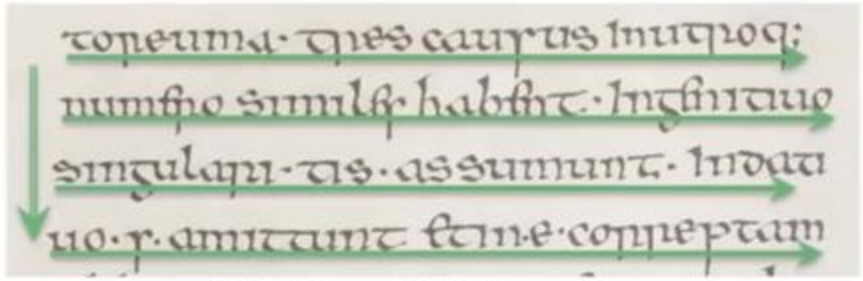
- **Use Geometric Analysis**
- **Apply Smoothing and Filtering**
- **Detect Local Minima/Maxima**
- **Fit Lines or Curves**



Baselines



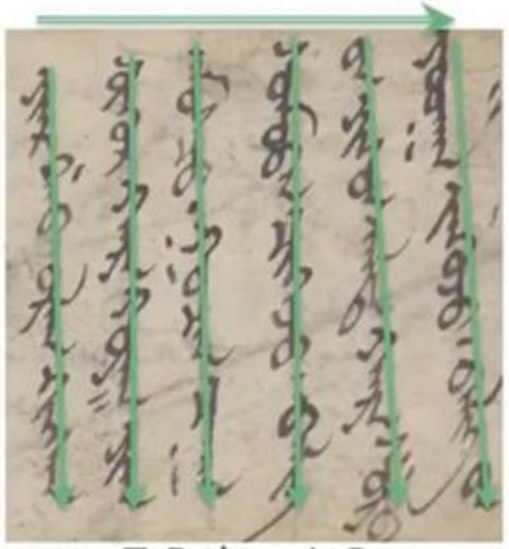
A ‘simple’ example: writing line & direction



LtR then TtB, baseline

RtL then TtB, topline

TtB then RtL,
L & R column lines

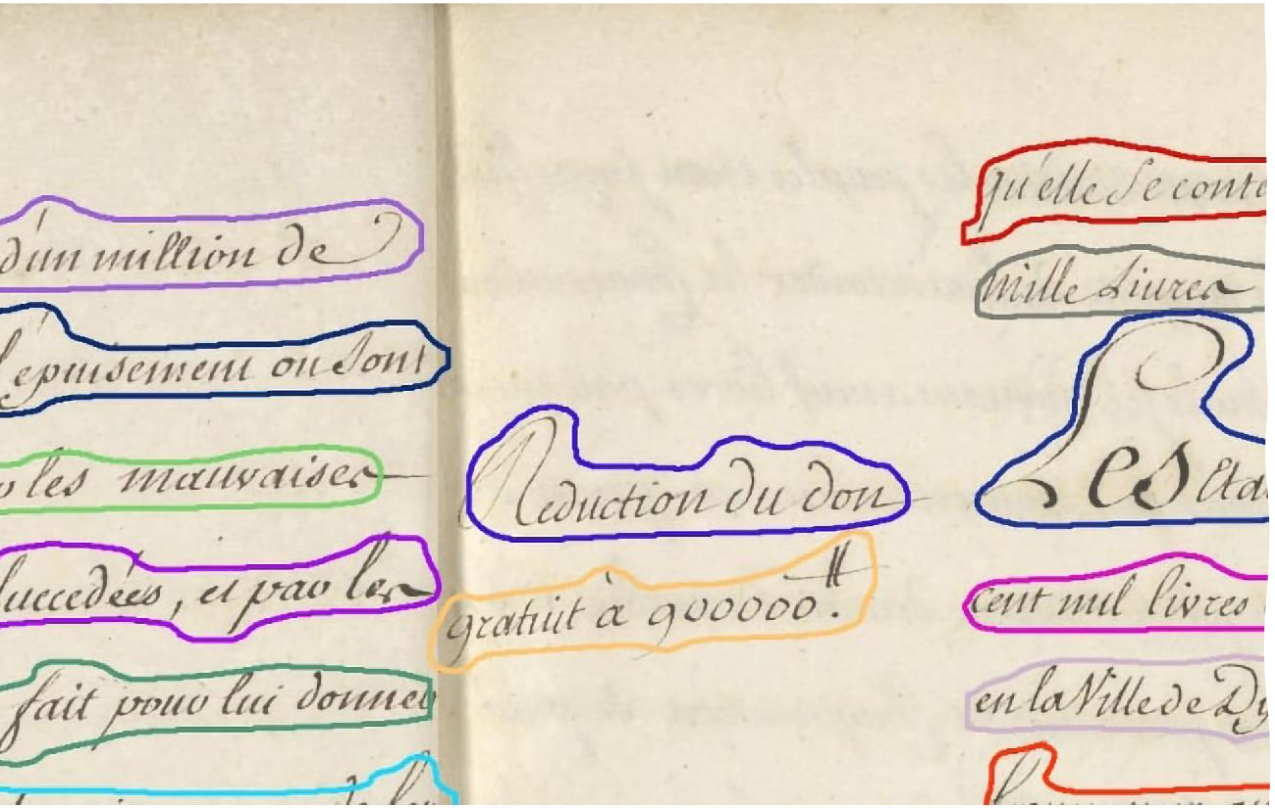


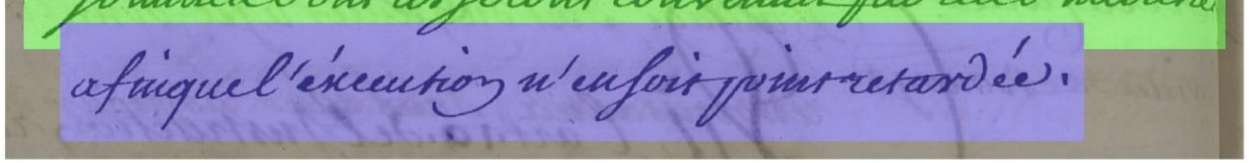
TtB then LtR,
vertical right ‘base’ line

BtT, implicit
column lines

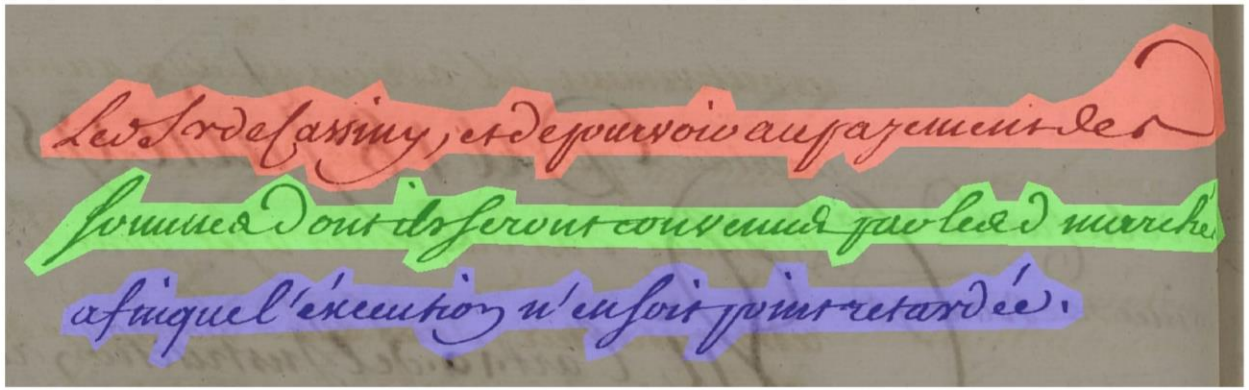
Clockwise,
internal baseline

- [01 OPEN HTR WITH ESCRIPTORIUM AND KRAKEN - Peter Stokes]





(b)



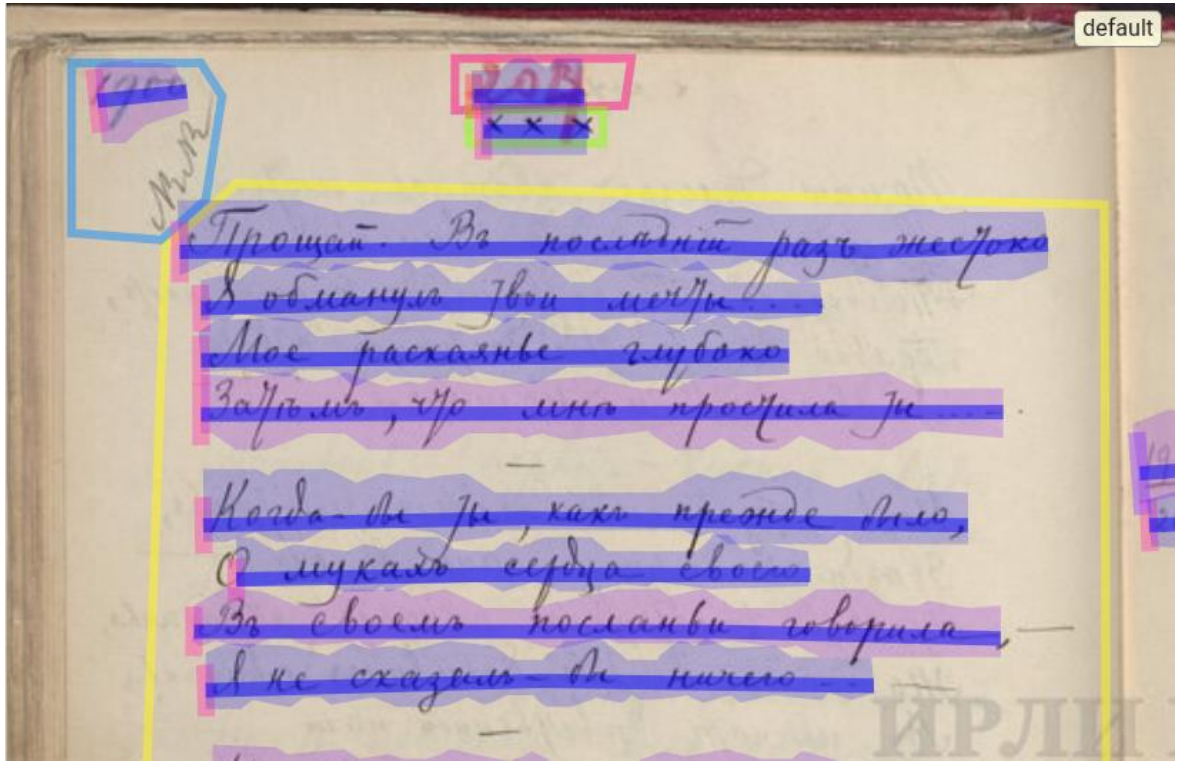
(c)

Line Masks

Fizaine FC, Bard P, Paindavoine M, Robin C, Bouyé E, Lefèvre R, Vinter A. Historical Text Line Segmentation Using Deep Learning Algorithms: Mask-RCNN against U-Net Networks. *Journal of Imaging*. 2024; 10(3):65. https://doi.org/10.3390/jimaging10030065

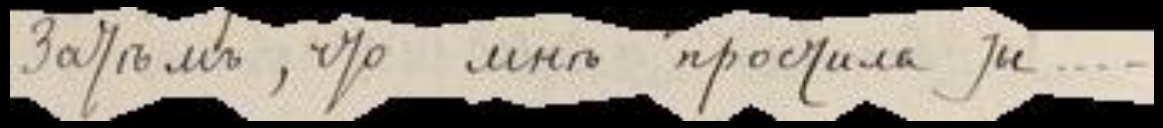
Look at the line masks

```
<TextBlock ID="eSc_dummyblock_">
            WIDTH="251"
            HEIGHT="30"></String>
  </TextLine>
  <TextLine ID="eSc_line_ce789720"
            TAGREFS="LT6"
            BASELINE="113 186 454 185"
            HPOS="113"
            VPOS="168"
            WIDTH="341"
            HEIGHT="33">
    <Shape><Polygon POINTS="113 186 113 194 118 194 128 201 128 201 129 2
    <String CONTENT="Затѣм, что мнѣ простила ты..."
            HPOS="113"
            VPOS="168"
            WIDTH="341"
            HEIGHT="33"></String>
  </TextLine>
  <TextLine ID="eSc_line_b430f30e"
            TAGREFS="LT6"
            BASELINE="573 178 651 174 704 168 792 162 820 158 910 154"
            HPOS="572"
            VPOS="141"
            WIDTH="338"
            HEIGHT="46">
    <Shape><Polygon POINTS="730 156 730 156 729 156 727 157 710 158 708 1
    <String CONTENT="Мнѣ Твой, о милая, чертогъ..."
```

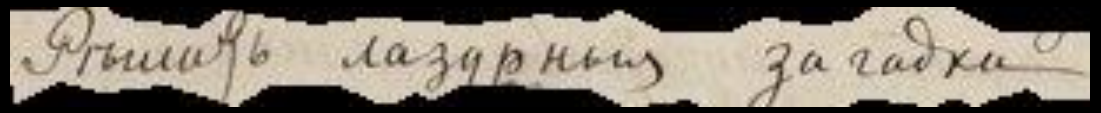


Manuscript from the Pushkin House Manuscripts Online, Institute of Russian Literature. Block, A. A. (1880-1921), poet. F. 654, Op. 1, Unit 1. P. 129.

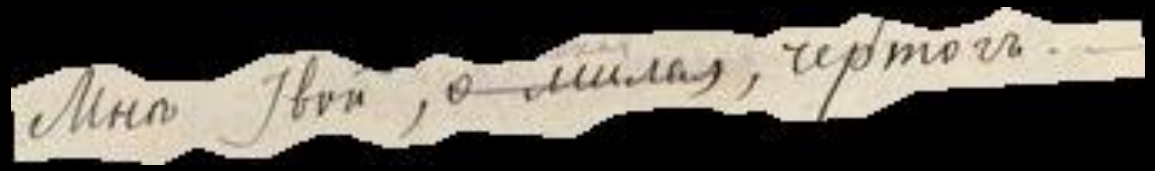
Cropped Line Masks



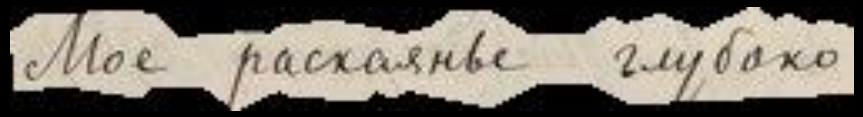
Затѣм, что мнѣ простила ты…



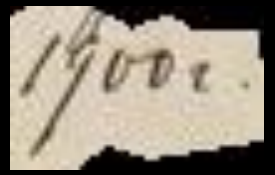
Рѣшить лазурныя загадки



Мнѣ Твой, о милая, чертогъ…



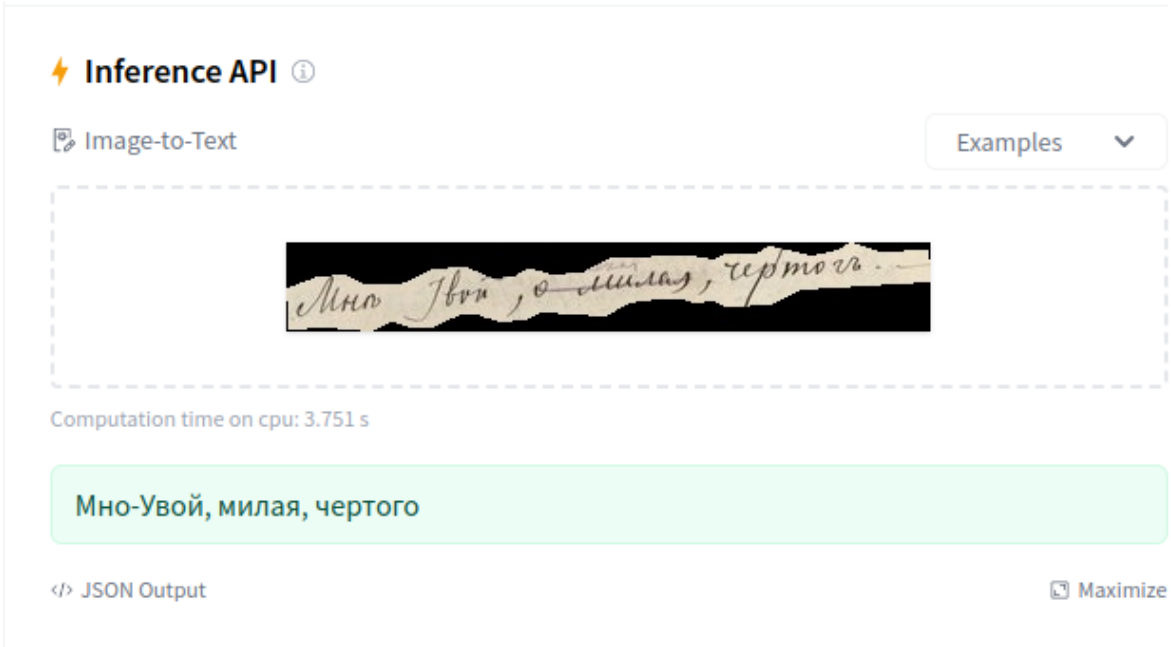
Мое раскаянье глубоко



1900 г.

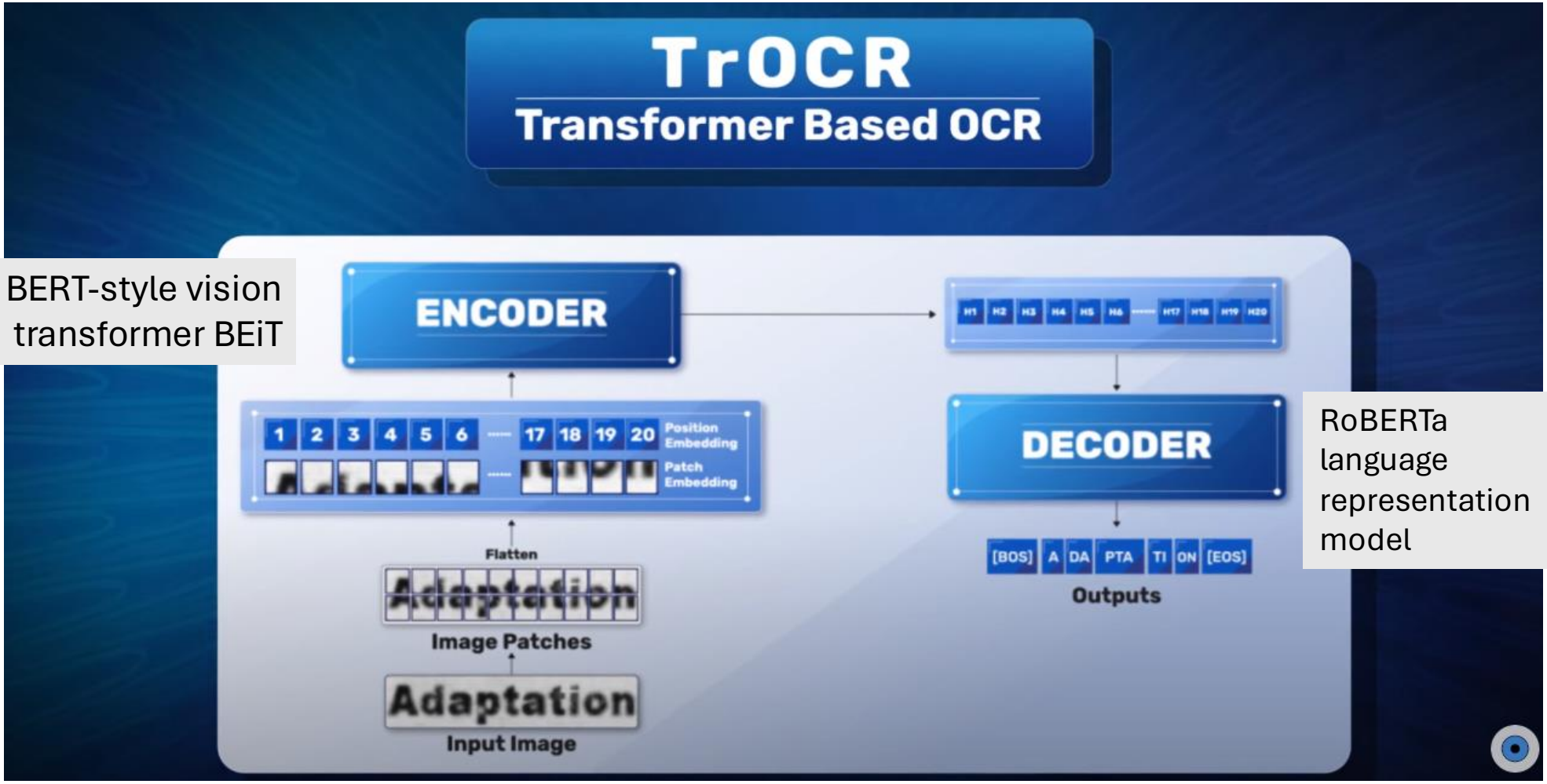
HuggingFace models (TrOCR-base-ru)

Мое раскаянье глубоко



МнѢ Твой, о милая, чертогъ...

TrOCR

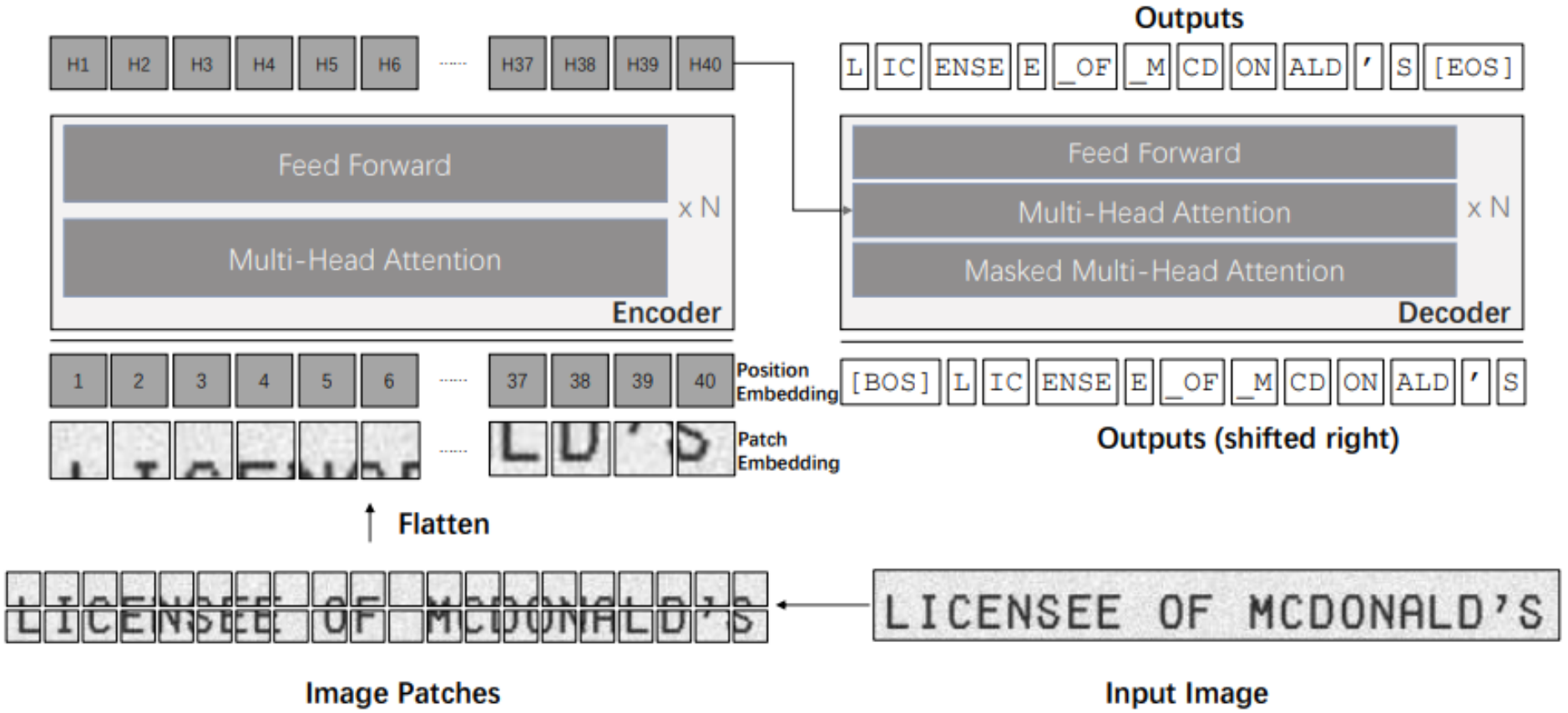


the BERT-style vision transformer BEiT

RoBERTa language representation model

Li, Minghao, Tengchao Lv, Lei Cui, Yijuan Lu, Dinei A. F. Florêncio, Cha Zhang, Zhoujun Li and Furu Wei. "TrOCR: Transformer-based Optical Character Recognition with Pre-trained Models." *AAAI Conference on Artificial Intelligence* (2021).

TrOCR



combines the BERT-style vision transformer BEiT with a RoBERTa language representation model.

Ströbel, Phillip Benjamin, Simon Clematide, Martin Volk, and Tobias Hodel. "Transformer-based HTR for historical documents." *arXiv preprint arXiv:2203.11008* (2022).

Reading order

Example of Baseline Segmentation *Singāsan Battīsī Naẓm* (1871)
Image: Universitätsbibliothek Heidelberg

Metric: CER (character error rate)

Leifert, Gundram, Christel Annemieke Romein, Achim Rabus, Phillip Benjamin Ströbel, Benjamin Kiessling, и Tobias Hödel. Evaluating State-of-the-art Handwritten Text Recognition (HTR) Engines; with Large Language Models (llms) for Historical Document Digitisation. Zenodo, 7 December 2023 г. https://doi.org/10.5281/zenodo.8102666

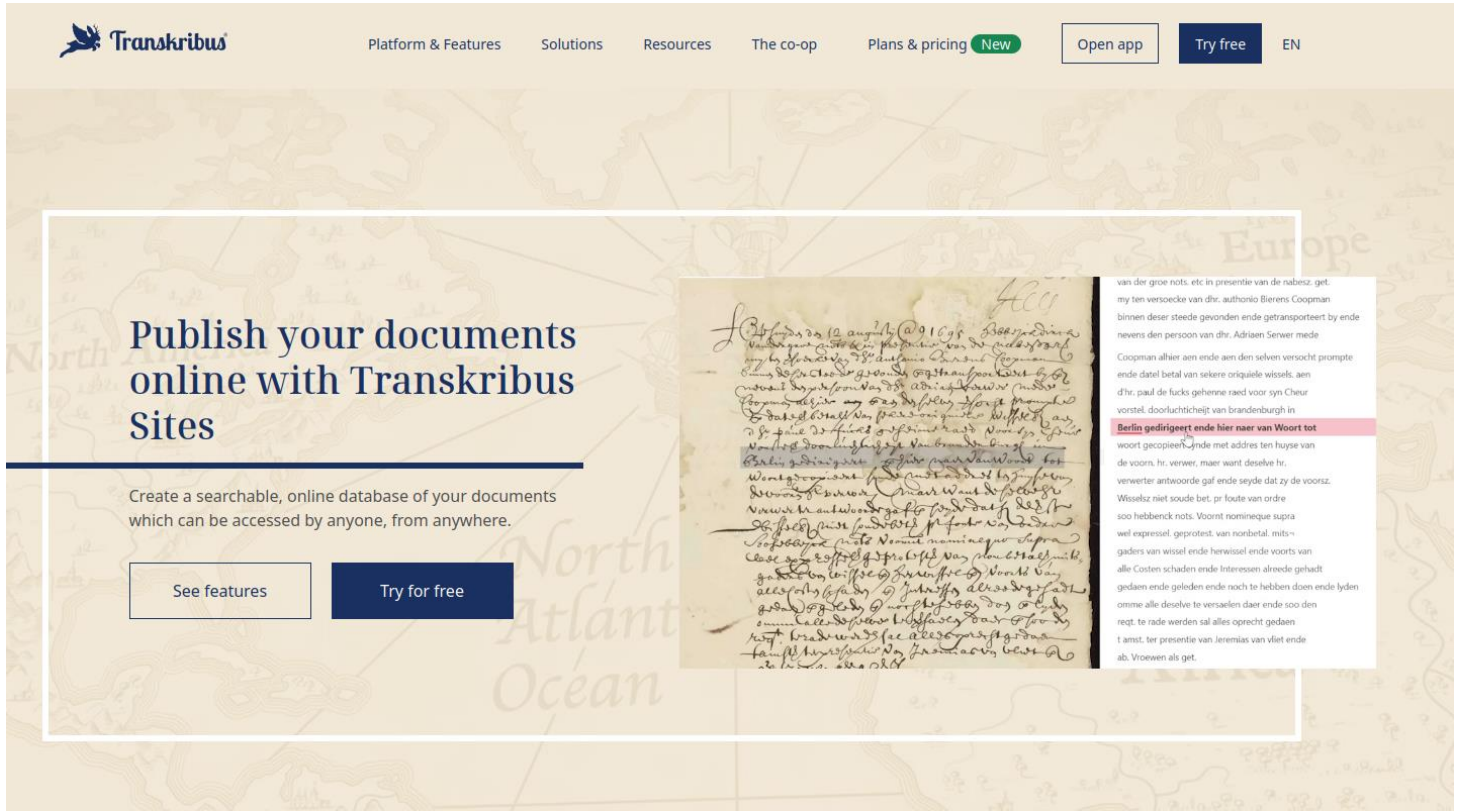
$$\mathrm{CER} = \frac{\sum_{i=1}^{n} \mathrm{dist}_c(\mathrm{pred}_i, \mathrm{true}_i)}{\sum_{i=1}^{n} \mathrm{len}_c(\mathrm{true}_i)}$$

CER - base

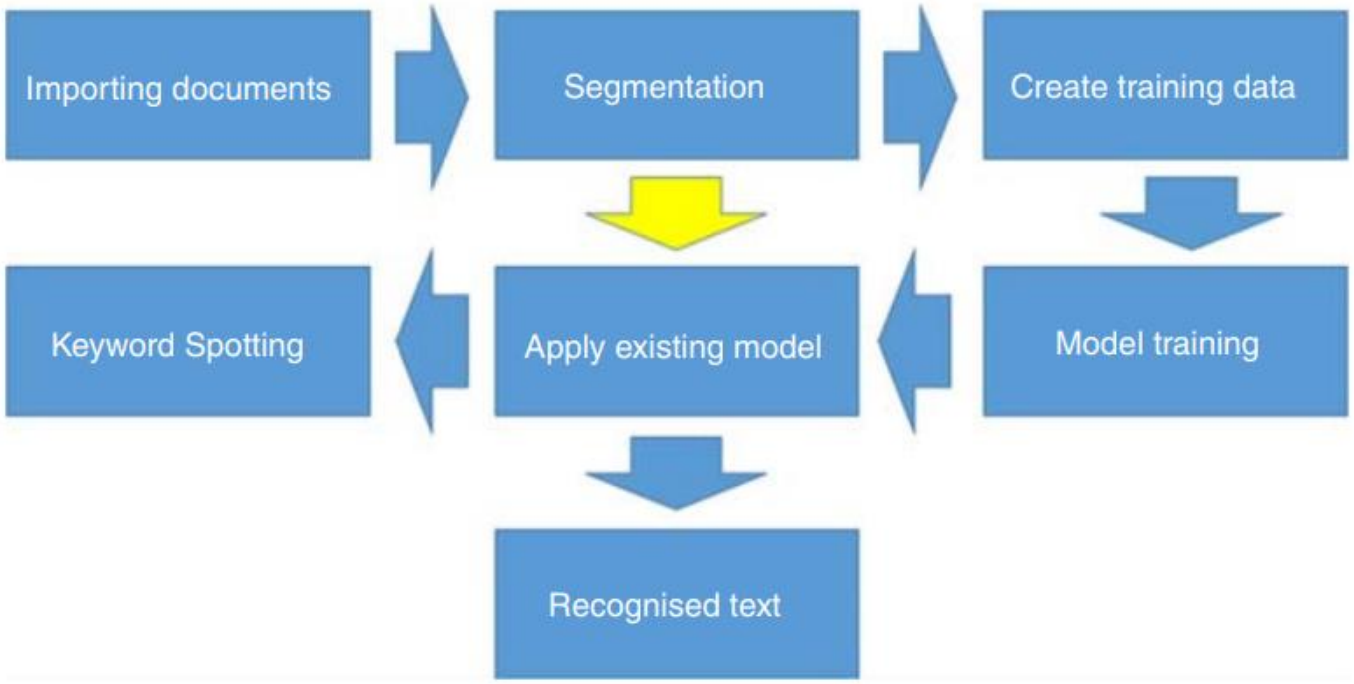
Dataset	Roman Type Print		Republic (7)		Glagolitic		Shorthand	
Engine	w/o LM	w/ LM	w/o LM	w/ LM	w/o LM	w/ LM	w/o LM	w/ LM
HTR+	**0.52%**		**0.62%**			5.65%		13.63%
Pylaia	2.02%		2.98%		6.32%	6.05%	11.91%	11.18%
IDA	2.74%	2.65%	2.69%	2.16%	4.83%	**4.54%**	9.53%	**9.29%**
TrOCR	4.1%		3.24%		4.67%		9.70%	

Transkribus

- **Archivists,**
- **humanities scholars,**
- **members of the public, all of whom are interested in the study and exploitation of historical documents**



Transkribus workflow



HTR+

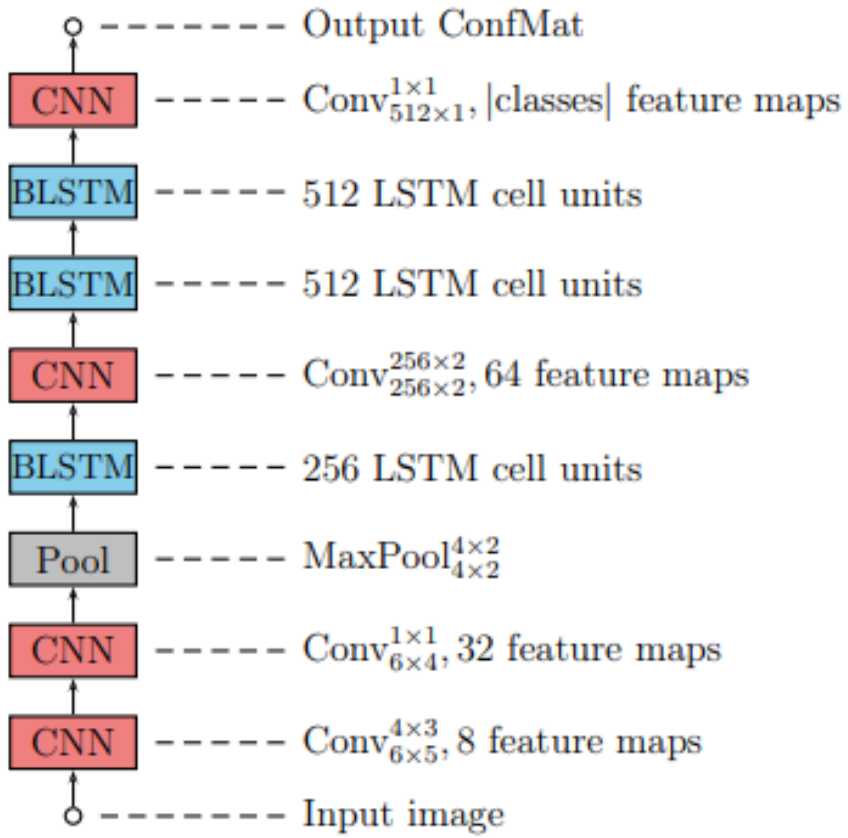
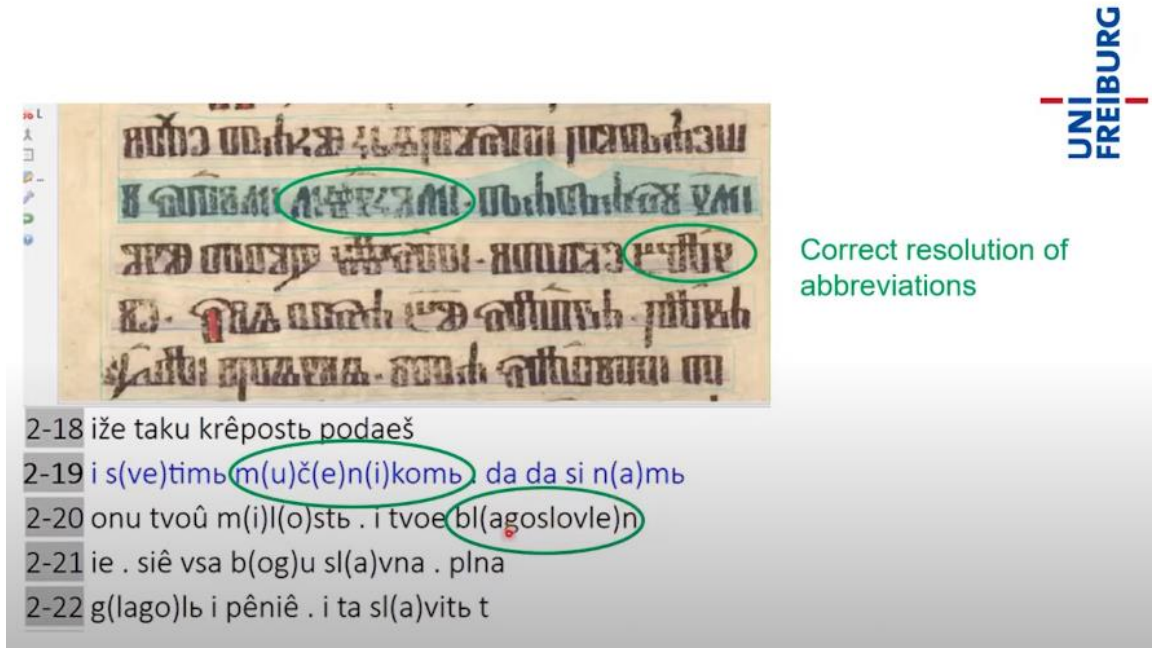


Figure 4: The proposed HTR model: A combination of various CNN and BLSTM blocks with one additional pooling layer. The subscript in $\text{Conv/MaxPool}_{\tilde{m}\times\tilde{n}}^{m\times n}$ describes the size of the kernel and the superscript defines the stride along both dimensions. Basically, |classes| is the number of characters present in the ConfMat.

Weidemann, M., Michael, J., Gruning, T., and Labahn, R. ¨ (2018). HTR Engine Based on NNs P2 Building Deep Architectures with TensorFlow. Technical report.

Transkribus "Smart" Models

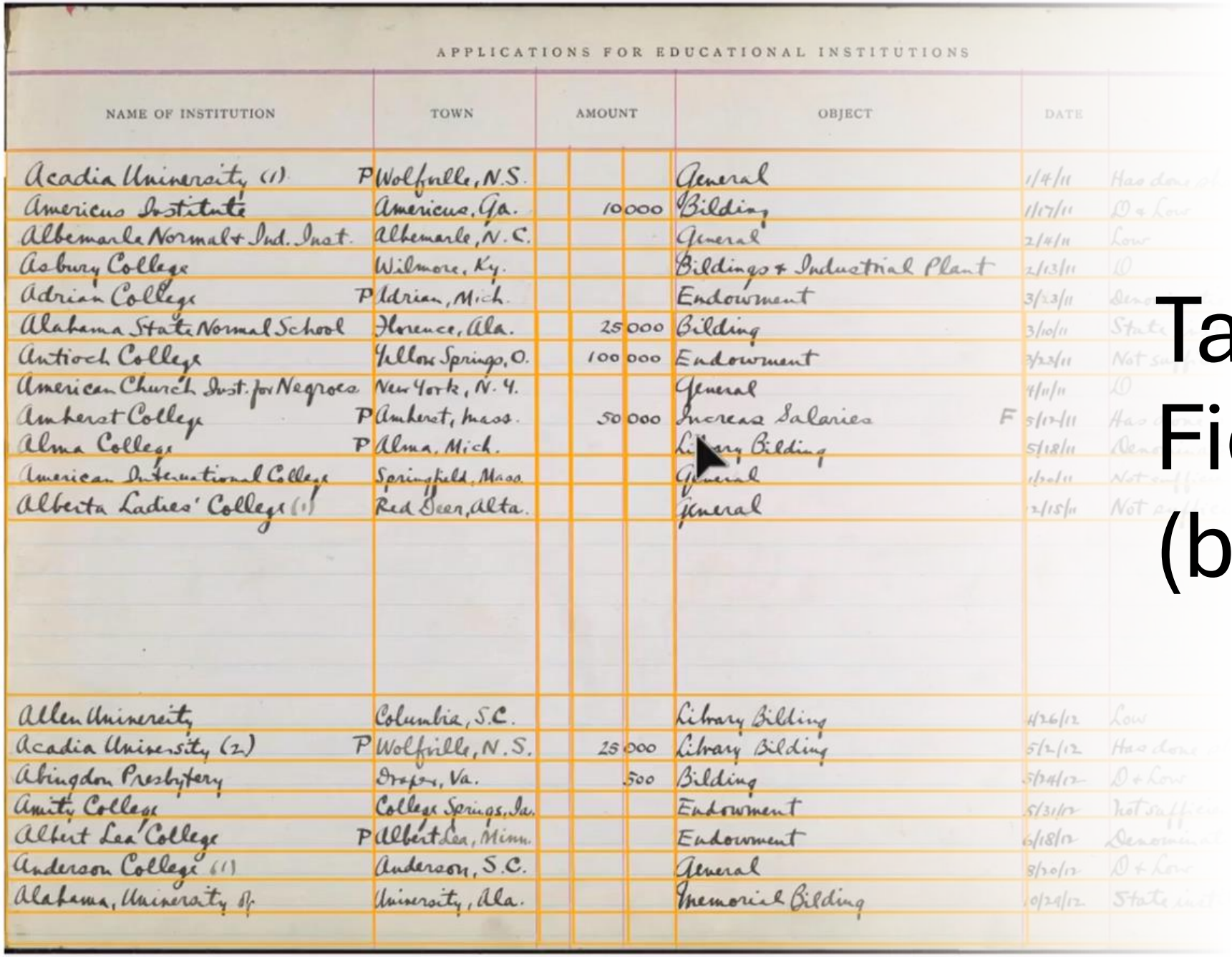
- Resolve abbreviations
- Bring superscripts down to the line
- Add missing punctuation or graphemes
- Modernize/Unify the orthography
- Transcribe into a different alphabet



Correct resolution of abbreviations

Smart models are for non-philological use!

- Quantitative studies
- Wider audience



APPLICATIONS FOR EDUCATIONAL INSTITUTIONS

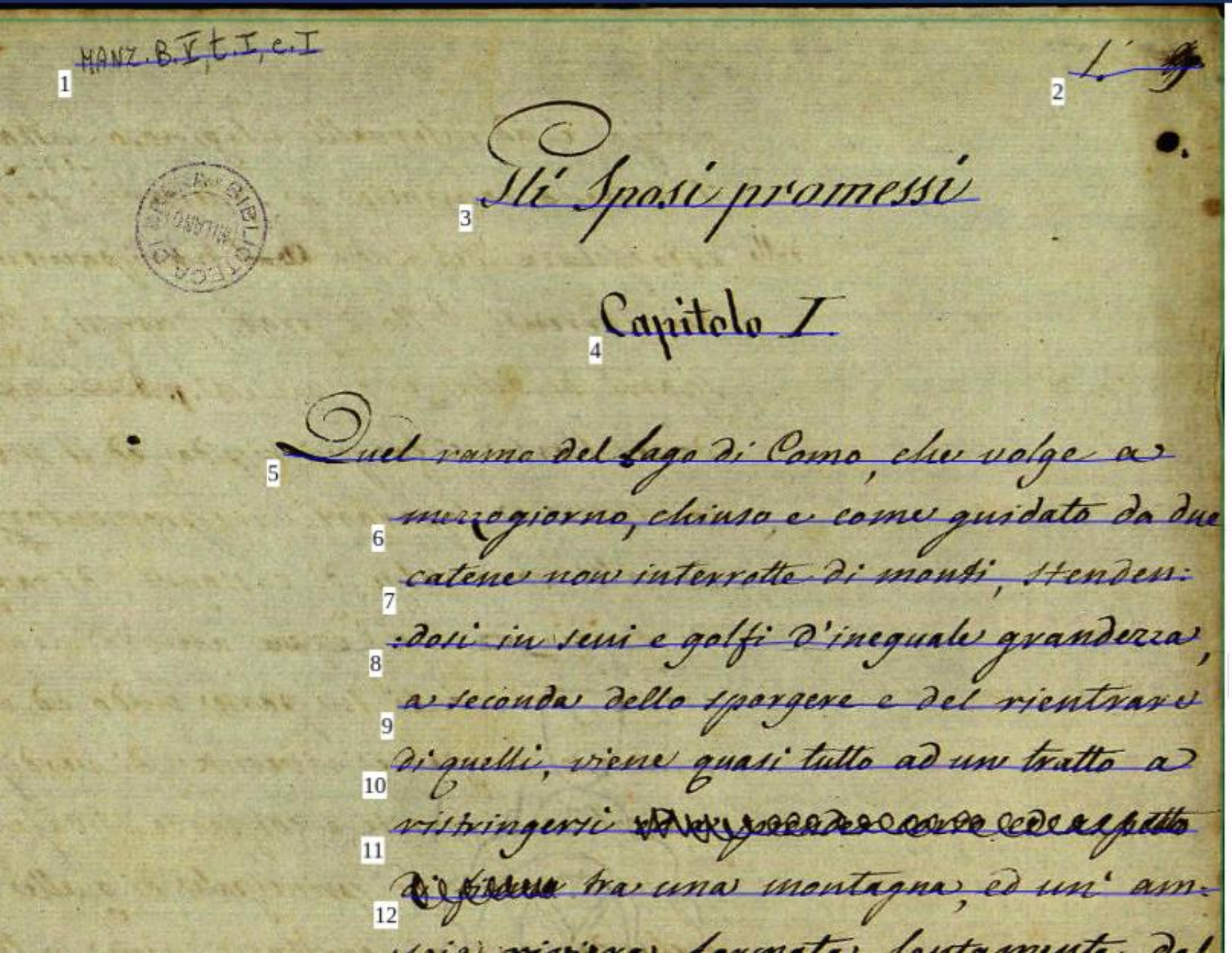
NAME OF INSTITUTION		TOWN	AMOUNT	OBJECT		DATE	
Acadia University (1)	P	Wolfville, N.S.		General		1/4/11	Has done
Americus Institute		Americus, Ga.	10000	Bilding		1/17/11	D + Low
Albemarle Normal + Ind. Inst.		Albemarle, N.C.		General		2/4/11	Low
Asbury College		Wilmore, Ky.		Bildings + Industrial Plant		2/13/11	D
Adrian College	P	Adrian, Mich.		Endowment		3/23/11	Den
Alabama State Normal School		Florence, Ala.	25000	Bilding		3/10/11	State
Antioch College		Yellow Springs, O.	100000	Endowment		3/23/11	Not suff
American Church Inst. for Negroes		New York, N.Y.		General		4/11/11	D
Amherst College	P	Amherst, mass.	50000	Increas Salaries	F	5/17/11	Has
Alma College	P	Alma, Mich.		Library Bilding		5/18/11	Den
American International College		Springfield, Mass.		General		1/24/11	Not suff
Alberta Ladies' College (1)		Red Deer, Alta.		General		12/15/11	Not
Allen University		Columbia, S.C.		Library Bilding		4/26/12	Low
Acadia University (2)	P	Wolfville, N.S.	25000	Library Bilding		5/2/12	Has done
Abingdon Presbytery		Draper, Va.	500	Bilding		5/24/12	D + Low
Amity College		College Springs, Ia.		Endowment		5/31/12	Not suffic
Albert Lea College	P	Albert Lea, Minn.		Endowment		6/18/12	Denomin
Anderson College (1)		Anderson, S.C.		General		8/20/12	D + Low
Alabama, University of		University, Ala.		Memorial Bilding		10/24/12	State inst

Table and Field models (beta)

Data is all you need

- For the text written in one hand:
  - 15,000 transcribed words (≈ 75 pages).
- For the printed text:
  - 5,000 transcribed words (≈ 25 pages).

	Character Error Rate	Training Data	Approximate time to create the training data
Printed text	0,5-2%	~ 5.000 words / 25 pages	15 hours
Single hand - simple writing	2-4%	+10.000 words / 50 pages	25 hours
Several hands – all seen during training	4-6%	several 10.000 words training data / several 50 pages	25 hours x each hand
Many hands from same period and region – not all seen during training	6-8%	+100.000 words / more than 500 pages	~ 250 hours

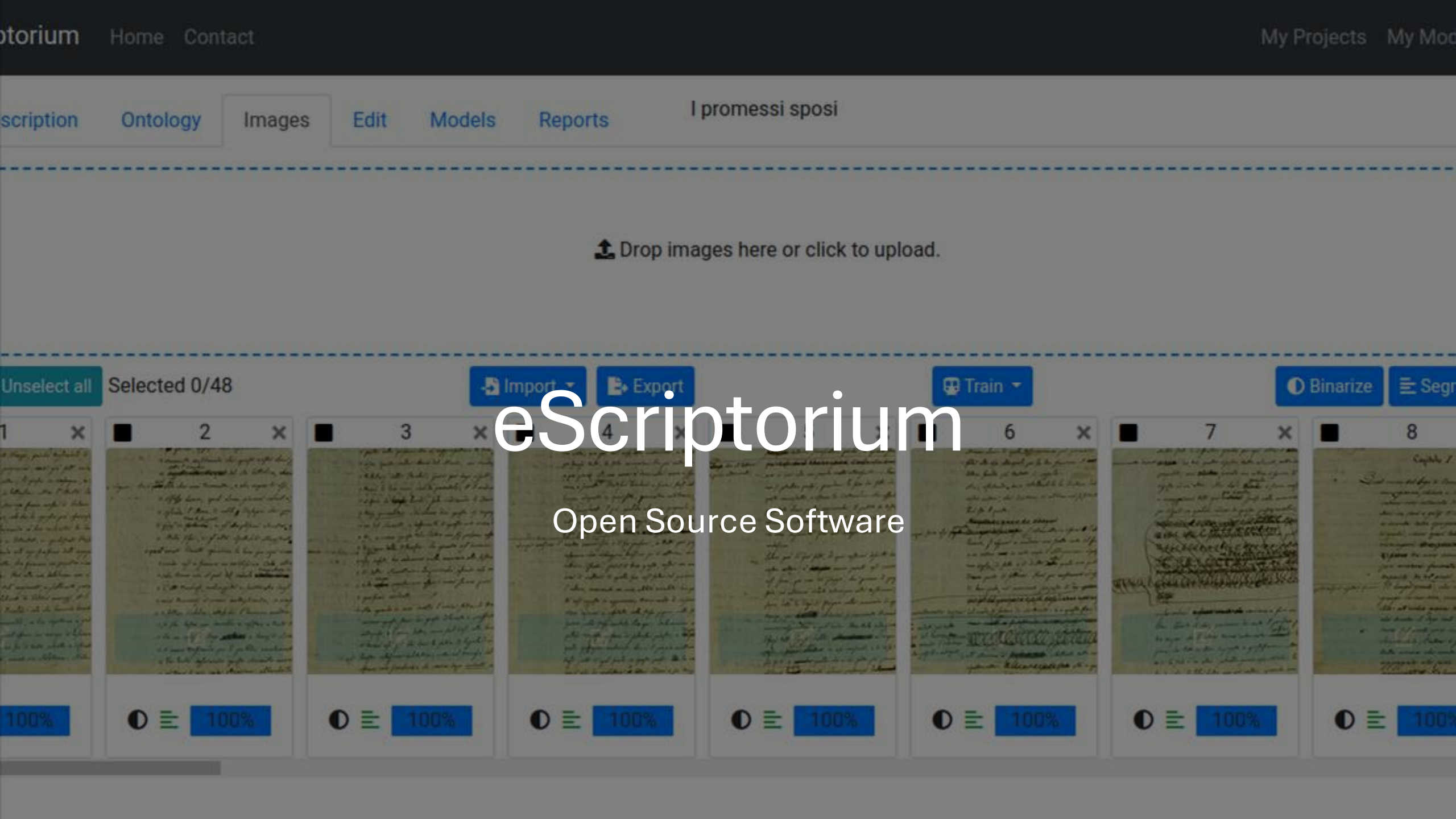


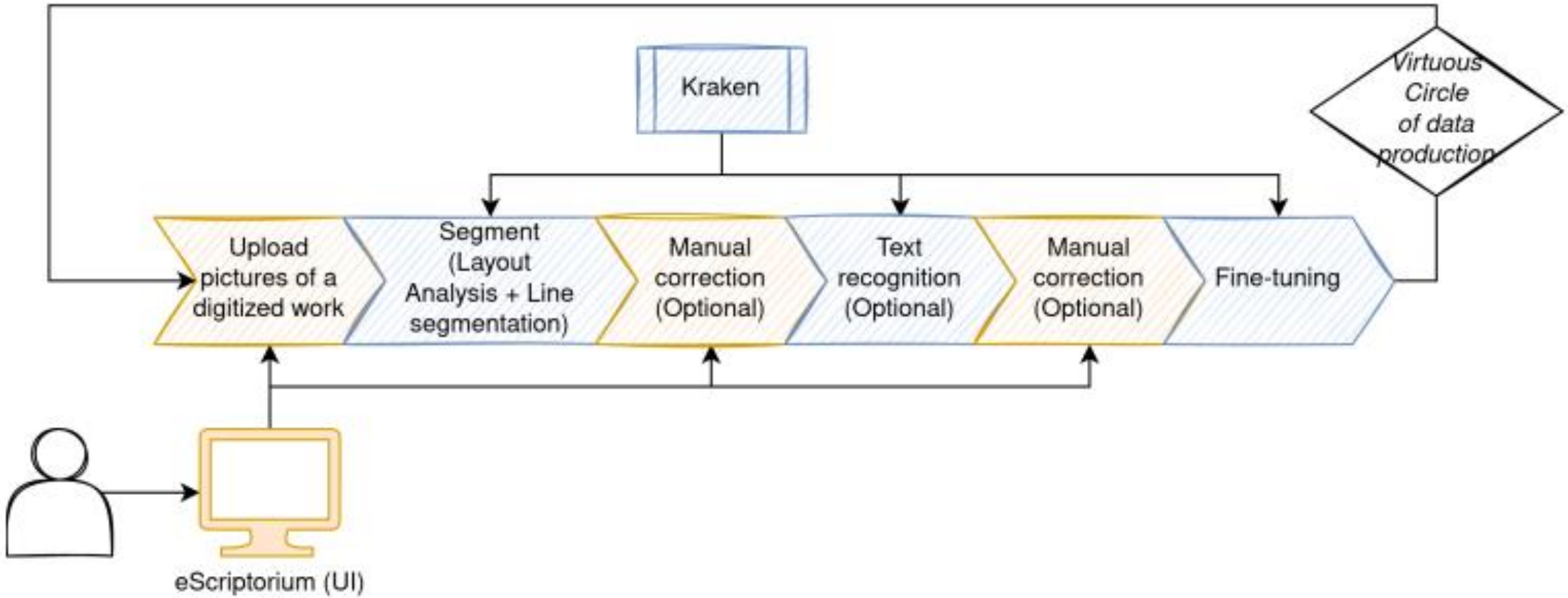
Gli Sposi promessi

Capitolo I

Quel ramo del lago di Como, che volge a mezzogiorno, chiuso e come guidato da due catene non interrotte di monti, stendendosi in seni e golfi d'ineguale grandezza, a seconda dello sporgere e del rientrare di quelli, viene quasi tutto ad un tratto a ristringersi [illegible] tra una montagna, ed un'ame-

1 Nurte
2 14
3 N Spose promesse
4 Capitolo
5 Quel ramo del lago di Como, che volge a
6 mezzo giorno, chiuso, e come guidato da due
7 catene non interrotte di monti, stenden¬
8 dosi in seni e golfi d'incquale grandezza
9 a seconda dello sporgere e del rientrare
10 di quelli, viene quasi tutto ad un tratto a
11 ristringeri
12 a una montagna ed un am¬
13 sia riviera formata sentamente dal
14 Deposito di tre grossi, e vicini torrenti,
15 Il lungo ponte, che in quel luogo con¬
16 Cepende qui corpo et aspetto di fine
17 giunge le due rive, vende ancor più senti¬
18 bile all'occhio questa trasformazione e par¬
19 che divida il lago dall'Alda. A diritta la


Home
Contact
My Projects
Ontology
Images
Edit
Models
Reports
I promessi sposi
Drop images here or click to upload.
Unselect all
Selected 0/48
Import
Export
Train
Binarize
eScriptorium
Open Source Software

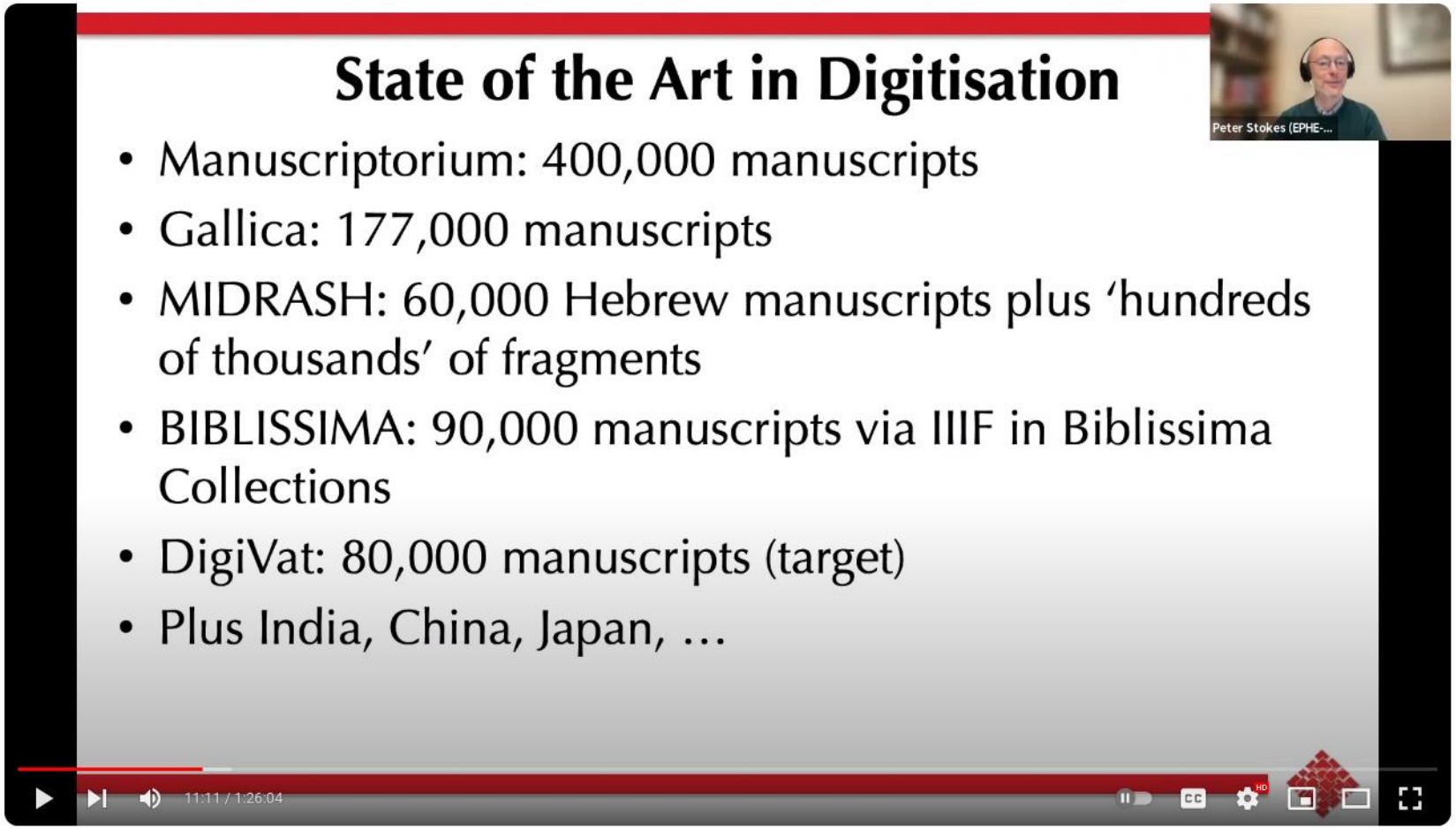


Clérice, Thibault. (2022). You Actually Look Twice At it (YALTAi): using an object detection approach instead of region segmentation within the Kraken engine. 10.48550/arXiv.2207.11230.

Transkribus vs. eScriptorium

• **Service** for a broad audience, including researchers, historians, and archival enthusiasts.	**Software** built and supported by the community specifically towards digital humanities specialists, scholars, and researchers
• Online usage	Requires local installation, which may be challenging
• Commercial product with a free credits	Open Source
• Unable to export models	Able to import and export models
• Exports to Txt, Page xml, ALTO, PDF, TEI, docx	Exports to Txt, Page xml, and ALTO formats
• HTR+, PyLaia, TrOCR models	Kraken models

Sharing the data



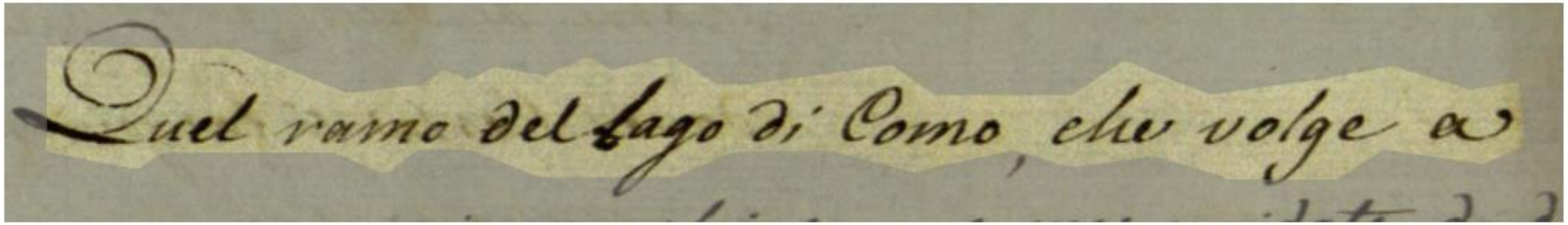
How to Transcribe a Million Manuscripts with eScriptorium

Kraken models for eScriptorium

- https://zenodo.org/communities/ocr_models/
- https://zenodo.org/communities/scriptnet

- Peter Stokes, Benjamin Kiessling. Sharing Data for Handwritten Text Recognition (HTR). Digital Humanities in Practice, In press. ffhal-04444641f
- **How to Transcribe a Million Manuscripts with eScriptorium**

Line #4



Duel ramo del Lago di Como ele volge a

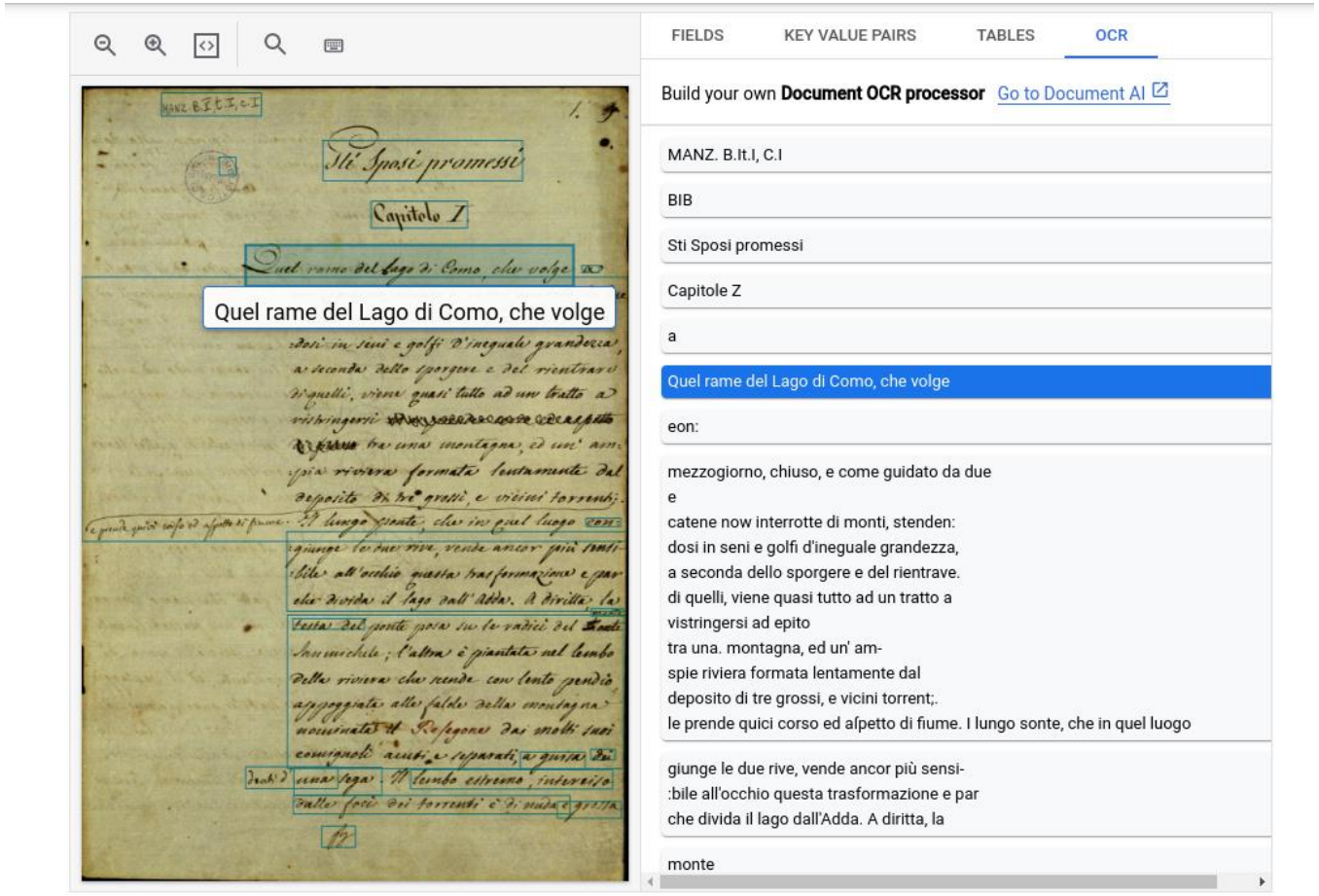
by admin (kraken:Tridis_Medieval_EarlyModern) on Mon Apr 22 2024 11:26:11 GMT+0100

-Toggle transcription comparison

?

DPuel rvamo del Llago di Como chele volge a	kraken:ManuMcFondue
DQueld ramoe deul Lagoávo dio CoGmo eoleu volgſe a&	kraken:Gallicorpora+_best
Duel ramo del Lago di Como ele volge a	kraken:Tridis_Medieval_EarlyModern (current)

Google Document AI



You

Can you read the text from this image?

ChatGPT

The text in the image reads "Рождения." and "Крещения.", which in English translates to "Births." and "Baptisms." respectively.

You

And this?

ChatGPT

The text from the image appears to say "е Мая", which likely corresponds to "of May" in English, considering the context and the style of the text, which suggests an older form of cursive.