Cartography M.Sc.

Toponym Extraction from Selected Historic Topographic Maps Using Deep Learning

Gongmingyue Tang

Supervisors:
Dr. Nikolas Prechtel
Prof. Dr. Markus Wacker

Cartography M.Sc.

Toponym Extraction from Selected Historic Topographic Maps Using Deep Learning

Gongmingyue Tang

Supervisors:
Dr. Nikolas Prechtel
Prof. Dr. Markus Wacker

Motivation

RQs

Literature Review

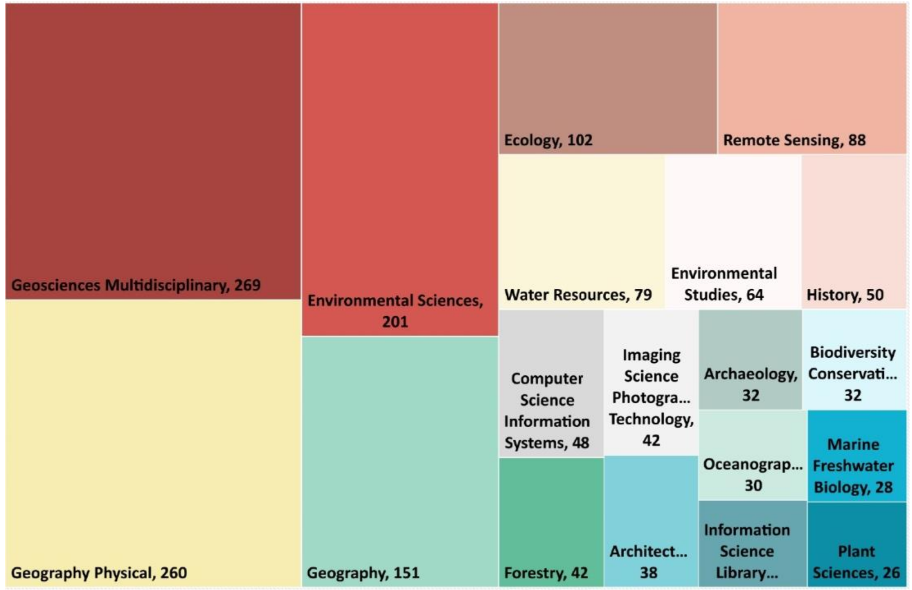
Material

Methods

Results

Conclusion

- Why historical maps?



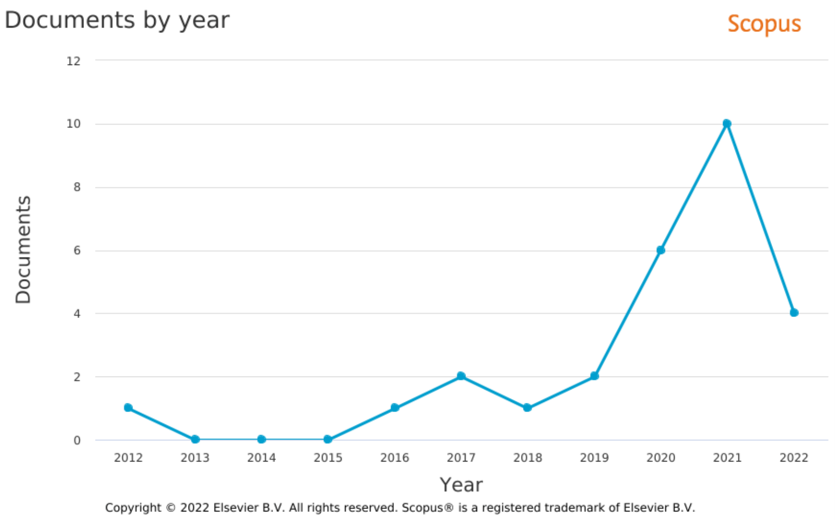
Analysis Report graphic derived from Clarivate *Web of Science*,

- Why historical maps?

- Why toponym?
  - Identify the location
  - Queryable
  - Understanding of geographical features
  - Linguistics and culture

(Jordan, 2009)

- Why historical maps?
- Why toponym?
- Why deep learning?
  - Scanned historical maps -> structured data



Research Objectives and Questions

Automated transfer to structured, tagged Geodata:

- Separation of text and graphic
  1. What are the existing deep learning pipelines for text extraction?
  2. How can synthetic training data help optimize the model training process?
  3. How well can toponyms be separated from the background using deep learning?
- Text recognition
  1. What are the existing text recognition models?
  2. How well is the performance of the adapted text recognizer?
- Evaluation
  1. How is the overall performance of the deep learning pipeline in toponym detection and recognition?

RQ1.1 & RQ2.1

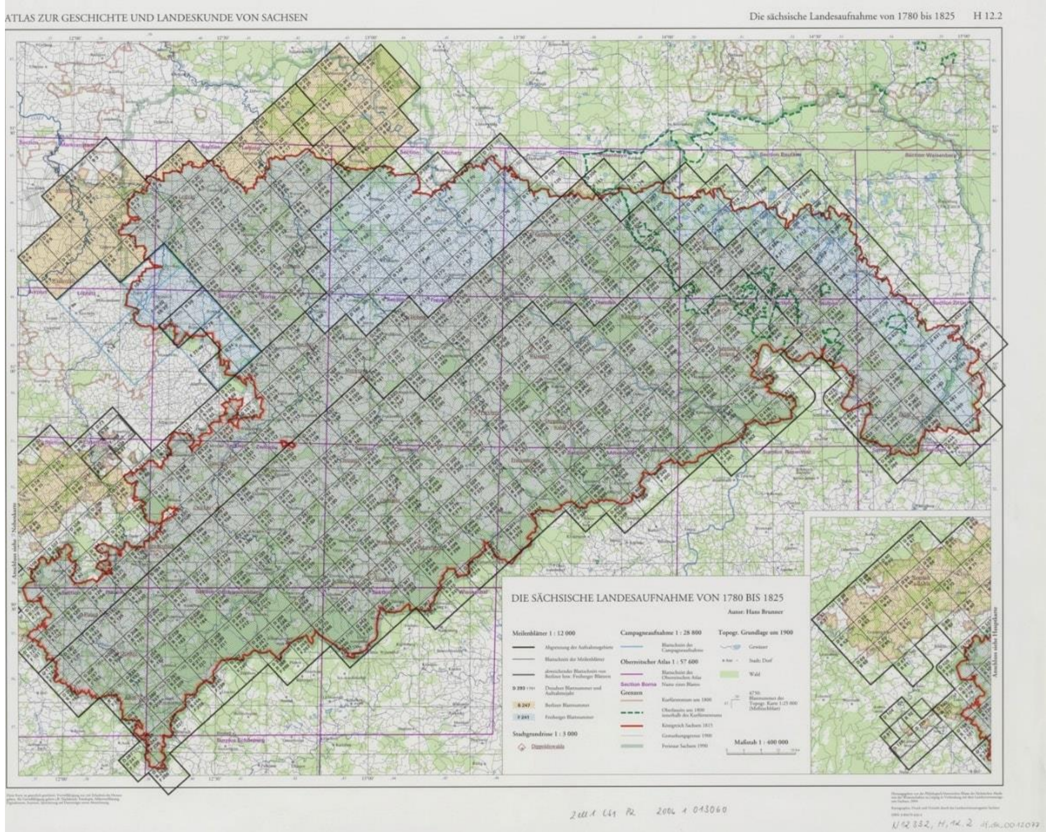
<u>Existing deep learning pipelines for text detection:</u>

- Proposal-based:
  CTPN (Tian et al., 2016), Seglink (Shi et al., 2017)…
- Segmentation-based:
  TextSnake (Long et al., 2018), DBNet (Liao et al., 2020)…
- Hybrid-based:
  FCE (Zhu et al., 2021)

<u>Existing deep learning pipelines for text recognition:</u>

- CRNN architecture
  HTR+ (Michael et al., 2020), CRNN+STN (Shi et al., 2016)
- Attention-based methods: ABiNet (Fang et al., 2021)

Selected Topographic Map



Overview Map of the Saxon Mile Sheets (Brunner, 2005)

- From Saxonian land survey between 1780-1826
- Scale: 1:12000
- 445 tiles
- Berlin Copies
- Accurate in geometry
- Very detailed

(Stams & Stams, 1981)

Toponym Hierarchy and Placement

Cities and regions

Towns, large villages, districts and rivers

Buildings, roads, landscapes

Motivation

RQs

Literature Review

Material

Methods

Results

Conclusion

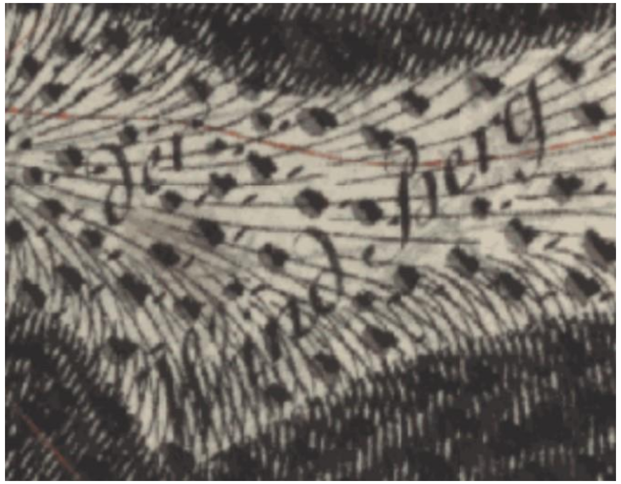
Toponym Hierarchy and Placement

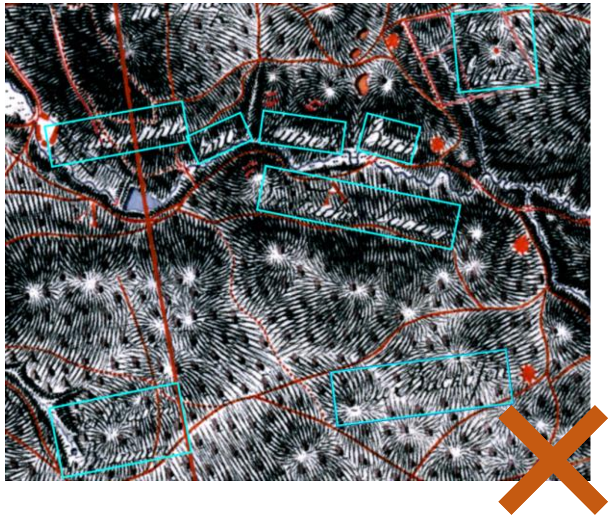
Cities and regions

Towns, large villages, districts and rivers

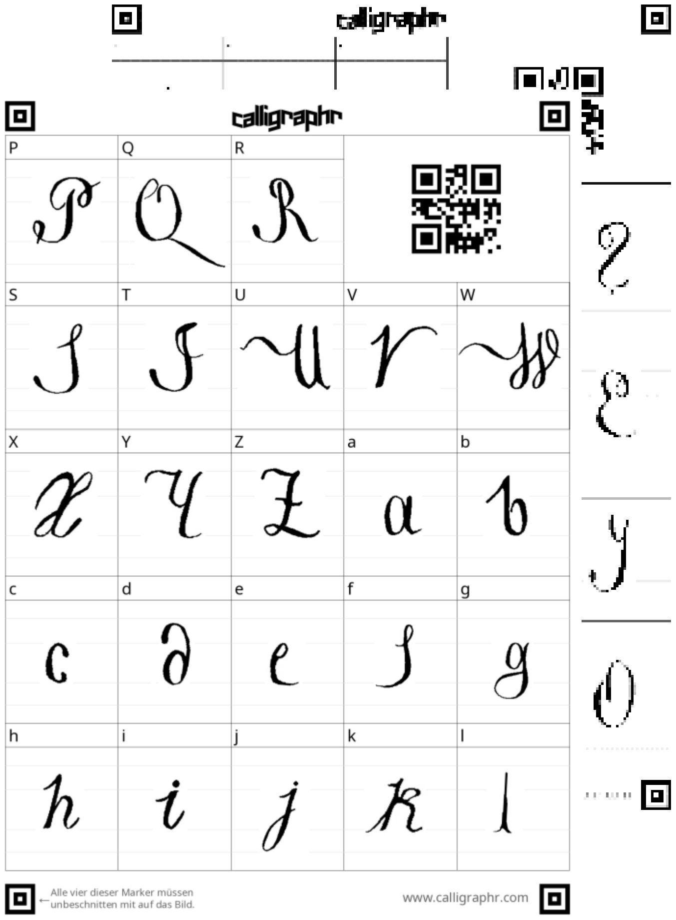
Buildings, roads, landscapes

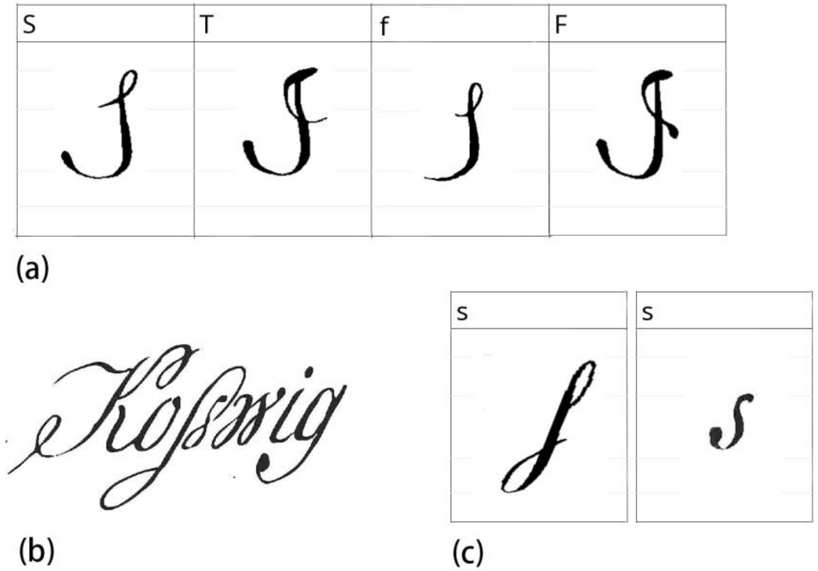
Toponym Hierarchy and Placement





Font


calligraphr
P Q R
S T U V W
X Y Z a b
c d e f g
h i j k l
Alle vier dieser Marker müssen unbeschnitten mit auf das Bild.
www.calligraphr.com


S T f F
(a)
(b)
s s
(c)

Motivation

RQs

Literature Review

Material

Methods

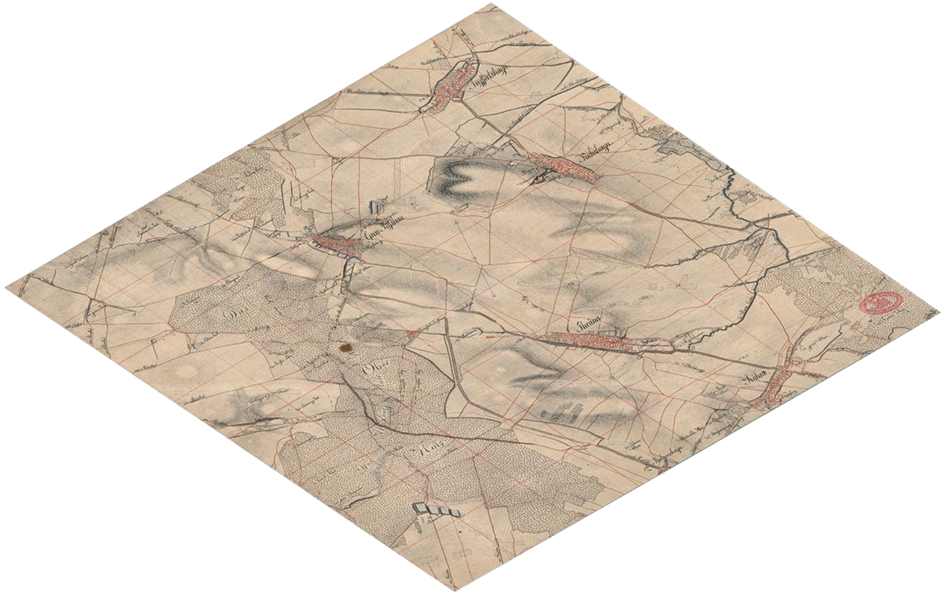
Results

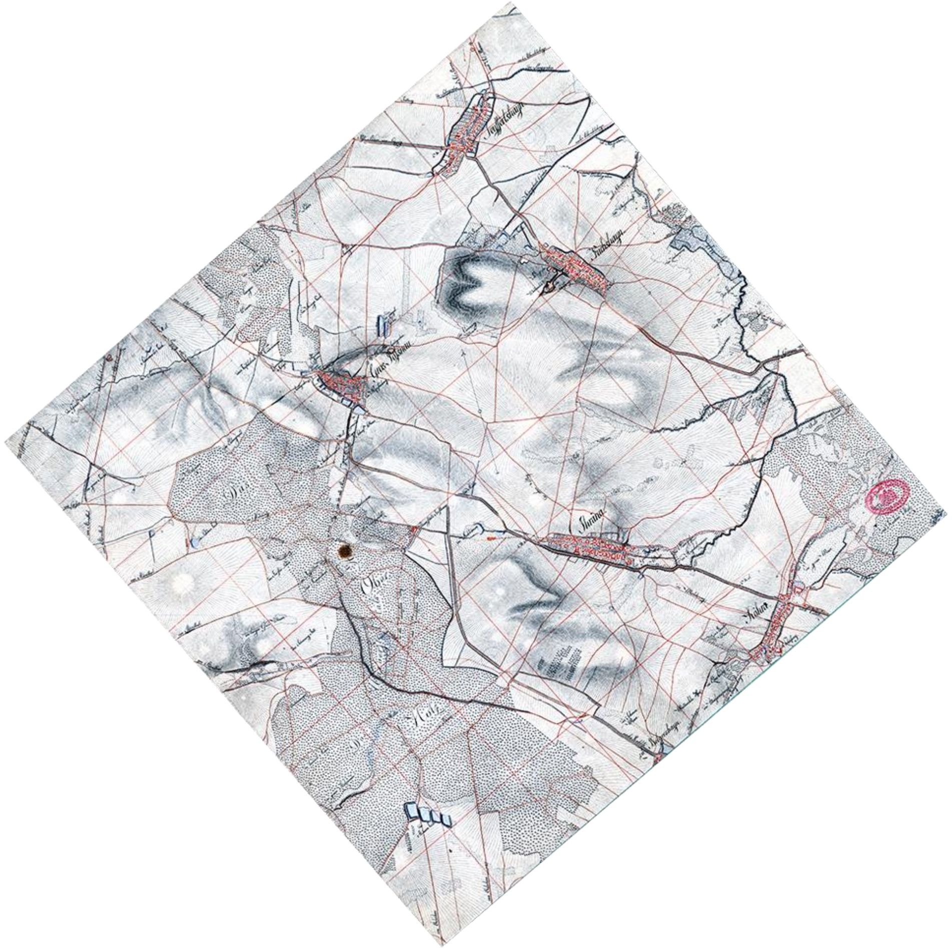
Conclusion

Pre-processing

1. Color Homogenization

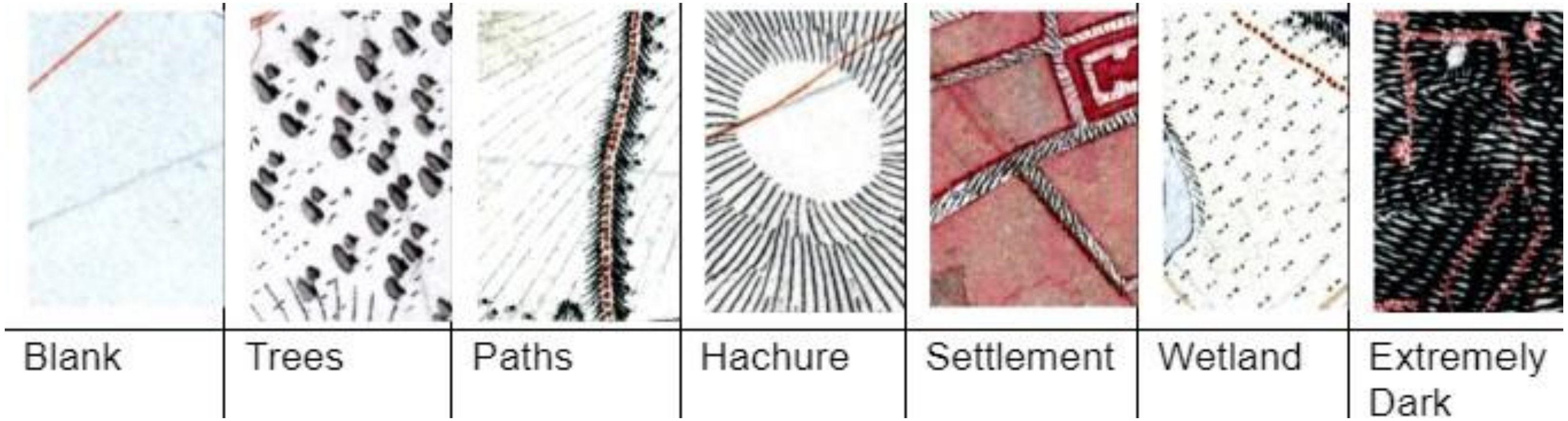
2. Re-projection



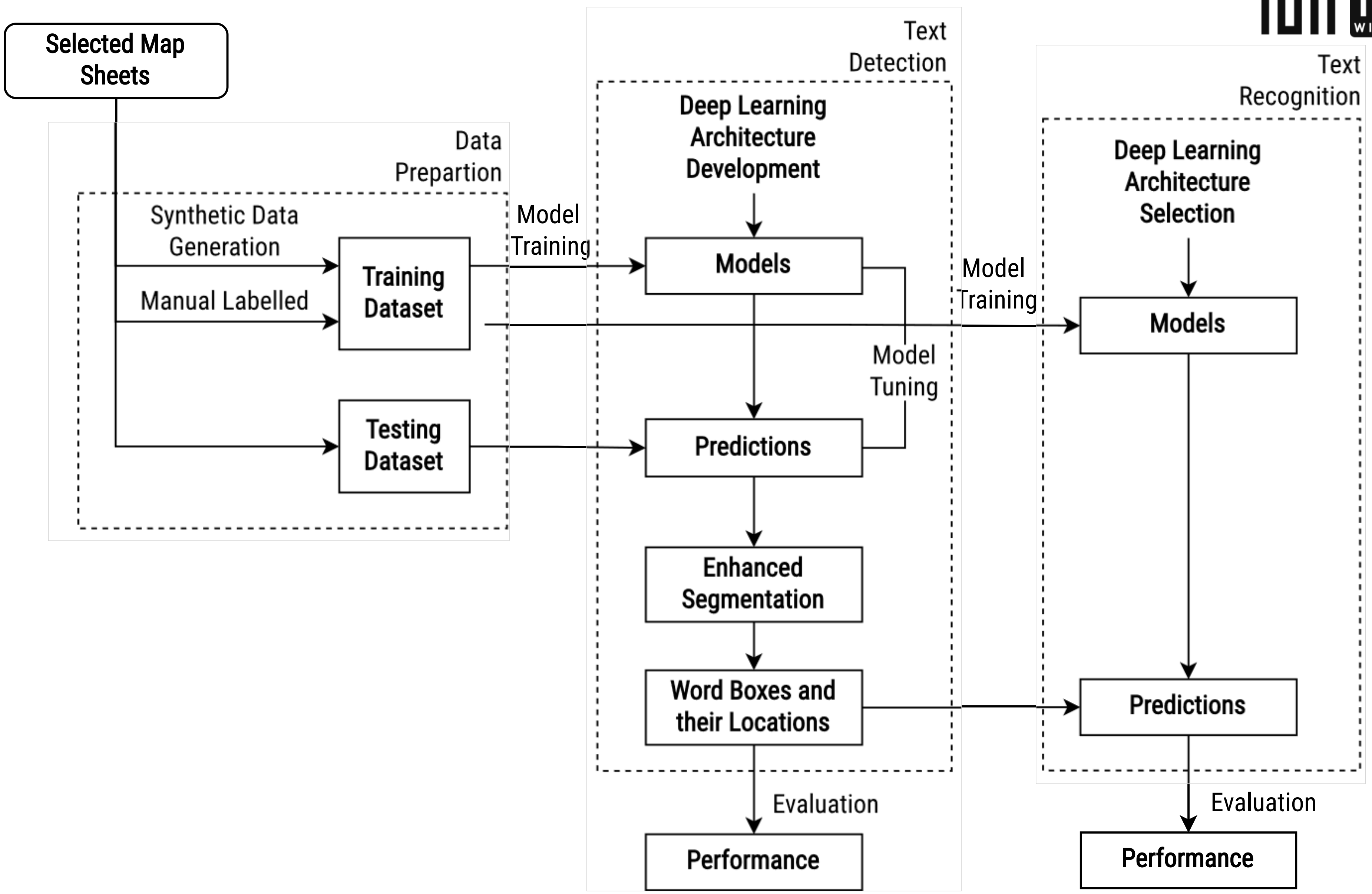


Pre-processing

1. Color Homogenization
2. Re-projection
3. Patches Extraction

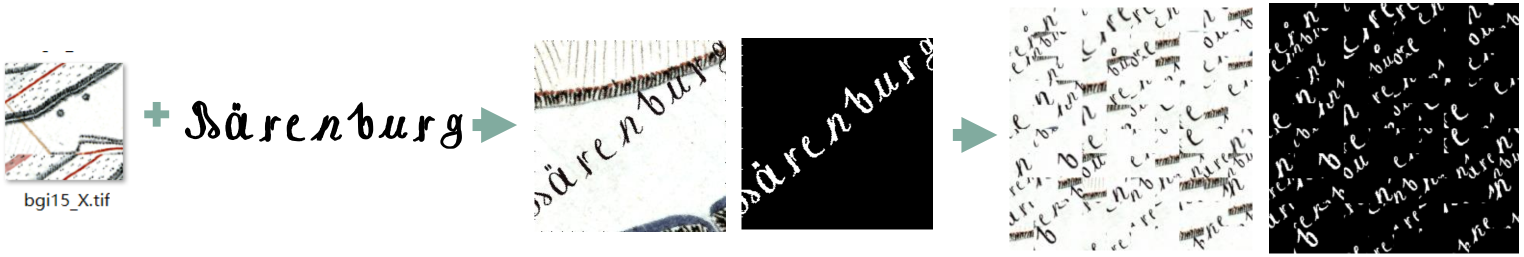


- Motivation
- RQs
- Literature Review
- Material
- **Methods**
- Results
- Conclusion

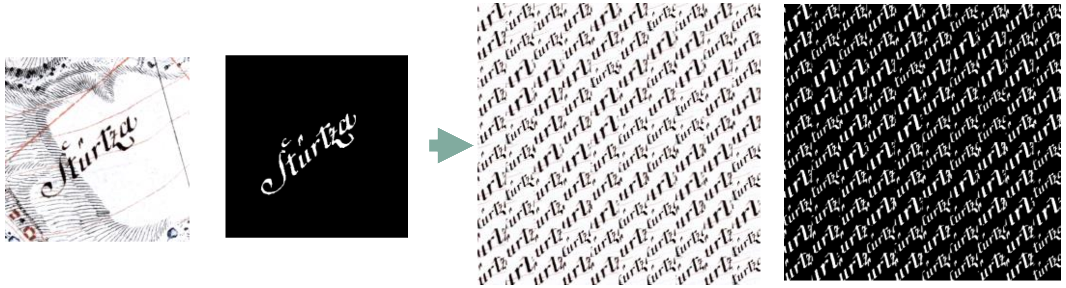


Motivation
RQs
Literature Review
Material
Methods
Results
Conclusion

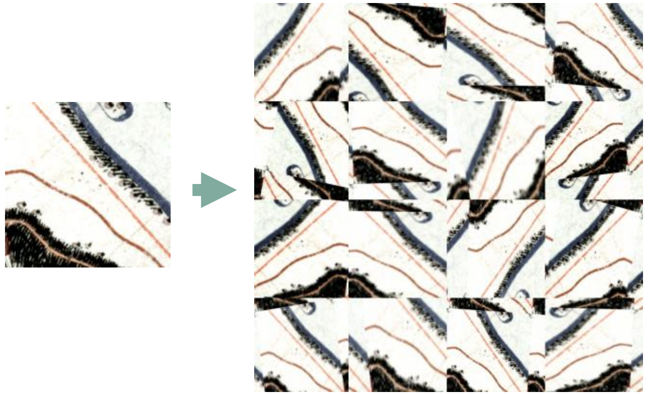
Synthetic training samples:



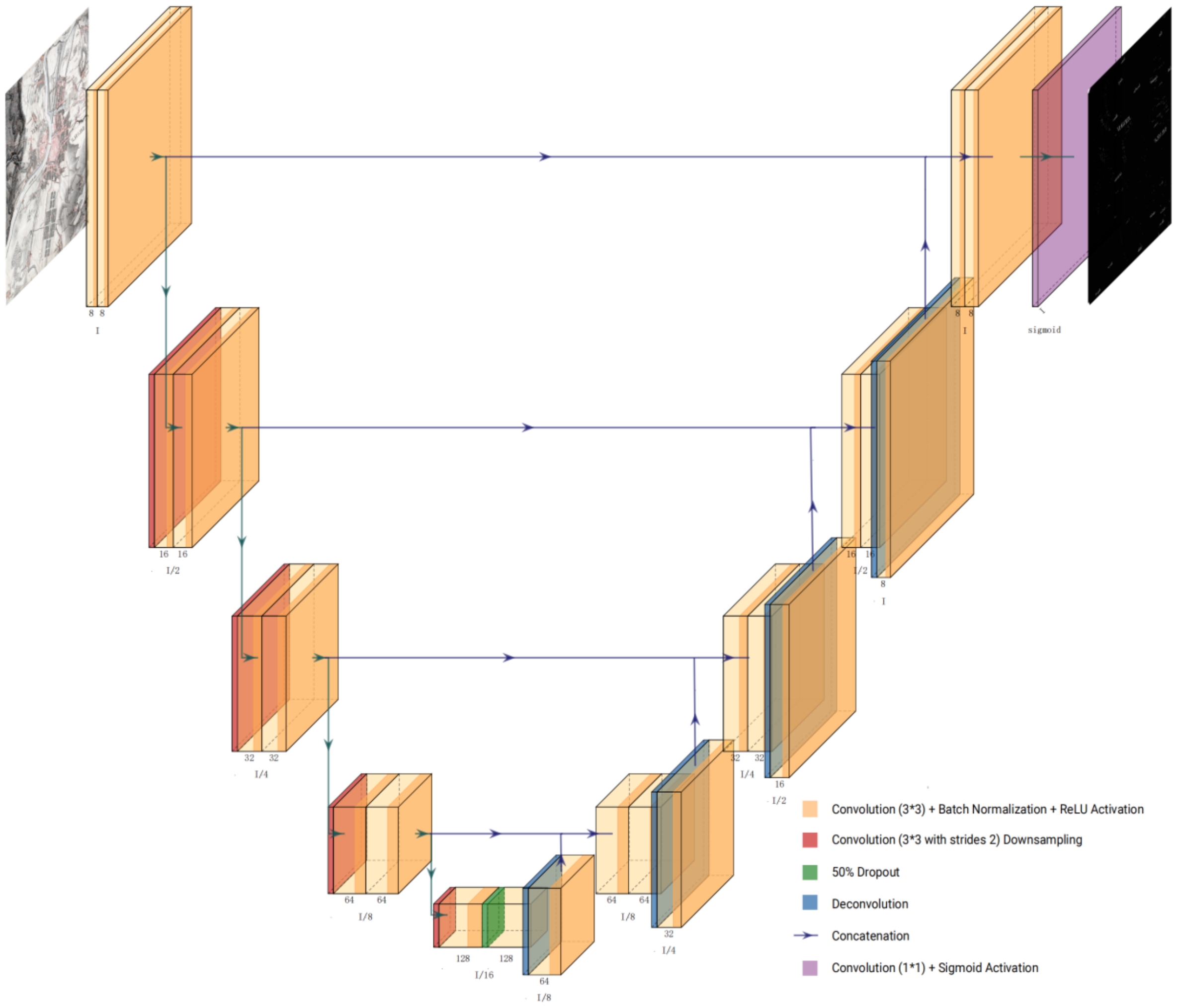
Manually labelled training samples:



Selected true negative patches:



U-Net Model


Convolution (3*3) + Batch Normalization + ReLU Activation
Convolution (3*3 with strides 2) Downsampling
50% Dropout
Deconvolution
Concatenation
Convolution (1*1) + Sigmoid Activation

Model Tuning

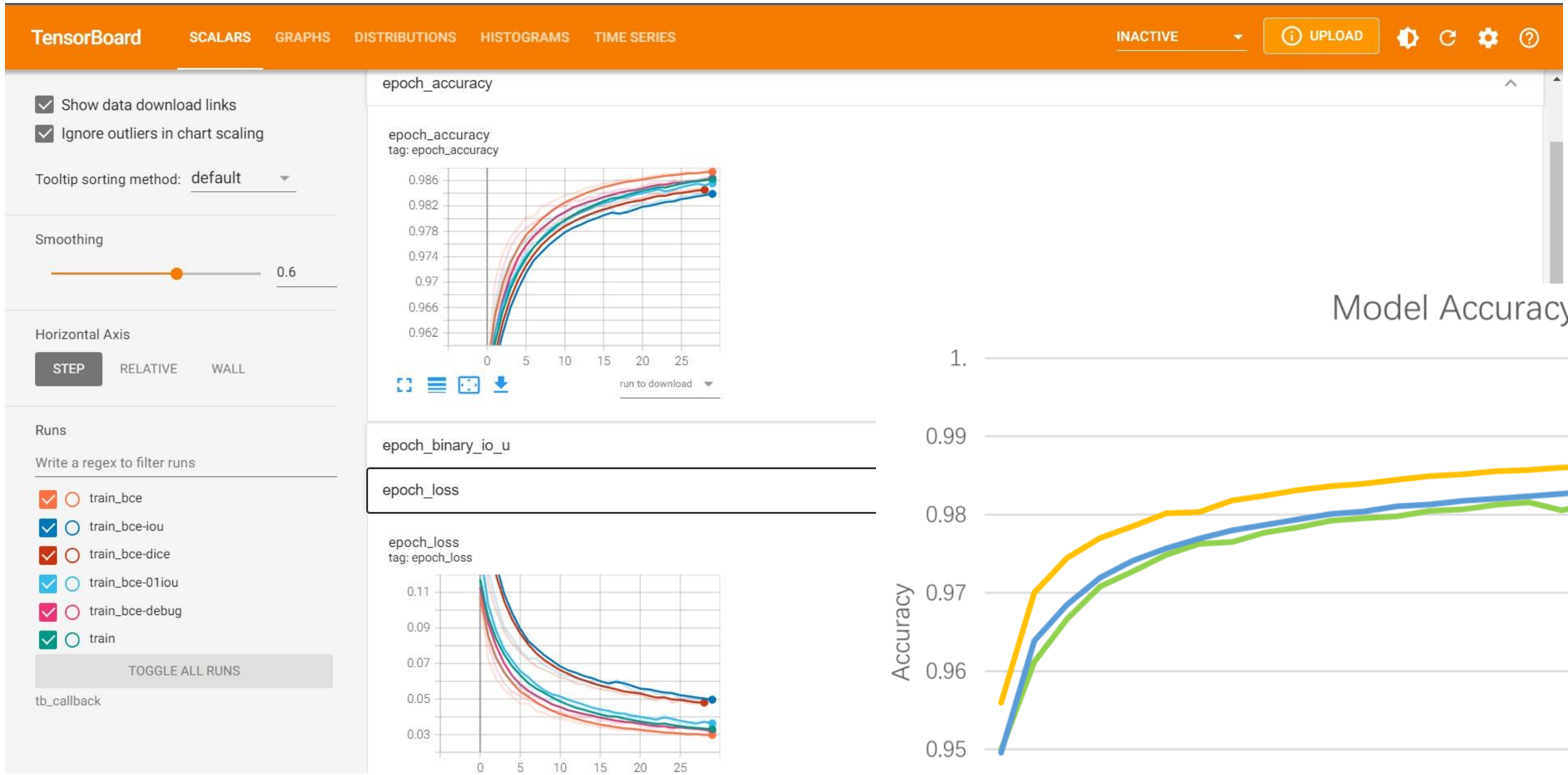
Experiment	1	2	3
Learning rate	1.00E-03	2.00E-03	1.00E-02
Batch size	5	5	5
Loss function	Binary Cross Entropy (BCE)	BCE	BCE
Input data size	(128, 128, 3)	(128, 128, 3)	(128, 128, 3)
Number of training patches (true positive + true negative)	15 + 15	15 + 23	15 + 25

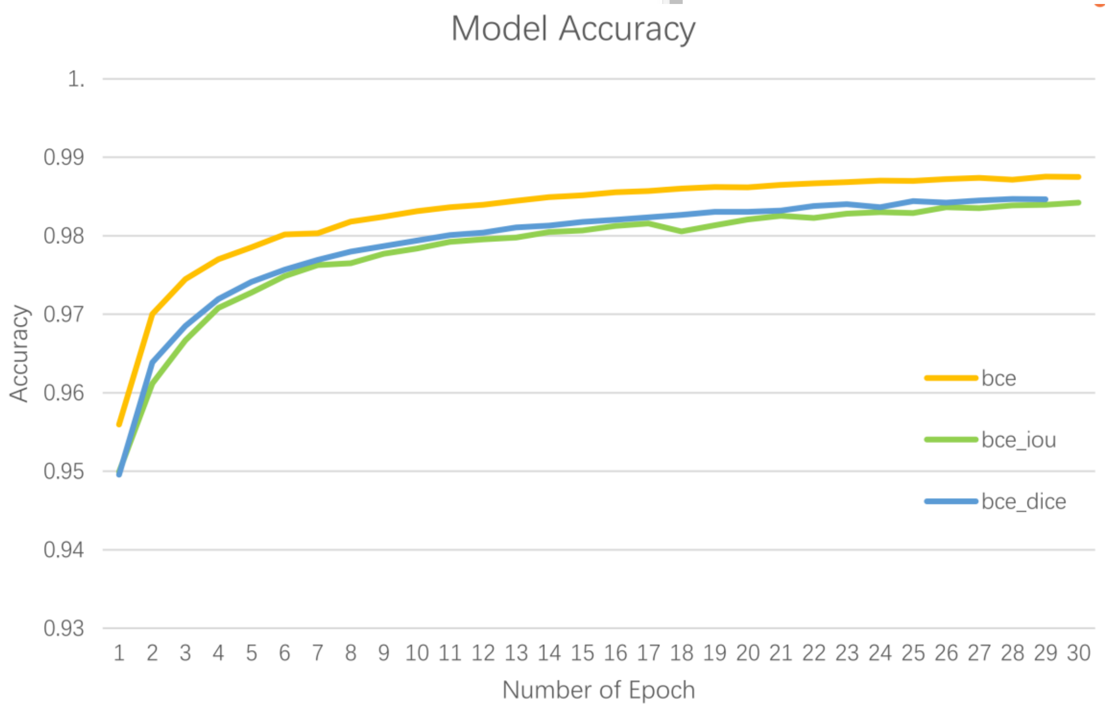
Experiment	4	5	6
Learning rate	5.00E-03	1.00E-03	1.00E-03
Batch size	5	3	5
Loss function	BCE	BCE	Intersection over Union (IoU)
Input data size	(128, 128, 3)	(128, 128, 3)	(128, 128, 3)
Number of training patches (true positive + true negative)	25 + 25	35 + 37	49 + 37

Experiment	7	8	9	10
Learning rate	1.00E-03	1.00E-03	1.00E-03	1.00E-03
Batch size	5	5	5	5
Loss function	Dice Loss	BCE + 0.2 * IoU	BCE + 0.2 * Dice	BCE + 0.2 * IoU
Input data size	(128, 128, 3)	(128, 128, 3)	(128, 128, 3)	(128, 128, 3)
Number of training patches (true positive + true negative)	49 + 37	49 + 37	49 + 37	49 + 37

Training schedule

Model Tuning





Post-processing

1. Ensemble predictions
2. Image Enhancement
   Morphological Opening + Closing
3. Word Boxes Localization and Extraction
   Morphological Closing
   Rotated Calipers
   Geocoding API, Nominatim

Evaluation

Intersection over Union:

$$IoU = \frac{|A \cap B|}{|A \cup B|}$$

Classification :

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN}$$

Text Recognition

Transkribus:

HTR+ engine

trained on:

- Synthetic data
- Manually labelled samples
- German historical documents from the National Library of Australia

MMOCR:

CRNN+STN model

ABiNet model

trained on: Syn90k dataset

Evaluation

Character Error Rate (CER) and Word Error Rate (WER)

$$F1 = \frac{errors}{correct + errors}$$

Confusion Metric for Error Analysis

Text detection – pre-trained weights

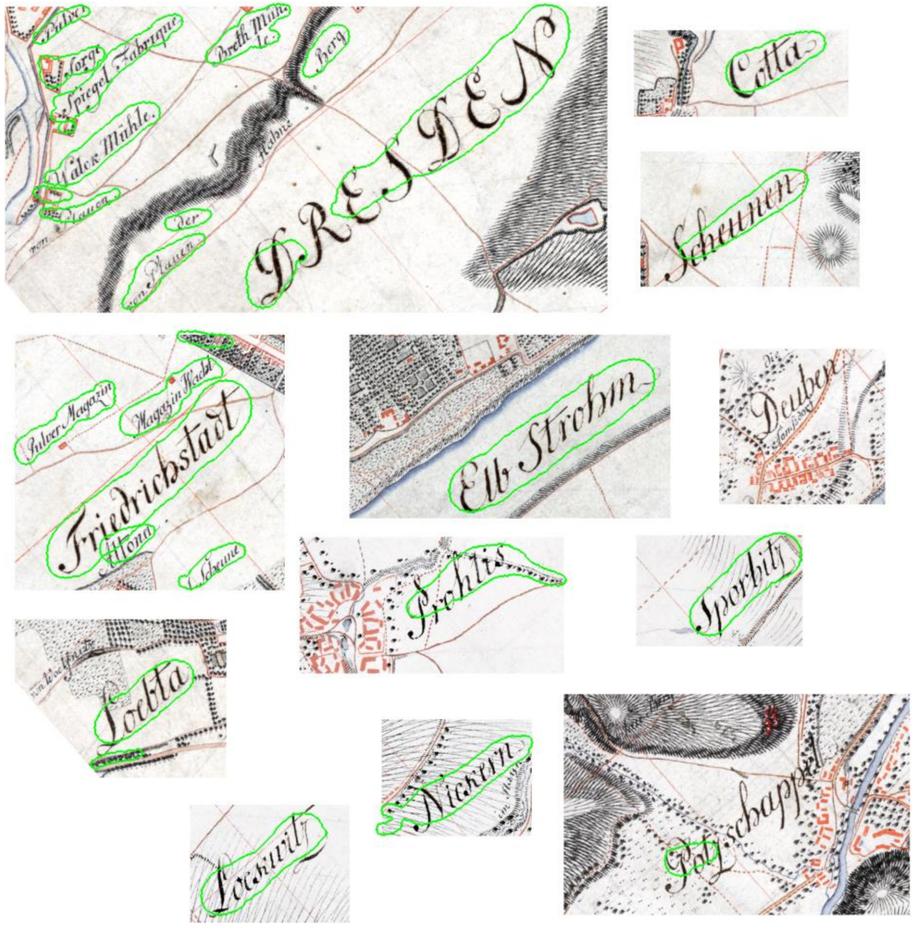
Prediction DBNet++

Prediction FCE

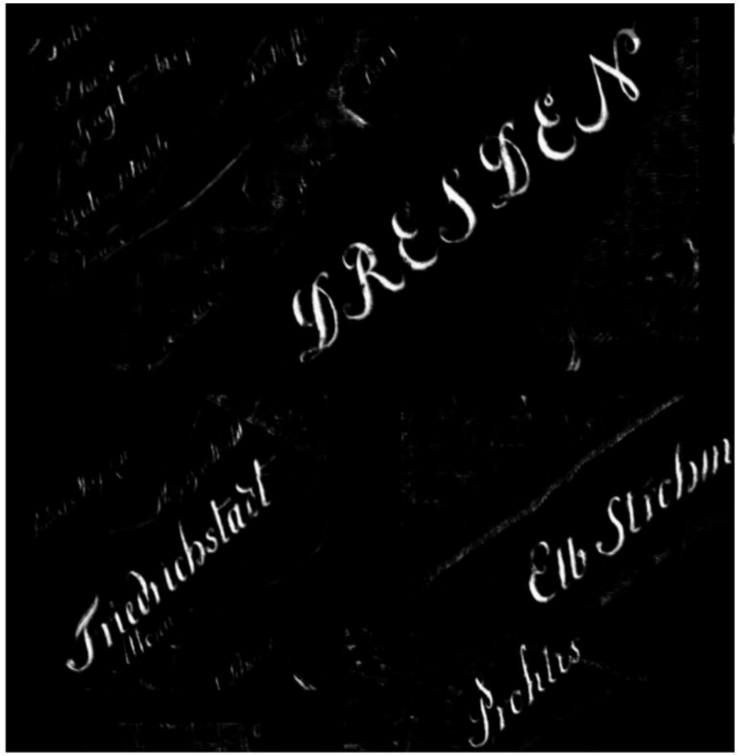
Text detection – pre-trained weights

Prediction DRRG

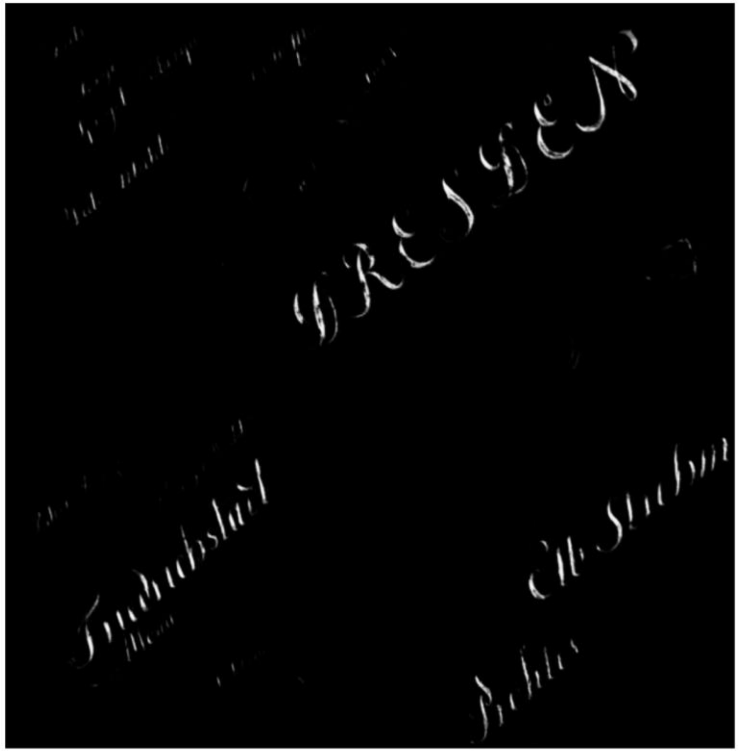
Prediction TextSnake



Segmentation results

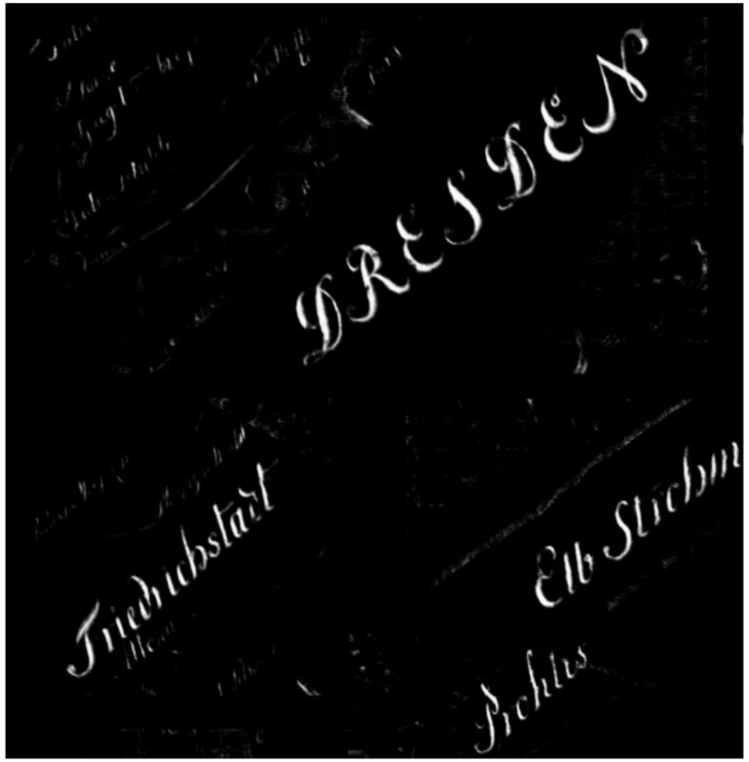


(a) BCE



(b) BCE + 0.2*IoU

Segmentation results – RQ1.2



(a) Synthetic training set

(b) Manually labelled ground truth

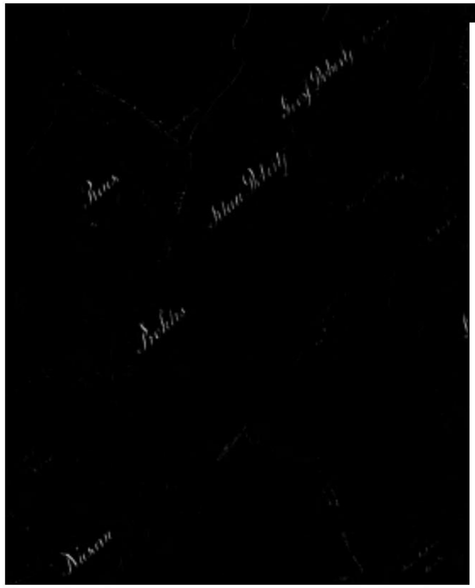
Ensemble result

(a) Prediction 1

(b) Prediction 2

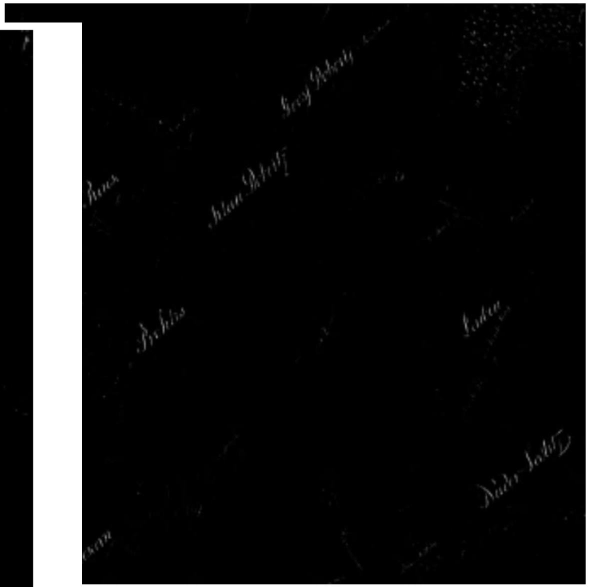
(c) Prediction 3

Ensemble result



(a) Prediction 1

(d) Ensemble prediction of weighted average



(c) Prediction 3

Motivation

RQs

Literature Review

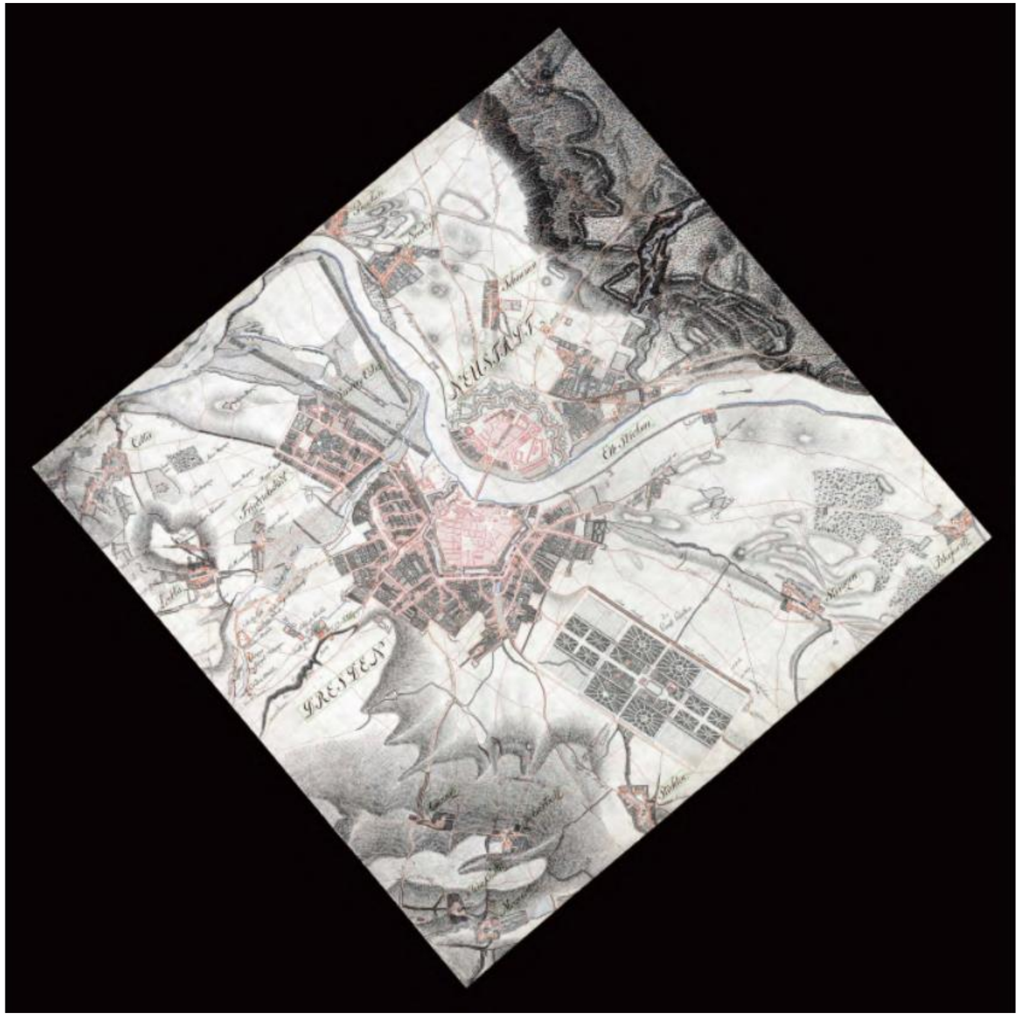
Material

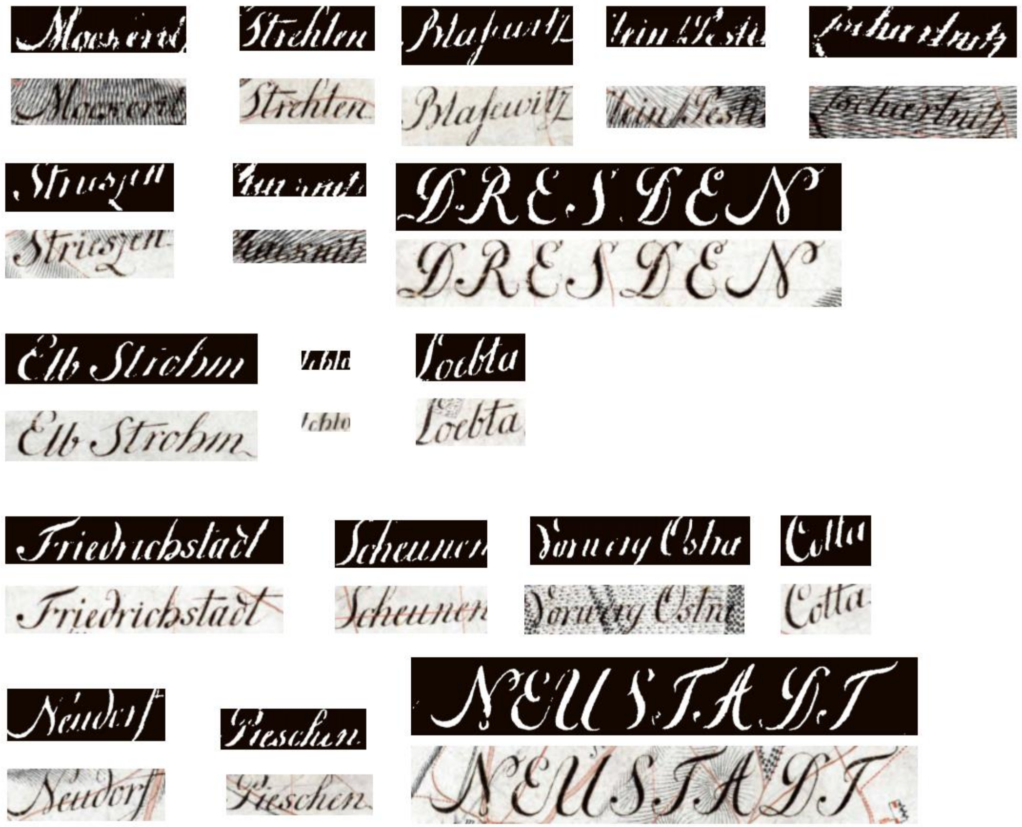
Methods

Results

Conclusion

Word boxes





Evaluation

Map	Precision (%)	Recall(%)	F1 (%)	IoU (%)
TestCrop	85.71	100	92.31	71.34
Neustadt	94.44	100	97.14	63.88
Woelckau	45.45	100	62.5	44.97
Nidersedlitz	62	100	76.54	57.19
Fuchshain	47.06	100	64	1.14

Text Recognition

Model	Input	CER (%)	WER (%)
CRNN+STN	RGB	16.24	76.92
	Extracted Binary	17.95	76.92
ABiNet	RGB	36.75	92.31
	Extracted Binary	41.03	92.31
HTR+	Extracted Binary (whole page)	88.89 -> 70.09 -> 52.51	100 -> 100 -> 95.65

Text Recognition

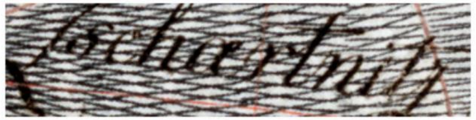
		Predicted Characters																									
		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
True Characters	a	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	2	0	0	0
	b	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
	c	0	0	7	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0	1	0	0	0	1	0	0	0
	d	0	1	0	6	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	1	0	0	0	0	0	0
	e	1	0	1	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	f	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	g	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	h	0	0	0	0	0	0	0	8	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	i	1	0	0	0	0	0	0	0	10	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	j	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	k	1	0	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	l	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	1	0	0	0	0	0
	m	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
	n	0	0	0	0	1	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0	0	0	0	0
	o	1	1	2	0	0	0	0	0	0	0	0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0
	p	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0
	q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	r	0	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	11	0	0	0	0	0	0	0	0
	s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	1	0	0	0	0	0	0
	t	0	0	0	0	1	0	0	0	1	0	0	4	0	0	0	0	0	0	1	9	0	0	0	0	0	0
	u	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
	v	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	w	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0
	x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	z	0	0	1	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	2

(a) HTR Predictions

		Predicted Characters																									
		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
True Characters	a	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	b	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	c	0	0	5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
	d	0	1	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	e	0	0	1	0	12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	f	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	g	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	h	0	0	0	0	0	0	0	3	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	i	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	j	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	k	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	l	0	0	0	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	m	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	n	0	0	0	0	1	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
	o	0	0	1	0	0	0	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	0	0	0	0	0
	p	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0
	q	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	r	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0
	s	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8	0	0	0	0	0	0	0
	t	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	6	0	0	0	0	0	0
	u	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
	v	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	w	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
	x	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	y	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	z	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0

(b) CRNN Prediction

Overall Performance

	ID	UTM	latlon	address	HTR+	CRNN
	word_1	(41198	(51.0180953594537, 13.745162075845785	Kleinpestitz/**Mockritz**, Zschertnitz, Plauen, Dresden	Bockerwitz	mouhered
	word_2	(41355	(51.02863364008097, 13.76727630737129	128a, Wiener Straße, Seevorstadt-Ost/Großer Garten, Alts	Wrehlen	strchlen
	word_3	(41633	(51.049619547963296, 13.8064412000307	19, Kretschmerstraße, **Blasewitz**, Dresden, Sachsen, 01	Blasewitz	bolafwng
	word_4	(41169	(51.02002789737003, 13.74105047686047	Räcknitz/Zschertnitz, Zschertnitz, Plauen, Dresden	Kein Gerlln	ppindursa
	word_5	(41225	(51.026328234310014, 13.7487423048190	Räcknitz/Zschertnitz, **Zschertnitz**, Plauen, Dresden	LnuKe Wdruth	ntlurehnlz
	word_6	(41524	(51.04635719422944, 13.79086641035864	Striesen-West, **Striesen**, Blasewitz, Dresden	Bausen	struaytn

Conclusion

RQ1.1 What are the existing deep learning pipelines for text extraction?

RQ1.2 How can synthetic training data help optimize the model training process?

RQ1.3 How well can toponyms be separated from the background using deep learning?

RQ2.1 What are the existing text recognition models?

RQ2.2 How well is the performance of the adapted text recognizer?

RQ3 How is the overall performance of the deep learning pipeline in toponym detection and recognition?

Conclusion

RQ1.1 What are the existing deep learning pipelines for text extraction?

RQ1.2 How can synthetic training data help optimize the model training process?

RQ1.3 How well can toponyms be separated from the background using deep learning?

RQ2.1 What are the existing text recognition models?

RQ2.2 How well is the performance of the adapted text recognizer?

RQ3 How is the overall performance of the deep learning pipeline in toponym detection and recognition?

Conclusion

RQ1.1 What are the existing deep learning pipelines for text extraction?

RQ1.2 How can synthetic training data help optimize the model training process?

RQ1.3 How well can toponyms be separated from the background using deep learning?

RQ2.1 What are the existing text recognition models?

RQ2.2 How well is the performance of the adapted text recognizer?

RQ3 How is the overall performance of the deep learning pipeline in toponym detection and recognition?

Conclusion

RQ1.1 What are the existing deep learning pipelines for text extraction?

RQ1.2 How can synthetic training data help optimize the model training process?

RQ1.3 How well can toponyms be separated from the background using deep learning?

RQ2.1 What are the existing text recognition models?

RQ2.2 How well is the performance of the adapted text recognizer?

RQ3 How is the overall performance of the deep learning pipeline in toponym detection and recognition?

Motivation

RQs

Literature Review

Material

Methods

Results

Conclusion

Conclusion

RQ1.1 What are the existing deep learning pipelines for text extraction?

RQ1.2 How can synthetic training data help optimize the model training process?

RQ1.3 How well can toponyms be separated from the background using deep learning?

RQ2.1 What are the existing text recognition models?

RQ2.2 How well is the performance of the adapted text recognizer?

RQ3 How is the overall performance of the deep learning pipeline in toponym detection and recognition?

Future Work

1. Improving the data synthesis

2. Testing more complex model on cloud-based server

3. Contextual analysis to correct the recognition results

4. An integrated tool to segment all features

References

Brunner, H. (2005). *Atlas zur Geschichte und Landeskunde von SachsenBeiheft zu den Karten H 12.1 und H 12.2 Die sächsische Landesaufnahme von 1780 bis 1825 / von Hans Brunner*. Verl. der Sächsischen Akad. der Wiss. zu Leipzig: Landesvermessungsamt Sachsen. https://katalog.slub-dresden.de/id/0-1473239567

Jordan, P. (2009). Some considerations on the function of place names on maps. Proceedings of the 24th international cartographic conference, Santiago, Chile (ICC 2009),

Liao, M., Wan, Z., Yao, C., Chen, K., & Bai, X. (2020, 2020). *Real-time scene text detection with differentiable binarization* Proceedings of the AAAI Conference on Artificial Intelligence,

Long, S., Ruan, J., Zhang, W., He, X., Wu, W., & Yao, C. (2018). TextSnake: A Flexible Representation for Detecting Text of Arbitrary Shapes. ECCV,

Shi, B., Bai, X., & Belongie, S. (2017, 2017). Detecting Oriented Text in Natural Images by Linking Segments.

Stams, M., & Stams, W. (1981). *Die große topographische Landesaufnahme von Sachsen 1780 bis 1811 und ihre Folgekarten*. Sektion Geodäsie u. Kartogr. d. Techn. Univ., Kartogr.-techn. Einrichtung.

THANK YOU!


ITC
UNIVERSITY OF TWENTE.
TECHNISCHE UNIVERSITÄT DRESDEN
Technical University of Munich
TUM
TU WIEN
TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology