

Finetune your OCR! Improving automated text recognition for early printed works by finetuning existing Tesseract models

fonen mit 2337812074 M urfichert bliekeven, davon 552246 Perfonen mit 1847622742 M bei ven 35 Se im Deutschen Reich. Der gesammte Zuwarters im Jahre 1877 stellt sich auf 15777 Perfonen mit 110873 820 M, d. h. 2,14 % der Zahl ver Versicherten und 4,98 % der Wachen, summe. Die Versichtigung der Unfälle, welche die Liquidation einiger Gesells Gesellschaften im Jahre 1877 zur Folge gehabt hat verwindert sich der Zuwachs auf 8542 (1 16%) der Versicherten

Jan Kamlah / Thomas Schmidt





- 1. OCR and work-specific training at UB Mannheim
- 2. Introduction
- 3. Training with synthetic ground truth (Approach A)
- 4. Training with real-world ground truth (Approach B)
- 5. Conclusion

Finetune your OCR! Improving automated text recognition for early printed works by finetuning existing Tesseract models



1. OCR and work-specific training at UB Mannheim

- 2. Introduction
- 3. Training with synthetic ground truth (Approach A)
- 4. Training with real-world ground truth (Approach B)
- 5. Conclusion

1. OCR and work-specific training at UB Mannheim



- since 2022: work-specific training (for eScriptorium + Kraken) with theological prints, 1860-72 (DigiTheo Ground Truth: <u>https://github.com/UB-Mannheim/DTGT</u>)
- since 2021: module project "work-specific training" (for Tesseract + Calamari) as part of the DFGfunded initiative "OCR-D"
- 2013–2019: digitisation / OCR "Aktienführer-Datenarchiv" (German stock exchange data 1979-99)
- 2019: work-specific training for "Weisthümer" (Jacob Grimm, 1840) <u>https://github.com/UB-Mannheim/Weisthuemer</u>
- 2011–2018: digitisation / OCR "Desbillons"-collection (French clergyman and scholar)





Jan Kamlah / Thomas Schmidt (Mannheim University Library): Finetune your OCR! Improving automated text recognition for early printed works by finetuning existing Tesseract models, June 9th, 2022

Finetune your OCR! Improving automated text recognition for early printed works by finetuning existing Tesseract models



1. OCR and work-specific training at UB Mannheim

2. Introduction

- 3. Training with synthetic ground truth (Approach A)
- 4. Training with real-world ground truth (Approach B)
- 5. Conclusion

Im Gangen waren im Sabre 1877 815272 Derfonen mit 2511 478 651 Ma versichert; hiervon schieden 13 109 Personen (35 906 127 M.) burch Tob und 48 754 Perfonen (137 760 445 M.) anderweitig aus, fo daß am Jahresschluß 753 409 Der. fonen mit 2 337 812 074 M. versichert blieben, davon 552 246 Perfonen mit 1 847 622 742 M. bei ben 35 Gefellichaften im Deutschen Reich. Der gesammte Bumachs im Jahre 1877 ftellt fich auf 15 777 Perjonen mit 110 873 820 16, b. 5. 2.14 % ber Babl der Bersicherten und 4,98 % ber Berficherungsfumme. Die Berücksichtigung ber Unfälle, welche bie Liquidation einiger Gesellschaften im Jahre 1877 zur Folge gehabt hat, vermindert fich ber Jumachs auf 8542 (1,16 %) ber Berficherten und 93 396 677 M. (4,19 %) ber Summe (gegen 4,35 bam. 6,57 %) in 1876; 5,30 bam. 8,09 % in 1875; 7,12 bam. 9,29 % in 1874; 8,84 bam. 11,44 % in 1873). Der Durchschnitt ber Summe belief fich pro Ropf im Jahre 1877 auf 3103 M, gegen 3021 M. in 1876. Den stärtiten Buwachs hatten im Sahre 1877 von ben 35 Gefellfcaften im Deutschen Reich Diejenigen ju Rarlerube (2 810 616 M.) und Gotha (1 167 800 M.), außerdem hatten noch 4 biefer Gefellichaften einen Buwachs, alle übrigen eine Abnahme ber Berficherungsfumme erlitten.

Deutscher Reichsanzeiger und Preußischer Staatsanzeiger, 1878, Issue 248

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



Im Gangen waren im Sabre 1877 815272 Personen mit 2511 478 651 M versichert; biervon schieden 13 109 Personen (35 906 127 M.) burch Lod und 48 754 Perfonen (137 760 445 M.) anderweitig aus, fo baf am Jahresichluß 753 409 Perfonen mit 2 337 812 074 M. versichert blieben, davon 552 246 Personen mit 1 847 622 742 Me bei ben 35 Gesellschaften im Deutschen Reich. Der gefammte Buwachs im Sabre 1877 ftellt fich auf 15 777 Perjonen mit 110 873 820 M. b. b. 2.14 % ber Babl ber Berficherten und 4,98 % Berficerungs. ber fumme. Die Berüchsichtigung ber Unfälle, melche Die Liquidation einiger Gefellschaften im Jahre 1877 zur Folge gehabt hat, vermindert fich ber Zumachs auf 8542 (1.16%) ber Berficherten und 93 396 677 M (4,19 %) ber Summe (gegen 4,35 bam. 6,57 %) in 1876; 5,30 bam. 8,09 % in 1875; 7,12 bam. 9,29 % in 1874; 8,84 bam. 11,44 % in 1873). Der Durchschnitt ber Summe belief fich pro Ropf im Jabre 1877 auf 3103 M, gegen 3021 M. in 1876. Den stärtiten Buwachs hatten im Sahre 1877 von ben 35 Gefellschaften im Deutschen Reich Diejenigen ju Karlsrube (2810 616 M.) und Gotha (1 167 800 .46), außerdem hatten noch 4 biefer Gefellichaften einen Buwachs, alle übrigen eine Abnahme ber Berficherungsfumme erlitten.

Deutscher Reichsanzeiger und Preußischer Staatsanzeiger, 1878, Issue 248

UNIVERSITÄTSBIBLIOTHEK MANNHEIM

Im Ganzen waren im Jahre 1877 815272 Perfonen mit 2511478 651 -ÄZ erfichert; hiervon fchieden 13 109 Perfonen (35 906 127 ìì) durch Tod und 48 754 11 (137 760 445 kìì anderweitig aus, fo daß am Jahresfchluß 753 409 Per≉ fonen mit 2337 812 074 &α verfichert blieben, davon 552 246 Perfonen mit 1847 622 742 bei den 35 re 157 fteh im Deutschen Reich. Der gesammte Zuwachs im Jahre 1877 ftellt (ich auf 15 777 Perionen mit 110 873 820 , d. h. 2,14 %0 der Zahl der Verficherten und 4,98 % der Wunche die fumme. Die Berückfichtigung der Unfälle, welche die Liquidation einiger Gefellfchaften im Jahre 1877 zur Folge gehabt hat, vermindert fich der Zuwachs auf 8542 (1,16%) der Verficherten und 93 396 677 α (4,19%) der Summe (gegen 4,35 bzw. 6,57 / in 1876; 5,30 bzw. 8,09 % in 1875; 7,12 bzw. 9,29 % in 1874; 8,84 bzw. 11,44% in 187, Der Durchfchnitt der Summe belief fich pro Kopf im Jahre 1877 auf 3103 α , gegen 302 σ in 1876. Den ftärk∫ten Zuwachs hatten im Jahre 1877 von den 35 Gefell≉ Chaften im DeutChen Reich diejenigen zu Karlsruhe (2810 616 und Gotha (1 167 800), außerdem hatten noch 4 dieſer Geſell≠ ſchaften einen Zuwachs, alle übrigen eine Abnahme der Verſiche≠ rungsfumme erlitten.

OCR result

Tesseract (tesseract 5.1.0-20) Model: frak2021 (UB Mannheim)

Jan Kamlah / Thomas Schmidt (Mannheim University Library):

Im Gangen waren im Sabre 1877 815272 Dersonen mit 2511 478 651 (M) versichert; biervon schieden 13 109 Personen (35 906 127 (M.) ourch Lod und 48 754 Perfonen (137 760 445 (M.) anderweitig aus, fo bak am Sabresichluft 753 409 fonen mit 2 337 812 074 (M.) persichert blieben, davon 552 246 Personen mit 1 847 622 742 (M.) bei ben 35 Gesellschaften im Deutschen Reich. Der gesammte Buwachs im Jahre 1877 ftellt fich auf 15 777 Per onen mit 110 873 820 (16, b. b. 2.14 % ber Bahl der Berficherten und 4,98 % ber Berficberunge-Die Berudfictigung fumme. Unfälle, ber melche Liquidation einiger Gefellschaften im Jahre 1877 zur Folge gehabt hat, vermindert fich ber Jumachs auf 8542 (1,16 %) ber Berficherten und 93 396 677 (M.) (4,19 %) ber Summe (gegen 4,35 bzw. 6,57 %) in 1876; 5,30 bin. 8,09 % in 1875; 7,12 bim. 9,29 % in 1874; 8,84 bzw. 11,44 % in 1873). Der Durchschnitt ber Summe belief fich pro Ropf im Jabre 1877 auf 3103 (M.) gegen 3021 (M.) in 1876. Den ftartiten Buwachs hatten im Sabre 1877 von ben 35 Gejellschaften im Deutschen Reich Diejenigen ju Karlerube (2 810 616 (M.) und Gotha (1 167 800 (M.), außerdem hatten noch 4 biefer Gefen. ichaften einen Buwachs, alle übrigen eine Abnahme ber Berficherungsfumme erlitten.

Deutscher Reichsanzeiger und Preußischer Staatsanzeiger, 1878, Issue 248



Im Ganzen waren im Jahre 1877 815272 Perfonen mit 2511478 651 -ÄZ erfichert; hiervon fchieden 13 109 Perfonen $(35\ 906\ 127\ 11)$ durch Tod und 48 754 11 (137 760 445 κ 1) anderweitig aus, fo daß am Jahresfchluß 753 409 Per# fonen mit 2337 812 074 & verfichert blieben, davon 552 246 Perfonen mit 1847 622 742 bei den 35 re 157 fteh im Deutschen Reich. Der gesammte Zuwachs im Jahre 1877 ftellt sich auf 15 777 Perionen mit 110 873 820, d. h. 2,14 %0 der Zahl der Verficherten und 4,98 % der Wunche die fumme. Die Berückfichtigung der Unfälle, welche die Liquidation einiger Gefellfchaften im Jahre 1877 zur Folge gehabt hat, vermindert-(ich der Zuwachs auf 8542 (1,16%) der Verficherten und 93 396 677 α (4,19%) der Summe (gegen 4,35 bzw. 6,57 / in 1876; 5,30 bzw. 8,09 % in 1875; 7,12 bzw. 9,29 % in 1874; 8.84 bzw. 11.44% in 187. Der Durchfchnitt der Summe belief fich pro Kopf im Jahre 1877 auf 3103 (α , gegen 302 σ in 1876. Den ſtärkſten Zuwachs hatten im Jahre 1877 von den 35 Geſell≤ ſchaften im Deutſchen Reich diejenigen zu Karlsruhe (2810 616 und Gotha (1 167 800)) außerdem hatten noch 4 dieſer Geſell≤ ſchaften einen Zuwachs, alle übrigen eine Abnahme der Verſiche≠ rungsfumme erlitten.

OCR result Tesseract (tesseract 5.1.0-20) Model: frak2021 (UB Mannheim)

Jan Kamlah / Thomas Schmidt (Mannheim University Library):

Im Gangen waren im Jahre 1877 815272 Perfonen mit 2511 478 651 (Ma) versichert; biervon schieden 13 109 Personen (35 906 127 (M.) ourch Lod und 48 754 Perfonen (137 760 445 (M.) anderweitig aus, fo bag am Sabresichluß 753 409 fonen mit 2 337 812 074 (M.) persichert blieben, davon 552 246 Perfonen mit 1 847 622 742 (M.) bei ben 35 Gefellichaften im Deutschen Reich. Der gesammte Buwachs im Sahre 1877 ftellt fich auf 15 777 Per onen mit 110 873 820 (16, b. b. 2.14 %) ber Bahl der Berficherten und 4,98 % ber Berficherungsfumme. Die Berücksichtigung ber Unfälle, melche Liquidation einiger Gefellschaften im Jahre 1877 zur Folge gehabt hat, vermindert fich der Zuwachs auf 8542 (1,16 %) ber Berficherten und 93 396 677 (M.) (4,19 %) ber Summe (gegen 4,35 bzw. 6,57 %) in 1876; 5,30 bin. 8,09 % in 1875; 7,12 bim. 9,29 % in 1874; 8,84 bzw. 11,44 % in 1873). Der Durchichnitt ber Summe belief fich pro Ropf im Jahre 1877 auf 3103 (M.) gegen 3021 (M.) in 1876. Den ftartiten Buwachs hatten im Sabre 1877 von ben 35 Gejellschaften im Deutschen Reich Diejenigen ju Karleruhe (2810 616 (M.) und Gotha (1 167 800 (M.), außerdem batten noch 4 biefer Gefett. schaften einen Buwachs, alle übrigen eine Abnahme ber Berficherungsfumme erlitten.

Deutscher Reichsanzeiger und Preußischer Staatsanzeiger, 1878, Issue 248



Im Ganzen waren im Jahre 1877 815272 Perfonen mit 2511 478 651 werlichert; hiervon [chieden 13 109 Perlonen] $(35 906 127 \mathcal{M})$ durch Tod und 48 754 14. $(137 760 445 \mathcal{M})$ anderweitig aus, ſo daß am Jahresſchluß 753 409 Per≠ fonen mit 2 337 812 074 (M verfichert blieben, davon 552 246 Perfonen mit 1 847 622 747 M bei den 35 Se im Deutschen Reich. Der gesammte Zuwachs im Jahre 1877 Stellt sich auf 15777 Per onen mit 110 873 820 M, d. h. 2,14 % der Zahl der Verficherten und 4,98 % der Wachen, fumme. Die Berückfichtigung der Unfälle, welche die Liquidation einiger Gefellfchaften im Jahre 1877 zur Folge gehabt hat, vermindert_(ich der Zuwachs auf 8542 (1,16%) der Verlicherten und 93 396 677 (M) (4,19 %) der Summe (gegen 4,35 bzw. 6,57 0% in 1876; 5,30 bzw. 8,09% in 1875; 7,12 bzw. 9,29 in 1874; 8,84 bzw. 11,44 % in 71. Der Durch[chnitt der Summe belief fich pro Kopf im Jahre 1877 auf 3103(*M*.) gegen 3021(*M*) in 1876. Den ftärkften Zuwachs hatten im Jahre 1877 von den 35 Gefelle chaften im Deutschen Reich diejenigen zu Karleruhe (2 810 616 س) und Gotha (1 167 800 M), außerdem hatten noch 4 diefer Gefell ſchaften einen Zuwachs, alle übrigen eine Abnahme der Verſiche rungsfumme erlitten.

OCR result

Tesseract (tesseract 5.1.0-20) Model: frak2021 (UB Mannheim) <u>after work-specific</u> <u>training</u>

Jan Kamlah / Thomas Schmidt (Mannheim University Library):

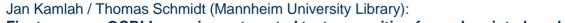


Reasons for a work-specific training ...

Jan Kamlah / Thomas Schmidt (Mannheim University Library): Finetune your OCR! Improving automated text recognition for early printed works by finetuning existing Tesseract models, June 9th, 2022

Reasons for a work-specific training ...

• improve text recognition quality







Reasons for a work-specific training ...

- improve text recognition quality
- expand the character set to be recognized → historical currency symbols, astronomical or mathematical signs, etc.

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



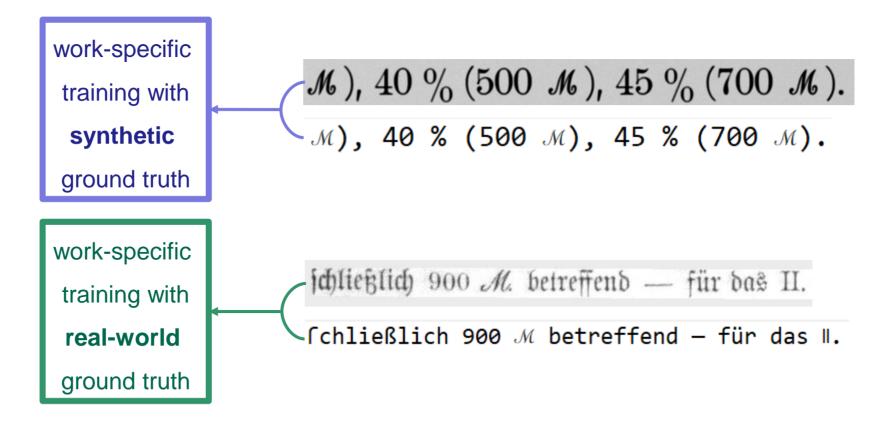
Reasons for a work-specific training ...

- improve text recognition quality
- expand the character set to be recognized → historical currency symbols, astronomical or mathematical signs, etc.
- specific mappings → for example: normalizations, resolution of abbreviations, dealing with white spaces (letterspaced text), etc.

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



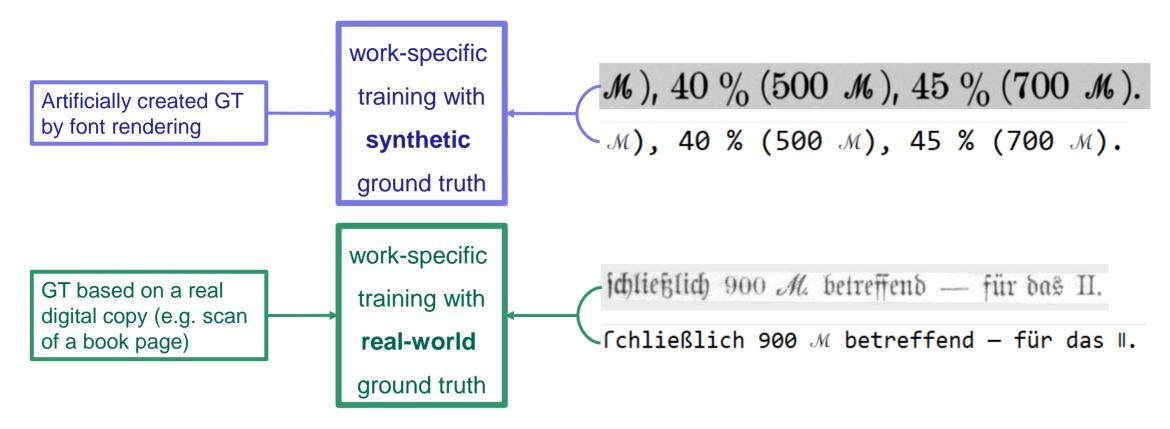
Two tested approaches:



Jan Kamlah / Thomas Schmidt (Mannheim University Library):



Two tested approaches:



Jan Kamlah / Thomas Schmidt (Mannheim University Library):



- 1. OCR and work-specific training at UB Mannheim
- 2. Introduction
- 3. Training with synthetic ground truth (Approach A)
- 4. Training with real-world ground truth (Approach B)
- 5. Conclusion



• **Text2Image** (Tesseract tool) renders text lines from text source and font

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



• Text2Image (Tesseract tool) renders text lines from text source and font

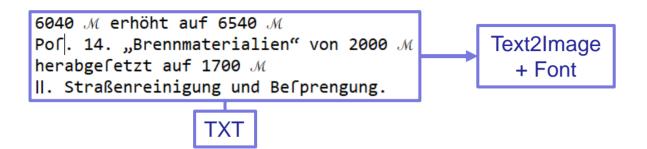
6040 *M* erhöht auf 6540 *M* Poſ . 14. "Brennmaterialien" von 2000 *M* herabgeſetzt auf 1700 *M* ||. Straßenreinigung und Beſprengung.



Jan Kamlah / Thomas Schmidt (Mannheim University Library):



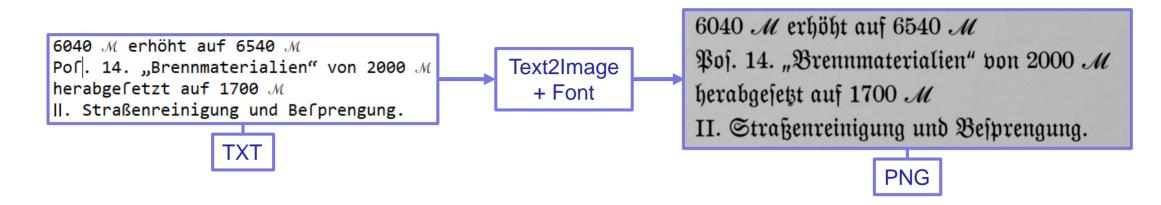
• Text2Image (Tesseract tool) renders text lines from text source and font



Jan Kamlah / Thomas Schmidt (Mannheim University Library):



• Text2Image (Tesseract tool) renders text lines from text source and font



Jan Kamlah / Thomas Schmidt (Mannheim University Library):



• 2 fonts used for rendering:

Gründerkrach=Fraktur	Verthold Mainzer Fraktur	
corresponds to the font used in the source material	resembles the font used in the source material	
large character set: long ſ, round א, $\mathcal M$ etc.	large character set: long ſ, round א, \mathcal{M} etc.	

Table 1: Fonts used for rendering

Jan Kamlah / Thomas Schmidt (Mannheim University Library):

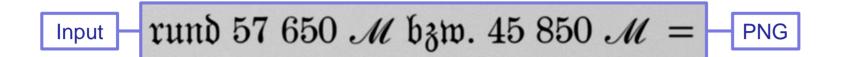


• Additional script to enrich 50 % of the image lines per data augmentation

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



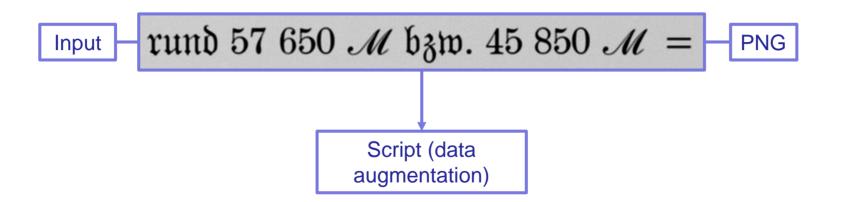
• Additional script to enrich 50 % of the image lines per data augmentation



Jan Kamlah / Thomas Schmidt (Mannheim University Library):



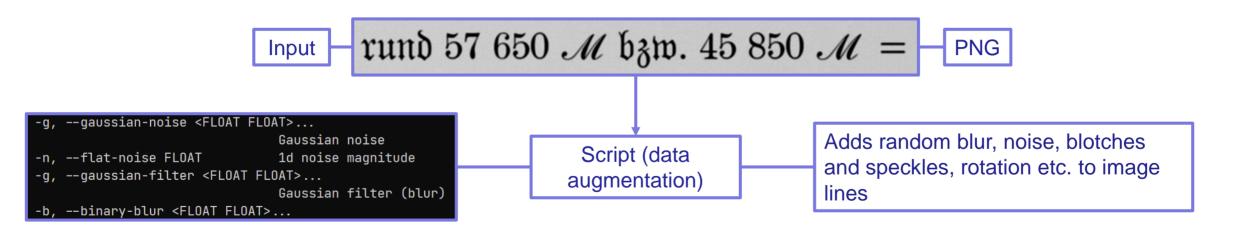
• Additional script to enrich 50 % of the image lines per data augmentation



Jan Kamlah / Thomas Schmidt (Mannheim University Library):



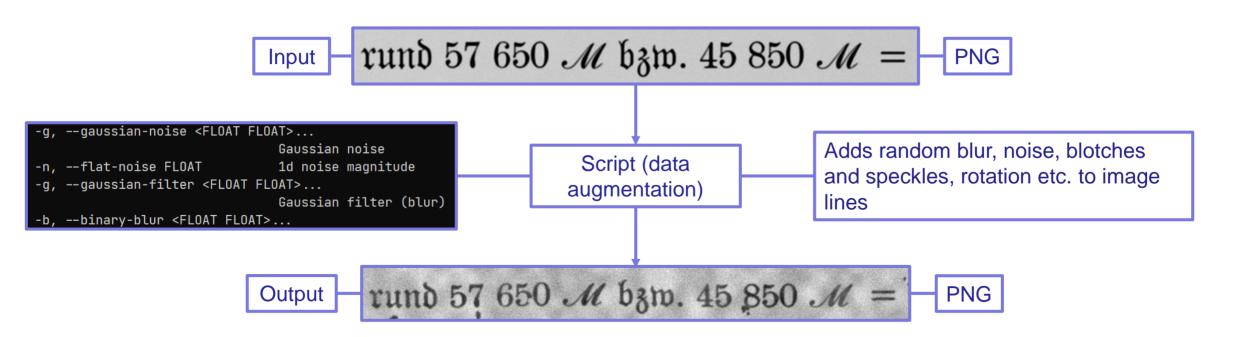
• Additional script to enrich 50 % of the image lines per data augmentation



Jan Kamlah / Thomas Schmidt (Mannheim University Library):



• Additional script to enrich 50 % of the image lines per data augmentation



Jan Kamlah / Thomas Schmidt (Mannheim University Library):



Table 2: Training sets with synthetic ground truth

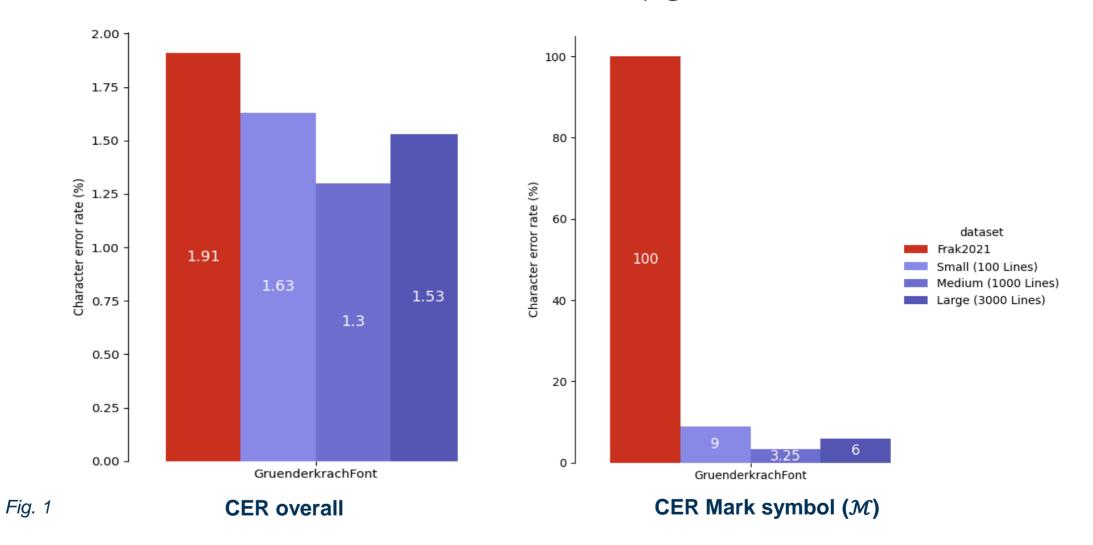
Work-specific training with synthetic ground truth (2 Fonts / 6 Datasets)						
Base model:	frak2021 (UB Mannheim)					
GT source:	"Charlottenburger Amtsschrifttum", 1879–1919 (ZLB)					
Fonts:	Gründerkrach=Fraktu	r	Berthold Mainzer Fraktur			
	© 2021 Philipp Poll	© Peter Wiegel				
Synthetic GT:	LARGE* 3.000 GT lines	MEDIUM[†] 1.000 GT lines		SMALL [†] 100 GT lines		
	PNG, grayscale, 50 % enriched by data augmentation					
Target:	Extension of the character set: \mathcal{M}					
Runtime:	60 epochs with termination criterion \rightarrow Actual runtime approx. 40–50 epochs					

* set contains lines with and without symbol \mathcal{M} ; * set contains only lines with symbol \mathcal{M}

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



Evaluation: Gründerkrach=Fraktur





2.00 100 1.75 80 1.50 Character error rate (%) Character error rate (%) 1.25 60 dataset Frak2021 1.98 1.00 1.91 100 97.2 Small (100 Lines) Medium (1000 Lines) 1.61 1.57 Large (3000 Lines) 0.75 40 0.50 20 0.25 22.6 11.06 0.00 ۲ ٥ MainzerFrakturFont MainzerFrakturFont

Evaluation: Verthold Mainzer Fraktur

Fig. 2

CER overall

CER Mark symbol (\mathcal{M})



- 1. OCR and work-specific training at UB Mannheim
- 2. Introduction
- 3. Training with synthetic ground truth (Approach A)
- 4. Training with real-world ground truth (Approach B)
- 5. Conclusion



• like the synthetic training, the source material for our ground truth production was the "Charlottenburger Amtsschrifttum", 1879-1919 (Zentral- und Landesbibliothek Berlin)

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



- like the synthetic training, the source material for our ground truth production was the "Charlottenburger Amtsschrifttum", 1879-1919 (Zentral- und Landesbibliothek Berlin)
- the transcription of the source material was done in Transkribus

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



- like the synthetic training, the source material for our ground truth production was the "Charlottenburger Amtsschrifttum", 1879-1919 (Zentral- und Landesbibliothek Berlin)
- the transcription of the source material was done in Transkribus
- in contrast to the synthetic training, only a grayscale normalization for the image contents was performed as data augmentation

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



- like the synthetic training, the source material for our ground truth production was the "Charlottenburger Amtsschrifttum", 1879-1919 (Zentral- und Landesbibliothek Berlin)
- the transcription of the source material was done in Transkribus
- in contrast to the synthetic training, only a grayscale normalization for the image contents was performed as data augmentation
- both number of ground truth lines as well as training runtime (in epochs) were identical to synthetic training

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



Table 4: Training sets with real-world ground truth

Work-specific training with real-world ground truth (3 Datasets)					
Base model:	frak2021 (UB Mannheim)				
GT Source:	"Charlottenburger Amtsschrifttum", 1881–1906 (ZLB)				
Ground Truth:	LARGE* 3.000 GT lines	MEDIUM[†] 1.000 GT lines	SMALL[†] 100 GT lines		
	PNG, grayscale (+ grayscale normalization)				
Target:	Extension of the character set: \mathcal{M}				
Runtime:	60 epochs with termination criterion \rightarrow Actual runtime approx. 40–50 epochs				

* set contains lines with and without symbol \mathcal{M} ; † set contains only lines with symbol \mathcal{M}

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



Evaluation: real-world ground truth

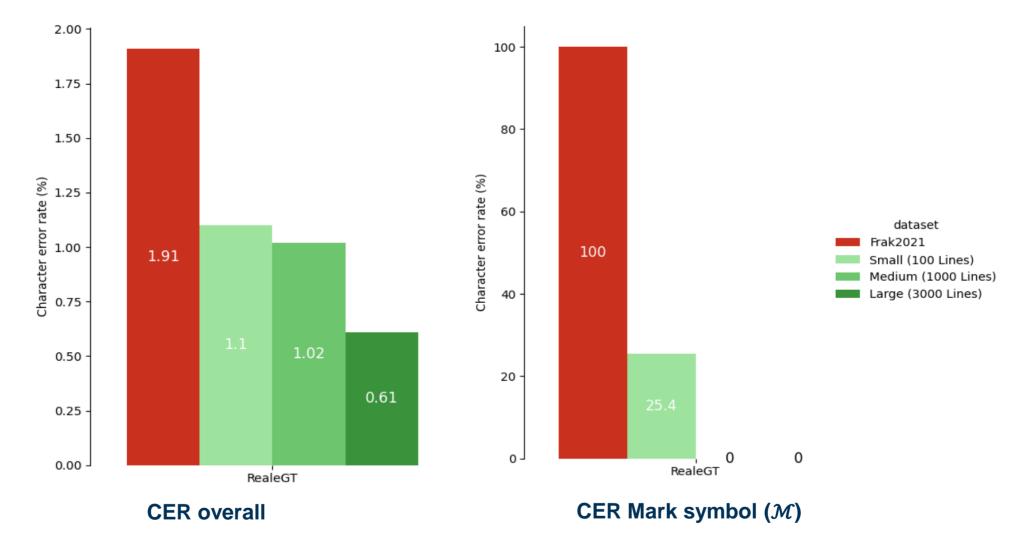


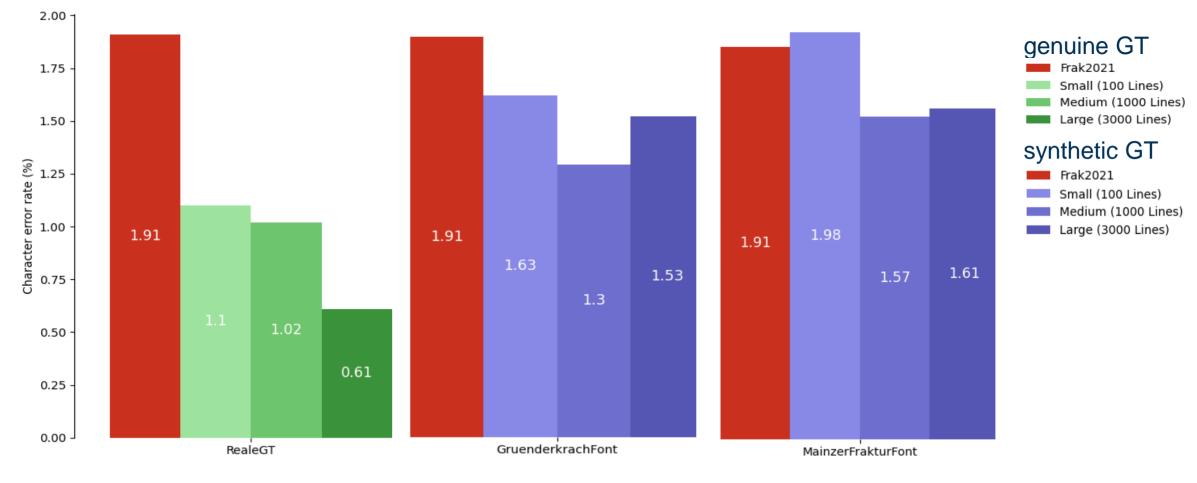
Fig. 3



- 1. OCR and work-specific training at UB Mannheim
- 2. Introduction
- 3. Training with synthetic ground truth (Approach A)
- 4. Training with real-wold ground truth (Approach B)
- 5. Conclusion



CER overall (synthetic and real-world GT)





CER Mark symbol \mathcal{M} (synthetic and real-world GT)

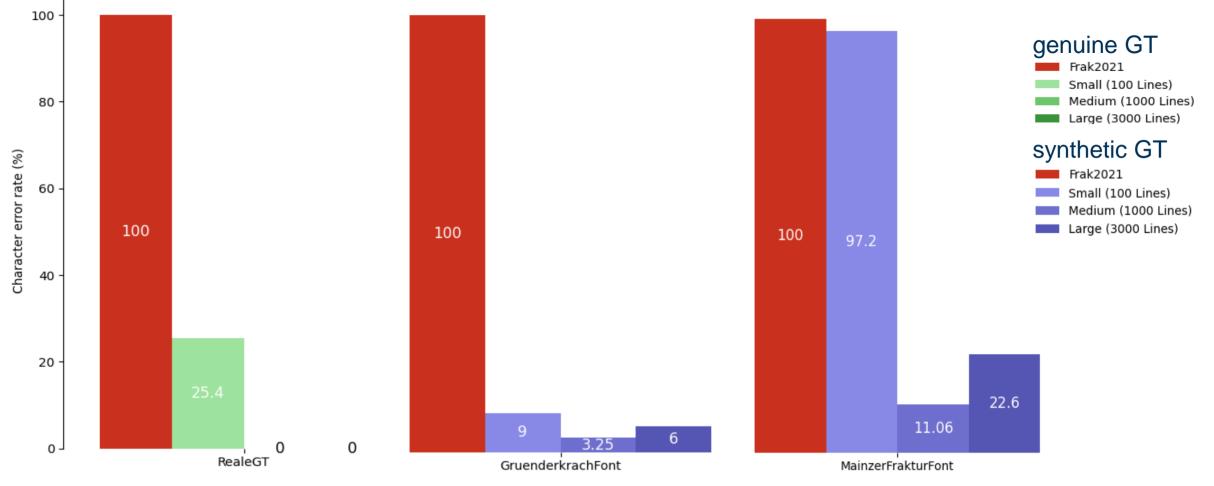


Fig. 5



Training with synthetic ground truth		Training with real-world ground truth	
Advantage	Disadvantage	Advantage	Disadvantage
Low time required to generate extensive GT sets	Fonts corresponding to source material are rarely available	Very good improvement of the base model	Relatively high effort to create the GT (selection of corresponding text passages with eligible glyphs + transcription + quality assurance)
Very good results when used font matches source material	Character set of available fonts mostly insufficient (especially for rare glyphs)	Trained model is superior to trained models with synthetic ground truth	
	Rare glyphs inconsistently encoded in different fonts		
	In general, fonts do not include glyph variations		

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



Berthold Mainzer Fraktur (synthetic ground truth):

Ende März 1892 auf 25 587 525,82 M gegen 23 606 652,22 M Ende

Gründerkrach Fraktur (synthetic ground truth):

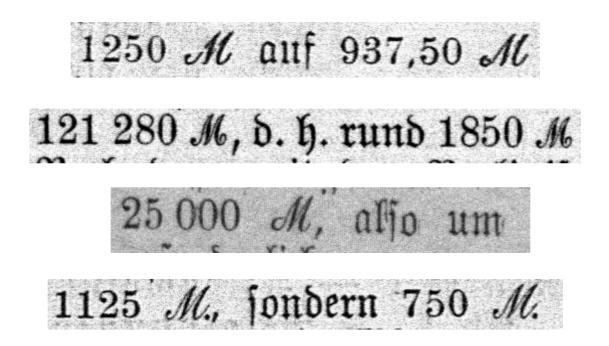
Ende März 1892 auf 25 587 525,82 M gegen 23 606 652,22 M Ende

"Charlottenburger Amtsschrifttum" (real-world ground truth): Ende März 1892 auf 25587525,82 *M.* gegen 23606652,22 *M.* Ende

Jan Kamlah / Thomas Schmidt (Mannheim University Library):



"Charlottenburger Amtsschrifttum" (real-world ground truth): more \mathcal{M} variants



Jan Kamlah / Thomas Schmidt (Mannheim University Library):



- Finetuning is an iterative and adaptive modern approach
- Besides pre- and post-processing, finetuning is an important tool to optimize OCR results
- Even a few hundred lines of ground truth can have a huge impact on training results
- Other projects following this approach:
 - Transkribus
 - eScriptorium
 - OCR4AII
 - PERO-OCR

Finetune Your OCR!

Jan Kamlah / Thomas Schmidt (Mannheim University Library):

Finetune your OCR! Improving automated text recognition for early printed works by finetuning existing Tesseract models



Ground truth data, further evaluations and additional information on GitHub: https://github.com/UB-Mannheim/charlottenburger-amtsschrifttum

Thank you!

Jan Kamlah (Development): jan.kamlah@bib.uni-mannheim.de Thomas Schmidt (Project management): thomas.schmidt@bib.uni-mannheim.de

References



- "Charlottenburger Amtsschrifttum", 1879–1919 (ZLB), https://digital.zlb.de/viewer/metadata/16046633/1/
- BertholdMainzerFraktur, Peter Wiegel, <u>http://www.peter-wiegel.de/MainzerFraktur.html</u>
- Weil, Stefan: Neue Frakturmodelle für Tesseract, 2019, https://madoc.bib.uni-mannheim.de/53748
- Weil Stefan: 126 Jahre Zeitung online Fundgrube f
 ür historisch Interessierte und Motor f
 ür die Bibliotheks-IT : 126 years of the newspaper online, 2018, <u>https://madoc.bib.uni-mannheim.de/46507/</u>
- Weil, Stefan und Kamlah, Jan: Forschungsdaten aus Digitalisaten, 2019, <u>https://madoc.bib.uni-mannheim.de/52204</u>