



IEEE ICIP 2023

Optical Character Recognition for Medical Records Digitization with Deep Learning

Author: Muhammad Ateeque Zaryab

Co-Author: ChuenRue Ng

Biomedical Engineering & Information Technology TU Ilmenau, Germany



TECHNISCHE UNIVERSITÄT
ILMENAU

Agenda

- ▶ Introduction
- ▶ Related Work
- ▶ Methodology
 - Dataset & Data Preprocessing
 - Training & Experimental Configurations
- ▶ Results
 - ROI Detection Model Results
 - OCR Model Results
- ▶ Discussion & Future Work

Introduction

Description:

- ▶ Optical character recognition (OCR) is a field of research in pattern recognition, artificial intelligence (AI), and computer vision (CV).
- ▶ OCR is a method of electronic translation of the handwritten, typewritten, or printed text into machine-translated images.
- ▶ forsorge aims to ensure optimized, self-determined, and barrier-free emergency treatment.

Motivation:

- ▶ Advancements in the field of AI, Machine Learning (ML), and Deep Learning (DL) techniques.
- ▶ Digitization of medical records.

Introduction

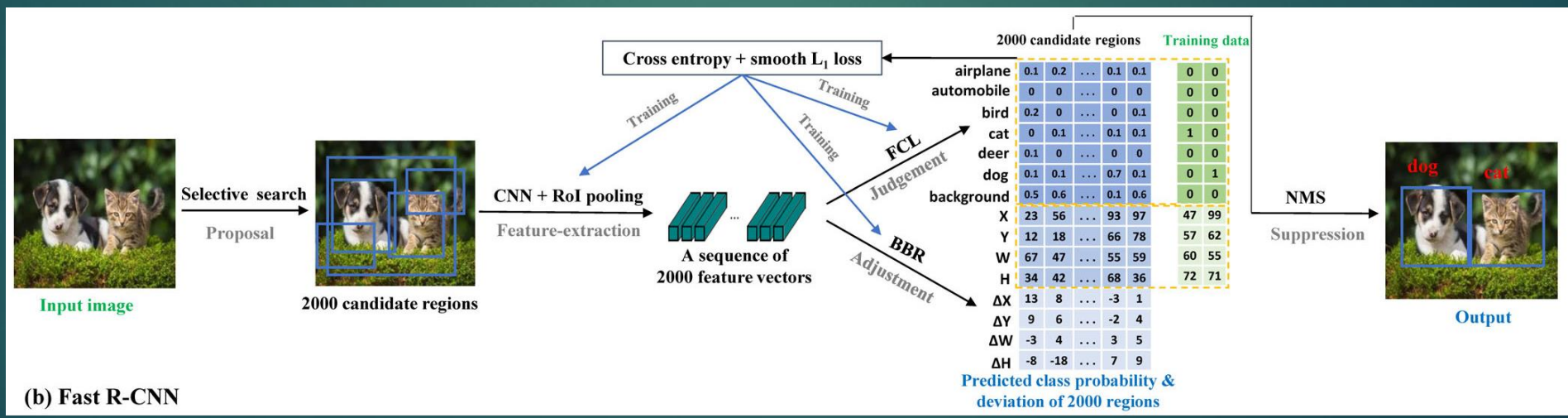
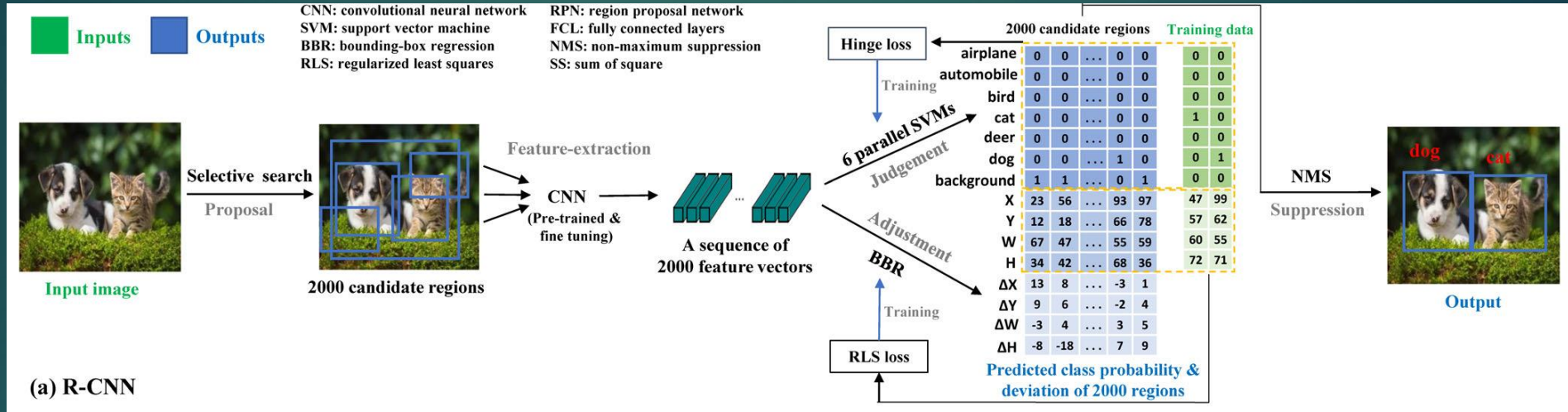
Objectives:

- ▶ To develop and compare the region of interest (ROI) detection algorithms for identifying the handwritten text and marked checkboxes in medical forms utilizing various neural network (NN) based methodologies.
- ▶ To train and evaluate several NN models for OCR that can also recognize German words and handwriting.
- ▶ Employing ROI detection and OCR to comprehensively digitize textual medical documents.

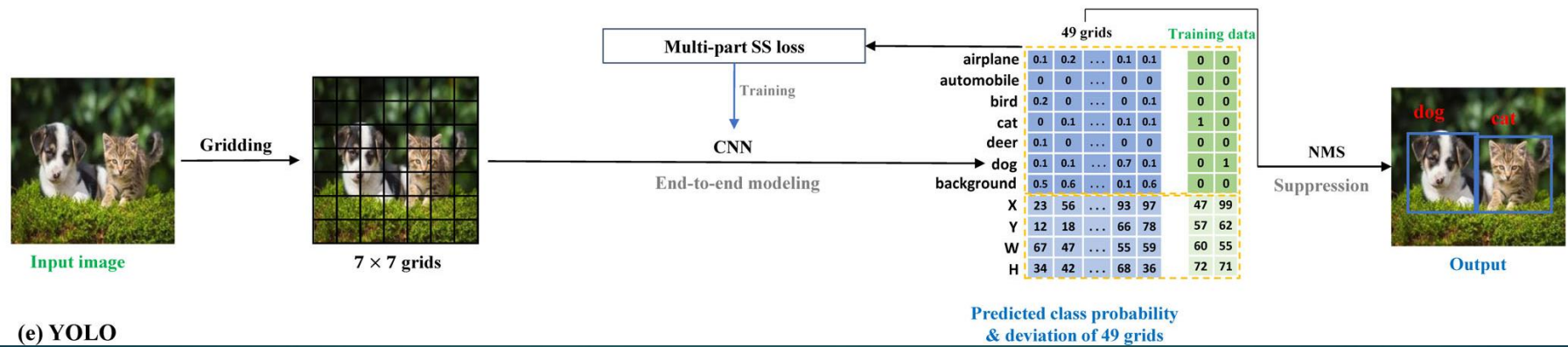
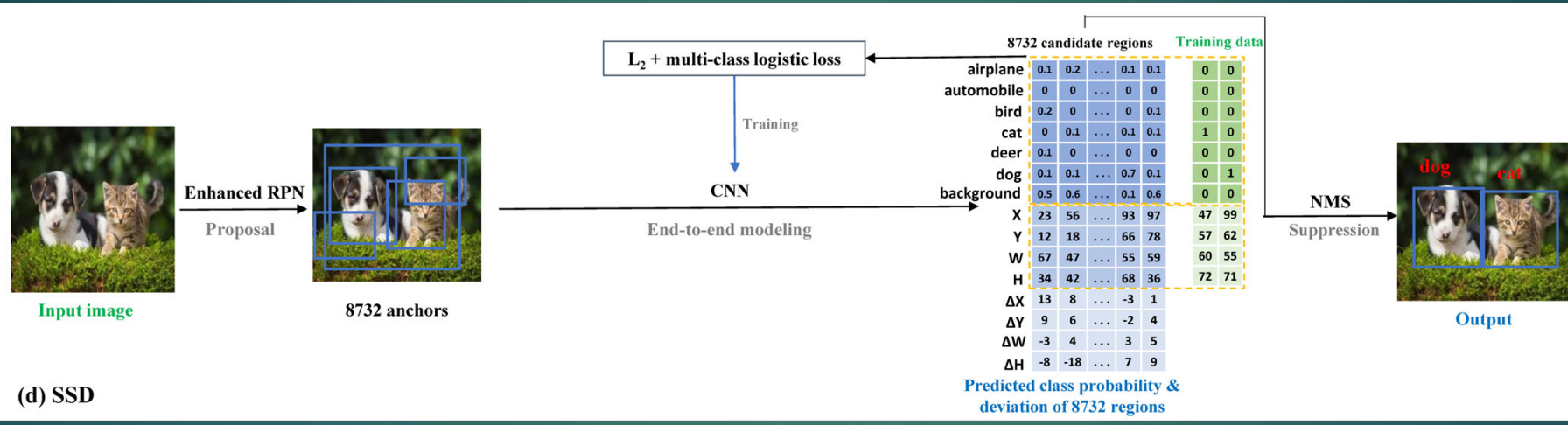
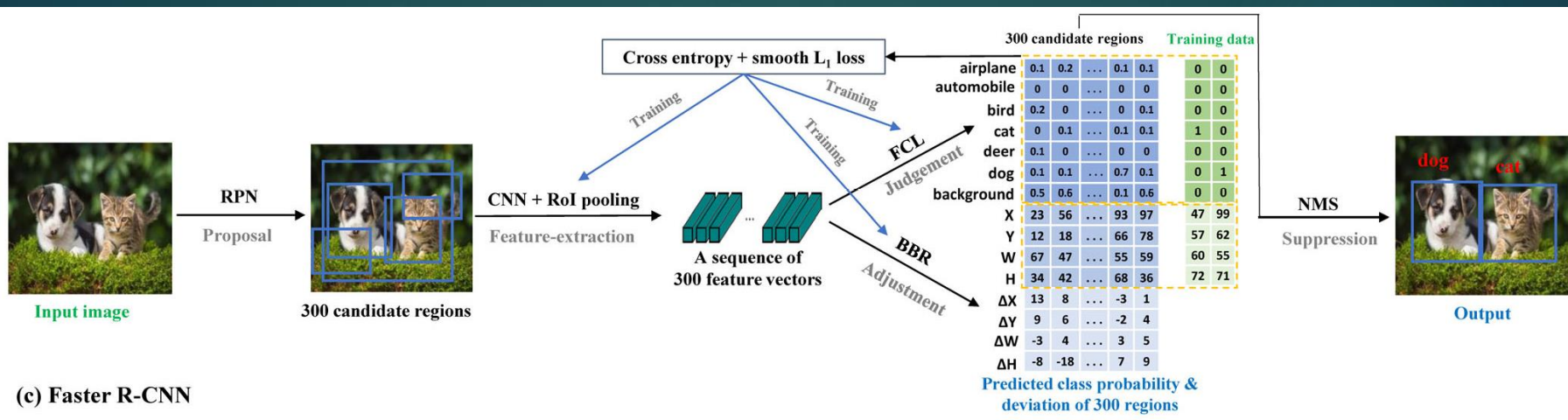
Related Work

Authors	Feature	Year
Y. Zhao, W. Xue, & Q. Li	Multi-scale CRNN for Chinese medical documents	2018
A. Nagarikar, R. S. Dangi, S. K. Maity, & A. Kuvelkar	DL approach using Yolov3 to detect input fields (text & checkbox)	2021
N. V. Rao, A. Sastry, A. Chakravarthy, & P. Kalyanchakravarthi	NN-based methodology for OCR and handwritten character recognition (HCR)	2016
E. Murphy, S. Samuel, J. Cho, & W. Adorno	ML and image processing techniques for checkbox detection on medical sheets to help the health department in African countries	2021
W. Yu, N. Lu, X. Qi, P. Gong, & R. Xiao	PICK, a framework for key information from medical invoices and tax receipts using Graph Learning Convolutional Networks	2021

ROI Detection Methodologies



Different methodologies for object detection



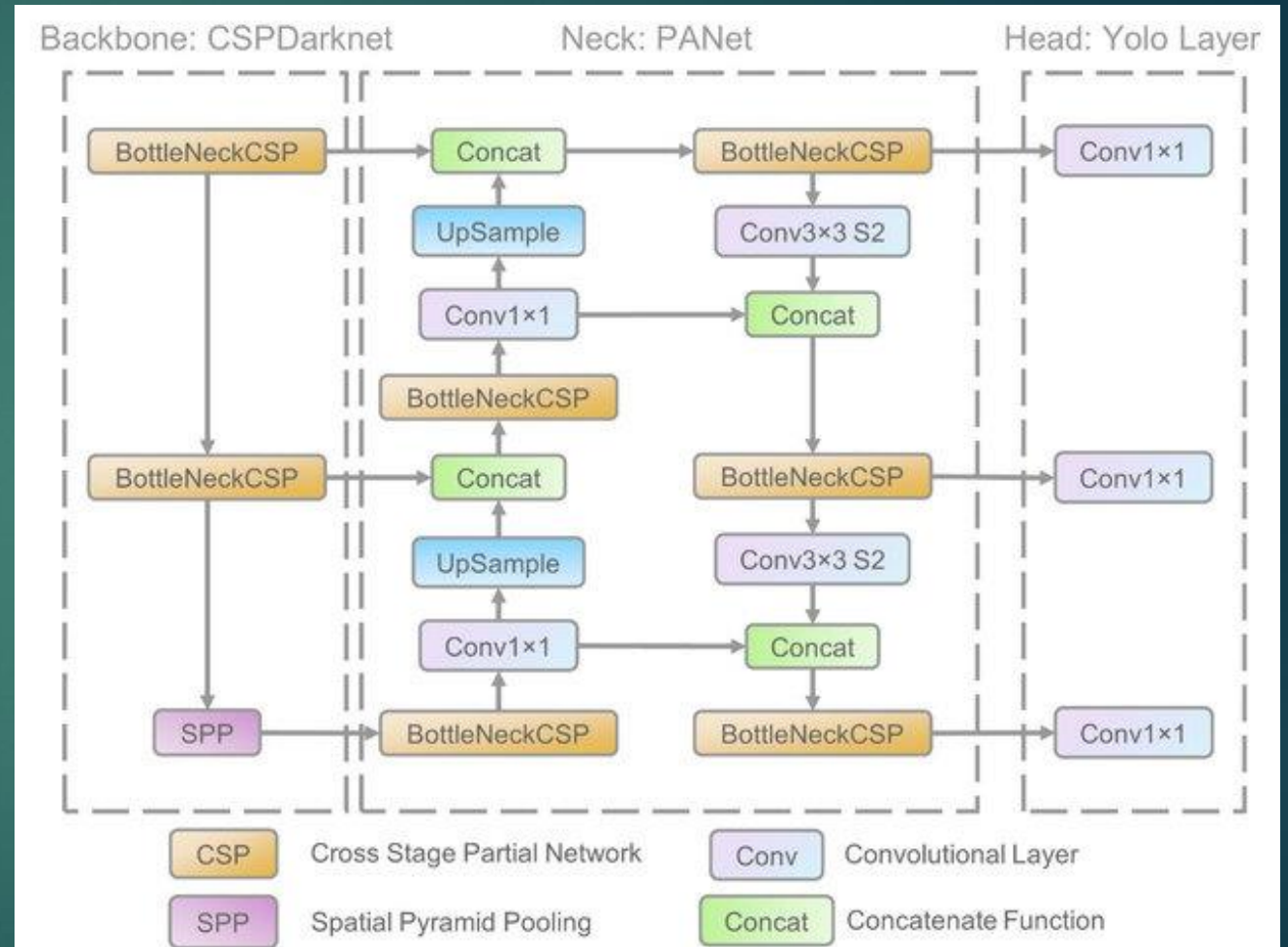
YOLO Model

▶ Yolov5 Detection Model Architecture

- Model backbone: CSPDarknet
- Model Neck: PANet
- Model Head: Yolo Layer
- 270 layers
- Binary cross-entropy
- SGD & Adam optimizer

▶ Model Parameters

- Learning rate
- Input shape



Yolov5 Network Architecture

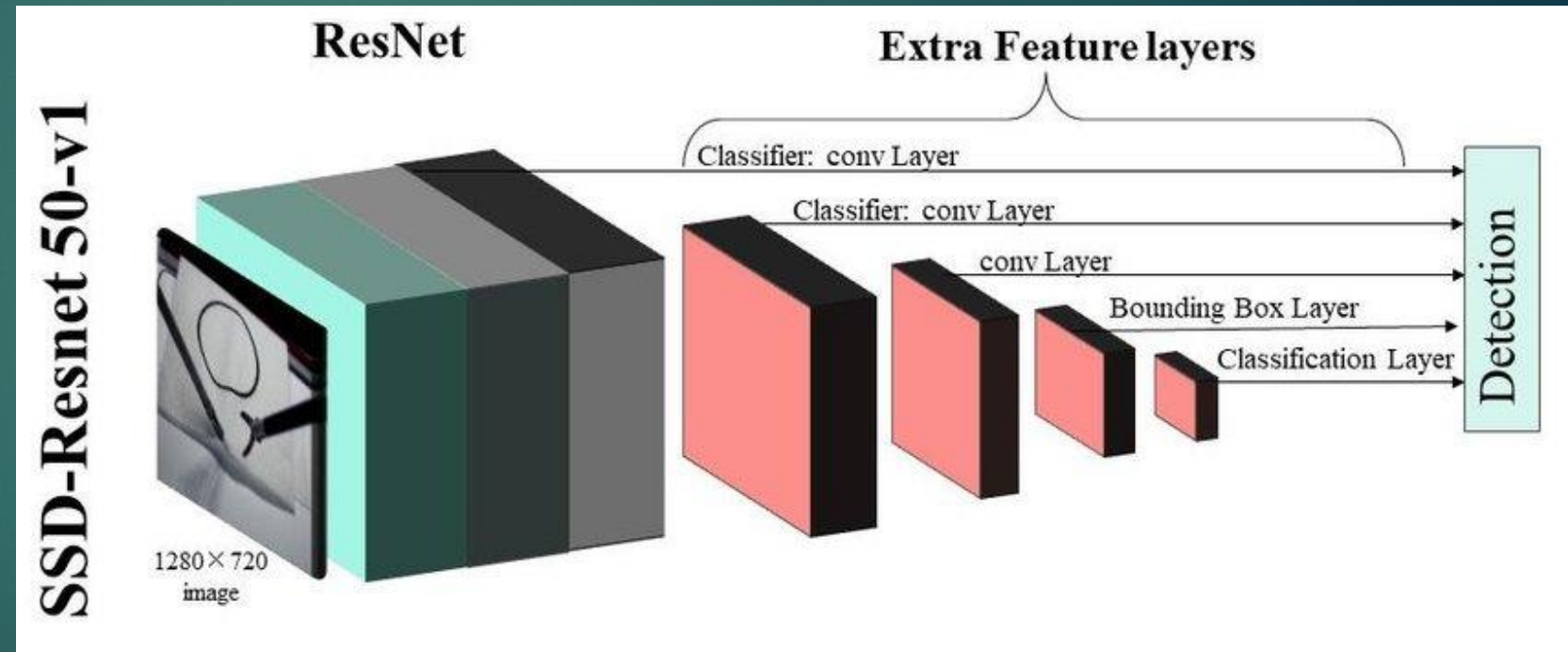
SSD-ResNet50 Model

▶ SSD-ResNet50 Detection Model Architecture

- Backbone network
- Two subnetworks
- Momentum optimizer
- 50 layers
- Binary cross-entropy

▶ Model Parameters

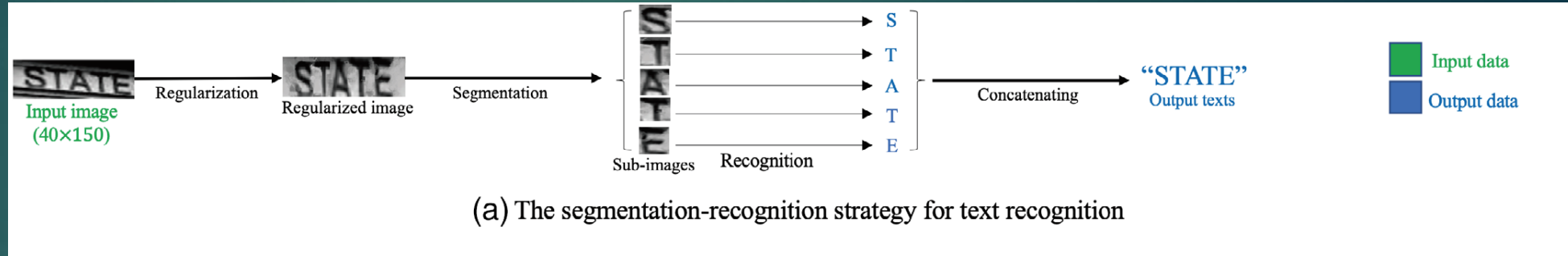
- Learning rate
- Input shape



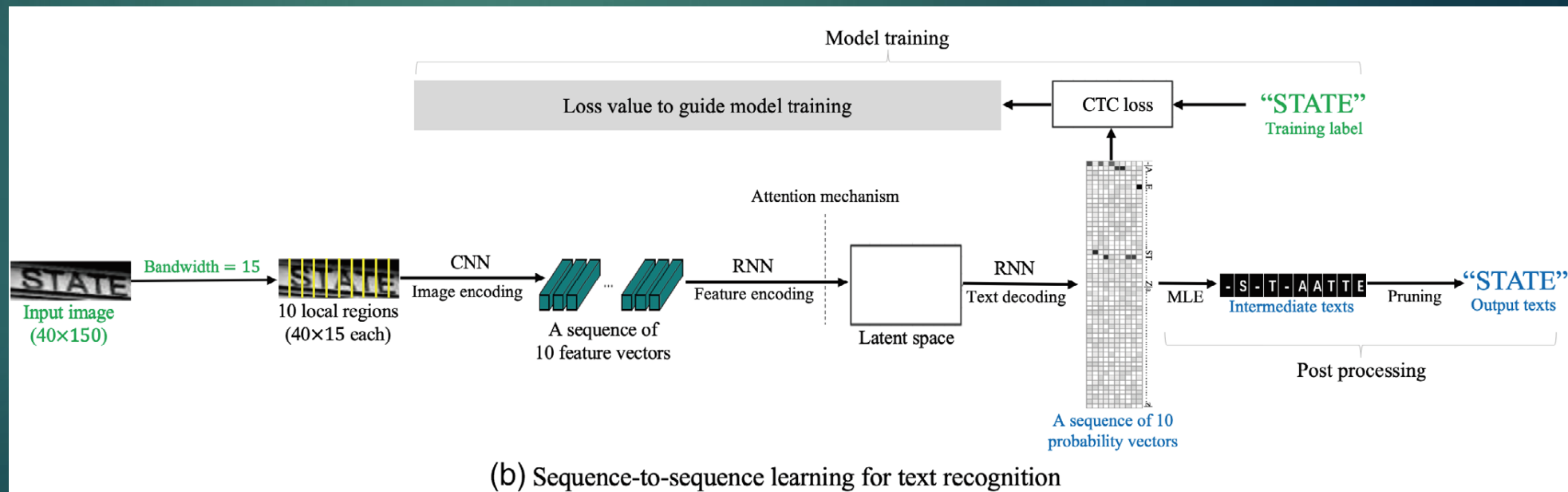
SSD-ResNet50 Network Architecture

OCR Methodologies

Segmentation-Recognition Approach

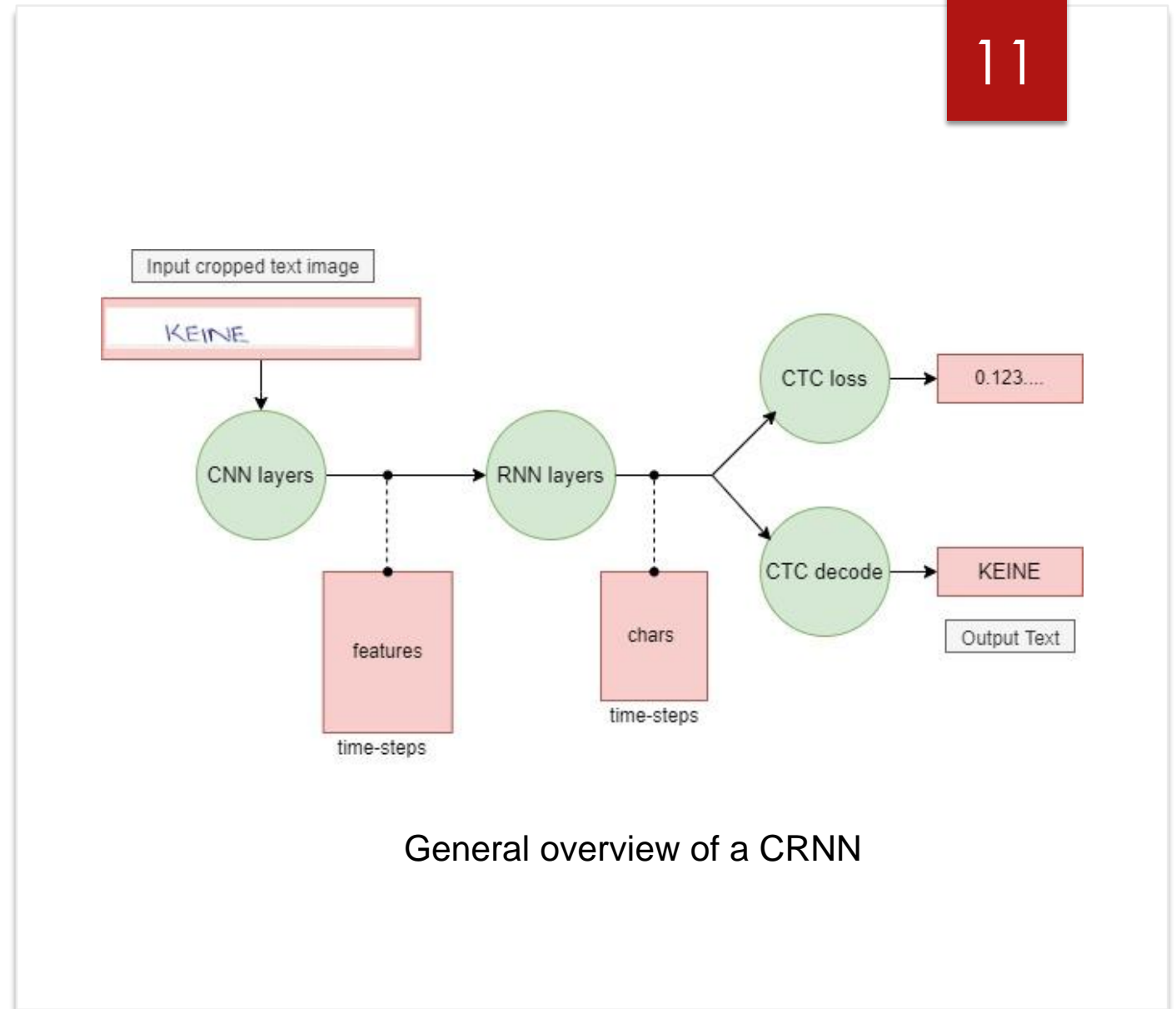


Sequence Learning Approach



Sequence-to-Sequence Learning Model

- ▶ Text recognition using a convolutional recurrent neural network (CRNN) architecture



Methodology

The project has the following major building blocks:

1. Dataset

- Data collection
- Data labeling
- Data preprocessing

2. ROI Detection

- Model training
- Test set results

3. OCR (Text Recognition)

- Model training
- Test set results

17. 28. Möchten Sie diese Patientenverfügung um eine individuelle Freitextangabe ergänzen? Damit können Sie z.B auf individuelle Wertevorstellungen eingehen.
Gegebenenfalls Frage 28-1 ausfüllen

Markieren Sie nur ein Oval.

- Ja (28-1 ausfüllen)
- Nein (28-1 überspringen)

18. 28-1. Bitte geben Sie Ihre individuellen Wertevorstellungen ein:
Nur ausfüllen, wenn Sie bei Frage 28 die Antwort "Ja" ausgewählt haben

YES. I AM FEELING HAPPY
TODAY. I ENJOY THE WEATHER.
IT'S SUMMERS.

19. 29. Möchten Sie eine medizinische Obduktion erlauben?

Markieren Sie nur ein Oval.

- Ja, ich wäre einverstanden
- Nein, ich möchte keinesfalls medizinisch obduziert werden

20. 30. Angenommen Sie leiden an schwerer Demenz (bspw. Alzheimer). Würden Sie den Tod vorziehen?

Markieren Sie nur ein Oval.

- Ja, der Tod wäre mir lieber
- Nein, ich würde weiterleben wollen
- Ich weiß es nicht

26. 35-3. Nachname

Berlinerstadt

27. 35-4. Straße und Hausnummer

Allgemeine

28. 35-5. PLZ

Wohnplatz

29. 35-6. Stadt

Kurz

30. 35-7. Telefonnummer

Allgemeinschafft

31. 35-8. Geburtsdatum und -ort

Wohnraum

32. 35-9. Ausweisnummer

Keine

Dieser Inhalt wurde nicht von Google erstellt und wird von Google auch nicht unterstützt.

Google Formulare

10. 7. Welche der folgenden Medikamente haben Sie in den letzten drei Jahren regelmäßig eingenommen: (Mehrfachauswahl möglich)
Gegebenenfalls Frage 7-1 ausfüllen

Wählen Sie alle zutreffenden Antworten an:

- Krebsmedikament (Chemotherapeutikum)
- Marcumar, Falithrom, Warfarin (Vitamin-K-Antagonisten)
- Plavix
- Aspirin (ASS) Acetylsalicylsäure
- Xarelto, Eliquis, Pradaxa
- Metformin, Insulin, Humalog, Actrapid
- Keine der genannten
- Ich weiß es nicht
- Andere (7-1 ausfüllen)

11. 7-1. Welche anderen Medikamente haben Sie in den 3 letzten Jahren regelmäßig eingenommen?

Nur ausfüllen, wenn Sie bei Frage 7 die Antwort "Andere" ausgewählt haben

Hi all

12. 8. Sind bei Ihnen Allergien bekannt?

Gegebenenfalls Beiblatt 8 ausfüllen

Markieren Sie nur ein Oval.

- Ja, ich weiß es (Beiblatt 8)
- Ja, ich vermute es (Beiblatt 8)
- Nein, mir ist nichts bekannt

13. 12. Gibt es spezifische Umstände, unter denen Sie nicht wiederbelebt werden möchten? (Mehrfachauswahl möglich)

Wählen Sie alle zutreffenden Antworten an:

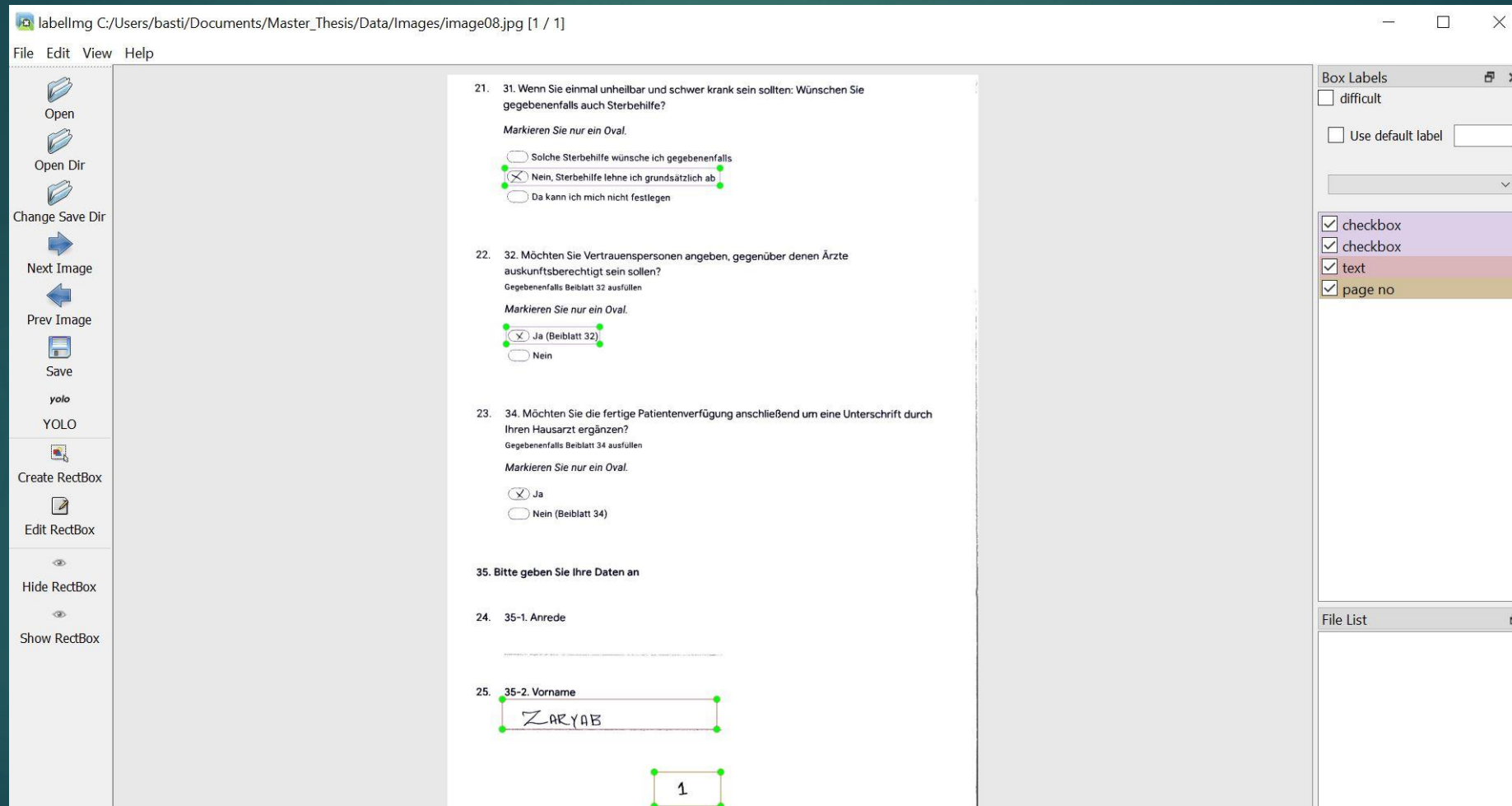
- Bei seit mindestens 10 Minuten bestehendem Herz-Kreislauf Stillstand
- Bei unbeobachtet eingetretenem Herz-Kreislauf Stillstand
- Nichts von alledem

Data Preprocessing

- ▶ Data Preparation
- ▶ Data Annotation Tools
 - **Labelling**
 - Amazon Sagemaker
 - VoTT (Visual Object Tagging Tool)
 - Label Studio
- ▶ Different format annotations (txt, XML, json) as different models architecture accept different input data.

Data Annotation

15



Dataset labeling sample using labellmg

Data Preprocessing

16

21. 31. Wenn Sie einmal unheilbar und schwer krank sein sollten: Wünschen Sie gegebenenfalls auch Sterbehilfe?

Markieren Sie nur ein Oval.

Solche Sterbehilfe wünsche ich gegebenenfalls

Nein, Sterbehilfe lehne ich grundsätzlich ab

Da kann ich mich nicht festlegen

22. 32. Möchten Sie Vertrauenspersonen angeben, gegenüber denen Ärzte auskunftsberechtigt sein sollen?

Gegebenenfalls Beiblatt 32 ausfüllen

Markieren Sie nur ein Oval.

Ja (Beiblatt 32)

Nein

23. 34. Möchten Sie die fertige Patientenverfügung anschließend um eine Unterschrift durch Ihren Hausarzt ergänzen?

Gegebenenfalls Beiblatt 34 ausfüllen

Markieren Sie nur ein Oval.

Ja

Nein (Beiblatt 34)

35. Bitte geben Sie Ihre Daten an

24. 35-1. Anrede

Macht auch Müti

25. 35-2. Vorname

Z

```
- <annotation>
  <folder>Images</folder>
  <filename>image25.jpg</filename>
  <path>C:\Users\basti\Documents\Master_Thesis\Images\image25.jpg</path>
- <source>
  <database>Unknown</database>
</source>
- <size>
  <width>2480</width>
  <height>3507</height>
  <depth>3</depth>
</size>
<segmented>0</segmented>
- <object>
  <name>checkbox</name>
  <pose>Unspecified</pose>
  <truncated>0</truncated>
  <difficult>0</difficult>
- <bndbox>
  <xmin>336</xmin>
  <ymin>439</ymin>
  <xmax>1252</xmax>
  <ymax>507</ymax>
</bndbox>
</object>
- <object>
  <name>checkbox</name>
  <pose>Unspecified</pose>
  <truncated>0</truncated>
  <difficult>0</difficult>
- <bndbox>
  <xmin>340</xmin>
  <ymin>1843</ymin>
  <xmax>508</xmax>
  <ymax>1907</ymax>
</bndbox>
</object>
```

XML file

Data Preprocessing

17

21. 31. Wenn Sie einmal unheilbar und schwer krank sein sollten: Wünschen Sie gegebenenfalls auch Sterbehilfe?

Markieren Sie nur ein Oval.

- Solche Sterbehilfe wünsche ich gegebenenfalls
 Nein, Sterbehilfe lehne ich grundsätzlich ab.
 Da kann ich mich nicht festlegen.

22. 32. Möchten Sie Vertrauenspersonen angeben, gegenüber denen Ärzte auskunftsberechtigt sein sollen?

Gegebenenfalls Beiblatt 32 ausfüllen

Markieren Sie nur ein Oval.

- Ja (Beiblatt 32)
 Nein

23. 34. Möchten Sie die fertige Patientenverfügung anschließend um eine Unterschrift durch Ihren Hausarzt ergänzen?

Gegebenenfalls Beiblatt 34 ausfüllen

Markieren Sie nur ein Oval.

- Ja
 Nein (Beiblatt 34)

35. Bitte geben Sie Ihre Daten an

24. 35-1. Anrede

Frau

25. 35-2. Vorname

James

image110 - Notepad

File Edit Format View Help

```
0 0.322984 0.157827 0.387903 0.023097
0 0.210685 0.370545 0.172177 0.023097
1 0.301411 0.778015 0.353629 0.032792
1 0.301210 0.872968 0.350806 0.033932
```

txt file

```
{"annotation": {"folder": "Images", "filename": "image110.jpg", "path": "C:\\Users\\basti\\Documents\\Master_Thesis\\Data\\Images\\image110.jpg", "source": {"database": "Unknown"}, "size": {"width": "2480", "height": "3507", "depth": "3"}, "segmented": "0", "object": [{"name": "checkbox", "pose": "Unspecified", "truncated": "0", "difficult": "0", "bndbox": {"xmin": "305", "ymin": "497", "xmax": "447", "ymax": "613"}}, {"name": "checkbox", "pose": "Unspecified", "truncated": "0", "difficult": "0", "bndbox": {"xmin": "320", "ymin": "2005", "xmax": "432", "ymax": "2097"}}, {"name": "text", "pose": "Unspecified", "truncated": "0", "difficult": "0", "bndbox": {"xmin": "401", "ymin": "2659", "xmax": "736", "ymax": "2870"}}, {"name": "text", "pose": "Unspecified", "truncated": "0", "difficult": "0", "bndbox": {"xmin": "405", "ymin": "2978", "xmax": "743", "ymax": "3155"}}]}
```

json file

Data Preprocessing

▶ CSV file initially generated

filename	width	height	name	xmin	ymin	xmax	ymin	ymax
image01.jpg	1071	1500	checkbox	126	160	188	212	
image01.jpg	1071	1500	text	201	406	647	487	
image01.jpg	1071	1500	checkbox	115	776	183	828	
image01.jpg	1071	1500	checkbox	115	1126	172	1169	
image02.jpg	1065	1500	checkbox	122	144	175	190	
image02.jpg	1065	1500	checkbox	125	508	168	545	
image02.jpg	1065	1500	checkbox	112	826	170	865	
image02.jpg	1065	1500	text	124	1119	261	1193	
image03.jpg	1050	1500	checkbox	95	360	147	396	
image03.jpg	1050	1500	checkbox	97	690	149	733	
image03.jpg	1050	1500	text	138	969	934	1062	
image04.jpg	1077	1500	checkbox	117	157	176	195	
image04.jpg	1077	1500	checkbox	118	641	153	675	
image04.jpg	1077	1500	checkbox	118	777	154	806	
image04.jpg	1077	1500	checkbox	120	1185	171	1224	
image05.jpg	2113	3000	checkbox	223	324	323	404	

A section of the prepared dataset CSV file

Data Preprocessing

- ▶ Ground truth file prepared

```
*transcription - Notepad
File Edit Format View Help
image814 D-o|y-o-u|e-v-e-r|h-a-v-e|p-a-i-n|i-n|y-o-u-r|c-h-e-s-t|o-r|h-e-a-r-t
image818 T-a-b|S-i-l-v-e-r|C-e-n-t-r-e-n
image819 A-g-e|a-t|d-e-a-t-h
image820 N-e-v-e-r
image823 D-e-c-e-a-s-e-d|
image824 O-f-t-e-n
image825 N-o
image828 C-a-n-t|r-e-m-e-m-b-e-r
image843 B-ö-s-a-r-t-i-g-e-r|G-e-h-i-r-n-s_mi-T-u-m-o-r
image836 C-o-n-g-e-n-i-t-a-l|h-e-a-r-t|d-i-s-e-a-s-e|e-x-i-s-t-i-n-g|a-t|b-i-r-t-h|b-u-t|n-o-t|h-e-r-e-d-i-t-a-r-y
image837 G-o-o-d
image848 D-i-a-b-e-t-e-s
image849 N-o
image857 N-e-i-n-s_cm|i-c-h|w-ü-r-d-e|i-m-m-e-r|u-n-d|u-n-t-e-r|a-l-l-e-n|U-m-s-t-ä-n-d-e-n|l-e-b-e-n|w-o-l-l-e-n-s_pt
image862 H-e-r-z
image872 M-a-r-c-u-m-a-r-s_cm|F-a-l-i-t-h-r-o-m-s_cm|W-a-r-f-a-r-i-n|V-i-t-a-m-i-n-s_mi-K-s_mi-A-n-t-a-g-o-n-i-s-t-e-n
image882 I-m|H-o-s-p-i-z
image888 S-t-r-o-k-e
image889 C-a-n-c-e-r
image942 I-m|K-r-a-n-k-e-n-h-a-u-s
```

Ground truth file for character recognition models

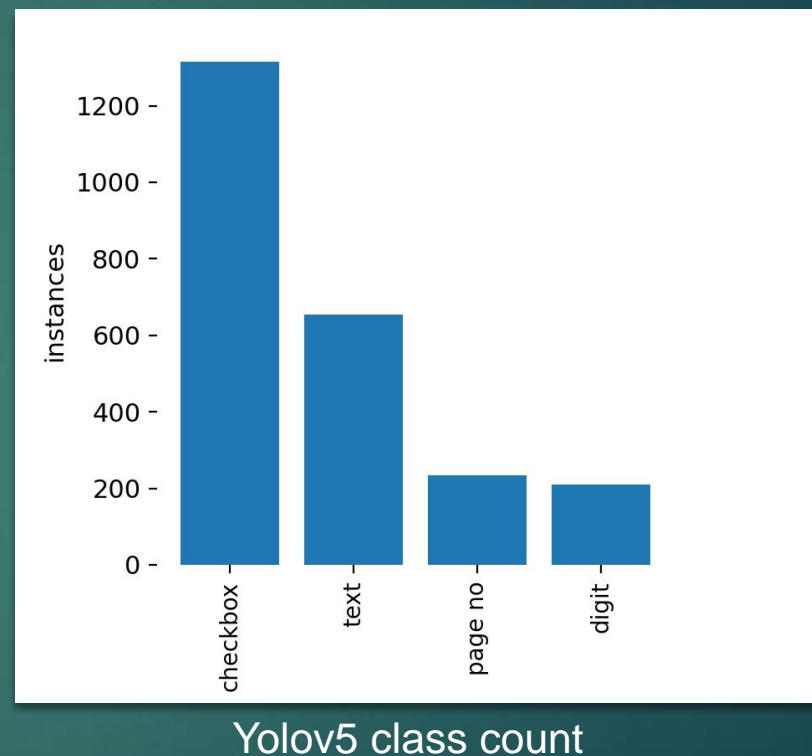
Data in a Glance



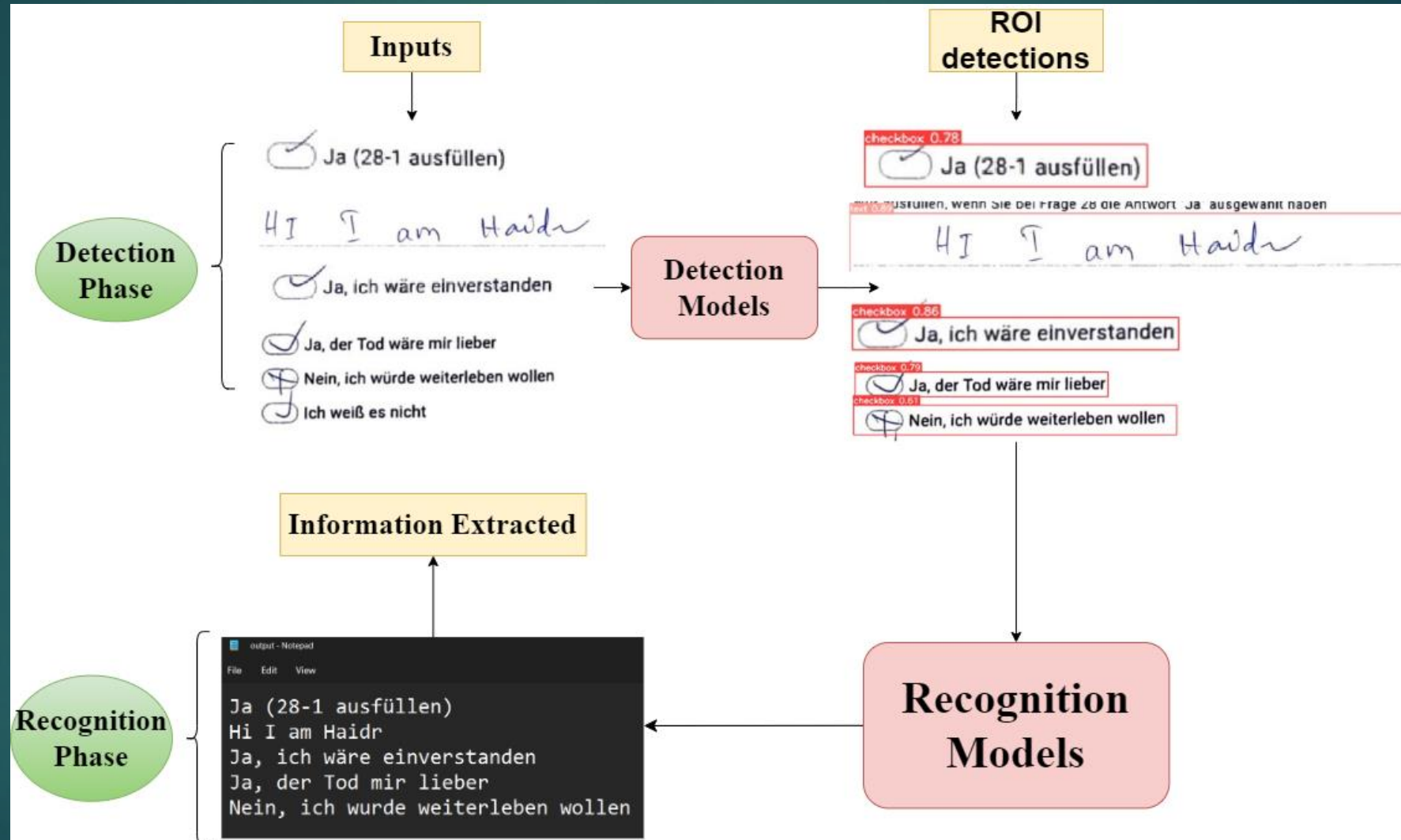
1,050 scanned
medical forms



Yolov5 training
classes



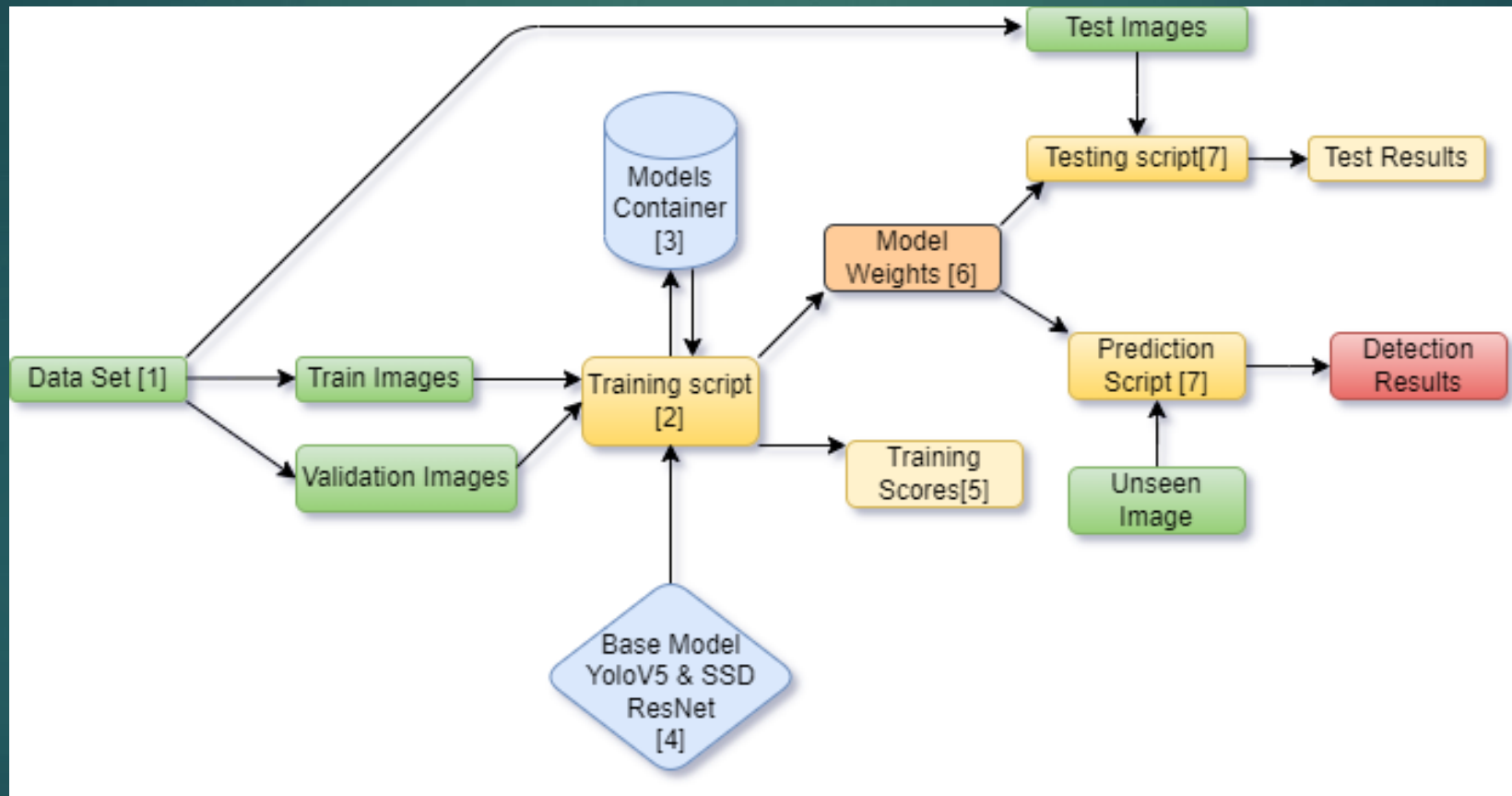
Methodology



General workflow of the proposed system

ROI Detection Methodology

22

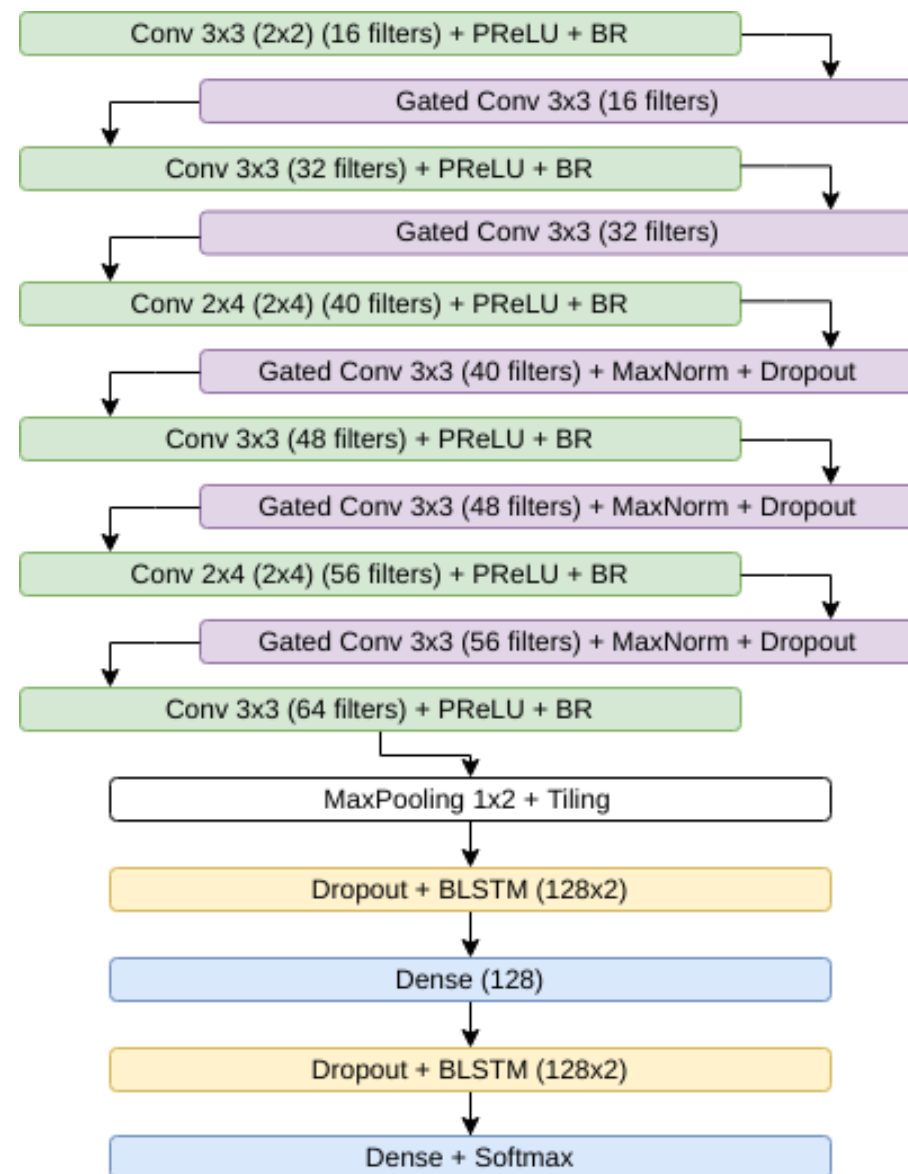


ROI detection workflow

OCR Methodology

OCR with Gated-
CNN-BLSTM
model

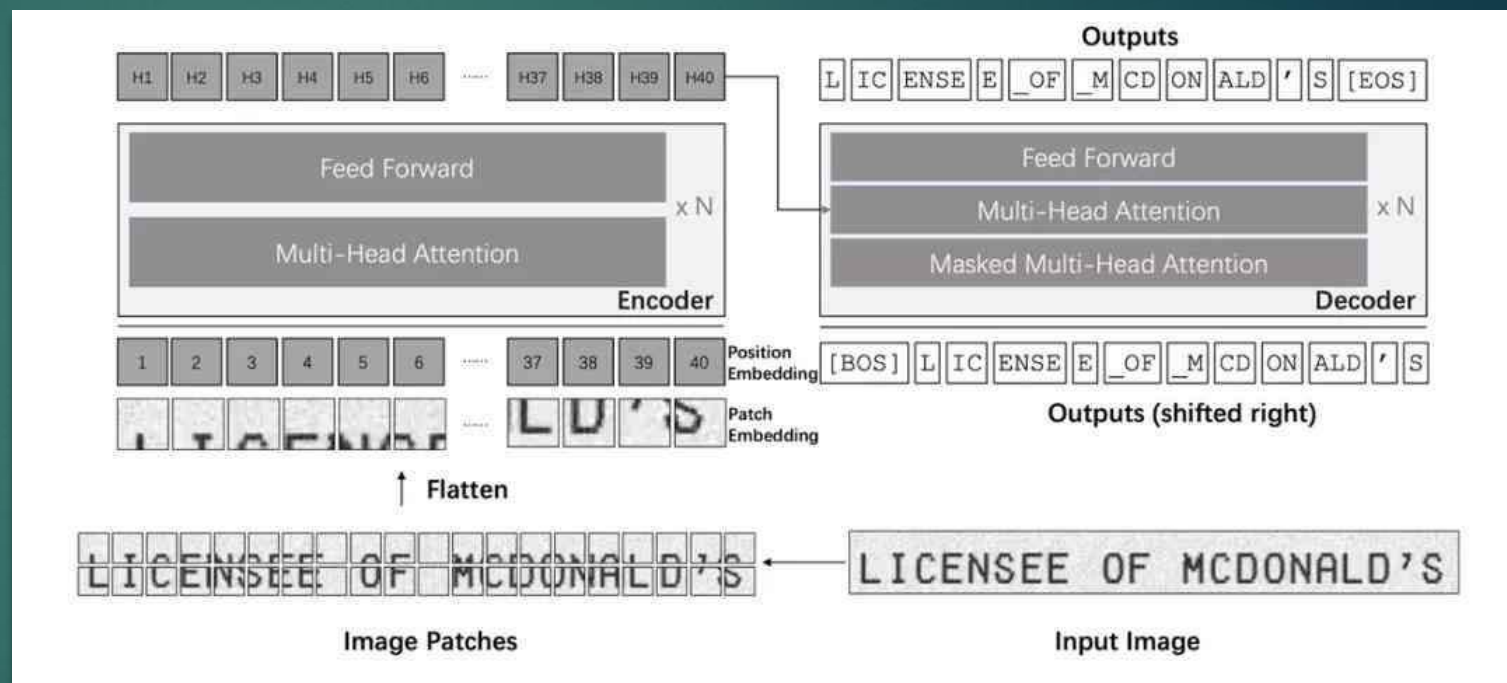
23



Gated-CNN-BLSTM model architecture

Transformer OCR Model

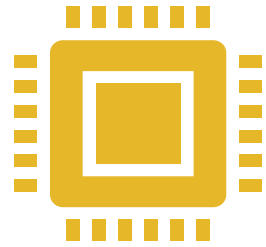
- ▶ Encoder-decoder architecture
- ▶ Transformers-based optical character recognition model [TrOCR](#)



TrOCR model architecture

Training and Experimental Configurations

25



Software Dependencies

Python 3.7
TensorFlow 2.3.1
PyTorch 1.10.0
NumPy v1.18.5



Hardware Dependencies

GPU	Google Colab
RAM	16 GB
OS	Linux

Training Hyperparameters Settings

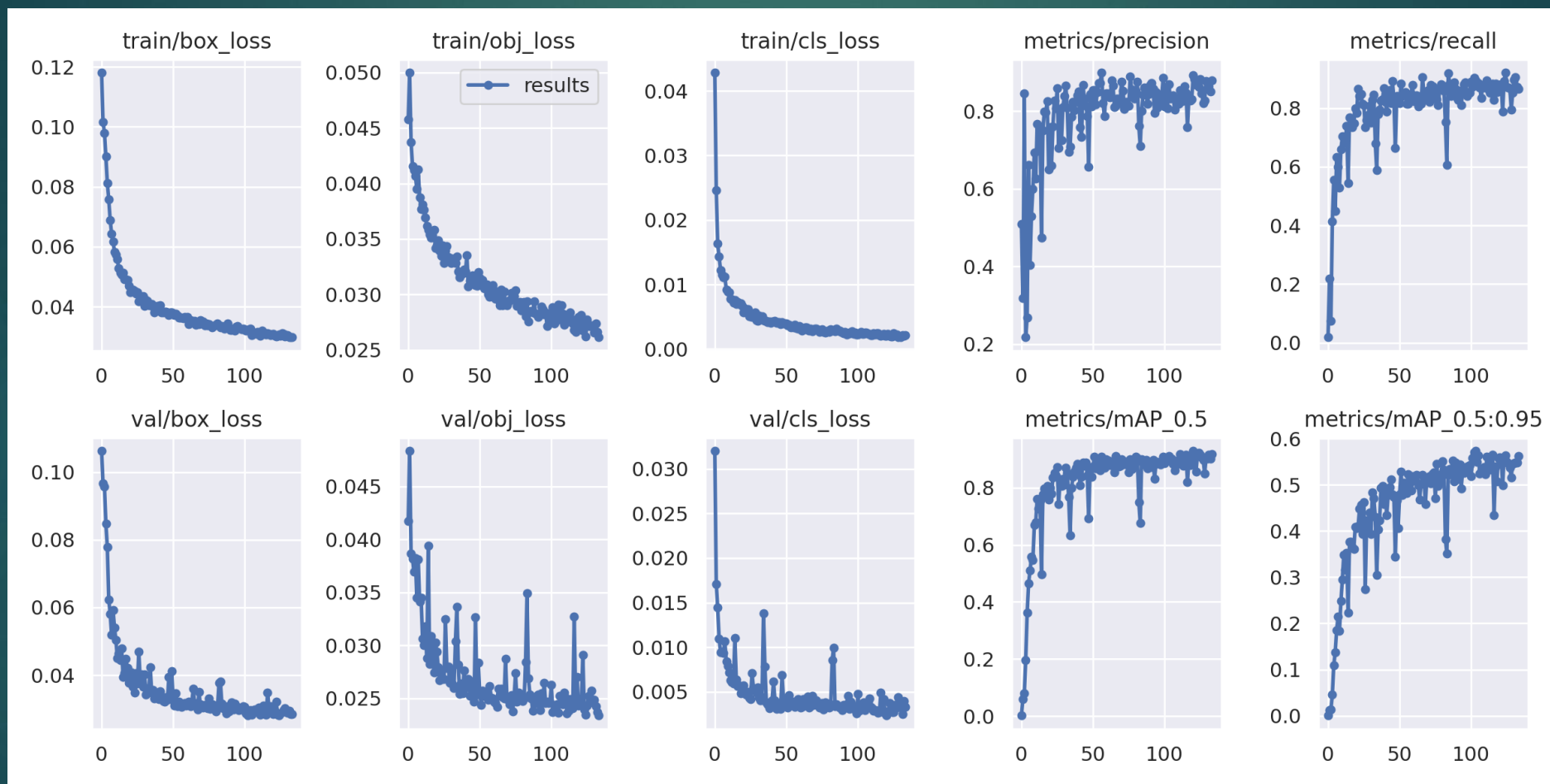
26

Model	Optimizer	Loss Function	Activation	Batch Size	Learning rate	Epochs
ROI detection model 1: Yolov5	SGD	Cross entropy	Sigmoid	16	1e-2	150
ROI detection model 2: SSD-ResNet	Momentum	Cross entropy	ReLU	16	1e-2-3e-2	150
OCR model 1: Gated-CNN-BLSTM	RMSProp	CTC	PReLU	32	1e-2	240
OCR model 2: Transformers	Adam	Cross entropy	GeLU	8	5e-2	100

Results

ROI Detection Model 1: Yolov5 Results

28



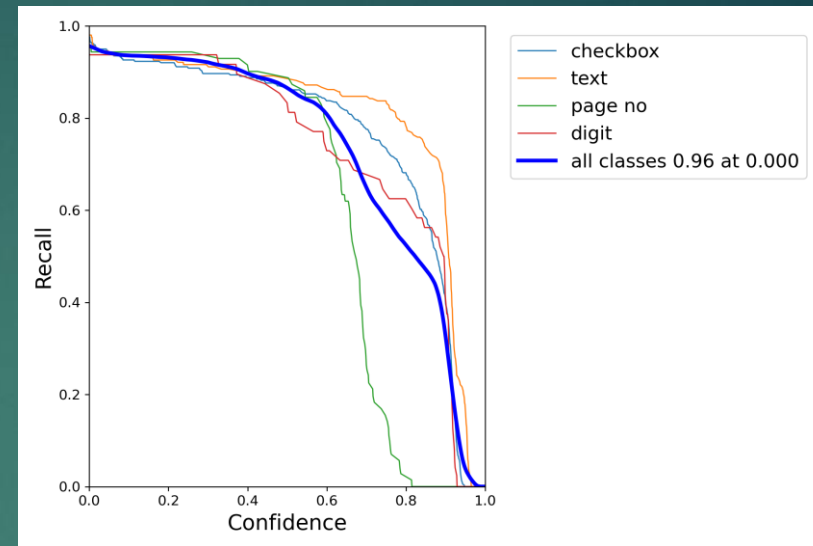
Yolov5 training results

ROI Detection Model 1: Yolov5 Results

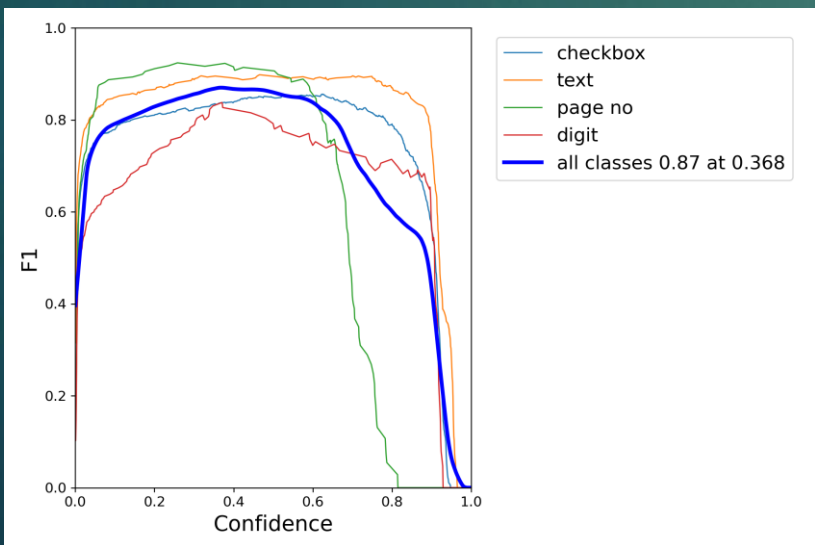
29



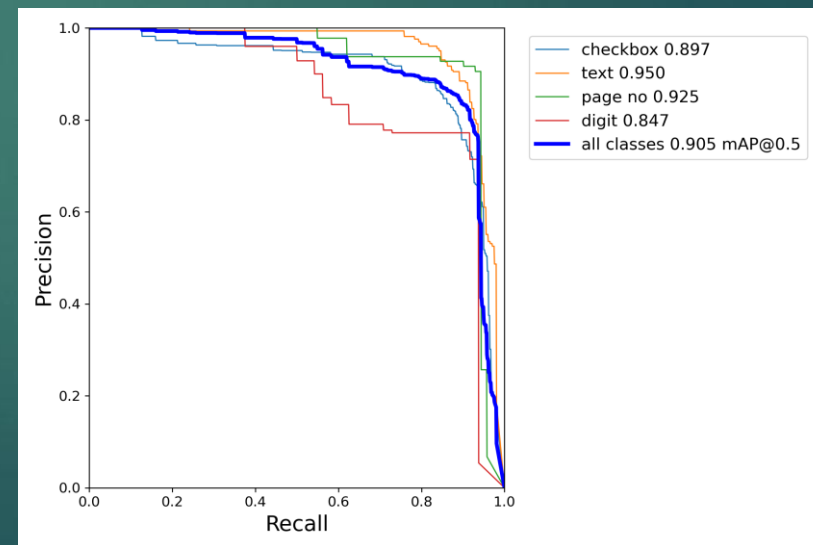
Precision curve



Recall curve



F1 curve



PR curve

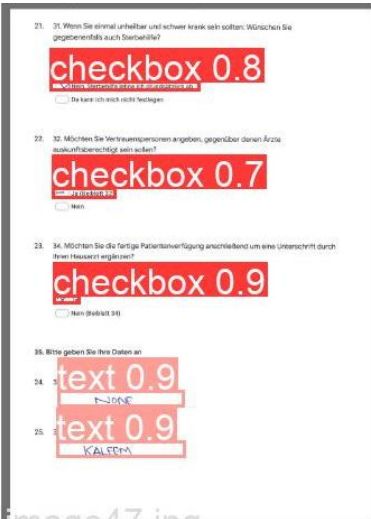


image47.jpg

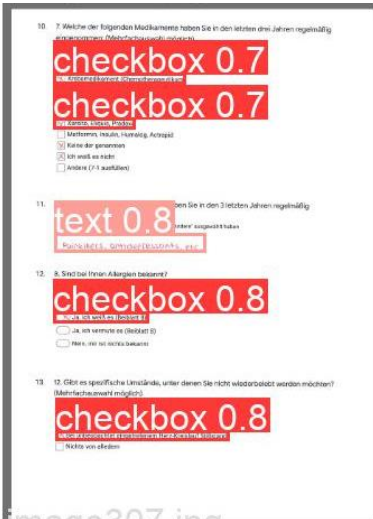


image307.jpg

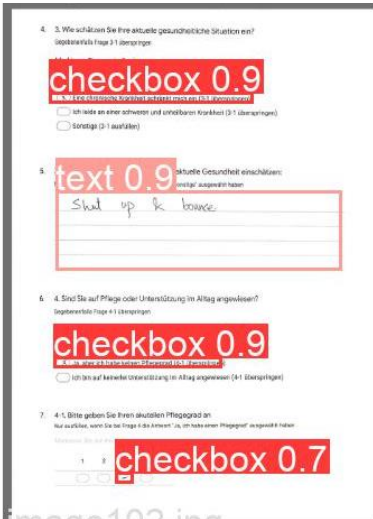


image102.jpg

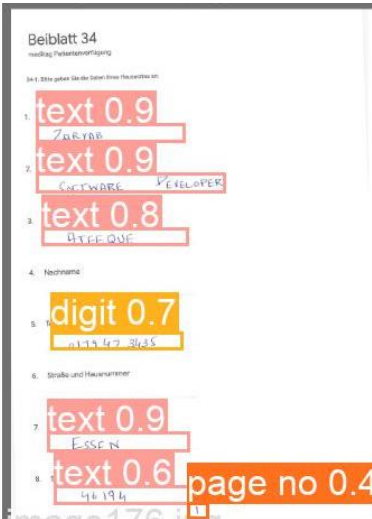


image176.jpg

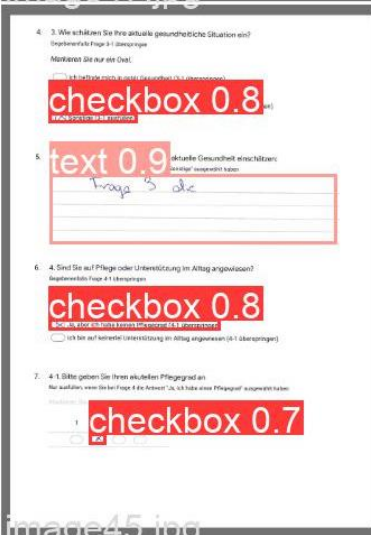


image45.jpg

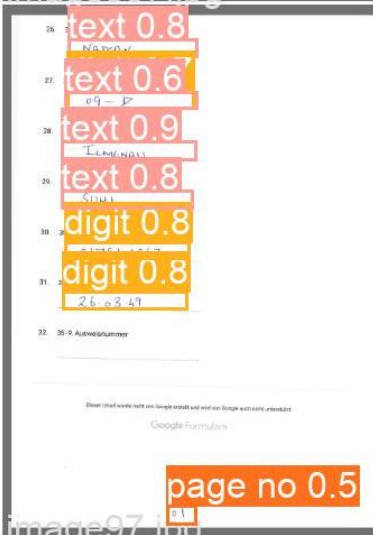


image07.jpg

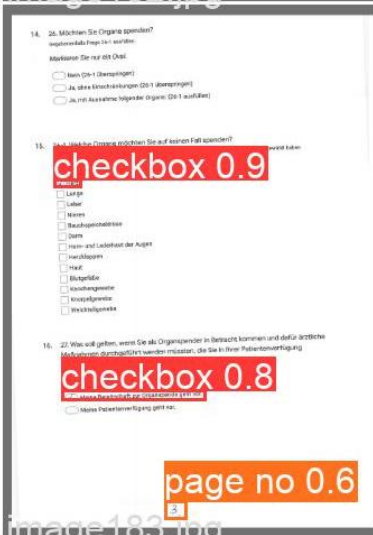


image183.jpg

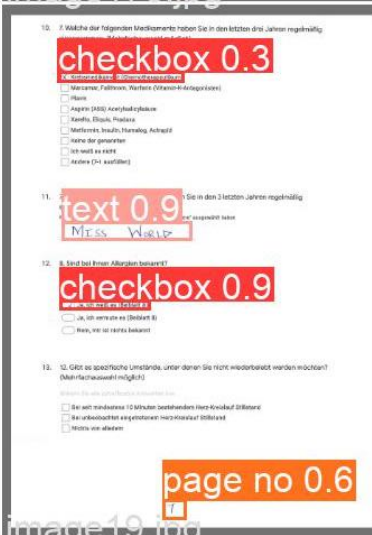


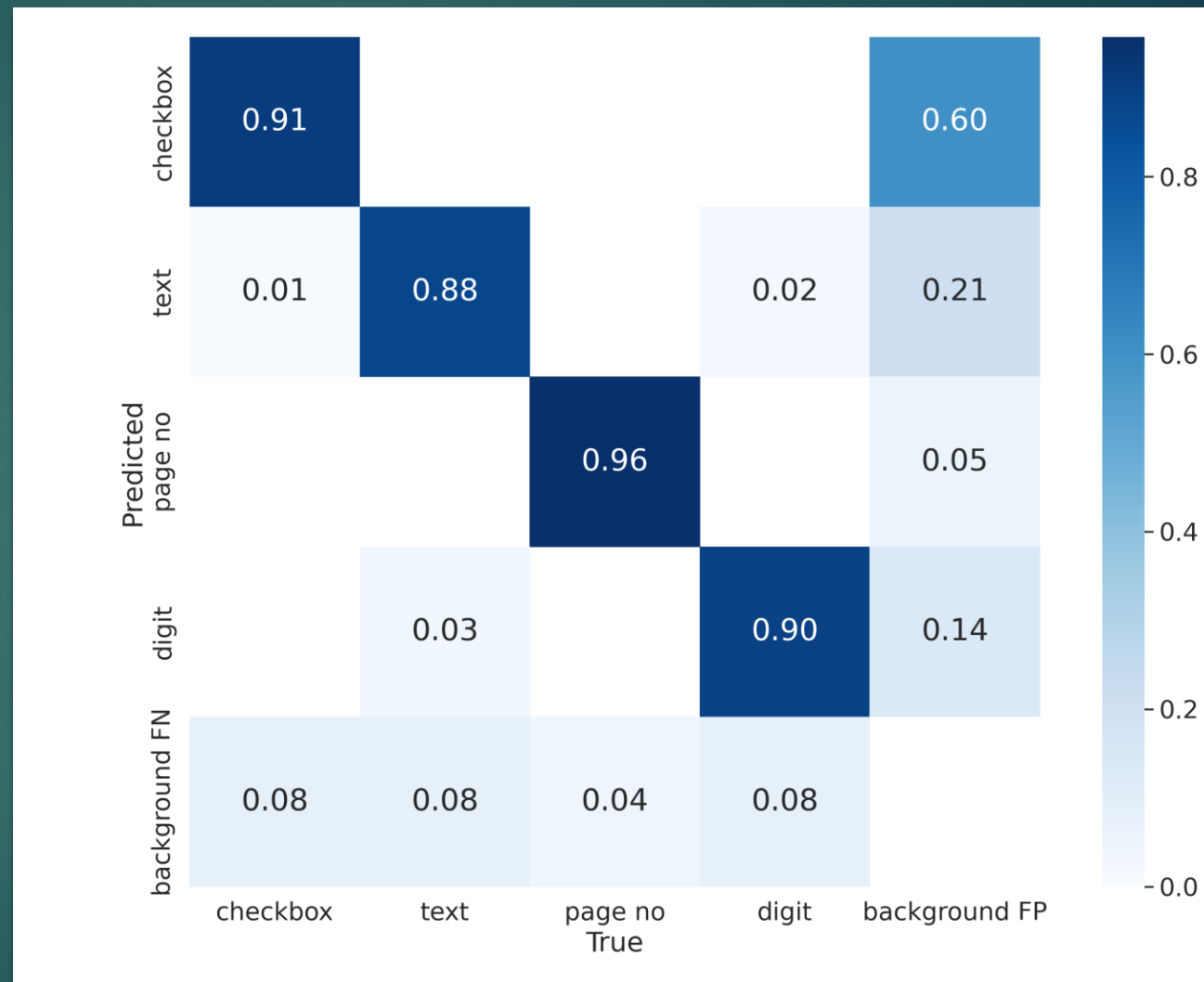
image19.jpg

Yolov5 Detection Results

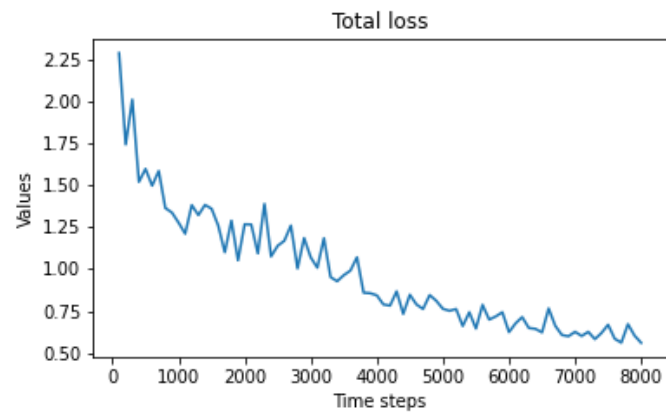
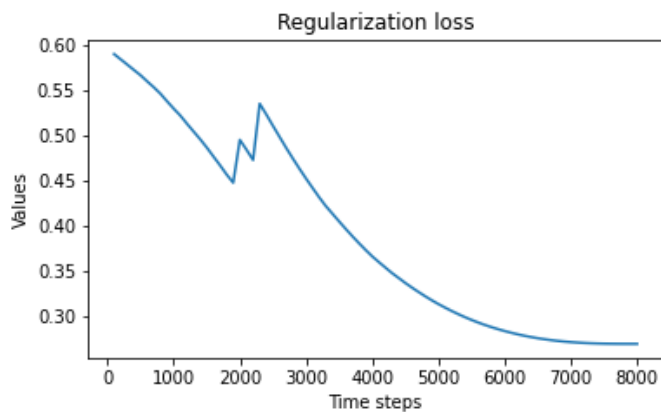
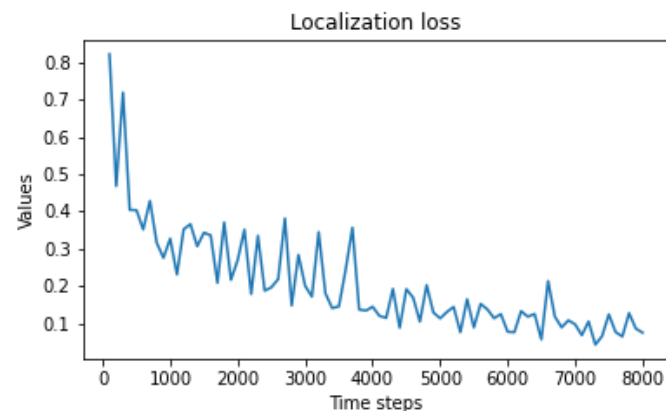
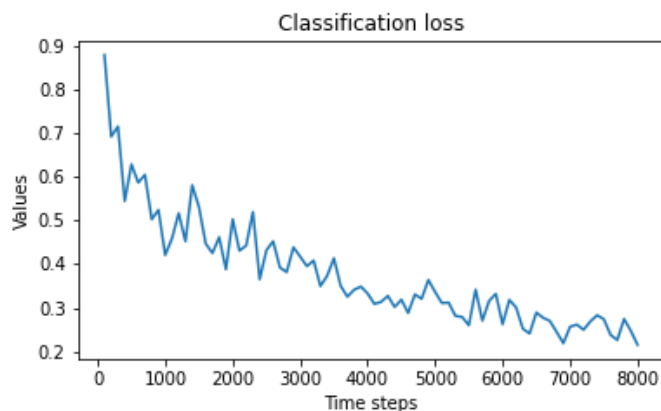
ROI detection results

Yolov5 Evaluation Matrix

Model	Precision	Recall	mAP	F1
Yolov5	0.878	0.866	0.918	0.894



Confusion Matrix



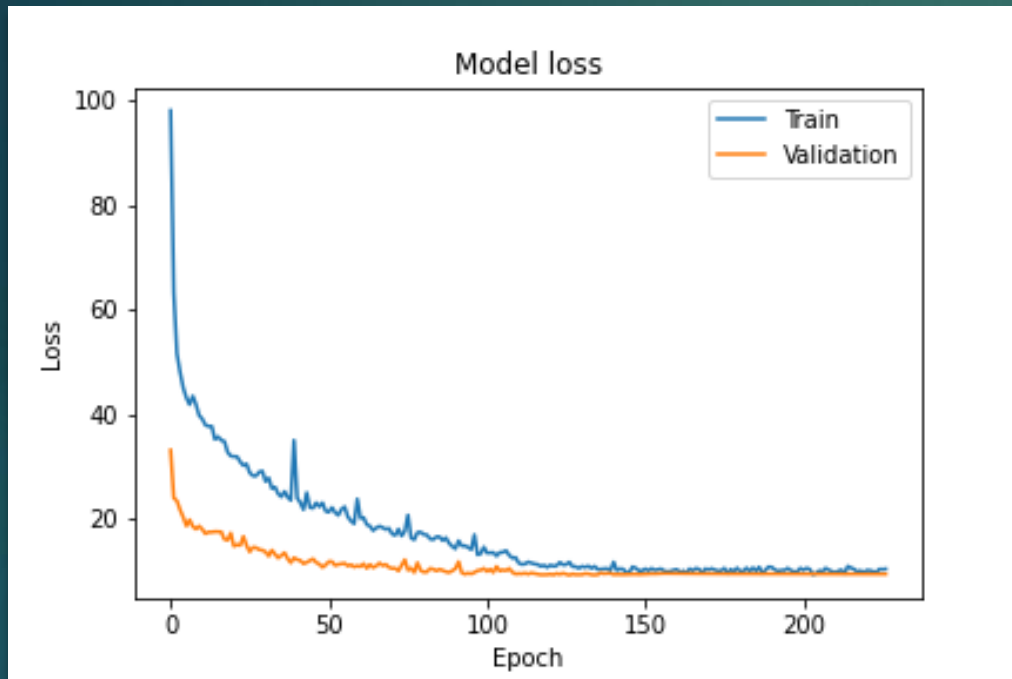
ROI Detection Model 2: SSD- ResNet50 Results

Model	Precision	Recall	mAP	F1
SSD-ResNet	0.817	0.805	0.845	0.831

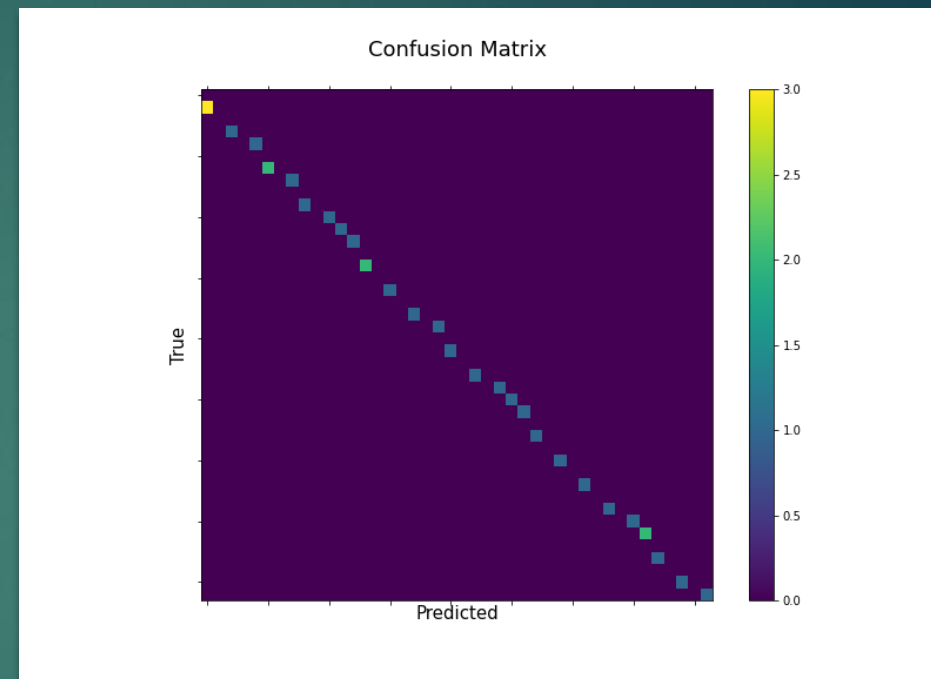
SSD-ResNet50 training results

OCR Model 1: Gated-CNN-BLSTM Results

33



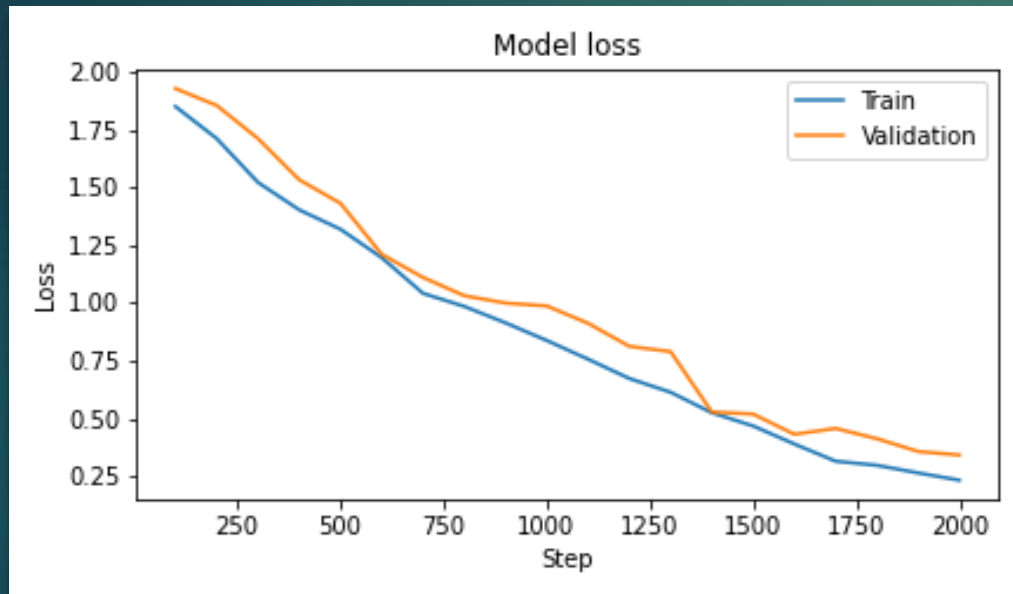
Gated-CNN-BLSTM training results



Confusion Matrix

Model	Parameters	CER	WER	SER
Gated-CNN-BLSTM	820k	0.09	0.13	0.19

OCR Model 2: Transformers Results



TrOCR training results

Model	Parameters	CER	WER	SER
TrOCR	5M	0.06	0.10	0.13

Results Comparison

ROI Detection Models

Model	Precision	Recall	mAP	F1
Model 1: Yolov5	0.878	0.866	0.918	0.894
Model 2: SSD-ResNet	0.817	0.805	0.845	0.831

Yolov5 performed better and ROI is cropped for OCR

OCR Models

Model	Parameters	CER	WER	SER
Model 1: Gated-CNN-BLSTM	820k	0.09	0.13	0.19
Model 2: TrOCR	5M	0.06	0.10	0.13

Transformers have less CER, but Gated-CNN-BLSTM has very less parameters

Results

36

17. 28. Möchten Sie diese Patientenverfügung um eine individuelle Freitextangabe ergänzen?
Damit können Sie z.B auf individuelle Wertevorstellungen eingehen.
Gegebenenfalls Frage 28-1 ausfüllen

Markieren Sie nur ein Oval.

- Ja (28-1 ausfüllen)
 Nein (28-1 überspringen)

18. 28-1. Bitte geben Sie Ihre individuellen Wertevorstellungen ein:
Nur ausfüllen, wenn Sie bei Frage 28 die Antwort "Ja" ausgewählt haben

HI I am Haidr

19. 29. Möchten Sie eine medizinische Obduktion erlauben?

Markieren Sie nur ein Oval.

- Ja, ich wäre einverstanden
 Nein, ich möchte keinesfalls medizinisch obduziert werden

20. 30. Angenommen Sie leiden an schwerer Demenz (bspw. Alzheimer). Würden Sie den Tod vorziehen?

Markieren Sie nur ein Oval.

- Ja, der Tod wäre mir lieber
 Nein, ich würde weiterleben wollen
 Ich weiß es nicht

17. 28. Möchten Sie diese Patientenverfügung um eine individuelle Freitextangabe ergänzen?
Damit können Sie z.B auf individuelle Wertevorstellungen eingehen.
Gegebenenfalls Frage 28-1 ausfüllen

Markieren Sie nur ein Oval.

- Ja (28-1 ausfüllen)
 Nein (28-1 überspringen)

18. 28-1. Bitte geben Sie Ihre individuellen Wertevorstellungen ein:
Nur ausfüllen, wenn Sie bei Frage 28 die Antwort "Ja" ausgewählt haben

HI I am Haidr

19. 29. Möchten Sie eine medizinische Obduktion erlauben?

Markieren Sie nur ein Oval.

- Ja, ich wäre einverstanden
 Nein, ich möchte keinesfalls medizinisch obduziert werden

20. 30. Angenommen Sie leiden an schwerer Demenz (bspw. Alzheimer). Würden Sie den Tod vorziehen?

Markieren Sie nur ein Oval.

- Ja, der Tod wäre mir lieber
 Nein, ich würde weiterleben wollen
 Ich weiß es nicht

Ja (28-1 ausfüllen)

HI I am Haidr

Ja, ich wäre einverstanden

Ja, der Tod wäre mir lieber

ROI cropped

output - Notepad
File Edit Format View Help
Ja (28-1 ausfüllen)
Hi I am Haidr
Ja, ich wäre einverstanden
Ja, der Tod wäre mir lieber

OCR output

Raw data sample

ROI detected document

Discussion and Future Work

Conclusion:

- ▶ ROI detection on 1,050 scanned medical documents and character recognition performed on 10,800 English and 3,800 German words.
- ▶ Yolov5 and SSD-ResNet models implemented for ROI detection. Yolo performed better and opted for the next step.
- ▶ Gated-CNN-BLSTM and TrOCR models are used to perform OCR.
- ▶ Yolov5 and Gated-CNN-BLSTM model combination is proposed for document digitization.

Future Work:

- ▶ Research may be utilized as an application for medical data digitization.
- ▶ Dataset increase and test it on other detection models.
- ▶ Hyperparameter tuning to reduce CER.
- ▶ OCR dictionary could be expanded and trained to recognize other Roman characters



Thank You

Questions?

References

- ▶ Yulei Zhao, Wenyuan Xue, and Qingyong Li, "A Multi-scale CRNN Model for Chinese Papery Medical Document Recognition," in IEEE Fourth International Conference on Multimedia Big Data (BigMM), 2018.
- ▶ Nagarikar, A., Dangi, R. S., Maity, S. K., Kuvelkar, A., & Wandhekar, S. (2021). Input Fields Recognition in Documents Using Deep Learning Techniques. REVISTA GEINTEC-GESTAO INOVACAO E TECNOLOGIAS, 11(4), 4405-4415.
- ▶ Rao, N. V., Sastry, A. S. C. S., Chakravarthy, A. S. N., & Kalyanchakravarthi, P. (2016). OPTICAL CHARACTER RECOGNITION TECHNIQUE ALGORITHMS. Journal of Theoretical & Applied Information Technology, 83(2).
- ▶ Emily Murphy, Swathi Samuel, Joseph Cho, William Adorno, Marcel Durieux, Donald Brown, and Christian Ndaribitse, "Checkbox Detection on Rwandan Perioperative Flowsheets Using Convolutional Neural Network," in 2021 Systems and Information Engineering Design Symposium (SIEDS), DOI: 10.1109/SIEDS52267.2021.9483723, 2021, pp. 1-6.
- ▶ Yu, W., Lu, N., Qi, X., Gong, P., & Xiao, R. (2021, January). Pick Processing key information extraction from documents using improved graph learning-convolutional networks. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 4363-4370). IEEE.
- ▶ Xu, R., Lin, H., Lu, K., Cao, L., & Liu, Y. (2021). A Forest Fire Detection System Based on Ensemble Learning. Forests, 12(2), 217.