

# **Segmentierung des Magen-Darm-Trakts anhand von MRT Bildern mithilfe von Deep Learning**

**Viktor Leonardo Krawutschke**

**Bachelorarbeit**

Beginn der Arbeit:	01. April 2010
Abgabe der Arbeit:	25. September 2022
Gutachter:	Univ.-Prof. Dr. S. Conrad Univ.-Prof. Dr. M. Kollmann



## **Erklärung**

Hiermit versichere ich, dass ich diese Bachelorarbeit selbstständig verfasst habe. Ich habe dazu keine anderen als die angegebenen Quellen und Hilfsmittel verwendet.

Düsseldorf, den 25. September 2022

---

Viktor Leonardo Krawutschke

## Zusammenfassung

Diese Bachelorarbeit beschäftigt sich mit der semantischen Segmentierung von MRT Bildern im medizinischen Bereich. Im Jahr 2015 wurde bei 5 Millionen Menschen Krebs im Gastro-Intestinaltrakt festgestellt, wovon die Hälfte für eine Strahlentherapie in Frage kommen. Diese dauert in der Regel 1 bis 6 Wochen und findet täglich für 15 Minuten statt. Dabei soll den Krebszellen gezielt eine hohe Strahlendosis verabreicht werden, gleichzeitig ist es zu vermeiden nicht betroffene Organe dabei in Mitleidenschaft zu ziehen. Problematisch ist, dass die Position der Organe und somit auch der Tumore von Tag zu Tag unterschiedlich ist. Dank Magnetresonanztomographen (MRT) und Linearbeschleunigern können Onkologen jedoch die genaue Position der Organe feststellen. Dies ist jedoch ein Zeitintensiver Prozess und dauert je nach Patienten zwischen 15 und 60 Minuten. Deep Learning könnte dabei helfen, diesen Prozess erheblich zu beschleunigen, um so Patienten zu entlasten und Ärzten die Chance zu geben mehreren Menschen im selben Zeitraum ihre Hilfe anzubieten.

Das UW-Madison Carbone Cancer Center hat im Rahmen einer Challenge auf Kaggle.com anonymisierte Scans von über 150 Patienten zur Verfügung gestellt, in denen manuell der Dickdarm, Dünndarm und der Magen von Fachleuten segmentiert worden ist. Ziel ist es, ein Modell zu entwerfen welches gegeben den Trainingsdaten dazu in der Lage ist in Bildausschnitten von vorher ungesehenen Patienten die Lokalität der genannten Organe in Form von einer Maske zu bestimmen. Baseline dieser Arbeit bildet ein Convolutional Neural Network, einem UNet Modell, welches auf 2 dimensional Daten antrainiert wurde, da es sich bewiesen hat Segmentierungen im Medizinischen Bereich vorzunehmen. Es besteht aus einem Encoder und einem "Decoder"Part, wobei wir verschiedene vortrainierte Encoder ausprobieren werden.

Auf einer detaillierten Datenanalyse aufbauend werden dann weitere Methoden angewendet, um die Performance zu verbessern. Es werden Trainingsdaten bereinigt und mit verschiedenen Eingabegrößen experimentiert. Ein weiterer Ansatz ist das intelligente beschneiden der Ausschnitte mit dem Ziel, dass sich das Modell besser auf die Aussagekräftigen Bereiche konzentrieren kann, sowie die Verwendung von "2.5 dimensionalen" Daten, auf die ich später genauer eingehe. Ein weiteres Feature ist das intelligente Zuschneiden des Inputs. Zuletzt erfolgt eine Auswertung der Ergebnisse und es werden weitere Ideen genannt, die es nicht in die Arbeit geschafft haben.

## Inhaltsverzeichnis

<b>1 Grundlagen</b>	<b>1</b>
1.1 Magen-Darmtrakt . . . . .	1
1.2 CNN . . . . .	1
1.3 U-Net . . . . .	1
1.4 RLE . . . . .	1
<b>2 Verwandte Arbeiten</b>	<b>1</b>
2.1 Semantische Segmentierung im medizinischem Bereich . . . . .	1
2.2 Intelligentes Zuschneiden des Inputs . . . . .	1
2.3 Loss Functions für Segmentierungsarchitekturen . . . . .	2
2.4 Encoder 1 . . . . .	2
2.5 Encoder 2 . . . . .	2
<b>3 Methodik</b>	<b>2</b>
3.1 Datenanalyse . . . . .	2
3.2 Metadatenextraktion . . . . .	3
3.3 Algorithmik . . . . .	3
3.4 2.5 dimensionale Daten . . . . .	3
3.5 Intelligentes Zuschneiden . . . . .	3
3.6 Experimente . . . . .	3
<b>4 Evaluation</b>	<b>3</b>
<b>5 Fazit</b>	<b>3</b>
5.1 Ausblick . . . . .	3
<b>Abbildungsverzeichnis</b>	<b>6</b>
<b>Tabellenverzeichnis</b>	<b>6</b>

## 1 Grundlagen

### 1.1 Magen-Darmtrakt

??

### 1.2 CNN

??

### 1.3 U-Net

U-Net wurde im Jahr 2015 von Olaf Ronneberger et al. vorgestellt als Segmentierungsarchitektur für den biomedizinischen Bereich und bildet die Baseline dieser Arbeit. Klassisch handelt es sich um ein Convolutional Neural Network bestehend aus 5 bis 6 'downsampling' (auch Encoder genannt) Schichten und der gleichen Menge an 'upsampling' (auch Decoder genannt) Schichten.

Eine 'downsample' Operation besteht zum einen aus einer 'Convolution operation' sowie einer 'Max pooling operation' an deren Ende jeweils eine ReLU Aktivierungsfunktion zum Einsatz kommt. In diesem Schritt lernt das Modell Eigenschaften auf Pixelebene, wie zum Beispiel Formen, Ecken, Kanten. Hierbei halbieren sich die Eingabedimensionen und die Tiefe des Bildes wird erhöht.

Im Decoder wird das Bild wieder auf seine ursprüngliche Größe mittels 'Transpose Convolution operations' auf seine ursprüngliche Größe gebracht. Hier ersetzt die Transponierung Operation die Maxpooling Operation (vergleiche Grafik). Dieser Part erlaubt dem Modell das Lokalisieren der zuvor gelernten Eigenschaften und gibt eine Maske zurück für jede Klasse, in der 0 der Hintergrund ist und 1 die jeweilige Klasse.

### 1.4 RLE

?

## 2 Verwandte Arbeiten

2 Seiten dazu

### 2.1 Semantische Segmentierung im medizinischen Bereich

### 2.2 Intelligentes Zuschneiden des Inputs

Paper dazu ist da

Index	ID	Class	Segmentation
1	case134_day0_slice_0085	large_bowel	NaN
2	case134_day0_slice_0085	small_bowel	41591 5 41599 7 41949 27 ...
3	case134_day0_slice_0085	stomach	NaN
4	case123_day0_slice_0001	large_bowel	35223 6 74352 7 32312 12 ...
5	case123_day0_slice_0001	small_bowel	63432 5 12354 7 41949 12 ...
6	case123_day0_slice_0001	stomach	NaN

Tabelle 1: Beispieldaten für zwe slices

### 2.3 Loss Functions für Segmentierungsarchitekturen

Paper dazu ist auch da

### 2.4 Encoder 1

Eventuell

### 2.5 Encoder 2

Eventuell

## 3 Methodik

### 3.1 Datenanalyse

Der Datensatz besteht aus 115 488 Zeilen und enthält drei features: ID, Klasse und einem Run-length codierten String, der die Maske enthält. Table 1. Jeder ID sind drei Zeilen gewidmet, jeweils für die drei Klassen. Zu jeder ID existiert ein Graustufen Bild, welche sich im train Ordner befinden:

```
input\uw-madison-gi-tract-image-segmentation\train\case101\
case101_day20\scans\slice_0001_266_266_1.50_1.50.png
```

Jedes slice enthält vier Zahlen (z.B. 266\_266\_1.50\_1.50.png), die ersten beiden stehen für die Auflösung des Bildes und die letzten beiden für den physischen Abstand der Pixel. Der Großteil der Aufnahmen stammt von Tag null oder tag eins Figure 2 und die durchschnittliche Anzahl an Bildern pro Fall beträgt X.

Beim betrachten der Verteilung der Segmentierungen fällt auf, dass das Vorkommen für jede Klasse stark variiert. Figure 5.

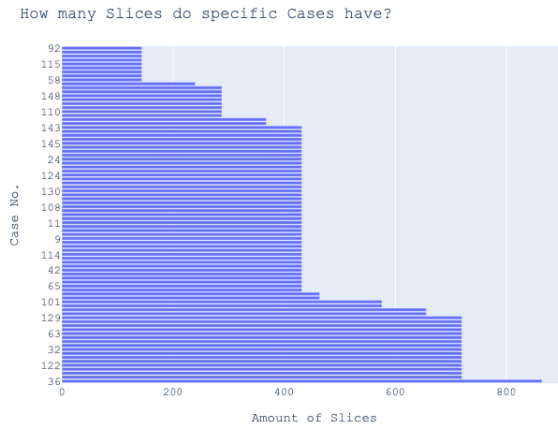


Abbildung 1: Anzahl Bilder pro Fall

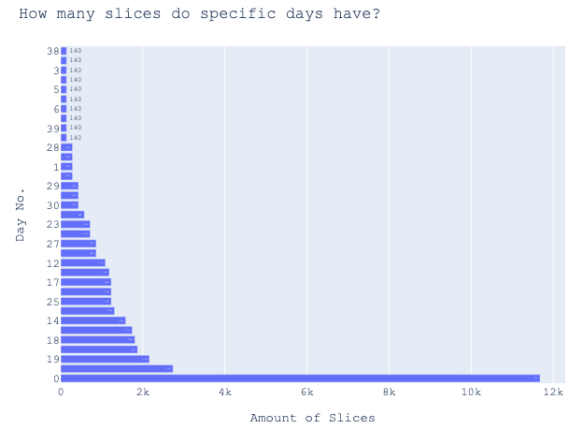


Abbildung 2: Anzahl Bilder pro Tag

### 3.2 Metadatenextraktion

### 3.3 Algorithmik

Was soll hier stehen?

### 3.4 2.5 dimensionale Daten

Was soll hier stehen?

### 3.5 Intelligentes Zuschneiden

Was soll hier stehen?

### 3.6 Experimente

Experimente.

## 4 Evaluation

## 5 Fazit

### 5.1 Ausblick

Ideen, die es nicht in die Arbeit geschafft haben oder nicht schaffen konnten.



**Con97** hat ein Buch geschrieben. Es gibt auch andere Arbeiten (**PeHe97**) die referenziert sind. In Abbildung ?? ist ein Sachverhalt dargestellt.

1 Autor: **Con97** (Con97)

2 Autoren: **IWNLP** (IWNLP)

3 Autoren: **liebeck-esau-conrad:2016:ArgMining2016**

(liebeck-esau-conrad:2016:ArgMining2016)

Online resource: **ILSVRC2016**

**quotes:**

Ein Beispiel für deutsche Anführungszeichen „quote“.

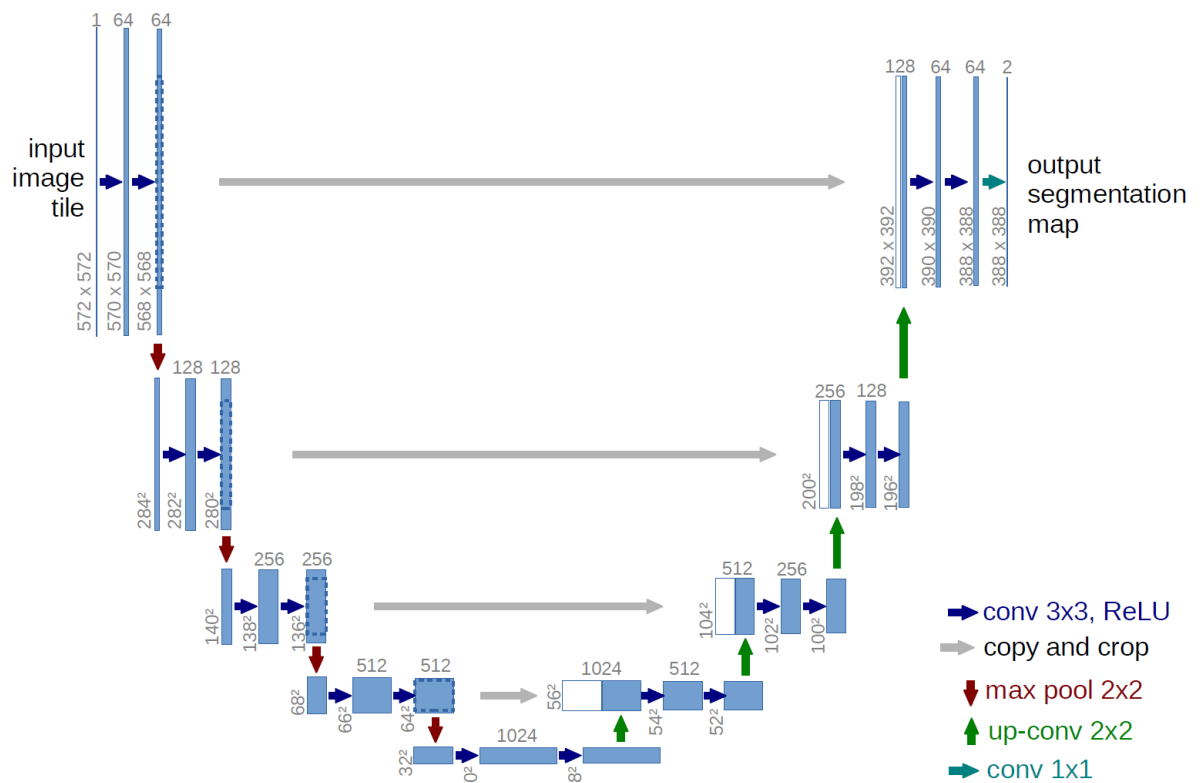


Abbildung 3: U-Net Architektur

Segmentation Distribution

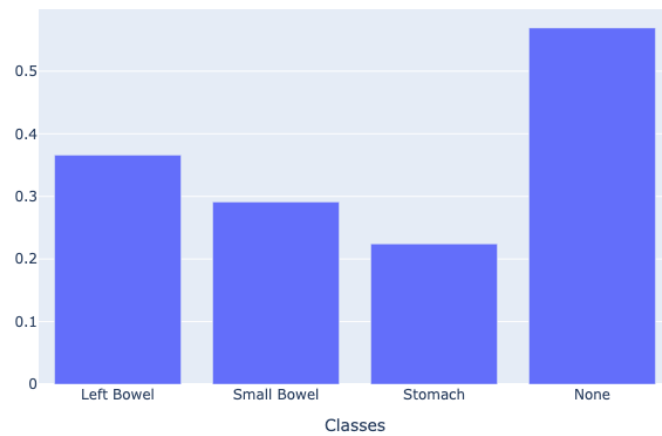


Abbildung 4: Verteilung der Klassen im Testset

Segmentation Distribution

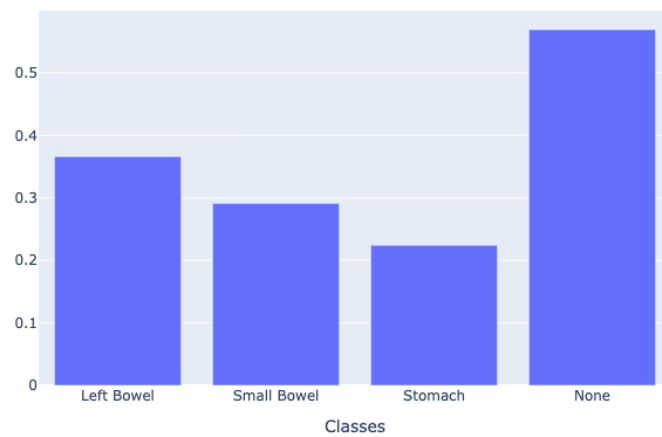


Abbildung 5: Verteilung der Klassen im Testset

## Abbildungsverzeichnis

1	Anzahl Bilder pro Fall . . . . .	3
2	Anzahl Bilder pro Tag . . . . .	3
3	U-Net Architektur . . . . .	4
4	Verteilung der Klassen im Testset . . . . .	5
5	Verteilung der Klassen im Testset . . . . .	5

## Tabellenverzeichnis

1	Beispieldaten für zwe slices . . . . .	2
---	--	---