



CNNs - nur Ideen weil ich gestresst bin und  
etwas schreiben lassen wollte

# PROJEKT-/STUDIEN-/BACHLEORARBEIT

für die Prüfung zum

Bachelor of Science

des Studienganges Informatik / Informationstechnik

an der

Dualen Hochschule Baden-Württemberg Karlsruhe

von

**chatGPT, by openAI**

Abgabedatum 1. April 2090

Bearbeitungszeitraum

12 Wochen

Matrikelnummer

4711

Kurs

tinf17b3

Ausbildungsfirma

Firmenname

Stadt

Betreuer der Ausbildungsfirma

Titel Vorname Nachname

Gutachter der Studienakademie

Titel Vorname Nachname

## Erklärung

Ich versichere hiermit, dass ich meine Projekt-/Studien-/Bachleorarbeit mit dem Thema: **?CNNs - nur Ideen weil ich gestresst bin und etwas schreiben lassen wollte?** selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Ich versichere zudem, dass die eingereichte elektronische Fassung mit der gedruckten Fassung übereinstimmt. \_\_\_\_\_

Ort      Datum

Unterschrift

*Sofern vom Dualen Partner ein Sperrvermerk gewünscht wird, ist folgende Formulierung zu verwenden:*

## Sperrvermerk

Der Inhalt dieser Arbeit darf weder als Ganzes noch in Auszügen Personen außerhalb des Prüfungsprozesses und des Evaluationsverfahrens zugänglich gemacht werden, sofern keine anderslautende Genehmigung vom Dualen Partner vorliegt.

## Abstract

Dieses L<sup>A</sup>T<sub>E</sub>X-Dokument kann als Vorlage für einen Praxis- oder Projektbericht, eine Studien- oder Bachelorarbeit dienen.

Zusammengestellt von Prof. Dr. Jürgen Vollmer <juergen.vollmer@dhbw-karlsruhe.de>  
<https://www.karlsruhe.dhbw.de>. Die jeweils aktuellste Version dieses L<sup>A</sup>T<sub>E</sub>X-Paketes ist immer auf der *FAQ-Seite* des Studiengangs Informatik zu finden: <https://www.karlsruhe.dhbw.de/inf/studienverlauf-organisatorisches.html> → *Formulare und Vorlagen*.

Stand \$Date: 2020/03/13 15:07:45 \$

# Contents

<b>1</b>	<b>Introduction</b>	<b>8</b>
1.1	Introduce the topic of deep neural networks and their applications in image segmentation/localization . . . . .	8
1.2	Outline the purpose of your paper and the approach you will use to train your model . . . . .	8
<b>2</b>	<b>Deep neural networks</b>	<b>10</b>
2.1	Explain the concepts of artificial neural networks and deep learning . . . . .	10
2.2	Discuss the architecture of convolutional neural networks (CNNs) and their role in image segmentation/localization . . . . .	11
2.3	Describe the different layers of a CNN and how they are used to extract features from images . . . . .	11
<b>3</b>	<b>Dataset and experimental setup</b>	<b>13</b>
3.1	Describe the dataset you used to train your model and the preprocessing steps you applied . . . . .	13
3.2	Explain the architecture of your CNN model and the hyperparameters you selected	13
3.3	Discuss the training process and any modifications you made to improve performance	13
<b>4</b>	<b>Results and analysis</b>	<b>14</b>
4.1	Present the results of your experiments, including metrics such as accuracy, precision, and recall . . . . .	14
4.2	Analyze the impact of different hyperparameters and training strategies on performance . . . . .	14
4.3	Visualize the output of your model and provide insights into its behavior . . . . .	14
<b>5</b>	<b>Comparison with other approaches</b>	<b>15</b>
5.1	Compare your model's performance with other approaches in the literature . . . .	15
5.2	Discuss the strengths and limitations of your model and the tradeoffs between different CNN architectures and training strategies . . . . .	15
<b>6</b>	<b>Applications and future directions</b>	<b>16</b>
6.1	Discuss the practical applications of deep neural networks for image segmentation/localization	16
6.2	Consider the limitations and challenges of deep learning for image segmentation/localization	16
6.3	Suggest possible directions for future research in the field . . . . .	16

<i>CONTENTS</i>	2
<b>7 Conclusion</b>	<b>17</b>
7.1 Summarize the key points of your paper and the main findings of your research .	17
7.2 Discuss the implications of your research and its potential impact on the field of computer vision . . . . .	17
7.3 Provide some final thoughts and suggestions for future work . . . . .	17
<b>Anhang</b>	<b>18</b>
<b>Index</b>	<b>18</b>
<b>Literaturverzeichnis</b>	<b>18</b>
<b>Liste der ToDo's</b>	<b>19</b>

# List of Figures

# List of Tables

# Liste der Algorithmen



# Formelverzeichnis

# Abkürzungsverzeichnis

# Chapter 1

## Introduction

### 1.1 Introduce the topic of deep neural networks and their applications in image segmentation/localization

Convolutional Neural Networks (CNN) sind zu einem grundlegenden Werkzeug in der Computer Vision und Bildverarbeitung geworden. Sie sind eine spezielle Art von tiefen neuronalen Netzen, die entwickelt wurden, um hierarchische und abstrakte Merkmale aus Bildern für Klassifizierungs- und Lokalisierungsaufgaben zu extrahieren. Die Kernidee hinter CNNs ist die Verwendung von Faltungsschichten mit lernfähigen Filtern um lokale Merkmale aus verschiedenen Regionen des Eingabebildes zu extrahieren. Diese Filter gleiten über das Eingabebild und berechnen Punktprodukte zwischen den Filterwerten und den entsprechenden Pixelwerten.

CNNs haben bemerkenswerte Erfolge in einem breiten Spektrum von Anwendungen erzielt, darunter Bildklassifikation, Objekterkennung und Segmentierung. Bei der Bildsegmentierung wird ein Bild in mehrere Segmente oder Regionen unterteilt, die jeweils einem bestimmten Objekt oder Hintergrund im Bild entsprechen. CNNs eignen sich besonders gut für Bildsegmentierungsaufgaben, da sie lernen können, komplexe Merkmale und Muster aus Bildern zu extrahieren und Objektgrenzen und räumliche Beziehungen zwischen Objekten zu erkennen.

In den letzten Jahren haben CNN-basierte Verfahren ihre Leistungsfähigkeit in verschiedenen Bildsegmentierungsanwendungen unter Beweis gestellt, darunter medizinische Bildgebung, Satellitenbilder und autonomes Fahren. In der medizinischen Bildgebung wurden CNNs beispielsweise zur Erkennung und Segmentierung von Tumoren, Hirnverletzungen und Anomalien in verschiedenen Organen eingesetzt. Im Bereich des autonomen Fahrens wurden CNN-basierte Modelle zur Erkennung von Fahrspuren und Hindernissen sowie zur semantischen Segmentierung der Umgebung eingesetzt.

### 1.2 Outline the purpose of your paper and the approach you will use to train your model

Dieser Beitrag gibt einen Überblick über die jüngsten Fortschritte bei CNN-basierten Methoden zur Bildsegmentierung und -lokalisierung. Wir werden die wichtigsten Ideen und Techniken hinter CNNs diskutieren, einschließlich der Verwendung von Faltungsschichten, Pooling-Schichten und vollständig verbundenen Schichten. Wir werden auch die Herausforderungen und Grenzen von CNNs aufzeigen, wie z.B. die Notwendigkeit großer Mengen annotierter Trainingsdaten und

erheblicher Hardware-Ressourcen für Training und Einsatz.

Um unser CNN-basiertes Modell zur Bildsegmentierung zu trainieren, werden wir einen großen annotierten Datensatz und ein Deep Learning Framework wie TensorFlow oder PyTorch verwenden. Wir werden die Eingabebilder vorverarbeiten, indem wir ihre Pixelwerte normalisieren und Datenanreicherungstechniken wie Rotation, Translation und Skalierung anwenden. Wir trainieren unser Modell mit einer Verlustfunktion, die den Unterschied zwischen der vorhergesagten und der tatsächlichen Segmentierung misst, wie z.B. Cross-Entropie-Verlust oder Dice-Verlust. Wir werden Backpropagation und Stochastic Gradient Descent (SGD) oder Varianten davon verwenden, um die Parameter unseres Modells zu aktualisieren und den Verlust zu minimieren.

Big bullshit incoming:

Um die Leistung unseres Modells zu bewerten, werden wir Standardmetriken wie Intersection over Union (IoU), Dice-Koeffizient und Pixelgenauigkeit verwenden. Außerdem werden wir die Leistung unseres Modells mit modernen Methoden vergleichen und seine Stärken und Schwächen analysieren. Abschließend werden wir zukünftige Forschungsrichtungen und potentielle Anwendungen von CNN-basierten Methoden in der Bildsegmentierung und -lokalisierung diskutieren.

## Chapter 2

# Deep neural networks

### 2.1 Explain the concepts of artificial neural networks and deep learning

Künstliche neuronale Netze (ANN) sind algorithmische Modelle, die das biologische Verhalten von Neuronen im menschlichen Gehirn nachahmen. ANNs bestehen aus vernetzten Knoten, die Informationen in mehreren Schichten verarbeiten. In einem einfachen neuronalen Feed-Forward-Netzwerk fließt die Information unidirektional von der Eingabeschicht zur Ausgabeschicht, wobei jede Schicht eine bestimmte Funktion ausführt. ANNs können trainiert werden, um Muster und Beziehungen in den Daten durch einen Prozess zu lernen, der als überwachtes Lernen bezeichnet wird, bei dem das Netzwerk mit markierten Trainingsbeispielen konfrontiert wird und seine internen Parameter anpasst, um den Fehler zwischen seiner Ausgabe und den korrekten Markierungen zu minimieren.

Deep Learning ist ein Teilgebiet des maschinellen Lernens, das ANNs mit mehreren verborgenen Schichten umfasst. Tiefe neuronale Netze (DNN) können hierarchische Darstellungen von Daten lernen, wobei jede Schicht abstraktere Merkmale lernt als die vorhergehende Schicht. DNNs haben sich bei verschiedenen Aufgaben als besonders leistungsfähig erwiesen, z. B. in den Bereichen maschinelles Sehen, Verarbeitung natürlicher Sprache und Spracherkennung. Insbesondere DNNs mit Faltungsschichten, so genannte Convolutional Neural Networks (CNNs), sind sehr effektiv für bildbezogene Aufgaben wie Bildklassifikation, Objekterkennung und Bildsegmentierung.

Der Erfolg des Deep Learning lässt sich auf mehrere Faktoren zurückführen, darunter die Verfügbarkeit großer Mengen an markierten Daten, Fortschritte bei der Rechenleistung und der Parallelverarbeitung sowie Verbesserungen bei Optimierungsalgorithmen wie dem stochastischen Gradientenabstieg. Deep Learning hat auch zur Entwicklung neuer Techniken geführt, z. B. Transfer Learning, bei dem zuvor trainierte Modelle für bestimmte Aufgaben verfeinert werden, und Generative Adversarial Networks (GANs), die realistische Bilder und Videos erzeugen können.

Zusammenfassend lässt sich sagen, dass die Konzepte der ANN und des Deep Learning den Bereich der künstlichen Intelligenz revolutioniert haben, indem sie Maschinen in die Lage versetzen, komplexe Aufgaben zu bewältigen, von denen man früher glaubte, dass sie ausschließlich dem Menschen vorbehalten seien.

## 2.2 Discuss the architecture of convolutional neural networks (CNNs) and their role in image segmentation/localization

Convolutional Neural Networks (CNNs) sind eine Art Deep Neural Networks, die den Bereich des maschinellen Sehens revolutioniert haben. CNNs sind darauf ausgelegt, visuelle Eingaben wie Bilder oder Videos zu analysieren und Merkmale zu extrahieren, die für eine bestimmte Aufgabe relevant sind. CNNs sind besonders effektiv bei bildbezogenen Aufgaben wie Bildklassifikation, Objekterkennung und Bildsegmentierung.

Die Architektur eines CNN besteht aus einer Folge von Schichten, die das Eingangsbild in ein vorhergesagtes Ausgangsbild umwandeln. Die erste Schicht eines CNN ist normalerweise eine Faltungsschicht, die eine Reihe von Filtern auf das Eingangsbild anwendet. Jeder Filter ist eine kleine Matrix von Gewichten, die mit dem Eingabebild gefaltet werden, um lokale Merkmale zu extrahieren. Faltungsschichten können lernen, grundlegende Muster wie Kanten, Ecken oder Flecken zu erkennen, die häufig in Bildern vorkommen.

Die Ausgabe einer Faltungsschicht wird dann durch eine Aktivierungsfunktion wie ReLU oder Sigmoid geleitet, die dem Modell Nichtlinearität verleiht. Die Aufgabe der Aktivierungsfunktion besteht darin, die Ausgabe der Faltungsschicht in eine komplexere und aussagekräftigere Merkmalsdarstellung umzuwandeln. Diese Merkmalsdarstellung wird dann durch eine Pooling-Schicht geleitet, die die räumlichen Dimensionen der Merkmalskarte reduziert, während die wesentlichen Informationen erhalten bleiben. Pooling-Schichten können auch eine Form von Translationsinvarianz einführen, was bedeutet, dass das Modell dasselbe Objekt unabhängig von seiner Position im Bild erkennen kann.

Die Ausgabe der Pooling-Schicht wird dann in einen weiteren Satz von Faltungsschichten, Aktivierungsschichten und Pooling-Schichten eingespeist. Die Anzahl dieser Schichten kann je nach Komplexität der Aufgabe und Größe des Datensatzes variieren. Schließlich wird die Ausgabe der letzten Pooling-Schicht abgeflacht und durch eine oder mehrere vollständig verbundene Schichten geleitet, die die endgültige Klassifizierung oder Segmentierung vornehmen.

Bei der Bildsegmentierung können CNNs lernen, jedes Pixel eines Bildes einer bestimmten Klasse oder Kategorie zuzuordnen. In der medizinischen Bildgebung können CNNs beispielsweise zur Unterscheidung von Tumoren oder Läsionen von gesundem Gewebe eingesetzt werden. Die Fähigkeit von CNNs, räumliche Merkmale aus Bildern zu lernen, macht sie besonders geeignet für Bildsegmentierungsaufgaben. Durch die Nutzung der räumlichen Kohärenz des Bildes können CNNs hochpräzise Segmentierungen durchführen, die herkömmliche Segmentierungsmethoden übertreffen.

## 2.3 Describe the different layers of a CNN and how they are used to extract features from images

Faltungsschichten sind das Herzstück von CNNs. Diese Schichten lernen, relevante Merkmale aus dem Eingabebild zu extrahieren, indem sie eine Reihe von adaptiven Filtern auf das Bild anwenden. Jeder Filter gleitet über das gesamte Eingabebild und führt eine elementweise Multiplikation mit anschließender Summation durch. Das Ergebnis dieser Operation ist ein einzelner Skalarwert, der den Grad der Ähnlichkeit zwischen dem Filter und der entsprechenden Region des Eingabebildes darstellt. Diese Skalarwerte werden dann zu einer Merkmalskarte angeordnet.

Nach jeder Faltungsschicht werden Aktivierungsschichten hinzugefügt, um dem Modell Nichtlinearität zu verleihen. Ohne diese Schichten wäre das Modell auf lineare Transformationen beschränkt, die nicht leistungsfähig genug sind, um komplexe nichtlineare Beziehungen in den Daten zu modellieren. Aktivierungsfunktionen wie ReLU, Sigmoid und tanh werden üblicherweise in CNNs verwendet.

Pooling-Layer werden verwendet, um die räumlichen Dimensionen der Merkmalskarten zu reduzieren, während die wesentlichen Informationen erhalten bleiben. Diese Schichten unterteilen die Merkmalskarten in nicht überlappende Regionen und wenden auf jede Region eine Reduktionsoperation an, wie z. B. Max- oder Average-Pooling. Dadurch wird die Anzahl der Modellparameter reduziert und das Modell recheneffizienter.

Voll verknüpfte Schichten werden normalerweise am Ende des CNN hinzugefügt, um das Bild zu klassifizieren oder die Segmentierungsmaske vorherzusagen. Diese Schichten nehmen die abgeflachten Merkmalskarten aus den vorhergehenden Schichten und lassen sie durch einen Satz vollständig verbundener Neuronen laufen, um die endgültige Vorhersage zu treffen. Im Fall der Bildklassifikation stellt die Ausgabe der vollständig verbundenen Schicht die Wahrscheinlichkeitsverteilung über die verschiedenen Klassen dar. Im Falle der Bildsegmentierung stellt die Ausgabe der vollständig verbundenen Schicht die Wahrscheinlichkeit dar, dass jedes Pixel zu dem interessierenden Objekt gehört.

Insgesamt ermöglicht die Kombination dieser verschiedenen Schichten den CNNs, komplexe hierarchische Repräsentationen des Eingabebildes zu erlernen, wodurch sie für Aufgaben wie Bildklassifikation, Objekterkennung und Segmentierung gut geeignet sind.

## Chapter 3

# Dataset and experimental setup

### **3.1 Describe the dataset you used to train your model and the preprocessing steps you applied**

We used the XYZ dataset, which contains X images of size YxZ. We applied normalization and data augmentation techniques to enhance the model's performance.

### **3.2 Explain the architecture of your CNN model and the hyperparameters you selected**

Our CNN model has X convolutional layers, Y pooling layers, and Z fully connected layers. We experimented with different hyperparameters, including learning rate, batch size, and number of epochs, to find the optimal settings.

### **3.3 Discuss the training process and any modifications you made to improve performance**

We used the Adam optimizer and the binary cross-entropy loss function to train our model. We experimented with different modifications, including dropout regularization and changing the number of filters in the convolutional layers, to improve performance.



## Chapter 4

# Results and analysis

### **4.1 Present the results of your experiments, including metrics such as accuracy, precision, and recall**

Our model achieved an accuracy of XWe also computed other metrics, such as F1 score and intersection over union (IoU), to evaluate the model's performance.

### **4.2 Analyze the impact of different hyperparameters and training strategies on performance**

We found that changing the learning rate and batch size had a significant impact on performance. We also observed that adding dropout regularization improved the model's generalization ability.

### **4.3 Visualize the output of your model and provide insights into its behavior**

We visualized the segmentation masks produced by our model and compared them with the ground truth masks. We observed that our model was able to accurately segment objects in the image, but struggled with small and complex shapes. We also visualized the feature maps produced by the convolutional layers to gain insights into how the model was extracting features from the images.

## Chapter 5

# Comparison with other approaches

### **5.1 Compare your model's performance with other approaches in the literature**

We compared our model's performance with other state-of-the-art approaches in image segmentation/localization. Our model achieved comparable or better results on the XYZ dataset, demonstrating its effectiveness.

### **5.2 Discuss the strengths and limitations of your model and the tradeoffs between different CNN architectures and training strategies**

Our model is a simple CNN architecture that achieved good performance on the XYZ dataset. However, more complex architectures, such as U-Net and Mask R-CNN, may perform better on more challenging datasets. The choice of hyperparameters and training strategies can also have a significant impact on the model's performance.

## Chapter 6

# Applications and future directions

### **6.1 Discuss the practical applications of deep neural networks for image segmentation/localization**

DNNs can be used in various applications, such as medical imaging, autonomous vehicles, and object recognition in images. Image segmentation/localization can provide valuable information for decision-making in these applications.

### **6.2 Consider the limitations and challenges of deep learning for image segmentation/localization**

Deep learning requires large amounts of labeled data, which can be challenging to obtain in some domains. Overfitting can also be a significant issue, particularly when dealing with small datasets.

### **6.3 Suggest possible directions for future research in the field**

Future research could focus on developing more efficient and accurate deep learning models for image segmentation/localization. Transfer learning and domain adaptation techniques could also be explored to overcome the data limitations in some domains.

## Chapter 7

# Conclusion

### **7.1 Summarize the key points of your paper and the main findings of your research**

We presented an approach to train a CNN model for image segmentation/localization. We experimented with different hyperparameters and training strategies and achieved good performance on the XYZ dataset.

### **7.2 Discuss the implications of your research and its potential impact on the field of computer vision**

Our research demonstrates the effectiveness of simple CNN architectures for image segmentation/localization. It has the potential to contribute to the development of more accurate and efficient models in the field of computer vision.

### **7.3 Provide some final thoughts and suggestions for future work**

Future work could explore more complex CNN architectures and novel training strategies to improve performance. The impact of deep learning on image segmentation/localization in various domains could also be further investigated.

# Änderungen

- 2020/03/13** Tippfehler korrigiert  
aktuelle Formulierungen aus der Prüfungsordnung Technik übernommen  
Formatdatei erklärt
- 2017/10/06** Anpassung an neuer Versionen diverse Pakete.
- 2016/03/16** Auf UTF-8 umgestellt, Indices.
- 2010/04/12** ToDo-Markierungen mit dem `\todo`-Kommando.
- 2010/01/27** Anhang (`appendix`), Selbständigkeits-Erklärung, `framed`-Paket.
- 2010/01/21** Abkürzungen (`acronym`), `table` und `tabular` benutzt, unübliche Pakete beigelegt.
- 2010/01/18** Code-Listings (`listings`), Literaturreferenzen `biblatex`)
- 2010/01/11** Initiale Version.

## Liste der ToDo's