

Master-Teamprojekt

Titel der Arbeit

Titel of the Thesis

Max Mustermann

Mat.Nr.: 420815

Betreuer:

Prof. Dr. rer. nat. E.-G. Haffner

Datum:

14. März 2022

Eidesstattliche Erklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit selbstständig verfasst habe und keine anderen als die im Literaturverzeichnis angegebenen Quellen benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder noch nicht veröffentlichten Quellen entnommen sind, sind als solche kenntlich gemacht. Die Zeichnungen oder Abbildungen in dieser Arbeit sind von mir selbst erstellt worden oder mit einem entsprechenden Quellennachweis versehen. Diese Arbeit ist in gleicher oder ähnlicher Form noch bei keiner anderen Prüfungsbehörde eingereicht worden.

Ort, Datum

Unterschrift

Danksagung

Ich bedanke mich bei all denjenigen, die mich während der Anfertigung dieser Arbeit unterstützt und motiviert haben. Ein ganz besonderer Dank gilt meinem Betreuer Prof. Dr. rer. nat. E.-G. Haffner. Ebenfalls möchte ich der Projektgruppe eTRainer der Hochschule Trier, sowie meinen Korrekturlesenden danken.

Abstract

This is a summary of all the important points and achievements of this work.

Zusammenfassung

Hierbei handelt es sich um eine Zusammenfassung aller wichtige Punkte und Errungenschaften dieser Arbeit.

Abkürzungsverzeichnis

ARS	Audience Response System
CGI	Common Gateway Interface
CSS	Cascading Style Sheets
HTML	Hypertext Markup Language
ML	Machine Learning

Inhaltsverzeichnis

1. Einleitung	1
2. Ein Kapitel	3
2.1. Ein Abschnitt	3
2.1.1. Ein Unterabschnitt	3
2.2. Pix2Pix	3
2.2.1. Pix2Pix-Kernkonzepte	4
3. Hauptteil - 1	7
4. Hauptteil - 2	9
5. Hauptteil - 3	11
6. Evaluation	13
7. Fazit und Ausblick	15
7.1. Fazit	15
7.2. Ausblick	15
A. Anhang - Code	17
B. Anhang - Dokumentationen	19
Verzeichnisse	21
Literaturverzeichnis	21
Abbildungsverzeichnis	23
Tabellenverzeichnis	25
Code-Auszugs-Verzeichnis	27
Glossar	29

1

Einleitung

Hier wird in die Arbeit eingeleitet.

2

Ein Kapitel

2.1. Ein Abschnitt

2.1.1. Ein Unterabschnitt

2.2. Pix2Pix

Pix2Pix, initiiert von Isola et al., hat sich als zentrales Framework für Bild-zu-Bild-Übersetzungen auf der Basis von bedingten generativen adversariellen Netzwerken (cGANs) etabliert. Es ermöglicht die Erstellung einer abstrakten Abbildung von einem Eingangsbild zu einem korrespondierenden Ausgangsbild und bewältigt dabei eine vielfältige Palette an Bildübersetzungsaufgaben, wie die Transformation von Skizzen in realistische Bilder oder die Konvertierung von Tages- zu Nachtaufnahmen.

Pix2Pix fungiert hier als Generative Adversarial Network (GAN), spezialisiert auf diverse Formen der Bildübersetzung. Darunter fallen die Umwandlung von Schwarz-Weiß-Fotos in Farbbilder, die Transformation von Skizzen in realistische Bilder, und relevant für diese Arbeit, die Konvertierung von Satellitenbildern in kartographische Darstellungen, ähnlich den Visualisierungen von Google Maps.

Die Architektur von Pix2Pix besteht aus einem Generator und einem Diskriminator. Der Generator, der eine U-Net-Architektur verwendet, besteht aus einem Encoder und einem Decoder. Der Encoder komprimiert das Eingangsbild schrittweise zu einer niedrigdimensionalen Repräsentation, während der Decoder diese dazu nutzt, das Ausgangsbild zu rekonstruieren. Skip-Verbindungen zwischen Encoder und Decoder helfen dabei, sowohl globale als auch lokale Informationen im generierten Bild zu bewahren.

Der Diskriminator nimmt die Form eines PatchGAN-Modells an und bewertet Patches des Bildes, indem er die Wahrscheinlichkeit für die Echtheit jedes Patches ausgibt. Dies ermöglicht die Anwendung des Diskriminators auf Bilder unterschiedlicher Größen. Im Zuge des adversariellen Trainingsprozesses passen sowohl der Generator als auch der Diskriminator ihre Fähigkeiten fortlaufend an. Während der Generator lernt, immer realistischere Übersetzungen zu erzeugen, wird der Diskriminator stetig besser darin, zwischen echten und generierten Bildern zu unterscheiden.

2.2.1. Pix2Pix-Kernkonzepte

Die Bildverarbeitung hat in den letzten Jahren durch den Einsatz tiefer neuronaler Netzwerke erhebliche Fortschritte gemacht. Im Mittelpunkt vieler dieser Fortschritte steht die U-Net-Architektur, die speziell für die Bildsegmentierung entwickelt wurde. Diese Architektur zeichnet sich durch ihre angeklügelte Kombination aus Encoder- und Decoder-Strukturen sowie durch den Einsatz von Skip-Verbindungen aus.

Bei der Encoder-Decoder-Struktur handelt es sich um einen Ansatz, bei dem das Eingangsbild zunächst durch den Encoder schrittweise reduziert wird. Dieser Prozess dient dazu, wesentliche Merkmale des Bildes zu erfassen. Anschließend wird das Bild durch den Decoder wiederhergestellt, indem die zuvor extrahierten Merkmale verwendet werden. Während dieser Prozesse besteht jedoch das Risiko des Informationsverlustes, insbesondere in den tieferen Schichten des Netzwerks. Um dieses Problem zu adressieren, führt die U-Net-Architektur Skip-Verbindungen ein. Diese direkten Verbindungen zwischen korrespondierenden Schichten des Encoders und Decoders sorgen dafür, dass Detailinformationen nicht verloren gehen. Genauer gesagt, ermöglichen diese Verbindungen den direkten Informationsfluss zwischen jeweils äquivalenten Schichten, wodurch die Rekonstruktion des Bildes im Decoder mit einer höheren Genauigkeit erfolgt.

Die Bedeutung von Skip-Verbindungen zeigt sich insbesondere in Anwendungen wie der Bild-zu-Bild-Übersetzung. Hier muss oft ein Bild mit niedriger Auflösung in ein Bild mit hoher Auflösung überführt werden, ohne dass Details verloren gehen. Die U-Net-Architektur, die angereichert mit diesen Verbindungen ist, ermöglicht daher eine feinere Rekonstruktion, die sowohl globale als auch lokale Informationen berücksichtigt.

Somit kann die U-Net-Architektur durch ihre Kombination aus Encoder-Decoder-Struktur und Skip-Verbindungen ein effektives Werkzeug für die Bildsegmentierung darstellen. Ihre Fähigkeit, sowohl globale Muster als auch feine Details zu berücksichtigen, macht sie zu einer bevorzugten Wahl für viele Bildverarbeitungsaufgaben.

In Abbildung 2.1 ist die typische U-Net-Architektur dargestellt. Die linke Seite des "U" repräsentiert den Encoder-Teil, der das Eingangsbild schrittweise reduziert und wesentliche Merkmale extrahiert. Die rechte Seite repräsentiert den Decoder-Teil, der das Bild mithilfe der extrahierten Merkmale rekonstruiert. Die horizontalen Linien repräsentieren die Skip-Verbindungen, die sicherstellen, dass Detailinformationen zwischen den korrespondierenden Schichten des Encoder und Decoders direkt übertragen werden.

In der Pix2Pix Technologie dient diese U-Net-Architektur als Generator. Er ist das zentrale Element, das für die Bild-zu-Bild-Übersetzung verantwortlich ist. Die Wahl der U-Net-Struktur für den Generator liegt in ihrer Fähigkeit, feinere Details und Kontextinformationen aus dem Eingangsbild beizubehalten, was für die Bild-zu-Bild-Übersetzung von entscheidender Bedeutung ist. Die Encoder-Decoder-Struktur des U-Net ermöglicht es dem Generator, den globalen Kontext des Bildes zu erfassen, während die Skip-Verbindungen sicherstellen, dass auch lo-

kale Details im resultierenden Bild berücksichtigt werden.

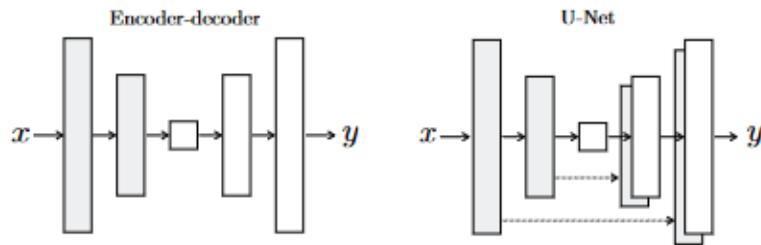


Abbildung 2.1.: Schematische Darstellung der U-Net-Architektur. Die Architektur besteht aus einem Encoder-Teil (links), einem Decoder-Teil (rechts) und Skip-Verbindungen zwischen korrespondierenden Schichten.

3

Hauptteil - 1

Hier wird der Inhalt der Arbeit präsentiert.

4

Hauptteil - 2

Hier wird der Inhalt der Arbeit präsentiert.

5

Hauptteil - 3

Hier wird der Inhalt der Arbeit präsentiert.



Evaluation

Hier wird die Arbeit evaluiert.

7

Fazit und Ausblick

Hier wird ein Fazit und ein Ausblick gegeben.

7.1. Fazit

Fazit.

7.2. Ausblick

Ausblick.



Anhang - Code

Hier sehen Sie den gesamten Quellcode!

B

Anhang - Dokumentationen

Hier sehen Sie die gesamten Dokumentationen zu den erstellten Programmen.

Literaturverzeichnis

- [Haf18] HAFNER, E.G.: *Lineare Algebra für Dummies*. Wiley, 2018 (Für Dummies). <https://books.google.de/books?id=XfWADwAAQBAJ>. – ISBN 9783527819430

Abbildungsverzeichnis

- 2.1. Schematische Darstellung der U-Net-Architektur. Die Architektur besteht aus einem Encoder-Teil (links), einem Decoder-Teil (rechts) und Skip-Verbindungen zwischen korrespondierenden Schichten. . . 5

Tabellenverzeichnis

Code-Auszugs-Verzeichnis

Glossar

- **C++:**
Hardwarenahe, objektorientierte Programmiersprache.
- **HTML:**
Hypertext Markup Language - textbasierte Auszeichnungssprache zur Strukturierung elektronischer Dokumente.
- **HTTP:**
Hypertext Transfer Protocol - Protokoll zur Übertragung von Daten auf der Anwendungssicht über ein Rechnernetz.
- **iARS:**
innovative Audio Response System - System mit zwei Applikationen (iARS-master-App; iARS-student-App), dass sich zum Einsetzen von e-TR-ainer-Inhalten in Vorlesungen eignet.
- **ISO:**
Internationale Vereinigung von Normungsorganisationen.
- **JavaScript:**
Skriptsprache zu Auswertung von Benutzerinteraktionen.
- **Konstruktor:**
Beim Erzeugen einer Objektinstanz aufgerufene Methode zum Initialisieren von Eigenschaften.
- **MySQL:**
Relationales Datenbankverwaltungssystem.
- **OLAT:**
Online Learning and Training - Lernplattform für verschiedene Formen von webbasiertem Lernen.
- **OOP:**
Objektorientierte Programmierung - Programmierparadigma, nach dem sich die Architektur eine Software an realen Objekten orientiert.
- **Open Source:**
Software, die öffentlich von Dritten eingesehen, geändert und genutzt werden kann.
- **PHP:**
Skriptsprache zur Erstellung von Webanwendungen.
- **Python:**
Skript- und Programmiersprache, die unter Anderem objektorientiertes Programmieren ermöglicht.

- **Shell:**
Shell oder auch Unix-Shell - traditionelle Benutzerschnittstelle von Unix-Betriebssystemen
- **Spyder:**
Entwicklungsumgebung für wissenschaftliche Programmierung in der Programmiersprache Python.
- **SymPy:**
Python-Bibliothek für symbolische Mathematik.