

# Classic Low-Level Image Processing with Deep Learning

## Metadaten

<b>Themenbereiche:</b>	Software-Erstellung
<b>Studierender:</b>	Etterlin Martin, Waller Dominik
<b>Dozent:</b>	Prof. Martin Jud, Prof. Dr. Thomas Koller
<b>Experte:</b>	Urs Gehrig
<b>Keywords:</b>	Image Processing, Edge Detection, Canny, Deep Learning, Artificial Neural Network

## 1. Aufgabenstellung

Die Erkennung von Kanten in einem Bild, hat eine grosse Bedeutung für die maschinelle Interpretation dessen. Aus diesem Grund gibt es eine Vielzahl von Algorithmen, die sich damit beschäftigen, möglichst genau Kanten zu detektieren. Ein weit verbreiteter Algorithmus ist dabei der sogenannte Canny-Algorithmus. In den letzten Jahren haben Machine-Learning Ansätze mit neuronalen Netzwerken, im Gebiet der Objektklassifizierung und Detektion, grosse Fortschritte gemacht. In dieser Arbeit wird nun untersucht, inwiefern die Kantenerkennung, durch solche neuronale Netzwerke umgesetzt werden kann. Der Fokus liegt darauf, eine Kantenerkennung zu ermöglichen, die der Wahrnehmung eines Menschen entspricht.

## 2. Ergebnisse

Im Rahmen der Bachelor-Diplomarbeit (BDA) werden folgende Ergebnisse erarbeitet:

- **Verschiedene Netzwerke für die Erkennung von Kanten:**  
Es werden zwei neuronale Netzwerke mit unterschiedlicher Architektur entworfen. Diesen Netzwerken wird die Kantenerkennung, wie sie ein Mensch machen würde, trainiert.
- **Daten zum Trainieren der Netzwerke:**  
Neben den Netzwerken an sich, wird in dieser BDA auch die Evaluation und Generation der Trainingsdaten vorgenommen. Dabei werden verschiedene Datensätze, welche von Menschen eingezeichnete Kanten beinhalten, vorgestellt. Einer dieser Datensätze ist der BSDS500-Datensatz, welcher auch für das Training der entworfenen Netzwerke verwendet wird.
- **Testen und Validierung der Netzwerke:**  
Die Qualität der trainierten Netzwerke wird mittels Metriken wie Precision, Recall, F-Score und Average Precision gemessen und die optimale Architektur anhand dieser Metriken evaluiert. Mit dem offiziellen Benchmarking des BSDS500-Datensatzes werden die generierten Netzwerke zum Schluss mit bereits existierenden Algorithmen und neuronalen Netzwerken verglichen.
- **Bericht:**  
Das komplette Vorgehen sowie die entstandenen Ergebnisse werden in einem Bericht dokumentiert.

### 3. Lösungskonzept

Nach einer initialen Einarbeitungsphase, bei der die Grundlagen zum Thema neuronale Netzwerke geschaffen und dokumentiert werden, werden in dieser Arbeit unterschiedliche neuronale Netzwerke kreiert, trainiert und evaluiert. Als Grundlage für das Training diente dabei der BSDS500-Datensatz, welcher insgesamt 500 Bilder mit von Mensch eingezeichneten Kanten enthält. Um die Performance der erstellten Netzwerke messen zu können, werden unterschiedliche Metriken definiert, welche nach jedem Trainingsvorgang evaluiert werden. Als Ergebnis dieser Evaluationen werden zwei unterschiedliche, finale Netzwerkarchitekturen definiert. Das erste neuronale Netzwerk ist ein einfaches Fully Convolutional Netzwerk (BDA-FCN-v2), bei dem mehrere Convolutional Layer hintereinandergelegt werden. Um zu prüfen ob andere, bereits etablierte Netzwerkarchitekturen besser abschneiden, wird für die zweite Architektur ein einfaches Encoder/Decoder-Netzwerk, das sogenannte U-Net, evaluiert und optimiert (BDA-UNet-v1). Diese finalen Architekturen werden zum Schluss, durch das vom BSDS500-Datensatz mitgelieferte Benchmarking, evaluiert. Das Resultat dieser Evaluation zeigt, dass die zwei in dieser Arbeit erstellten, einfachen Architekturen mit einem F-Score von 0.710 (BDA-FCN-v2) und 0.722 (BDA-UNet-v1), besser abschneiden als klassische Algorithmen, wie beispielsweise der eingangs erwähnte Canny-Algorithmus mit einem F-Score von 0.611. Abbildung 1 zeigt beispielhaft die (Zwischen-)Resultate, welche durch das BDA-UNET-v1-Netzwerke generiert werden. Des Weiteren konnte festgestellt werden, dass sehr intensiv an der Kantenerkennung mittels neuronalen Netzen geforscht wird und bereits einige sehr gute, spezifische Netzarchitekturen entstanden sind. In Abbildung 2 sind Precision Recall Kurven von den zwei trainierten Netzwerken wie auch von anderen Kantendetektionsverfahren zu sehen. Diese eignen sich für den Qualitätsvergleich der verschiedenen Verfahren.

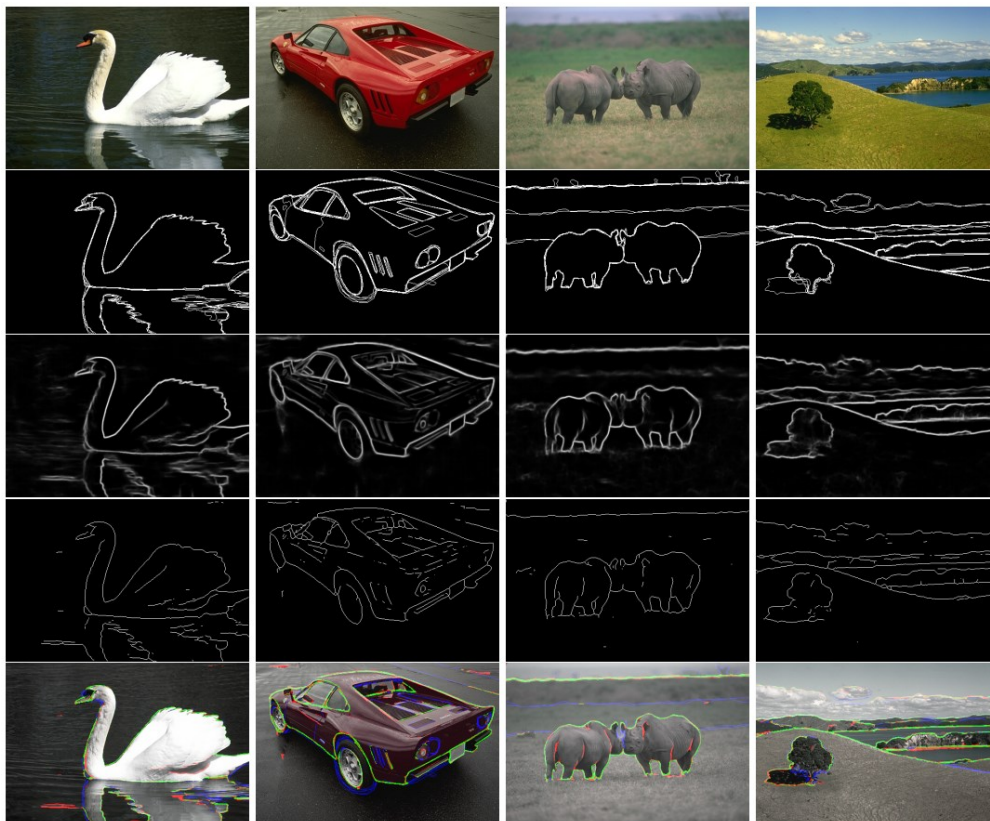


Abbildung 1: Bildvergleich mit Bildern die vom BDA-UNet-v1-Netzwerk generiert wurden. 1. Zeile: Originalbilder, 2. Zeile: überlagerter Ground Truth, 3. Zeile: geschätzte Kantenbilder, 4. Zeile: Binärbild mit verdünnten Kanten, 5. Zeile: überlagertes Vergleichsbild (Grün: Ground Truth und Netzwerk-Schätzung stimmt überein, Rot: Nur Netzwerk-Schätzung, Blau: Nur Ground Truth)

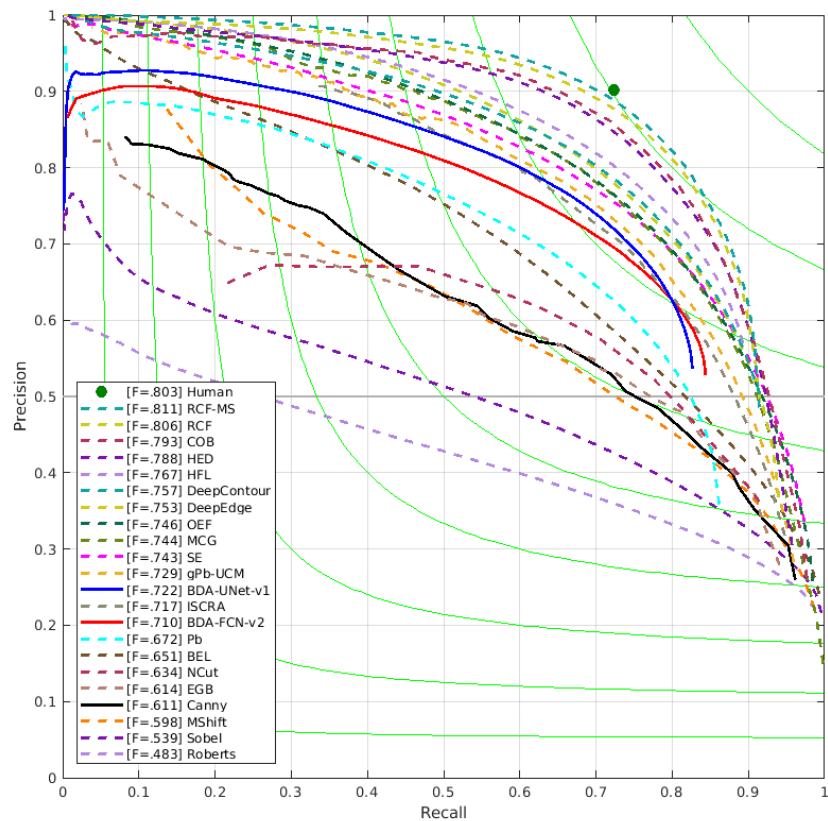


Abbildung 2: PR-Kurven der zwei finalen Netzwerke BDA-FCN-v2 und BDA-UNet-v1 im Vergleich zu anderen Kantendetektionsverfahren

## 4. Spezielle Herausforderungen

Eine grosse Herausforderung war die Einarbeitung in das Thema der neuronalen Netzwerke. Die richtigen und für das Projekt wichtigen Informationen zu erhalten war essenziell, aber nicht immer einfach. Dies liegt nicht zuletzt daran, dass sich das Fachgebiet der neuronalen Netzwerke in den letzten Jahren stark entwickelt hat. Das fachspezifische Know-how aufzubauen hat einen grossen Teil der Arbeit in Anspruch genommen.

Eine weitere Knacknuss ist die Vielzahl an übergeordneten Parametern der neuronalen Netzwerke, sogenannte Hyperparameter. Die Herausforderung ist, die Übersicht zu behalten, welcher Trainingsvorgang mit welchen Eigenschaften durchgeführt wurde. Für eine weitere Arbeit wäre deshalb eine von Anfang an klar definierte Naming-Convention hilfreich. Bezüglich Hyperparameter ist es ausserdem nicht trivial einzuschätzen, welcher Hyperparameter, welchen Einfluss auf andere Hyperparameter hat. Dieser Aspekt erfordert eine gewisse Erfahrung mit neuronalen Netzen, die während dieses Projektes gesammelt werden konnte.

Eine weitere Herausforderung ist, dass das Trainieren von neuronalen Netzwerken je nach Netz-Eigenschaften und Trainingsdaten lange dauern kann. Dies erfordert, dass allfällige Fehler respektive Bugs möglichst schnell aufgedeckt werden, um zeitintensive Trainingsvorgänge nicht mehrmalig wiederholen zu müssen. Regelmässiges und kritisches Begutachten der Metriken direkt nach einem Trainingsvorgang kann dabei helfen, allfällige Fehler schnell aufzudecken.

## 5. Ausblick

Die in dieser Arbeit entwickelten neuronalen Netze könnten in einer fortführenden Arbeit noch weiter verbessert werden. Dabei wäre es insbesondere interessant zu beobachten, inwiefern sich die hier entwickelten Netzwerke verhalten, wenn zusätzliche Trainingsdaten geschaffen werden würden. Die in dieser Arbeit evaluierten zusätzlichen Datensätze könnten dabei als Ausgangslage und in optimierter Form als Trainingsdaten dienen.

Die Arbeit hat jedoch auch gezeigt, dass bereits fortgeschrittene Modelle für die Kantenerkennung bestehen und sich die Frage stellt inwiefern noch bessere Architekturen überhaupt möglich sind. Grundsätzlich bietet sich hier an, an gewissen Aspekten der Kantenerkennung weiter zu forschen. So könnte untersucht werden, wie ein möglichst effizientes, ressourcensparendes Netzwerk aussieht, das auch auf Geräten mit niedriger Rechenleistung lauffähig ist oder wie, dass unterbrochene Liniensegmente durch neuronale Netzwerke ergänzt werden können.