



Marco Egger



Joel Hauser

|              |  |
|--------------|--|
| Diplomanden  | Marco Egger, Joel Hauser                 |
| Examinator   | Prof. Dr. Guido Schuster                 |
| Experte      | Gabriel Sidler, Eivycom GmbH, Zürich, ZH |
| Themengebiet | Artificial Intelligence                  |

## Embedded Real Time Computer Vision

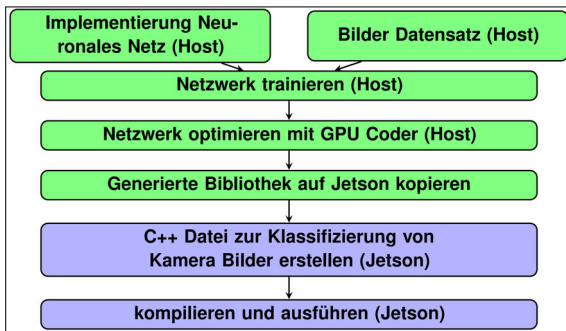


Nvidia Jetson TX2 (links), Xavier (rechts oben), Nano (rechts unten)  
Eigene Darstellung

**Einleitung:** Bei Computer Vision geht es darum, dass eine Maschine Bilder verstehen kann. Kombiniert man eine Kamera mit Deep Learning Algorithmen, verleiht man einem Rechner eine neue Form der Wahrnehmung. Anwendungen finden sich zum Beispiel in der Automatisierungs-, Verkehrs- oder Sicherheitstechnik.

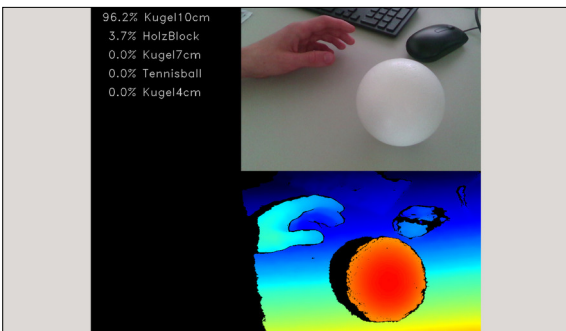
Computer Vision ist sehr rechenintensiv, was für gewisse Anwendungen eine Hürde darstellt. Jedoch dank den enormen Fortschritten in der Technik, öffnen sich neue Möglichkeiten. Leistungsfähige Rechner werden immer kleiner und somit sinken Platz-, Energie- und Kostenbedarf. Der Grafikkarten-Hersteller Nvidia veröffentlichte seit 2017 die Embedded Modelle Jetson TX2, Xavier und Nano speziell für KI-Anwendungen. Im Rahmen dieser Arbeit wollen wir diese eingebetteten Systeme für Computer Vision näher betrachten. Ausserdem schneidet diese Arbeit das Thema Tiefenbilder für Computer Vision an. Bei Tiefendaten handelt es sich um Bilder, bei denen der Wert eines Pixels der Distanz zwischen Kamera und fotografiertem Gegenstand entspricht. Mit diesen Daten erhält man weitere Informationen über die Umgebung und ermöglicht es Objekte zu unterscheiden, die man mit Farbbildern alleine nicht klassifizieren könnte. Zum Beispiel ist es in einem Farbbild schwierig, eine weisse runde Scheibe von einer weissen Kugel zu unterscheiden.

**Ziel der Arbeit:** Das Ziel dieses Projekts ist es, einen Ablauf zu schaffen, wie man von einem Neuronalen Netzwerk, das auf einem Personal Computer trainiert wurde, auf einen Jetson TX2 portieren kann. Es werden zwei Abläufe beschrieben: Beim ersten wird das neuronale Netzwerk mit Matlab und beim zweiten mit TensorFlow trainiert. Der Ablauf soll gut nachvollziehbar dokumentiert werden, damit dieser für spätere Anwendungen einfach übernommen und abgeändert werden kann. Zusätzlich soll der Jetson TX2 mit seinen Nachfolgermodellen Xavier und Nano verglichen werden, um das passende Gerät für die jeweilige Anwendung zu finden. Für Computer Vision sollen Bilder und Videostreams aufgenommen werden können. Dafür wird in dieser Arbeit mit der Realsense D415 Kamera von Intel gearbeitet. Diese Kamera ist fähig Farb- und Tiefenbilder mit bis zu 60 Frames pro Sekunde zu streamen. Ein neuronales Netzwerk soll so erweitert werden, dass es nicht nur Farb-, sondern auch Tiefenbilder verarbeiten kann. Eine Klassifizierung auf dem Nvidia Jetson soll in Echtzeit, dies bedeutet mit mindestens 10 Frames pro Sekunde, möglich sein.



Ablauf Portierung von Host (grün) auf Jetson (blau) mit Matlab  
Eigene Darstellung

**Ergebnis:** Es ist gelungen, einen Ablauf zu schaffen, in welchem man den Deep Learning Algorithmus auf den Jetson TX2 portieren kann. Dieser Ablauf ist Schritt für Schritt in den jeweiligen Projektordner beschrieben. Es kann je nach Anwendung Matlab oder Python's TensorFlow Framework verwendet werden. Auch wurde der Trade-off von Leistungsfähigkeit und Preisklasse der verschiedenen Jetson Maschinen verglichen, somit kann das bestmögliche Modell für die gewünschte Anwendung gebraucht werden. Wir haben es geschafft, die Intel Realsense D415 Kamera auf dem Nvidia Jetson in Betrieb zu nehmen und sie via Python und C++ anzusteuern. Eine mögliche Ansteuerung ist dokumentiert. Ausserdem haben wir ein neuronales Netz für Farb- und Tiefendaten erstellt, welches eine Kugel von einer runden Scheibe unterscheiden kann.



Computer Vision mit Farb- und Tiefendaten auf Nvidia Jetson TX2  
Eigene Darstellung