

Monocular Depth Estimation mit DeepLearning

Einleitung: Im Zeitalter von Robotern und Fahrzeugen, welche sich autonom bewegen, gewinnt das Thema der Distanzmessung zunehmend an Bedeutung. Während die meisten etablierten Methoden Distanzen mit Hilfe physikalischer Gegebenheiten (time of light, Triangulation, etc.) messen, verfolgt diese Arbeit einen neueren, innovativen Ansatz. Das Ziel ist es, mittels künstlicher Intelligenz aus einem RGB-Bild eine Tiefenkarte zu generieren.

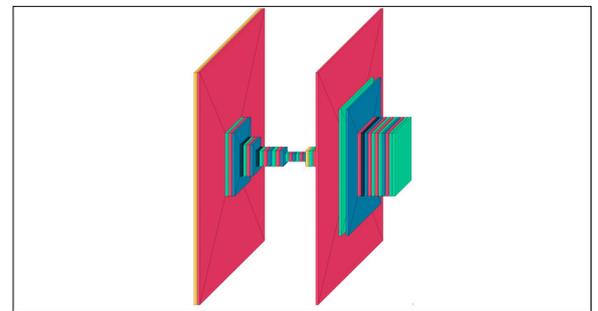
Die Aufgabe aus einem 2D-Bild Distanzen abzuschätzen, stellt für einen Menschen keine grosse Herausforderung dar. Durch das Verständnis der Szene ist es ein Leichtes zu entscheiden was nah und was fern ist. Dieses Verständnis soll via supervised learning mit einem neuronalen Netzwerk nachgebildet werden. Während die Distanz zu einem Objekt auf Stereobildern durch Triangulation einfach berechnet werden kann, ist die Distanzermittlung aus einem einzigen Bild weniger einfach. Für Triangulationsberechnungen reichen Kenntnisse über lokale Punkte. Für Monocular-Depth-Estimation sind sowohl ein lokales als auch ein globales Verständnis der kompletten Szene notwendig, um Proportionen von Objekten richtig einzuschätzen und daraus die Entfernung abzuleiten.

Vorgehen: Um das Ziel zu erreichen, wurden zwei unterschiedliche Modelarchitekturen trainiert, wobei der NYUv2 Datensatz mit Innenraumaufnahmen verwendet wird. Die erhaltenen Resultate wurden anschliessend mit etablierten Bewertungsmethoden auf deren Genauigkeit untersucht und mit bestehenden Monocular-Depth-Estimation Ansätzen verglichen. Durch den Aufbau eines Prüfstands wurden zudem Schwachstellen und Stärken des

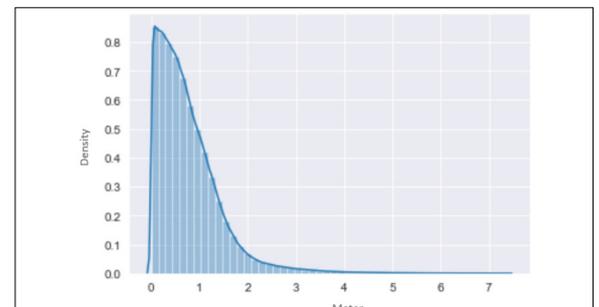
erarbeiteten Ansatzes gegenüber sensorbasierten Systemen evaluiert und ausgewertet.

Ergebnis: Die verwendeten Architekturen ermöglichen es Distanzen im Bereich von 0.1 Meter – 10 Meter vorherzusagen und bewegen sich mit ihrer Genauigkeit im Bereich anderer Monocular-Depth-Estimation Ansätze. Somit bilden die erarbeiteten

Autoencoder Netzwerk Architektur
Eigene Darstellung



Fehlerverteilung der Tiefenschätzung
Eigene Darstellung



Tiefenschätzung mit Hilfe einer monokularen Kamera
Eigene Darstellung



Referent
Prof. Dr. Klaus Frick

Korreferent
Prof. Dr. Stefan Rinner

Themengebiet
Computational
Engineering