

Einleitung

Eine, besonders für die industrielle Anwendung, sehr interessante Fragestellung ist, wie sich Daten von unterschiedlichen Sensoren in einem lernenden System verbinden lassen (Sensorfusion). Speziell wenn die Sensoren Daten sehr unterschiedlicher Dimensionalität (Bsp. Kraftsensor vs. Kamera) liefern, ist die Fusion anspruchsvoll. Während der Datenverlust beim Ausfall eines Pixels in einem Bild sehr gering ist, so wäre der Ausfall eines Sensors mit niedrigerer örtlicher Auflösung deutlich dramatischer.

Convolutional Neural Networks (CNN) haben sehr erfolgreiche Anwendung in der Verarbeitung von Bildern und Audiodaten gefunden, in der Sensorfusion sind sie jedoch noch deutlich weniger repräsentiert. Da diese Netzwerke generell eine niedrigdimensionale Darstellung der Eingangsdaten finden, sollte es möglich sein innerhalb des Netzes die Daten von unterschiedlichen Sensoren so in das Netz einzuführen, dass das Problem der unterschiedlichen Dimensionalität der Originaldaten gelöst wird. Visualisiert man sich ein CNN als Aneinanderreihung von mehreren Convolutional Layern, die pro Schicht die Dimensionalität der Eingangsdaten verringern, so ist es naheliegend die Sensordaten dort den Schichten zuzuführen, an denen die Daten des Sensors mit der größten Dimensionalität soweit reduziert wurden, dass sie der Dimensionalität der anderen Sensordaten entsprechen. Convolutional Neural Networks sind für diese Aufgabe besonders interessant, da es sich bei den relevanten Sensoren um jene dreht, die ein örtliches Bild der Umgebung liefern und die Daten somit räumlich korrelieren.

Beispiele:

- Kameras
- Taktile Näherungssensoren
- (Ultraschall-Arrays)

Hypothese

Die Dimensionsreduktion durch Convolutional Neural Networks kann dazu verwendet werden, Sensordaten unterschiedlicher Dimensionalität zu fusionieren um ein Klassifikationsproblem genauer lösen zu können, als es mit den Sensordaten einzeln möglich wäre.

Der Kerngedanke hierbei ist es, Sensordaten dort den Feature Maps des CNN zuzuführen, wo die Dimensionalität der Feature Maps der der Sensordaten entspricht und somit eine ausgewogene Informationsdichte über das komplette CNN zu schaffen.

Aufgabe

Nach einer umfangreichen Recherche zum Stand der Technik, in Bezug auf Sensorfusion mit Neuronalen Netzen, soll mit unterschiedlichen Sensoren (Kamera, taktile Näherungssensoren, Kraftsensoren) die Hypothese getestet werden, dass das Einspeisen von Daten unterschiedlicher Dimensionalität in tieferen Schichten eines CNN, zu einer Verbesserung von Klassifikations-Ergebnissen führt.

Um die nötigen Daten so gering wie möglich zu halten, soll hierzu ein Neuronales Netz verwendet werden, für das ein Checkpoint (Stichwort Transfer Learning) vorliegt. Hierzu eignen sich besonders Netze aus der TensorFlow Slim Bibliothek.

Bsp.:

- | | |
|-------------|-------------|
| - Alexnet | - ResNet |
| - Inception | - VGG |
| | - MobileNet |