

- Masterarbeit -
oder
- Vertiefende F&E Studien -
Benchmarking von Deep Learning Hardware

Motivation & Ziel

Deep Learning (DL) ist ein wichtiger Bestandteil in vielen Gebieten aktueller Forschung. Während Neuronale Netze (NN) i.d.R. auf leistungsstarken Serversystemen trainiert werden, stellt sich zu Beginn vieler Projekte meist die Frage, wie leistungsstark Hardware für Inference sein muss. Gerade im Bereich der Embedded Systeme ist dies keine einfach zu beantwortende Frage. Im Rahmen der Studien soll eine Möglichkeit für geeignete Standard-Benchmarks für Inference bzgl. Throughput, Latenz und Modellqualität auf Embedded Systemen untersucht werden. Zentral ist dabei die Frage, welche NN sich für diese Benchmarks eignen. Weiterhin soll untersucht werden, ob diese Benchmarks extrapoliert werden können, z.B. bei geänderter Auflösung von Bildern. Dies ermöglicht zu Projektbeginn Hardwareanforderungen in Abhängigkeit der Datenbeschaffenheit besser abschätzen zu können.

Aufgaben

- Konzeption einer Deep Learning Infrastruktur für Benchmarking auf Embedded Systemen (NVIDIA Jetson Nano, X1, Xavier)
- Auswahl von geeigneten Datensätzen und Parametern zur Durchführung von Benchmarks
- Untersuchung der Generalisierung bzgl. Extrapolation für geänderte Eingabedaten

Anforderung

Exzellente Kenntnisse in:

- Programmieren (Python & C/C++)
- Deep Learning und Frameworks (TensorFlow, PyTorch, CUDA)

Grundkenntnisse in:

- Linux für Embedded Systeme

Datum

Nach Absprache (anschließende Masterarbeit ist möglich)

Kontakt



stephan.gimbel@h-da.de
h_da
Faculty of Computer Science