|  |  |  |
| --- | --- | --- |
| **MA 3062** | | |
| Untersuchung und prototypische Umsetzung eines Lifelong Deep Neural Network Algorithmus | | |
| **Simon Kamm** | | |
|  | | |
| **Konzeption** | | |
|  | Prüfer: | Prof. Dr.-Ing. Michael Weyrich |
|  | Betreuer: | Benjamin Maschler, M.Sc. |
| Start: 29.04.2019 | | Abgabe: 29.10.2019 |
|  | |  |

**Dokument Versionsverwaltung**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Version | Autor | QS | Datum | Status | Änderungen |
| 0.1 |  |  | 27.05.19 | in Bearb. | Erstellung |
|  |  |  |  |  |  |

*Bei der Erstellung eines Dokuments ist folgendes zu beachten:*

* *Wurden alle Anforderungen in der Konzeption umgesetzt?*
* *Ist die Beschreibung der Konzeption detailliert genug?*
* *Entspricht das Ergebnis den Erwartungen des Auftraggebers?*
* *Ist auf Basis der erstellten Konzeption die Realisierung eines Prototyps möglich?*

# Inhaltsverzeichnis

0 Inhaltsverzeichnis 2

1 Beschreibung der Konzeption 3

1.1 Modul A 3

1.2 Modul B 3

2 Zusammenfassung 4

3 Literaturverzeichnis 5

# Beschreibung der Konzeption

In diesem Dokument wird die konkrete Konzeption des L DNN Algorithmus beschrieben. Dafür werden unterschiedliche mögliche Architekturen untersucht und verglichen, und schließlich eine konkrete Architektur ausgewählt, welche im weiteren Verlauf der Arbeit implementiert wird.

Der L DNN Algorithmus besteht aus zwei Modulen, dem Modul A und Modul B. Details zum Aufbau und den grundlegenden Ideen des L DNN Algorithmus sind in den „Grundlagen“ genannt. Für eine Übersicht ist in Abbildung 1 die Architektur des L DNN Algorithmus dargestellt.

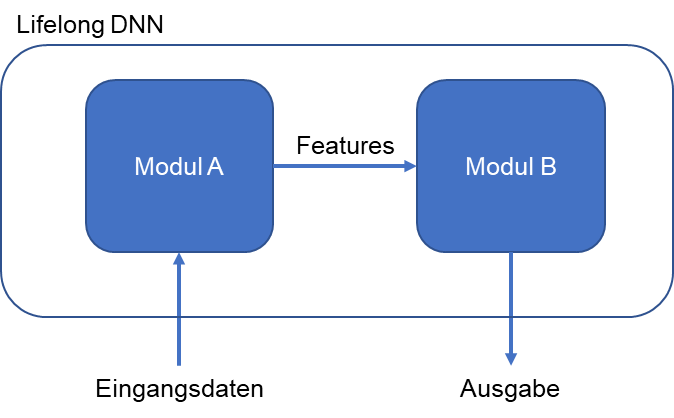


Abbildung : Graphische Darstellung des L DNN A

Im Folgenden werden nun unterschiedliche Architekturen für die einzelnen Module A und B verglichen und bewertet für die in dieser Arbeit gestellten Anforderungen.

## Modul A

Modul A ist das langsame oder nicht lernende Modul innerhalb des L DNN Algorithmus. Es ist ein klassisches DNN, welches zur Extraktion von relevanten Features auf Basis von Eingangsdaten genutzt wird. Im Rahmen dieser Arbeit werden vortrainierte, feste DNN-Architekturen genutzt, da das Trainieren solcher DNN-Architekturen sehr zeit- und rechenaufwändig ist.

Es gibt eine Vielzahl an DNN-Architekturen, die heutzutage zur Feature-Extraktion eingesetzt werden. Jede Architektur erfüllt dabei unterschiedliche Anforderungen, beziehungsweiße nutzt unterschiedliche Ansätze, die für den einen oder anderen Anwendungsfall besser geeignet sind. Für das Modul A innerhalb des L DNN Algorithmus ist zum einen wichtig, dass die extrahierten Features eine bestmögliche Klassifikation erlauben. Zum anderen soll das System auch fähig sein in Echtzeitsystemen auf mobilen Endgeräten (z.B. Smartphone) zu funktionieren [1], [2]. Dafür soll das Modul A Eingangsdaten schnell verarbeiten können und gleichzeitig darf der Speicher- und Rechenbedarf des Netzwerkes nicht zu groß sein, da sowohl Speicher als auch Rechenleistung auf mobilen Endgeräten limitiert vorhanden ist.

Im Folgenden werden bekannte Klassifikator-Netzwerke auf Basis von DNN Architekturen vorgestellt und diese verglichen auf deren Einsetzbarkeit in dieser Arbeit. Als Grundlage für den späteren Vergleich und genannte Metriken wird der Bilddatensatz ImageNet [3] genutzt. Mit ca. 1,2 Millionen Bildern aus 1000 verschiedenen Klassen stellt ImageNet ein sehr komplexes Problem dar. Vortrainierte Netzwerke, welche zur Extraktion von Features genutzt werden sind heutzutage auf diesem Datensatz vortrainiert, da er aufgrund seiner großen Anzahl an Trainingsbildern und Klassen eine gute Generalisierung der extrahierten Features des Netzwerks erzeugt.

### AlexNet

AlexNet war eine der ersten DNNs welche die Klassifikationsgenauigkeit auf dem ImageNet Datensatz signifikant erhöhte. Im Rahmen der ILSVRC (*ImageNet Large Scale Visual Recognition Competition*) erreichte AlexNet einen Top-5 Klassifikationsfehler von 15,3%. Der zweitplatzierte dieses Wettkampfes erreichte lediglich einen Top-5 Klassifikationsfehler von 26%. Dadurch konnte eine erhebliche Steigerung der Klassifikationsperformanz erzielt werden. Dabei ist AlexNet im Vergleich zu aktuellen Architekturen sehr simpel, mit lediglich fünf Convolutional Schichten, verschiedenen Max-Pooling Schichten sowie drei Fully Connected Schichten [4], [5]. Neuerungen im Vergleich zu damaligen DNNs war der Einsatz von ReLU-Aktivierungsfunktionen sowie die Regularisierung mithilfe von Dropout. Beide Methoden sind seitdem in nahezu allen DNN-Architekturen im Einsatz. Das Netzwerk besteht aus 60 Millionen Parametern und 650.000 Neuronen. In Abbildung 2 ist die Modellarchitektur des AlexNet graphisch dargestellt, bei welcher die fünf Convolutional und die drei Fully Connected Schichten zu sehen sind.

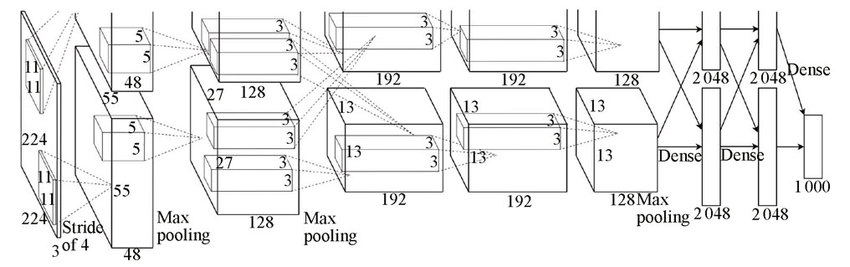


Abbildung : Modelarchitektur des AlexNet [5]

Die *Default*-Größe eines Parameters (Floating Point oder Integer) in TensorFlow ist mit 32-bit angegeben [6]. Unter der Annahme dieser Parametergröße benötigt AlexNet mit 60 Millionen Parametern 240 MB Speicher zur Sicherung des Modells.

### VGG

## Modul B

# Zusammenfassung

*Hier soll das Ergebnis der Konzeption dargestellt werden.*

# Literaturverzeichnis

|  |  |
| --- | --- |
| [1] | M. Luciw, S. Olivera, A. Gorshechnikov, J. Wurbs, H. M. Ames und M. Versace, „Systems and Methods to enable Continual, Memory-Bounded learning in Artificial Intelligence and Deep Learning Continuously operating Applications across networked Compute Edges“. United States of America Patent US 2018/0330238 A1, 15 November 2018. |
| [2] | Neurala Inc., „Lifelong Deep Neural Networks - Tech Summary,“ [Online]. Available: https://info.neurala.com/hubfs/docs/ Neurala\_LifelongDNNWhitepaper.pdf. [Zugriff am 7 Mai 2019]. |