

Machbarkeitsstudie: Smart Warehouse

Echtzeit-Objektdetektoren im Vergleich

Studienarbeit

im Rahmen der Prüfung zum
Bachelor of Science (B.Sc.)

des Studienganges Angewandte Informatik
an der Dualen Hochschule Baden-Württemberg Karlsruhe

von

Felix Hausberger und Robin Kuck

Oktober 2019 - Mai 2019

-Sperrvermerk-

Abgabedatum:	18. Mai 2020
Bearbeitungszeitraum:	30.09.2019 - 18.05.2020
Matrikelnummer, Kurs:	2773463, 4409176, TINF17B2
Ausbildungsfirma:	SAP SE Dietmar-Hopp-Allee 16 69190 Walldorf, Deutschland
Gutachter an der DHBW:	PD Dr.-Ing. Markus Reischl

Eidesstattliche Erklärung

Wir versichern hiermit, dass wir unsere Studienarbeit mit dem Thema:

Machbarkeitsstudie: Smart Warehouse

gemäß § 5 der „Studien- und Prüfungsordnung DHBW Technik“ vom 29. September 2017 selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt haben. Die Arbeit wurde bisher keiner anderen Prüfungsbehörde vorgelegt und auch nicht veröffentlicht.

Wir versichern zudem, dass die eingereichte elektronische Fassung mit der gedruckten Fassung übereinstimmt.

Karlsruhe, den 26. Oktober 2019

Gez. Felix Hausberger und Robin Kuck

Hausberger, Felix und Kuck, Robin

Abstract

- English -

In this thesis the object detectors *Regional Convolutional Neural Networks*, *You Only Look Once* and *Single Shot MultiBox Detector* are compared for precision, reactivity and training behaviour and examined for their potential for industrial use. The background scenario of the *Smart Warehouse* offers live video data of a drone with goods in a warehouse, which are to be classified and localized in real time. In the future, this should make it possible to carry out inventories and inventory analyses of a warehouse in a time- and cost-efficient manner conserving resources.

The goal of this feasibility study is to find out whether the *Smart Warehouse* scenario is technically feasible and economically reasonable. In addition, the focus is also on the object detectors themselves and their differences in architecture and behavior in the *Smart Warehouse* environment.

Abstract

- Deutsch -

In dieser Arbeit werden die Objektdetektoren *Regional Convolutional Neural Networks*, *You Only Look Once* und *Single Shot MultiBox Detector* nach Präzision, Reaktionsvermögen und Trainingsverhalten miteinander verglichen und auf deren Potential zum industriellen Einsatz untersucht. Das Hintergrundscenario des *Smart Warehouses* bietet dabei Live-Video Daten einer Drohne mit Warengegenständen in einem Warenhaus, die in Echtzeit klassifiziert und lokalisiert werden sollen. Dadurch sollen in Zukunft in der Industrie Inventuren und Bestandsanalysen eines Warenhauses zeit- und kostengünstig sowie ressourcenschonend ermöglicht werden können.

Diese Machbarkeitsstudie hat zum Ziel herauszufinden, ob das Szenario des *Smart Warehouse* technisch umsetzbar sowie wirtschaftlich sinnvoll ist. Zusätzlich liegt der Fokus ebenso auf den Objektdetektoren selbst und deren Unterschiede hinsichtlich Architektur und Verhalten im *Smart Warehouse* Umfeld.

Inhaltsverzeichnis

Abkürzungsverzeichnis	VI
Abbildungsverzeichnis	VII
Formelverzeichnis	VIII
Listenverzeichnis	IX
0 Vorwort	1
1 Einführung	2
1.1 Forschungsumfeld	2
1.2 Problemstellung und Motivation	3
1.3 Vorgehensweise und Zielsetzung	3
2 Grundlagen und Forschungsstand	5
2.1 Neuronale Netze	6
2.2 Hyperparameter	6
2.3 Objektdetektoren	6
2.4 Einsatz in der Industrie	6
2.5 Verwendete Bibliotheken	6
2.6 Compute Unified Device Architecture	6
3 Konzeption	8
3.1 Bewertungskriterien	8
3.2 Initialer Vergleich der Objektdetektoren	8
3.3 Echtzeitumgebung	8
4 Realisierung	9
4.1 Trainieren der Objektdetektoren	9
4.2 Drohnen Anbindung	9
4.3 Dashboard Entwicklung	9
5 Ergebnisse	10

6	Bewertung	11
7	Zusammenfassung und Ausblick	12
	Literaturverzeichnis	X
A	Anhang	XI

Abkürzungsverzeichnis

ANN	Artificial Neural Network
CNN	Convolutional Neural Network
COCO	Common Objects in Context
PascalVOC	Pascal Visual Object Classes
R-CNN	Regional Convolutional Neural Network
SSD	Single Shot MultiBox Detector
YOLO	You Only Look Once

Abbildungsverzeichnis

Formelverzeichnis

Listenverzeichnis

0. Vorwort

Besonderen Dank ist an unseren Betreuer PD Dr. -Ing. Markus Reischl auszusprechen, ohne den die folgenden Forschungsergebnisse nicht zustande gekommen wären. Auch dem Informatik Labor unter Enrico Hühneborg der DHBW ist für die nötige finanzielle Unterstützung zum Erwerb der Drohne zu danken.

1. Einführung

1.1. Forschungsumfeld

Einen Teilbereich des maschinellen Lernens (engl.: machine learning) stellt das *Deep Learning* dar, welches auf künstlichen neuronalen Netzen (engl.: artificial neural networks) (ANNs) basiert [1, S. 253]. Unter einer Vielzahl von Typen von ANNs wie Autoencodern, Deep Boltzmann Machines oder rekurrenten neuronalen Netzen befindet sich ebenso die Klasse der *Convolutional Neural Networks* (CNNs), welche hauptsächlich zur Lösung von Klassifikationsproblemen in der Audio-, Text- und Bildverarbeitung genutzt werden [2].

Ein Forschungsfeld im *Deep Learning* stellen Objektdetektoren dar, welche basierend auf CNNs neben Bildklassifikationsproblemen ebenso in der Lage sind, Lokalisationsprobleme zu lösen. Solchen Objektdetektoren werden in der heutigen Zeit immer mehr Bedeutung zugesprochen angesichts neuer Herausforderungen wie autonomen Fahren, automatisierter industrieller Verarbeitung oder aber auch staatlicher Überwachung. Verschiedene Ansätze werden zur Realisierung von Objektdetektoren verwendet, unter anderem Netzarchitekturen wie *You Only Look Once* (YOLO), *Single Shot MultiBox Detector* (SSD) oder *Regional Convolutional Neural Networks* (R-CNN).

Gerade in Zeiten des industriellen Wandels in Richtung *Industrie 4.0* können solche Objektdetektoren ein großes Optimierungspotential für bestehende Industrieszenarien bieten, beispielsweise in der Lagerhaltung und Logistik. Kombiniert mit einer autonomen Drohne können Objektdetektoren es ermöglichen, ohne menschliche Hilfe Inventuren und Bestandsprüfungen in einem Lager- oder Warenhaus durchzuführen. Start-up Unternehmen wie *doks. innovation* werben bereits mit ähnlichen Lösungen, die 80% Zeiteinsparung und 90% Kostensenkung versprechen [3]. Lösungen wie *inventAIRyX* beschränken sich allerdings speziell auf Lagerhäuser, in denen die verpackten Waren mittels Sensoren identifiziert werden, was Großhändler mit Warenhäusern wie *Baumarkt* oder *Selgros* ausschließt. Statt Waren mittels RFID Chips oder Barcodes zu identifizieren, soll in dieser Arbeit der Einsatz von Objektdetektoren für dieses Szenario evaluiert werden.

Wie sich die unterschiedlichen Objektdetektoren unter Echtzeitvoraussetzungen im Betrieb verhalten, soll anhand des Industriebeispiels *Smart Warehouse* innerhalb dieser Arbeit untersucht werden.

1.2. Problemstellung und Motivation

Das *Smart Warehouse* beschreibt ein Warenhaus, welches unter Einsatz einer Drohne in der Lage sei soll, Inventuren und Bestandsprüfungen weitgehend ohne menschliche Hilfe durchzuführen. Das Live-Bild der Drohne soll von den Objektdetektoren dazu genutzt werden, Warengegenstände zu lokalisieren und klassifizieren.

Neben der Frage, ob ein solches Industrieszenario überhaupt umsetzbar und wirtschaftlich sinnvoll ist, sollen die Objektdetektoren in diesem Anwendungsszenario nach verschiedenen Kriterien miteinander verglichen und beurteilt werden. Diese Kriterien lassen sich hauptsächlich in die Kategorien Präzision, Reaktionsvermögen und Trainingsverhalten untergliedern und werden später genauer eingeführt. Dadurch lassen sich Aussagen darüber treffen, ob nach dem momentanen Forschungsstand um Objektdetektoren solche das Potential bieten, industriell eingesetzt zu werden.

Falls die Machbarkeitsstudie des *Smart Warehouse* glückt, so kann der Industrie ein kostengünstiges, zeitsparendes und ressourcenschonendes Modell zur Inventurverwaltung eines Warenhauses angeboten werden.

1.3. Vorgehensweise und Zielsetzung

Zunächst muss sich mit den theoretischen Grundlagen von CNNs und Objektdetektoren auseinander gesetzt werden. Hierzu ist zunächst eine Einführung in neuronale Netz erforderlich, darunter zu Perzeptronen, dem Gradientenverfahren, dem Backpropagation Algorithmus und Hyperparametern zum Trainieren eines neuronalen Netzes.

Nachdem kurz auf den Grundbaustein moderner Objektdetektoren eingegangen wird, den CNNs, können anschließend die Funktionsweisen und Architekturen der drei miteinander verglichenen Objektdetektoren *YOLO*, *SSD* und Detektoren der *R-CNN* Familie erläutert werden. Hier ist zu bemerken, dass unterschiedliche Evolutionsstufen der drei

Detektoren zu betrachten sind.

Auch bisherige industrielle Einsatzfelder von Objektdetektoren sollen betrachtet werden, bevor technische Grundlagen zum Aufsetzen und Trainieren der drei Modelle näher betrachtet werden.

In der Konzeptionsphase sollen zunächst die Vergleichskriterien eingeführt werden und deren Metriken anschließend für initiale Benchmarkdatensätze für jeden Objektdetektor ermittelt werden. Hierzu wird auf die Datensätze *Pascal Visual Object Classes* (Pascal-VOC), *Common Objects in Context* (COCO) und ImageNet zurückgegriffen. Anschließend wird der Datensatz für das *Smart Warehouse* Szenario eingeführt.

In der Realisierung werden die Herausforderungen zur Steuerung und Anbindung der Drohne betrachtet und zudem die Objektdetektoren auf die realen Datensätze trainiert. Auch die Entwicklung der Webapplikation zur Visualisierung des Live-Bildes und der erkannten Objekte wird Bestandteil dieses Kapitels sein. Die Ergebnisse der Realisierungsphase werden im folgenden Kapitel dargestellt.

Ziel der Arbeit ist es Aussagen über die Fähigkeit von Objektdetektoren zum Einsatz in der Industrie zu treffen, indem eine Bewertung der Verhaltensweisen der Objektdetektoren nach den eingeführten Bewertungskriterien durchgeführt wird. Auch wirtschaftliche Gesichtspunkte werden in diesem Kapitel nicht außer Acht gelassen.

Zuletzt wird das Wesen der Arbeit nochmals kurz zusammengefasst und anschließend auf mögliche Verbesserungen und Ausblicke in die Zukunft aufmerksam gemacht.

2. Grundlagen und Forschungsstand

2.1. Neuronale Netze

2.2. Hyperparameter

2.3. Objektdetektoren

2.3.1. Convolutional Neural Networks

2.3.2. Mask Regional Convolutional Neural Network

2.3.3. You Only Look Once

2.3.4. Single Shot MultiBox Detector

2.4. Einsatz in der Industrie

2.5. Verwendete Bibliotheken

2.5.1. PyTorch

2.5.2. OpenCV

2.5.3. Django

2.6. Compute Unified Device Architecture

- Methoden

- Literaturrecherche
- Herausarbeitung des Neuheitswertes

3. Konzeption

3.1. Bewertungskriterien

- Kriterien für spätere Bewertung einführen

3.2. Initialer Vergleich der Objektdetektoren

- Verfahren incl. Variantendiskussionen
- Funktionsweise an Benchmarkdatensatz erklären
- Parameter d. Methode diskutieren und u.U. bestimmen

3.3. Echtzeitumgebung

- Richtigen Datensatz beschreiben

4. Realisierung

4.1. Trainieren der Objektdetektoren

4.2. Drohnen Anbindung

4.3. Dashboard Entwicklung

- Herausforderungen bei der Drohne
- Trainieren auf die richtigen Datensätze
- Webapplikation

5. Ergebnisse

- Ergebnisse auswerten

6. Bewertung

- Ergebnisse bewerten und diskutieren
- Was liefert die Arbeit, was bisher noch nicht bekannt war?

7. Zusammenfassung und Ausblick

- Klare Darstellung, was die Arbeit geliefert hat
- Ca. 2-4 Anpunkte: Zukünftige Ziele

Literaturverzeichnis

- [1] Aurélien Géron. *Machine Learning mit Scikit-Learn & TensorFlow: Konzepte, Tools und Techniken für intelligente Systeme: Übersetzung von Kristian Rother*. 1. Aufl. Heidelberg: dpunkt.verlag GmbH, 2018.
- [2] MathWorks. *Deep Learning: Drei Dinge, die Sie wissen sollten*. MathWorks, 2019. URL: <https://de.mathworks.com/discovery/deep-learning.html> (Einsichtnahme: 12.10.2019).
- [3] doks. innovation. *inventAIRyX*. doks. innovation Homepage, 2019. URL: <https://www.doks-innovation.com/inventairy-x/> (Einsichtnahme: 26.10.2019).

A. Anhang

Abbildungen