Berliner Hochschule für Technik Berlin
Fachbereich VI – Informatik und Medien

Bachelorarbeit

Erzeugung von Bildern mittels Neuronalen Netzen

Stefan Berger
Medieninformatik
Matrikel-Nr. 854184

Berlin, 6. April 2021

Betreut von Prof. Dr. F. Gers

Zusammenfassung

Im Experiment und im Inhalt dieser Bachelorarbeit werden die Fragen
1. ...
2. ...
...
beantwortet.

Inhaltsverzeichnis

1 Einleitung

1.1 Motivation

Text

1.2 Ziel der Arbeit

Text

1.3 Vorherige Arbeiten

2 Entwicklungsumgebung

2.1 Ubuntu Linux

2.2 Python

2.3 Tensorflow

2.4 CUDA

3 Neuronale Netze

3.1 Logistische Regression

3.2 Deep Neural Networks

3.3 Convolutional Neural Networks

Viele Computer-Vision-Aufgaben können durch ein Convolutional Neural Network (CNN) mit hoher Genauigkeit erfüllt werden. Eine Konvolution oder Faltung ist eine Operation auf zwei Funktionen mit je einem reellen Funktionsargument [1, S. 327-329].

Für Bilddaten wendet ein CNN diese Operation typischerweise auf zwei dreidimensionale Matrizen an, nämlich einerseits auf die Eingabedaten und andererseits auf einen sogenannten Filter oder auch Kernel. Die Eingabedaten sind in der ersten Schicht des Netzes die RGB-Pixelinformationen und in allen weiteren konvolutionalen Schichten die Ausgabe der vorherigen Schicht. Ein Filter ist eine Anzahl von trainierbaren Parametern, in diesem Fall auch eine . Beide Matrizen haben also die Form ($H\ x\ W\ x\ C$), wobei H die Höhe, W die Breite und C die RGB-Farbwerte repräsentiert.

Jede der drei Matriximensionen variiert üblicherweise zwischen den verschiedenen Schichten des Netzes. In fast allen CNNs ([1], [2], [3]) nimmt die Kardinalität zunächst ab. Die Reduktion kann durch die Konvolution selbst entstehen oder durch Pooling-Schichten. Beim Pooling werden aus benachbarten Matrixkoeffizienten meist das Maximum, seltener der Durchschnitt oder andere Aggregierungen gebildet. Auf diese Weise wird das neuronale Netz darauf trainiert die relevanten Informationen zu extrahieren. [1]

Den konvolutionalen Schichten folgt in einigen Anwendungsfällen eine voll vernetzte Schicht (engl. Fully Connected Layer, FC), in der für jedes Neuron mit jedem Neuron der vorherigen Schicht eine Verbindung besteht. Besonders für die Bilderkennung ist diese Architektur gut geeignet. Die Ausgabe des neuronalen Netzes ist dann ein Vektor, beispielsweise von Wahrscheinlichkeitswerten für das Vorhandensein bestimmter Objekte und gegebenenfalls Bildkoordinaten der erkannten Objekte. Im Fall der Bildgenerierung ist die Ausgabe aber wieder eine Matrix von RGB-Pixelinformationen in der Form ($H\ x\ W\ x\ 3$).

3.4 U-Net-Architektur

Die bis hierhin beschriebenen neuronalen Netze besitzen eine gradlinige Struktur, in der die Ausgabe einer Schicht nur an die nächste Schicht übergeben wird. Bei zunehmender Anzahl der Schichten verbessert sich die Performance neuronaler Netze mit diesem Aufbau zunächst, aber verschlechtert sich bei zu vielen Schichten wieder. In einem Residual Neural Network

(ResNet) verhindern zusätzliche Verbindungen zwischen nicht direkt aufeinanderfolgenden Schichten diesen Performanceverlust.

Ein U-Net ist eine spezielle Form eines ResNets. Es hat eine annähernd symmetrische Struktur, in der sich die Kardinalitäten der Matrizen zuerst verringern und anschließend wieder erhöhen. U-Nets erzielen selbst mit wenigen Trainingsdaten gute Ergebnisse und benötigen dafür vergleichsweise wenig Rechenleistung. [4]

3.5 Image- To Image-Translation

4 Durchführung des Experiments

4.1 Vorbereitung der Eingabedaten

4.2 Anwendung herkömmlicher Shader

4.3 Hyperparameter

4.4 Performancebeobachtungen

4.5 Zusammenfassung

Literatur

1. GOODFELLOW, I. u. a.: *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.
2. RONNEBERGER, O. u. a.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2015, Bd. 9351, S. 234–241. LNCS. Auch verfügbar unter: http://lmb.informatik.uni-freiburg.de/Publications/2015/RFB15a. (available on arXiv:1505.04597 [cs.CV]).
3. ISOLA, P. u. a.: *Image-to-Image Translation with Conditional Adversarial Networks*. 2018. Abgerufen unter arXiv: 1611.07004 [cs.CV].
4. HE, K. u. a.: *Deep Residual Learning for Image Recognition*. 2015. Abgerufen unter arXiv: 1512.03385 [cs.CV].

Bildnachweis

Abkürzungs- und Symbolverzeichnis

Abkürzungen

AC	Air Compressor, Luftverdichter
APH	Air Preheater, Luftvorwärmer
CC	Combustion Chamber, Brennkammer
EXP	Expander
HRSG	Heat Recovery Steam Generator, Abhitzekessel

Lateinische Symbole

c	Spezifische Kosten je Exergieeinheit, €/J_{ex}
$\dot{C}$	Kostenstrom, €/h
CC	Kapitalgebundene Kosten, €
cf	Capacity Factor, Jährliche Auslastung, –
e	Spezifische Exergie, J/kg
$\bar{e}$	Spezifisch molare Exergie, J/mol
$\dot{E}$	Exergiestrom, W
f	Exergoökonomischer Faktor, –
fc	Spezifische Brennstoffkosten, €/J
FC	Brennstoffkosten, €
h	Spezifische Enthalpie, J/kg
$\dot{H}$	Enthalpiestrom, W
HHV	Brennwert, J/kg

Griechische Symbole

Δ	Differenz
ε	Exergetischer Wirkungsgrad, –
η_s	Isentroper Wirkungsgrad, –
κ	Isentropenexponenten, –
λ	Luftzahl, –

Hoch- und tiefgestelle Indizes

0	Referenzpunkt, Thermodynamische Umgebung
a	Avarage, Mittlere
D	Destruction, Vernichtung
F	Fuel, Brennstoff, Aufwand
net	Netto

Abbildungsverzeichnis

Tabellenverzeichnis

Anhang