

Artificial Intelligence

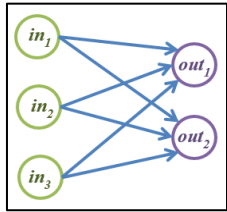
Raphael Appenzeller, Sophia Erni, Patrick Knüppel, Andrino Meli

Was ist Artificial Intelligence?

Artificial Intelligence bedeutet Künstliche Intelligenz. Das Ziel dieses Projekts war, ein „Convolutional Neural Network“ zu programmieren. Dabei handelt es sich um ein lernendes Programm, oder genauer um ein Netzwerk, welches Objekte in Bildern erkennt.

Konzept

Dieses Netzwerk soll Bilder von handgeschriebenen Ziffern erkennen. Die Programmausgabe ist also eine Ziffer zwischen 0 und 9. Die einzelnen Rechenschritte bis zum Resultat sind in der farbigen Bilderreihe



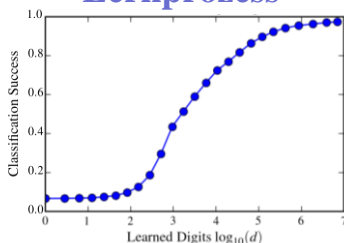
dargestellt. Darin kann der Weg des Bildes vom Input bis hin zur Bestimmung der Ziffer nachvollzogen werden. Die Skizze links veranschaulicht die Grundidee. Die Kreise stehen für die Pixel. Die blauen Pfeile verdeutlichen die durch Lernen anpassbaren Parameter im Netzwerk. Sie bestimmen, wie stark bestimmte Pixel gewichtet sind. Durch die Multiplikation der Inputpixel (grüne Kreise) mit den

Gewichten und durch Aufsummieren erhält man die Outputpixel (violette Kreise).

Wie erkennt das Netzwerk Ziffern?

Die Bestimmung von Objekten mithilfe von neuronalen Netzwerken beruht auf Training respektive Lernen. Dafür wird ein grosser Datensatz von Bildern dieser Objekte benötigt. Der Datensatz wird aufgeteilt in Trainings- und Testbilder. Die Trainingsdaten werden dem Netzwerk mitsamt den Lösungen übergeben. Während des Lernprozesses vergleicht das Programm seine Antworten mit den richtigen Lösungen. Die oben bereits erwähnten internen Parameter werden dabei verändert, damit das Netzwerk der Lösungen näher kommt. Der Lernvorgang wird mit allen Trainingsbildern durchgeführt und erfolgt automatisiert. Dem Netzwerk wird also nicht vorgegeben, auf welche Eigenschaften es achten soll. Auf dem Bild „Lernprozess“ kann der Trainingsvorgang nachvollzogen werden. Ein so trainiertes Netzwerk wurde für die eingefärbte Bilderreihe verwendet.

Lernprozess

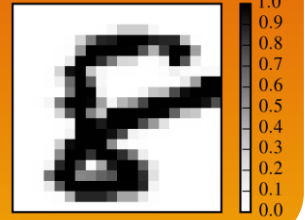


in der Abbildung rechts ersichtlich. Die erste Ziffer stellt dabei eine 3 dar und die zweite eine 7. Weiterhin konnte ein grosser Datensatz aus Baumblättern gesammelt werden. Das Baum-Netzwerk erreicht eine Erfolgsrate von 80%. Um diese komplexeren Bilder richtig zu klassifizieren, wäre weitergehende Programmierarbeit nötig. Insbesondere bei der Bildbearbeitung und Komprimierung gibt es noch Verbesserungspotenzial. Ein grösserer Datensatz könnte ebenfalls zu mehr Trainingserfolg führen.

Diskussion

Nach dem Training ist das Netzwerk in der Lage, 97% der unbekannten Testbilder der richtigen Ziffer zuzuordnen. Die Erfolgsrate ist also zufriedenstellend hoch. Das Ziel der erfolgreichen Programmierung eines funktionsfähigen Netzwerks wurde also erreicht. Unter den Bildern befinden sich jedoch auch sichtlich schwierige Exemplare. Davon sind zwei der falsch erkannten Ziffern

Input Digit



Preprocessed Digit

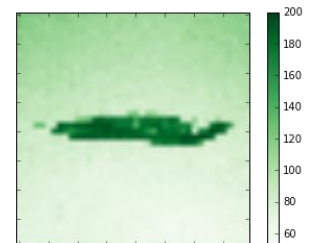


Baumblätter

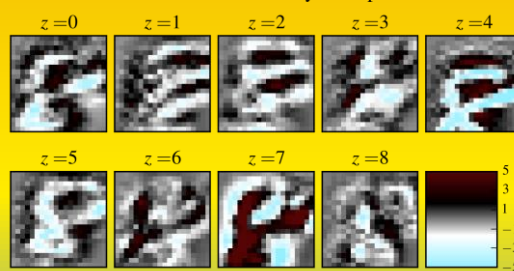
Im Lager wurde ein eigener Datensatz bestehend aus knapp 6'200 Bildern von Blättern von 7 verschiedenen Baumarten erstellt. Mit diesen Bildern wurde ein zweites

Netzwerk trainiert, um den Bildern die Baumart zuzuordnen. Die Erfolgsquote des Netzes liegt bei 80%. Dies bedeutet, bei 4 von 5 Bildern liegt das Programm richtig. In der Abbildung unten wird ein solches Baumblatt gezeigt. Es wurde bearbeitet und komprimiert, um es ans Netzwerk zu übergeben.

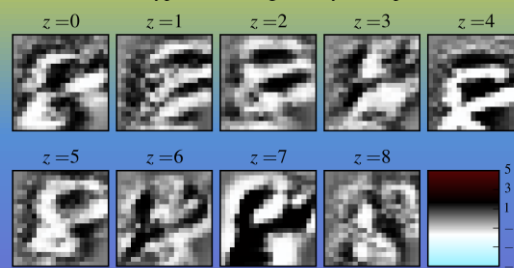
Bearbeitetes Bild eines Blattes



Convolutional Layer Output



Hyperbolic Tangent Layer Output



Fully Connected Layer Output

Digit 0	Digit 1	Digit 2	Digit 3	Digit 4
-0.79	-0.99	-1.02	-1.15	-1.26
Digit 5	Digit 6	Digit 7	Digit 8	Digit 9
-1.0	-0.42	-0.83	0.56	-1.14

Falsch erkannte Ziffern

