

Wie Computer träumen – Mustererkennung mittels künstlicher Intelligenz

Jonas Christian Stolzmann | Matrikel-Nr. 42733 | Wintersemester 2021/22 | Aktuelle Themen - KI

Abstract:

Um die Interaktion zwischen Menschen und Maschine zu ermöglichen, liegt eine Kernaufgabe darin, die menschliche Wahrnehmung nachvollziehen zu können. In diesem Paper wird gezeigt, inwieweit neuronale Netze ein Werkzeug sein können, um Eigenschaften der menschlichen Wahrnehmung zu übertreiben. Es wird die Mustererkennung mit Hilfe von neuronalen Netzen mit der Wahrnehmung des visuellen Kortex des menschlichen Gehirns verglichen. Es werden Phänomene der menschlichen Wahrnehmung betrachtet und in Zusammenhang zur künstlichen Intelligenz gestellt. Darüber hinaus wird erklärt, wie Mustererkennung bei neuronalen Netzen grundlegend funktioniert, anhand des neuronalen Netzwerkes „Deep Dream“ von Google.

Phänomene der menschlichen Wahrnehmung

Die Pareidolie bezeichnet ein Phänomen, bzw. eine Eigenschaft der menschlichen Wahrnehmung, in bestimmten Mustern vermeintliche Gesichter oder Figuren sehen zu können.¹ Eins in den späten siebziger Jahren sehr prominent gewordenes Beispiel ist das in Abbildung 1 gezeigte Foto der Marsoberfläche.



Abbildung 1: Marsgesicht

Das Foto wurde am 25. Juli 1976 von der NASA Raumsonde „Viking Orbiter 1“ fotografiert. Es zeigt, wie wir Gesichter auf Objekte projizieren können, wie in diesem

¹ (Payk, 2021)

Fall auf eine Gesteinsformation auf der Oberfläche des Mars.

Auch im alltäglichen Leben stoßen wir in vielen verschiedenen Situationen auf dieses Phänomen. Besonders typisch ist die Illusion, in Wolkenformationen Lebewesen und insbesondere Gesichter erkennen zu können. Was wir in Wolken sehen können ist dabei meist sehr individuell und hängt von den Verknüpfungen unserer Neuronen im visuellen Cortex ab.

Das visuelle System des Menschen ist ein Teil des Nervensystems, der sich mit der Verarbeitung von visuellen Informationen beschäftigt. Dazu zählen das Auge mit der Netzhaut (Retina), der Sehnerv, Teile des Thalamus (Ein Kernbereich des Zwischenhirns) und des Hirnstamms, sowie die Sehrinde.²

Während die Reize unserer Augen sich kaum voneinander unterscheiden, interpretiert das Gehirn einer jeden Person die eintreffenden Informationen auf Grund von unseren Erfahrungen und der Verknüpfungen im Thalamus unterschiedlich. Dadurch kommt eine unterschiedliche Interpretation solcher Phänomene zustande.

Neuronale Netze

Die Interpretation der Informationen findet in unserem Gehirn in etwa 100 Milliarden Nervenzellen statt und hängt davon ab, wie stark die einzelnen Nervenzellen miteinander verknüpft sind. In neuronalen Netzen versucht man dieses Prinzip der Gewichtung von einzelnen Nervenzellen zueinander nachzuahmen.³

Neuronale Netze bestehen aus mehreren Schichten, die jeweils aus vielen signalverarbeitenden Bausteinen bestehen. Diese Bausteine werden als Neuronen bezeichnet. Jedes neuronale Netz verfügt über eine Eingabeschicht, eine Ausgabeschicht und ein oder mehrere versteckte Schichten, dessen Neuronen jeweils wie in Abbildung 2 dargestellt miteinander verknüpft sind. Die Anzahl der Neuronen der Ein- und Ausgabeschicht sind variabel und hängen von der Dimension der Eingabe- und Ausgabevektoren ab.

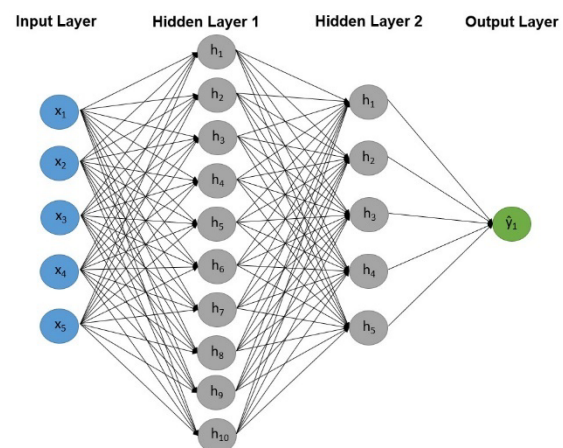


Abbildung 2: Beispielarchitektur eines neuronalen Netzes

Die Gewichtung zwischen den einzelnen Neuronen kann mit Hilfe von Trainingsbeispielen verändert werden. Mit Hilfe eines Lernalgorithmus wird das neuronale Netz auf Basis der Trainingsdaten angelehrt, was die Gewichte zwischen den Neuronen verändert. Dieser Vorgang wird so lange wiederholt, bis das neuronale Netz eine Erkennungsgenauigkeit erreicht, die zufriedenstellend ist.³

² (Leschnik, 2021)

³ (Dörn)

Mustererkennung durch neuronale Netze

Ein grundlegendes Beispiel für die Mustererkennung bei neuronalen Netzen ist die Erkennung von handgeschriebenen Zahlen.

Die Ausgabeschicht des neuronalen Netzes besteht aus zehn Neuronen. Ein Neuron für jede Ziffer im Zahlenbereich von null bis neun. Als Eingabe können beispielsweise Bilder mit einer vordefinierten Pixelgröße verwendet werden. Bei Bildern mit der Größe 64x64 Pixeln bekommt man so 4096 Eingabeneuronen.

Mit Hilfe von Trainingsbeispielen, bei denen neben den Eingabewerten auch die zugehörigen Ausgabewerte bekannt sind, wird das neuronale Netz trainiert. Mit jeder Iteration werden dadurch die Zahlen, in Form von Gewichten zwischen der Eingabeschicht, versteckten Schicht und Ausgabeschicht klassifiziert. Die Netzwerkparameter werden also schrittweise angepasst bis sie die gewünschte Klassifizierung liefern.⁴

Ein trainiertes Netz können wir nutzen, um unbekannte Eingabebilder nach diesen 10 Ausgabeneuronen zu klassifizieren. Dabei ist es jedoch unwichtig, ob wir Bilder von den Ziffern null bis neun als Eingabewerte nutzen, oder schauen wie viele Anteile der Ziffern sich beispielsweise in Wolken wiederfinden.^{5,6}

Deep Dream

Eine Herausforderung bei allen neuronalen Netzen ist es zu verstehen, was in den versteckten Schichten passiert. Typischerweise bestehen neuronale Netze aus 10-30 verschiedenen versteckten Schichten, die nacheinander iterativ durchlaufen werden. In der Ausgabeschicht werden sie schließlich nach ihren Merkmalen definiert.

Alexander Mordvintsev, Christopher Olah und Mike Tyka sind Mitarbeiter von Google und haben nach Möglichkeiten gesucht, um zu visualisieren, was in einem neuronalen Netzwerk vor sich geht.⁶

Dazu haben sie das Eingabebild, das aus Rauschen besteht iterativ so verändert, dass ein bestimmtes Ausgabeneuron bestmöglich angesprochen wird. Angenommen das angesprochene Ausgabeneuron klassifiziert ein Eingabebild als Banane. Das Eingabebild wird dann so verändert, dass sich die Klassifizierung des Ausgabeneurons maximiert. Idealerweise ergibt sich ein Bild, das zu 100% aus der Interpretation von dem Ausgabeneuron von einer Banane besteht. (siehe Abbildung 3)

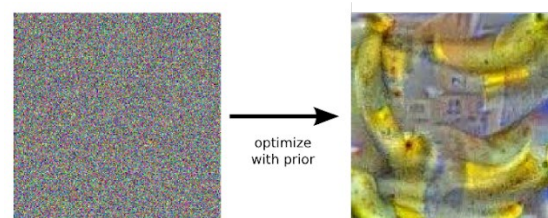


Abbildung 3: Veränderung des Eingabebildes nach dem Ausgabeneuron

⁴ (Dörn)

⁵ (Alexander Mordvintsev)

⁶ (Ertel, 2016)

Es entsteht eine angenäherte abstrakte Darstellung, wie das Ausgabeneuron eine Banane klassifizieren würde.

Eine weitere Möglichkeit, um herauszufinden, was in den versteckten Schichten passiert ist es, dem neuronalen Netzwerk die Entscheidung zu überlassen, wie es ein Bild interpretiert. Dazu geben wir dem neuronalen Netzwerk an der Eingabeschicht ein beliebiges Foto vor und lassen das Netzwerk all die Stellen verstärken, die es auf einer bestimmten Ebene in den versteckten Neuronen interpretiert. Dadurch zeigt sich, dass

niedrige Schichten eher Striche oder einfache Muster erzeugen, während höhere Schichten komplexere Gebilde bis hin zu grundlegende Merkmale verstärken.

Abbildung 4 zeigt, wie das neuronale Netzwerk „Deep Dream“ von Google ein einfaches Bild von Wolken überinterpretiert. Es zeigt abstrakte Darstellungen von vielen verschiedenen Tieren, die das neuronale Netz in den Wolken interpretiert. Da Deep Dream hauptsächlich mit Tieren trainiert wurde, neigt das Netzwerk natürlich dazu die Formen als Tiere zu interpretieren.

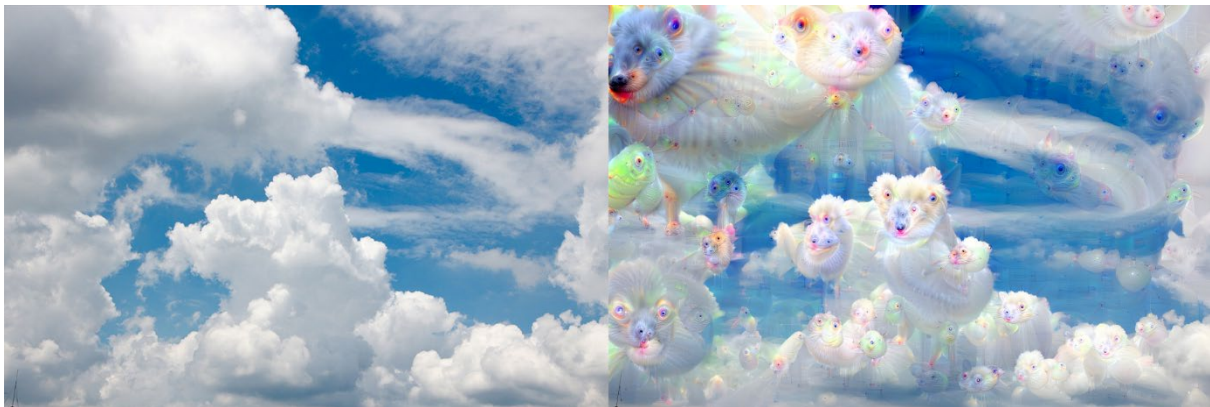


Abbildung 4: Interpretation von Wolkenbildern eines neuronalen Netzes

Fazit:

Wenn Menschen in eine Wolkenlandschaft blicken, dann beginnen sie mit der Zeit nach erkennbaren Mustern zu suchen. Unser Gehirn erkennt die Wolken in ihren grundlegenden Strukturen und „überinterpretiert“ die Umrisse durch von uns bekannte Dinge, wie bei dem Gesicht in den Felsstrukturen auf dem Mars.

Ein ähnliches Verhalten sehen wir bei „Deep Dream“. Es nutzt die Informationen aus den Trainingsbeispielen, um die grundlegenden Strukturen in Bildern zu analysieren und „überinterpretiert“ diese. Dadurch hilft es uns zu verstehen, wie neuronale Netze Muster in beliebigen Bildern erkennen und klassifizieren. Es zeigt uns auch wie die einzelnen Ebenen im neuronalen Netz arbeiten und kann dazu beitragen die Entwicklung von neuartigen neuronalen Netzen zu unterstützen. Dadurch bringt uns „Deep Dream“ der Frage, ob Androiden von elektrischen Schafen träumen ein kleines Stück näher.⁷

⁷ (Diverse)

Abbildungsverzeichnis:

Abbildung 1: Marsgesicht.....	1
Abbildung 2: Beispielarchitektur eines neuronalen Netzes	2
Abbildung 3: Veränderung des Eingabebildes nach dem Ausgabeneuron	3
Abbildung 4: Interpretation von Wolkenbildern eines neuronalen Netzes.....	4

Literaturverzeichnis

Alexander Mordvintsev, Christopher Olah, Mike Tyka. Googleblog. *Googleblog*. [Online] [Zitat vom: 27. März 2022.] <https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html>.

Diverse. Wikipedia. *Wikipedia*. [Online] [Zitat vom: 27. März 2022.] https://de.wikipedia.org/wiki/Tr%C3%A4umen_Androiden_von_elektrischen_Schafen%3F.

Dörn, Sebastian. <https://sebastiandoern.de/neuronale-netze/>. <https://sebastiandoern.de/neuronale-netze/>. [Online] [Zitat vom: 27. März 2022.] <https://sebastiandoern.de/neuronale-netze/>.

Ertel, Wolfgang. 2016. *Grundkurs Künstliche Intelligenz - Eine praxisorientierte Einführung*. Ravensburg-Weingarten : Springer Fachmedien Wiesbaden GmbH, 2016.

Leschnik, Andreas. 2021. *Wahrnehmung - Grundlagen, Clinical Reasoning und Intervention im Kinder und Jugendalter*. Großrosseln : Springer Fachmedien Wiesbaden GmbH, 2021.

Payk, Theo R. 2021. *Psychopathologie*. Bonn : Springer-Verlag GmbH, 2021.