

Woche 13 Praktische Anwendungsbeispiele: Bildverarbeitung (Teil 1)

Skript

Erarbeitet von
Dr. Jacqueline Klusik-Eckert

Lernziele	1
Inhalt	2
<i>Einstieg</i>	2
<i>Einsatzgebiete und Anwendungen</i>	2
<i>Wie funktioniert das?</i>	3
<i>menschliches Sehen vs. Computersehen</i>	4
Quellen	5
Disclaimer	6

Lernziele

- Übersicht über verschiedene Anwendungen von KI in der Bildverarbeitung erhalten
- Verständnis für die Vorteile und Limitationen von KI-basierten Lösungen im Vergleich mit traditionellen Methoden bekommen
- Verständnis für ethische und datenschutzrechtliche Aspekte des Einsatzes von KI in der Bildverarbeitung gewinnen
- Sensibilisiert für Biases durch Trainingsdaten werden

Inhalt

Einstieg

Manchmal hat man so schlecht geschlafen, dass einen die Gesichtserkennung am Smartphone noch nicht mal erkennt. Kennst du auch?

Da hat die KI mal wieder einen Fehler gemacht.

Oder besser gesagt: als du die Gesichtserkennung mit deinem Gesicht vertraut gemacht hast, warst du wahrscheinlich nicht krank. Womöglich noch gut frisiert. Mit guten Lichtverhältnissen. Woher soll die KI denn wissen, dass dieses verwuschelte, unausgeschlafene etwas, das an ein Gesicht erinnert, auch du bist?

Einsatzgebiete und Anwendungen

Die Technik, die hinter dieser sehr alltäglichen Anwendung steht, kommt aus dem Bereich der automatischen Bildverarbeitung. Damit wird die Forschung über Neuronale Netze zusammengefasst, die für spezielle Einsatzbereiche des maschinellen Sehens entwickelt wurden.

Schauen wir uns doch mal ein paar Einsatzbereiche dieser KIs genauer an. Wir finden schon heute CNNs, also diese speziell trainierten KIs für das Computersehen in allen Bereichen, in denen es um das automatische Klassifizieren von Bildern geht. Gesichtserkennung für das Öffnen des Smartphones bis hin zu Überwachungsanlagen mit automatischer Erkennung von Gefahrgütern wie Schusswaffen oder Messer. Das ist hilfreich für die Sicherheitskontrolle am Flughafen. Maschinelles Sehen ist auch für autonomes Fahren wichtig, denn schließlich soll die Karre ja auf der Straße bleiben.

Verschiedene Anwendungen:

- Automatisches Klassifizieren von Bildern (Identifizieren von Objekten, Gesichtern)
- Automatisches Erkennen und Verfolgen von Bewegungen in Videodaten
- Automatisches Analysieren von medizinischen Bildern, Medizinische Bildanalyse, Medizinische Diagnostik
- Bild- und Videobearbeitung
- Überwachungskameras
- Autonomes Fahren
- Verkehrsmanagement
- ...

Der Künstliche Intelligenz kommt bei der automatischen Bildverarbeitung auch dann zum Einsatz, wenn große Datenmengen schnell berechnet werden müssen. Sie ermöglichen erst die komplexen Open-World-Computerspiele mit super realistisch ausschauenden Landschaften oder atemberaubenden Special Effects in Blockbuster-Filmen.

Die verwendeten Verfahren sind auf die jeweiligen Anwendungsfelder angepasst und alle kann ich nicht erklären. Also konzentrieren wir uns mal in diesem Beispiel auf die automatische Bildverarbeitung, die für das Klassifizieren von großen Bildmengen verwendet wird.

Quelle [1, 2]

Wie funktioniert das?

Konvolutionale Neuronale Netze (CNNs) verwenden eine spezielle Art von maschinellem Lernen, um Bilder automatisch zu kategorisieren oder zu klassifizieren. Hierbei werden bestimmte Merkmale in einem Bild erkannt und dann zur Bestimmung der Kategorie verwendet.

Im ersten Schritt wird das Bild in kleine Teile unterteilt, die als Filter bezeichnet werden. Diese Filter werden über das gesamte Bild bewegt, um Merkmale wie Kanten, Farben und Formen zu erkennen.

The first layer bases (top) and the second layer bases (bottom) learned from natural images.

Quelle [3]

Dann werden diese Merkmale kombiniert und zu höheren abstrakten Konzepten weiterverarbeitet, die für die Klassifizierung von Bedeutung sind. Hierbei werden die Ausgaben der Filter in verschiedenen Schichten verknüpft, um eine Entscheidung über die Kategorie des Bildes zu treffen.

Das CNN lernt dabei aus einer großen Anzahl vorher trainierter Bilder, um Muster und Beziehungen zwischen Merkmalen und Kategorien zu erkennen. Dieses Wissen wird dann verwendet, um neue, ungesehene Bilder automatisch zu klassifizieren.

Die ganze Forschung in diesem Bereich, also das maschinelle Sehen, wird als Computer Vision bezeichnet. Die Einsatzgebiete dieser KI-Technik sind dann wieder unglaublich vielseitig. Immer dann, wenn es um Bilder geht, und das schließt auch Bewegtbilder mit ein, haben wir es mit Computer Vision zu tun. Dabei geht es nicht nur darum, den neuronalen Netzen beizubringen, Kanten und Kontraste zu erkennen. Man zielt in diesem Teilgebiet der Informatik darauf ab, die menschliche Wahrnehmung zu simulieren und sie zu komplementieren. Die Technik kommt also auch dann zum Einsatz, wenn es um unsere Wahrnehmungsgrenzen geht: wenn etwas unglaublich schnell passieren muss oder wenn die Bildmenge einfach viel zu groß ist.

Quelle [4, 5]

menschliches Sehen vs. Computersehen

Als gut trainierte Kunsthistorikerin kann ich mit einer Bilderflut schon ganz gut umgehen.

Einblendung von Bildern und Sortierungen

Ich habe gelernt, schnell zu sehen und nach Kategorien wie Datierung, Epoche und Stil Kunstwerke einzuordnen.

Epochen <http://katrin-hemmen.blogspot.com/2012/05/time-line.html>
OPEN HISTORY ARCHIVE. (o. J.). Abgerufen 2. März 2023, von
<https://www.openhistoryarchive.com/t-sne-map/>

Bei einer überschaubaren Menge von Bildern, und genügend Zeit, bekomme ich das hin. Nach Farben sortieren, nach Größe, nach Ähnlichkeiten. Alle heraussuchen, auf denen eine Frau abgebildet ist ...

Wenn ich aber jetzt aus 100.000 Bildern etwas sortieren müsste, bräuchte ich ewig. An Millionen Bilder erst gar nicht zu denken.

Dabei hat unser Sehen und das schnelle Lernen von Kategorien den KI's einiges voraus. Während eine Neuronales Netz über eine Million Bilder von Katzen braucht, um auf einem neuen, noch unbekannten Bild eine Katze zu erkennen, braucht ein Kleinkind oft nur wenige Beispiele.

<https://datainnovation.org/2014/08/10000-cat-pictures-for-science/>

Das Kind verknüpft mit dem Erscheinungsbild jedoch nicht nur die Form. Sehr schnell wissen Kinder, dass es sich dabei um ein Tier handelt, dass es Miau macht, dass sie flauschig sind und meistens schnell wegrennen, wenn man sie streicheln möchte. Wir Menschen sehen nicht nur einfache Formen. Wir verbinden alles Gesehene sofort mit Bedeutung, Sprache, Haptischem.

Bild menschliches vs. computersehen

Das Neuronale Netz verbindet mit einer trainierten Form eine Zeichenkette. Was diese Zeichenkette bedeutet, wird nicht verstanden.

Sie sind uns jedoch dann überlegen, wenn enorme Bildmengen in einer menschenmöglichen Gewichtigkeit verrechnet werden müssen. Sie sind besonders gut, wenn es zum Beispiel um

einen Vergleich von ganzen Bildern geht, in denen man Duplikate aus großen Datenmengen herausfiltern möchte. Auch das Auffinden von einzelnen Objekten in Bildern funktioniert super. Zeige mir alle Bilder mit einer Banane.

Was mit Bananen und Katzen anfängt, kann auch zu nützlichen Szenarien hin trainiert werden. Da gibt es beispielsweise das Projekt ChextNext. Hier wurde die KI trainiert, um Krankheiten wie Tuberkulose oder ähnliches in Röntgenaufnahmen des Thorax zu erkennen. Der Algorithmus kann dabei helfen die Mediziner*innen darauf aufmerksam zu machen, wenn irgendetwas abweicht und wenn es Ungereimtheiten gibt.

Quelle [6]

Durch die Errechnung von Unterschieden können auch kleinste Abweichungen festgestellt werden, oder hättet ihr mit bloßem Auge sehen können, wo der Fehler ist? Generell helfen diese KIs bei der Erschließung von Bilddaten, um sie zu strukturieren, zu clustern, um sie überhaupt sortierbar zu machen.

Bild von [shlomz](#) / [Pixabay](#)

Man kann sich das in etwas so vorstellen, wie ein Spürhund bei der Polizei arbeitet. Dem Hund ist es möglich, mit seiner speziellen Ausbildung für Substanzen, diese zu finden. Der Hund selbst hat aber nicht die gleiche Auffassung und Vorstellung über die Bedeutung der Substanzen. Wenn der Hund eine Ladung Drogen gefunden hat, ist ihm nicht bewusst, dass diese gegen das Betäubungsmittelgesetz verstoßen. Er weiß nur, dass er das finden soll, weil er darauf trainiert wurde. Es braucht die Kriminalkommissarin, um den Fund zu kontextualisieren.

Hund Flaticon

Das „Assistenzsystem“ Hund ähnelt dem Assistenzsystem Computer Vision, das ohne Verständnis für die Gesamtsituation und den Kontext für eine relativ spezifische Aufgabe eingesetzt wird. Die Funde, bzw. Befunde müssen dann vom Menschen interpretiert und eingeordnet werden.

Quelle [7]

Quellen

Quelle [1] Nischwitz, A., Fischer, M., Haberäcker, P., & Socher, G. (2019). Zusammenhang zwischen Computergrafik und Bildverarbeitung. In A. Nischwitz et al (Hrsg.), Computergrafik: Band I des Standardwerks Computergrafik und Bildverarbeitung (S. 6–24). Springer Fachmedien. https://doi.org/10.1007/978-3-658-25384-4_2

- Quelle [2]** Quelle Radke, R. J. (2013). *Computer vision for visual effects*. Cambridge University Press.
- Quelle [3]** Lee, H., Grosse, R., Ranganath, R., & Ng, A. Y. (2009). Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations. *Proceedings of the 26th Annual International Conference on Machine Learning*, 609–616. <https://doi.org/10.1145/1553374.1553453>
- Quelle [4]** Burrell, J. (2015). How the Machine „Thinks“: Understanding Opacity in Machine Learning Algorithms. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.2660674>
- Quelle [5]** Bell, P., & Offert, F. (o. J.). *Towards Open Computer Vision Methods: Visual Analytics of Machine Learning Models in the Digital Humanities* [Text]. Abgerufen 2. März 2023, von https://dh2020.adho.org/wp-content/uploads/2020/07/483_TowardsOpenComputerVisionMethodsVisualAnalyticsofMachineLearningModelsInTheDigitalHumanities.html
- Quelle [6]** Rajpurkar, Pranav et al., Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists, in: Aziz Sheikh (Hg.), *PLOS Medicine* 15/11, 2018, e1002686.
- Quelle [7]** Bell, P. (2022). Computer Vision und Visualisierung als didaktische Instrumente in der Kunstgeschichte. In H. Locher & M. Männig (Hrsg.), *Lehrmedien der Kunstgeschichte: Geschichte und Perspektiven kunsthistorischer Medienpraxis* (S. 384–401). De Gruyter. <https://doi.org/10.1515/9783422986251>

Disclaimer

Transkript zu dem Video „Woche 13 Praktische Anwendungsbeispiele: Bildverarbeitung“, Dr. Jacqueline Klusik-Eckert.

Dieses Transkript wurde im Rahmen des Projekts ai4all des Heine Center for Artificial Intelligence and Data Science (HeiCAD) an der Heinrich-Heine-Universität Düsseldorf unter der Creative Commons Lizenz [CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/) veröffentlicht. Ausgenommen von der Lizenz sind die verwendeten Logos, alle in den Quellen ausgewiesenen Fremdmaterialien sowie alle als Quellen gekennzeichneten Elemente.