



Wie KI bei der Erkennung von Hieroglyphen hilft

Marco Faisst, Matrikel-Nr. 42928

Einführung

Künstliche Intelligenz (KI) oder englisch „Artificial Intelligence“ (AI) wird nicht nur bei speziellen Anwendungsfällen eingesetzt, sondern begegnet uns zunehmend auch im Alltag und beruflichen Umfeld.

Die Entwicklung im Bereich KI, die ein Gebiet der Informatik ist, schreitet mit immenser Geschwindigkeit voran. Es vergeht fast kein Tag, an dem man nicht in einschlägigen IT-Nachrichtenportalen einen Artikel zu einem Thema mit KI-Bezug lesen kann. Neben autonomen Robotern und Fahrzeugen sticht auch das Gebiet der Bilderkennung (Musteranalyse) und Bilderzeugung hervor. Zu nennen ist hier bspw. „DALL-E“, ein Computerprogramm, das aus Texteingaben selbstständig computergenerierte Bilder in teilweise beeindruckender Qualität erzeugt.

Insbesondere im Zeitalter der digitalen Transformation und durch immer

leistungsfähigere Prozessoren spielt KI auch als Hilfsmittel in der Arbeitswelt eine wesentliche Rolle.

So auch in der Archäologie bei der Erkennung und Interpretation von antiken Schriften. Das vorliegende White Paper beschreibt die aktuellen Entwicklungen im Bereich der Hieroglyphen-Erkennung und gibt einen Überblick über aktuelle Software und Forschungsprojekte.

Von der Sprache zur Schrift

Die Erfindung der Schrift bedeutete in der Geschichte der Menschheit ein Umbruch. Bis zu diesem Zeitpunkt konnte Wissen nur mündlich weitergeben werden. Prinzipbedingt gingen dabei im Laufe der Zeit Informationen verloren, weil diese beim Weitererzählen vergessen, für nicht wichtig gehalten, absichtlich verändert oder vom Empfänger nicht richtig verstanden worden sind.

In der Informationstheorie spricht man dabei von einer Verminderung des Werts von Informationen während eines Informationsflusses einer Kommunikation.

Anfangs bestand die Schrift noch aus reinen bildlichen Darstellungen (Piktogrammen). Daraus entwickelten sich die

Keilschrift, verschiedene Hieroglyphen und weitere Schriften.

Zu den bekanntesten Hieroglyphen zählen die ägyptischen Hieroglyphen. Diese von ca. 3200 v. Chr. bis 394 n. Chr. verwendete Schrift ist sehr komplex, da sie sowohl Piktogramme enthält, als auch Zeichen, die Laute kennzeichnen (Phonogramme) sowie Zeichen, die eine Unterscheidung zwischen Piktogrammen und Phonogrammen bezwecken (Determinative).

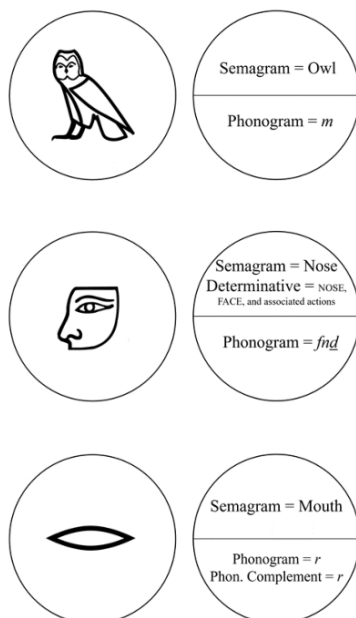


Abb. 1: Elemente ägyptischer Hieroglyphen

Anforderungen an die KI

Eine KI zur Erkennung von Hieroglyphen muss also auch die Bedeutung der jeweiligen Zeichen im Kontext „verstehen“ und berücksichtigen, dass es verschiedene Leserichtungen gibt: von links nach rechts, von oben nach unten, umgekehrt oder verschiedene Richtungen kombiniert.

Eine weitere Herausforderung stellt das Fehlen von Zeichen dar, wenn bspw. eine Schrifttafel beschädigt ist.



Abb. 2: Rekonstruktion einer beschädigten Inschrift

Die KI muss jedoch keine komplett unbekannte Schrift verarbeiten und entschlüsseln, sondern kann in der Regel auf bereits vorhandene Übersetzungen als Grundlage zurückgreifen.

In Bezug auf die ägyptischen Hieroglyphen gelang es dem Franzosen Jean-François Champollion vor 200 Jahren, im Jahr 1822, diese Hieroglyphen im Wesentlichen zu entziffern.

Am Anfang steht die Zeichenerkennung

Die Zeichenerkennung per OCR-Software (Optical Character Recognition) war anfangs darauf beschränkt, Zeichen für Zeichen, d.h. Pixel für Pixel, mit vorgegebenen Musterzeichen zu vergleichen und dann bei möglichst hoher Übereinstimmung den korrekten Buchstaben auszugeben. Dies setzt allerdings voraus, dass der von der Software zu verwendende Zeichensatz vorgegeben wird und die Software diesen aus mehreren Auswahlmöglichkeiten zumindest selbst erkennt. Im weiteren Verlauf der Entwicklung führte die Erkennung nicht nur einzelner Buchstaben, sondern kompletter Wörter zu weniger Fehlern. Hierzu werden Wörterbücher als Grundlage eingesetzt.

Auch abstraktere Formen wie bei individuellen Handschriften können heutzutage in

der Regel sehr gut erkannt werden. Bei der Handschrifterkennung leisten die im nächsten Abschnitt beschriebenen neuronalen Netze einen wichtigen Beitrag zur Verbesserung der Erkennungsrate.

Die KI kommt ins Spiel

Eine reine Zeichenerkennung und Digitalisierung der Hieroglyphen sind natürlich nicht ausreichend. Die verwendete Schrift eines Textes muss auch interpretiert werden und oftmals sind nur Bruchstücke des Ausgangsmaterials vorhanden.

Die fehlenden Informationen lassen sich von Menschen nur mit erheblichem Aufwand rekonstruieren. Ein weiterer Faktor ist, dass nicht alle Detailinformationen eines bereits bekannten Textes übersetzt worden sind, weil sie bisher als unwichtig betrachtet worden sind.

Verschiedene Software-Tools mit KI können hier Forschern viel manuelle Arbeit abnehmen.

Künstliche neuronale Netze stellen oftmals die Grundlage solcher Tools dar. Ein künstliches neuronales Netz orientiert sich an den Vorgängen im menschlichen Gehirn und hat als Anwendungsgebiet zum einen, die Funktionsweise des menschlichen Gehirns besser zu verstehen, zum anderen wird es dazu verwendet, konkrete Anwendungsprobleme zu lösen – wie die Erkennung von Hieroglyphen.

Ein künstliches neuronales Netz besteht aus mehreren Neuronen (auch als Units, Einheiten oder Knoten bezeichnet). Diese

dienen dazu, Informationen aus der Außenwelt oder von anderen Neuronen aufzunehmen und geben die Informationen an andere Units oder die Außenwelt in modifizierter Form weiter.

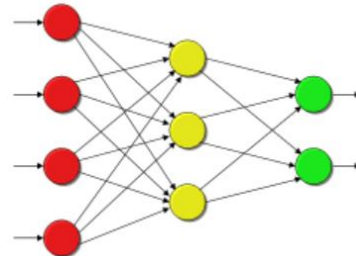


Abb. 3: Neuronales Netz, schematische Darstellung

Die Stärke der Verbindung zwischen zwei Neuronen wird durch ein Gewicht ausgedrückt. Je größer dies ist, desto größer ist der Einfluss einer Unit auf eine andere Unit.

Vereinfacht gesagt, ist das Wissen eines neuronalen Netzes in seinen Gewichten gespeichert und Lernen wird als Gewichtsveränderungen zwischen den Units definiert. Wie die Gewichtsveränderung genau erfolgt ist, wiederum von den jeweils verwendeten Regeln abhängig.

Ein spezielles neuronales Netz ist das „Convolutional Neural Network“ (ConvNet). Es eignet sich besonders gut für die Verarbeitung von Bild- oder Audiodaten. Einem Eingangsbild werden dabei verschiedenen Aspekten bzw. Objekten im Bild eine Bedeutung zugewiesen. Das Convolutional Neural Network ist in der Lage, diese voneinander zu unterscheiden. Gegenüber einem herkömmlichen neuronalen Netzwerk hat es Vorteile bei der Vorverarbeitung von Bildern, da es die Fähigkeit zum Lernen von Filtern hat, während diese bei

einfachen Methoden manuell erstellt werden müssen.

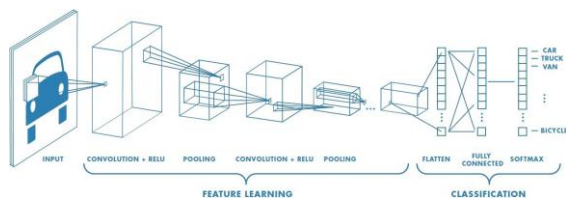


Abb. 4: Bildverarbeitung mit einem Convolutional Neural Network

Ithaca

Von dem Unternehmen DeepMind stammt das in der Programmiersprache Python geschriebene und als Open Source bei github veröffentlichte Tool „Ithaca“.

DeepMind wurde im September 2010 gegründet und ist seit 2014 ein auf KI spezialisiertes Tochterunternehmen von Google, jetzt Alphabet.

Im März 2022 erschien zu Ithaca ein Fachartikel in dem wissenschaftlichen Journal Nature. Der grundsätzliche Ablauf des Tools sieht wie folgt aus:

Ein lückenhafter Ausgangstext wird in das Tool eingegeben und Fehlstellen markiert. Als Ergebnis werden die wahrscheinlichsten Hypothesen zur Rekonstruktion des Textes geliefert. Des Weiteren präsentiert Ithaca eine Einschätzung zum Herkunftsort (84 mögliche Regionen) und zu der Ursprungszeit der Schrift. Um dies zu erreichen, greift das Tool, neben Verwendung verschiedener Algorithmen und mathematischer Funktionen (bspw. Sinusfunktion), auf ein tiefes neuronales Netz zurück, das in Abhängigkeit der Position der Wörter in dem Text ein Ergebnis generiert.

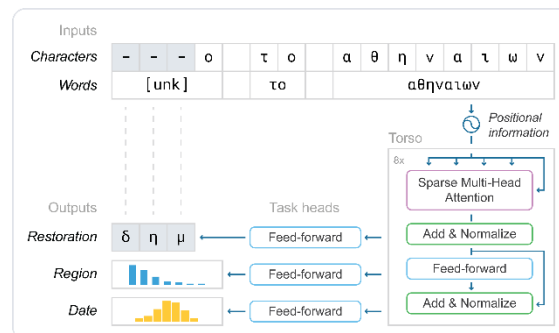


Abb. 5: Ablaufdiagramm Ithaca

Um das Tool zu trainieren, wurde ein Datensatz mit rund 180000 antiken Schriften genutzt. Diese sind bereits wissenschaftlich bearbeitet und transkribiert worden.

Der Trainingsprozess dauerte eine Woche und erfolgte auf 128 „Google TPU v4 Pods“ in der Google Cloud Platform.

Ithaca kommt auf eine Trefferquote von 62 Prozent, wohingegen Wissenschaftler nur bei rund 25 Prozent der Zeichen richtig lagen. Die besten Ergebnisse liefert eine Kombination aus beiden.

KI setzt Fragmente zusammen

Bei Ausgrabungen aufgefundene Tontafeln sind nicht selten in einem zerbrochenen Zustand und einzelne Teile können mit anderen, nicht zusammengehörigen Teilen, vermischt sein. Diese Einzelteile wie ein Puzzle zusammenzufügen, macht einen Großteil der Arbeit aus. Wissenschaftler von der Ludwig-Maximilians-Universität München möchten diesen Arbeitsaufwand verringern und entwickeln hierzu den Algorithmus „CuneiBLAST“.

Dieser basiert auf dem BLAST-Algorithmus (Basic Local Alignment Search Tool),

welcher im Bereich der Biologie genutzt wird, um bspw. DNA-Daten zu vergleichen.

Zunächst sollen die Fragmente fotografiert und nach Transkription, d. h. Überführung in das lateinische Alphabet, schließlich in einer Datenbank gespeichert werden. Der Algorithmus soll in der Lage sein, einzelne Textstücke mit der Datenbank abzugleichen und zusammengehörende Teile zu identifizieren.

Die Schwierigkeit liegt darin, dass ein einzelnes Zeichen je nach Kontext unterschiedliche Bedeutungen haben kann. Dies muss der Algorithmus berücksichtigen und erkennen.

Deep Learning zur Bildklassifizierung

Unterschiedliche Arten von Objekten in Bildern automatisch zu identifizieren, bezeichnet man als Bild- bzw. Objektklassifikation. Ob es sich bei den Objekten dabei um verschiedene Hieroglyphen oder bspw. Straßenschilder handelt, macht keinen Unterschied. In jedem Fall müssen der KI die noch „fremden“ Objekte von Menschen beigebracht werden, damit die KI später selbst sinnvolle Entscheidungen treffen kann.

Im Internet gibt es entsprechende Bilddatenbanken, die von vielen Menschen bearbeitet worden sind. Zahlreiche verschiedene bildliche Darstellungen bspw. eines Apfels müssen mit der Objektklasse „Apfel“ und evt. noch einer Kategorie „Obst“ versehen werden. Eine bekannte Datenbank ist „ImageNet“ mit über 20000 Objektklassen.

Diese Trainingsdaten dienen als Ausgangsbasis für maschinelles Lernen, ein Algorithmus, der sich selbst an veränderte Gegebenheiten anpasst. Ein Teilgebiet des maschinellen Lernens ist „Deep Learning“, das auf künstliche neuronale Netze zurückgreift und bei großen Datenmengen zum Einsatz kommt.

Das gleiche Prinzip wird bei der Erkennung von Hieroglyphen angewendet:

Forscher des Instituts für Angewandte Physik „Nello Carrara“ in Italien haben in der Fachzeitschrift IEEE Access im September 2021 eine Studie veröffentlicht, die mit Hilfe von Deep Learning und mit sehr hoher Genauigkeit die Möglichkeit der automatischen Klassifizierung von Bildern altägyptischer Hieroglyphen analysiert.

Ziel der Studie ist es auch aufzuzeigen, wie diese KI-basierenden Analyse-Tools die Forschung im ägyptologischen Bereich unterstützen können, indem sie in die Tätigkeiten der Archäologinnen und Archäologen einbezogen werden.

Eine Herausforderung laut Studie ist – wie auch bereits die Entwickler des CuneiBLAST-Algorithmus erwähnten –, dass Hieroglyphen aus einem breiten Spektrum von Ideogrammen bestehen, d.h. die Schriftzeichen vertreten nicht eine bestimmte Lautung, sondern einen kompletten Begriff. Sie sind allgemein etwa 26 Kategorien zuordenbar, die zu Wörtern und Lauten kombiniert werden.

Grundlage ist die über 700 Zeichen umfassende sogenannte Gardiner-Liste (engl.: Gardiner's sign list), die der britische Ägyptologe Sir Alan Gardiner 1927 angefertigt hat.

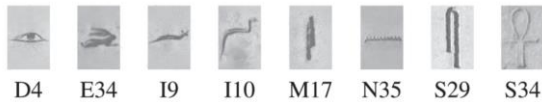


Abb. 6: Zeichenzuordnung gemäß Gardiner-Liste

Zudem wurden Hieroglyphen auf unterschiedliche Weise geschrieben, bspw. kursiv, in verschiedenen Richtungen und auf verschiedenen Trägermaterialien.

Für die Analysen standen zwei Datensätze mit Bildern zur Verfügung, welche zu einem Datensatz zusammengeführt worden sind.

Der erste Datensatz besteht aus 4310 Graustufen-Bildern mit Hieroglyphen von Fotos der Wände innerhalb der Unas-Pyramide. Jedes Bild hat eine Auflösung von 75 x 50 Pixeln und steht für eine einzelne Hieroglyphe. Der Datensatz enthält 172 Hieroglyphen.

Der zweite Datensatz enthält 1310 Farbbilder im RGB-Format mit variablen Auflösungen. Jedes Bild repräsentiert ein einzelnes Element, das zu 48 möglichen verschiedenen Hieroglyphen gehört.

Für die Klassifizierungsaufgabe befassten sich die Autoren u.a. mit „Inception v3“ – ein Convolutional Neural Network von Google zur Unterstützung bei der Bildanalyse und Objekterkennung – und der verbesserten Version „Xception“.

Die Autoren stellen in der Studie ihre Version eines für die Erkennung von Hieroglyphen optimierten neuronalen Netzwerks vor und nennen dieses „Glyphnet“.

Im Ergebnis übertrifft Glyphnet die anderen neuronalen Netzwerke sowohl in Bezug auf die Leistung als auch auf die Einfachheit des Trainings und die Rechenleistung.

Das Fabricius-Projekt

Fabricius ist ein im Jahr 2017 gestartetes gemeinsames Projekt von Google Arts & Culture, dem Australian Centre of Egyptology der Macquarie University und Ubisoft.

Es ist das erste digitale Tool, mit dem Hieroglyphen durch Unterstützung leistungsfähiger KI schnell entschlüsselt werden können. Der Quellcode befindet sich auf github unter <https://github.com/googleartsandculture/workbench>. Das webbasierte Tool steht in englischer und arabischer Sprache unter <https://fabriciusworkbench.withgoogle.com> zur Verfügung.

Über die Web-Oberfläche des Tools können direkt Hieroglyphen als Bilddatei hochgeladen werden. Die Bilder werden nach einer manuellen Markierung der relevanten Bereiche und weiteren Anpassungen automatisch analysiert und anschließend wahrscheinliche Übereinstimmungen mit bekannten Schriftzeichen (als Wort oder Satz) präsentiert. Falls gewünscht, kann durch Speichern eigener manueller Interpretationen von Hieroglyphen ebenfalls dazu beigetragen werden, dass die Erkennungsrate für die Allgemeinheit verbessert wird.

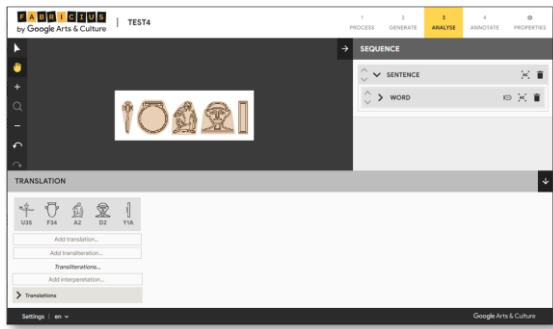


Abb. 7: Fabricius Workbench

Eine spielerische Variante, die auch Hintergrundwissen zum Thema ägyptische Hieroglyphen vermittelt, kann unter <https://arts-experiments.withgoogle.com/fabricius/en> ausprobiert werden.

Fazit

Zusammenfassend lässt sich festhalten, dass KI bei der Erkennung von Hieroglyphen einen beträchtlichen Beitrag leisten kann.

Durch Anpassung vorhandener neuronaler Netzwerke speziell an die Besonderheiten der Bilderkennung und -verarbeitung von Hieroglyphen gibt es bereits aussichtsreiche Möglichkeiten.

Vorteile bieten sich dabei sowohl bei der Erkennungsquote als auch durch Einsparung von Zeit.

Eine KI wird zwar die Tätigkeiten von Archäologinnen und Archäologen im Bereich Hieroglyphen nicht vollständig übernehmen können, aber eine Menge manueller Arbeitsaufwand kann durch KI-unterstützte Anwendungen vermieden werden.

Zu beachten ist dabei stets der Grundsatz: Je mehr eine KI mit menschlichen

Entscheidungen trainiert wird, d.h. je mehr Trainingsdaten zur Verfügung stehen, desto zuverlässiger und genauer arbeitet die KI bei der Erkennung von Hieroglyphen.

Quellen

Assael, Y.; Sommerschild, T.; Shillingford, B.; Bordbar, M.; Pavlopoulos, J.; Chatzipanagiotou, M.; Androutsopoulos, I.; Prag, J.; de Freitas, N.: Restoring and attributing ancient texts using deep neural networks. *Nature*, 09.03.2022, <https://www.nature.com/articles/s41586-022-04448-z> (Zugriffsdatum: 19.05.2022)

Barucci, A.; Cucci, C.; Franci, M.; Loschiavo, M.; Argento, F.: A Deep Learning Approach to Ancient Egyptian Hieroglyphs Classification. *IEEE Access*, 13.09.2021, <https://ieeexplore.ieee.org/document/9528382> (Zugriffsdatum: 19.05.2022)

Holland, M.: Puzzeln wie Profis: Mit KI antike Keilschrift-Texte entschlüsseln. *heise online*, 22.08.2022, <https://www.heise.de/newsticker/meldung/Puzzeln-wie-Profis-Mit-KI-antike-Keilschrift-Texte-entschluesseln-4502330.html> (Zugriffsdatum: 19.05.2022)

Kemling, L.: Hieroglyphen. *Planet Wissen*, 09.12.2020, <https://www.planet-wissen.de/geschichte/antike/hieroglyphen/index.html> (Zugriffsdatum: 21.05.2022)

Paaß, G.; Hecker, D.: Künstliche Intelligenz. Was steckt hinter der Technologie der Zukunft? Springer-Vieweg, Wiesbaden 2020

Rey, G. D. Rey; Wender, K. F.: Neuronale Netze. Eine Einführung, o. J., <http://neuralesnetz.de> (Zugriffsdatum: 10.06.2022)

Saha, S.: A Comprehensive Guide to Convolutional Neural Networks – the ELI5 way. *Towards Data Science*, 15.12.2018, <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53> (Zugriffsdatum: 10.06.2022)

o. V.: Hieroglyphen. Das Alphabet der Ägypter und wie es zu lesen ist. *GEO*, o. J., <https://www.geo.de/magazine/geo-epoche-kollektion/19064-rtkl-hieroglyphen-das-alphabet-der-aegypter-und-wie-es-zu-lesen> (Zugriffsdatum: 21.05.2022)

Bildnachweise

Titelbild: pixabay, <https://pixabay.com/de/illustrations/k%3%bcnstliche-intelligenz-netzwerk-3706562>, <https://pixabay.com/de/photos/%c3%a4gypten-schrift-zeichen-wand-gold-2226780>

Abbildung 1: *IEEE Access*, https://ieeexplore.ieee.org/mediastore_new/IEEE/content/media/62876399312710/9528382/argen2ab-3110082-large.gif

Abbildung 2: DeepMind, <https://github.com/deepmind/ithaca/blob/main/images/inscription.png>

Abbildung 3: *Towards Data Science*, <http://neuralesnetz.de/nbilder/neuronennetze/neuronennetz.gif>

Abbildung 4: *neuralesnetz.de*, <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>

Abbildung 5: DeepMind, <https://raw.githubusercontent.com/deepmind/ithaca/main/images/ithaca-arch.png>

Abbildung 6: *IEEE Access*, https://ieeexplore.ieee.org/mediastore_new/IEEE/content/media/6287639/9312710/9528382/argen4-3110082-large.gif

Abbildung 7: Google Fabricius Workbench, <https://fabriciusworkbench.withgoogle.com>