

# Erkennung von ägyptischen Hieroglyphen in historischen Büchern mit neuronalen Netzen

Philipp Lang

## Einleitung

Bei der Digitalisierung von gedruckten Büchern wird Schrifterkennung eingesetzt, um deren Text durchsuchbar zu machen. Schrifterkennung ist keine neue Technik. Mit Hieroglyphen sieht man allerdings alt aus. Während es bereits zahlreiche Methoden und Programme gibt, um Text zu erkennen, der im lateinischen Schriftsystem verfasst ist, so steckt das Erkennen von Hieroglyphen noch in den Anfängen.

In der Masterarbeit „Objekterkennung mit Hilfe von Convolutional Neural Networks am Beispiel ägyptischer Hieroglyphen“ von J. Nathanael Philipp wurde bereits das Klassifizieren von Hieroglyphen untersucht. Um diese Technologie auf die vorliegenden historischen Bücher anwenden zu können, müssen zunächst jedoch die Hieroglyphen auf den Buchseiten lokalisiert werden.

## Ziel

In dieser Arbeit soll ein künstliches neuronales Netz lernen, ägyptische Hieroglyphen auf eingescannten historischen Buchseiten zu lokalisieren. Die Buchseiten beinhalten jedoch nicht nur Hieroglyphen, sondern auch Abbildungen und Text im lateinischen Schriftsystem. Eine weitere Schwierigkeit liegt wie so oft im Fehlen der Trainingsdaten für das neuronale Netz.

## Vorgehen

Als Netzarchitektur dient das 2015 von Ronneberger, Fischer und Brox vorgestellte U-Net (siehe Abb. 1), welches gestaltet wurde, um mit wenigen Trainingsdaten umgehen zu können. Die U-Form besteht aus einem zusammenziehenden und einem erweiternden Zweig. Ersterer verkleinert das Bild stufenweise und überführt die Bildinformationen dabei in Kontext, während letzterer das Bild wieder vergrößert und dabei eine präzise Lokalisierung des Kontexts ermöglicht.

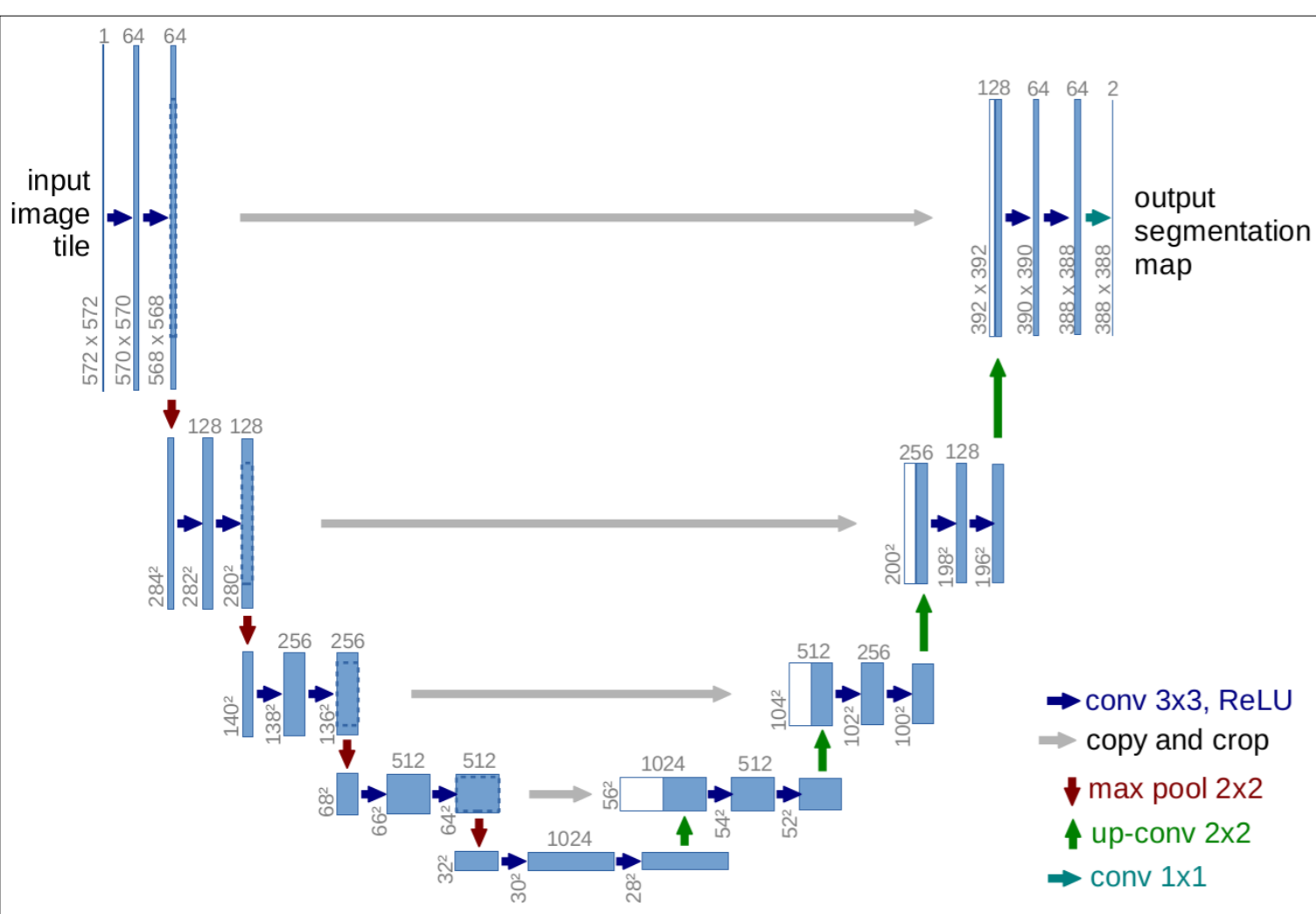


Abb. 1: Künstliches neuronales Netz namens U-Net aus Ronneberger, et al., 2015.

Das Netz soll lernen, erkennen zu können, wo sich auf einer Buchseite Hieroglyphen befinden. Dazu benötigt es Trainingsdaten. Diese bestehen aus beispielhaften Eingaben, in diesem Fall generierte Pseudobuchseiten, und den dazugehörigen, gewünschten Ausgaben, in diesem Fall ein Bild mit Markierungen. Eine beispielhafte Darstellung befindet sich unterhalb in Abb. 2.

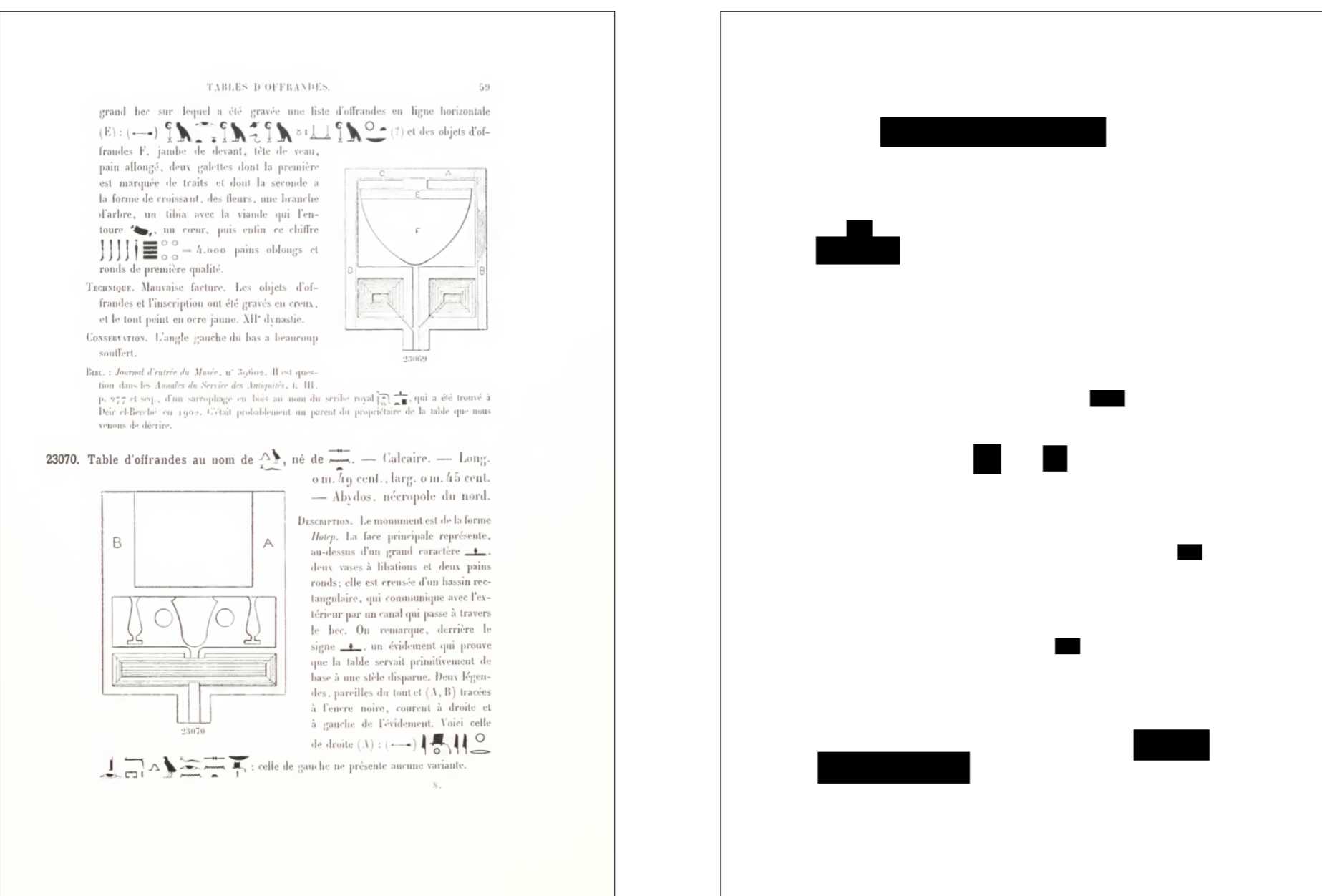


Abb. 2: Beispiel für eine gewünschte Ein- und Ausgaben des neuronalen Netzes

Mit einer Sammlung von ca. 6.000 Bildern aus Hieroglyphen der Gardiner-Liste in 3 bis 6 Varianten werden Pseudobuchseiten generiert, um diese als Trainingsdaten zu verwenden. Die Pseudobuchseiten werden kontinuierlich und zufällig aus einem Hintergrund, Pseudotext und Hieroglyphen zusammengesetzt. Dabei sind viele Parameter konfigurierbar und können teilweise mit einer Toleranz oder als Wertebereich definiert werden. Aus diesem wird zufällig ein Wert gezogen. Auf einem vorgegebenen Hintergrund wird ein durch Zufall generierter Pseudotext aus vorgegebenen Schriftzeichen und Wortlängen positioniert. Darauf wird eine zufällige Anzahl von Hieroglyphen ohne Überlagerung platziert. Die durch Zufall gewählten Platzierungen werden dabei benutzt, um das gewünschte Ausgabebild zu erzeugen.

Bei der Generierung der Pseudobuchseiten kommen folgende Konfigurationsparameter zum Einsatz:

- Textparameter:
  - Liste von Zeichen
  - Liste von zulässigen Wortlängen
  - Liste von Schriftarten
  - Randgrößen
  - Schrifthelligkeit und Toleranz des Wertes
  - Schriftgröße und Toleranz des Wertes
  - Rotation und Toleranz des Wertes
- Hieroglyphenparameter:
  - Größe der Hieroglyphen
  - Liste von Hieroglyphenbildern
  - Mindest- und Maximalanzahl der Hieroglyphen pro Seite
- Seitenparameter:
  - Seitengröße
  - Liste von Hintergrundbildern

Um zu sehen, ob und wie gut das Netz auf den Trainingsdaten funktioniert, wurden diese Stück für Stück ausgebaut und komplexer, siehe Abb. 3. In den ersten Versuchen bestanden die Trainingsdaten nur aus einem kleinen Hintergrund aus purem Weiß, auf welchem jeweils nur eine Hieroglyphe platziert wurde. Darauf aufbauend wurden verrauschtere Hintergründe und mehrere Hieroglyphen pro Bild eingeführt bis hin zu viel größeren Bildern mit Text, wie oberhalb beschrieben. In Abb. 4 wird ein Lernprozess des neuronalen Netzes auszugsweise visualisiert. Weitere Auswertungen auf den Trainingsdaten sowie Auswertungen auf den echten Buchseiten und gegebenenfalls Verbesserungen sind noch in Arbeit.

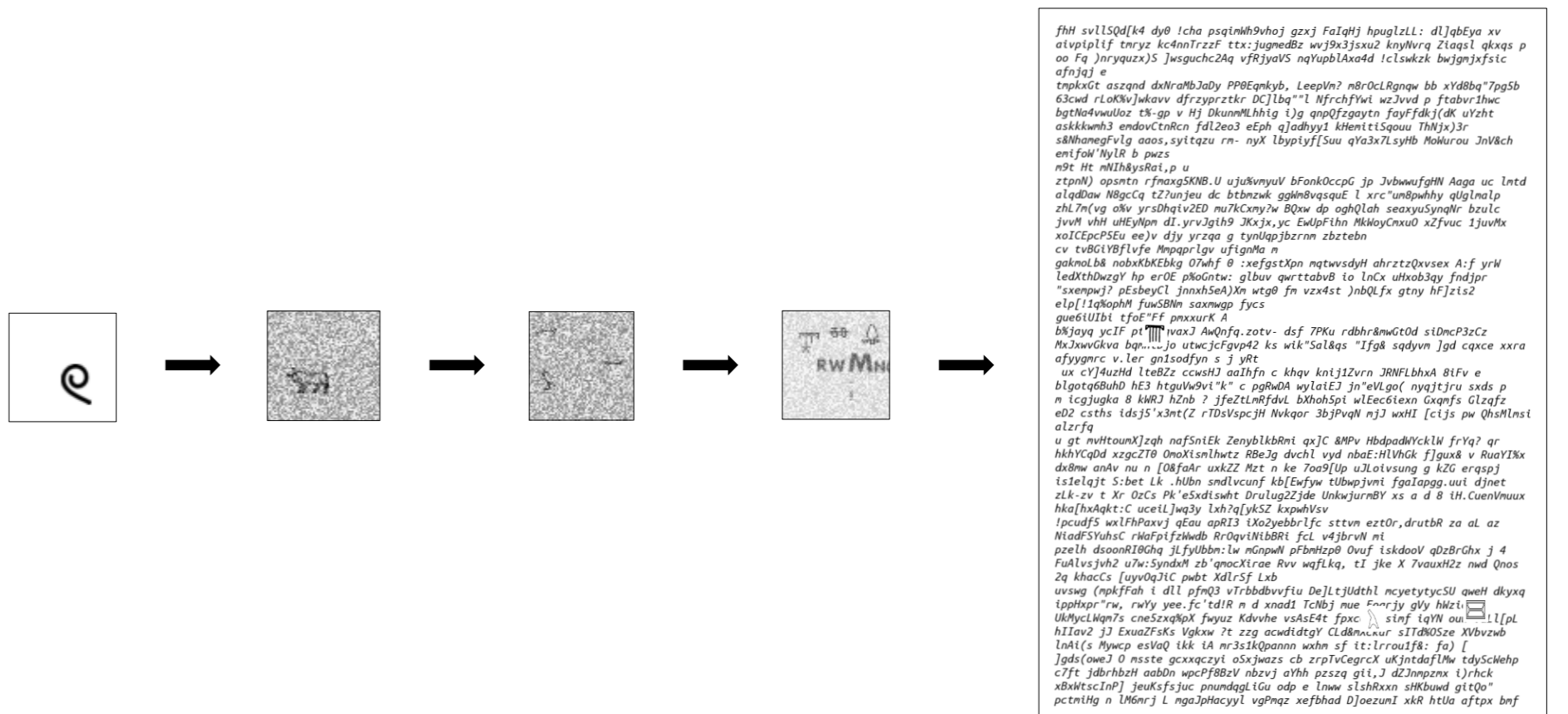


Abb. 3: Evolutionsschritte der Pseudobuchseiten

Epoche 0:									
Epoche 12:									
Epoche 24:									
Epoche 36:									

Abb. 4: Visualisierung des Lernprozesses des neuronalen Netzes über verschiedene Lernepochen hinweg. In den ersten Reihen befinden sich Eingabebilder, in den zweiten die Vorhersage des Netzes und in den dritten die ideale Vorhersage.