

Masterarbeit

KI-basierte Key Information Extraction aus gescannten Dokumenten mittels Optical Character Recognition

Framework



PYTORCH

Domäne



Bilderkennung



Computer Vision

Forschung



Applied AI

Ausgangslage

Im Zuge der digitalen Transformation von Unternehmen oder Behörden werden ganze Ordner an wichtigen Dokumenten händisch und manuell digitalisiert. In der Regel wird dafür das Dokument eingescannt und dann abgespeichert. Im nächsten Schritt werden die wichtigen Informationen aus den digitalen Kopien für die Weiterverarbeitung manuell abgelesen und händisch extrahiert.

Problemstellung

Ein Beispiel ist die Digitalisierung von Rechnungen. Diese wird i.d.R. händisch von Büroangestellten durchgeführt. Die manuelle Durchführung kostet Zeit und Geld und ist nicht mehr zeitgemäß. Heutzutage hat die automatisierte Bilderkennung im Bereich der Computer Vision teilweise bereits übermenschliches Niveau erreicht. Dabei ist die Erkennung von Schriftzeichen, Wörtern und ganzen Sätzen auf Bildern problemlos möglich. Für beliebige Rechnungen besteht die Herausforderung deshalb darin, aus den erkannten Zeichenketten die wichtigen Informationen zu extrahieren. Beispiele hierfür sind die Rechnungsnummer, Kundennummer, Absender oder Empfänger der Rechnung.



Vorgehensweise und Erwartete Ergebnisse

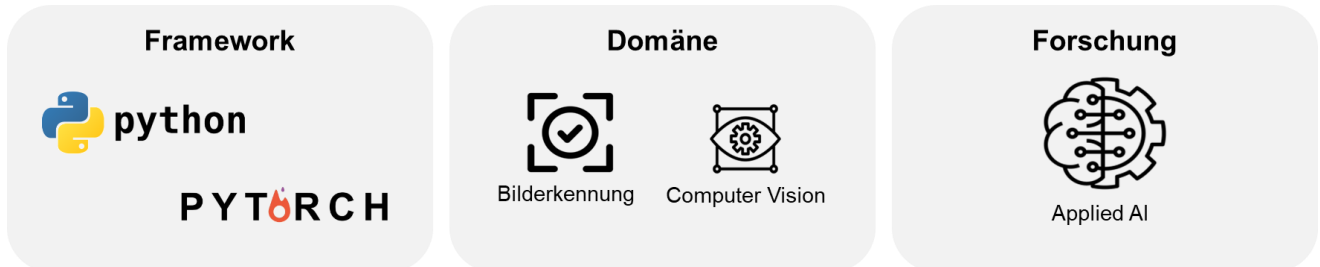
Zuallererst wird eine ausführliche Literaturrecherche zu Computer Vision und Algorithmen der optischen Zeichenerkennung mittels Deep Learning durchgeführt werden, um einen Überblick für den aktuellen Stand der Forschung zu erhalten. Dann werden mit Hilfe eines tiefen neuronalen Netztes für Optische Zeichenerkennung Zeichenketten aus beliebigen Rechnungen extrahiert. Aus den extrahierten Zeichenketten werden dann die wichtigen Informationen herausgefiltert. Das Verfahren wird mit klassischen Verfahren verglichen. Das Vorgehen und die Ergebnisse werden in einer schriftlichen Ausarbeitung dokumentiert und in einem Vortrag vorgestellt.

Ansprechpartner

Richard Meyes | E-Mail: meyes@uni-wuppertal.de

Master Thesis

Deep Learning-Based Key Feature Extraction of Scanned Documents with Optical Character Recognition Algorithms



Initial Situation

Digital transformation is a process many organizations carry out. Regarding important documents in this process; they are typically scanned and saved. For further processing, the key features of the digital copies need to be extracted manually.

Problem Definition

One example is the digitization of invoices which is usually carried out manually. Manual digitalization of invoices is labor-intensive, comes with high costs and takes its time. Nowadays, automated image recognition has reached superhuman performance. Thereby, the recognition of handwritten characters, strings, and whole sentences in images is feasible. The challenge here is to extract the key features from the recognized strings for any invoice. Examples of key features are the invoice number, customer number, sender, or recipient of the invoice.



Methods and Expected Results

First, an extensive literature research is conducted to get an overview of state-of-the-art methods in the field of computer vision and optical character recognition (OCR) with deep learning-based methods. Thereafter, the characters of arbitrary invoices are extracted via OCR deep learning-based methods. Then, the strings are assigned to the wanted key features with high accuracy. The approach is evaluated and compared against traditional OCR methods. Finally, the approach and results are written down in a thesis and afterward presented in a colloquium.

Contact Person

Richard Meyes | **E-Mail:** meyes@uni-wuppertal.de