

Bastian BRANDSTÖTTER, Philipp GABRIEL

Semante

Betreuer: Franz Reitinger

The purpose of this graduation project is to create a technology that can interpret images and form sentences to describe them. This will be implemented through a so-called machine-learning model which takes image data and converts it into text data (sentences). After said model is completed, it needs to be tested and optimized for better results. The resulting technology will be deployed as a programming library and a mobile app.

Aufgabenstellung:

Im Rahmen unserer Diplomarbeit beschäftigen wir uns mit Künstlicher Intelligenz, wobei mit Hilfe eines neuronalen Netzes Wissen angelernt wird. Das Anlernen wird mit einem Machine-Learning-Modell realisiert, wobei auf verfügbare Software-Infrastruktur zurückgegriffen wird. Als UseCase haben wir uns für die Interpretation von Bildern entschieden, die man dann in Sätzen beschreiben kann. Das zu entwickelnde Machine-Learning-Modell nimmt Bild-daten entgegen und liefert Textdaten in Form von Sätzen zurück. Die erzeugten Sätze können dann beispielsweise in Braille-Schrift transferiert werden, um blinden

und sehbehinderten Menschen die Möglichkeit zu geben, den Inhalt von Bildern zu erfassen.

Anschließend soll dieses Modell noch optimiert und zur Steigerung der Erkennungsgenauigkeit angepasst werden. Die daraus resultierende Technologie soll in Form einer Softwarebibliothek und einer mobilen App zur Verfügung gestellt werden.

Mobile App:

Die App soll es Benutzern ermöglichen, die Technologie auszuprobieren. Sie können Bilder aufnehmen, diese werden von der App lokal interpretiert und drei Sätze in absteigender Konfidenz generiert, die das

Bild möglichst akkurat beschreiben.



Abb. 1: Mobile App

Web-Interface:

Weiters wurde ein Web-interface für das Machine-Learning-Modell erstellt, um die Bildbeschreibungen über eine Schnittstelle im Internet abrufbar zu machen.