

Stream-Based Visual Analytics with Big Data Technology

Im Internet gibt es Unmengen an Informationen über Reisedestinationen in Form von Bildern. Zusätzlich zu den Bildern gibt es jedoch noch weitere nützliche Metadaten, z.B. Zeit und Ort der Erstellung des Bildes oder bestimmte Labels, die das Bild beschreiben wie z.B. «Strand», «Hütte», «Palme», «Sonne», etc.

Ziel der Arbeit ist es, Fragen der Art «Wo sind die schönsten Strände in Südamerika?» mittels klassischer Datenbanksuche und Bildanalyse beantworten zu können. In dieser Arbeit sollen unterschiedliche Lösungsansätze untersucht werden, um sowohl die Metadaten der Bilder als auch die Bilder selbst in natürlicher Sprache zu durchsuchen.

Zur Beantwortung der Frage wurde eine Lösung anhand der Big Data Engine *Apache Spark* und der Bibliothek für maschinelles Lernen *TensorFlow* implementiert. Dabei wurde regelmässig eine Sammlung von öffentlichen Webcams abgefragt. Die gewonnenen Bilder wurden dabei bezüglich Objekte und Szenen analysiert. Die Resultate wurden auf einer öffentlich zugänglichen Webseite dargestellt, auf der Ausgaben nach verschiedenen Kategorien geordnet (beispielsweise Seelandschaft, Alpenpanorama oder Hafenszene) auswählbar waren.

Die Ergebnisse zeigten ein durchwachsenes Bild. Einerseits haben gewisse Aspekte wie die Objekterkennung gut funktioniert, andererseits haben andere Funktionen wie die Szenenerkennung wesentlich mehr Probleme bereitet. Darunter waren Bilder von vernachlässigten Webcams, die nur ein Standbild in einer gewissen Farbe oder einen Textabschnitt als Platzhalter anzeigten, sodass die Kategorie eines Bildes immer dieselbe blieb.

Allgemein gestaltete sich, durch die Einbindung von TensorFlow in Form von benutzerdefinierten Funktionen in Spark, die Optimierung der Leistung als schwierig. Die Effizienz der Lösung war zum Schluss beeindruckend, jedoch konnte der letzte Schritt, nämlich die Zuordnung der kategorisierten Bilder in schöne und weniger schöne Szenarien, was die eigentliche Fragestellung dieser Arbeit ist, durch die Vielzahl an technischen und konzeptuellen Hindernissen nicht beantwortet werden.

Es konnten dennoch wertvolle Erkenntnisse über die verwendeten Werkzeuge gesammelt und eine gute Basis für die Weiterentwicklung der Lösung erarbeitet werden.



Diplomierende
Benjamin Jean Bollhalter
Daniel Djordjevic
Antonio Parenti

Dozierende
Helmut Grabner
Kurt Stockinger



Auswahl von gut erkannten Objekten und Szenen inklusive Voraussagewahrscheinlichkeit mit TensorFlow anhand der Modelle ResNet50 und VGG16



Durchschnittliche Prozessierungsgeschwindigkeit der zur Bildanalyse in TensorFlow verwendeten Modelle ResNet50 und VGG16 innerhalb eines Spark-Clusters mit fünf Maschinen