

THE „FIND WALTER“ PROBLEM

By Aaron, Bruno, Sandra, Timur

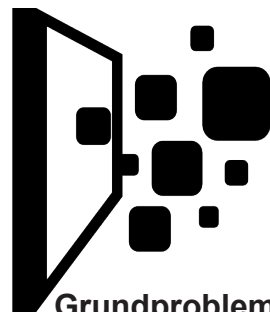


Prototyp A

Ein Framework mit einem vortrainierten Inception v2 Modell für MSCOCO Dataset (COCO ist ein Object detection Dataset).

Mehrheitlich erfolgreich, aber z.T. auch sehr falsch...
Wir gehen davon aus, dass wir zu wenig Trainingsdaten haben und die Bilder zu hochauflösend sind.

Trainiert wurde mit Vollaufgelösten Bildern (nicht mit Stücken wie in den Lösungen)



Grundproblem: Zu wenige Trainings Daten

Lösung: Random Bilder aus dem Internet in Schnipsel aufteilen. Diese sind dann die Negatives (not Wally).

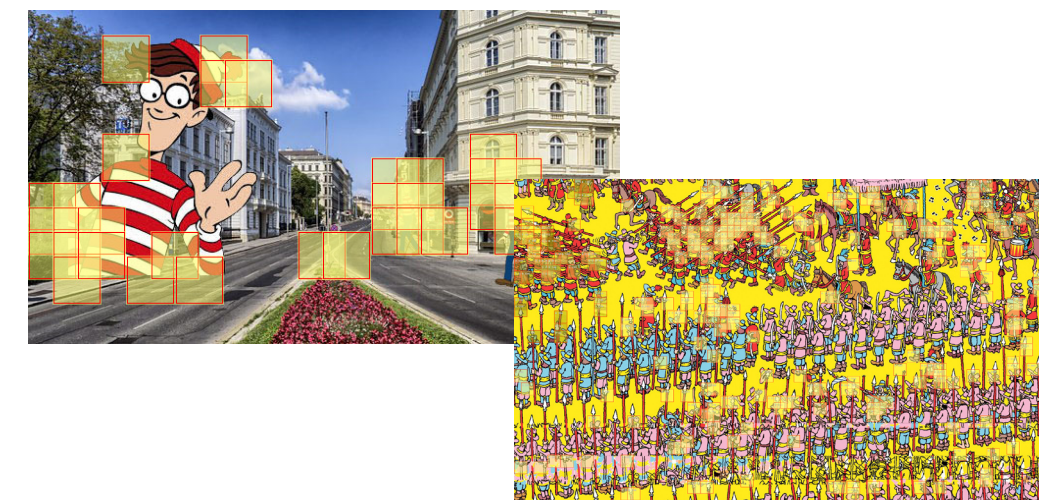
Für die Positives: Einige Kopf-Ausschnitte von Wally random mässig auf gewisse Negatives platziert und diese dann als Positives Klassifiziert.

Somit bestehen die Trainingsdaten nun aus rund 100 Tausend Bildauschnitten im JPEG-Format in der Grösse von 50 x 50 Pixel.



Lösung A

- Framework mit vortrainiertem Modell (Inception V3)
- Verwendung von „Bottleneck-Files“ um die finale Klasse der Trainingsbilder - von den vortrainierten Klassen zu unterscheiden.
- Training auf 50x50 Bilder (welche zusammengesetzt das Ausgangsbild ergeben)
- Treffermarkierung mittels gelben Boxen



Das Problem

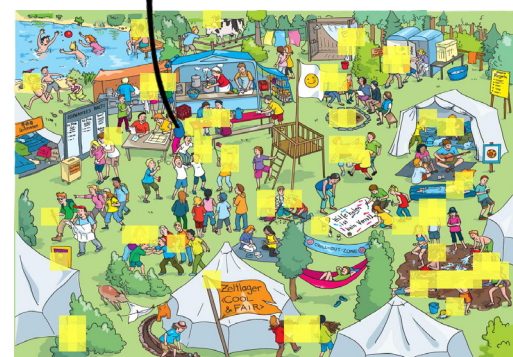
„Wo ist Walter?“ ist ein Kinderbuch mit gezeichneten Wimmelbildern auf denen die Figur Walter zu finden ist.

Beim Walter-Problem geht es A, zu lernen wie ein Objekt in verschiedenen Situationen identifiziert werden kann.

Ziel der vorliegenden Arbeit ist es, mit Hilfe eines neuronalen Netzes das Walter-Problem möglichst effizient zu lösen.

Dazu sollen drei möglichst eigenständige Modelle entwickelt werden.

Die „Wo ist Walter?“-Bücher wurden bisher ca. 43 Millionen Mal verkauft, allerdings ist die Anzahl der Wimmelbilder, aus naheliegenden Gründen, beschränkt.



Lösung B

1. Versuch: Vorverarbeiten der Bilddateien in Fließkommazahlen-Tensoren. Modell mit abwechselndem Conv2D(ReLu) und MaxPooling2D Layern.
-> Binärklassifizierung Waldo vs. not Waldo
- RMSprop optimizer und Sigmoid als Loss

Ergebniss: Zu wenig genau.

2. Versuch: Neue Trainingsdaten generiert, 400k Trainingsbilder und 100k Testbilder. (Davon 80k Positives)
- Dropout integriert in Modell

Ergebnis: Findet Waldo genauer, aber noch nicht eindeutig.

