

Objektdetektion des Inventars von Zivilschutzanlagen

Im Rahmen dieser Masterthesis wird die Objektdetektion des Inventars von Zivilschutzanlagen und Schutzbauten aus photographischen Bilddaten und Laser-Scanning Punktwolken untersucht. Ziel ist es, einen Prototyp zu implementieren, der eine möglichst automatische Erkennung inventarisierbarer Objekte innerhalb der unterirdischen Anlagen ermöglicht. Insbesondere neue Verfahren der künstlichen Intelligenz versprechen hierbei ein hohes Potential. Die Integrierbarkeit dieser Technologie in ein mobiles Handgerät ist erst in den letzten zwei Jahren durch die rapide Steigerung der Berechnungsgeschwindigkeit realistisch geworden. Verschiedene moderne Frameworks und Laufzeitumgebungen grosser Technologiekonzerne werden hierzu evaluiert. Eine prototypische Umsetzung einer spezifischen Objekterkennungs-App demonstriert die Skalierbarkeit der Technologie.

Projektrahmen

Der Schutz der Schweizer Bevölkerung vor Katastrophen und Nötfällen wird durch das Bundesamt für Bevölkerungsschutz gewährleistet (BABS, 2012). Landesweit gibt es rund 360'000 kleinere Personenschutzräume sowie 1'700 grössere Schutzanlagen. Derzeit befindet sich eine umfangreiche Digitalisierungsstrategie in der Entwicklungsphase.

Ziel

In diesem Projekt wurde einen Prototyp realisiert, der eine möglichst automatische Erkennung von 39 Klassen inventarisierbarer Objekte wie z.B. Panzertüren, Luftfilter, Explosionsschutzventile, Feuerlöscher oder Dieselgeneratoren in unterirdischen Zivilschutzanlagen ermöglicht. Es soll künstliche Intelligenz zur Anwendung kommen, welche die eigentliche Detektion (sog. "Inferencing") lokal auf einem Mobilgerät durchführen kann, ohne auf eine teure Serverinfrastruktur oder ausgelagerte Dienste angewiesen zu sein.

Welches Deep Learning Modell?

Single-Shot-Detektoren (SSD) sind Deep Learning Architekturen aus der Klasse konvolutionaler neuronaler Netzwerke (CNN). Diese Systeme sind besonders beim Training, aber auch bei der Anwendung auf neue Daten normalerweise enorm **rechenintensiv** und benötigen **teure Hochleistungshardware**. SSDs stellen eine einzigartige moderne Form schlanker, Performance-optimierter Netzwerke dar (**Serverless**) und stehen dabei für hohe Geschwindigkeit und Genauigkeit (Liu, 2016).

Umsetzung

Das quantisierte, unskalierte Modell wies hohe Detektions-raten auf und kann mit sehr performanten ca. **25 FPS** Inferenzgeschwindigkeit auf einem **Standard-Smartphone** ausgeführt werden. Zum Training wurden 5'065 Objekte manuell auf 2'086 Bildern erfasst. Die **Endnutzer-Detektionsrate liegt bei 84%**.

SSD MobileNet Modell	Original	Skaliert	Quantisiert	Quantisiert & Skaliert
Latenz	195 ms	145 ms	82 ms	74 ms
Modellgrösse	27 Mb	16 Mb	6 Mb	4 Mb
Detektionsrate (mAP@IoU>0.5)	74 %	66 %	71 %	68 %

Komplexe Szene	Ground Truth	Original	Skaliert	Quantisiert	Quantisiert & Skaliert
Standard Szene	Ground Truth				

Abbildung 1: Vergleich der DL-Modell. Grün = Verwendung in App

Quellennachweise

Aising, O. (2018). Mobile Object Detection using TensorFlow Lite and Transfer Learning.
BABS. (2012). Schutzbauten, Bundesamt für Bevölkerungsschutz (BABS). [28.9.2019].
Krizhevsky, Alex; Sutskever, Ilya and Hinton, Geoffrey E. (2012): ImageNet Classification with Deep Convolutional Neural Networks. S. 1097-1105.
Huang, J., 2018. Google AI Blog: Accelerated Training and Inference with the TensorFlow Object Detection API [WWW Document]. URL: <https://ai.googleblog.com/2018/07/accelerated-training-and-inference-with.html> (accessed 2.3.20).
Liu, W., Anguelov, D., Erhan, D., Szegegy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. In European conference on computer vision (pp. 21-37). Springer, Cham.

Etappe 1: Mobile Objektdetektion



Etappe 2: App-Entwicklung



Abbildung 2: Roadmap von der Datenbeschaffung bis zur fertigen Applikation. (Logos Eigentum der jeweiligen Gesellschaften)

Herausforderung: Optimierung für ein Smartphone

Waren vor wenigen Jahren noch grosse Rechenzentren für die Objekterkennung auf Bildern erforderlich (Krizhevsky, 2012), so hat die Rechenleistung von Smartphones derart schnell zugenommen, dass sich heute zum ersten Mal ein Potential für die günstige, lokale Anwendung von Deep Learning ergibt (Aising, 2018).

Das Modell muss jedoch speziell angepasst werden: Diesen Prozess nennt man **Quantisierung** (Huang, 2018). Hierbei werden die komplexen mathematisch-multidimensionalen Matrixstapel (sog. Tensoren) so komprimiert, dass ihre einzelnen Untereinheiten in ein 8-bit-Schema passen, statt auf Fließkommazahlen zu basieren. **Vorteil:** Sehr kleine Dateien, sehr flüssige Performanz.

Nachteil: Die Genauigkeit kann abnehmen, wenn das Modell nicht sehr genau auf den Datensatz abgestimmt wird. Auch eine **Faktor-Skalierung** (Verkleinerung) des Netzwerks ist eine Möglichkeit, Ressourcen zu sparen. **Vorteil:** Niedrige Latenzen. **Nachteil:** Nur wenig schneller, aber viel ungenauer.

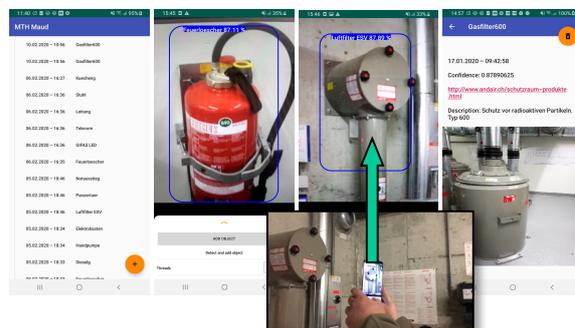


Abbildung 3: Screenshots aus der App (v.l.n.r.): Item-Liste mit detektierten Objekten (exportierbar); Detektion eines Feuerlöschers; Detektion eines Explosionsschutz-Ventils (mit Foto Live-Nutzung); Detailseite eines erkannten Gasfilters mit Confidence Score.

Fazit

Mit dem vorgeschlagenen Ansatz konnte dem Bundesamt für Bevölkerungsschutz ein **voll funktionsfähiger Prototyp** des Objektdetektors übergeben werden. Der Applikation wurden in Usability-Tests eine **hohe Effizienz, Stabilität und Nutzerfreundlichkeit** attestiert.

