

ABSTRACT

Heating and ventilation account for significant costs in buildings, both public and private. Inefficient heating wastes energy, inadequate ventilation may be detrimental to a person's well-being. In private buildings, factors such as a known work schedule can be used to turn off heating and ventilation when nobody is present. Public buildings and office spaces usually don't offer such possibilities to detect whether parts of a building or certain rooms are being used. Determining usage e. g. through timetables in universities is inaccurate and may not overlap with actual usage. Instead, classic occupancy detection sensors can be used. Simple detectors such as motion sensors can only detect whether people are present but cannot provide information about their quantity, while advanced sensors such as depth cameras are prohibitively expensive, making widespread use unattractive. Additionally, video-based counting techniques that transfer live video from each room to a central location for counting raise significant privacy concerns and offer an attack surface for cybercriminals to extract highly-sensitive data. To alleviate these concerns and to allow for widespread use, this thesis has the central theme of implementing an occupancy detection system using an overhead mounted camera. Cost-efficient embedded hardware shall be used to count people on-device using neural networks. The main research question of this work asks how such a system could be implemented on real hardware and which techniques and frameworks aid such an implementation. To answer this question, the specific challenges of running neural networks on resource-constrained embedded systems are illustrated based on a selection of devices. Existing neural network architectures are researched and chosen based on whether they fit these constraints. The selection of networks is then trained on several datasets and the networks are evaluated regarding both their object detection metrics and their performance on the counting task itself. Techniques such as quantization and teacher-student training are also evaluated in regards to whether they aid in preparing neural networks for the specific challenges of this resource-constrained context. Results of the experiments with these object detection networks show that training on a large, but not representative dataset results in better counting results on the used test dataset, but the significance of these results remains unclear due to the low size and variance of the given dataset. Quantization can be confirmed as a valuable tool to lessen resource demands on storage for network parameters, while teacher-student training does not offer benefits. The implementation on real hardware uses several different frameworks and runtimes to convert the neural networks and illustrates a strong fragmentation of the technologies used, while training specialized neural nets for this task is hindered by a lack of freely-available, representative datasets.

ZUSAMMENFASSUNG

Das Beheizen und Belüften von Gebäuden stellt einen sehr kostspieligen Aspekt des Betriebes öffentlicher als auch privater Gebäude dar. Ein unnötiges Heizen verursacht hohe Kosten und verschwendet Energie, während eine unzureichende Belüftung zu Problemen wie Konzentrationsstörungen bei anwesenden Personen führen kann. In privaten Wohngebäuden lässt sich die Gebäudeklimatisierung mit Konzepten des Smart Homes intelligent nach Bedarf steuern, in öffentlichen Gebäuden oder Büros ist es hingegen schwerer, die Auslastung eines Gebäudeteils oder einzelner Räume zu erkennen. Das Heizen nach Belegplänen ist ungenau, klassische Gebäudesensorik wie Bewegungsmelder können nur die Anwesenheit von Personen, aber nicht deren genaue Anzahl bestimmen. Vorhandene Systeme zur Personenzählung nutzen kostspielige Sensortypen wie Tiefenkameras, welche einen flächendeckenden Einsatz unattraktiv machen. Ebenfalls kommen bei Systemen, welche Videodaten zur Auswertung zu einer Zentrale übertragen, Datenschutzbedenken auf. In dieser Arbeit soll aus diesem Grund die Personenzählung durch eine zentral in einem Raum an der Decke montierte Kamera betrachtet werden. Hierbei soll die Zählung durch neuronale Netze auf kostengünstiger und somit leistungsschwacher eingebetteter Hardware durchgeführt und untersucht werden. Es wird sich der Frage gewidmet, wie neuronale Netze zur Personenzählung unter diesen Bedingungen implementiert werden können. Zur Beantwortung dieser Frage wird die Problemstellung anhand einer Hardwareauswahl mehrerer geeigneter Geräte illustriert. Durch eine Recherche existierender neuronaler Netze zur Objektdetektion werden solche identifiziert, welche unter den Rahmenbedingungen der geringen Leistung der Geräte einsetzbar wären und diese Netze mit geeigneten Datensätzen auf die Erkennung von Personen angepasst. Weiterhin werden problemspezifische Techniken wie Quantisierung und Teacher-Student-Training hinsichtlich ihrer Wirksamkeit zur Verkleinerung neuronaler Netze untersucht. Nach Evaluation der erhaltenen Netze hinsichtlich ihrer Zählqualität und Implementierung auf echter Hardware kann festgestellt werden, dass Netze, welche auf einem großen, aber für die Problemstellung nicht repräsentativen Datensatz trainiert wurden, eine ausreichend gute Personenzählung ermöglichen. Das Konzept der Quantisierung kann als nützliche Technik zur Verkleinerung neuronaler Netze bestätigt werden, während Teacher-Student-Training keine leistungsfähigeren Netze als ein direktes Training ergibt. Es kann identifiziert werden, dass eine Implementierung auf realer Hardware durch eine starke Fragmentierung des Marktes in viele verschiedene Laufzeitumgebungen und Frameworks herausfordernd ist, während das Training der Modelle durch einen Mangel an öffentlichen Datensätzen erschwert wird.