Poster Abstract: Camera-Assisted Training of Non-Vision Sensors for Anomaly Detection

Norah Albazzai
School of Computer Science and Informatics, Cardiff University
Cardiff, UK
AlbazzaiNA@cardiff.ac.uk

Omer Rana
School of Computer Science and Informatics, Cardiff University
Cardiff, UK
RanaOF@cardiff.ac.uk

Charith Perera
School of Computer Science and Informatics, Cardiff University
Cardiff, UK
PereraC@cardiff.ac.uk

ABSTRACT

Cameras are becoming pervasive and used for image classification and object detection in various applications, including anomaly detection. However, cameras pose a privacy threat and require significant power resources. To address these issues, researchers have explored non-vision sensors, but pre-training them for anomaly detection is challenging because anomalies are difficult to define and vary significantly across indoor environments. Thus, we propose a new approach to training non-vision sensors using a tiny camera and a pre-trained MobileNetV2 model. Data from non-vision sensors are labelled based on the image classification from the tiny camera, and an anomaly detection model is trained using these labelled data. The Random Forest model is used as the final model, achieving an accuracy of 95.58%.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in ubiquitous and mobile computing**.

KEYWORDS

camera, sensors, anomaly detection

ACM Reference Format:
Norah Albazzai, Omer Rana, and Charith Perera. 2023. Poster Abstract: Camera-Assisted Training of Non-Vision Sensors for Anomaly Detection. In *International Conference on Internet-of-Things Design and Implementation (IoTDI '23), May 9–12, 2023, San Antonio, TX, USA*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3576842.3589164

1 INTRODUCTION

Cameras are increasingly prevalent across various settings, including public spaces, businesses, and moving into households. Cameras are more advanced sensors having the potential to provide highly detailed and contextually rich sensory information. In addition, they offer various computer vision functions, such as image classification, that can run on resource-constrained devices. Computer vision has been utilised for multiple applications, including anomaly detection. However, despite their advantages, cameras are energy-hungry sensors. More importantly, they are considered an invasion of privacy, and individuals are hesitant to have visual monitoring, especially in indoor settings. In addition, computer vision systems require significant training data to function effectively.

Researchers have utilised non-vision sensors (e.g., temperature, humidity) for anomaly detection within built environments to address the camera's challenges. However, these sensors face challenges, such as difficulty in pre-training them due to the elusive nature of anomalies and the diversity of environments. They require on-premises training which can be impractical for multiple scenarios. Additionally, the need for a labelled dataset for accurate detection can be time-consuming, expensive, and prone to human error [4].

To overcome the issues above, we propose an approach that leverages the strengths of cameras and non-vision sensors. A tiny camera sensor is used to *train* commonly installed non-vision sensors to detect anomalies in indoor environments. The primary objective for the camera is to produce *labels*, which can then be utilised to train other sensors through supervised learning. During the deployment, the camera can be removed to preserve privacy, and only non-vision sensors contribute to making the final predictions.

In a similar study [1], the authors introduced a self-training sensor system for recognising human activities. For example, they employed a camera sensor to train an accelerometer.

2 METHODOLOGY

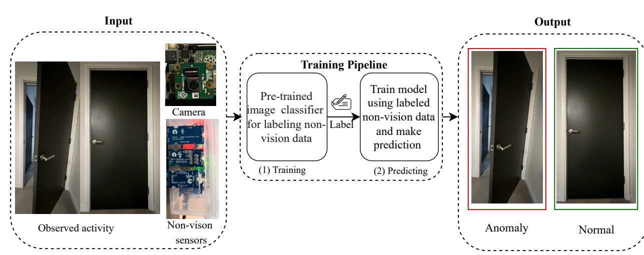

Figure 1: The proposed approach consists of (1) Training: using a camera to label non-vision data, and (2) Predicting: using a classification model trained on non-vision data and labels from the camera.

We propose a sensor-based anomaly detection approach that utilises a tiny camera to train non-vision sensors. Our approach consists of two phases: training and predicting (see Figure 1). During the training phase, the camera produces labels that annotate

IoTDI '23, May 9–12, 2023, San Antonio, TX, USA

ACM ISBN 979-8-4007-0037-8/23/05…$15.00
https://doi.org/10.1145/3576842.3589164

non-vision data. In the predicting phase, a model trained on non-vision data and camera labels detects anomalies. The camera can be removed during deployment, and predictions are based solely on non-vision data. We evaluate our approach by implementing a case study that detects open doors as anomalies due to their potential security risks, impact on Heating, Ventilation and Air Conditioning (HVAC) systems, and energy usage [2]. The system implementation requires a typical CPU with a minimum clock-speed of 2 GHz and at least 8.0 GB RAM. The implementation stages are as follows:

2.1 Data Acquisition

We recorded environmental properties and found that opening and closing doors caused fluctuations in temperature, humidity, and pressure. Humidity was stable when the door was closed and varied when it was opened, while temperature suddenly dropped or rose when the door was opened. The observations indicate that door status has a significant impact on environmental properties, and temperature and humidity sensors are often used in buildings to manage energy usage and maintain thermal comfort. Similarly, in [5], temperature sensors were used to detect abnormal behavior resulting from an open window. We used an environmental module (for temperature, humidity, and pressure), door sensor, and RTC module to collect data, as well as a tiny camera module to capture images and train non-vision sensors. Additionally, the DeepDoors2 dataset was used [3], and data were collected from a home and an office for two weeks. We found that a 1-second sampling rate was sufficient for detecting short-term fluctuations in environmental conditions.

2.2 Phase 1: Training

In this phase, the tiny camera acted as a *trainer* and produced labels to annotate the non-vision data. Using a tiny camera, we employed a pre-trained MobileNetV2 model with transfer learning to classify the door state and generate labels for non-vision data. We trained with 4300 images of 96x96 image size and achieved 97.5% accuracy. After training the model, we ran inferences on door images. Then, we cross-referenced the timestamp of the non-vision data with the classified image to assign predicted classes as *labels* to the non-vision data.

2.3 Phase 2: Predicting

Once all non-vision data from home and office were labelled, we trained the final anomaly detection model. The inputs of the model are labelled non-vision data only. To select the best classifier, we experimented with three supervised algorithms: Deep Neural Networks (DNN), Support Vector Machine (SVM), and Random Forest (RF). The RF scored the highest results, followed by the DNN and SVM. As a result, we used the RF model as our final model. Using binary classification; the model classifies input data into two categories - *Normal* (i.e., closed door) or *Anomaly* (i.e., opened door).

3 RESULTS

To assess the effectiveness of our approach, we evaluated the labelled data in two different ways. Firstly, we compared the camera-generated labels with the actual labels from the door sensor and found only **four** misclassified instances. Secondly, we implemented two Random Forest models: (1) the baseline model trained on the ground-truth dataset and (2) the camera-trained model trained on the dataset labelled by the camera - and compared their performance. Table 1 demonstrates the performance of the two models on a test dataset using 10-fold cross-validation, and we can see that the camera-trained model succeeded in achieving almost identical results to the baseline model.

Table 1: Performance of the Baseline model (trained on ground-truth data) and the Camera-trained model (trained on data labelled by the camera).

Model	Accuracy	Precision	Recall	F1-score
Baseline model	95.67%	96%	96%	96%
Camera-trained model	**95.58%**	**96%**	**96%**	**96%**

Creating annotated datasets is crucial for training supervised models, but manual data annotation can be complex and may require expertise. Automatic data labelling can overcome many challenges associated with human labelling techniques for sensor data and can label large amounts of data quickly and cost-effectively. Using a camera (trained with 4300 images) to label non-vision data automatically, we addressed the challenges of manual labelling. In less than forty minutes, we labelled our non-vision dataset, which had more than ten thousand data points.

Our approach can be scaled to detect indoor anomalies, including open windows, crowded places that lead to high CO_2 concentration, and falls. However, the approach is limited to cases that have visual change and can be detected through cameras.

There are multiple factors that affect our approach performance. Firstly, although cameras possess a high anomaly detection capability, their performance can be restricted by factors such as lighting conditions, camera placement, and occlusions. Secondly, cameras and non-vision sensors capture data that differ greatly in nature, necessitating diverse processing techniques. Finlay, certain anomaly detection tasks may require advanced modeling techniques and more labeled data for effective training due to varying levels of complexity.

REFERENCES

[1] Nicola Bicocchi, Matteo Lasagni, Marco Mamei, Andrea Prati, Rita Cucchiara, and Franco Zambonelli. 2010. Unsupervised learning in body-area networks. In *Proceedings of the Fifth International Conference on Body Area Networks - BodyNets '10*. ACM Press, New York, New York, USA, 164. https://doi.org/10.1145/2221924.2221955

[2] Aditya Dhakal and K. K. Ramakrishnan. 2017. Machine learning at the network edge for automated home intrusion monitoring. In *2017 IEEE 25th International Conference on Network Protocols (ICNP)*, Vol. 2017-Octob. IEEE, Toronto, ON, Canada, 1–6. https://doi.org/10.1109/ICNP.2017.8117594

[3] João Gaspar Ramôa, Vasco Lopes, Luís A. Alexandre, and S. Mogo. 2021. Real-time 2D–3D door detection and state classification on a low-power device. *SN Applied Sciences* 3, 5 (5 2021), 590. https://doi.org/10.1007/s42452-021-04588-3

[4] Yuji Roh, Geon Heo, and Steven Euijong Whang. 2021. A Survey on Data Collection for Machine Learning: A Big Data - AI Integration Perspective. *IEEE Transactions on Knowledge and Data Engineering* 33, 4 (4 2021), 1328–1347. https://doi.org/10.1109/TKDE.2019.2946162

[5] Dumidu Wijayasekara, Ondrej Linda, Milos Manic, and Craig Rieger. 2014. Mining Building Energy Management System Data Using Fuzzy Anomaly Detection and Linguistic Descriptions. *IEEE Transactions on Industrial Informatics* 10, 3 (8 2014), 1829–1840. https://doi.org/10.1109/TII.2014.2328291

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009