Poster: Smart Surveillance as an Edge Service for Real-Time Human Detection and Tracking

Seyed Yahya Nikouei[†], Yu Chen[†], Timothy R. Faughnan[‡]
[†]Dept. of Electrical & Computer Engineering, Binghamton University, SUNY, Binghamton, NY 13902, USA
[‡]New York State University Police, Binghamton University, SUNY, Binghamton, NY 13902, USA

I. Introduction

Monitoring for security and well-being in highly populated areas is a critical issue for city administrators, policy makers and urban planners. As an essential part of many dynamic and critical data-driven tasks, situational awareness (SAW) provides decision-makers a deeper insight of the meaning of urban surveillance. Thus, surveillance measures are increasingly needed. However, traditional surveillance platforms are not scalable when more cameras are added to the network. Internet of Things (IoT) technology enabled edge computing helps to resolve the issue by allowing real-time surveillance to be achieved as delay and uncertainty of data transfer are minimized. Also, security and privacy are promised with IoT because of data processing at the site.

On the other hand, new challenges require attention for smart surveillance to function properly. In an edge application based on the hierarchy architecture as shown by Fig. 1, the edge node is responsible for collecting data from one or more sensors, performing lightweight preprocessing, and sending the results to a fog node where fundamental processing takes place and decisions are made. Less data transmission yields faster processing and result generation, thus the more edge gets involved, the less delay is introduced. In utilizing this algorithm for human monitoring and surveillance, the sensors are cameras and any processing involves dealing with video. In order to distinguish normal scene and abnormal behavior, various methods have been studied in the surveillance community [4]. They all have one step in common: identify pedestrians in the scene, then based on the features extracted from their movement or pose, carry out individual or combined behavior analysis. Using edge node to execute feature extraction from the video prevents additional lag and allows for the most minimal delay possible in decision making.

In this work, a smart surveillance as an edge service has been proposed. To accomplish the object detection, identification, and tracking tasks at the edge-fog layers, two novel lightweight algorithms are proposed for detection and tracking respectively [2], [3]. A prototype has been built to validate the feasibility of the idea, and the test results are very encouraging.

II. Architecture and Algorithm

Today's Convolutional Neural Networks (CNN) are able to conduct human classification accurately. Smaller and more innovative models were recently introduced allowing implementation of CNN on mobile devices [3]. However, feature extraction for classification and CNNs are too computationally heavy for edge nodes. A trade off is required between the accuracy and speed of the algorithm on the edge device. It is achievable by downsizing the general purpose models.

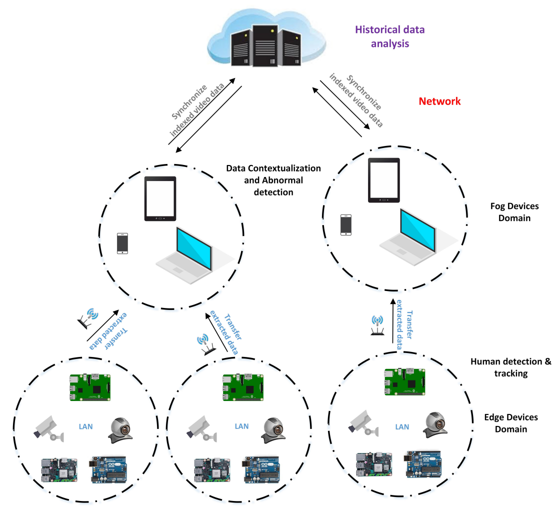

Fig. 1. Layered architecture for smart surveillance as an edge service.

Meanwhile, executing human detection algorithm on each frame is not helpful for feature extraction used in anomaly detection, even when it is computationally possible. The system needs to create a queue of the objects of interest and follow their movement such that the features used in abnormal behavior detection have a history of objects and their previous state. However, the detection achieved through today's contemporary methods only give positions in individual frames.

Based on these logistics, a tracking algorithm may be used alongside CNN in order to incorporate speed, accuracy, and a history of objects. Tracking algorithms such as Kernelized Correlation Filter (KCF) which are based on CPU implementation are best suited for the edge device because of their speed and lack of GPU on edge. Thus, the detection algorithm runs once in several frames and gives the position of all the pedestrians in the frame. The tracking algorithm will add the coordinates to its queue and refresh the position of each pedestrian in every incoming frame while deleting the objects who walk out of the frame. Figure 2 shows the steps the algorithm takes for human detection and tracking as a hybrid unified algorithm.

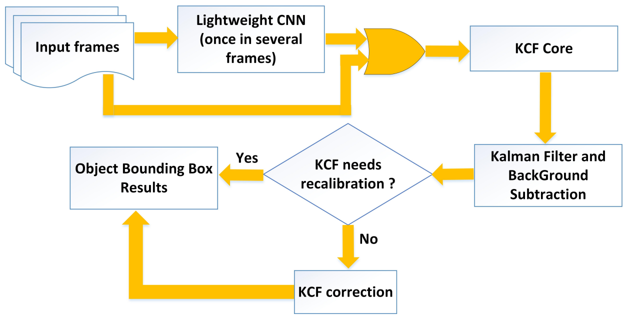

Fig. 2. Kerman: a hybrid human object detection and tracking algorithm.

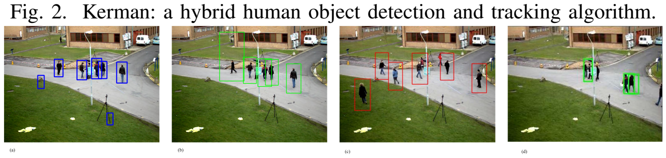

Fig. 3. Different Human Detection Algorithms implementation on edge. (a) Haar Cascaded. (b) HOG+SVM (c) SSD-GoogleNet. (d) L-CNN

III. Preliminary Experimental Results

Single Board Computers (SBC) are selected as edge devices. These boards have all the essential components of a standalone computer with a cost of less than \$80 per unit, making them the most efficient and cost effective for edge computation. The fog layer needs more computational power so devices such as tablets, laptops or smart phones can be utilized.

Applying the smart surveillance concept to the edge hierarchy should be evaluated using two criteria. Firstly, referring to either detection or tracking; the accuracy of the algorithm needs to be maintained while downsizing for the best performance of edge implementation. The second criterion considers how many Frames Per Second (FPS) the algorithm can achieve, in order to maintain a reasonable resolution of the features and prevent the system from crashing.

Figure 3 compares the detection performance obtained by implementing several different general purpose algorithms widely used today. Haar Cascaded is investigated because of its fast speed and its wide use for human face detection in cellphone cameras. HOG+SVM is considered because of its high accuracy in detection. And the GoogleNet is studied as an example of recent and well-performing CNN. We experimentally evaluated the feasibility of these algorithms to be applied for human detection task at the edge.

Figure 4 shows their overall performance on the edge device. The GoogleNet obviously is not a good candidate due to high memory consumptions and low speed, and our lightweight CNN (L-CNN) achieved higher speed with much less memory usage. Figure 5 compares the performance of our Kerman algorithm with several well-known tracking algorithms. The working flow of Keramn is illustrated in Fig. 2 which gives KCF more accurate performance. Figure 5 shows two scenarios: one scenario with approximately ten objects in the frame and the other scenario with only one or two objects in each frame to track. The average FPS is around eight after 30 seconds of runtime.

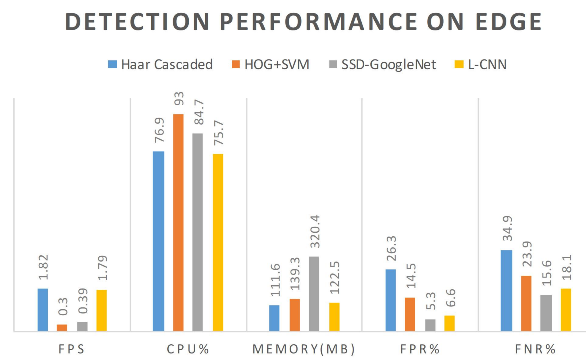

Fig. 4. Human Detection Performance on Edge lightweight CNN (L-CNN) is fast with good accuracy (FPR: False Positive rate, FNR: False Negative Rate)

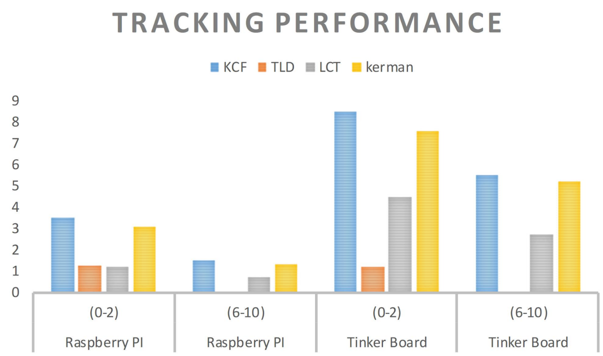

Fig. 5. Tracking algorithm performance on edge (TLD algorithm failed in case of many objects to track)

IV. Ongoing Work

The preliminary results show the feasibility to make smart surveillance as an edge service, more detailed discussions of our lightweight algorithms are available in [2] and [3]. Our current work is focused on two issues:

- following the detection and tracking, the next is anomalous behavior or activity detection. Under the umbrella of the edge-fog-cloud computing paradigm, new algorithms are being investigated to conduct the contextualization of features and human gestures by fog devices;
- beside performance requirements, security is among the top concerns of potential users. Lightweight but robust security network architecture is to be developed. One promising approach is a combination of blockchain technology and microservices architecture [1].

References

[1] D. Nagothu, R. Xu, S. Y. Nikouei, and Y. Chen, "A microservice-enabled architecture for smart surveillance using blockchain technology," *arXiv preprint arXiv:1807.07487*, 2018.

[2] S. Y. Nikouei, Y. Chen, and T. R. Faughnan, "Kerman: A hybrid lightweight tracking algorithm to enable smart surveillance as an edge service," *arXiv preprint arXiv:1808.02134*, 2018.

[3] S. Y. Nikouei, Y. Chen, S. Song, R. Xu, B.-Y. Choi, and T. R. Faughnan, "Intelligent surveillance as an edge network service: from harr-cascade, svm to a lightweight cnn," *arXiv preprint arXiv:1805.00331*, 2018.

[4] V. Tsakanikas and T. Dagiuklas, "Video surveillance systems-current status and future trends," *Computers & Electrical Engineering*, 2017.