INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Smart Surveillance System Real-Time Multi-Person Multi-Camera Tracking at the Edge

Shruti Khule[1], Supriya Jaybhay[2], Pranjal Metkari[3], Prof. Balasaheb Balkhande[4]

[1, 2, 3, 4]*Dept. of Computer Engineering, Mumbai University*

Abstract: Nowadays, new Artificial Intelligence (AI) and Deep Learning based processing methods are replacing traditional computer vision algorithms. On the other hand, the rise of the Internet of Things (IoT) and edge computing, has led to many research works that propose distributed video-surveillance systems based on this notion. Usually, the advanced systems process massive volumes of data in different computing facilities. Instead, this paper presents a system that incorporates AI algorithms into low-power embedded devices. The computer vision technique, which is commonly used in surveillance applications, is designed to identify, count, and monitor people's movements in the area. A distributed camera system is required for this application. The proposed AI system detects people in the monitored area using a MobileNet-SSD architecture. This algorithm can keep track of people in the surveillance providing the number of people present in the frame. The proposed framework is both privacy-aware and scalable supporting a processing pipeline on the edge consisting of person detection, tracking and robust person re-identification. The expected results show the usefulness of deploying this smart camera node throughout a distributed surveillance system.
Keywords: Edge Analytics, Person detection, Person re-identification, Deep learning, embedded systems, artificial intelligence.

I. INTRODUCTION

In recent years, deep learning has demonstrated significant advantages in a variety of domains, including finance, health care, E-commerce, automated modulation categorization in cognitive radios, and many more. Specifically, computer vision was the first domain on which deep learning was applied. As the crime rates have increased in recent years, there is a greater need to conduct accurate investigations and prevent potentially catastrophic incidents. Hence, surveillance systems are essential for public places like airports, metros, railway stations offices, restaurants, etc. This paper focuses on smart camera nodes that are distributed along a surveillance area. As software and hardware technologies have advanced, smart video systems are now capable of managing video feeds from a closed camera circuit as well as analysing and extracting information in real-time from the video streams. These embedded systems are able to detect a crowd's movement, the number of people in an area, and anomalous behaviours and can be installed in both public and private locations. In older systems, video footage was not processed automatically and was left to the operator's analysis.

The operator's analysis efficiency could be rendered as fatigue and boredom set in. An embedded system with a computer vision algorithm developed for real-time automated people detection, tracking, and counting is described in this paper.

Proposed system uses a realistic and large dataset to train and evaluate the surveillance system that has been employed to track people in the surveillance area throughout the entire day. The system is easy to set up and configure as an edge node in a distributed video monitoring system. Using an embedded platform called UpSquared2 with a VPU called Myriad-X, the proposed system allows real-time algorithms to work in parallel in reduced power consumption. The tracking and detecting of people have been designed using a MobileNet-SSD architecture. Additionally, a bank of Kalman filters allows tracking and counting people.

II. LITERATURE REVIEW

1) *Anurag M. (2003)[8]* presented a method that uses multiple synchronized cameras to track all the people in a cluttered scene while segmenting, detecting and detecting their movement at the same time. It introduces an algorithm based on region data that can be used to search for 3D points inside an object if we are aware of the regions within the object from two different viewpoints. people were constrained to move in only a small region.

2) *Koen Buys (2014)[1]* approach relies on an underlying kinematic model. This approach uses an additional iteration of the algorithm that segments the body from the background. It presents a method for RDF training, including data generation and cluster-based learning, that enables classifier retraining for different body models, kinematic models, poses or scenes.

3) *Hetal K. (2017)[13]* Tracking objects from any moving object has been successfully implemented using the Kalman filter. The system operates on indoor videos as well as outdoor environment taken using static camera as well as moving camera called PTZ camera under moderate to complex background conditions.

4) *Chandan G (2018)[12]* used an SSD and MobileNets based algorithms are for detection and tracking using python environment. CNN and deep learning is used for feature extraction. For image classification and counting, classifiers are used. Using the concepts of deep learning with a YOLO-based algorithm and a GMM model will provide good feature extraction and classification accuracy. Shallow layers in a neural network may not generate enough high level features for prediction for tiny objects. SSD is worse for smaller objects as compared to bigger objects.

III. RESEARCH GAP

The real-time object detection and tracking systems exists, wherein it is possible to detect and track people in real time but, this systems are not integrated with other technologies like AWS cloud and Edge computing.

In previously implemented surveillance systems database is not enforced. But, in our surveillance system, we are creating a database for specific campus/facility to identify people working/studying in that area.

Prior to the implementation of the proposed system, parallel processing of video streams rarely took place. In the proposed system we are using MobileNet-SSD for people detection, Kalman Filter Bank for people tracking in an embedded system. Along with this, the parallelization allows the processing of more than one video stream at the same time.

IV. OBJECTIVES

A. To build a fully functioning surveillance system that incorporates cutting-edge technology like Edge computing.

B. Incorporate AI algorithms in low-power embedded devices.

C. Reduce the manual workload of operator who is analyzing the footage's in real-time.

D. Improve awareness of security personal and decision makers by collecting real-time information.

V. PROPOSED METHODOLOGY

A. Embedded processing system

The embedded platform selected for the smart node is the UpSquared2 system, including the deep learning module of Intel Movidius Myriad X VPU, a System-On-Chip (SoC) that can be used for accelerating AI inference with a low-power footprint. To execute AI algorithms in the VPU, the OpenVino framework has been used, which eases the optimizing and deployment of CNNs. This framework has the advantage of being able to run on any device that meets the minimum requirements.

B. Image Processing To Detect And Track At The Edge Node

In the first step, referred to as data analytics, the image is pre-processed and the returned information from the AI inference engine was post-processed. Before the pre-processing of the image, different algorithms such as noise reduction, image edge detector or any image enhance low-level pixel processing could be applied.

C. MobileNet- SSD for people detection

The selected network is a MobileNet-SSD, a MobileNet architecture, which uses the SSD (Single Shot Detection) method for object detection.

The reasons for using this architecture: first being the need of a fast architecture, which could be implemented with algorithms like SSD; and secondly, the low resource consumption, because the devices used in this project are portable and the hardware is not as powerful as a high-performance devices.

At the beginning of the network there is the Mobile Net module, composed of 35 convolutional layers, responsible for the feature extraction. Tensorflow's Keras API is used for implementing the detection algorithm. After this network is trained, the Open Vino framework is used to optimize the model weights.

D. Kalman Filter Bank for tracking

The main reasons for the use of this filter in our system are: the robustness to overlapping and the low computational cost of the filter.
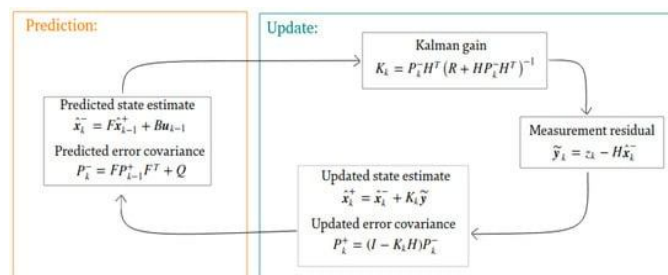
Fig 1 :Example of kalman filter

In the above diagram, the equations relate to tracking an individual using a Kalman filter. Thus, for each detection performed on the MobileNet- SSD, a Kalman filter will be tracking each object. In this tracking module, the bounding boxes of the detected objects will be received directly from the MobileNet-SSD.

E. Pipeline Operations

As part of the architecture, three major components are pre-processing, Kalman filtering, and MobileNet-SSD inference. Due to the separation of the inference processing from the CPU processing, this parallelization allows the processing of more than one video stream simultaneously. With the help of VPUs, this inference is processed. Devices like these are able to process multiple inferences simultaneously and more efficiently than CPUs.

VI. CONCLUSIONS

The proposed system is a portable video surveillance system with processing at the edge, which can detect and track people in a robust and reliable way. The system can works in real-time and can track multiple persons in the camera's field of view accurately with robust re-identification. In this paper, we described the framework of our video surveillance system and provided the algorithms and expected results of our system. Now- a- days everyone is aware of their surroundings and the potential hazards they face. Every employee is responsible for looking out for his or her own safety as well as that of their co-workers. The future work will focus on estimating the 3D trajectory of each moving object using multi-camera data fusion, analysing the multiple-person interaction, and detecting suspicious behaviours.

VII. ACKNOWLEDGMENT

REFERENCES

[1] K. Buys, C. Cagniart, A. Baksheev, T.-D. Laet, J. D. Schutter and C. Pantofaru, "An adaptable system for RGB-D based human body detection and pose estimation," Journal of visual communication and image representation, vol. 25, pp. 39-52, Jan 2014.

[2] A. Jalal, Y.-H. Kim, Y.-J. Kim, S. Kamal and D. Kim, "Robust human activity recognition from depth video using spatiotemporal multi-fused feature," Pattern recognition, vol. 61, pp. 295-308, 2017.

[3] B. Enyedi, L. Konyha and K. Fazekas, "Threshold procedures and image segmentation," in proc. of the IEEE International symposium ELMAR, pp. 119-124, 2005.

[4] A. Jalal, and S. Kamal, "Real-time life logging via a depth silhouettebased human activity recognition system for smart home services," in Proceedings of AVSS, Korea, pp. 74-80, Aug 2014.

[5] A. Sony, K. Ajith, K. Thomas, T. Thomas, and P. L. Deepa, "Video summarization by clustering using euclidean distance," in proc. of the SCCNT, 2011.

[6] A. Jalal and S. Kim, "The mechanism of edge detection using the block matching criteria for the motion estimation," in Proceedings of HCI Conference, Korea, pp. 484-489, Jan 2005

[7] L. Kaelon, P. Rosin, D. Marshall and S. Moore, "Detecting violent and abnormal crowd activity using temporal analysis of grey level cooccurrence matrix (GLCM)-based texture measures," MVA, vol. 28, no. 3, pp. 361-371, 2017

[8] Anurag Mittal and Larry S. Davis, M2Tracker: A Multi-View Approach to Segmenting and Tracking People in a Cluttered Scene. International Journal of Computer Vision. Vol. 51 (3), Feb/March 2003.

[9] J. Redmon and A. Farhad, "YOLOv3: An Incremental Improvement,"Retrieved from: https://pjreddie.com/media/files/papers/YOLOv3.pdf, 2018.

[10] H. Tao, H.S. Sawhney and R. Kumar, Dynamic Layer Representation with Applications to Tracking, Proc. of the IEEE Computer Vision & Pattern Recognition, Hilton Head, SC, 2000.

[11] S. Kamal and A. Jalal, "A hybrid feature extraction approach for human detection, tracking and activity recognition using depth sensors," Arabian Journal for science and engineering, 2016.

[12] G Chandan; Ayush Jain; Harsh Jain; Mohana Real Time Object Detection and Tracking Using Deep Learning and OpenCV,2018.

[13] Hetal K. Chavda; Maulik Dhamecha;Moving object tracking using PTZ camera in video surveillance system,2017

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 ⓢ (24*7 Support on Whatsapp)