Design and Implementation of Object Detection, Tracking, Counting and Classification Algorithms using Artificial Intelligence for Automated Video Surveillance Applications

Conference Paper · October 2022

CITATIONS
5

READS
209

1 author:

Mohana Mohana
R. V. College of Engineering
74 PUBLICATIONS **1,549** CITATIONS

SEE PROFILE

Design and Implementation of Object Detection, Tracking, Counting and Classification Algorithms using Artificial Intelligence for Automated Video Surveillance Applications

Mohana
Research Scholar, Electronics & Communication Engg.
Assistant Professor, Telecommunication Engineering
R. V. College of Engineering
Bengaluru, India
mohana@rvce.edu.in

Dr. H. V. Ravish Aradhya
Research Advisor, Professor
Electronics & Communication Engineering
R. V. College of Engineering
Bengaluru, India
ravisharadhya@rvce.edu.in

Abstract— Video Surveillance is very important and essential task for private and public organizations, sensitive areas to monitor safety, security issues and prevent crime. Especially in the field of surveillance system it has gained greater significance than ever before due to the recent terror activities taking place all over the world. Detection of objects precisely is vital for many applications such as person identification, abnormal activity detection, congestion analysis, military fields. Traditional surveillance was like capturing image or video from digital camera, store the data in a database, required information will be accessed manually from the database and it requires human operator to constantly monitor suspicious or threatening activities. Sometimes it may lead to loss of important information. In that particular cases automated video surveillance is very important. Current or automated surveillance system uses a digital technology to capture, store and process an image or video. Smart and intelligent surveillance is required to minimize the role of human operator and automatically detect, track, classify and monitor the suspicious or threatening activities. In this research work detection, tracking, counting and classification algorithms are implemented using matlab, DSP, FPGA and Artificial intelligence (AI) methods and described briefly. Feature of images is extracted using convolution neural network (CNN) using the concept of deep learning. YOLO based algorithm with GMM model by using the concepts of deep learning will give good accuracy for feature extraction and classification. AI combines SSD and Mobile Nets to perform efficient implementation of detection and tracking. This algorithm performs efficient object detection while not compromising on the performance. Further detection, tracking and classification algorithms are implemented for various applications under various constraints and environmental conditions.

Keywords— FPGA Zynq XC7Z020, DSP Embest Dev Kit 8500D, GPU, Modified Background subtraction, YOLO, YOLOv2, *Artificial intelligence, Convolution Neural Network, Mobile Nets, Single Shot Detector (SSD), COCO.*

I. INTRODUCTION

From past decade CCTV and survillance has increased its usage and popularity. Especially after several terrorist bomb attacks in different places of the world, the organizations have increased their security by strict authentication process at entrance and increasing number of surveillance cameras in all over the places. Video surveillance mostly refers observation of activities of human beings, vehicles or any other objects. Surveillance is helpful in monitoring threat or suspicious activities continuously, which maintains social control and stops criminal activities with one or more calibrated cameras. An intelligent surveillance system should be capable of detecting movement of objects in the scene, identifying the type of moving objects, tracking the moving objects in the scene, detecting the unusual events and warrant the security personnel to take preventive measures. Smart surveillance requires computing facility to capture frames from video, process images and take intelligent decision based on the logic provided in the software. It has a wide range of applications such as measuring traffic flow to crowd flux statistics, tracking people and interpreting their behavior, understanding movement of human body, detecting robbery at bank or secured places, identity checking or biometrics, suspect identification, etc. It can implement in three methods manually, semiautonomous and full autonomous. Presently used system in many applications is manual. It is less efficient and not feasible. Hence large amount of innovation and research is going on across the globe to automate the complete system. Some of the important parameters or constraints are cost, speed, efficiency of hardware and software required to achieve identified benchmark.

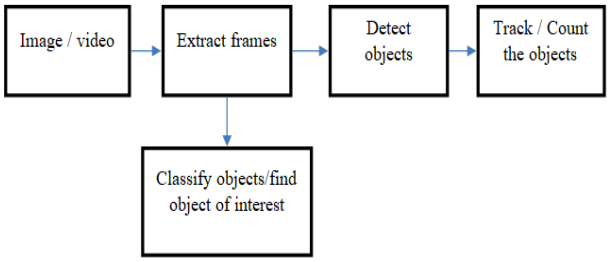

Fig.1. Block diagram of research work.

Figure 1 shows the basic flow of research work. Section II describes main motivation beyond the work. Section III gives the details of implementation methods briefly. Section IV gives the results and analysis work carried out so far, followed by Section V describes conclusion and future scope.

II. MAIN MOTIVATION

In defense, border and restricted area video surveillance is very essential to detect and track accurate position, movement information of unauthorized object. Understanding activities

of objects moving in a scene by using video is very fruitful domain with many promising applications. Main motivation is to design efficient algorithms for detection, tracking, counting and classification for various survillance applications and its performance analysis under various constraints. Classification of different types of objects, counting number of objects in a crowded environment, identifying region of interest, tracking and activity analysis, measuring the velocity, detection of suspicious or threatening behavior of objects and predict of their actions and interactions. Implementation of this algorithm to survillance applications and its performance analysis.

III. METHODS OF IMPLEMENTATION

This section explains the method of implementation such as matlab, DSP, FPGA, Artificial intelligence methods and NVIDA GPU technology.

A. Matlab implementation

The implementation was carried out in various environmental conditions with varying luminance to determine the efficiency of the system. Detection and tracking of objects are implemented using modified and adaptive background subtraction, fast RCNN, YOLO and SSD methods. Objects are detected in multiple regions either at a time or separately from single static camera; it allows using different thresholding for different regions so that objects can be detected more accurately by utilizing the properties of respective regions. It increases the speed of object detection and reduces the number of cameras and cost. To estimate the density of crowd objects are counted. Then single and multiple object tracking are implemented. To know the type of object, objects are classified into different classes [1] [4] [5].

B. DSP and FPGA implementation

Survillance applications require processing of huge amount of data. Implementation of this task serially is very difficult for real time applications because it gives a delay for all operations. And complexity also increases significantly. Digital signal processing (DSP) boards has fixed hardware structure. Means memory, peripherals and connections are fixed, predefined operations and handling in sequential manner. But in FPGA interconnections are determined by user and operations are not predefined. Here data will be processed parallely using HDL. It is implemented to optimize speed, memory required and usage of logic elements. Detection and tracking is implemented on Zynq XC7Z020 FPGA board. Camera used for implementation OV7670. It utilizes 652.8kB meory and 2884 logic elements for complete implementation. The camera captures 30 frames of video per second and it requires 0.36 seconds to processes all frames. Hence minimal delay between capture and display of video All the algorithms are implemented and verified on FPGA Zynq board as well as on a serial DSP based Embest Dev Kit 8500D further in depth analysis was carried out [2] [3].

C. Implementation using AI Methods

Objects are detected and tracked using SSD and YOLO algorithm. Objects are detected using SSD algorithm in real time scenarios. Additionally, SSD have results with considerable confidence level. Main Objective of SSD algorithm to detect various objects in real time video sequence and track them in real time. This model showed excellent detection and tracking results on the object trained and can further utilized in specific scenarios to detect, track and respond to the particular targeted objects in the video surveillance. This real time analysis of the ecosystem can yield great results by enabling security, order and utility for any enterprise.YOLOv2 model and YOLO9000, real-time detection systems for detecting and classifying objects video. YOLOv2 is agile and efficient in detecting and classifying the objects. The speed and accuracy were achieved with the aid of GPU functionalities and Anchor Box technique respectively. Furthermore, YOLOv2 can detect object movement in video with a proficient accuracy. YOLO9000 is a real-time framework which is able to optimize detection and classification and bridge the gap between them. The YOLOv2 model and YOLO 9000 detection system collectively are able to detect and classify objects varying from multiple instances of single objects to multiple instances of multiple objects. With faster run-time classification finds many applications in autonomous navigation and many such applications. Various classification algorithms are implemented using the concept of deep learning. Obtained results show that convolutional Neural Network will give 85.97 % accuracy for image classification. Further it can be used for various video survillance and security related applications for feature extraction and classification.

D. NVIDA GPU Technology

Currently working on algorithms with larger data set using NVIDA GPU hardware components to increase the accuracy and processing speed at NVIDIA Lab.

E. Sample Implementation Snapshots and results

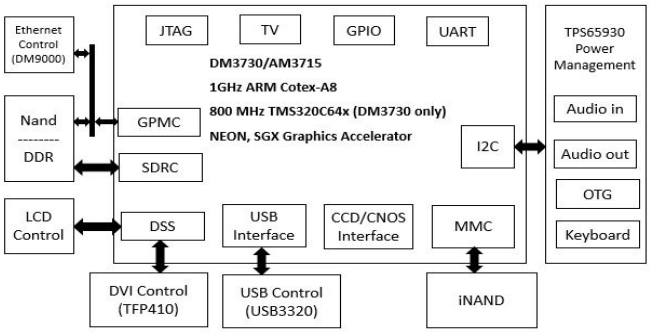

Fig.2. Implementation using DSP board Embest DEV kit 8500D.

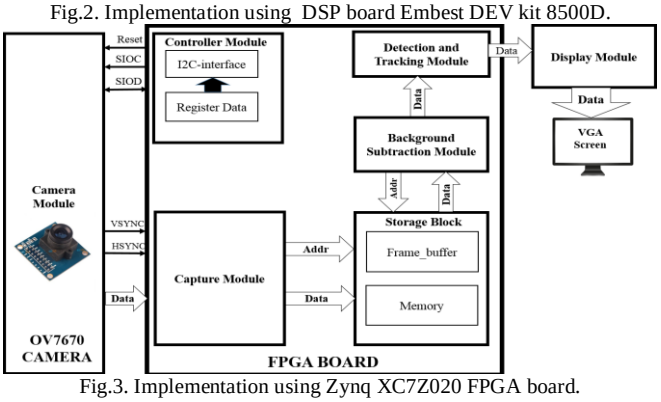

Fig.3. Implementation using Zynq XC7Z020 FPGA board.

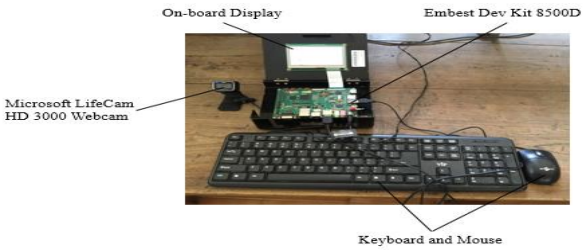

Fig.4. Experimental setup of DSP board Embest Dev kit 8500D.

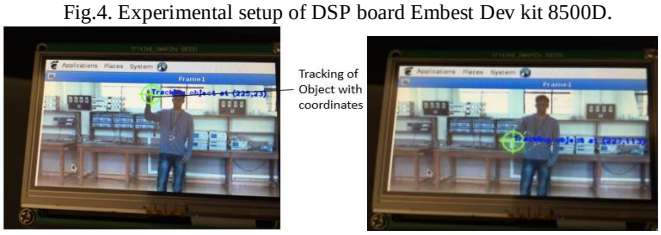

Fig.5.Object detection and tracking using Embest Dev kit 8500D.

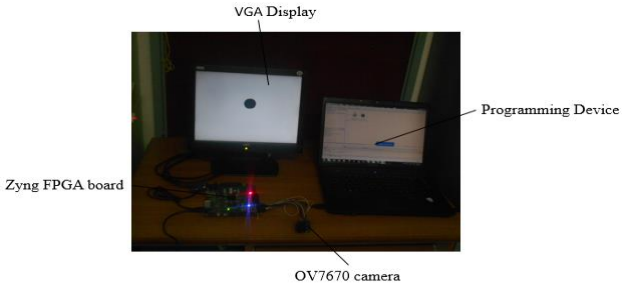

Fig.6. Experimental set up of Zynq XC7Z020 FPGA board.

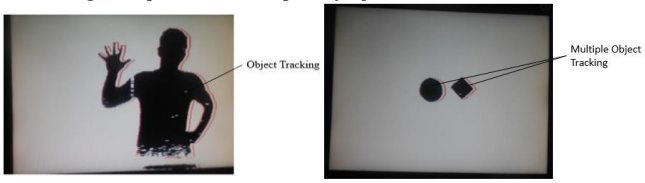

Fig.7.Object detection and tracking using Zynq XC7Z020 FPGA board.

Fig.8 Classification of objects (a) motor bikes (b) horses (using Matlab).

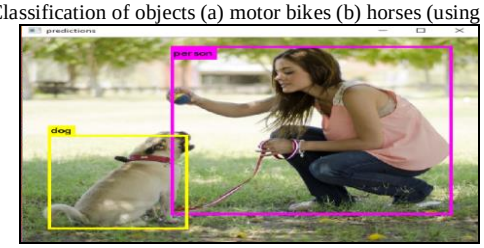

Fig. 9 Detection and labeling of objects (Using AI and Deep Learning).

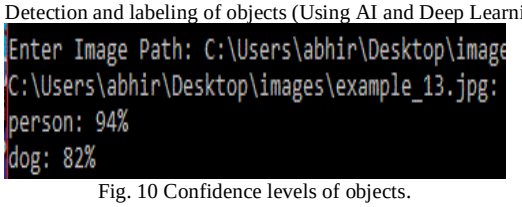

Fig. 10 Confidence levels of objects.

Fig. 11 Detection and labelling of multiple objects.

Fig. 12 Confidence levels of multiple objects.

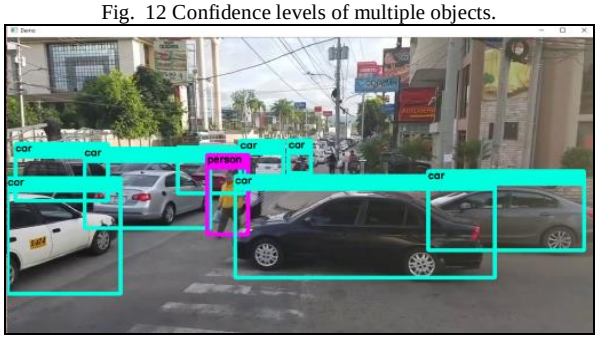

Fig. 13 Detection and Labeling of multiple instances of single object.

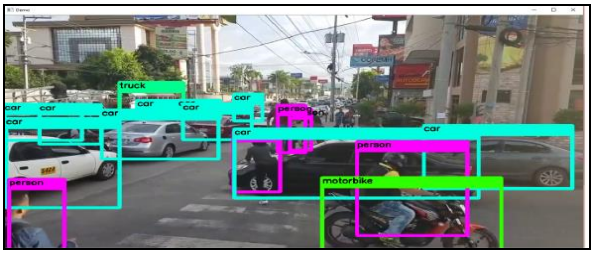

Fig. 14 Detection and Labeling of multiple instances of multiple objects.

Fig. 15 Real time objects detection and labeling of multiple classes of objects.

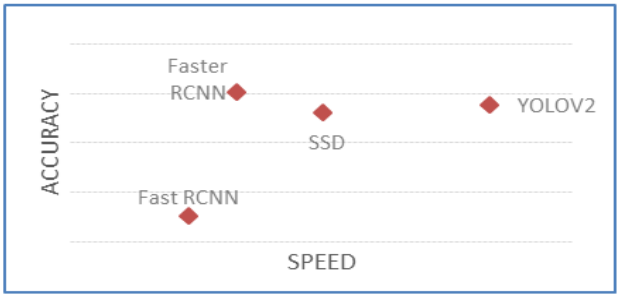

Fig. 16 Speed versus accuracy of object detection algorithms.

IV. RESULTS AND ANALYSIS

Object detection, tracking, counting, classification algorithms are simulated and implemented on matlab, FPGA and DSP kits.

TABLE. 1. DETECTION AND EFFICIENCY OF OBJECTS CLASSIFICATION (USING MATLAB)

Objects class	Number of images	Detected correct	Detected wrong	Efficiency (%)
Bus	30	25	5	83
Car	43	35	8	81.4
Motor bike	30	24	6	80
Horse	40	32	8	80

Table1 shows Detection and efficiency of objects classification (using MATLAB) for smaller data set [1].

TABLE.2. RESOURCE UTILIZATION SUMMARY (USING FPGA)

Slice Logic Utilization	Used	Available	utilization
Number of Slice Registers	286	106,400	1%
Number of Slice LUTs	343	53,200	1%
Number of IOBs	49	200	24%
Number of occupied Slices	125	13,300	1%
Number of RAMB36E1	120	140	85%
Number used as logic	326	53,200	1%
Number used as ROM	4	17,400	1%

Table 2 shows utilization of resources available in FPGA for detection and tracking.

TABLE 3: CLASSIFICATION ALGORITHM AND OBTAINED ACCURACIES

Classification Algorithm	*Accuracy in %*
KNN Classifier	28.2
SVM Classifier	37.4
SoftMax Classifier	34.1
Fully connected Neural Network	46.4
Convolutional Neural Network	*85.97*

Table 3 describes Classification algorithm and obtained accuracies for feature extraction. CNN using deep learning gives better accuracy.

A. Outcomes of research work

It will improve the security aspects in defense, military and public places. The system can be implemented in many surveillance applications such as public and commercial security, law enforcement, smart video data mining, can adopt in airports and railway stations to counter the terrorism, in highways (toll stations) to count and classify the number of vehicles or objects moving in a road, tracking and detecting accurate position of the object and movement of information of the object in military and restricted areas.

V. CONCLUSION AND FUTURE SCOPE

A. Conclusions

Started with an aim to develop a smart surveillance system for detection and tracking for various survillance applications. Currently system being employed in most of the domains is manually operated and is too cumbersome to implement it efficiently in real-time. Hence, development of an automated system, which operates without the help of human intervention and in real-time, has become inevitable to enhance its efficiency. The modified background subtraction and object detection algorithms are implemented using Xilinx ISE software in VHDL programming language. The FPGA board used is Zynq XC7Z020 and it uses OV7670 camera for capturing the video. FPGA systems are 5 to 6 times more efficient than the DSP board owing to the temporal and spatial parallelism it supports and also the flexibility to design the modules as per the requirements of the system. Further all the algorithms are implemented using deep learning and opencv and using AI methods. It gives good detection accuracy, classification efficiency with high speed to meet real time survillance applications.

B. Current work and Future work

Currently working on identifying the abnormal behaviour of objects in conjunction with detection and tracking. Improving the efficiency of classification of objects based on their contour. Once the classification is done accurately tracking of abnormal behavior of people in mass can be easily identified which gives enormous strength for the effective surveillance. Concept of partial face detection using multi key point descriptors can also be integrated to the system in order to increase the overall efficacy of the surveillance system. Extending work to detect ammunition and guns in order to trigger alarm in case of terrorist attacks. The model can be deployed in CCTVs, drones and other surveillance devices to detect attacks on many places like schools, government offices and hospitals where arms are completely restricted. Further working with larger data sets using GPU hardware components to increase the accuracy and processing speed at NVIDIA Lab. Finally integrating the system with artificial intelligence where no human intervention is required to monitor any of the surveillance process.

REFERENCES

[1] Mohana, H. V. Ravish Aradhya, "Elegant and Efficient Algorithms for Real Time Implementation of Object Detection, Classification, Tracking and Counting using FPGA Zynq XC7Z020 for Automated Video Surveillance Applications" *10th IEEE Advanced Networks and Telecommunications Systems (ANTS-2016)*, 6-9 November 2016, J.N. Tata Auditorium, Indian Institute of Science (IISc), Bangalore, India.

[2] Mohana, H. V. Ravish Aradhya, "Real-Time Implementation of object detection and Tracking on DSP for video survillance applications" *Proceedings of the 2016 IEEE International Conference on 'Recent Trends in Electronics, Information & Communication Technolog'* pp.1-5, may 18-20 2016, Bangalore, IEEE 2016.

[3] Mohana, H. V. Ravish Aradhya,"Real-Time Implementation of object detection and Tracking on FPGA for video survillance applications" *2016 IEEE International Conference on 'Distributed Computing,VLSI, Electrical Circuits and Robotics(DISCOVER)'* National Institute of Technology surathkal, Karnataka, NITK Surathkal 13-14 August 2016.

[4] Apoorva Raghunandan, Mohana, Pakala Raghav and H. V. Ravish Aradhya,"Object Detection Algorithms for video survilllance applications" *International conference on communication and signal processing (ICCSP)*, India, 2018, pp. 0570-0575.

[5] Akshay Mangawati, Mohana, Mohammed Leesan, H. V. Ravish Aradhya, "Object Tracking Algorithms for video survilllance applications" *International conference on communication and signal processing (ICCSP)*, India, 2018, pp. 0676-0680.

Cite This:

Mohana, HV Ravish Aradhya, **"Design and Implementation of Object Detection, Tracking, Counting and Classification Algorithms using Artificial Intelligence for Automated Video Surveillance Applications",** in Proceedings of ACCS 24th Annual International Conference on Advanced Computing and Communications (ADCOM 2018), pp. 119-122, 21-23 September 2018, at International Institute of Information Technology (IIIT-B), Organized by Advanced Computing and Communication Society (ACCS), India.