International Journal of Research in Advent Technology, Special Issue, ICATESM 2019
E-ISSN: 2321-9637
Available online at www.ijrat.org

Evolution of Deep Neural Network Algorithms Using Dynamic Camera and Objects on CPU AND FPGA

Modhia Urjit Yashwant[1], Monal Rakesh Patel[2], Prof. Yogesh Parmar[3], Prof. Carlos Valderrama[4]
Department of Electronics and Communication[1, 2, 3], Department of Microelectronics[4]
Parul University (India)[1, 2, 3], University of Mons (Belgium)[4]
Email: urjit58@gmail.com[1], monalrakeshpatel@gmail.com[2], yogesh.parmar8031@paruluniversity.ac.in[3], CarlosAlbertoVALDERRAMA@umons.ac.be[4]

Abstract- The moving object recognition is a standout amongst the most required research subject which is blasting among the image processing specialists. Artificial intelligence is a hypothesis that is utilized to make master framework that can perform exercises same as people. As of late, the amazing capacity with the feature learning and exchange learning includes developing interest inside the computer vision network. Detection of object in motion is the task of recognizing the physical development of object in particular area. From current years it has received much attention due to applications like video observation, human movement and traffic analysis, robot navigation and security. In this article we used techniques to distinguish moving items using dynamic camera and various Deep Learning Neural models. The abnormality score relies upon contrast in pixel intensities among the picture and the model. However, detection of genuine shape of moving object turns into a difficult task because of different difficulties like non-static scene changes, illuminated varieties, and shadow, disguise and bootstrapping issue. In order to conquer these issues, scientists have proposed number of new methodologies. The outline of this article is to use and have a look at current advancement in observation frameworks and object detection techniques. Their likelihood to get to sporadic lead and human object identification is the subject in different applications.

Keywords- Moving Object Detection, Moving Camera, Deep Learning, Background Subtraction, Optical Flow.

1. INTRODUCTION

Over the ongoing years, Moving object detection is an imperative and vital part in computer vision because of broad variety of usages like video perception, security at airport terminal, law requirement, compression of videos, identifying programmed target, surveillance under water and recognition of human actions. Surveillance frameworks have been increased more significantly because of expanding security demands. Investigating the whole place is a critical and troublesome undertaking. In this manner surveillance systems are generally utilized in numerous regions. Generally, moving item is identified by two kinds of cameras: static camera and moving camera. Static camera recognizes the object by the *background subtraction algorithm* and screens the similar area. Be that as it may, the supposition of the static camera has breaking points of the utilization of recognition algorithms because of the addition of moving camera stages, for instance, vehicles, robots and mobile phones. Moving cameras can capture broad areas in its way. The frames of a fixed camera possess same background in every frame; however it can't be the same for camera in motion frame since the purposes of the video frames will change their position on each move. Dynamic object identification is the act of sectioning non-stationary objects regarding surrounding region from the arrangement of given video frames. Assurance of the target in motion forms the essential step for classification and following procedure of object which is in motion. The fundamental point of moving object tracking action is to find frontal area moving targets either in each frame or at the first target appearance in video. Identifying the moving item winds up noteworthy step for any detection procedure. The problem discussed in this article is the recognition of several objects with the help of various deep neural network algorithms especially convolution neural networks. Object detection was previously done using only conventional deep convolution neural network whereas using regional based convolution network increases the accuracy and also decreases the time required to complete the program. The models used for detection are BNN, YOLO, and Libraries used are OpenCV and Tensorflow. The dataset used are COCO, Haarcascade, Cifar-10 and Keras. The datasets are very important in image recognition, object detection and other image processing problems. Supervised learning is also possible in implementing the problem using Decision trees but neural network works best in image processing because they can handle images in a great manner. With the end goal of evaluation, we will utilize Neural Networks which are already available with several datasets uniquely for the detection of vehicle, license plate, and face detection. We utilized the PYNQ board as hardware and we utilized few libraries improved with Python language to distinguish the objects.

Structure of paper is as follows

Section II will present the deep neural network models used for detection of moving object. Section III will focus on the introduction of Hardware. Section IV will represent the current research trends. Then Section V will present the results acquired during the evaluation and at last section VI concludes the paper. Section VII states the reference of papers I have used during the work.

2. DEEP NEURAL NETWORK

A neural network, in general, is an innovation worked to simulate the action of the human mind. Explicitly, design acknowledgment and the entry of contribution through different layers of reproduced neural associations.

Numerous specialists characterize Deep Neural Networks as systems that have an input layer, a output layer and no less than one hidden layer in the middle. Each layer performs explicit kinds of arranging and requesting in a procedure that some refer to as "Feature Hierarchy". One of the key employments of these refined neural Networks is managing unlabeled or unstructured information. The expression "Deep learning" is likewise used to describe these deep neural networks, as Deep learning represents an explicit type of machine learning where technologies utilizing parts of man-made intelligence try to characterize and arrange data in manners that go beyond basic information/yield conventions. There is couple of neural systems accessible for recognition like convolutional neural network, Binary Neural Network, Tensorflow and OpenCV.

We are evaluating these algorithms because these algorithms are of different types and used by many researchers. Binary neural network is a convolutional neural network but with binary weights and is small model and can be easily implemented on 32-bit FPGA which is our requirement. Tensorflow is a library which allows us to create and train network models. Yolo is a also a neural network model but is already trained with *coco* dataset and last but not the least OpenCV is a library which is used by many researchers in their detection work.

2.1. Binary Neural Network

According to a famous Deep learning framework, convolutional neural Networks has been commonly used in various assignments, for instance, picture grouping and object affirmation. Convolutional neural system abuses spatial connections in the photos by performing convolution assignments in nearby responsive fields. Convolutional neural systems are favored over totally related neural systems since they have fewer loads and are less requesting to plan. Many research works have been coordinated to lessen the computational multifaceted nature and memory requirements of convolutional neural system, to make it appropriate to the low-control introduced applications with limited recollections. This paper presents the convolutional neural system with parallel loads and incitation's, otherwise called BINARY NEURAL NETWORK, on a FPGA arrange. Loads and data enactments are binarized with only two qualities, +1 and -1. This diminishes all the settled point increment activities in convolutional layers and totally related layers to 1-bit XNOR errands.

2.2. Tensorflow

Tensorflow is made by the Google Brain group, It is an open source library for numerical figuring and broad scale machine learning. TensorFlow bundles together countless learning and Deep learning (known as neural Network organization) models and estimations and makes them profitable by strategy for a normal allegory. It uses python to give a favorable front-end API for building applications with the structure, while executing those applications in elite C++. TensorFlow can get ready and run Deep neural system for physically composed digit portrayal, picture acknowledgment, word embeddings, intermittent neural systems, arrangement to-grouping models for machine elucidation, characteristic dialect handling, and PDE (Partial differential condition) based diversions. Best of all, TensorFlow backings age forecast at scale, with comparative models used for getting ready.

2.3. Yolo

Yolo stands for You Only Look Once. It is a network for object detection task consisting the determination of the location on the images where some sort of objects is present, as well classifying those objects. There are some older methods for this type of task, like R-CNN which uses a pipeline to perform the task in multiple steps and this can be slow and also hard to optimize because every individual component must be trained separately. YOLO does everything with just a single neural network. So to keep it simple an image is needed as input and pass it through a neural network which looks similar to a normal CNN and the output is in form of vector of bounding boxes and class predictions in the output.

2.4. OpenCV

OpenCV (Open Source Computer Vision Library) is an open source Computer vision and machine picking up programming library. OpenCV was attempted to give a regular establishment to PC vision applications and to quicken the usage of machine acknowledgment in the business things. Being a BSD-approved item, OpenCV makes it basic for associations to utilize and change the code.

The library has more than 2500 improved estimations, which fuses a total game plan of both exemplary and bleeding edge PC vision and machine learning counts. These estimations can be used to recognize and perceive faces, distinguish objects, portray human exercises in recordings, track camera developments, track moving items, separate 3D models of articles, convey 3D point

cloud from stereo cameras, attach pictures together to make a high goals image of an entire scene, discover similar pictures from an image database, expelling red eyes from pictures taken using streak, pursue eye developments, perceive view and set up markers to overlay it with broadened reality, etc. OpenCV has more than 47 thousand people of customer organize and evaluated number of downloads outperforming 14 million. The library is used broadly in associations, ask about social occasions and by managerial bodies.

3. HARDWARE

Hardware is an important part of this paper because the evaluation is the comparison of neural network algorithms working on CPU as well as FPGA. Hence, here we have used PYNQ-Z1 as hardware to test the above mentioned algorithms because PYNQ is a 32-bit python productivity of Zynq board which is small but it provides python environment allowing the user to work with web architecture (Jupyter Notebook) along with Linux operating system.

3.1. PYNQ Board

PYNQ is an open-source venture from Xilinx that makes it simple to configuration implanted frameworks with Xilinx Zynq Systems on Chips (SoCs). Utilizing the Python language and libraries, creators have the advantages of programmable logic and chip in Zynq to manufacture more competent and energizing embedded systems. PYNQ clients would now be able to make high performance embedded applications with Parallel execution, High rate video handling, Fast Hardware calculations, Real-time flag preparing, High transfer speed between Input and Output, Low latency control.

4. CURRENT RESEARCH TRENDS

This section depicts some of continuous research approaches created for moving article identification to get execution with lessened missteps. Moving object recognition incorporates finding moving items in the casing of a video gathering. It has reliably been a testing task in the field of video taking care of results of moving object recognition strategies and is exceedingly affected by changing of background. Furthermore, contrast fit as a fiddle, development and speed of a moving target makes the assignment increasingly convoluted. Vast number of work is done to get exact results considering the referenced challenges.

As we have reviewed the papers mentioned below, we learned about the algorithms to detect Indian license plates using OpenALPRI, moving objects using moving cameras, strategy to detect missing objects, face recognition technique and optical flow based recognition system.

Indian License plates are not standardized and it is a disadvantage by considering recent situation as the system which must be prepared for numerous text styles. Convolution Neural Network approach recognizes specific text styles and is moderate in considering the sliding window method for pictures along with high goals. Sachin prabhu et al.[10]. introduced the Convolution Neural Network is a versatile and actualized productively neural network in comparison with other methodologies such as k-NN (LPRI) and OpenALPR approaches for Indian License Plates (OpenALPRI). LPRI can perceive a large portion of tag inside a picture considering still picture or the casing if there should be an occurrence of video contrasted with OpenALPRI and CNN. However, Convolution Neural Network and OpenALPRI have better acknowledgment of characters in comparison to LPRI. K-NN should be enhanced via preparation with some examples. GPU power is tackled to conquer this by preparation on server. CNN has most extreme exactness and is even more versatile, likewise it perceives tags from CCTV film with Full High Definition quality. Hence, license plate can be recognized from stationary picture or live streaming video.

Kimintun et al.[5]. Proposed an incorporated structure handling issues brought up in the identification of object utilizing dynamic cameras. The structure depends on dual-mode displaying with intentional inspecting and receives scene conditional adjustment of the model. Also, scene prior based adjustment is joined to adapt the testing issues from videos by dash camera. The proposed plan fulfills the constant execution which is essential to a real Application. Besides, the plan kept running in non supervised and online way, it was a less compelled and even less touchy compared to different techniques. Because of the ease of use of the techniques, strategies can be broadly used of numerous applications in the dynamic camera including drone, dash camera and many more.

Sincan et al.[2]. Proposed a technique for detection of moving items with dynamic camera. They discover the intrigue focuses in successive frames of video and tracking them by pyramidal Lucas-Kanade strategy. At that point they compute camera movement accepting most basic movement vectors has a place with camera movement. In the wake of disposing of camera movement, frame differentiation technique is utilized to recognize dynamic objects. They utilize versatile threshold considering distinctive conditions in a similar frame of video and tried to calculate with their datasets and inspected that the proposed technique identifies dynamic object with high precision and low fake cautions. Enhancement of the proposed strategy by getting exact object limits to classify objects accurately is done in future.

Rakumthong et al. [13].presents a strategy to detect the objects which are stolen. As begun by an initial step, set of frames procurement is used to import the video to complete the process in the subsequent stage. The next

step was procedure for the detection of the video to change ordinary picture to gray scaled picture and recognizes the dynamic object by the utilization of subtraction method with background reducing algorithm and current picture. After that there is separation between people or object by using the Haarcascade function, regularly utilized for human location, and segregate between stolen occasions by examining the boundaries of closer view areas. At last, the result will be displayed on the output screen. Although all procedure can identify and order sorts of items and occasions, the system still require productive and accuracy for more effective detection.

A face recognition system is proposed by huang et al. [6] utilizing a dynamic camera in an open area. To defeat the difficulties of multiple face identification using picture pyramid pattern, the proposed strategy confines the inquiry of face focuses on a restricted zone, which is accomplished utilizing the data of skin shading, edge, and face area estimation. Using this strategy, face competitors created, which additionally checked by the C-SVM classifier to frame confront items. The faces are then coordinated and followed utilizing the measurements of Euclidean separation. Test results demonstrate the proposed strategy can effectively identify a large portion of human countenances, showing the attainability of the proposed technique.

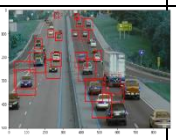
Younis et al.[4]. Presents an optical flow based recognition system which classifies the movement of a wearable camera continuously. Camera motion type (CMT) was either fixed or moving (left, right, up and down). A free-moving dynamic camera which was wearable and was used in the work as a advanced glasses modules that will be used in upcoming work by creating advanced assistive innovation to help living enduring vision misfortune which is used to keep them away from risks at the time they walk. The output of this work will be used in dynamic item identification and following the decrease of the required handling loads. Execution results were 84% accurate for CMT recognition.

5. RESULTS

The results we achieved by evaluating the above mentioned algorithms in terms of speed and accuracy we can state that among all this algorithms Yolo is the best algorithms to develop a system for the detection of moving objects using moving camera because as we know Yolo is already a pre-trained model with coco dataset available with 80 classes. It can easily implement on 32-bit FPGA as well as CPU also. We have evaluated all the algorithms on CPU with the help of anaconda navigator which provides virtual environment and various important and usable libraries such as scipy, pandas etc.

You can see the result achieved during the evaluation in Table 1. The result presented is of four types of different images and is in terms of speed per frame and accuracy of the model per frame.

Table 1 Comparison of algorithms:

Frame	OpenCV	BNN	Tensor flow	Yolo
 (a)	3.7954 Haarcascade classifier 54%	0.00035 cifar-10 classifier 71%	0.076 Keras (Cifar-100) classifier 69%	0.0012 Coco dataset 73%
 (b)	2.7875 Haarcascade classifier 62%	0.00051 cifar-10 classifier 73%	0.063 keras(Cifar-100) **68%**	0.0039 Coco dataset 70%
 (c)	2.1238 Haarcascade classifier 59%	0.00041 cifar-10 classifier 68%	0.067 keras (Cifar-100) classifier 70%	0.0025 Coco dataset 75%
 (d)	2.3514 Haarcascade Classifier 58%	0.00045 Cifar-10 classifier 71%	0.066 Keras (cifar-100) classifier 7%	0.0029 Coco datatset 71%

5.1. Difficulties faced during implementation and evaluation:

OpenCV: OpenCV is easy to implement in CPU as well as FPGA also. In CPU, OpenCV is used with the help of anaconda whereas in FPGA it is used with the help of Jupyter implemented in Python. But the only disadvantage of OpenCV is that it is quite slow in processing with comparison to other network models as shown in Table 1. The advantages of using OpenCV is that it is fully trained with many classifiers and datasets.

Binary neural Network (BNN): BNN is convolutional neural network implemented on FPGA. It is faster in comparison with OpenCV but the only disadvantage is that it is only trained with few classifiers or datasets like cifar-10, Mnist and Streetview. It is quite hard to train BNN with new datasets or classifiers.

Tensorflow: Tensorflow is the general based library specially powered with keras classifier and dataset. In

CPU, Tensorflow is evaluated by using anaconda tool. But it is hard to implement in FPGA because there is only one compatible version available to run in 32-bit FPGA i.e. 0.9.0.

Yolo: Yolo is the fully trained neural network model available with coco dataset. This dataset contains 80 classes like person, car, bus, bottle etc. this model is easy to implement in CPU using anaconda as well as is compatible to implement in FPGA also.

6. CONCLUSION

As there are numerous researchers taking a shot at object detection and following there is a requirement for proficient technique. Subsequently this paper gives survey of the researches based on moving object recognition. In numerous papers, Optical flow was used to distinguish and group the movement of mobile camera continuously. During studies it was recognized that the shadow brightening variety and non-static background were the significant issues which worked since these issues led to decrease the exactness of progressive steps of examination process. This paper proposed moving item recognition based Deep learning techniques appropriate to freely moving camera. The proposed strategy comprises of a neural system concentrating on the speed and availability of techniques committed to the FPGA. Upgrade of knowledge and the practicability of object recognition dependent on various algorithms and PYNQ Board are key point in future research. A strategy is proposed from different moving object detection and tracking system. The key elements will be Binary Neural Network, Tensorflow, Yolo, PYNQ, OpenCV and Python. These techniques can be generally used in numerous applications like moving camera observation framework, movement control, security reason, mischance cures and so forth.

7. REFRENCES

[1] "Appearance and Motion Based Deep Learning Architecture for Moving Object Detection in Moving Camera"- Byeongho Heo; Kimin Yun; Jin Young Choi, 2017 IEEE International Conference on Image processing (ICIP).

[2] "Moving object detection by a mounted moving camera"- Ozge Mercanoglu Sincan, Vahid Babaei Ajabshir, Hacer Yalim Keles, Suleyman Tosun; IEEE EUROCON 2015- International Conference on Computer as a Tool (EUROCON).

[3] "A review of object detection based on convolutional neural network"-Wang Zhiqiang, Liu Jun: 2017 36t Chinese Control Conference (CCC).

[4] "Real-Time Detection of Wearable Camera Motion Using Optical Flow"- Ola Younis, Waleed Al-Nuaimy, Fiona Rowe, Mohammad H. Alomari; 2018 IEEE Congress on Evolutionary Computation, At Rio de Janeiro, BRAZIL.

[5] "Situation-aware Framework From Moving Object Detection in Moving Camera"-Kimin Tun Jongin Lim, Jin young Choi; 8th International Conference of Pattern Recognition Systems (ICPRS 2017).

[6] "Real-Time Face Detection Using a Moving Camera"- Deng-Yuan Huang, Chao-Ho Chen, Tsong Yi Chen, Jian-He Wu, Chien-Chuan Ko; 2018.

[7] "Character Feature based Indian Vehicle License Plate Detection and recognition"- Sudhir K. Ingole, Shital B. Gundre; 2017 International Conference on Intelligent Computing and Control (I2C2).

[8] "Moving Object Detection and Tracking From Moving Camera"-Won Jin Kim, In-So Kweon; 2011, 8th International Conference on Ubiquitous Robots and Ambient Intelligent (URAI).

[9] "Object-Level Motion Detection from Moving Cameras"-Tao Cheng Shijan Lu; IEEE Transactions on Circuits and Systems for Video Technology, Nov. 2017.

[10] "Recognition of Indian License Plate Number From Live stream Videos"- Sachin Prabhu B. Subramanium Kalambur, Dinkar Sitaram; 2017 International Conference on Advance in Computing Communications and Informatics (ICACCI).

[11] "IOT-Driven Automated Object Detection Algotrithm For Urban Surveillence Systems in Smart Cities"-Ling Hu Qiang Ni; IEEE Internet of Things Journal, April 2018.

[12] "A Framework for abandoned Object Detection From Video Surveillance"- Rajesh Kumar Tripathi; Anand Singh Jalal; Charul Bhatnagar; 2013 Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG).

[13] "Unattended and Stolen Object Detection based on Relocating of existing object"- Waritchana.

[14] "Moving Object detection using Background Subtraction and Motion Depth Detection in Depth Image Sequences"- Jichan Lee; Sungsoo Lim; Jun-Geon Kim; Bomin Kim; Daeho Lee; The 18thIEEE International Symposium on Consumer Electronics (ISCE).

[15] "Moving Object Detection with a Freely Moving Camera via Background Motion Subtraction"- Yuanyuan Wu; Xiaohai He; Truong Q. Nguyen; IEEE Transactions on Circuits and Systems for Video Technology.