Identifying Moving Objects Using Background Subtraction, Optical Flow Method, Deep Neural Network for Camera in Motion Using PYNQ Architecture: Review of Recent Research Trends

[1]Patel Monal Rakeshbhai, [2]Arvind Yadav, [3]Carlos Valderrama
[1]PhD Student, [2]Professor, [3] Chief of Department
[1]Dpt. of Electronics and Communication,
[1]Parul University, Vadodara, India

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 with a similar knowledge that of people. As of late, the amazing capacity with the feature learning and exchange learning Binary Neural Network (BNN) includes developing interest inside the computer vision network. Moving object detection is the assignment of recognizing the physical development of a object in a given area. From most recent couple of years it has gotten much attraction because of applications like video observation, human motion analysis, robot navigation, event detection, traffic analysis and security. In this article we reviewed techniques to distinguish moving items dependent on optical flow and Background subtraction. The abnormality score is figured dependent on Background subtraction and relies upon contrast in pixel intensities between the current picture and the Background model. However, undertaking of detecting genuine shape of object in motion turns into a precarious because of different difficulties like dynamic scene changes, illumination varieties, presence of shadow, disguise and bootstrapping issue. To conquer these issues, scientists have proposed number of new methodologies. The outline of this article is to have a look at current advancement in observation frameworks and object detection. Their likelihood to get to irregular conduct and human object identification is the subject in different applications.

Index Terms **- Moving Object Detection, Moving Camera, Deep Learning, Background Subtraction, And Optical Flow.**

I. INTRODUCTION

Over the ongoing years, Moving object detection is an imperative and vital part in the field of computer vision because of its extensive variety of utilizations like video observation, checking of security at air terminal, law requirement, video compression, programmed target identification, marine surveillance and human action recognition. Surveillance frameworks have increased incredible significance 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 Camera: Static camera and Moving camera. Static camera recognizes the object by the Background subtraction Algorithm and screens the similar area. However, the supposition of the stationary camera has limits of the use of detection algorithm on account of increment of moving camera stages, for example, vehicles, robots and cell phones. Moving camera can screen bigger regions in its moving way. The video frames of a stationary camera will have a similar background, however it can't be the equivalent for moving camera frame since every one of the purposes of the video frames will change their position on each move. A video is a gathering of fundamental structural units, for example, scene, shot and frame related with sound information.

A frame is characterized as a single picture shot of film camera, driven by numerous progressive edges for consistent video. Moving object identification is the act of sectioning non-stationary objects of interest regarding surrounding area or region from a given arrangement of video frames. Assurance of the moving target forms the essential step for classification and following procedure of object in motion. The fundamental point of moving object detection and tracking action is to find frontal area moving targets either in each video frame or at first appearance of moving target in video. In any video examination movement there are three noteworthy stages: identification of the moving target (object), tracing of identified moving item in a given arrangement of video edges and investigation of the moving target (object) with the end goal to decide its behavior. Thus, identifying the moving item winds up noteworthy step for any investigation procedure. A few techniques have been proposed so forward for object detection, out of which Background Subtraction, Frame differencing, and Optical Flow are broadly utilized conventional methods. Along these lines, in this paper we review the strategy for the moving object detection utilizing moving camera. We will identify the moving article like vehicle and people by background subtraction algorithm and optical flow strategy. With the end goal to identify and recognize the moving item, we will utilize Binary Neural Network which will be exceptionally prepared with custom datasets uniquely for the detection of vehicle, license plate, and face detection. We will utilize PYNQ board for hardware purpose and we will utilize a few libraries improved with Python language to distinguish the objects. Structure of remaining paper is as follows. Section II classification of approaches of moving object detection. Section III focuses on deep neural network. Section IV focuses on the hardware introduction. Then Section V is based on recent research trends in moving object detection VI is results and last section VII concludes the paper.

II. CLASSIFICATION OF APPROCHES OF MOVING OBJECT

Traditional methodologies for moving item recognition can be extensively sorted into two structures as Background Subtraction, Optical Flow. The figure underneath demonstrates a short grouping of traditional methodologies.

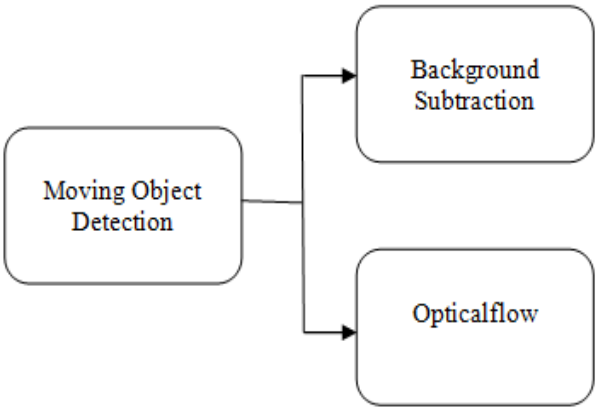


Figure 1: Grouping of traditional methodologies

Background Subtraction:

Background Subtraction Method is viewed as a standout amongst the most dependable technique for moving object identification. Background subtraction works by introducing a Background Model, at that point distinction between current frame and assumed Background Model is acquired by contrasting every pixel of the current frame with expected Background Model shading map. In the event that contrast between colors is more than threshold, pixel is viewed as having a place with closer view Performance of Traditional Background subtraction technique for the most part gets influenced when Background is dynamic, illumination changes or in presence of shadow. Various strategies have been created so forward to upgrade Background subtraction technique and overcome its disadvantages.

Optical flow:

Optical Flow approach of moving target identification depends on calculation of optical flow field of picture (or video frame).Clustering is performed based on the acquired optical flow circulation data got from the picture (video frame).This technique permits getting complete information about the movement of the object and is valuable to decide moving target from the Background. However, this strategy experiences some of downsides like extensive amount of figuring's required to acquire optical stream data and it is affectability to noise.

III. 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 so forth.

Binary Neural Network:

As a famous Deep learning system, convolutional neural Networks has been generally utilized in numerous tasks, for example, picture classification and object acknowledgment. Convolutional neural network misuses spatial relationships in the pictures by performing convolution tasks in local responsive fields. Convolutional neural networks are preferred over completely associated neural networks since they have less weight and are less demanding to prepare. Many research works have been directed to reduce the computational complexity and memory prerequisites of convolutional neural network, to make it pertinent to the low-power installed applications with restricted memories. This paper introduces the architectural structure of convolutional neural network with parallel weights and actuations, also known as BINARY NEURAL NETWORK, on a FPGA stage. Weights and information activations are binarized with just two values, +1 and - 1. This reduces all the settled point increase operations in convolutional layers and completely associated layers to 1-bit XNOR tasks.

Tensorflow*:*

Tensorflow is made by the Google Brain team, It is an open source library for numerical calculation and extensive-scale machine learning. TensorFlow packages together a large number of machine learning and Deep learning (known as neural Network administration) models and calculations and makes them valuable by method for a typical metaphor. It utilizes Python to give an advantageous front-end API for building applications with the structure, while executing those applications in high

performance C++.TensorFlow can prepare and run Deep neural network for manually written digit characterization, picture recognition, word embeddings, recurrent neural networks, sequence-to-sequence models for machine interpretation, natural language processing, and PDE (Partial differential equation) based recreations. The best part is that TensorFlow supports generation prediction at scale, with similar models utilized for preparing.

OpenCV:

OpenCV (Open Source Computer Vision Library) is an open source Computer vision and machine learning programming library. OpenCV was worked to give a typical foundation to computer vision applications and to accelerate the utilization of machine recognition in the business items. Being a BSD-authorized product, OpenCV makes it simple for organizations to use and change the code.

The library has in excess of 2500 optimized calculations, which incorporates a complete arrangement of both classic and cutting edge computer vision and machine learning calculations. These calculations can be utilized to distinguish and recognize faces, identify objects, characterize human activities in videos, track camera movements, track moving objects, separate 3D models of articles, deliver 3D point cloud from stereo cameras, fasten pictures together to create a high resolution picture of a whole scene, find comparative pictures from a picture database, removing red eyes from pictures taken utilizing flash, follow eye movements, recognize view and set up markers to overlay it with enlarged reality, and so forth. OpenCV has in excess of 47 thousand individuals of client network and assessed number of downloads surpassing 14 million. The library is utilized widely in organizations; inquire about gatherings and by administrative bodies.

IV. HARDWARE

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 equipment execution
- high outline rate video handling
- hardware quickened calculations
- real-time flag preparing
- high transfer speed IO
- low idleness control high bandwidth IO
- low latency control

Figure 2: PYNQ Board

V. RECENT RESEARCH TRENDS

Moving object recognition includes finding moving object in the frame of a video grouping. It has dependably been a challenging task in field of video handling since consequences of moving object recognition techniques are exceedingly influenced by varieties in illuminating and Background changes. Additionally, difference in shape, movement and speed of moving target makes the task more complicated. Large number of work is done to get precise outcomes thinking about the referenced difficulties. This segment portrays some of ongoing research approaches produced for moving object detection to acquire higher execution with diminished mistakes.

Sachin prabhu et al.[10]. introduced the CNN approach is versatile and can be actualized productively compared with other two methodologies i.e k-NN (LPRI) and OpenALPR for Indian License Plates i.e Open ALPRI. 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 other Open ALPRI and CNN. However, CNN and Open ALPRI has better acknowledgment of characters compared to LPRI. K-NN approach should be enhanced via preparing with various example character sets. Indian License plates are not standardized and this is a disadvantage considering present situation as the system should be prepared for numerous text styles. CNN approach recognizes specific text styles and is moderate considering the sliding window approach for picture with high goals. GPU power can be tackled to conquer this as preparing is done on server. CNN approach has most extreme exactness and is more versatile, likewise it perceives tags from CCTV film with Full HD quality. Hence, license plate can be recognized from still picture or live streaming video.

Kimin tun et al. [5]. Proposed an incorporated structure handling issues brought up in the moving object identification utilizing moving cameras. The structure depends on double 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 dash cam

videos. The proposed plan fulfills the constant execution which is essential to a real application for a pre-processing. Besides, since plan kept running in an unsupervised and online way, it was a less compelled and less touchy than different techniques. Because of the ease of use of the techniques, strategies can be broadly used of numerous applications in the moving camera including drone, dash cam and so forth.

Sincan et al. [2]. Proposed a technique for detecting moving items from a moving camera. They discover the intrigue focuses in successive video frames and track them with 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 difference technique is utilized to recognize moving objects. they utilize versatile thresholding considering distinctive lighting conditions in a similar video and tried to calculate with their Golbasi dataset and Hopkins dataset and saw that the proposed technique identifies moving object with high precision and low false cautions. Enhancing the proposed strategy by getting exact object limits to classify objects accurately is the work to be done in future.

Rakumthong et al. [13] presents a strategy for the unattended and stolen object detection is exhibited. As begun by an initial step, video procurement is utilized to import the video to process in the subsequent stage. The second step was the video detection procedure to change ordinary picture to grayscale picture and recognize the moving object by utilizing subtraction method with background and current picture. After that separate between people or object by utilizing the Haarcascade function, regularly utilized for people location, and segregate between unattended or stolen occasions by examining the boundaries of static closer view areas. At last, the result of detection will alarm on the output screen. Although all procedure can identify and order sorts of objects and occasions, the system still needs productive and accuracy for more reliable detection.

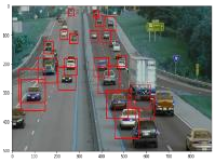
A real-time face detection system is proposed by huang et al. [6]. Utilizing a moving camera in an open space. To defeat the challenge of multi-scale face identification utilizing picture pyramid, the proposed strategy confines the inquiry space of face focuses in a restricted zone, which can be accomplished utilizing the data of skin shading, edge, and face area estimation. By this strategy, face competitors are created, which are additionally checked by the two-class C-SVM classifier with the HOG highlights to frame confront targets. The face targets 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 systems which classify the movement of a wearable camera continuously. Camera motion type (CMT) was named either stationary, moving left, moving right, moving up and moving down. A wearable, free-moving monocular camera was utilized in this work as one of a few brilliant glasses modules that will be utilized in future work in creating smart assistive innovation to help people enduring vision misfortune to stay away from risks while walking. The output of this will be utilized in moving object identification and following the decrease of the required handling loads. Promising execution results of 84% right states for CMT recognition were acquired.

VI. RESULTS

The results we achieved by evaluating the above mentioned algorithms in terms of speed and accuracy. 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	Tensorflow
 (a)	3.7954 Haar cascade classifier 54%	0.00035 cifar-10 classifier 71%	0.076 Keras (Cifar-100) classifier 69%
 (b)	2.7875 Haar cascade classifier 62%	0.00051 cifar-10 classifier 73%	0.063 keras(Cifar-100) **68%**
 (c)	2.1238 Haar cascade classifier 59%	0.00041 cifar-10 classifier 68%	0.067 keras (Cifar-100) classifier 70%

 (d)	2.3514 Haar cascade Classifier 58%	0.00045 Cifar-10 classifier 71%	0.066 Keras (cifar-100) classifier 7%

VII. 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 shadow brightening variety and dynamic background are the significant issues which are worked over since these issues lead to decrease in exactness of progressive steps of examination process. this paper proposed the Deep learning based moving item location system appropriate to freely moving camera. The proposed strategy comprises of a neural system concentrating on the appearance and additionally organize committed to the movement. The proposed deep learning approach accomplishes a strong execution against background contamination even with free movement of camera. Upgrade of knowledge and the practicability of object recognition dependent on BNN and PYNQ Board are enter point in future research. A strategy is proposed from different moving object detection and tracking system. The key elements will be: Background subtraction, Optical Flow, Binary Neural Network, PYNQ, OpenCV and Python. Our techniques can be generally used in numerous applications like moving camera observation framework, movement control, security reason, mischance cures and so forth.

REFERENCES

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 36th 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 mocing 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 Intelligence (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"- Sachi Prabhu B., Subramanium Kalambur, Dinkar Sitaram; 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI).

11. "IOT-Driven Automated Object Detection Algorithm For Urban Surveillance 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 Rakumthong ; Natpaphat Phetcharaladakun ; Wichuda Wealveerakup; Nawat Kamnoonwatana; 2014 Third ICT International Student Project Conference (ICT-ISPC).

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 18th IEEE 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.