IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-7, Issue-5, May 2019 E-ISSN: 2347-2693

Automation of Security System Using Machine Learning

Sam Sebastian[1*], Dipin S Nair[2], Aiswarya B Nair[3], Jeswin Elza Varghese[4], Sithu Ubaid[5]

[1,2,3,4,5]Dept. Of computer science and engineering, St. Thomas College Of Engineering And Technology, Chengannur, India

[1]*Corresponding Author: samsebastian1996@gmail.com Tel.: +919633825210*

DOI: https://doi.org/10.26438/ijcse/v7i5.17561761 | Available online at: www.ijcseonline.org

Accepted: 20/May/2019, Published: 31/May/2019

Abstract— This project mainly focuses on controlling and securing our home or office using machine learning technology. In this system we are implementing a system to monitor our home or office using a security camera system. Most thefts are done in the night and most of the times thieves cover their face during the theft time. So even though the home and office have those security camera systems, we are unable to identify the thief. In this project we use the machine learning technology in the security camera. It identifies each and every object in the visuals by real time. Whenever it identifies a person in the night, the system will send message to the concerned admin and nearby police station. Alarm will be produced as it automatically sends signal to our automation system. This system has the ability to identify the person, animals and all other objects. So system won't work when it identify an animal in the night.

Keywords—Machine Learning technology, Security Camera, Identify

I. INTRODUCTION

Human brain can easily understand the visual frames which are obtained by eyes, but computers cannot do it without any special methodology. If we compare human and computer, camera module is the eye of the computer and the computer vision is the way that how brain understand the image. So, we use the Machine Learning Technology to make our camera more intelligent. Computer vision is a field that includes some methods, processing and analyzing techniques for understanding images also known as image analysis, scene analysis and image understanding. Object tracking process can be simply defined in three steps with an order and these steps are controlled by the computer autonomously.

1. Detection of interested object in video stream.
2. Tracking target object in every single frame.
3. Analysing the tracking process to recognize the behaviour.

Objects can be detected using machine learning technology. The detection of stationary or moving targets, and tracking them on a real-time video streams is a very important and challenging task in order to protect fields from enemies. The enemies can be a human, an animal or even an object. Object detection compromises located objects in the frame of a video sequence. Every tracking method requires a detection method in every single frame. Object tracking is the process of detecting one or more objects using a camera. Background subtraction is the most common detection method used from simple object trackers. It is based on comparing two successive frames. Mean shift method classifier follows this. OpenCV libraries provide us these methods and have some RGB color detection algorithms. .

Rest of the paper is organized as follows, Section I contains the introduction of Object Detection, Section II contain the related work of detecting human interference using various methods, Section III contain the some methods of detecting human from a frame, the architecture and essential steps of SSD network we used to detect human motion, Section IV contain results and discussion of our project, Section V contain the conclusion and future scope.
.

II. RELATED WORK

In this section, the literature available within the preview of the objectives of the present study is revised and the need for the proposed work is discussed. Existing system was merely supported frames or we will say objects. Simple approach was utilized in existing system like capturing photos or frames with CCTV camera. After capturing frame it will calculate the difference between captured frames. Then it will calculate the threshold value by applying some algorithmic standards and it will detect the objects based on the motion of that object.

This research project aims at the design and construction of a Short Message. Service (SMS)-based intruder detection system. This system consists of controller, receiver and

device circuit. This system not like the standard magnetic switch alarms equipped on doors and windows has incorporated motion sensors so a brief message service, SMS is sent to the house owner on any attempt of a break. This project is constructed employing a programmed microcontroller interfaced with SIM 900 module, motion detectors and switches. The Passive Infrared (PIR) sensing element that is that the motion detector utilized in this project is placed at the roof of the example and a switch on the point of the door, so that when an intruder pass through the PIR or press the door switch a message is displayed on the liquid crystal display and SMS is sent to the phone numbers embedded within the C language program use to program the microcontroller [1].

A wireless security system wherever associate warning device is programmed in an exceedingly graphical program (GUI). The system is used to monitor the RFID reader, RFID tag and the GSM terminal. The information obtained from the tag is distributed to the server during a RF link that's exhibited. If the laptop is stolen from the covered region, the alarm system will start to draw attention. Meanwhile, the portable computer owner is notified by open source alert message. In addition, the alarm won't be stopped till the portable computer is restore within the lined region, or the program is stopped/terminated. RFID have been available for many years for reading bar codes RFID tag located several meters away. It is progressively getting used in alternative applications starting from inventory management to anti-counterfeiting protection [2].

Surveillance is very useful to governments and law enforcement to maintain social control, recognize and monitor threats, and prevent/investigate criminal activity. In this paper the human object is captured, a data base about the pixel values is trained to the system. Video Camera is fixed at a required place where security is needed. Whenever human movement is captured by the camera it is immediately detected and the object is tracked by background subtraction method, finally the system is processed to make the alarm to produce sound. This method detects moving objects in required environments under changing illumination conditions and in the presence of background dynamics. Also present a fast implementation of the method using an extension of integral images.

Background subtraction is a widely used for detecting moving objects. The ultimate goal is to "subtract" the background pixels in a scene leaving only the foreground objects of interest. A non-recursive technique uses a sliding window approach for background estimation. Non-recursive is highly adaptive as they do not depend on the history beyond those frames stored in the buffer. Some of the commonly used non-recursive techniques are: frame differencing, median filter, mean filter. Here in this work we use frame differencing method for background modelling. Frame Differencing Fundamental logic for detecting moving objects from the difference between the current frame and a reference frame is called background image subtraction and this method is called as frame differencing.

The object is detected from the live video and tracked using background subtraction, this system is proposed for real-time security purpose. In the live video 18 frames are processed at a unit time. Based on the camera's range the monitoring area may be increased. Object detection is made efficiently using the background subtraction technique and the frames processed per second are improved. For this purpose MATLAB 7.14(R2012A) tool is used [3].

This paper presents a new eHealth platform incorporating humanoid robots to support an emerging multidimensional care approach for the treatment of diabetes. The architecture of the platform extends the Internet of Things (IoT) to a web-centric paradigm through utilizing existing web standards to access and control objects of the physical layer. This incorporates capillary networks, each of which encompasses a set of medical sensors linked wirelessly to a humanoid robot (via the Internet) to a web-centric disease management hub (DMH). This provides a set of services for both patients and their caregivers that support the full continuum of the multidimensional care approach of diabetes. The platform's software architecture pattern enables the development of various applications without knowing low-level details of the platform. This is achieved through unifying the access interface and mechanism of handling service requests through a layered approach based on object virtualization and automatic service delivery [4].

III. METHODOLOGY

In the method of securing our system we take the video from the CCTV camera. The recorded video is analysed and each and every object in the visuals is identified by real time and whenever it identifies a person, the system will send message to the concerned admin and nearby police station. The alarm will be produced as it automatically sends signals to the automation system. The system has the ability to identify objects other than the human and will produce no action against objects other than human.

SSD is designed for object detection in real-time. Faster R-CNN uses a region proposal network to create boundary boxes and utilizes those boxes to classify objects. While it is considered the start-of-the-art in accuracy, the whole process runs at 7 frames per second. Far below what a real-time processing needs.

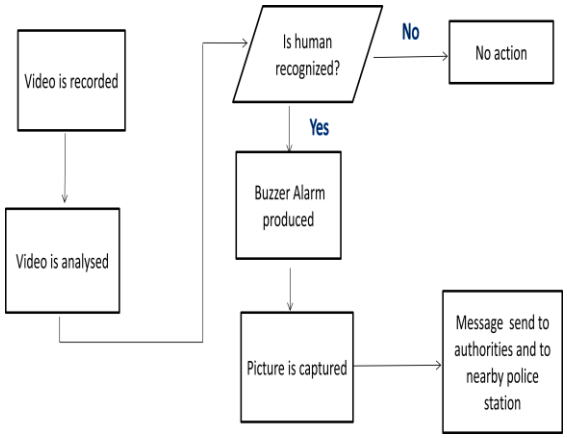

Figure 1. Overview

SSD speeds up the process by eliminating the need of the region proposal network. SSD applies a few improvements that include multi-scale features and default boxes to recover the drop in accuracy.

There are some key phases in this algorithm

- Object detection and its relation to classification
- Sliding window detection
- Reducing redundant calculations of Sliding Window Method
- Training Methodology for modified network
- Dealing with Scale of the object

Figure 2. Input Image

Here the proposals are extracted using some other computer vision technique and then resized to fixed input for the classification network, which acts as a feature extractor. SVM is trained to classify between object and background. A bounding box is trained that outputs some correction for proposal boxes.

Figure 3. Object Classification

After the classification network is trained, it can then be used to carry out detection on a new image in a sliding window manner. First, we take a window of a certain size (blue box) and run it over the image at various locations. Then crop the patches contained in the boxes and resize them to the input size of classification convolutional network. To obtain labels of the object, we feed these labels to the network. We repeat this process with smaller window size in order to be able to capture objects of smaller size. So the idea is that if there is an object present in an image, we would have a window that properly encompasses the object and produce label corresponding to that object.

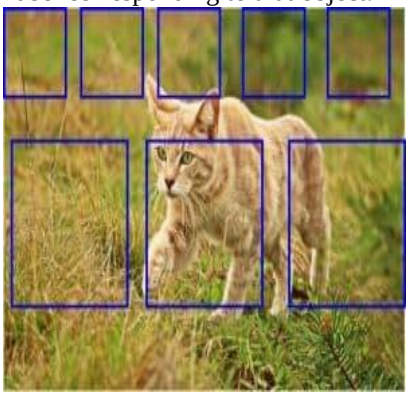

Figure 4. Sliding Window Detection

Here there is a lot of overlap between these two patches. This means that when they are fed separately into the network, the same set of calculations for the overlapped part is repeated. This can easily be avoided using a technique which was introduced in SPP-Net and made popular by Fast R-CNN. Let's take an example network to understand this in details. Calculating convolutional feature map is computationally very expensive and calculating it for each patch will take very long time. The main aim of this scheme is that we can avoid re-calculations of common parts between different patches. Here we are calculating the feature map only once for the entire image. And then since we know the parts on penultimate feature map which are mapped to different patches of image, we directly apply prediction weights (classification layer) on top of it. It is like performing sliding window on convolutional feature map instead of performing it on the input image**.** So this saves a lot of computation. The boxes which are directly represented at the classification outputs are called default boxes or anchor boxes. In the above example, boxes at center (6,6) and (8,6) are default boxes and their default size is 12X12.

In this method we took an example of a bigger input image, an image of 24X24 containing the cat. It is first passed through the convolutional layers similar to above example and produces an output feature map of size 6×6. During training SSD matches ground truth with anchors. Each

element of the feature map has a number of anchors associated with it. Any anchor with an IoUie. jaccard distance greater than 0.5 is considered a match.

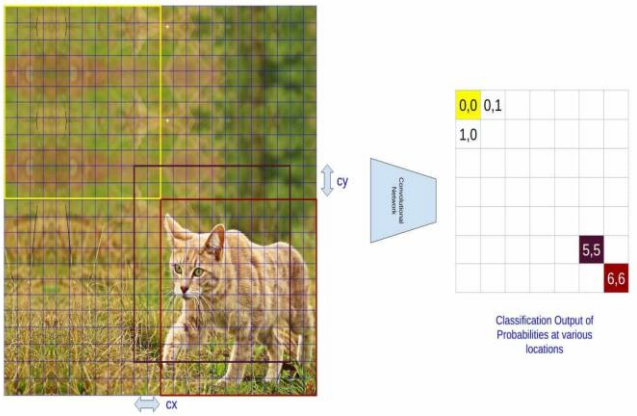

Figure 5. Training Image

For preparing training set, first of all, we need to assign the ground truth for all the predictions in classification output. Here we index the location at output map of 7,7 grid by (i,j). We already know the default boxes corresponding to each of these outputs. For reference, output and its corresponding patch are color marked in the figure for the top left and bottom right patch. Now since patch corresponding to output (6,6) has a cat in it, so ground truth becomes [1 0 0]. Since the patches at locations (0,0), (0,1), (1,0) etc do not have any object in it, their ground truth assignment is [0 0 1].
The patches for other outputs only partially contain the cat**.**

We can see that the object is slightly shifted from the box. The box does not exactly encompass the cat, but there is a decent amount of overlap. So for its assignment, we have two options: Either tag this patch as one belonging to the background or tag this as a cat. Tagging this as background will necessarily mean only one box which exactly encompasses the object will be tagged as an object. And all the other boxes will be tagged bg. This has two problems. Firstly the training will be highly skewed. Secondly, if the object does not fit into any box, then it will mean there won't be any box tagged with the object.

So we resort to the second solution of tagging this patch as a cat. But in this solution, we need to take care of the offset in center of this box from the object center. Let's say in our example, cx and cy is the offset in center of the patch from the center of the object along x and y-direction respectively. We need to devise a way such that for this patch, the network can also predict these offsets which can thus be used to find true coordinates of an object. So for every location, we add two more outputs to the network (apart from class probabilities) that stands for the offsets in the center**.** Let's call the predictions made by the network as ox and oy. And in order to make these outputs predict cx and cy, we can use a regression loss. Vanilla squared error loss can be used for this type of regression.

Now we have taken care of objects at different locations, let's see how the changes in the scale of an object can be tackled. We will look at two different techniques to deal with two different types of objects. One type refers to the object whose size is somewhere near to 12X12 pixels (default size of the boxes). The other type refers to the objects whose size is significantly different from 12X12.

Objects with size close to 12X12
For the objects similar in size to 12X12, we can deal them in a manner similar to the offset predictions. Let us assume that true height and width of the object is h and w respectively. So we add two more dimensions to the output signifying height and width(oh, ow). Then we again use regression to make these outputs predict the true height and width.

Objects far smaller than 12X12
Dealing with objects very different from 12X12 size is a little trickier. For the sake of argument, let us assume that we only want to deal with objects which are far smaller than the default size. The one line solution to this is to make predictions on top of every feature map (output after each convolutional layer) of the network as shown in figure. The prediction layers have been shown as branches from the base network in figure. This is the important idea introduced in Single Shot Multibox Detector**.**

Feature vector is the main component we use for identifying each object in our image frame. Feature engineering is a crucial step and basic step in the machine-learning pipeline.. Python packages including numpy, Pandas, Scikit-learn, and Matplotlib are used in code examples. The MHI feature vectors are obtained from human action video clips. The 2-D Haar wavelet transform is employed to extract spatial information within the MHI, along with that the temporal information was extracted by computing the histogram of the MHI. Then the two feature vectors are combined to produce a lower dimensional and discriminative feature vector. Finally, the linear SVM is used for the classification process.

An MQTT system consists of clients that communicate with a server, often called a"broker". A client can either be a publisher of information or a subscriber. Each client has the ability to connect to the broker. Information is organized in a hierarchy of topics. When a publisher sends a control message with the data to the connected broker it means that it has a new item of data to distribute. The broker then publishes the information to those clients that have subscribed to that topic. When publishing, the client first connects to the broker and it will set up a default message to

be sent to subscribers if the broker detects that the publishing client has unexpectedly disconnected from the broker. Clients only interact with a broker, but a system may contain several broker servers that exchange data based on their current subscribers' topics.

Once a connection is established, we can start to publish messages. To do this we use the publish method. The publish method accepts 4 parameters. The parameters are shown below with their default values. ***publish(topic,payload=None,qos=0,retain=False)***. The only parameters you must supply are the topic, and the payload**.** The payload is the message you want to publish. To subscribe to a topic use the subscribe method of the Paho MQTT Class object. The subscribe method accepts 2 parameters – A topic or topics and a QOS (quality of Service) as shown below with their default values. ***subscribe(topic,qos=0).***

The main feature of MQTT is that it decouples the publisher and subscriber spatially. To publish or receive messages, it is mandatory that publishers and subscribers need to know the hostname/IP and port of the broker. Another important feature is that MQTT decouples by time. Even though most MQTT use cases deliver messages in near-real time, if required, the broker has the ability to store messages for clients that are not online. MQTT works asynchronously is another salient feature of the same. As most of the client libraries work asynchronously and are based on call backs or a similar model, tasks are not blocked while waiting for a message or publishing a message. Another important thing to be mentioned is that MQTT is especially easy to use on the client-side. As most pub/sub systems have the logic on the broker-side and MQTT is really the essence of pub/sub when using a client library and that makes it a light-weight protocol for small and constrained devices.

The alert module includes producing buzzer alarm and sending messages in the form of text and electronic mail to the concerned admin. For sending electronic mails to the concerned admin we use SMTP protocol. SMTP is the Short form of S*imple Mail Transfer Protocol.* It is a protocol for sending messages between servers. Most commonly used system that sends mail over the Internet use SMTP to send messages from one server to another. POP or IMAP is used to retrieve messages with an e-mail client. After the SMTP client establishes a reliable channel to the SMTP server, the session is opened with a greeting by the server. Alarm monitoring can be defined as a quick and detailed communication between our security system and the central station of security provider. The control panel registers an emergency event and sends a signal to the central monitoring station, where the appropriate authorities are notified and sent to your home. Here after the object detection is carried out and if any human motion is detected in the restricted area then soon after that the signal is automatically passed to the security system and buzzer alarm produces which alarms the security officer present there.

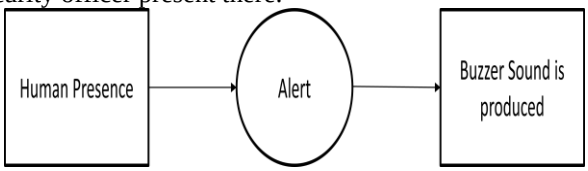

Figure 6. Alert Module

An alarm system is the most important method that alerts us when our property is unattended or invaded by thieves. Installing these alarm system is a smart and easy way to protect our homes and offices from burglars. There are many advantages of using it like constant protection, strong deterrent, uninterrupted functionality, can relocate. As a coin has two sides there are some disadvantages of installing alarm monitoring systems like getting false alarm, expensive or can be stolen. Under these circumstances our project plays an important role and vanishes these disadvantages. Using machine learning we detect human interference in the restricted area, so if human is not identified the no buzzer alarm is produced, so false alarm will not occur in our system. The implementation of this system is very cheap as it do not use any extra devices, by using the existing security cameras, we detect humans. As these cameras are placed in unnoticing areas, there are fewer chances of them to be stolen.

.

IV. RESULTS AND DISCUSSION

Testing can be defined as a process of analysing a software item to detect the differences between the existing and required condition and to evaluate the features of the software item. Here we implement our system by dividing our entire project into three different modules and each module undergoes unit testing. After this, all the modules are integrated and integration testing is carried out. After integration testing is done validation testing is performed to check the validity of our system. As the final stage of testing, the output testing is performed to check whether our system works as per the requirements.

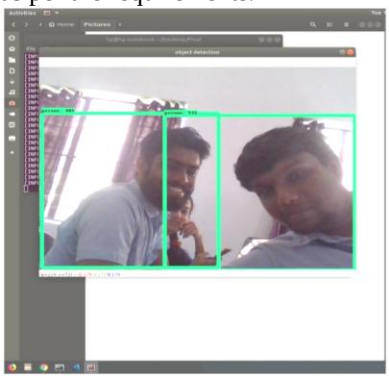

Figure 6. Image of Person

V. CONCLUSION AND FUTURE SCOPE

Object detection compromises located objects in the frame of a video sequence. Every tracking method requires a detection method in every single frame. Object tracking is the process of following one or more objects that found on detection process using a camera. Background subtraction is the most common detection method used from simple object trackers. It is based on comparing two successive frames. OpenCV libraries provide us these methods and have some RGB color detection algorithms. The prime motivation for developing this project is that, earlier methodologies only give alert but do not send the sms to the authorized person. The interesting part of this project is that when signal is passed to automation system it produce buzzer alarm, then the message is send to nearby police stations and authorized person. This method can be used in banks, offices and homes for providing better security.

The future enhancements can be identifying a burglar wearing different types of masks to cover his face. We can also include a module which has an interface with the IOT i.e. Internet of Things. This may be helpful to automatically switch on the lights and capture the image of the burglar. This image can be sent along with message to the authorities to enhance the security features. Using artificial Intelligence method this project can be implemented in higher platforms like defence, shopping, security systems and so on.

REFERENCES

[1] Bejammin Danielsson, "Real-time object detection and identification Using machine learning, training a model from a pretrained checkpoint".

[2] Human Motion Detection and Tracking for Real-Time Security System Pavithra ,Mahanthesh, Stafford Michahial1, Dr. M Shivakumar Asst. Professor, Department of IT1 Prof & Head, Department of ITI. International Journal of Advanced Research in Computer and Communication Engineering ISO 3297:2007 Certified Vol. 5, Issue 12, December 2016.

[3] Nwalozie G. C, Aniedu A. N, Nwokoye C. S, Abazuonu I.E "SMS Based Intruder Detection System" IJCSMC, Vol. 4, Issue. 6, June 2015, pg.1177 – 1184.

[4] Jorge Garcia, Alfredo Garden, Ignacio Bravo, Jose Luis Lazar and Miguel Martinez, (2013)."Tracking People Motion Based Extended Condensation .Algorithm, "IEEE Trans.Syst.,Man,Cybern. A, Syst., vol. 43, No. 3, May.

[5] Y.-L. Hou and G. K. H. Pang, (2011)."People counting and human detection in a challenging situation," IEEE Trans. Syst., Man, Cybern. A, Syst., Humans,vol. 41, no. 1, pp. 24-33, Jan.

[6] X. Zhao, E. Delleandrea, and L. Chen, (2009)."A people counting system based on face detection and tracking in a video," in Proc. IEEE Int. Conf. AVSS, pp. 67-72.

[7] Xiaowei Zhou, Can Yang, and Weichuan Yu, (2013). "Moving Object Detection by Detecting Contiguous Outliers in the LowRank Representation" IEEE Trans.

[8] Shi-qi Xing, Yong-zhen Li, Da-hai Dai, XueSongWang,(2012)."Three-Dimensional Reconstruction of ManMade Objects Using Polarimetric Tomographic SAR", IEEE Trans. Vol. 51, Issue: 6 Journal on Geoscience and Remote Sensing.

[9] Liyuan Li, Weimin HuangIreneYu-Hua Gu, and. QiTian, (2004). "Statistical Modelling of Complex Backgroundsfor Foreground Object Detection" IEEE Transon Image Processing, Vol. 13, NO. 11

[10] Weiming Hu, Xue Zhou, Wei Li, WenhanLuo, Xiaoqin Zhang, and Stephen Maybank,(2013)."Active Contour-Based VisualTracking by Integrating Colors, Shapes ,and Motions Image Processing", IEEETrans.onImageProcessing,Vol.22

[11] .Hsu-Yung Cheng, Chih-Chia Weng, Yi-Ying Chen, (2012). "Vehicle Detection in Aerial Surveillance Using Dynamic Bayesian Networks", IEEE Trans on Image Processing, Vol.21, Issue:4.

[12] Pushe Zhao, Hongbo Zhu, He Li, and Tadashi Shibata, (2013). "A Directional-Edge-Based Real-Time Object Tracking System Employing Multiple Candidate-Location Generation" IEEE Trans.vol. 23,no. 3, Mar.

[13] Wang. H and Suter.D, (2006). "A Novel Robust Statistical Method for Background Initialization and Visual Surveillance" Proc. Asian Conf. Computer Vision.

Authors Profile

Sam Sebastian currently the final year Computer Science and Engineering student of St. Thomas College Of engineering and technology, Chengannur Kerala, India. His areas of interests are Debugging, System Programming, Network Security and Robotics.

Dipin S. Nair currently the final year Computer Science and Engineering student of St. Thomas College Of engineering and technology, Chengannur Kerala, India. His areas of interests are System Programming, Debugging.

Aiswarya B. Nair currently the final year Computer Science and Engineering student of St. Thomas College Of engineering and technology, Chengannur Kerala, India. Her areas of interests include Networking, Image Processing.

Jeswin Elza Varghese currently the final year Computer Science and Engineering student of St. Thomas College Of engineering and technology, Chengannur Kerala, India. Her area of interests includes Networking, Image Processing and Artificial Intelligence.

Sithu Ubaid obtained her Bachelor of Technology (B.Tech) from Kannur University in Computer Science and Engineering and Master of Technology in Image Processing from Cochin University of Science and Technology. She is currently working as an Assistant Professor in the department of Computer Science and Engineering at St. Thomas College of engineering and technology, Chengannur, Kerala, India. Her areas of interests include Image processing, Artificial Intelligence.

