A Mini Project Report

ON

Smart Surveillance Using OpenCV

Submitted in partial fulfillment for the Degree of B. Tech

in

Artificial Intelligence

by

SHRI AKSHITA. G (19911A3547)

ALLAPURAM PRIYADHARSHINI (19911A3502)

Under the guidance of

Ms. K. Nirosha

Assistant Professor

Department of Artificial Intelligence

DEPARTMENT OF ARTIFICIAL INTELLIGENCE

VIDYA JYOTHI INSTITUTE OF TECHNOLOGY

(An Autonomous Institution)

[Accredited by NAAC & NBA, Approved by AICTE New Delhi & Permanently Affiliated to JNTUH]

2022 – 2023

VIDYA JYOTHI INSTITUTE OF TECHNOLOGY
AN AUTONOMOUS INSTITUTION

[Accredited by NAAC & NBA, Approved by AICTE New Delhi & Permanently Affiliated to JNTUH]

DEPARTMENT OF ARTIFICIAL INTELLIGENCE

CERTIFICATE

This is to certify that the project report entitled **SMART SURVEILLANCE USING OPENCV** submitted by **SHRI AKSHITA.G (19911A3547), ALLAPURAM PRIYADHARSHINI (19911A3502)** to Vidya Jyothi Institute of Technology, Hyderabad, in partial fulfillment for the award of the degree of **B. Tech in Artificial Intelligence** a bonafide record of project work carried out by us under my supervision. The contents of this report, in full or in parts, have not been submitted to any other Institution or University for the award of any degree.

Ms. K. NIROSHA
Supervisor
Department of Artificial Intelligence

Dr. A. OBULESH
HOD
Department of Artificial Intelligence

DECLARATION

We declare that this project report titled **SMART SURVEILLANCE USING OPENCV** submitted in partial fulfillment of the degree of **B. Tech in Artificial Intelligence** is a record of original work carried out by under the supervision of **Ms. K. Nirosha**, and has not formed the basis for the award of any other degree or diploma, in this or any other Institution or University. In keeping with the ethical practice in reporting scientific information, due acknowledgements have been made wherever the findings of others have been cited.

SHRI AKSHITA. G (19911A3547)

ALLAPURAM PRIYADHARSHINI (19911A3502)

ACKNOWLEDGEMENT

Any endeavor in the field of development is a person's intensive activity. A successful project is a fruitful culmination of efforts by many people, some directly involved and some others who have quietly encouraged and supported.

I would like to thank **Ms. K. Nirosha, Assistant Professor,** Department of Artificial Intelligence who has patiently guided and helped me throughout my project.

I wish to express profound gratitude to **Dr A. Obulesh**, Professor and Head of the Department, Artificial Intelligence, for his continuous encouragement to ensure the successful results in all my endeavors.

I wish to express my heart full thanks to the Principal **Dr. A. Padmaja** Madam, whose support was indispensable to me during the course.

I wish to express my heart full thanks to our Beloved Director **Dr E. Sai Baba Reddy**, whose support was indispensable to me during the course.

I wish to express my heart full thanks to our **College Management,** for providing all the infrastructure and facilities during the course.

Last but not least, I thank my family members and my friends who are the backbone and provided support in every possible way.

By

SHRI AKSHITA.G (19911A3547)

ALLAPURAM PRIYADHARSHINI (19911A3502)

ABSTRACT

Safety, and security are major concerns in the modern day. People and organizations can employ security mechanisms to safeguard their property be it home or company. Present security systems involve the utilization of assorted sensors for cameras for video surveillance. This project aims at providing one such idea to make sure the protection and security of one's property. This technique performs face recognition as an authentication procedure when an unknown face is detected by a snapshot of the unknown face. during this paper, we propose to supply a sensible CCTV closed-circuit television with intrusion detection using the LBPH-Local binary pattern histogram algorithm, SSIM-Structural Similarity Index Measure, Haar Cascade Classifier, Tkinter. By the utilization of intrusion detection, CCTV cameras not only record the real-time videos and process the video at the time of recording to search out the unwanted persons arriving within our surveillance area.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The development of technology associated with the processing using a computer is already growing rapidly, whereas image processing is using digital technology. As technology advances, graphic images are fully utilized to market the welfare of mankind. together with the event of creativity, the image cannot be separated from digital image processing. Most of the CCTV footage is employed for buffering purposes. If any issues/events occur CCTV Footage is used to determine the criminal. CCTVs will only capture the events/incidents but don't seem to be capable of recognizing a criminal. Now almost altogether places, CCTVs are used. they're only used for capturing live video occurring under its premises. In most cases, an individual is deployed to watch the video streaming. And in another case, video captured by CCTVs is recorded and it'll be checked given that any incident had occurred.

1.1 Benefits of Smart Surveillance

There was a time when surveillance techniques were utilized only in shopping centers and malls. Nowadays, you can notice the electric circuit televisions in almost any place you visit, from tiny low stores to homes and holy places. As a result, they guarantee greater security at fraction of the price.

1.2 Why Smart Surveillance?

In our busy lives, we do not have much time to observe and to stay a watch on everything. From every family, most of the members are working, or maybe in malls and hospitals so observing each and each space isn't possible. It’s the 21st century and we must think smartly to form our life better, easier, and more secure, so rather than sitting without delay place for an extended why not carry the protection in our pocket

CHAPTER 2

LITERATURE SURVEY

[1] In people verification, we have applied a hybrid approach of skin detection and body parts detection to intensify the accuracy level and reduce the error rate. During skin detection, the mask of the original image has been created and applied to the object detected to convert them into an RGB image. This RGB image is further converted to the YCbCr domain and pixel values are used to detect skin pixels in each object. For part detection; the head, hands, and feet of each silhouette are extracted, the distance between these parts is identified, and accustomed to verify the object as a human. People counting are implemented using the centroid of each silhouette that's extracted at each frame and moving objects are identified among consecutive frames. In body tracking, centroid and jacquard similarity are accustomed track people's movements between frames.

[2] Deep convolutional neural networks (DNNs) are used to detect objects within the regions. In the proposed implementation, the CNN algorithm is used to predict class labels and detect objects Methodology of implementation, consists of data Collection: The chosen dataset is a vehicle detection dataset for urban roads. it is important because the quality and quantity of data will directly determine how well the trained model will behave under various environmental conditions.

[3] In this paper, we developed a current robust smart television to trace multiple people's overcrowding counting and normal/ abnormal event detection scenarios via a monocular camera. An accurate people detection and verification model is utilized to detect people. These detected people are further processed to come back up with head templates. Template matching using normalized cross correlation with the assistance of the Jaccard similarity measure is utilized to trace and count the people. Finally, detected people are accustomed generate Gaussian clusters to research the gang for abnormal event detections.

[4] The latest and most effective technology used against robbery and theft is video surveillance and a monitoring system. But the worth of installation and maintenance of these systems become impossible for some people to afford. Arduino Mega 2560 could also be a microcontroller board that has the potential to become a camera system when its mini-module called Arducam gets attached thereto. Instead of implementing various complex algorithms, a replacement and effective methodology has been developed for motion detection which is to use Pyroelectric infrared (PIR) sensors. Whenever the motion is detected by the PIR sensor, whenever motion is detected by PIR sensor, it captures the image with the help of an attached Arducam and sends a push message to Android using the Firebase API.

[5] The purpose of the paper is to define a safekeeping alert device spending little handling power by the Internet of things which assists to observer plus alerts when gestures instead motion are there and then sends images to a cloud server. Besides, internet of things centered use is often used continuously to look at the action further and acquire a warning when gestures as an alternate indication are there. the images are shown straight to a cloud attendant, when the cloud attendant isn't accessible at that time the records are put in storage nearby on a Raspberry Pi. A MasterCard size Raspberry Pi with the advantage of Open-Source Computer Vision (Open-CV) software knobs the image processing, control algorithms used for the attentiveness then shows taken images to concerned person email by the utilization of Wi-Fi module. The system uses an ordinary webcam.

CHAPTER 3

METHODOLOGY

3.1 Structural Similarity Index Metric (SSIM)

Structural Similarity Index Metric (SSIM) for monitoring This feature is employed to seek out what's the thing which is stolen from the frame which is visible to the webcam. This means It constantly monitors the frames and checks which object or thing from the frame has been bumped off by the thief. This uses Structural Similarity to seek out the differences between the two frames. the 2 frames are captured first when noise wasn't happening and second when the noise stopped happening within the frame. SSIM is employed as a metric to live the similarity between two given images. As this system has been around since 2004, lots of fabric exists explaining the idea behind SSIM but only a few resources pass into the main points, that too specifically for a gradient-based implementation as SSIM is commonly used as a loss function.

SSIM extracts 3 key features from an image:

- Luminance
- Contrast
- Structure

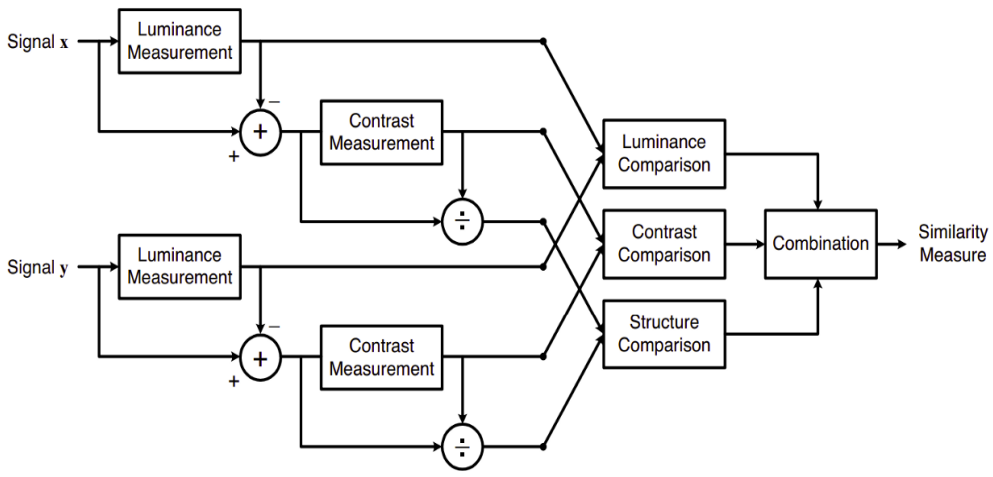


Fig. 3.1 The Structural Similarity Index Metric

The comparison between the 2 images is performed on the premise of those 3 features. This system calculates the Structural Similarity Index between 2 given images which could be a value between -1 and +1. a worth of +1 indicates that the two given images are very similar or identical while the worth of -1 indicates the two given images are very different. Often these values are adjusted to be within the range [0, 1], where the extremes hold the identical meaning

Luminance: It is measured by averaging over all the pixel values.

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i.$$

Its denoted by μ (Mu) and also the formula is given below

$$\sigma_x = \left(\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_x)^2\right)^{\frac{1}{2}}$$

Structure: The structural comparison is completed by employing a consolidated formula (more thereon later) but in essence, we divide the signaling with its variance so that the result has unit variance which allows for a more robust comparison.

$$(\mathbf{x} - \mu_x)/\sigma_x$$

Luckily, due to the image package in python, we shouldn't replicate all these mathematical calculations in python since the image has pre-built a feature that does all of those tasks for us by just calling its in-built function. We just must enclose two images/frames which we've captured earlier, so we just feed them in and it gives us out the masked image with the score.

3.2 Detecting faces in the frames

This is done through Haarcascade classifiers which are again built-in in the OpenCV module of python.

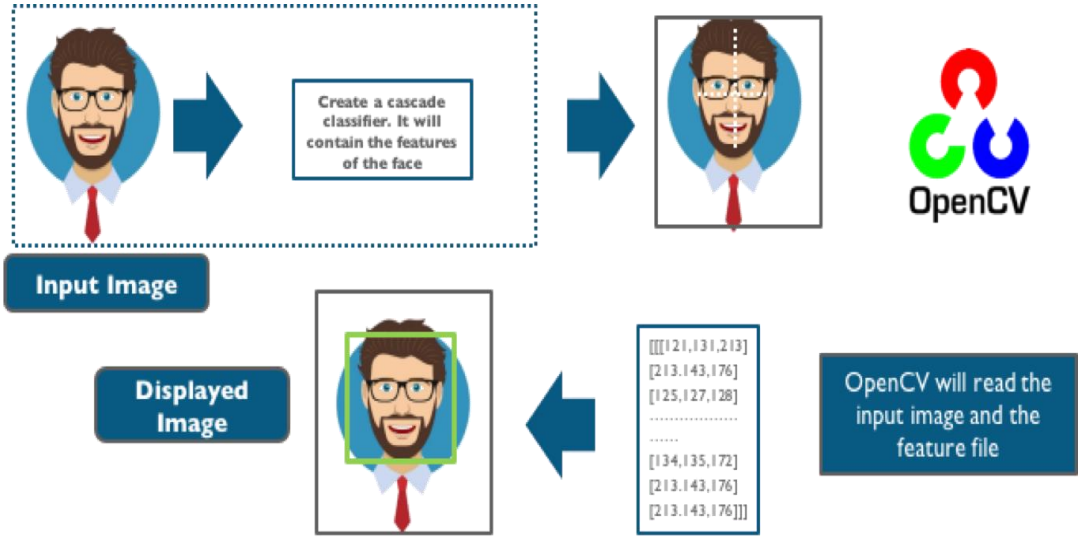


Fig. 3.2 Detecting faces in frame

Cascade classifier working with haar-like features, is a special case ensemble learning called boosting. It typically relies on Adaboost classifiers Cascade classifiers are trained on a few hundred sample images that contain the object we want to detect and other images that do not contain those images.

There are some common features:
a dark eye area compared to upper cheeks.

- a bright nose bridge region compared to the eyes.
- some specific locations of the eyes, mouth, and nose.

The feature extraction process will look like this feature are similar to this convolution kernel which is used to detect the presence of the feature in the given image. For doing all the stuff in the OpenCV module in python language has a built-in function called cascade classifier which we have used to detect faces in this frame.

3.3 Detect Noises in the frame

This feature is used to find the noises in the frames well this is something you would find in most of the CCTVs but in this module, we'll see how it works.

Talking in a simple way all the frames are continuously analyzed and checked for noises. Noise checked in the consecutive frames. We simply make a complete distinction between two frames, in this way the distinction between the two images is analyzed, and the edges (boundaries of motion are determined) and if there are no boundaries then no motion, and if there are any there is motion.

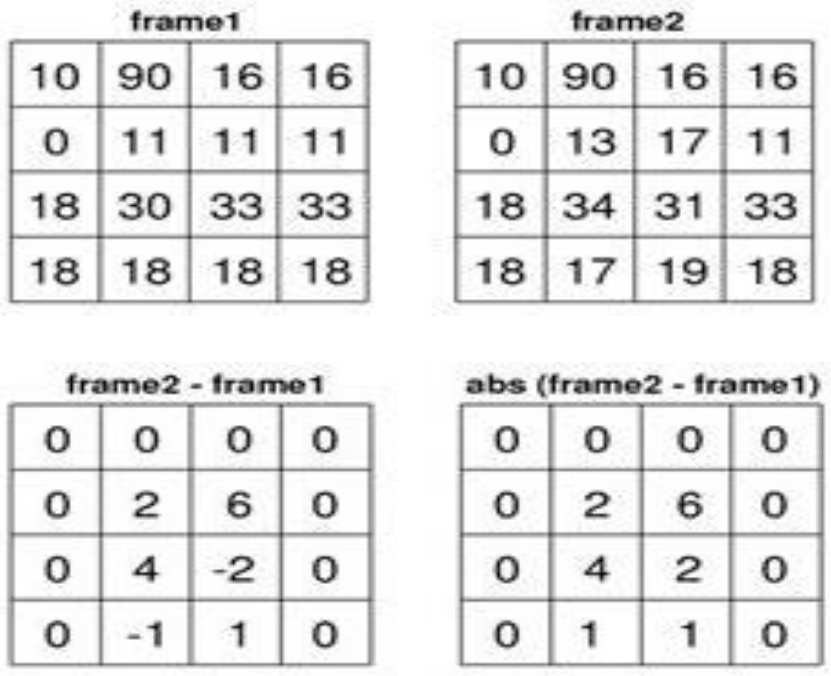


Fig. 3.3 Detecting noises in the frame

As you would know all images are just integer/ float values of pixels which tell the brightness of pixel and similarly every pixel has values of brightness. So, we are simply making a complete difference because the negative makes no sense.

3.4 Local Binary Pattern Histogram (LBPH)

So now we've detected faces within the frame and this can be the time to spot them and check if it's within the dataset which was accustomed to training our LBPH model.

The LBPH uses 4 parameters:

Radius: the radius is employed to make the circular local binary pattern and represents the radius around the central

Pixel. it's usually set to 1.

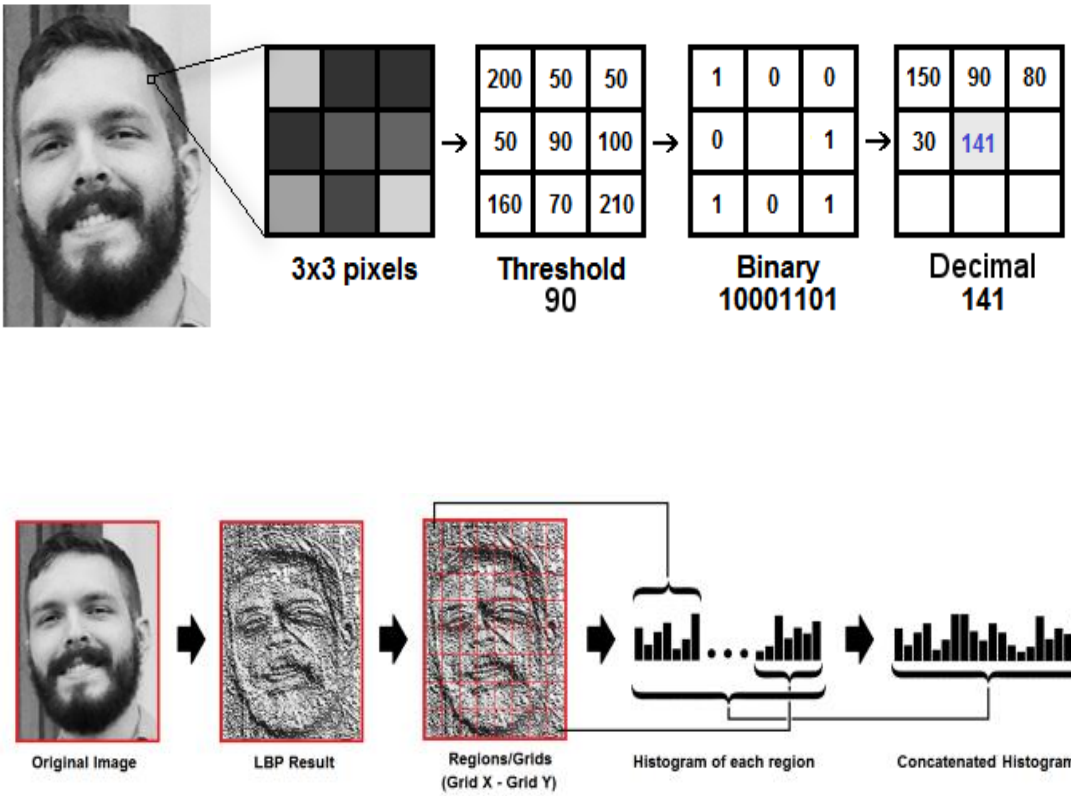


Fig. 3.4 Local Binary Pattern Histogram

Neighbors: the number of sample points to make the circular local binary pattern. detain mind: the more sample points you include, the upper the computational cost. it's usually set to eight.

Grid X: the number of cells within the horizontal direction. The more cells, the finer the grid, the upper the dimensionality of the resulting feature vector. it's usually set to eight.

Grid Y: the number of cells within the vertical direction. The more cells, the finer the grid, and the upper the dimensionality of the resulting feature vector. It is usually set to eight. The first step in calculating LBPH is to create an intermediate image that best describes the initial image, by

highlighting the facial characteristics. To do so, the algorithm uses an idea of a window, supported by the parameter's radius and neighbors. This is shown perfectly via the above

image.

Extracting the Histograms: Now, using the image generated within the last step, we can use the Grid X and Grid Y parameters to divide the image into multiple

grids. And on balance, the model is trained, and later on, when we would like to create predictions, the identical steps are applied to the make, and its histograms are compared with the already trained model in such a way this feature works.

3.5 Technologies

3.5.1 Python

Python is an interpreted, object-oriented, high-level programming language with dynamic semantics. Its high-level built-in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components together. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages, which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms, and can be freely distributed.

3.5.2 OpenCV

OpenCV is the huge open-source library for the computer vision, machine learning, and image processing and now it plays a major role in real-time operation which is very important in today’s systems. By using it, one can process images and videos to identify objects, faces, or even handwriting of a human. When it integrated with various libraries, such as NumPy, python is capable of processing the OpenCV array structure for analysis. To Identify image pattern and its various features we use vector space and perform mathematical operations on these features.

3.5.3 Tkinter

Tkinter is the standard GUI library for Python. Python when combined with Tkinter provides a fast and easy way to create GUI applications. Tkinter provides a powerful object-oriented interface to the Tk GUI toolkit.

Creating a GUI application using Tkinter is an easy task. All you need to do is perform the following steps −

- Import the *Tkinter* module.
- Create the GUI application main window.
- Add one or more of the above-mentioned widgets to the GUI application.

- Enter the main event loop to take action against each event triggered by the user.

3.5.4 Skimage

Scikit-image is an image processing Python package that works with NumPy arrays which is a collection of algorithms for image processing. Let's discuss how to deal with images into set of information and it's some application in the real world.

Important features of scikit-image:

- Simple and efficient tools for image processing and computer vision techniques.
- Accessible to everybody and reusable in various contexts.
- Built on the top of NumPy, SciPy, and matplotlib.
- Open source, commercially usable – BSD license.

CHAPTER 4

RESULTS AND DISCUSSIONS

GUI (GRAPHICAL USER INTERFACE)

Our **GUI** has different buttons supported with features as given below

- **Monitor** – allows to detect what thing is stolen from frame
- **Identify** – Finds the family members (it has to be trained first)
- **Action** – Finds the motions in the frame
- **In Out**- Finds who entered and left from the room.
- **Record**- It records the activities with date and time and gets saved in the folder
- **Rectangle –** It records the specific selected area and detects if there is any motion or not.

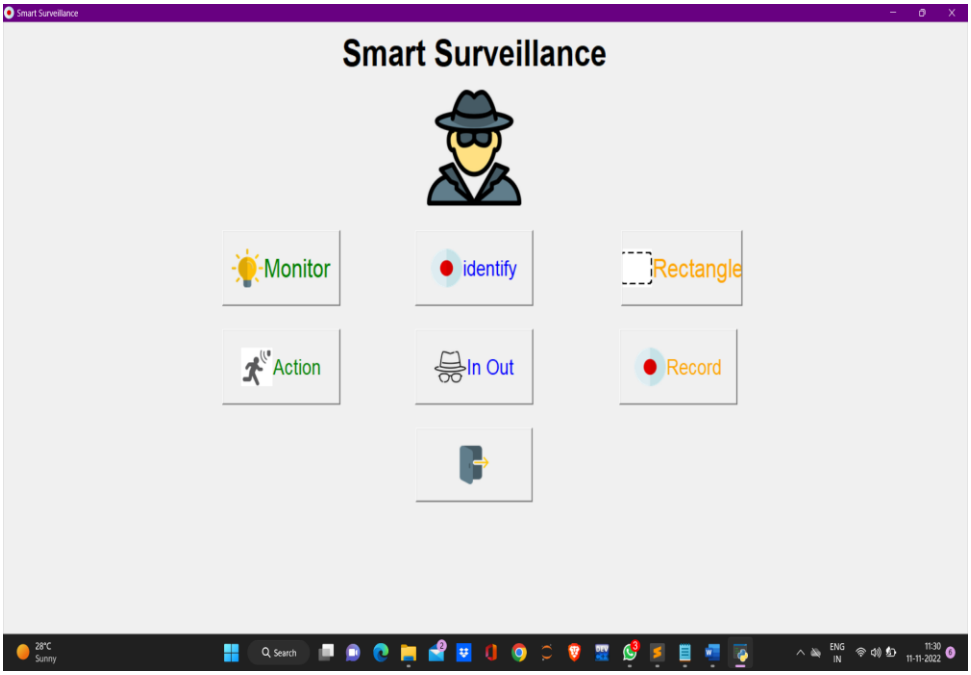


Figure 4.0 GUI

4.1 FEATURE 1 – MONITOR

Showing use of first feature or you can consider it as output for feature 1. as you can see that it is detecting bottle is stolen which is true

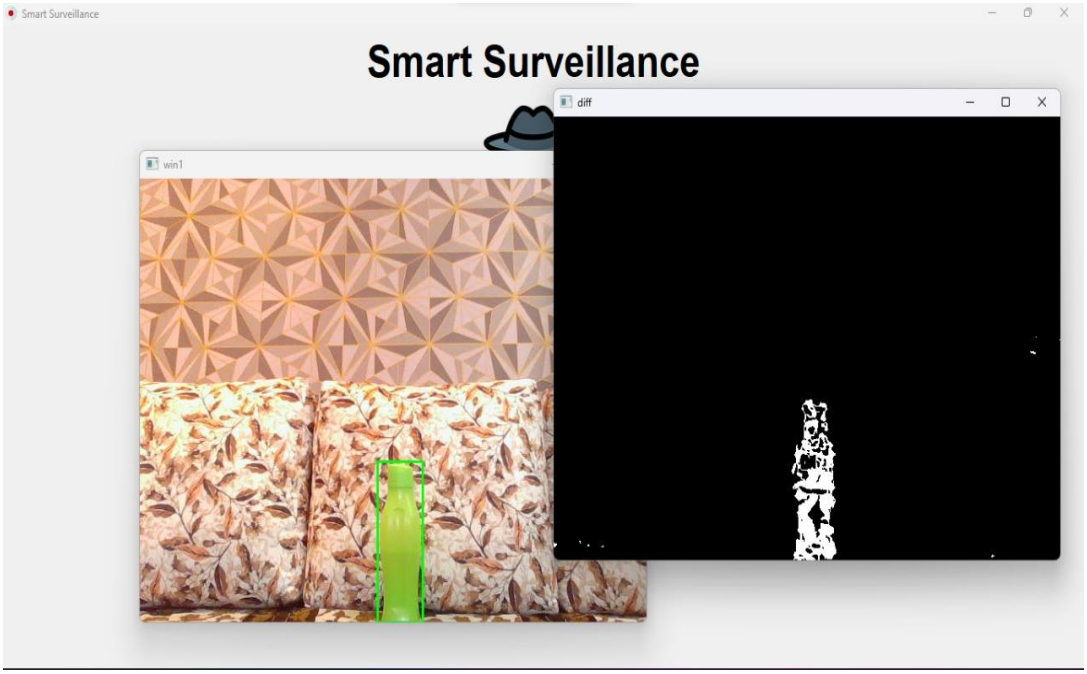
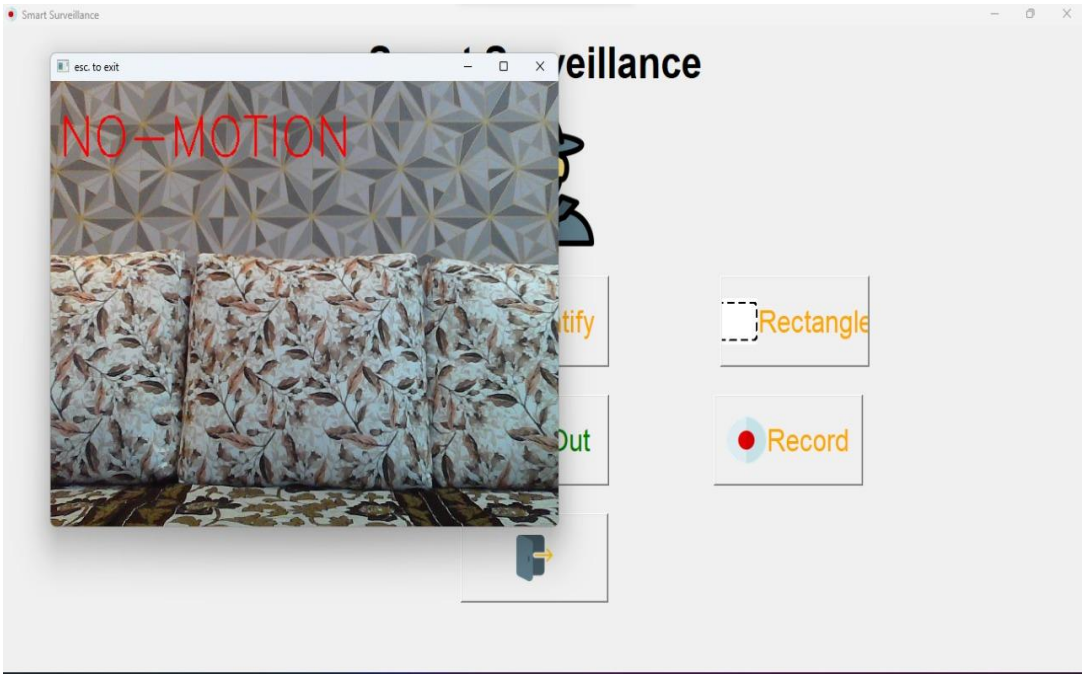


Fig. 4.1 Monitor

4.2 FEATURE 2 – ACTION DETECTION

This is working captured output for NO-Motion and Motion being detected by this application



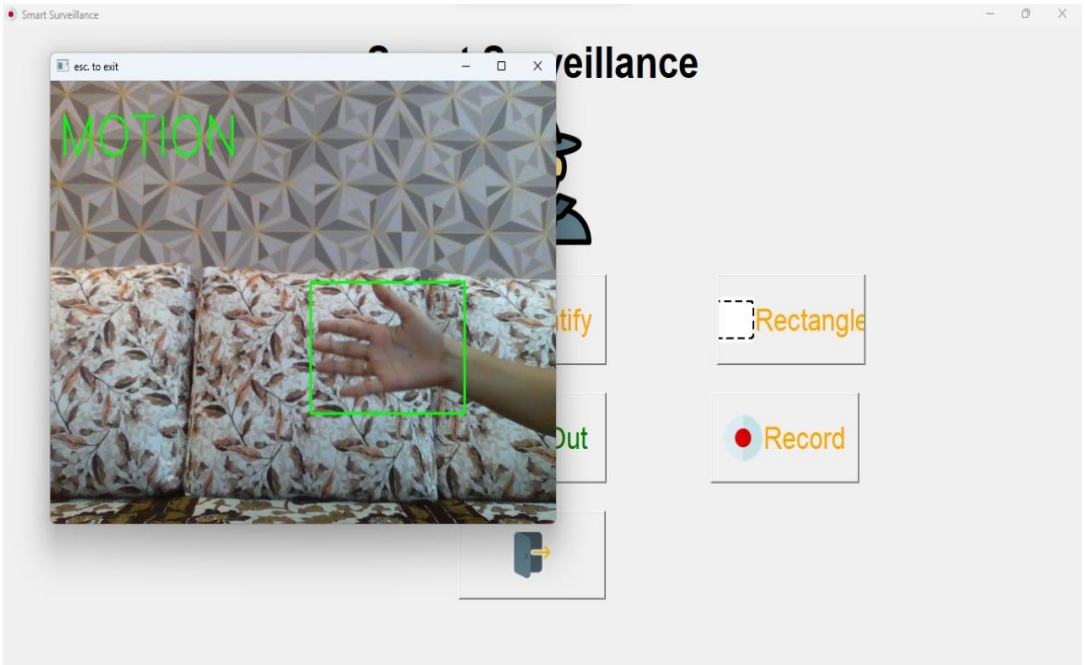


Fig. 4.2 Action Detection

4.3 FEATURE 3- RECTANGLE

This is working captured output for NO-Motion and Motion being detected by this application in the specified area.

Fig. 4.3 Rectangle

4.4 FEATURE 4 – RECORD

This Feature records the activities and save the video captures in the form of date and time.

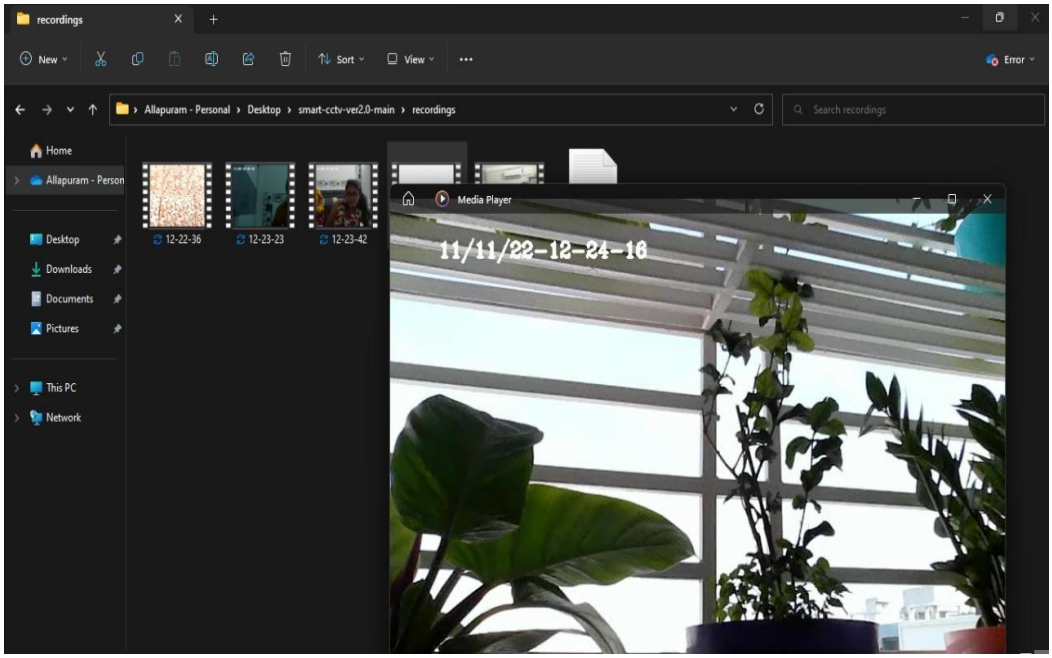


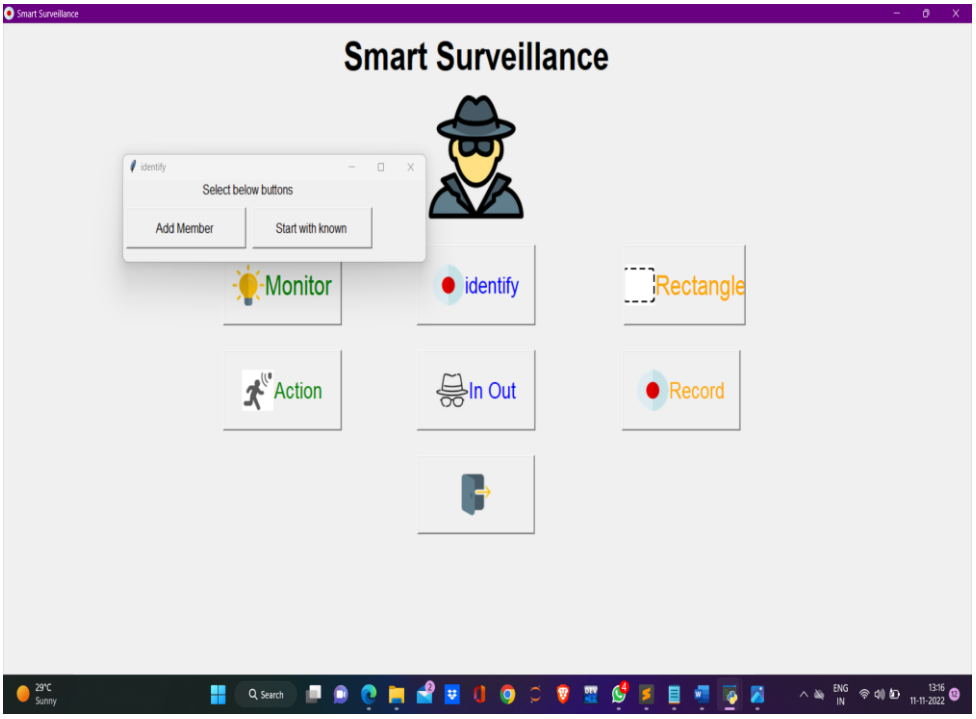
Fig. 4.4 Record

4.5 FEATURE 5- IDENTIFY

Identify is again divided into

4.5.1 Start with known – it detects the known person

4.5.2 Add New Member – we can even add new persons faces



We have trained with a dataset which includes famous personalities like Elon musk, Bill Gates, Steve Jobs, Jeff Bezos, Mark Zuckerburg.

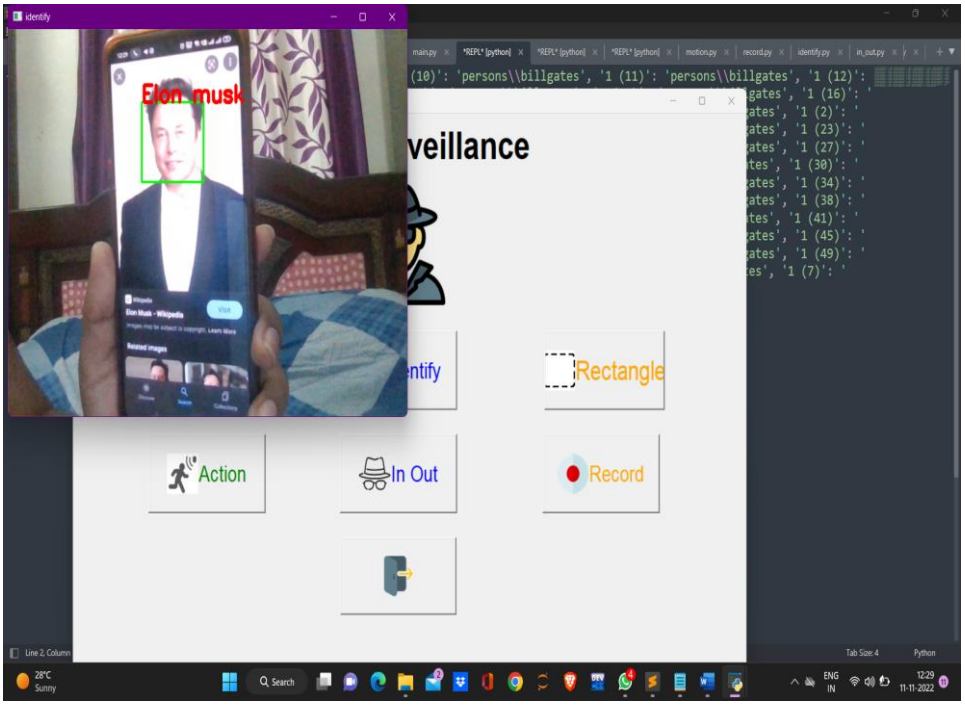


Fig. 4.5 Identify

4.6 FEATURE 6 – IN and OUT

This Feature detects the person entering into the room and leaving the room

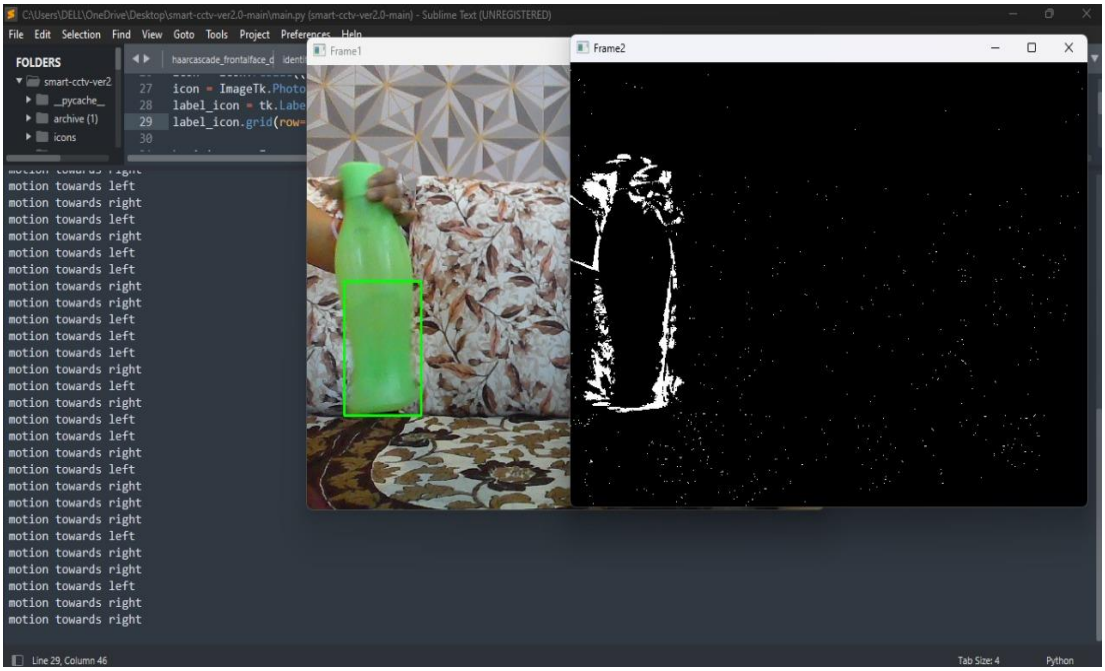


Fig. 4.6 IN and OUT

4.7 Future Scopes

Based On the technology improvements such being having the capability of small size but high processing power this project can be broadly used. Below are some future workouts on this project.

- Creating Portable CCTV.
- Adding in-built night vision capability.
- Adding deep learning if have high power device.

  More feature such as

  - Deadly weapon detection
  - Accident detection
  - Fire Detection
- Making a stand-alone application with no requirements such as python, etc.
- Making standalone device.

Adding DL support would create broad scope in this project such as with DL we would be able to add up much more functionality.

CHAPTER 5
CONCLUSION

In the end, we conclude that each person wants to be in an exceedingly better and secure world, this paper has covered the advantages and downsides of each paper that has been published to this point. To produce better security and safety new system was implemented using LBPH and SSIM and also which are cost-effective and form better further research goes on.

We can have future enhancements on this project such as creating Portable CCTV, Deadly weapon detection, Accident Detection, Fire Detection.

CHAPTER 6

REFERENCES

[1] A hybrid algorithm for multi people counting and tracking for smart surveillance, Mahwish Pervaiz, Ahmad Jalal, Kibum Kim,2021.

[2] Performance Analysis of Object Detection Algorithm for Intelligent Traffic Surveillance System, Chethan Kumar B, Punitha R, Mohana,2020.

[3] Multi-Person Tracking in Smart Surveillance System for Crowd Counting and Normal/Abnormal Events Detection, Ahsan Shehzad, Kibum Kim,2019.

[4] Multi-Person Tracking in Smart Surveillance System for Crowd Counting and Normal/Abnormal Events Detection, Ahsan Shehzad, Kibum Kim,2019.

[5] IoT Based Smart Surveillance Security System using Raspberry Pi, N Neha Patil, Shrikant Ambatka, r and Sandeep Kakde,2017.

[6] Smart Video Security Surveillance with Mobile Remote Control by Tasleem Mandrupkar Manisha Kumari Rupali Mane (2014).

[7] Video Monitoring and Motion Detection System based on arm-Linux platform and HTTP protocol with SMS capability Nava Jeevan Raju, O. and Praveen, P. (2014).