

Smart Surveillance System

¹Nikhil Kotrashetty, ²Omkar Ramachandra Hegde

^{1,2}Student, Department of Electronics and Communication, KLS Gogte Institute of Technology, Belagavi, India

Abstract - In recent times, it has become very necessary to secure each and every place (public/private), like stations, bus stands, agricultural fields and also other general areas from intruders. For this video surveillance is playing a very important role to secure the areas. In our project, we present an intelligent framework to detect the multiple events in the region of interest with this video surveillance. From the region of interest we get a set of variables, create classes like humans or vehicles, attributes of the classes like speed or locality, and create notions to detect and understand the activities in the circumstances. The objects may be pet, human or any vehicle and are detected on the basis of ratio of height and width. After calculation if objects falls into range then a boundary box is set around that object and a signal is given to the users. If this intruder system is used in the house then an additional algorithm can be set to check if a familiar face is detected, in this case system does not consider the person as intruder. Intruders are found out and marked by a boundary box or colored red in the frames given by video of CCTV or any camera footage.

I. INTRODUCTION

Image processing technology has come a long way in the last twenty years. Today, machines are able to automatically verify identity information for secure transactions, for surveillance and security tasks, and for access control to buildings. These applications usually work in controlled environments and recognition algorithms that can take advantage of the environmental constraints to obtain high recognition accuracy. Intensive research is going on in the field of intrusion detection. In our project we've managed to develop an intelligent framework to detect the multiple events in the region of interest with this video surveillance. Our project also includes algorithms for face recognition and another set of algorithms for voice interaction with the intruder.

II. MOTIVATION

Existing digital video surveillance systems provide the infrastructure only to capture, store and distribute video, while leaving the task of threat detection exclusively to human operators. Human monitoring of surveillance video is a very labor-intensive task. It is generally agreed that watching video feeds requires a bigger level of visual attention than most

everyday tasks. One of the conclusions of a recent study by the US National Institute of Justice into the effectiveness of human monitoring of surveillance video, is as follows "These studies demonstrated that such a task[manually detecting events in surveillance video], even when assigned to a person who is dedicated and well-intentioned, -not support an effective security system After only 20 minutes of watching and evaluating monitor screens, the attention of most individuals has degenerated to well below acceptable levels. Monitoring video screens is both boring and mesmerizing. There are no intellectually engaging stimuli, such as when watching a television program.".

Real Time Alerts: real time alerts can be generated using smart surveillance. In our project we have managed to generate couple of such alerts.

1. Motion Detection: This alert detects movement of any object within a specified zone.
2. Motion Characteristic Detection: These alerts detect a variety of motion properties of objects, including specific direction of object movement (entry through exit lane), differentiating between humans and pets with respect to their heights.

III. LITERATURE SURVEY

- Rama Moorthy H, Vijeth Upadhy, Vidyesh V Holla [1], proposes a CNN based Smart Surveillance System: A Smart IoT Application Post Covid-19 Era. The proposed system is for automatic detection and recognition of human faces for finding criminals/suspects/missing persons for surveillance. The intended method first detects a face in the video using a face detection algorithm and then it will search whether the face is present in the data centre. The method provides the ability to detect, extract features, and recognize a face from inputs taken by camera or video automatically.
- Namrata Singh, A. K. Daniel, Pooja Chaturvedi [2], proposes a Template matching for detection & recognition of frontal view of human face through MATLAB. Describes an approach of face detection technique that includes major characteristics such as lightening compensation based on luminance (Y) & chrominance (Cr), Color segmentation, skin tone statistics & eye-mouth region computation. A template matching algorithm using cross correlation method for

locating & recognizing a face has been applied on various face candidates to match the template with right face candidate. Thus, the presented work is divided into three steps: Face detection, Computation of template matching & Face recognition. The performance of given approach has been evaluated on the basis of run time & accuracy. The simulation result shows that the defined model is efficient in terms of accuracy which is 81% and the false alarms are reduced.

- Lisha Xiao, Qin Yan [3], proposes a Scene Classification with Improved AlexNet Model. AlexNet model is limited in image classification because of the large convolution kernel and stride in the first convolutional layer leading to over rapid decline of feature maps resolution and excessive compression of spatial information. This paper proposed an improved AlexNet model according to the design principle of convolutional neural networks (CNNs). The large convolution kernel is decomposed into a structure cascaded by two small convolution kernels with reduced stride. Another convolutional layer is added after the first one to enhance the integration process of the low level features or the spatial information. The asymmetric convolution kernel is applied in the last three convolutional layers. The experiments on two datasets show that the classification accuracy of the improved AlexNet model is higher than those of AlexNet model and ZFNet model for 23 categories of scene classification.
- Haris Papasaika-Hanusch [4], proposes techniques for digital image processing, explains general commands and data types, and images information's while dealing with digital images. Different arithmetic operations and their results on images.
- Sapana K. Mishra, Kanchan .S Bhagat [5], proposes a Human Motion Detection and Video Surveillance Using MATLAB. A human body detection algorithm based on the combination of moving information with shape information is proposed in the paper. Firstly, Eigen-object computed from three frames in the initial video sequences is used to detect the moving object. Secondly, the shape information of human body is used to classify human and other object. Furthermore, the occlusion between two objects during a short time is processed by using continues multiple frames.

IV. OBJECTIVES

To design and develop a smart surveillance system, with face recognition and object differentiation capabilities using MATLAB and CNN.

1. To detect a person in the region of interest.
2. To distinguish between humans from pets.

3. To check if the person is familiar, if not.
4. To assist the person via voice commands.

V. PRIOR ART SEARCH (IPR)

- Title: FACIAL RECOGNITION SYSTEM FOR SECURITY ACCESS AND IDENTIFICATION.

A method and apparatus for identifying individuals for the purposes of determining clearance access or Surveillance is characterized by enrolling an image of a person's face either Voluntarily or Secretly to be later used for comparison when the person Voluntarily desires clearance or is covertly detected. The System can recognize or identify individuals regardless of whether the individual is wearing eye glasses or attempted disguises. In one embodiment, the System allows an authorized operator to enroll an image of a person through a facial Scan for Subsequent clearance access. The System records the camera positioning, captures an image, Scales the image and records data from a region of interest within the Scanned image to a database for later comparison. Enrolment data and the corresponding image information are then associated with a personal identification number assigned to the person. Upon presentment before the system, and entering the assigned personal identification number, another facial Scan is taken of the person to be compared with the data from the regions of interest from the enrolment database of images to confirm the identity of the individual. In a Second embodiment, the System operator injects an image Secretly taken of an individual for later Surveillance and identification. In a Surveillance mode, the System automatically detects a person's presence, positions the camera through analysis of the image, captures an image, and then processes the image to determine if the person is enrolled in the enrolment database

- Title: SMART VIDEO SURVEILLANCE SYSTEM ENSURING PRIVACY

This invention describes a video surveillance system which is composed of three key components 1—smart camera(s), 2—server(s), Salient(s), connected through IP networks in wired or wireless configurations. The system has been designed so as to protect the privacy of people and goods under surveillance. Smart cameras are based on JPEG 2000 compression where an analysis module allows for efficient use of security tools for the purpose of scrambling, and event detection. The analysis is also used in order to provide a better quality in regions of the interest in the scene. Com- pressed video streams leaving the camera(s) are scrambled and signed for the purpose of privacy and data integrity verification using JPSEC compliant methods. The same bit stream is also protected based on JPWL compliant methods for robustness to transmission errors. The operations of the smart camera are optimized in order to provide the best

compromise in terms of perceived visual quality of the decoded video, versus the amount of power consumption. The smart camera(s) can be wireless in both power and communication connections. The server(s) receive(s), store(s), manage(s) and dispatch (es) the video sequences on wired and wireless channels to a variety of clients and users with different device capabilities, channel characteristics.

▪ Title: VIDEO SURVEILLANCE SYSTEM

A method and system of controlling a video surveillance system comprising providing a distributed video surveillance system having a plurality of devices, building a script having a parameter that is filled in when an event occurs, storing the script, detecting the occurrence of an event, loading data related to a first device from the plurality of devices in a first parameter to generate a first script, the first device being associated with the detection of the occurrence of an event, and executing the first script.

▪ Title: SYSTEM AND METHOD FOR FACE RECOGNITION USING SYNTHESIZED IMAGES

A system and method that includes a virtual human face generation technique which synthesizes images of a human face at a variety of poses. This is preferably accomplished using just a frontal and profile image of a specific Subject. An automatic deformation technique is used to align the features of a generic 3-D graphic face model with the corresponding features of these pre-provided images of the subject. Specifically, a generic frontal face model is aligned with the frontal image and a generic profile face model is aligned with the profile image. The deformation procedure results in a single 3-D face model of the specific human face. It precisely reflects the geometric features of the specific Subject. After that, Subdivision spline Surface construction and multi-direction texture mapping techniques are used to smooth the model and endow photometric detail to the specific 3-D geometric face model. This smoothed and texturized specific 3-D face model is then used to generate 2-D images of the Subject at a variety of face poses. These synthesized face images can be used to build a set of training images that may be used to train a recognition classifier.

VI. PROBLEM DEFINITION

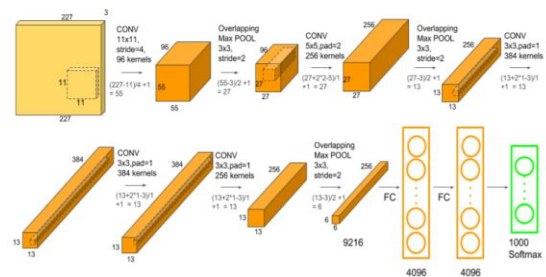
To design and develop a smart surveillance system, with face recognition and object differentiation capabilities using Matlab and CNN.

VII. TOOLS USED

AlexNet: AlexNet is classic convolutional neural network architecture. It solves the problem of image classification

where the input is an image of one of 1000 different classes (e.g. cats, dogs etc.) and the output is a vector of 1000 numbers. The *i*th element of the output vector is interpreted as the probability that the input image belongs to the *i*th class. Therefore, the sum of all elements of the output vector is 1. The input to AlexNet is an RGB image of size 256x256. This means all images in the training set and all test images need to be of size 256x256. If the input image is not 256x256, it needs to be converted to 256x256 before using it for training the network. To achieve this, the smaller dimension is resized to 256 and then the resulting image is cropped to obtain a 256x256 image.

AlexNet allows for multi-GPU training by putting half of the model's neurons on one GPU and the other half on another GPU. Not only does this mean that a bigger model can be trained, but it also cuts down on the training time. In our project we are using single GPU.



AlexNet consists of 5 Convolutional Layers and 3 Fully Connected Layers. Multiple Convolutional Kernels (a.k.a filters) extract interesting features in an image. In a single convolutional layer, there are usually many kernels of the same size. For example, the first Conv Layer of AlexNet contains 96 kernels of size 11x11x3. Note the width and height of the kernel are usually the same and the depth is the same as the number of channels.

Transfer Learning: It is a popular approach in deep learning where pre-trained models are used as the starting point on computer vision and natural language processing tasks given the vast compute and time resources required to develop neural network models on these problems and from the huge jumps in skill that they provide on related problems. We used transfer learning because we can take a pretrained network and use it as a starting point to learn a new task. In our project the task is face recognition. We have a data set of different persons (150 face photos of each individual of size 256x256) stored in the system. The data set is created manually and the system recognises these people. We did not use any data set from internet.

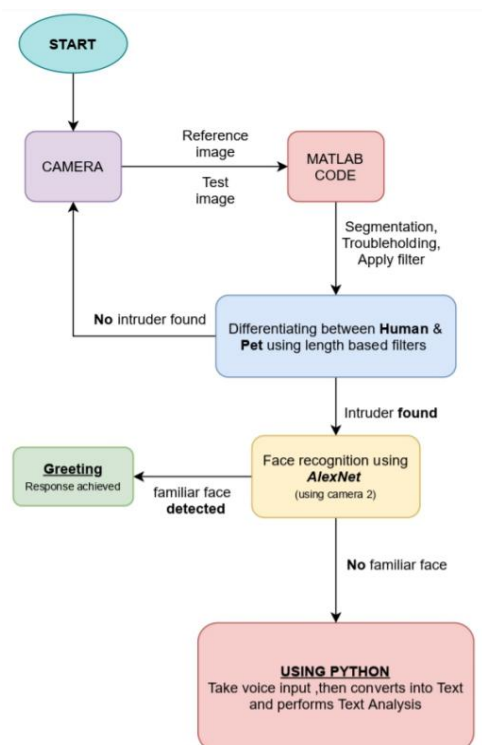
MATLAB: MATLAB has many toolkits and support packages to facilitate our project. It allows using multiple cameras in our project. Image processing is much faster and

efficient on MATLAB. Image Processing Toolbox provides a comprehensive set of reference-standard algorithms and workflow apps for image processing, analysis, visualization, and algorithm development. You can perform image segmentation, image enhancement, noise reduction, geometric transformations, image registration, and 3D image processing. Differentiating between humans and pets by taking in account of their height in pixels is very effective and can be efficiently done in MATLAB.

VIII. ALGORITHM

- Step 1: Take a snap of the surrounding which is to be monitored. This is our reference image using MATLAB.
- Step 2: Camera starts capturing photos (20-25 photos per second) and compares it with the reference image.
- Step 3: Differentiate humans and pets using length based filters. (Assuming the height of the pet animals will not exceed 80px we can differentiate between humans and animals)
- Step 4: Highlight the intruder if found and call face recognition script.
- Step 5: Face recognition- AlexNet algorithm is used to identify the person from the stored data set.
- Step 6: If known person is detected, the system greets the person. If unknown person is detected MATLAB calls python.
- Step 7: Python- takes in voice input from the person convert it into text and perform text analysis to understand the purpose of visit of the person and assist the person via appropriate voice response.

Flow chart:



IX. RESULTS AND DISCUSSION

For implementation and testing our work, it is decided that using a camera and creating a smart security system with image processing would be efficient, as it offers customizability and is much harder for an intruder to avoid. The number of cameras used in our project is 2. One for detecting an intruder and second camera for face recognition. The customizability of image processing will let us apply filters and functions to improve the accuracy of the system, thus satisfying the requirement of minimizing false positives and false negatives. Additionally, through image processing we will also be able to differentiate between single or multi objects in a frame, and also differentiate between humans and pets, and the additional feature of face recognition using AlexNet will help us identify if the intruder is someone familiar. Text analysis is more convenient using python. Python will use voice input to understand the purpose of visit of the unknown person. This is done using python packages for speech recognition. Voice will then be converted into text. Now python will perform text analysis and look for words such as delivery, packages, courier, and post. Once python finds any of these words it will give appropriate voice response to the person.

X. SNAPSHOTS

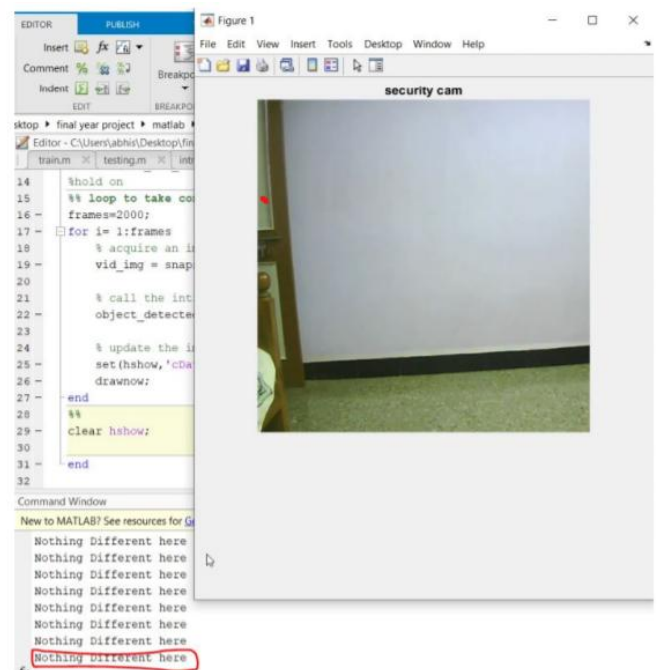


Figure 1: Security cam, Area to be monitored

In above fig, there is no moving object in the frame. So MATLAB displays the message nothing different here.

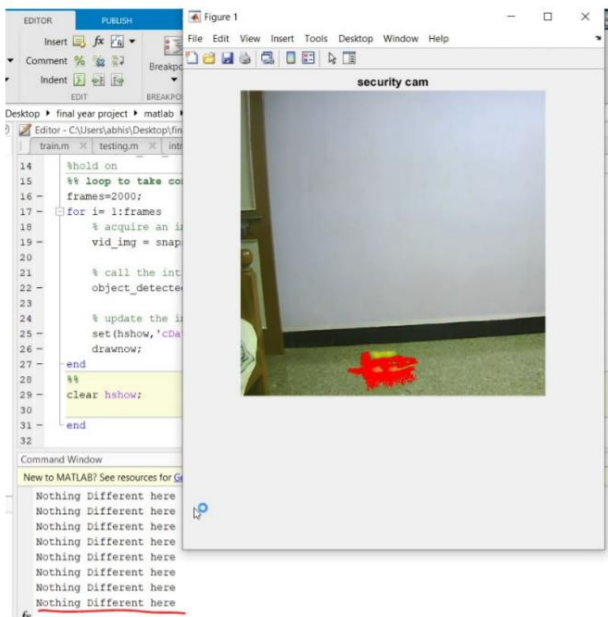


Figure 2: A cat in the frame highlighted

In the above fig a cat appears in the frame. The system differentiates the cat from human on the basis of height filter and ignores the cat. So the message displayed is still “Nothing Different here”.

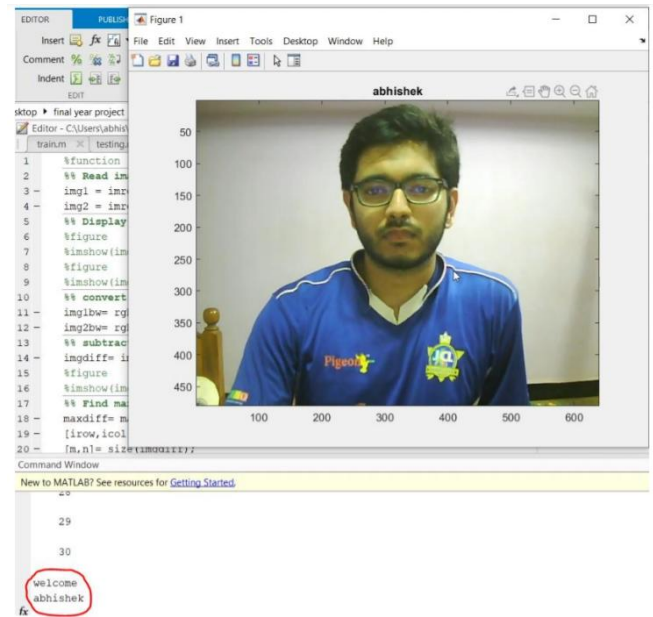


Figure 4: Systems recognizes the person

In the above fig, the person’s data set is present in the system so when the system successfully recognises the person the system greets the person with a voice note.

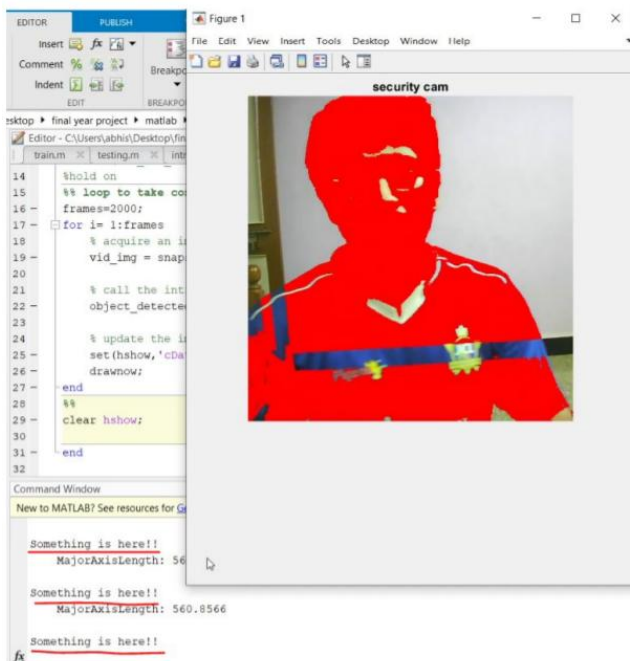


Figure 3: A human Intruder comes in the frame

In the above fig, a human comes in the frame, our system highlights the human in red and the message displayed is now “Something is here” At this point, Matlab calls face recognition script to identify the intruder.

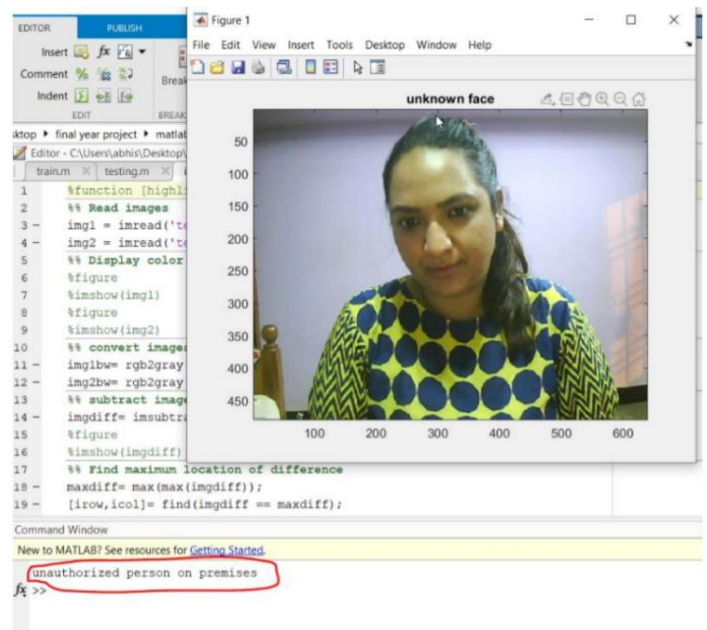


Figure 5: Systems fails to recognizes the person

In the above fig, the person’s data set is not present in the system so when the system fails to recognizes the person the system responses with a message –“unauthorized person on premises”, along with a voice note saying the same.

Non- functional requirements are the requirements that specify the performance parameters of the system Overview of SRS.

XI. CONCLUSION

Illegal entrance by the intruders on borders, house, office, etc and obstacle detection are challenging issues. So now a day, video surveillance using the concepts of image processing and machine learning is very effective in handling such challenges. In future, more work can be done to know the multiple functionalities of the intruder, like what the intruder is doing, if he is carrying any weapon, and also record the voice of the intruders to understand the intension of the intruder more clearly This project will provide number of benefits to users working in different fields. The fields which will be benefited are: - home, office, parking lot, border security, public places.

XII. SCOPE FOR FUTURE WORK

- Multiple cameras can be attached to the system for better face recognition.
- Using IOT the system can be re-designed to send alert messages along with the photo of the intruder to the user.
- Can add more default voice responses to some basic questions usually asked when a new person visits a place.

REFERENCES

- [1] Haris Papasaika-Hanusch “Digital Image Processing Using Matlab” Institute of Geodesy and Photogrammetry, ETH Zurich, 2018.
- [2] Lisha Xiao, Qin Yan, Shuyu Deng, “Scene Classification with Improved AlexNet Model”, 2017 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE).
- [3] Rama Moorthy H, Vijeth Upadhy, Vidyesh V Holla, Sunil S Shetty, Vinay Tantry. “CNN based Smart Surveillance System: A Smart IoT Application Post Covid-19 Era” Proceedings of the Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) IEEE Xplore Part Number: CFP20OSV-ART; ISBN: 978- 1-7281-5464-0.
- [4] Namrata Singh, A. K. Daniel, Pooja Chaturvedi, “template matching for detection & recognition of frontal view of human face through matlab”, international conference on information, communication & embedded systems (icices 2017).
- [5] Sapana K. Mishra, Kanchan. S Bhagat, “Human Motion Detection and Video Surveillance Using MATLAB”, International Journal of Scientific Engineering and Research (IJSER), ISSN (Online): 2347-3878, Impact Factor (2014): 3.05, “Probabilistic anomaly detection in natural gas time series data”, Int. J. Forecast., 32 (2015), pp. 948-956.
- [6] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks” <https://papers.nips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf> (27.5.2021).
- [7] Emmanuel Dare Alalade, Intrusion Detection System in Smart Home Network Using Artificial Immune System and Extreme Learning Machine Hybrid Approach, School of Information Technology University of Cincinnati Cincinnati, Ohio, United States 45221.
- [8] Akshay Bharadwaj K H, Deepak, Ghanavanth V, Harish Bharadwaj R, Uma R, Gowranga Krishnamurthy, “Smart CCTV Surveillance System for Intrusion Detection with Live Streaming”, 2018 3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology, MAY 18th & 19th 2018.
- [9] Ms Bulbul Bamne, Ms Neha Shrivastava, Mr Lokesh Parashar, Mr Upendra Singh, “Transfer learning-based Object Detection by using Convolutional Neural Networks”, Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC 2020), IEEE Xplore Part Number: CFP20V66-ART; ISBN: 978-1-7281-4108-4.
- [10] <https://www.mathworks.com/videos/introduction-to-matlab-with-image-processingtoolbox-90409.html> (11.11.2020).

Citation of this Article:

Nikhil Kotrashetty, Omkar Ramachandra Hegde, “Smart Surveillance System” Published in *International Research Journal of Innovations in Engineering and Technology - IRJIET*, Volume 5, Issue 12, pp 1-6, December 2021. Article DOI <https://doi.org/10.47001/IRJIET/2021.512001>
