Smart Security Surveillance

Sharan Patel[1] Nayan Ghodke[2] Tanishq Aher[3] Dinesh Kodwani[4] Mr. Avinash Taskar[5]
[5]Project Guide
[1,2,3,4,5]Department of Computer Science & Engineering (AIML)
[1,2,3,4,5]Sandip University, Nashik, Maharashtra, India

Abstract — In sensitive area where generally no one is allowed So first we will detect the motion after The motion detection it automatically revoke the functions for face detection. Around the world, putting on a mask, washing our hands often, and avoiding social contact have all become priorities. The World Health Organization (WHO) strongly advises against spreading the new coronavirus by using a mask that covers the mouth and nose. The primary objective of a computer vision system is to find moving objects. The performance of these systems is insufficient for many applications. One of the main causes is that moving object detection is more difficult when dealing with various constraints, including environmental changes. After object recognition, object counting is a process that is improved upon and made more reliable with the aid of OpenCV. OpenCV comes with a variety of helpful methods for item recognition and counting. Numerous industries, including transportation, health care, and environmental science, use object counting.

Keywords: Face Mask Detection, Motion Detection, Object Recognition, OpenCV, Image Processing

I. INTRODUCTION

With the growth of technology, there is a growing worry for safety and security everywhere nowadays. The number of security cameras has lately grown in order to address this problem. Despite this, it is challenging to manually store and monitor data [1] continuously. Unexpected effects of the coronavirus vary among people. Mild to severe sickness will be experienced by the majority of infected persons. and recuperate outside of the hospital. Fever, a dry cough, and fatigue top the list of usual symptoms[2]. Therefore, using masks and hand sanitizers has proven to be successful in preventing the transmission of the virus. As a result, a face mask detection system is needed, which will inform the public and aid in the fight against the epidemic. In this research, we investigate real- time face mask detection with deep learning and OpenCV. Deep learning is a branch of machine learning that uses algorithms with human brain-inspired design. In many domains, deep learning aids in the breakdown of issues. Convolutional neural networks, image detection, and picture classification are all provided by deep learning (CNN). Computer detection and classification tasks primarily utilise convolutional neural networks (CNN). In this study, deep learning algorithms are utilised to distinguish between faces wearing masks and those who are not. To incorporate the optimal amount of Convolutional Neural Layers for precise detection, Convolutional Neural Networks (CNN) are utilised.

A programming function library called OpenCV is primarily focused on real-time computer vision, machine learning, and image processing. The primary objectives of computer vision are to manipulate and extract data from a real-time source. It is applicable to automobiles with autonomous driving. The detection of faces, objects, and handwritings is done using OpenCV. It is crucial for both mask- and face-mask detection.

Many different moving object identification methods have been described in recent years. In contrast to human vision, a machine vision system's artificial method of detecting motion was based on statistical backdrop modelling. Our strategy was to periodically take a photo using the camera (make it the current picture), compare it to a prior picture, and keep both if there was a noticeable difference; otherwise, we would release memory from the old picture and make the new picture the current picture.

Motion detection encompasses the recognition and documentation of the complete movement procedure. Automatically detecting movement in a certain area is the aim of motion detection [3]. Finding movement in a certain region is the aim of motion detection. In camera geometry, this area is always represented by a zone of awareness known as the field of vision. It is also known as an environment that is monitored. In this instance, the environment with its moving objects and activity is the focus. A zone of interest, which is referred to as a moving object, might include a person, an animal, or an object.

II. RELATED REVIEW ON LITERATURE

Many different moving object identification methods have been described in recent years. A strategy based on statistical background modelling was recommended by Mahbub et al. [4]. To find moving objects, this technique compares each edge segment of the current frame with each edge segment of the backdrop. However, this method is unable to identify a moving edge segment that crosses a background edge segment. The goal of Geetha Priya S et al[5] .'s You Only Look Once (YOLO) approach for object detection is to find things. This methodology offers a lot of benefits over other object detecting strategies. While other algorithms, like Convolutional Neural Network and Fast Convolutional Neural Network, only partially examine the image, YOLO does so by using convolutional networks to predict the bounding boxes and the class probabilities for these boxes. As a result, YOLO detects the image more quickly than other algorithms.

The authors demonstrated the results of adding several filters to a picture, such as colour conversion, grey scale, and dilation.

To gauge how long the algorithms required to finish, the authors tested the features on three distinct smartphones. An instructional manual on how to use OpenCV for image recognition was written by the authors in [6].

For crucial industries like banks and enterprises, the authors want to create a good motion detection system. In order to determine the difference between the sampled photos, they start by extracting samples from a webcam that is capturing images. A counter is increased when they detect

movement, and when it reaches a certain level, they send a message to a mobile phone informing it that movement has taken place and sound a buzzer.

The Gaussian Mixture Model, a probability density function, served as the foundation for the authors' strategy to achieve this (GMM). The scientists changed the Gaussian parameters and component count of the traditional GMM method to shorten processing time. Each pixel in each frame was subjected to the authors' GMM method to create foreground-background segmentation. Once the foreground and background have been defined, the authors convert the frame to binary, with the foreground turning white (moving things) and the background turning black (none moving objects). The binary pictures are then subjected to filters by the authors to remove noisy elements as wind, light, and shadows.

III. COMPUTER VISION

Computer vision (CV) is a subfield of computer science that focuses on giving computers the ability to understand visual data. Gerald Jay Sussman was given the task by Martin Minsky in the early 1970s or late 1960s to connect a computer to a camera and have the device report what it saw.

The field of research known as "computer vision" (CV) strives to develop methods that will enable computers to "see" and decipher the content of digital images such as photos and videos. It appears to be simple since everyone, even very young toddlers, can figure out the computer vision issue. Despite this, the issue is still largely unresolved due to both a lack of knowledge on biological vision and the intricacy of visual perception in a dynamic physical universe that is practically infinitely changing.

A. Relationship between computer vision and AI

As seen in figure 1, computer vision is a field of study that focuses on helping computers see. It is a multidisciplinary discipline that falls under the umbrella of machine learning and artificial intelligence, and it may make use of both specialised methods and generic learning algorithms.

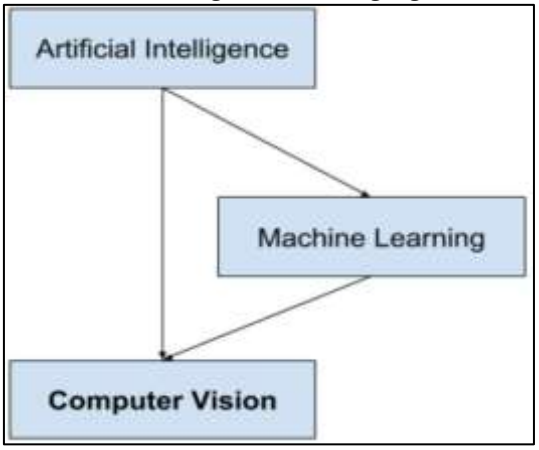


Fig. 1: The Relationship between Computer Vision & AI

B. Image Processing

Computer vision is not the same as image processing. Image processing is the act of transforming an existing image into a new one, typically by streamlining or improving the information. It is a type of digital signal processing that is not focused on deciphering the content of the images. For some computer vision systems, image processing, such as picture pre-processing, may be necessary.

Motion detection can be used to automatically carry out a job when motion is detected. Motion detection, for instance, can be used in security systems to identify illegal activity or to illuminate a room when people enter it.

Human bodies emit infrared radiation when they produce heat, which an infrared sensor can detect. Image processing can identify motion by comparing two pictures [6] to one another. By comparing pixels in the same positions in the two photographs, this is done. There is no difference between two pixels if the photographs are identical. The pixel difference between certain pixels will be more than zero if the photographs are not identical. Differences in the pixels where the people are in the image will be visible when comparing an image of a room that is empty to an image of the same room with people inside.

IV. OPENCV

The infrared radiation that human bodies release is detected using an infrared sensor. A free software library for computer vision and machine learning is called OpenCV (Open Source Computer Vision Library). Computer vision is essential to artificial intelligence or serves as its main foundation. Computer vision is vital to robots, self-driving automobiles, and image editing software. For taking, processing, and analysing photos, an Android application subsystem called OpenCV was used.

Additionally, these algorithms have MATLAB, Java, and Python bindings. It primarily concentrates on face identification, image processing, and video capture. The hardware acceleration of the underlying heterogeneous computing architecture benefits Open CV. Nearly 2,500 algorithms for various computer vision methods are available through Open CV. Because Open CV is a cross-platform library, it can run on any OS.

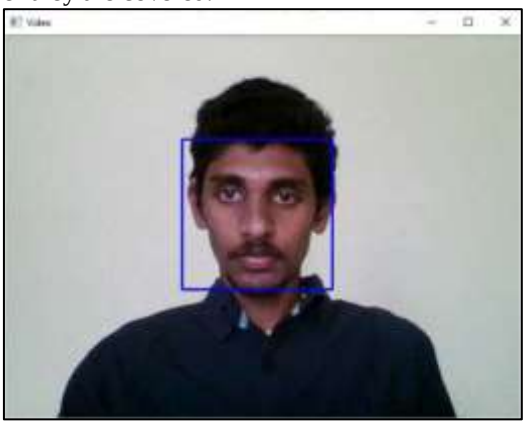
The library, which comprises a thorough blend of both traditional and cutting-edge computer vision and machine learning approaches, has more than 2500 optimised algorithms. For the purpose of processing images, OpenCV was developed. Each function and data structure was created with the image processing programmer in mind. Contrarily, Matlab is a fairly open-ended programming language. We can get just about anything in the world in the form of toolboxes. To do video tracking, an algorithm analyses successive video frames and outputs the movement of targets between the frames.

V. OPENCV

The detection of a face mask may be broken down into the following two steps: Face recognition and mask detection are the first two. Viola-Jones method, a machine learning object identification approach, and Haar feature-based cascade classifiers are used with OpenCV to accomplish face mask recognition. Rapid Object Detection Using a Boosted Cascade of Simple Features, a 2001 study by Paul Viola and Michael Jones, introduced the face detection technique. We must carry out face detection for each frame of the video in

order to achieve real-time face identification in a video. In figure2, face detection is displayed.

The image is now preprocessed and reduced in size to 224x224 pixels when the face is detected in it. Additionally, OpenCV spretrainedHaar course classifiers are used to identify the mouth and nose, which determines if a person is wearing a cover. If the person's lips and nose can be clearly seen, they are not wearing a veil, hence it is not certain whether they are covered.



VI. EXISTING PROBLEM

This paper's major objective is to enable a basic camera attached to a general-purpose computer to identify objects moving through it and to determine how long the object remained in the region where the camera is placed. Consequently, this programme may be used for monitoring. Figure 2 illustrates how to take the first frame using a camera. This frame will serve as the initial reference. To detect motion, the phase difference between the new frame with the item and the baseline frame will be computed. The brand-new frames will be called Delta frames. Your delta frame will then be adjusted using pixel intensity. The improved frame will be called Threshold. Then, to capture significant elements on the Threshold frame, you'll employ complex image processing techniques like Shadow Removal, Dilation, Contouring, and others. Here is an illustration of what you will achieve. The timestamps of objects entering and leaving the frame can be recorded. You'll be able to calculate the duration of screen time as a consequence.

Fig. 3: The Detected Object

VII. PROPOSED METHODOLOGY

In our project, we aimed to develop a surveillance system that, in addition to motion detection, could alert the user of an incursion, record video from the instant the motion was detected, and send SMS messages to the user's mobile device.

A. Comparing Phase

We compare the webcam's live photos to one another in order to detect changes in these frames and predict the presence of motion in order to determine whether any motion is present in the live Images.

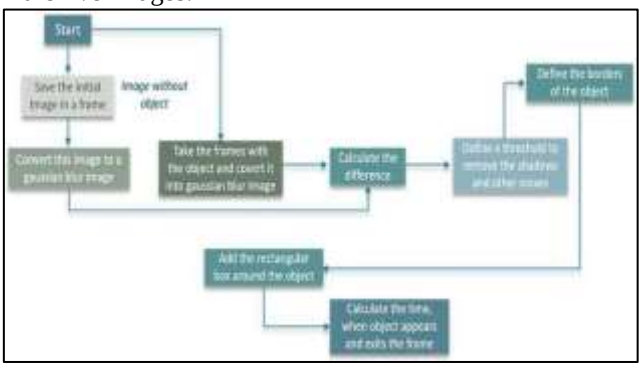


Fig. 4. Flow Diagram

B. Image Segmentation

The process of splitting a picture into parts, each with their unique set of attributes, in order to extract the target of interest is referred to as "image segmentation." The image segmentation technique used in this work is threshold segmentation. To put it another way, the foreground and background of an image are divided into two groups depending on the results of comparing a range in the picture to the threshold. The threshold segmentation procedure consists of two basic.

If we can locate a suitable threshold, we can modify the image for segmentation. An intelligent video surveillance system may be used to identify moving targets in a variety of different ways, such as frame difference and background removal techniques. Threshold segmentation is composed of two basic procedures. First, establish the threshold. The second step, which is the most crucial one in partition, compares the pixel value to the threshold value T to determine the threshold value.

Using the difference between the current picture and the background image, the background subtraction method is a method for identifying moving objects.

After subtracting the current image fk from the previously saved background image B, if the pixel difference above the bound threshold, the background pixel or the pixel to pixel on the moving target is identified. The background subtraction threshold must be properly selected in order to successfully identify motion. Motion detection's precision is crucial.

C. System Indicating Phase

The program contains an alarm mechanism because when motion is detected, the user may wish to get an SMS alert immediately away informing them that the software has detected an intrusion. This alarm system quickly creates an auditory alert signal in wav file format whenever any type of

motion is detected. This helps at the moment to avoid any kind of security compromise. When motion is detected, the user is provided a location with photos of the trespasser, and at the same time, an SMS is delivered to the user's mobile device.

VIII. OBJECT DETECTION

Every part of an object throws a shadow on the surrounding area or on other parts of the object. This can be confusing. For instance, the nose throws a shadow on your lips, any bigger stationary item casts a shadow on smaller surrounding objects, various light sources with variable luminous intensities, your room's curtains, the direction and viewing angle of the light source, etc. Some of the irregularities that have been identified in real-time collected frames are as follows.

We must filter the image as a consequence to lessen these types of disruptions. By altering the number of iterations, we may influence how smooth the Dilate function is.

Using the imshow function, which lets you display each frame in an own window, you may compare frames.

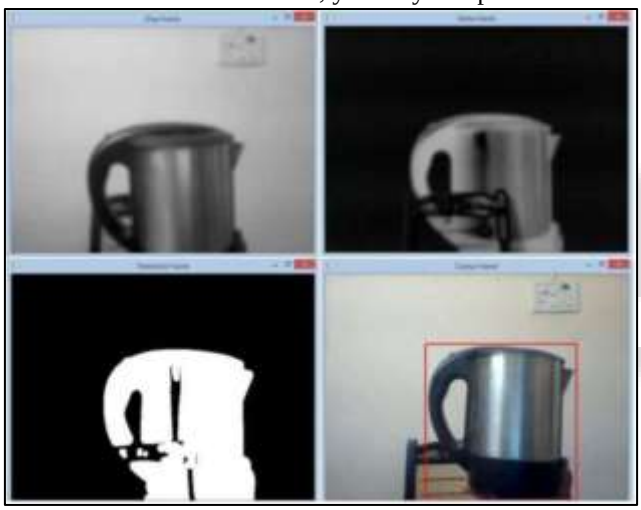


Fig. 5. The Frame with Detected Object

We use the waitKey function to postpone the operation till a key is pushed. In this instance, waitKey is used to obtain an ongoing live feed from the camera (1). To prevent the video from being recorded, just press the "Q" key on your keyboard.

IX. RESULTS AND DISCUSSION

A. Motion Detection

A home or business security system that may be very helpful in circumstances when security is a worry is the "Smart webcam motion detection surveillance system." The development of technology in the modern world has significantly increased the techniques employed by thieves and criminals to steal. As a result, surveillance systems need to develop in order to keep up with the world's changing needs. The most current tools used in the war against theft and devastation are video surveillance and monitoring.

The most important feature of digital video surveillance systems is motion detection. It enables the camera to just capture when necessary rather than constantly, which significantly reduces storage space. An alert may sound when unexpected motion is discovered. As a result, personnel are no longer necessary to constantly monitor. The motion detector satisfies the need for an inexpensive, minimal security system in daily life. The future holds great potential for computerised home security. Future is bright and made simpler with new technology.

Fig. 6: The Output Format

B. Mask detection

Using the Python computer language, real-time face mask identification with 98% validation accuracy was implemented. After a number of testing with a batch size of 32 and 20 repeats of the epoch, this is the rate that stands out the most. The results show how well the trained model performs in terms of accuracy and loss.

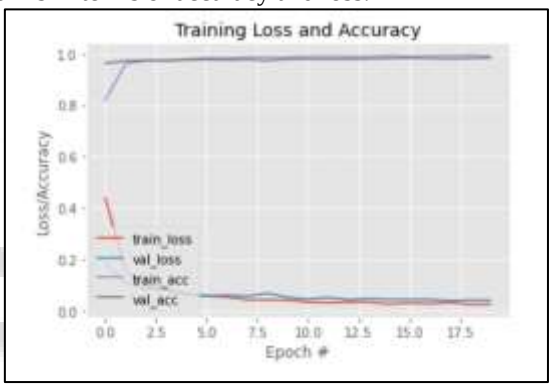


Fig. 7: Training loss and accuracy during model training.

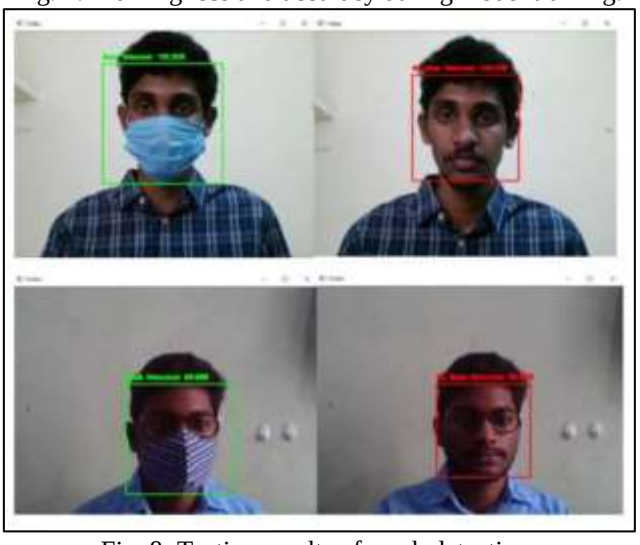


Fig. 8: Testing results of mask detection.

X. CONCLUSION AND FUTURE WORK

The most modern technologies used in the struggle against theft and destruction are video monitoring and surveillance. With the help of technology, it is possible to see and record

every square inch and passing second of the area of interest. The most important feature of digital video surveillance systems is motion detection. It enables the camera to just capture when necessary rather than always, which significantly reduces storage space. An alert may sound when unexpected motion is discovered. This frees up staff from ongoing supervision. The motion detector satisfies the need for an affordable, basic security system in daily life. The future of computerized home security is quite promising. With new technology, the future appears promising and simpler.

This application may be used for surveillance to secure any location or significant item from illegal access and to track how long an unauthorized person or object tries to breach the perimeter or enter the protected region. Additionally, more beneficial techniques and algorithms might be used to enhance the processes, such as those that track the object that generates the motion events or those that employ optical flows, also known as image flows. There are indications of human movements here. However, in the future, it could be desirable to recognize the human body posture or gesture that the prototype system's algorithm captured.

Based on the findings of the study presented in this article, real-time face mask identification using OpenCV and deep learning algorithms is a remarkable solution for straightforward facemask recognition. There are, however, a relatively small number of restrictions that may be readily solved in future work. The suggested approach is especially helpful for real- time applications that need face mask recognition, particularly in heavily crowded areas like schools, airports, and public areas. utilising a variety of IOT applications and deep learning approaches.

We can also use contactless temperature monitoring and mask detection, which can determine whether the mask is being worn properly and grant access only if both factors are met. If neither of these conditions is met, or if the body temperature is above normal, they can alert us, allowing us to take the necessary precautions to stop the spread of the Covid-19 virus indoors.

ACKNOWLEDGMENTS

Our first and foremost acknowledgment is to our supervisor and guide Prof. Shailesh P. Bendale. During the long journey of this study, he supported me in every aspect. He was the one who helped and motivated us to propose re-search in this field and inspired me with his/her enthusiasm on research, his experience, and his lively character.

We express true sense of gratitude to our guide Prof. Shailesh P. Bendale for his perfect valuable guidance, all the time support and encouragement that he gave us.

We would also like to thanks our head of department Prof. Shailesh P. Bendale, principal and management inspiring us and providing all lab and other facilities, which made this seminar presentation very convenient.

We are really thankful to all those who rendered their valuable help for the successful completion on the seminar presentation.

REFERENCES

[1] Real-Time Face Mask Detection using OpenCV and DeepLearning, Department of ECE, KoneruLakshmaiah Education Foundation, Andhra Pradesh, India.
[2] An Intelligent Motion Detection Using OpenCV. International Journal of Scientific Research in Science, Engineering and Technology Print ISSN: 2395-1990
[3] Y. You, S. Gong, C. Liu, "Adaptive moving object detection algorithm based on back ground subtraction and motion estimation", Int. J. Advancements in Computing Technology, vol. 5, no. 6, pp. 357-363, 2013 Conference,WWW2019,2019.
[4] 9M. Murshed, A. Ramirez, O. Chae, "Statistical Background Modeling: An Edge Segment Based Moving Object Detection Approach", Proc. of IEEE International Conf. on Advanced Video and Signal Based Surveillance, pp. 300-305, 2010
[5] Geethapriya. S, N. Duraimurugan, S.P. Chokkalingam, "Real-Time Object Detection with Yolo", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-8, Issue-3S, February 2019.
[6] Marengoni and Stringhini. High Level Computer Vision using OpenCV. 2011. Universidade Presbiteriana Mackenzie.

