A review on Different object Detection Techniques

*Shivam Kumar*

*II year department Computer Engineering,*

*Poornima College of Engineering Jaipur, India*

*Tripti Somani*

*II year department Computer Engineering,*

*Poornima College of Engineering Jaipur, India*

*Sonam Gour*

*Assistant Professor Department of Computer Engineering,*

*Poornima College of Engineering Jaipur,India*

***Abstract*—Scene interpretation, video surveillance, robotics, and self-driving cars are just a few in the various applications that have motivated substantial research in the field of computer vision throughout the last ten years. Visual recognition systems, which encompass image classification, localization, and identification, are at the heart of all of these techniques and have gotten a lot of interest in the scientific community. Due to significant developments in neural networks and particularly deep learning process these image recognition algorithms have attained exceptional performance. Object identification is important parameter where computer vision has had a lot of success. The purpose of this work is to conduct a methodical investigation into the relevance of object detection and its applications in the field of computer vision. Our work gives a thorough introduction to object detection, as well as various approaches, computer vision basics, and applications, which will be useful to the image processing and computer vision research communities.**

***Keywords—object detection, image processing, Deep learning, computer vision, Deep CNN.***

1. INTRODUCTION

Society has become increasingly comfortable with technology for information processing, storage, and transmission since the introduction of the electronic medium, notably the computer. Computers have a crucial part in every element of life and society in modern civilization. With the advancement of technology, man gets increasingly connected with computers as the modern age's leader, and a technological revolution has occurred all over the world as a result. It has ushered in a new era for humanity, ushering in a new world known as the technology world. Everyday life includes computer vision. Visual recognition skill equivalent to human is one of the most significant aims of computer vision. This review article seeks to familiarize the reader with the many object detection technologies accessible in today's world, as well as to inform users about the various approaches and the notion of object detection. The suggested research will take the reader on a fascinating voyage through the history of image processing and object detection, from old relics like Haar cascade detectors to today's buzzword deep learning. It is also clear that in the field of computer vision, object tracking is significant. Due to the availability of very complex processors and high-quality, low-cost cameras, object tracking algorithms have taken precedence. Vision systems and machine recognition have steadily become a popular topic in order to meet the needs of humans in the modern day. The paper also goes through some of the suggested object detection algorithms that use image processing and modification to monitor the motion of the discovered items. Many computer vision applications need the detection and segmentation of moving objects in video

streams, including video surveillance, person tracking, traffic monitoring, and semantic annotation of videos, all of which are demystified in the study from our perspective. Deep learning technology, as we all know, has been a buzzword in recent years as a result of cutting-edge breakthroughs in a range of sectors, including image classification, object identification, natural language processing, underwater object detection, object tracking, and many more. If we look at the popularity of deep learning for image processing in more depth, we can see that it is owing to the availability of large datasets and powerful graphics processing, both of which have been satisfied throughout this time period. The overview of the paper has been illustrated int the figure below.

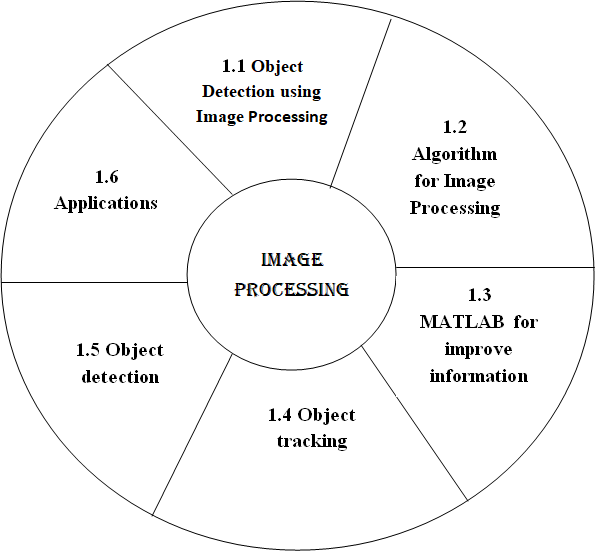


Figure 1: overview of the paper

1. IMAGE PROCESSING

In today's contemporary world, every pixel is a piece of data, and each frame has millions of them. However, extracting that data and putting it to good use is a difficult task, and we adopt image processing to solve this problem. Image processing is a technique in which we utilize digital computers to extract information from a digital image or frame using various techniques and approaches.

Image processing allows us to extract minute characteristics and information from a frame, which may then be used for other actions [1]. It helps us with picture improvement, compression, and restoration. These processes can aid us in

activities as simple as character recognition to as sophisticated as cancer cell identification in people.

1. *Object Detection using Image processing*

In this section we will discuss object detection through image processing with help of Haar cascades which was proposed by Viola and Jones. This algorithm was developed long before modern techniques like deep learning where developed.

Before discussing about the Haar cascade, let’s have a look at the method of image processing and how it is performed. Image processing is a methodology to perform some operations on image by converting it into digital form.

These operations can be of any type such as enhancing the quality, extracting some useful information or changing the tone of image.

Image processing can be categorized into three steps:

* Selecting or importing an image.
* Manipulating and analyzing the image.
* Output image.

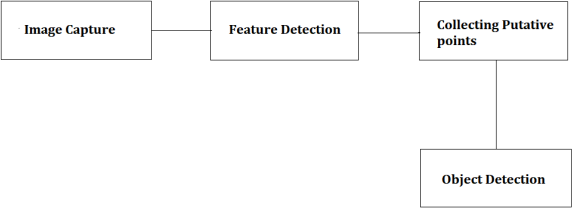


Figure 2.1: object detection steps

So, what exactly is the Haar Cascade? It is a face detection algorithm that detects a person's face in an image or real-time video. Viola and Jones suggested the detection process using a Cascade structure of boosted network. It was made public in 2001. The algorithm is trained by feeding with large number of positive and negative images.

To understand the working of Haar Cascades we we’ll apply it on a face/object detection system. We can categorize the working of a face/object detection system into three major categories:

* Detecting a face that needs to be tracked.
* Identifying facial features that will be tracked throughout.
* Tracking the face.
  + Detect a face to track - To track a face we need to detect it first and for this purpose we use cascade object detectors, the algorithm used by the object detectors is Viola Jones. The cascade detectors can detect a face very efficiently but if there is a change in the alignment of face then it is not able to detect the face in these successive frames.
* Identify facial features to track - The next important step is to detect a facial feature that will remain constant throughout the flame frames so that it is easy for us to track that face. That feature can be anything such as the color of the skin, shape or texture.
* Track the face - Once the feature is selected, we can track the face by distinguishing the geometrical coordinates of the face and the background.

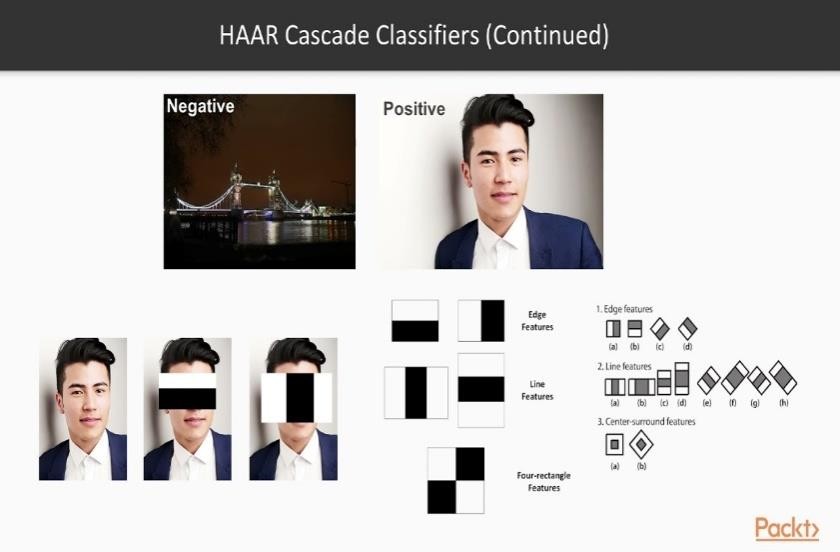


Figure 2.2: Haar Cascade Classification

The algorithm proposed by viola and Jones was one of the best real time face detectors and the factors contributing to the success of the detectors are: the integral image, the Adaboost and the attentional cascade.

* The integral image: The integral image is an algorithm for calculating the sum of intensity values in a rectangular slice of a picture quickly and efficiently. This approach was utilized by Viola and Jones to quickly compute a large number of Haar-like characteristics. The computing advantage allowed the system to be scaled up and saved time.
* The Adaboost: selecting features and training the classifier are done with Adaboost. In this, the weak learner must choose the feature that can easily distinguishes the weighted positive and negative training samples.
* Attentional Cascade: These are used for the integration of very intricate classifiers in a sequential fashion, enabling background sections. It was quickly rejected while computation is focused on probable object. The large majority of negative windows are rejected by the boot T. While allowing almost all positive windows to pass. After then, more complex and hence slower boosted classifiers with higher thresholds are used to reject the much minimal group of troublesome negative windows.

1. *Computer vision system interfacing using MATLAB:*

In this part, we'll look at how to use MATLAB to create an object detection prototype system. By tracking the motion of the identified item, we will employ an image processing technique together with modification in the output pin state of the Arduino board with AT mega 8 controller. A moving item will be detected and tracked by this prototype.

PROPOSED PROTOTYPE SYSTEM:

We employ several hardware and software components in the suggested prototype, such as a camera, Arduino board, AT mega 8 microcontroller, and MATLAB, to recognize and track objects in real time. For picture or video capture, we utilize a camera in the hardware configuration. To recognize and track an item from the provided input, a MATLAB image processing technique is used, and control signals are created and delivered to the Arduino board through serial connection. The position of the detected item is indicated by the state of LEDs linked to the AT mega 8 microcontroller's digital output pin.

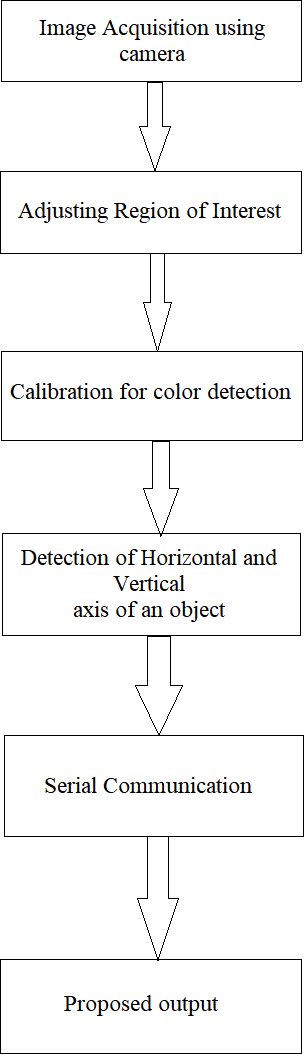


Figure 2.3: Proposed prototype approach

SOFTWARE IMPLEMENTATION**:**

The proposed system makes use of MATLAB as a platform for image processing. A camera can be either a built-in camera on a laptop or a USB camera. The MATLAB command (imaqhwinfo) is used to obtain information about the hardware device that is connected to the computer. Once the system is up and running, we'll have to import the footage from the attached camera. Further algorithms are applied when the video input is transformed into a sequence of frames. The MATLAB command (getsnapshot) is used for this. Once the frame has been set, the image's size is determined and calibrated using the thresholding process. Following the successful completion of calibration, features are taken from the picture, and object detection is performed using these features.

HARDWARE IMPLEMENTATION:

The camera is the most important and initial piece of image collecting equipment. An Arduino board is used in conjunction with a camera, as well as an AT mega 8 CPU to power the Arduino board. The image acquisition toolbox is the ideal tool for modifying the imaging equipment characteristics of a system t can be programmed using Arduino and the serial port can be adjusted precisely. The AT mega 8 is used to interpret the programming of the Arduino board, which is self- designed and self-made suited to the system's requirements. The Converter Adapter Module is used to establish connection between MATLAB and the Arduino board.

1. *Object detection and segmentation using Background Subtraction*

It is a method used in digital image processing and computer vision to convert a digital image into various groups called segment in order to minimize complexity and make processing easier. Image segmentation, to put it simply, is the process of allocating labels to pixels. All picture components or pixels that belong to the same category are assigned a similar level. It is the process of transforming an image's representation into something that is easier to analyze in order to obtain more information in the region of interest in an image, which helps with object scene annotation. To correctly identify the image's content, image segmentation is necessary. Edge detection is a fundamental mechanism for picture segmentation in this scenario and a number of general-purpose algorithms and strategies have been developed over a period of time.

During the segmentation the image is divided into a sequence of non-overlapping connected pieces, so that no two same are connected identical. This is a difficult endeavour because of the abundance and diversity of images and moving objects. Influencing factors include lighting, contrast, and frames. The most frequent segmentation approaches are threshold and template matching process. An

overview of object detection with the help of background subtraction has been discussed below**:**

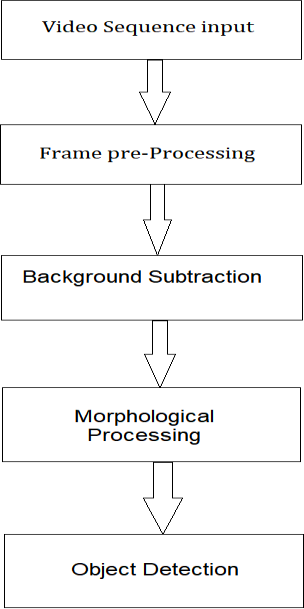


Fig. 2.4 Background subtraction for Object detection

Differentiation of the subject from the stationary is the most significant factor in object identification and picture segmentation. Background subtraction, which works by subtracting the background model, can be used to do this. After the subject has been isolated, noise is removed using a morphological process, and the image is segmented. following method:

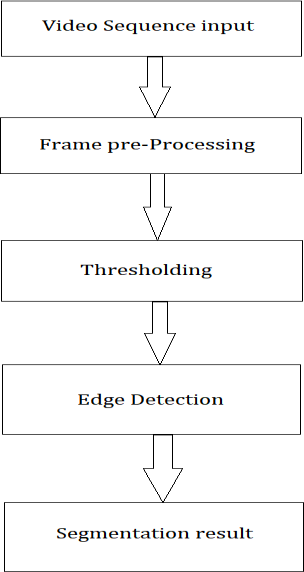


Fig. 2.5 Image Segmentation

Segmentation is the most crucial step in the image processing. The separation of an image into areas or categories that respond to a variety of items or parts of

objects is termed as image segmentation. It may be done in two ways. The image is first converted to binary images, and then elements are segmented using thresholding and edge detection. Each pixel in a picture is categorized into one of many groups. Typically, an effective segmentation is one in which has the following:

1. The grey scale of multivariate values in the same category of pixels is identical, forming a connected zone.
2. The values of neighboring pixels in various categories are varied.
3. *Object tracking*

The application which is used for tracking objects from frame-to-frame studying is driven by deep learning algorithm in which software takes multiple unique identifiers for each of the segments of the image. After that detection of object is done. The various steps to achieve object tracking are as follows:

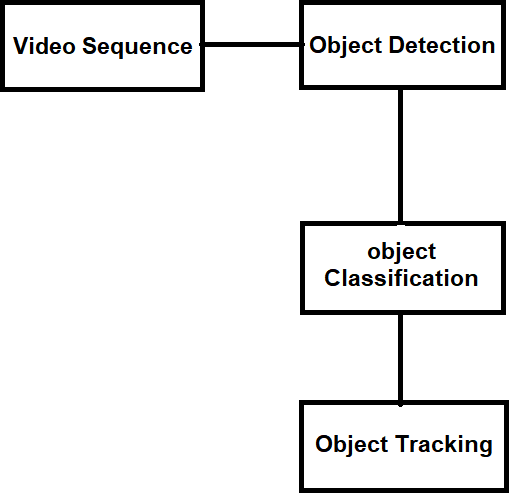


Figure 2.6: Object Tracking

Methods for detecting objects include: This is done either in every frame or when the item appears for the first time in the video. It deals with the removal of motionless background items in order to make room for the mobility of the object of concern.

Object Classification: we classify objects on the basis of their shape and features of the motion region.

Object Tracking: After object detection and categorization, this is the subsequent phase. This technique determines the direction of travel of object for its tracking. In comparison to frame difference and optical flow detection approaches, the background subtraction method is shown to be the simplest way for obtaining comprehensive data about the item. Kernel or contour-based tracking only requires detection when the item first shows on the screen, whereas detection at each frame will require point tracking. Object tracking can be achieved through different methodologies like:

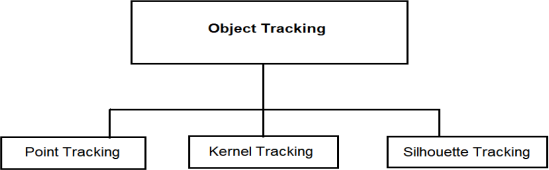


Fig. 2.7 Sub process for object tracking

1. *Object detection*

We disprove the value of deep learning algorithms based on neural networks for detecting object, which is one of the domains where computer vision has made significant progress. Deep learning-based object recognition systems and services are also explored. The use of deep learning algorithms in state-of-the-art object recognition systems is also discussed. Due to strong development in the consumer, robotics, and machine vision industries, the CV market is expected to reach $33.3 billion in 2019.

Convolutional neural networks have the benefit of not needing the development of feature extractors or filters. Instead, they teach themselves from the pixel level to the final object categorization. Deep CNNs have been widely employed in the detection of objects. CNNs are a sort of feed-forward human brain that works on the weight-sharing concept. CNN employs a number of different pooling layers. Depending on the type of research we are doing and how we want our CNN to work it can very such as**:**

* 1. Pulse Coupled Neural Network**:**

Researchers combined monkey visual processing approaches such assumption driven filtering, dependent on state modulation and temporal synchronization. It is related to the fusion picture network. PCNN also improves the accuracy.

Picture processing is used to remove unnecessary information of image for allowing a pattern recognition process to discovering and identifying the used item due to the increased signal-to-noise ratio. It is explained what function these biological events play in information fusion and the image fusion network. Because it accomplishes information connecting at the neural pulse level, the PCNN was chosen as the fusion network's design.

The first route mostly processes color information, whereas the second pathway primarily processes shape and motion. These ideas are utilized to construct an image fusion network that recognizes an item, combines characteristics, and isolates the entity from the rest of the picture in a Simplified Model of the Macaque Device.

The PCNN fusion network takes an original gray-scaled image and many filtered copies of it and produces a single image with the targeted elements brightest and hence clearly visible for object detection. The brightness of the pixels in the input picture that correspond to each neuron are called

feeding inputs. The PCNN's pulse-based connection algorithms segment the original picture using temporal synchronization. The outer PCNNs provide state-dependent modulation signals that are used to direct attention to certain locations.

The essential feature that separates the PCNN from other forms of neural networks in many ways is pulse-based synchronization. This synchronization is what gives the PCNN its picture segmentation ability. The PCNN is a network that connects pixels based on their similarity. In a unique way, the PCNN solves interneuron dependencies. Until the first neuron fires, there are no connecting signals. The brightest spots in a picture induce the earliest firing of their respective neurons. This firing transmits a inked way through the interconnect causing other neurons with similar inputs to activate as well. These pulsed inputs are coupled to the photograph using the center PCNN's linking inputs. These signals are features into neural response of the central PCNN to the object of interest.

PCNN network eliminated 46 percent of incorrect detections whereas only removing 7% of true detections. On the FLIR pictures, the fusion network increased accuracy more than on the mammography images. The network reduced the wrong alarm rate in the FLIR pictures from 8.2 to 0.6 per image, and in mammograms from 1.7 to 0.8 erroneous detections per correct detections.

During the fusion process, the PCNN channel does not add ed true detecting methods to the output; instead, it removes false detections. It also has a powerful computer structure for physiologically based fusion and other physiologically sensed pulse-based phenomena.

* 1. Region based convocational neural network**:**

Underwater exploration is very big because deep-sea exploration is so critical for the development and utilization of deep-sea resources, autonomous underwater operations are becoming more important to avoid the dangerous high- pressure deep-sea environment. The Convolution Neural Network (CNN) is recognized as the fastest detection approach in a range of study topics. Deep CNN has been widely utilized to solve classification issues since then, winning the ILSVRC (Image Net Large Scale Visual Recognition Challenge) and lowering the top 5 error rate to 15.3%.

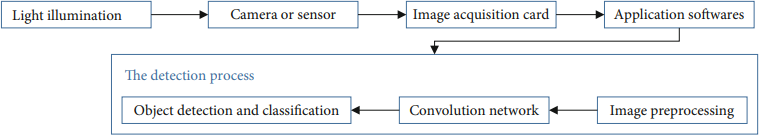


Figure 2.8 CNN base Detection Process

Since the early 2018, Redmon discussed the YOLO v3, which is regarded as fastest detection method, with considerably enhanced accuracy and detection process when compared with the previous approaches. The Convolutional Neutral Network is used to partition pictures into numerous nonoverlapping areas, and feature extraction is utilized for object recognition and categorization. Through this concept, the entire picture information is used for anticipating the bounding boxes while it is also classifying the objects. Mapping is immediately learned by the Convolutional Neural Network.

In a sum error function loss, the loss function form is a critical strategy for balancing the defects in the training process. The bounding box's width and height are swapped out for the expected boxes of various sizes. Dataset enrichment is a term that includes the process of adding more data to a Underwater datasets are difficult to create, aquatic photographs and video are hard to get on the internet, and pixels in the same location often have almost identical backgrounds, therefore the images in the dataset are comparable. The training output model is never effective when applied to different marine regions as a result of these variables. As a result, to make the deep learning model more generally applicable, the dataset must be changed and supplied. Other approaches include rotation, flipping, zooming, and shifting other techniques are used to supplement datasets.

1. *Applications*

Object detection is a technique for using a computer to simulate intelligence. It has a wide spectrum of uses in today's rapidly evolving technological world, including military, medicine, automation, environmental restoration, and so on. Spotting and detecting animals in sterile zones such as industrial areas, order to detect vehicles parked in

restricted areas, and sensing faulty electric wires when images are captured from drone cameras are all examples of day-to-day object detection application Sensors that are used for long-term monitoring generate petabytes of image data in span of just a few hours. These data are geospatially transformed and coupled with other information to create an insight into the current situation. This approach employs object detection to keep an eye on things like people, autos, and suspicious objects in raw image data. The detection not attended baggage is an important application of the object detection. Detecting things on the road, another use of object detection, would be critical in autonomous driving systems. Object detection might be used to detect a driver's fatigue on the highway and prevent an accident.

RCNN was also taken in use for the development of an underwater remotely operated vehicle (ROV) for the purpose of fishing marine objects. The robot is approximately 1 meter long, 0.8meter broad, and 90 kg in weight. The adsorption process is used to gather marine materials. The robot is controlled remotely, and many teams are working to rebuild the ROV to make it semi- autonomous, with the ability to detect and discover objects being the most significant technology. By merging the findings of numerous object detection methods, the PCNN fusion network enhances object detection accuracy. Wide variety of its uses are also evident in the detection of cancer cells and the interpretation of radar images.

Table no 1: Comparative analysis among different object detection approaches:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **S.**  **No** | **Paper title** | **Approach used** | **Advantage** | **Disadvantage** | **Finding** | **S/w and H/w Required** |
| 1 | Deep Learning for Object Detection [1] | Deep Learning | Large data sets are available with the enhancement of the technology. | With the current technology it is quite difficult to manage the large data sets. | State of art object detection systems used for Deep learning | Deep learning frame works and data sets. |
| 2 | Moving object tracking using image processing [2] | Object Trackin g | High quality cameras at minimal cost are available nowadays for better surveillance. | For doing real time detection it becomes quite difficult as high efficiency is not achieved till now. | Image processing converting image into digital inform by performing some operations on it. | Video surveillance technology. |
| 3 | Image Detection using PCCN [3] | PCNN for object detection | Many medical researches nowadays are incorporating new technology so that better results can be obtained to help the patients at early stage. | Practices and theories are still going on and the field being very vast still need work to be done. | Object detection techniques used for improving detection accuracy. | SCUD launcher and flash pods,  Dog filter, Morphological filtered, PCNN fusion network |
| 4 | Deep CNN Method for Underwater Image Processing [4] | RCNN  For Object detectio n | Real time detection done for the underwater Robot. Results explained that the robot achieved underwater working operation. | Dropout layers are not significant in proposed model | CNN method used for solving the illuminated problem. | Deep learning environment, data base for images and different objects. |
| 5 | Object Detection Using Image Processing [5] | Image processing using Haar- like features for object detection. | Very Efficient and fast for face tracking. | If alignment changes the negative result increases. | Tracking object by special features locking. | Open cv python. |
| 6 | Computer vision system interfacing using MATLAB [6] | Algorithm on MATLAB | High compatibility with hardware tools. | Execution is slow compared to the open CV. | Detection and comparison of image by conversion into binary image | MATLAB, Serial communicator, Arduino. |
| 7 | Moving Object Detection and Segmentation for Video Image Processing [7] | Detecting the moving object. | It has ability to segment an image with no well-defined relation between region and pixels. | Lack of clear definition of segmentation. | Separation and background subtraction. | Surveillance system. |

1. CONCLISION

Image processing and, more particularly object identification is a modern-day phenomenon that will do wonders for the globe in the future, as we can see from the ever-increasing technological achievements.

In this work, we have presented a thorough evaluation of object detection and associated technologies. We attempted to cover the fundamentals of object recognition and tracking. Object detection's foundations and applications have been investigated. The first stage in the development of self-driving automobiles and robots is object detection. We explored the function of CNN-based deep learning systems for object detection in this article. Object detection and the domains in which it can be used were discussed. The researchers compared deep learning- based object identification systems. Every strategy has its own set of benefits and drawbacks, but progress is being made at a breakneck pace, indicating the scope and dependability of object detection in the near future.

1. REFRENCES
2. Application of Deep Learning for Object Detection/ International Conference on Computational Intelligence and Data Science (ICCIDS 2018).
3. A SURVEY ON MOVING OBJECT TRACKING USING IMAGE PROCESSING/ 2017

11thInternational Conference on Intelligent Systems and Control (ISCO).

1. Physiologically Motivated Image Fusion for Object Detection using a Pulse Coupled Neural Network/ IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. NO 10, 3MAY 1999.
2. Underwater Image Processing and Object Detection Based on Deep CNN Method/ Hindawi Journal of Sensors Volume 2020, Article ID 6707328
3. Object Detection Using Image Processing/ Moscow Institute of Physics & Technology, Department of Radio Engineering & Cybernetics,2016
4. Interfacing of MATLAB with Arduino for Object Detection Algorithm Implementation using Serial Communication**/** International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 3 Issue 10, October- 2014.
5. Video Image Processing for Moving Object Detection and Segmentation using Background Subtraction**/** 2014 First International Conference on Computational Systems and Communications (ICCSC) | 17-18 December 2014 | Trivandrum
6. Transformer Transforms Salient Object Detection and Camouflaged Object Detection/ JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8, AUGUST 2015