“An overview of Object Detection”

A Project report submitted in partial fulfillment of the requirements of the award of the degree of

## Bachelor of Technology

## In

**Computer Engineering**

by

## Tripti Somani, PCE20CS188

## Shivam Kumar, PCE20CS175

Sahil Tank, PCE20CS162

Shivin Gupta PCE20CS177

Under the guidance of

### Mrs. Sonam Gaur



(Session 2021-22)

# Department of Computer Engineering

Poornima College of Engineering

ISI-6, RIICO Institutional Area, Sitapura, Jaipur –302022

**December, 2021**

**Department Certificate**

This is to certify that **Ms.Tripti Somani, Mr. Shivam Kumar, Mr. Sahil Tank & Mr. Shivin Gupta**, registration no.**PCE20CS188, PCE20CS175, PCE20CS162 & PCE20CS177**, of the **2ND Year** Department of Computer Engineering, has submitted this project report entitled “**An overview of Object Detection**” under the supervision of Dr. / Prof. **Sonam Gour**, working as Assistant Professor in department of Computer Engineering as per the requirements of the Bachelor of Technology program of Poornima College of Engineering, Jaipur.

Dr. Surendra Kumar Yadav Ms. Sonam Gour Head, Dept. of Computer Engineering Coordinator-Project

**CANDIDATE’S DECLARATION**

I hereby declare that the work which is being presented in this project report entitled **“An overview of Object Detection”** in the partial fulfillment for the award of the Degree of Bachelor of Technology in (Computer Engineering), submitted in the Department of Computer Engineering, Poornima College of Engineering, Jaipur, is an authentic record of my own work done during the period from July 2021 to December 2021 under the supervision and guidance of Prof. **Sonam Gour**, working as Assistant Professor in department of Computer Engineering.

We have not submitted the matter embodied in this project report for the award of any other degree.

Dated:

Place: Jaipur

**SUPERVISOR’S CERTIFICATE**

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

(Signature)

Dated: Mrs. Sonam Gour

Place: Jaipur Assistant Professor Department of CS, PCE

# ACKNOWLEDGEMENT

I would like to convey my profound sense of reverence and admiration to my supervisor Mrs. Sonam Gour, Assistant Professor **Department of Computer Engineering, Poornima College of Engineering,** for his intense concern, attention, priceless direction, guidance and encouragement throughout this research work.

I am grateful to **Dr. Mahesh Bundele**, Director of Poornima College of Engineering for his helping attitude with a keen interest in completing this dissertation in time.

I extend my heartiest gratitude to all the teachers, who extended their cooperation to steer the topic towards its successful completion. I am also thankful to non-teaching staff of the department to support in preparation of this dissertation work.

My special heartfelt gratitude goes to **Dr. Surendra Kumar Yadav, HOD,Department of Computer Engineering, Ms. Sonam Gour, Project Coordinator, Department of Computer Engineering, Poornima College of Engineering**, for unvarying support, guidance and motivation during the course of this research.

I would like to express my deep sense of gratitude towards management of Poornima College of Engineering including **Dr. S. M. Seth**, Chairman Emeritus, Poornima Group and former Director NIH, Roorkee, **Shri Shashikant Singhi**, Chairman, Poornima Group, **Mr. M. K. M. Shah**, Director Admin & Finance, Poornima Group and **Ar. Rahul Singhi**, Director Poornima Group for establishment of institute and providing facilities my studies.

I would like to take the opportunity of expressing my thanks to all faculty members of the Department, for their kind support, technical guidance, and inspiration throughout the course.

I am deeply thankful to my parents and all other family members for their blessings and inspiration. At last but not least I would like to give special thanks to God who enabled me to complete my dissertation on time.

**Sahil Tank, Department of Computer Engineering, PCE20CS162**

**Shivam Kumar, Department of Computer Engineering, PCE20CS175**

**Shivin Gupta, Department of Computer Engineering, PCE20CS177**

**Tripti Somani, Department of Computer Engineering, PCE20CS188**

# 

# LIST OF FIGURES

|  |  |  |  |
| --- | --- | --- | --- |
| **S. No.** | **Fig. No.** | **Description** | **Page No.** |
| 1 | 1 | Object Detection as Foremost step in visual recognition activity | 4 |
| 2 | 2 | Basic steps for object detection and tracking | 9 |
| 3 | 3 | Object tracking pathway | 10 |
| 4 | 4 | Haar Cascade Classifier | 14 |
| 5 | 5 | Object Tracking In MATLAB | 15 |
| 6 | 6 | Image Segmentation | 17 |
| 7 | 7 | Mango recognition Block Diagram | 27 |
| 8 | 8 | Contribution of each points to the accumulator space | 28 |
| 9 | 9 | Aerial & satellite surveillance system | 31 |
| 10 | 10 | Qualitative results of considered method over problematic video sequences | 34 |
| 11 | 11 | Result analysis using precision-recall curve | 34 |
| 12 | 12 | Result Analysis using ROC Curve | 34 |
| 13 | 13 | Proposed approach | 38 |

**LIST OF TABLES**

|  |  |  |  |
| --- | --- | --- | --- |
| **Serial Number** | **Figure Number** | **Description** | **Page Number** |
| 1 | 1 | Pooling layers used for object detection | 4 |
| 2 | 2 | Comparative study of object classification methods | 10 |
| 3 | 3 | Comparative study of object detection method | 10 |
| 4 | 4 | Comparison of tracking methodologies | 10 |
| 5 | 5 | Description of data set available | 33 |
| 6 | 6 | Performance Matrices | 33 |
| 7 | 7 | Average Analysis of Yadav, Zhou, Ng & Delp, Jung, Haque, Stauffer, Irani | 35 |
| 8 | 8 | Comparison Table | 36 |

**LIST OF ACRONYMS**

|  |  |  |
| --- | --- | --- |
| **Serial Number** | **ACRONYM** | **FULL FORM** |
| 1 | **CNN** | Convocational neural network |
| 2 | PCNN | Pulse coupled neural network |
| 3 | FLIR | Forward Looking Infrared |
| 4 | MAP | Mean Average Precision |
| 5 | UAV | Unmanned Aerial Vehicle |
| 6 | USB | Universal Serial Bus |
| 7 | HSI | Hue-Saturation-Intensity |
| 8 | RGB | Red Green Blue |
| 9 | IDE | Integrated Development Environment |
| 10 | HOG | Histogram of gradients |
| 11 | CBIR | Contract based image retrieval |
| 12 | AP | Average Precision |

# TABLE OF CONTENT

* Abstract
* Introduction
* Problem Statement &Objective
* Literature Reviews Paper[1-16]
* Comparison Table
* Proposed Approach
* Conclusion
* Future-Scope
* References

# Abstract:

Object detection has been a hot topic in the field of research and development for quite a time now. Object detection is any form of signal processing for which the input is an image such as photographs or frame of video and the output is the detection of a desired object. Object detection stems around the principal application of processing of image data for storage, transmission and representation for autonomous machine perception. In existing system there are many techniques which are available for object detection that we have discussed indetail.

# Introduction:

In today’s modern world every pixel is a data and each frame contains millions of it, but to extract that data and use it for a cause is the real challenge and to overcome this problem we take image processing in use.

Image processing is a technique in which we use digital computers to process a digital image or frame to extract information with help of different algorithms and methodologies.

Image processing helps us to extract minute information and details about a frame which can be taken in use for different operations. It assists us in enhancement, compression and restoration of image. These operations can help us to carry out tasks as simple as character recognition and as complex as identifying cancer cells in humans.

Deep learning technology has become a buzzword nowadays due to the state-of-the-art results obtained in the domain of image classification, object detection, natural language processing, underwater object detection, object tracking etc. The reasons behind popularity of deep learning is availability of large datasets and powerful Graphics Processing both of which requirements have already been satisfied in this current era. Image processing is used to reduce unwanted information from an image with the hope that the improved signal-to-noise ratio will allow a pattern recognition process to detect and possibly identify the desired object. To perform this information fusion, primate vision processing principles are used to design a pulse coupled neural network (PCNN) based image fusion network for the purpose of improved object detection also Object tracking plays a vital role in the field of computer vision. Object tracking algorithms have acquired priority due to the availability of highly sophisticated computers, good quality and inexpensive cameras. To satisfy the requirements of the human vision system and machine recognition has gradually become a hot issue. The paper explains proposed algorithm for object detection using image processing and manipulation of the output pin state of Arduino board with ATmega 8 controller by tracking the motion of the detected object. The object detection algorithm has been developed on MATLAB platform by the combination of several image processing algorithms. Detection and segmentation of moving objects in video streams is an essential process for information extraction in many computer vision applications, including video surveillance, human tracking, traffic monitoring and semantic annotation ofvideos.

# Problem Statement & Objective:

We are planning to develop an Object detection system that will take input from camera and the input frame will be processed using our proposed method and produce a required output. This output will be passed on in the form of voice command which will be achieved by the integration of a voice output unit with our system. This system in the coming future will help visually-aid people to create an artificial vision and will work like “artificial eyes”.

# Literature Review:

Paper 01:

## Application of Deep Learning for Object Detection (ICCIDS , 2018

By: Tripti Somani

## SUMMARY:

* Paper demystifies the role of deep learning techniques based on convolution neural network for object detection.
* Deep learning frameworks and services available for object detection are also enunciated. Deep learning techniques for state-of-the-art object detection systems are assessed in the paper.
* CV market would reach $33.3 billion in 2019 fostering the remarkable growth in the domains of consumer, robotics, and machine vision.
* The reasons behind popularity of deep learning are two folded, viz. large availability of datasets and powerful Graphics Processing Units both requirements have already been satisfied in this current era.
* The beauty of convolution neural networks is that they do not rely on manually created feature extractors or filters. Rather, they train per se (self) from raw pixel level up to final object categories.
* Deep neural architectures handles complex models efficiently than shallow networks. CNNs are less accurate for smaller data but show significant/ record breaking accuracy on the large image datasets. But, CNNs requires large amount of labeled datasets to perform computer vision related tasks (recognition, classification and detection).

1. Object Detection
   * Detecting single instance of class from image is called as **single class object** detection, whereas detecting the classes of all objects present in the image is known as **multi class object** detection.

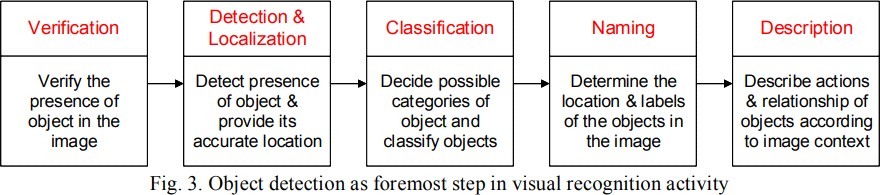
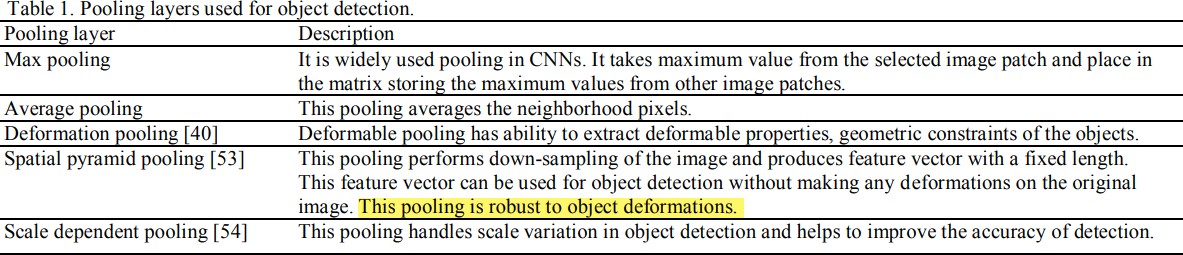


Fig1. Object Detection as Foremost step in visual recognition activity

* + Deep CNNs have been extensively used for object detection. CNN is a type of feed- forward neural network and works on principle of weight sharing. CNN employs different kinds of pooling layers:



1. Application Domains of Object Detection
   * Defense (surveillance), human computer interaction, robotics, transportation, retrieval, etc. Sensors used for persistent surveillance generate peta-bytes of image data in few hours. These data are reduced to geospatial data and integrated with other data to get clear notion of current scenario. This process involves object detection to track entities like people, vehicles and suspicious objects from the raw imagery data.

Spottinganddetectingthewildanimalsintheterritoryofsterilezoneslikeindustrial area, detecting the vehicles parked in restricted areas are also some applications of object detection. Detecting the unattended baggage is very crucial application of object detection.

* + For autonomous driving, detecting objects on the road would play important role. Detection of faulty electric wires when the image is captured from drone cameras is also application of object detection. Detecting the drivers’ drowsiness on the highway in order to avoid accident may be achieved by object detection.

### Object detection analytics can be performed offline, online or near real time.

Paper 02:

## Physiologically Motivated Image Fusion for Object Detection using a Pulse Coupled Neural Network(IEEE MAY 1999)

### By: Tripti Somani

**SUMMARY:**

* + This paper presents the first physiologically motivated pulse coupled neural network (PCNN)-based image fusion network for object detection.
  + Primate vision processing principles, such as expectation driven filtering, state dependent modulation, temporal synchronization, and multiple processing paths are applied to create a physiologically motivated image fusion network.
  + PCNN’s are used to fuse the results of several object detection techniques to improve object detection accuracy.
  + Image processing is used to reduce unwanted information from an image with the hope that the improved signal-to-noise ratio will allow a pattern recognition process to detect and possibly identify the desired object.

To perform this information fusion, primate vision processing principles are used to design a pulse coupled neural network (PCNN) based image fusion network for the purpose of improved object detection. The role these biological phenomena perform in information fusion and in the image fusion network is discussed. The PCNN is chosen as architecture for the fusion network because it performs information linking at the neuronal pulse level.

A BIOLOGICAL FOUNDATION FOR A FUSION NETWORK

* + The parvocellular pathway and the magno-cellular pathway & the former pathway predominantly processes color information, and the later processes form and motion.

1. A Simplified Model of the Primate Vision System: These theories are used to design an image fusion network that segments an object, combines features, and isolates the object from the rest of the image.
2. Temporal Synchronization Provides Object Segmentation and Fusion: Visual segments that are related in some fashion will synchronize and pulse in unison. These synchronized segments represent objects, or segments of objects within a visual scene. Through this synchronization, the visual image is represented as an ensemble of synchronously pulsing objects. Segment an image and fuse features into those segments. For object detection and recognition to be performed, a method is needed to select and extract a particular segment from the resulting image. This means the PCNN fusion network needs a method of focusing attention on individual groups of synchronously pulsing neurons.
3. State Dependent Modulation Provides Focus of Attention: State dependent signals are believed to be the stimulus that causes this preferential treatment. This phenomenon is called state dependent modulation and is a method for one area of processing to superimpose its findings, or expectations on another area. The modulator effect of state dependent modulations is believed to focus attention by elevating the perception of objects of interest effectively suppressing unneeded information in a visual scene. The PCNN fusion network uses this biological principle to focus attention on objects that best fit the criteria of a desired object.

A PCNN IMAGE FUSION NETWORK FOR OBJECT DETECTION

A. The PCNN Fusion Network: To perform object detection, the PCNN fusion network takes an original and several filtered versions of a gray-scaled image and outputs a single image in which the desired objects are the brightest and thus easily detected. Each neuron receives feeding inputs which are the intensity of the corresponding pixels in the input image. The pulse based linking mechanisms of the PCNN use temporal synchronization to segment the original image. The outer

PCNN’s provide state dependent modulation signals used to focus attention on segments of interest.

D. The Pulse Coupled Neural Network: The heart of the fusion network is the PCNN.

1. Pulse Coupling Performs Temporal Synchronization: Pulse based synchronization is the key characteristic that distinguishes the PCNN from other types of neural networks. The image segmentation property of the PCNN comes from this synchronization. The PCNN links pixels based on similarity. The actual PCNN solves the interneuron dependencies in a unique way. No linking signals are present until the first neuron fires. The brightest points within an image cause their corresponding neurons to fire first. This firing initiates a linking signal (linking wave) which travels through the multiplicative linking interconnects causing other neurons with similar inputs to fire. The actual PCNN solves the interneuron dependencies in a unique way. No linking signals are present until the first neuron fires. The brightest points within an image cause their corresponding neurons to fire first. This firing initiate a linking signal (linking wave) which travels through the multiplicative linking interconnects causing other neurons with similar inputs to fire.
2. Pulse-Based Multiplicative Linking Performs State Dependent Modulation: The neurons whose inputs most match the desired target would have the greatest modulator input, thus having the highest frequency output. This increased output effectively separates the neurons from the rest of the image. Segments with a greater number of desired features present will be more active than other segments; therefore the most active segments are those that fulfill more of the target criteria. These segments are easily separable from the rest of the image.
3. How Information Is Fused

This temporal synchronization groups the image pixels into individual, disjoint segmented regions (objects) that pulse in different time steps. The outer PCNN’s convert the filtered images into pulsed signals for use as state dependent modulation signals. These pulsed signals are linked to the original image using the center PCNN’s linking inputs. These signals both fuse features into its associated segment, and modulate the center PCNN’s neuronal response to the object of interest.

1. The PCNN Produces a Time Signal as an Added Benefit

PCNN forms a time signal that can be used as a translation, rotation, scale, distortion, and intensity invariant signature for each object in the image. The time signal was not utilized in this project, but could have been. The next stage of this research will use the signature property of the time signal to increase object detection accuracy

IV. DETECTION RESULTS ON CANCER AND RADAR IMAGES

One hundred FLIR images, from aircraft training runs, were used to calibrate and test the object detection capability of the PCNN fusion network. **When compared to the best filter result, the PCNN network removed 46 percent of the false detections while removing only 7% of the true detections**

The fusion network provided a greater accuracy increase on the FLIR images than on the mammogram images. **The network reduced the false alarm rate from 8.2 to 0.6 false alarms per image in the FLIR images and from 1.7 to 0.8 false detections per true detections in the mammograms**.

In the fusion process, the PCNN network does not add true detections to the output, but instead removes false detections. The PCNN also provides good computer architecture for implementing physiologically-based fusion and other pulse-based physiologically observed phenomenon.

Paper 03:

## A SURVEY ON MOVING OBJECT TRACKING USING IMAGEPROCESSING ( ISCO 2017)

### By: Tripti Somani

**SUMMARY:**

This paper reviews the various challenges and aspects of detection and tracking of objects.

1) Object Detection methods: This is done either in each and every frame or when the object first appears in the video. It handles with the elimination of stationary background objects form the oving of object of interest.

1. Frame Difference Method: The difference between two consecutive images is calculated, which further identifies the moving object.

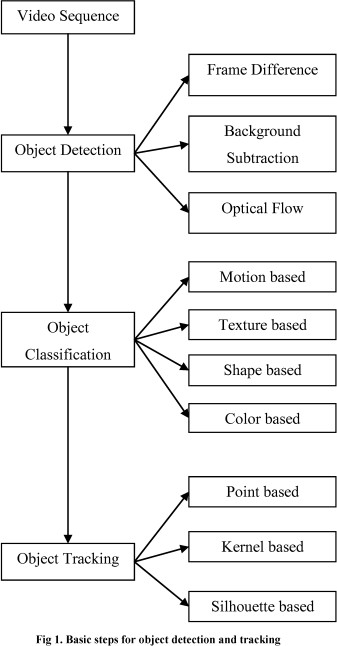
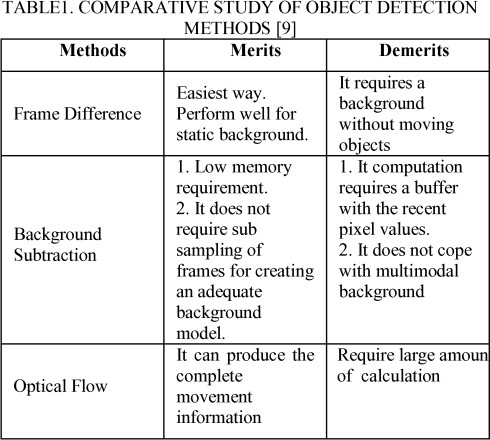
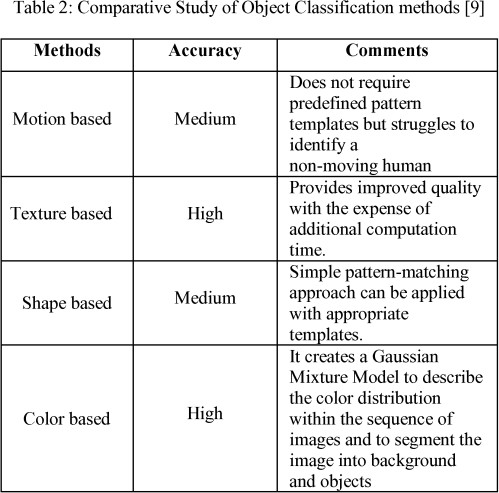


Fig2. Basic steps for object detection and tracking

1. Background Subtraction method: Background modeling is the first step of background subtraction. The two common approach available in background subtraction are:
2. Recursive Algorithm
3. Non-Recursive Algorithm
4. Optical Flow: The complete detection and movement information of object from the background can be obtained from this method.
5. Methods of Object Classification: classification of objects is done based on their shape features of the motion region.
6. Methods of Object Tracking: It is a technique used to track and also the travelling direction of objects



3

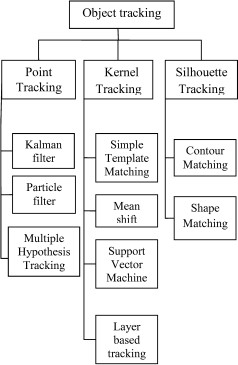
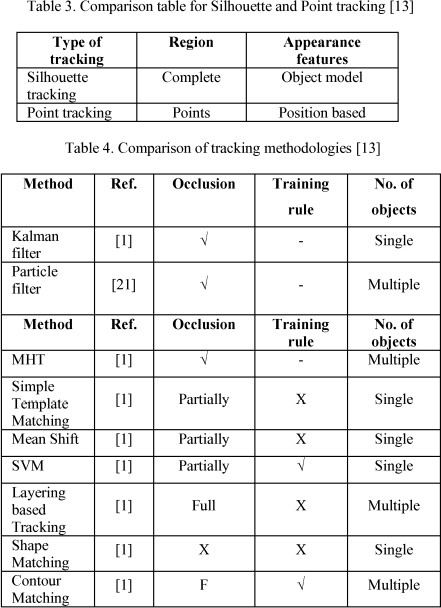
 

Fig3. Object tracking pathway

Paper 04:

## Underwater Image Processing and Object Detection Based on Deep CNN Method ( HINDAWI 2020)

### By: Tripti Somani

**SUMMARY:**

Due to the importance of underwater exploration in the development and utilization of deep-sea resources, underwater autonomous operation is more and more important to avoid the dangerous high-pressure deep-sea environment. After the image processing, a deep CNN method is proposed to perform the underwater detection and classification, according to the characteristics of underwater vision, two improved schemes are applied to modify the deep CNN structure. The detection speed is about 50 FPS (Frames per Second), and MAP (mean Average Precision) is about 90%. The results show that the methods are reasonable and effective, but the common

Weakness is that the processing is very time consumable, and it is difficult to achieve real-time detection.

The Convolution Neural Network (CNN) is recognized as the fastest detection method by many ways in different research fields; Krizhevskyet al. applied CNN method to deal with classification problem winning the champion of ILSVRC (Image Net Large Scale Visual Recognition Challenge), which reduce the top 5 error rate to 15.3%, from then on deep CNN has been widely applied. Earlier of 2018, Redmond put forward the YOLO v3, which is generally recognized as the fastest detection method, and the accuracy and the detection speed are greatly improved compared with the other methods.

Convolution Neutral Network is used to divide images into multiple non-overlapping regions; the basis of object detection and classification is based on feature extraction. By analyzing the current image processing algorithms, enhancement algorithms for underwater images are proposed in this paper. The Convolution Neural Network directly learns from an end-to-end mapping between dark and bright images. Low-light image enhancement in this paper is regarded as a machine learning problem. A weakly illuminated image is input, and a 32 ∗ 6 ∗ 6 convolution layer is applied to change the image into 32 channels; the 3-D view figure means multilayer’sfeaturemap, andthenc18∗6∗6and8∗1∗1convolutionlayersareaddedinthe network; the output is a one channel feature map. In order to improve the detection accuracy, the whole image information is used to predict the bounding boxes of the targets and classify the objects at the same time; through this proposal, the end-to-end real time targets detection can be realized.

Loss Function: In the process of training, the loss function form is a key technique; for the method proposed in this paper, a sum squared error loss is used to balance the errors. For the boxes in different size prediction, the width and height of the bounding box are substituted by the square root value; thus, the smaller box has a relatively large value offset to make the prediction more effective. In order to solve the phenomenon of gradient dispersion or explosion of the network, the better proposal is to change the layer-by-layer training of deep neural network to step-by-step training. Auxiliary loss is only used for training, not for the prediction process.

Dataset Augmentation. Underwater dataset is difficult to prepare, the underwater images and video are not easy to obtain on the internet, and for underwater images, the background is almost

the same in the same area, so the images in the dataset are similar, because of these factors the training output model is always not effective to be used in other sea areas.

The method proposed in this paper is going to be used on an underwater remote operated vehicle (ROV) for fishing marine products. The robot is about 1 m long, 0.8 meters wide, and weighs 90 kg. The method of collecting marine products is adsorption type; the design and real robot are shown in Figure 10. The robot is remote operated; our team is going to reconstruct the ROV to semiautonomous, so the key technology is how to detect and locate the objects. 5.1. Detection Comparison. The GPU used in these computations is NVIDIA GTX 1080ti, and the total number of images is 30000, which are labeled one by one artificially. And in deep learning, 8520 images are used for training, 8530 for validation, and 12950 for test. In object detection, Precision, Recall, and Mean Average Value are commonly used to assess the accuracy.

The underwater vision is in low quality, and the objects are always overlapped and shaded, so the original YOLO V3 method is not very effective for underwater detection; two methods are proposed to deal with these problems. Through detection results comparison with the other methods, the scheme 2 can give a better detection. The trained model is used to assist the ROV to detect underwater objects; although some of the objects are missed, the effectiveness and capability of the proposed method are obviously verified by the qualitative and quantitative evaluation results. The proposed method is suitable for our underwater robot to detect the objects, which is not better than the typical methods for the other dataset. And dropout layers and other technologies are not significant in this model; the reconstruction of the network by using a more complicated algorithm would be more effective.

Paper 05:

## Object Detection Using Image Processing (Moscow Institute of Physics &Technology, Department of Radio Engineering & Cybernetics,2016)

### By: Shivam Kumar

**SUMMARY:**

The aim of this paper is to is to develop an Open CV-Python code using Haar Cascade algorithm for object and face detection at a budget with reduced human involvement.

An Unmanned Aerial Vehicle (UAV) is an aircraft with no pilot on board. UAVs are currently used for a number of missions, including reconnaissance and attack roles. The FAA has adopted the acronym UAS (Unmanned Aircraft System) to reflect the fact that these complex systems include ground stations and other elements besides the actual air vehicles.

Using images and video to detect, classify, and track objects or events in order to "understand" a real-world scene is the field of computer vision. Object detection determines the presence of an object and/or its scope, and locations in the image. Computer vision powers applications like image search, robot navigation, medical image analysis, photo management and many more.

Manual intervention of a camera and joy stick will reduce the man work time and complexity of the work. This project will be useful in replacing the laser sensor and servile the location using cheaper systems. UAV is a very expensive vehicle which cannot be lost under blunders of non- detected objects and unprocessed faces so this project aims in compensating for such situations. Image processing is the process of converting an image into digital form and performing some operations on it, in order to get an enhanced image or to extract some useful information from it. It is a type of signal dispensation in which input is image, like video frame or photograph and output may be image or characteristics associated with that image.

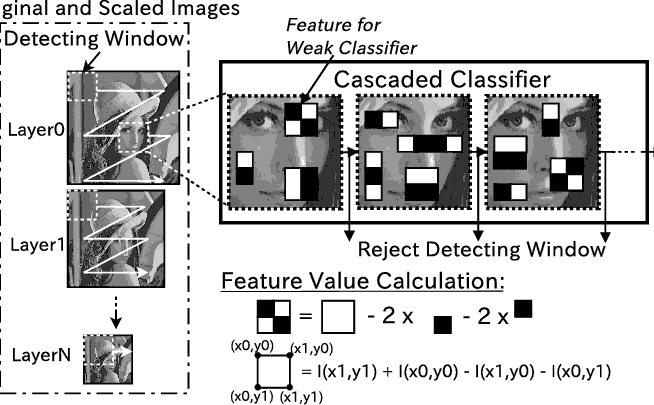


Fig 4: Haar Cascade Classifier

In this methodology we use Cascade Object Detector to detect the location of a face in a video frame. By default, the detector is configured to detect faces, but it can be configured for other objects. we choose a feature that is unique to the object and remains invariant even when the object moves and once the object in the video is identified, as a position which occupied in the output area in term of geometric coordinators, we can distinguish between the real face shape and its correspondent background.

Paper 06:

## Interfacing of MATLAB with Arduino for Object Detection Algorithm Implementation using Serial Communication (IJERT, 2014)

### By: Shivam Kumar

**SUMMARY:**

The aim of this paper is to propose a prototype model which both detects and tracks an object with distinct features and generates and sends a control signal to the hardware according to the position of the object. Digital computers open up the chances of using images and video frames as an input signal of the signal processing. The aim of this paper is to propose a prototype model which both detects and tracks an object with distinct features. It generates and sends a control signal to the hardware according to the position of the object.

Monitoring the motion of an object can be done by identifying and tracking distinct feature of the moving object. Here, a hardware set up is required which can be used as an image acquisition hardware such as camera. This camera is connected to a computer. An algorithm for object detection and tracking is written in MATLAB. This program detects the object in real time.

The proposed system uses MATLAB as a platform on which image processing algorithm has been developed and tested. Image acquisition toolbox of MATLAB can also be used for image acquisition purpose. Next part of the algorithm is to do the calibration for detecting the color of an object. After calibration, if value of pixel lies in that range, then that pixel is set white, otherwise it will be black.

Communication of object detection algorithm with Arduino board is done through serial data transfer. MATLAB function for serial port access is shown below. To access the serial port of a computer using MATLAB, couple of lines of coding should be done. A program has to be booted on ATmega 8 using Arduino and an input port as well as LEDs connected at digital output pins as an output.

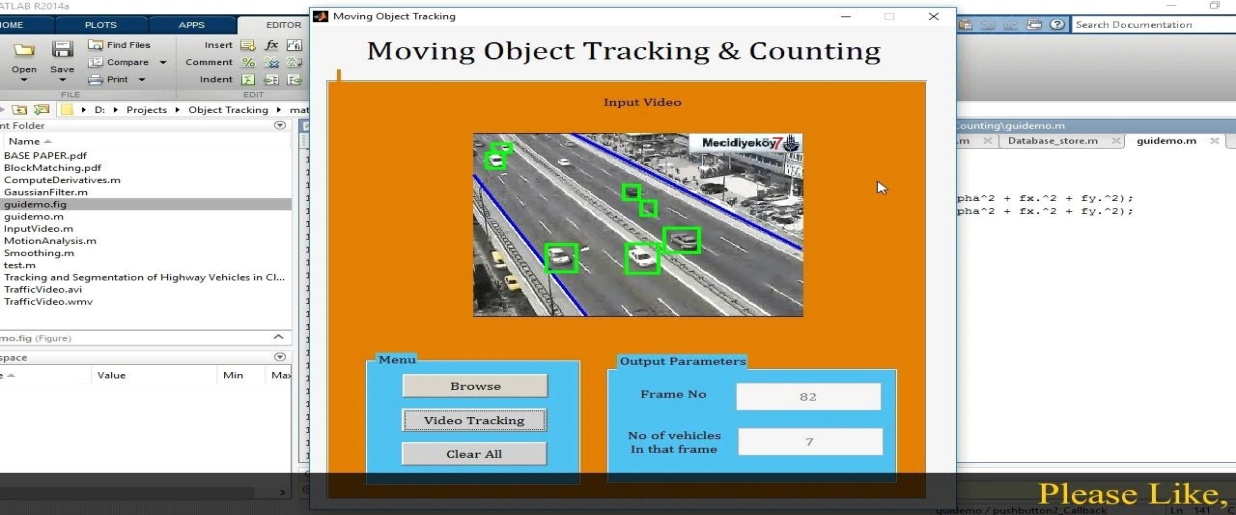


Fig 5: Object Tracking In MATLAB

Proper selection of a hardware is important for the effective working of the system. Camera is the first and the most important hardware used for image acquisition purpose an Arduino board is used (these are self-designed according to the requirements) which is power edby ATmega 8 microcontroller which can be programmed using Arduino IDE. The next thing we need is an Image acquisition toolbox, it is the best tool for the adjustment of parameters of an imaging device used in the system. A converter adapter module is also used, PL-2303HX is a convenient solution for the connection between RS- 232 like full duplex asynchronous serial device and Universal Serial Bus (USB) capable host.

Paper 07:

## Video Image Processing for Moving Object Detection and Segmentation usingBackground Subtraction (ICCSC, 2014)

**SUMMARY:**

### By: Shivam Kumar

Surveillance system uses cameras to monitor the activities of targets (human, vehicle, etc.) in a scene. One of the main problems is the presence of noise. There are three conventional approaches to moving object detection: background subtraction, temporal differencing and optical flow and Image segmentation is the partitioning of an image in objects of interest. Many computer vision methods have been developed for analyzing image motion. In this work, an algorithm for detection and segmentation of objects in the video frames is presented. The algorithm is based on object detection from background and segmentation using thresholding and edge detection.

The method is based on using background subtraction algorithm for separating moving objects from their background. Background subtraction finds moving objects by subtracting background model from input image. Background Subtraction Conventionally, assuming that the background is stationary, then the moving object can be determined by taking the difference between the background image and the input image. The operation shows the outline of the cell quite nicely, but there are still holes in the interior of the object. If the area of the holes is greater than 40% of the total area then the algorithm will combine this area with the total area enclosed by the boundary. For grey video stream, HSI (Hue-Saturation-Intensity) color space background model is used. After we are done with this, we move on to the next part that is image segmentation.

Segmentation is often the critical step in image analysis, Image segmentation is the division of an image. Every pixel in an image is/are located to one of a number of these categories. Segmentation is typically one in which. Pixels in the same category have similar grayscale values and form a connected region.

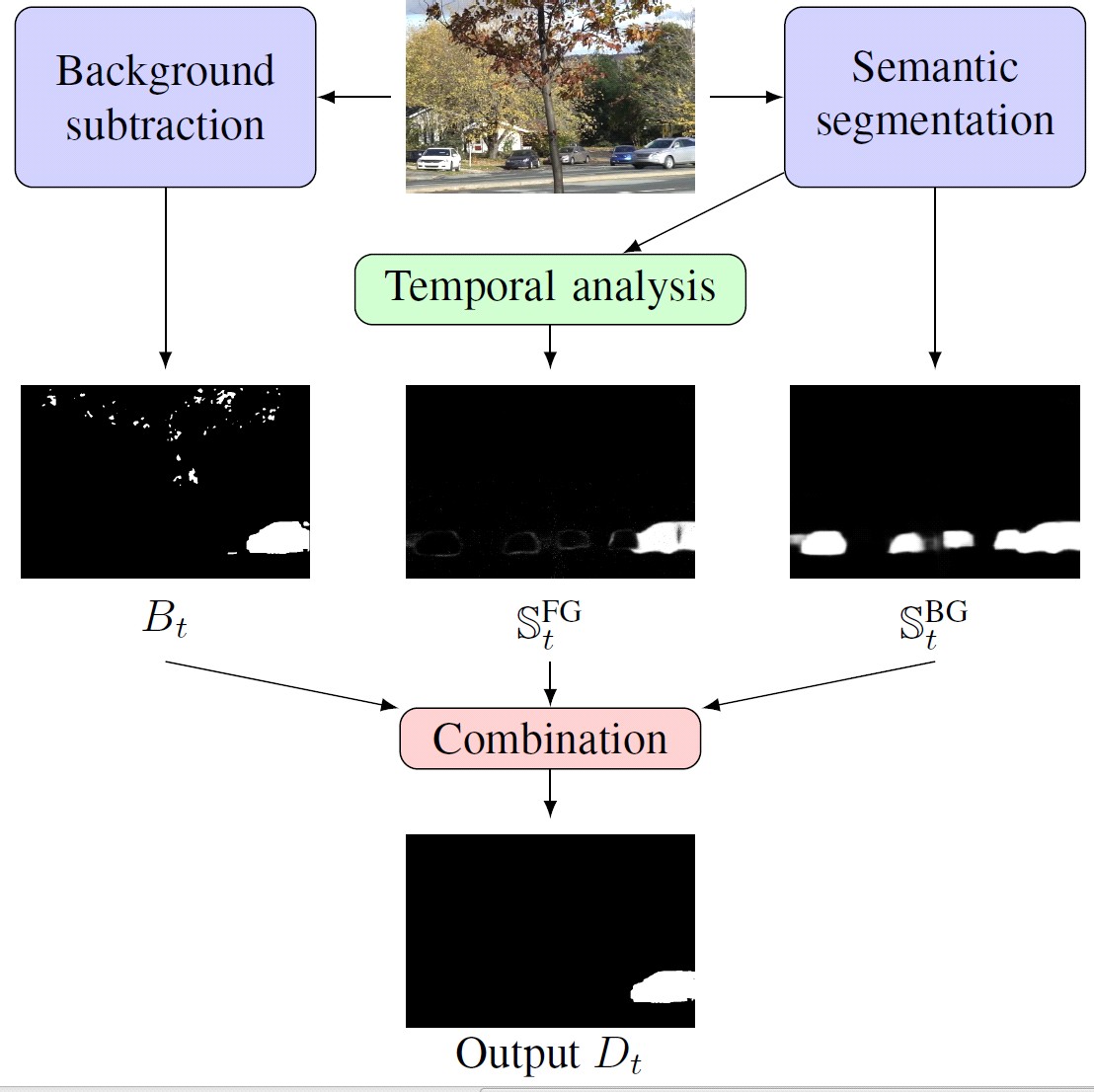


Fig 6: Image Segmentation

Image segmentation is the step at which we move from considering each pixel as a unit of observation to working with objects (or parts of objects) in an image. There are three general approaches to segmentation, termed thresholding, edge-based and region-based methods. The next important step in this process is the edge detection.

* Edges are significant local changes of intensity in an image. They typically occur on the boundary between two different regions in a digital image. Edges are detected to identify discontinuities in the image. There are various edge detectors that are used to segment the images. The Canny Operator is a sort of new edge detection operator. It has good performance of detecting edge.

Paper 08:

### Transformer Transforms Salient Object Detection and Camouflaged Object Detection(JOURNAL OF LATEX, 2015)

**By: Shivam Kumar**

**SUMMARY:**

This paper aims to localize the regions of an image that attract human attention. To, reduce the labeling effort, several weakly supervised salient object detection models have been proposed to learn saliency with image-level supervision, scribble supervision or learn saliency directly from noisy labeling.

Before the deep learning revolution, conventional salient object detection models used handcrafted features, which define saliency as the contrast between each pixel (or super-pixel) and the other pixels (or super-pixels).

In this way, the receptive fields of the conventional handcrafted feature-based models are the entire image, which is the global context, where saliency of the pixel colored in yellow depends on all the other pixels. The deep convolution neural network (CNN) based salient object detection models achieve significant performance improvement compared with those handcrafted-feature based techniques with more sophisticated features extracted from the deep network.

The conventional deep CNN based SOD network usually includes two main parts: an encoder to extract different levels of features, and a decoder to aggregate features from different levels of thenetworkforfinerprediction.Theencoderpartisadoptedfromthetrainedbackbonenetwork

on Image Net, e.g., VGG, ResNet and the most effort for SOD models have been put on designing an effective decoder for feature aggregation. The main issue with this backbone networks are that the larger receptive field is obtained with the loss of structure information as a sacrifice. In this way, a network with a larger receptive field without losing fine-grained information can be beneficial for context-based tasks e.g., salient object detection, to achieve effective context modeling. Inspired by and the accurate structure modeling ability of transformer, we present a “unified” transformer backbone-based SOD network to achieve fully-supervised RGB image- based SOD, RGB-D image pair-based SOD and weakly-supervised RGB image-based SOD with scribble annotation, leading to three new benchmark models.

We also observe that the positional encodings of vision transformer are less effective in modeling the accurate “spatial” information for dense prediction tasks. Then we investigate deep supervision and difficulty-aware learning within the transformer.

Our main contributions are:

1. We introduce a unified transformer backbone-based network for three static image-based salient object detection tasks, and one RGB image based camouflaged object detection task. We discover the superior performance of the transformer backbone for accurate structure modeling, which makes it powerful in learning from weak annotations;
2. We investigate two strategies, namely deep supervision and difficulty-aware learning, and illustrate the effectiveness of them for transformer backbone-based frameworks to generate stronger spatial supervision;
3. We compare feature of CNN backbones and transformer backbones and find that the superior performance of transformer backbones mostly lies in the accurate structure and semantic information encoding with the long-range dependency modeling mechanism.

Paper 09:

### Applications of Object Detection Systems (IRJET 2019)

**By: Sahil Tank**

**SUMMARY-**

The applications of ODS have been broken into multiple industries with use cases ranging from personal security to productivity in the workplace. Some of the current and future applications include:-

* + Optical Character Recognition

Optical character reader, often abbreviated as OCR, is the mechanical or electronic conversion of images of typed, handwritten or printed text into machine-encoded text whether from a scanned document, a photo of a document or a scene-photo.

* + Self-Driving Cars

In order for a self-driving car to decide what to do in next step whether accelerate, apply brakes or turn, it needs to know where all the objects are around the car and what those objects are & that requires object detection.

* + Tracking Objects

Object detection system is also used in tracking the objects, for example tracking a ball during a football match, tracking movement of a cricket bat, tracking a person in a video.

* + Facial Detection

Face Recognition is widely used in computer vision task. It’s also widely being used for mobile photography in detecting smiles of the subjects. Another use is for extracting faces off of videos and plaster them on different videos a.k.a. Deep faking.

* + Medical Imaging

Accurate, robust and fast tracking of deformable anatomical objects such as the heart, is a crucial task in medical image analysis.

* Object Recognition as Image Search

By Recognizing the objects in the images ,combining each object in the image and passing detected objects label in the URL we can make the object detection system as image search.

* + Manufacturing Industry

Object detection is also used in industrial processes to identify products. In the process of Quality Management, sorting, assembly line Object detection is a part of the process.

* + Robotics

Reliable object detection and recognition necessary for assistive robots to process visual data in real-time so that they can react adequately to the changes in their environment.

* + Object Counting

Object detection can be also used for people counting, it is used for analyzing store performance or crowd statistics during festivals.

* + Automatic Target Recognition

Automatic target recognition (ATR) is the ability for an algorithm or device to recognize targets or other objects based on data obtained from sensors.

* + Automated CCTV

A very important part of Surveillance is comprised of the abilities of the CCTV cameras in question to both detect and track activities.

* + Automatic Image Annotation

It’s also known as automatic image tagging or linguistic indexing & it is the process by which a computer system automatically assigns metadata in the form of captioning or keywords to a digital image.

* + Digital Watermarking

A digital watermark is a kind of marker covertly embedded in a noise-tolerant signal such as audio, video or image data. "Watermarking" is the process of hiding digital information in a carrier signal.

Paper 10:

### Object Detection Using Gabor Filters (Pergamum 1997)

**By: Sahil Tank**

**SUMMARY-**

This paper pertains to the detection of objects located in complex backgrounds. A feature-based segmentation approach to the object detection problem is pursued, where the features are computed over multiple spatial orientations and frequencies. The method proceeds as follows: a given image is passed through a bank of even-symmetric Gabor filters. A selection of these filtered images is made and each (selected) filtered image is subjected to a nonlinear (sigmoid like) transformation. Then, a measure of texture energy is computed in a window around each

Transformed image pixel. The texture energy (“Gabor features”) and their spatial locations are inputted to a squared-error clustering algorithm. This clustering algorithm yields a segmentation of the original image—it assigns to each pixel in the image a cluster label that identifies the amount of mean local energy the pixel possesses across different spatial orientations and frequencies. The method is applied to a number of visual and infrared images, each one of which contains one or more objects. The region corresponding to the object is usually segmented correctly, and a unique signature of “Gabor features” is typically associated with the segment containing the object(s) of interest. Experimental results are provided to illustrate the usefulness of this object detection method in a number of problem domains. These problems arise in IVHS, military reconnaissance, fingerprint analysis, and image database query.

Multi-Channel Filtering and Segmentation works by using Selecting a relatively large bank of such Gabor filters & results in a nearly uniform coverage of the spatial- frequency domain

Paper 11:

### Rapid Object Detection using a Boosted Cascade of Simple Features (2004)

**By: Sahil Tank**

**SUMMARY-**

This paper describes a machine learning approach for visual object detection which is capable of processing images extremely rapidly and achieving high detection rates. This work is distinguished by three key contributions. The first is the introduction of a new image representation called the “Integral linage” which allows the features used by our detector to be computed very quickly. The second is a learning algorithm, based on AdaBoost, which selects a small number of critical visual features from a larger set and yields extremely efficient classifiers. The third contribution is method for combining increasingly more complex classifjers in a “cascade” which allows background regions of the image to be quickly discarded while spending more computation on promising object-like regions.

In other face detection systems, auxiliary information, such as image differences in video sequences, or pixel color in color images, have been used to achieve high frame rates. Our system achieves high frame rates working only with the information present in a single grey scale image. These alternative sources of information can also be integrated with our system to achieve even higher frame rates.

Our object detection procedure classifies images based on the value of simple features. There are many motivations for using features rather than the pixels directly. The most common reason is that features can act to encode ad-hoc domain knowledge that is difficult to learn using a finite quantity of training data. For this system there is also a second critical motivation for features: the feature-based system operates much faster than a pixel-based system. The simple features used are reminiscent of Haar basis functions which have been used by Papa georgiouetal. Given a feature set and a training set of positive and negative images, any number of machine learning approaches could be used to learn a classification function. In our system a variant of AdaBoost is used both to select a small set of features and train the classifier. In its original form, the AdaBoost learning algorithm is used to boost the classification performance of a simple (sometimes called weak) learning algorithm.

All example sub-windows used for training were variance normalized to minimize the effect of different lighting conditions. Normalization is therefore necessary during detection as well. Since the final detector is insensitive to small changes in translation and scale, multiple detections will

usually occur around each face in a scanned image. The same is often true of some types of false positives. In practice it often makes sense to return one final detection per face. Toward this end it is useful to post process the detected sub-windows in order to combine overlapping detections into a single detection. In these experiments detections are combined in a very simple fashion.

The set of detections are first partitioned into disjoint subsets. Two detections are in the same subset if their bounding regions overlap. Each partition yields a single final detection. The corners of the final bounding region are the average of the corners of all detections in the set.

Paper 12:

### Object Detection Using YOLO Network (ITMEC 2018)

**By: Sahil Tank**

**SUMMARY-**

There are many problems with images in real-world shooting such as noise, blurring and rotating jitter, etc. These problems have an important impact on object detection. Using traffic signs as an example, we established image degradation models which are based on YOLO network and combined traditional image processing methods to simulate the problems existing in real-world shooting. After establishing the different degradation models, we compared the effects of different degradation models on object detection. We used the YOLO network to train a robust model to improve the average precision (AP) of traffic signs detection in real scenes.

In this paper, we simulated different degenerative processes of images for analysis and research. Firstly, we established the models of degraded images. We mainly used mathematical models to generate degraded images which are based on standard data sets. Then, we used these models to train the network to adapt to the complex real-world environment. Finally, we improved the ability of the model to generalize complex images. We took the traffic signs as the research object and used the YOLO neural network to analyze.

The YOLO neural network integrates the candidate boxes extraction, feature extraction and objects classification methods into a neural network. The YOLO neural network directly extracts candidate boxes from images and objects are detected through the entire image features. Traffic sign detection refers to applying the candidate box extraction technique to the input image to determine whether it contains traffic signs and output their location. The data set we used in this paper for training and evaluation comes from Image Net. None of the images are tagged and they are independent from the ones used in pre-training. We selected 1,652 images containing traffic signs as the data set. Then, 1,318 images were selected which form the standard training set, and the remaining 334 images were used as the test set. All traffic signs in the images are labeled. In the YOLO network, the images are divided into S×S grids. Candidate boxes are equally distributed on the X-axis and Y-axis. The candidate boxes have object detection and predict the

confidence of the existence of the object in each candidate box. Confidence reflects whether the images include the object or not, as well as the accuracy of the object’s position.

We evaluate the performance of a single degenerate model. Next, we select 80% of the data in the standard set which contains all images that went through the general degradation processing. We feed these images to the network to train the general degenerative model, while the remaining 20% of data are used for testing. Finally, we compare the general degenerative model with the standard model. We find the average precision (AP) of the general degenerative model (Ma) is better than the standard model of all test sets. We believe that the network which is trained with degraded images learns more features and can cope with more complex scenes. The model has better generalization ability and higher robustness.

Paper 13:

**Study of Object Detection Method and App. On Digital Images**

**By: Shivin Gupta**

**SUMMARY-**

Object Detection –

It is simply called as detecting objects in an image; it is used in the vehicle navigation etc. It detects objects such as – humans, buildings cars etc. Area of application -face detection, faces recognition and video object detection. Some of the applications are, tracking motion of the ball, tracking ball during the match, tracking person in a video.

Features –

1. Color- representing colors for detection.
2. Histogram of gradients – (HOG) widely used in the human detection. Movement does not affect the detection phase.
3. Edges- edges of the object through color features technique
4. Optical Flow – Motion Segment, done through the displacement vector

Challenges-

1. Lighting – This may delay the detection as lighting plays an important role
2. Scaling Method- the barrier of the scaling.

Machine Learning-

It’s the system derived from the --pattern recognition and computational learning based on artificial intelligence. Prediction based on data sets and algos.

1. Absence of the human expertise.
2. Unable to understand the human expertise such as speech recognition.
3. The solution method changes with the time.
4. Adaptations case based on biometric
5. Supervised ML- Already defined , relation b/w o/p and i/p
6. Unsupervised ML- the algo for the relation of i/p and o/p is unsupervised

Method of Object Detection-

1. Template Based – In this method, the small parts of the image can be recognized using the template image. This technique is also called as the template matching. The relation between the template image and the real image are detected through the geometrical parameters. The data image for the template matching use different iterations for the geometrical parameters.The geometrical parameters with the search images S where represent the coordinates of each pixel in the search image. The method is spatial filtering and the template is filter mask.
2. Part Based Object Detection- The collection of the deformable configuration for the representation of an object. Each part of the models is separately arranged with the deformed configuration and that are represented by the connections between the pairs of the parts.
3. Region Based – transforming the i/p image into directed graph by rules determined. Graph info. Gives idea about the object in the image.
4. Contour Based OD - The image data base determines the various types of the objects that are stored by single prototype images. The cameras are used where the robots are located and identified by the passing objects. This process based on the two phases. The first phase is the location of the individual objects and describes the polynomial shapes. The other phase is the prototypes and detects the type of the objects and calculates the relative orientation.
5. Appearance Based – deals with the 3d recognition of the object of occlusions.

* Local Features- located at a single point, gradient, color, will give the pixel.
* Global Features- Whole image is regarded. Histograms are used.

1. Background Subtraction- In this method the foreground objects are subtracted from the background organizations in the photo frames. The various methods maintain the background estimation based on the input frames and approximation of the median filter, Kalman filter.
2. Foreground Detection - In this the objects are separated from the background model from the background subtraction. The detection of the foreground image and compares the input video frame input video frames with the background model. The initial background models are compared with the shadow detection and the background subtraction using density and Kernel estimation.

Local Shape and Global Descriptor-

1. Local Shape- The local shape may include the noise and the sampling errors. The geometric shape of the function leads to the change in the shape that may be robust to the noise. The Local shape descriptor is compared to the functions and the scales are linearly changed that are robust the noise. The geometric properties are measured which are not changeable for handling and are the invariant to rotation and the translation such as length, volume and angle.
2. Global Shape- the Global shape descriptor is based on the observation method for each class of the objects. For instance, the extraction of the street points position approach and the observation between the fire hydrant and poles. This descriptor is used to find the find stimulating points and by the brain mapping community for recognition of the fold surfaces .

Application of Object Detection-

1. Biometric Detection- the security has increased so much due to this.
2. Surveillance Systems- objects Tracking
3. Inspection of industries
4. Contract based image retrieval- (CBIR)
5. Autonomous Robotics- Main issue in recent world
6. Analysis of medical – some shaped recognition
7. Document Recognition

Future Scope –

It is concluded that, machine learning in object detection is an important technique in dealing with the occlusion, positioning ,scale transformation and lighting. The machine learning method has shown the impressive performance on the various vision tasks such as image classification, object detection and object classification. The various domains of the object detection based on the different objectives and classified on specific and conceptual categories. The different methods in the object detection recognize the part based, region of the image and contour based image.

Paper 14:

**Digital Image Processing Techniques for Object Detection From Complex Background Image**

**By: Shivin Gupta**

**SUMMARY-**

Object detection and location is very useful invention. Goal is to detect the mangoes by the means of color and shape. The MATLAB will count the total number of mangoes on the Tree. Mat lab is the most popular image processing software as it has various libraries.

The Problems which occur during the image processing –

1. The object may be blocked by other object
2. The lightning condition may also effect
3. Overlapping
4. Methodology-

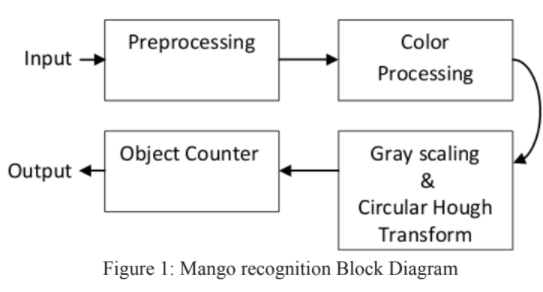


Fig7. Mango recognition Block Diagram

The image is resized in pre processing state -> then RGB adjustment ->color processing all the color is eliminated and only mango color is left. All the background removed and then the grey scale is adjusted for the shape detection of mango from leaves. After that the CHT are applied on the selected image to find the circular patterns within an image. It uses to transform set of feature points in the image into set of accumulated votes in the parameter space. Accumulated votes are in form of array and highest number of array indicates the presence of the shape.

1. Results-
2. Color Processing – It will only detect the matching RGB colors. It needs static lightning conditions. If the change is lightning is occurring then the RGB colors may change and the object may not detect accurately so we have to focus more on Shape Detection by using CHT. It colors of each pixel to detect.
3. Circular Hough Transform (CHT) – Due to Inconsistent lightning shape detection is more accurate and if we have the sphere like object in the image we can just run the math of circular formulas and the all the other objects other than sphere will be eliminated.

Lastly, CHT take place to detect the circular objects and display the total number of it. Lastly, the CHT may not exactly detect the circular object as sometimes it is connected with other object together and give an inaccurate result.

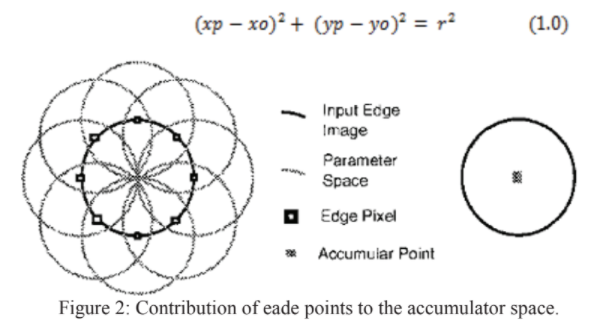


Fig8. Contribution of each points to the accumulator space

Paper 15:

**Object Detection Using Pulse Coupled Method**

**By: Shivin Gupta**

**SUMMARY-**

The union of every subset of segments in a cluster constitutes a potential tank if the union satisfies the following conditions.

This eliminates duplication by ensuring that any subset of segments that can be formed from the segments of a cluster cannot be formed from the segments of any other cluster.

• The union of segments must form a connected region.

OBS

The detection system finds potential objects in the input image and extracts a sub-image from each potential object area. These sub-images are further processed by the recognition system. This paper describes an object detection system which utilizes PCNN’s for smoothing and segmentation of digital images. The object detection system

1. System Organization- 5 modules
2. Smoothing module- educes the random noise
3. Segmentation module- forward the segment to control module
4. Control module – checks if the segment is available as a part of object
5. Detection module-
6. Knowledge base- it’s the knowledge of the attributes.

It repeats until the termination condition is attained

1. System Operation-

Step 1: Image smoothing using PCNN is accomplished by modifying the intensities of noisy pixels based on the neuron firing patterns. In general, the intensity of a noisy pixel is expected to be significantly different from the intensities of the surrounding pixels.

Image smoothing can be accomplished by adjusting the intensity of each pixel so that the corresponding neuron either captures its neighbors or is captured by them.

Step 2 - The general approach to segment images using PCNN is

to adjust the parameters of the network so that the neurons

corresponding to the pixels of a given region pulse together and the neurons corresponding to the pixels of adjacent regions

do not pulse together.

Step 3- Image segments to be discarded are determined. The knowledge base includes information to facilitate the evaluation of individual segments and the merging of adjacent segments. The actual size and shape of the tank in the image varies depending on the orientation, the imaging system and the imaging geometry.

1. Area Constraint
2. Elongation Constraint

If the number of active segments at the end of the current iteration is less than a pre specified value (25 segments), the segmentation process is terminated. The resulting segments are input to the detection module for further processing. Otherwise, the value of is increased by a constant amount and the process of segmentation (Step 2) and segment elimination (Step 3) is repeated.

Step 4- Detection of a group of connected regions that could form an object of interest. Each segment received by the detection module may be a tank or a part of a tank. The function of the detection module is to select groups of segments that could constitute a

Potential tank. This task is accomplished in two steps—cluster seeking and cluster analysis. First, in cluster seeking, the input segments are grouped into several clusters so that segments that belong to different clusters, together, cannot form a tank (based on the attributes in the knowledge base).The set of input segments are sorted in the non increasing order of their areas. The process continues till all the segments in the input set are processed. The largest possible object that can be assembled from the segments of a cluster consists of all the segments in the cluster. Therefore, if the sum of the areas of all the segments in a cluster is less than the minimum area of the tank, the cluster is eliminated.  
  
If the sum of the areas of all the segments remaining in the input set is less than the minimum area of the tank, the clustering algorithm is terminated. The boundary of each member segment of the cluster is highlighted. There are three member segments in the first cluster , zero member segments in the second cluster , three member segments in the third cluster and five member segments in the fourth cluster .

Cluster Analysis –

The union of every subset of segments in a cluster constitutes a potential tank if the union satisfies the following conditions.  
This eliminates duplication by ensuring that any subset of segments that can be formed from the segments of a cluster cannot be formed from the segments of any other cluster.  
• The union of segments must form a connected region.

Paper 16:

**Performance analysis of moving object detection using BGS techniques in visual surveillance**

**By: Shivin Gupta**

**SUMMARY-**

The motion based detection is profoundly used in the IT sector nowadays and in surveillance as well,

It predicts the trajectory of the moving object.

1. Background Initialization- takes few framed into consideration.
2. Foreground detection – compare the particular scene with the background frames
3. Background maintenance- Bugs Model updated periodically

2. State of ART

Models of Background Subtraction

1. Basic background model – it takes few initial frames into consideration. Best for static Background.
2. Gaussian Model- it is a mixture of Gaussian pixel intensity values, the color histogram and texture info.
3. Cluster BG model- this handle the illumination changes and the moving objects.
4. Neural and neuro-fuzzy based network BG model-weighted sum of network and network is trained on a clean frames
5. Estimation Model- divergate greatly pixel is known as foreground pixel.
6. Fuzzy logic Based Model- introduced

concept of fuzzy logic for object detection and object tracking to measure uncertainty and imprecision such as moving leaves, camera jitter, illumination variations, etc. and two algorithms, uncertainty over mean.

1. Principal Component analysis modeling approaches-
2. Subspace and Low rank Modeling approach- This algorithm iteratively accomplished incremental gradient descent compel to the Grassmann manifold of subspaces to get simultaneously evaluation of images which are decaying into a low-rank subspace, a sparse fraction of occlusions and foreground pixels, and a transformation such as image spinning.
3. Sparse and domain transform modeling approaches- The discrete cosine transform

(DCT) coefficients are calculated for the training and testing frame.

Application of Object detection –

1. Intelligent Visual surveillance- The main aim is to detect moving object for security purpose of particular area on traffic such as in airport, roads etc.
2. Optical Motion Capture-full capturing of human being.
3. Intelligent visual observation of forestry – surveillance system to check the behavior of animal in restricted areas.
4. Human machine interaction
5. Content based video coding
6. Medical surveillance – for critical people medical officers assign surveillance to the patient.
7. Bio-metric Identification – verification on the basis of psychological characteristics. Faces , fingerprint .
8. Aerial surveillance- digital imaging technologies have strengthen a lot in hardware of this surveillance such as micro-aerial vehicles, forward looking infrared (IR) and high resolution imagery that are capable of identifying the objects at extremely long distances as shown in Figure 4(a)
9. Satellite Surveillance- used for terrain visualization for terrain in both 2D and 3D



Fig9. Aerial & satellite surveillance system

Challenges and Issues-

1. Noise image- noisy image are result of poor grade picture source such as images taken from web cameras or images that are being compressed.
2. Camera jitter- camera is displaced due to motion of wind and causes nominal motion that result in detection of false images.
3. Illumination changes- sudden and gradual changes in light results in false detection.
4. Bootstrapping: during the training period in some environmental conditions background is not present. Then, it becomes very difficult to compute a background.
5. Foreground Aperture- if changes occur in uniform coloured objects then it will not be detected.
6. Moved background Objects- sometimes the background object is moved and should not considered as a part of foreground. So, both initial and current object position are detected without efficient background maintenance mechanism.
7. Camouflage- characteristic of foreground object are assumed on the basis of background model that often lead to miss interpretation of background and foreground pixel.
8. Inserted background object: in case a new background object is inserted then this inserted object cannot be considered as a foreground object because this object is detected without any robust maintenance mechanism.
9. Dynamic background: background is sometimes so cluttered it become very difficult to differentiate a pixel as a foreground or background. There are various types of background that can lead to false detection of pixel such as waving
10. Leaves, rippling of water, water surface, etc.
11. Sleeping foreground object: it become very difficult to distinguish non moving objects from background. Then they are considered as background.
12. Shadowing: shadow detection is an active research area itself and can be detected as

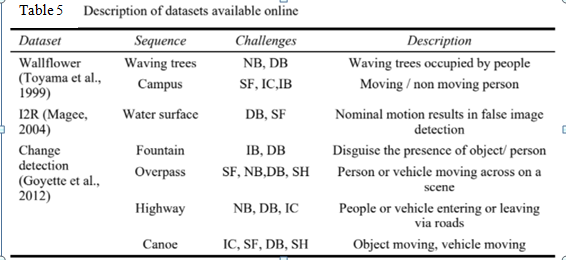
foreground comes from moving object or background.

Resources, Datasets and codes-

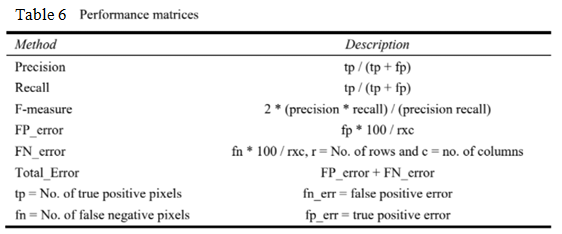
1. LIBRARIES
   1. Background subtraction websites- This website contains links to the datasets and list of references. Data is updated regularly and different background models are classified as per the information provided by this website.
   2. BGS libraries- A developed BGS library contains 29 BGS algorithm for background modeling
   3. IMBS libraries- IMBS library is a C++ library specifically designed for real time accurate foreground extraction. IMBS creates a multimodal model of the background in order to handle several challenging issues such as gradual or sudden illumination variations, activities of minute background rudiments, camera jitter, and change of background framework. A statistical analysis of the frames is performed for background modeling
2. DATASETS
   1. Wall Flower- It consists of full fledge details of seven different video sequences with obstacles that are likely to encounter like illumination or intensity changes, dynamic or cluttered background.
   2. I2R8 dataset- this dataset mainly consist of nine video sequences and was provided Each sequence has illumination effect or dynamic background with 176 × 144 pixels image size. The important benefit of this dataset provides 20 ground truth images for every single frame sequence. There ground truths are captured only when critical situation occurs.
   3. PETS- related to conference performance, evaluation of tracking and surveillances. These are mainly used for object tracking evaluation as compared to BGS
   4. Change detection.net- comprises of realistic, large scale video sequence nearly about 90,000 frames in31 video sequence that represent six different categories to cover vast orbit of challenging issues in two modularity’s ( color and thermal IR).
   5. 3DPeS- 3D people Surveillance dataset is specifically designed for re identification of human beings in case of multi camera systems with non-overlapping field of view.
   6. OTCBVS dataset: This dataset is freely or publicly available benchmark for testing and evaluating different computer vision algorithms.
   7. BMC 2012 dataset-BMC mainly focused on comparison of BGS technique with both real and synthetic videos.
   8. UCSD background subtraction dataset-this dataset consists of 18 video
   9. Sequences with their ground truth images in 3D array format starting fromframe1 of the sequence and each frame is in the format of JPEG but in some cases ground truth mask is smaller than the No.
   10. CSIR-CSIO (CSIR India) dataset: this dataset has been provided by Council for Scientific and Industrial Research (CSIR), India. This dataset is used for object detection on thermal video.
   11. Houston zoo dataset: this dataset comprises of both colored and gray scale video sequences. Main advantage of this dataset is that it comprises of training, testing frames with their ground truth video sequences.

Experimental Evaluation-

1. Statistical measurement for system performance



1. Performance analysis of considered state-of-the-art of the literature.



* 1. Benchmark evaluation

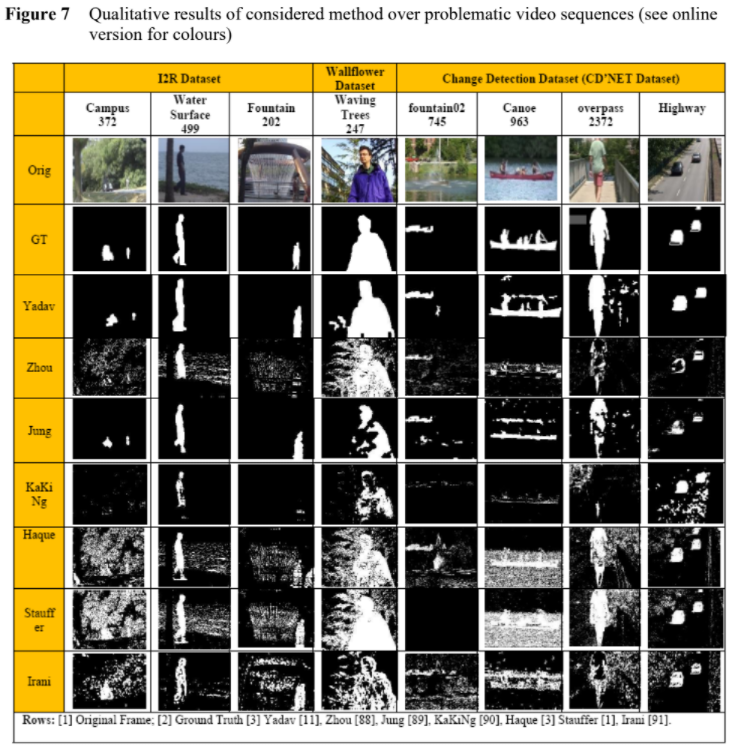


Fig10. Qualitative results of considered method over problematic video sequences

Fig10. Qualitative results of considered method over problematic video sequences

2.2 Quantitative analysis across combination of metrics-BGS algorithms on different datasets are evaluated byusing several statistical and combination of metrics

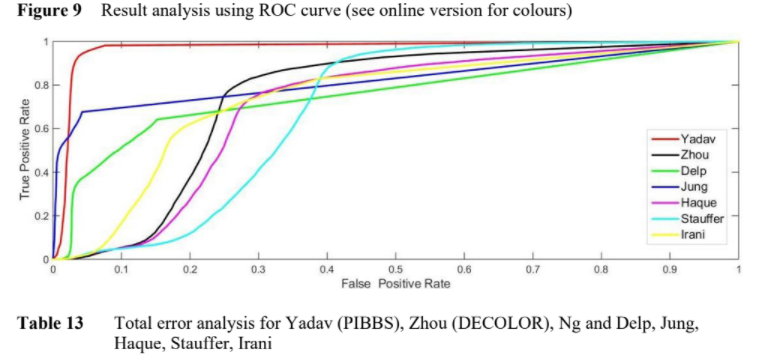
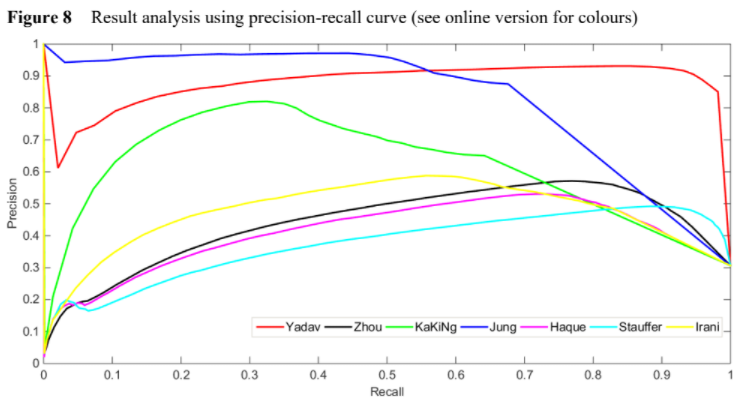


Fig12. Result Analysis using ROC Curve

Fig11. Result analysis using precision-recall curve

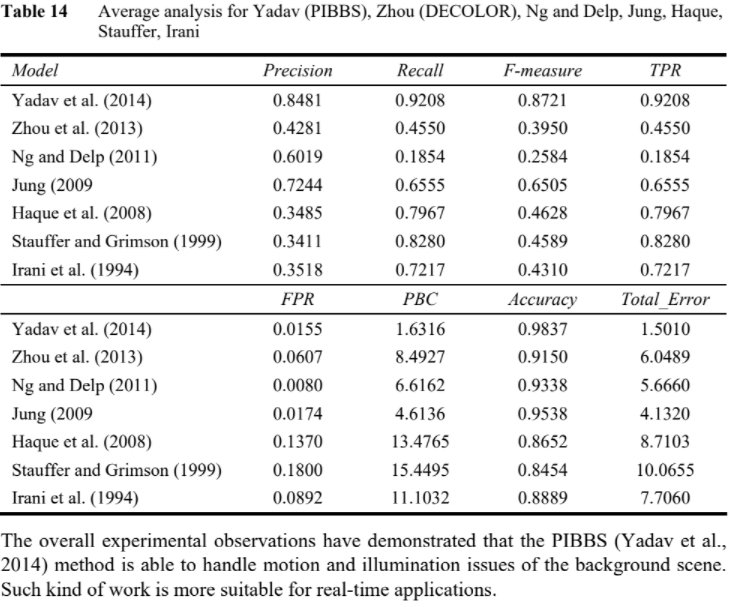


Table 7

# Comparison Table:

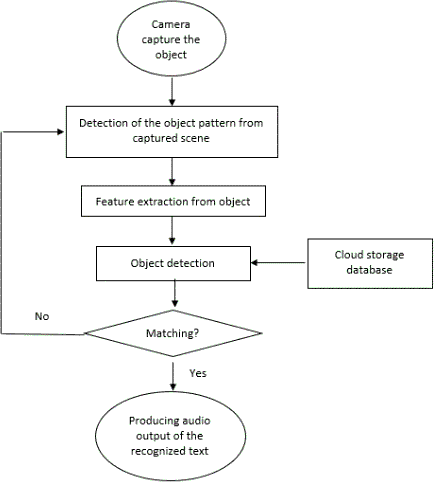
**Table 8:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **S.No** | **Paper title** | **Author’s Name** | **Year** | **Approach used** | **Finding** | **S/w and H/w Required** |
| 1 | Application of Deep Learning for Object Detection [1] | 1. Ajeet Ram Pathaka 2. Manjusha Pandeya 3. Siddharth Rautaraya | 2018 | Deep Learning | Deep learning techniques for state-of-the-art object detection systems are assessed. | Deep learning frame works and data sets. |
| 2 | A SURVEY ON MOVING OBJECT TRACKING USING IMAGE PROCESSING [2] | 1. S. R.Balaji 2. Dr. S. Karthikeyan | 2017 | Object Tracking | Image processing converting image into digital inform by performing some operations on  it. | Video surveillance technology. |
| 3 | Physiologically Motivated Image Fusion for Object Detection using a Pulse Coupled Neural Network | 1. Randy P. Broussard 2. Steven   K. Rogers | 1999 | PCNN for object detection | PCNN’s are used to fuse the results of several object detection techniques to improve object  Detection Accuracy | SCUD launcher and flash pods, DoG filter, Morphological filtered, PCNN fusion network |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 4 | Underwater | College of Shipbuilding Engineering, Harbin Engineering University | 2020 | RCNN | A | Deep |
|  | Image |  | For | combination | learning |
|  | Processing and |  | Object | of max-RGB | environment, |
|  | Object Detection |  | detection | method and | data base for |
|  | Based on Deep |  |  | shades of | images and |
|  | CNN Method |  |  | gray method | different |
|  |  |  |  | and then a | objects. |
|  |  |  |  | CNN method |  |
|  |  |  |  | for solving |  |
|  |  |  |  | the weakly |  |
|  |  |  |  | illuminated |  |
|  |  |  |  | problem. |  |
| 5 | Object Detection Using Image  Processing | Fares Jalled, Ilia Voronkov | 2016 | Image processing using Haar- like features for object detection. | Tracking object by special features locking. | Open cv python. |
| 6 | Interfacing of MATLAB with Arduino for Object Detection Algorithm Implementation using Serial Communication | Panth Shah, Tithi Vyas | 2014 | Algorithm on MATLAB | Detection and comparison of image by conversion into binary image | MATLAB,  Serial communicator, Arduino. |
| 7 | Video Image Processing for Moving Object Detection and Segmentation using Background | Anaswara S Mohan, Resmi R | 2014 | Detecting the moving object. | Separation and background subtraction. | Surveillance system. |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Subtraction |  |  |  |  |  |
| 8 | Transformer Transforms | Yuxin Maoz, Jing Zhangz, Zhexiong Wan, YuchaoDai,Aixuan Li, YunqiuLv, Xinyu Tian, Deng-Ping Fan  and Nick Barnes | 2015 | Localize images that | Transforming through | vision transformer |
|  | Salient Object |  | attracts | transformer. | architecture. |
|  | Detection and Camouflaged |  | human  attention. |  |  |
|  | Object |  |  |  |  |
|  | Detection |  |  |  |  |

**Proposed Approach**:Fig13.



**Conclusion:**

Object detection is considered as foremost step in deployment of self-driving cars and robotics. In this paper, we demystified the role of deep learning techniques based on CNN for object detection. Deep learning frameworks and services available for object detection are also discussed in the paper. Benchmarked datasets for object localization and detection released in worldwide competitions are also covered. The pointers to the domains in which object detection is applicable has been discussed. State-of-the-art deep learning-based object detection techniques have been assessed and compared. Image processing and more specifically object detection is the modern era thing, and in the coming future it will be doing wonders for the world if we look in to the day by day increasing technological advancement(s) we can easily understand that.

# Future Scope:

Future directions can be stated as follows. Due to infeasibility of humans to process large surveillance data, there is a need to bring data closer to the sensor where data are generated. This would result into real time detection of objects. Currently, object detection systems are small in size having 1-20 nodes of clusters having GPUs. These systems should be extended to cope with real time full motion video generating frames at 30 to 60 per second. Such object detection analytics should be integrated with other tools using data fusion. The main issue is how to integrate processing into a centralized, powerful GPU for processing data obtained from various servers simultaneously and performs near real time detection analysis.

If we talk about the future scope Object detection has the potential to free people from menial jobs that can be done more efficiently and effectively by machines.

The future of our project is that we want to become a bliss for the visually aid people

## References

1. Application of Deep Learning for Object Detection/ International Conference on Computational Intelligence and Data Science (ICCIDS2018)
2. A SURVEY ON MOVING OBJECT TRACKING USING IMAGE PROCESSING/ 2017 11 the International Conference on Intelligent Systems and Control(ISCO)
3. Physiologically Motivated Image Fusion for Object Detection using a Pulse CoupledNeural Network/ IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 10, NO. 3, MAY1999
4. Underwater Image Processing and Object Detection Based on Deep CNN Method/ Hindawi Journal of Sensors Volume 2020, Article ID6707328
5. Object Detection Using Image Processing/Moscow Institute of Physics &Technology, Department of Radio Engineering &Cybernetics,2016
6. Interfacing of MATLAB with Arduino for Object Detection AlgorithmImplementation using Serial Communication**/**International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 3 Issue 10, October-2014.
7. 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
8. Transformer Transforms Salient Object Detectionand Camouflaged ObjectDetection/ JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8, AUGUST2015
9. Applications of Object Detection System / International Research Journal of Engineeringand Technology (IRJET) Volume: 06 Issue: 04 | Apr2019
10. Object Detection Using Gabor Filters / Pergamon Pattern Recognition, Vol. 30, No. 2,1997
11. Rapid Object Detection Using a Boosted Cascade of Simple Features / Paul Viola - Cambridge, MA 02139 & Michael Jones - Cambridge, MA02142
12. Object Detection Based on YOLO Network / 2018 IEEE 4th Information Technologyand Mechatronics Engineering Conference (ITOEC2018)
13. Study of Object Detection Methods & Applications on Digital Image / May 2019 IJSDR Chandigarh Engineering of College Landran Mohali, India
14. Digital Image Techniques for Object Detection From Complex Background Image / International Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012)
15. Object Detection Using Pulse Coupled Neural Networks / IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 10, NO. 3, MAY 1999
16. Performance analysis of moving object detection using BGS techniques in visual surveillance/ Int. J. Spatio-Temporal Data Science, Vol. 1, No. 1, 2019

Review paper

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, robots, and self-driving systems are just a few of the many applications that have prompted extensive study in the field of computer vision in the last decade. Visual recognition systems, which include picture categorization, localization, and detection, are at the heart of all of these applications and have gathered a lot of research attention. These image identification algorithms have achieved extraordinary performance because to considerable advancements in neural networks, particularly deep learning. Object identification is one of these sectors 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.

# Introduction

With the advent of the electronic medium, particularly the computer, society has become increasingly reliant on computers for information processing, storage, and transmission. In modern civilization, computers play a significant role in every aspect of life and society. 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 item.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.

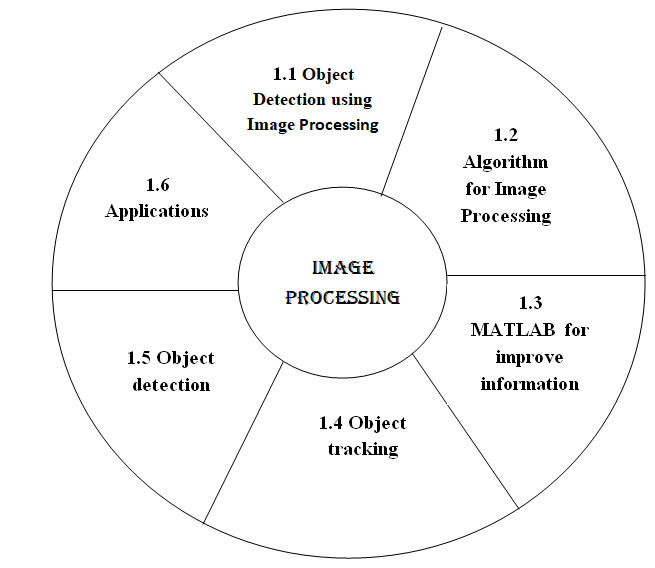


Figure 1: overview of the paper

# 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.

## 1.1 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 process of image processing and how image processing is done.

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.

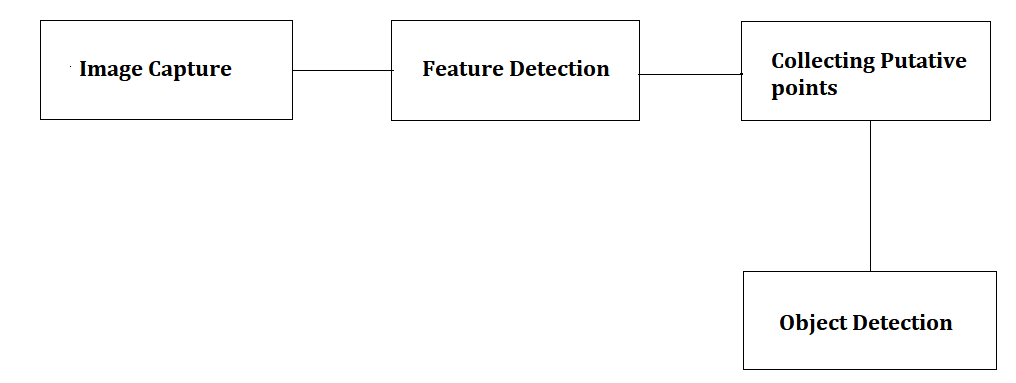


Figure 2.1: object detection steps

So, what is Haar Cascade? It is face detection algorithm which is used for the detection of face from an image or real time video. The algorithm uses edge or line detection features proposed by Viola and Jones in their research paper “Rapid Object Detection using a Boosted Cascade of Simple Features” published in 2001. The algorithm is rained by giving it a lot of positive images which consists of face and lots of negative images which doesn’t consists of face.

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:

* Detect a face to track
* Identify facial features to track
* Track 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, these object detectors use Viola Jones detection algorithm to detect a face. 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.

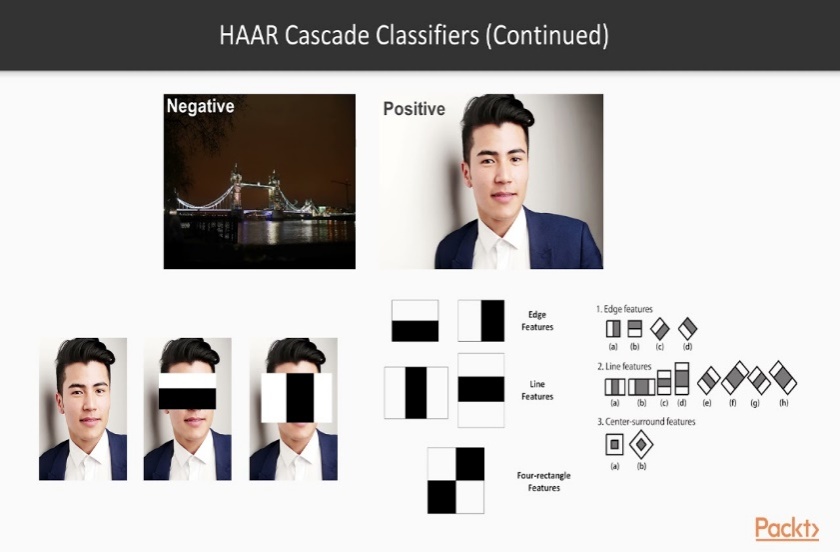


Figure 2.2: Haar Cascade Classification

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: Both selecting features and training the classifier are done with Adaboost. Here, the weak learner must choose the feature that best distinguishes the weighted positive and negative training samples.
* Attentional Cascade: They are used for sequential integration of highly complicated classifiers, allowing background areas of the picture to be swiftly dismissed while more computation is spent on potential object-like regions. Simpler (low T) boosted classifiers are employed first to reject the vast majority of negative windows while passing practically all positive windows. Then, to reject the much smaller number of problematic negative windows, more sophisticated and hence slower boosted classifiers having higher threshold are utilized.

## 1.2 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.

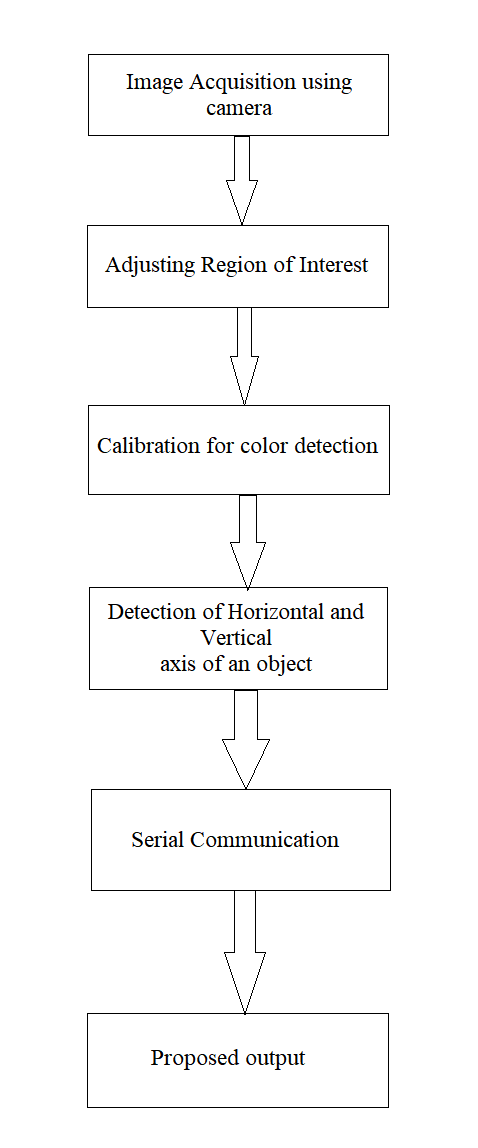


Figure 2.3: Proposed prototype approach

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.

## 1.3 Object detection and segmentation using Background Subtraction

Image segmentation is a technique in digital image processing and computer vision in which a digital image is broken down into various subgroups called image segments to help reduce the complexity of the image and make further processing or analysis of the image easier. In simplified way, image segmentation is the process of assigning labels to pixels. A common level is allocated to all image parts or pixels that belong to the same category. Image segmentation is the process of converting 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 aids in the annotation of an object scene. Image segmentation is required in order to accurately identify the image's content. Edge detection is a crucial technique for picture segmentation in this scenario. For picture segmentation, a number of general-purpose algorithms and approaches have been developed.

The image/frame is separated into a series of non-overlapping uniformly linked sections throughout the segmentation process, such that no two neighbouring ones are same. Due to the richness and diversity of pictures and moving objects, this is a challenging task. Illumination, contrast, and frames are all influencing elements. The majority of these segmentation techniques are based on similarity and difference, and may be classified as threshold, template matching, region growth, edge detection, and clustering.

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

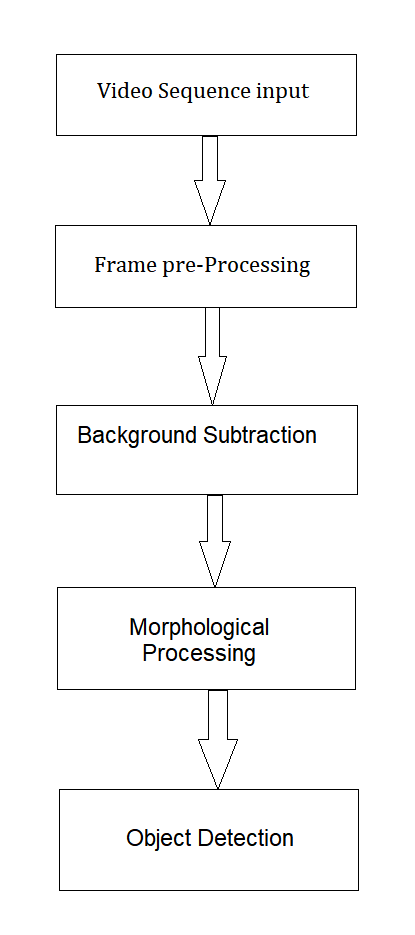


Fig. 2.4 Background subtraction for Object detection

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:

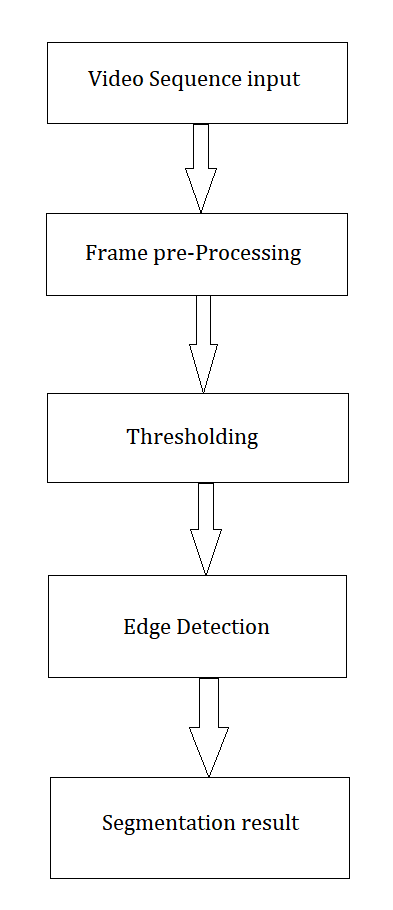


Fig. 2.5 Image Segmentation

In image processing, segmentation is the most crucial step. The split of an image into areas or categories that correspond to various items or parts of objects is known as image segmentation. There are two processes to picture segmentation. The image is first transformed to binary images, after which thresholding and edge detection are used to segment items. Each pixel in an image is assigned to one of many categories. A good segmentation is typically one in which:

***(a)***The grey scale of multivariate values in the same category of pixels is identical, forming a connected zone.

***(b)*** The values of neighboring pixels in various categories are varied.

## 1.4 Object tracking

Object tracking is a deep learning application in which the software takes a series of initial object detections and creates a unique identifier for each of them, then follows the detected objects as they move across frames in a clip. Object tracking is achieved through different steps such as:



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. It is a technique for tracking and determining the travel direction of objects. 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 point tracking necessitates detection at each frame. Object tracking can be achieved through different methodologies like:

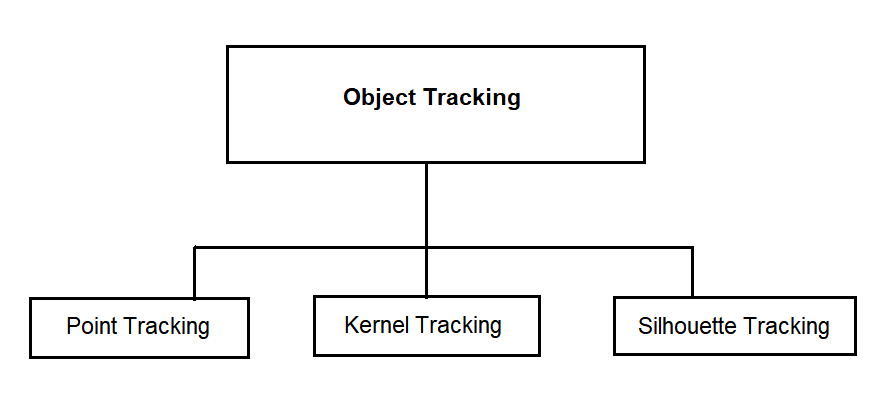


Fig. 2.7 Sub process for object tracking

## 1.5 Object detection

We debunk the importance of deep learning approaches based on convolutional neural networks in object recognition, which is one of these fields that has seen a lot of progress in computer vision. Object identification frameworks and services based on deep learning are also discussed. Deep learning approaches are also discussed for state-of-the-art object identification systems. In 2019, the CV market is expected to reach $33.3 billion, fueled by strong growth in the consumer, robotics, and machine vision sectors.

Convolutional neural networks have the advantage of not requiring feature extractors or filters to be manually developed. Rather, they self-train from the pixel level up to the final object categories. For object detection, deep CNNs have been widely employed. CNN is a sort of feed-forward neural network that operates on the weight-sharing concept. CNN makes use of a variety of 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**:**

Primate vision processing methods such as expectation driven filtering, state dependent modulation, temporal synchronization, and multiple processing channels are employed to construct a physiologically motivated image fusion network. PCNNs are used to aggregate the findings of several object detection algorithms to improve object detection accuracy.

Picture processing is used to remove unnecessary information from an image in the hopes of allowing a pattern recognition procedure to discover and maybe identify the intended 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 colour information, whereas the second pathway primarily processes shape and motion. In a Simplified Model of the Primate Vision System, these principles are used to develop an image fusion network that distinguishes an item, combines features, and isolates the object from the rest of the picture.

The PCNN fusion network takes an original and many filtered copies of a gray-scaled picture and returns a single image in which the targeted items are the brightest and therefore readily spotted to conduct object detection. Each neuron gets feeding inputs, which are the intensities of the pixels in the input picture that correspond to it. The PCNN's pulse-based connection methods segment the original picture via temporal synchronization. The outer PCNNs generate state-dependent modulation signals that are used to direct attention to certain regions.

The essential feature that separates the PCNN from other types of neural networks 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 sends out a linking signal (linking wave) across the multiplicative linking interconnects, forcing additional neurons with comparable inputs to fire. Using the centre PCNN's connecting inputs, these pulsed signals are connected to the original picture. These signals fuse features into their related segments while also modulating the neural response of the central PCNN to the object of interest.

The PCNN network reduced 46 percent of erroneous detections while only deleting 7% of real detections in comparison to the best filter result. On the FLIR pictures, the fusion network increased accuracy more than on the mammography images. The network lowered the false alarm rate in FLIR pictures from 8.2 to 0.6 per image, and in mammograms from 1.7 to 0.8 erroneous detections per correct detections.

The PCNN network does not add true detections to the output during the fusion process, rather it eliminates incorrect detections. It also has a strong computer architecture for physiologically-based fusion and other physiologically observed pulse-based phenomena.

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

Underwater exploration is so vital in the development and use of deep-sea resources, autonomous underwater operation is becoming increasingly important to avoid the dangerous high-pressure deep-sea environment. The existing technology has a common flaw that processing takes long time and real-time detection is difficult to achieve. In a variety of study domains, the Convolution Neural Network (CNN) is acknowledged as the quickest detection method. Since then, deep CNN has been frequently used to deal with classification problems, winning the ILSVRC (ImageNet Large Scale Visual Recognition Challenge) champion and reducing the top 5 error rate to 15.3 percent.

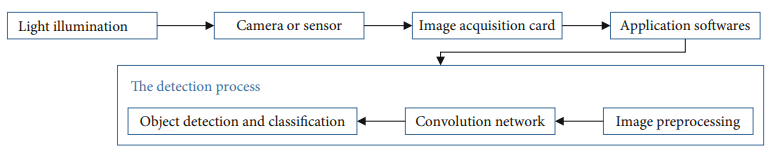


Figure 2.8 CNN base Detection Process

Since the early 2018, Redmon proposed the YOLO v3, which is widely regarded as the quickest detection technique, with considerably enhanced accuracy and detection speed when compared to previous approaches. The Convolutional Neutral Network is used to partition pictures into numerous non overlapping areas, and feature extraction is utilized for object recognition and categorization. Through this concept, the entire picture information is used to anticipate the bounding boxes of the targets while also classifying the objects. An end-to-end mapping between dark and bright pictures is immediately learned by the Convolutional Neural Network.

The loss function form is a crucial approach used in a sum squared error loss to balance the faults in the training process. The width and height of the bounding box are replaced for the boxes of varied sizes predicted. Augmentation of datasets, Underwater datasets are difficult to prepare, underwater images and video are difficult to come by on the internet, and underwater images have nearly identical backgrounds in the same area, so the images in the dataset are similar. As a result of these factors, the training output model is never effective when applied to other sea areas. As a result, the dataset needs be adjusted and supplemented to make the deep learning model more widely applicable. Rotation, flipping, zooming, shift, and other techniques are used to supplement datasets.

## 1.6 Applications

The purpose of object detection is to use a computer to imitate human intelligence. It has waste application in every single aspect of this growing technical world, may it be defense, medication, automation, nature reinstatement or any other. Like Surveillance, human computer interaction, robotics, transportation, retrieval, Spotting and detecting the wild animals in the territory of sterile zones like industrial areas, detecting the vehicles parked in restricted areas, Detection of faulty electric wires when the image is captured from drone cameras are all day-to-day applications of object detection. In just a few hours, sensors used for long-term monitoring create petabytes of picture data. These data are reduced to geospatial data and combined with other data to provide a clear picture of the present situation. Object detection is used in this procedure to monitor items such as people, cars, and suspicious objects in raw imaging data. Unattended luggage detection is a critical application of object detection. Detecting items on the road, which is also an application of object detection, would play an essential part in autonomous driving systems. Object detection may be used to identify the driver's tiredness on the highway in order to avert an accident.

RCNN also assisted in the development of an underwater remotely operated vehicle (ROV) for the purpose of fishing marine products. The robot is approximately 1 m long, 0.8 m wide, and 90 kg in weight. The adsorption technique is used to gather marine materials. The robot is controlled remotely, and multiple teams are working to rebuild the ROV to make it semiautonomous, with the essential technology being the ability to recognize and find items. By fusing the outputs of different object detection methods, the PCNN fusion network improved object detection accuracy. The fusion network outperformed the results of any single filtered output or the logical AND of all filter outputs in terms of accuracy. It’s application can also be seen in detection of cancer cells and 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. | Deep learning techniques for state-of-the-art object detection systems are assessed. | Deep learning frame works and data sets. |
| 2 | Moving object tracking using image processing [2] | Object Tracking | 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 Fusion for Object Detection using a Pulse Coupled Neural-Network [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. | Use of PCNN’ to fuse the results of several object detection techniques for improving the object detection accuracy. | SCUD launcher and flash pods,  Dog filter,  Morphological filtered, PCNN fusion network |
| 4 | Underwater Image Processing and Object Detection Based on Deep CNN Method [4] | RCNN For Object detection | The program was applied on an underwater robot for the real-time detection. Results showed that the robot achieved underwater working operation. | Dropout layers and other technologies are not significant in this model. The reconstruction of the network by using a more complicated algorithm would be more effective. | A combination of max-RGB method and shades of gray method and then a CNN method for solving the weakly 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 | Video Image Processing for Moving Object Detection and Segmentation using Background Subtraction [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. |

# Conclusion

As we can see from the ever-increasing technical breakthroughs, image processing and more precisely object identification is a modern-day phenomena that will do wonders for the world in the future.

We have provided a comprehensive examination of object detection and related technologies in this paper. We made an attempt to cover the basics of object identification and tracking. We've studied both the foundations and applications of Object detection. Object detection is the initial step in the development of self-driving cars and robotics. In this paper, we deciphered the function of CNN-based deep learning systems for object detection. The domains in which object detection can be applied were discussed. State-of-the-art object recognition systems based on deep learning have been examined and compared. 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.

# References

1. Application of Deep Learning for Object Detection/ InternationalConference on Computational Intelligence and Data Science (ICCIDS 2018)
2. A SURVEY ON MOVING OBJECT TRACKING USING IMAGE PROCESSING/ 2017 11 the International Conference on Intelligent Systems and Control (ISCO)
3. Physiologically Motivated Image Fusion for Object Detection using a Pulse Coupled Neural Network/ IEEE TRANSACTIONS ON NEURAL NETWORKS, VOL. 10, NO. 3, MAY 1999
4. Underwater Image Processing and Object Detection Based on Deep CNN Method/ Hindawi Journal of Sensors Volume 2020, Article ID 6707328
5. Object Detection Using Image Processing/ Moscow Institute of Physics & Technology, Department of Radio Engineering & Cybernetics,2016
6. 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.
7. 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
8. Transformer Transforms Salient Object Detection and Camouflaged Object Detection/ JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8, AUGUST 2015