Project Name: Object Detection through Deep learning

Report by- 3CSC-15

Tripti Somani (PCE20CS188)

Shivam Kumar (PCE20CS175)

**Bachelor of Technology in Computer Engineering - II Year**

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

# TABLE OF CONTENT

* Abstract
* Introduction
* Problem Statement & Objective
* Literature Reviews Paper[1-8]
* 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 in detail.

**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 of videos.

**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 Detectio(ICCIDS , 2018**

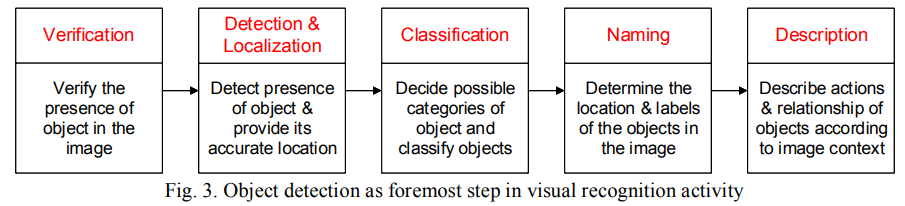
By: Tripti Somani

**SUMMARY:**

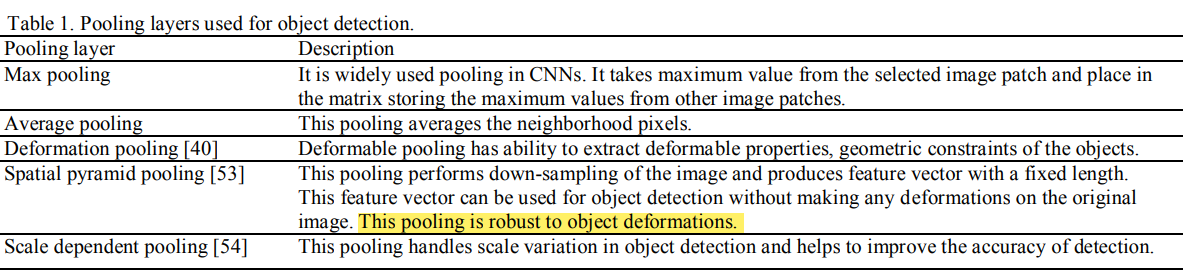
* Paper demystifies the role of deep learning techniques based on convolution neural network for object detection.
* Deep learning frameworks and services available for object detection are also enunciated. Deep learning techniques for state-of-the-art object detection systems are assessed in the paper.
* 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).

2. 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.



* 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:



5. 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.
* Spotting and detecting the wild animals in the territory of sterile zones like industrial 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.

A. 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.

B. 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.

C. 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 modulatory 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 initiates 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 modulatory 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.

4. 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**.

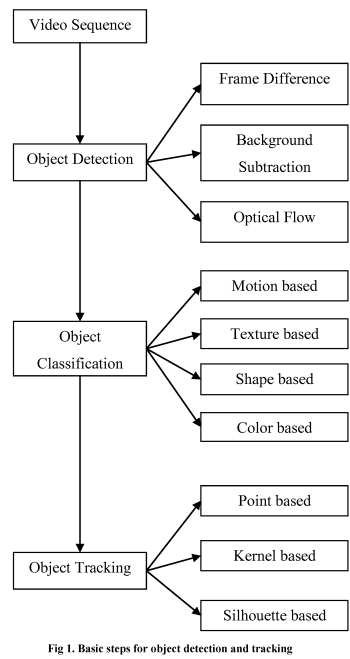
In the fusion process, the PCNN network does not add true detections to the output, but instead removes false detections. The PCNN also provides good computer architecture for implementing physiologically-based fusion and other pulse-based physiologically observed phenomenon.

Paper 03:

**A SURVEY ON MOVING OBJECT TRACKING USING IMAGE PROCESSING ( 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 moving of object of interest.

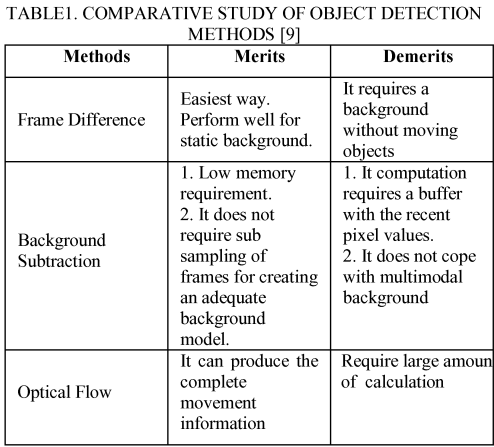
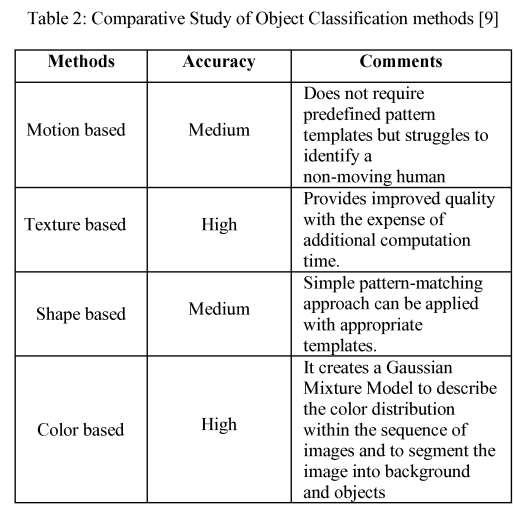
A. Frame Difference Method: The difference between two consecutive images is calculated, which further identifies the moving object.

B. Background Subtraction method: Background modeling is the first step of background subtraction. The two common approach available in background subtraction are:

i. Recursive Algorithm

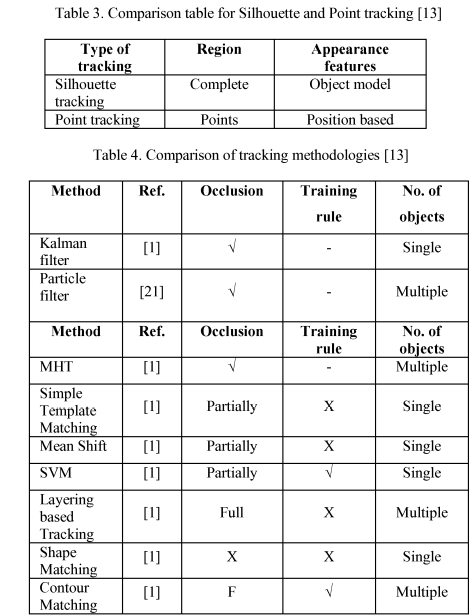
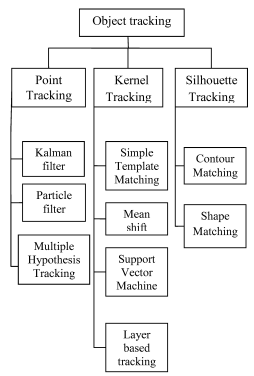
ii. Non-Recursive Algorithm

C. Optical Flow: The complete detection and movement information of object from the background can be obtained from this method.



2.Methods of Object Classification:classification of objects is done based on their shape features of the motion region.

3. Methods of Object Tracking: It is a technique used to track and also the travelling direction of objects



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 (ImageNet 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, Redmon 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’s feature map, and then c18 ∗ 6 ∗ 6 and 8 ∗ 1 ∗ 1 convolution layers are added in the 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:**

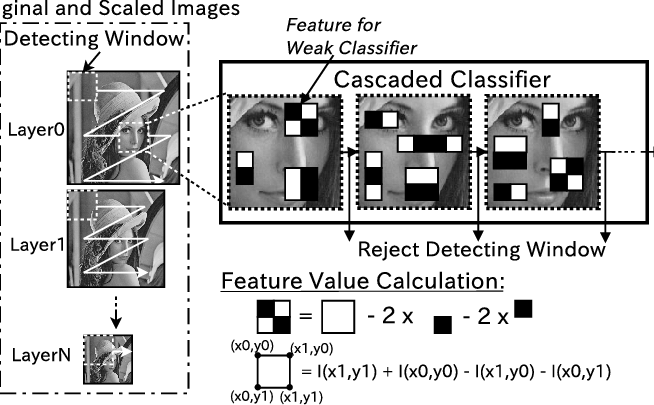
The aim of this paper is to is to develop an OpenCV-Python code using Haar Cascade algorithm for object and face detection at a budget with reduced human involvement.

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.



In this methodology we use Cascade Object Detector to detect the location of a face in a videoframe. 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**

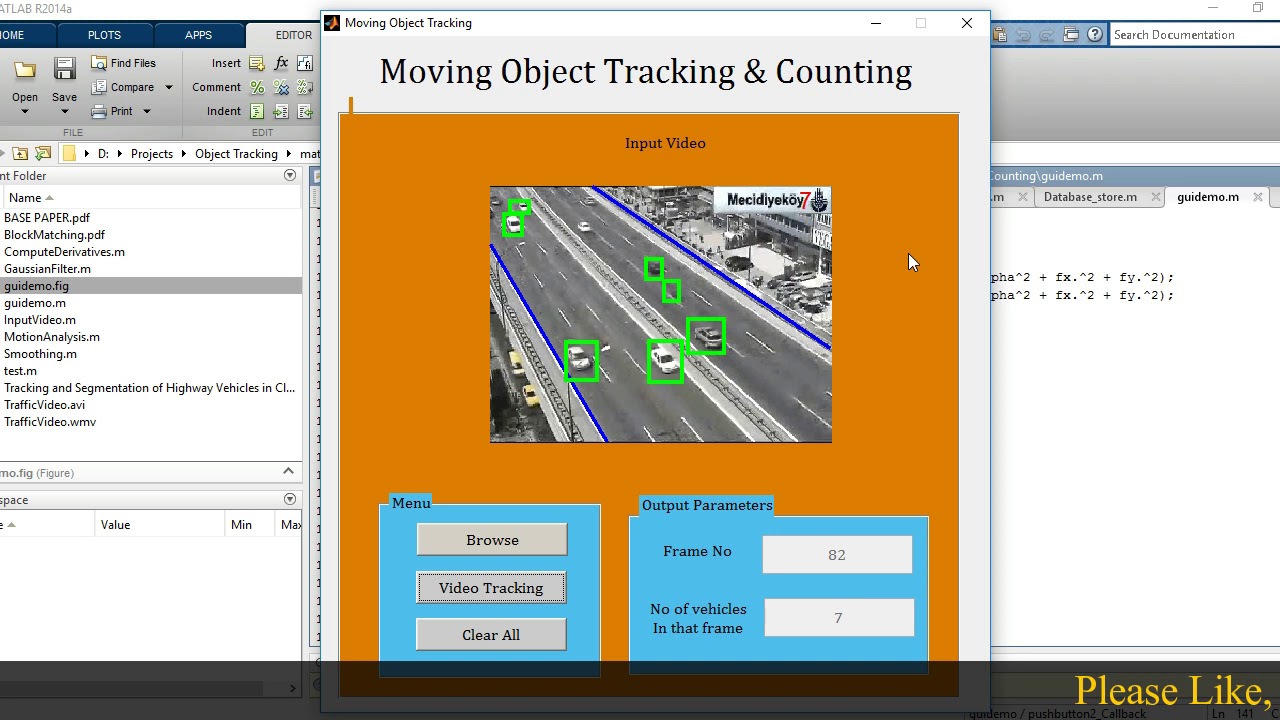
**SUMMARY:**

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.



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 powered by

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 using Background Subtraction (ICCSC, 2014)**

**By: Shivam Kumar**

**SUMMARY:**

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 grey scale. values and form a connected region.

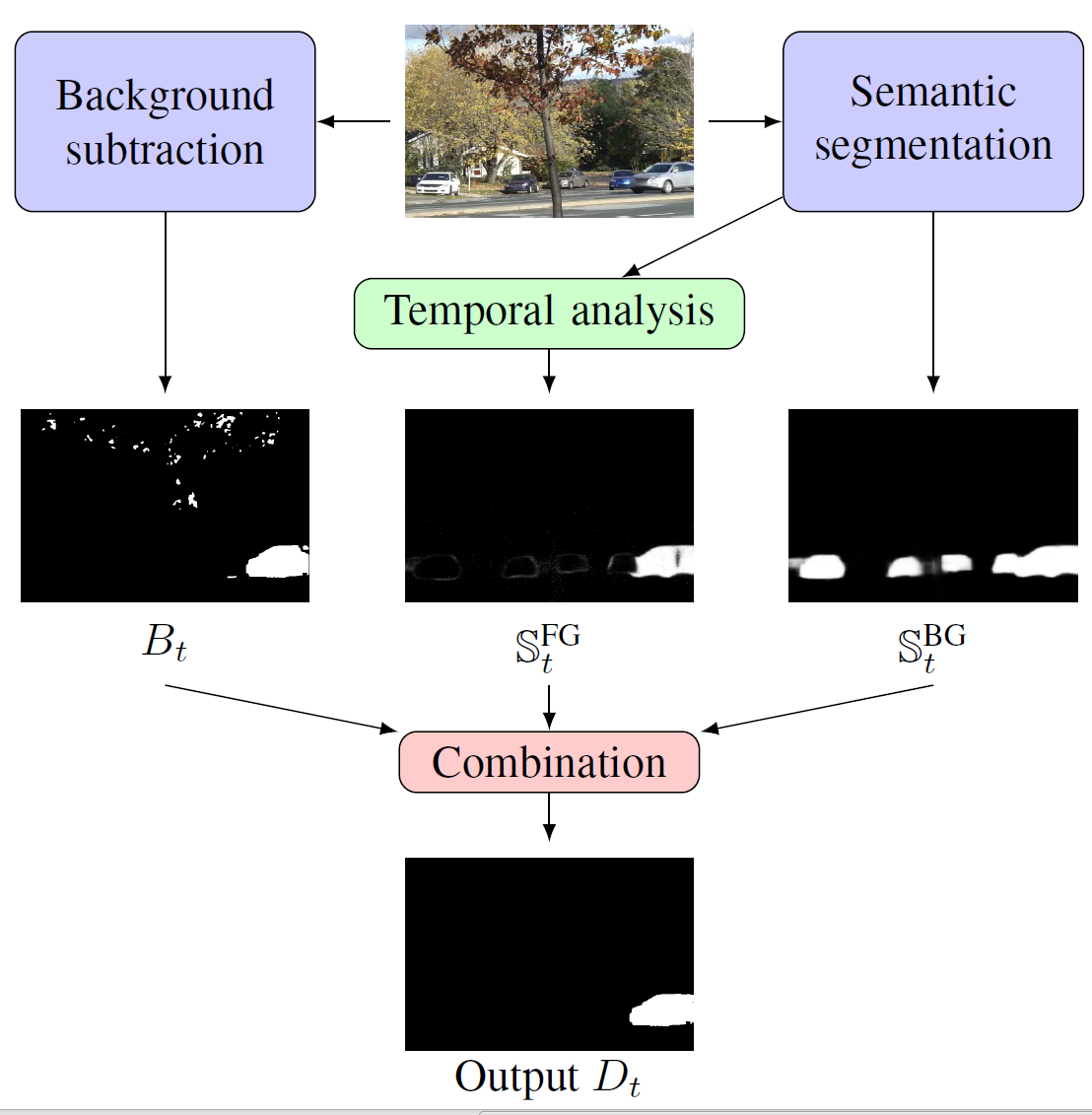


Image segmentation is the step at which we move from considering each pixel as a unit of observation to working with objects (or parts of objects) in an image. There are three general approaches to segmentation, termed thresholding, edge-based and region-based methods. The next important step in this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

effort, several weakly supervised salient object detection modelshave been proposed to learn saliency with image-level supervision

, scribble supervision or learn saliency directly from noisylabeling.

Before the deep learning revolution, conventional salient objectdetection modelsused handcraftedfeatures, which define saliency as the contrastbetweeneach pixel (or super-pixel) and the otherpixels (or super-pixels).

In this way, the receptive field of the conventional handcrafted feature-based models are the entireimage, which is the global context, where saliency of the pixel colored in yellowdepends on all the other pixels. The deep convolutionalneural network (CNN) based salient object detection modelsachieve significant performance improvementcompared with those handcrafted-featurebased techniques withmore sophisticated features extracted from the deep network.

The conventional deep CNN based SOD network usuallyincludes two main parts: an encoder to extract different levels offeatures, and a decoder to aggregate features from different levelsof the network for finer prediction. The encoder part is adoptedfrom the trained backbone network onImageNet, e.g., VGG,ResNet and the most effort for SOD models have been puton designing an effective decoder for feature aggregation. The main issue with this backbone networks are that thelarger receptive field is obtained with theloss of structure information as a sacrifice. In

this way, a network with a larger receptive field without losingfine-grained information can bebeneficial for context-based tasks,e.g., salient object detection, to achieve effective contextmodeling. Inspired byand theaccurate structure modeling ability of transformer, we present a“unified” transformer backbone-based SOD network to achievefully-supervised RGB image-based SOD, RGB-D image pair-based SOD and weakly-supervised RGB image based SOD withscribbleannotation, leading to three new benchmark models.

We also observe that the positional encodings of visiontransformer are less effective in modeling the accurate “spatial”information for dense prediction tasks. Then we investigate deepsupervisionand 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 camouflagedobject 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

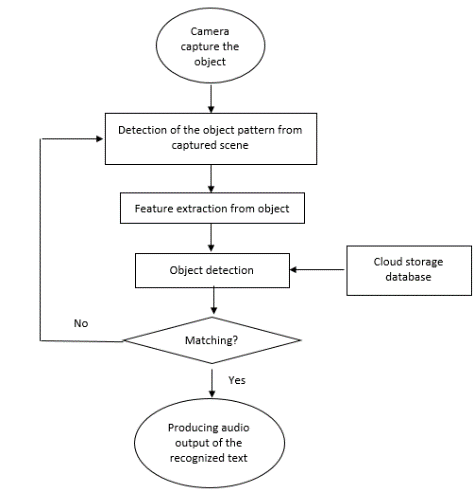
semantic information encoding with the long-range dependency

modeling mechanism.

# Comparison Table:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **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 Image Processing and Object Detection Based on Deep CNN Method | College of Shipbuilding Engineering, Harbin Engineering University | 2020 | RCNN For Object detection | 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 | 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 Subtraction | Anaswara S Mohan, Resmi R | 2014 | Detecting the moving object. | Separation and background subtraction. | Surveillance system. |
| 8 | Transformer Transforms Salient Object Detection and Camouflaged Object Detection | Yuxin Maoz, Jing Zhangz, Zhexiong Wan, Yuchao Dai,Aixuan Li, YunqiuLv, Xinyu Tian, Deng-Ping Fan and Nick Barnes | 2015 | Localize images that attracts human attention. | Transforming through transformer. | vision transformer architecture. |

**Proposed Approach:**

****

**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 (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 Detectionand Camouflaged Object Detection/ JOURNAL OF LATEX CLASS FILES, VOL. 14, NO. 8, AUGUST 2015