COMP 5435 – Computer Vision

Instructor-Dr.Shan Du

Fall 2019 Project Report

Title: - **OBJECT TRACKING**

Yatesh Laxman Kumar Gumma	1104076	COMP 5435 FA
Phaneendra Seethamraju	1097337	COMP 5435 FB
Manoj Kumar Bhuma	1101603	COMP 5435 FB

1.Introduction: -

By looking and seeing we come to know the world we live in. Environment that surrounds us is filled with endless type of objects and impressions. Vision in other words means to gain an understanding for the world around us. Even physiologist have been investigating from decades about how the visual system works. When speaking about replacing the living creature vision with the computational instrument, we have the broad and abstract expression computer vision. It can be summarized as the process of computers analysing digital images or videos and gain high level understanding from it.

One of the main objectives of the computer vision is to enable computers to replicate the basic functions of human vision such as motion perception and scene understanding. To obtain such kind of functionality researches have spent lot of efforts on Visual Object Tracking which also one of the most challenging and important topics in the computer vision. With usage of VOT in real-time applications like video surveillances, traffic monitoring, and human computer interaction it has gained more research attention. The main aim of Visual Object Tracking is to estimate the motion state of an object of interest in a given video. One of the live applications for visual object tracking is in the filed of traffic transportation where it is used to monitor the flow of traffic, pedestrian counting, accident detection and soon. Apart from that it is also used in MPEG-4 video compression technique where the object movement in the video is detected and high encoding bytes are assigned to it compared to that of redundant backgrounds.

Object Tracking can be defined as the problem of estimating the trajectory of an object in an image plane as it moves around a scene. The ability of object tracking depends on various parameters like information about the target object, type of objects that need to be tracked and even the format of the video. With the help of this object tracking many complex problems were easily solved one of such kind of example is MPEG-4 video compression. Even though there where lot of drastic changes in the way that the object tracking but still it is considered as one of the major complex problem. There are several important steps in the successful tracking of objects including the model selection to represent the object and the object tracking system suitable for the task.

2.Objectives:-

The question was how various objects are identified in a video where both rigid and non-rigid objects can be. Object Tracking may be used to automate the process, instead of having to manually search for an object in, for example, a product in a catalogue, to present information about a particular object showing itself in the video.

3.Literature Survey: -

First occurrence of the work related to the computer vision took place earlies in the year of 1970s. Researches always wanted to replicate the human intelligence into the computers is the one of the agenda during that period. Many of the pioneers in the field of artificial intelligence believed that solving the vision input problem is the easiest task but along the way it became more difficult and interesting. Even though digital image processing exists at that period, but the interpretation of the scene brought about the ability to use images to retrieve the world's three-dimensional structure which is regarded as a step towards the ultimate objective. First kind of scene understanding was done by edge extraction and then translating the 2D lines into 3D structures. Parallelly there were studies on some of the interpreting factors such as color intensities and shade variation for explaining the shadows and surface orientation of the image.

Over the next decade, researchers continued to study the relevant subjects, where some were more interesting than others. During this period a great many important work was carried out, particularly to track objects, when several tracking algorithms were dramatically improved. The work on the advancement of computer vision techniques in the 2000s continued. In the field of complex global optimization problems, amongst other fields, we have seen growing interest in developing more effective algorithms.

Visual objects have made considerable progress in certain classes, for example humans, heads and animals. Despite this tracking generic objects remains as a challenging task as they can change their visual appearance from one frame to another frame. While we still have a long way to go and many places to further explore, the technology for object tracking was greatly improved in recent years.

Vision based object tracking typically involves estimating the state of a moving object form a sequence of camera images. The state is typically includes the object pose and derivate parameters. Technically speaking object tracking starts with the object detection and then assigning the boundary boxes around them. The object tracking algorithm assigns id to every object identified in the video sequence and in the subsequent frames it tries to carry across the id and identify the new position of the same object.

Video analysis consists of three main steps: object detection, object tracking, and identification by evaluating the records of object activities. The tracking algorithms can be classified widely in board groups based on the tracking process. One approach is to split the processes into three main groups, namely point tracking, kernel tracking, and silhouette tracking. Other ways to categorize methods in object tracking are region-based tracking, contour-based tracking, and boundary-based tracking.

There are two main types of object tracking modes. Offline Tracking and Online Tracking. In case of offline tracking the object tracking in the pre-recorded video where all the information about the future frames are known in advance. And in case of online training object tracking is done the video is live streaming. One of the examples, of such kind is video surveillance. It more challenging because the algorithm must work faster, and it also doesn't have any information related to the future frames.

In general object tracking model consists system is composed of four different modules: Object Initialization, Appearance Modeling, Motion Estimation, and Object Localization.

- **Object Initialization** can be done either manually or automatically. Manual Initialization is done by the end user by drawing the bounding boxes around the object of interest in order to locate it in the next frames. In contrast we use object detectors like face or human detectors to detect the object in the given frame.
- **Appearance Modeling** usually classified into two categories visual representation and statistical modeling. Visual Representation is used to construct robust object descriptor using different type of visual features. Whereas in Statistical Modeling is used for building an effective mathematical model for object identification using statistical learning techniques.
- **Motion Estimation** is formulated as a dynamic state estimation problem. And the motion estimation task is usually completed with the help of predictors such as liner regression techniques, or Kalman filter.
- **Object Localization** is performed based on the motion estimation which is performed by greedy search or maximum posterior estimation.

Object detection has greatly developed over the last two decades, moving from standard statistical or machine learning approaches to CNN-based deep learning approaches. The introduction of deep learning enabled object detection as more accurate than others. As mentioned above, an extension of object detection is simply an object tracking method. The developers of the common algorithm called Simple Online and Realtime Tracking (SORT) say that modern algorithms of object detection can detect and re-identify objects in the next frames most of them.

Other developed robust object training algorithms that deal with object detection and apply fundamental learning techniques to bring an object into the next frame. Some of the challenges of object tracking compared to static object detection are:

- **Re-Identification** – Connecting an object in one frame to the same object in another frame.
- **Appearance and Disappearance** – Objects can move in or out of the frame unpredictably.
- **Occlusion** – Objects in some frames are partially or completely occluded when the any new object appears in front of it.
- **Identity Switches** – when two objects cross each other we need to discern which one is which.
- **Motion Blur –** Objects in the video may look different because of their own motion or else movement in the camera.
- **View-Points –** Objects may look different from different viewpoints, but we must consistently identify them irrespective of the viewpoint.
- **Scale Change –** Due to camera zoom option the size of an object can change drastically

- **Illumination –** Lighting changes in the video can affect the way that the object looks which makes it harder for object tracking.

Many other issues have made robust object tracking very challenging. Low quality camera sensors which results in low frame rate, low resolution, color distortion. Challenging factors such as non-rigid object tracking, tracking varying number of objects, and complicated pose estimation. Realtime processing requirements. Object tracking across cameras with non-overlapping views. Object variations in terms of its appearances. All the above factors may cause tracking degradation and even failures.

Research has proposed various appearance models with visual representation and statistical modelling techniques to tackle these challenges. In general, these appearance models focus on various problems in visual object tracking and therefore have different properties. Thus model-based tracking assumes that some pre-existing data from the tracked object can be used to generate an estimate observation in a predicted position as a model. By comparing the true observation with the foresight, the status of an object can be evaluated. Generally, the correlation involves different analysis of the corresponding features extracted.

Tracking an object in a video can be expressed as the process of finding the motion path of an object over time, by locating its position in every frame of the video. Object tracking has emerged as one of the most popular research subjects within the area of computer vision. Although it is well studied problem, it remains a challenge in many of the aspects. Building the tracker for the generic objects in one of the most difficult part in this part, because they change their appearance with the change in the illumination, noise in the video, complex motions or complex shapes of the objects.

Imagine looking at a moving arbitrary object that changes suddenly its movement pattern or aspect such as shape or color. This together with multiple other objects in the scenery that might behave in the similar way keeping track of the objects in such cases would be most difficult thing even for a human eye. Keeping this in mind it suggests that it is also the most impossible task for computer to manage. But it can be achieved by applying some constraints to the simple tracking algorithm such as motion of the object being tracked is smooth i.e., with no sudden changes. Another way of simplifying task is having the prior information about the object along with the constant velocity or acceleration.

The process of building an object tracker is usually divided into several steps, which are object representation, object detection, and object tracking. Even though our project deals with the object tracking but the object representation and object detection will be helpful in building successful object tracking algorithm.

- **Object Representation: -**

To able to track any object that the user is interested first need to be represented in a way that its computer can understand. Properties such as appearance and shape are usually used as basis or the representation. Either appearance representation or shape representation can be used alone to represent the object, but in some both of them are combinedly used. Some of the external factors like application domain, purpose, and goal determines how the object should be represented. Furthermore, it helps in suggesting the best suitable algorithm based on the object representation.

$$\boxed{\text{Object Representation} = \text{Shape Representation} + \text{Appearance Representation}}$$

- **Shape Representation: -**

There are several ways to represent an object's shape so that the measurement can be made for object detection and object tracking. However, some part of representation is suitable only for certain kind of objects only. Each and every model has it own advantage and disadvantage, and these models vary depending on the application domain and object type. Common ways to use shape representation includes points, geometric shapes, silhouettes, contour, articulated shape models, and skeletal models.

i. Points: -

The object of interest is represented either as a single point or set of points. These set of point representation can be difficult while representing multiple object in the same video frame which results in partial or full occlusion. This points representation can cause misdetection because it is difficult to keep track of such points. Finally, this point representation is more suitable for small objects that can be done by using a single point.

ii. Geometric Shapes: -

Primitive geometric shapes such as rectangle, ellipse are used as a shape representation. Both the rigid and non-rigid objects it is a common approach, even though it is more suitable for rigid object representation. Moreover, objects in the video are not as simple and exact as the type of shape. So, it is common that parts of the object are left outside, or part of the background included in the shape results in the tracking problem.

iii. Silhouette and contour: -

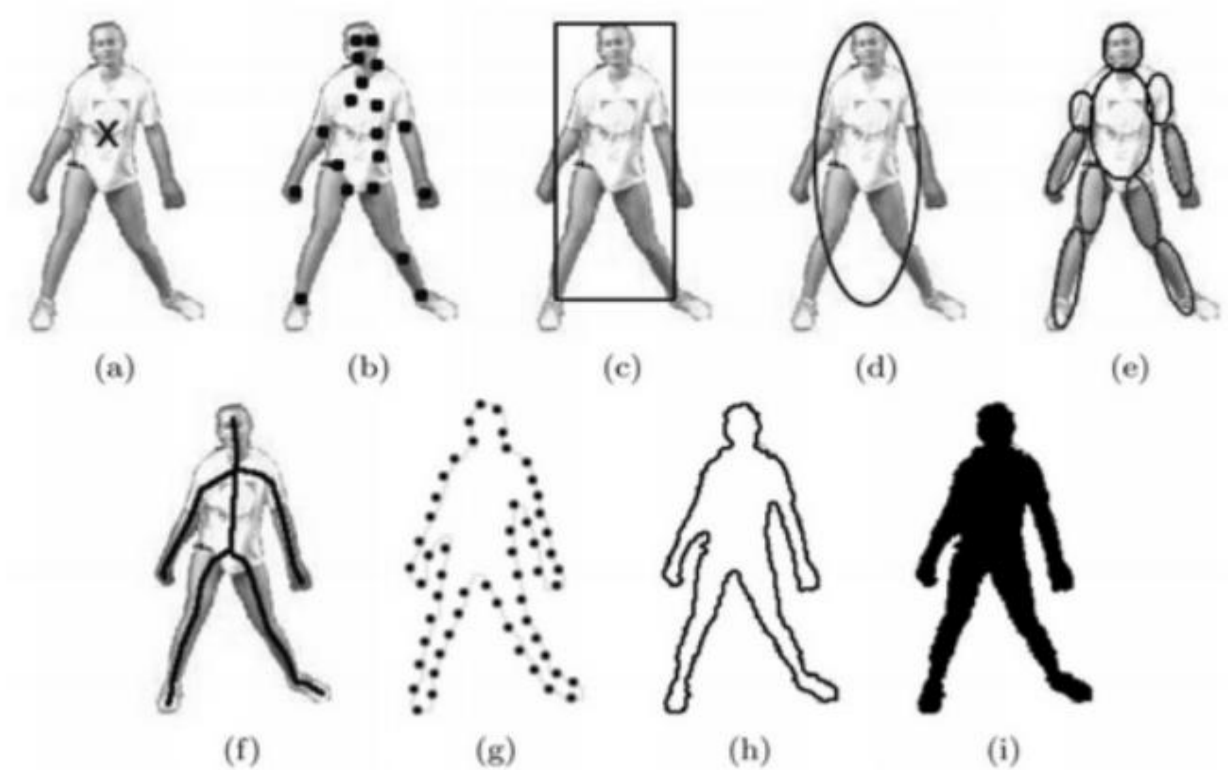
Representing the object by its outline or boundary is known as contour representation and silhouette means the region inside the contour. It is flexible model for representing many objects of different shapes even the non-rigid objects can also be represented easily.

iv. Articulated Shape Models: -

An object is called an articulated object, if it is structured by different parts that are held together by joints. For instance, it is used to describe a human being with parts such as face, arms, etc. Every element can be represented with a simple geometric form such as ellipse.

v. Skeletal Model: -

Using the silhouette of an object and applying medial axis transform makes it possible to extract the object skeleton. This is a model commonly used in object recognition, but it is very rarely in the literature part.



Different approaches in shape representation.

- ## Appearance Representation: -

Similar to the shape representation there are numerous methods for representing an object by appearances. Some common ways of representing an object are using probability densities, templates, active appearance models, and Multiview appearance models.

i. Probability densities of object appearance: -

A probability density function defines the probability that a given event will fall within that broad range of values. By using the interior region of an image defined by the shape model, an estimate of the probability densities of the appearance characteristics of an object can be determined, for example by a contour. The appearance features can be texture or a color. The probability density estimation can be either gaussian distribution or non-parametric such as histogram.

ii.Templates: -

For templates we use normal geometric shapes or silhouettes. The advantage of this template method is that it can carry both the spatial information and appearance information. But it has disadvantage as well it can create a problem with the object that look different when viewed from another view point since it encodes the objects from one point of view only. Another problem will arise when there are changes in the features while tracking the object one of such kind of example is illumination change.

iii. Active appearance models: -

A set of landmarks that reside on either the object boundary or inside the object region defines the shape of the object. The appearance of the object is parallelly modelled based on the vector that has been stored for each landmark. This model however requires the training phases where it learns the shapes and associated appearance form the sample datasets.

iv. Multiview appearance models: -

Unlike templates this model encodes objects from different views. It can be by using different approaches, one of them is by generating the subspace from the given views. Principal Component Analysis and Independent Component Analysis are used for this approach for obtaining the subspaces.

➢ Feature Selection: -

An important part in the object tracking is how to make an object distinguishable from other objects. Basically, this process is done manually. Object representation choice strongly affects the feature selection as both of them are closely related.

i. Edges: -

For a human eye it is normal to detect the boundaries between the object and background as there will be change in the intensities in an image at the particular boundary point. Edges used as representing features that allows the algorithms to track boundaries in an object to recognize the change in the intensity. Moreover, edges are less sensitive to changes in the illumination.

ii. Optical Flow: -

Refers to the apparent movement of the pattern of illumination in a visual scene, often referred to as a motion feature. Optical flow is a visual phenomenon that is witnessed every day by almost every human being. While driving a car and looking out the window, it appears that outside things like trees and building are moving backwards. This visible motion is the flow of optics. It is achieved by defining each pixel movement between frames to determine the apparent motion. In these calculations, the brightness constraint is used, which means that the corresponding pixels are consistent in brightness in different frames.

iii. Color: -

The information can be stored in several color combinations from various frames. In the image processing, color is usually represented by RGB, but sometimes by other colors like YCbCr and HSV. One problem with using color as a representation of the feature is the sensitivity to changes in illumination, since the illumination factor directly affects the apparent color of an object. Besides illumination, the apparent color is also influenced by the object's reflectance properties.

iv.Texture: -

Properties such as smoothness and consistency are important, and the surface intensity variance can be calculated. To generate the descriptors that come out in various forms, a processing step is needed. One of the texture descriptors is Laws texture measurements which contain philtres that match the level, edge, spot, wave and ripple.

➢ Object Detection: -

The video sequence must first be detected and identified in order to track an arbitrary object. Regardless of the method of monitoring, an object detection mechanism of any kind is always required. Two different approaches exist to this, however. The detection can be achieved either in each frame or when the object is first shown in the video. Many different methods for object detection are available.

i. Point Detectors: -

Interested points in an image are detected by using this method. Under any circumstances like changes in the illumination and camera view point the point of interest should not be changed. We generally use corner and interest point more in the definition which leads more confusion to the readers. Intersection of two edges is the one way of defining the corner. Similarly, point of interest can be termed as point in an image which can be easily detected. So, from both the definition above point interest can be on the line edges and the same point can also be corner. There are many different detection methods for points and corner, in which the Harris detector is one of the most popular methods. The detection is done by measuring the own values of a specific matrix that can be seen in a small image region as a scatter matrix of the color gradient. Corners are identified by using a minimum value measure. Although methods for so-called corner detection exist, they are generally applicable to interest points in practice and thus are classified as points detectors.

ii.Background Subtraction: -

Foreground Objects are separated from the background by using background subtraction technique. The basic approach is in this method is initially we build a background model that represents the completed scene and then it is used as a reference for further frames. Now each and every frame is compared with the background model in order to detect the changes in the frame. By comparing each and every frame moving objects are detected in terms of deviation from reference model.

Algorithms used for background subtraction are considered as simple and straight forward to use however they are very sensitive to change in the environment. It is further classified in two kinds recursive technique and non-recursive technique. In recursive technique the background model in every frame is updated recursively, so it can affect the input frames processed in a distant past.

Non-recursive technique stores a buffer with the last n video frames, the context template is then based solely on the frames stored in the buffer. The relevance of the context template is modified faster, but the memory demand increases as the buffer size increases. Recursive techniques requires less storage compared to non-recursive technique, but it leads to eventual errors in the background model for long period.

iii.Segmentation: -

Segmentation is the process of partitioning an image into multiple regions, and that particular segment will collectively cover the complete image this is helpful for locating objects and boundaries.

Figure 3: A segmented image [48]

There are different approaches in this method. One of the main examples of this kind is active contours and its main purpose is to find the object outline in an image. Active contour alone cannot solve the problem of finding contours in an image. The method relies on other processes, e.g. user communication. This is because awareness of the desired contour shape is needed in advance

iv.Supervised Learning: -

It is possible to obtain object detection by learning different object views automatically instead of storing a full set of models to represent different object views. Training under supervision uses a set of predetermined examples to create a function that maps inputs to the desired output. In other words, the computer is provided with the predefined examples which has defined answer given by the human, then computer uses it for predicting the output of the unknown models. There are several learning approaches such as neural networks, adaptive boosting, decision tree and support vector mechanism.

v.Temporal Differencing: -

Temporal differencing identifies difference between two or three consecutive frames. However, this method fails if the movement of the object is very slow as it fails for extracting all relevant pixels. It stops detecting the object if there is no movement in the object detected with the frames.

➢ Object Tracking: -

There are many algorithms that are used for tracking objects some one can only handle for tracking the simple or single objects while others are used for handling the occlusion occurrence in case of multiple object tracking. Algorithms can be categorized in to three kinds such as point tracking, kernel tracking and silhouette tracking. Other way of categorization for object tracking methods are region-based tracking, boundary-based tracking and contour-based tracking.

Region based tracking describes the process of object tracking based on the complete region of the object. Sometime it is also called as silhouette tracking. And similarly, both the boundary-based and contour-based tracking is done by using the border between the tracked object and the background.

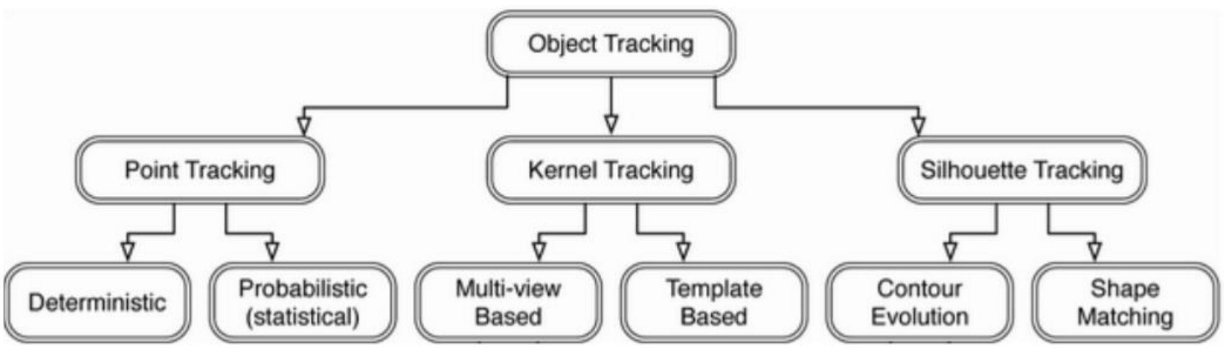


Figure 5: A common categorization of object tracking techniques [48].

i.Point Tracking: -

The point tracking approach is used when the object representation is done in the form of points. Tracking is always performed in terms of their position and motion by evaluating their state.

ii.Kernel Tracking: -

Kernel Tracking approach is based on the measurement of an object motion defined by the primitive region of the object. The next position of an object can be determined by the movement model, i.e. the movement from a frame to another frame. Different parts of the estimation is important for the tracking. For instance, when using the trajectory to analyze the object behavior only motion is needed. However, region near the object also becomes important while identifying the object is needed. The meaning of the kernel is always referred as look of the object and it is always represented with the help of geometric shapes like rectangle or ellipse.

Kernel based tracking can be further classified in to two kinds multi-view-based appearance models and template-based appearance models. The template-based appearance are also termed as single view-based approach.

Figure 9: Tracking objects using simple geometric shapes [40]

Template based appearance models are considered as a straight-forward to use and it is also not as complex as other models are this is what it made as one of most widely used appearance model. However, template-based tracking differ based on the number of objects being tacking that can be either a single object or multiple object.

Single Object Tracking is used for tracking simple and single objects, the most common approach is template matching. Different features like color or intensity are used to form this template. The basic function of this approach is for searching the specific template pattern which has been obtained from the previous pattern. This method is basically used because of its straightforward approach and flexibility. Based on the object in the previous frame it has search only limited area in the next frame. For an image of size n and template size as k it has to perform kn straight forward template matching implementation. Even there were development in the algorithms which can perform the task in n operations.

Multiple Object Tracking – Tracking multiple objects makes more complex tasks where the factors like interaction between the objects need to take into consideration. One kind of interaction involves an object occluding an object in part or entirely. Single object tracking modelling does not consider these problems. There were various suggestions to solve these problems. One of them is to consider the image as a set of layers where the number of layers depends on the number of objects being tracked it also contains the background layer and this background layer is used for compensating any background motion so that an object motion can be calculated from the compensated image. Another suggested way is by using Bayesian decision theory to track and detect the overlapping of the objects. This done by using color intensity and color histograms as a feature representation. The pair is considered a match when the matching value is greater than a certain threshold value and the templates that have been used to track are updated. If the score is below the threshold, the item will be checked further for the occlusion. And it doesn't work if the objects occluded completely. However, if one of the scores is low and other is high, one piece of object is still considered as a match although the model is not modified.

Multiview Based Appearances- For instance in the previous section we have used a template as a reference which is usually based on the latest observation of the object. Here the template is formed just form one point of view only, but this can be a complex problem for the objects that look different from other view-points. Occlusion and objects leaving the frame may also cause the problem. The approaches are used to consider the topic from a different perspective and thus to solve the problem. One way to do this is to use a subspace-based approach to calculate the transition from an image of an object to a model reconstruction with vectors of an object.

iii.Silhouette Tracking: -

It is also termed as region-based tracking. For some complex objects like hands or heads, representing objects with simple geometric forms may be insufficient. And tracking may be insufficient because of bad object representation, but by representation of an object such as a human with simple formations such as cylindrical and skeleton models will make the tracking more feasible. The object models used in the tracking are created using the previous frames and can be in the form of a color histogram, an object edge or object contour.

Figure 10: Tracking objects using object silhouettes [21]

Shape Matching is used for finding the similarities between two objects which means that similarity measure method is the mandatory in this part. This can be performed similar to the template base approach and algorithms used in this part depends on the required properties and matching problem.

Contour Evolution is also termed as boundary tracking as it uses the contour from the previous frame as an initial contour in the current frame. This method uses edge-based features which are insensitive to illumination changes.

4.Project Design: -

We have two types of execution in this project to select the method from the user

a.) Offline Tracking using local files

b.) Online Tracking using Webcam

a.) In offline tracking user has a wide range of choice to select the video from his desired local system and run the program to detect objects. In this method the file from the user is selected and objects or the people present in the video are detected and the moving objects in the video file are tracked all along the duration until it occludes or loses its path.

b.) In Online method the user selects the webcam to track the video from and in this, the objects present in front of the webcam are detected and , if there are any moving objects in front of the camera those objects are detected and tracked up to a certain distance.

5.Methodology:-

Object tracking for real-time video or web camera, based on the results from object detection, is aimed to estimate the optimal trace of the moving objects. If there is one moving object, then the tracking algorithm works flawlessly. But per suppose, if there are multiple moving objects, then overlapping may occur with each other; the tracking problem becomes more difficult. When various objects separate from each other after occlusion, we particularly need to find and match the correct location for each object, making it possible to detect and track multiple objects all the time.

CAM Shift:

It is briefed as Continuously Adaptive Mean Shift Tracking. It is based on an adaptation of Mean Change, which finds the mean (mode) of the distribution in the direction of maximum increase in probability density, given a probability density image. Like Mean Shift, which uses Static Distributions, it uses continuously adaptive distributions of probability. It is one of the simplest approaches which offers accurate and stable performance if the background colours are significantly different from those in the target object.

Steps included in this algorithm are stated as follows:

1) Determination of the interest region of the target image in current initial image.

2) Automatic selection of two colourimetric channels.

3) In each region constituting the target, calculation and affectation of the average colour to all the corresponding pixels.

4) The two-dimensional histogram is calculated with the colourimetric channels selected in 2.

5) Back-projection of this histogram with the initial image +1 in order to obtain the probability distribution image.

6) Application of the MeanShift algorithm on this image to determine the new target centre in the initial image +1.

7) Determination of regions constituting the target image +1 by using the same averages calculated in 3.

8) To take into consideration changes of the object, we may need to go to step 2; otherwise, go to step 4.

The advantages of mean-shift algorithms are:

(1) It makes no model assumptions, unlike Gaussian mixture models or K-means

(2) It can track any complex objects.

(3) The user needs only set only one initial parameter, the bandwidth, which has an intuitive physical meaning of local scale, and this determines the number of clusters automatically. This is often more convenient than having to select the number of groups explicitly.

(4) It has no local minima. Thus the clustering it defines is uniquely determined by the bandwidth, without the need to run the algorithm with different initializations.

(5) Outliers, which can be very problematic for Gaussian mixtures and K-means, do not overly affect.

The CamShift algorithm calls upon the MeanShift one to calculate the target centre in the probability distribution image. It is a matter of finding a rectangle presenting the same moments as those measured on the probability image. These parameters are given from the first and second moments.

The mean shift is a non-parametric feature space analysis technique. It is a method for finding local maxima of a density function from given discrete data samples.

Mean Shift:

The mean shift concept is a possible forward-tracking technique because it estimates the positions of the regions in the current frame from the previous frame. A mean-shift algorithm is a method to track an object of interest as it passes through a series of images. It is a gradient method that models the area of the image to be controlled by its histogram of colour. The mean shift is a tool for non-parametric analysis of space features. The mean shift technique is an independent application tool.

It is suitable for real data analysis because it does not assume any prior shape on data clusters, which we have used in our webcam real-time object tracking.

Kalman Filter:

R.E. Kalman wrote his renowned article describing the discrete-data linear filtering problem in 1960 as a recursive solution. From then, the Kalman filter has been the focus of extensive research and development, particularly in the area of autonomous or aided navigation, due to improvements in digital technology. Kalman filter has been the subject of extensive research and application, especially in the field of autonomous or assisted navigation, which is developed using a CV, AI, Deep Learning. This filter helped in large parts of developments in digital computing.

The design of Kalman filter is relatively simple and needs minimal computational power. Nonetheless, recognizing and enforcing the Kalman filter is still not easy for people who are unfamiliar with estimation theory. Kalman filter predicts the object position in the next frame by maximizing the posteriori probability of the object position based on the previous estimations. Then the result of the Kalman filter prediction is used to improve the performance of the tracking method. KF is widely used to track moving objects, with which we can estimate the velocity and even acceleration of an object with the measurement of its locations.

Kalman filtering (KF) is widely used for tracking moving objects that allow us to estimate the velocity and even the acceleration of an object by measuring its locations

1) The Kalman filter gives solution to find the state of time handled method which is controlled by the linear differential equation.

$$X_{k=} Ax_{k-1} + Bu_{k-1} + W_{k-1}$$

where $$Z_k = Hx_{k+} V_k$$

W_k- Process noise

V_k- Measurement noise

The Kalman filter estimates the system state at some stage and then obtains feedback in the form of measurements. The Kalman filter equations further into two groups i.e time updating equations and equations for measuring updates. The time change calculations are responsible for estimating the present state and error covariance projections for the next time phase forward to achieve a categorical estimate.

The equations of time updates can also be known as prediction equations, while the equations of metric updates can be considered as equations of corrector.

The time upgrade plans are determined ahead of time by the current state. At that time, the approximation adjustment changes the estimate expected by an actual measurement. Following are the equations to update time.

$$x_k = Ax_{k-1} + Bu_{k-1}$$

$$P_k = A_k P_{k-1} A^T{}_k + Q$$

The above equations further effect the covariance factors for time estimates k-1 to k. The below equations are used to update measurement values.

$$K_k = P_k H^T{}_K (H_k P_k H^T{}_k + V_K R_K V^T{}_K) -1$$

$$\hat{X}_k = \hat{x}_k + K_k(Z_k - H(\hat{x}_k, 0)$$

$$P_k = (I - K_k H_K) P_k$$

The covariance of the measuring noise is usually measured before the filter is worked. In addition, thus, we should be able to take certain off-line test measurements to evaluate the variability of the measuring noise.

The process is repeated with the previous a posteriori estimates used to forecast or model the latest a priori estimates after each period and calculation change pair. This recursive existence is one of the Kalman filter's rather attractive features.

The covariance R estimation noise is usually measured before the filter is performed. Measuring the covariance measurement error R is generally practical because the process must still be measured.

Terrain-referenced navigation (TRN), also referred to as terrain-aided navigation (TAN), offers positional information by matching terrain measurements with an electronic altitude model (DEM) mounted on an aircraft's on-board computer. The TRN algorithm combines an inertial navigation system (INS) navigation approach with the calculated terrain profile below the aircraft.

2) It first estimates the initial position of an object in the current frame by recovering the relative object motion between consequence frames. This filter predicts the object position in the next frame, which can help the tracking method to recover movement from the expected position. This filter works by predicting the next position of each object based on its estimated velocity when they depart. Using Kalman filter algorithm, we can detect the central locations of moving objects. Such a central position representation sometimes does not indicate the exact location of a moving object due to the noise and limitation of the detection method. If multiple objects exist, to track each object, the locations need to be further delineated.

Using statistical values of the system noise and measurement noise, the variables are used as input signal, and the estimation variables for tracking object that we need to know are the output of the filter. The whole filtering process is composed of a prediction equation and an

update equation, which also serve to portray the entire system, as defined by update equation.

$X(n) = F.X(n-1)+V_q(n-1)$ (prediction equation)

$Y(n) = H.X(n)+ V_p(n)$ (update equation)

The object estimation of Kalman filtering based of initial object position is calculated when

(1) the state model and measurement model are both linear;

(2) the model is designed well to fit the actual system;

(3) the noise is additive Gaussian noise.

During real-time object tracking as we proposed through webcam implementation, noise occurance is more which makes tracking algorithm complex. Due to which Kalman filtering has been more developed to overcome noise for accurate results. This developed Kalman filtering is known as unscented Kalman filter (UKF). UKF where X(n) and Y(n) are the estimated state variable and measurement variable, respectively. F is the state transition matrix and H is the measurement matrix.

$V_q(n)$ and $V_p(n)$ represent the system noise and measurement noise. Unscented transform (UT) is adopted in UKF to replace the linear transform, thus enabling Kalman filtering to deal with nonlinear systems directly.

UKF

UKF's core idea is to address the issue of non-linear structures and the non-Gaussian noise to some area.

While both KF and UKF are based on the same concept but makes it applicable to non-linear systems. It is worth noting that for many nonlinear applications, UKF can provide best results for all the conditions that KF would, however, applying to a much larger field than KF. Additionally UKF can often provide better estimate accuracy than KF can often, even if the noise is not distributed in Gaussian, UKF may still function, while KF will struggle to provide sub-optimal estimate results for state variables.

EKF:

A Kalman filter is referred to as an extended Kalman filter which linearize the existing mean and covariance. The major problem of the EKF is that, since performing their respective nonlinear transformations, the distributions of the different random variables are no longer natural.

The approximation was rendered in the extended Kalman filter based on a single point i.e. implies. This approximation might not be the best approximation available.

Estimating a system of non-linear relationships of variance and estimation, the following equations correspond to the initialization of finding the linear value.

$X_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}$
$Z_k = H x_k + v_k$

where,

X_k, Z_k are the actual and measurement vectors.
X_k -Posterior estimate

Time Update:
Projecting the state of ahead

$x_k = f(x_{k-1}, u_{k-1}, 0)$

Project the error covariance ahead

$P_k = A_k P_{k-1} A^T{}_k + W_k Q_{k-1} W^T{}_K$

Operation of the extended Kalman filter, combining the high-level

Measurement Update:

Kalman Gain calculation-

$$K_k = P_k\, H^T{}_K\, (H_k\, P_k\, H^T{}_k + V_K\, R_K\, V^T{}_K) -1$$

Z_k measurement updation

$$\hat{X}_k = \hat{x}_k + K_k(Z_k - H(\hat{x}_k, 0)$$

An essential feature of the EKF is that only the appropriate portion of the calculation data is correctly propagated by the Jacobian in the formula for the Kalman gain.

This method has been implemented and verified on our video data set. The experiment results show it successfully tracking objects and improve the efficiency of the tracking method.

6.Code (Functional Part): -

The function first captures the first frame from an image and converts to RGB and then searches for an object. If there is no movement of object in the frame the detection is stopped. And if any movement of object is observed the bounding box is appeared around the object and continues tracking.

i.) **Offline Tracking Method Function**: This part enables the browse option from the system to select the video files and track the desired objects from the required video using the kalman filter along with the gaussian detection for the objects inside the frame.

```
def uploadVideo():
    global filename
    filename = filedialog.askopenfilename(initialdir="videos")
    pathlabel.config(text=filename)
    text.delete('1.0', END)
    text.insert(END,filename+" loaded\n");
```

```
    vc = cv2.VideoCapture(filename)
    while True:
        frame = vc.read()
        frame = frame if filename is None else frame[1]
        if frame is None:
            break
        frame = imutils.resize(frame, width=720)
        (h, w) = frame.shape[:2]
        blob = cv2.dnn.blobFromImage(cv2.resize(frame, (240, 240)),0.007843, (240, 240),
127.5)
        net.setInput(blob)
        detections = net.forward()
        for i in np.arange(0, detections.shape[2]):
            confidence = detections[0, 0, i, 2]
            if confidence > 0.2:
                idx = int(detections[0, 0, i, 1])
                box = detections[0, 0, i, 3:7] * np.array([w, h, w, h])
                (startX, startY, endX, endY) = box.astype("int")
                if (confidence * 100) > 50://Kalman Filter
                    label = "{}: {:.2f}%".format(CLASSES[idx],confidence * 100)
                    cv2.rectangle(frame, (startX, startY), (endX, endY),COLORS[idx], 2)
                    y = startY - 15 if startY - 15 > 15 else startY + 15
                    cv2.putText(frame, "Object detected ",(startX,
y),cv2.FONT_HERSHEY_SIMPLEX, 0.5, COLORS[idx], 2)
                    text.insert(END,"Object detected "+"\n")


        cv2.imshow("Frame", frame)
        key = cv2.waitKey(1) & 0xFF
        if key == ord("q"):
            break
```

```
    vc.stop() if filename is None else vc.release()
    cv2.destroyAllWindows()
```

ii.) **Online Tracking Method Function**: This part activates the webcam button to record the video or capture the objects in the current frame and track the moving objects in it.

```
def webcamVideo():
    text.delete('1.0', END)
    webcamera = cv2.VideoCapture(0)
    time.sleep(0.25)
    oldFrame = None
    while True:
        (grab, frame) = webcamera.read()
        if not grab:
            break
        frame = imutils.resize(frame, width=500)
        gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
        gray = cv2.GaussianBlur(gray, (21, 21), 0)


        if oldFrame is None:
            oldFrame = gray
            continue
        (h, w) = frame.shape[:2]
        blob = cv2.dnn.blobFromImage(cv2.resize(frame, (300, 300)),0.007843, (300, 300), 127.5)
        net.setInput(blob)
        detections = net.forward()
        for i in np.arange(0, detections.shape[2]):
            confidence = detections[0, 0, i, 2]
            if confidence > 0.2:
```

```
            idx = int(detections[0, 0, i, 1])
            box = detections[0, 0, i, 3:7] * np.array([w, h, w, h])
            (startX, startY, endX, endY) = box.astype("int")
            if (confidence * 100) > 50:
                label = "{}: {:.2f}%".format(CLASSES[idx],confidence * 100)
                cv2.rectangle(frame, (startX, startY), (endX, endY),COLORS[idx], 2)
                y = startY - 15 if startY - 15 > 15 else startY + 15
                cv2.putText(frame, "Object detected in video", (startX, y),cv2.FONT_HERSHEY_SIMPLEX, 0.5, COLORS[idx], 2)
                text.insert(END,"Object detected in video"+"\n")

        cv2.imshow("Frame", frame)
        key = cv2.waitKey(1) & 0xFF
        if key == ord("q"):
            break

    #webcamera.stop()
    webcamera.release()
            cv2.destroyAllWindows()
```

7.Requirements :-

We have used Python 3.6 for this project to execute and have used other functions and libraries like tkinter and imutils and have imported OpenCv for object tracking.

- Initially we setup a python environment to access our files and install python in the system.

- After installing python, we import the necessary packages in order to our program to execute without errors.

- Once we import the required packages make sure that a python IDE is available to compile the program, so install a python IDE ex: PyCharm (We have used PyCharm with Python=3.6)

- Within an IDE select the python interpreter that chooses the version of python to run the required programs, you can have multiple interpreters.
- Tkinter is usually directly imported from the python packages by default.
- OpenCV, is imported from the python terminal using import commands or can be installed through selecting packages from an available environment. Ex: Anaconda.
- iMutils is also imported manually during the execution from the default packages present in python.

Python:

Python is a dialect for programming for general purposes. You can therefore use the programming language to build desktop and web applications. You can also use Python to develop complex applications in science and numbers. Python was designed to make data analysis and visualisation simpler with features.

Python had two types of classes before version 3.0: old-style and new style. The structure of both variants is the same, the exception being whether the class object is inherited directly or the other way.

tkinter:

Tkinter is an integrated Python framework used to create simple GUI applications. It is the most commonly used Python GUI module. To initialize tkinter, we have to create a root widget, which is a window with a title bar and other decoration provided by the window manager. The root widget has to be created before any other widgets, and there can only be one root widget.

imutils:
A collection of convenience features to enable simple image processing functions such as translation, rotation, resizing, skeletonization, and view for OpenCV and both Python 2.7 and Python 3.

cv2:
cv2 is the name that OpenCV developers chose when they created the binding generators. This is held to be compatible as the name of the import. It would also be difficult to change the import name or behaviour for experienced users who are used to importing cv2.

numpy:
NumPy is an array-processing programme for general purposes. It provides a multidimensional array object with high performance and tools to work with these arrays.

MobileNetSSD_deploy.prototxt:

We use a MobileNet pre-trained taken from https://github.com/chuanqi305/MobileNet-SSD/ that was trained in Caffe-SSD framework.

Here we are loaded above model through variable using cv2.dnn.readNetFromCaffe() function for reading a network model stored in Caffe framework with args for "prototxt "and "model" file paths.

Here .prototxt and .caffemodel files are part of the OpenCV. They are necessary for today's face detection + tracking method, but you could easily use another form of detection.

Syntax: cv2.dnn.readNetFromCaffe("MobileNetSSD_deploy.prototxt.txt","MobileNetSSD_deploy.caffemodel")

Here prototxt is for architecture model file and model is for pre-trained model.

To run caffe model we need to import imutils, Numpy without fail.

8.Procedure :-

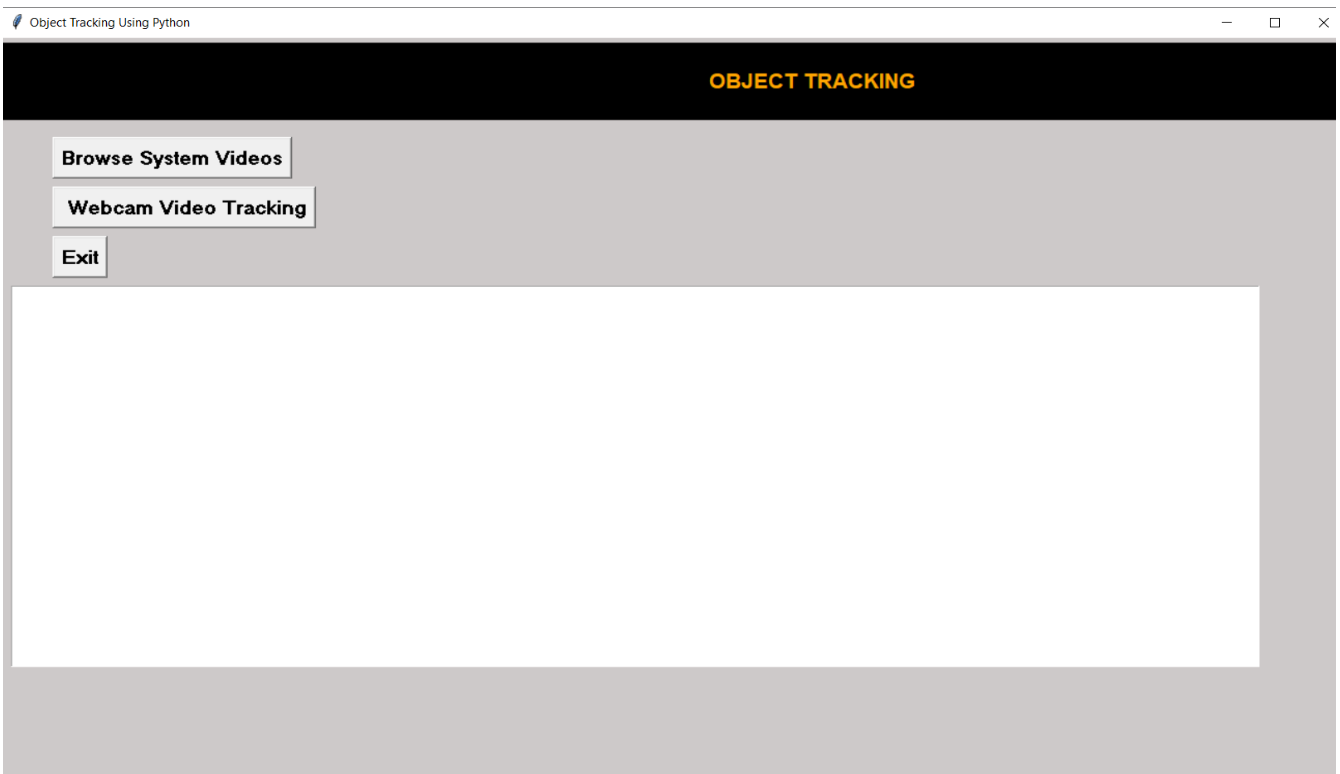
Steps to execute the Program:

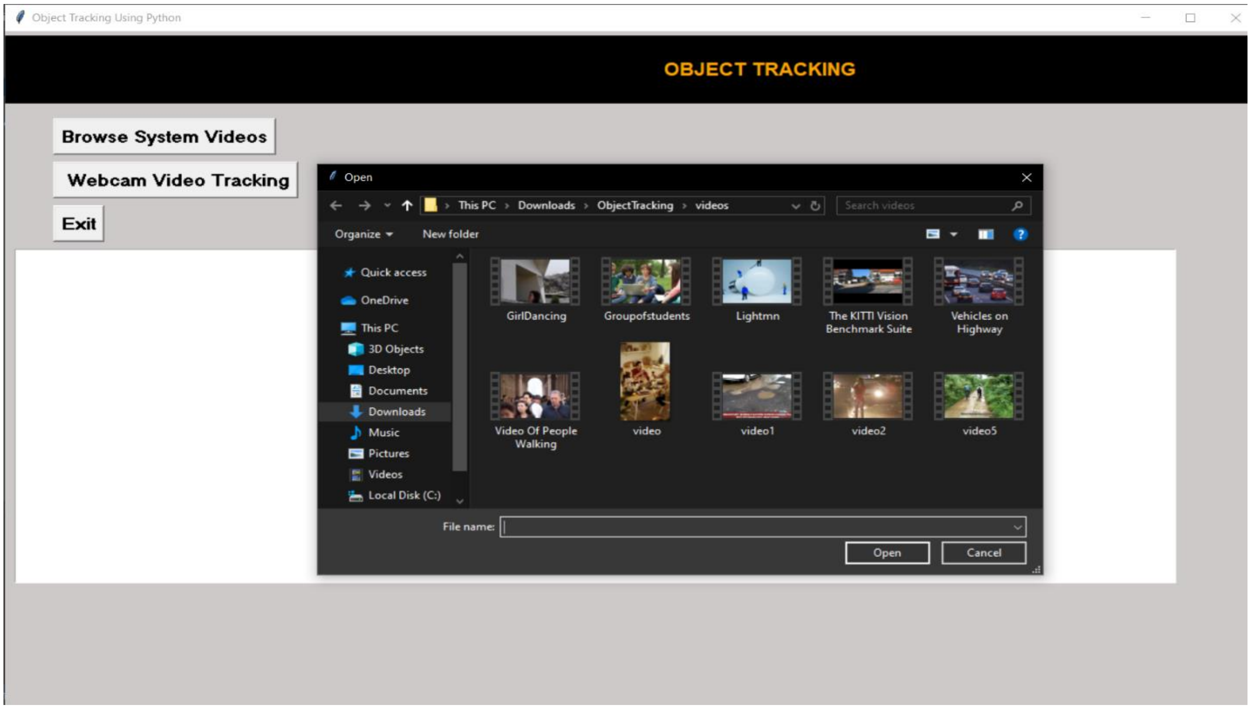
- Extract the files from the folder 'ObjectTracking'.
- After extraction, you can find python file and videos folder which are extracted along with trained data files which are mentioned below that are used by the opencv.
- Open main.py python file using any python Ide that you have installed using the steps described above.
- Once the python file is open in an IDE make sure that the python interpreter is selected to run the program.
- After selecting the interpreter i.e version of the python to run the program on select the python file and click on 'Run'.
- If everything is installed correctly and all the libraries are imported the program would run perfectly and output Python GUI will be displayed.

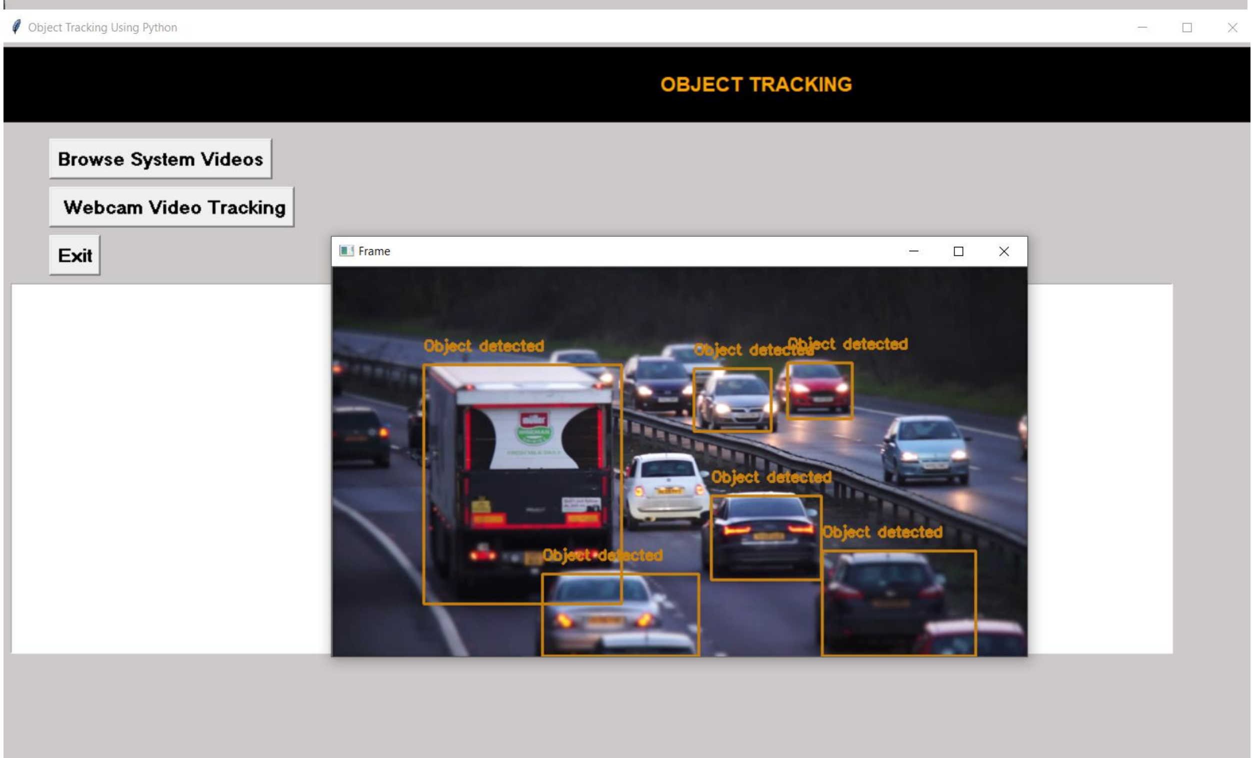
- Python GUI includes the type of output to choose from which is ‘Selecting a file from the System’ or ‘Webcam Video Tracking’ which opens and has acces to the web camera and tracks the detected objects.

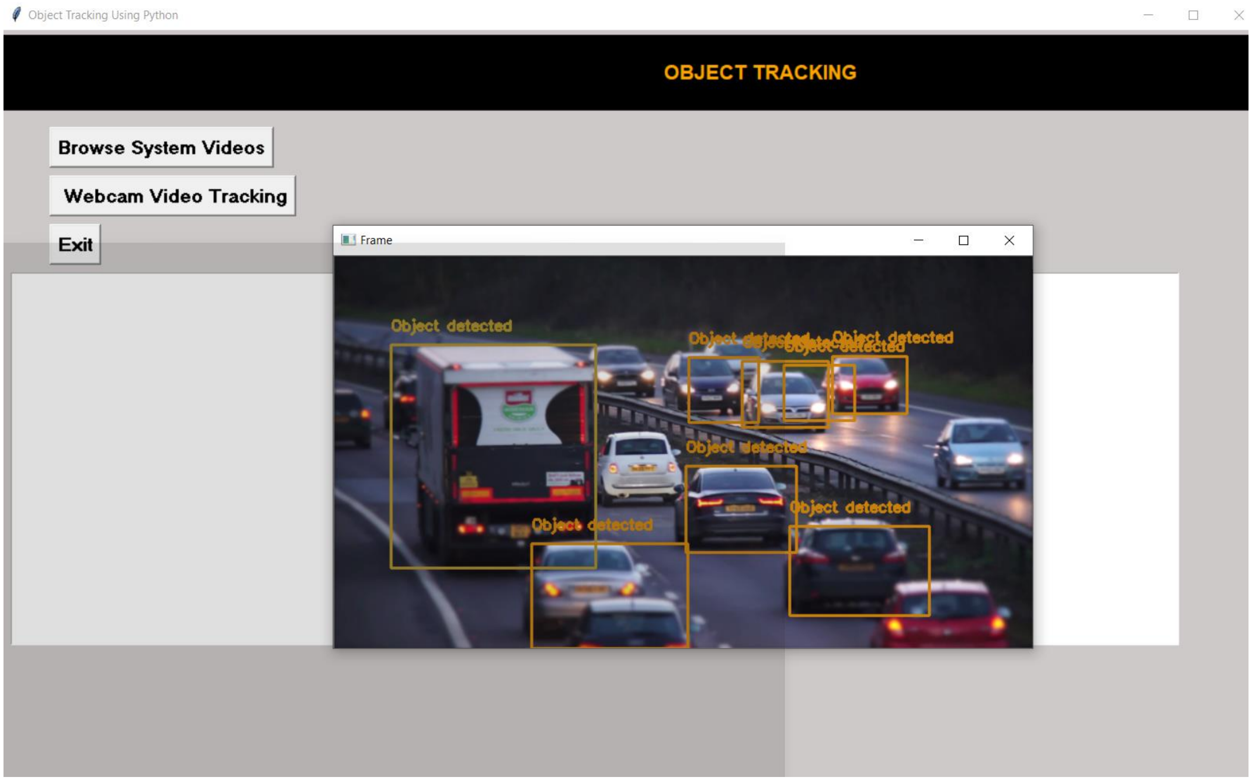
9.Results:-

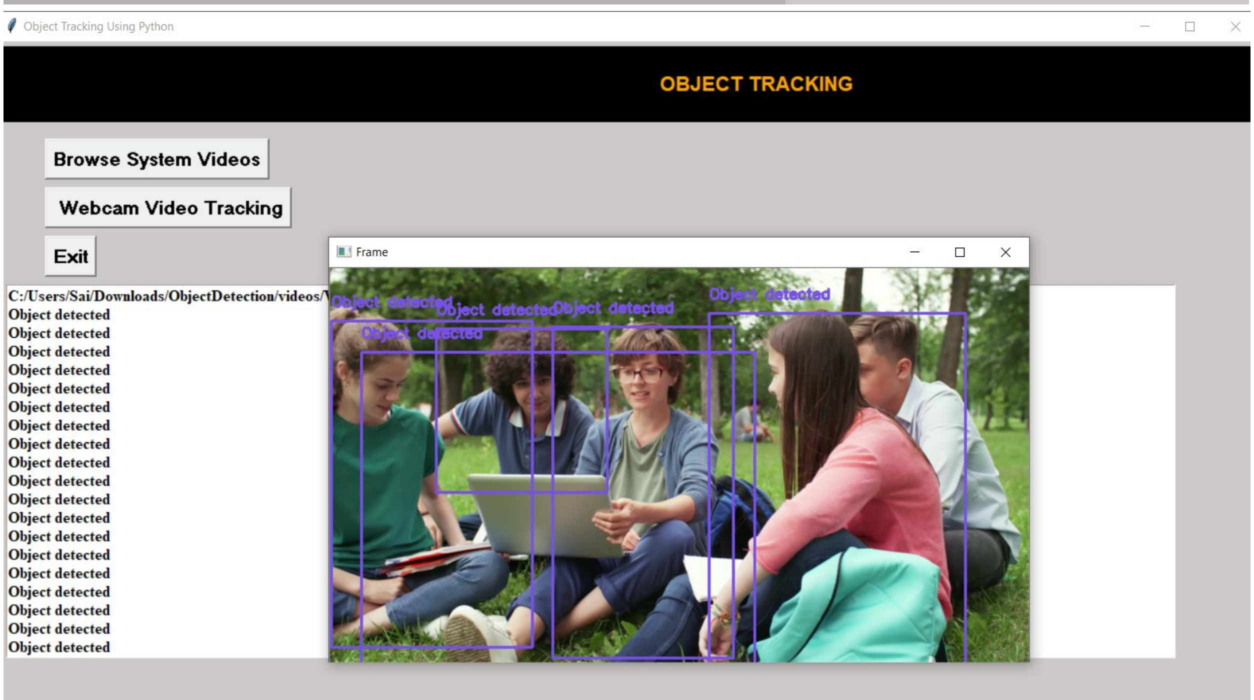
The interface consists of Three buttons one to select videos from the system to track, other to track using webcam and an ‘Exit’ button to close the interface after tracking the videos.

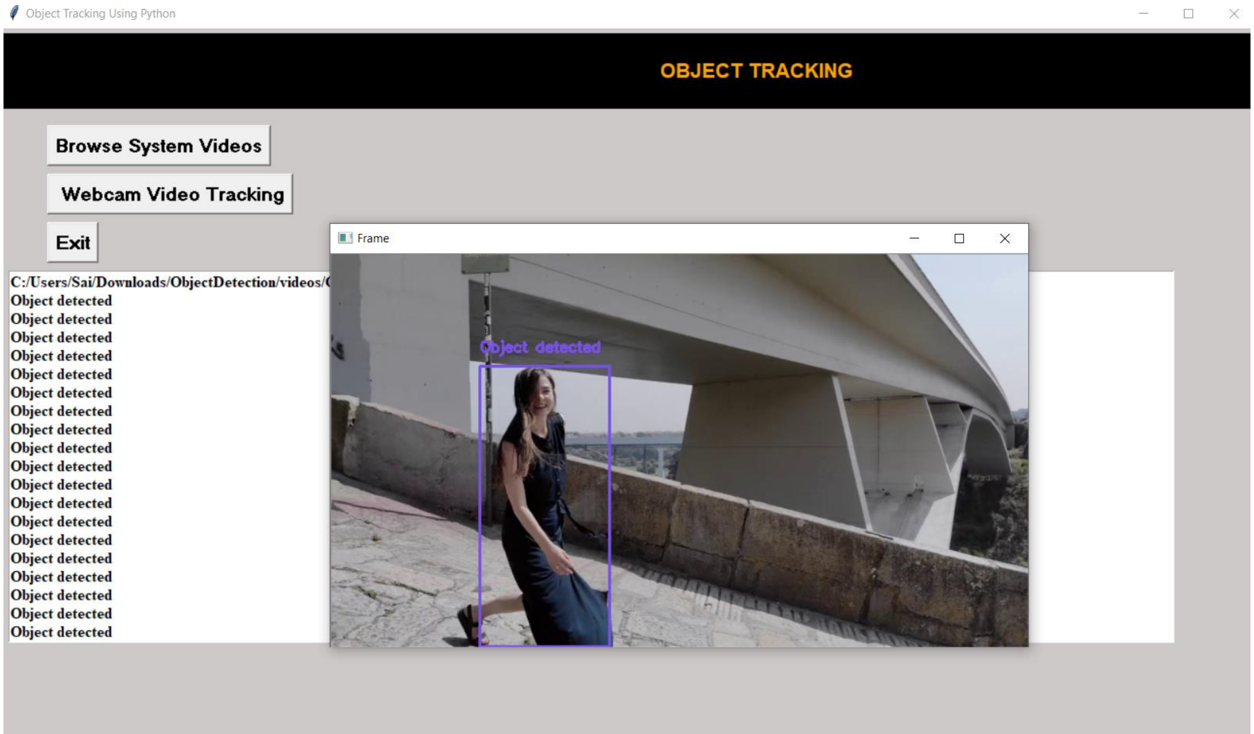


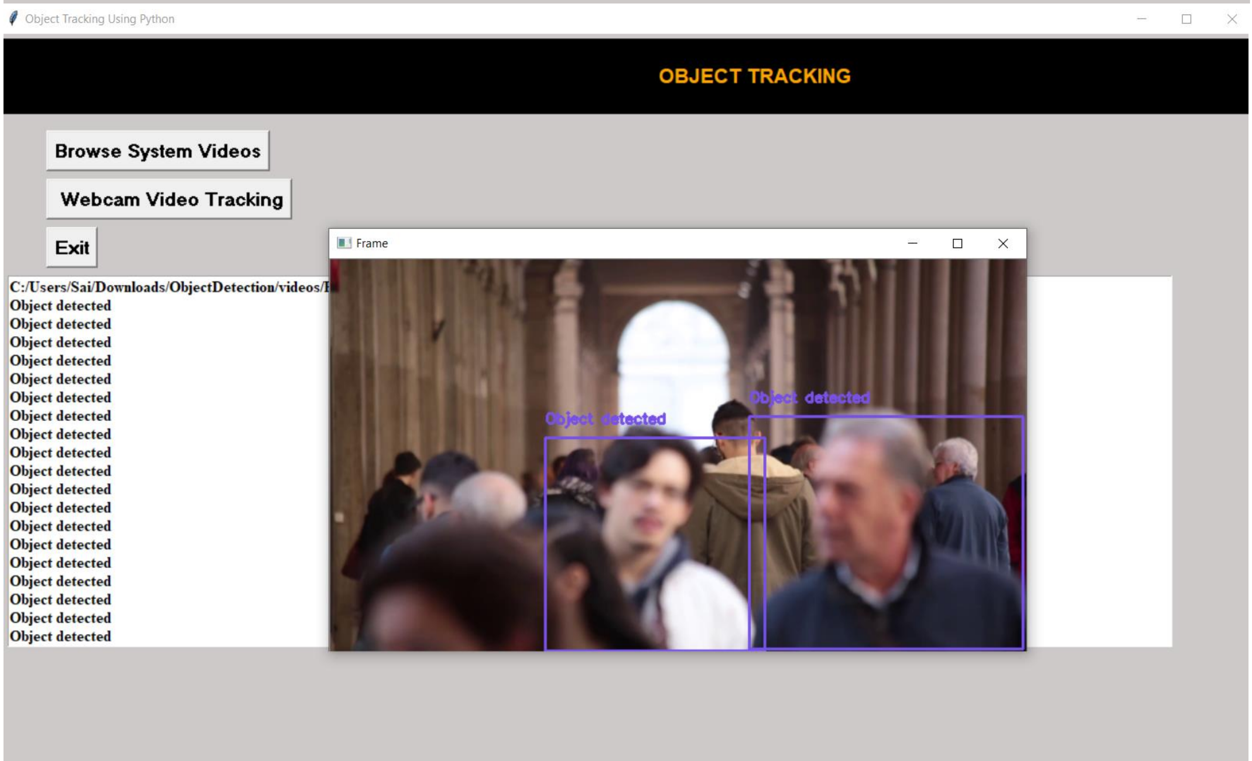

Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
Open
This PC › Downloads › ObjectTracking › videos
Search videos
Organize
New folder
Quick access
OneDrive
This PC
3D Objects
Desktop
Documents
Downloads
Music
Pictures
Videos
Local Disk (C:)
GirlDancing
Groupofstudents
Lightmn
The KITTI Vision Benchmark Suite
Vehicles on Highway
Video Of People Walking
video
video1
video2
video5
File name:
Open
Cancel

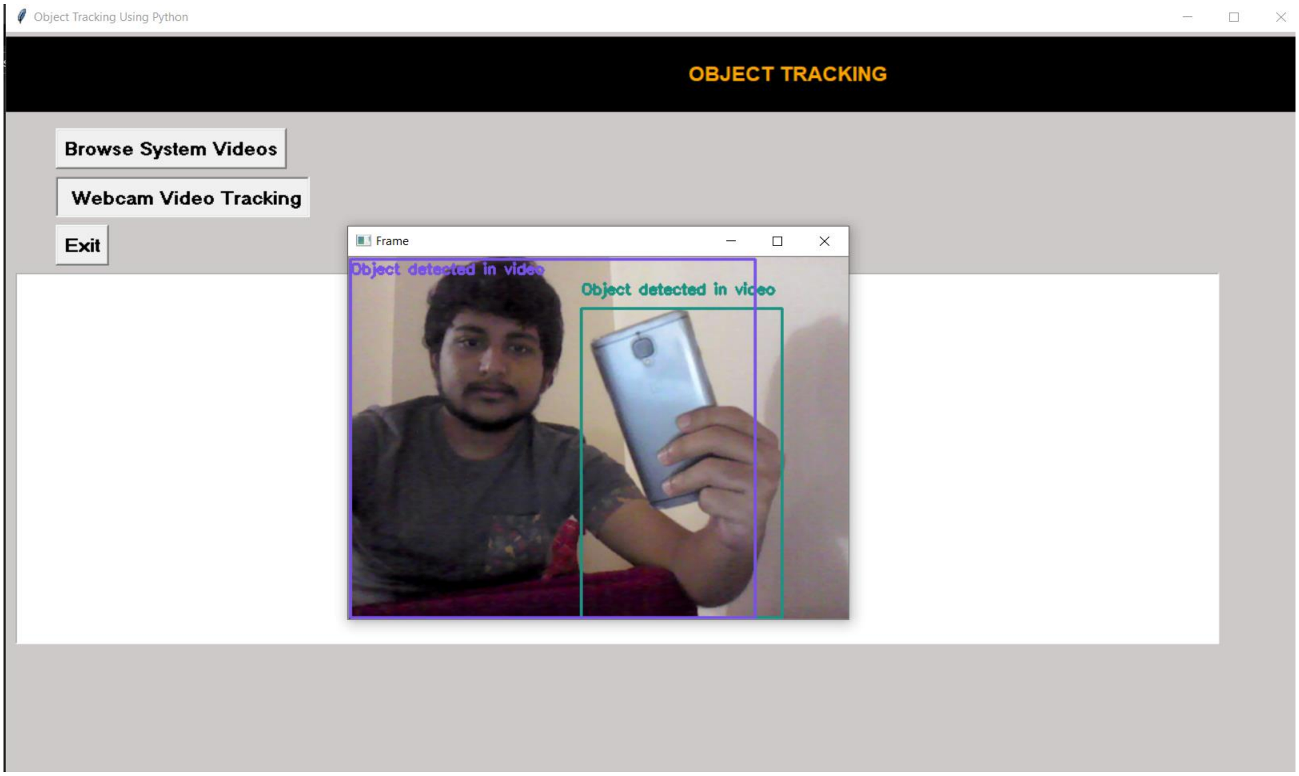

Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
Frame
Object detected
Object detected
Object detected
Object detected


Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
Frame
Object detected

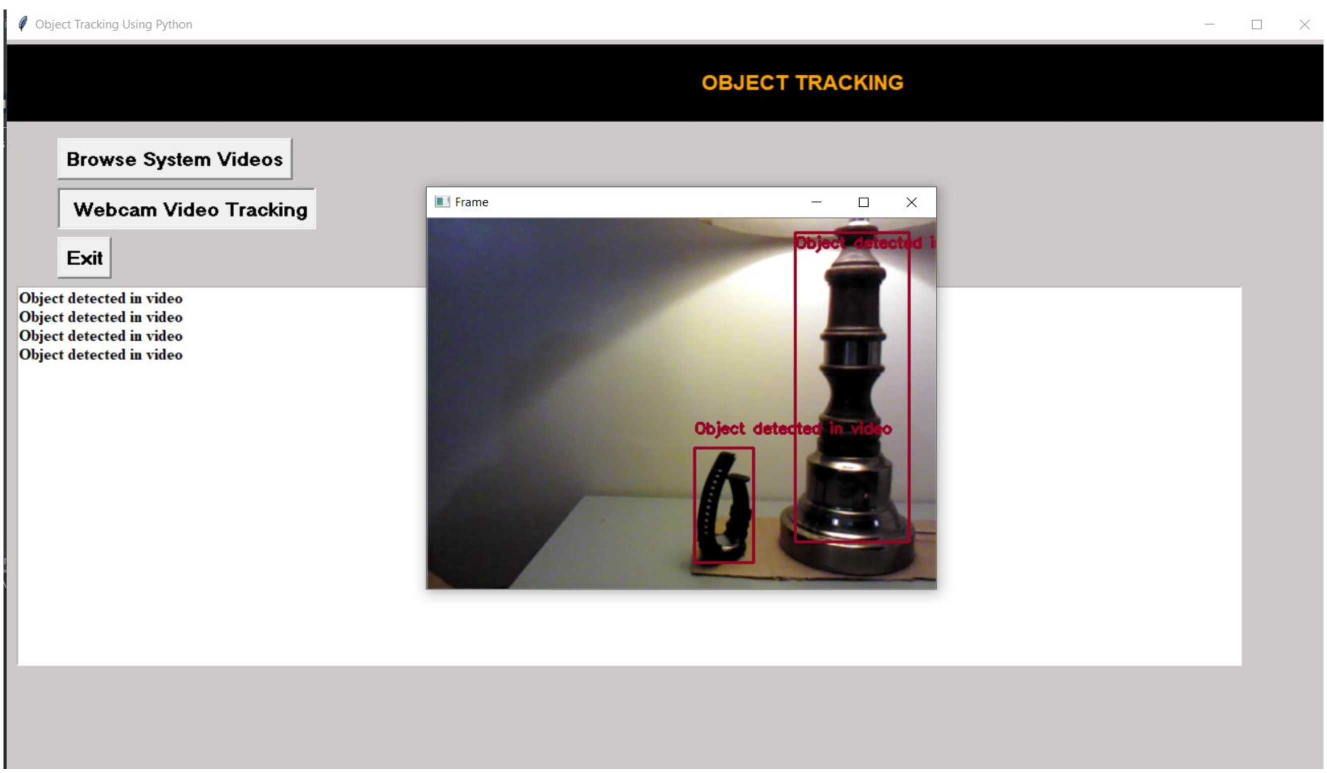

Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
Frame
C:/Users/Sai/Downloads/ObjectDetection/videos/
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected

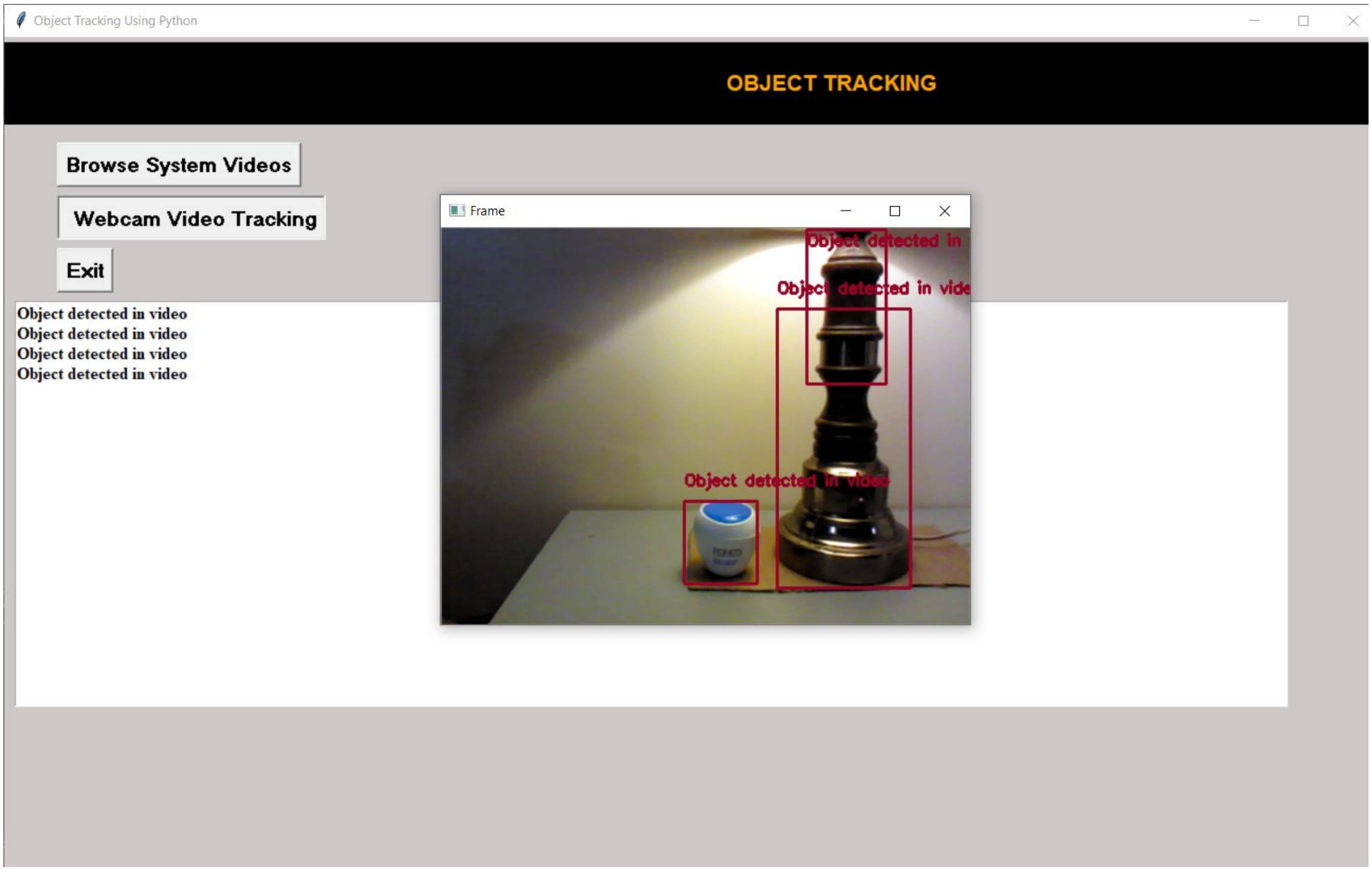

Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
Frame
C:/Users/Sai/Downloads/ObjectDetection/videos/
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected


Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
Frame
C:/Users/Sai/Downloads/ObjectDetection/videos/
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected



These images from the webcam are in low luminance conditions so the objects which are in the light are detected better. If the luminance is high the detection and tracking performance increases accordingly.



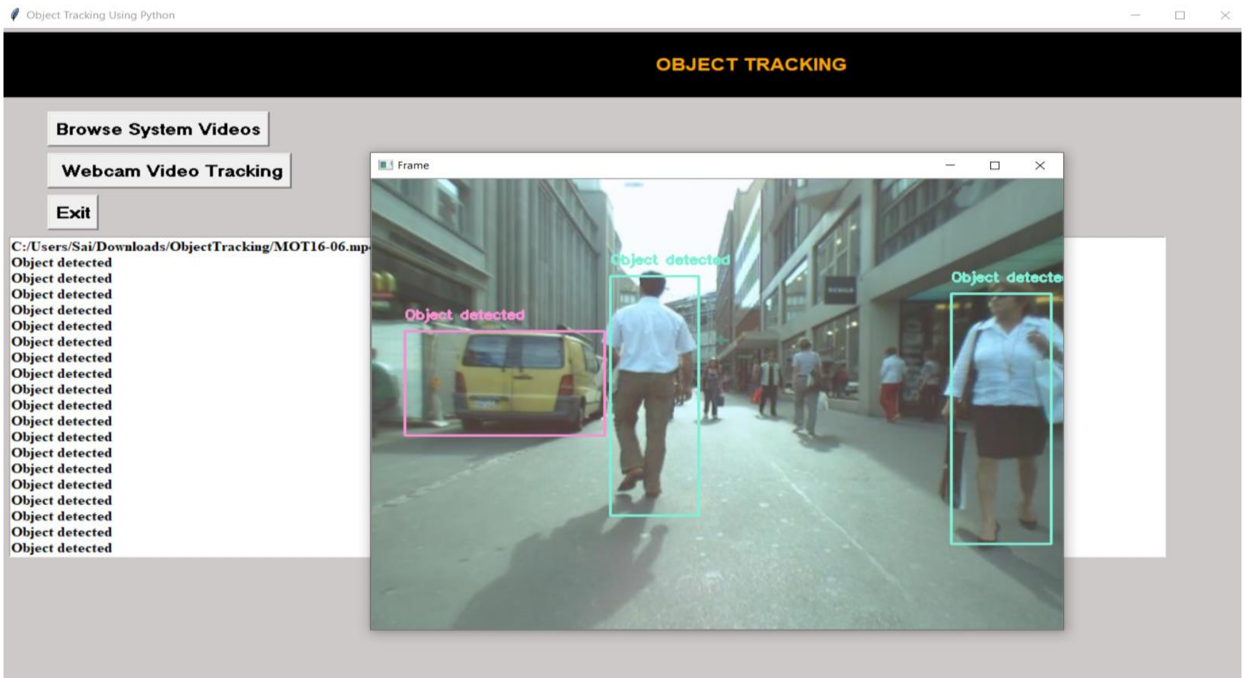


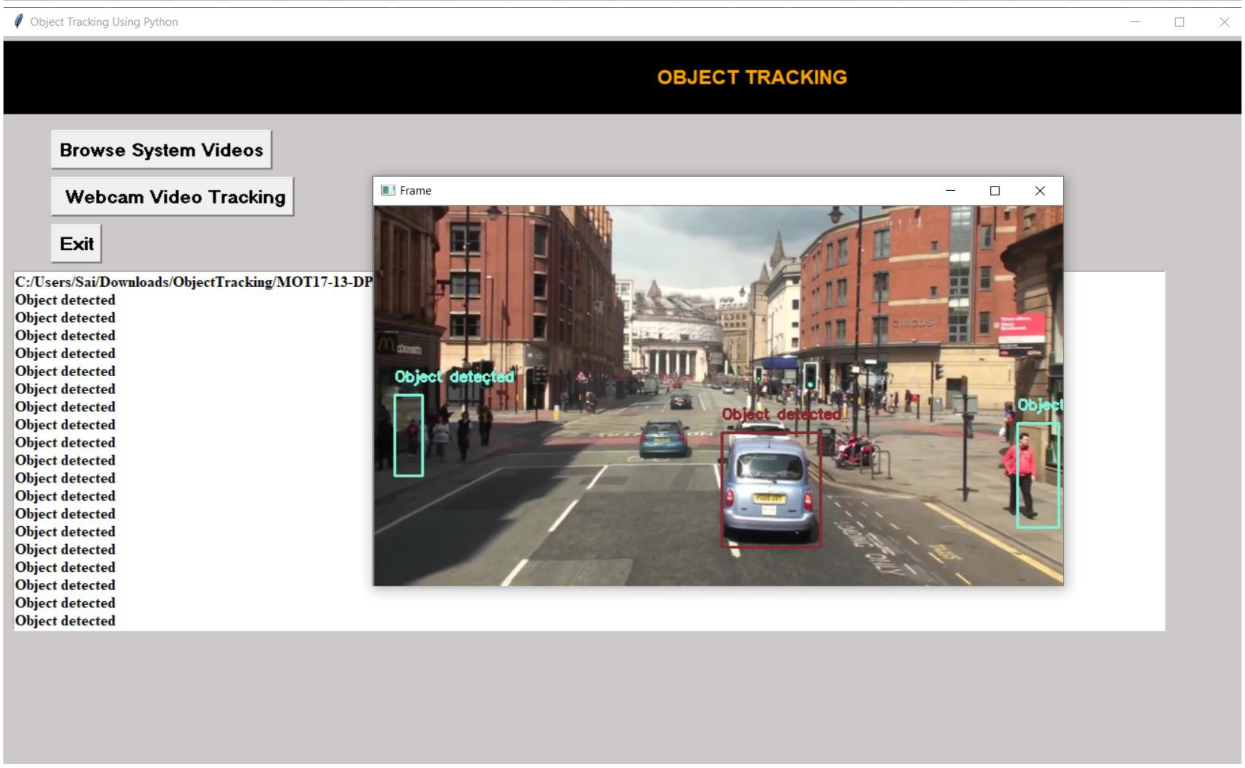
Benchmarking is a path to discover the best performance that can be achieved in a particular company, by a competitor or by a completely different industry. In order to achieve a competitive advantage, this information can then be used to identify gaps in the processes of an organization.

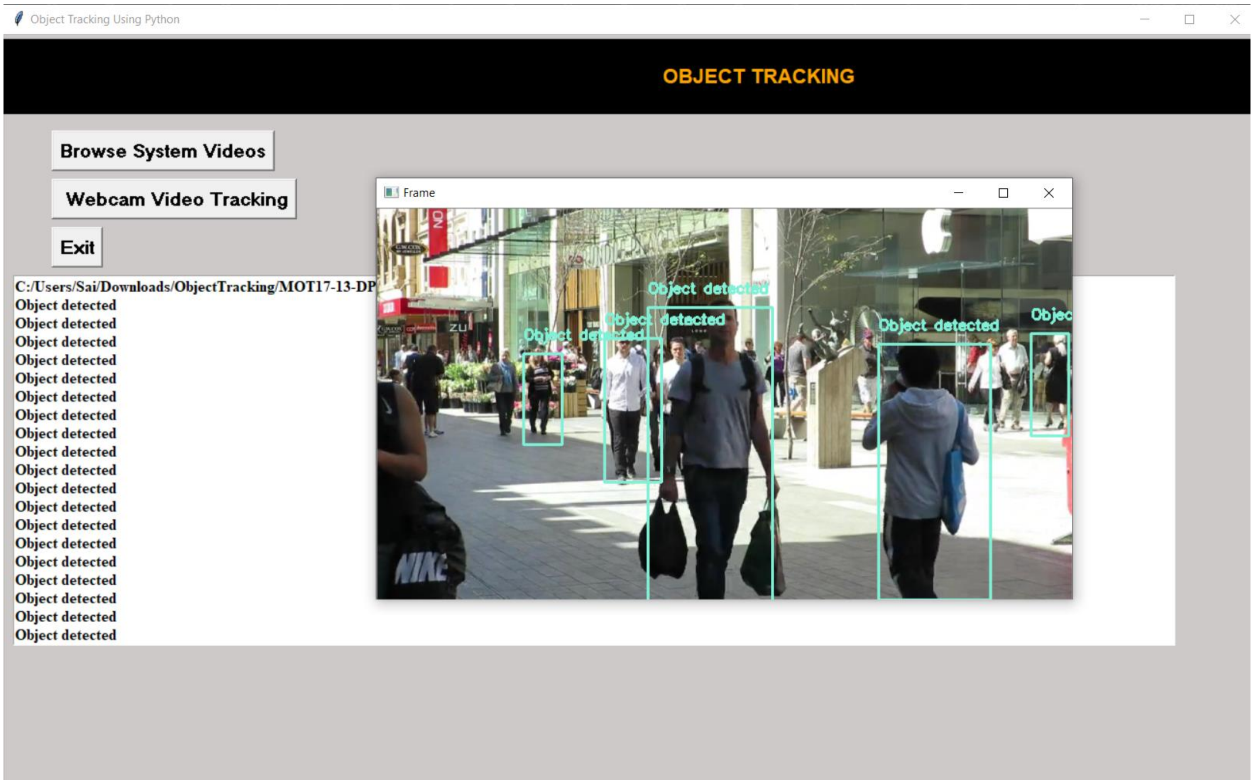
We have used the video dataset from https://motchallenge.net/ which has a large sets of video data including training and test data. The videos in the datasets are examined for various algorithms and the results are posted in their website for different methods. It is a benchmark for comparison of the output by evaluating between alternate run times.

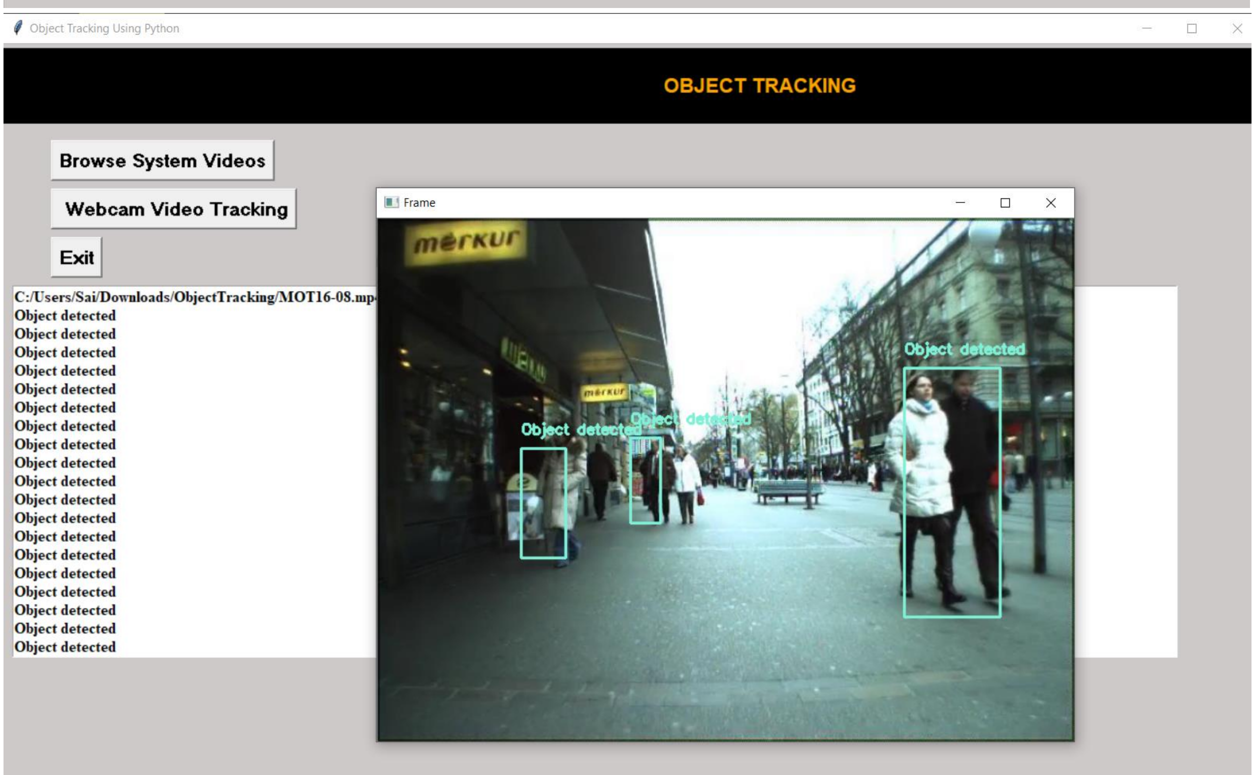
For most computer vision applications, standardized benchmarks are crucial. While leaderboards and ranking tables should not be over-claimed, benchmarks often provide the most objective performance measurement and are therefore important research guidelines. A new Multiple Object Tracking benchmark, MOTChallenge, has recently been launched to collect existing and new data and create a framework for standardized evaluation of multiple object tracking methods.

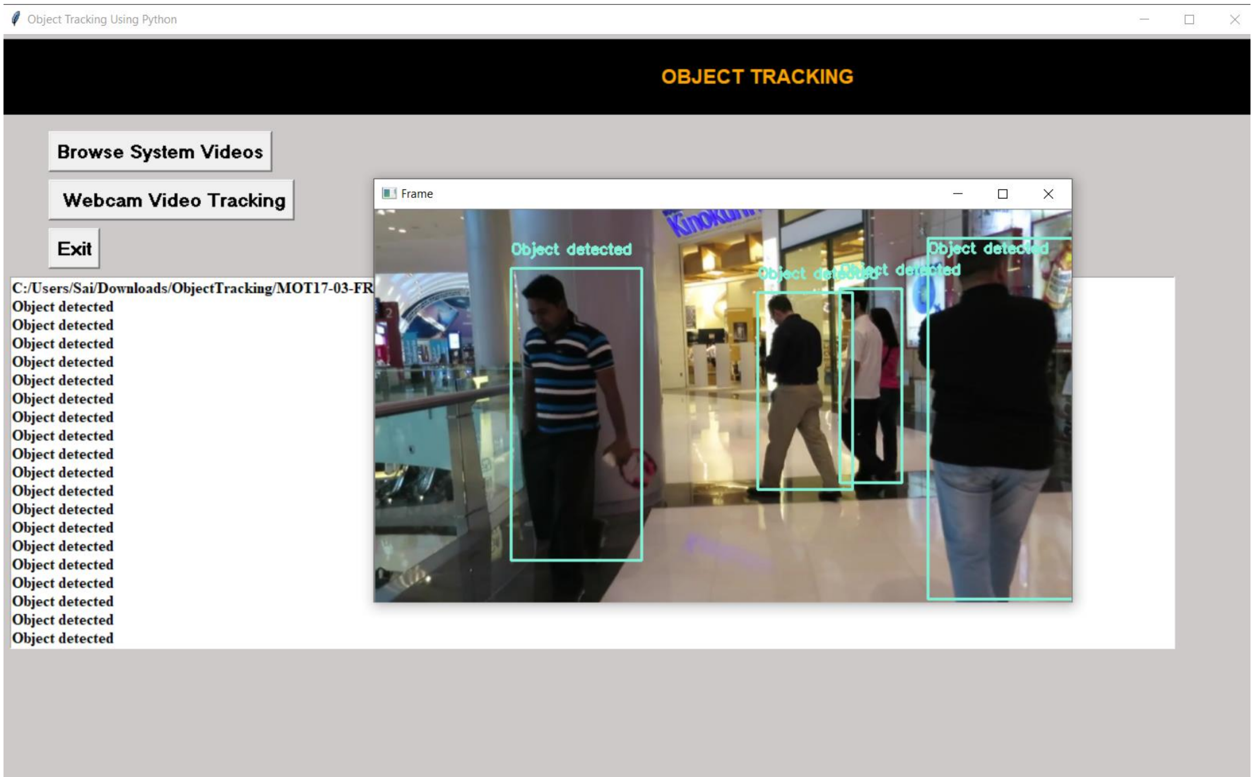
Output of the Benchmark Dataset Videos:

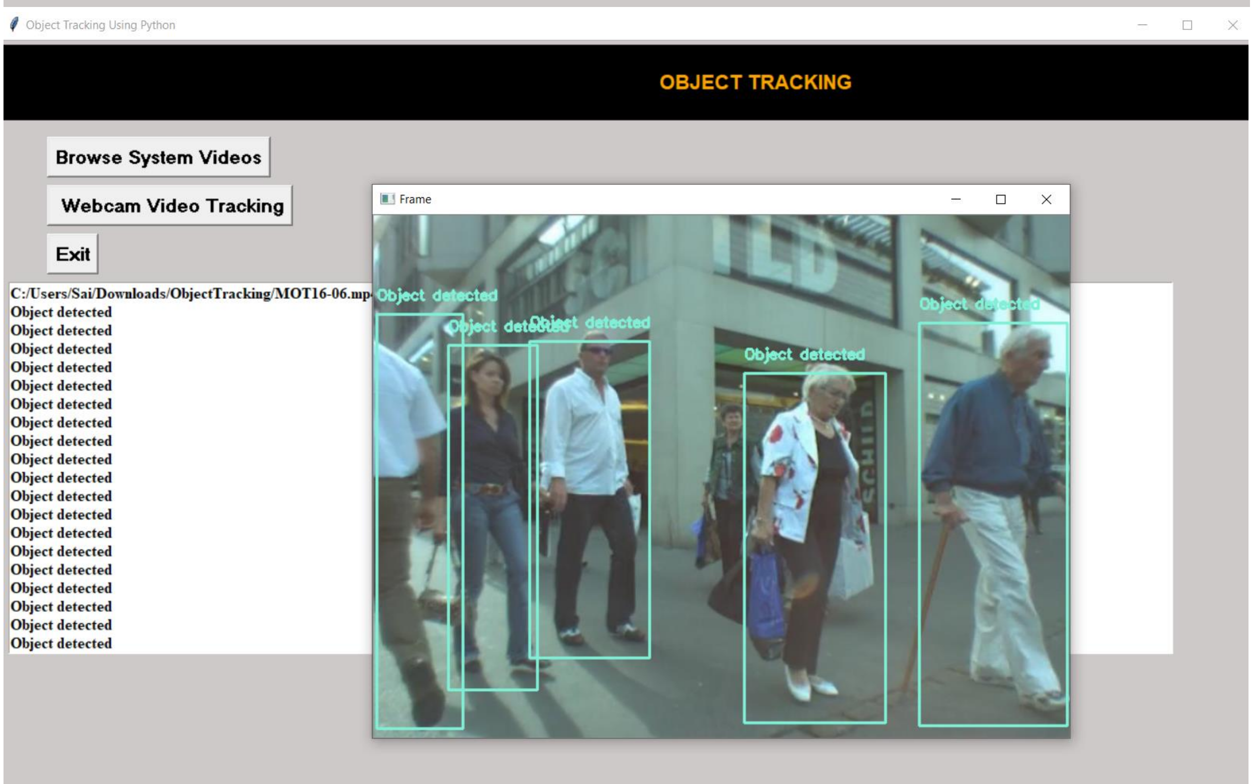


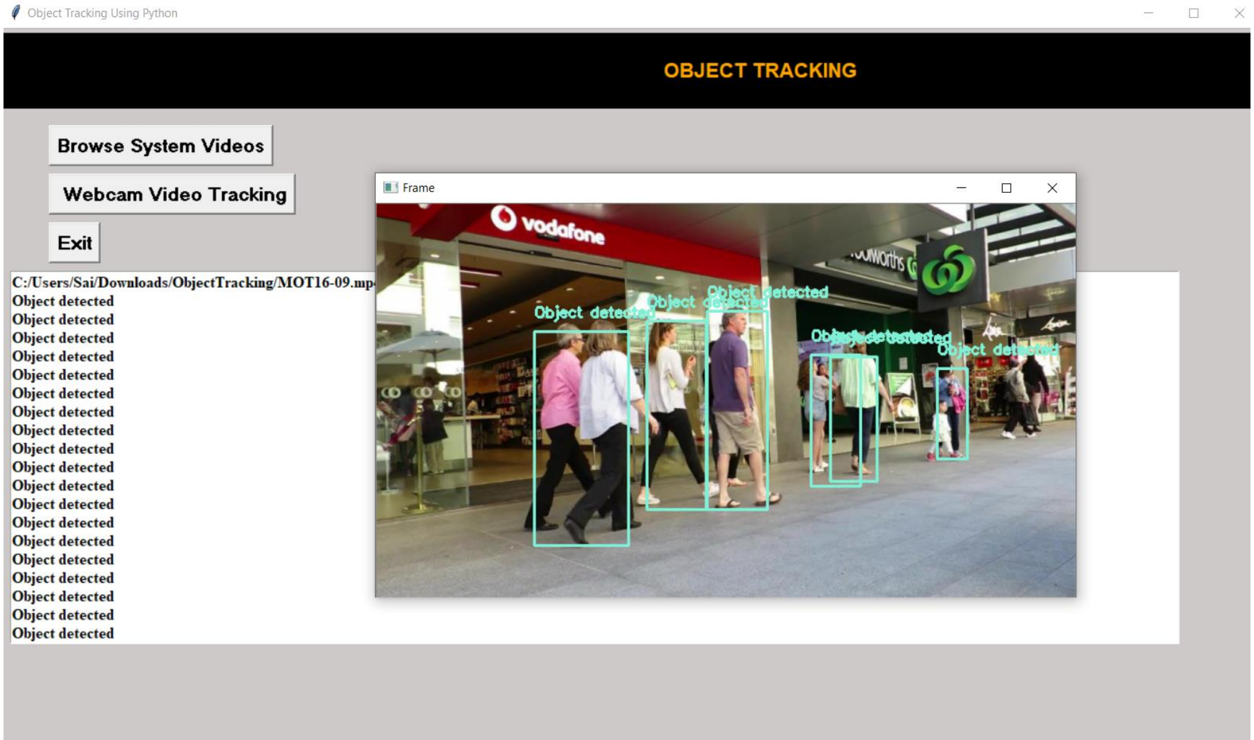


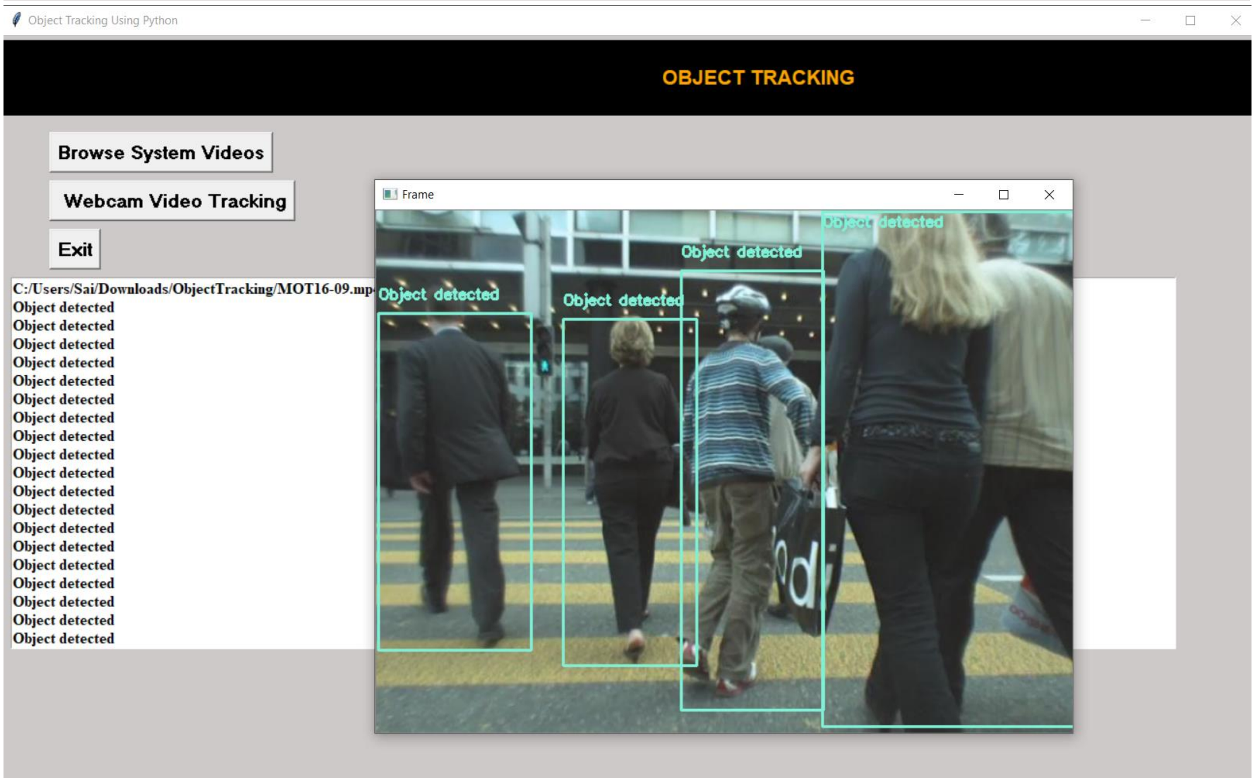

Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
Frame
C:/Users/Sai/Downloads/ObjectTracking/MOT17-13-DP
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
NIKE

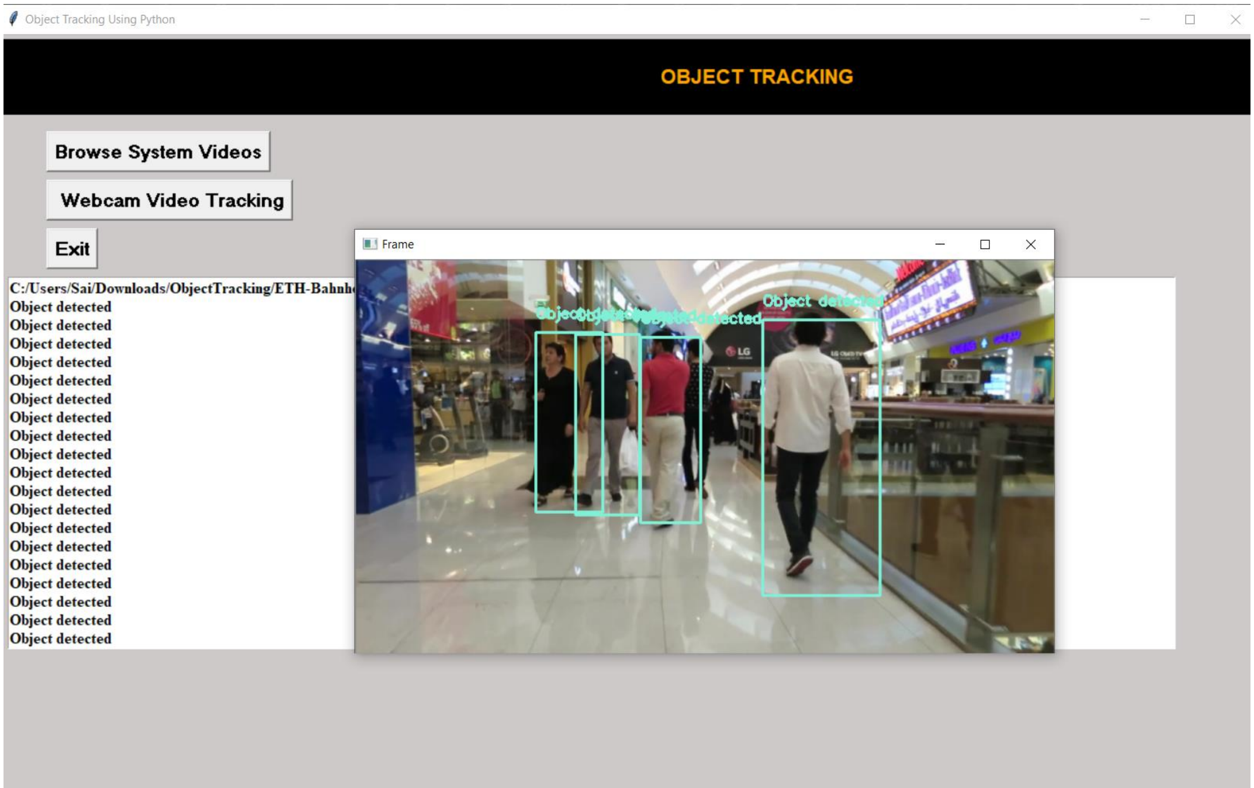

Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
Frame
C:/Users/Sai/Downloads/ObjectTracking/MOT16-08.mp
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
mërkur
Object detected
Object detected
Object detected

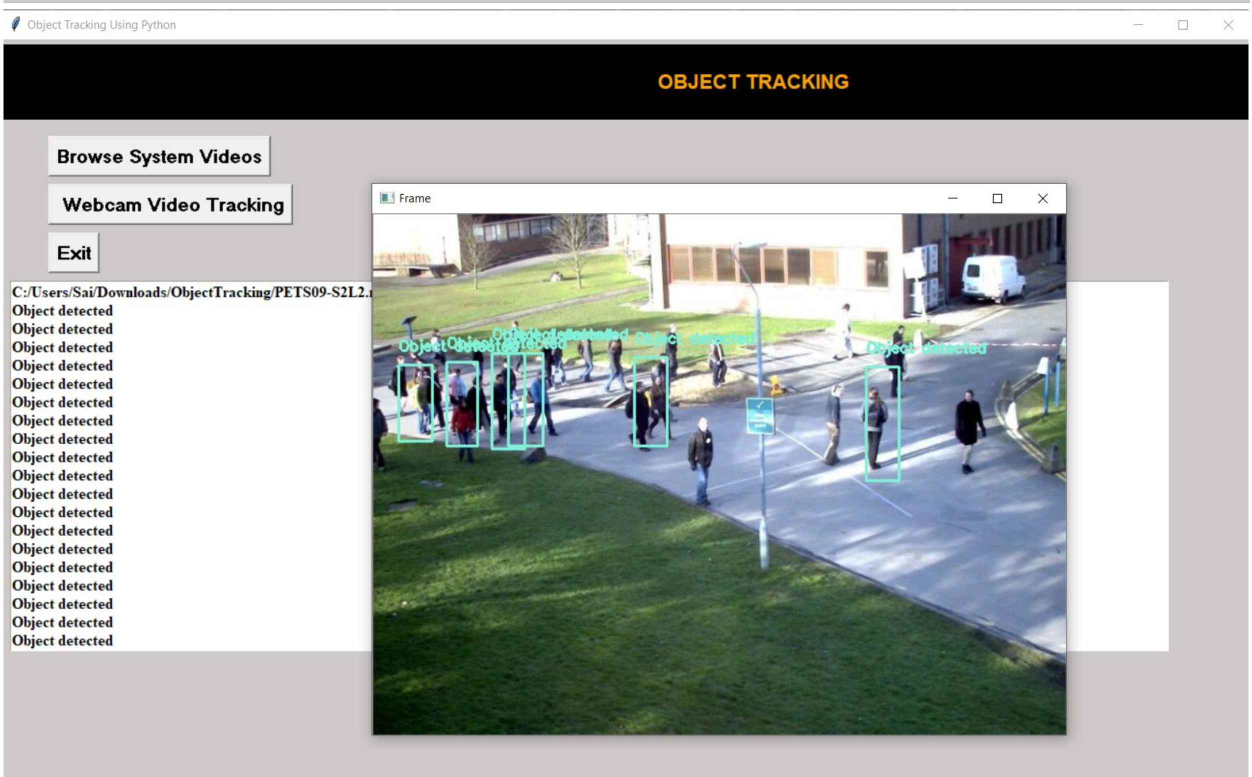

Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
C:/Users/Sai/Downloads/ObjectTracking/MOT17-03-FR
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Frame
Object detected

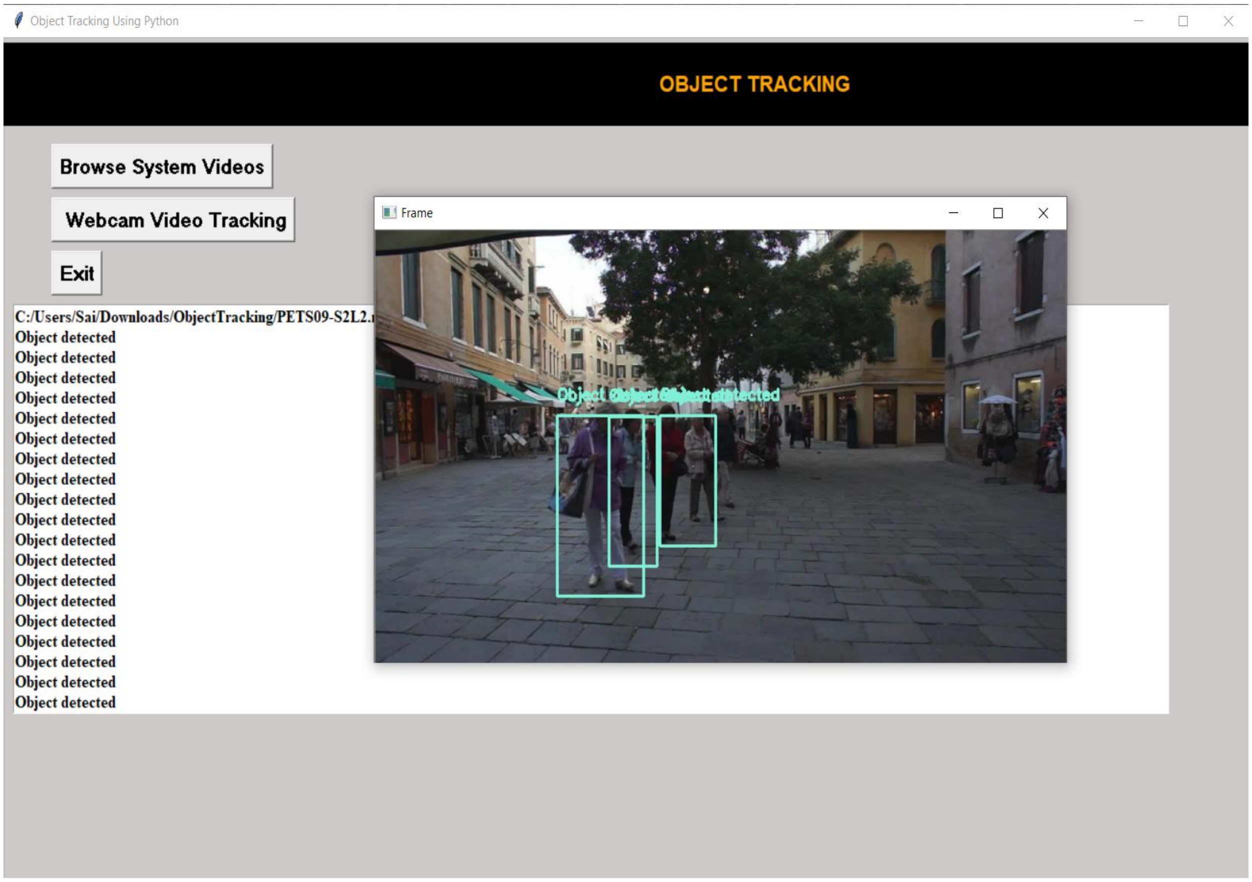

Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
C:/Users/Sai/Downloads/ObjectTracking/MOT16-06.mp
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Frame
Object detected
Object detected
Object detected


Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
C:/Users/Sai/Downloads/ObjectTracking/MOT16-09.mp
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Frame
vodafone
Object detected


Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
C:/Users/Sai/Downloads/ObjectTracking/MOT16-09.mp
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Frame
Object detected
Object detected
Object detected
Object detected


Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
C:/Users/Sai/Downloads/ObjectTracking/ETH-Bahnh
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Frame
Object detected


Object Tracking Using Python
OBJECT TRACKING
Browse System Videos
Webcam Video Tracking
Exit
C:/Users/Sai/Downloads/ObjectTracking/PETS09-S2L2.
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Object detected
Frame
Object detected



10. Why do we use Kalman Filter?

There are several benefits to the Kalman filter algorithm. This is a statistical technique that properly describes the experimental measurement random structure. This filter can consider quantities that are neglected in other techniques partially or completely.

When the results are compared to other algorithms the execution time taken by the Kalman filter was very low and compared algorithm (Yolo) required higher processing GPU to generate results along with very slow training rate.

The variation of the original state estimation and the design deviation variance. This provides information on the calculation value by including the uncertainty of the estimation error in comparison to the best estimate. The Kalman filter is ideal for electronic internet storage.

Its recursive structure enables it to be executed in real time without storing past observations or estimates.

- It will measure the countries affected by the external disruption correctly but cannot be used to forecast the condition of a non-linear environment.

- Kalman filter is a combination of estimating parameters with pattern recognition.
- Computational process increases along with the increase in total number of detectors.
- The parameters estimated follow along the path observing very nearer to the object.
- It doesn’t require information of derivatives to perform the operations.
- Can be easily computed without higher GPU and scaling is done with low processing.
- It can handle sensor silence for a small amount of time.
- Manages occlusion, when an object occludes the measurement of the occluded object is skipped and predicts the next detected object.
- Saves time and no need to calculate matrix values.

11. Conclusion :-

For real-time applications, we may need a fast high-performance algorithm; hence, we opted for mean and cam shift algorithms.

Achieved excellent results for real-tim offline object tracking using these algorithms through locally file selection. During multiple object tracking, the Kalman filter algorithm is successful and has many advantages over the state-of-the-art. The rollback loop added will eliminate the majority of over and under segmentation, resulting in increased accuracy in both object detection and tracking techniques.

In case of online recording of the video, the algorithm performed well in most cases and has detected the objects near it and tracking the objects while moving and if in any situation the object movement is halted then the detection will pause for that object and will resume when that object starts moving again.

Overall, the performance of the algorithm seemed more enthusiastic for the offline upload method due to already gathered information in the video, the conditions for tracking were well suited for tracking. Where in the other hand, the tracking of objects using the webcam resulted in a little less performance due to movement of objects and uneven distribution of light was also a major factor for the results to be positive i.e tracking the objects without occlusion.

This paper justifies a summary of the most common object tracking methods, which hopefully can give valuable and accurate results.

If you've ever played face detection with OpenCV, you know it works in real time and in every frame, you can quickly identify the eye. So, why do you first need to track? Let's discuss the various reasons for detecting objects in a clip and not just repetitive detections.

Tracking algorithms are usually faster than algorithms for detection. The explanation is clear. You know a lot about the object's presence when you watch an object that was observed in the previous frame. You also know the location and direction and speed of its movement in the previous frame. A successful tracking algorithm, while a detector algorithm still starts from scratch, will use all the information it has about the target up to that point.

Therefore, while designing an efficient system, an object detection is usually performed on each nth frame while the tracking algorithm is used between the n-1 frames. Sure, tracking benefits from the extra information it has, but you can also lose track of an object if it goes behind an obstacle for an extended period of time or if it runs so quickly that the tracking algorithm can't catch up. Tracking algorithms is also common to generate errors and the bounding box that tracks the target slowly drifts away from the object it tracks.

When you run a face detector on a video and an item occludes the face of the person, most likely the face detector will malfunction. On the other hand, a successful monitoring algorithm can manage any occlusion level. This makes a upside for tracking incase of failure in detection.

For Better Results

We have been facing some issues during low light and noise occurrence. To overcome this, we have been trying real time object tracking with different algorithms like centroid tracking algorithm, but comparatively, we achieved better results with mean, cam shift and Kalman filter.

Through improving methods for a more automated collection of devices, the accuracy of object tracking may potentially increase. From practise, we know that a human tendency makes more errors than a computer programme designed for a particular purpose.

Automatic feature selection based on pattern recognition where methods are divided into philtre methods and wrapper methods for this purpose. However, in the field of object tracking, where the selection of features is still mostly done manually, these have not received the same attention. By developing fast and accurate methods for automatic collection of features, there could be room for improvement in object tracking.

12. Future Work :-

Many measures are mostly done manually in the process of tracking objects, one example is the selection of apps. Through improving methods for a more automated selection process of items, the precision of object tracking may potentially increase. From experience, we know that a human tendency makes more mistakes than a computer programme optimised for a specific purpose.

Automatic interface selection has received attention in the field of pattern recognition, where methods are categorised into philtre methods and wrapper methods for this purpose. But, in the area of object tracking, where collection of features is still mostly done manually, these have not received the same attention. By developing fast and accurate methods, there could be room for improvement in object tracking.

13. References :-

1) A 3D State Space Formulation of a Navigation Kalman Filter for Autonomous Vehicles Alonzo Kelly CMU-RI-TR-94-19

2) R. G. Brown and P. Y. C. Hwang, “Introduction to Random Signals and Applied Kalman Filtering”, Wiley, 1992.

3) An Introduction to the Kalman Filter Greg Welch1 and Gary Bishop2 TR 95-041 Department of Computer Science University of North Carolina at Chapel Hill Chapel Hill, NC 27599-3175.

4) Model-Based 3D Object Tracking Using an Extended-Extended Kalman Filter and Graphics Rendered Measurements Hua Yang, Greg Welch.

5) 1960 Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. Journal of Basic Engineering, 1960(82), pp. 35-45.

6)IMPROVED OBJECT TRACKING WITH CAMSHIFT ALGORITHM OULD-DRIS Nouar*, GANOUN Ali*,* and CANALS Raphaël.

7) 2011 International Conference on Physics Science and Technology (ICPST 2011) Head Pose Estimation Based on Head Tracking and the Kalman Filter Wang Yua, b, Liu Gangc .

8) OPTIMIZED OBJECT TRACKING TECHNIQUE USING KALMAN FILTER Liana Ellen Taylor a,*, Midriem Mirdanies b, Roni Permana Saputra.

9) J Joshan Athanesious and P Suresh. Systematic survey on object tracking methods in video. International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) October, pages 242–247, 2012.

10) Real-time detection and tracking for moving objects based on computer vision method - Yi-You Hou ; Sz-Yu Chiou-2017.

11)Modeling from an Object and Multi-object Tracking System -Afef Salhi , Yacine Moresly , Fahmi Ghozzi -2016.

12) J. Meja Patel and Bhumika Bhatt. A comparative study of object tracking techniques. International Journal of Innovative Research in Science, Engineering and Technology, 4(3), 2015.

13) Grandham. Sindhuja and Renuka. Devi. A survey on detection and tracking of objects in video sequence. International Journal of Engineering Research and General Science, 3(2), 2015.

14) Performance comparison of extended Kalman filter and particle filter in satellite relative measurement-Zhang Qi ; Liu Rongke ; Mu Weiqing ; Wang Zijie-2017.

15) 'https://arxiv.org/ftp/arxiv/papers/1802/1802.01235.pdf'-Unscented Kalman Filter

16) Multi-object tracking using Kalman filter and particle filter-Chetan M. Bukey ; Shailesh V. Kulkarni ; Rohini A. Chavan-2017.