

# A WSM-based Comparative Study of Vision Tracking Methodologies

Sara Bouraya, Abdessamad Belangour  
Laboratory of Information Technology and Modeling  
Hassan II University, Faculty of Sciences Ben M'sik  
Casablanca, Morocco

**Abstract**—Vision tracking is a key component of a video sequence. It is the process of locating single or multiple moving objects over time using one or many cameras. The latter's function consists of detecting, categorizing, and tracking. The development of the trustworthy solution for video sequence analysis opens up new horizons for a variety of applications, including intelligent transportation systems, biomedical, agriculture, human-machine interaction, augmented reality, video surveillance, robots, and many crucial research areas. To make efficient models, there are challenges in video observation to deal with, such as problems with the environment, light variation, pose variation, motion blur, clutter, occlusion, and so on. In this paper, we present several techniques that addressed the issues of detecting and tracking multiple targets on video sequences. The proposed comparative study relied on different methodologies. This paper's purpose is to list various approaches, classify them, and compare them, using the Weighted Scoring Model (WSM) comparison method. This includes studying these algorithms, selecting relevant comparison criteria, assigning weights for each criterion, and lastly computing scores. The obtained results of this study will reveal the strong and weak points of each algorithm mentioned and discussed.

**Keywords**—Multiple object tracking; object detection; WSM method; computer vision; video analysis

## I. INTRODUCTION

Target/Object tracking is the challenge of determining the location, path, and attributes of objects of interest using sensor measurements [1]. A sensor could be any measuring equipment that collects information about targets in the environment, such as sonar, radar, infrared sensor, lidar, camera, ultrasound, microphone, or any other sensor. Object tracking is common goals include determining the number of targets, their states, and their identities, such as velocities, positions, and in certain situations their features. The radar monitoring of aircraft is a typical example of target/object tracking.

Based on measurements collected from radar, the object tracking issue, in this case, aims to determine the number of aircraft in a surveillance area, their types, such as commercial, military, or recreational, their speed, and their identities and positions. The object tracking problem contains a variety of sources of uncertainty, making it a difficult undertaking. Object motion, for example, is frequently subject to random disruptions; sensors could misdetect objects, and the number of objects/targets in a sensor's field of view could change at random. Sensor measurements are prone to random disturbances, and the number of measurements acquired by a

sensor could vary and be unpredictable from one look to the next. Objects may be closer, and the measurements acquired may not be able to distinguish among them properly. Sensors could offer data even when there is no object in the area of view.

Humans could track objects visually with relative ease. However, it is not obvious, and it is difficult for a computer to track a moving object under illumination variation, with different shapes and format, object-to-scene and object-to-object occlusions, as well as background, clutters, and with several appearances in-camera projective space, object appearance, and object disappearance. Multiple object tracking is a set of tasks to reach tracking starting from object detection, object classification then objects tracking. Tracking objects is following their trajectory with video and their different positions within frames. Lastly, Multi-object tracking or multi-target tracking is gaining widespread interest across different research areas ranging from autonomous cars and vehicles, video surveillance, human-machine interaction, virtual environment, biomedical analysis, and so on. These objects may be pedestrians, cars, fish, vehicles, motorcycles. Although object tracking has a wide range of areas, it is still suffering from problems that we are going to discuss carefully in the next coming sections. Most of the researchers are trying to tackle these challenges carefully; however, some of them fail to solve some of the obvious problems. That is why we are offering you this comparative study that aims to identify, analyze, and compare the different object tracking methods.

As a result, this work provides a weighted scoring model (WSM)-based comparison evaluation of different methodologies. Our comparative study begins with the extraction of crucial criteria for comparison and a description of each criterion (Section 5). These criteria have been gathered from relevant work comparative research (Section 3) in addition to some other criteria that are missed. The WSM approach is then used to calculate final scores for each algorithm, which requires weight attribution (Section 5). The results are depicted as a chart and carefully discussed (Section 6).

## II. RELATED WORK

Many scientific studies have been conducted to develop multi-object tracking systems and applications based on a wide range of algorithms of the deep learning era and other eras. Many researchers and scientists have put up scientific efforts in this direction to apply several algorithms. The authors provide

solutions in a variety of areas, including Smart City, Smart Vehicle, Industry, logistics, medical, surveillance, etc.

Several studies have compared multi-object tracking methods relied on different methods and directions. Each comparative study has its uniqueness and crucialness.

In [2], the authors picked up just three algorithms (Extended Kalman Filter [3], Gaussian mixture model GMM [4], and Mean Shift Algorithm [5]) and compared them. On the one hand, resulting in that GMM performs better when there is occlusion, contrary to Extended Kalman filter. On the other hand, the mean shift algorithm is best suited to single target tracking, at meanwhile, the results reveal that this approach is unable to detect many objects when even minor occlusion exists.

In [6], the authors are comparing multi-object tracking methods for sports events, choosing four algorithms; namely, Medianflow [7], boosting[8], multi-instance learning(MIL) [9], Kernelized correlation filter (KCF) [10], measuring as well as comparing their performances. Resulting from that the tracking of the object is efficient if the object's movement is constant when utilizing the MEDIANFLOW[7] algorithm. However, if the object's movement suddenly changes, it will be unable to follow the trajectory. For tracking many items in a sporting event, the KCF [11]algorithm, performed best, boosting had a greater tracking success rate than MIL, and it tracked roughly 5 times faster.

In [12], the authors are including an examination object detection method, object representation, and feature selection, and object tracking over many frames. By comparing seven different techniques (Kalman Filter[13], Particle Filter[14], Mean shift [7], CamShift[15], KLT tracker[16], template matching[17], Contour Tracking [18]) relied on number of object tracked, occlusion handling, optimal or not. Resulting from that Kalman Filter, KLT tracker, Particle Filter, Contour Tracking is handling occlusion and they are optimal strongly the last two ones could strongly track multiple objects.

In [19], the authors are presenting a paper that provides a brief overview of the numerous object detections, categorization, and tracking algorithms in the literature, as well as analysis and comparison of the many strategies utilized at various phases of tracking; reaching a comparison of the different techniques Point-Based Tracking, Kernel-Based Tracking, Silhouette Based Tracking. Concluding that single object tracking provides good accuracy for many types of movies with varying conditions, such as low resolution or weather changes.

In [20], the authors wrote a comparative examination of multiple vision tracker categories is conducted by the work to determine which one is the most successful in tracking construction resources. The benefits and limits of each kind of tracker are discussed, as well as the testing procedures for evaluating them. The methods are divided into the same categories used in the previous paper. The kernel-based ones are insensitive to illumination conditions and scale variation than point tracking-based methods as well as point tracking are effective under occlusion.

In [21], on the one hand, the authors examine various object tracking techniques in this research. The classification of these techniques is the same classification as in the previous papers kernel, silhouette, and point tracking-based. Their comparison is based on a wide range of criteria, concluding that point tracking methods are dealing better with occlusion; meanwhile, they are optimal. On the other hand, they present the advantage and the limitation of the different existing methods.

In [22] authors discuss Background subtraction, template matching, Frame difference, and shape-based approaches that are some of the most frequent approaches used in object tracking. In addition, they talk about topics like detection and tracking. At the end of this review on object detection and tracking methods, they summarize a table containing the advantages and disadvantages of usual object tracking methods. Concluding that the best method of object vision tracking vision is by combining different methods at the same time.

In [23], this comparison is evaluating 2D to 3D vision tracking method's performance, reaching a comparison also for the usual sub-categories kernel object tracking, point object tracking, silhouette object tracking. Concluding that the first findings of this comparison indicate that of the two types of trackers examined for construction-related applications, kernel-based approaches are more trustworthy.

In [24], the authors present a review of the existent approaches of moving object detection, their Challenges, and mentioning object tracking at the end. The object tracking methods that are mentioned in this article are Mean-shift, KLT, Condensation [25], TLD, Tracking Based on the Boundary of the Object. Mentioning that, TLD [26] is an award-winning, real-time method for tracking unknown objects in a video sequence.

### III. BACKGROUND

#### A. Vision Tracking

The difficulty of determining the trajectory of an object in the image plane as it moves across a scene is known as tracking. Various ways of tracking the object include kernel tracking, point tracking, and silhouette tracking.

Various ways of tracking the object include point tracking, kernel tracking, and silhouette tracking. The following categories could be used to categorize tracked tracking methods.

Because of the enormous variety of applications and the high-performance of vision tracking in many fields, object tracking has sparked a lot of interest and attention in the field of computer vision. Two key stages must be completed in order to achieve vision tracking. The initial stage is to detect targets of interest, which could be done either automatically or manually depending on the approach used. The second stage, also known as target localization, involves following the discovered objects and predicting their changes in terms of position and form in subsequent image frames.

### B. Target/Object Tracking Methods

As mentioned in Fig. 1 there are three main subclasses of object tracking [1]:

- Point tracking.
- Kernel tracking.
- Silhouette tracking.

Table I is summarizing an overview of object tracking methodologies their main goal, their advantages, their disadvantages, and some of their methods. Additionally, Table II illustrates the advantages and limitations of object tracking methodologies.

1) *Silhouette based tracking approach:* Because of their irregularity, many sorts of objects could not be successfully represented using simple geometric shapes due to their complexity. Silhouette-based approaches give a precise form description for different objects'/targets' shapes. Heads, hands, and shoulders are examples of composite shapes that are difficult to characterize with geometric geometry or shapes, that is what this category is made for; in other words, to shape non-geometric shapes. A suitable geometric shape for those objects/ targets will be given by this methodology. The main goal of this mechanism of tracking is to detect the target region in each frame using the assistance of a target model outputted from previous frames. It is possible to deal with a variety of object shapes, object split and merge and occlusions. The model is represented by object edges, a color hologram, or a contour. Shape matching and contour tracking are the two types of silhouette tracking that we use. These techniques are capable of handling a variety of problems such

as it could be used to manage a wide range of objects of various shapes means dealing with a wide range of target's shape, and it also handles occlusion, silhouette tracker may split and merge objects as well. This silhouette-based tracking could be divided into two main categories, i.e. contour-based tracking approaches and shape matching tracking approaches.

a) *Contour Tracking:* Iteratively evolving an initial contour based on object placement in previous frames to a new place in the current frame is how this strategy works. Object portions in the current frame must overlap the object in the previous frame for the contour to evolve. When tracking is based on contour evolution, two ways could be used. State-space models, which are utilized for motion and contour shape modeling, are used in the first method. The other technique generates the contour directly by using direct minimization methods to reduce contour energy. Consider gradient descent, one of the most appealing aspects of this approach is its ability to handle a wide range of object shapes. The silhouette could be represented intuitively by a function defined on a grid, and it could also be expressed clearly by a set of control points to indicate its border.

b) *Shape Matching:* In the kernel technique, shape matching performs similarly to template-based tracking. Another method for Shape matching is to look for silhouettes that are similar in two consecutive frames. The process of silhouette matching is similar to that of point matching. Background subtraction is used to conduct Silhouette detection. Density functions, silhouette boundaries, and object edges make up the Models object. Hough transform techniques will be used to handle occlusion and dealing with a single object.

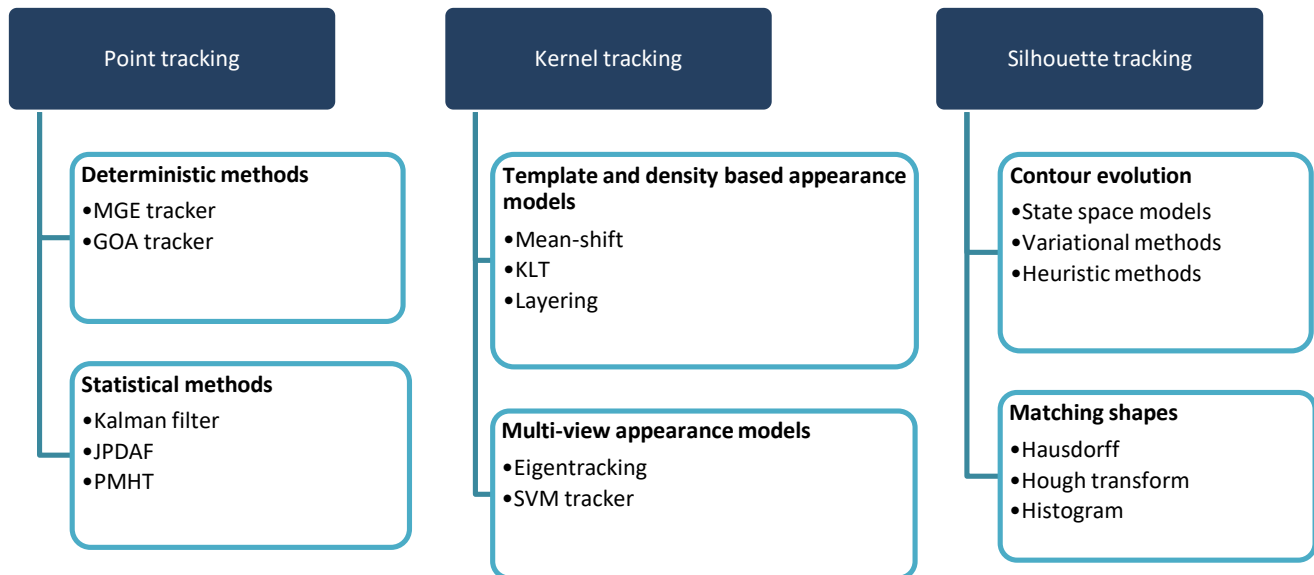


Fig. 1. Object Tracking Main Methodologies [1].

TABLE I. OBJECT TRACKING METHODS

	Kernel tracking		Silhouette tracking		Point Tracking	
<b>Main goal</b>	Estimating object motion		Tracking exact contour		tracking small objects	
<b>Advantages</b>	handling occlusion		Ability to represent a wide range of shapes, handle a large variety of object shapes		tracking small objects	
<b>Disadvantages</b>	handling multiple objects		handling occlusion and object merge and split		Handling occlusion at the meanwhile appearance and re-appearance	
<b>Categories</b>	Template based models	Multi-view models	Contour evolution	Matching Shapes	Deterministic	Statistical
<b>Some of the methods</b>	Mean-shift	SVM Tracker	Variational methods	Histogram	MGE Tracker	MHT filter

TABLE II. ADVANTAGES AND LIMITATIONS OF OBJECT TRACKING METHODS

Tracking Methods	Advantages	Limitations
<b>Kalman Filter</b>	<ul style="list-style-type: none"> <li>. Capable of tracking pictures that are noisy. We could not rely on the average of past values if our sensors deliver incorrect data or if they do not produce any data at all for a certain time. Outliers are dealt with the use of Kalman filter.</li> <li>. Work for linear models and Gaussian distribution.</li> </ul>	<ul style="list-style-type: none"> <li>. Kalman filter is unusual for non-gaussian problems.</li> </ul>
<b>Particle Filter</b>	<ul style="list-style-type: none"> <li>. An efficient algorithm used in the world of non-gaussian and non-linearity. In addition to the capability of multi-modal filter that why it is widely used.</li> <li>. Particle filter methods are extremely adaptable, simple to develop, parallelizable, and suitable in a wide range of applications.</li> </ul>	<ul style="list-style-type: none"> <li>. Particle filters do not have a strict proof of convergence</li> </ul>
<b>Mean shift</b>	<ul style="list-style-type: none"> <li>. When there are a lot of colors, this is a good option.</li> </ul>	<ul style="list-style-type: none"> <li>. Could not be applied in the case of complex scenes.</li> </ul>
<b>KLT tracker</b>	<ul style="list-style-type: none"> <li>. Handle Occlusion</li> </ul>	<ul style="list-style-type: none"> <li>. Multiple object tracking this method Becomes complex.</li> </ul>
<b>Template matching</b>	<ul style="list-style-type: none"> <li>. It is a lot easier to set up and utilize.</li> </ul>	<ul style="list-style-type: none"> <li>. A slow method for recognizing new variations of a pattern.</li> <li>. It only works if the object is always visible in the video; else, false detects will occur.</li> <li>. When there are a lot of objects/targets, it is difficult to keep track of them all.</li> <li>. Complex templates are not recommended.</li> <li>. When items leave the frame or become occluded, problems could arise.</li> </ul>
<b>Contour Tracking</b>	<ul style="list-style-type: none"> <li>. Complex models for non-rigid and rigid objects could be highly handled.</li> <li>. Illumination levels do not have a high impact</li> </ul>	<ul style="list-style-type: none"> <li>. It is difficult to deal with entry objects.</li> </ul>

2) *Kernel-based trackers*: Kernel-based tracker utilizes representations and appearance of the object/target of interest using ellipsoidal or rectangular shapes. It is possible to track a target or an object by tracking the motion of each kernel on the associated frame. The motion of a target or an object could be classified in different ways including rotation, translation, or transformation. The motion of an object could be categorized in different ways including affine transformations, rotation, and translation. Several algorithms could be utilized for this goal, which differs relying on the quantity of target tracking, object motion method, or object representation. For instance, in real-time applications, geometric shapes are frequently used to represent objects. One disadvantage of employing geometric shapes is that they may not completely encapsulate the target object, allowing background objects to be visible. The mean shift method, template matching, and CAMShift tracking are some available approaches that could be used for kernel tracking, relied on Kernel tracking methodology.

a) *Simple Template Matching*: Template matching is a computer vision program that finds related objects in photos. For mapping similar patterns between images, template matching is an important part of the image analysis process. Image processing is a computer-assisted technique for analyzing and manipulating images. It is mostly used to enhance the quality of an image, as well as to detect and highlight specific areas of the image. Image processing techniques are primarily used to improve the quality of an image and perform feature extraction and classification for a variety of purposes. In the automation process, template matching is used to recognize objects and improve the quality of the searching process. It is used to locate the target item from an image in astronomy, meteorology, medical imaging, remote sensing, and many other related domains. It is an image processing approach in which a specific object is chosen as a target, to be recognized and mapped to the original image. For images matching, a variety of techniques are utilized, with Edge-based matching and Greyscale-based matching being the most popular. Greyscale-based is an extension of correlation-based that works regardless of the

orientation of the image. The greyscale-based technique recognizes both the template location and orientation, allowing us to recognize images from a variety of angles. Edge-based matching is similar to greyscale matching, except that instead of computing the entire image, only the edges of the image are recognized and matched with surrounding pixels.

*b) Mean Shift Method:* This tracking method utilizes density-based appearance models to describe a target. A histogram is used to represent the appearance model (color, texture) within this algorithm. This method employs an iterative tracking approach, which entails identifying similar pattern distributions within a sequence of frames.

*c) Support Vector Machine (SVM):* SVM is a general classification approach that uses a set of positive and negative sample values to define it. Positive samples containing tracked visual objects are used in SVM.

The negative samples are made up of everything else that isn't being tracked. It could handle a single picture as well as partial occlusion of an object, although physical initialization and training are required. SVM finds the best separating hyperplane between two classes. In SVM as a tracker, the positive examples for SVM-based trackers are photos of the object to be tracked, while the negative examples are those objects that are not to be tracked. Negative examples are typically made up of background regions that could be mistaken for the object.

*d) Layering based tracking:* Multiple objects could be tracked using this kernel-based tracking approach. Each layer has an elliptical form, motion (such as translation and rotation), and layer appearance (based on intensity). Layering is accomplished, first by accounting for background motion such that the object's motion may be calculated using 2D parametric motion from the rewarded image.

*3) Point-based trackers:* It is a common computer vision task with a wide range of applications. During tracking, moving objects are carefully represented as feature points in an image structure. Point tracking is a difficult problem, especially when occlusions occur. Point Tracking could handle the following situations, suitable for tracking extremely small objects. Objects identified in successive frames are represented by points, which are linked together based on the prior object state, which could include object position and motion. This method necessitates the use of an external system to detect the objects in each frame.

These are the main point-based trackers' methods:

*a) Kalman Filter:* This is one of the most used techniques of point-based trackers [27]. In a real-world scenario where we are tracking a moving object from our car or vehicle, we could not rely on the average of prior values if our sensors transmit incorrect data or if they do not send any data at all. Outliers are dealt with with the use of the Kalman filter. It only evaluates one sensor data at a time and compares it to prior values, giving the previously estimated measurement more weight if it has a low error and giving the

newly taken value from the sensor more weight if it has a low error. Because it only analyzes a proportion of newly taken value at each time step, it avoids the problem of outliers. Well, the Kalman filter is used to anticipate most measurements where we acquire data from hardware, such as sensors, and we do not know how reliable the data is. The Kalman filter algorithm could be broken down into the following steps:

*1) Initialize state and covariance matrices:* When we receive the first sensor measurements, we initialize the state (position and velocity) of the moving target, such as a bicycle.

*2) Forecast step:* Based on certain prior data and the model, we produce a state prediction.

*3) Step of updating:* The predicted and measured locations are merged to produce an updated location. Depending on the uncertainty of each number, the Kalman filter will give greater weight to the projected or observed location.

*4) The process is repeated as the automobile receives new sensor readings.* They have relied on the algorithm of Optimal Recursive Data Processing.

*b) Particle Filter:* The versatility of particle filter technology is due to its efficiency in nonlinear and non-Gaussian systems. Furthermore, the particle filter's multimodal (we want to track, simultaneously, zero, one, or more than one object) processing capabilities is one of the reasons for its widespread use. Particle filtering has been used in a variety of fields around the world. It is based on the Markov Chain Monte Carlo improvement strategy, by generating Markov chains with good convergence; the approach generates samples from the target distribution. Particle filter methods are extremely adaptable, simple to develop, parallelizable, and suitable in a wide range of applications. The particle filter method can handle different problems such as occlusions.

*c) Multiple Hypothesis Tracking (MHT):* Several frames have been detected in the MHT algorithm for better tracking results. The MHT algorithm is an iterative process. Each Iteration starts with a set of track hypotheses that already exist. Each theory is made up of a group of disconnected tracks. A prediction of the object's position in the next frame is made for each hypothesis. Then, the predictions are compared, using a distance metric. The MHT could track several objects, deal with occlusions, and calculate optimal solutions. The algorithm can create new tracks for objects entering the field of view FOV and terminate tracks for objects exiting the field of view FOV.

### C. Research Areas

Due to the existence of several objects in our lives, for instance, humans being (pedestrians [28], sport players [6], shoppers), vehicles (cars [29], motorcycles, buses), animals (fishes [30], birds [31], cats), other (cells [32], insects), and so on; many scientific studies have explored implementing multiple object tracking using different algorithms. These initiatives are divided into research areas or domains; see Fig. 2, such as:

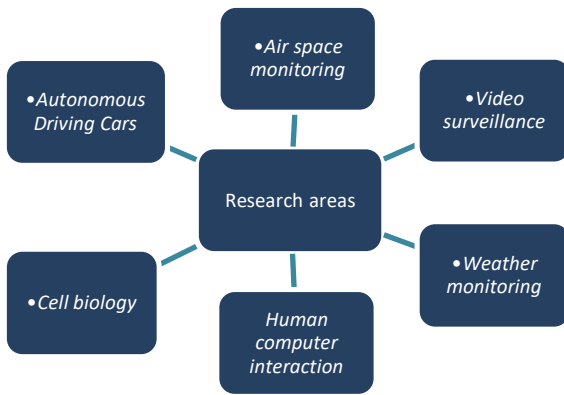


Fig. 2. Some Research Areas of Object Tracking.

- **Air space monitoring:** The tracking of aircraft using radar is a serious and crucial issue [33], for instance, air traffic control. Radar tracking is employed also in military surveillance systems to identify aircraft: identity, type, location, speed, and the item's potential intents to establish, for instance, if the object is a threat. Radar is capable of a wide range of observations, from high range measurements to simple-resolution imaging. Radar [34] utilizes radio waves to determine the object's distance, direction, and radial speed.
- **Video surveillance:** Surveillance [29] using digital video is becoming increasingly popular. Airports, banks, casinos, highways, stadiums, crowd gathering areas, buildings, streets, railway stations, department stores, and all government organizations now use video surveillance to reinforce their security. Video surveillance is employed in almost every aspect of society to deter criminal activity and improve public safety and security [35]. A large typical building in a major city has a vast network of cameras installed on main floors, entrances, huge gathering areas, hallways, offices, and labs.
- **Weather monitoring:** Weather bureaus employ a variety of approaches to give weather forecasts [36]. Tracking weather balloons, which offer data on high-altitude wind velocity, humidity, pressure, and temperature, is a widely used method. Each day, weather bureaus released 50 to 70 balloons at various times throughout the day. The frequency of releases rises during extreme weather because more data is required for good weather forecasting. Each weather balloon must be tracked to obtain weather-related parameters at various heights of the atmosphere.
- **Cell biology:** Pathologists and medical researchers commonly examine death and birth rates, as well as the mobility of biological cells [37], in research studies of humans, plants, insects, and animals. In immunology, the immune of organism response is linked and correlated with lymphocytes' life cycle. The division/death and birth rates of every generation of cells are the parameters of interest. The velocity or morphology of sperm cells is an interesting metric of infertility study.

- **Autonomous Driving Cars:** The self-driving [38] systems are also known as autonomous cars or driverless vehicles; a car is a vehicle that could sense its surroundings and move in designated lanes without the need for human intervention. In other words, object detection and tracking are required for autonomous vehicles to reliably recognize and localize dynamic targets in the environment surrounding vehicles that contain the tracking system. The principles of developing self-driving automobiles are completely based on perceiving their surroundings and automating tasks.
- **Human-computer interaction:** Gestures [35] have long been thought of as an inter-action approach that may help us communicate with our computers in more creative, natural, and intuitive ways [39]. Many modern apps rely on human-computer interaction are developed. Because of the interest of psychology and cognitive science, for example, understanding user behavior such as body motions, particularly facial expression recognition is one of the main uses of computer vision that are gaining a lot of interest due to its different uses.
- **Augmented reality:** Nowadays, this research area is gaining a crucial and wide range of interest. Augmented reality [40] is the fact of making users in a virtual environment using different virtual objects. By generating perceptual information to enhance real-world objects in a variety of ways, including aural, visual, somatosensory, haptic, and olfactory. Object tracking is implemented in augmented reality to reach sometimes the interaction between objects or targets.
- **Robotic:** Nowadays, this research area is gaining a lot of interest due to its importance in many cases. Robots may replace human tasks by doing these tasks efficiently or sometimes better than humans. Vision tracking is relying on detection or tracking reached by robots [41].

#### IV. WEIGHTED SCORING MODEL

The Weighted Scoring Model is a mechanism for comparing objects and picking up between them, based on a set of criteria. The WSM Method is used to compare the algorithms. The application of this strategy could be done in the following steps (Fig. 3):

- Step 1: Determine which criteria will be used to perform each method.
- Step 2: Granting weights to groups of criteria based on the crucialness of each one.
- Step 3: Create a table with nominal values for each criterion of each distribution.
- Step 4: Create a table with weighted values for each criterion. The weight is expressed as a percentage. The overall weight is 100 percent.
- Step 5: A calculation of the product score of weights and nominal values is elaborated for each method.



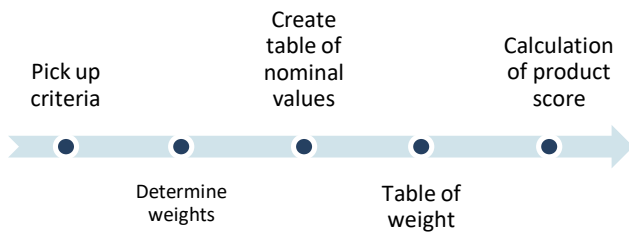


Fig. 3. WSM Process.

## V. COMPARISON OF MULTIPLE OBJECT TRACKING METHODS

### A. Comparison Criteria

The comparative criteria were chosen based on several video analysis investigations especially object tracking. Some criteria have been derived from the works cited below. Those are essentially the challenges are mentioned in Fig. 4 that all multi-object tracking methods and algorithms have in common to tackle them or to deal with them:

- **Occlusion:** Occlusion when the target is fully or sometimes partially or occluded. Complex interactions between objects result in both temporally and spatially occlusions, making object tracking a challenging problem [42]. In other words, in complicated scenes and crowded settings, it is common for an object's direction to be obscured, either by a background component such as a stationary scene or by other targets.
- **Illumination variation:** The illumination in the target area has been changed due to different issues such as obstacles that make birth of shadow, camera position and resolution, and the main source of illumination. Changes in illumination and posture are all essential elements that influence recognition rates, with the effect of light circumstances being particularly crucial. The direction of the light source may cause the image to be excessively bright or too dark, causing the algorithms to have difficulty in accurately obtaining crucial features [43].
- **Clutter:** The background in the vicinity of the target has the same color or texture as the target. A more complex scene results in more detection failures, which include undetected targets (false negatives) or spurious targets (false positives). As a result of the inadequate detection, data association accuracy is not reliable. Most algorithms are aiming at separating the background from targets [44].
- **Enter field of view:** Objects are moving in different directions with different motions. That makes the birth of more objects that enter the field of view and others leaving the field of view, and we are talking about object appearance and object disappearance, two different phenomena within a video for non-rigid objects.
- **Speed:** The motion or speed of the targets/ objects on the ground truth is large. The states of the targets, or

their positions, change with time. At first glance, we are unsure about their precise location. However, Things become more complicated when the object's direction changes. Each object has its speed and direction. The model should be able to deal with these fast changes [45].

- **Scale variation:** The bounding boxes shape of the object in different frames is ranging widely. The algorithm or model should be able to track the target from numerous perspectives and scales. One of the reasons for the disparity in performance is the large-scale variation amongst object instances, particularly the difficulty of recognizing very small objects [46].
- **Optimal:** This criterion is defining if the algorithm needs training or not, it needs resources or not.
- **Outlier:** some values are outside the range of what is expected, unlike the other data. We call these values outliers.

### B. Comparison Study

Table of nominal values is carried out. For each criterion, the value that corresponds to each method is assigned. These values are extracted from related work.

### C. Application of Weighted Scoring Model

The score of each challenge based on its existence is determined using the WSM method, as shown in Table III. The weight percentages are assigned based on the importance of the criterion. This collection of criteria is given priority due to their necessary requirement: number of objects tracked – Enter Field of view – Outlier. A weight of 12% is ascribed to each of their criteria. The following category of crucialness is given to the criteria Scale Variation – Illumination variation – Occlusion. A weight of 10% is ascribed to each of their criteria. And optimal, a weight of 20% is ascribed due to its importance and 14% associated to speed.

By applying the weight scoring model in Table IV, we are going to associate to each value a number relying on its crucialness  $n/a \rightarrow 0$ , all the criteria are ranging from 1 to 5, after getting these values we multiple each value by its weight as illustrated in Table III of weights. The result is mentioned in the table.

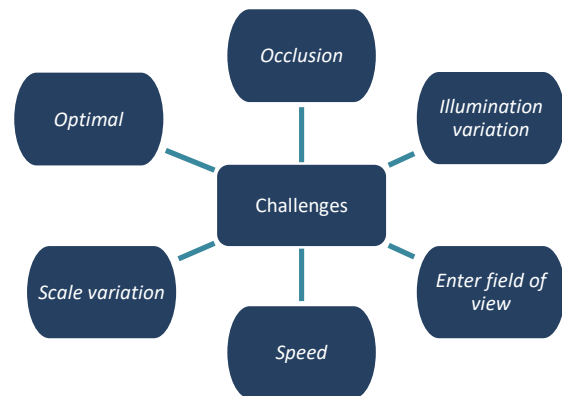


Fig. 4. Some Crucial Challenges of Object Tracking.

TABLE III. ASSOCIATED WEIGHTS TO EACH CRITERION

Criterion	Abbreviation	Proposed weights
Number of objects tracked	NO	12%
Illumination variation	IV	10%
Occlusion	OCC	10%
Enter Field of view	FOV	12%
Scale Variation	SV	10%
Speed	SP	14%
Optimal	OP	20%
Outlier	OUT	12%

TABLE IV. COMPARISON FOR DIFFERENT OBJECT TRACKING METHODS. (PT: POINT TRACKING, KT: KERNEL TRACKING, ST: SILHOUETTE TRACKING)

	TT	NO	OCC	IV	SP	SV	OP	FOV	OUT
Kalman Filter	PT	1	1	2	3	3	3	2	2
U Kalman Filter	PT	5	2	2	3	4	4	2	4
Particle Filter	PT	5	2	2	3	4	4	3	4
MGE tracker (D)	PT	5	2	1		4	1	2	4
GOA tracker	PT	5	2	2	n/a	3	2	1	n/a
JPDAF	PT	5	2	n/a	n/a	3	1	1	n/a
Multiple Hypothesis Tracking MHT	PT	5	2	n/a	n/a	4	1	3	4
Probabilistic Multiple Hypothesis Tracking PMHT	PT	5	2	3	3	3	1	3	n/a
Mean shift	KT	1	2	3	2	1	1	n/a	2
KLT tracker	KT	1	3	3	2	1	1		1
Simple Template matching	KT	1	1	1	2	1	2	1	1
SVM	KT	1	3	3	2	1	1	1	2
Layer/Layering based Tracking	KT	5	1	3	2	1	2	1	1
Eigentracking	KT	1	1	3	2	1	1	1	n/a
Shape Matching	ST	1	1	2	2	4	1	0	3
Contour Tracking	ST	5	1	3	2	4	1	1	3

#### D. Comparison

1) *SilhouettComparison of tracking methodologies based on the three categories kernel, silhouette, and point:* The spider Fig. 8 illustrates a comparison of silhouette, point, and kernel tracking category methodologies relied on the Number of objects tracked, Occlusion, illumination variation, speed, Scale variation, Enter field of view, Optimal, outliers.

2) *Comparison of the three categories:* According to Fig. 8 results, Multiple Hypothesis Tracking is the most favored algorithm within the three categories. Most of the advantages dealing with tracking multiple objects, handling occlusion, and give optimal solutions. Without ignoring that can handle object appearance and disappearance. There is a big lack of handling fast motion that should object-tracking

methodologies take into consideration. Regarding our Comparison of Silhouette tracking, methodologies see Fig. 6, they are a bit similar to each other and that describes that they belong to the same category.

Fig. 5 shows that SVM within Kernel tracking methodologies is dealing better with occlusion and outliers. Regarding point tracking, in Fig. 7 multiple MHT and PMHT are gaining the best results due to their speed and their way to handle outliers, occlusion.

3) *Comparison based on Table V:* The Table V illustrates see that contour tracking, particle filter, MGE tracker, GOA tracker, JPDAF, MHT, and PMHT are gaining the best results to the other methods.



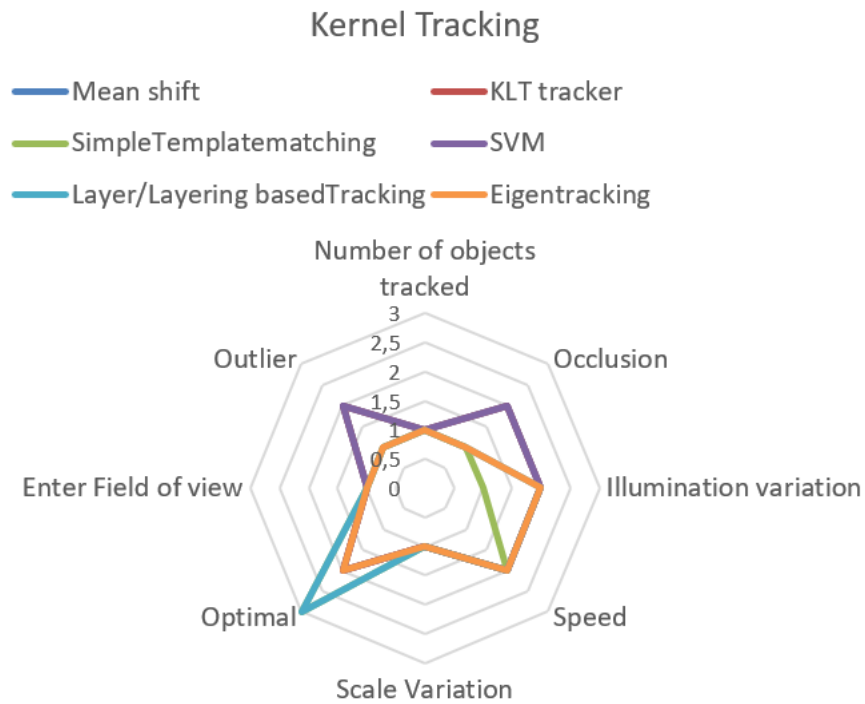


Fig. 5. Graph of Multicriteria Comparison of Kernel Tracking Methodologies.

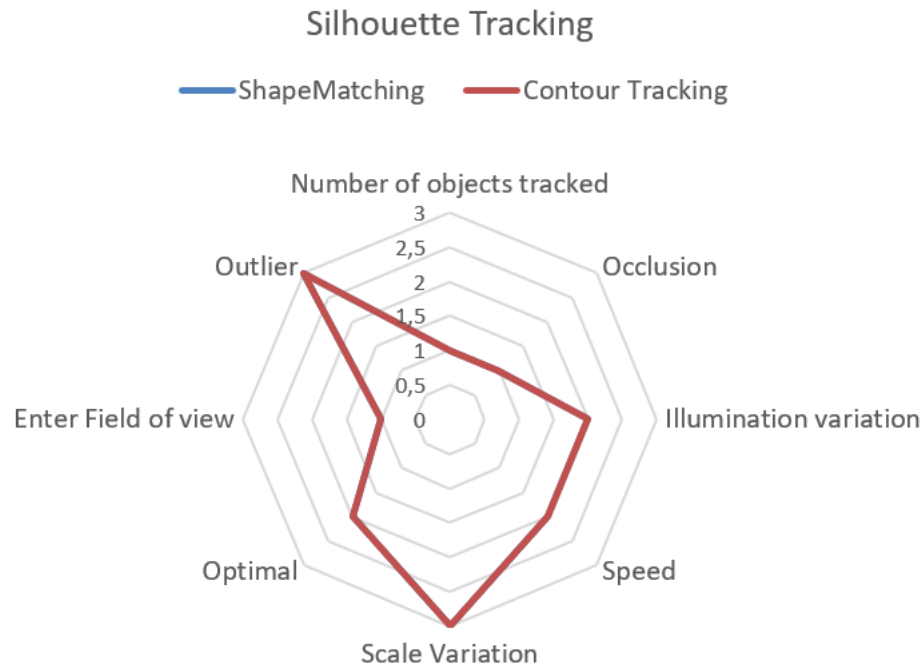


Fig. 6. Graph of Multicriteria Comparison of Silhouette Tracking Methodologies.

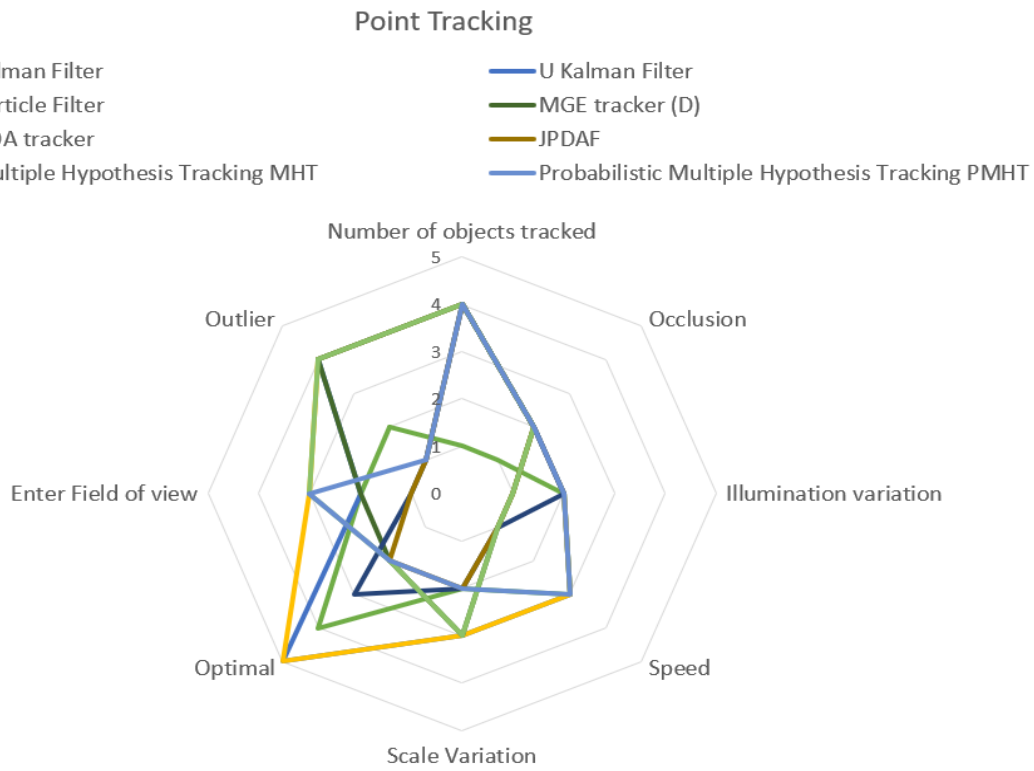


Fig. 7. Graph of Multicriteria Comparison of Point Tracking Methodologies.

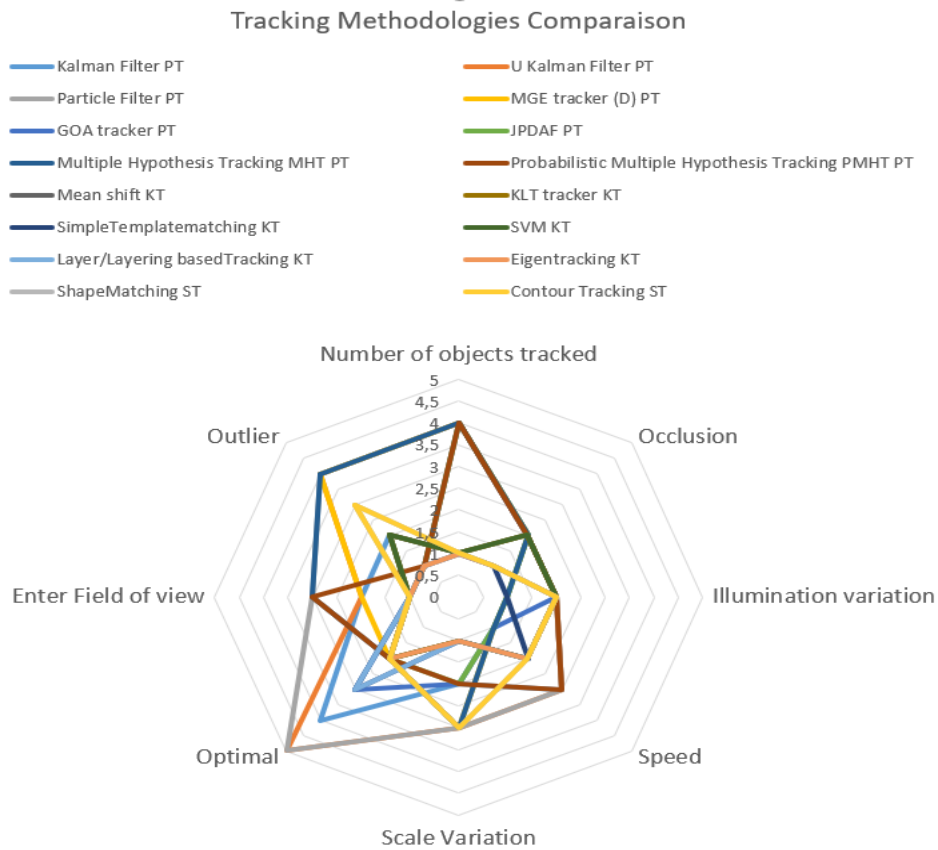


Fig. 8. Multicriteria Spider Graph.

TABLE V. COMPARING THREE OBJECT TRACKING CATEGORIES' METHODOLOGIES

Method	Score
Kalman Filter PT	2%
U Kalman Filter PT	10%
Particle Filter PT	10%
MGE tracker(D)	10%
GOA Tracker PT	10%
JPDAF	10%
MHT PT	10%
PMHT PT	10%
Mean shift KT	2%
KLT Tracker KT	2%
Simple Template Matching KT	2%
SVM KT	2%
Layer/ Layering based Tracking KT	10%
Eigen tracking KT	2%
Shape Matching ST	2%
Contour Matching ST	10%

## VI. DISCUSSION

According to the previous results, some algorithms are dealing better with multiple objects such as Particle Filter, MGE tracker, GOA tracker, JPDAF, multiple hypothesis tracking, and so on. Point and Silhouette methodologies deal better with scale variation. However, they do not handle fully occlusion. Regarding kernel tracking methodologies they handle occlusion but they do not go forward with handling multiple objects.

On the one hand, relying on the comparative study-based WSM method Point Tracking is better to deal with small objects and tracking multiple objects, however, is not good to deal with occlusion and object appearance and disappearance.

On the other hand, Kernel tracking methodologies are handling better occlusion at the meanwhile estimate the object motion but they don't deal with multiple objects also this method is not a good choice to handle multiple objects.

Regarding Silhouette Tracking methodologies are dealing better with tracking object based on silhouette. This method is accurate to track objects or targets because of its ability to track different shapes. But these techniques are not good to handle occlusion at the same time these methods are not better to handle multiple objects.

Multiple object tracking is still facing a wide range of challenges. We could not find one category or method that deals with all of these challenges. Each algorithm or method is mastering to handle a challenge. Hope to find a hybrid method that deal with all these issues.

## VII. CONCLUSION

A comparison of object tracking methods is has been presented in this research. This project has begun with the identification of a group of relevant works that are implementing these methods and the proposal of object

tracking methods that fits the needs of various areas. Then a list of research areas that are concerned by these works has been discussed. There is also a set of criteria that this comparison has been based on. Using the Weight Score Model, the scores for each algorithm evaluated have been obtained. Various scores or results not only have been assisting us in determining an overall ranking amongst these platforms, but they have also revealed their internal strengths and shortcomings concerning each criterion.

This research has identified a collection of failures that object tracking algorithms face. The main problems that platforms face are the number of targets and it is optimal. Researcher must provide an algorithm that considers optimal aspects to provide an accurate model or algorithm of object tracking to the wide range of the research areas, such as surveillance, robotic, and other fields. The researchers must create a model or algorithm that could tackle and operate the different challenges, dealing with the different problems carefully due to the crucialness of this such as the biomedical one. Finally, researchers must embrace the resources used, which is a set of tools for increasing product quality and lowering development costs while also ensuring optimal delivery for users.

Future work will focus on the measurement of the accuracy of each method under each challenge. Our main goal is to implement, test, compare and analyze the results of point tracking methods, kernel tracking methods, silhouette tracking methods that we are going to implement to track multiple objects/ targets. Our main aim is that we are going to decorticate models or algorithms relying on their architecture, iteration, practical use, and their uniqueness. At the end, a summary table of the comparison is developed.

## REFERENCES

- [1] Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Comput. Surv.*, vol. 38, no. 4, 2006, doi: 10.1145/1177352.1177355.
- [2] D. H. Santosh and P. G. K. Mohan, "Multiple objects tracking using Extended Kalman Filter, GMM and Mean Shift Algorithm-A comparative study," *Proc. 2014 IEEE Int. Conf. Adv. Commun. Control Comput. Technol. ICACCCT 2014*, no. 978, pp. 1484–1488, 2015, doi: 10.1109/ICACCCT.2014.7019350.
- [3] G. A. Einicke and L. B. White, "Robust extended Kalman filtering," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2596–2599, 1999, doi: 10.1109/78.782219.
- [4] R. Sicre and H. Nicolas, "Improved Gaussian mixture model for the task of object tracking," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 6855 LNCS, no. PART 2, pp. 389–396, 2011, doi: 10.1007/978-3-642-23678-5\_46.
- [5] S. W. Chou, C. H. Hsieh, B. J. Hwang, and H. W. Chen, "A modified mean shift algorithm for visual object tracking," *2013 Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. APSIPA 2013*, 2013, doi: 10.1109/APSIPA.2013.6694229.
- [6] S. Moon, J. Lee, D. Nam, H. Kim, and W. Kim, "A comparative study on multi-object tracking methods for sports events," *Int. Conf. Adv. Commun. Technol. ICACT*, pp. 883–885, 2017, doi: 10.23919/ICACTION.2017.7890221.
- [7] K. Kale, S. Pawar, and P. Dhulekar, "Moving object tracking using optical flow and motion vector estimation," *2015 4th Int. Conf. Reliab. Infocom Technol. Optim. Trends Futur. Dir. ICRITO 2015*, pp. 2–7, 2015, doi: 10.1109/ICRITO.2015.7359323.
- [8] T. K. Kim, T. Woodley, B. Stenger, and R. Cipolla, "Online multiple classifier boosting for object tracking," *2010 IEEE Comput. Soc. Conf.*

- Comput. Vis. Pattern Recognit. - Work. CVPRW 2010, pp. 1–6, 2010, doi: 10.1109/CVPRW.2010.5543889.
- [9] L. Wietzke and G. Sommer, “Robust object Tracking with online multiple instance learning,” *Comput. Vis.*, vol. 33, no. 8, pp. 1619–1632, 2009, [Online]. Available: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=5206784](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5206784).
- [10] L. Gao, Y. Li, and J. Ning, “Improved kernelized correlation filter tracking by using spatial regularization,” *J. Vis. Commun. Image Represent.*, vol. 50, no. July 2017, pp. 74–82, 2018, doi: 10.1016/j.jvcir.2017.11.008.
- [11] A. S. N. Rani, V. Maik, and B. Chithravathi, “Robust object tracking using kernelized correlation filters (KCF) and Kalman predictive estimates,” *RTEICT 2017 - 2nd IEEE Int. Conf. Recent Trends Electron. Inf. Commun. Technol. Proc.*, vol. 2018-January, pp. 587–591, 2017, doi: 10.1109/RTEICT.2017.8256664.
- [12] N. Almohaimeed and M. Prince, “A Comparative Study of different Object Tracking Methods in a Video,” *Int. J. Comput. Appl.*, vol. 181, no. 41, pp. 1–8, 2019, doi: 10.5120/ijca2019918470.
- [13] P. R. Gunjal, B. R. Gunjal, H. A. Shinde, S. M. Vanam, and S. S. Aher, “Moving Object Tracking Using Kalman Filter,” *2018 Int. Conf. Adv. Commun. Comput. Technol. ICACCT 2018*, pp. 544–547, 2018, doi: 10.1109/ICACCT.2018.8529402.
- [14] X. Wang, T. Li, S. Sun, and J. M. Corchado, “A survey of recent advances in particle filters and remaining challenges for multitarget tracking,” *Sensors (Switzerland)*, vol. 17, no. 12, pp. 1–21, 2017, doi: 10.3390/s17122707.
- [15] D. Exner, E. Bruns, D. Kurz, A. Grundhöfer, and O. Bimber, “Fast and robust CAMShift tracking,” *2010 IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. - Work. CVPRW 2010*, pp. 9–16, 2010, doi: 10.1109/CVPRW.2010.5543787.
- [16] D. Chatterjee and S. Chandran, “Comparative study of camshift and KLT algorithms for real time face detection and tracking applications,” *Proc. - 2016 2nd IEEE Int. Conf. Res. Comput. Intell. Commun. Networks, ICRCICN 2016*, pp. 62–65, 2017, doi: 10.1109/ICRCICN.2016.7813552.
- [17] F. Jurie and M. Dhome, “Real Time Robust Template Matching,” no. September 2002, pp. 10.1-10.10, 2013, doi: 10.5244/c.16.10.
- [18] C. I. Patel and R. Patel, “Contour Based Object Tracking,” *Int. J. Comput. Electr. Eng.*, no. July, pp. 525–528, 2012, doi: 10.7763/ijcee.2012.v4.549.
- [19] I. Pathan and C. Chauhan, “A Survey on Moving Object Detection and Tracking Methods,” vol. 6, no. 6, pp. 5212–5215, 2015.
- [20] M. W. Park, A. Makhmalbaf, and I. Brilakis, “Comparative study of vision tracking methods for tracking of construction site resources,” *Autom. Constr.*, vol. 20, no. 7, pp. 905–915, 2011, doi: 10.1016/j.autcon.2011.03.007.
- [21] J. J. Athanasiou and P. Suresh, “Systematic Survey on Object Tracking Methods in Video,” *Int. J. Adv. Res. Comput. Eng. Technol. Oct.*, vol. 1, no. 8, pp. 242–247, 2012.
- [22] P. Panchal, G. Prajapati, S. Patel, and H. Shah, “a Review on Moving Object Detection and Tracking Methods,” *Int. J. Adv. Eng. Res. Dev.*, vol. 3, no. 02, pp. 7–12, 2015, doi: 10.21090/ijaerd.c1021308.
- [23] A. Makhmalbaf, M. Park, J. Yang, I. Brilakis, and P. A. Vela, “2D Vision Tracking Methods ’ Performance Comparison for 3D Tracking of Construction Resources College of Architecture , Georgia Institute of Technology , Atlanta , GA . 30332-0355 ; School of Civil and Environmental Engineering , Georgia Institute of Tech,” pp. 459–469.
- [24] S. H. Shaikh, K. Saeed, and N. Chaki, “Moving object detection approaches, challenges and object tracking,” *SpringerBriefs Comput. Sci.*, vol. 0, no. 9783319073859, pp. 5–14, 2014, doi: 10.1007/978-3-319-07386-6\_2.
- [25] M. Isard and A. Blake, “CONDENSATION - Conditional Density Propagation for Visual Tracking,” *Int. J. Comput. Vis.*, vol. 29, no. 1, pp. 5–28, 1998, doi: 10.1023/A:1008078328650.
- [26] Z. Kalal, “Tracking Learning Detection, PhD Thesis,” no. April, 2011.
- [27] X. Li, K. Wang, W. Wang, and Y. Li, “A multiple object tracking method using Kalman filter,” *2010 IEEE Int. Conf. Inf. Autom. ICIA 2010*, vol. 1, no. 1, pp. 1862–1866, 2010, doi: 10.1109/ICINFA.2010.5512258.
- [28] P. Dendorfer et al., “MOT20: A benchmark for multi object tracking in crowded scenes,” arXiv, pp. 1–7, 2020. S. Sushmitha, N. Sathesh, and V. Kanchana, “Multiple car detection, recognition and tracking in traffic,” *2020 Int. Conf. Emerg. Technol. INCET 2020*, pp. 1–5, 2020, doi: 10.1109/INCET49848.2020.9154107.
- [29] H. E. D. Mohamed et al., “MSR-YOLO: Method to Enhance Fish Detection and Tracking in Fish Farms,” *Procedia Comput. Sci.*, vol. 170, no. 2019, pp. 539–546, 2020, doi: 10.1016/j.procs.2020.03.123.
- [30] A. Fernandes, M. Baptista, L. Fernandes, and P. Chaves, “Drone, Aircraft and Bird Identification in Video Images Using Object Tracking and Residual Neural Networks,” *Proc. 11th Int. Conf. Electron. Comput. Artif. Intell. ECAI 2019*, no. April, 2019, doi: 10.1109/ECAI46879.2019.9042167.
- [31] F. Delestro, F. Cloppet, P. Paul-gilloteaux, T. Pr, and F. Delestro, “A multiple cell tracking method dedicated to the analysis of memory formation in vivo To cite this version: HAL Id: tel-02178841 eat A multiple cell tracking method dedicated to the analysis of memory formation in vivo,” 2019.
- [32] A. Ellouze, M. Ksantini, F. Delmotte, and M. Karray, “Multiple Object Tracking: Case of Aircraft Detection and Tracking,” *16th Int. Multi-Conference Syst. Signals Devices, SSD 2019*, pp. 473–478, 2019, doi: 10.1109/SSD.2019.8893202.
- [33] B. Kovacevic, D. Ivkovic, and Z. Radosavljevic, “IMMPDAF Approach for L-Band Radar Multiple Target Tracking,” *2020 19th Int. Symp. INFOTEH-JAHORINA, INFOTEH 2020 - Proc.*, no. March, pp. 18–20, 2020, doi: 10.1109/INFOTEH48170.2020.9066328.
- [34] M. Fiaz, A. Mahmood, S. Javed, and S. K. Jung, “Handcrafted and deep trackers: Recent visual object tracking approaches and trends,” *ACM Comput. Surv.*, vol. 52, no. 2, 2019, doi: 10.1145/3309665.
- [35] D. Sudha and J. Priyadarshini, “An intelligent multiple vehicle detection and tracking using modified vibe algorithm and deep learning algorithm,” *Soft Comput.*, vol. 24, no. 22, pp. 17417–17429, 2020, doi: 10.1007/s00500-020-05042-z.
- [36] S. Challa, M. R. Morelande, D. Mušicki, and R. J. Evans, *Fundamentals of object tracking*, vol. 9780521876. 2011.
- [37] A. Yoganandhan, S. D. Subhash, J. Hebison Jothi, and V. Mohanavel, “Fundamentals and development of self-driving cars,” *Mater. Today Proc.*, vol. 33, no. xxxx, pp. 3303–3310, 2020, doi: 10.1016/j.matpr.2020.04.736.
- [38] S. S. Rautaray and A. Agrawal, “Vision based hand gesture recognition for human computer interaction: a survey,” *Artif. Intell. Rev.*, vol. 43, no. 1, pp. 1–54, 2015, doi: 10.1007/s10462-012-9356-9.
- [39] H. Sarmadi, R. Muñoz-Salinas, M. Álvaro Berbís, A. Luna, and R. Medina-Carnicer, “Joint scene and object tracking for cost-effective augmented reality assisted patient positioning in radiation therapy,” arXiv, 2020.
- [40] H. Chen, C. Xue, S. Liu, Y. Sun, and Y. Chen, “Multiple-object tracking based on monocular camera and 3-D lidar fusion for autonomous vehicles,” *IEEE Int. Conf. Robot. Biomimetics, ROBIO 2019*, no. December, pp. 456–460, 2019, doi: 10.1109/ROBIO49542.2019.8961438.
- [41] Y. Huang and I. Essa, “Tracking Multiple Objects Through Occlusions 2 . Related Work.”
- [42] G. Yan, J. Li, J. Li, Q. Ma, and M. Yu, “Illumination variation in face recognition a review,” *ICINIS 2009 - Proc. 2nd Int. Conf. Intell. Networks Intell. Syst.*, pp. 309–311, 2009, doi: 10.1109/ICINIS.2009.85.
- [43] D. Musicki and B. La Scala, “Multi-target tracking in clutter without measurement assignment,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 3, pp. 877–896, 2008, doi: 10.1109/TAES.2008.4655350.
- [44] K. Host, M. Ivašić-Kos, and M. Pobar, “Tracking handball players with the Deepsort algorithm,” *ICPRAM 2020 - Proc. 9th Int. Conf. Pattern Recognit. Appl. Methods*, pp. 593–599, 2020, doi: 10.5220/0009177605930599.
- [45] B. Singh and L. S. Davis, “An Analysis of Scale Invariance in Object Detection - SNIP,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 3578–3587, 2018, doi: 10.1109/CVPR.2018.00377.