ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 4, April 2023

A Study on Real Time Object Tracking

[1] Jaismeen, [2] Harjeet Singh, [3] Sangeeta

[1] [2] [3] Chandigarh University, Ajitgarh, Punjab, India.

Corresponding Author Email: [1] jaistiwana.231@gmail.com, [2] myengineering25@gmail.com, [3] sangeetae7531@cumail.in

Abstract*— One of the most crucial jobs in computer vision is object tracking, which has many real-world uses in fields like robotics, automated car tracking, and traffic monitoring. Using the camera in video sequences, object tracking is the proceeding of finding moving things over time. Object recognition is the first step in tracking and aims to identify or find the moving object in the frame. Following that, the observed object can be categorized as moving animals, people, trees that are swaying, birds, and some other objects. The object tracking method divides the area of interest in a picture, tracks its movements and position, and looks for any blobs. Numerous studies have been conducted recently, but this is due to a variety of difficulties, including shadowing, varying lighting conditions, rapid motion, etc. research in this field is still ongoing. The identification and monitoring of multiple objects in dynamic environments have lately attracted the attention of computer vision researchers. In this study, a tracking algorithm that incorporates the object characteristics mentioned above is reviewed and examined. The purpose of this article is to analyse and evaluate the prior approaches for object recognition and tracking using video sequences over various stages. Deep learning has received a lot of attention lately due to its excellent performance.

***Index Terms*—Deep Learning, Object Detection, Object Identification, Object Tracking.**

I. INTRODUCTION

One of the significant jobs in computer vision is object tracking, which aims to find and follow things in picture sequences. In order to monitor an item, it must be identified in the initial frame of the movie and then tracked in the succeeding frames. Object monitoring has a variety of uses. The technique of segmenting an object or group of objects from the recorded video with the intention of monitoring their motion, occlusion, and orientation is known as object tracking.

A technique for categorizing or forecasting the class of particular items in a video frame is called object classification. The things are divided into categories such as people, creatures, birds, cars, and moving items [1]. Object tracking, the last stage, is the identification of target objects over a series of video frames [2]. Determining the objects in a film and reading their paths correctly is what is referred to as object tracking.

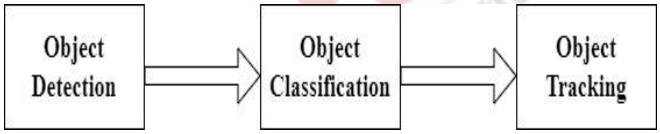


Figure 1. Stages of Object Tracking

Object tracking is useful in a variety of fields, including robotics (such as the ASIMO humanoid robot [3]), self-sustaining vehicle monitoring (such as path-tracking [4]), medical diagnosis systems (such as the medical instrument control [5]), traffic monitoring (such as traffic flow monitoring [6] and activity recognition. (such as recognition of human activity [7]). There are many difficulties in object tracking, which has prompted continuing study in this field. Low resolution, backdrop groups, occlusion, rapid motion, changing target location, and other difficulties are a few of these.

It is important to attempt to meet these standards as much as feasible when developing each tracking program because there are many parts of tracking that present significant challenges. Here is a summary of a few of these characteristics.

Robustness: When a tracking system is robust, it can continue to follow the target despite challenging circumstances like backdrop debris, occlusion, and changing lighting. Adaptability: The subject is also susceptible to changes, such as complicated and sudden movements, in addition to changes in the surroundings. The tracking device must have the capacity able to recognize and follow the target's present apparent traits in order to resolve this issue. Information processing in real time: A system that works with picture sequences needs to be able to handle information quickly. A high-performance method must therefore be used.

The various stages or crucial components of object tracking are object detection, object classification, and monitoring. Any method for tracking an item needs to follow some fundamental stages. Fig. 1 shows the basic steps. The following parts provide a detailed explanation of each step. The study of automated object recognition and monitoring changed due to the advancement of technology.

The paper is organised as follows: Section 1 is the introduction; Section 2 is the literature review; Section 3 is the methodology; Section 4 is the result and discussion; and Section 5 is the conclusion.

II. LITERATURE REVIEW

Various techniques to recognize and follow the items in the video frames have been suggested by numerous writers. There aren't many algorithms that use image processing

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 4, April 2023

methods, and with the latest advancements in technology, deep learning-based approaches are showing encouraging results. Several techniques for object recognition and object tracking are discussed in this paper.

a. Image processing techniques for Tracking and Detection of various Objects

Methods of Object detection: It is a necessary first stage in all object tracking in order to recognize things of interest in video sequences. The process of finding moving objects in a series of film frames is called object recognition. Fig 2 displays various kinds of Object Detection techniques.

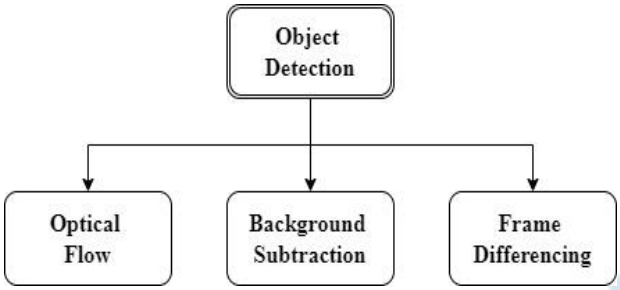


Figure 2. Different Object Detection methods

The author of [8] proposes an underlying subtraction-based image recognition method that identifies different objects with the greatest pixel-level difference among the frame and video framework before it. A backdrop with multiple models is incompatible. A sliding window method was used by the author in [9] to generate the background picture based on the temporal change of each and every pixel in the file containing L Video frames.

According to optical flow distribution characteristics of the picture, the author [10] suggested a technique for determining the flow of light through the area of a picture and conducting the clustering process. A Simplest backdrop Subtraction was recommended by Aldhaheri and Edirisinghe (2014) and Haritaoglu et al. (2000) [11]. The fundamental idea was that the backdrop frame is removed from the current frame. This approach provides precision ranging from modest to high.

The optical flow method, which makes use of the pixel-level optical flow distribution features, was proposed by Krishna et al. in 2011 [12]. This method provides all moving info but necessitates more computations. It is evident from the foregoing that the object identification relies on backdrop subtraction, optical movement, and frame difference.

Methods of Object Classification: Object recognition is finished once the first stage of monitoring is over. Any moving item can be spotted, including people, animals, swaying trees, birds, and vehicles. The second step of object tracking, Object classification can be used to classify this discovered item. Fig 3 displays various kind of Object Classification approaches.

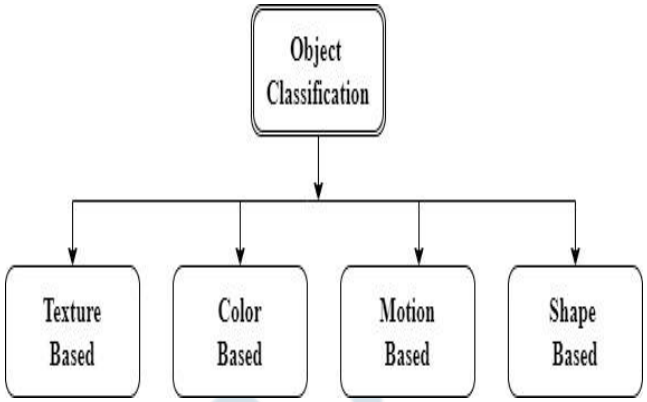


Figure 3. Different Object Classification methods

The author [13] has suggested a motion-based categorization method that doesn't need a template for pre-set patterns but struggles to recognize motionless objects. The categorization method based on movements is moderately accurate.

The idea of texture-based image categorization was put forth by the author in [14], who computed the gradient orientation that happens in particular areas of the picture. This method offers higher quality but requires more calculation time. Using a Gaussian Mixture Model, Hu et al. (2004) characterized the color distribution in a picture sequence and separated the backdrop and objects.

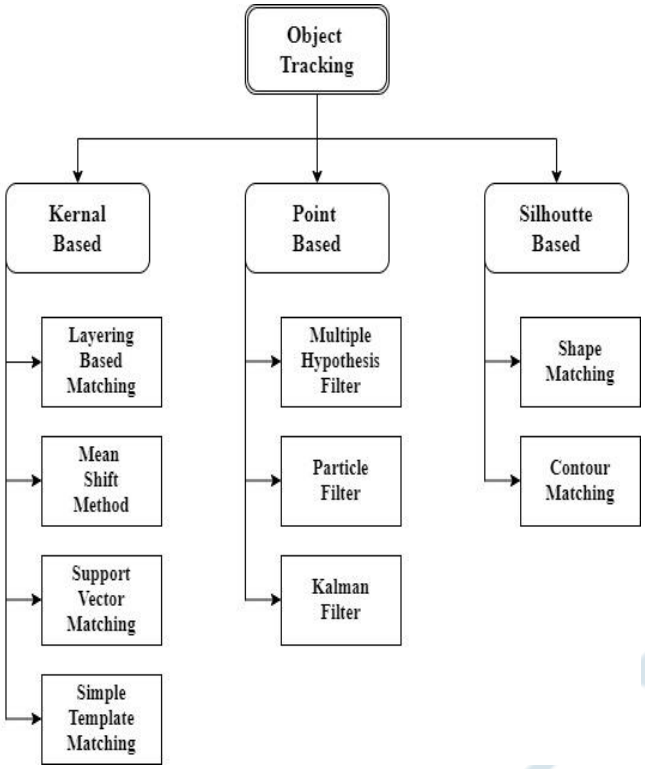
Color-based object recognition was suggested by X.Dong et al. [15]. With this technique, moving objects and their shadows can be extracted for further use by first being identified in advance based on luminance and chromatic distortion. The author [16] suggested a Shape-based categorization, a straightforward pattern matching method that can be used with graph matching methodology.

Methods of Object Tracking: The method of subdivision of an object that is interesting from a set of video sequences or sequential frames is known as object tracking. Due to information loss, complicated movement of objects or irregular object structure, noise in the picture, real-time requirements, and occlusion, tracking is a challenging issue in image processing. Fig 4 displays the various kinds of Object Classification approaches.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 4, April 2023



Figue 4. Different Object Tracking methods

The author of [17] provides a thorough overview of object tracking methods. The three different kinds of object monitors are point tracking, kernel tracking, and outline tracking. For the purposes of point tracking, focal points symbolize moving things. Oiwa et al. (2016) [18] proposed the probabilistic background model for tracking the object from video frames. When compared to previous methods, the simulation results show that this strategy is extremely accurate and effective. The goal of this study should be to improve object tracking accuracy while also accelerating object tracking speed.

A Particle Filter method was recommended by Hu et al. in 2004 [19]. The Recursive Bayes filtration method was employed. It produces positive outcomes for occultation and complicated backgrounds. A simple template matching technique was used by Cucchiara et al. in 2003 [20]. This makes use of the video technique known as Matching area of interest, which can handle partial occlusion.

Zhang et al. (2016) [21] proposed an approach combining frame difference and non-parametric method for video analysis traceability. The test findings showed that this approach outperformed GMM and traditional frame difference. It also, has the ability to eliminate background noise, which enables us to accurately recognise moving objects in applications like tracking analysis for food and agricultural goods. This study must increase the traceability system's capability and offer a visible supply chain in order to ensure the user's safety.

b. Deep Learning based techniques for Tracking and Detection of various Objects

Object detection and object tracking approaches which are based on deep learning, have gained a lot of attention in comparison to previous approaches due to the computer vision research's quick growth.

Object Recognition based on Deep-learning:

Two distinct methods exist: (i) single stage detectors and (ii) two stage detectors.

Contrasted with one stage detectors, two stage detectors are more precise but one stage detectors are quicker. The two step detectors include the Fast R-CNN, RCNN (Region Convolutional Neural Network), Granulated CNN, Mask R-CNN, Faster R-CNN, and Granulated RCNN. The single stage devices are the YOLO series, M2Det and SSD. Fig 5 displays the single stage and two stage methods based on deep learning techniques.

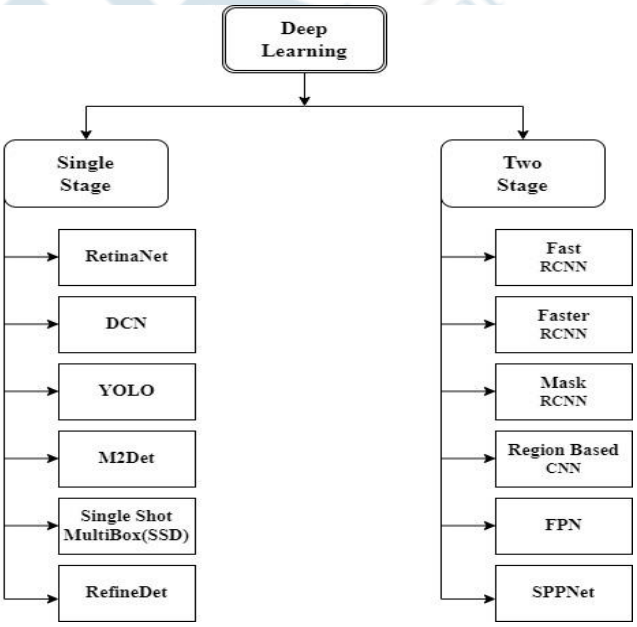


Figure 5. Various Single and Two stages methods

Object Tracking based on Deep-learning:

One of the oldest and most commonly used object monitoring frameworks and algorithms is Deep SORT. Bewley et al. [22] suggested a method using Simple Online Real-Time Tracking, (SORT). By assigning a distinctive id to each bounding rectangle, SORT keeps account of each discovery. When an object disappears because of obstruction, misidentification, or other reasons, the tracker gives a new id and starts following newly found objects.

The technique for tracking various objects with a human-centred emphasis was suggested by Fan et al. in [23]. In order to gather geographic and temporal data, neural networks also known as Convolutional neural networks are used during the offline training procedure. Some of top studies on deep learning-powered recognition and tracking of objects are discussed in [24].

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 4, April 2023

In order to track objects that move in surveillance footage, the GCNN technique known as Growing Competitive Neural Network created by Lobito et al. [25] was derived. This technique was basically derived from the traditional Competitive Neural Network viewpoint. In [26], the author suggested using the K mean algorithm for object identification and tracking. He applied the "good feature to track" technique to retrieve the characteristic data from each and every frame.

Table 1: Some of the Object Tracking approaches

Methods	Algorithm	Accuracy	Computational Time	Comments
Mean Shift Method	Optimal gradient decline	Low	Moderate	Having less calculations, so used in real time scenarios. (Tao Zhang et al., 2010)
Template selection strategy (Reinforcement Learning)	Novel tracking algorithm	Moderate to High	Moderate	Policy network that is capable of selecting the proper template from a template collection to monitor any frame (Janghoon Choi et al., 2018)
Particle Filter	Recursive Bayes filtering	Low to High	High	For complex background and occultation provides the better results (Hu et al., 2004)
Classical competitive neural network approach	Growing competitive neural network (GCNN)	Moderate	Moderate	Used for moving objects in video surveillance (Lobito et al., 2009)
Kalman Filter	Faster-RCNN Detector and KCF Tracker	Moderate	High	To identify pedestrian items, obtain their dimensions and scale from tracker outputs, suitable for multiple moving objects. (Fan Bu et al., 2016)
Shape matching	Hough Transform	High	High	Performance can be improved and less affected towards variance of glance (Karasulu, 2010)

III. METHODOLOGY

In software for computer vision Detecting objects and monitoring them go hand in hand. The procedure of finding a particular case of interest within a group of suspect frames is known as object detection. The process of figuring out an object's route or course across several frames is known as object tracking. Fig 6 below displays the structure for basic object recognition and tracking. The dataset is split into two parts in order to identify the items in the picture. 20% of the pictures are used for assessment, and the remaining 80% are used for training. In order to identify objects within an image, deep learning methods are used. The item has a bounding frame drawn around it. The neural network is given the identified bounding area as a reference to help them follow. Bounding box is tracked using object tracking across numerous frames.

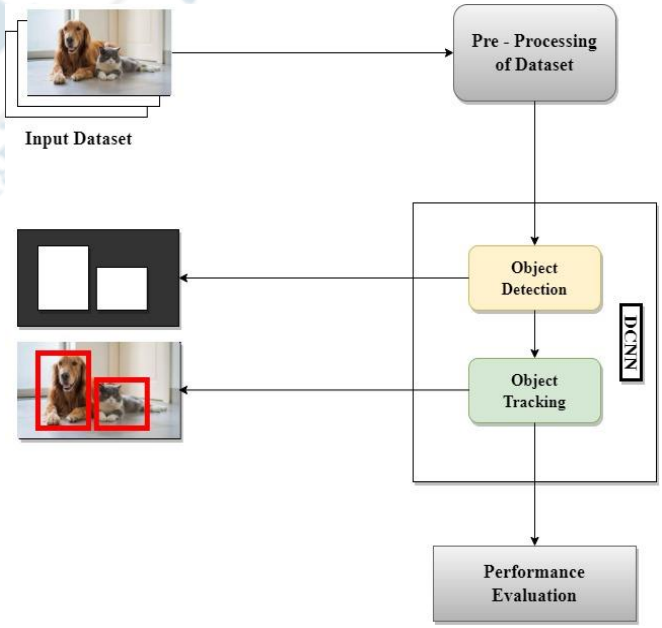


Figure 6. Object Detection & Tracking based on Deep Networks

IV. RESULT

This part includes a summary of popular detectors, sensors, and their combinations that have been written about by various writers. Weng et al. [27] used an organically

formed method to improve tracking accuracy for real-world items. However, it needs to be faster and more accurate by switching out the SURF algorithm for another highly effective method. A method for face recognition was devised by Sarkar et al. [28] using the abundance of color information. The model has the capability to identify features in surveillance video clips or low-resolution pictures. The suggested model also recognizes how the faces are aligned in the pictures.

By merging frame difference and non-parametric methods, Zhang et al. [29] created an innovative algorithm for enhancing video analysis traceability. To guarantee the safety of food and goods connected to agriculture, this research must improve the traceability system's capabilities and provide a visual supply chain for the general public. With a point level assessment range of 30-100 meters, Tianya Zhang et al.'s [30] suggested technique for LiDAR tracking & detection has total counting recall of 70.08%, precision of 97.69%, and accuracy of 94.74%. With six film segments,

Akansha Bathija et al. [31] used the SORT tracker and YOLO detector, yielding recall of 94.4%, precision of 95%, and accuracy of 85%. Using the MOT2015 dataset, Sankar K. Pal et al. [32] did research on multi-object identification and tracking and found that the combo of Deep SORT & Faster R-CNN is better to supplementary tracking evaluation measures. This shows that the deep learning-based algorithms are more appropriate to use in object tracking as compared to the image processing because it gives the more accuracy in different situations.

V. CONCLUSION

A short analysis of object detection and tracking methods based on deep learning and image processing is included in this piece of writing. Mainly illustrated the fundamental object recognition and tracking procedure while primarily concentrating on single stage & two stage object detectors. When object recognition and monitoring are combined, some important conclusions are drawn. By utilising Deep SORT tracking and two stage detectors, it is possible to more precisely extrapolate the features of the items from the findings provided by different writers. A number of writers used various datasets, including INRIA Person, MOT2015, the KITTI 3D object recognition benchmark data set, and the COCO Dataset, to demonstrate how well the network performed when compared to cutting-edge techniques. Because of the ongoing updating of powerful processing hardware, deep learning-based object recognition and tracking is rapidly flourishing. Deep learning networks consist of numerous levels and extract multiple characteristics of the object in each layer, allowing them to follow objects with greater precision than shallow networks due to their higher model complexity and number of layers. Deep learning may not always be the best option, but by understanding the benefits and drawbacks of each approach, one can determine which approach will be most effective for fixing problems. Since object detection comes before object tracking, object detection accuracy over the video frame is an extremely crucial element in tracking accuracy.

REFERENCES

[1] Himani S. Parekh, Darshak G. Thakore, Udesang K. JaJiya, " A Survey on Object Detection and Tracking Methods", International Joulllal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 2, February 2014.

[2] S. Ojha and S. Sakhare, "Image processing techniques for object tracking in video surveillance- A survey," 2015 International Conference on Pervasive Computing (ICPC), 2015, pp. 1-6.

[3] Y. Sakagami, R. Watanabe, C. Aoyama, S. Matsunaga, N. Higaki, and K. Fujimura, “The intelligent ASIMO: System overview and integration,” in IEEE/RSJ international conference on intelligent robots and systems, 2002, vol. 3, pp. 2478–2483.

[4] M. Menze and A. Geiger, “Object scene flow for autonomous vehicles,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3061–3070.

[5] S. Walker et al., “Systems and methods for localizing, tracking and/or controlling medical instruments,” Apr-2017.

[6] B. Tian, Q. Yao, Y. Gu, K. Wang, and Y. Li, “Video processing techniques for traffic flow monitoring: A survey,” in 2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2011, pp. 1103–1108.

[7] R. Bodor, B. Jackson, and N. Papanikolopoulos, “Vision-based human tracking and activity recognition,” in Proc. of the 11th Mediterranean Conference on Control and Automation, 2003, vol. 1.

[8] RupaJi S.Rakibe, Bharati D.Patil, "Background Subtraction Algorithm Based Human Motion Detection", Intelllational Journal of Scientific and Research Publications, May 2013.

[9] Abhishek Kumar Chauhan, Prashant Krishan, "Moving Object Tracking Using Gaussian Mixture Model and Optical Flow", International Journal of Advanced Research in Computer Science and Software Engineering, April 2013.

[10] Balaji, S.R.; Karthikeyan,S “A Survey on moving object tracking using Image processing” 11th International Conference on Intelligent Systems and Control (ISCO),2017,469-474

[11] Tiwari, Mukesh, and Rakesh Singhai. "A review of detection and tracking of object from image and video sequences." *Int. J. Comput. Intell. Res* 13.5 (2017): 745-765.

[12] Ram, M. Raghu, et al. "A novel approach for motion artifact reduction in PPG signals based on AS-LMS adaptive filter." *IEEE Transactions on Instrumentation and Measurement* 61.5 (2011): 1445-1457.

[13] Hitesh A Patel, Darshak G Thakore,"Moving Object Tracking Using Kalman Filter", International Journal of Computer Science and Mobile Computing,April 2013, pg.326 – 332.

[14] Jae-Yeong Lee; Wonpil Yu, "Visual tracking by partition-based histogram back projection and maximum support criteria," Robotics and Bio-mimetics (ROBIO), pp.2860,2865, Dec-2011.

[15] X. Dong, K. Wang and G. Jia, "Moving Object and Shadow Detection Based on RGB Color Space and Edge Ratio," 2009 2nd International Congress on Image and Signal Processing, 2009, pp. 1-5, doi: 10.1109/CISP.2009.5301770.

[16] Salve, Suhas G.; Jondhale, Kal, "Shape matching and object recognition using shape contexts" 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT 2010) ,pp.471–474,2010.

[17] Z. Soleimanitaleb, M. A. Keyvanrad and A. Jafari, "Object Tracking Methods:A Review," 2019 9th International Conference on Computer and Knowledge Engineering (ICCKE), 2019, pp. 282-288, doi: 10.1109/ICCKE48569.2019.8964761.

[18] D. Oiwa, S. Fukui, Y. Iwahori, T. Nakamura and M.K. Bhuyan, "Tracking with Probabilistic Background Model by Density Forests", *Proc. of IEEE/ACIS*, pp. 391-396, 2016.

[19] Hu, W., Tan, T., Wang, L. & Maybank, S., 2004. A Survey on Visual Surveillance of Object Motion and Behaviors. IEEE Trans. Syst. Man Cybern. Part C (Applications Rev. 34, 334–352. doi:10.1109/TSMCC.2004.829274

[20] Cucchiara, Rita, et al. "Detecting moving objects, ghosts, and shadows in video streams." *IEEE transactions on pattern analysis and machine intelligence* 25.10 (2003): 1337-1342.

[21] Zhang, J., Cao, J. & Mao, B., 2016. Moving Object Detection Based on Non-parametric Methods and Frame Difference for Traceability Video Analysis. Procedia Comput. Sci. 91, 995–1000. doi: 10.1016/j.procs.2016.07.132

[22] Kothiya, Shraddha & Mistree, Kinjal. (2015). A review on real time object tracking in video sequences. 1-4. 10.1109/EESCO.2015.7253705.

[23] J. Fan, W. Xu, Y. Wu, Y. Gong, Human tracking using convolutional neuralnetworks, IEEE Trans. Neural Netw. 21 (10) (2010) 1610–1623.

[24] S. Krebs, B. Duraisamy and F. Flohr, "A survey on leveraging deep neural networks for object tracking," 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), 2017, pp. 411-418.

[25] J.M.O.L. Lobato, R.M. Luque, D.L. Rodríguez, E.J. Palomo, Growing Competitive Network for Tracking Objects in Video Sequences,Springer-Verlag, Berlin, Heidelberg, 2009, pp. 109–118.

[26] A. Keivani, J. Tapamo and F. Ghayoor, "Motion-based moving object detection and tracking using automatic K-means," 2017 IEEE AFRICON, 2017, pp. 32-37,

[27] Weng, E.N.G., Khan, R.U., Adruce, S.A.Z., & Bee, O.Y., 2013. Objects Tracking from Natural Features in Mobile Augmented Reality. Procedia - Soc. Behav. Sci. 97, 753–760. doi: 10.1016/j.sbspro.2013.10.297

[28] Sarkar, R., Bakshi, S. & Sa, P.K., 2012. A Real-time Model for Multiple Human Faces Tracking from Low-resolution Surveillance Videos. Procedia Technol. 6, 1004–1010. doi: http://dx.doi.org/10.1016/j.protcy.2012.10.122 .

[29] Zhang, J., Cao, J. & Mao, B., 2016. Moving Object Detection Based on Nonparametric Methods and Frame Difference for Traceability Video Analysis. Procedia Comput. Sci. 91, 995–1000. doi: 10.1016/j.procs.2016.07.132

[30] Zhang, Tianya and Peter J. Jin. "Roadside Lidar Vehicle Detection and Tracking Using Range and Intensity Background Subtraction." (2022).

[31] Akansha Bathija, Prof. Grishma Sharma, 2019, Visual Object Detection and Tracking using YOLO and SORT, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 08, Issue 11 (November 2019)

[32] Pal, S.K., Pramanik, A., Maiti, J. et al. Deep learning in multi-object detection and tracking: state of the art. Appl Intell 51, 6400–6429 (2021). https://doi.org/10.1007/s10489-021-02293-7