International Journal of Scientific Research in Science and Technology
Available online at : www.ijsrst.com
Print ISSN: 2395-6011 | Online ISSN: 2395-602X
doi : https://doi.org/10.32628/IJSRST

A Survey of Available Techniques for Identification of Objects on the Road and Calculating Distance from DashCam

Shreya Jaiswal[1], Sonam Singh[2]
Research Scholar[1], Assistant Professor[2]
SHEAT College of Engineering, Varanasi, Uttar Pradesh, India

ARTICLEINFO

Article History:
Accepted: 01 Jan 2024
Published: 12 Jan 2024

Publication Issue :
Volume 11, Issue 1
January-February-2024
Page Number :
636-644

ABSTRACT

The rapid advancement of AI is driving continuous innovation in the area of computer vision. Object detection and object tracking are innovative technologies that enhance safety and improve company efficiency. Object identification and tracking play a crucial role in preventing hazardous situations on the roadways, recognizing traffic signals and cars, and performing several other tasks. There are other methods to implement object recognition and tracking, which we will discuss below. This article provides a comprehensive analysis of the distinctions between object detection and object tracking.

Keywords – Object Detection, Object Identification and Tracking, Dash Cam.

I. INTRODUCTION

A. Detection of Objects

Object detection is a computer vision technology used in software engineering to accurately recognize and locate items inside an image or video. Object detection precisely identifies and delineates the boundaries of observed items, accurately determining their location or movement inside a given scene.

B. Object Tracking.

Object tracking is a technique used to determine the precise position of an item either in recorded film or in real-time. item tracking algorithms monitor the movement of an item and generate precise data on its trajectory. The dataset preparation for object-tracking algorithms involves the task of picture labeling, where engineers annotate and categorize the objects. The algorithm's efficacy may be assessed by its ability to accurately assign item IDs.

C. Distinction between Object Detection and Object Tracking

Object tracking is the process of identifying and monitoring objects as they move through a sequence of frames in a video or film stream. Object detection is a component of the object tracking procedure, notably occurring in the beginning, when a neural network locates an item in a video or picture and recognizes it as the intended target.

Object detection and object tracking are used for the analysis of visual data in order to ascertain the precise positions of things. However, it is important to note that there exist notable distinctions between these two techniques. Object detection is the process of identifying specific objects inside an image or frame, whereas object tracking involves following the movement of a target item over consecutive frames. Object detection algorithms usually analyze each picture or frame separately, while object tracking techniques predict the position of the target in upcoming frames.

II. Three Types of Object Detection

To develop a precise object-detection system, you may use the two prevailing approaches: machine learning and deep learning. In this section of the essay, we will provide comprehensive information on those topics. Please continue reading for more details.

A. Object detection using machine learning algorithms

Data science engineers often choose for machine learning (ML) models to enhance the precision of object recognition. The primary concept of the ML technique is training a model using annotated pictures that include instances of the desired items. Once the model has been trained, it may be used to analyze fresh photos and identify items within them.

One machine learning approach is aggregate channel features (ACF), which allows for the identification of target objects using a training dataset. The deformable parts model (DPM) is a frequently used machine learning approach that identifies things by examining the connectivity of its pieces.

B. Artificial Neural Networks

Deep learning training approaches enhance the accuracy of algorithmic outputs for object identification. You have the option to develop a personalized algorithm and manually train it, or you may choose to use a pre-existing neural network. To perform object identification using deep learning, it is necessary to train convolutional neural networks (CNNs). They provide expedited and more accurate outcomes. It is important to consider that a high-performance graphics processing unit (GPU) and a much bigger training dataset are need for this task, compared to the standard requirements for mathematical modeling.

C. Object Tracking Categories

The use of object tracking is expanding across several domains, prompting engineers to develop more sophisticated methods for training neural networks to monitor specific objects in diverse scenarios. In the following discussion, we will explore the prevailing forms of visual object tracking, namely image and video object tracking.

1. Visual Tracking: Image tracking is used to precisely locate target items among a multitude of other objects or identify necessary alterations in the target objects. By using this method, users may assign the task of monitoring changes made to a certain item over several photos, resulting in important results. Image tracking, for instance, aids medical experts in the examination of tumor progression.

2. Video Tracking: In order to track objects in films, a neural network is required to progressively analyze a series of pictures. In order to monitor the movement of an item, engineers must train a neural network to forecast the general direction in which the object will travel. The system utilizes video frames to estimate the historical and present position of an item, and continues to follow it as long as it is visible.

3. Levels of Object Detection: Object detection, similar to object tracking, may be used throughout diverse domains, necessitating varying degrees of proficiency that are contingent upon the intricacy of the duties at hand. These are the three tiers of object detection that may be necessary for your solution.

4. Object Localization using Bounding Box Detection: This level focuses on the task of object detection, which involves identifying things and determining their precise positions by surrounding them with rectangular bounding boxes. Bounding box detection provides information on the spatial coordinates of an item, but does not provide any information about the specific shape or structure of the object.

5. Semantic Segmentation: Semantic segmentation is the process of assigning a semantic label to every individual pixel in a picture. At this stage, the object identification algorithm categorizes every pixel into distinct object classes, resulting in more accurate object positioning and revealing details about the shape and structure of the target item.

6. Instance Segmentation: Instance segmentation is a method that distinguishes between separate instances of objects belonging to the same class. The algorithm adds semantic labels to individual pixels and provides a unique identifier to each instance of an item. Instance segmentation is a method that operates on objects that resemble each other and allocates individual instances of each item from various classes.

D. Stages of Object Tracking

Our data science specialists will use several training methodologies for neural networks based on the specific job requirements and the quantity of objects to be tracked. There are two degrees of object tracking: single object tracking (SOT) and multiple object tracking (MOT), which will be discussed below.

1. Single Object Tracking (SOT): SOT, or Single item Tracking, is a technique that specifically aims to monitor and follow the movement of a solitary item inside a video recording or live stream. SOT methods often depend on the initial item placement, which may be determined either manually or automatically. Subsequently, the algorithm makes an estimation of the object's motion and extracts its distinctive characteristics in order to monitor its movement during a period of time. SOT algorithms may be used in several sectors, such as augmented reality (AR), surveillance, and autonomous vehicle management.

2. Object Tracking (MOT): Motion Object Tracking (MOT) algorithms are capable of concurrently tracking several objects in video or real-time data streams. Developing MOT algorithms is a complex task because to the need of accurately tracking objects with varying sizes and looks, while they move along diverse trajectories. Data science engineers use data association, trajectory prediction, and motion modeling approaches for MOT. MOT, or Multiple Object Tracking, is extensively used for the study of human activities, monitoring traffic, and managing crowds.

E. Intelligent Parking Management

In order to relieve the drivers from the task of finding an open parking place, parking lot owners have used object-detection algorithms to monitor the number of unoccupied parking spaces. In addition, intelligent parking systems are now included into almost all contemporary vehicles, aiding drivers in parking securely by providing them with information on impediments.

F. Automated Payment at Grocery Stores

Implementing object identification and tracking technology may effectively decrease waiting times in grocery shops. Customers are not required to make payments using cash or cards. Within these establishments, a patron has a mobile application with unique identification that allows them the ability to effortlessly choose the desired items and go from the premises without any more interaction. A computer vision system will analyze the goods that the consumer selects and process the payment using an application-based payment system.

G. Analysis of Human Behavior

Object identification and tracking techniques contribute to the accomplishment of improved safety. AI-powered surveillance cameras are capable of monitoring and documenting many forms of criminal activity. If necessary, a police department may use the recorded video footage to identify and apprehend an individual who has committed a legal infraction.

H. Management of the Warehouse

By using object identification and object-tracking algorithms, firms may easily monitor the movement of packages throughout a warehouse. AI technologies enable the identification of the package's status and storage circumstances, as well as the scanning of label information, functioning as an object tracker. By implementing an object detection or tracking system, the likelihood of losing a product is greatly reduced since the system is able to identify and monitor each package.

I. Biometrics and Face Recognition Technologies

Airport security systems use object detection and tracking algorithms. A computer vision system is used to authenticate the identity of passengers by comparing a passenger's facial features with the photograph in their passport, thereby preventing fraudulent activities. The technology further verifies that a passenger's conduct is not deemed suspicious.

J. Artificial Reality

Object detection is now a popular method used in the field of augmented reality (AR). For example, individuals have the option to download an application that allows them to virtually try on clothing. The app's technology is designed to identify individuals in images and, upon detection, overlay a clothing over the subject. AR applications for the smart home and furniture sectors are another example of object detection. Users have the ability to capture an image of their apartment or house and assess the visual appearance of a particular smart home gadget or piece of furniture inside their environment.

III. Implementing Object Detection And Tracking

The science of computer vision has extensively studied the process of recovering scene geometry from a single view [1, 3, 17]. Leibotwzki used geometric properties, such as parallelism and orthogonality, to recreate three-dimensional models from a single or dual viewpoint, even in the absence of explicit information about the metric attributes of the objects [2]. Lee et al. have shown that it is feasible to generate a comprehensive three-dimensional (3D) representation from a solitary image, even in the presence of visual complexity. This is accomplished by using geometric reasoning to analyze the line segments that are produced from planar surfaces [3]. Zaheer et al. conducted a recent work where they used orthogonal lines in a multi-planar single view to successfully accomplish a comprehensive 3D reconstruction of the scene [4]. These solutions are particularly fascinating for indoor artificial environments or for outdoor landscapes in close proximity to man-made structures. The reason behind this is because human architecture is defined by many parallel and perpendicular lines that stem from flat surfaces.

A. Selecting Best Appropriate ML Model

Our data science engineers evaluate the existing machine learning models that have previously undergone training to avoid the need for re-training a model and so save time. If there are no machine learning models that meet the client's needs, we will train a new model from the beginning using the dataset that has been gathered. This is the most often chosen option.

B. Engaging in Several Training Iterations

Our specialists instruct the neural network, evaluate the outcomes, get feedback on its precision, incorporate necessary modifications, and begin the subsequent training cycle. Neural networks often exhibit low accuracy initially, requiring a considerable amount of effort to get satisfactory results.

C. Field Testing

When a project necessitates field testing, our data scientists will personally visit the customer. In order to evaluate the efficacy of our weight-measuring algorithm, we conducted a field study on a farm to observe its performance under real-world situations.

D. Incorporating a ML Model into a Specific Device

For any data science project that necessitates object detection, we adhere to the following processes. The process of developing an object-tracking algorithm follows a similar sequence of phases. The primary distinction is in the methodology used for data gathering, procedures, and the development of algorithms.

IV. Challenges and Difficulties in Core Object Detection and Tracking with Solutions

Although there have been notable improvements in object identification and tracking technology, there are still several hurdles and limitations that remain. Engineers often encounter many significant obstacles when it comes to object identification and tracking. These issues may include:

A. Components and Details Related to Video Streaming

In order to accurately identify a target item and its properties, an algorithm must be able to clearly distinguish it while monitoring its movement. As the pace increases, the object becomes more fuzzy. It is important to understand that a camera does not immediately capture a frame with an object. Instead, it collects light on a sensor to create a frame. When the frame time is sufficiently long, the resulting picture will have enough lighting, but the object may seem blurred. A shorter frame time will result in less blurring of the target object, but it will also cause the final picture to seem dimmer. Data science engineers aim to achieve a balance or compromise. As an alternative, our team utilizes picture post-processing to identify and evaluate blurred frames, determining their suitability for further study.

As an instance, in our pig weighing object-detection method, we examined the degree of blurriness and instructed the neural network to recognize the image that is sufficiently clear to ensure the correctness of the outcomes. Furthermore, the speed of detection is greatly influenced by the specific hardware components that you choose. We can assist you in choosing the appropriate components for your embedded artificial intelligence system.

B. Varied Object Dimensions

Objects of interest might exhibit considerable variation in their dimensions, hence posing a challenge for the task of detecting and monitoring them. Furthermore, the system must precisely identify and monitor the objects, even as they approach or move away from the camera, resulting in variations in their actual dimensions.

C. Occlusion

Occlusion occurs when a target item is partly or totally hidden by other objects or the backdrop. At such point, the object identification or tracking algorithm may fail to identify a target item, leading to an imprecise output. In order to address occlusion, engineers must develop a sophisticated algorithm capable of precisely detecting or tracking objects, even in cases when they are partly obscured.

During the development of our pig weighing algorithm, we encountered the issue of occlusion. This difficulty arose when our solution had difficulties in accurately catching a pig that was in close proximity to other pigs. The breed of the pigs also had an impact on the accuracy of the detection. Therefore, we trained our neural network using the same pig breed since they did not exhibit significant variations. Consequently, our system successfully overcomes occlusion and delivers accurate outcomes.

D. Contextual Distractions

The things that need to be detected and tracked are often found in intricate surroundings with crowded backgrounds. Excessive visual elements in the background might introduce superfluous data or interference that may cause confusion for the object recognition or tracking system, resulting in imprecise outcomes. To address the issue of background clutter, it is necessary to use efficient methods for extracting relevant features and reducing noise. This is crucial in order to achieve precise identification and tracking of objects.

We suggest using neural network training to identify and monitor certain items amongst varying backdrops. The neural network is trained by our data science specialists utilizing diverse photos including distinct backgrounds in order to enhance the algorithm's precision.

In order to enhance the training of your neural network, it is advisable to gather a substantial dataset and capture images of target items against various backdrops and lighting conditions. This will augment the variety of the dataset. To achieve accurate detection and tracking of moving objects, it is crucial to take into account several parameters. The algorithm should possess the capability to operate well throughout a wide spectrum of conditions.

V. Technology Stack for Object Detection & Tracking

Pre-existing object identification and tracking algorithms are crucial for data science engineers specializing in computer vision since they have a broad variety of applications, including surveillance, autonomous driving, and robots. DeepSort, Optical Flow, and YOLO are the most widely used algorithms for object recognition and tracking. Continue reading to get more insight into these algorithms and their use in computer vision tasks.

A. DeepSort

DeepSort is a real-time item recognition and tracking method that use deep learning to track things. The DeepSort method utilizes a basic online and real-time tracking (SORT) algorithm as its foundation. It employs a deep neural network to recognize objects, determine their location, velocity, and size, monitor their movements, and precisely predict their future paths.

B. Optical Flow

OpenCV, a computer vision library, offers the Optical flow method for object recognition and tracking. This algorithm utilizes computer vision techniques to predict the movement of target objects in a given scene. Optical flow examines changes in pixel intensity inside a video frame and computes the motion's direction and magnitude. Optical flow is beneficial for capturing the dynamic motion in video recordings or streams, particularly when the target objects exhibit irregular movement.

C. YOLO (You Only Look Once)

YOLO is a computer system that using deep learning to accurately recognize and track objects inside images or video streams. The YOLO algorithm analyzes the input picture by dividing it into individual pixels and then calculates the likelihood of an item being present in each pixel. YOLO is often used as a fundamental basis for developing object-tracking algorithms.

Once the algorithm has been trained, it may be used in practical situations. Our algorithm for weighing pigs used DeepSort and Optical flow as its foundational components. Nevertheless, the image processing pipeline comprises several bespoke algorithms due to our client's need for a distinctive data science solution.

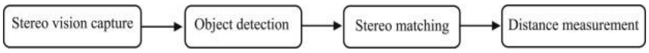

Fig. 1: Displays the comprehensive flow diagram of the suggested approach

	Ground Truth	Our Algo	Monodepth	Dorn	Radar
1	5.02	5.77	5.29	8.54	6.00
2	26.57	26.08	13.30	18.90	28.00
3	21.00	18.68	13.47	18.03	21.00

Fig. 2. Sample image from nuScenes [3].

The estimated horizon (shown by the green line) and calculated distances (measured in meters) for each car (1, 2, and 3) from the ego vehicle are shown. We compare our technique to depth estimate based on deep learning and radar. GT refers to the ground truth data, which is obtained from the high-resolution lidar technology. Warning systems have been developed to notify drivers before to dangerous moves.

Timely recording of this data is essential in this context and may assist fleet management in evaluating a driver's conduct. It is also a crucial component of autonomous driving, since it is used in determining the best course to take, avoiding obstacles and collisions, and assessing the overall safety of self-driving automobiles.

Radar and Lidar are often used for active sensing to measure distances to nearby barriers. Radars typically have a maximum range of 150 meters [5], however their resolution is frequently poor. However, lidars provide superior spatial resolution [6], but they are prohibitively costly (the least priced option costs US$ 7,500 [1]). An inherent benefit of using active sensing is its ability to operate in real-time, hence minimizing the need for extensive post-processing. Scene depth may be determined rapidly, and a simple thresholding technique may be sufficient for this task. Although the cost of lidars is anticipated to decrease by 2022 because to advancements in sensor technology and mass manufacturing [7], they still suffer from performance issues in unfavorable weather situations including rain, snow, and fog. Standard dashcams, on the other hand, provide affordable and high-resolution RGB photos of the scene. However, they need clever post-processing employing strong algorithms and more computing power on board.

VI. LITERATURE REVIEW

A. Overview

Within the realm of literature, there are only a limited number of works that have used methodologies comparable to ours. Park and Hwang [8] used monocular imagery and the average width of vehicles in the actual world to estimate the horizon. Additionally, the vertical distance between the bottom of the vehicle's detected bounding box and the horizon offered a clue about its approximate distance in relation to the observer. However, they made the assumption of a pre-established camera height, which might provide challenges when applied to several vehicles within a fleet. The winning response to TuSimple's velocity estimate competition in 2017 included advanced deep learning algorithms for predicting depth and detecting objects (namely vehicles). This combination of approaches allowed for the determination of relative velocities. [9] In light of this, we need to examine the research conducted in two areas: our suggestion of single view geometry for estimating depth, and the current cutting-edge deep learning approach for predicting depth.

B. Lateral Alignment

The lateral position is determined as the orthogonal distance between the left front wheel and the road marker on its side, as seen in Figure. The road marker serves as a point of reference due to its status as the only permanent element on the road that has clearly defined boundaries. In order to ensure uniformity in data, it is essential that this distance be consistently measured from a predetermined location. The lateral position data is considered positive while the vehicle is inside the lane boundaries. However, when the front left wheel crosses the center road marker, the position is recorded as negative. This situation may occur while a vehicle is navigating a curve, choosing a more direct route, or passing another vehicle.

C. Time-to-Line Crossing (TLC) and Additional Software Programs or Computer Applications

Many driver behavior research studies are interested in continuously measuring the lateral position of a vehicle. This measurement allows for the analysis of irregular lane positioning behavior and the location of the vehicle inside the lane. The vehicle's alignment with the road markings is crucial in several safe driving applications within the field of Intelligent Transport Systems (ITS). Driver assistance systems use lane position information to aid drivers in maintaining their place inside a lane. Examples of such systems include lane maintaining, steering control systems, and driver status monitoring, such as sleepy driving systems. AssistWare Technology has created a driver condition monitoring system called SafeTRAC [9] for commercial use. A driver condition monitoring system was created as part of the SAVE initiative for research purposes.

D. Optical Perception-Oriented Methodology

Throughout the years, many methods for measuring lateral position have been created, which may be categorized into infrastructure-based and vehicle-based strategies. An example of a common infrastructure-based method is the use of magnetic markers that are implanted in the lane. These markers work in conjunction with sensors on the vehicle to accurately measure the distance between the vehicle and the marker in the lane. This technique has shown exceptional precision, and the electromagnetic system has the benefit of being impervious to weather conditions. As a result, several snowplough guiding systems use this kind of technology in situations where visibility is often poor [10].

Here, h represents the intended distance between the item and the cameras. In order to determine the distance h, the following parameters are required:

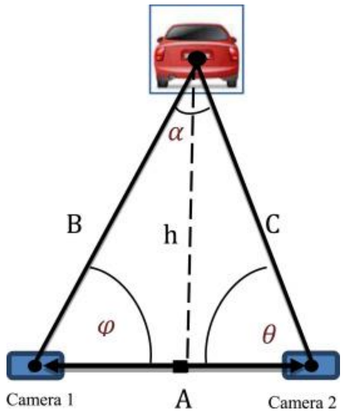

Fig. 3: Two Cameras Mounted as a Stereo Camera

VII. CONCLUSIONS AND FUTURE WORK

This study introduces a method for estimating the relative distance of road vehicles based on single view geometry. We have conducted a comparative analysis between our algorithm and data-driven deep learning (DL) techniques. Supervised deep learning achieved outstanding performance when applied to KITTI data. While data-driven initiatives are indeed the future, it is important to acknowledge that they also have their limitations. Instead, we suggested a geometrically-based alternative that is straightforward, intuitive, and not reliant on data. This alternative performs nearly as well on all the datasets that were examined, with the root mean square error (RMSE) consistently falling within a tight range of 6.10 to 7.31. The initialization phase of our approach utilizes the ADAS module for lane recognition, which is computationally intensive due to its reliance on deep learning methods. However, initialization is a single computation that occurs just once, and after that, our distance estimate operates in real-time. In this study, we have shown that both geometric and deep learning frameworks are viable alternatives to replace active sensors in ADAS systems and self-driving automobiles. This substitution would significantly contribute to cost reduction in future intelligent transport systems.

VIII. REFERENCES

[1]. V. Kesten, R. and Usman, M. and Houston, J. and Pandya, T. and Nadhamuni, K. and Ferreira, A. and Yuan, M. and Low, B. and Jain, A. and Ondruska, P. and Omari, S. and Shah, S. and Kulkarni, A. and Kazakova, A. and Tao, C. and Platinsky, L. and Jiang, W. and. Lyft Level 5 Dataset. https://level5.lyft.com/dataset/, 2019.

[2]. M. Khader and S. Cherian. An Introduction to Automotive LIDAR (Texas Instruments). Technical report, 2018.

[3]. H. Khan, A. Rafaqat, A. Hassan, A. Ali, W. Kazmi, and A. Zaheer. Lane detection using lane boundary marker network under road geometry constraints. In IEEE Winter Conference on Applications of Computer Vision, Snowmass, Co). IEEE, 2020.

[4]. I. Slimani, A. Zaarane, A. Hamdoun, Atouf I. April, Vehicle License Plate Localization and Recognition System for Intelligent Transportation Applications, IEEE (2019), pp. 1592-1597.

[5]. W. Al, A. Zaarane, I. Slimani, I. Atouf, M. Benrabh, Vehicular queue length measurement based on edge detection and vehicle feature extraction, J Theor Appl Inf Technol, 97 (5) (2019).

[6]. Martinez, F., Jacinto, E., Martinez, F.: Obstacle detection for autonomous systems using stereoscopic images and bacterial behaviour. Int. J. Energy Clean Environ. 10(2), 2164–2172 (2020).

[7]. Hu, Y., et al.: A new microscopic telecentric stereo vision system - calibration, rectification, and three-dimensional reconstruction. Opt Laser. Eng. 113, 14–22 (2019).

[8]. Tian, X., et al.: High quality 3D reconstruction based on fusion of polarisation imaging and binocular stereo vision. Inf. Fusion. 77, 19–28 (2022).

[9]. Wu, X., et al.: Simultaneous localisation and mapping of medical burn areas based on binocular vision and capsule networks. Soft Comput. 24(23), 18155–18171 (2020).

[10]. Romero, C., et al.: A validation strategy for a target-based vision tracking system with an industrial robot. J. Phys. Conf. Ser. 1547, 012–018 (2020).

[11]. Huang, L., et al.: Obstacle distance measurement based on binocular vision for high-voltage transmission lines using a cable inspection robot. Sci. Prog. 103(3), 1–35 (2020).

[12]. Munoz, D.F.T., Prieto, F., Correa, A.C.: Power lines detection from a stereo vision system. In: 2020 5th International Conference on Control and Robotics Engineering (ICCRE), pp. 242–245. Osaka (2020).

[13]. Sun, X., et al.: Distance measurement system based on binocular stereo vision. IOP Conf. Ser. Earth Environ. Sci. 252(5), 052051 (2019).

[14]. Emani, S., et al.: Obstacle detection and distance estimation for autonomous electric vehicle using stereo vision and DNN. In: J. Wang, et al. (eds), Soft Computing and Signal Processing, vol. 898, pp. 639–648. Springer Singapore, Singapore (2019).