Computation of TTC using Lidar and Monocular Camera Data for Automobiles

Ramakanthkumar P[1*], Vijay Raj[1], Rahul J[1], Pavithra H[1]

[1]Department of Computer Science & Engineering, RV College of Engineering, Bengaluru, Karnataka

Abstract

Automobile safety systems are ever evolving technologies that rely on multiple domains to achieve their goals. One of the most used safety systemsin commercial vehicles is collision avoidance system, which works on Time-to-Collision (TTC) estimation. OpenCV is the core library used to develop this system, the image onto YOLOV3 pre-trained neural net to detect and recognize object in the frame and tag them with bounding boxes. Data fusion was done on camera and LIDAR data to obtain LIDAR TTC of the vehicle on the ego lane and cluster the point cloud on bounding box. In addition, as for computation camera TTC, key point descriptors using Akaze and FLANN were adopted to match them between current and previous frame and hence to obtain ratios of their distances. TTC was computed based on both Lidar and monocular camera in real time and thus increasing the reliance on TTC estimation. The application was tested on multiple KITTI dataset and TTC estimation was done by using different sensors. Both the sensors were found to be reliable in different scenario, while LIDAR emerged to be better performer overall.

Keywords: *Time To collision, TTC, Data Fusion, safety critical systems.*

1.0 Introduction

One of the most crucial safety system functionality is time to collision estimation, which is used for Collision Avoidance System (CAS), assisted braking, lane changing assistance system, etc. Time-To-Collision (TTC) has proven to be a valuable method for determining the importance of crucial and normal behaviour and gauging the severity of traffic accidents [1]. Many vehicles are not equipped with fancy sensors that are present in autonomous vehicles; most vehicles are equipped with monocular camera like dash camera in the front [2]. Camera is one of the

*****Mail address: Ramakanthkumar P, Professor and Head, Department of Computer Science & Engineering, RV College of Engineering. Bengaluru – 560 059.*
Email: ramakanthkp@rvce.edu.in.Ph: 9886309520

most useful sensors which provide RGB image that can be used to detect and identify the object using YOLOV3 algorithm and estimate TTC using Open-CV key point descriptors(KPD) between two consecutive frames. However, camera alone cannot reliably estimate TTC, since it is limited to short distance and hence NVIDIA Jetson Nano 2 GB which is a powerful edge-computing device is used. This device is a powerful computer packed in a small for AI, IOT and embedded applications. It has the performance and capability to run workload in a fast and easy way. With the help of powerful tools and efficient detecting algorithms, a stand-alone working model is developed.

For obtaining a reliable TTC a second sensor is needed. Lidar as a second sensor, which projects point clouds to the scene to obtain depth measurement from a point is used [3]. Lidar can track TTC estimation of vehicle in a lane for longer distance than camera. Hence camera based calculated TTC estimation as a benchmark to Lidar based TTC estimation was adopted.

1.1 Computer Vision and Autonomous Vehicles

Computer vision focuses on recreating some of the complexity of the human visual system so that computers can recognize and analyze items in images and videos [4].

Computer vision researchers sought to create algorithms for such visual perception tasks such as:

(i) Recognize certain object in the image: object recognition,

(ii) Object detection to find instances of a particular class of semantic objects, and

(iii) Scene understanding to divide an image into useful chunks for analysis [5]. Computer vision is all about pattern recognition. Consequently, a large number of tagged pictures are sent to a computer and then subjected to different software approaches or algorithms that enable the computer to look for patterns in all the parts that relate to those labels as a way of teaching it to interpret visual data.

The foundation of autonomous vehicle technology is computer vision. In order to safely navigate the road, cars use object detection and identification algorithms with combination of sophisticated cameras and sensors to evaluate their surroundings in real time and identify objects like people, traffic signs, barriers, and other vehicles [6]. As per the SAE Levels of Automation scale, there are five levels of automation for autonomous cars [7]:

The foundation of autonomous vehicle technology is computer vision. In order to safely navigate the road, cars use object detection and identification algorithms with combination of sophisticated cameras and sensors to evaluate their surroundings in real time and identify objects like people, traffic signs, barriers, and other vehicles [6]. As per the SAE Levels of Automation scale, there are five levels of automation for autonomous cars [7]:

Level 0: Automated alerts and momentary assistance, such as lane departure warnings and emergency braking, are at Level 0.

Level 1: The car has a single autonomous driving aid system that can steer or accelerate (cruise control).

Level 2: This refers to ADAS, or advanced driver assistance systems. The car has steering and acceleration/deceleration controls. Because a human is seated in the driver's seat and has the ability to take over the vehicle at any time, this automation falls short of self-driving in this instance. Level two systems include Cadillac (General Motors) Super Cruise and Tesla Autopilot.

Levels 3 and 4: Under particular circumstances, these features allow the vehicles to function on their own. Level 3 features can call for the driver to take over steering.

Level 5: These characteristics are identical to Level 4 features, with the exception that they can operate in any kind of road situation.

1.2 Open CV

Open CV is a noteworthy open library for computer vision, machine learning, and image processing. Presently, it contributes significantly to real-time operation, which is essential in contemporary systems. It may be used to search for people, objects, and perhaps even human handwriting in images and videos. With the help of integrated libraries like NumPy, Python and other high level programming languages the OpenCV array structure can be used for analysis [8]. To recognize visual patterns and their numerous aspects, vector space is employed and mathematical operations on these features are performed.

1.3 YOLO V3

You only look once (YOLO) is a state-of-the-art, real-time object detection system. On a Pascal Titan X, it processes pictures at a speed of 30 frames per second, and on COCO test-dev, it has a mAP of 57.9. YOLOv3 is extremely fast and accurate (Fig 1 and 2). YOLOv3 is nearly four times quicker than Focal Loss in mAP measured at.5 IOU, but they are equivalent [9]. Furthermore, a balance between accuracy and speed can be achieved by simply changing the size of the model; no retraining is required.

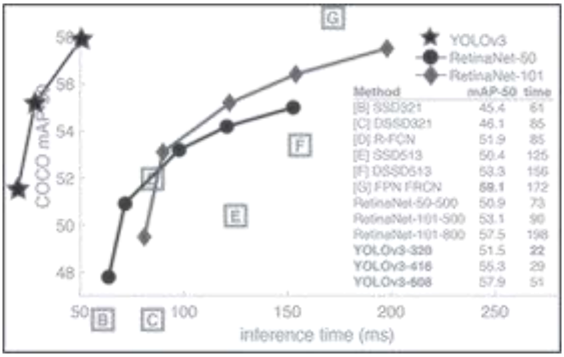

Fig.1. YOLO V3 performance on COCO dataset

1.4 Sensor fusion

Sensor fusion is the process of combining data from various sources, such as sensors, to produce information that is less uncertain than it would be if the sources were used separately [10]. The term "uncertainty reduction" in this context might mean either being more accurate, thorough, or dependable, or it can mean the outcome of a fresh viewpoint, like stereoscopic vision (calculation of depth information by combining two-dimensional images from two cameras at slightly different viewpoints) [11].

1) Lidar-Camera data fusion

For object detection, LiDAR and camera fusion techniques with varying levels of data fusion have been introduced. The fusion-technique establishes a connection between the point clouds from LiDAR and the object identified by a camera to expedite processing. This is possible because sensor fusion increases robustness and detection accuracy while making up for the shortcomings of the individual sensors [12].

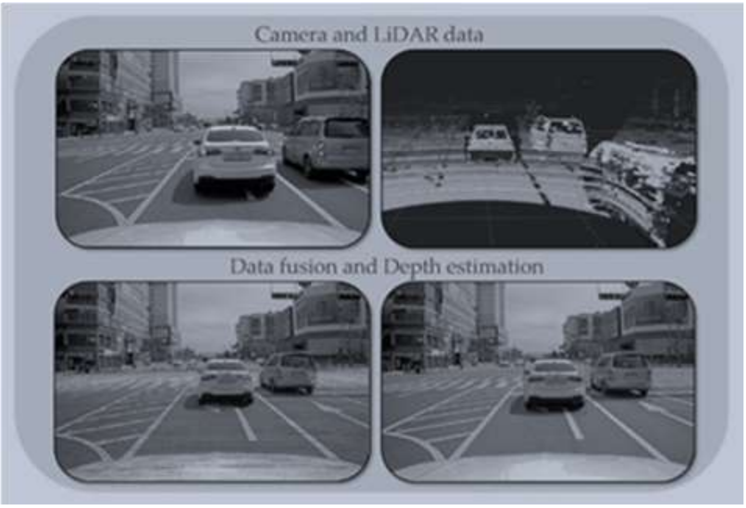

Fig. 2. YOLO V3 performance on COCO dataset

Fusing these two data can result in reliable TTC estimation [13]. For instance, a typical camera has a much higher resolution than LiDAR, but it has a smaller field of view and is less accurate at estimating distances to objects than LiDAR. When utilizing merely images, a camera is furthermore sensitive to changes in illumination and has advanced image processing, but LiDAR finds it difficult to classify objects and discriminate between colours in contrast to a camera. Data fusion, which combines sensor data from various sources, can be used to overcome the constraints of individual sensors and lower their ambiguity.

1.5 AKAZE and FLANN

AKAZE algorithm is the accelerated version of KAZE. As opposed to KAZE, the accelerated KAZE algorithm creates the non-linear scale-space using a quicker technique known astheFastExplicitDiffusion(FED) [14].Tofindandcompare key points between two images, AKAZE local features are used. The key points on a pair of images with the given homography matrix were matched to find the inliers, and count them (i.e. matches that fit in the given homography).

FLANN (Fig 3) was used to quickly find approximate neighbours in high-dimensional spaces. It includes a set of algorithms that were discovered to be the most effective for nearest neighbour searches as well as a system for automatically selecting the most effective algorithm and ideal parameters based on the dataset. Key point descriptors were compared and matched using the Euclidean distance utilizing traditional feature descriptors (SIFT, SURF, etc.) [15]. Histogram-based metrics can be used as alternatives to the Euclidean distance since SIFT and SURF descriptors provide the histogram of an oriented gradient in a neighbourhood.

Fig. 3. Flann example

1.6 Safety critical system in automobile

In automobile industry safety critical systems are required to follow many sets of international safety standards [16], such as estimating TTC for developing further safety systems like Intelligent Driver Assistant System (IDAS) or developing a diagnostic system for detecting malfunctioning component of the vehicle [17], in turn avoiding recurring malfunctions and improving traffic safety. There are two waves of safety systems, namely, First wave of safety systems and Second wave of safety systems.

First wave of safety systems: Modern passenger automobiles and commercial vehicles are generally equipped with the first wave of active safety systems. a) Antilock Braking System (ABS):One of the oldest safety features used in an automobile. It is essentially an active protection mechanism installed in the majority of autos that prevent the wheels from locking up while braking hard [18], b) Electronic stability control (ESC):

ESC was introduced in 1998. It aids avoiding skidding and the driver losing control of the vehicle while turning a corner. The brakes could be automatically applied using ESC technology to assist in steering the vehicle in the appropriate direction. A study was conducted to see the effect of ESC, and ESC was found to have a greater impact on single-vehicle crashes than multiple-vehicle crashes and crashes with fatal injuries than less severe crashes. [19].

Second wave of active safety systems: It was introduced using more advanced technologies such as stereo cameras, radar, GPS and Lidar [20]. It includes:

a) Lane keeping assistance (LKA):To reduce the effort of the driver without lowering the driving motivation, a novel idea of cooperative driving between the driver and the assistance system is developed. The system has received approval from the Japanese Ministry of Land, Infrastructure, and Transport to be deployed on expressways in Japan as a basic driver assistance system [21].

b) Intelligent speed assistance (ISA):ISA systems actively deter drivers from going over the posted speed limit using GPS-linked speed limit databases and road sign recognition cameras. Although this safety feature is still in trials due to its technical issue, it can be one of the most game changing safety features to save life.

1.7 Time-To-Collision

TTC is proven as a traffic conflict technique as an indicator for gauging the seriousness of traffic accidents and differentiating between problematic and typical behavior. The research findings support the direct use of TTC as a cue for traffic decision-making [1]. It can be implemented using only a camera or even with advanced sensors such as LIDAR and radar. TTC was further used to implement CAS,IDAS, lane changing assistance andABS. TTC and CAS are not only used in land vehicles but also is an application for unmanned aerial vehicles such as drones [22].

TTC with Camera:The ratio of the body's image size to its derivative time can be used to calculate TTC for a body moving in relation to the camera. The main objective is to compute this ratio using local scale change and motion data gathered from the identification and monitoring the feature points [23]. TTC from camera has its own limitation of less visibility like fog and it cannot calculate TTC beyond a short distance [24].

TTC with Lidar:Computing TTC with Lidar requires fusing of LIDAR and camera sensor data using object detection and a lean implementation approach, a 3D multi-target tracking method with a real-time. LIDAR point clouds were mapped on to the image or frame from the camera in real time, and distance was measured from ego lane vehicle, which is in front of the car to estimate TTC. Fused Lidar based TTC is proved to be more effective than camera TTC alone [25]. TTC can be estimated to a further distance that a camera cannot when there is less visibility.

1.8 Sensor Data Fusion

Data fusion is an emerging domain, which works on fusion of two different types of data, which may be unreliable into one type of reliable data. Data fusion is usually implemented using algorithms like use of Kalman filters, neural nets, Bayesian networks, KNN algorithm and fuzzy logic [21]. For navigation and control of an autonomous vehicle, multi-sensor data fusion was employed. Advanced sensors were employed to collect information from structured environments, including GPS, a digital compass, a laser scanner, and an ultrasonic sensor [22]. This research deals with combining camera and LIDAR data with different data generation. Therefore, an Asynchronous Data Fusion Algorithm was employed by incorporating a weight vector to the data source to increase the accuracy of the model[23]. This fused data was usedtocalculateLIDARbasedTTC.

1.9 KITTI dataset

The KITTI dataset was collected from a moving platform while travelling through and around Karlsruhe, Germany. It is made up of laser scans, digital images, extremely accurate GPS data, and IMU accelerations from a GPS/IMU system operating together. The main goal of this collection was to enhance robotic and computer vision systems for autonomous cars[26].

2.0 Methodology

Two main sensors namely, monocular camera and Lidar were used to capture synchronized data. The camera captures RGB images and each camera frame was made to run through YOLOV3 algorithm for recognizing the objects in the image and tag them with bounding boxes.

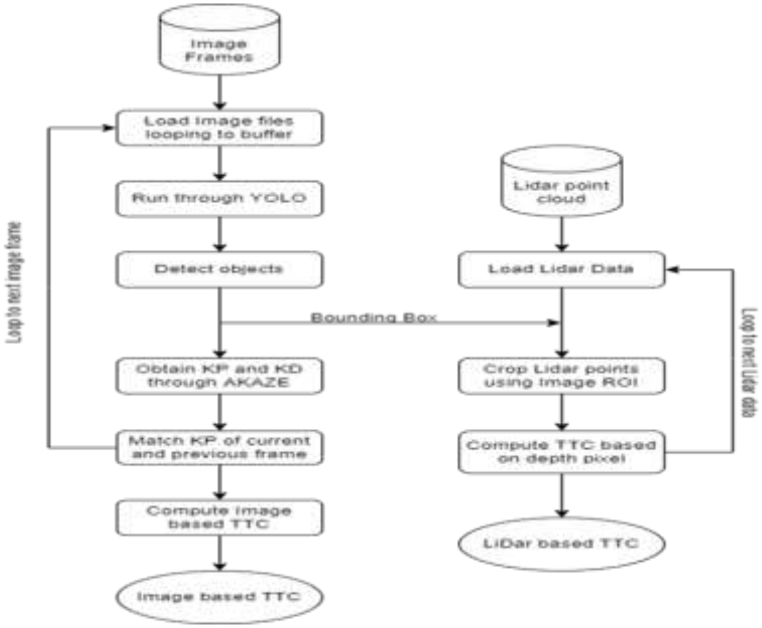

Fig. 4. Workflow Diagram

Fig. 4 represents the workflow used to calculate TTC from both the sensors. For calculating image based TTC, the image was first converted to grayscale. The data was processed to obtain key point descriptors using AKAZE. First, the image frame was passed through YOLOV3 object detection algorithm to identify objects or vehicles in frame and tag all of them with bounding boxes (BB). The image frame was tagged with BB passed through AKAZE (key point detection algorithm), for detecting key points in the ego lane BB. Subsequently, the frame was passed through FLANN for key point descriptor matching between current and previous frame. The distance ratio of the KPD between two frames was used to compute image based TTC.

Lidar TTC computation was obtained from both image and LIDAR point clouds. Lidar data and image were fused to overlay point clouds on to the image. Frame was used to detect objects and tag them with BB using YOLOV3. The detected image was focused on the ego lane-bounding box to cluster and crop LIDAR point clouds. The ego lane BB was obtained and was then used for cropping and clustering point clouds on to the ego lane BB. These LIDAR points were used to calculate distance from the vehicle, in return to calculate Lidar TTC computation.

3.0 Experimental Details

KITTI dataset is one of the most popular datasets for robotic systems and autonomous driving. It is composed of many hours worth of traffic events recorded using a variety of sensor modalities, including high-resolution RGB, grayscale stereo cameras, and 3D laser scanners. RGB image from monocular camera and LIDAR data was used for calculating TTC and the same was observed in different scenarios. The difference between the two sensors was used for calculating TTC.

TTC was calculated using equation of uniform velocity, in which the source vehicle(vehicle capturing the data) and the ego Lane Vehicle(vehicle that is in the same lane or in front of source vehicle) were used. If TTC is estimated zero means that both source and ego lane vehicle are travelling with same velocity. If TTC is positive, it implies that the ego lane vehicle velocity is greater than that of the source vehicle. Similarly, if TTC is positive, the velocity of the source vehicle is greater than that of the ego lane vehicle.

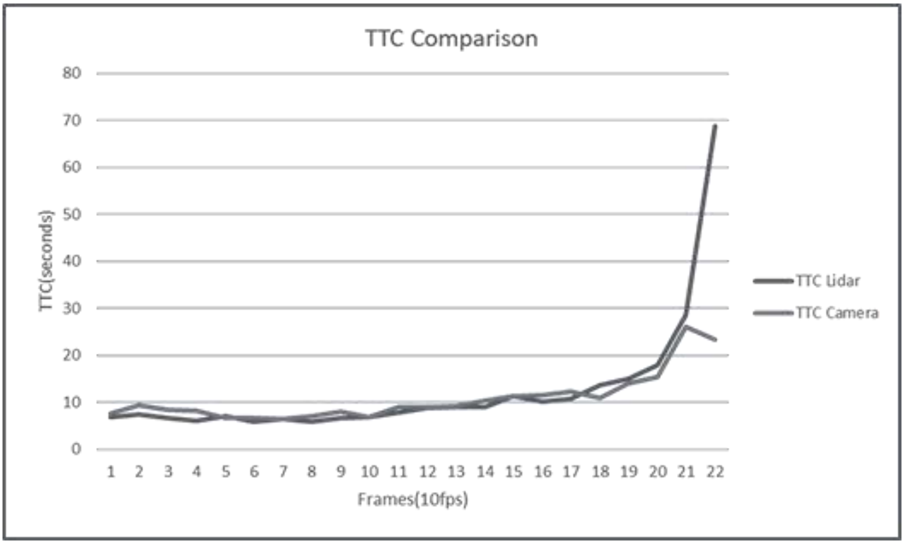

Fig. 5. TTC from KITTI dataset 1

In the first dataset both Lidar and camera performed similarly. This dataset was captured in heavy traffic scenario where the ego lane vehicle was close to the source vehicle. A spike in Lidar TTC is observed (Fig 5 and Table 1). It is because the source vehicle is very close to the ego lane vehicle causing difficulty in clustering the LIDAR point and velocity of source and ego lane vehicle is nearly zero.

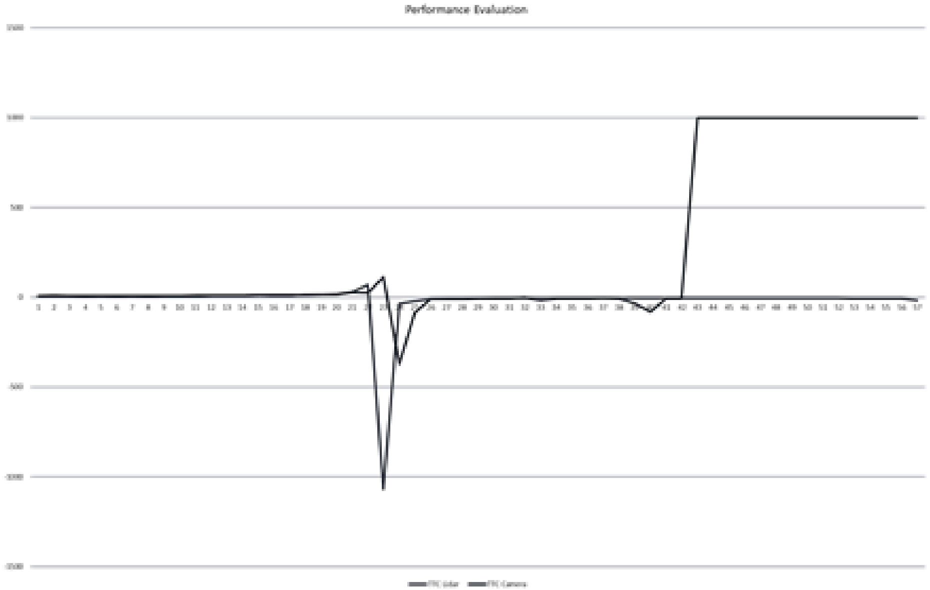

Fig. 6. TTC from KITTI dataset 2

The second dataset provided more complicated results. However, it showed clear difference in the ability to calculate TTC of the two sensors

(Fig 6). In this dataset both source vehicle and ego lane vehicle have varying velocity, which produces both positive and negative TTC. As observed in the previous dataset both sensors performed similarly when the vehicle is near to each other, but after certain distance camera is not able to track the key points to estimate TTC and hence resulted in a NaN value in case of the ego lane vehicle at certain distance from the source. However, Lidar was able to continue estimating TTC far beyond camera.

Table 1. TTC for comparison for dataset 1

TTC Lidar	TTC Camera
6.827779	7.609445
7.452518	9.440468
6.603506	8.301106
5.997915	8.11849
6.98247	6.700708
5.85468	6.660233
6.391715	6.452117
5.802376	7.088461
6.712941	8.062627
6.767212	6.892021
7.75002	8.993668
8.774075	9.001555
9.022193	9.182595
8.922193	10.355707
11.371478	11.286063
10.030941	11.444733
10.612555	12.293897
13.563763	10.942348
14.970927	13.914239
17.864875	15.37593
28.484072	25.993016
68.814718	23.390807

4.0 Conclusion

Safety critical systems of automobiles is facing challenges in maintaining price to safety benefit ratio. Advanced sensors such as stereo camera and LIDAR incur additional cost. Lidar based TTC was found to be more reliable than that of the camera based. While LIDAR TTC was able to estimate TTC beyond certain distance the Camera TTC showed limitations. Camera is very sensitive to light and can have further disadvantages. Also in rough weather, it is extremely difficult for these sensors to estimate a reliable TTC value, so computation of TTC using data fusion has higher accuracy in unconventional weather conditions such as rain ordust.Furthermore, a reliable TTC can be used to implement assisted braking system for vehicles, implementing self-driving cars with better TTC estimation, intelligent driver assistance system and other applications for better navigation in harsh climatic conditions.

References

01. Van Der Horst, Richard, and Jeroen Hogema. Time-to-collision and collision avoidance systems. (1993), 129-143.

02. Q Lim, Y He and U -X Tan, Real-Time Forward Collision Warning System Using Nested Kalman Filter for Monocular Camera, *IEEE International Conference on Robotics and Biomimetics (ROBIO*), 868-873, 2018, doi: 10.1109/ROBIO.2018.8665220.

03. Wei, P, Cagle, L, Reza, T, Ball, J and Gafford, J, LiDAR and camera detection fusion in a real-time industrial multi-sensor collision avoidance system. *Electronics,* 7(6),84. 2018

04. Szeliski, R, 2022. Computer vision: algorithms and applications. *Springer Nature*

05. Feng, X, Jiang, Y, Yang, X, Du, M and Li, X, Computer vision algorithms and hardware implementations*: A survey. Integration, 69*, 309-320, 2019.

06. Janai, J, G u¨ney, F, Behl, A and Geiger, A, Computer vision for autonomous vehicles: Problems, datasets and state of the art. *Founda- tions and Trends® in Computer Graphics and Vision,* 12(1–3), 1-308. 2020.

07. Miller, D, Sun, A and Ju, W, Situation awareness with different levels of automation. *IEEE International Conference on systems, man, and cybernetics (SMC) 688-693*, 2014,

08. I. Culjak, D Abram, T Pribanic, H Dzapo and M Cifrek, A brief introduction to OpenCV, *Proceedings of the 35th International Convention MIPRO,* 1725-1730, 2012

09. Zhao, L and Li, S, Object detection algorithm based on improved YOLOv3. *Electronics,* 9(3),537, 2020.

10. Elmenreich, W, An introduction to sensor fusion. *Vienna University of Technology, Austria*, 502, pp.1-28, 2002.

11. J Koci c´, N Jovicˇic´ and V Drndarevic´, Sensors and Sensor Fusion in Autonomous Vehicles, *26thTelecommunications Forum (TELFOR), 420-425, 2018, doi: 10.1109/TELFOR.2018.8612054*.

12. Kumar, GA, Lee, JH, Hwang, J, Park, J, Youn, SH and Kwon, S, LiDAR and camera fusion approach for object distance estimation inself-drivingvehicles.*Symmetry*,12(2*),324*, 2020.

13. Kotur, M, Luki c´, N, Krunic´, M and Lukacˇ, Zˇ, Camera and LiDAR Sensor Fusion for 3D Object Tracking in a Collision Avoid-

ance System. *Zooming Innovation in Consumer Technologies Conference (ZINC), IEEE, 198-202*, 2021.

14. S A K Tareen and Z Saleem, A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK, *International Conference on Computing, Mathematics and Engineering Technologies (iCoMET),* 1-10,2018doi:10.1109/ICOMET.2018.8346440.
15. Muja, M and Lowe, D, Flann-fast library for approximate nearest neighbors user manual. Computer Science *Department, University of British Columbia, Vancouver, BC, Canada,* 2009.
16. Maurya, A and Kumar D, Reliability of safety-critical systems: A state-of-the-art review. *Quality and Reliability Engineering International,* 36(7),2547-2568, 2020.
17. Dygalo, V, Keller, A and Evtiukov, S, Monitoring of vehicles' active safety systems in operation,*Transportation Research Procedia,* 50, 113-120, 2020.
18. Wiseman, Y, Revisiting the Anti-Lock Braking System. *Technical Report*.2021.
19. Farmer, CM, Effect of electronic stability control on automobile crashrisk,*Trafficinjuryprevention*,5(4),317-325,2004.
20. Wu, H, The Active Safety of the Automobile in the Future, *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference,AmericanSocietyofMechanicalEngineers*, 57106,V003T01A006, 2015.
21. Ishida, S and Gayko, JE, June. Development, evaluation and introduction of a lane keeping assistance system. *IEEE Intelligent Vehicles Symposium,* 943-944, IEEE,2004,
22. Yasin, JN, Mohamed, SA, Haghbayan, MH, Heikkonen, J, Tenhunen, H and Plosila, J, Unmanned aerial vehicles (uavs): *Collision avoidance systems and approaches. IEEE access,* 8, 105139-105155, 2020.
23. Pundlik, S, Peli, E and Luo, G, September, Time to collision and collision risk estimation from local scale and motion. *International Symposium on Visual Computing Springer, Berlin, Heidelberg*,728-737,2011.
24. Kilicarslan, M and Zheng, JY, Predict vehicle collision by TTC from motion using a single video camera. *IEEE Transactions on Intelligent Transportation Systems,* 20(2), 522-533, 2018.

25. S K Manghat and M El-Sharkawy, A Multi Sensor Real-time Tracking with LiDAR and Camera, *10th Annual Computing and Communication Workshop and Conference (CCWC)*, 0668- 0672, 2020,doi:10.1109/CCWC47524.2020.9031247.

26. Geiger A, Lenz P, Stiller C, Urtasun R Vision meets robotics: The KITTI dataset. *The International Journal of Robotics Research,* 32(11), 1231-1237,2013, doi:10.1177/0278364913491297.