Annotated Bibliography: Computer Vision for Autonomous Vehicles

Shanthan Kumar Bine

Lubar School of Business

University of Wisconsin Milwaukee

Author Note

Shanthan Kumar Bine, Lubar School of Business, University of Wisconsin Milwaukee. Prepared for course BUS ADM-744, Information Technology Strategy and Management, Professor H. Morris and submitted on (05/11/2023). Correspondence regarding this article should be directed to Shanthan Kumar Bine.

E-mail: (Sbine@uwm.edu)

Annotated Bibliography: Computer Vision for Autonomous Vehicles

Summary:

The sources presented in this annotated bibliography aim to explore the role of computer vision in autonomous vehicles to recognize and understand their surroundings in real-time, as well as how this technology can be used to increase safety and efficiency in the transportation industry. In recent years there have been significant advancements in the computer vision field, they have made it possible to achieve a high level of accuracy and reliability in autonomous driving applications.

The detection and identification of objects in a complex environment is the fundamental problem in computer vision for autonomous vehicles. Objects including other vehicles, people, and traffic signs are recognized and tracked using techniques like object recognition, tracking, and segmentation.

The computer vision for autonomous vehicles also includes mapping and localization, which involves creating a 3D model of the environment using information from sensors like LiDAR and cameras and then utilizing that model to precisely locate the vehicles in real-time.

Deep learning is essential to autonomous driving applications, it has completely changed the science of computer vision. Convolutional neural networks (CNN's) are useful for object identification and recognition. while Recurrent neural networks (RNN's) are used for tasks like object tracking and trajectory prediction.

The sources in this annotated bibliography will help in determining how computer vision has the potential to transform the auto industry and advance autonomous driving.

- Mao, Jiageng, Shaoshuai Shi, Xiaogang Wang, and Hongsheng Li. "3D Object Detection for Autonomous Driving: A Comprehensive Survey." arXiv preprint arXiv:2206.09474 (2022)

- **Link**: 2206.09474.pdf (arxiv.org)

The most recent study, "3D Object Detection for Autonomous Driving: A Comprehensive Survey," was released on arXiv in 2022. A thorough analysis of the advanced in 3D object identification for autonomous driving is provided by authors Mao, Jiageng, Shaoshuai Shi, Xiaogang Wang, and Hongsheng Li.

Annotated Bibliography: Computer Vision for Autonomous Vehicles

The authors discuss over the many methods and strategies used in 3D object detection, such as camera- and LiDAR-based methods as well as hybrid systems that include both. They give a thorough analysis of the performance, accuracy, and efficiency of these approaches, highlighting both their advantages and disadvantages. The authors additionally look into issues like occlusion, variable lighting, and the necessity for real-time performance when it comes to 3D object detection in autonomous vehicles. To meet these issues, they emphasize the significance of creating more effective algorithms and incorporating various sensors.

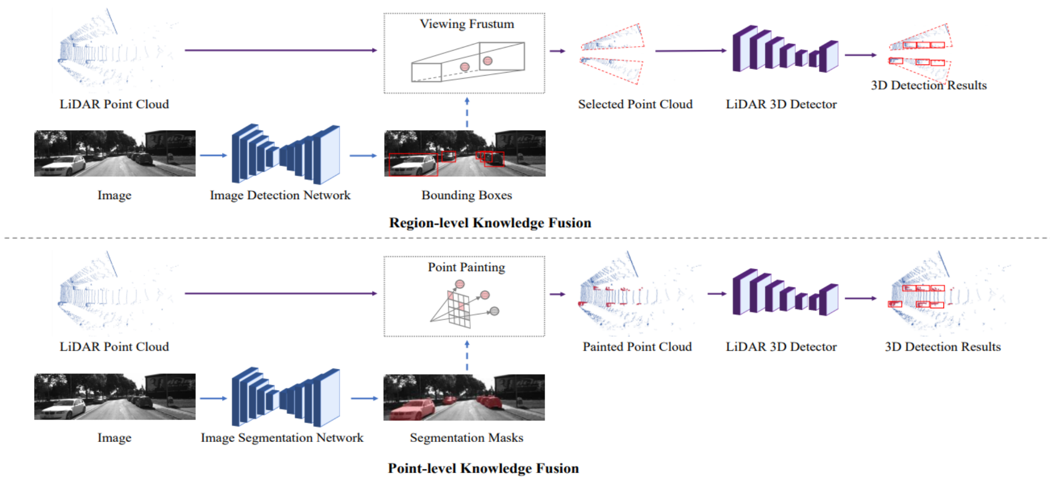


Fig. 17: An illustration of early-fusion based 3D object detection methods.

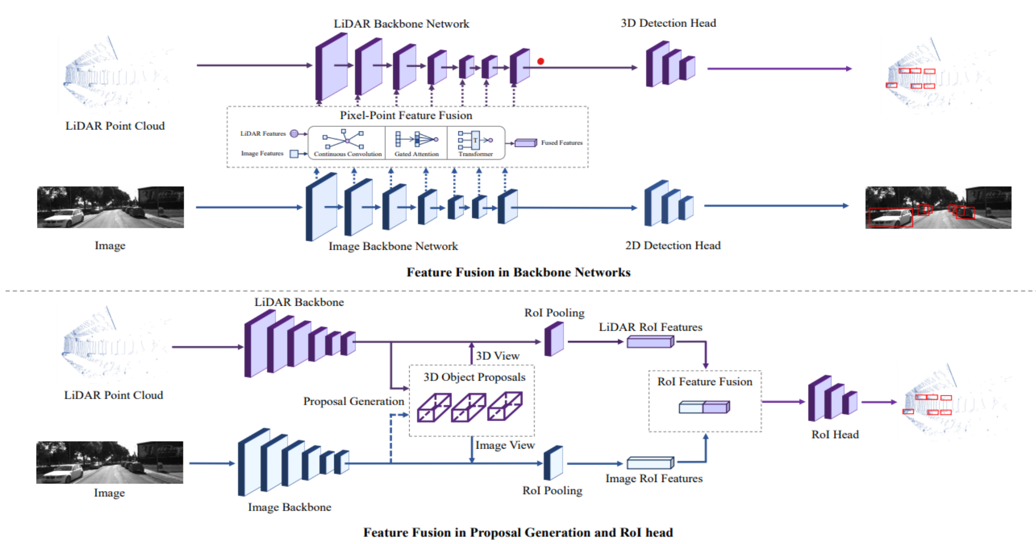


Fig. 18: An illustration of intermediate-fusion based 3D object detection methods.

The early-fusion frameworks are typically constructed in a step-by-step manner, starting with the use of 2D detection or segmentation networks to extract knowledge from images, followed by the transfer of that knowledge to a point cloud and the enhancement of the point cloud before feeding it to a LiDAR-based 3D object detector. Early fusion techniques can be categorized into two groups based on the fusion types: point-level knowledge fusion and region-level knowledge fusion. Figure 17 provides an illustration of the early-fusion based methods.

At the intermediate stages of a LiDAR-based 3D object detector, such as in backbone networks, at the proposal generation stage, or at the RoI refinement step, intermediate-fusion based approaches attempt to fuse image and LiDAR features. The stages of fusion can also be used to categorize these techniques. Figure 18 provides an example of intermediate-fusion based methods.

- J. Cahill et al., "Exploring the Viability of Bypassing the Image Signal Processor for CNN-Based Object Detection in Autonomous Vehicles," in IEEE Access, vol. 11, pp. 42302-42313, 2023, Doi: 10.1109/ACCESS.2023.3270710.

- **Link:** Exploring the Viability of Bypassing the Image Signal Processor for CNN-Based Object Detection in Autonomous Vehicles | IEEE Journals & Magazine | IEEE Xplore

From the title "Exploring the Viability of Bypassing the Image Signal Processor for CNN-Based Object Detection in Autonomous Vehicles," it is clear that the authors are interested in looking into a different method of object detection in autonomous vehicles.

They want to evaluate the viability and potential advantages of using Convolutional Neural Networks (CNNs) directly for object detection instead of the Image Signal Processor (ISP). The authors have thought that processing image data for object detection might be hampered, delayed, or limited by the conventional ISP pipeline used in autonomous vehicles. They may try to address these potential problems and enhance object detection performance by looking into the viability of circumventing the ISP.

According to their perspective, applying CNNs directly to the raw sensor data may produce object detection results that are more precise and trustworthy. They might contend that avoiding the ISP would streamline the processing pipeline, decrease latency, and possibly enhance the effectiveness of object detection systems in autonomous vehicles.

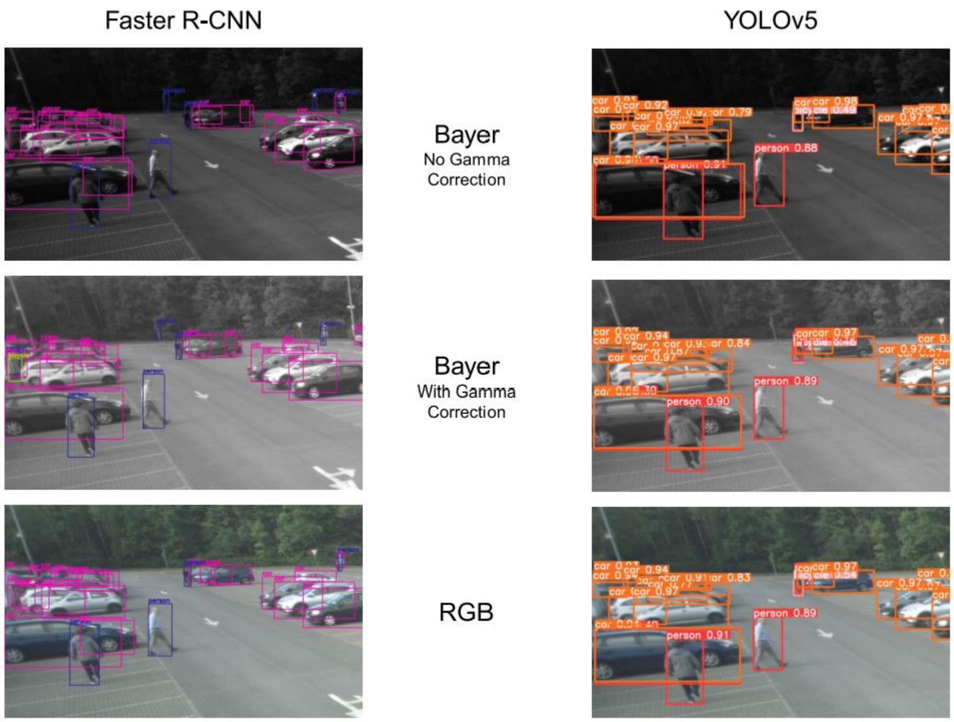


FIGURE 10. Figure showing three corresponding detections from each image type from the G-MIND dataset using both the Faster R-CNN and YOLOv5 object detection models. *Top:* Bayer images with no ISP processing; *Middle:* Bayer images with only gamma correction; *Bottom:* Full ISP processed RGB images.

- V. Dygalo, M. Lyashenko and P. Potapov, "Ways for Improving Efficiency of Computer Vision for Autonomous Vehicles and Driver Assistance Systems," 2019 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM), Sochi, Russia, 2019, pp. 1-5, doi: 10.1109/ICIEAM.2019.8743060.

- **Link**: [Ways for Improving Efficiency of Computer Vision for Autonomous Vehicles and Driver Assistance Systems | IEEE Conference Publication | IEEE Xplore]

In their study, Dygalo, Lyashenko, and Potapov explore strategies for enhancing the efficiency of computer vision in autonomous vehicles and driver assistance systems. They examine the characteristics of human vision and how they might affect the creation of computer vision systems.

Annotated Bibliography: Computer Vision for Autonomous Vehicles

The importance of factors like the field of vision, which can be influenced by a variety of factors and has a significant impact on driving safety, is emphasized by the authors. The requirements related to operational circumstances, such as variations in lighting and the capacity to detect motion in the dark, are also covered in the paper. The authors recommend a specific strategy that involves positioning two cameras in a way that promotes binocular vision.

The overall goal of the paper is to enhance computer vision methods in order to enhance the performance and safety of autonomous vehicles and driver assistance systems.

- M. Johnson-Roberson, C. Barto, R. Mehta, S. N. Sridhar, K. Rosaen and R. Vasudevan, "Driving in the Matrix: Can virtual worlds replace human-generated annotations for real world tasks?" 2017 IEEE International Conference on Robotics and Automation (ICRA), Singapore, 2017, pp. 746-753, doi: 10.1109/ICRA.2017.7989092.

- **Link**: 1610.01983.pdf (arxiv.org) / Driving in the Matrix: Can virtual worlds replace human-generated annotations for real world tasks? | IEEE Conference Publication | IEEE Xplore

The need for a significant amount of annotated data, according to the authors, is one of the biggest obstacles in the development of autonomous vehicles. Machine learning algorithms that enable autonomous vehicles to recognize and react to various objects and situations on the road require annotated data to be trained. Although it requires human expertise, annotating data is a time-consuming and expensive task.

The authors suggest using virtual worlds as a source of annotated data to solve this problem. They contend that creating large quantities of labeled data that closely resemble the real world can be done effectively and inexpensively using virtual worlds. The authors show how labeled data can be produced for object detection tasks using the virtual world of the video game Grand Theft Auto V (GTA V).

The authors demonstrate how a deep neural network can be trained to recognize objects in real-world images using labeled data created in a virtual environment. They demonstrate that the two networks perform similarly by comparing the performance of a neural network trained on virtual world data to that of a network trained on real-world data.

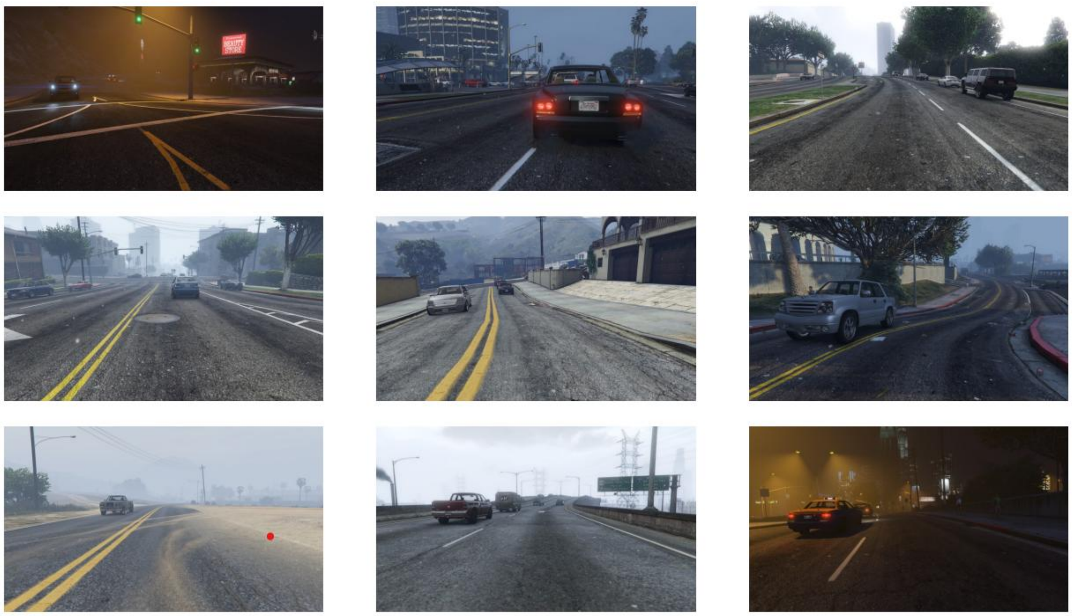


Fig. 1: Sample images captured from the video game based simulation engine proposed in this paper. A range of different times of day are simulated including day, night, morning and dusk. Additionally the engine captures complex weather and lighting scenarios such as driving into the sun, fog, rain and haze.

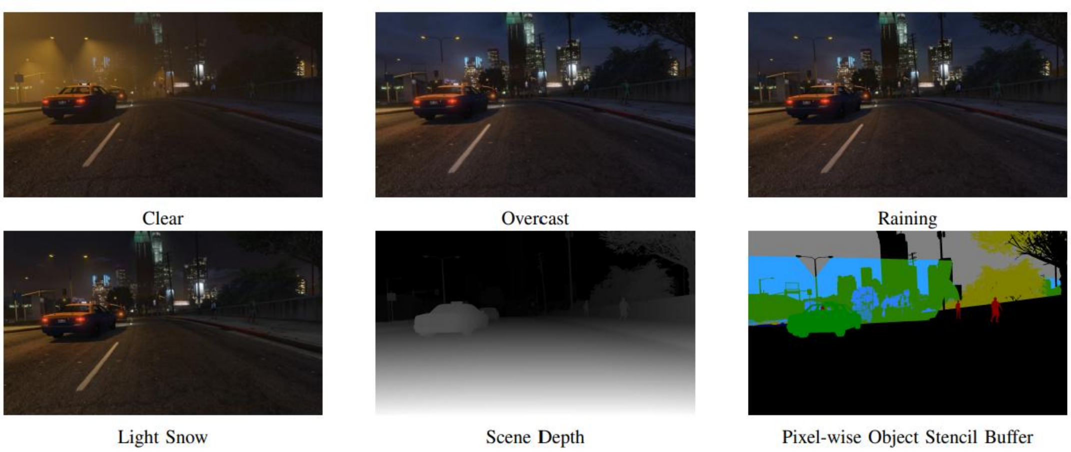


Fig. 2: Four different weather types appear in the training data. The simulation can be paused and the weather condition can be varied. Additionally note the depth buffer and object stencil buffer used for annotation capture. In the depth image, the darker the intensity the farther the objects range from the camera. In the stencil buffer, we have artificially applied colors to the image's discrete values which correspond to different object labels for the game. Note that these values cannot be used directly. The process by which these are interpreted is highlighted in Section III-B.

Annotated Bibliography: Computer Vision for Autonomous Vehicles

According to the authors, annotated data for tasks in the real world, like autonomous driving, can be found in virtual worlds. They contend that creating large quantities of labeled data for machine learning algorithms in virtual worlds can be done effectively and affordably. The quality of the labeled data produced in virtual worlds depends on the accuracy of the simulation, they also acknowledge the limitations of these worlds.

- H. Liu, C. Lin, B. Gong and D. Wu, "Automatic Lane-Level Intersection Map Generation using Low-Channel Roadside LiDAR," in IEEE/CAA Journal of Automatica Sinica, vol. 10, no. 5, pp. 1209-1222, May 2023, doi: 10.1109/JAS.2023.123183.

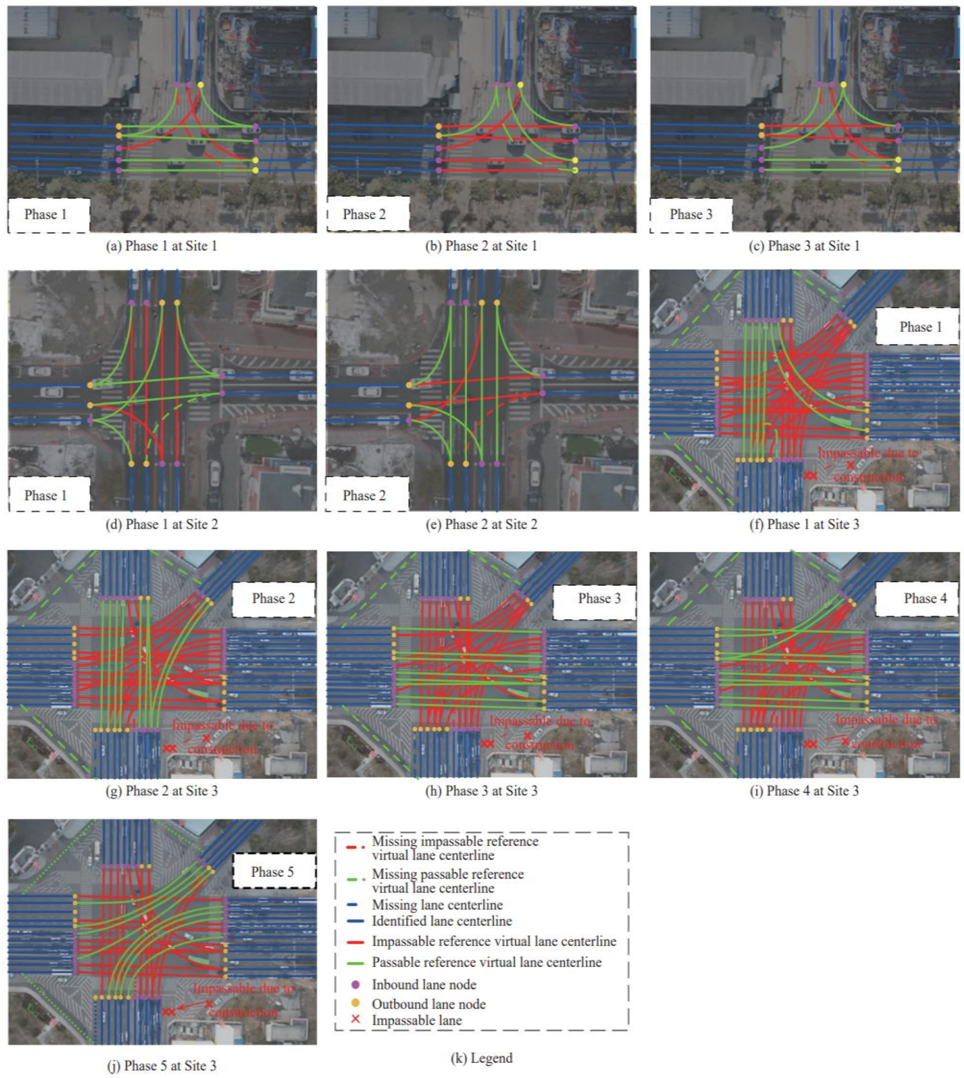
- **Link**: Automatic Lane-Level Intersection Map Generation using Low-Channel Roadside LiDAR | IEEE Journals & Magazine | IEEE Xplore

The paper “Automatic Lane-Level Intersection Map Generation using Low-Channel Roadside LiDAR” was published by IEEE/CAA Journal of Automatica Sinica in May 2023. The authors are Hui Liu, Ciyun Lin, Bowen Gong, and Dayong Wu.

Using a low-channel roadside LiDAR, the authors of the paper "Automatic Lane-Level Intersection Map Generation using Low-Channel Roadside LiDAR" suggest a novel technique for creating lane-level intersection maps. From their vantage point, traditional techniques for mapping an HD intersection are costly, time-consuming, and labor-intensive.

They contend that the approach they have suggested, which makes use of a low-channel roadside LiDAR, can automatically and dynamically produce an intersection at the lane level, complete with the signal phases, geometry, layout, and lane directions.

In comparison to other methods, such as satellite-based, MMS-based, and crowdsourcing-based lane mapping methods, they also assert that their method has a lower average lane location deviation and a shorter update period.



- A. Pidurkar, R. Sadakale and A. K. Prakash, "Monocular Camera based Computer Vision System for Cost Effective Autonomous Vehicle," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2019, pp. 1-5, doi: 10.1109/ICCCNT45670.2019.8944496.

- **Link**: [Monocular Camera based Computer Vision System for Cost Effective Autonomous Vehicle | IEEE Conference Publication | IEEE Xplore]

Annotated Bibliography: Computer Vision for Autonomous Vehicles

The authors, A. Pidurkar, R. Sadakale, and A. K. Prakash, have presented a research paper titled "Monocular Camera based Computer Vision System for Cost Effective Autonomous Vehicle" at the 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT) held in Kanpur, India.

The authors' main goal is to create a computer vision system that uses monocular cameras for autonomous vehicles. The use of monocular cameras indicates that an economical method of putting autonomous vehicle technology into practice is being investigated.

They are considering the difficulties of using pricey sensor systems like LiDAR or radar and working to create a more practical substitute. The authors' main argument is that sophisticated computer vision algorithms and monocular cameras can give autonomous vehicles adequate perception abilities. They might emphasize how using such a system could be advantageous in terms of both cost and utility.

Furthermore, they go over the difficulties encountered when trying to develop a trustworthy and accurate computer vision system based on monocular cameras, including depth estimation, object recognition, and tracking.

- M. Sugadev, Yogesh, P. K. Sanghamreddy and S. K. Samineni, "Rough Terrain Autonomous Vehicle Control Using Google Cloud Vision API," 2019 2nd International Conference on Power and Embedded Drive Control (ICPEDC), Chennai, India, 2019, pp. 244-248, doi: 10.1109/ICPEDC47771.2019.9036621.

- **Link**: Rough Terrain Autonomous Vehicle Control Using Google Cloud Vision API | IEEE Conference Publication | IEEE Xplore

The authors of this article are proposing the design and implementation of a rough terrain autonomous vehicle using the Rocker Bogie mechanism and the Google Cloud Vision API for obstacle avoidance. The Rocker Bogie mechanism is a suspension arrangement commonly used in vehicles that travel on rough surfaces or multiple terrains. The Google Cloud Vision API is a remote real-time image processing platform that can process, classify, and extract metadata from images and videos.

Annotated Bibliography: Computer Vision for Autonomous Vehicles

The authors highlight the importance of computer vision in automated vehicles, robotics, and industrial automation, as it allows machines to emulate human vision capabilities. They mention that computer vision involves tasks such as object detection, tracking, and recognition, which can be achieved using convolutional neural networks (CNNs) and deep learning techniques.

To address the limitations of standalone and portable computer vision systems, the authors propose leveraging remote cloud platforms like the Google Cloud Vision API for image analysis. They mention that the API provides models for image classification, object detection, face detection, and text recognition, enabling accurate and real-time analysis of captured images.

In the proposed work, the authors aim to integrate the Google Cloud Vision API with a rough terrain autonomous vehicle. They outline the functional blocks of the system, including the Pi Camera, Raspberry Pi (used as the control unit), dual power DC grid, DC motors and motor drivers for vehicle movement, and obstacle sensors for detecting and avoiding obstacles.

- A. Bewley et al., "Learning to Drive from Simulation without Real World Labels," 2019 International Conference on Robotics and Automation (ICRA), Montreal, QC, Canada, 2019, pp. 4818-4824, doi: 10.1109/ICRA.2019.8793668.

- **Link**: [Learning to Drive from Simulation without Real World Labels | IEEE Conference Publication | IEEE Xplore]

The article "Learning to Drive from Simulation without Real World Labels" describes a technique for training an autonomous driving system using only simulated data, without the need for manually annotated real-world data. A high-fidelity simulation environment, an adversarial imitation learning algorithm, and a novel attention mechanism that concentrates on significant areas of the input data make up the authors' novel approach, which combines these three essential components.

Annotated Bibliography: Computer Vision for Autonomous Vehicles

The suggested technique involves training the autonomous driving system in a highly realistic simulation setting that faithfully simulates the physics of the real world. Using an adversarial imitation learning algorithm, the system is trained to mimic the actions of a skilled driver.

The system can more effectively learn from the simulation data thanks to the authors' novel attention mechanism, which directs the system's attention to crucial areas of the input data like road edges and obstacles. On a variety of driving tasks, including lane following, lane changing, and merging onto a highway, the suggested method is assessed.

According to the results, a system that was only trained on simulated data can perform as well as or even better than a system that was trained on real-world data. In conclusion, the paper presents a novel method for training autonomous driving systems using only simulated data, which has the potential to drastically cut the cost and time needed to develop autonomous driving systems.

- Teeti, Izzeddin & Khan, Salman & Shahbaz, Ajmal & Bradley, Andrew & Cuzzolin, Fabio. (2022). Vision-based Intention and Trajectory Prediction in Autonomous Vehicles: A Survey. 10.24963/ijcai.2022/781.

- **Link**: Vision-based Intention and Trajectory Prediction in Autonomous Vehicles: A Survey (ijcai.org) / Vision-based Intention and Trajectory Prediction in Autonomous Vehicles: A Survey (ijcai.org)

The authors of the article "Vision-based Intention and Trajectory Prediction in Autonomous Vehicles: A Survey" seek to present a thorough review of recent research in the area of vision-based intention and trajectory prediction in autonomous vehicles. They go over a variety of methods, including data-driven techniques, probabilistic models, and rule-based strategies, for predicting future vehicular and pedestrian movements.

The authors also talk about problems like the lack of data, the difficulty in predicting people's behavior, and the demand for real-time predictions in autonomous vehicle systems.

Annotated Bibliography: Computer Vision for Autonomous Vehicles

For the safe and effective operation of autonomous vehicles, they emphasize the significance of accurate prediction.

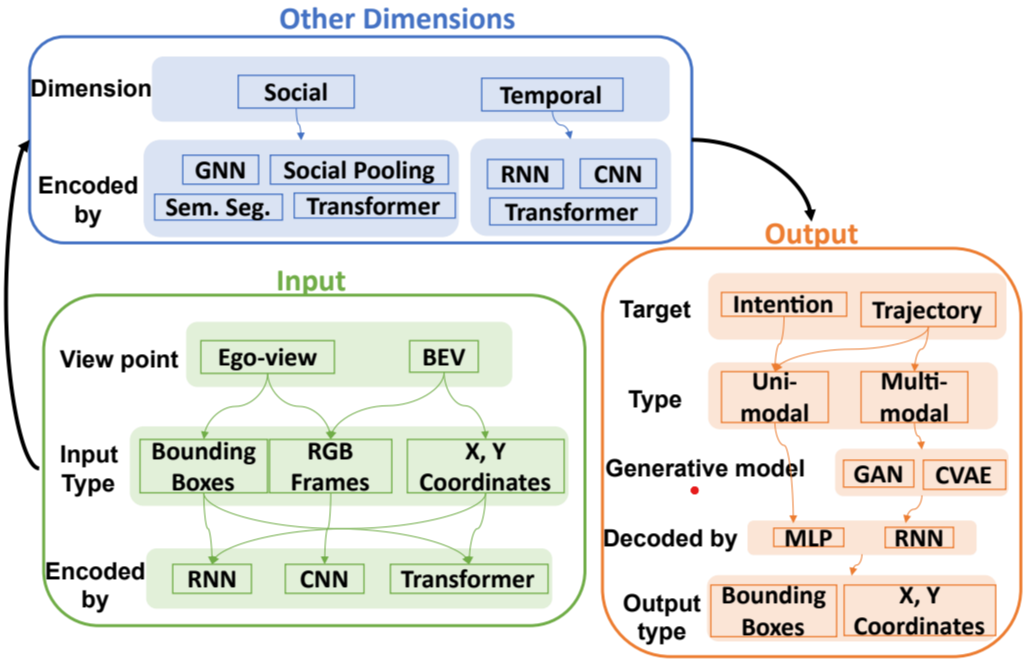


Figure 1. Prediction stack components and flow.

The goal of the authors is to lay out a plan for future research in this field and to promote the creation of reliable prediction algorithms that can aid in the secure deployment of autonomous vehicles.

- A. Geiger, P. Lenz and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, 2012, pp. 3354-3361, doi: 10.1109/CVPR.2012.6248074.

- **Link**: [Are we ready for autonomous driving? The KITTI vision benchmark suite | IEEE Conference Publication | IEEE Xplore]

Annotated Bibliography: Computer Vision for Autonomous Vehicles

The need for more difficult benchmarks that mimic real-world situations is highlighted in this paper, especially for visual recognition systems. In order to create new benchmarks for tasks like stereo, optical flow, visual odometry/SLAM, and 3D object detection, the authors make use of their autonomous driving platform.

The benchmarks include 389 stereo and optical flow image pairs, stereo visual odometry sequences of 39.2 km length, and more than 200k 3D object annotations captured in cluttered scenarios, which together make up a sizable amount of data. These benchmarks are designed to offer a more accurate assessment of visual recognition systems when they are used in situations outside of the lab.

Modern algorithms' findings show that methods outperform average in the lab when applied to real-world datasets outside of established datasets like Middlebury. This emphasizes how crucial it is to create benchmarks that can properly assess how well visual recognition systems perform in difficult real-world scenarios.

By offering difficult benchmarks with new difficulties, the authors hope to lessen bias in the computer vision community. The benchmarks are openly accessible online, allowing researchers and developers to test their algorithms and evaluate their performance in a more accurate environment.

- Huval, B., Wang, T., Tandon, S., Kiske, J., Song, W., Pazhayampallil, J., Andriluka, M., Rajpurkar, P., Migimatsu, T., Cheng-Yue, R., Mujica, F., Coates, A., & Ng, A. Y. (2015). An empirical evaluation of deep learning on highway driving. arXiv preprint arXiv:1504.01716.

- **Link**: [1504.01716.pdf (arxiv.org)]

An arXiv preprint of the article "An empirical evaluation of deep learning on highway driving" by Huval et al. was released in 2015. In their study, the authors demonstrate how deep learning algorithms can be used to identify and categorize objects in images captured while a car is moving down a highway. A deep convolutional neural network (CNN) was trained by the authors using a sizable dataset of images that included over 100,000 examples of objects that might be seen on a highway, including cars, trucks, motorcycles, and pedestrians. The performance of the trained network was then assessed using a different test set of images, which were also captured from a moving car on a highway.

4

Fig. 3: overfeat-mask

The study's findings demonstrated that the deep CNN outperformed several other cutting-edge techniques for the same task in terms of accuracy when it came to object detection and classification. The size of the network, the volume of training data, and the use of data augmentation techniques were just a few of the variables that the authors tested in their experiments.

Overall, the paper makes a significant contribution to the fields of computer vision and deep learning by showing how these methods can be used in practical situations like autonomous driving.

- Yi Xiao, Felipe Codevilla, Akhil Gurram, Onay Urfalioglu, and Antonio M. Lopez. "Multimodal End-to-End Autonomous Driving." IEEE International Conference on Robotics and Automation (ICRA), 2018, pp. 1-8.
- **Link**: 1906.03199v2 (arxiv.org)

A system for autonomous driving is proposed in the paper "Multimodal End-to-End Autonomous Driving" by Yi Xiao , Felipe Codevilla, Akhil Gurram, Onay Urfalioglu, and Antonio M. Lopez. That makes use of various sensors and modalities, such as lidar, camera, and radar, to increase the accuracy and robustness of the system. The suggested system is a complete neural network that receives unprocessed sensor data from various sources and generates steering, acceleration, and braking commands to control the vehicle. Using a mix of supervised and reinforcement learning methods, the neural network is trained on a sizable dataset of real-world driving scenarios.

The authors demonstrate that, on a number of metrics, including accuracy, robustness, and safety, their system performs better than other cutting-edge autonomous driving systems. Additionally, they show how well their system works in difficult driving conditions like at night and inclement weather.

Overall, the study makes a significant contribution to the field of autonomous driving by showing how multimodal end-to-end methods can increase the precision and robustness of such systems.

- Weng, Xinshuo & Man, Yunze & Cheng, Dazhi & Park, Jinhyung & O'toole, Matthew & Kitani, Kris. (2020). All-In-One Drive: A Large-Scale Comprehensive Perception Dataset with High-Density Long-Range Point Clouds. 10.13140/RG.2.2.21621.81122.

- **Link**: [(PDF) All-In-One Drive: A Large-Scale Comprehensive Perception Dataset with High-Density Long-Range Point Clouds (researchgate.net)]

The paper titled "All-In-One Drive: A Comprehensive Perception Dataset with High-Density Long-Range Point Clouds" by Xinshuo Weng, Yunze Man, Jinhyung Park, Ye Yuan, Dazhi Cheng, Matthew O'Toole, Kris Kitan, describes a new dataset that contains high-density long-range point clouds for comprehensive perception tasks in autonomous driving.

Although there are already a number of datasets available for autonomous driving perception tasks, the majority of them have restrictions on their coverage or the kinds of objects they include, according to the paper's authors. The authors present the All-In-One Drive dataset, which seeks to offer a thorough representation of the urban driving environment, in order to address this limitation. The dataset contains information gathered from a unique sensor suite made up of numerous lidar sensors, cameras, and GPS/IMU sensors. The ability of the lidar sensors to collect dense point clouds over long distances enables accurate perception of the surroundings.

There are many different types of objects in the dataset, including cars, bicycles, pedestrians, traffic signs, and road markings. Daylight and nighttime data are both included in the data collection process under various lighting and weather conditions.

The authors present findings from a variety of perception tasks, such as object detection and segmentation, semantic segmentation, and object tracking, to illustrate the potential of the All-In-One Drive dataset. The findings demonstrate that the dataset is suitable for the given tasks and can be used to test and develop perception algorithms for automated driving. The All-In-One Drive dataset offers an extensive and varied set of data for perception tasks, and as a whole, it contributes significantly to the field of autonomous driving. believe that the dataset will be an important tool for scientists and engineers who are researching and developing autonomous driving technology.

- B. Kanchana, R. Peiris, D. Perera, D. Jayasinghe and D. Kasthurirathna, "Computer Vision for Autonomous Driving," 2021 3rd International Conference on Advancements in Computing (ICAC), Colombo, Sri Lanka, 2021, pp. 175-180, doi: 10.1109/ICAC54203.2021.9671099.

- **Link**: cdap.sliit.lk

The paper "Computer Vision for Autonomous Driving" published by IEEE discusses the application of computer vision techniques in autonomous driving. The authors, Bimsara Kanchana, Rojith Peiris, Damitha Perera, Dulani Jayasinghe, and Dharshana Kasthurirathna, provide an overview of the various computer vision techniques used in autonomous driving, including object detection, lane detection, and pedestrian detection.

The introduction to autonomous driving and computer vision in it serves as the paper's introduction. Following that, the authors go over the various parts of a self-driving car and how computer vision is used in each one. They also discuss the difficulties in dealing with occlusion and handling various lighting conditions when creating computer vision algorithms for autonomous driving. The authors then go over the various computer vision methods that are employed in autonomous driving, beginning with object detection. They describe various methods for object detection, such as deep learning-based techniques and conventional computer vision techniques.

The authors then go on to talk about lane detection and pedestrian detection, giving a thorough overview of the various methods applied in each situation. scenario worlds.

The paper also discusses 3D object detection, which is used to identify objects in a 3D environment, and semantic segmentation, which is used to categorize the various elements in a scene. The use of computer vision in ADAS (advanced driving assistance systems) and the difficulties in implementing autonomous vehicles in practical situations are also covered by the authors.

"Computer Vision for Autonomous Driving" offers a thorough overview of the various computer vision methods used in autonomous driving as well as the difficulties in their creation. It is a helpful tool for academics and professionals involved in computer vision and autonomous vehicles.

- A. de la Escalera, L. E. Moreno, M. A. Salichs and J. M. Armingol, "Road traffic sign detection and classification," in IEEE Transactions on Industrial Electronics, vol. 44, no. 6, pp. 848-859, Dec. 1997, doi: 10.1109/41.649946.

- **Link**: Road traffic sign detection and classification | IEEE Journals & Magazine | IEEE Xplore

The paper titled "Road traffic sign detection and classification" was published in the December 1997 issue of the IEEE Transactions on Industrial Electronics. The authors of the paper are A. de la Escalera, L. E. Moreno, M. A. Salichs, and J. M. Armingol. It is one of the most cited article.

The method for identifying and categorizing road traffic signs using computer vision techniques is presented in the paper. The suggested approach uses a camera mounted on a moving car to take pictures of the road signs. After that, image processing algorithms are used to process the images in order to find and categorize the signs. The proposed method's various steps, such as image preprocessing, feature extraction, and classification, are described in detail in the paper. The authors divide the signs into various categories using various features, including color, shape, and texture.

Annotated Bibliography: Computer Vision for Autonomous Vehicles

The results of experiments done to gauge how well the suggested method works are also presented in the paper. A dataset of road signs gathered from real-world scenarios was used in the experiments. The findings demonstrate that the suggested method can detect and classify road signs with high accuracy.

Overall, the paper offers a thorough method for identifying and categorizing traffic signs on roads using computer vision techniques. By giving drivers real-time information about the road signs, the suggested method has the potential to increase road safety and lower the risk of accidents.

- Y. Li et al., "A Deep Learning-Based Hybrid Framework for Object Detection and Recognition in Autonomous Driving," in IEEE Access, vol. 8, pp. 194228-194239, 2020, doi: 10.1109/ACCESS.2020.3033289.

- **Link**: A Deep Learning-Based Hybrid Framework for Object Detection and Recognition in Autonomous Driving | IEEE Journals & Magazine | IEEE Xplore

In this paper, a hybrid deep learning framework for object detection and recognition in autonomous driving is presented. The strengths of the two well-known object detection algorithms You Only Look Once (YOLO) and Region-based Convolutional Neural Network (R-CNN) are combined in the proposed framework. While R-CNN is used for precise object recognition and classification, the YOLO algorithm is used to quickly identify objects in images.

A YOLO-based detector, an R-CNN classifier, and a fusion module that combines the outputs of the two components make up the three main parts of the proposed hybrid framework. Regions of interest in the image are found using the YOLO-based detector and are then forwarded to the R-CNN-based classifier for additional analysis. The final results of object detection and recognition are created by fusing the outputs of the two components. On the KITTI dataset, which is frequently used in autonomous driving research, the proposed framework was assessed.

The findings demonstrated that the suggested framework outperformed YOLO or R-CNN alone in terms of detection and recognition performance. Additionally, it was demonstrated that the suggested framework was computationally effective, making it appropriate for real-time applications.

Overall, the proposed hybrid framework offers a promising method for autonomous vehicle object detection and recognition and may be used to increase the dependability and safety of these vehicles.

- S. -Y. Yu, A. V. Malawade, D. Muthirayan, P. P. Khargonekar and M. A. A. Faruque, "Scene-Graph Augmented Data-Driven Risk Assessment of Autonomous Vehicle Decisions," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 7, pp. 7941-7951, July 2022, doi: 10.1109/TITS.2021.3074854.
- **Link**: 2009.06435.pdf (arxiv.org)

In order to evaluate the risk involved with autonomous vehicle decisions, this paper suggests a scene-graph enhanced data-driven approach. The authors begin by building a scene graph by examining the data coming from the lidar, radar, and camera sensors of the autonomous vehicle. Scene graphs show the spatial relationships between various scene elements like moving cars, people walking around, and obstacles.

The authors then train a risk assessment model based on the scene graph using a data-driven methodology. The model uses the scene graph as input and produces a risk score that shows how likely it is that a collision or other unsafe event will occur.

The authors used experiments on a real-world dataset and compared the outcomes with those attained using a conventional rule-based approach to validate the proposed approach. The outcomes demonstrate that, in terms of accuracy and robustness, the proposed approach performs better than the conventional approach.

Annotated Bibliography: Computer Vision for Autonomous Vehicles

Overall, the suggested method offers a promising approach for determining the risk involved in autonomous vehicle choices, which is essential for guaranteeing the safety of autonomous vehicles in practical situations.

- Rastgoftar, Hossein & Bingxin, Zhang & Atkins, Ella. (2018). A Data-Driven Approach for Autonomous Motion Planning and Control in Off-Road Driving Scenarios. 5876-5883. 10.23919/ACC.2018.8431069.
- **Link**: 1805.09951.pdf (arxiv.org)

A new method for autonomous motion planning and control in off-road driving scenarios is presented in the paper "A Data-Driven Approach for Autonomous Motion Planning and Control in Off-Road Driving Scenarios". The suggested approach creates a control policy that enables a vehicle to navigate through complicated and difficult off-road environments by combining machine learning techniques and conventional control methods.

The authors train a machine learning model with data from human-driven off-road driving scenarios, and the model is then used to forecast the best steering angle and speed for the vehicle based on its current condition and the surroundings. The final control signals are produced by combining the predicted control inputs with conventional control techniques, such as PID controllers.

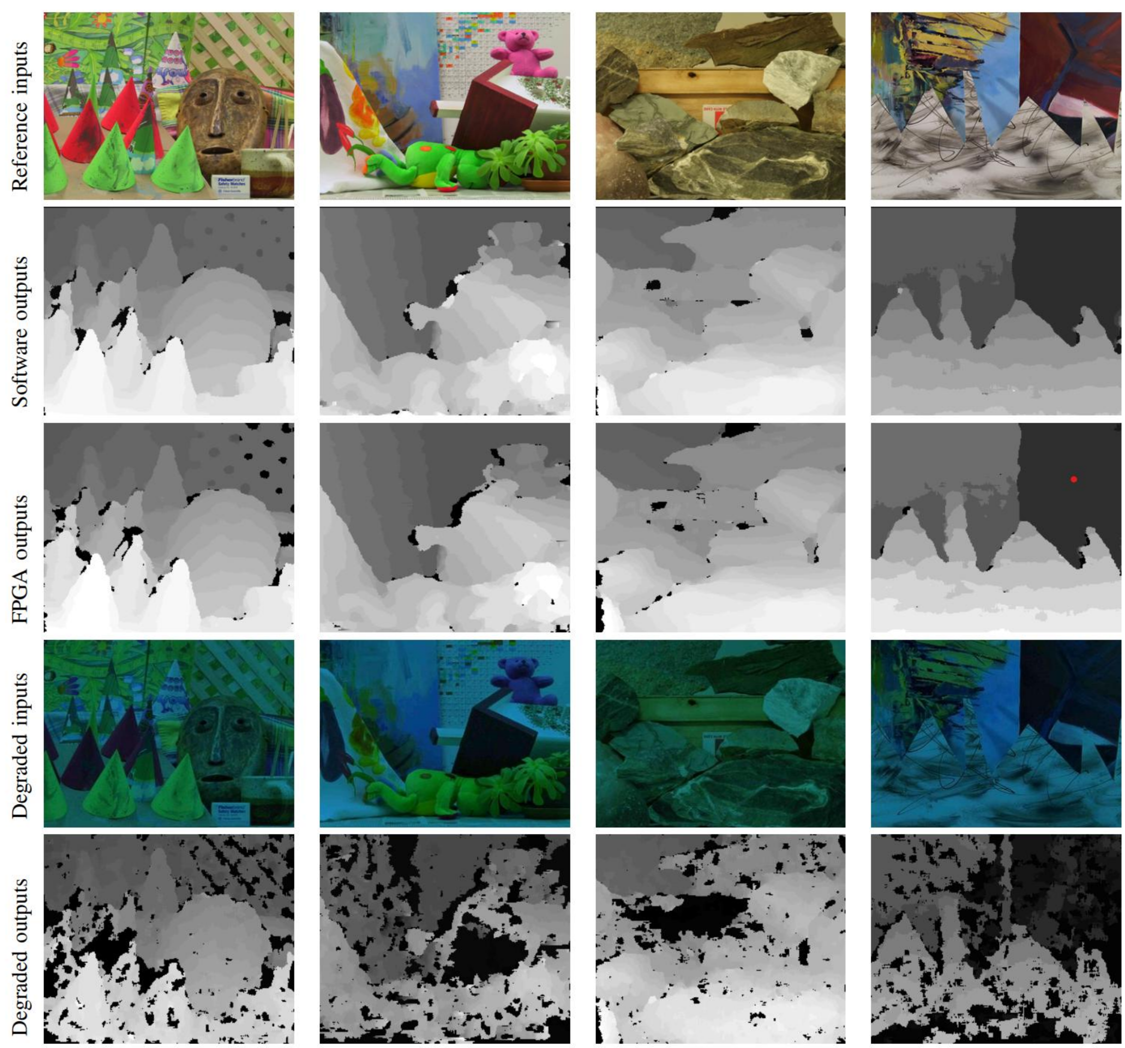
On a 1:10 scale off-road vehicle platform, simulations and actual experiments are used to assess the proposed approach. The findings demonstrate that the suggested method can successfully navigate through difficult off-road terrain, such as steep hills, rocks, and uneven terrain, while achieving high levels of performance and stability.

Overall, the paper presents a novel and practical method for autonomous off-road driving that combines the advantages of traditional control techniques and machine learning. The suggested method might make it possible to create stronger and more dependable autonomous off-road vehicles for a variety of uses, from mining and farming to military and defense.

Annotated Bibliography: Computer Vision for Autonomous Vehicles

- D. Gregorek, S. Srinivas, S. Nasrulla, S. Paul and R. Bachmayer, "Towards Energy-Optimized On-Board Computer Vision for Autonomous Underwater Vehicles," OCEANS 2022, Hampton Roads, Hampton Roads, VA, USA, 2022, pp. 1-6, doi: 10.1109/OCEANS47191.2022.9977107.
- **Link**: Towards Energy-Optimized On-Board Computer Vision for Autonomous Underwater Vehicles | IEEE Conference Publication | IEEE Xplore

The paper titled "Towards Energy-Optimized On-Board Computer Vision for Autonomous Underwater Vehicles" was presented at the OCEANS 2022 conference held in Hampton Roads, VA, USA. The authors of the paper are D. Gregorek, S. Srinivas, S. Nasrulla, S. Paul, and R. Bachmayer.



(a) Cone (b) Teddy (c) Rocks1 (d) Sawtooth

Fig. 8: Stereo test images for disparity computations and outputs

Annotated Bibliography: Computer Vision for Autonomous Vehicles

The development of an autonomous underwater vehicle (AUV)'s on-board computer vision system is covered in this paper. In order for AUVs to be used for tasks like underwater exploration, inspection, and maintenance, the authors emphasize the value of computer vision. The endurance and range of AUVs are, however, constrained by the high computational demands of computer vision algorithms, which can result in significant energy consumption.

The authors offer a framework for energy-optimized on-board computer vision that addresses this problem by considering the system's hardware, software, and algorithmic components. They discuss a number of energy-saving strategies, such as effective data transfer, low-power hardware, and improved algorithms. The authors also provide experimental findings that show how their strategy effectively lowers energy consumption while maintaining high performance.

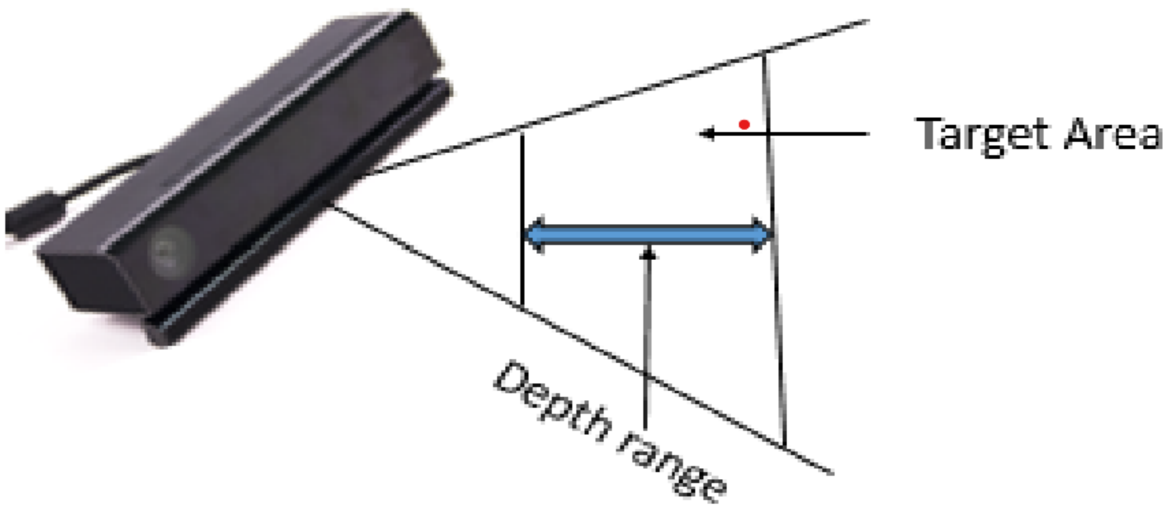
Overall, the paper offers insightful information on the creation of energy-efficient computer vision systems for AUVs, which can increase their range and endurance while allowing them to carry out more difficult tasks.

- N. Eric and J. -W. Jang, "Kinect depth sensor for computer vision applications in autonomous vehicles," 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN), Milan, Italy, 2017, pp. 531-535, doi: 10.1109/ICUFN.2017.7993842.

- **Link**: Kinect depth sensor for computer vision applications in autonomous vehicles | IEEE Conference Publication | IEEE Xplore

The paper "Kinect depth sensor for computer vision applications in autonomous vehicles" by N. Eric and J.-W. Jang was presented at the 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN) in Milan, Italy.

The use of the Kinect depth sensor in computer vision applications for autonomous vehicles is covered in the paper. The commercially available, low-cost Kinect sensor can deliver RGB data as well as depth information. The affordability, simplicity of integration, and capacity to deliver precise depth information are just a few of the benefits of using the Kinect sensor in the context of autonomous vehicles that are discussed by the authors.



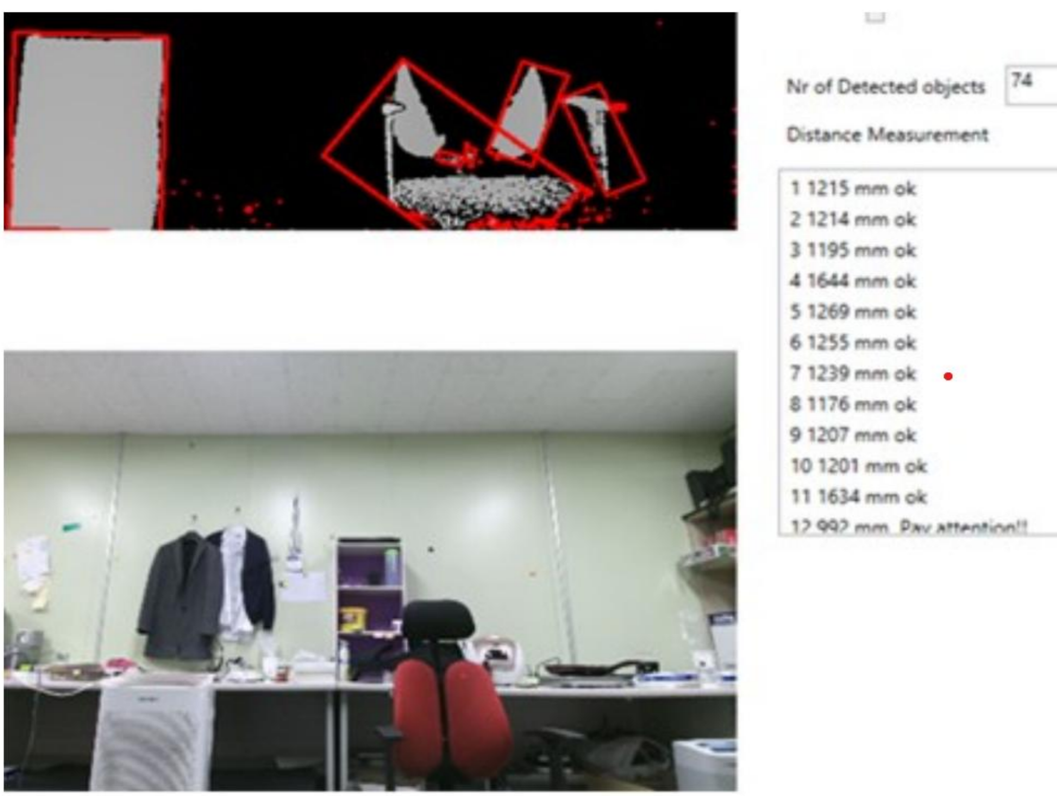


Fig.4 Corresponding color image

The experimentation in this paper shows that the Kinect sensor performs well in a number of computer vision tasks, including object detection and tracking, lane detection, and obstacle avoidance. The authors also go over the Kinect sensor's drawbacks, such as its short range and noise sensitivity. The Kinect depth sensor is a promising tool for computer vision applications in autonomous vehicles, and it can deliver precise depth information for a variety of tasks at a low cost, according to the paper's overall conclusion.