

# Navigation of Autonomous Vehicles Utilizing Hybrid Sensor Fusion Techniques

Dr. Ayaka Yamashita <sup>a</sup>

<sup>a</sup> *University of Tokyo, Japan*

---

## Abstract

Numerous sensor integration frameworks have been presented using multiple setups, sensor combinations, and fusion methodologies. Most studies focused on optimizing accuracy and efficiency, but research is still lacking in incorporating these methods into autonomous vehicles (AVs). Some fusion frameworks might perform exceptionally well in lab environments with abundant computing power; however, their use in integrated edge computing (EC) poses severe challenges due to cost and high computational needs. This study proposes a hybrid sensor fusion technique (HSFT) to enhance autonomous vehicle navigation systems based on deep Q networks (DQN). The approach features the control of flow for camera and LiDAR data and the synthesis of an independent algorithm for object identification based on sensor image data. Noise suppression has been implemented using a Federated Kalman Filter (FKF). Contrast enhancement has been achieved with the help of Homomorphic Filtering (HMF) and adaptive thresholding. Recognition and classification have been performed using the YOLO V7 model. The proposed work is evaluated based on requirements such as speed and accuracy rate.

**Keywords:** Directional Filters, Iris Image Indexing, Iris Recognition, Multi Scale and Multi Orientation Data Fusion.

---

## **Introduction**

Autonomous Vehicles (AV) employ sensors for two main purposes: geo-location, which helps the vehicle determine where on the road, and sensing the environment, which helps identify the surrounding objects and features. For the localization of the AV, the global positioning satellite system (GPSS) is used in conjunction with a Momentum Measuring Unit (MMU) and automobile odometry sensors (Li et al., 2019). Geo-location attainment is vital in knowing the position of the AV concerning the surroundings using spatial references. Subsequently, the environment is detected using LiDAR, cameras, and radar detectors. Understanding the environment is paramount for effective motion traversal within the region by the AV. Calculated geo-location and environmental perception outcomes are relayed to the AV route planner for actions predetermined by the route planner to be executed (Kolar et al., 2020). The actions are sent to AV movement control, which changes control signals to mechanical components to yield the appropriate actions.

This research article presents an HSFT architecture for driving autonomously. This modular architecture improves precision and combines the benefits of irregular state analysts like the Extended Kalman Filter with deep learning (DL) techniques such as DQN adapted to the sensors and their configuration (Butt et al., 2022). The AV sensor model applies this SF technique to form a complete understanding of the environment by providing information concerning the drivable road area and detecting or tracking objects within the scene. This technique was evaluated using an integrated in-vehicle device, and the outcomes were assessed against reference networks and legitimate data (Shahian Jahromi et al., 2019).

The AV environmental perception domain consists of AV sensors, analysis of the sensor information, and various fusion techniques. The sensors are categorized into 1) camera, 2) LiDAR, and 3) radar. Each grouping has a sub-group, SF, which utilizes state evaluators like Kalman filters (KF) and DL-based approaches, including DQN. Literature focuses on improving algorithmic accuracy; the AV algorithms did not focus on applications. More compact, flexible, and robust frameworks are required. Thus, SF control architecture balances the complexity of a fusion model and real-time operational scenarios while maximizing the perceived environment accuracy.

## **Proposed methodology**

The research investigates AV systems concerning the target sub-task of pre-processing and optimal route finding along with lane assignment. The KITTI database, which comprises image data: camera images, LiDAR images, meteorological images, and FoV images, is also analyzed. The main building blocks of the project's concept include sensors, EC servers, and the cloud. Cloud computing (CC) and EC servers offer relief in storage space and on-demand resources for computation. EC servers perform image processing and segmentation, yielding segmentation images. Redundant data debug systems help secure data and enhance server reliability while images are segmented in real-time.

This research aims to develop DQN models for multi-object recognition and surveillance at AVs (figure 1), which pose intricate challenges. This framework contains several core processes beginning with image refinement, which offers initial sensor data using dynamic thresholding, HMF-based contrast enhancement, and FKF-based noise reduction. These pre-processing steps are necessary to enhance the quality of the processed data. One of the filters for this phase, the Weighted Mean Filter (WMF), can obtain efficient image descriptions and reduce noise. WMF achieves improved filtering results by allocating increased weights to pixels likely to contain essential information, resulting in cleaner pictures than traditional filtering techniques.

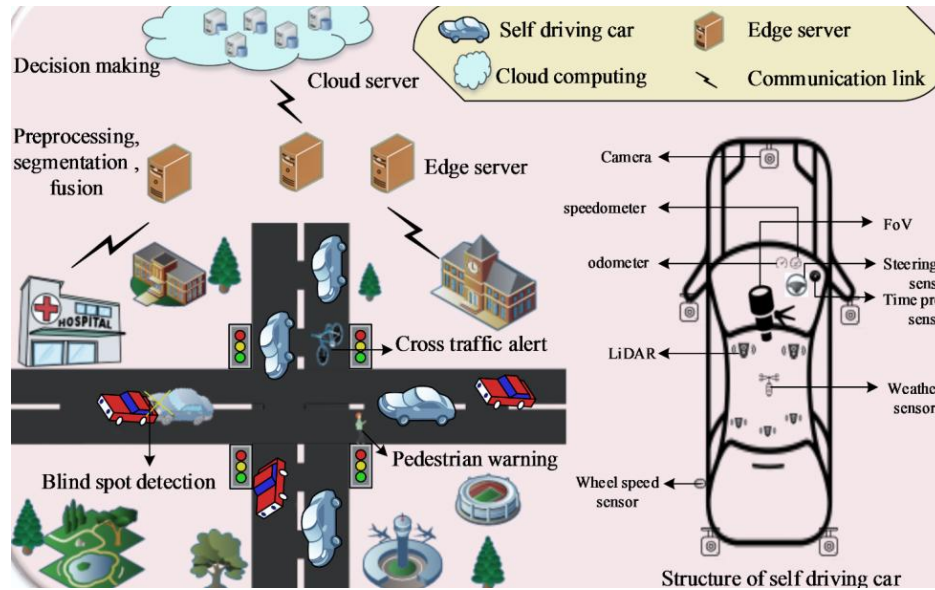


**Fig. 1: The overall architecture of the DQN is used in the navigation of AV.**

After this enhancement, the multi-segmentation step divides images into areas of distinction FoV and positions it to detect smaller or more difficult objects in their head streams. Segmentation is performed using fast MobileNet V1 of the respective networks. MobileNet V1 enables immediate segmentation in AV with a compact convolutional architecture with dense links, ensuring effective feature preservation and reduced processing overhead. HSFT is used to augment the segmented pictures from this phase. This procedure amalgamates information from many sensors, such as cameras and LiDAR, to provide a thorough fused depiction of the AV's environment. The YOLO V7 framework is used for flexible and exact real-time scheduling of routes, facilitating the identification and surveillance of objects for correct lane and route allocation.

#### **Hybrid Sensor Fusion Techniques pipeline in AV**

The three primary vision sensors utilized by AV possess distinct advantages and disadvantages. Consequently, their data must be integrated to get the most coherent understanding of the surroundings. This work presented SF for this purpose. The HSFT has two concurrent components, as seen in Fig. 2. Each section employs a specific arrangement of sensors and an SF mechanism optimally aligned with the respective fusion job.

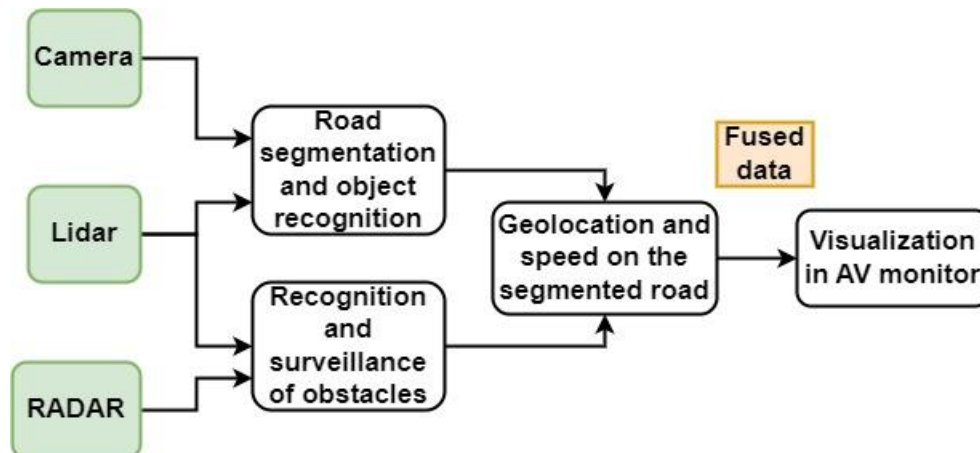


**Fig. 2: Hybrid Sensor Fusion Techniques pipeline in AV**

The first section addresses high-quality problems, including item categorization, localization, and conceptual road segmenting using camera (vision) and LiDAR. The unprocessed video frames from the camera and the depth data

from the LiDAR are integrated before being delivered to a DQN responsible for object categorization and road segmenting operations.

Integrating the camera and LiDAR with a DQN framework provides the optimal sensor-based coupling for categorization and segmenting tasks. This study emphasizes real-time efficiency and precision. Consequently, rather than aiming to divide the roadway scene into several tiny segments, this work conducted segmentation for two categories: the wide region (drivable area) of the road and the non-drivable area for the AV. A rapid and effective network such as YOLO V3 has been employed to identify impediments inside the segmented roadway. In AV driving, understanding the drivable area and categories of obstacles on the roadway is crucial for route development and choice-making.

The second section addresses the detection of objects and the monitoring of their statuses via LiDAR and radar detectors. The LiDAR information and radar images are handled and integrated at the component level. Analyzing LiDAR and radar images will provide collections of impediments on the roadway inside the ROI and their respective statuses. Late merging denotes the integration of collected sensor information at the component level. The Kalman Filter is the preferred approach for state estimation; nonetheless, it operates under the presumption of linear movement models. Due to the irregular movement of impediments such as vehicles, a modified variant of the Kalman Filter has been employed, namely the FKF. Understanding and monitoring the conditions of road impediments facilitates the anticipation and consideration of their conduct in the AV's route planning and choice-making processes. Ultimately, each fused result was superimposed and displayed on the AV monitor.

## Results

The HSFT using the DQN model in AV has been constructed and simulated with MATLAB R2019a (mathworks.com). To achieve excellent simulation outcomes, we optimized the system setups for both hardware and software.

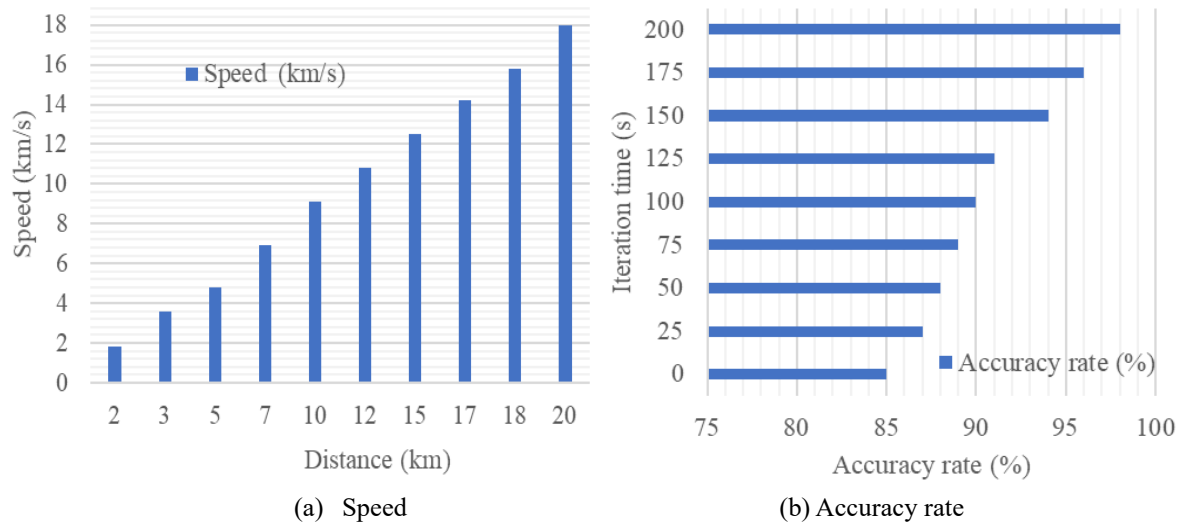


**Fig. 3: Performance analysis of the proposed HSFT using the DQN model in AV**

Fig. 3(a) depicts the speed (km/s) versus distance (km) covered by the AV using HSFT. At a distance of 2 km, the AV is moving with a speed of 1.8km/s, and this accelerates until it reaches a maximum of 18 km/s at 20 km. Such a pattern is observed that the approximately HSFT incorporating the DQN model improves the speed performance as the distance increases. The smooth scaling of speed demonstrates optimum control efficiency and decision-making system performance, which indicates that the model has a high potential for performance-enhanced navigation in autonomous vehicles.

The proposed model (Fig. 3(b)) achieved a heightened value of 0.98, obtaining maximum accuracy at 200 seconds for HSFT using DQN in AV. This marks the predicted strength of the learning model proposed within the framework of AV systems utilizing deep learning technologies. The model was operating at an 85% accuracy rate at zero iterations. Throughout iteration time, consistency was achieved, indicating that improvement to 98% accuracy was achieved with

further training. This outcome fully aligns with expectations, confirming that integrating HSFT with the DQN structure improves prediction accuracy and performance under sustained training scenarios.

## **Conclusion**

This study proposes a method for the navigation of Autonomous Vehicles (AVs) using Deep Q Networks (DQN) based on Hybrid Sensor Fusion Techniques (HSFT). The system proposed in this work integrates pathway management for the camera and LiDAR data and builds an independent solution for object detection by processing sensor images. A Federated Kalman Filter (FKF) has been used for noise suppression. Homomorphic Filtering (HMF) performed contrast improvement, and some filters were also applied using adaptive thresholding. For detection and classification purposes, the model used is YOLO V7. The proposed work is evaluated using measures of speed and accuracy rate. At a distance of 2 km, the AV moves at a speed of 1.8km/s, and this accelerates until it reaches a maximum of 18 km/s at 20 km.

## **References**

- Li, X., Chen, W., Chan, C., Li, B., & Song, X. (2019). Multi-sensor fusion methodology for enhanced land vehicle positioning. *Information Fusion*, 46, 51-62.
- Kolar, P., Benavidez, P., & Jamshidi, M. (2020). Survey of datafusion techniques for laser and vision based sensor integration for autonomous navigation. *Sensors*, 20(8), 2180.
- Butt, F. A., Chattha, J. N., Ahmad, J., Zia, M. U., Rizwan, M., & Naqvi, I. H. (2022). On the integration of enabling wireless technologies and sensor fusion for next-generation connected and autonomous vehicles. *IEEE Access*, 10, 14643-14668.
- Shahian Jahromi, B., Tulabandhula, T., & Cetin, S. (2019). Real-time hybrid multi-sensor fusion framework for perception in autonomous vehicles. *Sensors*, 19(20), 4357.
- Yeong, D. J., Velasco-Hernandez, G., Barry, J., & Walsh, J. (2021). Sensor and sensor fusion technology in autonomous vehicles: A review. *Sensors*, 21(6), 2140.
- Liao, Y., Xie, J., & Geiger, A. (2022). Kitti-360: A novel dataset and benchmarks for urban scene understanding in 2d and 3d. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(3), 3292-3310.
- <https://www.mathworks.com/products.html>