

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/365202635>

Artificial intelligence-empowered vision-based self driver assistance system for internet of autonomous vehicles

Article in *Transactions on Emerging Telecommunications Technologies* · November 2022

DOI: 10.1002/ett.4683

CITATIONS

7

READS

160

2 authors:



Oshin Rawlley

Birla Institute of Technology and Science Pilani

14 PUBLICATIONS 77 CITATIONS

[SEE PROFILE](#)



Shashank Gupta

Birla Institute of Technology and Science Pilani

80 PUBLICATIONS 2,317 CITATIONS

[SEE PROFILE](#)

Artificial intelligence-empowered vision-based self driver assistance system for internet of autonomous vehicles

Oshin Rawlley | Shashank Gupta

Department of Computer Science and Information Systems, Birla Institute of Technology and Science, Pilani, Rajasthan, India

Correspondence

Shashank Gupta, Department of Computer Science and Information Systems, Birla Institute of Technology and Science, Pilani, Rajasthan, India.
Email:
shashank.gupta@pilani.bits-pilani.ac.in

Funding information

CHANAKYA Fellowships of IITI DRISHTI CPS Foundation under the National Mission on Interdisciplinary Cyber Physical System (NM-ICPS) of Department of Science and Technology, Government of India.

Abstract

Artificial intelligence (AI) and edge computing have truly advanced in vehicular networks encouraging assessment of real-time traffic conditions using kinetic information of autonomous vehicles with the help of road side units (RSUs). However, regardless of numerous improvements in sensor fusion technologies the existing vision/LIDAR-based systems have found severe difficulties during perception on roads. In addition, the inter-vehicular communications are hampered due to inefficient RSU placement techniques causing high-latency issues during transmission of messages. Therefore, this article presents an AI-driven vision-based self driver assistance system (VSDAS) comprising a joint RSU deployment mechanism that utilizes enhanced memetic architecture-based optimal RSU placement (MARp) method and an object detection model that implements an improved Haar-cascade object detection algorithm for speedy identification of object. We have designed two varieties of genetic algorithm (GA) to solve optimal placement problem of RSUs: genetic architecture-based with random restart hill climbing (GARRH) and the other is MARp for efficient placement of RSUs. After our experimental results, we see that the MARp algorithm shows best possible RSU locations over different generations achieving significantly better fitness scores than the GAHR and GA ascribing to its local search process. In addition, Haar-cascade achieves better mean average precision than local binary pattern and histogram of oriented gradients by selecting key frames. The experimental outcomes of our model reveals that the proposed enhanced memetic algorithm reduces the transmission delay to a greater extent. Additionally, the outcomes of our investigations on two public datasets (KITTI and Panasonic) showed that our improved algorithm clearly enhances the object detection performance.

1 | INTRODUCTION

Vehicular ad hoc networks (VANETs) are recognized as self-organizing networks for disseminating information between AVs using wireless mode. VANETs transfer data among ego-vehicles in multiple hops. Moreover, the VANETs connected with the internet realizes the Internet of Autonomous Vehicles (IoAV) which refers to the inter-communication of the intelligent systems such as RSUs, other embedded processing units in the vehicles and various cyber-physical systems (such as sensors, LIDAR, RADAR, mobile devices, etc.) operating on advanced networking infrastructure.¹ This system

of IoAV enables exchange of information facilitating Vehicle-to-Anything (V2X) communication to contribute towards a safer and more efficient transportation system.

However, the very basic nature of the transportation system possesses great uncertainties. The changing dynamics of the fleeting vehicles with respect to their relative position and time causes the communication topology to alter accordingly. When the ego-vehicles breach the effective communication range between the vehicles, the multi-hop strategy for data communication becomes ineffective. This makes VANETs unsuitable and unstable for vehicular communications.

1.1 | Necessity for road side units (RSUs)

To facilitate the inter-communication of the vehicles, nowadays the autonomous vehicle (AV) cyber-physical systems are empowered by the capability to process the data locally and communicate with others timely through V2X communication. However, due to limited caching and computing resources, effective data utilization becomes a challenging task for vehicles and RSUs. Therefore, FC has become a promising paradigm to cater to these issues. FC enables time-critical applications such as autonomous driving, facilitates computation-intensive tasks such as local processing in AVs/RSUs and so forth. FC has also facilitated the development of smart vision-based driver support scheme and enhanced network capabilities facilitating IoAV technology using RSU for latency-sensitive areas in smart devices which are energy-constrained. Figure 1 depicts a generic framework of a fog-driven computing model towards IoAV. The RSU module depicts the RSU deployment on road intersections covering a radius of 200 m. The secluded areas are equipped with edge servers for better connectivity. The accurate object detection is aided by the knowledge base provided by the RSUs.² Since there is a constraint of RSUs having limited caching power, we need better strategies to install the RSUs in a way that is well within the communication range of the vehicles and is effective in low-density regions also. Owing to the decision-making competence of the RSUs for better roadside services it is required to improve the efficiency and quality of vehicle-to-infrastructure (V2I) communication and vehicle-to-vehicle (V2V) communication on the road. For this purpose, a greater number of vested interests should be encouraged to invest in the area of RSUs to improve upon the road network perception. Although, the investments in RSU deployment are crucial for better vehicular network connectivity, this is no hidden fact that RSUs can only serve a small number of vehicles.

Moreover, in comparison with the cellular networks, the RSU coverage is relatively restricted. Therefore, it is difficult to provide connectivity for continuous AV roaming. RSUs in IoAV system are primary in considering efficiency of

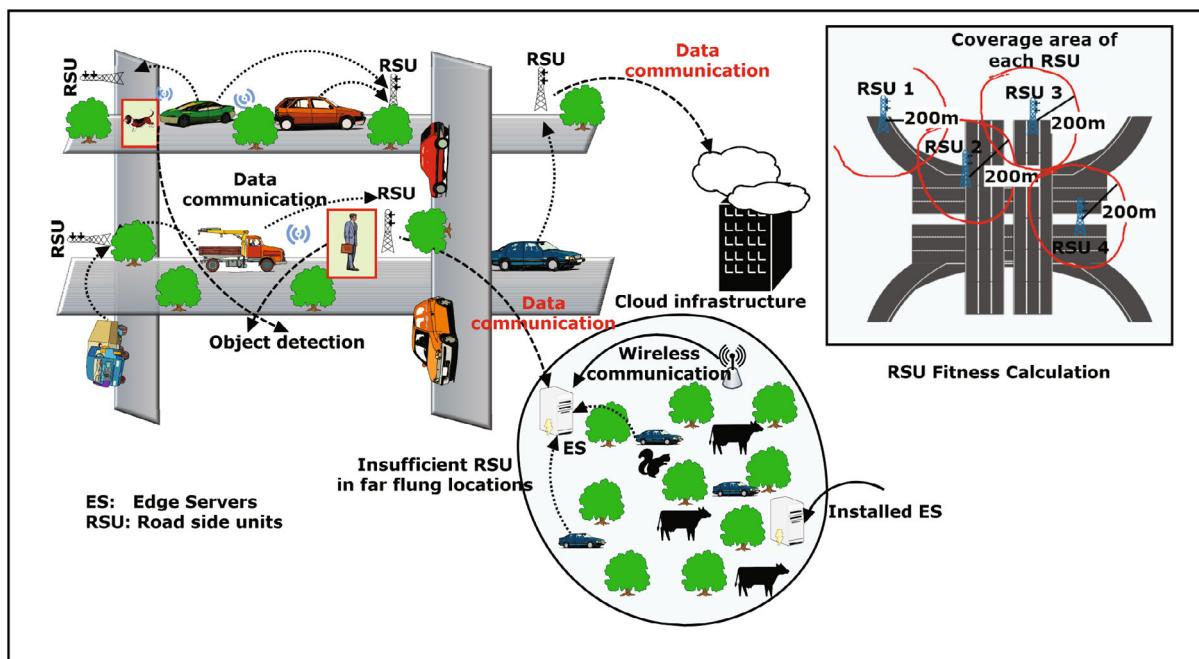


FIGURE 1 A fog-based computing model towards IoAV

network and provides other functions of safety. RSUs are infrastructure nodes that are critical in (a) broadcasting useful information to vehicles and (b) sending messages to intended receivers via internet connection.³ RSUs have acquired different roles namely, location servers, safety managers, content disseminators, proxies for services and so forth. Furthermore, RSUs are required to provision distributed and cooperative applications of IoAV which the vehicles can utilize to process and share the multimedia content. Therefore, we need to determine that with a limited budget what is the best way to choose the priority intersections of roads for RSU positioning in a broad spectrum of road pathways which will maximize the efficacy of RSUs. In addition, an algorithm for efficient placement of vehicular infrastructure to facilitate a seamless inter-vehicle and V2X interaction is required.

The other primary aspect in AVs is object localization and detection which monitors the environment and tracks several happenings in the surveillance system. Object detection localizes the target to accurately estimate the desired position of the AV and then detect it with a label. This process captures the ambient perception for smooth navigation and safe automated driving for appropriate knowledge of surroundings. Subcategories such as pedestrians, car, cycle and so forth are identified by selecting the informative region from the image by multi-scale slicing.

1.2 | Object detection techniques

Earlier methods of video surveillance such as scene segmentation have been playing a significant role in making the OD process faster. Enormous amount of video data simplifies the monitoring of the vehicles but at the same time causes data management issues.² If the long video streams can be converted into a set of relevant frames for focused region of interest, it will be very convenient for the object detection algorithms to process them faster and detect objects faster on roads. However, present deep-learning techniques have limitations of both memory and accuracy when processed on edge devices. To overcome challenges of exponential and redundant data, there is a need for efficient algorithms to process selective frames of the video chunk for further processing of the video scenes. Therefore, we need to ascertain that how to choose the important video frames from the long video sequence and apply appropriate object detection algorithm for the method to perform better in detecting objects.

Hence, in this article, we propose a resilient and decentralized joint RSU deployment and an object detection model used in vision-based self driver assistance system (VSDAS) for calculation of optimal RSU coverage area along with enhanced AV localization and detection. The system will employ the intelligence of placing the RSUs optimally, on road intersections in a particular area by utilizing machine-learning based memetic framework (MF) for achieving maximal coverage in low installation cost. The installed RSUs will help in on-the-spot processing of this data to provide insights into traffic conditions. Thus, RSUs will also play a vital role in realizing object detection in VSDAS model. It will allow V2X communications to occur in an integrated set of novel networking paradigms allowing VSDAS to be in spatial-temporal coherence with the RSUs. Additionally, VSDAS model will employ a machine-learning based Haar-cascade algorithm to accurately localize and detect the objects in varied light conditions. Therefore, motivated by these findings our main contributions in this article can be summarized as follows:-

- We propose an efficient and ultra-reliable memetic framework-based optimal RSU deployment (MARp) deployment mechanism for minimizing latency delay on the network. This optimization model will generate a RSU deployment strategy for optimal RSU deployment location. This model will be simulated on different scenarios using NS2 simulation platform to visualize different scenarios and design efficient scheme for installing RSUs in a particular geographical location.
- We also propose an improved low-latency Haar-cascade algorithm for AVs to increase detection performance and interpretability of roads. We will implement a relevant key frame detection technique for selecting key frames and further conduct object detection.
- Moreover, we will use simulation of urban mobility (SUMO), an open-source traffic simulation platform to evaluate various traffic management issues and develop various strategies to implement intelligent transportation system (ITS).
- We have compared our scheme with the existing state-of-the-art to prove the efficiency and feasibility of our proposed methods for RSU placement and OD by taking into consideration some performance parameters like response time, fitness score versus RSU count, mean average precision (mAP), precision-recall and so forth, respectively.

The organization of this article is structured as follows: Section 2 delineates the approach of our proposed work, Section 3 elucidates the implementation work and details the results obtained from it, Section 4 elaborates the evaluation metrics and performance parameters entailed in judging the scalability and practicability of the adopted method. Lastly, Section 5 summarizes this article.

2 | RELATED WORK

In this section we elaborate the two types of network connectivity of vehicles and the need for deploying more roadside infrastructure to facilitate a seamless communication between vehicles. We also explain the issues related to RSU installation at the priority intersections. Further, we elucidate the background for edge-based vision systems for low-latency object detection. The limitations related to the existing centralized ML systems used in object detection frameworks are addressed in this section. We strongly feel that both the aspects are required for a competent performance of the proposed VSDAS model to enhance the vehicular connectivity and its accurate localization.

2.1 | Intra-vehicle connectivity

The first type of communication which forms the basis of AV driving realizes the message communication by different electronic control units (ECU), sensors, and actuators. Intra-vehicle communication actualizes cutting-edge functionalities such as scene perception, path planning, fault diagnosis, and motion control in a unified system. There are two ways of coordinating control and exchanging information among diversified ECUs in present self-driving car on-board network, one is wired connection such as media oriented system transport (MOST), local interconnect network (LIN), controller area network (CAN), FlexRay, and other is wireless connection such as wireless fidelity (WiFi), Bluetooth 5.0, ZigBee, ultra-wideband (UWB). The existing technologies lack an advanced intra-vehicle connectivity with upgradable system compatibility and higher bandwidth support.

2.2 | Inter-vehicle connectivity

Owing to the dependence of AVs on environmental factors the inter-vehicle communication is of prime importance. This will further benefit in the advanced tasks such as planning, environment perception, and interaction between vehicular infrastructures of autonomous driving. Inter-vehicle communication requires road-side facilities having computing and storage capability and promising wireless communication ability. Moreover, these road-side facilities help in acquiring plentiful of information about the roads through lots of on-board sensors and information interaction among automotive components of AV domain. Since the functioning of AVs partly or wholly relies on a sensor generated real-time data and inter-vehicle communications there is a leading requirement for more integrated and high-speed processing and communication technologies. DSRC and LTE-V are the existing state-of-the art standards in AV networking community. DSRC has been recognized best for low delay communications and short distance transmissions however, certain limitations can be witnessed in volatile scenarios of AV driving and therefore RSUs are needed to be established.

The inter-vehicle communication and intra-vehicle communication forms an extensible connection which acts as a backbone structure for the AV communication system and therefore to have a low-latency communication system external vehicular infrastructures such RSUs are required for a 2-way transmission of vehicular messages. To determine an economical RSU deployment technique for low-latency communication has been a key problem in vehicular communications. In earlier works, the authors proposed a uniform distribution strategy for RSU placement in which RSUs are placed away from each other at a fixed distance. Though the strategy appears to be straight and simple, but it leads to recurrent disconnections. A suitable alternative to this problem is that, RSUs can be positioned in hot spot areas in a way that AVs passing by that particular area can communicate packets to the RSUs. Moreover, load balancing of the communication can be done by deploying RSUs in hot spot zones further alleviating the contention issue. Pan et al⁴ suggests a technique for accessing a single gateway meant for a set of access points, analogous to the placement of base stations in cellular systems. Nevertheless, this technique did not consider the interference constraint and the topology of the road. Therefore, it was rendered unsuited for real-time systems. Gao et al⁵ suggested a method DynLim, which applied a function of profit density for maximizing the utility of RSUs. On the other hand, Liang et al⁶ provided a mathematical solution

in the form of integer linear programming (ILP) model for reducing the delay in highway conditions. Ni et al⁷ resolved the RSU deployment concern as the shortest path problem with the help of the ILP model. The earlier deployment techniques failed to satisfy the requirement of optimal coverage area. Also, delay-critical data requiring high computing demand a novel orientation in AV owing to extreme mobility and fast changing dynamics of network topology for better deployment of RSUs. Hence, to conduct a massive installation of RSUs in a region, it necessitates an optimal RSU deployment technique that covers a well-defined coverage area. In addition, it should also decrease the delay in communication and the technique should facilitate the expansion of the span of the RSU coverage between IoV devices. Presently, it is economically infeasible to install adequate quantity of RSUs and also, it lacks widespread coverage area mechanism. Therefore, we intend to propose an algorithm which accommodates the needs and caters to the unattended aspects of the V2I and V2V connectivity, routing, and transmission delay.

2.3 | Enhancing perception of the AV systems

Various advanced sensors such as camera, LIDAR, radar are mounted on the AVs for better perception and decision-making. However, these sensors have certain limitations such as unreliability, infeasibility, inefficiency and so forth, which leads to lot of fatal accidents. Further, we need robust ML/DL oriented object detection algorithms to be deployed on the AVs for immediate detection of obstacles on road. In the AV systems, the majority of the road data about traffic conditions is collected from surveillance devices and then is preprocessed. However, to process such highly time-critical data with constrained computation resources bears more challenges. In this section we review existing state-of-the art vision-based technologies for the IoAV system namely, object detection algorithm for VSDAS.

The earlier researchers namely, Blaschke et al⁸ utilized a semiautomatic pattern matching method. Ranft and Stiller⁹ used vision techniques for numerous tasks of intelligent AVs, such as image isolation and mapping, understanding the automotive environment, and categorizing vehicular objects. Mukhtar et al⁴ examined 2D object detection techniques for ADAS concentrating on attributes such as motion and appearance by leveraging a conventional pipeline. Notably, the existing technologies have failed in achieving the ambient perception for automated object detection required in AVs.¹⁰ Moreover, the correct object detection is hampered by the higher compression rate of videos, which compromises the accuracy. Besides that, the redundant frames of the video take greater bandwidth to process, which increases the error rate. Further, the existing RADAR and LIDAR techniques have proven to be accurate, however, the on-board units (OBUs) on AVs embedded with object detection ML/DL frameworks attests to be more accurate.¹¹ Therefore, an improved object detection algorithm is required for achieving high-precision localization in self-driving assisted navigation for which significant key frames will be determined with less error rate. The existing state-of-the-art have explored the static cases and have significantly achieved better results, but have not taken into consideration the dynamic scenarios. In a vision-based driver support system, the proliferation of enormous redundant data causes latency issues and due to high compressibility of videos, the accuracy of object detection gets compromised. Considering such gaps, the evolving real-time object detection methods and infrastructure planning such as placement of RSUs will introduce a new opportunity for environment perception. High-precision localization, seamless communication between vehicles and other RSUs for better service continuity and enhanced navigation services is provided to AVs under the diverse environments.¹² The existing state-of-the-art can be seen in Table 1 evaluated over different parameters including the benefits of our proposed methods of RSU placement and object detection framework in the road network. The Table 1 discusses various existing algorithms and their techniques which have been used for optimal RSU deployment and scalable object detection techniques based on different parameters. The existing state-of-the-art has not been able to establish strong grounds for an optimal RSU deployments on different topologies, has not covered important research factors. Moreover the earlier object detection methods for better detection of AVs on roads has suffered many drawbacks such as high latency, unable to predict dynamic traffic flow and so forth. We have therefore highlighted our proposed methods for RSU placement and enhanced object detection technique based on the same parameters and claim that they have performed well in the results and evaluation section.

Despite of the various intra-vehicle and inter-vehicle communication the we feel that there is a need for low-latency communication between vehicles as much of the available spectrum of bandwidth remains unutilized. The future ITS will be requiring a greater attention on V2X applications for achieving direct communication with RSUs to facilitate immediate transmission of messages as and when vehicles move in-and-out of a connected zone. The optimal deployment of RSUs leads to high performance communication of vehicles with itself and with other vehicles which forms an efficient ITS. Moreover, with better connectivity of AVs AUTONOMOUS driving also require multiple perceptions for better navigation

TABLE 1 Comparative analysis of existing methods of RSU placement and object detection on road networks

Citation	Existing algorithms	Technique used	Research factors	Mobility trace	Simulators used	Drawbacks of related work
13	Capacity maximization strategy	ILP	Deployment budget	1250 m by 150 highways	NS-2, VanetMobiSim,	Completely analytic, no algorithm deployment and simulations used for verification
14	Randomized	Mathematics study	Time and connectivity probability	Stretch of 100 km highway	Specific	Lacked better strategy for RSU allocation
5	OptDynLim	D1RD	RSUs quantity	None taken	MATLAB	No testing was done on any real topology area. Do not have any mobility traces and does not consider any network performance metrics.
15	Voronoi diagram	Voronoi graph	Levels of QoS	Nashville, TN, USA	NS-2, SUMO	Have not considered traffic density and junction priority, lack feasibility of selected deployment locations, did not account for barriers like buildings and rivers.
16	Genetic algorithm	Geometric model	Transmission time for emergency communications	Madrid, Valencia	SUMO	Installation cost and information about accident are not ignored in this design
17	Binary differential evolution	0-1 variation Knapsack	Cost-restricted and delay-bounded	Zhengzhou, China	SUMO	Lacks consideration for QoS-guaranteed RSU placement in VANETs
18	Mathematical study	Clustering model	Transmission of alert messages delay bound	No real topology taken	Specific	Algorithm is tested on only on a single topology
19	Greedy and hill climbing	Gamma deployment approach	Threshold of contact time	Cologne, Germany	SUMO	Did not consider multi-hop communications for vehicular node connectivity and V2V communication
20	Genetic	Aggregation scheme	Cost-restricted	Brunswick, Germany	NS-2, VISSIM	Focusses more on information aggregation rather than data dissemination
21	Greedy, dynamic and hybrid	Intersection priority	Overlapping area	Seoul, South Korea	SUMO, NS-2	Did not consider realistic urban areas
22	Genetic and D-trimming	Optimal RSU distribution planer (ORDP)	RSU coverage and installation budget	Tamil Nadu, India	VISSIM	QoS parameters such as delay, speed, data transmission is not taken.
23	Artificial neural network (ANN)	Mathematical modeling of traffic	Speed, volume, average occupancy	NA	Fuzzy rule-based system	Do not consider the distribution of vehicles, not capable of identifying key features

(Continues)

TABLE 1 (Continued)

Citation	Existing algorithms	Technique used	Research factors	Mobility trace	Simulators used	Drawbacks of related work
23	Deep neural network (DNN)	Acoustic modeling techniques	Speed, flow, occupancy	NA	Pytorch, Tensorflow	Predicts short traffic flow
23	HOG	DWT (discrete wavelet transform) technique with HOG method, FPGA-based method and so forth	Complex occlusions and Pose variations	Roadways	ML/DL Platforms	Identifies linear distributions
Proposed VSDAS	MARP algorithm for optimal RSU deployment and Haar cascade low latency object detection	Simple hill climbing, AdaBoost training	Coverage range of RSUs and Number of RSUs to be installed High ratio of positive to negative samples for object detection	Murilipura and Connaught place (CP), India, KITTI dataset	SUMO, NS-2, ML/DL platforms	-

on roads. After reviewing the existing state-of-the-art we feel that there is a lack of accurate and prompt action model for detecting objects on road. Therefore, we aim to achieve faster object detection on account of ML architecture combined with key-frame selection method.

3 | AI-EMPOWERED VSDAS MODEL

The proposed VSDAS system employs an optimum RSU deployment method and an enhanced object detection technique in an IoAV network. It ascertains the optimal positions of roadside units on any given topography using a MARP algorithm. Unlike the conventional genetic algorithm, our proposed memetic framework involves a local search which makes it unique for implementing this intelligence. We present two variants of the hybrid genetic algorithm for RSU placement.

On the other hand, Haar-cascade object detection algorithm is proposed which works on the principal of choosing key frames from the video. It is intended to reduce the latency on the edge nodes in IoAV system. This method can considerably show improvements in reducing the latency at the cost of low detection precision. In this article, we conduct a comparative analysis of three methods namely, histogram of oriented gradients (HOG), local binary pattern (LBP), Haar-like traits over a public dataset.²⁴ In our proposed method, we suggest a weight coefficient and pixel difference algorithm to filter out the redundant frames and obtain relevant frames for detecting an object. Subsequently, an enhanced mechanism namely, Haar-like traits classification method is utilized for scalable object detection.²⁵

3.1 | Problem description

We consider a vehicular network over two locations. The road areas are divided into various segments and consists of multiple lanes. The distance of every segment is selected within the transmission range of each RSU. We aim to install RSUs in a way such that every intersection should have at most one RSU in the center of the segment. We consider RSUs to be high-end processing and communicating devices and are the only gateways having backhaul access to the Internet. Therefore, the positioning of RSUs is very beneficial for the ITS to have seamless connectivity in a cost-effective manner. Since the RSUs are very expensive and requires a tedious installation, we intend to propose an effective placement model for achieving end-to-end throughput. In addition, to aid low-latency OD we extract the key-frames from the video sequences to have a specific scene understanding. This will help the OD algorithms focus on important region of interest and perform their detection process. RSUs in conjunction with this can give real-time decisions such as collision alert, wrong lane departure information, diversions and so forth.

These functionalities are implemented across the fog nodes and are accessible via a web-based dashboard/controller. The cloud service providers, meanwhile, stores the encrypted data which is also accessible to the users. Lastly, some extensive simulations confirm the superiority of the suggested algorithm over the other obsolete schemes.

System model of RSU and object detection: The optimal placement of RSUs enhance the network efficiency and facilitates judicious utilization of network resources. This has been considered as a significant task in an urban vehicular environment. It also expands the coverage region among several IoAV nodes via optimal RSU placement. Moreover, the object detection also plays a vital role in accurate localization of objects and in lowering the latency. Figure 2 depicts the formal overview of the system model of our respective proposed scheme. In the same context, we present our modules as follows:

The roadside unit placement module: This module deploys the roadside units in the best possible locations by employing MARP algorithm resulting in an expansion of the coverage area of each RSU, increased network efficiency, reduced message delay, and prevents signal degradation. This makes it apt for an efficient and quick delivery of critical messages in an IoAV.

The object detection module: This module computes the importance coefficient of the objects and finds out the similar frames. The redundant frames will be removed to reduce the latency. Finally, object detection will be efficiently carried out using an improved Haar-cascade classifier.

3.2 | The roadside unit placement module and MARP mechanism

Since, a small number of RSUs are required to be installed on the priority road intersections, we aim to enable direct delivery of messages when a vehicle enters the transmission range of the RSU by efficient deployment of the communication

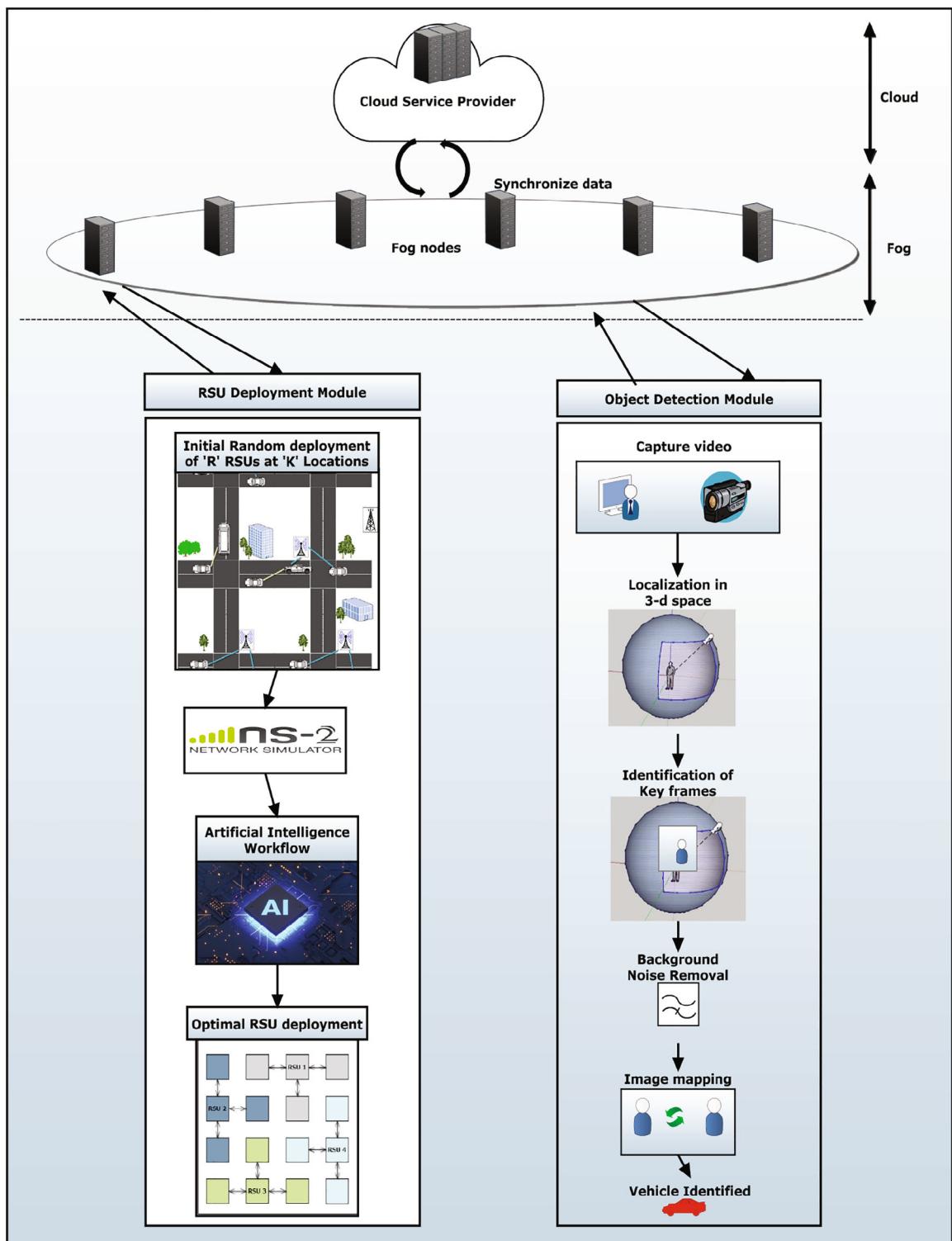


FIGURE 2 An overview of the system architecture

infrastructure. The main aim of this module is to establish a cluster of RSU in an IoAV ecosystem for any topography such that it delivers low-latency communication and covers larger area under it. Nevertheless, the costly installation and maintenance of these RSUs constraints their deployment. The existing state-of-the-art methods of deploying RSUs were unable to realize network efficiency and failed to be implemented under diverse traffic situations. Considering the above limitations, we propose the MARP algorithm which is based on the evolutionary concept in artificial intelligence.²⁶

To illustrate the RSU placement problem, let us consider a 4×4 road network shown in Figure 2 having 16 potential junctures. Every road juncture is a probable location where a RSU can be installed. If these infrastructure nodes are installed at the positions highlighted in green, then all the sites are covered with less overlapping within the RSU coverage area.

Motivated by Darwinian evolution, the search and optimization problems are solved by evolutionary algorithms (EAs) which apply natural selection. A criteria-specific selection procedure is implemented on the population to determine the fit individuals suitable for survival. The best fit individuals are passed to the next generation whereas, less fit generation is dropped from the population. This procedure is repeated till the termination condition is arrived.²⁷

GA builds a relationship and a respective set of solutions among the set of individuals in a natural population. These set of solutions are called as *phenotype*. This is achieved by *genotype* representing the information of every solution in a string. The evolution of genotypes happens over multiple generations using the genetic operators namely, parent selection crossover, mutation, and replacement. Yet, in a complex combinatorial search space the classic genetic algorithm (GA) fails to produce tuned results in a combinatorial search space.²⁸ While, if we integrate local search process in hybrid GA, it substantially enhances the search process efficiency which is termed as memetic algorithm.

The MARP algorithm uses a memetic framework which is an advanced version of conventional GA.²⁹ In memetic framework, this local search approach is implemented over every GA created individual for improving the fitness of the individual. It will facilitate faster generation of results as compared to GAs. The fact that if we explore the solution space in combination with local search space exploitation, it reduces premature convergence to local optima point thereby making them suitable for optimization problems.

For the RSU placement problem, the phenotype acts as a probable location where RSUs can be deployed. We represent the genotypes with a set of integers representative of RSU locations. According to our objective, we also take care that too many RSUs are not installed at one location to balance the installation costs. We designed two different techniques of hybrid GA which uses hill climbing meta-heuristic. Algorithm 1 termed as genetic architecture-based with random restart hill climbing (GARRH) utilizes the random restart method in the local search procedure, whereas the Algorithm 2 that is, MARP algorithm implements a simple hill climbing method. The input to these algorithms includes parameters such as: set of k possible sites, quantity of RSUs R to be deployed, generation number G_N , population size n . Another input parameter which is taken, that is, a target area map of different intersections and streets and the different traces of mobility which simulate the real-time traffic. These realistic maps have been produced through OSM and SUMO.

3.3 | Description of the algorithm

The Algorithm 1 is the GARRH algorithms that outputs set of R locations. The stepwise explanation starts by steps 1 to 4 declaring the initial solution space which consist of probable sites, likelihood, respecting indices and arbitrary mutation and crossover numbers respectively. Further in step 5 fitness score variable is introduced. The step 7 specifies the new generation variables and their off springs produced in each iteration. The subsequent step assesses the individuals on the basis of their fitness function.

Further, we sort the generated population in decreasing order based on fitness score. We would like to highlight that every problem has its own unique fitness function which is an indicator for the proximal solution. The result of this function is a numeral which should be conditionally either minimized or maximized reliant on the problem. Our proposed MARP method considers RSU coverage area Ψ as an objective function. The RSU coverage ratio Ψ finds the number of intersections under one RSU as these are considered potential RSU deployments. Let N_i and N_j signify the number of intersections under the communication range of RSUs R_i and R_j , respectively. Therefore, the optimum intersections covered under R_i and R_j are as: $R_i + R_j - (R_i \cap R_j)$ and the coverage ratio Ψ of an RSU in a given topography is expressed as:

$$\Psi = \sum_{i=1}^{t-1} R_i + R_j - (R_i \cap R_j), \quad (1)$$

where t denotes number of RSUs.

Algorithm 1. GARRH for optimal RSU deployment (GARRH algorithm)

Input: Probable Locations $C = C_1, C_2, C_3, \dots, C_k$, R : RSUs Quantity, size of the population n , Number of generations G_N .

Output: Probable locations R of RSU.

procedure Generate_RSU_sites:

```

1: Generate initial population P of n size where an entity has R characteristics
2:  $P_{sel}, P_{mut} \leftarrow$  Selection probability and mutation probability, respectively
3:  $Ind_{mut}, Ind_{cross} \leftarrow$  Mutation index and crossover index, respectively
4:  $Rand_m, Rand_c \leftarrow$  Random number produced between 0-1
5:  $\epsilon_s \leftarrow$  Score of the fitness
6: while not  $G$  do
7:    $\beta_n, \theta_n \leftarrow$  New population, blank list for every generation child
8:   calculate  $\epsilon_s$  for P and sort P on  $\epsilon_s$ 
9:   for  $i: 0$  to  $n * P_{sel}$  do
10:     $\beta_{ni} \leftarrow P_i$ 
11:   end for
12:   for  $i: n * P_{sel} + 1$  to  $n$  do
13:      $E_{c1}, E_{c2} \leftarrow$  Selection ( $\beta_n$ )
14:      $\phi_{ni} \leftarrow$  Crossover ( $E_{c1}, E_{c2}, I_c$ )
15:     if  $Rand_m < P_{mut}$  then
16:        $\theta_{ni}' \leftarrow$  Mutate (,  $\theta_{ni}, Ind_{mut}$ )
17:       Replace  $\theta_{ni}$ , by  $\theta_{ni}'$ 
18:     end if
19:   end for
20:    $\theta_n \leftarrow$  empty list (local search process chosen chromosomes).
21:   for  $i: 0$  to length ( $\theta_n$ ) - 1 do
22:      $\phi_{li} \leftarrow$  Local_search_replacement ( $\theta_n$ )
23:   end for
24:   Combine the new chromosomes obtained in  $\theta_l$  with the chromosomes in  $\beta_n$ 
25:    $P \leftarrow \beta_n$ 
26: end while
27: return  $P_0$ 
28: end procedure
29: procedure Local_search_replacement ( $\phi_n$ ):
30:   currentSolution  $\leftarrow$  Generate a random solution using  $\phi_n$ .
31:   neighbors  $\leftarrow$  Retrieve all the neighbors of the current solution.
32:   bestNeighbor  $\leftarrow$  Find the best neighbor among all the neighbors.
33:   while  $\epsilon_s(\text{currentSolution}) > \epsilon_s(\text{bestNeighbor})$  do
34:     update the current solution with the best neighbor.
35:   end while
36: end procedure
```

Algorithm 2. Memetic-based framework for optimal RSU deployment (MARP algorithm)

Input: Probable Locations $C = C_1, C_2, C_3, \dots, C_k$, R : RSUs Quantity, size of the population n , Number of generations G_N

Output: Probable locations R of RSU

procedure Generate_RSU_sites:

```

1: Generate initial population P of n size where an entity has R characteristics
2:  $P_{sel}, P_{mut}, P_{loc-mut} \leftarrow$  Selection probability, mutation and local mutation probability
3:  $Ind_{mut}, Ind_{cross} \leftarrow$  Mutation index and crossover index respectively
4:  $Rand_m, Rand_c, Rand_{loc-mut} \leftarrow$  Random number produced between 0-1.
```

```

5:  $\epsilon_s \leftarrow$  Score of the fitness.
6: while not  $G$  do
7:    $\beta_n, \theta_n \leftarrow$  New population, blank list for every generation child
8:   calculate  $\epsilon_s$  for  $P$  and sort  $P$  on  $\epsilon_s$ 
9:   for  $i: 0$  to  $n^* P_{\text{sel}}$  do
10:     $\beta_{ni} \leftarrow P_i$ 
11:   end for
12:   for  $i: n^* P_{\text{sel}} + 1$  to  $n$  do
13:      $E_{c1}, E_{c2} \leftarrow$  Selection ( $\beta_n$ )
14:      $\phi_{ni} \leftarrow$  Crossover ( $E_{c1}, E_{c2}, I_c$ )
15:     if  $Rand_m < P_{\text{mut}}$  then
16:        $\theta_{ni}' \leftarrow$  Mutate (,  $\theta_{ni}, Ind_{\text{mut}}$ ,)
17:       Replace  $\theta_{ni}$ , by  $\theta_{ni}'$ 
18:     end if
19:   end for
20:   Combine the new chromosomes obtained in  $\theta_I$  with the chromosomes in  $\beta_n$ 
21:    $\theta_I \leftarrow$  empty list (local search process chosen chromosomes)
22:   for  $i: 0$  to length ( $\beta_n$ )–1 do
23:     if  $R_l$  is less than  $P_l$ , do
24:       Local_search_replacement ( $\beta_{ni}$ )
25:     end if
26:   end for
27:    $P \leftarrow \lambda_n$ 
28: end while
29: return  $P_0$ 
30: end procedure
31: procedure Local_search_replacement ( $chromosome_i$ ):
32:    $existingSolu \leftarrow chromosome_i$ 
33:   local-genmax  $\leftarrow$  maximum generations required for local search process
34:   for  $i: 0$  to local-genmax–1 do
35:      $newSolu \leftarrow$  Mutate  $chromosome_i$ 
36:     if  $\epsilon_s(newSolu) > \epsilon_s(existingSolu)$  do
37:       update  $existingSolu$  with  $newSolu$ 
38:     end if
39:   end for
40: end procedure

```

We have confirmed that the potential RSU location R is evaluated only once avoiding any recalculations in further derivations of Ψ . The goal of the GARRH method is to maximize the function Ψ . Therefore, the respective fitness function for MARP is expressed as:

$$\alpha = \text{maximum } (\Psi). \quad (2)$$

Steps 9 to 11 in the procedure conducts selection process on the basis of P_{sel} (selection ratio) of the best individuals from the existing population. This is done by means of an elitist process and then they are inserted into the new population β_n . This ensures that the best individuals from the current generation are forwarded to the next generation sans any alteration. Random selection of best fit parents is done among these best individuals for their genes transmission in the succeeding population in step 13. In our other proposed variant GARRH this is achieved using a crossover process (single-point). While one parent lends information to the child from itself towards the left of I_c , the other parent does so towards the right of Ind_{cross} . Therefore, it is guaranteed that the child holds the genes coming from both the parents. Then on the basis of mutation probability P_{mut} the generated offspring is mutated in steps 15 to 18. This injects diversity into the population. Mutation ensures that the individuals in the population do not become too similar to each other which can have an adverse effect of not reaching the optimum solution. The local search process is employed on

the children produced by GA in steps 20 to 23. This process of local search helps in faster determination of the local optimum and in rapid convergence. We have leveraged the hill climbing method with random restart as its local search procedure.

The algorithm is initiated with a random configuration in each iteration and is treated as the current solution. The neighbors of the current solution are found by applying a transition operator following which the best neighbor is retrieved. The fitness of the best neighbor is evaluated and if it is found to be better than the current solution, then the best neighbor becomes the current solution for the next iteration. The entire process is delineated in lines 29 to 37 of **Algorithm 1**.

On the contrary, **Algorithm 2** which is termed as MARP algorithm, employs local search process³⁰ by means of a simple hill climbing technique. It differs from GARRH in two ways. First, the combination of chromosomes with the current ones generates the individual on which the local search is implemented. Second, process of local search is performed on the basis of local mutation probability. It can be said that no random configuration is utilized in the proposed MARP algorithm unlike the GARRH algorithm. The present chromosome is treated as the current solution. It is mutated to form a new solution and if the new solution is evaluated to be better than the current one, the new solution is treated as the current solution for the next iteration. This process is repeated for several iterations not exceeding *max_local_gen*. This notion is explored in steps 29 through 37 of MARP. Step 12 ensures only the least fit in the current population are replaced by the new offspring generated. MARP incorporates a local search process to avoid optimal solutions in the local search space. The next generated population undergoes the fitness evaluation once again till the stopping condition is reached. The MARP algorithm outputs a set of probable *R* locations having maximum RSU coverage and achieving minimal delay in the transmission.

Figure 3 depicts the procedure of MARP. First, there is a random deployment of RSUs at selected intersections and this assumption is modeled by a network simulator. The fitness function of MARP determines the fitness of the existing population by using genetic procedures such as selection, crossover, and mutation. The local search procedure is implemented on the offspring produced. Further, the algorithm substitutes a fraction of the existing population by the generated off springs. Subsequently, the stopping condition is checked if it is reached. The stopping condition in MARP equals the maximum generation number touched and if this condition is not fulfilled, we repeat the process. Several parameters need to be defined to guide the search process in a memetic framework.³¹ In case of MARP, some simulation parameters are assumed which will be discussed in the implementation section later.

3.4 | System model of object detection model

In this section, the second aspect of VSDAS system (ie, low-latency object detection) is aimed to support the motive of seamless connectivity by integrating this aspect with efficient RSU placement technique in the proposed system. Object detection helps in better tracking and monitoring of the vehicles for determining its positions on the road. It is considered to be one of the key components in the development of artificial intelligence-empowered secure vision-dependent driving assistance models in IoAV for efficiently locating the different semantic object patterns for a certain class of objects.³² A review and analysis of the existing machine-learning methods of vehicle detection is done in this article. It entirely elaborates the roadmap of implementing object detectors namely, HOG, LBP for a better comparison. However, our proposed method further improves the results by capturing the information using Haar-like traits using AdaBoost cascade classification. In this article, we have explored the link between the similar frames and latency drop-in real-time scenario for simulating the object detection. Hence, a novel scalable cluster object detection system is proposed that comprises enhanced object detection modules with the selection of key frames. Figure 4 illustrates the overall proposed framework of object detection.

3.5 | Significant video frames detection techniques

To achieve the accurate localization of AVs on the road, it is essential to equip vehicles with robust OD technique which is highly responsive in time-critical scenarios such as driving. Since the cameras capture long video sequences to monitor the traffic for detecting vehicles on road with higher precision, it is important to select the key frames from the long video sequences captured with the help of camera. It is a primary step in the processing process of a video as a small chunk of video has hundreds of frames in it.

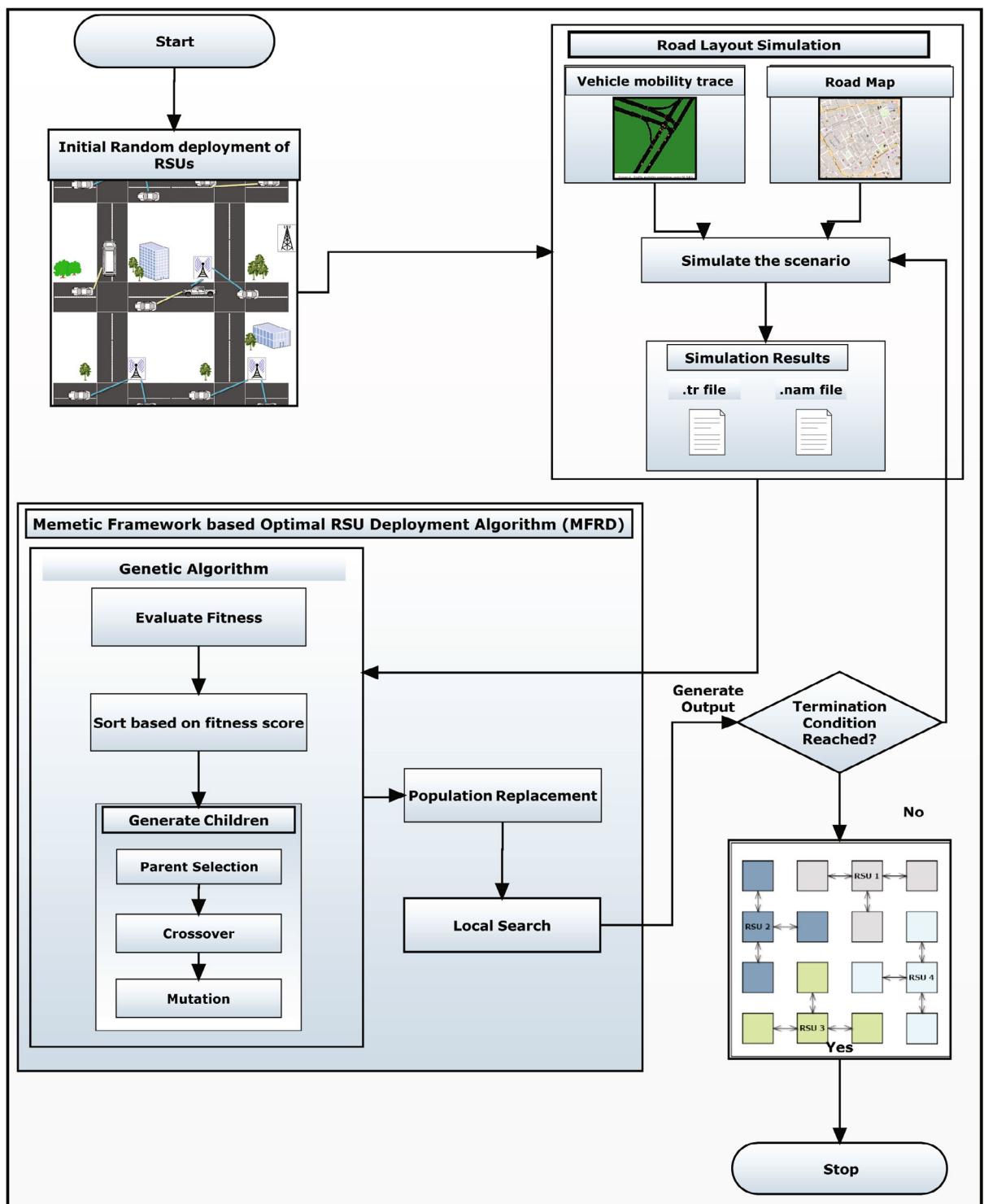


FIGURE 3 Memetic framework-based optimal RSU deployment progression

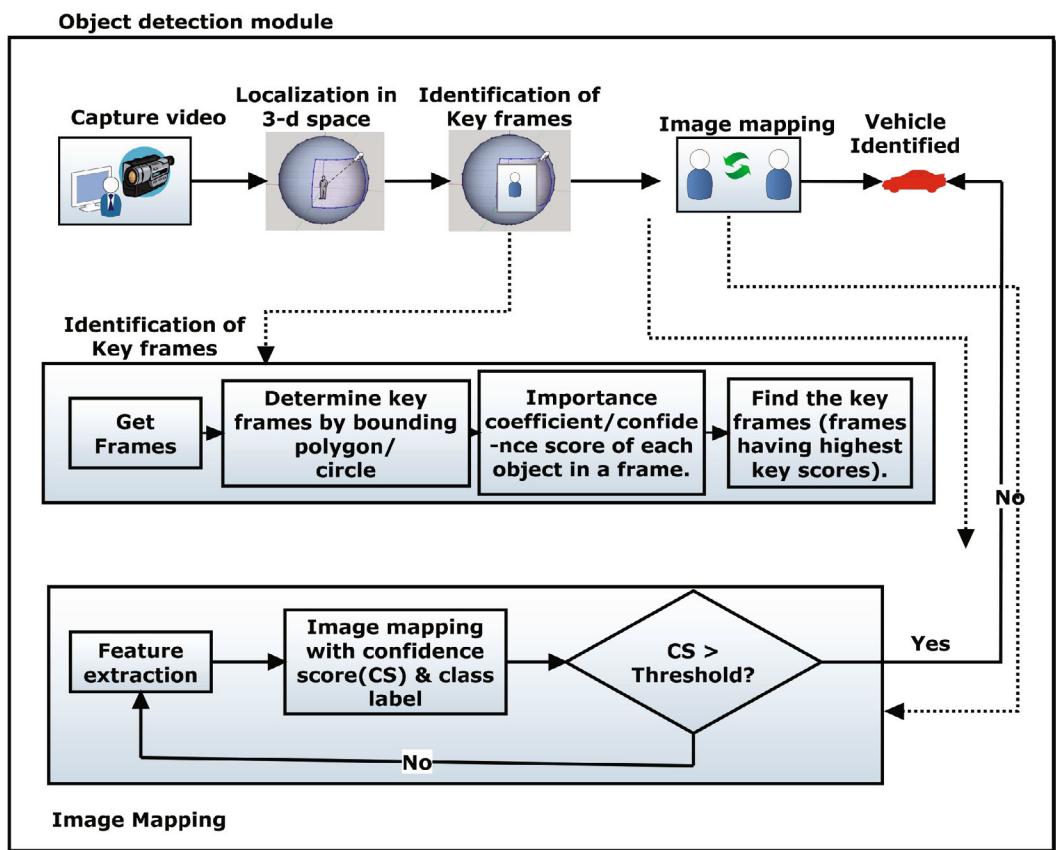


FIGURE 4 System design model of object detection model

There may be many dissimilar data frames to be processed due to which the system experiences low processing speed & compromised accuracy.^{33,34} Hence, it is crucial step to choose important frames having meaningful correlation among them. This will facilitate the process of feature extraction which is followed after selecting the key frames from the video chunk. Therefore, we design a key frame selection procedure that compares the two similar frames of the video. The similarity score of the frames which have high correlation is evaluated in this algorithm. The algorithm explanation is further explained in next section.

3.6 | Relevant object detection algorithm

In the real time video surveillance, the camera records hours of video data. Though, only some part of the video is significant to us which holds dominant information. Thus, to retrieve the key video frames from long videos, our proposed object detection technique is required. We studied that the object selection in a frame is in accordance with the number of objects present in the video. To emphasize on this fact, we can locate the object of interest in a video which appears to be more prolonged and has greater coverage area in the frame. Therefore, subjectively it will appear to be lingering in a central view giving an impression of a foreground image such that the object is more important and viewed more centrally.^{35,36} In this view, we take center point of an image as one of the parameters and the second parameter is object's coverage area in the image. The object's importance is determined by its closeness to the center point of an image. The object holds more weight as the object appears closer to the center point. To complete this process, we calculate the object's weight and afterwards evaluate the key score of every object present in the sequence of the video. Consequently, the relevant frames are determined to be considered in accordance with the key score.

Procedure: We consider an example image on which the center point of the object of interest Obj_I is shown as $Obj_I(x_1, y_1)$ and image's center point of the image is $I_{cent}(x_i, y_i)$. The weight coefficient of the image having height h_t and width w_d is denoted as W_{coeff} . There exists a negative correlation between the frame center and object center. The W_{coeff} can be evaluated as follows in Equation (3) below:

$$W_{coeff} = 1 - \frac{\sqrt{(x_I - x_i)^2 + (y_I - y_i)^2}}{\frac{\sqrt{(w_d^2 + h_t^2)}}{2}}. \quad (3)$$

The range of W_{coeff} can be interpreted from Equation (3) as maximum 1 and minimum 0 [0, 1]. The highest limit 1 denotes center of the image and lowest limit denotes the picture corners. The calculation of the key score can be shown in the Equation (4).

From Equation (3), we can interpret that range of weight coefficient is [0, 1], maximum limit being 1. Here, maximum limit 1 signifies the image center & minimum limit 0 signifies the four corners of the picture. Subsequently, the key score calculation can be done as follows in Equation (4):

$$\text{Key}_{\text{score}} = \sum_{i=0}^N \sum_{j=0}^M \langle W_{coeff}(i, j) \times w_d(i, j) \times h_t(i, j) \rangle, \quad (4)$$

where the constraints are explained as follows:

N : video frames summation

M : summation of all the objects determined in j frames

$W_{coeff}(i, j)$: weight coefficient

$w_d(i, j)$: object width (i, j)

$h_t(i, j)$: object height (i, j)

The share percentage of every identified object in k objects is shown in Equation (5) as follows:

$$\text{Per} = \frac{\text{Key}_{\text{score}}}{\sum_{i=0}^k \text{Key}_{\text{score}}(i)}. \quad (5)$$

The procedure of the relevant frame recognition algorithm in a video chunk is explained next in the **Algorithm 3** which helps in focussing on key frames for object detection. The algorithm results in producing set of important frames in the video. The procedure CalculatePercentage() determines the proportion of key score. This function calculates every object's proportion out of k objects and Equation (5) evaluates the percentage of the object proportion. Subsequently, CalculateKeyScore(N) calculates each object's score in the frame. The CalculateWeightCoefficient() **function** tells the importance of the object on the basis of calculating key score and percentage.

Algorithm 3. Relevant object frame selection algorithm (Selection of relevant object frame)

Input: video frames summation N

Output: set of key frames

procedure **CalculatePercentage**:

$\text{PercentageScore} \leftarrow$ array of size N

$T_h \leftarrow$ Threshold value for proportion

$\text{Key_Frames} \leftarrow$ set of key frames

1. $\text{TotalKeyScore}, \text{KeyScore} \leftarrow \text{CalculateKeyScore}(N)$ // TotalKeyScore is a variable while KeyScore is an array of size N
for $i = 0, 1, \dots, N-1$ **do**

$\text{PercentageScore}_i \leftarrow \text{TotalKeyScore}/\text{KeyScore}_i$

end for

for $i = 0, 1, \dots, N-1$ **do**

if $\text{ProportionScore}_i > T_h$ **then**

 Add Frame i to Key_Frames

end if

end for

 return Key_Frames

procedure **CalculateKeyScore (N)**:

$\text{TotalKeyScore} \leftarrow 0, \text{KeyScoresOfFrames} \leftarrow$ array of size N

```

for  $i = 0, 1, \dots, N-1$  do
     $M \leftarrow$  summation of all the objects determined in  $j$  frames
     $I(x, y_i) \leftarrow$  center point of image  $i$ 
     $w_d \leftarrow$  width of image  $i$ ,  $h_t \leftarrow$  height of image  $i$ 
    for  $j = 0, 1, \dots, M$  do
         $O_I(x_i, y_i) \leftarrow$  object of interest  $O_I$  center point of object  $j$ 
         $W_c(i, j) \leftarrow$  CalculateWeightCoefficient( $O_I(x_{OI}, y_{OI}), I(x, y_i), w_d, h_t$ )
         $w_d(i, j) \leftarrow$  object  $j$  width,  $h_t(i, j) \leftarrow$  Object  $j$  height
         $KeyScoresOfFrames_j \leftarrow W_c(i, j) \times w_d(i, j) \times h_t(i, j)$ 
         $TotalKeyScore \leftarrow TotalKeyScore + KeyScoresOfFrames_j$ 
    end for
end for
return  $TotalKeyScore, KeyScoresOfFrames$  to CalculatePercentage
procedure CalculateWeightCoefficient ( $PointO_I(x_{OI}, y_{OI})$ ,  $PointI(x, y_i)$ ,  $widthw_d$ ,  $heighth_t$ ):
     $x_{OI} \leftarrow O_I.x$ -coordinate
     $y_{OI} \leftarrow O_I.y$ -coordinate
     $x_i \leftarrow C.x$ -coordinate
     $y_i \leftarrow C.y$ -coordinate
     $W_c \leftarrow 1 - \sqrt{(x_{OI} - x_i)^2 + (y_{OI} - y_i)^2} / \sqrt{w_d^2 + h_t^2} / 2$ 
    Return  $W_c$  to CalculateKeyScore

```

After getting the key frames from the [Algorithm 3](#) we apply our proposed algorithm named enhanced Haar-cascade. The key frames contain the object of interest that is the positive sample where a feature of an object is present in the frame to be processed further. Our proposed algorithm then scans only the relevant frames instead of the whole sequence of video. This preprocessing step of key frame selection boosts the usage of multiple cascading classifiers since, there is a dedicated frame to process. Since, the Haar-cascade method is faster in detecting objects due to the utilization of integral images, the selected frames are now fed in the algorithm for even faster performance.

Haar-like feature: A robust algorithm named Haar-like feature changes the image pixels to grayscale values which makes easier to detect the objects. Haar can be said a series of resized square shape utilities. Those are features of a digital image which can be used in recognizing objects. It scans successive rectangular regions inside the detection window.³⁷ Then the summation of pixel intensities in adjacent regions are done. Haar-like features are simple rectangular frames which are scalable and relates the pixel intensities to each other. The advantage to use this algorithm is that it can reduce the dimension of the data and achieves a real time performance.³⁸ Therefore, in this section, we elaborate object detection algorithm based on Haar features. Here, there are two perspectives being observed:

- Haar features oriented computational design
- Design based on compound stage classifier

Haar features oriented computational design is generated based on two positions: one is an erected rectangle and other is rotated rectangles at 45°.

Haar-like attributes algorithm: This algorithm scans rectangular zones at different angles to create the Haar-like features. The rectangle position at an upright angle and at an angle of 45° are used for Haar-like trait generation.³⁹ We have explained the pseudocode of the algorithm in [Algorithm 4](#). We determine the presence of an object in a frame by measuring the (w_d, h_t) pixels. The total probable trait set of the intermediate Haar-attributes is depicted as follows in Equation (6):

$$\text{Attribute}_m \sum_{i \in m=1, \dots, N} \text{PS}(t). \quad (6)$$

In our proposed methodology, the rectangular region is expressed as a tuple $t = (p, q, w_d, h_t, \theta)$. The ranges of each parameter can be defined as: $0 \leq p, p + w_d \leq W, 0 \leq q, q + h_t \leq H, p, q \geq 0, w_d, h_t > 0$ and $\theta \in [0^\circ, 45^\circ]$.

The pixel sum is calculated as $\text{PS}(t)$. The upright rectangular zones are formulated as follows. The supplementary image is represented as totaled area in Equation (7):

$$\text{Totaled area}(T) = (p, q) \quad (7)$$

and the T can be well-defined as the summation of all the image pixels of an upright rectangular region as shown in Equation (8). This rectangular area ranges from left corner coordinate at $(0, 0)$ to the right corner coordinate at (p, q) .

$$T(p, q) = \sum_{p' \leq p, q' \leq q} I(p', q'). \quad (8)$$

This can be calculated with one iteration of navigation from left to right of all the pixels and again from top to bottom as seen in line 16 of [Algorithm 4](#) where the following parameters are defined as:

$$T(-1, q) = T(p, -1) = 0.$$

Algorithm 4. Haar-like traits algorithm for video-based object detection (Haar-like traits algorithm)

Input: Selected Key frame

Output: Pixel summation of image's rectangular subregion

```

1. procedure PS (Tuple tup, Image I):
2.   deg ← tup.deg
3.   If deg is 0, then
4.     return UprightSumRec (tup, I)
5.   else
6.     return RotatedSumRec (tup, I)
7. procedure UprightSumRec (Tuple tup, Image I):
8. // auxiliary grayscale_img is a 2-D matrix when each cell represents the pixel intensity.
9.   p ← tup.p, q ← tup.q, wd ← tup.wd, ht ← tup.ht
10.  nRows ← rows count in WC
11.  nCols ← columns count in WC
12.  T ← a 2-D matrix with nRows and nCols
13. // a single iteration over all image pixels (top to bottom) and (left to right)
14.  for i: 0, 1, ..., nRows – 1 do
15.    for j: 0, 1, ..., nCols – 1 do
16.      T(p, q) = T(p, q – 1) + T(p – 1, q) + I(p, q) – T(p – 1, q – 1)
17.    end for
18.  end for
19.  PS(tup) = T(p – 1, q – 1) + T(p + wd – 1, q + ht – 1) –
20.    T(p – 1, q + ht – 1) – T(p + w – 1, q – 1)
21.  return PS
22. procedure RotatedSumRec (Tuple tup, grayscale_img):
23.   p ← tup.p, q ← tup.q, wd ← tup.wd, ht ← tup.ht
24.   // grayscale_img is a 2-D matrix when each cell represents the pixel intensity.
25.   nRows ← rows count in grayscale_img
26.   nCols ← columns count in grayscale_img
27.   Trot ← a 2-D matrix with no of rows and columns as nRows and nCols respectively
28.   // first pass from left to right and top to bottom
29.   for i ← 0, 1, ..., numRow – 1 do
30.     for j ← 0, 1, ..., numCol – 1 do
31.       Trot(p, q) = Trot(p – 1, q – 1) + Trot(p – 1, q) + I(p, q) – Trot(p – 2, q – 1)
32.     end for
33.   end for
34.   // right to left second pass and bottom to top
35.   for i: 0, 1, ..., numRow – 1 do
36.     for j = 0, 1, ..., numCol – 1 do
  
```

```

36.       $T_{rot}(p, q) = T_{rot}(p, q) + T_{rot}(p-1, q+1) - T_{rot}(p-2, q)$ 
37.       $PS(tup) = T_{rot}(p+w_d, q+w_d) + T_{rot}(p, h_t, q+h_t) - T_{rot}(p, q) - T_{rot}(p+w_d-h_t, q+w+h_t)$ 
38.      end for
39.      end for
40.      return PS

```

Following that, the upright rectangle tuple $tup = (p, q, w_d, h_t, \theta)$ can be calculated in accordance with the pixel sum, and the corresponding Equation (9) is as follows.

$$PS(tup) = T(p-1, q-1) + T(p+w_d-1, q+h_t-1) - T(p-1, q+h_t-1) - T(p+w-1, q-1) \quad (9)$$

Though, for the rotated rectangle area at 45° , the corresponding image can be depicted in Equation (10) as follows:

$$T_{rot}(p, q) = \sum_{p' \leq p, p' \leq p-q-q} I(p', q'), \quad (10)$$

where $T_{rot}(p, q)$ can be evaluated by two iterations of passes on all pixels. The first pass starts from left coordinate to right coordinate and top coordinate to bottom coordinate is determined by Equation (11):

$$T_{rot}(p, q) = T_{rot}(p-1, q-1) + T_{rot}(p-1, q) + WC(p, q) - T_{rot}(p-2, q-1) \quad (11)$$

and $T_{rot}(-1, q) = T_{rot}(-2, q) = T_{rot}(p, -1) = 0$.

The second iteration from right corner to left corner and from bottom to top coordinate is depicted in Equation (12) as follows:

$$T_{rot}(p, q) = T_{rot}(p, q) + T_{rot}(p-1, q+1) - T_{rot}(p-2, q). \quad (12)$$

Hence, the rotated rectangle $tup = (p, q, w_d, h_t, 45^\circ)$ can be computed according to the line 37 in [Algorithm 4](#).

Multiple cascading classifiers: In our designed experiment, assuming that the detection rate of every Haar-like feature which is trained on the model is not very good, we therefore design a strong classifier having an improved detection rate.⁴⁰ Multiple classifiers are integrated as shown in the Figure 5. The object of interest is found out by navigating the moving window on the image sequentially and it is ensured by the training process. Nevertheless, if any scan is not able to detect any object, the remaining portion is left untouched.

Therefore, this section elaborates the approach of enhanced Haar-cascade object detection algorithm which is video-based key frame selection utilizing AdaBoost cascade classifier for improving its result. This will deliver better performance and conduct faster object detection in real time.⁴¹

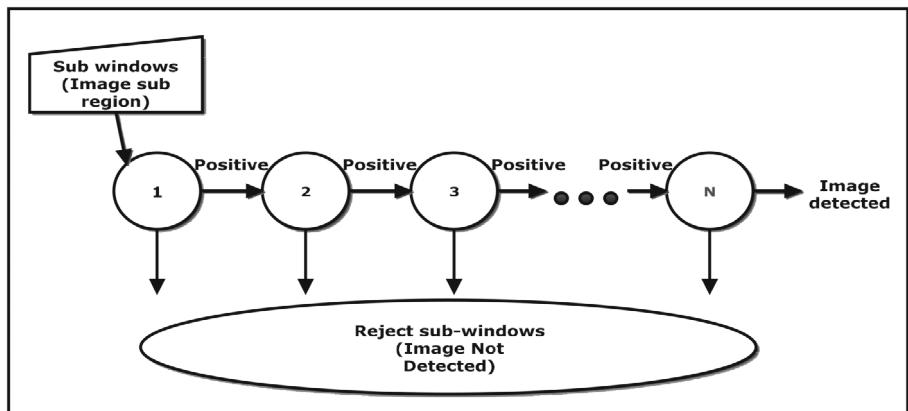


FIGURE 5 Multiple cascading classifiers

4 | IMPLEMENTATION SETUP AND EVALUATION OUTCOMES

Since the deployment of vehicular network is less feasible in regard with scalability factor of the AVs and the incurring expenses therefore, this section discusses the environment opted for simulating the proposed framework. This will then help in assessing the viability and scrutinizing the performance of our proposed framework.

4.1 | Experimental setup

For an exhaustive evaluation and diversity in results, we opted for two different topological locations, namely, Murlipura and Connaught Place (CP). We have considered different types of topographies to simulate the RSU deployment environment. The open-source software OSM (OpenStreetMap) sources the topographic maps of these two places. The second open-source software SUMO (Simulation of Urban MObility), a realistic traffic generator helps in generating mobility of vehicles for road network and route visualization.⁴² We have chosen different types of location for diversity. Murlipura, Jaipur is chosen for less dense area representation having lesser number of intersections while Connaught Place, New Delhi is chosen for representing dense road network having greater number of intersections on road. This provides better probabilities for RSUs placement locations. The OSM and SUMO road networks for Murlipura and Connaught Place are shown in Figures 6 and 7, respectively. As discussed in the previous section that various parameters are needed to be defined to guide the memetic framework search process for which they have been summarized in Table 2. Table 3 summarizes all the road and route information of the two preferred locations. The utilization of these locations aids in confirming the potential of proposed MARP algorithm in choosing the optimal locations for RSU under different conditions.

Table 2 provides an exhaustive list of parameters which we have considered in our work. Under the standard of IEEE 802.11p the Wireless Access in Vehicular Environments (WAVE) standard has been simulated. Even though there are many RSU deployment techniques been proposed we have chosen to do a comparative analysis of our approach against the GA algorithm⁷ because it delivers sufficient facts to conduct a fair evaluation. We then developed memetic framework which uses local search process based on a GARRH. It was then subsequently upgraded by utilizing a different local search process of simple hill climbing approach for fine tuning the search results ensuing in the MARP algorithm. We analyze the MARP performance to find the best possible RSU sites between all the available sites. The surety of results getting a representative value comes from running an average of five rounds spanning over 20 implementations. Further, MARP performance is compared against GA and GARRH algorithms.¹⁶



FIGURE 6 OSM and SUMO road network for Murlipura



FIGURE 7 OSM and SUMO road network for Connaught Place

TABLE 2 Simulation parameters

Parameter	Value
Quantity of vehicles	5-100
Quantity of RSUs	4-20
Type of direction	Bi-directional
Velocity of vehicle	10-50 m per second
Area covered	500 m × 700 m
Routing protocol used	AODV
Channel mode	Wireless
Propagation model	Two ray ground
Interface queue	Drop tail/Priority queue
Antenna mode	Omni-directional
MAC type	IEEE 802.11
Type of agent	TCP
Time of simulation	100 s
Initialization	Random
Genotype representation	Integer
Parent selection	Tournament using $k = 2$
Type of crossover	Single-point
Mutation probability	1 gene per individual
Size of population	8
Total generations	20

TABLE 3 Details of the area layouts

Area	Count of intersections ^a	Count of roads ^a
Murlipura, Jaipur	194	550
Connaught Place, New Delhi	1069	1890

^aAs derived from SUMO network visualization.

4.2 | Evaluation outcomes

The overall RSU placement phase in the road areas is connected to the flow of the traffic on the road segments. We have considered the behavior of road segments having higher traffic density for deploying RSUs. However, in view of reducing the latency of the data delivery over the network RSUs cannot be completely installed at intersections having large traffic. Therefore, we have also considered the number of RSUs factor in observing the fitness of our proposed MARP and GAHRC algorithm in comparison with the existing methods.

The MARP performance was assessed for the selected locations and is contrasted with the performance of GA and GARRH. The evolutionary process is followed by MARP procedure and it is represented as fitness scores of the best and average chromosomes. This signifies best RSUs deployments in two different topologies by taking different number of RSUs (3, 6, 9, and 12), respectively in Figures 9 to 11. MARP algorithm maximizes the function of fitness to select optimal RSU intersections as denoted in Equation (1). As it is shown in Figures 8 to 10, the search process in MARP algorithm is fine-tuned which is consequential of having a substantial improvement in the fitness score with the increase in the number of generations. MARP performs the best for both the chosen locations by showing the greater fitness scores. The GARRH outperforms GA in all the cases. Though, it shows equivalent results with GA in two cases. One of the cases is when the RSUs number is least (in our case 3) for both the topologies and other is when the RSUs number is maximum for Murlipura. These have been established in Figure 8A,B.⁴³

The fitness values increase initially as generation number increases and a gradual convergence is witnessed by going near to the global optimum value. Further, abrupt fitness score improvements are witnessed in some cases which can be attributed to the involvement of local search process and mutation operator as local search process facilitates probing in the solution space for new values of maximums. On the other hand, the GARRH algorithm shows better results against GA, and further it is experienced that the proposed MARP shows even better results than the GARRH algorithm. GA produces incompetent results against proposed MARP technique as shown in Figures 8 to 10 in terms of a poor fitness score. Even though the GA has faster convergence against both the GARRH and the MARP algorithms, still the GARRH and MARP achieves noteworthy improved scores of fitness. This happens because of the local search method implemented in the GARRH and the MARP algorithms which facilitate in discovering the local optimum value in a relatively more efficient way as in GA.

Further, it is evident from Figure 11A,B that to improve the fitness values of both the topographies the RSU count increases. This observation is distinctly seen and more noticeable in case of a denser region taken in our work that is, Connaught Place comprising of greater number of road intersections. This enables in providing a wider coverage having more number of intersections by deploying lesser RSUs with less coverage areas overlap. Further, we have assumed the coverage area of each RSU to be equidistant and varied the coverage area of an RSU from 10 to 90 m.

We also infer from our results that if we compare the fitness of our proposed algorithm MARP with the other proposed variant GAHRC and other existing state-of-the-art in relation to the RSU coverage area, MARP algorithm shows slightly more fitness values as the coverage area increases. This happens for both the locations. The variation in the coverage area is in sync with the amount of the area chosen to be simulated and it is distinguishable from our inferential results that how an increment in the RSU coverage area enhances the output of the target function.

To visualize the MARP evolutionary process in Figure 12 we simulate the two different topological scenarios taking 6 RSUs and 12 RSUs for Murlipura and CP, respectively. Moreover, to reveal the process of computation of MARP algorithm for choosing the optimal RSU locations over successive generations we choose three different generations in the algorithm (after Generation 0, after Generation 15, and after Generation 30).

Figures 12 and 13 depicts the evolution of the RSU deployments over three generations in the chosen topologies. Similarly, Figures 14 and 15 display the best locations in case of 12 RSUs over three generations at the chosen topologies.

We observe from the experimental studies that as we increase the number of RSU the RSU placements increases as is seen in case of 12 RSUs against 6 RSUs. Therefore, it requires a greater number of generations for converging. The increase in the number of RSUs tends to increase its search space. Hence it is inferred that we can find better potential deployments for installing RSUs in later generations owing to the algorithm's randomness. The Figures 13 to 15 illustrate the evolving nature of MARP algorithm which chooses best locations of RSU over increasing generation number. Both the proposed algorithms that is, MARP and GARRH locate best places of installing RSUs by showing better fitness values, but it is obtained at the stake of higher response time. GA consumed less time as opposed to GARRH and MARP algorithms for generating solutions as can be seen in Table 4.

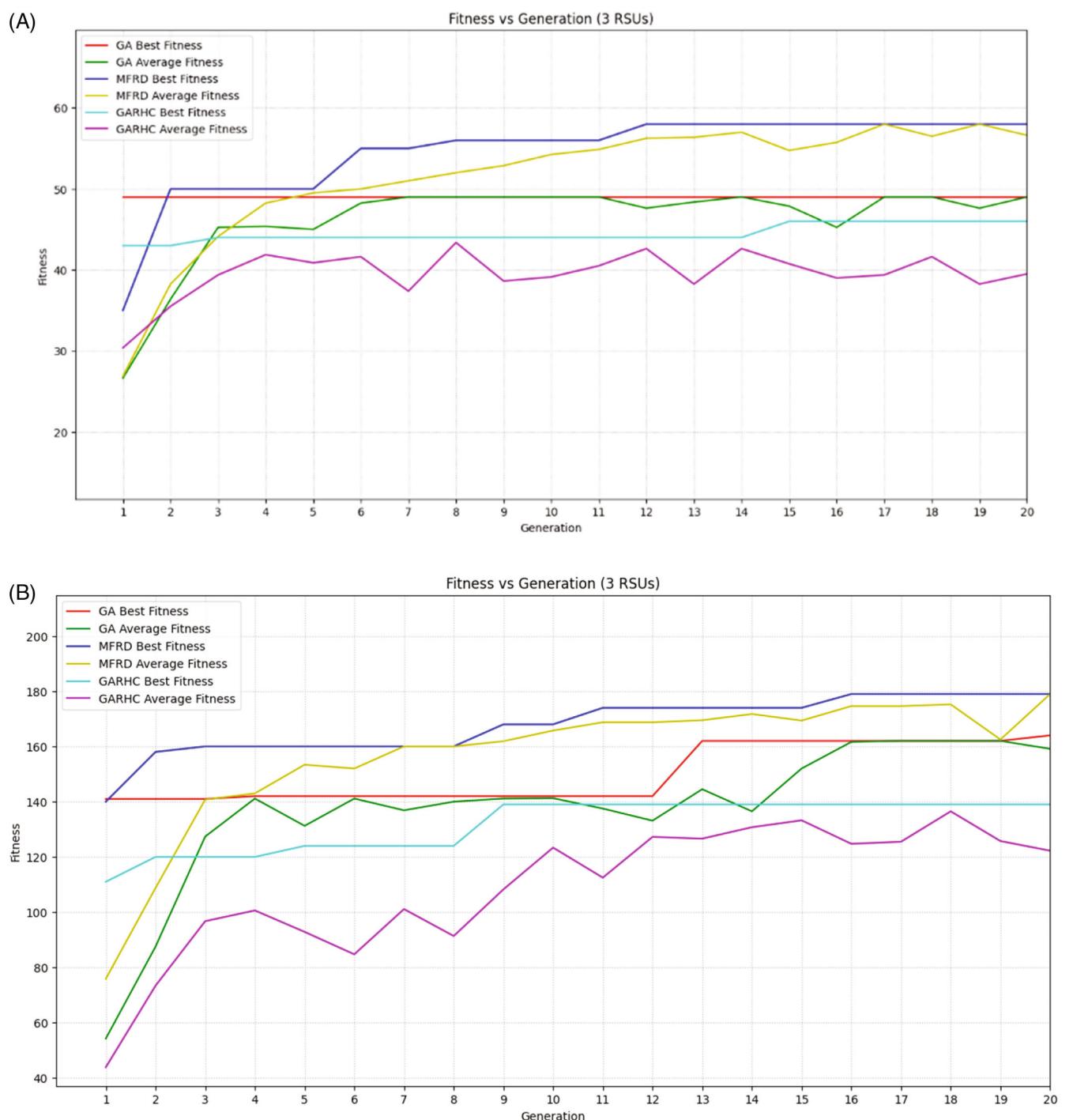


FIGURE 8 RSU placement algorithms fitness scores comparison. (A) 3 RSUs in Murlipura and (B) 3 RSUs in Connaught Place

5 | IMPLEMENTATION SETUP FOR OBJECT DETECTION

In this section, we examine the effectiveness of our proposed Haar-cascade model to see the improvement in object detection process of AVs. We also compare our proposed algorithm with obsolete methods namely, HOG, LBP.⁴⁴ Various setups like dataset selection, parameter settings, multiple stage procedures have been adopted to conduct the experiment.

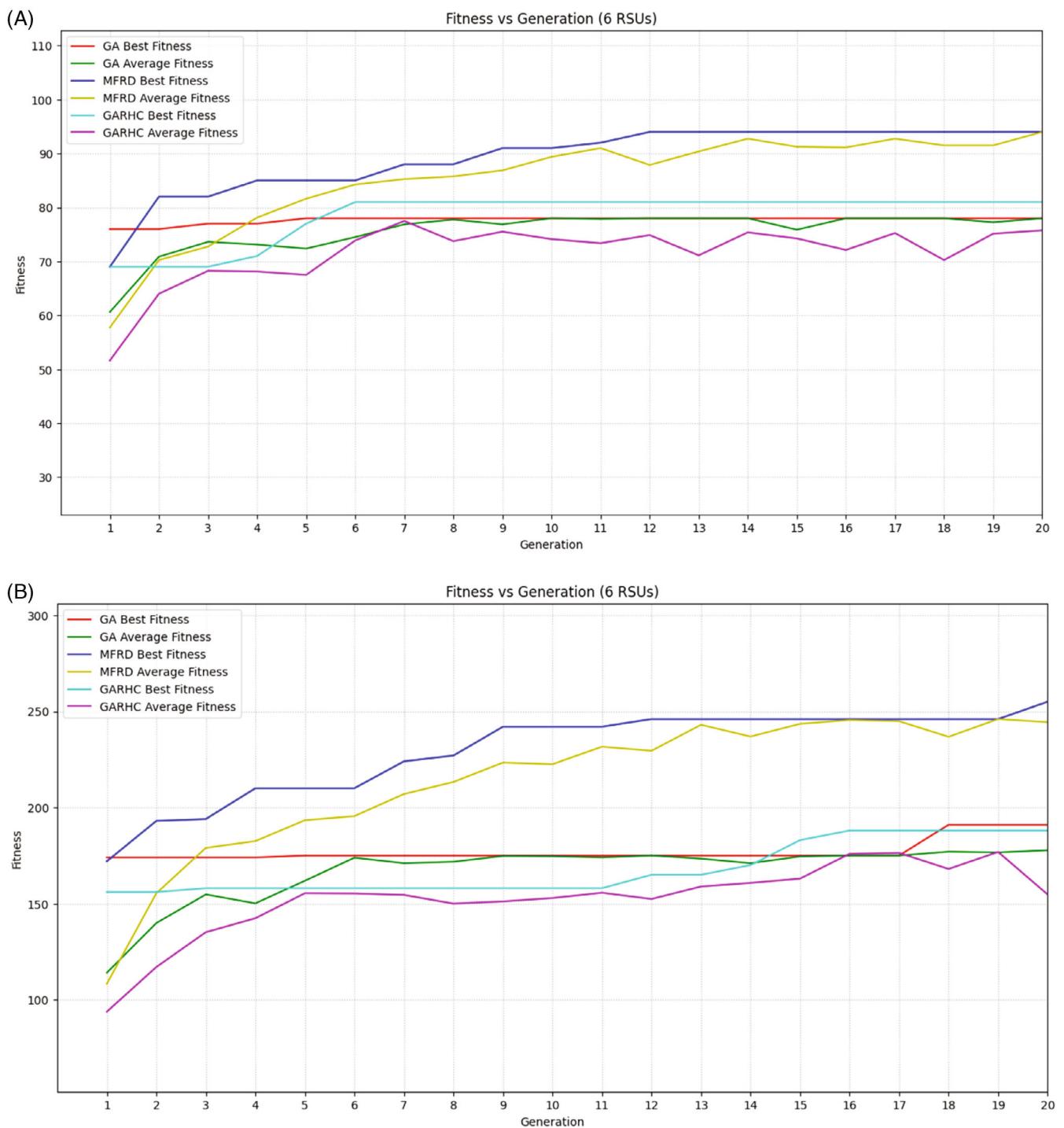


FIGURE 9 RSU placement algorithms fitness scores comparison. (A) 6 RSUs in Murlipura and (B) 6 RSUs in Connaught Place

5.1 | Experimental setup

We conduct certain experiments using two data sets, one is KITTI⁴⁵ dataset and the other is Panasonic⁴⁶ public dataset for further evaluation of our experiments. For leveraging various functions of video and image processing we have utilized the state-of-the-art OpenCV version 4.5.4 and utilized Ubuntu version 20.04 OS platform equipped with VGA compatible controller. The other functionalities were extracted by utilizing Matplotlib, NumPy, SciPy for visualizing the processed data as output.

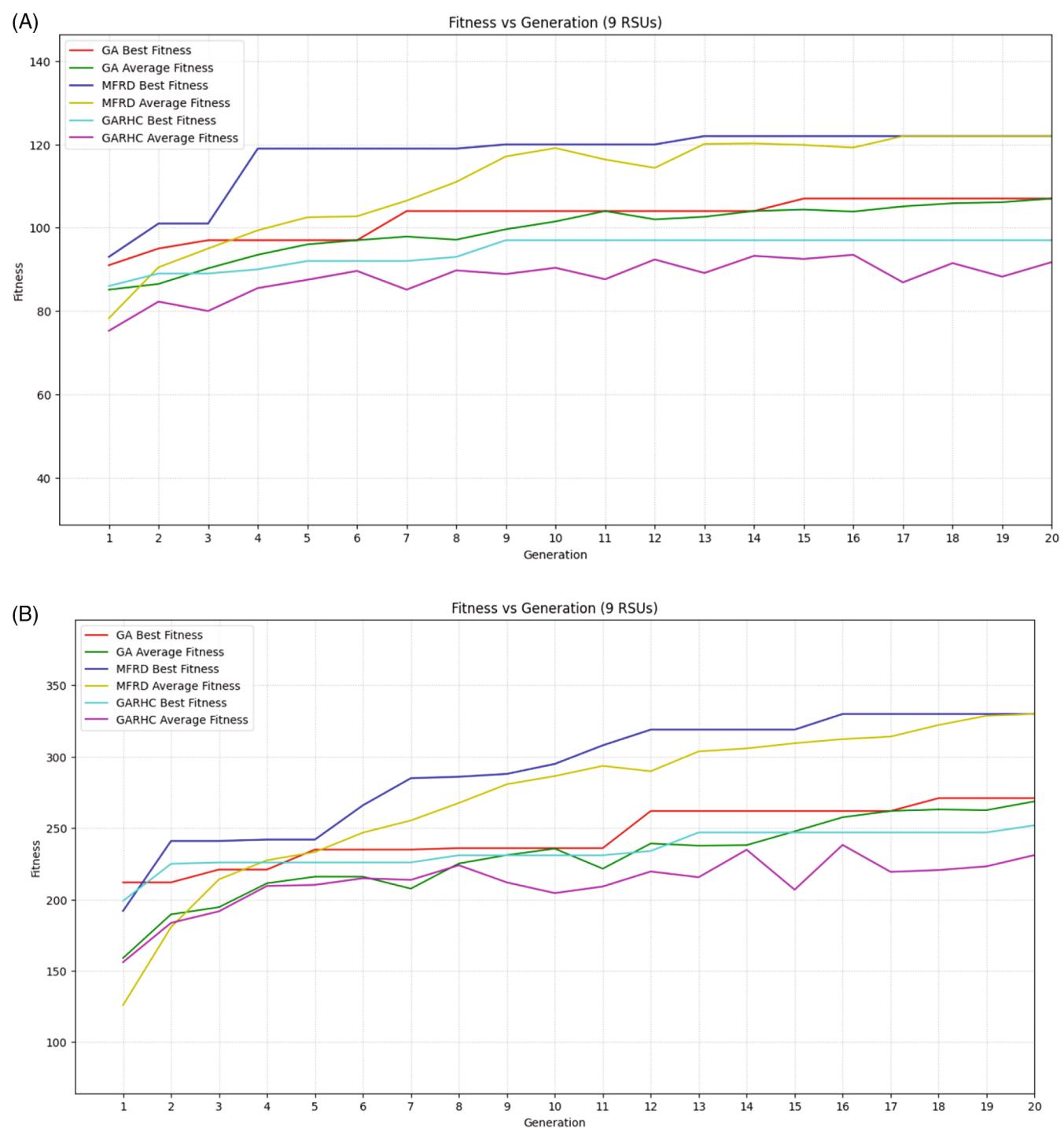


FIGURE 10 RSU placement algorithms fitness scores comparison (A) 9 RSUs in Murlipura and (B) 9 RSUs in Connaught Place

5.1.1 | Dataset selection

The quality of the dataset which we select directly affects the object detection performance. A good quality dataset should always be given consideration for the purpose of training the models. The dataset should comprise of vehicle images captured at various angles and in different orientations at different phases of the day in varying illumination conditions. The vehicle dataset needs to be varied consisting images in both urban traffic and different highway scenarios.⁴⁷ A dataset for training must contain both positive and negative samples of the images. Positive samples should be

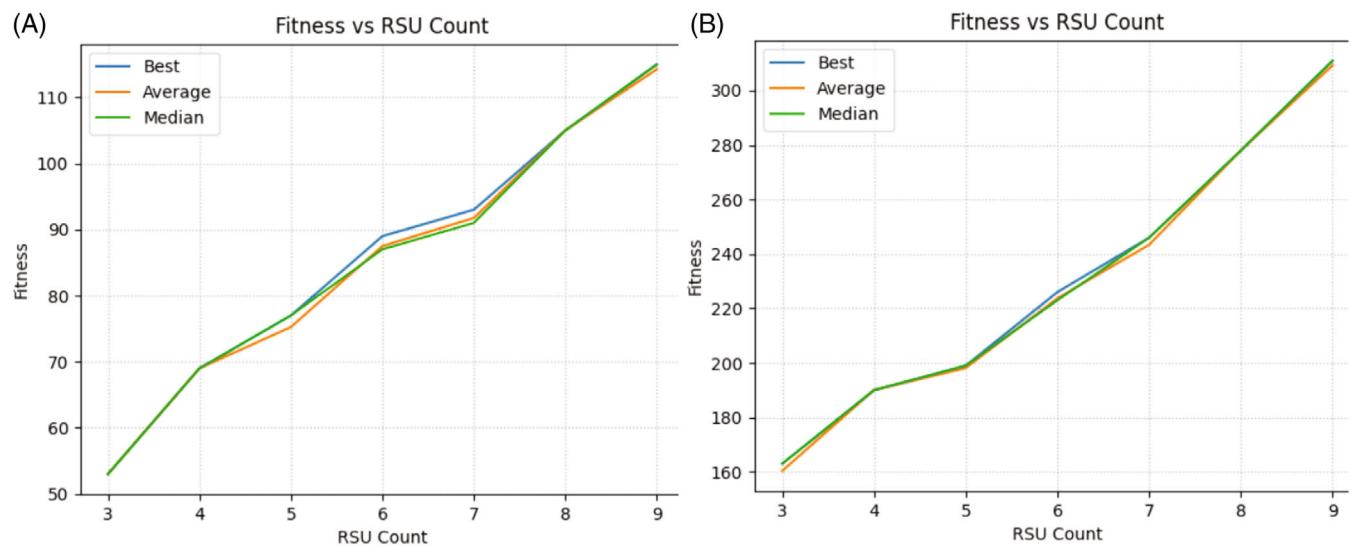


FIGURE 11 (A) Fitness versus number of RSUs in Murlipura. (B) Fitness versus number of RSUs in Connaught Place

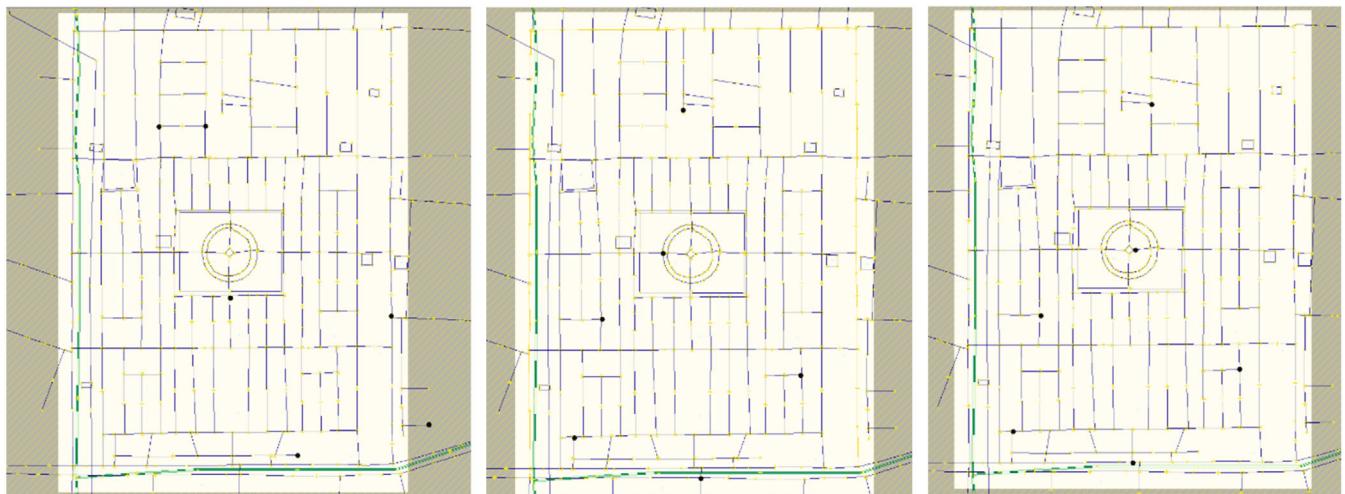


FIGURE 12 Evolution of the deployment of 6 RSUs in Murlipura after (A) Generation 0, (B) Generation 15, and (C) Generation 30

consisting of the objects of interest and negative samples should contain the background components like trees, roads, buildings/infrastructures and other less relevant things which are basically associated with our region of interest. This would help the classifier in differentiating between the foreground object of interest components like vehicles and the background components like non-vehicles while navigating the detector in a real time video sequence.

We perform our experiments first on the KITTI dataset which is a suite of computer vision functionalities and modular tasks based on autonomous driving platform. It contains 7000 training image and 7500 testing images. The KITTI vision benchmarking suite uses 2D anchor/bounding boxes to localize the objects to be identified.⁴⁸ Further, it computes the precision-recall curves for object detection for better comprehension. There were no bounding boxes provided by the KITTI dataset for the objects to be qualified for detection, therefore, rectangular boxes are computed to form an anchor box. Four boundary lengths are calculated of the image segment to form a bounding box to determine the location of the object in the image. We have set four classes for object detection namely, car, bus, pedestrian, bike. However, there were some categories which were not important for the detection of objects therefore, they are set as background classes. To further verify and validate the precision of our proposed model, we also carry out the experimentations on Panasonic public dataset.



FIGURE 13 Deployment evolution of 6 RSUs in CP after (A) Generation 0, (B) Generation 15, and (C) Generation 30

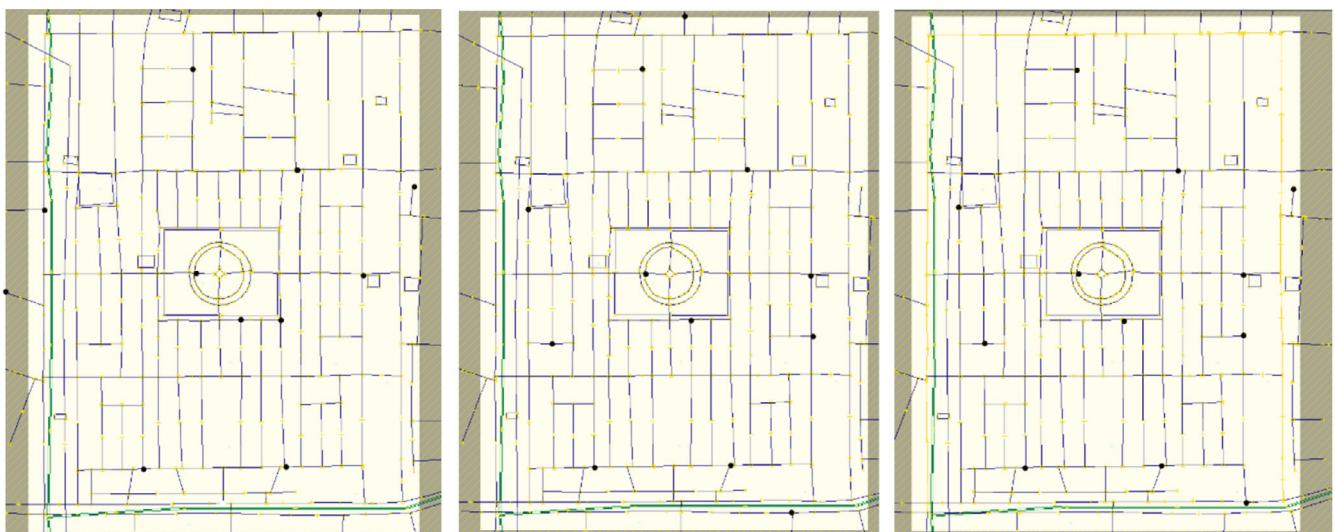


FIGURE 14 Evolution of the deployment of 12 RSUs in Murlipura after (A) Generation 0, (B) Generation 15, and (C) Generation 30

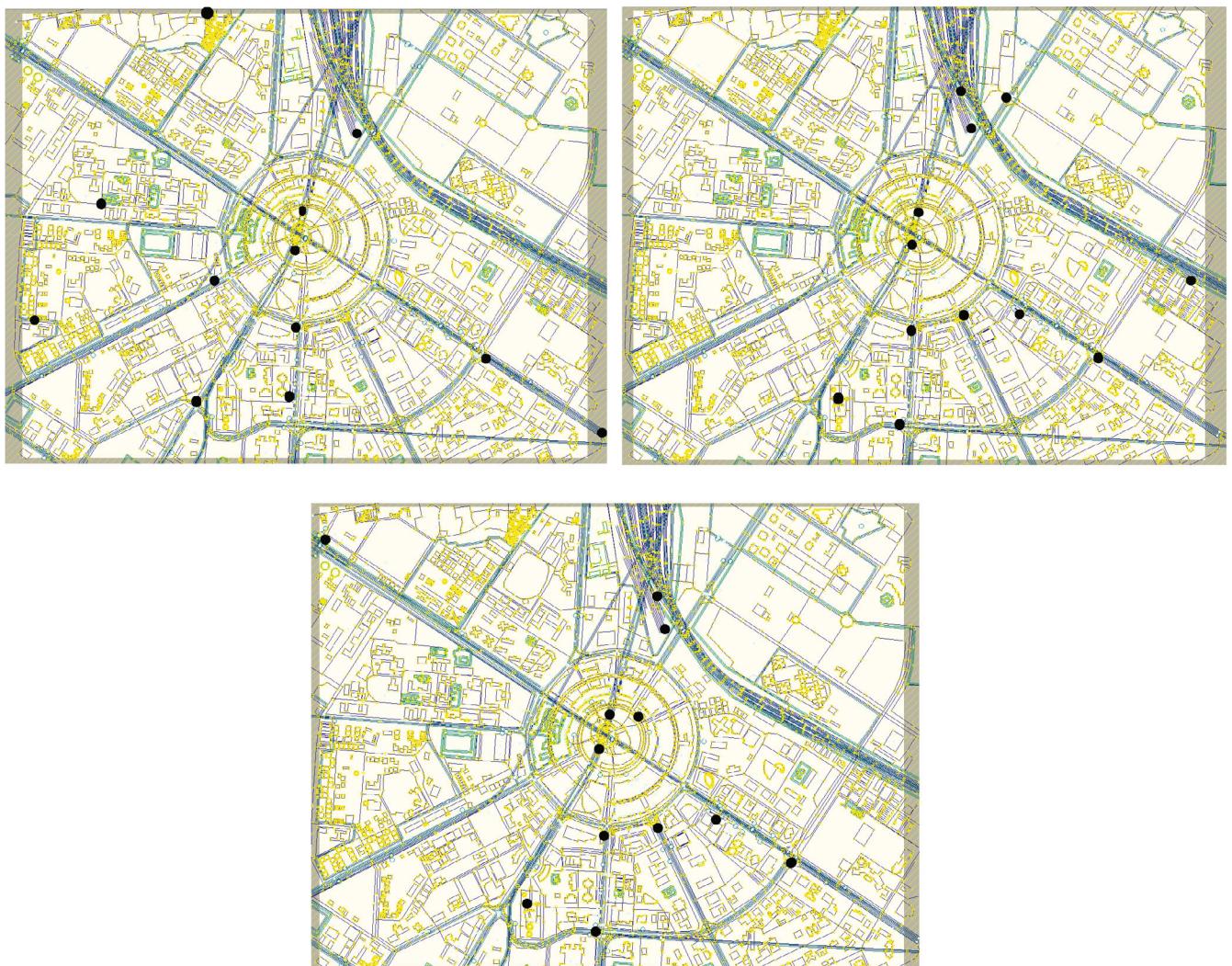


FIGURE 15 Evolution of the deployment of 12 RSUs in Connaught Place after (A) Generation 0, (B) Generation 15, and (C) Generation 30

TABLE 4 RSU placement algorithms response time

Number of generations	Response time of GA	Response time of GARRH	Response time of MARP
20	5.041	22.383	38.478
40	10.242	44.413	83.181
60	13.263	77.441	133.346
80	18.992	93.428	164.584

5.1.2 | Extracting positive and negative images

KITTI dataset is labeled so as to mine out the positive samples of images. This is done by cropping out vehicles that are our region of interest. The positive samples are cropped from KITTI dataset based on the labels provided and only the fully visible components are cropped.

The training data may also contain some images having occlusion when viewed from various angles. Further when the positive images are cropped, then they are rigorously analyzed to ascertain that they actually contain the object of

interest in varying driving conditions and illuminations. For negative image samples since the same dataset is utilized, therefore the intensity of the pixels in object of interest having vehicles is changed to 0 intensity value to blacken that region. This ensures that the negative images do not contain any vehicle component while including all the background images associated with the vehicle just as it is in the real environment.⁴⁹

The detector has multiple pretrained stages and the function classifier iterates the detection model over the negative samples before beginning to train a new stage. Any object of interest detected from this sample becomes the false positive, which are then used for next iteration training. As we proceed to add more stages, the number of false positives decreases causing difficulty for the detector to generate more negative samples. Hence, for this reason it is necessary to supply many negative images for training. In our proposed method the ratio of positive image to negative images is retained at 1:13.

5.1.3 | Cascade object detection

The cascading object detection comprises of stages, where each iteration is an ensemble of various weak learners. These are basic classifiers termed as decision stumps. Each phase is trained using boosting technique which exhibits capability of training a highly accurate classifier. The best threshold value is generated for each classifier thereby detecting the components accurately and classifying them into categories of positive and negative. Three kinds of classifiers called HOG, LBP, Haar are used for comparative analysis of detection results.⁵⁰

5.2 | Prerequisites in function parameter settings

1. **Number of phases:** The greater number of phases results in higher accuracy of the detector, while it increases the time taken to train the dataset. Every stage requires more training samples as some quantities of positive and negative images are dropped out in every iteration. Phases with lesser rate of false positives are more difficult as they have a greater number of weak learners. Phases with higher rate of false positives have less weak learners.
2. **Rate of true positives:** It is defined as number of positive images available for training every stage. The greater number of true positives intensifies the complexity of every stage thereby leading to more accurate detections though taking more training time for detection.
3. **False rate of alarm:** To achieve a considerable detection accuracy, a greater number of cascading stages are required for higher false alarm rate values. Lower values lead to increased training time and complexity.

5.2.1 | Training with HOG

Since the edges and corners, image gradients (variation in color or intensity) are useful this detector first learns the HOG-based features. These attributes possess large information about the image. HOG utilizes the distribution of the direction of some specific gradients as features. This helps to comprehend the overall object shape. In an image a histogram is computed over the pixels by selecting a bin on the basis of the direction. The gradient magnitude towards the summed value of the bin is calculated by every pixel.^{51,52} This histogram acts like a feature vector for that image. The trained detector is present in the format of .xml file which consists of information such as parameters for training and every stage's leaf and node values of the weak classifiers. Additionally, generation of some feature rectangles also takes place at every stage.

5.2.2 | Training with LBP

For AdaBoost training the same dataset is applied by conducting parameter training based on LBP features. Every pixel is recognized by a binary number which is of 8-digit that tells the intensity of the pixel and is dependent upon neighboring and the center pixel. Subsequently, the feature vectors are calculated by the binary patterns. The localized patterns are captured by LBP method. Hog takes into account only the direction where magnitude of the gradient is the highest while the LBP takes all 8 directional gradients. However, binning the magnitude of the gradients into 0 and 1 s is rough and tends to lose substantial information unlike previous method HOG which holds gradient magnitude.

5.2.3 | Haar training

Finally, the AdaBoost cascade training takes the features of Haar as the feature type.⁵³ There are several Haar features that are produced thereby making the process more time taking as opposed to the above discussed two types of features. This technique demands higher ratio of positive to negative samples. The results of object detection are shown in Figure 16A to D which is conducted by our method implementation depicts three stages. The first stage depicts the KITTI raw sample images, the second stage shows the localized objects in bounding boxes and the final stage states the detected objects with its classification. We extracted images from the KITTI dataset as raw images on which we implemented the Haar-cascade algorithm for defining the bounding boxes around the object of interest. Lastly, the final stage accurately detects the object by identifying the objects with its labels. We have assessed our proposed technique in diverse scenarios having variable illuminating settings.⁵⁴⁻⁵⁷

6 | EVALUATION ANALYSIS

Before evaluation analysis, we presented the implementation of our proposed VSDAS model implementing memetic-based framework and Haar algorithm, respectively. We intend to analyze the performance of the same through some key parameters like mean average precision,⁵⁸ precision, and recall.

6.1 | Model evaluation metrics

For detecting objects there are certain statistical measures which are adopted to measure the precision, recall of the model. mAP is mostly used to conduct the evaluation of the performance of the object detection model.⁵⁹⁻⁶⁴ The precision is calculated by taking the true positives and false positives in a frame of a video or in an image. It is calculated by the Equation (13) as follows:

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positive}}. \quad (13)$$

Different objects are evaluated by calculating mAP three methods such as HOG, LBP, and our proposed Haar-cascade procedure.⁶⁵ Table 5 after the assessment observes that if Haar technique is combined with AdaBoost, it outperforms all three approaches in terms of better precision results.^{5,13,14,66}

7 | EVALUATION OUTCOMES OF PROPOSED VSDAS

We observe from our experimental analysis that a pure GA is not able to produce better results in a difficult combinatorial search environment. To cater this, our proposed MARP mechanism employs a memetic framework by exploiting the local search space making it suitable for optimization problems. Our second proposed GARRH ([Algorithm 1](#)) uses the random restart method in its local search, while the MARP algorithm ([Algorithm 2](#)) leverages a simple hill climbing method. After our experiments, we see that MARP performs best as compared to GA, GARRH in both the locations Murlipura and Connaught Place giving higher fitness scores. Murlipura records fitness scores of 55 in 12th generation while Connaught Place records fitness scores of 180 in 16th generation when 3 RSUs were considered. In case of 6 RSUs, Murlipura records fitness scores of 95 in 12th generation whereas Connaught Place records fitness scores of 260 in 20th generation. Further, in case of 9 RSUs Murlipura accounts for 122 as its fitness score and on the other hand, Connaught Place shows 330 fitness score in 16th generation. Lastly, for 12 RSUs the Murlipura location displays 130 fitness score in 11th generation and Connaught Place records 390 in 16th generation. With this observation we can conclude that as we increase the number of RSUs MARP algorithm shows better fitness scores and this effect is more pronounced in the denser region that is, Connaught Place. Moreover, as we increase the number of generations there is a convergence



FIGURE 16 (A): Scenario 1: sample results of our method on KITTI and Panasonic datasets on each dataset, the raw samples, localized results and detected objects results are exhibited. (B) Scenario 2: sample results of our method on KITTI and Panasonic datasets on each dataset, the raw samples, localized results and detected objects results are exhibited. (C): Scenario 3: sample results of our method on KITTI and Panasonic datasets on each dataset, the raw samples, localized results and detected objects results are exhibited. (D): Scenario 4: sample results of our method on KITTI and Panasonic datasets on each dataset, the raw samples, localized results and detected objects results are exhibited

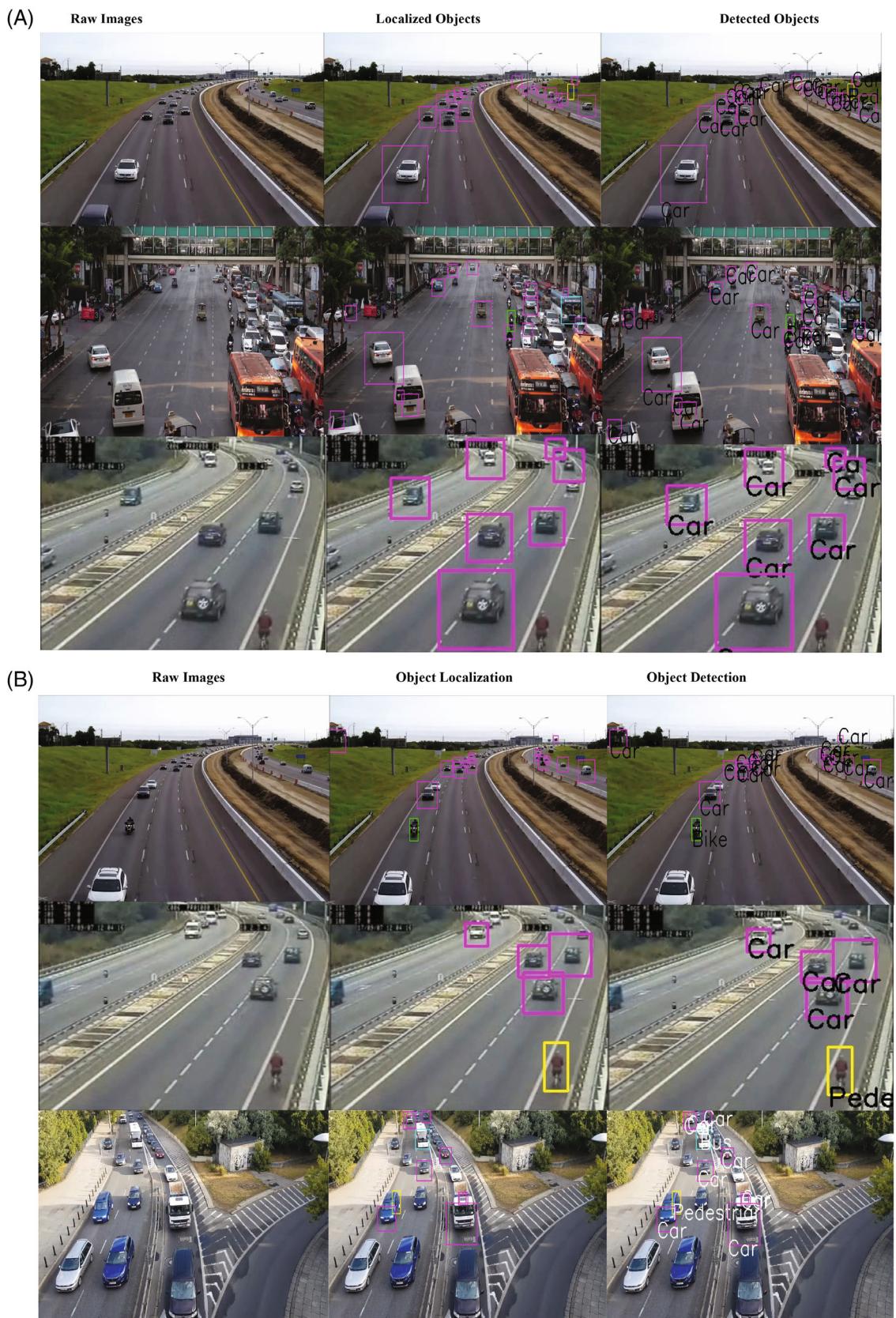


FIGURE 16 (Continued)

TABLE 5 Mean average precision of different objects for Haar, LBP, HOG

Classes/Models	Haar	LBP	HOG
Car	74.1%	69.9%	66%
Bus	89.8%	81%	83%
Pedestrian	66%	68%	62.3%
Bikes	94%	89.5%	86.4%
mAP	80.9%	77%	74.4%

Note: The bold value suggests the better performance value of haar method in comparison to LBP and HOG.

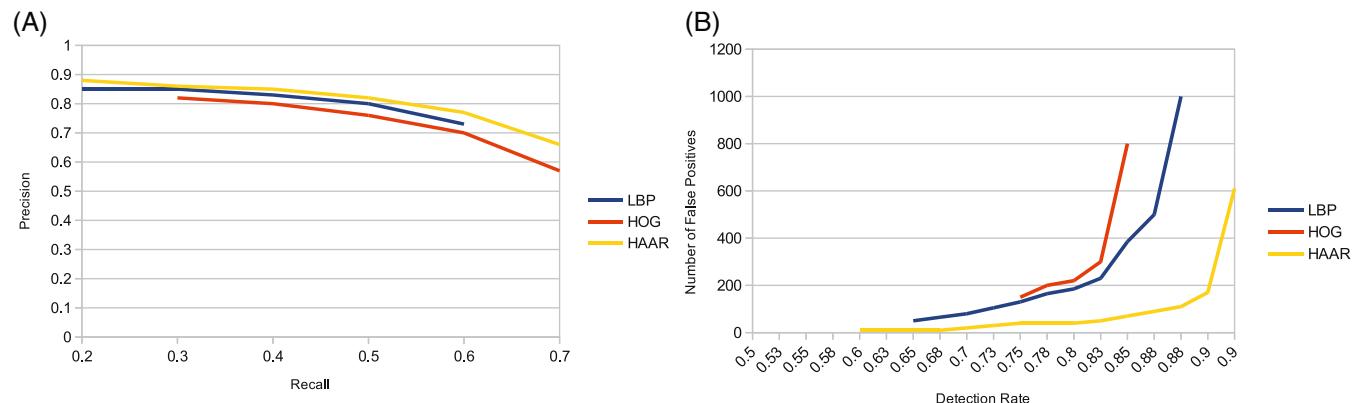


FIGURE 17 Experimental results of the obsolete feature types with the proposed Haar cascade algorithm (A) precision-recall curve, (B) false positives versus detection rate

of the fitness score towards global optimum. This higher fitness score accounts for the best possible locations in a particular geographical region with varied number of RSUs taken, as we can decide that how many number of RSUs we can deploy for the fitness function to maximize and deliver optimal result. Our goal of the article is to choose optimal locations of RSU to be deployed over a wide-range of geographical area. Therefore, we did simulations taking different number of RSUs at two different places having low and dense populace. Eventually, we are able to comprehend from our simulations that on the application of MARP algorithm employing local search technique achieves optimum results.

For the OD aspect, we proposed a latency-efficient object detection technique for self-driving cars which uses an enhanced Haar-cascade method for which we implemented a key frame detection technique for enhanced object recognition. The Figure 17A shows that our proposed method shows high precision nearing to 90% compared to existing state-of-the-art LBP and HOG with low recall. This means our method has delivered the most relevant results with correct identification of vehicular components on road in the adopted dataset and ignored the irrelevant samples. In VSDAS, it is essential for the AVs to get the most accurate results even if they are few in number with correct identification and disregard the erroneous identification to eliminate any extra diversions. Moreover, in Figure 17B, we witness a drop in the number of false positives with higher detection rate in our proposed Haar-cascade method while LBP and HOG shows higher false positives which is highly disadvantageous for AVs. Moreover, the mAP of the proposed method is highest showing 79.9% as compared to LBP ad HOG which is only 76% and 73.4%, respectively.

For assessing better detection rate, we have taken four values of the scale factor: 1.1, 1.3, 1.5, 1.7 over 10 image samples. A single set of 10 samples of images are taken over every value of scale factor. The Table 6 parameters TP, FP, FN are taken to judge that which scale factor gives better results in terms of higher precision and less recall. We can see in Table 6 which shows that scale factor having 1.1 value yields the best performance for our proposed method.

Hence, we validate our experimental work by taking different scaled images and witness a good performance in detecting objects scale factor = 1.1 on road by using our proposed Haar-cascade algorithm.

TABLE 6 Testing of rate of detection

Image	Scaling Factor	Definite total count	True positive (TP)	False positive (FP)	False negative (FN)	Recall	Precision
61	1.1	26	17	4	9	0.65	0.81
10	1.1	21	13	4	8	0.62	0.76
168	1.1	18	7	4	9	0.44	0.64
337	1.1	17	12	5	5	0.71	0.71
402	1.1	33	18	6	15	0.55	0.75
463	1.1	35	24	10	11	0.69	0.71
537	1.1	26	17	3	9	0.65	0.85
475	1.1	25	15	2	10	0.6	0.88
601	1.1	6	3	7	3	0.5	0.3
97	1.1	9	7	0	2	0.78	1
Average						0.62	0.74
61	1.3	26	4	1	22	0.15	0.8
10	1.3	21	8	2	13	0.38	0.8
168	1.3	18	3	1	15	0.17	0.75
337	1.3	17	5	1	12	0.29	0.83
402	1.3	33	10	3	23	0.3	0.77
463	1.3	35	7	2	28	0.2	0.78
537	1.3	26	12	1	14	0.46	0.92
475	1.3	25	7	0	18	0.28	1
601	1.3	6	2	4	4	0.33	0.33
97	1.3	9	1	0	8	0.11	1
Average						0.27	0.8
61	1.5	26	2	2	24	0.08	0.5
10	1.5	21	4	1	17	0.19	0.8
168	1.5	18	1	1	17	0.06	0.5
337	1.5	17	2	1	15	0.12	0.67
402	1.5	33	6	1	27	0.18	0.86
463	1.5	35	3	3	32	0.09	0.5
537	1.5	26	8	1	18	0.31	0.89
475	1.5	25	6	1	19	0.24	0.86
601	1.5	6	2	4	4	0.33	0.33
97	1.5	9	0	0	9	0	0
Average						0.16	0.59
61	1.7	26	1	4	25	0.04	0.2
10	1.7	21	1	0	20	0.05	1
168	1.7	18	0	1	18	0	0
337	1.7	17	2	0	15	0.12	1
402	1.7	33	2	0	31	0.06	1
463	1.7	35	3	0	32	0.09	1
537	1.7	26	6	0	20	0.23	1
475	1.7	25	4	0	21	0.16	1
601	1.7	6	1	0	5	0.17	1
97	1.7	9	0	0	9	0	0
Average						0.09	0.72

8 | CONCLUSION

The proposed VSDAS has presented a comprehensive system to intelligently deploy RSUs and conduct a low-latency vehicle detection. The simulation results of the proposed MARP algorithm show higher fitness scores with increase in number of RSUs (3, 6, 9) when compared to GA and GARRH. The fitness scores show an increase in the Connaught Place due to its denser road network. Since the RSU placement with greater scalability is expensive and even their maintenance cost is economically high, the random nature of the proposed MARP algorithm contributes in finding new potential deployments in larger coverage area and also by increasing the number of generations (Generation 0, Generation 15, and Generation 30). This achieves a balance between the RSU coverage and deployment costs. In addition, the performance of the upgraded Haar-cascade algorithm in detecting objects also shows a mAP of 79.9% as compared to HOG, LBP proving a significant reduction in the latency requirement of the edge IoAV system. Haar-cascade shows improved precision value of 90% as compared to HOG, LBP. Moreover, the detection rate is 90% with a smaller number of false positives as is required in the identification of vehicles.

ACKNOWLEDGMENTS

We are thankful to the anonymous reviewers who have given us all the suggestions to improve the quality of our manuscript. We are also thankful to the Editor-in-Chief of this journal for the timely processing of this article. This work is supported by CHANAKYA Fellowships of IITI DRISHTI CPS Foundation under the National Mission on Interdisciplinary Cyber Physical System (NM-ICPS) of Department of Science and Technology, Government of India.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

1. Storck CR, Duarte-Figueiredo F. A survey of 5G technology evolution, standards, and infrastructure associated with vehicle-to-everything communications by internet of vehicles. *IEEE Access*. 2020;8:117593-117614.
2. Sharma S, Kaushik B. A survey on internet of vehicles: applications, security issues & solutions. *Veh Commun*. 2019;20:100182.
3. Priyan MK, Devi GU. A survey on internet of vehicles: applications, technologies, challenges and opportunities. *Int J Adv Intell Paradig*. 2019;12(1-2):98-119.
4. Mukhtar A, Xia L, Tang TB. Vehicle detection techniques for collision avoidance systems: a review. *IEEE Trans Intell Transp Syst*. 2015;16(5):2318-2338.
5. Gao Z, Chen D, Cai S, Wu HC. Optdynlim: an optimal algorithm for the one-dimensional rsu deployment problem with nonuniform profit density. *IEEE Trans Industr Inform*. 2018;15(2):1052-1061.
6. Ni Y, He J, Cai L, Pan J, Bo Y. Joint roadside unit deployment and service task assignment for internet of vehicles (IoV). *IEEE Internet Things J*. 2018;6(2):3271-3283.
7. Liang Y, Wu Z, Tian Y, Chen H. Roadside unit location for information propagation promotion on two parallel roadways with a general headway distribution. *IET Intell Transp Syst*. 2018;12(10):1442-1454.
8. Blaschke T, Hay GJ, Weng Q, Resch B. Collective sensing: integrating geospatial technologies to understand urban systems—an overview. *Remote Sens*. 2011;3(8):1743-1776.
9. Ranft B, Stiller C. The role of machine vision for intelligent vehicles. *IEEE Trans Intell Veh*. 2016;1(1):8-19.
10. Zhang X, Li Y, Miao Q. A cluster-based broadcast scheduling scheme for mmWave vehicular communication. *IEEE Commun Lett*. 2019;23(7):1202-1206.
11. Jiang X, Yu FR, Song T, Ma Z, Song Y, Zhu D. Blockchain-enabled cross-domain object detection for autonomous driving: a model sharing approach. *IEEE Internet Things J*. 2020;7(5):3681-3692.
12. Kim H, Lee SH, Kim S. Cooperative localization with constraint satisfaction problem in 5G vehicular networks. *IEEE Trans Intell Transp Syst*. 2020;23(4):3180-3189.
13. Wu TJ, Liao W, Chang CJ. A cost-effective strategy for road-side unit placement in vehicular networks. *IEEE Trans Commun*. 2012;60(8):2295-2303.
14. Liya X, Chuanhe H, Peng L, Junyu Z. A randomized algorithm for roadside units placement in vehicular ad hoc network. Paper presented at: 2013 IEEE 9th International Conference on Mobile Ad-Hoc and Sensor Networks; 2013:193-197; IEEE.
15. Patil P, Gokhale A. Voronoi-based placement of road-side units to improve dynamic resource management in vehicular ad hoc networks. Paper presented at: 2013 International Conference on Collaboration Technologies and Systems (CTS); 2013:389-396; IEEE.
16. Fogue M, Sanguesa JA, Martinez FJ, Marquez-Barja JM. Improving roadside unit deployment in vehicular networks by exploiting genetic algorithms. *Appl Sci*. 2018;8(1):86.
17. Yang H, Jia Z, Xie G. Delay-bounded and cost-limited RSU deployment in urban vehicular ad hoc networks. *Sensors*. 2018;18(9):2764.
18. Liu C, Huang H, Du H. Optimal RSUs deployment with delay bound along highways in VANET. *J Comb Optim*. 2017;33(4):1168-1182.

19. Silva CM, Pitangui CG, Miguel EC, Santos LA, Torres KB. Gamma-reload deployment: planning the communication infrastructure for serving streaming for connected vehicles. *Veh Commun.* 2020;21:100197.
20. Lochert C, Scheuermann B, We wetzer C, Luebke A, Mauve M. Data aggregation and roadside unit placement for a vanet traffic information system. In proceedings of the fifth ACM international workshop on VehiculAr inter-NETworking; 2008:58-65.
21. Chi J, Jo Y, Park H, Park S. Intersection-priority based optimal RSU allocation for VANET. Paper presented at: 2013 Fifth International Conference on Ubiquitous and Future Networks (ICUFN); 2013:350-355; IEEE.
22. Sankaranarayanan M, Chelliah M, Mathew S. A feasible rsu deployment planner using fusion algorithm. *Wirel Pers Commun.* 2021;116(3):1849-1866.
23. Miglani A, Kumar N. Deep learning models for traffic flow prediction in autonomous vehicles: a review, solutions, and challenges. *Veh Commun.* 2019;20:100184.
24. Rawlley O, Gupta S. An upgraded object detection model for enhanced perception and decision making in autonomous vehicles. Paper presented at: 2022 IEEE International Conference on Communications Workshops (ICC Workshops); 2022:1201-1205; IEEE.
25. Zheng Z, Wang P, Ren D, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation. *IEEE Trans Cybern.* 2021;52(8):8574-8586.
26. Zhang W, Zhang Z, Chao HC. Cooperative fog computing for dealing with big data in the internet of vehicles: architecture and hierarchical resource management. *IEEE Commun Mag.* 2017;55(12):60-67.
27. Ahmed E, Gharavi H. Cooperative vehicular networking: a survey. *IEEE Trans Intell Transp Syst.* 2018;19(3):996-1014.
28. Dai P, Song F, Liu K, Dai Y, Zhou P, Guo S. Edge intelligence for adaptive multimedia streaming in heterogeneous internet of vehicles. *IEEE Trans Mob Comput.* 2021.
29. Dawkins R, Carlisle TR. Parental investment, mate desertion and a fallacy. *Nature.* 1976;262(5564):131-133.
30. Eiben AE, Smith JE. *Introduction to Evolutionary Computing.* Vol 53. Berlin: Springer; 2003:18.
31. Jiao L, Zhang F, Liu F, et al. A survey of deep learning-based object detection. *IEEE Access.* 2019;7:128837-128868.
32. Wang H, Yu Y, Cai Y, Chen X, Chen L, Liu Q. A comparative study of state-of-the-art deep learning algorithms for vehicle detection. *IEEE Intell Transp Syst Mag.* 2019;11(2):82-95.
33. Muhammad M, Kearney P, Aneiba A, Arshad J, Kunz A. RMCCS: RSSI-Based Message Consistency Checking Scheme for V2V Communications; 2021.
34. Rimol M. Gartner Forecasts More Than 740,000 Autonomous-Ready Vehicles to Be Added to Global Market in 2023. <https://www.gartner.com/en/newsroom/pressreleases/November 14, 2019-gartner-forecasts-more-than-740000-autonomous-ready-vehicles-to-be-addedto-global-market-in-2023> (дата обращения: 09.04. 2021).
35. Wang DZ, Posner I, Newman P. What could move? Finding cars, pedestrians and bicyclists in 3d laser data. Paper presented at: 2012 IEEE International Conference on Robotics and Automation; 2012:4038-4044; IEEE.
36. Azim A, Aycard O. Layer-based supervised classification of moving objects in outdoor dynamic environment using 3D laser scanner. Paper presented at: 2014 IEEE Intelligent Vehicles Symposium Proceedings; 2014:1408-1414; IEEE.
37. Behley J, Steinhage V, Cremers AB. Laser-based segment classification using a mixture of bag-of-words. Paper presented at: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems; 2013:4195-4200; IEEE.
38. Ren S, He K, Girshick R, Sun J. Faster r-cnn: towards real-time object detection with region proposal networks. *Adv Neural Inf Process Syst.* 2015;28:1-9.
39. Zhang Y, Pezeshki M, Brakel P, Zhang S, Bengio CLY, Courville A. Towards end-to-end speech recognition with deep convolutional neural networks. arXiv preprint arXiv:1701.02720; 2017.
40. Chen X, Jiao L, Li W, Fu X. Efficient multi-user computation offloading for mobile-edge cloud computing. *IEEE/ACM Trans Netw.* 2015;24(5):2795-2808.
41. da Xu L, He W, Li S. Internet of things in industries: a survey. *IEEE Trans Industr Inform.* 2014;10(4):2233-2243.
42. Khosravian A, Amirkhani A, Kashiani H, Masih-Tehrani M. Generalizing state-of-the-art object detectors for autonomous vehicles in unseen environments. *Expert Syst Appl.* 2021;183:115417.
43. Gunasekaran R, Siddharth S, Krishnaraj P, Kalaiaraslan M, Uthariaraj VR. Efficient algorithms to solve broadcast scheduling problem in WiMAX mesh networks. *Comput Commun.* 2010;33(11):1325-1333.
44. Lin CC, Deng DJ. Optimal two-lane placement for hybrid VANET-sensor networks. *IEEE Trans Industr Electron.* 2015;62(12):7883-7891.
45. Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? The kitti vision benchmark suite. Paper presented at: 2012 IEEE Conference on Computer Vision and Pattern Recognition; 2012:3354-3361; IEEE.
46. Marakby S, Loydl F Ford autonomous driving subsidiary.
47. Zhang Z, Zeng P, Pan B, Choo KKR. Large-universe attribute-based encryption with public traceability for cloud storage. *IEEE Internet Things J.* 2020;7(10):10314-10323.
48. Feng D, Harakeh A, Waslander SL, Dietmayer K. A review and comparative study on probabilistic object detection in autonomous driving. *IEEE Trans Intell Transp Syst.* 2021;23:9961-9980.
49. Soua R, Turcanu I, Adamsky F, Führer D, Engel T. Multi-access edge computing for vehicular networks: a position paper. Paper presented at: 2018 IEEE Globecom Workshops (GC Wkshps); 2018:1-6; IEEE.
50. Kiran BR, Sobh I, Talpaert V, et al. Deep reinforcement learning for autonomous driving: a survey. *IEEE Trans Intell Transp Syst.* 2021;23:4909-4926.
51. Moon S, Kim H, Hwang I. Deep learning-based channel estimation and tracking for millimeter-wave vehicular communications. *J Commun Netw.* 2020;22(3):177-184.

52. Liu K, Lee VCS, Ng JK, Chen J, Son SH. Temporal data dissemination in vehicular cyber–physical systems. *IEEE Trans Intell Transp Syst.* 2014;15(6):2419–2431.
53. Dai P, Liu K, Wu X, Liao Y, Lee VCS, Son SH. Bandwidth efficiency and service adaptiveness oriented data dissemination in heterogeneous vehicular networks. *IEEE Trans Veh Technol.* 2018;67(7):6585–6598.
54. Arnold E, Al-Jarrah OY, Dianati M, Fallah S, Oxtoby D, Mouzakitis A. A survey on 3d object detection methods for autonomous driving applications. *IEEE Trans Intell Transp Syst.* 2019;20(10):3782–3795.
55. Nguyen BT, Murphy L, Muntean GM. Energy-efficient QoS-based congestion control for reliable communications in wireless multimedia sensor networks. Paper presented at: 2018 IEEE International Conference on Communications Workshops (ICC Workshops); 2018:1–6; IEEE.
56. Wang J, Liu K, Xiao K, et al. Dynamic clustering and cooperative scheduling for vehicle-to-vehicle communication in bidirectional road scenarios. *IEEE Trans Intell Transp Syst.* 2017;19(6):1913–1924.
57. Zhang K, Wang J, Jiang C, Quek TQ, Ren Y. Content aided clustering and cluster head selection algorithms in vehicular networks. Paper presented at: 2017 IEEE Wireless Communications and Networking Conference (WCNC); 2017:1–6; IEEE.
58. Liu C, Liu K, Guo S, Xie R, Lee VC, Son SH. Adaptive offloading for time-critical tasks in heterogeneous internet of vehicles. *IEEE Internet Things J.* 2020;7(9):7999–8011.
59. Yang Y, Bagrodia R. Evaluation of VANET-based advanced intelligent transportation systems. Paper presented at: 6th ACM International Workshop on Vehicular InterNETworking; 2009:3–12.
60. Rawlley O, Gupta S. Achieving ambient intelligence in addressing the COVID-19 pandemic using fog computing-driven IoT. In: Saini K, Raj P, eds. *Advancing Smarter and More Secure Industrial Applications Using AI, IoT, and Blockchain Technology.* USA: IGI Global; 2022:56–92.
61. Mukherjee S, Gupta S, Rawlley O, Jain S. Leveraging big data analytics in 5G-enabled IoT and industrial IoT for the development of sustainable smart cities. *Trans Emerg Telecommun Technol.* 2022;e4618.
62. Ghosh D, Katehara H, Rawlley O, Gupta S, Arulselvan N, Chamola V. Artificial intelligence-empowered optimal roadside unit (RSU) deployment mechanism for internet of vehicles (IoV). Paper presented at: 2022 IEEE 23rd International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM); 2022:495–500; IEEE.
63. Watteyne T, Vilajosana X, Kerkez B, et al. OpenWSN: a standards-based low-power wireless development environment. *Trans Emerg Telecommun Technol.* 2012;23(5):480–493.
64. Dutta AK, Elhoseny M, Dahiya V, Shankar K. An efficient hierarchical clustering protocol for multihop internet of vehicles communication. *Trans Emerg Telecommun Technol.* 2020;31(5):e3690.
65. Bai Q, Nossek JA. Energy efficiency maximization for 5G multi-antenna receivers. *Trans Emerg Telecommun Technol.* 2015;26(1):3–14.
66. Guerrero-Ibañez J, Contreras-Castillo J, Zeadally S. Deep learning support for intelligent transportation systems. *Trans Emerg Telecommun Technol.* 2021;32(3):e4169.

How to cite this article: Rawlley O, Gupta S. Artificial intelligence-empowered vision-based self driver assistance system for internet of autonomous vehicles. *Trans Emerging Tel Tech.* 2023;34(2):e4683. doi: 10.1002/ett.4683