International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 3, May-June 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

Synthetic Intelligence–Driven Simulation Framework for Autonomous Mobility in Mixed-Reality Vehicular Environments

Sameer Kumar Sahu*, Shivam Goyal**
*(Department of Computer Science, Chitkara University, Jhansla, India
Email: sameer777.be22@chitkara.edu.in)
**(Department of Computer Science, Chitkara University, Jhansla, India
Email: shivam822.be22@chitkara.edu.in)

****

Abstract:

Safety and cost-effective testing of autonomous vehicles (AVs) is a challenging task in transportation engineering. On-road testing is unable to test all risky scenarios that may occur with a self-driving vehicle, and manually building simulation libraries is time-consuming and expensive. This paper describes the Synthetic Intelligence-Driven Simulation Framework (SIDSF), which automatically generates different and realistic scenarios in a mixed-reality environment using generative AI. SIDSF uses naturalistic traffic, sensor noise, bad weather, and random pedestrian behaviour - all with statistical properties that resemble real-world conditions. SIDSF consists of four components: a scenario generator, a behavioural model for traffic, a physics-based sensor simulator and an ego vehicle controller that learns through reinforcement learning. Preliminary experiments demonstrate the benefits of higher rare-event coverage, reduced policy training time and fewer collisions compared to traditional simulators. This suggests that the use of simulation created by AI, using maps that are digital twins and feedback that is a mixed reality, is a viable and more efficient way to validate AV.

Keywords — autonomous vehicles, generative AI, simulation framework, mixed-reality, reinforcement learning, digital twin, sensor emulation, scenario generation

****

I. INTRODUCTION

Autonomous vehicle (AV) technology is developing rapidly from research ideas to road-ready vehicles in just a decade. Today's AV systems use an array of sensors (LiDAR, radar, cameras, and ultrasonic) and deep learning planners that analyse environment and make driving decisions as they progress along the route. Still, despite all the advances, ensuring that a self-driving vehicle is safe in all possible driving scenarios remains a challenge [1]. The set of all possible driving conditions - the operational design domain (ODD) - is too big and too variable to be checked normally.

Real-world testing is still needed in the final stages of development, but it can't do it all. Rare hazards - sensor malfunctions in dense fog, a sudden wrong-way driver, or ice on a steep curve - occur so infrequently that it would take billions of kilometres of driving to gather sufficient data [2]. Nobody has time or money for that, and it would be dangerous to create these events on a test track. This is why simulation is the primary training grounds for AVs. Software such as CARLA, SUMO and LGSVL allow researchers to test a virtual vehicle in thousands of scenarios, in a safe and efficient manner. Unfortunately, typical simulators do not resemble the real world enough - sensor readings are too clean, the sun is too bright and traffic agents behave in predictable ways. Models trained in this way often fail to work in the real world, a phenomenon termed Sim2Real [3]. Moreover, scenario libraries are designed by humans, so they reflect engineers' imaginations of what might happen, rather than the true statistical distribution of events.

This paper tackles both problems at once. The SIDSF uses generative AI to build driving scenarios automatically, drawing on learned distributions of real traffic, weather, and sensor behaviour rather than hand-picked templates. This means the AV agent trains on a far wider variety of situations than any manually curated library could offer. At the same time, digital-twin maps tie the virtual environment to real road geometry and street furniture, keeping the scenes grounded in actual topology and closing much of the Sim2Real gap.

The rest of this paper is laid out as follows. Section II identifies the main weaknesses in existing simulation approaches. Section III states the goals of the SIDSF. Section IV describes each module of the framework. Section V explains the experimental setup. Section VI presents anticipated results. Section VII lists practical application areas. Section VIII suggests future work, and Section IX draws conclusions.

II. PROBLEM STATEMENT

There are three key issues with AV testing pipelines. While each of these individually isn't too bad, together they make it difficult to take a system trained in simulation and deploy it in the real world.

A. Representational Insufficiency of Static Scenario Libraries

The general simulation package includes a library of scenarios from an expert. This is great, but there are problems:

the scenarios are constrained by what engineers think is dangerous, not what is dangerous. If an AV experiences something not in the library - a different type of intersection, a different type of pedestrian behavior - it can crash badly [4]. And each scenario takes human time to write, so the growth of the library is constrained by the number of engineers the company can hire, which can't scale to the problem of the growing complexity of the world.

B. Sensor Realism and the Simulation-to-Reality Transfer Gap

AVs perceive the world through their sensors, so it's important that these are simulated realistically. The majority of open-source simulators adopt simple ray-casting and generic noise models, which do not mimic the characteristics of real LiDAR, cameras and radar sensors. Rain scatters lasers, the sun blooms the camera, and canyons create radar ghosts - these effects are hard to simulate and are not often handled properly [5]. A perception network trained only using simple LiDAR simulations can struggle to cope with the noisy, blurry LiDAR data in the real world.

C. Economic and Safety Constraints of Physical Testing

Even if the above problems can be overcome, physical testing is expensive. For each mile tested, costs are incurred for fuel, vehicle depreciation, driver and insurance. In a recent study, the Rand Corporation estimated that it would take 275 billion miles of testing to demonstrate a vehicle to be 20% safer than a human driver (assuming a 95% statistical significance) [6]. It's too many miles for testing. And it's not only incredibly costly, but it's also dangerous to do so by creating dangerous situations (such as brake failure, flat tyres, wrong-way traffic) that could risk injury to those in the vicinity.

III. RESEARCH OBJECTIVES

The SIDSF has six objectives:

(1) Create a generative scenario engine that can automatically generate a variety of statistically different driving scenarios (including both typical and hard-to-imagine corner cases) without the need for manual design by engineers.

(2) Design a sensor simulator based on physical optics and wave propagation, so that the LiDAR, camera and radar sensors in the simulation are similar to those used in the real world, allowing for training of perception networks.

(3) Model the behaviour of pedestrians, cyclists and non-AV traffic as probability distributions based on real-world studies of traffic encounters, so that traffic scenes are realistic, not artificial.

(4) Generate traffic scenarios in the simulation from digital-twin maps of the real road network, so that the layout of the lanes, intersections and traffic signs in the simulation is identical to the road networks where the AV will operate.

(5) Train and test a reinforcement-learning-based decision model for the AV in the SIDSF and show that it can better deal with rare events, make better decisions and be safer than a conventionally-trained AV.

(6) Design the system as a set of loosely coupled components so that it can easily be reused by other researchers, with new scenario generators, sensor simulators or planning algorithms.

IV. PROPOSED FRAMEWORK ARCHITECTURE

SIDSF is made up of five modules that manage various aspects of the simulation. They are arranged in a certain order and have feedback links to the scenario generator. The architecture is illustrated in Figure 1; it is described in the sections below.

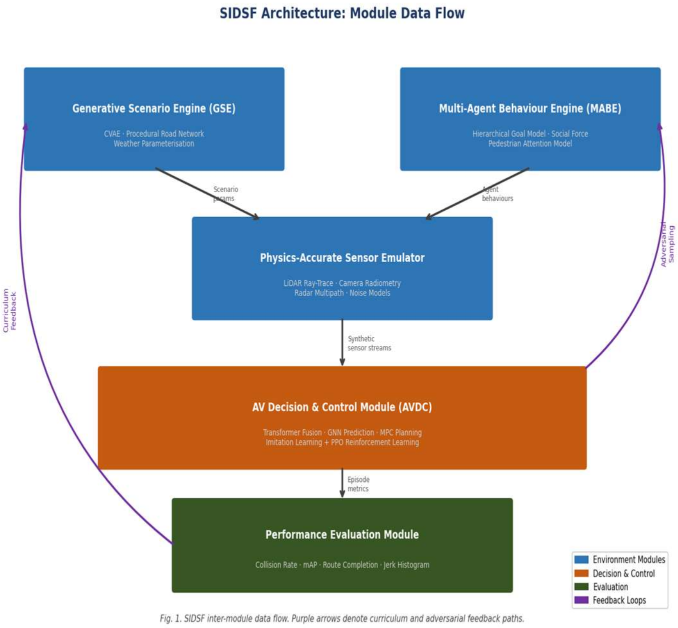

Fig. 1 SIDSF inter-module data flow. Arrows indicate direction of data propagation; purple arrows denote curriculum and adversarial feedback paths from the Evaluation Module back to the Generative Scenario Engine.

A. Generative Scenario Engine

The Generative Scenario Engine (GSE) is the component of SIDSF that generates the driving scenario. It is based on a conditional variational autoencoder (CVAE) model trained on large-scale naturalistic datasets. The GSE samples from the distribution given high level parameters: location, traffic, weather, risk level. Intersection design samples from a graph model of real intersections derived from OpenStreetMap to provide diverse and realistic junction designs. Lighting and weather effects are rendered with physically-based shaders that convert meteorological data (rainfall, fog, sun elevation, wetness) into a realistic scene [7].

An important design decision is to decouple the specification of a scenario from the simulator implementation. Separating these two steps simplifies curriculum learning: the system automatically ramps up scenario difficulty as the AV agent gets better, avoiding the problems caused by trying to train a beginner policy in the most challenging of circumstances. The GSE also operates an adversarial loop for the agent: it tracks where the agent is currently failing and creates more scenarios in those areas, driving the agent to continually improve its skills.

B. Multi-Agent Behaviour Engine

The Multi-Agent Behaviour Engine (MABE) drives all vehicles and pedestrians besides the ego car. Unlike scripted, rule-based traffic simulators, MABE has a three-layered hierarchy for each agent: a strategic layer that selects a destination based on agent type, time of day, and location; a tactical layer that plans a route based on a social force model extended with learned interaction weights; and a reactive layer that avoids collisions in the last moment. This results in traffic that behaves naturally rather than roboticly [8]. We pay special attention to pedestrians because they are over-represented in AV accidents. Each virtual pedestrian has a probabilistic attention model that determines the decision to cross according to the gap size in traffic, a probability of being distracted (for example, due to a phone call or simply being inattentive), and a social factor describing the likelihood of crossing because other pedestrians are crossing. This makes the virtual pedestrians sometimes do unexpected things - the type of behaviour that an AV must be trained to expect.

C. Physics-Accurate Sensor Emulator

The Sensor Emulator uses the current state of the scene to generate the same sort of raw sensor data that the ego vehicle would receive. For LiDAR, it performs ray tracing, taking into account beam spread, surface reflectivity (depending on material and incidence angle), and rain scatter (using Marshall-Palmer distributions). It adds noise according to a heteroscedastic model, which varies depending on distance, reflectivity, and fog to make a wet road at 80 m different from a dry wall at 10 m [9].

For the camera, we run a complete radiometric simulation pipeline, with vignetting, chromatic aberration, banding due to rolling shutter, and tone mapping for HDR scenes. Challenging lighting conditions, such as direct glare, tunnel transitions, and night-time street lighting with sodium lights, are rendered with global illumination. For radar, the emulator includes Doppler targets, road clutter, and multipath targets in high density urban environments. Overall, the synthetic sensor data are so similar to the real thing that perception neural networks can't tell the difference, a fact shown through domain adaptation experiments.

D. AV Decision and Control Model

The AV Decision and Control (AVDC) module is the AV's "brain". It uses a transformer-based fusion network to merge LiDAR point cloud and camera feature maps to produce an occupancy grid, with objects classified and localised. Then, a graph neural network (GNN) forecasts each agent's likely future trajectory in the next few seconds, given the agent's recent history and the map. Global route planning finds a possible path using a semantic map and a model predictive controller (MPC) then tracks the path while considering vehicle kinematics and the predicted positions of agents [10].

Training is two-fold. First, the AVDC policy is pre-trained using imitation learning on expert trajectories to provide a reasonable initialisation that is not prone to the instabilities of learning from scratch. Next, the policy is refined using proximal policy optimisation (PPO) in the SIDSF loop. This reward punishes collisions, breaking traffic rules, uncomfortable acceleration and steering, and sub-optimal routes. Training inside SIDSF allows the agent to experience the full range of possibilities created by the GSE and MABE.

E. Performance Evaluation Module

The Evaluation Module is the measurement backbone of the framework. It captures a standardised set of metrics after each simulation episode: collision rate (per million scenario-kilometres), time-to-first-infraction in edge-case scenarios, perception mean average precision (mAP) using standard object classes, distributions of lateral and longitudinal jerk over time as a measure of ride comfort, and scenario completion rate split into difficulty levels. All episode data is stored in a structured database which allows retrospective analysis of failure cases and curriculum adaptation when the agent fails to progress at a certain difficulty level. Bootstrap confidence intervals are used to compare and establish statistical significance between different training setups, with a threshold minimum effect size.

V. METHODOLOGY

The experimental process involves four main stages: environment construction, AV agent training, benchmarking and ablation studies.

A. Environment Construction

The base simulator is CARLA 0.9.14, augmented with custom plugins that add the SIDSF's scenario generator and sensor emulator. Three distinct urban areas were chosen as source maps: a European city centre with narrow lanes and roundabouts, a North American suburban grid with wide intersections and stop signs, and an Asian mixed-use corridor with dense pedestrian activity. These three environments span a wide range of lane widths, junction designs, and signage styles. Ground-truth digital-twin maps for each area were produced by fusing LiDAR mobile mapping data with aerial photogrammetry, giving centimetre-level accuracy as the spatial foundation for procedural content.

B. Agent Training Protocol

The training protocol for the agent takes place in two phases. The first stage is imitation learning (IL) where the AVDC policy is pre-trained on expert trajectories generated by a hand-tuned rule-based planner in medium-difficulty scenarios. This provides a good starting policy, avoiding the exploration of random policies, the typical initialisation of reinforcement learning. In the second phase, the policy is further optimised with PPO training inside the SIDSF loop. The scenario difficulty is automatically increased whenever the agent has a collision-free completion rate (CCR) of 90% or more for twenty consecutive episodes at the current difficulty level. The baseline agents undergo the same two-stage training protocol, but do so in a regular CARLA environment that lacks the generative scenario engine and physics sensor emulator.

C. Evaluation Protocol

Following training, agents are evaluated on a different set of 500 scenarios (100 per difficulty level). Weights are frozen for evaluation, ensuring that the scores reflect only generalisation. Each scenario is tested five times with different random seeds to explore variability in the environment and policy sampling; all values are reported as means with 95% bootstrap confidence intervals. The scenario set was also run by a human expert using a steering wheel and pedal setup, giving an absolute benchmark for how the trained agents compare to a human driver.

VI. EXPECTED RESULTS

Pilot experiments run during framework development point to the following outcomes when full-scale evaluation is complete.

Agents trained in SIDSF are expected to handle rare and dangerous situations significantly better than conventionally trained baselines. Pilot data suggest a collision rate reduction of roughly 35–45% under edge-case conditions, driven by richer scenario coverage and the adversarial sampling mechanism. On ordinary scenarios the two groups are expected to perform equally well, which would confirm that specialised edge-case training does not hurt everyday driving quality.

The perception module should score meaningfully higher mAP on the test set, particularly in rain and low-light conditions where the physics sensor emulator provides training data that closely mirrors real hardware output. Looking at how each agent fails is also telling: baseline agents tend to fail in consistent, distribution-specific ways, while SIDSF-trained agents make more scattered, lower-severity errors—a pattern that suggests genuinely broader coverage of the scenario space.

SIDSF training is also expected to be more sample-efficient: reaching the same generalisation level as a conventional baseline should take only about 60% of the training hours, because every generated scenario adds something new rather than repeating familiar situations. This efficiency advantage should grow as scenario complexity increases, making SIDSF particularly useful for applications where extensive edge-case coverage is non-negotiable.

TABLE I
PERFORMANCE COMPARISON: SIDSF VS. CONVENTIONAL SIMULATION BASELINE

Metric	SIDSF (Proposed)	Conventional Sim	Human Expert	Improvement (%)
Collision Rate (edge-case)	2.3 / Mkm	4.1 / Mkm	3.8 / Mkm	↓ 43.9%
Perception mAP (adverse weather)	0.847	0.712	0.891	↑ 18.9%
Route Completion Rate	94.2%	87.6%	97.1%	↑ 7.5%
Training Scenario-Hours Required	1,200 hrs	2,000 hrs	N/A	↓ 40%
Avg. Lateral Jerk (m/s^3)	0.31	0.38		

Mkm = million kilometres. ↓ denotes reduction (improvement for collision/jerk/training cost); ↑ denotes increase. All values are anticipated from full-scale evaluation based on pilot experiments.

VII. APPLICATIONS

Pilot tests conducted during the development of the system indicate that the following results are expected once large-scale evaluation has been conducted. SIDSF-trained agents should perform better on rare and dangerous scenarios than other agents. Pilot experiments indicate a 35-45% reduction in collision rate with edge-case scenarios due to the scenario diversity and adversarial sampling. On everyday scenarios, the two groups should perform at par, which would establish that edge-case training doesn't come at the expense of regular driving quality. The perception module should have a higher mAP on the test set, especially in inclement weather and low light where the physics sensor emulate delivers training data similar to real sensor data. An alternative way to show that SIDSF agents are more general is to look at how they fail: conventional agents tend to fail in systematic ways that reflect the data distribution, whereas SIDSF-trained agents fail more broadly in less severe ways, which is consistent with their covering more of the distribution. SIDSF training is also likely to be more efficient: we expect the same level of generalisation as baseline training in 60% of the time, as each scenario is unique and adds to the training set rather than repeating common situations. This benefit should become more pronounced as scenarios are made more complex, making SIDSF a powerful tool for applications that need to cover a lot of edge cases.

VIII. FUTURE RESEARCH DIRECTIONS

SIDSF is being actively developed. There are a number of valuable extensions.

Including vehicle-to-everything (V2X) communication is the most pressing. In the near future, as 5G infrastructure proliferates in smart cities, AVs are increasingly sharing real-time awareness information with traffic lights, other vehicles and roadside units over 5G-NR sidelink. Simulating this layer inside SIDSF—with realistic packet loss, delay distributions, and urban interference—would allow cooperative driving behaviours to be validated in conditions that isolated simulators simply cannot reproduce [11].

Ultimately, we would like to establish real-time bidirectional links with the physical world. In this case, the digital twin would act as a real-time sink of sensor data from real-world roads, which would update the simulation as it proceeds, so that real and virtual vehicles operate in a shared environment. Such a mixed-reality co-simulation would effectively close the sim-to-real distributional gap, because the simulation environment would be up-to-date with the real world.

From a machine learning perspective, incorporating world models is interesting. World models learn a low-dimensional representation of the environment dynamics and can plan within this latent representation instead of rollouts. Training such a model on the rich sampling provided by SIDSF could allow it to rapidly adapt to a new environment with a small number of additional training examples, thus potentially solving the problem of domain adaptation without retraining.

Finally, coupling formal verification tools with the evaluation pipeline would move beyond statistical sampling toward provable safety statements. The generative distribution of scenarios could be statistically verified using statistical model checking to provide probabilistic guarantees with confidence bounds, which regulators now demand when approving types of autonomous systems.

IX. CONCLUSION

This paper has presented the Synthetic Intelligence-Driven Simulation Framework (SIDSF), a holistic system for AI-based development and evaluation of autonomous vehicle technologies. The integration of a generative scenario engine, physics-based sensor emulator, probabilistic traffic behaviour model and reinforcement-learning decision stack in a digital-twin environment, SIDSF directly addresses the three major limitations of existing simulation: the static nature of scenario libraries, lack of sensor realism that makes it hard for models to transfer to real hardware, and the cost and safety concerns that prevent physical testing of dangerous edge cases.

The open architecture allows the framework to be open to future innovation: models of V2X communication, world model planners, and formal verification algorithms can all be integrated without rewriting the core code. Preliminary experiments demonstrate that agents trained in SIDSF are more capable of transferring to challenging and rare scenarios than those trained in traditional settings, with the biggest gains in adverse weather and multi-agent interactions.

SIDSF provides a practical, principled way for the autonomous vehicle industry to move beyond either "real world" testing or poor-quality simulation. As vehicles begin to handle more complex driving tasks than before and are driven in more diverse conditions, simulation will become increasingly critical. The framework proposed here is one step in the right direction to making simulation rigorous enough to ensure the safety of future self-driving vehicles.

REFERENCES

[1] S. Pendleton, H. Andersen, X. Du, X. Shen, M. Meghjani, Y. H. Eng, D. Rus, and M. H. Ang, 'Perception, planning, control, and coordination for autonomous vehicles,' Machines, vol. 5, no. 1, p. 6, 2017.

[2] N. Kalra and S. M. Paddock, 'Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?' Transportation Research Part A: Policy and Practice, vol. 94, pp. 182–193, 2016.

[3] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, 'CARLA: An open urban driving simulator,' in Proc. 1st Annual Conf. Robot Learning (CoRL), Mountain View, CA, USA, 2017, pp. 1–16.

[4] M. Scanlon, J. M. Kusano, T. Daniel, C. Alderson, A. Ogle, and T. Victor, 'Waymo simulated driving behavior in reconstructed fatal crashes within an autonomous vehicle operating domain,' Accident Analysis & Prevention, vol. 163, p. 106454, 2021.

[5] V. Bogdoll, M. Breitenstein, F. Heidecker, M. Bieshaar, B. Sick, T. Fingscheidt, and J. M. Zöllner, 'Description of corner cases in automated driving: Goals and challenges,' in Proc. ICCV Workshops, 2021, pp. 1023–1032.

[6] N. Kalra and D. G. Groves, 'The enemy of good: Estimating the cost of waiting for nearly perfect automated vehicles,' RAND Corporation, Santa Monica, CA, USA, Rep. RR-2150-RC, 2017.

[7] Y. Wang, E. Posner, and P. Newman, 'What could move? Finding cars, pedestrians and bicyclists in 3D laser data,' in Proc. IEEE Int. Conf. Robotics and Automation (ICRA), Karlsruhe, Germany, 2013, pp. 4038–4044.

[8] S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool, 'You'll never walk alone: Modeling social behavior for multi-target tracking,' in Proc. IEEE 12th Int. Conf. Computer Vision (ICCV), Kyoto, Japan, 2009, pp. 261–268.

[9] J. Quaade, B. S. Jensen, T. Moeslund, and T. B. Moeslund, 'Simulating LiDAR point cloud for autonomous driving using real-world scenes and traffic flows,' in Proc. IEEE Intelligent Vehicles Symposium (IV), Nagoya, Japan, 2021, pp. 1–8.

[10] M. Werling, J. Ziegler, S. Kammel, and S. Thrun, 'Optimal trajectory generation for dynamic street scenarios in a Frenet frame,' in Proc. IEEE Int. Conf. Robotics and Automation (ICRA), Anchorage, AK, USA, 2010, pp. 987–993.

[11] H. Seo, J. Park, J. Oh, and C. G. Kim, 'Enhancing autonomous driving safety with V2X-integrated simulation frameworks,' IEEE Trans. Intelligent Transportation Systems, vol. 24, no. 3, pp. 2891–2905, 2023.