

Real-Time High-Speed Autonomous Driving Via Lidar-Event Camera Fusion Using Attention-Enriched Spatio-Temporal Graph Neural Networks

Dr. Amol Raoji Madane¹, Dr. Ho Chin Kuan²

¹Post-Doc fellow, Asia Pacific University of Technology, Jalan Teknologi 5, Taman Teknologi, Malaysia, 57000

²Vice Chancellor and Post-Doc Research Guide, Asia Pacific University of Technology, Jalan Teknologi 5, Taman Teknologi Malaysia

Abstract

In autonomous vehicle driving, the use of diverse sensors to recognize surrounding vehicles at medium and long distances is essential for enhancing perception and ensuring safe high-speed maneuvers. Frame-based cameras suffer from motion blur and latency during rapid movements, while LiDAR, though highly accurate in 3D spatial mapping, has limited temporal resolution and struggles to detect small or low-reflectivity objects. To address these constraints, this research proposes a fusion approach that integrates LiDAR with event-based camera data to improve detection and execution of high-speed autonomous maneuvers. Event-based cameras asynchronously capture pixel-level brightness changes with microsecond latency, generating motion-rich, blur-free information that is ideally suited for dynamic environments. The data is preprocessed using noise reduction and normalization techniques to isolate relevant objects and filter out background noise. The proposed framework projects LiDAR point clouds into image space, aligns them with event data, and converts the fused information into bird's-eye view (BEV) coordinates for robust object detection and tracking using a Bacterial Colony-Driven Attention-Enriched Spatio-Temporal Graph Neural Network (BC-AST-GNN) algorithm, which is applied to model spatio-temporal relationships between objects, combining event-driven motion cues with LiDAR's 3D localization. The model further incorporates a state-estimation mechanism to refine object trajectories between LiDAR scans, ensuring stable tracking and predictive accuracy in high-speed environments, with the resulting perception outputs directly supporting a decision-making module that anticipates object interactions and selects safe and efficient maneuvers such as obstacle avoidance, lane changes, and emergency braking. Experimental results demonstrate that the BC-AST-GNN-based fusion approach significantly enhances detection accuracy, achieving 97.12% for Cars, 91.74% for Vans, 90.05% for Trucks, 92.63% for Pedestrians, and 94.02% for Cyclists, outperforming the baseline method in all categories. These advancements contribute directly to safer, more agile, and more reliable autonomous vehicle operation.

Keywords: Autonomous Vehicles, Bacterial Colony-Driven Attention-enriched Spatio-Temporal Graph Neural Network (BC-AST-GNN), Event-Based Camera, High-Speed Maneuvers, Object Detection and Tracking.

1. Introduction

Perception tasks involve a variety of sensors, including vision-based sensors for appearance cues and LiDAR for 3D spatial mapping. High-speed maneuvers like lane changes, obstacle avoidance, and emergency braking, self-driving cars depend on accurate and timely perception of their environment to enable safe navigation [1]. To achieve this objective, one must be able to identify, monitor, and forecast the movement of objects in the vicinity in a variety of environmental circumstances [2]. Road boundaries can be accurately located using LiDAR, but it has trouble with temporal resolution and dynamically shifting conditions [3]. Although vision-based sensors might detect appearance cues, they are susceptible to delay and motion blur. A promising substitute was provided by event-based cameras, which produce motion-rich data while collecting asynchronous luminance variations with millisecond latency [4-5]. Figure 1 shows LiDAR and event-based camera fusion enabling precise autonomous vehicle maneuvers.

Monitoring, identifying items, and semantic division in automatic driving have all made substantial use of traditional machine learning (ML) and deep learning (DL) models [6]. Convolutional neural networks (CNNs) and frame-based cameras are used in methods that obtain excellent accuracy under controlled conditions, but they degrade at fast speeds because of motion blur and excessive data latencies [7]. Performance is limited in small, fast-moving, or poorly reflecting environments by LiDAR-based DL pipelines, which provide spatial localization but fail to pick up on fine-grained temporal cues. Although synchronizing frames restricts performance in high-speed situations [8-9]. A fusion architecture was used for rapid autonomous maneuvering that combines LiDAR and event-based camera data. The system allows for integrated dynamical modeling by aligning LiDAR pinpoint clusters with incident sources.

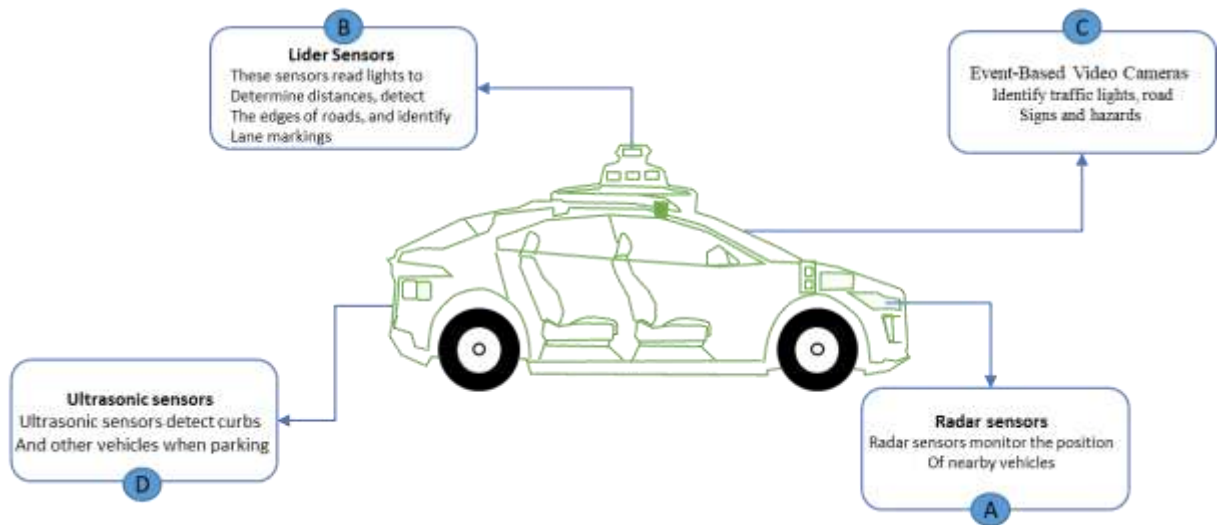


Figure 1: Sensor fusion framework for high-speed autonomous maneuvers

The research's aim is to improve object trajectories between scans by the BC-AST-GNN approach to record object associations. The suggested approach is used to enhance high-speed autonomous vehicle perception and maneuver execution by fusing LiDAR and event-based camera data. The research is organized as follows: Part 2 reviews related work on autonomous perception and sensor fusion. Part 3 covers the dataset, preprocessing, and BEV-based LiDAR–event fusion. Part 4 presents the results and discussion, and Part 5 concludes with future directions.

2. Related work

For object identification and monitoring, the suggested [10] Artificial Intelligence (AI) technique combined event-driven processing with frame-specific sensors. It included continuous detection for interface updates and dynamic boundary-box refining. With Higher Order Tracking Accuracy (HOTA) rising from 56.6% (frames alone) to 64.9% at 24 Hz and from 51.3% to 63.1% at higher temporal rates up to 500 Hz, the results demonstrate enhanced tracking performance. The AI approach incorporated a continuous detection pathway into Car Learning to Act (CARLA) and Robot Operating System (ROS), trained on the GEN1 dataset, and modified the You Only Look Once version (YOLOv8) for historical signals. Effectiveness was improved [11] for simulation situations by transfer learning. The results

reveal that it was feasible to use event cameras in vehicular perceptual methods with a mean average precision (mAP) of 70.1%, exceeding prior standards.

Event-based sensor-based object tracking is the main focus of the creation. Two techniques were used [12]: a YOLOv4-based detection with Kalman filtering and Hungarian allocation, and portable independent model trackers for device-embedded. With 82.88% mAP0.5 and

62.7% mAP0.5, respectively, YOLOv4 and YOLOv4-tiny demonstrated precise and timely tracking capabilities on a range of hardware configurations.

A variety of sensors, especially event-oriented Dynamic Vision Sensor (DVS), were investigated to assess [13] safety hazards in self-driving cars. CARLA was used to simulate seven different assault types, each of which was countered by customized barriers and an anomaly detector. While DVS demonstrated more resilience, the outcomes also revealed over 95% detection accuracy, full prevention of global positioning system (GPS) deception, and significant diminution of RGB and distance assault.

Event cameras are being investigated for the assessment and control of driverless ground vehicles. Ryder, a brand-new 1/10th scale device, was created [14] to evaluate their efficacy. Real-time control decisions were guided by event cameras' minimal latency, high dynamic flexibility, and high spatial quality. The findings solve the shortcomings of traditional cameras and show their applicability for self-navigating vehicles.

By adding event cameras and addressing the extrinsic calibration problem, the research [15] was enhanced by LiDAR-camera fusion. An automated checkerboard-based approach was

proposed to incorporate event-frame reconstruction and a coarse-to-fine optimization

strategy. When tested on 19 datasets, the method outperformed existing approaches and achieved globally optimal estimation, ensuring robust multimodal perception under varying lighting conditions.

A LiDAR-based system for autonomous cars to anticipate 3D scenes and identify irregularities was presented in the experiment [16]. A Markov Jump Particle Filter (MJPF) inferred upcoming data, while a 3D constrained encoding and decoding algorithm with transferred learning extracted traits from point clouds. Experiments conducted in controlled settings [18] verify that anomaly detection was more successful than other predictive algorithms.

2.1. Problem statement

Event-based approaches minimize motion blur and latency and increase tracking precision in dynamic driving applications [10]. They, however, have trouble giving credible spatial depth that was essential to high-speed maneuvers. LiDAR provides high-quality 3D

localization, but with low temporal resolution, limiting operating capability in dynamic conditions [16-19]. To fill these gaps, the proposed BC-AST-GNN combines event motion cues and LiDAR data, which guarantees strong spatio-temporal modeling and trajectory

refinement as well as predictive maneuvering.

3. Methodology

The process starts with gathering data from event-based camera sensors and synchronized LiDAR. Preprocessing uses wavelet denoising to reduce noise in event streaming and Z-score normalization to normalize characteristics [20]. For spatial consistency, LiDAR point clouds were projected into BEV dimensions after being filtered. After that, event-based motion cues are combined with the aligned BEV features to provide reliable object detection, trajectory refinement, and precise decision-making for fast autonomous maneuvers. Figure 2 shows high-speed autonomous driving via sensor fusion.

3.1. Dataset

A high-speed vehicle sensor fusion dataset was collected from Kaggle. This dataset contains multi-camera images collected for high-speed autonomous vehicle maneuver analysis. Each scenario includes synchronized views from six cameras, covering the full 360° surroundings of the vehicle. The dataset supports research in sensor fusion, object detection, tracking, and decision-making for safe and efficient driving maneuvers.

Source: <https://www.kaggle.com/datasets/zara2099/high-speed-vehicle-sensor-fusion-dataset/data>

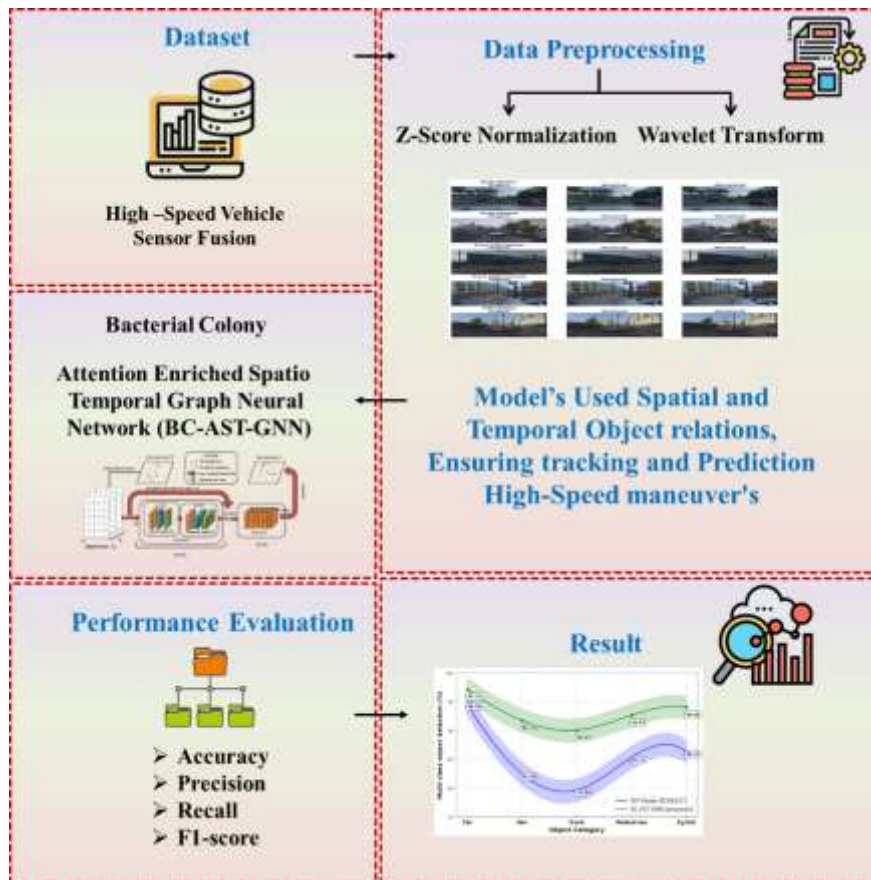


Figure 2: Flow diagram of Lidar and Event-Based Camera Data Fusion

3.2. Data preprocessing

3.2.1. Z-score normalization:

For the purpose of improving the reliability of fusion techniques and guaranteeing consistent scaling of sensor data, normalization is applied [21] to the elements that were extracted from LiDAR and event-based cameras. By altering attribute values according to their average and variance, this approach to statistics reduces the impact of outliers by Equation (1):

$$v' = \frac{v - \mu}{\sigma} \quad (1)$$

Where v' is the normalizing standardized value with a unit variation and zero mean. μ and σ stand for the attribute's average and standard deviations, respectively.

3.2.2. Wavelet transform denoising: It is used to improve component clarity for reliable high-speed navigation and precise object identification by [25] suppressing high-frequency noise while maintaining motion edges in camera events and LiDAR inputs. A nonlinear data processing method for multi-resolution evaluation in autonomous driving is the wavelet transform. By adjusting the window size to bandwidth for extremely moving instances, it extracts highly detailed temporal and spatial data from LiDAR point clusters and time-based camera streaming. Equation (2) was applicable to define the transformation for 1-D (dimensional) signals, such as periodic event streams:

$$\varphi(a) = \tan^{-1} \left[\frac{JnWf(b_{q,a},a)}{QfWf(b_{q,a},a)} \right] \quad (2)$$

where the actual value of the dimension elongation variable a with respect to the ridges position at location b is represented by $b_{q,a}$. Equations (3–4) can be used to represent the 1-D wavelet transforms of $e(w)$, $Wf(b, a)$ and the expanded wavelet $\Psi(w)$:

$$Wf(b, a) = \frac{1}{\sqrt{b}} \int_{-\infty}^{\infty} e(w) \Psi^* \left(\frac{w-a}{b} \right) cw \quad (3)$$

$$\Psi(w) = \frac{1}{\pi^{\frac{1}{4}}} \left(\frac{2\pi}{\gamma} \right)^{\frac{1}{2}} \exp \left[-\frac{\left(\frac{2\pi}{\gamma} \right)^2 w^2}{2} + i2\pi w \right] \quad (4)$$

Equation (5) gives the phase $\varphi(b)$ for 2-D WT (such as LiDAR range images or fused temporality mapping):

$$\varphi(a) = \tan^{-1} \left[\frac{JnWf(a, b_{q,a}, \theta_{q,a})}{JnWf(a, b_{q,a}, \theta_{q,a}) - e_p \left(\frac{\pi}{n} \right) \exp \left(-\frac{\pi^2}{n} \right)} \right] \quad (5)$$

where a is the 2-D translating associated to the wavelet's location, $b_{q,a}$ and $\theta_{q,a}$ are the scale and rotational parameters about the ridge point, θ is the rotation angle, and $b < 0$ is the scale component (wavelet width). Equations (6–7) can be used to express the 2-D wavelet transform of $e(w)$ and the resulting altered wavelet:

$$Wf(a, b, \theta) = b^{-m} \int_{-\infty}^{\infty} e(w) \Psi^* [b^{-1} q_{-\theta} (w - a)] c^2 w \quad (6)$$

$$\Psi(w) = \exp(-n|w|^2 + i2\pi w) \quad (7)$$

Where $q_{-\theta}$ is the standard 2×2 rotational matrix, m is a window-adjusting parameter, and $w = (w, z)$ are the 2-D coordinates for particular images (from LiDAR or event maps). The wavelet component at position a , scaled b , and rotational θ is represented as $Wf(a, b, \theta)$. $e(w)$ is the original signal, such as event streaming or LiDAR readings. The assessing filter is the wavelet value's Ψ^* imaginary conjugate. The conversion is performed using the parent wavelet, $\Psi(w)$. Due to the spatial localization provided by the $\exp(-n|w|)^2$ Gaussian wrapping, the wavelet is cognizant of local changes. To record spectral details, variations are introduced via the $i2\pi w$ compound exponential. Figure 3 shows the preprocessing of LiDAR and event-based camera images for autonomous maneuvers.

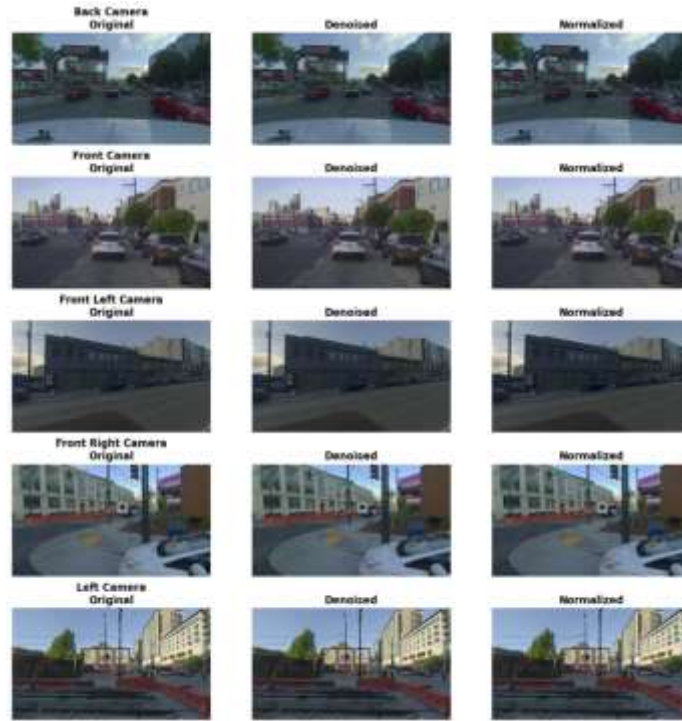


Figure 3: Back Camera Image Processing – Original, Denoised, Normalized

3.3. Birds-eye view (BEV)

By projecting LiDAR and events-camera data on a horizontal surface, the BEV presentation makes multi-object localization easier, streamlines spatial connections, and permits precise tracking and detection for fast moves [22]. Combining LiDAR and event-related camera data with BEV modification to enhance high-speed self-navigating maneuvers allows for accurate topographic perception of the surroundings. Coordinates that have been normalized: Using the camera-intrinsic matrix L , imaging values $p_i = (u, v, 1)^T$ are transformed into normalized coordinates $p_n = (X_n, Y_n, 1)^T$, given by Equation (8).

$$p_n = L^{-1}p_i \quad (8)$$

where p_n stands for parameters in a normalizing camera space, u, v are the original picture pixel coordinates, and L is the image-intrinsic matrix that contains focal durations and the main position. Finally, the BEV coordinates are projected onto the bird's eye view plane $p_b = (x, y, 1)^T$ using scaling factors and principal point correction, given by Equation (9).

$$\lambda p_b = \begin{bmatrix} S_x & 0 & ct_x \\ 0 & S_y & ct_y \\ 0 & 0 & 1 \end{bmatrix} \quad (9)$$

where ct_x, ct_y stand for the fundamental position deviations on the BEV plane, λ is a scaling factor, and S_x and S_y are scalability ratios that corresponds to the image's elevation and breadth. In contrast to the preceding process, the inverse transformation locates a point $p_i = (u, v, 1)^T$ in the original image that correlates to a specified location in bird's eye perspective, $p_b = (x, y, 1)^T$. Since both pre- and post-transformed data are stored in a structured data container, the inverse transformation can be efficiently implemented using a lookup table (LUT). The LUT stores precomputed correspondences between p_i and p_b , avoiding time-consuming matrix multiplications for every pixel, and allows direct projection of color and

intensity data. The CNN input is then created by combining the converted bird's eye view image with LiDAR point cloud data [23]. Two stages are taken to transform a disordered 3D point system into an organized 2D representation that can be processed by a CNN:

1. The point cloud's $x - y$ plane is projected onto a 2D grid, where each cell corresponds to a pixel. The cell size is set to $0.1 \times 0.1 \text{ m}$, covering a region $x \in [6,46]$ and $y \in [-10,10]$, resulting in a 200×400 -pixel image.
2. Four-channel features are computed for each cell: maximum height, number of points, maximum intensity, and altitude difference from surrounding cells, which are stored in image format.

This structured representation of fused BEV images and point cloud geometric features allows the CNN to accurately learn spatial relationships, improving perception and decision-making for safe high-speed autonomous maneuvers.

3.4. BC-AST-GNN

BC-AST-GNN was utilized to attain the goal of robust and accurate perception of high-speed autonomous maneuvers. It also resolves motion cues as well as 3D space by modeling spatio-temporal relationships between the objects presented in the fused LiDAR and event-based camera data. The attention mechanism identifies major dynamic objects, and the optimization of learning, which is driven by a bacterial colony, makes it stable. This allows accurate tracking of objects, prediction of trajectories, and the timely assistance of safe and agile decisions during the swift maneuvering of the vehicles.

3.4.1. AST-GNN

An innovative BC-AST-GNN framework is presented for maneuver execution employing LiDAR and event-driven camera data to overcome the shortcomings of traditional fusion techniques in collecting changing conditions during high-speed self-driving. The encoder consists of two modules: (1) an attention-based temporal graph module to represent temporal dependencies across motion sequences, and (2) an attention-oriented spatial graph module to capture spatial interactions among objects from fused LiDAR–event data. To facilitate informed choices for safe high-speed navigation, the decoder serves as a trajectory and movement predictor. For M surrounding vehicles observed across S_s time steps, the fused representation is denoted as Tin with dimensions $(M \times S_s \times D)$, where C includes both spatial coordinates and motion cues derived from LiDAR–event fusion. The spatial and temporal interactions of adjacent vehicles must be accurately represented to facilitate safe and dependable high-speed movements. A spatial graph Gt that depicts the relative statuses of the cars in the scene is created from combining sensor information gathered from LiDAR and event-driven cameras at each time step s .

$H_s = \{U_s, F_s\}$, where the interacting edges are represented by $\{f_s^{ji} \mid \forall j, i \in \{1, \dots, M\}\}$ and the nodes associated to the encircling vehicles are represented by $U_s = \{u_s^j \mid \forall j \in \{1, \dots, M\}\}$.

Each node u_s^j feature vector is generated using event-based movement cues and combined LiDAR point-cloud measurements.

• Graph Attention Mechanism for Adaptive Interaction Learning

This process continuously modulates attention weights to concentrate on important vehicle interactions, which allows the model to deal with the rapidly varying traffic conditions when performing high-speed maneuvers [24]. Instead of depending only on Euclidean distances, a Graph Attention Network (GAT) is used to continually build a weighted proximity matrix B_s to reflect the continuous influence across vehicles. The combined features are projected onto higher dimensions using the linear alteration expressed in equation (10).

$$e_{conv}^s = \{g_j^s \mid \bar{g}_j^s \in \mathbb{R}^{F \times F} \mid \mathbb{R}^F \quad (10)$$

where g_j^s is the vehicle's fused characteristic at time s . Equation (11) is used to calculate the attention factor $\alpha_{ji}^{\rightarrow s}$, which represents the impact of vehicle i on vehicle j .

$$\alpha_{ji}^{\rightarrow s} = \frac{\exp(\text{LeakyReLU}(b^S[xh_j^{\rightarrow s} || xh_i^{\rightarrow s}]))}{\sum_{l \in M_j} \exp(\text{LeakyReLU}(b^S[xh_j^{\rightarrow s} || xh_l^{\rightarrow s}]))} \quad (11)$$

Where the impact of vehicle i on vehicle j at time t is measured by $\alpha_{ji}^{\rightarrow s}$. M_j stands for node j 's neighborhood, $||$ for concatenation, and $b \in \mathbb{R}^{2\mathbb{F}}$ for the attention weighted vectors. Accurate estimation of trajectory and movement executed during high-speed navigation is supported by the learned adjacency weights, which enable the system to concentrate on vehicles bearing a substantial amount of influence (such as rapidly approaching or closely separated objects).

• Spatial GNNs for Vehicle Interaction

Spatial graph networks include relative distances and orientation of LiDAR-event fusion, which provides the accurate modeling of car interactions of the surrounding space. The format of the fused input tensor T_{in} is $(M \times S_o \times D)$, as previously mentioned, where M is the number of surrounded cars, S_o is the number of monitoring images, and D is the feature channels, which include 3D location, velocity, and event-related movement cues. The raw tensor is first run through a convolutional layer with a kernel count of 1×1 to create improved feature maps, or e_{conv}^s , that are used to learn significant spatial interaction information. To depict the degree of collaboration among vehicles at time step s , an estimated matrix of adjacency B_s is created using the GAT. Equation (12) is then used to carry out the spatial graph procedure:

$$e_{graph}^s = \sigma \left(\Lambda_s^{-\frac{1}{2}} \hat{B}_s \Lambda_s^{-\frac{1}{2}} e_{conv}^s X_s \right) \quad (12)$$

The spatial graph feature at time step s is represented by the e_{graph}^s , the adjacency matrices with self-looping is represented by $\hat{B}_s = B_s + J$, the diagonal degree vector of $\hat{B}_s$ is represented by Λ_s , a trainable linear transform is represented by X_s , and the stimulation coefficient (PReLU) is represented by σ .

• Temporal Graph Representation of a Vehicle

Encodes motion history combining LiDAR spatial states and event dynamics. A temporally graph H_m is created for the m -th vehicle with the purpose to simulate its mobility across time $H_m = \{U_m, F_m\}$. With the node set $U_m = \{u_j^m | \forall j \in \{1, \dots, S_o\}\}$ representing the car's states at various time stages and the edge set $F_m = \{f_{ji}^m | \forall j, i \in \{1, \dots, S_o\}\}$ encoding the periodic relationships among these stages, the temporal graph simulates the motion development of a single vehicle across observation frames. Weighted adjacency matrix B_m , built with the GAT technique, continually captures periodic interactions such as speeding, stopping, and steering, while every node feature correlates to the fused trajectory elements (w_j^m, z_j^m) derived from LiDAR and event-driven camera input.

• Temporal GNN

It represents the evolving temporal patterns for stable high-speed vehicle tracking. A 1×1 convolution is used to analyze the spatial graph characteristics e_{graph}^s that S-GNN has retrieved, resulting in temporal convolutional component maps e_{conv}^m . The temporal network H_m and associated relation matrix B_m are created using these. Equation (13) defines the temporal graph function.

$$e_{graph}^m = \sigma \left(\Lambda_m^{-\frac{1}{2}} \hat{B}_m \Lambda_m^{-\frac{1}{2}} e_{conv}^m X_m \right) \quad (13)$$

Where the temporal graph parameter of the m – th vehicle is denoted by e_{graph}^m , an adjacency array with self-loops is $B_m = B_m + J$, the horizontal degree dimension of B_m is Λ_m , X_m is a configurable linear conversion, and the symmetrical transformed Laplacian $\Lambda_m^{-\frac{1}{2}} \hat{B}_m \Lambda_m^{-\frac{1}{2}}$ is used to preserve the range of feature value.

• Trajectory Prediction Model

It makes future vehicle predictions based on fused spatial-temporal features that in turn assist in proactive maneuvers such as lane changes, obstacle avoidance, and emergency braking. As illustrated in Figure. 4, the encoder is utilized to extract time-spatial embeddings $\tilde{U}$ from the fused LiDAR and event-driven data tensor J_{in} . The spatial insertions e_{graph}^s , which depict inter-vehicle interactions at each time step, are first produced using the attention-based spatial graph module. The temporal embedded e_{graph}^m , represent the dependencies in vehicle motion over time, are then extracted via the attention-focused behavioral graph generator. To maintain low-level motion cues, a residual connection is provided between these two graph modules' input and output. Subsequent vehicle trajectories are then estimated using the direction forecasting method and the consequent embedding $\tilde{U}$.

• Time-Extrapolator Convolutional Neural Network

It extrapolates trajectories forward, reducing reaction latency in maneuvers. TXP-CNN functions along the temporal dimension after receiving the time-space graph embedding $\tilde{U}$ of the nearby vehicles. First, the embedding is reshaped to $(M \times S_e \times D_e)$, where the coefficients of a bidirectional Gaussian pattern are represented by $D_e = 5$. A final convolutional layer generates output features S_{out} of type $(M \times D_e \times e)$, where $S_e = 12$ was the prediction horizon, after five 1D convolutional layers with kernel size 3 and PReLU activity retrieve temporal characteristics. To allow high-speed maneuver planning, the outcome is reshaped to $(M \times S_e \times D_e)$ and combined with the multivariate Gaussian to anticipate upcoming car trajectories.

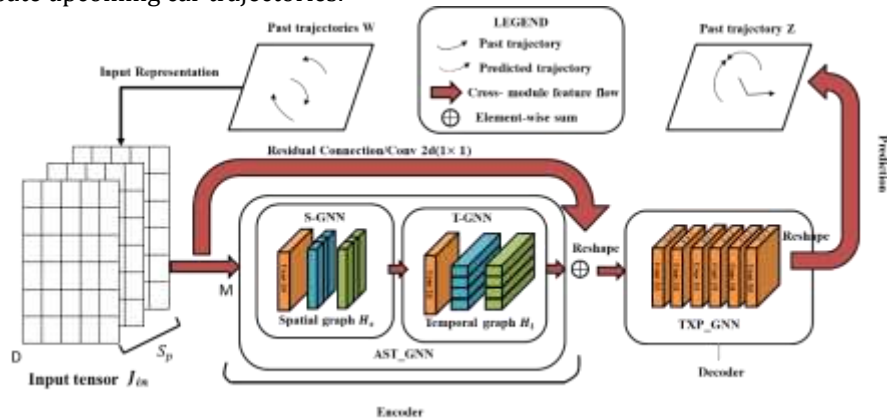


Figure 4: Architecture diagram of AST-GNN for using LiDAR and Event-Based Camera Data Fusion

3.4.2. Bacterial Colony Optimization (BCO)

The BCO algorithm was modified to employ fused LiDAR and historical camera data to optimize trajectory planning during high-speed autonomous maneuvers. The vehicle can investigate several possible trajectories during the metabolic and transmission phases, while migratory and removal-reproduction make sure that only the smoothest and most energy-efficient routes are kept. In the fused time-spatial domain, the continuous positioning update

equations (14-16) expressed the trajectory improvements. The following is the updating technique:

$$Position_j(S) = Position_j(S-1) + D(j) * [e_j \cdot (H_{best} - Position_j(S-1)) + (1 - e_j) * (O_{best_j} - Position_j(S-1)) + turb_j] \quad (14)$$

$$Position_j(S) = Position_j(S-1) + D(j) * [e_j \cdot (H_{best} - Position_j(S-1)) + (1 - e_j) * (O_{best_j} - Position_j(S-1))] \quad (15)$$

$$D(j) = D_{min} + \left(\frac{Iter_{max} - Iter_i}{Iter_{max}} \right) (D_{max} - D_{min}) \quad (16)$$

$turb_j$ introduces unpredictability through turbulence direction, $D(j)$ represents the chemical step size, and e_j regulates the ratio of local to global searching. O_{best} is the bacterial individual O_{best} , whereas H_{best} is the best solution in the world. $Iter_i$ is the present cycle, while $Iter_{max}$ is the greatest quantity of iterations. $Position_j(S)$ is the j -th bacterium's updated location at cycle S . $Position_j(S-1)$ the organism's prior location. Step size in $D(j)$ chemotaxis regulates motion. Obest is the best remedy for local bacteria. D_{min} is the smallest phase size permitted for fine-tuning. D_{max} the largest step size permitted for extensive exploring.

4. Result and Discussion

The experiments were conducted using Python 3.9. Through combining LiDAR and event-camera features, the existing technique, Regional Feature Fusion- Faster-region based convolution neural network (RFF-Faster R-CNN) [17], is used for high-speed object recognition.

This Figure 5 visualization depicts the ROC curve assessing the performance of the LiDAR, event-based camera data fusion and BC-AST-GNN model to perform high-speed autonomous maneuvers. The large AUC (0.980) indicates the model was accurate in telling the difference between true and false positives and confirms the applicability of the model in dynamic, real-time navigation applications.

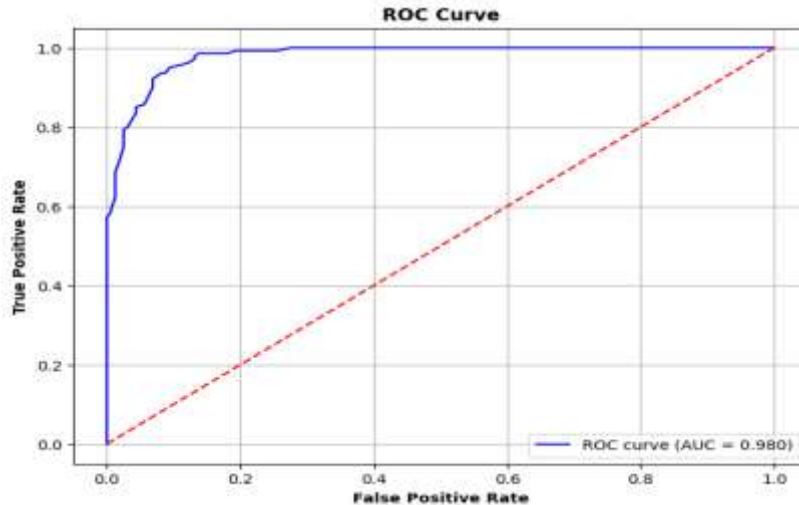


Figure 5: ROC Curve Analysis for BC-AST-GNN in High-Speed Autonomous Maneuvers

Figure 6 demonstrates the training and validation accuracy, as well as the loss of the proposed fusion-based DL architecture to high-speed autonomous maneuvers. It demonstrates with which epochs the model gradually becomes more efficient in learning, with the resultant validation accuracy and loss minimization becoming constant. This shows that LiDAR and event-based camera data fusion are effective in supporting the use of reliable real-time decision-making for autonomous navigation.

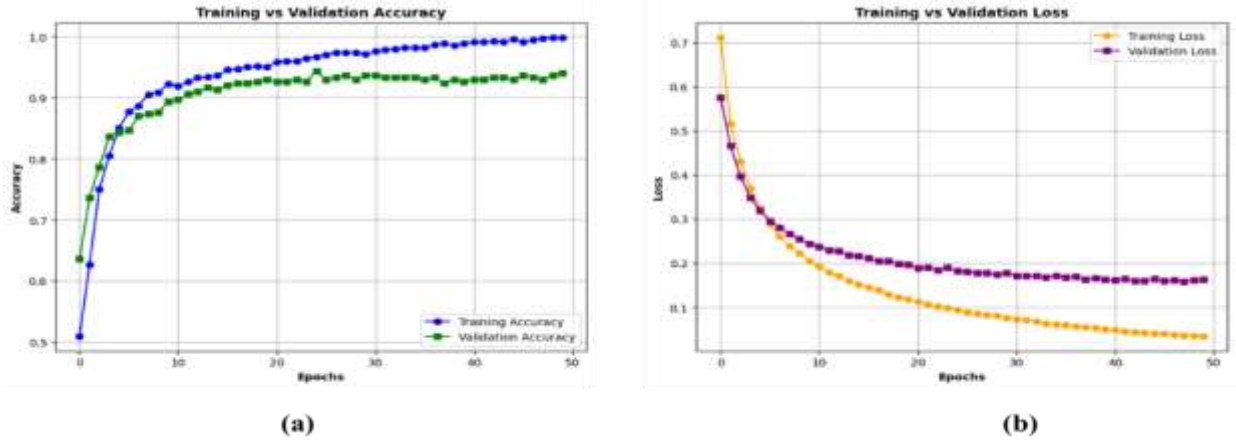


Figure 6: Performance Comparison of (a) Training loss vs (b) Validation Loss of BC-AST-GNN

Using event-based camera inputs merged with LiDAR data, Figures 7 (a) and (b) demonstrate the real-time identification of traffic signals, cars, and people in intricate urban settings. The suggested BC-AST-GNN ensures correct perception for the performance of high-speed autonomous maneuvers by enabling dynamic object tracking and robust feature extraction. The technique improves autonomous driving systems' safety and decision-making effectiveness by enhancing traffic signal identification, vehicle localization, and pedestrian detection in a variety of motion and illumination scenarios.

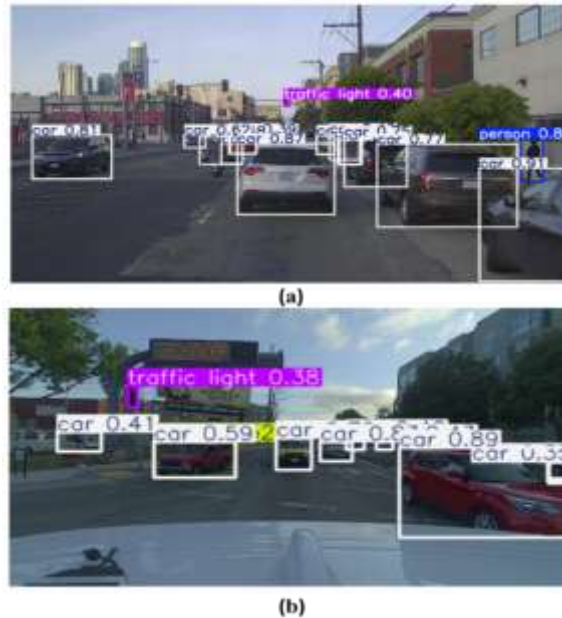


Figure 7:(a) Object detection in urban traffic with an event-based camera; (b) Scene understanding using LiDAR-camera fusion

4.1. Overview of Experimental Detection Performance

Table 1 and Figure 8 present the object detection accuracy (%) achieved by the baseline RFF-Faster RCNN and the proposed BC-AST-GNN across five object categories: Car, Van, Truck, Pedestrian, and Cyclist. The results indicate that BC-AST-GNN consistently outperforms the baseline, achieving accuracy gains of +2.53% for Cars, +9.24% for Vans, +10.45% for Trucks, +7.32% for Pedestrians, and +7.69% for Cyclists. These improvements demonstrate the effectiveness of the proposed LiDAR and event-based camera fusion strategy in enhancing detection robustness for high-speed autonomous maneuvers.

Table 1: Comparison of Baseline and Proposed Method in Multi-Class Object Detection

Method	Car	Van	Truck	Pedestrian	Cyclist
RFF-Faster RCNN [17]	94.59	82.50	79.60	85.31	86.33
BC-AST-GNN [Proposed]	97.12	91.74	90.05	92.63	94.02

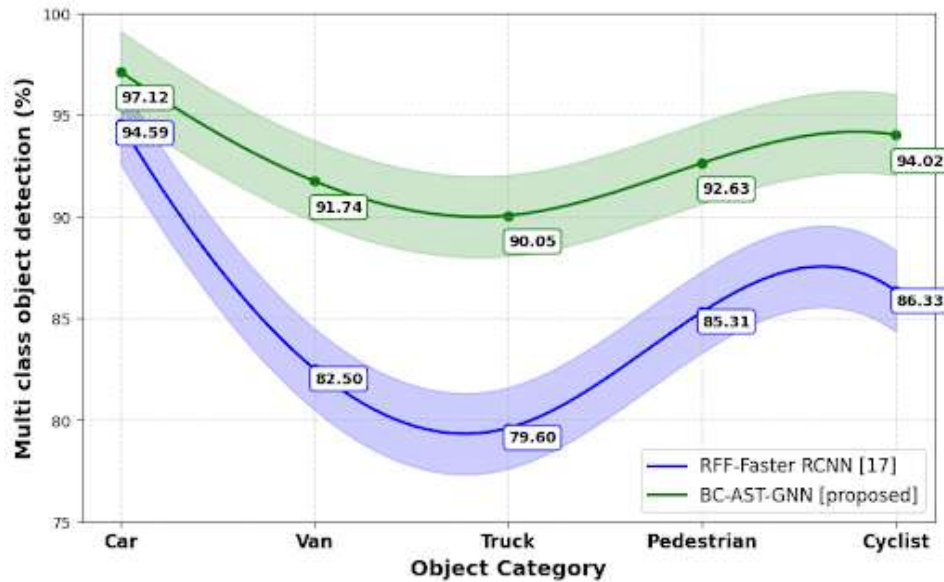


Figure 8: Class-wise Detection Performance of Proposed and Baseline Models

The RFF-Faster RCNN [17] method is a useful integration of camera and LiDAR, but it cannot cope with motion blur, time misalignment, and dynamic object following. The suggested BC-AST-GNN addresses these issues by combining event-based temporal information with LiDAR 3D localization to allow strong and fast maneuvering.

5. Conclusion

The combination of LiDAR and event-based camera images enables safer high-speed automated maneuvers, reliable tracking, and precise diagnosis. The suggested LiDAR and event-based camera fusion framework with BC-AST-GNN shows notable gains in autonomous vehicle perception at high speeds. The model accomplishes strong object detection and trajectory refinement by skillfully fusing temporal dynamics from event cameras with spatial precision from LiDAR. The experimental findings outperform the

baseline in every category, with detection accuracies of 97.12% for cars, 91.74% for vans, 90.05% for trucks, 92.63% for pedestrians, and 94.02% for cyclists. In dynamic, high-speed situations, these improvements guarantee safer maneuver planning, quicker reaction times, and more dependable autonomous driving. Real-world deployment issues and dataset diversity are the suggested framework's limitations. Large-scale testing, multiple-sensor flexibility, and domain adaptability are all part of the future potential for ensuring reliable high-speed vehicular autonomy in a variety of environments.

References

- [1]. Guo, A., Huang, J., Lv, C., Chen, L., and Wang, F.Y., 2024. Advances in autonomous vehicle testing: The state of the art and future outlook on driving datasets, simulators and proving grounds. *Authorea Preprints*. 10.22541/au.172446843.37817639/v1
- [2]. Cornwall, C., 2022. *Event-Based Obstacle Detection with Commercial Lidar* (Master's thesis, Utah State University).
- [3]. Xiao, C., Cao, Q., Zhong, Y., Lan, L., Zhang, X., Luo, Z. and Tao, D., 2024. Motiontrack: Learning motion predictor for multiple object tracking. *Neural Networks*, 179, p.106539. <https://doi.org/10.1016/j.neunet.2024.106539>
- [4]. Luo, G., Shao, C., Cheng, N., Zhou, H., Zhang, H., Yuan, Q. and Li, J., 2023. EdgeCooper: Network-aware cooperative LiDAR perception for enhanced vehicular awareness. *IEEE Journal on Selected Areas in Communications*, 42(1), pp.207-222. 10.1109/JSAC.2023.3322764
- [5]. Tejero-Ruiz, D., Solís-Martín, D., Pérez-Grau, F.J. and Galan-Paez, J., 2024. Unmanned Aerial Vehicle Navigation with Event Cameras. *Available at SSRN 4861406*. <http://dx.doi.org/10.2139/ssrn.4861406>.
- [6]. Su, H., Gao, L., Liu, T. and Kneip, L., 2024. Motion-Aware Optical Camera Communication with Event Cameras. *IEEE Robotics and Automation Letters*. 10.1109/LRA.2024.3517292
- [7]. Xing, J., Wei, D., Zhou, S., Wang, T., Huang, Y. and Chen, H., 2024. A comprehensive study on self-learning methods and implications to autonomous driving. *IEEE Transactions on Neural Networks and Learning Systems*. 10.1109/TNNLS.2024.3440498
- [8]. Siyu, L., 2024. LiDAR-Driven Drivable Space Detection with Deep Neural Network.
- [9]. Shalumov, A., Halaly, R. and Tsur, E.E., 2021. Lidar-driven spiking neural network for collision avoidance in autonomous driving. *Bioinspiration & Biomimetics*, 16(6), p.066016. 10.1088/1748-3190/ac290c
- [10]. El Shair, Z. and Rawashdeh, S., 2023. High-temporal-resolution event-based vehicle detection and tracking. *Optical Engineering*, 62(3), pp.031209-031209. <https://doi.org/10.1117/1.OE.62.3.031209>
- [11]. Feng, J., Zhao, P., Zheng, H., Konpang, J., Sirikham, A. and Kalnaowakul, P., 2025. Enhancing Autonomous Driving Perception: A Practical Approach to Event-Based Object Detection in CARLA and ROS. *Vehicles*, 7(2), p.53. <https://doi.org/10.3390/vehicles7020053>
- [12]. Kambale, A.W., 2021. Event-based camera for object tracking.
- [13]. Sakhai, M., Sithu, K., Oke, M.K.S. and Wielgosz, M., 2025. Cyberattack Resilience of Autonomous Vehicle Sensor Systems: Evaluating RGB vs. Dynamic Vision Sensors in CARLA. *Applied Sciences*, 15(13), p.7493. <https://doi.org/10.3390/app15137493>
- [14]. Queen, K.J., 2023. Event-Based Perception for Ground Vehicle Control (Doctoral dissertation, University of Pennsylvania).
- [15]. Jiao, J., Chen, F., Wei, H., Wu, J. and Liu, M., 2023. LCE-Calib: Automatic LiDAR-frame/event camera extrinsic calibration with a globally optimal solution. *IEEE/ASME Transactions on Mechatronics*, 28(5), pp.2988-2999. 10.1109/TMECH.2023.3259444

- [16]. Iqbal, H., Campo, D., Marin-Plaza, P., Marcenaro, L., Gómez, D.M. and Regazzoni, C., 2021. Modeling perception in autonomous vehicles via 3D convolutional representations on LiDAR. *IEEE Transactions on Intelligent Transportation Systems*, 23(9), pp.14608-14619. 10.1109/TITS.2021.3130974
- [17]. Wu, Q., Li, X., Wang, K. and Bilal, H., 2023. Regional feature fusion for on-road detection of objects using camera and 3D-LiDAR in high-speed autonomous vehicles. *Soft Computing*, 27(23), pp.18195-18213. <https://doi.org/10.1007/s00500-023-09278-3>
- [18]. Astha Modak, Samruddhi Paradkar, Shruti Manwatkar, Amol R. Madane, and Ashwini M. Deshpande. "Human head pose and eye state-based driver distraction monitoring system." In *Proceedings of 3rd International Conference on Computer Vision and Image Processing: CVIP 2018*, Volume 1, pp. 403-421. Singapore: Springer Singapore, 2019.
- [19]. Dinesh Kadam, Amol R. Madane, Krishnan Kutty, S. V. Bonde, "Rain Streaks Elimination using Image Processing Algorithms", *Signal & Image Processing: An International Journal (SIPIJ)*, Vol. 10, No. 3, pp. 21-32, Jun-2019.
- [20]. Apeksha Kamble, Akanksha Singh, Damini Thusoo, Amol R. Madane, "Design and Implementation of Surround View Monitoring System in View of Autonomous Vehicle", 6th International Conference on Data Management, Analytics and Innovation 2022 (ICDMAI 2022), India, pp. 509-519, 14-16 Jan 2022.
- [21]. Ashish Kathe, Pranav Teli, Amol R. Madane, "COVID-19 Regulations Check: Social Distancing, People Counting and Mask Wear Check", 6th International Conference on Data Management, Analytics and Innovation 2022 (ICDMAI 2022), India, pp. 467-484, 14-16 Jan 2022.
- [22]. Priyojit Chakraborty, Amol R. Madane, "Vehicle Queue Management System at Fuel Station", NIELIT's International Conference on Communication, Electronics and Digital Technologies (NICE-DT'23), New Delhi, India, 10-11 Feb 2023.
- [23]. Amol Raoji Madane, Dr. Ho Chin Kuan, "Convxf-Net: An Explainable AI Approach for Recognizing Pedestrian Behaviour Using Convolutional Xgboost Pyramid Fractional Network", *International Journal of Advances in Signal and image Sciences*, Nov 2025.
- [24]. Amol Raoji Madane, Dr. Ho Chin Kuan, "Adaptive Multi-Spectral Vision Fusion for Reliable Hazard Identification in Challenging Visibility Conditions", *International Journal of Applied Mathematics*, Vol 38, No. 12s, pp. 4241-4257, Dec 2025. (ISSN: 1311-1728 (printed version), ISSN: 1314-8060 (on-line version)).
- [25]. Amol R. Madane, Dr. Uttam D. Kolekar, "High Frequency Component based Addition Algorithm of Digital Image Watermarking", *International Journal of Computer Applications*, Vol. 50, No. 10, pp. 30-34, Jul 2012 (ISBN: 973-93-80869-59-3) (Impact Factor- 0.821).