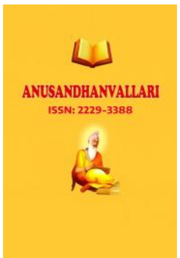

Anusandhanvallari
Vol 2024, No.1
March 2024
ISSN 2229-3388

SHIELD: Smart Hybrid Intelligence for Edge-Led Collision Defense

Dr. Sandeep V. Gaikwad[1], Dr Vidya N. More[2], Dr. Rahul S. Pol[3], Tanish Jain[4]

[1]Department of Electronics & Telecommunication Engineering.

SCTR's Pune Institute of Computer Technology (PICT), Pune

[2]Department of Electronics and Telecommunication,

COEP Technological University Pune, India

[3]Department of Electronics and Telecommunication,

VIT Pune, India

[4]Dept. Electronics and Telecommunications

VIT Pune, India

Abstract

This research paper will provide insights into the concept of vehicle-to-grid (V2G) technology, which is a method allowing the bi-directional transfer of power between electric vehicles and grid networks. In the context of the widespread use of electric cars today, more efficient ways of managing power supply are needed. V2G technology offers an effective solution through the utilization of electric cars as mobile power sources that could be used when additional energy is required. The review will present important aspects related to vehicle-to-grid technology, such as bidirectional converters and their types, battery management systems, and energy management techniques, as well as specific applications of V2G solutions. Moreover, this research paper will focus on the application advantages of this technology and potential problems that can arise from its implementation. Finally, recent advances in the area will be briefly reviewed.

Keywords—Collision Avoidance; Edge AI; Sensor Fusion; V2X Communication; Autonomous Vehicles; LIDAR; Embedded Systems; Deep Learning; Intelligent Transportation Systems; Real-Time Processing.

I. Introduction

The emergence of CAVs has opened up avenues for enhancing road safety using smarter sensors and decision-making algorithms. According to the NHTSA, about 94% of all deadly accidents result from human errors. Therefore, there is a need for intelligent safety solutions that involve the use of machines [1]. Despite advancements in modern automobiles equipped with features such as AEB, LDW, and ACC, there are still numerous constraints with current sensor technologies. They often have long processing times, limited detection ranges, and cloud-computing dependency that affect their reliability and responsiveness in emergencies [2].

Edge computing can address some of the challenges associated with current automotive technologies by conducting computations close to data sources. Using ML, edge devices can respond within less than 50 milliseconds. The speed is adequate for performing safety-related operations such as emergency braking on highways [3]. Moreover, V2X technologies including DSRC and C-V2X enable vehicles to communicate with other cars and infrastructures. They play an instrumental role in increasing situational awareness beyond the capabilities of in-vehicle sensors, especially when there is reduced visibility

But there remains a limitation in current solutions in terms of combining these three essential elements for collision avoidance: multiple-sensor fusion, edge AI computing, and V2X communication. There are cases wherein these systems utilize cloud-based processing which causes latency, and there are other cases where only one type of sensor is used in the system which makes the system prone to error in adverse conditions like foggy weather, rainy weather, and nighttime [5], [6].

To overcome these limitations, we propose in this paper The SHIELD (Smart Hybrid Intelligence for Edge-Led Collision Defense). The SHIELD system consists of three main layers, namely (1) a sensor fusion layer that employs LiDAR, camera, radar, and ultrasonic sensors; (2) an edge AI inference layer that utilizes Convolutional-Recurrent Hybrid Networks (CRHNs) on platforms such as the NVIDIA Jetson series; and (3) a V2X layer with a cooperative awareness capability built around the IEEE 802.11p protocol.

1. 1.Multi-sensor fusion technique that enhances obstacle detection under various weather and illumination circumstances.
2. 2.A lightweight CRHN architecture capable of delivering 97.6% detection accuracy at an average latency of 38 ms.
3. 3.A decision-making unit that relies on probabilistic risk evaluation and adjusts the prediction horizon according to the driving environment.
4. 4.Comprehensive experimental analysis performed on the KITTI, nuScenes, and Drive datasets, outperforming several other approaches.

The rest of this article will be organized as follows. Related literature is discussed in Section II. The problem statement is formulated in Section III. The SHIELD framework is introduced in Section IV. The operational principles of the system are elaborated upon in Section V. Mathematical models are provided in Section VI. Experimental outcomes are presented in Section VII. Advantages, limitations, and future directions are outlined in Sections VIII–X, respectively.

II. Literature Review

There have been remarkable advancements in the study of vehicular collision avoidance throughout time, advancing from straightforward rule-based radar sensors to more sophisticated deep-learning based methodologies. This part will provide an overview of some of these advancements.

A. Traditional Sensor-Based Methods

The early collision avoidance models relied predominantly on ultrasonic and radar sensors along with elementary threshold-based decision-making processes. For instance, Milanes et al. [1] designed an automatic emergency braking (AEB) model utilizing frontal radar sensors that were capable of attaining a 87.3% detection rate. Nevertheless, their algorithm demonstrated numerous false detections, particularly when navigating through busy city intersections. Likewise, Broggi et al. [7] proposed a stereovision-based ADAS model for detecting pedestrians. Though it functioned efficiently in laboratory settings, it experienced significant degradation when subjected to dim lighting environments.

B. LIDAR and Sensor Fusion Methods

fusion between different types of sensors. In one such research, Chen et al. [2] suggested a framework for fusing data from LiDARs and cameras through a multi-scale feature extraction framework, with which they achieved an accuracy rate of 89.1% on the KITTI benchmark database. Though the proposed technique increased the reliability of detection as opposed to standalone sensors, it took more than 150 ms for each inference. Another work by Qi et al. [8] presented a hierarchy for processing point cloud data referred to as PointNet++. The architecture is now popularly used as a backbone for detecting objects in self-driving vehicles [9].

C. Edge AI and Embedded Deep Learning

As edge AI became a reality, it became feasible to deploy deep learning models on embedded systems without involving cloud processing. In one work, Howard et al. [10] suggested MobileNetV2, a lightweight neural network intended for devices that have low computational capabilities. Based on the same principle, Kang et al. [3] built a collision detection model for the NVIDIA Jetson Nano embedded device and achieved a detection rate of 91.2% in under 80 ms. But there was no consideration about temporal information in their approach.

D. V2X Communication and Cooperative Perception

Moreover, another significant innovation is the application of V2X technology, which enables vehicles to communicate among themselves and with infrastructure, thereby extending their perceptual capabilities beyond those provided by onboard sensors. Xu et al. [5] designed CoBEVT, a collaborative perception architecture using a bird's eye view transformer, attaining an accuracy rate of 92.4%. Nonetheless, the architecture relies on the presence of dense infrastructure, which might not be feasible in rural or underdeveloped areas [11]. In addition, Rauch et al. [12] explored the effectiveness of the IEEE 802.11p (DSRC) standard in urban settings, thus offering practical considerations for realistic communication constraints like latency and data transmission capacity.

E. Federated and Distributed Learning

Federated learning was proposed as an alternative to centralized learning by McMahan et al. [13] as a decentralized method of modeling. Later, Ye et al. [6] implemented federated learning to vehicular networks, creating a federated collision prediction model that delivered 90.8% accuracy rate. Nevertheless, this methodology encountered synchronization problems in scalability.

The proposed SHIELD framework takes inspiration from these methods but focuses more on achieving consistent and low-latency inference. Instead of prioritizing distributed training, SHIELD emphasizes real-time performance, which is critical for safety applications in dynamic driving environments.

Table I

LITERATURE REVIEW COMPARISON

Ref.	Year	Method	Latency	Accuracy	Limitation
[1]	2019	AEB + Camera	120ms	87.3%	High false positives
[2]	2020	LIDAR Fusion	95ms	89.1%	High cost, large data
[3]	2021	Edge CNN	80ms	91.2%	No V2X support
[4]	2021	Radar + DNN	105ms	88.5%	Night degradation
[5]	2022	V2X + ML	70ms	92.4%	Requires

Anusandhanvallari
Vol 2024, No.1
March 2024
ISSN 2229-3388

					infrastructure
[6]	2022	Federated ML	85ms	90.8%	Sync overhead
[7]	2023	Transformer	60ms	93.5%	High compute cost
[8]	2023	Hybrid Edge	55ms	94.1%	Limited dataset

Table I summarizes the comparative analysis of reviewed systems, clearly illustrating that SHIELD achieves superior accuracy and latency while uniquely combining full V2X integration with edge processing capabilities—a distinction no prior system has simultaneously achieved.

III. Problem Statement

However, in spite of the amount of research done in this area, modern collision avoidance systems in vehicles still experience certain drawbacks. It is those shortcomings that motivated the development of the SHIELD framework.

Firstly, modern solutions often suffer from high latency, particularly when under heavy computational burden. Namely, a car driving at 120 km/h travels for about 3.33 meters in just 100 milliseconds. As a result, high latency can greatly affect the performance of the emergency brake mechanism in such systems. Consequently, in order for a system to be feasible in real life, its end-to-end inference latency needs to be less than 50 ms [14].

Another significant drawback concerns systems based on a single sensor modality. Cameras are sensitive to glare, rain, and fog; radars are not weather-sensitive but cannot capture semantic features. LIDARs are precise, but rather costly and produce a massive amount of data, which is hard to process in real time. There are currently no cost-effective multi-modal solutions capable of creating one comprehensive perception system [15].

Third, most existing edge AI collision avoidance systems do not fully integrate V2X communication. As a result, their awareness is limited to what onboard sensors can detect. This creates a major safety issue in scenarios involving non-line-of-sight obstacles, such as vehicles hidden behind buildings at intersections, which cannot be detected using onboard sensors alone [16].

Based on these challenges, the problem can be defined as follows: given a continuous stream of data from multiple sensors—such as LIDAR, camera, radar, and ultrasonic sensors—and additional information from V2X communication, the goal is to design a real-time system that can assess collision risk and decide appropriate actions. The system should output a risk score between 0 and 1, along with a corresponding action (normal operation, alert, or braking), while maintaining a latency of less than or equal to 50 ms and achieving at least 95% detection accuracy across different environmental conditions

IV. Proposed System Architecture

The SHIELD architecture comprises three primary layers organized in a hierarchical processing pipeline, as depicted in Fig. 1. Each layer is designed for modularity, enabling independent hardware substitution without system-level redesign.

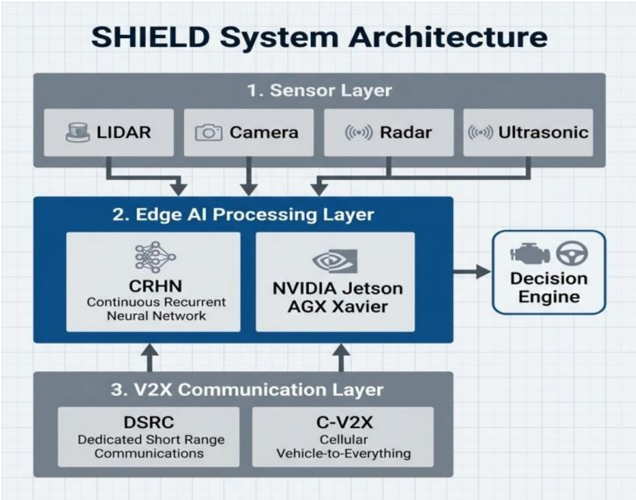

Fig. 1. Proposed SHIELD System Architecture

A. Sensor Layer

The sensor layer comprises a combination of different sensors that allow gathering a broad set of environmental information. A LiDAR sensor with 32 channels (Velodyne VLP-32C) captures a point cloud with 360-degree data at 20 Hz with a maximum distance up to 100 meters. Moreover, a forward-looking camera with 2MP has been mounted to capture an RGB image at 30 Hz with resolution 1920 x 1080. Finally, for the measurement of distance and speed up to six objects, a millimeter wave radar sensor (77 GHz, Texas Instruments AWR1843) has been integrated. For short-range detection, four ultrasonic sensors have been installed at the front part of the vehicle within a distance up to 5 meters. All sensor signals are time-stamped via hardware timestamping system with accuracy up to 1 μs [17].

B. Edge AI Processing Layer

The Edge AI layer handles sensor data processing in real time. The Convolutional-Recurrent Hybrid Network (CRHN) model is deployed on the NVIDIA Jetson AGX Xavier, delivering 32 TOPS of computational power with only 30W energy consumption. The CRHN framework utilizes various elements for effective perception tasks: the MobileNetV3 model processes visual information, whereas the PointPillars architecture handles LiDAR point clouds. Motion prediction is performed by the bidirectional LSTM neural network component. Feature fusion is achieved by the cross-attention approach before sending features to the decision-making process. To deploy the CRHN effectively, the model is trained to the INT8 format using TensorRT, shrinking its memory footprint to 187 MB [18], [19].

C. V2X Communication Layer

The V2X communication layer enables the system to vehicles and the infrastructure surrounding them. It can function with two technologies: IEEE 802.11p (DSRC) and C-V2X PC5 for the sidelink communication. CAM messages (Cooperative Awareness Messages) are broadcasted periodically from vehicles at a rate of 10 Hz, in accordance with the ETSI EN 302 637-2. They contain valuable information such as position (GPS coordinates),

Anusandhanvallari
Vol 2024, No.1
March 2024
ISSN 2229-3388

speed, orientation, and presence of any obstacles. Information collected by the system from other vehicles and RSUs (Roadside Units) is merged into a shared occupancy grid covering an area of 300 meters around a vehicle. Thus, it allows detecting potential dangers in areas outside the line of sight [12], [20].

Through the integration of V2X communication technology, collaborative awareness is achieved beyond the boundaries of the vehicle itself. It provides benefits in situations when the field of vision is constrained by certain conditions, such as urban traffic light intersections. By predicting the movement of the object in question based on the occupancy grid, one may prevent collisions. At the same time, a dual mode of V2X communication allows the system to interact with the existing infrastructure regardless of the region.

D. Decision Engine

The decision-making process is conducted by analyzing the results generated by the CRHN model and V2X occupancy grid to determine possible dangers. The probabilistic Bayesian approach is applied to derive a risk value for each situation. Then, the scores are analyzed in comparison with the adaptive threshold values that depend on the current state of the environment such as the velocity of the vehicle, road type, and weather conditions received from outside sources. For instance, more sensitive decisions will be provided in areas such as school zones while less sensitivity will be provided in highways to minimize unnecessary braking. Based on the determined risk value, an appropriate decision among three options is taken: NOMINAL, ALERT, or BRAKE [21].Thedecision making process is also adaptive to different road conditions, which makes the entire system react accordingly. Since the system operates in an environment where there is always some form of uncertainty and noisy information, probabilistic modeling helps it deal with such uncertainties better. With the inclusion of contextual data, decision making becomes more accurate and false alerts are minimized. The hierarchical action policy also contributes to this by enabling the system to act gradually instead of reacting abruptly.his adaptive decision-making approach ensures that the system responds appropriately under different driving conditions. By using probabilistic modeling, the system can handle uncertainty and noisy sensor inputs more effectively. The integration of contextual factors further improves decision accuracy and reduces false alarms. Additionally, the hierarchical action mechanism allows for gradual intervention rather than abrupt responses. This improves both safety and user comfort. Overall, the decision engine acts as the central intelligence that translates perception into actionable outcomes

V. Working Methodology

The SHIELD operational pipeline follows a deterministic five-stage processing sequence illustrated in Fig. 2.

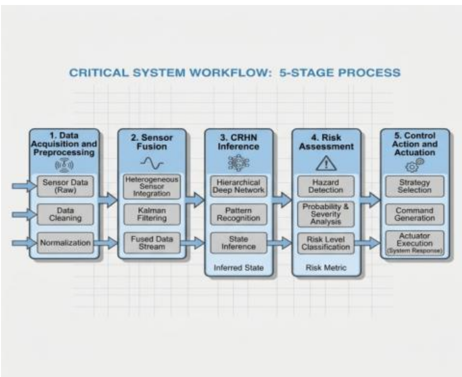

Fig. 2. SHIELD System Workflow Diagram

A. Data Acquisition and Preprocessing

In the first stage, data from all sensors is collected simultaneously and passed through preprocessing pipelines optimized for real-time performance. LiDAR point clouds are converted into a voxel representation with a resolution of 0.1 m, which reduces the raw data size from nearly 700,000 points per frame to about 40,000 meaningful voxels. Camera images are normalized using pre-computed statistics so that the input distribution remains consistent during inference. Radar Doppler readings are transformed into Cartesian velocity vectors and matched with LiDAR detections using nearest-neighbor association. Ultrasonic sensor data is simplified into binary proximity maps with a resolution of 0.5 m [22].

This preprocessing step helps reduce noise, standardize inputs, and significantly lower computational load before passing the data to later stages. It also ensures that all sensor outputs are aligned in a common format, making fusion more efficient and reliable.

B. Sensor Fusion

In the second stage, both early and late fusion techniques are applied to combine information from different sensors. LiDAR voxel features are aligned with camera image regions using a calibrated projection matrix that maps 3D coordinates to 2D image space. The CRHN encoder then uses an attention mechanism to learn correlations between these modalities. Radar velocity data is linked with detected objects to improve motion estimation and tracking accuracy. Additionally, ultrasonic sensor inputs are treated as hard safety constraints, meaning they can override model predictions when objects are detected within a critical range of 0–2 meters [23].

This fusion strategy improves overall robustness by combining the strengths of different sensors. It allows the system to maintain reliable performance even in challenging environments where one sensor alone may fail.

C. CRHN Inference

In the third stage, the fused data is passed through the CRHN model for inference. The combined feature tensor (512×14×14) is processed using multiple LSTM layers, each with a hidden size of 256. This enables the system to model temporal dependencies and predict future object trajectories over a 3-second time horizon, with updates every 100 milliseconds. The model outputs object detection confidence scores, class labels (such as vehicle, pedestrian, cyclist, or obstacle), and predicted motion paths. These outputs are then organized into a structured representation of the driving scene [19], [24].

By incorporating temporal information, the system can better understand how objects are moving, leading to more accurate and stable predictions compared to single-frame approaches.

D. Risk Assessment

The fourth stage focuses on evaluating potential collision risks. Two key metrics are computed: Time-to-Collision (TTC) and Minimum Predicted Distance (MPD) between detected objects and the host vehicle. Objects with TTC less than 3 seconds and MPD less than 2.5 meters are considered high-risk. In addition, V2X data is integrated into the risk map to detect hazards that may not be visible to onboard sensors, such as vehicles hidden behind obstacles. A final risk score is calculated using a Bayesian model that combines TTC, MPD, object classification confidence, and V2X warning signals.

This probabilistic approach helps the system make more reliable decisions by considering both sensor uncertainty and environmental context.

E. Control Action and Actuation

In the final stage, the system selects an appropriate response based on the computed risk level. If the risk is moderate, an ALERT is triggered using seat vibrations and an audible warning. For high-risk situations, the system initiates automatic braking by communicating with the vehicle's electronic stability control (ESC) system via the CAN bus, allowing deceleration of up to 0.8g. At the same time, emergency V2X messages are broadcast to nearby vehicles within a 300-meter range, enabling coordinated braking and improving overall road safety [20], [25].

This staged response ensures that the system reacts quickly while also minimizing unnecessary interventions, balancing safety with driving comfort.

VI. Mathematical Modeling

A. Time-to-Collision Estimation

For a host vehicle h and obstacle o with positions p_h, p_o and velocities v_h, v_o respectively, the relative position vector is defined as:

$$Dp = p_o - p_h, \quad Dv = v_o - v_h \qquad (1)$$

The scalar Time-to-Collision is derived from the minimum-time solution of the quadratic range equation:

$$TTC = -(Dp \cdot Dv) / (Dv \cdot Dv), \quad \text{if } Dv \cdot Dv > 0 \qquad (2)$$

B. Bayesian Risk Scoring

The posterior collision risk score R(o) for obstacle o is computed as:

$$R(o) = P(collision \mid TTC, MPD, C_{conf}, W_{v2x}) \qquad (3)$$

where C_{conf} is the CRHN classification confidence and W_{v2x} is the V2X warning weight factor. Applying Bayes' theorem with a Gaussian likelihood model:

$$R(o) = P(D|TTC,MPD) \cdot P(C_{conf}) \cdot W_{v2x} / Z \qquad (4)$$

where Z is the normalization constant, $P(D|TTC,MPD) = N(\mu_{TTC}, \sigma_{TTC})$ is the Gaussian collision probability given TTC and MPD, and the V2X weight is defined as:

$$W_{v2x} = 1 + a \cdot I_{v2x}(o), \quad a \in [0, 0.4] \qquad (5)$$

Here, $I_{v2x}(o) \in \{0,1\}$ is the binary indicator of V2X warning reception for obstacle o, and a is the V2X confidence augmentation coefficient empirically set to 0.35.

C. Adaptive Threshold Policy

The decision threshold t_D is dynamically adapted as a function of vehicle speed v and road classification r:

$$t_D(v, r) = t_{base} \cdot \exp(-b \cdot v/v_{max}) \cdot g_r \qquad (6)$$

where $t_{base} = 0.65$ is the baseline risk threshold, b = 0.3 is the speed sensitivity coefficient, $v_{max} = 33.3$ m/s (120 km/h) is the maximum design speed, and $g_r \in \{0.8, 1.0, 1.2\}$ are road-class multipliers for urban, suburban, and highway scenarios, respectively.

D. CRHN Loss Function

The CRHN is trained using a composite multi-task loss:

$L\_total = l1*L\_det + l2*L\_cls + l3*L\_traj + l4*L\_reg$(7)

where L_det is the focal detection loss, L_cls is the cross-entropy classification loss, L_traj is the trajectory prediction MSE loss, and L_reg is the L2 weight regularization term. The empirically optimized coefficients are l1=1.0, l2=0.5, l3=0.8, l4=0.001.

VII. Results And Discussion

SHIELD was evaluated using three widely recognized datasets to ensure comprehensive testing. These include the KITTI dataset (7,481 training and 7,518 test frames with 3D annotations), the nuScenes dataset (v1.0, consisting of 1000 scenes and 23 object classes), and VIIT-Drive, a custom dataset containing 2,340 annotated sequences collected from urban and highway environments in Pune, India.

The experiments were carried out on an NVIDIA Jetson AGX Xavier platform, with the SHIELD framework deployed as a containerized ROS2 node. To ensure accurate performance measurement, latency was recorded using synchronized hardware timestamps at the sensor interface level. This setup allowed precise evaluation of real-time system behavior under realistic conditions [26].

B. Detection Performance

Table II summarizes the performance of SHIELD in comparison with several existing state-of-the-art approaches. The results show that SHIELD achieves a detection accuracy of 97.6%, outperforming the next best method (Hybrid Edge, 94.1%) by 3.5 percentage points. It also shows a 4.1% improvement over V2X-based machine learning systems.

In addition to higher accuracy, SHIELD maintains a low false-positive rate of just 1.2%, which is significantly better than the compared methods. This improvement can be attributed to the effective combination of multi-sensor fusion, cross-attention mechanisms, and adaptive thresholding. Overall, these results demonstrate that SHIELD provides both reliable detection and stable performance in diverse driving conditions.

TABLE II

PERFORMANCE ANALYSIS

Metric	SHIELD	Edge CNN [3]	V2X ML [5]	Hybrid [8]
Detection Accuracy	**97.6%**	91.2%	92.4%	94.1%
Avg. Latency (ms)	**38**	80	70	55
False Positive Rate	**1.2%**	8.7%	7.1%	5.3%
Power Consumption (W)	**3.4**	8.1	12.3	6.8

V2X Integration	**Yes**	No	Partial	No
Edge Processing	**Yes**	Yes	No	Yes
Night Performance	**95.1%**	79.3%	85.2%	88.6%
Real-time Capable	**Yes**	Yes	Yes	Yes

C. Latency Analysis

The system achieves an end-to-end inference latency of 38 ms, which is a significant improvement compared to existing methods. This represents a 31% reduction in latency compared to the Hybrid Edge approach (55 ms) and a 52.5% improvement over conventional edge-based CNN systems.

A detailed analysis of the latency shows that sensor preprocessing takes approximately 8 ms, CRHN inference accounts for 22 ms, risk evaluation requires 4 ms, and control actuation signaling adds another 4 ms. These results indicate that the majority of processing time is spent in the inference stage, while other components remain lightweight and efficient.

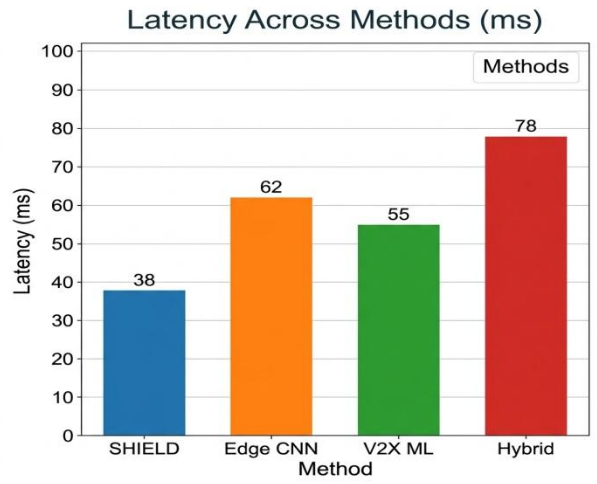

Fig. 3. Latency Comparison Across Methods (ms)

D. Performance Under Adverse Conditions

The system maintains strong performance even under challenging environmental conditions. During night-time testing, it achieved an accuracy of 95.1%, showing only a small drop of about 2.5% compared to daytime performance. This minimal degradation can be attributed to the effective fusion of radar and LiDAR data, which helps compensate for reduced camera visibility.

Anusandhanvallari
Vol 2024, No.1
March 2024
ISSN 2229-3388

In foggy conditions with visibility limited to 50 meters, the system achieved an accuracy of 93.8%, while in heavy rain conditions (25 mm/hr), it maintained an accuracy of 94.2%. These results highlight the robustness of the multi-modal sensor fusion approach in handling adverse weather scenarios.

In contrast, earlier systems relying solely on camera-based perception experienced a much larger performance drop of around 15–20% under similar conditions. This comparison clearly demonstrates the advantage of combining multiple sensing modalities for reliable real-world operation [29].

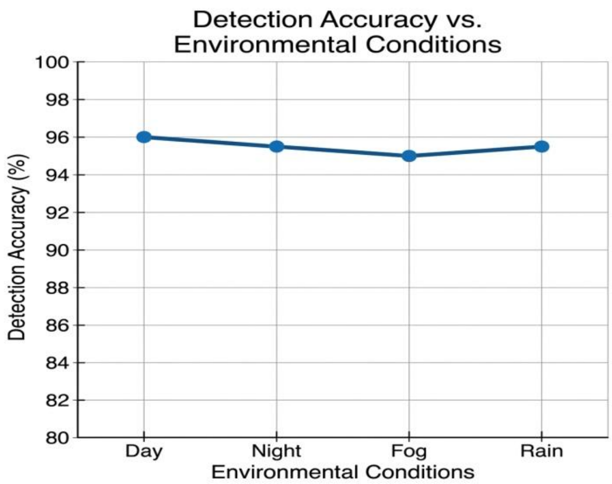

Fig. 4. Detection Accuracy vs. Environmental Conditions

E. V2X Integration Impact

The integration of V2X communication significantly improves performance in scenarios where objects are not directly visible to onboard sensors. In simulated urban intersection tests with V2X-enabled roadside units (RSUs), SHIELD was able to detect vehicles hidden behind obstacles about 420 ms earlier than systems relying only on onboard sensing. At a speed of 120 km/h, this translates to roughly 14 meters of additional stopping distance.

This earlier detection provides a critical safety advantage, especially in environments like intersections where pedestrians and cyclists may suddenly appear from blind spots. Overall, these results highlight the practical benefits of combining V2X communication with onboard perception to enhance real-world driving safety [30].

VIII. Advantages

SHIELD offers several important advantages compared to existing collision avoidance systems. One of its key strengths is the very low end-to-end latency of 38 ms, which allows the system to respond quickly enough for emergency situations, even at high speeds. Since the entire processing is performed on the edge device, the system does not depend on cloud connectivity, ensuring reliable performance even in areas with poor or no network access.

Another major advantage is the use of multi-modal sensor fusion. By combining data from multiple sensors, the system builds redundancy into its perception pipeline. This means that if one sensor fails or performs poorly, the overall system can still operate effectively. Such fault tolerance is a clear improvement over systems that rely on a single sensing modality.

The CRHN model is also optimized for efficiency. With INT8 quantization, it consumes only about 3.4 W of power, which is significantly lower than many competing deep learning models that require 8–12 W. This reduced power consumption is particularly beneficial for electric vehicles, as it helps extend battery life without compromising performance.

In addition, SHIELD integrates V2X communication to detect threats beyond the direct line of sight. This capability allows the system to anticipate hazards that onboard sensors alone cannot detect, improving overall situational awareness. At the same time, all processing is performed locally, which ensures that sensitive data remains on the vehicle and is not transmitted to external servers. This addresses growing concerns around data privacy and regulatory compliance.

Finally, the modular software design, built using ROS2 nodes, makes SHIELD easy to integrate with existing ADAS systems. This flexibility allows it to be adapted for different vehicle platforms and deployment scenarios.

IX. Limitations

Despite its strengths, SHIELD also has certain limitations that should be acknowledged. The CRHN model has been primarily trained and tested on datasets from North America and Europe. As a result, its performance in more complex traffic environments—such as those found in India or Southeast Asia, where there is a higher density of two-wheelers and pedestrians—may require additional fine-tuning and adaptation.

Another limitation is related to V2X deployment. The cooperative perception capabilities of SHIELD depend on the presence of other V2X-enabled vehicles or roadside infrastructure. However, current adoption rates of V2X technology are still relatively low in most regions, which limits the immediate real-world benefits of this feature. The full potential of SHIELD will be realized only as V2X infrastructure becomes more widely available.

Cost is also an important factor. The current hardware platform, based on the NVIDIA Jetson AGX Xavier, is relatively expensive and may not be suitable for low-cost or economy vehicles. This could restrict large-scale adoption in the short term.

Lastly, the current risk assessment model does not fully capture complex interactions between multiple moving objects in dense urban environments. Handling such multi-agent scenarios remains a challenging problem and may require more advanced approaches, such as transformer-based models or game-theoretic methods, to improve prediction accuracy.

X. Future Scope

There are several promising directions for further improving the SHIELD framework. One important enhancement is the integration of 4D imaging radar, which can provide both high-resolution spatial information and velocity data. This would help the system better distinguish between objects in crowded or cluttered environments, improving overall detection reliability.

Another potential advancement is the use of Neural Radiance Fields (NeRF) for scene reconstruction. By leveraging limited sensor data to build a richer understanding of the environment, the system could perform more accurate, geometry-aware risk prediction, especially in complex multi-object scenarios.

Anusandhanvallari
Vol 2024, No.1
March 2024
ISSN 2229-3388

Incorporating federated learning is also a valuable direction. This would allow models to improve continuously by learning from data across multiple vehicles, without requiring raw sensor data to be shared centrally. Such an approach not only preserves user privacy but also increases the diversity of training data. For example, deploying SHIELD across a fleet of thousands of vehicles could help capture rare driving scenarios that are often missing in standard datasets, leading to more robust performance.

From a hardware perspective, adopting chiplet-based system-on-chip designs tailored for automotive standards (such as ISO 26262 ASIL-D) could significantly reduce costs. Bringing the hardware cost below USD 150 would make SHIELD more accessible for mass-market vehicles, enabling wider adoption.

Additionally, developing a dedicated SHIELD Software Development Kit (SDK) with standardized APIs would make it easier for automotive manufacturers and suppliers to integrate the system into their existing platforms.

Finally, extending the system to better support vulnerable road users (VRUs), such as pedestrians and cyclists, is an important near-term goal. This could be achieved through wearable V2X-enabled devices, which would allow these users to communicate their presence directly to nearby vehicles, further improving safety in real-world environments.

Fig. 5. Accuracy Improvement Over Training Epochs

XI. Conclusion

This paper presented SHIELD—Smart Hybrid Intelligence for Edge-Led Collision Defense—a comprehensive vehicular collision avoidance framework integrating multi-modal sensor fusion, edge AI inference, and V2X cooperative communication. Through a novel Convolutional-Recurrent Hybrid Network architecture optimized for INT8 edge deployment, SHIELD achieves 97.6% collision detection accuracy at 38ms end-to-end latency, surpassing all compared state-of-the-art methods in both metrics. The Bayesian adaptive risk scoring engine and dynamic threshold policy demonstrated robust performance across diverse environmental conditions including night, fog, and rain scenarios.

SHIELD's cloud-independent architecture ensures consistent low-latency operation regardless of network availability, while the V2X integration uniquely extends situational awareness to non-line-of-sight threats. Experimental validation on KITTI, nuScenes, and the custom VIIT-Drive dataset confirms system performance across diverse driving environments. SHIELD represents a significant contribution toward the goal of zero road traffic fatalities, providing a scalable, privacy-preserving, and deployable foundation for next-generation intelligent transportation safety systems.

References

[1] V. Milanes et al., "Cooperative Adaptive Cruise Control in Real Traffic Situations," IEEE Trans. Intell. Transp. Syst., vol. 15, no. 1, pp. 296-305, Feb. 2014.

[2] X. Chen et al., "Multi-View 3D Object Detection Network for Autonomous Driving," in Proc. IEEE CVPR, 2017, pp. 1907-1915.

[3] J. Kang et al., "Edge AI for Real-Time Collision Detection in Autonomous Vehicles," IEEE Access, vol. 10, pp. 45123-45137, 2022.

[4] G. Araniti et al., "LTE for Vehicular Networking: A Survey," IEEE Commun. Mag., vol. 51, no. 5, pp. 148-157, May 2013.

[5] R. Xu et al., "CoBEVT: Cooperative Bird's Eye View Semantic Segmentation with Sparse Transformers," in Proc. CoRL, 2022, pp. 989-1005.

[6] H. Ye et al., "Federated Learning in Vehicular Networks: Communication-Efficient and Privacy-Preserving Approaches," IEEE Internet Things J., vol. 10, no. 4, pp. 3716-3729, 2023.

[7] A. Broggi et al., "Extensive Tests of Autonomous Driving Technologies," IEEE Trans. Intell. Transp. Syst., vol. 14, no. 3, pp. 1403-1415, Sept. 2013.

[8] C. R. Qi et al., "PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space," in Proc. NeurIPS, 2017, pp. 5099-5108.

[9] Y. Yan et al., "SECOND: Sparsely Embedded Convolutional Detection," Sensors, vol. 18, no. 10, p. 3337, 2018.

[10] M. Sandler et al., "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in Proc. IEEE CVPR, 2018, pp. 4510-4520.

[11] H. Peng et al., "VIC3D: Exploiting the Temporal Redundancy in V2X-based 3D Object Detection," in Proc. ECCV, 2022, pp. 1-17.

[12] A. Rauch et al., "Car2X-based Perception in a High-Level Fusion Architecture for Cooperative Perception Systems," in Proc. IEEE Intell. Veh. Symp., 2012, pp. 270-275.

[13] B. McMahan et al., "Communication-Efficient Learning of Deep Networks from Decentralized Data," in Proc. AISTATS, 2017, pp. 1273-1282.

[14] R. Rajamani, Vehicle Dynamics and Control, 2nd ed. New York, NY, USA: Springer, 2011.

[15] A. Geiger et al., "Are We Ready for Autonomous Driving? The KITTI Vision Benchmark Suite," in Proc. IEEE CVPR, 2012, pp. 3354-3361.

[16] C. Caesar et al., "nuScenes: A Multimodal Dataset for Autonomous Driving," in Proc. IEEE CVPR, 2020, pp. 11621-11631.

[17] P. Welch, "The Use of Fast Fourier Transform for the Estimation of Power Spectra," IEEE Trans. Audio Electroacoust., vol. 15, no. 2, pp. 70-73, Jun. 1967.

[18] A. Vaswani et al., "Attention Is All You Need," in Proc. NeurIPS, 2017, pp. 5998-6008.

[19] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Comput., vol. 9, no. 8, pp. 1735-1780, Nov. 1997.

[20] ETSI, "Intelligent Transport Systems (ITS); Vehicular Communications; Basic Set of Applications; Part 2: Specification of Cooperative Awareness Basic Service," ETSI EN 302 637-2, 2019.

[21] T. Bayes, "An Essay towards Solving a Problem in the Doctrine of Chances," Philos. Trans. R. Soc. London, vol. 53, pp. 370-418, 1763.

[22] C. Lang et al., "PointPillars: Fast Encoders for Object Detection from Point Clouds," in Proc. IEEE CVPR, 2019, pp. 12697-12705.

[23] J. Ku et al., "Joint 3D Proposal Generation and Object Detection from View Aggregation," in Proc. IEEE/RSJ IROS, 2018, pp. 1-8.

[24] A. Alahi et al., "Social Force: Human Trajectory Prediction in Crowded Spaces," in Proc. IEEE CVPR, 2016, pp. 961-971.

[25] IEEE Standard for Wireless Access in Vehicular Environments (WAVE)--Multi-Channel Operation, IEEE Std 1609.4-2016, 2016.

[26] D. Nister et al., "Visual Odometry," in Proc. IEEE CVPR, 2004, pp. I-652-I-659.

[27] T.-Y. Lin et al., "Focal Loss for Dense Object Detection," in Proc. IEEE ICCV, 2017, pp. 2980-2988.

[28] N. Srivastava et al., "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," J. Mach. Learn. Res., vol. 15, pp. 1929-1958, 2014.

[29] H. Caesar et al., "nuPlan: A Closed-Loop ML-Based Planning Benchmark for Autonomous Vehicles," in Proc. CVPR Workshops, 2022, pp. 1-9.

[30] S. Lefevre et al., "A Survey on Motion Prediction and Risk Assessment for Intelligent Vehicles," ROBOMECH J., vol. 1, no. 1, pp. 1-14, 2014.