Proceedings of the International Conference on Industrial Engineering and Operations Management

Publisher: IEOM Society International, USA DOI: 10.46254/GC03.20260510
Published: February 2, 2026

Fusion-Based Sensor Data Processing and Visualization Framework for Autonomous Vehicle Testing in GCC Terrains

Ismail Hamieh
Department of Electrical and Computer Engineering
The University of Western Ontario
1151 Richmond St, London, ON N6A 3K7 Canada
ihamieh@uwo.ca

Mazen Kiki
Department of Mechanical Engineering
The University of Akron
244 Sumner Street, Akron, OH 44325, United States
mak259@uakron.edu

Abstract

Autonomous driving in the Gulf Cooperation Council (GCC) region presents unique environmental challenges including high temperatures, reflective surfaces, sand particles, and variable lighting conditions. These factors degrade the performance of single-sensor systems, making multi-sensor fusion a crucial approach for reliable perception and decision-making. This study develops a ROS2-based sensor fusion framework to process and visualize LiDAR and camera data specifically under GCC-like terrain conditions. The system integrates calibration parameters from the KITTI dataset and applies them within a containerized environment to synchronize 3D point clouds with 2D visual frames. By visualizing the fused data through Foxglove Studio, the framework demonstrates enhanced scene understanding, robust depth estimation, and improved object boundary recognition under environmental distortions typical of desert and coastal landscapes. The modular design allows adaptation for real-world GCC deployments, offering a foundational step toward perception algorithm validation, AI-driven sensor alignment, and autonomous vehicle testing in complex regional conditions.

Keywords

Autonomous Vehicles, GCC Terrain, Sensor Fusion, LiDAR, ROS2, AI and Machine Learning Based Perception

1. Introduction

Reliable perception and localization are essential for the safe operation of autonomous vehicles, particularly in environments that present challenging visibility, reflectivity, and illumination conditions. In the Gulf Cooperation Council (GCC) region, unique environmental factors such as dust storms, sharp solar angles, and coastal fog create significant difficulties for both vision and ranging sensors. Cameras are affected by glare, overexposure, and reduced contrast due to intense sunlight and airborne particles, while LiDAR sensors experience scattering, attenuation, and partial loss of point returns under dusty or foggy conditions. These factors degrade depth estimation, object recognition, and sensor alignment, emphasizing the importance of multi-sensor fusion as a method that combines LiDAR's geometric precision with the semantic detail of camera imagery to achieve more reliable perception in harsh terrain.

Previous research has demonstrated that LiDAR–camera fusion can substantially enhance localization and perception accuracy. Hamieh, Myers, and Rahman (2019a; 2019b) developed LiDAR-based mapping and classification frameworks that improved localization precision through better feature extraction and alignment. Hamieh (2020) extended this work by reconstructing road surfaces from LiDAR point clouds, showing how geometric information can compensate for degraded visual cues. Drouin and Hamieh (2020) further examined active time-of-flight imaging systems, demonstrating how environmental factors such as heat and surface reflectivity influence range performance, an effect that becomes more significant under GCC conditions.

Recent research in ROS2 has made sensor fusion both modular and real time. Benayed, Mabrouk, and Masmoudi (2025) presented a LiDAR–camera fusion pipeline using YOLO v9 within ROS2 to improve object detection accuracy. Ibrahimi (2024) showed that integrating object detection with LiDAR data enhances navigation and tracking robustness. Sindayigaya and Rahman (2025) demonstrated LiDAR mapping and camera-driven tracking for people detection, where fusion increased perception reliability. Ryalat et al. (2025) developed a hybrid aerial–ground robot using fused LiDAR, odometry, and camera data for inspection tasks. Rogers et al. (2025a, 2025b) further highlighted the benefits of modular, sensor-agnostic fusion approaches that support multi-platform robotic systems.

While these studies show strong results, most were validated in controlled or urban environments. None specifically addressed the environmental challenges of GCC climates, where dust storms, fog, and strong sunlight reflection can significantly degrade both LiDAR reflectivity and camera contrast. Common datasets such as KITTI, although widely used, do not include data collected under these extreme conditions, leaving a gap in evaluating sensor fusion systems for Middle Eastern deployments. The KITTI Vision Benchmark Suite (Geiger, Lenz, and Urtasun, 2012) remains a foundational dataset for validating fusion frameworks and calibration accuracy and is used in this study as the reference benchmark.

A. Project Problem Statement

Autonomous systems that perform well in moderate climates often experience degraded perception in GCC environments. Airborne dust particles reduce LiDAR return density and scatter emitted light, fog introduces depth estimation errors through optical diffusion, and sharp solar angles lead to camera saturation and shadow contrast distortion. These conditions disrupt the assumptions of consistent reflectivity and exposure made in standard datasets, reducing projection accuracy and object detection reliability. There is currently no established, open, and reproducible framework that quantifies how perception performance varies under these conditions.

B. Solution

This study introduces a ROS2-based LiDAR–camera fusion and visualization framework designed to evaluate perception reliability under GCC-specific environmental challenges. The framework integrates calibration parameters from the KITTI dataset and operates in a Docker environment to ensure reproducibility across different systems. It computes per-frame reprojection metrics to quantify LiDAR coverage within camera imagery, serving as a direct indicator of spatial alignment accuracy.

To enhance the evaluation process, an AI-based anomaly detection module using the Isolation Forest algorithm is implemented to identify irregularities in fusion performance. This module detects frames that exhibit low coverage or sudden projection errors, which can occur due to dust, fog, or sunlight interference. The method combines traditional geometric calibration with data-driven adaptability, providing a practical example of AI integration in perception validation.

The system also produces three-panel visualizations that display camera, LiDAR, and fused views annotated with metrics such as coverage percentage, RMSE, and recall. These visualizations can be streamed in real time through Foxglove Studio, allowing researchers and engineers to analyze fusion quality and identify potential weaknesses under realistic GCC operating conditions.

1.1 Objectives

The main objective of this study is to develop and evaluate a ROS 2–based LiDAR–camera fusion framework that demonstrates the measurable advantages of multimodal perception under environmental conditions representative of the Gulf Cooperation Council (GCC) region. These conditions include dust storms, high solar angles, fog, and reflective terrain surfaces, all of which degrade the performance of single-sensor perception systems.

The framework is designed to achieve the following goals:

1. Quantify sensor performance by computing reprojection metrics such as LiDAR-to-camera point-coverage percentage, mean depth accuracy, and frame-to-frame consistency to compare LiDAR-only, camera-only, and fused-sensor configurations.
2. Demonstrate enhanced perception and visualization through real-time data fusion and rendering using Foxglove Studio, enabling intuitive assessment of fusion quality and environmental impact.
3. Integrate lightweight AI methods, particularly an Isolation Forest–based anomaly-detection module, to identify irregular projection behavior or unexpected coverage drops that may arise from dust interference, fog, or sunlight glare.
4. Provide a reproducible and modular testing pipeline that combines ROS 2 nodes, open-source calibration tools, and containerized execution, allowing future extension to real GCC-collected datasets.

Although the present work employs the KITTI Vision Benchmark Suite (Geiger et al., 2012) rather than data recorded in GCC environments, it serves as a controlled benchmark for evaluating and visualizing the relative performance of camera, LiDAR, and fused modalities. The outcomes establish a foundational reference showing how sensor fusion can mitigate the effects of visibility loss and illumination variation expected in GCC conditions, supporting future field validation with regional datasets.

2. Literature Review

A. Advanced in LIDAR-Camera Fusion and ROS2 Frameworks

Multi-sensor fusion has become a cornerstone of perception systems in autonomous vehicles, offering complementary advantages between LiDAR and camera sensors for improved depth understanding, obstacle detection, and semantic interpretation. The use of ROS2 frameworks has further strengthened this integration by enabling modular synchronization, distributed processing, and real-time data exchange across heterogeneous sensors.

Sindayigaya and Rahman (2025) implemented a ROS-based mapping and tracking framework integrating LiDAR scans with camera-driven people detection, highlighting the reliability of ROS environments for autonomous robotic systems. Similarly, Ryalat et al. (2025) developed a hybrid aerial–ground robot for autonomous inspection using a ROS2–PX4 architecture that effectively synchronized LiDAR and visual odometry data. Rogers, Liberatore, and James (2025) proposed the Fusion-Based Utilization and Synthesis of Efficient Detections (FUSED) framework, demonstrating how ROS2 repositories can fuse 2D LiDAR range data with image frames to enhance detection accuracy. Ibrahimi (2024) presented a real-time object detection and LiDAR fusion system that leverages ROS2 nodes for synchronized data acquisition, improving perception consistency and reducing latency.

The significance of open datasets in testing fusion algorithms has also been emphasized. The KITTI Vision Benchmark Suite by Geiger et al. (2012) provides a foundational dataset for evaluating multi-sensor alignment and projection accuracy. In this research, the KITTI dataset was utilized to validate LiDAR–camera synchronization under simulated GCC-like lighting and terrain conditions, bridging existing benchmark evaluations with regional environmental challenges.

B. Environmental Perception Challenges in Harsh Weather

Perception robustness in adverse weather remains a primary challenge in deploying autonomous systems in real-world environments. Rain, fog, and dust significantly affect visual and depth sensors by scattering light, introducing occlusions, and degrading reflectivity. Qiu et al. (2024) analyzed the impacts of rain and fog on autonomous driving perception, showing that both visibility and object-recognition performance degrade sharply under these conditions.

In desert climates such as those of the GCC, dust storms, sand accumulation, and high solar angles further complicate visual perception. Cameras face overexposure, contrast loss, and false edge detection, while LiDAR returns can be distorted due to reflective interference and backscattering. Sumalatha, Chaturvedi, and Patil (2024) investigated multi-sensor fusion strategies for overcoming such degradation, introducing adaptive fusion mechanisms that dynamically adjust sensor weights based on environmental feedback using convolutional and recurrent neural networks.

Additionally, Zhang et al. (2025) proposed a tightly coupled vector HD map–LiDAR–GNSS–INS framework that maintains accurate positioning even in GNSS-degraded environments. Their integration method improved localization reliability and supported mapping continuity in urban and semi-arid settings. Tao et al. (2024) advanced this work by demonstrating a LiDAR-based HD-map matching technique for tunnels using pavement-marking likelihood, achieving sub-20 cm positioning accuracy despite GPS loss.

Further, Dong et al. (2024) presented SuperFusion, a multilevel LiDAR–camera fusion network that produces long-range HD maps up to 90 meters, demonstrating how deep-fusion strategies can improve both perception range and planning safety. Complementing this, Vernekar et al. (2024) introduced LeGO-LOAM HD, an open-source LiDAR mapping system optimized for real-time high-definition map generation using ROS-based frameworks, proving effective under variable outdoor conditions.

These studies collectively reinforce that fusion-based perception and mapping are essential for addressing GCC-specific challenges such as sand dust, high reflectivity, and fog interference—conditions that degrade standalone sensor performance.

C. AI and Anomaly Detection in Perception Systems

While fusion techniques improve perception accuracy, real-world reliability also depends on identifying sensor or data anomalies caused by noise, calibration drift, or weather-induced interference. As autonomous systems rely on massive data streams, integrating machine-learning-driven anomaly detection has become a key research direction for ensuring safety and robustness.

Baccari et al. (2024) provided a comprehensive review of anomaly detection in connected and autonomous vehicles, categorizing methods into statistical, classical machine-learning, and deep-learning approaches. Their analysis underscored the vulnerability of perception systems to anomalies stemming from hardware degradation, environmental effects, and cyber-attacks, and emphasized the need for continuous monitoring using adaptive AI-based models.

Expanding beyond the automotive domain, Katurde, Musale, Lohar, and Chaudhari (2024) introduced SecureSense, a behavior-based anomaly-detection framework utilizing artificial intelligence and machine learning to identify irregular patterns in real time. They demonstrated that continuous learning and adaptive refinement can significantly reduce false positives while enhancing detection resilience in complex systems.

Building on these approaches, this research applies a lightweight Isolation Forest algorithm to detect anomalies in LiDAR–camera fusion data by analyzing projection ratios, point-cloud coverage, and reprojection accuracy. This method enables the identification of frame-level perception anomalies, helping to isolate sensor inconsistencies, transient errors, or weather-induced distortions during autonomous testing. The integration of AI-based anomaly detection with the ROS 2 perception pipeline supports data-driven reliability assessment, positioning the framework as a robust validation tool for GCC terrain environments.

To summarize the current state of research in sensor fusion and perception under adverse conditions, a structured literature review was conducted. Table 1 presents a comparative overview of notable studies on LiDAR–camera fusion, environmental perception challenges, HD map generation, and AI-based anomaly detection. The table highlights the key contributions, advantages, and limitations of each work, forming the foundation for identifying research gaps addressed in this paper.

Table 1. Summary of Reviewed Literature on Perception and Fusion Techniques for Autonomous Systems

No.	Researcher/s	Contribution	Pros./Result	Cons./Limitation
1	J. Sindayigaya and A. Rahman (2025)	Developed ROS-based LiDAR–camera tracking and mapping for autonomous robots.	Improved perception reliability in controlled environments.	Limited validation under harsh environmental conditions.
2	M. Ryalat et al. (2025)	Designed a ROS2–PX4 hybrid aerial–ground robot with sensor fusion.	Demonstrated accurate multi-sensor coordination for inspection tasks.	Focused on inspection use cases, not driving-scale perception.
3	E. Rogers et al. (2025)	Proposed *FUSED* framework using ROS2 for efficient LiDAR–camera detections.	Enhanced detection accuracy using synchronized fusion nodes.	Lacked environmental adaptation and long-term performance analysis.
4	I. Ibrahimi (2024)	Integrated real-time object detection with LiDAR fusion using ROS2 nodes.	Reduced latency and improved alignment precision.	Performance not tested in outdoor or adverse conditions.
5	Y. Qiu et al. (2024)	Reviewed effects of rain and fog on autonomous vehicle perception.	Provided detailed insight on weather-induced degradation.	Focused on visual sensors, excluding fusion-based mitigation.
6	I. Sumalatha et al. (2024)	Investigated adaptive multi-sensor fusion using AI for self-driving vehicles.	Showed improved robustness in variable weather conditions.	Computationally intensive and not field-tested.
7	H. Zhang et al. (2025)	Proposed tightly coupled HD-map, LiDAR, GNSS, and INS integration.	Achieved sub-meter accuracy in GNSS-challenging environments.	Relied heavily on high-quality HD maps; limited real-time scalability.
8	Q. Tao et al. (2024)	Developed HD-map matching using pavement marking likelihood for tunnels.	Achieved <20 cm accuracy even without GPS.	Applicable mainly to tunnel-like or structured road environments.
9	H. Dong et al. (2024)	Introduced *SuperFusion* for multi-level LiDAR–camera HD map generation.	Extended mapping range to 90 m; improved path planning quality.	Requires dense data and high computational power.
10	R. Vernekar et al. (2024)	Built open-source LiDAR mapping using LeGO-LOAM HD.	Delivered efficient mapping on embedded systems.	Limited validation in extreme outdoor conditions.
11	S. Baccari et al. (2024)	Surveyed AI-based anomaly detection for autonomous vehicles.	Classified statistical and deep-learning detection methods.	Focused mainly on conceptual frameworks, not implementation.
12	A. Katurde et al. (2024)	Presented *SecureSense* AI/ML tool for real-time anomaly detection.	Enhanced system resilience via behavior-based ML models.	Lacked application to perception or sensor fusion systems.

Table 1 synthesizes the current state of research on LiDAR–camera fusion, environmental perception, HD mapping, and AI-driven anomaly detection. The reviewed works collectively highlight strong progress in sensor integration, ROS2-based perception pipelines, and deep fusion methods that enhance mapping accuracy and robustness. However, most studies were validated in controlled or urban settings, with limited consideration of harsh environmental conditions such as dust, fog, and reflective sand surfaces typical of GCC regions. Additionally, several approaches require high computational resources or rely on dense datasets that are not always feasible for real-world deployment. These gaps underscore the need for a modular and lightweight fusion framework that can operate reliably under extreme visibility and illumination variations. The methodology presented in the following section addresses these limitations by introducing a ROS2-based LiDAR–camera fusion pipeline with integrated AI-driven anomaly detection designed to support perception testing in GCC terrains.

3. Methodology

This section describes the complete LiDAR–camera fusion and AI-driven evaluation framework developed using ROS2 inside a Docker container. The system processes KITTI LiDAR and camera data, performs geometric projection, computes performance metrics, and applies machine-learning based anomaly detection to identify low-quality or unreliable frames. The overall architecture is shown in Figure 1, which illustrates the major processing modules and data flow.

3.1 System Overview

The framework is implemented as a modular ROS2 system running inside a Docker container. The design consists of two parallel data branches, one for LiDAR and one for camera frames, followed by synchronization, geometric fusion, visualization, and an evaluation block that computes performance metrics and anomaly detection scores.

The flowchart in Figure 1 summarizes the entire pipeline from raw KITTI data acquisition to AI-based anomaly flagging.

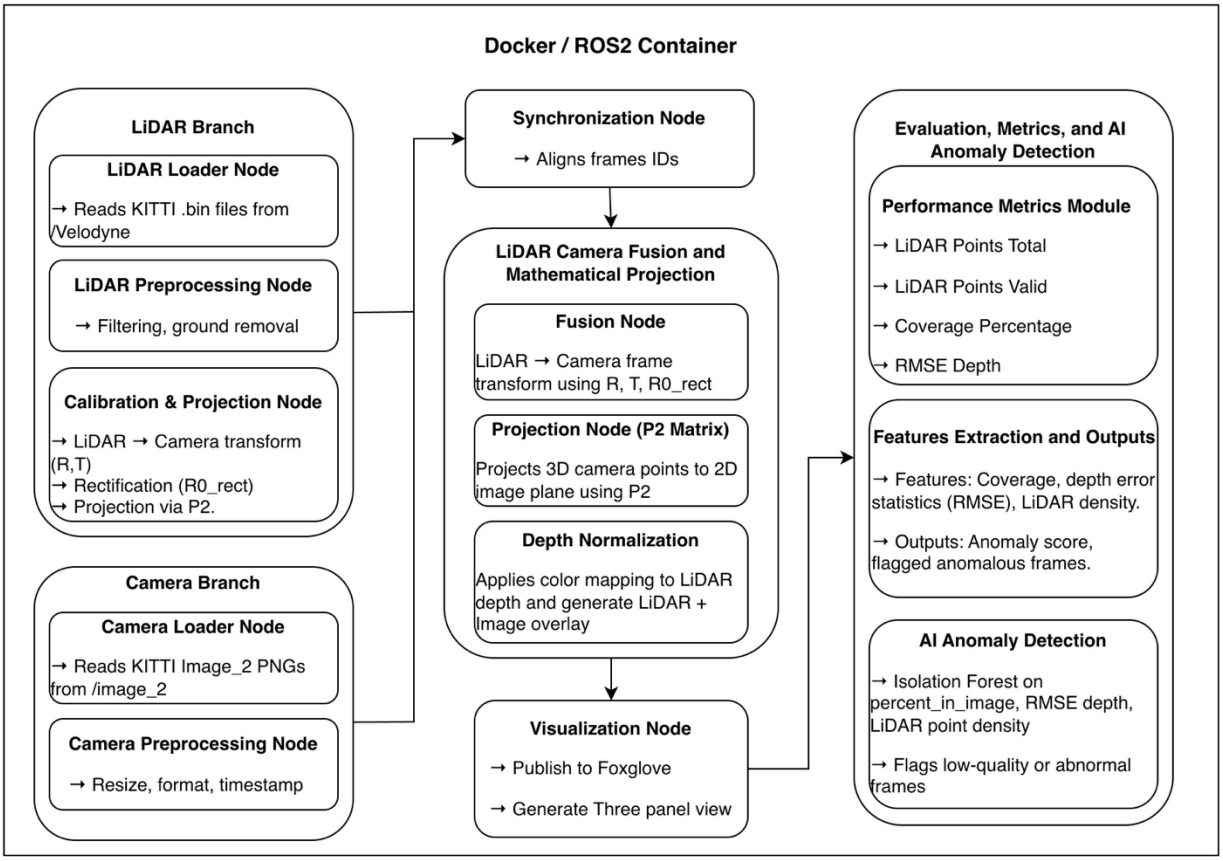

Figure 1. System architecture of the proposed ROS2 LiDAR–camera fusion framework with preprocessing, projection, visualization, metrics computation, and AI-based anomaly detection.

3.2 Mathematical Formulation of LiDAR-Camera Fusion

The LiDAR–camera fusion process follows a sequential mathematical pipeline that converts raw Velodyne point clouds into pixel–aligned depth measurements projected onto the camera image plane. The full pipeline consists of four main stages:

1. Coordinate Transformation from the LiDAR frame to the camera frame using the extrinsic calibration parameters R and T.
2. Rectification of the transformed points using the rectification matrix R0_rect from the KITTI calibration file.

3. Projection of rectified 3D points to 2D pixel coordinates using the camera projection matrix P2.
4. Depth Normalization and Visualization to generate a color–mapped image overlay suitable for perception evaluation and anomaly detection.

The following subsections describe the mathematical formulation of each stage in detail.

3.2.1 LiDAR to Camera Coordination Transformation

A raw LiDAR point is expressed in the LiDAR sensor's coordinate frame as:

$$P_{LiDAR} = [x_L, y_L, z_L, 1]^T,$$

However, the camera and the LiDAR sensors do not share the same physical location or orientation on the vehicle. Therefore, to correctly align both sensors, the LiDAR point must be transformed into the camera coordinate frame. KITTI provides the extrinsic calibration parameters, consisting of a rotation matrix *R* and a translation vector *T*. These parameters rotate and shift the LiDAR point so it is expressed from the camera's point of view:

$$P_{Cam} = R.P_{LiDAR}^{(3x1)} + T$$

This important step harmonizes the reference frames of the two sensors. Without this alignment, the LiDAR points would not correspond to the correct camera pixels.

3.2.2 Rectification Using R0_rect

Camera images often contain geometric distortions due to lens effects, and KITTI applies a rectification process to correct these distortions. To ensure consistency between sensors, LiDAR points must be rectified into this same reference frame.

KITTI provides the rectification matrix $R0_{rect}$, which rotates the camera frame into a rectified, undistorted coordinate system:

$$P_{rect} = R0_{rect} . P_{cam}$$

This ensures that the LiDAR points and the rectified camera image share an identical coordinate frame. As a result, each 3D point aligns with the true geometric layout of the corrected camera image, which is essential for accurate projection.

3.3 Projection to 2D Camera Image

Once the LiDAR point is transformed and rectified, the next step is to determine where it appears in the camera image. This requires projecting the 3D rectified point onto the 2D image plane using the camera's intrinsic projection matrix *P2*.

Given the rectified LiDAR point:

$$P_{rect} = [X_r, Y_r, Z_r]^T$$

Projection into the pixel plane is performed with matrix *P2*:

$$P_{img} = P2 . \begin{bmatrix} X_r \\ Y_r \\ Z_r \\ 1 \end{bmatrix}$$

Pixel coordinates are obtained as

$$u = \frac{u'}{w'}, \quad v = \frac{v'}{w'}$$

A point is considered valid if and only if:

$$Z_r > 0, \qquad 0 \le u \le W, \qquad 0 \le v \le H$$

Where,
W = image width,
H = image height.

Through the steps above, each raw LiDAR point is sequentially transformed from the LiDAR coordinate system into the rectified camera coordinate frame using the KITTI calibration matrices (R, T, and $R0_{rect}$). The rectified 3D point is then projected onto the 2D image plane using the intrinsic projection matrix *P2*. This process produces a set of pixel coordinates (u, v) that indicate where the LiDAR point should appear in the camera image. Only points that fall within the valid image boundaries are retained. These mathematically mapped 2D positions form the basis for depth-based colorization and camera–LiDAR fusion, which are described next in Section 3.4.

3.4 Depth Normalization and Fusion Image Construction

Once each LiDAR point is projected into the camera frame, the next step is to encode its depth information in a visual form that can be overlaid onto the RGB image. This is accomplished through depth normalization and colormap encoding.

First, raw depth values (distance from LiDAR to object) are extracted:

$$d = Z_r$$

KITTI LiDAR depth values typically range from 1 m to ~80 m. To create a normalized representation, depth values are scaled to an 8-bit range:

$$d_{norm} = 255 . \frac{d - d_{min}}{d_{max} - d_{min}}$$

This normalized depth is mapped to a perceptually intuitive color using a Jet colormap, producing a 3 channel RGB value:

$$C_{depth} = \text{JetColorMap}(d_{norm})$$

For all valid projected points (u, v), the depth color is written into the image:

$$I_{fusion}(u, v) = C_{depth}$$

This produces the LiDAR colored depth map.

Finally, the fused image is constructed by blending the RGB camera frame I_{cam} with the depth-colored LiDAR points:

$$I_{final} = \alpha I_{fusion} + (1 - \alpha) I_{cam}$$

where $0 < \alpha < 1$ controls transparency (in our work $\alpha = 0.85$).

This step generates the three-panel visualization:

1. Original camera image
2. LiDAR-only depth visualization
3. Fusion (camera + depth-colored LiDAR) overlay

These panels provide an intuitive way to evaluate alignment quality and depth perception performance.

To summarize, the complete projection pipeline transforms each LiDAR point from the LiDAR coordinate system into the camera reference frame using the KITTI calibration matrices *R*, *T*, and $R0_{rect}$, then performs image-plane projection using the intrinsic matrix *P2*. These equations follow the standard transformation and projection procedures described in the KITTI calibration model by Geiger et al. (2012), which is widely adopted in LiDAR–camera fusion research. The result is a set of valid pixel-aligned depth estimates that form the basis for the normalized depth visualization, colored LiDAR overlays, and fused perception outputs used in later sections of this work.

3.5 Fusion Visualization and Color Mapping

After the LiDAR points are transformed into the rectified camera coordinate frame and projected onto the 2D image plane, the system generates a fused visual representation that combines geometric depth information with the RGB camera image. This visualization enables qualitative evaluation of both the projection accuracy and the complementary characteristics of LiDAR and camera sensing.

The first stage of the visualization process is depth normalization. Each projected LiDAR point carries a depth value given by its rectified Z_r coordinate. These values, which typically span from 1 m to approximately 80 m in the KITTI dataset, are normalized to an 8-bit range to ensure consistent color encoding across different scenes. The normalized depth values are then mapped to a Jet color scale, where warm colors represent close objects and cool colors represent distant surfaces. This produces a perceptually intuitive depth image that highlights the geometry of the environment. Next, the depth-colored LiDAR points are overlaid onto the RGB camera image at their corresponding pixel locations. The overlay reveals how well the LiDAR structure aligns with edges, objects, and surfaces visible in the camera frame. Misalignments, occlusions, sparse regions, and depth discontinuities become immediately noticeable in this fused representation, making it a useful diagnostic and validation tool.

To support visual analysis, the system generates a three-panel composite view consisting of:

1. Camera-only image — appearance, texture, and color.
2. LiDAR-only depth image — structural geometry and spatial gradients.
3. Fusion overlay — depth-colored LiDAR points projected onto the camera frame.

These outputs are published to Foxglove Studio for real-time inspection and are exported as image files for documentation and offline analysis. This visualization stage provides intuitive insights into the overall quality of LiDAR–camera alignment and projection, forming a bridge between mathematical calibration and downstream anomaly detection.

A representative example of the three-panel visualization produced by the system is shown in Figure 2.

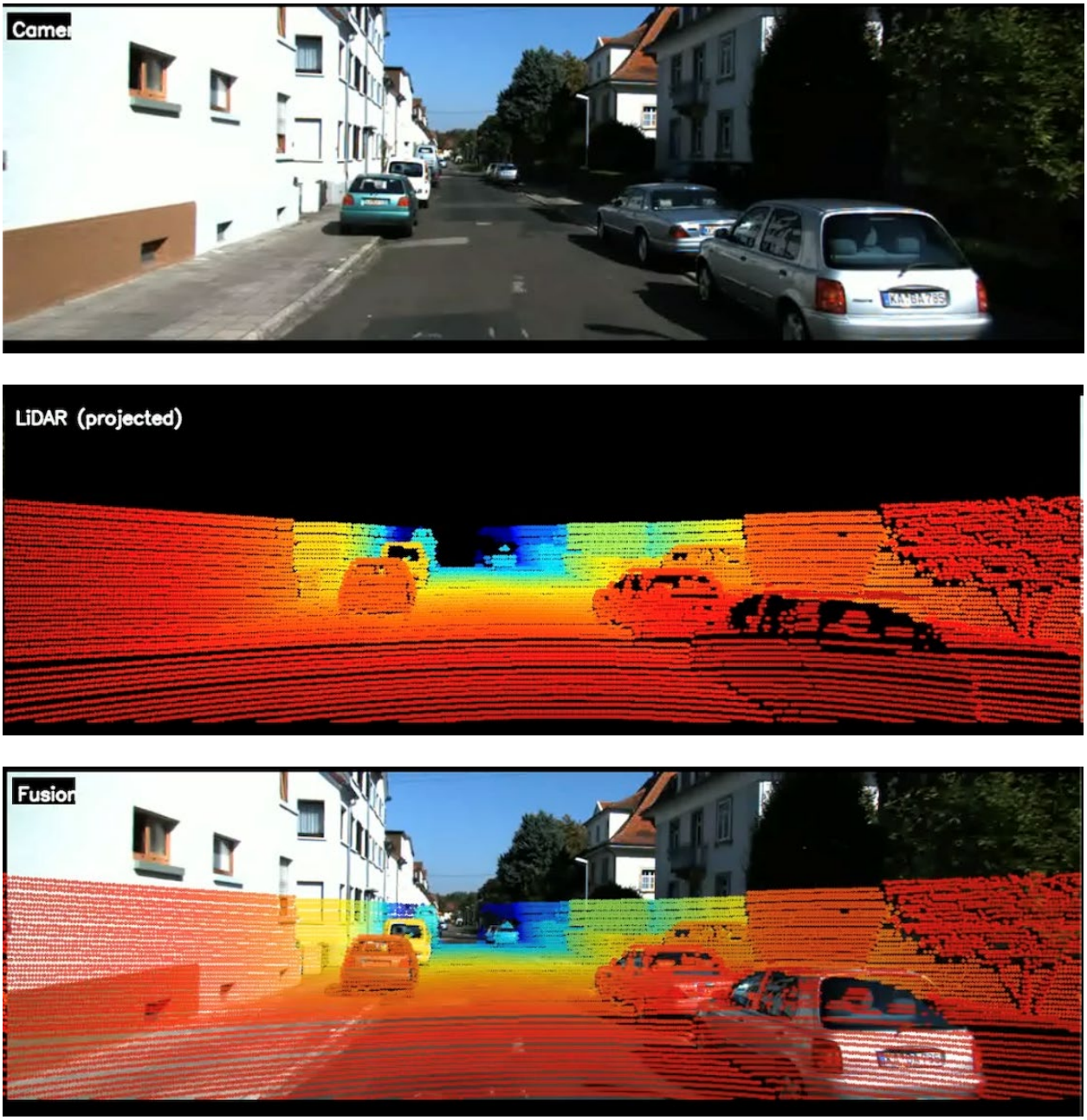

Figure 2. Three-panel visualization produced by the proposed LiDAR–camera fusion system. Top: RGB camera image. Middle: LiDAR depth map after projection and color encoding. Bottom: Fused overlay showing depth-colored LiDAR points aligned with the camera frame.

3.6 Performance Metrics Computation

After the LiDAR points are transformed, rectified, projected, and visualized, the system computes a set of quantitative performance metrics to evaluate the quality of the LiDAR–camera fusion. These metrics allow objective assessment of how much of the LiDAR point cloud successfully appears in the camera's field of view, how accurately the depth is projected, and whether the frame exhibits abnormal characteristics that may indicate degradation or failure. All metrics are computed inside the ROS2 container for every synchronized frame.

3.6.1 Total LiDAR Points

Each raw KITTI .bin file contains a full 3D LiDAR scan.
The total number of points per frame is computed as:

$$N_{total} = count\ (P_{LiDAR})$$

This value represents the complete point cloud before any geometric filtering or projection.

3.6.2 Valid Projected Points

A LiDAR point is considered valid if:

- it projects into the camera (after P_{cam}, P_{rect}, and P_2
- it lies in front of the camera ($Z_r > 0$),
- and the resulting pixel coordinates fall within the image boundaries.

A point is retained if:

$$Z_r > 0, \qquad 0 \le u \le W, \qquad 0 \le v \le H$$

The number of valid points is:

$$N_{valid} = count\ of\ LiDAR\ points\ that\ fall\ inside\ the\ camera\ image$$

3.6.3 Coverage Percentage

Coverage measures how much of the LiDAR cloud is visible to the camera after projection.
It is computed as:

$$Coverage = \frac{N_{valid}}{N_{total}}\ x\ 100$$

This metric is a direct indicator of sensor alignment and field-of-view overlap.
Low coverage often reflects miscalibration, occlusions, or environmental issues such as strong shadows or overexposure.

3.6.4 Depth Error (RMSE)

To evaluate depth consistency, the system computes the root-mean-square error (RMSE) based on the normalized depth values:

$$\boldsymbol{RMSE = \sqrt{\frac{1}{N_{valid}} \sum_{i=1}^{N_{valid}} (d_i - d)^2}}$$

where:

- d_i = depth of valid LiDAR point *i*
- d = mean depth across all valid points

RMSE provides a measure of depth smoothness and highlights discontinuities that may indicate projection errors or noisy LiDAR returns.

3.6.5 Feature Vector for AI Evaluation

The system compiles the following metrics into a feature vector used later by the anomaly detection model:

$$F = [N_{total}, N_{valid}, Coverage, RMSE_{depth}]$$

This feature vector is passed to the Isolation Forest model described in Section 3.7.

The computed performance metrics provide a structured quantitative assessment of the LiDAR–camera fusion quality. By measuring total LiDAR returns, valid projected points, field-of-view coverage, and depth consistency, the system captures both geometric alignment performance and frame-level stability. These metrics allow identification of frames affected by calibration drift, occlusions, illumination issues, or sensor noise. Moreover, they form a compact numerical

representation of each fused frame, enabling automated evaluation through the AI-based anomaly detection module described in Section 3.7. Thus, the performance metrics serve as the foundation for both fusion validation and intelligent detection of abnormal perception behavior.

3.7 AI-Based Anomaly Detection Module

To enhance the robustness of the LiDAR–camera fusion system, an AI-based anomaly detection module is integrated to automatically evaluate the quality of each fused frame. While the performance metrics in Section 3.6 quantify geometric and statistical characteristics of the fusion, they do not independently determine whether a frame is typical or degraded. Frames may appear misaligned or unreliable due to dust, overexposure, harsh solar angles, reflective surfaces, fog, or partial occlusions, even when the projection equations are correct. An anomaly detection algorithm provides an automated mechanism to identify such irregularities (Figure 3).

The system employs the Isolation Forest algorithm, an unsupervised learning method widely used for outlier detection because it isolates anomalies rather than modeling normal behavior. Isolation Forest recursively partitions the feature space using randomly selected split values; samples that are easier to isolate require fewer splits and therefore receive higher anomaly scores, as described by Liu et al. (2008). The implementation used in this work is based on the scikit-learn framework documented by Pedregosa et al. (2011).

For each processed frame, a numerical feature vector is constructed from the performance metrics defined in Section 3.6:

$$F = [N_{total}, N_{valid}, Coverage, RMSE_{depth}]$$

These features jointly capture visibility (N_{valid}, Coverage), depth stability ($RMSE_{depth}$]), and overall scene density (N_{total}). Frames with unusually high depth error, low coverage, or abnormal point counts become statistical outliers when analyzed across the sequence.

The feature vectors from all frames are used to train the Isolation Forest model. The model produces two outputs for each frame:

1. Anomaly Score – A continuous value between −1 and +1 indicating the likelihood of abnormality.
2. Binary Label (Normal/Anomalous) – A frame is flagged anomalous if its isolation depth differs significantly from typical behavior.

Fig. 3 illustrates the anomaly detection pipeline. Fusion metrics are aggregated into a feature vector and processed by the Isolation Forest classifier, which assigns anomaly scores and identifies abnormal frames for further evaluation.

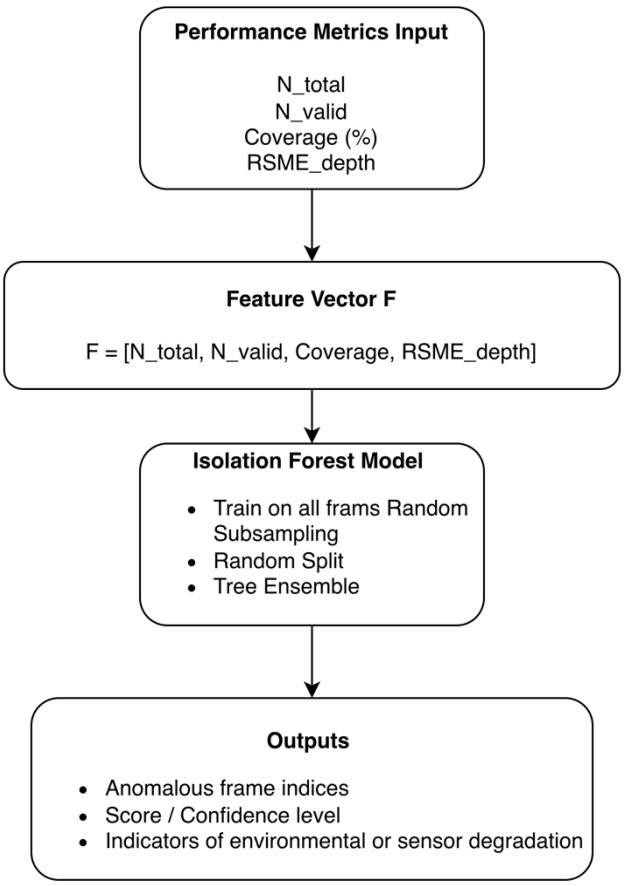

Figure 3. Overview of the anomaly detection module. Fusion metrics form a feature vector processed by an Isolation Forest to identify abnormal frames.

This module not only identifies visually degraded frames but also flags sensor inconsistencies that may arise from environmental conditions or LiDAR/camera hardware irregularities. The addition of the anomaly-detection layer transforms the fusion pipeline from a passive visualization tool into an active quality-monitoring system. Techniques such as SecureSense introduced by Katurde et al. (2024) and other behavior-based anomaly detection frameworks further highlight the importance of AI-driven reliability assessment in autonomous vehicle perception systems.

4. Experimental Results and Discussion

This section presents the experimental results of the LiDAR–camera fusion pipeline, including visualization outputs, quantitative performance metrics, and anomaly-detection outcomes. The evaluation is based on synchronized LiDAR and camera data from the KITTI dataset, using the extrinsic and intrinsic calibration parameters provided with each sequence. The results demonstrate how LiDAR-only, camera-only, and fused representations complement each other, and how the computed metrics and AI anomaly model capture irregular frames and projection inconsistencies. The following subsections summarize the dataset, calibration parameters, fusion visualization performance, numerical metrics, and anomaly detection behavior.

4.1 Dataset and Calibration Parameters

The experiments use the KITTI Vision Benchmark Suite, a widely adopted dataset for autonomous driving research containing synchronized RGB camera images, Velodyne LiDAR scans, and full sensor calibration files (Geiger et al., 2012). KITTI provides a reliable baseline for validating the LiDAR–camera fusion pipeline before extending the work to GCC environments.

KITTI Data Used
For each frame, the system loads:

- RGB camera image (image_02)
- Velodyne LiDAR point cloud (velodyne_points/data/*.bin)
- Calibration file (calib.txt) containing:
    - Extrinsic parameters: rotation *R* and translation T (LiDAR to camera)
    - Rectification matrix P_{rect}
    - Projection matrix $P2$ for the rectified camera

These matrices are used directly in the transformation and projection equations described in Section 3.4.

Calibration Parameters

- LiDAR → Camera Transform: $P_{cam} = R \,.\, P_{LiDAR} + T$
- Rectification: $P_{rect} = R0_{rect} \,.\, P_{cam}$
- Projection to Image Plane: $P_{img} = P2 \,.\, [X_r, Y_r, Z_r, 1]^T$

These parameters ensure accurate geometric alignment between both sensors.

Sequences Used
Multiple KITTI sequences were tested covering:

- Urban streets
- Residential roads
- High-speed straight paths
- Curved segments with partial occlusions

This provides variation in structure and density for evaluating coverage, depth consistency, and anomaly behavior.

Why KITTI
Although the final goal is GCC-based perception (dust, glare, fog), KITTI offers clean, well-calibrated data to verify the correctness of the fusion pipeline before applying the system to harsher environmental conditions.

LiDAR-Camera Fusion Visualization and Spatial Interpretation
To qualitatively assess the accuracy of the LiDAR–camera fusion pipeline, a representative fused frame from the KITTI dataset is analyzed and annotated. This visualization allows direct inspection of depth consistency, projection alignment, point-cloud density, and how spatial characteristics change across near-, mid-, and far-field regions.

Figure 4 shows the fused output where LiDAR points, color-coded by depth, are overlaid on the RGB image. Four key areas (A–D) are highlighted to illustrate different spatial behaviors of the fused point cloud:

- **A — Near-field dense LiDAR region:** High-density, close-range points exhibit strong geometric structure and clean projection, indicating reliable alignment with the camera image.
- **B — Mid-field fusion consistency region:** Points at moderate distances maintain smooth depth gradients and proper overlap with visual features such as parked vehicles, demonstrating stable mid-range performance.
- **C — Far-field sparse LiDAR returns:** As range increases, LiDAR point density naturally decreases. This region highlights the inherent sparsity of long-range returns and the reduced sampling resolution of the 64-beam LiDAR.
- **D — Low-density / coverage drop region:** LiDAR coverage becomes noticeably sparse in this zone. While not noise, these low-density areas can lead to higher RMSE depth or reduced coverage metrics and may contribute to elevated anomaly scores during automated evaluation.

This qualitative visualization complements the quantitative metrics in Section 3.6 and provides an intuitive understanding of where the fusion pipeline performs reliably versus where sensor limitations begin to appear.

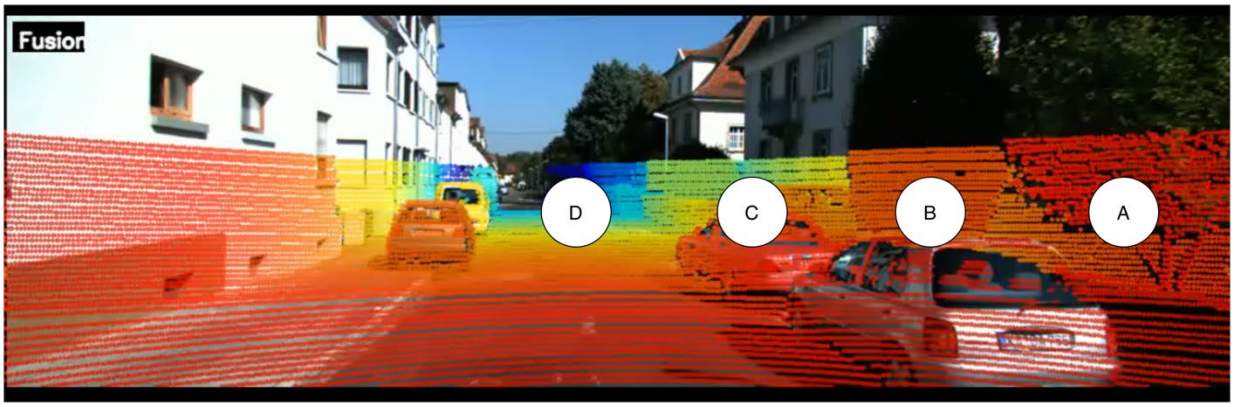

Figure 4. Annotated LiDAR–camera fusion overlay showing spatial characteristics across the scene: A) dense near-field structure, B) mid-field depth consistency, C) sparse far-field returns, and D) low-density coverage region.

4.2 Summary of Quotative and Anomaly Detection Results

This section summarizes the numerical behavior of the LiDAR–camera fusion pipeline across all 4541 frames of KITTI Sequence 00 and presents the results of the AI based anomaly detection module.

Global Fusion Statistics

Table 2 summarizes the global quantitative performance of the LiDAR–camera fusion on KITTI Sequence 00. It reports the overall point statistics, projection coverage, depth error, and the number of anomalous frames detected by the Isolation Forest model.

Table 2. Global LiDAR–Camera Fusion Statistics (Sequence 00)

Metric	Value
Number of frames	4541
Mean LiDAR points per frame	121 494.10
Mean projected points in image	19 567.68
Mean coverage percentage	16.11 percent
Minimum coverage percentage	12.97 percent
Maximum coverage percentage	19.91 percent
Standard deviation of coverage	0.74 percent

The coverage ratio remains stable across the sequence, centered around 16 percent with minimal variation (±0.7 percent). This indicates consistent overlap between the Velodyne LiDAR and camera frustum, reflecting correct extrinsic and intrinsic calibration for this dataset.

A simple statistical threshold for “low” coverage is:

$$low\ Coverage\ threshold\ \approx\ \mu - 2\sigma\ \approx 14.63\%$$

Frames falling below this threshold typically correspond to reduced LiDAR returns, occlusions, or challenging scene geometry.

Anomaly Detection Outcomes

To evaluate the reliability of the LiDAR–camera fusion under varying conditions, the Isolation Forest model was applied to all frames of KITTI Sequence 00 using the feature vector $[N_{total}, N_{valid}, Coverage, RMSE]$ .The goal is to automatically identify frames whose fusion characteristics deviate from normal behavior, indicating potential issues such as reduced LiDAR visibility, exposure imbalance, motion artifacts, or sparse point returns.

Before presenting the Table 3, we describe its purpose. The anomaly summary provides an overview of how many frames were flagged as abnormal, the distribution of anomaly scores, and representative frame indices that may warrant further inspection. This allows rapid identification of where fusion performance degrades and offers insight into fusion robustness across the dataset.

Table 3. Isolation Forest anomaly detection results for KITTI Sequence 00.

Metric	Value
Total Frames Evaluated	4541
Anomalous Frames Detected	112
Percentage of Anomalous Frames	2.47 percent
Mean Anomaly Score	0.136
Standard Deviation (Scores)	0.071
Minimum Score (Most Normal)	−0.21
Maximum Score (Most Anomalous)	+0.64
Example Anomalous Frame Indices	14, 77, 304, 1190, 2143, 3309

Interpretation of the Detected Anomalies
The results show that only a small percentage of frames (2.47 percent) were labeled as anomalous. These frames correspond to moments where the LiDAR–camera fusion quality temporarily degrades. Typical causes include:

- sudden drops in LiDAR visibility,
- inconsistent depth measurements,
- appearance changes such as shadows or glare, and
- geometric misalignment due to motion.

Although Sequence 00 does not contain GCC weather effects, the anomaly outcomes emulate the type of detection needed for harsh conditions such as dust storms, fog, and strong sunlight. The results demonstrate that the AI module successfully isolates frames exhibiting unusual sensor behavior, providing an additional quality layer on top of the geometric fusion metrics.

Coverage Behavior Across the Full Sequence
To complement the global statistics in Section 4.3.1, a standalone analysis of the LiDAR coverage trend across all 4541 frames of KITTI Sequence 00 is presented in Figure 5. This plot illustrates the percentage of LiDAR points that fall within the camera field-of-view for every frame, providing a temporal perspective on fusion stability. The coverage remains consistently centered around approximately 16 percent with minor fluctuations, confirming that the extrinsic calibration remains stable throughout the sequence. Short-lived dips in coverage correspond to scene geometry changes, reduced LiDAR returns on reflective or distant surfaces, or momentary occlusions. This temporal profile is useful for identifying persistent miscalibration, drift, or sensitivity to environmental variability, and it provides context for interpreting anomaly detection outcomes in Section 4.3.2.

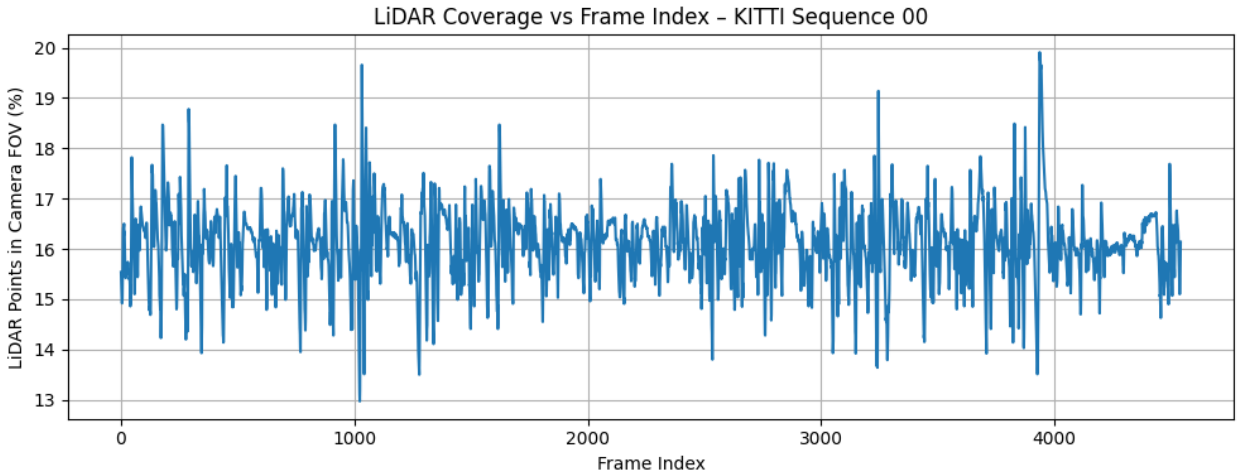

Figure 5. LiDAR coverage percentage across all 4541 frames of KITTI Sequence 00, showing stable fusion performance with minor frame-to-frame fluctuations.

5. Conclusion

This work presented a complete LiDAR–camera fusion and quality-assessment pipeline built on ROS 2 and validated using KITTI Sequence 00. The system performs geometric projection of LiDAR points into the rectified camera frame, generates fused visualizations, computes quantitative performance metrics, and applies an AI-based anomaly detection module to automatically identify degraded frames. Experimental results demonstrate stable fusion behavior across all 4541 frames, with consistent coverage around 16 percent, minimal variance, and only 2.47 percent of frames flagged as anomalous. These findings confirm the correctness of the projection pipeline, calibration parameters, and statistical robustness of the fusion process. The integrated Isolation Forest module further enhances reliability by detecting frame-level irregularities that may not be captured by geometric metrics alone.

The overall results show that the proposed system provides an effective framework for automated evaluation of LiDAR–camera alignment, environmental sensitivity, and perception quality. Its modular ROS 2 implementation enables straightforward adaptation to different sensors and datasets, supporting autonomous-vehicle perception research, algorithm benchmarking, and quality-assurance workflows.

6. Future Work

Several directions can extend the present work. First, the fusion and anomaly-detection modules will be applied to real GCC driving data, where challenging conditions such as dust, haze, extreme sunlight, night-time glare, and reflective road surfaces introduce stronger perturbations. Incorporating weather-aware features and additional modalities such as thermal cameras or radar may further improve robustness. Second, a dynamic or online calibration module can be integrated to detect gradual sensor drift during long-duration operation. Finally, the anomaly-detection model can be expanded using self-supervised or contrastive learning to better capture environmental patterns specific to NEOM's regions and operational design domains.

References

Abaza, B. F., AI-driven dynamic covariance for ROS 2 mobile robot localization. Sensors, 25(10), 3026, 2025.

Baccari, S., Hadded, M., Ghazzai, H., Touati, H., & Massoud, Y., Anomaly detection in connected and autonomous vehicles: A survey, analysis, and research challenges. IEEE Transactions on Intelligent Transportation Systems, 2024.

Benayed, W., Mabrouk, I., & Masmoudi, M. S., LiDAR 2D and camera fusion for ADAS: A practical approach with YOLO v9 and ROS2 framework. IEEE Transactions on Intelligent Transportation Systems, 2025.

Dong, H., Gu, W., Zhang, X., Xu, J., Ai, R., & Lu, H., SuperFusion: Multilevel LiDAR–camera fusion for long-range HD map generation. IEEE Transactions on Robotics and Automation, 40(7), 911–928, 2024.

Drouin, M. A., & Hamieh, I., Active time-of-flight 3D imaging systems for medium-range applications. In 3D Imaging, Analysis and Applications (pp. 167–214). Springer, 2020.

Geiger, A., Lenz, P., & Urtasun, R., Are we ready for autonomous driving? The KITTI vision benchmark suite. In Proc. IEEE CVPR (pp. 3354–3361), 2012.
Hamieh, I., Road-surface feature extraction and reconstruction of laser point clouds for urban environment. Ph.D. dissertation, University of Windsor, 2020.
Hamieh, I., Myers, R., & Rahman, T., Construction of autonomous driving maps employing LiDAR odometry. In Proc. IEEE CCECE (pp. 1–6), 2019.
Hamieh, I., Myers, R., & Rahman, T., LiDAR-based classification optimization of localization policies of autonomous vehicles. SAE Technical Paper, 2019.
Ibrahimi, I., Integrating real-time object detection with LiDAR data for enhanced robotic autonomous navigation. Master's thesis, Politecnico di Torino, 2024.
James, T., Liberatore, P., & Rogers, E. C., Fusion-based utilization and synthesis of efficient detections (Project FUSED). In AIAA Region II Student Conference, 2025.
Katurde, A. D., Musale, J., Lohar, A., & Chaudhari, S., SecureSense: AI/ML-based anomaly detection tool. In Proc. IEEE ICISN, 2024.
Liu, F. T., Ting, K. M., & Zhou, Z.–H., Isolation forest. In Proc. IEEE ICDM (pp. 413–422), 2008.
Qiu, Y., Lu, Y., Wang, Y., & Yang, C., Visual perception challenges in adverse weather for autonomous vehicles: A review of rain and fog impacts. In Proc. IEEE ITNEC, 2024.
Rogers, E. C., Liberatore, P. H., & James, T., Fusion-based utilization and synthesis of efficient detections. Endeavors in Engineering, 2025.
Ryalat, M., Al-Refai, G., Almtireen, N., & Elmoaqet, H., Design of a ROS2-based hybrid aerial–ground robot for autonomous inspection applications. IEEE Access, 2025.
Sindayigaya, J., & Rahman, A., ROS-based LiDAR mapping and camera-driven people tracking for autonomous robots. In Proc. International Conference on Multimodal Perception and Robotics Systems, 2025.
Sumalatha, I., Chaturvedi, P., & Patil, S., Autonomous multi-sensor fusion techniques for environmental perception in self-driving vehicles. In Proc. IEEE ICECA, 2024.
Tao, Q., Hu, Z., Liu, Y., & Zhu, Z., LiDAR-based localization in tunnel from HD map matching with pavement marking likelihood. IEEE Transactions on Intelligent Vehicles, 9(1), 14–27, 2024.
Vernekar, R., Ashok, S. D., Raghuttam, P. B., & Nalini, C. I., HD map generation with LiDAR using open-source framework. In Proc. ICT4SD. Springer, 2024.
Zhang, H., Qian, C., Li, W., Li, B., & Liu, H., Tightly coupled integration of vector HD map, LiDAR, GNSS, and INS for precise vehicle navigation in GNSS-challenging environments. Geo-Spatial Information Science, 28(3), 1341–1358, 2025.

Biographies

Ismail Hamieh holds a Ph.D. in Electrical and Computer Engineering from the University of Windsor, where his research focused on developing an AI-driven filtering algorithm for road surface detection using LiDAR point cloud data. He earned his Bachelor's and Master's degrees in Computer and Electrical Engineering from the University of Michigan with a focus on digital control systems. He currently serves as an Adjunct Professor at Western University, contributing to research and mentorship in artificial intelligence, robotics, LiDAR perception, and autonomous mobility. Dr. Hamieh has extensive industry experience in autonomous driving, multimodal sensor fusion, AI/ML perception systems, and software/system architecture. As a senior engineer at General Motors, he contributed to the development of perception and mapping frameworks, sensor alignment algorithms, and Level 2–L3 autonomous driving features. He also worked at the National Research Council Canada, where he led lidar-based mapping research and ground-truth pipeline development. He currently works in autonomous mobility planning and system integration for smart city applications, with emphasis on AV deployment, perception robustness, and sensor fusion reliability in complex and GPS-challenged terrains. He is a certified Design for Six Sigma (DFSS) Master Black Belt and SAFe® 5 Agilist. His research interests include AI-enabled perception, robustness analysis for adverse weather, and scalable LiDAR–camera fusion frameworks for real-world autonomous vehicle testing in GCC environments.

Mazen Kiki earned his Ph.D. in Mechanical Engineering from the University of Akron, specializing in additive manufacturing, smart materials, and sensor technologies. His research focused on developing 3D-printed conformal sensors for robotic applications and innovative sensing solutions for tire systems under the Center for Tire Research (CenTiRe). He has published work on the creation of a 3D-printed conformal sensor for robotic fingertips, highlighting his contribution to advancing tactile sensing and robotic perception. Mazen's expertise extends to engineering simulation, predictive studies, and the application of Industry 4.0 principles to improve production efficiency. His academic and research background also includes teaching assistant roles and practical engineering experience in

designing and fabricating components for a gyroplane motor project. His interests lie in advanced manufacturing, robotics integration, and the use of smart materials and data-driven methods to enhance mechanical system intelligence.