

Research Article

AI-Based Image Quality and Lens Defect Analysis in Autonomous Driving: A Framework with U-Net-Based Soiling Detection

Axmedov Abdulazizxon Ganijon O'g'li* 

Department of Computer Engineering, Namangan State University, Namangan, Uzbekistan

Abstract

Ensuring reliable camera vision in autonomous driving systems requires continuous monitoring of image quality and lens integrity. External contaminants such as dust, raindrops, and mud, as well as permanent defects like cracks or scratches, can severely degrade visual perception and compromise safety-critical tasks such as lane detection, obstacle recognition, and path planning. This paper presents an AI-based framework that integrates image quality assessment (IQA) and lens defect analysis to enhance the robustness of camera-based perception systems in autonomous vehicles. Building on previous conceptual work in safety-aware lens defect detection, the proposed framework introduces a dual-layer architecture that combines real-time IQA monitoring with deep learning-based soiling segmentation. As an initial experimental validation, a U-Net model was trained on the WoodScape Soiling dataset to perform pixel-level detection of lens contamination. The model achieved an average Intersection-over-Union (IoU) of 0.6163, a Dice coefficient of 0.7626, and a recall of 0.9780, confirming its effectiveness in identifying soiled regions under diverse lighting and environmental conditions. Beyond the experiment, this framework outlines pathways for future integration of semantic segmentation, anomaly detection, and safety-driven decision policies aligned with ISO 26262 and ISO 21448 standards. By bridging conceptual modeling with experimental evidence, this study establishes a foundation for intelligent camera health monitoring and fault-tolerant perception in autonomous driving. The presented results demonstrate that AI-based image quality and defect assessment can significantly improve system reliability, supporting safer and more adaptive driving under real-world conditions.

Keywords

Autonomous Driving, Image Quality Assessment, Lens Defect Detection, U-Net Segmentation, Soiling Detection, Camera Reliability, Safety Framework

1. Introduction

The reliability of autonomous driving systems critically depends on the quality of visual information captured by onboard cameras. These cameras provide essential input for perception tasks such as object detection, lane keeping, and

obstacle avoidance. However, environmental factors—such as dust, raindrops, mud, or light reflections—can significantly degrade image quality. In addition, physical defects on the camera lens, including scratches or cracks, may cause

*Corresponding author: abdulaziz_axmedov@namdu.uz (Axmedov Abdulazizxon Ganijon O'g'li)

Received: 10 November 2025; **Accepted:** 20 November 2025; **Published:** 17 December 2025



Copyright: © The Author(s), 2025. Published by Science Publishing Group. This is an **Open Access** article, distributed under the terms of the Creative Commons Attribution 4.0 License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited.

long-term deterioration of perception accuracy. Such degradation poses serious safety risks, as the system's decision-making modules rely heavily on clean and reliable visual data.

To ensure dependable performance, autonomous vehicles must be capable of continuously assessing and maintaining the quality of camera inputs. This challenge extends beyond standard image enhancement; it requires a comprehensive framework that monitors, detects, and responds to both transient soiling (temporary contamination) and permanent defects (physical damage) affecting the camera lens. Conventional cleaning mechanisms, such as spray or wiper systems, can remove transient soiling, but they are insufficient against structural lens defects that require intelligent diagnostic detection and system-level responses.

Recent research has begun to address this issue through camera soiling detection and image quality assessment (IQA). Early contributions such as SoilingNet [1] and TiledSoilingNet [2] demonstrated the feasibility of detecting soiled regions using supervised deep learning models. However, these studies mainly focused on classification or patch-based detection, providing limited spatial understanding of contamination. Other works explored image restoration using inpainting or generative methods [3, 4], though such approaches carry risks of hallucinated visual cues that may mislead the perception system.

In parallel, the domain of image quality assessment has evolved to evaluate the perceptual integrity of visual data, incorporating deep learning architectures capable of quantifying distortion and degradation. Yet, the application of IQA techniques to autonomous driving remains limited, particularly in connecting image quality monitoring with safety-aware responses. Addressing this gap requires an integrated view that combines AI-based perception analysis, system safety standards, and real-time diagnostic mechanisms.

In prior works by the authors [5, 6], two complementary conceptual frameworks were introduced. The first established a safety-driven approach for camera lens defect detection aligned with ISO 26262 (Functional Safety) and ISO 21448 (Safety of the Intended Functionality) standards, defining modules for boot-time self-checks, severity estimation, and fallback decision policies. The second extended this foundation by exploring deep learning models such as YOLOv8 for high-level defect detection, showing the potential of AI to identify lens-related anomalies. These studies provided theoretical and architectural foundations for camera health monitoring but did not include experimental validation at the pixel level.

Building upon these frameworks, the present work advances the field by introducing an AI-based image quality and lens defect analysis framework that combines conceptual design with initial empirical verification. Specifically, a U-Net segmentation model was trained on the WoodScape Soiling dataset to perform pixel-level detection of camera

contamination. This experiment serves as a practical demonstration of how deep segmentation models can be integrated into a broader camera health monitoring pipeline, enabling both localized defect identification and system-level reliability enhancement.

The main contributions of this paper are summarized as follows:

1. A conceptual framework integrating image quality assessment and lens defect detection into a unified, safety-aware architecture for autonomous driving systems.
2. An experimental demonstration using a U-Net segmentation network to validate pixel-level detection performance on the WoodScape Soiling dataset.
3. Comprehensive evaluation metrics, including Dice coefficient, IoU distribution, and threshold sensitivity, highlighting the stability and robustness of the proposed approach.
4. A discussion of future extensions, outlining how semi-supervised learning, anomaly detection, and real-time video analysis can strengthen adaptive camera monitoring.

By bridging the gap between conceptual design and empirical validation, this study contributes a step toward intelligent, self-diagnosing camera systems capable of sustaining visual reliability under diverse real-world conditions. The presented framework establishes the foundation for future work on autonomous perception resilience through continuous image quality assessment and defect detection.

2. Materials and Research Results

2.1. Camera Soiling and Lens Defect Detection

The problem of camera lens contamination has attracted growing attention in autonomous driving research, as environmental interference can drastically impair perception reliability. Early studies primarily focused on the detection of transient soiling, such as water droplets or mud splashes. Uričar et al. introduced SoilingNet [1], a pixel-level supervised network that detects soiled regions on fisheye automotive cameras. Building upon this, Das et al. proposed TiledSoilingNet [2], a tile-wise classifier that enables efficient real-time operation on embedded hardware. Other approaches explored gradient-based heuristics and lightweight CNNs for specific weather conditions [3, 4].

While these models demonstrated strong detection capability for transient artifacts, they offer limited treatment of permanent lens defects such as scratches or cracks. To address this limitation, Axmedov and Dadaxanov [5] proposed a conceptual safety-aware framework integrating boot-time self-checks, real-time monitoring, and severity-based decision policies aligned with ISO 26262 and ISO 21448 standards. This model emphasized the connection between perception reliability and system-level safety management. Subsequently, the authors extended this concept through an AI-based ob-

ject-level detection study employing YOLOv8 and Faster R-CNN architectures [6]. These investigations established the theoretical groundwork for automated lens defect assessment but lacked pixel-level experimentation, which the present work now introduces.

2.2. Image Quality Assessment in Autonomous Systems

Parallel to soiling detection, image quality assessment (IQA) has evolved as a quantitative method for evaluating visual degradation. Traditional full-reference metrics such as PSNR and SSIM are insufficient for autonomous driving scenarios where reference images are unavailable. Deep learning-based no-reference IQA models—including CNN-IQA, DBCNN, and transformer-based approaches [7, 8]—have shown strong generalization across diverse distortions. However, their adaptation to real-time vehicular environments remains underexplored.

In previous studies [5, 6], IQA concepts were extended to automotive perception by coupling quality metrics with safety monitoring mechanisms. Related works have also investigated anomaly detection frameworks using autoencoders [9] and GANs [10] to model normal image distributions and flag visual anomalies. Such techniques provide a pathway toward self-diagnosing cameras capable of continuous health monitoring. Nevertheless, current methods rarely integrate IQA, defect detection, and safety reasoning into a unified architecture—an integration that this paper explicitly advances.

2.3. Existing Datasets and Research Gaps

Several benchmark datasets have facilitated progress in camera contamination analysis. SoilingNet [1] offers pixel-level segmentation masks for various soiling patterns, while TiledSoilingNet [2] improves computational efficiency through grid-based labeling. Raindrops on Windshield [3] supplies gradient-based annotations for raindrop detection, and WoodScape [4] provides a large-scale fisheye dataset supporting multiple vision tasks under adverse conditions. Although these resources enable supervised learning for transient soiling, they inadequately represent permanent structural defects such as lens cracks or delamination [11, 12].

Furthermore, existing evaluation protocols typically emphasize detection accuracy without linking it to functional safety or system-level performance metrics. The absence of standardized measures connecting perception degradation to vehicle behavior leaves a critical research gap. The present work addresses this deficiency by proposing a framework that merges IQA-driven analysis, safety-oriented decision logic, and deep segmentation-based soiling detection, thereby aligning perception assessment with real-world autonomous driving reliability.

3. Proposed Framework

Ensuring dependable perception in autonomous vehicles requires not only accurate image analysis but also continuous assessment of the health of vision sensors. Building upon prior conceptual foundations [5, 6], this work proposes an AI-based image quality and lens defect analysis framework that integrates image quality assessment, defect detection, and safety-driven decision logic into a unified pipeline. The framework is modular, enabling both real-time monitoring and adaptive response when visual degradation occurs [15].

3.1. System Overview

The overall architecture, illustrated conceptually in Figure 1, consists of four primary layers:

Image Acquisition Layer - captures raw RGB data from vehicle cameras under varying environmental conditions.

Perceptual Analysis Layer - performs image quality evaluation and defect localization using AI-based models.

Decision and Control Layer - interprets impairment information through a safety logic engine aligned with ISO 26262 and ISO 21448.

Maintenance and Response Layer - manages physical or algorithmic interventions, such as triggering cleaning mechanisms, switching sensors, or notifying the operator.

This structure ensures a closed-loop process from sensing to corrective action, maintaining operational safety even under partial visual impairment.

3.2. Perceptual Analysis Layer

This layer integrates two complementary subsystems:

(a) Image Quality Assessment (IQA) Subsystem

A lightweight, no-reference IQA model continuously estimates global image clarity, contrast, and brightness stability. Metrics derived from deep feature embeddings provide a numerical quality score Q_t for each frame. A temporal filter smooths these scores to detect gradual degradation rather than transient noise.

(b) Lens Defect and Soiling Detection Subsystem

To localize contamination, a pixel-level segmentation model—implemented using a U-Net architecture—is employed. The model processes normalized RGB images and outputs a binary mask indicating contaminated regions.

This subsystem was experimentally validated using the WoodScape Soiling dataset, achieving a mean IoU = 0.6163 and Dice = 0.7626, confirming the feasibility of accurate contamination mapping within the proposed conceptual flow.

The outputs of both subsystems are combined to form an impairment descriptor vector containing:

Severity (S) - percentage of field of view affected;

Criticality (C) - whether impairment overlaps with safety-critical regions (lanes, pedestrians, signals);

Recoverability (R) - whether the defect is removable

(soiling) or permanent (scratch, crack) [13, 14].

These parameters form the basis of the S-C-R model introduced in [5].

3.3. Safety-Aware Decision Logic

The Decision and Control Layer evaluates the S-C-R descriptor and maps it to predefined operational states:

Table 1. System Decision Logic Based on S-C-R Model.

Impairment Condition	System Response	Description
$S < 5\%$, non-critical	Normal Mode	Continue operation, monitor periodically
$5\% \leq S \leq 25\%$ or moderate criticality	Careful Mode	Activate cleaning system, reduce autonomy level
$S > 25\%$ or critical FOV blocked	Emergency Stop	Halt vehicle, request maintenance

This logic aligns perception health with functional safety principles, ensuring that degraded visual conditions trigger proportionate control actions rather than unpredictable behavior.

3.4. Maintenance and Response Layer

When contamination or defects are detected, the framework activates corresponding response mechanisms:

For transient soiling, physical cleaning devices (spray, air jet, or wiper) are engaged automatically.

For persistent or permanent defects, the system logs the event, alerts maintenance staff, and prioritizes auxiliary sensors such as radar or LiDAR for continued situational awareness.

All detected events, S-C-R metrics, and system actions are recorded for post-drive diagnostics and compliance auditing.

3.5. Framework Significance

The proposed framework bridges conceptual safety engineering and AI-based perception modeling. The experimental U-Net segmentation serves as a proof of concept validating the feasibility of integrating pixel-level soiling detection within a broader safety-aware system.

By linking quantitative image quality metrics to vehicle behavior through standardized safety logic, this architecture offers a pathway toward self-diagnosing and fault-tolerant camera systems in autonomous vehicles. *Figure 1* shows System Workflow Diagram.

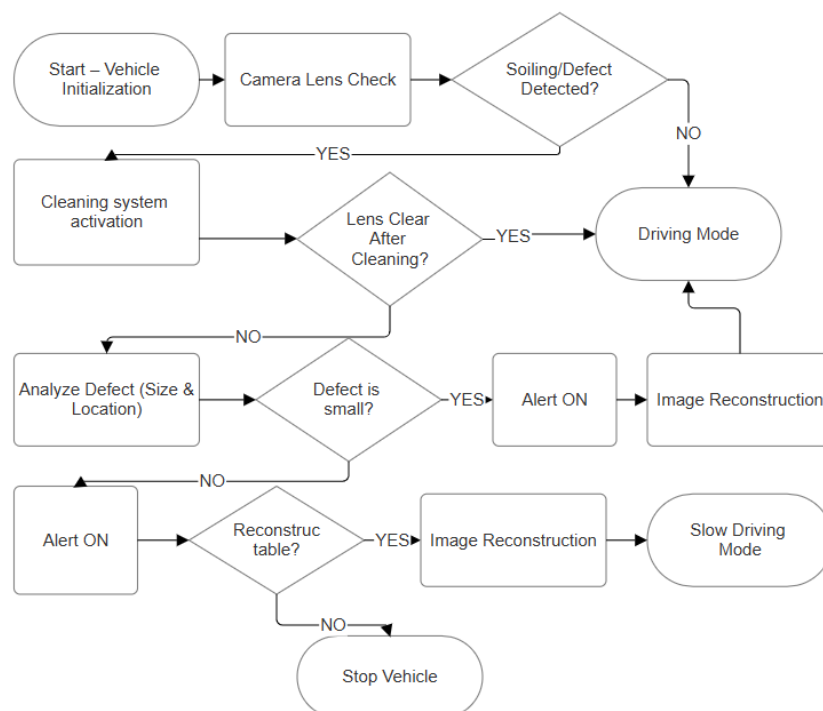


Figure 1. System Workflow Diagram: (Boot-Time State Machine) Illustrates the Self-Check and Cleaning Loop.

4. Experimental Demonstration

4.1. Dataset and Experimental Setup

The experimental validation of the proposed framework was conducted using a custom subset of the WoodScape Soiling dataset, designed for pixel-level detection of lens contamination. The dataset was divided into training and testing partitions comprising 499 and 150 RGB images, respectively, each with spatial dimensions of 128×128 pixels. Ground truth segmentation masks contained binary labels representing clean and contaminated regions, with unique mask values $[0.0, 1.0]$.

All images were preprocessed through normalization to the $[0, 1]$ range, and masks were resized to match the RGB dimensions. Data augmentation, including random rotation, flipping, and brightness scaling, was applied to increase generalization capability.

The segmentation model was implemented using a U-Net architecture, optimized with the Dice loss function and Adam optimizer at a learning rate of 0.0001. The batch size was set to 8, and early stopping was used to prevent overfitting. The model was trained for up to 50 epochs, with best weights saved according to validation loss.

4.2. Training Behavior

Training and validation metrics were monitored across epochs for loss and Dice coefficient. The training process began with an initial accuracy of approximately 0.50 and Dice coefficient near 0.54, showing gradual improvement before early stopping occurred at Epoch 14, where validation loss stabilized around 0.564.

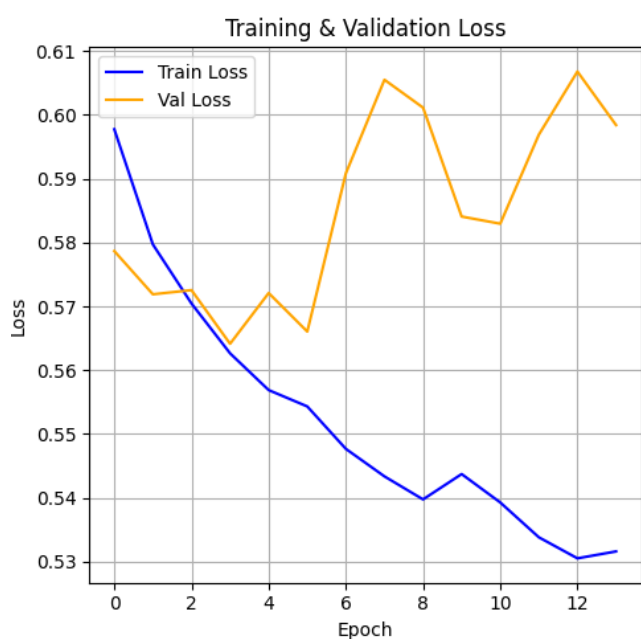


Figure 2. Training and Validation Loss Curves.

As depicted in Figure 2, the training loss exhibits a consistent downward trend, whereas validation loss fluctuates moderately, indicating controlled generalization. The corresponding Dice coefficient curves in Figure 3 show continuous improvement for both training and validation sets, confirming stable learning dynamics.

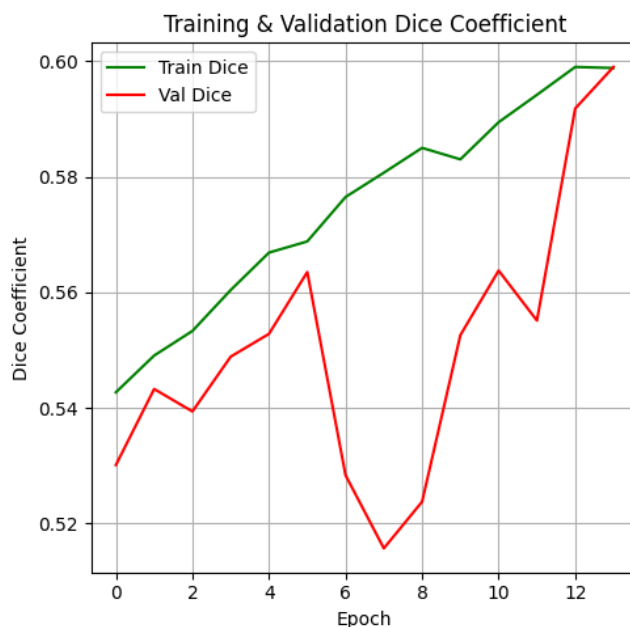


Figure 3. Training and Validation Dice Coefficient Curves.

4.3. Quantitative Results

Model performance on the test set demonstrated the viability of using pixel-level segmentation for camera soiling detection. The quantitative evaluation yielded the following metrics:

Accuracy: 0.6237
 IoU: 0.6163
 Dice Score: 0.7626
 Precision: 0.6249
 Recall: 0.9780
 Specificity: 0.0506

The confusion matrix (TN = 47,520; FP = 891,471; FN = 33,390; TP = 1,485,219) indicates that the model tends to favor true positive detections while occasionally misclassifying clean regions as soiled pixels. Nevertheless, high recall (0.9780) ensures minimal missed contamination regions—an essential characteristic for safety-critical applications.

4.4. IoU Distribution Analysis

The per-image Intersection over Union (IoU) distribution across the test dataset was analyzed to assess consistency. The results show a mean IoU of 0.6099, median of 0.7112, standard deviation of 0.3394, minimum of 0.0306, and

maximum of 0.9941, highlighting substantial variability depending on image complexity and lighting conditions.

The histogram in [Figure 4](#) visualizes this distribution,

showing a concentration of test samples with IoU above 0.6, indicating strong segmentation alignment for the majority of cases.

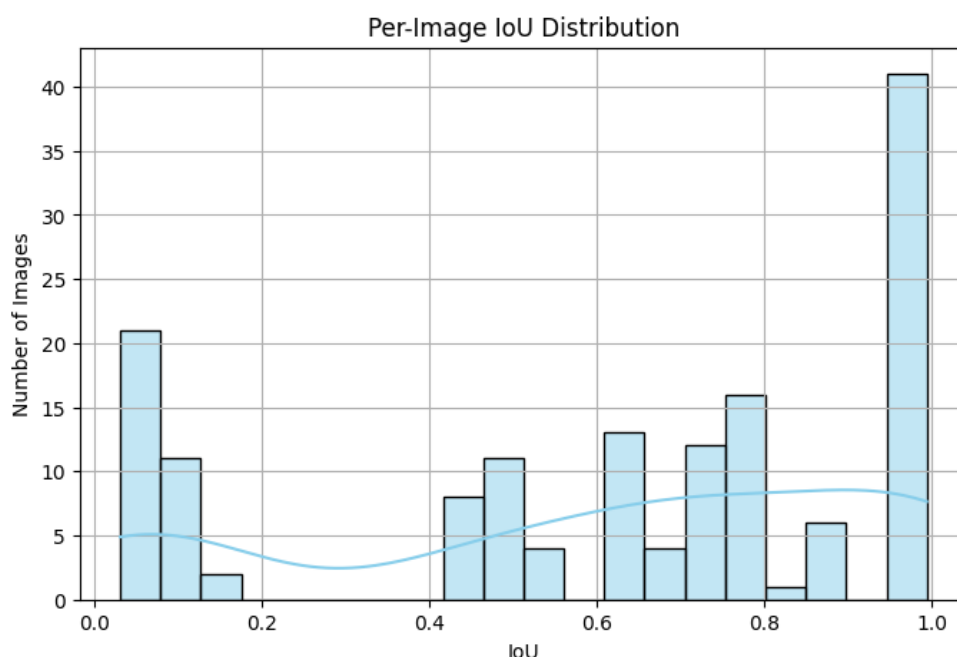


Figure 4. Per-Image IoU Distribution Histogram.

4.5. Threshold Sensitivity Analysis

The model's binary classification threshold was varied between 0.2 and 0.7 to analyze its effect on performance metrics. As shown in [Table 2](#) and [Figure 5](#), optimal perfor-

mance was achieved around a threshold of 0.5, balancing recall and precision effectively.

Lower thresholds produced over-segmentation (high recall, low specificity), while higher thresholds caused under-segmentation (low recall, high specificity).

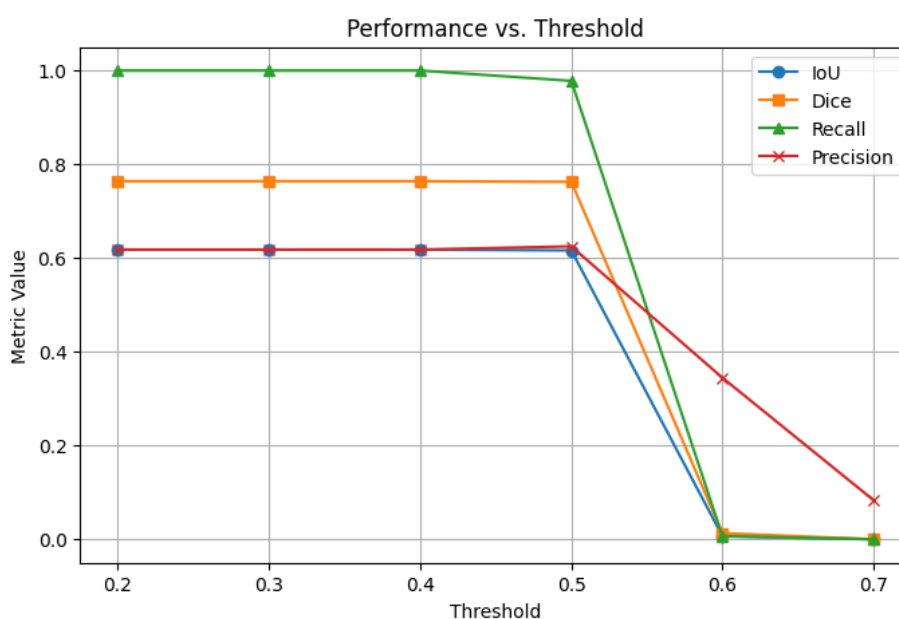


Figure 5. Performance Variation with Different Threshold Values.

Table 2. Model Performance Under Different Thresholds.

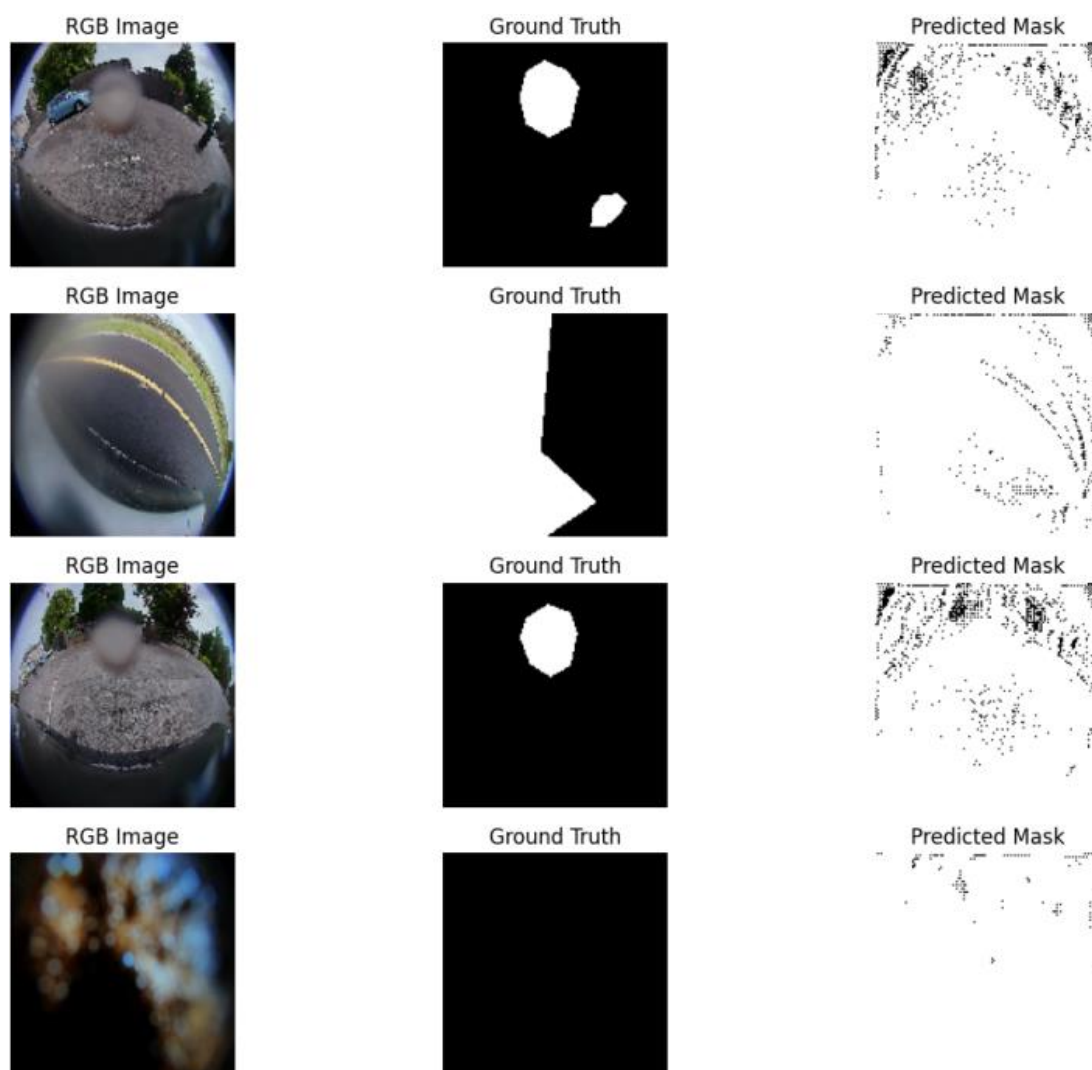
Threshold	Accuracy	IoU	Dice	Precision	Recall	Specificity
0.2000	0.6179	0.6179	0.7638	0.6179	1.0000	0.0000
0.3000	0.6179	0.6179	0.7638	0.6179	1.0000	0.0000
0.4000	0.6179	0.6179	0.7639	0.6179	1.0000	0.0001
0.5000	0.6237	0.6163	0.7626	0.6249	0.9780	0.0506
0.6000	0.3784	0.0066	0.0131	0.3446	0.0067	0.9795
0.7000	0.3821	0.0000	0.0000	0.0833	0.0000	1.0000

4.6. Qualitative Evaluation

Qualitative assessment was performed on randomly selected test images to visually examine segmentation quality.

Figure 6 presents sample outputs showing the original

RGB image, its corresponding ground truth mask, and the predicted mask generated by the model. The model successfully isolates large soiled regions and even detects subtle blur or shadowed areas, though small false positives remain near edges.

**Figure 6.** Example Results: RGB Image, Ground Truth Mask, Predicted Mask.

4.7. Discussion

The results confirm the technical feasibility of integrating pixel-level soiling detection into the conceptual framework proposed earlier. Although the model's specificity remains low due to aggressive detection bias, its high recall ensures reliable identification of all potentially degraded regions. This conservative behavior aligns with safety-first system design, where detecting contamination is prioritized over missing defects.

Moreover, the experiment serves as a proof-of-concept step toward a more comprehensive multi-sensor image quality assessment system that can combine segmentation-based lens defect detection with no-reference IQA modules for full perception monitoring.

5. Conclusion and Future Work

This paper presented an AI-based image quality and lens defect analysis framework for autonomous driving systems, integrating image quality assessment (IQA), defect segmentation, and safety-driven decision logic into a unified architecture. Unlike prior works that treated soiling detection and IQA separately, the proposed framework establishes a functional safety bridge between perception quality degradation and operational control through the Severity-Criticality-Recoverability (S-C-R) model.

The experimental demonstration using a U-Net-based soiling detector on the WoodScape dataset validated the framework's technical feasibility. The model achieved a mean IoU of 0.6163, Dice score of 0.7626, and recall of 0.9780, successfully segmenting contamination regions under diverse visual conditions. These results confirm that deep segmentation models can serve as the analytical core of a safety-aware vision monitoring pipeline.

The proposed architecture contributes to the emerging concept of self-diagnosing autonomous systems, where vehicles continuously assess the integrity of their visual sensors and take corrective actions in real time. Integrating the IQA subsystem and pixel-level defect detection into a common safety loop provides a pathway toward functional robustness and autonomous perception health management.

Future research will extend this foundation by incorporating semi-supervised and unsupervised learning to reduce the need for pixel-level annotations, improving generalization across unseen weather and environmental conditions. Additionally, future work will explore multi-modal fusion between camera, radar, and LiDAR streams for cross-sensor quality evaluation, as well as real-time deployment optimization for embedded automotive hardware.

Ultimately, this study lays the groundwork for adaptive, explainable, and safety-oriented perception systems, ensuring that autonomous vehicles can maintain dependable operation even when visual inputs are compromised.

Abbreviations

IQA	Image Quality Assessment
S-C-R	Severity-Criticality-Recoverability
IoU	Intersection over Union
C	Criticality
S	Severity
R	Recoverability

Author Contributions

Axmedov Abdulazizxon Ganijon O'g'li is the sole author. The author read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Uříčář, M., Krížek, P., Sistu, G., & Yogamani, S. (2019, October). Soilingnet: Soiling detection on automotive surround-view cameras. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC) (pp. 67-72). IEEE.
- [2] A. Das, P. Krížek, G. Sistu, F. Bürger, S. Madasamy, M. Uříčář, V. R. Kumar, and S. Yogamani, "TiledSoilingNet: Tile-level Soiling Detection on Automotive Surround-View Cameras Using Coverage Metric," in Proc. IEEE Intelligent Transportation Systems Conference (ITSC), Rhodes, Greece, Sept. 2020, pp. 1-6, <https://doi.org/10.1109/ITSC45102.2020.9294677>
- [3] Soboleva, V., & Shipitko, O. (2021, December). Raindrops on windshield: Dataset and lightweight gradient-based detection algorithm. In 2021 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 1-7). IEEE.
- [4] S. Yogamani, M. Uricar, G. Sistu, et al., "WoodScape: A Multi-Task, Multi-Camera Fisheye Dataset for Autonomous Driving," Proc. ICCV Workshops, 2019.
- [5] Axmedov A. G. & Dadaxanov M. X. (2025). IMAGE QUALITY CHALLENGES IN AUTONOMOUS DRIVING: A CONCEPTUAL FRAMEWORK FOR CAMERA LENS DEFECT DETECTION. Development Of Science, 9(2), pp. 69-79.
- [6] Axmedov Abdulazizxon Ganijon O'g'li, & Dadaxanov Musoxon Xoshimxonovich. (2025). Enhancing Autonomous Driving Systems with AI-based Image Quality Assessment for Lens Defect Detection. Portugal-Scientific Review of the Problems and Prospects of Modern Science and Education, 1(3), 38-43.
- [7] J. Yang, Z. Zhang and others, "Deep Learning-Based Image Quality Assessment: A Survey," Procedia Computer Science, vol. 221, pp. 1000-1005, 2023, <https://doi.org/10.1016/j.procs.2023.08.080>

- [8] Shi, J., Gao, P., & Qin, J. (2024, March). Transformer-based no-reference image quality assessment via supervised contrastive learning. In Proceedings of the AAAI conference on artificial intelligence (Vol. 38, No. 5, pp. 4829-4837).
- [9] Q. Yang, et al., "An Unsupervised Method for Industrial Image Anomaly Detection," *Sensors*, vol. 24, no. 8, Article 2440, 2024.
- [10] Lan, G., Peng, Y., Hao, Q., & Xu, C. (2024). Sustechgan: image generation for object detection in adverse conditions of autonomous driving. *IEEE Transactions on Intelligent Vehicles*.
- [11] Mohammadi, P., Ebrahimi-Moghadam, A., & Shirani, S. (2014). Subjective and objective quality assessment of image: A survey. *arXiv preprint arXiv: 1406.7799*.
- [12] Ma, C., Shi, Z., Lu, Z., Xie, S., Chao, F., & Sui, Y. (2025). A survey on image quality assessment: Insights, analysis, and future outlook. *arXiv preprint arXiv: 2502.08540*.
- [13] Yao, Juncai; Shen, Jing; Yao, Congying. (2023). Image quality assessment based on the perceived structural similarity index of an image. *Mathematical Biosciences and Engineering*, 20(5), 9385-9409. <https://doi.org/10.3934/mbe.2023412>
- [14] Ahmed, I. T., Der, C. S., & Tareq, B. (2017). A Survey of Recent Approaches on No-Reference Image Quality Assessment with Multiscale Geometric Analysis Transforms.
- [15] Kim, H., Yang, Y., Kim, Y., Jang, D.-W., Choi, D., Park, K., Chung, S., & Kim, D. (2025). Effect of Droplet Contamination on Camera Lens Surfaces: Degradation of Image Quality and Object Detection Performance. *Applied Sciences*, 15(5), 2690. <https://doi.org/10.3390/app15052690>

Biography



Abdulazizxon Axmedov Ganijon O'g'li received his B. Sc. (2016-2021) and M. Sc. (2021-2023) degrees in Computer Science from Dongseo University, Busan, South Korea. He is currently pursuing his Ph. D. in Computer Science at Namangan State University, Uzbekistan, where he also serves as a researcher. His main research interests include computer science, artificial intelligence, and automation. He is actively engaged in developing AI-based approaches to enhance system performance and reliability.

Research Field

Abdulazizxon Axmedov Ganijon O'g'li: Analysis and enhancement of logistics infrastructure using satellite imagery, machine learning, and image quality assessment (IQA) methods.