Camera Health-Aware Cooperative Perception via Controlled Degradation and Uncertainty-Based Reliability

Shiva Aher
Georgia Institute of Technology
saher8@gatech.edu

Abstract

Cooperative autonomous driving systems rely on multi-agent perception and V2X communication, but degraded camera inputs can silently reduce fused perception quality. Existing cooperative perception pipelines often treat shared visual observations as equally reliable, even when individual cameras are affected by fog, motion blur, glare, sensor noise, or other degradations.

This paper presents a camera health-aware cooperative perception framework that estimates sensor reliability before downstream perception and uses this signal during V2X fusion. The central component is the Global Sensor Health Index (GSHI), a scalar score in $[0, 1]$ *derived from learned uncertainty patterns. GSHI is embedded as lightweight metadata in V2X packets and used to assign reliability-aware fusion weights,* $w_i = H_i / \sum_j H_j$.

To support controlled training and evaluation, this work introduces ***DRIVE-C****, a driving corruption benchmark pairing clean clips with structured degradations and GSHI supervision targets. Experiments show that GSHI decreases with degradation severity and provides a useful pre-fusion reliability signal. In simulated multi-agent fusion, health-weighted aggregation reduces the influence of degraded agents compared with uniform fusion.*

1. Introduction

Camera-based perception is central to modern Advanced Driver Assistance Systems (ADAS) and autonomous driving pipelines. In real-world deployment, however, cameras are exposed to degradations such as motion blur, glare, low light, adverse weather, and sensor noise. These effects can reduce perception reliability and cause silent failures before downstream task performance visibly collapses. Proactively detecting such degradation remains an important open problem.

Large-scale driving datasets such as KITTI, nuScenes, and Waymo support perception research under natural driving conditions [1, 2, 6], but do not provide systematic control over degradation type, severity, or reliability labels. Controlled corruption benchmarks [3] partially address robustness evaluation, but are not designed for autonomous driving or cooperative perception, and lack paired supervision for learning camera health.

Cooperative perception and vehicle-to-everything (V2X) communication improve spatial coverage by allowing multiple agents to share perceptual information [4, 8, 9]. However, many fusion pipelines treat shared observations as implicitly reliable. In practice, one agent may observe a scene through fog, glare, blur, or sensor noise, while another observes the same region under clearer conditions. Without explicit reliability reasoning, degraded observations can reduce the quality of the collective perception output.

This motivates *pre-fusion* reasoning about camera reliability. Prior work on distribution shift [5, 7] improves generalization across domains, but does not directly quantify or communicate per-agent input quality at the V2X layer. Existing cooperative perception frameworks similarly improve information sharing, but typically do not treat sensor health as a first-class fusion signal.

This paper presents a unified *camera health-aware cooperative perception* framework built around three contributions:

- **Global Sensor Health Index (GSHI).** A scalar reliability estimate $H \in [0, 1]$ derived from learned uncertainty patterns. GSHI is computed *prior* to downstream perception and embedded as lightweight metadata in each agent's V2X packet.
- **DRIVE-C dataset.** A controlled driving corruption benchmark pairing clean driving clips with structured degradations, including fog, snow, and motion blur across multiple severity levels. Each corrupted sample is paired with a GSHI supervision target for training and evaluating camera health estimation.
- **Reliability-aware V2X fusion.** A plug-in fusion scheme that weights each agent's contribution according to its estimated sensor health, $w_i = H_i / \sum_j H_j$, re-

ducing the influence of degraded inputs on the fused perception output without requiring architectural changes to existing cooperative perception pipelines.

Experiments on DRIVE-C show that GSHI decreases with degradation severity and provides a useful pre-fusion reliability signal. In simulated heterogeneous multi-agent settings, GSHI-weighted fusion reduces the influence of degraded agents compared with uniform fusion. The current evaluation is intentionally controlled, providing an initial validation of camera health as lightweight V2X reliability metadata.

2. Framework Overview

This section introduces the proposed camera health-aware cooperative perception framework. The framework combines controlled degradation, pre-fusion sensor reliability estimation, and reliability-aware multi-agent fusion. The goal is to estimate the quality of each agent's camera input before downstream perception results are fused through V2X communication. The full pipeline is illustrated in Fig. 1.

Given an input image x, a corruption operator produces a degraded observation

$$x' = \mathcal{C}(x;\, t, s), \tag{1}$$

where t denotes the degradation type and s denotes its severity level. A camera health model estimates a scalar Global Sensor Health Index (GSHI),

$$H = f_{\text{health}}(x'), \qquad H \in [0, 1], \tag{2}$$

where larger values indicate more reliable camera inputs and smaller values indicate stronger degradation.

In parallel, a perception model processes the same degraded input and produces detections or intermediate perception features,

$$D = f_{\text{perception}}(x'). \tag{3}$$

Each agent then broadcasts the pair

$$\mathcal{O}_i = \{D_i, H_i\} \tag{4}$$

through V2X communication, where D_i is the perception output from agent i and H_i is its estimated camera health. The health score is treated as lightweight metadata rather than a replacement for perception output. In multi-agent settings, these health scores provide a pre-fusion reliability signal that allows the fusion module to reduce the influence of degraded agents under heterogeneous sensing conditions.

3. DRIVE-C Dataset

DRIVE-C is a controlled driving corruption dataset designed to study camera reliability under structured visual degradations. Starting from clean driving clips, a corruption pipeline applies common camera failure modes, including fog, snow, rain, motion blur, sensor noise, and exposure changes. Each corruption is applied at multiple severity levels, enabling systematic analysis of how degradation affects camera health and downstream perception.

Each corrupted sample is paired with its degradation type t, severity level s, and a scalar Global Sensor Health Index label $H = \text{GSHI}(t, s) \in [0, 1]$, which serves as the reliability supervision target. Higher values indicate cleaner and more reliable inputs, while lower values indicate stronger degradation. In cooperative settings, DRIVE-C provides a controlled testbed for assigning different agents different degradation levels and evaluating whether health metadata can reduce the influence of unreliable inputs during V2X fusion.

4. Sensor Health Estimation

Camera degradations can affect visual quality before downstream perception models fail, yet most cooperative perception systems do not explicitly estimate input reliability before fusion. To address this, the proposed framework uses a camera health model f_{health} that maps a degraded image x' to a scalar Global Sensor Health Index (GSHI) score,

$$H = f_{\text{health}}(x'), \qquad H \in [0, 1], \tag{5}$$

where higher values indicate cleaner and more reliable camera inputs, while lower values indicate stronger degradation.

Model. The health model uses a lightweight convolutional encoder to extract degradation-sensitive visual features from x', followed by a regression head that predicts the scalar health score H. The model is trained on DRIVE-C using mean squared error against the corresponding GSHI supervision targets:

$$\mathcal{L}_{\text{health}} = \mathbb{E}\big[(H - \hat{H})^2\big], \tag{6}$$

where $\hat{H}$ denotes the ground-truth GSHI label derived from the controlled degradation type and severity.

Unlike confidence scores derived from downstream detection outputs, GSHI is estimated directly from the input image before the perception model is used. This allows camera reliability to be communicated as lightweight metadata in the V2X packet and used by the fusion module before degraded observations have a strong influence on the collective perception output. In this work, the health model is evaluated under controlled DRIVE-C corruptions, providing an initial validation of GSHI as a pre-fusion reliability signal.

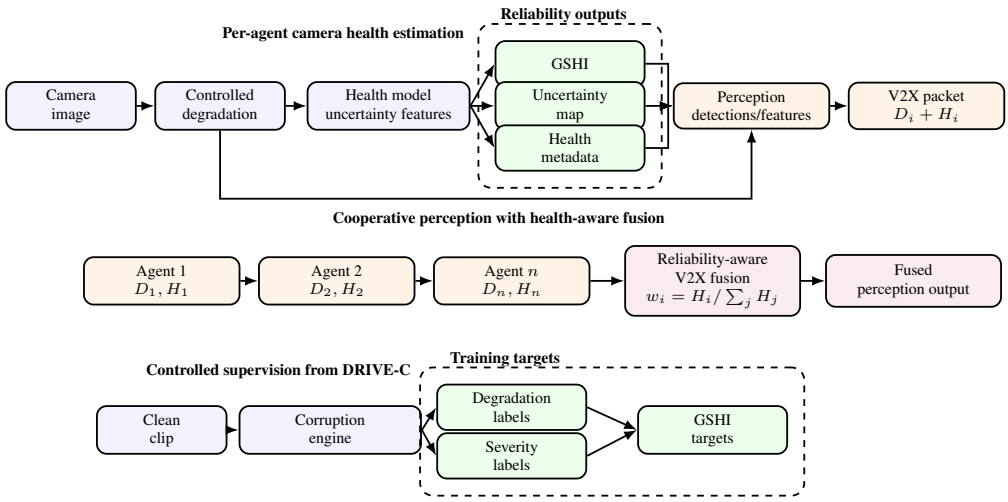

Figure 1. Camera health-aware cooperative perception framework. Each agent estimates camera reliability, shares perception outputs with GSHI metadata, and enables reliability-aware V2X fusion. DRIVE-C provides controlled degradation, severity, and GSHI supervision targets.

5. Camera Health-Aware Cooperative Perception

Cooperative perception pipelines improve spatial coverage by combining observations from multiple agents, but many do not explicitly model camera reliability before fusion. As a result, degraded observations caused by fog, blur, glare, or sensor noise may be fused with healthier observations and reduce collective perception quality.

The proposed framework embeds the Global Sensor Health Index (GSHI) as lightweight *camera health metadata* in each agent's V2X packet, providing a pre-fusion reliability signal for the fusion module.

Reliability-aware fusion. Each agent i broadcasts $\mathcal{O}_i = \{D_i, H_i\}$, where D_i denotes its perception output and $H_i \in [0, 1]$ denotes its estimated camera health score. Fusion weights are assigned in proportion to the health scores:

$$w_i = \frac{H_i}{\sum_j H_j}, \qquad D_{\text{fused}} = \sum_i w_i \cdot D_i. \tag{7}$$

Under this formulation, agents with higher estimated camera health contribute more strongly to the fused output, while agents with lower health scores are down-weighted. Agents with H_i below a threshold τ can also be excluded from fusion, reducing the influence of severely degraded observations. Unlike detection confidence scores, which are derived after running the downstream perception model, H_i is estimated directly from the camera input before fusion. This enables proactive reliability-aware aggregation rather than relying only on post-detection confidence. The formulation is lightweight and can be integrated with object-level or related cooperative perception pipelines without changing the underlying perception backbone.

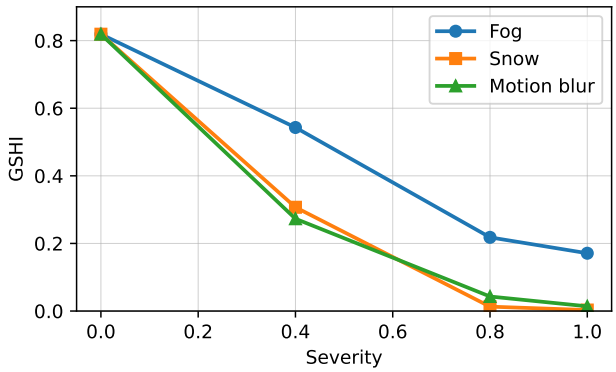

Figure 2. GSHI response under increasing corruption severity for fog, snow, and motion blur. Lower GSHI values indicate reduced estimated camera reliability.

6. Experiments

Setup. Experiments are conducted on DRIVE-C across fog, snow, and motion blur at five severity levels ($s \in \{0.2, 0.4, 0.6, 0.8, 1.0\}$). YOLOv5s is used as the detection backbone to evaluate downstream perception behavior under the same degradations. The health model uses a lightweight ResNet-18 encoder with a single regression head and is trained for 20 epochs using mean squared error loss with Adam optimiser (lr$= 1 \times 10^{-3}$). Cooperative fusion is evaluated with n=3 simulated agents under heterogeneous degradation, where each agent is independently assigned a degradation severity. Uniform averaging is com-

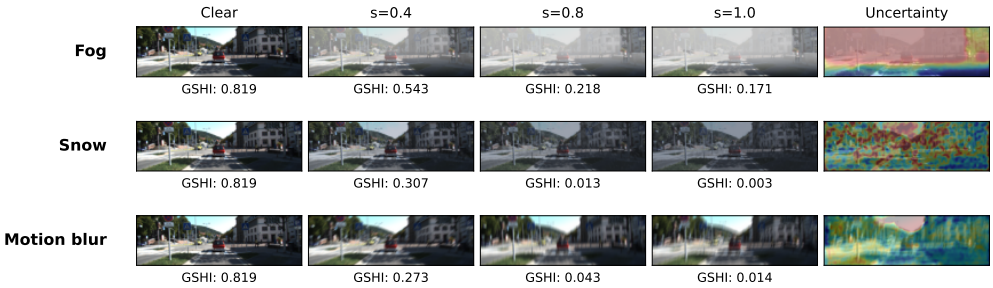

Figure 3. Qualitative degradation examples and corresponding GSHI estimates. Higher severity generally lowers GSHI and increases uncertainty response.

pared against GSHI-weighted fusion using Eq. 7, with an exclusion threshold of τ=0.1.

GSHI as a reliability signal. Fig. 2 shows that GSHI decreases as corruption severity increases for fog, snow, and motion blur. This trend indicates that the health model is sensitive to visual degradation and can provide a pre-fusion estimate of camera reliability. The health score also begins to decline at mild and moderate severity levels, suggesting its usefulness as an early reliability indicator.

The qualitative examples in Fig. 3 show the same trend: stronger corruptions produce lower GSHI values and more pronounced uncertainty responses. The degradation profiles differ across corruption types, with snow showing the sharpest reduction, motion blur declining more gradually, and fog following an intermediate trend.

Reliability-aware fusion. In the cooperative setting, GSHI-weighted fusion reduces the contribution of degraded agents by assigning lower weights to lower health scores. Agents below τ=0.1 are excluded from aggregation. Compared with uniform averaging, this provides a simple mechanism for limiting unreliable observations while preserving contributions from healthier agents.

Under clean or lightly degraded conditions, GSHI values remain higher and the weights approach uniform fusion. Under stronger heterogeneous degradation, the weighting mechanism shifts the fused output toward agents with more reliable camera inputs.

Limitations. The current evaluation is intentionally controlled and limited to synthetic degradations, one detection backbone, and simulated three-agent fusion. Future work will extend the evaluation to additional perception models, stronger fusion baselines, feature-level fusion, correlated sensor failures, and real-world multi-agent V2X datasets.

7. Conclusion

This paper presented a camera health-aware cooperative perception framework that estimates sensor reliability using GSHI and communicates it as lightweight V2X metadata. Experiments on DRIVE-C show that GSHI decreases with increasing degradation severity across fog, snow, and motion blur, supporting its use as a pre-fusion reliability signal. In simulated heterogeneous multi-agent settings, GSHI-weighted fusion reduces the influence of degraded agents compared with uniform fusion. Future work will extend the approach to real-world V2X datasets, additional backbones, stronger fusion baselines, multi-modal sensing, and adaptive exclusion thresholds.

References

[1] Holger Caesar et al. nuscenes: A multimodal dataset for autonomous driving. In *CVPR*, pages 11621–11631, 2020.

[2] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *CVPR*, pages 3354–3361, 2012.

[3] Dan Hendrycks and Thomas Dietterich. Benchmarking neural network robustness to common corruptions and perturbations. In *International Conference on Learning Representations (ICLR)*, 2019.

[4] Yue Hu et al. Where2comm: Communication-efficient collaborative perception via spatial confidence maps. In *NeurIPS*, pages 4874–4886, 2022.

[5] Da Li et al. Deeper, broader and artier domain generalization. In *ICCV*, pages 5542–5550, 2017.

[6] Pei Sun et al. Scalability in perception for autonomous driving: Waymo open dataset. In *CVPR*, pages 2446–2454, 2020.

[7] Antonio Torralba and Alexei A. Efros. Unbiased look at dataset bias. In *CVPR*, pages 1521–1528, 2011.

[8] Runsheng Xu et al. Opv2v: An open benchmark dataset and fusion pipeline for vehicle-to-vehicle perception. In *ICRA*, pages 2583–2589, 2022.

[9] Runsheng Xu et al. V2x-vit: Vehicle-to-everything cooperative perception with vision transformer. In *ECCV*, pages 107–124, 2022.