Work-in-Progress: Context and Noise Aware Resilience for Autonomous Driving Applications

Hamidreza Alikhani[1], Anil Kanduri[2], Pasi Liljeberg[2], Amir M. Rahmani[1], Nikil Dutt[1]
[1] Dept. of CS, University of California Irvine, USA, [2] Dept. of Computing, University of Turku, Finland

***Abstract*—Autonomous Vehicles (AVs) often use noise prone sensory data from cameras and LiDAR for perception. In specific noisy scenarios, different object detection models exhibit non-intuitive and varying degrees of resilience, necessitating adaptive model selection. In this work, we develop a context and noise aware framework for run-time adaptive configuration of objection models for high accuracy and low latency inference. We combine driving scene context and input data noise to prioritize among input modalities, followed by selection and configuration of most resilient object detection model appropriate for the context. Our evaluation for 2D object detection on nuScenes dataset provided average 1.83x speedup in latency compared to baseline while preserving average prediction confidence.**

I. Resilience of Objection Detection Models

Multi-modal sensory data (e.g., camera, LiDAR, RADAR) used in AVs are often perturbed with noisy components and environmental factors such as weather condition [1]. Object detection models used in AVs should provide inference at lower latency and higher accuracy while being resilient to input data perturbations. Existing approaches use deeper and complex models at higher noise levels to improve prediction accuracy, resulting in higher inference latency [2]. However, object detection models exhibit non-intuitive and diverse levels of resilience towards specific types of noises and environmental conditions. For example, a quantized model could provide a higher prediction accuracy and lower latency under a specific type of noise [3]. In this work, we explore diversity in resilience to adaptively configure object detection models at run-time for handling input data perturbations. We analyze input data quality and environmental context to (i) determine resilience of different object detection models under the given conditions, (ii) select the most resilient model, (iii) configure the selected model by prioritizing among input modalities.

Figure 1 illustrates resilience of various object detection models across multiple driving scenarios (driving maneuvers and traffic situations) using nuScenes dataset [1], with and without noise. The size of each circle in the plots shows the relative inference time for each model. Although the most complex model (faster-rcnn-FPN) outperforms other models with ideal input (Figure 1 a), relative performance of models changes with noisy input. Specifically, there are scenarios where the bigger models are not the best performing ones. For instance, the faster-rcnn-FPN-320 model (quantized faster-rcnn-FPN with 0.5x latency) outperforms the latter in all settings with noisy input (Figure 1 b, c, and d). Moreover, other models with faster inference like YOLOv8-l have same or better performance than faster-rcnn-FPN in Scenario 5 (night scene) in Figure 1 b, c with the advantage of 70% faster inference time per frame. This showcases non-intuitive resilience of models against different noise types in various driving scenarios, which can be even more unpredictable when utilizing multi-modal models with more modalities like

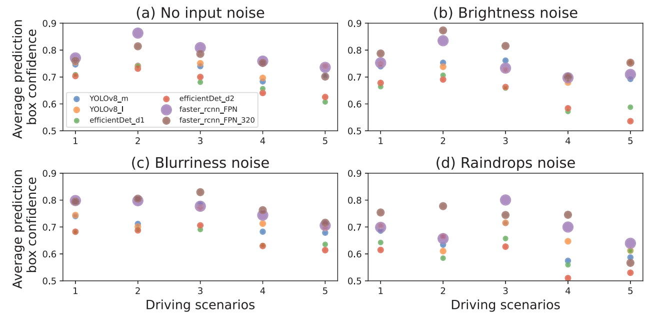

Fig. 1. Performance of object detection models in different driving scenarios with and without noise.

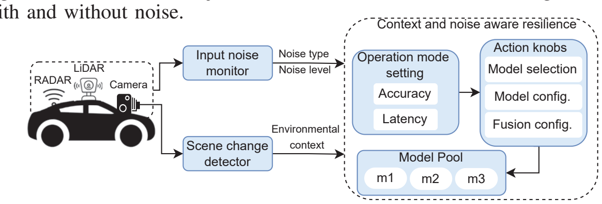

Fig. 2. Context and Noise Aware (CNA) Object Detection Framework

intelligent framework aware of such fluctuations in models' performance based on context and data quality, and choosing optimal action knobs accordingly for achieving robustness while preserving accuracy in end-to-end AV pipelines.

II. Preliminary results and future work

We design a Context and Noise Aware (CNA) Object Detection framework as shown in Figure 2. CNA is equipped with a scene change detector based on camera input, and input sensor noise monitoring responsible for detecting events for re-evaluating models performance. Based on a defined operation mode (set based on accuracy and latency demands), an appropriate action knob is activated by the CNA model. This action can vary from intelligent model selection from a pool of available models, model configuration of unified multi-modal models, or fusion configuration in multi-modal settings [4]. Our evaluation with an ad-hoc method which selects best performing model in certain noise type and driving scene (configuration with highest accuracy) achieves up to 1.83x speed-up in inference time while preserving prediction confidence compared to the baseline method (selecting best performing model with ideal input). However, this ad-hoc method relies on offline information about models' performance in different context and noise settings. We aim to automate this process by more intelligent and real-time input data monitoring and more effective action knobs.

References

[1] H Caesar et al. nuscenes: A multimodal dataset for autonomous driving. In *IEEE/CVF CVPR*, pages 11621–11631, 2020.
[2] Y Zhang et al. Carma: Context-aware runtime reconfiguration for energy-efficient sensor fusion. In *ISLPED 2023*, pages 1–6. IEEE, 2023.
[3] T Liang et al. Pruning and quantization for deep neural network acceleration: A survey. *Neurocomputing*, 461:370–403, 2021.
[4] T Baltrusaitis et al. Multimodal machine learning: A survey and axonomy. 41(2):423–443, feb 2019.

2832-6474/24/$31.00 ©2024 IEEE
DOI 10.1109/CODES-ISSS60120.2024.00010