Impact of Data Compression on AI Perception Tasks for Teleoperated Autonomous Vehicles

Nitin Varyani
varya001@umn.edu
University of Minnesota

Jason Carpenter
carpe415@umn.edu
University of Minnesota

Steven Sleder
slede001@umn.edu
University of Minnesota

Wei Ye
ye000094@umn.edu
University of Minnesota

Rostand A. K. Fezeu
fezeu001@umn.edu
University of Minnesota

Eman Ramadan
eman@cs.umn.edu
University of Minnesota

Zhi-Li Zhang
zhzhang@cs.umn.edu
University of Minnesota

ABSTRACT

Among 5G promises are higher throughput and lower latency, which can help new applications such as teleoperating autonomous vehicles. However, there is a huge disparity between the uplink (required to send data to the remote driving center) and downlink throughput. Hence, sending the raw data is not feasible. In this study, we analyze the tradeoff of using multiple compression algorithms and techniques to minimize the size of the data sent and enhance latency, and the degraded video quality which impacts object detection and can lead to wrong decisions made by the remote human operator.

KEYWORDS

5G, Low Latency, Autonomous Vehicles, Teleoperation, AI

1 INTRODUCTION

Teleoperating Autonomous Vehicles (AVs) plays a pivotal role in the transition towards fully Autonomous Vehicles. The potential of 5G technology for AV teleoperation is promising [1–3]. Compression techniques applied to video and LiDAR data not only reduce transmission delays but also degrade data quality. This compromise directly impacts AI perception tasks essential for effective AV teleoperation. Achieving a balance between delay and quality trade-offs is vital for the development of robust teleoperation systems.

The absence of adequate 5G throughput for uplink transmission necessitates compression of video and LiDAR data before streaming. However, such compression can significantly impact AI perception tasks crucial for AV teleoperation. The study investigates the impacts of compression on downstream AI tasks, such as object detection and recognition.

2 BACKGROUND

The original dataset consists of MJPEG video for the camera feed and full LiDAR point data. To address limitations in 5G throughput for uplink transmission, video compression techniques such as H.264 and H.265 are employed. Additionally, LiDAR downsampling is implemented using voxel sizes ranging from $0.1m^3$ to $1.0m^3$. These compression and downsampling methods aim to reduce data size for efficient transmission.

Our observations indicate that H.264 and H.265 outperform MJPEG by effectively reducing per-frame network delay, particularly noticeable in tail latency. This improvement suggests that these compression techniques facilitate smoother and more responsive data transmission.

However, the study seeks to evaluate the impact of these techniques on downstream AI tasks, particularly object detection and recognition.

3 EXPERIMENT SETUP & DATASET

In our experiment setup, the primary task involves predicting bounding boxes from mono-modal and multi-modal inputs, incorporating data from both cameras and LiDAR sensors, with subsequent computation of mean Average Precision (mAP) for all classes. Two datasets are utilized: MnCAV, a comprehensive 1085-mile driving campaign dataset offering diverse real-world driving scenarios, and NuScenes-Mini, a subset of NuScenes with ground truth labels. FocalFormer3D is chosen for NuScenes due to its versatility in supporting both mono- and multi-modal prediction, while YOLO is employed in MnCAV for real-time object detection, generating bounding boxes for various classes including vehicles and pedestrians.

4 RESULTS

MnCAV dataset results In our analysis, we found that H.265 compression resulted in fewer detections due to reduced data quality. Specifically, one car was missed in the H.265 compressed footage compared to the original MJPEG format. Additionally, while H.265 significantly decreased bitrate (e.g., from 87.97 Mbps to 4.946 Mbps) and latency, these improvements came at the cost of compromised AI detection performance. This trade-off highlights the complex interplay between compression efficiency and the accuracy of AI-based detection algorithms in teleoperation systems for autonomous vehicles.

Nuscenses Dataset Results: We present findings from our investigation into object detection performance in both mono-modality and multi-modality settings. For mono-modality object detection, we observed that both H.264 and H.265 compression techniques led to decreased object detections compared to MJPEG, attributed to lower video quality. Notably, H.265 exhibited a more pronounced negative effect on video-based detection performance than H.264. However, LiDAR-based detection remained resilient until a voxel size of $(0.5m)^3$ was reached. Moving to multi-modality object detection, we found that combining data from both cameras and LiDAR sensors significantly outperformed camera-only and LiDAR-only detection methods, demonstrating improvements of up to 171.0% and 39.5%, respectively.

5 CONCLUSION AND FUTURE WORK

In conclusion, our study underscores the complex relationship between data compression and AI perception tasks in teleoperated AVs. While compression techniques alleviate transmission delays, they can degrade data quality, adversely affecting object detection and recognition. The observed trade-offs emphasize the importance of carefully balancing compression efficiency and AI detection accuracy for robust teleoperation systems.

In future work, we aim to delve deeper into the effects of compression on various aspects of teleoperation systems. Specifically, we plan to investigate its impact on trajectory prediction, semantic segmentation, and entity tracking within AI frameworks. Additionally, we seek to explore the distribution of AI tasks among on-board, edge, and remote systems, optimizing the allocation for efficiency. Another focus will be on deploying streamlined AI model architectures tailored for real-time AV teleoperations. Furthermore, our research extends to evaluating AI models and perception tasks in real-time robot AV teleoperation scenarios.

REFERENCES

[1] 2017. Remote Operation of Vehicles With 5G (extract from the Ericsson mobility report) (June). https://www.ericsson.com/en/reports-and-papers/mobility-report/articles/remote-monitoring-and-control-of-vehicles. Ericsson, Stockholm, Sweden. Last Accessed May 2024.

[2] 5GAA. 2021. Teleoperated driving (ToD): System requirements analysis and architec- ture). https://5gaa.org/content/uploads/2021/09/5GAA_ToD_System_Requirements_Architecture_TR.pdf Last accessed May 2024.

[3] Lei Kang, Wei Zhao, Bozhao Qi, and Suman Banerjee. 2018. Augmenting Self-Driving with Remote Control: Challenges and Directions *(HotMobile '18)*. Association for Computing Machinery, New York, NY, USA, 19–24. https://doi.org/10.1145/3177102.3177104