ConnectGPT: Connect Large Language Models with Connected Automated Vehicles

Kailin Tong[1] and Selim Solmaz[1]

Abstract— **This paper explores the intersection of recent AI advancements and Intelligent Transportation Systems (ITS), specifically focusing on enhancing the capabilities of Connected and Automated Vehicles (CAVs) in dynamic traffic scenarios. While combinations of vehicular sensors and AI offer promising prospects for advanced environmental perception, challenges still persist in accurately identifying dangers during the transition to automated traffic. The ESRIUM project, funded by the EU Horizon 2020 Programme, aims to address these challenges by developing digital maps representing road deterioration and employing Vehicle-to-Everything (V2X) communication to generate infrastructure-assisted routing recommendations for CAVs. While the solutions for sending standardized safety messages and controlling enabled CAVs were demonstrated in the ESRIUM project, the solution for the automatic generation of V2X safety messages was not studied. In this paper, we propose a pipeline named "ConnectGPT", which connects Large Language Models (LLMs) with CAVs, utilizing GPT-4, to observe traffic conditions, identify conditions that can endanger the flow of traffic, and automate the generation of the corresponding standardized V2X messages, such as Decentralised Environmental Notification Message (DENM) about the actual safety problem. Practical experiments with ongoing development show potential for real-world applications, which can significantly improve traffic management efficiency and enhance the security of all traffic participants, marking a crucial advancement in the integration of AI tools in ITS.**

Fig. 1. The world's first GPT that connects Large Language Models (LLMs) with Connected Automated Vehicles (CAVs). Our proposed pipeline can be applied in real-life scenario and has the potential to revolutionize the Intelligent Transportation System (ITS). ConnectGPT is customized upon GPT-4 [7]. Logo source: DALL·E 3

I. Introduction

The recent progress in AI methodologies has opened up new opportunities for Intelligent Transportation Systems (ITS). Vehicular sensors are continuously improving, facilitating the development of Advanced Driver Assistance Systems (ADAS) or Automated Driving (AD) [1]. The recent rapid development of Large Language Models (LLMs) and Vision Foundation Models (VFMs) are new AI tools that can also assist Connected and Automated Vehicles (CAVs) in driving themselves by interpreting the environment and making choices like humans [2].

However, there are still challenges in successfully managing and directing the actions of CAVs in dynamic traffic situations, especially during the shift from conventional to automated traffic. An obstacle in the process of obtaining higher degrees of CAVs is the requirement to accurately identify potential dangers, plan optimal routes, and make informed choices in diverse situations [3], such as encountering lane closures or road damage. An Automated Driving System (ADS) lacking connection with infrastructure or other cars may face challenges in avoiding maintenance zones on the road due to the limitations of its deployed sensors in detecting small cones. The vulnerability of perception sensors to deterioration in adverse weather conditions is apparent. The inclement weather conditions also have a negative impact on the Global Navigation Satellite System (GNSS) [4]. Hence, V2X communication is crucial in attaining advanced levels of autonomy (SAE Level-4 and Level-5 [5]) since it allows CAVs to get real-time routing and driving suggestions.

The EU-H2020-funded project ESRIUM [6] draws inspiration from this, with the objective of enhancing the safety and resource efficiency of transportation on European roads. This will be achievable by developing a precise digital map that precisely represents the deterioration and destruction of road surfaces. The map will be employed to mitigate road construction and associated problems by efficiently overseeing traffic and regulating route utilization. Furthermore, it will furnish road operators with useful data for strategizing maintenance operations. In addition, connected vehicles will be provided with route and driving suggestions to reduce road surface deterioration and the frequency of necessary maintenance measures.

In the ESRIUM project, we have effectively executed practical experiments involving Vehicle-to-Everything (V2X) and a CAV on the Austrian road A2. Our CAV demonstration was thoroughly validated on a functional highway with heavy

[1] K. Tong and S. Solmaz are with Virtual Vehicle Research GmbH, Inffeldgasse 21a, 8010 Graz, Austria. {kailin.tong, selim.solmaz}@v2c2.at

traffic and poor weather conditions, testing all operational autonomous driving capabilities as well as the On-Board Unit (OBU) and Roadside Unit (RSU). The V2X messages were generated by the road operator, transmitted to our CAV demonstrator via the RSU, and received by our OBU, resulting in the appropriate maneuvering of an SAE Level-3 CAV in complex traffic scenarios.

Although the demonstration was successful, we have identified opportunities to further optimize the efficiency of our framework, specifically in regards to managing traffic hazards and maneuvering CAVs. The road operators are consistently monitoring the traffic conditions and the state of the route, including road damage, wear, and traffic accidents. While there is some automation involved, human traffic experts are primarily responsible for understanding and analyzing traffic scenes. A highway spanning hundreds of kilometers presents challenges in terms of prompt traffic management and necessitates a significant amount of manpower. Additionally, after a situation is recognized, the human traffic expert must formulate a request and communicate it to software programmers to generate standard V2X messages. RSU then disseminates these messages to inform other vehicles of the situation. However, this complex process presents difficulties in promptly responding to an unforeseen traffic incident, such as a stone falling onto the road.

To address this issue, we propose a novel pipeline that connects LLMs with CAVs, and develop a customized model called ConnectGPT, which is based on the latest GPT-4 1. ConnectGPT can observe the present traffic conditions, recognize road deterioration types, and identify dangerous circumstances. It can then automatically notify the road operator, who will confirm the information. In addition, it will automatically generate standard Cooperative Intelligent Transport Systems (C-ITS) messages for the road operators, depending on the actual traffic circumstances. ConneeGPT will significantly enhance the effectiveness of road operations and facilitate the timely identification of potential risks for traffic participants. CAVs can utilize the information produced by ConnectGPT to promptly implement precautionary measures or execute minimal-risk maneuvers, thereby augmenting the safety and efficiency of the ITS.

As far as we know, we are the first in the world to provide a pipeline that enables the transfer of data from infrastructure cameras to connected autonomous vehicles (CAVs) using LLM and V2X communication. Our pipeline is practical and readily applicable in real-world scenarios. Although the development of the pipeline incorporating ConnectGPT is still ongoing, we have conducted validation using a small-scale dataset and a specific form of C-ITS message. Additionally, there are interesting conclusions that arise, highlighting ways to optimize our proposed pipeline. Moreover our pipeline has significant potential to be extended and can increase the traffic management efficiency as well security of all traffic participants. .

The rest of this paper is structured as follows: First, Section II states the state-of-the-art on infrastructure-assisted CAVs and use of LLMs for connected mobility applications. Then in Section III the problem statement regarding establishment of an automated pipeline for generation of standardized routing recommendations is given, and then in Section IV our proposed solution approach is presented. The representative experiments together with the analysis of the initial results are accordingly presented in Section V, and finally the outcomes and future directions are stated Section VI.

II. Related Work

A. Development of CAVs and V2X

The emergence of CAVs and infrastructure-supported automated driving functions has garnered significant interest in recent years, as outlined in [4]. Schulte-Tigges et al. proposed a software architecture and logic for CAVs that effectively utilizes hazard notification and road signage information from V2X messages [8]. This approach allows for the management of Operational Design Domain (ODD) decisions and reactions in a predictable manner. Their proposed software architecture incorporates a maneuver planner that utilizes separate state machines to respond to various types of V2X information. The system generates target objectives for a motion planner and path controller. When compared to a basic autonomous vehicle (AV) model that only relies on sensors on board, the simulations show how the presented CAV solution is better. Furthermore, they conducted real-world test-track experiments to confirm the usefulness of the proposed logic. However, the test track is only straight and without the presence of real traffic scenarios. This compromised the strength of their validation.

Most recently, practical experiments of V2X and CAV were conducted on the Austrian route A2 as part of the ESRIUM project [6]. To our knowledge, there has not been a comprehensive validation undertaken on a functional highway that encompasses all operational autonomous driving capabilities, along with the On-Board Unit (OBU) and Roadside Unit (RSU). The CAV driving function employed an adaptive cruise control (ACC) module in conjunction with a PI controller to either sustain a specified velocity or a consistent time distance to the preceding vehicle. A rule-based trajectory planner was used to place Bézier-based reference paths in a Frenet frame relative to the center of the current lane to regulate lateral motion. A state feedback controller based on LQR was responsible for managing the reference path tracking. To obtain further information on the control methods, please refer to the following sources: [9]–[11].

Other recent research on CAVs and V2x has a different focus. It includes a decentralized protocol for CAVs' coordination [12], a vehicle-in-the-loop (ViL) test environment with V2X communication [13], ways to account for the unique role of infrastructure-assisted collective perception (ICP) [14], the accuracy and delay of V2X communication localization [15] and EGNSS-based path tracking [16]

B. Large Language Models for CAVs

Large Language Models (LLMs) show promise in activities requiring human-like reasoning, such as autonomous vehicles. Recent improvements have led to increased interest in Multimodal Large Language Models (MLLMs), which combine LLMs' sophisticated reasoning with picture, video, and audio data [17]. A recent in-depth study looks at the current state, problems, and future research that is needed for these LLM-based AI systems in the field of AD [2]. The researchers in the survey paper claim that MLLMs can enhance driving decision-making, navigation, safety, and efficiency by understanding traffic scenes, enhancing vehicle planning, and adapting to changing road conditions. In addition, LLMs can personalize driving experiences and in-vehicle entertainment, build trust in autonomous technology, and explain their actions to passengers. These models can learn from new data and adapt to driver preferences over time. However, their work primarily concentrates on the intelligence of individual vehicles while disregarding the broader scope of ITS and CAVs, which encompass infrastructural support for AD, V2X communication, and intelligent traffic management.

The researchers also emphasize the hardware constraints of implementing LLMs in a single automobile [2]. For autonomous vehicles to perform driving duties using Large Language Models (LLMs), these models must make rapid decisions with minimal latency in order to ensure safety. Consequently, the utilization of substantial computational resources in this manner has a direct impact on both the efficiency and energy consumption of the autonomous driving system. Conversely, LLMs designed for automobile navigation, maneuver instruction, and traffic management may exhibit slower processing speeds and infrequent reliance on requesting instructions, making them easier for computers to handle. Consequently, it is acceptable for these LLMs to operate from the infrastructure side. Moreover, modern V2X communication technologies provide accurate and effective long-range connections between infrastructure facilities and vehicles.

Other studies propose strategies for integrating AI with CAVs in the context of ITS. In 2019, a review paper examined multiple AI approaches, including swarm intelligence, natural language processing (NLP), machine learning, and deep learning, intending to improve the capabilities of V2X systems [18]. Nevertheless, LLMs were not widely recognized at that period, and their capacity to enhance V2X was not a subject of discussion. In 2022, Binisha et al. provided suggestions about the potential of using AI like NLP to solve crucial issues in V2X systems, such as ensuring efficient communication and safety measures for automobiles on the road [19].

In summary, while there are many discussions around the application of LLMs in autonomous driving, the actual implementation in the industry is few. Furthermore, the significance of utilizing LLMs to enhance V2X and ITS has not been extensively explored. This study is a pioneering effort to integrate LLMs with V2X communication technology and present a viable framework that can be immediately implemented in the existing intelligent traffic management system. This framework enhances traffic flow efficiency and promotes safer driving of CAVs.

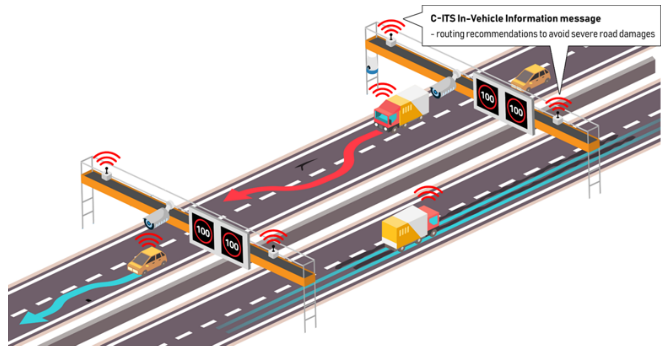


Fig. 2. Implemented scenarios: In-lane offset recommendation (cyan trajectory) and lane change recommendation (red trajectory).

III. Problem Statement

The primary approach and solution of the EU-H2020-funded project ESRIUM [6] involve developing and maintaining a precise digital map that represents the deterioration of road surfaces using sensor-equipped vehicles. This map is then utilized to regulate route utilization through simple in-lane positioning and lane-change maneuvers, as illustrated in Fig. 2. While ESRIUM successfully conducted practical public on-road experiments and validated infrastructure-assisted automated driving functions on the Austrian A2 motorway, the demonstrations also highlighted opportunities for framework optimization. However, challenges persist in managing traffic hazards and maneuvering CAVs, going beyond the road damage identification alone, necessitating improved automation and prompt response mechanisms for unforeseen incidents. Notably, the ESRIUM project leaves unanswered the question of how to generate a standardized V2X message for an imminent road hazard.

Building upon these initial observations, our current study is centered on investigating an automated pipeline tailored for generating standardized V2X messages. Our primary aim is to effectively alert oncoming CAVs (as well as connected vehicles in general) about impending dangers. In our methodology, we leverage infrastructure sensors, including cameras and potentially other perception sensors, in conjunction with a purpose-built LLM serving as a central decision support tool. These components are utilized to monitor specific road sections, collecting data on traffic conditions and potential obstacles that may disrupt traffic flow. Thus, the pivotal inquiry becomes: How can we establish an automated or semi-automated system that ensures the timely generation of standardized V2X safety messages, offering prompt warnings and guidance to oncoming traffic while ensuring a dependable safety guarantee?

IV. Our Approach

Fig.3 illustrates the proposed pipeline that links LLMs with CAVs in order to decrease the reaction time for event

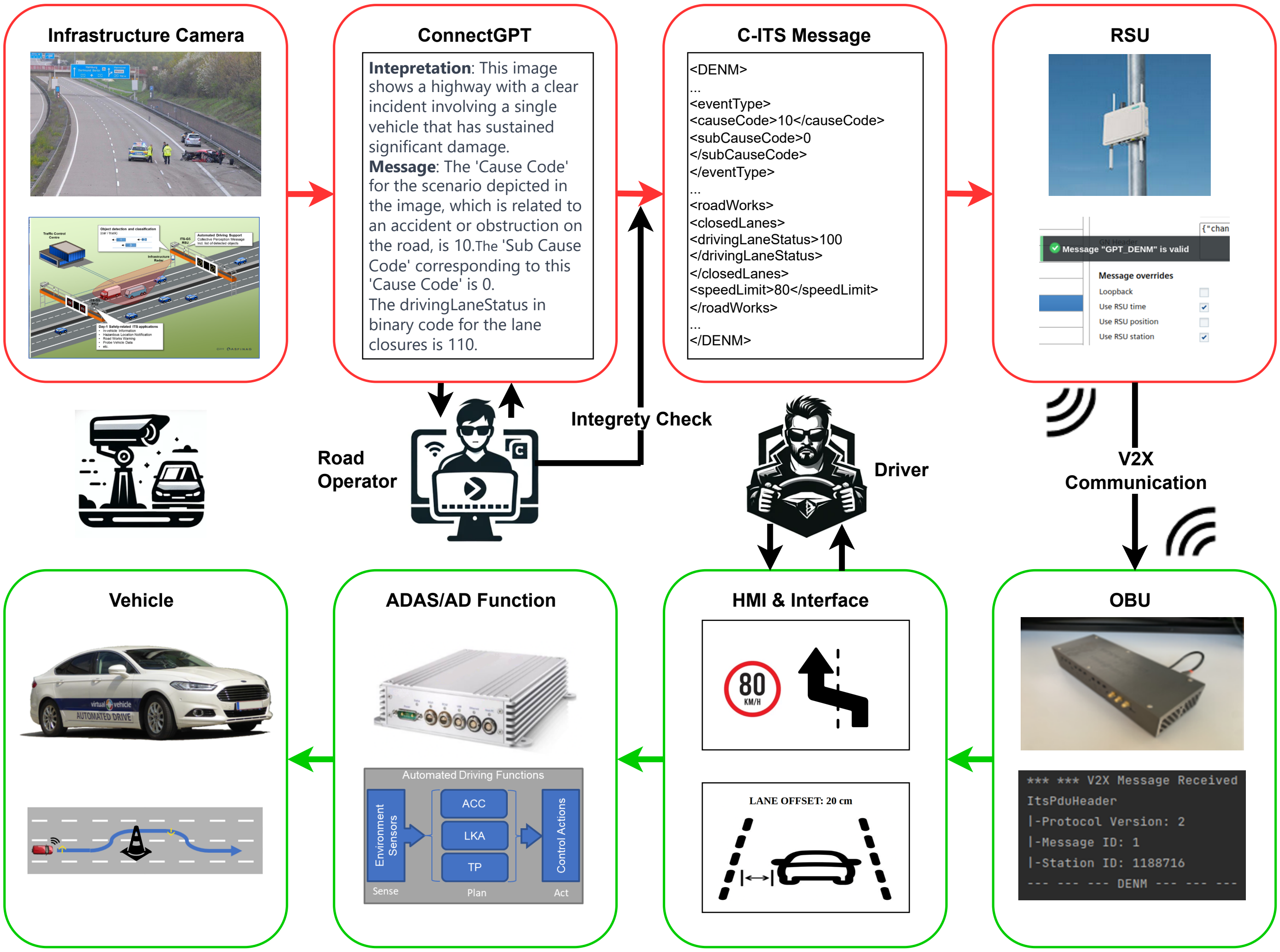


Fig. 3. Overview of our proposed pipeline that connects LLMs and CAVs. infrasture-related software and hardware are in red blocks , while the vehicle-related software and hardware are in green blocks.

management in traffic management and enhance driving safety and traffic flow efficiency. The following sections will provide detailed explanations for each block shown in the illustration.

A. Infrastructure Camera

This block signifies the initial phase, whereby cameras installed on infrastructure take images of road conditions. Images such as these are essential for identifying and recognizing events or hazards on the road. Furthermore, these infrastructure cameras have been implemented for the purpose of traffic management in Austria. The Austrian test site AlpLab, situated in the Graz region along the A2 highway between Graz West and Lassnitzhöhe, exhibits the significant implementation of cutting-edge infrastructure and sensory technology. This location spans a distance of 20 kilometers along the highway and is equipped with advanced and novel sensing equipment. A network including 26 cameras is established in both directions of the test location [20]. These cameras are crucial in detecting incidents, enhancing road safety, and facilitating effective traffic management. The video cameras record real-time images from various places and transmit them to ConnectGPT for additional processing.

B. Image Interpretation and C-ITS Message generation by ConnectGPT

ConnectGPT subsequently analyzes images obtained from the infrastructure cameras. We have devised several prompts to acquire the most accurate response. This analysis may entail examining the images in order to discern certain situations, such as accidents, road obstructions, or hazardous conditions. After identifying a particular scenario or potential situation, as depicted in Fig. 3, the connecGPT system interprets the images and produces C-ITS messages. In our work, these messages are specifically Decentralized Environmental Notification Messages (DENM) [21], but our proposed framework is also applicable to other types of C-ITS messages.

As seen in the ConnetGPT block of Fig. 3, according to the standard DENM (ETSI EN 302 637-3) [21], ConnectGPT analyzes the picture and subsequently outputs the cause code (10 in the given example) and sub cause code (0 in the given example). Considering the present circumstances, ConnectGPT proposes a lane closure on ”110”, indicating that the first two left lanes are open while the rightmost lane has been shut off. The generation of C-ITS messages requires an interaction between ConnectGPT and the road operator. Furthermore, it is required to do an integrity check by the road operator on the generated C-ITS message, as there is a possibility of LLMs making errors. The third block in 3 displays a code snippet of the whole DENM . The message provides precise information regarding the identified event or hazard, such as cause codes and driving lane status.

C. V2X communication between RSU and OBU

The RSU program can forward the C-ITS message after confirming that it is free of syntax errors. V2X communication encompasses the communication of information between the vehicle's OBU and RSU, guaranteeing the widespread distribution of the message to all pertinent road participants.

The On-Board Unit (OBU) installed in cars receives the DENM from RSU and analyzes them to deliver notifications or information to both the driver and the Automated Driving systems, therefore improving road safety. The vehicleCAPTAIN [15], an OBU along with an open-source library developed by Virtual Vehicle Research, receives C-ITS messages and performs message parsing in the first stage at the vehicle side. Subsequently, the essential information will be conveyed to the Human-Machine Interface (HMI) and the interface to the automated driving system.

D. Automated driving system

During the subsequent phase, human drivers get information on the state of the road through the Human-Machine Interface (HMI). The HMI efficiently notifies the driver of important information, such as speed limits, advice for changing lanes, or lane offset. Meanwhile, the extracted data from the C-ITS message is being analyzed. And The parsed information is then transmitted to a real-time ECU, which executes the ADAS/AD function, as seen in the HMI and Interface block in 3.

The AD/ADAS function regulates the vehicle's longitudinal and lateral movements. The vehicle's lateral position is controlled by the lane markings, detected by a commercially available Mobileye camera system. To regulate lateral motion, a rule-based trajectory planner generates Bézier-based reference routes in a Frenet frame relative to the current lane's midline. Subsequently, the task of navigating along this reference path was controlled by a LQR-based state feedback controller. The longitudinal control system uses an adaptive cruise control (ACC) module and a PI controller to keep the vehicle at a certain speed or at the same distance from the vehicle ahead of it. To obtain further information on the control methods, please refer to the following sources: [9]–[11].

The test vehicle we are using is designed to execute lane change or lane-keeping maneuvers, as well as speed adaptation. The test vehicle is a Ford Mondeo Hybrid that has been equipped with supplementary hardware for environmental perception, as seen in the vehicle block in Fig. 3.

The original pipeline demonstrated on a real highway in Austria [22] did not utilize LLMs, so camera images had to be tediously monitored by traffic engineers, and C-ITS messages were generated by hand using a professional software. Our proposed pipeline incorporating ConnectGPT, on the other hand, lets road operators quickly analyze a large amount of real-time data from infrastructure cameras and respond to traffic incidents or dangers. The process of generating C-ITS messages is also simplified.

In summary, the proposed pipeline presents a holistic strategy for enhancing road safety and optimizing traffic flow through the utilization of cutting-edge GPT-4 technology for instantaneous data analysis, efficient communication, and seamless interaction with both manually operated vehicles and connected automated vehicles. As the implementation of this type of intelligent transportation system becomes more prevalent in our everyday lives, the advantages of reducing labor are becoming more evident.

V. Experiments and Discussion

A. Customizing ConnectGPT

Starting on November 6, 2023, OpenAI enables users to generate customized versions of ChatGPT [7]. It is important to mention that customization often occurs at the application level rather than the model level. We formulated the prompts for GPT-4 and customized ConnectGPT by equipping it with knowledge from CAV and V2X literature, as well as C-ITS message standards. In order to enhance the likelihood of providing the desired response, we employed two sample photos to ”train” ConnectGPT. ConnectGPT is therefore customized to specialize in monitoring present traffic conditions, recognizing road deterioration, identifying incidents and dangerous circumstances, and promptly notifying road operators for verification purposes. This specialization involves several key capabilities: Traffic Condition Monitoring: ConnectGPT has the capability to continuously monitor current traffic conditions. This can include tracking traffic flow, detecting traffic jams, and assessing overall traffic patterns.

- Traffic Condition Monitoring: ConnectGPT has the capability to continuously monitor current traffic conditions. This can include tracking traffic flow, detecting traffic jams, and assessing overall traffic patterns.
- Incident and Hazard Detection: ConnectGPT can detect incidents such as accidents or dangerous conditions on the road. This could involve real-time analysis of traffic data, inputs from sensors or cameras, or integration with emergency reporting systems.
- Standard C-ITS Message Creation: The system is capable of creating standardized C-ITS messages.

Currently the generation of C-ITS message is not perfect, as discussed in the next section. But the successfully generated messages conform to specific formats and protocols to ensure compatibility for RSU, which can be directly used for V2X communication. Furthermore, our present study exclusively examines the DENM and ConnectGPT is used to generate standard DENM . But in the future work ConnectGPT will be used in a wider range of traffic management scenarios and support different types of C-ITS messages. This has the potential to greatly improve the efficiency of road operations by enabling the early identification of dangers and allowing quick responses to potential harm for all participants involved in traffic.

B. Validation

In order to validate the proposed pipeline, we created a mini data set of 50 images of traffic scenarios on the highway at different locations, which ConnectGPT has never seen before. In the dataset, we have 20 images with incidents (”positive” in this context) and 30 images without incidents (”negative” in this context”), which is from camera images on the highway in Czech Republic originally from the road damage dataset [23]. And the ConnectGPT shall implement the following three tasks:

1) **Image understanding**: As a professional traffic expert, ConnectGPT is tasked with analyzing uploaded camera images of a highway to identify any traffic incidents, hazards, or road damages. ConnectGPT will examine each image individually and provide interpretations, clearly indicating which image is being referred to in each analysis.
2) **Information extraction and decision-making**: As a traffic expert, ConnectGPT will analyze uploaded images to determine the number of lanes on the highway and suggest lane closures using a binary code system, where ’0’ indicates a lane to be closed and ’1’ indicates a lane to remain open. Based on the DENM (ETSI EN 302 637-3) standard [21], ConnectGPT will identify the causeCode and subCauseCode, save this information along with the drivingLaneStatus in a CSV file named after the image, and provide it for download.
3) **C-ITS message generation**: As a C-ITS message expert, ConnectGPT will generate a DENM in XML format based on the contents of an uploaded DENM example. This task in our current work involves updating the causeCode, subCauseCode, and drivingLaneStatus, saving the complete XML file without omissions, and naming the XML file after the uploaded CSV file for download.

In the following part, we will elaborate on the quantitative as well as qualitative analysis of our experiment results.

1) Quantitative analysis: The table I shows our experiments conducted with 50 images across very different traffic scenarios presented on the highway. The restuls are divided into three main categories, as follows:.

Image Understanding: This task focuses on the system’s ability to correctly identify incidents in images with a perfect recall rate, meaning that it identified incidents without false negatives, and a precision rate of 95% (the only image with a false negative incident shown in Fig. 4(i)), indicating that ConnectGPT has high success rates in scene understanding, and major advantages in incident detection in general.

Information Extraction and Decision-making: This task deals with extracting information from the images and making decisions. ConnectGPT is effective at determining the number of lanes with a 90% correctness rate. However, its effectiveness drops significantly when determining the driving lane status (50%) and is particularly low when determining the cause code (15%). But a different answer of the cuase code does not necessariy the answer from ConnectGPT is not reasonable. We will discuss this phenomenon in detail in the next section.

Code Generation: The final task involves generating the content and format of V2X messages. ConnectGPT is fairly accurate at generating the content, with a 75% correctness rate but less accurate at determining the correct message format, with a 55% correctness rate. We also investigate the cause of the content or syntax issues for C-ITS message generation. It is because ConnectGPT may lose certain tags or create additional tags that are not part of the DENM protocol. Also, it may have difficulty converting a binary code to a string, such as ”000”.

2) Qualitative analysis: Fig. 4 presents nine exemplary images from the dataset, each accompanied by human and GPT-generated responses noted at the bottom. In Fig. V-B.1, ConnectGPT’s responses are indistinguishable from human answers. Fig. 4(b) shows ConnectGPT providing the matching Cause and Sub Cause Code as the human annotator, though reversing the open and closed lane statuses. In Fig. 4(c), GPT incorrectly identifies the number of lanes—a challenging task even for humans—and misinterprets a ”shed load” of white powders as ”snow drift.” Fig. V-B.1 contrasts the human annotator’s precise identification of ”rockfalls” as the accident cause with GPT’s broader categorization as a ”hazardous location - obstacle on the road.” This GPT response, despite being also correct, is classified as incorrect in Table I.

Further, in Fig. 4(e), GPT accurately detects ”human presence on the road,” rather than identifying the more accurate ”hazardous location - surface condition.” A same error is repeated in Fig.s 4(f), V-B.1, and 4(h), where ConnectGPT refers to an incorrect C-ITS message standard, despite clear guidelines in the prompt. Lastly, Fig. 4(i)—the only image not positively identified by ConnectGPT—presents a scenario that is also initially confounding for humans.

In short, ConnectGPT demonstrates notable capability in understanding and interpreting images. However, its code generation is found to be unsatisfactory and unstable, as discussed in Section V-B.1. Also, it does not perform so well when choosing a cause code according to a document for a given traffic scene. Similar to the findings of Chen et al. in their study on ChatGPT [24], we also experienced the variability in GPT-4’s performance, which means fluctuations in the effectiveness of the model over short periods of time.

VI. Conclusion and Outlook

In this paper, we developed an automated pipeline tailored for generating standardized V2X messages for Connected Autonomous Vehicles (CAVs) utilizing ConnectGPT,

TABLE I
QUANTITATIVE ANALYSIS OF OUR EXPERIMENTS WITH 50 TEST IMAGES

Task	Image Understanding		Information Extraction and Decision-making			Code Generation	
Sub tasks	Incident Detection		Number of lanes	Driving Lane Status	Cause Code	V2X Message Content	V2X Message format
Metric	Recall	Precision	Correctness Rate	Correctness Rate	Correctness Rate	Correctness Rate	Correctness Rate
Result	100%	95%	90%	50%	15%	75%	55%

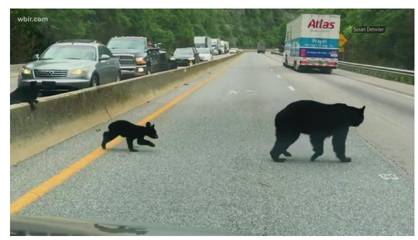
(a) Human: DLS(10), CC(9), SCC(0), text(hazardous location - surface condition); GPT: DLS(10), CC(9), SCC(0), text(hazardous location - surface condition).

(b) Human: DLS(01), CC(3), SCC(0), text("short-term stationary roadworks"); GPT: DLS(01), CC(3), SCC(0), text("short-term stationary roadworks").

(c) Human: DLS(001), CC(10), SCC(1), text("shed load"); GPT: DLS(01), CC(9), SCC(5), text("snow drifts").

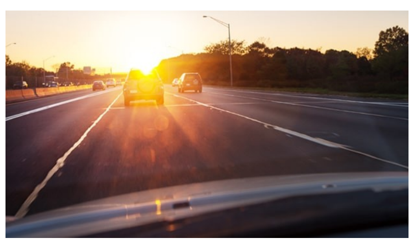
(d) Human: DLS(00), CC(9), SCC(1), text("rockfalls"); GPT: DLS(00), CC(10), SCC(0), text("hazardous location - obstacle on the road").

(e) Human: DLS(000), CC(9), SCC(0), text("hazardous location - surface condition"); GPT: DLS(100), CC(12), SCC(0), text("human presence on the road").

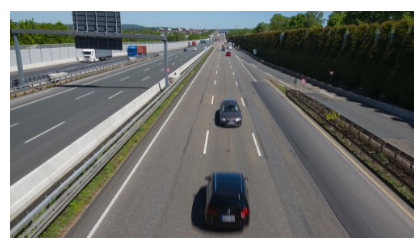


(f) Human: DLS(00), CC(11), SCC(0), text("hazardous location - animal on the road"); GPT: DLS(00), CC(91), SCC(0), text("refering to another standard").

(g) Human: DLS(111), CC(18), SCC(1), text(visibility reduced due to fog); GPT: DLS(111), CC(7), SCC(0), text("refering to another standard").



(h) Human: DLS(111), CC(18), SCC(6), text("visibility reduced due to low sun glare"); GPT: DLS(111), CC(5), SCC(0), text("refering to another standard").



(i) Human: DLS(101), CC(14), SCC(0), text("wrong way driving"); GPT: DLS(111), CC(0), SCC(0), text("no incident").

Fig. 4. Quantatitive analysis of answers of ConnectGPT. DLS stands for Driving Lane Status; CC stands for Cause Code; SCC stands for Sub Cause Code; the text is the label for each Cause Code with Sub Cause Code in DENM (ETSI EN 302 637-3) standard. The definition of Driving Lane Status is that'0' indicates a lane to be closed and '1' indicates a lane to remain open [21].

a proof-of-concept test LLM designed for V2X message generation. Our approach involves the use of infrastructure sensors, demonstrated in this paper specifically for camera images, in conjunction with ConnectGPT as a central decision support tool. The proposed pipeline generates standardized Decentralized Environmental Notification Message (DENM) messages to effectively alert oncoming CAVs about imminent dangers. Consequently, we analyzed the effectiveness of such a system in automatically generating safety-related V2X messages, providing prompt warnings and guidance for oncoming traffic while ensuring a reliable safety guarantee.

Currently we rely on GPT-4, which is one of the most powerful and well-trained LLMs, as the main source of the ConnectGPT. However, it is not sufficiently accurate for our specific ITS use cases. Current proof-of-concept implementation suffers from problems such as lacking proprietary knowledge, the risk of outdated information, and hallucinations (e.g., false positives). Our objective is to utilize an open-source LLM (such as Llama 2 [25]) to further train ConnectGPT, focusing on a reduced number of parameters while acquiring more specialized knowledge in the domain of CAV and ITS. Furthermore, the customized ConnectGPT will run in a local server at the infrasture side, so the potential data security concerns can be also addressed.

Future work involves integrating ConnectGPT with our existing computer vision pipeline for enhanced object detection, accurate bounding boxes, and estimating the incident location and orientation. This integration shall aim to enrich traffic management decisions with more detailed and reliable information from ConnectGPT with a possibility cross-verification of the results from different information sources. Additionally, we are exploring optimal pipeline configurations for automated C-ITS message generation, in tandem with our current manual workflow, to ensure 100% accuracy in both content and format of the standardized C-ITS messages. The fully established pipeline, including all the software and hardware components, will be developed and put through testing once again on an operating roadway in Austria.

VII. Acknowledgement

This work was conducted in the scope of the ESRIUM Project, which has received funding from the European Union Agency for the Space Programme under the European Union's Horizon 2020 Research and Innovation Programme and under grant agreement No 101004181. The content of this paper reflects only the authors' view. Neither the European Commission nor the EUSPA is responsible for any use that may be made of the information it contains. Virtual Vehicle Research GmbH has received funding within COMET Competence Centers for Excellent Technologies from the Austrian Federal Ministry for Climate Action, the Austrian Federal Ministry for Labour and Economy, the Province of Styria (Dept. 12) and the Styrian Business Promotion Agency (SFG). The Austrian Research Promotion Agency (FFG) has been authorised for the programme management.

The authors would like to extend their thanks to Mr. Kenan Mujkic for his contributions regarding the example implementations and mutual discussions on the C-ITS messages.

References

[1] D. Watzenig and M. Horn, *Automated driving: safer and more efficient future driving*. Springer, 2016.

[2] C. Cui, Y. Ma, X. Cao, W. Ye, Y. Zhou, K. Liang, J. Chen, J. Lu, Z. Yang, K.-D. Liao *et al.*, "A survey on multimodal large language models for autonomous driving," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2024, pp. 958–979.

[3] K. Tong, Z. Ajanovic, and G. Stettinger, "Overview of tools supporting planning for automated driving," in *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 2020, pp. 1–8.

[4] M. M. Rana and K. Hossain, "Connected and autonomous vehicles and infrastructures: A literature review," *International Journal of Pavement Research and Technology*, vol. 16, no. 2, pp. 264–284, 2023.

[5] SAE, "urface vehicle recommended practice–(r) taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles," SAE-J3016, 2021, this is an example standard citation.

[6] ESRIUM. (2022) EGNSS-enabled smart road infrastructure usage and maintenance for increased energy efficiency and safety on european road networks. [Online]. Available: https://esrium.eu/

[7] OpenAI, "Introducing gpts," https://openai.com/blog/introducing-gpts, 2023, accessed: Jan 31, 2024.

[8] J. Schulte-Tigges, M. Rondinone, M. Reke, J. Wachenfeld, and D. Kaszner, "Using V2X communications for smart odd management of highly automated vehicles," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, 2023.

[9] G. Nestlinger and M. Stolz, "Bumpless transfer for convenient lateral car control handover," *IFAC-PapersOnLine*, vol. 49, no. 15, p. 132–138, 2016. [Online]. Available: http://dx.doi.org/10.1016/j.ifacol.2016.07.721

[10] M. Rudigier, G. Nestlinger, K. Tong, and S. Solmaz, "Development and verification of infrastructure-assisted automated driving functions," *Electronics*, vol. 10, no. 17, p. 2161, Sep 2021. [Online]. Available: http://dx.doi.org/10.3390/electronics10172161

[11] M. Rudigier, S. Solmaz, G. Nestlinger, and K. Tong, "Development, verification and kpi analysis of infrastructure-assisted trajectory planners," in *2022 International Conference on Connected Vehicle and Expo (ICCVE)*, 2022, pp. 1–6.

[12] D. Maksimovski and C. Facchi, "Decentralized V2X maneuver sharing and coordination message for cooperative driving: Analysis in mixed traffic," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, 2023.

[13] S. Ha1, M. Lee, Y. Hwang, B. Kim, K. Yun, and H. Lee, "Connected vehicle-in-the-loop test environment with V2X communication and its applications for connected and automated vehicles," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, 2023.

[14] F. Schiegg, H. Tchouankem, A.-L. Rueeck, E. Xhoxhi, J. Gamerdinger, and G. Volk, "Accounting for the special role of infrastructure-assisted collective perception," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2023.

[15] C. Pilz, A. Steinberger, and G. Steinbauer-Wagner, "Positional accuracy provided by state-of-the-art cooperative awareness and collective perception," in *2023 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2023.

[16] S. Solmaz, G. Nestlinger, K. Diengsleder-Lambauer, R. Lesjak, S. Schweitzer, and J. M. V. García, "Real-life implementation and comparison of authenticated path following for automated vehicles based on galileo osnma localization," in *2023 IEEE International Automated Vehicle Validation Conference (IAVVC)*, 2023, pp. 1–6.

[17] S. Yin, C. Fu, S. Zhao, K. Li, X. Sun, T. Xu, and E. Chen, "A survey on multimodal large language models," *arXiv preprint arXiv:2306.13549*, 2023.

[18] W. Tong, A. Hussain, W. X. Bo, and S. Maharjan, "Artificial intelligence for vehicle-to-everything: A survey," *IEEE Access*, vol. 7, pp. 10 823–10 843, 2019.

[19] R. Binisha, P. Rajasekaran, K. Kalpana, T. ChithrakumarThangaraj, U. Ramani, and S. IrfanBasha, "Enhancement of v2x technology with artificial intelligence," in *2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)*, 2022, pp. 159–163.

[20] G. Harvas, W. Schildorfer, F. Hofbauer, M. Walch, and M. Neubauer, "D2.1 use case definition," ESRIUM, Tech. Rep., Nov. 2023. [Online]. Available: https://esrium.eu/wp-content/uploads/2023/11/Attachment_0-2.pdf

[21] TISA, "Safety related message sets: Datexii, denm, tpeg-tec, tmc," TISA, Tech. Rep., 2023. [Online]. Available: https://tisa.org/wp-content/uploads/ITSTF17001_SafetyrelatedMessage-Sets-DATEXII_DENM_TPEG-TEC_TMC_-v1.0.pdf

[22] S. Solmaz, F. Pölzlbauer, K. Lambauer, G. Nestlinger, K. Tong, M. R. Kenan Mujkic, R. Lesjak, S. Schweitzer, J. Vallet, and G. Allmer, "D5.3 test result analysis report," ESRIUM, Tech. Rep., Dec. 2023. [Online]. Available: https://esrium.eu/wp-content/uploads/2023/12/ESRIUM_DEL_VIF_2023-12-07_WP5_D5.3-Test-result-analysis-report_v2.0.pdf

[23] H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, "Road damage detection and classification using deep neural networks with smartphone images," *Computer-Aided Civil and Infrastructure Engineering*, vol. 33, no. 12, pp. 1127–1141, 2018.

[24] L. Chen, M. Zaharia, and J. Zou, "How is chatgpt's behavior changing over time?" *arXiv preprint arXiv:2307.09009*, 2023.

[25] H. Touvron, L. Martin, K. Stone, P. Albert, A. Almahairi, Y. Babaei, N. Bashlykov, S. Batra, P. Bhargava, S. Bhosale *et al.*, "Llama 2: Open foundation and fine-tuned chat models," *arXiv preprint arXiv:2307.09288*, 2023.