Deep Learning for Autonomous Vehicle Signal Recognition and Interpretation

By Dr. Kevin Whalen

Associate Professor of Cybersecurity, Purdue University (Branch outside normal colleges)

1. Introduction to Autonomous Vehicles and Signal Recognition

In particular, Deep Learning architectures have been successfully implemented in different classification tasks. A key element for the realization of infrastructure and vehicle classification systems is the ability to accurately classify the messages' contents in presence of various types of signal degradation. The various types of possible environment degradations may include strong background noise (e.g., a busy intersection with many vehicles and people; a noisy motorway), simulation of rain deposits on the camera sensor; and different forms of occlusions interfering with some or all relevant regions of the signals. All of these factors can contribute to a remarkable decrease in performance of recognition methodologies that are not specifically optimized with a view to robustness and minimal computational demand.

[1] [2]The development of technologies for autonomous driving has gained considerable interest from academics and industry, motivated by potential applications to the transportation system. Classification of Optical, Acoustic, and Lidar Signals for the Recognition and Interpretation of Safety Messages and Traffic Signs: Recently, the usage of Artificial Intelligence (AI) in autonomous vehicles has increased. This is because it is designed and developed with robustness and simplicity in mind. Specifically, Deep Learning (DL) has the potential to perform better and faster than its conventional counterparts. This has made Deep Learning techniques the main candidate in various tasks on both infrastructure and in-vehicle sensors.

1.1. Overview of Autonomous Vehicles

An intelligent self-driving car needs to understand and interpret the status of its environment, which requires dynamic scene detection, recognition of objects and indicating their activity.

Interpreting the dynamic and complex scenes encountered on the road requires very high-resolution representations. Therefore, it is interesting to use deep learning-based hierarchical representations. The setup to solve these problems is learnt end-to-end [1]. A significant number of research groups and automotive companies on autonomous and semi-autonomous driving car systems leverage deep-learningfancy techniques. Although engineers have created a suite of programs able to perform quite well in various navigating tasks is felt that deep-learning techniques will be a key factor to solve the remaining issues of traffic visual recognition and interpretation. However, deep-learning perfects irreplaceably in dealing with large amounts of data (more than 108 data points to be handled just to fairly call "large scale"), and it needs an extreme amount of thinking to teach the system, an automatic labeled and on the fly learning that mimics the development of the persons in the childhood.

As the focus of this report is on machine vision, we are going to describe system processing from the sensors (Section 1.2) to the application of intended driving commands in Section 1.2. The process starts with sensor data that need to be translated into semantic information by the use of perceptually meaningful intermediate representations (Section 1.2.1), which will allow the planning of driving manoeuvres based on human traffic code (Section 1.2.2) and feasibility constraints (Section 1.2.3). These will then provide the desired commands needed by the Control Unit to actuate the actuators.

1.2. Importance of Signal Recognition

However, while traffic-sign recognition handles traffic signs, it doesn't exclude non-standard objects conveniently fitting the traffic-sign recognition dataset. This means that other non-traffic-sign objects near or close to the car should be properly recognized and correctly judged. Meanwhile, efficient decision-making also includes classification of unnecessary or unhelpful traffic signs, which may appear on roads as artificial factors. In this way, drivers could focus on helpful signs and ignore unhelpful ones, which could be implemented even in advanced driver assistance systems (ADAS) [3].

[4] Deep learning has already achieved impressive results in computer vision tasks, including via Convolutional Neural Networks (CNN), which highlights feature extraction and automatically assigns weights to these features. Since deep learning can better target identification effects in complicated road scenarios, many approaches leverage it to handle traffic-sign recognition problems in intelligent vehicles. By automatically extracting valuable

features from high-dimensional images of traffic signs and turning them into compact linear separable forms, traffic sign identification can be famously treated as a binary classification task. Via deep-learning architectures such as LeNet, VGG, and GoogLeNet, researchers perform efficient image classification on traffic signs to send categorized knowledge to an upper-level decision-making process [5].

2. Fundamentals of Deep Learning

In order to embed intelligent capabilities in vehicles, it is important to design and evaluate strategies for obtaining signals that are representative of the signals seen by the driver while driving in different situations. Such signals require proper pre-processing, feature extraction, transformation, and representation so that they contain adequate information classifiable by learning models. Furthermore, fully convolutional models, even considering sequence history frames and encompassing spatial temporal dependency, do not suffice in the specific area of autonomous driving and intelligent vehicle systems. This will be explained through the analysis of driving environment and need for efficient non-end-to-end models that take into account long-term dependencies. .

Deep learning is moving the field of autonomous driving towards autonomous vehicles (AVs). Buses equipped with reprogrammable neural networks learned from large-scale real data to drive on predefined visual routes opened to traffic. The experiments involved more than 1 million kilo-meters of real traffic data collected yearly in the City of Geneva, including more than 27,000 kilometers recorded during ground truth data collection sessions. Such a neural distributed system based on plain RGB images led to satisfactory performances on a variety of challenging routes performed either by the official post bus of the City of Geneva company or the joint venture between the AV company PostBus and researchers from the POST T7 AI research group.

Deep learning's popularity arises from its ability to automatically learn visual feature representations, and to implicitly capture the complex interplay between features [6]. The three most commonly used deep learning approaches are Convolutional Neural Networks (CNNs), Recursive Neural Networks (RNNs), and Deep Reinforcement Learning (DRL). CNNs focus on processing spatial information; for this reason, they are mostly used for image recognition and development of autonomous vehicle perception modules [7]. Recurrent networks are particularly effective in handling temporal signals such as log data, and for sequence generation tasks, for example, they are used for recognizing complex driving behaviors including lane changing and safe-following separation distance.

2.1. Neural Networks

Motivated by this, we provide in this chapter a detailed overview of neural networks and deep learning theoretical frameworks that can effectively handle signal recognition and interpretation tasks in autonomous vehicles. As more sensor networks can be mounted onboard and manned missions can be changed into unmanned ones, signal recognition and interpretation in autonomous vehicles are becoming more and more important. To process data from different sensors, at present radar, Lidar, Stereo Camera, and acoustic sensors, would make maximizing use of data from different sensor information. The fusion of several sensing modalities has sometime yielded superior results than the data from single sensing modality, which is also known as sensor fusion.

[8]Recent years have witnessed increased interest in applying machine learning approaches, particularly deep learning technologies, in the realm of autonomous vehicle development to yield robust signal recognition and interpretation methods. Recognized for its exceptional ability to learn intricate and abstract representations from data, deep learning has exhibited superior performance in addressing many real-world signal processing, interpretation and recognition problems, specifically in undertaken real-time response tasks. Neural networks are essential ingredients in deep learning models and possess the capability to automatically learn relevant features in a high-dimensional space [9]. For autonomous vehicle technology, the capability to interpret and reason all types of collected information to make decisions has to be developed. The immediate need is to develop effective machine learning models to provide reliable, minimal intervention and safe decisions within a short time scale, while addressing issues of data quality and run-time performance.

2.2. Convolutional Neural Networks (CNNs)

CNNs mainly consist of convolutional layers which are directly followed by pooling layers (either max-pooling or average-pooling). The pooling layers are typically followed by a ReLU (Rectified Linear Unit) layer. In terms of the successes achieved on deep learning models, these have inspired researchers to further investigate and improve the performance of the models in many different aspects. The trend is to create a deeper network having more and more hidden layers. It has been experimentally validated that a deeper network will lead to more discriminative features and hence better performance [10]. Despite the improved performance, it raises significant challenges. It is well known that more layers will result in much slow training, while it faces more difficulties in selecting a suitable learning rate, and may result in the effect of gradient explosion or vanishing. That is, simply adding deeper layers would not result in the desired improvements.

Convolutional Neural Networks (CNN) are a specialized type of neural network that is specifically designed for recognizing and understanding signals that have grid-like topology. CNNs are primarily used for image recognition and interpretation, available in different networks and architectures. The success of CNNs is attributed to the fact that they exploit local spatial correlation between pixels in an image. Moreover, deep learning provides a promising method that might automatically abstract useful information from their raw data [11].

2.3. Recurrent Neural Networks (RNNs)

To address the above shortcomings, there are various improved RNN architectures with better memorization ability, such as Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) [12]. Among these improved RNNs, the LSTM units has been the most extended unit in practical applications. The salient proposition of LSTM lies in its ability to selectively choose which information from the past must be remembered and which information is least relevant to forget, which is very similar to human cognitive abilities. Theoretical and practical analysis on the memory content summarization process of LSTM has been conducted from the perspective of human cognitive science. Furthermore, LSTM is widely used in a variety of tasks. In this study, it is utilized to control autonomous vehicles to learn navigational policies. Following the introduction of the algorithm, an experimental framework that is designed to imitate real-world driving settings in a fictional environment is described.

A recurrent neural network (RNN) is a class of artificial neural network where connections between units form a directed cycle [13]. This creates a form of internal memory, making them ideal for tasks that must process sequential data. RNNs, which possess remarkable memory retention properties, are developed to serve as an important part of modern deep neural networks and are widely used for sequence to sequence learning, such as speech recognition, handwriting recognition, speech synthesis, and machine translation. Although they perform well for the above tasks, RNNs suffer from vanishing/exploding gradient problems, due to the sequential function composition, which makes training difficult.

2.4. Deep Learning Architectures for Signal Recognition

The end-to-end model, which directly maps inputs to output, is widely deployed in the autonomous navigation field. It is a comprehensive study of all the features extracted from the input signal X to build a direct mapping between Z1 and Z2. Specifically, in the localization task, end-to-end models extract key features and generate coordinates without intermediate steps. It is necessary to introduce some representative methods in deep learning, such as unsupervised pretraining and label smooth during supervised learning. In the specific task, apart from the above-mentioned model, attention mechanisms, mixed precision model training, model compression, distillation, and adversarial robustness can be utilized to achieve better generalization ability and accelerated training speed.

The recurrent neural network (RNN) is one of the oldest deep learning models and is widely used in signal processing, text analysis, image generation, and other fields. RNN serves as the foundation of more advanced networks, such as long short-term memory (LSTM) and gated recurrent units (GRUs) [9]. Unlike traditional RNNs, these advanced RNNs use a gating mechanism to selectively learn and omit information. In the application of signal recognition, feature domains, such as image, text, and audio signals, vary widely. CNNs, originally proposed for image processing, are also useful for other signal domains, including voice and natural language processing. In recent years, the Transformer model and its variants, which incorporate self-attention mechanisms, have made significant progress in the natural language processing field. The Transformer model, in most scenarios, eliminates the sequential learning process through parallel computation, while solving the long-distance learning problem. It is capable of handling multimodal tasks, including joint learning of images and text.

[ref: 7b68113e-abca-4fcd-85a9-10c079f4f5b7, 16b92040-a1ad-44c8-9180-42de851bb8bb]

3. Data Collection and Preprocessing

As the authors of [14] argue, the present approach differs from the existing literature by exploiting the audio correspond- ing to exhaust noise separately. The inspiration of this work is using the latent structured redundancy between spectrogram features and raw time-domain input to complement their respective limitations. This study's findings further the machine learning domain by introducing a deep learning method for bipedal sys- tems in a vehicle diagnostics application. The authors conclude that this combination of data transformation method and transfer learning technology can simultaneously address the limitations in classical feature engineering and the huge data investment needed by end-to-end architectures to deliver an improved vehicle fault diagnosis system. Adapting the proven identification capabilities of deep learning to a vehicle fault diagnosis system.

According to the authors of [15], audio signals can be more robust and convenient than images for categorizing vehicles. This is especially important for autonomous vehicles, where tough weather conditions can lead to changes in the visual scene. Extracting audio features can reduce the data required to identify a vehicle to a greater extent than image features. The identification of vehicles based on noise signals can be divided into two phases. In the first phase, the noise signals are separated into two types: engine noise and exhaust noise. Once this separation is achieved, the location of audible information can also be determined to classify the vehicle types by the frequency of the engine noise and the exhaust noise. This work allows identification of vehicles passing at different speeds and the time information corresponding to the designated frequency.

3.1. Sensor Data Acquisition

In the literature, the Automated Vehicles Signal Recognition and Interpretation problem is treated as independent object detection, segmentation and classification steps. Advances in research revealed that, spatial data localization, resolutional stride, data augmentation, transfer learning and se =mantic segmentation models that integrate long and short range environment observations are important for this process [16]. The preceding researches also focused on color image based IoU optimization of semantic segmentation networks such as U-Net or FPN. However in order to achieve efficient results in road user signal recognition and interpretation it is also crucial to evaluate temporal correlation between nearby observations, recurrent neuronal network restrictions of spatial information, model architecture, temporal memory improvement and the use of depth, lidar and radar sensors. In this study we aim to deal with these limitations and use multimodal sensor fusion in a whole End to End manner.

Using sensors in autonomous vehicles is crucial for dynamic environment perception, vehicle understanding, situation understanding, lateral and surrounding environment observation, learning, reasoning, and interacting with the environment [17]. Environment perception planning and control for automated vehicles are impossible without a reliable and elaborate environmental observation infrastructure. An ideal infrastructure corresponds well to the level of traffic environment complexity, and city volumes.

3.2. Data Annotation and Labeling

The data collected are of the type of today's commercial radar, delivering 3D point clouds (PCs) with an angle of 360° around the vehicle, together with their associated object signatures. Other configurable commercial automotive sensors (lidar, camera, etc.) can be used with EFTD. The variety of traffic light appearances in the EFTD datasets is relatively large, leading to a rich visual signature and object sophistication, depending on the manufacturer definitions, the obliquity of objects triggering the cameras, and the stature of the corpus. 3D modeling can also be used to guide annotations within data thanks to the PC.

[18] Autonomous vehicles (AVs) require high-level interpretation of sensor data, while processing perception robustness in very challenging environments. In fog or dusty conditions, the visual system is highly impacted, and the downstream perception effects can be catastrophic. An AV must be able to recognize traffic participants through the history of their movement as well, e.g., detecting signalized activities of other agents can ensure safe interaction with them. In order to cover a wide range of relevant scenarios and situations occurring in natural conditions, other vehicle systems need to rely on different types of sensors. Fusing of multiple sensor readings can increase perception reliability, using sensor diversity as dilution of failure modes.[19] Following a supervised training strategy, we need annotated public data reference. In the context of traffic light detection (TLD), several standard benchmark datasets have been annotated. The main difference with a fusion of data protocol or salient region proposal methods, both input to an image processing pipeline, is that Deep Learning (DL) directly reads sensor data and delivers interpreted representations as output in a joint algorithmic task. Unlike extensive PCs, we prefer to start the annotation process using EFTD only in order to feed it with additional diverse data for TLD. Moreover, the TLD annotations are associated from scratch.

4. Training Deep Learning Models

Advancements in deep learning methodologies are used in evaluating various driving events/activities and human driver behaviour. In particular, drivers' physical and physiological responses are analysed in several studies such as Model distortion score, model class distance or recalibration is observed with driving event recognition. However, the contributions in using deep learning are evident with reported results of cretasdingaddClassification performance in context of improved Figure and ideal signal combination [7]. AAAI says about formal education in the material, the reader is leading to a extensive bibliography and Deep neural networks have been used to model the relationship between states at distant times in a reinforcement setting without any exquisitely crafted inputs. Deep recurrent networks have been trained with reinforcement to learn to play multiple games such as Pong from raw pixel input, but training the continuous action actor-critic algorithm with a deep network results in poor performance in comparison to simple methods such as delayed actor learning.

Experiments have shown that a suitable frequency range is reasonable for effectively extracting the acoustic characteristics related to vehicle type. In this range, few vehicle sound events occurred, which may increase the acoustic emission of the considered vehicle types. Abstract features are extracted from the extracted temporal and time-frequency domain features. Consequently, conventional classifiers have been used to classify the vehicle sequences and some new candidate classifiers are tested at replacing support vector machine and logistic regression [15]. In order to keep classifications of each frame and obtain the final decision, typical frame-independent classifiers and some frame-dependent classifier are adopted. The VIVID dataset—a large dataset of in-cabin videos of driver behavior in a number of vehicles. This dataset is the first benchmark dataset for driving activity recognition using CNNs and its ultimate goal is to help spur research in the field, as well as to enable the transfer of benchmarking in published work to tangible products.

4.1. Model Selection and Evaluation

On the other hand, for Subtask 3—interpreting the recognized signals—the main goal is to create a data set with a lot of input signals within 50 ms and then predict a specific signal. In this task, creating an input data set is as vital as choosing a model. Since the input consists of vehicle speed and input signals, predicting the turn signal is a valuable problem to solve in understanding the intention of a vehicle. Thus, turns out to be classification-type output data consisting of four groups of [LEFT, RIGHT, BOTH, NONE]. The features themselves were used to input data in a 50 ms interval and other data such as road signs, lane, traffic light status, and distance of the detected signal were used to control the recognition diversity. Classification using signal and detection data no heavily bounded the possible model alternatives. Classifying only the vehicle turn signals for some models in the literature [10–12] makes the problem less specialized and more manageable. It also works with the assumption if vehicles are detectable well by the sensors, and that is an excessively optimistic assumption. The model chosen performs learn-immune processing, even if a single value or road sign not correctly predicts the output information.

Deep learning algorithms, like convolutional neural networks, recurrent neural networks, and deep spatial-temporal importance prediction, have been shown to work in camera sensor data for action recognition, human pose estimation using skeletal information, gesture recognition with both skeleton and depth data in three-dimensional analysis, and real-time multi-camera setup video data for human activity recognition and pose estimation, human pose estimation in the multi-camera setup, and vehicle detection in bird's eye view data. Besides optical sensors, the LiDAR and IMU sensors are also used. LiDAR uses scan data to detect objects and track them in the data, and IMUs can be used to estimate speed and traveling direction of vehicle. While object detection and tracking tasks are achieved using LiDAR sensor data and camera set-up by using deep learning techniques for lane detection, street sign classification and traffic light detection are also performed using image data from the camera sensors only. Another popular technique for machine learning is reinforcement learning. Here, the deep Q-network is used for input a data (either image or point clouds) to make an action (steer control or decide an action about turning on indicators) for autonomous vehicle control. After a while, it is used for a discrete action for lateral and longitudinal control using direct torque control from the steering and drive-by-wire system.

4.2. Hyperparameter Tuning

The efficiency of AutoML approaches used in the age of multimodal machine learning models is another component of model efficiency that must be carefully managed. Just as the performance of a model is influenced by sample quality and the number of samples available for training, AutoML methods include various hyperparameters that control the learning process itself. Although AutoML hyperparameters can significantly impact the quality of machine learning models, the combination of features and low data complexity makes model performance relatively robust against variations, resulting in the assumption of AutoML hyperparameters lacking a significant influence on model quality. However, the same heuristics established for AutoML on large language model applications may not be uniformly applicable to multimodal machine learning models used in autonomous vehicle perception tasks and may lead to suboptimal performance on large multimodal datasets. [20]

Recent improvements in the development of deep learning models have inspired the evolution of a variety of methods including hyperparameter tuning strategies to efficiently navigate the vast configuration spaces of these models. Achieving state-of-the-art performance requires the optimization of numerous hyperparameters such as data splitting, feature engineering, network architecture, regularization, optimization techniques, and the number of iterations. Consequently, automatic tools have been developed to address this optimization while minimizing expert intervention. However, AutoML methods have shown diminished performance for large multimodal datasets, including sensor data.

4.3. Transfer Learning

Simultaneous visual scene recognition and event anticipation methods using deep neural networks to recognize traffic lights and human traffic lights [9]. Transfer learning the images collected from the mixed scenario with the simulated images as source domain, the images collected in the new scenarios can be divided into three categories, and the model can predict the corresponding concurrent events in the mixed or real scenario, which will simulate the events in the real traffic scene and provide decision-making basis for automatic driving and intersection traffic lights. To solve incongruent event recognition in the autonomous vehicle and some corner cases, this transfer learning cooperative method is proposed with a context aware reinforcement learning model. In the new detection, the recognition probability heat map is targeted as the context of the simultaneous recognition event model, for most targets which can move, can be detected as following the second and a queue, and then the model on explains the global context, the probability for the second event is used to the effectiveness of the model in the autonomously vehicle vision recognition of the cascade event hypotheses and positive training.

The performance of traditional machine learning (ML) models can be greatly degraded when dealing with new problems in target domains and cannot be optimized. TL can well address the problem of knowledge sharing from a well-learned domain (source domain) to a new problem (target domain) where labeled data is difficult to obtain in a well-learned manner. Yunqiang Tang et al. proposed a deep transfer learning framework using domain adaptation with strategic attention and knowledge distillation to improve the recognition performance of autonomous vehicle signals in a different scene with limited labels [21]. In unsupervised DA model training, the model learns from source data to extract important spatial and channel features in target data, and the learned knowledge helps the model to improve generalization ability in target domain. Finally, TL framework is trained with the strategic attention and matching generation of synthetic data as the distillation knowledge. The experiment results show that the proposed method can effectively solve the problem of mismatch between the recognition of traffic lights or human lights in the target domain and the source domain due to the representation difference of few labeled data, and it can further improve the generalization ability of the model.

5. Signal Recognition Applications in Autonomous Vehicles

In order to ensure the development of biology and speed of athletes and enable the use of state-of-the-art data, it is important to know the main logical elements in the traffic management since the data effect means the field is very well researched as a source of inclusive data. The created database contains a total of 15,125 pieces, including light signals. SIFT local feature descriptors are used to detect the recognition elements necessary for the object in our simple skipping data processing, while the most accurate ensemble learning methods are used be visual ability to recognize a particular correlation of the objects identified in the different color spaces [22]. The recognition of traffic lights always depends on the quality of images and the quality of the data recorded within the database.

The recognition and interpretation of traffic signals in the environment by autonomous vehicles, in particular, artificial intelligence algorithms for real-time image and video data, continue to attract many researchers and engineers because of the importance in traffic safety, vehicle control and human trust. Using this reasoning, we described the general scheme and construction of the AI model for artificial recognition of light signals by a specially designed camera signal tuning system with a recording frequency of 25–30 frames per second [23]. The work also makes it original with the proposed camera's focus setting mechanism. Readers can also get an overview of the basic image processing algorithm using OpenCV, input image partitioning, and object recognition.

5.1. Traffic Light Recognition

In the literature, traffic light detection algorithms generally fall into two categories: image processing based and learning-based models. Image processing based models typically localize lights and classify state manually using color information. On the other hand, learning-based models are trained on a general set of traffic lights and are used to replace any image processing step to achieve real-time capability. The basic architecture of most object detectors in deep learning algorithms is the region-based CNN. One of the most famous and successful architectures of five- billion generation deep learning described is the Fast R-CNN, which uses the region composition network to calculate feature maps in places and assigns class probabilities to each region through a fully connected layer [24]. Other architectures have been used to speed up deep learning, for example, the use of feature pyramids or single shot detectors like YOLO or SSD. While these models improve real-time performance, they tend to struggle with multi-scale object detection, especially when it comes to large overlapping objects. So, the model can not be very good for localizing the training algorithm based on grid of the object and has one algorithm fix-sized objects in place and to predict class probabilities with the use of the full cornered later.

Deep learning is essential for traffic light recognition in autonomous vehicles as it can handle real-time conditions and difficult situations like over-exposure and occlusions [25]. It generally breaks down into traffic light detection, state classification, tracking, and light sequence inference. The term refers to the ability of an automobile to consult traffic signals to decide if it can continue without delaying the following vehicle. Deep learning is now an absolute essential component. A hierarchical multi-modal deep learning architecture to effectively detect and infer the light state was proposed. Each detection referred to an initial state [26]. Multi-modal fusion of image and depth camera data made it possible to map the light states to a door image, according to the observed orientation and distance, with high accuracy. The proposed approach significantly outperforms previous works due to a well-fitting module design and a mini-batch pruning method. Optical flow is also used to track the light signal. Thus, joint detection and tracking are considered in this work as well.

5.2. Pedestrian Crosswalk Detection

As discussed in this article, it is necessary for an autonomous car to perceive traffic signals and make well-organized interpretations while passing through road cross sections. This application targets self-driving vehicles for managing ADAS. The inbound information is sensed through a shield containing on a LIDAR sensor and 2 cameras, while controlling the warping, the intrinsic calibration, and the real-time pointed optimization. Incidentally, new entirely data has been introduced, which can practically represent the city driving patterns. It embraces the 124,000 images and more than ten thousand seconds of homogeneous traffic light fact images [27].

Building on top of the previously mentioned intersection-aware prediction [28], we are considering a Main Specified Transition in a distinct way of combining 2 different crowd counts together and their densities to represent the severity or density of pedestrians in the crosswalk, which is known as MS-T-DC. Creating this distinct feature makes this work to be more resilient for adaptation under contrasted climate conditions in light of the fact that the event of pedestrian crossing belongs to the verified activities by our primary detector [2]. Finalizing the whole project, we are subcontracting an international new testing scenario for running on the new data comparison, improving the detection accuracy. As can be observed, the primary goal of this study was to design and develop the autonomous signal recognition and interpretation in an intersection in hard-realtime constraint at the same time. As a result, the electrical system has been planed and installed based on the test track. Moreover, the Ava middleware has been developed in order to provide robust and plausible operation of: (i) Learning-based predictors for the ensemble of vehicles, persons and obstacles, and (ii) Model-based controllers for the self-driving vehicle of the AVs. The signaling system was evaluated in the proposed different scenarios.

5.3. Road Sign Recognition

In an “attention economy”, and autonomous vehicles as a local platform for future subscription services, our full focus will be controlled by different platforms. Even if the decision maker is a machine, either a human-like intelligent specialized application will try to control the majority of its user’s time for social, cultural, professional, educational or health reasons in manual or hands-free usage. It would be unlikely for millions of people during daily work, having repetitive or long activities to memorize and direct some of the dominant (self-declared) platforms of the 21stcentury (e.g. Facebook, Twitter, Instagram, Pinterest, LinkedIn) that want to be like "everything to everybody". That may turn this freedom to control our focus into a "cultural despotism". A reliable, high quality vehicle-oriented method cannot be left behind during development, especially if the expected use of traffical location will be more digital-like. The current method described could really improve the usual rate because during the examination normal participants are potential drowsy patients. That group, in the vicinity of the test, can be increased for recovering patients in care homes, or as we said before, to screen long-range truck drivers [29].

While there are many supporting strategies and technologies available to increase driving safety, very few can compare with the benefits that adoption of fully autonomous vehicles will bring. Our autonomous road vehicles can be used as a key element of the sustainable and smart cities project. Furthermore, the scientific literature has demonstrated that self-driving vehicles present at least a similar safety level as Uber and Lyft services do now. According to a March 18, 2020 Newsweek article, "In 2018, Alphabet's self-driving vehicles drove more miles (at an average of 11,017 of intervention every 11,109 miles on the public roads where they were allowed to drive) than 24 other companies operating in the state with testing permits; they had a total of 445 vehicles and 1,600 drivers by the end of 2018 [30]. The technology startup Nuro drove the second most self-driving miles, with its 62 vehicles driving 10,810 miles, followed by Aurora Innovation, Inc. and Zoox," while on the other hand, the generalized best self-driving companies have had relatively few collisions. Self-driving vehicles are also expected to have a positive effect on the motor vehicle ownership rate, To increase our confidence, we can refer to the projections by the Credit Suisse Institute, showing that "every 10 % increase in penetration of car sharing will lead to a decline in new car registrations for new and used cars over the period 2015/2016 to 2025" [4].

6. Challenges and Future Directions

[31] [32] Modern classification algorithms exhibit high precision and accuracy rates and hence have a wide range of applications for the intelligent decision-making system. However, the performance of algorithms is a major concern in terms of deploying them in real-time and meaningful signal recognition tasks. The drawbacks are curtailed by our next potential research direction that seeks to identify the computational-cost-effectiveness and embedded system integration techniques for the intelligent signal recognition module in autonomous-driving systems. The next research direction in this domain pertains to the detection of

external elements and variable environmental conditions which are currently in use to support and guide complex driving behavior. This is a critical investigation for autonomous vehicles in the deployed transportation system because the performance and precision rate of the detection model significantly influence the vehicle decision-making module. The future direction addresses the detection task in poor lighting and environmental conditions, e.g., night and fog, etc. Future research will address the possible techniques and methodologies for the self-driving vehicle to recognize and understand signals in real-life situations, altering drivers' behaviors and detecting the vehicles and cyclists' presence and making decisions.[33] The efficiency and quality of the detection task depend on the feature extraction mechanism and the number of samples used in the training and validation phase in the deep-recognition models. Therefore, LSTMs for recognizing difficult traffic signs due to their sequential connectivity-positive character are very promising. The detection of the traffic signals on the move is critical for any road and autonomous vehicles. The research also established that FL has a 100% precision rate in CVS and GTSRB detecting tasks, with reduced training case demand. This research has made several scientific contributions. Initially, the literature has not systematically elucidated the possible deep-learning solutions for the sign asset classifications in the traffic signal recognition domain, and this paper contributes to filling this gap.

6.1. Data Privacy and Security

In this work, privacy risks and protection methods are revisited in the context of AVs, with main contributions including a review of privacy risks and protection methods suited for deep learning-based autonomous vehicle applications and a discussion about future research directions [34]. In summary, this paper introduces the motivation behind investigating data privacy and security in deep learning-based AV ensemble models, reviews up-to-date research on privacy and security for AVs, provides readers with some prospect for future privacy protection, and provides future research options.

The increasing autonomy levels of autonomous vehicles (AVs) require increasingly data-driven, deep learning-based approaches. Unfortunately, deep learning networks require massive datasets and the large accumulation and centralized treatment of data raises data privacy concerns [35]. In this context, there are efforts towards federated or edge-computing with compressed data, private data synthesis, homomorphic encrypted data processing and privacy-preserving deep learning (PPDL) in general. Privacy issues have been framed as the level of information leakage, the identification of individual subjects and the perception of the external world by the AV, assessing privacy risks arising from the data handled by the AV at three levels as follows [36]: (1) the sensor data, (2) the view of the externally perceived world and (3) the decision-making process of the AV.

6.2. Robustness and Generalization

Smart industries host numerous examples of sensor data fusion between different sensors to achieve different objectives. This includes manufacturing companies that aim to assure component quality, detect failures of the production process, and monitor equipment status by collecting data from a wide variety of sensors including cameras, ultrasound, thermography, Vibration Analysis, Laser Inspection Systems and much more as in [article_id: be626e7a-d8c9-490a-9f78-6a7984dce2f2]. The process industries also extensively use distributed sensor systems to monitor the quality of the production process, health of the equipment, as a tool in process control, planning and optimization and for energy consumption monitoring. The sensors used in this setting include cameras, gas sensors, pH sensors, turbidity sensors, temperature sensors, pressure sensors, and many more. Another area where smart sensor systems are used is in the area of equipment and asset monitoring in heavy industries such as oil and gas, where the exhaust temperatures, wear and tear of the equipment, status and quality of the material, corrosion, leak detection etc., are monitored using all kind of sensors.

[9] [37]A research field that has recently gained popularity is that of intelligent sensor systems, where the fusion of data collected by different sensors leads to smarter and more robust systems. In [article_id: 01efb997-4cfa-4dc9-8793-7273b3d31273], for example, the authors present a combined video and lidar metric that leverage individual sensor strengths to model a more realistic scenario for vehicle trajectory prediction problem involving not only traffic data, but also high-definition maps. A second example is the domain of intelligent traffic management systems, where data from different sensors are used to monitor traffic conditions, manage traffic lights, inform drivers when and how to proceed, and supervise traffic when engaging in autonomous driving. A third example comes from control engineering as in the use of visual and GPS sensors that measure a power distribution system to infer the grid voltage and frequency without the use of any additional sensors.

References:

1. [1] S. Kuutti, R. Bowden, Y. Jin, P. Barber et al., "A Survey of Deep Learning Applications to Autonomous Vehicle Control," 2019. [PDF]
2. [2] A. Biglari and W. Tang, "A Review of Embedded Machine Learning Based on Hardware, Application, and Sensing Scheme," 2023. ncbi.nlm.nih.gov
3. [3] E. Hsueh-Chan Lu, M. Gozdzikiewicz, K. H. Chang, and J. M. Ciou, "A Hierarchical Approach for Traffic Sign Recognition Based on Shape Detection and Image Classification," 2022. ncbi.nlm.nih.gov
4. [4] R. Greer, J. Isa, N. Deo, A. Rangesh et al., "On Salience-Sensitive Sign Classification in Autonomous Vehicle Path Planning: Experimental Explorations with a Novel Dataset," 2021. [PDF]
5. Tatineni, Sumanth. "Enhancing Fraud Detection in Financial Transactions using Machine Learning and Blockchain." *International Journal of Information Technology and Management Information Systems (IJITMIS)* 11.1 (2020): 8-15.
6. Shaik, Mahammad, et al. "Enhancing User Privacy in Decentralized Identity Management: A Comparative Analysis of Zero-Knowledge Proofs and Anonymization Techniques on Blockchain Infrastructures." *Journal of Science & Technology*1.1 (2020): 193-218.
7. Vemoori, Vamsi. "Comparative Assessment of Technological Advancements in Autonomous Vehicles, Electric Vehicles, and Hybrid Vehicles vis-à-vis Manual Vehicles: A Multi-Criteria Analysis Considering Environmental Sustainability, Economic Feasibility, and Regulatory Frameworks." *Journal of Artificial Intelligence Research* 1.1 (2021): 66-98.
8. [8] F. Jibrin Abdu, Y. Zhang, M. Fu, Y. Li et al., "Application of Deep Learning on Millimeter-Wave Radar Signals: A Review," 2021. ncbi.nlm.nih.gov
9. [9] Y. Zhu, M. Wang, X. Yin, J. Zhang et al., "Deep Learning in Diverse Intelligent Sensor Based Systems," 2022. ncbi.nlm.nih.gov
10. [10] G. Roffo, "Ranking to Learn and Learning to Rank: On the Role of Ranking in Pattern Recognition Applications," 2017. [PDF]
11. [11] A. Pamuncak, W. Guo, A. Soliman Khaled, and I. Laory, "Deep learning for bridge load capacity estimation in post-disaster and -conflict zones," 2019. ncbi.nlm.nih.gov

12. [12] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili et al., "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," 2021. ncbi.nlm.nih.gov
13. [13] R. Trabelsi, R. Khemmar, B. Decoux, J. Y. Ertaud et al., "Recent Advances in Vision-Based On-Road Behaviors Understanding: A Critical Survey," 2022. ncbi.nlm.nih.gov
14. [14] C. S. Alex Gong, C. H. Simon Su, Y. E. Liu, D. Y. Guu et al., "Deep Learning with LPC and Wavelet Algorithms for Driving Fault Diagnosis," 2022. ncbi.nlm.nih.gov
15. [15] H. Chen and Z. Zhang, "Hybrid neural network based on novel audio feature for vehicle type identification," 2021. ncbi.nlm.nih.gov
16. [16] Łukasz Sobczak, K. Filus, A. Domański, and J. Domańska, "LiDAR Point Cloud Generation for SLAM Algorithm Evaluation," 2021. ncbi.nlm.nih.gov
17. [17] Z. Yan, L. Sun, T. Krajnik, and Y. Ruichek, "EU Long-term Dataset with Multiple Sensors for Autonomous Driving," 2019. [PDF]
18. [18] A. Ouaknine, "Deep learning for radar data exploitation of autonomous vehicle," 2022. [PDF]
19. [19] L. Tabelini Torres, T. M. Paixão, R. F. Berriel, A. F. De Souza et al., "Effortless Deep Training for Traffic Sign Detection Using Templates and Arbitrary Natural Images," 2019. [PDF]
20. [20] A. Tornede, D. Deng, T. Eimer, J. Giovanelli et al., "AutoML in the Age of Large Language Models: Current Challenges, Future Opportunities and Risks," 2023. [PDF]
21. [21] X. Liu, J. Li, J. Ma, H. Sun et al., "Deep Transfer Learning for Intelligent Vehicle Perception: a Survey," 2023. [PDF]
22. [22] F. Balducci, D. Impedovo, and G. Pirlo, "Detection and Validation of Tow-Away Road Sign Licenses through Deep Learning Methods," 2018. ncbi.nlm.nih.gov
23. [23] A. M. Nascimento, L. F. Vismari, C. B. S. T. Molina, P. S. Cugnasca et al., "A Systematic Literature Review about the impact of Artificial Intelligence on Autonomous Vehicle Safety," 2019. [PDF]
24. [24] J. Kaur and W. Singh, "Tools, techniques, datasets and application areas for object detection in an image: a review," 2022. ncbi.nlm.nih.gov
25. [25] F. El-zahraa El-taher, A. Taha, J. Courtney, and S. Mckeever, "A Systematic Review of Urban Navigation Systems for Visually Impaired People," 2021. ncbi.nlm.nih.gov

26. [26] A. D. Pon, O. Andrienko, A. Harakeh, and S. L. Waslander, "A Hierarchical Deep Architecture and Mini-Batch Selection Method For Joint Traffic Sign and Light Detection," 2018. [PDF]
27. [27] M. O. Sayin, C. W. Lin, E. Kang, S. Shiraishi et al., "Reliable Smart Road Signs," 2019. [PDF]
28. [28] N. Radwan, W. Burgard, and A. Valada, "Multimodal Interaction-aware Motion Prediction for Autonomous Street Crossing," 2018. [PDF]
29. [29] K. Alexandridis, "Machine Learning Computer Vision Applications for Spatial AI Object Recognition in Orange County, California," 2023. [PDF]
30. [30] K. Dhawan, S. Perumal R, and N. R. K., "Identification of traffic signs for advanced driving assistance systems in smart cities using deep learning," 2023. ncbi.nlm.nih.gov
31. [31] C. Corbière, "Robust Deep Learning for Autonomous Driving," 2022. [PDF]
32. [32] J. Kocić, N. Jovičić, and V. Drndarević, "An End-to-End Deep Neural Network for Autonomous Driving Designed for Embedded Automotive Platforms," 2019. ncbi.nlm.nih.gov
33. [33] Y. Chen and Z. Li, "An Effective Approach of Vehicle Detection Using Deep Learning," 2022. ncbi.nlm.nih.gov
34. [34] H. Cao, W. Zou, Y. Wang, T. Song et al., "Emerging Threats in Deep Learning-Based Autonomous Driving: A Comprehensive Survey," 2022. [PDF]
35. [35] F. Manfio Barbosa and F. Santos Osório, "Camera-Radar Perception for Autonomous Vehicles and ADAS: Concepts, Datasets and Metrics," 2023. [PDF]
36. [36] C. Xie, Z. Cao, Y. Long, D. Yang et al., "Privacy of Autonomous Vehicles: Risks, Protection Methods, and Future Directions," 2022. [PDF]
37. [37] D. Yu, H. Lee, T. Kim, and S. H. Hwang, "Vehicle Trajectory Prediction with Lane Stream Attention-Based LSTMs and Road Geometry Linearization," 2021. ncbi.nlm.nih.gov