Paper number ITS-TP18524

Federated learning to enable automotive collaborative ecosystem: opportunities and challenges

Lei Chen*, Martin Torstensson, Cristofer Englund

RISE Research Institutes of Sweden, Sweden

lei.chen@ri.se, martin.torstensson@ri.se, cristofer.englund@ri.se

Abstract

Despite the strong interests in creating data economy, automotive industries are creating data silos with each stakeholder maintaining its own data cloud. Federated learning (FL), designed for privacy-preserving collaborative Machine Learning (ML), offers a promising method that allows multiple stakeholders to share information through ML models without the exposure of raw data, thus natively protecting privacy. Motivated by the strong need for automotive collaboration and the advancement of FL, this paper investigates how FL could enable privacy-preserving information sharing for automotive industries. We first introduce the statuses and challenges for automotive data sharing, followed by a brief introduction to FL. We then present a comprehensive discussion on potential applications of federated learning to enable an automotive collaborative ecosystem. To illustrate the benefits, we apply FL for driver action classification and demonstrate the potential for collaborative machine learning without data sharing.

Keywords:

FEDERATED LEARNING, AUTOMOTIVE DATA SHARING, PRIVACY-PRESERVING

Introduction

Cars are becoming connected, with each other and with infrastructure. In addition to the traditional telematics services that collect limited types and amounts of data, today's connected vehicles are essentially part of the digital transport systems and generate large amounts of rich data in real-time. According to Frost & Sullivan (1), by 2020, 98% of all new cars will be connected, and each car could generate as much as 25GB data per hour. With connectivity, the enormous real-time vehicle data is becoming the new assets for many stakeholders including the original equipment manufacturers (OEMs) who can access the data primarily, and other stakeholders, e.g., Tier 1 suppliers, insurance companies, telematics providers, authorities, among others. Investigating the value of such assets through sharing data, building ecosystems and monetization is becoming an emerging and intensive development area.

Automotive data sharing ecosystem: from OEM exclusive access to multi-stakeholder sharing

With the advancement of vehicle automation and connectivity, car data becomes the new digital oil that potentially can contribute to many new vehicle applications and generate new business values for stakeholders. Realizing such potential, OEMs have been connecting their vehicles and collecting data to their cloud platforms. Essentially, each OEM maintains its own connected vehicle cloud and tries to

investigate the value of such data. Therefore, in most of the cases, OEMs are the exclusive actors that have access to the data (with customer agreement from the vehicle owners), while third parties such as after-market service providers have no choice but to buy such data from OEMs. This creates barriers for equal data access and rapid innovation in data-driven services. As another important factor, most of the car data relates to driver information, which is highly private. The European General Data Protection Regulation (GDPR) applies in such situations and therefore, both consensus from the vehicle driver and strong encryption methods are required to protect the driver privacy.

In view of the enormous potential of connected vehicle data and to create equal competition for innovative vehicle services and the data economy, the industry has been working on different initiatives to enable data sharing. One notable commercial solution is direct OEM to OEM data sharing such as the live data sharing between Volvo cars and Volvo trucks for improving traffic safety (2). Another solution is the on-going pilot project Data for Road Safety (3) for decentralized sharing of safety-related vehicle data between OEMs, service providers, as well as member countries' authorities.

In addition to safety-related data, other types of vehicle data could be used for creating data-driven services, stimulating innovation and business development. One potential way for sharing the data is through the neutral server concept as was described by C-ITS platform (4), and automotive industry consortiums (5) (6). Illustrated in Figure 1, neutral servers are provided by independent third parties and are neither operated nor financed by OEMs. The aim is to provide equal data access to automotive data for service providers.

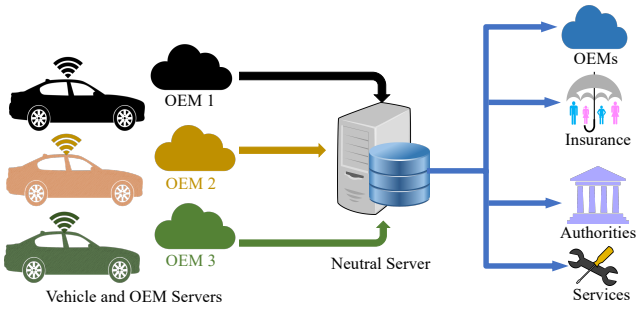


Figure 1 Neutral server to facilitate multi-stakeholder data sharing

Machine learning and federated learning: from data sharing to model sharing

The increasing availability of data and computing power has led to the booming of artificial intelligence (AI), which can be shown by the rapid development of computer vision, natural language processing and the autonomous driving. To train a machine learning (ML) model, especially in deep learning (DL), the common method is to collect large amounts of data to a powerful data server. For automotive industries, this means the cars need to send large amounts of driving and monitoring data to the cloud, either in real-time or through periodical data copying. The advancing of AI and the interests on data availability have triggered the discussion of data ownership, and the concern on data privacy and confidentiality. For vehicles, much of the real-time data is related to the drivers, which is highly private. While OEMs have access to those data

through consensus, sharing such data with other stakeholders is facing challenges, especially with the adoption of the GDPR within the EU and similar legislative actions in other regions.

The increasing awareness of data privacy, together with the development of advanced ML framework has led a new ML paradigm, Federated Learning (FL). With FL, instead of collecting large amounts of training data to a central server for global training, FL brings ML to the devices for local training without the need of data collection. In other words, while traditional ML brings data to the computing resource at central servers, FL brings computing to local devices, where data is generated. This essentially creates a distributed AI system of systems (SoS) for privacy-preserving collaboration. Since training is done locally where the data is generated and no raw data needs to be exchanged, the possibility of data leakage is reduced. This will ease the data privacy concern and encourage multi-stakeholder collaboration for rapid innovation.

FL (7) (8) was proposed and has been applied in applications such as Google's Gboard system for typing word auto-completion (9). It is a typical server-client FL framework, where an aggregation server exists for model parameter averaging and aggregation. Figure 2 illustrates such a server-client FL framework. It is acknowledged that this paper takes the server-client framework for discussing the key features of FL, while different frameworks exist such as peer-to-peer (P2P) (10) to accommodate different collaboration schemes and business strategies. FL in the shape of the server-client consists of many training participants and an aggregation server. The participants could be mobile phones, vehicles, or OEM clouds, and each maintains its own data and executes training locally. The aggregation server is usually a trusted 3rd party where the ML models are aggregated into a global model. A typical server-client FL consists of the following steps.

① The aggregation server choses participants to join a FL training process and distributes the initial model

② Each participant takes the model and uses its locally observed data to train a new model in parallel

③ All participants upload the newly trained models or gradients to the aggregation server

④ The aggregation server averages the models from the chosen FL participants into a global model

⑤ The newly aggregated global model is distributed to all participants for a new training round

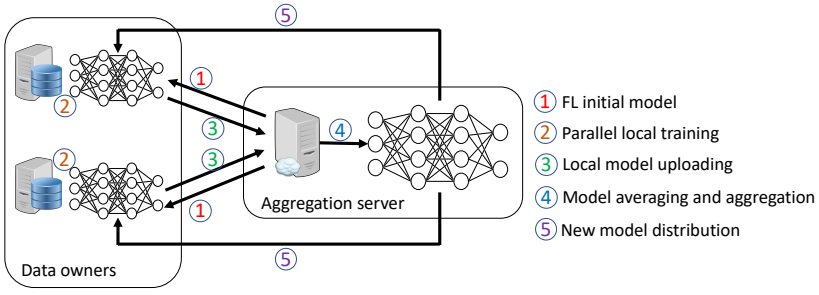


Figure 2 A server-client FL framework

Depending on the similarity of user samples, data features, as well as application areas, FL can be categorized into three types (11), the Horizontal FL (HFL), the Vertical FL (VFL), and the Federated Transfer Learning (FTL). Taking the most commonly used supervised learning as an example, HFL applies on scenarios where the data sets from different stakeholders share the same feature and label spaces, while

the sample space is different. VFL, on the other hand, applies on scenarios where the data sets share the sample space (e.g., same users) but the feature and label spaces are different. For FTL, the data sets are highly heterogeneous where only small sets of samples and features are common. We discuss the application potential of FL in the automotive industry with a focus on HFL and VFL.

FL for automotive collaborative ecosystem

The privacy-preserving nature of FL indicates strong potential in the automotive collaboration ecosystem, which may lead to solutions that eliminate the risk of intrusion in privacy during information sharing and stimulate innovative applications. In this section, we identify key application opportunities of FL in the automotive industry with detailed discussion on the FL framework, the procedure, as well as the challenges.

HFL for single stakeholder decentralized learning

As discussed in previous section, OEMs have exclusive access to the vehicle data within its brand. In the case where drivers have very strong needs of data privacy, OEMs could apply FL for training of certain vehicle functions such as autonomous driving (AD) without actually collecting the raw data. Similarly, for Tier 1 suppliers such as driver monitoring system providers, FL may allow them to train their system within the vehicle, thus skipping the process to get contracted data from OEMs. The same method applies to other service suppliers such as voice assistance systems. Since OEMs or Tier1 suppliers share very similar application spaces, this leads to a typical scenario for HFL that can help a stakeholder to train common models without collecting raw data from the end-users such as drivers.

In Figure 3 we present two use cases applying to a single OEM and a service provider, respectively. In Figure 3(a), FL can be used to support many AD function developments without collecting raw driving data. Many of the AD models require a collection of large amounts of data for training. Those data contain very private information, such as behaviors and actions of the drivers. In addition, external cameras record information of the surrounding environment which may contain information on other vehicle drivers, pedestrians or cyclists. This usually involves the GDPR and requires methods to remove sensitive information such as masking the faces. Similar situations are faced by Tier 1 or other service providers such as driver monitoring system providers. Normally, driver monitoring systems consist of cameras inside the vehicles for detecting the driver focus, drowsiness and other conditions. Those cameras may also capture images of other passengers within the vehicle. If suppliers need to collect that information for centralized model training, consensus with the driver will be needed and GDPR will apply. With FL, model training is done within the vehicle and the storage of data will be minimized, which will help to eliminate the privacy and GDPR concern. This will significantly ease the efforts to apply continuous model training for developing and improving the AD functions.

Applying HFL within a single stakeholder domain with the server-client framework is rather straightforward since the OEM cloud can serve as an aggregation server. The challenges lie in a proper strategy for distributed learning with consideration of resource availability at the vehicle side. For any neural network

training process, the aggregation server needs to initialize the training with a proper model and hyperparameter setting. The training process needs to consider communication performance to deal with unstable and limited wireless connectivity. Furthermore, as a common challenge for FL, a secure and trusted collaboration framework is needed to deal with potential cheating participants.

(a) FL applies to vehicles within the same OEM or fleet company for functions such as autonomous driving, driving behavior recognition

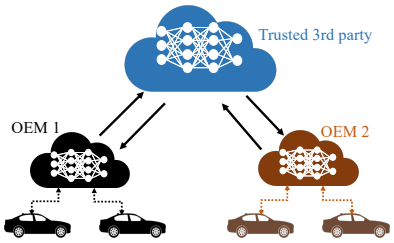
(b) FL applies to vehicles within the same Tier 1 or service suppliers such as driver monitoring systems

Figure 3 HFL application for single stakeholder

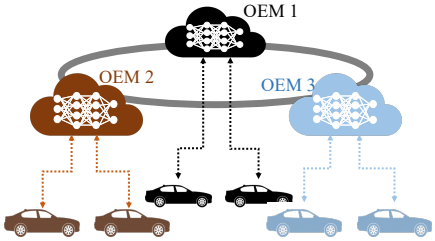
HFL for cross-OEM multi-stakeholder collaboration

Keeping data within the vehicles protects the driver and passenger privacy substantially, while the shortcomings are apparent. Vehicles have limited power and computational resources, therefore, participating in certain FL tasks brings extra power consumption, especially for electric vehicles. A much more energy and computationally efficient method is to move the computation to local edge servers or cloud servers. As mentioned, OEMs already own their cloud services where the neutral server concept provides a method to connect the OEM-owned data for cross-OEM collaboration. The same concept can be reused here to support FL. Instead of being the neutral party for data aggregation, the neutral server may play the role of model aggregator. Since model aggregation doesn't require raw data exchange, it becomes easier for the neutral server providers to handle GDPR while allowing multi-OEM collaboration. If a third party is not considered, with strong interests for collaboration between OEMs, they can also establish a P2P FL framework for certain training tasks. In both the use cases, we anticipate the application of HFL since different OEMs share the same applications (e.g., vehicle and mobility) with different user groups.

Figure 4 illustrates both the server-client case and the P2P case for FL application in the multiple OEM setting. In Figure 4(a), the framework leverages the similar architecture of a neutral server, where the trusted 3rd party can be neutral server providers or other trusted partners. The participating OEMs and the 3rd party agree on common model training tasks and data needs. The trusted party initializes FL and distributes an initial model to each of the OEM clouds. Each OEM then does parallel model training by using their locally available data. This follows exactly the same procedure as the single-OEM server-client training, while the training now can leverage much more powerful server capacities with no concerns on power limits. For P2P collaboration, OEMs can agree with each other directly for certain model training. Illustrated in Figure 4(b), each OEM is both a local trainer and an aggregator. It receives model updates from other peers, aggregates with the model trained on local data for a new model and updates the model to its peers. Since there is no 3rd partner, a secure P2P coordination protocol will be needed.



(a) A server-client HFL architecture for mutliple OEM collaborative learning.



(b) A peer-to-peer HFL architecture for multiple OEMs collaborative learning

Figure 4 HFL for cross-OEM multi-stakeholder collaboration

As can be interpreted, with the involvement of a 3rd party, the biggest challenge is to establish such a collaborative framework with strong business potential. The trusted 3rd party needs to work closely with each of the participating OEMs for the task definition, model training, as well as communication protocol. Collaboration incentives from each OEM are critical where benefits need to be justified through joining such a FL framework. Trustworthy and secure communication are even more challenging in comparison to single-stakeholder as external communication is involved.

VFL for cross-industry multi-stakeholder collaboration

HFL within the automotive industry depends on data sets with the same feature space since data mostly comes from vehicle sensors and the applications are driving-related. Going beyond the automotive industry, FL could also be used for cross-industry stakeholder collaboration. As introduced, VFL applies to data sets with different feature spaces but share the same users. A driver who owns a car usually has certain insurances under the same name. He/she may have data records within certain authorities such as the road administration and/or certain service providers e.g., social networks. Those different industries have different application areas with different business goals, and the domain data contains different feature and label spaces. The common cross-industry information is the user space, where common customers across the industries could enable cross-industry VFL.

Figure 5 illustrates the application of VFL for cross-industry federated learning beyond the automotive industry. The connected vehicle data contains rich information from simple telematics data to real-time vehicle sensor data and user-related behavior data. The external sensors also deliver rich environment data e.g., road information and road user information. By jointly building ML models that combines information from both the automotive data and its own domain data, other stakeholders may leverage the knowledge from the automotive domain for better services. For example, the insurance company may leverage information derived from driving data and deliver better models for insurance offers. The road administration can leverage vehicle sensor data for better traffic flow prediction and road condition maintenance. And social network providers can leverage the driver information for providing personalized information and services.

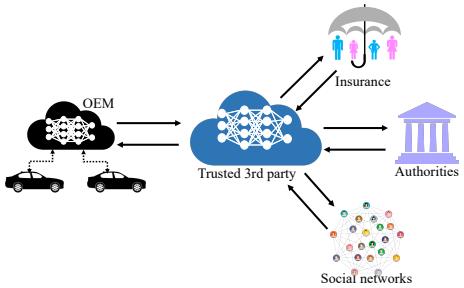


Figure 5 Cross-industry VFL for privacy preserving collaboration

In contrast to HFL, where feature spaces are the same and common models can be used, in VFL the only common information is the shared customers, e.g., a small set of sample spaces. This requires additional steps to securely identify the common samples, thus, much closer interaction between cross-industry stakeholders.

FL for driver action classification

Driver action classification is becoming an increasingly popular tool when studying driver behaviors for e.g., driver distraction detection and personalized services. In-vehicle cameras are commonly used for driver action classification when applying a machine learning approach. The image data collected for centralized model training is highly personal and requires complex procedures to satisfy GDPR. With FL, as already discussed, model training can be done directly within the vehicle and no images need to leave it. This provides a promising solution that supports continuous model training and improvement while protecting the drivers' privacy. Driver action classification is a typical application area for HFL, where the feature, label spaces and environment are the same, while drivers are different. To verify the potential of FL in the setting of driver-action recognition a series of experiments is conducted. A neural network architecture is chosen and trained on a dataset with an HFL setup. For comparison, the same neural network is also trained in a centralized setup. The two approaches are compared to demonstrate the potential of HFL for privacy-preserving driver action classification.

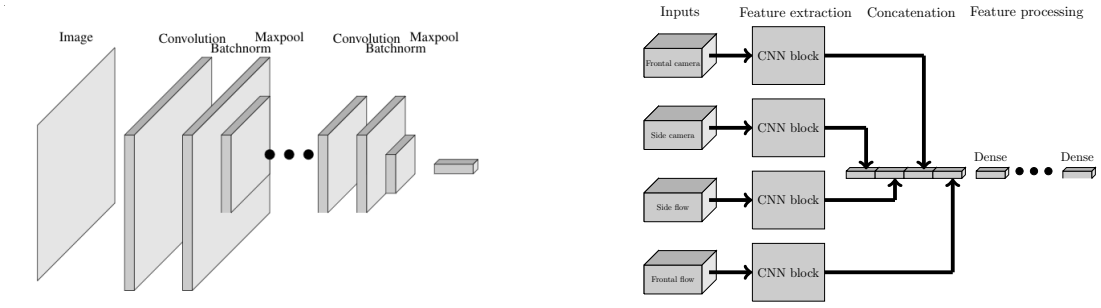
The dataset, preparation and the chosen model

The dataset utilized in this study is the Multimodal Multiview and Multispectral Driver Action Dataset (3MDAD) (12), which consists of two video streams from two Kinect cameras, one on the dashboard and the other on the door handle. It contains a set of 16 different action classes performed by a total of 50 participants. Both RGB and depth data are available, while we employ only the RGB data for FL training.

For the neural network training, the dataset was first split into a training and a testing set with 45 and 5 of the participants respectively. FL is a distributed learning process where parallel training processes are done on different nodes. For emulating such a process, the training set is again divided into 20 nodes where each node is given roughly the same amount of data from each class. The distribution of data to the nodes is done randomly from all the participants in the training set.

In HFL the different parties need to agree on a joint model to be trained. For driver action classification, we

modify a CNN-based image recognition model, which was developed in our previous studies (13). The model inputs consist firstly the two images from each camera, rescaled to 256x128 pixels. In addition, optical flow fields (14) are calculated for each of the images together with the corresponding previous image. In total, there are four input channels at each time point. For each of the input channels, data goes through a CNN block, shown in Figure 6(a), which contains four groups of series of single convolutional, batchnorm, and ReLU layers. After, the outputs of the four channels are concatenated into one vector, which then goes through the following dense, dropout with a keep ratio of 0.8 and finally a softmax layer. The whole process is illustrated in Figure 6(b).



(a) A CNN block showing two out of the four groups of convolutional, batchnorm and maxpool layers.

(b) A diagram of the complete trainable network model.

Figure 6 The chosen CNN model and the local training framework for each node

The experiments and results

With the dataset and the chosen model, HFL is designed by following the FL framework as shown in Figure 3. The experimental HFL framework consists of an aggregator for model aggregation and distribution, and 20 nodes for local model training. For each iteration, the aggregator first distributes the global CNN model to each of the nodes. Notice for the first iteration, an initialized model is created at the aggregator based on the chosen CNN model. After receiving the global model, each of the 20 nodes executes local training based on their locally allocated data and updates its respective model. This process is emulated in a single desktop sequentially. After the training, local models are sent to the aggregator for model aggregation. We employ FedAvg (8) which is the current state-of-the-art algorithm for FL model aggregation. Notice that transmitting models from the aggregator to the nodes and vice versa requires reliable communication support. Since we emulate the processes in one machine, communication performance is not the focus, instead, we test the performance in an environment with perfect communication. Experiments have been done on a desktop with a GeForce GTX 1080 Ti graphics card. Keras (15) with tensorflow (16) is used during both the training and testing process.

Figure 7 shows accuracies versus epochs for the implemented HFL and the centrally trained model. Here the accuracy is the average percentage of correctly classified images in the validation set. For HFL, an epoch refers to one iteration of FL which consists of one epoch of local training at each node and one global aggregation of the resulting models at the aggregator. For the centralized model, an epoch refers to a normal

epoch of training. The learning rate for the HFL is set as 10 times greater than that of the centralized model. Notice learning rate may affect the training convergence speed, which will be part of our future studies.

As can be seen, for the centralized model training, the maximum accuracy of 43.1% can be reached in fewer epochs in comparison to the HFL model with a maximum accuracy of 44.9%. It can, therefore, be demonstrated that with the current setup the HFL model is able to perform similarly to centralized training in regard to accuracy, at the cost of an increased training time.

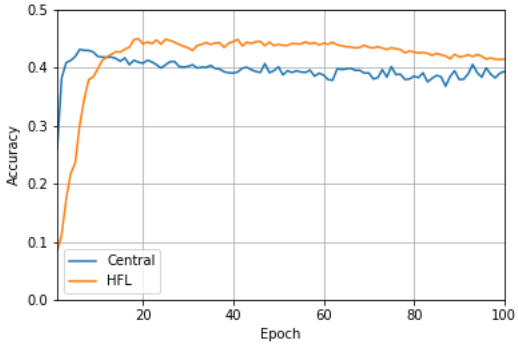


Figure 7 Accuracy of the proposed HFL and centralized model

Conclusions and future works

Motivated by the strong needs for automotive information sharing and with consideration on the stringent privacy regulation, this paper presents a privacy-preserving collaborative learning framework based on the emerging federated learning (FL). With FL, stakeholders can cooperate through collaborative model training without sharing raw data, thus, increasing the protection of data privacy and easing the effort to comply with regulations. The paper has discussed different alternatives to apply FL for building automotive collaborative ecosystems including horizontal FL for both single-OEM and cross-OEM collaborative model training, vertical FL for cross-industry collaboration, and their potential applications. To demonstrate the potential of FL, the paper also conducts experiments of FL on driver action classification. The results show that FL has a strong potential for privacy-preserving model training in automotive applications.

FL is emerging and extensive further research is anticipated. One natural next step for the drive action classification application is the model selection and hyper parameter optimization. This paper has presented initial results on the use case with FL based on one selected model with simple parameter setting. Experimentation with more models and optimizing the hyper parameters forms an interesting research direction for the use case. Another research question is the high requirements on communication reliability. A potential research direction is to apply the network slicing methods from the 5th generation (5G) telecommunication networks. Lastly, expanding the current experiments to more automotive applications by integrating new models for a comprehensive FL framework for automotive industries forms a research direction that is highly relevant to privacy-preserving collaboration.

Acknowledgements

The work is partially supported by the Swedish innovation agency Vinnova through the project SoSER – System of Systems for efficient Emergency Response and Urban Mobility.

References

1. Frost & Sullivan, "Otonomo, 2018 European Car Data Platform New Product Innovation Award," 2018.
2. "Volvo Cars and Volvo Trucks share live vehicle data to improve traffic safety," 2018. https://bit.ly/volvo-car-truck-datasharing.
3. "Data for road safety," 2020. https://www.dataforroadsafety.eu/.
4. European Comission, "C - ITS Platform Final report," January, 2016.
5. Verband der Automobilindustrie, "Access to the vehicle and vehicle generated data," 2016.
6. European Automobile Manufacturers Association, "Access to vehicle data for third-party services," 2016.
7. J. Konečný, H. B. McMahan, D. Ramage, and P. Richtárik, "Federated Optimization: Distributed Machine Learning for On-Device Intelligence," pp. 1–38, 2016, [Online].
8. H. Brendan McMahan, E. Moore, D. Ramage, S. Hampson, and B. Agüera y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, AISTATS 2017*, 2017.
9. A. Hard *et al.*, "Federated Learning for Mobile Keyboard Prediction," 2018, [Online]. Available: http://arxiv.org/abs/1811.03604.
10. A. Lalitha, O. C. Kilinc, T. Javidi, and F. Koushanfar, "Peer-to-peer Federated Learning on Graphs," 2019, [Online]. Available: http://arxiv.org/abs/1901.11173.
11. Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated Machine Learning: Concept and Applications," *ACM Trans. Intell. Syst. Technol.*, vol. 10, no. 2, pp. 1–19, 2019.
12. I. Jegham, A. Ben Khalifa, I. Alouani, and M. A. Mahjoub, "MDAD: A Multimodal and Multiview in-Vehicle Driver Action Dataset," in *Lecture Notes in Computer Science*, 2019, vol. 11678 LNCS, pp. 518–529.
13. M. Torstensson, B. Duran, and C. Englund, "Using recurrent neural networks for action and intention recognition of car drivers," in *ICPRAM 2019 - Proceedings of the 8th International Conference on Pattern Recognition Applications and Methods*, 2019, pp. 232–242.
14. G. Farneb, "Two-Frame Motion Estimation Based on," *Lect. Notes Comput. Sci.*, vol. 2749, no. 1, pp. 363–370, 2003.
15. C. François, "Keras." https://keras.io/.
16. M. Abadi *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems," 2016, [Online]. Available: http://arxiv.org/abs/1603.04467.