# Federated Learning for Lidar Super Resolution on Automotive Scenes

Alexandros Gkillas
Industrial Systems Institute,
Athena Research Center,
Patras, Greece

Gerasimos Arvanitis
Electrical and
Computer Engineering,
University of Patras, Greece

Aris S. Lalos
Industrial Systems Institute,
Athena Research Center,
Patras, Greece

Konstantinos Moustakas
Electrical and
Computer Engineering,
University of Patras, Greece

***Abstract*—In this paper, the problem of lidar super-resolution is explored under a federated learning perspective. The high cost of high-resolution lidar sensors is a major obstacle to the widespread adoption of connected and autonomous vehicles (CAVs). To reduce the cost, this study investigates the use of low-cost sensors in conjunction with super-resolution algorithms. Unlike previous studies that approach this problem with centralized solutions, this work employs federated learning to leverage private lidar data from different autonomous vehicles under different environmental conditions, resulting in more robust and diverse deep learning model. Extensive experiments on a real-world lidar odometry dataset highlight the merits and applicability of the proposed lidar super-resolution method with federated learning.**

***Keywords*—federated learning, lidar, super-resolution, deep learning, autonomous driving**

## I. Introduction

Connected and Autonomous Vehicles (CAVs) are a rapidly evolving area of machine learning that has the potential to revolutionize modern life [1]. However, widespread adoption of CAVs is currently constrained by two main factors, i.e., the high cost of the required sensor equipment, such as high-resolution lidar (light detection and ranging) systems, and the lack of understanding of data-driven machine learning methods, which poses significant trust issues for critical applications e.g., autonomous driving. It is worth noting that the cost of a 64-channel HDL-64E Lidar, which is typically used for autonomous driving, is approximately $85,000$$ [2]. However, the large-scale use of such systems will become possible only if the costs associated with the incorporation of this technology are significantly reduced. Two of the main factors responsible, at present, for the increased costs are the expensive sensing equipment (i.e., high-resolution Lidar systems) and the need for processing devices with increased memory and computation capabilities.

To address these challenging issues, in literature several studies have investigated the use of low-cost sensors, e.g, a 16-channel Lidar, in conjunction with super-resolution algorithms to enhance the data obtained from the sensor, thus aiming to replace the expensive high-resolution Lidar sensors and reduce the overall cost of deploying the CAV technology. In more detail, two main categories of approaches have been explored to improve the performance of Lidar odometry. The first category includes the integration of additional sensors, such as visual cameras [3] or inertial measurement units (IMUs) [4], [5], or a combination of both IMUs and cameras [6]. The second category involves the application of appropriate restoration methods to the noisy or low-resolution lidar data. In most cases, a super-resolution algorithm based on deep learning is used, either after the initial computation of range images [7] or directly in the point cloud domain [8], [9].

Despite the efficacy of deep learning approaches, they necessitate copious amounts of diverse training data. Consequently, data acquired from individual autonomous vehicles may prove inadequate for optimizing these models [10]–[12]. One viable solution is to gather data from multiple entities and develop more compact neural network models in a centralized manner. However, this approach adopted by the above mentioned works, may be hindered by privacy concerns, which may prevent autonomous driving entities from sharing their datasets, even if they are interested in using their own data and similar private datasets from others [13]. In addition, data sharing can lead to significant communication overhead, making the collection and transfer of data from different locations to a central location both time consuming and costly, especially for large amounts of data [14]. Federated Learning (FL) as a secure and distributed methodology can effectively address the challenges associated with high-resolution lidar for autonomous driving. FL allows multiple clients to collaboratively learn a deep learning model without having to share their own private datasets, ensuring privacy [14]. The privacy-preserving and decentralized nature of FL makes it well-suited for deployment at the network edge, where each autonomous driving unit can be treated as a client that includes a lidar sensor for data collection [14], [15].

In contrast to the existing literature, which primarily focuses on centralized settings, this study presents a prototype implementation of federated learning that aims to address the challenge of LiDAR super-resolution in automotive scenes. To the best of our knowledge, this problem has not been studied previously from the federated learning perspective. Our approach leverages the distributed nature of federated learning to utilize private lidar data collected from different dispersed autonomous vehicles in various environmental conditions, e.g., rural and urban areas. The proposed framework allows us to derive a more robust and diverse model that can handle a wider range of environmental settings. To sum up, the key contributions of this paper are the following:

This work has received funding from the European Union's research and innovation programme TRUSTEE under grant agreement No 101070214.

979-8-3503-3959-8/23/$31.00 

- A novel federated learning-system is proposed that tackles LiDAR super-resolution in automotive scenarios, leveraging distributed data from autonomous vehicles leading to a robust and versatile model.
- Extensive experiments conducted on a real-world lidar odometry dataset emphasize the merits of the proposed LiDAR super-resolution method utilizing federated learning.

## II. Preliminaries

### A. Lidar super resolution

Consider a high resolution point cloud derived from a 64-channel lidar sensor, which results in a high-resolution range image $\boldsymbol{X} \in \mathbb{R}^{C\times M}$, where $C$ denotes the vertical resolution (i.e., the number of channels or lasers such as $C = 64$) and $M$ represents the horizontal resolution of the range image. Based on the degradation model in [7], a corresponding low-resolution range image $\boldsymbol{Y} \in \mathbb{R}^{c\times M}$ with the same horizontal resolution as $X$ but only $c < C$ channels in the vertical resolution (e.g., $c = 16$) can be derived as follows

$$\boldsymbol{Y} = \boldsymbol{S}\boldsymbol{X} + \boldsymbol{N} \tag{1}$$

where $\boldsymbol{S} \in \mathbb{R}^{c\times C}$ denotes the downsampling operator that selects only the c channels from the high resolution range image and $N$ is a zero-mean Gaussian noise term. Thus, it belongs to the category of the highly ill-posed inverse imaging problems [16]. The objective is to estimate the high-resolution range image $\boldsymbol{X}$ given the low-resolution range image $\boldsymbol{Y}$, so that

$$\boldsymbol{X} = \mathcal{F}(\boldsymbol{Y}) \tag{2}$$

where $\mathcal{F}(\cdot)$ denotes the mapping function that can be model using some deep learning network. The desired high-resolution point cloud is derived by transforming the estimated high-resolution range image into 3D coordinates (see, Figure 1).

It is imperative to emphasize that the development of a deep-learning model specific to the LiDAR super-resolution problem requires an extensive and heterogeneous dataset consisting of high- and low-resolution range images representing a variety of environments and scenarios. To address this problem, we developed a federated learning system that leverages information gathered from a set of autonomous vehicles operating in different locations and environmental conditions. This methodology facilitates the creation of a robust and comprehensive model while bypassing the need to transmit large amounts of data to a central server, thereby mitigating potential communication and privacy issues.

## III. Federated Learning Lidar Super resolution

To mathematically formulate the considered lidar super-resolution federated learning problem, we define a set of $\mathcal{N}$ edge devices, where each device $n \in \mathcal{N} = \{1, 2, \dots, N\}$ contains a local private dataset, denoted as $\mathcal{D}_n = \{\boldsymbol{X}_{i,n}, \boldsymbol{Y}_{i,n}\}_{i=1}^{p_n}$, where $x_n^i$ is the high-resolution range image, and $Y_n^i$ is the corresponding low-resolution range image.

Given $\mathcal{D}_n$, each device $n$ aims to train a local deep learning model, whose weights are denoted as $\theta_n$. This can be achieved by minimizing a local objective $g_n(\theta; \mathcal{D}_n)$ that utilizes some loss function, denoted as $\mathcal{L}(.)$. In particular, the local objective of device $n$ is:

$$g_n(\theta_n; \mathcal{D}_n) = \frac{1}{p_n} \sum_{i=1}^{p_n} \mathcal{L}(\boldsymbol{X}_{i,n}, \boldsymbol{Y}_{i,n}) \ , \tag{3}$$

Under the FL framework, the devices aim to collaboratively train a global model, say $\theta_g$, in a manner orchestrated by a central server. Particularly, the FL minimizes the aggregation of the local objectives and entails a common output for all devices using the global model. The objective of FL is:

$$G(\theta_g) = \sum_{n=1}^{N} w_n g_n(\theta_n; \mathcal{D}_n) \tag{4}$$

where $w_n$ denote some weight coefficients.

### A. Edge device-side

On the edge-device side, each autonomous vehicle $n$ utilizes its private dataset $\mathcal{D}_n$ to optimize a local deep learning models to solve the lidar super-resolution problem, based on the following loss function:

$$L(\vartheta_n) = \frac{1}{p_n} \sum_{i=1}^{p_n} |\boldsymbol{Y}_{i,n} - \mathcal{F}_{\vartheta_n}(\boldsymbol{X}_{i,n})|, \tag{5}$$

where $p_n$ is the size of the local dataset and $\mathcal{F}_{\vartheta_n}$ represents the local model at device $n$ whose weights are denoted as $\vartheta_n$.

The aim of the deep learning model is to approximate the mapping function in equation (2). The deep learning model used in this study employs an encoder-decoder architecture proposed in [7]. Specifically, the encoder part of the network includes convolutional blocks and average pooling layers to down-sample feature spatial resolutions and increase filter bands. The decoder portion of the network uses transposed convolutions to increase the spatial resolutions of the features and produce high-resolution output. After each convolution within the convolutional blocks, batch normalization and ReLU activation functions are used. The final high-resolution range image is generated by the output layer, which uses a single convolution filter without batch normalization. Figure 2 illustrates the considered model.

### B. Server-side

On the server-side, the primary objective is to compute a global model based on the local models received from participating edge autonomous agents. Specifically, the server aggregates the local models ${\vartheta_n}_{n=1}^{N}$ into a new global model denoted as $\vartheta_g$ by applying a weighted average fusion rule, i.e.,

$$\vartheta_g = \frac{1}{N} \sum_{n=1}^{N} w_n \vartheta_i^m, \tag{6}$$

(6), where $w_n$ represents the size of the local dataset of the $n$-th device. After aggregating the local models, the centralized server sends the new global model to all devices, which then use it as the initial point to update their local models using their own data via the training procedure described in equation (4). This process is repeated for $T$ communication rounds until the global model converges.

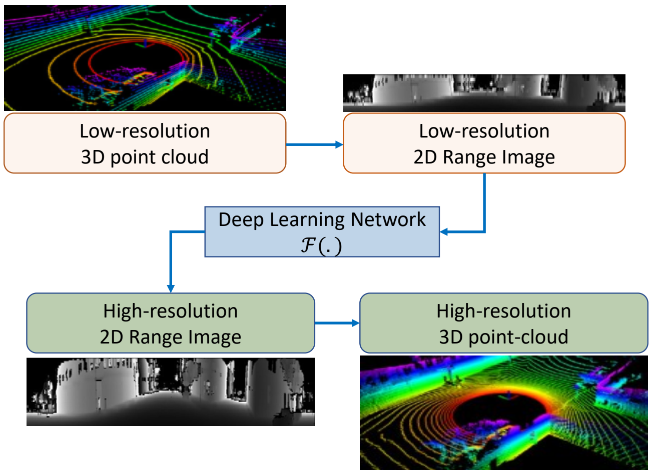


Fig. 1. Given a 3D point cloud from a 16-channel lidar, the considered framework involves projecting the low-resolution point cloud onto a 2D range image. This image serves as input for a deep learning model that estimates the high-resolution range image. The estimated image is then transformed back into 3D coordinates to generate the high-resolution point cloud.

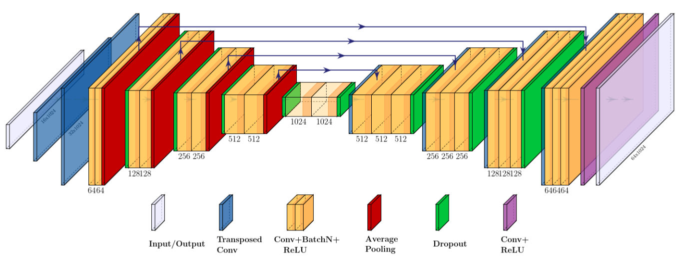


Fig. 2. The considered deep learning model employed from the edge autonomous vehicles for the proposed lidar super-resolution FL framework.

### C. Post processing method - Outlier Detection

A major problem of using convolutional operations on range images is the smoothing effect resulting in inaccurate object boundaries in the generated 3D lidar maps [7]. To tackle this problem, the Monte Carlo dropout (MC-dropout) [17] is employed to estimate the uncertainty of range predictions. In particular, the MC-dropout regularization approximates a Bayesian Neural Network (BNN) by performing multiple feed-forward passes with active the dropout layers during inference time, resulting in a distribution over outputs. Thus, given a low-resolution range image $X$, we derive the following expressions

$$\boldsymbol{Y}^* = \frac{1}{K}\sum_{k=1}^{K} \mathcal{F}_{\vartheta_g}(\boldsymbol{X})$$
$$\sigma(\boldsymbol{Y}) = \frac{1}{K}\sum_{k=1}^{K} (\mathcal{F}_{\vartheta_g}(\boldsymbol{X}) - \boldsymbol{Y}^*)^2 \tag{7}$$

where K defines the number of the forward passes, $\boldsymbol{Y}^*$ denotes the mean value of the estimated range image and $\sigma(Y)$ is the uncertainty. To remove the outliers i.e., points with high variance, a noise removal threshold, denoted as $\mu$ is introduced, scaling linearly with the predicted sensor range, to account for increasing noise levels with distance. Thus, the final output is given by

$$\boldsymbol{Y} = \begin{cases} \boldsymbol{Y}^*, & \text{if} \le \mu \boldsymbol{Y}^*. \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

Throughout the study, the scaling parameter was set 0.03 and an inference level of 50 were chosen for optimal results

## IV. Experimental Part

To validate the effectiveness of the proposed federated learning framework, a series of experiments were carried out on the real-world driving dataset named Ouster[1] . The objective was to upscale the data from a 16 to 64 channel lidar by a factor of 4. The experiments were conducted on the simulation framework that was developed in [18].

### A. Datasets

**Training Data**: Regarding the training, we employed the same dataset presented in study [7]. A 64-channel lidar, OS-1-64, was simulated in the CARLA Town 2 scene, matching the Ouster dataset field of view (33.2°). For the same scene, a 16-channel lidar, OS-1-16, was simulated to generate low-resolution point clouds. Both high- and low-resolution point clouds were projected onto range images, resulting in 7000 pairs of 64x1024 and 16x1024 images. The images were then normalized to a range of 0-1 for training. Note that Town 2 scene contains a variety of environmental settings, resulting in a rich dataset that simulates the diverse experiences clients may encounter. **Testing Data**: To validate the performance of the proposed FL architecture, the real-world Ouster lidar dataset was utilized. This dataset comprises 8825 scans collected over a 15-minute drive in San Francisco using an OS-1-64 3D lidar sensor. The high-resolution point clouds were converted into 64x1024 range images, and 16 rows were extracted to create 16x1024 low-resolution images. These data pairs were utilized to assess the architecture's performance in recovering high-resolution 3D point clouds from low-resolution inputs.

### B. Implementation Details

**Federated learning scenario**: We analyzed a network comprising of 5 autonomous vehicles (nodes). Consequently, we partitioned the aforementioned training data into 5 distinct blocks, with each block representing the local dataset of an individual client. During the local training on the edge devices, we employed 5 epochs, along with a learning rate and batch size of 1e-04 and 6, respectively. Moreover, the communication rounds between the central server and the edge devices were set to T = 50 after we thoroughly explored the parameter space to determine their optimality. The local models were trained using the Adam optimizer.

**Compared Methods**: We compared the proposed FL approach with: (i) *Centralized scheme*: Conventional approach where a central server gathers all available data from distributed edge devices to train the lidar super-resolution model. (ii) *Individual scheme*: When a single client works alone with limited data, without participating in the federated learning process. For the training process of the above scenarios, we used 50 epochs with learning rate equal to $1e-04$ utilizing the Adam optimizer.

[1]https://github.com/ouster-lidar/ouster_example

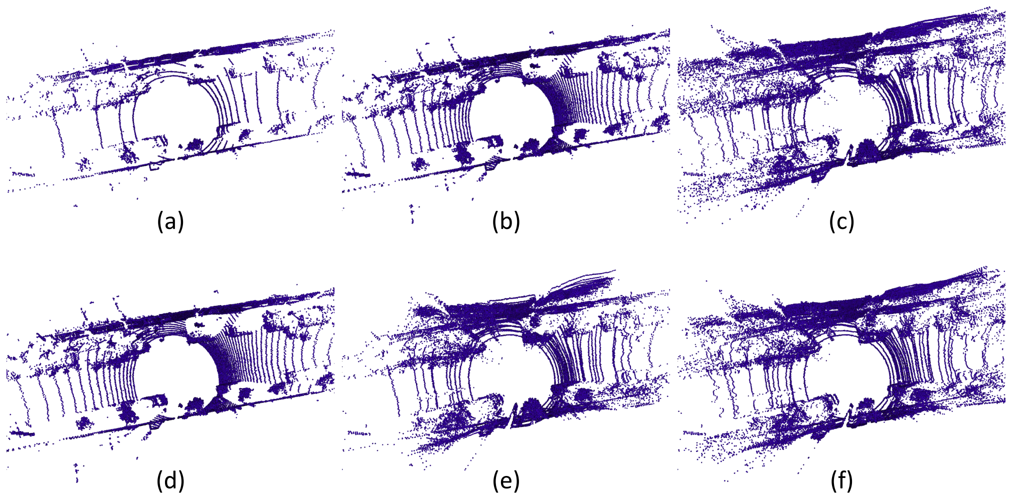


Fig. 3. 3D points clouds: (a) lidar-16, (b) FL-scheme with Monte Carlo dropout, (c) FL-scheme without Monte Carlo dropout, (d) ground truth lidar-64 (e) Centralized scheme and (f) Individual learning scheme.

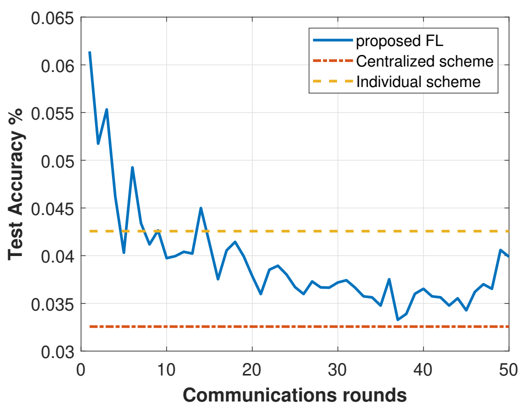


Fig. 4. Loss of the derived global model from the proposed FL scheme vs communication rounds along with the best accuracy achieved by the centralized and individual training schemes.

### C. Federated Learning - Lidar super-resolution performance on raw data

In this section, a comparison is performed between the proposed FL method and the other two centralized approaches. Table I summarizes the quantitative results comparing the reconstructed high resolution range images with the corresponding ground truth images based on the L1 loss. Additionally, Figure 4 illustrates the convergence of the global model obtained from the FL scenario, along with a comparison to the best accuracy achieved by the centralized and individual training schemes. It is evident that the proposed FL scenario is able to achieve competitive performance against the centralized solution. Although the centralized scheme performed better in terms of quantitative results, it requires the sharing of massive amounts of data, thus introducing a significant burden on the communication links between the edge devices and the central server, and also privacy concerns. Moreover, the proposed federated learning method provides a significant advantage over the individual training scheme when clients seek to train their model using only their private data. While a client may continuously gather new data, this data may be limited in scope and diversity, potentially resulting in a

TABLE I. QUANTITATIVE RESULTS.

| Dataset | Method | Data training size | L1 loss |
|---|---|---|---|
| Ouster | FL | 700 per client | 0.0357 |
| | Centralized scheme | 7000 | 0.0336 |
| | Individual scheme | 700 | 0.0427 |

less accurate and less generalizable model. The FL approach, however, enables clients to obtain more accurate models by leveraging information from diverse datasets obtained from various autonomous vehicles operating in different locations and environmental conditions. Additionally, Figure 3 illustrates the reconstructed point clouds derived from the considered approaches along with the ground truth 64-lidar and and the 16-lidar point clouds, providing strong evidences about the applicability of the proposed FL system in autonomous driving problems.

## V. CONCLUSION

In this study, the problem of lidar super-resolution was studied under a federated learning perspective. The federated learning approach demonstrated competitive performance compared to the centralized solution while mitigating privacy and communication overhead concerns. By using federated learning, clients were able to leverage different data sets from multiple sources, improving model accuracy. Future research directions should focus on extending privacy-preserving nature of federated learning including methods such as secure differential privacy, or homomorphic encryption. In addition, conducting experiments in a variety of more complex environments and weather conditions will provide further valuable insight into the benefits of the proposed federated learning methodology for lidar super-resolution.


## ACKNOWLEDGEMENT

This work has received funding from the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 101092875 - DIDYMOS-XR: Digital DynaMic and responsible twinS for XR.

# References


[1] S. Wang, C. Li, D. W. K. Ng, Y. C. Eldar, H. V. Poor, Q. Hao, and C. Xu, "Federated deep learning meets autonomous vehicle perception: Design and verification," *IEEE Network*, pp. 1–10, 2022.

[2] J. Wu, H. Xu, and J. Zhao, "Automatic lane identification using the roadside lidar sensors," *IEEE Intelligent Transportation Systems Magazine*, vol. 12, no. 1, pp. 25–34, 2020.

[3] R. Ishikawa, T. Oishi, and K. Ikeuchi, "Lidar and camera calibration using motions estimated by sensor fusion odometry," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 7342–7349.

[4] J. Zhang and S. Singh, "Visual-lidar odometry and mapping: low-drift, robust, and fast," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 2174–2181.

[5] H. Ye, Y. Chen, and M. Liu, "Tightly coupled 3d lidar inertial odometry and mapping," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 3144–3150.

[6] X. Zuo, Y. Yang, P. Geneva, J. Lv, Y. Liu, G. Huang, and M. Pollefeys, "Lic-fusion 2.0: Lidar-inertial-camera odometry with sliding-window plane-feature tracking," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 5112–5119.

[7] T. Shan, J. Wang, F. Chen, P. Szenher, and B. Englot, "Simulation-based lidar super-resolution for ground vehicles," *Robotics and Autonomous Systems*, vol. 134, p. 103647, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0921889020304875

[8] J. Yue, W. Wen, J. Han, and L.-T. Hsu, "3d point clouds data super resolution-aided lidar odometry for vehicular positioning in urban canyons," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 5, pp. 4098–4112, 2021.

[9] D. Tian, D. Zhao, D. Cheng, and J. Zhang, "Lidar super-resolution based on segmentation and geometric analysis," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–17, 2022.

[10] S. Wang, C. Li, D. W. K. Ng, Y. C. Eldar, H. V. Poor, Q. Hao, and C. Xu, "Federated deep learning meets autonomous vehicle perception: Design and verification," *IEEE Network*, pp. 1–10, 2022.

[11] D. Jallepalli, N. C. Ravikumar, P. V. Badarinath, S. Uchil, and M. A. Suresh, "Federated learning for object detection in autonomous vehicles," in *2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService)*, 2021, pp. 107–114.

[12] Z. Du, C. Wu, T. Yoshinaga, K.-L. A. Yau, Y. Ji, and J. Li, "Federated learning for vehicular internet of things: Recent advances and open issues," *IEEE Open Journal of the Computer Society*, vol. 1, pp. 45–61, 2020.

[13] T. Gafni, N. Shlezinger, K. Cohen, Y. C. Eldar, and H. V. Poor, "Federated learning: A signal processing perspective," *IEEE Signal Processing Magazine*, vol. 39, no. 3, pp. 14–41, 2022.

[14] S. Savazzi, M. Nicoli, M. Bennis, S. Kianoush, and L. Barbieri, "Opportunities of federated learning in connected, cooperative, and automated industrial systems," *IEEE Communications Magazine*, vol. 59, no. 2, pp. 16–21, 2021.

[15] Y. Li, X. Tao, X. Zhang, J. Liu, and J. Xu, "Privacy-preserved federated learning for autonomous driving," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 8423–8434, 2022.

[16] A. Gkillas, D. Ampeliotis, and K. Berberidis, "Connections between deep equilibrium and sparse representation models with application to hyperspectral image denoising," *IEEE Transactions on Image Processing*, vol. 32, pp. 1513–1528, 2023.

[17] Y. Gal and Z. Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning," in *international conference on machine learning*. PMLR, 2016, pp. 1050–1059.

[18] C. Anagnostopoulos, C. Koulamas, A. Lalos, and C. Stylios, "Open-source integrated simulation framework for cooperative autonomous vehicles," in *2022 11th Mediterranean Conference on Embedded Computing (MECO)*, 2022, pp. 1–4.