AN EFFICIENT DEEP UNROLLING SUPER-RESOLUTION NETWORK FOR LIDAR AUTOMOTIVE SCENES

Alexandros Gkillas[1,2], *Aris S. Lalos*[1], *Dimitris Ampeliotis*[1,3]

[1]Industrial Systems Institute, Athena Research Center, Patras Science Park, Greece
[2]AviSense.AI, Patras Science Park Building, Greece
[3]Digital Media and Communication Department, Ionian University, Greece

ABSTRACT

Considering the high cost of high-resolution Lidar sensors, in this work, a novel Lidar super-resolution method is proposed to improve the performance on numerous autonomous vehicle perception tasks, including that of a Lidar odometer. Specifically, we propose a regularized optimization problem employing a learnable regularizer (neural network) to capture the properties of the data. To efficiently solve this problem, a deep unrolling methodology is proposed, thus forming an interpretable and well-justified deep architecture. Extensive experiments on a real-world lidar odometry application highlight that the proposed model exhibits both superior performance as well as a significantly reduced number of trainable parameters i.e., $\mathbf{99.75}\%$ less parameters, as compared to other deep learning methods. The source code used for this work can be found at our repository: repository.

***Index Terms*—** super-resolution, lidar, deep unrolling, SLAM, lidar odometry

1. INTRODUCTION

Connected and Autonomous Vehicles (CAVs) constitute one of the most prominent application areas for the recent advancements of machine learning methods, offering the possibility to revolutionize the modern way of living [1]. However, the broad use of such systems will become possible only if several concerns are alleviated. Two of the main reasons that restrict the large-scale use of such systems are the expensive sensing equipment required (i.e., high-resolution light detection and ranging - Lidar - systems) and the fact that the operation of data-driven machine learning methods is not well understood, something that raises significant trust issues especially in critical applications such as autonomous driving. In order to have an idea about the cost of the sensors, the cost of a 64-channel HDL-64E Lidar is approximately $85,000\$$ [2].

With the ultimate goal of alleviating the concerns associated with the incorporation of the technology that enables CAVs by the car industry, this work proposes a method to replace the high-resolution lidar device by a low-resolution one. An efficient super-resolution method is developed to enrich the data coming from the low-cost sensor. This method is built upon one of the latest advancements in the area of deep learning, namely deep unrolling [3–7], that offers the possibility to combine deep learning methods with analytical optimization-based methods, thus offering the advantages of both worlds: remarkable performance, low-computational complexity and explainable model architecture [8].

In the literature, several approaches that try to improve the performance of Lidar odometry have been studied. In general, these approaches can be categorized into two categories. In the first category, methods that integrate additional sensors, such as visual cameras [9] or inertial measurement units (IMUs) [10], [11] or both IMUs and cameras [12], can be found. In the second category, some suitable restoration method is applied to the noisy/low-resolution lidar data. In most cases, a deep-learning based super-resolution algorithm is applied, either after first computing the respective range-images [13], or directly in the point-cloud domain [14, 15].

In this work, we develop a method that improves the performance of Lidar-based odometry, without resorting to the help from other sensors. Particularly, we propose a super-resolution method that can be applied to range images coming from a 16-channel Lidar system, to effectively predict range-measurements as if a 64-channel Lidar system was used. Different from the existing studies in literature, the proposed super-resolution method is not entirely based on deep-learning. In more detail, a novel well-justified regularized optimization problem is proposed incorporating in the cost-function a learnable regularizer in the from of a neural network. Utilizing the half quadratic splitting (HQS) method, an iterative solver is derived which can be leveraged by employing the deep unrolling framework [3, 8, 16], where a small number of iterations are unrolled to form a highly interpretable deep learning network. The proposed method notably outperforms other deep learning approaches, exhibiting superior results in real-world lidar odometry. A benefit that stems out of the proposed methodology is the fact that the deep unrolling network contains $\mathbf{99.75}\%$ less parameters compared to other state-of-the-art deep learning networks.

2. PROBLEM FORMULATION

Consider a high-resolution point cloud derived from a 64-channel lidar sensor and the respective high-resolution range image $\boldsymbol{X} \in \mathbb{R}^{C\times M}$, where C represents the vertical resolution (i.e., the number of channels or lasers such as $C = 64$) and M denotes the horizontal resolution of the range image. Mathematically, the 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 based on the following degradation model [13]

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

where $\boldsymbol{S} \in \mathbb{R}^{c\times C}$ denotes the downsampler operator that selects only the c channels from the high resolution range image and N is a zero-mean Gaussian noise term. Given the low-resolution range image $\boldsymbol{Y}$, our goal is to effectively estimate the corresponding high resolution range image $\boldsymbol{X}$. Thus, we propose the following regularized optimization problem

$$\arg\min_{\boldsymbol{X}} \ \frac{1}{2}\left\|\boldsymbol{Y} - \boldsymbol{S}\boldsymbol{X}\right\|_F^2 + \mu\mathcal{R}(\boldsymbol{X}) \,, \tag{2}$$

consisting of a data fidelity term, and an unknown learnable regularizer $\mathcal{R}(\cdot)$ that enforces some prior on the solution, trained to capture the structure of the range images. Also, μ denotes a regularization parameter. The high-resolution point cloud is derived by transforming the estimated high-resolution range image into 3D coordinates (see, Figure 1).

3. PROPOSED METHOD

3.1. HQS iterative solver

To effectively solve (2), the Half Quadratic Splitting (HQS) methodology [17] was employed. By introducing an auxiliary variable $Z \in \mathbb{R}^{C\times M}$ problem (2), can be written as a follows

$$\arg\min_{\boldsymbol{X}} \ \frac{1}{2}\left\|\boldsymbol{Y} - \boldsymbol{S}\boldsymbol{X}\right\|_F^2 + \mu\mathcal{R}(\boldsymbol{Z}) \quad s.t. \quad \boldsymbol{Z} - \boldsymbol{X} = 0,$$

The loss function that HQS aims to minimize is

$$\mathcal{L} = \frac{1}{2}\left\|\boldsymbol{Y} - \boldsymbol{S}\boldsymbol{X}\right\|_F^2 + \mu\mathcal{R}(\boldsymbol{Z}) + \frac{b}{2}\left\|\boldsymbol{Z} - \boldsymbol{X}\right\|_F^2 \tag{3}$$

where b is a penalty parameter. Based on equation (3) a sequence of individual sub-problems emerges:

$$\boldsymbol{X}^{(k+1)} = \arg\min_{\boldsymbol{X}} \frac{1}{2}\left\|\boldsymbol{Y} - \boldsymbol{S}\boldsymbol{X}\right\|_F^2 + \frac{b}{2}\left\|\boldsymbol{Z}^{(k)} - \boldsymbol{X}\right\|_F^2 \tag{4a}$$

$$\boldsymbol{Z}^{(k+1)} = \arg\min_{\boldsymbol{Z}} \mu\mathcal{R}(\boldsymbol{Z}) + \frac{b}{2}\left\|\boldsymbol{Z} - \boldsymbol{X}^{(k+1)}\right\|_F^2 . \tag{4b}$$

The sub-problem in (4a) constitutes a quadratic regularized least squares problem. The closed form solution is given as

$$\boldsymbol{X}^{(k+1)} = (\boldsymbol{S}^T\boldsymbol{S} + b\boldsymbol{I})^{-1}(\boldsymbol{S}^T\boldsymbol{Y} + b\boldsymbol{Z}^{(k)}) \,. \tag{5}$$

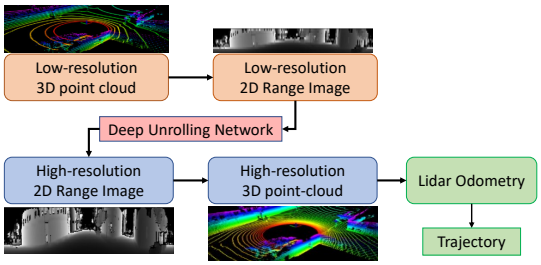


Fig. 1: Given a low-resolution 3D point cloud derived from a 16-channel lidar, it is project onto a 2D low-resolution range image. This image is provided as input to the proposed deep unrolling model to estimate the corresponding high-resolution range image. The estimated image is then transformed into 3D coordinates, thus producing the desired high resolution point cloud for the lidar odometry problem.7.

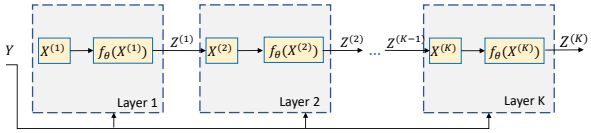


Fig. 2: The proposed end-to-end Deep unrolling model.

Moreover, from a Bayesian perspective, sub-problem (4b) corresponds to a Gaussian denoiser [18]. Hence, a neural network $f_\theta(\cdot)$ can be employed to act as denoiser, and learned from training data. Thus, equation (4b) can be written as

$$\boldsymbol{Z}^{(k+1)} = f_\theta(\boldsymbol{X}^{(k+1)}) \tag{6}$$

Thus, the HQS solver consists of the following update rules:

$$\boldsymbol{X}^{(k+1)} = (\boldsymbol{S}^T\boldsymbol{S} + b\boldsymbol{I})^{-1}(\boldsymbol{S}^T\boldsymbol{Y} + b\boldsymbol{Z}^{(k)}) \tag{7a}$$

$$\boldsymbol{Z}^{(k+1)} = f_\theta(\boldsymbol{X}^{(k+1)}) \tag{7b}$$

Note that the neural-network f_θ can be pre-trained employing pairs of generated noisy high resolution range images (using Gaussian noise, denoted as $\boldsymbol{N}$) and their corresponding ground truth versions denoted as $\{\boldsymbol{X}^i + N, \boldsymbol{X}^i\}_{i=1}^p$ via the loss function $\sum_{i=1}^p \left\|f_\theta(\boldsymbol{X}^i + N;\theta) - \boldsymbol{X}^i\right\|_F^2$.

3.2. Interpretable deep unrolling-based model

In this section an efficient methodology is proposed based on the deep unrolling framework. In particular, a small number of iterations, say K, of the proposed HQS solver in (7) are unrolled and treated as a deep learning architecture. Each iteration of the iterative solver is considered a unique layer of the proposed model, resulting in a K-layer deep learning architecture. In light of this, the parameters and the depth of the derived model are highly **interpretable** as they have direct correspondence to the underlying iterative algorithm. Figure 2 demonstrates the proposed deep unrolling network. A great benefit that stems from this strategy is the fact that the

learnable parameters of the proposed model (i.e., the neural network $f_\theta(\cdot)$ and the penalty parameter b) can be trained end-to-end by minimizing some loss function, as follows

$$l(\theta) = \sum_{i=1}^{p} \left\| \boldsymbol{Z}_i^{(K)} - X_i \right\|_F^2. \quad (8)$$

where $\boldsymbol{Z}_i^{(K)}$ is the output of the proposed deep unrolling network given a low-resolution range image $\boldsymbol{Y}_i$ and $\boldsymbol{X}_i$ denotes the i^{th} ground truth high resolution range image.

4. EXPERIMENTAL PART

To demonstrate the effectiveness of the proposed deep unrolling model, a series of experiments were conducted on a real-world driving dataset, referred as Ouster [1] aiming to perform a 4x upscaling from a 16 to 64 channel lidar. More importantly, we evaluated the benefits of the proposed method on a lidar SLAM system based on the LeGO-LOAM algorithm [19]. The lidar odometry experiments were performed on the developed simulation framework [20].

4.1. Datasets

Training data: Concerning the training data, we employed the same dataset provided by study [13]. In more details, since the Ouster dataset contains a lidar sensor with $33, 2^o$ FOV (field of view), a 64-channel lidar, OS-1-64 was simulated in CARLA Town 2 scene. The corresponding low-resolution point clouds were derived by simulating a 16-channel OS-1-16 lidar for the exact same scene. Having the high and low-resolution point clouds, we then projected the data onto range images, thus resulting pairs of ground truth high and low-resolution range images with resolution 64×1024 and 16×1024 respectively.

Testing data: To validate the performance of the proposed architecture, the real-world Ouster lidar dataset was utilized, which consists of 8825 scans over a span of 15 min of driving in San Francisco using an OS-1-64 3D lidar sensor. This 64-channel sensor provides us the ground truth high resolution point clouds for validation. Following, the experimental set-up of study [13], to generate the low-resolution range images, we first projected the high-resolution point clouds onto high-resolution range images (64×1024) and then we extracted 16 rows from the high-res range images, thus forming low-res range images with a resolution of 16×1024.

4.2. Implementation details

Neural Network Architecture:Regarding, the neural network $f_\theta(\cdot)$, we used a 5-layer CNN network, where each layer consists of 64 filters with size 3×3. Moreover, we employed the ReLU as activation and drop-out rate equal to 0.05. **Proposed Model-Parameter Setting**: Concerning the proposed lidar super-resolution model, the number of unrolling iterations was to $K = 6$, thus leading to 6-layer deep network architecture. During the end-to-end training procedure, we employed the Adam optimizer with batch size equal to 6. Furthermore, the learning rate was set to $1e-03$ and the number of epochs to 100. Finally, note that at inference time, similar to study [13], we employed a Monte-Carlo drop-out method [21] as a post-processing technique to remove unreliable range predictions.

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

Table 1: Quantitative results.

Dataset	Method	L1 loss	Number of parameters (milions)
Ouster	Linear	0.0324	-
	Cubic	0.0467	-
	SR-ResNet [22]	0.0231	$35M$
	SR-AE [13]	0.0214	$30M$
	Proposed	**0.0208**	$\mathbf{0.1M}$

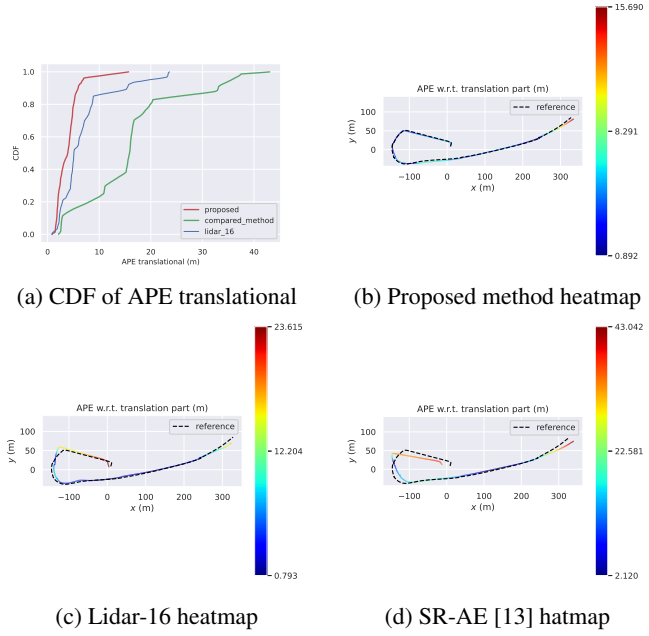
4.3. Lidar super-resolution performance on raw data

In this section, a comparison is performed between the proposed deep unrolling method and several methodologies including the baseline linear and cubic interpolation approaches, the state-of-the-art/well-established super-resolution SR-ResNet [22] model in classic image processing and the state-of-the-art lidar-based super-resolution SR-AE [13] model. Table 1 summarizes the reconstruction results. It is evident that the proposed deep unrolling method is able to provide better results in terms of L1-loss in comparison with the other approaches. Apart from the quantitative gains, the proposed method requires remarkably less parameters as compared to the deep learning methodologies, due to its well-defined architecture, derived from the proposed optimization problem in (2). *In particular, the proposed model contains* $\mathbf{99.75}\%$ *less parameters as compared to the SR-Resnet and SR-AE approaches, thus rendering it suitable for real-world applications with computational and storage limitations.*

4.4. Lidar super-resolution on lidar odometry

To fully validate the merits of the proposed model, we examined an real-world application, evaluating the performance of the LeGO-LOAM [19], a Lidar Odometry (LO) that provides real time six degree-of-freedom pose estimation. The experiments were conducted by utilizing the Ouster dataset. From the considered dataset, we used two sequences, i.e., (i) the first 2600 consecutive scans made while the vehicle is moving along a simple trajectory and (ii) the 6000 scans corresponding to a more challenging trajectory with short close loops.

The goal of the specific LO is to provide real-time six degree-of-freedom pose estimation for ground vehicles equipped with 3D lidar. This is achieved by extracting planar



(a) CDF of APE translational (b) Proposed method heatmap

(c) Lidar-16 heatmap (d) SR-AE [13] hatmap

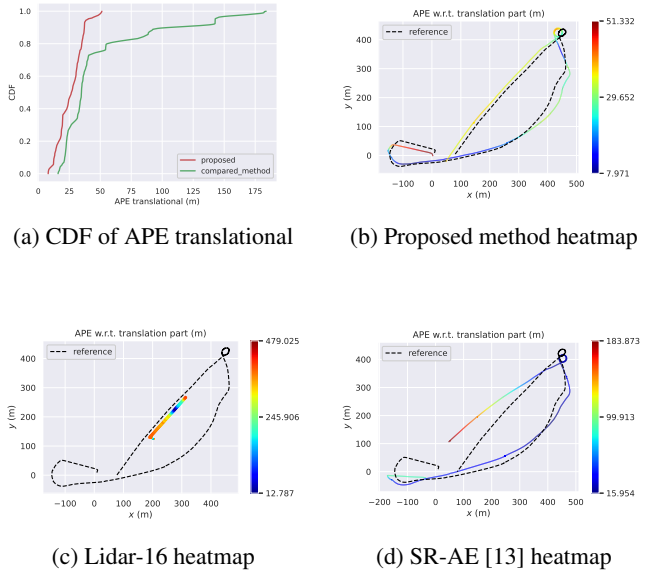
Fig. 3: CDF and heatmaps for the proposed method, the 16-channel Lidar and the method SR-AE [13] using as reference trajectory the path derived from the lidar with 64 channels.

and edge feature and then use the planar and edge features to solve different components of the six degree-of-freedom transformation across consecutive scans. To study the impact of the SR approach in such a LO, we compare the performance of the Lego Loam algorithm using as input (i) the reconstructed high-resolution 3D point clouds derived from the proposed method, (ii) the low resolution 3D point clouds generated by a 16-channel lidar sensor and (iii) the reconstructed point clouds derived from the SR-AE method [19]. By adopting the error metrics from study [23, 24], the output metrics along with the CDF (cumulative distribution function) of APE translational and trajectory heatmaps, using as reference pose (trajectory) the path derived from the lidar with 64 channels are depicted on Table 2 and Figures 3 and 4.

From Figure 3 and 4, we can deduce that the proposed method consistently outperforms the 16-channel lidar and the SR-AE [13] for the both considered trajectories, thus verifying that the reconstructed 64-channel lidar data derived from the proposed model are notably more accurate compared to the other state-of-the-art SR-AE [13] method. Additionally, although the 16-channel lidar provides satisfactory results during the simple trajectory (2600 scans, see Figure 3), in the more challenging trajectory (6000 scans), which contains close loops, the 16-channel lidar is not able to follow the reference trajectory (see Figure 4). This can be justified by the fact that the LeGO-LOAM method relies on the availability of the edge and planar features to estimate the vehicle transformation, and thus it fails to generate robust features

Table 2: Lidar Odometry: Absolute Pose Error w.r.t translation part (m).

Ouster Dataset	Method	mean	rmse	max
2600 scans	Lidar-16	6.84	8.52	23.61
	SR-AE [13]	15.97	18.52	43.04
	Proposed	**3.86**	**4.42**	**15.69**
6000 scans	Lidar-16	299.90	316.82	479.02
	SR-AE [13]	48.29	61.73	183.42
	Proposed	**26.06**	**27.85**	**51.33**



(a) CDF of APE translational (b) Proposed method heatmap

(c) Lidar-16 heatmap (d) SR-AE [13] heatmap

Fig. 4: CDF and heatmaps for the proposed method, the 16-channel Lidar and the method SR-AE [13] using as reference trajectory the path derived from the lidar with 64 channels.

from the sparse point cloud derived from the 16-channel lidar. The above remarks, provide strong evidences that the proposed method can be an effective and cost-efficient solution for real-world applications.

5. CONCLUSIONS

In this study, the problem of super-resolution on lidar data was considered. Taking into account the under-determined nature of the examined problem, a new optimization problem was proposed employing a learnable regularizer (CNN) to capture the inherent structure of the data. In light of this, an interpretable and concise deep unrolling model was developed to solve efficiently the proposed problem, containing $\mathbf{99.75}\%$ less parameters compared to other state-of-the-art deep learning networks. Extensive experiments on real-world lidar odometry highlighted its efficacy and applicability.

6. 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] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Processing Magazine*, vol. 38, no. 2, pp. 18–44, 2021.

[4] Y. Yang, J. Sun, H. Li, and Z. Xu, "Admm-csnet: A deep learning approach for image compressive sensing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 3, pp. 521–538, 2020.

[5] I. Y. Chun, Z. Huang, H. Lim, and J. Fessler, "Momentum-net: Fast and convergent iterative neural network for inverse problems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–1, 2020.

[6] Y. Li, M. Tofighi, J. Geng, V. Monga, and Y. C. Eldar, "Efficient and interpretable deep blind image deblurring via algorithm unrolling," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 666–681, 2020.

[7] Y. Yang, J. Sun, H. Li, and Z. Xu, "Admm-csnet: A deep learning approach for image compressive sensing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, no. 3, pp. 521–538, 2020.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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

[14] 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.

[15] 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.

[16] A. Gkillas, D. Ampeliotis, and K. Berberidis, "A highly interpretable deep equilibrium network for hyperspectral image deconvolution," in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1–5.

[17] D. Geman and C. Yang, "Nonlinear image recovery with half-quadratic regularization," *IEEE Transactions on Image Processing*, vol. 4, no. 7, pp. 932–946, 1995.

[18] K. Zhang, W. Zuo, S. Gu, and L. Zhang, "Learning deep cnn denoiser prior for image restoration," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

[19] T. Shan and B. Englot, "Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018, pp. 4758–4765.

[20] 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.

[21] Y. Gal and Z. Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning," in *Proceedings of The 33rd International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, M. F. Balcan and K. Q. Weinberger, Eds., vol. 48. New York, New York, USA: PMLR, 20–22 Jun 2016, pp. 1050–1059. [Online]. Available: https://proceedings.mlr.press/v48/gal16.html

[22] C. Ledig, L. Theis, F. Huszár, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, and W. Shi, "Photo-realistic single image super-resolution using a generative adversarial network," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 105–114.

[23] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of rgb-d slam systems," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012, pp. 573–580.

[24] N. Piperigkos, A. S. Lalos, and K. Berberidis, "Graph laplacian diffusion localization of connected and automated vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 8, pp. 12 176–12 190, 2022.