


Toward Physics-Aware Deep Learning Architectures for LiDAR Intensity Simulation

Vivek Anand¹^a, Bharat Lohani¹^b, Gaurav Pandey²^c and Rakesh Mishra²^d

¹Geoinformatics, Department of Civil Engineering, Indian Institute of Technology Kanpur, India

²Department of Engineering Technology & Industrial Distribution (ETID), Texas A&M University, U.S.A.

³Geodesy and Geomatics Engineering, University of New Brunswick, Canada

Keywords: Simulation, LiDAR, LiDAR Intensity, Autonomous Vehicle, Deep Learning.


Abstract: Autonomous vehicles (AVs) heavily rely on LiDAR perception for environment understanding and navigation. LiDAR intensity provides valuable information about the reflected laser signals and plays a crucial role in enhancing the perception capabilities of AVs. However, accurately simulating LiDAR intensity remains a challenge due to the unavailability of material properties of the objects in the environment, and complex interactions between the laser beam and the environment. The proposed method aims to improve the accuracy of intensity simulation by incorporating physics-based modalities within the deep learning framework. One of the key entities that captures the interaction between the laser beam and the objects is the angle of incidence. In this work, we demonstrate that adding the LiDAR incidence angle as a separate input modality to the deep neural networks significantly enhances the results. We integrated this novel input modality into two prominent deep learning architectures: U-NET, a Convolutional Neural Network (CNN), and Pix2Pix, a Generative Adversarial Network (GAN). We investigated these two architectures for the intensity prediction task and used SemanticKITTI and VoxelScape datasets for experiments. The comprehensive analysis reveals that both architectures benefit from the incidence angle as an additional input. Moreover, the Pix2Pix architecture outperforms U-NET, especially when the incidence angle is incorporated.


1 INTRODUCTION


Autonomous vehicles (AVs) have emerged as a transformative technology that promises to revolutionize transportation systems worldwide (Schwartz et al., 2018). These vehicles, equipped with advanced sensors and intelligent algorithms, are capable of navigating and making decisions without human intervention. One of the key sensors used in AVs is LiDAR (Light Detection and Ranging), which plays a crucial role in perceiving the surrounding environment and ensuring safe and efficient autonomous navigation (Royo and Ballesta-Garcia, 2019) (Li and Ibanez-Guzman, 2020). These sensors work on the principle of emitting laser beams and measuring the time it takes for the laser pulses to return after reflecting off objects in the scene (Li and Ibanez-Guzman,


2020). In addition to distance measurements, LiDAR sensors can also capture other attributes, such as intensity, which refers to the strength of the reflected laser signal. LiDAR intensity carries valuable information about the characteristics of objects and surfaces, making it an important parameter in various applications, including object recognition, scene understanding, obstacle detection, classification, and semantic segmentation (Wang et al., 2018) (Wang et al., 2019) (Yang et al., 2018a) (Yang et al., 2018b) (Wang and Shen, 2017) (Meyer et al., 2019) (Liang et al., 2020).

To train and validate LiDAR perception algorithms, large-scale datasets with ground truth intensity information are required. However, collecting data in the real world is a very expensive and time-consuming task, hence, simulation turns out to be the most promising alternative (Yang et al., 2021). Currently, LiDAR simulation methods heavily rely on physics-based models and assumptions about material properties and surface interactions. However, accurately simulating LiDAR intensity using these meth-

^a <https://orcid.org/0000-0001-7503-5484>

^b <https://orcid.org/0000-0001-8589-192X>

^c <https://orcid.org/0000-0002-4838-802X>

^d <https://orcid.org/0000-0001-6856-4396>

ods is a complex and computationally expensive task as it presents significant challenges due to the dependence on multiple factors, including incidence angle, material properties, and surface interactions among others (Dai et al., 2022)(Vacek et al., 2021). Among these factors, the incidence angle plays a vital role in determining the intensity of the reflected laser signal (Dai et al., 2022) (Tatoglu and Pochiraju, 2012). As the angle between the incident laser beam and the surface normal changes, the intensity observed by the LiDAR sensor also changes.

The physics-based approaches and learning-based approaches have quite complimentary challenges. For example, physics-based approaches require that the material property of the environmental objects is known whereas learning-based methods learn the material properties directly from the data but they struggle to capture the intricate interactions between laser pulse and the object surfaces accurately. These limitations hinders the realism and fidelity of simulated LiDAR data, impeding the development and evaluation of perception algorithms for autonomous vehicles. To overcome these challenges, we propose a hybrid learning-based approach that not only leverages the power of deep learning algorithms but also captures the complex interactions between the laser pulse and the object surface more effectively. By incorporating the incidence angle into the learning model, it becomes possible to leverage this contextual information and improve the realism and accuracy of LiDAR intensity simulation.

In this paper, we conduct a comparative study of two different deep learning architectures, namely U-NET(Ronneberger et al., 2015) and Pix2Pix(Isola et al., 2017a), for LiDAR intensity simulation. The novel methodology we propose involves incorporating the incidence angle as an input modality, enhancing both architecture's performance. We conduct extensive experiments using SemanticKITTI data (Behley et al., 2019), which is collected in the real world, and VoxelScape data (Saleh et al., 2023), which is simulated using physics-based methods, exploring both established and unexplored methods in the field. The results showcase the importance of the incidence angle in accurately predicting LiDAR intensity and demonstrate that the Pix2Pix architecture, a new addition to the literature for this task, outperforms the existing U-NET approach. Our findings pave the way for more realistic and reliable LiDAR perception in AVs, offering valuable insights into the comparative strengths and weaknesses of different deep learning methodologies in LiDAR intensity simulation.

2 RELATED WORK

2.1 Simulation for Autonomous Vehicles

Simulation plays a vital role in the development and evaluation of autonomous vehicles, particularly in mimicking real-world driving conditions, enabling comprehensive testing of perception algorithms and the assessment of their performance(Yang et al., 2021). It provides a controlled and repeatable environment for algorithm development and evaluation, allowing researchers to precisely control variables and compare different approaches. Additionally, simulation overcomes limitations in data availability of different sensors used in autonomous vehicles including LiDAR by generating large-scale datasets with diverse scenarios, supporting algorithm training and validation (Geiger et al., 2012) (Caesar et al., 2020) (Sun et al., 2020).

2.2 Physics-Based LiDAR Simulation

Physics-based LiDAR simulation is a widely used approach for generating synthetic LiDAR data to test and evaluate autonomous driving systems(Yang et al., 2021). This method involves modeling the physical processes of laser beam emission, propagation, and interaction with objects in the environment(Elmqvist and Negrut, 2020). By considering factors such as surface reflectance, material properties, and light scattering effects, physics-based simulations aim to accurately replicate the behavior of real-world LiDAR sensors. However, simulating LiDAR intensity poses significant challenges due to the complex nature of interactions between the laser beam and the objects in the scene(Vacek et al., 2021). Precisely modeling these interactions and accounting for various environmental factors are extremely difficult, computationally expensive, and may result in inaccuracies when compared to real-world LiDAR intensity measurements.

2.3 Learning-Based LiDAR Simulation

Learning-based approaches are emerging as a promising alternative for simulating LiDAR sensors. Instead of relying on complex physics-based models, these methods leverage the power of deep learning algorithms to learn the complexity of the factors involved from the data itself (Marcus et al., 2022). Learning-based LiDAR simulation offers several advantages, including improved computational efficiency, flexibility in handling different environmental conditions, and the potential to capture subtle nuances that may

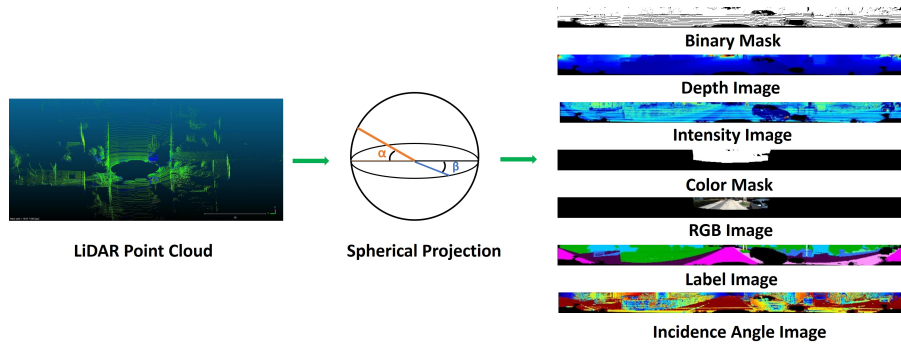


Figure 1: **Data Preparation:** The LiDAR point cloud is projected on a spherical surface to create LiDAR spherical projection images.

be challenging to model using physics-based methods. By training neural networks on large datasets of real-world LiDAR measurements, these models can capture the underlying patterns and dependencies between input features and intensity values. Recently, a generative adversarial network (GAN) for intensity rendering in LiDAR simulation was introduced by (Mok and Kim, 2021) where they transform the 3D point cloud to a 2D spherical image and project the intensity data on the pixel values to learn intensity based on unpaired real and synthetic data. Some of the recent works ((Nakashima and Kurazume, 2021), (Caccia et al., 2019) and (Manivasagam et al., 2020)) have also tried to predict the reflected signal (also called ray-drop) and estimate the intensity from the returned signal by applying some detection threshold and hence do not predict the intensity values directly. (Guillard et al., 2022) aims to simulate an enhanced point cloud with ray drops and intensity values but they project the LiDAR data and the corresponding intensity values on the RGB image space and the model learns and generates the value in the RGB image space itself and hence the intensity values can not be generated for all the points in the LiDAR data. (Vacek et al., 2021) utilize a fully supervised training approach with Convolutional Neural Networks (CNNs) using LiDAR data and its derived modalities like range, RGB, and label image to predict intensity values. They employ their trained model to generate synthetic LiDAR data with intensity by applying it to ray-casted point clouds generated in the gaming engine.

In this paper, we included the incidence angle in conjunction with range, RGB values, and semantic labels, building upon the work of (Vacek et al., 2021) and extended the comparative study to include Generative Adversarial Network (GAN) by implementing Pix2Pix architecture. By including incidence angle as an input, learning-based models can better account for the influence of surface orientation on LiDAR inten-

sity, leading to more accurate simulations. The rest of the paper is organized as follows: section 3 describes the proposed methodology, section 4 presents results on open source datasets and provides discussions about the findings, and finally section 5 provides concluding remarks.

3 METHOD

We propose to develop a hybrid learning-based approach that takes into account the physical interactions between the laser and the environment objects. We propose to use the physics-based models combined with the novel deep learning architectures to improve the realism and accuracy of LiDAR intensity. To achieve this, we employ a spherical projection technique (Li et al., 2016) to transform the 3D LiDAR points into a structured 2D image representation (Fig. 1). This involves mapping the LiDAR points onto a virtual spherical surface centered at the LiDAR sensor origin. Each LiDAR point’s coordinates are converted into spherical coordinates, including azimuth and elevation angles, which determine their positions on the spherical surface. This spherical projection preserves the geometric relationships and distances between LiDAR points, enabling the capture of spatial information in a structured format. The resulting LiDAR spherical projection image serves as a foundation for deriving various modalities used in the study.

The derived modalities from LiDAR and other complementary sensor (e.g., camera) data include depth image, logical binary Mask, segmented spherical image, RGB spherical image, LiDAR color mask, and LiDAR incidence angle image. The LiDAR depth image captures the distance information of LiDAR points from the sensor, providing a depth-based representation. The LiDAR logical binary mask indicates

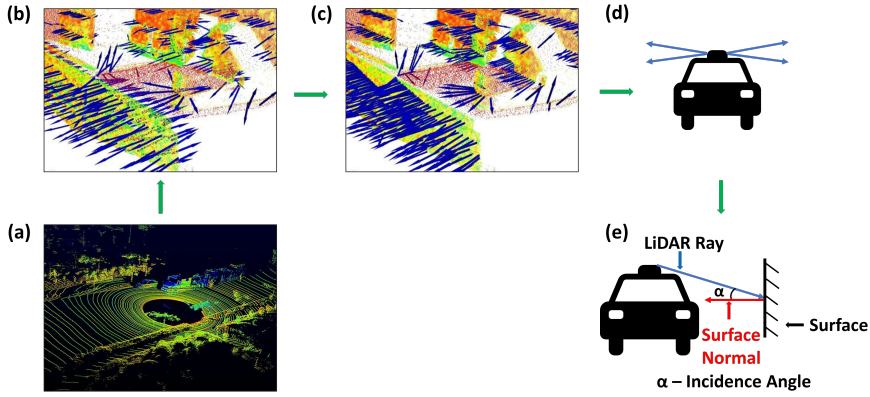


Figure 2: **Incidence Angle Calculation:** (a) LiDAR point cloud (b) Estimating surface normal (c) Orienting surface normal towards the sensor (d) Computing the direction vector of LiDAR rays (e) Computing the dot product between the direction and normal vectors of the point to get the incidence angle.

the presence or absence of returned LiDAR rays. The LiDAR segmented image represents the semantic segmentation of LiDAR points based on their assigned labels. The LiDAR RGB spherical image projects RGB values onto LiDAR points from corresponding camera images, providing color information. The LiDAR color mask is a binary mask indicating the presence or absence of RGB values for each LiDAR point. The LiDAR incidence angle image represents the angle of incidence of the laser beam at the given 3D coordinate point. Integrating these diverse modalities enhances the accuracy and realism of LiDAR intensity prediction.

Calculation of the incidence angle for each point in the LiDAR point cloud is a critical component of our proposed methodology, the entire process is shown in Fig 2. The incidence angle represents the angle between the LiDAR ray direction and the surface normal at a given point. To determine the local orientation of the underlying surface, we estimate the surface normal at each point in the point cloud. The estimated normal is then oriented towards the sensor origin to ensure consistent directionality. By considering the Euclidean distance from the sensor origin to each point, we calculate the radial position of the point. Next, we compute unit direction vectors from the sensor origin to each point by normalizing the direction vectors. The dot product between the direction vectors and the normal vector for each point is then calculated to determine the angle between them. To ensure the correct orientation of the normal vectors towards the sensor, we check if the dot product is negative for any point. If a negative dot product is found, indicating that the normal vector points away from the sensor, the normal vector is flipped to ensure it points toward the sensor. The dot product is recalculated after flipping the normal vectors. The incidence

angle for each point is obtained by taking the arccosine of the dot product. These calculated incidence angles serve as valuable information for understanding the interaction between the LiDAR sensor and the surfaces.

3.1 Network Architecture

In this study, we investigate two different categories of deep learning architectures for the task of LiDAR intensity prediction, namely U-NET and Pix2Pix, each offering distinct characteristics and capabilities.

3.1.1 U-NET Architecture

The widely used UNET architecture is employed in the first part of the study, popular for semantic segmentation tasks and recognized for its ability to capture fine-grained spatial details. To adapt UNET for this specific task, the last layer is removed, allowing the direct output of the predicted LiDAR intensity and bypassing non-linear transformations. Multiple modalities are fed into the network as separate channels which are described earlier along with the novel addition of LiDAR Incidence Angle image. This multi-channel input enables the network to utilize the diverse information captured by each modality during the prediction process. The output of the network is the predicted LiDAR intensity, which aims to accurately represent the intensity values associated with each corresponding point in the input LiDAR point cloud. The loss function used is an extended version of L2 loss called masked L2 loss. The masked L2 loss incorporates a binary mask (B) as an additional input channel in the loss calculation, enabling the model to

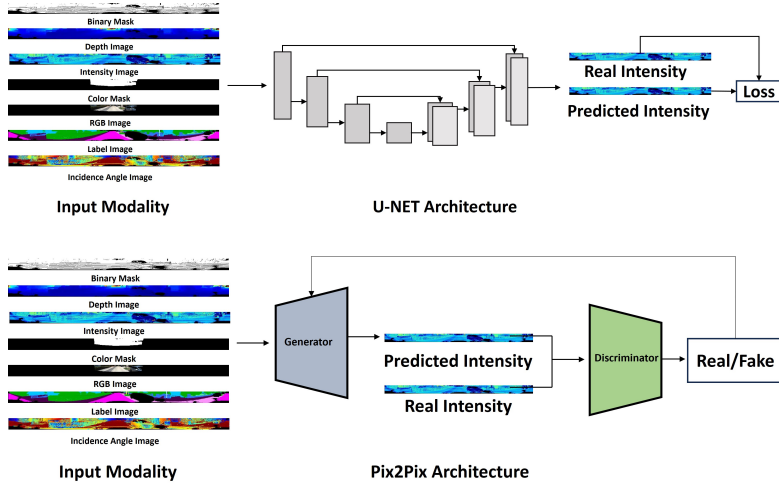

 Figure 3: **Training Pipeline:** LiDAR Spherical images are fed into the architecture to predict LiDAR intensity.

Table 1: Comparison of MSE Loss for U-NET Architecture [D - Depth, I - Incidence Angle, L - Label, RGB - Color].

Dataset	D	D+I	D+L	D+L+I	D+RGB	D+RGB+I	D+RGB+L	D+RGB+L+I
SemanticKITTI	0.353	0.349	0.339	0.321	0.317	0.301	0.298	0.234
VoxelScape	0.373	0.291	0.341	0.263	-	-	-	-

focus on the returned scan points during training.

$$L = \frac{1}{n} \sum_{i,j} (I_{i,j} - \hat{I}_{i,j})^2 \cdot B_{i,j} \quad (1)$$

It measures the difference between the real intensity values (I) and predicted intensity values (\hat{I}) for each pixel (i,j) in the spherical image. The binary mask (B) indicates the presence of returned scan points, and the loss is calculated by summing the squared errors and normalizing by the total number of successful rays. The loss quantifies the accuracy of intensity predictions in the image. To optimize the model, we employed the Adam algorithm with a learning rate of 0.003 and weight decay of 0.001 building upon the work of (Vacek et al., 2021). The experiments were performed on SemanticKITTI and VoxelScape datasets respectively and the training dataset was divided into 7500 training frames and 2500 validation frames. To evaluate the performance of the LiDAR intensity prediction model, test sets consisting of 2500 frames each were used. By using these test sets, we assessed the model’s accuracy and its ability to generalize to different datasets and scenarios.

3.1.2 Pix2Pix Architecture

In the second part of the study, we introduce the novel use of the Pix2Pix architecture for this specific task. The Pix2Pix architecture, unlike U-NET, leverages a conditional generative adversarial network (cGAN)

that consists of a generator and a discriminator. The generator aims to transform an input image into the desired output (LiDAR intensity in this case), while the discriminator’s task is to distinguish between real and generated images. For this study, the generator part of the Pix2Pix model was adapted to take in the same multi-channel inputs as the U-NET. The generator part employs a series of encoder-decoder layers to capture the relationship between the different input modalities and the LiDAR intensity. The discriminator which employs a PatchGAN structure (Isola et al., 2017b), in parallel, assesses the authenticity of the generated LiDAR intensity image. By being trained to recognize real-intensity images, it provides feedback to the generator, guiding it to produce more realistic predictions.

The loss function used in the Pix2Pix model is a combination of a conditional adversarial loss and an L1 loss (Isola et al., 2017a), as defined below:

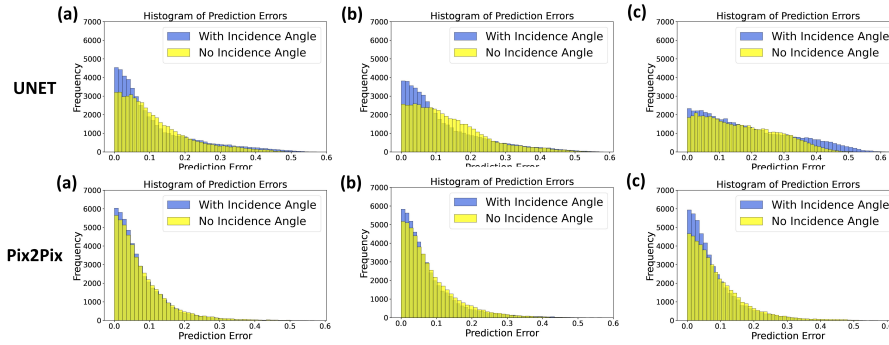
$$L = \lambda \cdot \text{L1 Loss} + \text{Adversarial Loss} \quad (2)$$

Here, the L1 Loss measures the absolute difference between the real intensity values (I) and predicted intensity values (\hat{I}) for each pixel (i,j) in the spherical image, and the Adversarial Loss quantifies how well the generator fools the discriminator.

The Pix2Pix architecture was optimized using the Adam algorithm with a learning rate of 0.0002, λ value of 100, and a gradient penalty term with a coefficient of 10. Similar to the U-NET model, the experiments were carried out on SemanticKITTI and Vox-

Table 2: Comparison of MSE Loss for Pix2Pix Architecture [D - Depth, I - Incidence Angle, L - Label, RGB - Color].

Dataset	D	D+I	D+L	D+L+I	D+RGB	D+RGB+I	D+RGB+L	D+RGB+L+I
SemanticKITTI	0.334	0.325	0.319	0.299	0.295	0.272	0.253	0.201
VoxelScape	0.247	0.192	0.211	0.182	-	-	-	-

Figure 4: **Error Histogram:** SemanticKITTI Data - Input combinations: (a) Depth + RGB + Label (b) Depth + RGB (c) Depth.

elScape datasets. The division of training and validation frames and the evaluation metrics were kept consistent with the U-NET model to ensure a fair comparison.

4 RESULTS AND DISCUSSION

The LiDAR intensity prediction experiments were conducted using two datasets: SemanticKITTI (Behley et al., 2019) and VoxelScape (Saleh et al., 2023). To assess the performance of different input modalities, the mean squared error (MSE) loss was calculated for both U-NET and Pix2Pix GAN architectures on the test set. Further, error histograms and heatmaps were generated to analyze the prediction errors and their spatial distribution. The experiments were performed on SemanticKITTI and VoxelScape datasets, respectively, and the training dataset was divided into 7500 training frames and 2500 validation frames. To evaluate the performance of the LiDAR intensity prediction model, test sets consisting of 2500 frames each were used.

4.1 Mean Squared Error (MSE)

In ablation studies involving both the SemanticKITTI and VoxelScape datasets, the performances of U-NET and Pix2Pix architectures were compared using various input modalities. As highlighted in Table 1 and Table 2, the results show that the Pix2Pix architecture outperforms U-NET in all instances. Importantly, adding the incidence angle across both architectures

consistently led to improved performance. This highlights the significance of incorporating the incidence angle as an informative input modality for LiDAR intensity prediction. The VoxelScape data does not contain RGB images; hence, the experiment results involving RGB information for VoxelScape are not present in Table 1 and Table 2.

4.2 Error Histogram

To further analyze the prediction errors between the ground truth and predicted intensity values, error histograms for each frame from the SemanticKITTI and VoxelScape datasets were generated. These histograms visualize the distribution of prediction errors across the intensity range. As shown in Fig 4 and Fig 5, it was observed that in the results from Pix2Pix architecture, the error distribution is concentrated more around zero error compared to the U-NET architecture, and when the incidence angle was included it was further improved indicating a better alignment between predicted and ground truth intensity values. This supports the notion that incorporating the incidence angle enhances the accuracy of the LiDAR intensity prediction.

4.3 Error Heatmaps

Furthermore, error heatmaps were generated to visualize the spatial distribution of prediction errors and assess the impact of incorporating incidence angle on LiDAR intensity prediction. Upon examining the

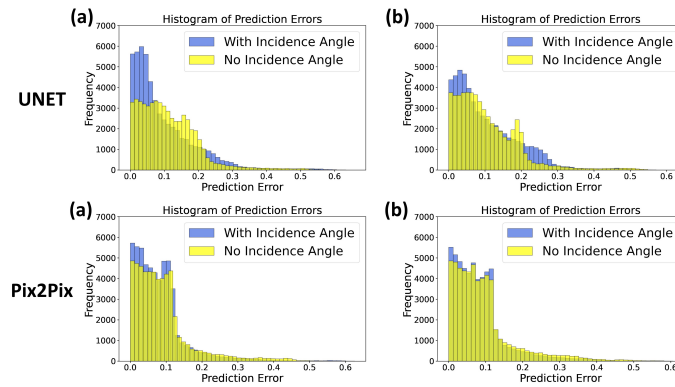


Figure 5: **Error Histogram:** VoxelScape Data - Input combinations: (a) Depth + Label (b) Depth.

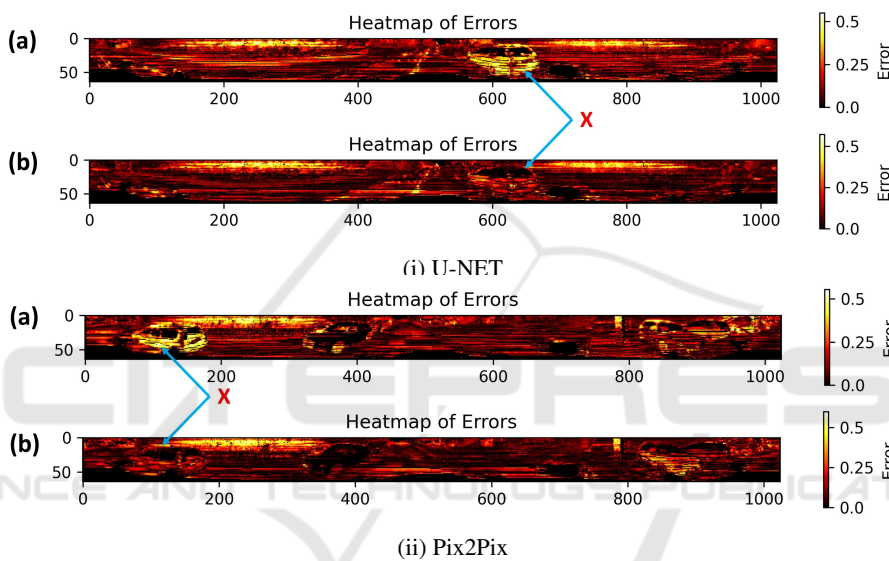


Figure 6: **Error Heatmap:** Illustrating the impact of Incidence Angle on LiDAR Intensity Simulation a) Without Incidence Angle b) With Incidence Angle.

heatmaps, a distinct pattern emerges, showcasing localized improvements in predicted intensity; as shown in Fig 6, the error is significantly reduced for the cars present in the image (marked as X). Integrating the incidence angle results in localized refinements, reducing prediction errors in distinct regions, thereby enhancing the spatial consistency and accuracy of the intensity prediction. This phenomenon can be attributed to the significant influence that incidence angle has on the intensity of LiDAR reflections. By incorporating this crucial factor, the model gains a deeper understanding of the underlying knowledge of the interactions between the LiDAR rays and the surface, capturing the variations in reflectivity based on surface orientation, resulting in more precise and contextually aware intensity predictions.

4.4 Qualitative Analysis

For a visual assessment, we looked closely at the images generated by both U-NET and Pix2Pix and compared them to the reference images from the SemanticKITTI and VoxelScape datasets as shown in Fig 7. In this comparison, it was clear that the Pix2Pix images are closer to the reference images than the U-NET ones. U-NET produces blurry output as also mentioned by (Isola et al., 2017b) and Pix2Pix did a far better job at capturing the details present in the scene.

While both the architectures share U-NET as the core architecture, Pix2Pix leverages a Generative Adversarial Network (GAN) training framework, setting it apart from U-NET’s reliance solely on pixel-wise loss (L2 loss). This distinction in loss objectives explains the Pix2Pix’s edge. U-NET minimizes pixel-wise errors, while Pix2Pix’s adversarial loss encour-

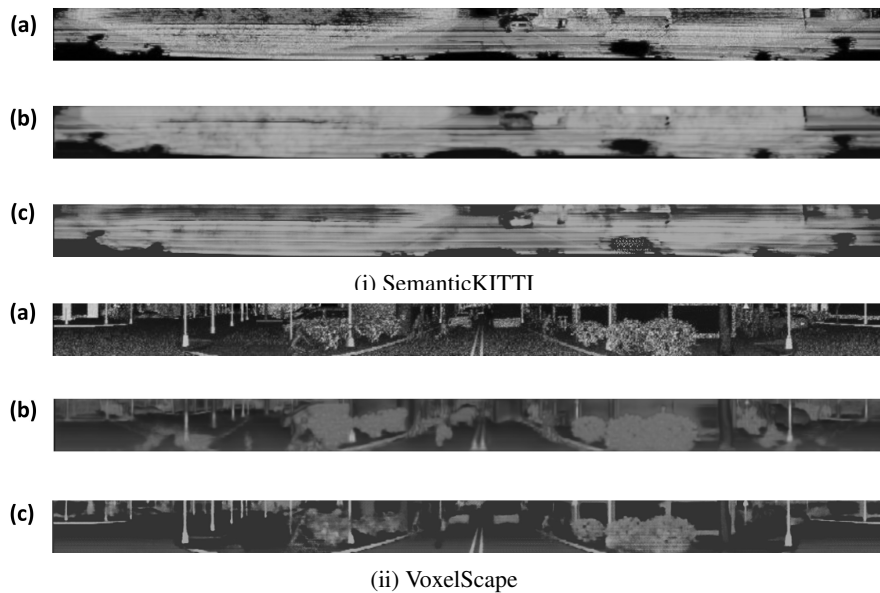


Figure 7: **Qualitative Analysis:** Visual assessment of results from U-NET and Pix2Pix architectures. a) Reference LiDAR intensity image b) U-NET architecture c) Pix2Pix architecture.

ages the model to match ground truth and generate outputs that deceive a discriminator network, potentially leading to more refined and realistic outcomes, hence outlining the importance of generative adversarial network (GAN) in simulating LiDAR data.

5 CONCLUSION

In this research paper, we presented a comprehensive study involving two different deep learning architectures U-NET and Pix2Pix for LiDAR intensity simulation. The core novelty of this approach lies in incorporating the incidence angle as an additional input modality, enhancing both the accuracy and context awareness of LiDAR intensity predictions. Through extensive experiments using both SemanticKITTI and VoxelScape data, we demonstrated the effectiveness of the U-NET and Pix-2Pix architecture with the inclusion of the incidence angle. We showed that this combination leads to improvements in prediction performance. Further, our implementation of the Pix2Pix architecture in this context, a novel approach in the literature, exhibited even better results, thus underscoring the potential of generative adversarial networks in LiDAR intensity simulation. Our findings consistently revealed that the integration of the incidence angle, along with other modalities led to superior performance across both architectures. This success not only addresses the inherent challenges posed by the complex nature of intensity simulation but also reflects a significant step forward in narrowing the gap

between simulated and real-world LiDAR data.

The successful integration of the incidence angle and the comparative analysis of U-NET and Pix2Pix represent innovative and meaningful contributions to the field of LiDAR intensity simulation. These advancements have tangible implications for various applications, including autonomous driving, robotics, and environmental mapping, where accurate LiDAR intensity plays a crucial role in perception, object recognition, and scene understanding. For future research, we will explore other factors, such as material properties, surface reflectance, and the development of network architectures and loss functions designed specifically for LiDAR intensity simulation and incorporating the physics of LiDAR intensity into it.

ACKNOWLEDGEMENTS

This work was supported by the Mitacs Globalink Research Award. We are also grateful to the University of New Brunswick, Canada, for providing access to *Compute Canada*, the national high-performance compute (HPC) system, for our experiments.

REFERENCES

- Behley, J., Garbade, M., Milioto, A., Quenzel, J., Behnke, S., Stachniss, C., and Gall, J. (2019). Semantickitti: A dataset for semantic scene understanding of lidar sequences. In *Proceedings of the IEEE/CVF inter-*

- national conference on computer vision*, pages 9297–9307.
- Caccia, L., Van Hoof, H., Courville, A., and Pineau, J. (2019). Deep generative modeling of lidar data. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5034–5040. IEEE.
- Caesar, H., Bankiti, V., Lang, A. H., Vora, S., Liong, V. E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., and Beijbom, O. (2020). nuscenes: A multimodal dataset for autonomous driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11621–11631.
- Dai, W., Chen, S., Huang, Z., Xu, Y., and Kong, D. (2022). Lidar intensity completion: Fully exploiting the message from lidar sensors. *Sensors*, 22(19):7533.
- Elmqvist, A. and Negrut, D. (2020). Methods and models for simulating autonomous vehicle sensors. *IEEE Transactions on Intelligent Vehicles*, 5(4):684–692.
- Geiger, A., Lenz, P., and Urtasun, R. (2012). Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE conference on computer vision and pattern recognition*, pages 3354–3361. IEEE.
- Guillard, B., Vemprala, S., Gupta, J. K., Miksik, O., Vineet, V., Fua, P., and Kapoor, A. (2022). Learning to simulate realistic lidars. In *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 8173–8180. IEEE.
- Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. (2017a). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134.
- Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A. A. (2017b). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Li, B., Zhang, T., and Xia, T. (2016). Vehicle detection from 3d lidar using fully convolutional network. *arXiv preprint arXiv:1608.07916*.
- Li, Y. and Ibanez-Guzman, J. (2020). Lidar for autonomous driving: The principles, challenges, and trends for automotive lidar and perception systems. volume 37, pages 50–61. IEEE.
- Liang, Z., Zhang, M., Zhang, Z., Zhao, X., and Pu, S. (2020). Rangercnn: Towards fast and accurate 3d object detection with range image representation. *arXiv preprint arXiv:2009.00206*.
- Manivasagam, S., Wang, S., Wong, K., Zeng, W., Sazanovich, M., Tan, S., Yang, B., Ma, W.-C., and Urtasun, R. (2020). Lidarsim: Realistic lidar simulation by leveraging the real world. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11167–11176.
- Marcus, R., Knoop, N., Egger, B., and Stamminger, M. (2022). A lightweight machine learning pipeline for lidar-simulation. *arXiv preprint arXiv:2208.03130*.
- Meyer, G. P., Laddha, A., Kee, E., Vallespi-Gonzalez, C., and Wellington, C. K. (2019). Lasernet: An efficient probabilistic 3d object detector for autonomous driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 12677–12686.
- Mok, S.-c. and Kim, G.-w. (2021). Simulated intensity rendering of 3d lidar using generative adversarial network. In *2021 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pages 295–297. IEEE.
- Nakashima, K. and Kurazume, R. (2021). Learning to drop points for lidar scan synthesis. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 222–229. IEEE.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18*, pages 234–241. Springer.
- Royo, S. and Ballesta-Garcia, M. (2019). An overview of lidar imaging systems for autonomous vehicles. *Applied sciences*, 9(19):4093.
- Saleh, K., Hossny, M., Abobakr, A., Attia, M., and Iskander, J. (2023). Voxelscape: Large scale simulated 3d point cloud dataset of urban traffic environments. *IEEE Transactions on Intelligent Transportation Systems*, pages 1–14.
- Schwarting, W., Alonso-Mora, J., and Rus, D. (2018). Planning and decision-making for autonomous vehicles. *Annual Review of Control, Robotics, and Autonomous Systems*, 1:187–210.
- Sun, P., Kretzschmar, H., Dotiwalla, X., Chouard, A., Patnaik, V., Tsui, P., Guo, J., Zhou, Y., Chai, Y., Caine, B., et al. (2020). Scalability in perception for autonomous driving: Waymo open dataset. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 2446–2454.
- Tatoglu, A. and Pochiraju, K. (2012). Point cloud segmentation with lidar reflection intensity behavior. In *2012 IEEE International Conference on Robotics and Automation*, pages 786–790. IEEE.
- Vacek, P., Jašek, O., Zimmermann, K., and Svoboda, T. (2021). Learning to predict lidar intensities. *IEEE Transactions on Intelligent Transportation Systems*, 23(4):3556–3564.
- Wang, W. and Shen, J. (2017). Deep visual attention prediction. *IEEE Transactions on Image Processing*, 27(5):2368–2378.
- Wang, Y., Shi, T., Yun, P., Tai, L., and Liu, M. (2018). Pointseg: Real-time semantic segmentation based on 3d lidar point cloud. *arXiv preprint arXiv:1807.06288*.
- Wang, Z., Fu, H., Wang, L., Xiao, L., and Dai, B. (2019). Snet: Subdivision coding network for object detection based on 3d point cloud. *IEEE Access*, 7:120449–120462.
- Yang, B., Liang, M., and Urtasun, R. (2018a). Hdnet: Exploiting hd maps for 3d object detection. In *Conference on Robot Learning*, pages 146–155. PMLR.

- Yang, B., Luo, W., and Urtasun, R. (2018b). Pixor: Real-time 3d object detection from point clouds. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 7652–7660.
- Yang, G., Xue, Y., Meng, L., Wang, P., Shi, Y., Yang, Q., and Dong, Q. (2021). Survey on autonomous vehicle simulation platforms. In *2021 8th International Conference on Dependable Systems and Their Applications (DSA)*, pages 692–699. IEEE.

