Point Cloud Up-Sampling and Domain Adaptation

Prabhakar Kumar Panday

Examiner : Prof. Dr.-Ing. Bin Yang, ISS

Supervisor : Mr. Ahmed Yousif, Valeo

University of Stuttgart

31.08.2023

Structure of the Presentation

- Introduction
  - Approaches
  - Dataset Creation
- Challenges
  - Scale
  - Data Heterogeneity
- U-Net
  - Metrics
  - Sample Outputs
  - Observations
- Pix2PixHD
  - Sample Outputs
  - Frobenius Norm
  - Range & Area
- Conclusion

Introduction

Manual Labelling and Bounding Box Generation is a tedious task and a bottleneck.

Time-Consuming and Resource-Intensive aspects of manual data labeling.

Transforming data from previous-generation LiDAR sensors to next generation can be a solution.

Build a Deep Learning model to Up-sample the Point-Cloud recorded.

Approaches

A sample Range Image

Grid-Based Architectures:

Involves 3D convolution and Voxelization

Point-Based Methods:

Works directly on raw point cloud data using pointwise FCC layers and Max pooling

Graph based Techniques:

Using Graph Neural Networks

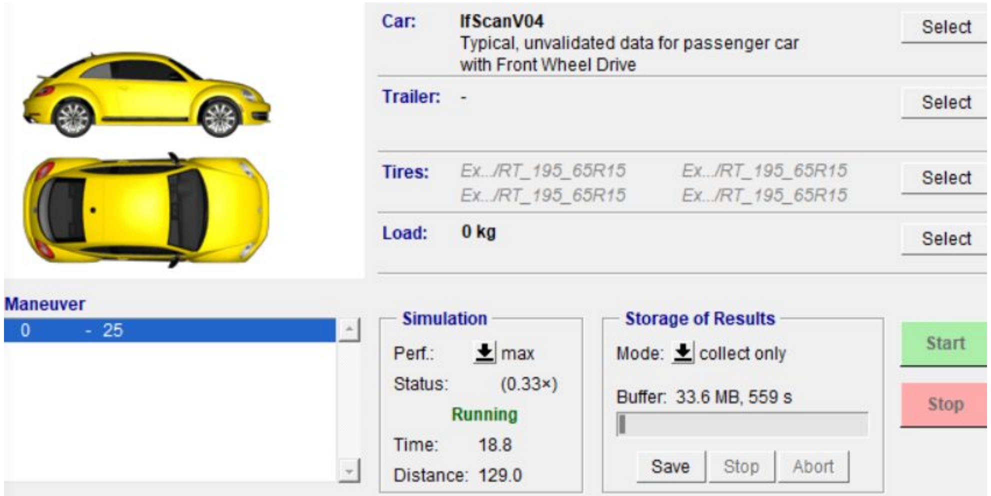
Examples: FB, Twitter, and so-on…

Minkowski Engine

Range Image based Approach

Dataset: Simulated Data Generation

- 43,000 Files for different Scenarios according to Daimler Test cases.

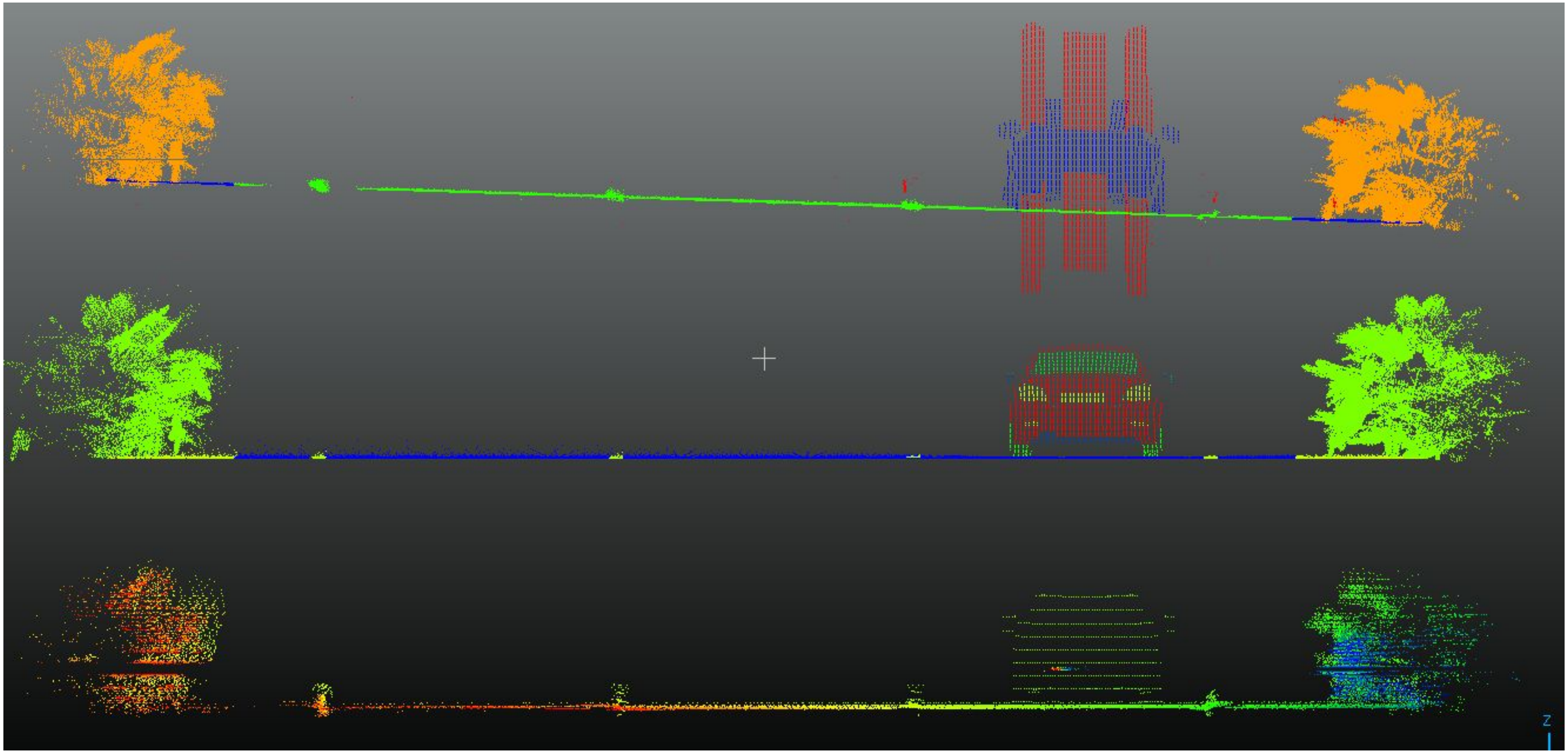


CarmakerWorking Directory.

```
# .PCD v.7 - Point Cloud Data file format
VERSION .7
FIELDS x y z EPW EchoNumber ObjectID threshold Layer.Point
SIZE 4 4 4 4 4 4 4 4
TYPE F F F F I I I F
COUNT 1 1 1 1 1 1 1 1
WIDTH 18183
HEIGHT 1
VIEWPOINT 0 0 0 1 0 0 0
POINTS 18183
DATA ascii
4.69488899 10.76537527 0.96300167 104.00000000 0 0 1 0.24
4.69329534 10.76172104 0.96267479 144.00000000 0 0 0 0.24
4.73752462 10.73506179 0.96212998 120.00000000 0 0 1 0.25
4.73752462 10.73506179 0.96212998 152.00000000 0 0 0 0.25
5.15374546 10.77478732 0.97934590 104.00000000 0 0 1 0.32
```

Sample PCD file

Dataset: Sensor Models: Ideal & Physical



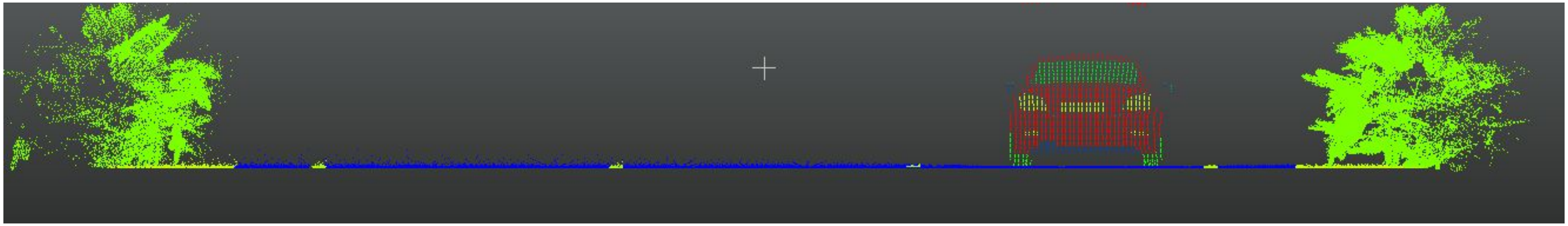
TOP: Trace generated from Physical Sensor Model.(SCALA Gen 3)

Middle: Trace generated from Ideal Sensor Model.(SCALA Gen 3)

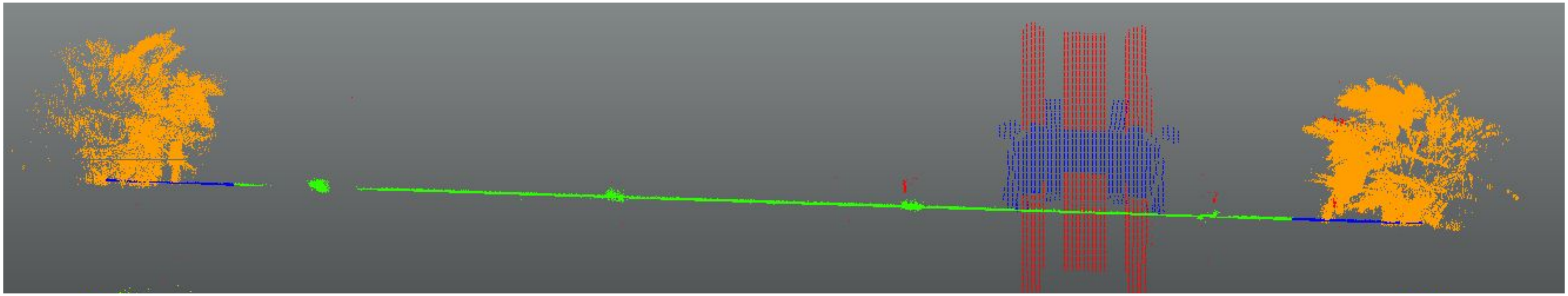
Bottom: Trace generated from Physical Sensor Model.(SCALA Gen 2)

Dataset: Sensor Models: Ideal & Physical

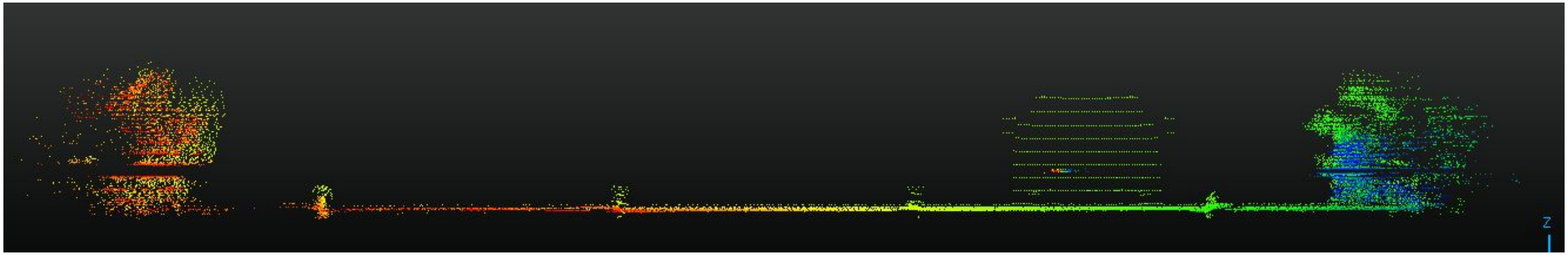
Trace generated from Ideal Sensor Model.(SCALA Gen 3)



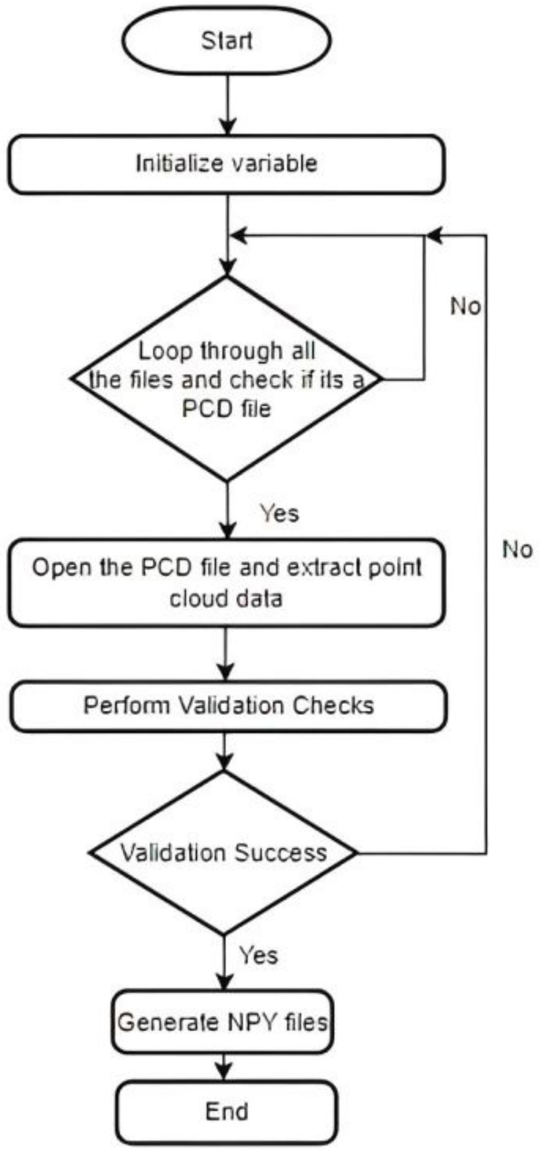
Trace generated from Physical Sensor Model.(SCALA Gen 3)



Trace generated from Ideal Sensor Model.(SCALA Gen 2)



Dataset: Generation of Range Images



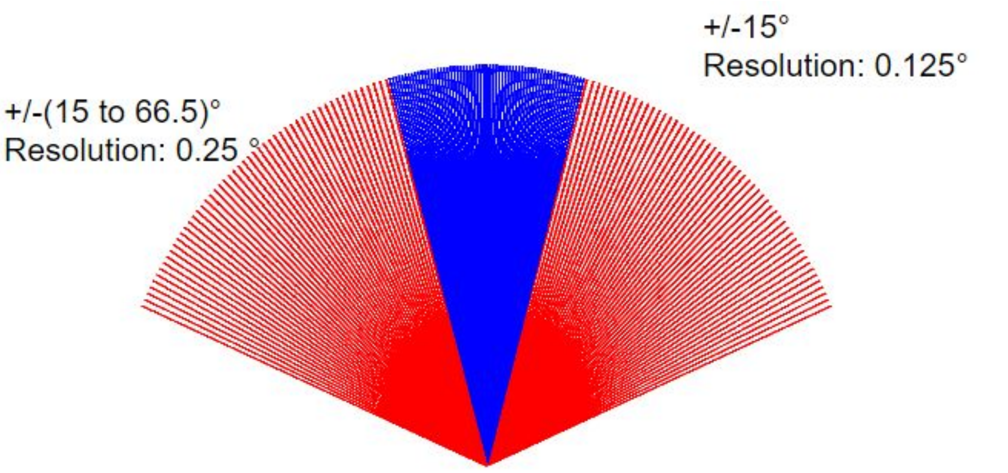
Flowchart of data-preprocessing

To validate that, the dataset for Scala2 and Scala3 are in Sync w.r.t time, mounting position on the car, and data columns.

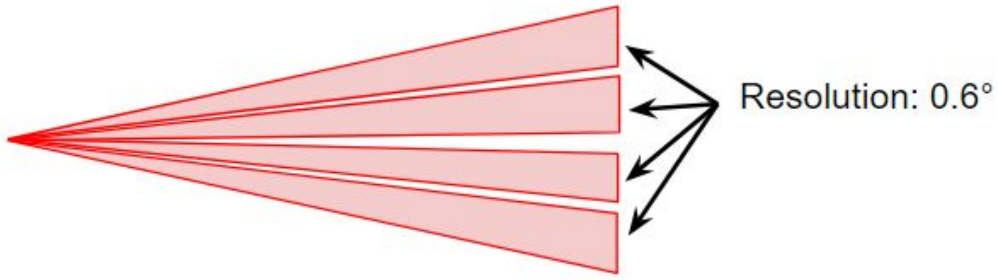
Generate NPY files of range images using Layer and Slots of the PCD file.

Challenges: Misalignment

- 133° (±66.5°) horizontal field of view, 0.25° resolution sides [1]
- Center hFOV (±15°) with double resolution to 0.125°
- 10° vertical field of view with 0.6° resolution



Top View of Scala Gen 2

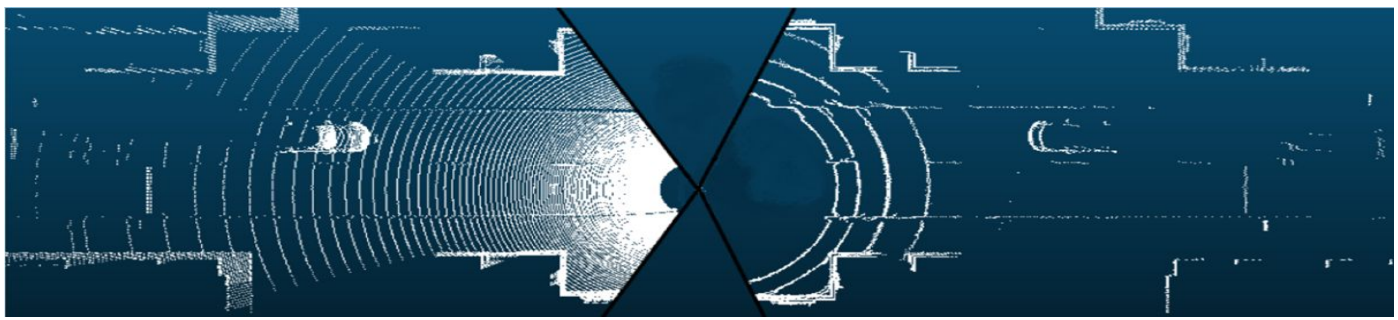


Lateral View of Scala Gen 2

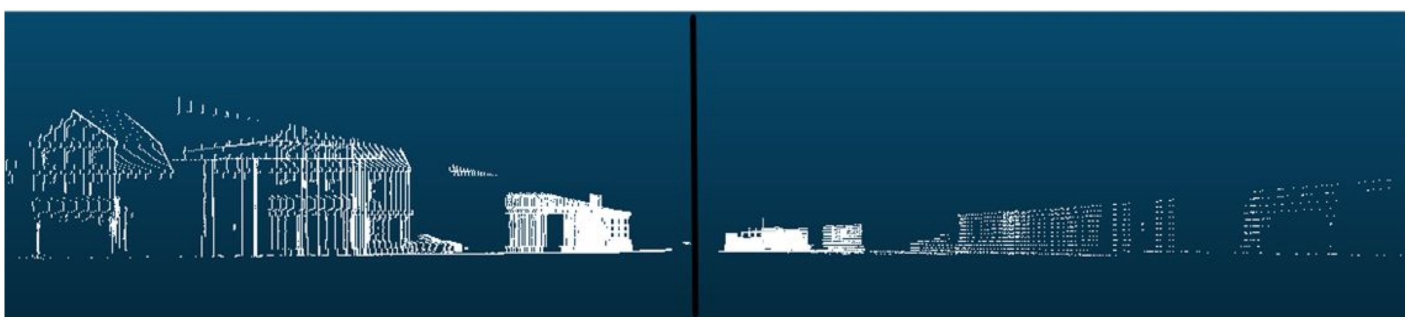
Challenges: X to Up-Sample

		Horizontal		Vertical		Maximum Number Of Points	X to Upsample Vertically
		Angle	Resolution	Angle	Resolution		
SCALA2[2]	Input	133°	701	10°	16	11216	**x8.5**
	Pruned	120°	600	10°	16	**9600**	
SCALA3[1]	Input	120°	600	***°	**	*****	
	Pruned	120°	600	10°	136	**81600**	

Comparison of Scala Gen 2 v/s Scala Gen 3 Sensors

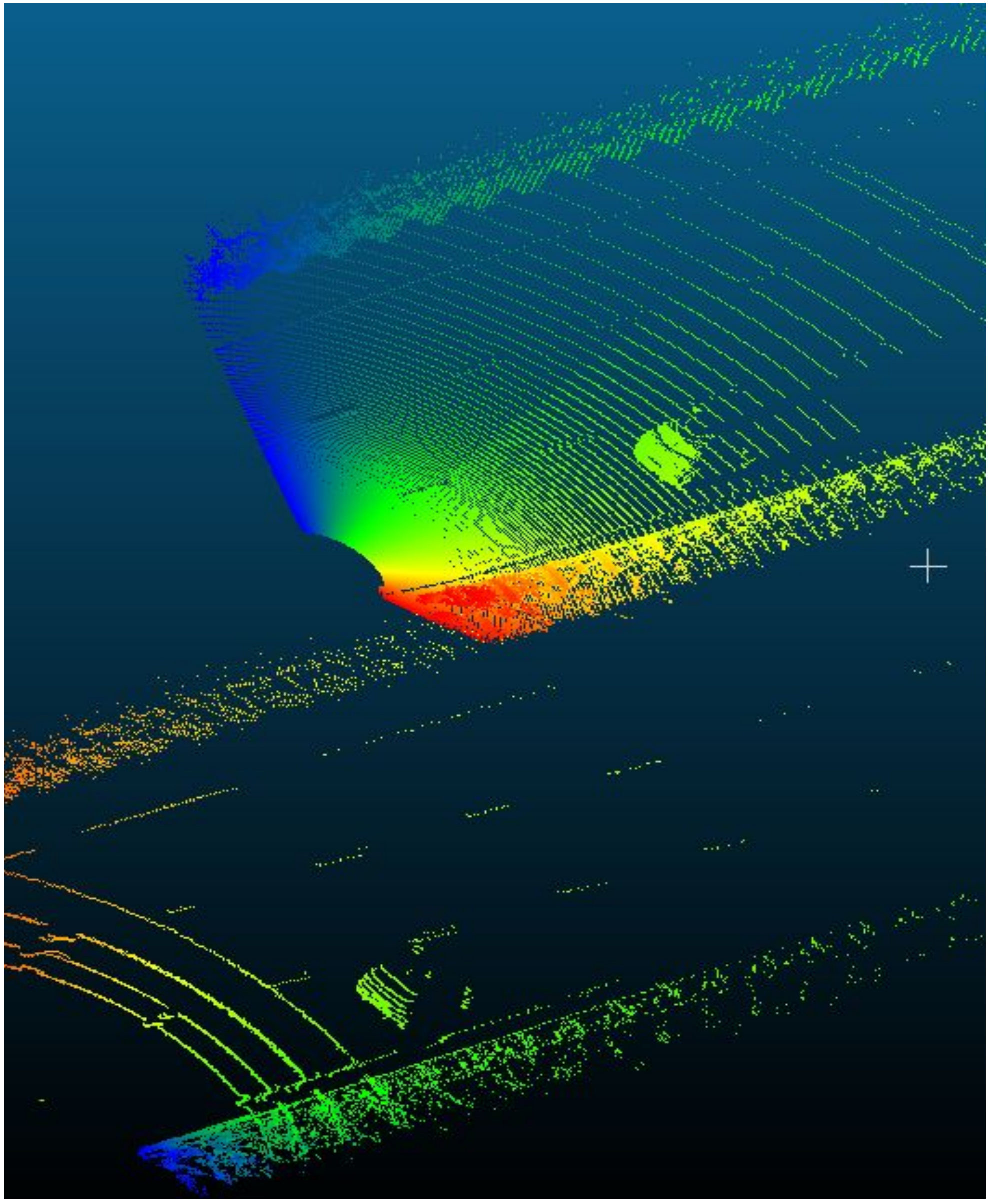


BEV Comparison Scala Gen 3 (Left), Scala Gen 2(Right)



Comparison Scala3(Left), Scala2(Right)

Challenges: Additional



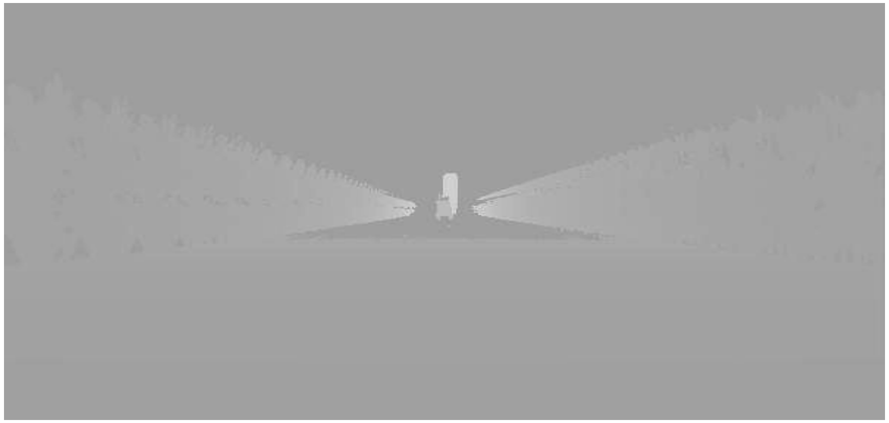
Output (Top)

Input (Bottom)

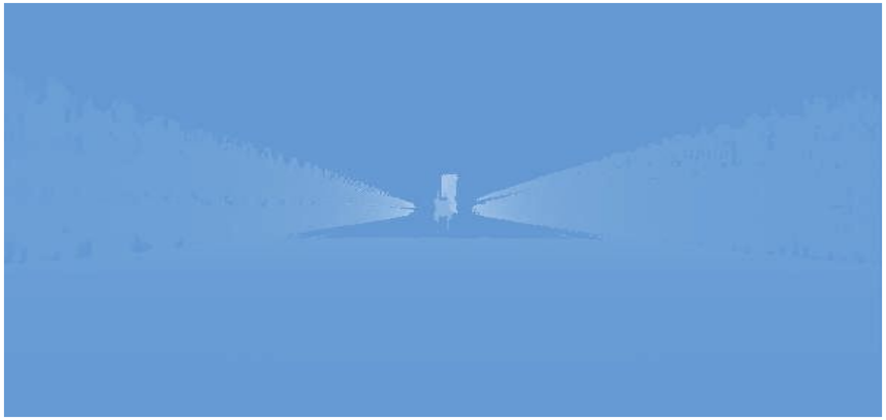
- Data Heterogeneity.
- Misalignment of up to 10°.
- Occlusion Handling.
- Noise and Artifacts like blooming.
- No Semantic Understanding.
- Computational Complexity.

Challenges: Visualization

Input to the Model

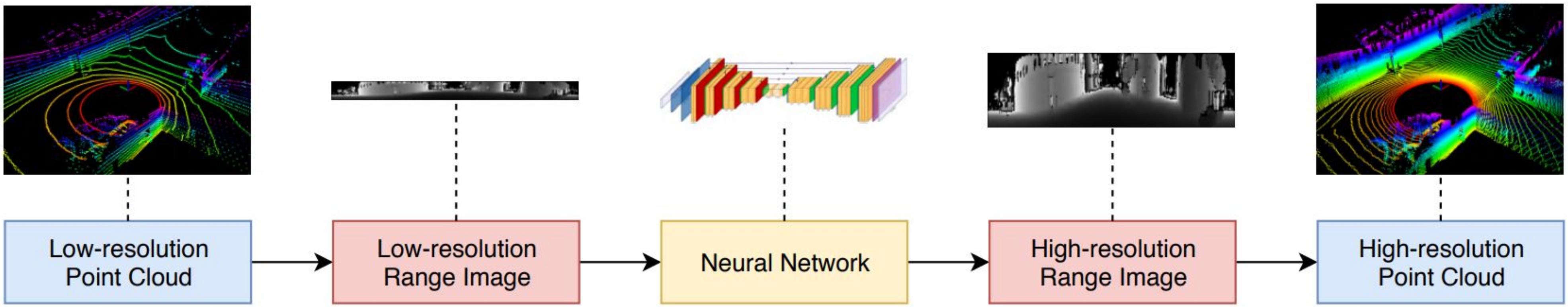


Ground Truth

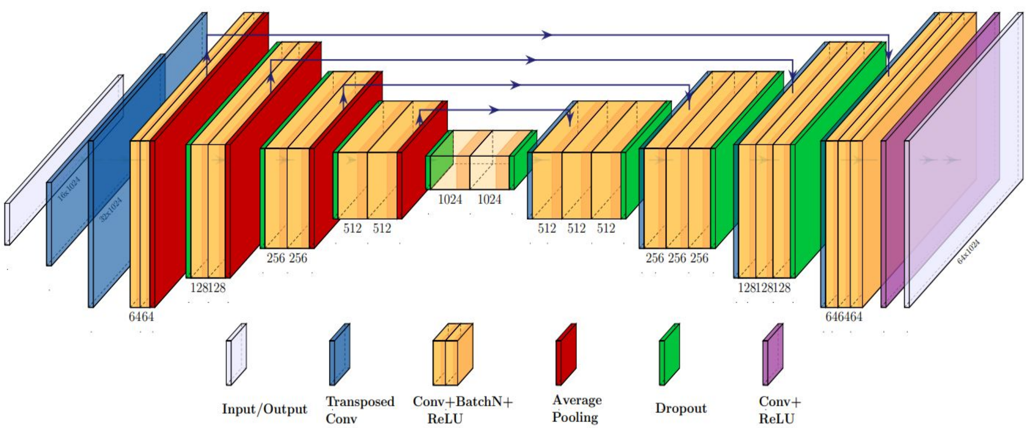


Prediction of Pix2PixHD with Original Losses

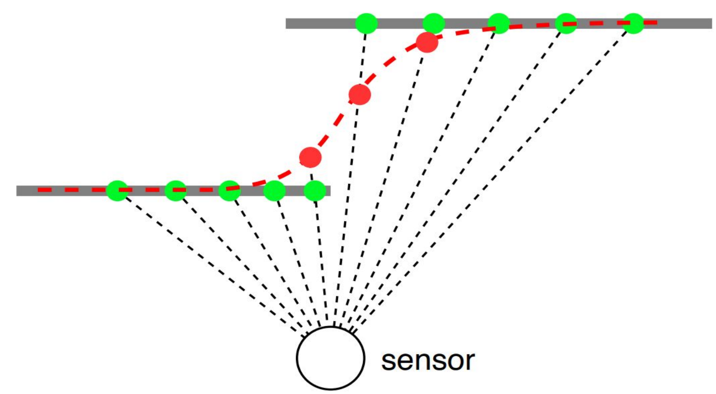
U-Net based model[3]



Workflow of the U-Net based model



U-Net architecture



Monte-Carlo Dropout[7]

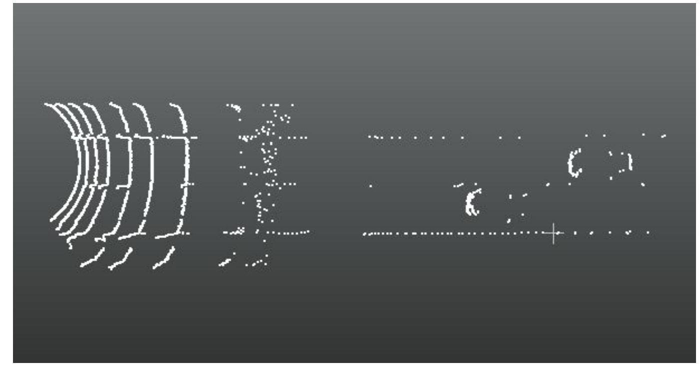
U-Net based model: Metrics

3D Metrics

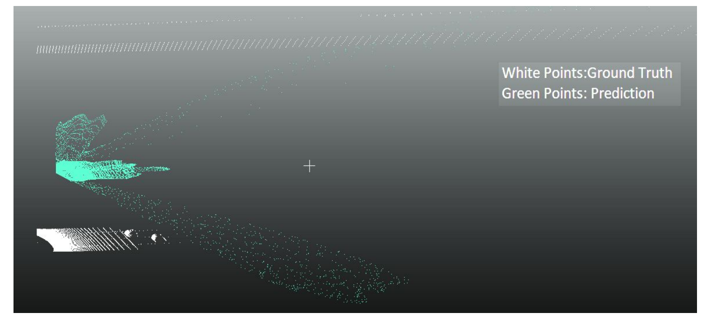
Loss Function Name	Hausdorff Distance ↓	Chamfer Distance ↓	IoU ↑	RMSE ↓	Spatial Distribution Difference
MAE	0.139	**0.016**	**0.92**	0.078	0.023
MSE	0.347	0.037	0.84	0.065	9.490
MDE	0.253	**0.016**	0.88	0.078	0.008
(MSE*0.1+ MDE*0.4)	0.198	0.019	0.90	0.079	0.008
SSIM	**0.185**	0.020	0.85	0.078	**0.007**
VNL	0.665	0.26	0.76	**0.055**	0.009

2D Metrics

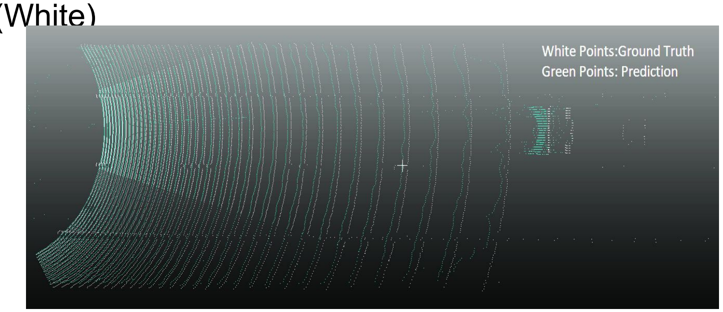
Loss Function Name	PSNR ↑	WSN Loss ↓	F1 Score ↑	EMD ↓	IOU↓
MAE	63.183	0.00014	0.96	0.0425	**0.875**
MSE	42.152	0.000329	0.75	0.0919	0.690
MDE	**64.248**	**0.00012**	**0.97**	**0.040**	0.885
(MSE*0.1+ MDE*0.4)	62.421	0.00019	0.92	0.059	0.84
SSIM	61.466	0.00017	0.92	0.0500	0.81
VNL	60.511	0.00015	0.92	0.0512	0.805



BEV of the Input point cloud



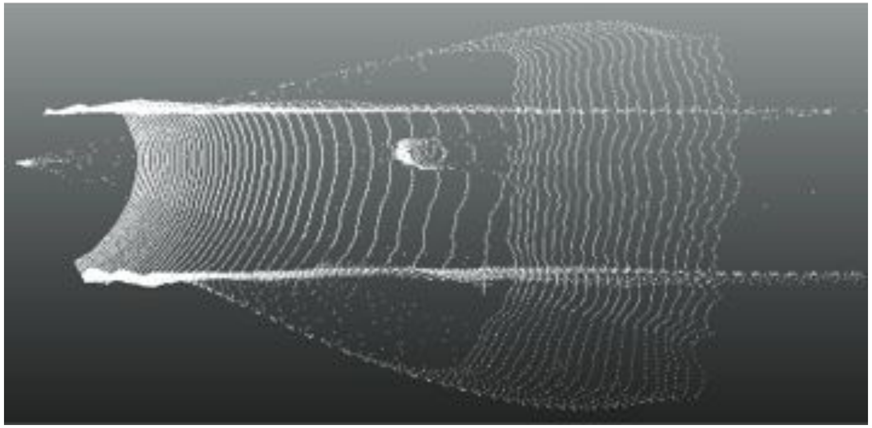
Lateral View of the prediction (Green) & Ground Truth (White)



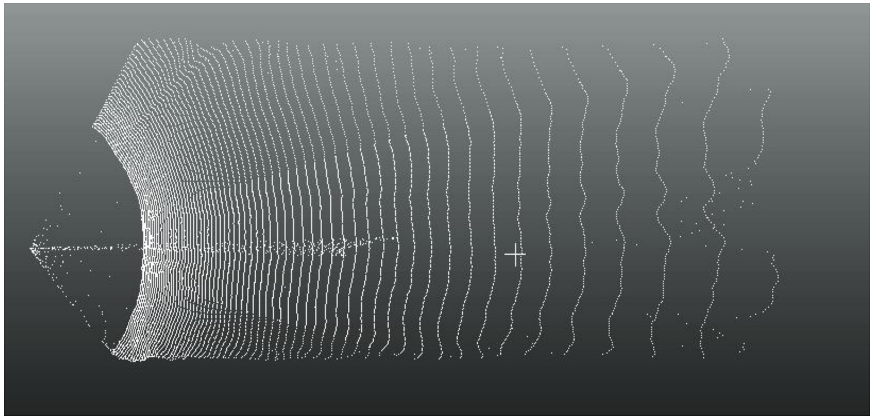
U-Net output with SSIM Loss

Note: All the Range images are 16 bit*

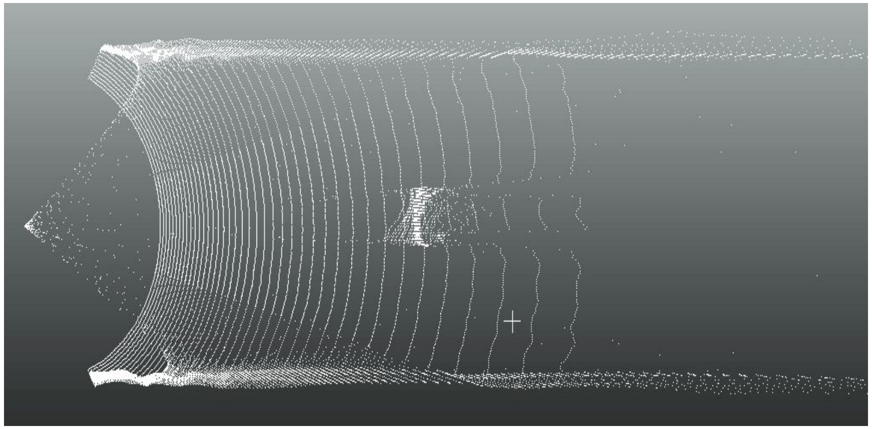
U-Net: Quantitative Analysis of Losses



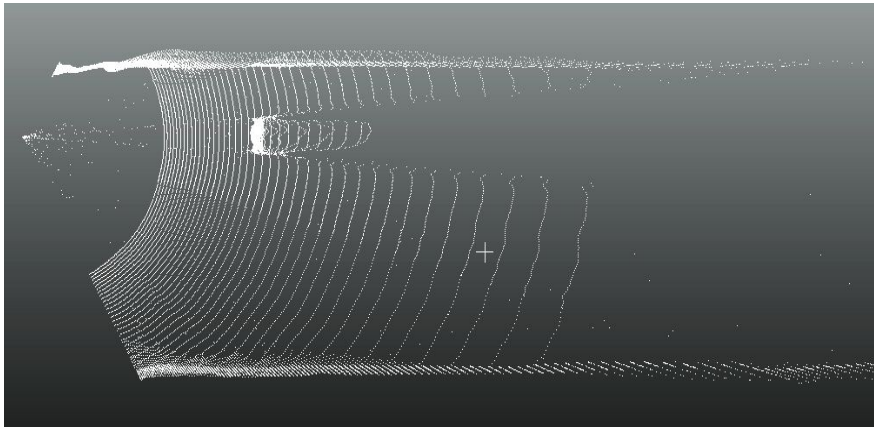
SSIM Loss



VNL Loss



MSE Loss



MAE Loss

U-Net based model: Observations

MSE

MSE loss Increased noise in the output.

$\mathcal{L}_{VNL}$

VNL Loss gave accurate ground reflection but traffic objects were poorly up-sampled.

MAE

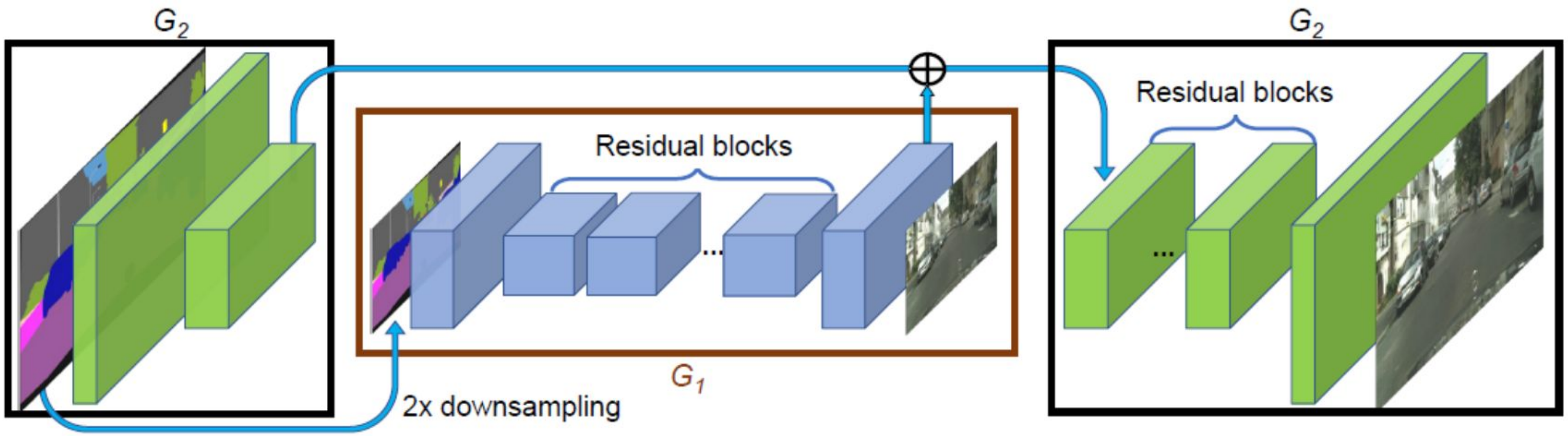
MAE Loss has less noise but far away objects were poorly up-sampled.

SSIM

SSIM loss improved the output but only for one traffic object .

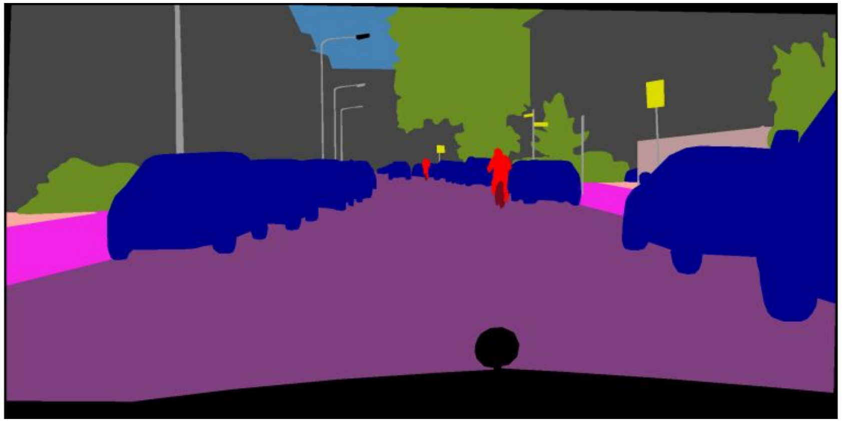
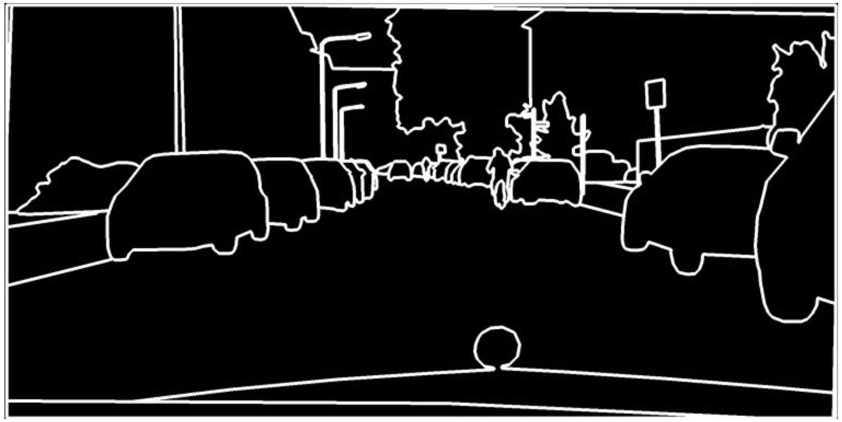
Pix2PixHD: Architecture[3]

- Pix2PixHD is based on conditional GANs.
- Pix2PixHD is designed for 3 channels RGB images.
- Generates high resolution images with the help of semantic label maps.
- It uses instance map to differentiate between different objects.



Pix2PixHD Architecture[3]

Pix2PixHD: Instance Maps

A typical semantic segmented scene where same label is used to label the connected cars (Left). On the right, the instance map for the same image is passed to the GAN model which is the edges which help separate different objects.


Comparison of the output of the model when trained with(right image) and without (left image) the instance maps.

Pix2PixHD: Instance Maps

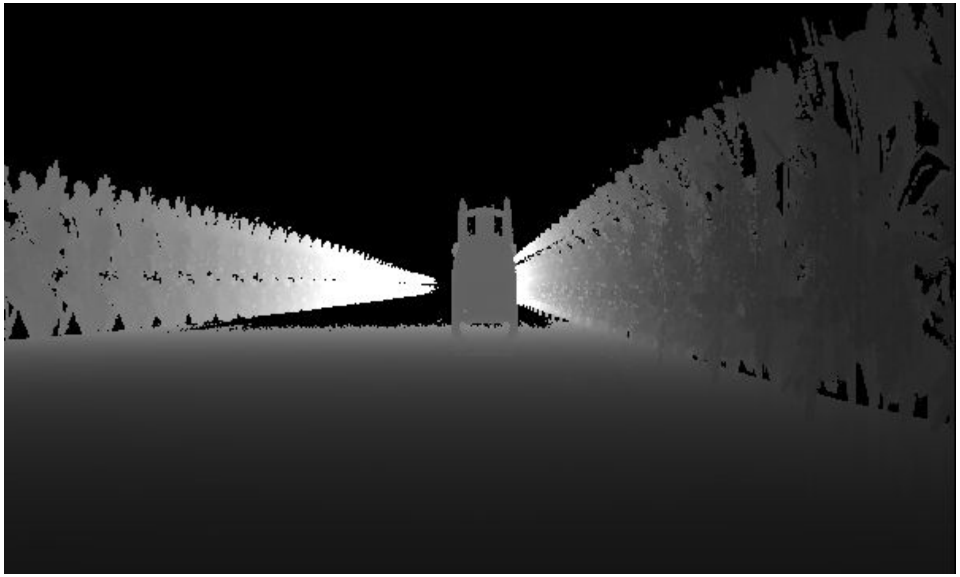


Figure: Range Image

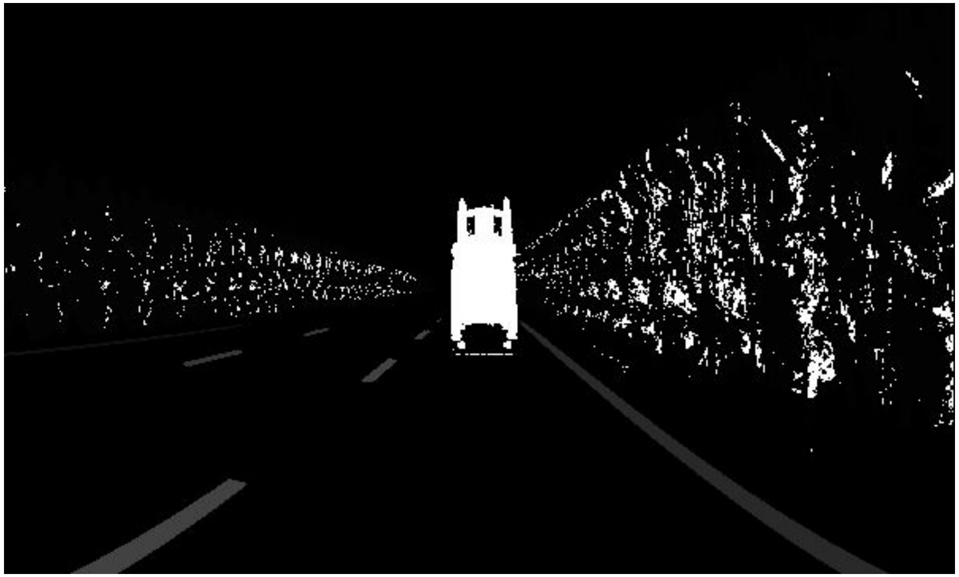
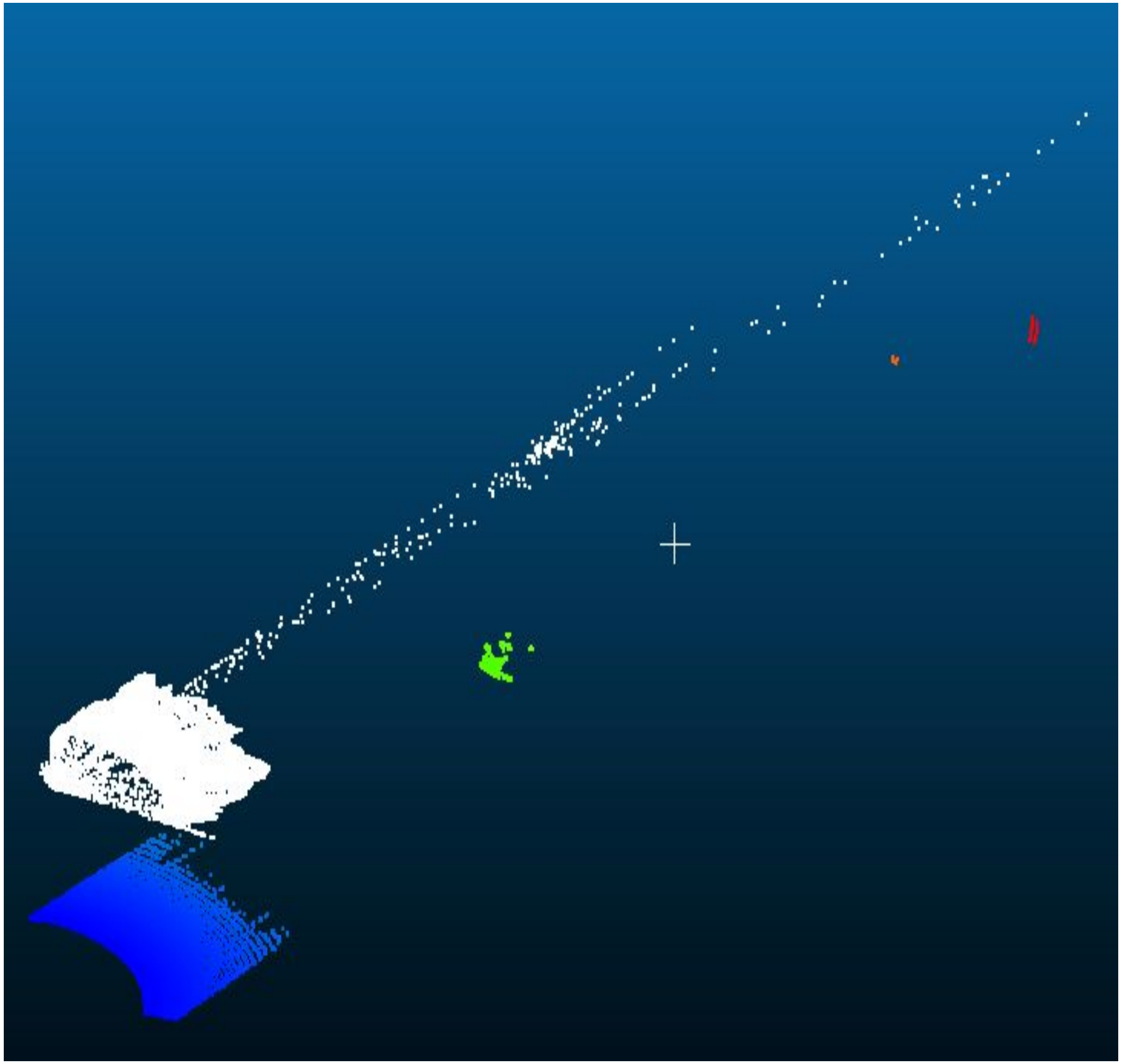


Figure: Edge Detector

Pix2PixHD: Original Loss Outputs

- Original Losses
  - MAE
  - GAN Feat Loss
  - Perceptual Loss

- Observation
  - Output is Noisy
  - The traffic Objects are not accurately up-sampled.
  - Not enough to build clusters.



Output of the Pix2PixHD model with Original Loss functions.

Top:Prediction, Bottom: Ground Truth

Pix2PixHD: Frobenius Norm[5]

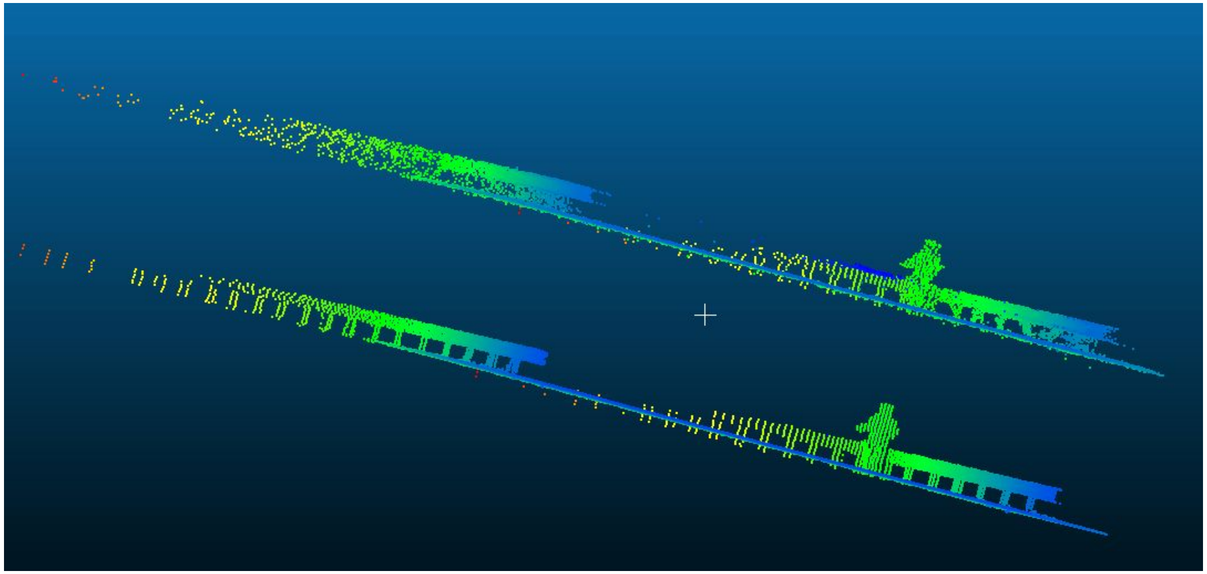
- Frobenius Norm quantifies the overall magnitude or "size" of a matrix.

$$\|A\|_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{n} a_{i,j}^2 \right)^{\frac{1}{2}}$$

Frobenius Norm equation for the Loss

Matrix

- MSE quantifies element-wise discrepancies between matrices.

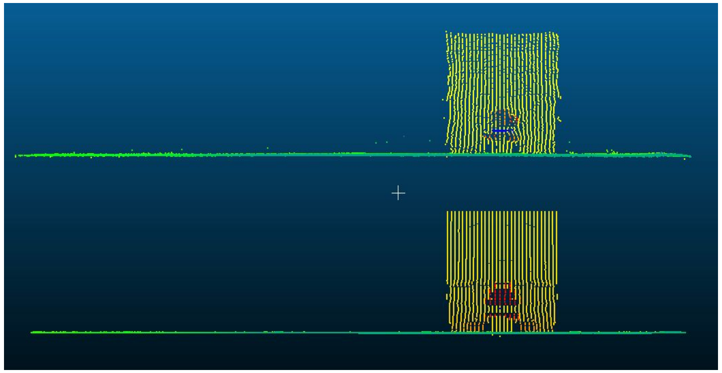


Pix2Pix output with Frobenius Norm along with Original Loss Functions

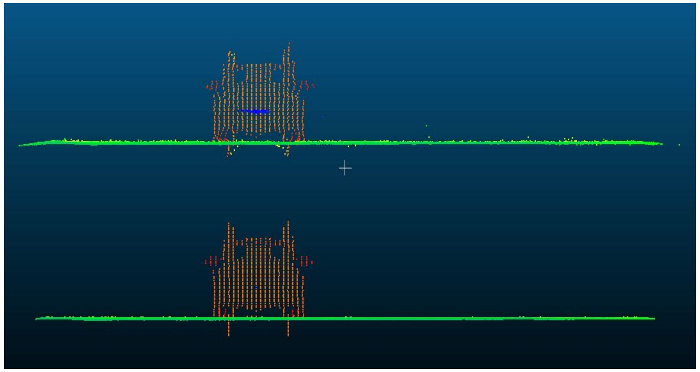
TOP: Prediction, BOTTOM: Ground Truth

Pix2PixHD: Outputs with Range Images Only

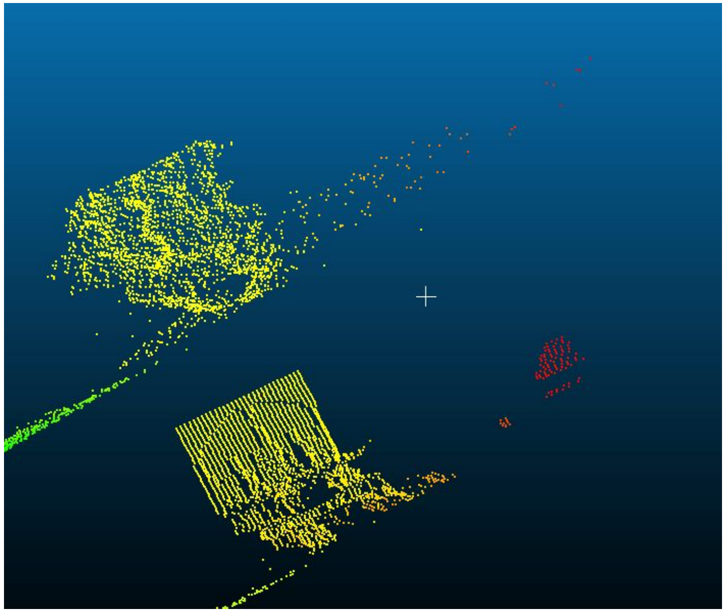
TOP: Prediction & Bottom: Ground Truth



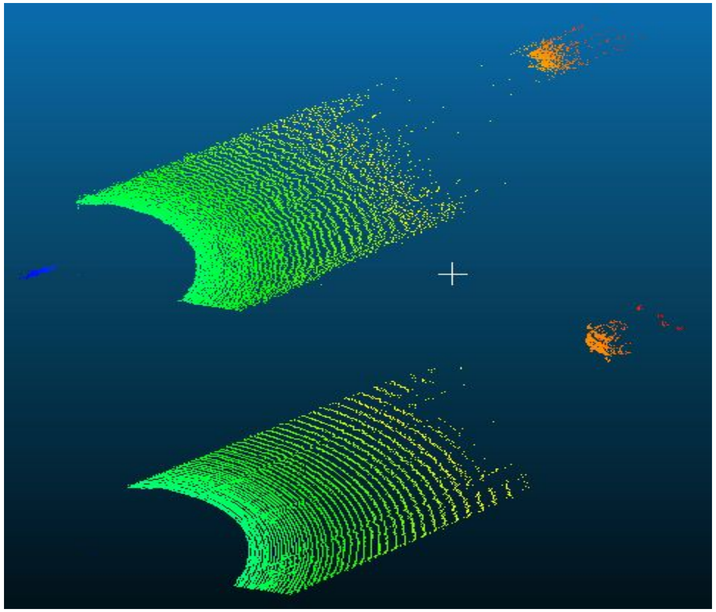
Ego view of the Truck



Ego view of the Car



Magnified Side view of the above Truck



Side view of the above trace

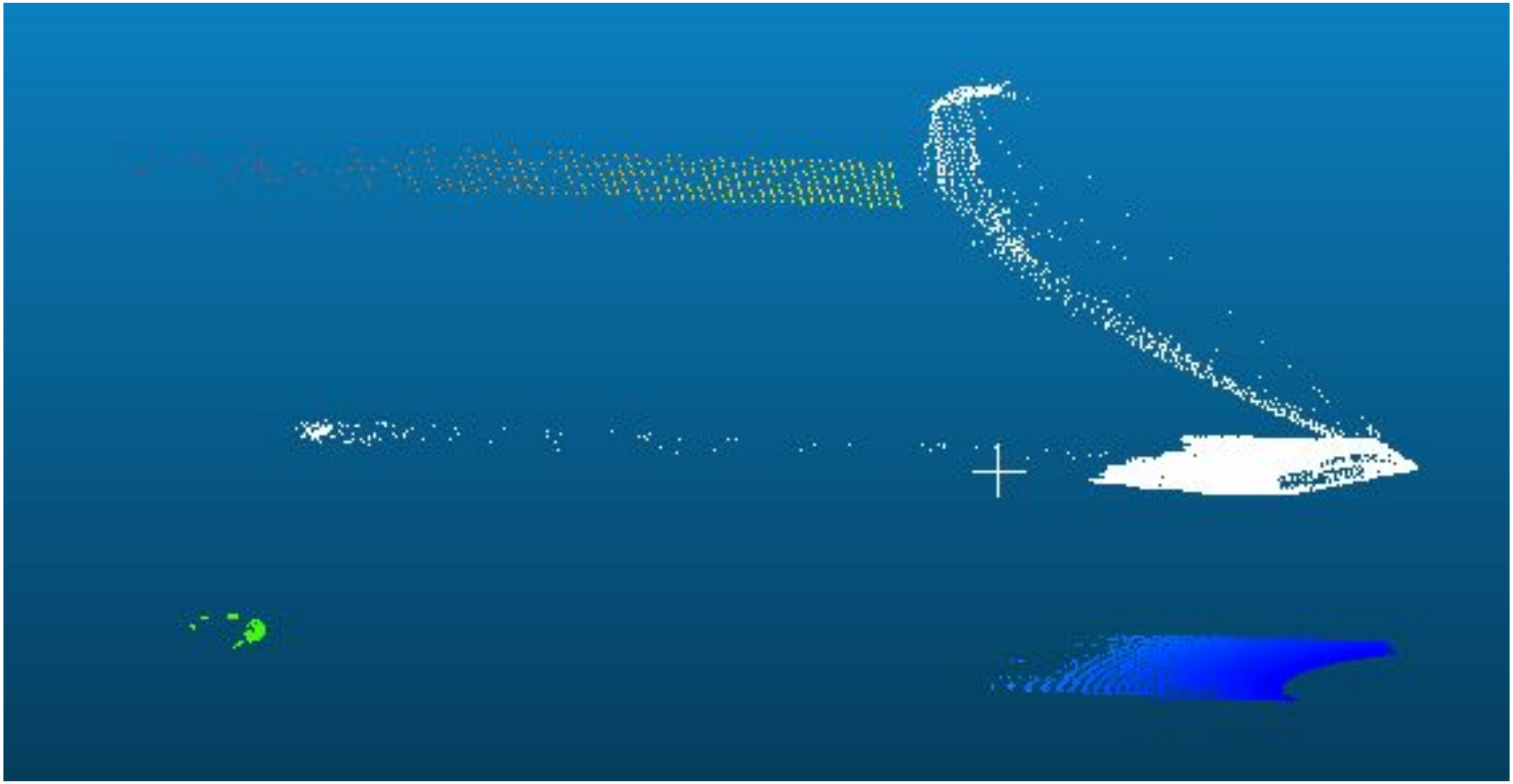
Pix2PixHD: Matrics

Loss Function Name	Hausdorff Distance ↓	Chamfer Distance ↓	RMSE ↓	Spatial Distri-bution Difference
Frobenius Norm + All Pix2PixHD Losses	0.431	0.0375	0.0304	0.268
Frobenius Norm Only	0.587	0.0687	0.085	0.549
Frobenius Norm + GAN Feat Loss (2 Channels)	0.285	0.024	0.111	0.156

3D Metrics for the Pix2PixHD Model

- Frobenius Norm improves the Output of the Pix2PixHD Model.
- Frobenius Norm alone is not effective.
- Frobenius Norm gives best results with GAN_Feat_Matching Loss.

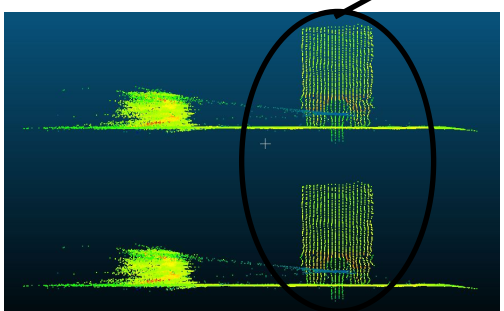
Frobenius Norm on U-Net



Top: Prediction of the U-Net model (White color points)

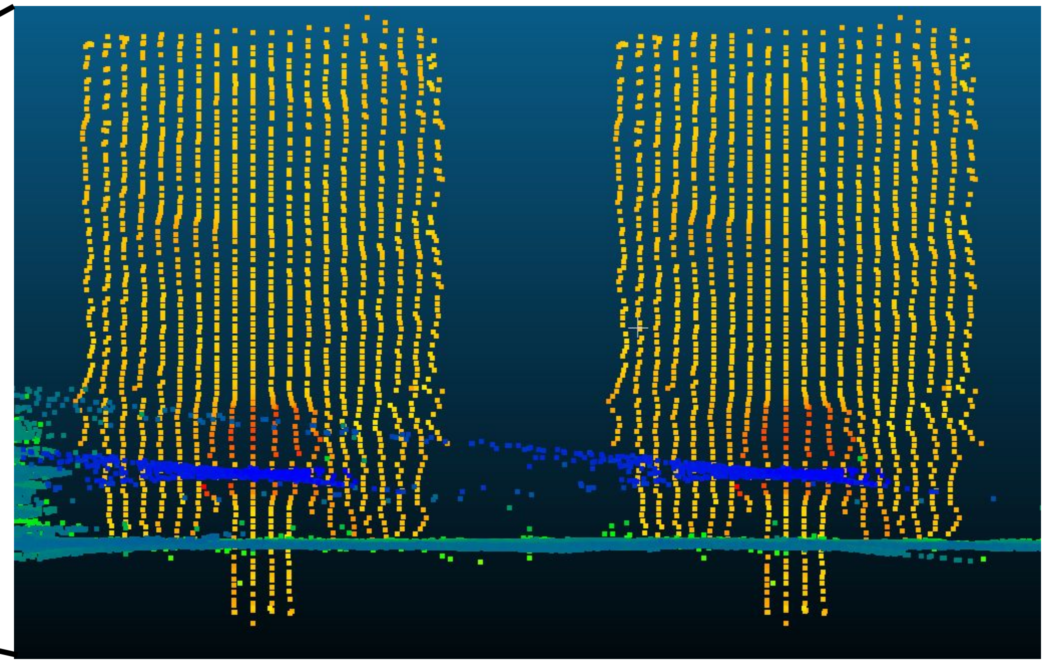
Bottom: Ground Truth (Other Color)

Pix2PixHD: Outputs: Range and Area



Top: Prediction

Bottom: Ground Truth

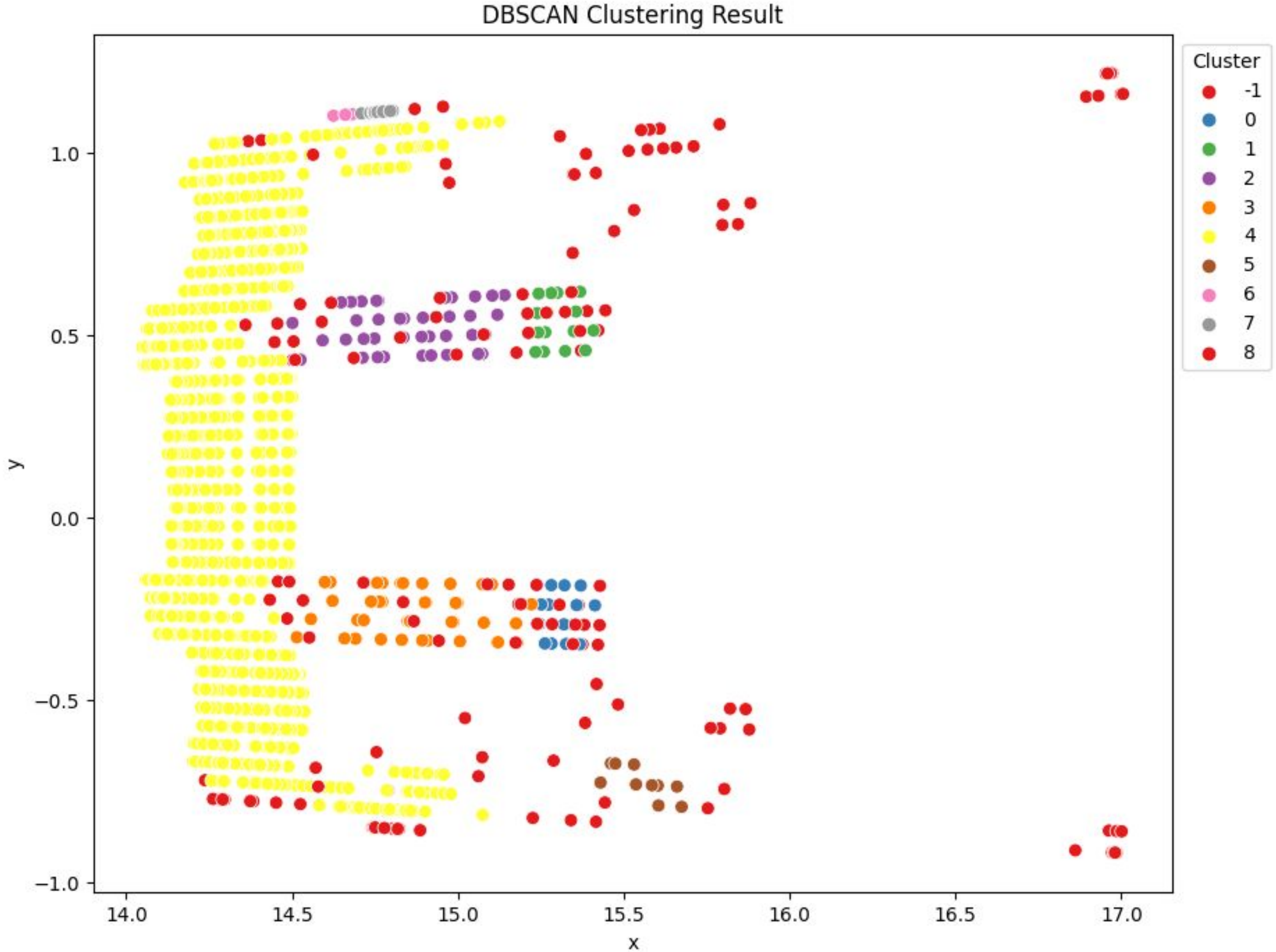


Left: Prediction

Right: Ground Truth

Output of the Pix2Pix model with 2 channels, range and Area. Left: Point Cloud, Right: Magnified view of the Truck.

DBSCAN Clustering on the output[6]



Clustering the output of the Pix2PixHD model for 2 channels, Range and Area

Conclusion

- Able to up-sample complex traffic objects using range images.
- Conducted extensive Literature survey for different techniques used for point cloud processing in 3D and 2D.
- Successfully trained two Deep Learning models and implemented various metrics to test their performance.
- Setup a system for creation and validation of the paired LiDAR point cloud dataset for up sampling.
- The model was successful in coping up with heterogeneous data from 2 different LiDAR sensors.
- Able to cluster the points based on the Area of the prediction thus able to identify material of the different objects in the point cloud.

Comparison U-Net v/s Pix2PixHD

Mode Name	Hausdorff Distance	Chamfer Distance	RMSE	Spatial Distribution of Points
U-Net	0.139	0.016	0.078	0.023
Pix2PixHD	0.285	0.024	0.111	0.152

Conclusion

Achievements:

- Complex Traffic Object Up-sampling (x8.5 times).
- Successfully accomplished Domain Adaptation.
- Deep Learning Model Training & Evaluation.
- Dataset Creation & Validation System.
- Material Identification via clustering.

Impact:

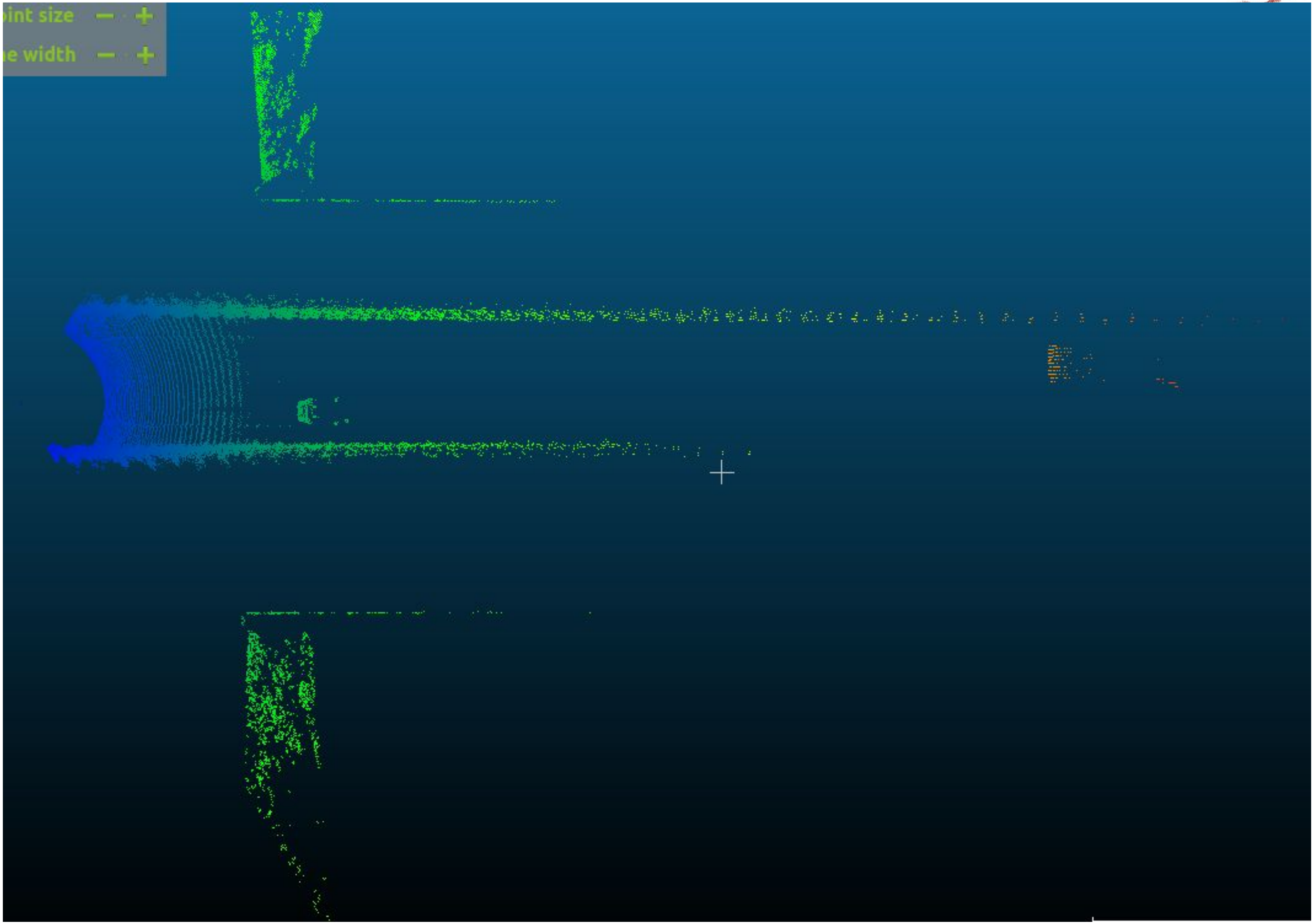
- Enhanced LiDAR Data Quality.
- Applicable to Diverse Domains.
- Improved Object Recognition.
- Material Composition Insights.

Future Directions:

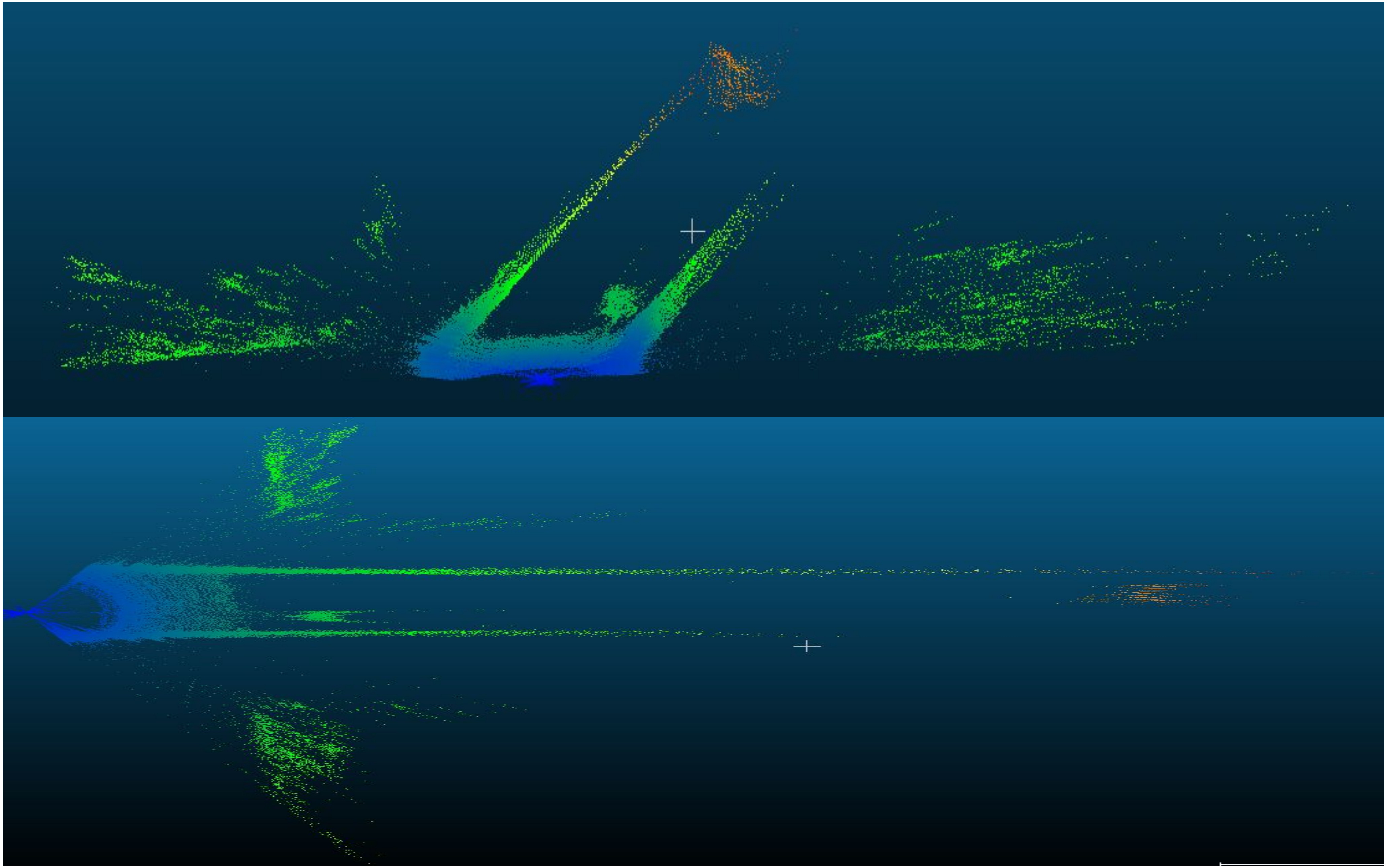
- Model Refinement
- Real-World Integration
- Wider Application Exploration

🙏Thank You

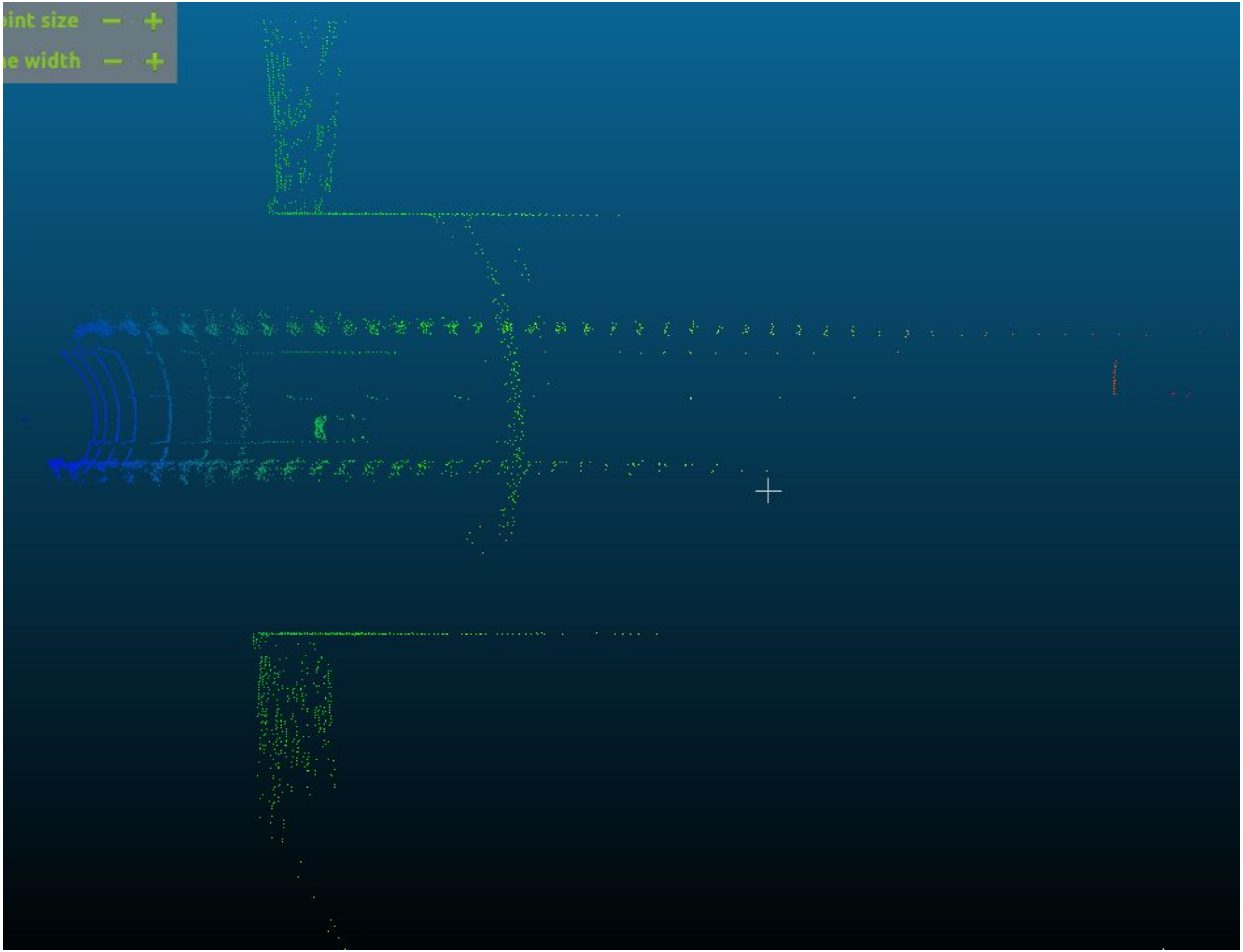
Sample Image 1: Scala Gen 3 Ground Truth



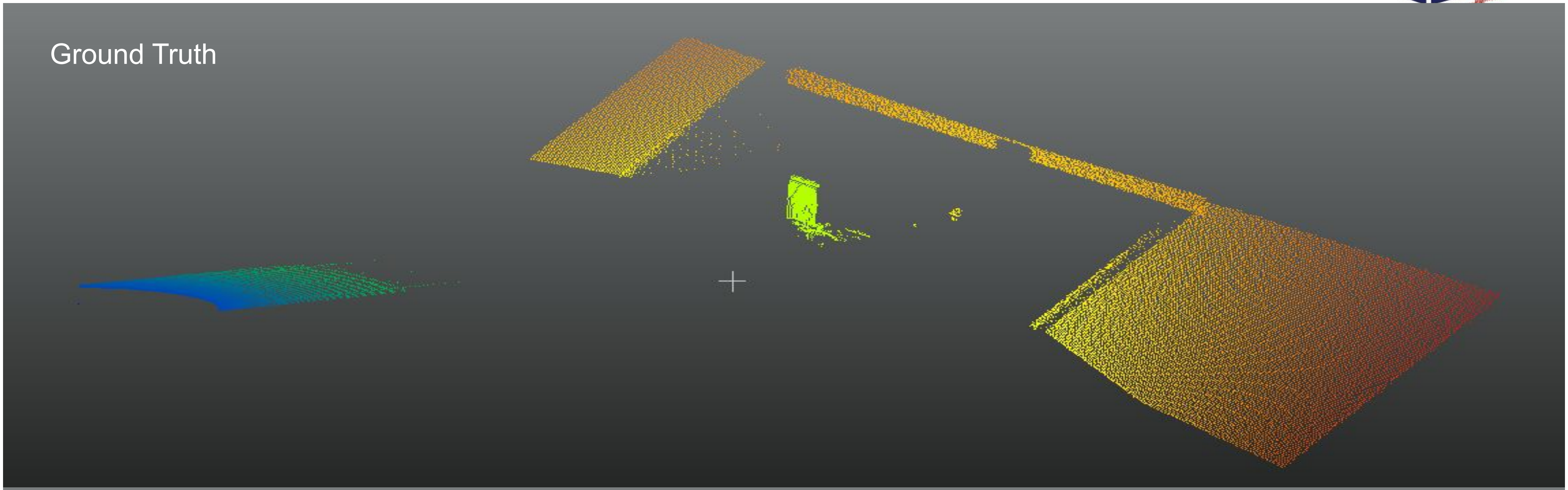
Sample Image 1: Scala Gen 3 Prediction

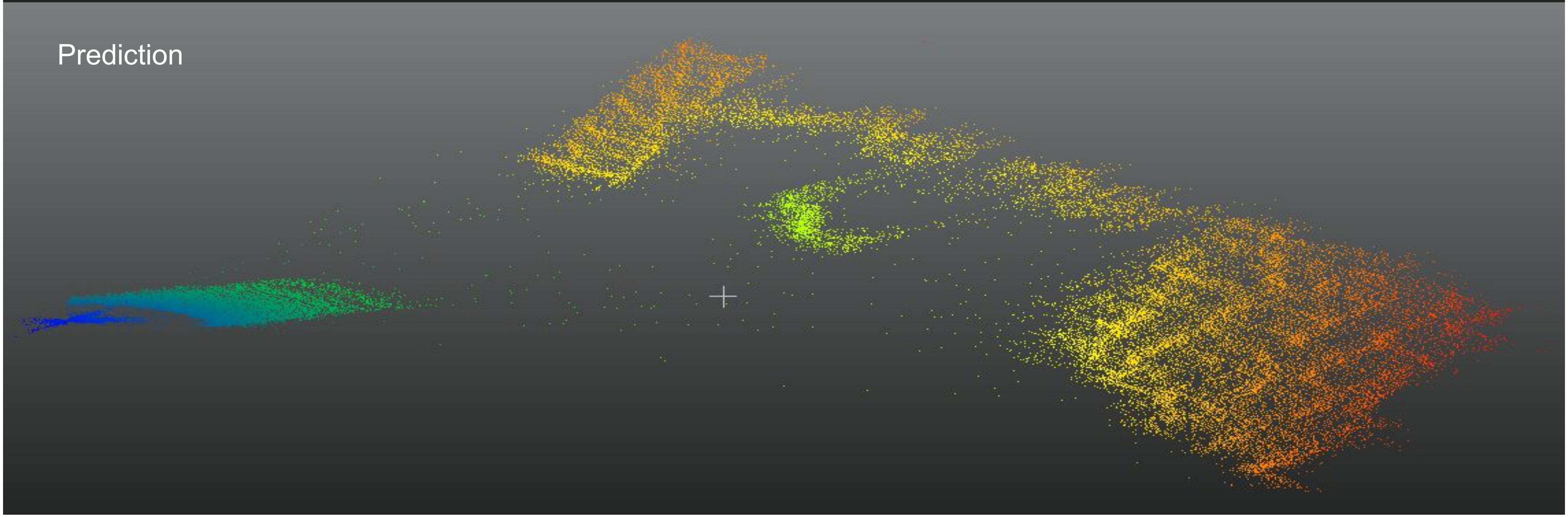


Sample Image 2: Scala Gen 2 (Input)



Sample Image 2: Scala Gen 3 GT & Prediction





References

1) https://autonomoustuff.com/-/media/Images/Hexagon/Hexagon%20Core/autonomousstuff/pdf/valeo-scala-gen-2-v17-datasheet-whitelabel.ashx?la=en&hash=3132D13FD3DF0446A785659CB0245F57
2) https://hexagondownloads.blob.core.windows.net/public/AutonomouStuff/wp-content/uploads/2020/10/valeo-scala-datasheet-whitelabel.pdf
3) https://arxiv.org/pdf/1711.11585.pdf
4) https://www.sciencedirect.com/science/article/abs/pii/S0921889020304875
5) https://www.sciencedirect.com/topics/mathematics/frobenius-norm
6) https://dl.acm.org/doi/10.5555/3001460.3001507
7) https://arxiv.org/abs/2007.10114