IRF GLOBAL | الهيئة العامة للطرق
Roads General Authority

Road Safety & Sustainability Conference

Sustainability AI Platforms for Roads Design, Construction, and Maintenance

Evaluating Semantic Segmentation-Based Scene Descriptions for Rural Highway Point Clouds

UNIVERSITY OF ALBERTA

Hesham Elmasry, MSc
Research Assistant
helmasry@ualberta.ca
University of Alberta

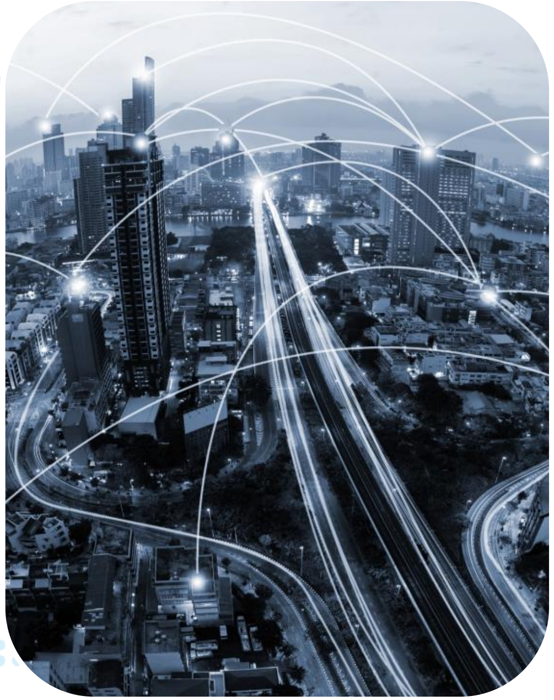
Honglin Jiang, MSc
Research Assistant
honglin@ualberta.ca
University of Alberta

Amr Sakr, MSc
Research Assistant
amsakr@ualberta.ca
University of Alberta

Karim El-Basyouny, PhD, PEng
Professor & Associate Dean
basyouny@ualberta.ca
University of Alberta

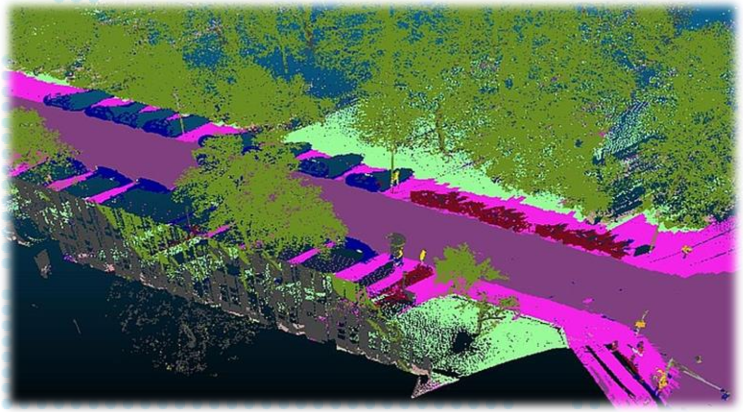
Background

- Highway scene description plays a crucial role in monitoring and managing infrastructure
- Efficient and accurate highway scene descriptions are essential for the development and deployment of autonomous driving technologies.
- The growing complexity of highway networks requires advanced tools to analyze and interpret vast amounts of data, reducing the need for manual inspection
- Scene description aids in compliance with transportation regulations and supports data-driven decision-making for infrastructure improvements



Motivation

- Automation and digitizing of rural highway infrastructure analysis is critical for improving safety, maintenance, and asset management
- Current methods of manual labeling, scene description, and report generation are time-consuming and prone to human error
- Transformer-based models show promise in classifying highway infrastructure components from point clouds, automating the process.
- Accurate scene descriptions can enhance the functionality of autonomous vehicles, improving their ability to navigate complex environments, and transportation infrastructure monitoring



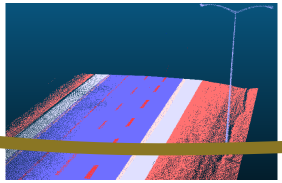
Research Gaps

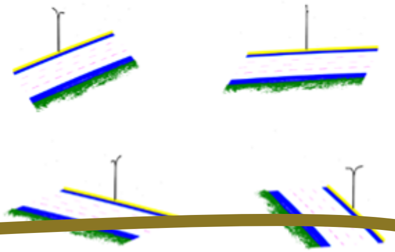
- Current models lack fully automated systems capable of generating accurate scene descriptions for complex rural highway environments.
- While current Transformer-based models show promise, further improvement is needed to enhance accuracy in classifying complex highway features extracted from LiDAR point clouds
- The suitable architecture for 3D semantic segmentation point cloud processing is still undiscovered

Research Objectives

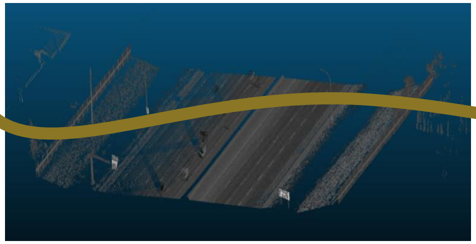
- Train and implement two Transformer-based models (Point Transformer v2 and a custom model with self-attention and cross-attention mechanisms) to label 11 distinct highway infrastructure elements
- Develop an automated process for generating descriptive text of rural highway scenes from semantically labeled point cloud data
- Employ a Generative AI model for Natural Language Processing (NLP) to produce accurate and detailed descriptions of highway segments
- Validate the accuracy of generated descriptions by comparing them against manual annotations of ground truth point clouds
- Evaluate the consistency and quality of descriptions produced by both models using NLP-based semantic similarity scores

General Framework





Unlabeled point cloud



ML Transformer models (CNN)

Labelled point cloud

Images generated & Description for each segment

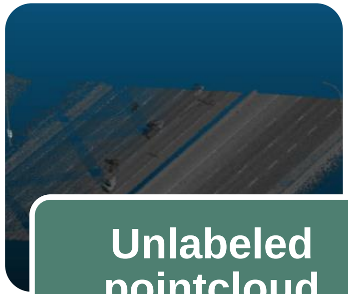
Computations for infrastructure assessment

Various scripts

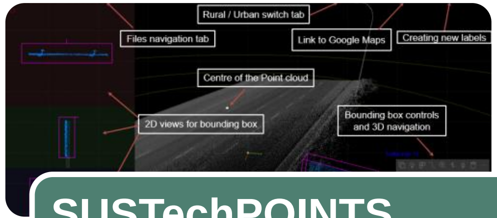
NLP/LLM

Qualitative & Quantitative Report

Detailed Framework

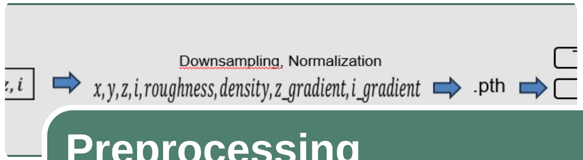


Unlabeled pointcloud form LiDAR



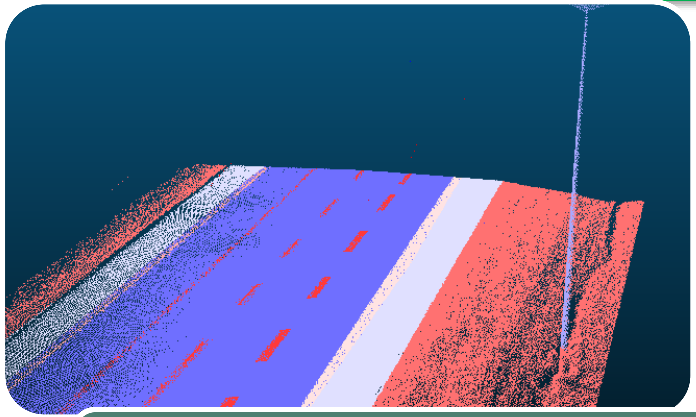
SUSTechPOINTS Annotation tool

- Creating labelled Ground Truth Dataset



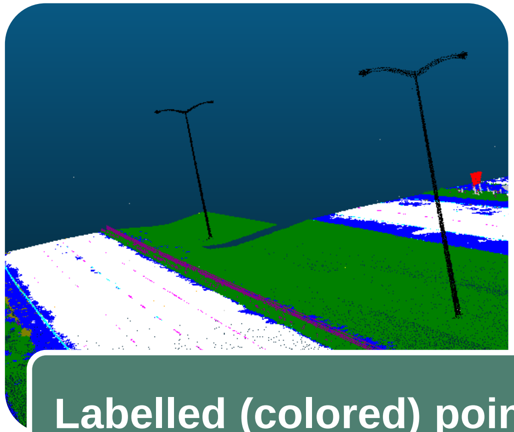
Preprocessing

- 50 meter segments
- Data augmentation / computation of extra features
- Down sampling / normalization
- Split: train, val, test



Semantic Segmentation ML models (2 models)

- Point Transformer v2
- Transformer-based Point Classification



Labelled (colored) point cloud

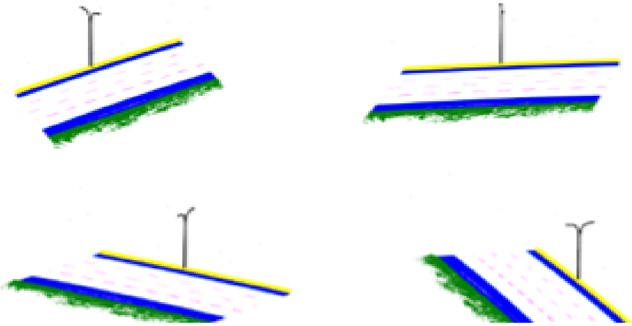
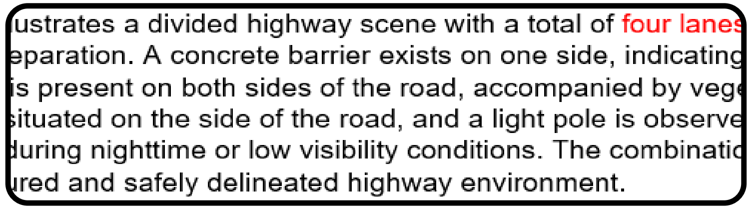


Image generation

- For each segment, several images are generated from different prespectives


ustrates a divided highway scene with a total of four lanes
eparation. A concrete barrier exists on one side, indicating
is present on both sides of the road, accompanied by veg
situated on the side of the road, and a light pole is observe
during nighttime or low visibility conditions. The combinatio
ured and safely delineated highway environment.

Scene Description using NLP models (2 models)

- Google Gemini
- GPT

Dataset

- Equipment: VMX-450 LiDAR scanner
- Collected Date: 2020
- Conditions: Regular traffic flow & speed up to 100km/h
- Data format: Compressed into 2 km sections LAZ files (x,y,z,i)

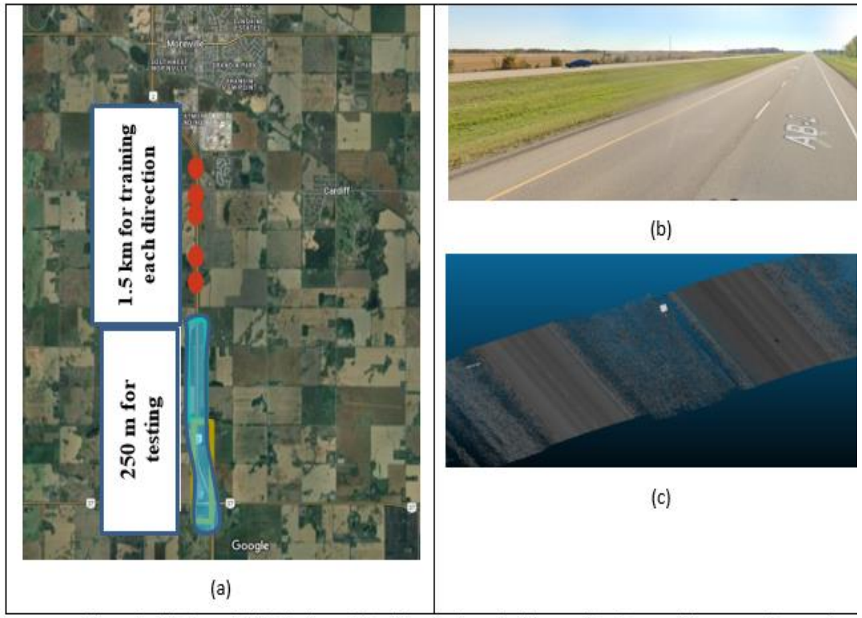


Figure 1. Highway 2 (AB-2) situated in Alberta, Canada. (a) a satellite image, (b) a street view, and (c) a point cloud sample representing a 50-meter segment

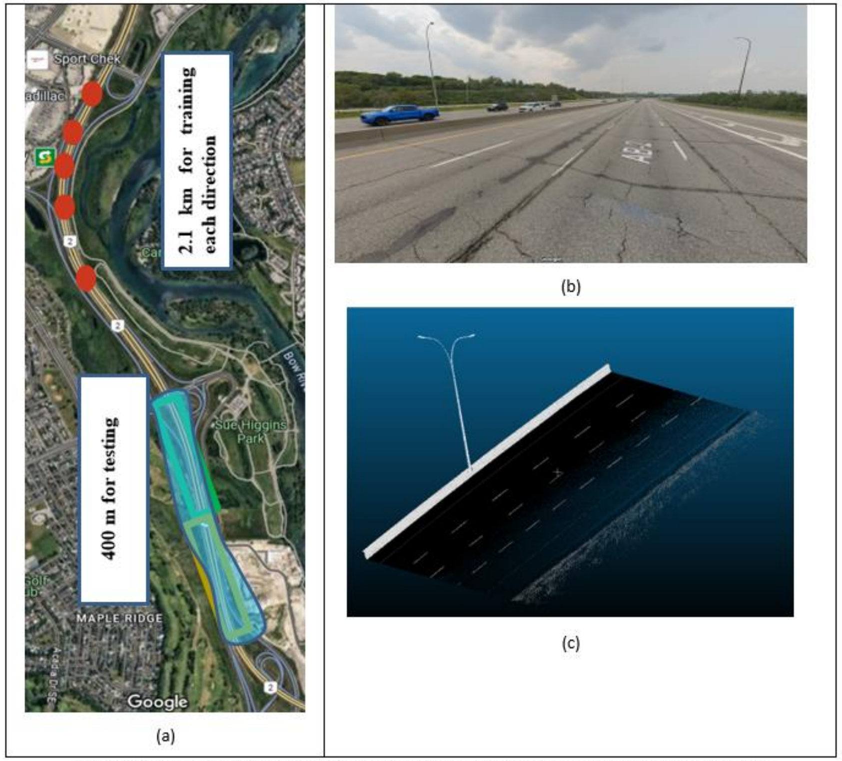


Figure 2. Deerfoot Trail situated in Alberta, Canada. (a) a satellite image, (b) a street view, and (c) a point cloud sample representing a 50-meter segment

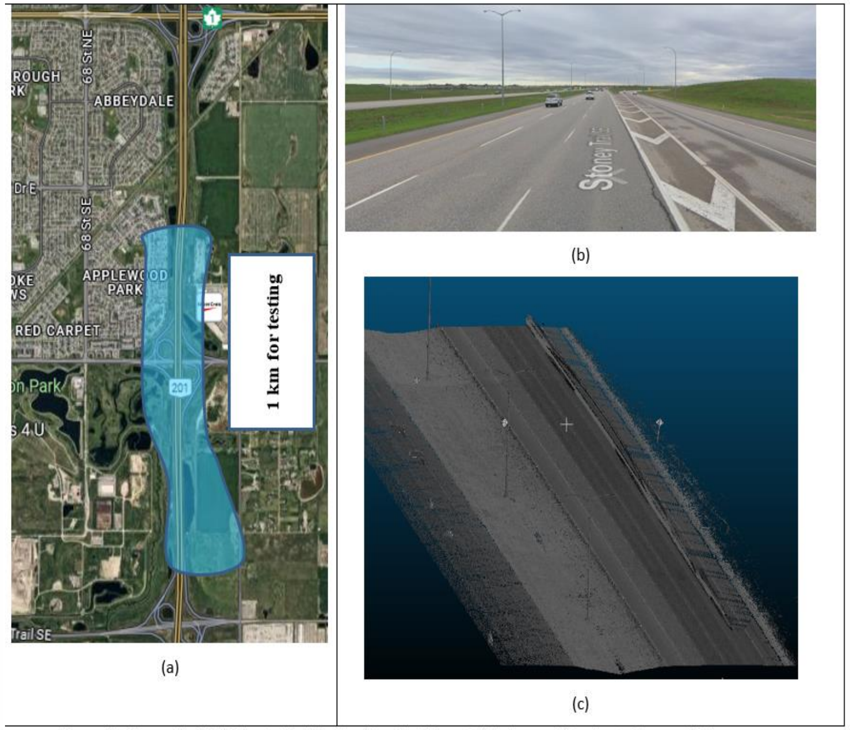
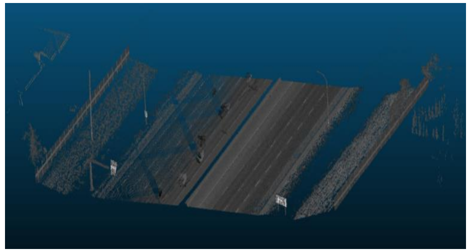
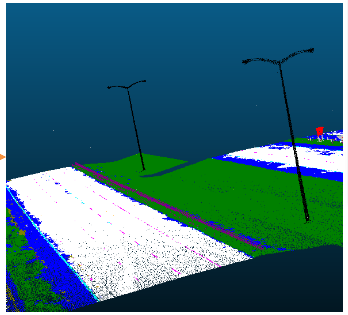


Figure 3. Stoney Trail SE situated in Alberta, Canada. (a) a satellite image, (b) a street view, and (c) a point cloud sample representing a 50-meter segment

Methodology Point Transformer v2 network





centres

labels

.las files

Class_name_# (.txt)
For example:
Solid-line_1.txt
Lane_1.txt

$\boxed{x, y, z, i}$

Downsampling, Normalization

$x, y, z, i, roughness, density, z\_gradient, i\_gradient$

.pth

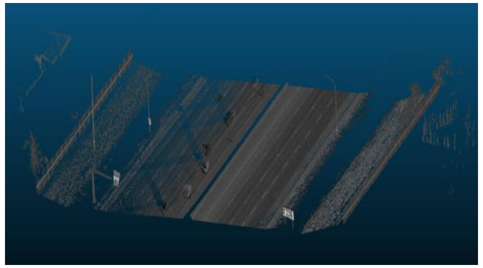
Train

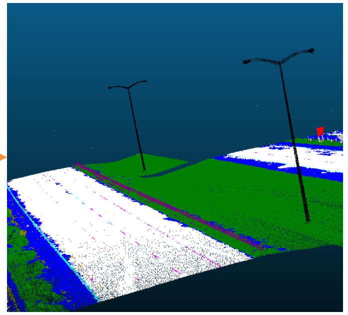
Val

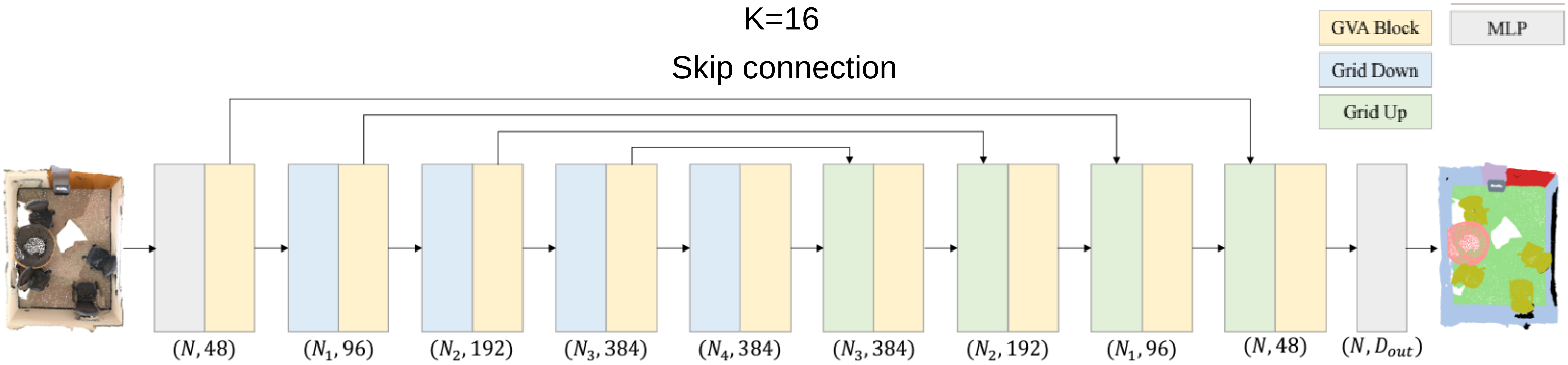
Test

Additional feature	Definition
Z-gradient	the gradient of the elevation of a point relative to its neighbours within a sphere of radius R
Roughness	the distance between a point and the best fitting plane to its neighbours within a sphere of radius R
Intensity gradient	the gradient of the intensity of a point relative to its neighbors within a sphere of radius R
Neighbour density	counts the number of neighbour points within a sphere of radius R

Methodology Point Transformer v2 network







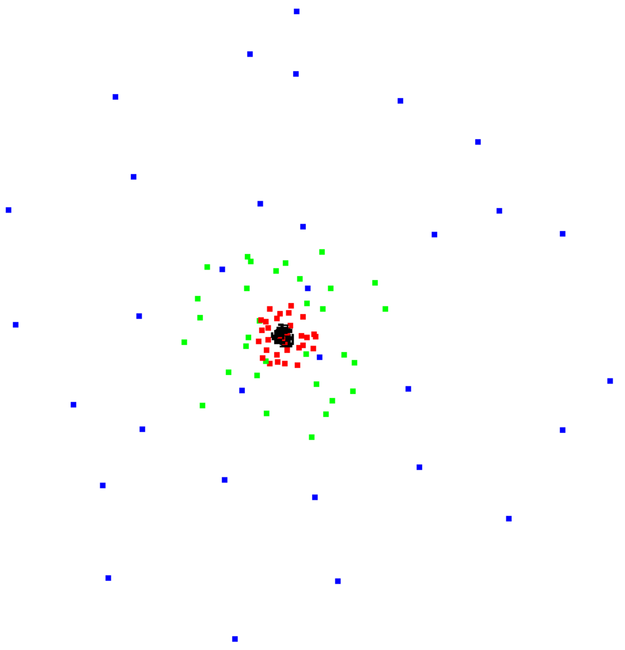
Point Transformer v2 networks for semantic segmentation

Encoding

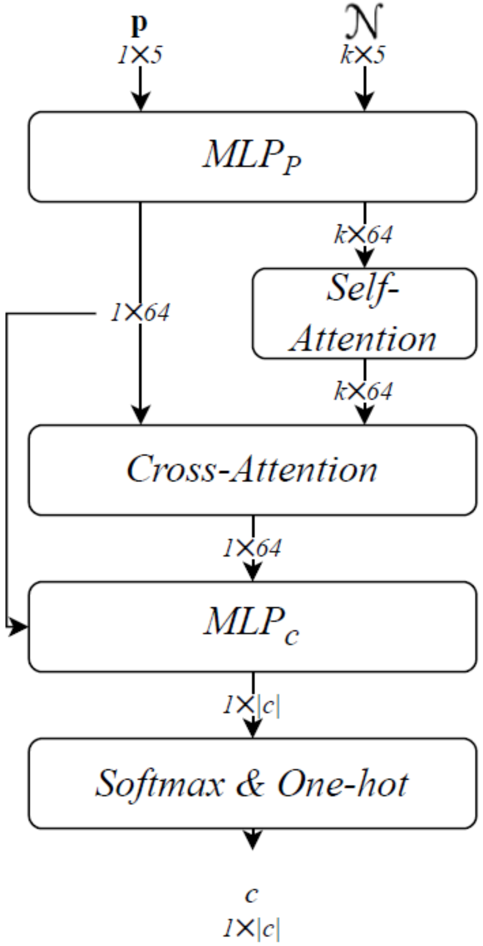
Decoding

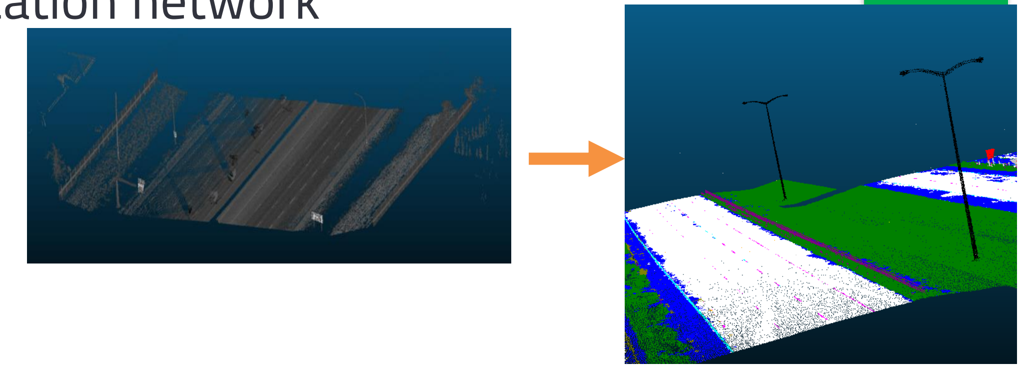
Methodology Transformer-based Point Classification network

K=32



0.1m, 0.3m, 1m, and 3m





$$\boxed{x, y, z, i, d}$$

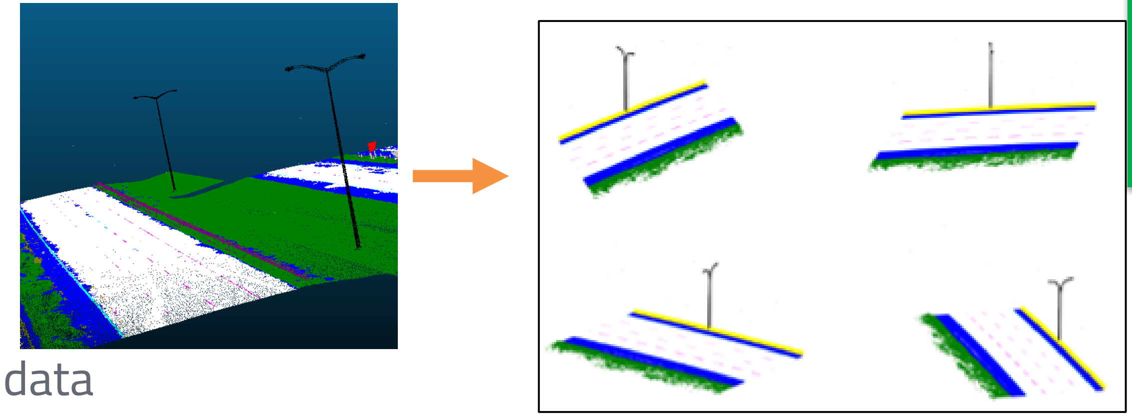
$$d = \frac{\#\ of\ points\ within\ voxel}{(voxel\ size * coefficient)^2}$$

$$coefficient = 50$$

$$N_{bal} = \frac{total\ \#\ of\ points\ for\ class\ x}{\sqrt{total\ \#of\ points\ for\ class\ x\ * coefficient}}$$

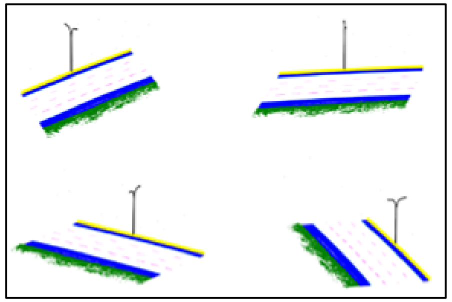
$$coefficient = 10$$

Methodology Image generation



- Utilize matplotlib to plot the 3D points from point cloud data
- The input data (colored point cloud) is stored in .txt files, where each row represents a point in 3D space, and columns contain values for x, y, z coordinates and possibly color data (r, g, b)
- The script reads each file using a custom function (read_point_cloud) that parses either space- or comma-delimited values into a structured array
- Generate four images for each input point cloud file, varying azimuth angles while keeping a fixed elevation angle
- Capture various perspectives to provide comprehensive visual representations of the scene

Methodology NLP models



The point cloud data illustrates a divided highway scene with a total of four lanes designated by three broken lines for lane separation. A concrete barrier exists on one side, indicating that the highway is divided. The shoulder is present on both sides of the road, accompanied by vegetation. Additionally, there are traffic signs situated on the side of the road, and a light pole is observed nearby, potentially providing illumination during nighttime or low visibility conditions. The combination of these elements suggests a well-structured and safely delineated highway environment.

Google Gemini

- Use Google's Gemini-1.5-pro model to generate narrative text descriptions based on visual features
- Focus on infrastructure elements (12 classes)
- Several prompts used: *Ex: "Describe ALL the divided multilane highway infrastructure elements that you can see in narrative paragraph text in the image among lanes, shoulders, .."*
- Each description file corresponds to the set of images derived from a specific point cloud file.

GPT

- Use GPT-4o LLM to generate text descriptions through view images
- One-shot Learning method is used to give model an example of the description " *The scene represents a multi-lane divided highway. A concrete barrier median separates the lanes for opposing traffic flows. Alongside the road, there is a light pole, and an overhead traffic sign provides guidance for drivers.*"
- Restrictive conditions: Role, Scenario Overview, Clarification, Additional information were specified

Methodology Evaluation parameters

Semantic Segmentation ML models prediction

- ✓ **Confusion matrix**
- ✓ **F1 score**
- ✓ **Accuracy**
- ✓ **mIoU**

$$Precision = \frac{TP_i}{TP_i + FP_i}$$

$$Recall = \frac{TP_i}{TP_i + FN_i}$$

$$IoU_i = \frac{TP_i}{TP_i + FP_i + FN_i}$$

$$F1\ score_i = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall}\right)$$

$$mIoU = \frac{1}{N}\sum_{i=1}^{N} IoU_i$$

$$mean\ F1\ score = \frac{1}{N}\sum_{i=1}^{N} F1\ score_i$$

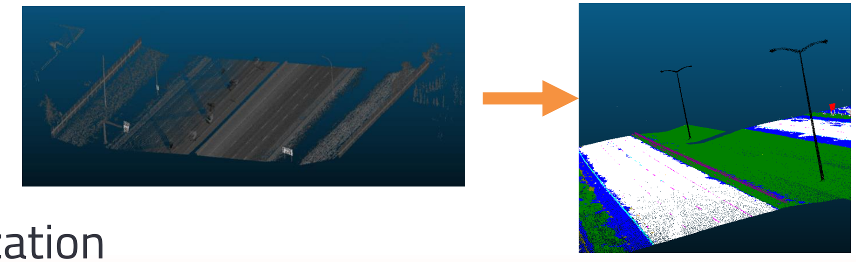
TP_i = the number of true positive predictions for the class i.
FP_i = the number of false positive predictions for the class i.
FN_i = the number of false negative predictions for the class i.
$F1\ score$ = the $F1\ score$ for each class i.
$Mean\ F1\ score$ = the mean F1 for all classes

NLP models

- ✓ **Semantic Similarity Score**
- BERT score
- BLEU score
- ROUGE

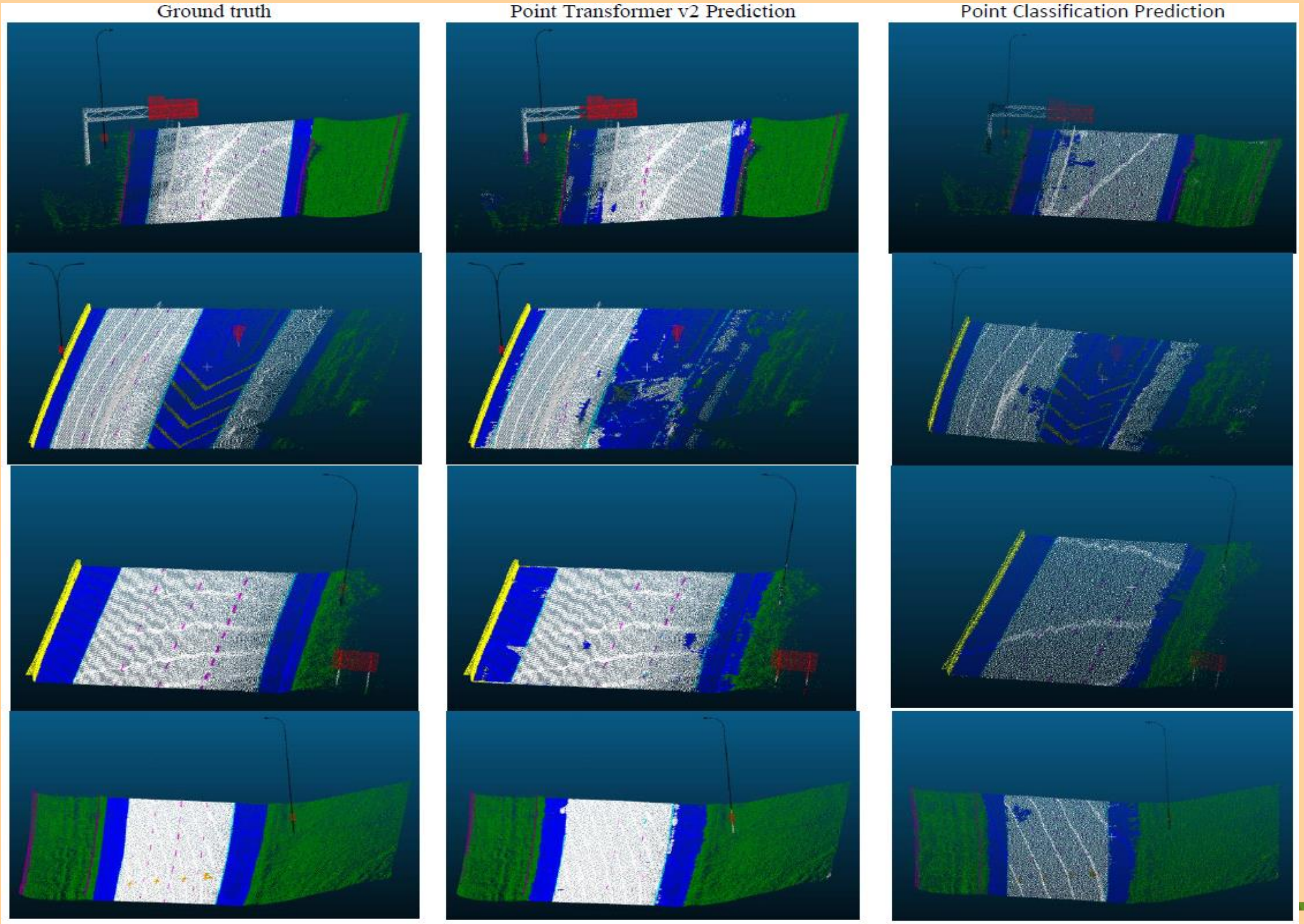
- Semantic similarity score measures how closely two pieces of text are in meaning. It quantifies the degree of resemblance between sentences, phrases, or documents by analyzing their underlying semantic content rather than their syntactic structure or exact words
- The semantic similarity score aims to capture meaning by comparing concepts, relationships, and the context of words in two texts. It goes beyond lexical similarity (word overlap) by considering words with similar meanings (synonyms) or shared contexts. For example, "car" and "vehicle" may not share surface-level similarity, but they have high semantic similarity because they represent similar concepts

See references

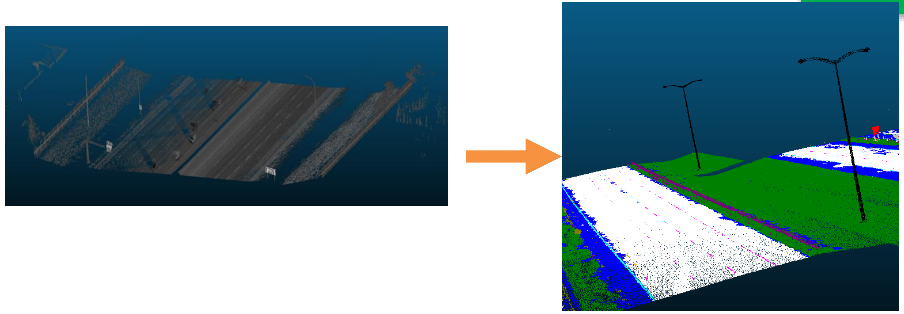


Results & Discussion Visualization

```
colors = {
    0: [255, 255, 255],  # white, lane
    1: [0, 0, 255],      # blue, shoulder
    2: [128, 128, 0],    # olive, chevrons
    3: [255, 0, 255],    # purple, broken-line
    4: [0, 255, 255],    # cyan, solid-line
    5: [255, 165, 0],    # orange, arrows
    6: [0, 128, 0],      # green, vegetation
    7: [255, 0, 0],      # red, traffic-sign
    8: [128, 0, 128],    # magenta, highway-guardrails
    9: [255, 255, 0],    # yellow, concrete-barriers
    10: [0, 0, 0],       # black, light-pole
    11: [192, 192, 192]  # silver, clutter
```



Results & Discussion Point Transformer v2



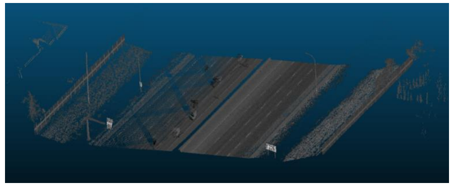
	lane	shoulder	chevrons	broken line	solid line	arrows	vegetation	traffic sign	highway guardrails	concrete barriers	light pole	clutter
lane	0.8946	0.0894	0.0000	0.0006	0.0050	0.0000	0.0025	0.0000	0.0001	0.0000	0.0000	0.0078
shoulder	0.3088	0.6363	0.0000	0.0001	0.0105	0.0000	0.0311	0.0000	0.0017	0.0032	0.0000	0.0082
chevrons	0.2148	0.6735	0.0000	0.0000	0.0761	0.0000	0.0270	0.0000	0.0000	0.0000	0.0000	0.0087
broken line	0.7422	0.0817	0.0000	0.1087	0.0611	0.0000	0.0012	0.0000	0.0000	0.0000	0.0000	0.0052
solid line	0.4134	0.2203	0.0000	0.0041	0.3479	0.0000	0.0084	0.0000	0.0000	0.0000	0.0000	0.0059
arrows	0.6551	0.3319	0.0000	0.0033	0.0098	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
vegetation	0.0323	0.1695	0.0000	0.0000	0.0003	0.0000	0.7433	0.0002	0.0059	0.0087	0.0003	0.0395
traffic sign	0.0000	0.0003	0.0000	0.0000	0.0000	0.0000	0.0079	0.9211	0.0077	0.0045	0.0160	0.0424
highway guardrails	0.0057	0.0056	0.0000	0.0000	0.0000	0.0000	0.0711	0.0003	0.5686	0.0042	0.0000	0.3445
concrete barriers	0.0101	0.0422	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.9441	0.0000	0.0036
light pole	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0178	0.0040	0.0212	0.0066	0.9248	0.0255
clutter	0.0403	0.0109	0.0000	0.0000	0.0002	0.0000	0.0161	0.0044	0.0044	0.0048	0.0088	0.9101

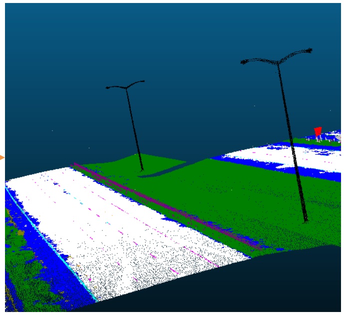
Mean accuracy	78.84%
Mean F1score	56.35%
mIoU	46.16%

Results & Discussion

Transformer-based Point Classification

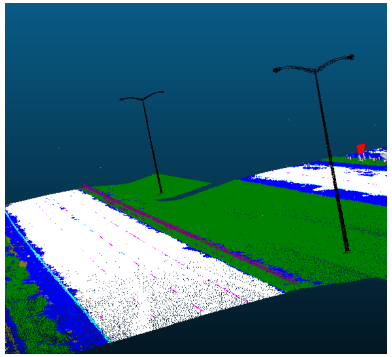
	arrows	broken line	chevrons	clutter	concrete barriers	highway guardrails	lane	light pole	shoulder	solid line	traffic sign	vegetation
arrows	0.5348	0.0125	0.1642	0.0000	0.0000	0.0000	0.1117	0.0000	0.0525	0.1243	0.0000	0.0000
broken line	0.0028	0.3557	0.0253	0.0012	0.0000	0.0000	0.3675	0.0000	0.0717	0.1755	0.0000	0.0003
chevrons	0.0000	0.0000	0.6601	0.0000	0.0000	0.0000	0.0021	0.0000	0.2196	0.0294	0.0000	0.0888
clutter	0.0003	0.0003	0.0022	0.8934	0.0016	0.0120	0.0205	0.0084	0.0073	0.0055	0.0198	0.0286
concrete barriers	0.0000	0.0000	0.0000	0.0004	0.9871	0.0000	0.0000	0.0005	0.0115	0.0004	0.0000	0.0001
highway guardrails	0.0000	0.0000	0.0000	0.0023	0.0000	0.8708	0.0000	0.0001	0.0019	0.0000	0.0003	0.1246
lane	0.0009	0.0013	0.0039	0.0038	0.0001	0.0002	0.8900	0.0000	0.0967	0.0021	0.0000	0.0011
light pole	0.0000	0.0000	0.0000	0.0027	0.0006	0.0000	0.0000	0.9921	0.0000	0.0000	0.0030	0.0015
shoulder	0.0002	0.0000	0.0169	0.0002	0.0021	0.0005	0.2085	0.0000	0.7110	0.0027	0.0000	0.0577
solid line	0.0141	0.0693	0.0669	0.0012	0.0000	0.0000	0.3487	0.0002	0.0820	0.4133	0.0000	0.0043
traffic sign	0.0000	0.0000	0.0000	0.0285	0.0000	0.0004	0.0000	0.0319	0.0000	0.0000	0.9388	0.0004
vegetation	0.0000	0.0000	0.0008	0.0004	0.0002	0.0025	0.0006	0.0008	0.0180	0.0004	0.0000	0.9763



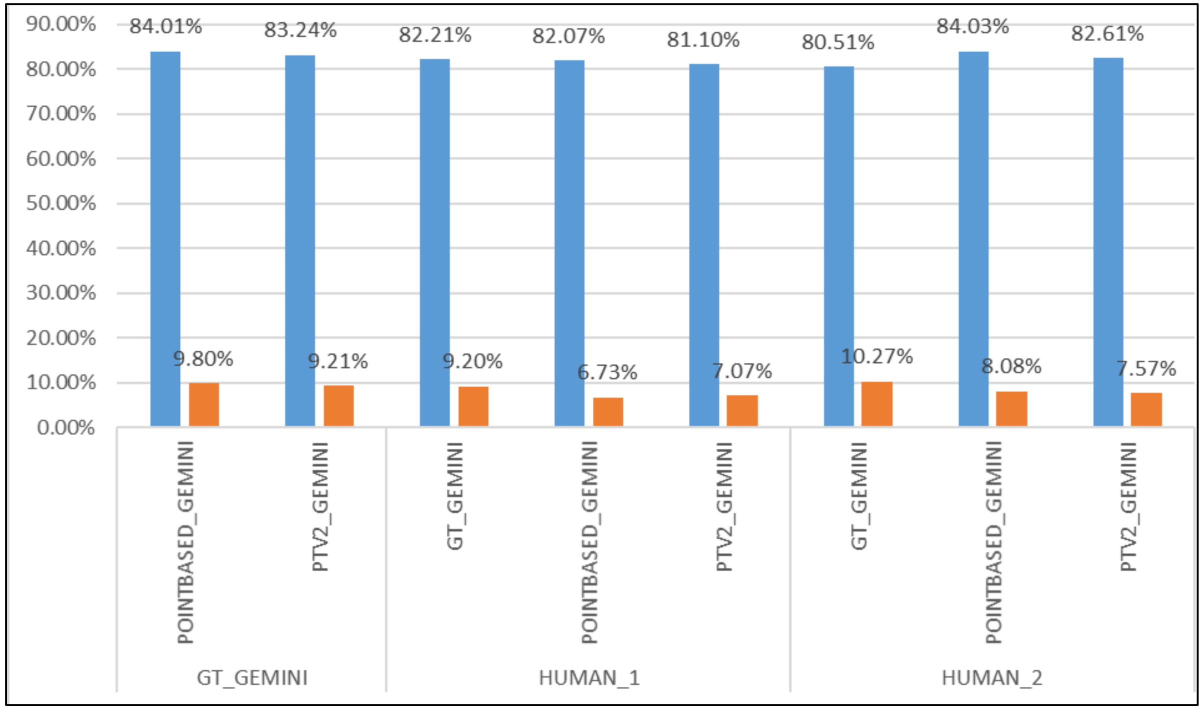


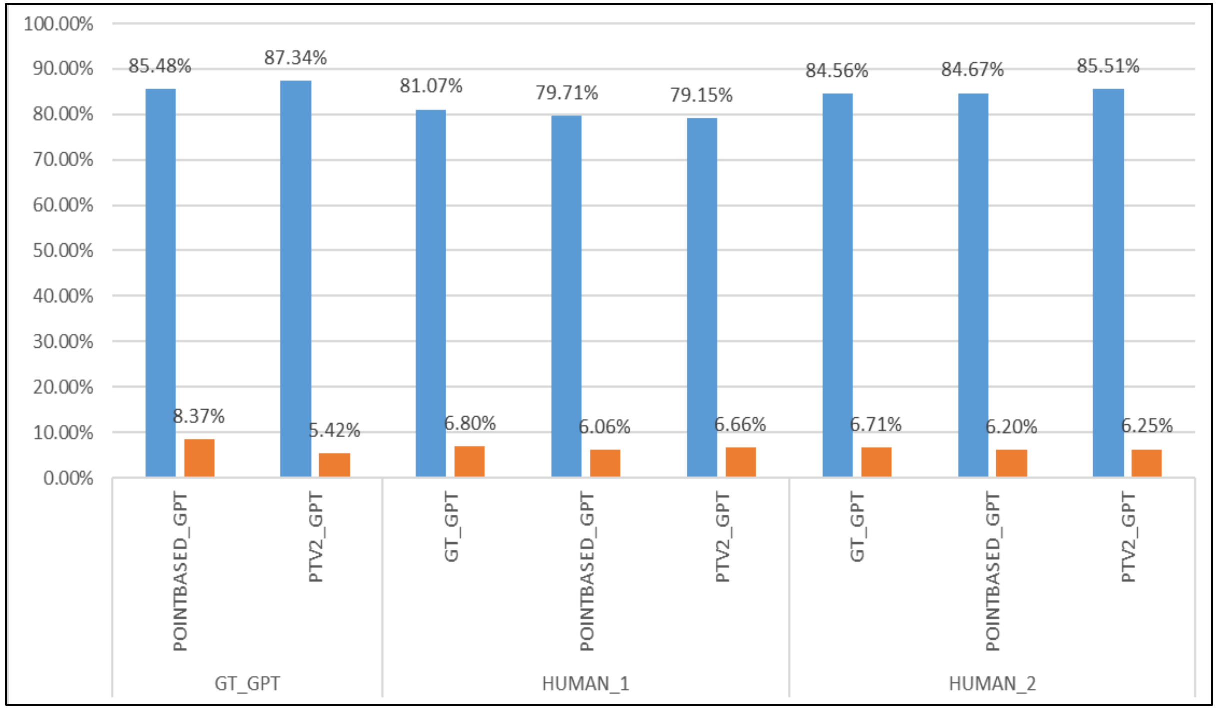
Mean accuracy	88.35%
Mean F1score	70.56%
mIoU	60.95%

Results & Discussion NLP models



The point cloud data illustrates a divided highway scene with a total of four lanes designated by three broken lines for lane separation. A concrete barrier exists on one side, indicating that the highway is divided. The shoulder is present on both sides of the road, accompanied by vegetation. Additionally, there are traffic signs situated on the side of the road, and a light pole is observed nearby, potentially providing illumination during nighttime or low visibility conditions. The combination of these elements suggests a well-structured and safely delineated highway environment.





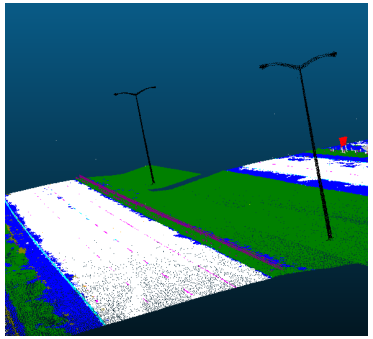
Average of AVG SEMANTIC SIMILARITY
Average of STD SEMANTIC SIMILARITY

Google Gemini

GPT

ROAD SAFETY & SUSTAINABILITY CONFERENCE

Results & Discussion NLP models

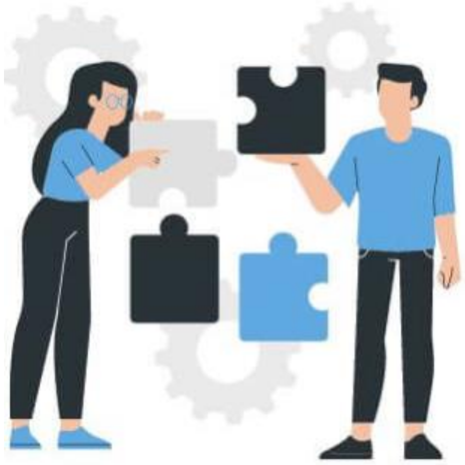


The point cloud data illustrates a divided highway scene with a total of four lanes designated by three broken lines for lane separation. A concrete barrier exists on one side, indicating that the highway is divided. The shoulder is present on both sides of the road, accompanied by vegetation. Additionally, there are traffic signs situated on the side of the road, and a light pole is observed nearby, potentially providing illumination during nighttime or low visibility conditions. The combination of these elements suggests a well-structured and safely delineated highway environment.

NLP model	Scenario	Normality Conclusion	Test Type	p-value	Overall Conclusion
Gemini	H1 vs ptv2 and PB	Model 1 data is not normally distributed (p-value=0.0373). Model 2 data is normally distributed (p-value=0.1255).	Wilcoxon signed-rank test	0.458006509	No significant difference between the two models.
	H2 vs ptv2 and PB	Model 1 data is not normally distributed (p-value=0.0115). Model 2 data is not normally distributed (p-value=0.0013).	Wilcoxon signed-rank test	0.608816501	No significant difference between the two models.
	GT vs ptv2 and PB	Model 1 data is normally distributed (p-value=0.1209). Model 2 data is not normally distributed (p-value=0.0006).	Wilcoxon signed-rank test	0.659182065	No significant difference between the two models.
GPT	H1 vs ptv2 and PB	Model 1 data is not normally distributed (p-value=0.0093). Model 2 data is normally distributed (p-value=0.0561).	Wilcoxon signed-rank test	0.479795904	No significant difference between the two models.
	H2 vs ptv2 and PB	Model 1 data is not normally distributed (p-value=0.0000). Model 2 data is not normally distributed (p-value=0.0006).	Wilcoxon signed-rank test	0.271868427	No significant difference between the two models.
	GT vs ptv2 and PB	Model 1 data is not normally distributed (p-value=0.0074). Model 2 data is not normally distributed (p-value=0.0002).	Wilcoxon signed-rank test	0.154726336	No significant difference between the two models.

Conclusions

- Transformer-based Point Classification network outperformed Point Transformer v2 in semantic segmentation predictions
- No significant difference between scene descriptions generated from Transformer-based Point Classification network and Point Transformer v2
- Google Gemini and GPT models generated very similar descriptions



Research Contributions

- Novel application of Transformer-based models for rural highway scene description
- Integration of NLP with 3D point cloud data
- Development of a scalable, automated framework for highway infrastructure analysis

Limitations & Future work

- Occluded object reconstruction
- Dataset expansion and analysis of moving vehicles
- Labelled point clouds computations (quantitative analysis)
- Improve processing time & pipeline the whole process

Acknowledgment

- **Sangwon Lim**
- **Aaditya Baruah**
- **Arindam Barman**
- **Vamshika Sutar**

References

- Agarwal, A., & Lavie, A. (2008). Meteor, m-bleu and m-ter: Evaluation metrics for high-correlation with human rankings of machine translation output. Proceedings of the Third Workshop on Statistical Machine Translation, 115–118.
- Altınel, B., & Ganiz, M. C. (2018). Semantic text classification: A survey of past and recent advances. Information Processing & Management, 54(6), 1129–1153. https://doi.org/https://doi.org/10.1016/j.ipm.2018.08.001
- Banerjee, S., & Lavie, A. (2005). METEOR: An automatic metric for MT evaluation with improved correlation with human judgments. Proceedings of the Acl Workshop on Intrinsic and Extrinsic Evaluation Measures for Machine Translation and/or Summarization, 65–72.
- Camacho-Collados, J., & Pilehvar, M. T. (2018). From word to sense embeddings: A survey on vector representations of meaning. Journal of Artificial Intelligence Research, 63, 743–788.
- Chandrasekaran, D., & Mago, V. (2021). Evolution of Semantic Similarity—A Survey. ACM Comput. Surv., 54(2). https://doi.org/10.1145/3440755
- Denkowski, M., & Lavie, A. (2014). Meteor universal: Language specific translation evaluation for any target language. Proceedings of the Ninth Workshop on Statistical Machine Translation, 376–380.
- Doddington, G. (2002). Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. Proceedings of the Second International Conference on Human Language Technology Research, 138–145.
- Guo, Y., & Hu, J. (2019). Meteor++ 2.0: Adopt syntactic level paraphrase knowledge into machine translation evaluation. Proceedings of the Fourth Conference on Machine Translation (Volume 2: Shared Task Papers, Day 1), 501–506.
- Hadj Taieb, M. A., Zesch, T., & Ben Aouicha, M. (2020). A survey of semantic relatedness evaluation datasets and procedures. Artificial Intelligence Review, 53(6), 4407–4448. https://doi.org/10.1007/s10462-019-09796-3
- Johnson, J., Karpathy, A., & Fei-Fei, L. (2016). Densecap: Fully convolutional localization networks for dense captioning. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 4565–4574.
- Jun Chen, Han Guo, Boyang Li, & Mohamed Elhoseiny. (2022). VisualGPT: Data-efficient Adaptation of Pretrained Language Models for Image Captioning. ArXiv Preprint ArXiv:2203.02637.
- Krishna, R., Zhu, Y., Groth, O., Johnson, J., Hata, K., Kravitz, J., Chen, S., Kalantidis, Y., Li, L.-J., Shamma, D. A., Bernstein, M. S., & Fei-Fei, L. (2017). Visual Genome: Connecting Language and Vision Using Crowdsourced Dense Image Annotations. International Journal of Computer Vision, 123(1), 32–73. https://doi.org/10.1007/s11263-016-0981-7
- Lastra-Díaz, J. J., & García-Serrano, A. (2015). A new family of information content models with an experimental survey on WordNet. Knowledge-Based Systems, 89, 509–526. https://doi.org/https://doi.org/10.1016/j.knosys.2015.08.019
- Lastra-Díaz, J. J., Goikoetxea, J., Hadj Taieb, M. A., García-Serrano, A., Ben Aouicha, M., & Agirre, E. (2019). A reproducible survey on word embeddings and ontology-based methods for word similarity: Linear combinations outperform the state of the art. Engineering Applications of Artificial Intelligence, 85, 645–665. https://doi.org/https://doi.org/10.1016/j.engappai.2019.07.010
- Leusch, G., Ueffing, N., & Ney, H. (2006). CDER: Efficient MT evaluation using block movements. 11th Conference of the European Chapter of the Association for Computational Linguistics, 241–248.
- Nießen, S., Och, F. J., Leusch, G., & Ney, H. (2000). An Evaluation Tool for Machine Translation: Fast Evaluation for MT Research. LREC.
- Papineni, K., Roukos, S., Ward, T., & Zhu, W.-J. (2002). Bleu: a method for automatic evaluation of machine translation. Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, 311–318.
- Qi, Z., Fang, Y., Sun, Z., Wu, X., Wu, T., Wang, J., Lin, D., & Zhao, H. (2024). Gpt4point: A unified framework for point-language understanding and generation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 26417–26427.
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., & Clark, J. (2021). Learning transferable visual models from natural language supervision. International Conference on Machine Learning, 8748–8763.
- Saadany, H., & Orasan, C. (2021). BLEU, METEOR, BERTScore: Evaluation of Metrics Performance in Assessing Critical Translation Errors in Sentiment-oriented Text. CoRR, abs/2109.14250. https://arxiv.org/abs/2109.14250
- Snover, M., Dorr, B., Schwartz, R., Micciulla, L., & Makhoul, J. (2006). A study of translation edit rate with targeted human annotation. Proceedings of the 7th Conference of the Association for Machine Translation in the Americas: Technical Papers, 223–231.
- Tillmann, C., Vogel, S., Ney, H., Zubiaga, A., & Sawaf, H. (1997). Accelerated DP based search for statistical translation. Eurospeech, 2667–2670.
- Tumuluru, A. K., Lo, C., & Wu, D. (2012). Accuracy and robustness in measuring the lexical similarity of semantic role fillers for automatic semantic MT evaluation. Proceedings of the 26th Pacific Asia Conference on Language, Information, and Computation, 574–581.
- Vinyals, O., Toshev, A., Bengio, S., & Erhan, D. (2015). Show and tell: A neural image caption generator. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3156–3164.
- Xu, D., Zhu, Y., Choy, C. B., & Fei-Fei, L. (2017). Scene graph generation by iterative message passing. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 5410–5419.
- Xu, K., Ba, J., Kiros, R., Cho, K., Courville, A., Salakhudinov, R., Zemel, R., & Bengio, Y. (2015). Show, attend and tell: Neural image caption generation with visual attention. International Conference on Machine Learning, 2048–2057.
- Xu, R., Wang, X., Wang, T., Chen, Y., Pang, J., & Lin, D. (2023). Pointllm: Empowering large language models to understand point clouds. ArXiv Preprint ArXiv:2308.16911.
- Yang, H., Li, S., Yin, X., Han, A., & Zhang, J. (2017). Recurrent Highway Networks with Attention Mechanism for Scene Text Recognition. 2017 International Conference on Digital Image Computing: Techniques and Applications (DICTA), 1–8. https://doi.org/10.1109/DICTA.2017.8227484

Hesham Elmasry, MSc
Research Assistant
helmasry@ualberta.ca
University of Alberta

Honglin Jiang, MSc
Research Assistant
honglin@ualberta.ca
University of Alberta

Amr Sakr, MSc
Research Assistant
amsakr@ualberta.ca
University of Alberta

Karim El-Basyouny, PhD, PEng
Professor & Associate Dean
basyouny@ualberta.ca
University of Alberta

IRF GLOBAL | الهيئة العامة للطرق Roads General Authority

THANK YOU