Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Ciência da Computação

Impacto da Compressão de Nuvem de Pontos em Algoritmos de Segmentação Semântica

Tiago de S. Fernandes

Monografia apresentada como requisito parcial
para conclusão do Curso de Engenharia da Computação

Orientador
Prof. Dr. Ricardo L. de Queiroz

Brasília
2024

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Ciência da Computação

Impacto da Compressão de Nuvem de Pontos em Algoritmos de Segmentação Semântica

Tiago de S. Fernandes

Monografia apresentada como requisito parcial
para conclusão do Curso de Engenharia da Computação

Prof. Dr. Ricardo L. de Queiroz (Orientador)
CIC/UnB

Prof. Dr. Diogo C. Garcia	Prof. Dr. Camilo C. Dórea
FGA/UnB	CIC/UnB

Prof. Dr. João Luiz A. de Carvalho
Coordenadora do Curso de Engenharia da Computação

Brasília, 5 de agosto de 2024

Agradecimentos

Agradeço o grupo de estudos UnBalloon, com quem estudei e competi em maratonas que me tornaram um estudante melhor, em especial ao meu companheiro de time e amigo, Alberto Duarte. Agradeço também a Universidade de Brasília, que me ofereceu ensino gratuito e de qualidade, pelo qual eu não teria condições de pagar. Por último, agradeço aos meus professores (Adriana Costa, Júlio Vargas, Guilherme Ramos, Ricardo Queiroz...), que me incentivaram na busca por conhecimento, aos meus amigos (Cristian Jonas, Rafaela Oliveira, Leonardo Riether, Gabriel Faustino, Euler Torres, Mateus Junqueira, Luiza Velozo...), pelos momentos de descontração, e à minha família, em especial à minha mãe, que sempre priorizou o estudo dos seus filhos.

Resumo

Veículos autônomos utilizam sensores lidar para gerar nuvens de pontos 3D, que posteriormente são processados utilizando segmentação e detecção de objetos. Como esses sensores geram grande quantidade de dados, a compressão é fundamental. Dito isso, alguns grupos de padronização estão desenvolvendo técnicas avançadas de compressão de nuvens de pontos. Neste trabalho desenvolvemos uma nova métrica de distorção para avaliar o impacto da compressão de nuvens de pontos no desempenho da segmentação semântica, o que pode afetar a segurança da navegação e os requisitos de largura de banda. Testamos dois algoritmos de compressão da MPEG (GPCC e L3C2) e dois dos principais algoritmos de segmentação semântica (2DPASS e PVKD) no conjunto de dados SemanticKITTI. Os resultados mostram que, para manter alta qualidade de segmentação, é necessária uma taxa de transmissão de aproximadamente 1 MB/s para o GPCC e 2,8 MB/s para o L3C2. Esses resultados são importantes para o planejamento de recursos de infraestrutura para a navegação autônoma.

Palavras-chave: nuvem de pontos, compressão, segmentação, métricas

Abstract

Autonomous vehicles rely on lidar sensors to generate 3D point clouds for accurate segmentation and object detection. As they generate a large amount of data, compression is essential. In effect, standardization groups are trying to develop advanced point cloud compression techniques. We developed a new (suitable) distortion metric to evaluate the impact of point cloud compression on semantic segmentation performance which can impact navigation safety and bandwidth requirements. Two of MPEG's compression algorithms (GPCC and L3C2) and two leading semantic segmentation algorithms (2DPASS and PVKD) were tested over the SemanticKITTI dataset. Results indicate that high segmentation quality requires communication throughput of approximately 1 MB/s for GPCC and 2.8 MB/s for L3C2. These results are important to plan infrastructure resources for autonomous navigation.

Keywords: point cloud, compression, segmentation, metrics

Sumário

Lista de Figuras

Capítulo 1

Introdução

Os sistemas autônomos representam uma inovação tecnológica, com a capacidade de operar sem intervenção humana [1]. No contexto dos carros autônomos, essa tecnologia combina abordagens de inteligência artificial com sensores avançados para interpretar o ambiente e tomar decisões em tempo real [2]. Um dos sensores mais importantes para os carros autônomos é o LiDAR, um sensor que emite pulsos de luz laser e mede o tempo que leva para esses pulsos refletirem de volta para o sensor [2]. Com base nesses dados, o LiDAR gera um mapa da superfície do ambiente ao redor do veículo, conhecido como Nuvem de Pontos ou Point Cloud.

A segmentação semântica desempenha um papel vital no processamento desse dado, ao utilizar de técnicas de inteligência artificial para identificar e classificar as informações contidas na nuvem de pontos, atribuindo rótulos semânticos para cada ponto, como veículo, humano e rodovia [3]. A integridade e a qualidade dos dados são cruciais para garantir a precisão dos algoritmos de segmentação, o que pode refletir na segurança do sistema de condução autônoma [3]. Contudo, as nuvens de pontos representam um grande volume de dados, exigindo vastos recursos de memória, processamento e transmissão[4].

A compressão desses dados torna-se, portanto, um passo essencial para otimizar o processamento, o armazenamento e a transmissão dessas informações [5], especialmente em cenários que dependem do processamento em nuvem, onde a eficiência na transferência de grandes volumes de dados é crítica [4].

O objetivo deste trabalho é avaliar o impacto da compressão de nuvens de pontos no desempenho dos algoritmos de segmentação semântica. Um ponto importante para tal avaliação é selecionar adequadamente uma métrica de distorção. Estudos semelhantes foram realizados [6], mas não propõem uma métrica adequada para o problema. Estamos interessados em manter a qualidade do algoritmo de segmentação, minimizando modificações nas classes segmentadas causadas pelos algoritmos de compressão.

As métricas convencionais de distorção de nuvens de pontos não são adequadas por muitos motivos. Após a compressão, pontos podem desaparecer ou mudarem de posição, isso dificulta a comparação com o resultado esperado. Uma visualização simplificada do problema pode ser vista nas Figuras 1.1 e 1.2. Na Figura 1.1 a comparação da segmentação com o resultado esperado é trivial, já na Figura 1.2, se perde a referência com o resultado esperado para comparação por conta das distorções causadas pela compressão.

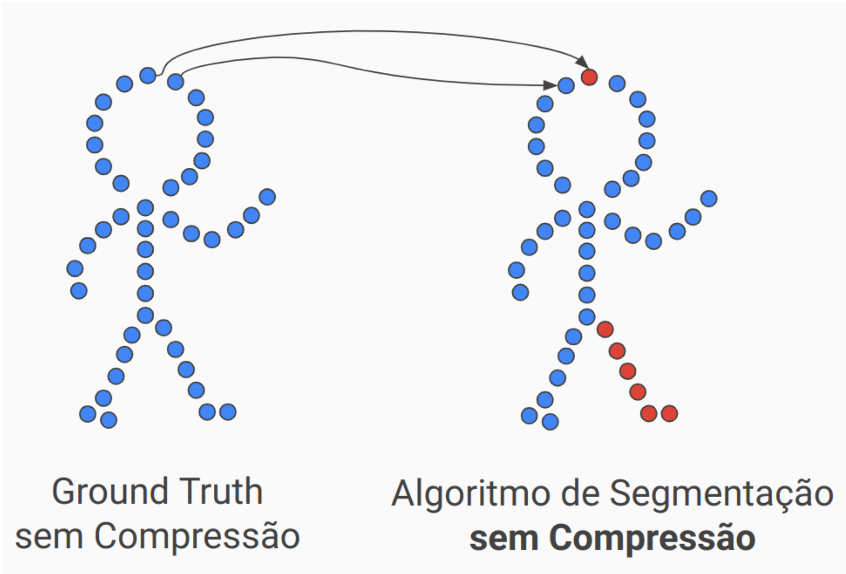

Figura 1.1: Comparação dos rótulos sem compressão.

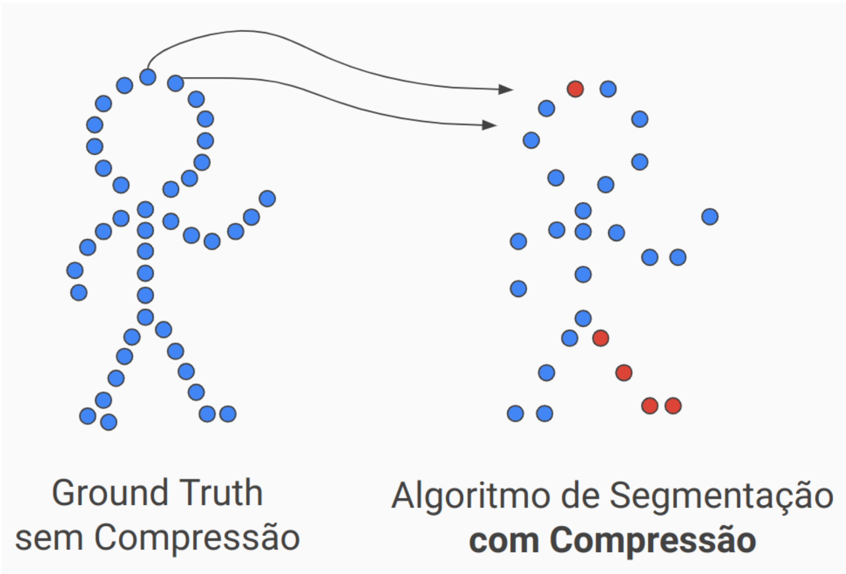

Figura 1.2: Comparação dos rótulos com compressão.

Também gostaríamos de considerar se humanos estão envolvidos na mudança de classe, dada a importância crítica de garantir a segurança das pessoas nesses ambientes [7]. Neste trabalho uma nova métrica é proposta, solucionando os problemas levantados e enfrentados por métricas convencionais, e uma análise é realizada no impacto da compressão nos algoritmos de segmentação semântica utilizando dessa nova métrica.

Usamos o conjunto de dados SemanticKITTI [8], uma base de dados amplamente reconhecida, dois algoritmos de compressão padrão (GPCC [9] e L3C2 [10]) e dois algoritmos de segmentação semântica (2DPASS [11] e PVKD [12]) para avaliar o nível máximo de compressão que pode ser aplicado antes de degradar substancialmente a qualidade da segmentação. A nova métrica é utilizada para avaliar quantitativamente os resultados, e alguns exemplos são analisados qualitativamente por meio de imagens das cenas.

Capítulo 2

Artigo

Este trabalho foi elaborado no formato de artigo científico, e será submetido para a IEEE Signal Processing Letters[1]. Uma versão preliminar do artigo é apresentado nas próximas onze páginas deste documento.

[1] `https://signalprocessingsociety.org/publications-resources/ieee-signal-processing-letters`

On the Impact of Lidar Point Cloud Compression on Semantic Segmentation Algorithms

Tiago de S. Fernandes and **Ricardo L. de Queiroz**

Abstract

Autonomous vehicles rely on lidar sensors to generate 3D point clouds for accurate segmentation and object detection. As they generate a large amount of data, compression is essential. In effect, standardization groups are trying to develop advanced point cloud compression techniques. We developed a new (suitable) distortion metric to evaluate the impact of point cloud compression on semantic segmentation performance which can impact navigation safety and bandwidth requirements. Two of MPEG's compression algorithms (GPCC and L3C2) and two leading semantic segmentation algorithms (2DPASS and PVKD) were tested over the SemanticKITTI dataset. Results indicate that high segmentation quality requires communication throughput of approximately 1 MB/s for GPCC and 2.8 MB/s for L3C2. These results are important to plan infrastructure resources for autonomous navigation.

Index Terms

lidar point cloud, compression, semantic segmentation, metrics

I. Introduction

Autonomous vehicles rely on light detection and ranging (LiDAR) sensors to perceive their surroundings[1]. These sensors generate detailed 3D Point Clouds (PC) which are subject to semantic segmentation and object detection. LiDAR sensors produce vast amount of data, complicating data processing, storage, or transmission without compression[2]. Standartization bodies such as the Moving Picture Experts Group (MPEG[1]) have looked into developing standard state-of-the-art PC compressors [3].

An important processing step for autonomous systems is semantic segmentation, a key computer vision technique for interpreting point clouds. This technique assigns a semantic class to each point (e.g., car, person, road), enabling systems to understand their environment and make critical real-time decisions. The integrity and quality of the data are crucial to ensure the accuracy of the semantic segmentation algorithms[4] which can reflect in the safety of the autonomous driving system. Our objective is to evaluate the impact of point cloud compression on the performance of semantic segmentation algorithms.

[1] www.mpeg.org

A key issue for such an evaluation it to properly select a distortion metric. Ideally one would like to reduce or eliminate fatalities, but we are not aware of studies relating segmentation quality to autonomous driving patterns yet. Hence, we are interested in maintaining the quality of the segmentation algorithm, i.e. we want to minimize modifications into the segmented classes caused by the compression algorithms. Conventional point cloud distortion metrics are not suitable for many reasons, for example because points can disappear or change class label, so that objects may disappear, appear, or change its aspect. We also want to take into account if humans are involved in the class change.

We use the SemanticKITTI dataset[5], a widely recognized benchmark, two standard compression algorithms (G-PCC[6] and L3C2[7]) and two state-of-the-art semantic segmentation algorithms (2DPASS[8] and PVKD[9]) to evaluate the maximum compression one can apply before substantially degrading segmentation quality.

II. Proposed Semantic Segmentation Metric

The D1 and D2 metrics [10] are commonly used to evaluate point cloud compression and reconstruction accuracy. The D1 metric (point-to-point error) measures the average distance between corresponding points in the original and reconstructed point clouds. The D2 metric (point-to-plane error) measures the average distance from each point in the reconstructed cloud to the plane defined by its corresponding point and neighbors in the original cloud, offering a more robust evaluation of surface smoothness and alignment.

The Dice-Sørensen Coefficient[11] (DSC) and Intersection over Union[12] (IoU) are metrics used to evaluate segmentation accuracy. The IoU assesses the overlap between predicted and ground truth segments, ranging from 0 to 1:

$$\text{IoU} = \frac{|A \cap B|}{|A \cup B|}. \tag{1}$$

The DSC metric also measures the similarity between predicted and ground truth segments, with 1 indicating perfect overlap:

$$\text{DSC} = \frac{2|A \cap B|}{|A| + |B|}. \tag{2}$$

Evaluating the quality of the segmentation by comparing it to the ground truth is non trivial. This is so because the compressed point cloud can be distorted and have less points than the original point cloud, so the reference of what is the right label for a given point is lost.

To find an effective metric for this problem, an initial approach was tested, which involve dividing the point cloud into its respective labels (e.g., car, road, person). For each set of points with the same label, the D1 metric was calculated in between the compressed and the original point clouds. This aim to measure the geometric difference between the sets of points with the same label. However, the distortion caused by the compression increases the D1 metric, even if the segmentation is perfect, thus mixing segmentation quality with the effects of distortion. Also, there are cases where labels disappear from the point cloud. The D1 metric is undefined for an empty set of points. Different workarounds were tested for this issue, with limited effectiveness.

Our approach was inspired by the Dice's coefficient metric (DSC). The idea is to adapt the definition of intersection. For each point in the uncompressed point cloud of a specific label, the closest Euclidean point in the compressed point cloud is found, and their labels are compared. The same idea can be applied by swapping the point clouds, and summing their results yields a value analogous to the doubled intersection term of the DSC metric. The union is simply the sum of the sizes of both point clouds.

To extend this idea, a weight is introduced to represent the importance of preserving the human label after compression. If the original point is labeled as a human, and the label changes after compression, it's assigned a weight higher than 1. An arbitrary weight value $\alpha = 2$ was used in this paper. A new term is also added in the denominator to normalize the function.

Given the uncompressed point cloud X, the compressed one Y, and a label L, let X_i^L be a point in X with label L and Y_i^L a point in Y with label L:

$$\lambda_1(X_i^L, Y, L) = \begin{cases} 0, & \text{if point } X_i^L \text{ and it's closest point} \\ & \text{in } Y \text{ have the same label } L; \\ \alpha, & \text{if their labels are different} \\ & \text{and L is a \textit{human} label;} \\ 1, & \text{otherwise.} \end{cases} \tag{3}$$

$$\lambda_2(Y_i^L, X, L) = \begin{cases} 0, & \text{if point } Y_i^L \text{ and it's closest point;} \\ & \text{in } X \text{ have the same label } L; \\ 1, & \text{otherwise.} \end{cases} \tag{4}$$

Then, our metric is:

$$\delta(X, Y, L) = \frac{\sum_i \lambda_1(X_i^L, Y, L) + \sum_i \lambda_2(Y_i^L, X, L)}{|X^L| + H_X(\alpha - 1) + |Y^L|} \tag{5}$$

Where H_X is the number of human-labeled points in X.

This metric addresses the problems with D1 metric. It can be defined for empty set of points and it does not mix geometry distortion with segmentation, since it does not use the distance between points. The final value also fits inside the $[0, 1]$ range. Using $\alpha = 2$ has the same effect as duplicating the human labeled points.

III. Results

The SemanticKITTI dataset is a widely recognized benchmark in the field of 3D semantic segmentation. It consists of point clouds captured using LIDAR sensors in urban environments. Each point in a point cloud scan has three coordinates, a reflectance and a label. The training labels consists of 28 classes, including road, sidewalk, buildings, vegetation, human, etc.

In order to evaluate the impact of compression on semantic segmentation, specific data from the SemanticKITTI dataset was selected for the experiment. Six sequences (11, 12, 13, 14, 15, and 18) were arbitrarily chosen. From

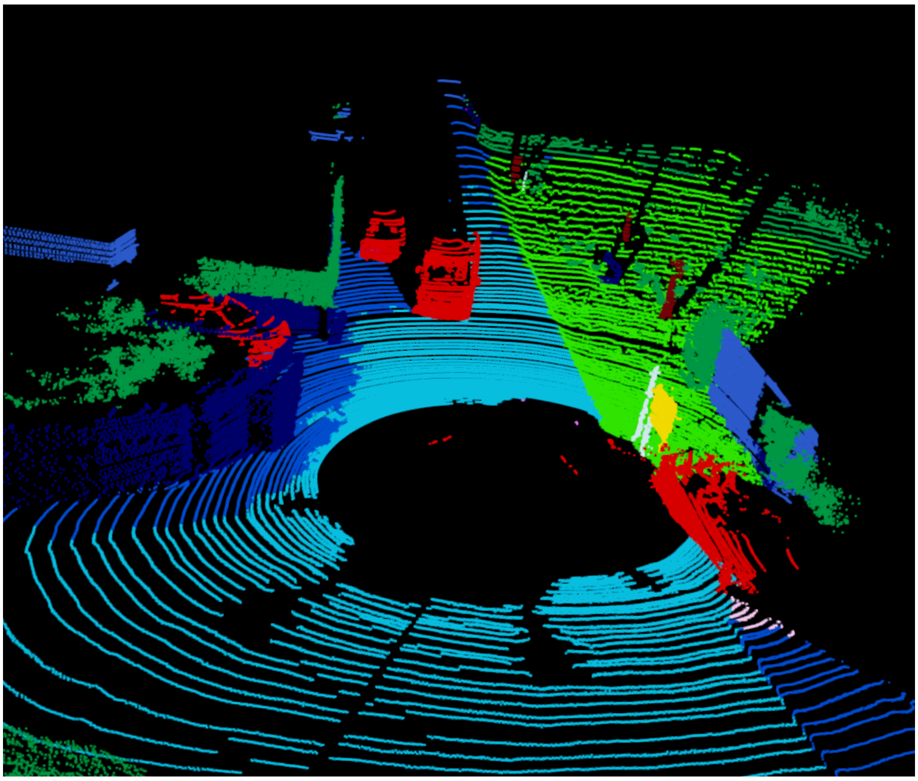

Fig. 1. Example of a SemanticKitti Point Cloud with labels generated by the 2DPASS segmentation algorithm.

each sequence, 101 scans were used, ranging from scan 0 to scan 100, resulting in a total of 606 different point clouds.

The Geometry-based Point Cloud Compression (G-PCC) is a standard developed by the MPEG for compressing 3D point clouds. It uses techniques such as octree decomposition and transforms such as RAHT, to reduce the amount of data required to represent a point cloud. G-PCC is particularly effective in scenarios where the point cloud exhibits sparse points with low geometric complexity and is widely used in applications like autonomous driving and 3D mapping.

The Low-Latency Low-Complexity Codec (L3C2) is an compression algorithm developed to address the specific needs of LiDAR point clouds. L3C2 focuses on providing efficient compression with minimal computational complexity and latency, making it ideal for real-time applications such as autonomous driving and robotic navigation.

MPEG provides six compression settings for both compressors, ranging from *R01*, the most compressed, to *R06*, the least compressed. Additionally, an *R07* rate is defined and used in the experiments, to introduce minimal geometric distortion to the point cloud. The Quantization ScaleI is associated to each compression setting and is used to scale each (x, y, z) coordinate. The geometric distortion caused by the compressors can be understood as multiplying the *(x, y, z)* coordinates of the points by the QS and rounding the result. This process is similar to pruning the Octree representation of the point cloud. The compressor's version this paper uses are from September 27, 2023 (version v23.0-rc2-0-ga3d15c5).

TABLE I

AVERAGE THROUGHPUT AND QUANTIZATION SCALE FOR EACH COMPRESSION RATE.

	Throughput (MB/s)		
Compression	**G-PCC**	**L3C2**	**Quantization Scale**
R01	0.0408	0.1661	0.0019
R02	0.1102	0.4125	0.0039
R03	0.5729	1.7662	0.0160
R04	1.0965	2.7951	0.0310
R05	2.2189	4.2298	0.1300
R06	2.8484	4.8205	0.2500
R07	4.0803	5.9432	0.9000
No compression	28.7465	28.7465	1.0000

The SemanticKITTI website hosts a Semantic Segmentation competition to compare results of the current best algorithms available. As of October 8, 2023[2], the top two algorithms were the The 2D Projection-based Aggregated Segmentation Scheme (2DPASS) and the The Point-Voxel Knowledge Distillation (PVKD).

The 2DPASS is a semantic segmentation algorithm that aims to enhance segmentation accuracy by leveraging 2D images alongside 3D LiDAR data. It achieved a 72.9 IoU metric in the SemanticKITTI competition.

The PVKD is a semantic segmentation algorithm that tackles the challenge of transferring knowledge from a large, complex teacher model to a smaller, more efficient student network. This process improves real-time processing capabilities without significantly compromising performance. It achieved a 71.2 IoU in the SemanticKITTI competition.

The version this paper uses for both segmenters are from October 20, 2023.

For each combination of point cloud, compressor, compression rate, and segmenter, the following process is performed: First, the point cloud is compressed using the specified compressor and compression rate. Next, the compressed point cloud is segmented by the specified segmenter. Finally, a metric is applied by comparing the results to the ground truth: the segmentation obtained without compression.

After compressing and segmenting each point cloud, calculating the average throughput and the proposed metric, five labels were arbitrarily chosen to visualize the results: car, road, sidewalk, terrain, and person. For each of these labels, the average the proposed metric for all point clouds was computed and plotted against the average throughput for the corresponding compression rate, in detail in Table I.

To better visualize the curves, the throughput was plotted on a logarithmic scale. Rates achieved by each setting are marked as vertical traced lines. The results using the L3C2 compressor are displayed in Figures 2 and 3, and those using the G-PCC compressor are shown in Figures 4 and 5.

Both the G-PCC and L3C2 compressors demonstrate similar results, with the proposed metric decreasing as throughput increases, as expected. However, the L3C2 requires slightly higher throughput to achieve comparable

[2]http://www.semantic-kitti.org/tasks.html#semseg

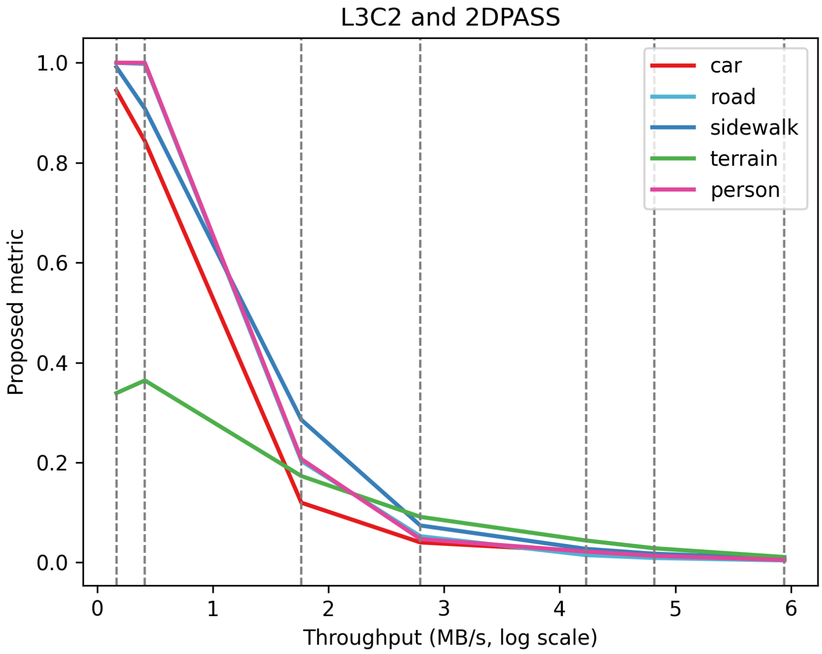

Fig. 2. The proposed metric results for the L3C2 compression and 2DPASS segmentation algorithm.

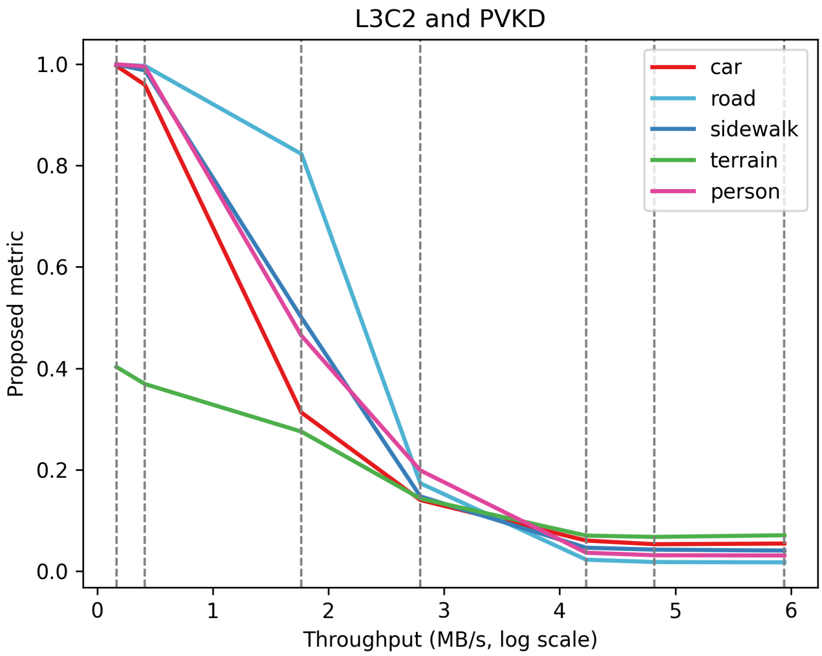

Fig. 3. The proposed metric results for the L3C2 compression and PVKD segmentation algorithm.

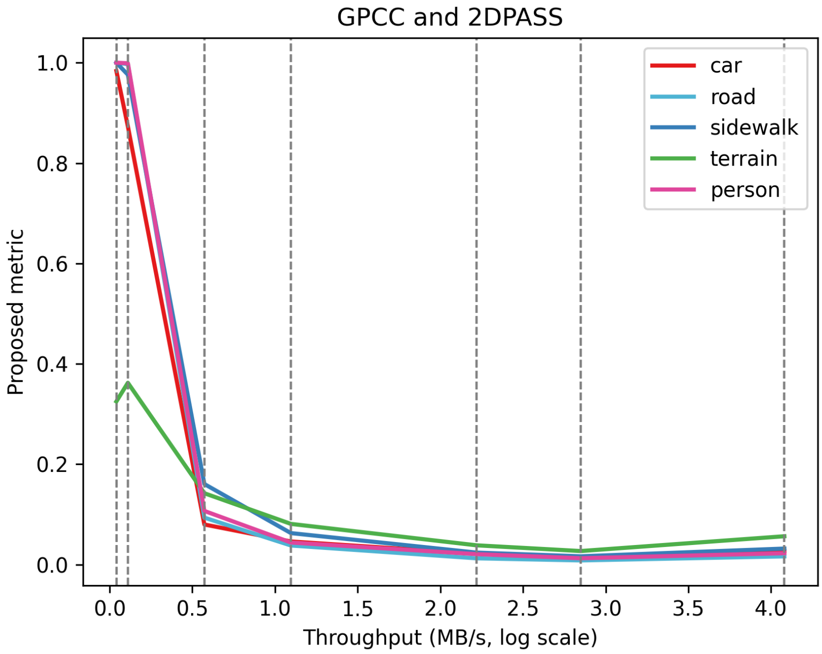

Fig. 4. The proposed metric results for the G-PCC compression and 2DPASS segmentation algorithm.

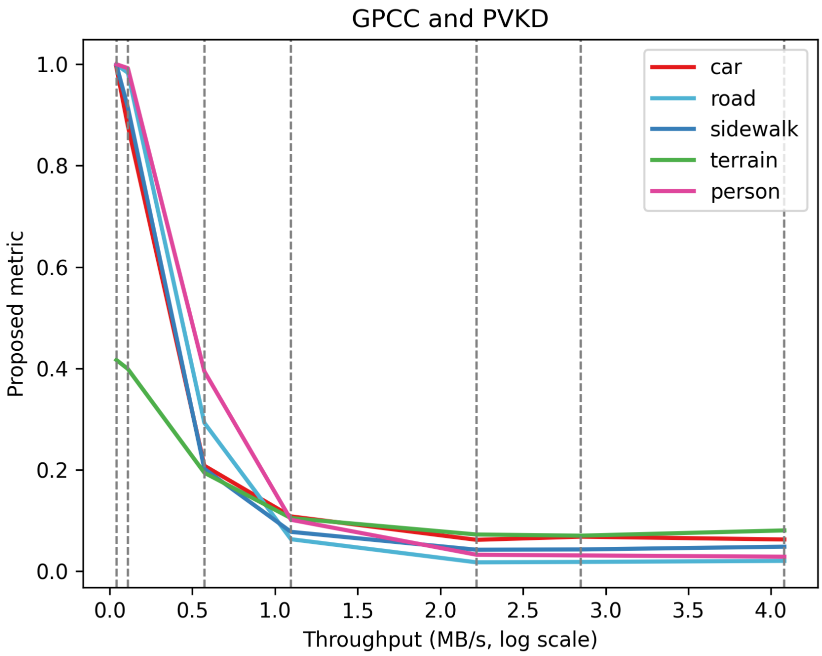

Fig. 5. The proposed metric results for the G-PCC compression and PVKD segmentation algorithm.

The proposed metric values to the G-PCC compressor, particularly noticeable at lower throughput levels. This suggests that the G-PCC compressor may be more efficient at preserving segmentation quality at lower throughputs. This result is expected since the L3C2 is supposed to be a simpler and faster compressor. Both 2DPASS and PVKD segmentation algorithms exhibit similar performance, but the 2DPASS seems to slightly surpass the PVKD with more consistent results across the labels.

Across all graphs, more stable results are observed for higher throughput values, where the proposed metric is close to zero, the lowest value it can reach. However, the metric increases rapidly at lower throughput levels. This increase mainly occurs when transitioning from R03 to R02, resulting in significantly degraded segmentation results. Examples of this degradation are shown in Figures 6 and 7, each illustrates three segmentations using G-PCC and 2DPASS: one without any compression, and the other two using R03 and R02.

Although the R03 compression does not reach the stable region in the graphs, it still produces reasonably good results compared to the uncompressed data. However, the quantization caused by the R02 compression is sufficient to erase many important features that the segmentation algorithms use to properly label the points. Consequently, most points are incorrectly assigned the terrain label, where there should be the road and people walking. The threshold between a compression rate that provides good segmentation and one that doesn't appears to lie between the R03 and R02 compression rates.

Therefore, a reliable throughput threshold value can be established at the R04 setting. For the G-PCC compressor, this threshold is approximately 1 MB/s, while for the L3C2 compressor, it is around 2.8 MB/s. Beyond these throughput values, segmentation quality stabilizes and remains high for most labels. This suggests that having a throughput at or above these levels ensures that those segmentation algorithms can perform effectively, preserving the integrity and accuracy of the segmented point clouds.

IV. Conclusions

We explored the impact of point cloud compression on semantic segmentation algorithms, focusing on the G-PCC and L3C2 compression methods and the 2DPASS and PVKD segmentation algorithms. Using a new and suitable metric, we evaluated how different compression rates affect segmentation quality within the SemanticKITTI dataset. The results indicate that maintaining high segmentation quality requires throughput values of approximately 1 MB/s for G-PCC and 2.8 MB/s for L3C2, which are important for ensuring efficiency in autonomous systems, specially in a scenario where the data is remotely processed.

The significant drop in segmentation quality observed when transitioning from the R03 to the R02 compression settings suggests that point density might be a key factor. The reduced point density at R02 may erase important features necessary for accurate labeling, leading to degraded results. Future work should investigate methods to retain important features while achieving effective compression.

Future research could extend the study of our new metric to other steps in the autonomous vehicle pipeline, such as object detection and decision making. Understanding how this metric affects these subsequent steps will provide a more comprehensive view of it's effectiveness. These findings offer valuable insights for researchers and developers in the field of lidar-based perception systems.

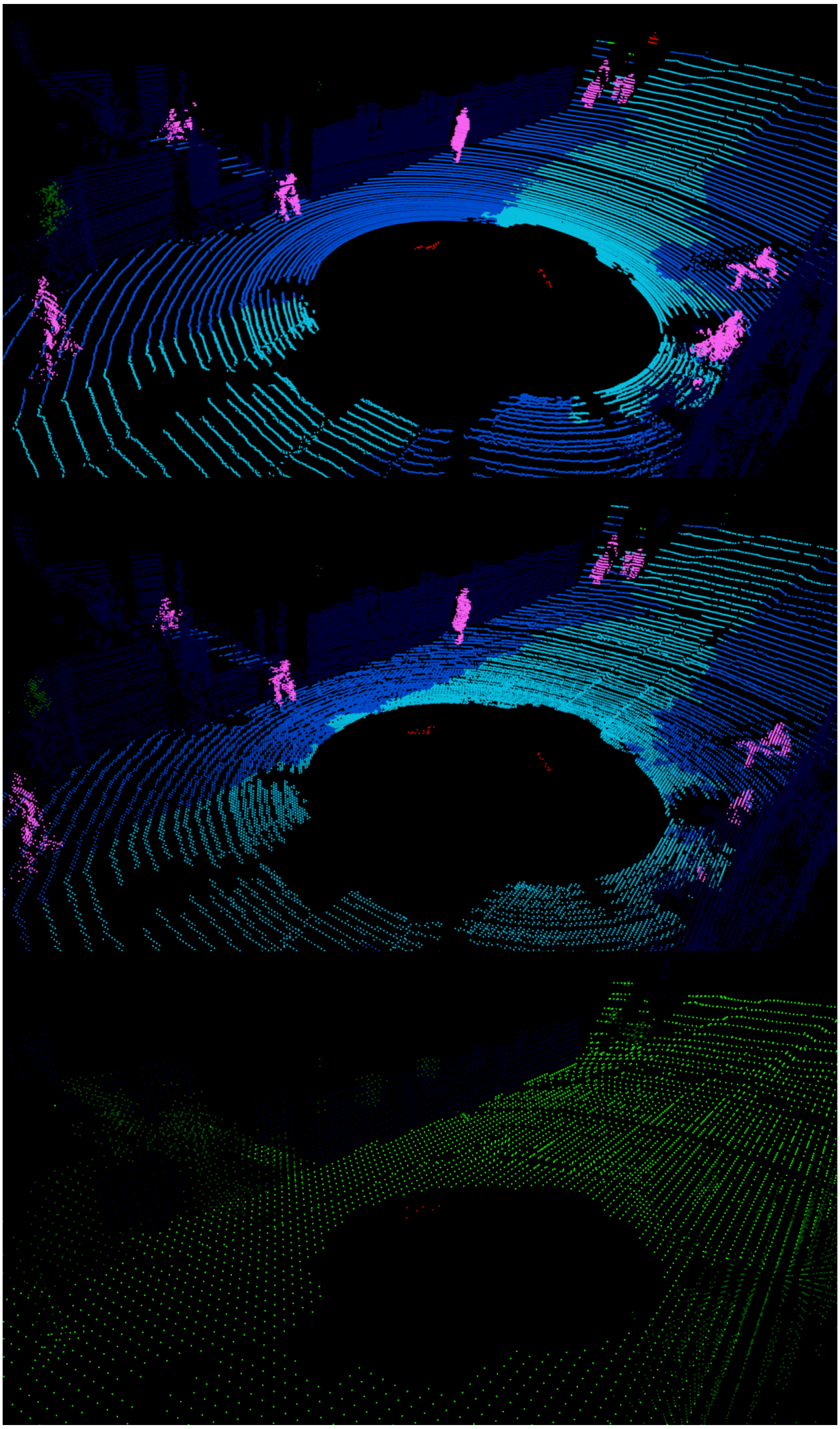

Fig. 6. G-PCC and 2DPASS on Sequence 18: without compression, R03 and R02, from top to bottom.

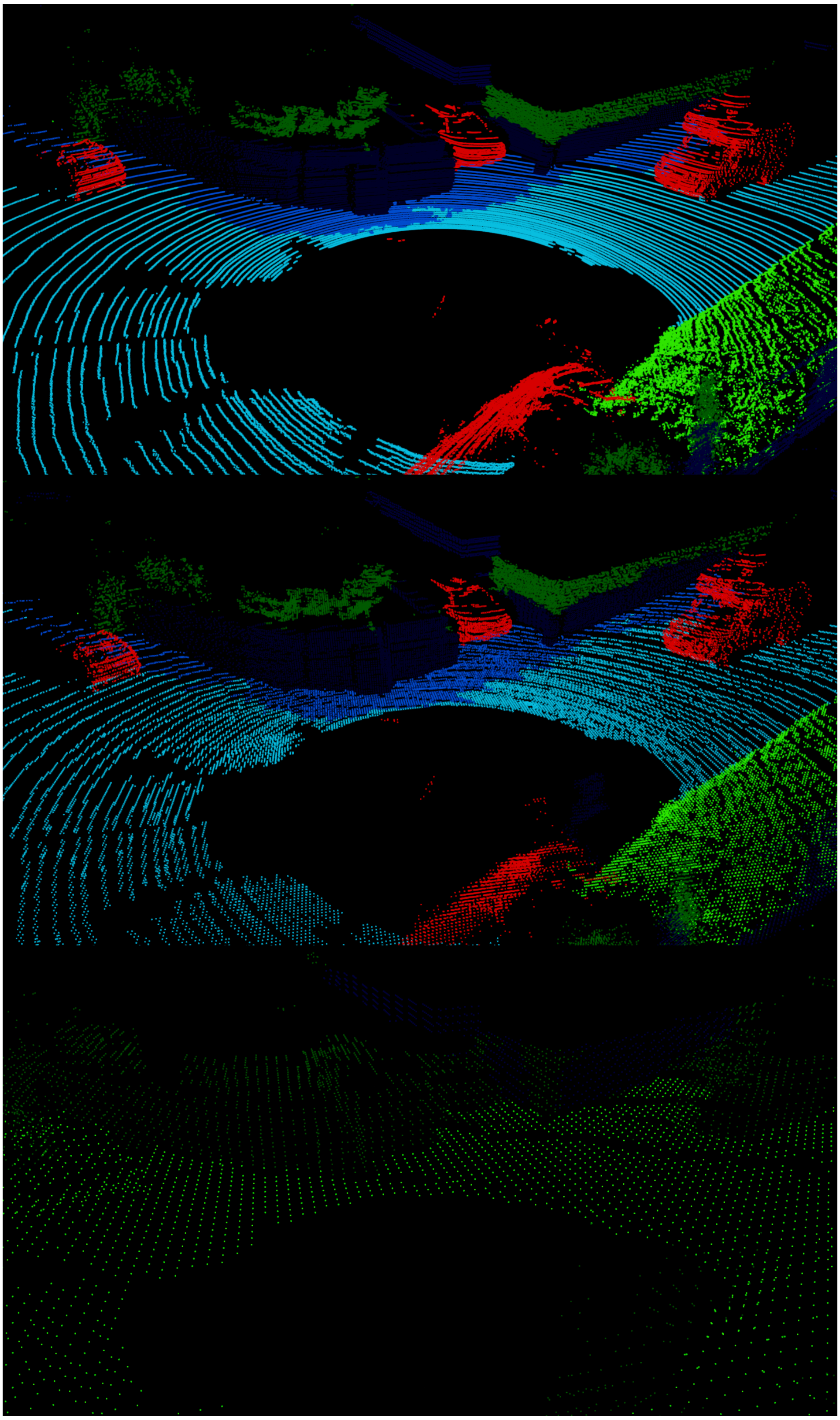

Fig. 7. G-PCC and 2DPASS on Sequence 11: without compression, R03 and R02, from top to bottom.

REFERENCES

[1] J. Zhang and S. Singh, "Loam : Lidar odometry and mapping in real-time," *Robotics: Science and Systems Conference (RSS)*, pp. 109–111, 01 2014.

[2] C. Tu, E. Takeuchi, A. Carballo, and K. Takeda, "Real-time streaming point cloud compression for 3d lidar sensor using u-net," *IEEE Access*, vol. 7, pp. 113 616–113 625, 2019.

[3] S. Schwarz, M. Preda, V. Baroncini, M. Budagavi, P. Cesar, P. A. Chou, R. A. Cohen, M. Krivokuća, S. Lasserre, Z. Li, J. Llach, K. Mammou, R. Mekuria, O. Nakagami, E. Siahaan, A. Tabatabai, A. M. Tourapis, and V. Zakharchenko, "Emerging mpeg standards for point cloud compression," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 9, no. 1, pp. 133–148, March 2019.

[4] E. Grilli, F. Menna, and F. Remondino, "A review of point clouds segmentation and classification algorithms," *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII-2/W3, pp. 339–344, 2017. [Online]. Available: https://isprs-archives.copernicus.org/articles/XLII-2-W3/339/2017/

[5] J. Behley, M. Garbade, A. Milioto, J. Quenzel, S. Behnke, C. Stachniss, and J. Gall, "SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences," in *Proc. of the IEEE/CVF International Conf. on Computer Vision (ICCV)*, 2019.

[6] D. Graziosi, O. Nakagami, S. Kuma, A. Zaghetto, T. Suzuki, and A. Tabatabai, "An overview of ongoing point cloud compression standardization activities: video-based (v-pcc) and geometry-based (g-pcc)," *APSIPA Transactions on Signal and Information Processing*, vol. 9, p. e13, 2020.

[7] J. T. S. LASSERRE, "A point cloud codec for lidar data with very low complexity and latency." 2021, m56477. ISO/IEC JTC 1/SC 29/WG 7.

[8] X. Yan, J. Gao, C. Zheng, C. Zheng, R. Zhang, S. Cui, and Z. Li, "2dpass: 2d priors assisted semantic segmentation on lidar point clouds," in *European Conference on Computer Vision*. Springer, 2022, pp. 677–695.

[9] Y. Hou, X. Zhu, Y. Ma, C. C. Loy, and Y. Li, "Point-to-voxel knowledge distillation for lidar semantic segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2022, pp. 8479–8488.

[10] D. Tian, H. Ochimizu, C. Feng, R. Cohen, and A. Vetro, "Geometric distortion metrics for point cloud compression," in *2017 IEEE International Conference on Image Processing (ICIP)*, Sep. 2017, pp. 3460–3464.

[11] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945. [Online]. Available: http://www.jstor.org/stable/1932409

[12] A. H. Murphy, "The finley affair: A signal event in the history of forecast verification," *Weather and Forecasting*, vol. 11, no. 1, pp. 3 – 20, 1996. [Online]. Available: https://journals.ametsoc.org/view/journals/wefo/11/1/1520-0434_1996_011_0003_tfaase_2_0_co_2.xml

Capítulo 3

Conclusões

Este trabalho explora o impacto da compressão de nuvens de pontos em algoritmos de segmentação semântica, propondo uma nova métrica de distorção para mensurar a qualidade dos resultados. Atualmente, existem poucos trabalhos na literatura que atacam este problema, e a definição de uma nova métrica é um diferencial.

Para testar a eficácia da abordagem, foram utilizados dois algoritmos de compressão do MPEG, o GPCC e o L3C2, e dois algoritmos estado da arte em segmentação semântica, o 2DPASS e o PVKD. A nova métrica busca mensurar o erro na segmentação de uma classe comparando-a com o ponto mais próximo na segmentação de referência. O trabalho utilizou o SemanticKITTI para realizar os testes e análises, uma base de dados de nuvens de pontos de lidar em rodovias urbanas, amplamente reconhecida e utilizada.

Os resultados indicaram que manter uma qualidade alta na segmentação das nuvens de pontos requer certos limiares de taxa de transferência do sistema. Os valores encontrados foram de aproximadamente 1 MB/s para o GPCC e 2.8 MB/s para o L3C2. Esses valores são importantes para garantir eficiência de sistemas autônomos, especialmente em cenários onde os dados são processados em nuvem.

A queda significativa na qualidade da segmentação observada em uma curta diferença de compressão, do R03 para R02, sugere que a densidade dos pontos pode ser um fator crucial. A densidade reduzida de pontos em R02 pode eliminar características importantes necessárias para a rotulagem precisa, resultando em uma degradação dos resultados. Trabalhos futuros devem investigar métodos para reter características importantes enquanto se obtém uma compressão eficaz.

Pesquisas futuras poderiam estender o estudo de nossa nova métrica para outras etapas no pipeline de veículos autônomos, como a detecção de objetos e a tomada de decisões. Entender como essa métrica afeta essas etapas subsequentes fornecerá uma visão mais abrangente de sua eficácia. Essas descobertas oferecem conhecimento valioso para pesquisadores e desenvolvedores no campo de sistemas de percepção baseados em lidar.

Referências

[1] King, David, David Jacques, Jeremy Gray e Katherine Cheney: *Design and simulation of a wide area search mission: An implementation of an autonomous systems reference architecture*. Em *2020 Winter Simulation Conference (WSC)*, páginas 540–551, 2020. 1

[2] Zhang, Ji e Sanjiv Singh: *Loam : Lidar odometry and mapping in real-time*. Robotics: Science and Systems Conference (RSS), páginas 109–111, janeiro 2014. 1

[3] Grilli, E., F. Menna e F. Remondino: *A review of point clouds segmentation and classification algorithms*. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W3:339–344, 2017. `https://isprs-archives.copernicus.org/articles/XLII-2-W3/339/2017/`. 1

[4] Tu, Chenxi, Eijiro Takeuchi, Alexander Carballo e Kazuya Takeda: *Real-time streaming point cloud compression for 3d lidar sensor using u-net*. IEEE Access, 7:113616–113625, 2019, ISSN 2169-3536. 1

[5] Schwarz, Sebastian, Marius Preda, Vittorio Baroncini, Madhukar Budagavi, Pablo Cesar, Philip A. Chou, Robert A. Cohen, Maja Krivokuća, Sébastien Lasserre, Zhu Li, Joan Llach, Khaled Mammou, Rufael Mekuria, Ohji Nakagami, Ernestasia Siahaan, Ali Tabatabai, Alexis M. Tourapis e Vladyslav Zakharchenko: *Emerging mpeg standards for point cloud compression*. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 9(1):133–148, March 2019, ISSN 2156-3365. 1

[6] Martins, Nuno A. B., Luís A. da Silva Cruz e Fernando Lopes: *Impact of lidar point cloud compression on 3d object detection evaluated on the kitti dataset*. J. Image Video Process., 2024(1), jun 2024, ISSN 1687-5176. `https://doi.org/10.1186/s13640-024-00633-4`. 1

[7] Cunningham, Mitchell e Michael Regan: *Autonomous vehicles: Human factors issues and future research*. setembro 2015. 2

[8] Behley, J., M. Garbade, A. Milioto, J. Quenzel, S. Behnke, C. Stachniss e J. Gall: *SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences*. Em *Proc. of the IEEE/CVF International Conf. on Computer Vision (ICCV)*, 2019. 3

[9] Graziosi, D., O. Nakagami, S. Kuma, A. Zaghetto, T. Suzuki e A. Tabatabai: *An overview of ongoing point cloud compression standardization activities: video-based*

(v-pcc) and geometry-based (g-pcc). APSIPA Transactions on Signal and Information Processing, 9:e13, 2020. 3

[10] S. LASSERRE, J. TAQUET: *A point cloud codec for lidar data with very low complexity and latency.*, 2021. m56477. ISO/IEC JTC 1/SC 29/WG 7. 3

[11] Yan, Xu, Jiantao Gao, Chaoda Zheng, Chao Zheng, Ruimao Zhang, Shuguang Cui e Zhen Li: *2dpass: 2d priors assisted semantic segmentation on lidar point clouds*. Em *European Conference on Computer Vision*, páginas 677–695. Springer, 2022. 3

[12] Hou, Yuenan, Xinge Zhu, Yuexin Ma, Chen Change Loy e Yikang Li: *Point-to-voxel knowledge distillation for lidar semantic segmentation*. Em *IEEE Conference on Computer Vision and Pattern Recognition*, páginas 8479–8488, 2022. 3