A Synthetic Data Generation Pipeline for Point-Cloud-Based Rebar Segmentation

Tao Sun[1], Yingtong Luo[2] and Yi Shao[1]

[1]Department of Civil Engineering, McGill University, Canada
[2]Department of Mechanical Engineering, McGill University, Canada
tao.sun@mail.mcgill.ca, yingtong.luo@mail.mcgill.ca, yi.shao2@mcgill.ca

Abstract

Automated rebar cage assembly and quality inspection require reliable rebar recognition. Although rebar segmentation from point clouds has been extensively studied, its generalizability remains limited. One key challenge is the scarcity of real data for training the segmentation models. To address this issue, we propose, for the first time, a pipeline for generating synthetic data for the rebar point cloud instance segmentation task. Using this pipeline, we applied the state-of-the-art Oneformer3d on rebar mesh instance segmentation. The model trained on our synthetic dataset achieved 92.1 mAP in real-world experiments, showing strong synthetic-to-real transfer capability. By eliminating the need for manual data collection and annotation, the proposed method facilitates advancements in automated rebar cage assembly and dimensional quality inspection technologies.

Keywords –

Point cloud instance segmentation; Rebar Recognition; Synthetic-to-Real; Computer Vision

1 Introduction

The automation of rebar cage assembly and dimensional quality inspection has garnered significant attention in the construction industry due to its potential to enhance productivity and reduce labor costs [1]. Accurate rebar recognition is a key technology that enables the automation of the above tasks. It serves as a critical foundation for many downstream tasks, such as rebar manipulation based on pose estimation and the dimensional measurement of rebar cages.

Existing studies have extensively explored rebar segmentation from images or point clouds to accurately identify the shapes and dimensions of rebars. Depending on the type of sensor data, these methods can be categorized into image-based rebar segmentation methods and point cloud-based segmentation methods. The advantage of the former lies in the ease of data collection and annotation and lower hardware costs. For instance, Kardovskyi et al. [2] applied the Mask R-CNN to segment rebars from images, which was then integrated with stereo vision techniques to measure rebar length and spacing. Chang et al. [3] present a hybrid rebar segmentation method to enable accurate rebar dimension measurement. This method first employs Mask R-CNN to segment rebars from images, followed by mapping the segmented results into 3D space, where point cloud processing techniques are used to cluster and identify 3D rebar instances.

Compared to object segmentation from images, direct segmentation from 3D point clouds offers two significant advantages: (1) it enables end-to-end processing of point clouds from laser scanners, and (2) it leverages prior knowledge of objects' 3D shapes more effectively. Consequently, many studies have focused on rebar segmentation from 3D point clouds [4], [5], [6]. They all relied on point cloud processing techniques, leveraging prior knowledge of rebar shapes to segment point clouds from complex scenarios. For example, Kim et al. [7] proposed a method based on the line RANSAC method and circular RANSAC to efficiently identify straight rebars within rebar cages. However, relying solely on point cloud processing to identify rebars faces limitations in generalizability [3].

Deep learning-based point cloud instance segmentation has demonstrated remarkable performance in various computer vision tasks, such as indoor scene understanding [8] and industrial part dimension inspection [9]. They typically rely heavily on large quantities of annotated data for training [10]. However, such large-scale data in the context of rebar segmentation is still lacking [3], which directly hinders the advancement of rebar segmentation. Moreover, the labor-intensive and time-consuming nature of manual annotation processes further exacerbates this challenge. Wang et al. [11] proposed a synthetic dataset using BIM and Lumion 11 rendering software, which contains automatically labeled 2500 synthetic images of rebar meshes. Mask R-CNN was trained and tested on their proposed dataset, showing high accuracy of bounding

box and mask prediction. However, to the best of the authors' knowledge, no point cloud dataset specifically designed for rebar instance segmentation has been proposed. Therefore, the applicability of advanced point cloud instance segmentation networks to rebar segmentation tasks also remains unexplored.

To address these critical gaps, we first develop a synthetic data generation pipeline for generating synthetic point clouds. The pipeline simulates realistic rebar point clouds from image-based 3D reconstruction and their corresponding 3D labels, enabling the training of advanced segmentation networks without the need for manual data collection and annotation. Using the proposed synthetic dataset, we apply a state-of-the-art point cloud instance segmentation network to the real-world task of rebar mesh recognition. The feasibility and accuracy of the developed pipeline are validated through real-world experiments. Additionally, the experiments show both the applicability and limitations of the state-of-the-art method in rebar segmentation tasks.

2 Method

Point cloud instance segmentation networks typically require a substantial amount of training data to enable the model to effectively learn the visual and geometric features of target objects. Therefore, we developed an end-to-end data generation pipeline for rebar instance segmentation based on the modeling and rendering software Blender, as shown in Figure 1. This pipeline efficiently simulates 3D point clouds reconstructed using Structure-from-Motion (SfM) and Multi-View Stereo (MVS) algorithms and automatically generates labels. The proposed pipeline consists of three stages: (1) scene generation, (2) point cloud reconstruction, and (3) model training.

2.1.1 Scene generation

We first developed a parameterized modeling method using Blender and Infinigen [12]. The method can generate large-scale rebar models with diverse shapes and appearances. Specifically, Blender is used to model rebars, while Infinigen converts the parametric geometry nodes into executable code for batch processing. Figure 2 shows different rebar models generated by our method. We use Cylinder primitives in Blender to represent straight rebar segments. To simulate different real rebars, our straight rebar model supports parameters required for typical design specifications, including rebar diameter, length, bending degree, and rib type. The rib structures are created using meshes, enabling high-fidelity and geometric details.

To simulate different rebar appearances, we randomize visual parameters such as base color, metallic properties, roughness, and texture to randomize the visual appearance of the rebar models. Additionally, we use a combination of Noise Texture and Musgrave Texture nodes to simulate surface defects on rebars, such as stains and corrosion.

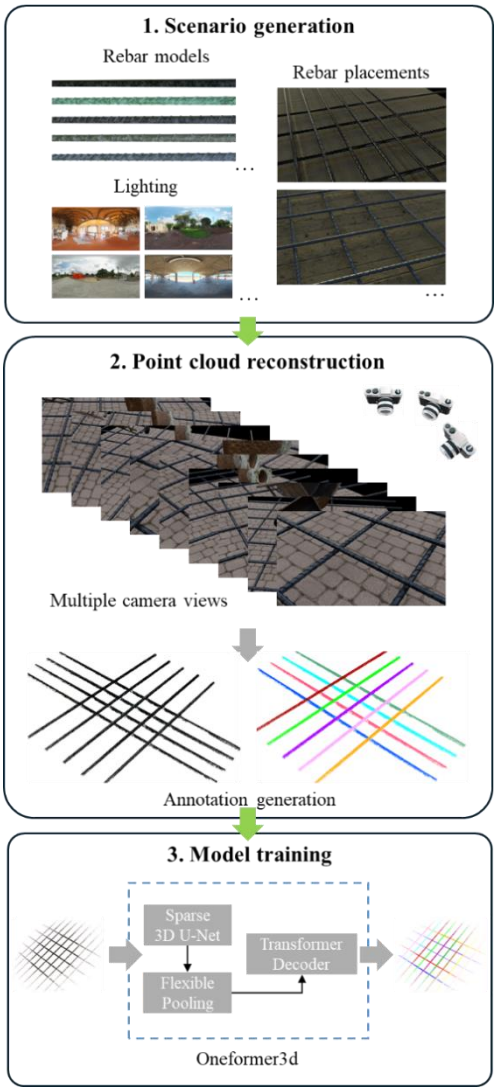

Figure 1. Proposed pipeline for rebar instance segmentation

By defining heuristic rules and graph structures, we can use automatically generated rebar models to establish various rebar cage configurations. In this study, we specify the positions and orientations of straight rebars to construct reinforcement mesh scenarios for slab production. For lighting, the approach of high dynamic range images (HDRI) is applied to illuminate the scenario. To capture a wide range of lighting conditions, 733 HDRI images are collected from the open-source asset library PolyHaven.

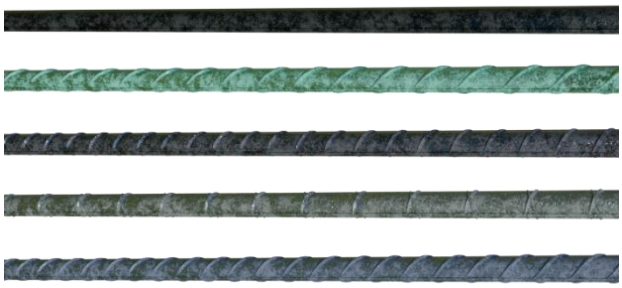

Figure 2. Examples of generated rebar models

2.1.2 Point cloud reconstruction

A straightforward method for generating point clouds from 3D scene models is to downsample the mesh directly to obtain the point cloud [13]. However, this approach struggles to effectively replicate the distribution of point clouds reconstructed from real images, leading to reduced segmentation accuracy [14]. To better simulate the point clouds derived from real image reconstruction, we sample cameras within the 3D scene and employ traditional SfM [15] and MVS [16] algorithms to generate synthetic point clouds. Blender's Cycles renderer is applied to generate photorealistic images of the scenarios. The Cycles renderer utilizes physically based ray tracing to simulate complex interactions between light, shadows, and material reflections.

To accurately label the reconstructed point cloud P_{re}, we first perform uniform downsampling on the rebar mesh model generated in the synthetic scenario, denoting it as P_{GT}. Subsequently, we align P_{re} and P_{GT} by utilizing the camera poses estimated via SfM and the ground-truth poses of the virtual cameras. Based on this alignment, the nearest neighbor for each point in P_{re} is identified in P_{GT}, allowing instance labels to be assigned to P_{re}.

2.1.3 Instance segmentation network

A SOTA transformer-based instance segmentation network, Oneformer3d [17], was used for the instance segmentation model. The method achieved the best performance on three different benchmarks (ScanNet, ScanNet200, and S3DIS), and excels in scenarios involving slender objects like chair legs, which share similar geometric characteristics as with rebars.

3 Experiment

To evaluate the adaptability and performance of the developed pipeline, a large-scale synthetic dataset was first generated. This dataset was designed to be diverse, incorporating randomized rebar models, spacing configurations, lighting conditions, camera numbers and camera poses. Based on this dataset, we trained the state-of-the-art point cloud instance segmentation model, Oneformer3d. To assess the model's syn-to-real generalization capability, real-world experiments were conducted on rebar mesh arrangements, simulating slab reinforcements that are commonly used in construction sites.

3.1 Dataset generation

Based on the proposed pipeline, a synthetic dataset containing 238 samples was reconstructed automatically, as shown in Figure 3(a-b). Each sample was annotated with detailed instance labels. The dataset was divided with an 8:2 ratio into a training set (190 samples) and a validation set (48 samples).

To evaluate syn-to-real transfer performance, we arranged slab rebar meshes to collect real-world point cloud data. A smartphone was used to capture 48 to 79 multi-angle images of each rebar placement, with a resolution of 4032 × 3024 pixels. The camera was positioned approximately 70 cm from the target. To introduce variability, we adjusted the number of rebars (8 to 14) and their placement intervals (5 cm to 25 cm), aiming to challenge the algorithm and assess its robustness under diverse spatial configurations. The final test set comprises 12 manually annotated samples. Figure 3(c-d) provides examples from our real-world test set.

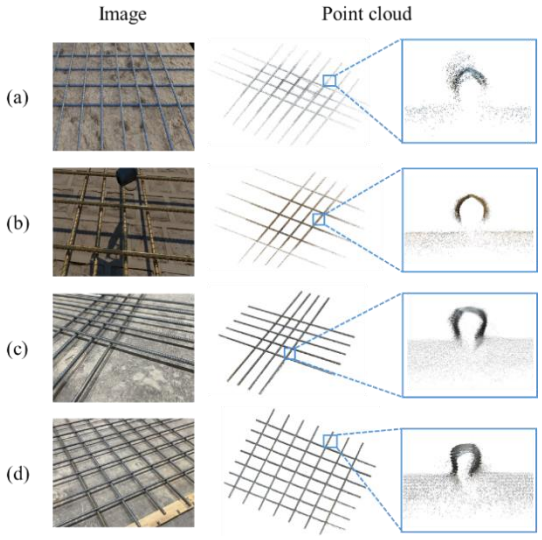

Figure 3. Examples of our synthetic training set and real-world test set: (a-b) synthetic training set; (c-d) real-world test set

3.2 Evaluation Metrics

We evaluated the performance of the 3D rebar instance segmentation using mean average precision (mAP), a standard metric in point cloud segmentation. The mAP was computed as the average precision over Intersection over Union (IoU) thresholds ranging from 50% to 95% in increments of 5%.

3.3 Training setting

All training settings of Oneformer3d are basically following the original implementation [17]. For the learning rate, we used an AdamW optimizer with an initial learning rate of 0.0001 and weight decay of 0.05. Taking computational efficiency into account, we employed a batch size of 2 and constrained the training to 160 epochs. Furthermore, we have experimentally observed that the scale diversity of the point cloud can significantly improve the sensitivity of the segmentation network to the scale of the point cloud. Therefore, we applied a larger range of random scaling, from 0.5 to 1.5, compared to the original implementation.

3.4 Results

Figure 4 illustrates the progression of the mAP on the validation set throughout training. The trend indicates that training on our synthetic dataset is stable and effective, with no signs of significant overfitting. Oneformer3d demonstrates strong segmentation performance on the synthetic dataset, achieving the highest mAP of 99.4. When evaluated on the real-world test set, the model achieved a high mAP of 92.1, indicating its ability to generalize from synthetic training data to real-world scenarios.

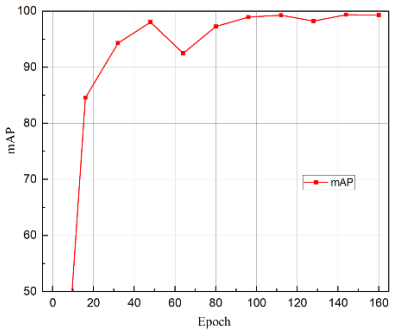

Figure 4. mAP variation with training epochs on the synthetic validation set

Figure 5 presents several results obtained from the real-world dataset. The majority of rebar outputs exhibit high segmentation accuracy with clearly defined instance boundaries However, a few cases exhibit poor segmentation performance. For instance, as illustrated by the red arrows in Figure 5(c), two rebars were incorrectly merged into a single instance. This may be attributed to the global context awareness of transformer-based models, which, in rare cases, can lead to objects with similar features being recognized as the same instance. This merged detection is also noted in indoor object segmentation [18] using a similar transformer-based model with Oneformer3d. Employing a distance-based clustering filter, such as DBSCAN[18], enhanced by shape priors, offers a potential solution to mitigate this challenge.

Additionally, missing points were observed in the segmentation results of certain rebars, as indicated by the blue circles in Figure 5(d). This can be attributed to the model's insufficient accuracy, particularly in handling incomplete edges in point clouds. This minor issue can be addressed using various post-processing techniques, depending on the downstream tasks. For instance, in the case of rebar 6D pose estimation, known geometric priors of the rebar can be utilized to accurately compute the pose even in the presence of missing points.

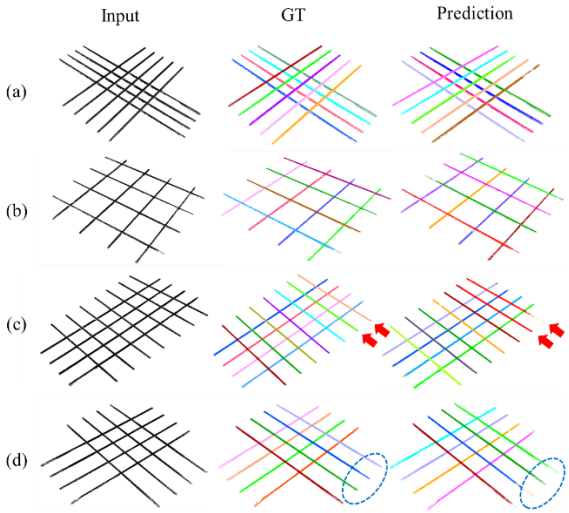

Figure 5. Examples of segmentation results on our real-world test set

4 Discussion

This study demonstrates the adaptability of synthetic data for training the rebar point cloud instance segmentation model. This validates that the approach can generate reliable rebar visual and geometric features, enabling deep neural networks to effectively learn and recognize real-world rebars. Therefore, it is expected that the proposed pipeline can automatically generate 3D data for a wide range of rebar cages. For example, as shown in Figure 6(a), our pipeline can generate complex rebar cage point clouds and annotations. Besides, the pipeline

shows promise for training networks to estimate 6D poses and 2D masks of rebars from 2D images. Furthermore, synthetic models can be integrated into physical simulations to train reinforcement learning policies for dexterous rebar manipulation in future work, as illustrated in Figure 6(d).

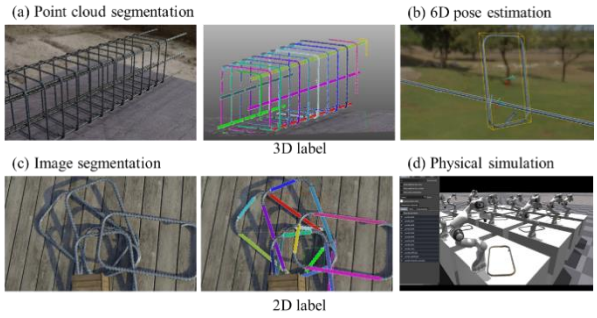

Figure 6. Our pipeline is promising to be extended in the future to: (a) 3D instance segmentation for rebar cage; (b) 6D pose estimation of rebars; (c) 2D image segmentation; (d) Physical simulation

This preliminary exploratory study has certain limitations. First, our point cloud segmentation experiments did not explicitly address the challenges posed by real-world variations, such as more complex rebar structures, occlusions, lighting conditions, and different scanner resolutions. While our synthetic dataset performs well in controlled settings, its generalization to diverse real-world environments remains an open question. Future work should focus on incorporating a broader range of 3D data that captures these variations, ensuring robustness across diverse rebar structures and scanning conditions. Additionally, while preprocessing techniques such as RANSAC [19], can effectively remove the ground, achieving end-to-end rebar segmentation in the presence of complex background interference using neural networks remains a valuable research direction.

Second, the proposed method faces challenges in achieving real-time running in practice due to the high computational cost of the point cloud segmentation network and the offline nature of the reconstruction method. In the future, adopting 3D cameras directly, combined with real-time reconstruction and lightweight segmentation or keypoint detection from 3D point clouds, will be a promising research direction. Third, future work could focus on comprehensive benchmarking to identify the optimal network architecture for rebar point cloud segmentation and facilitate model selection in real-world applications.

Finally, while this study focuses on demonstrating the effectiveness of synthetic data for rebar segmentation, a detailed ablation study on the impact of individual augmentation parameters, such as texture variations, occlusions, and lighting conditions, remains an important area for future work.

5 Conclusion

To address the challenge of lacking training data for rebar point cloud instance segmentation, we first developed a synthetic data generation pipeline. This pipeline enables high-fidelity annotated point clouds, facilitating the training of instance segmentation networks without the need for manual data collection and annotation. The adaptability of the pipeline was validated through a real-world experiment. The state-of-the-art point cloud instance segmentation network Oneformer3d trained on our synthetic dataset achieved a 92.1 mAP on the real-world test set, showing good syn-to-real capacity.

This work contributes to the advancement of automated rebar cage assembly and dimensional quality inspection technologies, paving the way for the broader adoption of intelligent construction solutions.

References

[1] A. Klarin and Q. Xiao, “Automation in architecture, engineering and construction: a scientometric analysis and implications for management,” Engineering, Construction and Architectural Management, vol. 31, no. 8, pp. 3308–3334, 2024.

[2] Y. Kardovskyi and S. Moon, “Artificial intelligence quality inspection of steel bars installation by integrating mask R-CNN and stereo vision,” Automation in Construction, vol. 130, p. 103850, Oct. 2021, doi: https://doi.org/10.1016/j.autcon.2021.103850.

[3] C.-C. Chang, T.-W. Huang, Y.-H. Chen, J. J. Lin, and C.-S. Chen, “Autonomous dimensional inspection and issue tracking of rebar using semantically enriched 3D models,” Automation in Construction, vol. 160, p. 105303, Apr. 2024, doi: 10.1016/j.autcon.2024.105303.

[4] H.-L. Chi, M.-K. Kim, K.-Z. Liu, J. P. P. Thedja, J. Seo, and D.-E. Lee, “Rebar inspection integrating augmented reality and laser scanning,” Automation in Construction, vol. 136, p. 104183, Apr. 2022, doi: 10.1016/j.autcon.2022.104183.

[5] X. Yuan et al., “Automatic evaluation of rebar spacing and quality using LiDAR data: Field application for bridge structural assessment,” Automation in Construction, vol. 146, p. 104708, Feb. 2023, doi: 10.1016/j.autcon.2022.104708.

[6] G. Hodge and J. M. Gattas, “Geometric and semantic point cloud data for quality control of bridge girder reinforcement cages,” Automation in Construction, vol. 140, p. 104334, Aug. 2022, doi:

10.1016/j.autcon.2022.104334.

[7] M.-K. Kim, J. P. P. Thedja, and Q. Wang, "Automated dimensional quality assessment for formwork and rebar of reinforced concrete components using 3D point cloud data," Automation in Construction, vol. 112, p. 103077, Apr. 2020, doi: 10.1016/j.autcon.2020.103077.

[8] Y. Sun, X. Zhang, and Y. Miao, "A review of point cloud segmentation for understanding 3D indoor scenes," Visual Intelligence, vol. 2, no. 1, p. 14, 2024.

[9] Y. Li, Y. Wang, and Y. Liu, "Three-dimensional point cloud segmentation based on context feature for sheet metal part boundary recognition," IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1–10, 2023.

[10] H. Jiang, F. Yan, J. Cai, J. Zheng, and J. Xiao, "End-to-end 3D point cloud instance segmentation without detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020, pp. 12796–12805.

[11] H. Wang, Z. Ye, D. Wang, H. Jiang, and P. Liu, "Synthetic Datasets for Rebar Instance Segmentation Using Mask R-CNN," Buildings, vol. 13, no. 3, p. 585, 2023, doi: https://doi.org/10.3390/buildings13030585.

[12] A. Raistrick et al., "Infinite Photorealistic Worlds using Procedural Generation," Jun. 26, 2023, arXiv: arXiv:2306.09310. Accessed: Jan. 10, 2024. [Online]. Available: http://arxiv.org/abs/2306.09310

[13] J. Shu, W. Li, C. Zhang, Y. Gao, Y. Xiang, and L. Ma, "Point cloud-based dimensional quality assessment of precast concrete components using deep learning," Journal of Building Engineering, vol. 70, p. 106391, Jul. 2023, doi: 10.1016/j.jobe.2023.106391.

[14] M. Chen et al., "Stpls3d: A large-scale synthetic and real aerial photogrammetry 3d point cloud dataset," arXiv preprint arXiv:2203.09065, 2022.

[15] P. Moulon, P. Monasse, R. Perrot, and R. Marlet, "OpenMVG: Open multiple view geometry," in International Workshop on Reproducible Research in Pattern Recognition, Springer, 2016, pp. 60–74.

[16] D. Cernea, "OpenMVS: Multi-View Stereo Reconstruction Library," 2020. [Online]. Available: https://cdcseacave.github.io/openMVS

[17] M. Kolodiazhnyi, A. Vorontsova, A. Konushin, and D. Rukhovich, "Oneformer3d: One transformer for unified point cloud segmentation," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2024, pp. 20943–20953.

[18] J. Schult, F. Engelmann, A. Hermans, O. Litany, S. Tang, and B. Leibe, "Mask3D: Mask Transformer for 3D Semantic Instance Segmentation," in 2023 IEEE International Conference on Robotics and Automation (ICRA), London, United Kingdom: IEEE, May 2023, pp. 8216–8223. doi: 10.1109/ICRA48891.2023.10160590.

[19] L. Li, F. Yang, H. Zhu, D. Li, Y. Li, and L. Tang, "An improved RANSAC for 3D point cloud plane segmentation based on normal distribution transformation cells," Remote Sensing, vol. 9, no. 5, p. 433, 2017.