Semantic Segmentation of Building Components from Drone-Based 3D Point Clouds for Damage Assessment

Shiori K. 1*, Daigo M.2 and Hidenori, Y.3

1Lecture, Institute of Education, Research and Regional Cooperation for Crisis Management Shikoku, Kagawa University, Japan

2Undergraduate Student, Faculty of Engineering And Design, Kagawa University, Japan

3Professor, Faculty of Engineering And Design, Kagawa University, Japan

*kubo.shiori.a6@kagawa-u.ac.jp

Abstract: *In recent large-scale earthquakes in Japan, delays in the issuance of Disaster Victim Certificates have become a major challenge for the prompt provision of public support. Currently, the damage assessment required to issue these certificates relies on on-site visual inspections by municipal staff, a process that is not only time-consuming and labor-intensive but also entails issues of safety and subjectivity. The ultimate goal of this research is to assess house inclination based on its structural components, for which component detection is an essential prerequisite. This paper, as a foundational step toward this goal, proposes and validates a method that utilizes the PointNet++ deep learning model to automatically extract structural elements from high-precision LiDAR point clouds acquired by a drone. The method performs semantic segmentation on the point cloud, classifying points into four classes: Roof, Wall, Ground, and Others, to extract the components necessary for subsequent inclination analysis. The proposed method was trained and validated using a point cloud dataset collected from areas affected by the 2024 Noto Peninsula Earthquake, and it demonstrated high classification performance with a mean Intersection over Union (mIoU) of 77.3% and an Overall Accuracy (OA) of 88.6%. Specifically, the horizontal surfaces of the Roof and Ground classes achieved a high IoU of approximately 85%. On the other hand, the classification accuracy for the Wall class, with an IoU of 59.4%, remains a future challenge, attributed to class imbalance in the dataset and the geometric constraints of data acquisition from the drone's nadir perspective. Nevertheless, the method was shown to be capable of recognizing the structure of not only houses with minor damage but also those that were tilted. This achievement demonstrates its effectiveness as a foundational method for the quantitative evaluation of three-dimensional damage conditions, a task difficult with conventional 2D imagery. We believe this work can serve as a foundational technology for future automated calculation of house inclination, contributing to rapid, objective damage assessment and the issuance of Disaster Victim Certificates.*

Keywords: *Deep learning, Disaster management, Lidar, Semantic Segmentation, UAV*

Introduction

In recent large-scale earthquakes in Japan, typified by the 2024 Noto Peninsula Earthquake, delays in the initial public support for victims' livelihood reconstruction have become a significant issue. One contributing factor is the extensive time required to issue Disaster Victim Certificates, which are official documents issued by local governments to certify the degree of damage to homes caused by natural disasters and serve as a prerequisite for various forms of assistance. The Disaster Victim Certificate is necessary to receive support, such as publicly funded demolition and repair of houses, and eligibility for temporary housing (Tanaka, 2023). The issuance process begins after an application is received from a victim, whereupon municipal staff first conduct on-site damage assessment surveys, visiting each damaged house to evaluate its condition in accordance with national guidelines. Based on these survey results, the damage level is determined, and the certificate is subsequently issued. The substantial personnel and time required for these on-site surveys are the fundamental cause of the delays in issuance. For example, a survey conducted after the 2024 Noto Peninsula Earthquake reported that approximately 85% of municipal staff in the affected areas experienced an increased workload, while about 70% reported a shortage of personnel (Ishikawa Prefecture Headquarters of Jichiro, 2024). Consequently, there are limitations to investigating the vast number of buildings dispersed over a wide area with limited staff, resulting in a bottleneck throughout the entire process.

Furthermore, the current damage assessment survey method itself has several issues. According to the guidelines from the Cabinet Office of Japan, damage assessment is based on the damage ratio of structural components such as roofs, walls, and foundations, as well as the overall inclination of the house. Inclination is a critical evaluation metric; notably, a house is classified as being in a "total collapse" state if its inclination exceeds 1/20. Despite the establishment of these objective criteria, the assessment relies on visual inspection by surveyors, making the results susceptible to their experience and subjectivity, which often leads to inconsistencies. Moreover, the risk of collapse often prevents surveyors from closely approaching damaged houses, forcing them to rely on limited external inspections from a safe distance. Such visual-based external surveys can identify superficial damage, such as cracks in walls, but they make it difficult to quantitatively assess critical metrics, like the overall inclination of the house, often leading to an underestimation of the actual damage (Kyushu Regional Administrative Evaluation Bureau, Ministry of Internal Affairs and Communications, 2018). As a result, victims may file objections to the assessment results, necessitating re-surveys that require additional personnel and time (Orihashi &

Urakawa, 2022). Additionally, conducting surveys near potentially collapsing houses is a hazardous task from a safety perspective, as aftershocks are frequent immediately after an earthquake.

These issues of delays and assessment accuracy have been repeatedly pointed out in past disasters. For instance, during the 2016 Kumamoto Earthquake, significant delays in issuing certificates were reported, caused not only by issues with assessment accuracy but also by a shortage of staff with specialized knowledge and an inadequate system for accepting support staff from other municipalities (Fujii *et al.*, 2012). Similarly, during the 2011 off the Pacific coast of Tohoku Earthquake, the vast and severe damage caused by the tsunami resulted in a severe shortage of survey personnel. Although support staff were accepted from other cities, this led to discrepancies in judgment and highlighted issues with the assessment criteria themselves (Sendai City, 2012). Against this backdrop, the establishment of a rapid, safe, and objective damage assessment method is crucial for the early recovery and reconstruction of disaster-affected areas.

In this study, we propose a method that utilizes a deep learning model to process high-precision 3D point cloud data acquired with a drone-mounted LiDAR system, enabling the automatic extraction of structural components of houses, such as roofs and walls. The ultimate goal is to expedite damage assessment surveys, reduce labor, and enhance objectivity by quantitatively and automatically calculating the inclination of houses from the extracted point cloud data of these structural components. This paper, as a foundational step, focuses on the construction of a model for the automated extraction of these structural parts and verification of its accuracy.

Literature Review

The detection of building damage has primarily been conducted using satellite and aerial imagery. Wang *et al.* (2022) applied deep learning to satellite imagery for the automatic classification of damaged buildings, demonstrating its effectiveness for wide-area damage assessment. Numerous studies have also utilized aerial imagery; for instance, Li *et al.* (2019) detected collapsed houses and those with minor damage from aerial imagery, Ye *et al.* (2014) used multiple classifiers, and Tu *et al.* (2017) leveraged the differences in pre- and post-disaster images to detect building damage. However, these images are subject to limitations in resolution and acquisition conditions, making it difficult to classify the damage level of individual buildings accurately. The recent proliferation of drones has advanced the utilization of high-resolution aerial imagery. For example, Attari *et al.* (2017)

proposed Nazr-CNN, a deep learning pipeline that integrates semantic segmentation and texture-based classification to assess building damage levels from UAV imagery. Similarly, Jozi *et al.* (2024) classified images of damaged and undamaged houses based on texture and edge features, while Jinno *et al.* (2024) proposed a method to quantitatively classify the degree of roof damage by performing instance segmentation to categorize roofs into four classes based on their level of damage. However, the 2D images obtained in these studies, regardless of their acquisition method, have a fundamental limitation in that they cannot capture the three-dimensional structural deformation of buildings. For instance, there is a risk of overlooking damage that is undetectable from nadir imagery, such as when an entire house is tilted even if the roof appears undamaged.

To overcome these limitations, the use of methods capable of acquiring 3D information, such as LiDAR and Structure from Motion (SfM), has garnered attention. Sun *et al.* (2024) applied deep learning to urban-scale point cloud data to achieve segmentation of roofs and walls and reconstruction of roof shapes. Additionally, Ntiyakunze & Inoue (2023) utilized the spatial dependencies in point clouds to automatically segment structural elements such as floors, columns, and beams. Furthermore, Xiu *et al.* (2020) detected collapsed buildings from airborne LiDAR point clouds using deep learning. These achievements are significant as they demonstrate the ability to capture building inclination and structural damage, overcoming a fundamental limitation of conventional 2D imagery: the inability to quantify 3D deformations. Although currently applied to urban planning and maintenance, we posit these 3D methods have substantial potential for the disaster management field, where they can be adapted for crucial tasks like rapid damage assessment and recovery planning.

This review of prior research reveals that while studies based on satellite and aerial imagery are effective for assessing wide-area damage and detecting roof damage, they have limitations in quantifying structural damage, such as the deformation and inclination of entire buildings. On the other hand, research based on 3D point clouds has demonstrated its effectiveness in component extraction and collapse detection. However, studies that use urban-scale drone point clouds to extract building components for the subsequent calculation of inclination, a critical step for damage assessment, remain limited. Therefore, this study develops a foundational method that targets high-resolution drone point clouds acquired at an urban scale to automatically classify components into Roof, Wall, Ground and Others, thereby establishing a foundational method for future applications in inclination calculation and damage assessment support.

Methodology

a. Study Area and Dataset:

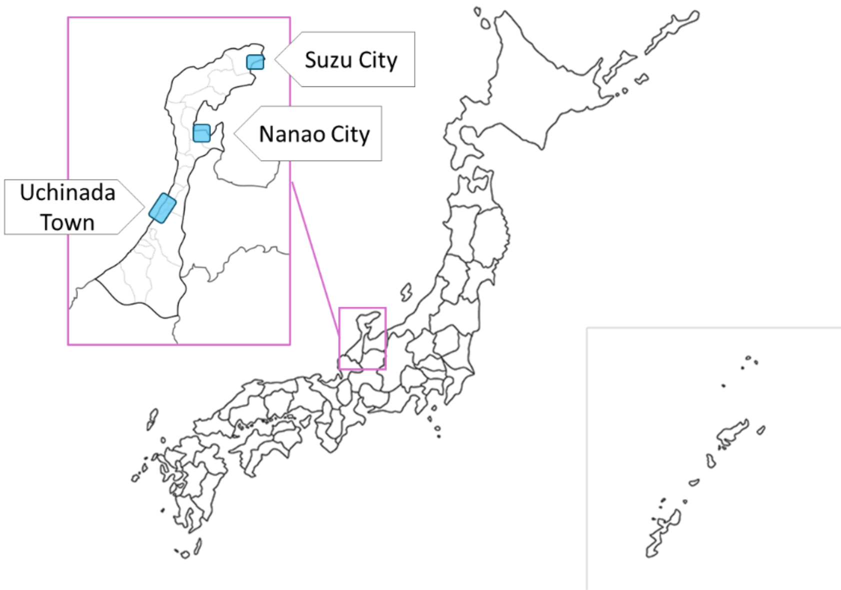

Figure 1: Study area

The dataset used in this study is a 3D point cloud acquired by a drone equipped with a LiDAR sensor. The study area comprises three municipalities affected by the 2024 Noto Peninsula Earthquake: Uchinada Town in Kahoku District, Suzu City, and Nanao City, all in Ishikawa Prefecture. The data were acquired from May 15 to 17, 2024, at an altitude of 25 m or 50 m.

For this research, a total of eight original data files from Nanao City and Uchinada Town were used for training, while four original data files from Suzu City were used for validation. The classification targets were divided into four classes: Roof, Wall, Ground, and Others, which includes objects such as utility poles and fences.

For the purpose of data augmentation, the training dataset was processed by applying a 5% random dropout, followed by a four-directional rotation (0, 90, 180, and 270 degrees) around the Z-axis. The total number of points in the training dataset is 63,217,352, with the class distribution being 27.0% for Roof, 10.1% for Wall, 35.5% for Ground, and 27.4% for Others, which includes utility poles, trees, and vehicles. The validation dataset consists of the original files from Suzu City without any data augmentation. The total number of points is 6,887,898, and the class distribution is 31.8% for Roof, 15.0% for Wall, 33.8% for Ground, and 19.4% for Others.

b. Proposed Method:

In this study, we adopt PointNet++ as the semantic segmentation method for 3D point cloud data. PointNet++ features an encoder-decoder architecture that extracts local features by downsampling the point cloud through multiple Set Abstraction (SA) layers and restores detailed point-wise features by upsampling through Feature Propagation (FP) layers. The model used in this study is composed of four SA layers and four FP layers.

For training, the Adam optimizer was used with a learning rate set to 0.001. The learning rate scheduler employs a Warmup strategy for the initial five epochs to stabilize training, followed by Cosine Annealing to decay the learning rate gradually. The CrossEntropyLoss function, with label smoothing applied, was used as the loss function. Furthermore, to address the class imbalance, we adjusted the weights by increasing the weight for the Wall class, which has a particularly small number of samples, from 1.520 to 1.976. The weights for the other classes were 1.095 for Roof, 1.000 for Ground, and 1.090 for Others. In urban point cloud data, the orientation of houses is consistent with respect to the direction of gravity; therefore, the random rotation during training, a data augmentation technique applied in the original model, was disabled.

c. Evaluation Metrics:

The performance of the constructed segmentation model is evaluated using Mean IoU (mIoU), Overall Accuracy (OA), and Mean Class Accuracy. Furthermore, for a more detailed analysis of each class's performance, supplementary metrics including Accuracy, Precision, Recall, and F1-score are also calculated. The formulas for each metric are shown in Equations (1)-(7), where *TP* (True Positive), *TN* (True Negative), *FP* (False Positive), and *FN* (False Negative) represent the number of true positive, true negative, false positive, and false negative points, respectively.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

$$\text{F1} - \text{score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{\text{Precision} + \text{Recall}} \tag{3}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$\text{IoU} = \frac{TP}{TP + FP + FN} \tag{5}$$

$$\text{mIoU} = \frac{1}{C}\sum_{i=1}^{C} \text{IoU}_i\,, \quad \text{where C is the number of classes} \tag{6}$$

$$\text{Overall Accuracy} = \frac{\sum_i \text{TP}_i}{\text{Total number of points}} \quad (7)$$

Results and Discussion

a. Segmentation Results:

Figures 2-7 show the ground truth labels and the segmentation results from our proposed method, visualized by class. Figures 2, 4, and 6 show the ground truth labels, while Figures 3, 5, and 7 show the corresponding segmentation results. In each figure, Roof are colored blue, Wall green, Ground red, and Others yellow.

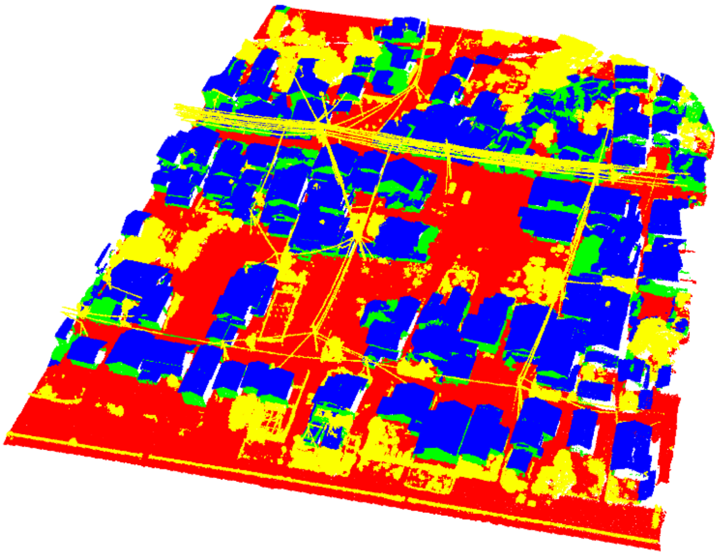

Figure 2: Ground truth labels for a representative scene in the validation area

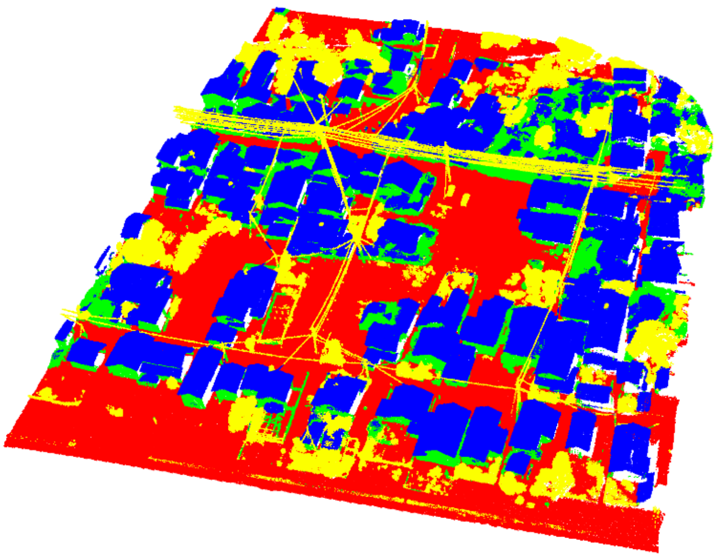

Figure 3: Segmentation results for a representative scene in the validation area

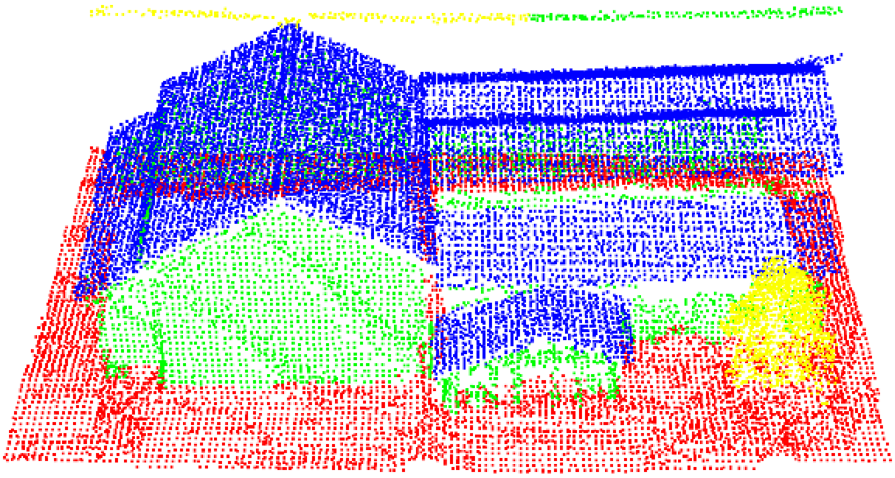

Figure 4: Ground truth labels of a house with minor damage

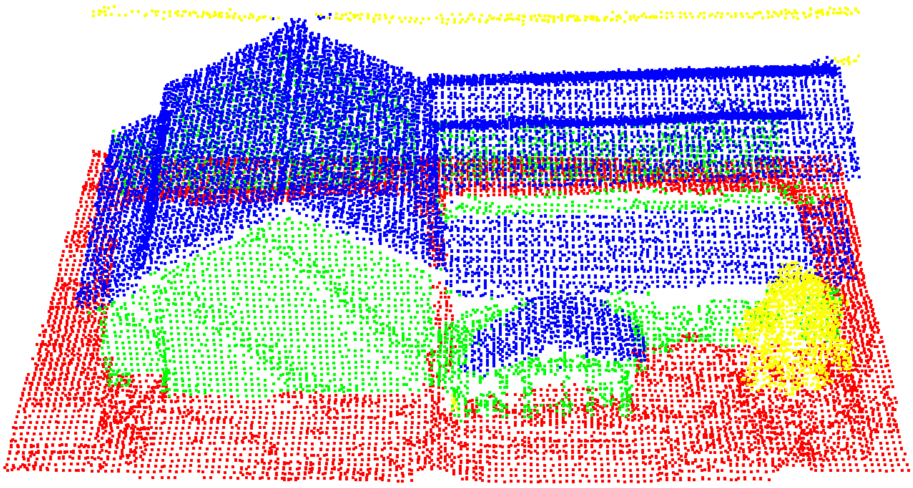

Figure 5: Segmentation results of a house with minor damage

Figures 2 and 3 show the results for a representative scene from the validation area. A comparison of the two figures reveals some misclassification of Wall (green) in the upper left of Figure 3. However, the remaining areas are largely consistent with the ground truth, demonstrating high fidelity. The study area contains numerous objects such as utility poles and fences, which were also classified with high accuracy as the Other classes.

Next, Figures 4 and 5 show a magnified view of a house with minor damage from the study area. The target house consists of multiple roof planes, and there are trees on the property, all of which are generally classified accurately. Notably, even small structures like the entrance porch roof (Geya) in the center foreground are appropriately classified as Roof for

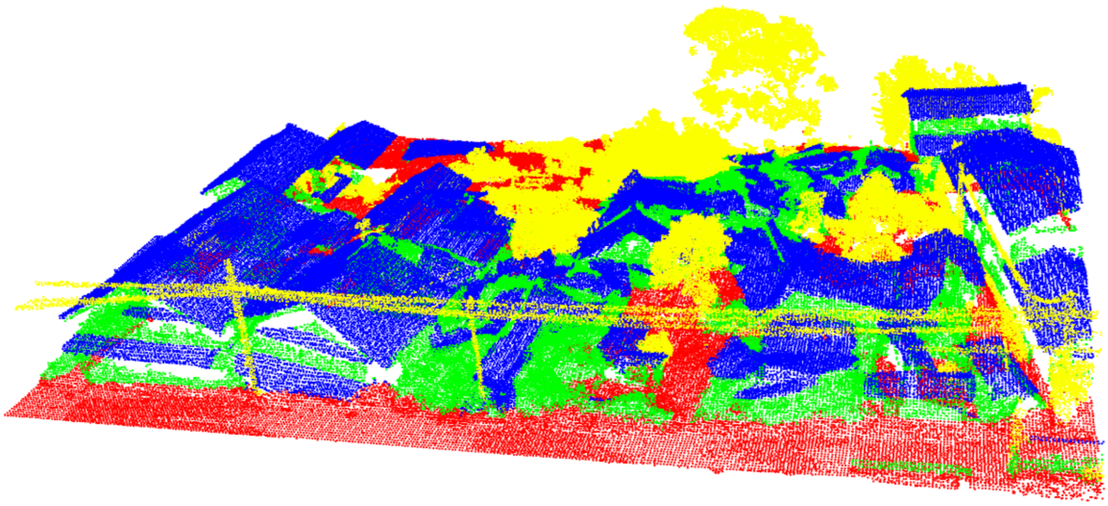

Figure 6: Ground truth labels of houses with severe damage

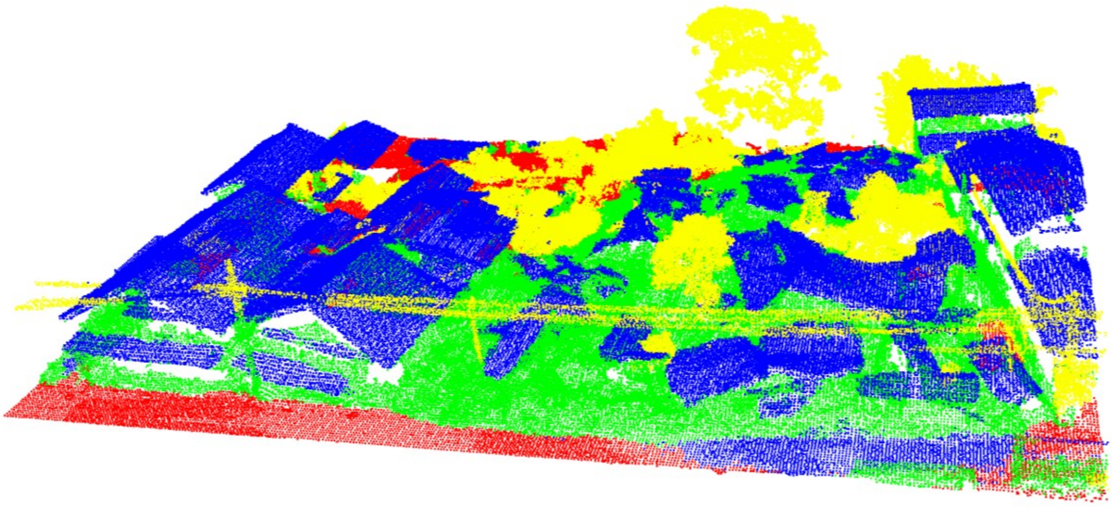

Figure 7: Segmentation results of houses with severe damage

the most part, although a portion is misclassified as Wall. This level of classification accuracy is considered sufficient for subsequent geometric analyses, such as the planned calculation of the house's inclination.

Subsequently, Figures 6 and 7 show a magnified view of an area containing collapsed and tilted houses. Within these figures, from left to right, are a house with minor damage, a house tilted due to the collapse of its first floor, and an area with numerous totally collapsed houses, representing a region with varying degrees of damage. The house with minor damage on the far left is accurately classified up to the boundaries of both its Roof and Wall. The tilted house in the center, whose first floor has collapsed, has its lower roof (Geya) partially misclassified as a Wall, but its main structure is correctly recognized. In the area on the right with densely packed, totally collapsed houses, while many structures are appropriately classified, a tendency to misclassify parts of the Ground as Wall was observed. This is presumed to be because the ground in that area was not flat and was scattered with debris. Furthermore, the high classification accuracy for the Other classes in this area is thought to be because many of the target objects were trees, and their characteristic color information contributed to the classification.

From these results, it was confirmed that the proposed method can extract structural components with high accuracy for houses with minor damage. However, for houses with large-scale damage close to total collapse, cases of misclassifying Roof or Ground as Wall were observed due to the influence of debris.

b. Quantitative Evaluation:

Next, the results presented in the previous section are quantitatively evaluated using performance metrics. Table 1 shows the overall performance metrics of the proposed method. For the entire validation dataset, the mIoU reached 77.3% and the OA was 88.6%. Furthermore, the Mean Class Accuracy, which is the average classification accuracy for each class, was also high at 87.1%, indicating that the proposed method possesses high classification performance.

Table 2 shows the detailed classification performance for each class. On a class-specific basis, Roof and Ground demonstrated exceptionally high IoU values of 84.8% and 87.0%, respectively. This is consistent with the visualization results in the previous section, where these classes were clearly distinguished. On the other hand, the IoU for Wall was low at

Table 1: Evaluation results based on performance metrics

Evaluation metric	Value (%)
mIoU	77.3
OA	88.6
Mean Class Accuracy	87.1

Table 2: Classification performance for each class

Class name	IoU (%)	Precision (%)	Recall (%)	F1-score (%)
Roof	84.8	92.2	91.4	91.8
Wall	59.4	68.7	81.8	74.6
Ground	87.0	94.2	91.9	93.0
Others	78.0	92.4	83.4	87.6

Figure 8: Confusion matrix

59.4%, quantitatively indicating that its classification was challenging. The F1-score for the Wall class was also the lowest among the four classes at 74.6%. Judging from the balance of Precision and Recall, this suggests that the model has more patterns of misclassifying other objects as Wall compared to other classes. This result aligns with the findings in the previous section, where instances of misclassifying parts of Roof or Ground as Wall were observed.

The primary factor for the lower classification performance of the Wall class compared to Others is considered to be the class imbalance in the training and validation datasets. As shown in the previous section, the proportion of Wall in the training dataset is 10.1%, which is less than one-third of the most frequent Ground class (35.1%). In this study, we applied weighting to the loss of the Wall class during training to mitigate this imbalance, but the IoU remained at 59.4%. Therefore, we believe that further data augmentation and the introduction of learning methods specialized for imbalanced data are necessary to improve the accuracy for the Wall class in the future.

To further analyze the misclassification trends between classes, the confusion matrix is shown in Figure 8. An examination of the misclassifications for the Wall class reveals that some points that were actually Roof or Others were classified as Wall. As previously mentioned, this is likely due to the misclassification of parts of roof eaves, adjacent fences, or debris as walls. It is considered that shape discrimination becomes difficult and misclassification is more likely to occur, especially in areas with sparse point clouds, such as at the eaves, or where debris is scattered on the ground.

c. Discussion:

The validation results demonstrated that the proposed method achieved a high IoU of approximately 85% for both the Roof and Ground classes, indicating its capability to classify major structural components with high accuracy from complex urban data. Notably, the method was also able to recognize and classify the structure of tilted houses to a certain extent, which is considered a significant step toward grasping the three-dimensional damage situations that are difficult to assess with conventional 2D imagery.

On the other hand, the classification accuracy for the Wall class was particularly low compared to other classes, with an IoU of 59.4%. This is attributed to the class imbalance, where the proportion of the Wall class in the training data is extremely small. Additionally, due to the nature of acquiring data by flying a drone in a nadir perspective, the number of points for vertical surfaces, such as walls, is inherently less than for horizontal surfaces, such as roofs and the ground. It is possible that there was insufficient data for the model to adequately learn the features of walls. Based on these results, it is necessary to improve the accuracy for the Wall class in the future by introducing learning methods specialized for class imbalance and models capable of capturing more localized information.

Conclusion and Recommendation

This study established a foundational method for automatically extracting structural components of damaged houses using deep learning on 3D point cloud data acquired by a drone. The validation results revealed that the proposed method demonstrates high performance in classifying Roof and Ground, and is capable of recognizing three-dimensional structures from complex scenes, including tilted houses. This achievement indicates the potential for directly evaluating the structural deformation of houses with drone point clouds, a task that is inherently difficult with conventional 2D imagery. On the other hand, the classification accuracy for the Wall class remains a significant challenge due to class imbalance in the dataset and the geometric constraints of data acquisition.

Future work should first address the challenge of improving the classification accuracy for the Wall class; improved data augmentation techniques and the introduction of learning methods specialized for imbalanced data are essential. Furthermore, an algorithm is required to separate and group the structural components belonging to individual houses by applying a mathematical method, such as clustering, to the results of the city-wide semantic segmentation achieved in this study. Based on this per-house component extraction, an algorithm will be constructed to quantitatively calculate the inclination of each house, which is the ultimate objective. Additionally, to achieve a more comprehensive damage assessment, the development of an integrated damage classification model is desired. This model would combine our 3D method with high-resolution 2D imagery acquired by a drone to capture fine-grained damage, such as the peeling of roof tiles, which is difficult to detect in point clouds. Ultimately, it will be necessary to validate this integrated model at different disaster sites to evaluate its generalization performance and practical utility.

Acknowledgements

This research was supported in part by the National Institute of Information and Communications Technology (NICT) under the program "Demonstrative R&D for Solving Social and Regional Issues through Promotion of Digitalization" (Project No. 233); by The Telecommunication Advancement Foundation through its 40th Anniversary Project research grant "Digital Society in the Age of AI (Technology and Challenges)"; by the Shikoku Create Association through the "FY2025 Research on Technological Development in Construction Projects"; and by the Institute of Education, Research and Regional Cooperation for Crisis Management Shikoku at Kagawa University through the "FY2025 Crisis Management Organization Project Activities and Research."

References

Attari, N., Ofli, F., Awad, M., Lucas, J., & Chawla, S. (2017). Nazr-CNN: Fine-grained classification of UAV imagery for damage assessment. *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 50–59. https://doi.org/10.1109/DSAA.2017.72

Fujii, M., Numada, M., Ohara, M., & Meguro, K. (2012). Study on implementation of building damage assessment of local governments after the 2011 off the Pacific Coast of Tohoku Earthquake. *Seisan Kenkyu*, 64(4), 433–437. https://doi.org/10.11188/seisankenkyu.64.433

Ishikawa Prefecture Headquarters of Jichiro (All Japan Prefectural and Municipal Workers Union). (2024). Survey results on mental health and other issues in municipalities affected by the Noto Peninsula Earthquake. (In Japanese). Retrieved September 28, 2025, from https://www.jil.go.jp/kokunai/topics/mm/20240911b.html

Jinno, M., Kubo, S., & Yoshida, H. (2024). Detection of damaged roofs using instance segmentation. Transactions of the Japan Society for Computational Methods in Engineering, 24, 97–105. https://doi.org/10.60443/jascome.24.0_97

Jozi, D., Shirzad-Ghaleroudkhani, N., Luhadia, G., Abtahi, S., & Gül, M. (2024). Rapid post-disaster assessment of residential buildings using Unmanned Aerial Vehicles. *International Journal of Disaster Risk Reduction,* 111, 104707. https://doi.org/10.1016/j.ijdrr.2024.104707

Kyushu Regional Administrative Evaluation Bureau, Ministry of Internal Affairs and Communications. (2018). Fact-finding survey on the issuance of disaster victim certificates in large-scale disasters: Focusing on the 2016 Kumamoto Earthquake. Report of findings. (In Japanese). Retrieved September 28, 2025, from https://www.soumu.go.jp/main_content/000528758.pdf

Li, Y., Hu, W., Dong, H., & Zhang, X. (2019). Building damage detection from post-event aerial imagery using single shot multibox detector. Applied Sciences, 9(6), 1128. https://doi.org/10.3390/app9061128

Ntiyakunze, J., & Inoue, T. (2023). Segmentation of structural elements from 3D point cloud. Sensors, 24(3), 1924. https://doi.org/10.3390/s23041924

Orihashi, Y., & Urakawa, G. (2022). A study on the present situation and issues of simplified building damage assessment methods for storm and floods based on interviews with local government officials. Journal of Natural Disaster Science, 41(S09), 237–252. https://doi.org/10.24762/jndsj.41.S09_237

Sendai City. (2012). Chapter 9: Disaster victim certificates and disaster relief benefits. Sendai City Earthquake Reconstruction Materials. (In Japanese). Retrieved September 28, 2025, from https://www.city.sendai.jp/shinsaifukko/shise/daishinsai/fukko/documents/06_c09-10.pdf

Sun, X., Guo, B., Li, C., Sun, N., Wang, Y., & Yao, Y. (2024). Semantic segmentation and roof reconstruction of urban buildings based on LiDAR point clouds. ISPRS International Journal of Geo-Information, 13(1), 19. https://www.mdpi.com/2220-9964/13/1/19

Tanaka, T. (2023). On the disaster victim support system and disaster victim certificate (1). Fukuoka University Law Review, 68(3), 303–326. https://fukuoka-u.repo.nii.ac.jp/record/2000134/files/J6803_0303.pdf

Tu, J., Li, D., Feng, W., Han, Q., & Sui, H. (2017). Detecting damaged building regions based on semantic scene change from multi-temporal high-resolution remote sensing images. International Journal of Geo-Information, 6(5), 131. https://doi.org/10.3390/ijgi6050131

Wang, Y., Chew, A. W. Z., & Zhang, L. (2022). Building damage detection from satellite images after natural disasters on extremely imbalanced datasets. Automation in Construction, 140, 104328. https://doi.org/10.1016/j.autcon.2022.104328

Xiu, H., Shinohara, T., Matsuoka, M., Inoguchi, M., Kawabe, K., & Horie, K. (2020). Collapsed building detection using 3D point clouds and deep learning. Remote Sensing, 12(24), 4057. https://doi.org/10.3390/rs12244057

Ye, S.-J., Nourzad, S., Pradhan, A., Bartoli, I., & Kontsos, A. (2014). Automated detection of damaged areas after Hurricane Sandy using aerial color images. In Computing in Civil and Building Engineering (pp. 1796–1803). https://doi.org/10.1061/9780784413616.223