

APPLICATION OF YOLOV8 FOR DETECTING SEDIMENT DISASTER PRONE AREAS FROM MAP IMAGES

OSUKE YAMASHITA¹, KEIGO MIWA¹, NORITAKA SHIGEI^{1,*}
SATOSHI SUGIMOTO², YOICHI ISHIZUKA² AND HIROMI MIYAJIMA¹

¹Graduate School of Science and Engineering
Kagoshima University
1-21-40 Korimoto, Kagoshima City, Kagoshima 890-0065, Japan
{ k3645249; k7306768; k2356323 }@kadai.jp

*Corresponding author: shigei@ibe.kagoshima-u.ac.jp

²Graduate School of Engineering
Nagasaki University
1-14 Bunkyo, Nagasaki City, Nagasaki 852-8521, Japan
{ s-sugi; isy2 }@nagasaki-u.ac.jp

Received November 2025; accepted January 2026

ABSTRACT. *In Japan, where approximately 70% of the land is mountainous, sediment-related disasters such as debris flows, landslides, and collapse of steep slopes occur frequently. To prevent such disasters, each prefectural government designates sediment disaster prone (SDP) areas; however, this process requires time-consuming and labor-intensive field surveys. Therefore, this study proposes a method for automatically detecting SDP areas using machine learning, with the aim of reducing the required effort. Specifically, the image segmentation model YOLOv8 was applied to map images. In addition, to improve detection accuracy, this study proposes the use of RRIM-BI images, which were created by integrating building information extracted from standard maps into Red Relief Image Maps (RRIM) that represent detailed landform morphology such as ridges and valleys. This is because building information, in addition to topographic features, is crucial for the designation of SDP areas. Furthermore, this study compares the performance of YOLOv8 and Mask R-CNN, and examines the influence of input image types, RRIM and RRIM-BI, on detection accuracy. The results showed that YOLOv8 outperformed Mask R-CNN, and that the use of RRIM-BI improved recall and F1-score for both debris flows and collapse of steep slopes. These findings indicate that the combination of YOLOv8 with RRIM integrated with building information is effective for detecting SDP areas.*

Keywords: Sediment disaster prone area, YOLOv8, Red Relief Image Map, Image segmentation

1. Introduction. In Japan, where approximately 70% of the land is mountainous, sediment disasters such as landslides, debris flows, and slope failures frequently occur due to typhoons, earthquakes, and heavy rainfall. For disaster prevention purposes, prefectural governments have designated sediment disaster prone (SDP) areas [1]. SDP areas refer to a region that is designated as being at risk of posing a threat to the lives and safety of residents in the event of sediment disasters. Although these areas are made public after fundamental surveys, the designation process requires on-site investigations by experts, demanding considerable time and effort.

To reduce this effort, deep learning has been applied to detecting hazardous areas from map data [2, 3, 4, 5, 6, 7, 8]. In [2], an attempt has been made to detect the location of collapse hazard areas on steep slopes from Red Relief Image Map (RRIM) [9] images and aerial photographs using convolutional neural network (CNN). In [3] using pix2pix,

it was found that using SDP areas as the ground truth, rather than specific points of risk of collapse of steep slope, resulted in greater accuracy of the detection. However, pix2pix tended to predict almost the entire image as a warning zone. In [4], by using Mask R-CNN, collapse of steep slopes and debris flows were detected through segmentation based on input images such as RRIM, aerial photographs, land-use maps, and standard maps [10]. By combining these maps as input to the model such as building information, it is possible to provide more information, and it has been shown that this approach tends to achieve higher accuracy compared to using RRIM alone. In addition, Ishii et al. developed deep learning models to automatically delineate SDP areas from digital elevation models (DEMs) and examined the effectiveness of different training datasets; they further extended this approach by incorporating social factors through building information to adjust warning zones so that they actually include conservation targets [5]. Unlike the method of [4], this approach uses only DEM data as input to the deep learning model. Beyond SDP areas, deep learning has also been applied to landslide surface deformation and relict landslide detection, for example, by integrating YOLOX and Mask R-CNN to identify landslide surface deformation [7] and by employing deep-learning segmentation architectures for relict landslide detection in rainforest regions [8], as well as to deep-seated landslide susceptibility mapping in Japan using automatically constructed CNN architectures with pre-slide DEMs [6]. This line of work, unlike that in [4] and [5], focuses on the detection and characterization of landslide topography.

Additionally, a comparative study of YOLOv8 and Mask R-CNN in segmentation tasks showed that YOLOv8 outperformed Mask R-CNN in terms of F1-score [11]. Despite these efforts, the optimal combination of a highly efficient instance segmentation model and effective feature representation, particularly when integrating both terrain and building information for SDP area detection, remains an open question in this domain.

In this paper, we aim to improve the detection accuracy of SDP areas by applying YOLOv8 as an instance segmentation model and by developing a novel method for creating input images that incorporate building information into RRIMs. We first recognize that the hue (H) channel of RRIM in HSV color space holds minimal topographic information. By replacing the H channel of RRIM images in HSV color space with building information extracted from standard maps, we propose a technique that enables the integration of terrain and building features in a single image. Through these approaches, this study seeks to contribute to the development of more effective and efficient methods for SDP area detection.

2. Detection of SDP Areas by Image Segmentation. In this study, sediment disaster prone (SDP) areas are detected from map images such as RRIMs and standard maps using image segmentation. Image segmentation is a technique to predict the class and location of objects within an image at the pixel level. In [4], Mask R-CNN model is used as a model for image segmentation. This paper proposes applying YOLOv8 to detecting SDP areas.

The following provide an overview of the SDP areas, the map images used for detection, and the YOLOv8 model employed for detection.

2.1. SDP area. SDP area refers to a region recognized as being at risk of harm to human life or health in the event of a sediment disaster, such as a collapse of steep slope. In Japan, the government has enacted the Sediment Disaster Prevention Act, under which each prefecture publicly designates areas with potential sediment disaster risks as SDP areas. These designations are made based on the “Act on the Promotion of Sediment Disaster Countermeasures in Sediment Disaster Prone Areas”, which came into effect on April 1, 2001. Figure 1 shows the types of sediment disaster that can lead to designation

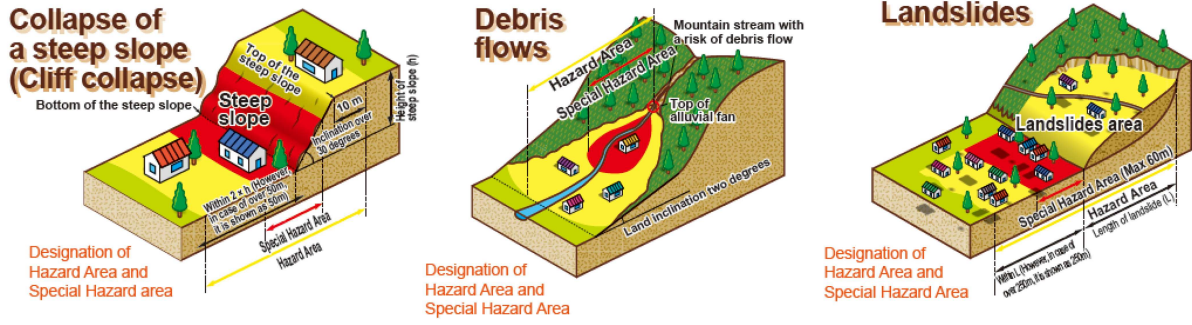


FIGURE 1. Disasters designated for SDP areas [12]

as an SDP area. There are three types of disaster for which an area may be designated: collapse of steep slope, debris flows, and landslides.

2.2. Map images. The maps used in this study are of three types: RRIMs, standard maps, and SDP area maps. RRIMs and standard maps are employed for the input of YOLOv8. Both images are generally provided as RGB color images. The RRIM, as shown in Figure 2(a), uses the saturation in red which indicates the amount of slope, and the brightness in red indicates the degree of the valley of the ridge, making it easier to visually understand the slope and valley of the terrain. The standard map, as shown in Figure 2(b), is a topographical map issued by the Geospatial Information Authority of Japan, where features such as terrain, contour lines, houses, rivers, and map symbols are depicted. Standard maps are provided in vector format, which allows easy selection of displayed information. In this study, such vector maps are utilized. The SDP area map, as shown in Figure 2(c), is used as the ground-truth and as a mask image for training. This map is issued by local governments based on administrative designations, showing areas designated as SDP areas. The SDP area map images can be provided in black and white.

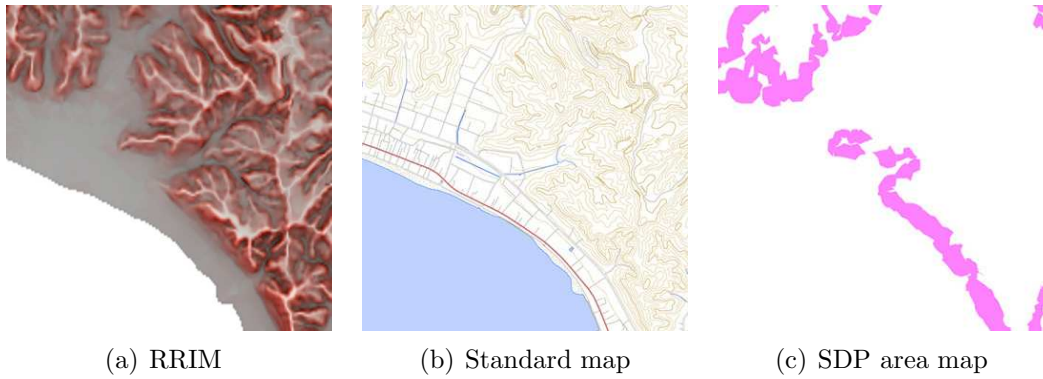


FIGURE 2. (color online) Three types of maps used in this study

2.3. YOLOv8. This paper employs the YOLOv8 model for detecting SDP areas. YOLOv8 is an extended version of the object detection model YOLO [13, 14]. In addition to object detection, it provides capabilities for classification, segmentation, and pose detection.

YOLOv8 introduces several architectural innovations to enhance performance in both object detection and instance segmentation. These include an anchor-free, center-based detection approach, the adoption of pseudo ensemble/pseudo supervision (PS) to improve prediction robustness, and the utilization of the Darknet-53 backbone. Furthermore, the model replaces the C3 module with the more efficient C2F module and reduces the number

of parameters by substituting 6×6 convolutional layers with 3×3 layers. Furthermore, YOLOv8 employs a decoupled head structure. Through these architectural improvements, YOLOv8 achieves improvements in both accuracy and inference speed.

The input for the YOLOv8 model is typically a 3-channel color image. Since both the RRIM images and the standard map images are provided as 3-channel color images, only one of them can be used at a time unless modifications are made. The next section presents a method for generating input images by effectively combining these images to enhance detection accuracy.

3. RRIM Images with Building Information. This section describes a method for creating input images for YOLOv8, in which layer information extracted from RRIM images and standard map images is combined to effectively detect SDP areas.

SDP areas are designated not only because of the potential for disasters such as debris flows, landslides, and collapse of steep slopes but also because there are buildings, such as residences, in those areas, posing a risk of physical harm to people. However, the use of RRIMs alone does not allow for obtaining information about buildings. The proposed method generates images that incorporate building information while preserving the information contained in the RRIMs. These images are used as input images for YOLOv8 in this application.

RRIM represents the amount of slope using red saturation and the ridge-valley characteristics using brightness. Therefore, to obtain slope and ridge-valley degree, it is reasonable to convert RRIM images in RGB format into the HSV color space and extract their saturation and value components.

HSV images represent the color space using three components: hue, saturation and value of brightness. Let $C_{\max} = \max(R, G, B)$ and $C_{\min} = \min(R, G, B)$. Then, the conversion formulas from RGB to HSV are as follows:

$$H = \begin{cases} 60 \times \frac{G - B}{C_{\max} - C_{\min}} & \text{if } C_{\max} = R \\ 60 \times \frac{B - R}{C_{\max} - C_{\min}} + 120 & \text{if } C_{\max} = G \\ 60 \times \frac{R - G}{C_{\max} - C_{\min}} + 240 & \text{if } C_{\max} = B \\ 0 & \text{if } (R, G, B) = (0, 0, 0) \end{cases} \quad (1)$$

$$S = \frac{C_{\max} - C_{\min}}{C_{\max}} \quad (2)$$

$$V = C_{\max} \quad (3)$$

$$C_{\max} = \max(R, G, B) \quad (4)$$

$$C_{\min} = \min(R, G, B) \quad (5)$$

Figure 3 shows the RGB-format images in which the H, S, and V components of the RRIM images are assigned to the R, G, and B channels, respectively, where the pixel values of the R, G, or B channels to which none of the H, S, or V channels are assigned are set to 0.

While all channels in RGB format contain information, the HSV representation reveals that the hue channel contains minimal information in RRIM images. In contrast, the primary terrain information is mainly captured in the saturation and value channels. Based on this observation, inserting building information into the hue channel enables the integration of building features into the model input without compromising the original terrain information of the RRIM.

In the proposed method, vector tiles, which are vector-based standard maps, are used to obtain building information. Unlike raster image formats, vector tiles store map elements

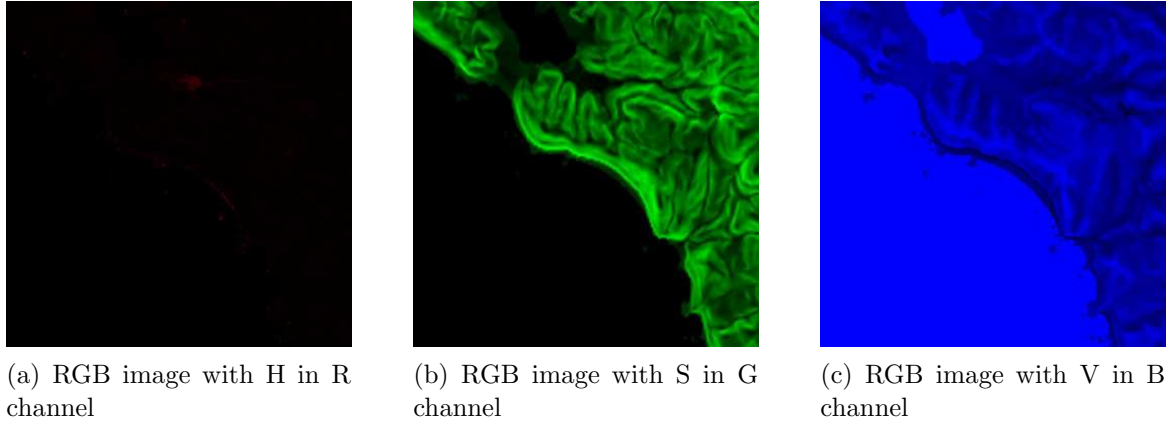


FIGURE 3. (color online) Visualization of each channel image in the proposed method

such as buildings, roads, rivers, and map symbols as text-based information rather than as images. Therefore, it is possible to extract the stored map elements easily and without degradation. The position and shape of buildings are extracted from the vector tiles, and a single-channel image is generated.

The proposed method consists of the following four steps.

Step 1: Extract building information from a standard map

From the vector tiles, the position and shape of buildings are extracted.

Step 2: Binarize the building information

Based on the extracted building data, a binary image is generated by assigning a pixel value of 255 to building areas and 0 to all other areas. This clearly indicates the presence or absence of buildings.

Step 3: Convert the RRIM to HSV format

The RRIM is converted from RGB to HSV color space based on Equations (1), (2) and (3).

Step 4: Combine the S and V channels of the HSV image with the binarized building image

The saturation (S) and value of brightness (V) channels are extracted from the converted RRIM. The binarized building image obtained in Step 2 is inserted in place of the hue (H) channel. The 3-channel image used in this study is composed of images similar to those shown in Figure 3(b). $\square$

For Steps 1 and 2, we provide additional details of our implementation. In Step 1, building information is first extracted from vector tiles [10] and rasterized into a single-channel grayscale map image of size 512×512 pixels. To generate the rasterized map image, all 256×256 map tiles required for a 512×512 image centered on the target coordinates are retrieved at Zoom Level 15. The tile containing the target coordinates and its eight neighboring tiles are retrieved, and a 512×512 map image centered on the target coordinates is generated from these tiles. Zoom Level 15 was selected because building locations cannot be identified reliably at Zoom Level 14 or lower. The RRIM images used in conjunction with the map images are also prepared at a spatial resolution equivalent to Zoom Level 15. Note that all input images are resized from 512×512 pixels to 640×640 pixels to match the input size required by YOLOv8.

For our implementation, the binary building image is stored in the first channel, the RRIM (S) channel in the second channel, and the RRIM (V) channel in the third channel. However, it is considered that the order of this channel arrangement does not affect the learning performance of the model.

4. Evaluation.

4.1. **Evaluation conditions.** The evaluation metrics are as follows:

$$\begin{aligned} \text{precision} &= \frac{TP}{TP + FP}, \\ \text{recall} &= \frac{TP}{TP + FN}, \\ \text{F1-score} &= \frac{2 \times \text{recall} \times \text{precision}}{\text{recall} + \text{precision}}, \end{aligned}$$

where TP , FP , and FN are the numbers of pixels for true positive, false positive, and false negative, respectively, in terms of the pixels in the mask images for the segmentation task.

Input images consist of patch images cropped from RRIM and standard maps at Zoom Level 15. A patch image at Zoom Level 15 corresponds to a region of approximately 1.3 km on each side in Japan. Additionally, among the disasters specified in the SDP areas, landslides are excluded due to the small number of cases and the fact that they occur underground. The two disasters considered for SDP areas detection are collapse of steep slope and debris flows.

The regions used for training, validation, and testing correspond to mesh numbers in the first-level mesh map: 4730, 4731, 4830, 4831, 4929, 4930, 4931, 4932, 4933, 5029, 5030, 5031, 5032, 5033, 5034, 5129, 5131, 5132, 5133, 5134, 5229, 5232, 5233, 5234, 5333, and 5334, which approximately correspond to the Kyushu and Chugoku regions in Japan. A mesh map refers to a grid-based regional framework defined on a map using latitude and longitude, established for the purpose of digitizing geographical information and collecting various types of statistical data. The first-level mesh map is based on the meridian at 100 degrees east longitude and the parallel at 0 degrees north latitude, and it divides the map into rectangular areas using each degree of longitude and latitudinal lines created by trisecting the intervals between even-numbered degrees of latitude. The training data includes regions with mesh numbers 4929, 4930, 4931, 4932, 4933, 5029, 5030, 5031, 5032, 5033, 5034, 5129, 5131, 5132, 5133, 5134, 5229, 5232, 5233, 5234, 5333, and 5334. The validation data uses regions with mesh numbers 4830 and 4831, while the test data uses regions with mesh numbers 4730 and 4731. The number of images is 49,610 for training, 4,510 for validation, and 4,510 for testing. The image size is 640×640 pixels, the learning rate is set to 0.01, and the number of training epochs is 100. Also, the confidence score threshold for detection was set to 0.05 for debris flows and 0.01 for collapse of steep slope. The confidence score represents the threshold used to determine whether a detected object is considered positive or negative. All other settings were kept at their default values in the experiments.

4.2. **Results.** This study conducts two types of comparison. The first is a comparison of detection models, in which Mask R-CNN and YOLOv8 are compared using RRIM as input images. The second is a comparison of input images, in which the performance of YOLOv8 is evaluated by comparing the use of RRIMs as input images with the use of RRIM-BI images generated by the proposed method.

The evaluation results are shown in Table 1 and Table 2. Tables 1 and 2 are for debris flows and collapse of steep slope, respectively.

The first comparison discusses the relative merits of Mask R-CNN and YOLOv8 in terms of accuracy when RRIMs are used as input images. In the case of debris flows, although recall is 7.49 percentage points (pp.) lower than that of Mask R-CNN, precision and F1-score improve by 3.67 pp. and 1.73 pp., respectively. For collapse of steep slope, improvements are observed in all three evaluation metrics: recall, precision, and F1-score,

TABLE 1. Evaluation results for debris flows

Model	Input	Recall (%)	Precision (%)	F1-score (%)
Mask R-CNN	RRIM	56.98	23.88	33.66
YOLOv8	RRIM	49.49	27.55	35.39
	RRIM-BI	78.89	24.10	36.92

TABLE 2. Evaluation results for collapse of steep slope

Model	Input	Recall (%)	Precision (%)	F1-score (%)
Mask R-CNN	RRIM	44.75	20.29	27.92
YOLOv8	RRIM	45.98	24.77	32.19
	RRIM-BI	76.17	27.13	40.01

by 1.23 pp., 4.48 pp., and 4.27 pp., respectively. These results suggest that YOLOv8 is generally effective as a segmentation model for SDP area detection.

The second comparison discusses the relative accuracy when using RRIMs and RRIM-BI images generated by the proposed method as input images for YOLOv8. For both debris flows and collapse of steep slope, the proposed method using RRIM-BI images achieves higher recall and F1-score values than the conventional method using RRIM, although precision is lower for debris flows. Specifically, for debris flows, recall increased by 29.40 pp. and F1-score by 1.53 pp., whereas precision decreased by 3.45 pp. Despite the decrease in precision, the substantial increase in recall resulted in a higher F1-score, indicating an overall improvement in detection performance. For collapse of steep slope, recall, precision, and F1-score increased by 30.19, 2.36, and 7.82 pp., respectively. These results indicate that building information is an important factor in SDP area detection and confirm the effectiveness of the proposed method.

5. Conclusion. This paper explores the application of YOLOv8 for realizing an effective method for detecting SDP areas using image segmentation. The results demonstrated that the proposed YOLOv8-based method outperformed the conventional Mask R-CNN-based model in terms of detection accuracy. Furthermore, in order to generate the input images called RRIM-BI, we propose to convert RRIM images to HSV format and replace the H channel with an image representing building information extracted from a standard map. The evaluation result also demonstrates the effectiveness of the proposed input images.

In summary, this approach avoids the fundamental modification of model input channels, which usually requires extensive effort and advanced techniques, thereby facilitating the application of complex disaster detection models and contributing to the acceleration of both research and practical implementation.

Despite these advantages, the proposed method still has limitations. First, it does not incorporate other geographical information essential for sediment disaster occurrence, such as rivers and contour lines. Second, because high-performance machine learning models typically require three-channel inputs, it is inherently difficult to fully integrate and represent all topographic information from RRIM, building information, and various other geographical features within this limited input space. Therefore, future work will explore new data integration and feature representation methods that can more effectively provide this multidimensional information to the model.

Acknowledgment. This work was partially supported by JSPS KAKENHI Grant Number JP25K15275.

REFERENCES

- [1] Ministry of Land, Infrastructure, Transport and Tourism, *Designation of Sediment Disaster Alert Areas*, <https://www.mlit.go.jp/mizukokudo/sabo/linksinpou.html>, Accessed on 14 May, 2025 (in Japanese).
- [2] Y. Yamashita, N. Shigei, S. Sugimoto, R. Takaesu and Y. Ishizuka, Estimation of landslide hazard areas from 3D maps and aerial photographs using CNN, *Proc. of Japan Society of Civil Engineers West Annual Meeting*, pp.339-340, 2019 (in Japanese).
- [3] Y. Nishi and N. Shigei, Estimation methods of slope failure hazard areas by using pix2pix, “*Hinokuni – Land of Fire*” *Information Processing Symposium*, vol.A11-2, 2022 (in Japanese).
- [4] K. Miwa, N. Shigei, S. Sugimoto, Y. Ishizuka and H. Miyajima, Estimation of sediment disaster prone areas from map images using mask R-CNN, *ICIC Express Letters, Part B: Applications*, vol.15, no.6, pp.575-584, 2024.
- [5] T. Ishii, M. Hirokane, K. Kuramoto and N. Nishihara, Considering social factor in setting of sediment disaster warning area using deep learning, *Journal of Japan Society of Civil Engineers, Ser. F6 (Safety Problem)*, vol.76, no.2, pp.I-193-I-199, 2020 (in Japanese).
- [6] T. Kikuchi, K. Sakita, S. Nishiyama and K. Takahashi, Landslide susceptibility mapping using automatically constructed CNN architectures with pre-slide topographic DEM of deep-seated catastrophic landslides caused by Typhoon Talas, *Natural Hazards*, vol.117, 2023.
- [7] C. Wan, J. Gan, A. Chen, P. Achaya, F. Li, W. Yu and F. Liu, A novel method for identifying landslide surface deformation via the integrated YOLOX and Mask R-CNN model, *International Journal of Computational Intelligence Systems*, vol.17, 255, 2024.
- [8] G. P. B. Garcia, L. P. Soares, M. Espadoto and C. H. Grohmann, Relict landslide detection using deep-learning architectures for image segmentation in rainforest areas: A new framework, *International Journal of Remote Sensing*, vol.44, no.7, pp.2168-2195, 2023.
- [9] Asia Air Survey Co., Ltd., *RRIM*, <https://www.rrim.jp/en/>, Accessed on 30 April, 2023.
- [10] *Geospatial Information Authority of Japan (GSI) Website*, <https://maps.gsi.go.jp/vector/>, Accessed on 18 April, 2025 (In Japanese).
- [11] R. Sapkota, D. Ahmed and M. Karkee, Comparing YOLOv8 and Mask R-CNN for instance segmentation in complex orchard environments, *Artificial Intelligence in Agriculture*, pp.84-99, 2024.
- [12] R. Abdulharis, A. P. Handayani, C. Isouchi and I. Meilano, Developing community disaster resilience in the Lembang Fault area, Indonesia: Lessons learned from Japanese experience, *Applied Sciences*, vol.12, no.3, 1271, 2022.
- [13] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, You only look once: Unified, real-time object detection, *arXiv Preprint*, arXiv: 1506.02640v5, 2016.
- [14] Ultralytics, *Explore Ultralytics YOLOv8*, <https://docs.ultralytics.com/models/yolov8>, Accessed on 19 November, 2025.