

Article

Proposal of a Disrupted Road Detection Method in a Tsunami Event Using Deep Learning and Spatial Data

Jun Sakamoto

Faculty of Science and Technology, Kochi University, Kochi 780-8520, Japan;
jsak@kochi-u.ac.jp; Tel.: +81-88-844-8092

Abstract: Tsunamis generated by undersea earthquakes can cause severe damage. It is essential to quickly assess tsunami-damaged areas to take emergency measures. In this study, I employ deep learning and develop a model using aerial photographs and road segment data. I obtained data from the aerial photographs taken after the Great East Japan Earthquake; the deep learning model used was YOLOv5. The proposed method based on YOLOv5 can determine damaged roads from aerial pictures taken after a disaster. The feature of the proposed method is to use training data from images separated by a specific range and to distinguish the presence or absence of damage related to the tsunami. The results show that the proposed method is more accurate than a comparable traditional method, which is constructed by labeling and learning the damaged areas. The highest F1 score of the traditional method was 60~78%, while the highest F1 score of the proposed method was 72~83%. The traditional method could not detect locations where it is difficult to determine the damage status from aerial photographs, such as where houses are not completely damaged. However, the proposed method was able to detect them.

Keywords: aerial photograph; deep learning; disrupted section; GIS; YOLO

Citation: Sakamoto, J. Proposal of a Disrupted Road Detection Method in a Tsunami Event Using Deep Learning and Spatial Data. *Sustainability* **2023**, *15*, 2936. <https://doi.org/10.3390/su15042936>

Academic Editors: Stefano Morelli, Veronica Pazzi and Mirko Francioni

Received: 14 January 2023

Revised: 2 February 2023

Accepted: 3 February 2023

Published: 6 February 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The Great East Japan Earthquake that occurred on 11 March 2011, caused severe damage over a wide area. The municipalities damaged by the tsunami could not assess, report, and transmit information because of the disruption of communication systems and the collapse of government buildings; in addition, the safety of their leaders and employees was threatened [1]. The more severely damaged areas were, the more difficult it was to transmit and collect information; hence, it was difficult knowing whom to contact to have countermeasures taken. The Nankai Trough earthquake, which has a 0.7–0.8 probability of occurring within the next 30 years, is expected to cause a massive tsunami of more than 10 m in height over a wide area along the Pacific coast from the Kanto region to the Kyushu region [2]. Methods are required for an early warning of tsunamis and a quick assessment of the damage caused by tsunamis.

Early tsunami warnings in coastal areas enable timely evacuation. Accurate and rapid prediction of impending tsunamis is essential to mitigate damage to human life and property [3,4].

Assessing the damage after a natural disaster provides essential information for determining rescue priorities, guiding victims to safe locations, and estimating the amount of damage [5]. Aerial photographs can provide a broader range of damage information in a smaller sample than land photographs [6,7]. Previous studies have considered many methods to identify damage from remote sensing images. These methods can be classified as multi-temporal and single-temporal assessment methods.

Multi-temporal assessment methods identify damage by detecting changes. The authors of [8] extracted earthquake damage information using high-resolution remote sensing images before and after the 2010 Yushu earthquake in Qinghai. The results showed that the object-oriented change detection method could extract damage conditions with high accuracy. The authors of [9] compared the roofs of buildings before and after the 2021 earthquake in Yangbi County, Dali Prefecture, and Yunnan Province. It was found that the investigation time to detect damage was significantly shorter than the manual investigation. However, it was a limited evaluation due to the angle and time constraints of capturing images [10].

Single-time assessment methods are less data constrained because they only analyze damage from post-earthquake remote sensing. The authors of [11] identified landslides by integrating nighttime light, multi-seasonal, and elevation data and by using neural networks to classify satellite imagery. However, factors such as noise and illumination in remote sensing images seriously affect detection accuracy [12], resulting in the inability to extract building information accurately.

Computer vision methods have been widely used to investigate the extent of disasters; they are mainly categorized into two types: stereotype and deep learning (DL) methods [13]. Stereotype methods usually rely on manually designed models from features such as color, texture, contours, and edges. However, these are highly subjective and often vary widely from scene to scene, which limits their applicability. Hence, DL has been brought into the spotlight [14]. Convolutional neural network (CNN) models can eliminate many processes involved in determining disaster damage. CNN can process low-level characteristics through deep structures to obtain high-level semantic information. Compared with hand-crafted features, the high-level information is more abstract and robust. Several studies have been conducted on post-disaster building damage detection using remote sensing images and DL. The authors of [15] proposed a method to extract damage information from a group of buildings in post-earthquake remote sensing images by combining CNN and geographic information system (GIS) data. The authors of [16] proposed a method to detect post-disaster building damage using only pre-disaster images of buildings. The authors of [17] proposed a method to detect objects on building roofs, vehicles, debris, and flooded areas from post-disaster aerial video footage. In addition, a building damage detection method has been proposed and demonstrated using a ground-based imagery dataset [18].

Despite numerous trials, automatic disaster detection still needs improvement for a number of reasons. First, methods that do not use aerial photographs can determine which areas in the photographs are affected but cannot decide which regions on the map are affected. Second, methods using aerial photographs cannot distinguish a road disruption if it is covered by water because it is unknown whether a road exists on the map. Finally, there are limited publicly available image datasets depicting structural damage from disasters [19]; moreover, the damage caused by a tsunami is more complex than a house collapsing, making it challenging to study.

In this study, I use aerial photographs to determine which roads were damaged by the Great East Japan Earthquake. I apply a learning model developed using aerial photographs of a tsunami-damaged area to aerial photographs of other sites and verify the model's fit. Then, I visualize the road disruption by determining the presence or absence of damage on a mesh-by-mesh basis. In other words, the features of this study are as follows: The first is to add road information to the disaster area by overlaying aerial photographs and road segment data. The second is to attempt to identify tsunami damage with high accuracy by learning and applying mesh-based learning to complex tsunami disasters.

2. Materials and Methods

The training data for a traditional object detection model comprise definitions of where objects exist in each image. However, when there are various types of damage, such

as damage caused by a tsunami, it may be challenging to identify the presence or absence of damage from such training data. The method proposed in this study identifies tsunami damage by developing a learning model using data from segmented photographs classified according to whether they were damaged by the tsunami. I demonstrated the significance of the proposed method by comparing its detection accuracy with that of a traditional method.

2.1. Method Flowchart

Figure 1 shows the flowchart of this study. Both the traditional and proposed methods use the same training and test images. After collecting aerial photographs of the study site (Step 1), the tsunami inundation status for the training data was defined by referring to the inundation estimation map published by the Geospatial Information Authority of Japan [20] (Step 2). The traditional method uses training photographs to label which areas are affected by the tsunami (Figure 2). Meanwhile, the proposed method divides the training photographs into 100 m image units and classifies each image as having tsunami damage.

Before setting the unit to 100 m, I examined the relationship between unit size and computation time and found that a smaller range increases computation time but improves accuracy. For example, on a Core i7 1195G7 (Tiger Lake)/2.9 GHz/4-core computer with 16 GB of memory, the computation time to handle 100 m mesh (2108 images) and 500 m mesh (84 images) images were 14 min 40 s, and 30 s, respectively. The F1 scores of the 100 m mesh (2108 images) and 500 m mesh (84 images) images were 85% and 50%, respectively.

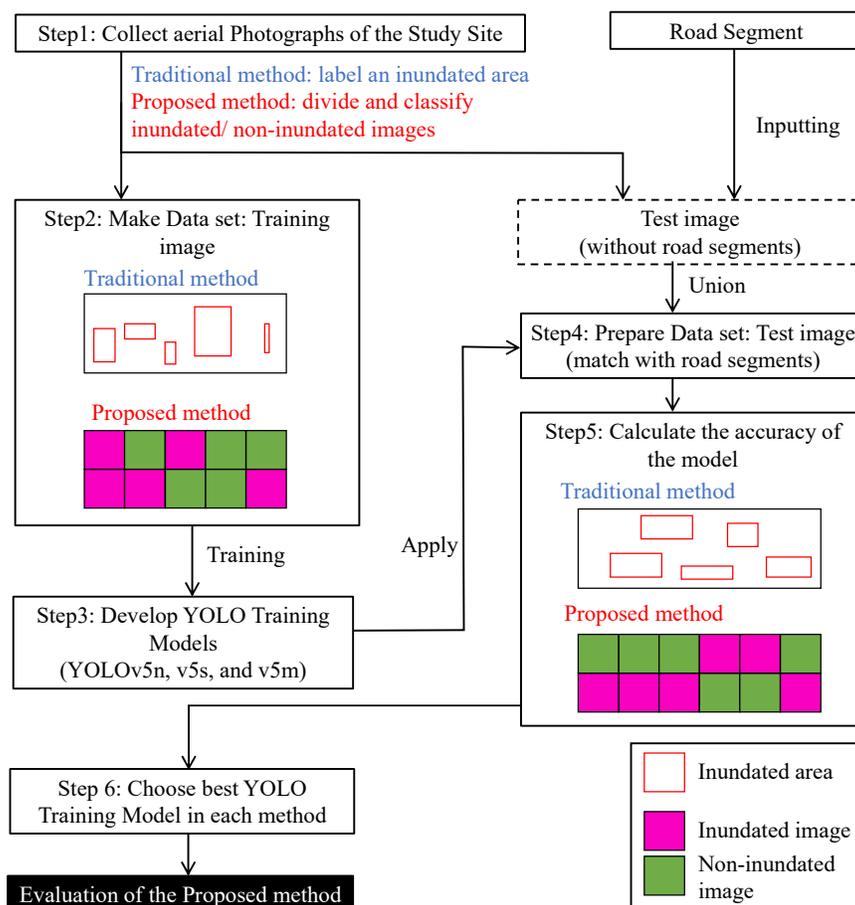


Figure 1. Method flowchart.

2.2. Outline of YOLOv5 Model

Joseph Redmon [23] proposed the YOLO target detection algorithm in 2015. It is an end-to-end network model that directly predicts a target's bounding box and category. YOLO considers object detection a single regression problem, replacing image pixels with bounding box coordinates and class probabilities. Using this, one only needs to look at an image once to predict what object is where.

In this study, I selected YOLOv5, which was released in 2020 [24]. It is lightweight and has good advantages for detecting small objects in terms of accuracy and speed. In addition, because it integrates the anchor box selection process, it can learn the best anchor box for a given dataset automatically and use it during training without considering the dataset as input. The anchor box described here is a list of predefined boxes that best match the desired objects. YOLOv5 network predicts bounding boxes as deviations from a list of anchor box dimensions [25]. YOLOv5 outperforms YOLOv4 and YOLOv3 in terms of accuracy [26].

Figure 4 shows the YOLOv5 architecture; it comprises a backbone, neck, and head [27]. The backbone extracts the essential features from the input images. The CSP1-x structure is incorporated into DarkNet to create CSPDarknet, the backbone of YOLOv5. CSPDarknet extracts feature from images comprising CSP1-x networks. Point Operations per Second (FLOPS) features can develop smaller model sizes while ensuring inference speed and accuracy. The neck is a series of network layers that mix and combine image features. The head predicts image features, generates bounding boxes for detection, and predicts the target object type. The CSP2-X structure used here enhances network feature fusion capabilities; for multiscale prediction, the head generates feature maps of three different sizes: 80×80 grid cells, 40×40 grid cells, and 20×20 grid cells. Detection results include class, score, location, and size [28].

YOLOv5 delivers various types of models, e.g., YOLOv5n (nano), YOLOv5s (small), YOLOv5m (medium), YOLOv5l (large), and YOLOv5x (large) [29]. As described later in Chapter 3, YOLOv5m (medium) took about 8.8 h to train using the proposed method. Prior attempts using YOLOv5l (large) and YOLOv5x (large) failed because the Google Colaboratory session timed out during the calculation. Considering the trial, I compare the training results of YOLOv5n, YOLOv5s, and YOLOv5m and then implement the best model on the test data.

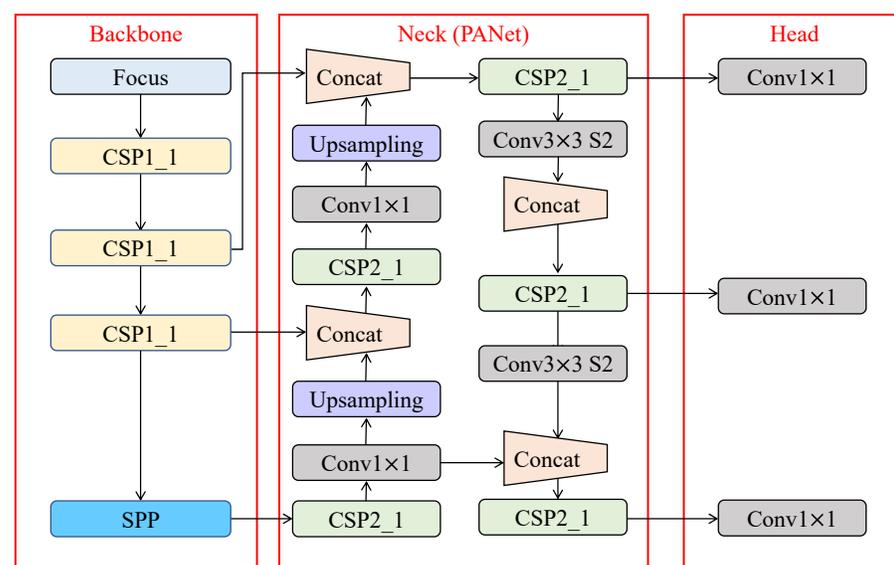


Figure 4. YOLOv5 architecture.

2.3. Outline of Google Colaboratory

Google Colaboratory is a service designed to educate and promote machine learning. Colaboratory notebooks are based on Jupyter and run as objects in Google Docs. The notebooks can be saved to the user's Google Drive or imported from GitHub; users can share Colaboratory notebooks such as Google Docs or Google Spreadsheets. The runtime stops after a particular time and all user data and settings are lost [30]. However, they can save the notebooks and transfer the files to the user's Google Drive.

The authors of [31] summarized the advantages and disadvantages of Google Colaboratory as follows. The advantages include fast computation; training a CNN is faster with Colaboratory's accelerated runtime than with 20 physical cores on a Linux server. Meanwhile, the disadvantages include the lack of CPU cores.

2.4. Data

2.4.1. Training and Test Images

It is essential to select training data similar to the test data to develop suitable model accuracy. Similarity conditions include the time of year (elapsed time since the disaster), weather, and scale of the disaster. I selected aerial photographs considering this policy.

The aerial photographs used for the training images are of Yamada Town, Iwate Prefecture; Miyako City, Iwate Prefecture; Minamisanriku Town, Miyagi Prefecture; and Watari Town, Miyagi Prefecture. The test images are of Rikuzentakata City, Iwate Prefecture and Kesenuma City, Miyagi Prefecture. The tsunami caused by the Great East Japan Earthquake in 2011 damaged these cities. Figure 5 shows the locations of the training and test images, and Table 1 shows the number of fatalities, etc., and damage to residential properties in the Great East Japan Earthquake.



Figure 5. Locations of training and test images.

Table 1. Number of fatalities, etc., and damage to residential properties in the Great East Japan Earthquake (as of 1 September 2014) [32].

Category	Municipality	Human Damage			Housing Damage		Non-Housing Damage	
		Death	Missing	Injured	Completely Destroyed	Partially Destroyed	Public Building	Others
Training data	Miyako City	473	94	33	2767	1331	70	
	Yamada Town	683	148	unknown	2762	405	65	unknown
	Minamisanriku Town	620	216	unknown	3143	178	14	220
Test data	Rikuzentakata City	1599	207	unknown	3805	240	61	unknown
	Keserumama City	1198	230		8483	2571	9605	

Source: Overview of the 2019 White Paper on Fire Service, Fire and Disaster Management Agency.

The photographs were obtained from Google Earth [33]; they were captured a few days after the earthquake. They were collected by selecting the area to include inundated/non-inundated areas and divided into 100 m image units for the proposed method. Then, as shown in Figure 6, each 100 m mesh image was classified into “inundated image” and “non-inundated image” by referring to the inundation estimation map published by the Geospatial Information Authority of Japan. I made this classification manually on a 100 m mesh, which can be challenging to determine. For example, if 20% of the mesh contains inundation, it is comprehensively classified as inundated/not inundated based on photographs and the inundation estimation map. It is a limitation of the accuracy verification of this study.

Table 2 shows the number of training and test images. Figures 7 and 8 show examples of inundated/non-inundated images in 100 m units for the proposed method. The traditional method develops a model after combining the images of the same area and predicts a merged test image of the same place.

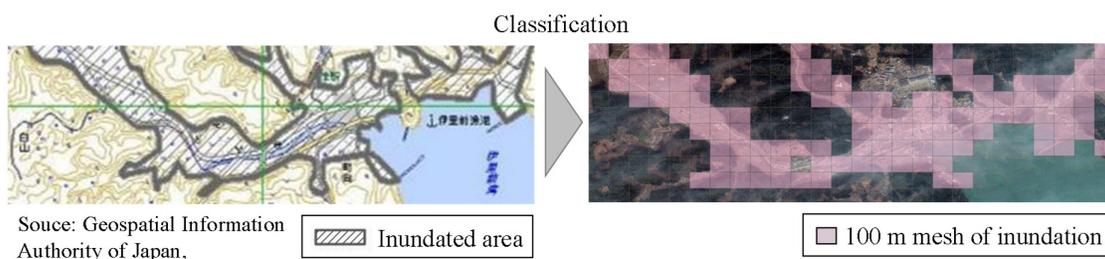


Figure 6. Classification of the tsunami inundation area.



Figure 7. Training images with inundation in the proposed method ($n = 1419$).

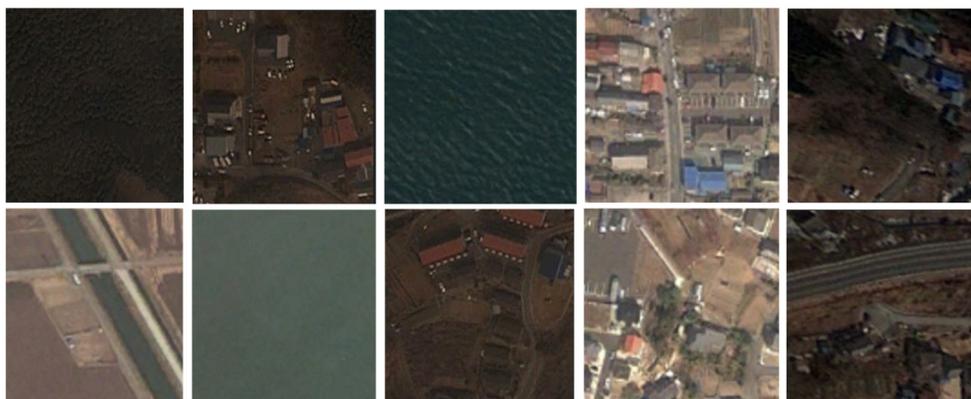


Figure 8. Training images with non-inundation in the proposed method ($n = 1689$).

Table 2. Number of images.

	Study Area	Number of Inundated Images	Number of Non-Inundated Images	Number of Total Images
Training images	Yamada Town	255 (32%)	551 (68%)	806
	Miyako City	153 (35%)	284 (65%)	437
	Minamisanriku Town	114 (39%)	176 (61%)	290
	Watari Town	897 (57%)	678 (43%)	1575
Test images	Rikuzentakata City	958 (34%)	1870 (66%)	2828
	Kesenuma City	494 (27%)	1366 (73%)	1860

2.4.2. Road Segment

The road segment data for the test images are from the Conservation GIS-consortium Japan [34]. The road data were constructed based on the situation as of 2006 (before the earthquake). The road segment data ranges from small roads to arterial roads. The road classification of these data consists of national roads, prefectural roads, municipal roads, expressway national highways, etc.

2.5. Evaluation Indicators

In the traditional method, when the training model is input to a test image, the locations with a certain probability of being inundated are marked. In this study, I set the probability to 0.5.

Moreover, in the proposed method, when the training model is input to a set of test images, the probabilities of “inundation” and “non-inundation” are output for each image. After comparing the probabilities, the decision result chooses the image with the higher value. For instance, if an image has a probability of 0.5 for “inundation” and 0.6 for “non-inundation,” the decision for this image is “non-inundation.” The probability here is the predicted probability of the target defined as intersection over union (IoU) [35], a standard indicator in target detection. The primary function is to determine positive and negative samples and evaluate the distance between the output box and the correct label.

In this study, I evaluate models using the following indicators: Precision, Recall, Specificity, and F1-score [36]. Precision is the True Positive (TP) divided by the detected objects, calculated in Equation (1). TP and True Negative (TN) are the indicators of correctly detected objects by a model and correctly missed objects by the model, respectively. False Positive (FP) and False Negative (FN) are the number of wrongly detected objects by the model and the number of wrongly missed objects by the model, respectively.

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

Recall is a ratio of correctly detected objects retrieved to the quantity of all detected objects.

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

Specificity is the percentage of true negatives correctly classified by a model.

$$Specificity = \frac{TN}{TN + FP} \quad (3)$$

F1-score is a measure of a model's overall accuracy considering Precision and Recall. It is the harmonic mean of Precision and Recall, which have contrasting characteristics.

$$F1\ Score = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall} \quad (4)$$

In this study, the above decisions (TP, TN, FP, FN) are made on the units of 100 m segments for both the traditional and proposed methods. Figure 9 shows an example of distinguishment as taking a location with 25 100 m image units. The mesh numbers of the inundation are 9, 11, and 13–24 (second from the left in the figure). The mesh numbers predicted by the traditional method are 15, 18, 19, 20, 23, and 24 (second from the right in the figure). The red boxes indicate the inundation zones, and the 100 m mesh overlapping these zones is the inundation mesh. The inundation meshes predicted by the proposed method are 2, 5, 7, 9–12, and 14–24 (first from the right in the figure). The discriminant results of the traditional method are TP = 6, TN = 11, FP = 0, and FN = 8. The discriminant results of the proposed method are TP = 13, TN = 16, FP = 5, and FN = 1.

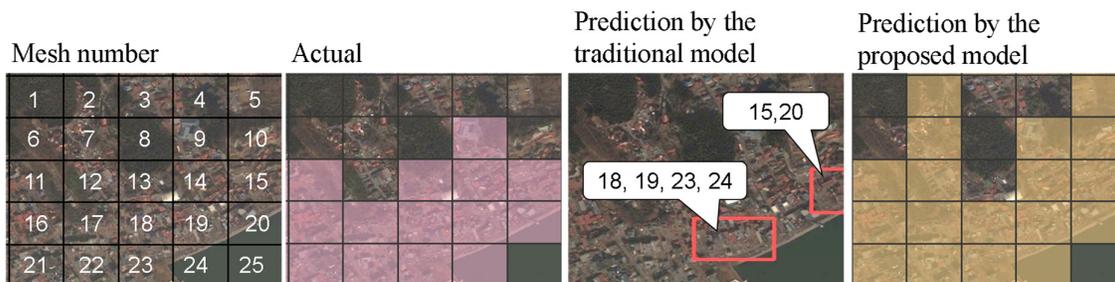


Figure 9. Example of distinguishment.

3. Results and Discussion

3.1. Training Result

To accurately ascertain the models' accuracies, I evaluated the models based on the loss function curve (train/box_loss) and average accuracy value (metrics/mAP_0.5) [37]. In the learning process, the loss function curve can intuitively reflect whether the network model converges stably with respect to the number of iterations.

The upper graph of Figure 10 shows the specific changes in the models' loss functions. The horizontal axis is the number of learning epochs, 1000 for the traditional models and 200 for the proposed models. The number of training epochs differs for each model to account for the computational time required for training. As described later, the proposed model structure took a long time to train, and Google Colaboratory timed out in the middle of the training. The figure shows that as the number of training cycles increases, the curves for both model structures gradually converge and the loss values decrease. The loss values of the proposed models are significantly smaller than those of the traditional models, proving the high accuracy of the proposed method.

The mAP measures the quality of a defect detection model. The higher the mAP value, the higher the average detection accuracy and the better the performance. The

lower graph of Figure 10 shows the training epoch trend with respect to mAP for all models; the mAP increases with the number of epochs.

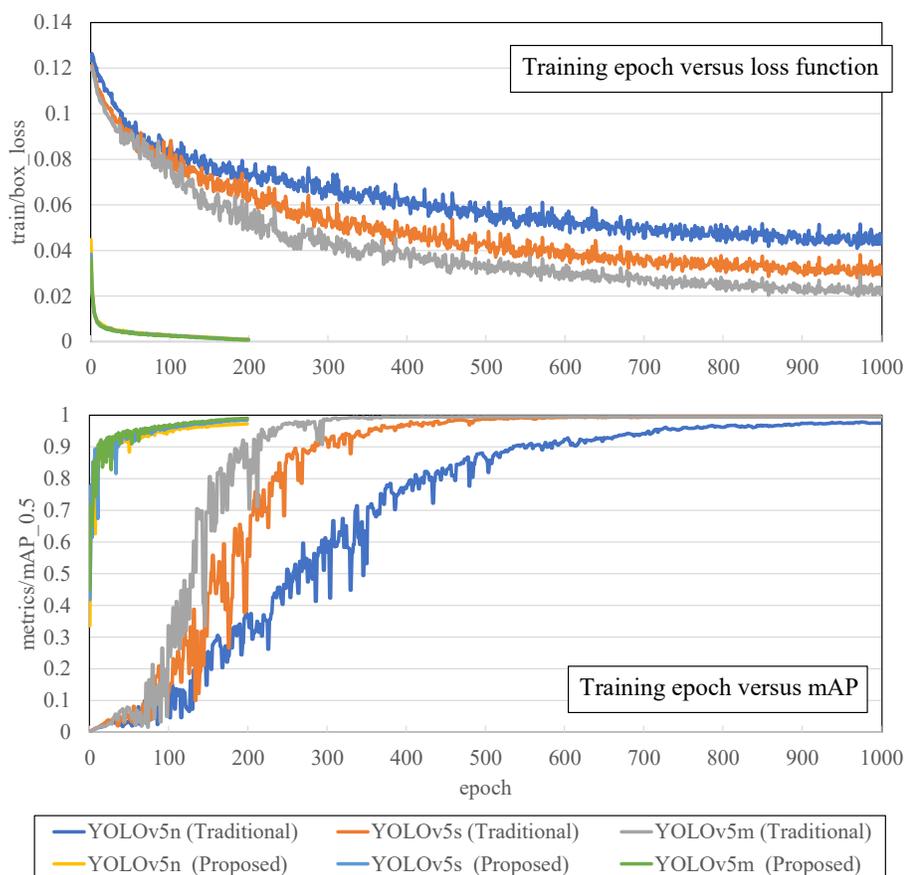


Figure 10. Training epoch versus loss function and versus mAP.

3.2. Comparative Analysis of Models

Model accuracy is verified by implementing the training models on the test images. Table 3 shows the accuracy results using the three types of YOLO training models. Rikuzentakata City is more accurate for both methods than Kesennuma City. For the traditional method, the highest F1-score for Rikuzentakata City is 78% (for YOLOv5s), whereas the highest for Kesennuma City is 60% (for YOLOv5s). For the proposed method, the highest F1-score for Rikuzentakata City is 83% (for YOLOv5m), and the highest F1-score for Kesennuma City is 72% (for YOLOv5s). The accuracy of the proposed method is better than that of the traditional method. Focusing on F1-score, the traditional models have values in the range of 59–78%, whereas the proposed models have values in the range of 66–83%, indicating an accuracy improvement.

However, the time to build a model is longer for the proposed method than for the traditional method. As stated earlier, I developed three training models each for the traditional and proposed methods. The model with the longest calculation time for the traditional method was YOLOv5m at 0.633 h. Meanwhile, the model with the longest calculation time for the proposed model was YOLOv5m at 8.800 h.

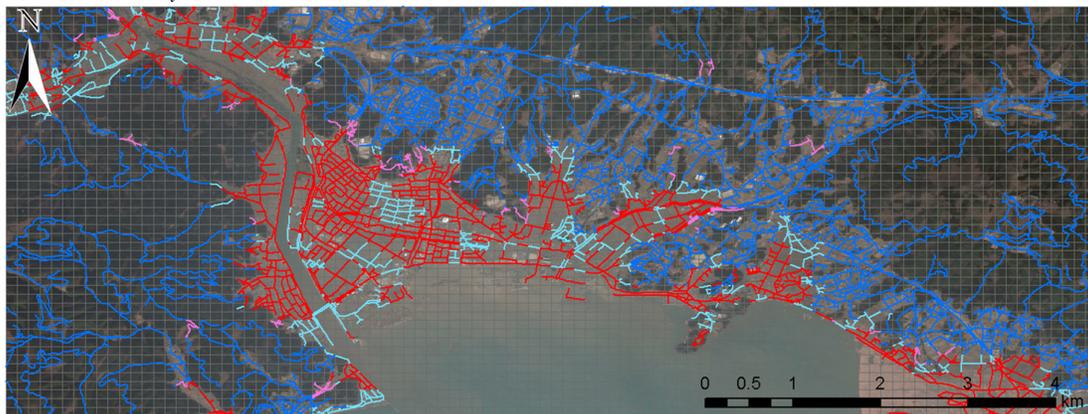
Figures 11 and 12 depict the results for Rikuzentakata City and Kesennuma City using the traditional and proposed methods (YOLOv5s), respectively. Figure 11 shows that the traditional method has good detection accuracy for no-damaged areas, such as inland areas (TN), but does not detect the damaged areas in coastal locations correctly. In particular, the FN in Kesennuma City stands out. Figure 13 shows that it can detect damaged

coastal areas with high accuracy. However, many wrong detections (FP) in the inland regions exist. The coastal areas of Rikuzentakata City and Kesennuma City were devastated, and the tsunami ran up rivers and caused extensive damage. The proposed method designated these locations as TPs. Figure 13 shows an enlarged view of a location in Kesennuma City, where FN is particularly abundant. The red boxes in the figure indicate the areas extracted as disaster-stricken areas by the traditional method. It is clear from the figure that the traditional method cannot detect most of the places as disaster-stricken areas where houses were not completely damaged.

Table 3. Comparative analysis of the developed YOLO models.

Method	Study Area	Model	TP	TN	FP	FN	Precision	Recall	Specificity	F1 Score	Training Time (h)
Traditional method	Rikuzentakata City	Yolov5n	517	1653	217	441	70%	54%	88%	61%	0.300
		Yolov5s	662	1795	75	296	90%	69%	96%	78%	0.350
		Yolov5m	429	1821	49	529	90%	45%	97%	60%	0.633
	Kesennuma City	Yolov5n	270	1214	152	224	64%	55%	89%	59%	0.300
		Yolov5s	237	1304	62	257	79%	48%	95%	60%	0.350
		Yolov5m	236	1299	67	258	78%	48%	95%	59%	0.633
Proposed method	Rikuzentakata City	Yolov5n	817	1603	267	141	75%	85%	86%	80%	2.867
		Yolov5s	856	1548	322	102	73%	89%	83%	80%	3.867
		Yolov5m	809	1681	189	149	81%	84%	90%	83%	8.800
	Kesennuma City	Yolov5n	441	966	400	53	52%	89%	71%	66%	2.867
		Yolov5s	440	1070	296	54	60%	89%	78%	72%	3.867
		Yolov5m	456	928	438	38	51%	92%	68%	66%	8.800

Rikuzentakata city



Model: Yolo v5s

Study Area	TP	TN	FP	FN
Rikuzentakata City	662	1795	75	296
Kesennuma City	237	1304	62	257

TP (True Positive) TN (True Negative)
 FP (False Positive) FN (False Negative)

Kesennuma city

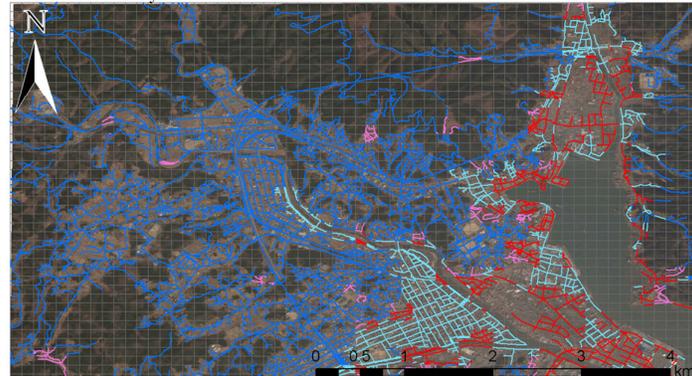


Figure 11. Visualization of the results of the traditional method (YOLOv5s).

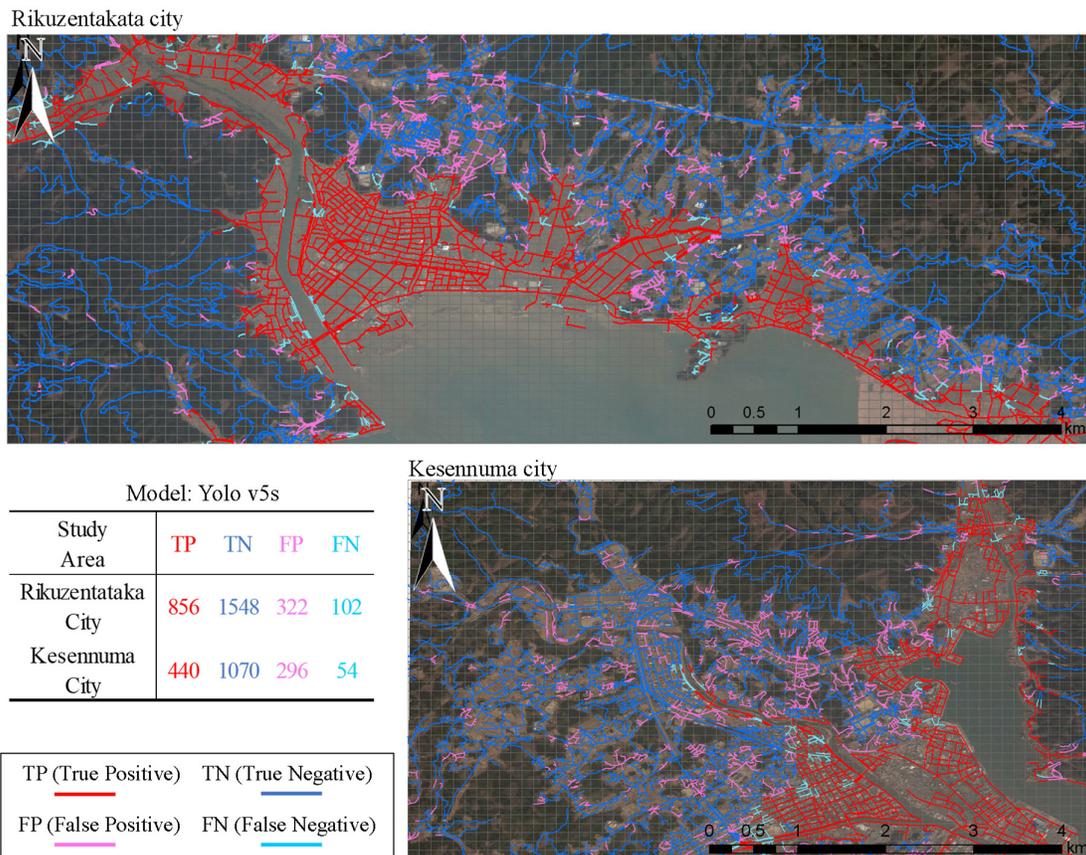


Figure 12. Visualization of the results of the proposed method (YOLOv5s).



Figure 13. Example of the detection results of the traditional method (YOLOv5s).

3.3. Accuracy Verification Focusing on the Number of Samples

When applying the proposed method in practice, it is necessary to understand how much training data is needed to accurately identify the disaster situation.

Table 4 shows the calculation results for validating the proposed method (YOLOv5s) by the percentage of samples in the training images. The 100% sample rate in the table means the total of the 3108 training images mentioned above, and the 75%, 50%, and 25% sample rates mean the corresponding extractions, chosen at random, from the total number of the training images.

A common feature of both districts is that Precision, Specificity, and F1-score improved as the number of samples increased. Precision, Specificity, and F1-score were the highest in both areas at 100%.

Table 4. Relationship between the number of samples and accuracy.

Study Area	Percentage of Sample	TP	TN	FP	FN	Precision	Recall	Specificity	F1 Score	Training Time (h)
Rikuzentakata City	25%	934	424	1446	24	39%	97%	23%	56%	0.883
	50%	898	1173	697	60	56%	94%	63%	70%	1.733
	75%	898	1429	441	60	67%	94%	76%	78%	3.467
	100%	856	1548	322	102	73%	89%	83%	80%	3.867
Kesenuma City	25%	467	257	1109	27	30%	95%	19%	45%	0.883
	50%	486	623	743	8	40%	98%	46%	56%	1.733
	75%	450	950	416	44	52%	91%	70%	66%	3.467
	100%	440	1070	296	54	60%	89%	78%	72%	3.867

4. Conclusions

DL is a novel technique to assess damaged situations quickly. In the case of complex damage situations such as tsunamis, it takes work to develop learning models. In this study, I used YOLOv5 to develop a learning model based on data from subdividing images and classifying them into tsunami-affected and tsunami-affected areas instead of labeling tsunami-affected regions from a single image. The proposed method can quickly identify damaged areas after a tsunami disaster. Once analyzers prepare the training model and road sections in units of 100 m mesh in advance, all that is required is to upload aerial photographs to Google Colaboratory for identification. In this study, 7092 aerial pictures were uploaded, including those without road segments, and the process took only a few seconds. In addition, it took approximately three minutes to classify inundation/non-inundation using the training model.

In addition, the proposed method could automatically identify the damaged areas more accurately than the traditional method. Therefore, if a road administrator develops road sections per mesh and a learning model in preparation for disasters, it will be possible to detect which road sections are damaged simply by applying aerial photographs taken after the occurrence of a disaster.

Nevertheless, this study has limitations in terms of aerial photographs and detection accuracy. Regarding the selection of aerial photographs, I set training images similar to the test images in this study, but clouds and brightness might have reduced the accuracy. It is necessary to examine the improvement in accuracy by eliminating such factors.

Regarding the improvement of the accuracy of the detection results, the proposed method determines the damaged section by mesh unit. Thus, it cannot be considered separately if the same mesh contains different types of roads: high-standard arterial highways and local roads. It is necessary to enhance the distinguishing ability by adding elevation data with each section, thereby improving accuracy.

Funding: This research was supported by JDC Foundation and JSPS KAKENHI; grant number 22K04580.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the author.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Disaster Management, Cabinet Office, White Paper on Disaster Management. 2011. Available online: <https://www.bousai.go.jp/kaigirep/hakusho/h23/bousai2011/html/honbun/index.htm> (accessed on 6 January 2023).
2. Wang, Y.; Tsushima, H.; Satake, K.; Navarrete, P. Review on recent progress in near-field tsunami forecasting using offshore tsunami measurements: Source inversion and data assimilation. *Pure Appl. Geophys.* **2021**, *178*, 5109–5128. <https://doi.org/10.1007/s00024-021-02910-z>.
3. Wang, Y.; Imai, K.; Kusumoto, S.; Takahashi, N. Tsunami early warning of the Hunga volcanic eruption using an ocean floor observation network off the Japanese Islands. *Seismol. Res. Lett.* **2022**. <https://doi.org/10.1785/0220220098>.
4. White paper on Ministry of Land, Infrastructure, Transport and Tourism in Japan. Available online: <https://www.mlit.go.jp/hakusyo/mlit/r01/hakusho/r02/html/n1222000.html> (accessed on 6 January 2023).
5. Eguchi, R.T.; Huyck, C.K.; Ghosh, S.; Adams, B.J.; McMillan, A. Utilizing new technologies in managing hazards and disasters. *Geospat. Tech. Urban Hazard Disaster Anal.* **2009**, 295–323. https://doi.org/10.1007/978-90-481-2238-7_15
6. Nex, F.; Duarte, D.; Steenbeek, A.; Kerle, N. Towards real-time building damage mapping with low-cost UAV solutions. *Remote Sens.* **2019**, *11*, 287. <https://doi.org/10.3390/rs11030287>.
7. Raffini, F.; Bertorelle, G.; Biello, R.; D’Urso, G.; Russo, D.; Bosso, L. From nucleotides to satellite imagery: Approaches to identify and manage the invasive pathogen *Xylella fastidiosa* and its insect vectors in Europe. *Sustainability* **2020**, *12*, 4508. <https://doi.org/10.3390/su12114508>.
8. Gong, L.; Li, Q.; Zhang, J. Earthquake building damage detection with object-oriented change detection. *IEEE Int. Geosci. Remote Sens. Symp-IGARSS* **2013**, 2013, 3674–3677. <https://doi.org/10.1109/IGARSS.2013.6723627>.
9. Dong, Z.; Zhang, M.; Li, L.; Liu, Q.; Wen, Q.; Wang, W.; Ji, W. A multiscale building detection method based on boundary preservation for remote sensing images: Taking the Yangbi M6. 4 earthquake as an example. *Nat Hazards Res.* **2022**, *2*, 121–131. <https://doi.org/10.1007/s11069-022-05320-z>.
10. Ge, P.; Gokon, H.; Meguro, K. Building damage assessment using intensity SAR data with different incidence angles and long-time interval. *J. Disaster Res.* **2019**, *14*, 456–465. <https://doi.org/10.20965/jdr.2019.p0456>.
11. Hu, Q.; Zhou, Y.; Wang, S.; Wang, F.; Wang, H. Improving the accuracy of landslide detection in “off-site” area by machine learning model portability comparison: A case study of Jiuzhaigou earthquake, China. *Remote Sens.* **2019**, *11*, 2530. <https://doi.org/10.3390/rs11212530>.
12. Weicong, D.; Longxu, J.; Guoning, L.; Zhiqiang, Z. Real-time airplane detection algorithm in remote-sensing images based on improved YOLOv3. *Opto. Electron. Eng.* **2018**, *45*, 180350–180351.
13. Demir, I.; Koperski, K.; Lindenbaum, D.; Pang, G.; Huang, J.; Basu, S.; Hughes, F.; Tuia, D.; Raskar, R. Deepglobe 2018: A challenge to parse the earth through satellite images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Salt Lake City, UT, USA, 18–22 June 2018; pp. 172–17209. <https://doi.org/10.1109/CVPRW.2018.00031>.
14. Miura, H.; Aridome, T.; Matsuoka, M. Deep learning-based identification of collapsed, non-collapsed and blue tarp-covered buildings from post-disaster aerial images. *Remote Sens.* **2020**, *12*, 1924. <https://doi.org/10.3390/rs12121924>.
15. Ma, H.; Liu, Y.; Ren, Y.; Wang, D.; Yu, L.; Yu, J. Improved CNN classification method for groups of buildings damaged by earthquake, based on high resolution remote sensing images. *Remote Sens.* **2020**, *12*, 260. <https://doi.org/10.3390/rs12020260>.
16. Tilon, S.; Nex, F.; Kerle, N.; Vosselman, G. Post-disaster building damage detection from earth observation imagery using unsupervised and transferable anomaly detecting generative adversarial networks. *Remote Sens.* **2020**, *12*, 4193. <https://doi.org/10.3390/rs12244193>.
17. Pi, Y.; Nath, N.D.; Behzadan, A.H. Convolutional neural networks for object detection in aerial imagery for disaster response and recovery. *Adv. Eng. Inform.* **2020**, *43*, 101009. <https://doi.org/10.1016/j.aei.2019.101009>.
18. Liu, C.; Sui, H.; Wang, J.; Ni, Z.; Ge, L. Real-time ground-level building damage detection based on lightweight and accurate YOLOv5 using terrestrial images. *Remote Sens.* **2022**, *14*, 2763. <https://doi.org/10.3390/rs14122763>.
19. Nex, F.; Duarte, D.; Tonolo, F.G.; Kerle, N. Structural building damage detection with deep learning: Assessment of a state-of-the-art CNN in operational conditions. *Remote Sens.* **2019**, *11*, 2765. <https://doi.org/10.3390/rs11232765>.
20. Geospatial Information Authority of Japan, Tsunami Inundation Map (Scale 1:25,000). Available online: <https://www.gsi.go.jp/kikaku/kikaku40014.html> (accessed on 6 January 2023).
21. Pratheesh Shivaprasad, A. Comprehensive Guide to Object Detection Using YOLO Framework? Part I. Available online: <https://towardsdatascience.com/object-detection-part1-4dbe5147ad0a> (accessed on 6 January 2023).
22. Google Colab. Google. Available online: <https://colab.research.google.com/notebooks/welcome.ipynb?hl=ja> (accessed on 6 January 2023).
23. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 779–788. <https://doi.org/10.1109/CVPR.2016.91>.
24. YOLOv5 Download Site. Available online: <https://github.com/ultralytics/yolov5> (accessed on 6 January 2023).
25. What Is YOLOv5? A Guide for Beginners. Available online: <https://blog.roboflow.com/yolov5-improvements-and-evaluation/> (accessed on 29 January 2023).
26. Ge, Z.; Liu, S.; Wang, F.; Li, Z.; Sun, J. Yolox: Exceeding Yolo Series in 2021, Computer Vision and Pattern Recognition. *arXiv* **2021**, arXiv:2107.08430.

27. SeekFire Overview of Model Structure about YOLOv5. Available online: <https://github.com/ultralytics/yolov5/issues/280> (accessed on 6 January 2023).
28. Hu, D.; Zhang, Y.; Xufeng, L.; Zhang, X. Detection of material on a tray in automatic assembly line based on convolutional neural network. *IET Image. Process.* **2021**, *15*, 3400–3409. <https://doi.org/10.1049/ipr2.12302>.
29. Jocher, G. Train Custom Data. Available online: <https://github.com/ultralytics/yolov5/wiki/Train-Custom-Data> (accessed on 6 January 2023).
30. Google Colaboratory: Frequently Asked Questions. Google. Available online: <https://research.google.com/colaboratory/faq.html> (accessed on 6 January 2023).
31. Carneiro, T.; Medeiros Da Nobrega, R.V.; Nepomuceno, T.; Bian, G.; De Albuquerque, V.H.C.; Filho, P.P.R. Performance analysis of Google Colaboratory as a tool for accelerating deep learning applications. *IEEE Access* **2018**, *6*, 61677–61685. <https://doi.org/10.1109/ACCESS.2018.2874767>.
32. Overview of the 2019 White Paper on Fire Service, Fire and Disaster Management Agency. Available online: <https://www.fdma.go.jp/publication/hakusho/h26/cat/740.html> (accessed on 6 January 2023).
33. Google Earth Download Site. Available online: <https://www.google.co.jp/intl/ja/earth/> (accessed on 6 January 2023).
34. Conservation GIS-Consortium Japan. Available online: <http://cgisj.jp> (accessed on 6 January 2023).
35. Jiang, B.; Luo, R.; Mao, J.; Xiao, T.; Jiang, Y. Acquisition of localization confidence for accurate object detection. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 816–832. https://doi.org/10.1007/978-3-030-01264-9_48.
36. Karaci, A. VGGCOV19-NET: Automatic detection of COVID-19 cases from X-ray images using modified VGG19 CNN architecture and YOLO algorithm. *Neural Comput. Appl.* **2022**, *34*, 8253–8274. <https://doi.org/10.1007/s00521-022-06918-x>.
37. Li, Z.; Namiki, A.; Suzuki, S.; Wang, Q.; Zhang, T.; Wang, W. Application of low-altitude UAV remote sensing image object detection based on improved YOLOv5. *Appl. Sci.* **2022**, *12*, 8314. <https://doi.org/10.3390/app12168314>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.