IDRiM (2022) 12 (1) **ISSN: 2185-8322**
DOI10.5595/001c.35885

Journal of Integrated Disaster Risk Management

Original paper

Rapid and Accurate Detection of Building Damage Investigation Using Automatic Method to Calculate Roof Damage Rate

Shono Fujita[1*] and Michinori Hatayama[2]

Received: 26/11/2021 / Accepted: 30/03/2022 / Published online: 24/05/2022

Abstract In the event of a natural disaster, local Japanese governments investigate the level of damage of the buildings and issue damage certificates to the victims. The damage certificate is used to determine the content of the support provided to the victims; hence, they must be issued rapidly and accurately. However, in the past, the investigation of damage was time-consuming, thus delaying the support provided to the victims. Additionally, while investigating the roof of the damaged building, it was difficult for the investigators to look at the entire roof and calculate the damage rate accurately. Therefore, we have developed an automatic method to calculate the damage rate of a roof using image recognition from aerial photos so that building damage investigation can be more accurate and rapid. We requested the staff in the disaster management division to evaluate the estimation results of this model and confirm its effectiveness. As a result, 80 % of the roof data obtained from this method was equal to or more accurate than the investigator checking from the ground. Additionally, we have developed an efficient flow for building damage investigation using the proposed system. In the future, we aim to investigate the system usage to ensure responsibility in data estimation.

Keywords: Deep learning, roof damage, damage certificate, image recognition

[1] Graduate School of Informatics, Kyoto University
* Corresponding author email: fujita.shono.32x@st.kyoto-u.ac.jp
[2] Disaster Prevention Research Institute, Kyoto University

1. INTRODUCTION

In the event of natural disasters, such as earthquakes, storms, and floods, the Japanese local government investigated the level of damage to each building and issued a damage certificate to the victims to prove that the buildings were damaged by the disaster. As this certificate is used to determine the content of their support, such as temporary housing, support money, and loans, it is necessary to reconstruct their livelihoods. Moreover, until they receive the actual certificate, they are unable to make plans to reconstruct their livelihoods, especially regarding their residence. Hence, they must be issued accurately and rapidly (Disaster Management, Cabinet Office, Government of Japan 2020). After the earthquake in Great East Japan, the Basic Act on Disaster Control Measures was revised, and local governments were instructed to issue damage certificates without delay (Disaster Management, Cabinet Office, Government of Japan 2020). However, in the past, the process of building damage investigation and issuing damage certificates was time-consuming and delayed the support provided to the victims.

The building damage investigation is conducted by the staff of the local government and has three stages: first, second, and reinvestigation. The results of this investigation require the agreement of the victims. If the victims are not convinced of the results of the first investigation, they can apply for a second investigation or a reinvestigation. The first investigation is comparatively simple because it is a visual inspection. The second investigation and reinvestigation are detailed. At each stage of the investigation, the investigators check the appearance, inclination of the building, or damage degree of each part, such as walls, foundations, and roofs, to calculate the level of damage to the entire building. An overall damage rate of more than 50 % is assigned to completely destroyed structures, 40–50 % to large-scale half destroyed structures, 30–40 % to middle-scale half destroyed structures, 20–30 % to half destroyed structures, 10–20 % corresponds to semi half destruction, and less than 10 % to partially damaged structures (Disaster Management, Cabinet Office, Government of Japan 2021).

However, in the roof damage investigation, the investigators cannot investigate the entire roof. Therefore, they look at the roof from a distance and investigate it within the range of vision from the ground. In addition, they calculate the degree of damage to each roof surface and need advanced expert knowledge. These inaccurate investigations result in victims' dissatisfaction with the results, and the number of second investigations or reinvestigations increases. In past earthquakes, investigations classified as unsatisfactory to resident caused much trouble between local governments and victims (Shigekawa *et al.* 2005).

After the Kumamoto earthquake in 2016, 135,959 first investigations, 37,807 second investigations, and 2,635 reinvestigations were conducted in Kumamoto City (Department of Crisis Management, Kumamoto Prefecture, 2020). One of the previous studies (Inoue *et al.* 2018) reported that the work on building damage investigation in one city of Kumamoto Prefecture took approximately 29,000 man-days, which was the largest number after the work

on the issue of damage certificates. Thus, improving the efficiency of building damage investigation may enable the local governments to allocate manpower to other tasks effectively.

In previous earthquakes in Japan, many buildings, especially wooden buildings, were damaged. Based on investigations to assess damaged buildings after the 2016 Kumamoto earthquake, 8,642 buildings were found to have been completely destroyed, 34,393 buildings were half destroyed, and 155,177 buildings were partially damaged (Department of Crisis Management in Kumamoto Prefecture, 2020). In Mashiki-town, which was the epicenter of the damage, 28.2 % of wooden buildings built before 1981 collapsed (Ministry of Land, Infrastructure, Transport, and Tourism 2016). In Japan, the ratio of wooden buildings is 56.97%, and this increases to 92.5 % for single houses (E-Stat 2018, Ministry of Internal Affairs and Communications 2018). Although changes in building standards have resulted in buildings more robust to earthquakes, many buildings may still be damaged. Traditional Japanese tile roofs tend to be damaged by earthquakes or typhoons. Many Japanese people select this tile roof for reasons such as their design, durability, and heat resistance. Prior to the change in the building standard in 1981, tile roofs of buildings did not have to be fixed to the roof base. Currently, the number of tile roofs that sustain damage in disasters is decreasing because of revised strict building standards, the use of light raw materials, and a decrease in the number of people using tile roofs. However, many old buildings, or buildings without sufficient countermeasures, may still be damaged in the future.

2. RESEARCH PURPOSE

Based on the aforementioned challenges, the problems encountered during building damage investigations are as follows:

(1). Building damage investigation is time consuming and delays the support provided to the victim's livelihood.
(2). In the roof investigation, the investigators cannot look at the entire roof from the outside; hence, the results are not accurate.
(3). The investigators must have prior technical knowledge owing to the complexity of investigation.

Therefore, the purpose of this study is to develop a system to automatically calculate the damage rate of roofs using image recognition from aerial photos so that building damage investigation can be more rapid and accurate. Additionally, using aerial photos enable the investigators to accurately investigate the part of the roof that cannot be seen from the outside. Moreover, if image recognition can calculate the precise damage rate of the roof, prior technical knowledge is not required and people other than experts in architecture can also investigate the roof accurately.

3. PREVIOUS STUDIES

Vetrivel *et al.* (2018) developed a damage-building detection model using deep learning and 3D point cloud features. They used high-resolution oblique aerial photos as the inputs for the model. Tu *et al.* (2017) developed a detection model to identify damaged regions using a support vector machine, which is a type of machine learning method. They used multi-temporal high-resolution remote sensing images to detect changes in the buildings owing to damage. This study uses aerial photos that are obtained by drones or aircraft at a low price. This provides easy access to training and estimation data during disasters.

Ji *et al.* (2019) estimated damaged buildings using texture analysis, which analyzes the state and pattern of the object surface, and a convolutional neural network (CNN), which is a type of deep learning for image recognition, based on differences in post-disaster and pre-disaster aerial photos. In addition, Fujita *et al.* (2017) used the information on differences in post-disaster and pre-disaster aerial photos to detect damage of buildings by tsunami using a CNN. However, the pre-disaster aerial images may be old, and the building information may be significantly different from the post-disaster images. Such buildings include not only damaged buildings but also new buildings or demolished buildings. Therefore, their system may detect buildings that are not damage. This problem was addressed in this study because only post-disaster aerial photos were used. Meloy et al. (2007) investigated the roof performance of new homes in Florida damaged by Hurricane Charley using oblique aerial photos with four angles. They used a special tool to obtain the area ratio of the damaged part by manual entry, and determined the damage level from this ratio. This study calculates automatically the damage rate of a Japanese building damage investigation from both the area ratio and type of damage. Moreover, this study does not use oblique aerial photos but ortho photos, which were taken from directly overhead. In Japan, many companies and organizations employ ortho arial photographs because it is easy to overlay them on maps, and are thus considered important by disaster response organizations. Considering their availability, ortho aerial photos are effective for this system.

Inoguchi *et al.* (2019) developed a detection system for buildings with blue sheets using a CNN for deep learning from the aerial photos captured by drones. Their data were generated from images of the roof immediately after they were captured manually. However, these methods are time-consuming during disasters. In this study, data are generated using the original trimming algorithm based on the location information of the building polygon, which is geospatial information. Miura *et al.* (2020) identified collapsed buildings and blue tarp-covered buildings using deep learning from aerial images. However, the focus was on the damage to the entire building, not just the roof.

According to previous studies, the efficiency of building damage investigations can be improved using various methods. Matsuoka *et al.* (2018) judged the level of damage to the building using a CNN from aerial photos and investigation field photos for the purpose of investigation. They also focused on the damage to the entire building and did not refer to the

usage of their system in the actual investigation field during a disaster. Tanaka *et al.* (2008) analyzed the processes of building damage investigation and proposed a self-inspection system that can be used by a non-expert. Fujiu *et al.* (2012) developed a remote judgment system for building damage investigations using smartphones and the Internet to improve the efficiency of the investigation and issue of damage certificates. To the best of our knowledge, there is no existing research on automatically calculating the degree of damage of the damage building investigation of individual parts of buildings, particularly roofs from the aerial photos, considering usage during actual disaster response, such as obtaining aerial photos and generating image data.

4. AUTOMATIC METHOD TO CALCULATE DAMAGE RATE OF ROOF

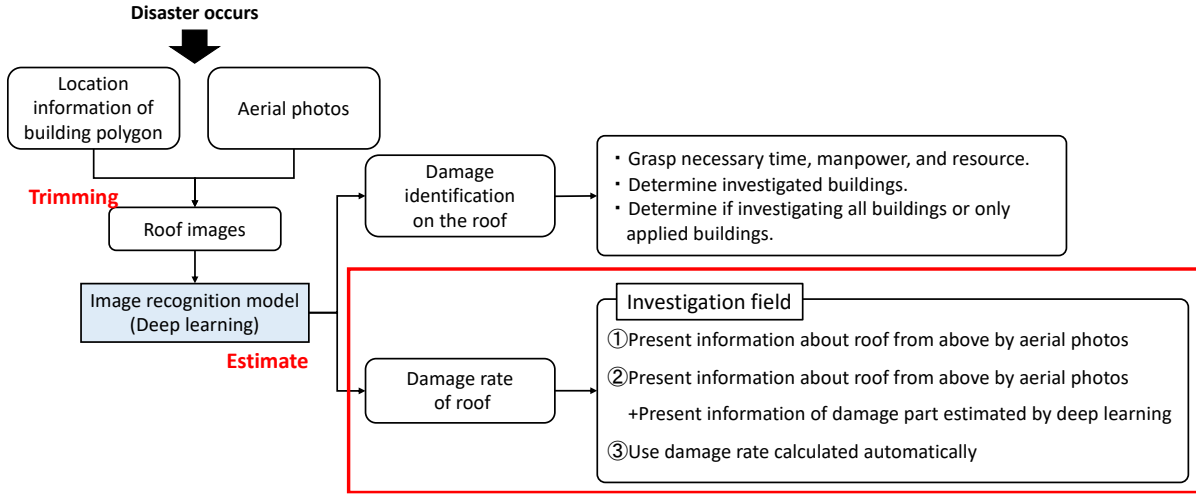

Figure 1. Overall structure of the proposed system

Table 1. Usage of the proposed system

	Use of aerial photos	Use of deep learning	Basis for judgement of investigator	Final calculation
①Present information from aerial photos	○	×	Field information +Aerial photos	Investigator
②Present information from aerial photos +Present information from deep learning	○	○	Field information +Aerial photos +deep learning	Investigator
③Use damage rate of roof by deep learning	○	○	-	Deep learning

4.1 Overall Structure of Proposed System

Figure 1 shows the overall structure of the proposed system. After a disaster occurs, the proposed system automatically generates images of the roof using a trimming algorithm, aerial

photos and geospatial information, such as the location information of building polygons. Then, each data point of the image of the roof is fed as an input to the image recognition model to estimate the damage to the roof. In our previous study (Fujita and Hatayama 2021), this image recognition model estimated the presence of roof damage to grasp the necessary time, manpower, and resources, determine the investigated buildings, and investigate all buildings or only the applied buildings. In this study, we developed an image recognition model to calculate the damage rate of the roof so that the building damage investigation can be more rapid and accurate to address the following usages:

(1) Present information about roof parts that cannot be seen from the outside using aerial photos.
(2) Present information about roof parts with high probability of damage and advise investigation to investigators.
(3) Calculate the damage rate of the roof directly using the model for building damage investigation.

Table 1 lists the details of these usages. In this table, “O” means “use” and “X” means “do not use”. In all three usages, the investigators are in the investigation field to investigate certain parts of the building other than the roof.

4.2 Trimming Algorithm

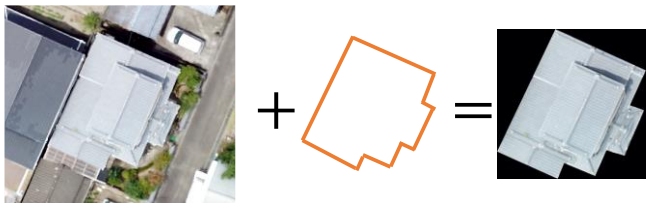

Figure 2. Trimming algorithm

The trimming algorithm can automatically generate data from the image of the roof using aerial photos and location information of the vertex of the building polygon as shown in Figure 2. Because trimming can extract each image from aerial photos, this model can reduce the time required to generate image data so that the level of damage can be estimated rapidly during a disaster. Subsequently, a large amount of training data is generated for deep learning. Moreover, this method uses building polygons; hence, the proposed model can estimate the damage not by area but by building and obtain detailed information. Moreover, painting areas of the building other than the roof with black can remove unnecessary information while calculating the damage rate of the roof, which increases accuracy.

4.3 Method to Calculate Damage Rate of Roof in Building Damage Investigation

During the investigation, the investigator calculates the damage rate of certain parts of the building, such as the wall, foundation, and roof and, determines the level of damage of the building from the total amount of damage. The damage rate of the roof is calculated by multiplying the degree of damage by the roof surface area rate. These values are then added as in equation (1), where *Si* is the area of roof *i*, S^e is the area of the entire roof, and *Di* is the damage degree of roof *i*. The damage degree is represented as a percentage based on the damage type and position. If one roof surface has a different degree of damage, then the degree of damage is calculated by the average weight of these areas (Disaster Management, Cabinet Office, Government of Japan 2020).

$$Damage\ rate = \sum_i \frac{S_i}{S^e} \times D_i \tag{1}$$

4.4 Problem of This Study

Our previous study (Fujita and Hatayama 2021) estimated the damaged roof and the roof covered with a blue sheet using deep learning from trimming roof data. Because many victims covered the damaged part of the roof with a blue sheet to prevent wind and rainwater, we identified the roof covered with a blue sheet as a damaged roof. Consequently, the accuracy of the estimation of the damaged roof was lower than that of the roof covered with a blue sheet. Based on this result, we concluded that the challenges to be addressed were the difficulty in extracting the features of the damaged part and the lower resolution of the aerial photos captured by the aircraft. It is necessary to use abundant training data to improve the accuracy of deep learning models. However, there are insufficient high-resolution aerial photos that include roofs damaged by earthquakes. The reasons for this are the low frequency of earthquakes, short time since the invention of drones that can capture aerial photos with high resolution, and time limitation, which forced us to take aerial photos of the roof covered with a blue sheet. Thus, we concluded that the problem in the estimation from aerial photos during disasters in our study was limited training data, which is necessary for the improvement of accuracy.

4.5 Proposed Method to Calculate Damage Rate of Roof

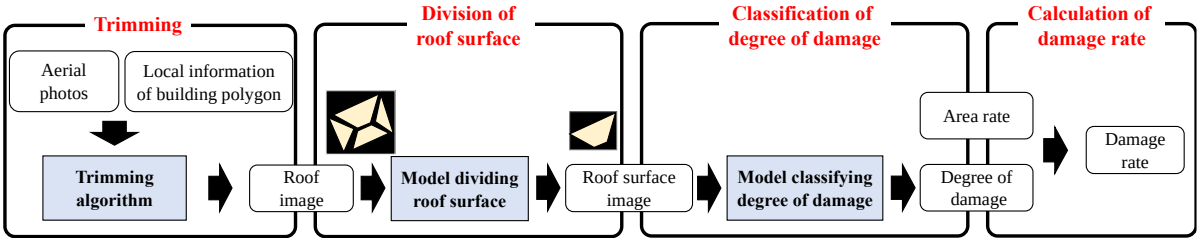

Figure 3. Flow of the proposed method to calculate damage rate of the roof

Based on the above problem, the trimming of roof images is divided into roof surfaces, thus increasing the amount of training data, as shown in Figure 3. Generating multiple roof images from one trimming roof image enables us to obtain training data several times. Then, the divided roof surface images are fed as inputs to the classification model to estimate the degree of damage. Finally, the degree of damage is multiplied by the area rate, and the values for all the roof surfaces are added to obtain the damage rate of the roof.

Ise *et al.* (2018) suggested a deep learning method using finely divided image data obtained from images as training data to classify moss and obtain a high accuracy. This shows that the division method can generate a large amount of training data from one image to improve accuracy. In terms of the building damage investigation, the proposed method follows the calculation method of the actual investigation. Therefore, the calculation result obtained in the proposed method is close to the actual result, and the skepticism of the victims regarding the accuracy of the estimated result may decrease. Moreover, this method obtains the degree of damage of each roof surface; hence, we can determine the actual damaged roof surface. Therefore, the second usage of the proposed system, which is to present information about roof parts with a high probability of damage, can be achieved. Additionally, in the third usage of our system, which is to use the damage rate of the roof directly, as calculated by the model during building damage investigation, the proposed model can indicate the basis for judging the calculation result to the investigators or victims. As mentioned above, the damage level of a building estimated during a building damage investigation determines the support content for the victims. Thus, when the victims are not convinced of the estimated level of damage or the calculated damage rate is located near the borderline position of the damage level class, many troubles with victims are likely to occur. In addition, in deep learning with high accuracy, numerous parameters act on each other. It is difficult to directly obtain an interpretation or explanation of the training or manual estimation. However, it is likely that providing a basis for judgment in addition to the high accuracy of the proposed method, will enables a smooth response without problems.

4.6 Necessary Data

To operate this system, data such as aerial photos for estimation, location information of building polygons, and training data (aerial photos and correct labels for deep learning) are necessary. Aerial photos for estimation during a disaster can be obtained for free from the Geospatial Information Authority of Japan or a non-profit organization, such as Drone Bird. The location information of building polygons can be downloaded from Fundamental Geospatial Data in Japan or Open Street Map. The approach to training data differs based on when the model is trained, whether before or after a disaster. If the model is trained before a disaster occurs, a general model that considers the features of the region, such as the roof format, must be constructed. For example, Naito *et al.* (2020) constructed a general model to estimate building damage by deep learning using aerial photos of several regions as training data. If the

model is trained after a disaster, a model that considers the features of a region can be constructed using part of the aerial photo of the estimation data as training data. However, in this case, the input of the correct answer label to the training data (annotation) must be conducted rapidly after a disaster. For example, cooperation by cloud sourcing or requesting input from outside people, such as those who are aware of the investigation process or experts is necessary.

Thus, aerial photos and location information of building polygons can be obtained from existing data, and the label for deep learning can be obtained from existing systems or staff with knowledge of the investigation. This indicates that the proposed system can be operated only by local governments, and ensures the feasibility of the system.

5. DIVISION OF ROOF SURFACE

5.1 Data Division Model

This study uses an instance segmentation model of deep learning that extracts the region of an object by classifying each pixel in an image to divide the roof surface because, a deep learning model can extract complicated features such as roof surfaces in images even with damage or dirt. In particular, this study uses Mask R-CNN (He *et al*. 2017), which is one of the most accurate instance segmentation models.

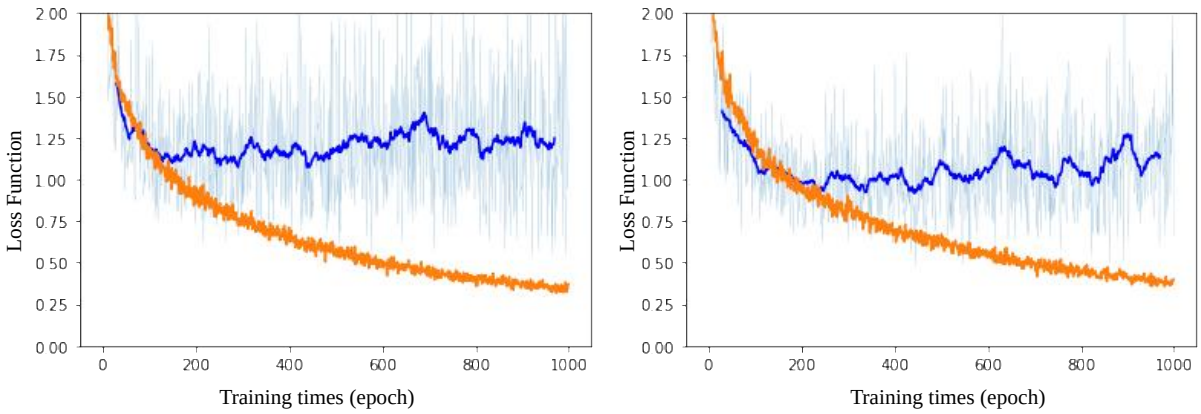

Figure 4. Left side shows transition of the loss function of the first experiment, and right side shows one of the additional experiments. The orange line represents one of the training data and blue line represents one of the validation data.

5.2 Training Method

Table 2. Result of dividing roof surface

	First experiment				Additional experiment		
	All images	Images with damage in roof surface	Images with damage in boundary	Images with leaf	All images	Images with damage in roof surface	Images with damage in boundary
Average of IoU	0.7580	0. 7341	0. 6874	0.7237	0. 7672	0. 6858	0. 7074
Average of AP	0. 6670	0. 6846	0. 5752	0. 3979	0. 6934	0. 6094	0.4628

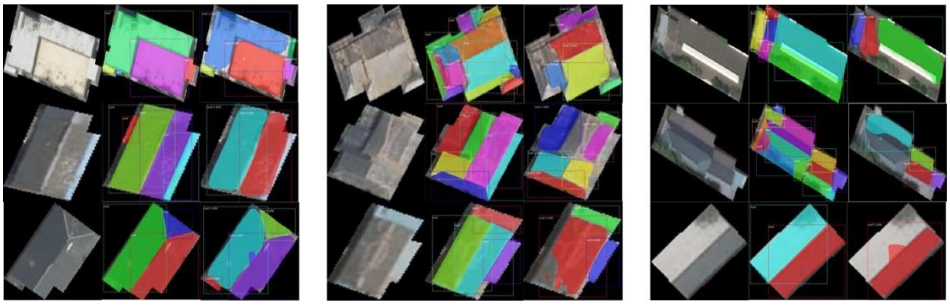

Figure 5. Left: images of the roof with damage in roof surface, center: images of the roof with damage in boundary, and right: images of the roof with leaf in them. Each row shows original image, correct region, and estimated region from left to right.

We generated images of the roof using a trimming algorithm and ortho aerial photos taken in Mashiki town of Kumamoto prefecture after the 2016 Kumamoto Earthquake. These aerial photos were taken by aircraft rather than drones, with a resolution of 20 cm and an elevation of 1396m. Although these are low-resolution photographs, they cover a very wide area in the damaged region. The location information of the building polygons was obtained from the Fundamental Geospatial Data of the Geospatial Information Authority of Japan. This annotation took approximately 1 min for one roof image and a total of 8–9 h. In this study, 2400 (300×8) images were used as training data to update the parameters, 800 (100×8) images were used as validation data to confirm overfitting, and 800 (100×8) images were used as test data to evaluate the accuracy of the model. The number of data points was increased eight times with horizontal flipping and rotation of each image. Because there is a gap between the location of the building polygon and the actual building, we excluded roof images whose gap was large or the roof surface could not be judged by observation from these datasets.

The loss function of the validation data was confirmed to prevent overfitting. When this value increased, we considered the increase to be overfitting and stopped the training. In this study, the classification classes were set to 2 (roof and background), the size of the input image to 256×256, the batch size to 2, the number of iterations in one epoch to 100, and the number of iterations in the calculation of the loss function of the validation data to 5. Tesla K40c and

GeoForce GTX 1060 6 GB of NVIDIA were used as the GPU in this experiment. The training of 1000 epochs took approximately 16 h. The left side of Figure 4 shows the transition of the loss function of the training and validation datasets. Because an increase in the loss function appeared in epoch 500, the training was stopped at epoch 500 using with the proposed model.

5.3 Result of Division in the First Experiment

The average intersection of union (IoU) by image in the first experiment was 0.7580 and the average of average precision (AP) was 0.6670, as shown in Table 2. The IoU indicates the degree of overlap between the correct and estimated regions. The AP indicates the degree of recall and precision of the regions estimated by image. As a result, images with high AP and IoU tended to have large and few regions of the roof surface in each image.

In the roof images with damaged regions, as shown in Figure 5, there is a difference between the accuracy of roof surface division between roof images with damaged regions on the roof surface (Figure 5, left) and roof images with damaged regions at the boundary of the roof surface (Figure 5, middle). Comparing the 40 image data with damage to the roof surface (five original images) with all the other images from Table 2, it was observed that the proposed model is capable of dividing the roof surface of these images as accurately as other roof images. Therefore, even if the color or texture of the roof surface is discontinuous, the model can accurately divide the roof surface. This result suggests that instance segmentation using deep learning is effective for image division. However, when the 40 images with damage at the boundary of the roof surface (five original images) were compared with all other images from Table 2, the model could not divide the roof surface of these images with the same accuracy as other roof images. Additionally, it was observed that roof images with a low average IoU and AP included roof images with leaves, as shown in Figure 5 (right).

5.4 Result of Division in the Additional Experiment

To handle images with the abovementioned features, we added 160 images with damage regions in the boundary (20 original images) and 160 images with leaves (20 original images) to train the data. The proposed model was trained using 2720 training data points, 800 validation data points, and 800 test data points. The right side of Figure 4 shows the transition of the loss function of the training and validation datasets. Because an increase in the loss function appeared at epoch 600, the training was stopped at epoch 600 with the proposed model. As a result, the average IoU of the roof images with damage regions in the boundary was lower by 0.0016, and the average AP was higher by 0.0342 compared to that in the first experiment, as shown in Table 2. The average IoU of the roof images with leaves was lower by 0.0163, and the average AP was higher by 0.0649 compared to that in the first experiment. In both roof images with damage regions at the boundary and with leaves, both IoU and AP did not increase.

However, the difference between the increase and decrease in both of these values indicates that the accuracy of the division of these images improved to some extent.

5.5 Image Processes after Division

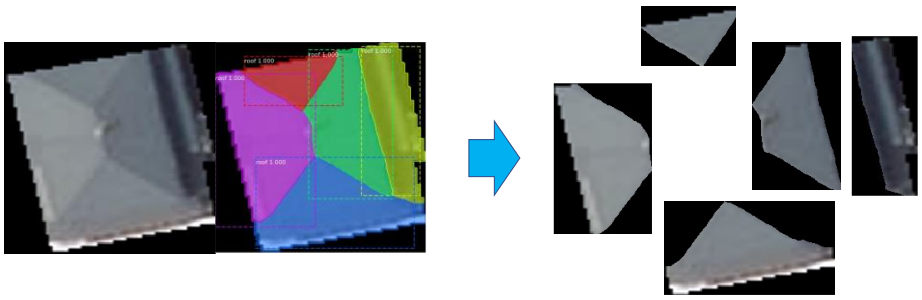

Figure 6. Example of divided roof surface data image

There were many overlapping regions with multiple estimated regions and overlooked regions that were not estimated to be roof surfaces.

In overlapping regions, many regions with large areas expanded to regions with small areas. To remove overlapping regions, we selected the smallest region of overlapping regions and excluded the others. To compensate for the overlooked regions, we expanded the estimated regions to all sides at the same speed after removing overlapping regions. Roof surface images were generated after these two processes, as shown in Figure 6.

6. CLASSIFICATION OF DEGREE OF DAMAGE

6.1 Used Data

Roof surface image data created by division of the roof surface have a low resolution because the image of the roof from which the roof surface is divided has a resolution of 20 cm. Therefore, it is difficult to determine the degree of damage of these images in detail. Thus, we have constructed a model that classifies five classes of roof surfaces: no damage, damage (-25 %), damage (25–50 %), damage (50–75 %), and damage (75%). The corresponding image data were generated as shown in Figure 7.

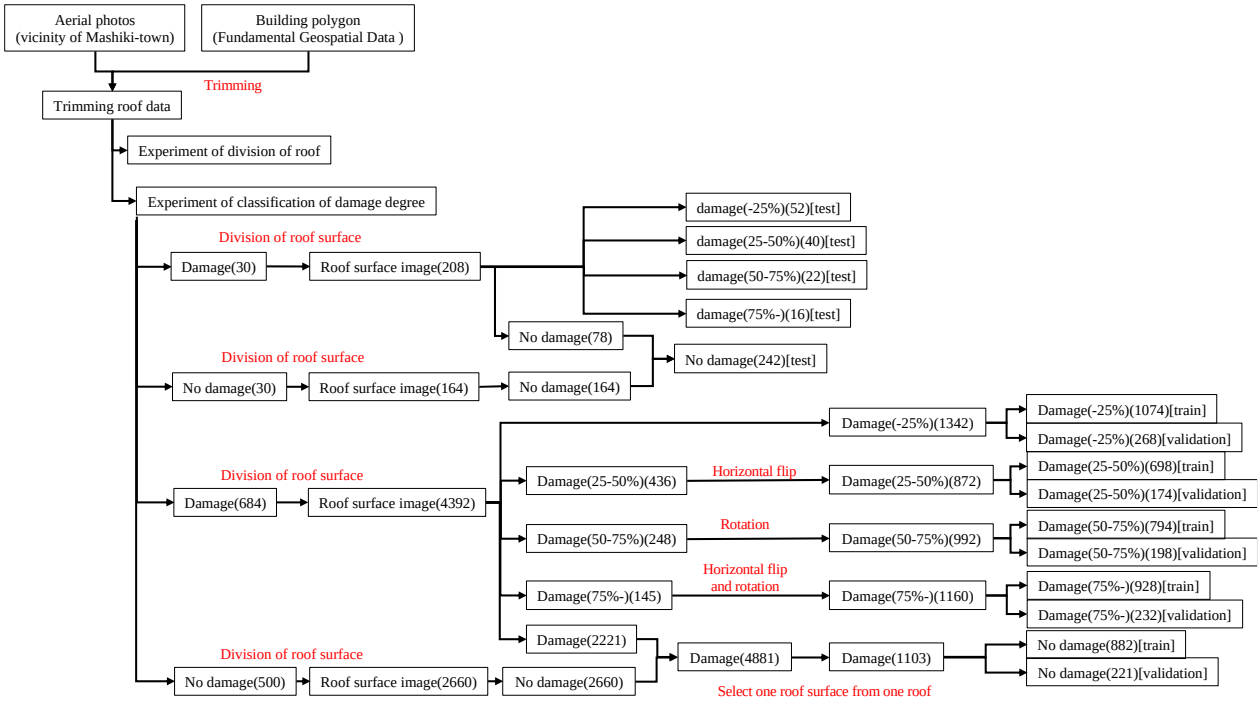

Figure 7. Breakdown of data used in classification of degree of damage

In this study, we used 30 roof images, each with and without damage, as the test data. Of the 30 roof images with damage, we selected roof images with different degrees of damage to obtain various damage types. Then, we divided the images with damage using the division model described in Section 4, and requested the staff in the Department of Crisis Management of Shimanto-town, Kochi Prefecture, who have an experience of building damage investigation, to input the correct label of damage degree to 208 divided images. In this study, 684 roof images with damage and 500 roof images without damage were used as the training and validation data. A total of 4,392 roof surface images were obtained from roof images with damage. These were fed as input labels of the degree of damage in reference to the labels of the test data and investigation manual. Consequently, this model could generate 2,171 roof surface images with damage from 684 roof images with damage, indicating that the training data could be increased by 3.174 times. In this study, the number of data points of each class in the training and validation data was equal by horizontal flipping and rotation. Additionally, in roof images without damage, we selected one roof surface from each roof to obtain various types of roof data. Figure 8 shows examples of images of damaged roof surfaces.

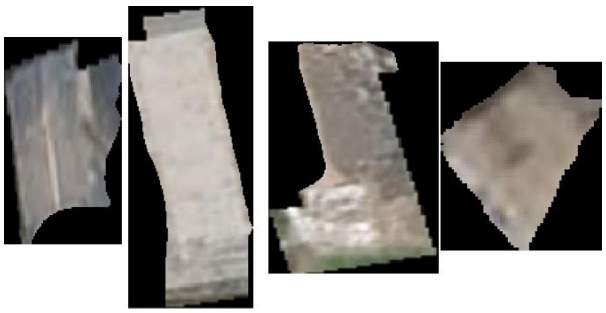

Figure 8. Examples of image of roof surf with damage (damage (-25 %), damage (25–50 %), damage (50–75 %), damage (75%) ordered from left to right)

DOI10.5595/001c.35885

6.2 Training Method

We used ResNet50 (He *et al.* 2015) to classify the degree of damage, which is one of the most accurate classification models. We trained ResNet50 using the data described in the previous subsection. The batch size was set to 16, the size of the input image to 256×256, and the loss function to cross-entropy loss. It took 5 h and 56 m to train this model using the GPU GeoForce GTX 1060 6 GB of NVIDIA for 200 epochs. Figure 9 shows the transition between the loss function and the accuracy. 1 epoch represents 274 times of training. This represents the total number of times training was provided to all data simultaneously. Figure 9 indicates that the training was stopped at epoch 500 when the accuracy was high and the loss function increased.

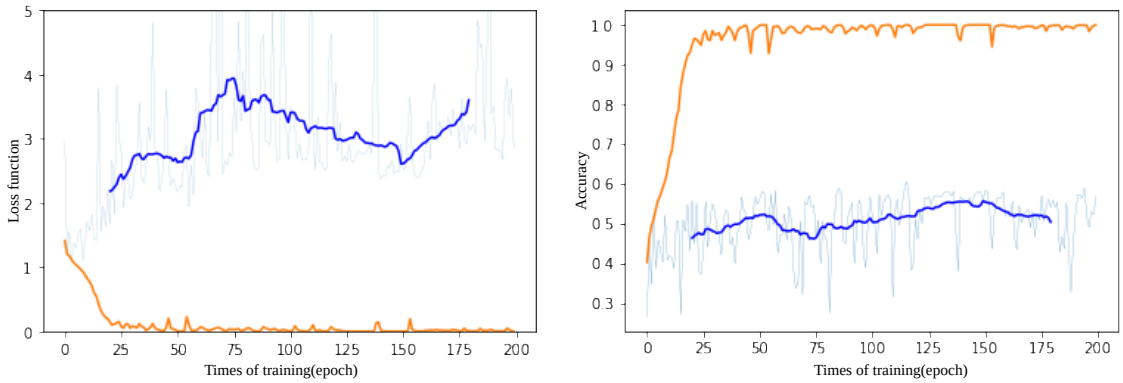

Figure 9. Left: loss function and right: accuracy. Orange line: transition of training data and blue line: transition of validation data (moving filter)

6.3 Classification Result

		True label					
		No damage	Damage (-25%)	Damage (25-50%)	Damage (50-75%)	Damage (75-%)	
Estimated label	No damage	181	15	9	4	0	
	Damage(-25%)	42	29	19	6	6	
	Damage(25-50%)	8	5	5	5	5	
	Damage(50-75%)	6	1	3	3	1	
	Damage(75%-)	5	2	4	4	4	Average Recall
	Recall	0.7479	0.5577	0.1250	0.1364	0.2500	0.3634

Table 3. Confusion matrix of classification of damage degree

Table 3 presents the confusion matrix of the estimation results of the test data. The recall of image data with damage, particularly for damage (25–50 %), damage (50–75 %), and damage (75 %), was lower than image data without damage. Additionally, degree of damage of image data with large damage, such as damage (25–50 %), damage (50–75 %), and damage (75%) were underestimated.

In these underestimated image data, there were many roof surfaces that reflected little damage in the image, but collapsed overall or reflected only the damaged part. These examples suggest that the data were judged not only from the roof surface of the target but also from other roof surfaces when the label of the damage degree of the test data was fed as input. Therefore, the model underestimated the degree of damage because it was trained and estimated based only on information from the roof surface of the target. Additionally, using aerial photos with low resolution as inputs to the label of the degree of damage probably resulted in the mixing of individual subjects in the judgment criterion. This implies that the difference in judgment criteria between the staff of the local government inputting labels of test data and training data caused low recall of image data with large damage.

7. CALCULATION OF DAMAGE RATE

7.1 Calculation Method of Estimated Damage Rate

In this study, the value of the degree of damage was assigned to each class classified in Section 5. The values of no damage was set to 0, damage (-25 %) to 0.125, damage (25–50 %) to 0.375, damage (50–75 %) to 0.625, and damage (75 %) to 0.875 as degree of damage. Then, the model multiplied these degrees of damage by the area rate and summed the values of every roof surface to calculate the damage rate of the roof. The area rate of the roof surface was calculated by dividing the number of pixels of each roof surface by the number of pixels of the entire roof.

7.2 Comparison of Correct Damage Rate and Estimated Damage Rate

The model estimated the damage rate of 30 images of the roof with damage and 30 images without damage to the test data using the above calculation method based on the degree of damage of each roof surface estimated in Section 5. The coefficient of determination between the correct and estimated damage rates was 0.3445. The coefficient of correlation was 0.6486, the average error was -5.401, and the average absolute value error was 11.07. Figure 10 shows a scatter diagram of the correct and estimated damage rates. The average error of the images of the roof with damage was -13.44 and that of the images without damage was 2.461.

For images of the damaged roof, the average error indicated that the model underestimated the damage rate. This was owing to an underestimation of the classification of the degree of

damage in Section 6. Similarly, for the images of the roof without damage, the average error indicated that the model estimated multiple data as little damage and only 12 of 30 images as no damage. This is because, the model incorrectly calculates the damage as a whole if the classification model estimates even one roof surface without damage as damage. Therefore, the model must increase the recall of no damage in the classification of the degree of damage.

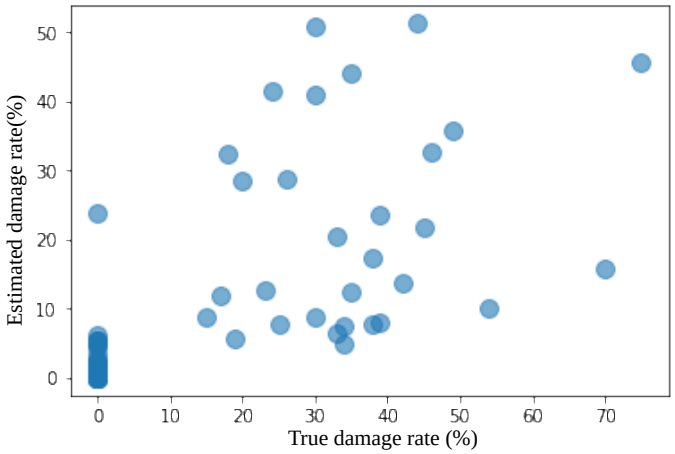

Figure 10. Scatter diagram of correct damage rate and estimated damage rate

8. EVALUATION OF SYSTEM EFFECTIVENESS

8.1 Result of Evaluation

Table 4. Question and answer to each roof data

Answer choices	It can.	It cannot.	total
By using aerial photos, can this system present the part which cannot be seen from outside?	60(100%)	0(0%)	60

Answer choices	It can advise perfectly and improve efficiency.	It can advise partially and improve efficiency.	It can estimate no damage roof to be ”No damage” class.	It has too much mistaken advice and cannot improve efficiency.	total
To improve efficiency in the investigation, can this system advise to investigate the part with high probability of damage?	16(26.67%)	32(53.33%)	12(20.00%)	0(0%)	60

Answer choices	It is as accurate as investigator judging from aerial photos and the ground.	It is more accurate than investigator judging from the ground.	It is as accurate as investigator judging from the ground.	It is more inaccurate than investigator judging from the ground.	total
How accurate is it?	15(25.00%)	3(5.000%)	30(50.00%)	12(20.00%)	60

To evaluate the effectiveness of the three usages (Table 1) of the proposed system in building damage investigation, we framed three questions for the staff of the Department of Crisis Management of Shimanto-town as shown in Table 4 for each of the 60 images of the roof surface estimated by the proposed system and obtained the relevant answers. The answers were based on the images of the roof shown in Figure 11, correct damage rate, and estimated damage rate. The image on the left of Figure 11 shows the image of the roof with a width of 3 m from the trimming roof image, and the right image shows the trimming image of the roof, which visualizes the estimated degree of damage of each roof surface.

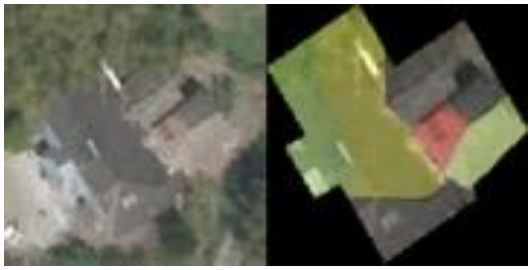

Figure 11. Roof data used for framing questions

The first question is “by using aerial photos, can this system present the portion that cannot be seen from outside?” The answer was “It can.” with 100 %. Therefore, this system can present information about the part of the roof that cannot be seen from outside by using aerial photos, which is the first usage of the system. This indicates that the proposed system can improve the efficiency and accuracy of investigations. The second question is “to improve the efficiency of investigation, can this system advise to investigate the part with high probability of damage?”. The answer was “the system provides incorrect advice and cannot improve efficiency” with 0 %. There were 18 roof data points without damage that the system estimated as a damage of a few percent by mistake. However, the answers of all of these 18 roof data were not “It has too much mistaken advice and cannot improve efficiency.” but “it can advise partially and improve efficiency”. According to the respondent, the system presented information about the roof surface without damage and could improve the efficiency of the investigation. Therefore, this system can present information about roof parts with a high probability of damage and advise investigation to investigators, which is the second usage of the system. This indicates that the proposed system can improve the investigation efficiency. In the evaluation of this model, it is desirable to quantitatively compare its accuracy with the accuracy of the investigation by an investigator who looks at the roof from the ground or from the ground and aerial photos. However, it is difficult to obtain these data. Thus, the staff were advised to classify the estimated accuracy of each roof data into four classes, such as “equal to investigators who look at the roof from ground and aerial photos,” “(lower than the above and) higher than the investigator who looks at the roof from ground,” “(lower than the above and) equal to the investigator who looks at the roof from ground,” and “lower than the investigator who looks at the roof from ground” to qualitatively evaluate the accuracy. This indicates that the proposed model can calculate the damage rate of 30 % of roof data more accurately than the investigators who look from the ground, which is a conventional judgment method, and that of 80 % of roof data is the same accurately or more than the investigators who look from the ground.

We asked additional questions about the overall system. First, “can this system show basis for judgment of calculation result by obtaining estimation of degree of damage of each roof surface?” The answer was “It can.” Therefore, this indicates that the proposed system can interpret the estimation using a deep learning model, which is a black box model, by calculating and visualizing the degree of damage to each roof surface. Next, “under which conditions can you use the estimated damage rate as the investigation result directly?” The answer was, “We cannot use the estimation directly if the roof is damaged because we need to understand the feelings of the victim. Hence, the presence of damage must be verified manually.” Another answer was that “investigating the roof in the field is necessary because it is difficult to judge damage from the data used by the system” and “estimated damage rate is useful for reference data.” Therefore, these results indicate that the calculated damage rate of the roof cannot be used for building damage investigation directly without a human check; however, the damage rate estimated by using the proposed system is useful for reference data.

8.2 Usage Flow of Our Proposed System

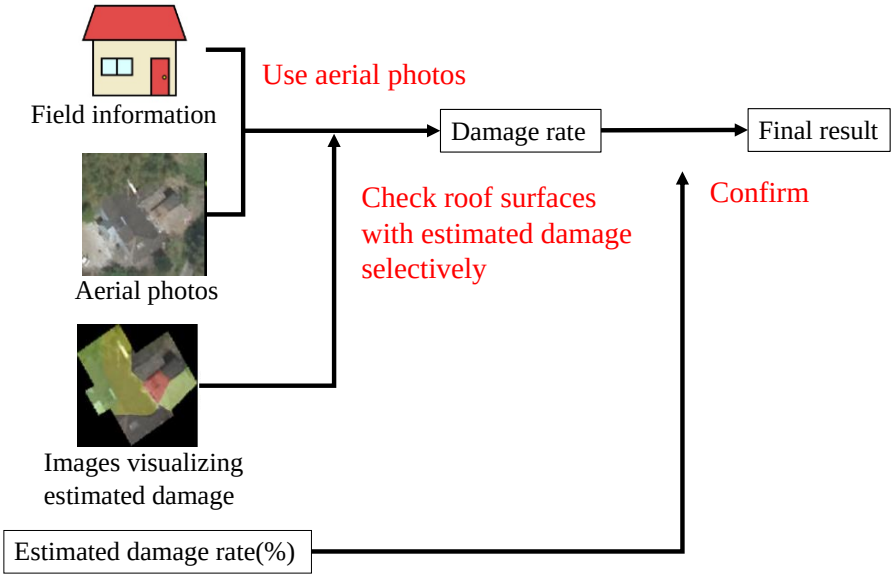

Figure 12. Usage flow of the system in the investigation field

Considering the results presented in the previous subsection, the usage flow, shown in Figure 12, incorporating the three usages mentioned in Table 1, can be considered. First, focusing on the roof parts with a high probability of having a high value of image visualizing the estimation of the system, the information from the field and aerial photos were investigated. Then, with reference to the damage rate estimated by the system, the calculated degree of damage was compared and confirmed to determine the final degree of damage. This indicates that the investigation can be more rapid and accurate in the red regions of Figure 12, using both the second and third usages as reference data.

8.3 Map of Damage Estimation

Figure 13. Left: map of building damage (Miura *et al.* 2020) and right: map of damage estimated by this system

We estimated the damage rate of the image of the roof of 3,513 buildings around the government office of Mashiki town in Kumamoto prefecture using the proposed system and plotted it on the map, as shown on the right side of Figure 13. Let d be the degree of damage estimated using this system. The value of $d \leq 10$ % is set to light blue, 10 %$< d \leq$20 % to yellow-green, 20 %$< d \leq$30 % to yellow, 30 %$< d \leq$40 % to orange, and 40 %$< d \leq$50 % to red. The larger the damage rate, the larger the point size set on the map. The left side of Figure 13 shows the actual damage map (Miura *et al.* 2020), and the data in the area surrounded by the dotted line reflect the results of the on-the-spot investigation by the Architectural Institute of Japan. These maps were compared to identify common areas, such as areas where the damage was concentrated. Because this study focused on roof damage and Miura *et al.* (2020) focused on whole building damage, an accurate comparison of the accuracy is not possible. However, according to Figure 13, rough spatial distributions between these maps are similar.

We believe that this map is effective in grasping the necessary resources, such as manpower for investigation or deciding the method of investigation. This indicates that the map is effective in grasping the overall damage in the entire area, as well as on the roof. Additionally, it is likely that this map is useful not only for building damage investigation but also for decision-making regarding the first response after a disaster or deciding the rehabilitation plan.

9. CONCLUSIONS AND FUTURE RESEARCH

In this study, we developed a system to automatically calculate the damage rate of a roof and suggested its usage to improve the speed and accuracy of building damage investigations.

A drawback of our previous study was that the aerial photos captured were insufficient as training data for the estimation of damage using deep learning from aerial photos (Fujita and Hatayama 2021). Therefore, this study divided the roof images into roof surfaces and increased the amount of training data. We developed an automatic method to calculate the damage rate of a roof from four processes: a trimming algorithm to automatically create roof images, dividing the roof surface by instance segmentation using the deep learning model, classification of the degree of damage by the image classification model of deep learning, and calculation of the damage rate from the degree of damage and area rate.

The effectiveness of this system was evaluated based on its capability to present information of the parts that could not be seen by the investigators using aerial photos, which is the first usage, and highlight parts with a high probability of damage to the investigators, which is the second usage. Moreover, this model could calculate the damage rate of 30 % of roof data more accurately than judgment from ground and that of 80 % of roof data with the same accuracy or more than judgment from ground. We then suggested the usage flow of this system in building damage investigation based on the answers provided by the staff to the questions framed by us. Additionally, we observed that the map shown on the right side of Figure 13 for plotting the estimation result of this system was effective for estimating the overall damage. This indicates that the proposed system is effective for both the estimation of local damage, representing the damage rate of one roof, and overall damage, representing damage to the entire area.

In the future, a more accurate system can be developed based on the following improvements:

- Information around the target roof surface is added to the input value of the model to classify the degree of damage.
- Increase the recall of roof data without damage in the classification of the degree of damage.
- Generate more general training data based on the label inputs of several individuals.
- Address roof data with a gap in the building polygon and low resolution.

In addition, the use of image data of roofs damaged by typhoons may be effective as training data. There may be common features, such as dropped tile roofs, between earthquakes and typhoons. However, the presence of different features may decrease the accuracy of deep learning models. Therefore, a fine-tuned model that uses earthquake image data after training with typhoon image data may be effective. This study needs to confirm the increase in accuracy by performing fine-tuning. In addition to image information, information such as earthquake acceleration, building type, and the raw material of the roof may be effective. Because neural networks can be trained using this information as input data, this study needs to consider the development of a multimodal model.

Moreover, we found that the estimation results of this system cannot be used directly in the field of building damage investigations. This indicates that it is difficult to use a deep learning

model for decision making because it lacks reliability and explainability. This is a limitation specific to deep learning or machine learning systems during disasters as well as a lack of training data. In the future, we aim to design a usage to improve the reliability and explainability of the model and ensure responsibility by combining the efforts of the systems and humans.

REFERENCES

Department of crisis management, Kumamoto Prefecture. (2020) Report about damage status of the Kumamoto earthquake in 2016 (308th reports). https://www.pref.kumamoto.jp/uploaded/attachment/126303.pdf (In Japanese)

Disaster Management, Cabinet Office, Government of Japan. (2020) Guidance of the Implementation System for Building Damage Investigation in Disasters. https://www.bousai.go.jp/taisaku/pdf/r203saigai_tebiki_full.pdf (In Japanese)

Disaster Management, Cabinet Office, Government of Japan. (2020) Guidelines of the Operation of Criteria for Building Damage Investigation in Disasters. https://www.bousai.go.jp/taisaku/pdf/r203shishin_all.pdf (In Japanese)

Disaster Management, Cabinet Office, Government of Japan. (2021) Guidelines of the Operation of Criteria for Building Damage Investigation in Disasters. https://www.bousai.go.jp/taisaku/pdf/r303shishin_all.pdf (In Japanese)

E-Stat (2018) Housing and Land Survey, Basic data on housing and households. https://www.e-stat.go.jp/dbview?sid=0003355582 (in Japan)

Fujita, A., Sakurada, K., Imaizumi, T., Ito, R., Hikosaka, S., and Nakamura, R. (2017) Damage Detection from Aerial Images via Convolutional Neural Networks. 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA), 5-8. https://doi.org/10.23919/MVA.2017.7986759

Fujita, S. and Hatayama, M. (2021) Estimation Method for Roof-Damaged Buildings from Aero-Photo Images During Earthquakes Using Deep Learning, *Information Systems Frontiers*. https://link.springer.com/article/10.1007/s10796-021-10124-w

Fujiu, M., Ohara, M., and Meguro, K. (2012) Development of Survey Management System for Building Damage Assessment. Proceedings of the International Symposium on Engineering Lessons Learned from the 2011 Great East Japan Earthquake, Tokyo, Japan. https://www.jaee.gr.jp/event/seminar2012/eqsympo/pdf/papers/168.pdf

He, K., Gkioxari, G., Dollar, P., and Girshick, R. (2017) Mask R-CNN. Proceedings of the IEEE International Conference on Computer Vision, 2961-2969. arXiv:1703.06870. https://doi.org/10.48550/arXiv.1703.06870

He, K., Zhang, X., Ren, S., and Sun, J. (2015) Deep residual learning for image recognition. Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition, 770-778. arXiv preprint arXiv:1512.03385. https://doi.org/10.1109/CVPR.2016.90

Inoguchi, M., Tamura, K., and Hamamoto, R. (2019) Establishment of Work-Flow for Roof Damage Detection Utilizing Drones, Human and AI based on Human-in-the-Loop Framework. In 2019 IEEE International Conference on Big Data (Big Data), 4618-4623. https://10.1109/BigData47090.2019.9006211

Inoue, M., Suetomi, I., Fukuoka, J., Onishi, S., Numata, M., and Meguro, K. (2018) Development of Estimation Formula of Disaster Response Work Volume based on the Kumamoto Earthquake. *Monthly journal of the Institute of Industrial Science*, University of Tokyo, *70*(4): 289-297. (In Japanese) https://doi.org/10.11188/seisankenkyu.70.289

Ise, T., Minagawa, M., and Onishi, M. (2018) Classifying 3 Moss Species by Deep Learning, Using the "Chopped Picture" Method, *Open Journal of Ecology*, *8*(3): 166-173. https://doi.org/10.4236/oje.2018.83011

Ji, M., Liu, L., Du, R., and Buchroithner, M.F. (2019) A Comparative Study of Texture and Convolutional Neural Network Features for Detecting Collapsed Buildings After Earthquakes Using Pre- and Post- Event Satellite Imagery. *Remote Sensing*, *11*(10): 1202. https://doi.org/10.3390/rs11101202

Matsuoka, M., Ishii, Y., Maki, N., Horie, K., Tanaka, S., Nakamura, R., Hikosaka, S., Imaizumi, T., Fujita, A., and Ito, R. (2018) Damaged Building Recognition Using Deep Learning with Photos Taken after the Kobe Earthquake. Proceedings of the 11th U.S. National Conference on Earthquake Engineering. https://www.11ncee.org/images/program/papers/11NCEE-001788.pdf

Meloy, N., Sen, R., Pai, N., Mullins, G. (2007) Roof damage in new homes caused by Hurricane Charley. *J Perform Constr Facil 2007*;21(2):97–107. http://dx.doi.org/10.1061/(ASCE)0887-3828(2007)21:2(97)

Ministry of Internal Affairs and Communications (2018) Housing and Land Survey, Approximate number of housing units. https://www.stat.go.jp/data/jyutaku/2018/pdf/g_gaiyou.pdf (In Japan)

Ministry of Land, Infrastructure, Transport and Tourism (2016) Key Points of the Report "Committee to Analyze the Causes of Building Damage in the Kumamoto Earthquake". https://www.mlit.go.jp/common/001155087.pdf (In Japanese)

Miura, H., Aridome, T., and Matsuoka, M. (2020) Deep Learning Based Identification of Collapsed, Non-Collapsed and Blue Tarp-Covered Buildings from Post-Disaster Aerial Images. *Remote Sensing*, 12(12): 1924. https://doi.org/10.3390/rs12121924

Naito, S., Tomozawa, H., Mori, Y., Monma, N., Nakamura, H., and Fujiwara, H. (2020) Development of the Deep Learning Based Damage Detection Model for Buildings Utilizing Aerial Photographs of Plural Earthquakes. *Journal of Japan Association for Earthquake Engineering*, *20*(7): 7_177-7_216, 2020. (In Japanese) https://doi.org/10.5610/jaee.20.7_177

Shigekawa, K., Tanaka, S., Horie, k., Hayashi, H. (2005) Some Regulational Issues on Building Damage Assessment ―A Case Study of Niigata-ken Chuetsu Earthquake－. *Journal of Social Safety Science*, No. 7, 133-140, 2005. (In Japanese) https://www.jstage.jst.go.jp/article/jisss/7/0/7_133/_article/-char/ja/

Tanaka, S., Shigekawa, K., and Takashima, M. (2008) Development of the Building Damage Self-Inspection System for Earthquake Disaster. The 14th World Conference on Earthquake Engineering, Beijing, China. http://www.iitk.ac.in/nicee/wcee/article/14_01-1028.pdf

Tu, J., Li, D., Feng, W., Han, Q., and 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

Vetrivel. A., Gerke, M., Kerle, N., Nex, F., and Vosselman, G. (2018) Disaster damage detection through synergistic use of deep learning and 3D point cloud features derived from very high- resolution oblique aerial images, and multiple-kernel-learning. *ISPRS Journal of Photogrammetry and Remote Sensing, 140:* 45-59. https://doi.org/10.1016/j.isprsjprs.2017.03.001