 doi:10.20944/preprints202204.0240.v1

Article

Damage Detection and Level Classification of Roof Damage After Typhoon Faxai Based on Aerial Photo and Deep Learning

Jinglin Xu[1], Feng Zeng[2], Wen Liu[3] and Toru Takahashi[4,*]

[1] Graduate Student, Department of Architecture, Chiba University, 263-8522, Chiba prefecture, Japan; ayfa6114@chiba-u.jp
[2] Lecturer, Department of Architecture, Xihua University, 610097, Sichuan Province, China;
[3] Research Associate, Department of Urban Engineering, Chiba University, 263-8522, Chiba prefecture, Japan;
[4] Professor, Department of Architecture, Chiba University, 263-8522, Chiba prefecture, Japan;
* Correspondence: takahashi.toru@faculty.chiba-u.jp (T.T) Tel:81-43-290-3145.

Abstract: Following the occurrence of a typhoon, quick damage assessment related to residents can facilitate quick dispatch of house repair and disaster insurance works. Employing a deep learning method, this study used aerial photos of the Chiba prefecture obtained following the Typhoon Faxai in 2019 to automatically detect and evaluate the roof damage. This study comprised three parts: training deep learning model, detecting the roof damage using trained model, and classifying the level of roof damage. The detection object comprised roof outline, blue tarps, and roof completely destroyed. The roofs were divided into three categories: roof without damage, roof with blue tarps and roof completely destroyed. the F value obtained using the proposed method was higher than those obtained using other methods. In addition, it can be further divided into 5 levels from level 0 to 4. Finally, the spatial distribution of the roof damage was analyzed using ArcGIS tools. The proposed method is expected to provide certain reference for the real-time detection of the roof damage after the occurrence of a typhoon.

Keywords: Deep Learning; Aerial photo; Typhoon Faxai; roof damage; detection; classification

1. Introduction

1.1. Purpose and significance

Reports on the field investigation following the occurrence of typhoons in the past [1,2] have reported houses, particularly low-rise wooden houses as being significantly affected by strong winds and often suffering serious damage. After the typhoon, quick damage assessment of residents' houses can facilitate quick dispatch of necessary assistance to house repair and the confirmation of disaster insurance to residents. Therefore, in March 2017, the Japan Cabinet Office revised "the guidelines for the identification and classification of residential house damage after disasters" [3] to improve the efficiency of filed investigation for house damage by disasters. In addition, local governments requested amendments to the "disaster relief act" and "act on support for reconstructing livelihoods of disaster victims" to propose investigation methods and evaluation criteria suitable for strong wind disasters. Recently, Typhoon Faxai in 2019 caused serious damage to houses in Japan. Consequently, the effective and rapid detection and classification of the damage has become an urgent topic in Japan.

In Japan, the assessment of the disaster situation regarding a house following a typhoon is mostly conducted by local government and disaster research groups through field investigations [4]. Thus, the result of the evaluation is released after a few days. The disaster field investigation is conducted not only for typhoons, but following the occurrence of earthquakes as well. For example, according to the "investigation on the situation of certificates issuing for disaster in case of large-scale disasters - centered on the Kuma-

 doi:10.20944/preprints202204.0240.v1

moto earthquake in 2016" [5] conducted by the Ministry of general affairs Japan, the certificates issuance for disaster was delayed by more than three months the Kumamoto earthquake occurred.

From November 10 to 13, 2019, a field investigation was conducted on the residential house damaged after Typhoon Faxai in 2019, considering Tateyama city, Kyonan town, and Minamiboso city in the south of Chiba Prefecture as the investigation area. The statistical result of field investigation is presented in Table 1, where more than half of the houses in the survey area were damaged, and roof with damage accounted for more than 80% of the damaged houses, far exceeding the damage proportion of other components of the house. Thus, it was found that the damage condition of the house can be speculated based on the damage condition of the roof.

Table 1. Statistics of Field investigation after typhoon Faxai in 2019 [4].

Investigation	Total	Percentage
Surveyed houses	1,976	-
Damaged house	1,133	57%
Damaged Roof	902	46%
Damaged Wall	228	12%

1.2 Previous research

After the 1960s, use of aerial photographs in the field investigation after typhoon, earthquake, tsunami and other disasters gained traction. Aerial photographs can provide local information with rich surface texture and spectral characteristics. Rapid and effective house damage extraction has been one of the most important research contents in the field of aerial photo processing. For example, Miura et al. [6] successfully measured the damaged area of houses in the 1995 Hanshin earthquake with aerial photos.

There are three extraction methods of roof damage detection following a disaster based on aerial photo.

The first method is visual inspection. The damage extraction is primarily performed manually by architects, emergency risk judgers, and other civil field experts. For example, Noda et al. [7] conducted a field investigation using a small aircraft after the typhoon JEBI in 2018. Considering the south of Osaka Prefecture and the north of Wakayama Prefecture as the survey areas, they investigated the distribution of blue tarps covered on residents' roofs. The results indicated a linear relationship between the number of blue tarps, such that, the number of damaged houses can be estimated using the number of blue tarps. Suzuki et al. [8] successfully detected houses with blue tarps covering the roof, using the aerial photos taken after tornadoes in Tsukuba, adhering to the criteria for damage identification issued by the Japan cabinet office. However, although this method offered high accuracy, it often requires several days.

The second method is image analysis, which is based on the pixel, threshold of the spectrum, brightness, and texture of aerial photos to judge house damage. When the threshold is exceeded, roof damage is concluded to have occurred. For example, Kono et al. [9] used the aerial photos taken after typhoon Jebi in 2018 to assess the roof damage of residential houses depending on whether blue tarps were covered, and used the results to infer the damage rate of the roof. In addition, Kono et al. [10] established a method of identifying the damage of residential roofs in a wide area using aerial photos acquired after Typhoon Faxai in 2019. In addition, it was compared to visual inspection results and its accuracy was confirmed as efficient. Naito et al. [11] used the aerial photos taken immediately after the Kumamoto earthquake, and used DSM analysis and texture analysis method to extract the roofs covered with blue tarps. However, this method does not provide clear technical points regarding threshold setting.

 doi:10.20944/preprints202204.0240.v1

The third method is deep learning [12], wherein a single house is considered as the extraction object. For example, Miura et al. [13] used the aerial photo of the 1995 southern Hyogo and 2016 Kumamoto earthquakes to train a deep learning model. The model can be used to automatically classify the damage grade of houses into three levels: collapsed houses, non-collapsed houses, and houses covered with blue tarps. Further, Miura verified the accuracy of the model at Kyonan town, Chiba prefecture after Typhoon Faxai in 2019.

As mentioned above, the method of detecting the roof damage using the distribution of blue tarps covered on the roof is widely used after disasters such as typhoons and earthquakes in Japan. Moreover, the results of above research show that with the increased blue tarps area, there is increased loss of residential roofs.

However, the current detection methods for roof damage in aerial photos suffer from two primary shortcomings. First, it is impossible to calculate the proportion of roof damage. Currently, the existing basic outline of house provided by Japan Institute of land and geography is used. However, in addition to new construction, reconstruction and expansion of the house, there are also certain completely destroyed houses caused by the typhoon, resulting in great changes in the roof outline. Therefore, the roof outline must be measured in real time. Second, roofs that are completely destroyed and not covered with blue tarps cannot be detected. If the house is seriously affected by typhoon, for example, collapses or damage due to a secondary fire disaster, it is impossible to continue to live in them. Consequently, many residents tend to move to another safe place temporarily and although the house is seriously damaged, there is no need to continue to cover using blue tarps. Therefore, this type of house damage has been omitted.

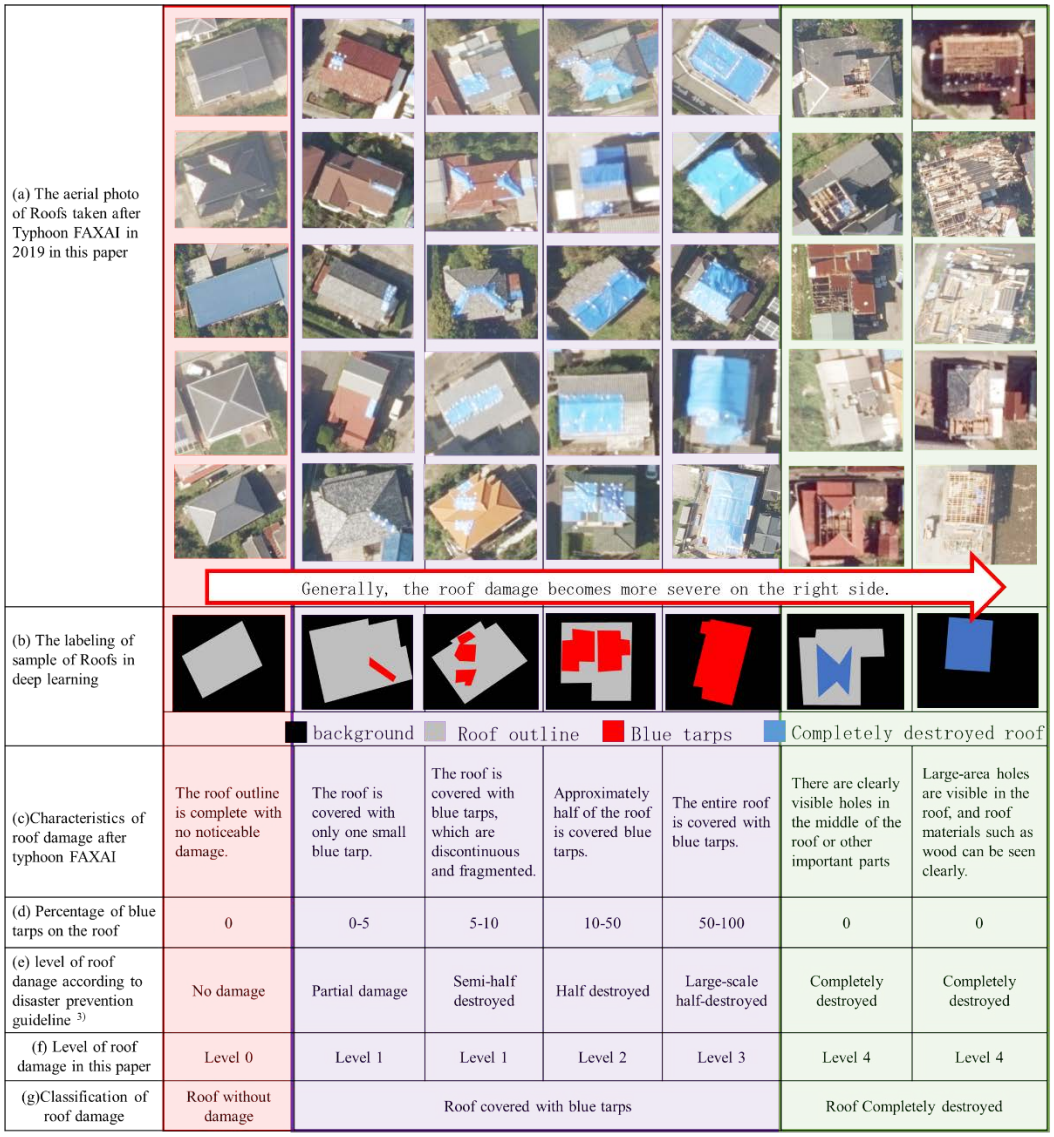
Thus, it is incorrect to only count the area of blue tarps and roofs completely destroyed must be detected separately. Consequently, the sum of the area of blue tarps and the area of the roof completely destroyed must be considered as the total damaged area of the roof in the investigation area. In addition, according to the previous deep learning model, only two or three types of the level of roof damage can be classified, for example, damaged or undamaged roofs. Moreover, there is a lack of specific classification basis for calculating the area of damaged roof.

Therefore, in this study, a deep learning model with segmentation function was used to detect roof outline, blue tarps, and roofs completely damaged simultaneously. Then, according to the detection result of this three objects, the grade of roof damage is divided into five levels: roof without damage (level 0), roof covered with blue tarps (level 1-level 3), and roof completely destroyed (level 4). In addition, level 1-level 3 is be classified according to the proportion of blue tarps in the roof.

2. Research methods

Figure 1 shows the specific appearance of the roof damage after a typhoon. In Figure 1 (a), the roof is clearly divided into three parts. Left represents the roof without damage, the middle is the roof covered with blue tarps (coverage area of blue tarps increases from left to right), and the right is the roofs completely destroyed (function of the roof was completely lost, thus, was not necessary to cover using blue tarps). Thus, from the left to the right, the degree of roof damage increases. In Figure 1 (f), according to "the guideline for disaster prevention" [3] by the Japan cabinet office, the level of roof damage after typhoon is divided into five levels: levels 0 ~ 4. The roof without damage is level 0. The roof covered with blue tarps is divided into levels 1 - 3 according to the proportion of blue tarps in the roof. Finally, the roof completely destroyed is classified as level 4. The specific distinction is shown in Figure 1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1



(a) The aerial photo of Roofs taken after Typhoon FAXAI in 2019 in this paper							
(b) The labeling of sample of Roofs in deep learning							
(c)Characteristics of roof damage after typhoon FAXAI	The roof outline is complete with no noticeable damage.	The roof is covered with only one small blue tarp.	The roof is covered with blue tarps, which are discontinuous and fragmented.	Approximately half of the roof is covered blue tarps.	The entire roof is covered with blue tarps.	There are clearly visible holes in the middle of the roof or other important parts	Large-area holes are visible in the roof, and roof materials such as wood can be seen clearly.
(d) Percentage of blue tarps on the roof	0	0-5	5-10	10-50	50-100	0	0
(e) level of roof danage according to disaster prevention guideline [3)]	No damage	Partial damage	Semi-half destroyed	Half destroyed	Large-scale half-destroyed	Completely destroyed	Completely destroyed
(f) Level of roof damage in this paper	Level 0	Level 1	Level 1	Level 2	Level 3	Level 4	Level 4
(g)Classification of roof damage	Roof without damage	Roof covered with blue tarps				Roof Completely destroyed	

Figure 1. Appearance and level of roof damaged in Typhoon Faxai in 2019.

In this study, it was found that dividing roofs (for instance, the roof which is neither covered with blue tarps nor completely destroyed, but with some holes,) into two types can greatly improve the accuracy of classification. One type is roof with small hole (less than approximately 5% of the roof area). The hole will not affect the normal life of residents and thus, it can be marked as "roof without damage." The other roof type can be marked as "roof completely destroyed" in the case of roof with huge holes (more than 80% of the roof area). This is because such holes are not covered with blue tarps. Such type of houses are completely unfit for living, or the residents have moved to another place during the field investigation. Such a classification method may lead to detection omission, such as, classifying an actual "roof with damage" as "roof without damage;" however, it is beneficial to the overall accuracy of "roof completely destroyed."

This study was mainly divided into three parts. First, based on the aerial photos taken after the typhoon Faxai, the images of roofs were obtained. Thereafter, the obtained images were used to render samples for training a deep learning model. Finally, the model was used to automatically detect three types of objects: blue tarps, roof outline, and roofs completely destroyed, and the level of roof damage was classified through the combination of three types of objects (Figure. 2).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1

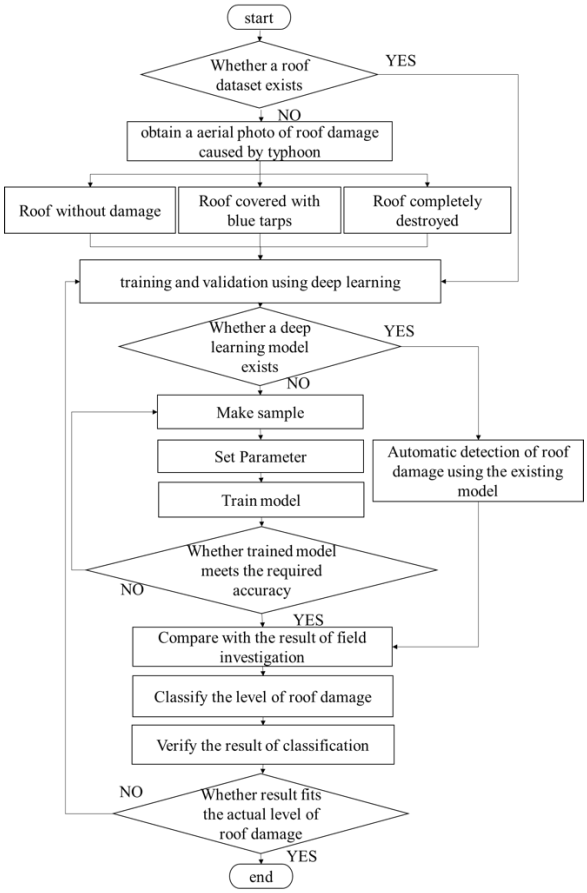


Figure 2. Flowchart of detection and classification of damage roof in this paper.

3. Materials (aerial photo)

3.1 Used aerial photos

When the damaged area is relatively large, generally, satellite images are used for photography of roof damage. In contrast, in the case of disasters with limited area, photos obtained via aircrafts offer better image resolution than satellite images [10]. This study selected the aerial photos of Typhoon Faxai in 2019 taken by International Airlines Company from September 19 to 20 in 2019. The area included a part of the Minamiboso and Tateyama cities. All aerial photos were acquired using a digital camera DMC for aerial survey built by Z / I Imaging Company. Simple orthophoto with terrain skew corrected and distortion of camera lens fixed were obtained. They were composed of 8-bit Blue, Green and Red bands. Finally, the aerial photos were resampled to 0.1 m/pixel.

Figure 3 shows a data crop at the south of Chiba prefecture, where the area with red circle is the field investigation area (G1 is a part of Minamiboso city, and G2 and G3 are a parts of Tateyama city). The black square area is the photographing area taken by aerial plane, which were purchased with scientific research fees. The area with yellow cross represent the manually labeling samples used for training the deep learning model (18 areas). The area with red 7-point star is the area used for testing the accuracy of the trained model, including k1-k5.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1

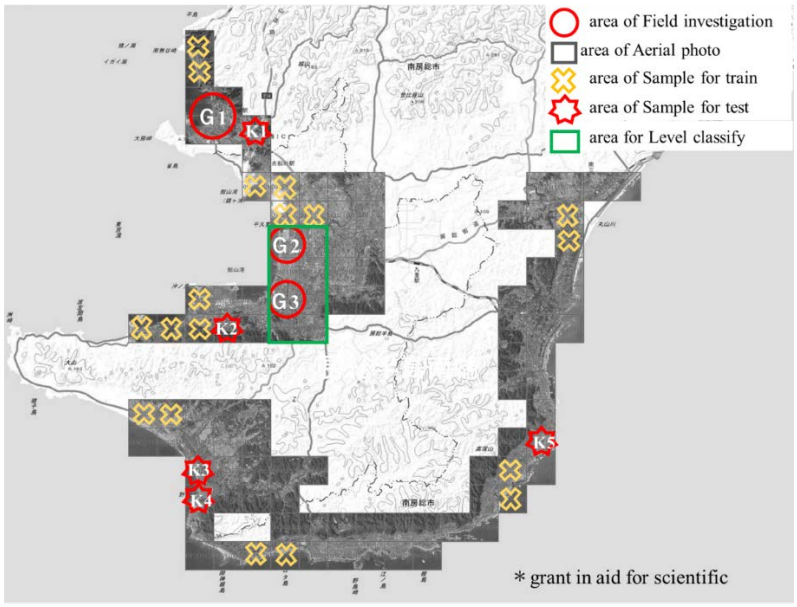


Figure 3. Area photo used (southern part of Chiba prefecture).

3.2 Labeling and cropping the aerial photo

The upper left corner block "a" (Figure. 4) was assumed as a piece of the aerial photo in Figure. 3, and used as an example to illustrate the method in this study.

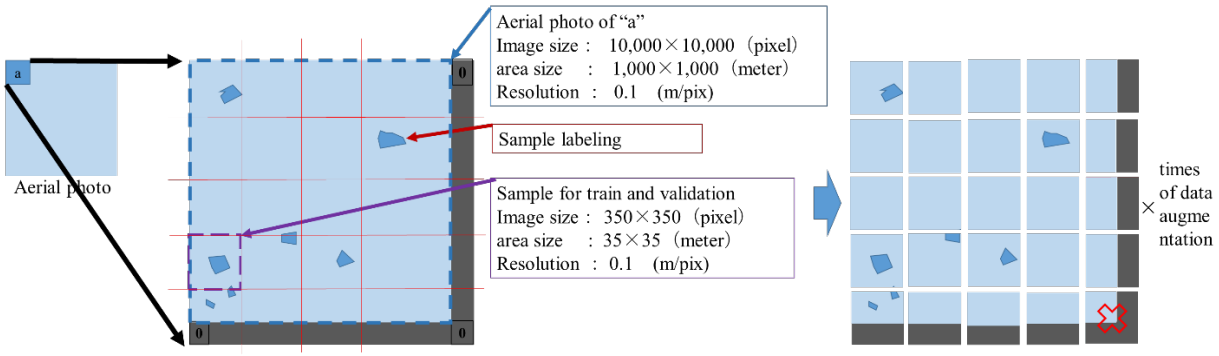


Figure 4. Schematic of labeling and cropping on the aerial photo.

The size of this aerial photo is 10000 × 10000 pixels. Further, the largest area in the manual labeling data is the roof profile with 260 × 240 pixels, whereas the smallest is blue tarps with 20 × 20 pixels.

First, through manual labeling, the three types of data, comprising roof outline, blue tarps, and roof completely destroyed, were labeled on the aerial photos, which were then used to create training samples for the deep learning. The labeling was mainly operated via visual inspection with error rate less than 5%. The operators of this study were professionals in architecture and civil from Chiba University. To avoid data obfuscation between the training and validation data, that can affect the validation accuracy, the labeling area was mainly set along the coastline with more damaged roofs. In the area for training sample, as in Figure. 3, 25,205 data were labeled for training in Section 4.2. In addition, 3,269 data were made in k1-k5 area of Figure. 3 to test the accuracy of the trained model. This study compared the detection results using the model in G1-G3 area of Figure. 3 with those obtained from the field investigation after typhoon Faxai in 2019. The statistics of the labeling data is presented in Table 2.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1

Table 2. Statistics of the labeling data on the figure 3.

Number of labeling	Roof out-line	Blue tarps	Roof completely destroyed	Total
Labeling at area T1-T18	21,600	3,221	384	25,205
Labeling at area KI-K5	2,697	506	66	3,269
Labeling at all area	24,297	3,727	450	28,474

3.3 Export data

First, the above areas were cropped to pieces with dimensions 350 × 350 pixels. The size can include at least one or two roof to reflect the overall characteristics of the roof and those of the surrounding environment that inflects on the neural network. However, a very small size will result in the roof outline not being well detected. Whereas, too large a size will drastically reduce the detection accuracy of blue tarps. Thereafter, a 150-pixel width area with 0 value was added around the original aerial photo "a", and the aerial photo was divided into 29 × 29 training samples. Among them, the sample with the lowest 0 value accounting for the majority were deleted (Figure 4, it is OK not to delete).

In the subsequent stage, sample augmentation was conducted. Each cropped image was converted as follows. There were three rotation, three mirror, and four color space transformations, resulting in a total of 10 transformations being conducted. In color space conversion, Gaussian noise, random noise, brightness, and contrast were generated via mapping functions. Figure. 5 is a sample augmentation of one of these photos. Equation (1) is a formula for calculating the number of training samples created from aerial photos. Consequently, 8,800 samples were obtained from "a" of Figure. 4, whereas 151,200 training images were obtained from the entire area in 18 regions of Figure 3 (18 regions are the training sample regions indicated by the yellow cross.)

$$N_t = \left(\frac{(W_k + W_t) \times (H_k + H_t)}{W_s \times H_s} - N_n \right) \times N_k \quad (1)$$

where N_t is Number of training samples, W_s is cropping width (pixel), W_k is width of aerial photo (pixel), H_s is Height of cropping (pixel), W_t is width of added 0 value (pixel), N_n is number of discarded sample, H_k is Height of aerial photo (pixel), N_k is times of data augmentation, and H_t is Height of added 0 value (pixel).

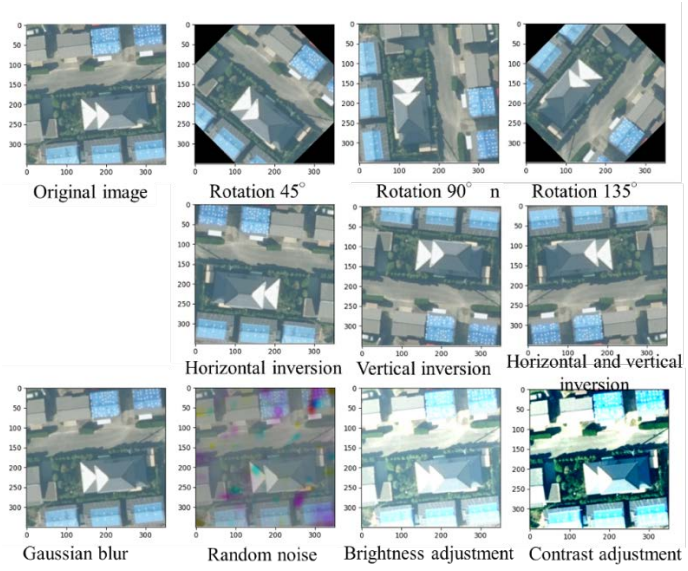


Figure 5. Data augmentation of the sample

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1

During the training stage, the training times can be calculated using Equation (2). Considering set A and D in Table 3 as an example. the training times were 17,010 times with a validation ratio of 10% and 8 batch sizes (8 samples trained simultaneously). In set D, the training times were 37,800 times with a validation ratio of 20%, 64 batch sizes (train 64 samples at the same time), and with 20 epochs. The epoch was set randomly in this study. The more the training times, the better; however, the training time increased. Furthermore, to inspect the performance of the model, the observation of the results was limited to 6000 times in this study.

$$\mathrm{T} = \frac{E \times N_t \times (1 - P)}{B} \tag{2}$$

where T is training times, P is the ratio of validation, E is train epoch, and B is batch size.

4. Training and detection using deep learning

This study employed the neural network Mask R-CNN [15] (CNN is the abbreviation of convolutional neural network,). It can automatically detect the different object with high precision and calculate the area, and outline of the objects. Mask R-CNN is a neural network improved by R-CNN [16], FPN [17], Fast R-CNN [18] and Faster R-CNN [19]. Mask R-CNN adds a full convolution module to the existing two last modules (object detection and classification) and be used for object detection, classification, and segmentation (classify each pixel separately).

A schematic of Mask R-CNN is shown in Figure. 6. Taking blue tarps as an example, 1) in the stage of Resnet, the input roof image was convoluted to obtain the feature map of the many feature. 2) In the stage of RPN, the candidate region (ROI), which is considered as blue tarps, was selected, and consequently, the location of blue tarps was obtained. 3) In the stage of ROI align, all candidate areas of blue tarps were unified into a certain size (7 × 7). 4) In the stages of L_{class}, L_{Bbox}, and L_{mask}, the candidate regions were sent to the neural network to output the results of object classification, detection, and segmentation.

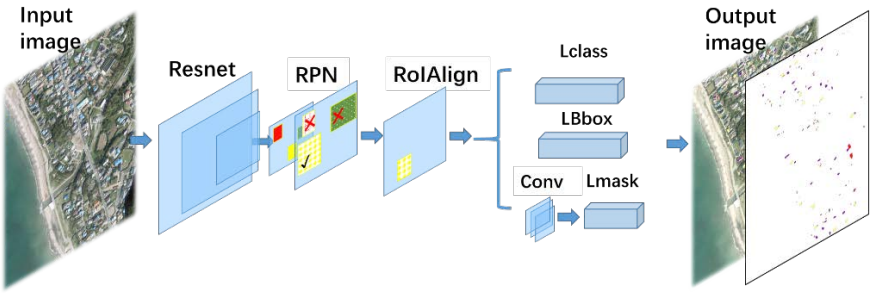


Figure 6. Framework of the deep learning

4.1 Making sample

Figure 7 shows a pie chart detailing each roof type and the level of roof damage in the sample. The left side is counted according to the type of roof, and the right side is counted according to the level of roof damage. Further, the training sample includes levels 0 - 4 in Figure 1 (a), as well all roof types identified in field investigation [4].

The parameters of the model can be understood as the weight composed of a series of values in deep learning. If there is a large number of samples, the weight of this type increases, which increases the probability of detection. In contrast, if the number of samples is small, the probability of failure to be detected is high. In this study, attempts were made to include different types of roof samples, such that the parameter of the model was balanced, and there were fewer cases where a certain type of roof was not be detected. In addition to locating the blue tarps in the center of the roof or on the eaves, samples of blue

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1

tarps hanging on the edges and corners were also marked as a supplement, which is relatively small, and may be neglected if not careful.

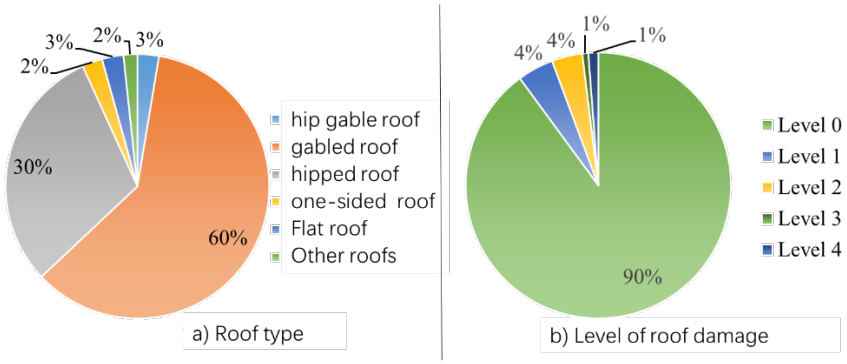


Figure 7. Composition of training sample

4.2. Setting parameters

When training the neural network, the initial learning rate, training times, batch size, and other parameters (Note 1 - Note 3) must be set. The manner in which to adjust the learning rate is an important stage to train a good model of deep learning.

The result and training model is affected by various parameters such as learning rate, proportion of validation, batch size and other factors. Currently, there exists no standard on the manner in which to obtain a set of best parameters. In the continuous experiment, the most appropriate parameters were selected in this study, and experiments were conducted using different learning rates and other parameters.

Table 3. Parameters for the model training

Parameters	Set A	Set B	Set C	Set D
Batch size (note1)	8	16	32	64
Ratio of training and validation	10%	10%	20%	20%
Backbone (note2)	Resnet 50	Resnet 50	Resnet 101	Resnet 101
Train time for evaluation	6,000	6,000	6,000	6,000
Initial learning rate (note3)	1.00E-02	1.00E-04	1.00E-03	5.00E-05

The training and the validation sample during the training acquired the ratio of 8:2 or 9:1. Further, Python language and Pytorch [20] were used to build a training framework using a workstation with 64 GB, windows 10 pro for workstations and 64bit Quadro RTX 5000 graphics card as the training environment. An entire training task required 2.5 h.

4.3 Training model

To eliminate the possible interference of cars and trees, a pre-trained model was used to transfer the parameters that have been trained in another task to improve the training efficiency. This study employed the model of the COCO data set [21] as a pre-trained model to improve the training effect.

4.3.1 Loss function in training

The loss function used in training comprises classification, detection, and segmentation losses (Equation (3)). The cross entropy of the real and predicted values of the object was used to calculate classification loss. Further, the real and predicted values of the four vertex coordinates of the rectangular box was used to calculated detection loss, and calculated using the smooth-L1 formula. Whereas, the binary cross entropy of mask's real

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1

and predicted values was used to calculate the segmentation loss. Furthermore, in each training, each loss value was summed to calculate the final loss value. The unit of the function loss is non-dimensional. In addition, the training quality of the model is determined by observing the decline of the loss curves later.

$$L = L_{class} + L_{Bbox} + L_{mask} \tag{3}$$

where L is training loss, L_{Bbox} is the loss of detection loss, L_{class} is classification loss, and L_{mask} is segmentation loss.

4.3.2 Optimization algorithm

Owing to the multitude of samples, the amount of calculation also increases. However, the loss value can be reduced using certain optimization algorithms. This study adopted the stochastic gradient descent (SGD) method as the optimization algorithm. The weight of the model was adjusted using Equation (4). Further, the learning rate Was adjusted using Equation (5) (inverse time decay method), such that the learning rate is inversely proportional to the current training times. The factor γ was set to 0.95.

$$W_{i+1} = W_i - R_{i+1} \times \nabla L \tag{4}$$

where W_{i+1} is updated weight in i+1 times, R_{i+1} is updated learning rate, W_i is *weight in i times*, and ∇L is gradient of loss value.

$$R_{i+1} = \frac{R_i}{1 + \gamma \times E} \tag{5}$$

where R_{i+1} is the updated learning rate which is expressed by Equation (5), R_i is the learning rate [Note 3)], γ is the factor which is set to 0.95.

4.3.3 Training process

Figure. 8 (a) shows the results of using set A of Table 3 for training. Initially, the loss value rapidly decreased. However, there was a sudden fluctuation in the loss value and it rebounded sharply at the 2500 training time. This is because there is a problem with the parameters, resulting in over fitting. Figure. 8 (b) shows the results using set B of Table 3. The loss value decreases gradually and is considered to be a good parameter that can be used for detecting. Further, Figure. 8 (c) shows the results for the parameters of set C of Table 3. The loss curve exhibits great changes in the initial stage and is stable in the later stage; however, the validation curve has two peaks. Finally, Figure 8 (d) shows the results using set D in Table 3. The parameters can be considered appropriate. The loss curve of training and validation in the training process decreases rapidly in the initial stage and tends to be stable in the last stage.

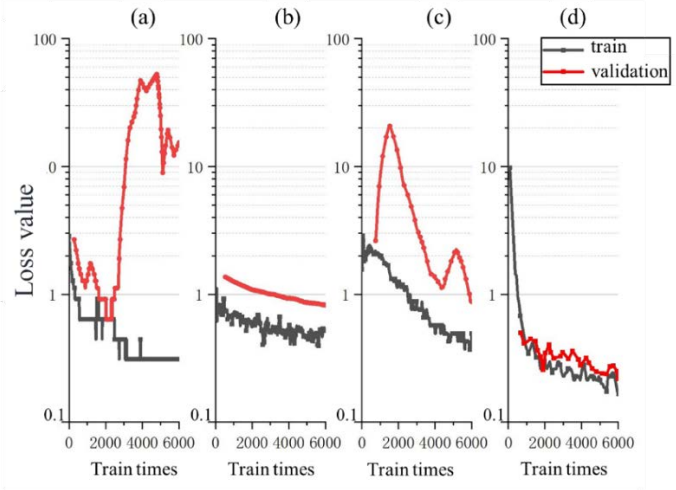


Figure 8. Loss value in the training and validation

 doi:10.20944/preprints202204.0240.v1

Parameters must be determined before model training and cannot be updated manually during training. Currently, there is no consensus on the selection of appropriate parameters. However, according to the various situations of loss curve and the user's personal preference, certain judgment conditions can be set for selection. The judgment condition used in this study was that the loss value reached less than 0.2 within 6,000 times without obvious rebound. According to this benchmark [22], set D was the best parameter, although set B can also be used according to personal preference.

4.4 Validation in training

In this study, it was determined that the overlap between the detected and actual targets was greater than 0.5 as the correct detection (T). Whereas, overlap less than 0.5 was considered as incorrect detection (F). Figure. 9 shows a calculation diagram of TP, FN, FP, and TN. Taking roof outline as an example, TP (true positive) is the situation that actual roof was detected as a roof. FN (false negative) is the situation that actual roof was not detected as a roof (missed detection). FP (false positive) is the situation where a roof is incorrectly detected in the absence of roof. TN (true negative) is the situation there is no actual roof and is not detected as one either. K in Equation (6) is the total number of roof outlines, blue tarps, and roofs completely destroyed. The accuracy of a general deep learning model is not clearly specified and is according to actual needs. The required accuracy of target segmentation is approximately 72% [20]. This study attempted to measure the specific area of roof damage, and at least 90% accuracy was required in the training stage. According to the adjustment of parameters, the learning accuracy was approximately 98%. Further, the *mIoU* (mean value of *IoU*) is usually used as an evaluation indicator. The calculation formula is shown in Equation (6).

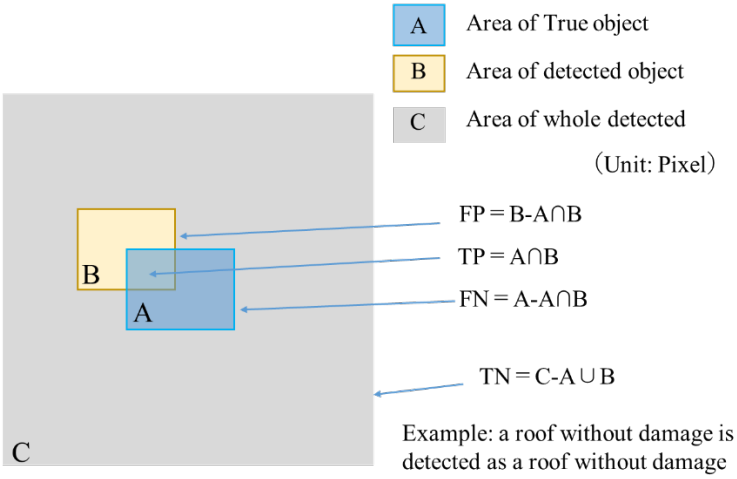


Figure 9. Validation in the training

$$mIoU = \frac{1}{k}\sum \frac{TP}{FP + FN + TP} \tag{6}$$

Figure. 10 is a result of validation in training. According to the results of roof outline, certain adjacent houses were mistaken for a whole object. The test result of blue tarps is the best. According to the results of roofs completely damaged, certain ships were mistakenly detected as roofs completely damaged. In the comparison between ground truth and the results of validation, it is found that the target region is accurate and satisfies the accuracy of segmentation. The mIoU of roof outline, blue tarps, and roofs completely destroyed was more than 98%.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1

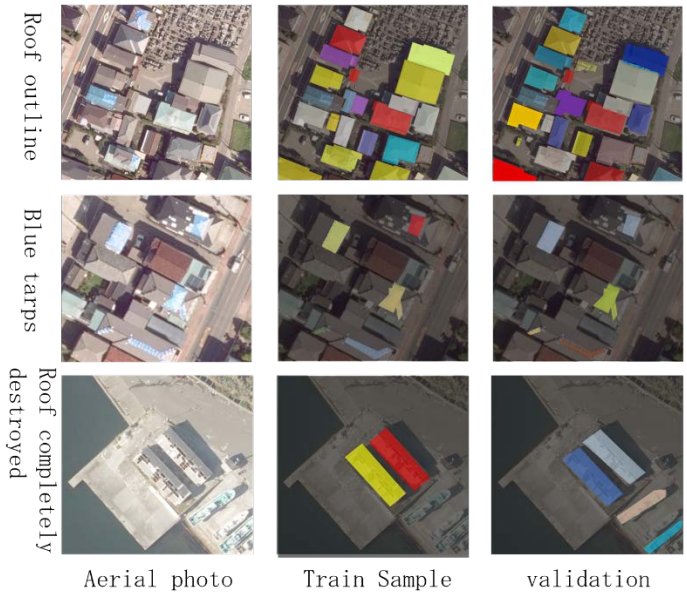


Figure 10. The validation in the training

4.5 Test after training

To test the model after training, k1-k5 area was selected as the test area (Figure. 3), and a confusion matrix [21] was used for evaluation. In the confusion matrix, the detection accuracy of each pixel was calculated separately. The result is shown in Table 4. The unit of the value is pixels. In this study, the pixel statistics of three detection targets were conducted. This is because the three detection target pixels always overlap. For example, blue tarps always lay on the roof, thus, the pixels on the roof always include pixels of blue tarps. Each row of Table 4 represents the results of each detection target, and the total number of each row is 10^8.

Table 4. Results of each detection target

Detected object	TP	FN	FP	TN
Roof outline	10,312,006	1,274,517	1,390,382	87,023,095
Blue tarps	257,733	22,909	28,637	99,690,721
Roofs completely destroyed	213,646	60,259	21,912	99,704,183

In addition, according to the value of the confused matrix, five indices: accuracy, precision, recall, specification prefecture, and F value, were selected, which are expressed in Equations (7) - (11), respectively. Evaluation results of the test results are shown in Table 5.

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \tag{7}$$

$$Precision = \frac{TP}{(TP + FP)} \tag{8}$$

$$Recall = \frac{TP}{(TP + FN)} \tag{9}$$

$$Specificity = \frac{TN}{(TN + FP)} \tag{10}$$

 doi:10.20944/preprints202204.0240.v1

$$F = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{11}$$

Table 5. Evaluation results of each detection target

Detected object	Accuracy	Precision	Recall	Specificity	F value
Root outline	0.973	0.881	0.890	0.984	0.886
Blue tarps	0.999	0.900	0.918	0.999	0.909
Roofs completely destroyed	0.999	0.907	0.780	0.999	0.839

Blue tarps have the highest F value, reaching 0.90 because of the large number of samples of blue tarps that were labeled in Section 4.1. The characteristics of blue tarps itself are simple, there are almost no changes other than shape and angle, resulting in easy detection. The F value of roof outline was 0.89 with accuracy of 0.97. This is considered to be because the use of COCO dataset as a pre-trained model excluded many other interfering objects. Finally, the F value of roofs completely destroyed was 0.84, which is relatively low. This is because the number of samples of roofs completely destroyed was small. Following the occurrence of the typhoon disaster, the number of roofs found to be completely destroyed was lower than the roofs covered with blue tarps. This study used the data augmentation method in Figure 5 to increase 384 samples of roofs completely destroyed to 2304. However, in the actual detection, the error rate was still very high. For example, four houses were missed during inspecting while five houses were mistakenly inspected (four agricultural greenhouses and one abandoned house). Owing to the large area of forest and sea in the study area in this research, there are many pixels that are not objects, and the value of TN of each detection target was very large. Therefore, the accuracy is determined to be higher than the actual perceived value. It is appropriate to use F value for final evaluation.

4.6 Comparison with field investigation

In Figure. 11, the automatic detection results by using deep learning are compared with the results of field investigation [4]. As a result of deep learning, it shows that the number and proportion of roof damage are almost consistent with the field investigation.

In addition, in the field investigation [4], the statistical results of the direction of the damage of roof are shown in Figure 12 (a). the south side of the roof suffered the most. The method adopted in filed investigation is that the investigators stand in one direction to visual inspection. For example, if a roof is found to been damaged in four directions, it will be counted as damaged in the south, east, west and north. For further discussion, this paper refines the calculation method of damaged direction from the results of deep learning. The center of gravity of the damaged roof is divided into 16 directions, and the central distribution of the damaged area is counted. The results are shown in Figure 12 (b). The results were slightly different from the results of the field investigation, and it is found that the southeast suffered the most.

Finally, the trained deep model was used to detect the roof damage in the south of Chiba prefecture (Figure. 3). The performance of PC was the same as that of previous training. Consequently, it was confirmed that it can be completed within 10 minutes without improving the hardware. Within the range of aerial photos in this study, the deep training model automatically detected 48,315 roofs. Among them, 4,838 roofs were damaged, and the rate was approximately 10%.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1

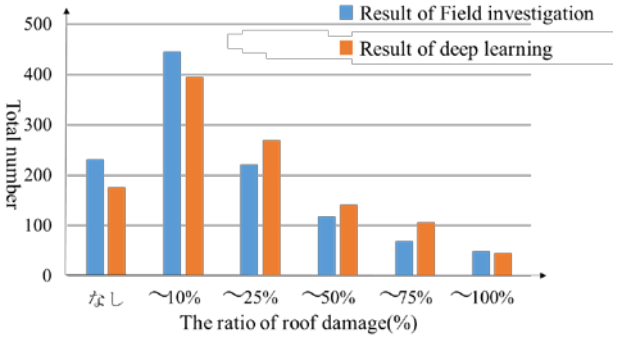


Figure 11. Comparison with field investigation in the proportion of roof damage

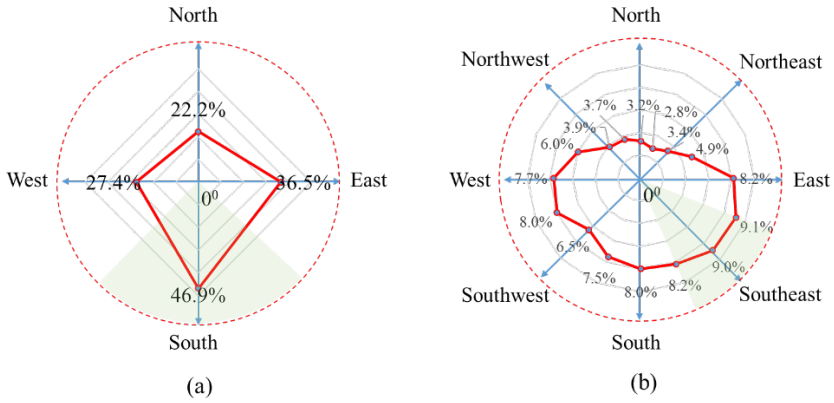


Figure 12. Comparison with field investigation in the location of roof damage

Figure. 13 is the fitting curve of the damaged roof detected through deep learning model. The horizontal axis in the figure is the proportion of the damaged area in the entire roof area, whereas, the black curve is the fitted probability density plot, showing a lognormal distribution. The statistical results show that the roof with damage ratio of 1 - 10% accounts for 44.3%. Further, the roof with a damage ratio of 10 - 30% accounts for 35.4%. Furthermore, the roof with 80 - 100% damage ratio is 3.6%. Thus, it was found that most roofs were partially damaged, and the proportion of roof completely destroyed was very small.

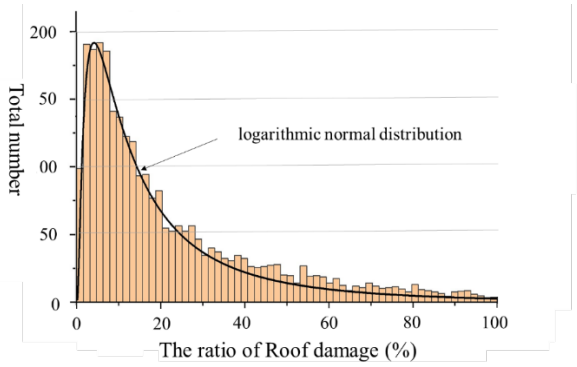


Figure 13. Statistics of detected roof damage

5. Classification of level of roof damage

The ultimate goal was to use the test model to classify the level of roof damage.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1

Figure. 14 shows a flow chart for classification that has been roughly divided into three parts. First is the detection of roof outline, which if not detected is determined to be Non-roof. Thereafter, if the roof completely destroyed is detected. whether blue tarps are detected later or not, it is directly classified as Level 4. Finally, If blue tarps. and roof outline are detected simultaneously in an area, the proportion of blue tarps in this area can be calculated using the formula provided in Figure 14. The levels 1, 2, and 3 were determined according to the proportion. If only roof outline was detected out and no blue tarps were detected out, it was classified as level 0.

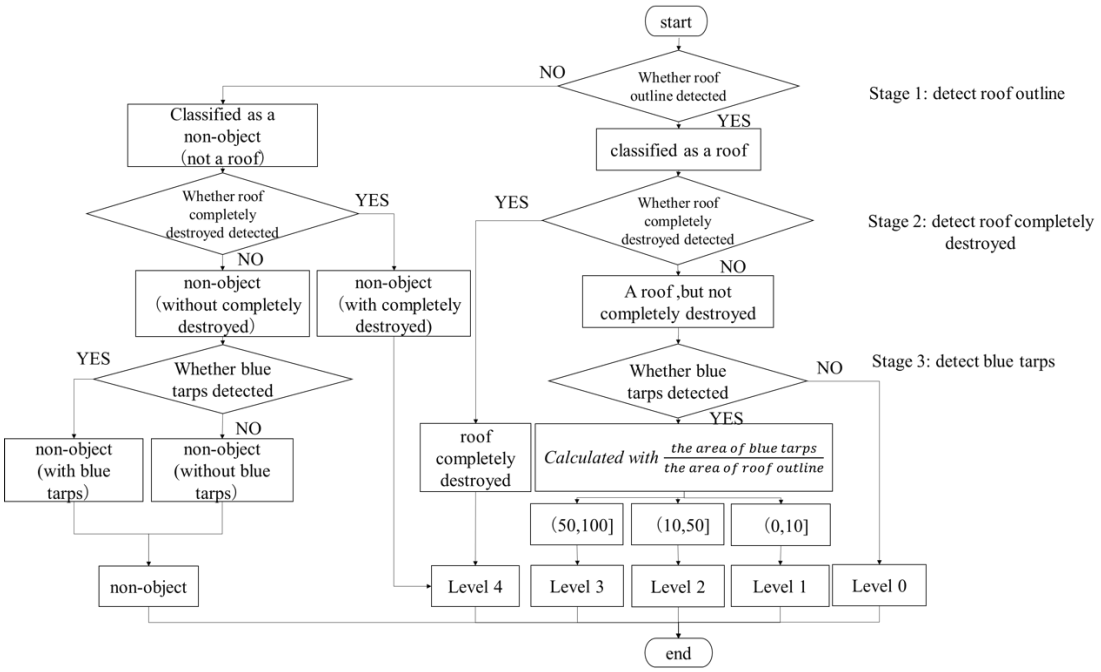


Figure 14. Flowchart of classification of the level of roof damage

Figure. 15 shows the statistical data of each classification stage. The first row is the data counted by the visual inspection of the area surrounded by the green box in Figure. 3. There are 7,649, 375, 318, 62, and 104 roofs of level 0,1, 2, 3, and 4, respectively, as well as several non-roofs. In the detection of contour of roof in the second row, 7,117 houses were correctly detected, 590 houses were missed (most were undamaged roofs, a total of 525 roofs), and 161 houses were wrongly detected (mainly automobiles, agricultural greenhouses, pools, etc.). Further, in the detection of roof completely destroyed in the third row, 82 roofs completely destroyed were correctly detected, 21 were missed, and 18 were wrongly detected (including outdoor warehouse). In the detection of blue tarps in the fourth row, the detection omission and error detection are displayed below the block. In particular, in row 5, owing to the detection accuracy of blue tarps and roof outlines, although certain objects were detected out successfully, the position deviation was considerably large, resulting in the wrong classification. For example, in column B of the row 5, because the area of the detected blue tarps is larger than the actual area, 33 roofs of level 1 are wrongly classified (5 and 29 roofs are wrongly classified as level 3 and 2, respectively). In row 6, the sum of results of final tests is presented. For example, in column C, there are 318 roofs of level 2 of which 247 roofs were correctly classified as level 2, 5 roofs are wrongly classified as level 0, 32 roofs are wrongly classified as level 1, 8 roofs are wrongly classified as level 3, 1 roof is wrongly classified as level 4, and 25 roofs are non-roof.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1

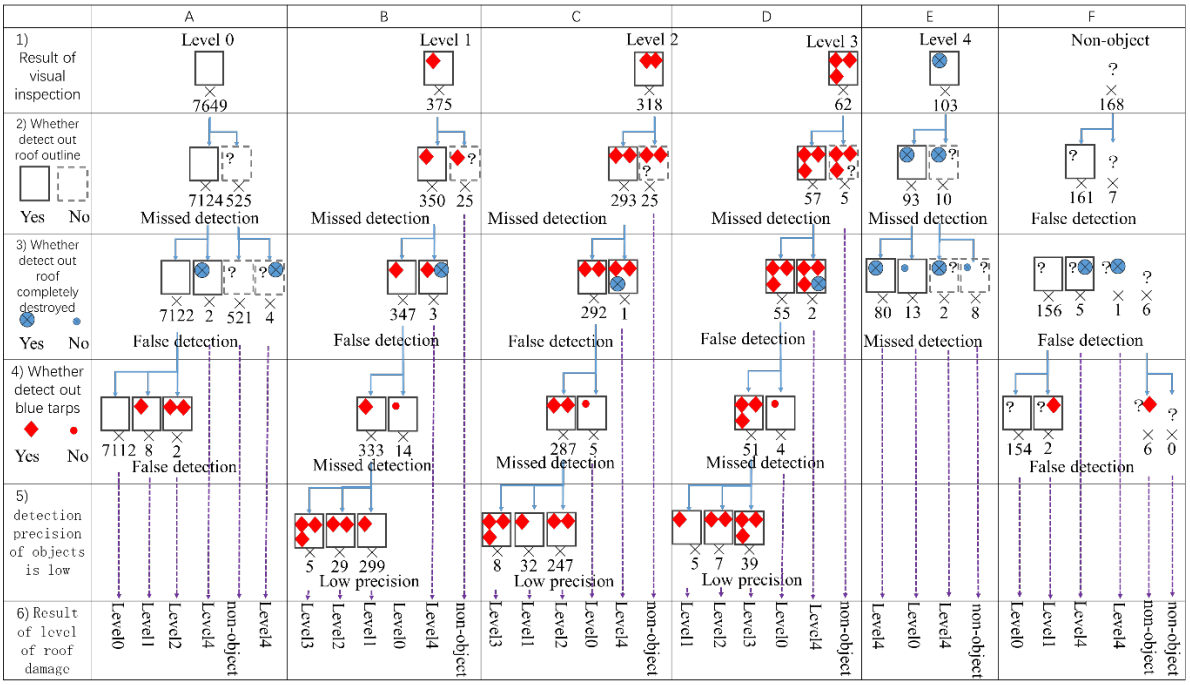


Figure 15. Statistical chart for classification of the level of roof damage

Other detailed statistical results are shown in Table 6. And Figure. 16 the map showing the result of each detected object and the level of roof damage is provided.

In Table 6, the accuracy and F value of the classification of each damage level are calculated separately. Then, the accuracy of each level and the average value of F value were considered as the mean accuracy and F value of classification, respectively. The results show that the overall accuracy was 0.97, and the overall F value was approximately 0.81. This is because the number of levels 0 was considerable at each level. Therefore, the accuracy can be evaluated based on the accuracy rate. However, when the number of different levels varies greatly, the accuracy rate could not be used as a good index. Therefore, the comparison in Chapter 6 mainly compares with other methods using the F value.

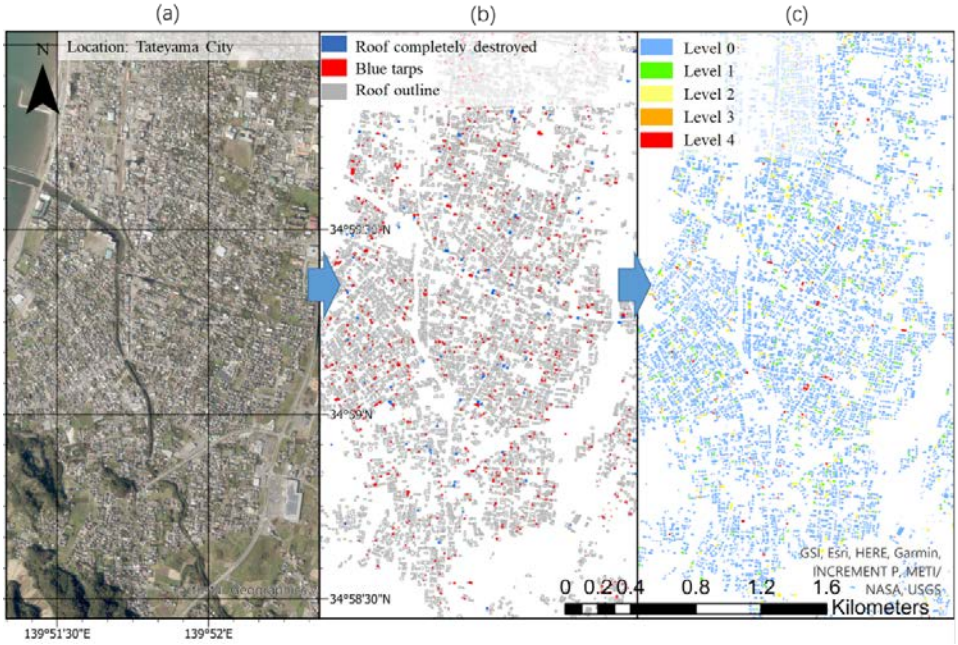


Figure 16. Results of the detection and classification for the green areas in Figure. 3

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1

Table 6. Evaluation of test results using confusion matrix

Classification using visual inspection	Classification using deep learning						
	Level 4	Level 3	Level 2	Level 1	Level 0	non-object	total
Level 4	82	0	0	0	13	8	103
Level 3	2	39	7	5	4	5	62
Level 2	1	8	247	32	5	25	318
Level 1	3	5	29	299	14	25	375
Level 0	6	0	2	8	7,112	521	7,649
non-object	6	0	0	2	154	6	168
total	100	52	285	346	7,302	590	8,675

To inspect the spatial distribution of blue tarps after typhoon Faxai a 250 m mesh was used to render the distribution map according to the benchmark of Bureau of statistics of Japan, and the blue tarps, roof outline, and the ratio of blue tarps were calculated for each mesh. Figure. 17 shows the distribution map, where (a) Is the total area of the detected blue tarps in each grid (sq m). Most of the blue tarps was distributed in Minamiboso City, the central part, near the Mera area and Tateyama city. Further, (b) Is the total area of the detected roof outline (sq m). The outline of the roof is mainly distributed on the coastline of Tateyama City, Mera area and Chikura area. (c) Is the distribution of ratio of blue tarps, and the unit is percentage. Equation (12) is used to calculate the rate of blue tarps in each mesh. In Tateyama City, it is evident that there exist areas with high ratio in the inland, rather than coastline.

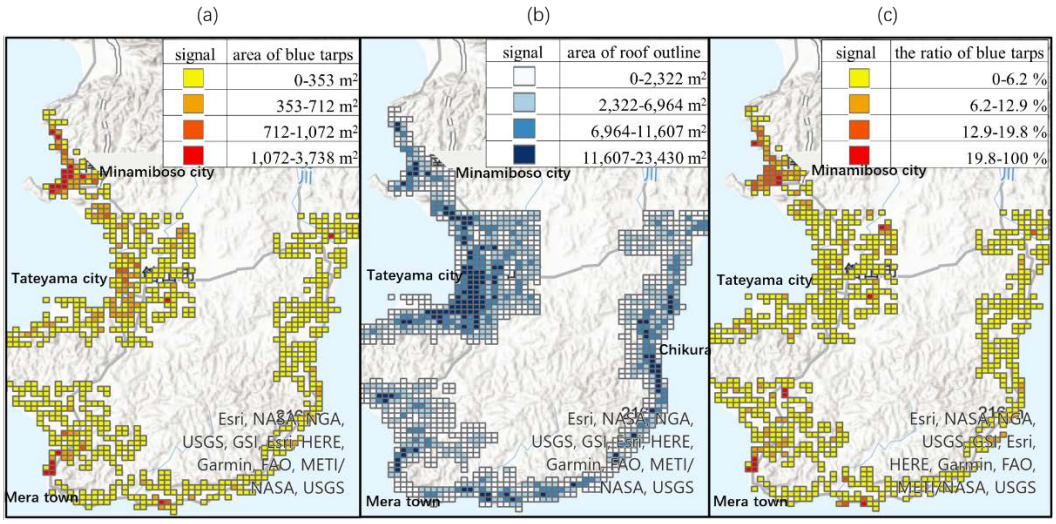


Figure 17. Distribution of blue tarps in southern Chiba Prefecture

6. Comparison with other classification methods

Through the comparison with other methods of damage roof detection the F value of this classification method was found to be good. The comparison studies considered the entire south of Chiba Prefecture after Typhoon Faxai in 2019 as the test area object and aerial photos. As a method of image analysis, it mainly refers to the methods of Liu et al. [26] and Kono et al. [10]. The method of CNN refers to the method of Miura et al.[13]. Table 7 shows the different classification methods and evaluation indexes compared.

In addition to the accuracy (accuracy and F value) determined, the time required for analysis and universality was also investigated.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 April 2022 doi:10.20944/preprints202204.0240.v1

First, although the method of visual inspection offers high accuracy, it is time consuming. The author initially extracted damaged roofs from aerial photos in the southern area of Chiba prefecture through visual inspection, which required more than 2 days. Next, employing the method of image analysis, the spectral characteristics of blue tarps were studied, and different parameters were used to extract blue tarps. This method also requires much time and lacks universality. The method of CNN spends time on the production of the model, while time spent on the automatic detection in the later of the automatically detection is neglected. However, in the comparison of various methods regarding the time, owing to the differences between the algorithms used and computer hardware, particularly, the calculation of the time to prepare the data set and the time to determine the error of which method is the best to not be determined, owing to the specific comparison being difficult.

The accessibility of this study, refers to the convenience with which it can be used by researchers later. In terms of the updatability of model parameters, the training in this study can be regarded as the first training, and subsequent users can take the model as the initial parameters for secondary and further training. Further, with the increase in training time, the accuracy of model parameters is gradually improved. It can provide initial parameters for other researchers' automatic detection research. Therefore, the method in this paper can be said to be accessible.

Thus, there are many difficulties in 100% automatic extraction of damaged roof from aerial photo, primarily owing to three factors. 1) Currently, owing to the different house materials of the roof, there are detection omissions or detection errors in extracting geometric or structural features. 2) The appearance of the roof of general public houses is a regular shape; however, the appearance of the roof of residential house is complex and changeable, with different shapes (such as oval, etc.). 3) Suburban homes are often covered by dense trees, while the lower houses are often covered by the shadow of high-rise houses, thereby rendering the extraction of complete roofs a challenge.

Table 7. Comparison with other classification methods

Compared article	Method of classification	Number of parts	Parts of classification	Accuracy	Average accuracy	F value	Average F value
Noda et al [7]	visual inspection	2 parts	roof without blue tarps	1.000	1.000	1.000	1.000
			roof covered with blue tarps	1.000		1.000	
Kono et al [10]	image analysis	2 parts	roof without blue tarps	0.746	0.800	0.738	0.791
			roof covered with blue tarps	0.929		0.844	
Liu et al [26]	image analysis	3 parts	roof without blue tarps	0.920	0.880	0.935	0.801
			roof partially covered with blue tarps	0.810		0.791	
			roof mostly covered with blue tarps	0.720		0.677	
Miura et al [13]	Deep learning (CNN)	3 parts	roof without damage	0.926	0.937	0.955	0.730
			roof with blue tarps	0.964		0.944	
			roof completely destroyed	0.833		0.291	
this paper	Deep learning (Mask R-CNN)	5 parts	roof without damage	0.916	0.976	0.951	0.818
			roof covered with 0-10% blue tarps	0.986		0.829	
			roof covered with 10-50% blue tarps	0.987		0.819	
			roof covered with 50-100% blue tarps	0.996		0.684	
			roof completely destroyed	0.996		0.808	

7. Summary

In recent years, frequent typhoon disasters have caused great damage to houses. The manner in which to quickly detect damaged roof has become a challenge. The integration of aerial photos and deep learning method in this study can provide a favorable means for the rapid detection of damaged roofs and classification based on level of roof damage

after typhoon. This study considered the residential roof in the southern part of Chiba Prefecture after typhoon Faxai in 2019 as the research area using aerial photo, and the sequence and principle of detection methods were described in detail. Specifically, it included the sample making, the parameter adjustment, the model training and validation. Subsequently, the detection results obtained using the model were compared with results of the field investigation after the typhoon Faxai. It proves the feasibility of this method. In addition, it also summarizes the advantages and disadvantages of this method with the field investigation. The advantage is fast detection in a wide search area, whereas, the disadvantage is still being interfered upon based on certain factors, such as trees and other details. However, with the increase of the data of roof damage and the improvement of deep learning model, the shortcomings of this method can be improved. In the future, this method is expected to be extended to many fields, such as to detect walls damage and other appearance parts damage of house following the occurrence of a typhoon.

Note 1) batch size: the number of data contained in each subset of training is called batch size. For example, when the dataset of 2000 roofs is divided into 100 subsets, the batch size is 20.
Note 2) backbone: the omitted form of backbone network, and its scale and type vary with the network. The backbone of the model in this paper is RESNET 50 and RESNET 101, which shows the number of the convolution layer.
Note 3) learning rate: learning rate is a parameter that adjusts the size of parameter changes in the optimization of deep learning. Excessive learning rate will lead to good results, and too small learning rate will prolong the training time.

Author Contributions: Conceptualization, J.L.; methodology, T.T.; validation, F.Z.; formal analysis, W.L.; investigation, J.L. and T.T.; resources, W.L. and T.T.; data curation, T.T.; writing—original draft preparation, J.L.; writing—review and editing, J.L.; funding acquisition, T.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by JSPS Grant-in-Aid for Special Purposes (2019) G19K24677, a field investigation on the effects of prolonged power outages caused by Typhoon Faxai (Principal Investigator: Yoshihisa Maruyama).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Part of this paper was supported by a field investigation after Typhoon Faxai (Principal Investigator: Yoshihisa Maruyama). Thank to everyone who accompanied us in the field investigation after Typhoon Faxai in 2019. We would like to thank Editage (www.editage.com) for English language editing.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. National Institute for Land and Infrastructure Management (NILIM); Building Research Institute (BRI). Disaster Report: Field Survey Report on Damage to Buildings Due to Strong Winds Caused by Typhoon Faxai in 2019 (Summary). *Building disaster prevention* **2019**, *503*, 57-66.
2. National Institute for Land and Infrastructure Management (NILIM). Field investigation report on damage of buildings by strong wind accompanied with typhoon No.21. http://www.nilim.go.jp/lab/bbg/saigai/ (accessed on 2018.11)
3. Cabinet Office (disaster prevention): disaster guidelines for disaster victims 2021. http://www.bousai.go.jp/taisaku//pdf/r303shishin_all.pdf (accessed on 2021.3)
4. Kazuyoshi, N.; et al. Survey report on damage and recovery after Typhoon Faxai in 2019 Part1-Part4. *Summaries of Technical Papers of Annual Meeting. Architectural Institute of Japan*. **2020,** (in japanese)

5. Ministry of Internal Affairs and Communications, Japan. Result report: Fact-finding survey on the issuance of disaster certificates in the event of a large-scale disaster, Focusing on the 2016 Kumamoto earthquakes. https://www.soumu.go.jp/main_content/000528758.pdf (accessed on 2021.3)
6. Mitomi, H.; et al. detection of the areas with building damage due to the kobe earthquake 1995using airborne MSS images. *The 11th Earthquake Engineering Symposium*. **2012,** 2323-2328. (in japanese)
7. Noda, M.; Tomokiyo, E.; Takeuchi, T. Aerial Survey of Wind Damage by T1821 in South Osaka and North Wakayama. *Summaries to Technical Papers of Annual Meeting, Japan Association for Wind Engineering (JAWE)* **2019**, *2019,* 105-106.
8. Suzuki, K.; Hanada, D.; Yamazaki, F. Building Damage Inspection of the 2012 Tsukuba Tornado from Field Survey and Aerial Photographs. *Institute of Social Safety Science* **2013**, *21*, 9-16.
9. Yuya, K.; Kazuyoshi, N.; Takashi, T.; Eriko, T.; Hiroshi, N. Summaries of Design Works of Annual Meeting. *Architectural Institute of Japan* **2019,** *1*, 91-92.
10. Kono, Y.; Nishijima, K. Analysis of roof damage on residential house with aerial photo. Summaries of Design Works of Annual Meeting. *Architectural Institute of Japan* **2019,** *2*, 89-99.
11. Shohei, N.; et al. Building damage detection of the kumamoto earthquake utilizing image analyzing methods with aerial photographs. *Journal of Japan Society of Civil Engineers* **2019**, 75, 218-237.
12. Kurihara, S. *The theory of the future of life, industry and society from the viewpoint of the practical use of AI revolution with human being.* NTS press. Tokyo, Japan, 2019; 49-113. (in japanese)
13. Hiroyuki, M.; Aridome, T.; Matsuoka, M. Deep Learning-based Identification of Collapsed, Non-collapsed and Blue Tarp-Covered Buildings from Post-Disaster Aerial photo. *Remote Sens*. **2020**, *12*, 100-108.
14. Ministry of Internal Affairs and Communications, Japan: About the damage situation caused by Typhoon Faxai in 2019. https://www.soumu.go.jp/r01_taifudai15gokanrenjoho/ (accessed on 2021.3)
15. Kaiming, H.; Gkioxari, G.; Dollar, P.; Girshick, R. Mask R-CNN. In Proceedings of IEEE International Conference on Computer Vision, Venice, Italy, 22-29 Oct. 2017, 2961-2969.
16. Girshick, R.; Donahueet, J.; Darrell, T.; et al. Region-based Convolutional Networks for Accurate Object Detection and Segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *38*, 142-158.
17. Lin, T.-Y.; Dollar, P.; Girshick, R.; et al. Feature Pyramid Networks for Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21-26 July 2017, 936-944.
18. Ross, R. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7-13 Dec. 2015, 1440-1448.
19. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans Pattern Anal. Mach. Intell.* **2016**, *39*, 1137-1149.
20. Paszke, A.; Gross, S.; Massa, F.; et al. PyTorch, An Imperative Style, High-Performance Deep Learning Library. *Adv. Neural Inf. Process. Syst*. **2019**, *32*, 8026-8037.
21. Lin, T.Y.; Maire, M.; Belongie, S.; et al. Microsoft COCO: Common Objects in Context. In Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science; Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T., Eds.; Springer, Cham, 2014; Volume 8693, pp. 740-755.
22. Chida, H.; Takahashi, N. Study on image diagnosis of timber houses damaged by earthquake using deep learning. *The Architectural Institute of Japan's Journal of Structural and Construction Engineering*. **2020,** *85*, 529-538.
23. Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.O.; et al. A review on deep learning techniques applied to semantic segmentation. *Comput. Vis. Pattern Recog*. **2017**.
24. Kanai, K.; Yamane, T.; Ishiguro, S.; Chun, P.-J. Automatic detection of slope failure regions using semantic segmentation. *Intell. Inform. Infra*. **2020**, *1*, 421-428.
25. Statistics bureau of Japan. About regional mesh statistics. https://www.stat.go.jp/data/mesh (accessed on 2020.12)
26. Liu, W.; Maruyama, Y. Damage estimation of buildings' roofs due to the 2019 typhoon Faxai using the post-event aerial photographs. *J. Jpn. Soc. Civ. Eng.* **2020**, *76*, 166-176.