TITLE:

# Estimation Method for Roof‐damaged Buildings from Aero-Photo Images During Earthquakes Using Deep Learning

AUTHOR(S):

Fujita, Shono; Hatayama, Michinori

京都大学
KYOTO UNIVERSITY

KURENAI 紅
京都大学学術情報リポジトリ
Kyoto University Research Information Repository

Check for updates

# Estimation Method for Roof-damaged Buildings from Aero-Photo Images During Earthquakes Using Deep Learning

Shono Fujita[1] · Michinori Hatayama[2]

## Abstract

Issuing a disaster certificate, which is used to decide the contents of a victim's support, requires accuracy and rapidity. However, in Japan at large, issuing of damage certificates has taken a long time in past earthquake disasters. Hence, the government needs a more efficient mechanism for issuing damage certificates. This study developed an estimation system of roof-damaged buildings to obtain an overview of earthquake damage based on aero-photo images using deep learning. To provide speedy estimation, this system utilized the trimming algorithm, which automatically generates roof image data using the location information of building polygons on GIS (Geographic Information System). Consequently, the proposed system can estimate, if a house is covered with a blue sheet with 97.57 % accuracy and also detect whether a house is damaged, with 93.51 % accuracy. It would therefore be worth considering the development of an image recognition model and a method of collecting aero-photo data to operate this system during a real earthquake.

**Keywords** Damage certification · Deep learning · Image recognition · Aero photo · GIS

## 1 Introduction

In Japan, when disasters, such as earthquake, flood, and wind damages occur, local governments investigate and certify the damage level of residents' buildings. Afterwards, the local governments issue damage certificates to the residents to prove that they are the victims of the disaster. The damage certificate is used as the basis for deciding the victims' support contents, for example, the price of support money, the order of priority of temporal houses corresponding to their buildings' damage level, and the necessity of livelihood rehabilitation. In the investigation, the buildings are divided into five damage levels like Fig. 1. In general, first, visual observations are performed outside the building, and second, detailed observations are performed inside the building. The latter is

performed if a victim disagrees with the former (Cabinet Office, 2013). Both of the investigation, at first, judge whether a building is complete collapse by the outside appearance and the inclination. When the building is not complete collapse, the investigators calculate the proportion of damage on the roof, the outer walls, and the foundations. Judging from the proportion, the building is divided into five damage levels.

After the Basic Act on Disaster Management in Japan was revised based on the information gathered from the Great East Japan Earthquake, local governments were obliged to issue damage certificates without any delays. Hence, they now must issue these certificates rapidly and exactly (Cabinet office in Japan, 2017). However, in Japan at large, house damage investigations and issuing of damage certificates take a long time in even now. Consequently, supports to victims of the disasters are delayed. According to a survey conducted after the 2016 Kumamoto Earthquake (Ministry of Internal Affairs and Communications, 2018), labor shortage, complexity of investigation, numerous reinvestigations, and difficulties of arranging a working environment, were the causes for these delays.

In such cases, the number of resources for disasters become limited and the investigation is not done adequately and efficiently.

A previous study (Inoue et al., 2018) analyzed work volume of disaster response in local governments based on the

✉ Shono Fujita
fujita.shono.32x@st.kyoto-u.ac.jp

Michinori Hatayama
hatayama@dimsis.dpri.kyoto-u.ac.jp

[1] Graduate School of Informatics, Kyoto University, 36-1, Sakyo-Ku, Kyoto, Kyoto 606-8501, Japan

[2] Disaster Prevention Research Institute, Kyoto University, Gokasho, Uji, Kyoto 611-0011, Japan

⧉ Springer

京都大学
KYOTO UNIVERSITY

352

Inf Syst Front (2023) 25:351–363

KURENAI 紅
京都大学学術情報リポジトリ
Kyoto University Research Information Repository

Kumamoto earthquake. The study defined the disaster response work based on the guideline of Japanese Cabinet office and classified the work into two types: maintenance type and achievement type. Maintenance type is the work requiring a fixed number of personnel for management. Achievement type is the work finished by achievement of specific objective. By increasing the number of personnel, the ending time of achievement type can be short, but the ending time of maintenance type cannot be short. In achievement type of the most seriously damaged local government, the work which require the most personnel is issuance of damage certificates, then house damage investigations. From this result, it is assumed that performing the two works rapidly can lead to supplement of personnel to other works.

Furthermore, the necessity of rapid damage investigations and the issuance of damage certificates increases owing to a forecast of large scale earthquake occurrences in Japan in the near future.
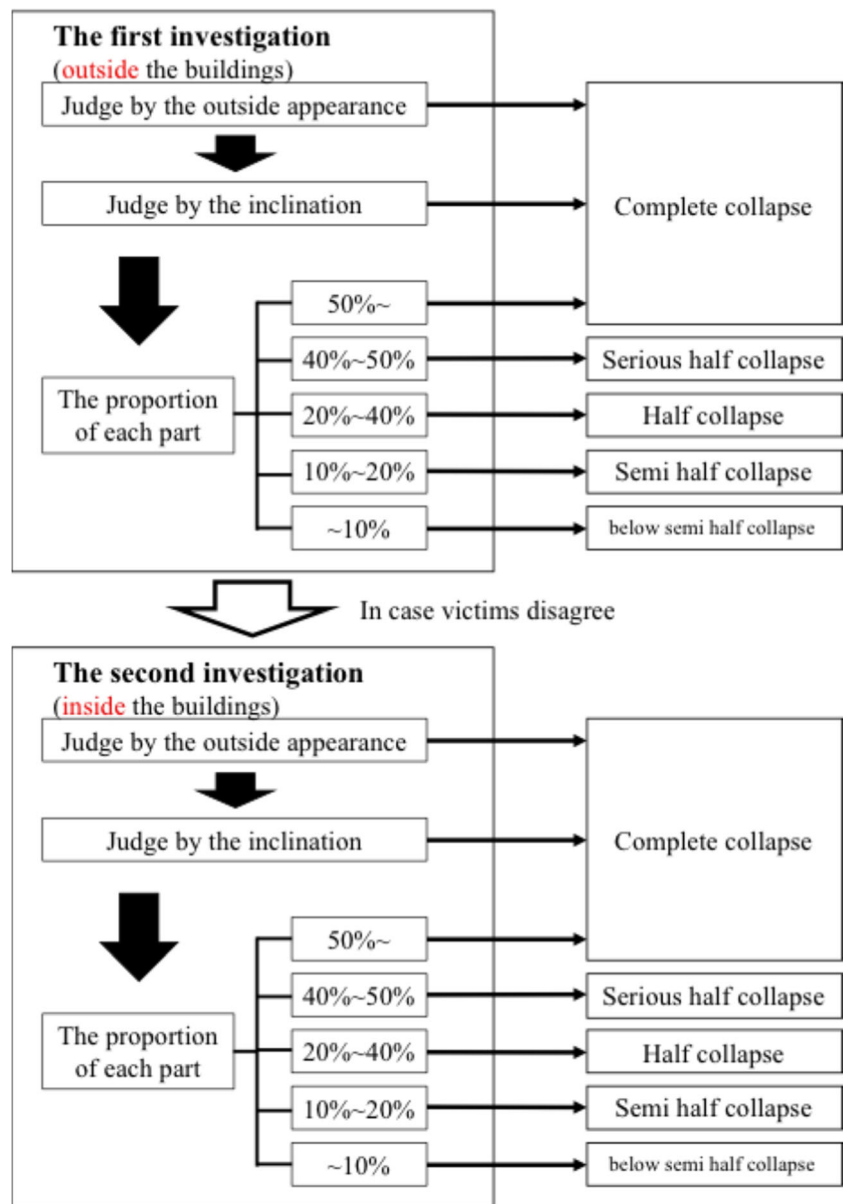
## 2 Research Background and Purpose

### 2.1 Research Background

#### 2.1.1 Revision of House Damage Investigation Manual

In the 2016 Kumamoto Earthquake, damage certificate issuance took a long time, and this led to a long delay in getting support to the victims. After the occurrence

**Fig. 1** Flow of house damage investigation

of this earthquake, the Cabinet Office in Japan revised the house damage investigation manual (Cabinet Office, 2018). The revision enabled local governments to make the necessary decisions based on the aero-photo images or the pictures taken by the victims. This means that the Japanese government requires a more efficient investigation mechanism and also recommends investigation using aero-photo images.

When the Northern Osaka's earthquake occurred in 2018, local governments determined the damage levels of the victims, based on the pictures taken by the victims. Moreover, when the Hokkaido Eastern Iburi Earthquake took place in 2018, local governments certified the damage level of the buildings based on the aero-photo images, under the consideration that the buildings were completely destroyed.

### 2.1.2 Roof Damage as a Result of Earthquakes

When earthquakes occur, many roofs get damaged, especially the tile roofs, in Japan. The tile roof is a Japanese traditional roof style, and it is vulnerable to the shakings of the earthquake. A previous study (Shimizu et al., 2019) analyzed the percentage of tile roof in each built date and each prefecture in Japan. Although the overall percentage in Japan is declining, the average of the percentage of each prefecture in current date (2011-) is 60.37 %. Considering that many Japanese buildings have tile roofs, counting the number of damaged roofs helps in the understanding of the overall damage of the buildings in the earthquake region.

Currently, observing the damage from the outside is the only way to investigate the damage level of the roofs. This method of investigation is difficult and has low accuracy of evaluation in the investigation of the building like Fig. 2.

### 2.1.3 Recent Improvements in Image Recognition Techniques

AlexNet utilizing DCNN (Deep Convolutional Neural Network) won the 2012 ILSVRC (IMAGENET Large Scale Visual Recognition Challenge 2012), which is a world image recognition competition. DCNN is one of the deep learning algorithms. After that competition, DCNN became the basic image recognition way and the image recognition accuracy became higher than other algorithms. In view of the development of large-scale image datasets, such as ImageNet (Deng et al., 2009) and the improvement of computers calculation capacity, such as GPU (Sharan Chetlur et al. 2014), image recognition techniques using deep learning is expected to lead to significant developments.

### 2.2 Research Purpose

The purpose of this study is to ensure good efficiency and rapidity in house damage investigations using aero-photo images and image recognition. Furthermore, if the proposed system can estimate the damaged roofs of buildings based on aero-photo images using DCNN as the image recognition algorithm, during an earthquake, local governments can obtain a clear understanding of the approximate number of houses that are damaged, and the regions with numerous damaged houses. Using image recognition enable to estimate many damage buildings automatically and rapidly even when there is little manpower during disaster. Hence, local governments can estimate the time required for the investigation, number of investigators, instruments and working places. Consequently, the estimations performed in this study can be used for planning house damage investigations in the future and also lead to making house damage investigations and issuing of damage certificates more rapid.

**Fig. 2** A building whose roof is damaged

## 3 Previous Studies

Some studies on the estimation of buildings with damaged roofs, using deep learning or aero-photo images, have already been performed.

Ogawa and Yamazaki (2000) visually determined damage buildings using aero photo images taken in the 1995 Kobe Earthquake. They concluded that damage of non-wooden buildings was difficult to be recognized using aero photo images at that time, but using aero photo images was effective tool for detecting overall damage distribution in a large area after natural disasters. However, this manual system needs a long time and many hands for detection and isn't practical during disaster response.

Suppasri et al. (2013) estimated probability of damaged building using fragility functions. Fragility functions lead probability of damaged building from input of some factors such as peak ground acceleration, types of building and so on. This fragility function could estimate damage buildings with high accuracy, but this estimation is not by building but by area.

Kashani and Graettinger (2015) detected damage buildings using 3D data and k-means algorithm of clustering method. Getting minute 3D data from ground-based laser scanning led to high accuracy. Matsuoka et al. (2017) determined whether a roof was damaged or not based on the temperature recorded by an UAV, using an infrared camera. In roof part under deformations owing to the shakings associated with earthquakes, foundation material under tile roof is exposed. This part is difficult to absorb sunlight and decrease temperature. The low temperature part was considered as damaging part. Kamagatani and Matsuoka (2017) used both aero-photo images taken from right above and oblique direction by a special camera and estimated house damage level judging from these images. Using two types of aero-photo images make the extraction of roof damaging features easy and also increase accuracy.

Our study use not special data such as 3D data or infrared camera data but aircraft camera data or drone data, which are cheaper and easier to take photos with, thereby ensuring that a large number of photo images can be taken with great ease, during an actual disaster. This also ensures that there is sufficient training data for our system. Additionally, although the damage state of roofs may vary over time owing to aftershocks, using aero-photo images by drones enable us to obtain damaged roof-photo data at any point in time.

In the abovementioned study, when the roof database of photo images was constructed for training and testing, they depicted each polygon fitting building (roof) shape manually. In this study, after getting aero-photo images and locating building shape polygons information from GIS, the trimming algorithm makes the database of roof images automatically. Furthermore, this algorithm enables us to get a lot of images for training data and reduces database making time after a disaster occurs. Additionally, this algorithm enables the proposed system to estimate damage level with respect to a building and not a region.

Ji et al. (2019) detected damage buildings from pre- and post-event very high-resolution (VHR) remote sensing imagery using deep learning. They could get high accuracy, but their system needs pre-event imagery data which may not exist in the damage area. This means that their system may not be able to validate in large area. Our system doesn't need pre-event aero photo images and it is assumed that Our system can be used in large area.

Hezaveh et al. (2017) extracted roof damage from aero photo images using deep learning in hail disaster. Their study showed that deep learning can estimate roof damage in hail disaster with high accuracy. Nex et al. (2019) estimate damage area from aero photo images using deep learning in earthquake disaster. Their study showed that deep learning can extract damage part of aero photo images in earthquake disaster.

Considering these above studies, our study uses deep learning, aero photo images and trimming algorithm to estimate damage buildings because using these methods can obtain an overview of earthquake damage in large area by building automatically and rapidly.

From the viewpoint of the damage certificate issuance, a previous study has developed and validated the "Damage Assessment Training System" in Niigata Prefecture Chuetsu Earthquake, 2004, to unify the criterion of investigators with prior training and ensure fairness (Horie et al., 2005). Moreover, Murakami et al. (2012) analyzed the issuance of damage certification work conducted in Uji city of Kyoto prefecture to examine a rapid and an efficient method of the work. Fujiu et al. developed database system using GIS to make building damage assessment more rapid (Fujiu et al., 2013). They showed that their system could assess their damage levels using uploaded photos of damaged house taken by residents or volunteer fire corps in damaged area. Although these studies approach a more rapid issuance of damage certification work from the viewpoint of the investigators training, operation flow or getting data, no study has approached the problem from aero-photo images.

## 4 Estimation System of Roof-damaged Buildings

### 4.1 Structure of the System

Figure 3 shows the structure of the system in this study. First, the system performs the trimming algorithm from aero-photo images and locates information on building polygons in GIS to make a roof-image database. Considering the

characteristics of a roof damage in an area, a part of the roof-image database is used for deep learning training data. Second, each label, whether the roof is damaged or not, is put to each image in the training data manually. This system uses supervised learning as image recognition model. Next, the system makes a DCNN model learn feature of damage roofs from the roof-image database and the label for training. Finally, the system enables the DCNN model to estimate damaged roofs from the remaining database and also estimate damaged buildings.

A few days after an earthquake occurs, many victims cover the damaged part of their roofs with blue sheets. Consequently, if it takes a long time to take aero photos or aero photos are taken again by aftershocks, many roofs are expected to be covered with blue sheets. In this case, the system must determine the damaged roofs based on the number of blue sheets present. Hence, the system involves the use of two DCNN models, one for estimating the damaged roofs, and another for estimating the blue sheets.

## 4.2 Application of the Proposed System

In this system, the estimation result is used in three ways to ensure that the house damage investigations are rapid.

(1) The first way is to obtain an overview of damaged buildings in the region, the required time for investigation and the number of investigators; instruments; and working places. In view of the fact that available resource is limited owing to an earthquake, it is necessary for local governments to obtain an overall damage overview and also make the disaster-response plan rapidly and effectively. In the current investigation, there are two ways to determine buildings which are targets for investigation like Table 1. Local governments select one of the two ways or both for each region. The appropriate way is

selected depending on the damage scale of the region because these ways have demerit and demerit respectively as in the Table 1. Consequently, the estimation of damaged buildings in the region is expected to enable local governments to select the appropriate way for each region and investigate more efficiently.

(2) The second way is to enable investigated buildings to efficiently make a round investigation. If this system can estimate and determine damaged buildings which need investigation before the application of damage certificates from victims, investigators need not include round buildings in the order of application. They can investigate in an order which considers the geographical condition of buildings efficiently. Additionally, they need not to investigate undamaged buildings, thereby reducing the number of unnecessary investigations. Therefore, it is assumed that this way of investigation using our system has both merits of the two current ways in Table 1.

(3) The third way is to directly use each roof estimation for damage score in house damage investigation. Although investigators must sometimes calculate roof damage proportion in the house damage investigation, they have no choice than to look up from the outside. Furthermore, it is assumed that the automatic calculation of the damage proportion from aero-photo images helps to investigate house damage more rapidly and exactly.

Although the third way can lead more rapid and accurate investigation, this study deals with both the first and the second ways at first. Because the third way needs more high accuracy and, fine classification or regression model, the way is not covered by this study. We consider the third way as ultimate goal.

This system cannot estimate buildings with completely crushed first floors and undamaged roofs. In Japan, quick
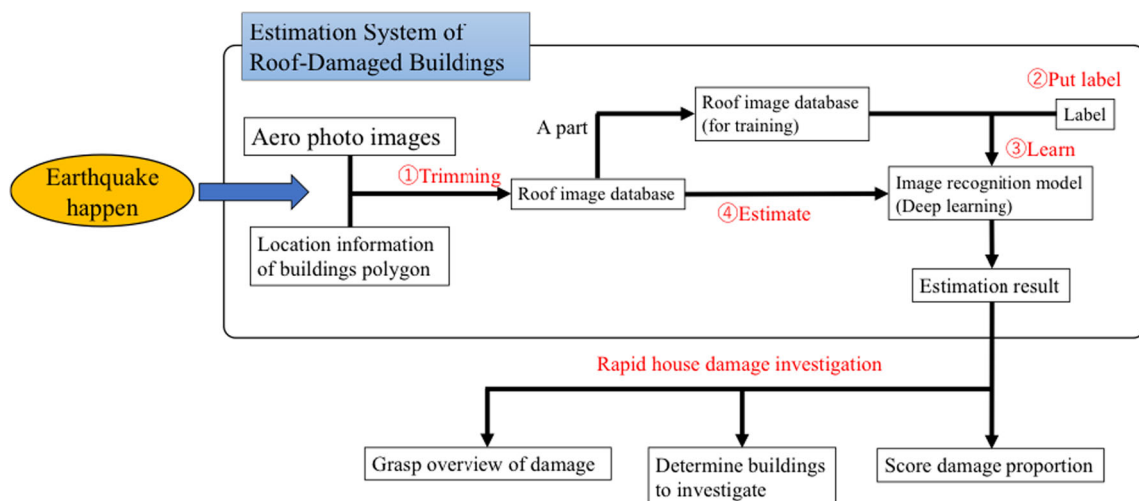


**Fig. 3** Structure of the proposed system

**Table 1** Two methods of investigating buildings

|  | Method | Suitable case | Merit | Demerit |
|---|---|---|---|---|
| All buildings in the area | Investigate→Apply→Issue | Severe damaged area | Investigators can make a round efficiently in the large earthquake. | Undamaged buildings must be investigated. The number of investigations increases. |
| Only applied buildings | Apply→Investigate→Issue | Little damaged area | Investigators can only damaged buildings. | Investigation may not be executed in the order adjoined geographically and efficiently. It is difficult to predict the number of investigated buildings and make plans of obtaining staff. |

post-earthquake inspections of damaged buildings are implemented to prevent secondary disasters (for example, buildings collapse and components falling by aftershock) and protect people safely. Hence, quick post-earthquake inspections of damaged buildings are implemented earlier and less exactly than the house damage investigation. As a result, this study's system does not have to estimate buildings with completely crushed first floors and undamaged roofs.

### 4.3 Trimming Algorithm

In this system, the trimming algorithm constructs the roof-image database automatically, with the use of aero-photo images and location information of building polygon vertex, similar to Fig. 4. These polygons can be gotten from some map such as Open Street Map.

First, the algorithm obtains both maximum and minimum values of longitude and latitude of the building polygon vertex from GIS and makes a circumscribed quadrangle. Next, the algorithm determines whether each pixel in the circumscribed quadrangle is in the building polygon. This judgement can be understood from counting crossing points of building polygon lines and lines drawn from the target point to the right end point like Fig. 5. If the number of cross points is even, the target point is out of the building polygon. In contrast, if the number is odd, the target point is inside the building polygon. This algorithm paints pixels outside building polygons with black and pixels inside with the color as it is.

This algorithm enables the reduction of learning and estimation time because time of making manually image data

where each building is reflected can be left out. Moreover, the system can estimate building damage one by one because of using building polygons. Additionally, removing parts except buildings in photo images is expected to prevent mistaken learning and estimation for deep learning. This study uses the trimming algorithm considering operation in an actual earthquake.

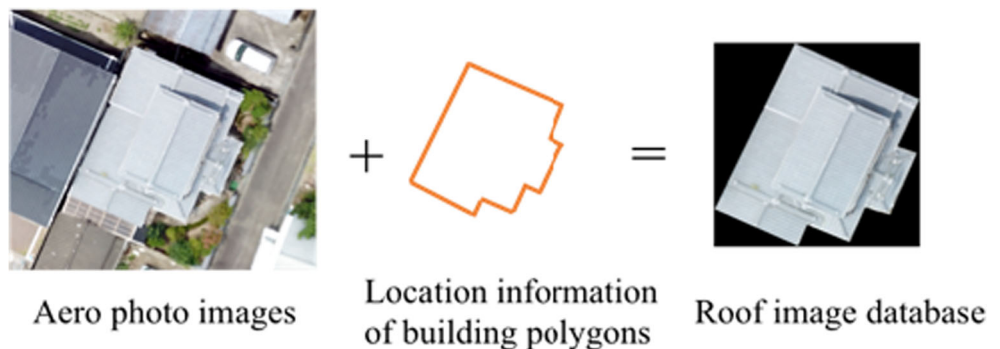### 4.4 Necessary Data in an Earthquake

In an earthquake, the system needs aero-photo images, location information of building polygon in the damaged area and label for deep learning as input data. However, because these are existing data, the system doesn't have to add new data and systems to get data. The system can be utilized with only existing data or system gathering data.

#### 4.4.1 Aero-Photo Images

After a disaster happens, the Geospatial Information Authority (GIA) of Japan takes aero-photo images of the damaged area as important data for initial response, understanding the entire damaged state, recovery and reconstruction (Geospatial Information Authority of Japan, 2020). People can get these aero-photo images for the purpose of disaster response or research.

The disaster rescue team "DRONE BIRD", which incorporated a nonprofit organization "Crisis Mappers Japan" administers, takes aero-photo images with drones having the latest technology for rescue or support activity in natural disasters or
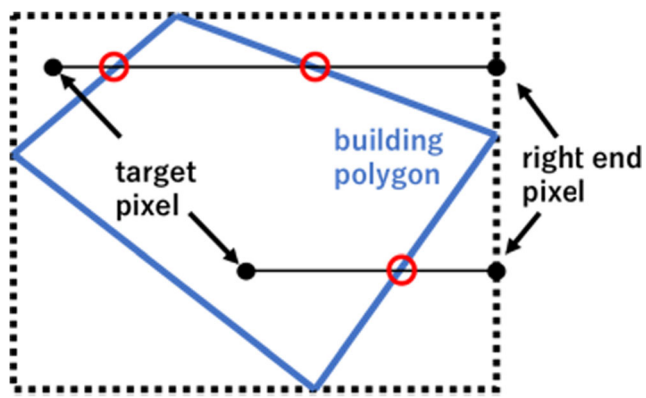
**Fig. 4** Trimming algorithm



Aero photo images     +     Location information of building polygons     =     Roof image database

**Fig. 5** Judging pixel



**Fig. 7** Breakdown of data in cross validation

political riots (DRONE BIRD, 2020). In view of the fact that their aero-photo images are uploaded on a website, people can download the images freely (Open Aerial Map, 2020).

Although aero-photo images taken by the GIA have lower resolution, they have a wider range area compared to the ones taken by the DRONE BIRD. The images have more data volume because of the wide range area and therefore cannot be uploaded on a website. In contrast, though aero-photo images taken by the DRONE BIRD have higher resolution, they have smaller range area compared to the ones taken by the GIA. Consequently, images have less data volume and can therefore be uploaded on a website and downloaded from a distant place immediately after an earthquake.

While the GIA aircraft images have 20 cm resolution, the drone images have several centimeters resolution. To detect tile roof damage, which is likely to be caused by an earthquake, the resolution requires at least 20-30 cm, which is the size of one piece of tile. Hence, it is assumed that the above aero-photo images have minimum necessary resolution.

### 4.4.2 Location Information of Building Polygons

Crisis Mappers Japan rapidly provides maps of damaged area all over the world when natural disasters or political riots arise
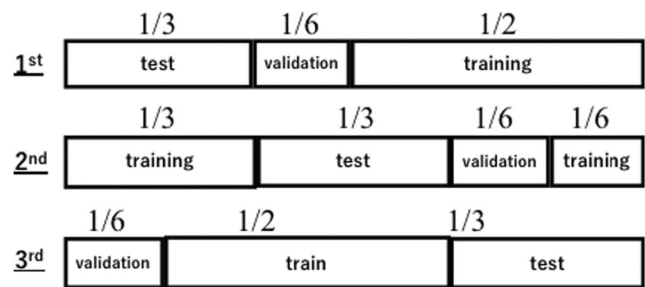
(Crisis Mappers Japan, 2020). This organization enables worldwide people to participate in drawing maps on the Internet using an Open Street Map (OSM). In view of the fact that geographic information in OSM is available to everyone, this system can use the location information of building polygons in the map (Kevin Curran et al. 2012). Although the building location condition change every day because of construction and destruction, maps drawn anew after a disaster provide the newest information of residence and deal with the gap owing to crustal movement by earthquake.

The accuracy of building polygons in an OSM depends on skills or the prudence of drawing people. When we used approximately 260 buildings in one area of Osaka prefecture and calculated the overlap proportion between OSM and basic map information of the Japan Geographical Survey Institute, the proportion was 59.62 %. This overlap proportion was the ratio of intersection area to union area between two maps. Currently aero-photo images, such as Bing-aero photos are used as background pictures when the OSM in a damaged region is drawn. However, these aero-photo images sometimes have low resolution or were taken a few years ago. Therefore, it is difficult to draw exact building polygons. Subsequently, it is assumed that this difficulty led to the reduction of the overlap proportion. Hence forth, if aero-photo images taken after an earthquake occurs are used as background pictures, OSM is expected to be more exact and reflect newest information.
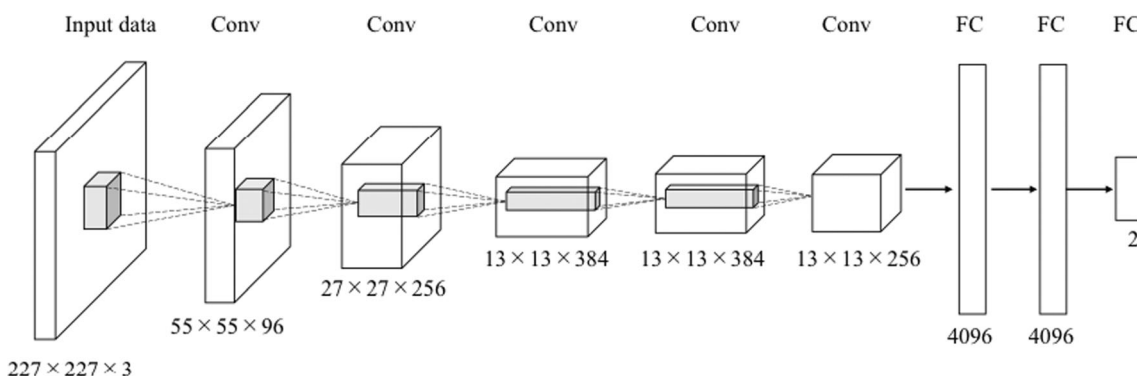


**Fig. 6** Structure of AlexNet

**Fig. 8** The gap between buildings and building polygons of OSM

### 4.4.3 Label for Deep Learning

Deep learning needs labeling, whether a roof is damaged or not, for training data. The labelling can be inputted by staffs of local governments or crowdsourcing. In recent years, volunteers input data online for information support activity to make and use geospatial information, which is necessary for disaster response like the above Crisis Mappers Japan or "National Network for Emergency Mapping" (National Network for Emergency Mapping, 2020).

## 5 Experiments of DCNN Model

### 5.1 DCNN Model

In this study, we used AlexNet as the DCNN model, which won 2012 ILSVRC (world image recognition competition) and the first deep learning model that exceeded the conventional machine learning models (Krizhevsky et al., 2012) accuracy. In view of the fact that models in the ILSVRC discriminated one thousand kinds of objects, it is assumed that AlexNet is a model for general use. Generally, increasing the number of layers in deep learning leads to the enrichment of the expressive power; however, the accuracy eventually reaches its peak. Consequently, volume of calculation

increases and the learning time becomes longer. Therefore, if we use the more complicated model that won the latest ILSVRC (Zeiler & Fergus, 2014; Szegedy et al., 2015; He et al., 2016; Hu et al., 2017), it may have high accuracy but longer time, which is insignificant in this study. Considering the above, in this study we use the Alexnet, which has the fewest layers and is the simplest among the models which won the ILSVRC. Figure 6 shows the structure of the Alexnet in reference to original paper (Krizhevsky et al., 2012). In Fig. 6, conv represents a convolutional layer and fc represents a full connection layer. We implemented AlexNet by python and tensorflow.

### 5.2 Learning Way

We used three datasets: training data to optimize parameters, validation data to determine training time and test data to evaluate the model. We did cross validation which changed these dataset three times for exact evaluation like that in Fig. 7. Furthermore, we did mini batch training with batch size 128 and changed the order of data in every epoch in the batch to reduce deviation of data in the mini batch. We used momentum optimizer, applying weight decay with $10^{-5}$ learning rate. Generally, excessive training time may lead to overfitting (Prechelt 1998). To restrain the overfitting, we confirmed the decline in the accuracy of the validation data and stopped the training when overfitting began.

In an actual earthquake, the location accuracy of aero-photo may not be exact and there may also be difference in mapping skills between mapped people. Hence, it is difficult to overlap actual buildings in aero-photo images and building polygons in an OSM completely and also there is expected to be somewhat gaps between them. From the above, although there are somewhat gaps in the data we used as building polygons like that in Fig. 8, we didn't revise the gaps and proceeded to train the data as it was.

### 5.3 Evaluation Criterion

In many damaged regions as a result of an earthquake, it is expected that there are more undamaged buildings than damaged ones. There is a gap between the number of data in two classes for estimation. If we use accuracy (1) as the evaluation criterion, estimating many data, as no damaged roof, increases accuracy easily. This means that the system cannot obtain an exact overview of damaged buildings in 4.2 (1) and may overlook the damaged buildings in 4.2 (2). Similarly, if we use recall (2) as the evaluation criterion, estimating many data, as damaged roof, increases recall easily. This means that the system cannot obtain an exact overview of damaged buildings in 4.2.(1) and may increase unnecessary investigations of undamaged buildings in 4.2 (2). To deal with imbalance between the number of data in the two classes, in this study, we used

**Table 2** Confusion matrix

| | | Actual | |
|---|---|---|---|
| | | Damage | No damage |
| Estimate | Damage | TP | FP |
| | No damage | FN | TN |

**Table 3**  The number of data

|  | Blue sheet | No blue sheet | Damage | No damage |
|---|---|---|---|---|
| Training data | 4416(276×16) | 4416(2208×2) | 7920(495×16) | 7920(3960×2) |
| Validation data | 1472(92×16) | 1472(736×2) | 2640(165×16) | 2640(1320×2) |
| Test data | 2944(184×16) | 2944(1472×2) | 5280(330×16) | 5280(2640×2) |

balanced accuracy (4), which does not depend on the proportion of data in the classes. Balanced accuracy is the mean of the recall and the specificity (3). TP, TN, FP and TP in the equations below are referenced by the confusion matrix in Table 2.

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN} \qquad (1)$$

$$recall = \frac{TP}{TP + FN} \qquad (2)$$

$$specificity = \frac{TN}{FP + TN} \qquad (3)$$

$$balanced accuracy = \frac{recall + specificity}{2} \qquad (4)$$

However, in the present experiment, we used the same number of test data in the two classes. Therefore, in this study, balanced accuracy equals accuracy.

Apparently, there is no clearly defined accuracy which is required in the actual work of the house damage investigation. Therefore, we set up a 90 % balanced accuracy as the required accuracy in the case of obtaining an overview of damaged buildings (4.2 (1)), a 90 % balanced accuracy and a 95 % recall as the required accuracy in the case of determining investigated buildings (4.2 (2)). Using recall implies that this system prioritizes the reduction of overlooking data (FN) in mistaken data (FP and FN). If this overlooking of investigative buildings happens, some supports cannot reach victims who need them. However, buildings of FP data are investigation targets because they are estimated as damage data by this system. Their mistaken (FP) can be revised at the

investigation. Additionally, many regions may have much undamaged buildings than damaged ones in general. Furthermore, if people revise mistaken data of this system's estimation to get more reliable data, they may revise the estimated data as damaged buildings (TP and FP), considering the smaller number of data. Therefore, the probability of revising FP data is higher than FN data. From the above, this study used not only balanced accuracy but also recall as a way of determining investigated buildings (4.4 (2)).

### 5.4 Estimation of Blue Sheets

We used aero-photo images taken by drone in the Ibaraki city of Osaka prefecture after the Northern Osaka's earthquake and location information of building polygons in OSM. We excluded some data which we cannot label because of low resolution, failure of ortho process and big gap between buildings and polygons.

The number of data used is shown Table 3. We increased data by rotation, horizontal reverse and contrast change of image to get a high accuracy and subsequently make a robust model. Furthermore, we made the number of training data in the two classes equal to have a high balanced accuracy. Putting label on each image took about 4 s. The total time was about 4 h.

This result is shown in Table 4. The mean of the balanced accuracy was 97.57 % and the mean of recall was 96.74 %. Stopped training times was 3000 times, and training time was about 65 min using GPU (GeoForce GTX TITAN X).

### 5.5 Estimation of Direct Damage

We used aero-photo images taken by aircraft in the Mashiki city of Kumamoto prefecture after the 2016 Kumamoto Earthquake and location information of building polygons in OSM. Additionally, we excluded some data which we cannot label even by human sight because of the low resolution of the aero photo images.

**Table 4**  Result of the estimation

|  | Blue sheet | | Direct damage | |
|---|---|---|---|---|
|  | Balanced accuracy | Recall | Balanced accuracy | Recall |
| 1st time | 98.30 % | 98.30 % | 92.63 % | 87.86 % |
| 2nd time | 97.12 % | 95.38 % | 94.46 % | 92.95 % |
| 3rd time | 97.28 % | 96.54 % | 93.43 % | 95.23 % |
| mean | 97.57 % | 96.74 % | 93.51 % | 92.01 % |

**Table 5**  Examples of used data

|  | Blue sheet | No blue sheet |
|---|---|---|
| Training data | 4896(306×16) | 4900(2450×2) |
| Validation data | 1232(77×16) | 1180(590×2) |
| Test data | 169 | 2702 |

**Table 6** Confusion matrix of used examples

|  |  | Actual | |
|---|---|---|---|
|  |  | Blue sheet | No blue sheet |
| Estimate | Blue sheet | 162 | 161 |
|  | No blue sheet | 7 | 2541 |

The number of data used is shown in Table 3. We also increased data by data augmentation. Furthermore, we also made the number of training data in the two classes equal. Putting label on each image took about 6 s. The total time was about 15 h.

This result is shown in Table 4. The mean of the balanced accuracy was 93.51 % and the mean of recall was 92.01 %. Stopped training times was 3000 times, and training time was about 65 min using GPU (GeoForce GTX TITAN X).

### 5.6 Used Example

We used the system in one area of Ibaraki city to show the used example. Table 5 represents the number of used data and Table 6 represents the confusion matrix of the result.

For obtaining an overview of damaged buildings (4.2 (1)), although the actual proportion of the damaged buildings was 5.886 % (169/2871), the estimation of the system was 11.25 % ((162 + 161)/2871). Consequently, it is assumed that this overestimation was caused by many mistaken data of the actual no blue sheet data. When an area has more undamaged roofs than damaged ones, the system is expected to overestimate.

When determining the investigated buildings (4.2 (2)), the system can reduce 2871 investigated buildings to 323 (11.25 %). However, this result means that the system overlooked 7 damaged buildings.

Furthermore, by referencing location information of each image data from GIS, the system can visualize the estimation result onto some maps or aero-photo images like as seen in Figs. 9 and 10. It is assumed that local governments can make an efficient plan for investigation from the visualized map.

## 6 Challenges of the System in Operation

### 6.1 Aero-photo Images

The system is expected to use aero-photo images taken by drones or aircrafts. Aircrafts can take photos in a large area in the form of one prefecture, approximately within one week. However, photos taken by the Geospatial Information Authority have 20 cm resolutions. Roughly 20 % of the roof-image database images, used in the estimation of direct damage, could not be categorized (damage and no damage) even by the naked eye; hence, they were removed from the training data in our experiments. However, drones of some organizations, such as the DRONE BIRD can take aero-photo images with a high resolution, in the order of several centimeters. However, drones take photos only in small areas,

**Fig. 9** Map visualizing estimation result on map (blue paint: actual blue sheet; red frame: estimated blue sheet)

**Fig. 10** Map visualizing estimation result on aero photo (blue point: actual blue sheet; red point: estimated blue sheet)



of approximately $10\sim 20$ square kilometers, owing to their smaller battery life.

The system therefore needs to improve its method of using aero-photo images. Example includes that photos taken by drones are used only as training data either in a severely damaged area, a thickly housed area, or the one in which changes have taken place due to aftershocks. Because drones and aircrafts have merit and demerit respectively, the system needs to use both photos properly like above examples.

### 6.2 Accuracy of Building Polygons

The accuracy of building polygons in an OSM is not high at present, as shown in 4.4.2 and other study (Girres & Touya, 2010). To improve this accuracy, mappers must be capable of creating precise drawings and the latest aero photo taken after earthquake must be used as the background map while drawing. According to Taichi Furuhasi, the trustee of Crisis Mappers Japan, in the 2016 Kumamoto Earthquake, approximately 17,000 building polygons in Mashiki city were drawn by a total of 122 mappers for approximately one day. Therefore, the system does not encounter any difficulties in terms of the time required to draw the building polygons.

### 6.3 People Putting Label for Deep Learning

If the labels for the data, used by the deep learning model, is input by the volunteers via crowdsourcing, a significant amount of labelled data can be generated. However, the accuracy of the labels cannot be completely ensured. Although some incorrect labels by the volunteers lead to inaccuracies in the estimation, no one may take responsibility.

Whereas, if staffs of the involved organization, such as local governments, provide the labels, the accuracy can be ensured to a certain extent, as they would be more responsible to toward the labelling task. However, a significant amount of labelled training data cannot be generated, because the number of that stuff is much smaller than volunteers.

Going forward, we need to consider the amount of necessary training data and time required for labeling in order to select an appropriate labelling method.

## 7 Discussion

In the blue sheet estimation, the mean of the balanced accuracy and recall were 97.57 % and 96.74 %, respectively. These values exceeded the level of accuracy required in obtaining an overview of the damaged buildings and determining the investigated buildings. In the direct damage estimation, the means of the balanced accuracy and recall were 93.51 % and 92.01 %, respectively. These values also exceeded the level of accuracy required with regard to obtaining an overview of damaged buildings but not in the determination of the investigated buildings because low recall means leading overlooking of damage buildings.

From the above result, it is assumed that local governments can know which area has large damage. This is good for not only local government but also victims because their buildings which are estimated as damage can be investigated preferentially. Additionally, it is assumed that high recall in the blue sheet estimation can prevent overlooking.

Subsection 5.6 presented a used example, where the system was operated in an actual disaster. The system was found to be

capable of estimating the buildings that needed investigation and also reduce the number of investigations efficiently, although the accuracy of estimation for direct damage may decrease. However, if the target area has more undamaged buildings than damaged ones, the system was found to overestimate damaged buildings. In the future work, we must consider the reduction of the overestimation to make the system more practical. Additionally, in both these applications, we must not only consider the damaged roofs of buildings but also other damaged parts. Furthermore, considering that these damaged buildings could not be detected by our system, we must use information of other investigations, such as quick post-earthquake inspections of damaged buildings or other estimation methods. Moreover, local governments can use map visualizing estimation result to make an efficient investigation plan.

The accuracy of the direct damage estimation was lower than that of the blue sheet in 5.4. and 5.5 experiments. It is assumed that this is caused by the lower resolution of the aero photo data and the difficulty in the extraction of direct damage features. While we used aero-photo images by drones in the estimation of blue sheet, aero-photo images by aircrafts were used in the estimation of direct damage. Generally, deep learning requires a considerable amount of training data to increase its accuracy. However, there are few high resolution aero-photo images, particularly aero-photo images including direct damage, owing to the very low frequency of earthquake disasters and the relatively short span of time since the development of drone techniques. Additionally, because many victims cover the damaged parts of their roofs with blue sheets few days after an earthquake occurs, it is extremely difficult for organizations from the outside of the damaged area to take aero-photo images. Therefore, some organizations, companies, and research institutes need a structure that can enable data to be shared among them such that sufficient training data is available for the system.

Further considerations will be needed to obtain more findings about the data, model, and learning methods to improve the accuracy of the system and address the challenges encountered during actual operations.

# References

Cabinet Office in Japan. (2013). Operation standard of damage certification for houses in disaster, pp. 1–3 (in Japanese). Available at: http://www.bousai.go.jp/taisaku/pdf/shishinall.pdf. Accessed 18 Feb 2020.

Cabinet Office in Japan. (2017). Operation manual of damage certification work for houses in disaster pp. 2, (in Japanese) Available at: http://www.bousai.go.jp/taisaku/pdf/saigai_tebiki_full.pdf. Accessed 18 Feb 2020.

Cabinet Office in Japan. (2018). Summary of revision in March 2018. (in Japanese) Available at: http://www.bousai.go.jp/taisaku/pdf/h3003kaitei.pdf. Accessed 18 Feb 2020.

Crisis Mappers Japan. (2020). Crisis Mappers Japan. (in Japanese) Available at: http://crisismappers.jp/about.html. Accessed 1 Apr 2021.

Curran, K., Crumlish, J., & Fisher, G. (2012). OpenStreetMap. *International Journal of Interactive Communication Systems and Technologies, 2*(1), 69–78.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009) ImageNet: A large-scale hierarchical image database. In Proc. CVPR, pp. 2–9.

DRONE BIRD. (2020). DRONE BIRD. (in Japanese) Available at: http://dronebird.org. Accessed 1 Apr 2021.

Fujiu, M., Ohara, M., & Meguro, K. (2013). Study on operation test of remote building damage assessment systemduring large scale earthquake disaster, Institute of Industrial Science, University of Tokyo, Bulletin of ERS, No. 46.

Geospatial Information Authority of Japan. (2020). About aero-photo images. (in Japanese) Available at: https://www.gsi.go.jp/gazochosa/gazochosa41006.html. Accessed 1 Apr 2021.

Girres, J.-F., & Touya, G. (2010). Quality assessment of the French OpenStreetMap dataset. *Transactions in GIS, 14*(4), 435–459.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778).

Hezaveh, M. M., Kanan, C., & Salvaggio, C. (2017). Roof damage assessment using deep learning, *2017 IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*.

Horie, K., Shigekawa, K., Maki, N., Tanaka, S., & Hayashi, H. (2005). Application of Damage Assessment Training System (DATS) to Ojiya City following the 2004 Niigata-ken Chuetsu Earthquake - through a disaster response support activity for issuing victim certificate- (in Japanese). *Journal of Social Safety Science, 7*, 123–132.

Hu, J., Shen, L., & Sun, G. (2017). *Squeeze-and-excitation networks*. arXiv preprint arXiv:1709.01507.

IMAGENET Large Scale Visual Recognition Challenge. (2012). ImageNet large scale visual recognition competition 2012 (ILSVRC2012). Available at: http://image-net.org/challenges/LSVRC/2012/results.html. Accessed 1 Apr 2021.

Inoue, M., Suetomi, I., Fukuoka, J., Onishi, S., Numada, M., & Meguro, K. (2018). Development of estimation formula of disaster response-work volume based on the Kumamoto earthquake (in Japanese). *Monthly Journal of the Institute of Industrial Science, University of Tokyo, 70*(4), 289–297

Ji, M., Liu, L., Du, R., & 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*, 1202.

Kamagatani, Y., & Matsuoka, M. (2017). Damaged building recognition of the 2016 Kumamoto earthquakes using deep learning with aerial photographs (in Japanese). Tono Research Institute of Earthquake Science "the report of Disaster Prevention Research Committee 2017". pp. 49–57.

Kashani, A. G., & Graettinger, A. J. (2015). Cluster-Based roof covering damage detection in ground-basedlidar data. *Automation in Construction, 58*, 19–27.

Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems 25* (NIPS 2012).

Matsuoka, Y., Fujiu, M., Takayama, J., Nakayama, S., Suda, S., & Sakaguchi, H. (2017). Development of new hazard map after a disaster using unmanned aerial vehicle (in Japanese), *54th Research Presentation of Infrastructure Planning and Management*.

Ministry of Internal Affairs and Communications. (2018). Study about the issue of damage certification in large scale earthquakes-focus on the 2016 Kumamoto Earthquake-. (in Japanese) Available at: https://www.soumu.go.jp/main_content/000528758.pdf. Accessed 1 Apr 2021.

Murakami, S., Hayashi, H., Tamura, K., Maki, N., Higashida, M., Horie, K., Hamamoto, R., & Komatsu, R. (2012). Analysis for effective operation of victim's certificate -in case of 2012 Kyoto-fu Nanbu flood disaster (in Japanese). *Journal of Social Safety Science, 23*, 1–10.

National Network for Emergency Mapping. (2020). National Network for Emergency Mapping. (in Japanese) Available at: https://www.n2em.jp. Accessed 18 Feb 2020.

Nex, F., Duarte, D., Tonolo, F. G., & Kerle, N. (2019). Structural building damage detection with deep learning: assessment of a state-of-the-Art CNN in operational conditions. *Remote Sensing, 11*, 2765.

Ogawa, N., & Yamazaki, F. (2000). Photo-interpretation of building damage due to earthquakes using aerialphotographs, Proceedings of the 12th World Conference on Earthquake Engineering.

Open Aerial Map. (2020). Open Aerial Map. Available at: https://openaerialmap.org. Accessed 1 Apr 2021.

Prechelt, L. (1998). Automatic early stopping using cross validation: quantifying the criteria. *Neural Networks, 11*(4), 761–767.

Sharan, C., Woolley, C., Vandermersch, P., Cohen, J., & Tran, J. (2014). cuDNN: efficient primitives for deep learning, arXiv preprint, arXiv: 1410.0759.

Shimizu, S., Komaru, Y., Wakaura, M., Tokizane, Y., Nakamura, H., & Fujiwara, H. (2019). Fragility function of wooden house considering the roofing type (in Japanese). *Journal of social safety science, 34* , 63–73.

Suppasri, A., Mas, E., Charvet, I., Gunasekera, R., Imai, K., Fukutani, Y., Abe, Y., & Imamura, F. (2013). Building damage characteristics based on surveyed data and fragility curves of the 2011 Great East Japan tsunami. *Natural Hazards, 66*(2), 319–341.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015)*Going deeper with convolutions*. In CVPR.

Zeiler, M. D., & Fergus, R. (2014). *Visualizing and understanding convolutional networks*. ECCV. 818–833. https://doi.org/10.1007/978-3-319-10590-1_53.

**Shono Fujita** is a Ph.D. student at graduate school of informatics in Kyoto University, Japan. He holds a Master of Science in Informatics from Kyoto University. His research interests include image recognition by deep learning and automatic method for disaster response.

**Michinori Hatayama** is a Professor of Disaster Prevention Research Institute in Kyoto University. He holds a doctoral degree in Department of Computational Intelligence and Systems Science from Tokyo Institute of Technology. His research interests include Spatial Temporal Information System, Disaster Information and Geographic Information Science.