Article

New Insights into Multiclass Damage Classification of Tsunami-Induced Building Damage from SAR Images

Yukio Endo [1],* , Bruno Adriano [2] , Erick Mas [3] and Shunichi Koshimura [3]

[1] Graduate School of Engineering, Tohoku University, Aoba-Ku, Sendai 980-8752, Japan
[2] Geoinformatics Unit, RIKEN Center for Advanced Intelligence Project, Chuo-ku, Tokyo 103-0027, Japan; bruno.adriano@riken.jp
[3] International Research Institute of Disaster Science, Tohoku University, Aoba-Ku, Sendai 980-8752, Japan; mas@irides.tohoku.ac.jp (E.M.); koshimura@irides.tohoku.ac.jp (S.K.)
* Correspondence: yukio.endo.r2@dc.tohoku.ac.jp; Tel.: +81-22-752-2084

check for
updates

Abstract: The fine resolution of synthetic aperture radar (SAR) images enables the rapid detection of severely damaged areas in the case of natural disasters. Developing an optimal model for detecting damage in multitemporal SAR intensity images has been a focus of research. Recent studies have shown that computing changes over a moving window that clusters neighboring pixels is effective in identifying damaged buildings. Unfortunately, classifying tsunami-induced building damage into detailed damage classes remains a challenge. The purpose of this paper is to present a novel multiclass classification model that considers a high-dimensional feature space derived from several sizes of pixel windows and to provide guidance on how to define a multiclass classification scheme for detecting tsunami-induced damage. The proposed model uses a support vector machine (SVM) to determine the parameters of the discriminant function. The generalization ability of the model was tested on the field survey of the 2011 Great East Japan Earthquake and Tsunami and on a pair of TerraSAR-X images. The results show that the combination of different sizes of pixel windows has better performance for multiclass classification using SAR images. In addition, we discuss the limitations and potential use of multiclass building damage classification based on performance and various classification schemes. Notably, our findings suggest that the detectable classes for tsunami damage appear to differ from the detectable classes for earthquake damage. For earthquake damage, it is well known that a lower damage grade can rarely be distinguished in SAR images. However, such a damage grade is apparently easy to identify from tsunami-induced damage grades in SAR images. Taking this characteristic into consideration, we have successfully defined a detectable three-class classification scheme.

Keywords: synthetic aperture radar; change detection; tsunami; buildings

1. Introduction

The use of satellite remote sensing within the domain of natural hazards and disasters has recently received widespread attention [1,2]. This novel sensing technology allows us to achieve a rapid damage assessment due to its fast response and wide field of view. Particularly, synthetic aperture radar (SAR), which can provide a high-resolution image irrespective of weather conditions, has the potential to detect damaged buildings with high accuracy [3]. The periodicity of satellites in orbit enables computing changes between a pair of SAR images [4,5]. Moreover, the recent emergence of freely available SAR image, such as Sentinel-1 data, has opened up the possibility for more users to

conduct urgent observations the Earth surface [6,7]. Several authors have addressed the problem of retrieving building damage information from SAR images to aid disaster relief activities.

The detection of damaged buildings from SAR images is divided into two processes: computing change features and constructing a classification model. Previous studies have suggested that a direct pixel comparison leads to poor change detection accuracy because it mainly focuses on the spectral values and mostly ignores the spatial context [8]. Calculating change features over a window is effective in considering the spatial distribution of image pixels [9]. Many authors have successfully applied change features derived from subimages to detect damage caused by natural disasters [10–14]. Moreover, using building footprint data and high-resolution SAR images enable detecting building damage at the building unit scale [15]. After computing change features, a classification model needs to be constructed. Many studies have proposed thresholds and parameters for damage detection models based on experience, with limited success [16]. A supervised machine learning approach enables overcoming such limitations. This approach allows us to construct a high-dimensional discriminant function [17,18]. The quality and quantity of training data play a pivotal role in supervised machine learning. These training data can be obtained through visual inspection of high-resolution optical images [19]. Moreover, by replacing training samples with probabilistic information derived from the spatial distribution of the hazard, databases of previous disasters can be used for constructing a discriminant function [20].

From the perspective of assessing monetary damages, damaged buildings should be sorted into detailed classes. For seismic events, multiclass classification of damaged buildings has been widely studied in the field of satellite remote sensing [21]. According to previous studies, the relationship between the damage grade and the building appearance in remote sensing imagery differs depending on the type of data used [22]. However, identifying lower damage grades at the building unit scale has been difficult, regardless of the sensor used [23–25]. Thus, these grades tend to be aggregated as one class to define a detectable classification scheme in the case of earthquakes. There are few reports about the multiclass classification of damaged buildings for tsunami-induced damage. Buildings in tsunami affected areas generally suffer damage on the lower part of the building sidewall, except for buildings that were washed away. Using optical sensor images, which allow us to inspect only roof conditions, it is difficult to conduct a detailed classification immediately after the disaster. Meanwhile, SAR is seemingly effective in detecting such changes due to its side-looking geometry [26]. However, an automatic classification of buildings with tsunami-induced damage in SAR images remains a major challenge. Although the multiclass classification model that considers the statistical relationship between the change ratios in areas with high backscattering and in areas with building damage has been proposed, lower dimensional feature space of the classifier has limited the generalization ability [27]. Unfortunately, previous studies have not provided consistent guidance on how to define a detectable classification scheme for SAR images. It is clear that further research is necessary to better understand multiclass classification for buildings with tsunami-induced damage.

The objective of this paper is to propose a new model for classifying building damage into detailed classes using pre- and post-event high-resolution SAR images. By evaluating the generalization ability of the model with various classification schemes, we demonstrate the limitations and potentials of multiclass classification of buildings with tsunami-induced damage. To perform satisfactory multiclass classification, the model considers change features derived from several sizes of pixel windows. In previous studies, a specific size of pixel window is selected due to the difficulty in considering several window sizes simultaneously. The selected window size differs depending on the region of the world and the quality of images [28,29]. For example, when analyzing building damage caused by the 2011 Great East Japan Earthquake and Tsunami from TerraSAR-X images, a 5×5 window is typically used [15,17]. However, using a single window size raises a question about the accuracy of derived change features. We would like to stress the need for constructing a model that can consider a high-dimensional discriminant function derived from several sizes of pixel windows. To this end, our classification model uses a support vector machine (SVM) for optimally weighting each change feature.

Due to its outstanding generalization ability, SVM has been used in the field of remote sensing [30]. To avoid the overfitting problem, which has been an obstacle to using a high-dimensional feature space, generalization parameters are determined based on the combination of a coarse grid search and a fine grid search. We tested the performance of building damage classification using pre- and post-event TerraSAR-X images taken in the 2011 Great East Japan Earthquake and Tsunami.

2. Study Area and Dataset

This study focused on the coastal area of Sendai city (Figure 1) in Miyagi prefecture, Japan. This city was devastated by the Great East Japan Earthquake and Tsunami on 11 March 2011. TerraSAR-X data acquired on 20 October 2010 and 12 March 2011 were used to detect damaged buildings. The images were captured in Stripmap mode and delivered as Enhanced Ellipsoid Corrected (EEC) products. The details of the acquisition of SAR images are shown in Table 1.

Table 1. Details of the acquisition of SAR images used in this study.

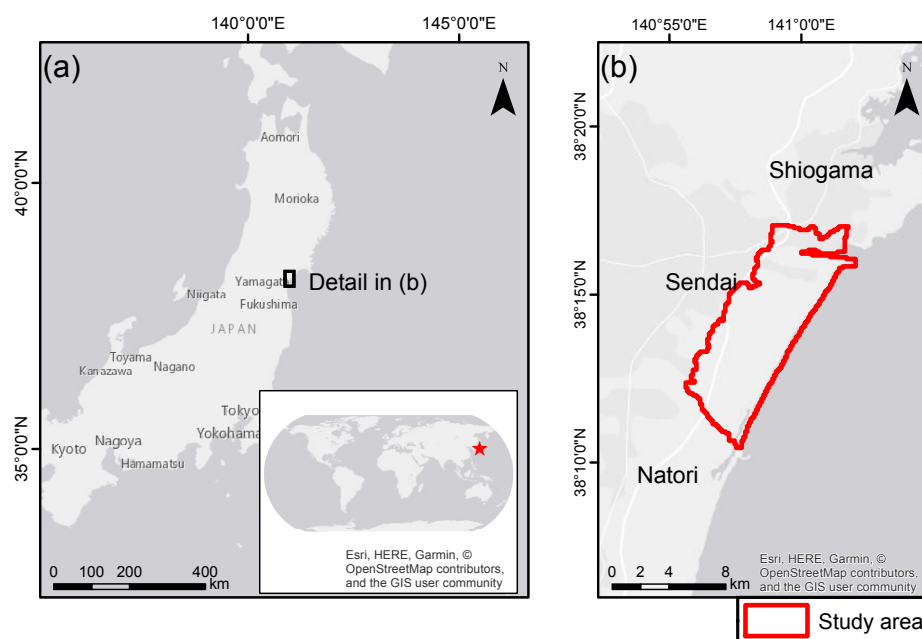| Sensor	Acquisition	Incidence Angle	Pixel Spacing	Path	Polarization	Band
TerraSAR-X	20 October 2010	37.3°	1.25 m	Descending	HH	X
TerraSAR-X	12 March 2011	37.3°	1.25 m	Descending	HH	X

Figure 1. Study area: (**a**) Tohoku region in Japan; and (**b**) coastal areas of Sendai.
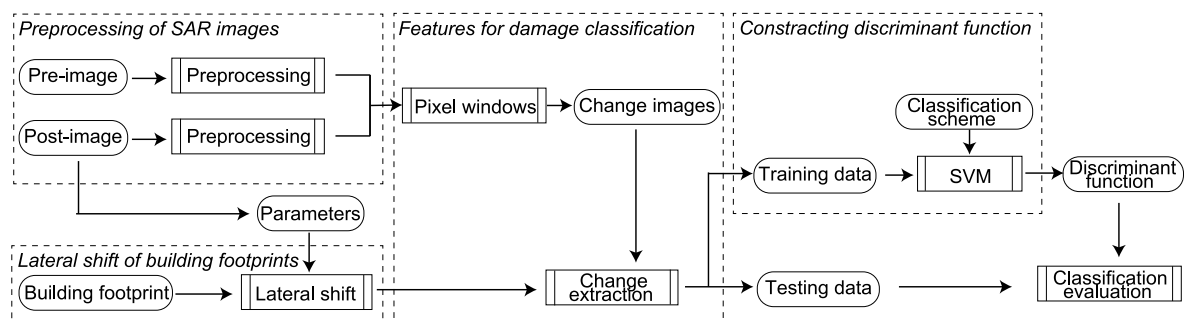
The ground-truth data (GTD) of building damage were provided by the Ministry of Land Infrastructure, Transport and Tourism (MLIT) [31]. These data were released on 23 August 2011, five months after the disaster occurred. The data were referenced at the building footprint and include 11,562 buildings described by seven damage grades (Table 2): "G7: no damage," "G6: minor damage," "G5: moderate damage," "G4: major damage," "G3: complete damage," "G2: collapsed," and "G1: washed away."

Table 2. Ground truth data provided by MLIT.

	Original Classification	
	Class	**Samples**
G1	Washed away	4114
G2	Collapsed	834
G3	Complete damage	615
G4	Major damage	1328
G5	Moderate damage	2716
G6	Minor damage	1130
G7	No damage	825

3. Methods

A schematic overview of the classification and performance evaluation framework is shown in Figure 2. In this study, preprocessing of the satellite images was performed with ENVI SARscape (ver 5.4). SVM was implemented in Python (3.6.3) using the scikit-learn library. All other processing and analysis steps were performed with ArcGIS (ver 10.5.1).

Figure 2. Schematic overview of the classification and performance evaluation workflow.

3.1. Preprocessing of SAR Images

EEC product was projected and resampled to the WGS84 reference ellipsoid by the vendor [32]. The distortion effects caused by varying terrain were also corrected using 90 m digital elevation data provided by Shuttle Radar Topographic Mission. Image pixel values represent brightness intensity of target objects on the earth. However, these values are influenced by the relative orientation of the illuminated resolution cell and the sensor, as well as by the distance in range between them. To make multitemporal images easily comparable, radiometric and geometric corrections were performed in the preprocessing steps.

First, brightness intensity was converted into backscattering coefficient, which represents the radar reflectivity per unit area in the ground range. Second, the images were coregistered using template matching to reduce the effect of the displacements caused by crustal movements. Finally, a Lee filter with a kernel size of 3×3 pixels was used to decrease the effect of speckle noise. Preprocessed SAR images are shown in Figure 3.
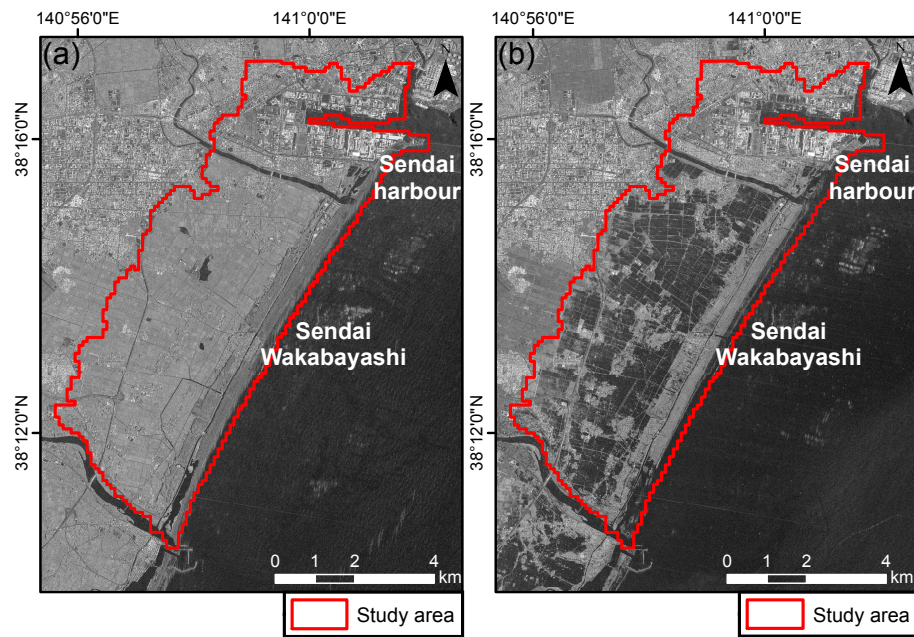
Figure 3. Preprocessed SAR images: (**a**) pre-event image; and (**b**) post-event image.

3.2. Lateral Shift of Building Footprints

Due to the side-looking geometry of SAR sensors, the radar beam reaches the roof-top and sidewalls of buildings before it reaches the bottom, as shown in Figure 4. In preprocessed images, the foot fringe of a building leans toward to the direction of SAR illumination. This distortion effect is called layover, and has to be taken into account in extracting building information. To match the building outlines in the SAR images, we shifted the building footprints by a portion of the layover. The length of the layover (L) is calculated by Equation (1).

$$L = \frac{H}{\tan \theta} \tag{1}$$

where H is the building height and θ is the incident angle. With $\theta = 37.3°$ and an assumed average building height of H = 6 m, the length of the layover is approximately 7.9 m. Thus, considering the path of the satellite (190.4° clockwise from north), a lateral shift of the building geometries could be decomposed into 7.8 m to the west and 1.4 m to the south.
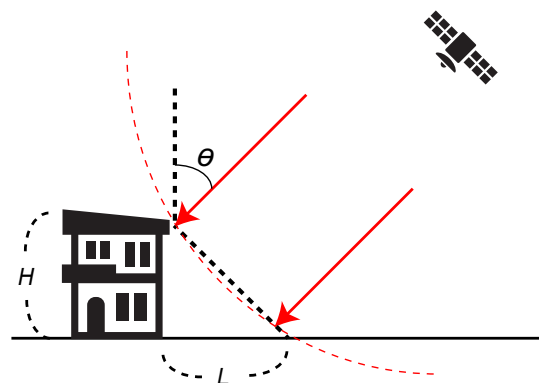

Figure 4. Schematic representation of SAR observations for buildings.

3.3. Features for Damage Classification

To compute change values per pixel in SAR images caused by the tsunami, the difference of the backscattering coefficients and the correlation coefficients were derived from the two images. The difference d and the correlation coefficients r are calculated by Equations (2) and (3), respectively.

$$d = \bar{I}b_i - \bar{I}a_i \tag{2}$$

$$r = \frac{N \sum\limits_{i=1}^{N} Ia_i Ib_i - \sum\limits_{i=1}^{N} Ia_i \sum\limits_{i=1}^{N} Ib_i}{\sqrt{\left(N \sum\limits_{i=1}^{N} Ia_i^2 - \left(\sum\limits_{i=1}^{N} Ia_i \right)^2 \right) \cdot \left(N \sum\limits_{i=1}^{N} Ib_i^2 - \left(\sum\limits_{i=1}^{N} Ib_i \right)^2 \right)}} \tag{3}$$

where i is the pixel number, Ia_i and Ib_i are the backscattering coefficients of the pre- and post-images, and $\bar{I}a_i$ and $\bar{I}b_i$ are the corresponding averaged values over the N pixel window surrounding the ith pixel. In calculating these change values, six pixel window sizes were used: 5×5, 9×9, 13×13, 17×17, 21×21 and 25×25.

The shifted building footprints were used for calculating the change features per building. The pixels within a footprint were considered to belong to the same segment. The averages of each change value were calculated over aggregated clusters of pixels. This model adopts these mean values as change features.

3.4. Discriminant Function

SVM is a supervised nonparametric statistical learning technique [33]. In this study, SVM was used to construct a model that predicts the damage grade of buildings in the testing set given only the change features. When constructing a classification model, SVM maps training data into a higher-dimensional space and finds an optimal separating hyperplane with the maximum margin. A kernel function plays a pivotal role in mapping the training data and is set before the learning process begins. Using a kernel function allows us to set a non-linear discriminant function. In our study, the radial basis function was used as the kernel function. SVM was originally designed for binary classification. To extend SVM for multiclass classification, binary classifiers per class pair are constructed. This strategy is generally called "one-against-one". At prediction time, the class that receives the most votes is selected. This method is generally slower than other methods due to constructing a number of classifiers, but it is suitable for practical use [34]. It is well known that SVM is sensitive to scaling of the individual features because attributes in larger numerical ranges may dominate those in smaller numerical ranges. To standardize change features, the feature database was normalized to zero mean and unit variance. Note that the mean and standard deviation computed from training samples were also used for transforming the features of the testing samples.

To train an SVM equipped with a radial basis function (RBF) kernel, users must specify two hyperparameters (kernel coefficient γ and regularization parameter C) [35]. Choosing a separating hyperplane depends not only on the training samples but also on these parameters. Thus, we have to tune these parameters such that the classifier can accurately predict unknown data. In this study, the optimal SVM parameters ewre tuned for each classification according to a grid search with a ten-fold cross-validation. A grid search consists of systematically enumerating all possible hyperparameters in a certain range and checking whether each candidate of hyperparameter provides a minimal error rate. This error rate was measured by ten-fold cross-validation on the training data. Since it takes time to complete a grid search in a wide range, this study initially considered a coarse grid. In a coarse grid, the candidates of γ and C are $\log_2 \gamma \in \{-15, -13, \cdots, 3\}$ and $\log_2 C \in \{-5, -3, \cdots, 15\}$, respectively. On the assumption that the optimal parameters lie in a peripheral region of the selected parameters in the first grid search, a finer grid search was conducted on that region. For example, assuming that

2^{-1} is selected as kernel coefficient and 2^7 is selected as regularization parameter in a coarse grid search, the candidates of γ and C in the finer grid would be $\log_2 \gamma \in \{-3.00, -2.75, -2.50, \cdots, 1.00\}$ and $\log_2 C \in \{5.00, 5.25, 5.50, \cdots, 9.00\}$, respectively. The SVM parameters were finally determined according to a fine grid search. This parameter selection was conducted for each classification based on sampled training data.

4. Classification Conditions

The following three analyses were conducted to show the effectiveness of our model and to understand the potentials and limitations of multiclass classification of buildings with tsunami-induced damage. In the analyses, the performance of the proposed classifier was evaluated with various conditions. Conditions of the analyses are summarized in Table 3. The classification performance was measured by comparing the estimated class of the testing samples with GTD. Precision, recall, and F-score were presented as standard accuracy measures in respective classes [36]. Precision is the ratio of the total buildings that are correctly classified as a specific class of interest to the total buildings that are classified as the class. Recall is the ratio of the total buildings that are correctly classified as a specific class of interest to the total buildings that are labeled as the class in GTD. F-score is the weighted average of Precision and Recall. Macroaveraging, which treats all classes equally, was used for calculating the average score of each accuracy measure [37].

Table 3. Classification conditions of the analyses.

	Reference Scheme	Windows	Training Samples	Testing Samples
Analysis 1	Seven-class	$5 \times 5 \sim 25 \times 25$	300/class	300/class
Analysis 2	Three-class	$5 \times 5 \sim 25 \times 25$	300/class	500/class
Analysis 3	Three-class Three-class	$5 \times 5 \sim 25 \times 25$ 5×5	50~400/class 50~400/class	400/class 400/class

The first analysis evaluated the performance of the proposed classification model with the original classification scheme, which divided damaged buildings into seven grades based on the field survey of MLIT. In total, 300 training samples per class were randomly selected from the study area. Additionally, 300 testing samples per class were randomly selected from the study area.

In the second analysis, reclassifications of original scheme were conducted to evaluated the effect of changing classification scheme. Damaged buildings were reclassified into three classes, as shown in Table 4. Note that the reason these five schemes were considered is described in Section 5.2. Generalization ability of the classifier was evaluated with the each scheme. The numbers of training and testing samples were set to 300 and 500, respectively.

Table 4. Reclassification of MLIT's classification.

	Washed Away	Moderate Change	Slightly Change
Scheme 1	G1	G2	G3, G4, G5, G6, G7
Scheme 2	G1	G2, G3	G4, G5, G6, G7
Scheme 3	G1	G2, G3, G4	G5, G6, G7
Scheme 4	G1	G2, G3, G4, G5	G6, G7
Scheme 5	G1	G2, G3, G4, G5, G6	G7

Finally, the effectiveness of the proposed method, which uses twelve change features derived from $5 \times 5, 9 \times 9, 13 \times 13, 17 \times 17, 21 \times 21$ and 25×25 pixel windows, was verified through comparison with a model that uses only two change features derived from a 5×5 pixel window. Both models constructed a discriminant function based on the process in Section 3.4, and referred to the reclassification scheme that was the most detectable scheme in the second analysis. The number of training samples that were

randomly selected in the study area was changed in the range from 50 to 400 per class. The classification ability of the function was tested with 400 independent testing samples per class. To evaluate the robustness of the classifier, this sequence of steps was repeated 50 times. The average and standard deviation of the classification ability were computed.

5. Results and Discussion

5.1. MLIT's Classification Scheme

The mapping results of MLIT's classification scheme is shown in Figure 5. The error matrix derived from classifying testing samples is shown in Table 5. Table 6 summarizes the generalization abilities in respective classes. The classification accuracy for G1 buildings is the highest overall (F-score = 59.3%). In general, washed away buildings are easy to identify [16,23]. Notably, the analysis suggests the possibility that G7 buildings are also detectable in SAR images compared with G2-G6 buildings. In total, 178 of 300 G7 buildings were correctly classified, corresponding to F-score of approximately 45.6%. As previously stated, undamaged buildings can rarely be distinguished from slightly damaged buildings in the case of earthquakes [24]. The result of this analysis reveals that multiclass classification of tsunami-induced damage has a different quality from that of earthquake-induced damage. This difference should arise from SAR's sensitivity to the moisture content of the surface. It is well known that flooded area shows the lowest backscatter among other forms of land cover. This characteristic of SAR images allows us to extract flooded areas [38,39]. Such geospatial features that is not directly derived from layover of damaged buildings allow us to distinguish G7 buildings that are located in non-inundated areas from other buildings that are severely inundated. There is a possibility that combining several sizes of pixel window allows our method to take advantage of both of sidewall changes and landcover changes in damage detection.
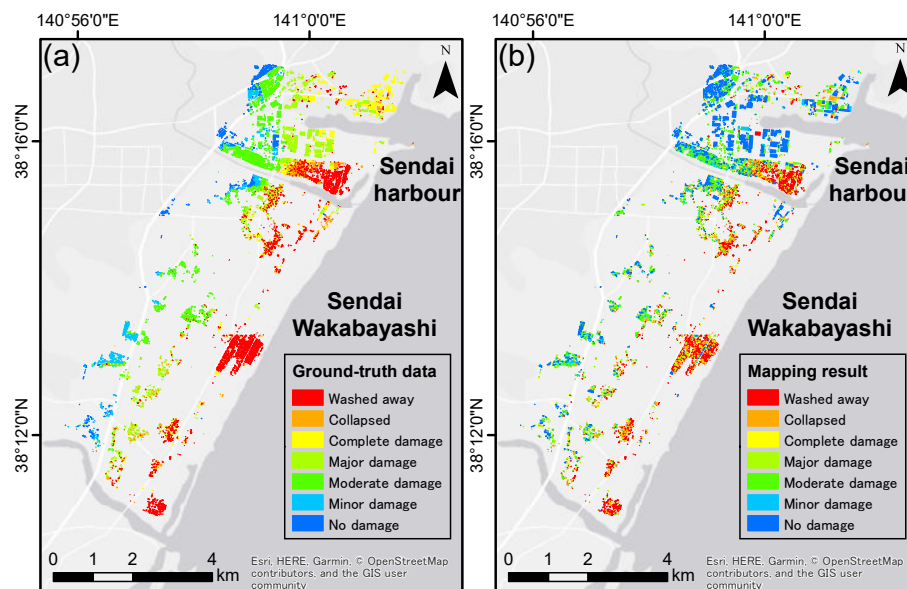
The model successfully classified G1 and G7 buildings, but it did not perform as well in the detection of other classes. In particular, the F-score for G3 buildings was extremely low. The factor that prevents detecting G3 buildings needs to be discussed. According to the official report of the field survey, G2 and G3 buildings are distinguished based on whether the buildings can be used after they are repaired. This definition highly depends on the subjectivity of the surveyor. In contrast, the other grades are defined based on the damage ratio, which is evaluated by a standardized measure in field surveys. Thus, such a vague definition of the grades may limit the generalization ability of a classifier. As indicated in Table 5, 225 buildings were classified as G2 and 99 buildings were classified as G3. This result can be interpreted as the classifier placed buildings that were included in either G2 or G3 into G2 to minimize misclassification costs.

Table 5. Error matrix for seven-class classification (MLIT's classification).

| | | Classification Result | | | | | | | |
		G1	G2	G3	G4	G5	G6	G7	Total
	G1	209	36	10	27	10	2	6	300
	G2	79	70	31	46	43	10	21	300
Ground	G3	79	70	31	46	43	10	21	300
truth	G4	41	32	19	81	52	29	46	300
data	G5	10	19	8	29	87	73	74	300
	G6	10	11	12	27	78	72	90	300
	G7	7	15	4	25	21	50	178	300
	Total	405	225	99	294	340	257	480	

Table 6. Accuracy evaluation of seven-class classification.

	Recall (%)	Precision (%)	F-score (%)
G1	69.7	51.6	59.3
G2	23.3	31.1	26.7
G3	5.0	15.2	7.5
G4	27.0	27.6	27.3
G5	29.0	25.6	27.2
G6	24.0	28.0	25.9
G7	59.3	37.1	45.6
Ave	30.9	33.9	31.3

Figure 5. Original classification scheme: (**a**) ground truth data; and (**b**) mapping result.

5.2. Three-Class Classification

In this analysis, the performance of the classifier was tested with various reclassification schemes in which damaged buildings are divided into three classes. As discussed in the previous analysis, washed away buildings seem to be easy to identify in SAR images. Thus, one class is composed of buildings that are originally classified into the "washed away (WA)" grade in reclassification schemes. The other buildings are divided into two classes, named "moderate change (MC)" and "slight change (SC)" (Table 4).

The mapping results of reclassification schemes are shown in Figures 6–10. Table 7 presents a comparison of the classification results. It is clear that the classifier referring to Scheme 5 outperformed the other classifiers. As discussed above, G1 buildings and G7 buildings have the potential to easily be identified in SAR images. This potential may also have a positive impact on Scheme 5. Aggregating damage grades (G2–G6) that can rarely be identified from SAR images probably increases the detectability of the defined three-class classification scheme. Thus, we can define the detectable three-class classification scheme by taking the characteristics of tsunami-induced damage into consideration. Fortunately, it has been shown that visual inspection of high-resolution optical images enables classifying damage buildings into three classes similar to Scheme 5 [40]. Meanwhile,

the classifier referring to Scheme 1 underperformed relative to the other classifiers. There is a possibility that the vague definition of the border between G2 and G3 decreases the detectability of Scheme 1.
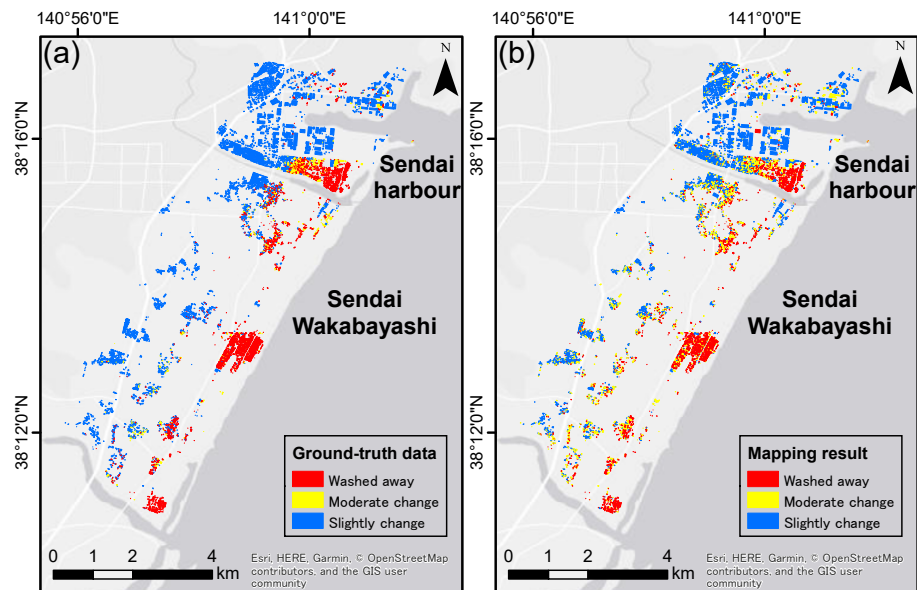

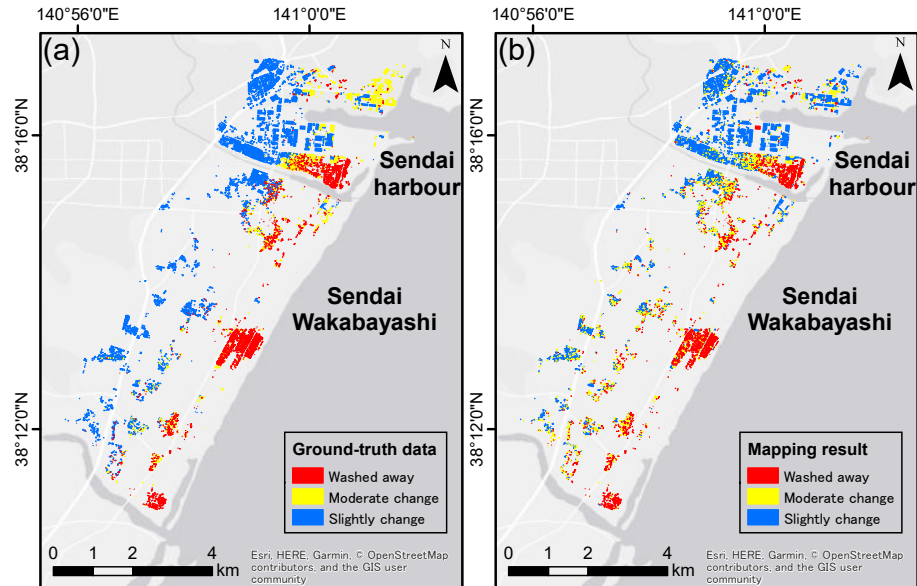**Figure 6.** Mapping result of Scheme 1: (**a**) ground truth data; and (**b**) mapping result.

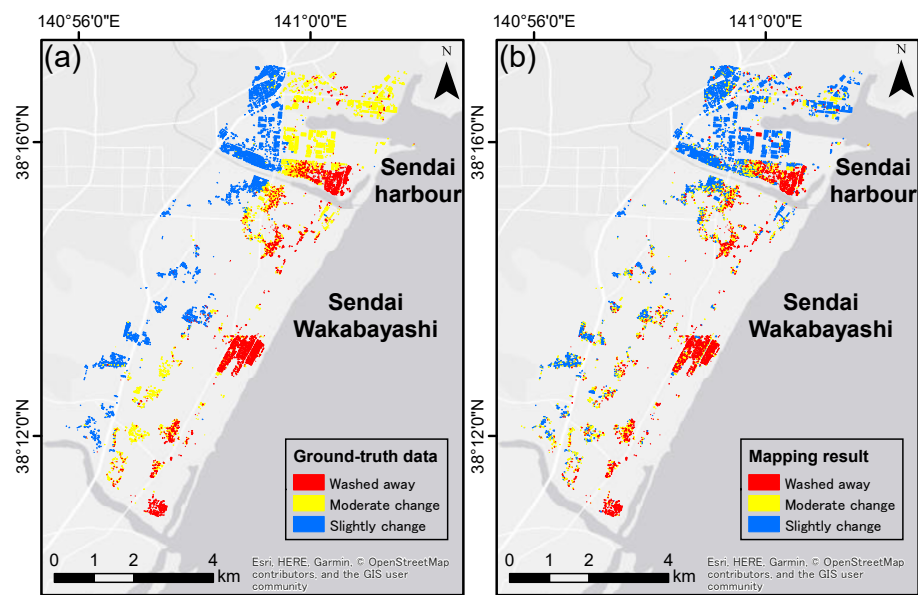**Figure 7.** Mapping result of Scheme 2: (**a**) ground truth data; and (**b**) mapping result.

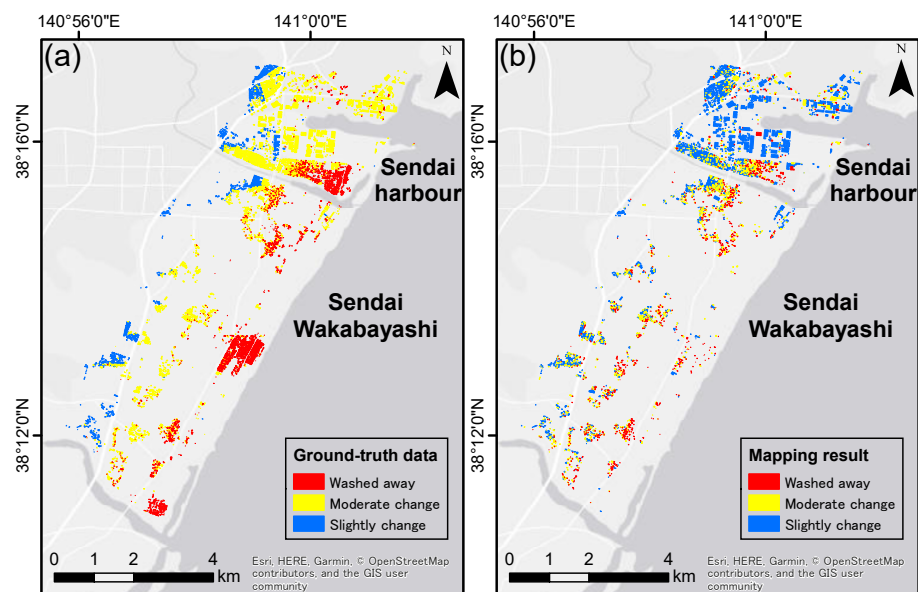**Figure 8.** Mapping result of Scheme 3: (**a**) ground truth data; and (**b**) mapping result.

Figure 9. Mapping result of Scheme 4: (**a**) ground truth data; and (**b**) mapping result.
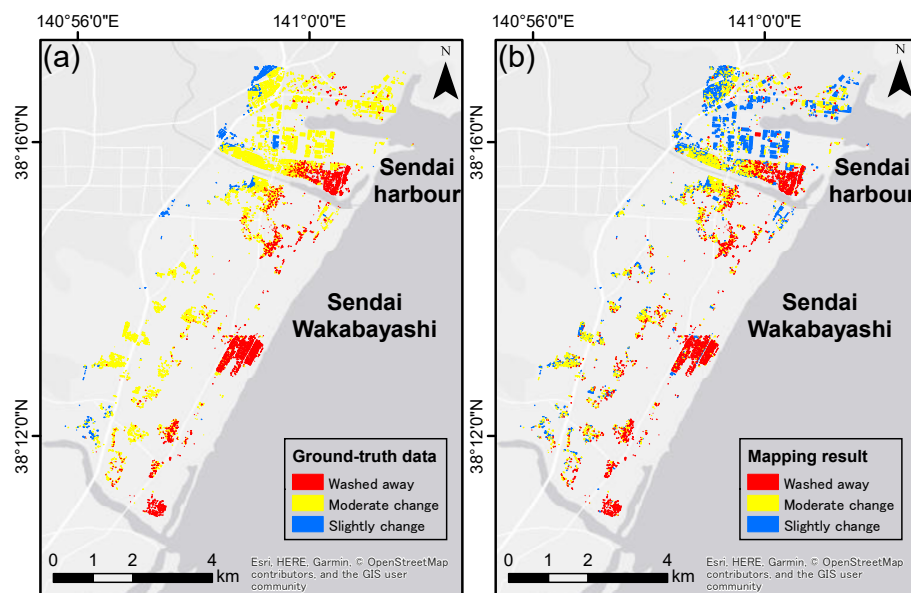
Figure 10. Mapping result of Scheme 5: (**a**) ground truth data; and (**b**) mapping result.

Table 7. Accuracy evaluation of three-class classification.

		Recall (%)	Precision (%)	F-Score (%)
Scheme 1	WA	70.0	67.7	68.8
	MC	46.2	52.2	49.0
	SC	71.2	65.8	68.4
	Ave	62.5	61.9	62.1
Scheme 2	WA	78.6	71.1	74.6
	MC	42.2	51.6	46.4
	SC	71.2	66.2	68.6
	Ave	64.0	62.9	63.2
Scheme 3	WA	77.0	76.5	76.8
	MC	41.0	54.8	46.9
	SC	78.4	62.9	69.8
	Ave	65.5	65.8	64.5
Scheme 4	WA	83.4	78.1	80.7
	MC	36.8	55.1	44.1
	SC	75.8	60.0	67.0
	Ave	65.3	64.4	63.9
Scheme 5	WA	85.2	83.5	84.4
	MC	55.2	61.9	58.4
	SC	73.8	67.8	70.7
	Ave	71.4	71.1	71.1

　　　To discuss the properties of our model in detail, we compared the mapping results of Scheme 1 and Scheme 5 (Figures 6 and 10). Although Scheme 5 provided the highest score, we could not confirm the superiority of Scheme 5 from a comparison of the mapping results. For Scheme 5, the difficulty lies in detecting the buildings in Sendai harbor, which is the northern industrial area of Sendai. Such misclassification may be from the difference in footprint area. In our model, change features are derived from the mean change values of pixels within each footprint. Thus, the more pixels that

a footprint includes, the less partial changes influence the mean values. Such tendency may cause misclassifying into "slightly change".

In reality, classifying damaged buildings into three or four classes is sufficient to aid the disaster relief and clean-up activities immediately after an earthquake or tsunami. Note that there is room for discussion on whether Scheme 5 is the best for disaster relief support. Therefore, the classifier has to be robust to various classification schemes. In previous studies, the generalization ability is not evaluated from the perspective of robustness to classification schemes due to the poor performance of multiclass classification. Notably, our method promises stable generalization ability (F-score > 60%) regardless of the defined three-class schemes.

5.3. Effectiveness of Using Multiple Pixel Windows

Finally, effectiveness of using multiple pixel windows was evaluated with Scheme 5, which was the most detectable classification scheme in the second analysis. To reduce the bias of training and testing samples, the average and standard deviation of the classification ability were computed over 50 iterations. The results of this analysis are summarized in Figure 11. This figure is a graphical representation of the relationship between the number of training samples and the macroaveraging of the F-score for each class. In the figure, solid lines represent the average of F-score. Colored areas (red and blue) are within the standard deviation range.

This figure clearly shows that the model considering twelve change features outperformed the model considering two change features. For both classifiers, the increase in F-score apparently began to slow at the classifier with 200 training samples per class. In general, SVMs that consider a higher-dimensional feature space considerably lose their generalization ability as the number of training samples decreases. Notably, our model provided a high generalization ability even with 50 training samples per class. This result apparently indicates that the double grid search has successfully controlled the hyperparameters of the classifier.
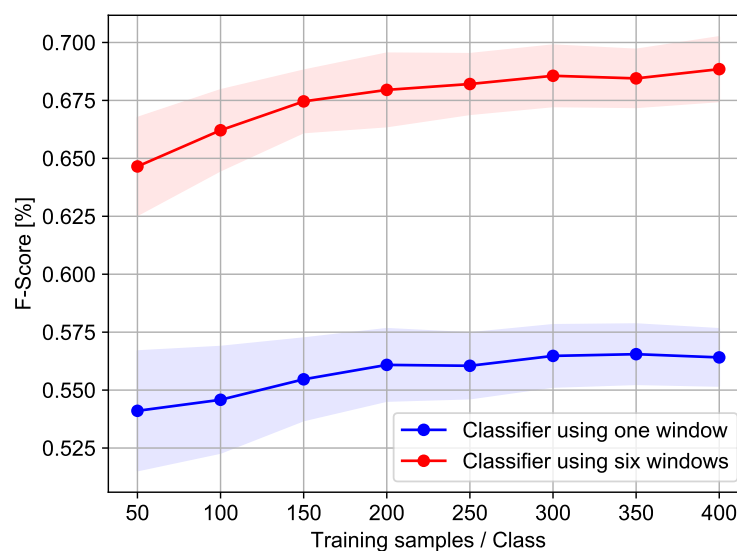

Figure 11. Learning curve for a comparison between SVM using one size of pixel window and SVM using six sizes of pixel windows. Solid lines represent the average of F-score. Colored areas (red and blue) are within the standard deviation range.

6. Conclusions

In this paper, we have developed multiclass building damage classification model for tsunami impact. This model considers a high-dimensional feature space derived from several sizes of pixel

windows. Taking advantage of SVM, we can construct a classification model from much smaller training samples.

To reveal the possibilities and limitations of multiclass classification of buildings with tsunami-induced damage, the proposed model was tested on the 2011 Great East Japan Earthquake and Tsunami. In the first analysis, classification performance was evaluated with the original seven-class classification scheme. Our analysis found that the proposed model can easily detect the most severely damaged buildings and undamaged buildings in tsunami affected areas. On the other hand, in the case of earthquake-induced damage, undamaged buildings can rarely be distinguished from slightly damaged buildings. This finding will advance understanding of the difference between tsunami-induced damage and earthquake-induced damage. The analysis also implies that detectability of other classes are limited due to the vague definition of the classes. Thus, we should apply a detectable classification scheme to ensure stable classification performance. To this end, the performance of the classifier was tested with various reclassification schemes in which damaged buildings were divided into three classes. Our study showed that aggregating damage grades (G2–G6) that can rarely be identified from SAR images increases the detectability of the defined three-class classification scheme. Notably, this classification scheme allows the proposed model to provide high generalization ability even with 50 training samples per class. The proposed method enables reducing the time required to retrieve training data in urgent situations.

Our findings open new possibilities not only for tsunami-induced damage detection but also for other change detection tasks. The proposed model, which extracts change features over several sizes of pixel windows and constructs a discriminant function using SVM, can be applied to other tasks, such as urban monitoring and earthquake damage detection. This possibility needs to be investigated in future studies.

Author Contributions: Y.E. and B.A. were responsible for the overall design and the study. Y.E. performed all calculations and drafted the manuscript. B.A. preprocessed the TerraSAR-X data. S.K. and E.M. contributed to designing the study. All authors read and approved the final manuscript.

References

1. Joyce, K.E.; Belliss, S.E.; Samsonov, S.V.; McNeill, S.J.; Glassey, P.J. A review of the status of satellite remote sensing and image processing techniques for mapping natural hazards and disasters. *Prog. Phys. Geogr.* **2009**, *33*, 183–207. [CrossRef]
2. Tralli, D.M.; Blom, R.G.; Zlotnicki, V.; Donnellan, A.; Evans, D.L. Satellite remote sensing of earthquake, volcano, flood, landslide and coastal inundation hazards. *ISPRS J. Photogramm. Remote Sens.* **2005**, *59*, 185–198. [CrossRef]
3. Plank, S. Rapid damage assessment by means of multi-temporal SAR-A comprehensive review and outlook to Sentinel-1. *Remote Sens.* **2014**, *6*, 4870–4906. [CrossRef]
4. Lu, D.; Mausel, P.; Brondizio, E.; Moran, E. Change detection techniques. *Int. J. Remote Sens.* **2004**, *25*, 2365–2407. [CrossRef]
5. Hachicha, S.; Chaabane, F. On the SAR change detection review and optimal decision. *Int. J. Remote Sens.* **2014**, *35*, 1693–1714. [CrossRef]
6. Raspini, F.; Bianchini, S.; Ciampalini, A.; Del Soldato, M.; Solari, L.; Novali, F.; Del Conte, S.; Rucci, A.; Ferretti, A.; Casagli, N. Continuous, semi-automatic monitoring of ground deformation using Sentinel-1 satellites. *Sci. Rep.* **2018**, *8*, 1–11. [CrossRef] [PubMed]
7. Olen, S.; Bookhagen, B Mapping Damage-Affected Areas after Natural Hazard Events Using Sentinel-1 Coherence Time Series. *Remote Sens.* **2018**, *10*, 1272. [CrossRef]

8.   Hussain, M.; Chen, D.; Cheng, A.; Wei, H.; Stanley, D. Change detection from remotely sensed images: From pixel-based to object-based approaches. *ISPRS J. Photogramm. Remote Sens.* **2013**, *80*, 91–106. [CrossRef]

9.   Caridade, C.M.R.; Marçal, A.R.S.; Mendonça, T. The use of texture for image classification of black and white air photographs. *Int. J. Remote Sens.* **2008**, *29*, 593–607. [CrossRef]

10. Liu, W.; Yamazaki, F. Extraction of collapsed buildings in the 2016 Kumamoto earthquake using multi-temporal PALSAR-2 data. *J. Disaster Res.* **2017**, *12*, 241–250. [CrossRef]

11. Matsuoka, M.; Yamazaki, F. Use of satellite SAR intensity imagery for detecting building areas damaged due to earthquakes. *Earthq. Spectra* **2004**, *20*, 975–994. [CrossRef]

12. Bovolo, F.; Bruzzone, L. A Split-Based Approach to Unsupervised Change Detection in Large-Size Multitemporal Images: Application to Tsunami-Damage Assessment. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 1658–1669. [CrossRef]

13. Sghaier, M.; Hammami, I.; Foucher, S.; Lepage, R. Flood Extent Mapping from Time-Series SAR Images Based on Texture Analysis and Data Fusion. *Remote Sens.* **2018**, *10*, 237. [CrossRef]

14. Karimzadeh, S.; Matsuoka, M.; Miyajima, M.; Adriano, B.; Fallahi, A.; Karashi, J. Sequential SAR Coherence Method for the Monitoring of Buildings in Sarpole-Zahab, Iran Sadra. *Remote Sens.* **2018**, *10*, 1255. [CrossRef]

15. Liu, W.; Yamazaki, F.; Gokon, H.; Koshimura, S.I. Extraction of tsunami-flooded areas and damaged buildings in the 2011 Tohoku-oki earthquake from TerraSAR-X intensity images. *Earthq. Spectra* **2013**, *29*, S183–S200. [CrossRef]

16. Uprety, P.; Yamazaki, F.; Dell'Acqua, F. Damage detection using high-resolution SAR imagery in the 2009 L'Aquila, Italy, earthquake. *Earthq. Spectra* **2013**, *29*, 1521–1535. [CrossRef]

17. Wieland, M.; Liu, W.; Yamazaki, F. Learning Change from Synthetic Aperture Radar Images: Performance Evaluation of a Support Vector Machine to Detect Earthquake and Tsunami-Induced Changes. *Remote Sens.* **2016**, *8*, 792. [CrossRef]

18. Bai, Y.; Adriano, B.; Mas, E.; Koshimura, S. Machine learning based building damage mapping from the ALOS-2/PALSAR-2 SAR imagery: Case study of 2016 kumamoto earthquake. *J. Disaster Res.* **2017**, *12*, 646–655. [CrossRef]

19. Frank, J.; Rebbapragada, U.; Bialas, J.; Oommen, T.; Havens, T.C. Effect of label noise on the machine-learned classification of earthquake damage. *Remote Sens.* **2017**, *9*, 803. [CrossRef]

20. Moya, L.; Marval Perez, L.; Mas, E.; Adriano, B.; Koshimura, S.; Yamazaki, F. Novel Unsupervised Classification of Collapsed Buildings Using Satellite Imagery, Hazard Scenarios and Fragility Functions. *Remote Sens.* **2018**, *10*, 296. [CrossRef]

21. Dong, L.; Shan, J. A comprehensive review of earthquake-induced building damage detection with remote sensing techniques. *ISPRS J. Photogramm. Remote Sens.* **2013**, *84*, 85–99. [CrossRef]

22. Chini, M.; Pierdicca, N.; Emery, W.J. Exploiting SAR and VHR optical images to quantify damage caused by the 2003 bam earthquake. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 145–152. [CrossRef]

23. Yamazaki, F.; Yano, Y.; Matsuoka, M. Visual damage interpretation of buildings in Bam city using QuickBird images following the 2003 Bam, Iran, earthquake. *Earthq. Spectra* **2005**, *21*, 1–5. [CrossRef]

24. Miura, H.; Midorikawa, S.; Matsuoka, M. Building damage assessment using high-resolution satellite SAR images of the 2010 Haiti earthquake. *Earthq. Spectra* **2016**, *32*, 591–610. [CrossRef]

25. Moya, L.; Yamazaki, F.; Liu, W.; Yamada, M. Detection of Collapsed Buildings from Lidar data due to the 2016 Kumamoto Earthquake in Japan. *Nat. Hazards Earth Syst. Sci.* **2018**, *18*, 65–78. [CrossRef]

26. Yamazaki, F.; Iwasaki, Y.; Liu, W.; Nonaka, T.; Sasagawa, T. Detection of damage to building side-walls in the 2011 Tohoku , Japan earthquake using high-resolution TerraSAR-X images. *Int. Soc. Opt. Eng.* **2013**, *8892*, 1–9.

27. Gokon, H.; Post, J.; Stein, E.; Martinis, S.; Twele, A.; Mück, M.; Geiß, C.; Koshimura, S.; Matsuoka, M. A Method for Detecting Buildings Destroyed by the 2011 Tohoku Earthquake and Tsunami Using Multitemporal TerraSAR-X Data. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1277–1281. [CrossRef]

28. Uprety, P.; Yamazaki, F. Use of high-resolution SAR intensity images for damage detection from the 2010 Haiti earthquake. *Int. Geosci. Remote Sens. Symp.* **2012**, *2009*, 6829–6832.

29. Matsuoka, M.; Yamazaki, F. Application of the damage detection method using SAR intensity images to recent earthquakes. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Toronto, ON, Canada, 24–28 June 2002; Volume 4, pp. 2042–2044.

30. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 247–259. [CrossRef]

31. Ministry of Land, Infrastructure, Transport and Tourism (MLIT). Results of the Survey on Disaster Caused by the Great East Japan Earthquake (First Report). Available online: http://www.mlit.go.jp/report/press/city07_hh_000053.html (accessed on 4 April 2018).

32. Breit, H.; Fritz, T.; Balss, U.; Lachaise, M.; Niedermeier, A.; Vonavka, M. TerraSAR-X SAR processing and products. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 727–740. [CrossRef]

33. Vapnik, V. An Overview of Statistical Learning Theory. *IEEE Trans. Neural Netw.* **1999**, *10*, 988–999. [CrossRef] [PubMed]

34. Hsu, C.; Lin, C. A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Netw.* **2002**, *13*, 415–425. [PubMed]

35. Hsu, C.; Chang, C.; Lin, C. *A Practical Guide to Support Vector Classification*; Department of Computer Science, National Taiwan University: Taipei, Taiwan, 2010.

36. Fawcett, T. An introduction to ROC analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [CrossRef]

37. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* **2009**, *45*, 427–437. [CrossRef]

38. Heremans, R.; Wiilekens, A.; Borghys, D.; Verbeeck, B.; Valckenborgh, J.; Acheroy, M.; Perneel, C. Automatic detection of flooded areas on ENVISAT/ASAR images using an object-oriented classification technique and an active contour algorithm. In Proceedings of the International Conference on Recent Advances in Space Technologies, RAST '03, Istanbul, Turkey, 20–22 November 2003; pp. 311–316.

39. Nakmuenwai, P.; Yamazaki, F.; Liu, W. Automated extraction of inundated areas from multi-temporal dual-polarization radarsat-2 images of the 2011 central Thailand flood. *Remote Sens.* **2017**, *9*, 78. [CrossRef]

40. Mas, E.; Bricker, J.; Kure, S.; Adriano, B.; Yi, C.; Suppasri, A.; Koshimura, S. Field survey report and satellite image interpretation of the 2013 Super Typhoon Haiyan in the Philippines. *Nat. Hazards Earth Syst. Sci.* **2015**, *15*, 805–816. [CrossRef]