

Article

Identifying Collapsed Buildings Using Post-Earthquake Satellite Imagery and Convolutional Neural Networks: A Case Study of the 2010 Haiti Earthquake

Min Ji, Lanfa Liu *  and Manfred Buchroithner 

Institute for Cartography, TU Dresden, 01062 Dresden, Germany; Min.Ji@tu-dresden.de (M.J.); Manfred.Buchroithner@tu-dresden.de (M.B.)

* Correspondence: Lanfa.Liu@outlook.com; Tel.: +49-0351-4633-2860

Received: 24 August 2018; Accepted: 24 October 2018; Published: 26 October 2018



Abstract: Earthquake is one of the most devastating natural disasters that threaten human life. It is vital to retrieve the building damage status for planning rescue and reconstruction after an earthquake. In cases when the number of completely collapsed buildings is far less than intact or less-affected buildings (e.g., the 2010 Haiti earthquake), it is difficult for the classifier to learn the minority class samples, due to the imbalance learning problem. In this study, the convolutional neural network (CNN) was utilized to identify collapsed buildings from post-event satellite imagery with the proposed workflow. Producer accuracy (PA), user accuracy (UA), overall accuracy (OA), and Kappa were used as evaluation metrics. To overcome the imbalance problem, random over-sampling, random under-sampling, and cost-sensitive methods were tested on selected test A and test B regions. The results demonstrated that the building collapsed information can be retrieved by using post-event imagery. SqueezeNet performed well in classifying collapsed and non-collapsed buildings, and achieved an average OA of 78.6% for the two test regions. After balancing steps, the average Kappa value was improved from 41.6% to 44.8% with the cost-sensitive approach. Moreover, the cost-sensitive method showed a better performance on discriminating collapsed buildings, with a PA value of 51.2% for test A and 61.1% for test B. Therefore, a suitable balancing method should be considered when facing imbalance dataset to retrieve the distribution of collapsed buildings.

Keywords: building damage assessment; convolutional neural networks (CNNs); earthquake; Haiti; SqueezeNet

1. Introduction

With the advance of sensor and space technology, remote sensing is able to obtain detailed temporal and spatial information at the target area, and has been widely used to detect, identify, and monitor the effect of natural disasters [1,2]. It has been adopted in various post-earthquake activities as remotely sensed images usually need minimal fieldworks, which is especially important to earthquake-affected areas that are difficult to access [3]. Building damage information is key to post-earthquake rescue and reconstruction. It has been demonstrated that remotely sensed data are capable to derive relatively accurate building damage information [4]. High-resolution remote sensing imageries are able to generate building-by-building damage maps by interpreting their damage states [5–7].

Building damage can be detected by using only post-event data with the help of the emergence of very high resolution (VHR) remote sensing imagery, which can provide detailed textural and spatial features of the damaged targets [8]. A wide range of remote sensing techniques is applicable to evaluate

post-earthquake damage, including optical satellite imagery, synthetic aperture radar (SAR), and light detection and ranging (LiDAR). With the rapid improvement of the spatial resolution of satellite optical sensors (such as WorldView-4, that has a GSD of 0.31 m in the panchromatic band), the utilization of optical data is a promising approach to detect earthquake damage. Visual interpretation, edge and textures, and spectral properties have been used to detect building damage when the post-event optical data are available. The distribution of damaged buildings was visually delineated using post-event optical images to support early emergency and rescue planning [9]. The semi-automated approach was applied to identify the region damage using spectral and textural information from an optical image after the earthquake [10]. The building damage was detected using the watershed segmentation of the post-event aerial images, assuming their shape information is available as a stored geographic information system (GIS) layer [11]. A refined example, automated damage detection method using optical data, was discussed in [12,13]. The advantage of using SAR data to assess damaged buildings is its independence from sun illumination and relative insensitivity to atmospheric conditions [14]. Compared with combined pre- with post-event data [15], single post-event PolSAR data is quicker and more convenient to assess building damage [16]. The potential of post-event SAR images has been demonstrated for building damage assessment after earthquakes [17–19]. LiDAR can provide a three-dimensional visualization of the damaged area, which is useful to automatically generate a damage map [20]. In addition, LiDAR has the ability to work at day and night, even in adverse conditions, such as in poor illumination or through clouds and smoke. A number of studies have used post-event LiDAR imagery to detect building damage [20–23]. There are also some studies that combined optical imagery with SAR imagery. Different damage types have been analyzed in [24] using post-event TerraSAR-X Spot-Light VHR SAR image and optical images as assistance to facilitate the analysis. To validate and analyze the results, a validation map was created based on optical imagery, and the result demonstrated that SAR data have potential for application in urban disaster monitoring and assessment [25].

Automatic and visual methods are common approaches to generate building damage maps using satellite or aerial imageries [26]. However, the visual method based on the study of manual sampling is time-consuming, which is disadvantageous for planning rescue [27]. By contrast, automatic methods are able to derive change information from satellite images efficiently. The classification model created by supervised learning can predict the class of other unclassified instances automatically, once the model is generated [28]. Nevertheless, it also requires time and effort to prepare a few training samples for supervised methods, which is a disadvantage for rapid damage assessment. Numerous scholars have paid attention to the use of machine learning methods to detect damaged buildings using post-event datasets, and carried out much fruitful work. The feature extraction was conducted by morphological profiles and texture statistics, then collapsed buildings were classified using support vector machine (SVM) [29]. The collapsed buildings were detected by methods based on object-based image analysis (OBIA), and SVM using post-event LiDAR data [30]. A support vector selection and adaptation (SVSA) method was applied to two small regions and the entire city of Port-au-Prince (Haiti), to assess the damage using the post-event satellite images [31]. A variety of algorithms and parameters were tested on post-event aerial imagery for the earthquake in Christchurch, New Zealand, and the results showed that object-based approaches can produce better results than pixel-based approaches in earthquake damage detection using remotely sensed images [32]. Random forest (RF), SVM, and *K*-nearest neighbor (*K*-NN) classifiers were applied to classify collapsed and standing buildings with the post-event SAR image and the building footprint map [33].

Convolutional neural networks (CNNs) have become hot research topics in the field of image recognition and speech analysis in recent years. A CNN is an alternative type of neural network architecture that can be used to model spatial and temporal correlations [34–36]. It can reduce the complexity of network models and the number of weights, due to its weight-sharing characteristic, which makes it more similar to a biological neural network. It has a high invariance in translation, scaling, incline. There are many kinds of CNNs, such as AlexNet [37], VGGNet [38],

and GoogleNet [39]. SqueezeNet [40,41] was developed by Forrest Iandola, and it can acquire the same accuracy of AlexNet-level on ImageNet, with 50 times fewer parameters. However, there are still limited studies using CNNs to obtain the damage information in earthquake-affected areas. In a recent study, deep learning was explored for building damage detection caused by earthquakes using oblique aerial images [42]. It demonstrated that CNN features performed better than 3D point cloud features using a multiple-kernel-learning approach for detecting damaged regions using VHR images. Besides, there is a significant challenge for remotely sensed imagery analysis due to the highly imbalanced class distribution [43,44]. To handle the imbalanced classification problems, various methodologies have been proposed, such as resampling, modification on classifier optimization problem, or introducing a new optimization task on top of the classifier [45]. A number of studies have been aimed to deal with imbalanced datasets acquired from remote sensing images. Infinitely imbalanced logistic regression (ILLR) was proposed to deal with remote sensing datasets [46]. The oil spills were able to be detected by applying the one-sided selection method (OSS) and satellite radar images [47]. To deal with the imbalance problem in the convolutional neural network, seven approaches were compared in [48], including random over-sampling, random under-sampling, and thresholding with prior class probabilities.

The objective of this study was to explore the performance of SqueezeNet on identifying collapsed buildings using single post-earthquake VHR satellite data. Completely collapsed buildings can be readily identified from disintegrated roof structures and associated texture features from VHR imagery, while lower damage grades are much harder to map, as such damage effects are largely expressed along the facade, which are not visible in such imagery. The dataset obtained after the 2010 Haiti earthquake was used in this study. As the distribution of building damage grades was imbalanced, three balancing methods were adopted to improve the accuracy of identifying collapsed buildings including random over-sampling, random under-sampling, and cost-sensitive approaches. The rest of this paper is organized as follows. Section 2 provides the descriptions of the study area. Section 3 briefly introduces basic concepts of convolutional neural networks, data balancing methods, and the metrics used to evaluate the performance. The workflow of using SqueezeNet to classify collapsed buildings caused by the earthquake is also included. Section 4 presents the experimental results and discusses the methodology used in the experiments. Finally, conclusions are drawn in Section 5.

2. Input Data

Massive buildings and infrastructures were damaged, and some of them even completely collapsed, after an earthquake struck Haiti on 12 January 2010. It was said that more than 300,000 people lost their lives, and about 105,000 houses were completely destroyed in Haiti earthquake [49]. The location of the study area is shown in Figure 1. Post-earthquake satellite images were captured on 15 January 2010 by the QuickBird satellite. The data were obtained via DigitalGlobe open data program having a resampled spatial resolution of 0.5 m, and the NIR band was not included. The building damage information was visually interpreted from high-resolution satellite images and aerial photos by UNITAR/UNOSAT [50]. The building damage level was classified into five categories based on the EMS-98 [51]; G5: Destruction; G4: Very heavy damage; G3: Substantial damage; G2: Moderate damage and G1: Negligible damage. Selected examples of damaged buildings were shown in Figure 2. Building footprints were manually extracted in the study area using ArcGIS 10.4. To train and validate the proposed method for collapsed building identification, the study area was further separated into three regions: train, test A and test B. In this study, the number of collapsed and non-collapsed buildings is 613 and 1857 for the training region, 129 and 454 for test A, and 322 and 553 for test B.

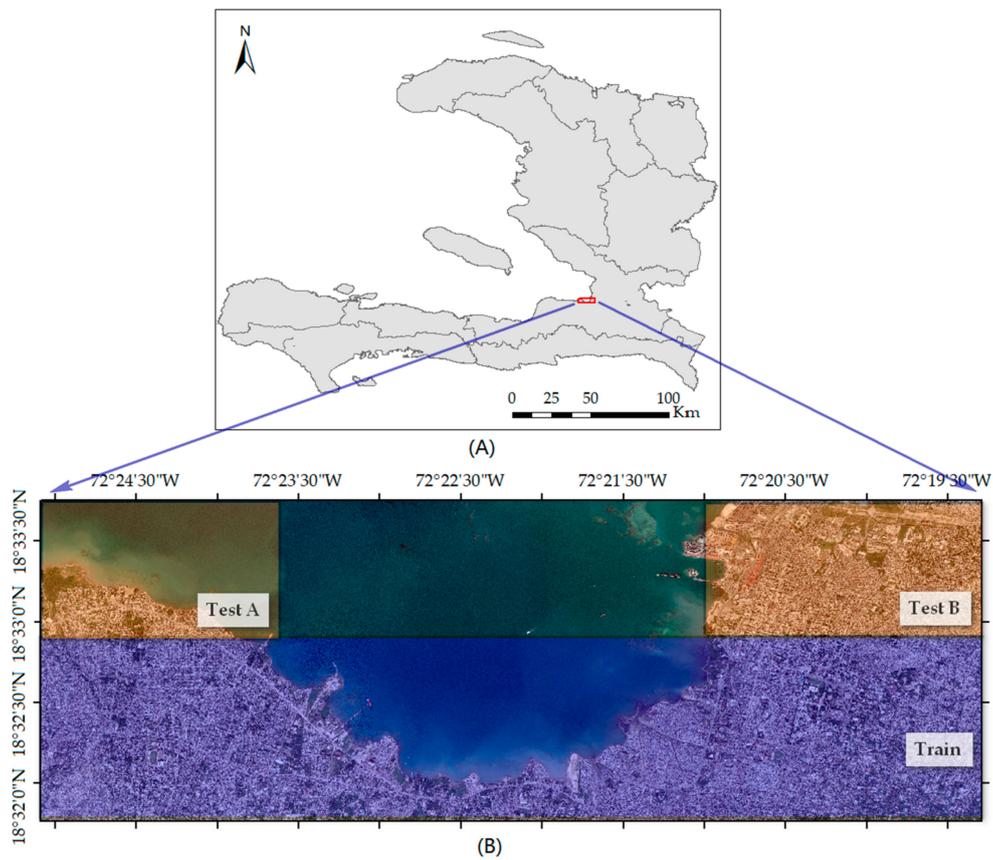


Figure 1. Location of the study area and regions for train and test datasets.

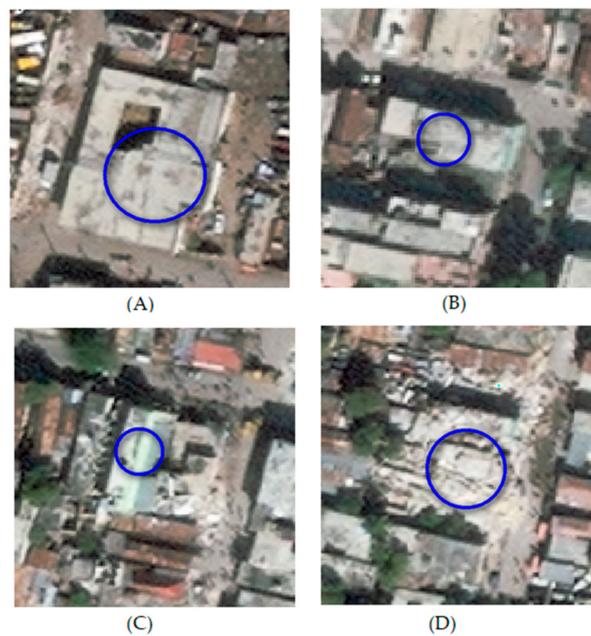


Figure 2. Examples of damaged buildings caused by the Haiti earthquake. (A) Grade 1; (B) Grade 3; (C) Grade 4; (D) Grade 5. Grade 2 was not presented in the dataset.

3. Methodology

In this study, a CNN-based approach was proposed for building collapsed assessment after the earthquake. The workflow can be seen in Figure 3. To make use of VHR satellite imagery,

a decomposition method should be used to split the large image into small processing patches [52]. Small building patches were extracted from satellite images according to the building boundary polygons. However, building patches have a different number of width and length pixels, which makes them unsuitable as inputs for the CNNs. We adopted a zero-padding operation to make the small building patches uniform, while discarding building patches with the width or length smaller than 10 pixels or larger than 96 pixels. The damage grades were reclassified into binary categories: collapsed (G5) and non-collapsed (G1–G4), which will be used as the labels for the corresponding building patches for further analysis. Non-collapsed buildings (2864) outnumbered collapsed buildings (1064), which caused an imbalance problem and may affect the classification results. Three balancing methods were considered and compared. Finally, a building collapsed map can be derived with the proposed workflow.

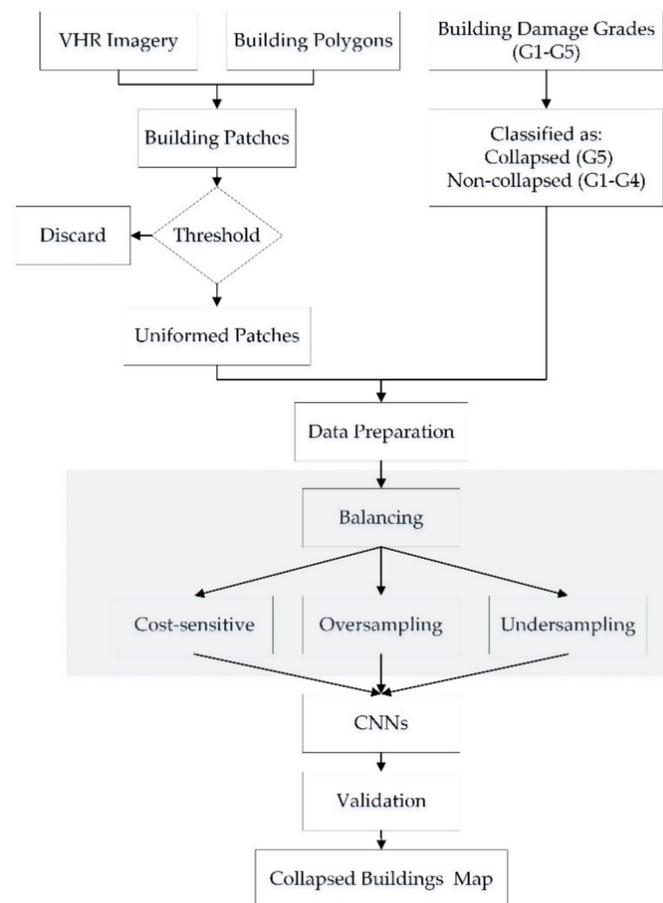


Figure 3. Workflow of mapping collapsed buildings using very high resolution (VHR) imagery and convolutional neural networks (CNNs).

3.1. Convolutional Neural Networks (CNNs)

CNNs can be viewed as multilayer neural networks, in which shift and distortion invariance can be ensured due to CNNs' special architectures: local receptive fields, shared weights and, sometimes, spatial or temporal subsampling. A typical CNN structure normally contains convolution, pooling, and activation function layers, as shown in Figure 4. The first layer represents input data, while the second layer means feature maps after the convolution process. The third layer contains features of activation maps after the process of the activation function. The fourth layer is the pooled feature map after the pooling process. Red squares in the figure represent filters, and the latter square is the output of the former one after the corresponding operation (convolution, ReLU activation, and pooling). The convolutional layer is meant to extract features, and filter weights can be shared across all pixels.

The spatial variation and correlation will be reduced in convolutional layers. There are many kinds of activation functions, such as sigmoid, tanh, and rectified linear unit (ReLU). ReLU activation function is able to avoid the vanishing gradient, and has less computation than tanh and sigmoid, as it involves simpler mathematical operations [37]. In the nonlinearity layer, ReLU is applied to each component in a feature map, as shown in Equation (1), in which x means the input to the activation layer. It is a half-wave rectifier function, which can significantly accelerate the training phase and prevent overfitting.

$$f(x) = \max(x, 0) \quad (1)$$

The main function of the pooling layer is to compress the feature graphs and reduce the dimensionality [53]. Common methods are to maximize or average the input values. The pooling layer can be viewed as down-sampling of the convolutional feature map [54]. A max operation is implemented over a small region G of each feature map.

$$P = \max_{i \in G}(f(x_i)) \quad (2)$$

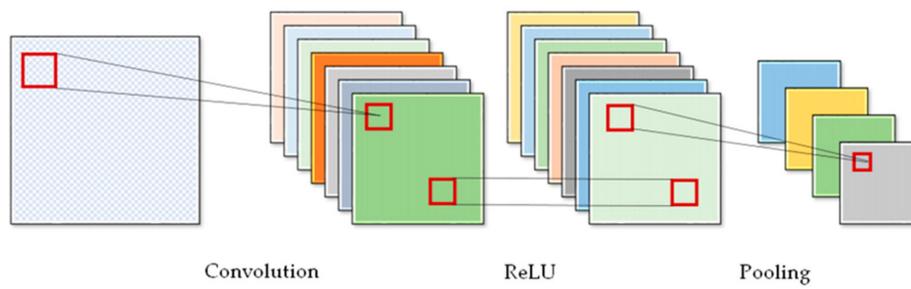


Figure 4. Illustration of typical layers of a CNN [55].

3.2. SqueezeNet

A small CNN architecture named SqueezeNet was proposed in 2016. Compared to AlexNet, SqueezeNet can get a similar accuracy of classification with $50\times$ fewer coefficients by using a compression methodology, and was proven on the ImageNet database [40]. The design goal of SqueezeNet is not to get the best CNN recognition accuracy, but to simplify the network complexity and attain the recognition accuracy of the public network. There are three main strategies in SqueezeNet architecture. For the first strategy, a fire module was proposed based on the use of 1×1 filters instead of 3×3 filters, as 1×1 filters have $9\times$ fewer parameters than 3×3 filters. For the second strategy, a squeeze layer was applied to decrease the number of input channels to 3×3 filters, instead of the 11×11 filters adopted by AlexNet. The first and second strategies are designed to reduce the number of parameters in a CNN, while maintaining similar inference accuracy. The last strategy is to get large activation maps in convolution layers and maximize the accuracy. For SqueezeNet, the stride for the first convolutional layer is 2×2 instead of 4×4 for AlexNet. As if early layers in the network have small strides, the following layers will have large activation maps.

The Fire module is the basic building block of SqueezeNet architecture, as shown in Figure 5. It consists of a squeeze convolution layer with 1×1 filters feeding an expanding layer with 1×1 and 3×3 filters. The number of filters per Fire module is gradually increased, from the beginning to the end of the network. Relu [56] is functioned as the activation function in all Fire modules. The SqueezeNet architecture utilizes global-average-pooling layer to replace the full-connected layer, which is easier to interpret and less prone to overfitting than full-connected layers.

The structure of the adopted CNN model is listed in Table 1. The input layer expected building patches with width, length, and band values as (96, 96, 3) followed by a convolutional layer with 64 filter kernels and a stride of 2×2 . Three Fire modules were adopted from SqueezeNet when

constructing the CNN model. A max-pooling operation was inserted between the fire modules. A dropout layer was also considered after the Fire modules. Finally, a global-average-pooling layer was used to replace the conventional flatten layer, followed by a softmax layer to classify if the input building collapsed or not. The model has a total number of 164,194 parameters, which were trained with the mentioned training dataset. The CNN model was implemented using Keras 2.1.5 with tensorflow 1.8 as backend. Computing was done using Google Cloud Platform with NVIDIA Tesla K80 GPU and 26 gigabyte memory.

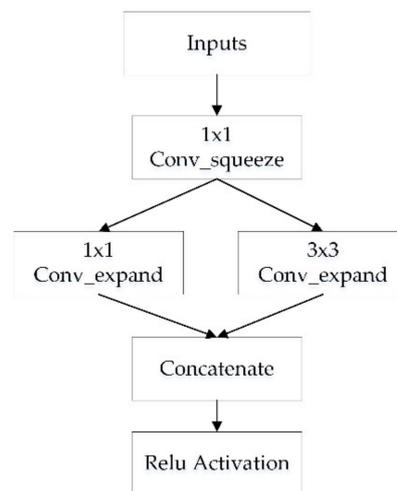


Figure 5. The structure of the Fire module used in SqueezeNet.

Table 1. The CNN structure adopted in this study.

	Layer	Shape (N, Width, Length, Bands)	Nr. of Parameters
	Input	(N, 96, 96, 3)	0
	Conv2D	(N, 47, 47, 64)	1792
	Relu	(N, 47, 47, 64)	0
	MaxPooling2D	(N, 23, 23, 64)	0
Fire Module	Conv2D_squeeze1 × 1	(N, 23, 23, 16)	1040
	Relu_squeeze1 × 1	(N, 23, 23, 16)	0
	Conv2D_expand1 × 1	(N, 23, 23, 64)	1088
	Conv2D_expand3 × 3	(N, 23, 23, 64)	9280
	Relu_expand1 × 1	(N, 23, 23, 64)	0
	Relu_expand3 × 3	(N, 23, 23, 64)	0
	Concatenate	(N, 23, 23, 128)	0
	MaxPooling2D	(N, 11, 11, 128)	0
	Conv2D_squeeze1 × 1	(N, 11, 11, 32)	4128
Fire Module	Relu_squeeze1 × 1	(N, 11, 11, 32)	0
	Conv2D_expand1 × 1	(N, 11, 11, 128)	4224
	Conv2D_expand3 × 3	(N, 11, 11, 128)	36,992
	Relu_expand1 × 1	(N, 11, 11, 128)	0
	Relu_expand3 × 3	(N, 11, 11, 128)	0
	Concatenate	(N, 11, 11, 256)	0
	MaxPooling2D	(N, 5, 5, 256)	0
	Conv2D_squeeze1 × 1	(N, 5, 5, 48)	12,336
	Relu_squeeze1 × 1	(N, 5, 5, 48)	0
Fire Module	Conv2D_expand1 × 1	(N, 5, 5, 192)	9408
	Conv2D_expand3 × 3	(N, 5, 5, 192)	83,136
	Relu_expand1 × 1	(N, 5, 5, 192)	0
	Relu_expand3 × 3	(N, 5, 5, 192)	0
	Concatenate	(N, 5, 5, 384)	0
	Dropout	(N, 5, 5, 384)	0
	Conv2D	(N, 5, 5, 2)	770
	Relu	(N, 5, 5, 2)	0
	Global-average-pooling	(N, 2)	0
	Softmax	(N, 2)	0
Total	-	-	164,194

3.3. Data Balancing Methods

There are mainly three approaches to deal with class imbalance problems, which can be classified as data-level, algorithm-level, and hybrid methods [57]. Data-level methods mainly use re-sampling methods to balance the class distribution in the training data. Random over-sampling methods increase the number of samples in the minority class by randomly replicating or generating a new minority. As opposed to the random over-sampling, majority class instances are randomly eliminated by the random under-sampling method, to balance class distribution until the minority and majority have the same number of instances. Algorithm methods modify existing classification algorithms to improve the sensitivity of the classifier towards minority classes. One of the most popular algorithm-level methods is the cost-sensitive approach [58], which assigns different cost to samples from different classes. In this study, we simply use the class proportion as the loss weight for different classes. The minority class samples have higher costs, thus giving them greater impact on the weight-updating in the neural network [59]. The hybrid method integrates previously mentioned approaches to improve the performance [60], which was not considered in this study.

3.4. Evaluation Metrics

The selected Haiti dataset was separated into the train, test A, and test B regions, as shown in Figure 1. Several metrics are used as the evaluation standards in this study, including producer accuracy (PA), user accuracy (UA), overall accuracy (OA), and Kappa, which are proposed based on the confusion matrix (Table 2). True positive (TP) is the number of positive examples correctly classified, false positive (FP) is the number of negative examples incorrectly classified as positive, false negative (FN) is the number of positive examples incorrectly classified as negative, and true negative (TN) is the number of negative examples correctly classified. OA (Equation (3)) is the percentage of examples correctly classified. OA is often used to measure the performance of learning systems. However, it is not appropriate when the dataset is imbalanced, since it tends to be biased toward the majority class while neglecting the minority class. PA is the probability that a value in a given class was classified correctly. UA is the probability that a value predicted to be in a certain class really is that class. The probability is based on the fraction of correctly predicted values to the total number of values predicted to be in a class. The Kappa coefficient of agreement developed by Cohen (1960) is a statistical measure of inter-rater agreement for categorical items. It can be calculated by Equation (4). P_0 is the observed proportion of agreement, and P_e is the proportion of agreement expected by chance. When dealing with imbalance dataset, it is important to pay attention not only to the overall accuracies but, also, the corresponding misclassification costs. Thus, Kappa would be a better performance measure than the OA when facing an imbalanced dataset. Kappa coefficients are interpreted using the guidelines outlined by Landis and Koch [61], who characterized values between 0.01 and 0.20 as slight, between 0.21 and 0.40 as fair, between 0.41 and 0.60 as moderate, between 0.61 and 0.80 as substantial, and between 0.81 and 1.00 as almost perfect.

$$OA = \frac{TP + TN}{TP + FP + FN + TN} \quad (3)$$

$$Kappa = \frac{P_0 - P_e}{1 - P_e} = \frac{\frac{(TN+TP)}{n} - \frac{f_1 * g_1 + f_2 * g_2}{n^2}}{1 - \frac{f_1 * f_2 + g_1 * g_2}{n^2}} \quad (4)$$

Table 2. Confusion matrix.

		Ground Truth		
		Collapsed	Non-Collapsed	Total
Predicted	Collapsed	True Positive (TP)	False Positive (FP)	g_1
	Non-collapsed	False Negative (FN)	True Negative (TN)	g_2
		Total		n
		f_1	f_2	

4. Results and Discussion

4.1. Identifying Collapsed Buildings Using CNNs

The number of width and length pixels for extracted building patches are different. The width and length vary from several to hundreds of image pixels. The length of the non-collapsed buildings ranges from 10 to 276 pixels, and the width ranges from 9 to 413 pixels. For collapsed buildings, the length falls within the range of 7–194 pixels, and the width within the range of 11–438 pixels. The distributions for collapsed and non-collapsed buildings have similar trends, mainly ranging from 20 to 40 pixels for both width and length, with long tails, as can be seen in Figure 6. Typical CNNs require fixed-size inputs, as pointed out by [62]. Convolutional layers do not require a fixed image size, but the fully connected layer needs to have fixed-size inputs, by definition. Although a global average pooling layer was used instead of a fully connected layer in the network, it is still difficult to implement CNNs using variable-size inputs. In this study, too large or small buildings were ignored by defining the thresholds, and the remaining buildings were padded by zero values to have the same dimensions. The width and length of building patches were limited from 10×10 to 96×96 pixels in this study. Building patches having pixel sizes outside of the mentioned range were filtered. For retained collapsed or non-collapsed buildings, a zero-padding operation was considered so that all building patches have a uniformed pixel size of 96×96 .

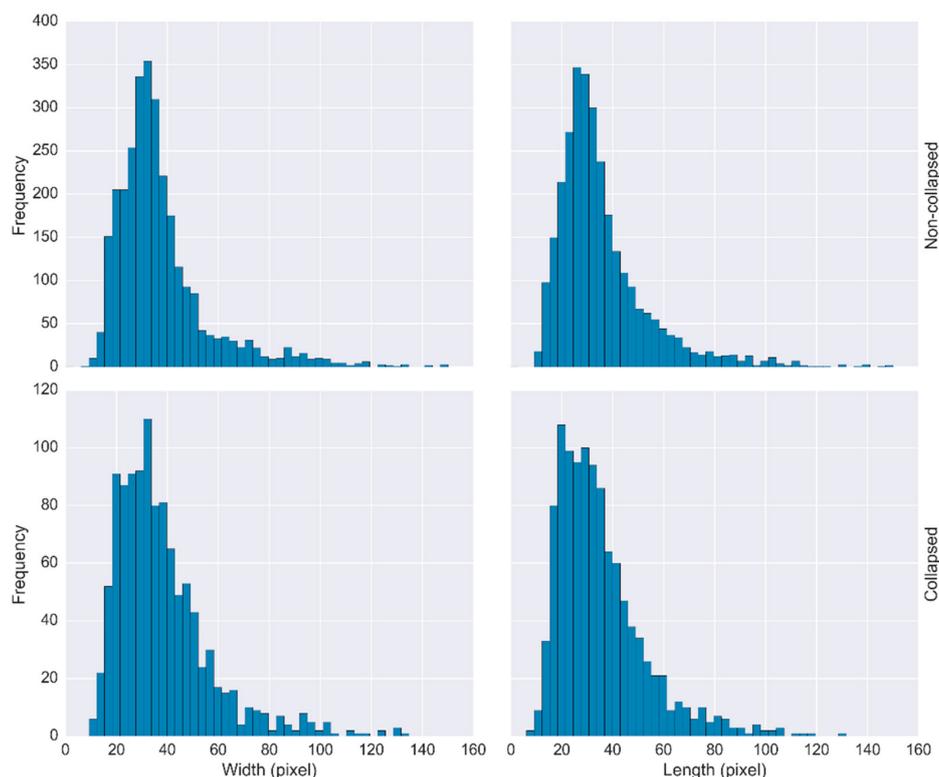


Figure 6. The distribution of width and length pixels for collapsed and non-collapsed building patches.

The results of using the mentioned CNN model can be seen in Table 3, to classify collapsed and non-collapsed buildings caused by the Haiti earthquake. SqueezeNet achieved Kappa values of 37.7% and 45.6% for discriminating collapsed and non-collapsed buildings using the post-earthquake VHR imagery on test A and test B, respectively. The PA values of non-collapsed buildings were very high (>90%) in the two test regions, which means the model performed well on classifying non-collapsed buildings. It can be seen that collapsed buildings were prone to being misclassified through the low PA values of 42.6% and 50.6%. The UA values for non-collapsed buildings are comparatively higher than

for collapsed buildings. The UA values for collapsed buildings were 58.5% and 78% for test A and test B, indicating that non-collapsed buildings were comparatively prone to being wrongly identified in test A region. One reason for the low PA value of collapsed buildings is that the classification method is prone to favor the majority class when dealing with the imbalanced dataset. The building structures should also be considered in the study area. It is prone to being correctly classified for concrete buildings having very prominent collapse or damage structures with totally broken down roofs. Steel or wooden frame buildings with metal sheet roofs, where the building was physically collapsed but there was no visible deformation or textural change to its roof structure, would be hard to correctly classify [63], which would decrease the number of correctly classified collapsed buildings.

Table 3. SqueezeNet performance on test A and test B dataset.

		Ground Truth			UA (%)	OA (%)	Kappa (%)
		Collapsed	Non-Collapsed				
Test A	Predicted	Collapsed	55	39	58.5	80.6	37.7
		Non-collapsed	74	415	84.9		
		PA (%)	42.6	91.4			
Test B	Predicted	Collapsed	163	46	78	76.6	45.6
		Non-collapsed	159	507	76.1		
		PA (%)	50.6	91.7			

PA: producer accuracy; UA: user accuracy; OA: overall accuracy.

The Kappa value for test A (37.7%) and test B (45.6%) also indicated that SqueezeNet performed better on test B to discriminate collapsed buildings, which could be partly caused by the difference in building structures in these two regions. Building structures in Test B include concrete structures with flat roofs of varying heights and sizes, wooden or steel frame buildings with corrugated metal sheet roofs, and low height metal sheet shelters (shanty housing) with very small-sized dwellings [63]. For test A, there is mainly a density of small buildings or even informal huts, as relatively poor residents lived here, and built mostly makeshift homes when the devastating earthquake struck Haiti. While such buildings were prone to be misclassified from the imagery, the extracted patches for small buildings were padded with more zero values in the pre-processing step, which will also affect the model’s performance. To demonstrate the achieved results, building-by-building evaluation maps were shown in Figure 7.

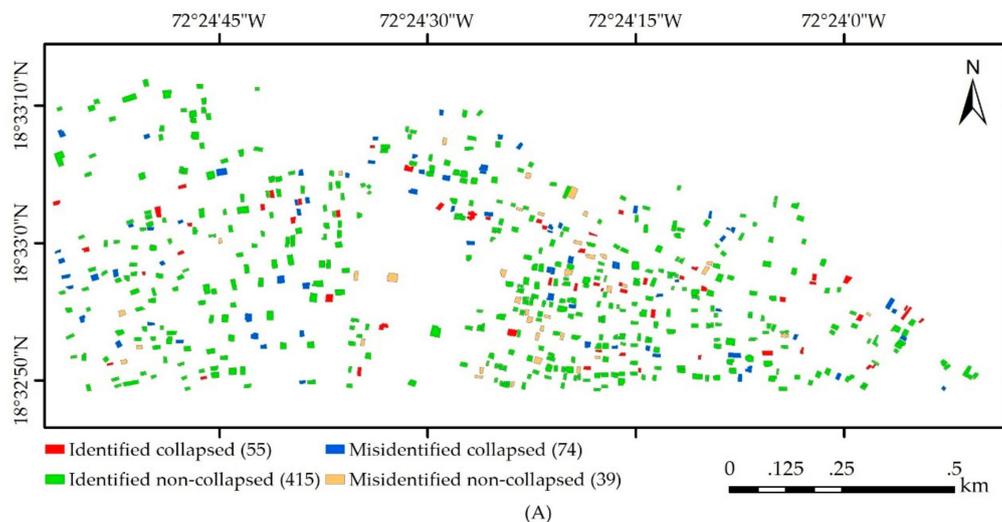


Figure 7. Cont.

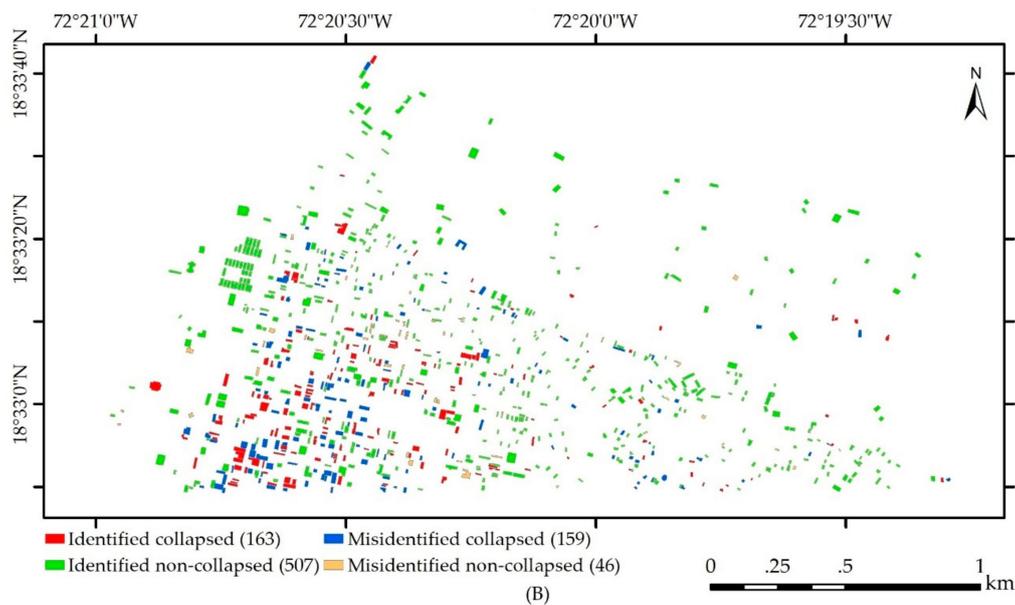


Figure 7. The performance of the CNN on identifying collapsed buildings using test A (A) and test B (B) dataset.

4.2. Performance of Balancing Methods for Identifying Collapsed Buildings

In the training dataset, the number of collapsed and non-collapsed buildings was 613 and 1857, respectively. The imbalanced distribution of training labels makes the classifier biased to the majority class, which results in the CNN model not performing well on identifying collapsed buildings compared to non-collapsed buildings. Random under-sampling, random over-sampling, and cost-sensitive methods were adopted to deal with the imbalance problem. Table 4 compared the overall performance of the three balancing methods. The highest PA values for collapsed buildings in regions of test A (61.2%) and test B (69.6%) were acquired by random under-sampling methods. Also, the UA values were lowest among these three methods, 47.6% and 65.7%, respectively. The higher PA for collapsed buildings means that more collapsed buildings were classified correctly. In addition, lower UA for collapsed buildings means larger number of FP. Random over-sampling achieved similar results with the cost-sensitive method. It can be seen that the highest OA and Kappa were acquired by the cost-sensitive method, 80.1% and 40.6% for test A, and 77.0% and 48.9% for test B. Therefore, the cost-sensitive method performed better in discriminating buildings. After balancing steps, the CNN model still achieved a better OA for test A than for test B, and the Kappa values for test B are comparatively higher. To demonstrate the achieved results, building-by-building evaluation maps were shown in Figure 8 for test A and test B, by considering the cost-sensitive method.

Table 4. The performance of balancing methods on test A and test B dataset.

Region	Method	Collapsed		Non-Collapsed		OA (%)	Kappa (%)
		PA (%)	UA (%)	PA (%)	UA (%)		
Test A	Cost-sensitive	51.2	55.4	88.3	86.4	80.1	40.6
	Random over-sampling	49.6	53.3	87.7	86.0	79.2	38.2
	Random under-sampling	61.2	47.6	80.8	88.0	76.5	38.1
Test B	Cost-sensitive	61.1	72.2	86.3	79.2	77.0	48.9
	Random over-sampling	60.9	71.7	86.1	79.1	76.8	48.5
	Random under-sampling	69.6	65.7	78.8	81.6	75.4	47.8

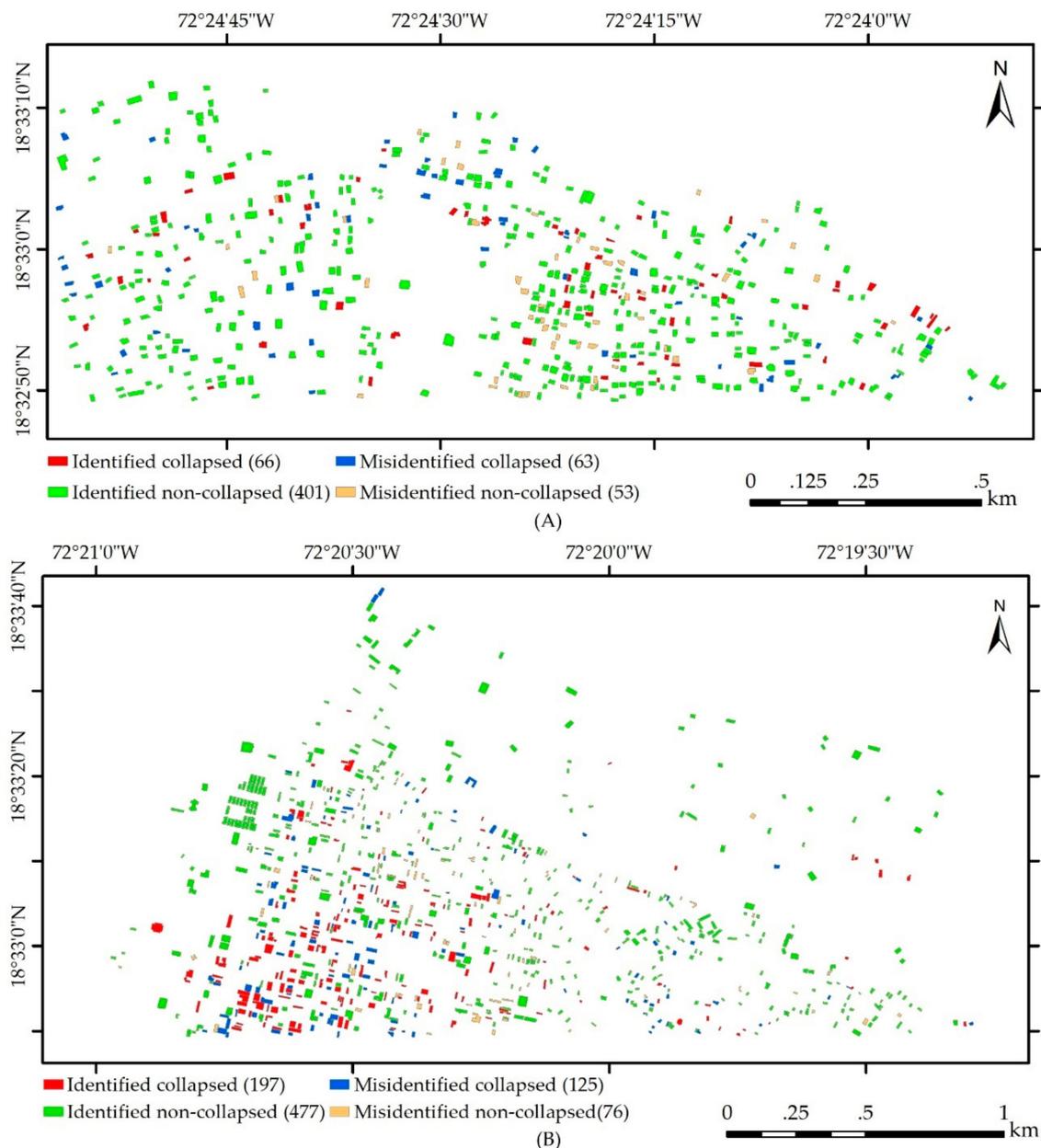


Figure 8. The performance of the balanced-CNN on identifying collapsed buildings using test A (A) and test B (B) dataset.

It can be seen that the PA values for non-collapsed buildings are higher than collapsed buildings with or without the balancing procedure. The PA values of collapsed buildings are increased after balancing, which means the balanced model has a better capability in identifying collapsed buildings, which is very important for planning rescue after an earthquake. For the random under-sampling method, it discarded a large number of non-collapsed building samples. An increasing number of collapsed buildings were correctly classified, and the PA values improved from 42.6% to 61.2% for test A, and from 50.6% to 69.6% for test B. However, the performance for non-collapsed buildings was severely damaged, thus, the OA decreased from 80.6% to 76.5% for test A, and from 76.6% to 75.4% for test B. In this study, random over-sampling and cost-sensitive methods achieved similar results, and the latter one has a slightly better performance. Although the overall accuracies did not improve with balancing methods, the Kappa values improved from 37.7% to 40.6%, and from 45.6% to 48.9% with the cost-sensitive method.

4.3. Intra-Class Analysis for Building Damage Assessment

To make the pixel size of building patches unified, a zero-padding operation was applied to the input data. If the original building patch has fewer pixel values, it will affect the model performance by padding too many zero values. To analyze the model performance on intra-class samples, test B dataset was re-classified according to the original width pixel numbers of extracted building patches. Test B dataset was used to explore the performance of the CNN and the cost-sensitive method on re-classified data. The distribution of width pixels for test B building patches was shown in Figure 9, mainly ranging from 20 to 40 pixels.

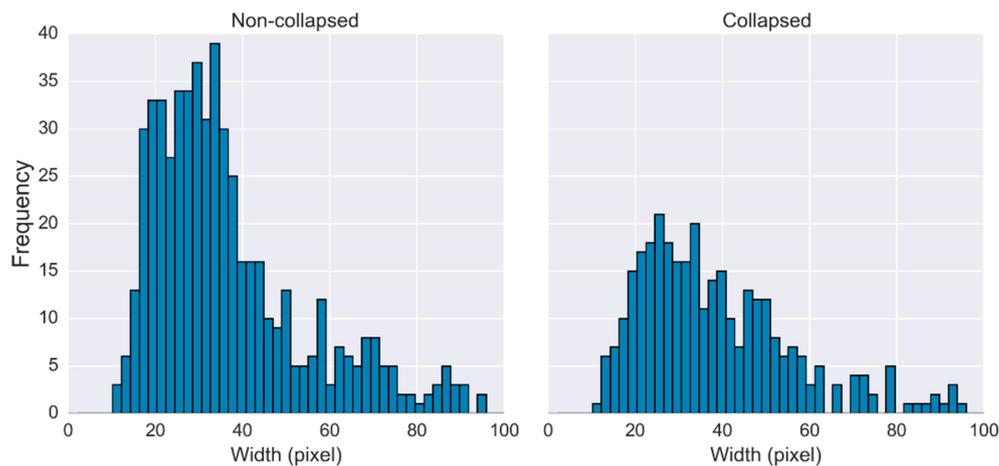


Figure 9. The distribution of width pixels for collapsed and non-collapsed building patches for test B dataset.

Building patches were classified into five categories according to the number of building width pixels, as shown in Table 5, making the number of training data as equal as possible for each category. It can be seen that Kappa values for the categories of “25–31” and “31–47” were comparatively low before balancing, and the highest Kappa was achieved for buildings with width pixels within the range of 37–46. After balancing, the Kappa values were improved for all categories, except for “<25”. For the category of building width larger than 46 pixels, the OA and Kappa values were improved from 76.3% to 80.2%, and 48.9% to 57.8%. For building patches with width pixels lower than 25, the Kappa value was decreased from 44.0% to 40.8%. While the balanced model achieved better Kappa values on building patches with width larger than 25 pixels, the balancing operation could deteriorate the performance for the left small buildings in this case study.

Table 5. SqueezeNet model performance on test B data re-classified by the number of building width pixels.

Width (Pixels)	Nr./Train	Nr./Test B	CNN		Balanced-CNN	
			OA (%)	Kappa (%)	OA (%)	Kappa (%)
<25	515	248	75.8	44.0	72.2	40.8
25–31	500	157	76.4	39.7	78.3	48.7
31–37	529	135	77.7	39.8	79.2	45.4
37–46	456	128	77.3	51.7	77.3	52.6
>46	470	207	76.3	48.9	80.2	57.8

The confusion matrix values obtained using re-classified test B dataset were plotted on Figure 10. TN has the highest values among the confusion matrix, as non-collapsed buildings outnumbered collapsed buildings, and the classifier performed well on classifying non-collapsed buildings. The aim of this study is to identify collapsed buildings, which plays a key role in post-event rescue and

reconstruction. After balancing, the number of TP samples increased, showing that the balanced classifier had a better performance on identifying collapsed buildings. However, the TP values were still relatively low. One reason is that only post-event data was considered in this study. The accuracy would be better when pre-earthquake data and LiDAR data were considered. The zero-padding pre-processing operation also affected the performance on small buildings. Furthermore, the number of training samples was still too small. Considering that it is not easy to prepare large number samples for such a task, transfer learning based on CNN could be considered, by fine-tuning large-scale dataset-derived complex CNN models with relatively less new data.

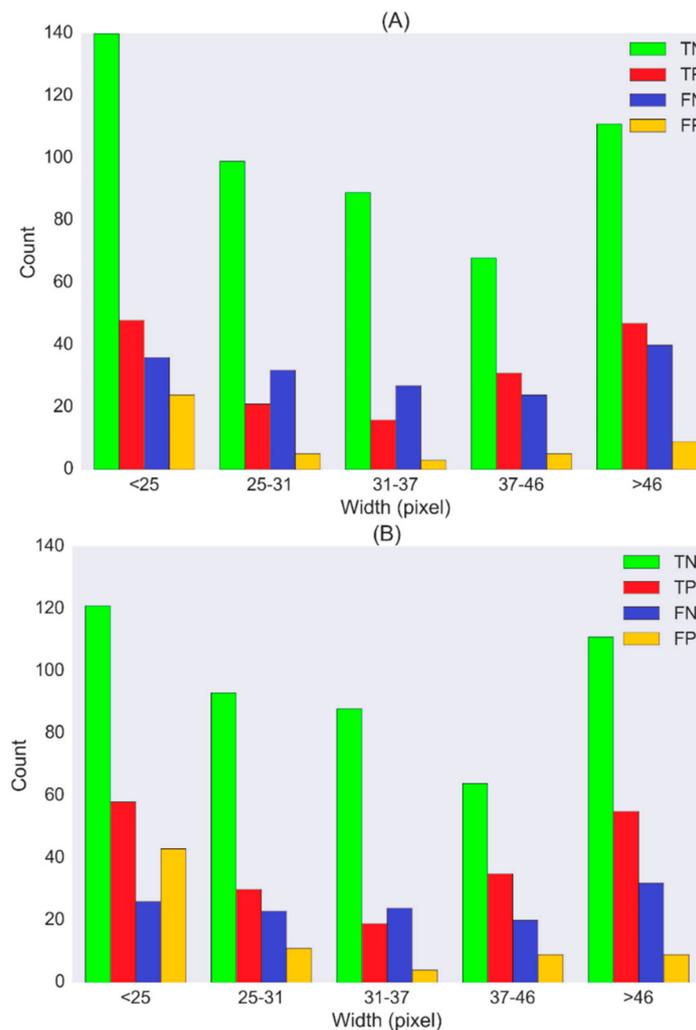


Figure 10. Plots of the confusion matrix values obtained using re-classified test B dataset: (A) CNN (B) Balanced-CNN. FP (non-collapsed buildings misclassified as collapsed ones); FN (collapsed buildings misclassified as non-collapsed ones); TP (collapsed buildings classified correctly); TN (non-collapsed buildings classified correctly).

The model was prone to favor collapsed buildings after balancing, which would increase the number of misclassified small buildings for non-collapsed buildings. It can be observed from Figure 10 that the number of false positives was increased, along with the increase of the number of true positives for buildings with width pixels smaller than 46. Similar results were also observed in [64]. This is a disadvantage and undesirable for rapid damage assessment. It is demonstrated that the number of false positives could be reduced by taking advantage of ensemble learning [65]. However, it is still a challenge to properly deal with an imbalanced dataset [57]. Besides, the building structures in the

study area should be considered when using balancing methods for identifying collapsed buildings after an earthquake, which could deteriorate the performance for small buildings.

4.4. CNNs for Identifying Earthquake-Induced Collapsed Buildings

There are several existing studies using VHR satellite imagery for mapping earthquake-induced collapsed or affected buildings after the 2010 Haiti earthquake. Texture and structure features were derived from pre- and post-earthquake VHR satellite imagery for the city of Port-au-Prince (Haiti), and obtained overall accuracies of 74.1–77.3% and Kappa values of 30.6–40.2% using artificial neural networks (ANN), radial basis function neural network (RBFNN) and RF [1]. A support vector selection and adaptation (SVSA) approach was carried out, to classify the post-earthquake QuickBird data into eight land-use classes, and 92 damaged buildings were correctly identified from the total 145 damaged samples [31]. The road and building classes were confused with the damage class due to pixel-based classification. A moderate result was achieved in this study using the CNN approach, which can take advantage of features learnt via the trained network, and requires no extra feature extraction. However, the learnt features are difficult for interpretation. When the post-event LiDAR point cloud data were considered, a better result was achieved in [66] by combining spectral, texture, and height information, and implemented at the object level. A one-class support vector machine (OCSVM) method was adopted to extract collapsed buildings and obtained an OA of 88.3% and Kappa value of 70.8%. LiDAR data can characterize building roof changes through the accurate and precise measurement of height information. A 3D shape descriptor was further developed in [3], based on building contour clusters derived from airborne LiDAR point cloud data, and achieved an OA of 87.3% and Kappa value of 73.8%.

Kappa was used in this study as one of the evaluation metrics, and it is also commonly adopted by many previous studies related to earthquake-induced collapsed buildings [2,67]. Kappa statistics showed significant response to the class distribution, while OA is not a useful measure when evaluating classifiers learned on imbalance datasets [68,69]. Furthermore, it provides a mean to make the achieved result comparable with previous studies. However, it is pointed out that there are some drawbacks for Kappa [70]. The major limitations are that randomness may mislead the accuracy of assessment, and may incorporate problems in computation and analysis [71]. The allocation agreement and quantity agreement were further proposed by Pontius and Millions [70] to replace Kappa, and will be considered in further studies.

In this study, grade 4 damage was classified as non-collapsed, as in [1]. However, it is pointed out that grade 4 damage is difficult to identify from remotely sensed images [6]. Very heavy damaged buildings might have failure of walls, or partial structural failure of roofs and floors. Pre- and post-event imagery are needed to distinguish them from collapsed buildings. With the availability of aerial oblique imagery, a detailed building damage map for each grade might be achievable [72]. Besides, building footprints, in this study, were manually extracted from the imagery, which is time-consuming and also disadvantageous for planning rescue after earthquake. In [65], a fractal net evolution approach (FNEA) algorithm was adopted to delineate the image into objects. It showed that collapsed buildings and other objects (e.g., intact buildings, vegetation, and shadow areas) were well segmented, and a collapsed building detection method was further proposed. A region-growing-with-smoothness-constraint approach was suggested to segment damaged and undamaged buildings from airborne LiDAR data [73]. Out of 1953 validation buildings, 1890 were correctly segmented, with only 0.03% errors of commission and 0.03% errors of omission. Accurate automatic segmentation methods should be considered to extract building footprints from remotely sensed data.

It should be mentioned that the achieved result, in this study, was not satisfied for rapid damage assessment, as indicated by the Kappa values. On the one hand, only post-event imagery was used in the study, and pre-event data and LiDAR data are also crucial for identifying collapsed buildings. On the other hand, the performance is also affected by the building structures in the study area. There are many small buildings, making it difficult to correctly classify them. Furthermore, the size

of training dataset was too small for the CNNs. However, this study demonstrated that the CNN method, to some extent, is able to distinguish collapsed buildings from non-collapsed buildings using single post-event satellite imagery. The performance is expected to be improved when more training data are available or using more advanced deep learning structures. Besides, rotating the input image will not significantly impact the output as CNNs are capable of learning invariant features due to the special architecture, including local receptive fields, shared weights, and the spatial subsampling. The model transferability to other areas was not explored in this study. It is worth pointing out that the CNN model is supposed to be transferrable to new datasets when a few training samples are available. The pretrained network weights can be fine-tuned by training the network with new data. Once an accurate CNN model was built, which can be viewed as the pretrained model, the fast availability of post-earthquake VHR imagery would be crucial for rapid damage assessment [74]. Interpretation of the imagery or field observation should also be involved, to prepare for a situation of little training data, which will be then used to fine-tune the pretrained model. The transferability of CNNs using remotely sensed data has been demonstrated in case studies of land-use classification, SAR target recognition, and soil clay content mapping using airborne hyperspectral data [36,75,76]. Furthermore, it is also possible to fine-tune CNN models (VGG, Xception, ResNet), trained by existing large-scale dataset (like ImageNet) to identify collapsed buildings caused by earthquakes.

5. Conclusions

Supervised classification algorithms have been widely used in damage assessment after an earthquake. In this study, the convolutional neural networks were proposed for identifying collapsed buildings after the Haiti 2010 earthquake using single post-earthquake VHR satellite imagery. The SqueezeNet method achieved OA values of 80.6% and 76.6% on classifying collapsed and non-collapsed buildings using test A and test B dataset. The damage grade distribution of earthquake-affected buildings is often imbalanced, as collapsed buildings are normally less than non-collapsed buildings after an earthquake. Three balancing methods were considered, integrating with the CNN model. Although the overall accuracies did not improve significantly for the two test regions, the model's capability of identifying collapsed buildings was enhanced with balancing methods. The SqueezeNet-similar CNN model achieved Kappa values of 40.6% and 48.9%, for test A and test B dataset, with the cost-sensitive method. The zero-padding operation for preparing building patches can improve the model performance on the intra-class dataset. The balanced model achieved the Kappa value of 57.8% for buildings with width larger than 46 pixels. However, the number of false positives were increased, along with the increase of the number of true positives, especially for small buildings. Thus, the building structures in the study area should be considered when using balancing methods for identifying collapsed buildings after earthquakes. The efficiency is crucial for emergency mapping. Apart from the preparation for training data, it took 95.3 s to train the model using Google Cloud Platform in this study. When more data and complex deep learning models with even millions parameters were involved, the model training time will be expected to be increased. It would be practical to consider a pretrained model, which requires fewer samples to fine-tune it, instead of starting from scratch after the earthquake.

Author Contributions: All authors contributed in a substantial way to the manuscript. M.J. conceived, designed and performed the research and wrote the manuscript. L.L. made contributions to the design of the research and data analysis. All authors discussed the basic structure of the manuscript. M.B. reviewed the manuscript and supervised the study at all stages. All authors read and approved the submitted manuscript.

Funding: This research received no external funding.

Acknowledgments: We acknowledge support by the German Research Foundation and the Open Access Publication Funds of the TU Dresden. The first author wants to express her acknowledgment to the China Scholarship Council (CSC) for providing financial support to study at TU Dresden.

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

References

1. Cooner, A.J.; Shao, Y.; Campbell, J.B. Detection of urban damage using remote sensing and machine learning algorithms: Revisiting the 2010 Haiti earthquake. *Remote Sens.* **2016**, *8*, 868. [[CrossRef](#)]
2. 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](#)]
3. He, M.; Zhu, Q.; Du, Z.; Hu, H.; Ding, Y.; Chen, M. A 3D shape descriptor based on contour clusters for damaged roof detection using airborne LiDAR point clouds. *Remote Sens.* **2016**, *8*, 189. [[CrossRef](#)]
4. Menderes, A.; Erener, A.; Sarp, G. Automatic detection of damaged buildings after earthquake hazard by using remote sensing and information technologies. *Procedia Earth Planet. Sci.* **2015**, *15*, 257–262. [[CrossRef](#)]
5. Ghosh, S.; Huyck, C.K.; Greene, M.; Stuart, P.; Bevington, J.; Svekla, W.; Desroches, R. Crowdsourcing for rapid damage assessment: The global earth observation catastrophe assessment network (GEO-CAN). *Earthq. Spectra* **2011**, *27*, S179–S198. [[CrossRef](#)]
6. Corbane, C.; Saito, K.; Oro, L.D.; Bjorgo, E.; Gill, S.P.D.; Piard, B.E.; Huyck, C.K.; Kemper, T.; Lemoine, G.; Spence, R.J.S.; et al. A comprehensive analysis of building damage in the 12 January 2010 Mw 7 Haiti earthquake using high-resolution satellite-and aerial imagery. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 997–1009. [[CrossRef](#)]
7. Gokon, H.; Koshimura, S. Mapping of building damage of the 2011 Tohoku earthquake tsunami in Miyagi Prefecture. *Coast. Eng. J.* **2012**, *54*, 1250006. [[CrossRef](#)]
8. 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](#)]
9. Saito, K.; Spence, R. Rapid damage mapping using post-earthquake satellite images. In Proceedings of the 2004 IEEE International Geoscience and Remote Sensing Symposium, Anchorage, AK, USA, 20–24 September 2004; Volume 4, pp. 2272–2275.
10. Rathje, E.M.; Crawford, M.; Woo, K.; Neuenschwander, A. Damage patterns from satellite images of the 2003 Bam, Iran, earthquake. *Earthq. Spectra* **2005**, *21*, 295–307. [[CrossRef](#)]
11. Turker, M.; Sumer, E. Building-based damage detection due to earthquake using the watershed segmentation of the post-event aerial images. *Int. J. Remote Sens.* **2008**, *29*, 3073–3089. [[CrossRef](#)]
12. Mitomi, H.; Saita, J.; Matsuoka, M.; Yamazaki, F. Automated damage detection of buildings from aerial television images of the 2001 Gujarat, India earthquake. In Proceedings of the IEEE 2001 International Geoscience and Remote Sensing Symposium (IGARSS 2001), Sydney, NSW, Australia, 9–13 July 2001; Volume 1, pp. 147–149.
13. Aoki, H.; Matsuoka, M.; Yamazaki, F. Automated detection of damaged buildings due to earthquakes using aerial HDTV and photographs. *J. Jpn. Soc. Photogramm. Remote Sens.* **2001**, *40*, 27–36. [[CrossRef](#)]
14. Brunner, D.; Lemoine, G.; Bruzzone, L. Earthquake damage assessment of buildings using VHR optical and SAR imagery. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 2403–2420. [[CrossRef](#)]
15. Park, S.E.; Yamaguchi, Y.; Kim, D.J. Polarimetric SAR remote sensing of the 2011 Tohoku earthquake using ALOS/PALSAR. *Remote Sens. Environ.* **2013**, *132*, 212–220. [[CrossRef](#)]
16. Zhai, W.; Shen, H.F.; Huang, C.L.; Pei, W.S. Building damage information investigation after earthquake using single post-event polsar image. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 7338–7341.
17. Shi, L.; Sun, W.; Yang, J.; Li, P.; Lu, L. Building collapse assessment by the use of postearthquake Chinese VHR airborne SAR. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 2021–2025. [[CrossRef](#)]
18. Timo, B.; Liao, M. Building-damage detection using post-seismic high-resolution SAR satellite data. *Int. J. Remote Sens.* **2010**, *31*, 3369–3391.
19. Zhai, W.; Huang, C. Fast building damage mapping using a single post-earthquake PolSAR image: A case study of the 2010 Yushu earthquake. *Earth Planets Space* **2016**, *68*, 86. [[CrossRef](#)]
20. Rastiveis, H.; Eslamizade, F.; Hosseini-Zirdoo, E. Building damage assessment after earthquake using post-event LiDAR data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2015**, *40*, 595. [[CrossRef](#)]
21. Labiak, R.C.; Van Aardt, J.A.N.; Bepalov, D.; Eychner, D.; Wirch, E.; Bischof, H.-P. Automated method for detection and quantification of building damage and debris using post-disaster lidar data. In Proceedings of the Laser Radar Technology and Applications XVI, Orlando, FL, USA, 25–29 April 2011; Article No; Volume 8037.

22. Dou, X.; Ma, Z.; Huang, S.; Wang, X. Building damage extraction from post-earthquake airborne LiDAR data. *Acta Geol. Sin. (Engl. Ed.)* **2016**, *90*, 1481–1489.
23. Yu, H.; Mohammed, M.A.; Mohammadi, M.E.; Moaveni, B.; Barbosa, A.R.; Stavridis, A.; Wood, R.L. Structural identification of an 18-story RC building in Nepal using post-earthquake ambient vibration and Lidar data. *Front. Built Environ.* **2017**, *3*. [[CrossRef](#)]
24. Wu, F.; Gong, L.; Wang, C.; Member, S.; Zhang, H. Signature analysis of building damage with TerraSAR-X new staring spotLight mode data. *IEEE Geosci. Remote Sens. Lett.* **2016**, *13*, 1696–1700. [[CrossRef](#)]
25. Li, X.W.; Guo, H.D.; Zhang, L.; Chen, X.; Liang, L. A new approach to collapsed building extraction using RADARSAT-2 polarimetric SAR imagery. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 677–681.
26. Tong, X.; Hong, Z.; Liu, S.; Zhang, X.; Xie, H.; Li, Z.; Yang, S.; Wang, W.; Bao, F. Building-damage detection using pre-and post-seismic high-resolution satellite stereo imagery: A case study of the May 2008 Wenchuan earthquake. *ISPRS J. Photogramm. Remote Sens.* **2012**, *68*, 13–27. [[CrossRef](#)]
27. Rastiveis, H.; Samadzadegan, F.; Reinartz, P. A fuzzy decision making system for building damage map creation using high resolution satellite imagery. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 455. [[CrossRef](#)]
28. Batista, G.E.; Carvalho, A.C.; Monard, M.C. Applying one-sided selection to unbalanced datasets. In Proceedings of the Mexican International Conference on Artificial Intelligence, Acapulco, Mexico, 11–14 April 2000; Springer: Berlin/Heidelberg, Germany, 2000.
29. Li, L.; Li, Z.; Zhang, R.; Ma, J.; Lei, L. Collapsed buildings extraction using morphological profiles and texture statistics—A case study in the 5.12 Wenchuan earthquake. In Proceedings of the 2010 IEEE International Geoscience and Remote Sensing Symposium, Honolulu, HI, USA, 25–30 July 2010; pp. 2000–2002.
30. Yu, H.; Cheng, G.; Ge, X. Earthquake-collapsed building extraction from LiDAR and aerophotograph based on OBIA. In Proceedings of the 2nd International Conference on Information Science and Engineering, Hangzhou, China, 4–6 December 2010; pp. 2034–2037.
31. Taskin Kaya, G.; Musaoglu, N.; Ersoy, O.K. Damage assessment of 2010 Haiti earthquake with post-earthquake satellite image by support vector selection and adaptation. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 1025–1035. [[CrossRef](#)]
32. Bialas, J.; Oommen, T.; Rebbapragada, U.; Levin, E. Object-based classification of earthquake damage from high-resolution optical imagery using machine learning. *J. Appl. Remote Sens.* **2016**, *10*, 036025. [[CrossRef](#)]
33. Gong, L.; Wang, C.; Wu, F.; Zhang, J.; Zhang, H.; Li, Q. Earthquake-induced building damage detection with post-event sub-meter VHR TerraSAR-X staring spotlight imagery. *Remote Sens.* **2016**, *8*, 887. [[CrossRef](#)]
34. LeCun, Y.; Bengio, Y. Convolutional networks for images, speech, and time-series. In *The Handbook of Brain Theory and Neural Networks*; MIT Press: Cambridge, MA, USA, 1995; Volume 3361.
35. Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324.
36. Liu, L.; Ji, M.; Buchroithner, M. Transfer learning for soil spectroscopy based on convolutional neural networks and its application in soil clay content mapping using hyperspectral imagery. *Sensors* **2018**, *18*, 3169. [[CrossRef](#)] [[PubMed](#)]
37. Krizhevsky, A.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. In Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.
38. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the International Conference on Learning Representations, San Diego, CA, USA, 7–9 May 2015.
39. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
40. Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv*, 2016; arXiv:1602.07360.
41. Bai, Y.; Gao, C.; Singh, S.; Koch, M.; Adriano, B.; Mas, E.; Koshimura, S. A framework of rapid regional tsunami damage recognition from post-event terraSAR-X imagery using deep neural networks. *IEEE Geosci. Remote Sens. Lett.* **2018**, *15*, 43–47. [[CrossRef](#)]

42. Vetrivel, A.; Gerke, M.; Kerle, N.; Nex, F.; Vosselman, G. Disaster damage detection through synergistic use of deep learning and 3D point cloud features derived from very high resolution oblique aerial images, and multiple-kernel-learning. *ISPRS J. Photogramm. Remote Sens.* **2018**, *140*, 45–59. [[CrossRef](#)]
43. Zhang, X.; Song, Q.; Zheng, Y.; Hou, B.; Gou, S. Classification of imbalanced hyperspectral imagery data using support vector sampling. In Proceedings of the 2014 IEEE Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 2870–2873.
44. Puertas, O.L.; Brenning, A.; Meza, F.J. Balancing misclassification errors of land cover classification maps using support vector machines and Landsat imagery in the Maipo river basin (Central Chile, 1975–2010). *Remote Sens. Environ.* **2013**, *137*, 112–123. [[CrossRef](#)]
45. Syafiq, M.; Pozi, M.; Sulaiman, N.; Mustapha, N. A new classification model for a class imbalanced data set using genetic programming and support vector machines: Case study for wilt disease classification. *Remote Sens. Lett.* **2015**, *6*, 568–577.
46. Owen, A.B. Infinitely imbalanced logistic regression. *J. Mach. Learn. Res.* **2007**, *8*, 761–773.
47. Provost, F. Machine learning for the detection of oil spills in satellite radar images. *Mach. Learn.* **1998**, *30*, 195–215. [[CrossRef](#)]
48. Buda, M.; Maki, A.; Mazurowski, M.A. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Netw.* **2018**, *106*, 249–259. [[CrossRef](#)] [[PubMed](#)]
49. 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](#)]
50. UNITAR/UNOSAT; EC Joint Research Centre; World Bank. Haiti Earthquake 2010: Remote Sensing Damage Assessment. Available online: <http://www.unitar.org/unosat/haiti-earthquake-2010-remote-sensing-based-building-damage-assessment-data> (accessed on 10 May 2017).
51. Grünthal, G. *European Macroseismic Scale 1998*; Cahiers du Centre Européen de Géodynamique et de Séismologie, Conseil de l'Europe, Ed.; Centre Européen de Géodynamique et de Séismologie: Luxembourg, 1998.
52. Huang, B.; Zhao, B.; Song, Y. Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery. *Remote Sens. Environ.* **2018**, *214*, 73–86. [[CrossRef](#)]
53. Ferreira, A.; Giraldo, G. Convolutional neural network approaches to granite tiles classification. *Expert Syst. Appl.* **2017**, *84*, 1–11. [[CrossRef](#)]
54. Guidici, D.; Clark, M. One-dimensional convolutional neural network land-cover classification of multi-seasonal hyperspectral imagery in the San Francisco Bay Area, California. *Remote Sens.* **2017**, *9*, 629. [[CrossRef](#)]
55. Sameen, M.I.; Pradhan, B.; Aziz, O.S. Classification of very high resolution aerial photos using spectral-spatial convolutional neural networks. *J. Sens.* **2018**, *2018*, 7195432. [[CrossRef](#)]
56. Nair, V.; Hinton, G.E. Rectified linear units improve restricted Boltzmann machines. In Proceedings of the 27th International Conference on Machine Learning (ICML10), Haifa, Israel, 21–24 June 2010; pp. 807–814.
57. Krawczyk, B. Learning from imbalanced data: Open challenges and future directions. *Prog. Artif. Intell.* **2016**, *5*, 221–232. [[CrossRef](#)]
58. Hou, Z.H.; Liu, X.Y. On multi-class cost-sensitive learning. *Comput. Intell.* **2010**, *26*, 232–257.
59. Kukar, M.; Kononenko, I. Cost-sensitive learning with neural networks. In Proceedings of the 13th European Conference on Artificial Intelligence, Brighton, UK, 23–28 August 1998; pp. 445–449.
60. Wozniak, M. *Hybrid Classifiers-Methods of Data, Knowledge, and Classifier Combination*; Springer: Berlin/Heidelberg, Germany, 2014; Volume 519.
61. Landis, J.R.; Koch, G.G. The Measurement of Observer Agreement for Categorical Data Published by: International Biometric Society Stable. *Biometrics* **1977**, *33*, 159–174. [[CrossRef](#)] [[PubMed](#)]
62. Fleet, D.; Hutchison, D. Spatial pyramid pooling in deep convolutional networks for visual recognition. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; Springer: Cham, Switzerland, 2014; pp. 346–361.
63. Ural, S.; Hussain, E.; Kim, K.; Fu, C.-S.; Shan, J. Building extraction and rubble mapping for city Port-au-Prince post-2010 earthquake with GeoEye-1 imagery and Lidar data. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 1011–1023. [[CrossRef](#)]
64. Gao, M.; Hong, X.; Chen, S.; Harris, C.J. A combined SMOTE and PSO based RBF classifier for two-class imbalanced problems. *Neurocomputing* **2011**, *74*, 3456–3466. [[CrossRef](#)]

65. Fernández-gómez, M.J.; Asencio-cortés, G.; Troncoso, A.; Martínez-álvarez, F. Large earthquake magnitude prediction in Chile with imbalanced classifiers and ensemble learning. *Appl. Sci.* **2017**, *7*, 625. [[CrossRef](#)]
66. Wang, X.; Li, P. Extraction of earthquake-induced collapsed buildings using very high-resolution imagery and airborne lidar data. *Int. J. Remote Sens.* **2015**, *36*, 2163–2183. [[CrossRef](#)]
67. Tong, X.; Lin, X.; Feng, T.; Xie, H.; Liu, S.; Hong, Z.; Chen, P. Use of shadows for detection of earthquake-induced collapsed buildings in high-resolution satellite imagery. *ISPRS J. Photogramm. Remote Sens.* **2013**, *79*, 53–67. [[CrossRef](#)]
68. Leichtle, T.; Geiß, C.; Lakes, T.; Taubenböck, H. Class imbalance in unsupervised change detection—A diagnostic analysis from urban remote sensing. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *60*, 83–98. [[CrossRef](#)]
69. Chawla, N.V. Data mining for imbalanced datasets: An overview. In *Data Mining and Knowledge Discovery Handbook*; Springer: Boston, MA, USA, 2009; pp. 875–886.
70. Pontius, R.G.; Millones, M. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment. *Int. J. Remote Sens.* **2011**, *32*, 4407–4429. [[CrossRef](#)]
71. Nijhawan, R.; Raman, B.; Das, J. Proposed hybrid-classifier ensemble algorithm to map snow cover area. *J. Appl. Remote Sens.* **2018**, *12*. [[CrossRef](#)]
72. Gerke, M.; Kerle, N. Automatic structural seismic damage assessment with airborne oblique Pictometry[®] imagery. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 885–898. [[CrossRef](#)]
73. Axel, C.; van Aardt, J. Building damage assessment using airborne lidar. *J. Appl. Remote Sens.* **2017**, *11*, 046024. [[CrossRef](#)]
74. Voigt, S.; Schneiderhan, T.; Twele, A.; Gähler, M.; Stein, E.; Mehl, H. Rapid damage assessment and situation mapping: Learning from the 2010 Haiti earthquake. *Photogramm. Eng. Remote Sens.* **2011**, *77*, 923–931. [[CrossRef](#)]
75. Zhao, B.; Huang, B.; Zhong, Y. Transfer learning with fully pretrained deep convolution networks for land-use classification. *IEEE Geosci. Remote Sens. Lett.* **2017**, *14*, 1436–1440. [[CrossRef](#)]
76. Huang, Z.; Pan, Z.; Lei, B. Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data. *Remote Sens.* **2017**, *9*, 907. [[CrossRef](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).