

Article

# Improved CNN Classification Method for Groups of Buildings Damaged by Earthquake, Based on High Resolution Remote Sensing Images

Haojie Ma <sup>1,2,†</sup>, Yalan Liu <sup>1,\*</sup>, Yuhuan Ren <sup>1,†</sup>, Dacheng Wang <sup>1</sup>, Linjun Yu <sup>1</sup> and Jingxian Yu <sup>1,2</sup>

<sup>1</sup> Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China; mahj@radi.ac.cn (H.M.); renyh@radi.ac.cn (Y.R.); wangdc@radi.ac.cn (D.W.); yulj201831@aircas.ac.cn (L.Y.); yujx@radi.ac.cn (J.Y.)

<sup>2</sup> University of Chinese Academy of Sciences, Beijing 100049, China

\* Correspondence: liuy1@aircas.ac.cn; Tel.: +86-139-1103-2598

† These authors contributed equally to this work.

Received: 3 December 2019; Accepted: 5 January 2020; Published: 11 January 2020



**Abstract:** Effective extraction of disaster information of buildings from remote sensing images is of great importance to supporting disaster relief and casualty reduction. In high-resolution remote sensing images, object-oriented methods present problems such as unsatisfactory image segmentation and difficult feature selection, which makes it difficult to quickly assess the damage sustained by groups of buildings. In this context, this paper proposed an improved Convolution Neural Network (CNN) Inception V3 architecture combining remote sensing images and block vector data to evaluate the damage degree of groups of buildings in post-earthquake remote sensing images. By using CNN, the best features can be automatically selected, solving the problem of difficult feature selection. Moreover, block boundaries can form a meaningful boundary for groups of buildings, which can effectively replace image segmentation and avoid its fragmentary and unsatisfactory results. By adding Separate and Combination layers, our method improves the Inception V3 network for easier processing of large remote sensing images. The method was tested by the classification of damaged groups of buildings in 0.5 m-resolution aerial imagery after the earthquake of Yushu. The test accuracy was 90.07% with a Kappa Coefficient of 0.81, and, compared with the traditional multi-feature machine learning classifier constructed by artificial feature extraction, this represented an improvement of 18% in accuracy. Our results showed that this improved method could effectively extract the damage degree of groups of buildings in each block in post-earthquake remote sensing images.

**Keywords:** earthquake; damaged groups of buildings; classification; remote sensing images; Convolution Neural Network (CNN); block vector data

## 1. Introduction

During the rescue and recovery phases following an earthquake, damaged buildings may indicate the locations of trapped people [1]. Hence, building damage maps are key to post-earthquake rescue and reconstruction. The use of traditional manual field survey methods to obtain building damage information presents relatively high accuracy and confidence. However, there are still some shortcomings such as large workload, low efficiency, high costs and unintuitive information, meaning that these methods cannot meet the requirements for the fast acquisition of building damage information [2]. With the progress of sensors and space technology, remote sensing can now provide detailed spatial and temporal information for target areas, whilst usually requiring little field work. Therefore, remote sensing has been widely used in various post-disaster rescue operations, being particularly important for earthquake-stricken areas where it is often difficult to conduct field surveys

for the first time [3]. To some extent, previous studies have proved that relatively accurate information of building damage can be obtained from remote sensing data [4].

Many methods to extract information about building damage, caused by earthquakes, from remote sensing images have been presented [5–7]. These can be divided into single-temporal and multi-temporal evaluation methods according to the number of images being used. In single-temporal evaluation methods, only one kind of post-earthquake image is used for the extraction of information. Multi-temporal evaluation methods, in turn, use at least two-temporal images. Due to the influence of data acquisition, sensor revisiting cycle, and filming angle and time, multi-temporal evaluation methods are difficult to be applied in practice [8]. Single-temporal evaluation methods are less restricted and have become an effective technical means to directly extract and evaluate the damage information of buildings using remote sensing images after earthquakes [9]. Chen [10] used object-oriented methods to segment remote sensing images, and classified image objects by Classification and Decision Tree (CART), Support Vector Machine (SVM), and Random Forest (RF) in the machine learning methods. The results showed that, among the three machine learning methods, RF was the best in extracting information about damaged buildings. Janalipour et al. [11] used high spatial resolution remote sensing images as a background to manually select and extract features based on the fuzzy genetic algorithm, establishing a semi-automatic detection system for building damage. This system has higher robustness and precision compared with machine learning methods such as the RF and SVM. In the single-temporal evaluation methods, the spectral, textural and morphological features of the image are mainly used for its classification. This process is often based on the object-oriented classification algorithm for the extraction of information about building destruction [12–14]. However, these object-oriented classification methods present problems such as difficulties in feature space selection and unsatisfactory image segmentation.

In recent years, deep learning technology has achieved great success in the image application field [15], becoming more and more popular in the applications of remote sensing [16]. Convolutional Neural Network (CNN) is a common method of deep learning. Since LeNet5 [17] has achieved satisfactory results in handwritten number recognition tasks compared with traditional methods, a large number of deeper and more complex CNN, such as AlexNet [18], VGGNet [19], Inception V3 [20], and ResNet [21], have made a great breakthrough in large-scale image classification tasks. These CNN models went deeper and deeper, reaching 152 layers by the time of ResNet. A deeper network layer means that deeper image features can be extracted, making the image classification results more accurate. Compared with traditional object-oriented classification methods, CNN-based methods can select and extract classification features automatically, presenting a strong self-learning ability and robustness [22]. Meng Chen et al. [23] proposed a method combining image segmentation with CNN to extract building damage information from remote sensing images, and effectively extracted the damage information of buildings after earthquakes. However, its robustness is largely affected by the accuracy of image segmentation, which affects the effectiveness of the successive tasks and limits practical applications.

With the continuous improvement of basic Geographic Information System (GIS) data in recent years, their applications for the extraction of richer and clearer disaster information have become more popular [24,25]. Ye et al. [8] combined the block information derived from urban road vector data in post-earthquake remote sensing images and constructed a multi-feature classification model with building blocks as its unit. The results showed that this method had high accuracy for the classification of the damage degree of groups of buildings. Therefore, GIS data can be applied to the damage assessment of buildings. In other words, we can accurately achieve the boundary segmentation of the groups of buildings by overlaying the analysis of GIS data and remote sensing images instead of using image segmentation.

To overcome the problem of feature selection and image segmentation in object-oriented classification, this paper proposes a new strategy to extract the damage information of groups of buildings via remote sensing images by combining CNN and GIS data. Using block vector data,

all buildings in each block were treated as a group of buildings unit, and the Inception V3 network in CNN was used as the basic classification network to classify the damaged groups of buildings, which was compared with traditional machine learning methods. The rest of this paper is organized as follows. Section 2 describes the details of the method. Section 3 provides the descriptions of the study area. Section 4 presents the analysis of the experimental results. Finally, Section 5 concludes this paper.

## 2. Method

### 2.1. The Basic Architecture of Inception V3

The basic architecture of CNN usually includes alternating overlapping convolution and pooling layers, full connection layer, input layer, and output layer. The convolution layer usually includes convolution operation and nonlinear transformation [26]. In the Inception V3 network, besides these basic structures, there are some special operations such as the addition of the well-designed Inception Module to replace the last fully connected layer with global average pooling and the addition of the BN (Batch Normalization) [27] method.

The Inception Module in Inception V3 can improve the efficiency of parameter utilization. One of the structures is shown in Figure 1. It is similar to a small network in a large network, and its structure can be repeatedly stacked to form a larger network. In the Inception Module structure, the convolution of  $1 \times 1$  can organize information across channels, improve the expression ability of the network, and raise or lower the dimensions of the output channel. Moreover, the idea of spatial asymmetric convolution [20] is introduced to split a large two-dimensional convolution into two smaller one-dimensional convolutions. This helps to reduce many parameters, accelerate the operation, and reduce the overfitting; on the other hand, it adds a layer of nonlinearity, increases the diversity of features and expands the expression ability of the model and its ability to deal with more and richer spatial features.

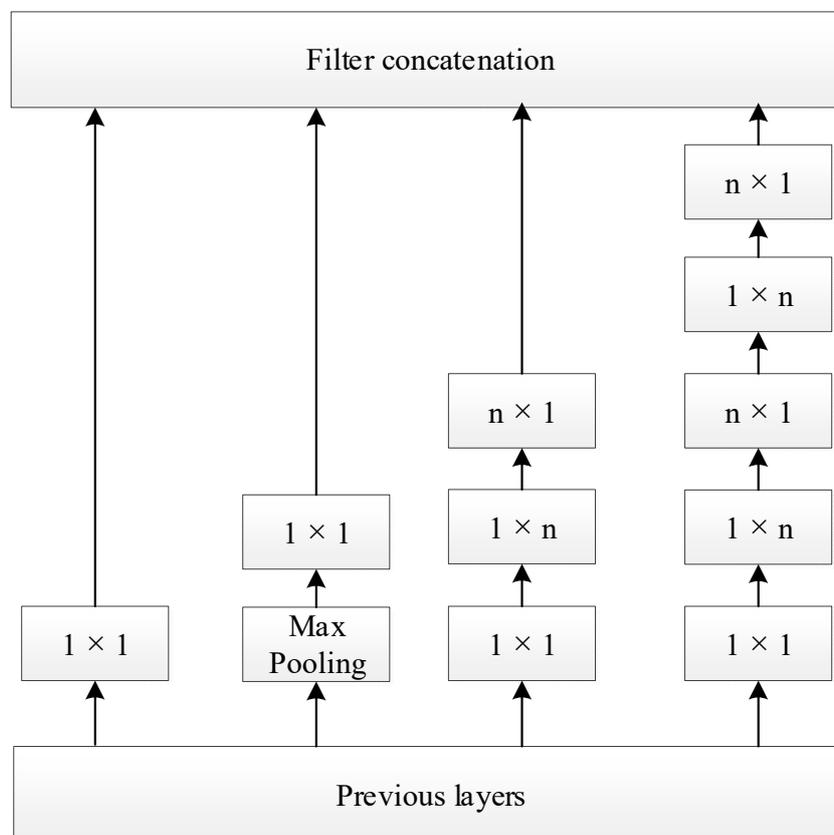


Figure 1. Inception Module structure in Inception V3.

Inception V3 removes the last fully connected layer and replaces it with a global average pooling layer. In AlexNet and VGGNet, the full connection layer accounts for almost 90% of all the parameters, which causes overfitting. Therefore, by removing the full connection layer, the model can be trained faster, and overfitting can be reduced.

The BN method added to Inception V3 is one of the most effective regularization methods. When BN is used for a certain layer of the neural network, it will standardize the internal processing of each mini-batch of data, to normalize the output to  $N(0,1)$  and reduce the changes in the distribution of internal neurons. It can considerably accelerate the training speed of large convolutional networks and greatly improve the classification accuracy after convergence.

## 2.2. Method Flow

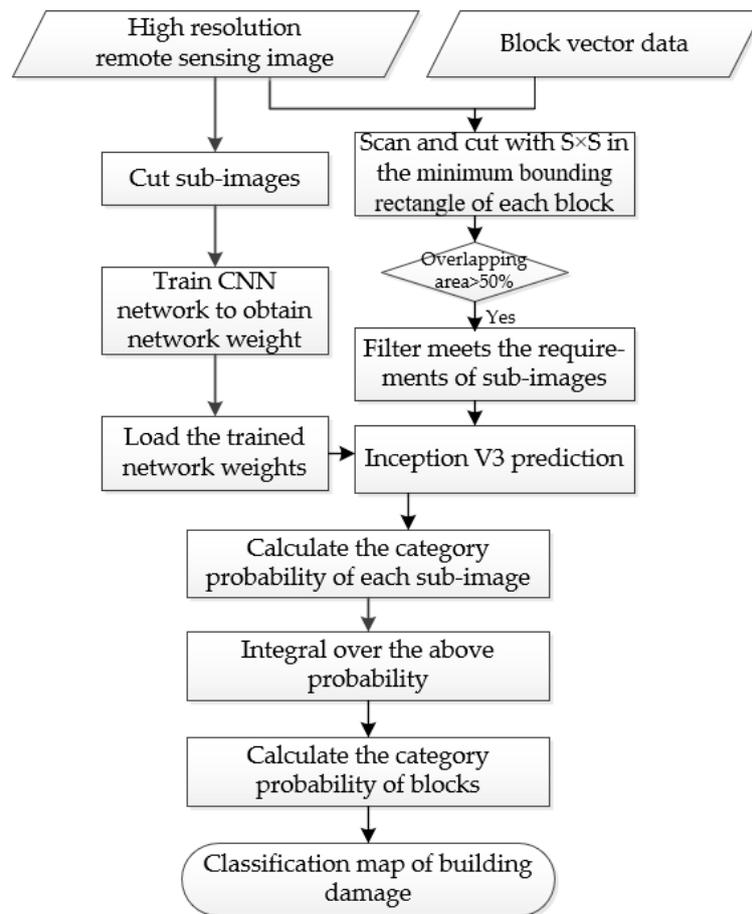
In this study, to utilize Inception V3, the remote sensing image was cropped into sub-images of  $S \times S$  pixels, and these sub-images with fixed size were used as an input for the Inception V3. Each sub-image was processed with Inception V3 to output the probability of each category of damage for groups of buildings. The higher the value, the greater the probability of belonging to the damaged category.

Although CNN can well predict the damage category of groups of buildings [23,28,29] in rectangular images with a fixed size, for groups of buildings in each block, it often results in a larger error to predict directly using CNN due to irregular shape and different sizes of blocks. Therefore, this paper proposed a method to avoid the bigger error when using CNN to predict the categories of blocks presenting irregular shapes and different sizes. The main idea of the method was to cut the rectangular image of fixed size in the minimum bounding rectangle of the block according to a specific step size, and the obtained rectangular image could directly predict its category by using CNN. However, many of the obtained rectangular images were outside the block range, so the threshold of the overlap ratio had to be set to filter out rectangular images outside the block range. Finally, the prediction results of all rectangular images within each block were averaged to obtain the category of the block.

By using the method described above combined with block vector data, the groups of buildings in post-earthquake remote sensing images were classified by the basic processes shown in Figure 2.

The first part was to cut the remote sensing images into rectangular sub-images with a fixed size and train the Inception V3 network to obtain the trained network weights.

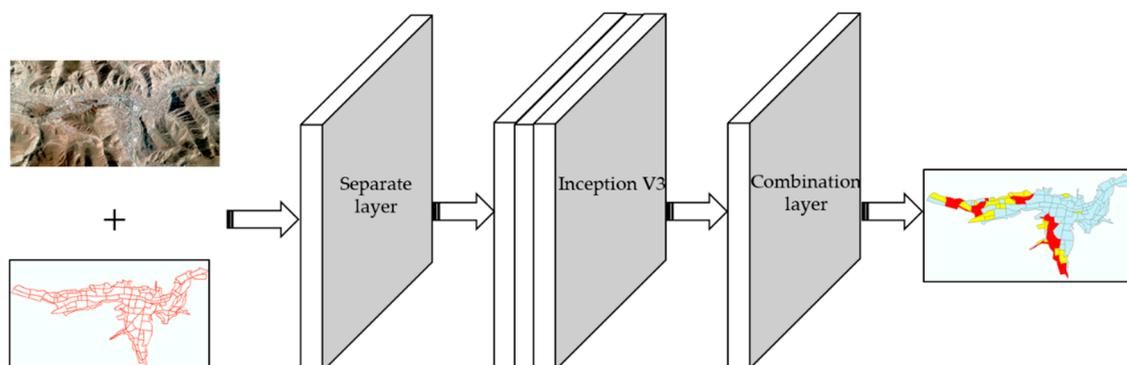
The second part consisted of three steps. First, an  $S \times S$  window was used to scan the images contained in the minimum bounding rectangle of each block with the step size  $S$ , and a certain number of  $S \times S$  sub-images were obtained. Then, the overlap ratio between each sub-image and its block was calculated, and the sub-images with an overlap area greater than 50% were used as inputs for the Inception V3 in order to predict their category probability. Finally, the category probabilities of all effective sub-images in each block were integrated to obtain the category probability of the block.



**Figure 2.** Classification process of damaged groups of buildings by remote sensing images combined with Convolutional Neural Network (CNN) and Geographic Information System (GIS).

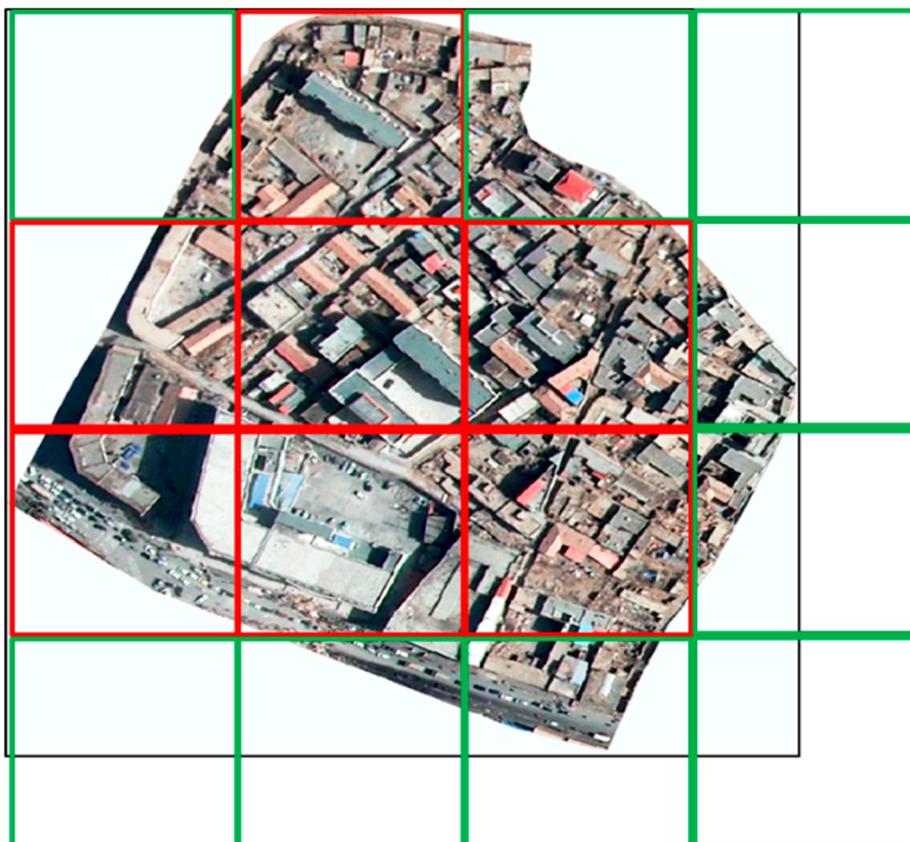
### 2.3. The Improved Convolutional Neural Network (CNN)

The remote sensing image usually has a larger width than that of the natural image, so they cannot be directly input into traditional CNN for processing. In order to enable the classification network to more easily process large remote sensing images for application and considering that the method proposed in this paper took city blocks as the basic processing unit, superimposition of block vector data for processing was required. Thus, this paper added Separate and Combination layers to Inception V3, as shown in Figure 3.



**Figure 3.** Improved CNN classification framework for damaged groups of buildings after earthquakes.

The role of the Separate layer here was to use a sliding window of  $224 \times 224$  to scan and cut with a step size of 224 within the minimum bounding rectangle of each block. Taking block 1 as an example, as shown in Figure 4, 16 sub-images of  $224 \times 224$  pixels were cut by scanning block 1. Then, the overlap area between these 16 sub-images and block 1 was calculated, and the 7 sub-images in which the overlap ratio was greater than 50% were selected as the valid sub-images, that is, the sub-images belonging to block 1. After the Separate layer, a number of  $224 \times 224$  pixels effective sub-images were generated for each block, and these sub-images were used as the input for the Inception V3 network for classification.



**Figure 4.** Sub-image cutting schematic diagram of block 1 (red boxes are valid sub-images).

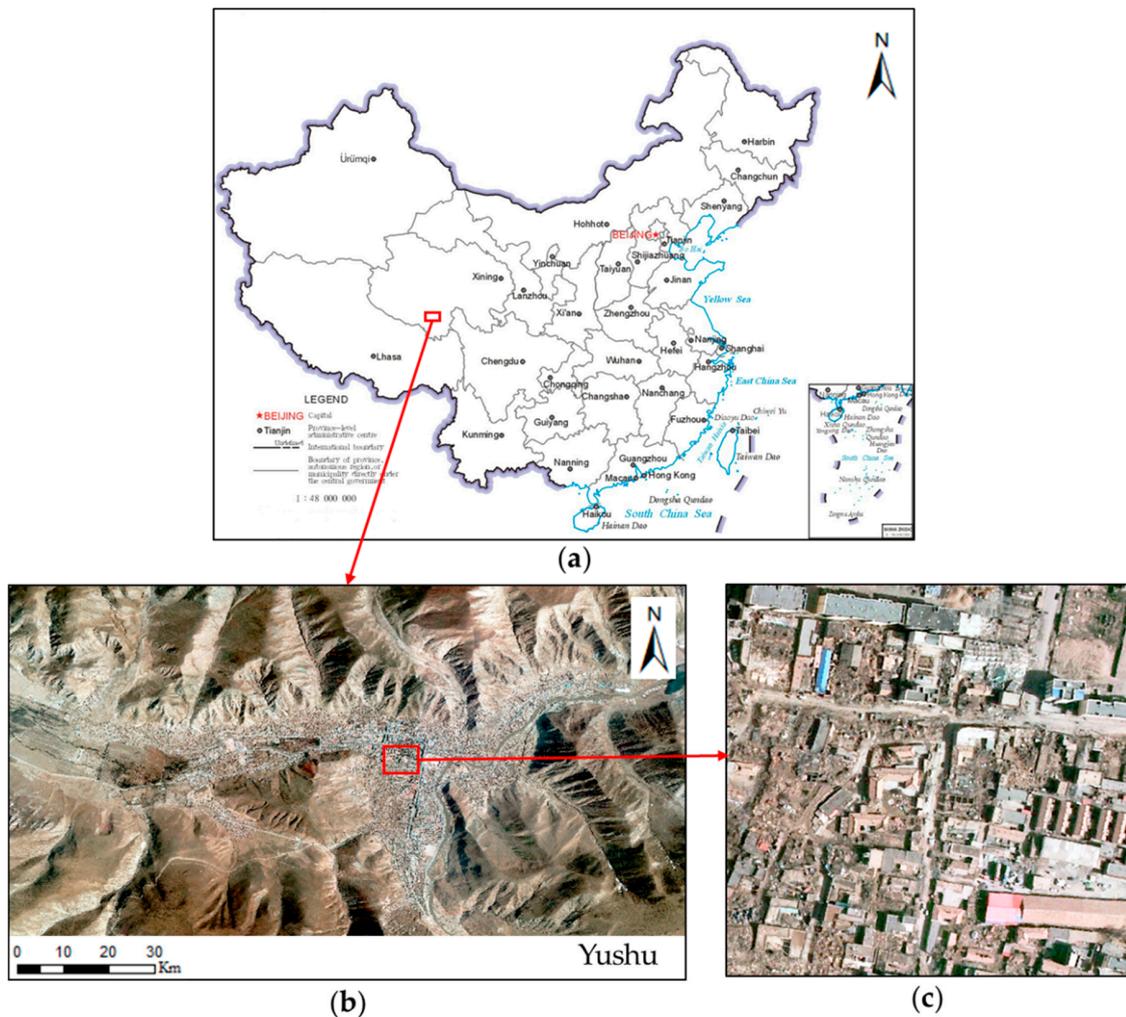
The purpose of the Combination layer was to combine the classification results of all valid sub-images in each block. For example, the function to integrate the probability values of 7 sub-images in block 1 is expressed as:

$$P_{1,j} = \frac{1}{7} \sum_{i=1}^7 A_{i,j} \quad (1)$$

where  $A_{i,j}$  is the probability that the  $i$  valid sub-image in block 1 is classified into  $j$  class, and  $P_{1,j}$  is the probability that block 1 is divided into  $j$  class.

### 3. Data

In order to test the effectiveness of this method, 0.5 m-resolution aerial remote sensing images were selected. These images were acquired on the second day after the 7.1-magnitude earthquake in the Yushu Tibetan Autonomous Prefecture of Qinghai Province on 14 April 2010. The studied area was severely affected by the disaster and there were a large number of collapsed buildings (Figure 5).



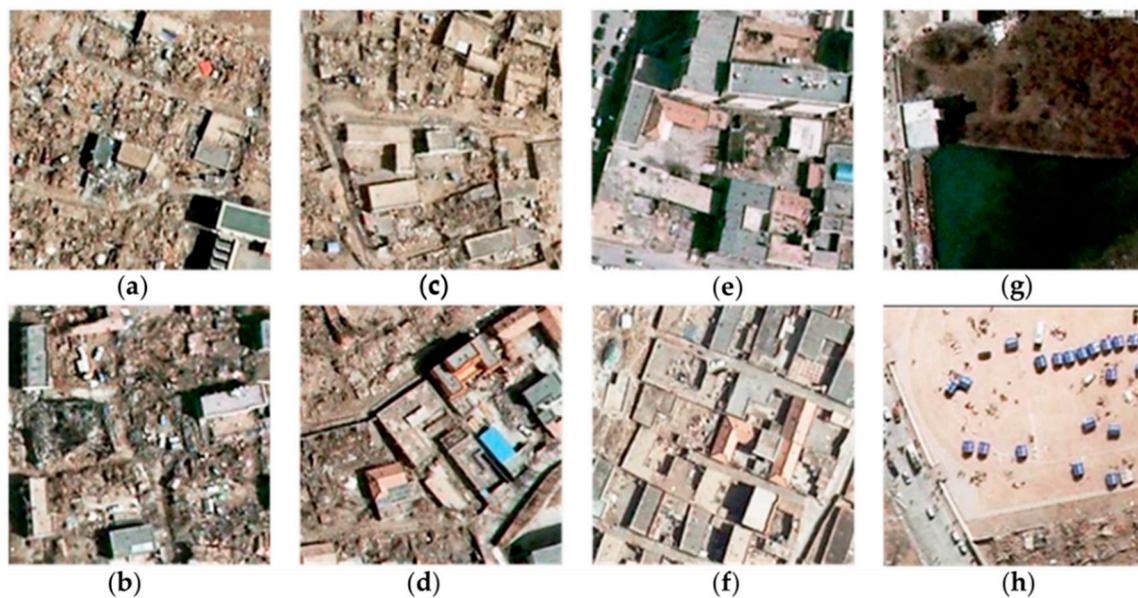
**Figure 5.** Location of the study area and the remote sensing image: (a) China map, (b) the image of Yushu Tibetan Autonomous Prefecture, and (c) part of the enlarged image.

The earthquake damage of buildings was divided into five grades according to the “Guidelines for Earthquake Damage and its Loss Assessment” formulated by the China Earthquake Administration and the “Classification Standard for Building Earthquake Damage Levels” formulated by the Ministry of Housing and Urban-Rural Development of China. In remote sensing images, classification is based mainly on the overall and detailed image characteristics of the buildings after the earthquake. Additionally, in the classification process, the sub-study area method is adopted to make a general assessment of the damage degree of all buildings in the sub-study area. In this paper, the sub-study area could be divided into a sub-image of  $224 \times 224$  pixels, or a block. Referring to previous research results [8,30,31], and according to the post-earthquake remote sensing images, this paper divided the damage of the groups of buildings into three levels: serious damage (all destroyed or most collapsed), moderate damage (about half collapsed), and slight damage (generally intact or a small part collapsed). The specific classifications are shown in Table 1. The collapse rate,  $c$ , was the ratio between the number of collapsed buildings (or collapsed building area) and the total number of buildings (or total building area).

**Table 1.** Classification of damaged groups of buildings in post-earthquake remote sensing images.

Category	Character Description	Collapse Rate
Serious damage	All destroyed or most collapsed	>70%
Moderate damage	About half collapsed	30–70%
Slight damage	Generally intact or a small part collapsed	<30%

After some necessary processing, the above experimental images can be trained within the CNN. The sub-images of  $224 \times 224$  pixels were clipped from the remote sensing images. Then, one category was assigned to each sub-image by visual interpretation. The samples for each category of Table 1 are shown in Figure 6 (where negative samples are open space and water bodies are near buildings, etc.).



**Figure 6.** Typical classification samples for damaged groups of buildings: (a,b) serious damage, (c,d) moderate damage, (e,f) slight damage, (g,h) negative samples.

Data used for training usually needs to be enhanced, similar to natural images, and processes such as rotation, flipping, and color conversion can be used [32]. However, this enhancement is different from that applied to natural images. Most objects in natural images can only be rotated at very small angles, yet, the buildings can be rotated at any angle in this study. In addition, remote sensing images are often displayed after being stretched, so data enhanced by stretching could make the Inception V3 model more robust. Through the process of enhancing, a total of 16,803 samples were obtained. These were divided into three groups: 10,764 samples were used as a training set, 4599 as a validation set, and 1440 as a test set.

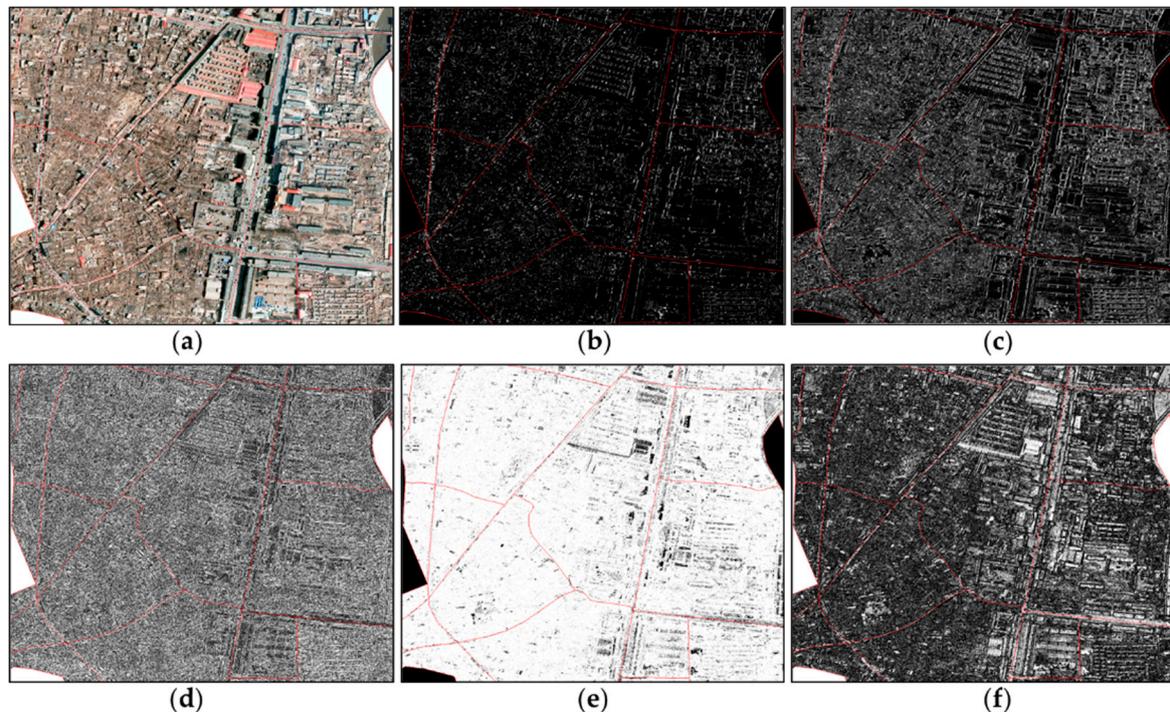
## 4. Results and Discussion

### 4.1. Multi-Feature Machine Learning Classification Method

To compare the accuracy of the results obtained by the method described in this paper with that of the results of traditional machine learning methods, the experiment involving the machine learning method was firstly carried out. First, various features of the image were extracted manually, and then, according to the features extracted from each block, the damage categories of the groups of buildings were classified by the machine learning method.

The features selected in this study were: contrast, dissimilarity, correlation, entropy, and homogeneity. The extracted image features are shown in Figure 7. It can be seen from the original

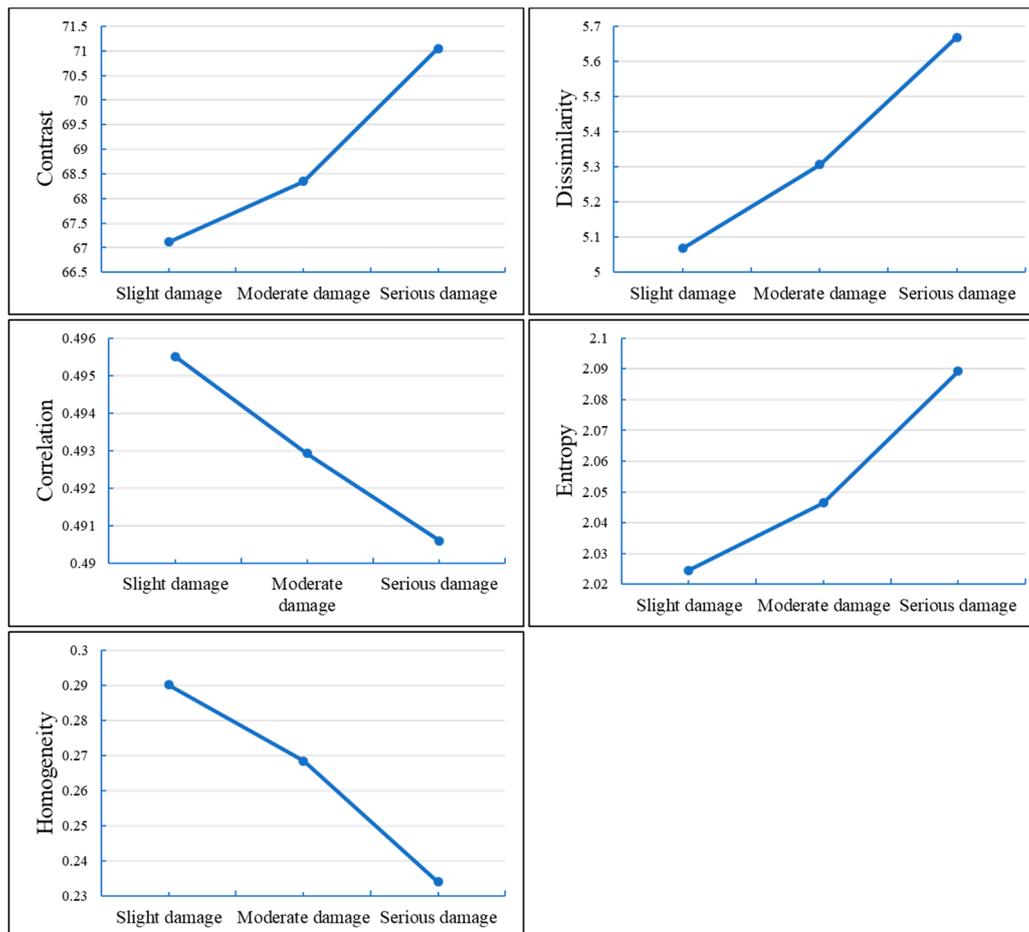
raster image in Figure 7 that the buildings in the left half of the image were completely collapsed, while the buildings in the right half were less collapsed. Comparing the five characteristic diagrams in Figure 7, it can be seen that the difference between the collapsed buildings on the left side and the non-collapsed buildings on the right side is clearly reflected in the feature map.



**Figure 7.** Feature maps: (a) raster images, (b) contrast, (c) dissimilarity, (d) correlation, (e) entropy, and (f) homogeneity.

In order to see more easily the range of values of each feature for the different damage categories, a quantitative statistical analysis was conducted for the feature maps of each category of damage, and the results are shown in Figure 8. Changes in building damage grade cause the values of the five features to change accordingly, which can reflect the damage grade of groups of buildings. Therefore, these five features can be used to classify the damage of groups of buildings.

After completing the feature extraction, the feature maps on each band were superimposed with the block vector data, and the average value of different features in each block were calculated by the ArcGIS software. The results of statistical calculations for multiple features were used as the feature vector of the block to classify the damage of groups of buildings, and the multi-feature classification model was established. According to the statistical results of the aforementioned features, SVM, CART and RF in machine learning methods were used to obtain the classification accuracy of the damage of groups of buildings.



**Figure 8.** Quantitative statistical analysis curves for different features of each damage category.

The final classification results obtained by using the three machine learning classifiers are shown in Table 2. SVM had the highest classification accuracy, reaching 72%. The parameter C corresponding to the optimal accuracy of SVM was 1, which was the penalty coefficient for the classification error item in this classifier. The larger C corresponded to a greater punishment degree for the misclassification of samples. Therefore, the higher the accuracy in training samples, the lower the generalization ability. Conversely, the smaller C corresponded to the more misclassified samples that were allowed in the training sample then the stronger the generalization ability.

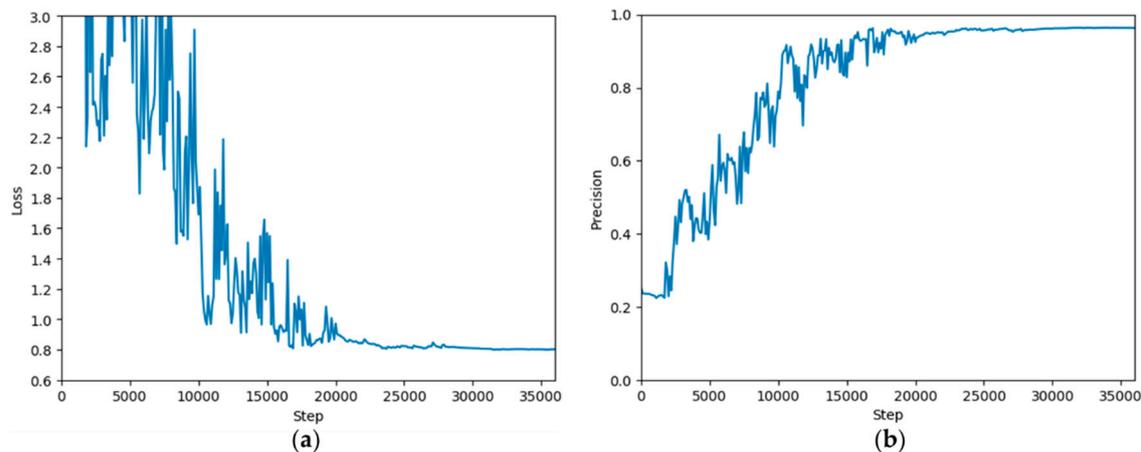
**Table 2.** Comparison of three machine learning classifiers for groups of buildings damage classification.

Classifier	Optimal Precision	Parameter Corresponding to the Optimal Precision
SVM	72.5%	Penalty factor for error terms: C = 1
CART	70%	Maximum depth of the tree: max_depth = 6
RF	60%	Number of trees in RF: n_estimators = 3

#### 4.2. Improved CNN Classification Method

In this study, the deep learning library TensorFlow was used to replicate the Inception V3 network model. To facilitate the model training, the Separate and the Combination layers shown in Figure 3 were removed. During the training process, the parameters were gradually optimized and adjusted, and the final parameters were as follows: the optimizer was Stochastic Gradient Descent with mini batches, the batch size was 32, the dropout ratio was 0.5. The learning rate started from 0.01 and dropped to 0.001 when the training reached 20,000 steps, and to 0.0001 at 28,000 steps.

After the training, the accuracy of the Inception V3 model finally reached 96.39% for the verification set. The loss value and accuracy change for the verification set during this training process are shown in Figure 9.



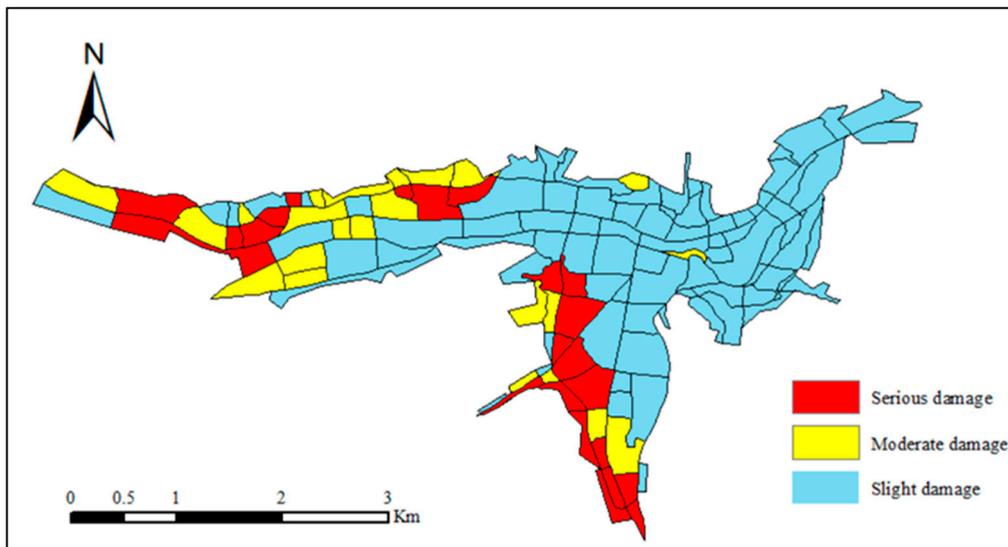
**Figure 9.** Loss and accuracy changes for the validation set during training: (a) loss curve, (b) accuracy curve.

The trained model was then tested on the test set of 1440 rectangular images, with an accuracy of 92.22%. The confusion matrix between the test results on the test set and the ground truth is shown in Table 3. The confusion matrix showed that the number of wrong judgements between the serious and the moderate damage was 56 and the number of wrong judgements between the moderate and the slight damage was 52, and there was almost no misjudgment between the serious and the slight damage.

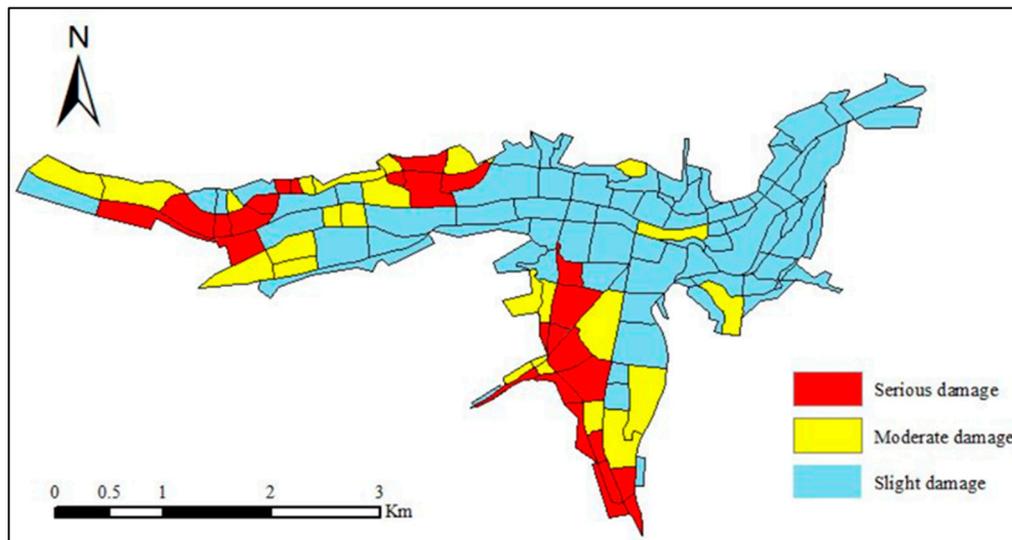
**Table 3.** Confusion matrix of ground truth and test results of the CNN model.

Category	Ground Truth				Total	User Accuracy
	Serious Damage	Moderate Damage	Slight Damage	Negative Samples		
Serious damage	457	29	0	0	486	94.03%
Moderate damage	27	249	15	0	291	85.57%
Slight damage	1	37	531	0	569	93.32%
Negative samples	1	0	2	90	93	96.77%
Total	486	315	549	90	1440	
Producer accuracy	94.03%	79.05%	96.72%	100.00%		92.22%

Finally, the Separate and Combination layers were added to the Inception V3 network to directly input aerial images of the Yushu earthquake and the vector map of urban blocks, and to output the vector map of building damage. After completion of the process described above, the building damage information was extracted from the aerial images after the Yushu earthquake (Figure 10). To compare the results extracted by the improved CNN method with the ground truth, visual interpretation was carried out by combining the pre-earthquake and post-earthquake images. The visual interpretation results were taken as the ground truth reference results. The results obtained by visual interpretation is shown in Figure 11. Together, Figures 10 and 11 show that the extraction result was basically consistent with the visual interpretation result, although there were still some misjudgments in some blocks.



**Figure 10.** Damage classification map for groups of buildings based on the improved CNN method.



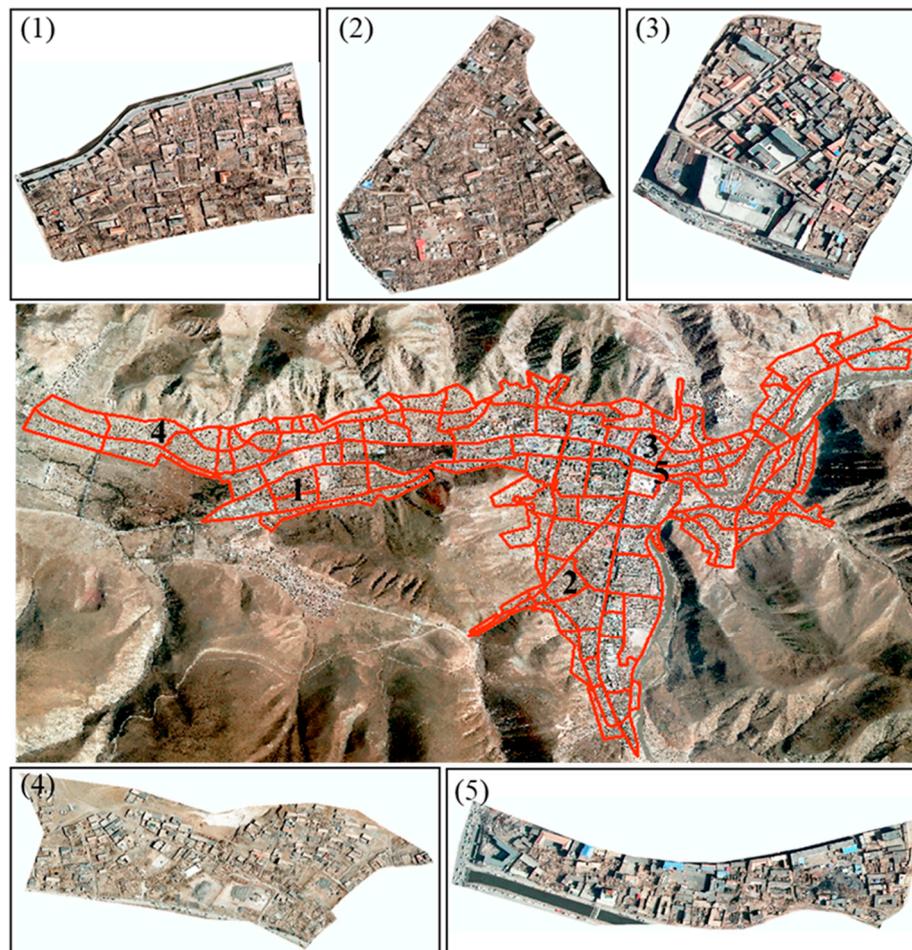
**Figure 11.** Damage classification map for groups of buildings based on visual interpretation.

To quantitatively compare the results of CNN extraction and visual interpretation, the confusion matrix table is shown in Table 4. The mis-detected blocks mainly resulted from wrong judgement between the moderate and the slight damage, and the number of wrong judgements between the two categories was 7. Finally, the overall accuracy was 90.07%, and Kappa Coefficient was 0.81.

**Table 4.** Confusion matrix of visual interpretation and extraction results of the improved CNN method.

Visual Interpretation Results					
Category	Serious Damage	Moderate Damage	Slight Damage	Total	User Accuracy
Serious damage	21	1	1	23	91.3%
Moderate damage	2	19	2	23	82.61%
Slight damage	2	5	78	85	91.76%
Total	25	25	81	131	
Producer accuracy	84%	76%	96.29%		90.07%

Figure 12 shows examples of some typical blocks. Blocks (1), (2) and (3) are the correct classification blocks, while (4) and (5) are the wrong classification blocks. Table 5 shows the probability distribution of the classification results of the five blocks in Figure 12. From the probability of each category, we concluded that the probability difference between the misclassified category and the correct category was not significant, although the classification category was wrong. This indicated that the final classification category was not the only judging criterion for the blocks with difficulty to distinguish collapse types. It was also possible to determine the confidence degree of each collapse type according to the probability of each category, to evaluate the collapse type of each block more accurately.



**Figure 12.** Block building classification example: (1) moderate damage, (2) serious damage, (3) slight damaged, (4) moderate damage, (5) moderate damage.

**Table 5.** Comparison of the test results of the improved CNN method with the results of visual interpretation.

Block	Categories Obtained by Visual Interpretation	Test Results of the Improved CNN Method			
		Categories	Probability of Serious Damage	Probability of Moderate Damage	Probability of Slight Damage
(1)	Moderate damage	Moderate damage	29.29%	48.35%	11.79%
(2)	Serious damage	Serious damage	79.12%	11.09%	0.01%
(3)	Slight damage	Slight damage	0.02%	0.23%	96.30%
(4)	Moderate damage	Serious damage	27.64%	23.91%	17.55%
(5)	Moderate damage	Slight damage	0.83%	25.50%	56.42%

### 4.3. Discussion

In this study, some measures were taken to prevent overfitting in the training of the CNN. In fact, there was not enough data on building damage to meet the requirements. When these limited samples were used to train the large CNN, it tended to over fit. For this reason, the data set was enhanced and expanded to increase the diversity of samples. Furthermore, in the training process of the CNN, when the loss function value did not decline after a certain number of steps, the training was terminated in advance to avoid excessive learning.

In this study, the classification of damaged groups of buildings achieved high accuracy through training, but there were also some wrong classifications, which were mainly attributed to the following two reasons. First, the background environment of the high-resolution remote sensing images after an earthquake are far more complex than those of the natural images, so the background environment has a greater impact on the classification results. For example, the characteristics of collapsed adobe house and bare soil are very similar, thus, it is easy to judge bare soil as collapsed adobe house. Second, because the labeling of sample categories is judged by visual interpretation and lacks the support of ground survey data, this may cause some error labels.

The improved CNN approach proposed in this study can be extended to other CNNs. With the continuous advancement of deep learning, a CNN with higher accuracy and better performance will be developed in the near future. Adding Separate and Combination layers to better network architectures may allow the achievement of a better classification effect.

## 5. Conclusions

By combining an improved CNN approach with GIS data, this paper proposed a new strategy to extract the damage information of groups of buildings in remote sensing images after earthquakes. From our experiment, we found that CNNs could effectively solve the problem of difficult feature selection, which is an advantage over traditional object-oriented classification methods. Compared with the traditional multi-feature machine learning classification method constructed by artificial feature extraction, accuracy is greatly improved, and a satisfactory effect can be achieved. Block vector data in GIS can form a meaningful boundary for groups of buildings, effectively replacing image segmentation and avoiding its fragmentary and unsatisfactory results. At the same time, our method was able to avoid the big error that arises when the CNN is used to predict irregular shapes and different sizes of blocks.

However, due to the limited number of samples used for training in CNN and the confusion between collapsed buildings and bare ground, classification errors for the groups of buildings can be easily caused, meaning that there are still some errors in comparison with the ground truth.

Therefore, extending the training data set, including remote sensing images of different types and resolutions, is future work to be tested for improvement. The methods combining multi-classifiers including CNN should be considered to improve the classification accuracy of groups of buildings.

**Author Contributions:** All authors contributed in a substantial manner to the manuscript. H.M. conceived, designed and performed research and wrote the manuscript. Y.L. and Y.R. made contributions to the design of the research and data analysis. All authors discussed the basic structure of the manuscript. All authors read and approved the submitted manuscript.

**Funding:** This research was funded by the National Key Research and Development Program, Project NO. 2017YFC1500902.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Dell'Acqua, F.; Gamba, P. Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives. *Proc. IEEE* **2012**, *100*, 2876–2890. [[CrossRef](#)]
2. Chen, W. *Research of Remote Sensing Application Technology Based on Earthquake Disaster Assessment*; China Earthquake Administration Lanzhou Institute of Seismology: Lanzhou, China, 2007.
3. He, M.; Zhu, Q.; Du, Z. 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. Gong, L.; Li, Q.; Zhang, J. Earthquake building damage detection with object-oriented change detection. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium-IGARSS, Melbourne, Australia, 21–26 July 2013; pp. 3674–3677.
6. Janalipour, M.; Mohammadzadeh, A. Building Damage Detection Using Object-Based Image Analysis and ANFIS from High-Resolution Image (Case Study: BAM Earthquake, Iran). *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *9*, 1937–1945. [[CrossRef](#)]
7. Nie, J.; Yang, S.; Fan, Y. Building losses assessment for Lushan earthquake utilization multisource remote sensing data and GIS. In Proceedings of the Mippr: Automatic Target Recognition and Navigation, Enshi, China, 31 October–1 November 2015; p. 98120J.
8. Ye, X.; Wang, J.; Qin, Q. Damaged building detection based on GF-1 satellite remote sensing image: A case study for Nepal MS8.1 earthquake. *Acta Seismol. Sin.* **2016**, *38*, 477–485.
9. 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](#)]
10. Chen, J. *Research on Extraction Methods of Damaged Buildings after Earthquake Based on Optical Remote Sensing*; China Earthquake Administration Lanzhou Institute of Seismology: Lanzhou, China, 2018.
11. Janalipour, M.; Mohammadzadeh, A. A fuzzy-ga based decision making system for detecting damaged buildings from high-spatial resolution optical images. *Remote Sens.* **2017**, *9*, 349. [[CrossRef](#)]
12. Zhang, L.; Zhang, J. Research on building collapse rate calculation method based on object-oriented classification. *Earthquake* **2009**, *29*, 139–145.
13. Li, F.; Ma, C.; Zhang, G. Rapid assessment of earthquake damage to buildings based on object-oriented classification. *J. Henan Polytech. Univ. Sci.* **2011**, *30*, 55–60.
14. Wang, Y.; Wang, X.; Dou, A. Building damage detection of the 2008 Wenchuan, China earthquake based on object-oriented classification method. *Earthquake* **2009**, *29*, 54–60.
15. Nogueira, K.; Penatti, O.A.B.; Santos, J.A.D. Towards better exploiting convolutional neural networks for remote sensing scene classification. *Pattern Recognit.* **2017**, *61*, 539–556. [[CrossRef](#)]
16. Zhu, X.X.; Tuia, D.; Mou, L. Deep learning in remote sensing: A comprehensive review and list of resources. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–36. [[CrossRef](#)]
17. Lecun, Y.; Bottou, L.; Bengio, Y. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [[CrossRef](#)]
18. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, CA, USA, 3–6 December 2012; pp. 1097–1105.
19. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
20. Szegedy, C.; Vanhoucke, V.; Ioffe, S. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 27–30 June 2016; pp. 2818–2826.
21. He, K.; Zhang, X.; Ren, S. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 27–30 June 2016; pp. 770–778.
22. Haykin, S. *Neural Networks and Learning Machines*, 3/E; Pearson Education: Upper Saddle River, NJ, USA, 2010.
23. Chen, M.; Wang, X.; Dou, A. The extraction of post-earthquake building damage information based on convolutional neural network. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *42*, 161–165. [[CrossRef](#)]

24. Wang, L.; Wang, X.; Ding, X. Study on loss assessment of construction earthquake damage based on remote sensing and GIS. *Earthquake* **2007**, *27*, 77–83.
25. Samadzadegan, F.; Rastiveisi, H. Automatic detection and classification of damaged buildings, using high resolution satellite imagery and vector data. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2008**, *37*, 415–420.
26. Li, Y.; Hao, Z.; Lei, H. Overview of convolutional neural networks. *J. Comput. Appl.* **2016**, *36*, 2508–2515.
27. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv* **2015**, arXiv:1502.03167.
28. Duarte, D.; Nex, F.; Kerle, N. Satellite image classification of building damages using airborne and satellite image samples in a deep learning approach. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *4*, 89–96. [[CrossRef](#)]
29. Ji, M.; Liu, L.; Buchroithner, M. Identifying collapsed buildings using post-earthquake satellite imagery and convolutional neural networks: A case study of the 2010 Haiti earthquake. *Remote Sens.* **2018**, *10*, 1689. [[CrossRef](#)]
30. Cao, D.; Shi, X.; Zhang, J. Study on the statistical characteristics of seismic disaster information of building in remote sensing image. *Remote Sens. Land Resour.* **2001**, *13*, 42–46.
31. Wei, Z.; Huanfeng, S.; Chunlin, H. Building earthquake damage information extraction from a single post-earthquake PolSAR image. *Remote Sens.* **2016**, *8*, 171.
32. Perez, L.; Wang, J. The effectiveness of data augmentation in image classification using deep learning. *arXiv* **2017**, arXiv:1712.04621.



© 2020 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/>).