remote sensing

MDPI

Article

Transfer Learning for Improving Seismic Building Damage Assessment

Qigen Lin [1], Tianyu Ci [2,3], Leibin Wang [4], Sanjit Kumar Mondal [1], Huaxiang Yin [1] and Ying Wang [2,3,*]

[1] Institute for Disaster Risk Management, School of Geographical Sciences, Nanjing University of Information Science & Technology, Nanjing 210044, China; linqigen@mail.bnu.edu.cn (Q.L.); sanjitmondal@nuist.edu.cn (S.K.M.); yinhuaxiang@nuist.edu.cn (H.Y.)
[2] Key Laboratory of Environmental Change and Natural Disaster of Ministry of Education, Beijing Normal University, Beijing 100875, China; city@mail.bnu.edu.cn
[3] Academy of Disaster Reduction and Emergency Management, Ministry of Emergency Management and Ministry of Education, Beijing 100875, China
[4] College of Resources and Environment Science, Hebei Normal University, Shijiazhuang 050024, China; leibin.wang@hebtu.edu.cn
* Correspondence: wy@bnu.edu.cn; Tel.: +86-010-5880-5486

Abstract: The rapid assessment of building damage in earthquake-stricken areas is of paramount importance for emergency response. The development of remote sensing technology has aided in deriving reliable and precise building damage assessments of extensive areas following disasters. It is well documented that convolutional neural network methods have superior performance in earthquake building damage assessment compared with traditional machine learning methods. However, deep learning models require a large number of samples, and sufficient numbers of samples are usually not available in the newly earthquake-stricken areas rapidly enough. At the same time, the historical samples inevitably differ from the new earthquake-affected areas due to the discrepancy of regional building characteristics. For this purpose, this study proposes a data transfer algorithm for evaluating the impact of a single historical training sample on the model performance. Then, beneficial samples are selected to transfer knowledge from the historical data for facilitating the calibration of the new model. Four models are designed with two earthquake damage building datasets and the performance of the models is compared and evaluated. The results show that the data transfer algorithm proposed in this work improves the reliability of the building damage assessment model significantly by filtering samples from the historical data that are suitable for the new task. The performance of the model built based on the data transfer method on the test set of new earthquakes task is approximately 8% higher in overall accuracy compared with the model trained directly with the new earthquake samples when the training data for the new task is only 10% of the historical data and is operating under the objective of four classes of building damage. The proposed data transfer algorithm has effectively enhanced the precision of the seismic building damage assessment in a data-limited context. Thus, it could be applicable to the building damage assessment of new disasters.

Keywords: building damage; transfer learning; earthquake; deep learning; convolutional neural networks

check for updates

Citation: Lin, Q.; Ci, T.; Wang, L.; Mondal, S.K.; Yin, H.; Wang, Y. Transfer Learning for Improving Seismic Building Damage Assessment. *Remote Sens.* **2022**, *14*, 201. https://doi.org/10.3390/rs14010201

Academic Editors: Wei Chen, Paraskevas Tsangaratos and Haoyuan Hong

Received: 25 November 2021
Accepted: 31 December 2021
Published: 2 January 2022

1. Introduction

Major earthquakes are one of the most destructive and devastating natural disasters, often triggering extensive secondary disasters that result in widespread building collapse and infrastructure damage [1–5]. In this context, the extensive building damage triggered by earthquake disasters is an overwhelming cause of human casualties and loss of assets [6,7]. Rapidly assessing the damage of the building during the emergency response period following the earthquake is an essential process, which would help to mitigate the loss of life and property [8–10]. Although traditional post-disaster field investigations are more accurate for damage mapping, it is time-consuming and less secure. In comparison, satellite or UAV-based imagery allows for faster, more precise, and more extensive coverage of the disaster-affected area, which could be analyzed for building damage assessment purposes, which could support the emergency response [11,12].

Earthquake building damage detection and assessment are performed based on pre- and post-earthquake images or a single post-earthquake image using various types of remote sensing data, including optical, LiDAR, and SAR [13,14]. In terms of detection and classification methods, visual interpretation [15], transformation-based methods [16], and post-classification comparison from multiple periods have been used [17,18]. In recent years, machine learning and deep learning methods have also been employed for earthquake building damage assessment [19–21]. Among these, the deep learning methods (Convolutional Neural Networks, CNNs) have been demonstrated to be more effective at assessing the seismic building damage classes, with their superior accuracy owing to their automatic feature extraction capability that allows them to extract the high-level features more effectively [22–24]. CNNs can not only automatically learn low-level and mid-level features from the raw images, but also automatically extract higher-level features with discriminating representation according to the problem [25]. The low-level CNNs could extract the basic features of buildings, such as the color, the edges, and the corners, and the deep-level CNNs derives high-level features, such as problem-specific features related to the new task [26]. Such automatic feature extraction abilities are an advantage compared to traditional machine learning models that require setting the features through expert experience for remote sensing image building detection [27,28]. The powerful learning capability of deep CNNs is mainly due to the use of multiple feature extraction stages, which can automatically learn representational features from the raw images according to the target task. Recent studies have demonstrated the impressive performance of deep CNNs for automatic feature extraction and the detection of buildings [29–31]. For instance, Vetrivel et al. (2016) compared the results of two supervised classification methods (expert knowledge extraction of image features and convolutional neural network automatic feature extraction) for collapsed building detection. The results showed that the convolutional neural network that automatically extracted features performed better. Moreover, the authors noted that with the increase in the remote sensing image spatial resolution, the performance of the convolutional neural network method for feature extraction will have more significant advantages [32]. Ci et al. (2019) built an earthquake building damage class assessment model based on the combination of a convolutional neural network and an ordered regression. The results demonstrated that the classification accuracy of the proposed model was significantly higher than that of traditional machine learning methods [33]. However, the established deep learning-based seismic building damage assessment models require a large amount of sample data for training, which is usually difficult to obtain during the emergency response period in a new earthquake-affected area. At the same time, these models were trained based on damaged building data within a single earthquake-stricken area, which affects the extrapolation capability of the trained models when they are applied to new earthquake scenarios that lack data. In addressing the issue of insufficient training data, transfer learning has been mostly used in the literature. For instance, Yu et al. (2019) found that the performance of the accuracy of image-based tumor diagnosis could effectively be improved by fine-tuning the pre-trained image deep learning model [34]. Qin et al. (2021) demonstrated that transfer learning outperforms traditional convolutional neural networks for landslide detection [35]. Few works in the literature use transfer learning to assess and recognize structural and building damage, but most of these works focus on feature-based and knowledge-based transfer learning [36,37]. Only a small amount of attention has been focused on the transfer of sample instance data. Among these, there are very limited seismic building damage studies based on data transfer.

To address this issue, this study proposes a data transfer algorithm based on instance evaluation to be able to maximize the use of historical heterogeneous data. Based on this knowledge, trained models, and historical data, the beneficial samples from historical data are effectively screened out to be used for improving the new task prediction. We attempt to solve the issue that only very limited labeled samples are available in the emergency response phase after the occurrence of a new earthquake.

2. Materials and Methods

2.1. Post-Earthquake High-Resolution Remote Sensing Images

In this study, high-resolution remote sensing images following two earthquakes were used to develop a transfer learning model for assessment of seismic building damage classes. These two earthquakes were the Yushu earthquake [38] in Qinghai Province in 2010 and the Ludian earthquake [39] in Yunnan Province in 2014. On 14 April 2010, an earthquake with a magnitude of 6.9 Mw occurred in the Yushu Tibetan autonomous prefecture, Qinghai Province, with an epicenter near 33.2°N and 96.6°E and a depth of 14 km. The Yushu earthquake destroyed a large number of buildings in Jiegu town and the surrounding areas in the Yushu Tibetan Autonomous Prefecture. The remote sensing images following the Yushu earthquake were 0.1 m spatial resolution aerial image data acquired on 16 April 2010, containing R, G, and B bands, and the collapsed buildings images near the Jiegu town are shown in Figure 1a. On 3 August 2014, an earthquake with a magnitude of 6.1 Mw occurred near 27.1°N and 103.3°E in the Ludian County, Zhaotong City, Yunnan Province, and the epicenter had a depth of 12 km. The remote sensing images following the Ludian earthquake were 0.2 m spatial resolution UAV image data acquired on 7 August and 14 August 2014, which contained R, G, and B bands, and the post-quake collapsed buildings near Longtoushan town are shown in Figure 1b.


(a) Yushu Earthquake (b) Ludian Earthquake

Figure 1. Damaged building on remote sensing images following the (**a**) Yushu earthquake and (**b**) Ludian earthquake.

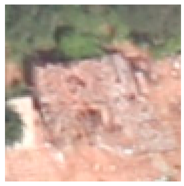
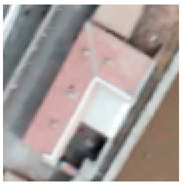
2.2. Labeling of Damaged Buildings in Remote Sensing Images

The visual interpretation approach was used to vectorize the building outlines on the post-earthquake remote-sensing images. Based on the analysis of the length-to-width ratio of multiple vectorized buildings and the spatial resolution of remote sensing images, it was found that producing the building samples in a standard image of size 88×88 could better present the characteristics of various buildings. The building damage classes were set into four categories, including intact damaged (no observable damage), lightly damaged, heavily damaged, and collapsed, by examining the building features, including outline, geometry, texture, and the relationship with the surrounding area, on the remote sensing images and by incorporating the criteria for classifying earthquake-induced building damage from the literature. The damage classes of the building samples in the images were evaluated one by one through visual interpretation, with the field investigations after the 2014 Ludian earthquake as a reference. The numbers of building samples of each damage class for the Ludian and Yushu earthquakes are shown in Table 1 after visual interpretation, and the morphological samples of the damaged buildings in remote sensing images are shown in Table 2.

Table 1. The sample size of damaged buildings in the Ludian dataset and the Yushu dataset [33].

Damage Classes	Description	Ludian Dataset	Yushu Dataset
D0	No observable damage	2680	778
D1	Light damage	5013	918
D2	Heavy damage	2807	665
D3	Collapse	3280	1140
Total		13,780	3501

Table 2. Sample of damaged buildings on remote sensing images in the Ludian and Yushu datasets.

Dataset	No Observable Damage	Light Damage	Heavy Damage	Collapse
Ludian dataset				
Yushu dataset				

The dataset obtained from the damaged buildings of the Ludian earthquake is referred to as the "Ludian dataset". The dataset obtained from the damaged buildings of the Yushu earthquake is referred to as the "Yushu dataset". The Ludian dataset contains a total of 2680 buildings with D0 damage level, 5013 buildings with D1 damage level, 2807 buildings with D2 damage level, and 3280 buildings with D3 damage level, for a total of 13,780 buildings. The Yushu dataset contains a total of 778 buildings with D0 damage, 918 buildings with D1 damage, 665 buildings with D2 damage, and 1140 buildings with D3 damage, for a total of 3501 buildings. In general, the sample size of the building samples of each damage level is relatively balanced within the two earthquake datasets, and the Ludian dataset is more abundant.

Although both the "Ludian dataset" and the "Yushu dataset" contain remote sensing image data acquired following the earthquake, the building damage classes and label evaluation criteria are identical in these two datasets. There are still differences between the two datasets due to various reasons, such as their geographical regions, imaging periods, aerial platforms, sensor parameters, local customs, etc., which are reflected by the different distribution of DN values of pixels in the remote sensing images. Such differences are tangible in the actual seismic building damage assessment. This study attempts to investigate the extrapolation of the model to different seismic scenarios, different regions, and the effect of the different amount of training data on the accuracy.

2.3. Transfer Learning

Transfer learning is a method of transferring the knowledge trained or learned from deep learning models to a new pending task. The advantage of transfer learning over traditional machine learning models is that the trained models and knowledge from historical data can be transferred to the new task instead of retraining a new model. This significantly reduces the number of training samples required for a new model, while at the same time it can improve the accuracy of the new model with the help of existing models, knowledge, and historical data [40].

Regarding the training for the Ludian dataset and the Yushu dataset, the Ludian dataset, with a more abundant sample size, is chosen as the historical earthquake scenario, and the Yushu dataset, with a relatively smaller sample size, is used as the new task scenario. Such a design is similar to the practical new earthquake scenario; i.e., there are more historical earthquake building damage samples, while new samples are usually relatively limited for the new earthquakes. In this case, the challenge for the building damage assessment of new earthquake scenarios is that there are already a small number of training samples and test samples with the same distribution. However, the number of these samples is too small, so it may not be possible to derive a reliable model by relying on these limited training and testing samples. It is also difficult to increase the number of training samples of new earthquakes, which would take more time and would not be sufficient for the rapid damage assessment for the emergency response. At the same time, there are a larger number of auxiliary samples (historical data). However, the distribution between these auxiliary samples and the training samples of the new earthquake scenario is not consistent.

The problem to be solved here is how to derive, with the help of these auxiliary samples, a new model that has the highest possible classification accuracy when applied to the new test samples. For this purpose, the following model framework is designed in this study (Figure 2).

Firstly, a convolutional neural network (CNN) combined with an ordered regression model is employed to build a seismic building damage class assessment model based on the Ludian dataset. This deep learning model architecture consists of a CNN feature extractor, i.e., the visual geometry group network (VGG), and an ordered regression classifier, which has been demonstrated to outperform other model architectures in seismic building damage class assessment. In this study, the VGG-OR network is employed to develop a VGG-OR(LD) model for the Ludian dataset. Then, four schemes are designed to evaluate the performance of different transfer learning schemes through comparative tests. The first scheme is to directly transfer the trained model, i.e., VGG-OR(LD), to the testing set of the "new" Yushu dataset for prediction. The second scheme is to retrain a new model based on the "new" Yushu dataset and the VGG-OR architecture, i.e., VGG-OR(YS). The third scheme is to utilize the training set of the "new" Yushu dataset to fine-tune the parameters of the VGG-OR(LD) model to derive a new model, i.e., VGG-OR(FT). The last scheme is to propose a new data transfer algorithm. The proposed data transfer algorithm is designed to identify potential samples of the historical data that are beneficial to the Yushu dataset, and these beneficial historical samples are used jointly with the training samples of the Yushu dataset to fine-tune the parameters of the VGG-OR(LD) model to derive a new model, i.e., VGG-OR(DT).

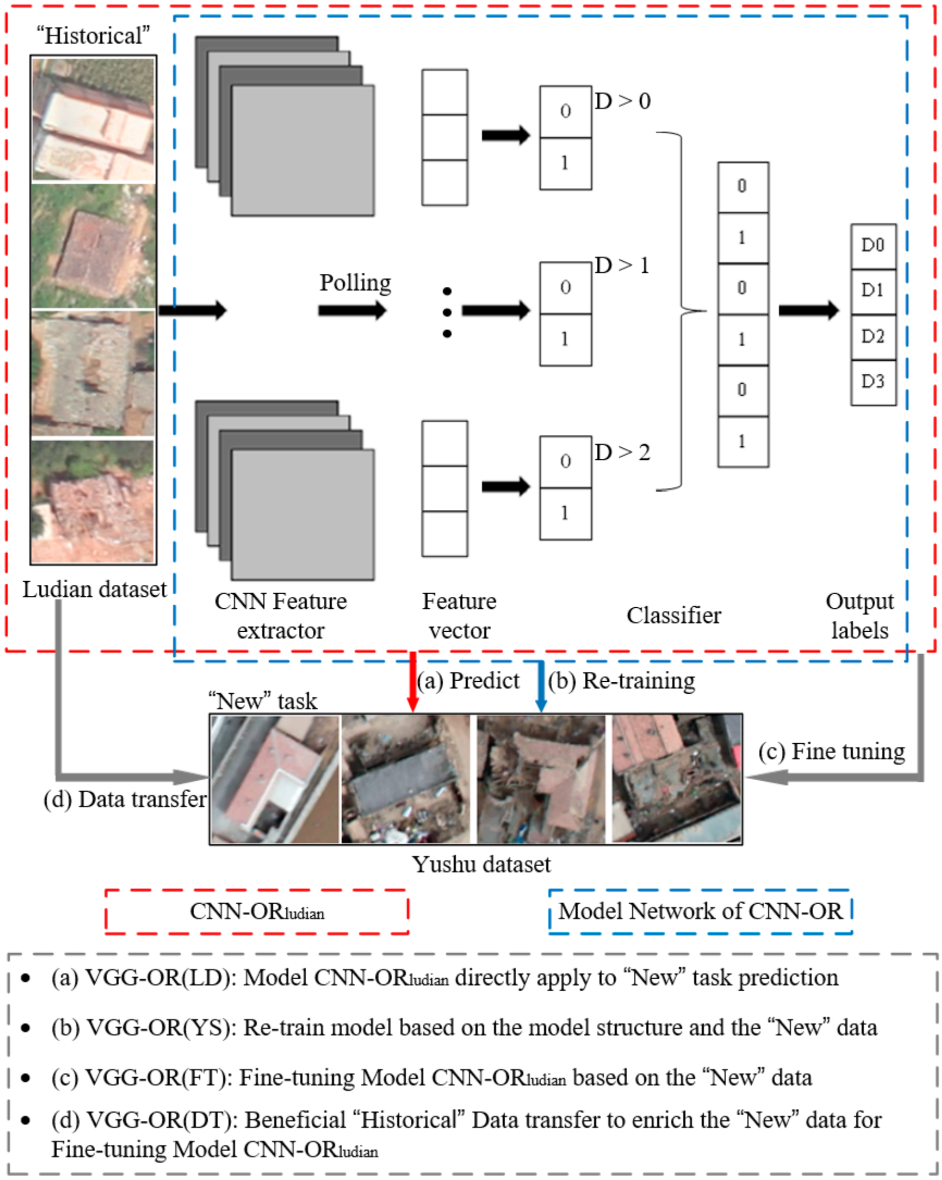


Figure 2. Model framework performed in this study.

The basic idea of the data transfer algorithm proposed in this study is that although the historical samples and the target training samples are more or less different from each other, there should still be a portion of the historical sample that is suitable for assisting in training the model that could be adapted to the test sample of the new task. Therefore, the critical implementation of this data transfer algorithm is to build an evaluation model that retains the eligible historical data and removes the historical data that differs significantly from the target data. These retained historical samples would be used as added data to be used with the target training samples to improve the reliability of the model after filtering. The data transfer algorithm process proposed in this study is illustrated in Table 3 and Figure 3. The effect of the historical sample on the accuracy of the model in the test set of the new task is evaluated using the loss function proposed by Koh and Liang (2017), and the beneficial samples are selected for improving the model with the training set of the new task [41].

Table 3. Data transfer algorithm processes.

Input
Z_S: Historical dataset (Ludian dataset)
$Z_{T,train}$: Training set from the new task dataset (Yushu dataset)
$Z_{T,val}$: Validation set from the new task dataset (Yushu dataset)
i, j: subscript
f_θ : f refers to the CNN model and θ refers to the parameter of the CNN model
$I_{loss}(z_i, z_j)$: loss function
Start
θ_{pre}: Pre-trained parameters obtained by training based on the historical dataset Z_S
$f_{\theta_0} \overset{Z_S}{\rightarrow} f_{\theta_{pre}}$
Processes
1. For each $z_i \in Z_S,\ z_j \in Z_{T,val}$, calculate the loss function on the validation set:
For z_i in Z_S
For z_j in $Z_{T,val}$
Calculating $I_{loss}(z_i, z_j)$
If $\sum\limits_j I_{loss}(z_i, z_j) > 0,$ then remove z_i from Z_S
2. Beneficial samples Z'_s
4. Combined $Z_{T,train} \cup Z'_s$ *as* the *training set*, Fine-tuning the pre-trained parameters θ_{pre} *to* θ_T
$f_{\theta_{pre}} \overset{Z_{T,train} \cup Z'_s}{\rightarrow} f_{\theta_T}$
Output
Z'_s: Beneficial samples after selection
θ_T: *Optimized model parameters*

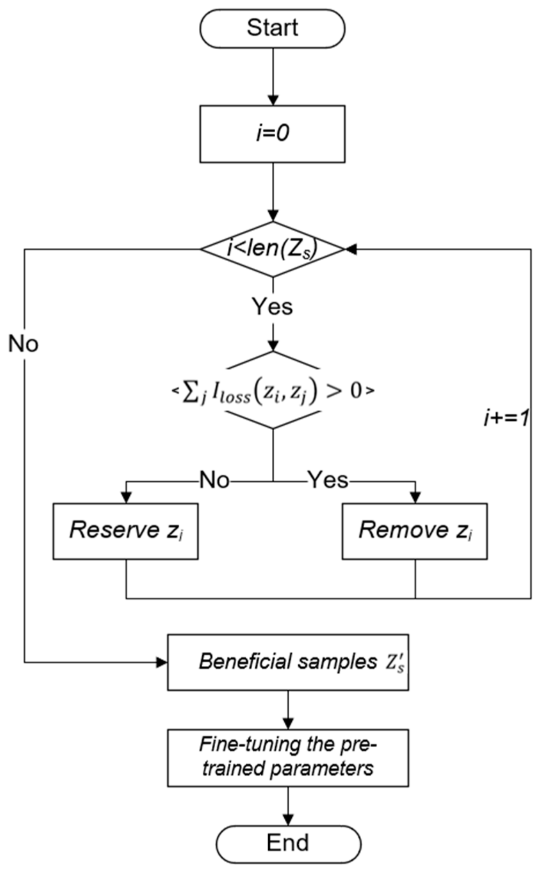


Figure 3. The flowchart of the data transfer algorithm proposed in this study.

In the existing transfer learning studies, it is found that the best transfer learning effectiveness is achieved when the amount of data in the new task is close to 10% of the amount of the historical data. When the data size of the new task is more than 20% of the historical data size, the effectiveness of transfer learning decreases and may even be weaker than that of direct training based on only the new training data [42,43]. Therefore, this study limits the data size of the new task to approximately 10% of the historical data size when comparing the performance of the models, and also explores the effect of the training sample size for the new task on the performance of the different models.

2.4. Evaluation Methods

The confusion matrix, overall accuracy, kappa coefficient and mean squared error (MSE) were used to evaluate the performance of the four models. The training procedure of all the models in this study randomly divides the dataset into the training set, validation set, and test set. The three sets (training set, validation set and test set) are not intersected with each other. The training set is used to train the parameters of the deep learning model. The validation set is used for model selection, where the optimal model is selected based on the performance of the model with the validation set, or the models trained with different hyperparameters, including model structure, learning rate, number of iterations, etc., are selected based on the performance of these hyperparameters with the validation set. The test set is employed to examine the extrapolation performance of the final selected optimal model. Since the sample size of the Ludian dataset is comparatively sufficient for the training of the VGG-OR(LD) model, the Ludian dataset is randomly divided into three parts according to the ratio of 80%:10%:10%, mapping to the training set, the validation set, and the test set, respectively. For a more detailed explanation of the VGG-OR(LD) model, refer to the literature [33]. For the new earthquake scenario, i.e., the Yushu dataset, 60% of the samples are randomly divided into the training set, 20% into the validation set, and the remaining 20% into the test set, and the evaluation metrics are based on the test set. Thus, for the VGG-OR(LD) model, the test set of the Yushu dataset was directly employed to evaluate the performance of the model. For the VGG-OR(YS) model, the model is trained using 60% of the Yushu training dataset combined with the CNN-OR model structure, then 20% of the Yushu validation dataset is utilized to fine-tune the parameters of the VGG-OR(YS) model, and finally the performance of the VGG-OR(YS) model is evaluated using 20% of the Yushu test dataset. For the VGG-OR(FT) model, the training and validation sets of the Yushu dataset were employed to fine-tune the parameters of the VGG-OR(LD) model, and then the performance of the VGG-OR(FT) model was evaluated with the Yushu test dataset. For the VGG-OR(DT) model, the data transfer algorithm proposed in this work is used to identify the beneficial samples in the Ludian dataset, and these beneficial samples are used together with the training and validation sets of the Yushu dataset to fine-tune the VGG-OR(LD) model, and then the Yushu test dataset is employed to evaluate the performance of the VGG-OR(DT) model. The details of each evaluation method are described as follows.

The confusion matrix is widely used to evaluate the agreement of the observation and the model's classification, which is one of the underlying methods for evaluating the performance of remote sensing image classification algorithms. This is achieved by counting the number of algorithmic classification result labels corresponding to the true category labels on a sample-by-sample basis. A confusion matrix is represented using C. Assuming that the samples have a total of K classes, the matrix C is a K-row and K-column matrix. C (i, j) represents the total number of samples whose observation category is i and predicts category j.

The overall accuracy is the rate of consistency between the results classified by a model and the observation. It can be derived using the following formula:

$$\mathrm{OA} = \frac{\sum_{\mathrm{i}}^{\mathrm{K}} \mathrm{C}(\mathrm{i},\mathrm{i})}{\sum_{\mathrm{i}}^{\mathrm{K}} \sum_{\mathrm{j}}^{\mathrm{K}} \mathrm{C}(\mathrm{i},\mathrm{j})} \tag{1}$$

Kappa coefficients are commonly used as a measure of classification accuracy [44]. The kappa coefficient would usually be between 0 and 1, and the larger the value, the better the consistency. It is derived using the following formula:

$$p_e = \frac{\sum_i^K \left(\sum_j^K C(i,j) * \sum_j^K C(j,i) \right)}{N^2} \tag{2}$$

$$\text{Kappa} = \frac{OA - p_e}{1 - p_e} \tag{3}$$

where N denotes the total number of samples.

Mean squared error (MSE) is the mean value of the sum of squares of errors between the predicted data and the actual value. The formula is:

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 \tag{4}$$

where y_i refers to the actual value, $\hat{y}_i$ refers to the predicted value, and N denotes the total number of samples.

3. Results and Discussion

3.1. Model Trained by Ludian Dataset Can Be Transferred to Yushu Dataset

The VGG-OR(LD) model trained on the "Ludian dataset" was directly applied to the test set of the "Yushu dataset" to evaluate the performance of directly transferring the established "historical" model to the new task. The overall accuracy and kappa coefficients of the VGG-OR(LD) model for the test set of the "Yushu dataset" are shown in Table 4 in three sets of damage classes, including two classes (D0–2, D3), three classes (D0–1, D2, D3), and four classes (D0, D1, D2, D3). In general, the direct transfer model performs relatively well with the new dataset. The overall accuracy of the VGG-OR(LD) model in the new task was above 90% with a kappa coefficient value of 0.8 in the dichotomous classification (Set 1: nearly intact and damaged), which outperforms the existing related studies. The performance of the VGG-OR(LD) model in the new task decreases significantly when the damage classes are divided into three categories (Set 2: nearly intact, severe damage, and complete collapse), yet the overall accuracy is 74% and the kappa coefficient value is 0.6, which is comparable to the performance of the existing models in the literature [28,45,46]. Since the performance of the "historical" model is directly applied to the "new" dataset from a different geographical region, without using the training samples from the new dataset in the model, this is a good demonstration that the VGG-OR(LD) model performs very well in the test set of the new task.

Table 4. Performance of the VGG-OR(LD) model applied to the Yushu dataset.

Sets	Subclass	Damage Grade	Overall Accuracy	Kappa Coefficient
Set 1	Nearly intact Damaged	D0, D1, D2 D3	90.14%	0.80
Set 2	Nearly intact Severe damage Complete collapse	D0, D1 D2 D3	74.43%	0.60
Set 3	No observable damage Light damage Heavy damaged Collapse	D0 D1 D2 D3	64.28%	0.49

The overall accuracy of the VGG-OR(LD) model with the Yushu test set is approximately 64% when the damage classes are divided into four categories (Set 3: no observable damage, light damage, severe damage, and complete collapse), which is significantly lower than the overall accuracy of the VGG-OR(LD) model with the "Ludian dataset", which was 78%, indicating that the performance of the direct transfer model in more detailed classification is still somewhat inadequate. The confusion matrix of the VGG-OR(LD) model directly applied to the Yushu testing dataset with four classifications of building damage levels is shown in Table 5. From Table 5, the classification accuracy of the model for collapsed buildings (D3 level) is significantly higher than that of other building damage levels, with the producer's accuracy reaching 88% and the user's accuracy reaching 90%. At the same time, the values of the lower left and upper right corners of the confusion matrix are both zero, which suggests that the model is less likely to misclassify across damage levels. As this "historical" model is implemented straightforwardly on the "new" dataset from a geographically distinct region, this indicates that the straightforward transfer model performs fairly well. However, the direct application of the VGG-OR(LD) model to the Yushu test set still has more misclassification and omission in the two damage levels of D0 and D1, and further improvement of the model is needed. For this reason, this study further explores the use of data transfer methods to attempt to enhance the classification accuracy.

Table 5. Confusion matrix of the VGG-OR(LD) model applied to the Yushu dataset.

Observation	Prediction				
	D0	D1	D2	D3	Total
D0	46	46	19	0	111
D1	25	74	56	10	165
D2	8	27	48	21	104
D3	0	4	34	282	320
Total	79	151	157	313	700

3.2. Data Transfer Improves the Performance of a New Task with Insufficient Data

To further explore the applicability of the data transfer algorithm for improving the model performance, three model schemes are designed in this study, including VGG-OR(YS), VGG-OR(FT), and VGG-OR(DT). VGG-OR(YS) uses the VGG-OR model architecture to retrain the model with the training data in the Yushu dataset. In VGG-OR(FT), the VGG-OR(LD) model is used as the pre-trained model, and the training data in the "Yushu dataset" are used to improve the model through fine-tuning (FT). For VGG-OR(DT), the data transfer (DT) algorithm proposed in this study is used to select samples from the "Ludian dataset" that are beneficial to the new task, and these beneficial historical samples are used together with the training samples from the "Yushu dataset" to fine-tune the parameters of the VGG-OR(LD) model.

Figure 4 indicates the accuracy of the VGG-OR(YS), VGG-OR(FT), and VGG-OR(DT) models when the training sample of "Yushu dataset" accounts for 10% of the sample size of the "Ludian dataset". The results in Figure 4 showed that VGG-OR(DT) performed better in the test set of the Yushu dataset with an overall accuracy of 74%, which was significantly improved compared with the other three models, especially compared with the VGG-OR(LD) model, where the overall accuracy was improved by approximately 10%, the kappa coefficient was improved by approximately 0.12, and the MSE was lowered by approximately 0.15. This demonstrates that the data transfer algorithm proposed in this study is effective at improving the performance of the model for prediction in new tasks.

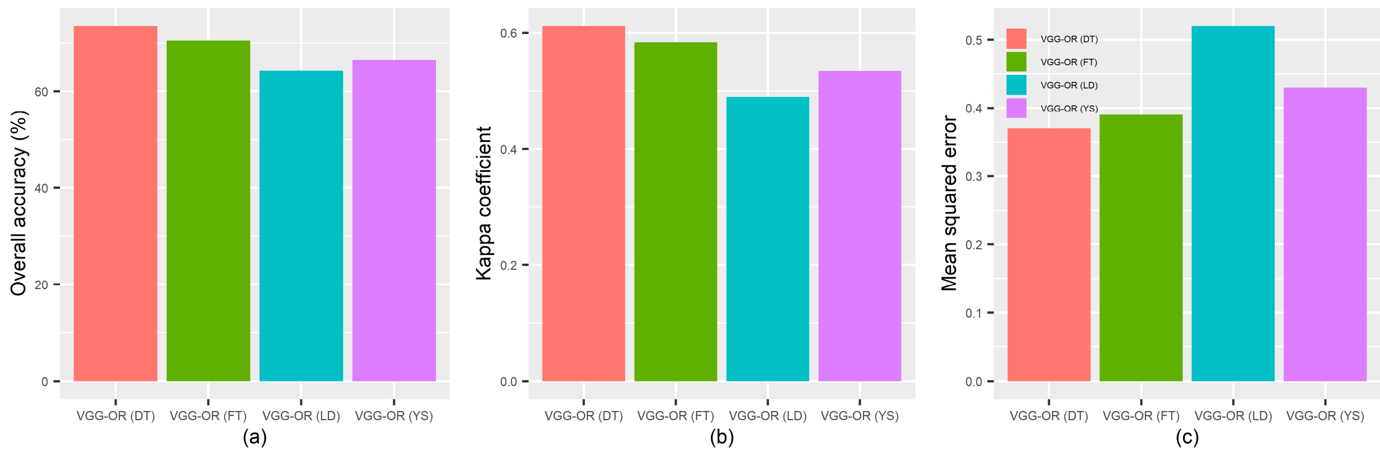


Figure 4. Performance of the transfer learning models when the sample data size of the new task is 10% of the historical data size.

It is also worth noting that the VGG-OR(FT) models also showed remarkable improvements compared to the VGG-OR(LD) model, with the overall accuracy increasing from 64% to approximately 71%, the kappa coefficient increasing from 0.49 to 0.58, and the MSE decreasing from 0.52 to 0.39. The fairly good performance of the VGG-OR(FT) models sufficiently demonstrates that the new target data could significantly enhance and improve the performance of the transfer model with the new task. In addition, both the VGG-OR(YS) and VGG-OR(FT) models show a certain degree of performance improvement compared to the VGG-OR(LD) model that was directly applied to the test set of the Yushu dataset. This suggests that using a portion of the samples from the new task for model training or parameter fine-tuning could improve the classification accuracy of the models in the test set of the new task. Therefore, it is greatly important to acquire the new dataset as soon as possible after a new disaster occurs, which would help to improve the classification accuracy of the model.

In order to explore the influence of the training data sample size on the accuracy of the transfer learning models in the new task, this study selects the data size of the "Yushu dataset" equal to 2–20% of the "Ludian dataset" as the training data, and the remaining data are used as testing data to evaluate the classification accuracy of the models. It is shown in Figure 5 that VGG-OR(DT) and VGG-OR(FT) always performed better than VGG-OR(YS). In particular, when the data size of the new task is quite limited, the improvement of the classification accuracy of the VGG-OR(DT) model built with the data transfer algorithm proposed in this study is more significant for the new task test set. When the data size approaches 20%, the classification accuracy of the three models gradually improves. This would be very helpful for the emergency response assessment when new disasters occur. It is usually inconvenient to acquire a large number of new data samples for the affected area immediately after a disaster, and the data transfer approach developed in this study could be employed to select samples from the existing historical data that are beneficial to the new task, which can be combined with a small amount of new sample data from the new disaster scenario to significantly improve the classification accuracy of the model.

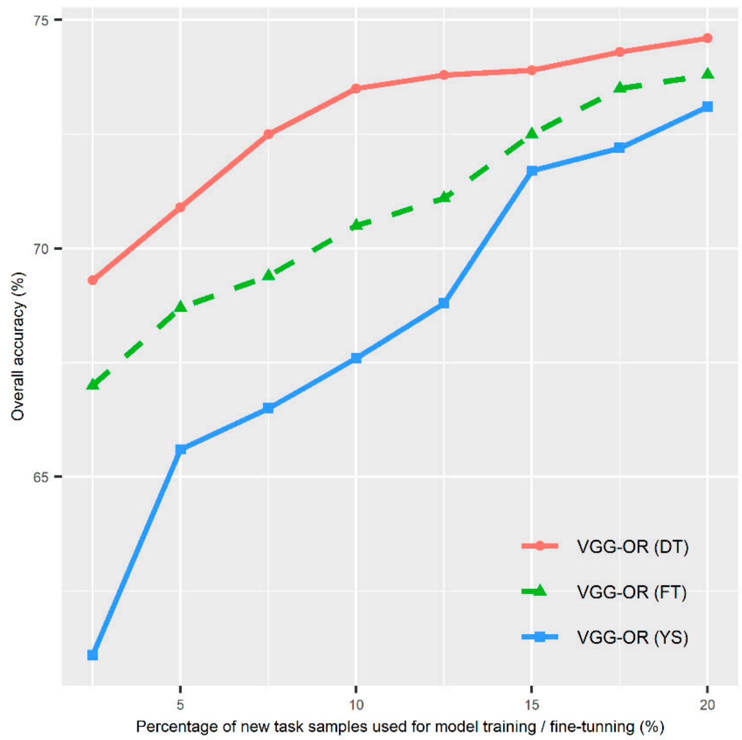


Figure 5. The relationship between the number of training samples in the new dataset and the accuracy of the transfer learning models.

3.3. The Applicability of Transfer Learning in New Scenarios

The data transfer methods proposed in this study could efficiently improve the accuracy of seismic building damage assessment. When new earthquake events occur, a building damage class assessment model applicable to the region of new earthquake events can be derived from a small number of labeled damaged building samples from the stricken region, combined with the already established model and historical dataset. The results of this study show that when the training set of the new task reaches about 10% of the historical data, i.e., 1000 samples, combined with the data augmentation method, the model has been able to achieve an overall accuracy of about 74% for the VGG-OR(DT) model derived from transfer learning in four building damage categories (Set 3 in Table 4: no observable damage, light damage, severe damage, and complete collapse). For the same number of new datasets, the overall accuracy is 66% when used to retrain the model VGG-OR(YS). The improvement of the classification accuracy of the transfer learning approach compared with the re-training model is remarkable. With this amount of data, it only requires a relatively short time to complete the manual labeling, which can greatly reduce the time required for disaster damage assessment in the disaster emergency response phase. Therefore, we believe that the reasonable use of historical data through transfer learning methods could greatly improve the efficiency of remote sensing-based damage assessment and reserve more time for emergency response decisions.

In addition to the implementation of the data transfer model in the new seismic scenario, the data transfer model may be further extrapolated in other research areas in the future to address the challenge of a lack of data samples in areas where modeling is not possible. For example, for landslide and debris flow susceptibility modeling in the mountainous areas of the Tibetan Plateau, where data are not available, the transfer of applicable landslide samples or experiences from more abundant regions may be explored. Related studies in regions with less data, such as in Africa, may also attempt to assist modeling with applicable samples from other data-rich regions.

Moreover, local site effects can have a significant impact on building damage in different locations [47]. Site effects lead to the amplification of seismic waves through stratigraphic and topographic effects, which may lead to a concentration of building damage in certain locations [48–50]. It would be beneficial for the building damage assessment in this study if site effects are taken into account. However, the purpose of this study is to propose a new data transfer model to derive beneficial samples from existing historical data through evaluation to assist in assessing rapid building damage assessment in areas where new seismic hazards have occurred. In this case, the goal of transferring learning to new areas may not be achieved by considering site effect factors. In future research on high-accuracy building damage assessment modeling in a given region, the site effect factor will be incorporated to improve the accuracy of the model.

4. Conclusions

This study investigates the effective use of historical data and the extrapolation performance of the existing trained models to new or unknown data through transfer learning approaches, attempting to address the limitations of deep learning models in practical seismic building damage assessment applications. For this purpose, a data transfer algorithm is proposed and applied to two historical earthquake building damage datasets (the Ludian earthquake and the Yushu earthquake). The main findings of this study are as follows: (1) The model built based on historical data can be effectively transferred to the new task, and could assist in addressing the situation where there is not yet data available for the new task. (2) The data transfer algorithm proposed in this study can filter data samples suitable for the target task from historical data, and using these historical data could significantly improve the reliability of the classification model. (3) The data transfer algorithm can effectively improve the classification accuracy of the model in the case of limited data, especially in the case of limited new data in the affected area during the disaster emergency response phase, and the data transfer model would be helpful for improving the efficiency and accuracy of rapid damage assessment.

Author Contributions: Conceptualization, Q.L., T.C. and Y.W.; methodology, T.C., Q.L. and L.W.; data curation, H.Y. and T.C.; writing—original draft preparation, Q.L.; writing—review and editing, Q.L., T.C., L.W., S.K.M., H.Y. and Y.W.; supervision, Y.W.; funding acquisition, Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Second Tibetan Plateau Scientific Expedition and Research Program, grant number 2019QZKK0906, 2019QZKK0606 and the Startup Foundation for Introducing Talent of NUIST.

Data Availability Statement: The dataset and the code are freely available online and can be found at https://github.com/city292/build_assessment (accessed on 26 June 2021).

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

References

1. Kassem, M.M.; Nazri, F.M.; Farsangi, E.N. Development of seismic vulnerability index methodology for reinforced concrete buildings based on nonlinear parametric analyses. *MethodsX* **2019**, *6*, 199–211. [CrossRef]
2. Kassem, M.M.; Nazri, F.M.; Farsangi, E.N. The efficiency of an improved seismic vulnerability index under strong ground motions. *Structures* **2021**, *23*, 366–382. [CrossRef]
3. Kassem, M.M.; Nazri, F.M.; Farsangi, E.N.; Tan, C.G. Comparative seismic RISK assessment of existing RC buildings using seismic vulnerability index approach. *Structures* **2021**, *32*, 889–913. [CrossRef]
4. Kassem, M.M.; Nazri, F.M.; Farsangi, E.N.; Ozturk, B. Improved Vulnerability Index Methodology to Quantify Seismic Risk and Loss Assessment in Reinforced Concrete Buildings. *J. Earthq. Eng.* **2021**, 1–36. [CrossRef]
5. Lin, Q.; Lima, P.; Steger, S.; Glade, T.; Jiang, T.; Zhang, J.; Liu, T.; Wang, Y. National-scale data-driven rainfall induced landslide susceptibility mapping for China by accounting for incomplete landslide data. *Geosci. Front.* **2021**, *12*, 101248. [CrossRef]
6. Li, Y.; Wang, Y.; Zhang, Y.; Zhou, X.; Sun, H. Impact of economic development levels on the mortality rates of Asian earthquakes. *Int. J. Disaster Risk Reduct.* **2021**, *62*, 102409. [CrossRef]
7. Wu, J.; He, X.; Li, Y.; Shi, P.; Ye, T.; Li, N. How earthquake-induced direct economic losses change with earthquake magnitude, asset value, residential building structural type and physical environment: An elasticity perspective. *J. Environ. Manag.* **2019**, *231*, 321–328. [CrossRef]

8. Nex, F.; Duarte, D.; Tonolo, F.G.; Kerle, N. Structural building damage detection with deep learning: Assessment of a state-of-the-art cnn in operational conditions. *Remote Sens.* **2019**, *11*, 2765. [CrossRef]
9. Kalantar, B.; Ueda, N.; Al-Najjar, H.A.; Halin, A.A. Assessment of convolutional neural network architectures for earthquake-induced building damage detection based on pre-and post-event orthophoto images. *Remote Sens.* **2020**, *12*, 3529. [CrossRef]
10. Khodaverdizahraee, N.; Rastiveis, H.; Jouybari, A. Segment-by-segment comparison technique for earthquake-induced building damage map generation using satellite imagery. International journal of disaster risk reduction. *Int. J. Disaster Risk Reduct.* **2020**, *46*, 101505. [CrossRef]
11. Burke, M.; Driscoll, A.; Lobell, D.B.; Ermon, S. Using satellite imagery to understand and promote sustainable development. *Science* **2021**, *371*, eabe8628. [CrossRef] [PubMed]
12. Mavroulis, S.; Andreadakis, E.; Spyrou, N.I.; Antoniou, V.; Skourtsos, E.; Papadimitriou, P.; Kasssaras, I.; Kaviris, G.; Tselentis, G.A.; Voulgaris, N.; et al. UAV and GIS based rapid earthquake-induced building damage assessment and methodology for EMS-98 isoseismal map drawing: The June 12, 2017 Mw 6.3 Lesvos (Northeastern Aegean, Greece) earthquake. *Int. J. Disaster Risk Reduct.* **2019**, *37*, 101169. [CrossRef]
13. 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]
14. 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]
15. Saito, K.; Spence, R.J.; Going, C.; Markus, M. Using high-resolution satellite images for post-earthquake building damage assessment: A study following the 26 January 2001 Gujarat earthquake. *Earthq. Spectra* **2004**, *20*, 145–169. [CrossRef]
16. Pesaresi, M.; Gerhardinger, A.; Haag, F. Rapid damage assessment of built-up structures using VHR satellite data in tsunami-affected areas. *Int. J. Remote Sens.* **2008**, *28*, 3013–3036. [CrossRef]
17. Chini, M.; Cinti, F.R.; Stramondo, S. Co-seismic surface effects from very high resolution panchromatic images: The case of the 2005 Kashmir (Pakistan) earthquake. *Nat. Hazards Earth Syst. Sci.* **2011**, *11*, 931–943. [CrossRef]
18. Shi, W.; Zhang, M.; Zhang, R.; Chen, S.; Zhan, Z. Change detection based on artificial intelligence: State-of-the-art and challenges. *Remote Sens.* **2020**, *12*, 1688. [CrossRef]
19. Mangalathu, S.; Sun, H.; Nweke, C.C.; Yi, Z.; Burton, H.V. Classifying earthquake damage to buildings using machine learning. *Earthq. Spectra* **2020**, *36*, 183–208. [CrossRef]
20. Xie, Y.; Ebad Sichani, M.; Padgett, J.E.; DesRoches, R. The promise of implementing machine learning in earthquake engineering: A state-of-the-art review. *Earthq. Spectra* **2020**, *36*, 1769–1801. [CrossRef]
21. Abdi, G.; Jabari, S. A Multi-Feature Fusion Using Deep Transfer Learning for Earthquake Building Damage Detection. *Can. J. Remote Sens.* **2021**, *47*, 337–352. [CrossRef]
22. Xiong, C.; Li, Q.; Lu, X. Automated regional seismic damage assessment of buildings using an unmanned aerial vehicle and a convolutional neural network. *Autom. Constr.* **2020**, *109*, 102994. [CrossRef]
23. Lu, X.; Xu, Y.; Tian, Y.; Cetiner, B.; Taciroglu, E. A deep learning approach to rapid regional post-event seismic damage assessment using time-frequency distributions of ground motions. *Earthq. Eng. Struct. Dyn.* **2021**, *50*, 1612–1627. [CrossRef]
24. Yu, Y.; Wang, C.; Gu, X.; Li, J. A novel deep learning-based method for damage identification of smart building structures. *Struct. Health Monit.* **2019**, *18*, 143–163. [CrossRef]
25. Khan, A.; Sohail, A.; Zahoora, U.; Qureshi, A.S. A survey of the recent architectures of deep convolutional neural networks. *Artif. Intell. Rev.* **2020**, *53*, 5455–5516. [CrossRef]
26. Yang, J.; Xiong, W.; Li, S.; Xu, C. Learning structured and non-redundant representations with deep neural networks. *Pattern Recognit.* **2019**, *86*, 224–235. [CrossRef]
27. 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]
28. Anniballe, R.; Noto, F.; Scalia, T.; Bignami, C.; Stramondo, S.; Chini, M.; Pierdicca, N. Earthquake damage mapping: An overall assessment of ground surveys and VHR image change detection after L'Aquila 2009 earthquake. *Remote Sens. Environ.* **2018**, *210*, 166–178. [CrossRef]
29. Duarte, D.; Nex, F.; Kerle, N.; Vosselman, G. Multi-resolution feature fusion for image classification of building damages with convolutional neural networks. *Remote Sens.* **2018**, *10*, 1636. [CrossRef]
30. Huang, J.; Zhang, X.; Xin, Q.; Sun, Y.; Zhang, P. Automatic building extraction from high-resolution aerial images and LiDAR data using gated residual refinement network. *ISPRS J. Photogramm.* **2019**, *151*, 91–105. [CrossRef]
31. Ji, S.; Wei, S.; Lu, M. A scale robust convolutional neural network for automatic building extraction from aerial and satellite imagery. *Int. J. Remote Sens.* **2019**, *40*, 3308–3322. [CrossRef]
32. Vetrivel, A.; Kerle, N.; Gerke, M.; Nex, F.; Vosselman, G. Towards automated satellite image segmentation and classification for assessing disaster damage using data-specific features with incremental learning. In Proceedings of the GEOBIA 2016: Solutions and Synergies, Enschede, The Netherlands, 14–16 September 2016.
33. Ci, T.; Liu, Z.; Wang, Y. Assessment of the degree of building damage caused by disaster using convolutional neural networks in combination with ordinal regression. *Remote Sens.* **2019**, *11*, 2858. [CrossRef]

34. Yu, S.; Liu, L.; Wang, Z.; Dai, G.; Xie, Y. Transferring deep neural networks for the differentiation of mammographic breast lesions. *Sci. China Technol. Sci.* **2019**, *62*, 441–447. [CrossRef]
35. Qin, S.; Guo, X.; Sun, J.; Qiao, S.; Zhang, L.; Yao, J.; Cheng, Q.; Zhang, Y. Landslide Detection from Open Satellite Imagery Using Distant Domain Transfer Learning. *Remote Sens.* **2021**, *13*, 3383. [CrossRef]
36. Gao, Y.; Mosalam, K.M. Deep transfer learning for image-based structural damage recognition. *Comput.-Aided Civ. Infrastruct. Eng.* **2018**, *33*, 748–768. [CrossRef]
37. Xu, S.; Noh, H.Y. PhyMDAN: Physics-informed knowledge transfer between buildings for seismic damage diagnosis through adversarial learning. *Mech. Syst Signal. Process.* **2021**, *151*, 107374. [CrossRef]
38. Xie, S.; Duan, J.; Liu, S.; Dai, Q.; Liu, W.; Ma, Y.; Guo, R.; Ma, C. Crowdsourcing rapid assessment of collapsed buildings early after the earthquake based on aerial remote sensing image: A case study of yushu earthquake. *Remote Sens.* **2016**, *8*, 759. [CrossRef]
39. Fan, Y.; Wen, Q.; Wang, W.; Wang, P.; Li, L.; Zhang, P. Quantifying disaster physical damage using remote sensing data—A technical work flow and case study of the 2014 Ludian earthquake in China. *Int. J. Disaster Risk Sci.* **2017**, *8*, 471–488. [CrossRef]
40. Torrey, L.; Shavlik, J. Transfer learning. In *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*; Olivas, E.S., Ed.; IGI Global: Hershey, PA, USA, 2009; pp. 242–264.
41. Koh, P.W.; Liang, P. Understanding black-box predictions via influence functions. In Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; JMLR: Sydney, Australia, 2017; Volume 70, pp. 1885–1894.
42. Zheng, L.; Liu, G.; Yan, C.; Jiang, C.; Zhou, M.; Li, M. Improved TrAdaBoost and its application to transaction fraud detection. *IEEE Trans. Comput. Soc. Syst.* **2020**, *7*, 1304–1316. [CrossRef]
43. Saha, B.; Gupta, S.; Phung, D.; Venkatesh, S. Multiple task transfer learning with small sample sizes. *Knowl. Inf. Syst.* **2016**, *46*, 315–342. [CrossRef]
44. Cohen, J. A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* **1960**, *20*, 37–46. [CrossRef]
45. Li, P.; Xu, H.; Guo, J. Urban building damage detection from very high resolution imagery using OCSVM and spatial features. *Int. J. Remote Sens.* **2010**, *31*, 3393–3409. [CrossRef]
46. 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]
47. Pitilakis, K. Site Effects. In *Recent Advances in Earthquake Geotechnical Engineering and Microzonation. Geotechnical, Geological, and Earthquake Engineering*; Springer: Dordrecht, The Netherlands, 2004; Volume 1, pp. 139–197. [CrossRef]
48. Stamatopoulos, C.A.; Bassanou, M.; Brennan, A.J.; Madabhushi, G. Mitigation of the seismic motion near the edge of cliff-type topographies. *Soil Dyn. Earthq. Eng.* **2007**, *27*, 1082–1100. [CrossRef]
49. Sextos, A.; De Risi, R.; Pagliaroli, A.; Foti, S.; Passeri, F.; Ausilio, E.; Cairo, R.; Capatti, M.C.; Chiarabando, F.; Chiaradonna, A.; et al. Local site effects and incremental damage of buildings during the 2016 Central Italy earthquake sequence. *Earthq. Spectra* **2018**, *34*, 1639–1669. [CrossRef]
50. Mayoral, J.M.; De la Rosa, D.; Tepalcapa, S. Topographic effects during the September 19, 2017 Mexico city earthquake. *Soil Dyn. Earthq. Eng.* **2019**, *125*, 105732. [CrossRef]