

# QuickQuakeBuildings: Post-Earthquake SAR-Optical Dataset for Quick Damaged-Building Detection

Yao Sun<sup>1</sup>, Yi Wang<sup>1</sup>, *Graduate Student Member, IEEE*, and Michael Eineder<sup>2</sup>, *Fellow, IEEE*

**Abstract**—Quick and automated earthquake-damaged building detection from postevent satellite imagery is crucial, yet it is challenging due to the scarcity of training data required for developing robust algorithms. This letter presents the first dataset dedicated to detecting earthquake-damaged buildings from postevent very high-resolution (VHR) synthetic aperture radar (SAR) and optical imagery. Utilizing open satellite imagery and annotations acquired for the 2023 Turkey–Syria earthquakes, we deliver a dataset of co-registered building footprints and satellite images of both SAR and optical data, encompassing more than 4000 buildings. The task of the damaged building detection is formulated as a binary image classification problem, that can also be treated as an anomaly detection problem due to extreme class imbalance. We provide baseline methods and results to serve as references for comparison. Researchers can utilize this dataset to expedite algorithm development, facilitating the rapid detection of damaged buildings in response to future events. The dataset and codes together with explanations and visualization are made available at <https://github.com/ya0-sun/PostEQ-SARopt-BuildingDamage>.

**Index Terms**—Building damage, deep learning, earthquake, geographic information system (GIS), large scale, openstreetmap (OSM), remote sensing, synthetic aperture radar (SAR).

## I. INTRODUCTION

EARTHQUAKES can result in substantial structural and infrastructural damage, often with significant socio-economic consequences. After an event, fast and accurate detection of earthquake-damaged buildings in remote sensing imagery is of great importance. Remote sensing technologies can effectively improve the efficiency of disaster management and have been employed to estimate the extent of earthquake damage to buildings [1], [2], [3]. Very high-resolution (VHR) optical images are easier to interpret, making them a preferred choice for many studies. However, acquisition of cloud-free optical images depends on weather conditions and often needs to wait. In contrast, synthetic aperture radar (SAR) imagery is particularly suitable for rapid disaster response scenarios

as it can be acquired regardless of cloud coverage and sun illumination conditions. Moreover, the enhanced resolution of contemporary SAR satellite images enables the extraction of information at the individual building level [4], [5], [6], comparable to VHR optical data.

Over the past years, many researchers have developed algorithms utilizing SAR data to detect earthquake-damaged buildings. Most works detect changes using both pre- and postevent SAR images to acquire building damage information [7], [8]; however, preevent high-resolution SAR imagery is generally unavailable in most locations. A few works simulate an SAR image using building shapes extracted from a preevent optical image and acquisition parameters of a postevent SAR image and detect earthquake-damaged buildings by comparing the simulated and real SAR images [9]. Utilizing only a single postevent SAR image, some researchers detect damaged buildings by analyzing signatures of destroyed buildings in high-resolution SAR data [10], [11], [12]. However, the study areas often comprise a limited number of isolated buildings that do not depict the typical conditions in densely populated urban areas, where partially occluded buildings and geometric distortions in SAR images, i.e., foreshortening, layover, and shadowing, commonly exist. Therefore, the question of whether a single VHR SAR image acquired after an event can effectively identify damaged buildings remains to be addressed. In this regard, benchmark datasets play a pivotal role in the development and comparative assessment of diverse methodologies aimed at addressing the following questions: To what extent can a single VHR SAR image, acquired postevent, allow to identify damaged buildings and with which accuracy? Additionally, how do the outcomes derived from a single postevent SAR image compare with those of an optical image?

From a practical standpoint, there are several challenges in creating such a dataset: 1) limited availability of VHR SAR images in disaster-affected areas; 2) absence of labels for damaged buildings; and 3) lack of accurate terrain models, without which aligning the two is a complex task. Currently, such a dataset does not exist in the remote sensing field.

This work presents the first dataset for detecting earthquake-damaged buildings in postevent VHR SAR and optical satellite imagery. We integrate publicly accessible satellite imagery and annotations obtained following the 2023 Turkey–Syria earthquakes and construct a dataset comprising over 4000 buildings, each with satellite image patches of both postevent SAR and optical data and its footprint coregistered

Manuscript received 24 March 2024; accepted 13 May 2024. Date of publication 30 May 2024; date of current version 11 June 2024. The work of Yi Wang was supported by the Helmholtz Association through the Framework of Helmholtz AI. (*Corresponding author: Yao Sun.*)

Yao Sun is with Data Science in Earth Observation, Technical University of Munich, 80333 Munich, Germany (e-mail: yao.sun@tum.de).

Yi Wang is with Data Science in Earth Observation, Technical University of Munich, 80333 Munich, Germany, and also with the Remote Sensing Technology Institute, German Aerospace Center, 82234 Wessling, Germany (e-mail: yi4.wang@tum.de).

Michael Eineder is with the Remote Sensing Technology Institute, German Aerospace Center, 82234 Wessling, Germany (e-mail: michael.eineder@dlr.de).

Digital Object Identifier 10.1109/LGRS.2024.3406966

with the corresponding image patches. We formulate the problem of damaged building detection as an image classification task and benchmark a set of popular machine learning and deep learning methods as a baseline reference.

In the remaining part of the letter, Section II introduces the dataset generation approaches, Section III presents the baseline methods and results, and Section IV concludes the letter.

## II. DATASET GENERATION

This work aims at addressing building-level damages. The tasks of identifying pixel-level damages and classifying the types of damages are beyond the scope of this dataset.

Considering the increasing availability of building footprint information across various geographic locations, we integrate preevent building footprints with postevent satellite images and assess whether buildings at those locations are damaged. The problem of detecting damaged buildings in postevent imagery is therefore formulated as an image classification task with two classes: damaged and intact buildings. We generate and deliver a dataset comprising postevent VHR SAR and optical image patches for each building, building footprint masks corresponding to the image patches, and labels indicating whether each building is damaged or intact. The image patches and masks serve as inputs for the algorithm, while the labels represent the ground truth.

### A. Study Area and Data Sources

The study area is chosen in the city of Islahiye, located in southeastern Turkey near the northwestern border of Syria. On February 6, 2023, a magnitude 7.8 earthquake struck Kahramanmaras, Turkey, followed by a 7.5 magnitude aftershock nine hours later. The earthquakes inflicted widespread destruction, leading to significant damage to buildings, injuries, and loss of life. Islahiye was one of the most affected areas.

After the earthquakes, a set of high-resolution satellite data were released under CC BY 4.0 license<sup>1</sup> to support rescue operations by commercial remote sensing companies, such as Maxar and Planet Lab for optical data, and Capella Space in the SAR domain. During the humanitarian relief efforts, communities across the globe, such as OSM and UN mappers, organized labeling events and identified and validated a significant number of damaged buildings.

We utilize the spotlight SAR image from Capella Space, acquired on February 9, 2023. The SAR image is of type Geocoded Terrain Corrected (GEO),<sup>2</sup> with a pixel spacing of 0.35 m in both the azimuth and the range direction. The incidence angle of this SAR image is 43.1°. Fig. 1 shows the image coverage and zoomed-in views of selected areas. The optical image covering the same area was obtained from Maxar analysis-ready data (ARD) under Maxar's open data program. The image was acquired on February 7, 2023 by WorldView-3, with a ground sampling distance of 0.31 m and an incidence angle of 83.1°. In addition, we obtained



Fig. 1. SAR image coverage and zoomed-in views of three areas in the colored boxes in the SAR image, respectively.

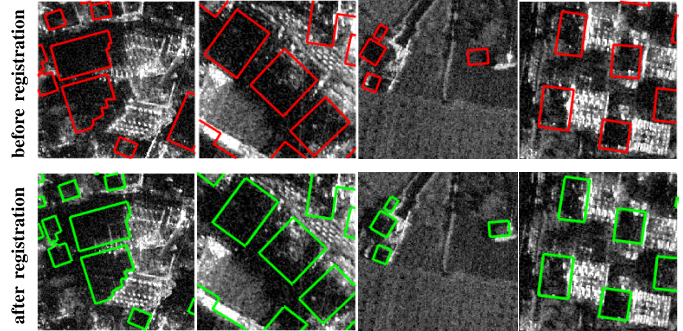


Fig. 2. Examples of four study areas in the SAR image. Building footprint polygons before and after registration are plotted in red and green, respectively.

postevent building footprints and labels of destroyed buildings in the study area from Humanitarian OpenStreetMap Team<sup>3,4</sup>. Since we consider both SAR and optical images, all used data are chosen or projected to the Universal Transverse Mercator (UTM) coordinate system so that they can be processed uniformly. The SAR image is logarithmically scaled in dB for further processing.

### B. Coregistration of Building Footprints and Satellite Imagery

Building-level analysis requires accurate registration of 2-D building footprints with satellite images. The ARD optical image aligns well with building footprints, requiring no additional registration. For the GEO SAR image, inspection shows that building polygons are not well-matched with the SAR image, as shown in the first row of Fig. 2, and further registration is needed.

In urban areas, the geocoding errors in SAR data are often caused by inaccurate terrain heights, as illustrated in Fig. 3(a). A height error  $\delta H$  causes an error of  $\delta L$  in the slant range and a shift of  $\delta G$  on the ground. For the used SAR image, the incidence angle is 43.1°; thus, a height error of 10 m results in an error of 10.69 m on the ground, causing errors of  $\delta x$  and  $\delta y$  in the geocoded image related to the flight direction, as shown in Fig. 3(b). The height error  $\delta H$  is usually inconstant over the observed area by the SAR sensor; hence, so are the geocoding errors.

To improve the alignment of building polygons and the SAR image, we apply the algorithm developed in [13], which relies on the corresponding building features representing the bottom of sensor-visible walls in both the two data, i.e., double bounce lines in the SAR image and near-range boundaries of 2-D

<sup>1</sup><https://creativecommons.org/licenses/by/4.0/>

<sup>2</sup><https://support.capellaspace.com/hc/en-us/articles/360039702691-SAR-Data-Formats>

<sup>3</sup>[https://data.humdata.org/dataset/hotosm\\_tur\\_buildings](https://data.humdata.org/dataset/hotosm_tur_buildings)

<sup>4</sup>[https://data.humdata.org/dataset/hotosm\\_tur\\_destroyed\\_buildings](https://data.humdata.org/dataset/hotosm_tur_destroyed_buildings)

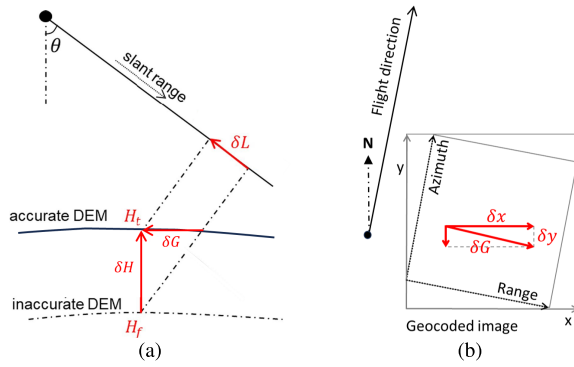


Fig. 3. Geocoding error from inaccurate height. (a)  $H_t$  and  $H_f$  are the accurate height and inaccurate height of a point, and  $\theta$  is the incidence angle. The height error  $\delta H$  results in an error of  $\delta L = \delta H \cos \theta$  in the slant range and  $\delta G = \delta H \cot \theta$  on the ground. (b) Geocoded image,  $\delta G$  is decomposed to  $\delta x$  and  $\delta y$  in the image coordinate system.

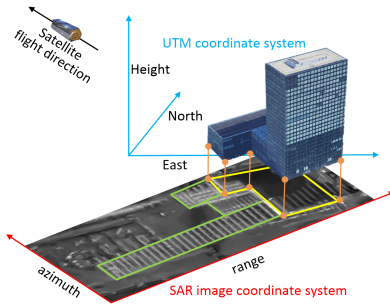


Fig. 4. Building signature in SAR imagery and building's geometric correspondence between SAR and building polygon: the near-range side of the building footprint corresponds to the double bounce line in the SAR image, which is the far-range side of the facade signatures [13].

building polygons, as illustrated in Fig. 4. As the majority of buildings remain upright, with expected double bounce line signatures on the SAR image, the algorithm is applicable. Interested readers are referred to [13] for more details. Next, we briefly explain the main steps of the algorithms.

1) *Extracting Corresponding Features*: In the SAR image, double bounce lines correspond to the far-range side of the bright building signature. The SAR image is first segmented using Potts model [14]. Then, an intensity threshold and an area threshold are applied to select building wall segments. Subsequently, the boundary of wall segments is extracted, and the visibility check is performed to extract the double bounce lines caused by the wall segment boundary.

In the building polygons, the segments in the near-range side of building polygons are extracted from each footprint, which in 3-D represent the bottom of illuminated or partially illuminated walls.

2) *Registering Corresponding Features*: two point sets, GIS points (from building polygons) and SAR points, are then sampled in the extracted features from both data, and the registration problem is reduced to determine the correspondence and the underlying spatial transformation between two point sets.

Feature registration consists of three progressive steps: global registration, subarea registration, and polygon registration, and the rigid registration in each step is solved with the iterative closest point (ICP) algorithm [15]. Global registration

uses rough height values for an initial alignment of the two data, ensuring that the residual shift falls within a manageable range. Then, a set of grids, i.e., subareas, is evenly distributed over the whole region. The distance between one GIS point and its closest SAR point is calculated for all points. If the distribution of all distances within one subarea shows a clear center, the  $\delta H$  in each grid is considered to be constant, and subarea registration is conducted. When the distribution of all distances does not show a clear center, the constant  $\delta H$  assumption does not hold, and the registration proceeds to the polygon level, i.e., finding a rigid transformation for each polygon.

Since the SAR image, i.e., the GEO product, has been terrain corrected, we neglect global registration and perform subarea and polygon registration. Afterward, manual validation and editing with expert knowledge are conducted to ensure precise registration, relying on identifying double bounce lines and the assumption of no abrupt terrain changes for editing buildings with unclear double bounce lines [13]. For individual buildings, we estimate a maximum registration error of 5–6 pixels (approximately 2 m) that is sufficient for locating individual buildings. Fig. 2 shows building polygons on the SAR image in exemplary areas before and after registration. As can be seen, the registration procedure effectively aligns the two datasets. Note that buildings in the mountainous area on the west side of the city (e.g., in or near the red box in Fig. 1) are excluded due to unclear signals in the SAR image, posing challenges for registration verification.

### C. Patch Generation

For each building, we crop the SAR image and the optical image based on the area of the building, considering including the target building area, i.e., footprint, wall, and roof, as well as the possible ruins around the target building and excluding surrounding buildings.

For side-looking SAR images, layover areas of buildings extend from building footprints toward the near-range direction. Therefore, when cropping SAR image patches, a buffer size of 10 pixels, i.e., around 3.5 m, is counted for far-range sides, and for near-range sides, an additional buffer size is counted to include layover areas of the target building in the image patch. Specifically, the layover length corresponding to 50 m building height in the ground range direction is decomposed to the image  $x$ - and  $y$ -directions and added to the buffer size. The optical image has a small off-nadir angle of  $6.3^\circ$ , resulting in a minor offset between the roof outline and the corresponding 2-D building footprint when the image is not perfectly orthorectified. To ensure that the image patches include the entire building roof, we apply a 16-pixel buffer to compensate for the offset, which is approximately 5 m. This buffer is slightly larger than the one used when cropping the SAR image from the near-range side, and it is used to crop the optical image around the bounding box of the building footprint polygons.

In addition, for each building, its footprint mask is generated corresponding to the SAR patch and the optical patch, respectively. For side-looking SAR data, it is necessary to include the footprint mask to help locate the target building, as the

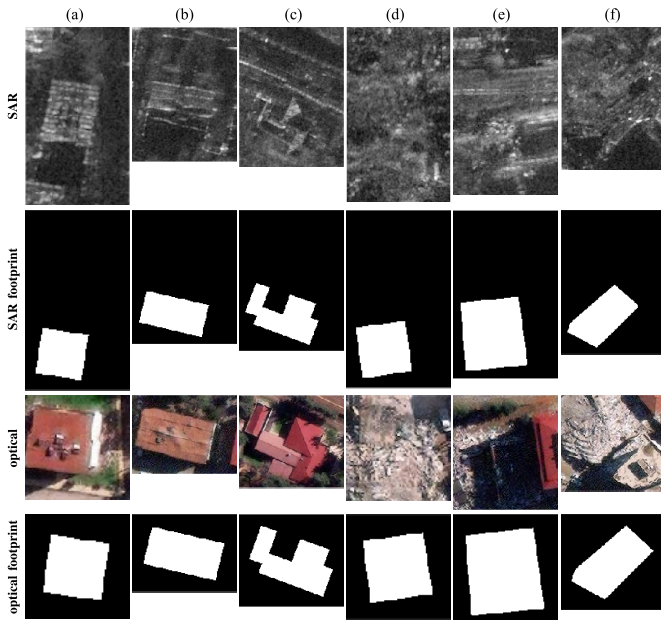


Fig. 5. Examples of the dataset (a), (b), (c) are *intact* buildings, and (d), (e), (f) are *damaged* buildings.

SAR patch may include signals of surrounding buildings. For nadir-looking optical data, footprint masks are included for a fair comparison with the SAR data.

Among the 4029 buildings, 169 are damaged, and the remaining 3860 buildings are intact. Consequently, the dataset contains 169 damaged buildings and 3860 intact buildings, and each of them has four patches: an SAR image patch, an SAR footprint patch, an optical image patch, and an optical footprint patch. Fig. 5 shows examples of the dataset on six buildings, of which three are intact and three are damaged.

### III. EXPERIMENTAL RESULTS

We benchmark an image classification task with two classes: damaged and intact buildings. Due to the significant class imbalance, it can also be viewed as an anomaly detection task, i.e., detecting damaged buildings within the entire dataset.

#### A. Baseline Approaches

Four models are introduced to establish a benchmark: support vector machine (SVM), random forest (RF), three-layer convolutional neural network (CNN), and ResNet-18 [16].

SVM and RF are selected for their good performance in many applications, including classifying collapsed and standing buildings from postevent SAR imagery as reported in [11]. To extract features as the input, we mask out nonbuilding pixels with footprints, and follow the setup proposed in [11], which employs four first-order statistics, i.e., mean, variance, skewness, and kurtosis, and eight second-order image statistical measures, i.e., mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation. For a detailed explanation, the readers are referred to [11].

A simple three-layer CNN and a ResNet-18 are selected as the deep learning backbones. The simple CNN consists of three convolution-ReLU-maxpool blocks, followed by average pooling, a linear layer with dropout, and a final classification layer. We stack image and building footprint as the input. The

ResNet-18 follows the standard design in [16], with which we benchmark both early and late fusion results of images and building footprints. For late fusion, images and building footprints are encoded by separate encoders, and the feature vectors are concatenated together to a following linear layer with dropout and a final classification layer. Apart from single-modal results, we also conduct a late fusion experiment with both SAR and optical data as a multimodal reference.

#### B. Evaluation Metrics

To evaluate the performance of the baseline methods, we report the precision, recall, and  $F_1$  scores

$$P = \frac{tp}{tp + fp}, \quad R = \frac{tp}{tp + fn}, \quad F_1 = 2 \cdot \frac{P \cdot R}{P + R} \quad (1)$$

where  $P$  and  $R$  denote the precision and recall, and  $tp$ ,  $fp$ ,  $tn$ , and  $fn$  represent true positives, false positives, true negatives, and false negatives for buildings, respectively. The thresholds for positive/negative are determined by best  $F_1$  scores.

In addition, we report the area under the receiver-operator curve (AUROC), a standard metric in anomaly detection tasks that well reflects the model's efficiency in distinguishing between classes. The AUROC score summarizes the ROC curve into a single number that describes the performance of a model for multiple thresholds at the same time.

#### C. Implementation Details

We conduct cross-fold experiments for a robust evaluation on the relatively small dataset. Specifically, we split the dataset into five folds with a balanced number of damaged and intact buildings, and run each experiment five times with four folds training and one fold testing. The mean and standard deviation are calculated and reported for every evaluation metric.

We preprocess the images by removing 2% pixel outliers and normalizing the pixel values to the range [0, 1]. For simple CNN, we randomly initialize the model; for ResNet-18 early fusion, we use pure ImageNet weights; for ResNet-18 late fusion, we use ImageNet weights for optical images and footprints, and SAR-HUB [17] weights for SAR images. We use random resized crop and random horizontal and vertical flip as data augmentations. To deal with the significant class imbalance, we give a bigger weight to damaged buildings and a smaller weight to intact buildings during data sampling. We optimize binary cross entropy loss with AdamW optimizer for 30 epochs. The learning rate follows a cosine-decay schedule starting from 0.0001. Batch size is set to 32.

For the statistic features used by the two machine learning methods, we calculated the features of image patches based on the implementation of or PyFeats.<sup>5</sup> For calculating the evaluation metrics, we use the implementation in scikit-learn.<sup>6</sup>

#### D. Performance Comparison

Table I illustrates variations in performance across different models on the dataset. In general, deep neural networks

<sup>5</sup><https://github.com/giakou4/pyfeats>

<sup>6</sup>[https://scikit-learn.org/stable/modules/generated/sklearn.metrics.roc\\_auc\\_score.html](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.roc_auc_score.html)

TABLE I  
BENCHMARK RESULTS ON THE DATASET: FIVE-FOLD MEAN(STD). THE HIGHEST VALUES OF DIFFERENT METRICS ARE HIGHLIGHTED IN **BOLD** FOR EACH SET OF UTILIZED DATA. EARLY FUSION IS NOTED AS EF, WHILE LATE FUSION IS NOTED AS LF

Image	SAR					optical					SAR+optical
	SVM	RF	CNN (EF)	ResNet18 (EF+ImageNet)	ResNet18 (LF+SAR-HUB)	SVM	RF	CNN (EF)	ResNet18 (EF+ImageNet)	ResNet18 (LF+ImageNet)	ResNet18 (LF+SAR-HUB+ImageNet)
Precision	0.096(0.021)	0.223(0.090)	0.167(0.028)	0.155(0.067)	<b>0.237(0.116)</b>	0.163(0.044)	0.611(0.215)	0.489(0.154)	<b>0.666(0.121)</b>	0.622(0.213)	<b>0.746(0.062)</b>
Recall	<b>0.495(0.202)</b>	0.344(0.208)	0.371(0.073)	0.276(0.080)	0.427(0.104)	0.352(0.148)	0.354(0.098)	0.432(0.078)	0.539(0.147)	<b>0.646(0.143)</b>	<b>0.615(0.117)</b>
$F_1$	0.154(0.037)	0.231(0.071)	0.228(0.032)	0.184(0.050)	<b>0.282(0.077)</b>	0.207(0.057)	0.421(0.106)	0.449(0.096)	0.581(0.100)	<b>0.605(0.120)</b>	<b>0.670(0.072)</b>
AUROC	0.653(0.068)	0.670(0.079)	0.739(0.026)	0.653(0.043)	<b>0.769(0.032)</b>	0.723(0.047)	0.810(0.062)	0.853(0.045)	0.938(0.022)	<b>0.941(0.026)</b>	<b>0.962(0.023)</b>

outperform SVM and RF in both SAR and optical scenarios. For the SAR image, ResNet-18 underperforms simple CNN regarding all four metrics with early fusion and ImageNet weights, indicating the optimization challenge of complex SAR data. This issue is resolved with proper weight initialization from SAR-HUB (around 10% improvement in AUROC and  $F_1$  scores), highlighting the importance of SAR pretrained models such as provided in [17].

For the optical image, ResNet-18 with late fusion stands out with the highest recall,  $F_1$  score, and AUROC. It outperforms other models across these three metrics. ResNet-18 with early fusion gave best precision. The other deep learning model, CNN, demonstrates a balance between precision and recall, resulting in a high  $F_1$  score and an impressive AUROC of 0.853. RF attains high precision at 0.611 but a relatively lower recall rate. In contrast, SVM yields less favorable outcomes compared to other models.

Comparing different modalities, we can see that SAR images are more challenging for all models and generally show lower performance. However, well-designed SAR models, e.g., ResNet-18 with SAR-HUB weights, outperform inferior optical models such as SVM.

In addition, the fusion of SAR and optical imagery provides further improvement compared to each single modality, confirming the complementary information across different modalities.

#### IV. CONCLUSION

Detecting earthquake-damaged buildings in postevent satellite imagery is essential yet challenging. This study introduces a dataset designed to address the issue and to foster the development of robust algorithms. The dataset combines postevent SAR and optical satellite images with labels of damaged and intact buildings, and the problem is formulated as an image classification task. We provide a benchmark on both modalities with different baseline methods and a baseline of fusing optical and SAR data. Results show that detecting damage from postevent SAR images is valuable and possible but more challenging than optical images. Such findings call for further research on improved methods, in particular on SAR images. In addition, the performance gain through simple SAR-optical fusion verifies the potential in using multimodal data when they can be acquired following an event.

Constrained by limited data quantity and data imbalance, this dataset serves as a starting point. The dataset will undergo expansion and updates as new data emerges in the future. We hope that the research community will engage in further algorithm development for post-earthquake-damaged building

assessment in SAR images and, where feasible, share their data, thus expediting the identification of postdisaster damaged structures.

#### REFERENCES

- [1] F. Dell'Acqua and P. Gamba, "Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives," *Proc. IEEE*, vol. 100, no. 10, pp. 2876–2890, Oct. 2012.
- [2] L. Dong and J. Shan, "A comprehensive review of earthquake-induced building damage detection with remote sensing techniques," *ISPRS J. Photogramm. Remote Sens.*, vol. 84, pp. 85–99, Oct. 2013.
- [3] D. Contreras, S. Wilkinson, and P. James, "Earthquake reconnaissance data sources, a literature review," *Earth*, vol. 2, no. 4, pp. 1006–1037, Nov. 2021.
- [4] Y. Sun, Y. Hua, L. Mou, and X. X. Zhu, "CG-Net: Conditional GIS-aware network for individual building segmentation in VHR SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5201215.
- [5] J. Chen, X. Qiu, C. Ding, and Y. Wu, "CVCMMFF Net: Complex-valued convolutional and multifeature fusion network for building semantic segmentation of InSAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2021, Art. no. 5205714.
- [6] Y. Sun, L. Mou, Y. Wang, S. Montazeri, and X. X. Zhu, "Large-scale building height retrieval from single SAR imagery based on bounding box regression networks," *ISPRS J. Photogramm. Remote Sens.*, vol. 184, pp. 79–95, Feb. 2022.
- [7] H. Miura, S. Midorikawa, and M. Matsuoka, "Building damage assessment using high-resolution satellite SAR images of the 2010 Haiti earthquake," *Earthq. Spectra*, vol. 32, no. 1, pp. 591–610, Feb. 2016.
- [8] R. Guida, A. Iodice, and D. Riccio, "Monitoring of collapsed built-up areas with high resolution SAR images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2010, pp. 2422–2425.
- [9] D. Brunner, G. Lemoine, and L. Bruzzone, "Earthquake damage assessment of buildings using VHR optical and SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2403–2420, May 2010.
- [10] T. Balz and M. Liao, "Building-damage detection using post-seismic high-resolution SAR satellite data," *Int. J. Remote Sens.*, vol. 31, no. 13, pp. 3369–3391, Jul. 2010.
- [11] L. Gong, C. Wang, F. Wu, J. Zhang, H. Zhang, and Q. Li, "Earthquake-induced building damage detection with post-event sub-meter VHR TerraSAR-X staring spotlight imagery," *Remote Sens.*, vol. 8, no. 11, p. 887, Oct. 2016.
- [12] S. Kuny, K. Schulz, and H. Hammer, "Signature analysis of destroyed buildings in simulated high resolution SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2013, pp. 903–906.
- [13] Y. Sun, S. Montazeri, Y. Wang, and X. X. Zhu, "Automatic registration of a single SAR image and GIS building footprints in a large-scale urban area," *ISPRS J. Photogramm. Remote Sens.*, vol. 170, pp. 1–14, Dec. 2020.
- [14] M. Storath, A. Weinmann, J. Frikel, and M. Unser, "Joint image reconstruction and segmentation using the Potts model," *Inverse Problems*, vol. 31, no. 2, Feb. 2015, Art. no. 025003.
- [15] Y. Chen and G. Medioni, "Object modelling by registration of multiple range images," *Image Vis. Comput.*, vol. 10, no. 3, pp. 145–155, Apr. 1992.
- [16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [17] H. Yang, X. Kang, L. Liu, Y. Liu, and Z. Huang, "SAR-HUB: Pre-training, fine-tuning, and explaining," *Remote Sens.*, vol. 15, no. 23, p. 5534, Nov. 2023.