Rapid Fine-Grained Damage Assessment of Buildings on a Large Scale: A Case Study of the February 2023 Earthquake in Turkey

Zhonghua Hong ⓘ, *Member, IEEE*, Hongyang Zhang ⓘ, Xiaohua Tong ⓘ, *Senior Member, IEEE*,
Shijie Liu ⓘ, *Member, IEEE*, Ruyan Zhou ⓘ, Haiyan Pan ⓘ, Yun Zhang ⓘ, Yanling Han ⓘ, Jing Wang ⓘ,
and Shuhu Yang ⓘ

Abstract—High-resolution stereo satellite images (HRSSIs) have the potential to provide the accurate height and volume information, playing a crucial role in assessing building collapses during various natural disasters. However, the time-consuming process of three-dimensional (3-D) reconstruction, inadequate vertical accuracy of digital surface model (DSM), and concentrated clustering of buildings pose challenges for collapse assessment focused on buildings. Therefore, we present an improved approach for rapid fine-grained assessment of building collapses. First, the accurate and consistent positioning parameters for HRSSIs are obtained through the combined block adjustment using laser altimetry points, ensuring the generation of DSMs with vertical accuracy exceeding 2 m. Next, a set of rapid 3-D reconstruction techniques is introduced, achieving a significant eightfold improvement in generating DSMs. Subsequently, we deploy an automated workflow for batch processing and registration of open-source building footprints, enabling the accurate extraction of building height changes from dual-time DSMs. Finally, based on the building change image, a large-scale GIS image of building floor-level collapses is generated using connected component detection and threshold classification strategies. These findings have far-reaching implications for postdisaster emergency response, damage assessment, and expeditious reconstruction efforts. In our study, we processed an 800 km² area in Kahramanmaras Province, Turkey, generating dual-time DSMs within 1 h. This enabled the assessment of floor-level collapses for a total of 48 092 buildings within the area. Validation was conducted on 361 houses in the city center, utilizing Google Street view images as ground truth. Remarkably, our approach achieved a high accuracy rate of 93.27% in floor-level assessment.

Index Terms—Building damage assessment, change detection, Gaofen-7 (GF-7) high-resolution satellite stereo images, three-dimensional (3-D) reconstruction.

Zhonghua Hong, Hongyang Zhang, Ruyan Zhou, Haiyan Pan, Yun Zhang, Yanling Han, Jing Wang, and Shuhu Yang are with the College of Information Technology, Shanghai Ocean University, Shanghai 201306, China (e-mail: zhhong@shou.edu.cn; m210901467@st.shou.edu.cn; ryzhou@shou.edu.cn; hy-pan@shou.edu.cn; y-zhang@shou.edu.cn; ylhan@shou.edu.cn; wangjing@shou.edu.cn; shyang@shou.edu.cn).
Xiaohua Tong and Shijie Liu are with the College of Surveying and Geo-Informatics, Tongji University, Shanghai 200092, China (e-mail: xhtong@tongji.edu.cn; liusjtj@tongji.edu.cn).
Digital Object Identifier 10.1109/JSTARS.2024.3362809

I. INTRODUCTION

THE seismic event that transpired on February 6, 2023 resulted in extensive devastation spanning across Turkey, Syria, and adjoining nations [1]. The toll inflicted upon architectural structures is profound, as over a million constructions suffered varying degrees of impairment, with an excess of 50 000 succumbing to complete collapse [2]. In light of the calamitous aftermath, these dilapidated edifices constitute the primary catalyst for both human casualties and substantial financial burden [3]. Consequently, the expeditious and precise evaluation of inflicted damage assumes a paramount role in the subsequent endeavors of postdisaster search and rescue operations, comprehensive documentation, and seamless reconstruction efforts [4].

High-resolution satellites offer significant benefits, such as expansive coverage and brief revisit cycles [5], [6]. The increasing availability of very high resolution (VHR) satellite images coupled with reduced acquisition complexity facilitates rapid detection and evaluation of building damage resulting from earthquakes [7], [8], [9]. One of the most important methods for building damage detection involves the use of spectral images acquired pre- and postearthquake to extract relevant information [10], [11]. Earthquakes can cause significant alterations in both buildings and their immediate environs, leading to noticeable variations in morphology, spectrum, and contour features as observed in VHR remote sensing images [12]. In recent times, deep learning techniques, particularly the superior performance of U-net and full convolutional network in image recognition and segmentation tasks, have contributed to the detection and classification of building damage with commendable results [13], [14], [15]. Nonetheless, precise three-dimensional (3-D) information, such as the height and volume of building collapse, is challenging to obtain from spectral data alone [16], [17]. By analyzing alterations in building elevation and volume in 3-D products pre- and postearthquake, a more accurate assessment of building damage can be achieved [18], [19], [20], [21]. Moreover, this approach can enable swift and effective postdisaster rescue operations based on casualty assessments and corresponding rescue operations difficulty [22], [23].

Progress has been made in the methods of obtaining damage assessment results by detecting the height variation of buildings

before and after an earthquake. Turker utilized aerial imagery to generate dual-phase DSMs with a resolution of 0.4 m for change detection of buildings, enabling damage assessment [24]. Tong et al. [19] employed dual-phase IKONOS stereo images to implement two different methods for 3-D damage detection of buildings: one based on building corner points for detecting changes in building height at the floor level, and the other based on dual-phase DSMs for detecting the affected area and conducting pixel-based assessment of building collapse. Tian et al. [18] proposed a method that combines multispectral imagery and dual-phase DSMs generated from GeoEye-1 data, enabling simultaneous detection of collapsed buildings, newly constructed buildings, and temporary shelters. Wang and Li [21] utilized pre-earthquake stereo imagery to extract height variation features by combining DSM generated from postearthquake LiDAR data and performed damage assessment on buildings by first masking nonbuilding areas.

Various methods mentioned above have achieved several main classifications of building damage: complete collapse, partial collapse, intact, and newly constructed. However, such rough categorization does not fully exploit the potential of high-resolution stereo satellite images (HRSSIs) at the submeter level. Tong et al. [19] demonstrated the feasibility of accurate height change detection based on building corner points using 1 m resolution IKONOS imagery, proving its suitability for floor-level damage assessment. Nevertheless, the experiment conducted for the Dujiangyan earthquake in Sichuan involved the use of highly accurate DSMs generated from pre-earthquake total station measurement systems and ground control points (GCPs) measured with postearthquake RTK GPS, which are challenging to obtain in postearthquake areas. Consequently, the integration of LAPs, namely the geoscience laser altimeter system (GLAS) and the advanced topographical laser altimeter system (ATLAS), as reference data to enhance the geometric precision of the generated DSMs derived from stereo imagery, has garnered considerable attention. Notably, Tang et al. [25], [26] undertook a combined block adjustment of GLAS LAPs and Ziyuan-3 (ZY3) stereo images, successfully augmenting the vertical accuracy of ZY3 stereo images to align with the specifications stipulated for China's 1:50 000 scale mapping requirements. Furthermore, China's deployment of the GF-7 satellite on November 3, 2019 introduced a novel configuration incorporating a stereo optical camera, a laser altimeter, and a laser footprint camera. By combining the forward-viewing imagery with a resolution of 0.8 m, backward-viewing imagery with a resolution of 0.65 m, and the acquired sparse LAPs with their footprint imagery, it is feasible to enhance the vertical accuracy of DSMs to reach an approximate threshold of 1 m [27].

The efficacy and efficiency of DSM generation exert a substantial influence on the framework employed for building collapse assessments relying on HRSSIs. Numerous open-source algorithms, such as S2P, RSP, ASP, and CARS, have been successfully employed for high-resolution DSM generation [28], [29], [30], [31]. These algorithms utilize epipolar rectification based on feature matching and semiglobal matching (SGM) to achieve dense matching. Through a consistent processing

workflow, they facilitate high-quality 3-D reconstruction for a majority of terrains. However, the efficiency and memory usage for 3-D reconstruction present significant challenges due to the billions of pixels in a single GF-7 imagery. For instance, CARS completed line correction and DSM reconstruction on a Pleiades stereo pair covering London, with a size of $38\,881 \times 21\,197$ pixels, in 9 min and 16 min, respectively. Nonetheless, this performance was achieved with a hardware configuration of 24 CPUs, 120 GB RAM, and advanced distributed systems, leveraging cluster cooperation and distributed computing to improve computational efficiency. In the case of SPOT7 imagery in the mountainous regions of the French Alps with notable terrain variations, DSM generation in such a large-scale hardware cluster required as long as 2.5 h [30]. This highlights the need for further efficiency improvements in disaster response. Noteworthy advancements have been achieved in recent years with regard to expediting feature matching and dense matching through the utilization of diverse methodologies. First, in the realm of expediting real-time processes, feature matching has emerged as a vital technique, finding wide application in demanding domains, such as autonomous driving and SLAM engineering. Prominent methods, such as ORB and its refined variants, have played a central role in driving these advancements. [32]. Second, extensive research has been conducted on SGM algorithms accelerated by SIMD and GPU techniques, enabling real-time performance in generating disparity maps with consistent quality [33], [34]. These advancements serve as a foundation for the rapid 3-D reconstruction of HRSSIs.

In addition, the proficient utilization of building masks can significantly bolster the efficiency and automation prowess of evaluation. Given the substantial volume of collapsed buildings to locate and classify, relying solely on manual labor is not a viable option. To overcome this challenge, the authors in [21] and [35] utilized image segmentation methodologies to accomplish automated extraction and batch processing of building footprints. Initially, nonbuilding regions, such as vegetation, bare land, and shadows, were extracted, followed by the masking of these regions to retain only the building areas and achieve more accurate assessment of building damages. This approach is motivated by the fact that roads, vegetation, open land, and even lakes, which are nonbuilding areas, are not the main focus during building damage assessment. By effectively masking these irrelevant areas through the use of building masks, false detections can be avoided. In addition, the extraction of buildings and generation of polygonal building masks using semantic segmentation not only achieves the same effect but also provides independent building footprints [36]. By employing building footprint masks to obtain DSM difference images that contain solely the elevation change information related to the buildings, further techniques, such as threshold segmentation, clustering, and graph cuts, can be directly applied to extract pertinent information from each building footprint for collapse assessment [37]. This enables the possibility of simultaneously locating and classifying damaged buildings through the processing of large-scale satellite images containing a multitude of buildings.

In summary, the assessment of floor-level collapse in large-scale buildings located in seismic areas using dual-temporal

HRSSIs before and after earthquakes presents two major challenges. First, the acquisition of GCPs in a timely manner during emergencies, such as earthquakes, poses difficulty, thereby limiting the accuracy of HRSSIs' rational polynomial coefficients (RPCs) positioning for precise building positioning and collapse assessment. By integrating open-source ICEsat-2 LAPs for combined adjustment, enhancements can be made to the accuracy of the horizontal and vertical dimensions of dual-phase DSMs. However, limited research exists on the capabilities of HRSSIs in detecting building collapses. Second, existing methods for building detection and height change assessment, employing multitemporal HRSSIs, entail intricate execution steps, extensive manual editing, and generate high-resolution DSMs. Consequently, their practicality and responsiveness in the context of sudden earthquakes are significantly undermined. Hence, it is imperative to devise a streamlined process based on input data, aiming to achieve automated and expedited detection and evaluation of collapsed buildings.

Therefore, an improved method for building damage assessment based on HRSSIs before and after a seismic event is proposed in this article. The innovative aspects of this method are as follows.

1) Employing open-source LAPs in conjunction with GF-7 HRSSIs, our integrated approach effortlessly accomplishes the accurate geolocation and classification of extensive architectural structures. This comprehensive investigation convincingly corroborates its exemplary capacity for ascertaining the vertical levels of building floors.

2) Our swiftly advanced 3-D reconstruction methodology affords a remarkable improvement of more than eightfold in the efficiency of DSMs generation. Through the use of building masks to streamline the process, we achieve complete automation and marked improvement in the efficiency of building damage evaluation.

The rest of this article is organized as follows. Section II provides an introduction to the research domain and the associated dataset, while Section III elucidates the proposed methodology in intricate detail. Section IV showcases the experimental outcomes, while Section V critically analyzes the method's performance. Finally, Section VI concludes this article.

II. STUDY AREA AND DATA

A. Study Area

In the early hours of February 6, 2023, a devastating earthquake struck Kahramanmaraş in southeastern Turkey, measuring 7.8 on the Richter scale. Remarkably, within 9 h of this event, another earthquake measuring 7.6 occurred in the same region at a depth of approximately 10 km. These two earthquakes caused unimaginable destruction in Turkey and its neighboring countries, and the resulting building collapses and loss of life surpassed all expectations. As of March 20, 2023, the combined death toll has already exceeded 57 000 people and continues to rise, making it the deadliest earthquake in modern Turkish history. Particularly alarming is the fact that over one million buildings in Turkey alone have been damaged, with over 56 000 collapsing or suffering severe damage. More than six million

people have been displaced as a result, and the number of lives lost due to collapsed buildings is truly distressing [2], [38].

The capital of Kahramanmaraş province, severely impacted by both earthquakes, as illustrated in Fig. 1, was selected as the study area for investigation. This region is one of the hardest hit areas in Turkey, with a significant number of collapsed buildings visibly identifiable from high-resolution satellite imagery. The city of Kahramanmaraş is densely populated with over 70 000 houses in the urban area alone, featuring a highly complex distribution of buildings that intertwines numerous high-rise structures exceeding ten floors with a multitude of low-rise buildings. Moreover, the city's terrain exhibits variations of over 300 m, resulting in a considerable number of structures built on slopes. These challenges pose significant difficulties for automated extraction and damage assessment of buildings.

B. Data and Preprocessing

The GF-7 imagery used in this experiment consists of a pre-earthquake stereo pair captured on July 2, 2021 and a postearthquake stereo pair captured on the ninth day after the earthquake (February 15, 2023). The cloud coverage on the images was less than 5% and they exhibited excellent quality. The pre-earthquake imagery exhibited an overlap percentage of 94.16%, surpassing the 86.71% overlap of the postearthquake stereo imagery. The combined coverage of these four satellite images exceeded 600 km^2, almost encompassing the entire urban area of Kahramanmaraş. Additionally, the RPC file corresponding to each satellite image, as well as the coacquired LAPs, was provided.

For the external source of laser point data used in the stereo adjustment, the open-source ATLAS LAPs were employed. A total of 62 laser tracks within the image coverage area were selected from the website. The corresponding data attributes were extracted from the HDF5-formatted LAPs files based on the official documentation of ICESat-2 [39]. Initial filtering of the LAPs was performed with reference to the latitude and longitude shapefile of the study area.

The urban building footprints for Kahrmanmaraş city were obtained from the Microsoft GlobalMLBuildingFootprints datasets [https://github.com/microsoft/GlobalMLBuilding-Footprints (accessed on 19 May 2022)]. The extraction of buildings involves two main steps. The first step is semantic segmentation, where deep neural networks are used to recognize building pixels in aerial images. Once the building pixels are identified, the second step, known as polygonization, is carried out. Polygonization involves converting the detected building pixels into polygons, which represent the shape and boundaries of the buildings accurately. The accuracy and recall rate in the European region were 94.13% and 85.9%, respectively, with the footprints all being in the WGS-84 coordinate system. Upon the acquisition of Turkish JSON format footprints, the buildings were then clipped to retain only those within the study area, resulting in a total of 48 092 building footprints. The distribution of these building footprints, along with the pre-earthquake and postearthquake stereo satellite imagery, can
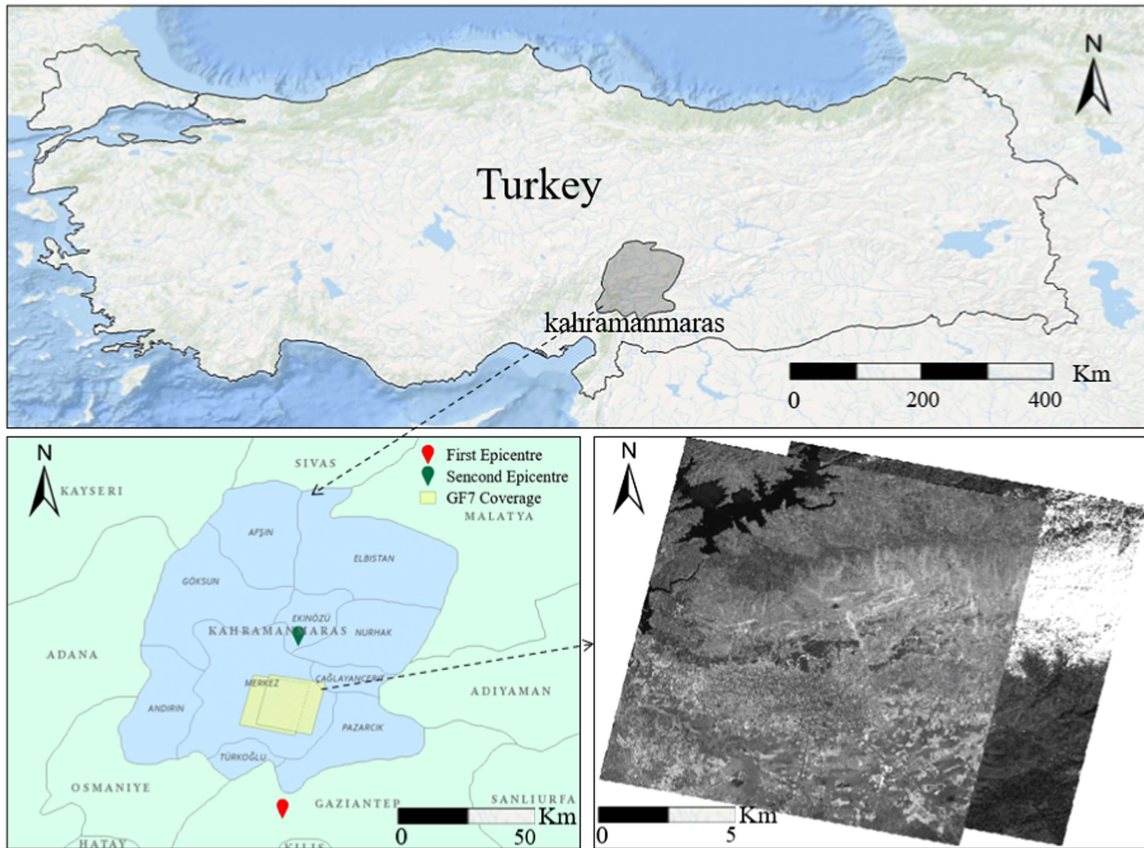
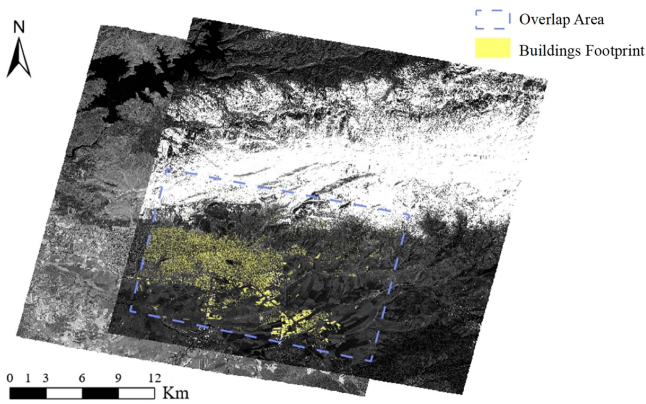Fig. 1.   Location map and dual-temporal HRSSIs of the study area.

Fig. 2.   Distribution of 48 092 building footprints on HRSSIs.

be seen in Fig. 2. The experimental data implemented in this study have been delineated in Table I.

III. PROPOSED METHOD

The entire process of rapid 3-D floor-level damage assessment for a large number of buildings utilizing high-resolution dual-temporal DSMs and building footprint masks is demonstrated in Fig. 3.

A. Rapid and Registered DSMs Generation

1) Combined Bundle Adjustment: The synchronized stereo imagery and LAPs obtained from the GF-7 satellite possess unique advantages, including minimal relative planimetric error, simultaneous data acquisition, and overlapping geographical coverage. However, the number of GF-7 LAPs is limited and unevenly distributed within the study area. In contrast, the ATLAS LAPs collected over the years exhibit higher ground resolution but contain a substantial amount of erroneous points. The overall bias of ATLAS LAPs is much larger than that of GF-7 LAPs, requiring strict selection. Therefore, this article adopted a combined bundle adjustment strategy, primarily using GF-7 LAPs as the main data source supplemented by ATLAS LAPs, to achieve the reorientation of pre- and postearthquake dual-temporal stereo pairs [27].

The rational function model is employed as the adjusted model for LAPs, which is based on the third-order rational polynomial transformations to accomplish the conversion between image space (row and column) and object space (*X, Y,* and *Z*). In order to address the errors arising from imprecise satellite scanning, attitude, and orbit positioning, the field of photogrammetry often relies upon the adoption of an affine transformation model. This widely employed model aims to mitigate these errors by effectively compensating for any misalignment present in the rows and columns of the satellite images, thereby facilitating their rectification. The construction of the adjustment model can

TABLE I
DETAILS AND SPECIFICATIONS OF TURKEY DATASETS

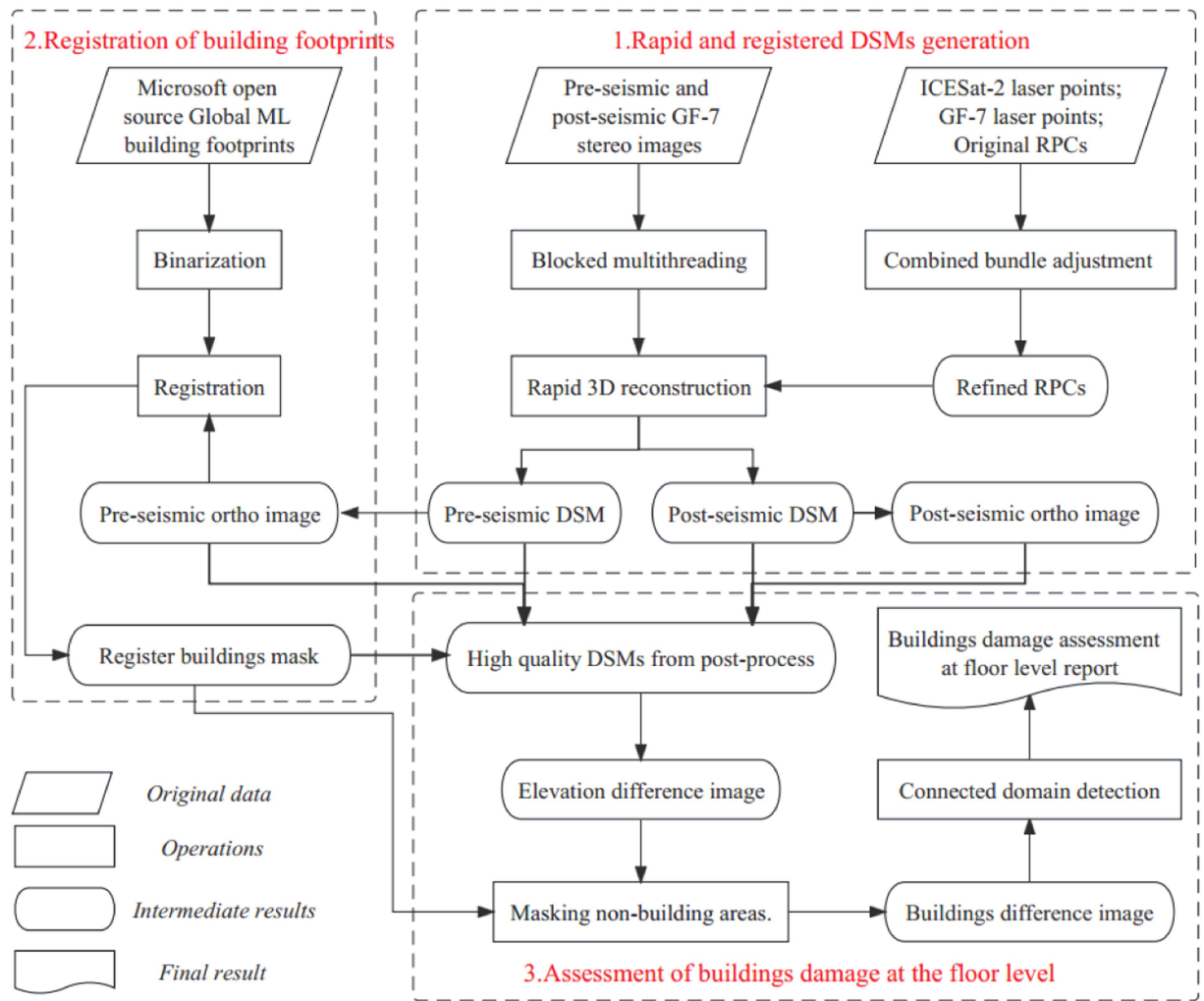| Datasets	Date	Time Phase	Number	Spatial Resolution (m)	laser beam diameter (m)
GF-7 HRSSIs	2021.7, 2023.2	Pre- and postseismic	4	0.65−0.8	\
Footprint images of GF-7 LAPs	2021.7, 2023.2	Pre- and postseismic	11	3.2	\
GF-7 LAPs	2021.7, 2023.2	Pre- and postseismic	11	\	30
ICESat-2 ATLAS	2018-2022	Preseismic	37 192	\	17.5
Buildings footprints	2015-2022	Preseismic	48 092	\	\

Fig. 3. General framework of the method proposed in this article. 1, 2, and 3 denote three successive steps mentioned in the proposed method.

be described as follows [40]:

$$V = AX_1 + BX_2 - L \qquad (1)$$

In the equation, V represents the residual term, A denotes the partial derivative of the undetermined solution coefficients in the affine transformation compensation model, and X_1 corresponds to the compensating quantity for the undetermined coefficients. B represents the partial derivative of the object coordinates at that point, while X_2 represents the square of the corresponding material compensation quantity. L refers to the residual of image
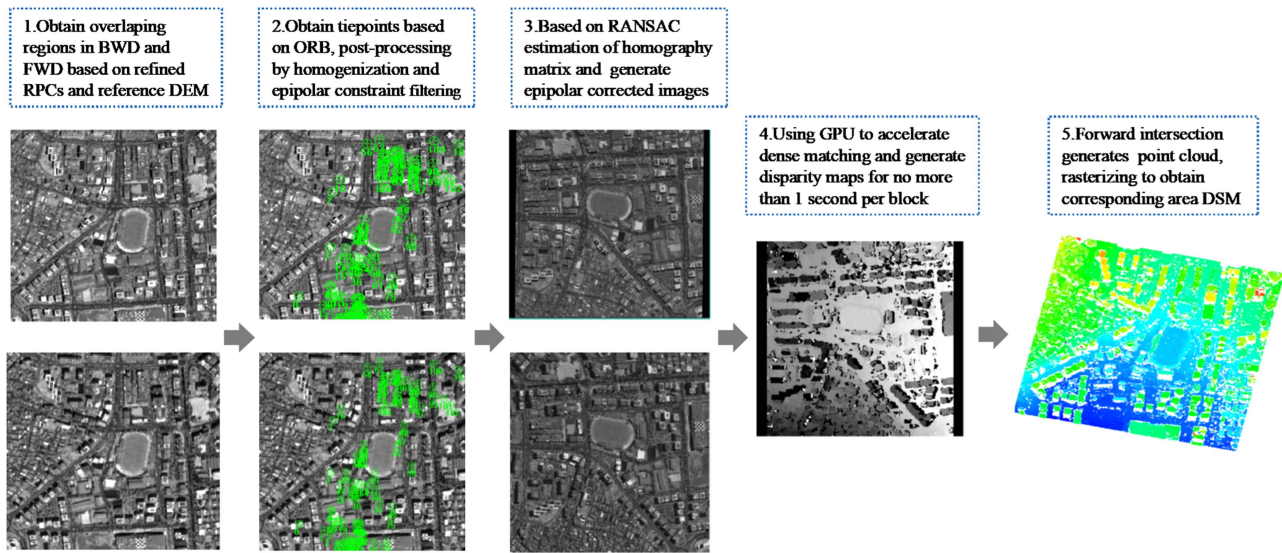
Fig. 4. Rapid 3-D reconstruction process proposed in this article.

coordinate differences between the model and observed values. The aforementioned equation represents the general form of the adjustment model. In practical research, adjustments are often required for different points, such as plane control separation and elevation. Within this article, the adjustment model primarily encompasses the following three types [25]:

$$\begin{cases} V_{tp} = A_{tp}\,X_{aff} + B_{tp}X_{tp} - L_{tp}, P_{tp} \\ V_{vcp} = A_{vcp}\,X_{aff} + B_{vcp}X_{vcp} - L_{vcp}, P_{vcp} \\ V_{las} = A_{las}\,X_{aff} + B_{las}X_{las} - L_{las}, P_{las}. \end{cases} \quad (2)$$

The first category is the junction point error equation, the second category is the virtual control point error equation, and the third category is the LAPs error equation. X_{aff} represents the undetermined compensation coefficients, X_{tp} is the object coordinate of the junction points, X_{vcp} represents the object coordinate of the virtual control points, and X_{las} denotes the object coordinate of the LAPs. P represents the weight of the points. Different types of points have different unknowns to solve in X.

2) Rapid 3-D Reconstruction: Currently, mainstream open-source software for satellite-based 3-D reconstruction follows a similar processing workflow: feature matching, epipolar line constraint, dense matching, forward intersection, and grid-based DSM generation [30]. Due to the large format of satellite imagery and the approximate hyperbolic form of epipolar lines, block-based multithreaded processing is commonly employed [41]. In the realm of blockwise processing of stereo image pairs, feature matching and dense matching emerge as the most time-intensive steps. Within the scope of this study, we put forth a rapid 3-D reconstruction approach that seeks to evaluate the algorithmic viability of these two aforementioned steps while also presenting potential alternative solutions. Furthermore, certain improvements are made to ensure the accuracy of DSM generation. Fig. 4 illustrates the processing workflow of the proposed fast 3-D reconstruction method for each block.

Before executing the 3-D reconstruction process, the accurate overlap areas need to be obtained for the stereo image pairs. In this method, the aft-looking image block serves as the reference image. By utilizing the shuttle radar topography mission (SRTM) with a resolution of 30 m as the elevation control and employing the refined RPCs with deviations within 10 pixels to establish the mapping relationship [42], then forward-looking image blocks with an intersection over union (IoU) exceeding 90% of reference image blocks are obtained. A sufficiently large IoU significantly enhances the efficiency and accuracy of subsequent feature matching and dense matching stages [43].

Next, feature matching is performed on the acquired stereo image block pairs. Currently, open-source libraries typically employ the scale-invariant feature transform (SIFT) or even slower Affine-SIFT (ASIFT) methods to extract the corresponding points. While these methods yield accurate and robust matches in various scenarios, they tend to be time-consuming. A comprehensive comparison of different feature matching algorithms revealed their relative efficiency, with the ranking from highest to lowest as follows: ORB > BRISK > SURF > AKAZE > SIFT > KAZE [44]. ORB generally exhibits a speed improvement of two magnitudes compared with SIFT, but its robustness and accuracy are significantly inferior. Moreover, ORBs distribution of feature points is uneven, with a strong concentration in regions with distinct features or edges. However, considering the study area characterized by heavily damaged urban zones, which encompass abundant building corners, edges, and textures, ORB can still extract numerous feature points. By subsequently applying strict selection criteria, high-quality feature points can be obtained.

Therefore, after obtaining the initial set of corresponding feature points using ORB, the following operations are conducted: based on the refined RPCs and the elevation range provided by the SRTM, a uniform distribution of virtual control point pairs is sampled on the stereo image block. The density of virtual control points is positively correlated with the elevation range.

The fundamental matrix estimated with virtual control points is then utilized to eliminate a considerable number of erroneous matches. To ensure a sufficient number of remaining matches, an adaptive distance threshold is employed. Furthermore, an Octree is employed to homogenize the distribution of the selected point pairs. Finally, RANSAC is used to estimate the epipolar line constraint and generate stereo image pairs of the same size with corrected epipolar lines. For cases where the number of ORB correspondences exceeds 50, and after the aforementioned steps, no less than 20 matches remain, accurate epipolar line constraints can be obtained. In challenging landscapes, such as mountains, lakes, or low-texture flat regions, where acquiring an adequate number of corresponding point pairs is arduous, a two-step approach is utilized. First, globally valid correspondences are identified, followed by the estimation of a global fundamental matrix to rectify the image blocks [45].

Regarding the most time-consuming process, dense matching, the mainstream SGM algorithm has achieved tremendous improvements in efficiency through hardware acceleration, such as FPGA and GPU. In this study, improvements are made on the GPU-accelerated SGM method proposed by Turker and Cetinkaya [24]. To meet the input requirements of the method, the first step involves optimal stretching to convert the 16-bit integer remote sensing images into 8-bit integer images. Then, to eliminate the illumination differences among high-resolution stereo image pairs captured from different viewing angles, color consistency adjustment using the Walis method is employed. Finally, an image pyramid is constructed to retrieve the disparity results at each level. Based on the disparity result from the upper level, the search range for pixel disparities in the lower level is reduced, further improving the efficiency of dense matching and reducing GPU memory consumption [46].

Subsequently, the refined RPCs and the dense correspondence relationship provided by the disparity map are utilized in the forward intersection to obtain corresponding object coordinates. Each image block yields point clouds in LAS format, featuring certain overlap regions between blocks. To obtain a consistent and accurate DSM, a weighted average fusion is performed on the overlapping regions of point clouds while eliminating the number of erroneous points. The point cloud is then resampled based on its density to meet the required initial DSM specifications [30]. For holes in the DSM and the problem of roughness on building surfaces, conventional postprocessing techniques, such as inverse distance interpolation and image-guided filtering, are employed in this method. The final result is a high-quality DSM.

B. Buildings Mask Registration

The directly downloaded Microsoft building footprint data cannot be directly used for extracting and calculating building height change information due to two main issues. The first issue is that the original format of building outlines can be used for content extraction, but it is challenging to statistically analyze elevation change information for a large number of buildings in automated procedures due to differences in resolution and data format compared with the rasterized DSM. Therefore, GDAL library is employed to read the building footprint file in GeoJSON format and resample it based on the resolution of DSM, resulting in a corresponding binary mask image. It is worth noting that, during the binarization process, overlapping and intersecting building footprints are excluded or merged based on the overlapping area proportion.

The second issue is the significant planar deviation between the open-source building footprints and generated DSM due to differences in positional accuracy and methods used for adjustment. This can be addressed by estimating the affine transformation relationship between the two. The widely used RANSAC method for estimating the affine transformation matrix requires a sufficient number of corresponding points and an inlier percentage greater than 50%. The building mask can directly utilize the polygon vertices (also the building corners) as feature points, whereas extracting corners on the DSM is challenging. In this article, the orthorectified image generated using the high-resolution DSM as the elevation reference is used as an intermediary. Given that the orthorectified image and DSM are inversely mapped to each other, aligning the building mask with the orthoimage can be considered as aligning it with the DSM. The methodology encompasses the following steps: first, obtain the building corners on the mask image, then utilize the FAST method within the range centered around each corner to acquire the building corners on the orthoimage, and apply nonmaximum suppression to retain the most effective points [47]. If the resulting number of points remains above one, the corner is discarded, which effectively avoids misalignment in densely populated building areas. Next, RANSAC is used to estimate the homography matrix for all corner points. Finally, based on the affine matrix, the alignment between the mask image and orthoimage is achieved [48]. The homogeneous transformation formula from the target pixel (x, y) to the reference pixel (u, v) is shown as follows:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & t_1 \\ a_3 & a_4 & t_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (3)$$

where the translation amounts are denoted as t_1 and t_2. Meanwhile, the parameters a_i encompass various transformations, including image rotation and scaling. By multiplying the homogeneous coordinates of each pixel on the target image by the affine matrix, the resulting pixel coordinates after transformation can be obtained, effectively accomplishing image registration.

C. Buildings Floor-Level Damage Assessment

Buildings can be extracted by masking nonbuilding areas in the elevation change map, resulting in an information map that only retains elevation changes within building regions. By analyzing the change information within each building polygon, object-based damage assessment reports can be generated. Therefore, connected component analysis is employed in this study to extract information for each irregular block and its internal contents [35].

After obtaining the information within each block, evaluating the extent of building damage based on the statistical analysis

TABLE II
THRESHOLD USED FOR SCREENING ICESAT-2 LAPS AND THE RESULTS

Filter Criteria	Difference from DEM	Urban	Slope	Cloud amount	Uncertainty
Threshold	$|H_{laser} - H_{dem}| < 10$ (m)	false	< 0.3	< 20%	< 5
Remaining points after filtering	26 382	22 393	10 627	3739	867

of change information requires an initial rough classification according to the European 1998 earthquake building damage grading standards. In this study, three main categories are considered: intact, partially collapsed, and completely collapsed. For intact buildings, classification can be achieved based on two distinct features: maximum collapse height and average collapse height, which are obtained after the first round of filtering. Partially collapsed buildings require further differentiation by utilizing the ratio of collapsed area to the total area of remaining buildings after the first round of filtering. The damaged buildings are then classified as either partially collapsed or completely collapsed. Finally, due to the presence of elevation calculation noise, we only consider the collapsed storeys to be credible if the proportion of collapsed pixels above fixed floors exceeds 30% of the total number of pixels in the building. The 30% threshold comes from the comparison results of small-scale testing. This is accomplished by iteratively summing the storeys from top to bottom, resulting in the final classification of collapsed storey levels. The detailed calculation and threshold values used for classification are provided in the following steps.

1) The maximum and average collapse heights of each building were computed, and those with a maximum collapse height of 3 m, as well as an average collapse height of 2 m or less, were directly categorized as structurally intact.
2) The total footprint area, collapsed area, and ratio of the collapsed area were computed. Buildings with a collapsed area ratio of less than 50% were classified as partially collapsed, while the rest were classified as completely collapsed.
3) Using Google Street view imagery, the approximate height of each floor in local buildings was estimated to be 3 m. The height change values in partially collapsed and completely collapsed buildings were divided by 3 m and rounded to determine the number of collapsed floors.
4) Starting from the most collapsed floor, collapsed pixel ratio was iteratively summarized for each collapsed floor until 30% was exceeded The final result is the number of collapsed floors corresponding to the pixel ratio exceeding 30% for the first time.

IV. EXPERIMENTS AND ANALYSIS

A. Results of Combined Bundle Adjustment

The indicators and thresholds used for filtering ICESat-2 LAPs are presented in Table II, which includes the difference with the SRTM 50 m reference DEM, whether it is an urban area, slope size, cloud amount, and uncertainty. Under these threshold settings, out of the initial 37 192 LAPs, only 867 were retained after filtering. The filtered LAPs were then projected onto GF-7

forward and backward images using RPC projection, and precise image coordinates were obtained by least square matching based on the coordinates from the backward image. During this process, matching points with a least square matching confidence of less than 0.8 and their corresponding LAPs were eliminated, resulting in a total of 276 reliable LAPs.

Combined bundle adjustment was performed on the dual-temporal stereo image pair using 12 GF-7 LAPs and 16 evenly distributed ICESat-2 LAPs selected from them. The remaining 260 LAPs were used as check points to validate the adjustment results.

In Table III, the MID, RMSE, MEAN, and MAX represent the median, root-mean-square error, mean, and maximum values of elevation and plane accuracy obtained using checkpoints, respectively. It can be observed that both the free network adjustment and the combined adjustment were able to optimize the planimetric accuracy from greater than two pixels to within one pixel. The RMSE of the elevation positioning errors before and after the earthquake, obtained through the combined adjustment, remained within 1 m. This level of accuracy is sufficient to achieve precise damage detection of buildings based on dual-temporal DSMs.

B. Efficiency Comparison of DSM Generation

The 0.8-m resolution DSM was generated from the GF-7 HRSSIs using rapid 3-D reconstruction workflow, as described in Section III-A. Following the positive correlation between block size and bundle adjustment error stated in the S2 bundle adjustment paper, the backward image with a width of 35 864 pixels and height of 40 000 pixels was divided into 1600 small blocks of size 1000 × 1000 pixels each. These blocks had overlapping regions to prevent gaps in the final generated entire DSM.

The rapid generation of the DSM primarily focuses on accelerating two steps: sparse matching and dense matching. Table IV compares the processing time of different sparse matching and dense matching methods for individual blocks and the generation of the entire DSM. All timing experiments were carried out on a computer system comprising an Intel i5-11400X CPU @ 2.70 GHz, 16 GB RAM, 12 cores, and an NVIDIA GeForce RTX 3050 Laptop 2 GB GPU. All methods were implemented using C++. In the case of processing all blocks, different sparse matching and dense matching methods required varying amounts of memory, resulting in different numbers of threads used, which are indicated in the last row of Table IV.

According to Table IV, it can be observed that using ORB had improved the processing speed by approximately 7.7 times compared with SIFT used in S2P and CARS for single image

TABLE III
EXPERIMENTAL RESULTS OF THE THREE METHODS ON FOUR DATASETS

Adjustment		Height (m)				Residual (pixel)			
scheme		MID	RMSE	MEAN	MAX	MID	RMSE	MEAN	MAX
Pre-seismic	Before adjustment	6.34	6.67	6.42	9.12	1.88	1.82	1.78	3.51
	Free network adjustment	6.24	6.62	6.18	9.56	0.17	0.23	0.15	2.77
	Adjustment with ICESat-2, GF-7	0.71	0.87	0.75	2.41	0.56	0.35	0.62	1.77
Post-seismic	Before adjustment	3.08	3.15	2.94	−5.12	1.21	2.24	1.23	4.12
	Free network adjutment	3.30	3.32	3.28	−6.17	0.26	0.17	0.29	0.83
	Adjustment with ICESat-2, GF-7	0.81	0.97	0.91	2.79	0.49	0.42	0.63	1.23

TABLE IV
DIFFERENT METHODS FOR SPARSE MATCHING AND DENSE MATCHING TAKE TIME AND NUMBER OF THREADS USED

	Sparse matching			Dense matching		
	SIFT	SURF	ORB	SGM	SGBM	SGM-GPU
Single-block time consumption (s)	5.19	3.27	0.67	35.12	6.85	0.22
ALL blocks time consumption (min)	23.56	12.47	4.78	184.98	32.83	8.81
Number of threads used	6	8	10	6	8	4

block processing. Furthermore, the speed for processing the entire mosaic consisting of 1600 blocks had increased by more than 4.9 times. The increase in speed for the entire mosaic was lower than the increase for single image block processing due to two reasons. First, the ORB single image block matching time was less than 1 s, but the time taken for image and homologous point file reading and saving cannot be ignored. Second, the time-consuming process of filtering homologous points based on the estimated fundamental matrix using virtual control points also contributed to this difference. The time taken for these two parts was also approximately 1 s.

The use of SGM-GPU resulted in even more significant speed improvements. The traditional SGM occupied excessive memory, which led to fewer threads and obvious time consumption. However, after GPU acceleration, the overall processing time for the entire image could be reduced from 3 h to less than 10 min. Most of the 10-min processing time was attributed to image reading and writing as well as process switching. Additionally, the limited GPU memory of the experimental laptop also restricts the number of threads.

Apart from the time spent on sparse matching and dense matching, the total time for image preprocessing, epipolar image generation, forward intersection, and rasterization to generate the DSM was 13 min. Therefore, the total time for generating the complete DSM was 27 min and generating the dual-temporal DSMs before and after an earthquake took less than an hour using the experimental laptop. This will significantly save data
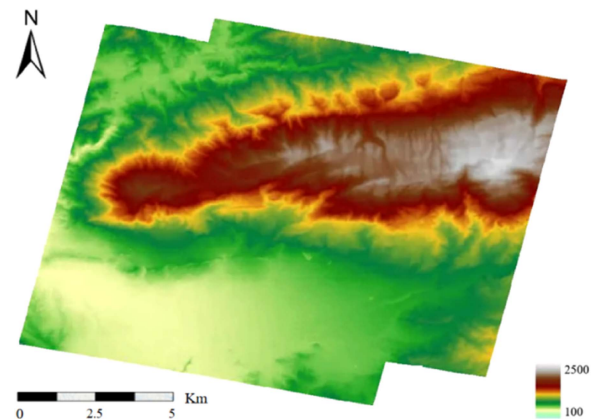

Fig. 5.   Pre- and postseismic DSMs from rapid 3-D reconstruction.

processing time for postdisaster rescue operations. The results of generating the dual-temporal DSMs before and after the earthquake are shown in Fig. 5.

C. Results of Buildings Mask Registration

First, the binarization method, as proposed in Section III-B, was used to rasterize the original format building footprints into a binary mask. The resulting mask, in an unsigned 8-bit data format, represented building and nonbuilding regions as
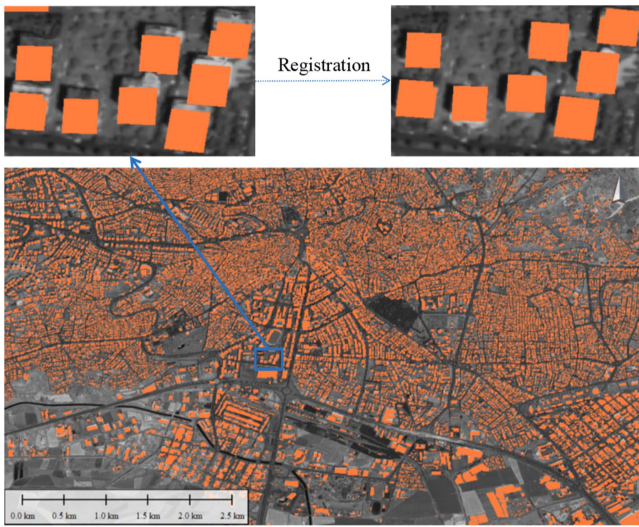
Fig. 6. Offset between buildings mask and DOM before and after registration.

0 and 255, respectively. Moreover, this mask shared the same resolution and frame as the orthorectified image generated based on our high-resolution DSM.

Next, the FAST method was employed to detect corner points in the binary mask image. A total of 48 092 building detections resulted in 230 000 corner points. However, it was determined that not all of these points were necessary for registration. Therefore, the octree algorithm was applied to uniformly select just over 3000 points evenly distributed across the entire image. Within the neighborhoods of these selected points, corresponding matching points were searched on the orthorectified image. As shown in Fig. 6, it could be observed that the uncorrected building footprints generally shifted toward the lower left direction. To account for this, the search window for matching points was artificially adjusted to a 35×35 window in the upper right direction.

Subsequently, the filtering method, as mentioned in Section III-B, was used to retain 567 pairs of accurately matched points and estimate the corresponding affine transformation matrix. The offsets between the preregistered and postregistered masks and their corresponding orthorectified images were illustrated in Fig. 6. The registered building footprints exhibited excellent alignment with the buildings depicted in the orthorectified image, resulting in an optimal congruence between the building mask image and the dual-temporal DSM. This alignment guarantees the efficacy of subsequent procedures in extracting building regions.

D. Results of Building Damage Assessment

In order to substantiate the accuracy of our proposed methodology, a preliminary small-scale experiment was undertaken in selected areas, where the ground truth values of building floor numbers were acquired through Google Street view [49]. The 3-D reconstruction process for this region was illustrated in Fig. 4, and the dual-temporal DSMs were generated by our

method. Evaluation was performed using 30% scaling parameters proposed in Section III-C and the confusion matrix was computed by comparing the predicted collapsed states with the actual floor numbers. The results, as shown in Table V, indicated the accuracy in predicting collapsed buildings, which included four partially collapsed buildings.

The evaluation method, as introduced in Section III-D, allowed for the assessment of collapse at the floor level, going beyond the mere determination of whether a building has collapsed. However, achieving floor-level assessment posed greater difficulty. As illustrated in Table VI, six intact buildings actually collapsed without being detected, while ten buildings were wrongly classified as collapsed when they were actually intact. These 16 buildings primarily consisted of low-rise structures, and inaccuracies in the reconstruction results were attributed to issues, such as occlusion and shadows. In terms of floor-level assessment for collapsed buildings, errors in detecting low-rise buildings were more prevalent than those in detecting high-rise buildings. The precision and recall rates for buildings over five stories were 85.17% and 91.30%, respectively, higher than the rates for buildings below five stories, which stood at 84.61% and 75.86%. This is partly because high-rise buildings exhibit more pronounced height variations, making them easier to detect using our method. Additionally, within this area, there were 46 collapsed high-rise buildings, outnumbering the 29 collapsed low-rise buildings. The final accuracy of the floor-level damage assessment for the 361 detected buildings was 93.27%.

We acquired the assessment results of floor-level collapse for the buildings, as illustrated in Fig. 7. Through the dual-phase images pre- and postearthquake, as displayed in Fig. 7(a) and (b), it is evident that the selected region is highly representative and satisfies the evaluation requirements. First, this region encompasses over 300 buildings with their corresponding collapse states and the accurate count of collapsed floors, unlike previous studies, where only a few or a dozen buildings were considered. Second, the area, as depicted in Fig. 7, consists of diverse building types, including closely interconnected midrise buildings highlighted in yellow, compact and low residential structures marked in white, expansive shopping complexes indicated in green, and scattered high-rise buildings enclosed in orange, which essentially encapsulate the comprehensive range of 48 092 building types. Finally, this particular area exhibits one of the most severe collapse scenarios within the entire urban area of Kahramanmaras, making it a focal point for the application of this methodology.

While there is currently a paucity of pertinent research comparing the accuracy of building floor-level collapse assessment results, our study has made substantial improvements compared with Wang's findings. The results, as presented in Table VII, show an increase of 4.06% and 3.74% in our OA and Kappa coefficients, respectively, in contrast to Wang's methodology. This significant progress underscores the critical role played by precise building localization and altitude information in the assessment of building collapses. Remarkably, for intact buildings, the producer accuracy and user accuracy reached levels of 95.55% and 97.90%, respectively. This achievement is attributable to the utilization of high-resolution DSM, as well as

TABLE V
ASSESSMENT RESULTS OF LOCAL BUILDINGS COLLAPSE

State of the buildings		Reference		
		Collapsed	Not collapsed	Total
Estimated value	Collapsed	65	10	75
	Not collapsed	6	280	286
	Total	71	290	361
User's accuracy		86.67%	97.90%	92.29%
Produce's accuracy		91.55%	96.55%	94.05%
Overall accuracy				95.56%
Kappa coefficient				86.27%

TABLE VI
ASSESSMENT RESULTS OF LOCAL BUILDINGS COLLAPSE AT FLOOR LEVEL

Confusion matrix for building damage floor-level assessment		Reference value (damaged floors)											
		0	1	2	3	4	5	6	7	8	9	10	total
Estimated value (damaged floors)	0	280	3	1	1	1	0	0	0	0	0	0	286
	1	3	3	1	0	0	0	0	0	0	0	0	7
	2	1	0	1	0	0	0	0	0	0	0	0	2
	3	0	0	0	0	1	0	0	0	0	0	0	1
	4	1	0	0	0	3	1	0	0	0	0	0	5
	5	2	0	1	0	0	4	0	0	0	0	0	7
	6	2	0	0	0	0	0	7	1	0	0	0	10
	7	1	0	0	0	0	0	0	17	1	0	0	19
	8	0	0	0	0	0	0	0	1	8	0	0	9
	9	0	0	0	0	0	0	1	0	0	6	0	7
	10	0	0	0	0	0	0	0	0	0	0	4	4
	total	290	4	4	1	5	5	8	19	9	6	4	357

TABLE VII
BUILDING DAMAGE EXTRACTION RESULTS USING DIFFERENT METHODS (ALL IN PERCENT)

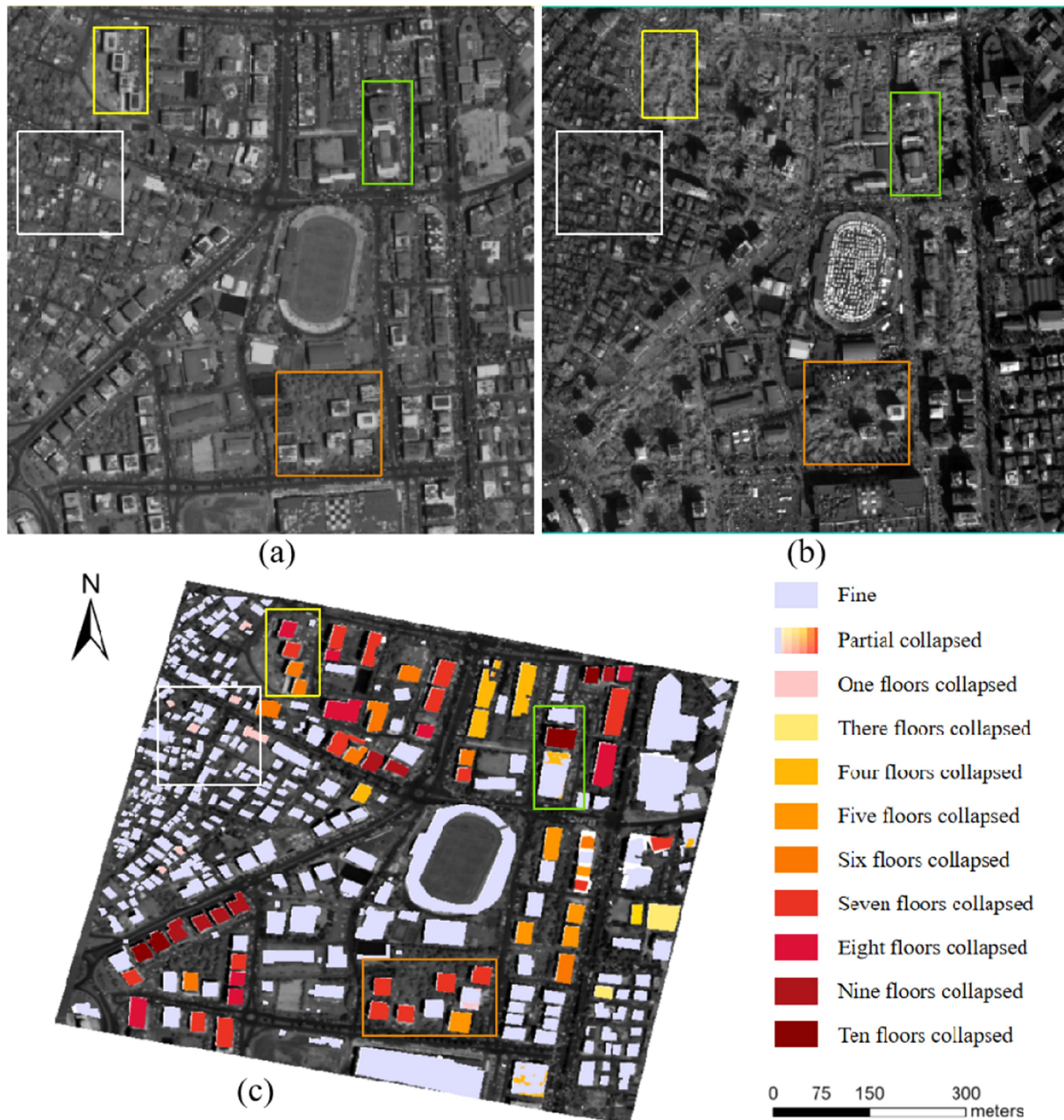| Methods	OA	Kappa	building damage		Nonbuilding damage	
			PA	UA	PA	UA
Wang's method [21]	91.53	82.53	88.56	90.96	93.67	91.92
Proposed method	95.56	86.27	91,55	86.67	96.55	97.90

Fig. 7. (a) and (b) Original images of the local area pre- and postearthquake. (c) Distribution map of floor-level assessment results for buildings collapse in local area.

the application of building masks to filter out shadows and minimize interference from ground-level data. It is imperative to note that further research is essential to validate and extend our findings to different scenarios. Moreover, the adoption of advanced technologies, such as drones and advanced sensors, should be considered to acquire superior quality building data, thereby enhancing the accuracy of collapse assessment methodologies. Ultimately, these improvements have significant implications for bolstering public safety and facilitating more effective disaster response measures.

Given that the collapse assessment results of buildings in the local area are consistent with expectations, the same parameters can be employed to carry out the same procedures for the entire stereo imagery. The ultimate outcome is outlined as follows. A total of 2709 collapsed buildings were detected, consisting of 2685 fully collapsed and 24 partially collapsed structures. Among the collapsed buildings, the distribution of the number of floors affected is as follows: 1 story: 1611; 2 stories: 524; 3 stories: 232; 4 stories: 137; 5 stories: 66; 6 stories: 61; 7 stories: 47; 8 stories: 15; 9 stories: 11; and above 9 stories: 5. These numbers are consistent with the figures provided by the Turkish authorities (https://www.sbb.gov.tr/turkiye-earthquakes-recovery-and-reconstruction-assessment).

TABLE VIII
COMPARISON OF DIFFERENT MATCHING METHODS RESULTS

Feature match scheme	Proportion of points' number			Distribution rate of points			Accuracy of points(pixel)		
	>500	50–500	<50	MEAN	MIN	MAX	MEAN	MID	MAX
ORB	77.7%	19.2%	3.1%	94.7%	8.0%	100%	1.13	0.78	13.12
SIFT	79.6%	19.6%	0.8%	97.8%	13.0%	100%	0.47	0.41	1.89
Refined ORB	77.1%	18.7%	4.2%	92.4%	1.0%	100%	0.62	0.49	2.37

Additionally, we calculated the collapsed area and volume as two indicators. There were 1732 buildings with a collapsed area exceeding 200 m^2 and 795 buildings with a collapsed volume exceeding 500 m^3. The total collapsed volume reached 5 million m^3. These indicators can help identify buildings with severe damage, providing valuable guidance for postdisaster rescue and reconstruction efforts. Furthermore, the collapsed volume, which is directly proportional to the number of casualties, serves as an important 3-D indicator, enabling a rough estimation of the approximate number of casualties. This information is crucial for subsequent international humanitarian assistance concerning food, water, medical supplies, and other resources.

V. DISCUSSION

A. Feature Matching Quality Assessment

The quality of the correspondences obtained from feature matching is a critical factor that affects the quality of the epipolar line constraint and dense matching. While using ORB as a replacement for SIFT provides nearly a tenfold improvement in speed, the lower quality of ORB correspondences results in approximately one-fourth of the 1600 image blocks of single-pair HRRSIs failing the epipolar line correction, with discrepancies exceeding one pixel in the estimated corresponding epipolar lines [32]. By applying the postprocessing approach proposed in this article, erroneous points in ORB correspondences are significantly removed, thereby enhancing the accuracy of individual epipolar lines. Furthermore, by propagating information from neighboring image blocks, even a small fraction of image blocks where the epipolar line correction failed can achieve robust fundamental matrix estimation, leading to a noticeable improvement in the overall accuracy of the epipolar lines across the entire image dataset.

The quality assessment of feature points includes the following three indicators. First, the number of matches was quantified, with the maximum number of points attainable from various matching techniques being established at 1000. The final number of points is divided into three categories: greater than 500 (sufficient for epipolar line estimation), 50 to 500 (may introduce significant errors or mistakes in epipolar line estimation after filtering), and less than 50 (can hardly achieve accurate epipolar line estimation). Second, the distribution of feature points, whether they are evenly distributed or not, also affects the quality of epipolar line identification. For each image

block, a division into an $n \times n$ grid is performed, wherein the distribution rate represents the ratio of grids containing points to the total number of grids. In this context, the value of n is determined as 10, taking into account the dimensions of the image block. Regarding the most important indicator, feature matching accuracy, we applied S2P to 1600 pairs of image blocks for epipolar line constraints. We manually processed the image blocks with epipolar line constraint errors exceeding 1 and ultimately obtained 1600 fundamental matrices with an epipolar line error less than 0.6. Based on these fundamental matrices, we calculated the reprojection errors of the correspondences to evaluate their accuracy. Table VIII presents a comparison among the original ORB, SIFT, and postprocessed ORB with the MIN, MAX, MID, and MEAN values indicating the minimum, maximum, median, and mean values of each indicator. The results indicate that ORB and SIFT have similar numbers of feature points and distribution. However, the optimized ORB slightly underperforms in these two indicators compared with the other two methods. The key focus lies in the accuracy of the correspondences, where it can be observed that the postprocessed ORB, incorporating virtual control point filtering, RANSAC, and regularization, significantly improves the quality of correspondences. It outperforms the original ORB with a 45% increase in average accuracy, which is comparable to SIFT and both within 0.7 pixels. It is worth noting that some hard-to-detect erroneous points that severely impact the results of epipolar line constraints have been successfully filtered out.

B. Impact of Image Resolution on Evaluation Results

In the proposed method for assessing building collapse at the floor level in this article, satellite image resolution is a key factor influencing the quality of DSM reconstruction and evaluation accuracy. However, there have been limited studies on building reconstruction using satellite stereo imagery, particularly with an unclear standard for the required image resolution in engineering applications related to reconstruction. In this study, we utilized multiscale data obtained from different downsampled resolutions of GF-7 HRSSIs and applied the proposed workflow for reconstruction. Subsequently, we conducted a comparative evaluation of accuracy and efficiency in the same study area, providing a robust basis for selecting usable data and practical data processing in future research.

TABLE IX
IMPACT OF IMAGE RESOLUTION ON EVALUATION RESULTS

| Different resolution HRSSIs	DSM elevation accuracy from ICESat-2 GCPs(m)			Accuracy of building floor-level collapse assessment (%)				Rapid 3-D reconstruction time consuming(min)	
	MAX	MEAN	STD	ACC	PRE	REC	F1	Feature match	Dense match
0.8 m	5.15	1.74	0.97	93.27	84.91	87.63	86.24	4.78	8.81
1.0 m	6.12	1.67	1.05	93.15	84.47	87.26	85.84	3.76	7.02
1.2 m	4.27	1.81	1.17	91.39	82.34	84.37	83.34	3.22	5.78
1.5 m	6.14	1.65	0.95	90.22	80.27	79,58	79.92	2.49	4.62
2.0 m	5.27	1.69	1.01	83.54	73.57	67.56	70.43	1.87	3,45
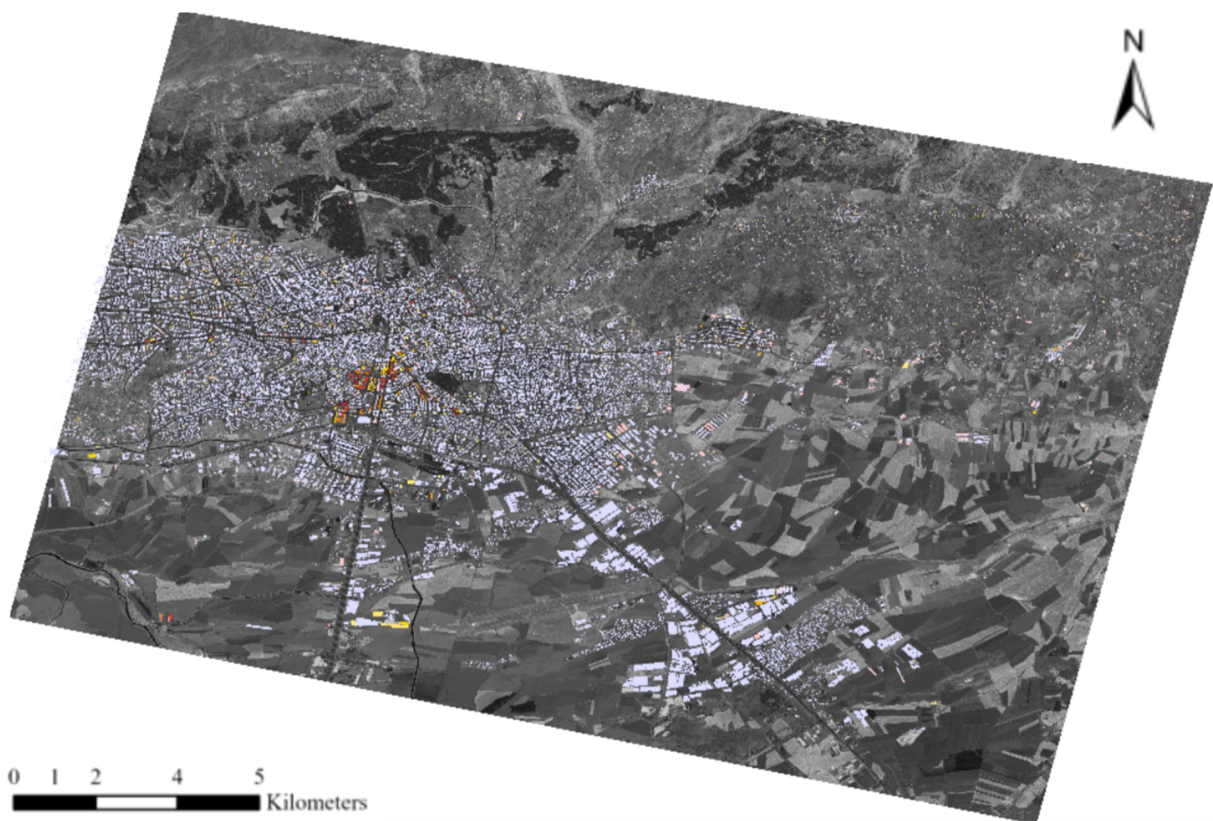
Fig. 8. Floor-level assessment results for buildings collapse in all overlap areas.

Initially, we resampled the initial 0.65-m resolution backward-looking image to 0.8-m resolution to match the forward-looking image. Then, we individually downscaled the forward- and backward-looking images to create reference image groups with resolutions of 1, 1.2, 1.5, and 2 m. Stereoscopic adjustment, fast 3-D reconstruction, and building collapse assessment were performed on each of these image groups. The accuracy of DSM elevation was evaluated using the global ICESat-2 laser point data, the true number of building floors was evaluated for collapsed buildings, and the execution efficiency was assessed by measuring the time required for fast 3-D reconstruction.

In Table IX, standard deviation, accuracy, precision, and recall are represented by STD, ACC, PRE, and REC, respectively. It can be observed that the evaluation of DSM elevation accuracy using LAPs is consistent across different resolutions. This is due to the careful selection of LAPs, which excludes those on buildings and steep terrains. Thus, the impact of resolution on reconstruction in flat areas is minimal. However, the assessment of collapsed building floor levels is significantly influenced by different image resolutions.

As a threshold, the accuracy of assessing collapsed building floor levels remains relatively unchanged for submeter resolutions of 0.8 and 1 m. This suggests that the resolutions of 1 m

and higher can provide equally satisfactory results, ensuring the accurate reconstruction and capturing precise information about vertical changes for various types of buildings. However, when the image resolution exceeds 1 m, even with identical processing workflows, the accuracy of assessing collapsed floor levels rapidly decreases. Through the observation of evaluation results, it can be attributed to a significant decline in accuracy when evaluating smaller sized buildings. Additionally, Table IX clearly indicates a positive correlation between the efficiency of 3-D reconstruction and the square of image resolution. Therefore, it is possible to adopt different downsampling ratios for image data according to specific accuracy requirements, enhancing efficiency, especially when dealing with large-frame or large-scale images. For example, in this case, downsampling the original image to 1-m resolution can improve efficiency by over 20% while maintaining almost unchanged accuracy in assessing collapsed building floor levels.

VI. Conclusion

To address the challenges of lengthy DSM generation, complex imaging conditions, viewpoint differences, and dense building distribution in accurate building collapse assessment, we propose an improved method for fast assessment of building collapses at the floor level. This method achieves the first-time simultaneous positioning and classification of large-scale collapsed buildings, elevating building collapse classification to the floor level by fully exploiting the vertical accuracy of subpixel resolution HRSSIs.

Experimental results in the earthquake-stricken region of Kahramanmaras, Turkey, demonstrate that we obtain HRSSIs positioning parameters with vertical accuracy of around 1 m by using ICESat-2 LAPs to construct a combined adjustment model. The proposed fast 3-D reconstruction method not only achieves high-precision submeter resolution DSMs but also improves DSM generation efficiency by more than eight times. To assess the accuracy of our method effectively, we selected severely damaged areas in the city center for validation, including 361 densely distributed buildings. Using Google Street view images as ground truth, the floor-level assessment achieved a high accuracy rate of 93.27%. Finally, floor-level collapse assessment was conducted for 48 092 buildings in the experimental area, resulting in a large-scale distribution map of building collapses, as shown in Fig. 8. Additionally, through automated detection, the quantity of collapsed buildings (2709 buildings), collapsed floor count, collapsed area, and collapsed volume was determined in one go. This provides significant geospatial and data support for postdisaster emergency response, damage assessment, and rapid reconstruction.

However, it is important to note some limitations of the proposed automated method. For instance, it will fail if there are missing or misidentified building footprints since the accurate identification of buildings heavily relies on building footprints. Although such occurrences are rare, cautious selection and examination of footprint data should be increased. Using satellites from other sources and long time series of satellites for detailed building detection and assessment will also be a direction for our future research.

References

[1] L. Dal Zilio and J.-P. Ampuero, "Earthquake doublet in Turkey and Syria," *Commun. Earth Environ.*, vol. 4, no. 1, 2023, Art. no. 71.

[2] W. Chen, G. Rao, D. Kang, Z. Wan, and D. Wang, "Early report of the source characteristics, ground motions, and casualty estimates of the 2023 Mw 7.8 and 7.5 Turkey earthquakes," *J. Earth Sci.*, vol. 34, no. 2, pp. 297–303, 2023.

[3] I. Ahmed, "Key building design and construction lessons from the 2023 Türkiye–Syria earthquakes," *Architecture*, vol. 3, no. 1, pp. 104–106, 2023.

[4] A. Menderes, A. Erener, and G. Sarp, "Automatic detection of damaged buildings after earthquake hazard by using remote sensing and information technologies," *Procedia Earth Planet. Sci.*, vol. 15, pp. 257–262, 2015.

[5] 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, 2013.

[6] C. Corbane et al., "A comprehensive analysis of building damage in the 12 January 2010 MW7 Haiti earthquake using high-resolution satellite and aerial imagery," *Photogrammetric Eng. Remote Sens.*, vol. 77, no. 10, pp. 997–1009, 2011.

[7] 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.

[8] M. Ji, L. Liu, and M. Buchroithner, "Identifying collapsed buildings using post-earthquake satellite imagery and convolutional neural networks: A case study of the 2010 Haiti earthquake," *Remote Sens.*, vol. 10, no. 11, 2018, Art. no. 1689.

[9] J. Jeong and T. Kim, "Analysis of dual-sensor stereo geometry and its positioning accuracy," *Photogrammetric Eng. Remote Sens.*, vol. 80, no. 7, pp. 653–661, 2014.

[10] A. Cooner, Y. Shao, and J. Campbell, "Detection of urban damage using remote sensing and machine learning algorithms: Revisiting the 2010 Haiti earthquake," *Remote Sens.*, vol. 8, no. 10, 2016, Art. no. 868.

[11] X. Wang and P. Li, "Extraction of earthquake-induced collapsed buildings using very high-resolution imagery and airborne lidar data," *Int. J. Remote Sens.*, vol. 36, no. 8, pp. 2163–2183, 2015.

[12] J. Liu and P. Li, "Extraction of earthquake-induced collapsed buildings from bi-temporal VHR images using object-level homogeneity index and histogram," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 8, pp. 2755–2770, Aug. 2019.

[13] C. Wu et al., "Building damage detection using U-net with attention mechanism from pre- and post-disaster remote sensing datasets," *Remote Sens.*, vol. 13, no. 5, 2021, Art. no. 905.

[14] S. Ahadzadeh et al., "Detection of damaged buildings after an earthquake using artificial neural network algorithm," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 37, pp. 369–372, 2008.

[15] D. Duarte, F. Nex, N. Kerle, and G. Vosselman, "Satellite image classification of building damages using airborne and satellite image samples in a deep learning approach," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. IV-2, pp. 89–96, 2018.

[16] S. Klonus, D. Tomowski, M. Ehlers, P. Reinartz, and U. Michel, "Combined edge segment texture analysis for the detection of damaged buildings in crisis areas," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 4, pp. 1118–1128, Aug. 2012.

[17] X. Tong et al., "Use of shadows for detection of earthquake-induced collapsed buildings in high-resolution satellite imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 79, pp. 53–67, 2013.

[18] J. Tian, A. A. Nielsen, and P. Reinartz, "Building damage assessment after the earthquake in Haiti using two post-event satellite stereo imagery and DSMs," *Int. J. Image Data Fusion*, vol. 6, no. 2, pp. 155–169, 2015.

[19] X. Tong et al., "Building-damage detection using pre- and post-seismic high-resolution satellite stereo imagery: A case study of the May 2008 Wenchuan earthquake," *ISPRS J. Photogramm. Remote Sens.*, vol. 68, pp. 13–27, 2012.

[20] J. Tu, H. Sui, W. Feng, and Z. Song, "Automatic building damage detection method using high-resolution remote sensing images and 3D GIS model," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. III-8, pp. 43–50, 2016.

[21] X. Wang and P. Li, "Extraction of urban building damage using spectral, height and corner information from VHR satellite images and airborne LiDAR data," *ISPRS J. Photogramm. Remote Sens.*, vol. 159, pp. 322–336, 2020.

[22] A. Vetrivel, M. Gerke, N. Kerle, F. Nex, and G. Vosselman, "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.*, vol. 140, pp. 45–59, 2018.

[23] A. Vetrivel, M. Gerke, N. Kerle, and G. Vosselman, "Identification of damage in buildings based on gaps in 3D point clouds from very high resolution oblique airborne images," *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 61–78, 2015.

[24] M. Turker and B. Cetinkaya, "Automatic detection of earthquake-damaged buildings using DEMs created from pre- and post-earthquake stereo aerial photographs," *Int. J. Remote Sens.*, vol. 26, no. 4, pp. 823–832, 2006.

[25] G. Li, X. Tang, X. Gao, H. Wang, and Y. Wang, "ZY-3 block adjustment supported by GLAS laser altimetry data," *Photogrammetric Rec.*, vol. 31, no. 153, pp. 88–107, 2016.

[26] P. Zhou, X. Tang, Z. Wang, N. Cao, and X. Wang, "Vertical accuracy effect verification for satellite imagery with different GCPs," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 8, pp. 1268–1272, Aug. 2017.

[27] X. Tang, P. Zhou, L. Guo, and H. Pan, "Integrating stereo images and laser altimetry points derived from the same satellite for high-accuracy stereo mapping," *Remote Sens.*, vol. 15, no. 4, 2023, Art. no. 869.

[28] R. A. Beyer, O. Alexandrov, and S. McMichael, "The Ames stereo pipeline: NASA's open source software for deriving and processing terrain data," *Earth Space Sci.*, vol. 5, no. 9, pp. 537–548, 2018.

[29] C. de Franchis, E. Meinhardt-Llopis, J. Michel, J. M. Morel, and G. Facciolo, "An automatic and modular stereo pipeline for pushbroom images," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. II-3, pp. 49–56, 2014.

[30] J. Michel et al., "A new satellite imagery stereo pipeline designed for scalability, robustness and performance," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. V-2-2020, pp. 171–178, 2020.

[31] R. Qin, "RPC stereo processor (RSP)–A software package for digital surface model and orthophoto generation from satellite stereo imagery," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. III-1, pp. 77–82, 2016.

[32] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vis.*, 2011, pp. 2564–2571.

[33] I. Haller and S. Nedevschi, "GPU optimization of the SGM stereo algorithm," in *Proc. IEEE 6th Int. Conf. Intell. Comput. Commun. Process.*, 2010, pp. 197–202.

[34] D. Hernandez-Juarez, A. Chacón, A. Espinosa, D. Vázquez, J. C. Moure, and A. M. López, "Embedded real-time stereo estimation via semi-global matching on the GPU," *Procedia Comput. Sci.*, vol. 80, pp. 143–153, 2016.

[35] J. Tian, X. Zhuo, X. Yuan, C. Henry, P. d'Angelo, and T. Krauss, "Application oriented quality evaluation of Gaofen-7 optical stereo satellite imagery," *ISPRS Ann. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. V-1-2022, pp. 145–152, 2022.

[36] R. Liu, H. Zhang, K. H. A. Yip, J. Ling, Y. Lin, and H. Huang, "Automatic building height estimation with shadow correction over heterogeneous compact cities using stereo Gaofen-7 data at sub-meter resolution," *J. Building Eng.*, vol. 69, 2023, Art. no. 106283.

[37] C. Zeng, J. Wang, W. Zhan, P. Shi, and A. Gambles, "An elevation difference model for building height extraction from stereo-image-derived DSMs," *Int. J. Remote Sens.*, vol. 35, no. 22, pp. 7614–7630, 2014.

[38] E. Hussain, S. Kalaycioglu, C. W. D. Milliner, and Z. Cakir, "Preconditioning the 2023 Kahramanmaras (Turkiye) earthquake disaster," *Nature Rev. Earth Environ.*, vol. 4, no. 5, pp. 287–289, 2023.

[39] T. A. Neumann et al., "The ice, cloud, and land elevation satellite-2 mission: A global geolocated photon product derived from the advanced topographic laser altimeter system," *Remote Sens. Environ.*, vol. 233, Nov. 2019, Art. no. 111325.

[40] Z. Xiong and Y. Zhang, "Bundle adjustment with rational polynomial camera models based on generic method," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 1, pp. 190–202, Jan. 2011.

[41] P. Liao et al., "A linear pushbroom satellite image epipolar resampling method for digital surface model generation," *ISPRS J. Photogramm. Remote Sens.*, vol. 190, pp. 56–68, 2022.

[42] E. Rodriguez, C. S. Morris, and J. E. Belz, "A global assessment of the SRTM performance," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 3, pp. 249–260, 2006.

[43] D. Youssefi et al., "CARS: A photogrammetry pipeline using dask graphs to construct a global 3D model," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2020, pp. 453–456.

[44] J. Ma, X. Jiang, A. Fan, J. Jiang, and J. Yan, "Image matching from handcrafted to deep features: A survey," *Int. J. Comput. Vis.*, vol. 129, no. 1, pp. 23–79, 2020.

[45] C. De Franchis, E. Meinhardt-Llopis, J. Michel, J.-M. Morel, and G. Facciolo, "On stereo-rectification of pushbroom images," in *Proc. IEEE Int. Conf. Image Process.*, 2014, pp. 5447–5451.

[46] M. Rothermel, K. Wenzel, D. Fritsch, and N. Haala, "SURE: Photogrammetric surface reconstruction from imagery," in *Proc. LC3D Workshop*, 2012, vol. 8, no. 2.

[47] A. Li, W. Jiang, W. Yuan, D. Dai, S. Zhang, and Z. Wei, "An improved FAST+SURF fast matching algorithm," *Procedia Comput. Sci.*, vol. 107, pp. 306–312, 2017.

[48] D. Barath and L. Hajder, "A theory of point-wise homography estimation," *Pattern Recognit. Lett.*, vol. 94, pp. 7–14, 2017.

[49] D. Anguelov et al., "Google street view: Capturing the world at street level," *Computer*, vol. 43, no. 6, pp. 32–38, Jun. 2010.
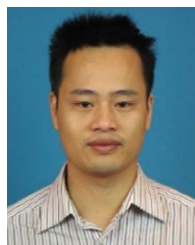
Zhonghua Hong (Member, IEEE) received the Ph.D. degree in geographic information system from Tongji University, Shanghai, China, in 2014.

He has been an Associate Professor with the College of Information Technology, Shanghai Ocean University, Shanghai, China, since 2022. His research interests include satellite/aerial photogrammetry, high-speed videogrammetry, planetary mapping, 3-D emergency mapping, GNSS-R, deep learning, and processing of geospatial big data.

Hongyang Zhang received the B.S. degree in computer science and technology from the School of Computer Science and Communication Engineering, Jiangsu University, Zhenjiang, China, in 2019. He is currently working toward the M.S. degree in computer science and technology with Shanghai Ocean University, Shanghai, China.

His research interests include satellite image 3-D reconstruction and computer geometry.

Xiaohua Tong (Senior Member, IEEE) received the Ph.D. degree in geographic information system from Tongji University, Shanghai, China, in 1999.

From 2001 to 2003, he was a Postdoctoral Researcher with the State Key Laboratory of Information Engineering in Surveying, Mapping, and Remote Sensing, Wuhan University, Wuhan, China. He was a Research Fellow with Hong Kong Polytechnic University, Hong Kong, in 2006, and a Visiting Scholar with the University of California, Santa Barbara, CA, USA, from 2008 to 2009. His research interests include photogrammetry and remote sensing, trust in spatial data, and image processing for high-resolution satellite images.

Shijie Liu (Member, IEEE) received the Ph.D. degree in cartography and geographic information engineering from Tongji University, Shanghai, China, in 2012.

He is currently a Professor with the College of Surveying and Geo-Informatics, Tongji University. His research interests include geometric exploitation of high-resolution remote sensing and its applications.

Ruyan Zhou received the Ph.D. degree in agricultural bioenvironment and energy engineering from Henan Agricultural University, Zhengzhou, China, in 2007.

From 2007 to 2008, she was with the Zhongyuan University of Technology. She is currently with Shanghai Ocean University, Shanghai, China.

Yanling Han received the B.E. degree in mechanical design and manufacturing and the M.E. degree in mechanical automation from Sichuan University, Chengdu, China, in 1996 and 1999, respectively, and the Ph.D. degree in engineering and control theory from Shanghai University, Shanghai, China, in 2005.

She is a Professor and currently working with Shanghai Ocean University, Shanghai, China. Her research interests include the study of ocean remote sensing, flexible system modeling, and deep learning.

Haiyan Pan received the Ph.D. degree in surveying and mapping from Tongji University, Shanghai, China, in 2020.

She is currently a Lecturer with the College of Information Technology, Shanghai Ocean University, Shanghai, China. Her research interests include multitemporal remote sensing data analysis, change detection, and multispectral/hyperspectral image classification.
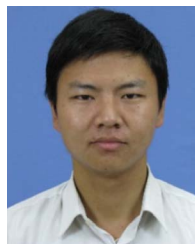

Jing Wang received the Ph.D. degree in biomedical engineering from the Department of Biomedical Engineering, Shanghai Jiaotong University, Shanghai, China, in 2014.

She has been a Lecturer with the College of Information Technology, Shanghai Ocean University, Shanghai, China, since 2015. Her research interests include computer vision and medical image processing.

Yun Zhang received the Ph.D. degree in applied marine environmental studies from the Tokyo University of Maritime Science and Technology, Tokyo, Japan, in 2008.

Since 2011, he has been a Professor with the College of Information and Technology, Shanghai Ocean University, Shanghai, China. His research interests include the study of navigation system reflection signal technique and its maritime application.

Shuhu Yang received the Ph.D. degree in physics of physics from the School of Physics, Nanjing University, Nanjing, China, in 2012.

Since 2012, he has been a Lecturer with the College of Information Technology, Shanghai Ocean University, Shanghai, China. His research interests include the evolution of the Antarctic ice sheet, hyperspectral remote sensing, and the use of navigational satellite reflections.