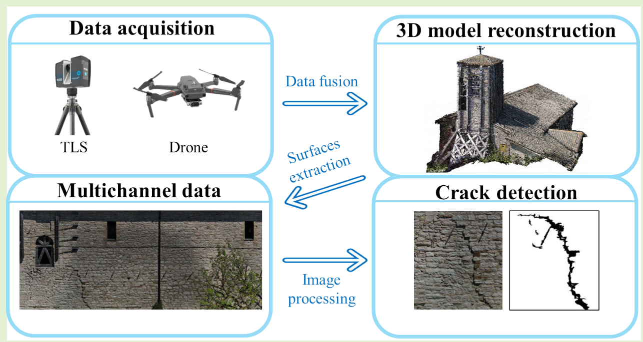
A Multisensor-Based Measurement Procedure for Seismic Damage Identification in Buildings

Giovanni Salerno, Maria Teresa Calcagni, Gloria Cosoli, Stefano Chiappini, Adriano Mancini, *Member, IEEE*, Alessandra Mobili, and Gian Marco Revel

***Abstract*—Effectively and precisely identifying damages after a seismic event is fundamental to ensuring timely intervention and optimizing the building life cycle. The combination of multisensor inspection and artificial intelligence (AI) technologies is powerful in this context. This work presents the results of the development of a multisensor measurement procedure relying on drone-embedded visible and thermal cameras and a terrestrial laser scanner (TLS) to acquire data on a masonry church damaged by an earthquake. These data were carefully aligned into a unified reference system and enabled the reconstruction of high-resolution 3-D models of the building, from which multichannel orthophotos were extracted, with visible, thermal, and depth data, to identify surface lesions. The analysis procedure involved three main stages: 1) preprocessing using principal component analysis (PCAs) to improve feature separability and reduce redundancy between channels; 2) lesion detection using a deep learning U-Net segmentation model trained to identify surface cracks; and 3) morphological postprocessing to refine the predicted masks and eliminate (or reduce) false positives predictions. The combination of multimodal data improves lesion detection, highlighting cracks that are not immediately visible, particularly when thermal gradients or geometric discontinuities provided complementary evidence. The results are promising and demonstrate both the feasibility of lesion identification and the validity of the proposed approach in the context of structural health monitoring. The use of sensors that can acquire data remotely and without contact offers significant advantages in terms of safety in a postearthquake context.**



***Index Terms*—3-D reconstruction, crack identification, damage identification, drone-based monitoring, masonry structures, measurement, structural health monitoring (SHM), terrestrial laser scanner (TLS).**

I. Introduction

CIVIL engineering structures are often exposed to multiple agents, such as loads, both operational and accidental, environmental degradation, and unexpected hazards, like earthquakes, floods, and fire, just to name a few. If structures are not adequately designed, the development of cracks and more serious damage can compromise the serviceability and safety of artifacts. Crack detection and mapping are pivotal, especially to evaluate the service condition accurately [1]. Routine inspections are essential to maintain structures/infrastructures' safety, and cracks assessment is imperative to predict and avoid their possible premature failure [2]. Traditional techniques used for cracks detection are often based on manual (mostly visual) methods, which are inefficient, labor-demanding, localized, and susceptible to errors and omissions [3]. This differs from the promptness and accuracy required in real-world engineering scenarios [4]. The timely identification of damages in buildings is fundamental to optimizing the building's life cycle as well as the needed interventions. Recent research has highlighted the importance of classifying seismic damage in reinforced concrete structures. For example, Ma et al. [5] proposed a probabilistic model based on Bayesian theory to identify the failure modes for bending, shear, and bending shear of reinforced concrete columns, achieving good accuracy in predicting collapse mechanisms under seismic loading, which is very important in the postearthquake context. Along the same lines, in [6], a method based on fragility curves was developed for reinforced concrete buildings with base isolation subjected to earthquakes near faults, demonstrating

Received 8 October 2025; revised 21 November 2025 and 23 November 2025; accepted 23 November 2025. Date of publication 4 December 2025; date of current version 14 January 2026. This work was supported by the Framework of the reCITY project "Resilient City–Everyday Revolution" (reCITY)–PON R&I 2014–2020 e FSC "Avviso per la presentazione di Ricerca Industriale e Sviluppo Sperimentale nelle 12 aree di Specializzazione individuate dal PNR 2015–2020," under Grant ARS01_00592. The associate editor coordinating the review of this article and approving it for publication was Dr. Wei Xu. *(Corresponding authors: Gloria Cosoli; Giovanni Salerno.)*

Please see the Acknowledgment section of this article for the author affiliations.

Digital Object Identifier 10.1109/JSEN.2025.3637312

how probabilistic seismic demand analysis can effectively capture the probability of structural damage. Indeed, De Sitter's law of Fives underlines the importance of adopting proper maintenance strategies on buildings to avoid a five-fold increase in costs for repair per decade [7]. In the last decades, many different technologies have been applied for damage identification, like visual inspection, crack gauges, vision systems, nondestructive testing (e.g., ultrasound-based sensors), analysis of vibrations, electrical impedance measurement, etc. In recent years, using drones for inspections in buildings, bridges, and other types of constructions has become common [8], thanks to the relatively fast, cheap, safe, and flexible acquisition of images [9]. The continuous evolution of digital tools, computer hardware, and intelligent algorithms has contributed to the affirmation in research and applied sciences of machine learning and deep learning crack detection methods [10]. Disposing of regular information on a building allows for proper designing of retrofitting operations and this is crucial for climate change resilience [11] as well as to timely act in consequence to extreme natural events such as earthquakes. They develop in seismic areas, such as the Apennine region, which spans across Italy. In this area, the most used building typology consists of masonry structures. Masonry is a heterogeneous material, commonly constituted of units and joints. There are several types of masonry; indeed, units can be made of natural stones or ceramic bricks, whereas joints can be either low-strength lime-based mortars or good-quality cementitious mortars. The emergence and development of cracks in masonry structures is often a precursor to accidents and greatly affects the performance of the structure itself [12]. The development of efficient, accurate, and large-scale inspection-oriented crack detection methods is paramount [13]. A newly automated measurement system based on the acquisition of red, green, blue - depth (RGB-D) image to localize and measure the geometric features of cracks has been used [14]. This tool, used previously for concrete structures, has been extended to masonry structures, which are characterized by a higher presence of nondesired elements, such as boundaries between bricks and mortar joints, which could be mistaken for masonry cracks and defects. Moreover, the same building typology (i.e., masonry) and a similar seismic risk of Italy can also be found in Greece, Cyprus, and Turkey. In this regard, not only is the local impact huge, but it can also be extended to many other Mediterranean areas. Additionally, many masonry artifacts belong to the cultural heritage, and hence it is essential that we all have a responsibility to preserve these monuments as long as possible [15]. However, at present, periodic inspections still prevail over regular monitoring strategies [16].

The seismic series that has affected the center of Italy from August 2014 to January 2017 was prolonged and destructive [17] and resulted in several structural failures and damages affecting different types of structures, e.g., reinforced concrete, masonry, and steel ones. The usage of remotely observed images for inspection purposes has been explored in literature and proved to be promising [18]. Gathering information from the combination of sensing and digital technologies can play a pivotal role in an enhanced design of resilient structures; open databases like the observed damage database (Da.D.O) [19] represent a valuable contribution to the research related to postearthquake inspections. In fact, these types of data can also be exploited to train artificial intelligence (AI) models for risk predictions that could also be fine-tuned and scaled to face different types of hazards (e.g., floods [20]). Objective data gathered on site are pivotal to make better choices in case of catastrophic events, also prioritizing interventions when needed, considering the different importance levels of structures and infrastructures. Multidomain sensors can provide synergies enabling to thoroughly depict the effective status of a certain structure [21]; in this context, this article concerns a multisensor framework providing multidomain data to exhaustively inspect a masonry structure struck by an earthquake. The building is the "San Pietro" church located in Pieve Torina, which is a small village in the center of Italy affected by the seismic swarm of summer 2016. A variety of technologies were employed for a more accurate and innovative reconstruction of the effective structural status of the building, going from drones equipped with visible and thermal cameras to terrestrial laser scanners (TLS) through data fusion strategies. The activities were performed in the framework of the reCITY project, aiming at developing different types of solutions to enhance resilience at the urban level. The analysis of the data collected on site were exploited for two different 3-D reconstructions of the damaged building, performed using the photogrammetry technique, one with visible images and one with thermal ones, aligned with TLS data for increased reconstruction accuracy. With the aim of recognizing lesions on surfaces, the information obtained after data processing with an AI-based algorithm can be useful to support decision-making processes in terms of reconstruction.

The contribution of this work beyond the state of the art can be summarized in the following points.

1) *Data Fusion:* The combined use of drones (with dual sensors) and TLS, as well as an in-depth processing pipeline including photogrammetry and advanced image analysis techniques, represents an improvement on the state of the art, which does not usually involve such in-depth analysis.
2) *Use of Multidomain AI:* AI algorithms are applied to multidomain data in an innovative way.
3) *Support in Intervention Management and Inspection Optimization:* especially after an extreme event, it is essential to prioritize interventions, and this study presents a method that could support this choice by exploiting remote inspections, which also guarantee a greater degree of safety for operators.

II. Materials and Methods

This section details the sensing equipment employed for data acquisition and describes the data processing pipeline aiming at the 3-D reconstruction of the building and the consequent damage identification, focusing on cracks.

A. Drone-Based Data Acquisition

In April 2024, a remote aerial survey was carried out on a masonry church in Pieve Torina (42°58'40.9" N

13°04’45.6" E). The drone flight was conducted at approximately 13:00 CET with clear sky and minimal wind, a condition specifically chosen to minimize the negative effects of shadows, which, when the sun is low, can change from the beginning to the end of the flight, and therefore, lead to difficulties in 3-D reconstruction. Obviously, the total absence of shadows (for example, on a cloudy day) would be preferable. The drone used was a DJI Mavic 2 Enterprise Advanced (M2EA, DJI, Shenzhen, China), equipped with both RGB and thermal cameras (Table I); both sensors are stabilized by a three-axis gimbal and synchronized at the factory so that each visible frame is matched with a thermogram.

The flight path was executed manually, and also due to the large size of the church, data acquisition was performed both in nadir mode and at different angles; oblique angles were manually changed between horizontal views (≈90° off-nadir) and intermediate angles of about 45°. The stand-off distance also varied from around 3 m to values above 10 m, both to better capture details of the building texture and to capture a large part of the building in one frame. The built-in timer triggered the shot every 3 s, resulting in consecutive images overlapping in the range 50%–80%. A total of 404 RGB frames and 404 corresponding thermal frames were collected. The visible camera acquired at full resolution (8000 × 6000) and automatic exposure and focus were maintained throughout the flight, while the thermal camera operated with default emissivity ($\varepsilon_r = 1$); radiometric R-JPEGs were subsequently converted to 16-bit GeoTIFF to preserve temperature information.

The masonry church was additionally surveyed using the FARO Focus 3-D MS 120 TLS (FARO Technologies Inc., Lake Mary, FL, USA) [22]. This instrument employs a phase shift distance measurement principle at a wavelength of 905 nm, operates over a range of 0.60–120 m, and delivers an accuracy of ±2 mm under standard lighting conditions. Its field of view spans 320° vertically and 360° horizontally, with an angular resolution of 0.009°. With an acquisition rate of 976 000 pts/s and a beam divergence of 0.19 mrad, it generates high-density point clouds in minimal time. These capabilities, combined with its compact form factor (5.2 kg; 240 × 200 × 100 mm) and IP54 protection rating, ensure reliable operation in confined spaces and under variable environmental conditions. The point-cloud density of the FARO Focus 3-D MS 120 was configured prior to each acquisition via the scanner control panel.

The selected density and accuracy settings directly affect scan duration: increasing acquisition time typically yields higher-quality point clouds with reduced noise. A network of reference targets was deployed on the church façades to guarantee a minimum of three common targets between consecutive scans.

The combined use of TLS and remotely sensed images significantly reduced the time (compared to traditional surveys) to obtain a 3-D “copy” of the church.

The choice of sensors is due to the need of extracting different characteristics from the data, which can often be complementary and allow for a more complete analysis of the structure under examination. The thermal camera, in particular, is able to highlight certain lesions in a different and sometimes better way than the visible camera. Furthermore, for this study, the sensors used are complementary in the data acquisition phase, since the TLS is certainly limited by the type of structure and what is around it, it is unable to acquire data from parts of the building hidden by others, or it may have difficulty detecting the highest parts of some buildings. The opposite is valid for drones, as they are usually freer to move around the structure and acquire data very well in the highest parts of the structure, whereas they may struggle more with the lower parts.

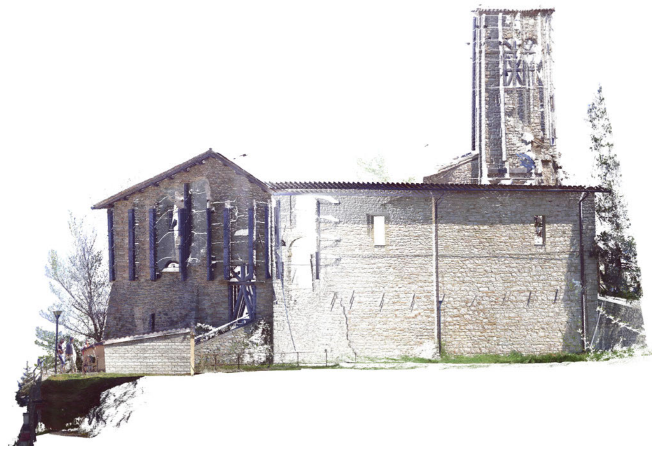


Fig. 1. Point cloud of TLS (FARO Focus 3-D) block.

Furthermore, these types of sensors allow data to be acquired remotely, particularly for drones, which could be very important in the context of postearthquake inspections from a safety perspective.

B. 3-D Model Reconstruction

The raw TLS dataset was imported into Autodesk ReCap v.23.1.000 [23], manually aligned, and then cleaned from noise and outliers using predefined distance and angle thresholds. In this initial processing stage, no automatic filters were applied, resulting in an E57 file of ≈2.2 × 108 points. Following segmentation to isolate the study area, the point cloud was reduced to ≈9.3 × 108 points, each annotated with intensity and RGB values derived from the scanner-integrated imagery. The postprocessed cloud is shown in Fig. 1.

A separate photogrammetric workflow was carried out for the RGB images and the co-registered thermal frames acquired with the drone. The processing was performed with Agisoft Metashape v 2.1.2 [24] on a workstation equipped with an Intel i9-13900K CPU (32 threads), 128 GB RAM, and an NVIDIA RTX 4090 GPU. The first step was the alignment of the images, which was performed at medium quality with predefined pair preselection. The key point and tie point limits were set to 60 000 and 6000 for the visible block and to 40 000 and 1000 for the thermal block.

The procedure generated 556 676 tie points for RGB analysis and 171 652 tie points for the IR one. Five geometric control markers (ground control points) extracted from the TLS survey were then placed on the images and manually cross-checked on them; other markers (13 check points) were

Fig. 2. Point cloud of visible (RGB) block.

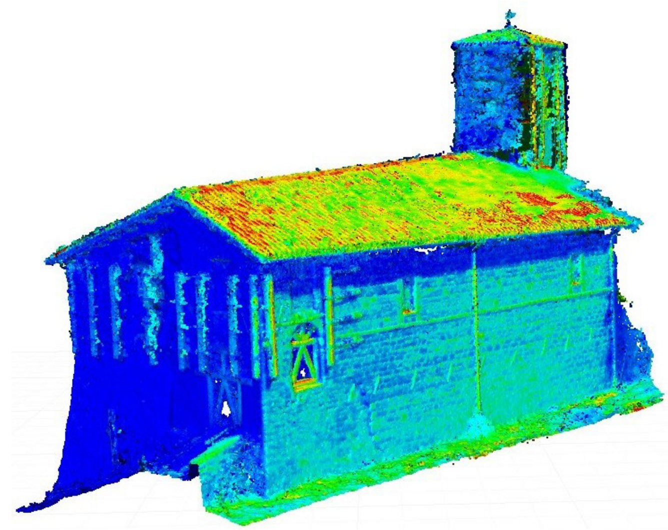


Fig. 3. Point cloud of thermal (IR) block.

also placed manually on the model and checked (and adjusted on the images, if necessary) to improve accuracy. These coincided with key points of the building that were easily recognizable as certain edges. To assess the quality of the reconstruction, the values of the residuals were observed; the control points residuals were 0.039 m (1.9 px) for RGB block and 0.075 m (1.2 px) for IR one, and the check points residuals were 0.061 m (1.8 px) for RGB block and 0.076 m (1.8 px) for IR one, keeping the overall reprojection error within acceptable values considering the actual building size.

Depth maps were processed with medium and ultrahigh quality for the RGB and IR block, respectively, and with the mild filter for both. The dense cloud phase was performed from the depth maps at high quality for the RGB block and ultrahigh for the IR block, resulting in ≈ 1.80 × 108 points (RGB) and 5.76 × 106 points (IR), after manual cleaning of noninterest points (i.e., those furthest from the building).

The differences between qualities used in different blocks were determined by the computational power required. Figs. 2 and 3 show the point clouds obtained from the two blocks, i.e., visible and thermal, after the overall processing.

After the co-registration of the two point cloud blocks (RGB, IR) in a common local reference system defined by the markers exported from the TLS point cloud, the point clouds of the RGB and IR blocks were manually segmented in Agisoft Metashape to isolate the façades of interest. The north, south (including the bell tower), and east façades were retained and subsequently analyzed; the west façade was discarded because dense vegetation obscured most of its surface. For each façade of interest, two independent 3-D surface models were derived, one from the RGB dense cloud and one from the IR dense cloud.

Meshes were generated from the depth maps; triangular meshes were constructed with the high-quality preset and hole filling (interpolation) enabled, resulting in highly detailed geometric representations.

To obtain the orthophotos of the façades, a plane was defined for each of them through a series of markers manually positioned (minimum three per façade), some coinciding with previously defined markers.

The orthophotos were then constructed from the 3-D models, disabling the interpolation of the holes. Pixel sizes were defined at the maximum resolution available, i.e., 1.67 mm/px (RGB) and 1.54 cm/px (IR), resulting in orthophotos with dimensions of 7475 × 7257 px (north), 10 795 × 9798 px (south), and 13 115 × 6864 px (east) for the visible block, and 750 × 713 px, 1044 × 912 px, and 1530 × 676 px, respectively, for the thermal block. RGB and thermal orthophotos were exported as 32-bit GeoTIFF; the visible ones in three bands (R, G, B) and the thermal ones with storage of measured temperature (°C).

To also have a depth reference, digital elevation models (DEMs) of each façade of the visible block were produced. They were generated with point clouds as the data source and with interpolation disabled to avoid artifacts. For the definition of the reference plane, the same markers were used as for the generation of the corresponding orthophotos.

The resolution was set as high as possible, resulting in 3.07 mm/px and images of 4021 × 3762 px (north), 5715 × 5364 px (south), and 7385 × 3342 px (east). The DEMs were exported in 32-bit GeoTiff format.

The facade orthophotos obtained via RGB and IR cameras, and the DEM obtained from the RGB camera, were merged into a single five-band raster using QGIS's Build Virtual Raster tool v 3.44.1 [25]. The output is multichannel data in TIFF format with consistent spatial resolution and coverage for further analysis.

C. Crack Identification

The reconstruction of 3-D models based on the combination of several sensors resulted in five-channel images containing depth (DEM), temperature (IR), and visible (R, G, B) information. These data were the input for the next stage of processing pipeline (depicted in Fig. 4), i.e., the inspection for cracks/structural fractures with AI-based algorithm.

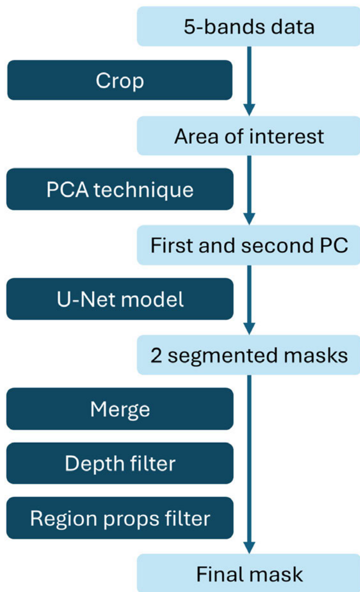


Fig. 4. Overall pipeline of the crack identification.

The algorithm involves a preprocessing step, the application of an AI model for crack detection and a postprocessing to refine the resulting mask.

The analysis was focused on three specific areas of the façades of interest as, due to their large size, it was not possible to analyse entire images as a single piece of data. The preprocessing phase included the manual cropping of the zone of interest, a normalization, and the application of the principal component analysis (PCAs) algorithm, reducing the dimensionality of a dataset (from five to three in this case) by transforming a set of correlated variables into a new set of uncorrelated variables, i.e., the principal components (PCs) [26]. The normalization was executed between the minimum and maximum values present for depth and thermal bands to transform them into greyscale images and allow subsequent AI-based analysis, and to fill the image with missing data (for the DEM and IR data there are no-data pixels). Hence, the AI model was applied to identify the cracks. Specifically, the model was U-Net [27], a semantic segmentation model trained with 750 manually annotated images of cracks on concrete and brick walls, collected both from different buildings with damage and online, additional details on the model used are available on [28] and [29]. The dataset was split into training, validation, and test subsets according to the ratio: 80:10:10; the training lasts 130 epochs, and the loss function was defined by summing Dice Loss and Focal Loss, obtaining a final value of 0.135. The metric used to evaluate the model (on the test sub-set) was the intersection over union (IoU) score and the value achieved was 0.991.

Inference with the model is performed by taking the image in consideration as input (resized according to computational capabilities) and providing as output a mask with a confidence value for each pixel (0 to 1), and these masks were binarized using the threshold set at 0.5. In order to avoid any possible recognition mistakes by the U-Net model due to the fact that the dataset consists mainly of images of cracks on concrete surfaces or more or less regular brick walls, significant post-processing was carried out.

The first two components obtained from the PCA were segmented using the U-Net model and the two binary masks were subjected to a morphological denoising operation, consisting of erosion followed by dilation, implemented using the OpenCV library [30] to reduce small-scale noise, and then merged; this was done to include a greater variance in the identification of cracks than using RGB images alone.

Then, the postprocessing started using the data obtained from the depth band (DEM) to filter out plane portions in the segmented mask (false positives) through a depth-based filter. Indeed, in the collected data, there are wooden beams, ties, and a canopy that can be confused with cracks due to color and shape. The filter takes the depth data as input, divides it into patches with a width of 500 pixels (experimentally determined value; height does not matter), calculates the average depth of the patch, and removes all the pixels with a minimum difference of 0.15 m from the average value by creating a mask (definitely a filter applicable only on vertical surfaces). Areas smaller than 40 pixels are removed from the mask to avoid noise. Finally, this depth-band-based mask was applied to the mask obtained from U-Net. A final filtering step based on region properties was applied to the resulting binary mask, relying on the scikit-image Python library [31]. The regions were filtered according to several criteria: 1) a minimum area and perimeter threshold to remove small artifacts; 2) aspect ratio constraints to discard long and narrow regions that are likely to represent nonstructural elements, such as shadows or edges; and 3) an eccentricity threshold to exclude highly elongated shapes. The resulting mask retained only plausible fracture regions, eliminating false positives and improving the accuracy of the structural analysis.

In order to conclude the analysis with quantitative evaluation metrics, the three areas of interest were also analyzed by an expert in the field, who provided a manually drawn segmentation mask. This mask was considered as the ground truth, and the traditional AI model metrics were calculated for the three areas of interest: precision, recall, F1-score, IoU score, and accuracy. The metrics were calculated by considering the individual labeled pixels as true/false positive/negative (T/F P/N).

III. Results and Discussion

A. 3-D Model

Figs. 5–7 show orthophoto and DEM from the RGB block and the orthophoto from the thermal block (thermal and depth data normalized) for the three façades analyzed; for the thermal orthophoto and the DEM of the south façade, the background is also shown (the lowest values – i.e., the same used for normalization—were assigned for visualization purposes).

For thermal and depth data, the figures show a normalized, greyscale visualization of the data, which does not allow small variations to be visually appreciated, especially in the case of the DEM; nevertheless, these variations are actually present in the data.

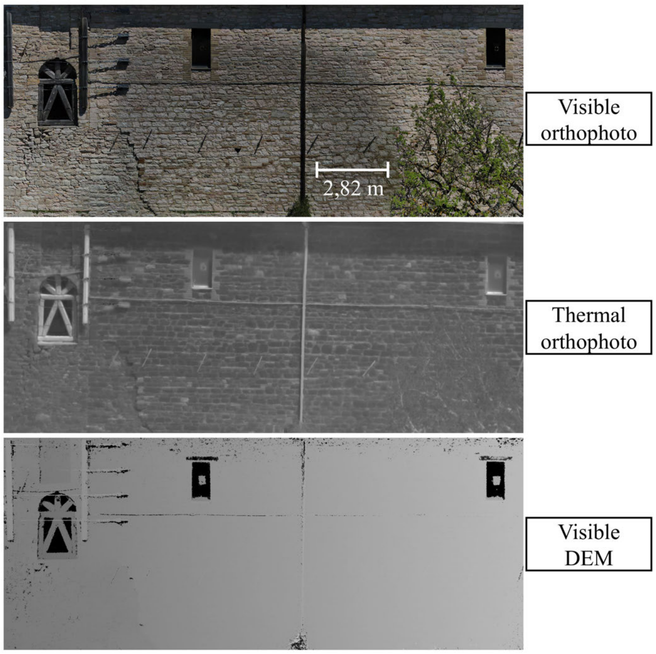


Fig. 5. East façade: visible orthophotos and DEM, and thermal orthophoto.

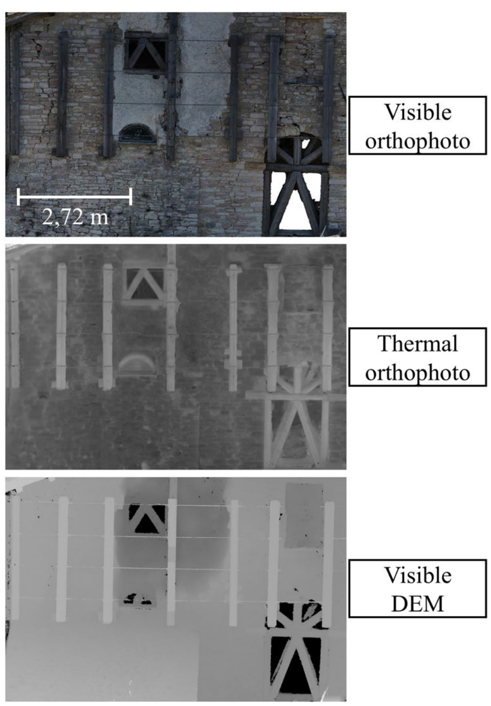


Fig. 6. North façade: visible orthophotos and DEM, and thermal orthophoto.

B. Crack Identification Results

From the orthophotos obtained from the 3-D reconstruction of the façades, the three areas of interest were extracted, shown in Fig. 8, where the structural damages were most severe. The PCA applied to the three areas has produced the images shown in Figs. 9 (top), 10 (left), and 11 (top), as first and second PCs.

The variance explained by the first two components, for the three areas of interest, is 96.3% and 1.9%, 77.1% and 15.2%,

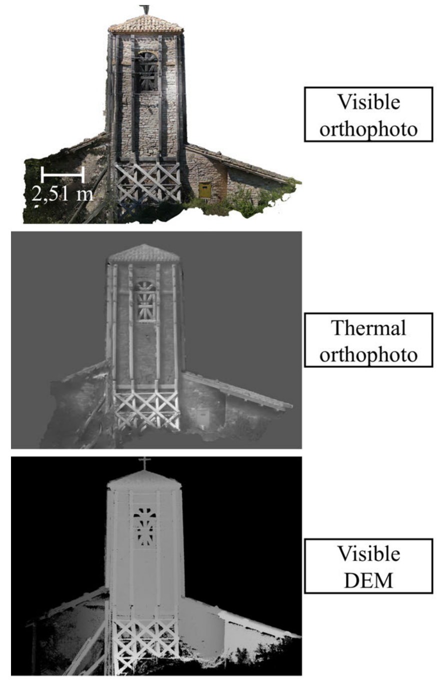


Fig. 7. South façade: visible orthophotos and DEM, and thermal orthophoto.

Fig. 8. Three areas of interest extracted from orthophotos: on the east façade (left) and on the north façade (center and right).

60.4% and 29.9%, respectively. Table II lists the corresponding loads, i.e., the coefficients of the linear combination of the five input bands. Positive and negative signs indicate a direct or inverse correlation, respectively, whereas the magnitude reflects the relative contribution of each band. The first component (PC1) explains the greatest part of the variance of the data, represented essentially in terms of appearance and morphology (color and texture); in fact, it is mainly composed of the visible bands (coefficient modulus). On the other hand, the second component (PC2) can better explain the variance of thermal or geometric anomalies (in terms of depth), and mainly consists of the thermal and depth bands, which are less heterogeneous than the visible band. It is worthy to note that the sensor in the visible range has a better resolution than the thermal; additionally, the depth band provides more homogeneous values. This separation reflects the high correlation of the visible bands and the orthogonal contribution of thermal and depth data, which is essential for highlighting cracks.

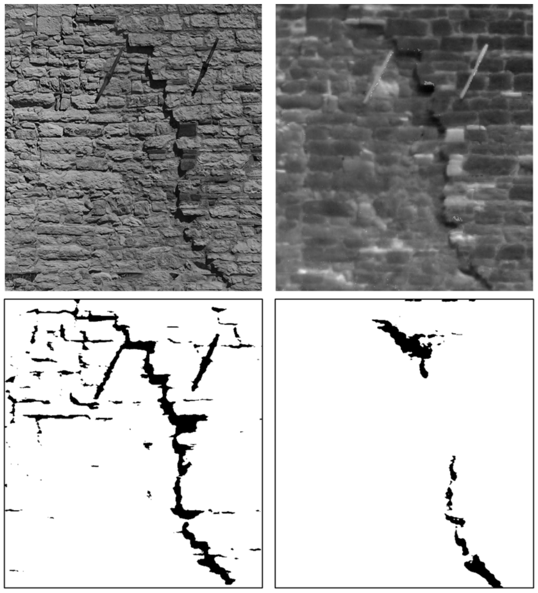


Fig. 9. PCs number 1 and 2 (top) of the first area of interest, and related segmented masks (bottom).

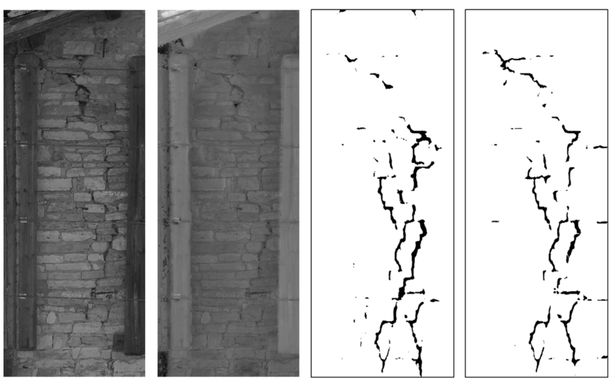


Fig. 10. PCs number 1 and 2 (left) of the second area of interest, and related segmented masks (right).

TABLE I
SPECIFICATIONS OF DRONE, RGB, AND THERMAL CAMERA AND LENS

Drone	DJI Mavic 2 Enterprise Advanced (M2EA)
RGB camera sensor	1/2.3" CMOS, Effective pixels: 48M, 8000x6000 px
RGB Lens	FOV: approx. 84°, 35 mm format equivalent: 24 mm, Aperture: f/2.8 Focus: 0.5 m to ∞
Thermal camera sensor	Uncooled VOx Microbolometer, Spectral band: 8-14 µm, Resolution: 640 x 512.
Thermal lens	Focal length: ≈ 9 mm, 35 mm format equivalent: ± 38 mm
Accuracy of thermal temperature	±2 °C or ±2%

The masks segmented by the U-Net model from the two PCs of the three areas of interest are shown in the same figures, i.e., Figs. 9 (bottom), 10 (right), and 11 (bottom).

The two masks obtained, for each area, were then merged to comprehend the variability of the two PCs; the results are

TABLE II
PCA ANALYSIS: COEFFICIENTS OF THE LINEAR COMBINATION OF THE FIVE INPUT BANDS

Area	Component	Depth	IR	Red	Green	Blue
1	1	-0.007	0.002	0.585	0.583	0.563
	2	-0.208	0.970	0.072	0.017	-0.098
2	1	-0.074	-0.328	0.541	0.540	0.550
	2	0.546	0.776	0.150	0.181	0.211
3	1	-0.089	-0.204	0.567	0.569	0.553
	2	-0.493	-0.829	-0.027	-0.119	-0.235

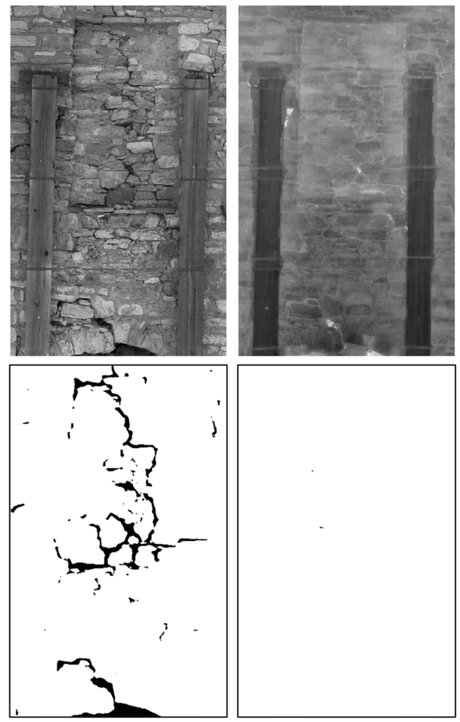


Fig. 11. PCs number 1 and 2 (top) of the third area of interest, and related segmented masks (bottom).

shown in Fig. 12. The combination of the two masks shows a better generalization in identification, in fact in the case of area number 1 the mask of the first component is more complete and better identifies the lesions, whereas in the case of area number 2 the second component shows a more continuous identification in the upper part of the image (the opposite at the bottom). Concerning the third area, the identification in the second component is almost void, and the first two PCs show a smaller explained variance, i.e., 90.3%, compared to 98.2% and 92.3% in the other two cases.

The application of the depth-based filter led to the removal of areas of out-of-plane wall; Fig. 13 compares the upper part of the north façade before and after the filter application.

For each façade, the white areas were excluded from the merged masks previously obtained, being identified as non-crack pixels due to their distance from the mean depth plane, and the resulting masks for the three areas of interest are shown in Fig. 14. On the newly produced masks, the final filter based on region properties was applied, as explained in Section II-C.

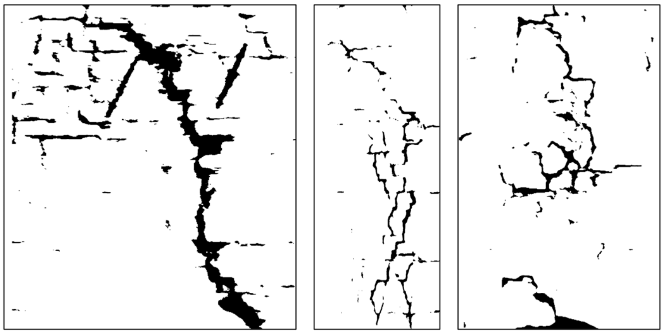


Fig. 12. Merged masks of two PCs for the three areas of interest.

Fig. 13. Upper part of the north façade (visible bands, top), and the same after the application of the depth-based filter (bottom).

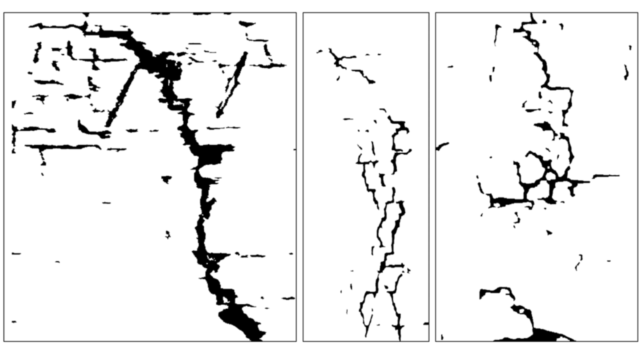


Fig. 14. Resulting masks after depth-based filter.

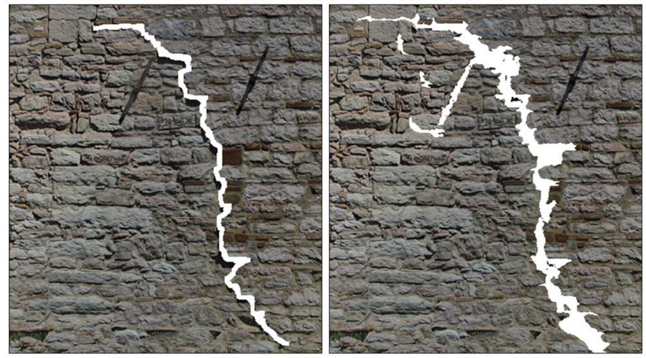


Fig. 15. Overlaps of manually produced mask (left) and algorithm mask (right) on visible image of first area of interest.

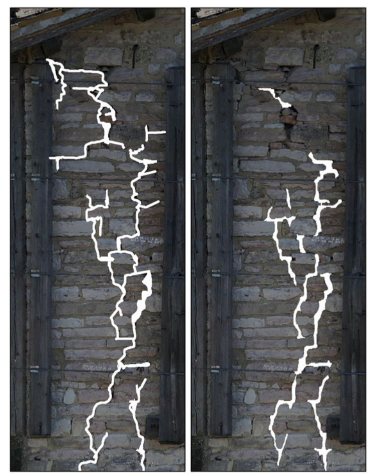


Fig. 16. Overlaps of manually produced mask (left) and algorithm mask (right) on visible image of second area of interest.

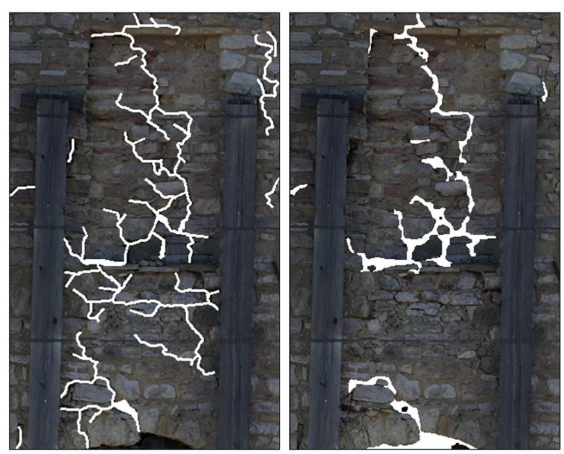


Fig. 17. Overlaps of manually produced mask (left) and algorithm mask (right) on visible image of the third area of interest.

The final masks obtained after the application of the entire algorithm were compared to masks manually produced by an expert on the field, in Figs. 15–17 the overlaps of the two types of masks (algorithm on the right and manual on the left) on the visible image of the area of interest are shown. As can be seen from the figures, the masks are certainly similar, but there are some differences.

The algorithm proves to be more conservative, producing reduced masks compared to the ground truths, particularly for the third area. From the first area, however, it can be noticed that the algorithm is misled by shadows, which the AI model often identifies as cracks, as they are basically the same color. Furthermore, the algorithm struggles to recognize very thin cracks, which can be observed at the top of the second area and in many parts of the third area.

TABLE III
COMPARISON RESULTS ON THE THREE AREAS OF INTEREST: EVALUATION METRICS

Area	Precision	Recall	F1-score	Accuracy
1	0.35	0.90	0.51	0.95
2	0.61	0.41	0.49	0.96
3	0.42	0.31	0.36	0.94
Mean	0.46	0.54	0.45	0.95

Finally, the evaluation metrics mentioned in Section II-C were calculated for each mask produced by the algorithm. Table III shows the results of the metrics for the three areas of interest.

The metrics confirm the difficulties in distinguishing shadows (or generally very dark colors) and thin cracks. In fact, for the first area (full of shadows), there is a very high recall and lower precision, which means that the algorithm produced few FNs and many FPs, mainly due to the presence of shadows that confused the model, causing it to produce a wider mask; and the same occurs in the lower part of the third area.

To reduce the problem of shadows, the method proposed could be used on cloudy days. Accuracy is the highest metric, as it considers both TP and TN, and backgrounds constitute a large part of the images; hence, it is very high. Looking at the metrics for area number 3, the difficulties with thin cracks can be highlighted (in fact, many of these were not recognized), but also the masks often produced were thicker for large lesions.

This is a fundamental aspect because, although the algorithm may not recognize some thin cracks, which may be less dangerous, it certainly highlights the thicker ones more decisively, which are generally more dangerous. Certainly, a better prediction of fine cracks could be obtained by using higher resolution images, which could be obtained if captured from closer. In this case, having to reconstruct entire building façades, the resolution decreases, and their detection becomes less accurate.

IV. CONCLUSION

This article presented the results of the work carried out within the framework of the reCITY project concerning the development of a multisensor measurement procedure for damage identification based on data fused from cameras embedded on a drone and data from a TLS. 3-D models of the surveyed building (a masonry church) were obtained by extracting multichannel orthophotos, where damage (in particular, superficial lesions) was properly identified using an AI-based algorithm with preprocessing and postprocessing. The results demonstrate the validity of the proposed approach and underline the strength of combining sensing technologies with AI-based models to support the management of structures and infrastructures in the case of natural extreme events like earthquakes. It is worth noting that the method is valid and useful as a support to operators, not as a total replacement for them; from this perspective, this approach could be used remotely (the distance for the drone depends on the drone itself but may be hundreds of meters, whereas the TLS needs to be closer to the building but still tens of meters away), which is essential from a safety point of view, to provide key information to operators who can then go into the field only where necessary. The approach could also exploit a second acquisition of the drone only in the areas most affected by damage, taking closer images that could help in identifying the finest cracks, a very interesting potential future work. The time required is reduced compared to traditional visual inspection (∼30–60 and ∼60 min for drone and TLS acquisitions) when considering a single building, but this is even more valid when considering a larger scale, since a single acquisition (especially for TLS) produces valid data for multiple buildings. Although data processing is time-consuming (considering the entire pipeline, a few hours per building), it can be automated and parallelized, running on remote machines 24 h a day. The cost of sensors is not usually very low, especially for TLS, but in the long term, it is quite reasonable.

The procedure has therefore proved to be efficient in terms of time for both data acquisition and processing and can be easily scaled and fine-tuned to different application fields in support to municipalities and stakeholders needing to prioritize interventions based on objective assessment of actual conditions, especially after massive destructive events.

ACKNOWLEDGMENT

Giovanni Salerno, Maria Teresa Calcagni, and Gian Marco Revel are with the Department of Industrial Engineering and Mathematical Sciences, Università Politecnica delle Marche, 60131 Ancona, Italy (e-mail: g.salerno@pm.univpm.it; m.t.calcagni@pm.univpm.it; gm.revel@staff.univpm.it).

Gloria Cosoli is with the Department of Theoretical and Applied Sciences, eCampus University, 22060 Novedrate, Italy (e-mail: gloria.cosoli@uniecampus.it).

Stefano Chiappini is with the Department of Construction, Civil Engineering and Architecture, Università Politecnica delle Marche, 60131 Ancona, Italy (e-mail: s.chiappini@staff.univpm.it).

Adriano Mancini is with the Department of Information Engineering, Università Politecnica delle Marche, 60131 Ancona, Italy (e-mail: a.mancini@staff.univpm.it).

Alessandra Mobili is with the Department of Science and Engineering of Matter, Environment and Urban Planning, Università Politecnica delle Marche, 60131 Ancona, Italy (e-mail: a.mobili@staff.univpm.it).

REFERENCES

[1] J.-L. Xiao, J.-S. Fan, Y.-F. Liu, B.-L. Li, and J.-G. Nie, "Region of interest (ROI) extraction and crack detection for UAV-based bridge inspection using point cloud segmentation and 3D-to-2D projection," *Autom. Construct.*, vol. 158, Feb. 2024, Art. no. 105226, doi: 10.1016/j.autcon.2023.105226.

[2] H. Zoubir, M. Rguig, M. E. Aroussi, R. Saadane, and A. Chehri, "Pixel-level concrete bridge crack detection using convolutional neural networks, Gabor filters, and attention mechanisms," *Eng. Struct.*, vol. 314, Sep. 2024, Art. no. 118343, doi: 10.1016/j.engstruct.2024.118343.

[3] C. Fan, Y. Ding, X. Liu, and K. Yang, "A review of crack research in concrete structures based on data-driven and intelligent algorithms," *Structures*, vol. 75, May 2025, Art. no. 108800, doi: 10.1016/j.istruc.2025.108800.

[4] Y. Liu, J. Fan, J. Nie, S. Kong, and Y. Qi, "Review and prospect of digital-image-based crack detection of structure surface," *Tumu Gongcheng Xuebao/China Civil Eng. J.*, vol. 54, pp. 79–98, Jan. 2021.

[5] Y. Ma, D. Wang, H. Cheng, and Y. Shi, "Bayesian theory-based seismic failure modes identification of reinforced concrete columns," *J. Earthq. Eng.*, vol. 26, no. 13, pp. 6703–6723, Oct. 2022. [Online]. Available: https://www.tandfonline.com/doi/abs/10.1080/13632469.2021.1927905

[6] C. Liu, D. Fang, and Z. Yan, "Seismic fragility analysis of base isolated structure subjected to near-fault ground motions," *Periodica Polytechnica Civil Eng.*, vol. 65, no. 3, pp. 768–783, Mar. 2021, doi: 10.3311/ppci.15276.

[7] W. R. de Sitter, "Costs of service life optimization 'The law of fives,'" in *CEB-RILEM Workshop on Durability of Concrete Structures*. Copenhagen, Denmark: Comité Euro-International du Béton, May 1984, pp. 131–134.

[8] T. Rakha and A. Gorodetsky, "Review of unmanned aerial system (UAS) applications in the built environment: Towards automated building inspection procedures using drones," *Autom. Construct.*, vol. 93, pp. 252–264, Sep. 2018, doi: 10.1016/j.autcon.2018.05.002.

[9] D. Kang and Y.-J. Cha, "Autonomous UAVs for structural health monitoring using deep learning and an ultrasonic beacon system with geo-tagging," *Comput.-Aided Civil Infrastruct. Eng.*, vol. 33, no. 10, pp. 885–902, Oct. 2018, doi: 10.1111/mice.12375.

[10] H. Yu, W. Yang, H. Zhang, and W. He, "A UAV-based crack inspection system for concrete bridge monitoring," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2017, pp. 3305–3308, doi: 10.1109/IGARSS.2017.8127704.

[11] A. Keshmiry, S. Hassani, U. Dackermann, and J. Li, "Assessment, repair, and retrofitting of masonry structures: A comprehensive review," *Construct. Building Mater.*, vol. 442, Sep. 2024, Art. no. 137380, doi: 10.1016/j.conbuildmat.2024.137380.

[12] B. Liu, V. Sarhosis, and J. V. Lemos, "Quantification of the crack propagation and global failure mechanism of single- and multi-ring masonry arch bridges," *Eng. Struct.*, vol. 306, May 2024, Art. no. 117805, doi: 10.1016/j.engstruct.2024.117805.

[13] Y. Zhou, Y. Wang, R. Wang, W. Zhang, and H. Yang, "Automatic crack detection and segmentation of masonry structure based on deep learning network and edge detection," *Structures*, vol. 76, Jun. 2025, Art. no. 108850, doi: 10.1016/j.istruc.2025.108850.

[14] N. Giulietti, P. Chiariotti, and G. M. Revel, "Automated measurement of geometric features in curvilinear structures exploiting Steger's algorithm," *Sensors*, vol. 23, no. 8, p. 4023, Apr. 2023, doi: 10.3390/s23084023.

[15] A. Soleymani, H. Jahangir, and M. L. Nehdi, "Damage detection and monitoring in heritage masonry structures: Systematic review," *Construct. Building Mater.*, vol. 397, Sep. 2023, Art. no. 132402, doi: 10.1016/j.conbuildmat.2023.132402.

[16] N. Giulietti et al., "Continuous monitoring of the health status of cement-based structures: Electrical impedance measurements and remote monitoring solutions," *ACTA IMEKO*, vol. 10, no. 4, p. 132, Dec. 2021, doi: 10.21014/acta_imeko.v10i4.1140.

[17] A. Masi, L. Chiauzzi, G. Santarsiero, M. Liuzzi, and V. Tramutoli, "Seismic damage recognition based on field survey and remote sensing: General remarks and examples from the 2016 central Italy earthquake," *Natural Hazards*, vol. 86, no. S1, pp. 193–195, Mar. 2017, doi: 10.1007/s11069-017-2776-8.

[18] R. Nascimbene, "Investigation of seismic damage to existing buildings by using remotely observed images," *Eng. Failure Anal.*, vol. 161, Jul. 2024, Art. no. 108282, doi: 10.1016/j.engfailanal.2024.108282.

[19] M. Dolce et al., "Observed damage database of past Italian earthquakes: The da. Do WebGIS," *Bollettino di Geofisica Teorica ed Applicata*, vol. 60, no. 2, pp. 141–164, 2019.

[20] R. Bentivoglio, E. Isufi, S. N. Jonkman, and R. Taormina, "Deep learning methods for flood mapping: A review of existing applications and future research directions," *Hydrol. Earth Syst. Sci.*, vol. 26, no. 16, pp. 4345–4378, Aug. 2022, doi: 10.5194/hess-26-4345-2022.

[21] K. Aminfar et al., "Data fusion in structural health monitoring: Integrating computer vision and strain gauges for enhanced inspection," in *Bridge Maintenance, Safety, Management, Digitalization and Sustainability*, 1st ed., London, U.K.: CRC Press, 2024, pp. 709–716, doi: 10.1201/9781003483755-82.

[22] FARO Knowledge Base. *Performance Specifications for the Focus3D*. Accessed: Jul. 17, 2025. [Online]. Available: https://knowledge.faro.com/Hardware/Focus/Focus/Performance_Specifications_for_the_Focus3D

[23] *Autodesk ReCap Pro Software | Get Prices & Buy Official ReCap Pro 2026*. Accessed: Jul. 23, 2025. [Online]. Available: https://www.autodesk.com/products/recap/overview

[24] *Agisoft Metashape*. Accessed: Jul. 17, 2025. [Online]. Available: https://www.agisoftmetashape.com/

[25] *7.1. Lesson: Working With Raster Data—QGIS Documentation Documentation*. Accessed: Jul. 17, 2025. [Online]. Available: https://docs.qgis.org/3.40/en/docs/training_manual/rasters/data_manipulation.html

[26] H. Abdi and L. J. Williams, "Principal component analysis," *WIREs Comput. Statist.*, vol. 2, no. 4, pp. 433–459, Jul. 2010, doi: 10.1002/wics.101.

[27] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015* (Lecture Notes in Computer Science). Cham, Switzerland: Springer, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[28] G. Salerno, M. T. Calcagni, M. Martarelli, and G. M. Revel, "Metrological evaluation of an AI-based vision computing model for crack detection on masonry structures," in *Proc. MATEC Web Conf.*, vol. 403, 2024, p. 4002, doi: 10.1051/matecconf/202440304002.

[29] M. T. Calcagni et al., "Concrete defect identification and measurement in buildings," *Acta IMEKO*, vol. 14, no. 3, pp. 1–11, Sep. 2025, doi: 10.21014/actaimeko.v14i3.1965.

[30] *OpenCV*. Accessed: Jul. 15, 2025. [Online]. Available: https://opencv.org/

[31] S. van der Walt et al., "Scikit-image: Image processing in Python," *PeerJ*, vol. 2, p. e453, Jun. 2014, doi: 10.7717/peerj.453.

Giovanni Salerno was born in Foggia, Italy, in 1997. He received the B.S. and M.S. degrees in mechanical engineering from the Università Politecnica delle Marche, Ancona, Italy, in 2019 and 2022, respectively, where he is currently pursuing the Ph.D. degree in industrial engineering, with a focus on mechanical and thermal measurement.

His research interests include the use of noncontact measurement techniques, especially vision-based techniques (visible, thermal, hyper-spectral) applied in construction and industrial sectors.

Maria Teresa Calcagni was born in Ascoli Piceno, Italy, in 1995. She received the B.S. and M.S. degrees in mechanical engineering and the Ph.D. degree in industrial engineering from Università Politecnica delle Marche (UNIVPM), Ancona, Italy, in 2018, 2020, and 2025, respectively.

She is currently a Postdoctoral Research Fellow with the Department of Industrial Engineering and Mathematical Sciences (DIISM), UNIVPM, with a focus on mechanical and thermal measurements. Her research interests include noncontact measurement techniques, especially vision-based techniques (visible, thermal, hyper-spectral) applied in construction and industrial sectors.

Gloria Cosoli was born in Chiaravalle, Ancona, Italy, in 1989. She received the B.S. degree in biomedical engineering, the M.S. degree in electronic engineering, and the Ph.D. degree in mechanical engineering from Università Politecnica delle Marche (UNIVPM), Ancona, Italy, in 2011, 2013, and 2017, respectively.

From 2016 to 2023, she was a Postdoctoral Research Fellow with the Department of Industrial Engineering and Mathematical Sciences (DIISM), UNIVPM. From 2023 to August 2025, she was a Research Collaborator at DIISM, UNIVPM; since May 2024, she has been an Associate Professor with the Department of Theoretical and Applied Sciences (DiSTA), eCampus University, Novedrate, Italy. She is the author of 40 articles, 60 conference proceedings, and two national patents. Her research interests include noninvasive physiological measurements, numerical modeling, mechanical measurements, signal processing, and non-destructive testing (NDT).

Prof. Cosoli received the IEEE MeMeA 2015 Best Poster Award and the IEEE MetroLivEnv 2023 Best Paper Award.

Stefano Chiappini was born in San Benedetto del Tronto, Ascoli Piceno, Italy, in 1990. He received the bachelor's and master's degrees in building systems engineering, and the Ph.D. degree in agricultural sciences from the Università Politecnica delle Marche (UNIVPM), Ancona, Italy, in 2013, 2017, and 2023, respectively, with a focus on rural constructions and agroforestry land management.

He is currently a Postdoctoral Research Fellow with the Department of Construction, Civil Engineering and Architecture (DICEA), UNIVPM, where his scientific work focuses on geomatics methodologies applied to geodetic and spatial engineering. His main research interests include georeferencing, land surveying, and spatial control, with particular emphasis on cartographic restitution in numerical, technical, and thematic formats, and the use of geographic information systems (GIS). His expertise also covers remote sensing in all its forms, including proximal sensing, and the application of artificial intelligence for the processing, analysis, and interpretation of spatial data.

Adriano Mancini was born in Fermo, Italy, in 1982. He received the B.S. and M.S. degrees in computer and automation engineering and the Ph.D. degree in artificial intelligence systems from Uniersità Politecnica delle Marche, Ancona, Italy, in 2004, 2006, and 2010, respectively.

He is currently an Associate Professor with the Department of Information Engineering (DII), Università Politecnica delle Marche, Ancona. He is also involved in several national and international projects (EU and industry-funded projects) in the field of precision/smart agriculture, remote sensing, robotics, cloud-based technologies, and big data. He is the author of more than 200 publications in international peer-reviewed conferences and journals. His research interests include computer vision and artificial intelligence, with applications in robotics, video analysis, human behavior analysis, and the automatic classification of images.

Alessandra Mobili was born in Turin, Italy, in 1987. She received the M.S. degree in building engineering–architecture and the Ph.D. degree in engineering of matter, environment, and territory from Università Politecnica delle Marche (UNIVPM), Ancona, Italy, in 2012 and 2016, respectively.

From 2015 to 2021, she was a Postdoctoral Research Fellow with the Department of Science and Engineering of Matter, Environment and Urban Planning (SIMAU), UNIVPM. From 2021 to October 2022, she was a Research Collaborator, and from 2022 to October 2023, she was a Technician at SIMAU, UNIVPM. Since March 2024, she has been a Researcher with SIMAU. She authored 65 publications and three national patents. Her research focuses on the development of multifunctional and sustainable building materials and on the development of smart and durable materials to increase the service life of buildings/infrastructures.

Gian Marco Revel was born in Ancona, Italy, in 1970. He received the M.S. degree in mechanical engineering and the Ph.D. degree in measurement science for engineering from Università degli Studi di Ancona, Ancona, Italy, in 1995 and 1999, respectively.

He is currently a Full Professor in mechanical and thermal measurement with the Università Politecnica delle Marche, Ancona, where he also plays the role of Rector Delegate for European Research. He is the author of more than 200 publications and holds scientific responsibility for several projects in the area of technologies for energy efficiency and materials in buildings. His research interests include sensors and measurement technologies for industrial applications, with particular focus on the construction sector, building energy efficiency, in-field material characterization, and nondestructive testing.

Dr. Revel is a Coordinator of the Materials and Sustainability Committee of ECTP and a Coordinator of the Italian Construction Technology Platform.