Sandro Parrinello
Anna Dell'Amico

Salvatore Barba
Andrea di Filippo

editors

D-SITE

Drones - Systems of Information on Cultural Heritage for a spatial and social investigation

PaviaUniversityPress

D-SITE, Drones - Systems of Information on Cultural Heritage for a spatial and social investigation / Sandro Parrinello, Salvatore Barba, Anna Dell'Amico, Andrea di Filippo (edited by) - Pavia: Pavia University Press, 2022. - 684 p.: ill.; 21 cm.

(Prospettive multiple: studi di ingegneria, architettura e arte)

ISBN 978-88-6952-159-1
ebook 978-88-6952-160-7

The present publication is part of the series "Prospettive multiple: studi di ingegneria, architettura e arte", which has an international referee panel. "D-SITE, Drones - Systems of Information on Cultural Heritage for a spatial and social investigation" is a scientific text evaluated and approved by double blind peer review by the Scientific Editorial Board.

Translation of chapters and treatment of citations and bibliography are due to the respective authors.

Pavia University Press
Edizioni dell'Università degli Studi di Pavia
info@paviauniversitypress.it
www.paviauniversitypress.it

Via Salasco, 5 - 20136 Milano
Tel. 02/5836.5751 - Fax 02/5836.5753
egea.edizioni@unibocconi.it
www.egeaeditore.it

Editors
Sandro Parrinello, Salvatore Barba, Anna Dell'Amico,
Andrea di Filippo

Graphic project
Anna Dell'Amico, Francesca Picchio, Anna Sanseverino

On cover: Drawing by Francesca Picchio and Sandro Parrinello
First edition: june 2022.

Stampa: Logo S.r.l. – Borgoricco (PD)

Organizer Committees

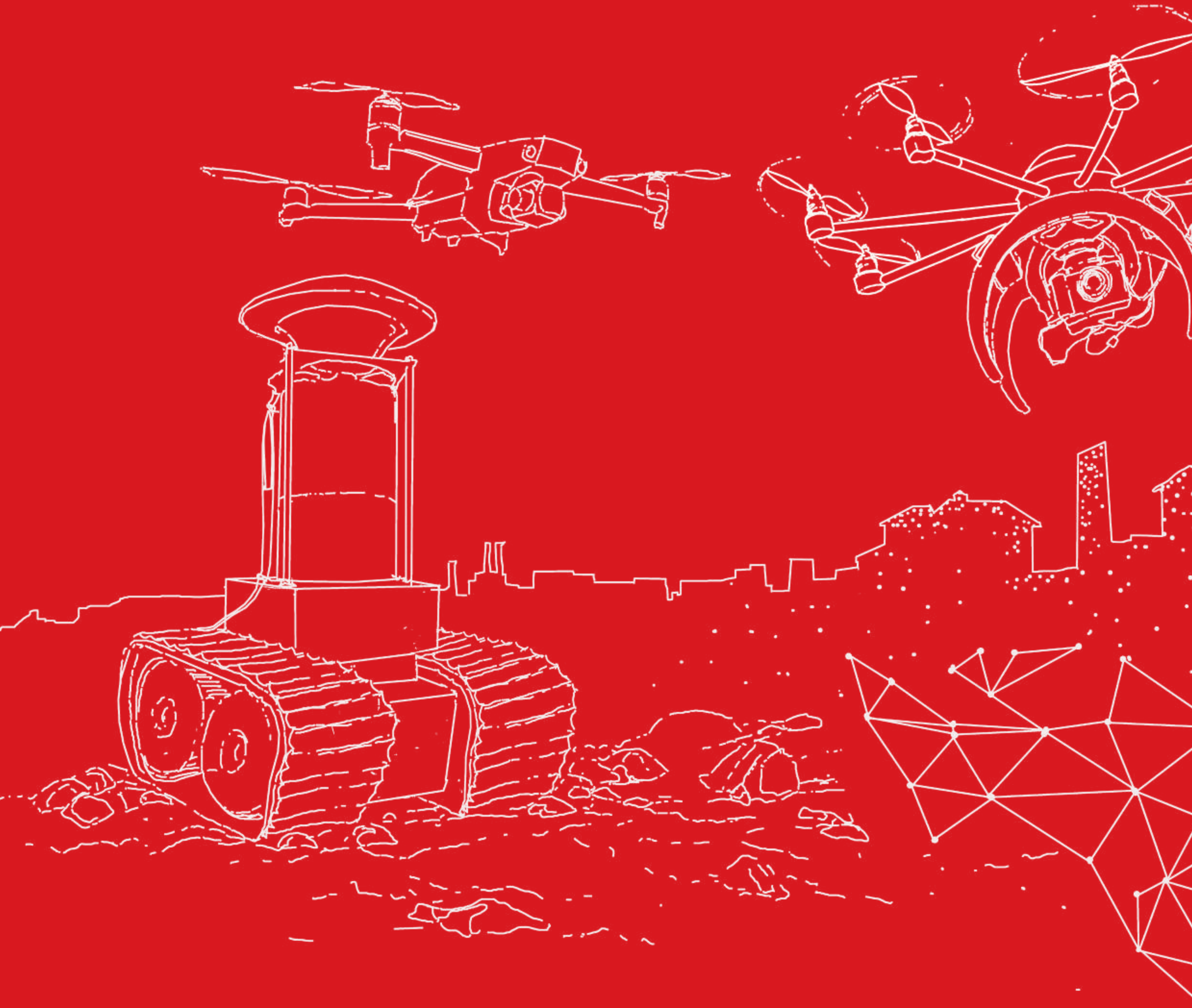
Sandro Parrinello	University of Pavia - Italy
Marco Limongiello	University of Salerno - Italy

Scientific Committees

Marcello Balzani	University of Ferrara - Italy
Salvatore Barba	University of Salerno - Italy
José Antonio Barrera Vera	University of Seville - Spain
Stefano Bertocci	University of Florence - Italy
Carlo Bianchini	La Sapienza, University of Rome - Italy
Mirco Boschetti	IREA, CNR - Italy
Enrico Borgogno Mondino	University of Turin - Italy
Stefano Campana	University of Siena - Italy
Massimiliano Campi	University of Naples Federico II - Italy
Gabriella Caroti	University of Pisa - Italy
Filiberto Chiabrando	Polytechnic of Turin - Italy
Gherardo Chirici	University of Florence - Italy
Antonio Conte	University of Basilicata - Italy
Krzysztof Cyran	Silesian University of Technology - Poland
Francesco Fassi	Polytechnic of Milan - Italy
Francesca Fatta	University of Reggio Calabria - Italy
Juan José Fernández Martín	University of Valladolid - Spain
Margherita Fiani	University of Salerno - Italy
Andreas Fritz	University of Freiburg - Germany
Diego González-Aguilera	University of Salamanca - Spain
Armin Gruen	ETH Zurich Faculty of Architecture - Swiss
Pierre Grussenmeyer	Institut National des Sciences Appliquées - France
Sorin Hermon	The Cyprus Institute - Cyrus
Xianfeng Huang	Wuhan University - Hubei China
Marinos Ioannides	Cyprus University of Technology - Cyprus
Falko Kuester	University of California - USA
Francesco Mancini	University of Modena and Reggio Emilia - Italy
Niccolò Menegoni	University of Pavia - Italy
Luis M. Palmero Iglesias	Polytechnic of València - Spain
Francesca Picchio	University of Pavia - Italy
Lorenzo Pollini	University of Pisa - Italy
Fabio Remondino	Bruno Kessler Foundation - Italy
Fulvio Rinaudo	Polytechnic of Turin - Italy
Mario Santana Quintero	Carlton University - Canada
Tesse D. Stek	Leiden University - Netherlands
Lina Tang	Chinese Academy of Sciences - Cina
Dieter Tengen	Technical University Braunschweig - Germany
Fabio Giulio Tonolo	Polytechnic University of Turin - Italy
Kyriacos Themistocleous	Cyprus University of Technology - Cyprus
Rebeka Vital	Shenkar College of Engineering and Design - Israel
Francesco Zucca	University of Pavia - Italy

Scientific Secretariat

Anna Dell'Amico (University of Pavia), Andrea di Filippo (University of Salerno), Silvia La Placa (University of Pavia)



Fausta Fiorillo[1], Luca Perfetti[1], Giuliana Cardani[2]

[1] DABC - Department of Architecture, Built environment and Construction engineering, Polytechnic of Milan
[2] DICA - Department of Civil and Environmental Engineering, Polytechnic of Milan

Milan, Italy

fausta.fiorillo@polimi.it
luca.perfetti@polimi.it
giuliana.cardani@polimi.it

Keywords:
UAVs, Machine Learning, Image Segmentation, Built Heritage, Damage Survey.

ABSTRACT

The paper describes an efficient workflow wherein UAV photogrammetry is combined with other 3D survey techniques (terrestrial photogrammetry, laser scanning and total station) to provide compre-hensive documentation of a historical building. The output orthoimage of the tiled roof allowed high-lighting the covering damage state. The research aims to test and evaluate the feasibility of auto-matically mapping roof damage using an image classification procedure based on supervised ma-chine learning. The methodology was validated on a historical building, now suffering from a seri-ous state of neglect.

Aerial-photogrammetric survey for supervised classification and mapping of roof damages

1. Introduction

Unmanned aerial vehicles (UAVs) have wide-ranging applications in the Cultural Heritage field (Barba et al. 2020). The flexibility and ease-of-use of modern consumer devices allow the visual inspection and digital documentation of even difficult-to-access or dangerous areas by enabling remote image acquisition and close-range aerial photogrammetry (Ronchi et al 2020). Indeed, a consistent and comprehensive photogrammetric survey project ensures accurate metric restitutions to support conservation and maintenance activities. UAVs are also useful in an integrated survey approach, supplementing other 3D measurements that typically miss the roofing data. The paper proposes an efficient survey workflow where UAV photogrammetry complements other 3D survey techniques such as terrestrial photogrammetry, laser scanning and total station survey in the comprehensive documentation of a historical building. The specific focus of this paper is to test and evaluate on a selected case study a semi-automatically mapping of roof damages through automatic image classification based on supervised machine learning (Grilli et al. 2019, Russo et al. 2021). This methodology could easily be used for the maintenance of the built heritage, especially when the critical conditions of a building do not allow access to roof structures from below.

2. Case study

The selected case study is *Palazzo Littorio* in Caronno Pertusella (VA), a representative building that housed the local branch of the Italian National Fascist Party in Italy.

After World War II, the building was used as a House of the People and then as a police station until it was closed and abandoned in the mid-80s, now showing serious signs of damage, particularly to the roof. The style of the building partly follows the dictates of Italian rationalism and partly the Art Nouveau style, especially the internal theatre space, and it presents two lateral additions to

Figure 1. Aerial photos of the case study: Palazzo Littorio (Caronno Pertusella - VA).

Figure 2. Schematic model of the Palazzo Littorio in Caronno Pertusella: main body in the middle, and the two 1-storey additions on either side of the main body. In green, the volume added to the east (bathrooms for the heliotherapy center), and in lilac, the one added to the west (after-hours club). The added volumes use the original surrounding wall and in red, in brilliant green and pink, the infills of the doors closed for safety reasons in different years.

the original volume that do not alter the impressiveness of the 3-storey central body, built-in 1930. The additions date only a few years after construction (Balin et al. 2021).

The building state of abandonment and degradation for over 30 years now necessitates intervention by the local administration to its preservation, reuse and valorization due to its location in the village center. The main problem of the building is the state of damage to the roof structures, especially in the area of the stairs, which makes an inspection from below too dangerous at the moment. Even in the absence of a definitive idea for reuse of the building, there is an urgent need to protect the existing roof structures, repairing provisionally the roof covering and so protecting the entire building from damages caused by water infiltration. The correct and detailed assessment of the state of preservation needs to be carried out in safety, and the use of the UAV is an ideal method to carry out an investigation that should not only be qualitative. The UAV also proves to be optimal for complementing surveys carried out with more traditional techniques from ground level, that inevitably leave undetectable areas.

3. Digital survey

An integrated digital survey based on range-based and image-based techniques was designed to comprehensively document this heritage building.

A TLS (Terrestrial Laser Scanner - Leica RTC3601) was used for the complete geometrical survey of the building exterior and interior. For creating 360° spherical panoramas, 3 HDR (High Dynamic Range) cameras allow for 5 bracketing exposures2. The VIS (Visual Inertial System) uses the other 5-cameras to track the laser scanner path. The integration of VIS and IMU platform enables real-time raw alignment between pairs of scans during on-site capturing. This raw registration ensures real-time control over the minimal needed overlap among scans and the completeness of surveyed areas. Therefore, the number and position of scans were planned ac-cording to a target-less acquisition mode. The survey consists of 84 scans (34 indoor) with a greater than 50% overlap to ensure proper cloud-to-cloud alignment to optimize the raw registration. The TLS acquisitions were generally set with a spatial sampling

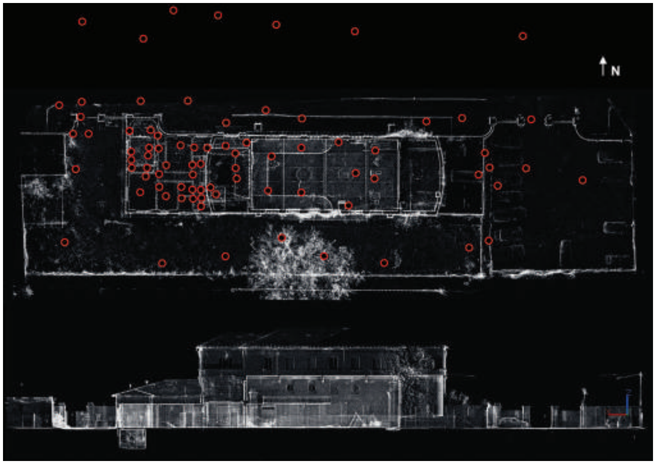


Figure 3. TLS scan stations and point cloud axial longitudinal section northward.

Figure 4. TLS panoramic images of the north facades and of theatre in the inner salon.

of 6mm@10m, ensuring the 1:50 graphic representation scale (1cm plotting error). The TLS workflow is performed according to the standard following steps: i) cleaning raw single scans, removing objects that can affect the ICP algorithm efficiency (moving automobiles, vegetation,

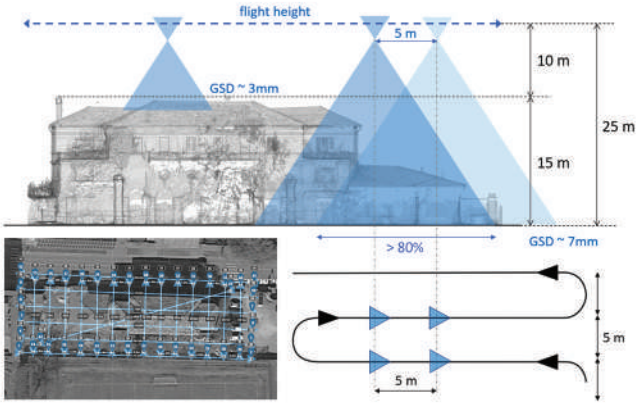


Figure 5. Aerial photogrammetric survey: GSD and overlap calculation and shot positions.

people, etc.); ii) scans alignment optimization; iii) in-deep individual scans filtering (removing noise, distant points and points measured with a sub-optimal incidence angle with the surface). The result is a measurable point cloud that represents the complete geometric model of the building.
The aerial photogrammetric survey was, on the other hand, essential for the roof measurement, which is the focus of this research.
The drone acquisitions were designed with 3 main goals: i) obtain an orthoimage of the tiled roof covering (1:50 scale); ii) quantify the extension and localization of the damage tiles; iii) integrate the terrestrial photogrammetry survey of the vertical facades with inclined shots. In addition, detailed photographic documentation was acquired to gather qualitative information on some critical areas. The instrumentation used is a Phantom 4 pro V.2 with an integrated camera of 8.8mm focal length and 2.6µm pixel size3. The flight mission was designed to meet the survey criterion (i): the flight height was set at 25m to ensure a 3mm GSD on the roof, and the distance among the photos was set at 5m to achieve more than 80% coverage on the ground level (overlap and sidelap).

The drone acquisition was supported and integrated by terrestrial photogrammetry, terrestrial laser scanning and total station measurements. A terrestrial photogrammetric acquisition was also carried out to produce orthoimages of the building facades, useful for materials and decay mapping (5mm GSD). The workflow follows the standard steps of the photogrammetric processing: i) images orientation by structure from motion; ii) tie points filtering and camera calibration optimization; iii) absolute orientation (scale and referencing in a local coordinate system); iv) dense image matching (dense cloud elaboration); v) mesh model creation; (vi) orthoimages generation.

4. Roof orthoimage classification

The drone images allowed an initial visual and qualitative inspection to understand the damage level and the possible risks of the roof elements. Together with the aerial photos, the roof orthoimage provided an overall view of the roof covering, combining these initial qualitative considerations with quantitative data such as the position and size of the most damaged areas and the most damaged roof tiles. It is noted that there are more tiles with holes to repair on the north pitch of the roof than on the south pitch. This identified critical area should always be kept under control during routine maintenance. The reason could be identified in the recent heavy hailstorms, as well as the low quality of the local clay tiles. This justification would also explain why several tile replacement works have been made over the years, as noted by the survey.
The tiles found are all of the interlocking tile typologies (*Marseillaise* tiles) with approximately 24x42 cm. With UAV survey, it was possible to identify 3 main categories of roof tiles. The 30s original clay tiles are very dark/black. The faded red roof tiles belong to the early 60s when the major change of use from the House of the People to the police station took place and to the later years for maintenance, now covered by a very variable grey patina dirt and biological growth. Lastly, the newer ones used during the last urgent repair works in 2020 have an intense red color, partly because they are cleaner.
These three roof tile types and the holes were detected and classified on the orthoimage thanks to an automatic

Figure 6. Roof photo of the north side towards the west corner before (left) and after (right) the urgent repair works in 2020. The 3-types of tiles are easily recognizable: the oldest dark ones, the new brilliant red ones and the other dull red ones.

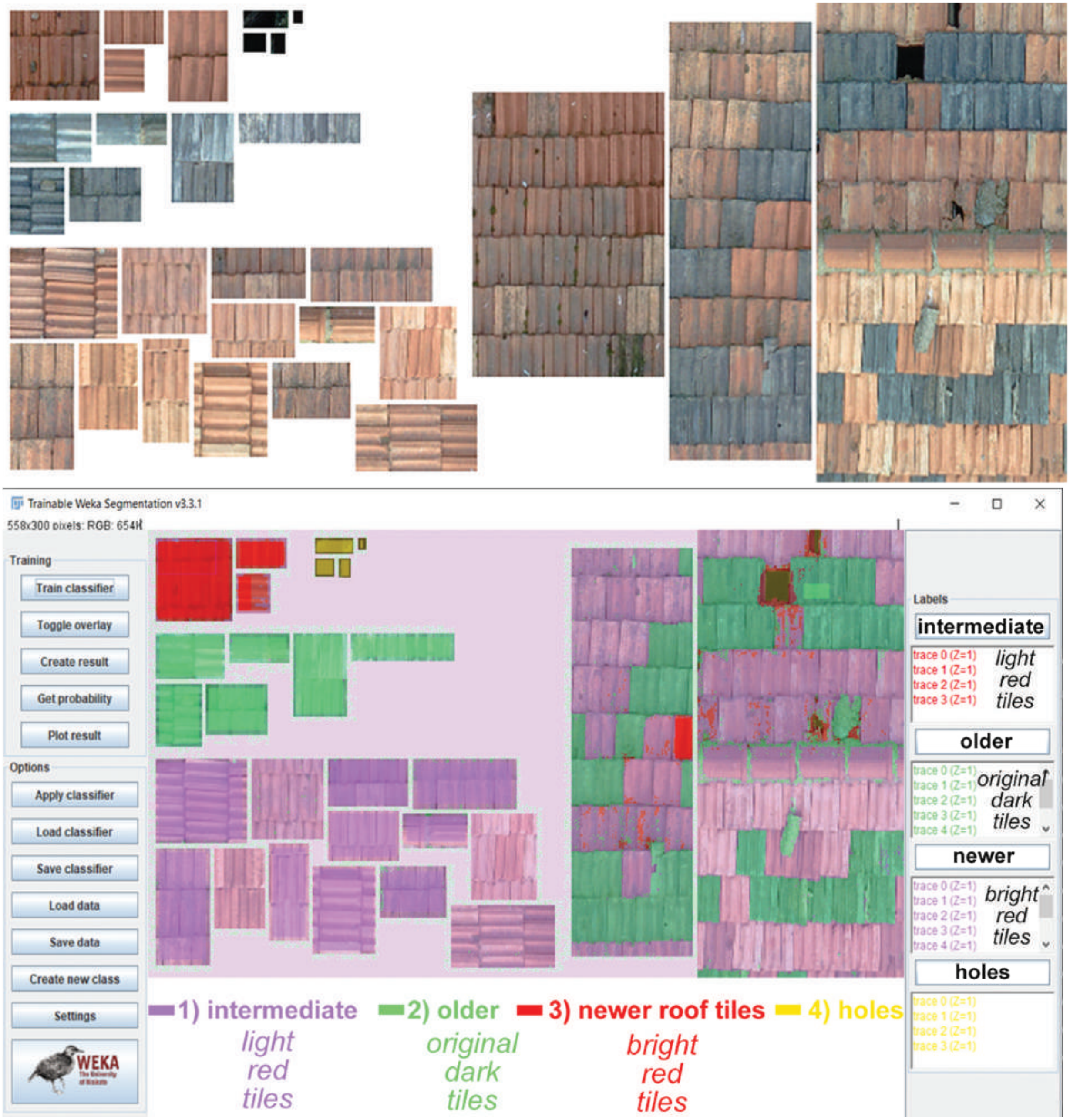


Figure 7. Mosaic composed of meaningful samples of the original orthoimage.

supervised procedure of an image segmentation system based on a trainable classifier (Grilli et al. 2018). Fiji4 (Schindelin et al. 2019), an open-source image analysis and processing software package that exploits WeKa5 (Frank et al. 2016) as an engine for machine learning models, was used for these experiments. An initial dataset of manually annotated images where the classes are defined is used to train the algorithm in a supervised way. In this case, for the classifier training, a mosaic composed of meaningful samples of the original orthoimage is used, where the following 4-classes are visible: 1) light red clay tiles - *intermediate* reference period between 1930 and 2020; 2) dark clay tiles - *older* original ones from 1930; 3) bright red clay tiles- *newer* of the latest 2020 repairs; and 4) *holes*. Each pixel in the image mosaic of the samples has been manually labelled with its corresponding class.

This solution was adopted to facilitate the computational capacity of the image-processing package. Indeed, the software supports images with a maximum resolution of 2 gigapixels, while the entire roof orthoimage has a 4662x1546 pixels resolution (1cm GSD). For the same reason, the orthoimage was than divided into 75 tiles of around 310x310 pixels, to be automatically classified as single images. Two python scripts were used to automatically divide the orthoimage into identical tiles and to automatically reassemble them after classification. In training the classifier, different sets of image feature parameters were computed and tested to identify the most effective ones in our case. Moreover, in order to assess the automatic procedure performance, the classification was also performed manually on the same orthoimage,

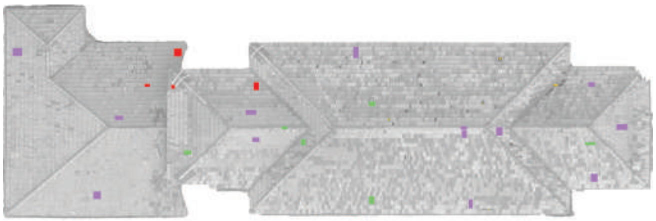
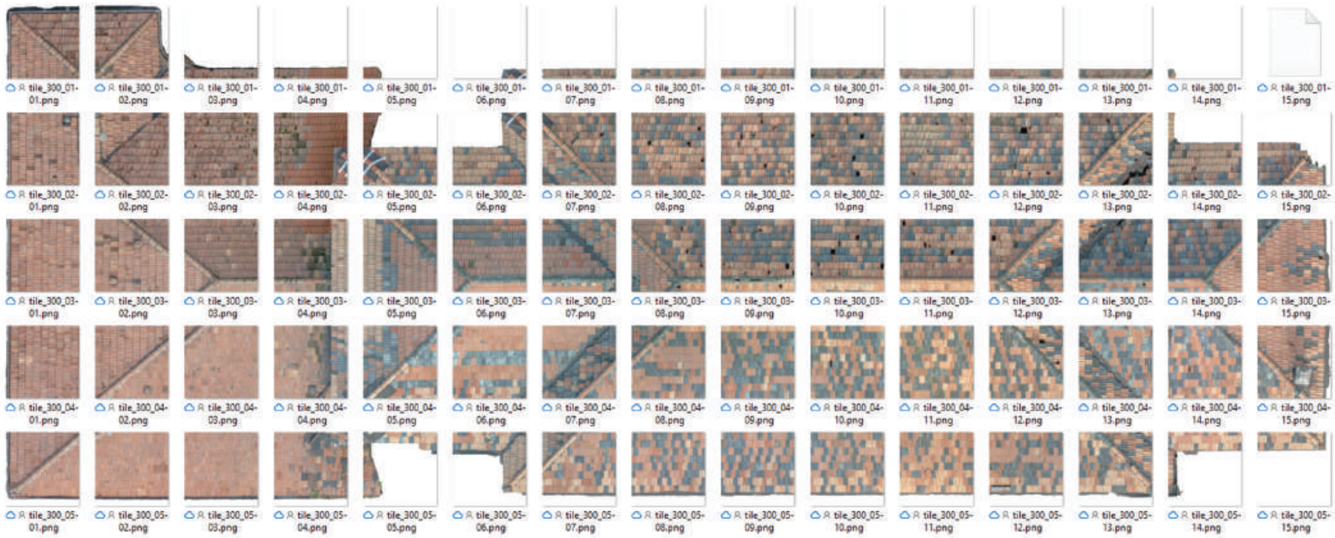


Figure 8. Position on the orthoimage of the chosen samples used for the classifier training.

Figure 9. Orthoimage decomposition in 75 tiles (5 rows and 5 columns) of about 310x310 pixels resolution.



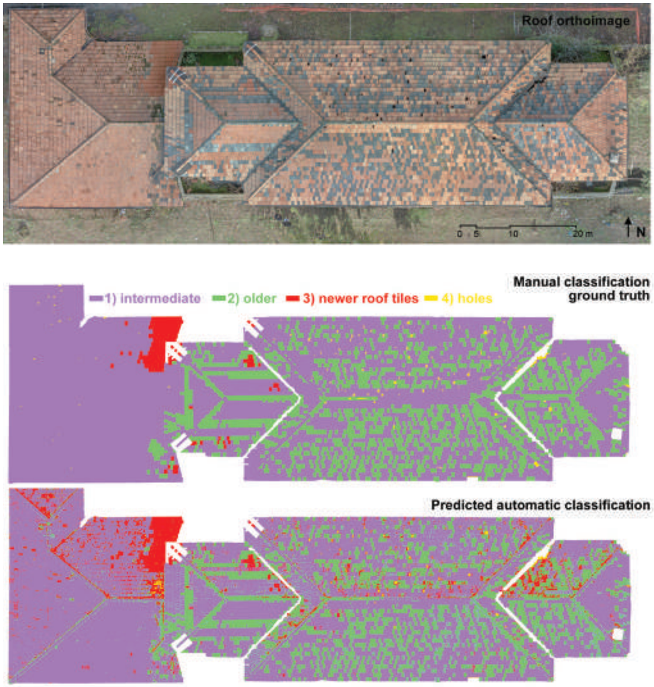


Figure 10. Original roof orthoimage and its segmentations (manual and automatic) in the chosen classes.

relying on the support of photos taken before and after the 2020 roof repairs and limited information from the municipality about maintenance works. The automatic classifier (Fast Random Forest) results were compared against this ground truth reference.

Comparing for each point the label provided by the classifier with the same manually annotated, for each class, true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) are determined at the pixel level and reported in a confusion matrix (Markoulidakis at al. 2021).

The TPs, i.e. the number of pixels that truly belong to a class, are reported on the diagonal; the FPs are in the columns, while the FNs are in the line. For example, 3.3% of the pixels classified as *newer* tiles are actually of the *intermediate* category. Simultaneously, 2.9% of pixels classed as *intermediate* should actually be in the *older* tile class.

In particular, the precision, recall and F1-score (Mohanchandra et al. 2015) calculated for each class were taken into consideration. The report with these data allows the following considerations. Precision and recall of the *intermediate* and *older* roof tile class are

confusion matrix (pixel numbers)		PREDICTIONS				TRUTH		
		intermediate	older	newer	holes	pixels	area [m^2]	tile numebers
TRUTH	intermediate (light red tiles)	3524826 66,8%	97432 1,8%	173661 3,3%	3138 0,1%	3799057	379,9	5836
	older (original dark tiles)	155669 2,9%	1147038 21,7%	51888 1,0%	281 0,0%	1354876	135,5	2081
	newer (bright red tiles)	9975 0,2%	97 0,0%	90839 1,7%	70 0,0%	100981	10,1	155
	holes	3980 0,1%	2820 0,1%	5860 0,1%	11682 0,2%	24342	2,4	37
PREDICTIONS	pixels	3694450	1247387	322248	15171			
	area [m^2]	369,4	124,7	32,2	1,5			
	tile numebers	5675	1916	495	23			

	precision	recall	f1-score
intermediate (light red tiles)	95,4%	92,8%	94,1%
older (original dark tiles)	92,0%	84,7%	88,2%
newer (bright red tiles)	28,2%	90,0%	42,9%
holes	77,0%	48,0%	59,1%

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 - score = 2\ \frac{Recal * Precision}{Recall + Precision}$$

Figure 11. On the upper part, there is the confusion matrix and related calculation of area and tile numbers for each class; on the bottom, the tab with the parameters for evaluating classifier reality.

very high; which means that the classifier is reliable for these classes. On the other hand, for examples, the precision of the *newer* tiles class is very low (~28%), which means that there are many FP, the majority of which are of the *intermediate* class; however, the recall is very high; meaning that there are few FN so that most of all new tiles have been identified correctly by the classifier, but many tiles of another class have ended up in this one. Finally, the precision of the *holes* class is acceptable (~77%); the majority of the holes have been correctly identified (few FP). Nevertheless, the recall is low (~48%), so many holes have not been identified and ended up in other classes (*older* class).

Therefore, this statistical data is helpful for both evaluating machine learning classifier performance and extracting practical information. Indeed, the sum along the lines of the confusion matrix gives us the number of pixels that belongs to that class according to the ground truth; while the sum on the columns is the number of pixels predicted by the classifier. From this information, it is possible to calculate the corresponding area and therefore the number of corresponding tiles for both true and predicted class. For example, the original tiles (1354876 real pixels and 369445 predicted) occupy an area of about 380m^2 and 370m^2 have been predicted. These values correspond to 6332 real tiles and 6157 predicted ones, with an error of about 10 tiles. The *holes* class is unquestionably one of the most significant, yet it also has the lowest recall. The classifier has predicted 1.5 m^2 of holes, but there were 2.4m^2 in total; therefore, 23 instead of 37 tiles to be replaced has been predicted. Future research will aim to improve the metrics used to classify these areas.

In summary, the workflow used consists of the following steps: 1) Orthoimage generation; 2) Identification of the classes; 3) Manual categories classification to create the ground truth; 4) Ad-hoc mosaic image creation with all classes visible, composed by orthoimage samples, used for the algorithm training; 5) Orthoimage splitting in equal tiles; 6) Automatic classification of all image-tiles; 7) Re-composition of the classified roof orthoimage; 8) Classifier validation; 8) metric useful information extraction.

Naturally, once the algorithm and the entire procedure have been validated, the same workflow can be applied to other case studies, omitting steps 3 and 8. Indeed, manual classification is a time-consuming process, and the research goal is to have the same results automatically.

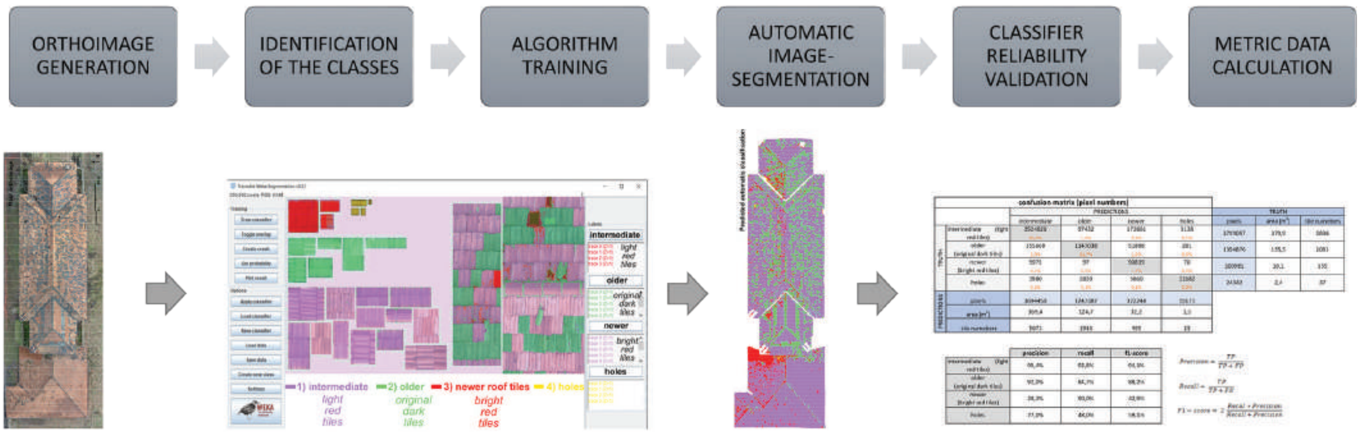


Figure 12. Workflow of the proposed methodology for damage mapping of roof covering based on supervised image segmentation.

5. Conclusions

A trainable automatic image classifier was employed on the roof orthoimage to determine the amount of damage accrued over time. The methodology employed proved to be quick and practical to be used as support for the municipal administration, providing not only an accurate geometric survey but at the same time also estimating the urgency and expense of the roofing works.

The proposed approach yields consistent and repeatable results and is effective for digital documentation and conservation activities, providing metric data such as those relating to the damaged areas to be repaired. The supervised image classification allows recognizing, localizing, and measuring the extent of the areas occupied by each labelled category. Therefore, it is possible to repeat the same procedure for other case studies and categories, such as semi-automatically map materials and decay on the facades.

Notes

1 For architectural scale survey, the instrument has excellent performance and specifications: 0.5-130m acquisition range, 360° (horizontal rotating base) x 300° (vertical rotating mirror) 360° Field of View, 4 mm at 10 m estimated noise range, and 2 million points per second max acquisition speed.

2 Each camera station captures 36 images, each with a resolution of 4000 × 3000 pixels.

3 Camera sensor size of 12.65 X 9.49 mm and resolution of 4864 X 3648 pixels.

4 Fiji Is Just ImageJ distribution of ImageJ and ImageJ2 which includes many useful plugins contributed by the community (https://fiji.sc/).

5 Waikato Environment for Knowledge Analysis, collection of machine learning algorithm (https://www.cs.waikato.ac.nz/ml/index.html).

Bibliography

Balin S., Cardani G., Fiorillo F. (2021). *Scan-to-HBIM: integrated survey and information modelling for the documentation of structural damage. Dienne - Building Information Modeling, Data & Semantics*, 9, pp. 8-21, 2021. ISSN 26108755

Barba S., di Filippo A., Ferreyra C., Limongiello M. (2020). *A pipeline for the integration of 3D data on aerophotogrammetric frameworks. The case study of Villa Rufolo*. In S. Barba S. Parrinello M. Limongiello A. Dell'Amico, *D-SITE, Drones - Systems of Information on Cultural Heritage. For a spatial and social investigation*. Pavia: Pavia University Press, 2020.

Frank E., Hall M. A., Witten I. H. (2016). *The WEKA Workbench*. In *Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques"*, Morgan Kaufmann, Fourth Edition, 2016.

Grilli E., Dininno D., Petrucci G., Remondino F. (2018). *From 2D to 3D Supervised Segmentation and Classification for Cultural Heritage* Applications. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2,* pp. 399-406, 2018. https://doi.org/10.5194/isprs-archives-XLII-2-399-2018.

Grilli E, Remondino F. (2019). *Classification of 3D Digital Heritage*. *Remote Sensing*, 11 (7), 847, pp. 1-23, 2019. https://doi.org/10.3390/rs11070847

Markoulidakis I, Rallis I, Georgoulas I, Kopsiaftis G, Doulamis A, Doulamis N. (2021). *Multiclass Confusion Matrix Reduction Method and its Application on Net Promoter Score Classification Problem*. *Technologies*, 9(4), 81, pp. 1-22, 2021. https://doi.org/10.3390/technologies9040081

Mohanchandra K., Saha S., Murthy K. S., Lingaraju G. M. (2015). Distinct adoption of k-nearest neighbour and support vector machine in classifying EEG signals of mental tasks. *International Journal of Intelligent Engineering Informatics,* 3(4), pp. 313-329, 2015.

Ronchi D., Limongiello M. (2020). *Landscape survey and vegetation filtering for archaeological cartography. A UNESCO World Heritage site in Cerveteri: "Banditaccia" necropolis and the "Via degli Inferi"*. In S. Barba, S. Parrinello, M. Limongiello, A. Dell'Amico, *D-SITE, Drones - Systems of Information on Cultural Heritage. For a spatial and social investigation*. Pavia: Pavia University Press, 2020.

Russo M., Grilli E., Remondino F., Teruggi S., Fassi F. (2021). *Machine Learning for Cultural Heritage Classification*. In Giordano A. Russo M. Spallone R. *Augmented Reality and Artificial Intelligence in Cultural Heritage and Innovative Design Domain*. Milano: FrancoAngeli, 2021.

Schindelin J., Arganda-Carreras I., Frise E. et al. (2019). *Fiji: an open-source platform for biological-image analysis. Nat Methods*, 9, pp. 676-682. https://doi.org/10.1038/nmeth.2019.