

RESEARCH

Open Access



Using deep learning for enrichment of heritage BIM: Al Radwan house in historic Jeddah as a case study

Yehia Miky^{1,2*} , Yahya Alshwabkeh³ and Ahmad Baik¹

Abstract

Building information modeling (BIM) can greatly improve the management and planning of historic building conservation projects. However, implementing BIM in the heritage has many challenges, including issues with modeling irregular features, surveying data occlusions, and a lack of predefined libraries of parametric objects. Indeed, surface features can be manually distinguished and segmented depending on the level of human involvement during data scanning and BIM processing. This requires a significant amount of time and resources, as well as the risk of making too subjective decisions. To address these bottlenecks and improve BIM digitization of building geometry, a novel deep learning based scan-to-HBIM workflow is used during the recording of the historic building in historic Jeddah, Saudi Arabia, a UNESCO World Heritage site. The proposed workflow enables access to laser scanner and unmanned aerial vehicle imagery data to create a complete integrated survey using high-resolution imagery acquired independently at the best position and time for proper radiometric information to depict the surface features. By employing deep learning with orthophotos, the method significantly improves the interpretation of spatial weathering forms and façade degradation. Additionally, an HBIM library for Saudi Hijazi architectural elements is created, and the vector data derived from deep learning-based segmentation are accurately mapped onto the HBIM geometry with relevant statistical parameters. The findings give stakeholders an effective tool for identifying the types, nature, and spatial extent of façade degradation to investigate and monitor the structure.

Keywords HBIM, Deep learning, Hijazi architecture, Laser scanner, Photogrammetry

Introduction

The use of BIM in historical buildings has increased in recent years because it allows for accurate modeling by utilizing interactive parametric objects that contain information about the object's geometry. The informed

model becomes an important part of the process of evaluating and managing historical heritage structures due to the various historical information needed by different stakeholders, such as documents, structural, restoration, and monitoring information, as well as the current state of the building in a 3D environment that is required for conservation [1–5]. Building information models (BIMs) are created using large libraries of predefined parametric objects found in most BIM software packages. This makes modeling more efficient because 3D geometry does not need to be created from scratch. These library objects' parameters are changed to fit the necessary dimensions and property for a project. The main challenge for the Heritage BIM (HBIM) is the lack of predefined parametric objects able to respond to the unique

*Correspondence:

Yehia Miky
yhhassan@kau.edu.sa

¹ Geomatics Department, Faculty of Architecture and Planning, King Abdulaziz University, 21589 Jeddah, Saudi Arabia

² Civil Engineering Department, Faculty of Engineering, Aswan University, Aswan 81528, Egypt

³ Department of Conservation Science, Queen Rania Faculty of Tourism and Heritage, The Hashemite University, P.O. Box 330127, Zarqa 13133, Jordan



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

and heterogeneous structures of historic buildings, as the native primarily adapted to new constructions.

Developing HBIM with resourceful libraries necessitates detailed and accurate surveys of architectural elements [6]. The method called scan-to-BIM is used to transform existing survey data into parametric modeling [7, 8]. Scan-to-BIM procedures include acquiring data, producing point clouds, importing and recognizing them in a BIM environment, tracing and modeling the structure, and obtaining associated attributes [9].

The acquisition of as-built data has recently depended on the reality capture technologies like terrestrial laser scanners (TLS) and photogrammetry. A laser scanner device captures the geometry of an object directly, producing 3D point clouds that are dense, dependable, and colorful. The scanner operates using either time-of-flight (TOF) or phase-based principles. The travel time of the emitted infrared laser pulse is used in a time-of-flight scanner to calculate distance and object coordinates. Whereas the phase-based laser scanner calculates the distance between the scanner and the scene surface by measuring the waveform between the emitted and returned signals [10]. Multiple scans from various angles are usually required to fully cover the surface area, depending on the size and shape of the scanned object [11]. A unified point cloud that represents the entire site is produced by aligning these scans with a single reference system. Despite its potential benefits, TLS cannot be used alone to model the entire historical scenes [12]. Difficulties may arise as a result of the complexity of the architectural elements, ambient and self-occlusions. Some field campaigns have roof accessibility issues in upper portions and gaps in the survey data result from a lack of high platforms [13, 14]. Most TLS allow acquiring colored point clouds by taking digital photos with an integrated camera in the system. However, color data may not be sufficient to detect surface features because the ideal camera position does not always coincide with the scanner position, resulting in unreliable recognition and BIM parametric reconstruction [15].

Photogrammetric techniques, on the other hand, use images taken from various perspectives with enough overlap to record the object's geometry. The fundamental principle of photogrammetry, known as triangulation, is that if a point is depicted in at least two images, the corresponding 3D object coordinates can be mathematically determined. Currently, photogrammetry is an efficient and cost-effective technique for recording historical structures. Portable cameras and those mounted on an unmanned aerial vehicle (UAV) could be utilized for collecting the data [16]. Photographs taken at optimal positions and times provide an accurate representation of color and façade characteristics. The processing pipeline

has developed into an efficient solution for 3D reconstruction with the latest developments in photogrammetry algorithms and computer vision with automated techniques [17]. Limits on image quality, the camera networks around the surveyed object, shadows and model scale are the core issues, which could have an impact on photogrammetric processing and the final 3D outputs [18]. Recently, different research presented effective approaches combining laser scanner and photogrammetry to accurately recording the geometry and the morphological data of large and complex heritage sites for HBIM applications [19–21].

HBIM models should become more useful for heritage applications by integrating high levels of structural information that represent surface features, damage state and crack propagation. Such information may use in-depth knowledge of the building aids in the optimization of the historical building's management, maintenance, and conservation processes [22]. For example, [23, 24] attached photos and orthophotos to BIM objects as additional data for heritage preservation. Most Scan-to-BIM approaches rely on human interpretation for tracing building surface features, which results in time-consuming, laborious processes and the risk of making subjective decisions [25]. Recently, there has been an increase in demand for automatic or semi-automatic procedures for segmenting and classifying cultural heritage features (e.g., windows, columns, walls, cracks, etc.). Deep learning image segmentation techniques are changing the way heritage experts interpret, recognize, and classify building components. Deep learning-based fully automated feature extraction methods from point clouds are currently at the cutting edge [25]. There are only a few applications in literature that use deep learning to classify 3D point clouds in different objects from cultural heritage scenes [26]. The main challenge is in the complexity and variety of point clouds that result from irregular sampling, variable object density, various object type and availability of labelled datasets [27].

Therefore, the paper proposed multi-sensor survey workflow to enrich HBIM for conservation purposes. The workflow takes the benefits of both TLS and UAV photogrammetric survey data, while deep learning is used to augment HBIM with damage state and surface decay information. Data fusion was used during the recording of the Al Radwan historic building in historic Jeddah, Saudi Arabia, a UNESCO World Heritage Site. The workflow depicted in Fig. 1 uses the UAV data to complement the TLS data in the upper occluded parts of the Al Radwan building using realistic orthophoto generated from SfM photogrammetry. The orthophoto can reveal details about the surface damage and texture with precise spatial distribution. In our approach,

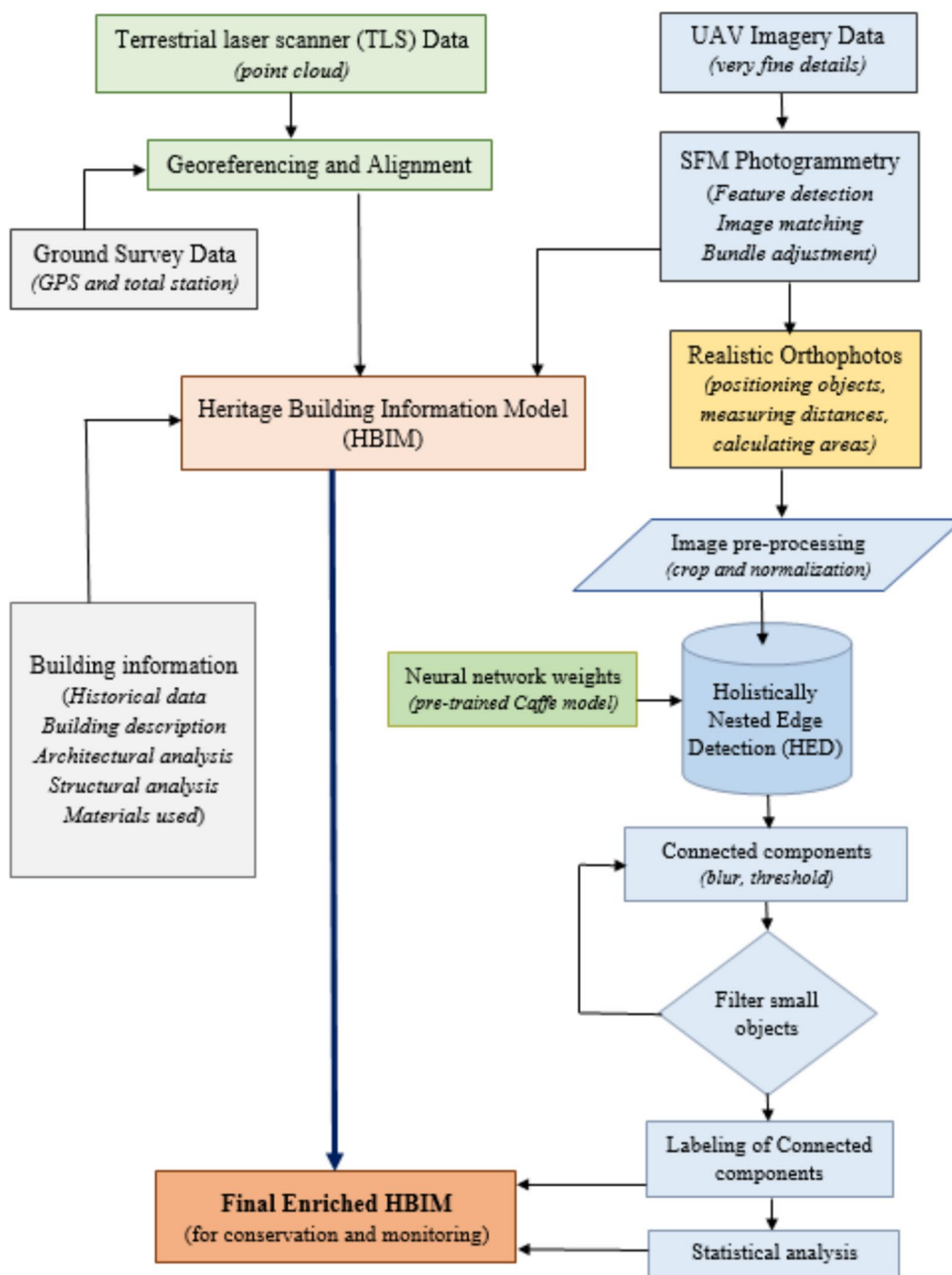


Fig. 1 HBIM conservation workflow

the orthophotos were used for HBIM texturing to provide realistic rendering and decay mapping. And used I combined with deep learning segmentation for quantifying the surface decay automatically and extract the feature components and their statistical analysis. The proposed algorithm provides a robust segmentation result with minimal computational power compared

with those applied in point cloud segmentation. The results are then mapped correctly to the BIM model for analysis of all types of surface features and weathering forms, including their extent, typology, cause, and conservation activities required.

To summarize, the following contributions were covered in this paper:

1. A fusion-based method that improves digitizing HBIM geometry with comprehensive mapping data by utilizing the results of TLS and UAV photogrammetry surveys.
2. Using deep learning in combination with ortho-photo to segment surface features and quantify weathering forms and facade degradation.
3. Creating a new HBIM library with high-detail parametric objects for Saudi Hijazi architectural elements.

The following sections comprise the paper's structure: Sect. "Background" presents a general background on deep learning. Sect. "Experimental examples of applying HED model" discusses the experimental examples of using the HED model. Sect. "Case study description: Al Radwan House in Ancient Jeddah City" contains the case study description. Sect. "Data Acquisition and Pre-processing" describes data acquisition and pre-processing. Sect. "HBIM and Creating Library of Architectural Elements" discusses the use of HBIM and the development of a library of Hijazi architectural elements. Sect. "Deep Learning process" addresses the outcomes of deep learning algorithms. Sect. "Discussion" discusses the results, while Sect. "Conclusion" presents the conclusions.

Background

Edge detection

Edge detection refers to the process of identifying and delineating the boundaries or edges within images acquired from different sensors. It is applied to images to enhance and extract meaningful information. The application of computer vision-based edge detection technology for heritage buildings has been widely studied. Various techniques such as threshold segmentation [28, 29], morphology [30], wavelet transform [31], and filter-based algorithms [29, 32, 33] have been explored in the field of edge detection. The main limitation of these algorithms is their focus on analyzing specific features of an image, with parameters designed for a particular dataset. Therefore, researchers have proposed the use of machine learning in image processing tasks [34, 35]. The introduction of Convolutional Neural Networks (CNNs) by Lecun et al. [36] marked a significant milestone, demonstrating the application of CNNs in handwritten character recognition. The main advantage of CNNs is its ability to automatically learn features and identify edges in images. Subsequently, studies have suggested the use of CNNs for tasks such as image classification [37–39], object detection [40–42], semantic segmentation [43–45], coastal vegetation edge detection [46], and crack detection [47–49].

Edge detection, segmentation, and crack detection are interconnected processes in image analysis. In one hand, segmentation involves partitioning an image into meaningful regions that represent objects or parts of objects based on certain criteria, such as color, intensity, or texture. Edge detection is a critical step in segmentation. By identifying edges, the boundaries of segments can be delineated more accurately. On the other hand, crack detection specifically focuses on identifying cracks or fractures in materials, structures, or surfaces. It often relies on edge detection techniques to identify the linear features characteristic of cracks. It can be considered a specialized form of edge detection tailored for identifying discontinuities in a material. In recent studies on crack detection and segmentation, various deep learning architectures have been applied [50–52].

Deep learning models, particularly those for semantic segmentation, have been extensively studied and have shown promising results in edge detection. Research in edge detection based on deep learning has proposed techniques such as image classification, object recognition, and semantic segmentation. For example, the encoder-decoder structure, a common framework used in various computer vision tasks, can be utilized without explicitly identifying edges or edge regions beforehand. It consists of an encoder network that extracts features using convolution, pooling, and activation layers, and a decoder network that upscales low-resolution feature maps to high-resolution maps for pixel-wise classification. Several encoder-decoder methods have been developed for crack detection on surfaces [53–61]. The objectives of these studies rely on investigation of the effectiveness of these methods in crack detection, segmentation tasks, and contributing to advancements in automated crack analysis.

When comparing deep learning-based edge detection techniques with classical edge detection algorithms like the Canny edge detector, Sobel operator, and Prewitt operator, several advantages can be observed in terms of methodologies, applications and objectives. First, classical methods utilize predefined rules and mathematical operations to identify edges, while deep learning techniques learn features directly from data during a training phase. Deep learning models can extract hierarchical representations, including edges, from input images without the need for explicit feature engineering [63]. Second, deep learning models can adapt and generalize well to different edge patterns and variations in data. They learn to recognize complex patterns and are often more robust to variations in input conditions [62]. Third, unlike classical methods which typically operate at the pixel level and may not effectively capture contextual information, deep learning models can capture contextual information

by considering larger spatial contexts. This enables them to understand and distinguish edges based on surrounding patterns, leading to more robust results, such as the calculation of area and perimeter for a certain feature [54]. Fourth, classical methods are often task-specific and may not readily transfer to different applications without significant modifications. Conversely, pre-trained deep learning models, especially in transfer learning scenarios, can offer features that are transferable across different domains.

However, deep learning-based techniques require a training phase and may be computationally more intensive compared to classical edge detection methods, which can be applied directly to an image in a single pass [63]. Researchers are motivated to employ deep learning-based techniques for edge and object boundary identification due to the remarkable, superior, and dynamic outcomes they achieve compared to conventional approaches across various domains [64–68].

Transfer learning

Deep learning technique requires enough data to complete the model training phase. However, the current dataset of images collected from heritage sites cannot meet the needs of such model training, and thus, the transfer learning method was adopted [69–73]. Transfer learning is a powerful technique in deep learning that allows models to leverage knowledge learned from one task and apply it to another related task. With transfer learning, pre-trained models that have been trained on large-scale datasets and complex tasks can be used as a starting point for solving new problems. Instead of training a model from scratch, transfer learning enables the transfer of knowledge and representations from the source task to the target task, which significantly reduces the amount of training data and computational resources required. By utilizing the learned features and weights from the pre-trained model, the model can generalize better, extract meaningful features, and achieve higher performance on the target task. In recent studies on image analysis and segmentation, transfer learning has been employed as an effective approach. Xu et al. [74] used the transfer learning approach to automatic segment and label histopathology images. Zhang et al. [75] applied transfer learning for identifying and extracting crack information from images of earthen heritage sites.

The researchers were able to train a more effective model and automate the process of crack extraction, which could be useful for conservators. In this research, we adopted the concept of transfer learning and used a pre-trained model: Holistically nested edge detection (HED). It is implemented using the publicly available Caffe library, this Caffe model is encoded in two files: a

text Caffe JSON file with the model definition and a text file with the neural network weight [76].

Deep learning model: holistically nested edge detection (HED)

Holistically nested edge detection (HED) is a convolutional neural network (CNN) that employs a unique approach to achieve multi-scale image convolution [77]. Like many other networks, its core is based on VGG16. VGG16 is a 16-layer deep neural network with a total of 138 million parameters. The architecture of VGG16 includes input layer, convolutional layers, activation layer where a function is used to provide a matching output for positive inputs and outputs zero for negative inputs, hidden layers, pooling layers which help reduce the dimensionality and the number of parameters of the feature maps created by each convolution step, and finally fully three connected layers. The main modification of HED is that after the last pooling layers the fully connected layers are cut off and only the convolutional layers are kept as shown in Fig. 2. In addition, unlike other CNNs that increase kernel size, HED progressively reduces image resolution through five sets of convolutional layers, each using 3×3 kernels, which are separated by max pooling layers of 2×2 .

After each set of convolutional layers, a side output layer is produced, five side outputs are formed. The first side output captures local boundary details but is prone to noise and false inland boundaries, while the fifth side output detects salient boundaries and is more robust to image noise. To obtain the final output, which predicts the likelihood of each pixel being an edge, these five side output layers are optimally fused. The training process of the HED model follows a structured approach involving feed-forward and backpropagation stages within each epoch. During the feed-forward stage, the HED model employs its internal weights to predict the locations of edges using the raw input image. By comparing the predicted edge positions with the

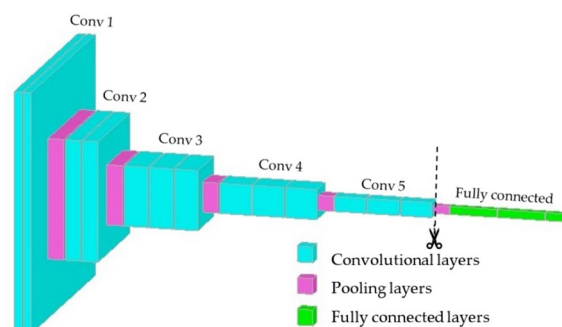


Fig. 2 Typical VGG16 architecture

ground-truth binary image, the discrepancy is back-propagated through the hidden layers, resulting in the update of the model's internal weights. These updated weights are then utilized in the subsequent epoch's feed-forward stage, continuing the iterative training process of the HED model. For a detailed overview of the HED architecture and its functionality, refer to Xie and Tu, Kokkinos [77, 78].

Experimental examples of applying HED model

Before applying the proposed methodology to the case study, an initial experiment was conducted to evaluate the efficiency of the algorithm used for edge detection, feature extraction, and image segmentation under various conditions. The results were then compared with traditional methods such as the Canny and Sobel operators, as depicted in Fig. 3. In this experiment, several examples of different historical building facades in KSA and Jordan were utilized. These examples were captured at different times and under diverse conditions. The selection of these images aimed to ensure diversity, enhance contrast, and present challenges for automated segmentation. For instance, example (A) represents a part of the main facade of Amra palace in Jordan, a UNESCO World Heritage site, built in the desert between 705 and 715 A.D. during the reign of the Umayyad Caliph Walid I. This example exhibits irregular stones with varying sizes and colors, as well as irregularities in architectural features. In example (B), the facade of Al-Abidit palace in Jordan, the palace was constructed in the second century BC by Hyrcanus during the reign of Seleucids IV, has large stones and different types of cracks. This example also demonstrates variations in tone for the same feature from one location to another in the image. Example (C) illustrates a UAV image of a building in Al Balad, historic Jeddah, Saudi Arabia, highlighting the irregularity and diversity of features and materials used, such as wood, cement, and bricks with significant lighting.

Figure 3 demonstrates the rich extraction efficiency of edges for traditional edge detection algorithms like Canny and Sobel. On the other hand, the pre-trained network algorithm that has been suggested enhances the ability to detect cracks and can offer labeled and contextual semantic information about edges and cracks, including area, length, and premier, which can be utilized for material conservation analysis. Thus, based on the experimental examples of applying HED model, using deep learning technique to provide such informative and semantic segmentation in the digital cultural heritage domain can help recognize the different architectural elements, cracks and surface condition at an appropriate level of detail.

Case study description: Al Radwan house in historic Jeddah City

Historic Jeddah city is located on the west coast of Saudi Arabia. Its architectural style was influenced by social, religious, cultural, and climatic factors. Jeddah shares many architectural features with Red Sea coastal cities such as Yanbu in Saudi Arabia and Suakin in Sudan, which at first look appears to be a twin of old Jeddah. Several historic buildings in the city are over 300 years old. The city was officially listed as a UNESCO world heritage site in 2014. The architectural elements of Jeddah houses known as Islamic Hijazi style. These elements include Roshan's windows, plasters, and various gate shapes. Even though many traditional buildings have been renovated and made public, many historic structures in Jeddah remain endangered.

For restoration and maintenance, the building structures require more data and engineering analysis. Among the most famous and oldest existing buildings so far is Al Radwan house depicted in Fig. 4a. Al Radwan house is located in Al Mazloun Neighborhood in the heart of historic Jeddah as shown in Fig. 5. The house was named Radwan House, after its owner, Mr. Abdel Salam Radwan. The house has four levels in addition to the roof. The total land area is 145 m², and the total building space is 744 m². Radwan's house is notable for its several Roshans, which are rich in detail. The Roshan façade, depicted in Fig. 4b, is made up of manjour, which are flat wooden panels, some of which are solid and others decorative. The manjour craft is widely used in the architecture of old structures around the country. It is made up of longitudinal hardwood pieces that are cut in a precise way and coupled with each other by connecting the joints of these pieces at an angle of generally 45 degrees and sometimes 90 degrees to generate small apertures with a variety of geometric shapes. The building's interior has also been carefully planned, with larger rooms in the middle and smaller living areas on either side. The massive twin staircases are located in the middle of the rear of the building. In the back are the quarters for the aides and servants. Because each story of the building has a different architectural design, the interior of the structure is just as distinctive as the exterior.

Data acquisition and pre-processing

Terrestrial laser scanning

The Al Radwan building's exterior facade and interior spaces were scanned using the Faro 350 s scanner. The scanner can provide up to 976,000 points per second. The scanner's data collection range is between 0.6 and 350 m. Its field of view is 360° × 300°. The system performs automatic point cloud registration, which improves scan control and field productivity. The



Fig. 3 Experimental examples of applying HED model on buildings facades: **A** Amra palace in Jordan. **B** Al-Abidit palace in Jordan. **C** Al Balad, historic Jeddah, Saudi Arabia

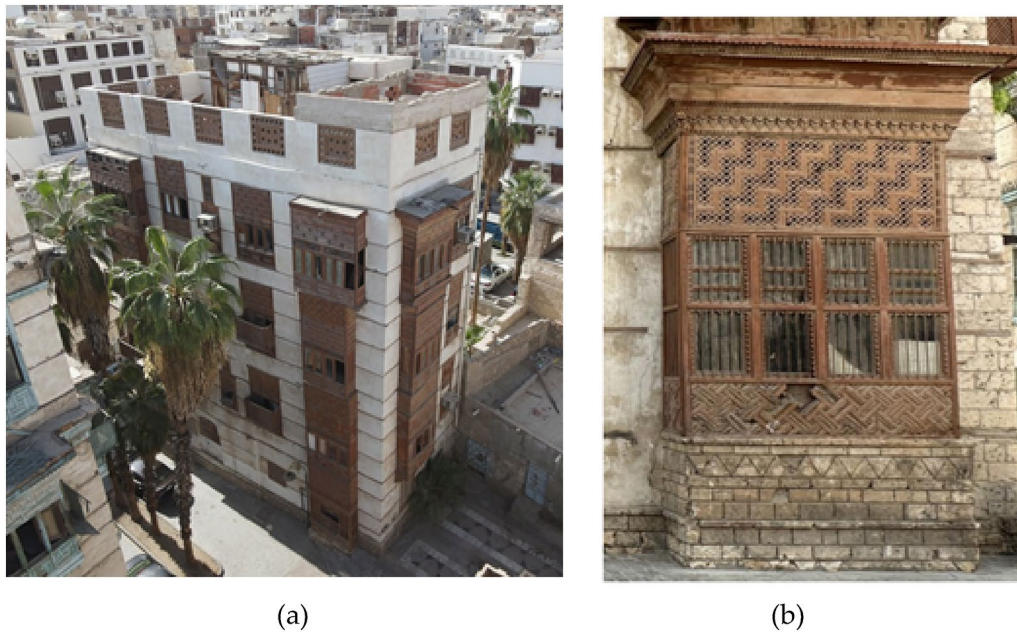


Fig. 4 a AL Radwan House b The Roshan



Fig. 5 The location of AL Radwan House, Jeddah

scanner has an integrated camera that has 165 Mega pixels to color the point cloud. The TLS positions were carefully chosen to ensure sufficient overlap between the adjacent point clouds for automatic registration. The number of scans required is affected by the amount of self-occlusion and obstacles, as well as the size of the object in relation to the sensor range. Using terrestrial laser scanners to collect data in confined spaces such

as the AL Radwan building, which has multiple stories and many rooms, is a tedious and time-consuming process. TLS locations were carefully planned to provide complete coverage of the building while also ensuring sufficient overlapping between the different scans. In some cases, the initial scanner position plan had to be changed to improve the results or to ensure safety. We completed 107 interior and exterior scans for our

project. Figure 6 shows a scanner workplan and an example of an exterior scan.

To create a point cloud model, the raw data must first be processed. Noise data elimination and scan registration are two of these procedures. The presence of noise and outliers in scanned data, caused by factors such as human interference, environmental conditions, and sensor error, may affect the model's accuracy. Noise data was manually removed using the Autodesk ReCap software editing tool to ensure that the data accurately represented the areas surveyed. The Autodesk ReCap software was used to automate the scan registration step, which aligns all scans in a unified coordinate system. The software-generated registration report reveals that over 98% of the overlapping points from two data sets have less than 6 mm corresponding deviations as shown in Fig. 7.

The final model was georeferenced with Ground control points (GCPs) designated in accordance with KSA local reference know as Ain el Abed (UTM 36 North). A Leica Total Station TCR 1201 with a reflector (0.3 mgon, 1 mm + 1.5 ppm precision) and Leica GNSS Viva GS15 receivers (3 mm + 0.5 ppm accuracy) were among the instruments used to establish the GCPs. With a Leica Runner 24 level, which offered an accuracy of 2.0 mm per kilometer double-run leveling, the orthometric height values of these GCPs were computed. According to the registration report generated by Autodesk Recap software, over 98% of the overlapping points are within 6 mm of corresponding features. Figure 8 shows a 3D model of the building colored with the TLS camera, as well as examples of an interior scene, a long section, and a cross section. The model contains over 1.2 billion points and

has an average Ground Sampling Distance (GSD) of 3 mm.

Despite the potential of the prior TLS, the data generated by these systems is not suitable for full 3D modeling of heritage applications. Texturing and feature interpretation of the scene using the color data obtained from TLS attached camera may not be of enough texture quality. It is possible that the camera position that yields the best quality images for texture mapping is not the same as the device viewpoints. Different lighting conditions from multiple scans taken at different times of the day may result in non-homogeneous color texturing and color jumping between the collected scans. This will obviously disturb the appearance of the final model, as seen in Fig. 9a. It is still challenging to interpret and recognize surface features and pathologies in BIM modeling, such as cracks, decay, and surface damage. Furthermore, there may be difficulties to survey roofs and upper portions of the building facade due to the absence of high platforms results in missing and gaps in the surveyed data as shown in Fig. 9b.

UAV photogrammetric data

The use of TLS to survey the entire structure of Al Radwan building is limited due to the city's narrow corridors. Additionally, terrestrial scanning capability cannot reach the upper zone of the building. These factors led to the implementation of a second UAV survey campaign. Aerial photogrammetry with increased mobility was used to supplement the collected point data in shadow areas and produce reliable orthophoto that improve surface feature interpretation. Aerial surveys were carried out using

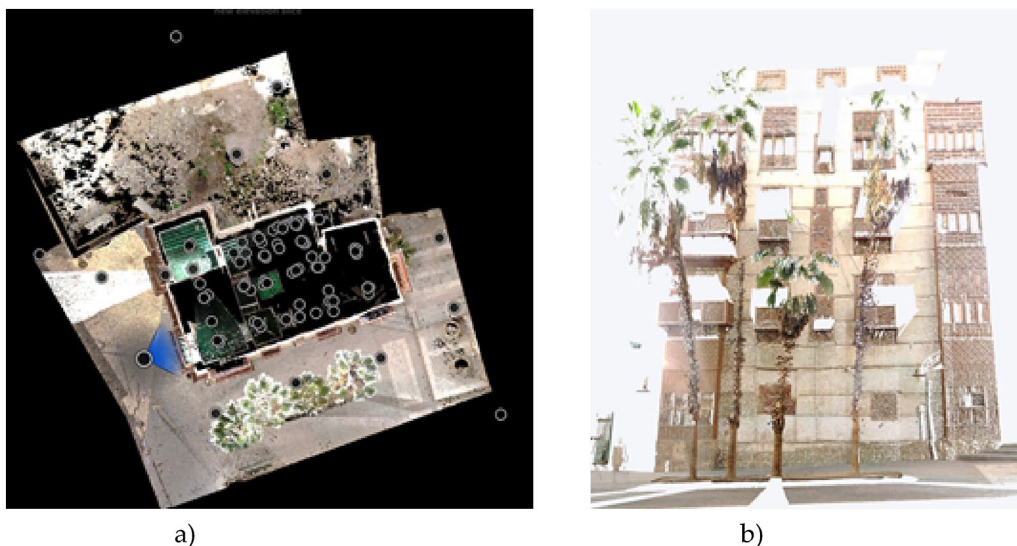


Fig. 6 a) TLS scanning positions b) an example of the exterior scan

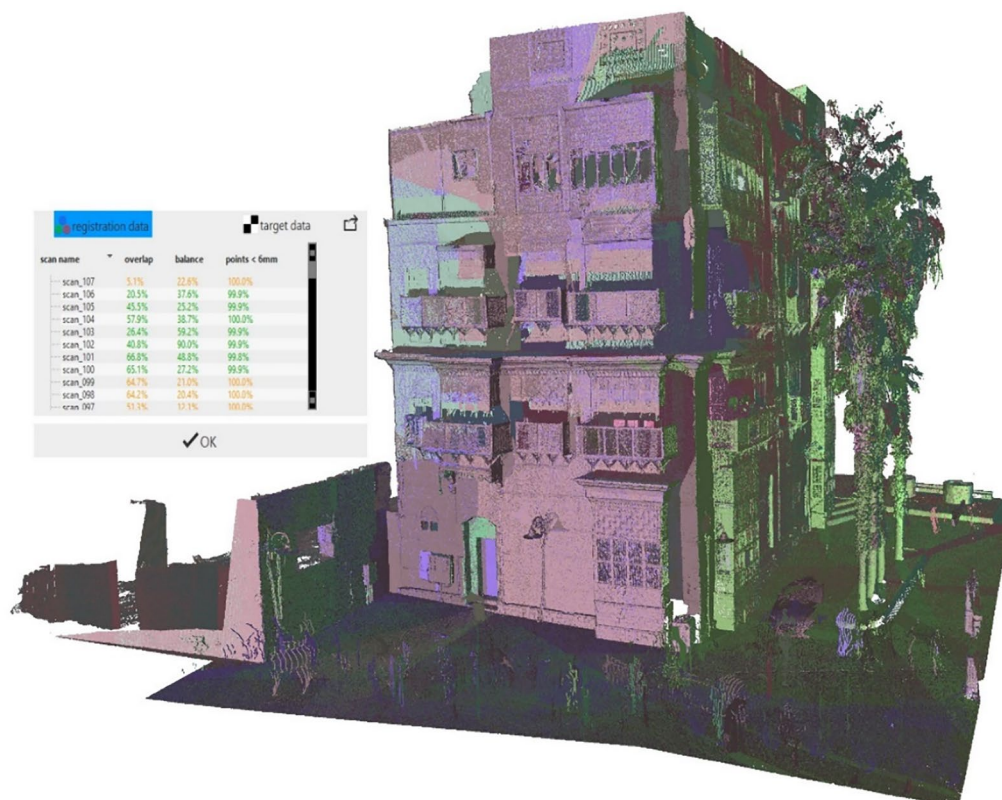


Fig. 7 Merging the raw scans into a single coordinate system with the scan registration report

the DJI Phantom 4. The rotary wings of the drone allow for takeoffs and landings in constrained areas. It weighs 1380 g and has a 48 mega pixel (8000×6000) digital camera with wide-angle lenses and a focal length of 35 mm. Almost all of the photos were taken within a specific time frame to ensure proper lighting.

The placement of cameras within the flight plan is critical to the success of photogrammetry processing. The camera in our project took images in a circle at regular intervals, with 80% overlap. The on-board GPS antenna's positioning data was georeferenced using world geodetic system (WGS84). Fixing the camera focal length during data collection allows for efficient results when using the structure from motion algorithm (SfM), while the camera was self-calibrated during subsequent data processing. The UAV data captured 47 images with a manually controlled trajectory. The UAV photogrammetric pipeline typically consisted of several stages. In our project, the Agi soft software was used to compute the camera positions using 48 thousand tie points as shown in Fig. 10a. The point cloud for each overlapping pixel was then densified using an automatic matching algorithm. Figure 10b illustrates the 248 million points that the densify results produced. Figure 11 depicts the final 3D textured model, which shows that the missing TLS data in

the occluded areas of the building's upper part have been sampled with UAV point clouds. The tall palm trees surrounding the building and the wind's ability to alter the UAV flight path present the main obstacles during data collection.

Data fusion (cloud to cloud registration)

By combining data from UAVs and TLS, the proposed workflow, shown in Fig. 12, aims to improve the quantity and quality of the collected point clouds that will be used for HBIM modeling. The UAV point cloud improved the TLS dataset by filling gaps and missing points in shadow and occluded areas, especially on upper levels of buildings. Furthermore, the UAV's ability to capture high-resolution imagery helps to create highly realistic orthophotos that are georeferenced to the TLS model. The 48-megapixel camera-generated orthophoto is used to interpret cracks and linear surface features, as well as quantify surface decay, using deep learning algorithms. Aligning and converting TLS and UAV point clouds is the most difficult aspect of combining data from multiple sensors. Typically, registration occurred in two ways: manually or automatically [79]. Research into fully automatic, reliable co-registration of TLS and



Fig. 8 a The final model of the building b long section c Cross section d Interior scanning

photogrammetric-based point clouds is still ongoing and faces numerous challenges [80, 81].

In our experiment, the homologous points are manually defined using a set of corresponding building points to calculate the necessary translation and rotation parameters between the two clouds, as shown in Fig. 13. Following the registration process, the true orthophoto for the building's main facades is then produced using the meshed and georeferenced model as shown in Fig. 14. The building photogrammetric orthophoto was created to improve the visual interpretation of damaged areas and to draw the exact contour of the cracks with deep learning algorithms, as will be discussed in the following sections.

HBIM and creating library of architectural elements

The TLS point cloud model of the Al Radwan house, as well as the generated true orthophoto, were imported into Autodesk Revit, a standalone Building Information

Modelling (BIM) software, as depicted in Fig. 15a. Revit is a robust and adaptable modelling tool that facilitates the creation of detailed parametric models. However, the available libraries that meet the specific needs of heritage BIM projects are limited. Therefore, it was necessary to model Al Radwan building elements and create a new library for the Hijazi architectural style.

The geometric survey, performed using Terrestrial Laser Scanning (TLS), produced a high-resolution point cloud model of the Al Radwan house. The point cloud was used to extract precise dimensions and spatial relationships between architectural elements. Scan-to-BIM is the process of converting the scanned data (point cloud) into a BIM model. In this project, the imported point cloud served as the basis for creating accurate and detailed models of the building's architectural elements. The process involved importing the point cloud, aligning it within the Revit environment to match the building's actual orientation and dimensions, and then using it as

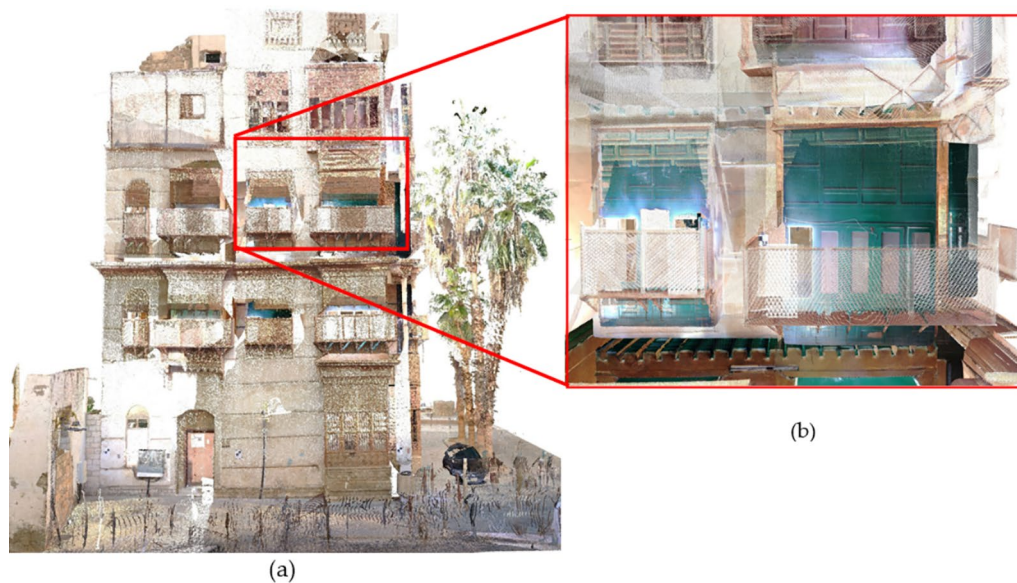


Fig. 9 **a** Multiple scans with varying lighting conditions disrupt the appearance of the façade. **b** Unsampld roof and upper zone facades obscured from TLS



Fig. 10 UAV **a** camera position and the initial tie points **b** point cloud densification

a guide to trace and model the geometric features of the building.

Two phases were defined for the BIM modelling of the Al Radwan house. The first phase focused on modelling basic architectural elements such as walls, columns, and simple façades. These elements are typically uniform and regular in shape, making them easier to model accurately. The second phase aimed at creating new families of repeating Hijazi architectural elements, such as windows, doors, arches, and Roshans (traditional wooden lattice windows). These elements were modelled with shared attributes to maintain consistency and accuracy.

The initial step in the modelling process involved establishing different levels, depicted in Fig. 15b. In BIM terminology, levels are horizontal planes that act as

references for placing elements such as floors, ceilings, and roofs. Levels are essential for organizing the vertical components of a building and ensuring accurate placement and alignment. Next, various longitudinal and cross sections were established at different levels to define and locate the positions and dimensions of architectural objects such as doors, columns, and walls. Sections are vertical cuts through the building that provide detailed views of its internal structure, allowing precise placement of elements.

To model the standard architectural elements, depicted in Fig. 16a, closed polygons were used to outline the features within the point cloud. Closed polygons help in defining the boundaries of these features accurately. The wall thickness, column dimensions, and window



Fig. 11 Final photogrammetric model of a UAV with texture

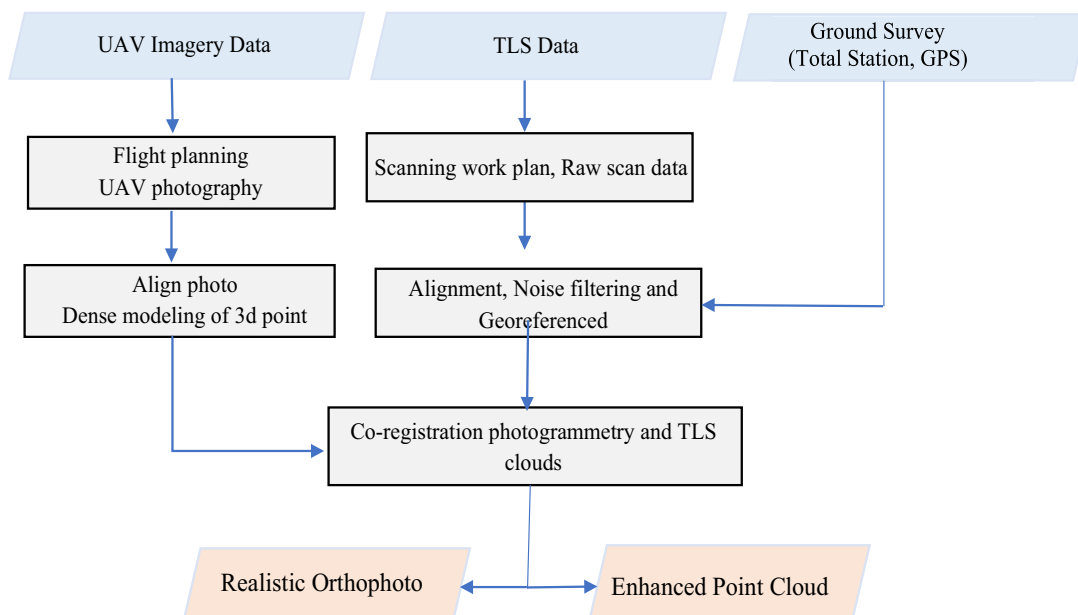


Fig. 12 Data fusion workflow

measurements (length, width, and depth) were critical for creating the parametric functions that allow for the modification of each component. Parametric functions enable changes to the model’s geometry by adjusting the parameters, ensuring flexibility and accuracy in the design process.

Due to the absence of unique designs for many historical building components in existing BIM libraries, new architectural component families were created. In

Revit, a family is a group of elements with a common set of properties and a related graphical representation. For heritage buildings, loadable families were created to capture the unique and intricate details of the Hijazi architectural elements. Loadable families are reusable components that can be loaded into different Revit projects. The creation process included modelling unique elements, depicted in Fig. 16b, such as Roshans, arches, and decorative columns as separate family files, which

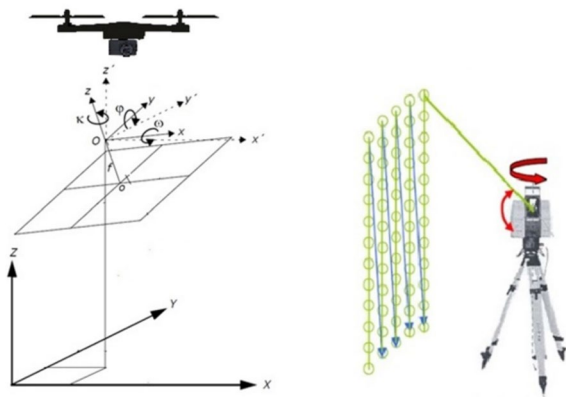


Fig. 13 TLS-UAV registration parameters

can then be imported into any project where these elements are needed. These families were designed with parametric capabilities, allowing for adjustments in size, shape, and other attributes to fit various contexts and requirements within the project.

The new families were designed to be parametric, meaning that their dimensions and properties could be easily adjusted to suit different applications. For instance, a Roshan family could be resized to fit various window

openings by changing its height, width, and depth parameters. This parametric design ensures that the elements can be modified without redrawing them from scratch, saving time and maintaining consistency across different projects. Parameters were set up within the family editor in Revit, where geometric constraints and formulas could be applied to control the behaviour of the family components.

Although loadable families are tailored to specific project requirements, they can be adapted and reused in other heritage projects in historic Jeddah. The newly created Hijazi architectural libraries are exportable, allowing the custom families to be exported from the Al Radwan project and imported into other Revit projects. This facilitates their reuse in modelling other heritage buildings in Jeddah. Since the families are parametric, they can be easily adjusted to fit different buildings with similar architectural styles, ensuring consistency and efficiency in modelling Hijazi architecture across multiple projects.

The new libraries were linked to the Heritage Building Information Modelling (HBIM) database, allowing users to modify the various parametric elements of the building components. This integration provides an efficient solution for modelling the repetitive Hijazi architecture style, ensuring that the models are both accurate and easily



Fig. 14 True orthophoto using UAV photogrammetry

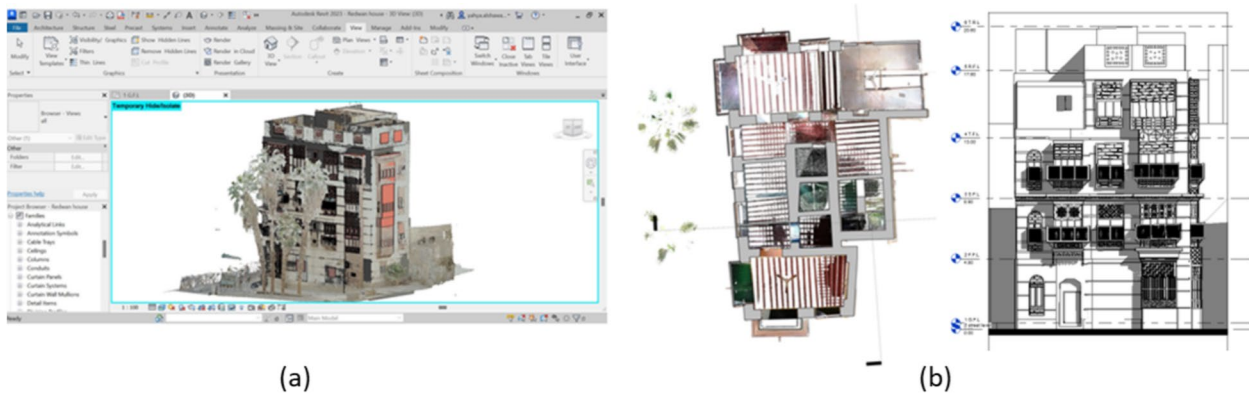


Fig. 15 **a** upload the point clouds in BIM platform. **b** Segment the point clouds for modeling

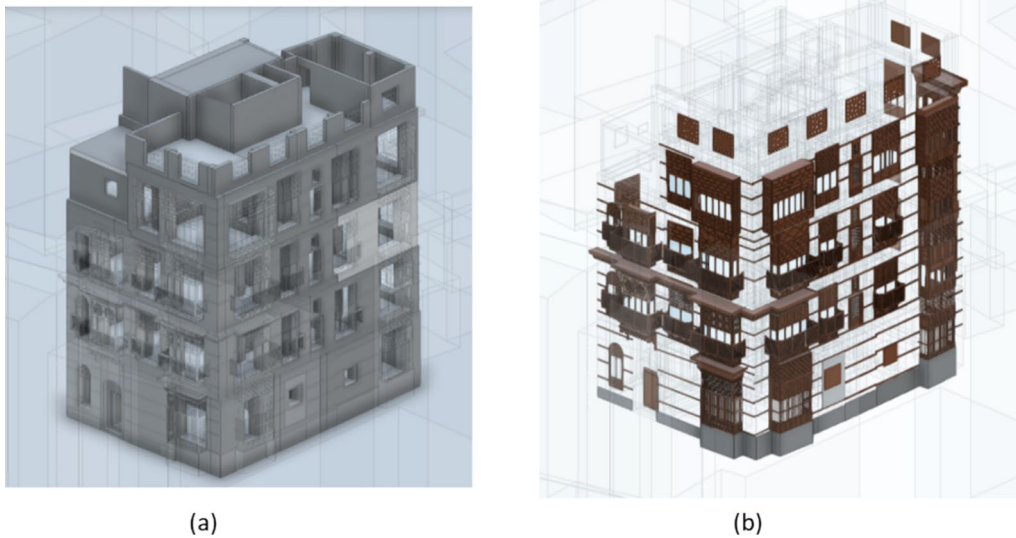


Fig. 16 **a** Modeling standard architectural elements **b** Create new families for Hijazi architecture



Fig. 17 BIM Model of Al Radwan historical Building

modifiable. Figure 17 depicts the final BIM model with all architectural elements, while Fig. 18 shows the BIM model in its georeferenced position relative to other historical structures in old Jeddah. Georeferencing ensures



Fig. 18 The final BIM of Al Radwan house georeferenced with other historic buildings in Jeddah city

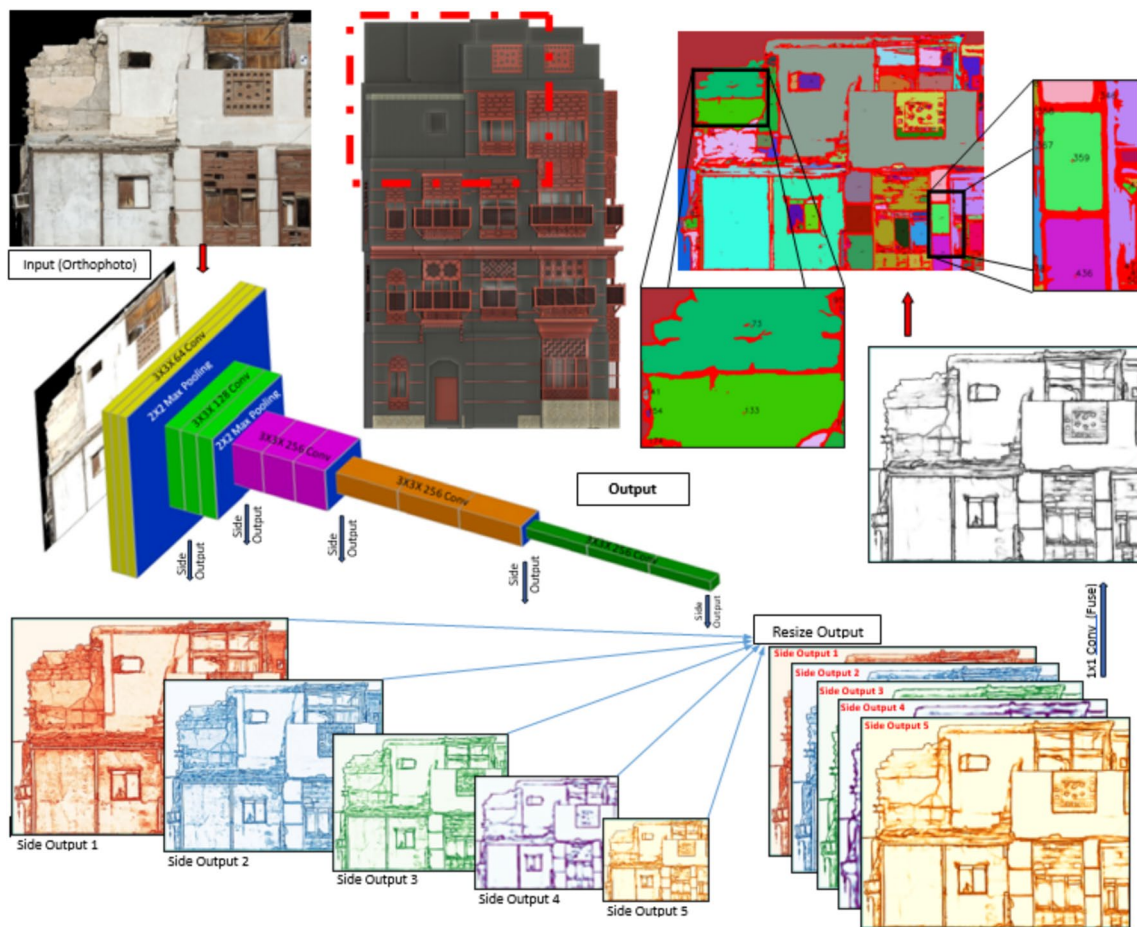


Fig. 19 detailed overview of the HED architecture

Table 1 A part of statistical results of labeled segments

Label	Area (cm ²)	Equivalent diameter (cm)	Perimeter (cm)	Mean intensity in red band	Mean intensity in green band	Mean intensity in blue band	Solidity
1	778.2	31.5	151.6	192.6	211.0	220.4	0.85
3	445.4	23.8	81.1	174.6	193.8	201.8	0.98
4	2332.0	54.5	409.1	151.2	167.6	177.6	0.58
8	30.5	6.2	26.1	80.9	92.3	106.5	0.87
9	64.8	9.1	31.4	193.2	212.8	224.5	0.92
10	609.1	27.8	118.8	91.4	104.0	112.9	0.89
11	38.1	7.0	33.8	85.7	103.5	120.5	0.77

that the BIM model is accurately positioned within the real-world coordinates, facilitating better analysis and integration with other geographic data.

Deep learning process

Using deep learning in point clouds offers many advantages but also presents several disadvantages and challenges includes sensitivity to noise and outliers, high

computational complexity, large data requirements, and complexity of data representation. On the other hand, applying deep learning to structured 2D ortho-photos produces segmentation results in metric units and helps overcome the limitations of using non-metric photos and 3D point clouds. In this approach, a deep learning algorithm is used to process the true ortho-photos UAV photogrammetric data integrated with

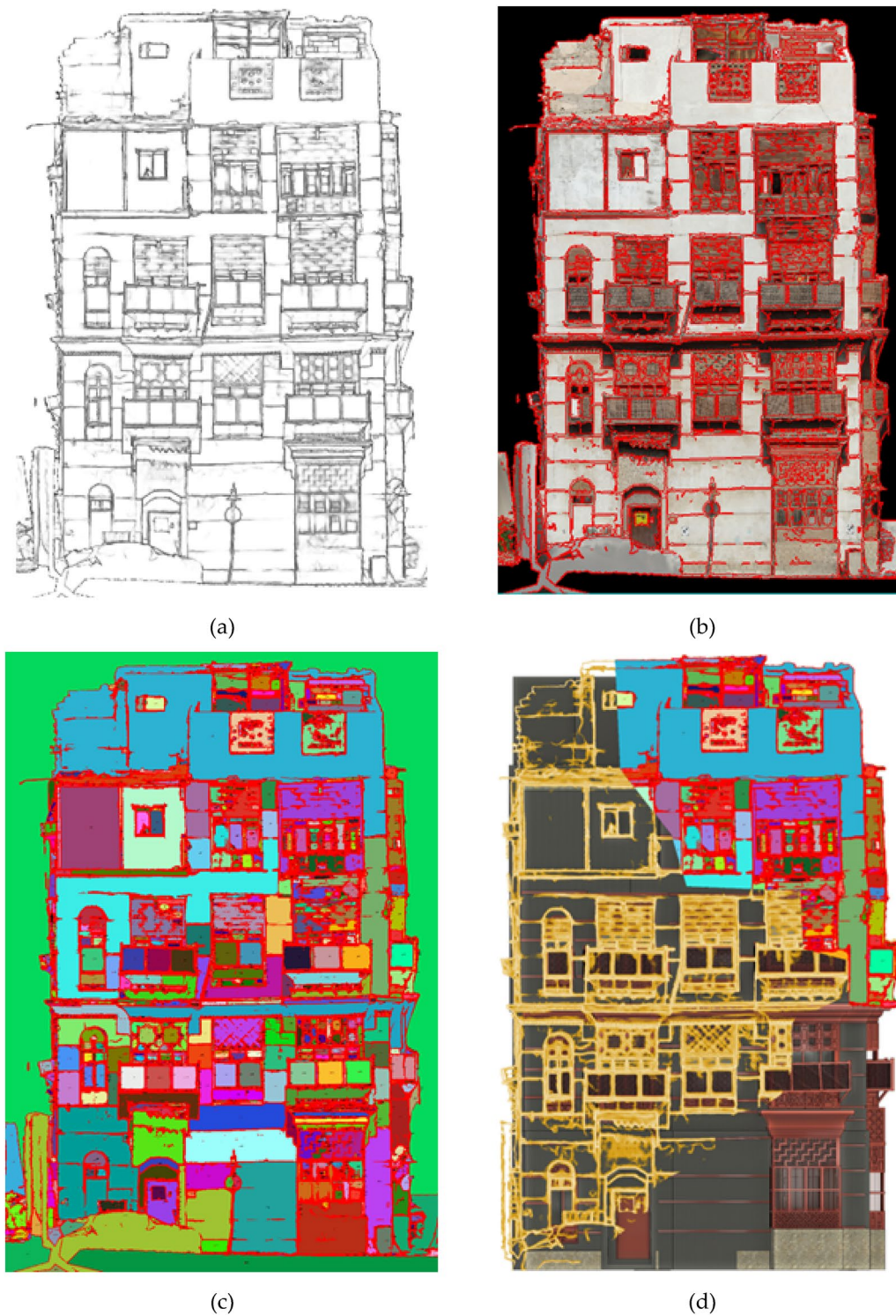


Fig. 20 **a** The HED model output; **b** The projection on orthophoto; **c** The main façade segmentation; **d** The enriched HBIM of Al Radwan house

TLS clouds. True orthophotos offer numerous advantages, including metric accuracy and radiometric information, which can be used for both quantitative and qualitative analyses of various data sets.

As mentioned in Sect. "Data Acquisition and Pre-processing", the production of the orthophotos of the main facades of Al Radwan House were generated and used as input for the HED model. OpenCV's deep

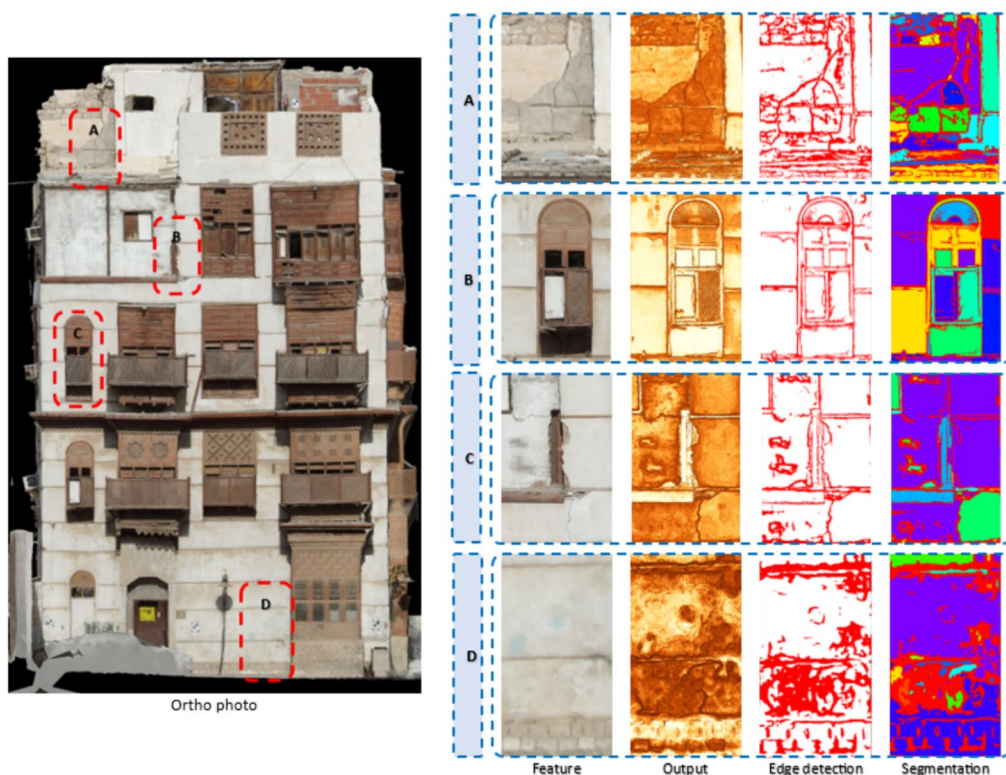


Fig. 21 The graphical representation of features on the main façade of Al Radwan house

neural network module was utilized for image pre-processing and preparing them for segmentation using pre-trained deep learning models. The HED model was implemented using Python 3.9, with the neural network weight calculated by [76] on a machine of 8 G B of RAM, processor i5 11th generation and a graphic card Nvidia GeForce.

As mentioned in Sect. "[Deep Learning Model: Holistically nested edge detection \(HED\)](#)", implementing the HED model involves several steps, from setting up the environment to training and testing the model. In setting up the environment, the necessary libraries must be installed. In our case, Caffe and OpenCV libraries are installed. The next step is to download the pre-trained HED model and the corresponding text file that contains the weights of the model. In the data preparation step, the images undergo preprocessing steps, including cropping, normalization, constructing a blob, and presenting the data in a format that can be read by OpenCV. Finally, the pre-trained HED model is loaded, and edge detection is performed on the images. In the post-processing step, a reasonable threshold should be applied to the predicted edge map to obtain binary edge maps. Figure 19 illustrates the entire process of applying the HED model for image segmentation.

In the figure, it can be observed that at the end of each step of the model, a side output is extracted based on the scale. The connected component step is performed after resizing each side output to the original scale, considering the features. Gaussian blur and Otsu's threshold in OpenCV are used to connect the components. After filtering out small objects, the remaining components are labeled. Statistical analysis can then be conducted for each labeled segment, including calculations for area, equivalent diameter, perimeter, mean intensity, and solidity. Table I presents a partial summary of the statistical results for the labeled segments.

Figure 20 showcases the main façade of Al Radwan house, along with the output of the HED model and the final segmentation of the orthophoto. More detailed information can be found in Fig. 21, which highlights selected features. The vector data obtained from the output segmentation will be mapped into the HBIM geometry to enhance the model for conservation and monitoring purposes, As shown in Fig. 22. The generated orthophoto and the HBIM surfaces are of identical dimensions, allowing for accurate and consistent texture and deep learning outcomes that warp throughout the BIM geometry. The findings provide a clear interpretation of surface features and are a useful tool for accurately



Fig. 22 Mapping realistic texture and learning segmentation results to HBIM

depicting the various types, characteristics, and spatial distributions of façade degradation.

Discussion

Surface cracks and material deterioration must be detected and monitored on a regular basis in built heritage. Although the scanner’s model contains high resolution and density points that represent surface features, the mixed pixel effect, which appears in the crack borders and edge outline, can have an impact on the coordinates of the collected data. This issue involves incorrect data interpretation and crack drawing during HBIM tracing and modeling. For these reasons, Lanzara et al. [82] and Conti et al. [83] used orthophoto as a visual aid for digitization of surface cracks and deterioration during the HBIM modeling, whereas Others included the images and orthophotos as supplementary information for the preservation of cultural heritage with the BIM objects [24, 84]. Our workflow proposed a deep learning approach for automatically detecting surface features and cracks with informative attributes from orthophotos created by combining TLS and UAV imagery data. Area, premier, and length are examples of attributes that can be used in HBIM modeling to monitor heritage and plan conservation efforts. These surface feature attributes were usually measured using gypsum strips, mechanical extensometers, and electrical sensors, all of which provide discrete point measurements at specific locations and require long-term bonding to the building surface. The proposed methodology used deep learning on true

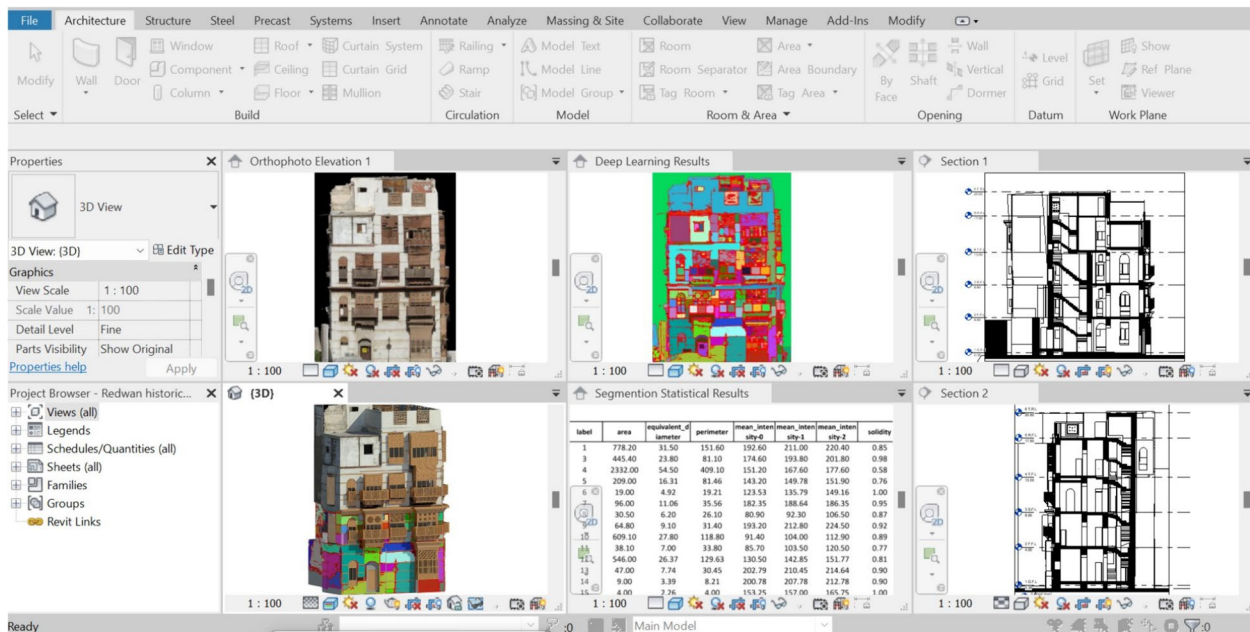


Fig. 23 Final HBIM enriched with deep learning results

ortho images rather than point clouds. This method was chosen to address specific challenges associated with semantic segmentation of point clouds, such as point discontinuity and the low reliability of extracting parametric objects alone. Furthermore, using deep learning on true ortho images helped overcome noise and outlier issues, which are common when working with point clouds [12]. Concerning processing time, we initiated the process using pre-trained weights of the (HED) network. These weights were obtained from a benchmark dataset, thus saving considerable time compared to manually labeling numerous scenes. Unlike most approaches that use deep learning models, this approach eliminated the need to wait for the training phase, which typically spans from a few hours to several days. The results obtained from applying the proposed methodology demonstrated its effectiveness in accurately identifying non-parametric parts, irregular shapes, and free-form elements as can be depicted in Fig. 21. Moreover, the methodology facilitated the determination of their specific characteristics, such as area and perimeter. This successful outcome addressed the challenges faced by previous studies [34] when identifying features like doors, arches, and windows.

Engineers, conservators, architects, and virtual tourists will benefit greatly from having an HBIM database that includes color information and parametric geometry. Our method produces realistic HBIM texture mapping by combining UAV and laser point data to create an orthophoto. Their warping over HBIM geometry, as well as the deep learning segmentation results, will be accurate and reliable, as shown in the Fig. 22, because the orthophoto will have the same dimensions as the corresponding HBIM surfaces. As illustrated in Fig. 23, contextual, statistical, and other derived data are linked into the Historic Building Information Model (HBIM) to offer a unified platform for all parties participating in the conservation process.

The most difficult challenges during data processing were massive data files that consumed computer capacity and took days to process. Another issue that arose when using a pre-trained model was the discrepancy between the images it was trained on and archaeological building facades. Furthermore, the model was trained on images with dimensions of 224×224 pixels, so using different image dimensions may result in incorrect results. Thus, some images were divided into smaller patches and processed separately. The outputs were then reassembled to produce accurate and complete results.

Conclusion

The study was motivated by the need for an efficient workflow for as-built BIM recording of large and complex historical buildings, as well as automated feature extractions to produce reliable and as-built BIM models. The fusion-based data sensor uses TLS and photogrammetry to produce realistic UAV orthophotos that are then used with a deep learning algorithm to automatically detect and draw the outline of cracks and surface features using the appropriate statistical parameters. The data is then remapped back into the BIM model for analysis of all surface features and weathering forms, including their extent, typology, cause, and conservation needs. The results can be used to guide conservation efforts and develop adaptable, successful monitoring programs that assess the extent of damage and predict the defect's potential spread. Further research will focus on using the presented approach to improve and optimize semantic segmentation on 3D data for heritage applications.

Acknowledgements

The authors gratefully acknowledge the technical and financial support provided by the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

Author contributions

Conceptualization, Y.M. and Y.A.; methodology, Y.M. and Y.A.; software, Y.M. and Y.A.; validation, Y.M. and Y.A. and A.B.; formal analysis, Y.M.; investigation, Y.A.; resources, A.B.; data curation, Y.M.; writing—original draft preparation, Y.M. and Y.A.; writing—review and editing, Y.M. and Y.A. and A.B.; visualization, A.B. All authors have read and agreed to the published version of the manuscript.

Funding

This research work was funded by Institutional Fund Projects under Grant No. IFPIP-1203–137-1443 provided by the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

Availability of data and materials

The data that support the findings of this study are available on request from the corresponding author.

Declarations

Competing interest

The authors declare no competing interests.

Received: 15 May 2024 Accepted: 18 July 2024

Published online: 26 July 2024

References

1. Costantino D, Pepe M, Restuccia AG. Scan-to-HBIM for conservation and preservation of cultural Heritage building: the case study of San Nicola in Montedoro church (Italy). *Appl Geomat*. 2023;15:607–21. <https://doi.org/10.1007/s12518-021-00359-2>.
2. Baik A. The use of interactive virtual bim to boost virtual tourism in heritage sites Historic Jeddah. *ISPRS Int J Geoinf*. 2021. <https://doi.org/10.3390/ijgi10090577>.
3. Di Stefano F, Gorreja A, Malinverni ES, Mariotti C. Knowledge modeling for heritage conservation process: from survey to HBIM implementation.

2020. *Inter Arch Photogramm Remote Sens Spatial Inf Sci*. <https://doi.org/10.5194/isprs-archives-XLIV-4-W1-2020-19-2020>.
4. Martinelli L, Calcerano F, Adinolfi F, Chianetta D, Gigliarelli E. Open HBIM-IoT monitoring platform for the management of historical sites and museums an application to the bourbon royal site of Carditello. *Int J Arch Heritage*. 2023. <https://doi.org/10.1080/15583058.2023.2272130>.
 5. Bouzas Ó, Cabaleiro M, Conde B, Cruz Y, Riveiro B. Structural health control of historical steel structures using HBIM. *Autom Constr*. 2022;140: 104308.
 6. López F, Leronés P, Llamas J, Gómez-García-Bermejo J, Zalama E. A review of heritage building information Modeling (H-BIM). *Multimodal Technol Int*. 2018;2:21.
 7. Murphy M, McGovern E, Pavia S. Historic building information modelling (HBIM). *Struct Surv*. 2009;27:311–27. <https://doi.org/10.1108/02630800910985108>.
 8. Andriasyan M, Moyano J, Nieto-Julián JE, Antón D. From point cloud data to building information modelling: an automatic parametric workflow for heritage. *Remote Sens*. 2020;12:1094.
 9. Aricò M, Lo BM. From scan-to-BIM to heritage building information modelling for an ancient Arab-Norman church. 2022. *Int Arch Photogramm Remote Sens Spatial Inf Sci*. <https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-761-2022>.
 10. Alonso SJ, Rubio M, Martín F, Fernández G. comparing time-of-flight and phase-shift. the survey of the royal pantheon in the basilica of san Isidoro (león). <http://www.uva.es/davap> Accessed 20 Jun 2024.
 11. Barrile V, Fotia A, Bilotta G. Geomatics and augmented reality experiments for the cultural heritage. *Appl Geomat*. 2018;10:569–78.
 12. Gines JLC, Cervera CB. Toward hybrid modeling and automatic planimetry for graphic documentation of the archaeological heritage: the Cortina family pantheon in the cemetery of Valencia. *Int J Arch Herit*. 2020;14:1210–20.
 13. Chiabrandò F, Sammartano G, Spanò A, Spreafico A. Hybrid 3D models: when geomatics innovations meet extensive built heritage complexes. *ISPRS Int J Geoinf*. 2019;8:124.
 14. Costantino D, Rossi G, Pepe M, Leserri M. Experiences of TLS, terrestrial and UAV photogrammetry in Cultural Heritage environment for restoration and maintenance purposes of Royal Racconigi castle, Italy. In: *Proceedings of the 2022 IMEKO TC4 International Conference on Metrology for Archaeology and Cultural Heritage*. Budapest: IMEKO; 2023 pp. 438–43.
 15. Alshwabkeh Y, Baik A, Miky Y. Integration of laser scanner and photogrammetry for heritage BIM enhancement. *ISPRS Int J Geoinf*. 2021;10:316.
 16. Themistocleous K, Agapiou A, Hadjimitsis D. 3D Documentation and BIM modeling of cultural heritage structures using uavs: the case of the foinkaria church. 2016. *Int Arch Photogramm Remote Sens Spatial Inf Sci*. <https://doi.org/10.5194/isprs-archives-XLII-2-W2-45-2016>.
 17. Mikita T, Balková M, Bajer A, Cibulka M, Patočka Z. Comparison of different remote sensing methods for 3D modeling of small rock outcrops. *Sensors*. 2020;20:1663.
 18. Tscharf A, Rumpler M, Fraundorfer F, Mayer G, Bischof H. On the use of uavs in mining and archaeology GEO-accurate 3d reconstructions using various platforms and terrestrial views. 2015. *ISPRS Ann Photogramm Remote Sens Spatial Inf Sci*. <https://doi.org/10.5194/isprs-annals-II-1-W1-15-2015>.
 19. Rocha M, Fernández F. A scan-to-BIM methodology applied to heritage buildings. *Heritage*. 2020;3:47–67.
 20. Castilla FJ, Ramón A, Adán A, Trenado A, Fuentes D. 3D sensor-fusion for the documentation of rural heritage buildings. *Remote Sens (Basel)*. 2021;13:1337.
 21. Yaagoubi R, Miky Y. Developing a combined light detecting and ranging (LiDAR) and building information modeling (BIM) approach for documentation and deformation assessment of historical buildings. *MATEC Web of Conf*. 2018;149:02011.
 22. Cuca B, Barazzetti L. Damages from extreme flooding events to cultural heritage and landscapes: water component estimation for Centa River (Albenga, Italy). *Adv Geosci*. 2018;45:389–95.
 23. Fregonese L, Taffurelli L, Adami A, Chiarini S, Cremonesi S, Helder J, et al. Survey and modelling for the BIM of basilica of san marco in venice international archives of the photogrammetry remote sensing and spatial information sciences ISPRS archives. 2017. *Int Soc Photogramm Remote Sens*. <https://doi.org/10.5194/isprs-archives-XLIII-2-W3-303-2017>.
 24. Brumana R, Condoleo P, Grimoldi A, Banfi F, Landi AG, Previtali M. HR LOD based HBIM to detect influences on geometry and shape by stereotomic construction techniques of brick vaults. *Appl Geomat*. 2018;10:529–43.
 25. Sánchez-Aparicio LJ, del Blanco-García FL, Mencías-Carrizosa D, Villanueva-Llauradó P, Aira-Zunzunegui JR, Sanz-Arauz D, et al. Detection of damage in heritage constructions based on 3D point clouds a systematic review. *J Buil Eng*. 2023;77:107440.
 26. Yang S, Hou M, Li S. Three-dimensional point cloud semantic segmentation for cultural heritage: a comprehensive review. *Remote Sens*. 2023;15:548.
 27. Maalek R, Lichti DD, Ruwanpura JY. Automatic recognition of common structural elements from point clouds for automated progress monitoring and dimensional quality control in reinforced concrete construction. *Remote Sens*. 2019;11:1102.
 28. Oliveira H, Correia P. Automatic road crack segmentation using entropy and image dynamic thresholding. 2009
 29. Talab A, Huang Z, Xi F, HaiMing L. Detection crack in image using Otsu method and multiple filtering in image processing techniques. *Optik Int J Light Elec Optics*. 2015. <https://doi.org/10.1016/j.jilleo.2015.09.147>.
 30. Mishra R, Chandrakar C, Mishra RS. Surface defects detection for ceramic tiles using image processing and morphological techniques. 2012. <https://api.semanticscholar.org/CorpusID:17740783> Accessed 7 Apr 2024.
 31. Abdel-Qader I, Abudayyeh O, Kelly ME. Analysis of edge-detection techniques for crack identification in bridges. *J Comput Civ Eng*. 2003;17:255–63.
 32. Salman M, Mathavan S, Kamal K, Rahman M. Pavement crack detection using the Gabor filter. In: *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*. Las Vegas: IEEE; 2013 pp. 2039–44.
 33. Medina R, Llamas J, Gómez-García-Bermejo J, Zalama E, Segarra M. Crack detection in concrete tunnels using a gabor filter invariant to rotation. *Sensors*. 2017;17:1670.
 34. Matrone F, Grilli E, Martini M, Paolanti M, Pierdicca R, Remondino F. Comparing machine and deep learning methods for large 3d heritage semantic segmentation. *ISPRS Int J Geoinf*. 2020;9:535.
 35. Billi D, Croce V, Bevilacqua MG, Caroti G, Pasqualetti A, Piemonte A, et al. Machine learning and deep learning for the built heritage analysis: laser scanning and UAV-based surveying applications on a complex spatial grid structure. *Remote Sens*. 2023;15:1961.
 36. Lecun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Proc IEEE*. 1998;86:2278–324.
 37. Pal KK, Sudeep KS. Preprocessing for image classification by convolutional neural networks. In: *2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*. Las Vegas: IEEE; 2016. p. 1778–81.
 38. Jmour N, Zayen S, Abdelkrim A. Convolutional neural networks for image classification. *Int Conf Adv Syst Elec Technol*. 2018;2018:397–402.
 39. Basha SHS, Dubey SR, Pulabaigari V, Mukherjee S. Impact of fully connected layers on performance of convolutional neural networks for image classification. *Neurocomputing*. 2020;378:112–9.
 40. Kang K, Ouyang W, Li H, Wang X. Object detection from video tubelets with convolutional neural networks. In: *2016 IEEE conference on computer vision and pattern recognition (CVPR)*. Las Vegas: IEEE; 2016. pp. 817–25.
 41. Galvez RL, Bandala AA, Dadios EP, Vicerra RRP, Maningo JMZ. Object detection using convolutional neural networks. *tencon 2018—2018 IEEE region 10 conference*. Las Vegas: IEEE; 2018. pp. 2023–7.
 42. Li K, Ma W, Sajid U, Wu Y, Wang G. Object Detection with Convolutional Neural Networks. 2019. <http://arxiv.org/abs/1912.01844> Accessed 7 Apr 2024.
 43. Kampffmeyer M, Salberg A-B, Jenssen R. Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks. In: *2016 IEEE conference on computer vision and pattern recognition workshops (CVPRW)*. Las Vegas: IEEE; 2016. p. 680–8.
 44. Mohammdimanesh F, Salehi B, Mahdianpari M, Gill E, Molinier M. A new fully convolutional neural network for semantic segmentation of polarimetric SAR imagery in complex land cover ecosystem. *ISPRS J Photogramm Remote Sens*. 2019;151:223–36.

45. Yang H, Yu B, Luo J, Chen F. Semantic segmentation of high spatial resolution images with deep neural networks. *ISPRS Remote Sens.* 2019;56:749–68.
46. Rogers MSJ, Bithell M, Brooks SM, Spencer T. VEdge_detector: automated coastal vegetation edge detection using a convolutional neural network. *Int J Remote Sens.* 2021;42:4805–35.
47. Cha Y, Choi W, Büyükköztürk O. Deep learning-based crack damage detection using convolutional neural networks. *Comput Aided Civil Infrast Eng.* 2017;32:361–78.
48. Zhao X, Li S. Convolutional neural networks-based crack detection for real concrete surface. In: Sohn H, editor. *Sensors and smart structures technologies for civil, mechanical, and aerospace systems 2018*. Washington: SPIE; 2018. p. 143.
49. Dung CV, Anh LD. Autonomous concrete crack detection using deep fully convolutional neural network. *Autom Constr.* 2019;99:52–8.
50. Ni F, Zhang J, Chen Z. Zernike-moment measurement of thin-crack width in images enabled by dual-scale deep learning. *Comput-Aided Civil Infrast Eng.* 2019;34:367–84.
51. Kang D, Benipal SS, Gopal DL, Cha Y-J. Hybrid pixel-level concrete crack segmentation and quantification across complex backgrounds using deep learning. *Autom Constr.* 2020;118: 103291.
52. Wei F, Yao G, Yang Y, Sun Y. Instance-level recognition and quantification for concrete surface bughole based on deep learning. *Autom Constr.* 2019;107: 102920.
53. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. United States: Springer; 2015.
54. Badrinarayanan V, Kendall A, Cipolla R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell.* 2017;39:2481–95.
55. Kendall A, Badrinarayanan V, Cipolla R. Bayesian SegNet: Model Uncertainty in Deep Convolutional Encoder-Decoder Architectures for Scene Understanding. In: *Proceedings of the British Machine Vision Conference 2017*. British Machine Vision Association. 2017.
56. Chaurasia A, Culurciello E. LinkNet exploiting encoder representations for efficient semantic segmentation. In: *2017 IEEE visual communications and image processing (VCIP)*. Las Vegas: IEEE; 2017. p. 1–4.
57. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *2016 IEEE conference on computer vision and pattern recognition (CVPR)*. Las Vegas: IEEE; 2016. p. 770–8.
58. Haznedar B, Bayraktar R, Ozturk AE, Arayici Y. Implementing pointnet for point cloud segmentation in the heritage context. *Herit Sci.* 2023;11:2.
59. Szegedy C, Ioffe S, Vanhoucke V, Alemi A. Inception-v4, inception-Resnet and the impact of residual connections on learning. *Proc AAAI Conf Art Int.* 2017. <https://doi.org/10.1609/aaai.v31i1.11231>.
60. Milletari F, Navab N, Ahmadi S-A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In: *2016 Fourth International Conference on 3D Vision (3DV)*. Las Vegas: IEEE; 2016. p. 565–71.
61. Loverdos D, Sarhosis V. Automatic image-based brick segmentation and crack detection of masonry walls using machine learning. *Autom Constr.* 2022;140: 104389.
62. Liu F, Shen C, Lin G, Reid I. Learning depth from single monocular images using deep convolutional neural fields. *IEEE Trans Pattern Anal Mach Intell.* 2016;38:2024–39.
63. Yosinski J, Clune J, Bengio Y, Lipson H. How transferable are features in deep neural networks. *Adv Neural Inf Proc Syst (NIPS)*. 2014;27:3320.
64. Liu Y, Lew MS. Learning Relaxed deep supervision for better edge detection. In: *2016 IEEE Conference on computer vision and pattern recognition (CVPR)*. Las Vegas: IEEE; 2016. pp. 231–40.
65. Sarkar S, Venugopalan V, Reddy K, Ryde J, Jaitly N, Giering M. Deep learning for automated occlusion edge detection in RGB-D frames. *J Signal Process Syst.* 2017;88:205–17.
66. Yu Z, Feng C, Liu M-Y, Ramalingam S. CASENet: deep category-aware semantic edge detection. In: *2017 IEEE Conference on computer vision and pattern recognition (CVPR)*. Las Vegas: IEEE; 2017. p. 1761–70.
67. Liu Y, Cheng M-M, Fan D-P, Zhang L, Bian J-W, Tao D. Semantic edge detection with diverse deep supervision. *Int J Comput Vis.* 2022;130:179–98.
68. Muntarina K, Mostafiz R, Khanom F, Shorif SB, Uddin MS. MultiResEdge: a deep learning-based edge detection approach. *Intell Syst Appl.* 2023;20: 200274.
69. Huang Z, Pan Z, Lei B. Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data. *Remote Sens.* 2017;9:907.
70. Alinsaif S, Lang J. Histological image classification using deep features and transfer learning. In: *2020 17th Conference on Computer and Robot Vision (CRV)*. Las Vegas: IEEE; 2020. p. 101–8.
71. Shin H-C, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans Med Imaging.* 2016;35:1285–98.
72. Le Van C, Nguyen Bao L, Puri V, Thanh Thao N, Le D-N. Detecting lumbar implant and diagnosing scoliosis from Vietnamese x-ray imaging using the pre-trained API models and transfer learning. *Comput Mater Continua.* 2020;66:17–33.
73. Ahmed S, Shaikh A, Alshahrani H, Alghamdi A, Alrizq M, Baber J, et al. Transfer learning approach for classification of histopathology whole slide images. *Sensors.* 2021;21:5361.
74. Xu Y, Jia Z, Wang L-B, Ai Y, Zhang F, Lai M, et al. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinform.* 2017;18:281.
75. Zhang Y, Zhang Z, Zhao W, Li Q. Crack segmentation on earthen heritage site surfaces. *Appl Sci.* 2022;12:12830.
76. Ashukid Ashutosh Chandra. HED model weights. <https://github.com/ashukid/hed-edge-detector>. Accessed 7 Apr 2024.
77. Kokkinos I. Pushing the Boundaries of Boundary Detection using Deep Learning. 2015. <http://arxiv.org/abs/1511.07386> Accessed 7 Apr 2024.
78. Xie S, Tu Z. Holistically-Nested Edge Detection. 2015.
79. Zhang J, Lin X. Advances in fusion of optical imagery and LiDAR point cloud applied to photogrammetry and remote sensing. *Int J Image Data Fusion.* 2017;8:1–31.
80. Xu N, Qin R, Song S. Point cloud registration for LiDAR and photogrammetric data: a critical synthesis and performance analysis on classic and deep learning algorithms. *ISPRS Open J Photogramm Remote Sens.* 2023. <https://doi.org/10.1016/j.ojphoto.2023.100032>.
81. Partovi T, Dähne M, Maboudi M, Krueger D, Gerke M. Automatic integration of laser scanning and photogrammetric point clouds: from acquisition to co-registration. 2021. *Inter Arch Photogramm Remote Sens Spatial Inf Sci.* <https://doi.org/10.5194/isprs-archives-XLIII-B1-2021-85-2021>.
82. Lanzara E, Scandurra S, Musella C, Palomba D, di Luggo A, Asprone D. Documentation of structural damage and material decay phenomena in h-bim systems. 2021. *Inter Arch Photogramm Remote Sens Spatial Inf Sci.* <https://doi.org/10.5194/isprs-archives-XLVI-M-1-2021-375-2021>.
83. Conti A, Fiorini L, Massaro R, Santoni C, Tucci G. HBIM for the preservation of a historic infrastructure: the Carlo III bridge of the Carolino Aqueduct. *Appl Geomat.* 2022;14:41–51.
84. Fregonese L, Taffurelli L, Adami A, Chiarini S, Cremonesi S, Helder J, et al. Survey and modelling for the BIM of basilica of san marco in venice. 2017. *Int Arch Photogramm Remote Sens Spatial Inf Sci.* <https://doi.org/10.5194/isprs-archives-XLII-2-W3-303-2017>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.