International Journal of Architectural Heritage
Conservation, Analysis, and Restoration

ISSN: 1558-3058 (Print) 1558-3066 (Online) Journal homepage: www.tandfonline.com/journals/uarc20

Defect Augmented Heritage-BIM For Masonry Structures Leveraging Deep Learning-Based Segmentation

Xinyu Tong, Cheng Zhang, Jingyu Zhou & Luigi Di Sarno

To cite this article: Xinyu Tong, Cheng Zhang, Jingyu Zhou & Luigi Di Sarno (23 Feb 2026): Defect Augmented Heritage-BIM For Masonry Structures Leveraging Deep Learning-Based Segmentation, International Journal of Architectural Heritage, DOI: 10.1080/15583058.2026.2633538

To link to this article: https://doi.org/10.1080/15583058.2026.2633538

Published online: 23 Feb 2026.

INTERNATIONAL JOURNAL OF ARCHITECTURAL HERITAGE
https://doi.org/10.1080/15583058.2026.2633538

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

Defect Augmented Heritage-BIM For Masonry Structures Leveraging Deep Learning-Based Segmentation

Xinyu Tong[a,b], Cheng Zhang[a,b], Jingyu Zhou[a], and Luigi Di Sarno[c]

[a]Department of Civil Engineering, Design School, Xi'an Jiaotong-Liverpool University, Suzhou, China; [b]Design School Intelligent Built Environment Research Centre, Xi'an Jiaotong-Liverpool University, Suzhou, China; [c]Department of Civil and Environmental Engineering, University of Liverpool, Liverpool, UK

ABSTRACT

Defects on masonry built heritage require regular monitoring and effective management to reduce structural risks. Deep learning techniques such as CNN have been employed for automated defect identification; however, previous models underperform in complex defect scenarios due to inadequate sample size, unbalanced categories, and unclear classification standards of the training dataset. Moreover, previous mapping of defect information onto H-BIMs is superficial and not deeply integrated with specific components, which is not conducive to subsequent numerical simulation. To address these challenges, this study establishes quantifiable classification rules and introduces data augmentation and k-fold cross-validation in dataset preparation and model training to improve identification accuracy. A multi-layer masonry H-BIM is proposed, where the masonry arrangement is extracted automatically, individual masonry unit components with realistic geometries are generated parametrically, and defect information is registered with and integrated into corresponding components. A case study was conducted on an ancient city wall in Suzhou, China. The trained CNN model accurately segmented various defects, and the generated multi-layer masonry H-BIM incorporates defect visualisation, lightweighting, and semantic enrichment, laying a foundation for subsequent numerical simulation.

ARTICLE HISTORY
Received 15 November 2025
Accepted 13 February 2026

KEYWORDS
Built heritage; convolutional neural network (CNN); defect mapping; defect segmentation; Heritage Building Information Modelling (HBIM); masonry structure

1. Introduction

Masonry structural built heritage is affected by various defects over time, which compromise the integrity of the structure and pose risks to its safety. Regular inspection, therefore, is necessary to identify defects as early as possible, and subsequent management schemes need to be taken to mitigate risks in vulnerable positions.

Conventional defect inspection and documentation processes are time-consuming and labour-intensive (Keshmiry et al. 2024). In recent years, computer vision technologies have been introduced for surface defect identification, UAV photogrammetry is used for efficient large-scale data acquisition, and deep learning models such as convolutional neural networks (CNNs) are trained to automatically identify defects from the data (Mishra and Lourenço 2024). The quality of the training dataset is closely related to the CNN model identification accuracy (Carvalho Ottoni and Cordeiro Ottoni 2025). An ideal dataset should encompass an adequate sample size of multiple, relatively balanced defect categories with standardised and consistent data labelling to ensure key feature identification, avoid overfitting or neglecting certain categories, and minimise interferences. However, these requirements are challenging to meet during dataset preparation due to built heritage scarcity, diverse defect distributions, and a lack of quantifiable classification indicators, resulting in compromised applicability of previous trained models within complex scenarios.

Defect information is integrated into Heritage Building Information Modelling (H-BIM) of the historical masonry structure for a more comprehensive management in a higher dimension compared to conventional fragmented paper or electronic documents, facilitating timely health status monitoring and risk responding (Chen, Lu, and Liu 2024; Nieto-Julián et al. 2023). However, when mapping the identified defects onto the H-BIM, previous research just simply displays them on meshes without distributing them in detail into corresponding component, which is not conducive to numerical simulations. If defects are assigned as semantic information to individual masonry units, their influence on each masonry unit can be analysed more accurately. In addition, previous H-BIMs lack

CONTACT Cheng Zhang cheng.zhang@xjtlu.edu.cn Department of Civil Engineering, Design School, Xi'an Jiaotong-Liverpool University, Suzhou, Jiangsu 215123, China

automated methods to generate masonry units that reflect reality, which also affects the accuracy of subsequent structural analysis.

The present paper proposes innovative methods to deal with the challenges in defect management of masonry structural built heritage, ancient city wall as represented, covering the entire process from defect identification to H-BIM data fusion. Quantifiable classification rules are established, and data augmentation and k-fold cross-validation are introduced to optimise the data preparation and training process of the CNN model, which improves the identification accuracy in complex defect scenarios. A framework of multi-layer H-BIM generation and defect mapping is proposed, which includes generating realistic individual masonry components and accurately assigning defect attributes, enhancing digital archiving and supporting subsequent numerical simulation.

2. Related works

2.1. CNN-based masonry defect identification

CNN is a representative architecture of deep learning, primarily designed for processing computer vision-related data. Table 1 summarises the previous research on CNN-based surface defect identification of masonry or similar-structured built heritage. The trained models in previous research are considered to have achieved satisfactory prediction results. However, their ability to handle complex scenarios with multiple defects still remains challenging. Specifically, this is due to two kinds of limitations: First, some research aims to enhance the identification accuracy of single defect categories, such as crack (Huang et al. 2024), detachment (Wang et al. 2019, 2020; Yan et al. 2024), biological colonization (D'Orazio et al. 2024) and so on. Although their trained models have achieved high accuracy in one defect category, they are insufficient to get control of the comprehensive defect condition of the whole built heritage. Second, while some other research involves multiple defect categories, including erosion, efflorescence, biological colonization, missing part, crack, etc, their trained models are only applicable in scenarios lacking complexity, such as walls with uniform background colours (Hatir, Barstuğan, and İnce 2020; Karimi, Valibeig, and Rabiee 2024; Li et al. 2023; Mansuri and Patel 2022; Meklati et al. 2023; Mishra, Barman, and Ramana 2024; Wang et al. 2018) or sparse defects (Karimi, Mishra, and Lourenço 2024; Mansuri and Patel 2022; Mishra, Barman, and Ramana 2024; Yang et al. 2023). These limitations of single or limited categories and scenarios lacking complexity hinder the trained models from achieving high accuracy when encountering more complex scenarios, such as dense or cluttered defects and irregular masonry units.

2.2. Masonry defect training dataset preparation

The limitations mentioned in Section 2.1 in previous trained models mainly stem from insufficient comprehensiveness and representativeness of training datasets (Li et al. 2018; Wan, Zhang, and Liu 2021). As mentioned above, an appropriate dataset requires an adequate sample size, balanced multi-category coverage, and standardiszed, consistent labelling. However,

Table 1. Previous research on CNN-Based defect identification of masonry or similar-structure built Heritage.

References	Scenario	Target Defect(s)	Shortcoming
Huang et al. (2024)	ancient city wall	crack	**single category**
Wang et al. (2019)	ancient city wall	detachment	**single category**
Wang et al. (2020)	roof tiles	detachment	**single category**
L. Yang et al. (2024)	roof pedestals	detachment	**single category**
D'Orazio et al. (2024)	old house (supplemented with laboratory images)	biological colonization	**single category**
Wang et al. (2018)	ancient city wall	efflorescence, crack, detachment	**uniform background**
Hatir, Barstuğan, and İnce (2020)	monument	erosion, efflorescence, crack, biological colonization, etc.	**uniform background**
Li et al. (2023)	ancient city wall	erosion, efflorescence biological colonization, crack	**uniform background**
Meklati et al. (2023)	old house	efflorescence, crack, detachment	**uniform background**
Karimi, Valibeig, and Rabiee (2024)	bridge	erosion, efflorescence, crack, detachment	**uniform background**
Yang et al. (2023)	old house	erosion, efflorescence, biological colonization, crack, missing part	**sparse defects**
Karimi, Mishra, and Lourenço (2024)	wall tiles	crack, detachment, etc.	**sparse defects**
Mansuri and Patel (2022)	tomb	efflorescence, crack, detachment	**uniform background, sparse defects**
Mishra et al. (2024)	tomb	efflorescence, crack, detachment	**uniform background, sparse defects**

during the preparation of masonry heritage defect datasets, these requirements are challenging to meet due to various factors.

During data acquisition, the sample size is often constrained by the scarcity of built heritage, unlike modern buildings that are more readily available. Additionally, achieving a balanced distribution among categories is difficult. Certain prominent defect categories, like efflorescence and biological colonization, may dominate a significant portion of the facades, appearing in almost every captured image, as shown in Figure 1. Despite efforts to balance the dataset by intentionally capturing more instances of less common categories, it can be hard to mitigate their prevalence.

During data labelling, ambiguous definitions and a lack of quantifiable indicators in guidance documents lead to unclear and inconsistent classification, causing confusion in model training. Previous research often refers to specific guidance documents when categorising defects in data labelling, such as the "*Illustrated glossary on stone deterioration patterns*," compiled by the International Scientific Committee for Stone of the International Council on Monuments and Sites (ICOMOS-ISCS) (Vergès et al 2008). These documents typically define defects by providing a most representative description with an example image. In practice, however, the same defect category may exhibit various appearances due to different damage extents. For example, emerging erosion, superficial erosion, and deep erosion may display different roughness and colours, as shown in Figure 2. If they are classified under the same label in the training set, it may disrupt the model from extracting key features. Therefore, to accurately identify defects without omissions, the category needs to be further subdivided based on the damage extent.

Defect evolution from mild to severe, however, constitutes a continuous progression with no clearly demarcated boundaries between stages. The absence of quantifiable indicators in guidance documents makes it difficult to assess the damage extent. As shown in Figure 3, distinguishing between superficial and deep erosion could be ambiguous without quantifiable indicators. Although some research has extracted quantifiable indicators for certain defects (Valero et al. 2019; Yue et al. 2025), research on leveraging these indicators to enhance CNN model accuracy in complex multiple-defect scenarios remains scarce.

Figure 1. Masonry structures covered by a large area of efflorescence, plant vines and erosion.

2.3. Masonry H-BIM generation and defect mapping

The identified defect information is often integrated with H-BIM to assist in safety monitoring and risk assessments. Previous models of masonry structural built heritage still have shortcomings in reflecting realistic geometry and defect position, which undermines the reliability of subsequent numerical simulation. On the one hand, although previous simulation models treat walls as assemblies of discrete bricks to better reflect the masonry unit separation from mortar (Loverdos et al. 2021), they adopt idealised (Abu-Haifa and Lee 2023; Alforno et al. 2024; Du et al. 2022; Loverdos et al. 2021; Pereira et al. 2023) or randomly generated (Ehab and Massart, 2025; Milani and Lourenço 2010; Shaqfa and Beyer 2022) brick

Figure 2. Erosion defect with different damage extent showing different appearance.

Figure 3. Unclear boundaries of No-risk - Superficial erosion (yellow box) and Superficial - Deep erosion (red box).

arrangements that do not capture morphological alterations caused by defects, which brings discrepancies between simulation results and real structural behaviour (Loverdos et al. 2021; Pantoja-Rosero et al. 2023; Sun et al. 2019; Tiberti and Milani 2020). On the other hand, previous research has explored the automated defect mapping process, typically based on the conversion among coordinates such as geographic coordinate, planar photo coordinate, camera coordinate and 3D model coordinate, to find the corresponding points between images and models (Cavalagli et al. 2019; Chen, Lu, and Lou 2023; Cuadros-Rojas et al. 2024; Liu et al. 2021; Luo and Wang 2024; Nieto-Julián et al. 2025; Pantoja-Rosero, Achanta, and Beyer 2023; Tan et al. 2022; Yan and Wang 2025; Yang et al. 2024; Ye et al. 2026). However, they mainly focus on modern buildings, such as concrete walls, and merely overlay defects on the model surface without deep integration with components. Such mapping process is imprecise for masonry built heritage where the facades are subdivided by more detailed subcomponents of masonry units, which is not conducive to subsequent analysis.

3. Problem statement and objectives

3.1. Problem statement

According to the Related Works on masonry structural built heritage, research gaps exist in defect identification and H-BIM establishment. First, regarding defect identification, the preparation of training dataset is affected by inadequate sample size, unbalanced categories and unclear classification rules. This leads to trained CNN models being unsuitable for complex scenarios. Second, regarding H-BIM establishment, in the defect mapping process, defects are only represented superficially without deeper fusion with the masonry unit components. The masonry unit components also fail to reflect realistic geometry, which hinders subsequent numerical simulation.

3.2. Objectives

The present paper aims to address the challenges of inadequate sample size, unbalanced categories, and unclear classification rules in defect dataset preparation, to improve the CNN model training quality and enhance identification accuracy in complex scenarios. Meanwhile, it also aims to deeply integrate defect information into more precise corresponding masonry units, thereby enhancing model reliability for subsequent numerical simulation and making the masonry H-BIM not only defect-visualised and lightweight but also semantically enriched.

4. Methodology

The proposed method provides an optimised workflow including defect and masonry unit identification, and masonry H-BIM generation with defect augmentation, as shown in Figure 4. UAV photogrammetry is employed for capturing defects and masonry units. Multiple CNN models are developed to identify defects and masonry units separately. To address the aforementioned challenges of the ancient masonry defect dataset, the present paper proposes quantifiable classification rules, and introduces data augmentation and k-fold cross-validation to establish clearer classification rules, enrich the sample size and mitigate the impact of category imbalance. Meanwhile, individual masonry units are identified after a similar data labelling and training process, and their contours are extracted. The identified defects and masonry units are then used to construct the masonry H-BIM, which consists of three independent but interrelated layers. The three layers focuses on visualisation display, lightweighting, semantic enrichment and numerical simulation, respectively, which can be chosen as needed in practical applications. The main process of establishing these layers covers mesh model reconstruction, projection of the mesh texture

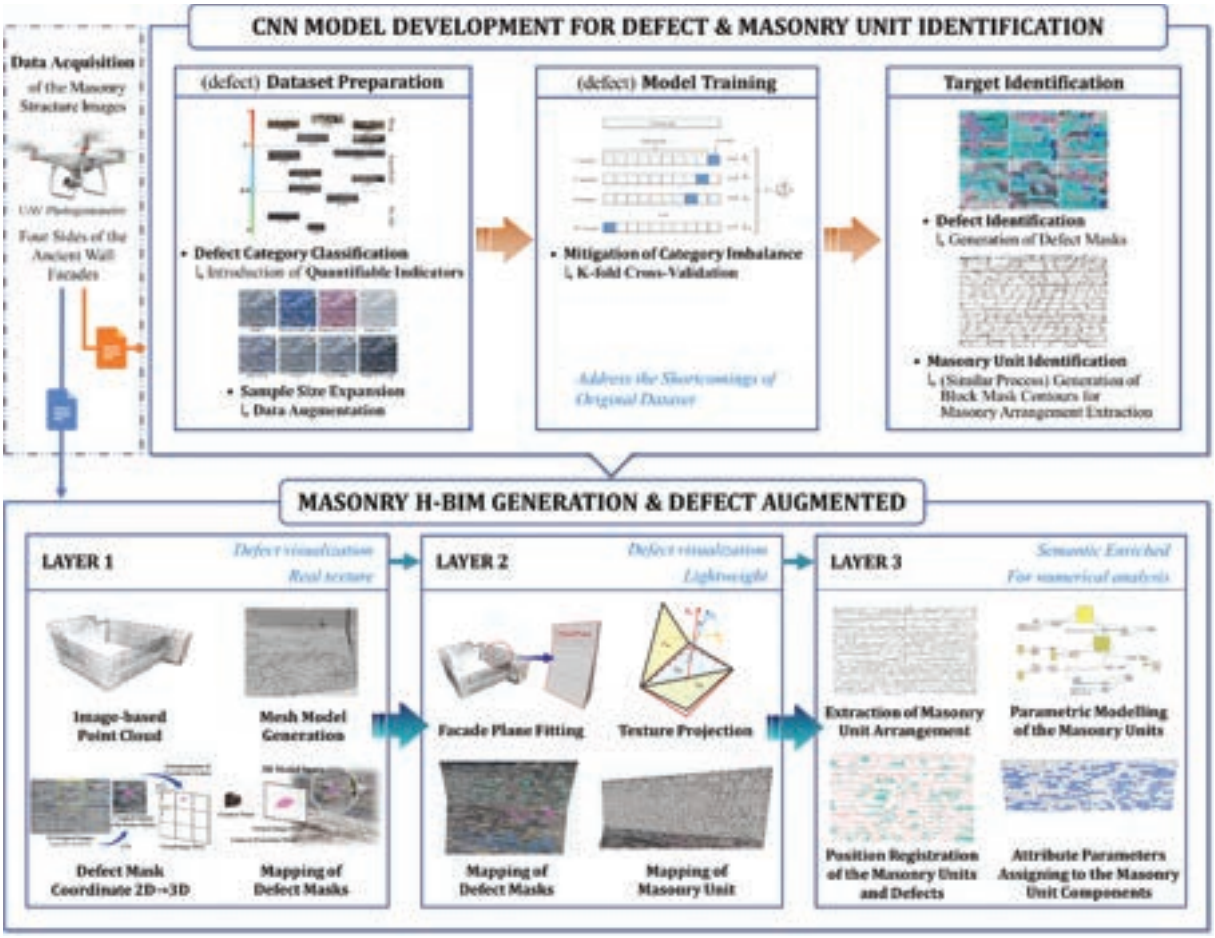

Figure 4. The Framework of the proposed method.

on the fitted model surface, parametric masonry unit component generation and the corresponding defect attribute parameters assigning.

4.1. CNN model development for defect and masonry unit identification

4.1.1. Quantifiable rules for defect classification

As mentioned above, there is no clear boundaries between different damage extents within a defect category. In order to improve labelling consistency and enable accurate identification across the whole defect stages, classification rules with quantifiable indicators are proposed.

Taking erosion as an example, erosion is divided into three subcategories by damage extent, which are No risk, Superficial erosion and Deep erosion. A two-screening classification process is set up, as shown in Table 2.

Firstly, in addition to referencing the original definitions and visual appearance features in existing guidance documents, the proportion of area is established as the boundary, used for the preliminary screening of the three erosion damage extents. Through preliminary screening, the majority of erosion bricks can be classified. Secondly, for those bricks that are still difficult to classify, the gray-level co-occurrence matrix (GLCM) and image

Table 2. Two-screening classification rules for erosion bricks with quantifiable indicators.

Damage Extent of Erosion (Subcategories)	Preliminary Screening		Secondary Screening
	Appearance Features	Proportion of Defects in Brick Area	GLCM & Gradient Quantifiable Indicators
No risk (no labelling)	Flat, smooth original surface	Defect<30%	Average normalised value<0.35
Superficial	Slightly rounding contour and roughening surface	30%<Defect<70%	0.35<Average normalised value<0.7
Deep	Obvious irregular contour or honeycomb-like uneven surface.	Defect>70%	Average normalised value>0.7

gradient are introduced as calculation indicators for extracting image features to provide a more detailed classification basis, and the secondary screening is conducted. Specifically:

Gray-Level Co-occurrence Matrix: GLCM is a second-order statistical texture analysis method used to examine the spatial relationship among pixels. It defines $\boldsymbol{P}(\boldsymbol{i},\boldsymbol{j})$ as the frequency of a combination of pixels present a specific gray level $\boldsymbol{i},\boldsymbol{j}$ in the directions of $\theta = 0°$, 45°, 90° and 135° and the optimal generation step $\boldsymbol{d}$ (Bhagat, Choudhary, and Singh 2019; Haralick, Shanmugam, and Dinstein 1973). Several statistical features based on GLCM are utilised to describe and differentiate texture characteristics. In the present paper, the following features are employed to quantify the damage extent:

Contrast: Contrast elucidates the disparity between light and dark regions in an image. High contrast implies a distinct texture structure, while low contrast signifies relatively uniform texture.

$$\text{Contrast} = \sum_{i,j} (i-j)^2 P(i,j) \tag{1}$$

Entropy: Entropy denotes the randomness of image texture, depicting whether the texture distribution exhibits uniformity or complexity. Greater complexity in texture corresponds to higher entropy values.

$$\text{Entropy} = \sum_{\boldsymbol{i},\boldsymbol{j}} \boldsymbol{P}(\boldsymbol{i},\boldsymbol{j}) \log(\boldsymbol{P}(\boldsymbol{i},\boldsymbol{j}) + \varepsilon) \tag{2}$$

where ε is a small constant to prevent logarithmic zero.

Homogeneity: Homogeneity reflects the proximity of pixel gray values, indicating the smoothness of the texture. The smoother the image texture, the higher the homogeneity.

$$\text{Homogeneity} = \sum_{\boldsymbol{i},\boldsymbol{j}} \frac{\boldsymbol{P}(\boldsymbol{i},\boldsymbol{j})}{1+|\boldsymbol{i}-\boldsymbol{j}|} \tag{3}$$

Angular Second Moment (ASM): ASM represents the repetitiveness and smoothness of the texture. A more uniform grayscale distribution in the image leads to a higher ASM value.

$$\text{ASM} = \sum_{\boldsymbol{i},\boldsymbol{j}} \boldsymbol{P}(\boldsymbol{i},\boldsymbol{j})^2 \tag{4}$$

Image gradient refers to changes in image intensity or colour, serving as an indicator for describing texture edges. The Sobel operator (Sobel 1970) is employed to approximate the image gradient. This operator consists of two 3×3 convolution kernels used to compute the image gradients $\boldsymbol{G_x}$ and $\boldsymbol{G_y}$ in the ×and y directions, respectively. Consequently, the Sobel gradient $\boldsymbol{G}(\boldsymbol{i},\boldsymbol{j})$ at position $(\boldsymbol{i},\boldsymbol{j})$ in the image is calculated as

$$G(i,j) = \sqrt{G_x^2 + G_y^2}$$

Therefore, the average gradient of the image can be expressed as,

$$\text{Average Gradient} = \frac{\sum_{i,j} G(i,j)}{\sum_{i,j} 1} \tag{5}$$

Due to the disparate scales of the aforementioned indicators, they are subjected to normalisation. Contrast, Entropy, and Average Gradient are positively correlated with the damage extent of erosion and are normalised using Equation (6). Homogeneity and ASM are inversely related to the damage extent and are normalised using Equation (7) to ensure consistent correlation among the indicators. The normalisation computation is as follows.

$$X_i^{'} = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \tag{6}$$

$$X_i^{''} = 1 - \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \tag{7}$$

The screening rules in Table 2 were developed adapting to the current dataset condition. For secondary screening, over 200 bricks were selected and cropped into individual images, with GLCM and image gradient indicators computed, to ensure the reliability of boundary value settings. The rules also refer to Valero et al. (2019) and Yue et al. (2025), and were formulated following discussions with experts and site managers.

To save labelling time, the indicators for secondary screening are only applied to particularly difficult-to-distinguish defects, while bricks with clear distinctions can be simply classified through preliminary screening. After the two rounds of screening based on the proposed rules, consistency in labelling the extent of erosion by different personnel can be assured.

4.1.2. Sample size expansion by data augmentation

Since the on-site collected data is limited, data augmentation is employed to obtain adequate training samples after the collected images are labelled, including Saturation adjustment, Random colour shift, Random noise, Colour jitter, Sharpen, Gaussian blur, Brightness change, etc. The expanded dataset has better diversity and flexibility, enhancing the identification accuracy.

4.1.3. Mitigation of category imbalance by K-fold cross-validation

A CNN-based YOLO model (Jocher and Qiu 2024) is trained using the prepared dataset because of its advantage of being adaptable to targets of various sizes and shapes

with good detection speed (Casas et al. 2024). K-fold cross-validation is used to mitigate the effects of data category imbalance. The original dataset is divided into K subsets, with K-1 subsets used iteratively for training and the remaining subset for validation. Among K rounds of independent training, the averaged and the best result are taken as outputs (Arlot and Celisse 2010). It leads to more representative training and more robust performance as it ensures that each separate training contains samples from those rare categories, reducing randomness in single trainings (Dutschmann et al. 2023).

4.1.4. Performance metrics for trained model Evaluation

Precision (P), Recall (R), mean Average Precision at IoU threshold 0.5 (mAP50), and mean Average Precision across multiple IoU thresholds from 0.5 to 0.95 (mAP50-95) (Jocher and Qiu 2024) are used as performance metrics to evaluate the segmentation tasks with different training sets. Specifically:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{8}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{9}$$

where ***TP*** (True Positives) is the number of correctly predicted bounding boxes; ***FP*** (False Positives) is the number of incorrectly predicted bounding boxes; ***FN*** (False Negatives) is the number of ground-truth objects not detected by the model.

$$\text{mean Average Precision (mAP)} = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} \mathrm{AP_i} \tag{10}$$

where ***N*** is the number of object classes; $\boldsymbol{AP_i}$ is the average precision for class ***i***, which is the area under the Precision-Recall curve. The mAP50 and mAP50-95 are mAP values computed at specific IoU thresholds; IoU is the ratio of the area of overlap to the area of union between the predicted and ground-truth bounding boxes.

The trained YOLO model is used to identify defects in multiple categories. Meanwhile, another model for masonry unit segmentation is obtained through a similar process. Subsequently, the identified defects and the masonry units are mapped onto the masonry H-BIM in 3D space.

4.2. Masonry H-BIM generation and defect augmented

An H-BIM with multiple layers is developed. Layer-1 focuses on defect visualisation, which incorporates real textures to facilitate the intuitive display of defects. Layer-2 builds upon Layer-1 but reduces the mesh for a more lightweight model. Layer-3 generates parametric components of masonry unit embedded with defect information for numerical simulation.

4.2.1. Layer-1

The Layer-1 is built on the mesh model obtained through image-based 3D reconstruction. Defects identified by the trained CNN model are mapped onto the mesh model, enabling the visualisation of defect information in 3D space. Defect positions are transformed from 2D image coordinate system to 3D model space coordinate system. The pipeline typically uses Structure from Motion (SfM) and Multiple-View Stereopsis (MVS) (Seitz et al. 2006) techniques to analyse the camera poses of a set of images based on the position information in single images and the overlaps between adjacent images, and align them in 3D space. Then, based on projection principles or ray casting techniques, the corresponding positions of points in the 2D image are found on the mesh model. In this pipeline, a new issue is identified which requires the following explanation. The alignment of image poses uses uncropped high-resolution original images. However, for CNN training, images are often cropped to smaller sizes (e.g., 640 × 640 pixels) to avoid excessive computational resources and time consumption. The position information of original images is lost in the cropping process, causing that the identified defect masks on the cropped small images cannot be directly project to the 3D space. Therefore, an additional coordinate transformation process is proposed as follows: each original image is cropped into several small images of m rows and n columns for defect prediction, and the defect coordinates on each small image need to be concatenated to the coordinates of the original image. The conversion is straightforward as shown in Figure 5, involving simple addition based on the row and column. The coordinates of defect masks from cropped small images are transformed into the original image coordinate system. Subsequently, the transformed defect coordinates are projected onto the mesh model based on the pose of the original image.

4.2.2. Layer-2

The Layer-2 H-BIM preserves the original colours of the real world while reducing the meshes, resulting in a more lightweight representation. The Layer-2 model is constructed from point cloud via image alignment. The image point cloud of facade surface is clustered and fitted into planes based on the Polyfit algorithm (Nan

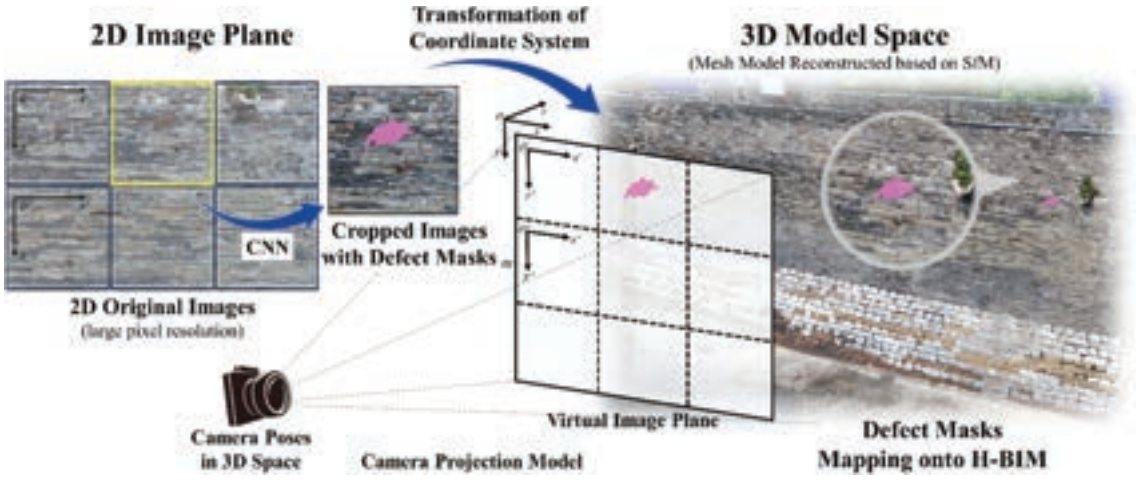

Figure 5. Layer-1: camera projection model with proposed coordinate transformation method.

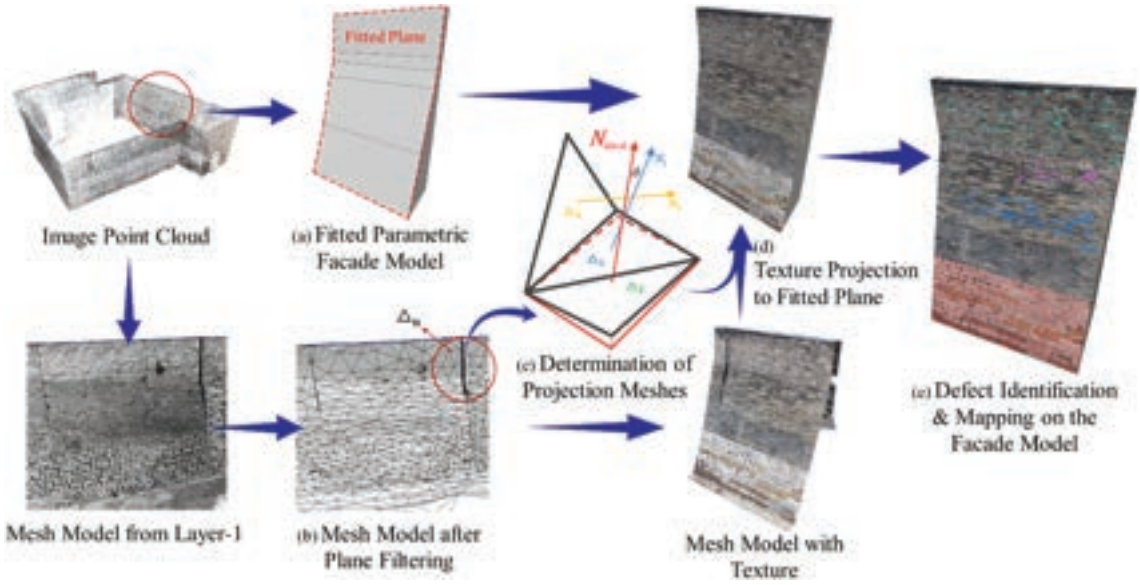

Figure 6. Layer-2: lightweight parametric model generation and texture projection.

and Wonka 2017), as shown in Figure 6(a), which are then fine-tuned to align with the boundaries of the Layer-1 mesh model. The texture mapping of the fitted model is shown in Figure 6, where the textures on the meshes of the Layer-1 model are projected onto the Layer-2 fitted planes. Specifically, the meshes of the Layer-1 model are initially filtered to smooth the facade surface for mitigating potential interference from uneven meshes that could affect the subsequent determination process, as shown in Figure 6(b). Each mesh is then evaluated to determine if it belongs to the fitted plane, as shown in Figure 6(c), based on the differences in normal vectors and distances between the two, with reference to Z. Du et al. (2014) and Tian, Ruan, and Wang (2002). The normal vectors of each mesh in the model, as well as the normal vectors of the fitted plane, are extracted, and then the angle between them is calculated. If the angle between a mesh and the fitted plane is less than a specified value, and the distance between them does not exceed a specific range, it is determined that the mesh belongs to the wall facade. Conversely, if the angle is too large or the distance is too far, the mesh may represent a corner or belong to another structure. The pixels on meshes are subsequently projected onto the corresponding fitted planes using the render-to-textures method to generate textures, as shown in Figure 6(d). Since both layers of the models originate from the same image point cloud, they share the same envelope, enabling precise texture projection. Once the texture of the fitted plane is generated, defects are

identified and displayed on the model, as shown in Figure 6(e) with different colours. The Layer-2 model will also be employed to construct the Layer-3 model.

4.2.3. Layer-3

The Layer-3 H-BIM contains more precisely located semantic information for numerical simulation. Unlike the preceding layers, the Layer-3 facade model is composed of individually arranged parametric masonry unit components, which are "embedded" into the base model and assigned corresponding defect information, as opposed to merely overlaying defect masks on the surface. Therefore, with the preceding two layers of models generated, the facade surface needs further division according to the masonry arrangement.

The masonry arrangement hinges on the contours derived from identified masonry unit masks. The masks identified from the Layer-2 texture of the fitted planes are used instead of the masks mapped on the Layer-1 mesh surface. This choice is based on two considerations. Firstly, the heightened complexity of spatial relationships among contour points in 3D space compared to a 2D plane. Images for generating the mesh model are overlapping, meaning that a single masonry unit might appear in two original images, leading to duplicate masks in overlapping regions when mapping images separately. Consequently, the calculation of the homography transformation is required to merge the masks on a coplanar surface for additional projection (Pantoja-Rosero et al. 2022), which is inconvenient. Secondly, the uneven mesh surface in Layer-1 does not reflect the realistic construction principles for bricklaying. In contrast, assembling masonry on planes adheres more closely to practical logic.

Based on the masonry arrangement derived from the Layer-2 masks, the division process of facade surface is as follows. As the masks are known to be located within the fitted planes, each plane can be easily divided into individual masonry unit elements according to the mask contours. Subsequently, each element is assigned a thickness value to parametrically extrude the contour into a solid, as shown in Figure 7(a). The thickness values are determined based on the actual block thickness to represent the structure in reality.

After generating the parametric components of individual masonry units on the facade, corresponding defect attributes are assigned to each block. The determination of the relative positions between defects and masonry units refers to Franklin (2006), which assesses whether defect points lie within the contours of masonry units. If the point falls within the contour, the corresponding defect is credited as being present in that masonry unit, and then the attributes and properties related to that particular defect category are assigned to that masonry unit component as parameters. By sequentially comparing the positions between all defects and the masonry units, the assignment of each defect to a specific masonry unit is established, as shown in Figure 7(b), and the relevant attributes and properties of each defect category are assigned as parameters to those specific masonry unit components, as shown in Figure 7(c), where the masonry units with defects are highlighted in blue. At this point, the Layer-3 H-BIM is finally generated. The

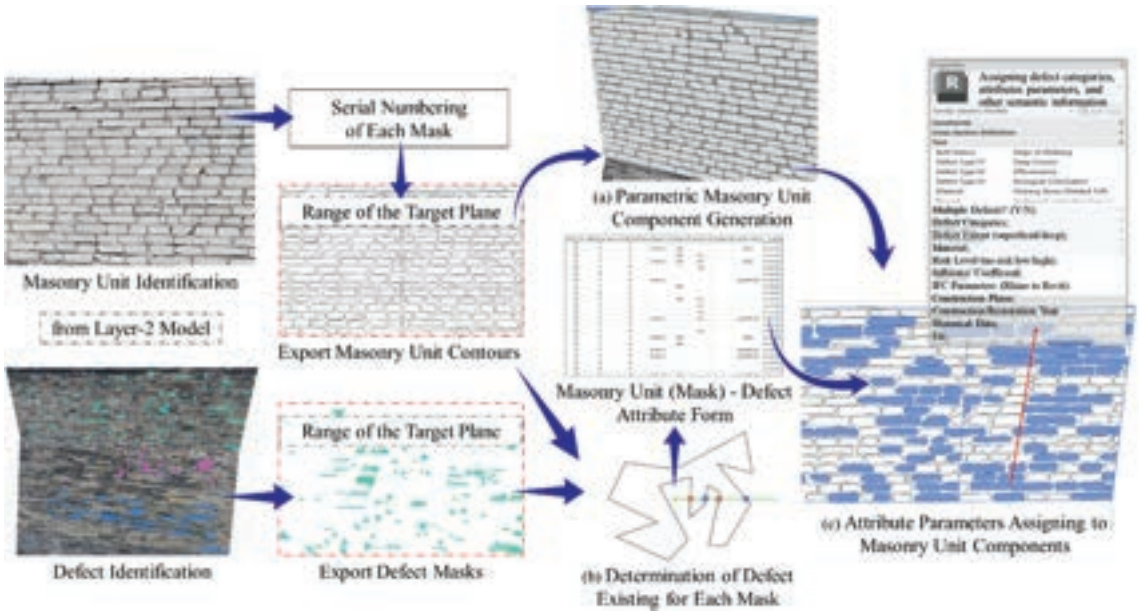

Figure 7. Layer-3: parametric masonry units generation and defect information embedment.

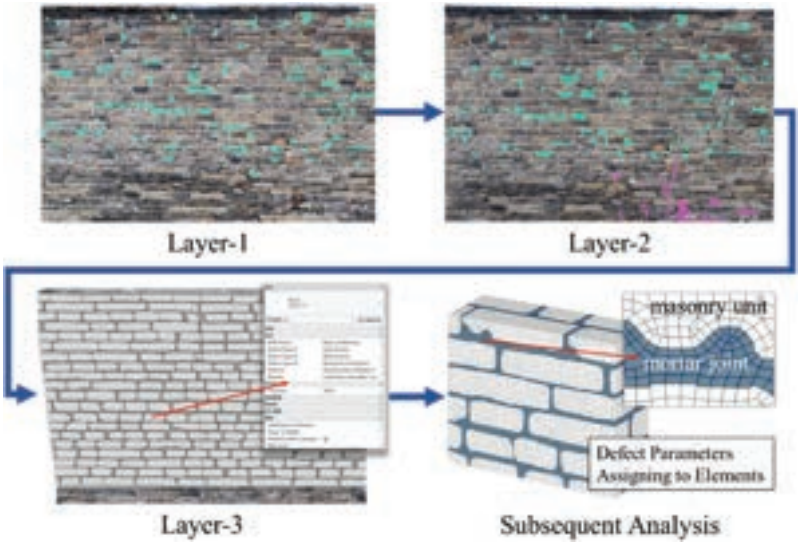

Figure 8. The defect augmented three-Layer masonry H-BIM.

Layer-3 model reflects each individual masonry unit parametrically with realistic geometry while embedding defect information, thereby achieving the fusion of defect information with the masonry model.

The three-layer masonry H-BIM is overall generated as shown in Figure 8. In the management of the historical masonry structures, Layer-1 or Layer-2 provide an intuitive display of the general positions of defects on masonry structures. After determining a specific area of interest, more specific information can be viewed in the corresponding position of the Layer-3 model, allowing for the retrieval of defect information on each individual masonry unit for further detailed analysis. For example, a conceptual illustration of numerical simulation for the subsequent analysis is presented in Figure 8. Based on the Layer-3 model where defect attributes are assigned to masonry units with realistic geometries, it is possible to develop finite element models that reflect the actual structural conditions.

5. Case study

A case study was carried out at the Pan Gate ancient wall, Suzhou, China, as shown in Figure 9. The current structure was built in 1351 and underwent reinforcement, restoration, and reconstruction between 1976 and 1981. The construction materials consist of bricks, stones, earth, mortar, and a small amount of wood with diversity in type, shape, and size. Multiple categories of defects are present and causing damage, including erosion, efflorescence, biological colonization, and missing part, as shown in Figure 10. Among these, erosion and efflorescence are distributed over large areas with irregular shapes; plant vines extend irregularly and are tangled over other defects, complicating identification. The damage extent of each brick also varies, which is difficult to classify when labelling. The complexity of this ancient wall presents significant challenges in defect management, thus highlighting the feasibility of the proposed methods for defect identification and masonry H-BIM data fusion.

5.1. Data acquisition and training set preparation

The ancient walls were captured through UAV photogrammetry. The shooting method is called nap-of-the-object photogrammetry, which involves capturing the target surface at a close distance, allowing to capture high-resolution images for the creation of detailed 3D models. The UAV automatically collected images from the four sides of the walls through flight trajectory planning. Each image was taken perpendicularly at a distance of 3–5 meters from the facade to ensure that the masonry arrangement and the defects

Figure 9. Ancient city walls of Pan Gate Historical site.

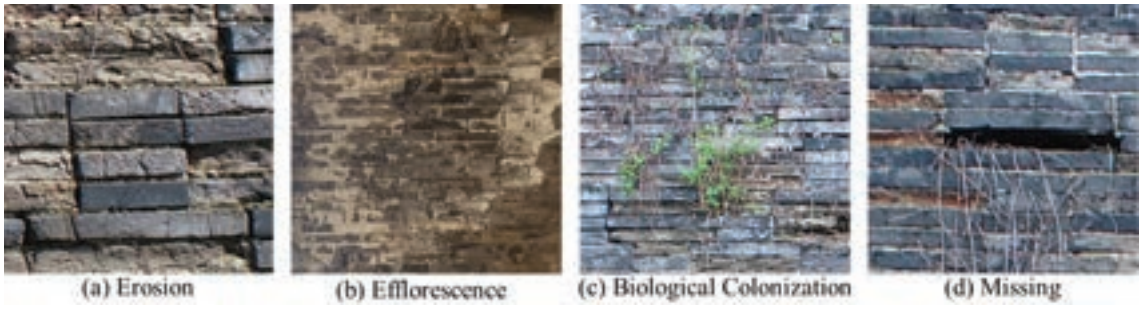

Figure 10. Defects on Pan Gate.

were clearly visible. Because the images were to be used for 3D reconstruction of the Layer-1 model, their quality had to be ensured. To avoid blurry images, data acquisition was carried out in clear weather without excessive sunlight to prevent overexposure. At the same time, there was no strong wind, which would cause the UAV to shake and blur the images. The UAV's built-in real-time kinematic (RTK) technology enables real-time dynamic positioning, so each image contains the position information at the shooting point.

A total of 191 orthoimages with a resolution of 5472×3648 pixels were obtained. The obtained images were divided into two parts: the majority was used to establish the YOLO training and validation dataset. Another small portion, selected form part of each side of the facade, was reserved for testing defect and masonry unit identification, as well as the reconstruction of the facade model, thus, this portion of images was not used for training. For training and validation dataset preparation, the part of original images were uniformly cropped to a size of 640×640 pixels and duplicate images taken at the same location were also removed to prevent interference from inconsistent labelling. After filtering, a total of 500 photos were produced.

Data labelling of defects and masonry units were performed separately on the pre-processed images. The labelling of defects covered more steps due to their diversity and complexity, while the labelling of masonry units was comparatively less demanding and required only similar simplified steps.

Defect categories, including erosion, efflorescence, biological colonization, missing part and crack were labelled by personnel under unified classification rules. In addition to referencing definitions from the guidance documents, the proposed classification rules incorporate more distinct appearance features and quantifiable indicators, ensuring consistent labelling among different personnel. The classification rules for each category are as follows:

(1) No-risk erosion: The area of erosion within the brick is less than 30% of the entire brick. The average normalized value of GLCM and gradient is less than 0.35. Since nearly all bricks exhibit even the slightest traces of erosion because of the long age of the wall, to reduce the labelling work, bricks that pose no risk to the structure are not labelled.
(2) Superficial Erosion: Slight rounding contour and roughening surface. The area of erosion within the brick is greater than 30% but less than 70% of the entire brick. The average normalized value of GLCM and gradient is between 0.35 and 0.7. The entire brick should be selected as the mask when labelling.
(3) Deep Erosion: Irregular brick contour or pronounced honeycomb-like unevenness. The area of erosion within the brick is greater than 70% of the entire brick. The average normalised value exceeds 0.7. The entire brick should be selected as the mask when labelling.
(4) Efflorescence: Powdery or whisker-like deposits on the surface, characterised by distinct whitish spots. Only the white portions should be selected in the mask.
(5) Biological Colonization: Vegetal living being, having root, stem, and leaves though consisting sometimes only of a single leafy expansion.
(6) Missing Part: Empty space, located in the place of some formerly existing stone part. Characterised by black voids.
(7) Crack: Separation and splitting of masonry units. Characterised by thick black vertical fissure.

Specially, for bricks which have ambiguous damage extent of erosion, the indicators of GLCM and image gradient were computed for individually cropped images of bricks, and some calculation results are presented in Figure 11. The average of the normalised value of each indicator was taken as the final result. Based on the screening rules in Table 2, a value less than 0.35 is

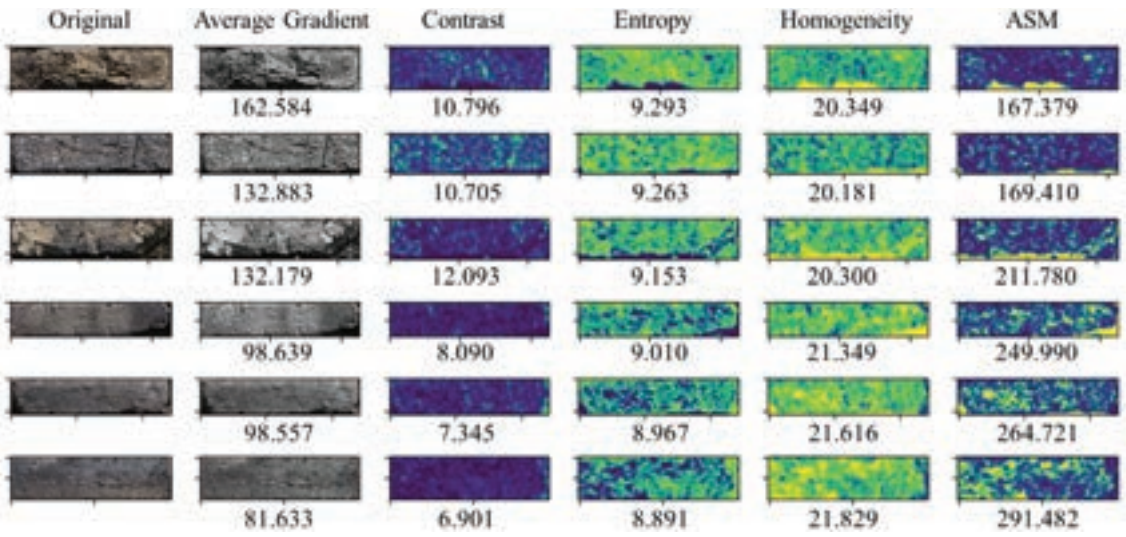

Figure 11. Calculation results of GLCM features and image gradient for individual bricks.

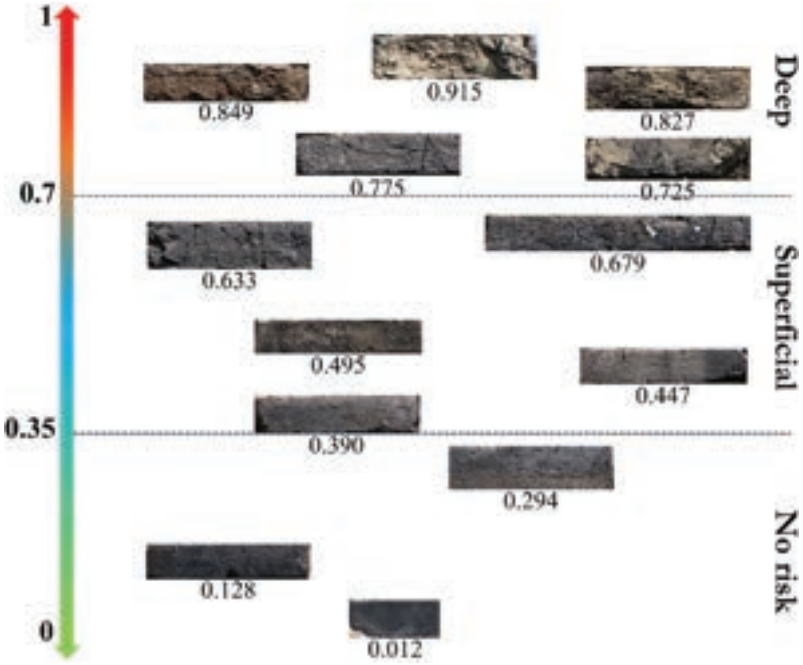

Figure 12. Bricks with different damage extent classified by the normalised value.

classified as "No risk," a value between 0.35 and 0.7 is classified as "Superficial erosion," and a value above 0.7 is classified as "Deep erosion". Some of the classification results are illustrated in Figure 12, showing that the established rules clearly classified bricks with different damage extents.

5.2. CNN model training and defect identification

After images were captured and labelled based on the quantifiable classification rules, the generated original training set was processed using data augmentation methods including two types of saturation adjustment, random colour shift, random noise, colour jitter, sharpening, Gaussian blur, and two types of brightness changes. Figure 13 shows an example of the augmented images. Each original image, along with its augmented version, was a combination of 10 images, expanding the dataset from 500 to 5000 images.

The augmented dataset of 5000 images was used to train YOLO11-segment models. To identify the optimal training strategy and achieve optimised model weights, model performance was compared across various training strategies, including different sample sizes, with or

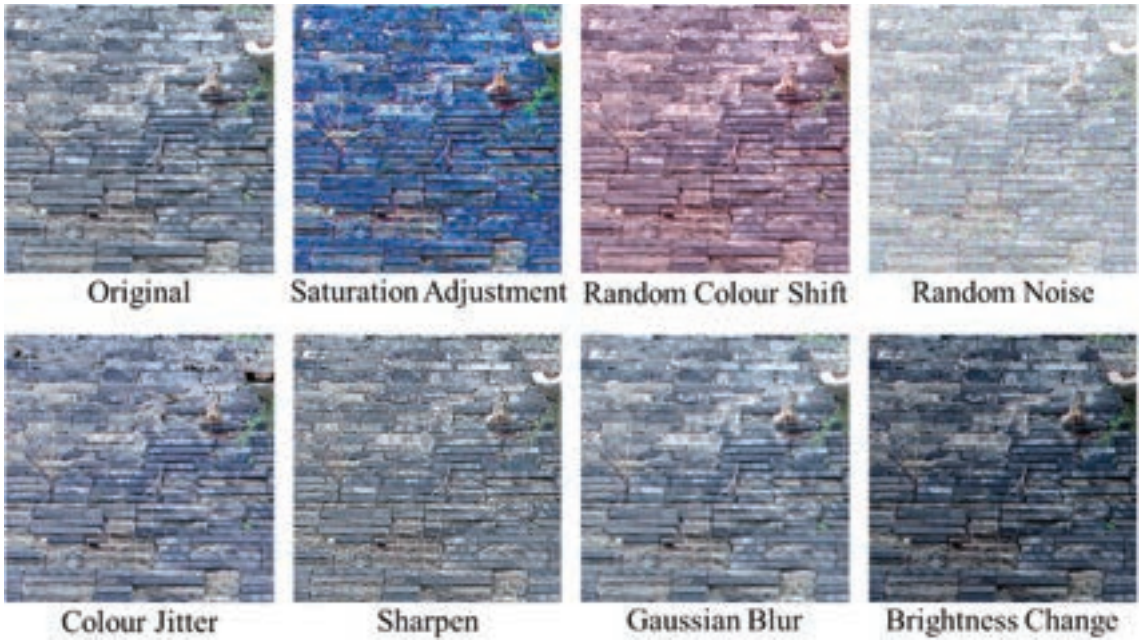

Figure 13. Example of the augmented images.

without k-fold cross-validation, and with different pre-trained models (n, s, m, l, x). Precision (P), Recall (R), mAP50 and mAP50-95 were used as performance metrics to evaluate the tasks.

Firstly, training sets with different sample sizes and corresponding validation sets at a fixed ratio were constructed to explore the sample size that achieves the accuracy requirement. Specifically, training sets with 400, 1000, 2000, 3000, and 4000 images were established, and corresponding validation sets were built at a 4:1 ratio, with 100, 250, 500, 750, and 1000 images, respectively, to compare performance. The detailed division process is as follows. The 5000-image augmented dataset was randomly divided into a 4:1 ratio for constructing the training set and validation set respectively. This means that 4000 images were only used to construct the training set, and 1000 images were only used to construct the validation set. It is important to note that images in the same combination (an original image and its 9 augmented versions) would not appear in both the training set and the validation set to avoid inflated results. From the 4000 images for constructing training set, 400, 1000, 2000, 3000, and 4000 images were randomly selected, respectively Validation sets of corresponding proportion were selected from the other 1000 images. The training sets off different sample sizes were trained and validated using YOLO11n-seg pre-trained weights and without k-fold cross-validation. The overall results of all categories are shown in Table 3, whereby the training set with 4000 augmented images achieved the highest accuracy.

Subsequently, with the sample size of 4000 images, the training performance under k-fold cross-validation and different pre-trained weights was further investigated. The 5000-image dataset were split into training and validation sets in the same 4:1 ratio, with each split randomly dividing 4000 images for training and 1000 images for validation. Images from the same combination would not be separated into two sets. With k-fold cross-validation, different pre-trained models from YOLO11, including

Table 3. Results of model validation metrics for different training sets in segmentation tasks.

	Dataset Size & Pre-trained Model	Precision (P)	Recall (R)	mAP50	mAP50-95
without KFCV	400pic YOLO11n-seg	0.477	0.278	0.236	0.143
	1000pic YOLO11n-seg	0.428	0.344	0.308	0.187
	2000pic YOLO11n-seg	0.564	0.442	0.442	0.270
	4000pic YOLO11n-seg	0.743	0.655	0.688	0.416
with KFCV	4000pic YOLO11n-seg	0.761	0.674	0.709	0.426
	4000pic YOLO11s-seg	0.929	0.874	0.906	0.589
	4000pic YOLO11m-seg	0.953	0.927	0.944	0.676
	4000pic YOLO11l-seg	0.955	0.932	0.953	0.698
	4000pic YOLO11x-seg	0.966	0.948	0.964	0.740

Note. KFCV refers to k-fold cross-validation; mAP50 refers to mean Average Precision at an IoU threshold 0.5; mAP50-95 refers to mean Average Precision across multiple IoU thresholds from 0.5 to 0.95. For the calculation expressions, please refer to Section 4.1.4.

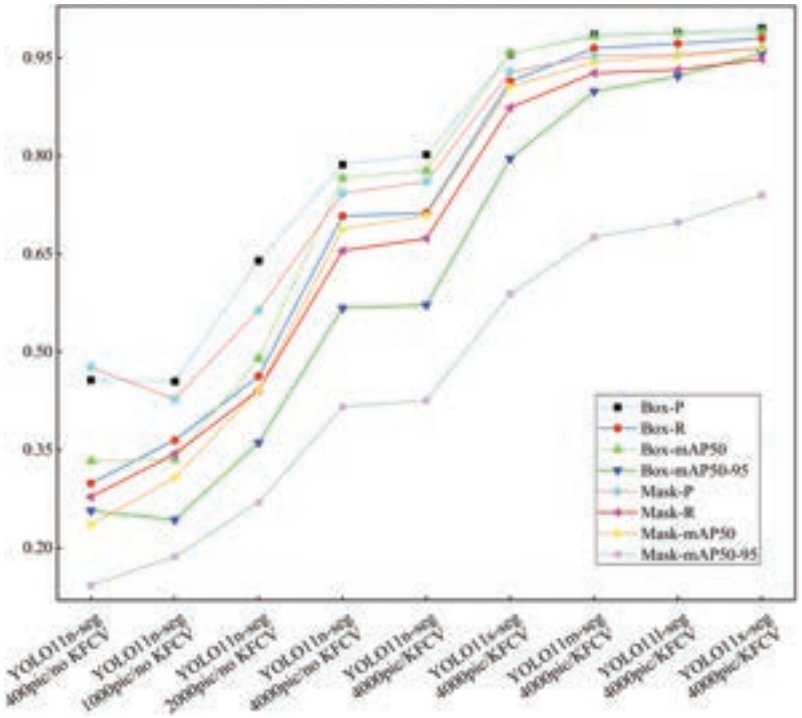

Figure 14. Comparison of the metrics among different training strategies in Detection and segmentation tasks.

nano (n), small(s), medium(m), large(l) and extra-large (x), were investigated under the 4000-image training set to obtain the optimal model. The overall results for all categories in segmentation tasks are shown in Table 3, and the comparison of the metrics among different training strategies in detection and segmentation tasks are shown in Figure 14, which demonstrates that the augmented dataset effectively improves training accuracy, addressing the issue of inadequate samples in the original dataset. Larger-scale pre-trained models also improve accuracy, but come with increased computational costs. Comparing the results from k-fold cross-validation to a single training revealed an average improvement of 1.98%.

The best weights obtained from k-fold training on the 4000-image dataset using YOLO11x-seg were selected for defect identification. The other portion of images reserved for model reconstruction was used to test the trained model. The results, as shown in Figure 15, depicting deep erosion in light blue, superficial erosion in white, efflorescence in light green, biological colonization in pink, missing part and crack in dark blue. The results indicate that the trained model generated masks covering most defect profiles, distinguished different damage extents of erosion, accurately identified the shape of plant vines, and captured efflorescence of varying shapes and sizes. A very small number of plant vines and lighter-coloured efflorescence were not fully covered, but this basically did not affect the subsequent position assignment with the masonry unit.

The labelling work of masonry units was also on those collected images. To improve the prediction accuracy, bricks and stones were classified into different categories, and bricks of different colours were also divided into multiple categories for labelling. After training, the YOLO model could accurately identify the masonry arrangement and generate masks. Figure 16 shows that the masks fit the original contours of the masonry units and reflect the masonry arrangement.

5.3. Masonry H-BIM generation

Part of the captured images were used for generating the H-BIM of the Pan Gate ancient wall facades. The distortions in the images were corrected according to the camera intrinsic parameters, and images were aligned based on the calculation of the camera poses to generate the image point cloud and create the Layer-1 mesh model, as shown in Figure 17(a). The Polyfit algorithm was employed to cluster facades from the image point clouds and fit them into planes. The fitted planes were then fine-tuned to align with the boundaries of the Layer-1 mesh model, thus generating the Layer-2 facade model. Textures were subsequently projected onto the

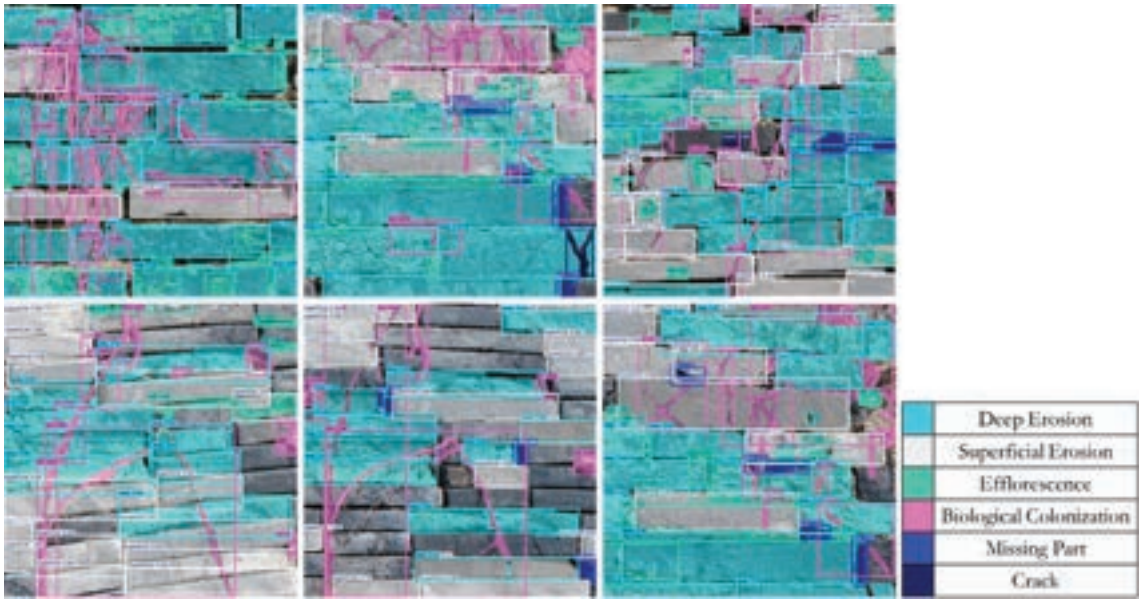

Figure 15. Examples of defect identification results.

facade model, which were then used for defect and masonry unit identification; the resulting masks were then mapped onto the surface, as shown in Figure 17(b). Above processes were implemented through programming on Python. To generate the Layer-3 model, the masonry unit masks located in the same plane range were extracted, and their contours were parametrically converted into individual components, as shown in Figure 17(c). The generation of masonry unit components was performed in Rhino, with the built-in Grasshopper plugin used for programming. This allowed the identified masonry unit contours to be automatically numbered and converted into individual components. Subsequently, the relative positions between defects and masonry units were determined based on the principle outlined in the Methodology for assessing the defect condition of each masonry unit. The semantic information for each masonry unit component was stored in a form file, encompassing various attributes including defect information, such as:

- Defect Exists? Value: Y/N;
- Multiple Defects? Value: Y/N;
- Defect Category 1: Name; Defect Category 2: Name; Defect Category 3: Name;
- Defect Extent? Value: Superficial/Deep (for erosion);
- Risk Level. Value: No risk/Low/High;
- Influence Coefficient. Value: Number. (Currently, this parameter is blank, and in future work, it is planned to be used for finite element numerical calculations.);
- The above parameters are related to defects. In addition, some other parameters beyond defects were also established for supplementing heritage knowledge, including:
- Material. Value: Text;
- Construction Phase. Value: Text;
- etc.

Figure 16. Masks of masonry units.

The algorithm could read the form file, filter specific attributes, and display the corresponding masonry unit with a particular colour in Rhino. Figure 17(d) shows bricks with efflorescence being displayed in blue for easy identification of their positions in H-BIM. The model was also attempted to be transferred into Revit, and each masonry unit was set as a separate type. The transferred attributes were checked and organised by personnel to ensure they display correctly in Revit. Components with specific defects could be filtered out using the filter tool, as shown in Figure 17(e).

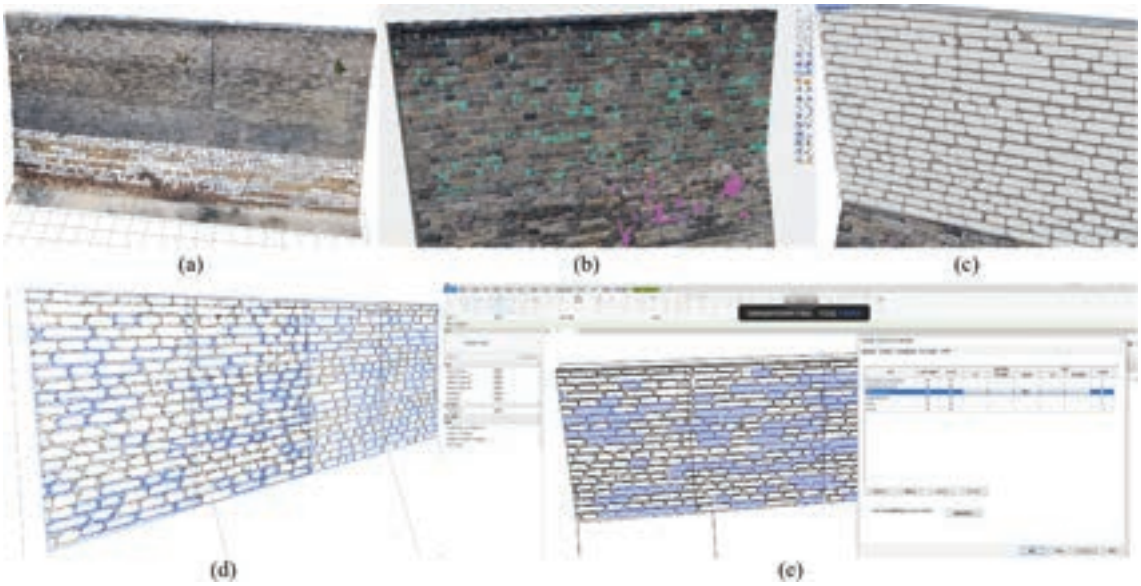

Figure 17. (a) Layer-1 mesh model; (b) Layer-2 lightweight model with mapping defects; (c) parametric masonry unit components in Layer-3; (d)&(e) bricks with specific defect are highlighted in H-BIM.

Figure 18 shows some locations of the Layer-1, Layer-2 models with mapped defect masks, along with the Layer-3 models at corresponding locations. The examples illustrate three different geometries of the bricks, each with varying sizes, contours, and arrangements. In locations (I) and (II), the bricks are arranged in a neat fashion with one long and one short, but with inconsistent sizes. In locations (III) and (IV), the bricks are the most regular, featuring an arrangement of four long and four short in an orderly manner. Location (Ⅴ) presents bricks that are small, irregular, and arranged in a mixed horizontal and vertical orientation. The contours of these bricks can be accurately identified by the trained model and converted into parametric components. The thickness of the brick components can be adjusted via parameters, with (I) demonstrating a portion of the contours extending outward, (II) illustrating the bricks extending inward, making the bricks embedded within the wall, and (III), (IV), (Ⅴ) showing the brick contours converted into mesh models and combined with the wall surface. If the thickness of individual bricks and stones are measured during field surveys, a specific thickness can be assigned to each block. The examples indicate that the parametric brick components not only accurately reflect the realistic masonry arrangement, but also can identify and reflect the shapes of missing elements and cracks (as shown at the positions circled in red in (III) and (IV)). A small number of positions with lower accuracy are primarily affected by plants covering the wall, causing difficulty in accurately capturing brick contours, which requires manual adjustments. Furthermore, the examples demonstrate that the defects match the corresponding bricks. To clearly display defect positions, one type of defect is selected for demonstration, where models (I) and (II) are efflorescence, models (III) and (IV) are deep erosion, and model (Ⅴ) illustrates biological colonization. The defect masks in Layer-2 are registered with corresponding bricks and assigned as parameters in Layer-3, highlighted in blue.

6. Discussion

The methodology constitutes a workflow for defect identification and fusion with masonry H-BIM: Firstly, CNN models are trained to identify defects in the collected images, with quantifiable classification rules established to enhance the consistency of dataset labelling, and then with data augmentation and k-fold cross-validation introduced to improve training quality. Second comes the generation of a masonry facade H-BIM that reflects the realistic geometry of masonry units, including 3D reconstruction in Layer-1, model lightweighting and texture mapping in Layer-2, CNN-based masonry unit contour extraction, parametric modelling of masonry unit components, and defect information assignment in Layer-3. The final H-BIM will be used for subsequent numerical simulation of the masonry structural built heritage in future work.

The proposed method was implemented at Pan Gate ancient wall, a representative historical masonry structure in China. Several wall sections were chosen, showcasing complex defect scenarios and various forms of masonry arrangements, to validate the methodology's applicability.

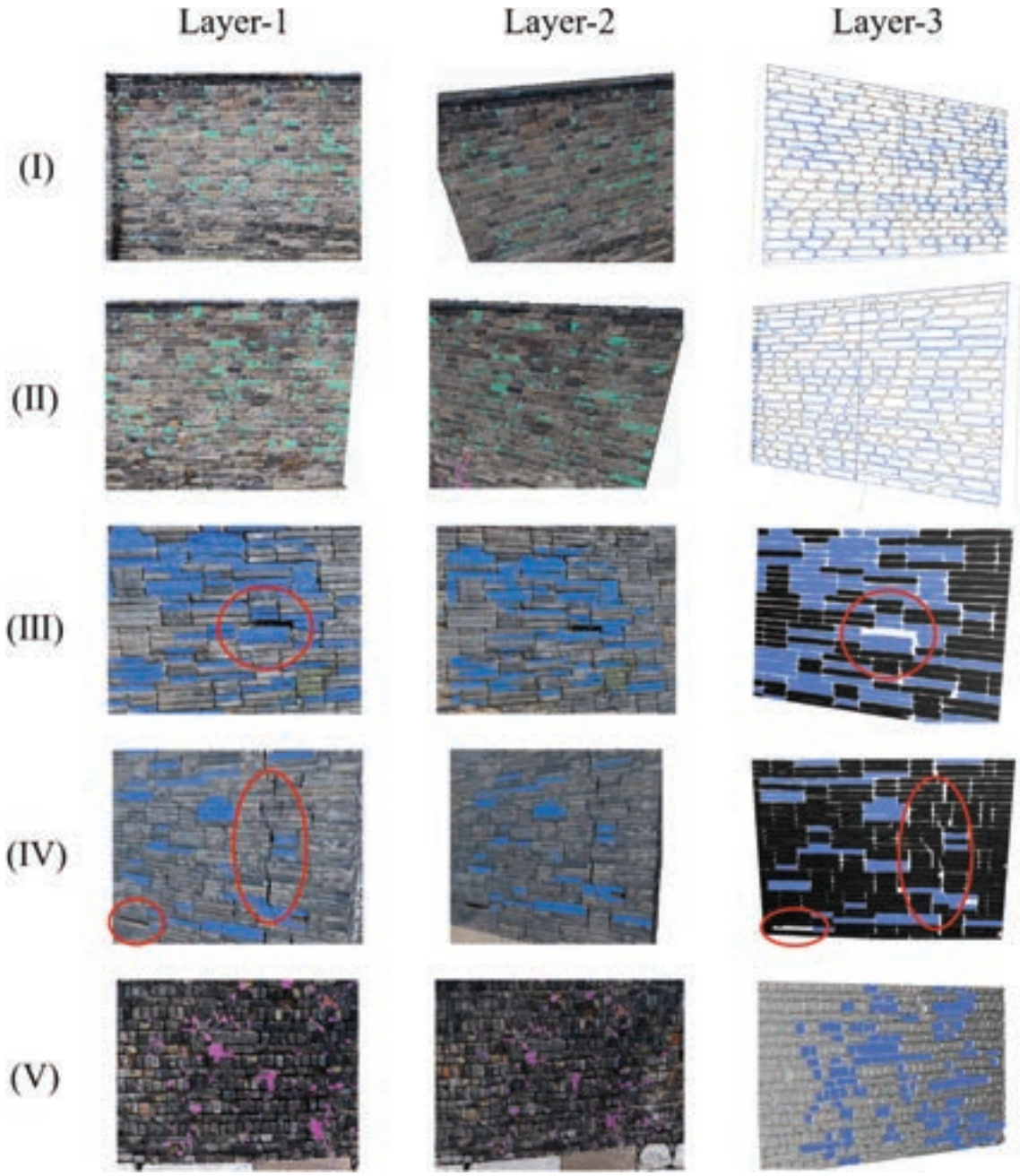

Figure 18. Result of the defect augmented three-Layer masonry H-BIM.

UAV photography was employed for efficient data acquisition, and the collected images were used to train CNN-based YOLO models. In training set establishment, quantifiable classification rules were developed, including a two-step screening process that integrates multiple defect characteristics into the criteria. Specifically, image features such as GLCM and gradient were extracted, ensuring that labelling was not reliant on subjective decisions. Previous research on defect labelling often only referenced visual appearance features from guidance documents and failed to specify the criteria, resulting in ambiguous defect classification. In contrast, the quantifiable classification rules proposed in the current study ensure consistent labelling across different personnel. Currently, these quantifiable classification rules have been validated in labelling the Pan Gate dataset. Due to the limited relevant existing research available for parallel comparison at present, if these specific criteria are applied to label new datasets in future applications, some boundary values may need to be exploratory adjusted to achieve better applicability.

Data augmentation and k-fold cross-validation were introduced to address the challenge of inadequate sample size and category imbalance in defect datasets of

masonry structural built heritage such as Pan Gate. Various strategies were tested on the training sets, involving different sample sizes, with or without k-fold cross-validation, and utilising different pre-trained models. The results indicate that data augmentation and k-fold cross-validation enhanced training quality. While larger pre-trained models can also lead to improvements, they demand more computational resources and longer training times. The YOLO model, trained using the proposed method's training set and strategy, demonstrates evaluation results for performance metrics as follows: Precision, Recall, mAP50, and mAP50-95 reach 0.966, 0.948, 0.964, and 0.740, respectively, achieving a comparable level to previous research. The trained model has delivered satisfying outcomes at Pan Gate, successfully identifying and segmenting most defects, distinguishing extents of erosion, shapes of plant vines, and efflorescence. Compared to previous research, the model's ability to handle multi-category and complex scenarios has been validated.

H-BIM models were created for various sections of the Pan Gate ancient wall using the proposed methodology. The H-BIM has multiple presentation forms, among which Layer-3 represents the innovation of the current study. A similarly trained YOLO model was used for masonry unit identification; except for a small number of bricks obscured by plants, almost all brick contours were successfully extracted and automatically converted into parametric brick components. Meanwhile, the previously identified defects were assigned to the corresponding bricks, enabling a more detailed distribution of defect information. After generating masonry arrangements, joint morphology, crack conditions, and missing parts could be displayed accordingly. The thickness of each brick and stone can also be adjusted based on on-site survey measurements. Compared with masonry numerical simulation models from previous research, the model developed in the current study features masonry units with more realistic geometry and pre-embedded defect information, thereby helping subsequent numerical simulations to be more reliable.

Although the current case study was conducted on planar walls, it is believed that the proposed process for creating masonry units through texture mapping, contour identification, parametric modelling, and attribute assignment is not limited to planar structures but has the potential to be extended to more complex structures, such as the curved surfaces of smokestacks and the spherical surfaces of Islamic-style domes, which will be further explored in future work.

The limitations of the current study stem mainly from the fact that UAV photogrammetry can capture only surface information. A few bricks exhibit inaccuracies in their identified contours due to plant coverings. Potential improvements include deploying additional penetrating instruments to assess the geometry of bricks beneath plants and employing algorithms to approximate the curvature of brick contours. Additionally, the current masonry H-BIM only incorporates facade information. Therefore, future work will supplement the wall's internal structure to make the H-BIM more comprehensive. Given the inaccessibility of internal structures and the non-destructive requirement for built heritage, non-destructive techniques such as ground-penetrating radar (GPR) may serve as viable detection instruments, in conjunction with referencing ancient documents to extract design principles for deducing internal structures.

7. Conclusion

This paper presents an optimised workflow to generate a defect augmented H-BIM for masonry structural built heritage which covers automated defect identification, masonry H-BIM establishment and defect information integration, facilitating subsequent numerical simulations for safety monitoring and risk assessment. The method establishes quantifiable classification rules and introduces data augmentation and k-fold cross-validation in dataset preparation and model training to enhance identification accuracy. A multi-layer masonry H-BIM is proposed, where individual masonry unit components are parametrically generated, and enriched with corresponding defect information.

The case study shows that the established quantifiable classification rules ensure labelling consistency, and the introduced methods effectively improve training quality. The developed model accurately segmented defects of diverse geometry and damage extent in complex scenario. The proposed masonry unit generation method straightforwardly establishes the masonry arrangement on the wall facade, reflecting the realistic masonry unit geometries as well as the cracking and missing conditions. The identified defect information is distributed in detail to each corresponding masonry unit for subsequent numerical simulations. The combined method of vector projection and texture identification for brick extraction and parametric modelling also has the potential to be applied to non-planar masonry structures, and further exploration will be conducted on more diverse structures.

In future work, additional data acquisition instruments and technologies will be employed to address the aforementioned limitations, including the use of penetrable instruments, referencing historical design

documents, and introducing new algorithms to infer the internal structure of the ancient wall. Supplemental information is planned to be added into the H-BIM model, including internal materials, wall loads, structural deformation, and the surrounding environment, to establish a comprehensive model with internal and external information for simulation.

Acknowledgement

The authors would like to thank the support from the Design School Intelligent Built Environment Research Centre, Xi'an Jiaotong-Liverpool University.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

No funding was received.

Data availability statement

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

References

Abu-Haifa, M., and S. J. Lee. 2023. "Image-Based 3D Modeling-to-Simulation of Single-Wythe Masonry Structure via Reverse Descriptive Geometry." *Journal of Building Engineering* 76:107125. https://doi.org/10.1016/j.jobe.2023.107125

Alforno, M., A. Monaco, F. Venuti, C. Calderini, G. Sammartano, G. Patrucco, and A. Spanó. 2024. "An Application of a Reality-Based Approach for the Generation of Masonry Cross Vaults FE Block Models." *Structures* 65:106638. https://doi.org/10.1016/j.istruc.2024.106638

Arlot, S., and A. Celisse. 2010. "A Survey of Cross-Validation Procedures for Model Selection." *Statistics Surveys* 4 (none). https://doi.org/10.1214/09-SS054

Bhagat, P. K., P. Choudhary, and K. M. Singh. 2019. "A Comparative Study for Brain Tumor Detection in MRI Images Using Texture Features." In *Sensors for Health Monitoring*, edited by N. Dey, J. Chaki and R. Kuma, 259–287. Elsevier. https://doi.org/10.1016/B978-0-12-819361-7.00013-0

Carvalho Ottoni, A. L., and L. T. Cordeiro Ottoni. 2025. "A Deep Learning Approach for Cultural Heritage Building Classification Using Transfer Learning and Data Augmentation." *Journal of Cultural Heritage* 74:214–224. https://doi.org/10.1016/j.culher.2025.06.010

Casas, E., L. Ramos, C. Romero, and F. Rivas-Echeverría. 2024. "A Comparative Study of YOLOv5 and YOLOv8 for Corrosion Segmentation Tasks in Metal Surfaces." *Array* 22:100351. https://doi.org/10.1016/j.array.2024.100351

Cavalagli, N., A. Kita, V. L. Castaldo, A. L. Pisello, and F. Ubertini. 2019. "Hierarchical Environmental Risk Mapping of Material Degradation in Historic Masonry Buildings: An Integrated Approach Considering Climate Change and Structural Damage." *Construction and Building Materials* 215:998–1014. https://doi.org/10.1016/j.conbuildmat.2019.04.204

Chen, J., W. Lu, and D. Liu. 2024. "Built Environment Defect Mapping, Modeling, and Management (D3M): A BIM-Based Integrated Framework." *Journal of Intelligent Construction* 2 (1): 1–15. https://doi.org/10.26599/JIC.2024.9180008

Chen, J., W. Lu, and J. Lou. 2023. "Automatic Concrete Defect Detection and Reconstruction by Aligning Aerial Images onto Semantic-Rich Building Information Model." *Computer-Aided Civil and Infrastructure Engineering* 38 (8): 1079–1098. https://doi.org/10.1111/mice.12928

Cuadros-Rojas, E., S. Saloustros, N. Tarque, and L. Pelà. 2024. "Photogrammetry-Aided Numerical Seismic Assessment of Historical Structures Composed of Adobe, Stone and Brick Masonry Application to the San Juan Bautista Church Built on the Inca Temple of Huaytará Peru." *Engineering Failure Analysis* 158:107984. https://doi.org/10.1016/j.engfailanal.2024.107984

D'Orazio, M., A. Gianangeli, F. Monni, and E. Quagliarini. 2024. "Automatic Monitoring of the Bio Colonisation of Historical Building's Facades Through Convolutional Neural Networks (CNN)." *Journal of Cultural Heritage* 70:80–89. https://doi.org/10.1016/j.culher.2024.08.012

Du, H., J. Yu, Y. Wang, Y. Zhu, Y. Tang, and H. Wang. 2022. "Visualized Failure Prediction for the Masonry Great Wall." *Buildings* 12 (12): 2224. https://doi.org/10.3390/buildings12122224

Du, Z., P. Luo, X. Zhu, Y. Zhang, and Y. Zhu. 2014. "Texture Optimization Methodology for 3D Building Based on Super Face." *Geomatics and Information Science of Wuhan University* 39 (12): 1401–1405. http://ch.whu.edu.cn/en/article/id/3131.

Dutschmann, T.-M., L. Kinzel, A. Ter Laak, and K. Baumann. 2023. "Large-Scale Evaluation of K-Fold Cross-Validation Ensembles for Uncertainty Estimation." *Journal of Cheminformatics* 15 (1): 49. https://doi.org/10.1186/s13321-023-00709-9

Ehab Moustafa Kamel, K., and T. J. Massart. 2025. "Towards Automated Image-Based Cohesive Zone Modeling of Cracking in Irregular Masonry." *Mathematics and Mechanics of Solids* 30 (1): 116–136. https://doi.org/10.1177/10812865241228825

Franklin, W. R. 2006. "Pnpoly-Point Inclusion in Polygon Test. *Web Site:* WWW.Ecse.Rpi.Edu/Homepages/Wrf/Research/Short_Notes/Pnpoly.Html.

Haralick, R. M., K. Shanmugam, and I. Dinstein. 1973. "Textural Features for Image Classification." *IEEE Transactions on Systems, Man and Cybernetics SMC*-3 (6): 610–621. https://doi.org/10.1109/TSMC.1973.4309314

Hatir, M. E., M. Barstuğan, and İ. İnce. 2020. "Deep Learning-Based Weathering Type Recognition in Historical Stone Monuments." *Journal of Cultural Heritage* 45:193–203. https://doi.org/10.1016/j.culher.2020.04.008

Huang, H., Y. Cai, C. Zhang, Y. Lu, A. Hammad, and L. Fan. 2024. "Crack Detection of Masonry Structure Based on Thermal and Visible Image Fusion and Semantic Segmentation." *Automation in Construction* 158:105213. https://doi.org/10.1016/j.autcon.2023.105213

Jocher, G., and J. Qiu. 2024. Ultralytics YOLO11. https://github.com/ultralytics/ultralytics.

Karimi, N., M. Mishra, and P. B. Lourenço. 2024. "Deep Learning-Based Automated Tile Defect Detection System for Portuguese Cultural Heritage Buildings." *Journal of Cultural Heritage* 68:86–98. https://doi.org/10.1016/j.culher.2024.05.009

Karimi, N., N. Valibeig, and H. R. Rabiee. 2024. "Deterioration Detection in Historical Buildings with Different Materials Based on Novel Deep Learning Methods with Focusing on Isfahan Historical Bridges." *International Journal of Architectural Heritage* 18 (6): 981–993. https://doi.org/10.1080/15583058.2023.2201576

Keshmiry, A., S. Hassani, U. Dackermann, and J. Li. 2024. "Assessment, Repair, and Retrofitting of Masonry Structures: A Comprehensive Review." *Construction and Building Materials* 442:137380. https://doi.org/10.1016/j.conbuildmat.2024.137380

Li, Q., L. Zheng, Y. Chen, L. Yan, Y. Li, and J. Zhao. 2023. "Non-Destructive Testing Research on the Surface Damage Faced by the Shanhaiguan Great Wall Based on Machine Learning." *Frontiers in Earth Science* 11:11. https://doi.org/10.3389/feart.2023.1225585

Li, S., W. Song, H. Qin, and A. Hao. 2018. "Deep Variance Network: An Iterative, Improved CNN Framework for Unbalanced Training Datasets." *Pattern Recognition* 81:294–308. https://doi.org/10.1016/j.patcog.2018.03.035

Liu, D., X. Xia, J. Chen, and S. Li. 2021. "Integrating Building Information Model and Augmented Reality for Drone-Based Building Inspection." *Journal of Computing in Civil Engineering* 35 (2). https://doi.org/10.1061/(ASCE)CP.1943-5487.0000958

Loverdos, D., V. Sarhosis, E. Adamopoulos, and A. Drougkas. 2021. "An Innovative Image Processing-Based Framework for the Numerical Modelling of Cracked Masonry Structures." *Automation in Construction* 125:103633. https://doi.org/10.1016/j.autcon.2021.103633

Luo, S., and H. Wang. 2024. "Digital Twin Research on Masonry–Timber Architectural Heritage Pathology Cracks Using 3D Laser Scanning and Deep Learning Model." *Buildings* 14 (4): 1129. https://doi.org/10.3390/buildings14041129

Mansuri, L. E., and D. A. Patel. 2022. "Artificial Intelligence-Based Automatic Visual Inspection System for Built Heritage." *Smart & Sustainable Built Environment* 11 (3): 622–646. https://doi.org/10.1108/SASBE-09-2020-0139

Meklati, S., K. Boussora, M. E. H. Abdi, and S.-A. Berrani. 2023. "Surface Damage Identification for Heritage Site Protection: A Mobile Crowd-Sensing Solution Based on Deep Learning." *Journal on Computing and Cultural Heritage* 16 (2): 1–24. https://doi.org/10.1145/3569093

Milani, G., and P. B. Lourenço. 2010. "A Simplified Homogenized Limit Analysis Model for Randomly Assembled Blocks Out-of-Plane Loaded." *Computers & Structures* 88 (11–12): 690–717. https://doi.org/10.1016/j.compstruc.2010.02.009

Mishra, M., T. Barman, and G. V. Ramana. 2024. "Artificial Intelligence-Based Visual Inspection System for Structural Health Monitoring of Cultural Heritage." *Journal of Civil Structural Health Monitoring* 14 (1): 103–120. https://doi.org/10.1007/s13349-022-00643-8

Mishra, M., and P. B. Lourenço. 2024. "Artificial Intelligence-Assisted Visual Inspection for Cultural Heritage: State-of-the-Art Review." *Journal of Cultural Heritage* 66:536–550. https://doi.org/10.1016/j.culher.2024.01.005

Nan, L., and P. Wonka. 2017. "Polyfit: Polygonal Surface Reconstruction from Point Clouds." *Proceedings of the IEEE International Conference on Computer Vision*, Venice, Italy, October 22–29, 2017, 2353–2361.

Nieto-Julián, E., M. D. Robador, J. Moyano, and S. Bruno. 2025. "Semantic HBIM for Heritage Conservation: A Methodology for Mapping Deterioration and Structural Deformation in Historic Envelopes." *Buildings* 15 (12): 1990. https://doi.org/10.3390/buildings15121990

Nieto-Julián, J. E., J. Farratell, M. Bouzas Cavada, and J. Moyano. 2023. "Collaborative Workflow in an HBIM Project for the Restoration and Conservation of Cultural Heritage." *International Journal of Architectural Heritage* 17 (11): 1813–1832. https://doi.org/10.1080/15583058.2022.2073294

Pantoja-Rosero, B. G., R. Achanta, and K. Beyer. 2023. "Damage-Augmented Digital Twins Towards the Automated Inspection of Buildings." *Automation in Construction* 150:104842. https://doi.org/10.1016/j.autcon.2023.104842

Pantoja-Rosero, B. G., R. Achanta, M. Kozinski, P. Fua, F. Perez-Cruz, and K. Beyer. 2022. "Generating LOD3 Building Models from Structure-from-Motion and Semantic Segmentation." *Automation in Construction* 141:104430. https://doi.org/10.1016/j.autcon.2022.104430

Pantoja-Rosero, B. G., S. Saloustros, R. Achanta, and K. Beyer. 2023. "Image-Based Geometric Digital Twinning for Stone Masonry Elements." *Automation in Construction* 145:104632. https://doi.org/10.1016/j.autcon.2022.104632

Pereira, M., A. M. D'Altri, S. de Miranda, and B. Glisic. 2023. "Automatic Multi-Leaf Nonperiodic Block-by-Block Pattern Generation and Computational Analysis of Historical Masonry Structures." *Engineering Structures* 283:115945. https://doi.org/10.1016/j.engstruct.2023.115945

Seitz, S. M., B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. 2006. "A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms." *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition* 1 (CVPR'06): 519–528. https://doi.org/10.1109/CVPR.2006.19

Shaqfa, M., and K. Beyer. 2022. "A Virtual Microstructure Generator for 3D Stone Masonry Walls." *European Journal of Mechanics - A/Solids* 96:104656. https://doi.org/10.1016/j.euromechsol.2022.104656

Sobel, I. E. 1970. *Camera Models and Machine Perception*. stanford university.

Sun, Z., J. Xie, Y. Zhang, and Y. Cao. 2019. "As-Built BIM for a Fifteenth-Century Chinese Brick Structure at Various LODs." *ISPRS International Journal of Geo-Information* 8 (12): 577. https://doi.org/10.3390/ijgi8120577

Tan, Y., G. Li, R. Cai, J. Ma, and M. Wang. 2022. "Mapping and Modelling Defect Data from UAV Captured Images to BIM for Building External Wall Inspection." *Automation in Construction* 139:104284. https://doi.org/10.1016/j.autcon.2022.104284

Tian, X., X. Ruan, and Y. Wang. 2002. "Mesh Simplification Based on Super-Face and Genetic Algorithm in Reverse Engineering." *International Journal of Advanced Manufacturing Technology* 20 (4): 303–312. https://doi.org/10.1007/s001700200156

Tiberti, S., and G. Milani. 2020. "3D Voxel Homogenized Limit Analysis of Single-Leaf Non-Periodic Masonry." *Computers & Structures* 229:106186. https://doi.org/10.1016/j.compstruc.2019.106186

Valero, E., A. Forster, F. Bosché, E. Hyslop, L. Wilson, and A. Turmel. 2019. "Automated Defect Detection and Classification in Ashlar Masonry Walls Using Machine Learning." *Automation in Construction* 106:102846. https://doi.org/10.1016/j.autcon.2019.102846

Vergès-Belmin, V. 2008. *Illustrated Glossary on Stone Deterioration Patterns*. Icomos.

Wan, X., X. Zhang, and L. Liu. 2021. "An Improved VGG19 Transfer Learning Strip Steel Surface Defect Recognition Deep Neural Network Based on Few Samples and Imbalanced Datasets." *Applied Sciences* 11 (6): 2606. https://doi.org/10.3390/app11062606

Wang, N., Q. Zhao, S. Li, X. Zhao, and P. Zhao. 2018. "Damage Classification for Masonry Historic Structures Using Convolutional Neural Networks Based on Still Images." *Computer-Aided Civil and Infrastructure Engineering* 33 (12): 1073–1089. https://doi.org/10.1111/mice.12411

Wang, N., X. Zhao, P. Zhao, Y. Zhang, Z. Zou, and J. Ou. 2019. "Automatic Damage Detection of Historic Masonry Buildings Based on Mobile Deep Learning." *Automation in Construction* 103:53–66. https://doi.org/10.1016/j.autcon.2019.03.003

Wang, N., X. Zhao, Z. Zou, P. Zhao, and F. Qi. 2020. "Autonomous Damage Segmentation and Measurement of Glazed Tiles in Historic Buildings via Deep Learning." *Computer-Aided Civil and Infrastructure Engineering* 35 (3): 277–291. https://doi.org/10.1111/mice.12488

Yan, L., Y. Chen, L. Zheng, and Y. Zhang. 2024. "Application of Computer Vision Technology in Surface Damage Detection and Analysis of Shedthin Tiles in China: A Case Study of the Classical Gardens of Suzhou." *Heritage Science* 12 (1): 72. https://doi.org/10.1186/s40494-024-01185-6

Yan, Y., and Y. Wang. 2025. "Deterioration Investigation and Numerical Stability Analysis of High Pedestal in Chinese Ancient City Towers: A Case Study of the Square Wall at the Ming Dynasty Xianling Tomb." *Case Studies in Construction Materials* 22:e04705. https://doi.org/10.1016/j.cscm.2025.e04705

Yang, L., K. Liu, R. Ou, P. Qian, Y. Wu, Z. Tian, C. Zhu, S. Feng, and F. Yang. 2024. "Surface Defect-Extended BIM Generation Leveraging UAV Images and Deep Learning." *Sensors* 24 (13): 4151. https://doi.org/10.3390/s24134151

Yang, X., L. Zheng, Y. Chen, J. Feng, and J. Zheng. 2023. "Recognition of Damage Types of Chinese Gray-Brick Ancient Buildings Based on Machine Learning—Taking the Macau World Heritage Buffer Zone as an Example." *Atmosphere* 14 (2): 346. https://doi.org/10.3390/atmos14020346

Ye, Z., M. Mozafarian, P. A. R. Cavallaro, K. Altinay, V. Villa, and J. Ninić. 2026. "Automated Multi-Category Tunnel Damage Detection and Report Generation from Ultra-High-Resolution Panoramic Laser Images." *Tunnelling and Underground Space Technology* 168:107194. https://doi.org/10.1016/j.tust.2025.107194

Yue, J., Q. Yang, Z. Luo, W. Chen, and M. Yue. 2025. "Evaluation of Surface Weathering Degree of Brick Historical Buildings Based on GLCM-Entropy Weight TOPSIS: A Case Study in the Kaifeng City Wall." *International Journal of Architectural Heritage*: 1–18. https://doi.org/10.1080/15583058.2025.2455382