Deep Learning-Based Restoration of Historic Building Images with Blockchain-Driven Maintenance Data Management

Zhaokun Tang
School of Art, Shangluo University, Shangluo 726000, Shaanxi, China
E-mail: PaulBryant7871@outlook.com

Keywords: Building extraction; Blockchain; Morphology; Convolutional neural network; U-Net

Received: Septrmber 25, 2025

Ancient buildings, steeped in millennia of human history, stand as precious legacies requiring meticulous preservation and restoration. With this premise, this study proposes a restoration method for damaged ancient structures based on a deep neural network. A curated dataset of ancient building images — including diverse damage patterns and material textures-was established to train and validate the model, ensuring robustness across varied architectural styles. The core network architecture is an enhanced U-Net, in which all convolutional layers are replaced with partial convolution layers. The sliding window of partial convolutional layers exclusively performs convolution operations in the relevant region of the image, enabling accurate prediction of building structures with irregular apertures. To further enhance feature extraction and generalization, a dual-transfer learning mechanism is introduced: structural features learned from large-scale architectural datasets are transferred to fine-tune restoration on heritage-specific samples. Consequently, the semantic richness of restoration results is reconstructed. Simultaneously, in accordance with the functional requisites of the ancient building maintenance data storage and sharing system, a dedicated blockchain-driven data storage platform is designed. This platform caters to diverse stakeholders, including lay users, professional restorers, museums, and governmental bodies. The platform employs Hyperledger Fabric technology and a traditional database to ensure secure and reliable data management. Model performance is quantitatively assessed using standard image restoration metrics—Learned Perceptual Image Patch Similarity (LPIPS = 0.725), Peak Signal-to-Noise Ratio (PSNR = 26.16), and Structural Similarity Index Measure (SSIM = 0.57)—demonstrating competitive reconstruction quality. It is worth noting that the proposed building restoration method demonstrates significant potential for cost savings, while also achieving a 25% improvement in efficiency. This potential is particularly relevant within the context of restoring old buildings.

Povzetek: Študija razvije globokonevronski model z izboljšano U-Net arhitekturo in blockchain podporo, ki omogoča natančnejšo ter do 25 % učinkovitejšo obnovo poškodovanih starodavnih stavb.

1 Introduction

Historic buildings serve as indispensable symbols of human history and civilization, capturing the abundant heritage of human endeavor. In the case of China, these structures surpass their intended purpose as preservers of traditional culture, instead embodying the essence of spiritual pillars for cities with precious artistic and scientific values [1, 2]. The preservation of historic buildings has gained increasing global attention. The protection procedures in China have evolved over several decades through the assimilation of foreign experience in protection and the gradual development of a deeper comprehension of cultural heritage. Awareness and competence in heritage conservation have continuously augmented. Motivated by methods and guidelines proposed in [Zhong Ban Fa (2018) No.54] by the General Office of the State Council, People's Republic of China, there has been a growing emphasis on leveraging the potential of big data, the Internet, cloud computing, and other technical resources to fortify scientific and technological support [3]. Information technology, represented by artificial intelligence, has engendered novel avenues for presenting and exploiting cultural heritages, fostering innovation through integrating different elements. Nowadays, heritage preservation has seen continuous promotion. However, the progress of time requires a shift from traditional conservation methods to a more structured, scientific, and comprehensive approach to managing heritage information. This evolution encompasses in-depth excavation, analysis, and an all-encompassing display of heritage value, culminating in dynamic monitoring and early warning of both the heritage sites and their surrounding environment. This phenomenon illustrates the transformation from a reactive approach focused on rescue protection to a proactive and systematic approach centered on preventive conservation, emphasizing the urgent need for enhanced preservation and management of cultural heritage. In the field of

historic building restoration, the convergence of the construction industry with digital technology has led to a surge of digital works in architectural domains. In the current context, the protection of digital property rights has become a pressing problem [4-6].

Numerous authors have conducted notable investigations in architectural heritage conservation, contributing significantly to the discipline. Fadli and Al Saeed (2019) [7] explored Qatar's existing urban conservation strategies, architectural heritage, and modern digitization and archival principles. They aimed to establish a sustainable and interactive archival platform, namely an information modeling platform, for historic buildings in Qatar. Müller et al. (2025) [8] proposed depth inpainting under perspective occlusion environments. If the mask M_0 is provided by user input or through object detection and segmentation, the object can be removed via the inpainting process. This alignment was observed through the establishment of a two-tier database according to Midas standards. Masciotta et al. (2021) [9] explored the significant role of digitization in heritage conservation, drawing insights from several programs focused on the monitoring and preventative protection of historical and cultural heritage. Their endeavor cultivated an integrated and technologically advanced approach, enhancing the process of documenting historical structures and encouraging the active participation of property owners in the preservation and restoration efforts. The defects found in ancient house murals were examined by Li (2021) [10], who identified their origins in both natural and manmade sources. The research improved existing approaches for restoring damaged murals by employing a multifaceted approach involving fracture identification, texture-plane segmentation, and the application of the discrete difference algorithm. Wang and Zhao (2022) [11] advocated the application of 3D images and Virtual Reality (VR) technologies in art, design, and the digital preservation and restoration of ancient buildings. Their work focused on an integrated system for restoring ancient buildings, which incorporated the use of 3D stereoscopy and VR. This technology holds substantial potential for wider implementation in art and design domains. Siountri et al. (2020) [12] underscored the transformative impact of Building Information Modeling (BIM) technology. However, the interplay of Internet of Things (IoT) devices and interconnections between mobile components and stakeholders in a modular environment have given rise to security concerns. The study examined the integration of the application of IoT and blockchain technologies in the construction industry, highlighting their cohesiveness and interoperability within the system architecture and emphasizing building applications. Nawari and Ravindran (2019) [13] explored the increasing adoption and impact of blockchain technology, lauded for its pervasiveness across diverse industries in both public and private sectors. Their study examined how blockchain technology could impact the construction and engineering sectors, evaluating its potential in addressing post-disaster reconstruction challenges. Dey et al. (2023) [14] employed machine learning techniques to segment building roofs using airborne light detection and range (LiDAR) point cloud data. They aimed to enhance the efficiency of LiDAR data processing through machine learning, facilitating automatic identification and segmentation of building roofs. They introduced a novel approach with applications in the remote sensing domain. Gharineiat et al. (2022) [15] comprehensively reviewed automated methods using machine learning for terrain and surface feature recognition in LiDAR data. The study summarized recent research outcomes, focusing on the application of machine learning techniques in LiDAR data processing, providing valuable insights for geography and remote sensing. Mirzaei et al. (2022) [16] thoroughly reviewed machine learning-based 3D point cloud data processing in architecture and infrastructure applications. The research reviewed the applications of machine learning techniques in architecture and infrastructure, emphasizing their significance in data processing and applications, offering profound insights into the field of engineering informatics. Khan et al. (2023) [17] introduced a deep learning framework for extracting building outlines from aerial images. The authors utilized an encoder-decoder structure to achieve building outline extraction through deep learning technology, aiming to enhance the efficiency and accuracy of this process in remote sensing image processing. Vacchio and Bifulco (2022) [18] explored the application of blockchain technology in the cultural heritage domain, providing a comprehensive literature review of its potential value and challenges. The authors summarized the role of blockchain technology in existing research concerning digital reconstruction, cultural heritage preservation, and digitalization of cultural heritage. However, the study did not delve into a detailed analysis of the methods and outcomes of the existing research. Trček (2022) [19] summarized the opportunities and challenges of blockchain technology in heritage preservation. The author emphasized the potential uses of blockchain technology, including cultural heritage digital reconstruction, data protection, and traceability. While the literature offers some practical cases showcasing the application of blockchain technology in cultural heritage preservation, it does not provide detailed technical details or specific operational methods. Although the above research has made significant progress in the field of architectural heritage protection, overall, the existing deep learning and blockchain frameworks still have limitations in meeting the three requirements of "multi-stakeholder

collaboration, security assurance and low-latency transmission".

At the same time, the introduction of blockchain technology is not only for the secure storage and sharing of data, but also for its ability to address the issues of trust and traceability in the maintenance data of historical buildings, especially in multi-party participation, long-term management, and cross-institutional collaboration. Traditional databases find it hard to ensure the true origin and non-tamperability of data. In contrast, the decentralized and tamper-proof mechanisms of blockchain can ensure the transparency and credibility of the data flow process, thus establishing a long-term basis of trust for the protection of architectural heritage. Therefore, this study fills the gaps in previous research concerning the application of deep learning and blockchain technology, offering new techniques and methods for protecting and maintaining historical buildings, holding both academic and practical significance.

2 Design of building restoration model(U-Net and the improved U-Net)

The U-Net is a classical Convolutional Neural Network (CNN) introduced at the 2015 conference. Since its inception, extensive discourse among researchers about its network architecture has led to numerous studies focusing on enhancing the U-Net structure. The U-Net architecture features a symmetrical U-shaped configuration comprising contracting and expanding paths. The contracting path embodies a conventional CNN design with a repetitive structure involving two convolutional operations and a pooling operation in each iteration. The convolutional layer utilizes a uniform kernel size of 3×3, applies the Rectified Linear Unit (ReLU) for activation functions, and incorporates maximum pooling. After each down-sampling operation, the number of feature maps is doubled. In contrast, each stage within the expanding path begins with up-sampling, resulting in a reduction of the number of feature maps by half while doubling their dimensions. Simultaneously, the outcome is superimposed with the appropriate feature map of the respective region from the contracting path. To ensure congruence, trimming is conducted before overlaying due to the slight variance in feature map sizes. The final stage involves performing two 3×3 and one l×l convolution operations on the overlaid feature map. However, the network faces two significant challenges. Firstly, the efficiency of the network is degraded due to computations being performed iteratively for each neighborhood, including repetitive calculations for overlapping neighborhoods. Additionally, the network encounters a trade-off between achieving accurate localization and gathering contextual information. Choosing a larger patch size requires the inclusion of extra maximum pooling layers, compromising the precision of object location. Conversely, opting for a smaller neighborhood results in a reduction of contextual information. Figure 1 illustrates the architectural representation of the U-Net network.

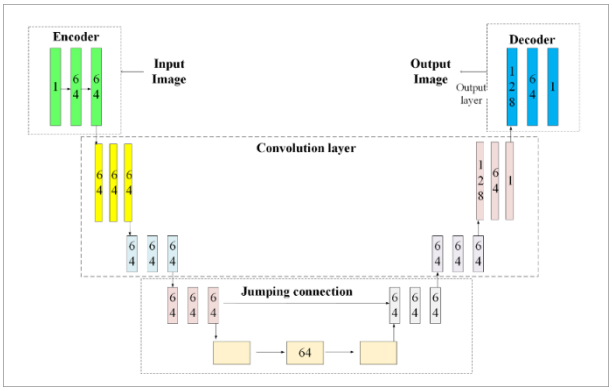

Figure 1: The network structure of U-Net

When employing the U-Net network structure for feature extraction and modeling design in architectural images, Figure 1 illustrates the architecture of this network. In this context, the input image represents architectural images or related data in the field, with the primary research objective being semantic segmentation—separating buildings from the background to explore the features, structures, and properties of the buildings. This study selects the U-Net network structure due to its exceptional performance in image segmentation tasks, simultaneously preserving detailed and semantic information. This provides a robust tool for achieving the research goals. Through this structure, this study aims to present a transparent and organized approach, elucidating how this algorithm extracts features from architectural images and its practical applications in architectural modeling and design.

Building upon the previously discussed content, this study aims to optimize the effectiveness of the U-Net network in capturing intricate texture details and global semantic structures efficiently. The objective is to achieve accurate and comprehensive image restoration by incorporating content loss functions and dual-transfer operations into the U-Net network. The dual-transfer operation is a mechanism introduced to enhance the restoration of significantly missing architectural images while preserving textures and global semantic structures. The core concept involves repairing missing regions by leveraging information from known regions to obtain updated estimations of image information. Unlike the original U-Net, which only performs inter-layer skip connections to transfer local features, the dual transfer mechanism performs feature transfer and re-encoding between two parallel networks to achieve cross-scale semantic and texture dual fusion. Within the U-Net network structure, the dual-transfer operation encompasses the establishment of transfer connection layers, capturing encoder and decoder features, identifying missing regions, and conducting nearest-neighbor searches. This operation facilitates the sharing of

feature representations between two parallel main structures, ensuring that estimated values for missing regions receive ample information while retaining crucial image details. Ultimately, this approach enhances the quality of restoration. The optimization process of the U-Net network using content loss functions and dual-transfer operations is delineated as follows: The U-Net framework incorporates two separate transfer connection layers.

The aforementioned layers have a parallel structure that mirrors the topology of the L layer in U-Net. The provided formulation denotes $\boldsymbol{\Phi_l(I)}$ and $\boldsymbol{\Phi_{L-l}(I)}$ as the encoder features that collect information from the $\boldsymbol{l}$-th and $(\boldsymbol{L-l})$-th layers, respectively. The objective is for $\boldsymbol{\Phi_l(I)}$ and $\boldsymbol{\Phi_{L-l}(I)}$ to encompass the majority of the relevant information contained in $\boldsymbol{\Phi_l(I^{gt})}$. For every arbitrary position $\boldsymbol{y \in \Omega}$, it follows that $\boldsymbol{(\Phi_l(I))_y \approx 0}$. Consequently, $\boldsymbol{(\Phi_{L-l}(I))_y \approx 0}$, indicating that it provides the same information as $\boldsymbol{\Phi_l(I^{gt})_y}$. The transfer connection layers $\boldsymbol{\Phi^{shift}_{L-l}(I)}$ and $\boldsymbol{\Phi^{shift}_{L-l+1}(I)}$ utilize $\boldsymbol{\Phi_l(I)}$ and $\boldsymbol{\Phi_{L-l}(I)}$, as well as $\boldsymbol{\Phi_{L-l}(I)}$ and $\boldsymbol{\Phi_{L-l+1}(I)}$, to derive an updated estimation of $\boldsymbol{\Phi_l(I^{gt})}$. The procedure involves the identification of the unknown region $\boldsymbol{\Omega}$ and the known region $\boldsymbol{\bar{\Omega}}$. Eq. (1) can be utilized to perform a nearest neighbor search (NNS) based on $\boldsymbol{\Phi_l(I)_x}$, where $\boldsymbol{x \in \bar{\Omega}}$, for each $\boldsymbol{(\Phi_{L-l}(I))_y}$, with $\boldsymbol{y \in \Omega}$.

$$\boldsymbol{X^*(y) = \arg\max_{x \in \bar{\Omega}} \frac{<(\Phi_{L-l}(I))_y, (\Phi_l(I))_x>}{\|(\Phi_{L-l}(I))_y\|_2 \|(\Phi_l(I))_x\|_2}} \quad (1)$$

For each $\boldsymbol{(\Phi_{L-l+1}(I))}$, with $\boldsymbol{y \in \Omega}$, an NNS based on $\boldsymbol{\Phi_{l-1}(I)_x}$ is conducted, where $\boldsymbol{x \in \bar{\Omega}}$, as delineated by Eq. (2).

$$\boldsymbol{X^*(y) = \arg\max_{x \in \bar{\Omega}} \frac{\langle(\Phi_{L-l+1}(I))_y, (\Phi_{l-1}(I))_x\rangle}{\|(\Phi_{L-l+1}(I))_y\|_2 \|(\Phi_{l-1}(I))_x\|_2}} \quad (2)$$

The estimations of $\boldsymbol{\Phi_l(I^{gt})_y}$ and $\boldsymbol{\Phi_{l-1}(I^{gt})_y}$ are subsequently revised through spatial rearrangements concerning the encoder features $\boldsymbol{\Phi_l(I)_x}$ and $\boldsymbol{\Phi_{l-1}(I)_x}$ [20]. The mathematical formulation of this can be expressed as follows:

$$\boldsymbol{\left(\Phi^{\text{shift}}_{L-l}(I)\right)_y = (\Phi_l(I))_{y+u_y}} \quad (3)$$

$$\boldsymbol{\left(\Phi^{\text{shift}}_{L-l+1}(I)\right)_y = (\Phi_{l-1}(I))_{y+u_y}} \quad (4)$$

Figure 2 portrays the U-Net structure with two special transfer connection layers.

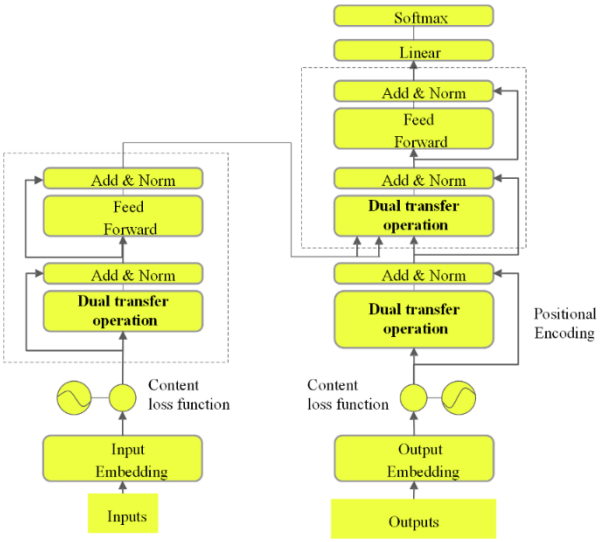

Figure 2: U-Net with two special transfer connection layers

A content loss function based on advanced image features is introduced to enhance the network's capability to capture intricate texture details. This approach involves extracting the high-level feature representations of both the decoder and encoder features corresponding to the original image's missing region. Subsequently, the mean squared error (MSE) is computed. The encoder examines the feature information of the image and extracts local features from it, while the decoder utilizes these local features to produce complex spatial and detailed structural information for the fragmented areas. The integration of encoders and decoders is facilitated through a hopping connection, which combines predictions from various feature layers. This mechanism allows for concurrently predicting local and global information, thereby enhancing the decoder's ability to effectively restore missing regions.

This study proposes the content loss function to depict the association between $\boldsymbol{\Phi_{L-l}(I)_y}$ and $\boldsymbol{\Phi_l(I^{gt})_y}$. The mathematical expression for this relationship is presented in Eq. (5).

$$\boldsymbol{L_{content} = \frac{1}{2}\sum_{y \in \Omega}\left((\Phi_{L-l}(I))_y - \left(\Phi_l(I^{gt})\right)_y\right)^2} \quad (5)$$

Eq. (6) describes the relationship between $\boldsymbol{(\Phi_{L-l+1}(I))_y}$ and $\boldsymbol{\Phi_{l-1}(I^{gt})_y}$.

$$\boldsymbol{L_{content} = \frac{1}{2}\sum_{y \in \Omega}\left((\Phi_{L-l+1}(I))_y - \left(\Phi_{l-1}(I^{gt})\right)_y\right)^2} \quad (6)$$

Given that $\boldsymbol{\Phi_l(I)_x \approx \Phi_l(I^{gt})_x}$ for every $\boldsymbol{x \in \bar{\Omega}}$, and $\boldsymbol{\Phi_{l-1}(I)_x \approx \left(\Phi_{l-1}(I^{gt})\right)_x}$, the content loss is confined to $\boldsymbol{y \in \Omega}$ to ensure $\boldsymbol{(\Phi_{L-l}(I))_y \approx \left(\Phi_l(I^{gt})\right)_y}$, and $\boldsymbol{\Phi_{L-l+1}(I)_y \approx \left(\Phi_{l-1}(I^{gt})\right)_y}$. The fusion of $\boldsymbol{\Phi_l(I)}$ and $\boldsymbol{\Phi_{L-l}(I)}$ captures nearly all the information in $\boldsymbol{\Phi_l(I^{gt})}$,

while the combination of $\boldsymbol{\Phi_{l-1}(I)}$ and $\boldsymbol{\Phi_{L-l+1}(I)}$ approximates all the information in $\boldsymbol{\Phi_{l-1}(I^{gt})}$ [21].

Furthermore, this study employs the foundational U-Net network model. Subsequently, an enhancement is introduced wherein conventional convolutional layers are replaced with partial ones. Specifically, the sliding window of these partial convolutional layers exclusively executes convolution operations within the pertinent image area. This innovation empowers the prediction of architectural constructions that incorporate asymmetrical gaps and simplifies the creation of renovation results featuring meaningful semantic characteristics.

3 Design of digital building file data manager system(blockchain-based cloud storage system)

The preceding section introduces the U-Net network and its enhancements. U-Net, a well-established convolutional neural network employed for image segmentation, incorporates compression and expansion paths. However, challenges related to efficiency and the balance between localization and contextual information are encountered. To address these challenges, an enhanced U-Net structure featuring specialized transfer connection layers is introduced in this study to elevate the precision and quality of image restoration. The subsequent section provides an in-depth exploration of the detailed design of the digital architectural document data management system, which adopts a blockchain-based cloud storage system. A comprehensive overview of the blockchain's structure and principles is provided, elucidating how it ensures data security and immutability. Simultaneously, the structure and operational principles of the blockchain-based cloud storage system are introduced.

The blockchain serves as a distributed transaction data structure [22, 23], facilitating the documentation of transaction details among nodes and ensuring data integrity through algorithms. It establishes a decentralized shared ledger, facilitating the verification and traceability of ledger information. The blockchain validates and stores data within a chain structure by employing cryptographic algorithms. Consensus mechanisms are employed to generate and update data across a distributed network, frequently utilizing smart contracts for data manipulation and coding. Consequently, blockchain technology achieves decentralized management and ensures equal rights across all blockchain nodes, mitigating delays and storage risks inherent in centralized systems [24-27]. The fundamental data unit in the blockchain is referred to as a block, forming a unidirectional chain structure by linking parent block hashes sequentially. Each block comprises a header and a body. The header contains metadata, and data block authenticity is ensured through hash verification. The interconnection between blocks is established by the parent block's hash, with the block's hash generated through the SHA-256 function. Notably, modifications in one block impact subsequent blocks, including the header and Merkle root hashes, rendering any attempt to introduce false data into the blockchain by maliciously altering node information unfeasible [28, 29]. The feasibility of achieving such an accomplishment is hindered by the substantial size of network manipulation necessary, which necessitates the utilization of around 50% of the network's processing power. The block's structural composition within the blockchain framework is illustrated in Figure 3.

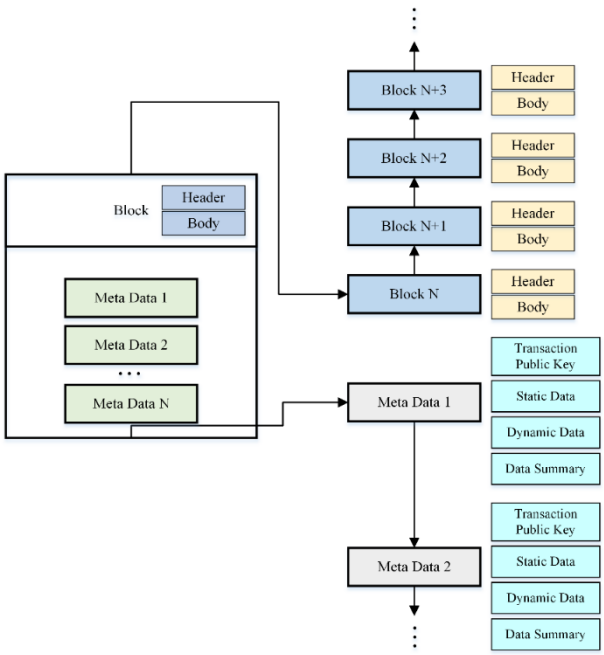

Figure 3: Block structure of the blockchain

Figure 3 illustrates that the block structure within the blockchain comprises two primary components: the header and the body. The header encompasses metadata elements, including timestamps and the hash value of the preceding block, establishing crucial links between consecutive blocks. The block's integrity is upheld through hash verification, wherein any tampering with block data alters its hash value, compromising the blockchain's overall coherence. Meanwhile, the body section accommodates actual transaction data or other relevant information, with the specific content contingent upon the blockchain's intended application. The blocks within the blockchain are systematically interconnected in chronological order, giving rise to a chain-like structure. This interconnection ensures both the immutability and sequential arrangement of data, endowing the blockchain with elevated security and traceability.

Based on blockchain technology, the envisaged cloud storage system encompasses four distinct roles: museums, governments, ordinary users, and professional restorers. Furthermore, this ecosystem involves the participation of blockchain nodes, as depicted in Figure 4.

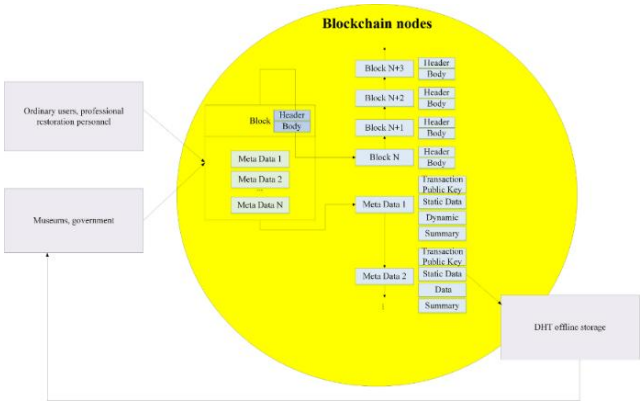

Figure 4: The cloud storage system based on blockchain

4 Enhancing texture and contextual feature restoration

This section discusses the implementation of the approach to enhance the restoration of texture and contextual features in cultural heritage images.Data Preprocessing: Before initiation, essential preprocessing steps are performed on the input data. These include image resizing, color space conversion, and noise removal, ensuring that the input data is optimized for network processing.Neural Network Architecture: The chosen neural network model is the optimized U-Net architecture, featuring an encoder for extracting image features and a decoder for generating the restored image.Loss Function: The Content Loss function is selected, employing mean squared error loss to gauge the dissimilarity between the generated and original images. The objective is to minimize the Content Loss, compelling the network to learn the feature representation of the original image, encompassing texture, and contextual features.Dual-Transfer Operation: This entails sharing feature representations within the backbone structure through parallel connections and nearest neighbor search. Nearest neighbor search aligns features from known regions with features from missing regions, enhancing the accuracy of estimates.Image Generation and Output: Following the completion of network training, the image to be restored is input into the network to generate the restored image.Drawing upon the content discussed in the preceding sections, this study integrates the improved U-Net neural network with a blockchain-based data storage and sharing system to realize the comprehensive algorithmic flow for image restoration, as depicted in Figure 5.

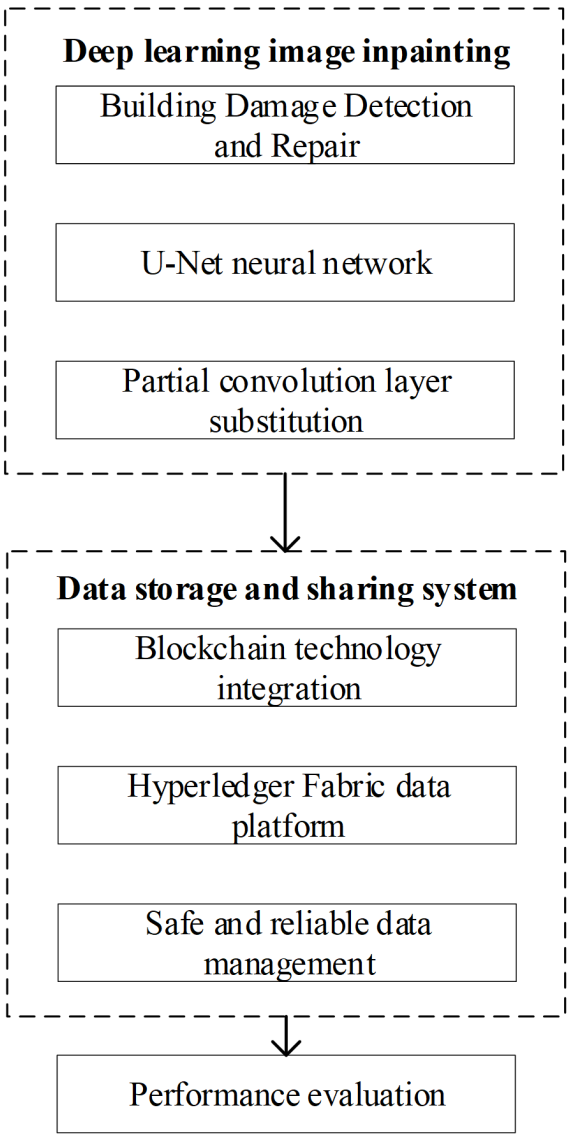

Figure 5: Algorithmic flow

5 Experimental data and evaluation index

This section outlines the experimental configuration aimed at assessing the performance of the previously proposed model. The experimental design encompasses the data source, data preprocessing, experimental environment, and the establishment of performance evaluation metrics.

The experimental data employed in this study originates from the WHU building dataset created by the Ji Shunping Team at Wuhan University. This dataset constitutes a collection of internationally standardized datasets designed for building extraction (available at http://study.rsgis.whu.edu.cn/pages/download/). The

building dataset consists of aerial images with a spatial resolution of 0.3 meters. These photos include three visible bands—Red, Green, and Blue (RGB). The dataset covers an expansive area of 400 square kilometers and includes a diverse array of 180,000 buildings, exhibiting variations in size, color, and function. Figure 6 depicts sample images extracted from the dataset employed in these experiments.

Figure 6: Pictures of dataset samples

The utilization of these metrics aids in evaluating the model's robustness to noise, ensuring accuracy and reliability in the extraction and modeling of building features. The specific details of the evaluation metrics are outlined below:

The Structural Similarity Index Measure (SSIM) is employed to quantify the similarity between the actual and generated images in deep learning-based image restoration. This approach aligns with human perceptual evaluation standards for assessing image quality by evaluating brightness, contrast, and structural information. The technique of contrast estimation involves the use of the standard deviation (SD), while brightness estimation relies on the mean value. SSIM is quantified using covariance. To compute the similarity between two input images, denoted as x and y, the SSIM is formulated as shown in Eq. (7).

$$SSIM(x,y) = l(x,y) \cdot c(x,y) \cdot s(x,y) \tag{7}$$

In Eq. (7), l(x, y) represents the luminance similarity factor, assessing the likeness of image luminance to ensure the robustness of the results against luminance variations. Similarly, c(x, y) denotes the contrast similarity factor employed to gauge the similarity of image contrast, ensuring resilience to contrast variations. Finally, s(x, y) signifies the structural similarity factor utilized to measure the similarity of image structure, ensuring the robustness of the results to structural variations. The methodologies for these three factors are elucidated as follows:

$$l(x,y) = \frac{2u_x u_y + c_1}{u_x^2 + u_y^2 + c_1} \tag{8}$$

$$c(x,y) = \frac{2\sigma_x \sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \tag{9}$$

$$s(x,y) = \frac{\sigma_{xy} + c_3}{\sigma_x \sigma_y + c_3} \tag{10}$$

where $\boldsymbol{u_x}$ and $\boldsymbol{u_y}$ represent the mean of images x and y, and $\boldsymbol{\sigma_x}$ and $\boldsymbol{\sigma_y}$ represent their SDs; $\boldsymbol{\sigma_x^2}$, $\boldsymbol{\sigma_y^2}$, and $\boldsymbol{\sigma_{xy}}$ stand for the variance and covariance of images x and y, respectively; Additionally, the constants $\boldsymbol{c_1}$, $\boldsymbol{c_2}$, and $\boldsymbol{c_3}$ are included to prevent division by zero. Usually, $\boldsymbol{c_1}$, $\boldsymbol{c_2}$, and $\boldsymbol{c_3}$ are defined as $\boldsymbol{(k_1 l)^2}$, $\boldsymbol{(k_2 l)^2}$, and $\boldsymbol{c_2/2}$, respectively. The interrelation between these quantities is illustrated in Eq. (11).

$$\mathrm{SSIM}(x,y) = \frac{(2u_x u_y + c_1)(2\sigma_{xy} + c_2)}{(u_x^2 + u_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{11}$$

The process involves extracting N×N windows from the image and calculating their respective averages. The global SSIM value is determined by utilizing the resulting average [30].

The PSNR is utilized as a quantitative measure for evaluating photographs objectively. It calculates the disparity at the pixel level between the generated images and the ground truth images [31]. The PSNR is determined by calculating the MSE between the current image x and the reference image y.

$$\boldsymbol{MSE = \frac{1}{h \times w} \sum_{i=1}^{h} \sum_{j=1}^{w} [x(i,j) - y(i,j)]^2} \tag{12}$$

In Eq. (12), *h* and w stand for the image's height and width, respectively; the PSNR is defined in Eq. (13) [32].

$$PSNR = 10 \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right) \tag{13}$$

Additionally, this study employs the Learned Perceptual Image Patch Similarity (LPIPS) metric to quantify image similarity. LPIPS is a perceptual image patch similarity metric designed to evaluate perceptual differences between images. Specifically, LPIPS utilizes a neural network to learn the perceptual features of image patches, calculating the distance between them to quantify the perceptual similarity of images, expressed as follows:

$$LPIPS(x,y) = \sum_{i=1}^{n} \omega_i \cdot \varphi_i(x) \cdot d_i\big(\varphi_i(x), \varphi_i(y)\big) \tag{14}$$

In Eq. (14), x and y represent two images, $\boldsymbol{\varphi_i(x)}$ denotes the representation of the i-th feature layer, $\boldsymbol{d_i}$ signifies the distance measurement function on that feature layer, and $\boldsymbol{\omega_i}$ denotes the weight for the corresponding feature layer.

The suggested prediction model is validated through simulation tests on a specialized platform. After gathering, preparing historical data, and creating a validation set, the proposed model and the training models are applied to both the training and validation sets to calculate losses. The training process concludes when the model's loss on

the validation set no longer shows a decline. The subsequent simulations are conducted using a Windows 10 operating system, featuring a 3.0GHz CPU, 4GB of RAM, and a central processing unit core-i7-7700HQ.

6 Experimental results and discussion

This chapter offers a comprehensive presentation and analysis of the performance results obtained from the enhanced algorithmic model. The primary objective is to validate the practicality and reliability of the model through a meticulous examination of its outcomes.

6.1 Test and analysis of the proposed improved algorithm

The primary objective of the developed algorithm is to restore building images. Figure 7 depicts the performance of the algorithm in restoring building pictures within the dataset of Wuhan University buildings.

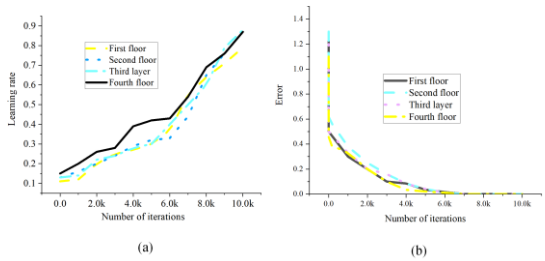

Figure 7: Verification results of the improved model ((a) depicts the verification results of the model improvement layer's learning rate, and (b) showcases the error analysis verification of the same layer.)

The observations from Figure 7 indicate notable advantages of the model. Firstly, the learning rate of the model shows a consistent increase throughout the iterations. Notably, after around 10,000 iterations, the learning rate settles at approximately 0.9, highlighting the model's robust learning capacity. Moreover, in the context of error verification, the model's error progressively decreases as the number of iterations increases. When the number of iterations reaches 10,000, the model's error exhibits significant convergence towards zero.

Figure 8 illustrates the correlation between the iteration count of the enhanced model and the PSNR value.

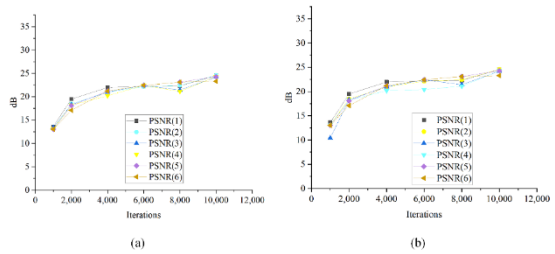

Figure 8: Iteration count vs. PSNR value (a. First Test; b. Second Test)

In Figure 8, the PSNR values corresponding to algorithm iterations of 1,000 and 10,000 are 24.21 and 26.16, respectively. It is observed that the quality of image generation improves with increased iterations. Moreover, beyond 10,000 iterations, the generated images approach benchmark quality.

6.2 Performance validation of the model

(1) Evaluation of PSNR, SSIM, and LPIPS values of the model

The quantitative evaluation involves assessing and comparing various image restoration algorithms by quantifying PSNR, SSIM, and Learned Perceptual Image Patch Similarity (LPIPS) metrics. Fifty validation images are randomly selected for this purpose, and the average values of SSIM, LPIPS, and PSNR are computed. The PSNR, SSIM, and LPIPS values for different methods are presented in Table 1. In this context, LPIPS distinguishes itself from conventional image quality assessment metrics like PSNR and SSIM. LPIPS places heightened importance on scrutinizing perceptual details within images, encompassing considerations such as structure, color distribution, and texture. In the realm of image restoration, LPIPS proves invaluable for appraising the perceptual similarity between generated or restored images and their original counterparts, delivering a quality assessment more attuned to the intricacies of the human visual system. Consequently, LPIPS enjoys broad applicability in the domains of deep learning and computer vision, particularly in tasks where emphasis is placed on perceptual quality, allowing for a more precise evaluation of the efficacy of image generation algorithms.

Table 1: Quantitative evaluation

Method (Model iteration)	PSNR	SSIM	LPIPS	Total Variation
Patch Match [33]	20.57	0.31	0.891	0.125
Context Encoder [34]	22.71	0.41	0.861	0.114
Globally and Locally [35]	21.21	0.35	0.851	0.121
Shift Net [36]	21.78	0.28	0.794	0.135
U-net [37]	21.32	0.34	0.723	0.132
Res Net [38]	24.82	0.52	0.798	0.112
Generative Adversarial Network [39]	25.94	0.55	0.783	0.109
The proposed algorithm	26.16	0.57	0.825	0.102

Table 1 illustrates that the proposed algorithm exhibits the highest PSNR value compared to the other comparative algorithms. This observation implies that the proposed algorithm generates the least error between corresponding pixels within the restored image compared to the original image. Moreover, the proposed algorithm minimizes texture distortion in the missing region, resulting in an image that closely resembles the original. Notably, the SSIM value for the proposed algorithm also outperforms other algorithms, indicating superior

restoration of brightness, contrast, and semantic structure in the image.

Moreover, as illustrated in Table 1, the algorithm presented here demonstrates comparatively diminished performance on the LPIPS evaluation metric in comparison to Patch Match, Context Encoder, Globally and Locally, and Shift Net. This discrepancy can be attributed to LPIPS functioning as a perceptual similarity measure, mirroring the subjective way in which humans perceive image similarity. Its scores are influenced by factors such as image structure, color distribution, and texture. Notwithstanding the lower scores on LPIPS in this study, the algorithm exhibits commendable performance on traditional evaluation metrics like PSNR and SSIM, indicating an overall satisfactory outcome.

(2) Feature extraction performance of the model

The algorithm proposed here demonstrates its proficiency in feature extraction when applied to the two target images as illustrated in Figure 6. The resulting output is presented in Figure 9.

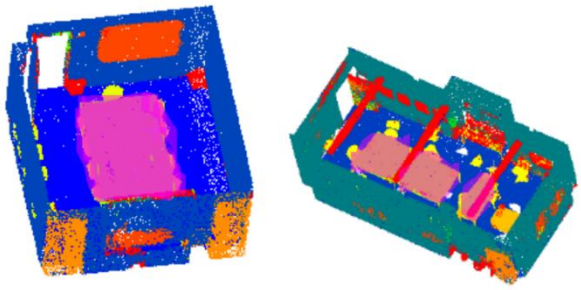

Figure 9: Output results of the model

Figure 9 illustrates the exceptional performance of the proposed algorithm in indoor scene recognition, with a specific focus on detailing the model's outputs in images containing furniture. Firstly, the model exhibits a high degree of pixel-level accuracy in recognizing subtle differences in chair designs. Secondly, the model excels in precisely identifying and differentiating images on walls, even in scenarios with numerous visual elements. This capability is pivotal for effectively handling the diverse range of visual elements present in indoor scenes. Furthermore, the model performs remarkably well in the broader context of semantic segmentation and excels in the precise instance segmentation of complex sofa point clouds. The ability to intricately understand spatial relationships enables the model to accurately delineate various instances of sofas within semantic scenes. These detailed and nuanced visual results underscore the model's exceptional performance and enhance the understanding of its capabilities in addressing diverse indoor scenes, offering a more comprehensive perspective on its overall efficacy in indoor scene recognition tasks.

6.3 Comparison of storage system performance

The performance of the system proposed here is compared to that of an existing cloud storage system. Two systems are considered in this analysis. The system outlined in Reference [40] entails a full-text retrieval mechanism for safeguarding the privacy of encrypted data within cloud storage. Meanwhile, the approach presented in Reference [41] outlines a cost-effective cloud storage scheme involving mixed strategies. The comparative performance analysis between the proposed system and the existing counterparts is depicted in Figure 10.

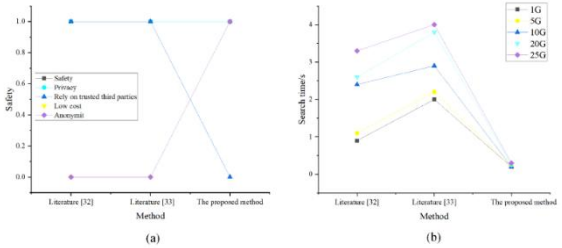

Figure 10: Performance comparison results between the proposed system and the existing cloud storage system (a: Basic performance; b: Data storage performance)

Figure 10(a) depicts the proposed blockchain-based data storage system designed to preserve historical structures. This system provides secure, private, cost-efficient, and anonymous storage and retrieval services, eliminating the necessity for a trusted middleman. Notably, its performance outpaces that of the other two cloud storage systems. Moving to Figure 10(b), it becomes apparent that the search time for data across the three storage systems increases as the data storage capacity increases. Within varying levels of stored data, the proposed data storage system exhibits shorter search times compared to the other two systems. This advantageous outcome is attributed to the system's utilization of blockchain as the underlying framework for offline data storage, authorization, access, and search token generation. This strategic employment ensures data security while concurrently curtailing search times, thereby underscoring the effectiveness of the proposed system.

6.4 Data storage cost and efficiency analysis

Figures 11(a) and 11(b) effectively depict the upward trajectory of both observed and projected expenditure values during the conservation process of historical building restoration data in the XX province. The provided figures accurately depict the changes observed before and after the introduction of the proposed system. Figure 11(a) demonstrates a noticeable increase in expenditure for the preservation project of historic building repair data in the XX province over the years 2016, 2017, 2018, 2019, and 2020. The recorded values for the years 2015, 2016, 2017,

2018, and 2019 were 549,400, 647,400, 781,100, 1,007,800, and 1,124,800 CNY, respectively. Figure 11(b) presents a predictive study that anticipates the incremental spending values for linked projects in the XX province from 2016 to 2020. The projected values for each year are as follows: 249,400 CNY for 2016, 347,400 CNY for 2017, 481,100 CNY for 2018, 607,800 CNY for 2019, and 724,800 CNY for 2020. These values represent the increase in expenditure compared to the preceding year. This data illustrates a notable improvement of 25%, 24%, 26%, 25%, and 25% in efficiency compared to the original approach. The methods for building restoration presented in this study are expected to result in an annual cost reduction of 1 million CNY.

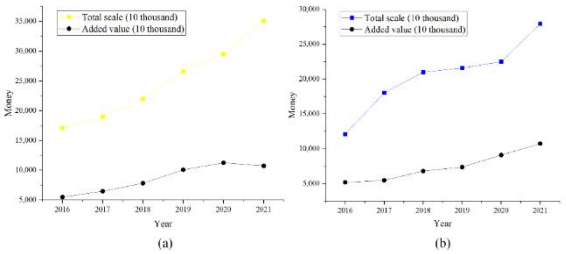

Figure 11: Statistics of restoration costs of historic buildings in the xx province (a. actual value; b. predicted value.

7 Conclusion

The preservation of ancient architectural designs holds considerable national significance, leading to the adoption of contemporary computer technologies for digitization and restoration. This study presents a novel picture restoration model for ancient buildings utilizing the U-Net architecture. The methodology employs the U-Net architecture as the fundamental network model, later enhancing it by substituting all convolutional layers with partial equivalents. Moreover, the partial convolutional layer employs a sliding window mechanism to perform convolutions only inside relevant portions of the picture. This allows for the accurate prediction of building structures containing irregular voids and the development of restoration results with semantic significance. The present study proposes the development of a platform utilizing blockchain technology to effectively store and manage historical building maintenance data. This platform aims to fulfill the specific requirements of a data storage and sharing system for the aforementioned purpose. The experimental results illustrate the effectiveness of the method in successfully reconstructing photographs of historical structures, minimizing MSE, and exhibiting robust practicality and possible use in real-world situations. The neural network evaluations demonstrate that the algorithm under consideration achieves SSIM, LPIPS, and PSNR scores of 0.57, 0.725, and 26.16, respectively. Future investigations aim to replace the existing backbone network with a residual network and enhance the model's up-sampling process to reduce the impact of injected noise. These endeavors are expected to improve the model's overall efficacy and performance.

Future research will further introduce Transformer backbone structures (such as Swin-UNet and ConvNeXt-UNet) to enhance the global modeling capabilities of feature expression. Simultaneously, we will explore diffusion-based image restoration methods to improve complex textures and semantic consistency. Furthermore, we will study blockchain-based model provenance and tracking mechanisms to ensure the verifiability and credibility of the restoration process and data management.

Funding

Research on the Integration and Creation of Intangible Cultural Heritage in the Inheritance of Qinling Culture.

References

[1] Paya-Zaforteza, I., & Hospitaler, A. (2023). Fire in heritage and historic buildings, a major challenge for the 21st century. Developments in the Built Environment, 13, 100102. doi:10.1016/j.dibe.2022.100102

[2] Liu G, Reda F A, Shih K J, Wang T C, Tao A, Catanzaro B. Image inpainting for irregular holes using partial convolutions. InProceedings of the European conference on computer vision (ECCV), 2018; 85-100. doi: 10.1007/978-3-030-01252-6_6

[3] Kong, X., & Hucks, R. G. (2023). Preserving our heritage: A photogrammetry-based digital twin framework for monitoring deteriorations of historic structures. Automation in Construction, 152, 104928. doi: 10.1016/j.autcon.2023.104928

[4] Ji Z, Zhu H, Xu Y, Wu M. Optimization Design of Permanent Magnet Assisted Single Winding Bearingless Synchronous Reluctance Motor. IEEE Transactions on Applied Superconductivity, 2020; 30(4):1-1. doi: 10.1109/TASC.2020.2971937

[5] Ascione F, N Bianco, Mauro G M, Vanoli G P. A new comprehensive framework for the multi-objective optimization of building energy design: Harlequin. Applied Energy, 2019; 241(1):331-361. doi: 10.1016/j.apenergy.2019.03.028

[6] Fa A, Nb A, Gmm B. Building envelope design: multi-objective optimization to minimize energy consumption, global cost and thermal discomfort. Application to different Italian climatic zones. Energy, 2019; 174:359-374.doi: 10.1016/j.energy.2019.02.182

[7] Fadli F, AlSaeed M. Digitizing vanishing architectural heritage; The design and development of Qatar historic buildings information modeling [Q-HBIM] platform. Sustainability, 2019; 11(9):2501. doi: 10.3390/su11092501

[8] Müller, S., Sentosa, W., Kolb, D., Müller, M., & Kranzlmüller, D. (2025). D-LaMa: Depth Inpainting of Perspective-Occluded Environments. Proceedings Copyright, 255, 266.doi: 10.5220/0013090000003912

[9] Masciotta M G, Morais M J, Ramos L F, Oliveira D V, González-Aguilera D. A digital-based integrated methodology for the preventive conservation of cultural heritage: the experience of HeritageCare project. International Journal of Architectural Heritage, 2021; 15(6):844-863. doi: 10.1080/15583058.2019.1668985

[10] Li H. Intelligent restoration of ancient murals based on discrete differential algorithm. Journal of computational methods in sciences and engineering, 2021; 1(3):21.doi: 10.3233/JCM-215195

[11] Wang T, Zhao L. Virtual Reality-Based Digital Restoration Methods and Applications for Ancient Buildings. Journal of Mathematics, 2022; 1:2022.doi: 10.1155/2022/2305463

[12] Siountri, K., Skondras, E., & Vergados, D. D. (2020). Developing smart buildings using blockchain, internet of things, and building information modeling. International Journal of Interdisciplinary Telecommunications and Networking (IJITN), 12(3), 1-15.doi: 10.4018/IJITN.2020070101

[13] Nawari N O, Ravindran S. Blockchain and Building Information Modeling (BIM): Review and Applications in Post-Disaster Recovery. Buildings, 2019; 9(6):149.doi: 10.3390/buildings9060149

[14] Dey E K, Awrangjeb M, Tarsha Kurdi F, Stantic B. Machine learning-based segmentation of aerial LiDAR point cloud data on building roof. European Journal of Remote Sensing, 2023; 56(1): 2210745.doi: 10.1080/22797254.2023.2210745

[15] Gharineiat Z, Tarsha Kurdi F, Campbell G. Review of automatic processing of topography and surface feature identification LiDAR data using machine learning techniques. Remote Sensing, 2022; 14(19): 4685.doi: 10.3390/rs14194685

[16] Mirzaei K, Arashpour M, Asadi E, Masoumi H, Bai Y, Behnood, A. 3D point cloud data processing with machine learning for construction and infrastructure applications: A comprehensive review. Advanced Engineering Informatics, 2022; 51: 101501.doi: 10.1016/j.aei.2021.101501

[17] Khan S D, Alarabi L, Basalamah S. An encoder–decoder deep learning framework for building footprints extraction from aerial imagery. Arabian Journal for Science and Engineering, 2023; 48(2): 1273-1284.doi: 10.1007/s13369-022-06768-8

[18] Vacchio E D, Bifulco F. Blockchain in Cultural Heritage: Insights from Literature Review. Sustainability, 2022; 14(4): 2324.doi: 10.3390/su14042324

[19] Trček D. Cultural heritage preservation by using blockchain technologies. Heritage Science, 2022; 10(1): 6.doi: 10.1186/s40494-021-00643-9

[20] Guo C, Lu J, Tian Z, Guo W, Darvishan A. Optimization of critical parameters of PEM fuel cell using TLBO-DE based on Elman neural network. Energy Conversion and Management, 2019; 183(5):149-158.doi: 10.1016/j.enconman.2018.12.088

[21] Mclane Y, Pable J. Architectural Design Characteristics, Uses, and Perceptions of Community Spaces in Permanent Supportive Housing. Journal of Interior Design, 2020; 45(4):121. doi:10.1111/joid.12165

[22] Laaroussi Y, Bahrar M, El Mankibi M. Occupant presence and behavior: A major issue for building energy performance simulation and assessment. Sustainable Cities and Society, 2020; 63:102420. doi: 10.1016/j.scs.2020.102420

[23] Hu S, Yan D, Qian M. Using bottom-up model to analyze cooling energy consumption in China's urban residential building. Energ Buildings, 2019; 202:109352.doi: 10.1016/j.enbuild.2019.109352

[24] Deng Z, Chen Q. Simulating the impact of occupant behavior on energy use of HVAC systems by implementing a behavioral artificial neural network model. Energ Buildings, 2019; 198:216-227.doi: 10.1016/j.enbuild.2019.06.015

[25] Day J K, McIlvennie C, Brackley C, Tarantini M, Tarantini C, Hahn J, O'Brien W, Rajus V S, Simone M D, Kjærgaard M B, Pritoni M, Schlüter A, Peng Y, Schweiker M, Fajilla G, Becchio C, Fabi V, Spigliantini G, Derbas G, Pisello A L. A review of select human-building interfaces and their relationship to human behavior, energy use and occupant comfort. Building and Environment, 2020; 178:106920.doi: 10.1016/j.buildenv.2020.106920

[26] Serrano-Jiménez A, Lizana J, Molina-Huelva M. Decision-support method for profitable residential energy retrofitting based on energy-related occupant behaviour. Journal of Cleaner Production, 2019; 222:622-632.doi: 10.1016/j.jclepro.2019.03.089

[27] Sarkar A, Bardhan R. Optimal interior design for naturally ventilated low-income housing: a design-route for environmental quality and cooling energy saving. Advances in Building Energy Research, 2020; 14(4):494-526.doi: 10.1080/17512549.2019.1626764

[28] 28. Del, M. S. T. T., & Tabrizi, S. K. (2020). A methodological assessment of the importance of physical values in architectural conservation using Shannon entropy method. Journal of Cultural Heritage, 44, 135-151.doi: 10.1016/j.culher.2019.12.012

[29] Aloisio A, Antonacci E, Fragiacomo M, Alaggio R. The recorded seismic response of the Santa Maria di Collemaggio basilica to low-intensity earthquakes. International Journal of Architectural Heritage, 2021; 15(1):229-247. doi: 10.1080/15583058.2020.1802533

[30] Milošević M R, Milošević D M, Stanojević A D, Martinho G. Fuzzy and interval AHP approaches in sustainable management for the architectural heritage in smart citiesc. Resources, Conservation & Recycling, 2021; 9(4):304. doi: 10.3390/math9040304

[31] Wang T, Zhao L. Virtual Reality-Based Digital Restoration Methods and Applications for Ancient Buildings. Journal of Mathematics, 2022; 2:1-10.doi: 10.1155/2022/2305463
[32] Spiridonov A, Umniakova N. Problems of Restoration of Historical Fenestration for Providing the Normalized Parameters of the Microclimate of Premises in Ancient Buildings. IOP Conference Series: Materials Science and Engineering, 2020; 753(2):022064.doi:10.1088/1757-899X/753/2/022064
[33] Yu D, Ji S, Liu J. Automatic 3D building reconstruction from multi-view aerial images with deep learning. ISPRS Journal of Photogrammetry and Remote Sensing, 2021; 171: 155-170. doi: 10.1016/j.isprsjprs.2020.11.011
[34] Li J, Li B, Lu Y. Deep contextual video compression. Advances in Neural Information Processing Systems, 2021; 34: 18114-18125. doi: 10.7551/mitpress/7503.003.0116
[35] Lundberg S M, Erion G, Chen H. From local explanations to global understanding with explainable AI for trees. Nature machine intelligence, 2020; 2(1): 56-67.doi: 10.1038/s42256-019-0138-9
[36] Ang Y Q, Berzolla Z M, Reinhart C F. From concept to application: A review of use cases in urban building energy modeling. Applied Energy, 2020; 279: 115738.doi: 10.1016/j.apenergy.2020.115738
[37] Wagner F H, Dalagnol R, Tarabalka Y. U-net-id, an instance segmentation model for building extraction from satellite images—case study in the joanópolis city, brazil. Remote Sensing, 2020; 12(10): 1544.doi: 10.3390/rs12101544
[38] Mohammed H, Tannouche A, Ounejjar Y. Weed Detection in Pea Cultivation with the Faster RCNN ResNet 50 Convolutional Neural Network. Revue d'Intelligence Artificielle, 2022; 36(1):3-4.doi: 10.18280/ria.360102
[39] Abdollahi A, Pradhan B, Gite S. Building footprint extraction from high resolution aerial images using generative adversarial network (GAN) architecture. IEEE Access, 2020; 8: 209517-209527.doi: 10.1109/ACCESS.2020.3038225
[40] Hmood K, Jumaily H, Melnik V. Urban architectural heritage and sustainable tourism. WIT Transactions on Ecology and the Environment, 2018; 227:209-220. doi: 10.2495/ST180201
[41] Noardo F. Architectural heritage semantic 3D documentation in multi-scale standard maps. Journal of Cultural Heritage, 2018; 32:156-165.doi: 10.1016/j.culher.2018.02.009