AI-based Digital Documentation as Sensors of Heritage Morphological Value: Recording Plane Forms of Historical Rural Settlements Hierarchically

Nanshuo Mu[1,2], Yijun Xiao[1,2], Yuxin Yin[3,4], Houwei Fu[5], Sinan Yuan[1,2]

[1] School of Architecture, Tianjin University, Tianjin 300072, China. - (ericmu_14, xiaoyijun, yuansinan)@tju.edu.cn
[2] Key Laboratory of Information Technology for Architectural Cultural Heritage (Tianjin University), Ministry of Culture and Tourism, Tianjin 300072, China.
[3] School of Architecture and Design, Harbin Institute of Technology, Harbin 150001, China. - 24B334010@stu.hit.edu.cn
[4] Key Laboratory of Cold Region Urban and Rural Human Settlement Environment Science and Technology (Harbin Institute of Technology), Ministry of Industry and Information Technology, Harbin 150001, China.
[5] The Bartlett School of Planning, University College London, London WC1H 9BT, United Kingdom. - ucbvhfu@ucl.ac.uk

Keywords: Cultural Heritage Places, Rural Ivestigation Workflow, Settlement Plane Form, Computer Vision, Building Regularisa-tion, Geospatial Registration.

Abstract

Historical Rural Settlements (HRS) are recognised by *ICOMOS-IFLA* as tangible cultural heritage, where plane forms preserve multi-layered morphological value accrued over successive eras. These plane forms not only reflect past settlement planning and community interactions but also reveal contrasts between vestigial historical cores and more recent modern developments. Although digital documentation has become a mainstream conservation approach, it often overlooks the interpretive dimension that connects raw geometric data with cultural significance. To address this gap and manage the complexity of large-scale image acquisition, annotation, and analysis, this study treats image-based artificial intelligence (AI) as a cultural sensor for HRS morphology. A re-peatable, open-source workflow that integrates high-resolution remote sensing, expert-driven annotation, hierarchical segmentation (*Mask R-CNN*), and building(polygon) regularisation was proposed. In the Taihang Baxing (THBX) region of China, 778 HRS samples were processed through a two-stage model with several supportive techniques: the first differentiates historical and modern regions, and the second extracts building footprints within each region. Validation on THBX demonstrates reliable differentiation of hierarchical plane forms and efficient generation of a vectorised digital documentation dataset. By embedding geospatial registra-tion and building regularisation tools, the workflow ensures downstream usability for use cases like geospatial statistics, quantitative morphology and evolutionary mechanisms, providing insights for the rural planning consequently. The proposed workflow emphas-ises heritage morphological value and digital documentation as guiding principles, demonstrating that AI as technical agents can perceptually interpret cultural significance. It can serve as a tool for extracting and highlighting the historical value of HRS in the formulation of conservation strategies and development plans, contributing to sustainable and inclusive rural heritage management.

1. Introduction

Rural settlements are clusters of buildings in rural areas that provide multifunctional spaces for rural living and play an important role in managing the local population and shaping regional and cultural identity (Gong et al., 2022). The formation of rural settlements is the result of long-term interaction between human society and the natural environment (Shi et al., 2022). This interaction has evolved and accumulated over time, endowing rural settlements with significant cultural significance and historical value (Tarrafa Pereira da Silva and Pereira Roders, 2012). According to the principles of *ICOMOS-IFLA*, rural settlements, as a physical aspect of rural areas, should be regarded as tangible cultural heritage. Different regions of rural settlements are linked to different historical periods, so when they are discussed as heritage, their types and degrees vary (ICOMOS-IFLA, 2017). Historical rural settlements (HRS) are those that were formed in earlier periods and retain traces of the built environment and social activities of different periods, representing a type of heritage with historical significance and value (Yuan et al., 2024). Unlike typical tangible built heritage, such as archaeological sites and historic buildings, HRS is a type of living, dynamic heritage that is constantly operating and involves multiple stakeholders (Du and Shi, 2019). Therefore, rural settlements are constantly subject to internal or external interventions. Currently, industrialisation and urbanisation threaten rural settlements by both neglect and overexploitation, making rural heritage increasingly vulnerable to abandonment, loss, or drastic transformation (Bai et al., 2023). Faced with such challenges, the historical value and cultural identity that HRS has accumulated over many years appear even more fragile and require conservation measures. With the emergence of digital twin perspectives and related technologies, the digital documentation of tangible cultural heritage has become an novelly effective means of conservation (Argyridou et al., 2023). Although there is still debate about relative authenticity, the advantages of digital documentation over invasive physical methods are that it does not harm the essence of the heritage, is highly efficient, and can be combined with more interpretive information (Galeazzi, 2018). There are already research and engineering initiatives that have applied concepts and techniques from the digital documentation of small-scale tangible heritage to the conservation of HRS (Nancarrow et al., 2021). Specifically, the popularity of open remote sensing image services and the improvement of UAV orthophoto and street view image capture provide a solid foundation for digital documentation.

Plane form is an attribute of HRS that suit for digital documentation. They are 2-D illustrations of the spatial relationships between buildings within HRS from a bird's-eye-view (Zhu and Liu, 2023), reflecting the macro-level orderly integration of settlement construction, mostly from an individual micro-level perspective. Pre-existing plane forms in HRS are stable and resistant to fundamental alters, so it also has the features of "palimpsest". The plane forms corresponding to areas developed in different periods are overlapped like layers, which is also a part of the built environmental value of HRS as cultural heritage. According to $The Burra Charter$, the first step in investigating and managing places of cultural significance (cultural heritage places) is to understand their significance (Australia ICOMOS, 2013). By studying the plane forms associated with different periods in HRS, the wisdom of settlement planning and the social interaction patterns of rural communities in different eras can be discovered. Thus, identifying plane forms related to early stages of development in historical rural settlements through human knowledge and experience helps judge and reveal their cumulative cultural significance and provides a perceptible basis for their historical and morphological value (Ding, 2021). Recently, with the development of artificial intelligence (AI), image analysis based on deep learning has contributed to the trend of HRS digital archiving and has already been linked to plane forms. Deep learning-based image analysis has not only enhanced the automation of image interpretation but also accelerated the preliminary screening of spatial features in rural settlements, significantly reducing the cost and time required for subsequent manual analysis and field surveys (Aamir et al., 2023). Reviewing relevant research, deep learning has already performed well in such field, with generating datasets of rural plane forms with broad adoption across diverse research and intervention contexts (Deng et al., 2024). However, different AI models are mostly used to capture the overall plane forms of villages (Zhang et al., 2021) or to classify the physical characteristics of buildings (Wang et al., 2023), rather than to detect the layered features of the internal plane forms of HRS from a cultural perspective. Limited studies have demonstrated the potential of AI to detect plane forms in historical regions at the box-level, but has not explained its segmentation capabilities at the mask-level (Tao et al., 2023). Therefore, intelligent hierarchical recording of HRS region-level plane forms based on AI is a feasible goal.

To bridge the gap, this paper first aims to demonstrate the cultural experience of AI-based image analysis as a sensor in the recognition of plane forms in HRS, i.e., the ability to distinguish between regions with obvious historical value in terms of plane forms and regions that exhibit a higher degree of modernisation. A complete and robust digital documentation workflow based on expected positive interpretation results is established to manage the extracted hierarchical data of HRS plane forms, using several technical steps to optimise AI performance. Open source and standardisation are the principles of this workflow, ensuring broad optimisation possibilities and low barriers to use. A case study at Taihang Baxing (THBX) in China was performed, generating a hierarchically local plane forms dataset while demonstrating potential use cases.

2. Methodology

2.1 Workflow

Figure 1 illustrates the key steps included in the digital documentation workflow. Three technical parts were involved, which are distinguished by colour for visualisation. Generating image data based on deep learning is the most critical step (Red), which manifests itself in the form of performing instance segmentation tasks in computer vision. It includes two successive deep learning modelling (DLM) steps. The first step generates region-level hierarchical data, recorded as DLM(A), while the second step infers the building footprint locations and masks within the region, defined as DLM(B). To support two DLMs respectively, the input data needs to undergo different preprocessing (Blue). The basic image data was collected from open remote sensing map services, and before being input into DLM(A), historical regions and modern regions were annotated based on knowledge and evidence. For DLM(B), only the building footprints are labelled, which does not require relevant survey experience, but requires the identification of representative remote sensing images of HRS. After executing two DLMs, post-processing also helps to produce optimised results (Green). After completing DLM(A), instances in the two classes are merged. Remote sensing images will use the two types of semantic boundaries after fusion as a reference for segmentation, outputting semantic sub-images that highlight historical regions, modern regions, and backgrounds, and then performing DLM(B) data pre-processing and inference based on them. Subsequently, the DLM(B) inference results will be reorganised, and at this point, the building footprints in different sub-images will be automatically categorised. Finally, in order to ensure that the digitally archived dataset can be used more widely, building regularisation (Li et al., 2022b) and geospatial registration (Loudon et al., 1980) will be embedded.

2.2 Data Collection

China has a long history of agricultural development, which has given rise to a large number of rural settlements (Li et al., 2022a). As history progressed, the plane forms of Chinese rural settlements have undergone a transition from micro-scale self-organisation during the imperial period to macro-scale institutionalised construction in modern times, reflecting the progress of construction technology and the extension of the state's centralised governance logic into rural areas (Jin, 1983). In the early days of Reform and Opening-up era, rural settlement construction was further standardised and unified with policy support (Ye and Huang, 2016). During this period, plane forms of HRS were regarded as "inefficient" and "disorder" spatial forms sometimes (Min, 1994), since their heritage value was largely overlooked. Nation-led rural settlement construction practices gradually followed common strategies such as unified design, regular planning, centralised procurement, and mass implementation, with the plane forms became increasingly regular and grid-like consequently (Wang et al., 2017). In many HRS, this trend led to a morphological rupture between newly constructed areas and the historical cores. Logically, reviewing the changes in the construction of China's HRS, historical regions in the HRS are defined as areas built during the Chinese imperial era, and their plane forms often have self-organising features. In contrast, modern regions refer to areas in the HRS that were built in modern China, with more geometric and regular plane forms.

Fortunately, due to cost-saving considerations and practical constraints, many historical core regions were not demolished (Zhao, 1982), making them crucial samples for research on investigation and conservation of rural cultural heritage places. In order to preserve traditional culture, China recently requires the construction of revitalised villages with local identities, while discovering the morphological value retained in the plane forms of Chinese HRS is crucial to this (Zhou et al., 2019). Therefore, conducting this study in China has practical significance.

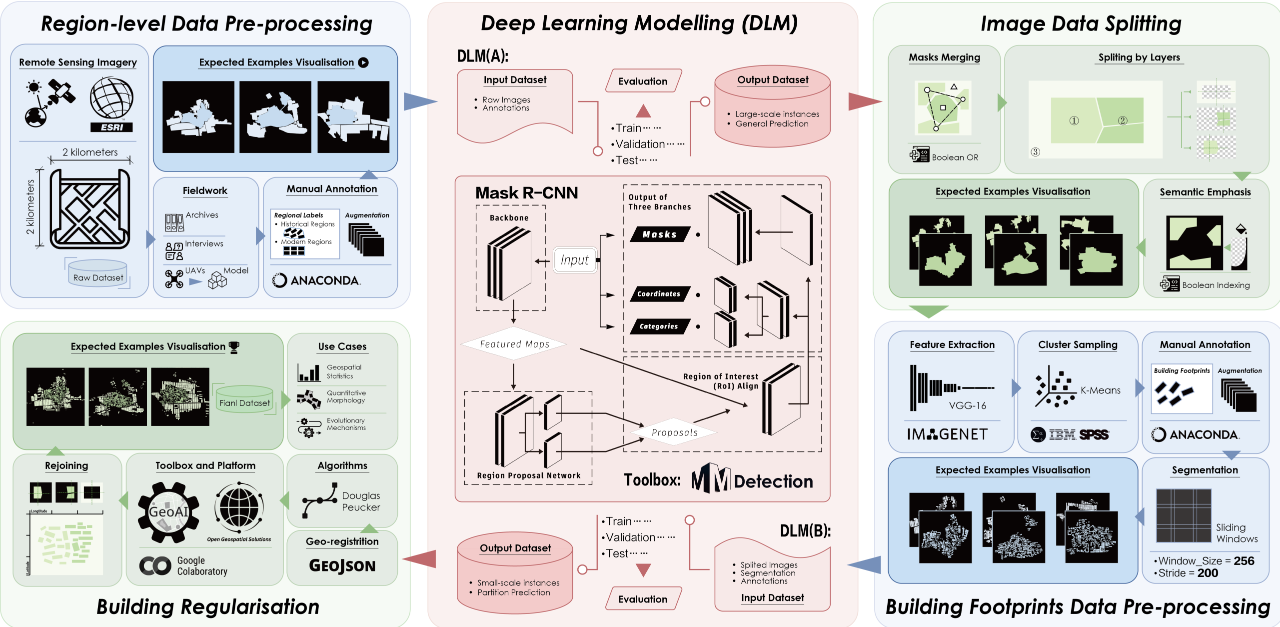

Figure 1. The digital documentation workflow of recording plane forms of HRS hierarchically

THBX has historically served as an important transportation route for communication between its neighbouring areas for several centuries (Figure 2a). In previous studies, THBX has been shown to retain significant amounts of HRS (Yuan et al., 2024). The HRS dataset used in this study is the "List of Traditional Chinese Villages (TCV)" published by Ministry of Housing and Urban-Rural Development of the People's Republic of China since 2012. The HRS on the list are recognised by a committee of experts in the conservation and development of rural settlement, and the latest entries are updated every two years. As illustrated in Figure 2b, they are usually rich in historical information and cultural landscapes, with their plane forms, retaining significant historical identities (Nie et al., 2023). Totally, there are 778 HRS in the list located in THBX.

With the development of open high resolution remote sensing imagery services, it is not a hardship to collect raw plane forms data of the HRS. More and more service providers are making their customised global remote sensing imagery freely available, which makes it possible to obtain remote sensing imagery

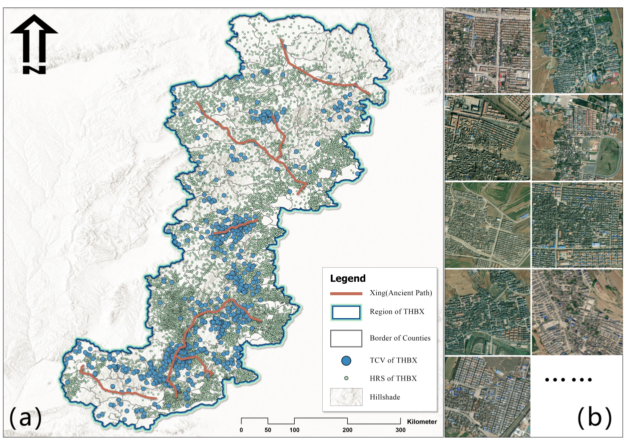

Figure 2.
(a) Spatial dataset of HRS in THBX (Yuan et al., 2024)
(b) Difference of plane forms in THBX's HRS

for a defined range at specific geographic coordinates through crawler-like tools. Thus, the plane form of a particular HRS can be read by examining remote sensing imagery within a specific area around its coordinate. After attempts and statistics, 4 square kilometres is a reasonable range as it can cover the complete plane form of each HRS in THBX as well as limit the invalid information within the extracted remote sensing images. *ESRI* World Imagery Basemap (up to one-meter) was selected as the base remote sensing dataset because it ensures image quality and geospatial accuracy. The map works seamlessly with *ArcGIS Pro*, significantly optimising the technical efficiency of the entire process from data acquisition to analysis.

2.3 Data Pre-processing

The proposed workflow involves two parts of customised deep learning of remote sensing images, and the data they require needs to be pre-processed differently. DLM(A) needs to differentiate and label the historical regions in HRS from the modern ones in the remote sensing image. The knowledge accumulated from field research in THBX and the HRS high-resolution orthophoto with modelled point cloud acquired via UAV provided the groundwork for this task, enabling accurate boundary determination between regions. The reason for labelling the modern region is to provide a contrast to the annotation of historical regions in order to highlight its typical character. Totally, 31 investigated HRS were labelled as samples and Figure 3a shows an example with essential supporting evidence and process.

In order to record a complete categorised documentation of HRS plane forms, the extraction of building footprints is indispensable. Due to the vast extent of THBX, the availability of the remote sensing images from different areas varies from time to time, resulting in extreme differences in the colour style of the extracted ones. Previous attempts at solving this problem have triggered topic of colour consistency in remote sensing images (Zuo et al., 2024). However, to ensure the generalisation of this workflow, the *VGG-16* model pre-trained on the *ImageNET* dataset is used for feature analysis of extracted remote sensing

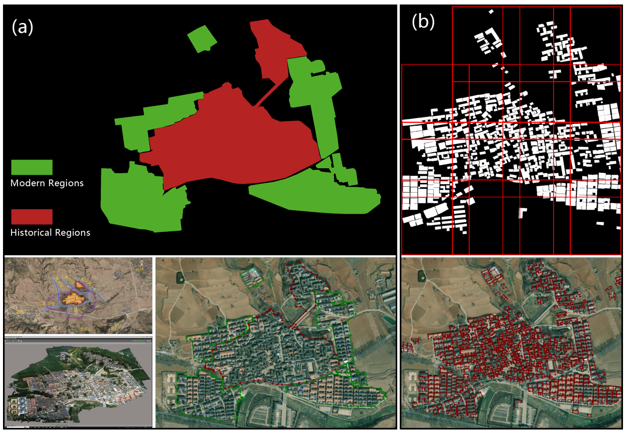

Figure 3. (a) Region-level annotations
(b) Building footprints annotations

images (Simonyan and Zisserman, 2015). This approach ensures access to information that is purely reflective of visual content rather than specific to a particular task. Afterwards, feature clustering is performed for the remote sensing images to obtain the centrally located 30 HRS for building footprint annotation. It is worth noting that buildings are small-scale instances in the 4 square kilometres of remote sensing images compared to the region-level annotation. As shown in Figure 3b, a complete image and its annotations need to be segmented into small patches of a certain size and ensure the overlap ratio between neighbouring ones before inputting it into DLM(B). In this paper, the remote sensing images used for training are partitioned according to a size of 256*256 pixels, which balances computational cost with information completeness. Eventually, regular data augmentation methods are also performed, including random rotations, random zooms, horizontal/vertical flips, and photometric transformations (Mumuni and Mumuni, 2022).

2.4 Deep Learning Model

In the proposed workflow, the technical steps strongly rely on the classical *Mask R-CNN* whose capability and performance has been proved in the past (He et al., 2017, Deng et al., 2024) . The implementation of deep learning modelling (DLM) relies on the *MMDetection* toolbox. DLMs training is implemented on *Ubuntu 22.04* with an *Intel(R) Xeon(R) Platinum 8352V CPU* and an *NVIDIA RTX 4090 GPU (24GB)*, using PyTorch 2.5.1 and CUDA 12.6. *ResNeXt101* is consistently adopted for its strong feature extraction capability, while the batch size is set to 2 for DLM(A) and 4 for DLM(B) to accommodate GPU memory constraints. 100 full epochs are performed with a learning rate of 0.001. As usual, the annotated samples are divided into training set, validation set and in the ratio of 7:3 for model training to ensure that it has reliable robustness. The dividing principle is at intra-sample level, i.e., at least one of the series of images from each HRS is included in each sets, in order to allow the DLM to maximise the learning and validation of annotated features of each remote sensing images. The testing set is no longer created separately, but all images in HRS are used for inference to directly express the generalisation ability of the DLMs. Accuracy and Loss Function were used to measure the convergence of model training, while Precision, Recall, f1 score, and mean Average Precision (mAP) were used to quantify the intrinsic performance of the model. They are commonly used evaluation metrics (Deng et al., 2024).

The evaluation of the inference quality represents the authentic stability and adaptability of the model.

2.5 Data Post-processing

The first step of data post-processing is located between DLM(A) and DLM(B). The results expected to be obtained after DLM(A) inference are at the region-level. So this workflow will segment the remotely sensed image of HRS based on the mask generated by DLM(A) detection. Although directly classifying and labelling building footprints (historical and modern) is valuable for the discovery of building-level rural heritage (Wang et al., 2023), it must be recognised that the renewal of building roof materials poses a risk to era judgement, which in turn affects the distinguishing of the region-level plane forms of the HRS. Specifically, using a pre-trained instance segmentation model from *MMDetection*, masks of historical and modern regions are extracted and applied to retain only target pixels while setting all background to black for accurate foreground visualisation. To prevent loss of semantically important but unsegmented areas, all predicted mask are merged via a pixel-wise logical Boolean *OR* operation and set to black, allowing background regions to be explicitly preserved for more complete and robust analysis.

The second step of data post-processing begins with geospatial registration, using affine transformation to convert raster spatial coordinates to geospatial coordinates (Loudon et al., 1980). Meanwhile, for accomplishing the task of building footprint regularisation, the *GeoAI*, which has risen to prominence in recent years, provides a great help. As a concept, Geospatial artificial intelligence (GeoAI) is an abstract workflow that applies AI techniques fused with geospatial data, science, and technology to accelerate real-world understanding (Janowicz et al., 2020). *GeoAI* is a Python package that built following the workflow, hosted by the Open Geospatial Solutions (*Opengeos*) GitHub organisation. A variety of widely validated building footprint regularisation methods have been included, as well as interactive and visualised user interface. Consequently, utilising *GeoAI*'s services not only makes the proposed workflow more concise and useful, but also allows for the future output of HRS plane forms that can be adapted to geospatial analyses (Wu, 2024).

3. Results

3.1 Evaluation of DLM Performance

The evaluation of model training and performance begins with the learning curves. As illustrated in Figure 4, both models converged after 100 training epochs, with steadily increasing accuracy and significantly decreasing loss, indicating effective and stable training. DLM(A) reached a final accuracy of 98.93%, with a loss stabilizing around 0.18. DLM(B) performed better, achieving 99.21% accuracy, with loss reduced to around 0.15. The total training times for the two models were 5 hours 23 minutes 21 seconds and 54 hours 27 minutes 9 seconds, respectively, both within an acceptable range, considering the mass input small-scale instances of DLM(B). To further assess the performance, four standard metrics were computed: Precision, Recall, F1-score, and mAP, as summarised in Table 1. The evaluation results indicate that DLM(B) outperformed DLM(A) across most metrics, with an mAP of 0.758 vs. 0.699, Precision of 0.816 vs. 0.733, and F1 Score of 0.784 vs. 0.759. The performance gap may be attributed to the clearer and more

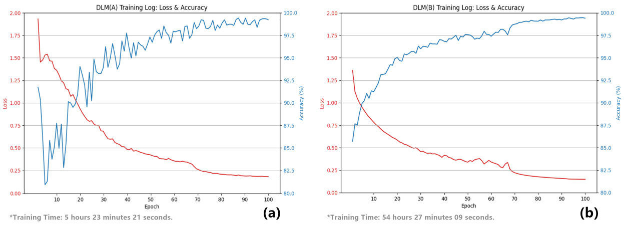

Figure 4. Accuracy and Loss of (a) DLM(A), (b) DLM(B)

DLM	Perfomance Metrics			
	mAP	Precision	Recall	F1 Score
DLM(A)	0.712	0.733	0.786	0.759
DLM(B)	0.748	0.816	0.725	0.768

Table 1. Comparison of DLM performance metrics

structured features of building footprints. Notably, DLM(A) achieved a relatively higher Recall score (0.786 vs. 0.755), suggesting greater sensitivity in capturing fuzzy or spatially diffuse historic regions. With comparison to existing research (Deng et al., 2024), both DLMs demonstrated solid performance in their respective tasks, providing reliable support for downstream feature recognition and general objects extraction.

3.2 Review of Generalisation Test

To further evaluate the practical capacity of the model and advance the digital documentation of HRS, the well-performing model from the previous section was applied to all HRS samples through a full-scale automated pipeline, resulting in the vectorised plane forms dataset of HRS in THBX with geographical coordinates. Figure 5 presents several representative examples, where the model successfully differentiates between historical and modern regions with coherent planning layouts, demonstrating its practical utility and accuracy for digital documentation. To verify consistency between the model results and the initial sample selection logic, all detection outcomes were categorised and analysed: (1) no detected regions, (2) historical regions only, (3) modern regions only, and (4) coexisting historical and modern regions (best matches the sampling reason). As shown in Figure 6, the category (4) comprises the highest proportion, indicating that the samples generally exhibit hybrid spatial characteristics and validating their representativeness in documenting HRS diversity.

Figure 5. HRS plane forms hierarchical record dataset (Partial)

3.3 Use Case: Geospatial Statistics

To demonstrate the practical utility of the produced dataset, this section presents an initial use case from a geospatial perspective. First, the region ratio between historical and modern regions was calculated for each HRS sample, and the results were visualised as a spatial distribution map. As shown in Figure 7a, the proportion of historical region varies across samples and displays spatial clustering or dispersion patterns across the THBX, indicating geospatial sensitivity and analytical relevance. Second, to assess the suitability of the ratio data for spatial analysis, a Global Moran’s I test was conducted. The results indicate that there is a statistically significant spatial clustering effect (Figure 7b), representing that the ratio change is dependent on spatial location.

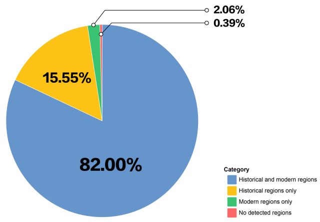

Figure 6. Proportion of different coexisting categories

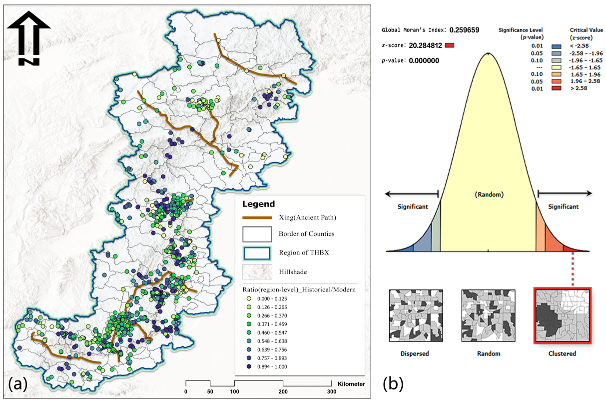

Figure 7. (a) Geo-visualisation of the historical/modern ratio (b) Global Moran’s Index

4. Discussion

4.1 Insights and Contributions

With a considerable amount of exploration of the concept of deep learning-based image analysis as a sensor in the field of built environment research, it is clear that it has the potential to expand the dimensions of the first-hand data collected. Compared with general rural settlements, the plane forms of HRS retain more complex historical information, which is particularly important for understanding and highlighting the heritage morphological value. Initially, it is important to extract the historical information condensed in the plane forms from the current built environment from the HRS, which is like a palimpsest. However, the trend toward intelligentisation of related technologies has not been fully integrated into such research topics. Therefore, in this paper, the contribution of deep learning technology to the extraction of historical information from HRS plane forms is explored, and a procedural and automated workflow is built around this core. The proposed workflow clarifies the availability of open high-resolution remote sensing images and provides a temporally morphological interpretation in the HRS digital documentation.. Given that changes in the built environment and even social relations in the development of rural settlements can be reflected in their plane forms, the HRS plane forms need to be interpreted in a chronological manner. In China, this interpretation can be preliminarily implemented in the hierarchical classification of historical regions and modern regions. Therefore, the plane forms of HRS are not only represented by single current characteristics, but also by the extent and pattern of change. This perspective on change is more appropriate for HRS because it is more in line with HRS's positioning as a kind of living heritage. It is hoped that this workflow, and the resulting digital documentation dataset of HRS plane forms, will contribute to and provide insights into a variety of domains.

First, compared with the previous research paradigm of treating HRS and even rural settlements as single entities in plane forms, and the application paradigm of using computer vision to directly extract building footprints, this study verifies the ability of AI to classify region-level plane forms in HRS through pixel feature classification. Observed from remote sensing images of HRS, the certified historical regions (often officially designated conservation areas) have unique plane forms. However, although judging historical morphological information based on regional characteristics can avoid the biases from roof materials renewal, it is highly dependent on the quality or knowledge of the sensors (often experts in the fields of architecture or urban and rural planning). Fortunately, the experiments in this paper verify that features recognised based on human professions can also be identified impartially by AI, which has extraordinary significance in terms of both concept and application. It demonstrates that AI has broader potential in the investigation of plane forms in HRS, becoming a cognitive co-constructor in understanding heritage, and also produces acceptable explanations from the perspective of technical agents, verifying that historical regions in HRS do have special visual features.

Second, based on the new advantages of AI mentioned above, a digital documentation workflow that bridges the gaps between different technologies was produced in this study. The proposed workflow is interlinked and can truly output a HRS plane forms hierarchical digital dataset format that is applicable to most research with novel temporal perspective. The region-level detection results based on deep learning can be regarded as labels for building footprints in this workflow. By segmenting the remote sensing image patches corresponding to the region mask, representing irrelevant areas in black and maintaining the original image size, multiple extraction results from the same HRS can be easily re-merged. Building regularisation in the workflow can simplify and vectorise the pixelated mask boundaries of building footprint extraction, reducing redundant information while optimising the visualisation of the dataset. It is clear that this workflow originated from straightforward, practical interdisciplinary heritage morphology research, can be directly replicated to support HRS plane form surveys, and can be optimised by incorporating other algorithmic technologies according to different needs. GeoAI was integrated into the final step of the workflow, which retained the geo-information of the output dataset, allowing it to be used for larger-scale geospatial analysis. Moreover, another contribution of this paper is the hierarchical dataset of plane forms generated using THBX's HRS as an example. The current interest in researching THBX rural heritage does not match its confirmed historical and cultural value, with most focusing on geospatial distribution that already integrated with temporal perspective rather than settlement morphology. Therefore, the customised dataset from this workflow is a commitment to specific research cases in this paper, which may inspire more morphological research on THBX's HRS, which can, in turn, promote this recording format more widely, forming a recognised HRS digital recording format and form a publicly recognised standard.

4.2 Limitations and Future Opportunities

It must be noted that this paper still has many limitations, which also represent opportunities for future research on this topic. First, the THBX results need to be further examined to summarise the unidentified and misidentified issues that arise. Unlike typical AI-based computer vision tasks, where high accuracy rates can be used to determine the excellence of the output model, in heritage research, especially when it comes to confirming critical information, human effort is needed to check the results, because what matters is simply "yes" or "no", rather than the correction rate. That is why digital technology is more suitable as an aid to heritage investigation rather than as a dominant force; fieldwork remains the core (Moullou et al., 2024). In the dataset produced by this study, another point that can be continuously optimised in the future involves both the acquisition of remote sensing images and the recording of plane forms. Due to the close proximity or direct adjacency of different HRS, building footprints from other settlements that are not of interest need to be manually removed when recording data. One possible solution is based on the widespread historical regions in HRS and their special plane forms. First, such historical regions can be identified, and then continuous or close locations can be summarised through distance clustering methods like *DBSCAN* algorithm to derive the specific property rights scope of HRS (Zhang et al., 2021). For China, a national-level geospatial distribution dataset of rural buildings has been fully established (Deng et al., 2024), and the intelligent identification of property boundaries for specific rural settlements is an important data dimensional expansion initiative similar to AI-based identification and digital documentation of HRS plane forms.

Secondly, since this paper is conducted from a structural perspective, no comparison between deep learning models was conducted. Although completing this part of the work is common in the core content of other research on this topic, it is undeniable that the core concept it reflects is also an emphasis on the quality of results (Chen et al., 2021). However, this part of the work is time-consuming and deviates from the main focus of this paper, so it has been left for future research. The proposed workflow does not embed DLM in code and scripts, so the relevant technical modules are always open to be optimised based on evidence from sources relevant to this topic. Third, the morphological information that needs to be extracted from HRS in layers is diverse. For example, the external space of buildings, as represented by road forms, reflects the social logic of settlements. Because this characteristic, like historical plane forms, is difficult to change completely, it is also applicable to the workflow proposed in this article, but requires changes in annotations (Jia et al., 2023). The constraint encountered in this study in achieving this task is the quality of remote sensing images. Higher resolution and unobstructed remote sensing images would be very useful, as they would generate high-quality annotations that are critical for supervised learning. In addition, 3-D geo-information represented by terrain also significantly affects the emergence and evolution of HRS plane forms, which is a key point in settlement morphology. The THBX in this paper is located in a predominantly mountainous terrain, and previous studies have also proven that the generation of HRS in THBX is affected by multi-scale terrain (Yuan et al., 2024), which means their plane forms can be documented and visualised better supported by integrating 3-D reconstruction and geo-information. Therefore, future opportunities may lie in identifying and validating whether specific 3-D geo-information datasets can be integrated (Jurado-Rodríguez et al., 2024). The final limitation is also related to 3-D geo-information, too. Specifically, the use of point cloud models in this study is only an empirical reference, while other studies have already incorporated them into deep learning modelling and achieved satisfactory results in rural building automatic classification (Meng et al., 2022). Such fusion learning of multimodal data may provide new insights into future workflow updates and result quality for this research topic.

5. Conclusion

With the close integration of cultural heritage conservation and advanced technology, the digital documentation is being seriously considered and widely adopted. As for built heritage, data-driven digital documentation captures and preserves its physical existence, and digitises heritage information from a digital twin perspective, which is suitable for further analysis. As a living type of built heritage, rural settlements are constantly dynamic, generating and accumulating similar or different information. On the recommendation of *ICOMOS-IFLA*, rural settlements are considered to be a highly valuable type of heritage due to their unique landscapes. In the impact of over-development and non-development, support is needed to protect the physical existence of rural settlements represented by HRS and to extract their potential heritage value. The plane forms examined in this study is not merely a visual representation of spatial distribution in rural settlements, but a spatial carrier of heritage value. As HRS have undergone various stages of social development, the plane forms shaped in different periods exhibit both differentiation and coexistence through cumulative layering, thus forming the composite spatial structure of present-day HRS in alignment with the “palimpsest” criteria for rural heritage as defined by *ICOMOS-IFLA*.

Therefore, this paper introduces the idea of using AI-based tools as heritage history information sensors, and uses this to create a workflow that is close to automation, allowing HRS, a representative of rural cultural heritage places, to be digitally recorded in the plane form data of publicly available remote sensing images, while at the same time identifying important historically interpretive information. At THBX, an exploratory case study conducted in this historically and culturally significant region of China showed that the proposed workflow revealed the potential of AI model in perceiving plane form features that are often judged based on human knowledge. In this workflow, the identification of plane forms of historical regions in HRS based on AI technical agents creates a cultural experience dimension for the plane form data of HRS, which is ultimately integrated into a vectorised, regularised, layered record data set that summarises and highlights the morphological value at the core of HRS. Such a latent aspect can promote the in-depth development of rural heritage conservation and research, and unlock the potential of digital technology to efficiently interpret more heritage information. Despite its limitations, this procedural workflow can still serve as a useful tool for the hierarchical archiving of HRS plane forms, helping to advance the structuring and regularisation of morphology information in rural cultural heritage places. The presented work demonstrates the practicality of transferring techno-optimism to a more social and cultural focus in data-driven heritage conservation. Regarding future plans, use cases based on customised digital documentation datasets will be explored, and the possibility of identifying and extracting more types of heritage information will be integrated.

6. Acknowledgements

Sincere thanks to the members of *VRLab* and *HBE Studio* for their discussions, to Mr. Huapeng Tang, Mr. Jiawei Zhang, and Mr. Linxi Zhang for their assistance in data annotation and to Ms. Shirley Wong for her technical support in *Mask R-CNN*. The research is supported by NSFC general programmes: Quantitative Research on the Formation Composition of Traditional Villages in the Taihang Mountains (No. 51978441); Qing Official Architectural Drawings as a Knowledge Base (No. 52378028).

References

Aamir, Z., Seddouki, M., Himmy, O., Maanan, M., Tahiri, M., Rhinane, H., 2023. Rural settlements segmentation based on deep learning U-NET using remote sensing images. *ISPRS Archives*, XLVIII-4/W6-2022, 1–5.

Argyridou, E., Efstathiou, K., Hadjiathanasiou, M., Ioannides, M., Karaoli, A., Karittevli, E., Mateou, M., Panagi, I., Samara, P., 2023. Digital holistic documentation of cultural heritage: challenges and risks, towards shaping the future. *ISPRS Archives*, XLVIII-M-2-2023, 95–102.

Australia ICOMOS, 2013. The Burra Charter: The Australia ICOMOS Charter for Places of Cultural Significance. https://australia.icomos.org/publications/burra-charter-practice-notes/, Last accessed on 10 June 2025.

Bai, N., Nourian, P., Pereira Roders, A., Bunschoten, R., Huang, W., Wang, L., 2023. Investigating rural public spaces with cultural significance using morphological, cognitive and behavioural data. *Environment and Planning B: Urban Analytics and City Science*, 50(1), 94–116.

Chen, J., Wang, G., Luo, L., Gong, W., Cheng, Z., 2021. Building area estimation in drone aerial images based on Mask R-CNN. *IEEE Geoscience and Remote Sensing Letters*, 18(5), 891–894.

Deng, W., Xu, W., Huang, Y., Li, X., 2024. A large-scale multipurpose benchmark dataset and real-time interpretation platform based on Chinese rural buildings. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17, 10914–10928.

Ding, W., 2021. Morphology and Typology: A Village as a Cultural and Environmental Process. V. Oliveira (ed.), *Morphological Research in Planning, Urban Design and Architecture*, Springer International Publishing, Chambridge, 117–139.

Du, X., Shi, D., 2019. Rural heritage: value, conservation and revitalisation: from the perspective of the human-Land relationship. *Built Heritage*, 3(2), 1–6.

Galeazzi, F., 2018. 3-D Virtual Replicas and Simulations of the Past “Real” or “Fake” Representations? *Current Anthropology*, 59(3), 268–286.

Gong, J., Jian, Y., Chen, W., Liu, Y., Hu, Y., 2022. Transitions in rural settlements and implications for rural revitalization in Guangdong Province. *Journal of Rural Studies*, 93, 359–366.

He, K., Gkioxari, G., Dollar, P., Girshick, R., 2017. Mask R-CNN. L. O'Conner (ed.), *IEEE International Conference on Computer Vision (ICCV), Venice, ITALY, October 22-29, 2017, Conference Track Proceedings*, 2980–2988.

ICOMOS-IFLA, 2017. ICOMOS-IFLA Principles Concerning Rural Landscapes as Heritage. https://rm.coe.int/1680083746, Last accessed on 10 March 2025.

Janowicz, K., Gao, S., McKenzie, G., Hu, Y., Bhaduri, B., 2020. GeoAI: spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. *International Journal of Geographical Information Science*, 34(4), 625–636.

Jia, Y., Zhang, X., Xiang, R., Ge, Y., 2023. Super-resolution rural road extraction from Sentinel-2 Imagery using a spatial relationship-informed network. *Remote Sensing*, 15(17), 4193.

Jin, O., 1983. The changing rural development. *Architectural Journal*, 10, 27–31. (In Chinese).

Jurado-Rodríguez, D., Latorre-Hortelano, P., René-Dominguez, L., Ortega, L., 2024. 3D modeling of rural environments from multiscale aerial imagery. *Computers & Graphics*, 122, 103982.

Li, H., Yuan, Y., Zhang, X., Li, Z., Wang, Y., Hu, X., 2022a. Evolution and transformation mechanism of the spatial structure of rural settlements from the perspective of long-term economic and social change: A case study of the Sunan region, China. *Journal of Rural Studies*, 93, 234–243.

Li, Q., Zorzi, S., Shi, Y., Fraundorfer, F., Zhu, X., 2022b. Reg-GAN: an end-to-end network for building footprint generation with boundary regularization. *Remote Sensing*, 14(8), 1835.

Loudon, T., Wheeler, J., Andrew, K., 1980. Affine transformations for digitized spatial data in geology. *Computers & Geosciences*, 6(4), 397–412.

Meng, C., Song, Y., Ji, J., Jia, Z., Zhou, Z., Gao, P., Liu, S., 2022. Automatic classification of rural building characteristics using deep learning methods on oblique photography. *Building Simulation*, 15(6), 1161–1174.

Min, C., 1994. Village constructionprojects are booming and urgently need to be strengthened by scientific guidance. *Housing Science*, 12, 39–40. (In Chinese).

Moullou, D., Vital, R., Sylaiou, S., Ragia, L., 2024. Digital tools for data acquisition and heritage management in archaeology and their impact on archaeological practices. *Heritage*, 7(1), 107–121.

Mumuni, A., Mumuni, F., 2022. Data augmentation: a comprehensive survey of modern approaches. *Array*, 16, 100258.

Nancarrow, J., Yang, C., Yang, J., 2021. An emotions-informed approach for digital documentation of rural heritage landscapes: Baojiatun, a traditional tunpu village in Guizhou, China. *Built Heritage*, 5(1), 11.

Nie, Z., Chen, C., Pan, W., Dong, T., 2023. Exploring the dynamic cultural driving factors underlying the regional spatial pattern of Chinese Traditional Villages. *Buildings*, 13(12), 3068.

Shi, Z., Ma, L., Zhang, W., Gong, M., 2022. Differentiation and correlation of spatial pattern and multifunction in rural settlements considering topographic gradients: Evidence from Loess Hilly Region, China. *Journal of Environmental Management*, 315, 115127.

Simonyan, K., Zisserman, A., 2015. Very deep convolutional networks for large-scale image recognition. Y. Bengio, Y. LeCun (eds), *International Conference on Learning Representations (ICLR), San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 1–14.

Tao, J., Li, G., Sun, Q., Chen, Y., Xiao, D., Feng, H., 2023. An approach for identifying historic village using deep learning. *SN Applied Sciences*, 5(1), 13.

Tarrafa Pereira da Silva, A., Pereira Roders, A., 2012. Cultural heritage management and heritage (impact) assessments. K. Michell, P. Bowen, K. Catell (eds), *International Conference on Facilities Management, Procurement Systems and Public Private Partnership, Edinburgh, Scotland, UK, January 23-25, 2012, Conference Track Proceedings*, 375–382.

Wang, J., Zhao, J., Wu, T., Tang, L., 2017. New-type intensive rural settlements in China based on the unified and self-construction mechanisms of spatial organization. *International Journal of Sustainable Development and Planning*, 12(6), 1073–1084.

Wang, W., Shi, Y., Zhang, J., Hu, L., Li, S., He, D., Liu, F., 2023. Traditional village building extraction based on improved Mask R-CNN: a case study of Beijing, China. *Remote Sensing*, 15(10), 2616.

Wu, Q., 2024. GeoAI: A powerful Python package for integrating Artificial Intelligence with geospatial data analysis and visualization. https://opengeoai.org/acknowledgments, Last accessed on 10 June 2025.

Ye, L., Huang, Y., 2016. Design down to villages again: a study of rural construction in the early years of Reform and Opening Period, 1978-1994. *Architectural Journal*, 11, 10–15. (In Chinese).

Yuan, S., Mu, N., Bai, J., Zhang, H., 2024. Spatial heterogeneity analysis on distribution patterns and underlying determinants of historical rural settlements: a case study in Taihang Baxing. *Heritage Science*, 12(1), 395.

Zhang, M., Li, Z., Wu, X., 2021. Semantic segmentation method accelerated quantitative analysis of the spatial characteristics of traditional villages. *ISPRS Archives*, XLVI-M-1-2021, 933–939.

Zhao, P., 1982. Several practices for saving land in rural housing construction. *Agricultural Economy*, S1, 59. (In Chinese).

Zhou, M., Chu, S., Du, X., 2019. Safeguarding traditional villages in China: the role and challenges of rural heritage [reservation. *Built Heritage*, 3(2), 81–93.

Zhu, Q., Liu, S., 2023. Spatial morphological characteristics and evolution of traditional villages in the mountainous area of southwest Zhejiang. *ISPRS International Journal of Geo-Information*, 12(8), 317.

Zuo, Z., Li, Y., Zhang, T., Mo, F., 2024. A NeRF-based color consistency correction method for remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 17, 6805–6817.