

Establishing a Historical Fine-Resolution Urban Database from KH-9 HEXAGON MCS and PCS Imagery Including an Urban Landscape Map, Stereo Anaglyph Image and Digital Surface Model

Amir Reza Shahtahmassebi¹, Golnaz Shahtahmassebi², Nathan Moore³, Peter M. Atkinson^{4,5,6}

¹Architectural Engineering School, Qingdao Hengxing University of Science and Technology, Qingdao 266100, China

²Department of Physics and Mathematics, School of Science and Technology, Nottingham Trent University, UK

³Department of Geography, Michigan State University, East Lansing, MI 48823, USA

⁴ Lancaster Environment Centre, Lancaster University, Bailrigg, Lancaster LA1 4YW, UK

⁵ Geography and Environmental Science, University of Southampton, Highfield, Southampton SO17 1BJ, UK

⁶ College of Surveying and Geo-Informatics, Tongji University, No.1239, Siping Road, Shanghai, PR China, 200092

Telephone: 0086550-88982272

*Corresponding authors: amirreza208@hotmail.com;

AmirReza Shahtahmassebi (amirreza208@hotmail.com); Golnaz Shahtahmassebi (golnaz.shahtahmassebi@ntu.ac.uk); Nathan Moore (moorena@msu.edu). Peter M. Atkinson (pma@lancaster.ac.uk)

Establishing a Historical Fine-Resolution Urban Database from KH-9 HEXAGON MCS and PCS Imagery Including an Urban Landscape Map, Stereo Anaglyph Image and Digital Surface Model

Abstract: Historical, fine-scale spatial information on urban regions such as heights, spatial patterns and categories are critical for various disciplines. However, obtaining such historical information prior to the year 2000 is challenging. Although declassified analog images from the KH-9 HEXAGON (hereafter, KH-9)'s mapping camera system (MCS) and panoramic camera system (PCS) at meter to sub-meter resolutions during 1970s are available, the absence of camera meta parameters (e.g., focal length, extrinsics, lens distortion) and the single film-based panchromatic image format pose significant limitations. Herein, a multifaceted analytical framework is developed to reconstruct a historical fine-scale urban database from KH-9 MCS and KH-9 PCS images which relies on minimal technical information. For the KH-9 MCS, a deep learning classification framework is proposed to retrieve historical urban landscape features from KH-9 MCS imagery. Also ablation studies are proposed to examine the impacts of deep learning hyperparameters on historical urban landscape classification and to assess the efficiency of deep learning models in comparison with benchmark machine learning techniques. For the KH-9 PCS, first a framework is developed using SuperPoint+SuperGlue (SPSG) models to create historical ground point controls (GCPs) from KH-9 PCS for accuracy assessment purposes and estimating the exterior orientation of the stereo cameras. Structure from motion (SfM) techniques are adopted to create an ultra-fine spatial urban stereo-anaglyph image and digital surface model (DSM). Finally, ablation investigations are conducted to compare the performance of the developed techniques with benchmark approaches. Our results suggest that KH-9 MCS and PCS images, when analyzed with appropriate methodologies, can significantly enhance the reconstruction of a historical fine-scale urban database. Moreover, both KH-9 MCS and PCS images can contribute to revealing super-urbanization (i.e., unprecedented urban development) including super-new urban development and super-urban redevelopment as well as super-urban no-change (i.e., stable over the four decades, 1975-2023).

Keywords: KH-9 HEXAGON; mapping camera system (MCS); panoramic camera system (PCS); deep learning; SuperPoint+ SuperGlue; Historical GCPs; stereo anaglyph; historical DSM; super-new urban development, super-urban redevelopment and super-urban no-change

1. Introduction

1.1 Background, importance, research questions and research objectives

A historical urban benchmark dataset at fine spatial resolution (FSR) prior to the launch of contemporary FSR satellite sensors in the 2000s (e.g., IKONOS, QuickBird) is useful for sustainable urban planning and informed decision-making, but is currently not readily available to the majority of researchers (Niang et al. 2020). Declassified analog panchromatic imagery from KH-9 HEXAGON (hereafter, KH-9) cameras can potentially fill this gap, opening up new avenues for land use and land cover science such as urban research (Sertel et al. 2024). The KH-9 is one of the satellite missions of the United States of America (USA) Keyhole (KH) reconnaissance program operated during the Cold War period. Among the many advantages of the KH-9 missions for urban research are: captured during 1972-1981, ultra fine (0.6-1.2 m) and fine (6-9 m) spatial resolution, stereoscopic cameras, surveyed nearly all of the Earth's surface (excluding Greenland, Antarctica and Australia) at multiple times and large geographic coverage (Hammer et al. 2022).

KH-9 HEXAGON included two cameras: the mapping camera system (MCS) and the panoramic camera system (PCS). MCS is a moderate spatial resolution camera (20 to 30 feet, or about 6-9 m, spatial resolution), whose images were released in 2002 (Dehecq et al. 2020). The PCS used a fine spatial resolution stereo-panoramic camera (2 to 4 feet or about 0.6 to 1.2 m spatial resolution), whose images were declassified in 2011, and have been made available to the public recently (Hammer et al. 2022). Although KH-9 (MCS and PCS) images have been used in non-urban research (Zhou et al. 2018, Zhou et al. 2021, Fowler 2022, Marzolf et al. 2022), limited research exists utilizing KH-9 camera data (MCS and PCS) in urban investigations. This paucity of information is presumably due to various bottlenecks including: (a) lack of sufficient classification frameworks for generating thematic maps from a single panchromatic image (Sertel et al. 2024), (b) absence of metadata for reconstructing stereo anaglyph imagery and digital elevation model (DSMs) (Hammer et al. 2022), (c) noise and contrast degradation (Shahtahmassebi et al. 2023), (d) lack of historical ground control point (GCP) data for analytical purposes such as labeling, accuracy assessment and DSM estimation and (e) geometric distortions (Surazakov and Aizen 2010). This research focuses on developing frameworks for classification, establishing a historical GCP database, generating stereo anaglyphs and reconstructing DSMs at FSR from historical KH-9 data (MCS and PCS) over urban regions.

In terms of classification, KH-9 MCS are suitable of generating thematic maps over urban regions at FSR (6-9 m) in which such thematic maps are compatible with contemporary fine to medium spatial resolution data such as Sentinel sensor data (10 m) and Landsat sensor data (30 m multispectral and 10 m panchromatic data) for urban landscape change detection and urban growth modelling. With recent advances in computer software and hardware, many deep learning models have been developed to classify land use and land cover (LULC) from remotely sensed data (Ma et al. 2019, Neupane et al. 2021, Yuan et al. 2021, Li et al. 2024). Deep learning models employ a range of components such as several hidden layers (e.g., convolutional neural network (CNN)), skip connection among layers, activation layers, fully connected layers, loss function and pooling layer to extract the most salient properties of remotely sensed images and, thus, increasing the accuracy of pattern recognition and semantic segmentation (Albarakati et al. 2024, Ruhab et al. 2024). Despite these advantages, the use of deep learning models for classifying urban landscapes from KH-9 MCS panchromatic imagery introduces challenges including: (a) classifying urban landscapes from a single panchromatic band is a non-trivial task due to the complexity of urban landscapes which can include spectral similarity across different classes (Pacifi et al. 2009). For example, urban features such as dark impervious surfaces, soils and vegetation cover all possess similar spectra in panchromatic imagery, causing potential misidentification; (b) while the current literature has provided valuable information for tuning the hyperparameters of deep learning models for contemporary remotely sensed data acquired by digital sensors (particularly multispectral data), such guidance is missing for historical panchromatic remotely sensed photographs, especially declassified analog panchromatic KH-9 MCS images (and KH-9 PCS images) (Mboga et al. 2020, Sertel et al. 2024); (c) many labels are required to train and validate deep learning models efficiently and achieve satisfactory performance (Albarakati et

al. 2024, Ruhab et al. 2024). However, establishing such a database particularly for urban mapping purposes based on KH-9 MCS imagery or other historical remotely sensed imagery is challenging partly due to lack of information and partly due to rapid land use and land cover changes; and (d) there also exists a lack of robust knowledge of the performance of state-of-the-art machine learning techniques against deep learning in the context of urban landscape mapping from KH-9 MCS imagery (with respect to the aforementioned problems).

Focusing on stereo anaglyph imagery and DSM, KH-9 PCS images are valuable resources for retrieving such information at ultra fine spatial resolution (UFSR) (0.6-1.2 m) as they were acquired by stereoscopic cameras. Many photogrammetric frameworks in current remote sensing software have been examined to generate stereo anaglyph imagery and DSMs from earlier versions of KH data such as CORONA (Altmaier and Kany 2002, Galiatsatos et al. 2008, Watanabe et al. 2017, Nita et al. 2018, Ghuffar et al. 2022) and KH-9 PCS (Zhou et al. 2021). To do so, however, such frameworks require two crucial components. First, meta parameters (e.g., focal length, extrinsics, lens distortion) of remote sensing cameras or sensors are essential to estimate the position of sensors or cameras accurately for subsequent procedures such as alignment of stereo images, bundle adjustment and creating point clouds. Second, historical GCPs play a central role in such frameworks to assess the accuracy of outcomes. However, the limited availability of technical parameters for KH-9 PCS cameras and an absence of FSR historical survey data providing GCPs can impede the utilization of such algorithms in reconstructing historical stereo anaglyph imagery and DSMs, particularly over urban regions. Generally, many urban regions have undergone rapid land use and land cover changes (LULCCs) which can limit collecting historical ground survey data. Moreover, collecting FSR GCP ground data by means of a field campaign involving Global Navigation Satellite Systems (GNSS) data would be challenging in many countries.

Considering the advantages and shortcomings of analog panchromatic KH-9 imagery (MCS and PCS) for urban investigations, five major research questions arise:

(a) KH-9 MCS:

- (1) Analyzing historical urban landscapes often focuses on a limited region rather than assessing entire land-use and land-cover patterns. How can appropriate deep learning models be developed for retrieving historical thematic urban information from KH-9 MCS?
- (2) Do the hyperparameters of deep learning models impact the performance of urban landscape classification from KH-9 MCS?
- (3) How do existing machine learning techniques compare to deep learning models for historical urban landscape classification?

(b) KH-9 PCS:

- (4) How can historical urban GCPs be established by considering that urban regions may be affected by unprecedented LULCCs?
- (5) How can ultra-FSR stereo anaglyph imagery and DSMs be created from stereoscopic KH-9 PCS imagery with minimal camera meta-parameters?

Thus, the aim of this research was to provide a comprehensive and systematic investigation to develop better analytical approaches and merits for reconstructing a historical fine-scale urban database. The database includes historical urban features, historical GCPs, urban stereo anaglyphs and urban DSMs in an area characterized by a highly heterogeneous urban landscape comprising a complex mixture of water bodies (e.g., lakes, ponds and river networks), impervious surfaces, shadow (e.g., topographic and urban), urban forested regions, semi-natural surfaces (e.g., urban green spaces) and non-impervious surfaces (e.g., farmlands). The database was built through employing both a single KH-9 MCS image in 1975 and stereo pairs of KH-9 PCS images in 1974. Moreover, sufficient data were available with which to implement the research and support our results. The main contributions of this research were: with respect to KH-9 MCS (a) to propose a deep classification framework for retrieving historical urban landscapes from KH-9 MCS imagery, (b) to assess the impacts of hyperparameters of the deep learning model on reconstructing the historical urban landscape from KH-9 MCS, and (c) to conduct a thorough ablation investigation to identify the challenges and effectiveness of the proposed deep learning model in comparison to existing benchmark machine learning techniques. With respect to KH-9 PCS (a) to devise an approach by leveraging a deep learning model and graph neural network to establish historical urban GCPs through comparison between FSR imagery in 2023 and KH-9 PCS in 1975(1975) (b) to develop Structure from Motion (SfM) approaches for creating a stereo anaglyph image and DSM (c) to design appropriate metrics for stereo anaglyph image quality assessment and (d) to perform a detail ablation study to scrutinize the performance of the proposed approaches in comparison with benchmark techniques. This research also demonstrated the potential for application of the generated results from KH-9 MCS and PCS imagery to urban investigation.

1.2 Related work

1.2.1 Literature on historical land use and land cover (LULC) classification using KH data

Several scholars have used machine learning techniques to classify LULC from KH data (i.e., all KH cameras). Song et al. (2015) proposed a framework for classifying forested and non-forested regions from CORONA KH 4A and 4B (the earliest of KH camera model) between 1966 and 1967 using integration of gray level co-occurrence matrix (GLCM) measures and a support vector machine (SVM). Although the established forest and non-forest map achieved an overall accuracy of about 90%, this research highlighted the need to develop more automated approaches for classifying KH data. Saleem et al. (2018) used a combination of textures and a maximum likelihood classifier (MLC) to classify CORONA in 1969 into five major classes: water bodies, built-up, bare land, forest and agriculture. Moreover, their benchmark analysis studies showed that the MLC achieved an accuracy of above 85% which was comparable with artificial neural network (ANN) and random forest (RF) approaches. Rendenieks et al. (2020) mapped forest cover from CORONA (KH-4B) in 1967 using object based image analysis (OBIA). This research suggested that inter- and intra-segment heterogeneity can be used for selecting an optimal scale. In addition, this investigation showed that

OBIA was advantageous for panchromatic CORONA imagery. Deshpande et al. (2021) developed two supervised classification approaches to classify historical LULC from CORONA imagery into four major classes: agriculture, built-up land, barren land and water body. The first method employed a two-dimensional convolutional neural network (2D-CNN) classifier while the second approach integrated geometric moments (GM)-based texture features and pixel-wise RF classification. Both the 2D-CNN and GM-RF approaches achieved overall accuracies of >90%. Shahbandeh et al. (2023) provided a systematic and comprehensive review of applications of CORONA imagery for LULC changes in which a part focused on techniques for classifying historical LULC from CORONA data.

Although CORONA imagery was used frequently in LULC research, particularly for forest mapping (Shahbandeh et al. 2023), the literature paid scant attention to evaluating applications of KH-9 data (PCS and MCS) for historical LULC and urban investigations (Sertel et al. 2024). In this view, Sertel et al. (2024) assessed the accuracies of nine deep learning models for classifying KH-9 PCS images between 1977 and 1980 into eight LULC classes: tree cover, shrub land, grass land, built-up, bare vegetation, permanent water bodies and permanent crop land. This research achieved the highest F1-score accuracy of 0.88 using U-Net++ with an SE-ResNeXt50 backbone. Additionally, this investigation demonstrated that KH-9 imagery holds invaluable potential for a range of applications such as creating historical benchmark LULC data, and monitoring urbanization and environmental changes at FSR. At the same time, deep learning modeling can play a central role in analyzing KH-9 imagery (Sertel et al. 2024). Shahtahmassebi et al. (2024) developed a deep learning framework based on the SegNet model for classifying urban landscapes from KH-9 MCS in 1975. The proposed framework achieved an overall accuracy of 72%, and it established a fine spatial resolution urban landscape map which included seven urban classes: *bright_impervious* surfaces (e.g., roads) *dark_impervious* surfaces (e.g., built-up roofs), *non_impervious* surfaces (e.g., farmlands), water bodies (e.g., rivers), shadow (e.g., topographic shadows), *semi_natural* (e.g., recreational) and forest (e.g., urban forest).

1.2.2 Literature on creating historical stereo anaglyph image and historical DSM using KH data

Stereoscopic capabilities of KH images are immensely useful to create historical FSR stereo products (such as stereo anaglyph imagery, digital elevation models (DEMs), DSMs and orthophotos) which could yield historical FSR stereo benchmark data for many areas such as glaciology (Ghuffar et al. 2023), earthquake hazards (Zhou et al. 2021), archaeology (Watanabe et al. 2017) and forestry (Nita et al. 2018). However, generating stereo products using KH stereo images is challenging due to: (a) limited access to technical parameters (e.g., extrinsics, lens distortion) for the KH stereoscopic cameras and (b) lack of historical GCPs (used for bundle adjustment and image registration) and control points (for accuracy assessment)(Galiatsatos et al. 2008).

Altmaier and Kany (2002) used ERDAS Imagine OrthoBASE Pro to automatically create DSMs of a mountainous region of Morocco with a ground resolution of 1.83 m from CORONA (KH-4B) images in 1972. The generated DSM had a vertical accuracy of about 10 m and planimetric accuracy of about 3 m. To tackle the challenges of GCPs and control points, this research conducted a field survey using a Differential Global Positioning System (DGPS). Regarding the lack of camera parameters, OrthoBASE Pro was implemented which demanded minimal inputs, namely focal length, flight height, pixel size, GCPs and control points. Casana and Cothren (2007) proposed an approach for stereo anaglyph, DEM extraction and orthorectification of CORONA imagery of archaeological landscapes. To reduce the complexity of CORONA imagery, this investigation treated a subset of a CORONA image as a frame camera exposed at a single instant in time due to the long focal length of the CORONA panoramic cameras, instead of using the full frame of CORONA. In this view, the subset image can be oriented to the ground using an approach known as space-resection, and then, utilized to create a DEM via standard forward space-intersection procedures. To derive coordinates and heights for GCPs and control points, this research used SPOT imagery (extracting coordinates) and 90 m SRTM topographic data (extracting heights). Stereo procedures (creating anaglyph imagery, DEM and orthorectification) were undertaken in Leica's Photogrammetry Suite of ERDAS Imagine Stereo Analyst while three dimensional analysis was conducted in ArcGIS and ArcScene.

Galiatsatos et al. (2008) created a high quality DEM from CORONA imagery for archaeological investigations in Syria. Given the local restrictions on field campaigns and the use of DGPS, this research utilized horizontal ground control from IKONOS imagery and vertical ground control from map-based contour lines to create the GCPs and control points required for bundle adjustment and accuracy assessment, respectively. Considering time differences and landscape changes, this research registered CORONA images and map-based contour lines to IKONOS for identifying common points (i.e., GCPs and control points) between these images. The non-metric camera model of ERDAS OrthoBase Pro was adopted because the minimum requirements of this model were an approximate focal length and the scanning resolution used to convert CORONA films to digital images (Galiatsatos et al. 2008).

Bolch et al. (2008) made use of the remote sensing software package Graz (RSG) (an add-on for ERDAS Imagine) which can estimate parameters of CORONA imagery using GCPs. They created a digital terrain model (DTM) based on the estimated parameters for glacier change detection. Maurer and Rupper (2015) developed a fully-automated pipeline for DEM extraction and image orthorectification from KH-9 MCS. The proposed pipeline was implemented in MATABL, and included stereo rectification, stereo matching, estimating relative camera poses, bundle adjustment and triangulation, creating a DEM and image orthorectification. Notably, this research employed the Shuttle Radar Topography Mission (SRTM) global DEM to correct the scale factor and absolute orientation of the point cloud from KH-9 MCS. In addition, this research estimated relative camera poses using the focal length, principle point, image scan resolution and fundamental matrix.

Watanabe et al. (2017) and Nita et al. (2018) employed SfM procedures in AgiSoft for generating DSMs and orthorectified images from CORONA imagery over archaeological sites and forested regions, respectively. For the bundle adjustment step, Watanabe et al. (2017) collected GCPs from Google Earth due to the restriction of conducting GPS measurements in the study region, while Nita et al. (2018) derived the coordinates and heights of GCPs from an aerial image and STRM data, respectively. Dehecq et al. (2020) established an automated framework for analyzing global elevation change analysis from KH-9 MCS. This

framework used integration of image correction procedures for KH-9 MCS scanned images and a stereo reconstruction phase following the open-source NASA Ames Stereo Pipeline. To compensate for the lack of historical GCPs and control points, this research utilized several external DEM products, namely SRTM, 90 m TanDEM-X global DEM and SwissAlti3D DEM. Zhou et al. (2021) proposed patch-based orthorectification using KH-9 PCS imagery to determine earthquake deformation. This research treated each tile of KH-9 PCS as an individual frame photograph that had an imaginary optical center instead of mosaicking the patches into a single image, and then, determined the rigorous optical center and camera model. Moreover, coordinates and heights of GCPs were collected from Sentinel-2 (10 m panchromatic) and ALOS PRISM (30 m) images, respectively.

Ghuffar et al. (2022) proposed an automated CORONA stereo pipeline (CoSP) for establishing a DEM from this stereo imagery for monitoring glacier elevation changes over large areas between 1962 and 1972. The CoSP consisted of four parts: (a) a feature matcher model called SuperPoint+SuperGlue (SPSG) in the Python programming environment to match feature points automatically between a CORONA image and contemporary remotely sensed imagery, thus, deriving GCPs, (b) camera modelling and bundle adjustment in MATLAB, (c) establishing dense matching, tie extraction and epipolar resampling in MicMac and (d) DEM coregistration in the LSM module of OPALS. The coordinates and heights of the extracted GCPs were derived from PlanetScope imagery and SRTM, respectively. Moreover, Ghuffar et al. (2023) utilized CoSP in creating the DEM from the KH-9 PCS during 1971-1984 for glacier investigations. In this research, the coordinates and heights of the GCPs were derived from 30 m Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and 30 m ALOS World DEM (AW3D30) AW3D30, respectively.

1.2.3 Summary of related work

In summary, promising results have been obtained in LULC and urban landscape classification from KH data using deep learning models. In addition, KH data provide stereo-pairs imagery which could create important historical databases (e.g., creating DSM, anaglyph imagery and DEM) for a range of applications. Moreover, the above literature review showed that (a) research to assess the efficiency of KH-9 data (either PCS and MCS) for historical urban investigations is largely missing; (b) most of the aforementioned methods were applied to CORONA images rather than KH-9 images; and (c) lack of further evaluation of deep learning models, ablation studies, establishing historical GCPs and control points pipelines and creating DSMs and anaglyph imagery could limit the future use of KH-9 data in LULC research, particularly in urban investigations. These three points motivated this research.

2. Study area and datasets

2.1. Study area

Hangzhou City (29°11'–30°33' N, 118°21'–120°30' E), capital of Zhejiang Province, China was chosen as the study site (Fig.1). Hangzhou is one of the most prosperous and historical megacities in China and has witnessed rapid urbanization. The initial reason for selecting this study area was to test the proposed methodology on highly heterogeneous urban landscape categories as it is composed of urban forested regions, water bodies, semi-natural covers (e.g., recreational regions) and impervious surfaces. Moreover, sufficient data were available with which to implement the research and support our results.

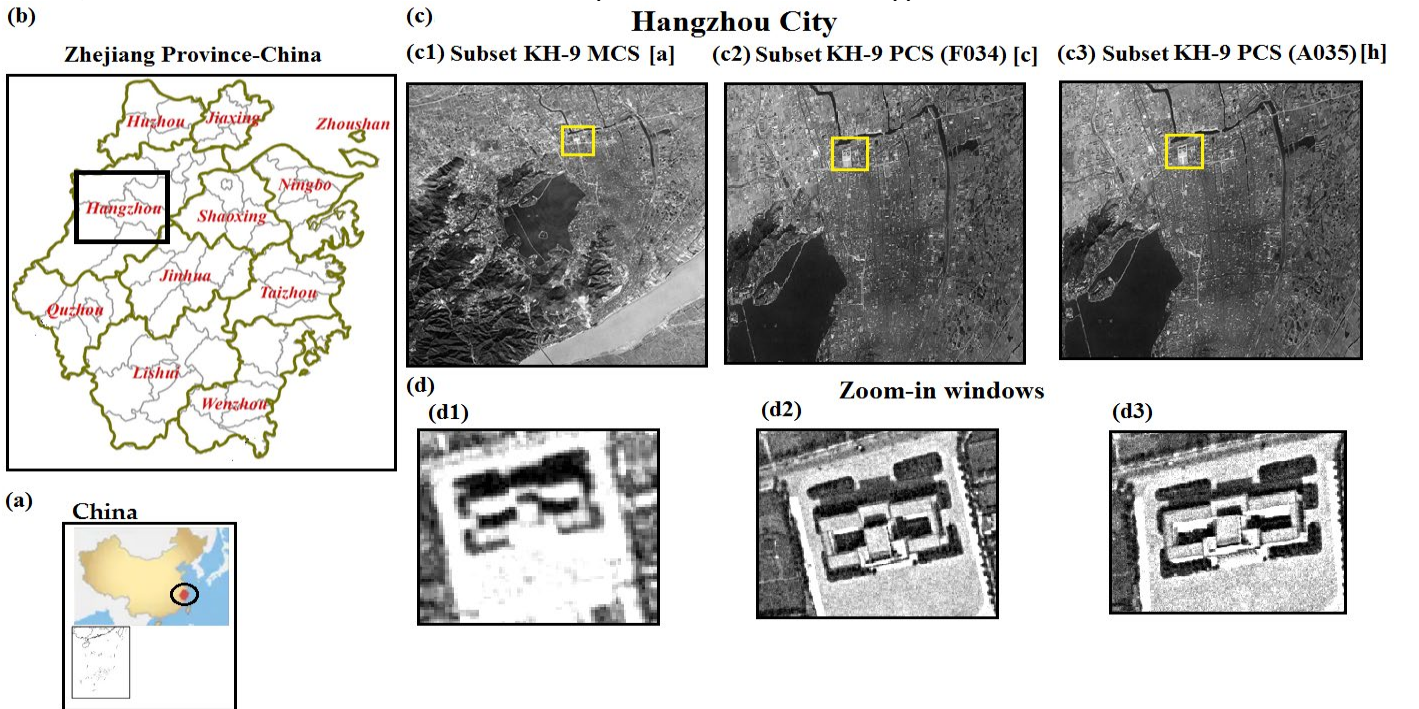


Fig.1 Study area, location of Hangzhou City on (c1) KH-9 MCS [a]: ‘a’ represents the tile; (c2) KH-9 PCS (F034) [c]: ‘F’ and ‘c’ represent the forward camera and tile, respectively and (c3) KH-9 PCS (A035) [h]: ‘A’ and ‘h’ represent backward (aft) camera and tile, respectively. Zoomed-in windows d1, d2 and d3 show a built-up region. The coordinate centers of Hangzhou City is 30°17'16.94"N, 120°8'29.36"E and Zoom-in windows is 30°17'16.94"N, 120°8'29.36"E.

2.2 Datasets and data preparation

A full frame of KH-9 MCS (including two tiles in compressed format) and two full frames of KH-9 PCS (including ten tiles in compressed format) were downloaded from the ‘Declass 2’ and ‘Declass 3’ groups of EarthExplorer (<https://earthexplorer.usgs.gov>), respectively (Table 1). The original KH-9 photographs were recorded on long films, each of which were scanned and divided into two tiles for MCS and ten tiles for PCS (labeled alphabetically) (Zhou et al. 2021). Both the KH-9 MCS and KH-9 PCS frames were produced by high performance photogrammetric film scanners at 7 microns (3,600 dpi) by the United States Geological Survey (USGS) Earth Resources Observation and Science (EROS) Archive (Declassified_Data_2 2018, Declassified_Data_3 2018). Some technical information on scanning procedures and its resolutions can be found in Declassified_Data_3 (2018), Zhou et al. (2021), Hammer et al. (2022), Marzolf et al. (2022) and Ghuffar et al. (2023). For generating ground truth and accuracy assessment, Hangzhou City Maps (1975) (see supplementary: 1. Datasets), Google Earth™ (2023) and a FSR aerial image (2023) were used.

Table 1. Summary of basic properties of the KH-9 imagery used.

Camera	Date	Frame	mission	tile	Spatial resolution(m)	Entity ID
MCS	December 18,1975	7	211-5	a	6-9 m	DZB1211-500049L007001
PCS (Right image)	November 21,1974	35	1209-2	h	0.6-1.2 m	D3C1209-200157A035
PCS (Left image)	November 21,1974	34	1209-2	c	0.6-1.2 m	D3C1209-200157F034

A close visual inspection showed that Hangzhou City was located on the tile ‘a’ of KH-9 MCS while tiles ‘h’ and ‘c’ of KH-9 PCS covered this city. The study site (covered about 100km²) was clipped from the KH-9 MCS and PCS tiles using its administrative boundary on 1975 map of Hangzhou City due to the availability of historical reference data only for Hangzhou city (Fig.1(c1), Fig.1(c2) and Fig.1(c3)). The subset images were then used for urban landscape classification and stereo procedures.

The stereo procedures (e.g., establishing a DSM) require a full frame of KH-9 PCS (ten tiles) to estimate the camera parameters. As mentioned above, a full frame of KH-9 PCS includes ten tiles. Therefore, the first step in stereo procedures is to mosaic these images so as to establish the full frame of KH-9 PCS. However, these images were contaminated by errors (as mentioned in the Introduction), thus, mosaicking such images could amplify the errors. Casana and Cothren (2007) and Zhou et al. (2021) suggested that, for CORONA and KH-9 PCS images, respectively, small portions of the full frame can be treated as a frame camera. The reason could lie in the long focal length of some KH cameras such that the subset image may be treated as a frame camera exposed at a single instant in time (Casana and Cothren 2007). For example, the focal lengths of CORONA and KH-9 PCS are 609.6 mm and 1,523.314 mm, respectively (Marzolf et al. 2022). Additionally, Zhou et al. (2021) demonstrated that GCPs can be used to estimate the “equivalent orientation parameters” of each tile of a full frame of KH-9 PCS though the estimated equivalent model is dissimilar to the actual model, it yields a more accurate outcome (e.g., orthoimage) as it is less affected by film distortions. Ultimately, the subsets of KH-9 PCS images (tiles h and c) over Hangzhou city were used for stereo procedures (Table 1).

A wavelet transform and multi-resolution Top-hat filter pipeline was applied to the subset KH-9 MCS image for de-noising and contrast enhancement, respectively (Shahtahmassebi et al. 2023). The subsets of KH-9 PCS tile ‘h’ and tile ‘c’ were assumed as the right and left images, respectively, for stereo procedures. These images were not subject to de-noising and contrast enhancement so as to assess the effects of stereo anaglyph procedures on their qualities. Additionally, those images were used for creating DSMs without using the aforementioned pre-processing techniques as suggested by Shahtahmassebi et al. (2024).

3. Methods

The overall methodology is divided into two parts: KH-9 MCS and PCS analytical frameworks (Fig. 2). A deep learning classification framework was designed to classify the historical urban landscape from KH-9 MCS imagery. Also this research assessed the impact of the hyperparameters of deep learning on the historical urban landscape classification. A detailed ablation study was conducted to examine the limitations and effectiveness of the proposed method. With respect to KH-9 PCS, a pipeline was developed for establishing historical urban GCPs. In addition, this investigation developed frameworks to create a stereo anaglyph image and DSM with minimal technical information. A thorough ablation investigation was performed to assess the performance of the proposed techniques.

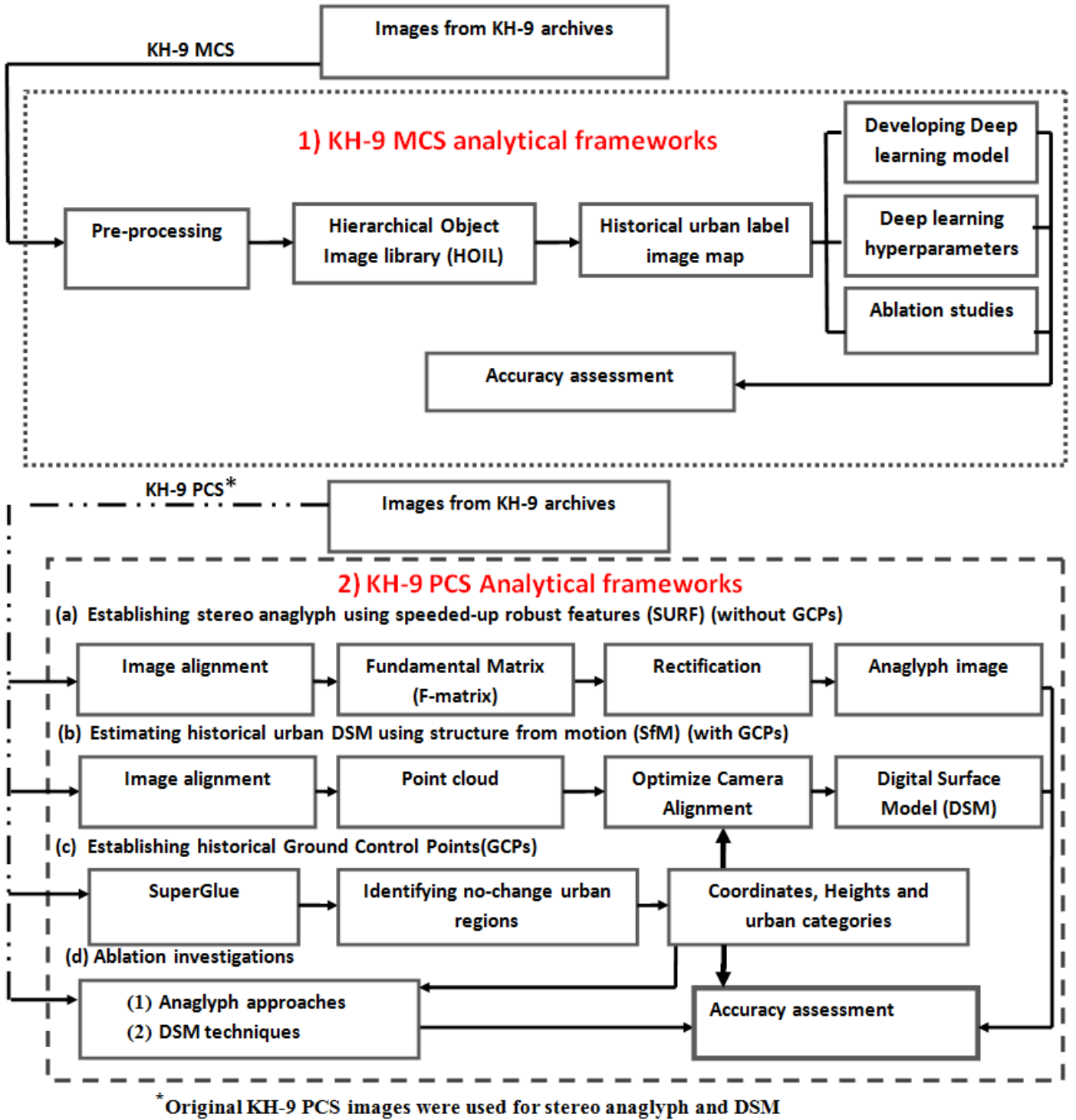


Fig.2 KH-9 PCS and KH-9 MCS analytical frameworks

3.1. KH-9 MCS analytical framework

3.1.1. Establishing a deep learning classification system

The deep learning urban landscape classification framework in this research comprises the following steps: (a) image segmentation- Hierarchical Object Image Library (HOIL) layers, (b) generating historical urban label image map, (c) deep learning models and hyper parameters, (e) ablation investigations and (f) accuracy assessment.

(a) Image segmentation- Hierarchical Object Image Library (HOIL) layers

The chief challenges associated with urban landscape classification from a single fine spatial resolution panchromatic image arise primarily from the fine spatial resolution and lack of spectral bands (Pacifi et al. 2009). To address these challenges, image segmentation can be used to segment such imagery into image-objects and the properties of image-objects (e.g., shape, texture, contextual, geometry and neighborhood) can be extracted to compensate for the lack of spectral bands (Zhang et al. 2018). Image segmentation can be done by various algorithms such as watershed, multiresolution and mean shift (Hossain and Chen 2019). To capture fine-scale characteristics of fine spatial resolution remotely sensed images, image segmentation techniques can be applied at multiple scales ranging from fine- to coarse- scales (Blaschke 2010). Take, for example, segmenting impervious surfaces from fine spatial resolution panchromatic imagery. Fine-scale segmentation of impervious surface pixels on such imagery generates sub-objects which include single buildings, roads and parking lots while coarse-scale segmentation divides pixels in the same image into two major groups or super-objects: impervious surfaces and non-impervious surfaces (Benz et al. 2004, Karakiş et al. 2006, Laliberte et al. 2007, Belgiu and Drăguţ 2014, eCognition_Reference_Book 2014, eCognition_User_Guide 2014). A range of segmented images from fine- to coarse- scales can be generated by setting different scale values. The generated segmented images (which include image objects) are placed into a library which can be used for object-based image classification (e.g., Geiss et al. 2016, Hussain and Shan 2016, Zhang et al. 2018).

No matter which of the segmentation methods is selected, establishing image segmentation at multiple scales can be done by two frameworks: (a) Independent Image Segmentation (IIS) (e.g., Zhang and Xie 2012, Zhang et al. 2018) and (b) Hierarchical Image Segmentation (HIS) (e.g., Hofmann 2001, Haya et al. 2003, Benz et al. 2004, Laliberte et al. 2004, Karakiş et al. 2006, Laliberte et al. 2007, Duro et al. 2012, Xiao et al. 2018).

For IIS: First, the segmentation technique (with fine-scale value) is applied to the target image (pixel-level) to generate a fine-scale segmented image (object-level) which includes sub-objects. Then, the segmentation approach (with coarse-scale value) is applied to the same image (pixel-level) to produce a coarse-scale segmented image (object-level) which consists of super-objects (e.g., Zhang and Xie 2012, Zhang et al. 2018). Within-object features can also be derived such as textures, shape, geometry, spectral and neighborhood from segmented images following the IIS framework (Zhang and Xie 2012). However, between-object features (e.g., distance, connectivity) cannot readily be extracted from such segmented images as no hierarchical relationship exists between the fine-scale (sub-objects) and coarse-scale (super-object) segmented images.

For HIS: First, the target image (per-pixel) is subjected to the segmentation technique (with fine-scale value) to produce a fine-scale segmented image (object-level) that includes sub-objects. Next, the generated fine-scale segmented image (object-level) is subjected to the same segmentation technique (with coarse-scale value) to generate a coarse-scale segmented image which includes super-objects. Given that the coarse-scale segmented image is created based on the fine-scale segmented image, there exists a hierarchical relationship between the segmented images in which super-objects in the coarse-scale segmented image include multiple sub-objects at the fine-scale. For example, forested regions (a super-object) in the coarse-scale segmented image may be composed of trees species (sub-objects) in the fine-scale segmented image. Therefore, HIS not only provides segmented images and their corresponding within-object features, but one can also extract between-object features (e.g., contiguity and distances) in terms of the relationship between the fine-scale (including: sub-objects) and coarse-scale (including: super-objects) segmented images (e.g., Hofmann 2001, Haya et al. 2003, Benz et al. 2004, Laliberte et al. 2004, Karakiş et al. 2006, Laliberte et al. 2007, Duro et al. 2012, eCognition_Reference_Book 2014, eCognition_User_Guide 2014, Xiao et al. 2018). Given the advantages of HIS, this framework was adopted for KH-9 MCS. This research used the eCognition® Developer 9.0.1 software (hereafter eCognition) to implement image segmentation. The proposed image segmentation comprises the following steps: Step1, establish HIS; Step 2, generate image-objects features; Step3, create Hierarchical Object Image Library (HOIL) layers. The details of these steps are as follows:

Step 1, establishing HIS

The multiresolution segmentation algorithm in eCognition was employed to delineate the image objects in the KH-9 MCS image (eCognition_Reference_Book 2014). This algorithm is based on the “bottom-up” image segmentation technique that begins with delineating one-pixel image objects, and merges neighboring objects together until a homogeneity threshold is satisfied (Baatz and Schäpe 2000, Benz et al. 2004, Belgiu and Drăguţ 2014, eCognition_Reference_Book 2014). The homogeneity threshold is determined by three user-defined parameters: scale (controlling the relative size of the image objects), color/shape weight (determining effects of spectral and shape information on segmentation) and smoothness/compactness weight (determining compactness or non-compactness of segmentation). For our study area, seven hierarchical segmented images were generated using the multiresolution segmentation technique based on seven different scale parameters, color/shape weights and smoothness/compactness weight (see Supplement 2: Hierarchical Object Image Library (HOIL), Table 1). Considering each segmented image was generated using a former segmented image, the seven generated segmented images inherited hierarchical characteristics. In these seven hierarchical segmented images, the scale ranged from 15 (fine scale) to 150 (coarse scale) while color/shape weights were set to decline gradually so that shape information was taken into account for segmentation. Moreover, smoothness/compactness was set to decrease gradually in favor of compactness. Visual comparison between the original KH-9 MCS and seven hierarchical segmented images showed that segmented image No.4 (scale: 60, shape/colour:0.2 and smoothness/compactness: 0.5) and segmented image No.6 (scale: 90, shape/colour:0.1 and smoothness/compactness: 0.3) captured appropriately fine-(sub-objects) and coarse-scale (super-objects) characteristics of urban landscapes in the KH-9 MCS image, respectively. Therefore, these two segmented images were selected for the next step.

Step 2, generating image-object features

Image-object features were generated for the selected segmented images (i.e., No.4 and No.6). The image-object features were divided into two groups: within-object and between-object features. Based on visual comparison between the segmented images (No.4 and No.6) and original KH-9 MCS images, 10 image-object features were extracted from segmented image No.4 and five image-object features were selected from segmented image No.6. The 10 image-object features of segmented image No.4 were: (a) the two between-object features: spatial hierarchy (super object relationship) and topological relationship (mean difference neighborhood); (b) the eight within-object information features: four shape features (shape index, roundness, compactness and rectangular fit), one positional feature (maximum latitude), and three GLCM measures (entropy, correlation, dissimilarity).

The five image-object features of the segmented image No.6 were: (a) the four within-object features: one positional feature (maximum latitude) and three GLCM measures (entropy, correlation, dissimilarity); and (b) one between-object features: topological relationship (mean difference neighborhood).

In total, 15 image-object features were established. These features were computed by image-object modules in the eCognition software. Detailed descriptions of these features can be found in the eCognition software reference book (eCognition_Reference_Book 2014).

Step3, creating Hierarchical Object Image Library (HOIL) layers

Finally, the two HIS segmented images (No.4 and No.6) and 15 image-object features were inserted into a new multi-band file which included 17 images in total. Considering that this file included image objects, their object properties and hierarchical characteristics, the created file was named as Hierarchical Object Image Library (HOIL) layers (hereafter HOIL). To create the HOIL, HIS segmented images (No.4 and No.6) and 15 image-object features were imported from the eCognition software to ENVI® V.5.5 (hereafter, ENVI). These files were then placed in a “Save File As-ENVI Standard” module as a new multi-band file with “img” format (ENVI standard format). Later, the created file was converted to “TIFF” to use in Python and Matlab.

(b) Generating a historical urban label image map

Generating an image label map (training samples) is a crucial step in supervised classification using either deep learning or conventional techniques. Such maps are used to train classification techniques (Mather and Koch 2011, Jensen 2015, Ma et al. 2019, Yuan et al. 2021). However, obtaining the necessary large number of historical training samples, particularly over urban landscape labels in KH-9 MCS imagery, or other legacy remotely sensed imagery is a challenging task (Mboga et al. 2020). This is because most landscapes, particularly urban regions, have been affected by changes, thus, making it very difficult to identify accurate landscape labels. Labeling processes can be performed in two ways: manual digitizing (Deshpande et al. 2021) and semi/fully automated approaches (Mboga et al. 2020, Kirillov et al. 2023, Sertel et al. 2024). While visual manual digitizing is one of the most accurate approaches, it can be time-consuming, difficult to replicate and subjective, whereas semi/fully automated techniques can tackle the limitations of visual manual digitizing (Cheng et al. 2018).

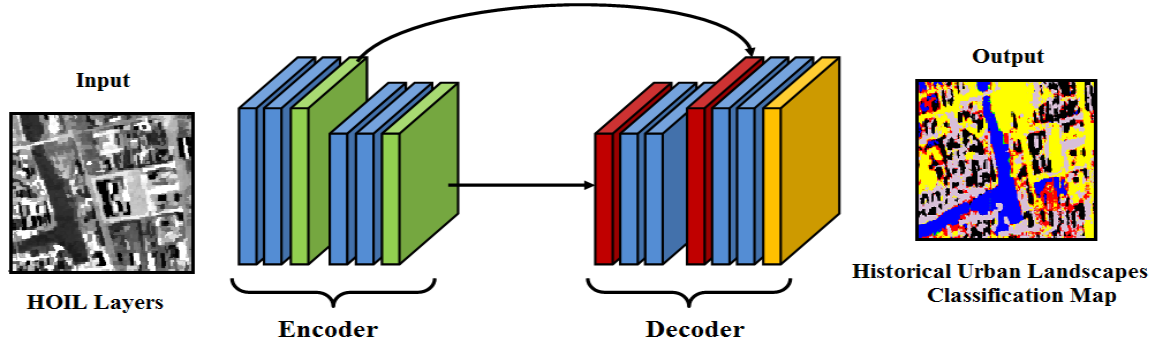
This research utilized a semi-automated supervised image annotation (SASIA) workflow. This workflow included five steps. First, 15 urban classes were identified carefully in Hangzhou City in the KH-9 MCS image through comparison with the 1975 Hangzhou City Map (see supplementary: 1. Datasets, Fig.1). The 15 identified urban classes were dark impervious surfaces (residential areas, business district and rural areas), bright impervious surfaces (railways, roads and parking lots), water bodies (ponds, river networks, lakes), forested areas (natural forest and urban forest), shadow (built-up shadow and topographic shadow), semi-natural areas (urban green spaces) and non-impervious surfaces (farmlands). Second, the characteristics of 15 identified urban classes in HOIL were scrutinized. The result showed that the segmented images (No.4 and 6), and their corresponding object-features in HOIL, highlighted appropriate characteristics (e.g., geometry and textures) of the identified urban classes. Third, the identified urban classes were then used to train the SVM classifier to annotate urban features in HOIL. The proposed SVM classifier used a radial basis function (RBF) kernel, with maximum iteration (5000) and cost parameter (500). Fourth, the initial accuracy of the generated label image map was 80%. Given the importance of annotated urban landscape labels for deep learning models and accuracy assessment, the generated map was revised using the ‘Interactive Class Tool’ option in the ENVI (ENVI® V.5.5) software, and with the aid of the 1975 Hangzhou City map. The final overall accuracy of this map was thereby increased to 95%. Fifth, the obtained historical urban landscape labels map and HOIL were cropped separately into image patches with a size of 100×100 pixels to training, validation and testing the deep learning model. 543, 186 and 81 patches were chosen randomly for training, validation and testing, respectively. The proposed historical urban landscape labels map also served as ground truth labels in training benchmark classifiers and accuracy assessment steps.

(c) Deep learning models and hyperparameters

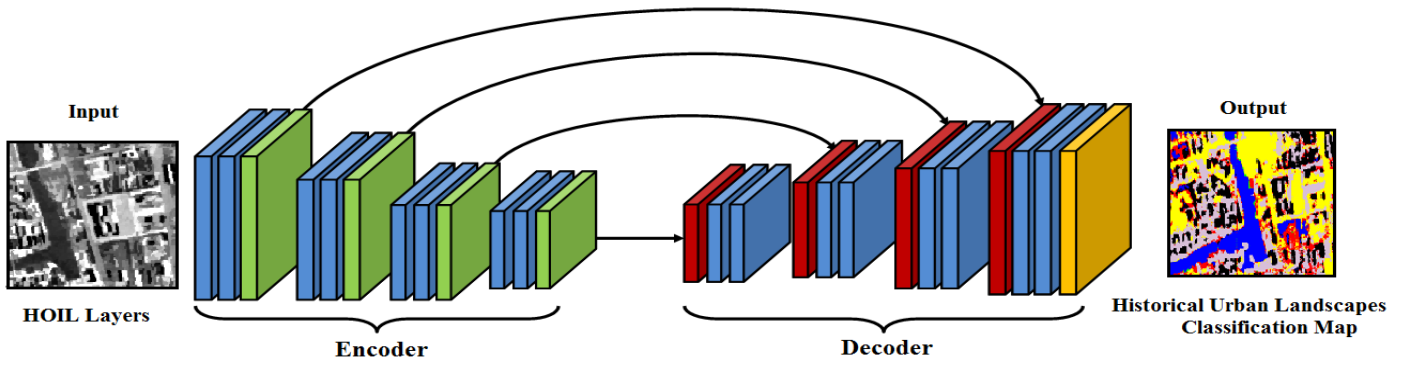
Considering the advantages of deep learning models for historical aerial photos in general (Mboga et al. 2020) and KH-9 data (and other KH family data) in particular (Deshpande et al. 2021, Sertel et al. 2022) as well as the paucity of documented research assessing the impacts of deep learning hyperparameters on the performance of urban classification from KH-9 MCS imagery, a supervised deep learning model architecture using the HOIL (17 layers) was proposed. The selected deep learning model was SegNet based on a directed acyclic graph (DAG) network, which utilizes cut connections for increased efficiency (Badrinarayanan et al. 2017). The current research focused on two hyperparameters, encoder depth (EnD) and minibatch (MB) size while other hyperparameters were held constant. As such, this framework is abbreviated to SegNet_EnD(value)_MB(value). This research examined a range of encoder depths (EnD) including 2 (Fig.3 (a)), 4 (Fig.3 (b)), and 6 (Fig.3(c)). This investigation also scrutinized a range of minibatch (MB) sizes: 16, 32, 64 and 128. It is noteworthy that the SegNet model employs an encoder network to extract depth information, thus, enhancing the performance of semantic segmentation (Badrinarayanan et al. 2017). Minibatch size

can also impact the performance of the deep learning model. The hyperparameters for training the proposed SegNet model were HOIL (17 layers), urban classes (15), a learning rate of 0.05, solver for training the neural network (stochastic gradient descent with momentum (SGDM)), momentum (0.9), weight decay (0.0001), mini-batch (MB) size (16, 32, 64, 128), encoder depth (EnD) (2, 4, 6) and maximum number of epochs (25) (see supplement: 3.1Hyperparameters, Table 2). The cross-entropy was adopted as the loss function throughout this research. Image augmentation was not implemented in this research. Class weighting based on the inverse frequency weighting was applied in the pixel layer classification due to the imbalanced nature of sparse urban land classes.

(a) SegNet Architecture, Encoder Depth=2



(b) SegNet Architecture, Encoder Depth=4



(c) SegNet Architecture, Encoder Depth=6

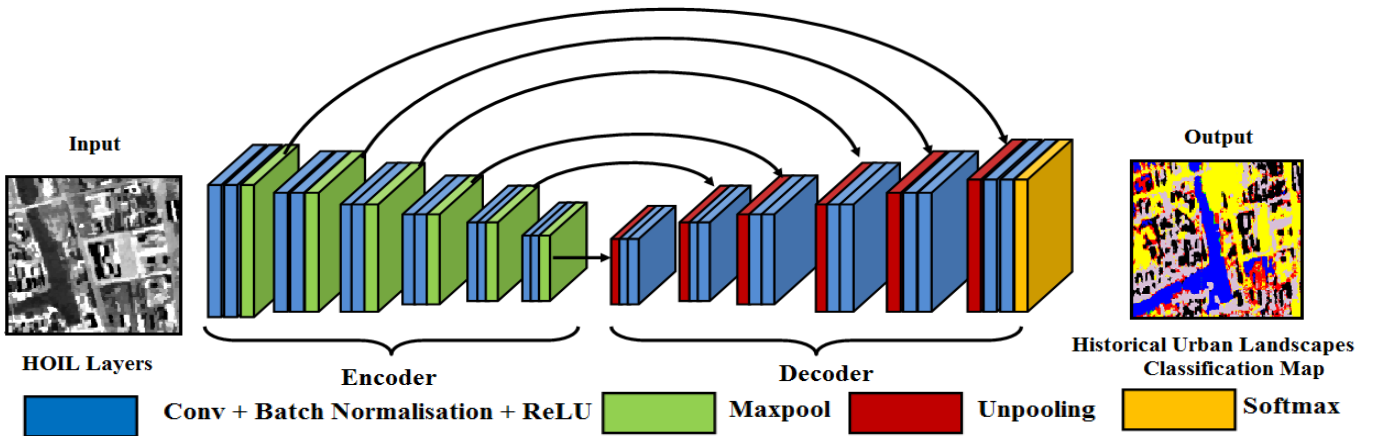


Fig.3 An illustration of the SegNet architecture. (a) SegNet Architecture, Encoder Depth=2 (Total layers:31 deep learning layers; Covolutional layers: 4 encoder layers and 4 decoder layers) (b) SegNet Architecture, Encoder Depth=4 (Total layers: 59 deep learning layers; Covolutional layers: 8 encoder layers and 8 decoder layers), (c) SegNet Architecture, Encoder Depth=6 (Total layers: 87 deep learning layers; Covolutional layers: 12 encoder layers and 12 decoder layers). Convolutional (Conv); Rectified Linear Unit (ReLU); Details of Convolutional layers can be found in Matlab® (MathWorks Inc., Natick, MA, USA. Release: 2022b)

(d) Ablation investigations

To provide a systematic and in-depth analysis, performance of proposed deep learning SegNet framework was evaluated against six existing benchmark machine learning methods, namely the linear support vector classifier (LinearSVC) which employs penalization and loss function, multi-layer perceptron (MLP), random forest (RF), logistic regression (LR), SVM and majority voting (MV) classifier (Pedregosa et al. 2011).

The LinearSVC was constructed by fixing the amount of regularization ($C=100$), penalty (11) and number of iterations (15000). The MLP included hidden layer sizes (100), activation ('relu'), solver (stochastic gradient descent (SGD)), initial learning_rate (0.001), max_iteration (10000) and batch_size (50). The LR consisted of penalty (12), solver ('saga') and max_iteration (15000). The RBF kernel and max_iteration (15000) was used for the SVM. The RF included the number of estimators (100) and minimum samples leaf (2). MV accommodated the aforementioned classifiers. For the sake of brevity only the most important parameters of the classifiers were mentioned here, while the rest were moved into the supplementary materials (see supplement: 3.2 Machine learning). The training the benchmark classifiers were established based on the generated historical urban landscape labels map in the deep learning classifier step (Subsection 3.1.1, step (b): Generating historical urban label image map). The generated map was subjected to stratified random sampling with 70% for training and 30% for the validation of model. The collected samples were converted to points of interest for training the above classifiers. Careful visual inspection was conducted to assess the quality of the sampling points. The results showed that the selected points were sufficiently homogeneous, dense and adequate which covered all classes.

(e) Accuracy assessment

Lack of sufficient ground reference data is a chief challenge to assessing the accuracy of generated maps from KH-9 imagery or other historical remotely sensed images. To tackle this problem, this research used the generated label image map for the accuracy assessment (Subsection 3.1.1, step (b): Generating historical urban label image map). Intentionally, the entire label map (rather than random sampling) was used partly to ensure sufficient image objects for accuracy assessment and partly to ensure that the boundary of image objects could be matched with corresponding regions on the produced urban landscape maps. Overall accuracy (OA), precision (PR) and recall (RC) measures were implemented for accuracy assessment. Moreover, Jaccard similarity metrics including Jaccard Micro Index (JDMI), Jaccard Weighted Index (JDWI) and Jaccard Macro Index (JDMAI) were adopted to gauge the degree of matching between two objects. For the sake of generalization, classified urban features produced by the deep learning models and machine learning techniques were accommodated into the seven major categories. These are: 'Bright_impervious surfaces (Bright_IS)' - e.g., parking lot, 'Dark_impervious surface (Dark_IS)' - e.g., residential areas, 'Non_impervious surfaces (Non_IS)' - e.g., vegetation covers such as agricultural regions, 'Water' - e.g., ponds, lakes, rivers, 'Shadow' - e.g., urban and topographic shadows, 'Semi_natural' - e.g., urban green spaces, and 'Forest' - e.g., forested regions.

3.2. KH-9 PCS analytical framework

3.2.1. Developing historical urban ground control points (GCPs) and check points

To estimate the exterior orientations of KH-9 PCS stereoscopic cameras, given the absence of their meta parameters, GCPs must be accurate and precise. In addition, very accurate control points are needed to assess the accuracy of stereo products such as DSMs. However, acquiring historical GCPs and control points in urban regions which include coordinates, heights and types of urban features is a major challenge particularly given rapid LULCCs. Although wall-to-wall visual interpretation of earlier and present remotely sensed imagery can be used to identify no-change urban regions, this procedure is time consuming and relies heavily on experience and the knowledge of the photo interpreter. Automatic feature matching such as scale-invariant feature transform (SIFT) or other feature matching algorithms can be utilized for such tasks yet their application can be challenging because of LULCCs in the scene and technical differences between current and earlier remotely sensed imagery. Applications of deep learning models for feature matching have suggested promising outcomes in the context of matching historical remotely sensed imagery (Zhang et al. 2021, Farella et al. 2022). Recently, SuperPoint+SuperGlue (hereafter, SPSG) models, which are based on deep learning models, have been used for matching historical aerial photographs (Zhang et al. 2021, Ghuffar et al. 2022, Maiwald et al. 2023).

In this view, the SPSG model was tested for matching CORONA (one of the KH satellite missions) and contemporary satellite sensor imagery over mountainous regions, thus, identifying GCP matches between them (Ghuffar et al. 2022). The SPSG includes the SuperPoint step, a self-supervised framework for point detectors and descriptors suitable for a large number of multiple-view geometry problems (DeTone et al. 2018). Next, SuperGlue matches two sets of local features by jointly finding correspondences and rejecting non-matchable points (Sarlin et al. 2020). SuperPoint uses a fully-convolutional neural network while SuperGlue employs an attentional graph neural network and optimal matching layer. SuperPoint can also identify feature points under illumination changes (DeTone et al. 2018, Sarlin et al. 2020).

Detailed mathematical descriptions of SuperPoint and SuperGlue can be found in DeTone et al. (2018) and Sarlin et al. (2020), respectively. As mentioned, this framework was tested on CORONA imagery over mountainous regions (Ghuffar et al. 2022), with less assessment of its suitability for application to urban regions. Our study extends SPSG to KH-9 PCS imagery in 1974 and contemporary FSR aerial imagery in 2023. Herein, the research contributes on two important aspects: (1) SPSG was tested over an urban region which witnessed unprecedented urban development and (2) a framework was proposed to extract geographic coordinates, determine types and estimate the heights of buildings. The proposed historical urban GCP process is as follows:

- (a) Patch size: In general, SPSG can be trained on relatively small patch size (default patch size is 640×480 pixels). Our empirical investigation showed that this patch size yielded reasonable outcomes. The subset of KH-9 PCS images had a

size of 8000×6000 pixels which may slow down conversion of this size to the default of SPSG. Thus, this image and corresponding urban landscapes geographic locations from reference FRS aerial imagery in 2023 (spatial resolution 1.6 m) were divided into patches. For both the KH-9 PCS and FSR aerial images, a range of patch sizes with respect to similar geographic location were tested, including: 2000×2000 , 1000×1000 , 1920, 1440, 2146×2225 and 1562×1445 pixels. A patch size of 2146×2225 pixels for the FSR aerial image and a patch size of 1562×1445 pixels were found to be appropriate for feature matching over urban regions. Later, these patch sizes were converted automatically into the default size (640×480) by SPSG.

- (b) Parameter setting: Our research used predefined parameters and original weights of the SPSG as tested by Sarlin et al. (2020) and Ghuffar et al. (2022).
- (c) Accuracy of identified points: SPSG identified a large number of match points between the KH-9 PCS and FSR aerial imagery. These points were extracted using a Python script provided by Sarlin et al. (2020). Among these, nearly 100 points were selected where the points were located in no-change features or close to those objects in corresponding locations in KH-9 PCS and FSR aerial images. Of these 100 points, 50 points located in no-change features were selected manually. Given that these points were located in no-change regions, their spatial accuracies were assumed to be 100%. However, this assumption should be approached with caution due to the intrinsic limitation imposed by the pixel size.
- (d) Extracting geographic coordinates and estimating the heights of no-change urban regions: It was difficult to use Real-Time Kinematic (RTK) Global Navigation Satellite System (GNSS) measurements for estimating latitude and longitude in our study area because of restrictions. Acquiring very accurate coordinates and survey data (e.g., historical base maps) was also challenging. To tackle this limitation, previous research derived coordinates of historical GCPs from contemporary remotely sensed data (e.g., Casana and Cothren 2007, Galiatsatos et al. 2008, Watanabe et al. 2017, Nita et al. 2018). A similar approach was applied to extract the coordinates of historical GCPs for KH-9 PCS in urban regions from FSR aerial imagery in 2023. First, the color composite FSR 2023 aerial imagery with a spatial resolution of 0.6 m was utilized. The FSR aerial imagery was geometrically corrected previously. The reported accuracy was 2 m root mean square error (RMSE) while other details such as flight path, geometric correction techniques and camera parameters were not available. Then the ground latitude and longitude for 50 identified no-change points in urban regions were derived from this FSR aerial imagery.
- (e) Estimating built-up height: Similar to the problem of acquiring RTK GNSS data, another obstacle was obtaining accurate built-up heights either from estimated DSMs by Light Detection and Ranging (LiDAR) or optical stereo sensors. Moreover, obtaining such data for nearly 50 years ago exacerbated this limitation. A framework based on Baidu street view imagery was adopted to minimize this limitation. Its performance was tested in previous research (Cao and Huang 2021, Cai et al. 2023). The geographic locations of 50 no-change GCPs in urban regions were transformed to Baidu street view images (<http://map.baidu.com>, access 6 July 2024). Then, the number of floors of each building located in both the KH-9 PCS image and Baidu map was identified manually. The number of floors was converted to building height assuming the height of each floor is 3 m (Cao and Huang 2021, Cai et al. 2023). Also the built-up height and number of floors data from Open Street Map (<https://download.geofabrik.de/>) in 2019 were used to test the accuracy of the built-up height estimates. To do so, 50 no-change urban points with estimated height and floors based on a proposed Baidu street view images framework (Cao and Huang 2021, Cai et al. 2023) were placed in the corresponding location in Open Street Map shape file. All 50 points were located accurately in open street map. Then corresponding built-up heights and floors were extracted from Open Street Map shape file. The result showed no differences between the counted built-up floors based on the proposed framework (Cao and Huang 2021, Cai et al. 2023) and those in the open street map. The built-up heights showed an average difference about 1.5 m, which satisfied our expected accuracy.
- (f) Determining the type of built-up land: The type of built-up land was identified during estimating the heights of built-up regions from Baidu street view images. Accordingly, the built-up locations were classified as architectural heritage sites as well as old residential and public built-up regions.
- (g) Establishing historical GCPs and check point libraries: The accuracy of GCPs plays a vital role in estimating (i.e., bundle adjustment stage) DSMs while accurate check points are crucial for assessing the accuracy of the DSMs. As mentioned above, obtaining precise GCPs and control points using RTK GNSS (extracting longitudes and latitudes) and LiDAR (deriving heights) was one of the chief challenges of this research. Given that the approximate geometric accuracy of FSR aerial imagery in 2023 (step (e): ~ 2 m) and the estimated built-up heights from Baidu street view image (step (f): ~ 1.5 m) met our requirements, these points were used to establish historical GCPs and control points. The geographic coordinates, heights and types (including their images) of 50 no-change urban points were inserted into a geographic information system (GIS) file. These points were then divided randomly into two groups: 38 points served as GCPs and 22 points were used as check points.

3.2.1.1 Assessing processing time

To compare the processing time of the proposed no-change urban identification using the SPSG procedure and visual interpretation, five experts in remote sensing were trained in analyzing the KH-9 images. The training included into two parts: (a) theoretical aspects which consisted of an introduction to KH-9 PCS imagery (panchromatic), characteristics, challenges and advantages, and (b) experimental aspects composed of feature identification on KH-9 PCS and detection of no-change regions in both KH-9 PCS and contemporary panchromatic FSR aerial imagery. The training stage took nearly 2 months. Afterwards, both Hangzhou city images from the KH-9 PCS image in 1975 and the FSR aerial imagery in 2023 were distributed among five trained experts to assess how long it would take to identify no-change urban regions on both images.

3.2.2. Generating stereo anaglyph image and DSM

3.2.2.1. Creating ultra FSR stereo anaglyph image

(a) Fully automatic SfM-Anaglyph (without GCPs and meta parameters of camera)

A SfM framework was developed to rectify automatically stereo-pair KH-9 PCS images without using GCPs and meta parameters of cameras. This procedure was designed based on Maurer and Rupper (2015)'s research and recent SfM algorithms for stereo uncalibrated cameras in Matlab® (MathWorks Inc., Natick, MA, USA. Release: 2022b). This framework consisted of four steps: First, speeded-up robust features (SURF) was applied to identify and collect keypoints from both stereo images. Second, descriptors (SURF feature vectors) are calculated from regions around each identified point. Those points located close to the border of the image or outside the image were removed from the descriptors. Then matched points between two stereo images were estimated using an "exhaustive" technique that computes the pairwise distance between features vectors in two stereo images. This technique used a feature matching measure, called the sum of absolute difference "SAD", that ensured the selection of high quality matches (see supplement: 4. SURF-matching points). Third, the correctly identified matching points were used to estimate the fundamental matrix (F-matrix) and inliers index (the corresponding selected match points in two stereo images were used to calculate the F-matrix). However, matching points may be contaminated by outliers which could affect the performance of estimating the F-matrix. To do so, random-sample consensus (RANSAC) was used to eliminate outliers that could introduce errors into the F-matrix. Moreover, the 'isEpipoleInImage' algorithm was adopted to examine whether the first stereo image associated with the estimated F-matrix (F) and the second stereo image associated with the estimated F-matrix (F') contained epipole and epipole location. In the case of errors, the feature detecting and matching steps were repeated. If the results satisfied the aforementioned requirements, the inliers index was used to extract separately the coordinates of corresponding points in the two stereo images. Fourth, given the absence of meta parameters (intrinsic and extrinsic) for KH-9 PCS cameras, the 'estimateStereoRectification' approach was adopted to produce a projective transformation to rectify the stereo images using the F-matrix and coordinates of the corresponding points in the two stereo images (e.g., KH-9 PCS-left and KH-9 PCS-right). Then the estimated projective transformations of the stereo images were used in the 'rectifyStereoImages' algorithm to rectify the KH-9 PCS stereo images. Finally, the 'stereoAnaglyph' function was applied to the rectified stereo images to establish a red-cyan anaglyph from the stereo pair. Detailed mathematical description of the aforementioned function can be found in Matlab® (MathWorks Inc., Natick, MA, USA. Release: 2022b) and Maurer and Rupper (2015) .

(b) Ablation investigations: Benchmark methods

1- Fully automatic SuperPoint-SuperGlue (SPSG)-Anaglyph (without GCPs and meta parameters of the camera)

Estimating the F-matrix, reconstructing epipolar imagery, and stereo rectification (as used in the proposed SfM-anaglyph) are necessary steps for establishing anaglyph images. Herein, a benchmark technique that accommodated SPSG, homography matrix (H-matrix) and image warping was devised to assess the performance of the proposed SfM-Anaglyph. As mentioned above, SPSG procedures can offer highly accurate matching points between two stereo images with the help of a deep learning model and graph neural network. So far, less attention has been given to assessing the applicability of SPSG procedures matching points in creating a stereo anaglyph image, particularly stereo-pairs of KH-9 PCS images. The first image in 1974 (right image: D3C1209-200157A035) was considered as a reference image while the second image in 1974 (left image: D3C1209-200157A034) was assumed to be the distorted image (Table 1). These images were then subjected to the SPSG procedures algorithm. A range of patch sizes with respect to similar geographic locations in both images were examined. Those patches were: 2000×2000, 1000×1000, 1920, 1440, 2146 ×2225 and 1562 ×1445 pixels. In addition, entire images (i.e. 8000×6000 pixels) without reducing the size to 640 ×480 pixels (default value of SPSG procedures) were tested. The investigation revealed that the original size yielded reliable results (see supplement: 5. SPSG procedures-matching points). A script was then written in Python to extract the generated accurate stereo matching points and corresponding local descriptors, as suggested by SPSG procedures (Sarlin et al. 2020). The extracted information was next subjected to the homography matrix (H-matrix) to compute the points in one stereo image to the corresponding points in the other image (perspective correction). Additionally, RANSAC was adopted to eliminate outliers that could introduce errors into the H-matrix. Finally, the parameters of the H-matrix were used in image warping to register the left image to the right image. Then those images were subjected to the developed anaglyph function to generate a stereo anaglyph image.

(2) Semi automatic approach using rational polynomial coefficients (RPCs)

Unlike for modern FSR satellite sensor data, rational polynomial coefficients (RPCs) of KH-9 PCS cameras are not available. In addition, most camera models (e.g., panoramic) in photogrammetric software are appropriate for contemporary FSR satellite sensor data. To tackle these problems, some remote sensing and photogrammetric software such as ERDAS Imagine and ENVI have the potential for building RPCs of KH-9 PCS cameras. Previous research examined the performance of ERDAS Imagine for estimating RPCs of CORONA imagery (the earliest of KH camera model), thus, creating anaglyph imagery and DSMs (Casana and Cothren 2007, Galiatsatos et al. 2008). The current research chose ENVI because less attention has been given to assessing its potential for such purposes particularly in the context of KH-9 PCS in urban regions.

The ‘Build RPCs’ module in ENVI® V.5.5 (hereafter ENVI) can be used to compute RPCs of KH-9 PCS cameras. This module includes two pre-processing parts to create sensor geometry: interior orientation (which transforms the pixel coordinate system to the camera coordinate system) and exterior orientation (which determines the position and angular orientation parameters associated with the image) (ENVI_help 2024). In addition, RPCs in this module are computed using a digital photogrammetry technique that uses a collinearity equation to construct sensor geometry, where the object point, perspective center, and image point are all on the same space line (ENVI_help 2024). The technique involves a series of transformations involving pixel, camera, image-space and ground coordinate systems (ENVI_help 2024). Although the ‘Build RPCs’ module has no panoramic camera model, it offers several camera models, among which ‘Digital (Frame Central)’ is the only option to use for KH-9 PCS cameras. In the ‘Digital (Frame Central)’, the minimum requirements are an approximate focal length, principal points, sensor pixel size and GCPs (for building exterior orientation).

The task of reconstructing RPCs for KH-9 PCS cameras using the ‘Build RPCs’ module was divided into three steps. Prior to using the ‘Build RPCs’ module, the first step was to register stereo-pairs of KH-9 PCS images to the FSR aerial image (WGS 84, Zone 51) in 2023 using the collected historical GCPs. This is because it is straightforward to allocate GCPs in the registered images of KH-9 PCS using ‘Build Exterior Orientation’ in the ‘Build RPCs’ module. The root mean square errors (RMSEs) of the registered images were less than 0.77 m. In the second step, approximate values of focal length, principal points and sensor pixel sizes in ‘Build RPCs’ module should be determined. The approximate focal length of KH-9 PCS cameras is 1,523 mm (Zhou et al. 2021, Marzolf et al. 2022, Ghuffar et al. 2023). For constructing a stereo anaglyph image, a range of focal lengths were tested, consisting of: 1500 mm, 1523 mm, 1546 mm, 1569 mm, and 1600 mm. A focal length of 1600 mm was found to be appropriate. The input sensor pixel size was set to 0.007×0.007 mm (7 micron, scanner resolution for converting from analog film to digital image). The principal point was set to the default value (0) as the principle points of KH-9 PCS cameras were not available. In the third step, the ‘Build RPCs’ module offers computing ‘Build Exterior Orientation’ dialog which requires allocating GCPs points in the images. Unlike image registration, the accuracy of each GCP used for estimation of the exterior orientation plays a central role in locating the position of the aerial camera. Hence, the GCPs were located in the stereo KH-9 PCS images with the utmost care, spread evenly across the image and including all four corners. For each image, six GCPs were collected with a root mean square less than 2 m. The collected points were used to compute the exterior orientation parameters (XS, YS, ZS, Omega, Phi and Kappa), along with the rotation angles and rotation systems (see supplement: 6. RPCs-anaglyph). This process was performed separately for each stereo image.

The stereo images with estimated RPCs were used to create a stereo anaglyph image in the ‘Topographic Toolbox’ of ENVI. To do so, the images were subjected initially to the ‘Select Stereo Tie points’ procedure. This research performed manual tie point selection as the automatic procedure did not provide satisfactory results. The maximum ‘Y-parallax’ was less than 2.49. Then the ‘Build Epipolar Images’ module was applied to the stereo images with estimated tie points and RPCs to build the epipolar images. Finally, a stereo anaglyph image was created using the epipolar images in the ‘Stereo 3D Measurement’ module. Herein, it should be noted that inappropriate RPCs could lead to high ‘Y-parallax’ and error in building epipolar images. A high ‘Y-parallax’ hampers selecting tie points while epipolar images contaminated by errors prevent establishing a stereo anaglyph image.

(c) Accuracy assessment of Anaglyph images

Understanding the quality and quantity of the stereo anaglyph image plays an important role in subsequent procedures such as 3-Dimensional measurements. Besides, although previous research showed that de-noising and contrast enhancement can increase the quality of the stereo anaglyph images of KH-9 PCS, it is unclear whether stereo anaglyph procedures (e.g., epipolar estimation, rectification) can also enhance the quality of the original stereo-paired KH-9 PCS images. Therefore, the original stereo-paired KH-9 PCS images were used to assess the degree to which stereo image anaglyph procedures can increase the quality of such images. Additionally, several measures were developed to assess the quality and quantity of stereo generated anaglyph images.

First, an image histogram was adopted to measure image quality before and after stereo anaglyph procedures. Second, normalized mutual information (Pedregosa et al. 2011) and geometric transformations using Matlab® (MathWorks Inc., Natick, MA, USA. Release: 2022b) measures were computed between the two rectified stereo images to quantify the degree of matching. The time to process the stereo anaglyph images was also measured. Thirdly, stereo anaglyph images may generally generate ‘ghosting effect or leaking of red and cyan of one image to another’. To capture this effect, those affected regions were selected as training samples in the generated stereo anaglyph images. Then training samples and the SVM classifier was applied to those images to generate maps of the ‘ghosting effect’ for each image. The proposed SVM classifier used the RBF kernel, penalty parameter (100) and gamma kernel function (0.167). Each map had three classes, red, cyan and normal region, with an overall accuracy of above

90%. Then the percentage of red and cyan classes was computed. Finally, GLCM measures were implemented as they can gauge changes in the spatial structure of image brightness values (BVs) (Hall-Beyer 2017). Two GLCM measures were selected: contrast (indicator of local variations) and dissimilarity (indicator of intensity levels). Given that the stereo anaglyph presents in Red-Green-Blue (RGB) mode include two images and one image is a repetition of the first or second image, the quality of both images were measured together rather than individually. To do so, the layer stack of the two stereo anaglyph images was subject to Principal Components Analysis (PCA). Then the GLCM measures were applied to the first band of PCA. The co-occurrence metrics were applied with a window size of 3×3 pixels. Finally, the standard deviation was computed for each texture.

Two experimental images were generated separately by stacking two identical images (demonstrating perfect matching) and stacking the original left and right stereo-paired images (demonstrating no-matching and ghosting effects).

3.2.2.2 Creating historical ultra FSR urban DSM

(a) SfM-DSM framework (with GCPs)

Establishing a historical ultra-FSR urban DSM (i.e. 1970s) using stereo-pair analog panchromatic images from the KH-9 PCS archives is a non-trivial task due to the problems discussed above. To tackle these problems, Watanabe et al. (2017) proposed a fully automatic SfM photogrammetry workflow in Agisoft Photoscan Pro 1.2 (hereafter, Photoscan) to extract an ultra-fine spatial digital surface model (UFSDSM) over archaeological landscapes from stereo-pairs of CORONA (KH-4A and KH-4B). In addition, Nita et al. (2018) used a SfM photogrammetry workflow in Photoscan to extract UFSDSM over a forested region from stereo-pairs of CORONA images. Our research expands their work by (1) proposing this algorithm in a historical heterogeneous and dense urban landscape from KH-9 PCS, (2) integrating historical GCPs, (3) conducting an ablation investigation and (4) accuracy assessment. The SfM-DSM workflow with GCPs was divided into five steps.

First, the sensor pixel size and focal length (both in mm) were set as 0.007×0.007 and 1600 mm in the camera calibration window of Photoscan, respectively. These values were selected based on the procedure described in Section 3.2.2.1 (b): part 2. It is noteworthy that Photoscan can also reconstruct three-dimensional scene in the absence of focal length and sensor pixel size. To do this, Photoscan assumes the focal length equals 50 mm (35 mm film equivalent) and tries to align the photos in accordance with this assumption (Photoscan Manual 2016). Initially, we implemented this assumption, but discarded it in favour of the solution above (the Section 3.2.2.1 (b): part 2) as the image alignment failed.

Second, the “Align Photos” module in the Photoscan software was used to align automatically the stereo-pair KH-9 PCS images. The setting parameters of “Align Photos” were “Accuracy:high” and “Key point limit: 40000”.

Third, a sparse point cloud was established using the “Build Mesh” module. Then, 30 GCPs were placed evenly on both stereo-pairs of KH-9 PCS images to optimize the camera positions and orientation data (see supplement: 7. SfM-DSM). Next, the RMSE of the located GCPs was checked aiming for an average location error of 10 m for the entire image, among which 15 points were selected that generated a smaller RMSE. If the error was larger than the threshold, the procedure was corrected by revising the number and position of GCPs and re-computing the camera positions and orientation data.

Fourth, a dense point cloud was established using the “Build Dense Cloud” module in the Photoscan software. The parameters of this module were “Quality:high” and “Depth filtering: Mild”. It should be noted that “Depth filtering: Mild” preserves the complexity of urban landscapes, as recommended by the Photoscan Manual (2016).

Finally, the generated dense cloud was subjected to Inverse Distance Weighting (IDW) interpolation in the “Build DEM” module to generate the DSM. In addition, the generated point cloud and DSM were placed into the ENVI® V.5.5 Lidar 3D viewer for visualization purposes. Our experiment was repeated several times to select carefully the hyperparameters of modules in Photoscan for our study area. However, using these exact hyperparameters for other study areas should be approached with caution.

(b) Ablation investigation: RPCs-DSM framework (with GCPs)

To evaluate the performance of the proposed SfM framework, the automatic DSM photogrammetric module in ENVI® V.5.5 was utilized. In this workflow, stereo-paired images KH-9 PCS were registered to an FSR aerial image in 2023 with RMSE less than 0.77 m (projection system: WGS 84, Zone 51). Then the RPCs of the stereo-pair images of KH-9 PCS were estimated similar to the anaglyph procedure based on the five most accurate GCPs with an RMSE less than 2 m (Section 3.2.2.1 (b): part 2). The focal lens (1600 mm) and sensor pixel size (0.007×0.007 mm) were used (Section 3.2.2.1 (b): part 2). The estimated RPCs coefficients (see supplement: 8. RPCs-DSM) and corresponding stereo-pair images were subjected to automatic point cloud module based on the default values. Finally, the generated point cloud and DSM were accommodated in ENVI® V.5.5 Lidar 3D viewer for visualization purposes.

(c) Accuracy assessment of the estimated historical urban DSM

The created check points (Subsection 3.2.1: step (g)) were utilized to assess the accuracy of the generated DSMs. The check points included 22 points distributed evenly across the study areas. Both the positional and height accuracy were assessed using these check points. The measures of accuracy were the root mean square error (RMSE), mean square error (MSE), bias, standard deviation (Std) and coefficient of determination (R-square).

3.3 Application assessment

Prior to developing any urban digital change detection techniques, rigorous visual comparisons between the generated products from KH-9 (MCS and PCS) and contemporary FSR remotely sensed data were conducted to reveal potential types of urban changes as well as their applications.

4. Results

4.1. KH-9 MCS analytical framework: Mapping the historical urban landscape

The results of the obtained maps are described below in two parts: deep learning assessment and ablation investigations. The deep learning assessment presents the impacts of the hyperparameters on urban landscape mapping based on KH-9 MCS imagery while the ablation investigation evaluates the performance of the proposed deep learning technique benchmarked against several state-of-the-art machine learning methods.

(a) Deep learning assessment

This experiment assessed the impacts of two hyper parameters, namely the number of deep learning layers and minibatch size. A desirable classification result was obtained by using the proposed HOIL_SegNet_EnD(2)_MB(16). In terms of visual assessment, careful comparison between the proposed SegNet architecture and benchmark architectural methods (Fig.4 and Fig.5) indicated that the proposed SegNet achieved the smoothest visual outcomes with precise boundaries. In particular, the semantic contents of heterogeneous urban landscapes (e.g., dark impervious surfaces, water bodies, etc) were detected effectively, and the urban shaped features including built-up and street were characterized with high geometric fidelity. For example, urban water bodies (Fig. 4: dark blue colour) were distinguished corrected by the proposed SegNet in comparison with the other deep learning models. In particular, HOIL_SegNet_EnD(6)_MB(16) and HOIL_SegNet_EnD(2)_MB(128) led to high misclassification. Moreover, those techniques generated pixelized distributions for the urban landscape classes. In general, an increase in depth of a deep learning architecture can increase the accuracy of deep learning models (Badrinarayanan et al. 2017, Gao et al. 2024), but our investigation produced contrary outcomes. Due to training sample constraints, increasing the depth of the deep learning network did not generate reliable information (i.e., a feature map) capturing the semantic context of the image, thus leading to a pixelized distribution for the urban landscape classes. However, our research showed that the collected training samples provided sufficient information for a smaller encoder depth (SegNet_EnD(2)) to perform the learning procedure, ultimately, generating reliable classified maps. Also, the results illustrated that smaller minibatch size (MB(16)) led to reliable outcomes. Masters and Luschi (2018) demonstrated that the use of small mini-batch sizes offered more up-to-date gradient computations, resulting in more stable and reliable training. Considering the lack of historical training samples and the paucity of information on the application of deep learning to KH imagery, our observations should be approached with caution. Further research is needed to reach a consensus on selecting the appropriate hyper parameters of deep learning models for KH imagery.

Focusing on accuracy assessment, the HOIL-SegNet overall accuracy (OA) started at around 38% for the urban landscape classifications with encoder depth 6 (HOIL_SegNet_EnD(6)_MB(16)), and increased rapidly to approximately 58% with encoder depth 4 (HOIL_SegNet_EnD(4)_MB(16))(Tables 2). A similar pattern can be observed for minibatch size. As shown in Table 3, the SegNet with minibatch size 128 achieved the lowest OA of 48%, a JDMI of 0.32, a JDMAI of 0.27 and a JDWI of 0.32. However, the proposed SegNet (HOIL_SegNet_EnD(2)_MB(16)) had an overall accuracy of 70%, a JDMI of 0.54, a JDMAI of 0.50 and a JDWI of 0.57(Tabel 2). The PR, RC and F1-score of all urban classes achieved by the proposed SegNet technique were higher than for the other deep learning models. Our results imply that modifications of minibatch size and learning layer depth can affect the performance of deep learning models in the context of urban land cover mapping using KH-9 MCS data with spatial resolution of 6-9 m.

Table 2, Comparison between accuracy of proposed HOIL_SegNet_EnD(2)_MB(16) and a range of encoder depth

Class	Methods								
	HOIL_SegNet_EnD(2)_MB(16)			HOIL_SegNet_EnD(4)_MB(16)			HOIL_SegNet_EnD(6)_MB(16)		
	PR	RC	F1-score	PR	RC	F1-score	PR	RC	F1-score
Bright_IS	0.53	0.57	0.55	0.46	0.46	0.46	0.36	0.28	0.32
Dark_IS	0.48	0.31	0.37	0.27	0.05	0.08	0.21	0.01	0.02
Non_IS	0.82	0.71	0.76	0.64	0.63	0.63	0.61	0.12	0.2
Water	0.82	0.84	0.83	0.59	0.76	0.66	0.47	0.58	0.52
Shadow	0.51	0.75	0.61	0.35	0.58	0.44	0.44	0.33	0.37
Semi_natural	0.46	0.60	0.52	0.35	0.38	0.37	0.27	0.39	0.32
Forest	0.95	0.89	0.92	0.86	0.85	0.85	0.38	0.66	0.48
OA	70%			58%			38%		
JDMI	0.54			0.41			0.23		
JDMAI	0.50			0.36			0.19		
JDWI	0.57			0.43			0.21		
Time-minutes	3.22			6.53			2.23		

Table 3, Comparison between accuracy of proposed HOIL_SegNet_EnD(2)_MB(16) and a range of minibatch size

Class	Methods								
	HOIL_SegNet_EnD(2)_MB(32)			HOIL_SegNet_EnD(2)_MB(64)			HOIL_SegNet_EnD(2)_MB(128)		
	PR	RC	F1-score	PR	RC	F1-score	PR	RC	F1-score
Bright_IS	0.53	0.55	0.54	0.46	0.49	0.47	0.39	0.50	0.44
Dark_IS	0.48	0.21	0.29	0.39	0.07	0.12	0.25	0.03	0.06
Non_IS	0.81	0.68	0.74	0.73	0.54	0.62	0.66	0.21	0.32
Water	0.76	0.85	0.80	0.67	0.84	0.74	0.44	0.81	0.57
Shadow	0.51	0.70	0.59	0.44	0.67	0.53	0.27	0.60	0.37
Semi_natural	0.43	0.62	0.51	0.38	0.58	0.46	0.31	0.38	0.35
Forest	0.95	0.89	0.92	0.92	0.89	0.90	0.83	0.71	0.76
OA	68%			63%			48%		
JDMI	0.52			0.45			0.32		
JDMAI	0.48			0.41			0.27		
JDWI	0.55			0.48			0.32		
Time-minutes	1.35			2.7			2.38		



Fig.4 Comparison between the results of applying the proposed HOIL_SegNet_EnD(2)_MB(16), a range of encoder depth (b1-b3) and a range of minibatch size (c1-c3). (a1) KH-9 MCS image and (a2) Ground Truth Urban Label Map. The coordinate center of image is 30°16'29.84"N,120°9'30.74"E

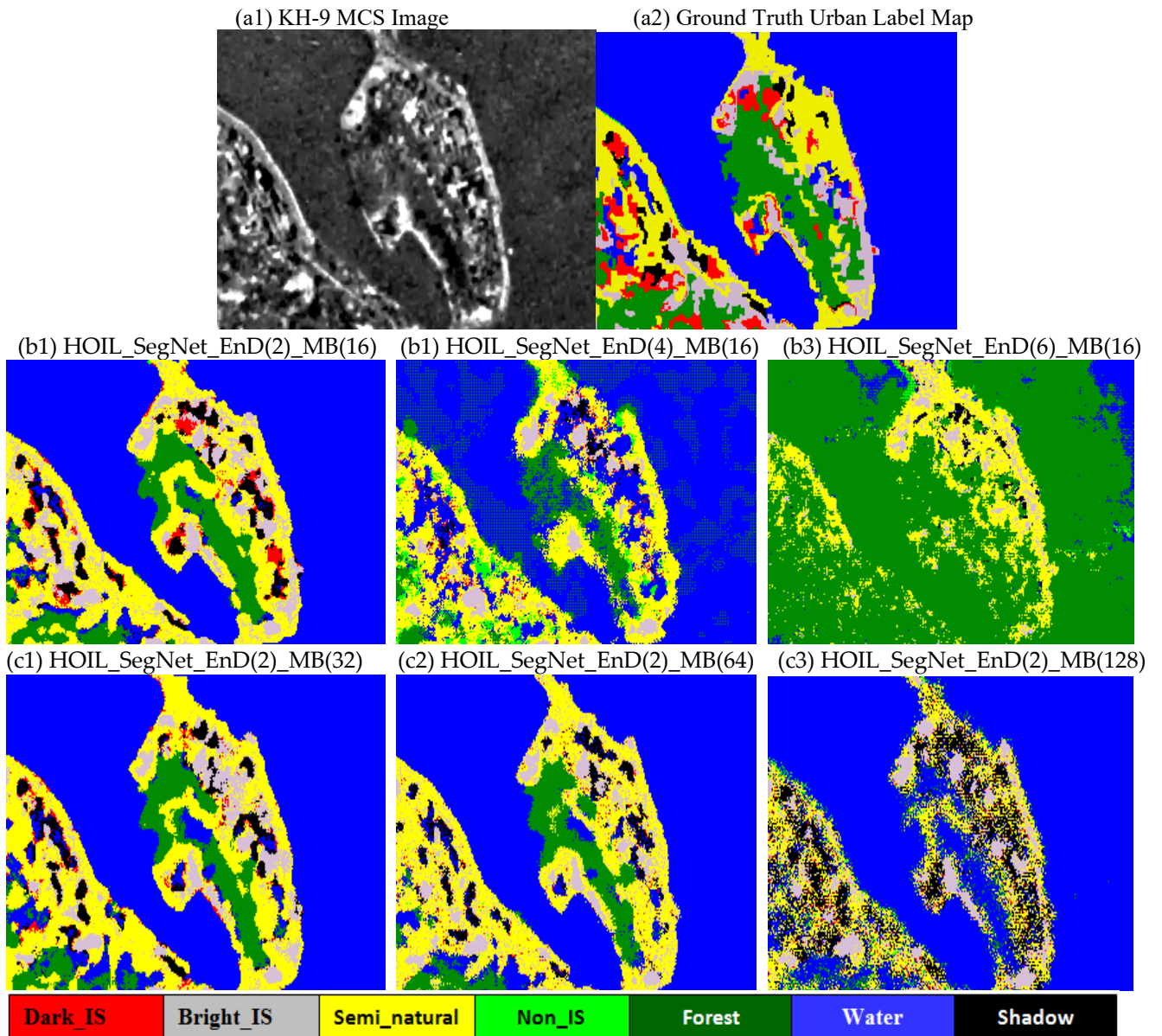


Fig.5 Comparison between the results of applying the proposed HOIL_SegNet_EnD(2)_MB(16), a range of encoder depth (b1-b3) and a range of minibatch size (c1-c3). (a1) KH-9 MCS image and (a2) Ground Truth Urban Label Map. The coordinate center of image is 30°15'14.73"N,120°8'17.27"E

(b) Ablation investigation:

Overall accuracies (OAs) were >70% for all benchmark machine learning techniques (Tables 4 and 5). When HOIL_LR and HOIL_MV were used, a significant increase was observed, with an overall accuracy between 75% and 78% (Table 5). Further enhancement of overall accuracy was achieved when the HOIL_SVM, HOIL_RF and HOIL_MLP were selected, leading to high accuracy from 80% to 82% (Tables 4 and 5). The benchmark machine learning method generated a JDWI from 0.56 to 0.7. With HOIL_SVM, HOIL_RF and HOIL_MLP methods, an increase of 0.3-0.11 was observed, resulting in JDWI values ranging from 0.68 to 0.71 (Tables 4 and 5). The HOIL_MLP and HOIL_RF generated superior accuracy metrics in comparison with the other machine learning techniques and the proposed SegNet (HOIL_SegNet_EnD(6)_MB(16)). More importantly, HOIL_MLP and HOIL_RF improved substantially the PR, RC and F1-score of the Bright_IS and Dark_IS classes. For example, HOIL_MLP led to PR of 0.72, RC of 0.67 and F1-score of 0.69 for Bright_IS. However, HOIL_SegNet_EnD(2)_MB(16) generated PR of 0.53, RC of 0.57 of and F1-score of 0.55 for Bright_IS. Additionally, a constant increasing trend PR, RC and F1-score was observed for the other urban classes.

Details of the spatial distribution of the urban landscapes in the predicted output maps are shown in Fig. 6 and Fig. 7. To compare the differences, visually-checked reference maps are given in Fig. 6 (a2) and Fig. 7 (a2). These maps show a subset of the study area within urban regions. In these subsets, all models predicted the spatial patterns of the most common urban classes relatively well in comparison to the deep learning models (Fig. 6 (a3) and Fig. 7 (a3)). A considerable increase in prediction accuracy and ensuing reduction in class confusion was observed in the patterns of urban classes, especially when MLP and RF are included in the classification in comparison with the proposed SegNet. For example, the forested regions generated by the proposed Segnet (Fig. 6(a3)-dark green colour-red rectangle) were misclassified as water (Fig. 6(a3)-dark blue colour-red rectangle).

Table 4, Accuracy of benchmark machine learning approaches: HOIL_LinearSVC, HOIL_SVM and HOIL_RF

Class	Methods								
	HOIL_LinearSVC			HOIL_SVM			HOIL_RF		
	PR	RC	F1-score	PR	RC	F1-score	PR	RC	F1-score
Bright_IS	0.55	0.63	0.59	0.72	0.60	0.66	0.74	0.63	0.68
Dark_IS	0.59	0.44	0.50	0.64	0.72	0.68	0.62	0.75	0.68
Non_IS	0.77	0.74	0.76	0.78	0.90	0.83	0.82	0.90	0.86
Water	0.72	0.83	0.77	0.92	0.85	0.88	0.94	0.87	0.91
Shadow	0.55	0.81	0.65	0.70	0.83	0.76	0.74	0.82	0.77
Semi_natural	0.51	0.51	0.51	0.66	0.61	0.64	0.70	0.64	0.67
Forest	0.95	0.85	0.90	0.96	0.93	0.94	0.95	0.95	0.95
OA	70%			80%			82%		
JDMI	0.54			0.66			0.69		
JDMAI	0.51			0.64			0.66		
JDWI	0.56			0.68			0.71		
Time-Minutes	5			60			10		

Table 5, Accuracy of benchmark machine learning approaches: HOIL_LR, HOIL_MV and HOIL_MLP

Class	Methods								
	HOIL_LR			HOIL_MV			HOIL_MLP		
	PR	RC	F1-score	PR	RC	F1-score	PR	RC	F1-score
Bright_IS	0.65	0.50	0.57	0.71	0.59	0.64	0.72	0.67	0.69
Dark_IS	0.51	0.64	0.57	0.59	0.71	0.65	0.71	0.72	0.72
Non_IS	0.72	0.86	0.79	0.76	0.90	0.82	0.84	0.87	0.86
Water	0.85	0.81	0.83	0.92	0.83	0.87	0.92	0.88	0.90
Shadow	0.65	0.75	0.69	0.67	0.84	0.75	0.73	0.79	0.76
Semi_natural	0.58	0.44	0.50	0.66	0.54	0.59	0.70	0.72	0.71
Forest	0.93	0.93	0.93	0.95	0.93	0.94	0.95	0.95	0.95
OA	74%			78%			82%		
JDMI	0.58			0.64			0.70		
JDMAI	0.55			0.61			0.67		
JDWI	0.60			0.66			0.71		
Time-Minutes	10.2			70.5			15.2		

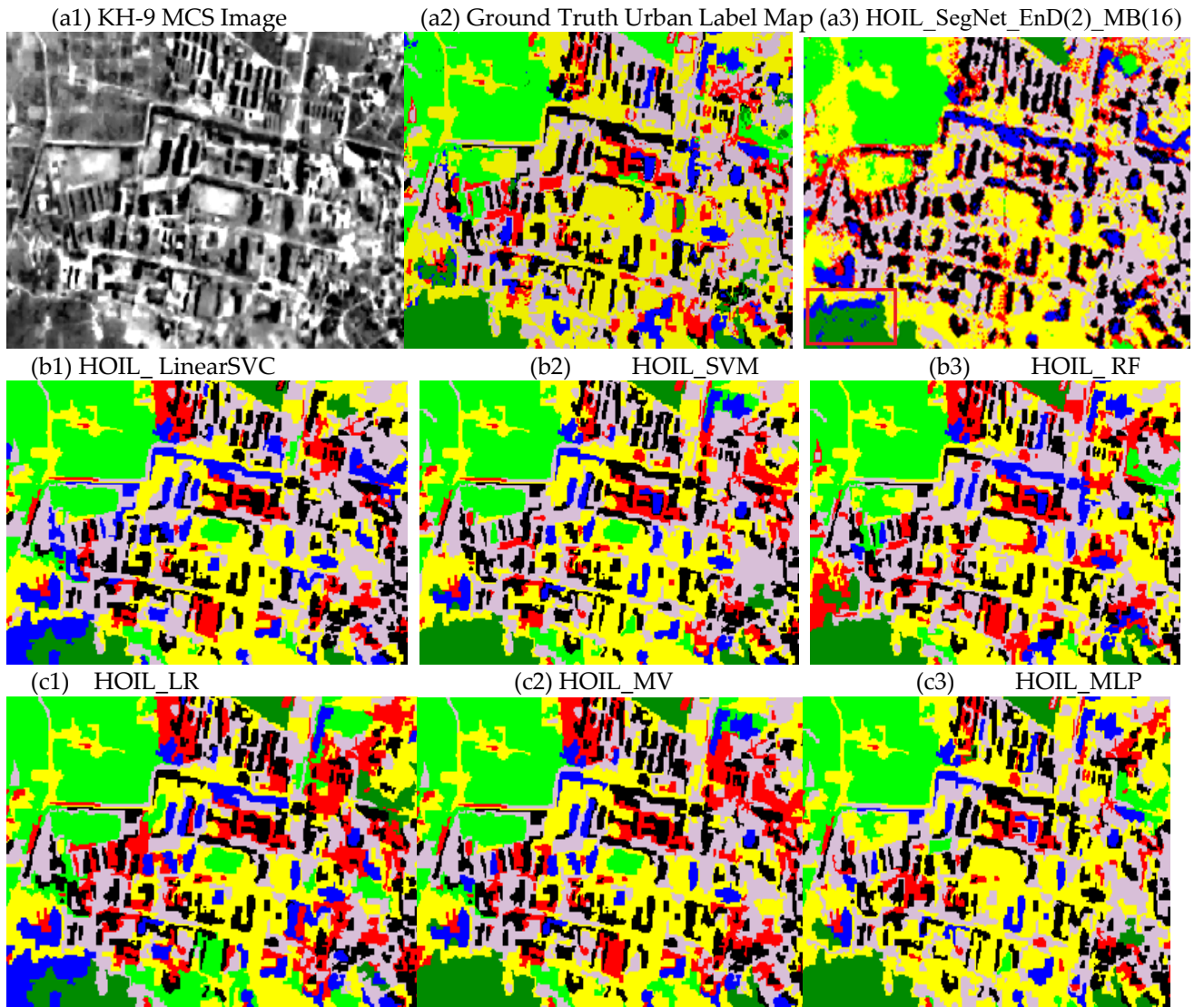


Fig.6 Comparison between the results of applying the proposed HOIL_SegNet_EnD(2)_MB(16) (a3) and benchmark machine learning techniques (b1-b3 and c1-c3). (a1) KH-9 MCS image and (a2) Ground Truth Urban Label Map. The coordinate center of image is 30°15'57.17"N,120°7'10.56"E

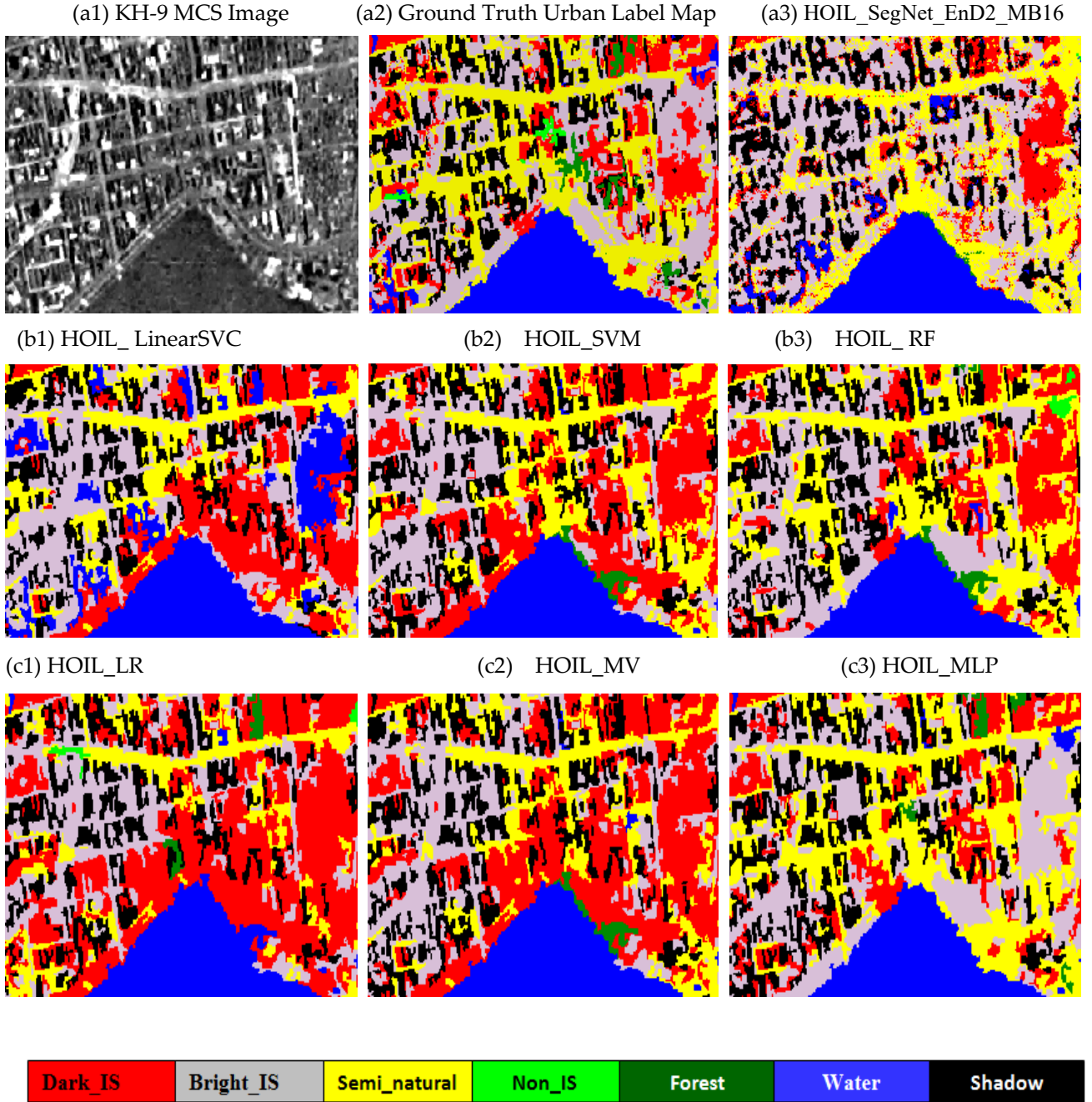


Fig.7 Comparison between the results of applying the proposed HOIL_SegNet_EnD(2)_MB(16) (a3) and benchmark machine learning techniques (b1-b3 and c1-c3). (a1) KH-9 MCS image and (a2) Ground Truth Urban Label Map. The coordinate center of image is 30°15'12.31"N,120°9'28.72"E

4.2 KH-9 PCS analytical framework

4.2.1 Developing historical urban ground control points (GCPs)

Using SPSG, a large number of no-change urban points were identified between the stereo-pair KH-9 PCS images in 1974 and contemporary FSR aerial imagery in 2023 (Fig.8 (a1) and (b2)). Of these, we selected 100 points with an overall accuracy of 98%. After a meticulous visual assessment, 50 no-change urban points with an overall accuracy of 100% were extracted from the 100 points. SPSG not only was superior in identifying many architectural heritage objects within urban regions, but also detected many no-change non-urban objects (e.g., lake and rivers) (Fig.8 (b1) and (b2)). The identified urban no-change regions were used to establish historical GCPs (Fig.8 (c1) and (c2)). The geographic locations of no-change urban regions were transformed to Baidu street view images (<http://map.baidu.com>, access 6 July 2024) (Fig.8 (c3)), and the number of floors of each building was then identified manually on the Baidu street map. The number of floors was converted to building height assuming the height of each floor is 3 m (Cao and Huang 2021, Cai et al. 2023).

The time taken to identify no-change urban regions was also estimated. First, five photo interpreter experts with local knowledge of Hangzhou City were trained for visual interpretation of KH-9 PCS imagery for two months. Next, both KH-9 PCS imagery in 1974 and contemporary FSR aerial imagery in 2023 were distributed among them. The identifying no-change urban regions from those images took nearly another two months (3 weeks for identifying no-change regions, one week for field survey within Hangzhou city, 2 weeks for refining the results, 2 weeks for digitizing no-change regions) using visual interpretation despite the experts having sufficient local knowledge and experience. However, whole SPSG procedures took one hour (30 minutes for image preparation, 2 minutes for identifying points and 28 minutes for refining) to identify no-change regions in both the KH-9 PCS imagery in 1974 and contemporary FSR aerial imagery in 2023.

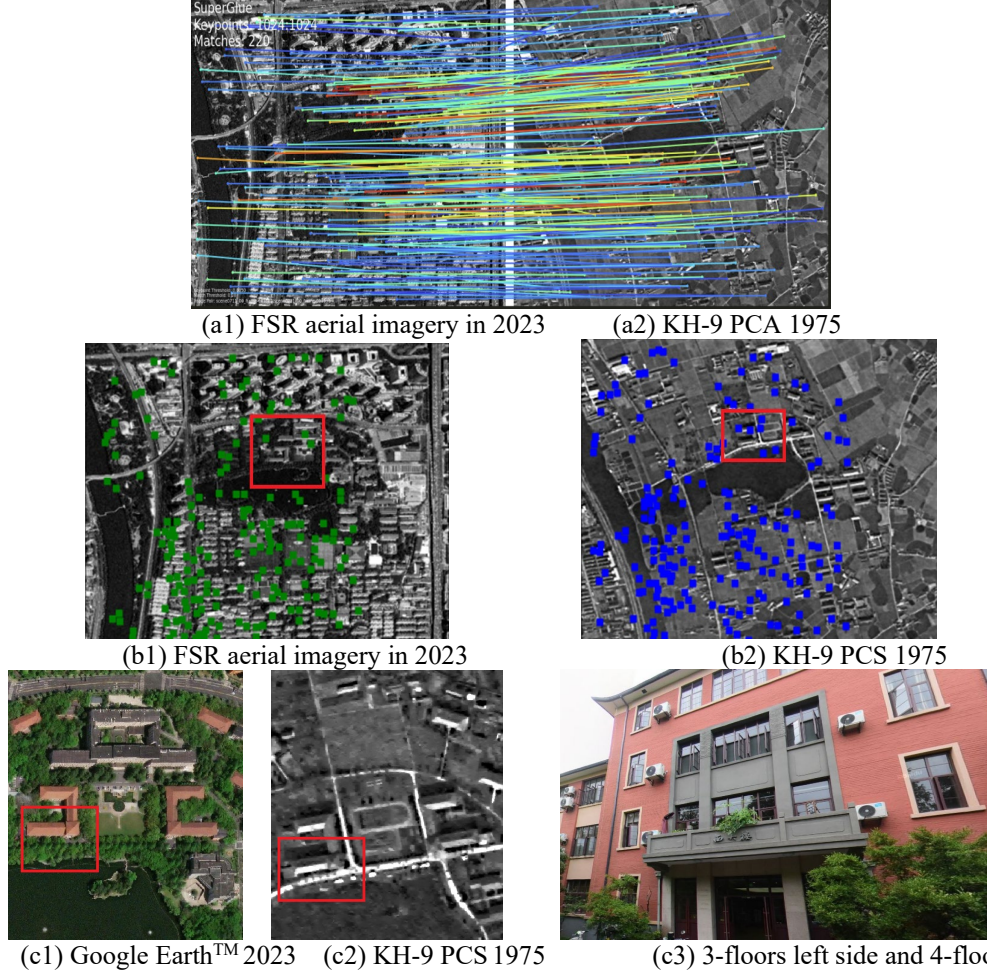


Fig.8 (a1) and (a2) identified matching points between KH-9 PCS in 1975 and FSR aerial imagery 2023; (b2) and (b2) best matches selected between these images; (c1) and (c2) an identified region on Google Earth™ 2023 and KH-9 PCS 1975; (c3) Count of number of floors using Baidu steet map (<http://map.baidu.com>, access 6 July 2024) for height estimation; the coordinate center of the images (a-f) is 30°16'30.10"N, 120°11'26.50"E.

4.2.2 Generating the stereo anaglyph image

The original stereo-paired panchromatic images from the KH-9 PCS archives were utilized intentionally to determine whether anaglyph procedures had any pre-processing effects (e.g., de-noising and contrast enhancement). The image histogram analysis demonstrated that the stereo anaglyph procedure can increase the quality of KH-9 PCS (Fig. 9, for the sake of brevity only left images show herein). In particular, the proposed SfM_anaglyph and benchmark SPSG_anaglyph techniques generated appropriate outcomes in comparison to the original image and the RPCs-anaglyph technique.

In the context of matching, both the proposed SfM_anaglyph and benchmark SPSG_anaglyph methods outperformed the RPCs_anaglyph technique (Table 6). For example, the mutual information and geometric transformation of the proposed SURF_anaglyph were 1.04 and (-0.20, -1.81), respectively. However, the RPCs_anaglyph technique generated unsatisfactory mutual information (1.01) and geometric transformation (5.98, 3.94) values. Visual inspection (Fig. 10) and image quantity measures (Table 7) also confirmed this finding. The ghosting effects produced by the RPCs_anaglyph technique were large, reflected in the ghosting effects percentage. In addition, dissimilarity and contrast of the stereo anaglyph image declined sharply using the RPCs_anaglyph technique as evidenced in the GLCM-dissimilarity and contrast measures (Table 7). Both the SfM_anaglyph and benchmark SPSG_anaglyph techniques yielded promising performance, yet the SPSG_anaglyph stereo approach generated a larger red-ghosting effect compared to the proposed SfM_anaglyph.

Focusing on the processing time, SPSG_anaglyph took two hours to process the stereo images with 8000×6000 pixels to generate the stereo anaglyph image while the proposed SfM_anaglyph process was performed within only two minutes. The RPCs (with GCPs) approach demanded one week for establishing the anaglyph image.

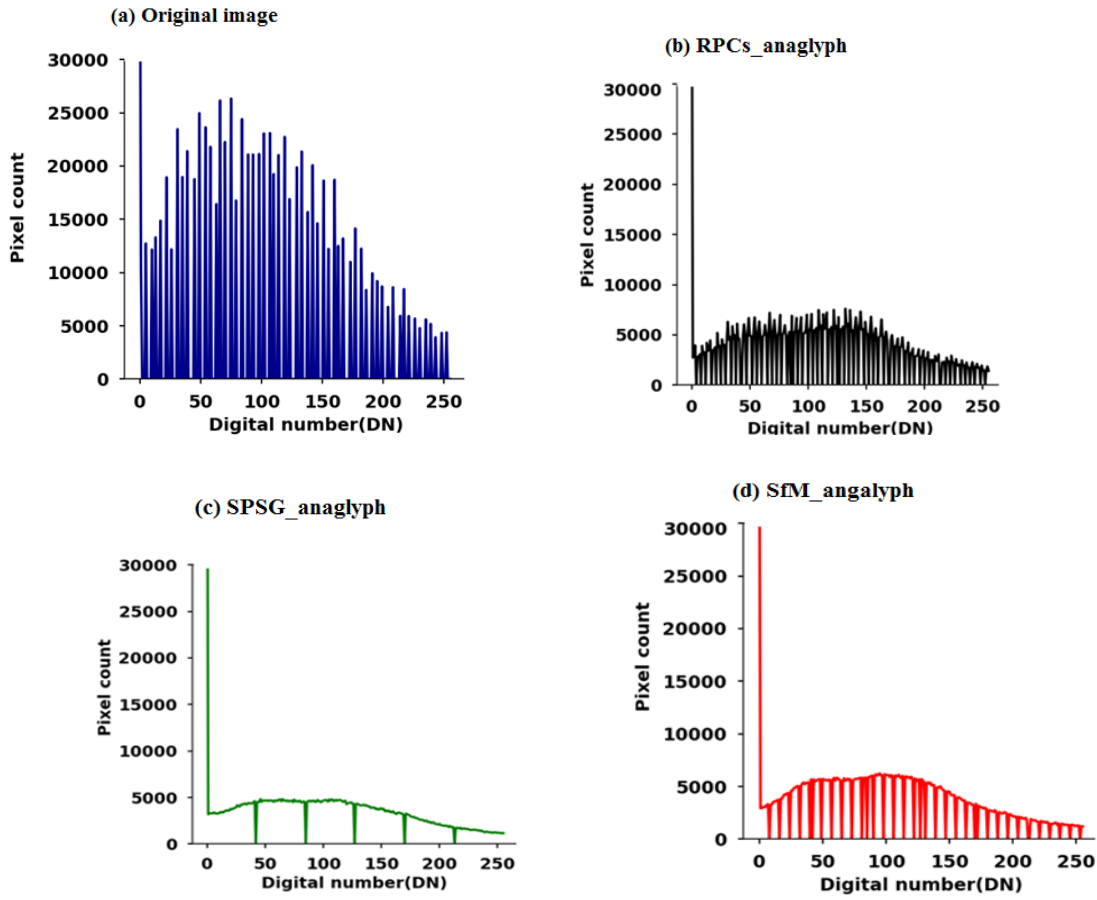


Fig.9 Comparison between image histograms of the original KH-9 PCS image in 1974 and the results of the three stereo anaglyph techniques. For the sake of brevity only the left images are shown.

Table 6 Image stereo matching quality assessment and time of processing

Stereo approaches	Normalized Mutual information	Geometric transformation (t_x, t_y)	Time of Processing
Perfect match	2.0	[0,0]	-----
SPSG_anaglyph	1.05	[0.23, 0.11]	2 hours
SfM_anaglyph	1.04	[-0.20, -1.81]	2 min
RPCs_anaglyph	1.01	[5.98, 3.94]	1 to 3 week(s)
Original images	1.0	[381.37, -374.68]	-----

Table 7 Image stereo quality assessment

Stereo approaches	Ghosting effect-Red (%)	Ghosting effect-Blue (%)	Std_Cont (PCA1) ¹	Std_Diss (PCA1) ²
Original images	25.33	28.88	38.30	1.69
RPCs_anaglyph	22.67	20.20	23.48	1.56
SPSG_anaglyph	30.50	12.60	54.01	2.26
SfM_anaglyph	14.50	15.64	49.79	2.92

1: Standard deviation of Contrast_Principal Components Analysis, component No.1(Std_Cont(PCA1))

2: Standard deviation of Dissimilarity_Principal Components Analysis, component No.1(Std_Diss(PCA1))

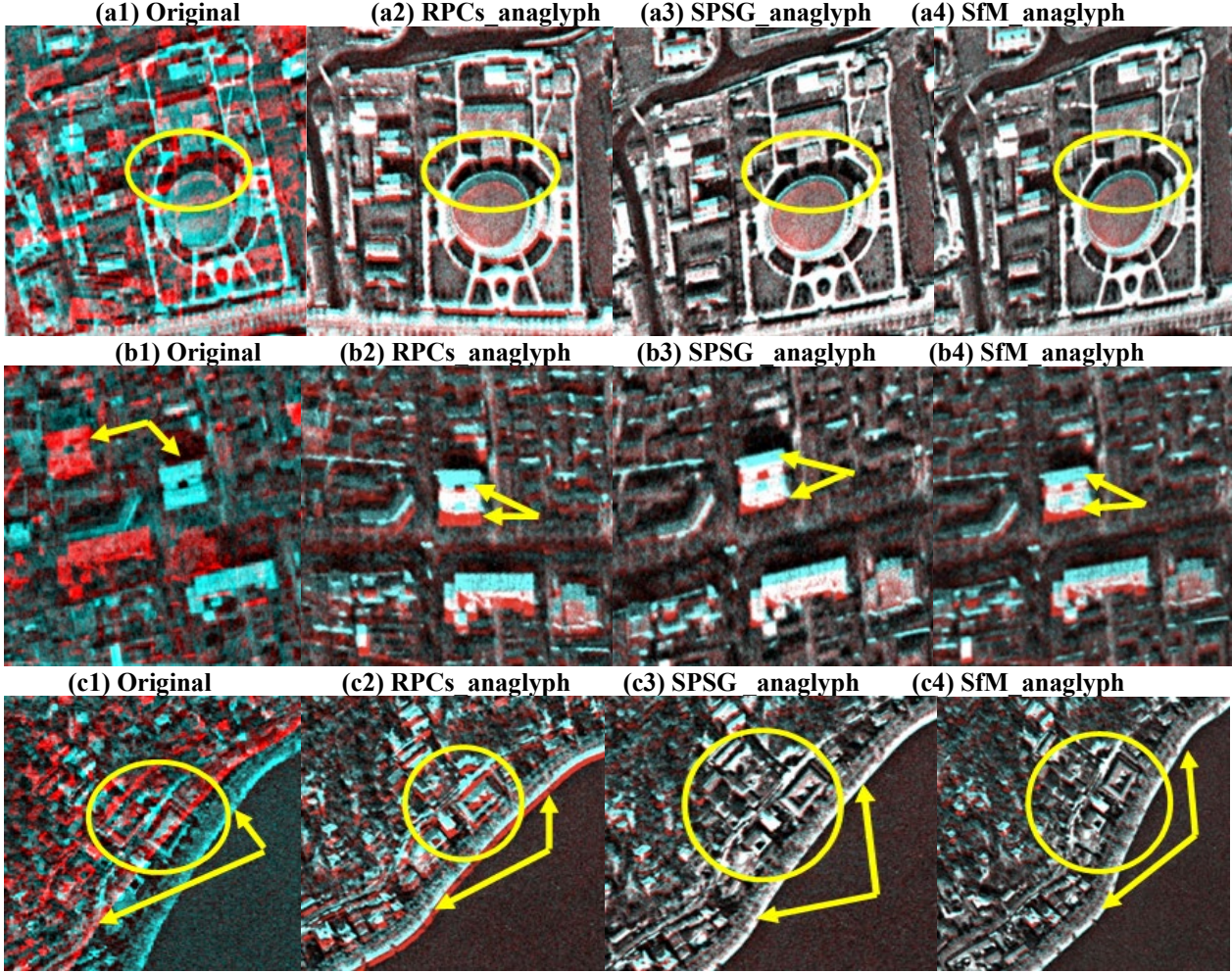


Fig.10 Comparison between original stereo-paired of KH-9 PCS in anaglyph modes and outcomes of stereo anaglyph techniques. The coordinate center of images are (a1-a4) $30^{\circ}16'22.75''N, 120^{\circ}10'5.70''E$; (b1-b4) $30^{\circ}15'10.85''N, 120^{\circ}9'58.43''E$; (c1-c4) $30^{\circ}15'35.23''N, 120^{\circ}8'26.79''E$

4.2.3 Creating a historical ultra FSR urban DSM

The accuracies of the proposed and benchmark DSM approaches were evaluated by the corresponding RMSE, MSE, bias, Std and R-Square (Table 8). Overall, the accuracy measures for the proposed SfM-DSM workflow outperformed the benchmark RPCs-DSM approach. For example, RMSE for latitude (X), longitude (Y) and height (Z) of the proposed SfM-DSM approach were about 1.84 m, 3.08 m and 0.85 m, respectively. However, RPCs-DSM approach yielded an RMSE of 2.0 m, 0.57 m and 3.86 m, for X , Y , Z , respectively.

In terms of visual inspection, the proposed SfM-DSM technique performed well where urban features were readily distinguishable with respect to their height from non-urban features (Fig. 11). In addition, the SfM approach revealed many urban features which could not be identified by a single panchromatic image. For example, this technique identified a small human-made island not identifiable using the DSM-based RPCs-DSM technique (Fig. 11 (a3)). Regarding the processing time, the SfM workflow including accuracy assessment required less time (2 hours) and it was fully automatic while the RPCs approach demanded more time (more than a week). This is because RPCs depend on many parameters such as estimating the RPCs and point cloud which were determined manually.

Table 8 Assessing accuracy of proposed SfM-DSM and benchmark RPCs-DSM.

Indicators	SfM-DSM			RPCs-DSM		
	X (m)	Y (m)	Z (m)	X(m)	Y (m)	Z (m)
RMSE	1.84	3.08	0.85	2.009	0.57	3.86
MSE	3.40	9.50	0.73	4.03	0.33	14.94
Bias	-1.04	2.41	-0.34	-1.51	0.54	-2.60
Mean	-0.85	2.57	-0.41	-0.73	0.89	-1.11
Std	2.60	3.16	2.67	0.81	1.15	3.88
R-square	0.99	0.99	0.97	1	0.99	0.94

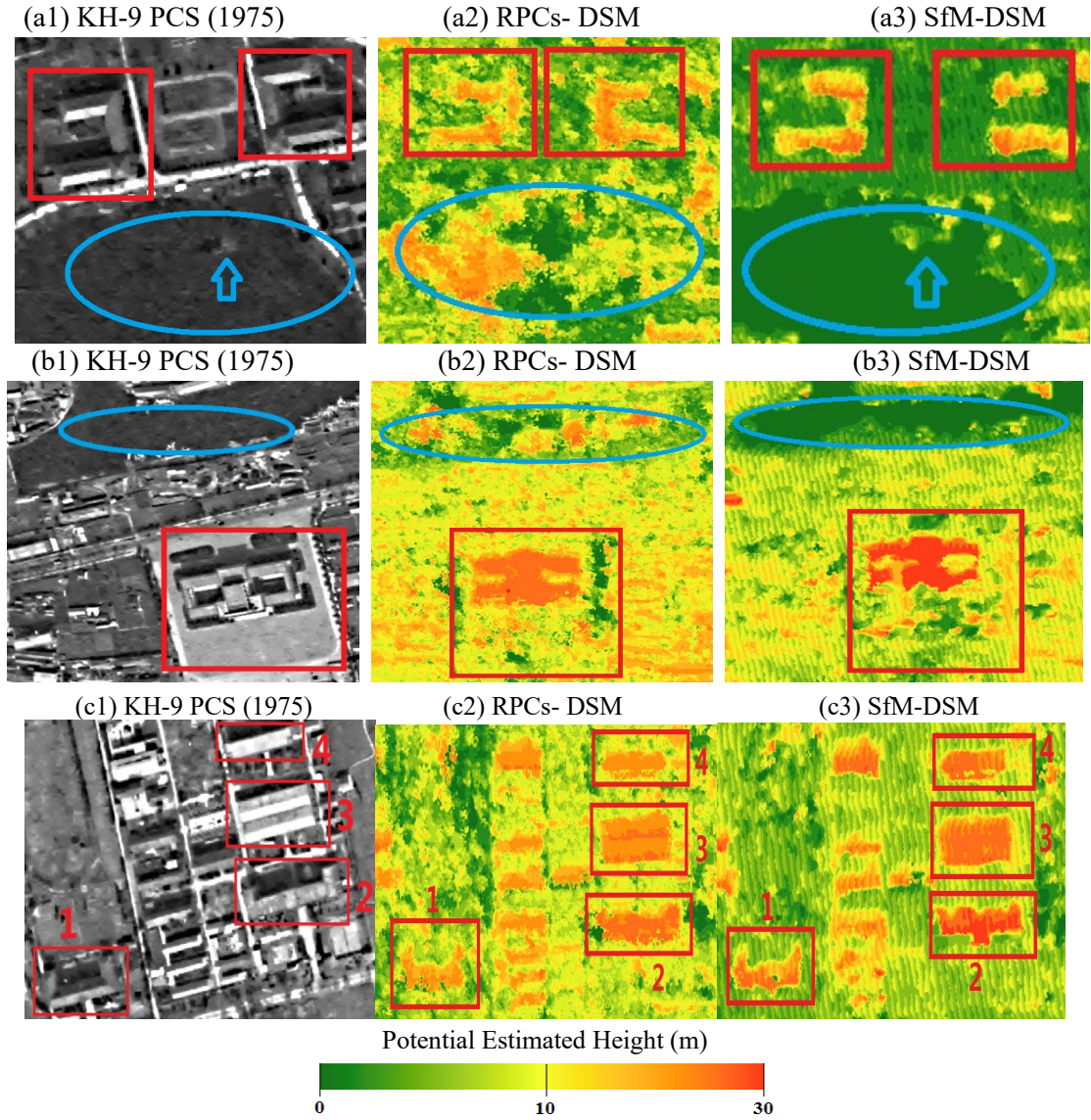


Fig. 11 Comparison between SfM-DSM and RPCs-DSM. The spatial resolution of the DSM is 1.6m. The coordinate centers of the images are (a1-a4) 30°16'30.10"N, 120°11'26.50"E; (b1-b4) 30°16'24.46"N, 120°9'30.81"E; (c1-c4) 30°16'50.82"N, 120°8'26.99"E.

4.3 Application assessment: Historical urban DSM and super-urban development

A crucial gap in urban research is the lack of an ultra FSR DSM representing four-to-five decades ago. The proposed SfM-DSM was used to derive a historical DSM with a spatial resolution of 1.6 m from stereo-paired analog panchromatic KH-9 PCS images in 1974. The generated historical DSM provides micro-topography of urban regions with fine-scale details (Fig. 12). Moreover, wall-to-wall comparison of generated historical DSM with contemporary remotely sensed data can provide valuable information about change in the micro-topography of urban regions, thus, facilitating monitoring for sustainable urban development (Fig. 12 (a) and (b)).

Many studies based on medium resolution imagery (e.g., Landsat) have quantified change in urban landscapes, yet under many circumstances the types of changes and no-changes were unknown. Interestingly, the developed historical FSR urban GCPs has an enormous potential to link KH-9 PCS and MCS products (thematic map, stereo anaglyph and height) to contemporary remotely sensed data. Our framework can quantify “super” details of urban regions which have witnessed unprecedented urban development (Fig. 13).

Many water bodies (wetlands) and agricultural areas in the periphery of Hangzhou City were converted into urban features which led to unprecedented or super new urban development (Fig. 13 (a1) and (a2)). In 1974, Hangzhou was composed of old and dense urban structures, as evidenced in stereo KH-9 PCS image.

Comparison between the products of the proposed framework and Google Earth™ showed that Hangzhou had experienced super redevelopment or the replacement of dense and old urban structures with new ones within existing urban regions (Fig. 13(b1) and (b2)). The redevelopment resulted in mixed urban density, with high-rise built-up areas. Moreover, our investigation indicated that several urban regions (potentially architectural heritage sites) were not affected by changes, which can be considered as super no-change (Fig. 13 (c1) and (c2)). Thus, integrating the proposed frameworks and the generated KH-9 products can enhance our understanding about the categories of LULC prior to urban development and the consequences of such developments.

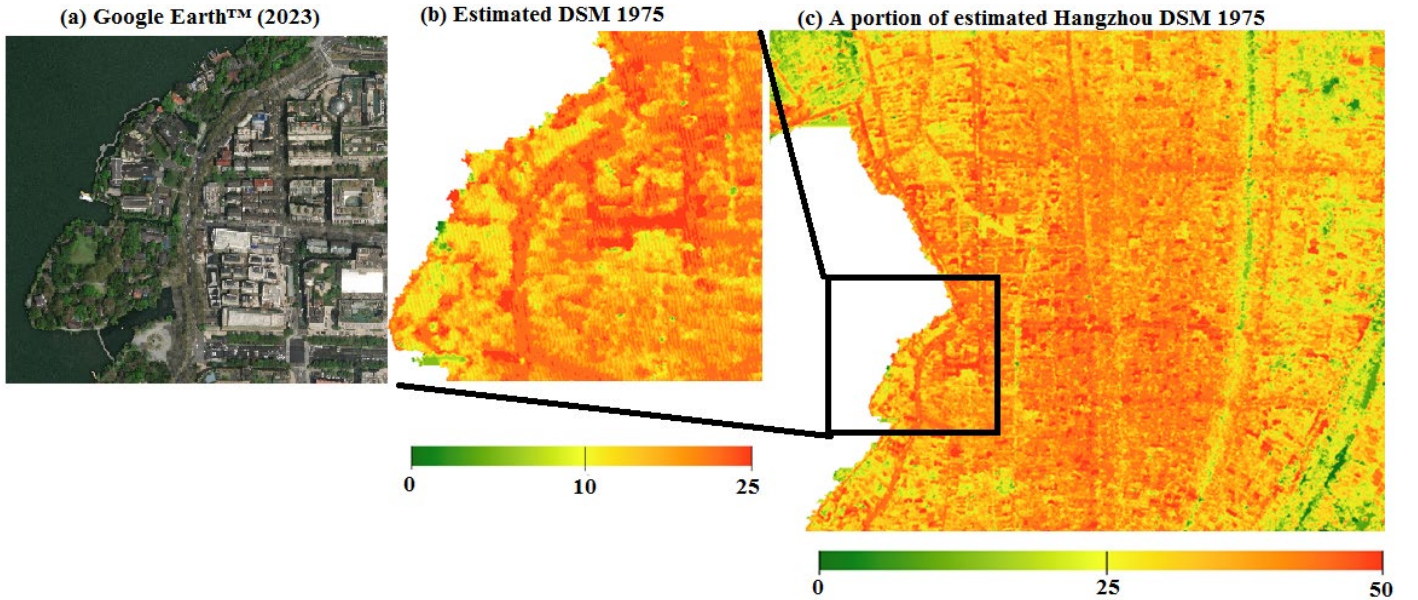


Fig. 12 Comparison between the conditions of urban landscapes on Google Earth™ 2023 and the proposed SfM-DSM. The spatial resolution of the DSM is 1.6 m. The coordinate centers of the images are: (a-b) 30°15'1.24"N, 120°9'17.75"E; (c) 30°15'10.72"N, 120°10'1.91"E.

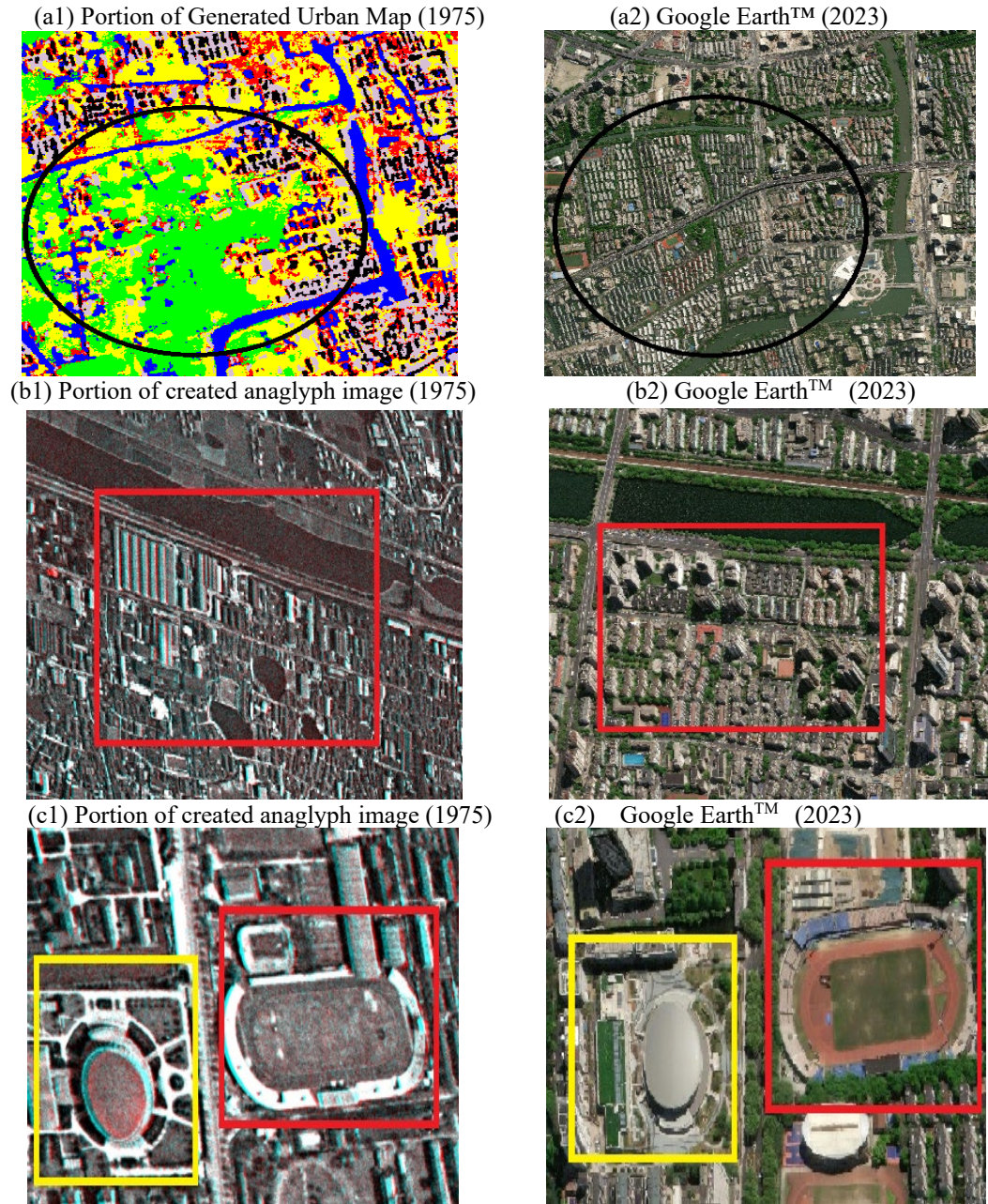


Fig.13 (a1) and (a2) demonstrates super urban new development; (b1) and (b2) shows super urban redevelopment; (c1) and (c2) presents super urban no-change. The coordinate center of images (a1-a2) are 30°16'59.57"N,120°9'52.12"E; (b1-b2) 30°15'48.18"N,120°10'57.58"E; (c1-c2) 30°16'22.85"N,120°10'5.67"E

5. Discussion

5.1. KH-9 MCS analytical framework

The KH-9 MCS consisted of a single panchromatic band with medium spatial resolution (6-9 m) and, hence, these characteristics may be insufficient for identifying urban features. To compensate for this problem, a framework was proposed to synthesize new bands based on hierarchical segmentation procedures. The new bands were inserted into a library called HOIL. A deep learning model was then applied to the HOIL to classify the historical urban landscape in a highly heterogeneous area, comprising a mixture of impervious surfaces, non-impervious surfaces, urban forested regions, water bodies and semi-natural regions. Assessment of the deep learning model hyperparameters and ablation studies were conducted.

The results suggested that minibatch size and encoder depth (number of deep learning layers) had important effects on the performance of the SegNet deep learning classifier applied to KH-9 MCS imagery. For example, the SegNet with encoder depth 6 (HOIL_SegNet_EnD(6)_MB(16)) achieved an OA of 38% , JDMI of 0.23, JDMAI of 0.19 and JDWI of 0.21. Similarly, SegNet with minibatch size 128 (HOIL_SegNet_EnD(2)_MB(128)) obtained an OA of 48%, JDMI of 0.32, JDMAI of 0.27 and JDWI (0.32). However, SegNet with encoder depth 2 and minibatch size 16 (HOIL_SegNet_EnD(2)_MB(16)) yielded reliable urban landscape classification with an OA of 70%, JDMI of 0.54, JDMAI of 0.50 and JDWI of 0.57. Also all urban features in this classifier were classified with higher accuracy in terms of PR and RC and F1-score.

The findings of this research highlight the need for greater understanding of the impacts of other hyperparameters of deep learning models on the reconstruction of historical urban landscape classes from KH-9 MCS imagery. It is noteworthy that recent research on classifying historical landscapes from KH-9 imagery achieved an OA greater than 90% (Sertel et al. 2024). However there are two major reasons for the differences in classification accuracy. First, the aforementioned research focused on land use and land cover with wide geographic scale which could provide sufficient training samples for deep learning models while our research concentrated on the urban landscape only (a portion of land use and land cover) for which it is more difficult to obtain sufficient training samples. The urban landscape in our research consisted of heterogenous urban features with mixed boundaries which amplified mixed pixel problems among the urban classes (Fig. 14). Second, the former research utilized KH-9 PCS imagery while this research employed KH-9 MCS imagery. Although both MCS and PCS were onboard the same satellite, they possess different technical characteristics.

In general, deep learning models demonstrate promising performance compared to conventional machine learning models. Our research suggested that conventional machine learning techniques achieved superior accuracy compared to the proposed SegNet deep learning model. For example, HOIL_MLP generated an OA of 82%, JDMI of 0.70, JDMAI of 0.67 and JDWI of 0.71 whereas SegNet (HOIL_SegNet_EnD(6)_MB(16)) produced an OA of 70%, JDMI of 0.54, JDMAI of 0.50 and JDWI of 0.57. One potential explanation could relate to the training samples. Deep learning models rely on large numbers of training samples for modelling. However, major bottlenecks were encountered here when trying to obtain large numbers of historical training samples particularly over urban regions. Thus, conventional machine learning techniques such as MLP and RF could be an alternative choice for handling the limited number of historical urban training samples for reconstructing historical urban landscapes from KH-9 MCS imagery.

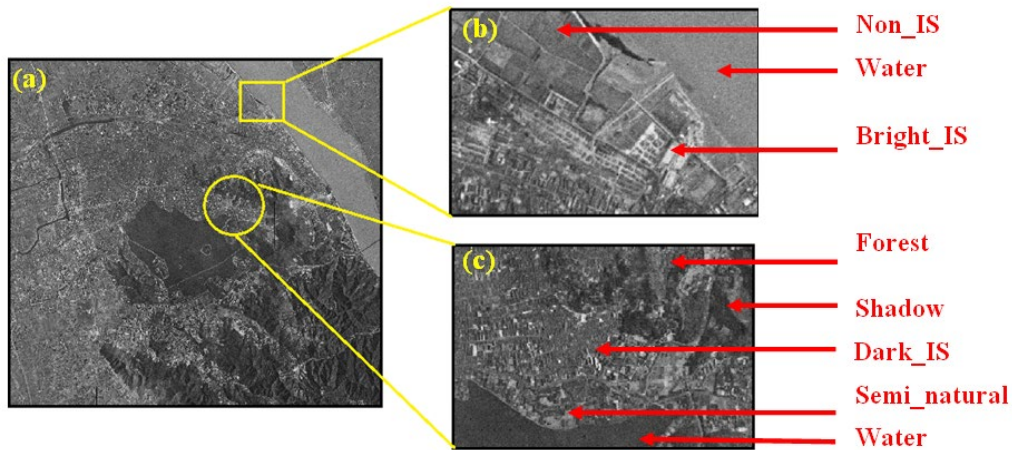


Fig.14 Spectral confusion in analogue panchromatic KH-9 MCS imagery over urban landscapes; The coordinate centers of the images are: (a) 30°15'10.72"N, 120°10'1.91"E; (b) 30°14'30.00"N, 120°12'38.63"E; (c) 30°14'14.30"N, 120°9'1.34"E.

5.2. KH-9 PCS analytical frameworks- methodological implication

5.2.1. Establishing historical urban fine spatial resolution (FSR) GCPs

Historical GCPs play an overarching role in creating stereo anaglyph images and DSMs from KH-9 PCS and other KH data series. This is because GCPs can be used to estimate RPCs photogrammetric models to compensate for the lack of meta parameters of cameras, and they are also necessary for accuracy assessment purposes. Given the ultra FRS of the KH-9 PCS (0.6-1.2 m), it is difficult or nearly impossible to obtain historical, precise and FSR GCP data (e.g., for 1974) in urban regions with respect to the coordinates, type and height of urban features. Although wall-to-wall visual interpretation of KH-9 PCS and contemporary FSR remotely sensed images can be used, it can be subjective, difficult to replicate, time-consuming and non-cost efficient, thus, leading to inconsistent interpretations. In fact, because of the difference in time, land surface and information content between the

historical KH-9 PCS and contemporary remotely sensed data, it is a challenging task to extract similar identifiable points. For example, nearly two months were spent to train a professional photo interpreter on KH-9 PCS images while another two months were used to extract urban features from these images. However, the resulting performance was unsatisfactory and, additionally, the task of individuals was challenging due to looking at panchromatic images for a long time.

To tackle the above problem, the deep learning SPSG was utilized to identify no-urban changes in both the KH-9 PCS image in 1975 and FSR aerial image in 2023. The whole SPSG procedure took about one hour to identify those regions with an OA higher than 90%. Subsequently, the coordinates and types of urban features were extracted from the identified no-urban change regions. Moreover, those regions were linked to Baidu street map (<http://map.baidu.com>, accessed 6 July 2024) to estimate the approximate height of buildings. Despite the KH-9 PCS film-based panchromatic images being acquired four-to-five decades ago, our research demonstrated that advances in deep learning image matching models, such as SPSG procedures in this research and previous investigations, can facilitate identifying no-change features between those images and contemporary remotely sensed data.

5.2.2. Creating the stereo anaglyph image

The absence of meta parameters for the KH-9 PCS cameras is one of the most important obstacles for establishing stereo anaglyph imagery. However, KH-9 PCS images may benefit from recent advances in stereo image rectification techniques, image warping pipelines and developed historical GCPs. In terms of the image stereo anaglyph, the proposed SfM-anaglyph approach utilized the SURF algorithm to identify automatically matching points between the two stereo images. Then, the acquired matching points were used to estimate the F-matrix, thus, aligning the two stereo images. Also two benchmark techniques were developed for ablations investigations. In the first benchmark, SPSG procedures-anaglyph was used to extract automatically matching points between stereo-pairs of KH-9 PCS images. At the same time, the extracted points were inserted into the H-matrix and warping procedure to register one stereo image to the other. In the second benchmark (i.e., RPCs-anaglyph), historical urban GCPs were placed manually into stereo-pair KH-9 images for building the RPCs. Then the stereo images with RPCs were subjected to manual image matching and rectification. Among the implemented anaglyph techniques, the proposed SfM-anaglyph yielded the most promising outcomes in terms of automatic procedure, matching, rectification and image quality relative to an appropriate set of benchmark techniques. Our investigation demonstrated that:

- (1) The stereo anaglyph image procedures based on SURF and SPSG were performed fully automatically while the RPCs procedure was manual.
- (2) The matching stereo-pair KH-9 PCS images with size of 8000×6000 pixels were completed in less than two minutes whereas similar procedures took 2 hours using SPSG -anaglyph. In terms of the RPCs-anaglyph procedure, although an automatic image approach is accommodated in this procedure, this approach could not generate reliable results due to very high ‘Y-parallax’ in the context of KH-9 PCS. Hence, this research used manual matching so as to reach a ‘Y-parallax’ of less than 10 pixels. The manual procedure required one week in our case.

In this research, the original stereo-paired images from KH-9 PCS archives were adopted intentionally for anaglyph procedures to examine whether this could improve the quality of the images. The necessity of image pre-processing for such images should be noted. Our results showed anaglyph procedures can increase the quality of these images to some extent, as evidenced in their image histograms. This could be due to interpolation techniques within the image anaglyph procedure. For example, nearest neighborhood was implemented in stereo rectification in SfM-anaglyph and SPSG-anaglyph while bilinear (automatic choice) was utilized in RPCs-anaglyph.

To comprehensively and systematically assess the performance of the stereo anaglyph technique it is necessary to gauge any pixels mismatching between the stereo images and quantify the quality of those images. Visual inspection may not be sufficiently sensitive for this task. In addition, conventional image quality assessment (e.g., Peak Signal-to-Noise Ratio (PSNR) and structural similarity index (SSIM)) may not offer interpretable information about ghosting effects and change to the local spatial variation of brightness values (BV) as a result of the stereo anaglyph procedures. Therefore, this research adopted the normalized mutual information and geometric transformation measures to quantify the performance of matching. Moreover, ghosting effects and quality of image appearance were assessed by the percentage of ghosting effects using the SVM classifier and GLCM co-occurrence metrics. In particular, GLCM co-occurrence metrics are based on spatial neighborhood and, thus, they capture the response of stereo-pairs of KH-9 PCS images to anaglyph techniques with respect to change in local intensity levels (dissimilarity measure) and local variation (contrast measure).

5.2.3. Reconstructing historical ultra FSR urban DSM

This research mined the potential of the SfM-DSM and RPCs-DSM algorithms to derive historical ultra FSR urban DSMs from stereoscopic KH-9 PCS imagery with respect to the absence of meta parameters (except focal length and scanning resolution of converting analog film to digital image) and minimal historical GCPs. Both SfM-DSM and RPCs-DSM generated urban DSM maps with spatial resolution of 1.6 m in 1974. Of these approaches, the SfM-DSM algorithm generated the most promising performance in comparison to the RPCs-DSM. Our investigation showed that:

- (1) The proposed SfM-DSM algorithm is fully automatic except for placing GCPs on the image to optimize photo alignment. Additionally, default parameters of workflows in SfM-DSM of AgisSoft including ‘Align Photos’, ‘Build Dense Cloud’, ‘Build Mesh’ and ‘Build DEM’ can generate satisfactory outcomes.

- (2) The RPCs-DSM algorithm requires manual intervention, and a range of parameters should be set so as to obtain optimal outcomes. Moreover, this algorithm is sensitive to the pattern of GCPs for building RPCs. If the pattern of GCPs on the stereo-pair images is not selected appropriately even though the RMSE of GCPs may be very low, the built RPCs in the point cloud procedure may not be able to generate the DSM.

5.3. Capability of KH-9 PCS and MCS images for urban studies

The KH-9 PCS and MCS images provided useful information on the Earth's surface, as suggested in previous research. In terms of urban research, these images are capable of extracting highly detailed historical urban features, quantifying precise shapes, locations and heights, as shown in this investigation. For example, the generated products in this research contributed to revealing super no-change, super redevelopment and super new development. Thus, this makes images from KH-9 PCS and MCS archives appealing for application in urban research, such as time-series analysis, urban growth modelling, cultural heritage assessment and urban redevelopment planning. Moreover, KH-9 images offer unique documentation of the 1970s for most urban regions in the world, particularly developing countries, which may not have sufficient FRS aerial photographs or which may not be available publicly. Additionally, such images can contribute in developing a local, regional, national and global fine-scale database (e.g., urban features and land use and land cover (LULC)), which is one of the objectives of United Nations Sustainable Development Goals (UN SDGs).

5.4. Obstacles and suggestions

This research not only enhances the limited existing documented research in applications of KH-9 PCS and MCS in urban investigations, but also offers novel analytical frameworks and ablation studies in this context. Despite these outcomes, some open issues exist which should be addressed in future research. These obstacles and recommendations with respect to the camera types are as follows:

- (a) KH-9 MCS and the proposed analytical deep learning approach for historical urban landscape classification
 - (1) For comprehensive and in-depth investigations, this research focused on only one city located within a KH-9 MCS image. This limited the number of training samples for deep learning models. Ultimately, future research may consider multiple cities which would increase the number of training samples;
 - (2) This investigation utilized the SegNet model for classifying historical urban landscapes from KH-9 MCS panchromatic imagery. In addition, the effects of encoder depth and minibatch size on classification were studied. Therefore other deep learning models and their hyper parameters should be considered in future investigations (Sertel et al. 2024). Moreover, optimization techniques, such as the Bayesian optimizer, should be of great benefit for selecting appropriate hyperparameters for deep learning models in the context of historical urban classification from KH-9 MCS imagery (Albarakati et al. 2024, Ruhab et al. 2024).
- (b) KH-9 PCS and the produced historical GCPs, stereo anaglyph image and DSM
 - (1) Obtaining accurate DSM and GNSS data were challenging in the study area. To tackle this problem, a novel approach was developed to build historical urban GCPs. In the established GCPs, the heights of buildings were estimated based on the height of each floor as suggested by previous research (Cao and Huang 2021, Cai et al. 2023). Future research may assess the merits of other open source databases such as ALOS World 3D data (5 m) and open street map (<https://download.geofabrik.de/>) for this purpose.
 - (2) This research developed a fully automatic benchmark technique using integration of deep learning SPSG procedures matching points, Homography matrix (H-matrix) and image warping in the Python environment to reconstruct a stereo anaglyph image. Future investigations should assess existing stereo workflows, or develop new workflows in the context of KH-9 PCS imagery.
 - (3) Our investigation could be one of the first to test RPCs and point-cloud modules in ENVI for such purposes based on stereo-paired KH-9 PCS images. However, these modules were implemented manually which was cumbersome and time-consuming, potentially affecting accuracy. Future research may, therefore, develop and exploit automatic procedures to further increase time efficiency and accuracy.
 - (4) To focus only on the Hangzhou City and minimize complexity of full frame (ten tiles) of stereo-pairs KH-9 PCS imagery, this research used a subset of these images to create anaglyph images and DSMs, as suggested by Casana and Cothren (2007) and Zhou et al. (2021). We recommend designing automated workflows for establishing historical anaglyph images and DSMs using full frame of stereo-pairs KH-9 PCS which can be used by researchers and practitioners from any field. In this way, the entire frame of KH-9 PCS could be analyzed at once with smaller errors and distortions, and potentially many frames could be batched processed (e.g., Maurer and Rupper 2015, Ghuffar et al. 2022, Ghuffar et al. 2023).
 - (5) Given the absence of camera meta parameters (e.g., extrinsics, lens distortion) for KH-9 PCS imagery, this research adopted the 'Digital (Frame Central)' model of ENVI, which requires minimum parameters including focal length and scanner resolution for digitizing the film. However, the 'Digital (Frame Central)' model may not be optimal as it is based on the conventional pinhole camera model. Further investigations will, therefore, be needed to develop new photogrammetric models for KH-9 PCS imagery for estimating meta parameters (e.g., Schenk et al. 2003, Sohn et al. 2004), thus, reliably establishing anaglyph images and DSMs.

6. Conclusion

This research examined the unique potential of analog panchromatic images from declassified archives of KH-9 MCS and PCS data for reconstructing an historical urban fine-scale database, particularly when utilized in historical urban landscapes that are characterized by highly heterogeneous land use and land cover. Both KH-9 MCS and PCS images share three major shortcomings: (1) availability of a single panchromatic band which can impede the precision of pattern recognition, feature matching and classification, (2) absence of technical parameters of cameras which could limit stereo procedures such as creating anaglyph images and DSMs and (3) lack of historical Ground Control Points (GCPs) data for analytical purposes such as accuracy assessment. Such problems can limit application of KH-9 MCS and PCS images for reconstruction of a historical urban fine-scale database.

A multi-stage framework was devised to tackle the above obstacles, which consisted of two components: KH-9 MCS analytical and KH-9 PCS analytical frameworks. The efficiency of the both frameworks was scrutinized comprehensively and systematically using a range of ablation investigations in Hangzhou City, capital of Zhejiang Province, China, as a representative historical urban region with heterogeneous urban features.

The integration of HOIL and the proposed SegNet (HOIL_SegNet_EnD(2)_MB(16)) increased the overall accuracy. This result is encouraging considering the level of urban land-use complexity in the KH-9 MCS panchromatic band. However, a combination of HOIL and other benchmark machine learning techniques such as MLP and RF produced the greatest classification accuracy compared to the proposed HOIL_SegNet_EnD(2)_MB(16). This difference in accuracy could be due to the lack of sufficient training samples for deep learning as our study focused only on one city.

Regarding KH-9 PCS analytical approaches, this research developed a technical pipeline using SPSG to establish a historical GCPs database. Such database can be used to estimate extrinsic orientation stereoscopic camera parameters and for accuracy assessment purposes. A fully automatic SfM-anaglyph approach was then developed to generate a stereo anaglyph image in the absence of camera parameters. Moreover, a SfM-DSM approach was proposed to create a historical ultra-FSR urban DSM. An ablation investigation revealed that (1) established historical GCPs facilitated estimating extrinsic orientation stereoscopic KH-9 PCS camera parameters, and this database showed potential application for accuracy assessment of the DSM and (2) both the SfM-DSM and SfM-anaglyph techniques demonstrated promising performance in comparison with benchmark techniques.

The proposed KH-9 MCS and PCS frameworks for establishing a historical urban fine-scale database have enormous potential to support urban research in a very wide range of applications. For example, our application investigations showed that Hangzhou City experienced three types of changes: super new development, super redevelopment and super no-change. Moreover, the comparison between the established database and Google Earth™ revealed details of LULCC such as conversion of many small ponds into impervious surfaces.

This research provides a bridge between analog panchromatic KH-9 images and advances in digital image processing techniques for creating a historical urban database. Further research should build on the results presented here and especially focus on exploration of the integration of KH-9 images with further state-of-the-art digital image processing techniques, at a time when such investigations are paramount to attainment of the UN SDGs, thus supporting urban sustainable development.

Statements and Declarations

Competing Interests: The authors declare no conflict of interest.

References

- Albarakati H. M., Rehman S. U., Khan M. A., Hamz A., Aftan J., Alasiry A., Marzougui M. (2024) A Unified Super-Resolution Framework of Remote-Sensing Satellite Images Classification Based on Information Fusion of Novel Deep Convolutional Neural Network Architectures. *Ieee J-Stars*, 17.
- Altmaier A., Kany C. (2002) Digital surface model generation from CORONA satellite images. *Isprs J Photogrammetry*, 56, 221-235.
- Baatz M., Schäpe A. 2000. Multiresolution segmentation—an optimization approach for high quality multi-scale image segmentation. In *Angewandte Geographische Informations-Verarbeitung XII*, eds. J. Strobl, T. Blaschke, G. Griesebner. Wichmann Verlag, Karlsruhe: 12–23.
- Badrinarayanan V., Kendall A., Cipolla R. (2017) SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *Ieee T Pattern Anal*, 39, 2481-2495.
- Belgiu M., Drăguț L. (2014) Comparing supervised and unsupervised multiresolution segmentation approaches for extracting buildings from very high resolution imagery. *Isprs J Photogrammetry*, 96, 67-75.
- Benz U. C., Hofmann P., Willhauck G., Lingenfelder I., Heynen M. (2004) Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry & Remote Sensing*, 58.
- Blaschke T. (2010) Object based image analysis for remote sensing. *Isprs J Photogrammetry*, 65, 2-16.
- Bolch T., Buchroithner M., Pieczonka T., Kunert A. (2008) Planimetric and volumetric glacier changes in the Khumbu Himal, Nepal, since 1962 using Corona, Landsat TM and ASTER data. *Journal of Glaciology*, 54.

- Cai B.,Shao Z.,Huang X.,Zhou X.Fang S. (2023) Deep learning-based building height mapping using Sentinel-1 and Sentienl-2 data. *Int J Appl Earth Obs*, 122.
- Cao Y.Huang X. (2021) A deep learning method for building height estimation using high-resolution multi-view imagery over urban areas: a case study of 42 Chinese cities. *Remote Sens Environ*, 264.
- Casana J.Cothren J. (2007) Stereo analysis, DEM extraction and orthorectification of CORONA satellite imagery: archaeological applications from the Near East. *Antiquity*, 82, 732 - 749.
- Cheng Q. M.,Zhang Q.,Fu P.,Tu C.Li S. (2018) A survey and analysis on automatic image annotation. *Pattern Recogn*, 79, 242-259.
- Declassified_Data_2. 2018. USGS EROS Archive - Declassified Data - Declassified Satellite Imagery - 2. Earth Resources Observation and Science (EROS) Center
- Declassified_Data_3. 2018. USGS EROS Archive - Declassified Data - Declassified Satellite Imagery - 3. Earth Resources Observation and Science (EROS) Center
- Dehecq A.,Gardner A. S.,Alexandrov O.,McMichael S.,Hugonnet R.,Shean D.Marty M. (2020) Automated Processing of Declassified KH-9 Hexagon Satellite Images for Global Elevation Change Analysis Since the 1970s. *Front. Earth Sci*, 8.
- Deshpande P.,Belwalkar A.,Dikshit O.Tripathi S. (2021) Historical land cover classification from CORONA imagery using convolutional neural networks and geometric moments. *Int J Remote Sens*, 42.
- DeTone D.,Malisiewicz T.Rabinovich A. 2018. SuperPoint: Self-supervised interest point detection and description. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*.
- Duro D. C.,Franklin S. E.Dubé M. G. (2012) A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. *Remote Sens Environ*, 118, 259–272.
- eCognition_Reference_Book. 2014. *eCognition® Developer 9.0.1 software*
- eCognition_User_Guide. 2014. *eCognition® Developer 9.0.1 software* Germany.
- ENVI_help. 2024. Build RPCs. ed. nv5geospatialsoftware.
- Farella E. M.,Morelli L.,Remondion F.,Mills J. P.,Haala N.Crompvoets J. (2022) The EuroSDR TIME benchmark for historical aerial images. *International archives of the photogrammetry, remote sensing and spatial information sciences*, 43, 1175-1182.
- Fowler M. (2022) The archaeological potential of KH-9 HEXAGON satellite photographs: the Roman temporary camps and siege works at Machaerus and Masada. *AARG News*, 64, 37-53.
- Galiatsatos N.,Donoghue D. N. M.Philip G. (2008) High resolution elevation data derived from stereoscopic CORONA imagery with minimal ground control: an approach using Ikonos and SRTM data. *Photogramm Eng Rem S*, 74, 1093–1106.
- Gao Z. J.,Wang Q.,Mei T. Y.,Cheng X. H.,Zi Y.Yang H. W. (2024) An Enhanced Encoder-Decoder Network Architecture for Reducing Information Loss in Image Semantic Segmentation. *arXiv*.
- Geiss C.,Klotz M.,Schmitt A.Taubenbock H. (2016) Object-Based Morphological Profiles for Classification of Remote Sensing Imagery. *Ieee T Geosci Remote*, 54, 5952-5963.
- Ghuffar S.,Bolch T.,Rupnik E.Bhattacharya A. (2022) A Pipeline for Automated Processing of Declassified Corona KH-4 (1962-1972) Stereo Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 60.
- Ghuffar S.,King O.,Guillet G.,Rupnik E.Bolch T. (2023) Brief Communication: Glacier mapping and change estimation using very high resolution declassified Hexagon KH-9 panoramic stereo imagery (1971-1984). *The Cryosphere Discuss*, 17, 1299-1306.
- Hall-Beyer M. (2017) Practical guidelines for choosing GLCM textures to use in landscape classification tasks over a range of moderate spatial scales. *Int J Remote Sens*, 38, 1312-1338.
- Hammer E.,FitzPatrick M.Ur J. (2022) Succeeding CORONA: declassified HEXAGON intelligence imagery for archaeological and historical research. *Antiquity*, 96 , Issue 387 , June 2022 , pp., 679 - 695.
- Haya G. J.,Blaschke T.,Marceau D. J.Bouchard A. (2003) A comparison of three image-object methods for the multiscale analysis of landscape structure. *ISPRS Journal of Photogrammetry & Remote Sensing*, 57.

- Hofmann P. 2001. Detecting urban features from IKONOS data using an object-oriented approach. In *1st Annual Conf RSPSoc 2001*. London.
- Hossain M. D.Chen D. (2019) Segmentation for Object-Based Image Analysis (OBIA): A review of algorithms and challenges from remote sensing perspective. *Isprs J Photogramm*, 150, 115-134.
- Hussain E.,Shan J. (2016) Urban building extraction through object-based image classification assisted by digital surface model and zoning map. *International Journal of Image and Data Fusion*, 7, 63-82.
- Jensen J. R. 2015. *Introductory digital image processing : a remote sensing perspective*. Upper Saddle River, N.J.: Prentice Hall.
- Karakış S.,Marangoz A. M.Büyüksalih G. 2006. ANALYSIS OF SEGMENTATION PARAMETERS IN ECOGNITION SOFTWARE USING HIGH RESOLUTION QUICKBIRD MS IMAGERY. In *ISPRS Workshop on Topographic Mapping from Space (with Special Emphasis on Small Satellites)*. Ankara, Turkey.
- Kirillov A.,Mintun E.,Ravi N.,Mao H.,Rolland C.,Gustafson L.,Xiao T.,Whitehead S.,Berg A. C.,Lo W. Y.,Dollár P.Girshick R. 2023. Segment Anything. In *IEEE/CVF International Conference on Computer Vision (ICCV)*. Paris, France.
- Laliberte A. S.,Fredrickson E. L.Rango A. (2007) Combining Decision Trees with Hierarchical Object-oriented Image Analysis for Mapping Arid Rangelands. *Photogrammetric Engineering and Remote Sensing*, 73, 197–207.
- Laliberte A. S.,Rango A.,Havstad K. M.,Paris J. F.,Beck R. F.,McNeely R.Gonzalez A. L. (2004) Object-oriented image analysis for mapping shrub encroachment from 1937-2003 in southern New Mexico. *Remote Sensing of Environment*, 93, 198–210.
- Li J. Y.,Cai Y. X.,Kou M. Y.Zhang T. X. (2024) A review of remote sensing image segmentation by deep learning methods. *International Journal of Digital Earth*, 17.
- Ma L.,Liu Y.,Zhang X. L.,Ye Y. X.,Yin G. F.Johnson B. A. (2019) Deep learning in remote sensing applications: A meta-analysis and review. *Isprs J Photogramm*, 152, 166-177.
- Maiwald F.,Feurer D.Eltner A. (2023) Solving photogrammetric cold cases using AI-based image matching: New potential for monitoring the past with historical aerial images. *Isprs J Photogramm*, 206, 184-200.
- Marzolf I.,Kirchhoff M.,Stephan R.,Seeger M.,Ait Hssaine A.Ries J. (2022) Monitoring Dryland Trees With Remote Sensing. Part A: Beyond CORONA-Historical HEXAGON Satellite Imagery as a New Data Source for Mapping Open-Canopy Woodlands on the Tree Level. *Front. Environ. Sci*, 10:896702.
- Masters M.Luschi C. (2018) Revisiting Small Batch Training for Deep Neural Networks. *arXiv*.
- Mather P. M.Koch M. 2011. *Computer processing of remotely sensed images : an introduction*. Chichester, West Sussex, England ; Hoboken, NJ: John Wiley & Sons.
- Maurer J.Rupper S. (2015) Tapping into the Hexagon spy imagery database: A new automated pipeline for geomorphic change detection. *Isprs J Photogramm*, 108, 113-127.
- Mboga N.,Grippa T.,Georganos S.,Vanhuysse S.,Smets B.,Dewitte O.,Wolf E.Lennert M. (2020) Fully convolutional networks for land cover classification from historical panchromatic aerial photographs. *Isprs J Photogramm*, 167.
- Neupane B.,Horanont T.Aryal J. (2021) Deep learning-based semantic segmentation of urban features in satellite images: A review and meta-analysis. *Remote Sens*, 13.
- Niang A. J.,Hermas E.,Alharbi O.Al-Shaery A. (2020) Monitoring landscape changes and spatial urban expansion using multi-source remote sensing imagery in Al-Aziziyah Valley, Makkah, KSA. *Egypt J Remote Sens*, 23, 89-96.
- Nita M. D.,Munteanu C.,Gutman G.,Abrudan I. V.Radeloff V. C. (2018) Widespread forest cutting in the aftermath of World War II captured by broad-scale historical Corona spy satellite photography. *Remote Sens Environ*, 204, 322-332.
- Pacifici F.,Chini M.Emery W. J. (2009) A neural network approach using multi-scale textural metrics from very high-resolution panchromatic imagery for urban land-use classification. *Remote Sens Environ*, 113, 1276-1292.
- Pedregosa F.,Varoquaux G.,Gramfort A.,Michel V.,Thirion B.,Grisel O.,Blondel M.,Prettenhofer P.,Weiss R.,Dubourg V.,Vanderplas J.,Passos A.,Cournapeau D.,Brucher M.,Perrot M.Duchesnay E. (2011) Scikit-learn: Machine Learning in Python,, . *Journal of Machine Learning Research*, 12, 2825-2830.
- Photoscan_Manual. 2016. Agisoft PhotoScan User Manual: Professional Edition, Version 1.2. ed. A. PhotoScan.

- Rendenieks Z.,Nita M. D.,Nikodemus O.Radeloff V. C. (2020) Half a century of forest cover change along the Latvian-Russian border captured by object-based image analysis of Corona and Landsat TM/OLI data. *Remote Sens Environ*, 249.
- Ruhab S.,Khan M. A.,Hamaz A.,Albarakati H. M.,Saidani O.,Alshardan A.,Alasiry A.,Marzougui M.Nam Y. Y. (2024) A Novel Network-Level Fusion Architecture of Proposed Self-Attention and Vision Transformer Models for Land Use and Land Cover Classification From Remote Sensing Images. *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING*, 17.
- Saleem A.,Corner R.Awange J. (2018) On the possibility of using CORONA and Landsat data for evaluating and mapping long-term LULC: Case study of Iraqi Kurdistan. *Appl Geogr*, 90, 145-154.
- Sarlin P.-E.,DeTone D.,Malisiewicz T.Rabinovich A. 2020. SuperGlue: Learning Feature Matching with Graph Neural Networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*.
- Schenk T.,Csatho B.Shin S. W. 2003. Rigorous panoramic camera model for DISP Imagery. In *High Resolution Mapping From Space*.
- Sertel E.,Ekim B.,Ettehad Osgouei P.Kahadai M. E. (2022) Land use and land cover mapping using deep learning based segmentation approaches and VHR Worldview-3 images. *Remote sens*, 14.
- Sertel E.,Kabadayi M. E.,Sengul G. S.Tumer I. N. (2024) HexaLCSeg: A historical benchmark dataset from Hexagon satellite images for land cover segmentation [Software and Data Sets]. *Ieee Geosc Rem Sen M*, 12, 197 - 206.
- Shahbandeh M.,Kaim D.Kozak J. (2023) Using CORONA Imagery to Study Land Use and Land Cover Change. *Remote Sens*, 15.
- Shahtahmassebi A. R.,Liu M.,Li L.,Wu J.,Zhao M.,Chen X.,Jiang L.,Huang D.,Hu F.,Huang M.,Deng K.,Huang X.,Shahtahmassebi G.,Biswas A.,Moore,N.Atkinson P. M. (2023) De-noised and contrast enhanced KH-9 HEXAGON mapping and panoramic camera images for urban research, . *Science of Remote Sensing*, 7.
- Shahtahmassebi A. R.,Shahtahmassebi G.,Moore N.Atkinson P. M. (2024) Identifying fine-scale archaeological features using KH-9 HEXAGON mapping and panoramic camera images: evidence from Liangzhu Ancient City, . *International Journal of Journal of Remote Sensing*, 45, 5544-5576.
- Sohn H. G.,Kim G. H.Yom J. H. (2004) Mathematical modelling of historical reconnaissance corona KH-4B imagery. *The Photogrammetric Record*, 19, 51-65.
- Song D. X.,Huang C. Q.,Sexton J. O.,Channan S.,Feng M.Townshend J. R. (2015) Use of Landsat and Corona data for mapping forest cover change from the mid-1960s to 2000s: Case studies from the Eastern United States and Central Brazil. *Isprs J Photogramm*, 103, 81-92.
- Surazakov A.Aizen V. (2010) Positional Accuracy Evaluation of Declassified Hexagon KH-9 Mapping Camera Imagery. *Photogramm Eng Rem S*, 76, 603-608.
- Watanabe N.,Nakamura S.,Liu B.Wang N. Y. (2017) Utilization of Structure from Motion for processing CORONA satellite images: Application to mapping and interpretation of archaeological features in Liangzhu Culture, China. *Archaeol Res Asia*, 11, 38-50.
- Xiao P.,Zhang X.,Zhang H.,Hu R.Feng X. (2018) Multiscale Optimized Segmentation of Urban Green Cover in High Resolution Remote Sensing Image. *Remote Sens*, 10.
- Yuan X. H.,Shi J. F.Gu L. C. (2021) A review of deep learning methods for semantic segmentation of remote sensing imagery. *Expert Systems With Applications*, 169.
- Zhang C.,Sargent I.,Pan X.,Li H. P.,Gardiner A.,Haree J.Atkinson P. M. (2018) An object-based convolutional neural network (OCNN) for urban land use classification. *Remote Sens Environ*, 216 57–70.
- Zhang C. Y.Xie Z. X. (2012) Combining object-based texture measures with a neural network for vegetation mapping in the Everglades from hyperspectral imagery. *Remote Sens Environ*, 124, 310-320.
- Zhang L.,Rupnik E.Pierrot-Deseilligny M. (2021) Feature matching for multi-epoch historical aerial images. *Isprs J Photogramm*, 182, 176–189.

- Zhou Y. S.,Chen G.,Qiao X.Lu L. J. (2021) Mining High-Resolution KH-9 Panoramic Imagery to Determine Earthquake Deformation: Methods and Applications. *Ieee T Geosci Remote*, 60.
- Zhou Y. S.,Li Z. W.,Li J.,Zhao R.Ding X. L. (2018) Glacier mass balance in the Qinghai Tibet Plateau and its surroundings from the mid-1970s to 2000 based on Hexagon KH-9 and SRTM DEMs. *Remote Sens Environ*, 210, 96-112.