Towards an AI-Based Time Machine

Mapping the Morphogenesis and Metamorphosis of Urban Fabrics and Blocks: Nanjing 1929

Marco Trisciuoglio, Didem Turk[1]

Abstract: The aim of the paper is to describe an experiment of remote sensing – remote in time rather than remote in space – about the Southern part of Nanjing (China), as portrayed by a 1929 aerial photography. This image is conserved at the Library of Congress Geography and Map Division, Washington D.C. (20540-4650 USA), and is entitled «Nanking, China». The image was taken by the United States Aircraft Squadrons in September 1929, during a survey campaign of the Republic of China to describe its main cities, including the new Capital city. The aerial photography, comprising six boards, shows the urban morphology of Nanjing, which is largely lost today, due to the great transformations that have taken place in the almost last 100 years of urban history. The remote sensing experiment has been driven within the framework of research aimed at understanding urban fabrics and their morphogenesis, involving the investigation and mapping of the origins of urban blocks and their constitutional elements. The project recognizes the urban block as a dynamic structure. The dynamic structure of urban blocks caused adjustments in the formation of an urban block, in some cases showing the metamorphosis of urban fabrics both in social and physical contexts. By mapping the change in urban blocks, the study aims to understand the temporal transition. To map changes in urban blocks, cases are selected in a different context. In the case of Nanjing, historical photography was selected as the main source to understand this change, facilitated by the possibility to compare it with contemporary satellite views. Accessibility of information from historical images is limited. To conduct quantitative and geometrical analysis on settlement characteristics, it is essential to acquire relevant spatial information by automatic geographical feature detection and extraction. The project uses satellite image processing technology to automatically digitize and extract information from historical images, addressing the challenge of vectorizing limited-quality historical maps. The main goal is to reconstruct past spatial characteristics to trace and understand urban patterns. The defined steps included annotating the building footprint using GIS and creating GeoJSON data and masks to serve as input for training the AI model. The preliminary results consisted of the automatic detection of building footprints. This will be used as input to conduct further analysis with parametric design.

Keywords: Urban Morphology, Artificial Intelligence, Remote Sensing, Historical Maps.

1. Background

Urban morphology enables the examination of transformations in settlements, spanning from their initial development to their contemporary formations, across different peri-

1. Politecnico di Torino.

DOI: 10.36158/979125669182111

ods, locations and contexts. The field of urban morphology has evolved and expanded from various perspectives, providing a transformative and evolutionary framework for research. According to Wang *et al.* (2022), mapping urban morphology has played a crucial role in comprehending cities. In recent years there has been a shift in the perception of cities and mapping methods, particularly in terms of their complex form, leading to the new approach that considers cities as data (Rhee *et al.*, 2019).

In a concurrent manner, recent years have witnessed the accelerated adaptation of systematic and more quantitative approaches in urban morphology, aligned with the primary theories in the field. Quantitative morphological analysis commonly involves converting urban elements into numerical indices (Chen *et al.*, 2021). Extensive research has been dedicated to quantifying urban form and its characteristics, establishing quantitative relationships among spatial variables, and defining indicators of urban form (Fleishmann, 2020; Marshall, 2005; Yu, 2014). Furthermore, the tools used for examining morphological patterns vary across different levels of analysis (D'Acci & Batty, 2019).

In light of the dynamic nature of urban form, there is a focus on understanding how changes in urban form can be traced over time. An essential aspect of comprehending and interpreting changes in urban form lies in analyzing the historical evolution of cities. In this context, historical maps play a critical role as invaluable tools for understanding such transformations. However, extracting precise information from historical maps through vectorization poses significant challenges. These challenges include limited geographical information and the absence of metadata unless it can be sourced from archival data (Ekim *et al.*, 2021).

Access to historical maps is often limited, with them typically available only in the form of scanned images, necessitating further processing to enable quantitative and geometrical analysis (Ekim *et al.*, 2021). Among these sources, aerial photography and satellite imagery have become crucial sources for retrieving geographic information. However, vectorizing aerial photography or historical maps can be time-consuming and labour-intensive. Recent advancements have emerged in reducing this labour through the utilization of computers and machines. Rhee *et al.* emphasize the transformation of computers from mere tools to generators of information. According to him, "the possibilities of AI technologies in architectural design are amplified when thoroughly dealing with a data space called a city with complex and innumerable relationships" (Rhee *et al.*, 2019).

The advancement in AI technologies has become invaluable in extracting information from aerial photography, satellite imagery, and other sources. Deep learning models, in particular, can be used in object detection, segmentation, and classification (Wang *et al.*, 2022; Moosavi, 2016). The effectiveness of the model's application depends on good quality input, such as image quality, the level of desired information, completeness and comprehensiveness of the dataset, among others.

This study employs a deep learning model (U-Net) to analyze and read urban forms in Nanjing, China. The study takes historical maps as input for the model to extract relevant information from aerial photographs captured in 1929, which, to our knowledge, is considered the most representative historical image available for the area. Due to rapid urbanization in Nanjing, it has become exceedingly challenging to trace the historical typology of the city. Therefore, the study's objective is to understand and analyze the changes in urban form by examining the information extracted from aerial photography taken in 1929 by United States Aircraft Squadrons.

Applying an AI model to extract urban form involves several essential steps. In the context of this study, these steps were carried out with the assistance of students participating in the Design Studio "Urban morphology, architectural typology, contemporary settlement patterns", held in Spring-Fall 2023 at the Southeast University of Nanjing (School of Architecture), led by Bao Li and Marco Trisciuoglio. The specific topic of the studio was "Urban

Re-generation in Diaoyutai (钓鱼台) Traditional District, Nanjing". The process encompassed the utilization of GIS (Geographic Information System) and a deep learning model to read and compare urban form in Nanjing.

2. Case Study Area

Nanjing, China, has been chosen as the case study area to investigate the transformation of its urban form, primarily due to the substantial changes that have occurred over time, making it nearly impossible to discern historical patterns.

The aerial photography used in this study was provided by aircraft and preserved in the Library of Congress Geography and Map Division, Washington, D.C. 20540-4650 USA, dating back to 1929. The photography covers a vast area within the city of Nanjing. For this work, a specific zone, known as Qinhuai District (named after the water canal that crosses it) or Old Town, has been selected as the focus of the experiment. This selected area exhibits diverse urban forms, particularly in terms of urban blocks (see Figure 1). As described by Shane (2020), changes in social and economic aspects significantly impact the modification or complete transformation of urban form, often observed within the urban blocks.

Nanjing, particularly the Qinhuai District, exhibits a combination of historical patterns and new formations of urban blocks that have recently dominated the area. The city image from 1929 provides a glimpse into the morphological structure of the city, with the block primarily consisting of courtyard houses and narrow surrounding streets. Architectural features during that period were characterized by low-rise courtyard houses. However, a significant transformation becomes evident when comparing the contemporary satellite image taken from Google Earth to the historical one (see Figure 1d).

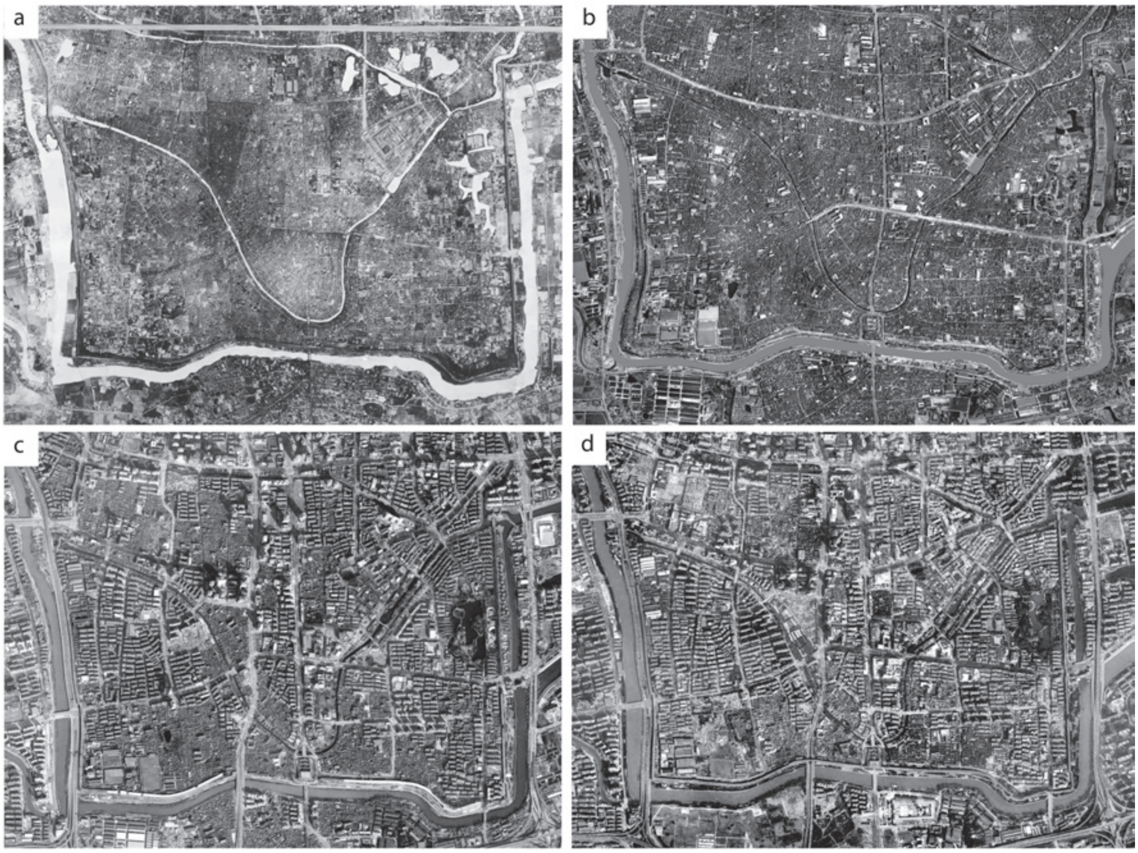

Figure 1. *a. Aerial photography of Old Town 1929. b. Aerial photography of Old Town 1976. c. Aerial photography of Old Town 2005. d. Aerial photography of Old Town age 2017.*

In 1929, the role of the Qinhuai River, the city walls and the great North-South Road, leading to the South Gate of the City (Zhonghuamen) was still prominent. However, by 1976, the dense urban fabric of courtyard houses underwent the first erosion due to the establishment of small/medium factories and service infrastructures. Following the economic reform from the 1990s, the same urban sector underwent heavy redevelopment, as evidenced by the 2005 satellite photo which showed a significant replacement of the original fabric of the Ming and Qing era with the construction of stick buildings, ranging from five to eight floors, characterized by a clear functional imprint. Additionally, a second north-south artery, Zhongshan Lu, was opened, disregarding the pre-existing settlement forms. Subsequent satellite representations from 2017 onward depict the progress of a redevelopment process that seeked to reorganize urban spaces, the housing system and activities, although not always successfully.

3. The Methodology and Application Process

In this project, urban form elements are used as examples to train a model that can automatically extract information from satellite imagery. Wang *et al*. (2022) highlight the natural connection between mapping urban morphology and remote sensing (RS), emphasizing how knowledge about cities can be obtained through this approach. Earth Observation (EO) technology advancements have significantly enhanced our capacity to observe cities from an aerial perspective. Analyzing tangible physical characteristics captured in imagery data using remote sensing techniques makes it possible to identify and study various aspects of the urbanization process at different scales and levels (Wang *et al.*, 2022). Regarding morphological analysis, crucial elements discernible from remote sensing data include the primary building typology, the arrangement of courtyards, and the street network. It is important to note that the quality and scale of the images play a vital role in effectively extracting such information.

As part of the master studio project, a workshop was conducted, encompassing two primary steps: teaching and experimenting. The initial phase involves introducing the notion and concept of urban form and urban block elements to the students. They were familiarized with various typologies observed in the area, which enabled them to utilize this knowledge in a subsequent extraction process. In this case, the understanding and digitization of urban form elements became a critical role as the initial step of the project.

The experiment's second step involved demonstrating the AI model's working mechanism. AI models require information to process and learn from, referred to as supervised training. Ekim states that "supervised classification approaches require reference data to be used in the training stage, which is also called ground truth masking, for the corresponding input images" (Ekim, 2021: 5). With the advancements in data science methods, particularly machine learning and deep learning applications, predictions (such as automatically detecting urban form) are increasingly derived from images by leveraging their inherent features (Wang *et al.*, 2022).

In the context of this project, specific information is provided to the model in a particular format, enabling the model to learn and recognize similar patterns when encountered. This learning process is known as training, wherein the model is exposed to a certain amount of data. Once trained, the model can be tested on unseen data with similar structures. In this case, the data presented in the model comprises aerial photography from 1929. A part of the Old Town area is selected on the map for training purposes.

Clarifying the data structure assumes great importance in this project. To enable the model to learn from aerial imagery, it is necessary to provide the model with both the image and a corresponding mask. The mask is created using provided labels that contain the specific information to be detected. In order to generate these labels, a manual process is followed in the QGIS environment, with the assistance of the students involved in the project (see Figure 2 for digitized building footprints).

The selected area is divided into 16 sections, with each student assigned a specific section to label the buildings using QGIS software. A step-by-step introduction is made to the students to ensure a comprehensive understanding of the process of vectorizing, which involves labelling and annotating the building footprints.

Following the labelling process, the creation of the mask is carried out using Python scripts along with the images and polygon in GeoJSON data format, where geographic information is stored with polygons. "This process is called reprojection and assists in matching the coordinate system of the image and the mask vector so that each pixel in the image matches with its corresponding point in the mask or vice versa" (Ekim *et al.*, 2022: 5).

Once the images and corresponding masks are prepared, they are introduced to the model for training and learning based on the dataset. The masks are defined with binary classification, meaning only two classes are introduced: buildings and unlabeled areas (see Figure 3). This binary classification allows the model to differentiate between the presence and absence of buildings in the images during the training process.

Training the model requires substantial computing resources, making it challenging to train the model using a single large image due to memory limitations. To address this issue, the commonly adopted approach is to use smaller patches of images of 256x256 pixels for training purposes. The data[2] (total amount of patches) is divided into three main sections: training, validating and testing.

In the case of the map image in this project, due to its resolution, the map is divided into smaller patches of 128x128 px (see Figure 3). A total of 64 patches are created, each with its corresponding mask. The validation dataset comprises 20% of the images, while the remaining images are used for the training dataset. The test dataset is provided separately for evaluating the trained model's performance.

To effectively train the model, a large amount of data is required. One method to increase the dataset size is through augmentation techniques. These techniques artificially expand the dataset by applying basic image processing techniques such as flipping, cropping, adding noise, adjusting brightness, blurring, etc. By augmenting one image with different variations, such as rotating to the different angles or above mentioned techniques, a single image is transformed into eight images (see Figure 4). In this project, augmentation techniques are applied, resulting in a total of 459 images created from the original 51 images in the training dataset.

The U-net deep learning model is employed for the training process. "UNet, evolved from the traditional convolutional neural network, was first designed and applied in 2015 to process biomedical images. As a general convolutional neural network focuses its task on image classification" (Zhang, 2019). Due to its effectiveness in capturing detailed spatial information and accurately delineating object boundaries, the model can be easily adapted to analyze satellite imagery. Its architecture allows for extracting meaningful features from the images, enabling the model to identify and classify different urban form elements.

The training process is carried out in a virtual server environment provided by the CRIB platform from the University of Twente, Faculty of Geo-Information Science and Earth Observation. The choice is made because training deep learning models with requires robust and powerful servers to handle the computational load efficiently.

2. The data defined here is structured in two formats: images, which are generated by dividing the map into smaller patches, and masks for these images, created through manual vector labeling of buildings and blocks. Both image and mask formats are used during the training process.

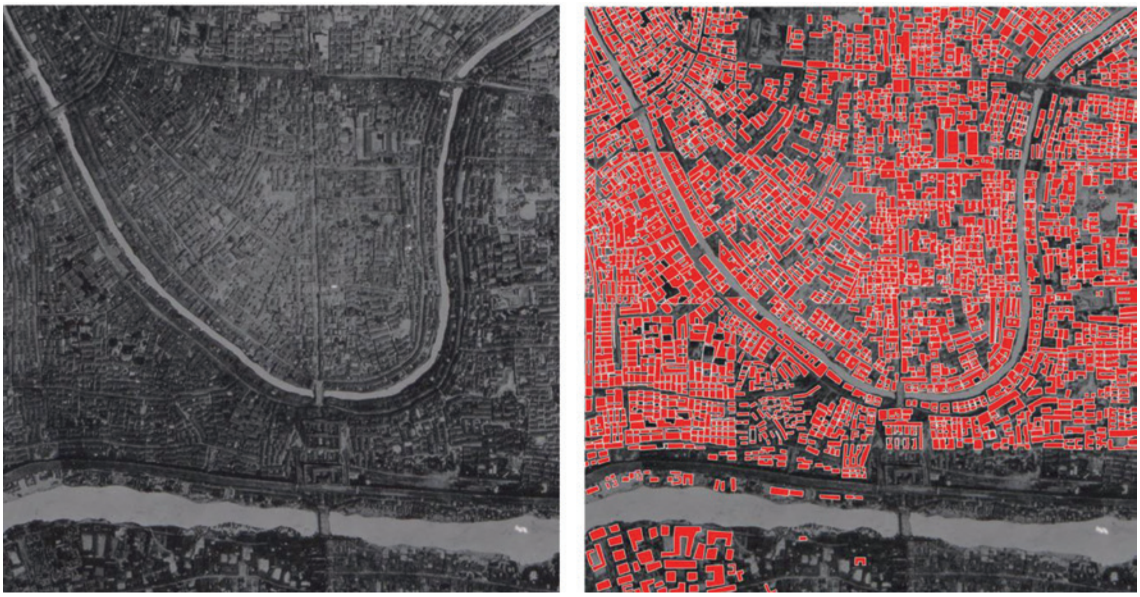

Figure 2. *Building labels drawn by the students.*

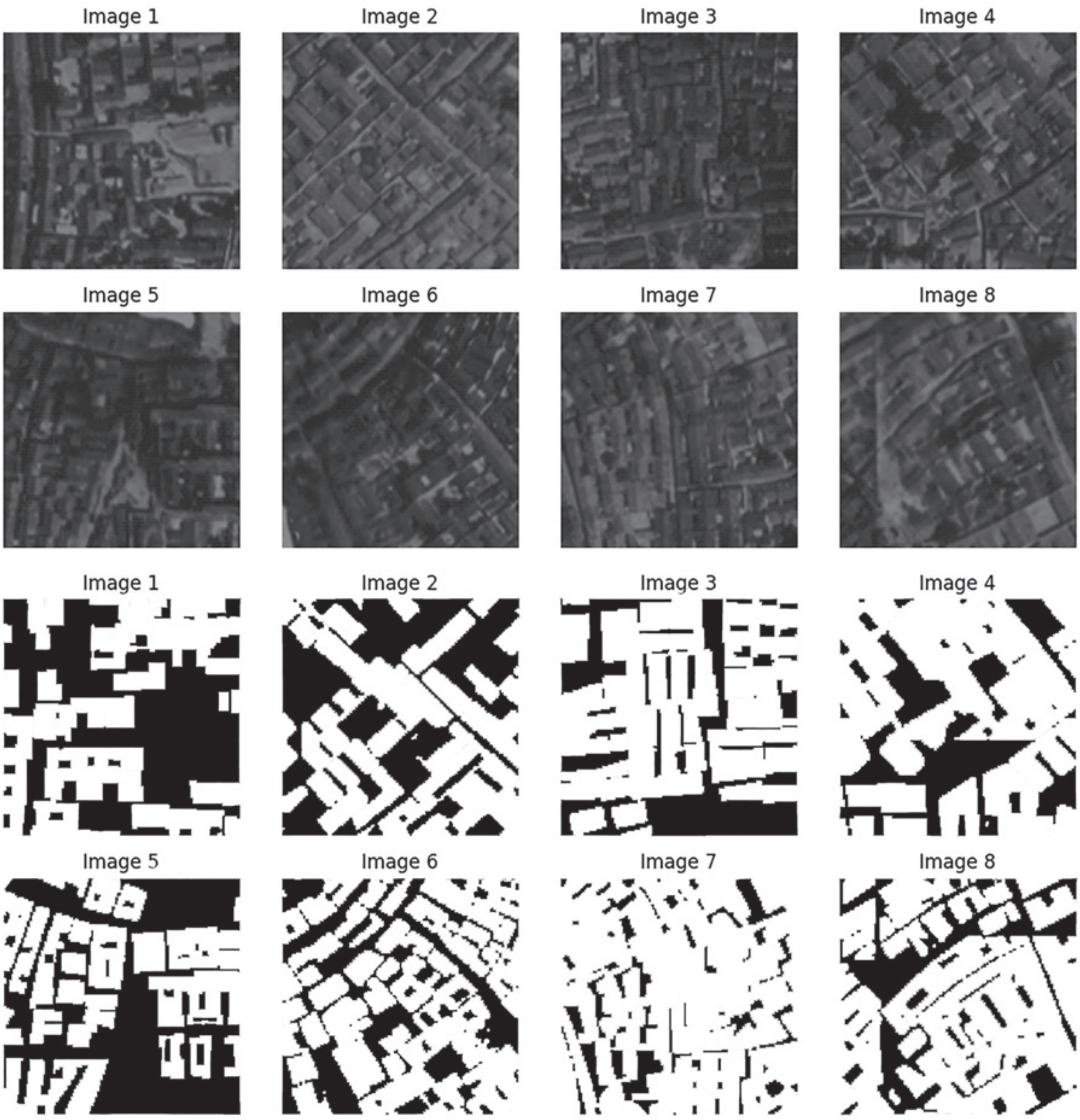

Figure 3. *Creating masks from images based on label: white color – buildings, black color – unlabeled.*

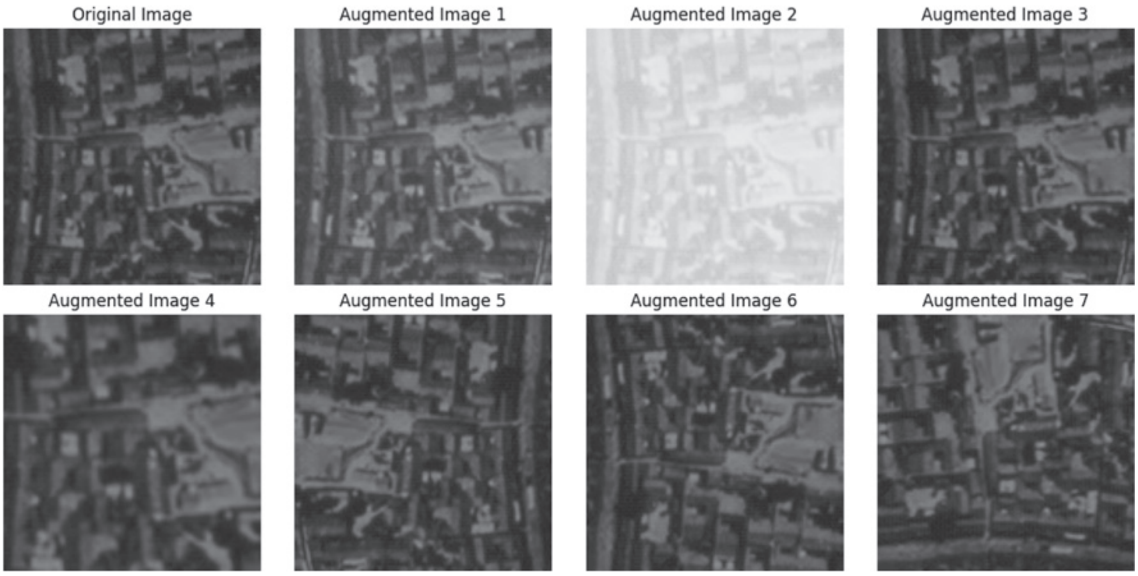

Figure 4. *Augmentation of the images.*

To enhance the performance of the model, pre-trained weights are utilized. This involves incorporating pre-trained Convolutional Neural Networks (CNNs) into the semantic segmentation architecture. By leveraging pre-trained weights obtained from models trained on large-scale datasets like ImageNet, the model benefits from the knowledge and feature extraction capabilities learned during the pretrained training process. This is particularly advantageous for extracting high-level features from the input data, such as aerial imagery, and improving the overall performance and accuracy of the model (Ekim *et al.*, 2021).

4. Results and Step Ahead

The training results are evaluated using graphs and accuracy metrics to assess the model's performance. Although the training results demonstrate slightly lower quality than expected, likely due to resolution limitations, it is observed that the model is capable of detecting building footprints (Figure 5)[3].

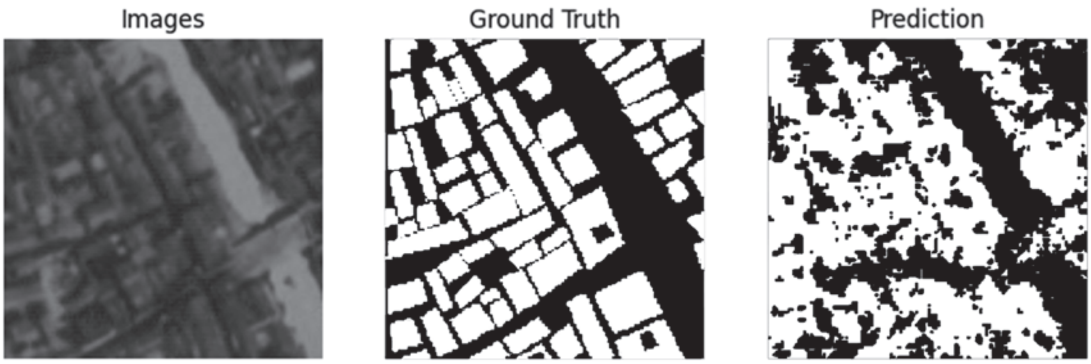

Figure 5. *Training result, ground truth (expected detection), prediction (detection by the model).*

To validate the model's performance, it is tested on a separate set of 24 test data images to predict the building footprints. Post-processing techniques are applied to the test data

3. The image illustrates the utilization of labeled data and the model's predictive results. As previously mentioned, the image, along with its corresponding mask, is used for training the model. Once the model has seen the image and its corresponding mask, it can learn from them. Consequently, when the model encounters the same image or similar versions, it can make predictions and successfully detect buildings.

to refine the results. Firstly, the individual automatically detected test images are merged into a single TIFF file and then detected pixels are converted into shapefiles to process further. Secondly, since the model's predictions are pixel-based, a simplification process is conducted using the QGIS platform to ensure smooth and accurate lines for the detected building footprints. The goal of this post-processing step is to enable the repeated testing of the model on different parts of the dataset, ultimately automating the generation of the entire map.

While the model's performance requires fine-tuning, the preliminary results are promising. The model could predict an area it had never seen (see Figure 6, upper left). We received lower detection quality due to color contrast and low resolution of images (both for training and testing). Nevertheless, it is still possible to observe recognition of form and, in some cases, clear continuity in block structure between training and test images (see Figure 6). Since the model was not trained on river or street elements, there was no detection or confusion between these elements and building footprints. As a result, the river path is visible alongside the detected buildings. The following steps will involve adjusting the training data to achieve better prediction results to read and measure the pattern.

We selected the predicted test area to observe the transition and the changes in block typologies over time. The maps (see Figure 7) illustrate the stages of transformation in the area in four stages, from 1929 to 2017. The process resulted in the loss of the organic block structure and the implementation of the modern block structure.

The change in density can be observed, but it is not possible to measure and compare them with raster images. In addition to the aerial photo of 1929, the cadastral map of Nanjing from 1936 reveals the city's primary structures, which are incredibly rare in the Chinese context. This map was created as part of a process that commenced with the establishment of the Republican government in 1912. It enables us to analyze changes in the area from the 1929 aerial photo to more recent maps from 2005. However, many details,

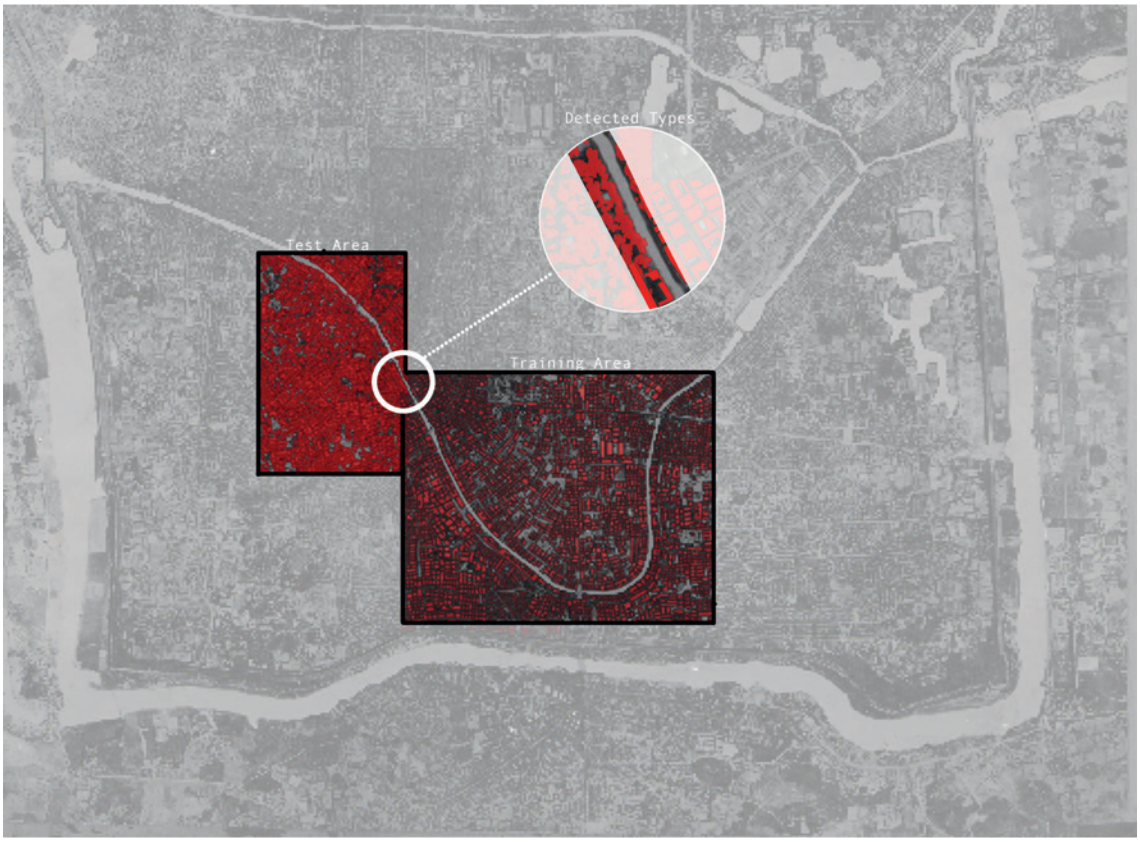

Figure 6. *Training and test area.*

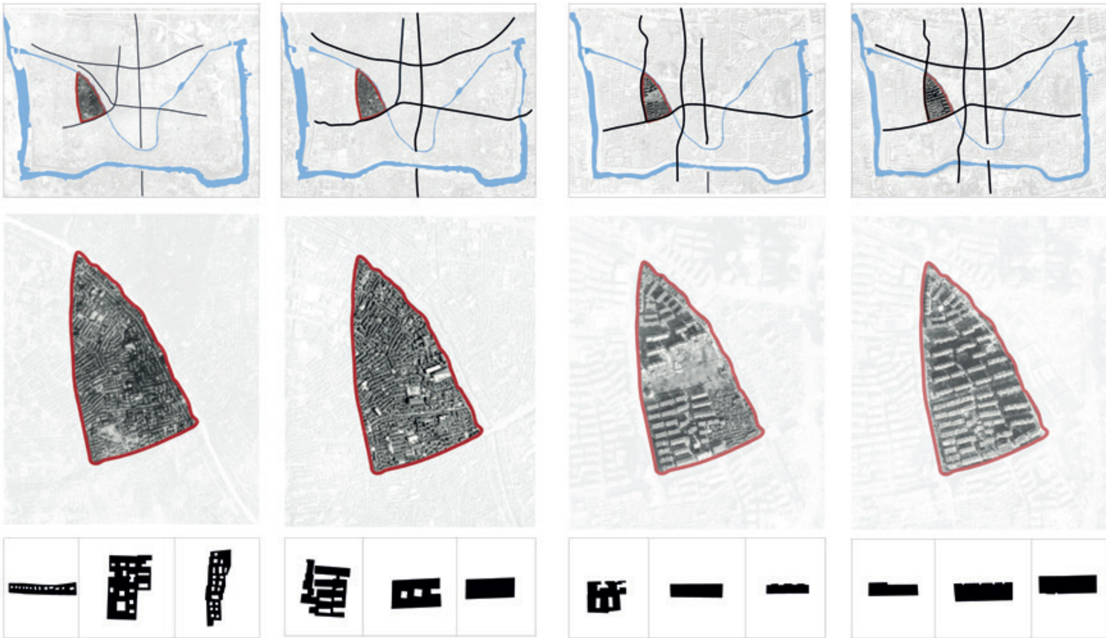

Figure 7. *The morphological transition of the predicted area, from left to right, from the map of 1929, 1976, 2005 and 2007.*

such as plot divisions, land use, and block structures from 1929, remain unavailable. This highlights the importance of applied methods. By detecting building footprints in the area, further metric analysis can be conducted to better understand the organic pattern and its transformation in time.

Conclusion

The research highlights the gap in generating information from historical maps. The research presents an AI approach to extract information from historical maps. The morphological process and transformation of the Old Town in Nanjing are presented. The remote sensing approach is applied to map and read urban forms. The project draws a new perspective on reading the transition of urban form with the application of new technologies.

References

Bao, L. *et al.* (2017), *Typological Permanencies and Urban Permutations. Design Studio of Re-generation in Hehua Tang Area*, Nanjing: SEU Press.

Chen, F., Thwaites, K. (2013), *Chinese urban design: The typomorphological approach*, London: Routledge.

Chen, W., Wu, A., Biljecki, F. (2021), "Classification of urban morphology with deep learning: Application on urban vitality", *Journal of Computers, Environment and Urban Systems*, https://doi.org/10.1016/j.compenvurbsys.2021.101706.

D'Acci, L., Batty, M. (2019), *The Mathematics of Urban Morphology*, New York: Springer International Publishing.

Ekim, B., Sertel, E., Kabadayı, M.E. (2021), "Automatic road extraction from historical maps using deep learning techniques: A regional case study of turkey in a German World War II Map", *ISPRS International Journal of Geo-Information*, 10 (8): 492.

Fleischmann, M., Romice, O., Porta, S. (2021), "Measuring urban form: Overcoming terminological inconsistencies for a quantitative and comprehensive morphologic analysis of cities", *Environment and Planning B: Urban Analytics and City Science*, 48: 2133-2150.

Jiang, L. (2019), *Morphological Research of the Historical Urban Boundary. The Inner Fringe Area of Nanjing*, Nanjing/Torino: "Transitional Morphologies" Joint Research Unit / Southeast University – Politecnico di Torino.

Malfroy, S., Caniggia, G., Powell, M. (2021), *A Morphological Approach to Cities and Their Regions*, Zürich: Triest.

Moosavi, V. (2017), "Urban morphology meets deep learning: Exploring urban forms in one million cities, town and villages across the planet", *arXiv*, https://doi.org/10.48550/arXiv.1709.02939.

Oliveira, V. (2016), *An introduction to the study of the physical form of cities*, Cham: Springer.

Rhee, J., Llach, D.C., Krishnamurti, R. (2019), "Context-rich Urban Analysis Using Machine Learning", *City Information Modelling and GIS*, 3: 353-352.

Shane, D.G. (2020), "Block, Superblock and Megablock: A Short Morphological History", in Johnson, J., Brazier, C., Lam, T., *The China Lab Guide to Megablock Urbanisms*, New York: Actar Publishers & Columbia University GSAPP: 161-207.

Trisciuoglio, M., Barosio, M., Ricchiardi, A., Tulumen, Z., Crapolicchio, M., Gugliotta, R. (2021), "Transitional Morphologies and Urban Forms: Generation and Regeneration Processes – An Agenda", *Sustainability*, 13, 6233.

Trisciuoglio, M. (2021), "Note on the transitional urban morphologies as a criticism of urban studies in the Chinese context. The typological map on Nanjing Hehua Tang", *U+D. Urbanform and Design*, 14: 14-27.

Wang, J., Georganos, S., Kuffer, M., Abascal, A., Vanhuysse, S. (2022), "On the knowledge gain of urban morphology from space", *Computers, Environment and Urban Systems*, 95, 101831.

Ye, Y., Van Nes, A. (2014), "Quantitative Tools in Urban Morphology: Combining Space Syntax, Spacematrix and Mixed-Use Index in a GIS Framework", *Urban Morphology*, 18(2): 97-118.