The Methods of Deep Learning and Big Data Analysis in Promoting Sustainable Architecture

H Yazdi¹,*, I Vukorep¹ and H Bazazzadeh²

¹ Faculty of Architecture, Brandenburg University of Technology, Germany
² Faculty of Architecture, Poznan University of Technology, Poland

*hadi.yazdi@b-tu.de

Abstract. These days, sustainability in different aspects has been among the main discussions of architecture and building science. At the same time, historic architecture has evolved over centuries and has adapted to environmental conditions, it can be a great source of inspiration in using smart ways to achieve sustainable architecture. A good illustration of this adaptation can be found in using vernacular materials, the spatial configuration according to climate conditions, and different elements of historic architecture that have helped to improve the occupant’s comfort. In response, one plausible solution for improving the sustainability of architecture is translating the concept of the sustainable elements and features of historic architecture to be used in contemporary architecture. Therefore, these elements need to be studied thoroughly to comprehend their features and characters. There are several studies, investigating sustainable historic architecture to find and measure sustainable solutions by using conventional methods. Although the accuracy of studying the sustainable historic elements has been fairly high, the number of features and variety of these elements in historic architecture have made this task highly challenging. It has been suggested to study and evaluate a considerable number of these elements in different historic architecture to reduce the errors and increase the reliability of results. Since the conventional methods are labor-intensive, time-consuming, and costly, this paper proposed a robust AI method to study the sustainable elements of historic architecture by using Deep Learning. In this study, by introducing and developing a new method for detecting sustainable elements in historic architecture, their features were comprehensively extracted by means of mining meaningful data from areal images of historic cities to produce big data. The proposed method has a sophisticated workflow starting from subdividing the High-Resolution Aerial Images to detecting the sustainable elements and using data science to analyze the extracted features of the segmented objects. Results of a sample analysis of this method showed its high accuracy and its applicability in analyzing sustainable elements of historic architecture, by which designers are expected to design more sustainable buildings inspired by historic architecture.

1. Introduction

For centuries, people have adapted and improved their housing to the climate. A central courtyard house is one of the types of housing that proved to be suitable for hot climates. It is characterized by the presence of rooms surrounding a yard, separating the living area from harsh environmental conditions and can be traced back to ancient civilizations in Iran, but also other parts of the Middle East and North...
Africa[1], [2]. In newer days modern courtyard houses can be found all around the globe. La Hacienda Jardín in Mexico from Práctica Arquitectura, House in Ecuador by Estudio Felipe Escudero, Hotel San Miguel Allende in Chile by architect Ian Pablo Amores are some examples of contemporary courtyard houses that deal with clime issues in their countries.

In Iran, a study by Soflaei et al.[3], found that courtyard houses also contain additional natural features of about 17%, which can be divided into water (8%) and plant (9%) elements. Studying six houses from Kerman and Isfahan, the authors found fourteen parameters that determine the design pattern of courtyard houses. Further, they provide recommendations for future designs of courtyard houses based on the geometrical proportion logic of the existing courtyard houses. In addition, linear fitting was used to determine the best line fitting of the scatter plot. Using the calculated ratio based on observation and prior research, these houses are more comfortable for the occupants. A passive cooling principle based on geometry and orientation along with embedded natural elements was the subject of their quantitative study. The authors have combined the results into one optimized central courtyard model based on its physical-environmental behavior.

Computing techniques have dominated the domains of structural analysis, design, and urban development in various fields. One of these techniques involves creating digital data by analyzing visually the surrounding and transforming the captured image. This technology’s biggest advantage is its ability to automatically classify objects from images. Over the last few years, machine learning (ML) and the development of deep neural networks has pushed this development. For example, AlexNet was developed in 2012 [4]. There are new research fields arising from airborne and satellite data, especially in the field of archaeology, that are helping to provide deeper information and discover patterns that were previously unknown [5], [6]. Even though the technology is available, there is still a void in architectural research. We are confronted with the reality of analog labeling and visual inspection when the idea of automatized data collection and analysis from areal images is considered. As a consequence, big data is a problem for researchers, since they cannot process large volumes of data using classical methods. Traditional techniques cannot keep up with the increasing quality of new information, as academics employ the same strategy for all images, such as making visual choices based on their prior knowledge of the data. As a solution to these issues, imaging operations associated with extracting characteristics from existing designs have to be automated [7]–[9].

In this paper, we introduce methods and a workflow for automatic feature extraction of sustainable architectural elements or typologies. Additionally, a case study that is presented here is used to test the proposed techniques. We describe single steps of how the deep learning method is implemented, so the technique can be applied to other types of buildings and urban structures. Through the use of these AI-aided methods, the concept of sustainable elements and characteristics of historical architecture can be translated into contemporary architecture, with the aim of improving the sustainability of future buildings. Through the use of these methods, further investigation of buildings can be undertaken to help with the future development of new courtyard architecture based on vernacular patterns.

2. Methods
In studying ancient cities, automatic identification and segmentation of elements in aerial and satellite photographs using ML techniques has become increasingly popular. There are number of researches and industrial fields that implements neural networks (NN) for image classification, including aerial photographs, gait recognition, medical images microorganisms classification, diverse objects and urban environment recognition [10]. In architectural heritage utilizing techniques such as retrieval [11], pattern detection [12], computer vision algorithms [13], Gabor filters, and support vector machines [14], local feature learning and clustering[15]. [16] is getting more common. Nevertheless, automated feature extraction does not eliminate the need for researchers. Its purpose is to serve as an alternative method
for scholars to rapidly assemble a collection of characteristics of vernacular architecture elements and gain methods to improve sustainable architecture and urbanism.

The proposed process is based on our previous research [17] and can be divided into five stages and four processes.

Figure 1. The processes involved in applying machine learning to the case study are: 1) courtyard detection, 2) courtyard segmentation, 3) extraction of features, and 4) data analysis. Designed by miro website [17].

2.1. Primary source data collection
A valuable primary source of input data is a collection of very high resolution (VHR) aerial images. As such data comes in diverse formats a preprocessing procedure is necessary as described in [17]. The unified images will finally be divided into subset of smaller images that vary between 1800*1700 to 3300*3000 pixels. This is necessary due to limiting the size of processed images during the training and categorization process. Subdividing the main aerial images has a limitation. It would be the missing samples that are located at the borders of the cropped images. For example, there are some samples exactly on the cutting line, so they will be missed in the subdivided images. Further, an image labeling system has to be developed to enable subsequent processing and any necessary image copies and transformations to be traced and geographically located. In each step, the images are cropped and named automatically by a Python script according to the location of the cropping function. For example, the main aerial image (XXX.jpg) may be subdivided into 6 columns and 10 rows, so the cropped images of row 2 and column 3 will be named XXX_2_3.jpg. Furthermore, the 7th detected element in the subdivided image will be named XXX_2_3_7.jpg. Accordingly, each element has a specific name for further recognition.

2.2. Detection Process
VHR aerial images differ in many ways from natural photography in detecting objects because of occlusion, shadows, and low resolution. According to the previous studies, there are two main methods for detecting the objects in VHR images. One is the traditional methods that used handcrafted factors [18] and second approach is using deep learning method [4][19]. We are using annotation tools like Remo_app [20] for the training and evaluation tasks as the natural features in courtyard houses are easily detectable and annotatable. Specifically, we used PyTorch toolkit, which is an open source machine learning python framework that is fundamentally a mathematics toolkit that enables fast computation and automatic differentiation of graph-based models. To expand our training dataset, we also had to apply data augmentation. In the augmentation task, the machine increases the dataset variety by rotating, flipping, shifting, and scaling the images. Furthermore, the training dataset should include variations with distinct characteristics of the observed elements, such as elements from different geographic locations, to enhance recognition capabilities.
We describe the training details in [17] where the object detection is done with the COCO-train2017 pre-trained model and further trained on FasterRCNN-ResNet-50 Backbone. This process is explained in [21].

2.3. Segmentation Process
The segmentation is necessary as it will extract each found element into its own image. The feature extraction is often a multilevel task. The element will be observed in micro and macro scale. For the urban constellation neighboring images are important as well as for recognition of sub elements inside or near the observed object. The order of the segmentation process depends on these micro and macro dependencies. So one workflow scenario can be a primary search on some very specific identifier, their extraction and widening the frame around the object for identifying other adjunct features (figure 2). As previously mentioned, a reasonable file naming is crucial for a good post processing.

![Figure 2](image.png)

**Figure 2.** A demonstration of the data in both datasets after cropping the images [17].

Moreover, the segmentation was carried out using the Mask R-CNN model. The suggested approach recognizes items in an image in a short amount of time while concurrently constructing a segmentation mask. When compared to Faster R-CNN, which operates at a rate of five frames per second, Mask R-CNN is simple to train and offers very little in the way of additional burden. Another benefit of using the identical foundation to estimate human posture is that Mask R-CNN is readily broadly applicable to other applications [22].

2.4. Feature extraction
After a high number of data has been gathered a feature extraction can be performed. This is a process that uses classical image processing algorithms for geometric features detection like size, area, volume, proportion, direction. The segmented polygons of each object should be merged to have one polygon for each object and some tools like the minimum bounding box will find the minimum rectangle which
is fitted to the object. This tool is helpful for rectangular objects like the courtyard to find the accurate direction and dimensions of the object.

Some other features can be further extracted by the previously described segmentation method to recognize and analyze containing elements. In the later described case study, known features were extracted out of single courtyard images as shadows (for height estimation), green and water areas, and the surrounding house. Outlines with geometrical information are further automatically executed and added.

The geographical, climatic, and socio-economic factors that are directly associated with each of these elements are hugely important. For example, in the case study of this research, the climatic information of each city is extracted from the 14 years of weather data and added to the courtyard samples according to their city label. Furthermore, it is possible to extract anonymized socio-economic data of each city or neighborhood from the yearly statistics by governments. For instance, the population, family size, family wealth, jobs of occupants.

2.5. Data analysis

The preliminary processing of data would be the first stage in the analytical procedure. For the duration of the preprocessing of the data, it is necessary to locate and remove any outliers or characteristics that were not properly detected. For the purpose of illustration, elements that have a surface area that is smaller than some predetermined threshold will be eliminated from the dataset when these components are detected. Wrong data can be caused by many things, but in this case, a 10% mistake in object recognition and a 10% mistake in picture segmentation led to wrong data in about 20% of the dataset.

During the whole of the phase devoted to the analysis of the data, an effort was made to identify connections between the characteristics of the dataset. In order to find the strongest link amongst different variables, we used correlation analysis. Here's how to figure out Pearson's correlation coefficient, which is used to measure how closely two variables are linked:

\[
r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}} = \frac{s_{xy}}{s_x s_y},
\]

The covariance of x and y is denoted by the symbol Sxy, while the standard deviations of x and y are denoted by the symbols Sx and Sy in that order. This formulation transforms the linear correlation coefficient, r, into a value between 1 and -1. In scatter plots, data points cluster around a line when r is close to one, which indicates the two variables are highly correlated. As r increases, the data points become more dispersed. Data points are mainly distributed horizontally when r is close to zero, suggesting that the variables do not have a linear relationship [23]. Finally, a mathematical model for the highly correlated variables in a dataset could be found. The most suitable approach for determining the model is Linear Regression, which involves fitting a line to the scatter diagrams. Several supervised and unsupervised machine learning methods are used to discover relations between features, classify them, cluster them, and build a prediction model. For example, the Sklearn library includes SVM, Naive Bayes, Logistic Regression, Decision Tree, Random Forest, and K-means. When a dataset contains more samples, the prediction model's output will be more accurate.

3. Case Study

The investigation of prehistoric Iranian inner courts is the subject of numerous of ongoing research projects in Iran. The investigators have conducted a systematic review in which they are contrasting several aspects of inner courts and residences, such as their alignment and proportion [3], [24]–[27]. On the basis of the most recent research [17] a situation is described in which the authors used machine learning algorithms to analyze aerial VHR photography using data gathered from inner courts. This
research collected and evaluated a substantial quantity of information on the features of inner courts in order to identify heretofore unrecognized meteorological and geometrical factors that had a role in the planning of inner courts. As part of this investigation, a case study was conducted in which observations were made of nine cities located throughout Iran and categorized given the Koppen climate classification system. Due to the fact that each of these cities has a unique environment, the samples will exhibit a wide range of characteristics depending on the climatic zone in which they were collected. Arid/steppe and arid/desert climatic zones (i.e., BS and BW, respectively) divide these cities. BS and BW climates each feature two subclass climatic zones known as "h" and "k" that fall under the chilly and hot classifications, correspondingly. As a result, there are 4 climate categories: Bsh, Bsk, Bwh, and Bwk.

3.1. Central courtyard detection

For the detection it was necessary to subdivide aerial images and prepare them for the training step. In a nutshell, the process starts with exceptionally high-resolution aerial photos and the Fast RCNN approach for locating the inner courts. The system is trained using statistics from all nine districts. The outcome of the object recognition was determined established upon the assessment dataset, and that finding shows a reliability of almost 90% on the assessment dataset. Figure 3 illustrates several examples of the test dataset after they have been subjected to the final trained model. After detecting courtyards in large datasets, the entire dataset was cropped to single courtyard images. As described in the methods chapter, the first step was to identify the courtyards and in the second step the belonging house. For this, an offset of the image was necessary. To end with, the court recognition and crop function are responsible for the production of two distinct datasets: the first dataset contains just the courts, while the second dataset, which includes the home that belongs to the court, is referred to as the offset dataset. As further courtyard segmentation can only be performed on the pure courtyard file, and the offset file contains unnecessary data, both images have to be saved.

Figure 3. The illustrations of the authentication dataset after the court recognition mission [17].
3.2. **Image segmentation**

When the model is trained using the two datasets of courtyards and courtyards offset, the model is then assessed using a test dataset that constitutes 20% of the total. The results of the examination show that courts, shadows, and homes attain a reliability of more than 90 percent when it comes to categorization, bounding-box, and masking missions. Water and vegetation, on the other hand, had an 80 percent reliability rate. Figure 4 depicts many instances of assessment datasets that were generated after the segmentation operation was completed.

![Image](image.png)

**Figure 4.** These images show the final test of the model a) The courtyard (blue), shadow (red), green regions (yellow), and water places (green) in the court dataset. b) The segmentation of courtyard (red) and house (blue) in the courtyards_offset datasets [17].

3.3. **Feature extraction and Dataset gathering**

The last phase was to extract features and gather datasets in producing the desired complete datasets, which consisted of approximately 26,000 samples and 76 features. The minimal boundary box for courts, residences, and water sites was then established for additional measuring [28]. Furthermore, the size and orientation of the minimal bounding box were calculated in order to extract the geometrical characteristics of the item that was picked. The dataset currently contains many meteorological characteristics for each one of city, in place of the geometric elements that were previously included. The meteorological data for every single city was then extracted from the EnergyPlus Weather Format (EPW) files. Environmental Protection Agency (EPW) records comprised 14 years' worth of climate data for every one of the cities (2004-2018).
4. Conclusion
This paper is presenting one possible approach towards discovering patterns for sustainable architecture from already existing architectural and urban systems with deep learning techniques and big data analysis beyond classical methods. We present a workflow for collecting data of a specific typology on a large scale independent of its geographic distribution, extracting their geometric features, and connecting these with local information, such as the geography and economy. Furthermore, a case study of the creation of an exemplary dataset of courtyard houses in Iran illustrates the applicability of this approach. Unexpectedly, the research found that such large-scale pattern recognition in architecture and urbanism are understudied and that no publicly available datasets exist.

The obtained data can help in analyzing specific sustainable typologies, finding unknown relationships, as well as in predicting the best configuration for buildings in a given region and based on specific use requirements. Due to the model's reliance on existing architectural types, urban structures, and other available information, it is essential to develop a diversity of attached features. Further work will concentrate on the systematic development of an expandable dataset framework so researchers can enrich the existing dataset with their own findings. Additionally, a connection can be tested between extraction of geometrical features and more detailed layers of information, such as floor plans. Eventually, such trained models can be used to develop an automatic floor layout recommendation system.

In order to reduce errors and increase the reliability of the results, we propose a further study and evaluation of a large number of elements in different historical architectures that have a high sustainability impact, thus reducing error rates, so that results are more reliable. With the proposed deep learning method combined with traditional image analysis tools, we have developed a sophisticated workflow, subdividing the high-resolution aerial images, detecting the sustainable elements, and then analyzing the extracted features using data science. Furthermore, a possible future research project could be using our dataset and adding exact simulated microclimate information of each courtyard. By this new dataset, it is possible to generate a data-driven prediction model for courtyard design counting a thousand microclimatic data in addition to geometric features. In this way, we hope to make sustainable buildings in harmony with historic vernacular architecture more accessible to architects and designers. The software and datasets are publicly available on GitHub (https://github.com/hadi-yazdi/yazd).

References
[1] P. Keshkaran, “Harmonization Between Climate and Architecture in Vernacular Heritage: A Case Study in Yazd, Iran,” Procedia Eng., vol. 21, pp. 428–438, Jan. 2011, doi: 10.1016/j.proeng.2011.11.2035.
[2] G. Memarian, “House typology in Iran (with special reference to Shiraz),” Ph.D., The University of Manchester, 1998. Accessed: Nov. 09, 2020. [Online]. Available: https://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.505868
[3] F. Soflaei, M. Shokouhian, and S. M. Mofidi Shemirani, “Traditional Iranian courtyards as microclimate modifiers by considering orientation, dimensions, and proportions,” Front. Archit. Res., vol. 5, no. 2, pp. 225–238, Jun. 2016, doi: 10.1016/j.faro.2016.02.002.
[4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in Advances in Neural Information Processing Systems, 2012, vol. 25.
[5] R. Lasaponara and N. Masini, Eds., Satellite Remote Sensing: A New Tool for Archaeology. Springer Netherlands, 2012. doi: 10.1007/978-90-481-8801-7.
[6] S. J. Leisz, “An Overview of the Application of Remote Sensing to Archaeology During the Twentieth Century,” in Mapping Archaeological Landscapes from Space, D. C. Comer and M. J. Harrower, Eds. New York, NY: Springer, 2013, pp. 11–19. doi: 10.1007/978-1-4614-6074-9_2.
[7] R. Bennett, D. Cowley, and V. D. Laet, “The data explosion: tackling the taboo of automatic feature recognition in airborne survey data,” Antiquity, vol. 88, no. 341, pp. 896–905, Sep. 2014, doi: 10.1017/S0003598X00050766.
[8] K. Lambers, W. B. Verschoof-van der Vaart, and Q. P. J. Bourgeois, “Integrating Remote Sensing,
Machine Learning, and Citizen Science in Dutch Archaeological Prospection,” Remote Sens., vol. 11, no. 7, Art. no. 7, Jan. 2019, doi: 10.3390/rs11070794.

[9] A. Traviglia, D. Cowley, and K. Lambers, “Finding common ground: human and computer vision in archaeological prospection,” AARGNews Newsl. Aer. Archaeol. Res. Group, vol. 53, p. 11, Sep. 2016.

[10] L. Liu, H. Wang, and C. Wu, “A machine learning method for the large-scale evaluation of urban visual environment,” ArXiv160803396 Cs, Aug. 2016, Accessed: Nov. 03, 2020. [Online]. Available: http://arxiv.org/abs/1608.03396

[11] A. Goel, M. Juneja, and C. V. Jawahar, “Are buildings only instances? exploration in architectural style categories,” in Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing, New York, NY, USA, Dec. 2012, pp. 1–8. doi: 10.1145/2425333.2425334.

[12] W.-T. Chu and M.-H. Tsai, “Visual pattern discovery for architecture image classification and product image search,” in Proceedings of the 2nd ACM International Conference on Multimedia Retrieval, New York, NY, USA, Jun. 2012, pp. 1–8. doi: 10.1145/2324796.2324831.

[13] N. Oses, F. Dornaika, and A. Moujahid, “Image-Based Delineation and Classification of Built Heritage Masonry,” Remote Sens., vol. 6, no. 3, Art. no. 3, Mar. 2014, doi: 10.3390/rs6031863.

[14] M. Mathias, A. Martinovic, J. Weissenberg, S. Haegler, and L. Van Gool, “Automatic architectural style recognition,” ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci., vol. XXXVIII-5/W16, pp. 171–176, Sep. 2012, doi: 10.5194/isprsarchives-XXXVIII-5-W16-171-2011.

[15] G. Shalunts, Y. Haxhimusa, and R. Sablatnig, “Architectural Style Classification of Building Facade Windows,” in Advances in Visual Computing, Berlin, Heidelberg, 2011, pp. 280–289. doi: 10.1007/978-3-642-24031-7_28.

[16] Z. Xu, D. Tao, Y. Zhang, J. Wu, and A. C. Tsoi, “Architectural Style Classification Using Multinomial Latent Logistic Regression,” in Computer Vision – ECCV 2014, Cham, 2014, pp. 600–615. doi: 10.1007/978-3-319-10590-1_39.

[17] H. Yazdi et al., “Central Courtyard Feature Extraction in Remote Sensing Aerial Images Using Deep Learning: A Case-Study of Iran,” Remote Sens., vol. 13, no. 23, Art. no. 23, Jan. 2021, doi: 10.3390/rs13234843.

[18] D. Dai and W. Yang, “Satellite Image Classification via Two-Layer Sparse Coding With Biased Image Representation,” IEEE Geosci. Remote Sens. Lett., vol. 8, no. 1, pp. 173–176, Jan. 2011, doi: 10.1109/LGRS.2010.2055033.

[19] R. Girshick, “Fast R-CNN,” 2015, pp. 1440–1448. Accessed: Aug. 28, 2021. [Online]. Available: https://www.cv-foundation.org/openaccess/content_iccv_2015/html/Girshick_Fast_R-CNN_ICCV_2015_paper.html

[20] Remo.ai, Remo.ai: Image Datasets management. Rediscovery.io, 2019. Accessed: Aug. 18, 2021. [Online]. Available: https://rediscovery.com/remoai-remo-python

[21] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks?,” ArXiv14111792 Cs, Nov. 2014, Accessed: Aug. 16, 2021. [Online]. Available: http://arxiv.org/abs/1411.1792

[22] K. He, G. Gkioxari, P. Dollár, and R. Girshick, “Mask R-CNN,” ArXiv170306870 Cs, Jan. 2018, Accessed: Aug. 18, 2021. [Online]. Available: http://arxiv.org/abs/1703.06870

[23] Hartmann, K., Krois, J., and Waske, B., “E-Learning Project SOGA: Statistics and Geospatial Data Analysis.,” Berlin, 2018. [Online]. Available: https://www.geo.fu-berlin.de/en/v/soga/Basics-of-statistics/Descriptive-Statistics/Measures-of-Relation-Between-Variables/Correlation/index.html

[24] G. Memarian and F. E. Brown, “CLIMATE, CULTURE, AND RELIGION: ASPECTS OF THE TRADITIONAL COURTYARD HOUSE IN IRAN,” J. Archit. Plan. Res., vol. 20, no. 3, pp. 181–198, 2003.

[25] P. Amiriparyan and Z. Kiani, “Analyzing the Homogenous Nature of Central Courtyard Structure in Formation of Iranian Traditional Houses,” Procedia - Soc. Behav. Sci., vol. 216, pp. 905–915, Jan. 2016, doi: 10.1016/j.sbspro.2015.12.087.

[26] F. Soflaei, M. Shokouhian, and S. M. Mofidi Shemirani, “Investigation of Iranian traditional
courtyard as passive cooling strategy (a field study on BS climate),” Int. J. Sustain. Built Environ., vol. 5, no. 1, pp. 99–113, Jun. 2016, doi: 10.1016/j.ijsbe.2015.12.001.

[27] M. Hajian, S. Alitajer, and M. Mahdavinejad, “The Influence of Courtyard on the Formation of Iranian Traditional Houses Configuration in Kashan,” Armanshahr Archit. Urban Dev., vol. 13, no. 30, pp. 43–55, May 2020, doi: 10.22034/aaud.2020.133667.1554.

[28] W. Rusnack, MinimumBoundingBox. 2021. Accessed: Aug. 29, 2021. [Online]. Available: https://github.com/BebeSparkelSparkel/MinimumBoundingBox