

Article

A Heritage DNA (HDNA) Framework for Quantitative Classification of Historic Facades

Onur Kılıç ^{1,*} and Harun Diler ²¹ Department of Interior Architecture, Cukurova University, 01330 Adana, Türkiye² Department of Interior Architecture and Environmental Design, Afyon Kocatepe University, 03032 Afyon, Türkiye; hdiler@aku.edu.tr

* Correspondence: okilic@cu.edu.tr

Abstract

Historic facade classification is generally based on qualitative assessments, limiting reproducibility and quantitative comparison. This study introduces the Heritage DNA (HDNA) framework, a descriptor-based approach integrating Morphological DNA (MDNA) and Material DNA (MaDNA) for the quantitative characterization of historic facades. Thirty-one facades from the Tepebağ Historic District (Adana, Türkiye) were documented using close-range photogrammetry and represented through a nine-dimensional HDNA vector. Principal Component Analysis (PCA) and hierarchical clustering were employed to identify facade patterns and typological relationships. The first two principal components explained 50.38% of the total variance, while the four-cluster solution produced a silhouette coefficient of 0.256. Hierarchical clustering identified four facade clusters, which were subsequently interpreted as four facade families characterized by distinct morphological and material compositions. The results demonstrate that facade identity can be represented through measurable descriptor combinations, providing a reproducible framework for facade classification. The proposed HDNA approach offers a potentially transferable descriptor-based methodology for comparative heritage studies and data-driven heritage management, subject to further validation across different historic contexts.

Keywords: heritage DNA (HDNA); historic facade classification; descriptor-based classification; facade morphology; material composition; photogrammetric orthophotos; principal component analysis; hierarchical clustering

Academic Editor: Emilia Vasanelli

Received: 13 June 2026

Revised: 12 July 2026

Accepted: 13 July 2026

Published: 14 July 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Historic buildings constitute an important component of cultural heritage, reflecting the architectural, social, and technological characteristics of past societies. The documentation and analysis of these structures are therefore fundamental to conservation, restoration, and heritage management processes. Among building components, facades are particularly significant because they preserve visible evidence of architectural form, construction techniques, material composition, and subsequent interventions. In recent decades, advances in digital heritage documentation, particularly through photogrammetry, laser scanning, and heritage building information modeling (HBIM), have significantly improved the recording, visualization,

and management of historic built environments [1–3]. The rapid development of digital surveying technologies has transformed heritage documentation practices over the last two decades [4,5]. Among these technologies, close-range photogrammetry has emerged as a widely adopted method due to its ability to generate accurate, high-resolution, and cost-effective three-dimensional records of historic structures [2–4]. Similarly, HBIM has extended the reach of documentation in relation to heritage by incorporating geometrical, semantic, and management information into one integrated digital environment [6–8]. The HBIM concept further enables the integration of documentation, analysis, and conservation information within semantically enriched digital heritage models [6–9]. However, documentation does not automatically lead to classification, as modern digital technologies can provide increasingly accurate models of historical buildings but do not have an intrinsic capacity to classify them [8–10]. Nevertheless, much work related to heritage digitization studies has focused mainly on documenting, reconstructing, visualizing, and managing data. Recent developments in HBIM and photogrammetry have led to highly accurate digital representations of historic environments; nevertheless, their use in developing systematic approaches to classifying facades has not been exploited. In addition, the process of translating facade information into quantifiable parameters for comparison has not received significant attention. Most facade studies depend on qualitative analyses of facades through descriptive interpretations, architectural observations, and case-based typological categorizations. Therefore, the need for approaches that translate facade information into measurable parameters for comparison has become increasingly imperative [10–12]. This need is consistent with recent discussions in urban morphology, which emphasize that quantitative and qualitative approaches should be understood as complementary rather than opposing traditions in the analysis of built form [13]. At the scale of historic districts, recent studies have also shown that measurable morphological parameters can support evidence-based conservation and renewal strategies for historic urban interfaces [14].

The use of facade typologies for analysis has traditionally been practiced in order to examine the characteristics of architecture and construction practices, as well as urban identity in historic towns [15,16]. In general, the traditional typological approach is often undertaken through the categorization of facades by means of composition, architectural elements, style, or construction material [17,18]. Such classifications can indeed be very informative in terms of the development of the historic townscape; however, the typological analyses are often performed from a descriptive point of view and thus depend on the interpreter's expertise [19,20]. As a result, the criteria used for classification may vary between studies, making systematic comparison and replication difficult across different contexts [21].

Quantitative and data-driven classification has nevertheless been explored at several architectural scales. Fleischmann et al. [21] established a numerical taxonomy of urban form based on morphometric characterization and hierarchical classification. At the building scale, Benni et al. [22] represented historic rural buildings through five quantitative formal parameters and used PCA and k-means clustering to identify groups with shared architectural characteristics. Most directly relevant to facade classification, Cheng et al. [23] quantified 325 traditional-village facades through nine geometric indices, used PCA-based factor analysis to reduce the descriptor space, and subsequently applied cluster analysis to derive five facade composition modes. A more recent study of historic-town facades in the Chengdu Plain likewise combined factor analysis and cluster analysis for systematic facade classification [24]. Beyond morphology-based typologies, multidimensional heritage assessment tools have also been developed to quantify and classify urban heritage buildings for monitoring and management purposes [25].

The DNA metaphor has already been used across urban morphology and heritage studies through concepts such as cultural DNA, architectural heritage DNA, urban DNA, morphological DNA, cultural genes, spatial genes, and historic urban landscape genes [26–34]. At the urban scale, Votsis and Haavisto operationalized urban DNA through a fixed set of numerical growth coefficients and used clustering to derive a behavioral taxonomy of cities [28], while Kaya and Bölen described a multi-parameter representation of urban pattern as morphological DNA [32]. Studies on cultural genes, historic urban landscape genes, and spatial genes have similarly used genetic metaphors to interpret identity, continuity, transformation, and morphology in historic districts and traditional settlements [29–31].

In heritage research, DNA- and gene-based concepts have increasingly moved beyond purely metaphorical interpretation toward structured analytical representations. Fan et al. [35], for example, combined cultural landscape gene theory with multivariate analysis, including PCA and clustering, to identify and interpret distinct cultural gene types. More recent studies have continued to use architectural DNA and morphological DNA as organizing concepts in urban heritage conservation and typology-led design [33,34]. These developments demonstrate the increasing use of DNA- and gene-based frameworks for describing complex built-environment identities across different spatial scales.

Building on this conceptual background, the present study adapts the DNA metaphor to the facade scale. In biology, DNA carries genetic information through the sequence and combination of its constituent nucleotides rather than through any single element [36,37]. Similarly, facade identity is here interpreted as the result of relationships among multiple morphological and material descriptors. In contrast to approaches that identify facade characteristics separately, the proposed HDNA framework treats these characteristics as interrelated descriptors organized into a fixed quantitative vector. From this perspective, the current research proposes the heritage DNA (HDNA) framework to describe historic facades quantitatively.

The HDNA model considers the facade's identity through two dimensions of its nature—namely, morphological DNA (MDNA) and material DNA (MaDNA). In other words, MDNA is concerned with the formal features of the facade represented by descriptors dealing with openings and projections, whereas MaDNA describes its materials and their diversity. As such, the proposed framework presents a comprehensive view of the facade's identity that involves not only its morphology but also its materiality. Thus, the application of the HDNA model to historical buildings allows turning architectural knowledge into measurable data for analysis. Previous studies have already demonstrated the value of quantitative architectural descriptors, multivariate classification, and DNA- or gene-inspired conceptualizations across urban, building, facade, and heritage scales [21–35]. However, these approaches generally address either morphological classification, multidimensional heritage assessment, DNA-inspired identity interpretation, or facade typology as separate analytical directions. In the literature reviewed, no previous study was identified that represents each individual historic facade through a fixed quantitative vector explicitly structured into morphological and material descriptor subspaces and recombines these subspaces within an integrated HDNA–MDNA–MaDNA framework. In light of this, the current study aims to present a heritage DNA (HDNA) approach to classify historic facades quantitatively. For this purpose, a dataset including 31 historic facades in the Tepebağ Historic District of Adana, Türkiye, documented through photogrammetric orthophotos, was used. Facade features, including morphological and material properties derived from the annotated facade components, were coded into HDNA vectors unique to each facade and were included in an

HDNA matrix. Later on, PCA and hierarchical clustering were utilized to explore the similarities and typology of facades.

The originality of the proposed framework does not reside in the use of DNA terminology, quantitative facade descriptors, PCA, or clustering individually, all of which have precedents in related architectural and heritage research [21–35]. However, it is the precise organization of these components into the form of a single historic facade that constitutes the contribution of the work. Indeed, the present framework allows facade identity to be represented as a nine-dimensional vector, which, by explicit separation of morphological and material-surface components as MDNA and MaDNA, respectively, combines them back into HDNA. As a result, each facade is allowed to become a quantitatively comparable identity profile and not merely a typologically identified facade belonging to a certain type. Thus, PCA and clustering can be viewed as an instrumental application rather than a novel approach in itself. The workflow of the study is shown in Figure 1.

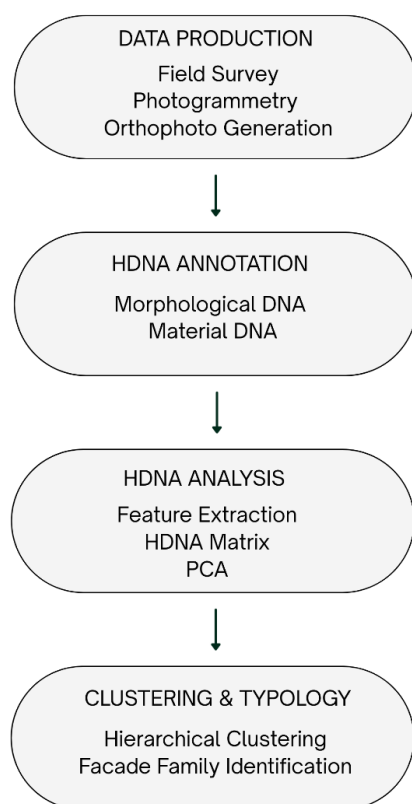


Figure 1. General workflow of the Heritage DNA (HDNA) framework. The workflow consists of four main stages: data production, HDNA annotation, HDNA analysis, and clustering-based typology generation. Historical facade characteristics were encoded through morphological and material DNA descriptors, transformed into a quantitative HDNA matrix, and subsequently analyzed using principal component analysis (PCA) and hierarchical clustering to identify facade families within the Tepebağ historic district. Source: Created by the authors.

The main contributions of this study are as follows:

- An integrated facade-level representation framework that structures historic facade identity into two explicit quantitative subspaces—morphological DNA (MDNA) and material DNA (MaDNA)—within a unified heritage DNA (HDNA) vector.

- A fixed nine-dimensional descriptor representation that enables each individual historic facade to be encoded as a comparable morphology–material identity profile derived from photogrammetric orthophotos.
- A reproducible annotation and descriptor-construction protocol that transforms visually observable facade characteristics into a structured numerical matrix suitable for multivariate analysis and comparison.
- A proof-of-concept demonstration showing how the proposed HDNA representation can support facade-family identification, robustness assessment, heritage inventories, monitoring, and future integration with digital heritage information environments.

In addition to the classification of facades based on typological criteria, the developed HDNA system is designed to provide support for heritage inventories, comparisons of facades, conservation planning, and analysis of the development of facades over time. Through the transformation of observable properties of facades into numbers, the developed system offers an organized way of recognizing patterns among facade conditions and atypical transformations.

2. Materials and Methods

2.1. Study Area

The study was conducted in the historic Tepebağ district located in Adana, Türkiye. Tepebağ represents one of the most significant surviving traditional urban settlements in the region and contains a diverse collection of late Ottoman and early Republican residential architecture. The district is characterized by narrow streets, attached and detached housing blocks, and a wide variety of facade configurations reflecting different construction techniques, material compositions, and architectural adaptations developed over time. Figure 2 illustrates the location of Adana Province within Türkiye and the boundaries of the Tepebağ Historic District, which constitutes the study area of this research.

The site of Tepebağ is chosen as the case study place since it consists of a heterogeneous group of facade types and creates appropriate situations for studying variations in terms of morphology and material in the context of the historical urban fabric. The presence of masonry, timber-framed, plastered, and partially deteriorating buildings permits the observation of several facade features. The analyzed dataset represents all photogrammetrically suitable historic facades identified within the surveyed part of the Tepebağ Historic District. Facades affected by severe occlusions, inaccessible viewpoints, or insufficient image acquisition conditions were excluded from the analysis. A total of 31 historic facades were included in the HDNA dataset (Figure 3).

The objective of the present study was not to develop a statistically representative regional facade inventory, but to establish and demonstrate a reproducible descriptor-based classification framework. Consequently, the complete set of photogrammetrically suitable historic facades identified within the study area was analyzed in its entirety. This approach ensured that the proposed HDNA framework was evaluated using the full available historic facade population rather than a sampled subset, providing a proof-of-concept implementation under real-world heritage documentation conditions.

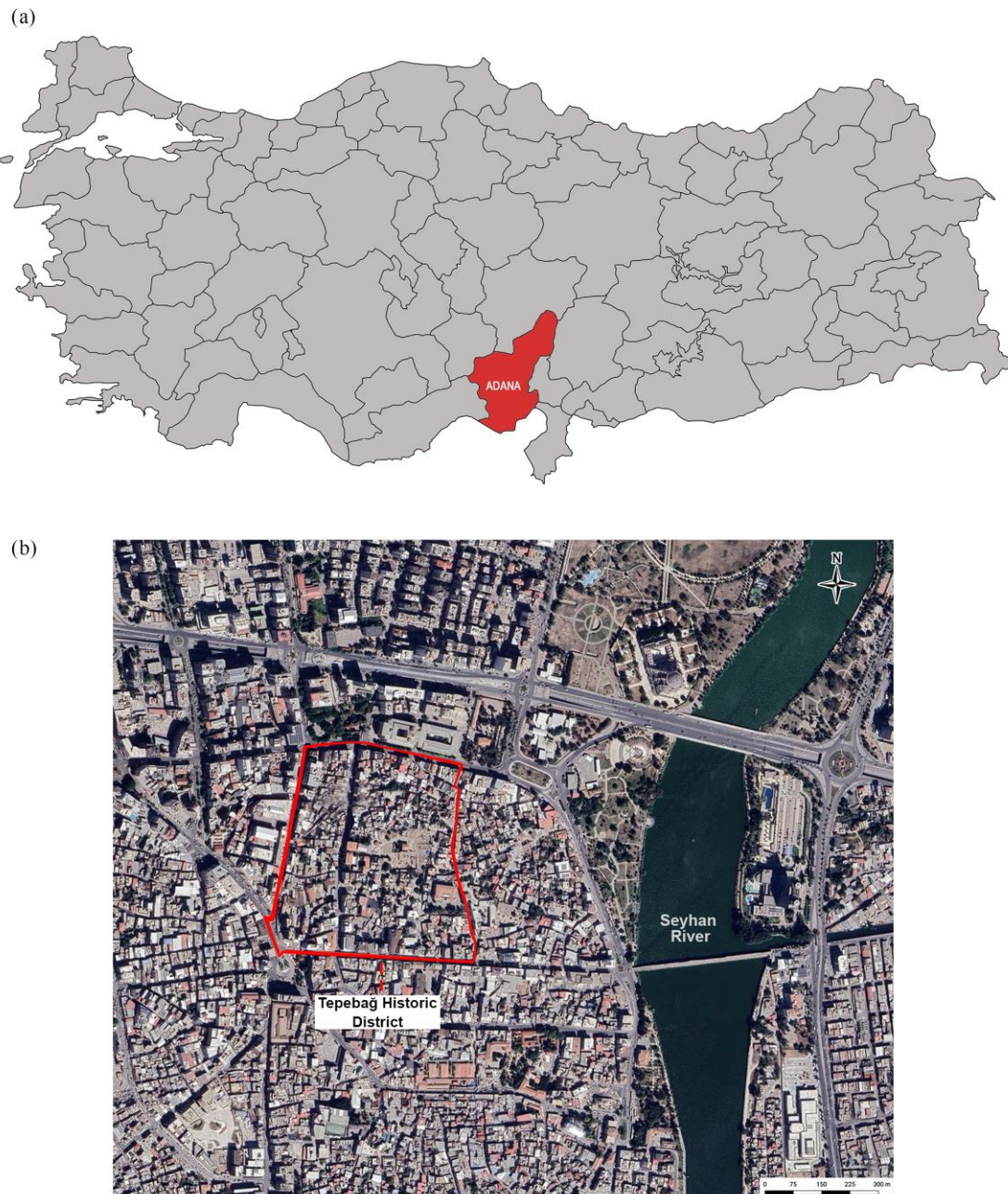


Figure 2. Study area. (a) Location of Adana Province within Türkiye. (b) Satellite view of the Tepebağ Historic District showing the study area boundary and its relationship to the Seyhan River. Source: Adapted by the authors from Wikimedia Commons and Google Earth Pro 7.3 satellite imagery (accessed March 2026).

Figure 3 presents all facade orthophotos used in the study. The dataset includes 31 facades representing different levels of preservation, material compositions, opening configurations, and architectural articulation observed within the Tepebağ Historic District.



Figure 3. Overview of the 31 historic facades included in the HDNA dataset. Source: Created by the authors.

2.2. Data Acquisition and Orthophoto Generation

A facade survey was done by means of digital photogrammetry at close range. The pictures were captured with a Canon EOS 2000D digital camera equipped with a Canon EF-S 18–55 mm lens (Canon Inc., Tokyo, Japan), providing an image resolution of 24.1 megapixels. Image acquisition was conducted during day time through a ground-based surveying technique. To provide good coverage and proper image matching for photogrammetry purposes, an overlap of about 60–80% was kept between successive images. The images were captured mainly in front of the building from a distance of 4–10 m from the facade to minimize the perspective deformation. About 720–800 photos were taken in the study area, corresponding to approximately 23–26 photos per facade. Photos obtained were then processed using WebODM 3.0 software in order to create dense point clouds, texturing models in three dimensions, and orthophotos of every single facade. The generation of orthophotos allowed for the creation of quantitatively analyzable, geometrically corrected two-dimensional projections of photographic data.

Recent reviews have also emphasized the role of photogrammetry as an accurate and scalable documentation method for cultural building heritage, particularly where detailed visual and geometric records are required for subsequent analysis [38]. Photogrammetry-based condition assessment studies further indicate that high-resolution ortho images are particularly useful for visual inspection, manual mapping, and semi-automated or automated detection of surface decay in heritage facades [39]. MeshLab 2022.02 was then used for further inspection and processing of the created models. The orthophotos created provided a unified standard for annotations, measurements, and identifications of elements of the facades in question required for HDNA

analysis. Orthophotos were visually inspected beforehand to check geometric consistency and image quality. They became a basis for further quantification of morphological and material characteristics of facades (Figure 4).

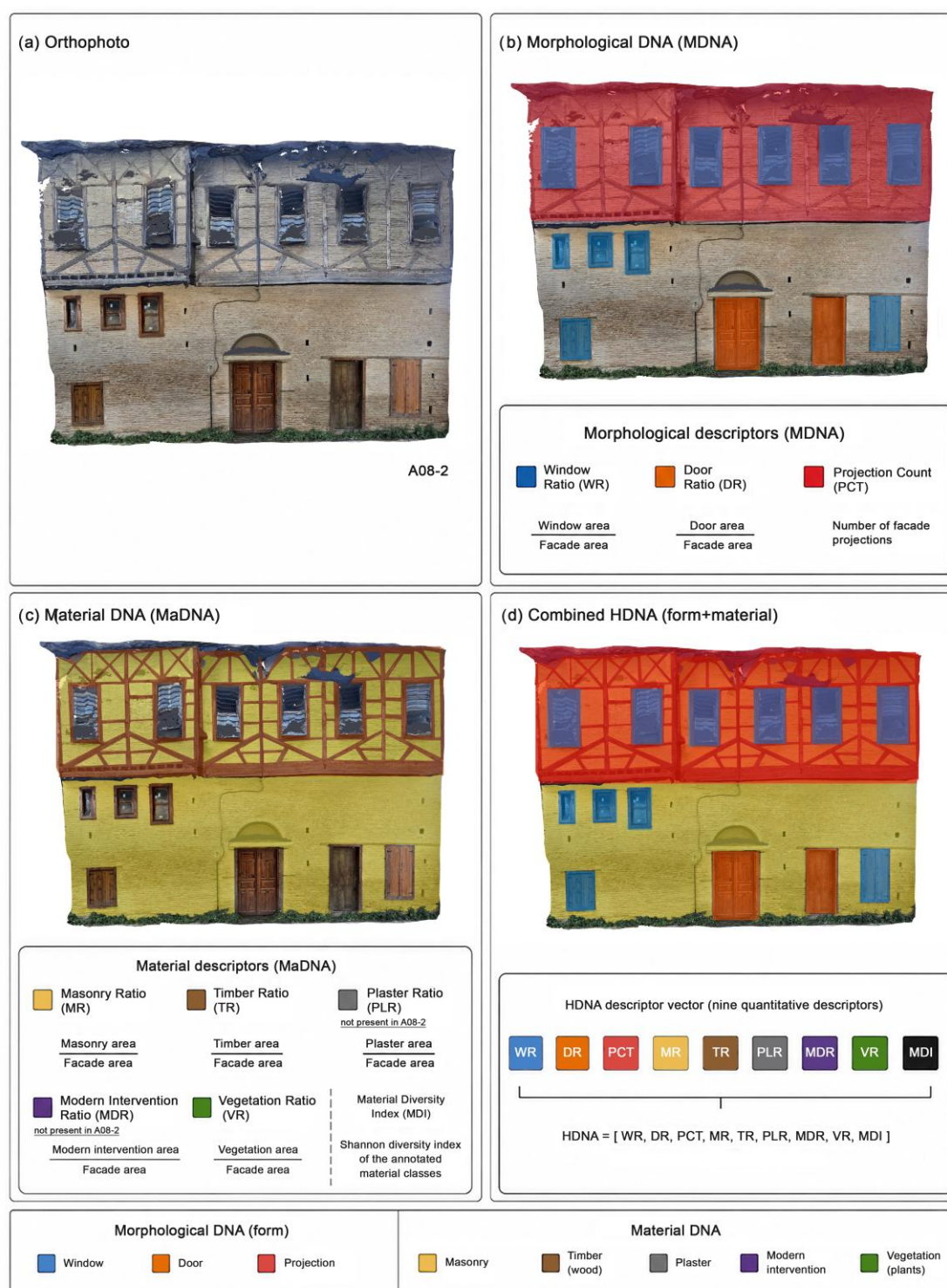


Figure 4. Graphical definition of the HDNA framework using the representative facade A08-2. (a) Orthophoto. (b) Morphological DNA (MDNA) annotation illustrating the extraction of Window Ratio (WR), Door Ratio (DR), and Projection Count (PCT). (c) Material DNA (MaDNA) annotation illustrating the extraction of Masonry Ratio (MR), Timber Ratio (TR), Plaster Ratio (PLR), Modern Intervention Ratio (MDR), Vegetation Ratio (VR), and Material Diversity Index (MDI). PLR and MDR are included in the fixed HDNA descriptor set but are absent in the representative facade shown. (d) Combined HDNA representation and the corresponding nine-dimensional descriptor

vector used for principal component analysis and hierarchical clustering. Source: Created by the authors.

Descriptor Selection Strategy

The descriptor set was selected according to three criteria: (i) architectural relevance, ensuring that each variable represents an observable characteristic commonly used in facade description and comparative architectural analysis; (ii) reproducibility, allowing for consistent extraction from photogrammetric orthophotos without requiring subjective interpretation; and (iii) transferability, enabling application across historic facades documented under similar photogrammetric conditions. Rather than attempting to capture every architectural attribute, the proposed descriptor set was designed as a compact yet representative framework capable of describing both morphological composition (MDNA) and material composition (MaDNA) using standardized quantitative variables.

Specific aspects of facade design, such as decorative features, time of construction, style features, and interventions in history, have been consciously left out because they involve subjective interpretation or cannot be reliably quantified solely based on the analysis of orthophotography. These variables constitute potential avenues for future research.

2.3. Heritage DNA (HDNA) Framework

The heritage DNA (HDNA) methodology was specifically designed to encode the features of facades into a structured quantitative format for classification and comparison purposes. Using information-coding logic of biological DNA, the HDNA methodology views the facade identity as an aggregation of several measurable descriptors rather than the mere existence of particular architectural elements. As per this methodology, each particular facade would be characterized by its own HDNA profile based on morphological and material properties.

As shown in Figure 5, the HDNA methodology comprises two major components of morphological and material aspects: morphological DNA (MDNA) and material DNA (MaDNA). While MDNA refers to the formal and spatial characteristics of a facade, MaDNA characterizes the material composition of the facade under consideration. Thus, the MDNA and MaDNA are combined to create an integrated representation of facade identity.

It should be noted that the list of nine descriptors was constrained by three main criteria: (i) measurability from photogrammetric orthophotos; (ii) potential transferability across various historic facade types; and (iii) relevance to both architectural and material composition. Among the nine descriptors considered, WR, DR, and PCT have been used as proxies for the most important morphological characteristics (openness and articulation). At the same time, MR, TR, PLR, MDR, and VR were introduced as proxies for the most significant material and environmental features observed in the study area. The Material Diversity Index (MDI) was included in order to measure the degree of diversity beyond just material proportion.

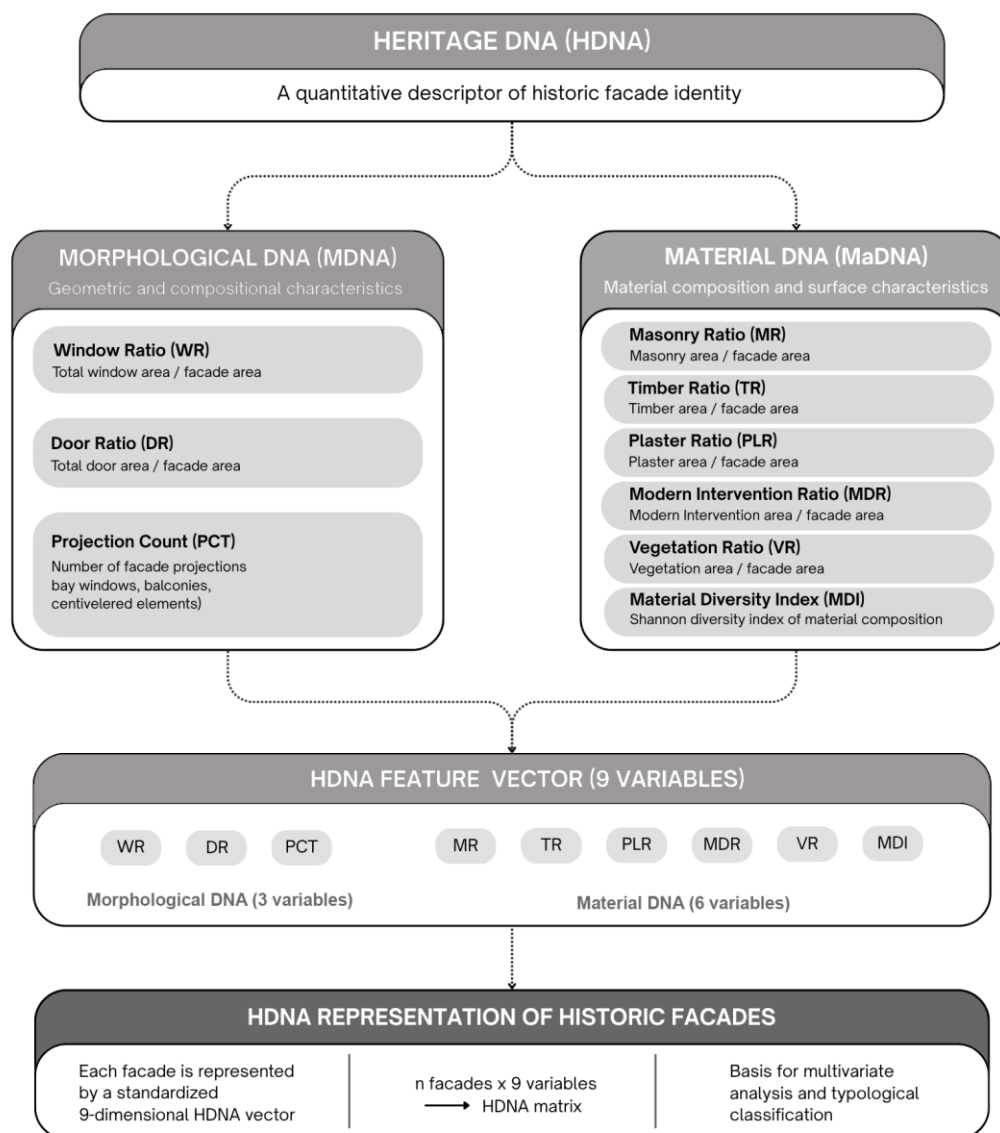


Figure 5. Conceptual structure of the Heritage DNA (HDNA) framework. The framework consists of Morphological DNA (MDNA) and Material DNA (MaDNA). MDNA includes Window Ratio (WR), Door Ratio (DR), and Projection Count (PCT), while MaDNA comprises Masonry Ratio (MR), Timber Ratio (TR), Plaster Ratio (PLR), Modern Intervention Ratio (MDR), Vegetation Ratio (VR), and Material Diversity Index (MDI). Together, these nine descriptors form the HDNA feature vector used for quantitative facade characterization and typological classification. Source: Created by the authors.

There are four major steps in the process of HDNA: (i) orthophotos production; (ii) facade labeling and descriptors calculation; (iii) creation of HDNA vector representation and HDNA matrix; and (iv) analysis by means of principal component analysis (PCA) and hierarchical cluster analysis. Every facade studied here was considered an independent case study and described by a vector of numerical descriptors in order to make a comparison between the analyzed facades possible. The obtained matrix of HDNA served as the basic dataset for further multivariate analyses. With the help of the described methodology, it is possible to identify similarities and typological connections between historic facades through numerical descriptors.

2.3.1. Morphological DNA (MDNA)

MDNA descriptors quantify the geometric organization and compositional structure of the facade independently from its material composition. MDNA consists of those properties of the historic facade that can be described through a series of geometric descriptors. Contrary to traditional typological analysis that employs qualitative descriptions, MDNA makes use of quantitative descriptors derived from orthophotos to code the morphological composition of facade architecture. The aim of using MDNA is to quantify architectural form and enable comparison across different buildings.

Three morphological descriptors that were chosen in this case include Window Ratio (WR), Door Ratio (DR), and Projection Count (PCT). Such a selection was made to ensure that only measurable variables are used to characterize the morphological composition of facades through geometric descriptors that would define the number of windows and doors, as well as the extent to which the facade was articulated. Measurement of all morphological data was conducted based on annotated orthophotos. Area values of the geometric descriptors were measured based on polygonal annotations, whereas the count and dimensions of the variables were collected based on the facade datasets.

Morphological Descriptors

Three numerical measures were utilized for describing the shape characteristics of the facades under consideration. The chosen parameters were supposed to reflect the key features of the facade structure that could be distinguished by visual inspection. All parameters were measured using annotated orthophotos and subsequently standardized before multivariate analysis.

The Window Ratio (WR) describes the proportion of facade area occupied by windows and is calculated as:

$$WR = Aw / Af$$

where Aw denotes the total window area and Af denotes the total facade area.

The Door Ratio (DR) represents the proportion of facade area occupied by doors:

$$DR = Ad / Af$$

where Ad is the total door area.

The Projection Count (PCT) represents the number of volumetric projection elements observed on each facade, including bay windows, balconies, and cantilevered upper-floor elements. Decorative moldings, cornices, roof eaves, and superficial facade profiles were excluded from the calculation. Together, these three descriptors constitute the morphological DNA (MDNA) component of the proposed HDNA framework.

2.3.2. Material DNA (MaDNA)

MaDNA descriptors quantify the relative material composition and surface characteristics observable on facade orthophotos. While morphological DNA is a description of facade geometry and morphology, material DNA (MaDNA) is the material composition described by a set of quantitative surface characteristics. Material composition is a key feature of facade identity, defined by the building traditions, available materials, material maintenance, environmental impact, and changes over time. For that reason, MaDNA was created to measure the relative presence of dominant facade materials and interventions applied to the facade throughout its existence.

The following six material variables were used: Masonry Ratio (MR), Timber Ratio (TR), Plaster Ratio (PLR), Modern Intervention Ratio (MDR), Vegetation Ratio (VR), and Material Diversity Index (MDI). This selection was made to characterize the material

composition and material diversity of facades, along with signs of biological colonization and subsequent interventions.

Materials were calculated based on facade annotation from the orthophotos. The manual delineation of all material classes was done in relation to the entire facade area, resulting in the ratio between the material area and the total facade area. Moreover, the Shannon Diversity Index [40] was calculated to quantify material diversity, thereby enabling differentiation of the facades with comparable materials but varied material diversity. Together, these six descriptors constitute the material DNA (MaDNA) component of the proposed HDNA framework.

Material Descriptors

Six descriptors were employed to characterize the material properties of the documented facades. All ratios were calculated relative to the total facade area using manually annotated orthophotos.

The Masonry Ratio (MR) represents the proportion of exposed masonry materials, including brick and stone surfaces:

$$MR = Am / Af$$

where Am denotes masonry area and Af denotes total facade area.

The Timber Ratio (TR) represents the proportion of visible timber structural or architectural elements:

$$TR = At / Af$$

where At denotes timber area.

The Plaster Ratio (PLR) describes the proportion of plastered facade surfaces:

$$PLR = Ap / Af$$

where Ap denotes the plaster-covered area.

The Modern Intervention Ratio (MDR) quantifies the proportion of facade surfaces affected by modern materials or later alterations, including metal sheets, concrete repairs, modern coatings, and contemporary infill elements:

$$MDR = Ami / Af$$

where Ami denotes the modern intervention area.

The Vegetation Ratio (VR) represents the proportion of facade area covered by biological growth, including climbing plants, shrubs, and spontaneous vegetation:

$$VR = Av / Af$$

where Av denotes vegetation-covered area.

Finally, the Material Diversity Index (MDI) measures material heterogeneity within each facade and was calculated using the Shannon Diversity Index based on the proportional distribution of identified material categories.

2.4. Construction of the HDNA Matrix

Following descriptor extraction, each facade was represented by a numerical heritage DNA (HDNA) vector composed of morphological and material variables. The HDNA vector integrates three morphological DNA (MDNA) descriptors and six material DNA (MaDNA) descriptors into a unified quantitative representation of facade identity.

For each facade i , the HDNA vector was defined as:

$$HDNA_i = [WR, DR, PCT, MR, TR, PLR, MDR, VR, MDI]$$

where WR, DR, and PCT represent morphological DNA descriptors, while MR, TR, PLR, MDR, VR, and MDI represent material DNA descriptors.

The descriptor vectors obtained from all documented facades were subsequently assembled into an HDNA matrix. Rows of this matrix corresponded to individual facades, whereas columns corresponded to the selected descriptors. Given that the number of facades in our dataset was 31, while that of descriptors was 9, the final HDNA matrix contained 31 rows and 9 columns. Before performing multivariate analysis, all descriptor values were scaled via z-score normalization in order to equalize their influence in statistical calculations. The standardized HDNA matrix was used as the primary input for PCA and hierarchical clustering. The resulting matrix provided a compact and reproducible representation of facade identity, enabling the systematic comparison of historic facades based on both morphological and material characteristics.

The HDNA descriptor model is not intended to be considered a machine-learning algorithm; instead, it provides a descriptive model that can be used as an input to various other types of analysis. PCA and hierarchical clustering were chosen to use in this current study because the analysis was exploratory in nature, and the dataset did not contain any pre-defined facade classes and was also somewhat small. Quantitative architectural descriptors with dimensional reduction and cluster analyses have been used before in taxonomy and facade studies [21–24]. PCA provides an interpretable representation of the dominant descriptor combinations, whereas Ward's hierarchical clustering enables transparent examination of nested similarities through a dendrogram without requiring predefined class labels. More data-intensive supervised machine-learning approaches were therefore not considered appropriate for the present unlabeled proof-of-concept dataset.

Annotation Quality Control

Because the HDNA descriptors were derived from manual orthophoto annotations, an annotation quality control procedure was applied to reduce potential delineation errors and improve reproducibility. All facade annotations were carried out using the same descriptor definitions and the same class hierarchy across the dataset. After the initial annotation process, each orthophoto was re-examined to verify the consistency of facade boundaries, opening polygons, projection counts, and material class delineations. Ambiguous boundaries, especially between exposed masonry, plastered surfaces, modern interventions, and vegetation-covered areas, were checked against the original image sets and photogrammetric outputs.

In addition, the resulting descriptor matrix was inspected for outliers, zero values, and inconsistent ratios before standardization and multivariate analysis. Although a formal inter-rater reliability test was not conducted in the present proof-of-concept study, the use of explicit descriptor definitions, repeated visual checking, and matrix-level consistency control was intended to minimize annotation-related inconsistencies. Future studies should further validate the HDNA annotation protocol through independent annotators and inter-rater reliability statistics.

2.5. Principal Component Analysis (PCA)

The Principal Component Analysis was carried out to detect the patterns of variations within the HDNA database and discover the combinations of the main descriptors which contribute the most to the facade differentiation process. Principal component analysis represents a popular and frequently applied multivariate statistical method aimed at transforming correlated descriptors into orthogonal principal components [41,42].

The standardized HDNA matrix was chosen to perform PCA. In particular, this technique was applied to the correlation matrix of the nine HDNA descriptors. Subse-

quently, the principal components were ranked by the amount of variance they explain, thus facilitating the detection of the predominant tendencies in facade variability.

PCA was implemented for two reasons. Firstly, PCA helped reduce the dimensionality of the problem and condense descriptors into several components. Secondly, this technique was used to quantify the process of comparing facades within the reduced multidimensional space of component scores. PCA scores were used to visualize and interpret the multivariate distribution of facades and to compare the clustering structure with the reduced feature space.

In order to detect the HDNA descriptors that contribute the most to certain PCs, the descriptor loadings were analyzed. As a result of this analysis, it became clear what characteristics of facades contributed to their variability. The ratio of explained variance of each component has been calculated to detect the proportions of the total HDNA variability represented by principal components. Moreover, the loading values of PC1 and PC2 have been specified.

2.6. Hierarchical Clustering

A hierarchical clustering approach was used to detect sets of facades sharing similar features concerning the HDNA properties and to derive a data-driven facade grouping. Different from partition-based techniques, hierarchical clustering does not require an a priori setting of the number of groups. In addition, hierarchical clustering provides a graphical representation of similarities among the objects in the form of a dendrogram.

Firstly, hierarchical clustering of facades was carried out based on the standardized HDNA matrix. Similarities were computed via Euclidean distance, a commonly used metric for measuring dissimilarity between descriptor vectors in multidimensional feature spaces [43]. Clustering was then performed using Ward's linkage, which minimizes the increase in within-cluster variance during agglomerative clustering [44].

Clustering of the facades into four groups was carried out using Ward's linkage and Euclidean distance. The dendrogram was cut at a linkage distance of approximately 7.0, resulting in four facade clusters. To avoid relying solely on visual inspection of the dendrogram, alternative cluster solutions between $k = 2$ and $k = 6$ were also evaluated using three complementary internal cluster validity indicators. The silhouette coefficient evaluates within-cluster cohesion relative to between-cluster separation, with higher values indicating more clearly defined clusters [45]. The Calinski–Harabasz index compares between-cluster dispersion with within-cluster dispersion, with higher values indicating stronger partitioning [46]. The Davies–Bouldin index evaluates the similarity of each cluster to its most similar counterpart, with lower values indicating better cluster separation [47].

Although the $k = 5$ solution produced a slightly higher silhouette coefficient and a lower Davies–Bouldin index, it resulted in additional very small subgroups and reduced the architectural interpretability of the classification. Thus, the four-cluster solution was chosen as it represented the optimal balance between dendrogram distinction, internal validity, cluster distribution, visual cohesiveness through PCA, and interpretation of types. The value of silhouette for the four-cluster solution is 0.256. This value was interpreted jointly with the Calinski–Harabasz and Davies–Bouldin indices, cluster-size distribution, PCA structure, and architectural interpretability rather than as a standalone decision criterion.

To assess the robustness of the final cluster structure, the four-cluster solution was further compared with alternative clustering procedures applied to the same standardized nine-dimensional HDNA matrix. These included k-means clustering [48], Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH) [49], and agglomerative hierarchical clustering using average and complete linkage [43]. Agreement between

each alternative solution and the final Ward partition was quantified using the Adjusted Rand Index (ARI), a chance-adjusted measure of agreement between partitions [50]. K-means was performed with $n_init = 100$ and $random_state = 42$. BIRCH was run with $n_clusters = 4$ and $threshold = 0.5$. Average- and complete-linkage agglomerative clustering used Euclidean distance, and all alternative procedures were specified to produce four clusters.

Because no external class labels were available, supervised cross-validation was not applicable; cluster stability was therefore evaluated through data-perturbation and resampling procedures. First, a leave-one-out (delete-one) stability analysis following the logic of jackknife resampling was conducted [51]. One facade was removed at a time, the remaining data were re-standardized, and Ward clustering with $k = 4$ was repeated, resulting in 31 runs. Second, repeated 80% subsampling was performed over 1000 iterations by randomly selecting 25 of the 31 facades without replacement, re-standardizing each subset, and repeating the clustering procedure. Subsetting-based perturbation provides an established approach for assessing the stability of clustering solutions [52]. Stability in both analyses was quantified using the ARI relative to the original Ward partition restricted to the retained observations.

Family-level recovery under subsampling was additionally evaluated using maximum Jaccard similarity. Jaccard similarity measures the overlap between two sets as the size of their intersection divided by the size of their union, ranging from 0 for no overlap to 1 for complete overlap [53]. Because cluster labels are arbitrary across resampling iterations, each original HDNA family represented in a subsample was compared with every cluster obtained in that subsample, and the largest Jaccard value was retained as the family-level recovery measure. This maximum-matching procedure follows the cluster-wise stability principle of comparing each original cluster with its most similar cluster in the perturbed dataset [52]. All statistical and multivariate analyses were conducted using Python 3.13.9 in conjunction with the scientific computing libraries, such as NumPy, pandas, SciPy, and scikit-learn. The workflow using Python was employed for PCA, hierarchical clustering, evaluation of cluster validity internally, comparative clustering, and robustness analysis. The additional analysis files including the matrix of descriptors of HDNA, PCA related files, cluster files, and robustness analysis files are included as Supplementary Materials.

2.7. Facade Family Identification

The last phase of the HDNA methodology focused on the determination of facade families according to the obtained clustering results. Herein, a facade family refers to an architectural family composed of facades that exhibit similar morphological and material profiles, reflected in their HDNA descriptors.

After hierarchical clustering of the dataset, the facades belonging to each cluster were considered together in order to examine their common descriptor pattern. The mean value of the descriptors, the distribution of PCA coordinates, and other specific characteristics of the clusters helped in recognizing common architectural tendencies in each group of facades, thus allowing for the interpretation of the clusters as real facade types rather than mere statistics.

Therefore, facade family determination was based on both the numerical structure given by PCA and hierarchical clustering and the interpretative potential provided by HDNA descriptors. The former determined the basis for grouping facades, whereas the latter allowed characterizing each architectural type with its typical morphological/material properties.

Thus, facade families serve as the outcome of the proposed framework. These architectural types of facade identity are reproducible groups that can be used for comparing historic facades in heterogeneous historic urban areas.

3. Results

3.1. HDNA Descriptor Statistics

The HDNA sample set involved 31 facades described by nine quantitative measures covering morphological and material features. There was great variation amongst the facades concerning the proportions of openings, materials used, vegetation on the facades, and diversity of materials. Facade window ratio demonstrated large variation among the sample set, showing variation in the degree of openings in the building facades. Likewise, there was a large variation in timber ratio, masonry ratio, and plaster ratio, demonstrating differences in the building methods and material preservation status. There was significant variation in vegetation ratio, where some facades had considerable vegetation cover, while others were free of vegetation. Variation in material diversity also characterized the sampled facades. Overall, it is evident that the documented facades demonstrated wide variation in their morphological and material attributes, forming a strong base for multivariate analysis.

Table 1 summarizes the descriptive statistics of the HDNA variables used in the analysis. Masonry Ratio (MR) and Plaster Ratio (PLR) exhibited the highest mean values, indicating the predominance of traditional masonry and plaster surfaces across the investigated facades. Vegetation Ratio (VR) and Modern Intervention Ratio (MDR) showed substantial variability, reflecting different levels of abandonment, biological colonization, and contemporary interventions.

Table 1. Descriptive statistics of HDNA descriptors used in the study. Source: Created by the authors.

Descriptor	Mean	SD	Min	Max
Window Ratio (WR)	0.091	0.077	0.000	0.276
Door Ratio (DR)	0.023	0.027	0.000	0.100
Projection Count (PCT)	0.355	0.608	0.000	2.000
Masonry Ratio (MR)	0.315	0.298	0.000	0.911
Timber Ratio (TR)	0.034	0.052	0.000	0.260
Plaster Ratio (PLR)	0.314	0.257	0.000	0.740
Modern Intervention Ratio (MDR)	0.054	0.083	0.000	0.346
Vegetation Ratio (VR)	0.049	0.098	0.000	0.462
Material Diversity Index (MDI)	0.441	0.213	0.117	0.806

Note: SD = standard deviation.

3.2. Principal Component Analysis Results

PCA was conducted in order to find out the dominant patterns of variation in the HDNA data and determine the descriptors with the greatest influence on facade differentiation. According to PCA output, there are two main axes that account for the variability of facades—facade articulation (PC1) and facade transformation (PC2). While PC1 accounts for 27.34% of the total variation, PC2 makes up 23.04%. The first two PCs together account for 50.38% of the total variation in the standardized HDNA dataset.

As seen from the PCA output, PC1 is mainly determined by Window Ratio, Door Ratio, and Projection Count. Thus, this axis can be said to reflect to what extent a facade is open and articulated. High PC1 scores are usually indicative of the openness of a facade and the number of its projections, while low (negative) scores are associated with denser

and more enclosed facades. Masonry Ratio (MR) showed an inverse relationship with PC1 (loading = -0.486), indicating that higher PC1 scores are associated with lower masonry dominance, whereas more masonry-dominated facades tend to have lower PC1 scores.

On the other hand, the second principal component (PC2) is mainly associated with positive loadings of Modern Intervention Ratio (MDR; 0.552) and Material Diversity Index (MDI; 0.533), together with a substantial negative loading of Masonry Ratio (MR; -0.417) and a positive loading of Plaster Ratio (PLR; 0.336). Thus, PC2 can be interpreted as reflecting a contrast between more transformed and materially diverse facade compositions and more masonry-dominated compositions (Table 2). Positive PC2 scores are associated with higher modern intervention, material diversity, and plaster presence, whereas negative PC2 scores are associated with more masonry-dominated and less materially diverse compositions.

Table 2. PCA loadings of HDNA descriptors. Source: Created by the authors.

Descriptor	PC1	PC2
Window Ratio (WR)	0.530	−0.194
Door Ratio (DR)	0.410	−0.054
Projection Count (PCT)	0.371	−0.248
Masonry Ratio (MR)	−0.486	−0.417
Timber Ratio (TR)	0.148	−0.140
Plaster Ratio (PLR)	0.357	0.336
Modern Intervention Ratio (MDR)	−0.030	0.552
Vegetation Ratio (VR)	−0.116	0.054
Material Diversity Index (MDI)	−0.117	0.533

The PCA plot (Figure 6b) shows a wide scatter of facades in the space defined by the reduced components, which indicates considerable diversity of both morphology and materials used in the examined site. Some of the facades appear as isolated samples, thus having unique architectural features, while others can be seen to have formed groups that show some typological similarities. This was the starting point for conducting the cluster analysis.

3.3. Hierarchical Clustering Results

Hierarchical clustering identified four main clusters in the analyzed dataset (Figure 6). The structure of the dendrogram indicated different levels of similarity between the facade types, which were determined by morphological and material differences. Dendrogram partitioning was performed according to the 7.0 level of linkage distance, which resulted in four main clusters.

The final cluster composition included four clusters of unequal size. Cluster 1 contained 10 facades, Cluster 2 had 14 facades, Cluster 3 had two facades, and Cluster 4 had five facades. Thus, some facade configurations are more frequent within the analyzed dataset, whereas others are represented by smaller groups. The four clusters were subsequently interpreted as four context-specific facade families sharing common morphological and material characteristics. To further evaluate the robustness of the selected cluster solution, alternative cluster numbers between $k = 2$ and $k = 6$ were compared using internal validity indices (Table 3). The $k = 4$ solution did not maximize a single statistical criterion; however, it provided a balanced compromise between statistical separation, cluster size distribution, dendrogram structure, and architectural interpretability. Although $k = 5$ produced a slightly higher silhouette coefficient, it divided the dataset

into additional small subgroups, which were less meaningful for typological interpretation in the context of this proof-of-concept study.

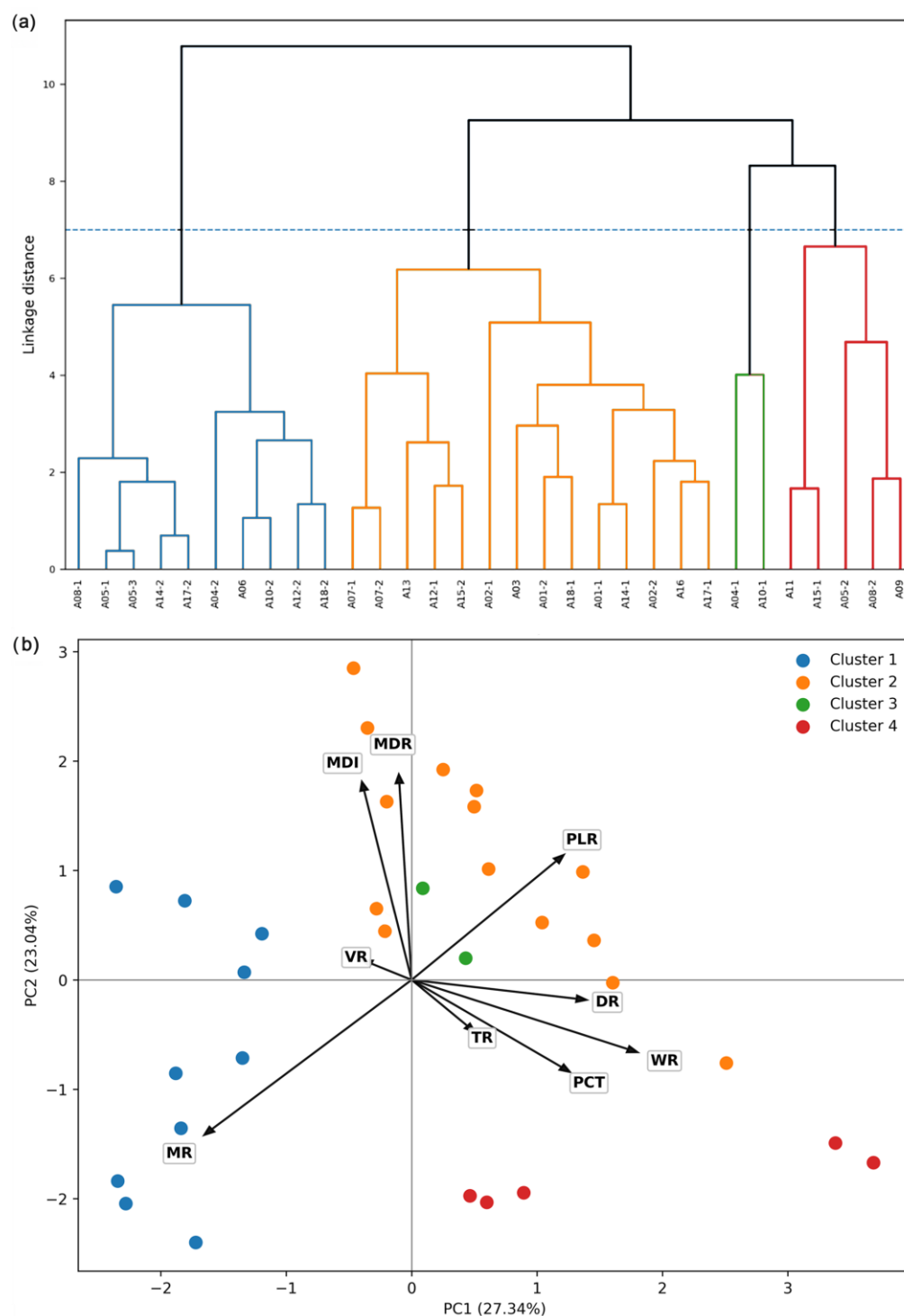


Figure 6. HDNA multivariate analysis of the investigated facades. (a) Hierarchical clustering dendrogram generated from the standardized MDNA and MaDNA descriptors, showing the four-cluster structure obtained at a linkage distance of approximately 7.0. The blue dashed horizontal line indicates the dendrogram cut-off threshold used to obtain the four-cluster solution. (b) PCA biplot illustrating the distribution of facades in the reduced feature space and the contribution of HDNA variables to cluster differentiation. The first two principal components explain 50.38% of the total variance (PC1 = 27.34%, PC2 = 23.04%). Abbreviations: WR = Window Ratio; DR = Door Ratio; PCT = Projection Count; MR = Masonry Ratio; TR = Timber Ratio; PLR = Plaster Ratio; VR =

Vegetation Ratio; MDR = Modern Intervention Ratio; MDI = Material Diversity Index. Source: Created by the authors.

Table 3. Internal cluster validity indices for alternative cluster solutions based on the standardized nine-dimensional HDNA matrix. Source: Created by the authors.

Number of Clusters (k)	Cluster Sizes	Silhouette Coefficient	Calinski–Harabasz Index	Davies–Bouldin Index
2	10, 21	0.183	7.63	1.53
3	10, 14, 7	0.223	7.93	1.63
4	10, 14, 2, 5	0.256	8.50	1.22
5	10, 14, 2, 2, 3	0.287	8.44	1.02
6	10, 5, 9, 2, 2, 3	0.259	8.63	1.08

The final four-cluster solution was chosen after taking into account the measures of internal validity as well as the dendrogram structure, distribution of the cluster sizes, visualization by means of PCA, and the architecture-friendly interpretation of the clusters. For the final four-cluster solution, the silhouette coefficient was equal to 0.256. This measure was not used as the sole criterion for selecting the solution; instead, it was considered together with other criteria, such as Calinski–Harabasz index, Davies–Bouldin index, dendrogram structure, PCA distribution, cluster sizes balance, and visual inspection of the resulting facade families.

In order to provide further evidence for interpreting the results of the clustering analysis, all of the facades belonging to the same HDNA family have been displayed together in Figure 7. This figure allows one to make visual comparisons within each cluster and provides complementary evidence to the quantitative analysis shown in the dendrogram and PCA biplot.

The alternative clustering comparison provided additional evidence that the core four-family structure was not specific to Ward’s linkage alone. Under the specified BIRCH configuration ($n_clusters = 4$, $threshold = 0.5$), the resulting partition was identical to the Ward solution ($ARI = 1.000$). K-means also showed high agreement with Ward ($ARI = 0.909$) and, after matching cluster labels for descriptive comparison, differed for only one facade (A13). By contrast, average- and complete-linkage solutions showed substantially lower agreement with the Ward partition and produced strongly imbalanced cluster-size structures. The low Davies–Bouldin value obtained for average linkage was interpreted cautiously because it accompanied a highly imbalanced 27–2–1–1 partition dominated by a single cluster and three very small clusters (Table 4). These findings show that the core partition was recovered by BIRCH under the specified configuration and closely approximated by k-means, whereas the other hierarchical linkage methods produced less concordant and more imbalanced solutions. Ward’s method was retained because it also provided a directly interpretable hierarchical representation through the dendrogram.

Cluster stability was further examined using leave-one-out and repeated subsampling analyses. In the leave-one-out analysis, the mean ARI was 0.881 and the median ARI was 1.000 across 31 runs, with an interquartile range of 0.826–1.000. The original partition of the retained facades was reproduced exactly in 19 of the 31 runs (61.3%). Under the more demanding 80% subsampling procedure, repeated over 1000 iterations, the mean ARI was 0.728 and the median ARI was 0.755, with an interquartile range of 0.561–0.887; an ARI of at least 0.75 was obtained in 51.1% of the runs. Mean family-level maximum Jaccard recovery ranged from 0.728 to 0.856. Together, these results indicate high local stability under single-facade deletion and lower, but still substantial, agreement under the more demanding removal of approximately 20% of the dataset. The wider variability observed under subsampling is consistent with the sensitivity expected

from a small proof-of-concept dataset containing clusters of unequal size, including one two-member family.

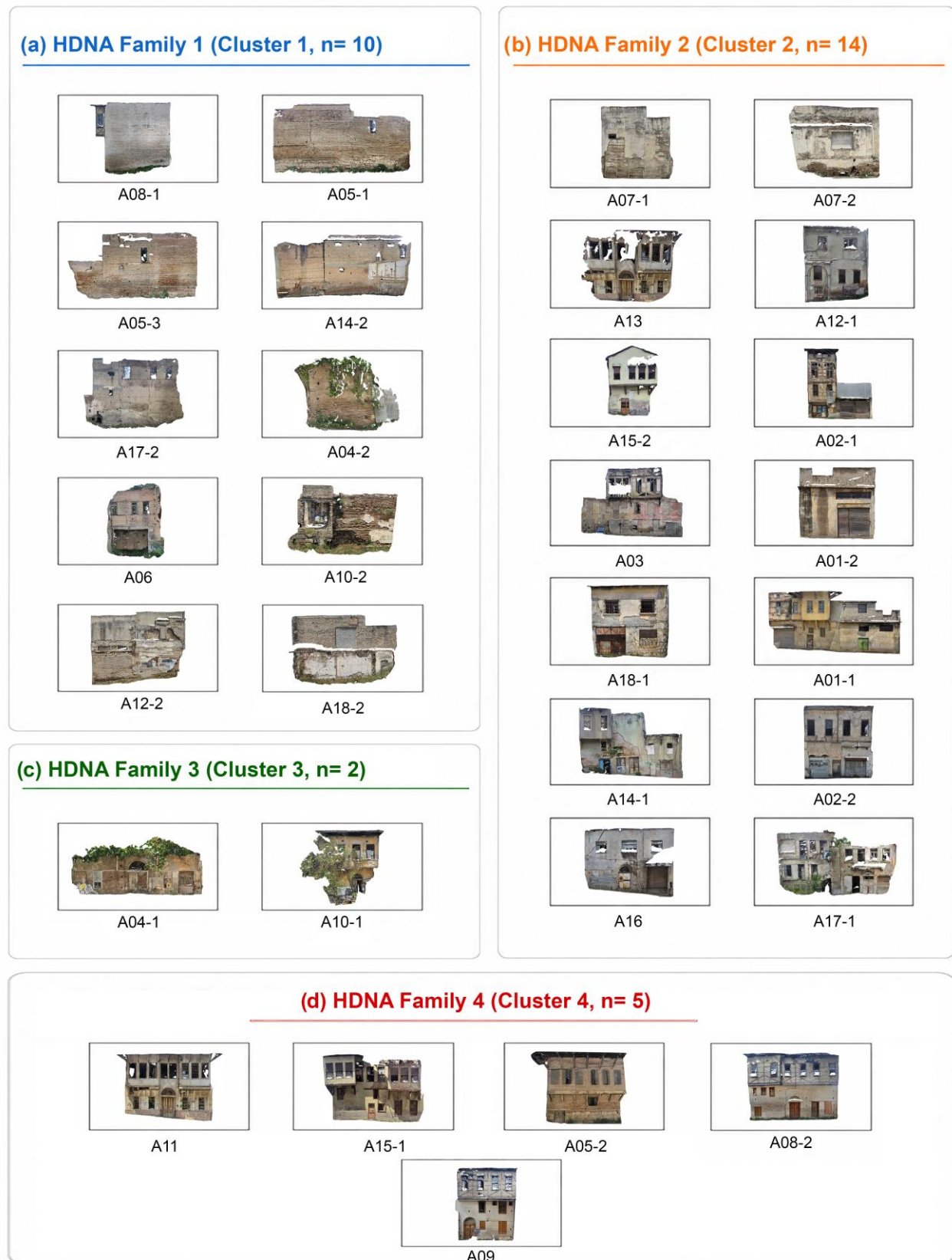


Figure 7. Collective display of all facades assigned to the four HDNA families identified through hierarchical clustering. Each panel shows the complete set of facades belonging to one cluster, with facade codes corresponding to the study dataset. The grouping provides a visual comparison of

intra-family similarities and complements the quantitative classification results presented in the dendrogram and PCA biplot. Colors correspond to the cluster colors used in Figure 6. Source: Created by the authors.

Table 4. Comparison of alternative clustering procedures for the final four-cluster HDNA solution. Source: Created by the authors.

Clustering Procedure	Cluster Sizes	Silhouette Coefficient	Calinski–Harabasz Index	Davies–Bouldin Index	ARI vs. Ward
Ward hierarchical	14, 10, 5, 2	0.256	8.50	1.22	1.000
K-means	13, 10, 6, 2	0.259	8.93	1.23	0.909
BIRCH	14, 10, 5, 2	0.256	8.50	1.22	1.000
Average linkage	27, 2, 1, 1	0.209	3.57	0.80	0.135
Complete linkage	22, 5, 2, 2	0.226	5.63	1.25	0.348

Note: ARI = Adjusted Rand Index. Higher silhouette and Calinski–Harabasz values and lower Davies–Bouldin values indicate better internal cluster validity. ARI values closer to 1 indicate greater agreement with the final Ward partition. K-means was performed with $n_{init} = 100$ and $random_state = 42$; BIRCH was run with $n_clusters = 4$ and $threshold = 0.5$. Cluster sizes are reported in descending order and therefore do not imply correspondence between cluster labels across methods.

3.4. Facade Families

PCA and hierarchical clustering analysis made it possible to define four different groups of facades that make up the families. Each family consists of morphological and material features specific to each one and defining an individual architectural identity. These facade families were defined in accordance with Clusters 1–4 derived through hierarchical clustering analysis and thus constitute the interpretative representation of clusters in terms of typology.

Table 5 summarizes the mean descriptor values of the four facade families identified through HDNA analysis. Family 1 is characterized by the highest masonry ratio (MR = 0.648), indicating facades dominated by exposed masonry surfaces. Family 2 exhibits the highest plaster ratio (PLR = 0.519) together with the highest modern intervention ratio (MDR = 0.103), reflecting plaster-covered facades affected by later modifications. Family 3 presents the highest vegetation ratio (VR = 0.337) and material diversity index (MDI = 0.721), representing facades with extensive biological colonization and high material diversity. Family 4 displays the highest window ratio (WR = 0.209), timber ratio (TR = 0.094), and projection count (PCT = 1.200), indicating more articulated and architecturally complex facade compositions.

Table 5. Mean HDNA descriptor values of the identified facade families. Source: Created by the authors.

Descriptor	Family 1	Family 2	Family 3	Family 4
WR	0.030	0.097	0.060	0.209
DR	0.000	0.025	0.075	0.041
PCT	0.100	0.214	0.500	1.200
MR	0.648	0.099	0.161	0.314
TR	0.010	0.029	0.049	0.094
PLR	0.116	0.519	0.154	0.201
MDR	0.019	0.103	0.019	0.000
VR	0.064	0.011	0.337	0.010
MDI	0.391	0.498	0.721	0.267

3.4.1. Family 1: Masonry-Dominated Facades

The first family is noted for its high masonry ratio and facade with a comparatively small amount of openness. The facades that belong to this category normally feature a very compact composition with low areas covered by windows and doors, thus giving rise to highly enclosed elevations. The domination of masonry material indicates a high level of connection with traditional methods of construction (Figure 8).

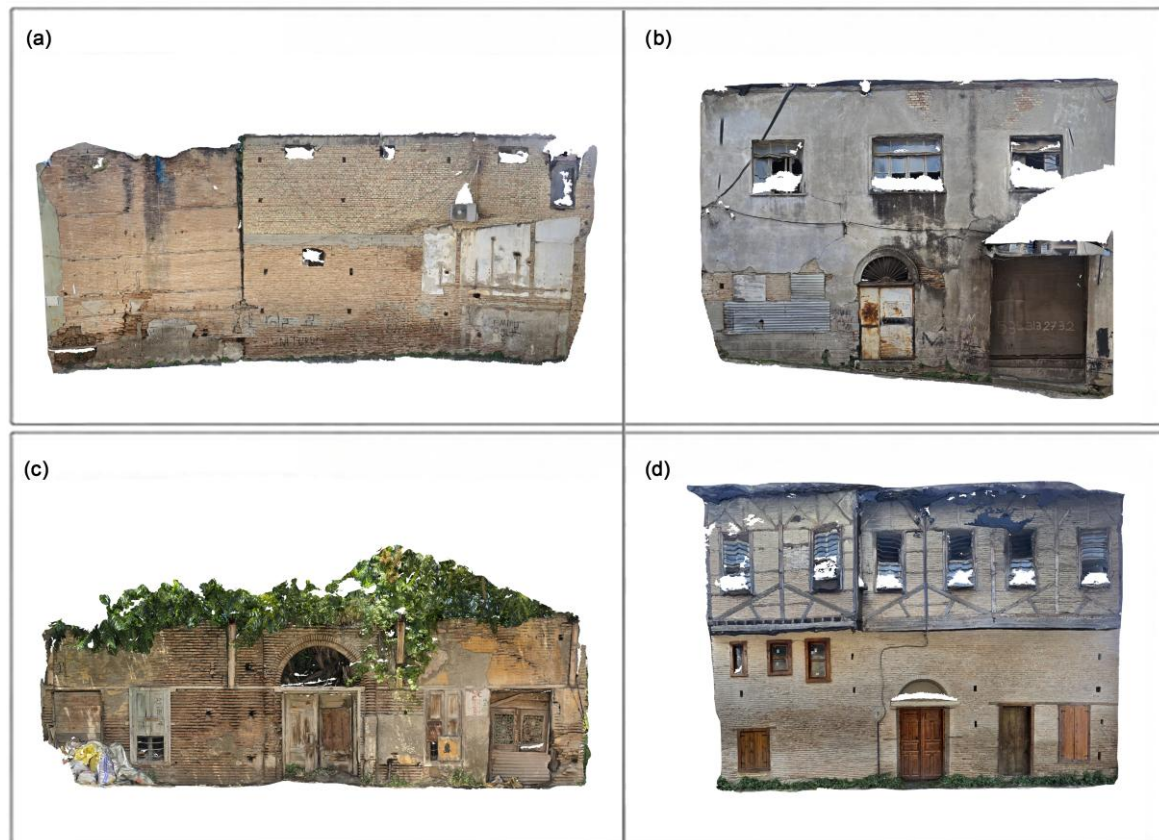


Figure 8. Representative examples of the four facade families identified through HDNA-based hierarchical clustering. (a) Masonry-Dominated Facades, (b) Plaster-Rich Residential Facades, (c) Vegetation-Affected Facades, and (d) Articulated Timber-Framed Facades. Source: Created by the authors.

3.4.2. Family 2: Plaster-Rich Residential Facades

The second family is distinguished by the high ratio of plastered surfaces and moderate facade articulation. The facades featured in this category look rather homogeneous and belong to residential buildings in which plastering plays a prominent part. This family demonstrates more pronounced material transformation than the first one, without significant changes in morphology.

3.4.3. Family 3: Vegetation-Affected Facades

Thirdly, the group is distinguished mainly based on high vegetation ratios. The occurrence of climbers and high vegetation cover is highly influential in determining the appearance of the facade, thus playing a role in creating unique features for the facade. Despite the differences in the inherent architecture for each facade, vegetation forms an important part of the description.

3.4.4. Family 4: Articulated Timber-Framed Facades

The fourth facade family contains the most articulated facades, with high Window Ratio (WR), a high number of projections, and a prevalence of timber elements. The facades in this group feature rich compositions with notable upper-floor projections and a high level of architectural articulation. In the HDNA classification scheme, the fourth facade family is considered the most articulated and compositionally varied facade type.

Through the proposed methodological approach, it becomes possible to classify historic facades based on their morphology and material characteristics. Instead of employing visual judgments alone, the use of a standardized set of morphological and material descriptors in the HDNA classification system makes it possible to identify relationships between facade types (Table 6).

Table 6. HDNA-based facade families and dominant characteristics. Source: Created by the authors.

Family	Name	n	Dominant Descriptors
Family 1	Masonry-Dominated	10	High MR
Family 2	Plaster-Rich Residential	14	High PLR, MDR
Family 3	Vegetation-Affected	2	High VR, MDI
Family 4	Articulated Timber-Framed	5	High WR, TR, PCT

4. Discussion

4.1. Interpretation of HDNA-Derived Facade Families

The clustering results demonstrate that the historic facades of the Tepebağ Urban Conservation Area do not form a homogeneous architectural group but instead exhibit multiple recurring facade identities. The four HDNA-derived facade families identified in this study reflect different combinations of material composition, facade articulation, vegetation presence, and opening configurations. Rather than representing stylistic categories in the conventional architectural sense, these families capture measurable similarities among facades based on their morphological and material characteristics.

Family 1 (Masonry-Dominated Facades) has predominantly extensive masonry finishes, minimal articulation, and low opening ratios. The facade type can be said to represent wall-dominant elevations, which are highly simplistic, with the architectural expression largely dependent on masonry continuity. Family 2 (Plaster-Rich Residential Facades) has higher proportions of plaster and a more homogeneous composition, thus representing another phase of material transformation and facade treatments. Family 3 (Vegetation-Affected Facades) exhibits significantly high levels of biological growth, which may be associated with abandonment or reduced maintenance. The same trends have been observed within cultural heritage literature, where vegetation and biological growth have been seen to play a crucial role in facade aesthetics as well as degradation processes [54,55]. Family 4 (Articulated Timber-Framed Facades) represents highly complex facades with timber-framed upper floors and high opening ratios.

4.2. Descriptor-Based Interpretation of Facade Formation Tendencies

In addition to explaining the most prevalent descriptor sets of the respective HDNA families, the results of the clustering analysis also provide some clues regarding the potential facade formation characteristics in the case of Tepebağ. The dominance of the masonry set implies that the facades of the neighborhood are formed by wall continuity and exposed surfaces of construction materials. This type of facade formation can be linked with the lack of architectural articulation, material degradation, or the continued presence of load-bearing masonry walls as the main visual element of the facade. On the

contrary, the predominance of the plaster set indicates that surface treatment and subsequent material homogenization play a major role.

The vegetation-affected family reveals another formation tendency in which biological colonization and reduced maintenance become visible and measurable components of facade identity. This family may reflect processes of abandonment, prolonged exposure, and environmental change. By contrast, the Articulated Timber-Framed family features high ratios of openings, timber, and projection elements, indicating that architectural articulation and upper-floor projection systems continue to play an important role in facade identity.

Nonetheless, since the current research does not make use of information from the archives on chronology, building usage, ownership, and socioeconomic factors, these conclusions are meant to be taken as morphological interpretations based on HDNA descriptors and not definitive explanations from the point of view of history. Future research that makes use of HDNA descriptors together with archival information would have a clearer look at this connection.

4.3. Relationship Between Morphological and Material Characteristics

The clustering results indicate that facade similarity cannot be explained solely by morphology or material composition. Facades exhibiting comparable formal characteristics may differ substantially in material structure, whereas materially similar facades may display different morphological organizations.

It is particularly evident in the case of the differentiation between facade families, when one can distinguish an interconnection between the material information and specific facade properties. For instance, while the difference between Masonry-Dominated and Plaster-Rich facade types relates to distinct material properties of the buildings' outer surfaces, the differentiation of the Articulated Timber-Framed family depends to a great extent on morphological aspects of facade identity—articulation level, opening shape, and facade complexity. Moreover, the Vegetation-Affected family provides additional insight into the way that environmental processes can play a significant role in defining facade identity.

From these data, it follows that the identity of historic facades can be considered as a result of complex interactions among the discussed categories. Instead of isolated groups, facade families should be understood as multifaceted combinations of various facade features. The findings suggest that integrating morphological and material descriptors provides a broader representation of facade identity than considering either dimension independently.

4.4. Methodological and Practical Contributions of the HDNA Framework

Previous research establishes three closely related lines of work. The first uses DNA- or gene-based concepts to interpret urban and heritage identity [26–35]. The second develops numerical taxonomies and multivariate classifications from quantitative architectural or morphological descriptors [21,22,25]. The third applies factor- or PCA-based dimensional reduction and cluster analysis directly to facade datasets [23,24]. These precedents demonstrate that DNA-inspired conceptualization, quantitative descriptor systems, and multivariate classification each have established foundations in the literature.

The relationship between these precedents and the present study is therefore one of methodological convergence rather than direct duplication. Studies of urban DNA and morphological DNA have shown that built-environment identity can be expressed through combinations of quantitative or structural attributes [28,32]. Cultural-gene studies have further demonstrated that DNA- or gene-inspired concepts can be com-

bined with multivariate analysis and clustering [35]. At the same time, quantitative facade studies have already used geometric indices, dimensional reduction, and cluster analysis to identify facade composition types [23,24]. These studies establish important precedents for the individual components of the present workflow.

The distinction of HDNA lies in the level and structure of representation. Rather than using DNA as a general metaphor for urban or cultural identity, or using quantitative facade indicators solely as inputs to classification, HDNA defines each individual historic facade as a fixed morphology–material identity vector. The framework explicitly separates this vector into MDNA and MaDNA subspaces and recombines them as a single facade-level representation before statistical analysis. In the literature reviewed, this specific combination of facade-level analysis, fixed-length morphology–material representation, and explicit HDNA–MDNA–MaDNA structuring was not identified.

Within this broader methodological landscape, the HDNA framework operates at the facade scale and organizes measurable morphological and material information within a single descriptor representation. The primary advantage of HDNA is that it provides the opportunity to turn the complicated characteristics of historic facades into a quantitative and organized set of descriptions. Through combining morphological and material features into one analysis model, HDNA helps identify similarities and differences among historic facades along with architectural patterns in general. Compared with purely visual or descriptive typologies, the proposed framework provides a reproducible numerical representation of facade characteristics. According to the clustering outcomes, the unique identity of historic facades depends on several factors in combination, rather than only individual ones.

Recent digital heritage studies increasingly emphasize that documentation workflows should move beyond geometric recording and visualization toward information-rich, interoperable, and analytically useful heritage data environments [56]. Built heritage modeling and HBIM studies have shown that photogrammetry, laser scanning, GIS, ontology-based information systems, and HBIM are increasingly integrated to connect geometric representation with semantic and management information [10,56]. Recent HBIM research also highlights the importance of incorporating non-geometric data, conservation knowledge, monitoring information, and interoperable data structures into digital heritage workflows [57]. In this context, the HDNA framework contributes to this broader shift by translating facade-level morphological and material information into a standardized descriptor vector that can be compared, clustered, monitored, and potentially integrated into digital heritage inventories or HBIM-based information systems.

At the facade scale, recent studies have proposed integrated approaches for analyzing architectural composition and modeling heritage urban facades within HBIM environments [58]. Similarly, HBIM–WebGIS research shows that semantic and geometric heritage information can be reused, managed, shared, and visualized through accessible digital platforms [59].

Taken together, these characteristics position HDNA as a reproducible bridge between quantitative facade classification and digital heritage management. Its practical value is therefore considered below in relation to structured facade inventories, conservation prioritization, longitudinal monitoring, and potential integration with AI- or HBIM-supported workflows.

From a heritage management perspective, the HDNA framework can support several practical tasks beyond typological classification. In terms of heritage management, the HDNA methodology can be used for a range of purposes apart from facade classification. Firstly, the HDNA approach can be applied to the creation of inventory databases of the facades using their standardized numerical descriptors that include information about the morphological and material characteristics. Thus, heritage managers will have

the opportunity not only to describe facades verbally but also to measure some parameters in relation to them. Secondly, the numerical values of the facade descriptors can be used to identify facades that should be prioritized for conservation. For instance, high MDR values might indicate facades that were subjected to considerable interventions at a later stage, whereas high VR values may reflect biological colonization of the facade, its abandonment or potential problems with maintenance.

This interpretation is also consistent with recent facade vulnerability assessment research, in which facade conservation state, openings, geometry, orientation, and connected elements are treated as parameters for identifying conservation vulnerability and prioritization needs [60].

Finally, the HDNA approach can be employed to monitor the evolution of a facade in time. In case one records the same facade at several time moments, variations in the HDNA vector can indicate increasing amounts of vegetation, more modern alterations, deterioration of materials, and changes in the patterns of openings and projections. In such a way, HDNA can serve as a quantitative baseline for continuous monitoring and heritage preservation-focused documentation. Moreover, since all facades are characterized by fixed-length descriptor vectors, such information can be readily incorporated into digital inventories of cultural heritage, heritage databases based on geographic information systems, or HBIM environments in the future. Hence, the real practical significance of HDNA is not in substituting expert preservation judgment, but rather in providing a comparative data layer for inventories.

4.5. Limitations and Future Research

The present study should therefore be interpreted as a methodological proof-of-concept rather than a universally applicable facade taxonomy. While the descriptor system itself is designed to be potentially transferable, the distribution of numeric values of descriptors and the corresponding facade groups can be assumed to differ depending on traditions in architecture, building materials, climates, and local building customs. This should be further tested in subsequent research.

It is crucial to distinguish between the transferability of the HDNA descriptors system and the transferability of the individual facade families determined in this study. The nine-descriptor HDNA system is designed as a potentially transferable representation system, as it consists of measurable attributes of facades that can be collected based on orthophoto documentation. At the same time, it must be mentioned that the four facade families discovered in the Tepebağ project cannot be regarded as universal categories of facades. They are site-specific groups resulting from certain conditions in architecture, materials, climate, and preservation.

Accordingly, transferability in the present study refers to the applicability of the descriptor-based workflow rather than to the universal validity of the resulting cluster labels. In other historic settlements, the same HDNA descriptors may produce different value distributions, different dominant variables, or different cluster structures. Future cross-context applications should therefore test whether the descriptor set remains meaningful across different architectural traditions and should refine or expand the descriptor ontology where necessary.

Future research may also examine how HDNA descriptors could be incorporated into HBIM-supported smart heritage and heritage digital twin environments. As stated in recent systematic reviews, HBIM has been described as a useful technique in the process of the digitalization of architectural heritage that is used for restoration, management, exposition, educational purposes, and even the whole set of operations known as smart heritage [61]. At the same time, the literature on heritage digital twin reveals that the concept of digital heritage environments is evolving into dynamic systems which will

integrate monitoring data, conservation data, semantic data, analytical results, and so forth [62].

Recent openBIM-based heritage studies similarly emphasize that semantic interoperability and structured data dictionaries can improve the consistency, usability, and long-term preservation value of heritage information models [63]. Related research on orthophoto-based digital twin monitoring also demonstrates that automatically generated orthophotos and semantic segmentation can be linked to digital twin environments for periodic monitoring of building fabric components [64].

Several limitations of the present study should be acknowledged. First, the HDNA framework was developed and evaluated using a dataset of 31 historic facades located within a single urban conservation area. Although the selected facades represent a broad range of morphological and material characteristics, the resulting facade families inevitably reflect the specific architectural conditions of Tepebağ. Consequently, the transferability of the identified family structures should be tested through applications in different historic settlements, cultural contexts, and architectural traditions. The resampling analysis also showed that the cluster solution was more sensitive to 80% subsampling than to single-case deletion; therefore, the reported stability estimates should be interpreted within the context of the small and unevenly distributed proof-of-concept dataset.

The second limitation pertains to the methodology of the descriptor selection procedure, which entails a certain level of subjective approach. While all variables were based on observable and quantifiable features of the facades, another set of descriptors might uncover other patterns between facades. Therefore, future research can develop the descriptor system to include other descriptors such as ornaments, colors, signs of degradation, structural systems, architectural elements, and transformations over time.

Another drawback refers to the current need for manual annotation and extraction of descriptors. While manual annotation provides transparency, scalability remains limited for large datasets. Moreover, because the present study did not include a formal inter-rater reliability assessment, possible annotator-dependent variation cannot be fully excluded. Future applications should therefore test the annotation protocol with multiple independent annotators and report agreement statistics for both morphological and material descriptor extraction. Future applications of HDNA may consider introducing automated pipelines that would use semantic segmentation and object recognition with the aid of machine learning in order to detect features.

Recent studies on facade and heritage segmentation support this direction. Automated material segmentation has been used to identify facade materials and colors for urban conservation [65], while newly developed photogrammetric image–point cloud datasets aim to accelerate semantic segmentation and scan-to-BIM workflows for heritage buildings [66]. In parallel, facade segmentation datasets developed for irregular architectural compositions demonstrate the growing importance of annotated facade data for training and evaluating automated recognition models [67].

A further constraint lies in the fact that the analysis process remains two-dimensional. Although the use of orthophotos provides an accurate and user-friendly representation of facades, they fail to convey any form of volume, spatial depth, or three-dimensional nature of the architecture. Contemporary studies have highlighted the fact that documenting heritage sites is not only about the geometrical aspect but also incorporates semantically rich knowledge-based information environments that can represent geometrical, historical, environmental, and analytical data in a single digital framework [68,69]. It would be pertinent for future research to study the incorporation of the HDNA descriptors with point clouds and HBIM environments.

Despite these limitations, the study demonstrates the feasibility of representing facade identity through measurable descriptor combinations. With the possibility of combining morphological and material descriptors into one representation system, HDNA provides the basis for future heritage analysis and contributes to a method for classifying facades that would be transferable to future digital heritage platforms.

The present study does not claim that the nine-descriptor set outperforms existing qualitative typologies or alternative quantitative feature schemes. No direct benchmark against expert-based classifications or alternative descriptor sets was conducted. The contribution of HDNA lies instead in proposing a compact, interpretable, and standardized representation framework. Future studies should compare the proposed descriptor ontology with alternative feature sets and established typological classifications.

5. Conclusions

This study introduced the heritage DNA (HDNA) framework as a quantitative methodology for the classification of historic facades through the integration of morphological and material descriptors. The new approach, which differs from traditional facade typologies that rely predominantly on qualitative assessment, enables facade attributes to be transformed into quantitative data for systematic comparison.

Principal Component Analysis revealed that the first two principal components explained 50.38% of the total variance within the dataset, indicating that a substantial proportion of facade variability can be interpreted through a reduced descriptor space. Hierarchical clustering based on the standardized HDNA matrix identified four distinct facade clusters. Subsequent interpretation of their morphological and material characteristics enabled the definition of four facade families representing different combinations of material composition, facade articulation, vegetation presence, and opening configurations.

The final four-cluster solution produced a silhouette coefficient of 0.256. Rather than being interpreted as a standalone decision criterion, this value was evaluated together with the Calinski–Harabasz and Davies–Bouldin indices, dendrogram structure, PCA distribution, descriptor profiles, and collective facade display. Additional robustness analyses showed that the specified BIRCH configuration reproduced the Ward partition exactly ($ARI = 1.000$), while k-means showed high agreement with Ward ($ARI = 0.909$). Resampling analyses further indicated high local stability under single-case deletion and lower, but still substantial, agreement under the more demanding 80% subsampling procedure.

The identified facade families—Masonry-Dominated Facades, Plaster-Rich Residential Facades, Vegetation-Affected Facades, and Articulated Timber-Framed Facades—demonstrate that facade identity can be represented quantitatively through measurable descriptor combinations rather than solely through descriptive architectural interpretation. The results further indicate that historic facade similarity emerges from the interaction of morphological and material characteristics rather than from any single architectural attribute.

Methodologically, the HDNA contribution is not in the DNA metaphor per se, but in the application of PCA and clustering techniques, as such precursors existed for interpreting the built environment based on the analogy to DNA, as well as for multivariate classification using quantitative methods. The contribution is in creating a framework where individual historical facades are defined as quantitative identity profiles where the morphological and material descriptors are explicitly divided into MDNA and MaDNA subspaces and recomposed into an integrated HDNA format. Such a format creates a common format for facade description that can be compared, classified, monitored, and used for possible integration with future digital heritage information systems. The

framework offers a potentially transferable analytical approach that can be used for comparative research across various historic settlements and buildings, assuming the validation process in other contexts. Moreover, the descriptor-based framework of HDNA can enable future integration with the HBIM environment and heritage management processes. Practically, HDNA can be used as a quantitative support layer for the creation of heritage inventories, prioritizing conservation, facade condition monitoring, and future integration with a GIS/HBIM digital heritage management system.

Overall, the findings demonstrate that historic facade identity can be encoded, analyzed, and classified through quantitative descriptors. The proposed HDNA approach thus provides an innovative methodology to support the systematic study of facade typologies, moving toward an analysis-based approach in heritage studies. Instead of prescribing a rigid classification system for all historic cities, the proposed HDNA framework creates a standardized approach to describing facades in order to facilitate validation. Future projects utilizing larger datasets and various historic districts will help refine the ontology of descriptors and validate their generalizability.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/buildings16142799/s1>.

Author Contributions: Conceptualization, O.K.; methodology, O.K. and H.D.; software, O.K.; validation, O.K. and H.D.; formal analysis, O.K. and H.D.; investigation, O.K.; resources, H.D.; data curation, O.K.; writing—original draft preparation, O.K. and H.D.; writing—review and editing, O.K. and H.D.; visualization, O.K.; supervision, O.K. and H.D.; project administration, O.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data presented in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

HDNA	Heritage DNA
MDNA	Morphological DNA
MaDNA	Material DNA
WR	Window Ratio
DR	Door Ratio
PCT	Projection Count
MR	Masonry Ratio
TR	Timber Ratio
PLR	Plaster Ratio
MDR	Modern Intervention Ratio
VR	Vegetation Ratio
MDI	Material Diversity Index
PCA	Principal Component Analysis
HBIM	Heritage Building Information Modeling
GIS	Geographic Information System
WebGIS	Web Geographic Information System
SD	Standard Deviation
ARI	Adjusted Rand Index
BIRCH	Balanced Iterative Reducing and Clustering using Hierarchies

References

- Çakıcı, F.Z.; Kaçdı, R. Systematic Analysis of the Digital Technologies Used in the Documentation of Historical Buildings. *Cult. Herit. Sci.* **2023**, *4*, 69–77. <https://doi.org/10.58598/cuhes.1344379>.
- Yastıklı, N. Documentation of Cultural Heritage Using Digital Photogrammetry and Laser Scanning. *J. Cult. Herit.* **2007**, *8*, 423–427. <https://doi.org/10.1016/j.culher.2007.06.003>.
- Dore, C.; Murphy, M. Integration of Historic Building Information Modeling and 3D GIS for Recording and Managing Cultural Heritage Sites. In Proceedings of the 18th International Conference on Virtual Systems and Multimedia, Milan, Italy, 2–5 September 2012; pp. 369–376. <https://doi.org/10.1109/VSM.2012.6365947>.
- Remondino, F.; Campana, S. (Eds.) *3D Recording and Modelling in Archaeology and Cultural Heritage: Theory and Best Practices*; BAR International Series 2598; Archaeopress: Oxford, UK, 2014.
- Stylianidis, E.; Remondino, F. (Eds.) *3D Recording, Documentation and Management of Cultural Heritage*; Whittles Publishing: Dunbeath, UK, 2016.
- Diara, F.; Rinaudo, F. Building Archaeology Documentation and Analysis through Open Source HBIM Solutions via NURBS Modelling. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2020**, *43*, 1381–1388. <https://doi.org/10.5194/isprs-archives-XLIII-B2-2020-1381-2020>.
- Bruno, N.; Roncella, R. A Restoration-Oriented HBIM System for Cultural Heritage Documentation: The Case Study of Parma Cathedral. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2018**, *42*, 171–178. <https://doi.org/10.5194/isprs-archives-XLII-2-171-2018>.
- Lovell, L.; Davies, R.; Hunt, D. The Application of Historic Building Information Modelling (HBIM) to Cultural Heritage: A Review. *Heritage* **2023**, *6*, 6691–6717. <https://doi.org/10.3390/heritage6100350>.
- Murphy, M.; McGovern, E.; Pavia, S. Historic Building Information Modelling (HBIM). *Struct. Surv.* **2009**, *27*, 311–327. <https://doi.org/10.1108/02630800910985108>.
- Penjor, T.; Banihashemi, S.; Hajirasouli, A.; Golzad, H. Heritage Building Information Modeling (HBIM) for Heritage Conservation: Framework of Challenges, Gaps, and Existing Limitations of HBIM. *Digit. Appl. Archaeol. Cult. Herit.* **2024**, *35*, e00366. <https://doi.org/10.1016/j.daach.2024.e00366>.
- Zhang, Z.; Zou, Y. Research Hotspots and Trends in Heritage Building Information Modeling: A Review Based on CiteSpace Analysis. *Humanit. Soc. Sci. Commun.* **2022**, *9*, 394. <https://doi.org/10.1057/s41599-022-01414-y>.
- Fai, S.; Graham, K.; Duckworth, T.; Wood, N.; Attar, R. Building Information Modelling and Heritage Documentation. In Proceedings of the XXIII International CIPA Symposium, Prague, Czech Republic, 12–16 September 2011.
- Oliveira, V.; Porta, S. Quantitative and Qualitative Analysis in Urban Morphology: Systematic Legacy and Latest Developments. *Proc. Inst. Civ. Eng. Urban Des. Plan.* **2025**, *178*, 75–87. <https://doi.org/10.1680/jurdp.24.00047>.
- Huang, K.; Kang, P.; Zhao, Y. Quantitative Research of Street Interface Morphology in Urban Historic Districts: A Case Study of West Street Historic District, Quanzhou. *Herit. Sci.* **2024**, *12*, 226. <https://doi.org/10.1186/s40494-024-01351-w>.
- Kropf, K. *The Handbook of Urban Morphology*; Wiley: Chichester, UK, 2017. <https://doi.org/10.1002/9781118747711>.
- Oliveira, V. *Urban Morphology: An Introduction to the Study of the Physical Form of Cities*; Springer: Cham, Switzerland, 2016. <https://doi.org/10.1007/978-3-319-32083-0>.
- Jabłońska, J.; Telesińska, M.; Adamska, A.; Gronostajska, J. The Architectural Typology of Contemporary Facades for Public Buildings in the European Context. *Arts* **2022**, *11*, 11. <https://doi.org/10.3390/arts11010011>.
- Wan Ali, W.N.; Ahmad, A.G. Taxonomy of Traditional and Modern Architectural Styles of Heritage Shophouse Facade at Ipoh, Perak, Malaysia. *Int. J. Acad. Res. Bus. Soc. Sci.* **2020**, *10*, 1426–1437. <https://doi.org/10.6007/IJARBS/v10-i11/8366>.
- Saylan, S.; Gürer, T.K. Typomorphology: A Methodological Approach to Context Analysis in Architectural Design Education. *ICONARP Int. J. Archit. Plan.* **2024**, *12*, 189–210. <https://doi.org/10.15320/ICONARP.2024.281>.
- Sonne, W. Urban Typology, Morphology, Iconography. In *Metropolitan Research: Methods and Approaches*; Gurr, J., Parr, R., Hardt, D., Eds.; Transcript Verlag: Bielefeld, Germany, 2022. <https://doi.org/10.1515/9783839463109-002>.
- Fleischmann, M.; Feliciotti, A.; Romice, O.; Porta, S. Methodological Foundation of a Numerical Taxonomy of Urban Form. *Environ. Plan. B Urban Anal. City Sci.* **2022**, *49*, 1283–1299. <https://doi.org/10.1177/23998083211059835>.
- Benni, S.; Carfagna, E.; Torreggiani, D.; Maino, E.; Bovo, M.; Tassinari, P. Multidimensional Measurement of the Level of Consistency of Farm Buildings with Rural Heritage: A Methodology Tested on an Italian Case Study. *Sustainability* **2019**, *11*, 4242. <https://doi.org/10.3390/su11154242>.
- Cheng, G.; Li, Z.; Xia, S.; Gao, M.; Ye, M.; Shi, T. Research on the Spatial Sequence of Building Facades in Huizhou Regional Traditional Villages. *Buildings* **2023**, *13*, 174. <https://doi.org/10.3390/buildings13010174>.

24. He, Y.; Cheng, B. Research on the Spatial Sequence of Building Facades in Historic Towns in the Chengdu Plain Region of China. *Buildings* **2026**, *16*, 838. <https://doi.org/10.3390/buildings16040838>.
25. Ornelas, C.; Sousa, F.; Guedes, J.M.; Breda-Vázquez, I. Monitoring and Assessment Heritage Tool: Quantify and Classify Urban Heritage Buildings. *Cities* **2023**, *137*, 104274. <https://doi.org/10.1016/j.cities.2023.104274>.
26. Goodarzi, P.; Ansari, M.; Mahdavinnejad, M.; Russo, A.; Haghighatbin, M.; Pour Rahimian, F. Morphological Analysis of Historical Landscapes Based on Cultural DNA Approach. *Digit. Appl. Archaeol. Cult. Herit.* **2023**, *30*, e00277. <https://doi.org/10.1016/j.daach.2023.e00277>.
27. Le, V.A.; Nguyen, T.K.N. Discussion of So-Called “Architectural Heritage DNA” via a Case Study of the Conservation of the Nara Palace Site, Japan. *Buildings* **2024**, *14*, 132. <https://doi.org/10.3390/buildings14010132>.
28. Votsis, A.; Haavisto, R. Urban DNA and Sustainable Cities: A Multi-City Comparison. *Front. Environ. Sci.* **2019**, *7*, 4. <https://doi.org/10.3389/fenvs.2019.00004>.
29. Ye, Y.; Zhao, T.; Shi, X.; Liu, P.; Fei, T. The Identification of Cultural Genes in Historic Districts and Their Influences on Cultural Perception: Case Study in Central Street in Harbin, China. *J. Asian Archit. Build. Eng.* **2025**, *24*, 2430–2446. <https://doi.org/10.1080/13467581.2024.2373823>.
30. Sun, J.; Shao, L. Interpretation of Historic Urban Landscape Genes: A Case Study of Harbin, China. *Land* **2024**, *13*, 1988. <https://doi.org/10.3390/land13121988>.
31. Zhang, W.; Yang, H. Quantitative Research of Traditional Village Morphology Based on Spatial Genes: A Case Study of Shaanxi Province, China. *Sustainability* **2024**, *16*, 9003. <https://doi.org/10.3390/su16209003>.
32. Kaya, H.S.; Bölen, F. Urban DNA: Morphogenetic Analysis of Urban Pattern. *ICONARP Int. J. Archit. Plan.* **2017**, *5*, 10–41. <https://doi.org/10.15320/ICONARP.2017.15>.
33. Do, D.T. Uncovering the Architectural DNA of Da Lat: A Pattern Language Approach to Urban Heritage Conservation. *J. Asian Archit. Build. Eng.* **2026**, 1–19. <https://doi.org/10.1080/13467581.2026.2629652>.
34. Wang, Z.; Yu, H.; Wang, J.; Zhang, Z.; Song, C. Typology-Led Innovation in Urban Furniture: A Framework for Reconciling Cultural Heritage and Functional Modernization in Macau. *Buildings* **2026**, *16*, 1371. <https://doi.org/10.3390/buildings16071371>.
35. Fan, D.; Maliki, N.Z.B.; He, C.; Bi, Y.; Yu, S. Cultural Gene Characterization and Mapping of Traditional Tibetan Village Landscapes in Western Sichuan, China. *npj Herit. Sci.* **2025**, *13*, 317. <https://doi.org/10.1038/s40494-025-01877-7>.
36. Minchin, S.; Lodge, J. Understanding Biochemistry: Structure and Function of Nucleic Acids. *Essays Biochem.* **2019**, *63*, 433–456. <https://doi.org/10.1042/EBC20180038>.
37. Crick, F.H.C.; Barnett, L.; Brenner, S.; Watts-Tobin, R.J. General Nature of the Genetic Code for Proteins. *Nature* **1961**, *192*, 1227–1232. <https://doi.org/10.1038/1921227a0>.
38. Xing, Y.; Yang, S.; Fahy, C.; Harwood, T.; Shell, J. Capturing the Past, Shaping the Future: A Scoping Review of Photogrammetry in Cultural Building Heritage. *Electronics* **2025**, *14*, 3666. <https://doi.org/10.3390/electronics14183666>.
39. De Fino, M.; Galantucci, R.A.; Fatiguso, F. Condition Assessment of Heritage Buildings via Photogrammetry: A Scoping Review from the Perspective of Decision Makers. *Heritage* **2023**, *6*, 7031–7066. <https://doi.org/10.3390/heritage6110367>.
40. Shannon, C.E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423.
41. Jolliffe, I.T. *Principal Component Analysis*; Springer: New York, NY, USA, 2002.
42. Abdi, H.; Williams, L.J. Principal Component Analysis. *Wiley Interdiscip. Rev. Comput. Stat.* **2010**, *2*, 433–459. <https://doi.org/10.1002/wics.101>.
43. Everitt, B.S.; Landau, S.; Leese, M.; Stahl, D. *Cluster Analysis*, 5th ed.; Wiley: Chichester, UK, 2011.
44. Ward, J.H. Hierarchical Grouping to Optimize an Objective Function. *J. Am. Stat. Assoc.* **1963**, *58*, 236–244. <https://doi.org/10.1080/01621459.1963.10500845>.
45. Rousseeuw, P.J. Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [https://doi.org/10.1016/0377-0427\(87\)90125-7](https://doi.org/10.1016/0377-0427(87)90125-7).
46. Caliński, T.; Harabasz, J. A Dendrite Method for Cluster Analysis. *Commun. Stat.* **1974**, *3*, 1–27. <https://doi.org/10.1080/03610927408827101>.
47. Davies, D.L.; Bouldin, D.W. A Cluster Separation Measure. *IEEE Trans. Pattern Anal. Mach. Intell.* **1979**, PAMI-1, 224–227. <https://doi.org/10.1109/TPAMI.1979.4766909>.
48. MacQueen, J. Some Methods for Classification and Analysis of Multivariate Observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*; University of California Press: Berkeley, CA, USA, 1967; Volume 1, pp. 281–297.

49. Zhang, T.; Ramakrishnan, R.; Livny, M. BIRCH: An Efficient Data Clustering Method for Very Large Databases. In Proceedings of the 1996 ACM SIGMOD International Conference on Management of Data, Montreal, QC, Canada, 4–6 June 1996; pp. 103–114. <https://doi.org/10.1145/233269.233324>.
50. Hubert, L.; Arabie, P. Comparing Partitions. *J. Classif.* **1985**, *2*, 193–218. <https://doi.org/10.1007/BF01908075>.
51. Miller, R.G. The Jackknife—A Review. *Biometrika* **1974**, *61*, 1–15. <https://doi.org/10.1093/biomet/61.1.1>.
52. Hennig, C. Cluster-Wise Assessment of Cluster Stability. *Comput. Stat. Data Anal.* **2007**, *52*, 258–271. <https://doi.org/10.1016/j.csda.2006.11.025>.
53. Jaccard, P. The Distribution of the Flora in the Alpine Zone. *New Phytol.* **1912**, *11*, 37–50. <https://doi.org/10.1111/j.1469-8137.1912.tb05611.x>.
54. D’Orazio, M.; Gianangeli, A.; Monni, F.; Quagliarini, E. Automatic Monitoring of the Biocolonisation of Historical Building’s Facades through Convolutional Neural Networks (CNN). *J. Cult. Herit.* **2024**, *70*, 80–89. <https://doi.org/10.1016/j.culher.2024.08.012>.
55. Paiva, D.S.; Fernandes, L.; Portugal, A. Fungi and Stone Heritage Conservation: Friend, Foe, or a Bit of Both. *J. Fungi* **2026**, *12*, 128. <https://doi.org/10.3390/jof12020128>.
56. Yang, X.; Grussenmeyer, P.; Koehl, M.; Macher, H.; Murtiyoso, A.; Landes, T. Review of Built Heritage Modelling: Integration of HBIM and Other Information Techniques. *J. Cult. Herit.* **2020**, *46*, 350–360. <https://doi.org/10.1016/j.culher.2020.05.008>.
57. Parente, M.; Bruno, N.; Ottoni, F. HBIM and Information Management for Knowledge and Conservation of Architectural Heritage: A Review. *Heritage* **2025**, *8*, 306. <https://doi.org/10.3390/heritage8080306>.
58. Hussein, K.A.; Abraham, A.R.; Mohammed, A.S.; Ali, K.A. Digital Preservation of Heritage Urban Facades: An Integrated Approach Using Historic Building Information Modeling and Shape Grammar Analysis. *Digit. Appl. Archaeol. Cult. Herit.* **2024**, *34*, e00383. <https://doi.org/10.1016/j.daach.2024.e00383>.
59. Dionizio, R.F.; López-Chao, V. Integrating HBIM and WebGIS for the Documentation, Visualization and Management of Modern Architectural Heritage Sites. *Int. J. Archit. Herit.* **2025**, *19*, 3530–3549. <https://doi.org/10.1080/15583058.2025.2502528>.
60. Fontenele, A.; Campos, V.; Matos, A.M.; Mesquita, E. A Vulnerability Index Formulation for Historic Facades Assessment. *J. Build. Eng.* **2023**, *64*, 105552. <https://doi.org/10.1016/j.job.2022.105552>.
61. Cui, H.; Wu, J. How Architectural Heritage Is Moving to Smart: A Systematic Review of HBIM. *Buildings* **2025**, *15*, 2664. <https://doi.org/10.3390/buildings15152664>.
62. Arsalan, H.; Heesom, D.; Moore, N. From Heritage Building Information Modelling Towards an ‘Echo-Based’ Heritage Digital Twin. *Heritage* **2025**, *8*, 33. <https://doi.org/10.3390/heritage8010033>.
63. Argasiński, K.; Tomczak, A. Enhancing Semantic Interoperability of Heritage BIM-Based Asset Preservation. *Heritage* **2025**, *8*, 410. <https://doi.org/10.3390/heritage8100410>.
64. Li, J.; Tao, B.; Bosché, F.; Lu, C.X.; Wilson, L. Automated Generation and Semantic Segmentation of Roof Orthophoto for Digital Twin-Based Monitoring of Slated Roofs. *Autom. Constr.* **2024**, *167*, 105725. <https://doi.org/10.1016/j.autcon.2024.105725>.
65. Xie, J.; Li, M.; Wu, J.; Zhang, X.; Zhang, J. Semantic Segmentation of Building Facade Materials and Colors for Urban Conservation. *npj Herit. Sci.* **2025**, *13*, 378. <https://doi.org/10.1038/s40494-025-01888-4>.
66. Pellis, E.; Masiero, A.; Betti, M.; Tucci, G.; Grussenmeyer, P. A Photogrammetric Image-Point Dataset for the Semantic Segmentation of Heritage Buildings. *Data Brief.* **2025**, *60*, 111661. <https://doi.org/10.1016/j.dib.2025.111661>.
67. Wei, J.; Hu, Y.; Zhang, S.; Liu, S. Irregular Facades: A Dataset for Semantic Segmentation of the Free Facade of Modern Buildings. *Buildings* **2024**, *14*, 2602. <https://doi.org/10.3390/buildings14092602>.
68. Costa-Jover, A.; Núñez-Andrés, M.A.; Buill Pozuelo, F. Advances in Geomatics Data Acquisition for Architectural Heritage: A Critical Review Toward Sustainable, Data-Driven Conservation. *Sustainability* **2026**, *18*, 3003. <https://doi.org/10.3390/su18063003>.
69. de Oliveira, E.; Sousa, M.N.P.; Correa, F.R. Towards Digital Twins for Heritage Buildings: A Workflow Proposal. *Int. J. Archit. Comput.* **2023**, *21*, 712–729. <https://doi.org/10.1177/14780771231168226>.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.