Journal of Archaeological Science: Reports 55 (2024) 104521

Contents lists available at ScienceDirect

Journal of Archaeological Science: Reports

journal homepage: www.elsevier.com/locate/jasrep

Beyond the naked eye: Employing machine learning and computer vision to explore Iberian oculated idols decorative patterns

Joaquín Jiménez-Puerto

Department of Prehistory, Archaeology and Ancient History (PREMEDOC Research Group), University of Valencia, Av. Blasco Ibanez, 28, 46010 Valencia, Spain

ARTICLE INFO

Keywords:
Computer Vision
Machine Learning
Copper Age
Iberian Peninsula
Oculated Idols on Long Bones

ABSTRACT

The study of the Copper Age in the Iberian Peninsula uncovers many enigmas, among them the oculated idols. Characterized by their distinctive eye-shaped motifs on long bones, these idols provide a window into the Chalcolithic communities' symbolic world. However, conventional archaeological techniques have fallen short in fully deciphering their narratives.

This study uses advanced machine learning techniques to examine the decorative patterns on these oculated idols. We introduce a toolkit called ImageClusterVis, which utilizes computer vision algorithms to extract features more effectively. By applying Gaussian Mixture Models, Spectral Clustering, and Self-Organizing Maps, we can identify and classify patterns that are not visible to the human eye. Our approach helps uncover artistic variations and regional differences in the idols that have not been recognized in past research, showcasing the potential of machine learning to reveal new dimensions of ancient art.

Lay Summary:

The analysis of visual patterns in material artefacts has always been a central part of archaeological practice, but it has typically relied on qualitative methods. With the advent of digital archaeology, there has been a growing interest in developing more objective and quantitative techniques to compare and classify artefacts based on visual attributes. In this article, we present a novel methodology that employs computer vision algorithms and machine learning to extract robust visual descriptors and cluster archaeological images in an unsupervised manner. This study explores the world of Copper Age Iberian oculated idols on long bones by examining their decorative patterns, to uncover new insights into the regional distinctions, and cultural significances of these enigmatic figures, revealing a previously unseen depth to Copper Age society and artistry. This novel research not only advances our understanding of prehistoric Iberian cultures but also sets a new case study for the interdisciplinary application of AI in archaeology.

1. Introduction

Positioned at the southwestern edge of Europe, the Iberian Peninsula has historically been a fertile ground for the convergence of diverse communities. This strategic confluence of cultures fostered a unique melting pot of cultural and technological innovations, particularly noticeable during the Copper Age. Among the myriad of artifacts from this era, the oculated idols are particularly notable. Found mainly in the southern and central regions, these idols, with their unique eye motifs carved on long bones, date back to the 3rd millennium BCE and reflect the intricate spiritual and artistic life of the early Iberian people (Maicas-Ramos, 2007; Pascual-Benito, 2010; Soler-Diaz, 2017).

Despite extensive research, the full significance and purpose behind these idols remain largely speculative. Traditional archaeological methods have primarily focused on their physical attributes and contexts, offering limited insight into their deeper meanings. However, the complex decorative patterns etched onto these idols suggest a level of artistic sophistication and symbolic importance that merits a deeper investigation. Recognizing this gap, our study introduces a novel approach, leveraging the power of modern ML and CV technologies to explore the intricate designs of these ancient artefacts.

This paper represents a novel approach to test the applicability of various advanced algorithms, such as Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), Scale-Invariant Feature Transform (SIFT), and Contour-Edge Detection (CED), specifically on the decorative motifs of the oculated idols. Our methodology, centered around the script ImageClusterVis—a sophisticated Python-based toolkit—aims not only to navigate the complex visual patterns of the idols but also to

E-mail address: joaquin.jimenez@uv.es.

https://doi.org/10.1016/j.jasrep.2024.104521
Received 5 February 2024; Received in revised form 20 March 2024; Accepted 5 April 2024
Available online 10 April 2024

critically evaluate the advantages and disadvantages of each algorithm's application in this unique archaeological context.

Furthermore, by applying various clustering algorithms, including Gaussian Mixture Models (GMM), Spectral Clustering, and Self-Organizing Maps (SOM), we categorize and interpret the visual elements of the idols, shedding light on the rich tapestry of Copper Age artistic expression. This analysis is instrumental in understanding the efficacy and limitations of each technique, providing valuable insights into their suitability for decoding the intricate designs of prehistoric artifacts. This analysis not only contributes to our understanding of prehistoric artistry but also demonstrates the potential of digital technologies to enhance archaeological research, enabling a deeper appreciation of our ancient heritage. Moreover, this study underscores the importance of integrating traditional archaeological inquiry with cutting-edge technological methodologies, offering a comprehensive framework for future research in the field.

By focusing on the methodological innovation and the critical assessment of computational techniques, this study aims to provide a small contribution to bridge the gap between conventional archaeological methods and the forefront of technological advancements, fostering a more profound and nuanced exploration of our ancient past. By employing advanced unsupervised machine learning (ML) (Bickler, 2021) and computer vision (CV) techniques (Szeliski, 2022), we aim to unlock the secrets hidden within the intricate designs of these ancient artefacts. The core of our methodology is ImageClusterVis, a Python-based toolkit designed for the clustering and visualization of image data.

ImageClusterVis navigates through the complex visual patterns of the idols using some advanced algorithms like Histogram of Oriented Gradients (HOG) (Dalal and Triggs, 2005), Local Binary Patterns (LBP) (Ojala et al., 1996), the Scale-Invariant Feature Transform (SIFT) (Lowe, 2004) and the Canny edge detectio (CED) (Canny, 1986) for feature extraction. This approach allows us to identify and analyze the key visual elements that define each idol, providing insights that go beyond what is visible to the naked eye. Our journey continues with the application of various clustering algorithms, including Gaussian Mixture Models (GMM) (Bishop and Nasrabadi, 2006), Spectral Clustering (Von Luxburg, 2007), and Self-Organizing Maps (SOM) (Kohonen et al., 2012). Each of these algorithms offers a unique perspective on the categorization and interpretation of the idols' designs.

As we get deeper into our analysis, the ML algorithms begin to reveal patterns and groupings in the idol decorations that were previously unseen. These findings suggest a rich tapestry of artistic expression and cultural connectivity, pointing to intricate regional variations among the idols. Furthermore, our study demonstrates the transformative potential of digital technologies in archaeological research, opening new avenues for understanding and interpreting the artistic legacy of Copper Age societies.

Despite extensive research, the full significance and purpose behind these idols remain largely speculative. Traditional archaeological methods have primarily focused on their physical attributes and contexts, offering limited insight into their deeper meanings. Recognizing this gap, our study introduces a novel approach, leveraging the power of modern ML and CV technologies to explore the intricate designs of these ancient artefacts. This paper represents a novel approach to fuse traditional archaeological inquiry with cutting-edge technological methodologies. By employing our script ImageClusterVis, a sophisticated Python-based toolkit, we aim to navigate the complex visual patterns of the idols using advanced algorithms such as HOG, LBP, SIFT and CED. This innovative method allows us to unearth previously unseen patterns and connections within the idols' designs, offering fresh insights into their artistic and cultural significance. Furthermore, through the application of various clustering algorithms, including Gaussian Mixture Models (GMM), Spectral Clustering, and Self-Organizing Maps (SOM), we will categorize and interpret the visual elements of the idols, shedding light on the rich tapestry of Copper Age artistic expression. This analysis not only contributes to our understanding of prehistoric artistry but also demonstrates the potential of digital technologies to revolutionize archaeological research, enabling a deeper appreciation of our ancient heritage.

1.1. Machine learning in archaeology: pioneering applications

Our study would like to highlight the invaluable role of computational tools in expanding the archaeological analytical framework. By integrating ML and CV into our research, we would like to open new paths for interpreting the enigmatic symbols of our ancestors, contributing to the broader field of archaeological science. In recent years, ML has significantly impacted various facets of archaeological research. Image analysis techniques have become invaluable tools in archaeology, enabling discoveries and detailed study of artefacts, sites, and landscapes. For example, pattern recognition algorithms have been applied to analyze large collections of ancient pottery sherds, allowing for quicker classification of styles and motifs compared to manual methods (Brahim et al., 2023; Chetouani et al., 2020; Karasik and Smilansky, 2008; Makridis and Daras, 2013). This has provided insights into production techniques, trade networks, and cultural trends in civilizations reliant on pottery, like ancient Greece. In the realm of architecture, 3D laser scanning and photogrammetry have been revolutionary in digitally preserving ruins and monuments that are deteriorating (McCarthy, 2014; Remondino and El-Hakim, 2006; Yastikli, 2007). The detailed 3D models created allowed researchers to reconstruct and analyze architectural features far more accurately than with 2D images or drawings. Adams (2014) focused on the wear and tool marks on artefacts, utilizing ML to offer insights into the usage and craftsmanship of prehistoric tools. At a broader level, satellite imagery and aerial photography have opened up entirely new avenues in landscape archaeology by revealing buried structures across large regions. This has led to discoveries of lost settlements and even entire cities that were not visible from the ground (Canuto et al., 2018; Lasaponara and Masini, 2011; Parcak, 2009). These diverse applications underscore the transformative role of ML in advancing archaeological research.

Building on the established groundwork, this study aims to extend the application of ML in archaeology to the subtle domain of oculated idols and their decorative patterns, by recognizing its potential in uncovering nuanced details in prehistoric artefacts. Our methodology, inspired by successful applications in pottery analysis, site reconstruction, and artefact examination, is designed to identify and analyze the intricate designs of these idols. Our approach focuses on CV and unsupervised ML algorithms to tackle the oculated problem. This approach can be especially useful for archaeological environments where the absence of clear typologies and classifications makes impossible the employment of labels for supervised ML. By doing so, we hope to offer new insights into their artistic and cultural significance, thereby contributing to the broader understanding of these enigmatic symbols from the past.

2. Idols in context

Oculated idols, often carved on long bones, sometimes on stones or pottery, are thought to have played a significant role in the spiritual and daily life of the Copper Age communities. The term "oculated," derived from the Latin "oculus" meaning eye, aptly describes the most striking feature of these idols — their prominent eye-like designs (Ballester, 1946). These motifs, ranging from simple concentric circles to intricate geometric patterns, are more than mere decorative elements; they are believed to be imbued with significant symbolic meaning, perhaps serving as talismans or objects of veneration, protective emblems, or representations of deities (Almagro-Gorbea, 1969; Gimbutas, 1974; Gonçalves, 1999; Lillios, 2008; Maicas-Ramos, 2007).

The appearance of oculated idols coincides with a period of significant transformation in the Iberian Peninsula, which falls within the Late Neolithic and Chalcolithic periods, dating approximately 4500–2500

BCE (Hurtado, 2008). This era saw communities transitioning from a predominantly nomadic lifestyle to more settled, agrarian societies. Moreover, the emergence of metallurgy (Montero-Ruiz and Murillo, 2016), plough agriculture, pastoralism, and secondary products revolutionized daily life and production (Sherratt, 1981). Permanent settlements with more complex architecture replaced scattered farming hamlets (Diaz del Río, 2011). In death, collective burial in megalithic tombs and caves emerged, in contrast to individual burials in pits or cists before (Scarre, 2002). This shift brought about changes in social structures, trade practices, and technological advancements. The idols, therefore, emerge not only as spiritual symbols but also as markers of a society in flux, grappling with the challenges and opportunities of a new age. The widespread distribution of these idols across various archaeological sites suggests a shared cultural or religious practice that transcended local boundaries. This observation raises intriguing questions about the extent of cultural interactions and trade networks during this period. The idols' stylistic variations and motifs could potentially map these interactions, revealing patterns of influence and exchange among different communities.

Oculated idols occur across regions of Iberia during this era, from Los Millares in the Southeast (Arribas et al., 1983) to the Tajo River estuary in the West (Boaventura, 2011). Their geographic distribution mirrors that of megalithic architecture and specialized pottery styles (Harrison, 1980). The increasing presence of exotic raw materials like ivory and sea shells indicates the growth of long-distance trade networks (Murillo-Barroso et al., 2018). Some settlements with metal production and fortification like Los Millares and Zambujal reflect local hierarchies and competition (Chapman et al., 1990). The idol carvings thus arise within a transformational era involving new economic bases, architectures, burial practices, and trade connections. Regional styles suggest they expressed group identities and their deposition mainly as grave goods implies ritual functions. As symbolic artefacts, oculated idols provide insight into how Late Neolithic/Chalcolithic societies negotiated these changes through shared ideologies and belief systems (Bradley, 1999).

In summary, the oculated idols of the Iberian Peninsula are not mere relics of a bygone era; they are potent symbols that encapsulate the complexities and nuances of Copper Age societies. Their study provides a deeper understanding of the spiritual inclinations, artistic endeavours, and cultural interactions that shaped the early history of the Iberian Peninsula, offering a fascinating glimpse into a world that was as complex as it was enigmatic.

2.1. *Distinctive Styles: Exploring the Pastora and Almizaraque types*

In this section, we focus on two distinct regional styles of ocular idols on long bones that stand out in Southern Iberia: the Pastora and Almizaraque types – the Pastora type concentrated in eastern Spain and the Almizaraque type in the southeast. These defined types (Ballester, 1946; Bernabeu, 1981; Siret, 1908; Soler-Diaz, 1985) are not only representative of the artistic traditions of their respective regions but also provide a unique window into the cultural and spiritual practices of Copper Age societies. We will now explore the main traits, archaeological contexts, and cultural significance of these two types of idols, offering a better understanding of their role in Iberian prehistory.

The Pastora type (see Fig. 1) is named after the site of Cova de la Pastora near Alcoi, where the largest assemblage has been found (Ballester, 1946). Pastora, as a natural cave burial site, housed between 50 and 70 individuals, buried at various times and accompanied by a rich

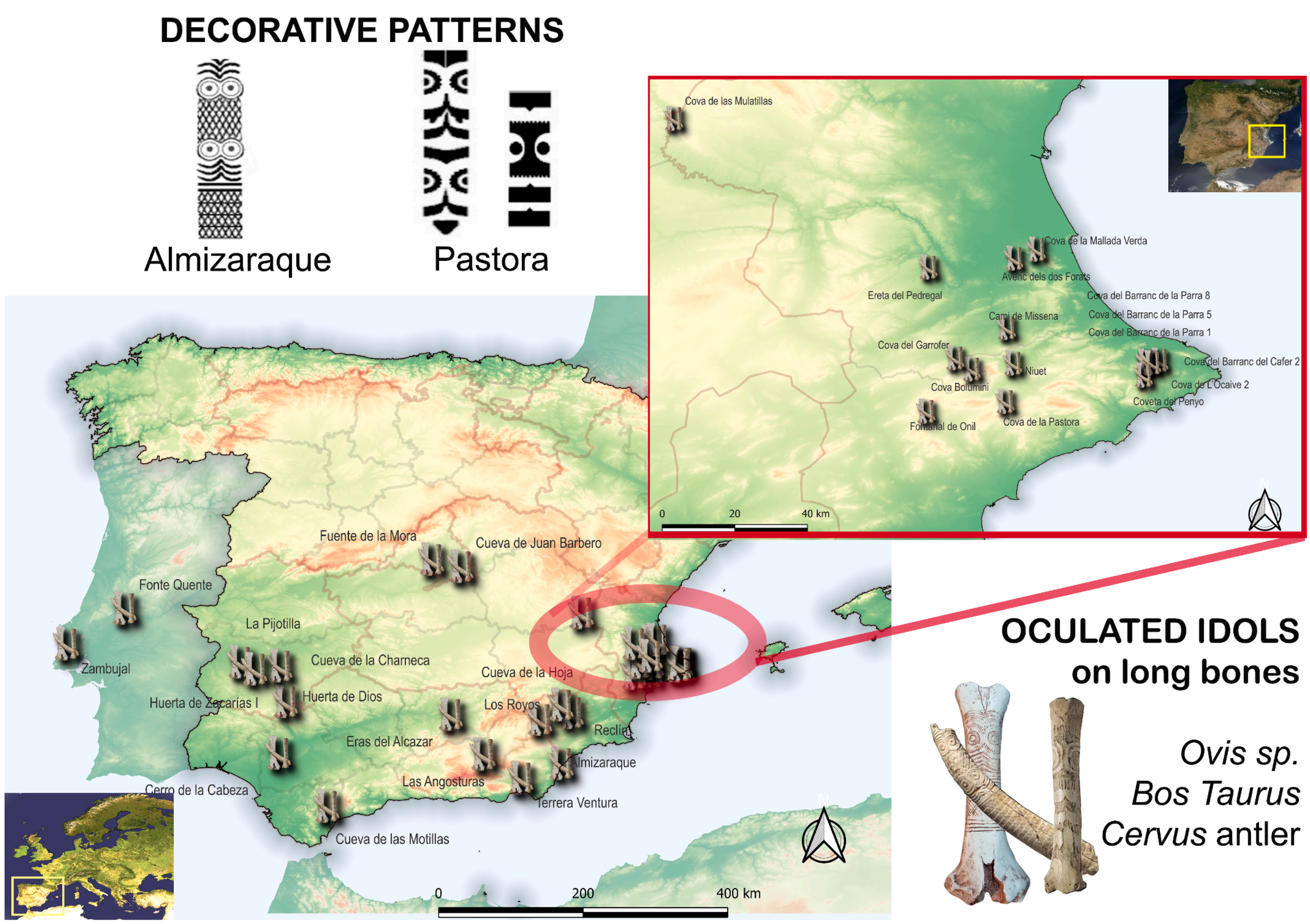

Fig. 1. Distribution of the Oculated Idols on long bones.

array of grave goods, reflecting the rich cultural and ritualistic tapestry of these early communities. Primarily found in Late Neolithic/Chalcolithic contexts, Pastora-type idols are mostly associated with funerary practices, forming part of grave goods in collective burials within natural caves and, to a lesser extent, in domestic contexts. The idols, including both complete and incomplete examples, show a decorative uniformity that aligns with oculated representations in pottery from Almería and other artefacts from the Southwest. They were often found alongside bundles of human bones, sometimes including multiple skulls, suggesting a significant role in burial rites and possibly in second burials. These artifacts are carved on ovicaprid radii and some deer and cattle bones (Pascual-Benito, 2010). The bone diaphysis is covered in painted or engraved decorative bands dividing the idol into zones representing stylized facial features. Arched "eyebrows" appear at the top, wide circular "eyes" in the middle, and sets of 2–4 parallel curved "tattooing" bands on the bottom (Bernabeu, 1981, 1979). Some also have additional geometric motifs below (Pascual-Benito, 1998). The eyes may be connected by a vertical bar and some have two stacked sets of eyes (Ballester, 1949). Dated to the late 4th – early 3rd millennium BC, Pastora-type idols occur in collective burials in caves and rock shelters in eastern Spain (Soler-Diaz, 2002).

The Almizaraque type (see Fig. 1) is named for the large assemblage recovered from the settlement of Almizaraque in Almeria, southeast Spain (Siret, 1908). The settlement of Almizaraque, located in Almería in southeastern Spain, was a major site of the Los Millares culture during the late 4th and early 3rd millennium BC. Excavations by the Siret brothers in the late 19th century revealed a large settlement composed of densely packed circular stone structures adjacent to a fortified citadel. The site was a centre of copper metallurgy, with evidence of smelting slag and metal objects. Among the abundant finds were over 60 engraved stone plaques and oculated bone idols deposited in household contexts. The site's material culture shows connections to both the Southeast and Western regions of Iberia. Almizaraque was a key hub for the exchange of copper, ivory, and other exotica during its peak occupation. Its size and complexity reflect the emergence of social stratification and far-reaching trade networks at this time in southeastern Iberia. Almizaraque-type idols are carved on various long bone types from domestic animals and other species (Maicas-Ramos, 2007). The bone surface is densely covered in precisely engraved banding that delineates stylized facial features like the Pastora type but with more elaborate motifs. This includes eyes with multiple concentric circles and radial lines, and many parallel lines for the eyebrows and tattooing bands. Found in domestic and funerary sites of the Los Millares culture like Los Millares and Las Angosturas, they date contemporaneously with the Pastora type (Arribas et al., 1983; Escoriza, 1991; Nocete et al., 2010) [8,9]. The Almizaraque style is more formal in its engraved motifs and composition compared to Pastora idols.

2.2. Geographic framework

The Oculate idols on long bones are concentrated in the central and southern regions of the Iberian Peninsula, especially in the basins of the Júcar, Segura, and Guadiana rivers, provide a window into the social structures, territorial boundaries, and cultural practices of early farming communities in Southern Iberia (Ballester, 1949, 1946; Hurtado, 2008; Pascual-Benito, 2010; Soler-Diaz, 2017).

The earliest examples of these idols come from the southeastern region dating back to the late 4th and early 3rd millennium BC (Arribas et al., 1983; Nocete et al., 2010). Made from various materials like bone, ivory, and stone plaques (Gonçalves et al., 2004; Lillios, 2008), they are often found in funerary contexts in natural caves and megalithic tombs, as well as in domestic contexts associated with habitation structures (Hurtado, 2008). In the Guadiana basin, for instance, ocular idols made of ovicaprid radii appear in the early 3rd millennium BC at sites like La Pijotilla and Huerta de Dios (Enríquez, 1983; Hurtado, 1980), often deposited collectively near tomb entrances, sometimes alongside skull bundles.

In eastern Spain, a significant concentration of these idols is found in the upper Serpis river basin and nearby areas, with the largest assemblage from Cova de la Pastora near Alcoi (Ballester, 1949, 1946; Soler-Diaz, 2017), dating to the late 4th/early 3rd millennium BC. More is known from sites like Ereta del Pedregal (Fletcher et al., 1961; Juan-Cabanilles, 2006), Cova Bolumini (Pascual, 1957), Cova del Garrofer (Bernabeu, 1981), El Fontanal (Soler-Diaz, 1985), Avenc dels Dos Forats (García-Puchol et al., 2010), Coveta del Penyó and Cova Negra de Xalò (Soler-Diaz, 2002), Barranc del Cáfer and Barranc de la Parra (Soler-Diaz, 2017), Niuet (Bernabeu et al., 1994) and Camí de Missena (Bernabeu Aubán et al., 2017; Pascual-Beneyto et al., 2005). Slightly to the north, sites such as Mallà Verda and Cueva de las Mulatillas (Pedraz Penalva and Molina Burguera, 2000) are also found, but none is located further north than the Júcar River. Also, idols are present in the Segura river basin, which seem more related to the Almizaraque style (see Fig. 1), in sites such as Los Royos, Cueva de la Hoja and Sierra de la Puerta (Ayala-Juan, 1985). The distribution continues to the southeastern Mediterranean litoral with the workshop site of Almizaraque (Almagro-Basch, 1965) (with more than 60 idols of Almizaraque style and 3 of Pastora style), and in the interior with the sites of Terrera Ventura (Gusi and Olaria, 1990; Von den Driesch and Morales, 1977), Angosturas de Gor (Mateu, 1990) and Eras del Alcázar (Nocete et al., 2010). Others can be found in areas far from the southeastern core, in distant locations such as near the Tajo River in the north, including sites like Juan Barbero (Martínez-Navarrete et al., 1984) and Fuente de la Mora (Vigil-Escalera and Martín, 2003), and in Badajoz, at places like Huerta de Zacarías and Cueva de la Charneca (Enríquez-Navascués et al., 2013).

The idols, carved on ovicaprid radii and some deer antlers and cattle bones, are found in various contexts, mostly funerary, where they form part of the grave goods in collective burials. Distinctive regional styles of ocular idols on stone cylinders are found farther west, with simple forms in Extremadura dating back to the early 3rd millennium BC (Escoriza, 1991; Hurtado, 2008). In regions like the Algarve, more elaborate curved "eyebrow" designs are prevalent (Gonçalves et al., 2004), while in the South, cylindrical types with incised eye designs are common (Molina et al., 2004). The Middle Guadiana basin is known for its unique rectangular stone plaques with oculated designs, appearing in the late 3rd millennium BC (Lillios, 2008).

In conclusion, the study of ocular idols on long bones across Southern Iberia, as explored through various archaeological findings and scholarly analyses, reveals a complex and rich tapestry of cultural and spiritual life in the early farming communities. Overall, the styles of these ocular idols seem to cluster regionally, suggesting they conveyed symbolic identities linked to specific territories.

2.3. Chronology of oculated idols on long bones

In the Southeast, oculated bone idols first appear in the Los Millares culture of Almería by the late 4th millennium BCE. At Los Millares itself, three contexts with idols or other symbolic artefacts fall between c. 3080–2660 BCE (Molina et al., 2004). Contemporary levels at Terrera Ventura also contained examples (Gusi and Olaria, 1990). In the Western Valencian region, Cova de la Pastora has the largest assemblage of bone idols, with dates ranging from 3330 to 2980 BCE (García-Puchol et al., 2013) when calibrated.

Toward the interior, engraved limestone idols occur at Perdigões within the first half of the 3rd millennium BCE (Valera, 2013; Valera and Godinho, 2010). La Pijotilla in the Guadiana basin, with over a dozen bone idols from the tholos Tomb 3, has a calibrated date range of 2860–2470 BCE. The Portuguese sites of Pomar and Palmeira also yielded late 4th-early 3rd millennium BCE dates associated with idols (Boaventura, 2009). In contrast, the stone slab idols of the Alentejo heartland appeared earlier, since the megalithic architecture associated with them initiated around the mid-4th millennium (Waterman et al., 2014). However, the slabs continued in use through the 3rd millennium BCE in secondary collective burials (Valera and Godinho, 2010).

Overall, the radiocarbon evidence suggests oculated idol carving emerged in southeastern Iberia by the late 4th millennium BCE before spreading inland to central and western regions during the 3rd millennium. However, the early carved stone plaques of the Tagus and Sado valleys indicate related symbolic imagery arose independently earlier in those areas. Regional divergence in raw materials and stylistic execution subsequently developed as oculated idols were adopted by diverse Chalcolithic groups across southern and western Iberia.

Nevertheless, it must be pointed out that most of the previously described radiocarbon determinations proceed from contexts that are not related to oculated idols or come from long-lived samples such as charcoal. Direct radiocarbon dating on short-lived organic samples provides the most reliable method for establishing the chronology. Fortunately, some contexts with these artefacts have yielded suitable for 14C analysis, like charred seeds or bone collagen. When calibrated, these dates provide a likely delimitation for the overall span and regional timing of production and use for this idol type in Iberia. They can be seen in Table 1, and have been used to create a model employing the software Chronomodel (Lanos and Dufresne, 2019).

The model (see Fig. 2) proposes an interval between −2814 to −2559 BCE (Confidence Interval 95 %), with a phase lasting 255 years. It spans from the appearance of the copper metallurgy (Gauss, 2013; Nocete et al., 2011) to the initial moment of the bell-beakers phase (Harrison and Heyd, 2007; Jiménez-Puerto and Bernabeu-Aubán, 2023). This places them firmly within the Chalcolithic period, overlapping with early copper metallurgy and pre-dating the bell-beaker culture. Moreover, regional variations in idol styles and materials provide evidence for the cultural diversity of this important symbolic tradition.

The spread of oculated idol imagery and use occurred against a backdrop of major social and economic changes in southern Iberia during the 3rd millennium BCE. The Chalcolithic period saw increasing social complexity, with some scholars arguing for the emergence of a proto-state centered around large settlements like Los Millares and Valencina de la Concepción (Chapman, 2008; García-Sanjuán and Murillo-Barroso, 2013). These sites featured monumental architecture, population nucleation, and evidence of craft specialization and hierarchical social organization.

At the same time, metallurgy, while still sporadic, became an important status marker, as copper objects like daggers and punches appeared in elite burials (Hunt Ortiz, 2003; Murillo-Barroso and Montero-Ruiz, 2012). Maritime networks were developed to facilitate metals trade between coastal regions, enabling later the spread of raw materials and ideas. Megalithic architecture for collective burials also peaked during this period (Aranda Jiménez et al., 2013; Costa Caramé et al., 2010). Not the case in the Júcar river basin, where the megaliths are absent, signalling a different funerary ritual for that area, which is associated with caves (Soler-Diaz, 2002). In that area is where the Pastora-type idols seem to have their origin, more specifically in the Serpis river basin. The study of paleodemography for that area, through

Table 1
Radiocarbon determinations, on short-lived samples, from contexts related to oculated idols on long bones.

Site	Level	Sample ID	Date	SD	Bibliography
Camí de Missena	Est 51	Beta331021	4030	30	Bernabeu, Orozco & Pascual 2017
Eras del Alcázar	F18	UA26019	4070	45	Nocete et al., 2010
Avenc dels dos Forats	UE1018	UCIAMS6631	4115	20	García-Puchol et al., 2010
Ereta del Pedregal	II-K5	Beta327998	4150	30	Pérez-Jordá 2013
Eras del Alcázar	F19A	Beta229716	4170	50	Nocete et al., 2010

the study of radiocarbon determinations as a proxy (see Fig. 3), indicates a big growth for that moment. It seems to be synchronous with the appearance of metallurgy and lasts until the introduction of bell-beaker pottery. It is within this transforming economic and social landscape that the tradition of oculated bone idols emerged and flourished. Their purpose remains unclear, but the idols perhaps served as emblems of identity, status, and belief for emergent elites and specialized craftspeople in the diverse prehistoric communities of Chalcolithic southern Iberia.

3. Material and methods

This section details a sophisticated approach to analyzing prehistoric artefacts using modern CV techniques. It focuses on a unique sample of JPEG images of decorated oculated idols, carved on long bones and sourced from archaeological sites across the Iberian Peninsula. This collection includes diverse regional variations of idol motifs. The images underwent a rigorous process of digitization, standardization, and enhancement to maintain quality and uniformity. This preparation was crucial for the subsequent application of advanced ML algorithms. The section outlines the technical steps undertaken to process the images. The comprehensive methodology employs three established algorithms for extracting detailed features from the idol images. These features, encapsulating the shapes, textures, and motifs of the idols, form the basis for a nuanced analysis using unsupervised ML, which will be further explored in the Methodology section.

3.1. The sample

The sample analyzed in this study consisted of 100 high-resolution JPG image files showing decorated oculated idols carved on long bones from 23 archaeological sites across Iberia (see Fig. 1). This comprehensive sample captures the diverse regional variations in carved idol motifs that emerged throughout southern and western Iberia during the Chalcolithic period from approximately 3000–2500 BCE. The data collection phase for this study involved a meticulous process of sourcing and acquiring images of oculated idols from the Copper Age in the Iberian Peninsula. These images were primarily gathered from archaeological repositories, museums, archaeological excavation reports and publications (Pascual-Benito, 2010, 1998; Soler-Diaz, 2017, 2002, 1985). High-resolution photographs of idols were digitized, ensuring high quality and uniformity for analysis.

The sites span coastal and inland locations in the southern half of Iberia. The extensive geographic distribution and number of samples provide a robust basis to identify both shared visual features and localized stylistic differences among the regional bone idol carving traditions using CV techniques. In addition, 21 images have been discarded due to the fragmentation of their decorations which can't be followed properly, leaving a final sample of 79 oculated idol decorations. Upon collection, the images underwent a process of preparation and standardization to ensure consistency and accuracy in the ML analysis.

A key preliminary step in our computational analysis pipeline was to standardize and preprocess the idol image dataset. First, all images were converted to the same resolution and format to enable consistent feature extraction across the sample. We chose a standard image resolution of 300 dpi and used a lossless PNG file format for maximum preservation of detail. Resolution and format standardization are crucial for removing potential biases when applying ML algorithms.

Next, various image enhancement techniques were applied to preprocess the idol images. Contrast enhancement and adaptive histogram equalization were used to improve clarity and amplify the visibility of carved patterns. Bilateral filtering helped reduce noise while preserving edges around the intricate decorations. Finally, unsharp masking sharpened the images to bring out fine details in the carvings. Such preprocessing and enhancement steps are essential to highlight the

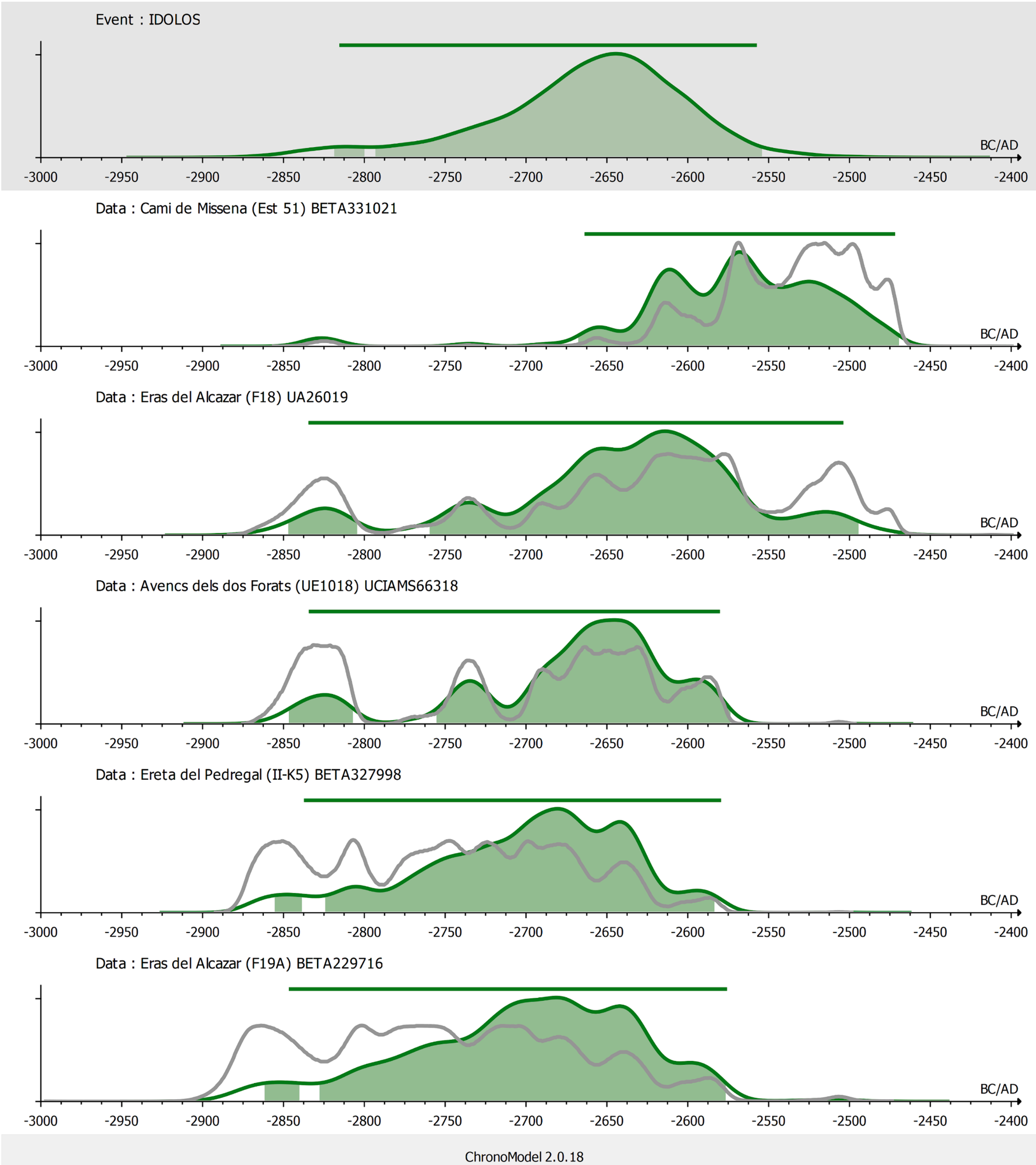

Fig. 2. Chronological model for oculated idols on long bones from the Iberian Peninsula.

subtle decorative elements on the idols, which contain key information for the feature extraction phase.

Along with this technical preprocessing, the images were manually annotated with relevant metadata. Each idol image was catalogued with known provenance information, estimated dates based on contextual archaeological evidence, and any available historical or cultural background. Annotating these details provides valuable context about each artefact to assist in interpreting results across different idol groups during the analysis process. Combined with the image standardization and enhancement, this comprehensive preprocessing battery generated an optimal dataset to feed into our feature extraction and ML workflow for deciphering visual relationships between Chalcolithic Iberian idols.

Four well-established algorithms – LBP, HOG, SIFT and CED – were employed to extract descriptive features encoding the shapes, textures, and motifs depicted on the idol specimens. These complementary feature descriptors provide detailed quantitative representations of the

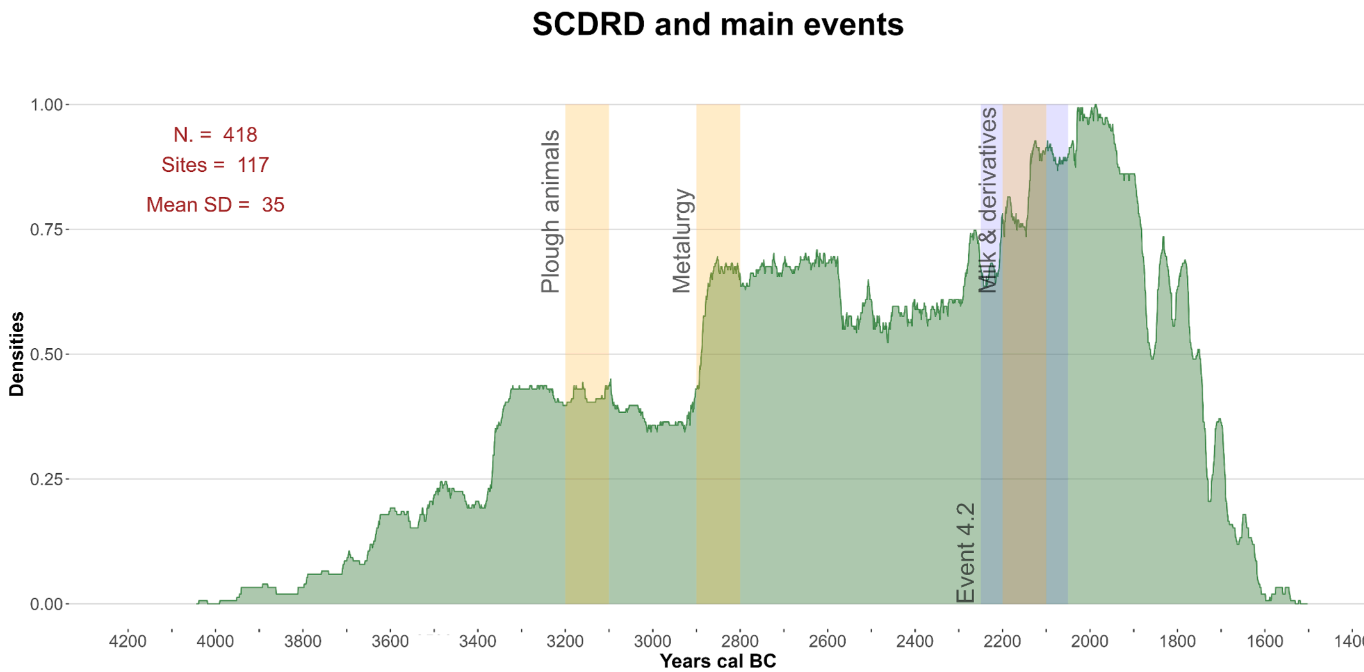

Fig. 3. SCDRD for the Mediterranean Jucar-Segura-Sur basins, employing only short-lived radiocarbon determinations (Jiménez-Puerto, 2021).

carved visual designs and will be discussed in the Methodology section.

Quality assurance and data processing for the image database

To ensure the reliability and consistency of our computational analysis, our study employs a rigorous Quality Assurance/Quality Control (QA/QC) protocol for the selection and processing of image data. The database utilized comprises 100 high-resolution JPEG images of oculated idols, which were carefully curated from 23 archaeological sites across Iberia. These images were standardized to a uniform resolution of 300 dpi and converted to PNG format to preserve detail, serving as our primary variables for running the model.

Understanding that image quality can vary and potentially influence data processing, our QA/QC measures address these discrepancies to maintain the strength of our results. Images were evaluated for clarity, resolution, and integrity of the decorative motifs; those compromised by fragmentation or poor quality—resulting in 21 images being discarded—were excluded to ensure consistency across the dataset. Our preprocessing pipeline, including contrast enhancement, adaptive histogram equalization, and bilateral filtering, further harmonizes the dataset by improving image quality and ensuring that each idol's decorative details are accurately captured for feature extraction.

The quality and uniformity of the images are pivotal for the reliability of our feature extraction and subsequent machine learning analysis. By adhering to these QA/QC protocols, we mitigate the impact of variable image quality and fortify the credibility of our findings, offering a robust foundation for identifying visual patterns and regional distinctions among the Chalcolithic period oculated idols of the Iberian Peninsula.

Data limitations and potential biases

Our study acknowledges several limitations inherent to the dataset utilized for analyzing the decorative patterns on Copper Age Iberian oculated idols. Despite efforts to collect a comprehensive sample, the finite quantity and variable quality of available high-resolution images impose constraints on our analysis. The breadth of our dataset is limited, potentially affecting the robustness of our model's findings. Additionally, there is an inherent bias due to the overrepresentation of idols from certain regions or archaeological sites, which might skew our interpretations towards these more prominently featured styles.

Such imbalances in data representation could inadvertently influence our conclusions, emphasizing the need for caution when extrapolating our findings to other datasets or broader archaeological contexts. The study's insights are primarily applicable to the subset of oculated idols examined and may not fully encapsulate the diversity and richness of Copper Age Iberian artistic expression. This limitation underscores the importance of further research to mitigate biases and enhance the generalizability of our results across different archaeological settings.

3.2. Methodology

The utilization of sophisticated computational techniques for the analysis of archaeological artefacts marks a significant shift in research methodologies within the field of archaeology. This shift is characterized by the innovative adaptation of CV and ML technologies, which have opened new avenues for extracting insights from archaeological datasets (van der Maaten et al., 2007). These technologies facilitate a novel approach towards the classification, comparison, and pattern recognition in collections of artefacts, particularly in those with intricate visual features such as pottery. This section explores a methodological framework that leverages CV algorithms alongside unsupervised ML clustering techniques to analyze archaeological artefact images. This framework enables the quantitative categorization of artefacts through the analysis of visual feature similarities, thereby fostering comparisons across various objects, styles, and prehistoric periods. Furthermore, it places this methodology within the larger context of computational archaeology, evaluating its benefits, constraints, and future research directions. The objective is to highlight the interdisciplinary application of computer science and engineering methods to archaeological questions, while critically assessing their role within the broader archaeological research toolkit.

Traditionally, the analysis of material culture's visual patterns has been a cornerstone of archaeological practice, albeit relying heavily on qualitative methods (Gilboa et al., 2013). The emergence of digital archaeology has sparked a heightened interest in adopting more objective and quantitative approaches for artefact comparison and classification based on visual attributes. Initial methods focused on simple geometric morphometrics to analyze artefact shapes. However, newer strategies aim to capture more complex feature representations, leveraging CV and ML principles to classify artefacts by visual similarities without the need for pre-defined categories (Gilboa et al., 2004).

Technological breakthroughs such as the SIFT (Lowe, 2004) have revolutionized CV tasks by introducing robust local feature descriptors capable of withstanding variations in rotation, scale, and lighting. These advancements have shown considerable promise in archaeology for artefact classification by quantifying unique visual markers, much like the human ability to recognize stylistic similarities. Unsupervised ML techniques, particularly clustering algorithms, further enhance this capability by grouping data points based on feature similarity, thus offering a data-driven method for artefact classification.

This chapter outlines a comprehensive computational workflow (see Fig. 4) designed for clustering archaeological image data, utilizing advanced local features and various clustering techniques. The process encompasses image loading and preprocessing, feature extraction using SIFT, HOG, LBP, CED, and clustering through algorithms such as GMM, Spectral Clustering, and SOM. This approach is not only quantitatively rigorous but also provides intuitive visual insights into the dataset structure and clustering outcomes. By elaborating on the technical specifics and archaeological relevance of these steps, implemented via Python (Van Rossum and Drake, 2014) for efficient data processing, this discussion underscores the potential of CV algorithms in identifying distinct archaeological patterns under varied conditions.

Computer vision algorithms

SIFT is a milestone CV algorithm that detects and describes local features in images (Lowe, 2004). It identifies key points likely to be distinct despite viewing variations, making it powerful for archaeological images captured under different conditions. A keypoint descriptor is a 128-D vector summarizing gradient patterns within its local neighbourhood. This descriptor proves to be robust to illumination, rotation, scale, and viewing angle changes.

In our implementation, we augment SIFT by incorporating HOG (Dalal and Triggs, 2005) and LBP (Ojala et al., 1996) for enhanced feature extraction. HOG analyzes the distribution of edge directions within an image region, providing additional texture information to complement the gradient data from SIFT. It captures repetitive local patterns and shapes characteristic of many artefacts. LBP encodes pixel-level image texture into bit string patterns based on comparisons with neighbouring pixels. It is computationally efficient and adds rotation-invariant texture description capacity. The complexity descriptors for images are generated using the CED algorithm (Canny, 1986). This method identifies the edges within each image by detecting areas of rapid intensity change. It operates by applying a set of filters to the image to identify intensity gradients, then traces along these gradients to enhance the edges. The complexity descriptor for an image is quantified by calculating the ratio of the sum of edge pixels to the total number of pixels in the image. This ratio represents the image's complexity, with a higher value indicating a more complex image due to the presence of more edges. This approach provides a numerical measure of image complexity, which can be used for further analysis or as part of feature sets in image processing tasks. By analyzing the distribution and quantity of complex visual elements within a design, one can assess its balance and harmony, offering a metric to distinguish simpler patterns from more elaborated ones.

By combining SIFT, HOG, CED and LBP, our feature extraction stage generates diverse descriptors incorporating gradient, edge, and texture data. This expanded characterization of local visual properties facilitates more robust matching and classification of archaeological artefacts exhibiting complex designs, carvings, complexity or surface patterns. The distribution of features, rather than just the shape, contains information about an artefact's style and provenance. Matching descriptors across objects provides a quantitative indicator of similarity, enabling

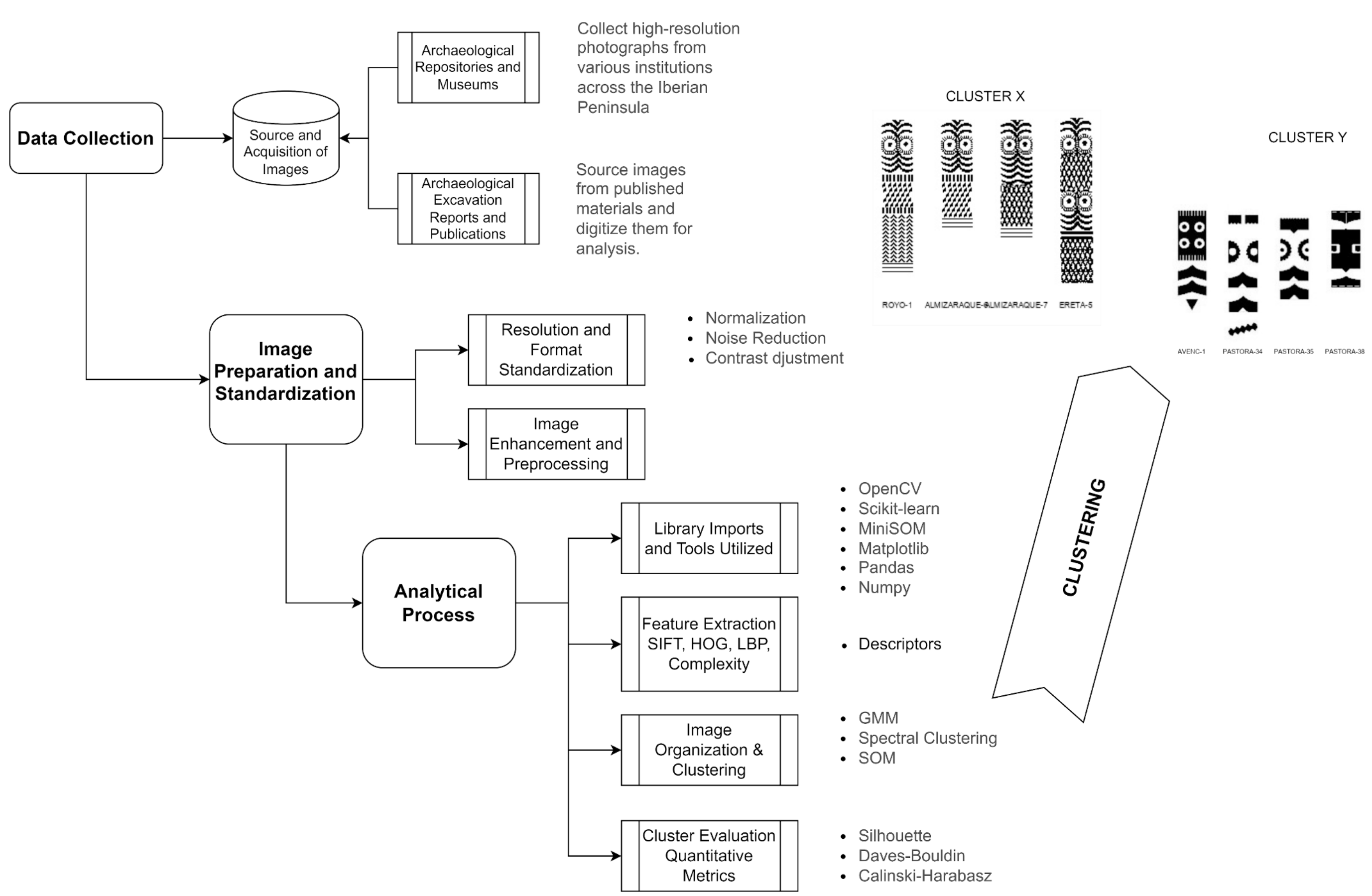

Fig. 4. Workflow of the ImageClusterVis toolkit.

more objective classification. Our enhanced SIFT-based descriptors are well-suited as input to clustering, supporting the discovery of stylistic groups based on local visual patterns.

Clustering algorithms

We employ three standard clustering algorithms – Gaussian Mixture Models (GMM), Spectral Clustering, and Self-Organizing Maps (SOM) – to explore different ways of grouping images based on their SIFT descriptors. GMM is a probabilistic model that assumes data is generated from a mixture of Gaussian distributions (Reynolds, 2009). Each image descriptor is assigned to the component it is most likely associated with, partitioning the data. GMM excels at finding clusters of varying shapes and sizes, useful for complex archaeological data. In Spectral Clustering, eigenvectors of similarity matrices between data points are analyzed to segment the dataset (Von Luxburg, 2007). This algorithm focuses on connectivity rather than distance. It can successfully detect non-convex and elongated clusters, common in stylistic artefacts. SOM projects high-dimensional data onto a two-dimensional organized map, preserving topological relationships between input data points (Kohonen, 1982). On the map, descriptors are grouped in a visually intuitive way, facilitating the discovery of patterns within the archaeological data. By testing multiple clustering algorithms, we can overcome individual limitations and gain a richer perspective on stylistic relationships between artefacts. The algorithms require some parameter tuning but generalize well without extensive optimization.

Various visualization techniques are employed to interpret the clustering results, each tailored to the nuances of different clustering algorithms. For Self-Organizing Maps (SOM), the script utilizes t-SNE (t-distributed Stochastic Neighbor Embedding) (Cieslak et al., 2020) to transform the high-dimensional feature space into a more comprehensible two-dimensional plot. This visualization technique is instrumental in demonstrating how SOM clusters images within this reduced space, thereby simplifying the identification of clusters and interrelations that might be less obvious in a higher-dimensional context. Similarly, the script applies t-SNE for visualizing the outcomes of Spectral Clustering, with each point representing an image. The plot often uses varied colours to demarcate different clusters, offering a clear visual distinction among the various groups.

Quantitative evaluation

Clustering quality was assessed using standard metrics – Silhouette Coefficient (Rousseeuw, 1987), Davies-Bouldin Index (Davies and Bouldin, 1979), and Calinski-Harabasz Index (Caliński and Harabasz, 1974). The Silhouette Coefficient measures how well samples fit within their assigned cluster versus neighbouring clusters. Davies-Bouldin quantifies intra-cluster cohesion versus inter-cluster separation. Calinski-Harabasz evaluates dispersion between and within clusters. Together, these metrics provide numerical measures of clustering compactness, isolation, and stability from different geometrical perspectives. They allow systematic benchmarking to determine which clustering scheme produces superior results for a particular archaeological application. The script inverts the Davies-Bouldin metric, due to the fact in this particular case lower is better. Finally, it normalizes metrics for scale uniformity and calculates a weighted average score for each algorithm. The outcome is a set of bar charts showing algorithm performance and the script prints the name of the best algorithm. However, it must be noted that numerical metrics alone do not guarantee optimal clustering or archaeological relevance, underscoring the need for qualitative evaluation by experts.

The ImageClusterVis Script: A Digital compass

The methodology indicates how image analysis and clustering can assist archaeological inquiry and interpretation. SIFT, HOG and LBP capture visual elements important for distinguishing between artefacts and styles, while clustering provides an unbiased way to quantitatively assess similarities based on these features. The algorithms can handle variations in image scale, rotation, and quality, making them well-suited for classifying real-world archaeological data.

However, there are some limitations to consider. Interpreting the output clusters requires expertise, and clustering algorithms may detect distinctions not meaningful to archaeology. Moreover, Image quality issues can impede feature extraction. The algorithms involve parameters and design choices influencing outcomes. Ongoing critical assessment of algorithmic methods and collaboration with archaeologic information is essential.

Nonetheless, this approach holds potential for numerous archaeological applications. It could support typology development, provenance studies, investigation of cultural transmission patterns, analysis of artistic themes, and other quantitative visual style research. Interesting directions include combining image data with other archaeological evidence, developing specialized feature extractors and cluster evaluation metrics for artefacts, and creating interactive in Methodological Limitations and Model Selection Challenges.

Methodological Limitations and Model Selection Challenges.

In developing our methodological framework, we encountered several challenges that required specific adjustments to our models. One significant hurdle was the selection and tuning of computer vision algorithms and unsupervised machine learning techniques best suited for analyzing the diverse and complex visual patterns of archaeological artefacts. Due to the heterogeneous nature of archaeological datasets, not all algorithms performed uniformly across different types of artefacts. For example, the SIFT algorithm, while robust in identifying key points in high-contrast images, required adjustment when applied to artefacts with more subtle or eroded decorative patterns. Similarly, the choice and tuning of clustering algorithms like GMM and Spectral Clustering had to be carefully balanced to optimize for both accuracy and computational efficiency.

The methodological limitations of our approach primarily revolve around the specificity of the artefact dataset under study. The effectiveness of the chosen computer vision and machine learning techniques in extracting and analyzing visual features may not be directly transferable to artefacts with significantly different visual attributes or to datasets with a broader variety of artefact types. Furthermore, the generalization of our findings to other archaeological datasets or artefact types may be limited by these methodological constraints. While our models and algorithms have demonstrated potential in uncovering previously unnoticed patterns and classifications within the dataset of oculated idols, their applicability to other datasets or broader archaeological inquiries requires further validation. This emphasizes the importance of ongoing critical assessment and adaptation of our computational methods to ensure their relevance and effectiveness in diverse archaeological contexts.terfaces for exploring clustering results.

4. Results and discussion

Embarking on a journey through the analytical landscape shaped by ImageClusterVis, this chapter transcends the mere technicalities of the script to dive into the interpretative results and insightful observations. By transitioning from the mechanics of image processing and clustering algorithms, we weave a narrative that integrates these computational processes with the thematic richness of our dataset. This dataset, a digital tapestry of images, serves not only as a collection of pixels but as a storyteller of history, culture, and artistic expression. Through clustering, we uncover subtle dialogues between these images, revealing connections and distinctions that might otherwise elude the naked eye.

4.1. Unveiling clustering narratives

The application of GMM, Spectral Clustering, and SOM through ImageClusterVis has not only grouped similar images but has also unveiled clusters of narratives and shared histories. Each algorithm, with its unique approach, peels away different layers of these stories, offering varied perspectives on the relationships between images.

In our clustering analysis involving spectral clustering, GMM, and SOM, we elected to use four clusters. Our decision was informed by the

Elbow Method and the assessment of cluster stability assignments across different runs of the algorithm with varying initial conditions. This method involves plotting the within-cluster sum of squares (WCSS) against the number of clusters and identifying the 'elbow point' (Han et al., 2022). The elbow point is where the rate of decrease in WCSS slows down significantly, indicating a natural grouping in the data.

When it comes to the GMM, the script presents two distinct types of visualizations. The first, known as Component Visualization, leverages Principal Component Analysis (PCA) to reduce the data to a two-dimensional representation. Overlaid on this are the GMM components, elucidating how the Gaussian clusters are spread and encompass the individual data points. The second type, the Cluster Scatter Plot, also based on PCA, displays the clustered images as distinct scatter points. Each cluster in this visualization is usually represented by a unique colour, aiding in the assessment of how well the clusters are separated and how cohesive they are within the GMM framework. Each visualization method has a specific role in the interpretation of the clustering results (see Fig. 5). The SOM visualizations, for instance, are particularly valuable for grasping the topological relationships among data points. In contrast, the Spectral Clustering and GMM visualizations shed light on the grouping dynamics, focusing on similarity and the distribution of Gaussian components. Collectively, these visual tools provide a holistic view of the clustering outcomes, enhancing the understanding of the intricate patterns embedded within the dataset.

4.2. *Beyond Metrics: Interpreting results*

In the evaluation of clustering algorithms using ClusterEvalVis script, the analysis revealed intriguing insights into the performance of GMM, Spectral Clustering, and SOM. The script provides a visualization of these metrics through bar charts (see Fig. 6), offering a clear comparative perspective on each algorithm's effectiveness. Upon analyzing the Silhouette Coefficient, Davies-Bouldin Index, and Calinski-Harabasz Index, it was observed that each algorithm had its strengths and weaknesses. For instance, Spectral Clustering has excelled in maximizing the Silhouette Coefficient and Calinski-Harabasz Index, indicating well-separated and distinct clusters. In contrast, SOM has shown superior performance in minimizing the Davies-Bouldin Index and performs very well in Calinski-Harabasz, suggesting tighter, more cohesive clusters. Although by a narrow margin, the best clustering algorithm for this data

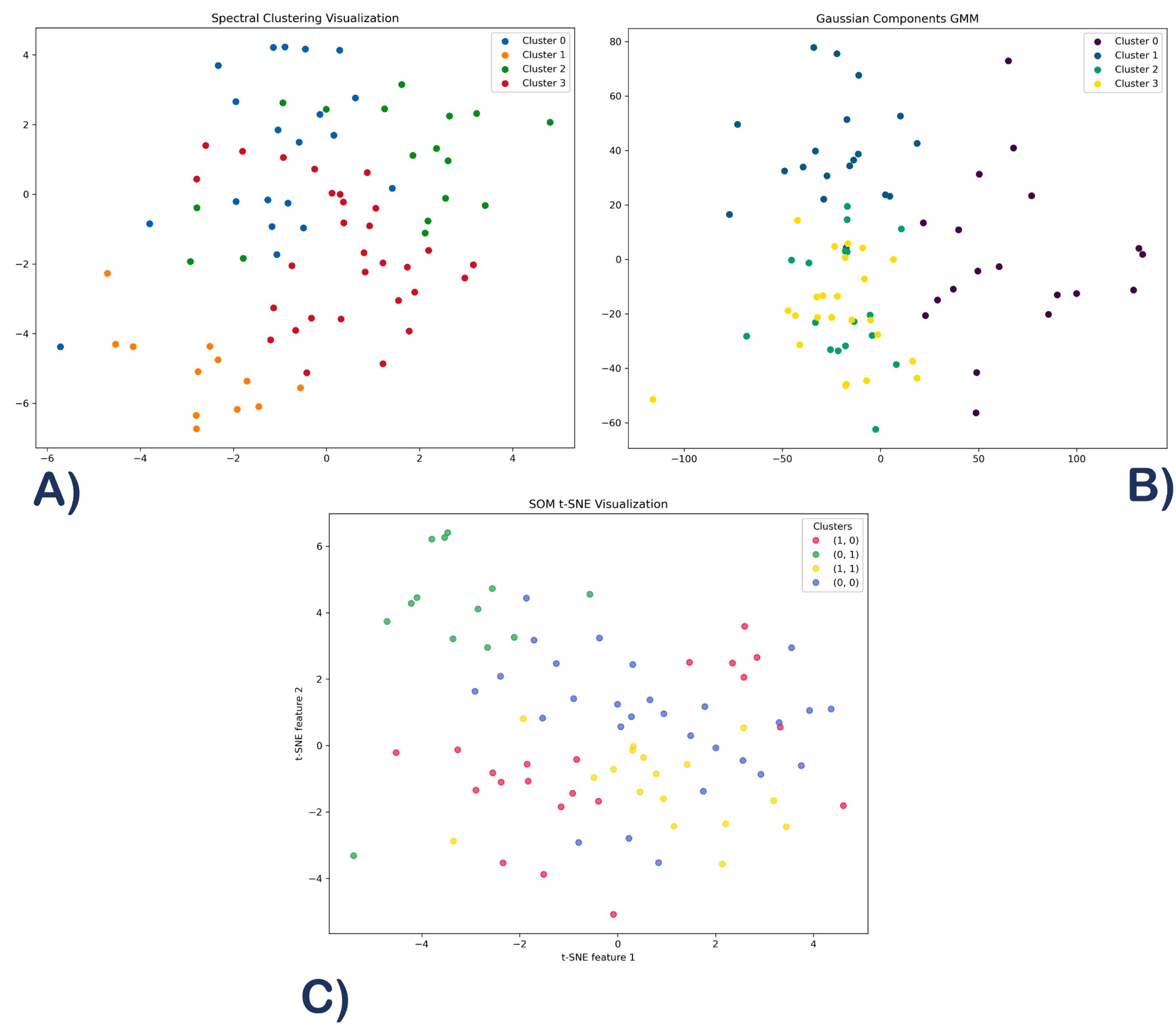

Fig. 5. Clustering plots: (A) Spectral Clustering, (B) Gaussian Components GMM, (C) SOM t-SNE visualization.

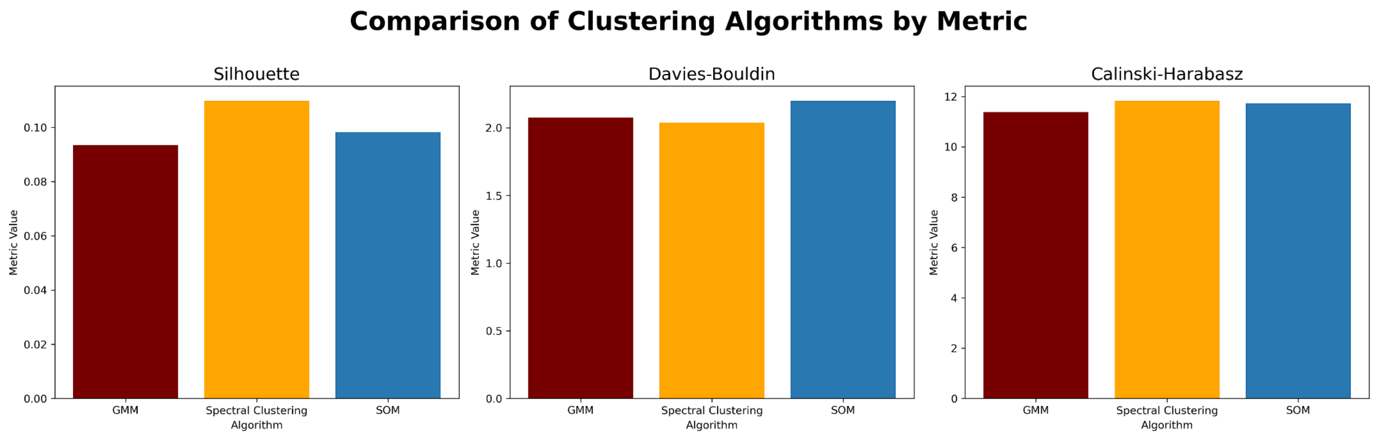

Fig. 6. Comparison of clustering Algorithms by metric.

set turns out to be Spectral Clustering. The script's ability to invert and normalize these metrics for a balanced comparison was crucial in identifying the overall best performer. Calculating a weighted average score for each algorithm, provided a holistic view of their performance across different metrics. The bar charts, with their distinct colour coding for each algorithm, visually articulated these comparisons, making it easier to grasp the relative strengths of each clustering method.

The script's computation revealed the best-performing algorithm based on the aggregated metric scores. This outcome, while highlighting the efficacy of one algorithm over the others, also underscored the importance of considering multiple metrics in evaluating clustering algorithms. The chosen 'best' algorithm, therefore, represented a balance of performance across various aspects of clustering quality, rather than excelling in just one metric. This comprehensive analysis provided valuable insights for selecting the most suitable clustering approach for specific types of datasets, demonstrating the script's utility in guiding data-driven decision-making in clustering tasks.

5. Discussion and conclusion

As this work shows, the visualizations can be considered as windows to data soul, and represent the perfect confluence of science and narrative. The final part of the workflow consisted of the display of the previously suggested clusters. The results challenge traditional interpretations of oculated idols, suggesting a complex narrative of cultural exchange rather than straightforward stylistic differences. This analysis indicates significant regional distributions and connections, underscoring the role of geographical and trade networks in cultural

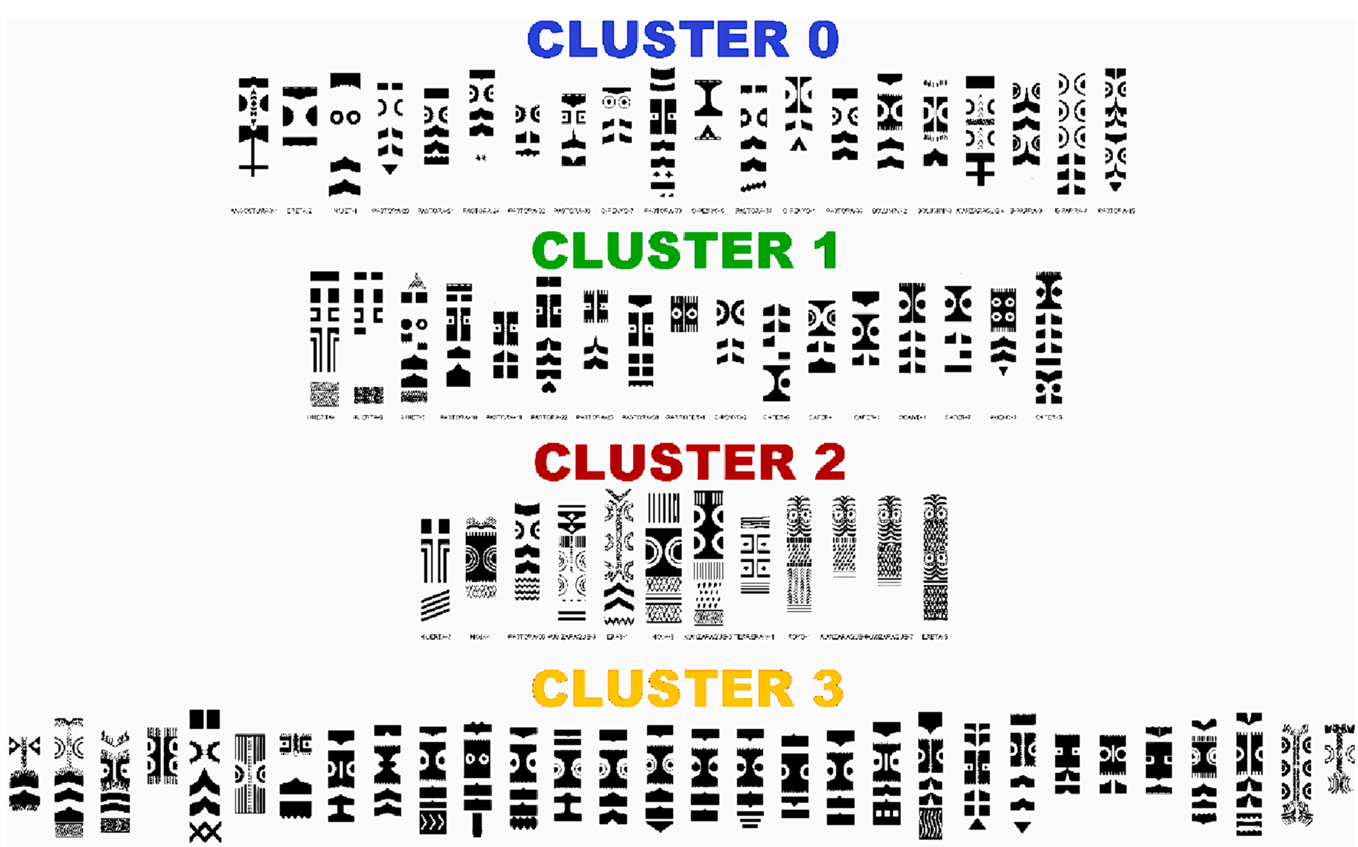

Fig. 7. Spectral Clustering results including the associated images.

transmission. As can be seen in Fig. 7, the clustering results are quite interesting.

In the first place, the chronology resolution doesn't allow to study the possibility of chronological differences between the found clusters, an issue that should be addressed in future works. In addition, the clusters do not seem to strongly support the idea that there are significant differences between oculated idols with multiple pairs of eyes versus those with simple eyes. Instead, the detected groupings indicate three main areas of distribution, with the Serpis river basin as a possible original focus from which this cultural manifestation would have spread through two axes: the Júcar-Tajo basin (which would be the northern frontier of this cultural phenomenon) and the Segura-Guadiana basin. It also shows a descending appearance gradient east–west.

Moreover, the clustering was capable of grouping all the manifestations of the Almizaraque type which seems to be restricted to the southeastern part of the peninsula (Maicas-Ramos, 2007). Furthermore, cluster 0 highlights the significant connections between the Serpis and the Vera basins. Previous studies have revealed the existence of copper ore trade networks, originating from Almería, at the La Vital site located at the mouth of the Serpis in the early Chalcolithic period (Pérez-Jordá et al., 2011), making it plausible to confirm the intense relationships between the communities of the Serpis and the Los Millares culture (López-Padilla, 2006).

It should also be mentioned the anomaly found in cluster 1 (see Fig. 8), which links the idols from La Huerta de Zacarías (Badajoz) (Enríquez-Navascués et al., 2013) with the core of the Serpis. Initially, this could be considered an outlier, although its angular motifs and the complex lower motifs are very similar to those of the idol from Niuet (Alqueria d'Asnar) (Bernabeu et al., 1994). Furthermore, based on the similarities found between the idols from the Southeast and those from Badajoz, cluster 2 also suggests the existence of a connection with the western part of the peninsula, probably through the mountains of Sierra Mágina in Jaén, where Eras del Álcazar (Úbeda) (Nocete et al., 2010) can be found. The link between the strong connections identified among the Tajo river sites and those within the Júcar and Serpis basins, and the tentative interpretation regarding the non-determining role of eye decoration in the diffusion of cultural representations, suggests a broader narrative of cultural exchange and interaction across these regions. The matching in cluster 2 of sites such as Juan Barbero (Tielmes) (Martínez-Navarrete et al., 1984), Fuente de la Mora (Leganés) (Vigil-Escalera and Martín, 2003) and Mallá Verda (Corbera) (Pascual-Benito, 2010), with the Serpis core by CV algorithms underscores the robust interconnections bridging the Mediterranean coast and the Tajo river basin, likely facilitated by the Júcar riverbed.

This connectivity could be reflective of more complex social dynamics than previously thought, where the physical geography of river systems served as vital conduits for the movement of ideas, goods, and people. These findings imply that the diffusion of cultural traits, such as the representation of eyes on idols, might have been influenced less by stylistic preferences and more by the interactions and exchanges occurring along these crucial riverine networks. This perspective challenges the conventional understanding of cultural transmission during this period, highlighting the need for a deeper exploration of how geographical and environmental factors shaped the development and

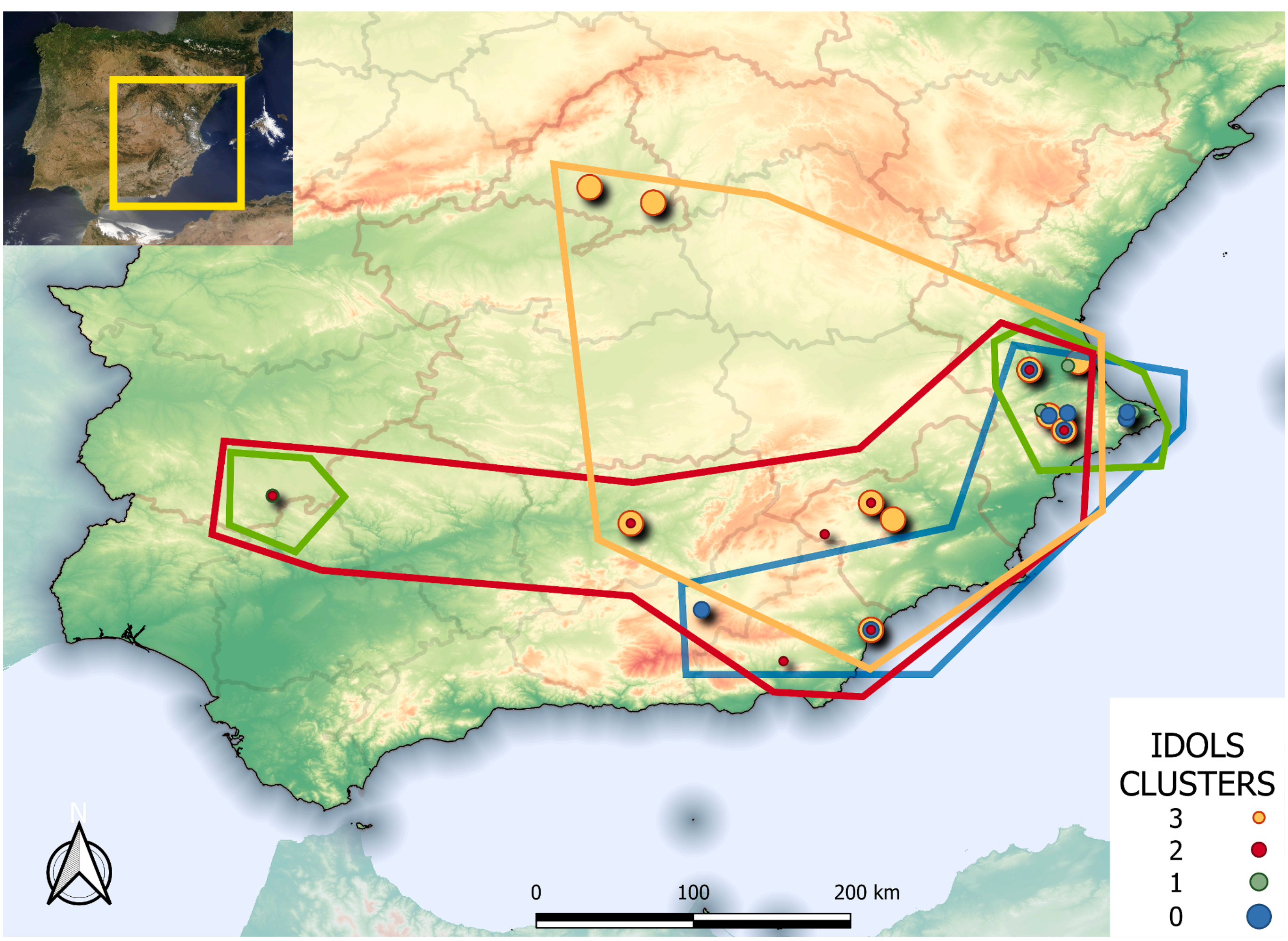

Fig. 8. Oculated idols on long bone Spectral Clustering results.

spread of Neolithic societies across the Iberian Peninsula.

This proposal goes beyond simply presenting findings; it embodies a synthesis of scientific rigor and historical storytelling. As we navigate through the clusters of images, we embark on a journey that blends the precision of ML with in-depth interpretive analysis, creating new pathways to understand the nuances of our dataset. The emphasis on the Copper Age—a pivotal period that connects the Neolithic with the Bronze Age—highlights the importance of this study. This era, marked by significant technological advancements, the emergence of extensive trade networks, and the development of complex societal structures, serves as a crucial point for exploring the evolution of hierarchical societies in the Iberian Peninsula.

The analysis of oculated idols from this time is especially revealing, shedding light on the social and cultural dynamics of the era. By employing ML to classify and examine these idols' decorative patterns, we uncover insights into the societal organization and cultural interplay of the Copper Age. These patterns not only reflect changes in social roles and religious practices but also illustrate the complex interactions among different communities, driven by trade and migration. Our study underscores the importance of stylistic variations in idols as a window into the cultural exchanges and influences of the time.

Utilizing ML, we expose patterns indicative of a rich tapestry of artistic and cultural traditions, offering fresh viewpoints on the interactions among various groups across the Iberian Peninsula. This innovative method not only augments traditional archaeological techniques but also paves the way for future investigations, particularly in transformative periods like the Copper Age. Moreover, our research enriches the narrative by exploring the symbolism of eye decorations on idols, suggesting that these motifs represent more than artistic choices—they signify social distinctions and roles. The identification of "artistic schools" and regional workshops, each with distinct stylistic approaches to eye depiction, allows for a deeper understanding of cultural expression influenced by geographic and social networks. This focus on personalized artistic exchanges and the influence of local myths on visual representations deepens our understanding of the cultural milieu that unified these societies.

In summary, applying computer vision (CV) techniques and generating visual descriptors without relying on class labels offers a promising strategy for exploring large, unlabeled datasets in various fields. This approach, grounded in unsupervised learning, uncovers natural patterns within the data, potentially revealing new categories or phenomena. It enhances scalability by minimizing the need for extensive manual labeling. However, extracting meaningful and robust visual descriptors is critical, as the absence of labels complicates the interpretability and validation of identified patterns. From a methodological standpoint, the differential weighting of visual descriptors—such as shape, texture, and pattern intricacy—plays a critical role in achieving accurate typological classification. Our analysis demonstrates that certain objective descriptors provided by tested algorithms such as SIFT, and not biased by the researcher considerations, have the potential to provide significantly grouping of idols, which reflect their cultural and artistic relevance.

The success of this methodology depends on sophisticated algorithms and careful tuning to distinguish meaningful connections from superficial similarities. Thus, while CV and unsupervised learning techniques hold significant promise for advancing our comprehension of complex datasets, they also require careful consideration of their limitations and challenges. Achieving a balance in leveraging these techniques' strengths and addressing their weaknesses is key to unlocking new insights from unlabeled data. It is essential to acknowledge that our conclusions are context-specific and require validation through further research. Future studies could extend this analytical framework to other artefacts and regions within the Iberian Peninsula, thereby testing the robustness of our findings across varied archaeological contexts. Such investigations could unveil new questions regarding Copper Age societies, such as variations in cultural practices and artistic expressions across different communities. By embracing a broader scope of analysis and incorporating diverse datasets, future research can continue to expand our knowledge of ancient communities, leveraging both existing and emerging methodologies to offer deeper insights into our shared human past.

Integrating these findings with evidence of inter-regional connectivity through river systems offers a comprehensive view of cultural exchange and its driving forces. This narrative not only connects cultural divides across the Iberian Peninsula but also sheds light on the complex web of influences that shaped these societies' historical trajectories. By merging technological innovations with archaeological evidence, this holistic approach opens new perspectives in the study of prehistoric idols, highlighting the intricate dynamics of cultural transmission and interaction in ancient societies.

Funding

This study has been funded under the research project Prometeo 2021/007: Neonets. A social network approach to understanding evolutive dynamics of neolithic societies (c. 5600–2000 cal BC), financed by the Valencian Ministry of Innovation, Universities, Science and Societies.

CRediT authorship contribution statement

Joaquín Jiménez-Puerto: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

I am very grateful to Joan Bernabeu for his comments. I would also like to express my gratitude to the anonymous reviewers whose corrections have ostensibly improved this work. Any errors or omissions are our own.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jasrep.2024.104521.

References

Adams, J.L., 2014. Ground stone analysis: a technological approach. University of Utah Press.
Almagro-Basch, M., 1965. El poblado de Almizaraque de Herrerías (Almería).
Almagro-Gorbea, M.J., 1969. Los ídolos-cilindros del bronce I en la península ibérica. Trab. Prehist. 26, 221.
Aranda Jiménez, G., García Sanjuán, L., Lozano Medina, Á., Costa Caramé, M., 2013. Nuevas dataciones radiométricas del dolmen de Viera (Antequera, Málaga). la colección gómez-Moreno. Menga. Rev. Prehist. Andal. 4, 235–249.
Arribas, A., Molina, F., Saez, L., de la Torre, F., Aguado, P., Nájera, T., 1983. Nuevas excavaciones en Los Millares (1978-1981). XVI CNArq (Murcia-Cartagena, 1982) 147–166.
Ayala-Juan, M.M., 1985. Aportación al estudio de los ídolos calcolíticos de Murcia. Anales De Prehistoria y Arqueología. 23–32.
Ballester, I., 1946. Ídolos oculados valencianos. servicio de investigacion prehistorica.
Ballester, I., 1949. Excavaciones en Cova de la Pastora (Alcoi). La labor del SIP y su Museo en los años 1940 a 1948 41–65.
Bernabeu, J., 1979. Los elementos de adorno en el eneolítico valenciano. SAGVNTVM. Papeles Del Laboratorio De Arqueología De Valencia 14, 109–126.
Bernabeu Aubán, J., Orozco-Köhler, T., Pascual Beneyto, J., 2017. A propósito de tres nuevas fechas. el camí de missena y los poblados con fosos en el país valenciano. MARQ. ARQUEOLOGÍA Y MUSEOS 8, 57–61.

Bernabeu, J., Pascual Benito, J., Orozco Köhler, T., Badal Garcia, E., Fumanal García, M. P., García Puchol, O., 1994. Niuet (L'Alqueria d'Asnar). Poblado del III Milenio aC. Recerques del Museu d'Alcoi 9–74.

Bernabeu, J., 1981. La Cova del Garrofer, Ontinyent (Valencia).

Bickler, S.H., 2021. Machine learning arrives in archaeology. Adv. Archaeol. Pract. 9, 186–191. https://doi.org/10.1017/aap.2021.6.

Bishop, C.M., Nasrabadi, N.M., 2006. Pattern recognition and machine learning. Springer.

Boaventura, R., 2009. As antas e o Megalitismo da região de Lisboa.

Boaventura, R., 2011. Chronology of Megalithism in South-Central Portugal. Exploring Time and Matter in Prehistoric Monuments: Absolute Chronology and Rare Rocks in European Megaliths. 159–190.

Bradley, R., 1999. The significance of monuments: on the shaping of human experience in Neolithic and Bronze Age Europe, Routledge. ed. J. Bellows., London and New York.

Brahim, K., Treuillet, S., Exbrayat, M., Jesset, S., 2023. Facsimiles-based deep learning for matching relief-printed decorations on medieval ceramic sherds, in: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1613–1622. Doi: 10.1109/ICCVW60793.2023.00176.

Caliński, T., Harabasz, J., 1974. A dendrite method for cluster analysis. Communications in Statistics-Theory and Methods 3, 1–27. https://doi.org/10.1080/03610927408827101.

Canny, J., 1986. A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. PAMI-8, 679–698. https://doi.org/10.1109/TPAMI.1986.4767851.

Canuto, M.A., Estrada-Belli, F., Garrison, T.G., Houston, S.D., Acuña, M.J., Kováč, M., Marken, D., Nondédéo, P., Auld-Thomas, L., Castanet, C., 2018. Ancient lowland Maya complexity as revealed by airborne laser scanning of northern Guatemala. Science 361, eaau0137. https://doi.org/10.1126/science.aau0137.

Chapman, R., 2008. Producing inequalities: regional sequences in later prehistoric southern Spain. J. World Prehist. 21, 195. https://doi.org/10.1007/s10963-008-9014-y.

Chapman, R., Chapman, C.R., Renfrew, C., et al., 1990. Emerging complexity: the later prehistory of south-east Spain. Cambridge University Press, Iberia and the west Mediterranean.

Chetouani, A., Treuillet, S., Exbrayat, M., Jesset, S., 2020. Classification of engraved pottery sherds mixing deep-learning features by compact bilinear pooling. Pattern Recogn. Lett. 131, 1–7. https://doi.org/10.1016/j.patrec.2019.12.009.

Cieslak, M.C., Castelfranco, A.M., Roncalli, V., Lenz, P.H., Hartline, D.K., 2020. t-distributed stochastic neighbor embedding (t-SNE): a tool for eco-physiological transcriptomic analysis. Mar. Genomics 51, 100723. https://doi.org/10.1016/j.margen.2019.100723.

Costa Caramé, M.E., Díaz-Zorita Bonilla, M., García Sanjuán, L., Wheatley, D.W., 2010. The copper age settlement of Valencina de la Concepción (Seville, Spain): demography, metallurgy and spatial organization. Trab. Prehist. 67, 85–117. https://doi.org/10.3989/tp.2010.10032.

Dalal, N., Triggs, B., 2005. Histograms of oriented gradients for human detection, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). Ieee, pp. 886–893.

Davies, D., Bouldin, D., 1979. A cluster separation measure: IEEE transactions on pattern analysis and machine intelligence. itpidj 0162-8828, pami-1, 2 224–227. Crossref Web of Science.

Diaz del Río, P., 2011. Labor in the making of Iberian Copper Age Lineages., in: Comparative Archaeologies: The American Southwest (AD 900-1600) and the Iberian Peninsula (3000-1500 BC). Oxbow Books, pp. 37–56.

Enríquez, J.J., 1983. Dos ídolos sobre hueso largo procedentes de la Huerta de dios. Trab. Prehist. 40, 293.

Enríquez-Navascués, J.J., García-Cabezas, M., Martín-Rodríguez, P., Rodríguez-Hidalgo, A., 2013. Los depósitos rituales calcolíticos de Huerta Zacarías I (Almendralejo, Badajoz). In: VI Encuentro De Arqueología Del Suroeste Peninsular., pp. 877–896.

Escoriza, T., 1991. La formación social de los Millares y las" producciones simbólicas". Cuadernos De Prehistoria y Arqueología De La Universidad De Granada 16, 135–165. https://doi.org/10.30827/cpag.v16i0.1301.

Fletcher, D., Plá, E., Llobregat, E., 1961. La ereta del Pedregal (Navarrés, Valencia). Archivo De Prehistoria Levantina 79–96.

García-Puchol, O., Cotino Villa, F., Miret Estruch, C., Pascual Benito, J.L., McClure, S.B., Molina Balaguer, L., Alapont, L., Carrión Marco, Y., Morales Pérez, J.V., Blasco Senabre, J., 2010. Cavidades de uso funerario durante el Neolítico final/Calcolítico en el territorio valenciano: trabajos arqueológicos en Avenc dels Dos Forats o Cova del Monedero (Carcaixent, Valencia). Archivo de Prehistoria Levantina, 2010, vol. XXVIII, p. 139-206.

García-Puchol, O., McClure, S.B., Blasco Senabre, J., Cotino Villa, F., Porcelli, V., 2013. Increasing contextual information by merging existing archaeological data with state of the art laser scanning in the prehistoric funerary deposit of pastora cave, eastern Spain. J. Archaeol. Sci. 40, 1593–1601. https://doi.org/10.1016/j.jas.2012.10.015.

García-Sanjuán, L., Murillo-Barroso, M., 2013. Social Complexity in Copper Age Southern Iberia (ca. 3200–2200 Cal bc): Reviewing the 'State'Hypothesis at Valencina De La Concepción (Seville, Spain), in: The Prehistory of Iberia. Routledge, pp. 119–140.

Gauss, R., 2013. The development of metallurgy on the Iberian Peninsula: technological and social patterns of a long-term innovation process. Metal Matters: Innovative Technologies and Social Change in Prehistory and Antiquity 209–229.

Gilboa, A., Karasik, A., Sharon, I., Smilansky, U., 2004. Towards computerized typology and classification of ceramics. J. Archaeol. Sci. 31, 681–694. https://doi.org/10.1016/j.jas.2003.10.013.

Gilboa, A., Tal, A., Shimshoni, I., Kolomenkin, M., 2013. Computer-based, automatic recording and illustration of complex archaeological artifacts. J. Archaeol. Sci. 40, 1329–1339. https://doi.org/10.1016/j.jas.2012.09.018.

Gimbutas, M., 1974. The gods and goddesses of old Europe: 7000 to 3500 BC myths. Legends and Cult Images. Univ of California Press.

Gonçalves, V., 1999. Reguengos de Monsaraz, territórios megalíticos. Instituto Português De Museus. https://doi.org/10.13140/2.1.1215.6480.

Gonçalves, V.S., ANDRADE, M., Pereira, A., 2004. As placas de xisto gravadas da gruta artificial de S. Paulo 2 (Almada). Revista Portuguesa de Arqueologia 73–96.

Gusi, F., Olaria, C., 1990. El yacimiento calcolítico de terrera Ventura (Tabernas, Almería) y sus relaciones con la cultura de los Millares. Archéologie En Languedoc 299–306.

Han, J., Pei, J., Tong, H., 2022. Data mining: concepts and techniques. Morgan Kaufmann Publishers.

Harrison, R.J., 1980. The beaker folk: copper age archaeology in western Europe. Thames and Hudson.

Harrison, R., Heyd, V., 2007. The transformation of Europe in the third millennium BC: the example of`le petit-chasseur I+ iII'(Sion, Valais, Switzerland). Praehistorische Zeitschrift 82, 129–214. https://doi.org/10.1515/PZ.2007.010.

Hunt Ortiz, M.A., 2003. Prehistoric mining and metallurgy in south west Iberian Peninsula. BAR Publishing.

Hurtado, V., 1980. Los ídolos calcolíticos de la pijotilla (Badajoz). Zephyrus 30, 165–203.

Hurtado, V., 2008. Ídolos, estilos y territorios de los primeros campesinos en el sur peninsular. Acercándonos Al Pasado. Prehistoria En 4, 1–11.

Jiménez-Puerto, J., Bernabeu-Aubán, J., 2023. Linking up Bell Beakers in the Iberian Peninsula. J Archaeol Method Theory A Complex Past: Theory and Applications. Doi: 10.1007/s10816-023-09625-6.

Jiménez-Puerto, J., 2021. Base de datos de niveles con C14 de vida corta. Doi: 10.5281/zenodo.4967346.

Juan-Cabanilles, J.J., 2006. La Ereta del Pedregal (Navarrés, Valencia), in: Arqueología En Blanco y Negro: La Labor Del SIP. 1927-1950. Diputación de Valencia, pp. 189–195.

Karasik, A., Smilansky, U., 2008. 3D scanning technology as a standard archaeological tool for pottery analysis: practice and theory. J. Archaeol. Sci. 35, 1148–1168. https://doi.org/10.1016/j.jas.2007.08.008.

Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. Biol. Cybern. 43, 59–69. https://doi.org/10.1007/BF00337288.

Kohonen, T., Huang, T.S., Schroeder, M.R., 2012. Self-organizing maps, 3rd ed. Springer, Berlin / Heidelberg.

Lanos, P., Dufresne, P., 2019. ChronoModel version 2.0.

Lasaponara, R., Masini, N., 2011. Satellite remote sensing in archaeology: past, present and future perspectives. J. Archaeol. Sci. 9, 1995–2002. https://doi.org/10.1016/j.jas.2011.02.002.

Lillios, K.T., 2008. Heraldry for the dead: memory, identity, and the engraved stone plaques of neolithic Iberia. University of Texas Press.

López-Padilla, J.A., 2006. Consideraciones en torno al "Horizonte campaniforme de transición". Archivo De Prehistoria Levantina 26.

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. 60, 91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94.

Maicas-Ramos, R., 2007. Industria ósea y funcionalidad: neolítico y calcolítico en la Cuenca de Vera (Almería). Editorial CSIC-CSIC Press.

Makridis, M., Daras, P., 2013. Automatic classification of archaeological pottery sherds. Journal on Computing and Cultural Heritage (JOCCH) 5, 1–21. https://doi.org/10.1145/2399180.2399183.

Martínez-Navarrete, M.I., Barrera, J.L., Rovira Llorens, S., Sanz Najera, M., Molero Gutierrez, G., Brea López, P., Bustos Pretel, V., Bermúdez de Castro, J.M., Pérez Pérez, P., 1984. El comienzo de la metalurgia en la provincia de Madrid: la cueva y el cerro de Juan Barbero (Tielmes).

Mateu, T.E., 1990. ídolos de la Edad del Cobre del yacimiento de Las Angosturas (Gor, Granada). Zephyrvs 43.

McCarthy, J., 2014. Multi-image photogrammetry as a practical tool for cultural heritage survey and community engagement. J. Archaeol. Sci. 43, 175–185. https://doi.org/10.1016/j.jas.2014.01.010.

Molina, F., Cámara, J.A., Capel, J., Nájera, T., Sáez, L., 2004. Los Millares y la periodización de la prehistoria reciente del sureste. Simposios De Prehistoria Cueva De Nerja 2, 142–158.

Montero-Ruiz, I., Murillo, M., 2016. Los inicios de la metalurgia y el valor social del metal. Menga 15–29.

Murillo-Barroso, M., Montero-Ruiz, I., 2012. Copper ornaments in the iberian chalcolithic: technology versus social demand. J. Mediterr. Archaeol. 25, 53–73. https://doi.org/10.1558/jmea.v25i1.53.

Murillo-Barroso, M., Peñalver, E., Bueno, P., Barroso, R., de Balbín, R., Martinón-Torres, M., 2018. Amber in prehistoric Iberia: new data and a review. PLoS One 13, e0202235.

Nocete, F., Lizcano, R., Peramo, A., Gómez, E., 2010. Emergence, collapse and continuity of the first political system in the Guadalquivir Basin from the fourth to the second millennium BC: the long-term sequence of Úbeda (Spain). J. Anthropol. Archaeol. 29, 219–237. https://doi.org/10.1016/j.jaa.2010.03.001.

Nocete, F., Sáez, R., Bayona, M., Peramo, A., Inacio, N., Abril, D., 2011. Direct chronometry (14 C AMS) of the earliest copper metallurgy in the Guadalquivir Basin (Spain) during the third millennium BC: first regional database. J. Archaeol. Sci. 38, 3278–3295. https://doi.org/10.1016/j.jas.2011.07.008.

Ojala, T., Pietikäinen, M., Harwood, D., 1996. A comparative study of texture measures with classification based on featured distributions. Pattern Recogn. 29, 51–59. https://doi.org/10.1016/0031-3203(95)00067-4.

Parcak, S.H., 2009. Satellite remote sensing for archaeology. Routledge.

Pascual-Beneyto, J., Barberà, M., Ribera, A., 2005. Camí de Missena (La Pobla del Duc): un interesante yacimiento del III milenio en el País Valenciano. pp. 803–814.

Pascual-Benito, J.L., 1998. Utillaje óseo, adornos e ídolos neolíticos valencianos. Serie de trabajos varios, Diputación Provincial de Valencia, Servicio de Investigación Prehistórica, Valencia.

Pascual-Benito, J.L., 2010. Ídolos oculados sobre huesos largos en las cuencas del Júcar y del Segura.

Pedraz Penalva, T., Molina Burguera, G., 2000. Nuevo aporte al eneolítico valenciano: la cueva de las mulatillas I (Villargordo del Cabriel, Valencia). Anales De Prehistoria y Arqueología 16, 2000.

Pérez-Jordá, G., Bernabeu Aubán, J., Carrión, Y., García-Puchol, O., Molina Balaguer, Ll., Gómez Puche, M., 2011. La Vital (Gandía, Valencia). Vida y muerte en la desembocadura del Serpis durante el III y el I milenio AC. Serie de Trabajos Varios del Servicio de Investigación Prehistórica.

Remondino, F., El-Hakim, S., 2006. Image-based 3D modelling: a review. Photogram. Rec. 21, 269–291. https://doi.org/10.1111/j.1477-9730.2006.00383.x.

Reynolds, D.A., 2009. Gaussian mixture models. Encyclopedia of biometrics 741.

Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65. https://doi.org/10.1016/0377-0427(87)90125-7.

Scarre, C., 2002. Monuments and landscape in Atlantic Europe: perception and society during the neolithic and early bronze age. Psychology Press.

Sherratt, A., 1981. Plough and pastoralism: aspects of the secondary products revolution. Cambrdge University Press.

Siret, L.M.J., 1908. Religions néolithiques de l'Ibérie. Vigot frères.

Soler Diaz, J.A., 1985. Los ídolos oculados sobre huesos largos del enterramiento de el fontanal (Onil, Alicante). Lucentum, IV 1985, 15–35.

Soler-Diaz, J.A., 1985. Los ídolos oculados sobre huesos largos del enterramiento de el fontanal (Onil, Alicante). Lucentum, IV 1985, 15–35. https://doi.org/10.14198/LVCENTVM1985.4.02.

Soler-Diaz, J.A., 2002. Cuevas de inhumación múltiple en la comunidad valenciana. Bibliotheca Archaeologica Hispana, Real Academia de la Historia, Madrid.

Soler-Diaz, J.A., 2017. De nuevo sobre los ídolos oculados" Tipo Pastora": A propósito de la presentación del conjunto de elementos ideomorfos identificado en el" Fondo Arqueológico La Marina 1995", in: Nuevos Datos Para El Conocimiento de La Prehistoria En La Comarca de La Marina Alta, Alicante. Museo Arqueológico de Alicante-MARQ, pp. 283–370.

Szeliski, R., 2022. Computer vision: algorithms and applications. Springer Nature.

Valera, A.C., Godinho, R., 2010. Ossos humanos provenientes dos fossos 3 e 4 e gestão da morte nos perdigões. Apontamentos De Arqueologia e Património 6, 29–39.

Valera, A.C., 2013. Cronologia absoluta dos fossos 1 e 2 do porto Torrao e o problema da dataçao de estruturas negativas "Tipo Fosso". APONTAMENTOS PONTAMENTOS 9.

Maaten, L. van der, Boon, P., Lange, G., Paijmans, H., Postma, E., 2007. Computer vision and machine learning for archaeology.

Van Rossum, G., Drake, F.L., 2014. The python language reference. Wilmington, DE, USA, Python Software Foundation.

Vigil-Escalera, A., Martín, A., 2003. El recinto calcolítico de fuente de la Mora. Conservar y Restaurar 58–59.

Von den Driesch, A., Morales, A., 1977. Los restos animales del yacimiento de terrera Ventura (Tabernas, Amería). Cuadernos De Prehistoria Y Arqueología De La Universidad Autónoma De Madrid. 4 https://doi.org/10.15366/cupauam1977.4.002.

Von Luxburg, U., 2007. A tutorial on spectral clustering. Stat. Comput. 17, 395–416. https://doi.org/10.1007/s11222-007-9033-z.

Waterman, A.J., Peate, D.W., Silva, A.M., Thomas, J.T., 2014. In search of homelands: using strontium isotopes to identify biological markers of mobility in late prehistoric Portugal. J. Archaeol. Sci. 42, 119–127. https://doi.org/10.1016/j.jas.2013.11.004.

Yastikli, N., 2007. Documentation of cultural heritage using digital photogrammetry and laser scanning. J. Cult. Herit. 8, 423–427. https://doi.org/10.1016/j.culher.2007.06.003.