Color-based discrimination of color hues in rock paintings through Gaussian mixture models: a case study from Chomache site (Chile)

Enrique Cerrillo-Cuenca[1]* , Marcela Sepúlveda[2] , Gloria Cabello[3] and Fernando Bastías[4]

Abstract

The article explores advanced image processing techniques for pigment discrimination in rock art paintings, emphasizing color separation using RGB (red, green, blue) and LHC_{Uv} (Luminance, Hue, Chroma) imagery. It highlights the use of dimensionality reduction methods such as Principal Components Analisys PCA and Independent Component Analysis (ICA), with a focus on Gaussian Mixture Models (GMM) for probabilistic classification of image elements. This approach, applied to the Chomache archaeological site on the northernmost coast of the Atacama Desert in Chile, reveals previously undetected motifs and details, offering a nuanced perspective in rock art documentation and analysis. This proposal reinforces the value of rock art panel not only as a finished product but as a process.

Keywords Rock paintings, Coastal chile, Image processing, Gaussian mixture models, Color spaces, Color discrimination, Digital imagery techniques

Introduction

In the analysis of rock paintings, color-based differentiation of pigments is essential for analyzing the overlapping of motifs, unraveling the sequences of painting, and ultimately understanding the underlying narratives depicted in panels, among other aspects. Traditionally, the task of analyzing rock art panels has relied on the indispensable practice of visual analysis [1–4]. Furthermore, advances in computing techniques have greatly enhanced the ability to separate pigments based on color in rock art imagery, which is helpful in discerning superimpositions.

Hyperspectral imagery is notably effective in rock art studies, as it discriminates pigments by analyzing reflectivity across wider spectra than visible light [5–11]. Nonetheless, RGB (Red, Green, Blue) digital imagery analysis is common in these studies. Tools like DStretch simplify using statistical techniques with minimal input [8–10], while fulfilling the needs of most archaeologists for digital image analysis. Additionally, computing implementation of statistic techniques, such as Principal Component Analysis (PCA) or Independent Component Analysis (ICA), have been explored, showing effectiveness in separating color tones [12–15], though they come with inherent limitations. For instance, PCA or ICA in standard imagery can isolate only as many components (new grayscale images) as there are channels in the original image, typically three in RGB images. This limitation makes it challenging to separate colors in panels with more than two or three colors and tones.

Thus, the development of new approaches for enhanced visualization and color separation results in more precise

*Correspondence:
Enrique Cerrillo-Cuenca
enriqcer@ucm.es
[1] Dpt. of Prehistory, Ancient History and Archaeology, Facultad de Geografía e Historia, Complutense University of Madrid, C/Profesor Aranguren, s/n, 28040 Madrid, Spain
[2] Social Sciences Dept, Social Sciences Faculty, Universidad Tarapacá, Domingo Bondi 990/758-0319 Las Condes, Santiago de Chile, Chile
[3] Centro de Estudios Interculturales e Indígenas; Estudios Aplicados, Facultad de Ciencias Sociales, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Macul. (Campus San Joaquín), 7820436 Santiago de Chile, Chile
[4] Estudios Aplicados, Pontificia Universidad Católica de Chile, La Reconquista 889, Las Condes, 7580140 Santiago de Chile, Chile

and detailed interpretations of the superimpositions in painted motifs. However, the process of tracing involves the interpretation of culturally significant elements and filtering out those regarded as less relevant [1], creating a binary framework that reflects the observer's subjective choices in highlighting specific features. It is crucial to recognize that this practice has persisted in the sphere of digital methodologies [8, 16, 17].

Data science introduces alternatives to this paradigm, notably through the increasing popularity of machine learning techniques. These techniques allow for the analysis of data from a probabilistic perspective [18], indicating that data (pixels in images) should no longer be viewed as belonging to fixed categories (color or motifs). Instead, they are entities whose association with a particular category can be determined based on probabilities. While the use of probabilities is not new in Archaeology, its application in the digital analysis of rock art could still be deepened and expanded [19]. Therefore, it is feasible to interpret images resulting from the statistical analysis of rock art images as probability maps, where each pixel has varying probabilities of belonging to different classes (motifs or colors). This approach doesn't negate the value of traditional 'tracings' or basic digital image enhancement; instead, it offers a renewed perspective in interpreting image processing outcomes. This does not imply a formal abandonment of the usefulness of "tracings" or digital enhancement but rather the opportunity to incorporate a new perspective in understanding the outcomes of image processing. This could involve, for instance, making decisions based on the probability of a certain group of pixels belonging to the color of a pigment.

This approach aligns more closely with the formation and preservation of a painted panel, where the information is influenced by numerous factors, such as post-depositional changes [20], reuse of the painted supports, and their inherent characteristics. Like any physical process, these factors introduce a variable amount of noise, limiting our ability to recognize and isolate information, such as painted motifs. Noise also has an intrinsic component in every digital process [21], such as that produced by photographic lenses and sensors, or by the transformation processes the images undergo after their recording, like the creation of orthoimages. Consequently, statistical treatment of noise and the use of probability might offer a way to address the complexity inherent in any rock paintings panel.

This article focuses on enhancing pigment differentiation in RGB orthoimages by introducing new channels through color space transformation. It also examines the effectiveness of dimensionality reduction techniques, PCA and ICA, in handling the expanded band count in the model. Furthermore, the study delves into the use of unsupervised classification methods to produce images with probabilistic categorization of pixels. Building on the goals outlined, this article investigates various techniques to evaluate their efficacy in documenting the paintings at the Chomache site [22, 23]. The implementation of these methods facilitates a discussion on novel analytical approaches and contributes to the identification of previously undetected motifs, obscured by support degradation and the overlapping of painted motifs.

Methods and materials
Study case and data capture
Rock art studies in Chile have received substantial contributions in recent years, which have helped to highlight the extensive sequences of painted panels and their chronologies [24]. A significant site is Chomache or Chomache-58, located 40 km north of the mouth of the Loa River, on a point located 1.7 km from the current San Marcos cove where fishing communities reside and 120 km south of the Iquique city. The only large Chomache panel is on the Pacific coastline, a few meters from the ocean (Fig. 1). Its conservation is particularly exceptional despite being affected by the cloudiness and coastal humidity, but also by the bird guano that partially covers the panel since various species tend to rest on the top of the rock. The panel (Fig. 2) is approximately 4 m high and wide, arranged on a flat surface of a granite block facing east, facing the imposing slope of the Cordillera de la Costa that rises in this sector by more than 1,000 m. Previously published works on the site paintings associate it with the groups of the Pica-Tarapacá complex (1000–1450 cal AD) [25, 26]. The red, white, orange, and purple motifs correspond to camelids, fish, and geometric motifs along with only one anthropomorphic motif [19, 20]. In prior publications [25, 26], the authors have noted the presence of bird guano and paint drippings that make it difficult to identify the figures with the naked eye, so they highlighted the most defined shapes.

Data was captured using a D-5300 Nikon camera equipped with a 24×16 mm DX sensor, shooting in the camera's native RAW format with 14-bit color depth. The 'Datacolor SpyderCheckr' card was used for data acquisition, and Darktable application for color calibration and correction, producing TIF images for subsequent post-processing in photogrammetric restitution. For the photogrammetric restitution, targets were placed around the decorated support to serve as a scale. This process involved the using of Metashape software, which generated orthoimages with a resolution of 0.05 cm.

Fig. 1 Location of Chomache site (Chile)

Comprehensive overview of the workflow

The proposed workflow (Fig. 3) for this study involves exploring the possibility of using an additional color space (LHC$_{uv}$, which stands for Luminance, Hue and Color), into which the image is transformed before undergoing statistical analysis. To clarify, a color space is a system for interpreting color that organizes the colors in an image using a structured procedure [27], sometimes through a mathematical model that arranges colors in coordinates.

Transforming the image into this color space serves to introduce three extra bands alongside the existing RGB ones, culminating in a composite 6-band matrix. This procedure increases the number of variable inputs for ICA or PCA analysis. Consequently, there is a corresponding rise in the number of available output components, facilitating clearer color-based separation of pigments. It is important to note that the results of both procedures, color space transformation and components, are interpretable primarily through visual analysis.

Cerrillo-Cuenca *et al. Heritage Science*   (2024) 12:169

Page 4 of 17

Fig. 2 General photograph of the painted panel of Chomache. The overlapping of figures is appreciated. The elements painted in white are clearly identifiable, as opposed to others in red which are more complex to discern

Finally, as an alternative to the previous methods, this study applies the Gaussian Mixture Model (GMM) to the 6-band matrix for generating unsupervised probability maps. These maps gauge the likelihood of each pixel's association with a particular class. A notable advantage of this approach is its ability to produce a higher number of components. Distinctly, unlike prior procedures, the GMM's outcomes are quantified in terms of the probability that a pixel belongs to a class as statistically determined by the algorithm.

The code was developed in the Python programming language, in which all the functions were implemented, which is available at the Github repository (see "Availability of data and materials" section).

Image transformation from RGB to LHC_{uv} colorspace

The objective of this research includes increasing the number of analyzable bands within the image. This is achieved by using additional RGB channels and transferring the same information to another color space, so that, as we have indicated before, we can have a 6-band matrix. A color space is a specific organization of colors that allows for the reproduction and management of color in various applications. Each color space

can define colors through a unique set of coordinates or dimensions, ensuring consistent color reproduction across different mediums [28]. To achieve this, we convert the RGB image to the LHC_{uv} color space. LHC_{uv} or CIE LCh_{uv} (Comission Internationale de l´Éclairage, Luminance, Chroma, hue) is the cylindrical version of the CIELUV [22] which has been used since the 1970s in television and video. LHC_{uv} allows for an individual extraction of chroma (or saturation) values, hue, and luminance in three new different bands. As other colour spaces, CIELUV is an alternative for the classic representation of RGB images, and, consequently, other direct transformation procedures from RGB to CIELuv (Comission Internationale de l´Éclairage, Luminance, *u*, *v*) do exist. To carry out this transformation, we use the OpenCV library in Python [29], whose cvtColor function has the capability to transform images from RGB bands to CIELuv. This process generates a new image with three bands: *L*, *u*, and *v*. *L* corresponds to luminance or brightness, aligned with human color perception, while *u* and *v* represent chromaticity, with *u* on the red/green axis and v on the blue/yellow axis. The LCh_{uv} transformation converts CIELuv band data into three new bands: *L* for luminance, *C* for chroma, and *H* for hue. This was achieved using specific formulas on matrices derived from the CIELuv transformation [29]:

$$C = \sqrt{u^2 + v^2}$$

$$H = \frac{\arctan2(v, u) \times 180}{\pi}$$

For luminance, the *L* band from the RGB to CIELuv transformation was kept, whereas chroma and hue were derived using the mentioned formulas. *L* and *C* values are within 0 to 100, and *H* is measured in degrees. This transformation offers dual advantages: LHC_{uv} alignment with human color perception, and the consolidation of hue information into a single band or matrix. This aids in initially discerning color differences in paintings, enhanced by chroma values that indicate color intensity or saturation. Finally, the image's three bands were combined into a single matrix along with the RGB bands using the Numpy library [30]. This resulted in a new matrix comprising R, G, B, L (luminance), C (chroma), and H (hue) bands.

An essential aspect in all classification models is the calculation of correlation between variables. To address this, the Spearman's ρ of the six matrices was computed to determine the extent to which this process can reduce the high correlation typically found among bands of a conventional RGB image in rock paintings [32, 33].
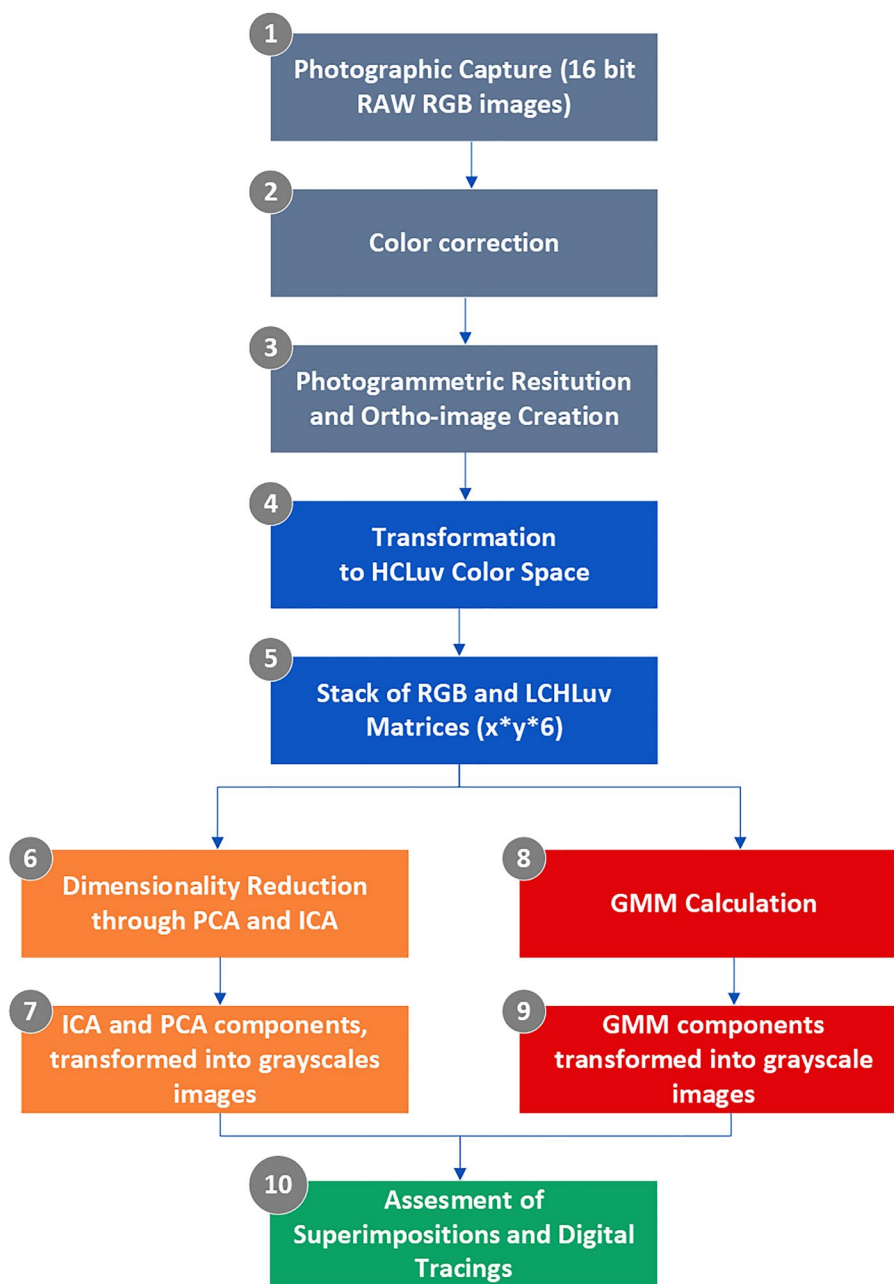
Fig. 3 Overview of the workflow implemented in this study

Reduction of variance and maximization of Independence: PCA and ICA

Decorrelation techniques in rock painting images aim to enhance the identification of painted motifs by statistically maximizing visually the relevant information of the paintings [14, 32–37]. These techniques typically employ PCA, which maximizes information separation among RGB bands by reducing variance. In other words, it possesses the capability to reproject the input channels of the RGB image into components that, in turn, can be transformed into independent images. Recently, ICA has also been introduced [6, 15], decomposing the original image into components with the highest possible independence in their information content [38]. ICA and PCA have been implemented using the scikit-image library for Python [39]. The differences between both methods are well established [15]. While PCA reduces variance among image pixels, ICA maximizes component independence,

Cerrillo-Cuenca *et al. Heritage Science*   (2024) 12:169

Page 6 of 17

influencing their digital image analysis roles. While PCA identifies orthogonal components based on variance, ICA excels in separating mixed signals into independent sources. This distinction is crucial for extracting complex features from images, where independence reveals nonlinear interactions rather than mere uncorrelatedness. Therefore, PCA enhances pigment color differentiation, whereas ICA separates these colors more clearly and efficiently. A more detailed explanation of ICA and its comparison with PCA, can be found in our studies previously undertaken in similar contexts in northern Chile [15]. This article utilizes both PCA and ICA methods to compare their outputs with those obtained from the described classification procedures in point 2.5.

Classification through variational bayesian gaussian mixture model (VB-GMM)

The procedures for classification of rock paintings often come from the domain of remote sensing, where non-supervised approaches are a common tool for classification large amounts of information [40]. Among the techniques from this field, K-means is a very straightforward technique [35], that classifies a dataset into a previously set number of classes. Nonetheless, in this study a probabilistic strategy was sought to improve the classification. Although, Machine Learning techniques are currently offering a new framework for developing such kind of approach [16, 21], in this case a non-supervised classification technique was desired to avoid the impact of subjective decision-making during the training of models. In other words, we sought to find a technique that would allow for the autonomous, unsupervised identification of pigment groups, taking into account the values of the 6-band matrix. Gaussian Mixture Models (GMM) techniques is an unsupervised classification procedure that have been successfully used in archaeological analysis [41, 42], thanks to the probabilistic outcomes and the accountability for robust testing. GMM is a statistical technique for modeling a dataset as a combination of several Gaussian distributions. This approach allows for the modeling of complex data structures by analyzing the presence of subpopulations within the overall population [43].

GMM operates on the premise that dataset information arises from a finite mix of Gaussian-distributed factors [43], each correlating with a specific group or factor. A painted surface, indeed, is an amalgamation of elements, including the painting itself with its spectrum of color tones, and the assorted factors of degradation that it undergoes. These elements are seen as processes or influential factors in the analysis. The application of GMM is geared toward categorizing each process based on the RGB values of the image, along with its luminance, chroma, and hue. This approach involves leveraging the matrix created in the previous step for several purposes: (1) to isolate factors based on the specified variables, (2) to generate several images corresponding to the statistically identifiable processes (or classes), and (3) to allocate a probability value to each pixel in these images, indicative of their association to a particular class.

In GMM, as with K-means, the number of resulting classes is predefined by the user, with K-means often serving as the initial training step. Our approach utilizes the Variational Bayesian Gaussian Mixture Model (VB-GMM), an extension of GMM that incorporates Bayesian inference [44, 45]. VB-GMM uses an automatic way to determine the optimal number of clusters in the data, unlike traditional GMM which requires this number to be set in advance. Primarily, it employs variational inference to estimate the probability distributions of the factors. Its main advantage is providing a robust way to determine the appropriate number of components, an approach that can be more complex in a standard GMM model, especially when dealing with high-dimensional data.

VB-GMM was applied to images using stacked RGB and LCHuv bands. This approach enabled effective grouping based on color values, but also in attributes such as luminance, hue, and chroma. The VB-GMM algorithm, implemented via the scikit-learn statistical package [46] in Python, was set up with up to 15 components and a maximum of 2000 iterations. Convergence was determined based on a tolerance threshold of 0.001, beyond which the iterative process would halt. The K-means algorithm was employed to initially weight the factors. Convergence in this context refers to the point where the algorithm has satisfactorily optimized the variational approximation of the posterior distribution of the model's parameters. Due to the image's substantial size, only a selected area of the painted zone and its surroundings were considered for training, with a random sample of 100,000 pixels. The model's results were subsequently applied to the entire image to produce the components.

The outputs of this procedure include a matrix or image indicating the most probable class for each pixel, and a series of matrices, one for each class, showing the probability of pixel membership in these classes, coded within a [0, 1] probability range. Furthermore, the weights of each model component were scrutinized, focusing particularly on whose weights contribute to the overall model exceeding the threshold of 0.05.

Results

Transformation from RGB to LCHuv color space and 6-band matrix creation

The initial observation in the study is that converting the image to the LHC_{uv} color space reveals painting details

Cerrillo-Cuenca *et al. Heritage Science*   (2024) 12:169

Page 7 of 17

Fig. 4 Orthoimage of the painted panel of Chomache site (Chile)

Fig. 5 Detail of the transformation of the RGB orthoimage to LHC$_{uv}$ color space. **A** Original orthoimage. **B** Luminance derived from the CIELUV color space. **C** Chroma from the LHC$_{uv}$ color space. **D** Hue from the LHC$_{uv}$ color space. The levels of images from **B** to **D** have been equalized to enhance their visualization

not easily discernible in the RGB format (Fig. 4). Specifically, extracting chroma data accentuates the distinctions in white pigments, which display greater saturation than others (Fig. 5). Additionally, the inspection of luminance and hue bands substantially improve the image, effectively bringing the motifs' contours into prominence. However, it is important to note that these enhancements stem more from the perceptual aspects of image visualization rather than from an intentional, color-based separation of the pigments.

The creation of a matrix incorporating data from the six bands facilitates the analysis of correlations among the various layers of the image. The Spearman correlation coefficient (Spearman's ρ), calculated for the RGB channels and the results from the LHC$_{uv}$ space conversion (Table 1), reveals several important observations. There is a discernible dissociation of hue and chroma from other bands, along with a strong positive correlation between luminance and RGB bands. This introduction of differential correlations assists in addressing the covariance among the variables, suggesting that the dataset contains decorrelated information, which is essential for an effective color-based separation of pigments. Moreover, it should be considered that the collinearity of variables in an RGB image, for instance, can impact the design of certain machine learning models. Therefore, the low correlation coefficients between L, C, and H might be an alternative when implementing these models.

PCA and ICA decomposition

The matrix was decomposed into principal and independent components. The results of both methods can be seen in Figs. 6 and 7. The PCA results show that despite an increase in input v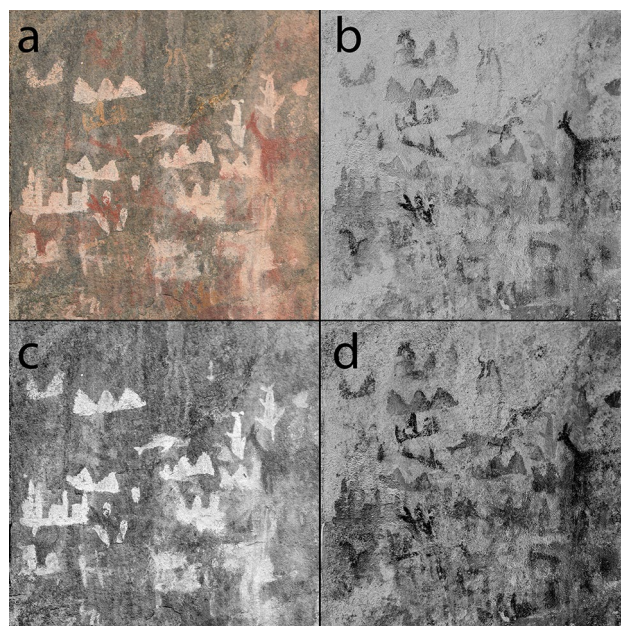ariables, pigment separation is not effective. For instance, in the first and fourth components (Fig. 6b and e), pigments of different shades are combined. The fifth component (Fig. 6f) effectively isolates white-toned paintings, but traces of red pigment are still visible. These appear in the sixth component (Fig. 6g), but not as distinctly as required. The information in these last two components has minimal variances within the model, consistent with the existing knowledge of RGB image processing of rock paintings [32].

He results from applying ICA, however, are more promising. It is important to remember that, unlike PCA, ICA components are not ordered by explained variance but randomly [31]. Thus, the discussion of the results pertains to the order returned by the algorithm in this case of study, as shown in Fig. 7. The white pigments, perhaps being the most prominent, are represented in components 1, 3, and 5 (Figs. 7b, d, and f), with greater clarity in component 1 (Fig. 7b). The red pigments, which are very faded in the original image, are clearly visible in the second component (Fig. 7c) and absent in the other components, indicating effective separation. The same inference is drawn from the fact that orange pigment traces are isolated in the sixth component (Fig. 7g). Finally, the fourth component reflects the texture of the wall.

Overall, these findings support the capability of ICA to separate colors more effectively than conventional PCA. However, the overly general nature of this color

Cerrillo-Cuenca *et al. Heritage Science*   (2024) 12:169

Page 8 of 17

Table 1 Correlation matrix (Spearman's ρ) of the 6 bands

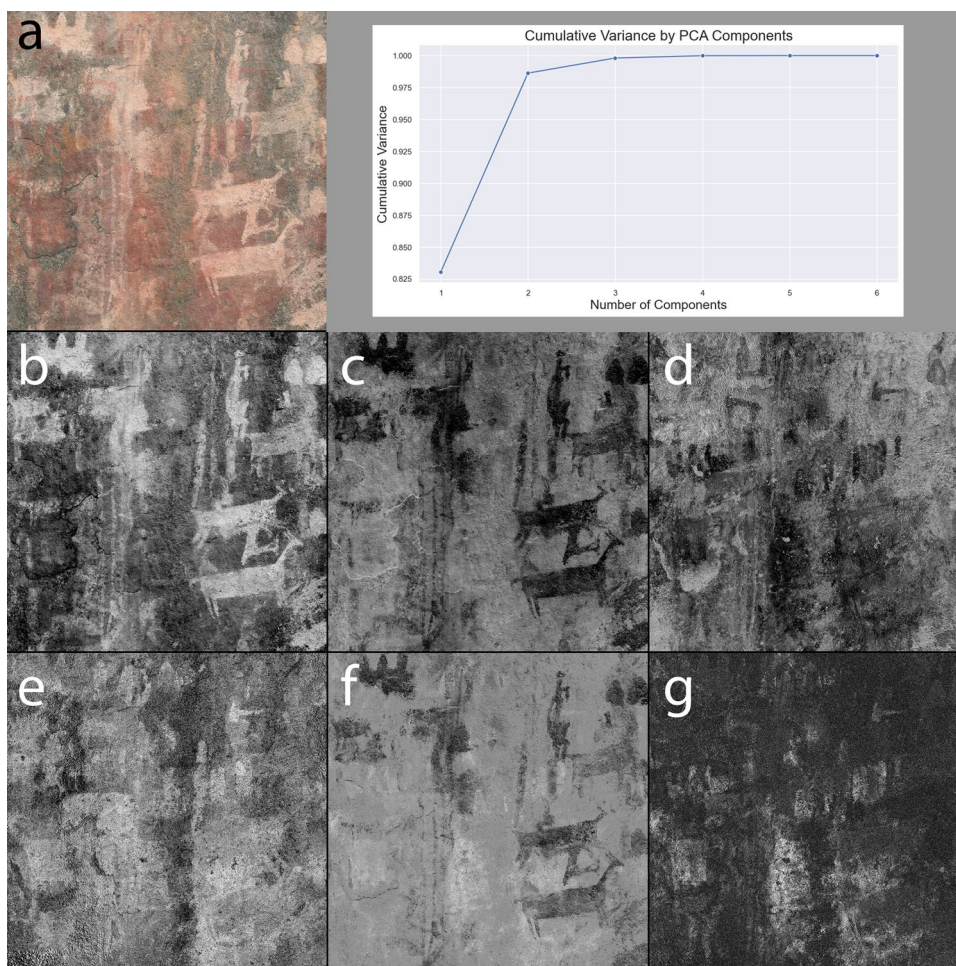|	R	G	B	L	C	H
R	1	0.97	0.88	0.97	− 0.23	− 0.03
G	0.97	1	0.93	0.99	− 0.32	− 0.01
B	0.88	0.92	1	0.94	− 0.62	− 0.33
L	0.97	0.99	0.94	1	− 0.39	− 0.04
C	− 0.23	− 0.32	− 0.62	− 0.34	1	0.8
H	− 0.02	− 0.01	− 0.33	− 0.04	0.8	1

Fig. 6 Results of PCA application with cumulative variance graph. **A**. Original Orthoimage. **B** Component 1. **C** Component 2. **D** Component 3. **E** Component 4. **F** Component 5. **G** Component 6"

separation does not allow for the identification of different shades, which is one of the reasons for seeking more specific solutions, such as those based on GMM.

VB-GMM (variational Bayesian Gaussian mixture model)
With the input parameters outlined in Sect. "Classification through Variational Bayesian Gaussian Mixture Model (VB-GMM)", the VB-GMM model converged satisfactorily after 1315 iterations. The results imply that the 15 returned components adequately capture the complexity of the image's color distributions, indicating successful isolation of key data features. The components are not arranged based on a specific index, but the weights of their contributions to the model
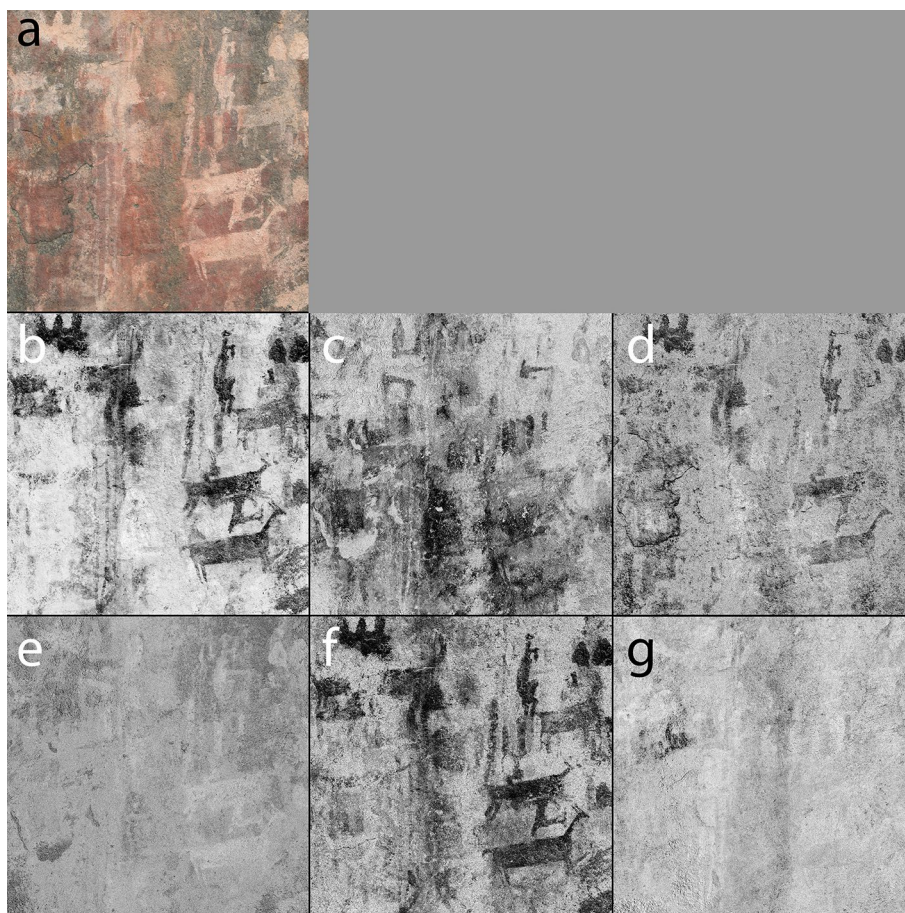
Fig. 7 Results of ICA application with cumulative variance graph. **A**. Original Orthoimage. **B** Component 1. **C** Component 2. **D** Component 3. **E** Component 4. **F** Component 5. **G** Component 6

(Table 2) show that components 1, 5, 10, 12, 13, and 14 have contributions over 0.05. After analyzing all 15 components, their interpretations have been noted in the same table. This classification reveals that components with higher weights correspond to variations in the wall's texture, which predominates in the components generated by the model.

Notably, the effect of image overlay from the photogrammetric restitution process is evident in component 4, and the model's ability to separate wall cracks is seen in component 6. Regarding pigments, the model demonstrates an excellent capacity to separate red shades in components 2 and 11, and white in components 9 and 15. This result might be attributed to the introduction of the 'chroma' variable into the model, effectively distinguishing different saturations of similar tones. It should be noted that some white-painted figures appear to have a reddish veil due to pigment fading, which may affect color classification. However, the model fails to isolate orange motifs, which are represented in components 9 and 11 and associated with classes identified as red and

Table 2 Weight contribution of each component in the VB-GMM Model and proposed identification

Component	Weight	Proposed identification
1	0.144	Variations in wall texture
2	0.020	Faded red pigments
3	0.048	Variations in wall texture
4	0.007	Variations in wall texture, influence of orthoimage creation process
5	0.183	Variations in wall texture
6	0.005	Wall cracks with dark tones
7	0.042	Variations in wall texture
8	0.018	Variations in wall texture and dark-toned cracks
9	0.032	Intense white pigments, orange pigment
10	0.161	Variations in wall texture
11	0.009	Intense red pigments, orange pigment
12	0.075	Variations in wall texture
13	0.133	Variations in wall texture
14	0.080	Variations in wall texture
15	0.046	White pigments

Cerrillo-Cuenca *et al. Heritage Science*   (2024) 12:169

Page 10 of 17

white. Their misidentification could be attributed to their limited representation in the input dataset.

The availability of probability estimates gives the interpreter of the tracings an additional criterion for producing these tracings. It is crucial to note that the obtained probability degree signifies the likelihood of a specific pixel belonging to the class of a particular component. That is, a pixel or a group of pixels may have varying probabilities across different components. Figure 8 illustrates the distinctions between component 2 and component 11, both reflecting the presence of pigments with reddish tones. Component 11 (Fig. 8a) displays more intense reddish tones with higher probability values. Component 2 (Fig. 8b) reveals more faded reddish figures with lower probability values, correlating perhaps with their poorer state of preservation. This last result enables the identification of figures that were absent in previous studies.

Regarding the motifs painted in white, their isolation introduces unique difficulties. Component 9 accurately captures the figures in this shade, including some that are faded, which become more pronounced in the image generated by this procedure (Fig. 9a). Component 15 also provides information about white pigments, but here it exclusively reflects those that apparently possess greater intensity (Fig. 9b).

Identification of superimpositions and creation of tracings

The complexity of the Chomache paintings required a detailed analysis of the motifs, which was made possible using the components generated by VB-GMM, and sometimes by combining them with some from ICA. For instance, as noted, the orange pigments were better isolated with ICA than with the components resulting from VB-GMM. The motifs identified in the imagery were manually traced using a vector design software. This process enabled the creation of tracings that are separable by color tones, as shown in the figures from 10, 11 and 12.

Discussion

Existing literature on the digital analysis of rock paintings underscores the importance of developing straightforward methods that align with archaeologists' training [9]. While this approach may be agreeable, there are case
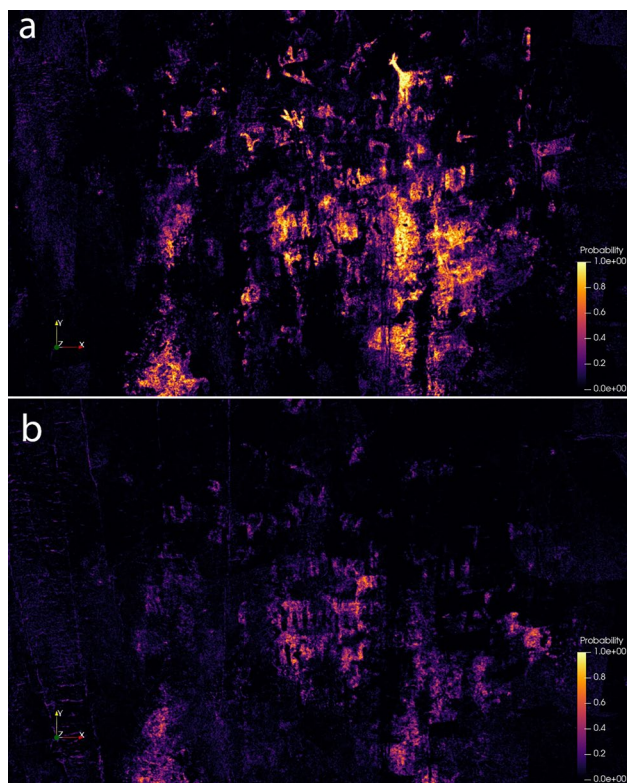

Fig. 8 Images resulting from the application of the VB-GMM algorithm on the stacking of the RGB and LCHuv bands. Probability values are shown for each pixel without applying any type of threshold. **A** Component 11 and **B** Component 2. Both images represent the same extent
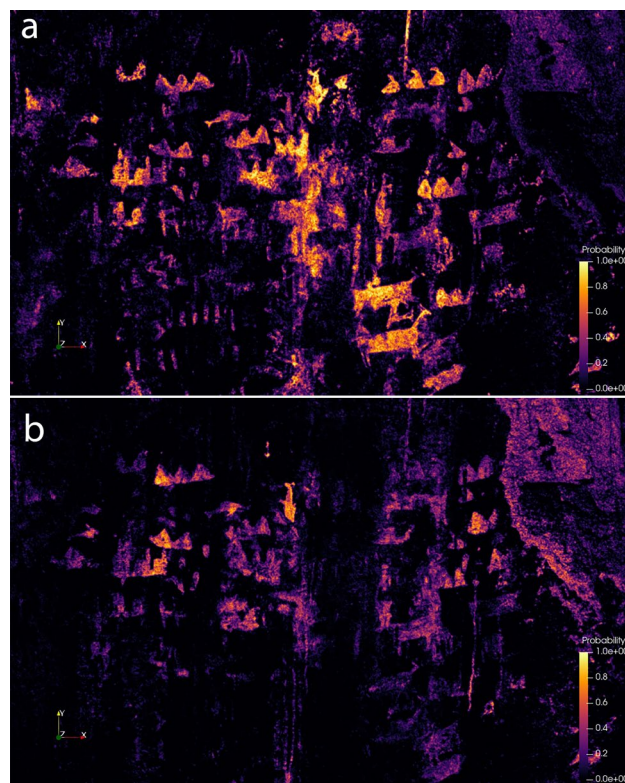

Fig. 9 Images resulting from the application of the VB-GMM algorithm on the stacking of the RGB and LCHuv bands. Probability values are shown for each pixel without applying any type of threshold. **A** Component 9 and **B** Component 15. Both images represent the same extent

Fig. 10 Tracing generated for the orange-colored motifs

Fig. 12 Tracing generated for the red-colored motifs, with different shades and preservation

Fig. 11 Tracing generated for the white-colored motifs

studies whose complexity necessitates the design of ad hoc solutions. In the case of Chomache, the overlay of motifs and the deterioration of the surface required the application of a specific strategy. This approach began with the use of traditional RGB images and techniques for color space transformation, information separation, and classification. Each of these methods offers undeniable advantages and complexities, but most importantly, they can be employed individually to tackle the documentation of panels where more common digital analysis techniques encounter limitations.

Transformation to color spaces

LHC_{uv} color space proves to be an effective transformation for visual image analysis, mainly in the analysis of the H band, which represents the difference in hues. It may not be the only useful color space, but a significant advantage of LHC_{uv} is that it helps avoid the problem of high correlation typically found between RGB bands, which in principle enhances the success of automatic classification procedures. Transforming images into a new color space can become an independent approach for the visual analysis of images. One of the main reasons is the wide variety of color spaces available for transformation from an RGB image [22], and there is also software that allows for the implementation of these transformations in a simple and mechanical way. Some authors have already successfully applied these transformations in rock paintings using the discontinued software retroReveal [47].

Information separation and classification

Both PCA and ICA are dimensionality reduction techniques effective in pigment separation, but they primarily focus on reducing covariance rather than direct data classification. Consequently, reprojecting the resulting components into a new space result in visually expressive new images. The results align with findings from previous studies [15]: while PCA enhances information separation, independent components prove more effective in separating colors in painted motifs. This suggests that ICA might be the most effective method for unmixing the color values in images of painted panels, especially when documented with multispectral or hyperspectral cameras. An interesting approach could involve reducing the output component number in ICA, so it returns fewer but more representative components corresponding to the color shades. However, this procedure should be evaluated tentatively.

In contrast, unsupervised classification techniques allow for the grouping of pixels based on affinities. The main advantage of the proposed classification method is that the number of components resulting from a VB-GMM analysis exceeds the limitation of restricting the number of components to the input channels. While VB-GMM is more time-consuming compared to PCA or ICA, the ability to use probabilities increases the interest of the method. Employing images with associated probability scales facilitates the automatic creation of tracings, achievable using thresholds of probability to distinguish between noise and relevant information. Nevertheless, making these decisions is not straightforward. Figure 13 illustrates how, at a low probability threshold, certain areas and figures cease to be represented in the images. This example highlights that pixels with low probabilities of class membership can still hold significant information for panel analysis.

Superimposition of figures and its implication for Chomache interpretation

The process facilitated the interpretation of motif overlays in various tones, crucial for deciphering the sequence of panel paintings. Challenges arise due to the panel's size, the difficulty in discerning certain overlays, and the complexity in arranging painting events chronologically. The interpretation relies on vector tracings and orthoimage analysis, with the latter essential for deducing figure superpositions. It is generally noted that orange-toned motifs tend to overlay other panel elements (Fig. 14). Although uncommon, these motifs frequently overlap with both white and red motifs (Figs. 15 and 16). White motifs are predominant, often overlapped by red figures, which vary from faded symbols to vivid, figurative camelids (Fig. 14). The arrangement of white figures appears cohesive and structured, yet in one case, an animal figure is seen overlapping a probable anthropomorphic figure (Fig. 15), further layered over a vague reddish pigment. The proposed sequence of painting is tentatively as follows:
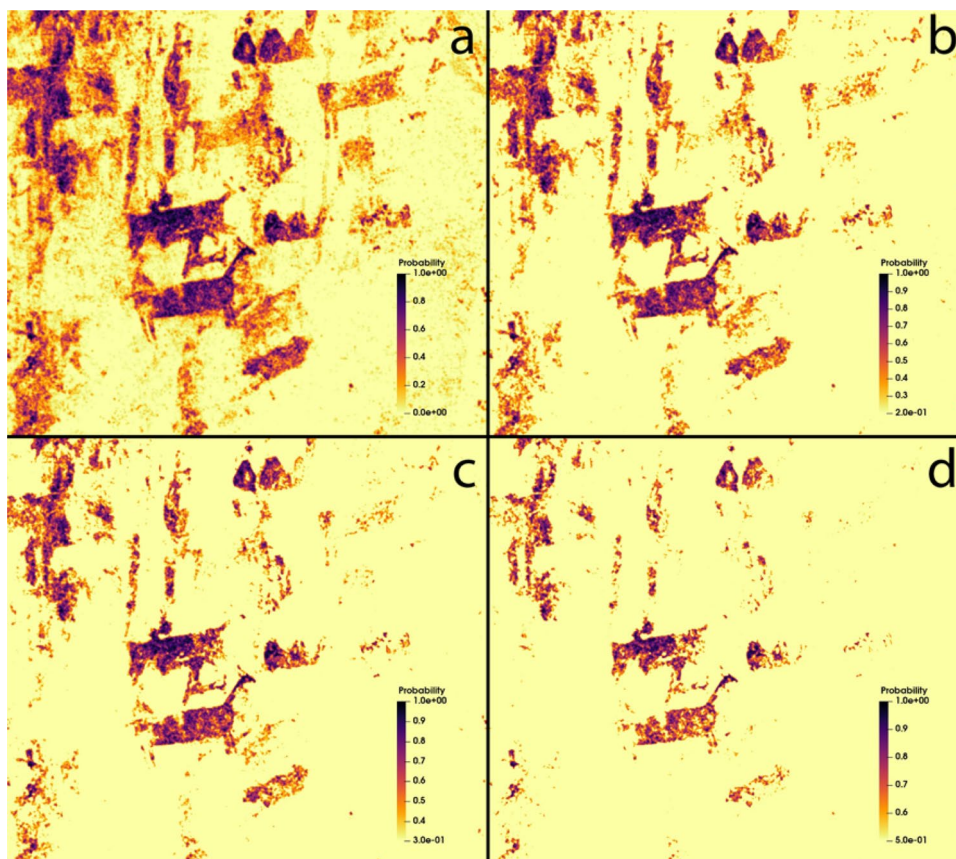

Fig. 13 Example of applying probability thresholds to produce automatic tracings from VB-GMM components, in this case, Component 11. **A** Unfiltered probability thresholds, **B** Probability values below 0.2 filtered out, **C** Values below 0.3 filtered out, **D** Values below 0.5 filtered out
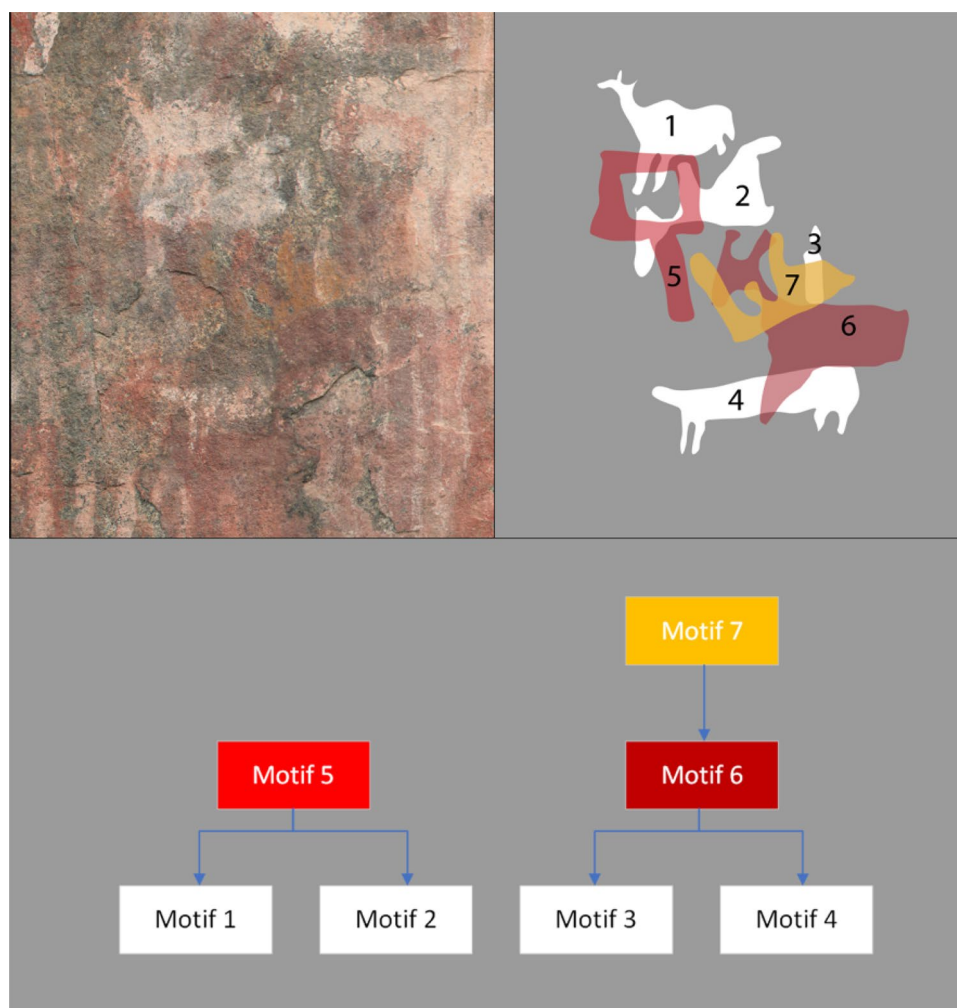
Fig. 14 Analysis of the superimposition of figures in different shades (white, red, and orange) from the Chomache site. Upper left, the original image, upper right, a proposal of the superposition of figures. Below, a matrix indicating the sequence of superimpositions

1. An initial layer of reddish pigment with indistinguishable figures.
2. White motifs, encompassing both figurative and symbolic elements, typically exhibiting a single layer of overlaps. Sporadic anthropomorphs and predominant animal figures like camelids, fish, and birds are included.
3. Reddish figures with diverse tones and preservation levels, superimposed on white, often augmenting other motifs in this hue. Figures whose referents we may or may not identify in nature are represented.
4. Rare orange or yellow motifs, superimposed on both white and red figures.
5. White animal motifs over red figures, animals, or undefined stains.

However, it is important to note that this sequence is neither linear nor unique. While it represents the most common pattern observed, other sections of the panel may exhibit slight variations.

Perspectives on the use of the described procedures

Individually, each of the methods described in the previous sections represents an improvement in the analysis of figures painted on rock art panels. Several considerations are worth making regarding their application in other sites and to other more widespread digital documentation procedures, as DStretch.

As for the first point, application in other contexts, the method may find limitations depending on the preservation conditions of the pigments, something that is evident in the field of rock art. Consequently, poor preservation would influence the study of overlays. Techniques such as PCA have been specifically employed to
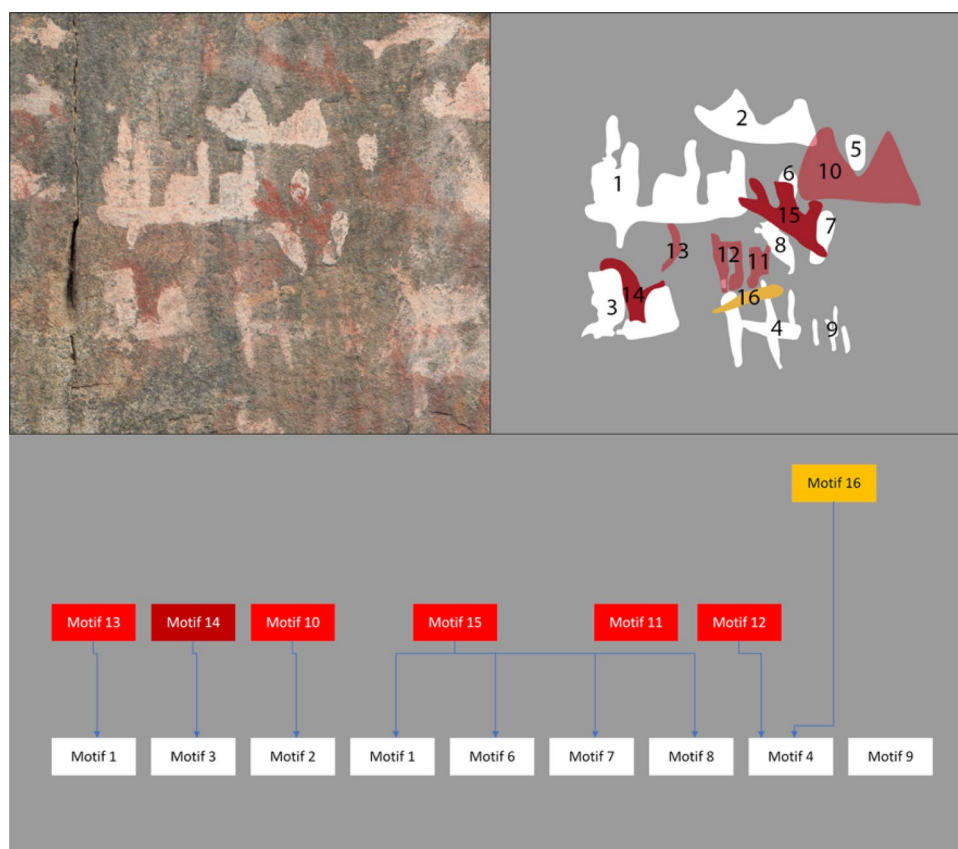
Fig. 15 Analysis of the superimposition of figures in different shades (white, red, and orange) from the Chomache site. Upper left, the original image, upper right, a proposal of the superposition of figures. Below, a matrix indicating the sequence of superimpositions

enhance the visibility of faded paintings and form the basis of this study. Nonetheless, in instances of severe deterioration, the effectiveness of these techniques—including DStretch—might be less impactful. Each painted surface presents its own technical complexities, and in such cases, turning to more advanced methods of characterization, like multispectral imaging or a combined approach of multiple techniques, could prove more beneficial. However, since the formation and preservation of panels result from a combination of factors, predicting how any technique will perform is always complex and speculative.

Secondly, technical solutions like DStretch can serve an "exploration" function for these supports, not forgetting that its development is based precisely on PCA and the transformation between color spaces [14]. This solution is effective in many cases but cannot isolate sets of figures with similar tonalities. It is worth evaluating whether the time investment and steep learning curve of solutions like the one described are efficient. In panels with no overlaps, where preservation is excellent and chromatic diversity is minimal, simple, and direct solutions like DStretch might be an effective choice. However, in situations with

significant challenges, such as those found in Chomache, crafting custom solutions can help to unveil and understand more intricate compositions. The techniques described in this paper possess all the requisites for adaptation across various painted rock panels; however, their efficacy is contingent upon the previously referenced factors. Essentially, the unique characteristics of the panels and the investigative questions we pose should guide our investment in time and learning towards the application and development of a methodological strategy. For our team, the developed approach facilitates a fresh interpretation of the panel, shedding light on sequences of painting. Consequently, the effort expended in applying these techniques is justified by the site's significance, its distinctive cultural setting, and the opportunity to assess the panel as an evolving narrative rather than a static product.

Conclusions

The creation of specific frameworks for the analysis of painted panels is now possible thanks to the implementation of ad hoc solutions from fields like image processing and automatic classification. However, these techniques
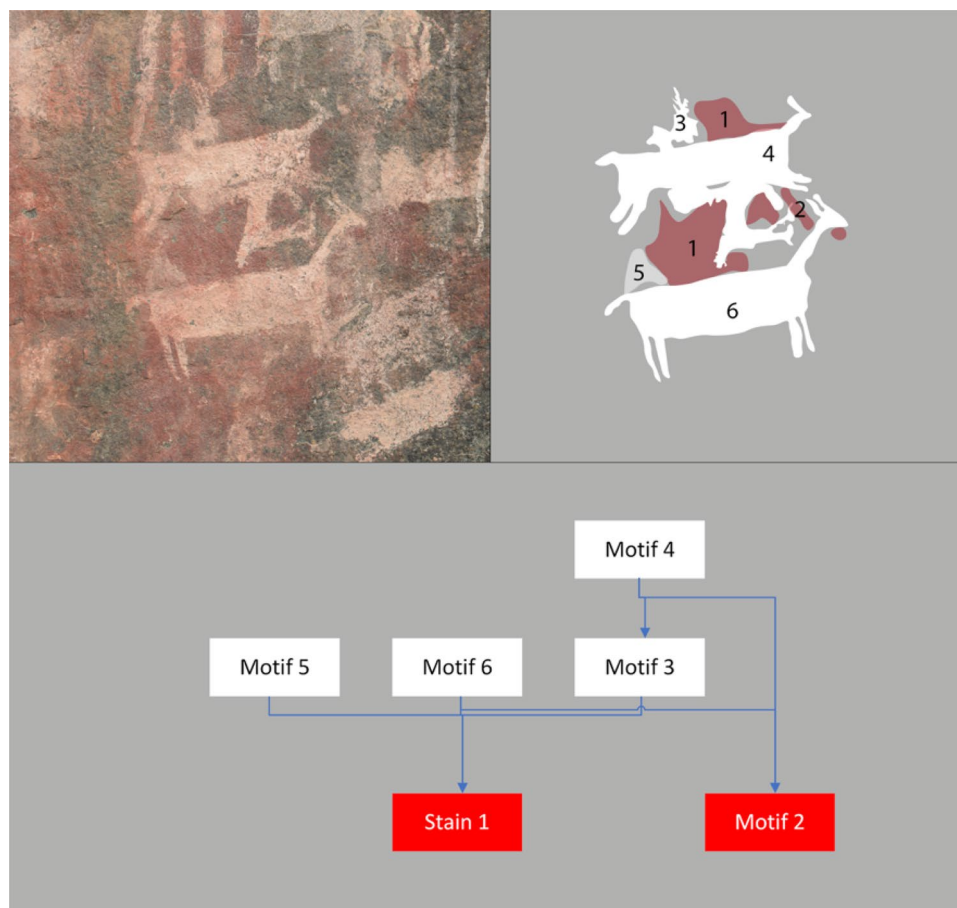
Cerrillo-Cuenca *et al. Heritage Science* (2024) 12:169

Page 15 of 17

Fig. 16 Analysis of the superimposition of figures in different shades (white and red,) from the Chomache site. Upper left, the original image, upper right, a proposal of the superposition of figures. Below, a matrix indicating the sequence of superimpositions

are not universally applicable in rock art documentation, nor is there a single solution that addresses all documentation needs of a panel. Method selection should be based on the specific questions of each case, and having a repertoire of validated solutions enhances each team's interpretative capability.

In Chomache, the application of these methods has unraveled a complex sequence of superimpositions and motifs. The methodological proposal discussed in this paper is based on conventional RGB image processing, extendable to cases using multispectral or hyperspectral images. Conclusions from the applied techniques include:

1. Transforming RGB images to other color spaces, like LCh_{uv}, can highlight motifs difficult to discern in the original images. This space, by encoding hue variations in one coordinate, facilitates component separation and classification, avoiding the high correlation typical of RGB bands of rock paintings.

2. ICA is confirmed effective in separating signals of different color tones, allowing for more accurate interpretations of superimpositions.

3. GMM models, particularly VB-GMM, are effective in separating color tones, overcoming previous methods' limitations like the number of components. Statistically robust and useful in terms of probability, their implementation can be time intensive.

These results point toward a new way of assessing color group membership, increasing the ability to identify superimpositions or faded figures. The analytical digital model's application is especially interesting and relevant for panels that are poorly preserved and feature numerous, often hard-to-distinguish overlays. The panel's history and its defining compositions reveal multiple overlays of colors and shapes added in various, not necessarily sequential ways, associating specific figures with particular colors. Many hands contributed to its creation, confirming that diverse human imaginaries, likely from

Cerrillo-Cuenca *et al. Heritage Science*   (2024) 12:169

Page 16 of 17

inland and coastal areas [19, 20], might have converged at this site.

Chomache was undoubtedly a favored space, highly appealing to different action paintings on a single panel. In this sense, this article is an invitation to look beyond the known forms that allow us to find references and make comparisons, typical of the history of Western art. We give importance to the processes involved in the intervention of a place: the action of making art and the effects of color, to the sensations that this creates, seeing rock art not only as a product but as a process, continuous and constant, that can directly and lastingly influence human beings [48–51]. Something that seems to not only be an exception among the marine hunting and gathering populations of the Pacific Ocean coast of the southern cone, according to recent evidence [52, 53]. In definitive, we reinforce methodologically the study of rock art in dynamic making and not as a static and finish construction.

Abbreviations

C	Chroma
CIE LCh$_{uv}$	Comission Internationale de l´Éclairage, Luminance, Chroma, hue
CIELuv	Comission Internationale de l´Éclairage, Luminance, *u, v*
GMM	Gaussian mixture model
H	Hue
ICA	Independent component analysis
L	Luminance
LHC	Luminance, Hue, Chroma
PCA	Principal component analysis
RGB	Red, Green, Blue
VB-GMM	Variational Bayesian Gaussian mixture model

Author contributions
ECC recorded the paintings, designed the analytical procedure, wrote the Python code and was a contributor in writing the manuscript. MS surveyed the site, recorded the paintings, interpreted the results and was a contributor in writing the manuscript. GC surveyed the site, recorded the paintings, interpreted the results and was a contributor in writing and revising the manuscript. FB surveyed the site, recorded the paintings, and performed the photogrammetric restitution. All authors read and approved the final manuscript.

Availability of data and materials
The code generated during the current study is available in the Github repository, https://github.com/ecerrillo/Chomache

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References

1. Domingo Sanz I. Rock art recording methods: from traditional to digital. In: Smith C, editor. Encyclopedia of global archaeology. New York: Springer; 2014. p. 6351–7.
2. Fiore D. The economic side of rock art: concepts on the production of visual images. Rock Art Res. 2007;24(2):149–61.
3. Janik L. Seeing visual narrative. New methodologies in the study of prehistoric visual depictions. Archaeol Dialogues. 2014;21(1):103–26. https://doi.org/10.1017/S1380203814000129.
4. Robb J. Art (Pre)history: ritual, narrative and visual culture in neolithic and bronze age Europe. J Archaeol Method Theory. 2020;27(3):454–80. https://doi.org/10.1007/s10816-020-09471-w.
5. Bayarri V, Sebastián MA, Ripoll S. Hyperspectral imaging techniques for the study, conservation and management of rock art. Appl Sci. 2019;9(23):5011. https://doi.org/10.3390/app9235011.
6. Bayarri V, Castillo E, Ripoll S, Sebastián MA. Improved application of hyperspectral analysis to rock art panels from el castillo cave (Spain). Appl Sci. 2021;11(3):1292. https://doi.org/10.3390/app11031292.
7. Schmitt B, Souidi Z, Duquesnoy F, Donzé F-V. From RGB camera to hyperspectral imaging: a breakthrough in Neolithic rock painting analysis. Herit Sci. 2023;11(1):91. https://doi.org/10.1186/s40494-023-00940-5.
8. Teira L, Bayarri V, Ontañón R, Castillo E, Arias P. Geometric and radiometric recording of prehistoric graphic expression: the case of Peña Tu (Asturias, Spain). Archaeol Anthropol Sci. 2024;16:32. https://doi.org/10.1007/s12520-023-01932-z.
9. Defrasne C, Massé M, Giraud M, Schmitt B, Fligiel D, Le Mouélic S, Chalmin E. The contribution of VNIR and SWIR hyperspectral imaging to rock art studies: example of the Otello schematic rock art site (Saint-Rémy-de-Provence, Bouches-du-Rhône, France). Archaeol Anthropol Sci. 2023;15:116. https://doi.org/10.1007/s12520-023-01812-6.
10. Domingo I, Chieli A. Characterizing the pigments and paints of prehistoric artists. Archaeol Anthropol Sci. 2021;13:196. https://doi.org/10.1007/s12520-021-01397-y.
11. Bayarri V, Latova J, Castillo E, Lasheras JA, De Las Heras C, Prada A. New documentation and of rock art analysis through multispectral techniques at Cueva de Altamira. (In Spanish: Nueva documentación y estudio del arte empleando técnicas hiperespectrales en la Cueva de Altamira). ARKEOS | perspectivas em diálogo, nº 37. In XIX International Rock Art Conference IFRAO 2015. Symbols in the Landscape: Rock Art and Its Context. Conference Proceedings; Instituto Terra e Memória: Tomar, Portugal. 2015;2293–2307.
12. Le Quellec J-LL, Harman J, Defrasne C, Duquesnoy FF, Defrasne C. DStretch® and the enhancement of digital images: applications to the Archaeology of rock images. (In French: DStretch® et l'amélioration des images numériques: applications à l'archéologie des images rupestres). Les Cahiers de l'AARS. 2013;16(2–3), 177–198.
13. Le Quellec JL, Duquesnoy F, Defrasne C. Digital image enhancement with DStretch®: is complexity always necessary for efficiency? Digit Appl Archaeol Cult Herit. 2015;2(2–3):55–67. https://doi.org/10.1016/j.daach.2015.01.003.
14. Quesada Martínez E, Harman J. A step further in rock art digital enhancements. DStretch on Gigapixel imaging. Digit Appl Archaeol Cult Herit. 2019;13:e00098. https://doi.org/10.1016/j.daach.2019.e00098.
15. Cerrillo-Cuenca E, Sepúlveda M, Guerrero-Bueno Z. Independent component analysis (ICA): a statistical approach to the analysis of superimposed rock paintings. J Archaeol Sci. 2021;125:105269. https://doi.org/10.1016/j.jas.2020.105269.
16. Monna F, Rolland T, Magail J, Esin Y, Bohard B, Allard A-C, Wilczek J, Chateau-Smith C. ERA: a new, fast, machine learning-based software to document rock paintings. J Cult Herit. 2022;58:91–101. https://doi.org/10.1016/j.culher.2022.09.018.

17. Yihao F, Yue C, Jun W, Xiaoyu Z, Lu L, Baheti Z, Peng J. Secrets on the rock: analysis and discussion of the Dunde Bulaq rock art site. Herit Sci. 2024;12:38. https://doi.org/10.1186/s40494-024-01146-z.

18. Barbuti P, Fuentes O, Konik S, Pinçon G. Le relevé interdisciplinaire d'art pariétal paléolithique en trois dimensions: intérêt, méthode et premiers résultats. Humanités Numériques. 2023. https://doi.org/10.4000/revuehn.3410.

19. Fritz C, Willis MD, Tosello G. Reconstructing paleolithic cave art: the example of marsoulas cave (France). J Archaeol Sci Rep. 2016;10:910–6. https://doi.org/10.1016/j.jasrep.2016.05.012.

20. Murphy KP. Probabilistic machine learning: an introduction. Cambridge: MIT Press; 2022.

21. Horn C, Ivarsson O, Lindhé C, Potter R, Green A, Ling J. Artificial intelligence, 3D documentation, and rock art—approaching and reflecting on the automation of identification and classification of rock art images. J Archaeol Method Theory. 2022;29(1):188–213. https://doi.org/10.1007/s10816-021-09518-6.

22. Chalmin E, Hœrlé S, Reiche I. Taphonomy on the surface of the rock wall. In: David B, McNiven IJ, editors. The oxford handbook of the archaeology and anthropology of rock art. Oxford: Oxford University Press; 2017. p. 1–35.

23. Rowlands DA. Physics of digital photography (Second edi). Bristol: IOP Publishing; 2020.

24. Troncoso A, Moya F, Sepúlveda M, Carcamo JJ. First absolute dating of Andean hunter-gatherer rock art paintings from North Central Chile. Archaeol Anthropol Sci. 2015;2(9):223–32.

25. Cabello G, Gallardo F, Odone C. The coastal rock paintings of Chomache and its social and economic context (Tarapacá region, Northern Chile) (In Spanish: Las pinturas costeras de Chomache y su contexto económico-social (Región de Tarapacá, Norte de Chile)). Boletín Del Museo Chileno de Arte Precolombino. 2013;18(1):49–66. https://doi.org/10.4067/S0718-68942013000100004.

26. Núñez L, Briones L. Assessment of Chomache's painting on the arid coast of Atacama desert (Northern Chile) (In Spanish: Valoración de las pinturas de Chomache en la costa árida del desierto de Atacama (Norte de Chile)). Chungará. 2021;53(2):261–82. https://doi.org/10.4067/S0717-73562021005000802.

27. Kuehni RG. Color space and its divisions. Color Res Appl. 2001;26(3):209–22. https://doi.org/10.1002/col.1018.

28. Malacara D. Color vision and colorimetry. Theory and applications. 2nd ed. Bellingham: SPIE; 2011.

29. Bradski G. The OpenCV Library. Dr. Dobb's Journal of Software Tools. 2000.

30. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, van Kerkwijk MH, Brett M, Haldane A, del Río JF, Wiebe M, Peterson P, Oliphant TE. Array programming with NumPy. Nature. 2020;585(7825):357–62. https://doi.org/10.1038/s41586-020-2649-2.

31. Hyvarinen, A. Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans Neural Netw. 1999;10(3):626–34. https://doi.org/10.1109/72.761722.

32. Cerrillo-Cuenca E, Ortiz-Coder P, Martínez-del-Pozo JÁ. Computer vision methods and rock art: towards a digital detection of pigments. Archaeol Anthropol Sci. 2014;6(3):227–39. https://doi.org/10.1007/s12520-013-0147-2.

33. Rogerio-Candelera MÁ. Digital image analysis based study, recording, and protection of painted rock art. Some Iberian experiences. Digit Appl Archaeol Cult Herit. 2015;2(2–3):68–78. https://doi.org/10.1016/j.daach.2014.11.001.

34. Carrión-Ruiz B, Lerma JL. Análisis de componentes principales de imágenes multiespectrales en el ámbito del arte rupestre. Proc 1st Congress Geomatics Eng CIGeo. 2017. https://doi.org/10.4995/CIGeo2017.2017.6597.

35. Cerrillo-Cuenca E, Sepúlveda M. An assessment of methods for the digital enhancement of rock paintings: the rock art from the precordillera of Arica (Chile) as a case study. J Archaeol Sci. 2015;55:197–208. https://doi.org/10.1016/j.jas.2015.01.006.

36. Domingo I, Carrión B, Blanco S, Lerma JL. Evaluating conventional and advanced visible image enhancement solutions to produce digital tracings at el Carche rock art shelter. Digit Appl Archaeol Cult Herit. 2015;2(2–3):79–88. https://doi.org/10.1016/j.daach.2015.01.001.

37. Vicent García JM, Montero Ruíz I, Rodríguez Alcalde ÁL, Martínez Navarrete MI, Chapa Brunet T, Brunet TC. The application of multispectral imagery to the study and conservation of post-palaeolithic rock art (In Spanish: Aplicación de la imagen multiespectral al estudio y conservación del arte rupestre postpaleolítico). Trab Prehist. 1996;53(2):19–35. https://doi.org/10.3989/tp.1996.v53.i2.390.

38. Hyvärinen A, Oja E. Independent component analysis: algorithms and applications. Neural Netw. 2000;13(4–5):411–30. https://doi.org/10.1016/S0893-6080(00)00026-5.

39. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay E. Scikit-learn: machine learning in python. JMLR. 2011;12:2825–30.

40. Duda T, Canty M. Unsupervised classification of satellite imagery: choosing a good algorithm. Int J Remote Sens. 2002;23(11):2193–212. https://doi.org/10.1080/01431160110078467.

41. Göhring A, Grupe G, Kröger P. Application of gaussian mixture model clustering for the resolution of multi-isotope fingerprints. In: Douglas Price T, editor. Isotopic proveniencing and mobility: the current state of research. Cham: Springer International Publishing; 2023. p. 199–239.

42. Kvapil LA, Kimpel MW, Jayasekare RR, Shelton K. Using Gaussian mixture model clustering to explore morphology and standardized production of ceramic vessels: a case study of pottery from Late Bronze Age Greece. J Archaeol Sci Rep. 2022;45:103543. https://doi.org/10.1016/j.jasrep.2022.103543.

43. Pan L, Li Y, He K, Li Y, Li Y. Generalized linear mixed models with gaussian mixture random effects: inference and application. J Multivar Anal. 2020;175:104555. https://doi.org/10.1016/j.jmva.2019.104555.

44. Corduneanu A, Bishop CM. 2001. Variational Bayesian Model Selection for Mixture Distributions. In: William A. Gale (Eds). Artificial intelligence and Statistics. 27–34.

45. McGrory CA, Titterington DM. Variational approximations in Bayesian model selection for finite mixture distributions. Comput Stat Data Anal. 2007;51(11):5352–67. https://doi.org/10.1016/j.csda.2006.07.020.

46. Van Der Walt S, Schönberger JL, Nunez-Iglesias J, Boulogne F, Warner JD, Yager N, Gouillart E, Yu T. Scikit-image: Image processing in python. PeerJ. 2014;2:e453. https://doi.org/10.7717/peerj.453.

47. Andrews TD, Brink JW. Using retroReveal as a complement to DStretch for enhancing red ochre pictographs. Canadian J Archaeol. 2022;46(1):1–15. https://doi.org/10.51270/46.1.1.

48. De Marrais E, Robb y J. Art makes society: an introductory visual essay. World Art. 2013;3(1):3–22. https://doi.org/10.1080/21500894.2013.78233.

49. Moro-Abadía O, González M. Art in the making: recent developments in the study of Pleistocene and Holocene images. J Archaeol Method Theory. 2020;27(3):439–53. https://doi.org/10.1007/s10816-020-09479-2.

50. Quesada M, Gheco L. Times, caves and paintings. Reflections on the polychromy of rock art at Oyola (Catamarca province, Argentina) (In Spanish: Tiempos, cuevas y pinturas. Reflexiones sobre la policromía del arte rupestre de Oyola (Provincia de Catamarca, Argentina)). Relaciones, Revista de la Sociedad Argentina de Antropología XL. 2015;2:455–76.

51. Troncoso A, Armstrong F, Vergara F, Ivanovic F, Urzúa P. Nurturing and balancing the world: a relational approach to rock art and technology from north central Chile (Southern Andes). Cambridge Archaeol J. 2020;30:239–55. https://doi.org/10.1017/S0959774319000581.

52. Gallardo F, Cabello G, Sepúlveda M, Ballester B, Fiore D, Prieto A. Yendegaia Rockshelter, the first rock art site on Tierra del Fuego island and social interaction in southern Patagonia (South America). Lat Am Antiq. 2022;34(3):532–49.

53. Bastías F, Cabello G, Gallardo F. Painted rocks on the Loa river mouth (Atacama Desert, Chile) (In Spanish: Piedras pintadas en la desembocadura del río Loa (desierto de Atacama, Chile)). CUHSO. 2023;33:63–94.

Publisher's Note