



OPEN Using digital archaeology and machine learning to determine sex in finger flutings

Andrea Jalandoni^{1,3}✉, Robert Haubt^{2,3}, Calum Farrar¹, Gervase Tuxworth¹,
Zhongyi Zhang¹, Keryn Walshe⁵ & April Nowell¹⁶

One of the earliest and most enigmatic forms of rock art are finger flutings and previous methods of studying them relied on biometric finger ratios from modern populations to make assumptions about the people who left the flutings, which is theoretically and methodologically problematic. This work is a proof-of-concept for a paradigm shift away from error-prone human measurements and controversial theories to computational digital archaeology methods for an innovative experimental design using a tactile, virtual, and machine learning approach. We propose a digital archaeology experiment using a tactile and virtual approach based on multiple samples from 96 participants. We trained a machine learning model on the known data to determine the sex of the person who made the fluting. While the virtual dataset did not provide sufficiently distinct features for reliable sex classification, the tactile experiment results showed potential for the identification of the sex of fluting artists, but more samples are needed to make any generalization. The significant contribution of this study is the development of a foundational set of methods and materials. We provide a novel digital archaeology approach for data creation, data collection, and analysis that makes the experiment replicable, scalable, and quantifiable.

Keywords Virtual reality, Experimental archaeology, Computational archaeology, Rock art, Digital tracing

Artistic expression can be a powerful means of exploring how early humans understood the world around them and how they engaged with their environments. Among the earliest forms of art, finger flutings (Fig. 1) (also known as digital tracings) — offer a window into the cognitive and cultural practices of prehistoric societies. These distinctive markings, made by pressing or scraping fingers against soft sediment lining the walls, ceilings and floors of limestone caves are found at sites across Western Europe and Australia during the late Middle to Upper Paleolithic period, ca. 60,000–12,000 years before present (BP). Not only are finger flutings one of the earliest types of art associated with *Homo sapiens* but they are also one of very few types of art that was clearly made by them and Neandertals¹.

Flutings have the potential to reveal information about age, sex, height, handedness and idiosyncratic mark-making choices among unique individuals who form part of larger communities of practice³. However, previous methods for making any determination about the individual artist from finger flutings have been shown to be unreliable⁴. Accordingly, we propose a novel digital archaeology approach to begin understanding this enigmatic form of rock art by leveraging machine learning (ML) as a tool for uncovering patterns from two datasets, one tactile and one virtual, collected from a modern population. We aimed to determine whether ML can reveal subtle differences in the sex of the artist based on their finger-fluted images.

The results of this study are significant because sex is one variable that can be used to study how identities were constructed in the past. Its intersectionality with other variables such as age or gender allows archaeologists to identify what might have been meaningful social categories in ancient societies. Further, in the history of archaeology, women's roles in society including in the production of art were often understudied^{5,7}. The method proposed here, if successfully applied to the archaeological record, could be a means of rendering women more visible in the past with concomitant implications for how women are viewed today^{5,8,9}. There has been a long history of research using morphometrics for rock art classification^{10–15}. Recent work on using geometric

¹Place, Evolution, Rock Art Heritage Unit, Griffith Centre for Social and Cultural Research, Griffith University, Gold Coast, QLD 4215, Australia. ²The School of Computer Science & Information Technology, SAE University College, Brisbane, Australia. ³Australian Research Centre for Human Evolution, Griffith University, Nathan, QLD 4111, Australia. ⁴School of Information and Communication Technology, Griffith University, Nathan, QLD 4111, Australia. ⁵Faculty of Arts, Anthropology, University of Auckland, Auckland, New Zealand. ⁶Department of Anthropology, University of Victoria, 3800 Finnerty Rd, Victoria, BC V8P 5C2, Canada. ✉email: a.jalandoni@griffith.edu.au



Fig. 1. Finger flutings from Koonalda Cave, Australia (adapted from)².

morphometrics to classify age and sex in hand stencils, demonstrates the potential and current trends in the use of algorithmic methods for rock art analysis¹⁶. Our interdisciplinary approach combines experimental archaeology and machine learning, opening new avenues for understanding prehistoric art-making processes and human behavior. This paper provides a first step towards understanding the potential of ML for analysing finger flutings as a proof-of-concept model that would need refinement to be applied to ancient sites with potentially different physical characteristics.

Literature review

Prehistoric finger flutings

Finger flutings are impressions created by dragging one or more fingers across a soft, compactable surface such as moonmilk, a calcium carbonate that covers the floors, walls and ceilings of some limestone caves. Historically, these markings were misinterpreted as “parasite lines” (i.e., lines that detracted from “real art”) or the result of animal activity rather than human interaction^{3,17}. By the 1960s, research shifted to confirming their anthropogenic origins and discussing their cultural significance^{18–22}. These markings may have held symbolic or ritualistic significance, possibly related to early forms of communication or shamanistic practices, connecting humans to the spiritual or supernatural world^{23,24}. Similarly, the work of Clottes on cave art explores the concept of art for ritual purposes, proposing that finger flutings were integral to the sensory and experiential nature of prehistoric art²⁵. The study of finger flutings initially aimed to affirm their human origin, focusing on patterns or repetitive sequences that might signify proto-language or mnemonic devices^{20,21,26–29}. These investigations were influenced by Marshack’s work on symbolic marks and Jungian interpretations of psychograms³⁰. While earlier interpretations emphasized codes and symbolic meaning, later research suggested these flutings might also represent playful or exploratory activities by children³¹.

The role of children in creating finger flutings became a significant focus, particularly following Bednarik’s hypothesis that children contributed to Paleolithic art³². This theory had been supported by experimental techniques correlating fluting width with age^{28,33,34} and continues to be cited³⁵. However, this approach has been shown to be unreliable⁴. A recent study of finger flutings in Koonalda Cave, a > 30,000-year-old site in southern Australia has looked to ethnographic data to understand tracings in a specifically Australian context arguing that the repetitive motif of the tracings at Koonalda is most similar to markings created in the context of propagation ceremonies^{36,37}. This relationship between finger flutings and ceremony has recently been further affirmed in the Australian context with a clear link identified between archaeological evidence at a Victorian finger fluting site and local oral histories²⁴. These studies underline the potential of finger flutings to offer insights into symbolic and cultural practices of early to contemporary humans and neanderthals. Further, the study of sex in prehistoric art creation, particularly finger flutings, raises intriguing questions about the role of physical attributes in artistic expression and the relationships between identity and cultural practice.

Previous methods

Attempts to determine the sex of the makers of Paleolithic art have focused primarily on two categories of mark making—(1) finger flutings and (2) hand stencils and handprints. A third method of using fingerprint analysis remains novel within the literature^{38,39}. A common approach to finger flutings and hand stencils is the application of the 2D:4D ratio. This ratio describes the relationship between the length of a person’s second digit (or index finger) and their fourth digit (or ring finger). The 2D:4D ratio is predetermined in utero through exposure to estrogen and testosterone. Ratios of less than 1.0 (i.e., the index finger is shorter than the ring finger) reflect greater testosterone exposure and are said to be characteristic of males while ratios greater than 1.0 are described as female⁴⁰. This ratio has been applied to prehistoric finger flutings in cases where the tips of the middle three fingers of either hand could be determined⁴¹. However, variability due to the pressure applied when fluting, arm height, palm/wrist angle relative to the fluting surface, and humidity of fluting matrix in conjunction with the fact that the flutings tend to widen over time²² mean that this method cannot be used to determine the sex of fluters with any degree of accuracy⁴.

The ratio has also been applied to handprints (when a hand is dipped in pigment and pressed against a cave wall) and hand stencils (when pigment is blown around the hand leaving a negative imprint of it). In these cases, experimental studies using North American subjects found their samples masculinized (i.e., both males and females patterned as males)³⁰. Some researchers have had greater success using additional ratios (3D and 4D) or other morphological data in conjunction with size measurements even though hand size between males and females can overlap by as much as 85%^{42–44}. However, it should be noted that neither handprints nor hand stencils are a precise reflection of the soft tissue hand. For example, applying pressure with the palm will often make fingertips of handprints “invisible” while the height/angle at which pigment is blown around a hand will introduce error to a hand stencil. Other factors such as the natural topography of a cave wall and the level of expertise/motor control of the mark maker can also introduce error⁴⁰.

Digital archaeology: virtual reality & machine learning

The application of Virtual Reality (VR) and machine learning (ML) in archaeology has grown in recent years, offering promising new tools for analyzing ancient artifacts, human remains, and cultural practices. Virtual Reality has been of use to archaeologists since the 1990s as an immersive research communication tool^{45–47} and has increasingly been used as a platform for experimental archaeology, including experiential analysis of rock art⁴⁸ but continues to be on the periphery of archaeological practice^{49–51}. The proliferation of VR technology in recent years has further improved the fidelity and accessibility of VR as a platform for experimental archaeology making it one that is both engaging for participants and a productive research tool.

Machine learning algorithms, particularly Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), have been increasingly used to analyze patterns in archaeological data, such as the classification of artifacts, sex determination from skeletal remains⁵², or gender biases in cultural heritage catalogues⁵³. In rock art research, there has been some progress towards using ML to detect rock art but also to classify rock art motifs and styles^{54–57}.

A growing area of interest is the integration of tactile and motion-based data for understanding human behavior in prehistoric contexts. ML's potential to identify subtle, individualized features in human motion patterns has been explored: ML was used to analyze fingerprint data to determine the identity of individuals⁵⁸; and the ability of ML to predict demographic attributes from biometric data such as fingerprints was explored⁵⁹. These studies suggest that ML, when applied to data sets like flutings, can identify patterns that may not be immediately visible to the human eye. However, ML's integration into archaeological research remains in its infancy, particularly in terms of dealing with complex, multimodal datasets like tactile and VR data. The use of VR in cultural heritage highlights both the potential and challenges of using virtual environments to simulate and analyze archaeological information⁶⁰.

Methodology

Experimental design

Our study consisted of two approaches: a tactile and a virtual reality (VR) experience that collected flutings from a modern population. The data were used to train and test a ML model that could classify finger flutings based on biometric attributes. The aim was to test if these approaches could be used to provide information about the artists.

Participant sampling and data collection

Ninety-six participants volunteered to contribute both tactile and virtual finger fluting data. Data collection was conducted in 2024 at the Australian Archaeological Association Conference, Griffith University, and SAE University College in the Gold Coast and Brisbane, Australia. There were no predetermined criteria for sex or height, but individuals were required to be over 18 years old. This age restriction introduced a bias, but as this was a pilot project we wanted to limit the scope and not further complicate the dataset with the variability of children's hand size. Further, the sample is biased toward a demographic of those who attend higher education and Australian archaeology conferences⁶¹. The definition of sex used in our study is a binary sex categorization (male/female) which was self-reported by participants. While this binary does not reflect the diversity of biological sex or gender, for the sake of the methodology participants were limited to this binary choice. The obvious bias is that our dataset is from a modern population collected from two universities and an Australian Archaeology Conference, which might have skewed the results towards the dominant demographics of these venues (e.g. “white”, Australian, women, and well-educated). However, this bias does not affect the validity of our study.

The physical attributes of each participant—hand measurements (palm width, 3 finger width, hand span), age, sex, handedness (left, right, ambidextrous), and height—were recorded. This data served as a foundation for analyzing the experimental outcomes. All methods were carried out in accordance with relevant guidelines and regulations. Ethical clearance was obtained through Griffith University and identified as protocol number 2023/667. Informed consent was obtained from all subjects. Participants were provided with a project information form which identified potential risks to both themselves and their data before being asked for written consent.

Instructions for participants

The instructions for the fluting experiment were designed to capture a wide range of flutting motions that prehistoric artists may have used (Fig. 2). This included specific combinations of hand motions, such as forehand and backhand strokes, to simulate possible different techniques used for flutting. Since finger flutings are an esoteric form of rock art, we attempted to give participants a better understanding by providing a printed 3D model of finger flutings from a prehistoric cave to see and touch.

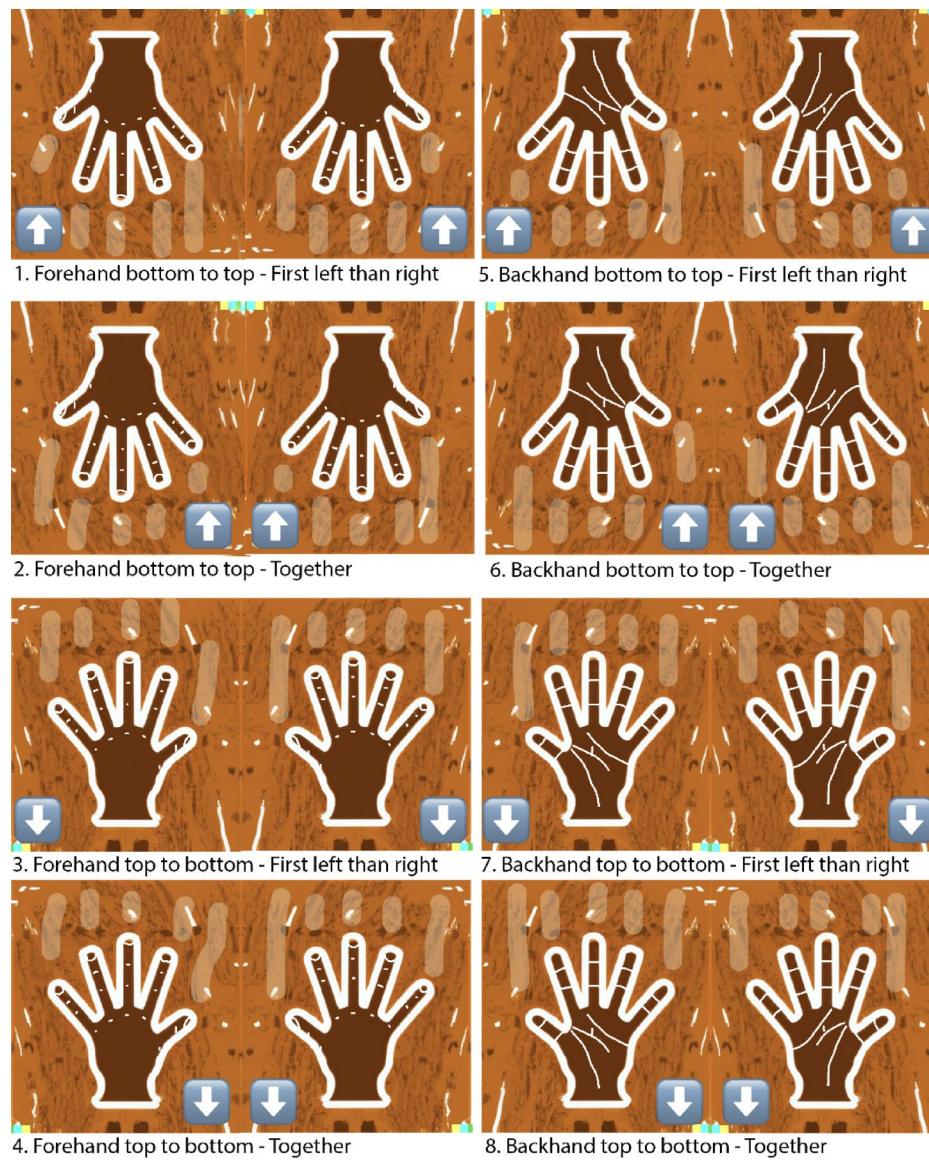


Fig. 2. Instructions for participants of tactile approach.

Tactile approach

The tactile experiment installation: moonmilk simulacra Due to the unavailability of large quantities of moonmilk, a substitute material was sought. The key criteria for the substitute were:

- **Structural Integrity:** The material needed to maintain its form during and after the fluting process.
- **Adherence:** It had to stick to a vertically erected canvas, simulating a cave wall.
- **Texture:** Similar look and feel to moonmilk, leaving a similar imprint (fluting).
- **Resetting:** The canvas needed to be reset between flutes to facilitate hundreds of data points (images).

Previous experimental studies utilised various mediums (e.g., plaster of Paris, finger paints and clay) to simulate finger-fluting creation^{33,34}. After testing the different materials previously used, it was clear that none of these fit the criteria. Therefore, a substitute material was developed in consultation with Danielle Clarke, a master potter (Appendix A). This material was designed to replicate moonmilk's texture and properties. The substitute was applied to a canvas approximately 5 cm deep with an effective drawing area of 86 cm by 56 cm accounting for the frame. The frame was mounted on an easel with the top of the frame reaching a height of 175 cm above the ground.

Fluting and image capture Each participant was asked to perform eight predefined flutings based on structured instructions, followed by one freehand fluting (Figs. 2 and 3). The instructions included a set of forehand and backhand motions in different sequences (one hand a time or together), designed to cover a range of possible flutting techniques. Flutings were mostly captured using a Panasonic DC-GH5 camera mounted on a tripod to



Fig. 3. Tactile (left) and virtual (right) data collection.

ensure high-quality, consistent images (Appendix B). At times there were issues with the camera and a Samsung Flip 5 was used to capture the images, in order not to delay participants. Notes were taken on observations made about some participant's behaviour and stance.

Virtual reality approach

Data collection through a bespoke Virtual Reality (VR) program was pursued for two reasons. First, it would provide a consistent experimental medium and environment between participants as well as producing a well-controlled data output in the form of born-digital images. Secondly, the VR platform allowed for multiple other kinds of data to be gathered unobtrusively and inexpensively such as finger, hand and head positions. The program was designed for the Meta Quest 3, which at the time had the most affordable and accurate consumer grade inbuilt hand tracking system. Furthermore, the Meta Quest 3 continues to be well supported for independent development allowing easy use of bespoke software on the hardware.

The primary functional requirements of the VR approach were to allow users to create virtual finger flutings with natural hand movements and to save each finger fluting to local/network storage. Secondary to this was the creation of a user interface (UI) which could independently instruct and guide users through the experiment. The Unity Game Engine was chosen for its range of both official and unofficial VR support and integration. OpenXR was used to manage VR integration as it provided better support for required add-ons and simpler customisation of tracked hand skeletons compared to the Oculus plugin, the officially supported integration plugin for the Meta Quest 3. Additionally, OpenXR allows for greater interoperability with other VR platforms if desired in the future.

Both primary requirements were largely met using the add-on *Drawing Board VR* available from the Unity Store, which provides assets and scripts to draw virtually and to save the images. Although designed for PCVR, it proved completely functional on the standalone Meta Quest 3. Adapting the prefab assets from Drawing Board VR was trivial, capsules were attached to the hand skeleton anchor points (see Fig. 4) with each capsule having the ability to leave an impression on the virtual board. After some testing and manual adjustments, the capsules accurately translated real world finger position and angles. The virtual board was scaled to match the tactile board (frame size 90 cm by 60 cm).

To avoid having participants alternate between hand tracking and controller interaction with UI elements, all UI was interactive through *poke* or *raycast* interactions created by hand pose and movement. Participants were prompted to press buttons to move through the experiment (Fig. 5). It was anticipated that many participants would have no prior VR experience, particularly with hand tracking exclusive interaction, therefore, the initial instructions acted to allow participants to become familiar with the feel of the UI interaction. This integrated tutorial approach was further developed by having a test board which users could finger flute on without it being recorded. When ready, the users were instructed to press the large start button on their left which would then start the recorded experiment. Each of the nine instructions came with textual and animated instructions.

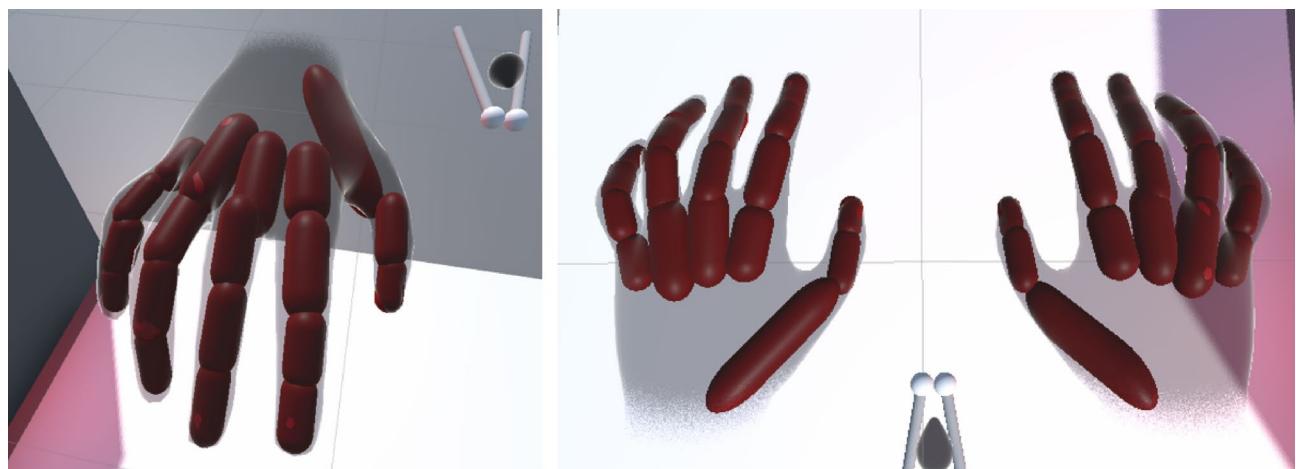


Fig. 4. An early development version of the drawing capsules attached to hand skeleton anchors. Not visible in the final version.

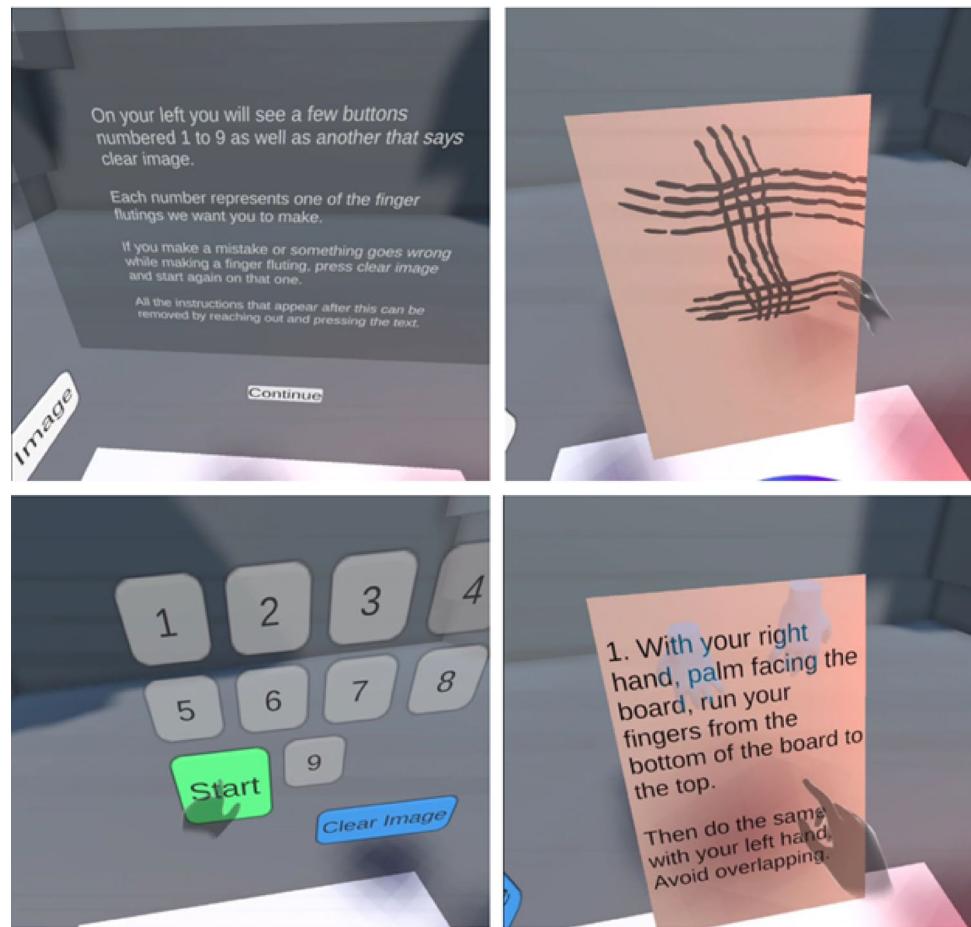


Fig. 5. (Top left) Introductory instructions, (bottom left) Image saving UI and start button, (top right) The virtual test board, (bottom right) Animated phantom hands demonstrating the hand movements wanted from the participants. (Note: High angle point-of-view does not reflect typical participant view, head position of user in demonstration images was ~ 2 m).

The animated instructions, which were phantom hands demonstrating the desired movement, were added to avoid any confusion that might arise from the wording of the text. Following the completion of an instruction, users were directed to press the corresponding number to their left, all other numbers were disabled to avoid misselection. Participants performed the same eight predefined flutings and one freehand fluting in the VR environment as they did in the tactile experiment.

The images were saved locally onto the Quest 3 at a resolution of 4096p x 6144p and were manually transferred via link-cable connected to a laptop during the experiment. In practice, this proximity to the participant allowed them to be actively monitored for both safety and guidance throughout the experiment. A limitation of this approach was the lack of tactile feedback for participants as real-world objects could not be effectively and consistently tracked into the virtual environment to provide tactility. To partly mitigate the lack of tactile feedback, pseudo-depth was added to the virtual canvas, meaning users' hands could only sink approximately 5 cm into the virtual board, mimicking the tactile experiment. This meant that rather than users' virtual fingers skimming the surface of the virtual board, they were able to sink them in and drag them across. For detailed instructions on how to recreate the virtual experiment and a link to the github repository, see Appendix C.

Dataset curation

Data from both the tactile and VR experiments were processed for analysis and used to train neural networks designed to identify correlations between participants' physical attributes and their fluting techniques.

VR images did not require manual cropping, as they did not contain redundant backgrounds. Tactile images underwent a semi-automated process to remove the background using the segmentation model SAM2⁶². SAM2 employs click points as input prompts to guide the segmentation process. When initial segmentation results were suboptimal, additional click points were applied iteratively to refine the output. Following automated segmentation, all SAM2-segmented images were manually reviewed to ensure the complete removal of personal information while preserving the integrity of the primary's content. In rare cases (approximately 5%), when SAM2 failed to achieve satisfactory segmentation, manual cropping was performed as an alternative measure.

The dataset consisted of both virtual (63 female, 29 male) and tactile (56 female, 23 male) images. To maximize data utility, the dataset was split into training and test sets in an 8:2 ratio at the individual level, ensuring that no participant appeared in both sets. For virtual images, the training set included 666 images (463 female, 203 male), while the test set contained 152 images (108 female, 44 male). For tactile images, the training set comprised 573 images (411 female, 162 male), with the test set consisting of 126 images (90 female, 36 male).

Machine learning approach

We employed two deep learning models, ResNet-18⁶³ and EfficientNet-V2-S⁶⁴, due to their strong classification performance and relatively small parameter counts, making them well-suited for the dataset's limited size. ResNet-18 is a lightweight convolutional neural network (CNN) architecture from the ResNet family, consisting of 18 layers. Its residual learning framework enhances feature extraction while mitigating vanishing gradient issues, making it particularly effective for smaller datasets. EfficientNet-V2-S is a more recent CNN model designed to optimize both computational efficiency and classification accuracy. Compared to ResNet-18, EfficientNet-V2-S provides enhanced feature representation with fewer parameters, making it a robust choice for classification tasks involving limited data availability.

Model training was conducted on a Linux workstation (Ubuntu 18.04) with an NVIDIA RTX 3060 GPU using PyTorch. To determine the optimal learning rate, two different training settings were applied: one with 200 epochs at a learning rate of 1×10^{-5} and another with 1000 epochs at a reduced learning rate of 2×10^{-6} . Input images were automatically resized to match the pretrained model requirements, with ResNet-18 using 224×224 pixels and EfficientNet-V2-S using 384×384 pixels.

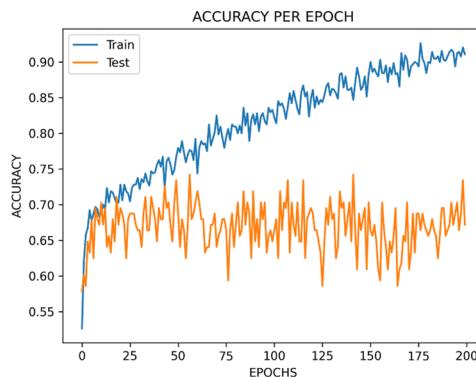
For training, a batch size of 32 was used for ResNet-18, whereas EfficientNet-V2-S was trained with a batch size of 16. Input images were normalized using the ImageNet mean [0.485, 0.456, 0.406] and standard deviation [0.229, 0.224, 0.225] to align with the pretrained model input distributions. To enhance model robustness and generalization, data augmentation techniques were applied, including random rotation ($\pm 10^\circ$), horizontal and vertical flipping ($p=0.5$), perspective distortion (scale=0.6), and Gaussian blur (kernel size=5 \times 9, $\sigma=0.1-5$). Model accuracy was computed as the ratio of correctly predicted instances to the total number of predictions. The code is publicly accessible at <https://github.com/johnnydfci/FingerFluting-SexClassification>.

Statistical analysis

Model performance was evaluated using three key metrics: Area Under the Curve (AUC), accuracy, and F1 score. AUC was calculated to assess the model's ability to distinguish between male and female-generated finger flutting patterns, with higher values indicating better discrimination. Accuracy was defined as the proportion of correctly classified samples among all predictions, providing an overall performance assessment. F1 score, which balances precision and recall, was used to quantify classification reliability, particularly in handling class imbalances.

Performance metrics were computed for each model configuration, including different learning rates (1×10^{-5} and 2×10^{-6}) and architectures (ResNet-18 and EfficientNet-V2-S). The statistical significance of differences in model performance across learning rates and architectures was analyzed to determine the optimal training configuration.

ResNet-18

Learning rate
 1×10^{-5} 

EfficientNet-V2-S

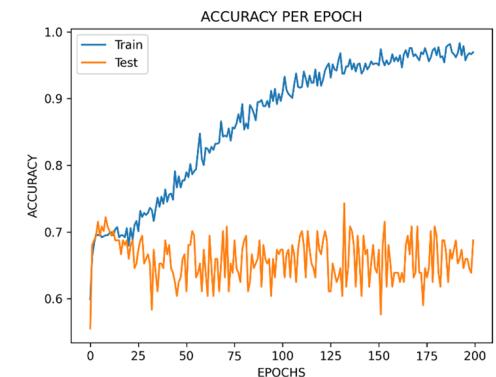
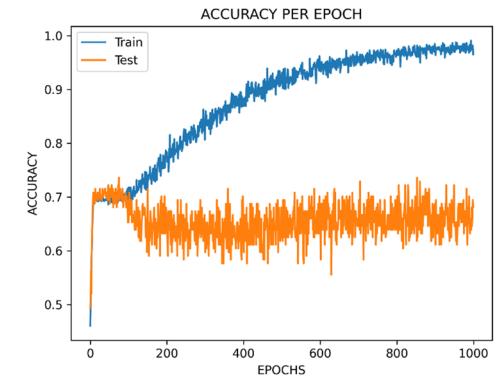
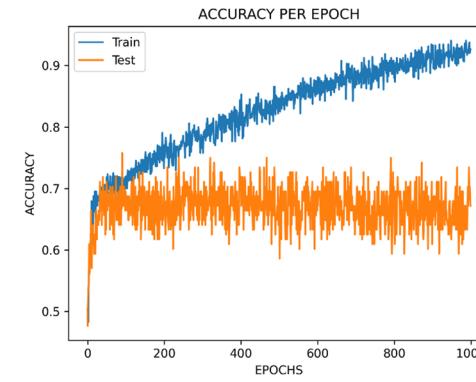
Learning rate
 2×10^{-6} 

Fig. 6. Training and testing accuracy of ResNet-18 and EfficientNet-V2-S on virtual images using two different learning rates (1×10^{-5} and 2×10^{-6}). The accuracy of both the training and test sets is plotted throughout the training process. Training was conducted for 200 epochs at a learning rate of 1×10^{-5} and 1000 epochs at 2×10^{-6} . The Y-axis represents accuracy, calculated as the number of correct predictions divided by the total number of predictions.

Model	Learning Rate	Accuracy	AUC	F1 Score
EfficientNet	2×10^{-6}	0.736	0.6566	0.0417
ResNet	2×10^{-6}	0.758	0.6156	0
EfficientNet	1×10^{-5}	0.743	0.7051	0.5289
ResNet	1×10^{-5}	0.742	0.5683	0.1644

Table 1. Performance metrics of ResNet-18 and EfficientNet-V2-S trained on virtual images under two different learning rates. AUC = Area under the Curve.

Results

Neural network training on the virtual images

We trained two deep learning models, ResNet-18 and EfficientNet-V2-S, using two different learning rates (1×10^{-5} and 2×10^{-6}), resulting in four separate training conditions (Fig. 6). The model weights achieving the highest accuracy were selected for AUC and F1 score calculation, with the results presented in Table 1.

ResNet-18, despite achieving the highest overall accuracy (0.758) at a lower learning rate (2×10^{-6}), struggled with AUC (0.6156) and yielded an F1 score of 0, indicating poor classification balance. The same model at a higher learning rate (1×10^{-5}) exhibited a slight drop in accuracy (0.742) but an increase in F1 score (0.1644), reflecting marginal improvements in handling class imbalances. EfficientNet-V2-S demonstrated more consistent performance across learning rates, with a minimal drop in accuracy but notable improvements in AUC and F1 score at the higher learning rate.

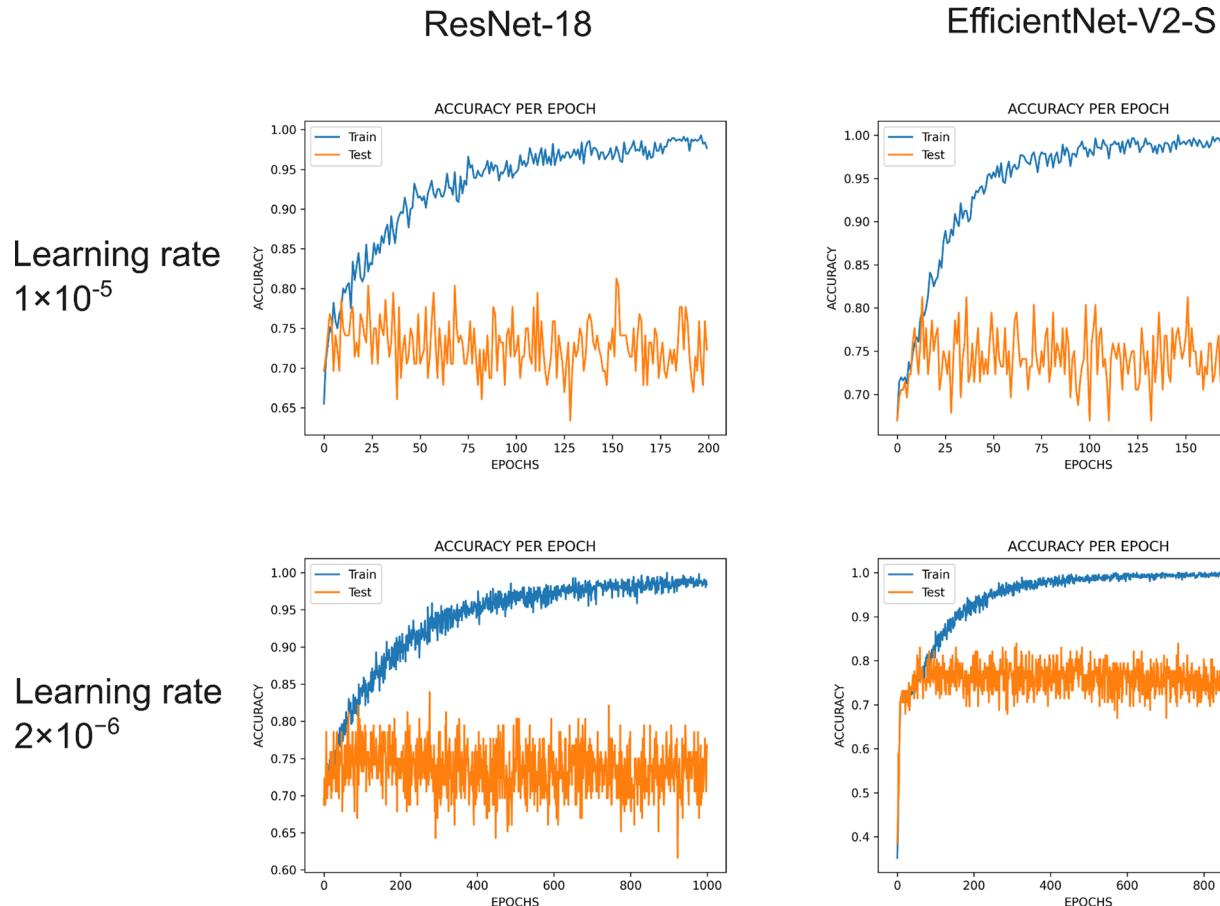


Fig. 7. Training and testing accuracy of ResNet-18 and EfficientNet-V2-S on tactile images using two different learning rates (1×10^{-5} and 2×10^{-6}). The accuracy of both the training and test sets is plotted throughout the training process. Training was conducted for 200 epochs at a learning rate of 1×10^{-5} and 1000 epochs at 2×10^{-6} . The Y-axis represents accuracy, calculated as the number of correct predictions divided by the total number of predictions.

model	Learning Rate	Accuracy	AUC	F1 Score
EfficientNet-V2-S	2×10^{-6}	0.839	0.7051	0.5289
ResNet-18	2×10^{-6}	0.839	0.8731	0.6087
EfficientNet-V2-S	1×10^{-5}	0.813	0.8667	0.4643
ResNet-18	1×10^{-5}	0.813	0.7892	0.5432

Table 2. Performance metrics of ResNet-18 and EfficientNet-V2-S trained on tactile images under two different learning rates. AUC = Area under the Curve.

Neural network training on the tactile images

Similar to the virtual image experiments, for the tactile experiments we trained two deep learning models, ResNet-18 and EfficientNet-V2-S, using two different learning rates (1×10^{-5} and 2×10^{-6}), resulting in four separate training conditions (Fig. 7). The model weights achieving the highest accuracy were selected for AUC and F1 score calculation, with the results presented in Table 2.

The results on the tactile image dataset were significantly better than the virtual image dataset. Both models achieved the highest accuracy (0.839) when trained with a lower learning rate (2×10^{-6}), with ResNet-18 demonstrating the highest AUC (0.8731) and F1 score (0.6087). EfficientNet-V2-S, while maintaining the same accuracy, showed a lower AUC (0.7051) and F1 score (0.5289). At the higher learning rate (1×10^{-5}), both models exhibited a slight drop in accuracy (0.813), with EfficientNet-V2-S attaining a higher AUC (0.8667) but a lower F1 score (0.4643) compared to its lower learning rate counterpart. ResNet-18, at the higher learning rate, had a decreased AUC (0.7892) but maintained an F1 score of 0.5432.

Incidental observations during the tactile approach

There were incidental observations made during the tactile approach that provided valuable insights into the finger fluting techniques and behavior of participants. These observations can be used to inform future avenues of exploration. Participants demonstrated a wide range of hand movement techniques. Most notably, participants exhibited different thumb placement techniques during the fluting process. Some did not involve their thumb, leaving only four marks on the surface, while others dragged the thumb across the canvas, which is not typical in cave flutings. Also, there was a noticeable distinction between the forehand and backhand techniques used. Forehand movements were typically more controlled and produced precise flutings, while backhand motions often resulted in broader, less defined strokes.

Finally, there was an obvious correlation between height and reach. Shorter participants faced specific challenges fluting backhand starting from below going upward, due to the top of the board being approximately 175 cm above the ground, often resulting in shorter flutings. The behaviour of the participants also influenced the markings. The position of a sample of participants' feet were noted to have had an influence on the symmetry of the flutings and the standing position impacted the direction and depth of the flutings. For example, those with one dominant foot forward tended to shift their weight, which in turn affected the final markings. Another example is that those who adopted a crouching position often produced deeper and more pronounced flutings, suggesting that posture influenced the final markings. The freehand flutings revealed a significant variation in artistic intent. While some participants focused on symmetry and precision, others leaned toward more abstract and expressive designs. This variation in creativity demonstrated the unique ways individuals interpreted the fluting task.

Discussion and future recommendations

Machine learning results

Overall, the deep learning models achieved high accuracy during training, with AUC values exceeding 0.85 for certain tactile image conditions. These results suggest that the models effectively learned patterns within the tactile dataset and demonstrated strong discrimination between male and female-generated finger fluting images. However, the relatively lower AUC values for virtual images, coupled with their unstable test accuracy, indicate that they do not provide sufficiently distinct features for reliable sex classification. This discrepancy highlights the greater robustness of tactile images over virtual images in capturing relevant classification features.

Despite the promising performance on tactile images, deep learning models exhibited a pronounced disparity between training and test performance. While training accuracy consistently increased, reaching near-perfect levels in the later epochs, test accuracy remained unstable and showed no substantial improvement over time. This pattern indicates overfitting, where the models effectively learn dataset-specific features but fail to generalize to unseen test data. The instability in test accuracy further suggests that the models struggle to extract robust and generalizable patterns from the finger fluting images, ultimately limiting their reliability for sex classification.

A possible contributing factor to this challenge could be individual variation in hand size and fluting characteristics. For example, some females may have larger hands and exhibit stronger fluting patterns resembling those of males, while some males may have smaller hands and display lighter, less pronounced fluting strength. This variability could confuse the model, making it difficult to accurately differentiate between sexes and ultimately hindering its performance on the test set.

These results underscore the critical need to increase the dataset size to alleviate overfitting and improve the model's generalizability. Moreover, the inherent variability in finger fluting images may impose fundamental limitations on the feasibility of using deep learning for sex classification, suggesting that alternative approaches or additional contextual data may be necessary to enhance classification accuracy.

The limited success of the tactile data in sex prediction underscores the importance of material-based approaches in understanding finger flutings. While the VR data failed to provide useful results, it opens up new and exciting possibilities for exploring the dynamic aspects of fluting and artistic intent in the future. While a modest achievement, this study highlights the potential of ML to enhance traditional archaeological methods.

Implications for finger fluting research

The traditional methods of using ratios described earlier in the literature review were flawed in their experimental method and the theory they were based on is contentious. For example, traditional methods when measuring had to be offset from the finger flutings to avoid damaging the rock art, introducing human error. Also, the ratios are not universally accepted because they are not consistent between modern populations nor proven applicable to ancient populations. In combination these issues cast doubt on the results of these traditional methods.

In contrast, our digital archaeology method addressed human error by introducing quantifiable methods through ML and the contention around ratios by adopting a theoretically agnostic approach. The ML model analyzed the photograph itself, not the physical characteristic of the hand. This is an important distinction between traditional methods that used hand measurements and inferred the results of the hand measurements onto the flutings, while our method simply classified the patterns in the photograph. Another advantage of the theoretically agnostic approach afforded ML is that it allows for the discovery of new theories that can be tested. The use of photography and computer vision as ways of remote sensing and measuring finger flutings makes the study scalable, replicable, and quantifiable, ultimately making it more robust than previous methods. Our study innovated in all aspects of the experimental design: the toolkit, the activity and the measurement.

Toolkit

As part of the tactile approach, an important contribution of this study is the recipe for a moonmilk substitute (Appendix A) that is a substantial improvement over the materials used in previous experiments with finger

flutings. Creating this simulacrum of moonmilk that can be easily replicated enables other researchers to undertake more realistic tactile finger fluting experiments.

This is the first known attempt of collecting finger fluting data through VR and the first use of ML to analyze finger flutings. The VR approach provides a convenient experimental environment, allowing it to be infinitely replicable. Furthermore, it has the ability to control, monitor and measure all aspects of the experiment. This multidimensionality produces rich observational data that is accurately and consistently recorded. However, it lacks fundamental realistic elements which are present in the tactile experiment.

Furthermore, we designed the finger fluting instructions used in both approaches to encourage different hand and body movements. While lacking in previous publications, this study produced a baseline instructional toolkit for finger flutings that is scalable and reproducible and can be used and improved upon by future researchers. Lastly, we developed a machine learning pipeline for finger fluting data that is made available on Github: <https://github.com/johnnydfci/FingerFluting-SexClassification>.

Activity

The design of the tactile and VR approach allowed for observations of modern populations' flutings. This provided insights into body balance, foot placements, reach, use of thumb etc., which previous studies may have noticed but did not publish. We also created a novel VR experiment space, while not successful, it has revealed other forms of data which can be captured during the experiment, such as exact finger, hand and arm positions throughout the activity. Furthermore, more VR data can be collected, for example where the participant looks (eye-gaze), providing further insight into the subtleties of the activity might be improved by future studies.

Analysis

Previous methods relied upon human judgement, which could have introduced variability in techniques, interpretation, or inherent biases. The experiments were often not designed to be agile enough to accommodate any other questions or to be expanded on. Furthermore, these methods often made assumptions that particular measurements were relevant to understanding physical characteristics.

In contrast, our method surpasses human capabilities and can uncover subtle unnoticed distinctions. Furthermore, it does not rely on the assumption that the measurements have a specific relationship to the artist's attributes; rather we are using machine learning with a computer vision approach that was not trained on these previous methods. Ideally, it would treat all potential avenues as equal initially, however, because we used transfer learning there may be residual biases. But these biases are different from human judgement biases and are computational and measurable. An example is the overfitting in our results where the models may have learned dataset-specific features but did not translate to the test data. Therefore, our data-driven approach is not only reproducible and consistent but improves the overall accuracy of the analytical model applied, making any potential shortcomings measurable.

The greater potential of this ML method is scalability and efficiency by feeding more data into the model and testing its accuracy. The dataset is also agile and can easily be used for a variety of other applications. For example, our binary approach for male and female can easily be expanded to right or left handedness. Another example is that third party researchers can take our toolkit and test the 2D:4D ratio theory and other traditional methods in a more rigorous way. Our current dataset was insufficient but showed promise, which can be further tested by adding more data. We can adapt the method in the future, by for example, making small changes to the variables in the code, while continuing to reuse the original dataset. The expansion of this data is enabled by the replicable toolkit we have designed.

Limitations and challenges, insights and potential

A major limitation is our sample size. Ninety-six participants producing 699 tactile data points and 818 virtual data points was not sufficient to make a definite determination of sex. Additionally, the lack of external validation further constrains the generalizability of the findings. At present, our model was trained and evaluated using data from a single center. While this provides internal validation, it may not fully reflect how the model performs on data from different imaging centers and populations, even when following similar photography standards.

In machine learning for image classification, a strong model is typically expected to also be validated on external datasets — for example, images collected from another center under similar standards. This helps demonstrate that the model's accuracy is stable and not simply the result of overfitting. Here, overfitting means the model learns patterns that are too specific to the training data. These patterns may include noise or unique characteristics, such as the lighting setup, camera settings, or background features in photographs, rather than the actual finger fluting patterns we aim to identify. As a result, an overfitted model performs well on the training data but poorly on unseen data. There are many cases showing that accuracy drops when moving from internal to external validation, even under similar photography standards. Therefore, including external validation is generally considered a more rigorous evaluation of model generalizability.

Another limitation of this experiment was that it did not intend to capture the environmental context. For example, fluting on a canvas is inherently different from fluting on a cave wall. Other differences include the moonmilk substitute, the humidity of the cave, and the lighting. This could also shed more light on the cultural context or intent of finger flutings, which was absent from this experiment.

The instructions were limited to eight vertical movements, which do not reflect the real-world range of finger flutings. Future experiments may need to include superimposition, more varied hand and arm movements, and body positions. Participants may have been influenced by observing other participants, which needs to be controlled in future experiments. For example, some participants were gouging the moonmilk simulacra instead of fluting, which was then copied by the next participant.

The tactile approach proved to be very time-consuming, impacting the quantity of samples. The tactile approach is also not as easily scalable as the VR approach because it requires more material and personnel time.

The VR approach was limited by both the capabilities of the hardware and design choices in the development of the application. The Meta Quest 3, while a very capable VR device, is reliant on camera detection of finger and hand position, with only limited ability to manage occlusion. This limited the accuracy of the virtual finger flutings of some hand positions, particularly the back handed movements. Furthermore, the virtual hand movements could not be easily matched with tactile feedback, i.e., an augmented reality approach, with the software available at the time, though this has since changed.

Design choices in the development of the VR application also posed significant limitations on the utility of the virtual data. For example, finger flutings were recorded as a 2D texture, providing no evidence of depth. This could be resolved using pseudo depth or 3D deformation of virtual surfaces and while the latter is more accurate it is more computationally intensive. Another design issue observed during the experiment was the difficulty a small number of participants had with the user interface elements, particularly the poke interactions with the virtual buttons, requiring significant guidance. This seemed to correlate to limited prior experience with VR but should be resolved by more intuitive UI to improve participant experience and accessibility.

While our intention was to develop a proof-of-concept to determine the sex of finger fluting artists based on modern populations, future researchers cannot assume a modern population has the same biomechanics as the ancient population that made the in-situ finger flutings. Future research into paleoanthropology for understanding the biomechanics of ancient populations is needed.

The novel combination of methods utilised in this study to understand the production of finger flutings has demonstrated several limitations and challenges, but also a range of insights into the application of these methods. The tactile approach captured nuances in the finger fluting that transferred to the images, which were computed by the ML model. The VR approach could be improved by adding motion capture and exploring alternative VR devices that could address the current limitations. For example, using haptic gloves to capture nuanced hand movements.

The methodologies developed in this study hold promise for a range of disciplines beyond archaeology, such as forensic science, human-computer interaction, and art history. AI-driven analysis of physical behavior and artistic intent could transform the way we study and understand ancient cultures, and the insights generated could have applications in modern fields such as user experience design and psychological research.

Conclusion

Our study makes an important contribution to experimental archaeology by using digital archaeology to understand if it is possible to determine the sex of the artist from the images of finger flutings taken from tactile and VR approaches. This study establishes a foundation for a paradigm shift from traditional analog methods that relied heavily on human-derived measurements (e.g., 2D/4D ratios) towards using purely computational digital archaeology methods, including ML, computer vision, remote sensing, for finger fluting analysis. While the tactile approach initially demonstrated promising performance, there was a pronounced disparity between training and test performance, likely the result of overfitting. The overfitting can potentially be remedied by increasing the sample size.

Another significant contribution of our study is the development of a quantifiable and scalable toolkit for finger fluting analysis. The toolkit can be used by future researchers for the entire lifecycle of the experiment, from planning to collecting and the tools for analyzing the data are available on github. It also includes the recipe for the moonmilk simulacra that was developed specifically to replicate the characteristics of moonmilk. The study paves the way for future research that integrates interdisciplinary approaches to cultural heritage studies with applications extending into diverse fields like forensics, psychology, and human-computer interaction.

Data availability

The data used in this study are openly available in github (<https://github.com/johnnydfci/FingerFluting-SexClassification>). The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

Received: 16 April 2025; Accepted: 29 August 2025

Published online: 16 October 2025

References

1. Marquet, J.-C. et al. The earliest unambiguous neanderthal engravings on cave walls: La Roche-Cotard, Loire valley, France. *PLoS ONE*. **18** (6), e0286568. <https://doi.org/10.1371/journal.pone.0286568> (2023).
2. Jalandoni, A., Haubt, R., Walshe, K., Nowell, A. C. & Photogrammetry Revolutionizing 3D modeling in low light conditions for archaeological sites. *J. Field Archaeol.* **50** (2), 132–144 (2024).
3. Nowell, A., Van Gelder, L. & Disentangled: The role of finger flutings in the study of the lived lives of upper paleolithic peoples. *J. Archaeol. Method Theory*. **27** (3), 585–606 (2020).
4. Walshe, K., Nowell, A. & Floyd, B. Finger-fluting in prehistoric caves – a critical analysis of the evidence for children, sexing and tracing of individuals. *J. Archaeol. Method Theory*. <https://doi.org/10.1007/s10816-024-09646-9> (2024).
5. Nowell, A. & Chang, M. L. Science, the media, and interpretations of upper paleolithic figurines. *Am. Anthropol.* **16** (3), 562–577 (2024).
6. Hays-Gilpin, K. Ambiguous Images. *Gender and Rock Art (Walnut Creek 2004)*.
7. Conkey, M. W. Mobilizing ideologies: Paleolithic art, gender trouble, and thinking about alternatives. In *Women in Human Evolution* (eds Schaller, M. B. & McCreery, J. G.) 173–207 (Routledge, 2005).
8. Moen, M. & Pedersen, U. (eds) *The Routledge Handbook of Gender Archaeology* (Taylor & Francis, 2024).
9. Moen, M. Gender and archaeology: where are we now? *Archaeologies* **15**, 206–226 (2019).
10. Gunn, R. G. Hand sizes in rock art: interpreting the measurements of hand stencils and prints. *Rock. Art Res.* **23**, 97–112 (2006).

11. Nelson, E., Hall, J., Randolph-Quinney, P. & Sinclair, A. Beyond size: the potential of geometric morphometric analysis of shape and form for the assessment of sex in hand stencils in rock Art. *J. Archaeol. Sci.* **78**, 202–213 (2017).
12. Rabazo-Rodríguez, A. M. et al. New data on the sexual dimorphism of the hand stencils in El Castillo cave (Spain). *JAS Rep.* **14**, 374–381 (2017).
13. Cobden, R. et al. The identification of extinct megafauna in rock Art using geometric morphometrics. *J. Archaeol. Sci.* **87**, 95–107 (2017).
14. Carden, N. & Blanco, R. Measurements and Replications of Hand Stencils. In *Paleoart and Materiality*, Archaeopress. (2017).
15. Hayes, S. & Van den Bergh, G. Cave art, art and geometric morphometrics. In *The Archaeology of Sulawesi* (eds O'Connor, S. et al.) 43–60 (ANU, 2018).
16. Fernández Navarro, V. et al. Decoding palaeolithic hand stencils: age and sex identification through geometric morphometrics. *J. Archaeol. Method Theory.* **32** (24). <https://doi.org/10.1007/s10816-025-09693-w> (2025).
17. Breuil, H. L'âge des caverns et roches Ornées de France et d'espagne. *Revue Arch.* **19**, 193–234 (1912).
18. Breuil, H. & Berger-Kirchner, L. Franco-Cantabrian rock art. In *Art of the Stone Age: Forty Thousand Years of Rock Art* 1st edn, (ed. Bandi, H. G.) 15–70 (Crown Publishers Inc., 1961).
19. Edwards, R. & Maynard, L. Prehistoric Art in Koonalda cave. *Proc. Royal Soc. Australasia.* **68**, 11–17 (1968).
20. Gallus, A. Parietal Art in Koonalda cave, nullarbor plain, South Australia. *J. Australasian Cave Res.* **6** (3), 43–49 (1968).
21. Gallus, A. Schematisation and Symbolling. In *Form in Indigenous Art: Schematisation in the Art of Aboriginal Australia and Prehistoric Europe* (Ed (ed Ucko, P. J.) (1st ed.) 370–386. (Duckwork Co. Ltd., 1977).
22. Maynard, L. & Edwards, R. Wall Markings. In *Archaeology of the Gallus site* (Ed. Wright, R. V. S.) (1st ed.) 61–80. (Australian Institute of Aboriginal Studies, 1971).
23. Lewis-Williams, D. *The Mind in the Cave: Consciousness and the Origins of Art* (Thames and Hudson, 2002).
24. Kelly, M., David, B. & Vilá, R. Finger flutings at new Guinea II cave, lower Snowy river Valley (Victoria), GunaiKurnai country. *Australian Archaeol.* 1–31. <https://doi.org/10.1080/03122417.2025.2529627> (2025).
25. Clottes, J. Shamanic practices in the painted caves of Europe. In *Spiritual Information: 100 Perspectives on Science and Religion* (Ed (ed Harper, C. L.) 279–285. (Templeton Foundation, (2005).
26. Sharpe, C. *Newly Discovered Art Sanctuary in Koonalda Cave, South Australia. D5/5/2; D5/8 628* (South Australia Museum Archives, 1973a).
27. Sharpe, C. *Report on a New Art Form in Koonalda Cave, South Australia. AR39* (South Australia Museum Archives, 1973b).
28. Sharpe, K., Lacombe, M. & Fawbert, H. Investigating finger flutings. *Rock. Art Res.* **19** (2), 109–116 (2002).
29. Sharpe, C. & Sharpe, K. A preliminary survey of engraved boulders in the Art sanctuary of Koonalda cave, South Australia. *Mankind* **10** (3), 125–130 (1976).
30. Marshack, A. *The Roots of Civilisation* (McGraw-Hill, 1972).
31. Van Gelder, L. The role of children in the creation of finger flutings in Koonalda cave, South Australia. *Child. Past.* **8** (2), 149–160 (2015).
32. Bednarik, R. G. Parietal finger markings in Europe and Australia. *Rock. Art Australia.* **3** (1), 30–61 (1986).
33. Sharpe, K., Lacombe, M. & Fawbert, H. Externalism in order to communicate. *Artefact* **21**, 95–104 (1998).
34. Sharpe, K. & Van Gelder, L. Evidence for cave marking by Palaeolithic children. *Antiquity* **80**, 937–947 (2006). (2006).
35. Assaf, E., Kedar, Y. & Barkai, R. Child in time: children as liminal agents in upper paleolithic decorated caves. *Arts* **14** (2), 27. <https://doi.org/10.3390/arts14020027> (2025).
36. Walshe, K. & Nowell, A. Rising up: digital traces, storytelling and performative Indigenous culture in Australian rock art. *Antiquity* (2025). <https://doi.org/10.15184/ajqy.2025.16>
37. Walshe, K. & Koonalda: Cave Nullarbor plain, South Australia: issues in optical and radiometric dating of deep karst caves. *Geochronometrika* **44**, 366–373 (2017).
38. Alvarez-Alonso, D. et al. More than a fingerprint on a pebble: A pigment-marked object from San Lázaro rock-shelter in the context of neanderthal symbolic behavior. *Archaeol. Anthropol. Sci.* **17** (6), 1–21 (2025).
39. Martínez-Sevilla, F. et al. Who painted that? The authorship of schematic rock Art at the Los machos rockshelter in Southern Iberia. *Antiquity* **94** (377), 1133–1151 (2020).
40. Nowell, A. *Growing Up in the Ice Age: Fossil and Archaeological Evidence of the Lived Lives of Plio-Pleistocene Children* (Oxbow Books, 2021).
41. Van Gelder, L. & Sharpe, K. Women and girls as upper paleolithic cave 'artists': Deciphering the sexes of finger fluters in rouffignac cave. *Oxf. J. Archaeol.* **28** (4), 323–333 (2009).
42. Snow, D. R. Sexual dimorphism in upper paleolithic hand stencils. *Antiquity* **80** (308), 390–404 (2006).
43. Chazine, J.-M. & Noury, A. Sexual determination of hand stencils on the main panel of the Gua Masri II cave (East-Kalimantan/ Borneo – Indonesia). *Int. News. Rock. Art.* **44**, 21–26 (2006).
44. Snow, D. R. Sexual dimorphism in European upper paleolithic cave Art. *Am. Am. Antiq.* **78** (4), 746–761 (2013).
45. Reilly, P. Towards a virtual archaeology. In *Computer Applications and Quantitative Methods in Archaeology CAA90* (eds (eds Rahtz, S. & Lockyear, K.) 132–139. (Tempus Reparatum, (1991).
46. Forte, M. *Virtual Archaeology: Re-Creating Ancient Worlds* (Harry N. Abrams, Inc., 1997).
47. Haubt, R. A. Virtual heritage archives: Building a centralized Australian rock Art archive. *Int. Arch. Photogramm Remote Sens. Spat. Inf. Sci.* **XL-5/W2**, 319–324. <https://doi.org/10.5194/isprsarchives-XL-5-W2-319-2013> (2013).
48. Wisher, I. & Needham, A. Illuminating palaeolithic Art using virtual reality: A new method for integrating dynamic firelight into interpretations of Art production and use. *J. Archaeol. Science: Rep.* **50** <https://doi.org/10.1016/j.jasrep.2023.104102> (2023).
49. Huggett, J. & Disciplinary Issues, Challenging the research and practice of computer applications in archaeology. In *Archaeology in the Digital Era* (eds Earl, G. et al.) 13–24 (Computer Applications and Quantitative Methods in Archaeology, 2013).
50. Huvila, I. et al. Archaeological information work and the digital turn. In *Archaeology and archaeological information in the digital society* (ed. Huvila, I.) 143–158. (Routledge 2018).
51. Morgan, C. L. Current digital archaeology. *Annu. Rev. Anthropol.* **51** (1), 213–231 (2022).
52. Bewes, J., Low, A., Morphett, A., Pate, F. D. & Henneberg, M. Artificial intelligence for sex determination of skeletal remains: application of a deep learning artificial neural network to human skulls. *J. Forensic Leg. Med.* **62** <https://doi.org/10.1016/j.jflm.2019.01.004> (2019).
53. Havens, L., Beatrice, A. & Terras, M. Confronting gender bias in heritage catalogues: A natural language processing approach to revisiting descriptive metadata. In *The Routledge Handbook of Heritage and Gender* (Ed. Ashton, J.C.) Part 5, Chap. 25Routledge., (2025).
54. Turner-Jones, R. N., Tuxworth, G., Haubt, R. A. & Wallis, L. Digitising the deep past: machine learning for rock Art motif classification in an educational citizen science application. *J. Comput. Cult. Herit.* **17** (4), 1–19 (2024).
55. Horn, C. et al. 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.* **29**, 188–213 (2022).
56. Jalandoni, A., Zhang, Y. & Zaidi, N. A. On the use of machine learning methods in rock Art research with application to automatic painted rock Art identification. *Journal Archaeol. Science.* **144**, 105629, 1–14 (2022).
57. Haubt, R. A. The global rock Art database: developing a rock Art reference model for the RADB system using the CIDOC CRM and Australian heritage examples. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci.* **II-5/W3**, 89–96. <https://doi.org/10.5194/isprsannals-II-5-W3-89-2015> (2015).

58. Nguyen, H. T. & Nguyen, L. T. Fingerprints classification through image analysis and machine learning method. *Algorithms* **12** (11), 241. <https://doi.org/10.3390/a12110241> (2019).
59. Ogechukwu, N. et al. Gender and age group classification from multiple soft biometrics traits. *Int. J. Biometrics*. **11** (4). <https://doi.org/10.1504/IJBM.2019.102883> (2019).
60. Forte, M. & Virtual Reality Cyberarchaeology, Teleimmersive Archaeology. In *3D Recording and Modelling in Archaeology and Cultural Heritage: Theory and best practices* (eds Remondino, F. & Campana, S.) BAR International Series 2598BAR publishing. (2014).
61. Mate, G. & Ulm, S. Working in archaeology in a changing world: Australian archaeology at the beginning of the COVID-19 pandemic. *Australian Archaeol.* **87** (3), 229–250 (2021).
62. Ravi, N. et al. Sam 2: Segment anything in images and videos. Preprint at (2024). <https://arxiv.org/abs/2408.00714> (2024).
63. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition 2016* 770–778, (2016). <https://doi.org/10.1109/CVPR.2016.90>
64. Tan, M., Le, Q. & Efficientnet Rethinking model scaling for convolutional neural networks. In *International conference on machine learning*, PMLR, 6105–6114. (2019). <http://proceedings.mlr.press/v97/tan19a.html>

Acknowledgements

This project was funded by a 2023 Griffith University AEL Research Project Grant, Australia Research Council Discovery Early Career Research Award (DE240100030), and Social Sciences and Humanities Research Council of Canada Insight Grant (#435-2019-0656). Many thanks to the participants who contributed samples of finger flutings for the tactile and virtual experiments. We would also like to acknowledge the Mirning People and Koonalda Cave for inspiring this work.

Author contributions

Conception and design: AJ, RH, CF; Acquisition, analysis or interpretation of data: AJ, RH, CF, GT, ZZ, KW, AN; Creation of new software: CF; Drafted or revised manuscript: AJ, RH, CF, GT, ZZ, KW, AN.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-18098-4>.

Correspondence and requests for materials should be addressed to A.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025