Utilizing machine-learning-based 3D image analysis for classifying charred grape seeds to the varietal level

Vlad Landa

Ariel University

Yekaterina Shapira

Ariel University

Michal David

Bar-Ilan University

Avshalom Karasik

Israel Antiquities Authority

Ehud Weiss

Bar Ilan University

Yuval Reuveni

Ariel University <https://orcid.org/0000-0002-8902-5540>

Elyashiv Drori (✉ droris@ariel.ac.il)

Ariel University

Article

Keywords: machine learning, grape seeds, 3D image analysis

Posted Date: October 7th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-85313/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Utilizing machine-learning-based 3D image analysis for classifying charred grape seeds to the varietal level

Vlad Landa¹, Yekaterina Shapira², Michal David³, Avshalom Karasik⁴, Ehud Weiss³, Yuval Reuveni^{5*} and Elyashiv Drori^{2,6*}

Author Information

Affiliations

¹ Department of Computer Science Ariel University, Israel

² Department of Chemical Engineering, Biotechnology and Materials, Ariel University, Ariel, 40700, Israel

³ Archaeobotanical laboratory and National Natural History Collection of Plants` Seeds and Fruits, Institute of Archaeology, Martin (Szusz) Department of Land of Israel Studies and Archaeology, Bar-Ilan University, Ramat-Gan, 5290002 Israel

⁴ The National Laboratory for Digital Documentation and Research in Archaeology, Israel Antiquities Authority, Jerusalem, Israel.

⁵ Department of Physics, Faculty of Natural Sciences, Ariel University, Science park, Ariel, Israel, 40700; Remote Sensing Lab, Eastern R&D Center, Ariel, Israel, 40700

⁶ The Wine Research Center, Eastern regional R&D Center, Ariel 40700, Israel.

Corresponding author*

Correspondence and requests for materials should be addressed to E.D. (email Droris@ariel.ac.il) or Y.R. (email: yuvalr@ariel.ac.il)

Yuval Reuveni- <https://orcid.org/0000-0002-8902-5540>

Elyashiv Drori- <https://orcid.org/0000-0003-3353-0465>

Abstract:

Grapevine (*Vitis vinifera* L.) is an essential part of the oldest group of fruit trees around which horticulture evolved, currently includes thousands of cultivars, grown at numerous climatic conditions. Discrimination between these varieties has been traditionally conducted using ampelography, and in recent decades mostly by genetic analysis. However, when aiming to identify archaeobotanical remains, which are mostly charred- with extremely low genomic preservation, the application of the genomic approach is rarely successful. As a result, variety-level identification of most grape remains is currently prevented. Because grape pips are highly polymorphic, several attempts were made to utilize their morphological diversity as a classification tool, mostly using 2D image analysis technics, aiming to utilize these methods for the identification of fresh and archaeological specimens. Here, we present for the first time a highly accurate varietal classification tool, using an innovative and accessible approach for 3D seed scanning. The suggested classification methodology is machine-learning-based, using a complete set of 3D data obtained for each seed, applied with the Iterative Closest Point (ICP) registration algorithm and the Linear Discriminant Analysis (LDA) technique. This methodology achieved classification results of ca. 90-99% accuracy when trained by fresh seeds to test unknown fresh seeds. Moreover, the classification of charred seeds reached up to 100% accuracy when trained by charred seeds. Based on this approach, our long-term aim is to develop a computerized classification tool for the identification of grape and possibly other species and varieties. Such a tool can significantly improve the fields of archaeobotany, as well as general taxonomy.

Introduction

Grapevine is one of the classical fruits of the Old World and an essential part of the oldest group of fruit trees around which horticulture evolved at the Mediterranean basin¹. This specie includes thousands of known cultivars, grown at a wide array of climatic conditions, as well as its wild progenitor (*Vitis vinifera* ssp. *sylvestris*)². Discrimination between grape varieties has been done traditionally using ampelography³, a field of classification by the shape and color of leaves, bunches and berries. In recent

decades, grape variety identification dramatically evolved, exploiting the development of DNA analysis methods by AFLP^{4,5}, SSR^{6,7}, and SNPs⁸⁻¹⁰. These techniques are very straight-forward and accurate when fresh plant material is available.

In archaeobotany (aka paleoethnobotany), the scientific study of plant remains from archaeological sites for reconstructing and interpreting past environments and human-plant relationships- single-species identification is fundamental. Although it may involve much time and great effort, it is of utmost importance as meaningful interpretations and reconstructing reliant on well-identified species¹¹. However, this goal is not always achieved, and this bottleneck hampers the researcher's ability to answer fundamental research questions. One of the main reasons for the mentioned bottleneck is the fact that high genomic preservation is typically found in rare desiccated or waterlogged plant remains¹². At the same time, most archaeobotanical assemblages worldwide went through the charring process – which badly influences genomic preservation^{13,14,15}. Current methods are limited in producing high quality and quantity of aDNA from charred seeds due to low endogenous DNA content, short DNA fragments with high rates of nucleotide damage, and high rates of modern DNA contamination, leading to the yield of insufficiently reliable genetic data¹⁶. These facts pose a high barrier preventing the identification of most grape remains in current archeological repositories. Therefore, seeking to identify grape varieties in archaeobotanical grape remains, a different approach is urgently needed.

Grape pips are highly polymorphic¹⁷. Exploiting this fact, several attempts were reported recently utilizing the diversity in fresh grape pip morphology as a diagnostic tool, using image analysis techniques, aiming to utilize these methods for the identification of fresh and archaeological specimens. First reports in this field appeared in the last decade- in the work of Terral *et al.*^{18,19}, where geometrical analysis (using elliptic Fourier transform method) was applied to analyze the 2D outlines of fresh grapevine pips. Using this methodology, a morphological key was created for pips of approximately fifty French grapevine varieties. Also, a significant correlation between pip morphology and the taxonomic relationship was demonstrated. Furthermore, an innovative approach for the investigation of archaeological remains by combined 2D morphometric and genetic methods was recently developed, showing promising results in melon seeds^{20,21}.

Over the last decade, 3D scanning technology has advanced dramatically. Besides its significant role in the industry, modern academic studies harnessed this technology to

explore and investigate new questions that were never accessible without the 3D acquired information. Various identification techniques were developed, combining the potential of highly accurate 3D scanning and imagery technology, with mathematical and statistical classification methods and innovative tools from the field of computer sciences^{22–26}.

Furthermore, current developments in cloud-based big-data technologies enable data-driven solutions, applied with increasing numbers of scientific computing studies^{27,28}. Machine learning (ML) is undoubtedly the most common data-driven solution approach which finds complex mathematical patterns and relationships inside the data and uses them to bring considerable datasets to the surface^{29,30}. Commonly, ML techniques consist of two learning types: Supervised learning and Unsupervised learning. Supervised learning means that every trained data sample has a known label. The ML model outcome can be categorical or continuous, depending on the nature of the problem. A categorical output can be laid in a simple 1/0 label, and its methodical term is referred to as binary classification. In the case of a continuous outcome, the methodical term is referred to as regression. The algorithms that are used to tackle classification and regression problems include linear regression, Random Forest (RF) decision trees, Support Vector Machines (SVM), and Linear Discriminant Analysis (LDA)³¹.

On the other hand, Unsupervised Learning aims to deduce hidden internal data features and patterns without the need for assigned labels. This type of learning is commonly used for hidden features-based clustering. ML algorithms such as K-Nearest Neighbors (KNN), Principal Component Analysis (PCA), and Self Organizing Map (SOM) are all examples of Unsupervised Learning models³². Therefore, using multivariate data analysis methods, adopted from the ML discipline for classifying agricultural, as well as archeological geometrical^{33–35} and structural features, can be extremely valuable^{36,37}.

In a previous publication³⁸ we described our efforts in developing a 3D tool for grape variety identification by grape pip structure. It was clearly demonstrated that the 3D method described is a promising tool for grape variety identification using fresh grape pips, as it enabled the separation between different *Vitis vinifera* varieties with high statistical certainty. This method made use of a set of planar curves extracted from a

full 3D scan of the seed, which represents its key features. The novelty of this method lies in the combination of scanning methods and the right selection of Fourier coefficients and their weights.

Here, we present a highly accurate varietal classification of charred grape seeds, using an innovative and accessible 3D scanning method, combined with a machine-learning-based classification technique, using the complete set of 3D imagery data. We also suggest an innovative way of upscaling the analysis to a broader set of varieties. This breakthrough is the first step in developing a computerized classification tool for the identification of grape, and possibly other species of archaeobotanical seeds, at the variety level.

Experiments & Results

Visualization of grapevine pips

A set of 3D grape seed scans were used for evaluating the performance of the alternative 3D classification. We selected pips of four grape varieties for scanning, followed by classification of the pips to their initial classes (varieties). Cabernet Sauvignon (N = 15) is a highly esteemed international variety. The three other varieties: '292' (N = 15), '13' (N = 15), and '9003' (N=15) are *Vitis vinifera* ssp. *sativa* lines that were collected from the Israeli endogenous grape varieties collection in Ariel^{39,40}. The 3D image scan was then converted into a 3D points cloud representation as described in materials and methods (Figure 1B, C.)

Variety 9003

Variety 13

Cabernet Sauvignon

Variety 292

Figure 1: The procedure used for grape pip 3D data acquisition. **A-** High-quality focused image; **B-** 3D image scan, and **C-** 3D points cloud representation of four varieties' fresh pips. 9003, 292, and 13 - Israeli endogenous varieties and Cabernet Sauvignon - international variety.

Linear discriminant analysis of fresh grape pips

In the first analysis, we evaluated the effectiveness of our classification method accuracy, using fresh pips from 4 grape varieties. For every variety, ten representative uncharred (fresh) pips were chosen as a training set (40 in total, Figure 2A), and five pips were selected as a test set (20 in total, Figure 2B), where each test sample pip was classified independently. The result of an LDA classification of grape pips with an accuracy of 90 % is presented in Figure 2B.

Figure 2: Classification of fresh pips, using a train set of fresh pips. The actual variety label of the pip is presented as a ring, and a dot marks the predicted label. **A** - Visualization of LDA separability learned during a training set with 10 out of 15 pips in each class; **B** - Visualization of LDA classification of the remaining five pips (out of the 15), in each class as a test set (accuracy 90.00%); **C** - Accumulative accuracy of LDA classification after performing 100 training and 100 independent test evaluations for each pip. At each run, the same ten representative pips from the assemblage were selected as a constant training group, and the other five pips were classified independently; **D** - Classification accuracy histogram - 100 training [10 samples] and tests [5 samples], gaining mean accuracy of 95%. Cabernet Sauvignon* - 1024.

One hundred evaluations were conducted. In each test, the same ten samples were chosen as a variety representative set for training, while five were kept out for the test and were classified independently. This was done in order to statistically characterize the impact of the ICP algorithm initialization randomness⁴¹, i.e., when initiating the comparison of two-point clouds by the ICP, a random point is selected from the target cloud and matched to the closest corresponding points at the source cloud. Figure 2C presents the classification statistics for each variety over the multiple tests. The number of times a pip was classified at any group is displayed as a yellow bar. If a pip was always classified into the same group, then a single bar is plotted, and it covers the full height of the corresponding group (i.e. 100% classification). However, when classified into several groups, corresponding bars are displayed, and their heights are proportional to the classification's percentages. The highest classification accuracy of 98.8% was obtained for Cabernet Sauvignon, the lowest accuracy of classification of 89.8% was obtained for the 9003 variety. In comparison, classification accuracies of 94% and 89.8% were accepted for 13 and 292 varieties, respectively. Figure 2D presents the total accuracy distribution over all 100 tests, with a mean value of 95% and a standard deviation of 5%.

Linear discriminant analysis of charred grape pips

Archaeobotanical seeds are preserved for many years due to charring. Unfortunately, seeds become deformed as a result of exposure to heat^{15,42}. Nevertheless, we hypothesized that our suggested 3D morphological classification method might overcome the limitation posed by the deformities and yield good classification results. We started with acquiring high-quality charred seed scans, in which the stereomicroscope proved to be a good selection, as opposed to scanning by a high-resolution 3D scanner - 'PT-M' (not shown). The resulting high-quality scans enabled transforming the data into cloud point representations, similar to those achieved for fresh seeds (Figure 3).

Figure 3: acquisition of 3D data for charred grape pips. A- High quality focused 3D image; B- height map, and C- points cloud of two varieties of charred pips.

Hence, we set to evaluate the classification accuracy of charred pips from 4 grape varieties, using 13 fresh pips from each variety (i.e., 52 pips in total), in order to characterize the burning effect. This was done by scanning the fresh pips for training purposes, then charring them, and re-scanning them in charred form for test purposes. As previously explained, 100 consequent evaluations were conducted. Figure 4A presents the training result of 13 fresh pips for each variety. The result of one classification procedure test using all charred pips from all four varieties is shown in Figure 4B. It can be noticed that the classification accuracy of the 13 charred pips, which is $\sim 88\%$, is lower than the classification accuracy for fresh pips presented above (Figure 2B). Figure 4C presents the classification statistics for each variety. The highest classification accuracy of 94.5% was obtained for Cabernet Sauvignon, while the lowest classification accuracy of 77.3% was obtained for variety 13, in which two

evident outlier seeds can be noticed. Classification accuracies of 91.7% and 91.3% were obtained for varieties 9003 and 292, respectively. Figure 4D represents the distribution of accuracies over all 100 tests, with a mean value of 89% and a standard deviation of 8%.

Figure 4: Classification of charred pips using fresh pips as a train set. The actual variety label of the pip is presented as a ring, and a dot marks the predicted label. **A** - LDA training - Fresh samples, 13 from each variety (1024=Cabernet Sauvignon); **B** - LDA testing - Burned samples, 13 from each variety (accuracy 88.46%); **C** - Accumulative distribution of LDA classification after running 100 tests. At each run, 13 representative pips from the assemblage were selected as the constant training group, and the 13 burned pips were classified independently; **D** - Classification accuracy histogram - 100 training [13 fresh samples] and independently 100 tests for each pip [13 burned samples].

Finally, we conducted an experiment to assess the classification efficiency when using charred pips to train the machine, followed by testing unclassified pips originating from the same varieties. As illustrated in Figure 5, the classification accuracy of this final

approach was the best so far, gaining 100% of success in the test phase (Figure 5B). Figure 5C presents classification statistics for each variety over 100 separated runs. Perfect classification accuracy of 100% was received for both Cabernet Sauvignon and variety 292, while varieties 13 and 9003 gained very high accuracies of 98.8% and 98.6%, respectively.

Figure 5: classification of charred pips using a charred training set. The actual variety label of the pip presented as a ring, and the predicted label marked by a dot. **A** - LDA training - burned samples, 10 out of 15 from each class (1024=Cabernet Sauvignon); **B** - LDA testing - Cleaned burned samples, the remaining independently five from each variety (100.00% accuracy); **C** - Accumulative distribution of LDA classification after running 100 tests. At each run, ten pips from the assemblage were selected as the constant training group, and the other five pips were classified. **D** - Classification score histogram - 100 training [10 burned samples] and independently 100 tests for each pip [5 burned samples].

Towards the classification of a broader set of varieties.

In order to examine the scalability of the suggested methodology described above, we performed an additional classification experiment involving eight grape varieties. We added four other charred varieties (98, 192, 236, 276) to the four charred varieties used in our previous classifications (292, 9003, 13, 1024). Ten representative pips from each variety were chosen as a training group, while five were kept out for testing. In total, we conducted 100 training and classification evaluations of LDA, in which a total of 80 charred samples were used as a train set and 40 charred samples as the test set. Figure 6 shows the classification statistics.

204

Figure 6: Classification of charred pips from 8 varieties. **A** - Classification accuracy (y-axis) over 100 iterations on each test sample (x-axis) **B** - Classification accuracy score histogram of 8 varieties - 100 training [10 charred samples] and independently 100 tests for each pip [5 charred samples].

As shown in Figure 6, the classification accuracy for five of the participating varieties was over 90%, while three varieties showed lower accuracy scores, resulting in an average accuracy level of 88%. Compared to the high classification accuracy achieved for four varieties, these results indicate that the classification accuracy degrades when increasing the number of classes. Obviously, to develop a practical method that enables

the classification of a seed of unknown origin, the reference population will contain a large number of varieties. As a result, the signal magnitude of each group will be reduced dramatically relative to the entire collection, and no learning can be gained based on the training set. Thus, to improve the accuracy for a broader variety set, we suggest implementing a “tournament” classification flow. This innovative approach is based on classifying an unknown seed by conducting multiple LDAs, trained with group sets of a small number of varieties (for example, four). The varieties selected from each group are forwarded to the next stage, in which a new group of varieties is trained into group sets of four that are used for the classification of the tested pip. This configuration reduces the number of remaining candidate varieties by a factor of 4 at each stage. Finally, a last group of the highest-ranking candidate varieties will be used as for a final LDA, producing the variety best-fitted to the tested pip, out of the entire tested population (for an illustration of the proposed process, see Figure 8).

Testing the classification accuracy of eight varieties by a Tournament classification flow.

To demonstrate our suggested methodology for classifying a large variety number (in this specific case, eight), we implemented a tournament classification flow experiment, aimed to distinguish between 8 different charred varieties, by initially training two individual LDA classifiers (Figure 7C). Classifier A (Figure 7A) was trained (using ten representative pips) to distinguish between varieties 292, 9003, 13, and 1024, as previously described. Additionally, classifier B (see Figure 7B) was trained in the same way to distinguish between varieties 98, 192, 236, and 276. The remaining 40 samples (5 from each variety) were kept out of the training as a test group. We then evaluated the classification accuracy of each test sample over 100 iterations in the following way: In the first stage of the tournament, each test sample was classified by classifier A, and then by classifier B. In the second stage, we trained a new LDA, which aimed to classify the test sample between the two “winning” results, which were selected during the first stage (Figure 7C). Finally, the selected variety was reported by the LDA of the second stage (see Figure 7 D, E) as the tested pip’s variety. Figure 7D shows the classification results. It can be noted that except for variety 292, which had one outlier, all other varieties were classified by this method with an accuracy ranging between 99-100%, with a total average of 99%.

A

B

C

D

E

Figure 7: Classification of charred pips from 8 varieties by the “tournament” flow method. **A** – Statistic of classifier A – trained to distinguish between 292,9003,13,1024 charred classes, **B** – Statistic of classifier B – trained to differentiate between 98, 192, 236, and 276 varieties in charred form. **D** - Accumulative distribution (y-axis) of LDA classification of 8 varieties of the “tournament” flow method, after running 100 iterations for each test pip sample (x-axis). **E** – Classification accuracy score histogram of 8 varieties –100 iterations for each independent test sample.

Discussion:

In this work, we demonstrated a successful varietal classification of charred and fresh grape seeds using an accessible 3D scanning method. Various grape pips positions, plate surfaces materials, and stereoscopic focus intervals were tested to achieve optimal conditions for accurate scanning. We used a Nikon SMZ25 stereo microscope to create scan sets of grape pips and designed the conversion of the resulting “wrl” files into a cloud point representation. An automated stereoscopic microscope is an available tool found in many labs, making our method of using the full data set gained by 3D data for separating the varieties by their morphological trait, available and approachable. This new approach is the next step in the journey for grape variety identification for agriculture and archaeobotanical purposes, started by the development of the traditional morphometric methods and the elliptic Fourier transform method^{18–20,43} later applied for charred and archaeological findings by analysis of surface morphology^{44–49}. Our study utilizes the LDA algorithm classification advantages for classifying preprocessed stereoscopic grape seed images. Our meticulously tailored preprocessing phase constitutes data representation that emphasizes the morphological distance between seeds type, based on a cloud point data structure. We define such distance as the minimum Mean Square Error (MSE) of the Euclidean distance between two cloud points measured by the Iterative Closest Point (ICP) algorithm. Our method discriminates between and accounts for three types of errors (the error reduction method is discussed later): 1. Human error – positioning of each sample by hand may introduce a lack of homogenous scanning, 2. ICP algorithm converge error - Applying ICP from sample A to B might be different from applying it from B to A due to the cost function defined in the ICP algorithm, and 3. The ICP algorithm introduces randomness, i.e., using ICP from A to B might differ in every iteration. Concerning the above mentioned possible introduced errors, we implemented the following steps: 1. We developed a scanning protocol while designing an alignment technique that reduces human error possibility, 2. We forced the ICP algorithm to include all data points in a cloud data structure representation to calculate minimum distance; such constrain minimizes the difference between the two way ICP evaluation, and thus emphasizes morphological similarities, and 3. We construct a proper statistical analysis, based on a vast amount of evaluations to examine our classification accuracy distribution under the suggested methodology. In addition to the errors mentioned above, which could lead to possible

misclassifications, trivial biological causes such as natural deformations can also lead to misclassification of seeds, as shown in Figure 4C for class 13C.

Furthermore, our suggested classification approach presents many advantages, which enhances its applicability for variety identification by seed, as compared with similar works. For example, Karasic et al.³⁸ used Fourier transform coefficients as a Machine Learning input and perform PCA for clustering visualization applied with 3D pip scans, which demand a meticulous scanning process and high accuracy of positioning. Bouby et al.⁴³ used Elliptical Fourier Transform (EFT) to extract dominant features from dorsal and lateral image outlines utilizing LDA for classification. The last study presented the results based on the “leave-one-out” folding method with posterior classification ($P > 0.75$). Such statistical analysis is informative for one sample classification but makes it difficult to generalize the model’s ability to classify the entire given group-set. In contrast, our suggested scanning method is less sensitive to inaccurate positioning of seeds; this simplifies and speeds-up data acquisition. Additionally, our results indicate good classification accuracy along with adequate ML model data generalization, given the selected representative sets. This was partially achieved since the ICP methodology increases the homogeneity in feature comparison between two given samples due to its algorithm nature. Thus, our proposed method might be suggested as suitable for big data analysis.

Implementing our methodology for classifying fresh and charred grape seeds of four varieties, we achieved a mean accuracy level of 88% for fresh (train) vs. charred (test) pips, mean accuracy of 94% for fresh vs. fresh pips, and unpredictably, the highest classification rate of 99% was achieved when charred pips were used as a training set to identify charred pips. These results emphasize that burned samples show increased morphological similarity compared to fresh ones, possibly due to the removal of various fresh soft tissues present on the seeds surface, which reduces “structural noise”.

These results also indicate that a future methodology for archaeobotanical remains identification might be best implemented on the construction of a reference database of a large set of charred grape pips of the populations of reference varieties, for the accurate classification of unknown archaeological charred grape pips. The applicability to other important plant species is yet to be determined, as the deformities caused by

charring may differ between species. In addition, the relatively high classification rate achieved when using a fresh seeds training set to classify a charred test set suggests that although typical morphological changes occur in grape pips following charring^{15,42,43}, our 3D identification approach can overcome this barrier, and suggests that identification results may also be achieved by using fresh pips reference data set towards possible identification of charred seeds, without the need for an empirically calculated charring compensation key.

Although the ICP based morphological classification method shows comparable and distinguishable results, the need for a more robust and deterministic algorithm still exists. One way towards such improvement is to take advantage of the scanned sample rigid structures, represented as a 3D surface. Such representation captures spatial morphological features, in contrast to cloud points representation based on discrete Euclidean points. Several previous studies attempted to find different metrics for surface similarities. A most recent work by Lipman et al.⁵⁰ used conformal mapping to map 3D scanned teeth of mammals, found on an archeological site, to a unit disk probability space, and defined the Wasserstein distance as the metric between them. Utilizing Differential Geometry in future works can be an essential key for building a robust, efficient, and accurate classifier that can handle hundreds and thousands of grape varieties.

Our method, which demonstrates promising results, still requires handling the need for identification of an unidentified pip against a broad set of reference varieties. Currently, classification results are dramatically better when a small group of varieties is classified at one point, dropping with the addition of varieties. We propose to address this issue by either implementing the suggested “tournament” methodology analysis (see Figure 8), in which the seed is identified by classification against pre-trained and “on flow” trained multiple models, such that each model will be trained to distinguish between four different classes; the classification flow will be divided into $\log_4 n$ stages, where n denotes the number of varieties. As a first stage, the sample will be classified by $n/4$ pre-trained classifiers. Then, in every following step, a new LDAs will be trained based on the output of the previous stage and trained only with the initial training subsets. As a final stage, a four varieties (or less) LDA classifier will be trained and output the final result. We recommend evaluating each sample with at least ten iterations in order to

gain reliable statistics. Figure 8 shows the general case of such implementation.

Figure 8: General demonstration of a “tournament” classification flow and its internal stages of training and classifications down to the final identification result. The test sample may be an unidentified charred grape pip, recent or archaeological. The whole reference population, which is believed to include the tested sample’s variety, is divided into random groups of four varieties, trained as separated machines. At the end of the first step, the selected variety of each machine will be elevated to the next stage, where again, the newly established set will be randomly divided into groups of four varieties, and so on, until the final stage will recognize the most probable identification result.

In Addition, we suggest that future approaches will utilize advanced ML algorithms such as Deep Learning (DL). Nevertheless, this approach requires a vast amount of training samples for each verity, which will demand an even more robust and less time-consuming scanning method. We are currently exploring the mentioned above approaches towards implementation in the identification of charred archaeological grape seeds.

Conclusions

The presented innovative 3D classification method shows good classification results for fresh grape seeds and even higher accuracy levels for charred ones. To our knowledge, this is the first application of a 3D classification tool which makes use of the full 3D data set. This tool can be further developed to accurately identify charred archeological seeds, and may present a breakthrough in any plant taxonomy-related field.

Materials and methods

Plant material

A total of 60 seeds from 8 cultivars were sampled. Grapes from the endogenous Israeli varieties 9003 (Dabuki M.), 13 (Marawi), 292 (Tzuriman S.), 98 (Homra Pisga), 192 (Banias 1), 236 (Samach Harduf), and 276 (Banias Shaar) were collected from the endogenous varieties vineyard collection^{39,40} and Cabernet Sauvignon grapes were collected from the European varieties collection vineyard, both at Ariel, Israel. Mature seeds were extracted from ripened grapes, collected from at least three different grapevines. The seeds were washed by water to discard any residual pulp tissue and air-dried for two days, then stored in a closed vial until used. Before scanning, each seed was carefully cleaned by brushes and needles from any external tissues coating its crevices, to enable an effective scan of the seeds' topography. This was done following experiments showing that a scan of uncleaned seeds does not capture a substantial part of their structure (sup. Figure 1). Fifteen seeds were prepared from each variety for the scans.

Imaging by stereo microscope

An aluminum foil was chosen as the optimal surface for accurate scan results, out of various tested surfaces: white paper, black paper, glasses with different coatings, etc.. Grape seeds were placed on a microscope glass slide wrapped with aluminum foil, set under the microscope and illuminated with LED spots. A polylactide (PLA) light cap was designed and printed using ULTIMAKER to achieve a uniform and equal luminescence, needed for high-quality scans. Four LED spots were glued inside the self-designed cap (sup. Figure 2). Images were taken with a Nikon SMZ25 stereo-microscope (Nikon, Tokyo, Japan) equipped with a Nikon DS-Ri2 microscope camera.

Sixty digital micrographs (resolution: 4908×3264 pixels), with each step about $50 \mu\text{m}$, were taken at different focal planes and compiled to a single image using ND2-NIS elements software with an Extended Depth of Focus (EDF) patch (Nikon Instruments, Japan). To construct a 3D image scan of the seeds, sequential imaging of the ventral side positioned horizontally was conducted by the stereo microscope, with intervals of approximately $50 \mu\text{m}$. These images were then transformed into a single high quality focused image in a “wrl” file format using the dedicated stereo microscope’s software (Figure 1A).

Heating experiment

Fresh grape pips were heat-treated to study the impact of charring on morphological changes in pips. For this purpose, clean grape pips were scanned by a stereo-microscope before and after charring. Pips were heated in batches of 15, at a temperature of 230°C for 2 hours under low oxygen availability (covered with a thick layer of sea sand) to prevent the burning of the pips^{15,51}.

Data preparation

Given a scanned data sample in the “wrl” file format, the first step is converting it to a cloud point representation. The conversion was achieved by multiplying every coordinate ($\vec{c}_i$)

$$\vec{c}_i = (x_i, y_i) \in F$$

inside the “wrl” file (F), excluding coordinates that represent the surface plate itself, by the inverted intrinsic matrix (K^{-1}).

$$K_{3 \times 3} = \begin{bmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

The intrinsic matrix (K) contains parameters that describe the visual sensor characteristics: (f_x, f_y) are the focal-lengths, (c_x, c_y) are the sensors center coordinates in pixels and (s) is the skew parameter⁵². Then, the coordinate ($\vec{c}_i$) is scaled by its corresponding height map value ($z_i \in F$ - z-axis coordinate).

$$\begin{aligned} \vec{V}_{c_i} &= K^{-1} \cdot z_i \cdot [c_i \quad 1] \\ &= K^{-1} \cdot z_i \cdot [x_i \quad y_i \quad 1] \end{aligned}$$

Intrinsic matrix parameters were determined manually based on the actual surface dimensions of the scanned image sample, along with the distance between the surface and the visual sensor. As a next step, we defined the training and test sets size in the following way: 10 constant variety representative pips as the training sets and the remaining five pips as the test sets for multi-class classification scenario and 13 fresh (uncharred) pips as the training set versus 13 charred pips for the heating classification scenario. The training set is described as an M by M matrix, where M is the total number of training samples (i.e., ten pips from each variety for the multi-class classification, and 13 pips from each variety for the heating classification). The matrix is constructed such that every i, j matrix entry represents the Mean Square Error (MSE) of the applied Iterative Closest Point (ICP) algorithm on points-cloud sample i and points-cloud sample j (i.e., for the multi-class classification scenario with four different varieties, we get 40 by 40 matrix, and for the heating classification scenario with four different varieties, we get 52 by 52 matrix). The test set matrix was constructed in the same manner as the training set, except that in this case, the test matrix size has N by M dimensions, where N is the total number of test samples (i.e., $N = 20 = 5*4$ for the multi-class classification scenario, and $N = 52 = 13*4$ charred pips for the heating classification), and M is the total number of training samples (i.e., $M = 40 = 10*4$ the multi-class classification scenario, and $M = 52 = 13*4$ fresh (uncharred) pips for the heating classification). An example of the training and test matrix representation for the multi-class classification scenario is shown in Figure 9:

(A)

(B)

Figures 9: Train and test matrices (sets), each index inside the matrices represents the MSE value reported by the ICP algorithm between pip points cloud i and pip point cloud j . A-Train matrix (set) of ICP scores with 40x40 dimensions, 10 constant representative pips (y-axis and x-axis) from each class of 292f,9003f,13f and 1024f classes. B - Test matrix (set) of ICP scores with 20x40 dimensions, 5 remaining pips (y-axis) from each class of 292f,9003f,13f and 1024f classes and the 10 constant representative pips (x-axis).

For the case of classifying individual pip, the test matrix becomes a single vector with a size of 1 by M. As such, for different classification scenarios (multi-class classification and heating classification), a different LDA model was trained based on the scenario training matrix (set) as an input. Each LDA model was evaluated based on the test matrix (set) as an input according to the classification scenario. We note that the evaluation of a test pip set is equivalent for evaluating individual pips one by one since the test pips are treated independently in the test set matrix.

LDA analysis:

Linear Discriminant Analysis (LDA) is a supervised machine learning technique used for classification problems⁵³. Similar to PCA, LDA uses dimensionality reduction as a preprocessing step, but in contrast to PCA, LDA considers data labels. The LDA

method creates a projection of high dimension features onto low dimensional space in three necessary steps:

1. It calculates the separability between different classes.

$$S_b = \sum_{i=1}^g N_i (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})^T$$

2. It calculates the distance between the mean-variance of each class (“within-class variance”).

$$S_w = \sum_{i=1}^g (N_i - 1) S_i = \sum_{i=1}^g \sum_{j=1}^{N_i} (x_{i,j} - \bar{x}_i)(x_{i,j} - \bar{x}_i)^T$$

3. It constructs a low-dimensional space such that it maximizes the mean-variance for each class (“between class variance”) and minimizes the mean-variance between different classes (“within-class variance”). Let P be lower dimensional space projection, which is called Fisher’s criterion.

$$P_{ida} = \arg \max_P \frac{|P^T S_B P|}{|P^T S_W P|}$$

Furthermore, we used the “discriminant_analysis” python package from the “sklearn” library to train and evaluate the LDA performance [https://scikit-learn.org/stable/modules/generated/sklearn.discriminant_analysis.LinearDiscriminantAnalysis.html]. The train set matrix in the LDA were trained on its default values and then assessed on the test set matrix.

ICP:

Iterative Closest Point (ICP) is a registration algorithm designed to find a transformation between two given point clouds that minimizes any arbitrary objective function^{54–57}. We used the Open3D library to perform ICP on the samples. [Open3D: A Modern Library for 3D Data Processing].

References

1. Zohary, D., Hopf, M. & Weiss, E. *Domestication of Plants in the Old World: The origin and spread of domesticated plants in Southwest Asia, Europe, and the Mediterranean Basin*. (Oxford University Press on Demand, 2012).
2. This, P., Lacombe, T. & Thomas, M. R. Historical origins and genetic diversity of wine grapes. *Trends Genet.* **22**, 511–519 (2006).
3. Garcia-Muñoz, S., Muñoz-Organero, G., de Andrés, M. T. & Cabello, F. Ampelography - An old technique with future uses: the case of minor varieties of *Vitis vinifera* L. from the Balearic Islands. *OENO One* **45**, 125–137 (2011).
4. Sensi, E., Vignani, R., Rohde, W. & Biricolti, S. Characterization of genetic biodiversity with *Vitis vinifera* L. Sangiovese and Colorino genotypes by AFLP and ISTR DNA marker technology. *Vitis* **35**, 183–188 (1996).
5. Cervera, M.-T., Cabezas, J. A., Sancha, J. C., Martínez de Toda, F. & Martínez-Zapater, J. M. Application of AFLPs to the characterization of grapevine *Vitis vinifera* L. genetic resources. A case study with accessions from Rioja (Spain). *Theor. Appl. Genet.* **97**, 51–59 (1998).
6. Grassi, F. *et al.* Evidence of a secondary grapevine domestication centre detected by SSR analysis. *Theor. Appl. Genet.* **107**, 1315–1320 (2003).
7. Emanuelli, F. *et al.* Genetic diversity and population structure assessed by SSR and SNP markers in a large germplasm collection of grape. *BMC Plant Biol.* **13**, 39 (2013).
8. Cabezas, J. A. *et al.* A 48 SNP set for grapevine cultivar identification. *BMC Plant Biol.* **11**, 153 (2011).
9. Ganai, M. W., Altmann, T. & Röder, M. S. SNP identification in crop plants. *Curr. Opin. Plant Biol.* **12**, 211–217 (2009).
10. Deleu, W. *et al.* A set of EST-SNPs for map saturation and cultivar identification in melon. *BMC Plant Biol.* **9**, 1–9 (2009).
11. Weiss, E. & Kislev, M. E. Plant remains as a tool for reconstruction of the past environment, economy, and society: Archaeobotany in Israel. *Isr. J. Earth Sci.* **56**, 163–173 (2007).
12. Mascher, M. *et al.* Genomic analysis of 6,000-year-old cultivated grain illuminates the domestication history of barley. *Nat Genet* **48**, 1089–1093 (2016).
13. Ramos-Madrigal, J. *et al.* Palaeogenomic insights into the origins of French grapevine diversity. *Nat. Plants* **5**, 595–603 (2019).
14. Wales, N. *et al.* The limits and potential of paleogenomic techniques for reconstructing grapevine domestication. *J. Archaeol. Sci.* **72**, 57–70 (2016).
15. Charles, M., Forster, E., Wallace, M. & Jones, G. “Nor ever lightning char thy grain”1: establishing archaeologically relevant charring conditions and their effect on glume wheat grain morphology. *Sci. Technol. Archaeol. Res.* **1**, 1–6 (2015).
16. Nistelberger, H. M., Smith, O., Wales, N., Star, B. & Boessenkool, S. The efficacy of high-throughput sequencing and target enrichment on charred archaeobotanical remains. *Sci. Rep.* **6**, 37347 (2016).

- 787 17. This, P., Lacombe, T. & Thomas, M. Historical origins and genetic diversity of wine grapes.
788 *Trends Genet* **22**, (2004).
- 789 18. Terral, J. Quantitative anatomical criteria for discriminating wild grape vine (*Vitis vinifera* ssp.
790 *sylvestris*) from cultivated vines (*Vitis vinifera* ssp. *vinifera*). *Br. Archaeol. Reports*
791 (*International Ser.* **1063**, 59–64 (2002).
- 792 19. Terral, J.-F. *et al.* Evolution and history of grapevine (*Vitis vinifera*) under domestication: new
793 morphometric perspectives to understand seed domestication syndrome and reveal origins of
794 ancient European cultivars. *Ann. Bot.* **105**, 443–55 (2010).
- 795 20. Bacilieri, R. *et al.* Potential of combining morphometry and ancient DNA information to
796 investigate grapevine domestication. *Veg. Hist. Archaeobot.* **26**, 345–356 (2017).
- 797 21. Sabato, D. *et al.* Molecular and morphological characterisation of the oldest *Cucumis melo* L.
798 seeds found in the Western Mediterranean Basin. *Archaeol. Anthropol. Sci.* **11**, 789–810 (2019).
- 799 22. Grosman, L., Karasik, A., Harush, O. & Smilansky, U. Archaeology in Three Dimensions in
800 Archaeological Research. *J. East. Mediterr. Archaeol. Herit. Stud.* **2**, 48–64 (2014).
- 801 23. Grosman, L., Sharon, G., Goldman-Neuman, T., Smikt, O. & Smilansky, U. Studying post
802 depositional damage on Acheulian bifaces using 3-D scanning. *J. Hum. Evol.* **60** 398–406 (2011).
- 803 24. Karasik, A. & Smilansky, U. 3D scanning technology as a standard archaeological tool for
804 pottery analysis: practice and theory. *J. Archaeol. Sci.* **35**, 1148–1168 (2008).
- 805 25. Razdan, A., Liu, D., Bae, M., Zhu, M. & Farin, G. Using geometric modeling for archiving and
806 searching 3d archaeological vessels. *CISST, Las ...* (2001).
- 807 26. Leymarie, F. F. *et al.* The SHAPE Lab: New technology and software for archaeologists. *Bar*
808 *Int. Ser.* **931**, 79–90 (2001).
- 809 27. Barrile, V., Cacciola, M., Morabito, F. C. & Versaci, M. TEC measurements through GPS and
810 artificial intelligence. *J. Electromagn. Waves Appl.* **20**, 1211–1220 (2006).
- 811 28. Liu, Z. & Sullivan, C. J. Prediction of weather induced background radiation fluctuation with
812 recurrent neural networks. *Radiat. Phys. Chem.* **155**, 275–280 (2019).
- 813 29. Liu, J. Y. *et al.* Pre-earthquake ionospheric anomalies registered by continuous GPS TEC
814 measurements. *Ann. Geophys.* **22**, 1585–1593 (2004).
- 815 30. Nagarajan, A. Explorations into Machine Learning Techniques for Precipitation Nowcasting.
816 *Masters Theses* (2017).
- 817 31. Sathya, R. & Abraham, A. Comparison of Supervised and Unsupervised Learning Algorithms
818 for Pattern Classification. *Int. J. Adv. Res. Artif. Intell.* **2**, 34–38 (2013).
- 819 32. Ghahramani, Z. Unsupervised Learning. in *Advanced Lectures on Machine Learning. ML 2003.*
820 *Lecture Notes in Computer Science* (eds. Bousquet, O., von Luxburg, U. & Rätsch, G.) (Springer,
821 2004).
- 822 33. Hörr, C., Lindinger, E. & Brunnett, G. Machine learning based typology development in
823 archaeology. *J. Comput. Cult. Herit.* **7**, (2014).
- 824 34. van der Maaten L. *et al.* Computer vision and machine learning for archaeology. *Proc. Comput.*
825 *Appl. Quant. Methods Archaeol.* 112–130 (2006).
- 826 35. Oonk, S. & Spijker, J. A supervised machine-learning approach towards geochemical predictive

- modelling in archaeology. *J. Archaeol. Sci.* **59**, 80–88 (2015).
36. Nitze, I., Schulthess, U. & Asche, H. Comparison of machine learning algorithms random forest, artificial neuronal network and support vector machine to maximum likelihood for supervised crop type classification. *Proc. 4th Conf. Geogr. Object-Based Image Anal. – GEOBIA 2012* 35–40 (2012).
37. Jamuna, K. S. *et al.* Classification of seed cotton yield based on the growth stages of cotton crop using machine learning techniques. *ACE 2010 - 2010 Int. Conf. Adv. Comput. Eng.* 312–315 (2010). doi:10.1109/ACE.2010.71
38. Karasik, A., Rahimi, O., David, M., Weiss, E. & Drori, E. Development of a 3D seed morphological tool for grapevine variety identification, and its comparison with SSR analysis. *Sci. Rep.* **8**, 6545 (2018).
39. Drori, E. *et al.* Ampelographic and genetic characterization of an initial Israeli grapevine germplasm collection. *Vitis - J. Grapevine Res.* **54**, 107–110 (2015).
40. Drori, E. *et al.* Collection and characterization of grapevine genetic resources (*Vitis vinifera*) in the Holy Land, towards the renewal of ancient winemaking practices. *Sci. Rep.* **7**:44463, (2017).
41. Smith, E. R., King, B. J., Stewart, C. V. & Radke, R. J. Registration of combined range-intensity scans: Initialization through verification. *Comput. Vis. Image Underst.* **110**, 226–244 (2008).
42. Smith, H. & Jones, G. Experiments on the effects of charring on cultivated grape seeds. *J. Archaeol. Sci.* **17**, 317–327 (1990).
43. Bouby, L. *et al.* Back from burn out: are experimentally charred grapevine pips too distorted to be characterized using morphometrics? *Archaeol. Anthropol. Sci.* 1–12 (2016). doi:10.1007/s12520-016-0425-x
44. Abder Khalik, K. & van der Maesen, L. J. G. Seed morphology of some tribes of Brassicaceae (implications for taxonomy and species identification for the flora of Egypt). **47**, 363–383 (2002).
45. Mangafa, M. & Kotsakis, K. A New Method for the Identification of Wild and Cultivated Charred Grape Seeds. *J. Archaeol. Sci.* **23**, 409–418 (1996).
46. Bruno, M. C., Pinto, M. & Rojas, W. Identifying Domesticated and Wild Kañawa (*Chenopodium pallidicaule*) in the Archeobotanical Record of the Lake Titicaca Basin of the Andes. *Econ. Bot.* **72**, 137–149 (2018).
47. Pagnoux, C. *et al.* Inferring the agrobiodiversity of *Vitis vinifera* L. (grapevine) in ancient Greece by comparative shape analysis of archaeological and modern seeds. *Veg. Hist. Archaeobot.* **24**, 75–84 (2014).
48. Segarra, J. G. & Mateu, I. Seed morphology of *Linaria* species from eastern Spain: identification of species and taxonomic implications. **135**, 375–389 (2001).
49. Al-Ghamdi, F. A. & Al-Zahrani, R. M. Seed morphology of some species of *Tephrosia* PERS. (Fabaceae) from Saudi Arabia identification of species and systematic significance. *Feddes Repert.* **121**, 59–65 (2010).
50. Lipman, Y. & Daubechies, I. Conformal Wasserstein distances: Comparing surfaces in polynomial time. *Adv. Math. (N. Y.)* **227**, 1047–1077 (2011).

51. Styring, A. K. *et al.* The effect of charring and burial on the biochemical composition of cereal grains: Investigating the integrity of archaeological plant material. *J. Archaeol. Sci.* **40**, 4767–4779 (2013).
52. Weng, J., Cohen, P. & Herniou, M. Camera calibration with distortion models and accuracy evaluation. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**, 965–980 (1992).
53. Fisher, R. Linear Discriminant Analysis. 1936 (1936). doi:10.4018/9781591408307.ch003
54. Zhou, Q.-Y., Park, J. & Koltun, V. Open3D: A Modern Library for 3D Data Processing. (2018).
55. Park, J., Zhou, Q. Y. & Koltun, V. Colored Point Cloud Registration Revisited. *Proc. IEEE Int. Conf. Comput. Vis.* **2017-Octob**, 143–152 (2017).
56. Chen, Y. & Medioni, G. chen-medioni-ICP.pdf. *Proceedings - IEEE International Conference on Robotics and Automation* 2724–2729 (1991).
57. Besl, P. J. & McKay, N. D. A Method for Registration of 3-D Shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**, 239–256 (1992).

Acknowledgments

This work was funded by the Israeli Science Foundation grant 551/18 to E.W., E.D., and A.K., and by the Israeli Ministry of Science, Technology & Space grant 3-16808

Author Contributions

All authors made significant contributions to the manuscript. V.L. processed the scanned seeds image data, developed and managed the mathematical analysis, prepared most figures, and wrote part of the manuscript. Y.S. collected part of the seed collection, developed the scanning method, carried out the scanning of seeds, and wrote part of the manuscript. M.D. collected and cleaned part of the seed collection. E.W. participated in the conceiving and the revising of the manuscript. A.K. helped develop the upscale method. Y.R. and E.D. conceived and designed the work, helped develop the algorithm, analyzed the data and results, and wrote the manuscript.

Competing interests

The authors declare no competing interests.

Figures

Figure 1

The procedure used for grape pip 3D data acquisition. A- High-quality focused image; B- 3D image scan, and C- 3D points cloud representation of four varieties' fresh pips. 9003, 292, and 13 - Israeli endogenous varieties and Cabernet Sauvignon - international variety.

Figure 2

Classification of fresh pips, using a train set of fresh pips. The actual variety label of the pip is presented as a ring, and a dot marks the predicted label. A - Visualization of LDA separability learned during a training set with 10 out of 15 pips in each class; B - Visualization of LDA classification of the remaining five pips (out of the 15), in each class as a test set (accuracy 90.00%); C - Accumulative accuracy of LDA classification after performing 100 training and 100 independent test evaluations for each pip. At each run, the same ten representative pips from the assemblage were selected as a constant training group, and the other five pips were classified independently; D - Classification accuracy histogram - 100 training [10 samples] and tests [5 samples], gaining mean accuracy of 95%. Cabernet Sauvignon* - 1024.

Variety 13

Variety 292

Figure 3

acquisition of 3D data for charred grape pips. A- High quality focused 3D image; B- height map, and C- points cloud of two varieties of charred pips.

Figure 4

Classification of charred pips using fresh pips as a train set. The actual variety label of the pip is presented as a ring, and a dot marks the predicted label. A - LDA training - Fresh samples, 13 from each variety (1024=Cabernet Sauvignon); B - LDA testing - Burned samples, 13 from each variety (accuracy 88.46%); C - Accumulative distribution of LDA classification after running 100 tests. At each run, 13 representative pips from the assemblage were selected as the constant training group, and the 13 burned pips were classified independently; D - Classification accuracy histogram - 100 training [13 fresh samples] and independently 100 tests for each pip [13 burned samples].

Figure 5

classification of charred pips using a charred training set. The actual variety label of the pip presented as a ring, and the predicted label marked by a dot. A - LDA training - burned samples, 10 out of 15 from each class (1024=Cabernet Sauvignon); B - LDA testing - Cleaned burned samples, the remaining independently five from each variety (100.00% accuracy); C - Accumulative distribution of LDA classification after running 100 tests. At each run, ten pips from the assemblage were selected as the constant training group, and the other five pips were classified. D - Classification score histogram - 100 training [10 burned samples] and independently 100 tests for each pip [5 burned samples].

Figure 6

Classification of charred pips from 8 varieties. A - Classification accuracy (y-axis) over 100 iterations on each test sample (x-axis) .B - Classification accuracy score histogram of 8 varieties - 100 training [10 charred samples] and independently 100 tests for each pip [5 charred samples].

Figure 7

Classification of charred pips from 8 varieties by the “tournament” flow method. A – Statistic of classifier A – trained to distinguish between 292,9003,13,1024 charred classes, B – Statistic of classifier B – trained to differentiate between 98, 192, 236, and 276 varieties in charred form. D - Accumulative distribution (y-axis) of LDA classification of 8 varieties of the “tournament” flow method, after running

100 iterations for each test pip sample (x-axis). E – Classification accuracy score histogram of 8 varieties – 100 iterations for each independent test sample.

Figure 8

General demonstration of a “tournament” classification flow and its internal stages of training and classifications down to the final identification result. The test sample may be an unidentified charred grape pip, recent or archaeological. The whole reference population, which is believed to include the tested sample’s variety, is divided into random groups of four varieties, trained as separated machines. At the end of the first step, the selected variety of each machine will be elevated to the next stage, where again, the newly established set will be randomly divided into groups of four varieties, and so on, until the final stage will recognize the most probable identification result.

(A)

(B)

Figure 9

Train and test matrices (each index inside the matrices represents the MSE value reported by the ICP algorithm between pip points cloud $\mathcal{P}_i$ and pip point cloud $\mathcal{P}_j$) A Train matrix (of ICP scores with 40x40 dimensions 10 constant representative pips (y axis and x axis) from each class of 292f,9003f,13f and 1024f classes B Test matrix (of ICP scores with 20x40 dimensions 5 remaining pips (y axis) from each class of 292f,9003f,13f and 1024f classes and the 10 constant representative pips (x axis)

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation.docx](#)