

Artificial Intelligence-Based Low-Terahertz Imaging for Archaeological Shards' Classification

Flora Zidane¹, Vanna Lisa Coli, Jérôme Lanteri², *Member, IEEE*, Julien Marot, Laurent Brochier, Didier Binder, and Claire Migliaccio³, *Member, IEEE*

Abstract—In order to map the migration and introduction of farming into Europe during the seventh and sixth millennia Before Common Era, archeologists have made a connection between the study of pottery and farming migration. We are interested here in the classification of pottery into coiling and spiral types based on their manufacturing techniques. To distinguish between these two techniques, we look for the lines formed by air bubbles embedded in the pottery samples. Current methods make use of bulky systems, such as computerized tomography scanners or synchrotrons. Microwave acquisition and processing offer an interesting alternative, due to the possibility to have compact and portable systems. In this article, we investigate the classification of pottery based on low-terahertz measurements in the D-band. We process the measurements with 3-D fast Fourier transform. The resulting matrix is classified with an artificial neural network, multilayer perceptron, which is optimized with the gray wolf optimizer, a bioinspired algorithm. The first results show that the accuracy reaches up to 99% using all the acquired spatial and frequency measurements. Then, we optimize the millimeter-wave (mm-Wave) measurement system with a critical criterion on accuracy in two different scenarios. In the first scenario, we reduce the spatial acquisition but maintain the wideband operation and the results show that the accuracy is between 85% and 96%. In the second one, we reduce the spatial acquisition and use a single frequency. For this second scenario, we achieve a classification accuracy, which is between 77% and 100%.

Index Terms—3-D fast Fourier transform (3-D FFT), artificial neural network (ANN), bioinspired optimization algorithm, computerized tomography (CT) scan imaging, gray wolf optimizer (GWO), low-terahertz (low-THz) imaging, low-THz measurements.

I. INTRODUCTION

ARCHAEOLOGY responds to the fundamental need of human beings to understand their origins, their traditions, and their culture. The invention of writing, between the third and fourth millennia Before Common Era (BCE), was a fantastic breakthrough that enabled information to be transmitted to future generations. However, long before writing, rock art

testified to the human will to leave traces of places and ways of life. Other ways of investigating the evolution of civilizations and traditions prior to the invention of writing include the study of architecture and settlement patterns, ceramics and stone tool technology, food production, and paleogenetics. We can therefore trace the migration of humans through time and space. The introduction of farming into Europe during the seventh and sixth millennia BCE is still poorly understood. This is a fundamental issue since agriculture is the first step toward a sedentary lifestyle, which shapes landscapes, cities, and new social behaviors. Due to the study of large datasets, archeologists have shown that, after a formative stage in the Aegean, the introduction of farming into western Europe followed two paths: the Danube route and the Mediterranean one. The Central European route is characterized by pottery manufactured by coiling, whereas the Mediterranean one is characterized by spiral pottery techniques. In Central Europe, within the Danubian world, the pottery manufacturing techniques based on coiling display a limited variability, which points to a large degree of homogeneity of cultural traditions originating from the Aegean and the Balkans. In contrast, in the Central and Western Mediterranean, where the Neolithic dispersal is partly based on early sailing [1], two different methods of pottery manufacturing were identified: the first, on the Adriatic side, uses coiling techniques and is similar to pottery from the Balkans and Central Europe, but the second, on the Tyrrhenian side, uses spiraled patchwork techniques (SPTs) whose origins are not yet identified [2]. Finding intact pieces of pottery dating back more than 8000 years is very rare, almost miraculous. Hence, archeologists have to work with pottery samples, which we refer to as shards throughout this article. Unfortunately, the characterization of fabrication techniques using macro-traces, which are visible on the fragments, is highly challenging. One way to determine the manufacturing method of pottery shards is to trace the air bubbles, known as “pores,” that were embedded in the pottery during the fabrication process. A linear distribution of pores corresponds to the superimposition of coils, while a curvilinear distribution indicates the use of SPT. Thus, we are looking for small features and require a high-resolution imaging system. Current methods make use of computerized tomography (CT) scans [3] or synchrotron imaging [4]. The use of CT scans in archeology has several applications, including autopsies on mummies [5], 3-D reconstruction and visualization of material and human remains [6]–[8], and deciphering the manufactur-

Manuscript received 7 July 2021; revised 7 June 2022; accepted 15 June 2022. Date of publication 8 July 2022; date of current version 8 September 2022. (Corresponding author: Flora Zidane.)

Flora Zidane, Jérôme Lanteri, Laurent Brochier, and Claire Migliaccio are with Université Côte d’Azur, CNRS, LEAT, UMR 7248, 06903 Sophia Antipolis, France (e-mail: flora.zidane@univ-cotedazur.fr).

Vanna Lisa Coli and Didier Binder are with Université Cote d’Azur, CNRS, CEPAM, UMR 7264, 06300 Nice, France.

Julien Marot is with Aix Marseille Université, CNRS, Institut Fresnel, UMR 6133, 13013 Marseille, France.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TAP.2022.3189553>.

Digital Object Identifier 10.1109/TAP.2022.3189553

0018-926X © 2022 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

ing techniques of Neolithic clay artifacts [9]. The strengths of CT scan imaging are its super-resolution, and low noise and distortion: even the smallest details can be identified. However, a CT scan requires a bulky system, which is expensive. As an alternative to CT scans, microwave systems can be mobile and compact at a lower cost. Like the CT scan, microwave imaging is a nondestructive evaluation and testing (NDE-NDT) method that can penetrate objects without changing any of their intrinsic properties. In general, microwave imaging is sensitive to contrasts in dielectric permittivity, so it is relevant to this application in archeology, where we wish to trace air bubbles in ceramic materials. The choice of the frequency band is of primary importance with respect to the intrinsic resolution and penetration depth: high frequencies provide the highest lateral resolution but a poorer penetration into the material. We are looking for tiny, superficial features and therefore choose to work in the low-terahertz (low-THz) frequency region, more precisely in the D-band. Detailed information on the nature or state of the object can be determined by solving an inverse problem but requires a high computational load. Since we are looking for a simple answer, that is, whether the manufacturing technique was coiling or spiral, and not for the exact value of the permittivity, we process the measured data with machine learning algorithms (MLAs) instead of solving a complex inverse problem. We have previously used this approach in [10], which is a proof of concept of the NDT method to sort healthy apples from damaged ones using millimeter-wave (mm-Wave) imaging and support vector machine (SVM) with 94% accuracy. In this previous work, we also show that we speed up the characterization of the object under test. In [11], we went further into the matter while sorting mm-Wave images of fruits (apples and peaches) by using a multiclass SVM to discriminate the nature and state of fruits with an accuracy of 98%. Although mm-Wave imaging exhibits poor resolution compared to CT scan imaging, many studies combining a microwave system with MLA have shown the potential of this approach for a large set of applications, for instance, in forward/inverse (EM) scattering problems [12]–[15] or for target detection [16], [17], solving medical issues [18]–[20], antenna modeling [21], [22], and detection of buried objects [23], [24]. In the last two decades, artificial neural networks (ANNs) have attracted increasing interest and have been adapted in various applications, such as antenna design [25], [26], RF and microwave design [27], microwave transistor noise modeling [28], EM-based optimization of microwave circuits [29], and locating fault elements in antenna arrays [30]. For performance enhancement, recent studies have tended to replace SVM with ANNs, for example, to classify human activities (walking, running, and so on) [31], [32] or for microwave image reconstruction [33], [34], including free-space estimation for self-driving systems using mm-Wave radar [35] and also security and surveillance solutions through convolutional neural networks (CNNs) [36], [37].

In this article, we introduce a new imaging modality for archeological classification of spiral and coiling shards. We are making use of low-THz measurements and investigate the optimized MLA solution. The results will be compared with those of CT scan imaging processed with MLA, which is the



Fig. 1. Example of coiling (left) [39] and home-made spiraled (right) potteries.

gold standard. Although it is not the primary focus of this article, we would like to point out that MLA has rarely been combined with CT scan imaging in the archeological domain, though ANNs have been used to determine the gender of skeletal remains [38].

This article is organized as follows. Section II defines the problem we address. First, we explain how we choose the shards, introducing the notion of experimental and archeological shards, and we describe the measurement system and the choice of its settings, with a focus on spatial and frequency diversity. In a second step, we process the measurements with a 3-D fast Fourier transform (3-D FFT) and show the preliminary results of classification with SVM. To improve these results, we switch from SVM to an ANN, as described in Section III. Although ANNs are very efficient, the number of possibilities for setting up their intrinsic parameters, such as the number of neurons in the hidden layer, is quasi-infinite. Thinking of an exhaustive search of all parameters could increase drastically the problem's complexity. To overcome this issue, we adapt a bioinspired optimization algorithm, which typically targets this problem. Finally, we present our complete methodology, which combines a multilayer perceptron (MLP) ANN with the gray wolf optimizer (GWO). Section IV discusses the results when making use of the whole diversity of measurements, while Section V presents the optimization of the low-THz measurement system using a critical criterion on the fitness. The conclusions are drawn in Section VI.

The following notations are used throughout this article. Manifolds are denoted by blackboard bold, \mathbb{A} , matrices by boldface uppercase roman, \mathbf{A} . Vectors are denoted by boldface lowercase roman, \mathbf{a} , and scalars by lowercase or uppercase roman, a , b , or A . The K scalar components of a vector \mathbf{a} are accessed via a_1, a_2, \dots, a_K such that $\mathbf{a} = [a_1, a_2, \dots, a_K]^T$.

II. ANALYSIS OF MEASURED DATA

As explained in Section I, the identification of the manufacturing technique of pottery shards is of crucial importance for tracing the introduction of farming into Europe. There are two manufacturing techniques: the coiling technique, as to be seen in Fig. 1 on the left, and the spiral technique on the right. Starting from the pottery pieces, we aim to find out how they were made, which is typically an inverse problem. Its solution is trivial in case of Fig. 1, and however, the archeological samples found on excavation sites are rarely entire pieces but most frequently small shards. Fig. 2 shows two examples, the coiling sample found in *La Cabre*, South of France, and



Fig. 2. Examples of archeological samples. (a) Coiling shard C. (b) Spiral shard I.

the spiral one in *Abri Pendimoun* on the French South-East coast of the Mediterranean sea. It is noteworthy that the manufacturing using coiling is easy to recognize on Fig. 2(a), with its straight lines, whereas the spiral technique cannot be identified visually. Before sophisticated imaging techniques were available, archeologists addressed this problem by becoming familiar with the different manufacturing techniques and gestures used in the Neolithic age, in the hope that they would leave different prints on the final samples depending on how they were made. To this end, they made their own shards, called experimental shards, using the same geological materials as those found on the excavation sites. This simple process has been the sole solution for many years, but its success rate is limited for several reasons, the most obvious one being the deterioration of real archeological shards that are several millennia old. One useful stable characteristic over time is the shape of the lines of air bubbles (pores) that were formed during the assembling of the pottery. Of course, we are looking for tiny details, which explains why archeologists are in need of high-resolution imaging systems, such as the CT scan. The pores are distinctly visible on a CT scan. However, CT scanners are bulky and expensive systems and so are rarely dedicated to a single research department, which limits their availability for archeological experiments. It is therefore of great interest for archeologists to have an alternative imaging technique. Due to its natural high resolution within the microwave spectrum, low-THz imaging is a potential candidate for this application, provided that we define the relevant measurement and processing schemes as discussed throughout this article.

Prior to the measurements, we present hereafter the shards under investigation. The archeologists of the Cultures et Environnements Préhistoire, Antiquité, Moyen Âge (CEPAM Laboratory), Université Côte d'Azur, Sophia Antipolis, France, provided us with the 15 samples in Figs. 3, six of which were archeological and nine of which were experimental. All archeological shards were found in *La Cabre* and *Abri Pendimoun*. They include one coiling and five spirals. Among the nine experimental shards, three are coiling and six are spiral samples. Table I summarizes the materials of the shards and Table II lists the relative permittivity (ϵ') values of the different materials [40]. A range of values is given instead

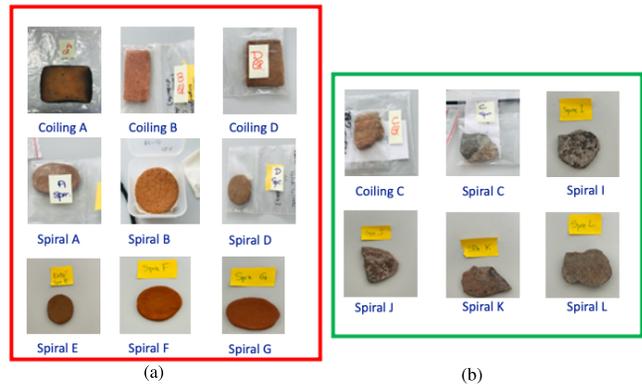


Fig. 3. Pictures of the shards sorted by type. (a) Experimental shards. (b) Archeological shards.

TABLE I
SHARDS SORTED BY TYPE AND MATERIALS

Samples	Type	Materials
Coiling A	experimental	Volcanic material
Coiling B	experimental	Granites, quartz, mica and feldspar
Spiral B	archaeological	
Spiral C	archaeological	Rhyolite (rich in silica and quartz, feldspar)
Spiral I	archaeological	
Coiling C	archaeological	Granites, quartz, mica, feldspar and glauconite
Spiral D	experimental	
Spiral A	experimental	
Spiral H	archaeological	
Spiral J	archaeological	Clay
Spiral K	archaeological	
Spiral F	experimental	Clay
Spiral G	experimental	
Coiling D	experimental	Glauconite
Spiral E	experimental	

TABLE II
RELATIVE PERMITTIVITY OF MATERIALS

Materials	Relative permittivity ϵ'
Clay	2-6
Feldspar	5-7
Glauconite	12.7
Granite	5-6.5
Mica	6-6.5
Quartz	4.4-4.5
Rhyolite	5.46
Volcanic material	6.68-7.68

of a single value for some materials (e.g., clay) because the permittivity is strongly dependent on the moisture content. However, considering that the shards are kept in a dry place, the permittivity is closer to the lower value of each estimated range. Low-THz imaging is sensitive to the dielectric contrast in general, and thus, we aim to detect the contrast between the air bubbles and the geological materials.

A. D-Band Measurement Settings

Measurements were conducted with the 3-D scanner installed and developed in the Laboratoire d'Electroniques Antennes et Telecommunications (LEAT) for measurements at mm-Wave and above, used in its monostatic configuration [41]. The complete setup is shown in Fig. 4. The shards were placed on a Rohacell tower whose relative permittivity is 1.02 at 1 GHz but which is not transparent to low-THz frequencies. This will be considered when choosing the postprocessing. The probe antenna is a WR7 standard gain horn that rotates around the shard with a radius of 585 mm. The system operates

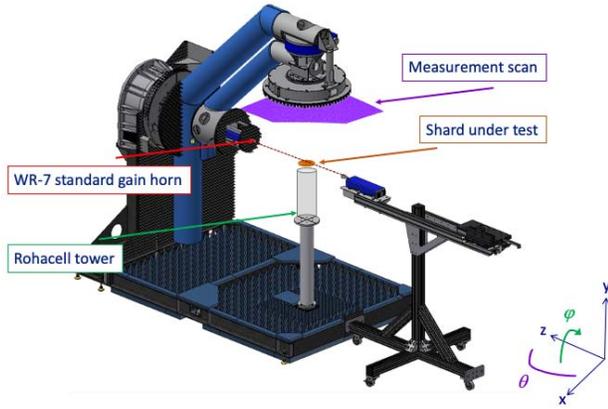


Fig. 4. Measurement setup.

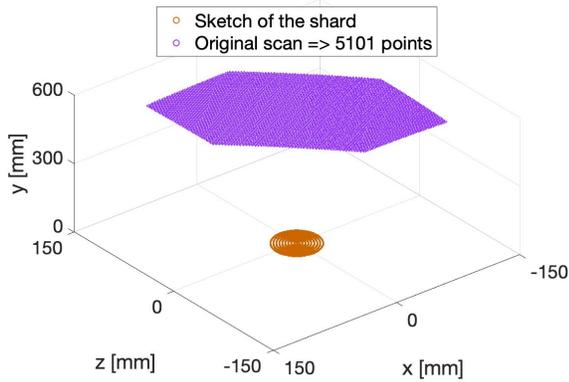


Fig. 5. Measurement scan and measurement points.

TABLE III
MEASUREMENT SETTINGS

Parameter	Range	Step
Frequency [min, max]	[110,170] [GHz]	10 [GHz]
Angle θ [min, max]	[80,100] [°]	0.2 [°]
Angle φ [min, max]	[80,100] [°]	0.2 [°]

in the far field, which contributes to the ill-posedness of the inverse problem since we cannot measure the evanescent waves. To solve this problem, we need to increase the content of information and therefore introduce spatial and frequency diversity. We have already worked on W-band measurement diversity for the detection of defects in fruits [10], [11]. However, this measurement scheme cannot be strictly transposed to the present application because we are looking for tinier features. The pores result from an aggregation of air bubbles. They are, for instance, negatives of carbonized plants. The maximum size of a pore can reach up to 1 mm, whereas some aggregates have a size of few micrometers. We, therefore, move to a higher frequency band, the D-band, and use the whole available frequency bandwidth, i.e., 60 GHz, for better range resolution. Spatial diversity is obtained by moving the probe antenna above the shard, as shown by the purple area on the sketch in Fig. 5. Table III sums up the measurement settings. Angles (θ, φ) are the classical spherical coordinates. Note that Oy is the vertical axis of the measurement setup, and hence, measurements above the shard are centered around $(\theta, \varphi) = (90^\circ, 90^\circ)$. The term “sample” bears several significations throughout this article.

- 1) *The Shard Sample*: A shard sample is a piece of pottery. We have a total of 15 shard samples. A–D work in pairs so we have two shard samples in A, B, C, and D. From E to K, we have one shard sample.
- 2) *The EM-Field Sample*: An EM-field sample corresponds to one spatial measurement point. According to the range of the scan as shown in Fig. 5 and to the scan step, we have a total of 5101 spatial measurements points per shard sample.

The total time for measuring one shard at all frequencies over an area, whose maximal range variation in (θ, φ) is $(20^\circ, 20^\circ)$, is 2.20 h and 20 min. Since our objective is to build a dataset for postprocessing, we have to tailor the measured data into classifier samples according to the following criteria.

- 1) Every shard sample must have several classifier samples in the dataset.
- 2) The classifier samples must reflect the frequency and spatial diversity of the measurements.
- 3) Redundancy between the classifier samples is an advantage, in particular for removing the noise and/or unwanted fixed objects such as the Rohacell tower.

Hence, we split the measurement area into small patches that will be the classifier samples of the dataset. The values given in Table III are given as follows: the scan step is 0.2° and the values of θ and φ are, respectively, the minimal and maximum values used in the scan. However, we do not scan continuously between the minimal and maximal values and the scan that we do here corresponds to the purple zone in Fig. 5, which is not a “square” scan. This scan has 5101 measurement points. The samples of the dataset correspond to a sliding patch of size $10^\circ \times 10^\circ$, i.e., 51×51 measurement points. This patch slides one measurement point over the diagonal leading to 51 patches over the purple area in Fig. 5, each of which has measurements at seven frequencies. Thus, we have $51 \times 7 = 357$ classifier samples per shard in the dataset.

The dataset comprises all shard samples measured over two measurement campaigns staggered in time, which is an advantage for testing the robustness of the whole process over time.

- 1) *First Campaign*: July 2019—measurements of shards A, B, C, and D (coiling and spiral).
- 2) *Second Campaign*: November 2020—measurements of shards E, F, G, H, I, J, and K (spiral).

Throughout this article, the training dataset, denoted \mathbf{X}_{tr} , contains the shard samples of AB, which includes two coiling and two spiral shards. This choice is the most useful from the user’s point of view because it includes only experimental shards, so we completely master the dataset. The cardinality of any training dataset or test dataset composed of any combination of shard samples can be easily deduced from the number of shard samples in the training or test dataset. For instance, the training dataset “AB,” which includes the pairs of shard in A and in B, has four shard samples, and the number of classifier samples is $4 \times 357 = 1428$. However, one has to remember that a classifier sample has 51×51 measurements points (complex values of S_{11}) so that the cardinality of the “AB” training dataset belongs to $\mathbb{C}^{1428 \times 51 \times 51}$.

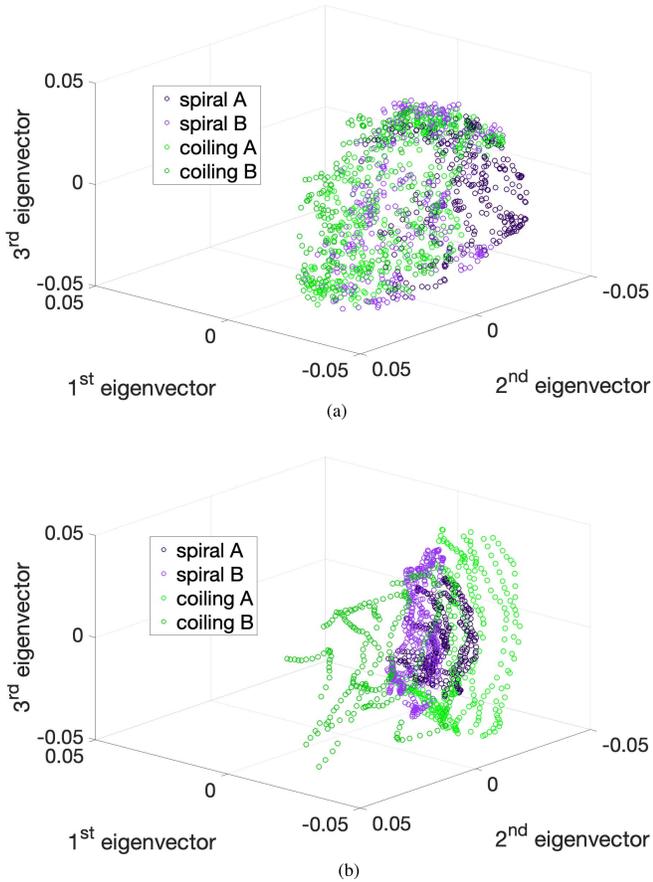


Fig. 6. Training dataset AB: comparisons of the three most significant eigenvectors obtained with the PCA of low-THz measurements preprocessed with 2-D FFT. (a) With binarization. (b) Without binarization.

B. 3-D FFT: mm-Wave Preprocessing and Corresponding SVM Classification Results

The gold standard for the classification of archeological shard samples is the result that we obtained with CT-scan measurements and the SVM. For the training, we consider the shard samples of AB, the other shard samples being in the test dataset. We reached an accuracy of 100% for every tested shard.

To start with the low-THz measurements, we used the same process as in [10] and [11], that is, the segmentation with an Otsu threshold [42] followed by a nonlinear SVM [43]. The classification accuracy drops to 59% and 70%. As the segmentation “kills” the small items of the images that are relevant for our application, we therefore decided to suppress the segmentation stage and process the images obtained after the 2-D FFT directly. Note that the accuracy is obtained by calculating the proportion of well-classified samples. More details are given in Section III-D. To illustrate the distribution of the classifier samples in the dataset, Fig. 6 shows the projection of the training dataset AB on the three most significant eigenvectors after using a principal component analysis (PCA) [44]. Results with [Fig. 6(a)] and without [Fig. 6(b)] binarization are shown for comparison purpose. Note that these three eigenvectors explain 90% of the total variance for the binarized scenario and 80% otherwise. With segmentation (i.e., with binarization), the classifier samples are completely

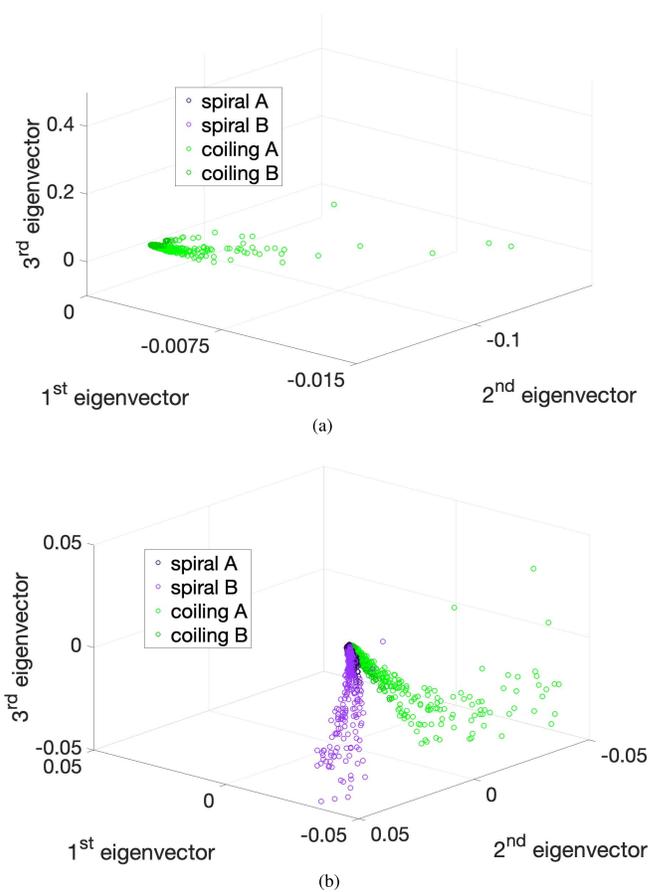


Fig. 7. Training dataset AB: comparisons of the three most significant eigenvectors obtained with the PCA of low-THz measurements preprocessed with 3-D FFT. (a) With binarization. (b) Without binarization.

mixed. Without segmentation, the samples of the shards are still superimposed, but this time they are gathered in distinct groups [Fig. 6(b)]. To separate them, we apply an FFT along the first dimension of the training dataset, i.e., along the 1428 classifier samples. Considering the 2-D FFT performed on the two other dimensions (51×51) of the measured data, the whole process is equivalent to a 3-D Fourier transform. From a practical point of view, we implement the 3-D FFT as follows.

- 1) We gather the measurement points into a 3-D matrix $\mathbf{M} \in \mathbb{C}^{1428 \times 51 \times 51}$.
- 2) We downsize $\mathbf{M} \in \mathbb{C}^{1428 \times 51 \times 51}$ to $\mathbf{M}_d \in \mathbb{C}^{1428 \times 2601}$.
- 3) We compute the 2-D FFT of \mathbf{M}_d .
- 4) \mathbf{M}_d is the new training dataset.

As shown in Fig. 7, which represents the projection of the data on the three most significant eigenvectors of the new training dataset, coiling shards (in light and dark green) are separated from the spiral ones (in light and dark purple) [Fig. 7(b)], whereas the 3-D Fourier transform is not efficient on the binarized images [Fig. 7(a)]. These three eigenvectors explain 99% of the total variance for the binarized scenario and 49% otherwise.

C. Test Dataset

We now precede the test dataset, in which the classifier has to assign the proper class (positive or negative) to unknown

TABLE IV
SVM CLASSIFICATION WITH 3-D FFT AND BINARIZATION
OMITTED ON mm-WAVE MEASUREMENTS

Training dataset	Shards under test	Accuracy (%) with 3D-FFT mm-Wave measurement	Accuracy (%) with CT scan images
<i>AB</i>	<i>C</i>	84.45%	100%
<i>AB</i>	<i>CD</i>	91.5%	100%

classifier samples. To do this, we have to go back to the interpretation of 3-D FFT and more precisely the FFT along the dimension of the samples. The latter simplifies the separation of the coiling from spiral shards, which implicitly supposes that both types of shards samples are present in the dataset. The approach we propose and describe hereafter is inspired by the historical method used by archeologists, whose experimental shards help identify unknown shards. We ensure that both categories of shard samples are present in the dataset by introducing one experimental coiling and one experimental spiral shard in the test dataset. The experimental shards are used as a reference, namely, the coiling *D* and the spiral *D*, and ensure the natural separation with FFT along the dimension of the samples. We recall here that the reference shards are not labeled, which means that we do not assign them to the positive or negative class prior to the classification. Finally, for every test, we built a test dataset with three samples: the two experimental shards *D* and the shard samples to be tested. To the best of our knowledge, this approach inspired by traditional archeological methods, consisting in using unlabeled reference samples in the test dataset for classification purposes, has never been used before. Fig. 8 shows the projection of the results for shards samples spiral C and spiral E after the 3-D FFT. As the training dataset which the coiling and the spiral shards samples are separable, we processed the training dataset with a linear SVM. At this stage, we consider only shard *C*, including shard *D* as a reference for the purpose of comparison with the results obtained in [45]. The results are presented in Table IV. Note that the accuracy is computed using only the classifier samples of *C* shards (the first line of Table IV), as it is the shard being classified. If we compute the accuracy using all the samples of the test dataset, i.e., including the samples of shards *D*, the accuracy reaches 91% (the second line of Table IV). By removing the segmentation step and directly processing the 3-D FFT, we improved the accuracy, though it remains below the accuracy reached with CT scan images. Consequently, we propose to replace the SVM with an ANN of MLP-type MLP, which has been shown to outperform the SVM for complex datasets [46].

III. ANN CLASSIFICATION: METHODOLOGY

The new methodology falls into three blocks.

- 1) The low-THz measurements are described in Section II-A.
- 2) The 3-D FFT is described in Section II-B.
- 3) Classification with an MLP ANN and optimization with the GWO method are defined in Sections III-A and III-B, respectively.

A. Multilayer Perceptron

The MLP is a widespread ANN architecture for classification in use today. The MLP is composed of several layers,

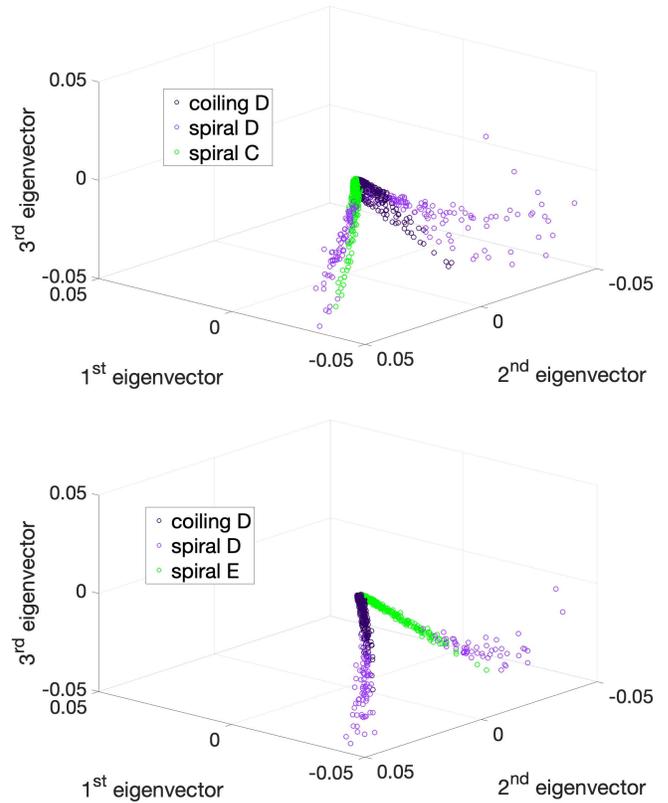


Fig. 8. Test datasets for spiral shard sample of *C* and *E* including the reference shard samples of *D* for separation purposes: the three most significant eigenvectors obtained with the PCA of mm-Wave measurements preprocessed with 3-D FFT.

as shown in Fig. 9. The first layer is called the input layer. The input layer represents the input feature and the number of neurons in this layer is equal to the dimension of samples in the training dataset. The last layer is the output of the network. The number of neurons in this layer is equal to the number of classes for a multiclass problem and to one neuron for a binary problem.

The layer between input and output layers is called the hidden layer. The hidden layer contains m neurons. The higher the number of neurons, the more complex the network becomes. Each neuron in the MLP has two purposes: the summation and the activation function, also called the transfer function. The connection is unidirectional and follows the path from the input to the output layer [47], [48]. To go from the input layer to the hidden layer, the summation function adds the weighted inputs of the neuron as given in (1). For $j = 1, \dots, m$

$$S_j = \sum_{i=1}^n \omega_{ij} I_i + \beta_j \quad (1)$$

where S_j is the output of neuron j , n is the number of input neurons, ω_{ij} is the weight between input neuron i and hidden neuron j , I_i is the contribution of input neuron i , and β_j is bias term associated with hidden neuron j .

Equation (1) shows that S_j is a weighted combination of the input features. The weighting coefficients ω_{ij} are estimated during a so-called “learning phase,” as described next. Then, S_j is provided to the activation function. This function plays

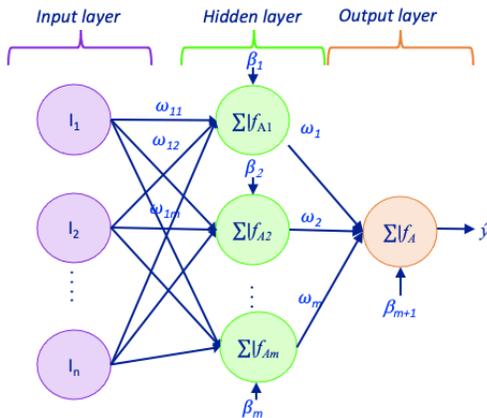


Fig. 9. MLP architecture and parameters.

the role of a threshold, at which the neuron will emit a signal. There exist several activation functions and the most popular is sigmoid, which is a nonlinear activation function, as given in (2). For $j = 1, \dots, m$

$$f_{A_j} = \frac{1}{1 + e^{-s_j}}. \quad (2)$$

The final output can be calculated as follows:

$$\hat{y} = \beta_{m+1} + \sum_{j=1}^m \omega_j f_{A_j} \quad (3)$$

where m is the number of neurons in the hidden layer and ω_j is the connection weight between the hidden neuron j and the output neuron; β_{m+1} is a bias term.

Learning Phase: It determines the weight values for each neuron in the network in order to bring the output of the network as close as possible to the desired output. Initially, the weights are assigned randomly, while the training dataset of features is in the input layer. The difference between the desired and the predicted output is calculated at every iteration and the network readjusts the weights to converge toward the solution. This is typically an optimization process for which we use the widespread backpropagation (BP) algorithm based on the gradient descent algorithm [49]. Note that this BP algorithm is distinct from the BP commonly used in electromagnetism, although the terminology is the same.

There is no single MLP architecture that addresses all classification issues. Instead of searching one by one for the best MLP architecture, we use an optimization method. Among the numerous possibilities, we chose the GWO because of its ability to avoid being trapped in a local minimum. Moreover, it has demonstrated its efficiency for training the weights of MLP [50]. Here, we plan to use it for a different purpose, namely, to optimize the number of neurons in the hidden layer.

B. Gray Wolf Optimizer

The GWO is a nature-inspired optimizer based on the observation of the social life of gray wolves in nature [51]. The GWO algorithm simulates the common behavior and hunting strategies of gray wolves in their environment. It is an agent-based algorithm where each search agent corresponds to

a wolf. The gray wolves have a strict social hierarchy. There is one leader which is called “the alpha” α . The second level in the hierarchy is formed by the beta β wolves, which are the subordinates of the α . Then, the δ wolves are subordinate to α and β but superior to the ω wolves, which comprise the remainder of the wolf pack. Mathematically, α represents the best solution; β and δ are the second and third best solutions, respectively, and Q is the total number of wolves, which are called search agents in the following. The GWO algorithm mimics the encircling and prey killing mechanisms. Its mathematical implementation is given as follows.

We assume that P parameters should be estimated: K_1, K_2, \dots, K_P , where $P \geq 1$. The following notations will be used.

- 1) P is the number of expected parameters, which are indexed with i .
- 2) iter denotes one iteration and T_{\max} denotes the maximum allowed number of iterations.
- 3) $f()$ is the function to be minimized, also called the criterion. It depends on the P parameters mentioned above.
- 4) $\mathbf{x}_q(\text{iter})$ is a vector corresponding to one search agent $q = 1, \dots, Q$, at iteration iter . It takes the form of a vector with a P -tuple of tested values $\mathbf{x}_q(\text{iter}) = [K_1, K_2, \dots, K_P]^T$.

The seminal GWO searches a continuous space. As mentioned above, among the search agents, there are three leaders α , β , and δ . All other agents are the ω wolves. The updated position is calculated as

$$\mathbf{x}_q(\text{iter} + 1) = \frac{1}{3}(\mathbf{y}_{\alpha,q}(\text{iter}) + \mathbf{y}_{\beta,q}(\text{iter}) + \mathbf{y}_{\delta,q}(\text{iter})). \quad (4)$$

It results from the equal contribution of the α , β , and δ wolves. These contributions are computed as follows, for instance, for α :

$$\mathbf{y}_{\alpha,q}(\text{iter}) = \mathbf{x}_\alpha(\text{iter}) - \mathbf{b} \cdot \mathbf{d}_{\alpha,q}(\text{iter}) \quad (5)$$

with $\mathbf{d}_{\alpha,q}(\text{iter}) = |\mathbf{c} \cdot \mathbf{x}_\alpha(\text{iter}) - \mathbf{x}_q(\text{iter})|$.

The vectors \mathbf{b} and \mathbf{c} are calculated as $\mathbf{b} = 2\mathbf{a} \cdot \mathbf{r}_1 - \mathbf{a}$ and $\mathbf{c} = 2 \cdot \mathbf{r}_2$. In these expressions, the components of vector \mathbf{a} are all equal to a , a scalar value that is a key parameter in the algorithm. The value of a decreases from 2 to 0 during the iterations. Vectors \mathbf{r}_1 and \mathbf{r}_2 have random components between 0 and 1, generated from a normal distribution.

During the hunt, the wolves first diverge from each other to search for the prey or, equivalently, to encircle it. Second, they converge to kill the prey. This is mathematically modeled through the deterministic vector \mathbf{a} . When $a > 1$, the search agents are obliged to diverge from the prey: this is the exploration phase. Conversely, when $a \leq 1$, the search agents are obliged to attack toward the prey: this is the exploitation phase. In the seminal version of GWO [51], the key parameter a decreased regularly from 2 to 0: $a = 2(1 - (\text{iter}/T_{\max}))$, where iter is the iteration index and T_{\max} is the maximum number of iterations. In more recent works, various expressions have been proposed for a such as a quadratic [52] or adaptive [53] function. Whatever the version [52], [53] the exploration phase

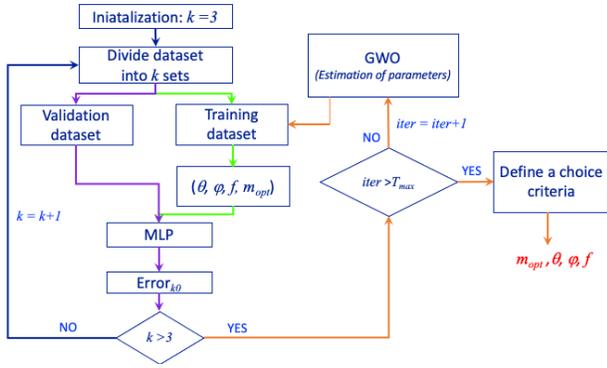


Fig. 10. Training workflow.

lasts until $a = 1$, then the exploitation phase lasts from $a = 1$ to $a = 0$.

C. Learning and Optimization Steps

Many theoretical and experimental works have shown that a single hidden layer with a nonlinear activation function is sufficient to approximate a complex nonlinear decision boundary, which separates between two or more classes [54]. Therefore, we choose to work with three layers in the MLP architecture, that is, a single hidden layer. The training dataset is divided into three subsets, on which we apply a threefold cross validation in order to avoid overfitting [55]. Two subsets are used for the training phase and the third one for validation. During validation, we compute an error as defined in (6). Each of the three subsets is used once for validation. At the end, we compute the average error, which is called the fitness [see (7)]

$$error_{k_0} = \frac{1}{N} \sum_{n_0=1}^{N-1} (y_{n_0} \neq \hat{y}) \quad (6)$$

where N is the number of samples

$$fitness = \frac{1}{k} \sum_{k_0=1}^k error_{k_0} \quad (7)$$

with $k = 3$ for threefold cross validation.

As mentioned previously, the GWO algorithm is used to look for the optimal values of the P parameters. In our application, we can optimize up to four parameters.

- 1) The angles of measurement, θ and φ , which correspond to the number of measurement points in the sliding patch, hence the number of neurons in the input layer.
- 2) The frequency f . The number of frequencies influences the number of samples per shard.
- 3) The number of neurons in the hidden layer m_{opt} .

Optimizing each of these parameters is of interest. Decreasing the number of values of θ , φ , and m_{opt} will improve the classification time, while the optimization of f reduces the bandwidth, which decreases the system's complexity.

The training workflow is shown in Fig. 10. We note here that the training dataset contains the low-THz measurements of the experimental coils and spiral shards samples AB after being processed according to Section II-B. The GWO algorithm terminates when it reaches the maximum number of iterations,

TABLE V
DEFINITION OF THE CONFUSION MATRIX CONTENT

Classes	Actual Negative	Actual Positive
Assigned Negative	True Negative (TN)	False Negative (FN)
Assigned Positive	False Positive (FP)	True Positive (TP)

T_{max} , defined before running the algorithm. We also define critical fitness, which is the fitness that we should not exceed. The choice of the optimal solution is made afterward, when considering what we call the “choice criteria” in Fig. 10. The latter is determined by considering the matching of the fitness computed at iteration $iter$ with the critical fitness and the values of the parameters at iteration $iter$ that lead to a minimum number of measurement points and/or frequencies and/or number of neurons in the hidden layer. At this stage, we have optimized the MLP. One of the significant advantages of the GWO is that optimizing one or four parameters does not increase the complexity of its implementation. However, it will increase the optimization time, i.e., the time it requires to reach a given fitness value. In all cases, the optimized MLP is noted MLP* throughout this article. In the following, we first wish to approve the proposed methodology to classify unknown spiral and coiling shards using the full spatial and frequency diversity, i.e., $(\theta, \varphi, \text{ and } f)$ unchanged.

D. Evaluation of the Classifier on the Test Dataset

To judge the performance of the MLP*, we make use of the following metrics: accuracy, sensitivity (SENS), and specificity (SPEC) [56], extracted from the confusion matrix.

- 1) *Confusion Matrix*: Comparison of actual and predicted class. Table V explains the content of the confusion matrix for a binary classification problem.
- 2) *Accuracy*: It computes the proportion of well-classified samples (8)

$$Accuracy = \frac{TN + TP}{TN + TP + FN + FP}. \quad (8)$$

- 3) *Sensitivity (SENS)*: It computes the probability that the samples are truly positive (9)

$$SENS = \frac{TP}{TP + FN}. \quad (9)$$

- 4) *Specificity (SPEC)*: It computes the probability that the samples are truly negative (10)

$$SPEC = \frac{TN}{TN + FP}. \quad (10)$$

IV. RESULTS KEEPING FULL MEASUREMENT DIVERSITY

As explained in Section II-A, the size of the training dataset AB is $\mathbf{X}_{tr} \in \mathbb{C}^{1428 \times 2601}$, which has to be classified into two classes: the positive class representing the spiral shards and the negative class the coiling ones. We therefore have one neuron in the output layer. Here, we optimize only the number of neurons m in the hidden layer, that is, $P = 1$. We set the maximal number of iterations of the GWO to $T_{max} = 10$, the number of search agents (number of wolves) to $Q = 6$, and the search space for m between 2 and 30. The results obtained by GWO are shown in Fig. 11. The fitness (orange curve)

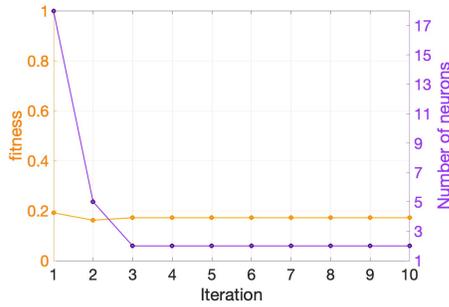


Fig. 11. GWO optimization results: fitness (in orange) and number of neurons (in purple) in the hidden layer.

TABLE VI
EXAMPLE OF RESULT: CONFUSION MATRIX OF COILING C

Shards	Coiling	Spiral
Coiling	354	0
Spiral	3	0

converges as early as the third iteration, which corresponds to a fitness of 0.17157 and a number of neurons in the hidden layer of 2 (purple line). The optimization time is 4 min. Once the number of neurons of the hidden layer is obtained, we compute the final weights ω_{ij} and ω_j to get the optimal MLP*. This step takes 5.53 s. After optimization, we launch the test of the unknown shards. We perform all our classification experiments on a PC, which is equipped with an Intel Core i9-vPro CPU @2.3 GHz, 4.8 GHz Turbo, and a “NVIDIA Quadro T2000” graphics card, including a 4Go GDDR5 and 1024 CUDA cores. In Table VI, we present the confusion matrix of the coiling shard under test. Results are presented in Table VII. The accuracy is improved compared to the results obtained with the SVM in [45] because the worst value is now above 99%. Moreover, we have reached the same accuracy as the CT scan, which was our main goal. Although the measurement time is not really critical to this application since we have no industrial constraint, it is still interesting to reduce it because it increases the measurement accuracy due to the decrease of the shift in phase caused by the drift of the network analyzer [57]. To overcome this issue, we have to minimize the number of acquisition points, as discussed in Section V.

V. OPTIMIZATION OF THE SYSTEM

We aim to optimize the system toward two scenarios. First, we reduce the number of spatial measurement points and keep all frequencies. Second, we set the frequency at the central value of the D-band and try to reduce the number of measurement points. Finally, we discuss and compare the two scenarios.

A. Reduction of the Number of Spatial Acquisitions With Wideband Operation

So far, the measurement time is 2 h and 20 min for each shard scanned over the whole D-band with a frequency step of 10 GHz. As explained in Section II-A, we have 51 sliding patches of 51×51 measurement points each, corresponding to a total number of points of 132 651. As there is an overlap of points between the patches, the real number of measurement points is 5101. Reducing this number will also lower the

TABLE VII
CLASSIFICATION RESULTS FOR ALL SHARDS USING ALL MEASUREMENTS

Shard under test	Test time	Accuracy	SPEC	SENS
Coiling C	3.35ms	99.16%	0.9916	0
Spiral C	3.3ms	99.44%	0	0.9944
Spiral E	1.55ms	99%	0	0.9972
Spiral F	1.6ms	100%	0	1
Spiral G	1.55ms	100%	0	1
Spiral I	1.58ms	100%	0	1
Spiral J	1.58ms	100%	0	1
Spiral K	1.6ms	100%	0	1
Spiral L	3.35ms	99.71%	0	0.997

views and hence increase the ill-posedness of the inverse problem. Therefore, we have to relax the constraints of the optimization and decide to set the critical fitness value to 0.2. Here, we simultaneously optimize the number of points in the sliding patch and look for the optimal architecture of the MLP (MLP*), by estimating the optimal number of neurons m_{opt} in the hidden layer. In order to optimize the number of points per sliding patch, we define a step that corresponds to the removal of points in lines and columns in the 51×51 matrices. While seeking an elevated step value, we decrease the number of values for θ and φ .

This will also automatically reduce the number of sliding patches since we use the same step for the slide and dropping points inside the patch. The settings of the GWO algorithm are now given as follows: $T_{max} = 20$ iterations, and we keep $Q = 6$ search agents. The step ranges between 0.6° and 5° and the search space for the number of neurons in the hidden layer is between 2 and 30.

When these $P = 2$ parameters are optimized, an agent q at iteration iter is given as follows: $\mathbf{x}_q(\text{iter}) = [S_\theta, m]^T$, where S_θ denotes the step between θ (and also φ) values. The cardinality of the training dataset becomes too small to use K-fold cross validation. To prevent underfitting [55], we used the GWO algorithm on 90% of the training dataset and the remaining 10% for the validation. The fitness is computed with (6) and (7). The number of neurons (purple line) and the fitness convergence curve (orange line) with respect to the iterations are shown in Fig. 12(a), and Fig. 12(b) represents the step values (green line) and the fitness convergence curve with respect to the iterations. The optimization time is 34.3 s. Any combination of step values and the number of neurons yields a fitness, which is lower than the critical value of 0.2 (red line). The GWO optimizer reaches the minimum fitness at the sixth iteration. At this stage of the optimization process, $\mathbf{x}_\alpha = [0.6^\circ, 16]^T$. In other words, the optimal value of S_θ is 0.6° and the optimal value m_{opt} of m is 16. However, this is not the best solution in terms of acquisition time. Therefore, we choose the parameters of the third iteration, i.e., the largest step of 0.6° (yielding the smallest number of points in the patch) and 16 neurons in the hidden layer. This corresponds to a fitness value of 0.083. The training time of the MLP* is 0.83 s. The optimized number of measurement points drops to 64, which means that, for each shard sample, we measure only four patches containing 16 measurement points. Removing the overlap between the points, the number of distinct measurement points is 37, while the accuracy remains above 85% for

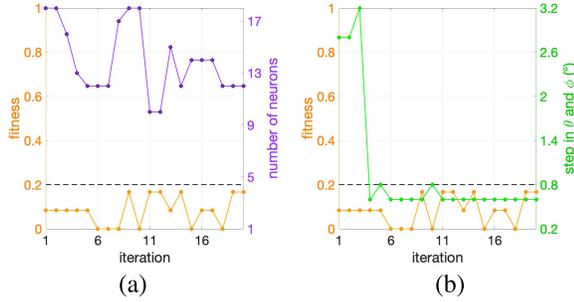


Fig. 12. GWO optimization results: fitness convergence curve in orange, number of neurons in the hidden layer in purple, and spatial step for wideband operation in green. The critical fitness is shown in black. (a) Fitness and number of neurons as a function of iterations. (b) Fitness and step as a function of iterations.

TABLE VIII
CONFUSION MATRIX OF COILING C

Shards	Coiling	Spiral
Coiling	24	0
Spiral	4	0

TABLE IX
CLASSIFICATION RESULTS FOR ALL SHARDS WITH 37 SPATIAL ACQUISITIONS AND SEVEN FREQUENCIES

Shard under test	Test time	Accuracy	SPEC	SENS
Coiling C	0.26ms	85.7%	0.857	0
Spiral C	0.3ms	96.4%	0	0.964
Spiral E	0.1ms	85.7%	0	0.857
Spiral F	0.09ms	89.28%	0	0.8928
Spiral G	0.1ms	96.4%	0	0.964
Spiral I	0.1ms	92.8%	0	0.928
Spiral J	0.1ms	96.4%	0	0.964
Spiral K	0.1ms	92.85%	0	0.9285
Spiral L	0.1ms	92.8%	0	0.928

all shards. As an example, Table VIII shows the confusion matrix of coiling C. By optimizing the measurement points, we also optimize the MLP training and test times, together with the measurement time, which drops dramatically from 2 h and 20 min to approximately 2 min. This is at the expense of the accuracy, which moves from 99% to 92%, on average (see Table IX). We notice that, in Fig. 12, the fitness convergence curve is oscillatory. This is due to the random aspect of the fitness function: the initial weights of the neural network are set randomly, in a different manner for each trial performed in the optimization process. What happens for instance from iteration 8 to iteration 9 is that the score of the best search agent (α) becomes worse, but still the best among all agents. It is then selected as the ninth value in the convergence curve.

B. Reduction of the Number of Spatial Acquisitions With Single-Frequency Operation

In this section, we investigate the ability of the MLP to classify the coiling and spiral shards with the simultaneous reduction of frequency and spatial diversity. As in Section V-A, we set the crucial fitness to 0.2. In contrast to Section V, we work at a single frequency, 140 GHz, which is the central frequency of the D-band. The GWO algorithm is now implemented as follows: we again set $T_{max} = 20$ iterations to estimate the step and the number of neurons in the hidden

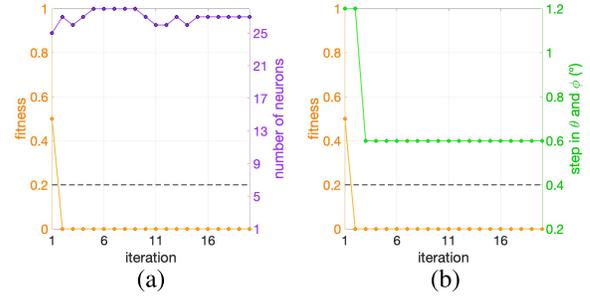


Fig. 13. GWO optimization results: fitness convergence curve in orange, number of neurons in the hidden layer in purple, and spatial step for single-frequency operation (140 GHz) in green. Critical fitness shown in black. (a) Fitness and number of neurons as a function of iterations. (b) Fitness and step as a function of iterations.

TABLE X
CLASSIFICATION RESULTS FOR ALL SHARDS WITH 217 SPATIAL ACQUISITIONS AT 140 GHZ

Shard under test	Test time	Accuracy	SPEC	SENS
Coiling C	0.28ms	100%	1	0
Spiral C	0.41ms	100%	0	1
Spiral E	0.1ms	100%	0	1
Spiral F	0.09ms	88%	0	0.88
Spiral G	0.1ms	88%	0	0.88
Spiral I	0.11ms	88%	0	0.88
Spiral J	0.10ms	77%	0	0.77
Spiral K	0.1ms	100%	0	1
Spiral L	0.1ms	88%	0	0.88

layer, that is, we optimize $P = 2$ parameters, with $Q = 6$ agents. The training dataset should contain sufficient data for learning. Since we have only one frequency, we need to reduce the search space of the step to between 0.6° and 1.6° while keeping that of the hidden neurons unchanged between 2 and 30. The results of the GWO are shown in Fig 13. We choose the largest step, which is $S_\theta = 1.2^\circ$ at the second iteration corresponding to $m_{opt} = 27$ neurons in the hidden layer and a fitness of 0.0. At this iteration, $\mathbf{x}_\alpha = [1.2^\circ, 27]^T$. The total optimization time is 9.95 s. The number of measurement points drops to 729, that is, nine patches of 81 measurement points. Removing the overlap, we have 217 unique measurement points, which corresponds to a measurement time less than 10 min. The MLP* structure has three layers, with 81, 27, and one neuron in the input, hidden, and output layers, respectively. The unknown shard classification results are shown in Table X. All shards samples are within the expected fitness of 0.2 (80% of accuracy) except for *spiral J* at 77% (see Table X).

C. Discussion

In this paragraph, we have pushed the system's parameters (the number of measurement points and frequencies) to their limit while keeping at least 80% accuracy. To give a clear picture of the performance obtained in the classification of archeological shards with an MLP optimized with the GWO, we summarize the results of the different studies in Table XI and illustrate the new measurement scans in Fig. 14. As expected, any simplification of the system (or degraded scenario) that involves a reduction of the spatial or frequency diversity decreases the classifier performance. Starting from a quasi-100% accuracy with dense mapping of the scan area

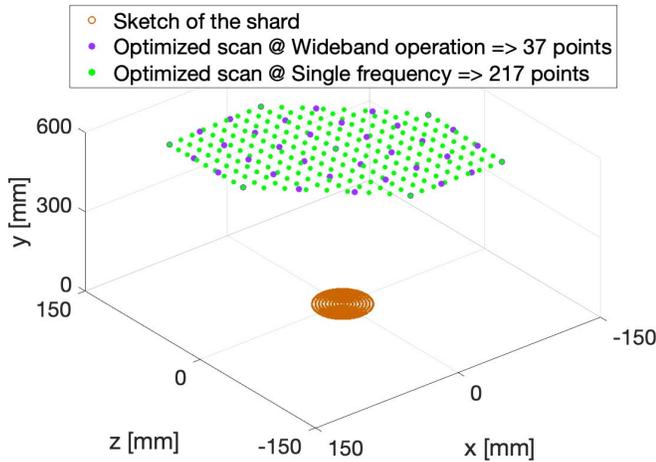


Fig. 14. Comparisons between original and optimized scans.

TABLE XI
SUMMARY OF RESULTS

Characteristics	Full measurements	Optimized scan (wideband)	Optimized scan (single frequency)
Number of frequencies	7	7	1
Patch*Points per patch	51×2601	4×16	9×81
Number of unique points	5101	37	217
Measurement time	2h20'	2'	10'
Accuracy [min-max]	[99 - 100] %	[85.7 - 96.4]%	[77 - 100]%

and the frequency, we have shown that it is possible to maintain at least 80% accuracy with a drastic reduction in the number of measurement points of the scan (337 times fewer points) while keeping the wideband. If we work at a single frequency, the reduction in the number of measurement points is less substantial (154 times fewer points). Comparing the mean values and the standard deviations in the three cases, we observe a remarkable stability with the dense mapping with a mean accuracy of 99.7% and a standard deviation of 0.4. The mean values of the accuracy for the degraded scenarios are very close to each other with 92% and 92.1% for the wideband and the single frequency, respectively. Conversely, the classification results of the wideband scenario exhibit a better stability, with a standard deviation of 4.3 compared with 8.2 for the single-frequency scenario. In addition to the accuracy, we have to consider the system's complexity. As already noted above, it is not a critical point for our application, but the method can be generalized to a larger set of NDE methods for industrial applications, such as food safety. The system's complexity becomes a key issue and both degraded scenarios have their advantages and drawbacks.

- 1) *The Wideband Scenario*: With only 37 measurement points, the system would require 37 antennas (see the purple points in Fig. 14). Conversely, we need seven frequencies ranging from 110 to 170 GHz, so the wideband is a challenge for building a compact transmitter/receiver.
- 2) *The Single-Frequency Scenario*: The system would require 81 antennas (see the green points in Fig. 14), but the transmitter/receiver would be narrowband.

The time is not really the key issue here, but we cannot afford 10 min or even 2 min in an industrial setting. The antenna

would therefore have to be driven electronically, which makes the number of antennas the most critical parameter. Considering this last argument and the better homogeneity of the results with the wideband scenario, we recommend this configuration.

VI. CONCLUSION

In this article, we have proposed and demonstrated a new imaging modality for the classification of archeological pottery shards. We have sorted coiling from spiral shards with an automatic and nondestructive method based on low-THz measurements and artificial intelligence. The difference between the two categories lies in the shape of the air bubbles embedded in the ceramic materials during the different manufacturing techniques. The setup operates at low-THz, i.e., the D-band, because of the small size of the pores. As very few studies have focused on the automatic classification of pottery samples, we had to benchmark the low-THz classification results with those obtained from CT scan images. Using exactly the same postprocessing with low-THz and CT scan images, that is, a segmentation with the Otsu method followed by an SVM classification, we obtain accuracies of 100% and 59% to 70% for the CT scan and low-THz imaging systems, respectively. This preliminary study shows that the following conditions hold.

- 1) The classification of coiling and spiral pottery shards is 100% successful with CT scans.
- 2) Once segmented, the training dataset is linearly separable in the case of the CT scan, unlike the low-THz.
- 3) The training dataset can be relatively small and can be composed of experimental shards only, which is a significant advantage because we completely master the dataset.

These results are taken as the gold standard that must be achieved with the low-THz system. To this end, we completely changed the classification schemes. We processed the measurements with a 3-D FFT, which transformed the dataset, so it became separable, at least for 50% of its spectral density (the first three eigenvectors). In order to use an MLP for the classification, we searched for the optimal parameter values using a bioinspired algorithm, the GWO. We assessed the proposed methodology by using the full spatial and frequency diversity, which corresponds to 35 707 measurement points by cumulating seven frequencies. The accuracy ranged from 99% to 100% for the nine shard samples under test. We then optimized the measurement system based on two scenarios. First, we reduced the number of spatial acquisitions while keeping the wideband operation. Second, we worked at a fixed frequency and investigated the minimal number of acquisitions. In both scenarios, we set the critical fitness at 20% (80% accuracy), which is the fitness value that the final classification should not exceed.

- 1) In the first scenario, the accuracy was between 85% and 96% with 37 measurement points operating on wideband frequency.
- 2) In the second scenario, the accuracy was between 77% and 100% with 217 measurement points operating at a single frequency.
- 3) Both scenarios have their advantages and drawbacks, as discussed in this article, but both of them offer

substantial time reduction and system simplification compared to the original one.

In the near future, we may investigate the following.

- 1) The improvement in accuracy in the case of the reduction of the spatial and/or frequency diversity by optimizing the hyperparameters of the MLP in addition to the number of neurons in the hidden layer.
- 2) The simplification of the measurement system, namely, the classification of the shards with amplitude-only measurements instead of using the amplitude and the phase as we currently do.

REFERENCES

- [1] M. Gabriele *et al.*, "Long-distance mobility in the north-western Mediterranean during the neolithic transition using high resolution pottery sourcing," *J. Archaeol. Sci., Rep.*, vol. 28, Dec. 2019, Art. no. 102050. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352409X19302615>
- [2] L. Gomart *et al.*, "Spiralled patchwork in pottery manufacture and the introduction of farming to Southern Europe," *Antiquity*, vol. 91, no. 360, pp. 1501–1514, 2017.
- [3] S. Hughes, "CT scanning in archaeology," in *Computed Tomography-Special Applications*, L. Saba, Ed. Rijeka, Croatia: InTech, 2011, pp. 57–70.
- [4] U. Bergmann, P. L. Manning, and R. A. Wogelius, "Chemical mapping of paleontological and archeological artifacts with synchrotron X-rays," *Annu. Rev. Anal. Chem.*, vol. 5, no. 1, pp. 361–389, Jul. 2012, doi: 10.1146/ANNUREV-ANCHEM-062011-143019.
- [5] M. Marx and S. H. D'Auria, "CT examination of eleven Egyptian mummies," *RadioGraphics*, vol. 6, no. 2, pp. 321–330, Mar. 1986.
- [6] A. Gaboutchian *et al.*, "3D reconstruction and image processing of anthropological archaeological findings," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. 43, pp. 845–850, Aug. 2020.
- [7] N. Lynnerup, "Medical imaging of mummies and bog bodies—A mini-review," *Gerontology*, vol. 56, no. 5, pp. 441–448, 2010.
- [8] I. Cornelius, L. Swanepoel, A. D. Plessis, and R. Slabbert, "Looking inside votive creatures: Computed tomography (CT) scanning of ancient Egyptian mummified animals in Iziko museums of South of Africa: A preliminary report," *Akroterion*, vol. 57, no. 1, pp. 129–148, 2012.
- [9] T. Ignat, R. Bugoi, F. Constantin, V. Parnic, and C. Lazăr, "Identifying the chaîne opératoire of prehistoric clay figurines using experimental archeology and imaging methods," *Int. J. Mod. Phys., Conf.*, vol. 48, Jan. 2018, Art. no. 1860107, doi: 10.1142/S2010194518601072.
- [10] F. Zidane, J. Lanteri, L. Brochier, N. Joachimowicz, H. Roussel, and C. Migliaccio, "Damaged apple sorting with mmWave imaging and nonlinear support vector machine," *IEEE Trans. Antennas Propag.*, vol. 68, no. 12, pp. 8062–8071, Dec. 2020.
- [11] F. Zidane *et al.*, "Nondestructive control of fruit quality via millimeter waves and classification techniques: Investigations in the automated health monitoring of fruits," *IEEE Antennas Propag. Mag.*, vol. 62, no. 5, pp. 43–54, Oct. 2020.
- [12] D. Erricolo *et al.*, "Machine learning in electromagnetics: A review and some perspectives for future research," in *Proc. Int. Conf. Electromagn. Adv. Appl. (ICEAA)*, Sep. 2019, pp. 1377–1380.
- [13] L. Li, L. G. Wang, F. L. Teixeira, C. Liu, A. Nehorai, and T. J. Cui, "DeepNIS: Deep neural network for nonlinear electromagnetic inverse scattering," *IEEE Trans. Antennas Propag.*, vol. 67, no. 3, pp. 1819–1825, Mar. 2019.
- [14] Y. Zhou, Y. Zhong, Z. Wei, T. Yin, and X. Chen, "An improved deep learning scheme for solving 2-D and 3-D inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 69, no. 5, pp. 2853–2863, May 2021.
- [15] Z. Wei and X. Chen, "Physics-inspired convolutional neural network for solving full-wave inverse scattering problems," *IEEE Trans. Antennas Propag.*, vol. 67, no. 9, pp. 6138–6148, Sep. 2019.
- [16] H. Zong, L. Bao, B. Liu, and J. Qiu, "Application of convolutional neural network in target detection of millimeter wave imaging," in *Proc. IEEE Int. Symp. Antennas Propag. USNC/URSI Nat. Radio Sci. Meeting*, Jun. 2018, pp. 1217–1218.
- [17] C. Kettlewell *et al.*, "Experimental microwave target identification using machine learning," in *Proc. IEEE Int. Symp. Antennas Propag. USNC/URSI Radio Sci. Meeting*, Jul. 2019, pp. 227–228.
- [18] A. Kerhet, M. Raffetto, A. Boni, and A. Massa, "A SVM-based approach to microwave breast cancer detection," *Eng. Appl. Artif. Intell.*, vol. 19, no. 7, pp. 807–818, Oct. 2006. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0952197606000996>
- [19] M. Salucci, G. Oliveri, and A. Massa, "Real-time electrical impedance tomography of the human chest by means of a learning-by-examples method," *IEEE J. Electromagn., RF Microw. Med. Biol.*, vol. 3, no. 2, pp. 88–96, Jun. 2019.
- [20] J. Saluja, J. Casanova, and J. Lin, "A supervised machine learning algorithm for heart-rate detection using Doppler motion-sensing radar," *IEEE J. Electromagn., RF Microw. Med. Biol.*, vol. 4, no. 1, pp. 45–51, Mar. 2020.
- [21] G. Angiulli, M. Cacciola, and M. Versaci, "Microwave devices and antennas modelling by support vector regression machines," *IEEE Trans. Magn.*, vol. 43, no. 4, pp. 1589–1592, Apr. 2007.
- [22] D. R. Prado, J. A. López-Fernández, M. Arrebola, and G. Goussetis, "Efficient shaped-beam reflectarray design using machine learning techniques," in *Proc. 15th Eur. Radar Conf. (EuRAD)*, 2018, pp. 525–528.
- [23] E. Bermiani, A. Boni, S. Caorsi, and A. Massa, "An innovative real-time technique for buried object detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 4, pp. 927–931, Apr. 2003.
- [24] Y. Zhang, D. Orfeo, D. Huston, and T. Xia, "Software defined Doppler radar for landmine detection using GA-optimized machine learning," in *Proc. IEEE Int. Radar Conf. (RADAR)*, Apr. 2020, pp. 208–213.
- [25] V. V. Thakare and P. K. Singhal, "Bandwidth analysis by introducing slots in microstrip antenna design using ANN," *Prog. Electromagn. Res.*, vol. 9, pp. 107–122, 2009.
- [26] L. Xiao, W. Shao, F.-L. Jin, and B.-Z. Wang, "Multiparameter modeling with ANN for antenna design," *IEEE Trans. Antennas Propag.*, vol. 66, no. 7, pp. 3718–3723, Jul. 2018.
- [27] Q.-J. Zhang, K. Gupta, and V. Devabhaktuni, "Artificial neural networks for RF and microwave design—From theory to practice," *IEEE Trans. Microw. Theory Techn.*, vol. 51, no. 4, pp. 1339–1350, Apr. 2003.
- [28] H. Kabir, Y. Wang, M. Yu, and Q.-J. Zhang, "Neural network inverse modeling and applications to microwave filter design," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 4, pp. 867–879, Apr. 2008.
- [29] J. E. Rayas-Sanchez, "EM-based optimization of microwave circuits using artificial neural networks: The state-of-the-art," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 1, pp. 420–435, Jan. 2004.
- [30] A. Patnaik, B. Choudhury, P. Pradhan, R. K. Mishra, and C. Christodoulou, "An ANN application for fault finding in antenna arrays," *IEEE Trans. Antennas Propag.*, vol. 55, no. 3, pp. 775–777, Aug. 2007.
- [31] Y. Kim and H. Ling, "Human activity classification based on micro-Doppler signatures using an artificial neural network," in *Proc. IEEE Antennas Propag. Soc. Int. Symp.*, Jul. 2008, pp. 1–4.
- [32] Y. Kim and Y. Li, "Human activity classification with transmission and reflection coefficients of on-body antennas through deep convolutional neural networks," *IEEE Trans. Antennas Propag.*, vol. 65, no. 5, pp. 2764–2768, May 2017.
- [33] W. Shao and Y. Du, "Microwave imaging by deep learning network: Feasibility and training method," *IEEE Trans. Antennas Propag.*, vol. 68, no. 7, pp. 5626–5635, Jul. 2020.
- [34] G. Chen, P. Shah, J. Stang, and M. Moghaddam, "Learning-assisted multimodality dielectric imaging," *IEEE Trans. Antennas Propag.*, vol. 68, no. 3, pp. 2356–2369, Mar. 2020.
- [35] R. Aihara and Y. Fujimoto, "Free-space estimation for self-driving system using millimeter wave radar and convolutional neural network," in *Proc. IEEE Int. Conf. Mechatron. (ICM)*, Mar. 2019, pp. 467–470.
- [36] Z. Meng, M. Zhang, and H. Wang, "CNN with pose segmentation for suspicious object detection in MMW security images," *Sensors*, vol. 20, no. 17, p. 4974, 2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/17/4974>
- [37] B. Choi and D. Oh, "Classification of drone type using deep convolutional neural networks based on micro-Doppler simulation," in *Proc. Int. Symp. Antennas Propag. (ISAP)*, 2018, pp. 1–2.
- [38] J. Bewes, A. Low, A. Morphet, F. D. Pate, and M. Henneberg, "Artificial intelligence for sex determination of skeletal remains: Application of a deep learning artificial neural network to human skulls," *J. Forensic Legal Med.*, vol. 62, pp. 40–43, Feb. 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1752928X18304219>
- [39] *Passion Poterie*. Accessed: Mar. 4, 2011. [Online]. Available: <http://passion-poterie.com/article/comment-couper-un-colombin-56.html>

- [40] G. R. Olhoeft, "Tables of room temperature electrical properties for selected rocks and minerals with dielectric permittivity statistics," Geological Survey, U.S. Dept. Interior, Washington, DC, USA, Tech. Rep. 79-993, 1979.
- [41] F. Nsengiyumva, C. Migliaccio, L. Brochier, J.-Y. Dauvignac, I. Aliferis, and C. Pichot, "New W-band scattering measurement system: Proof of concept and results for 2-D objects," *IEEE Trans. Antennas Propag.*, vol. 66, no. 12, pp. 7224–7236, Dec. 2018.
- [42] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
- [43] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, 1995.
- [44] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics Intell. Lab. Syst.*, vol. 2, nos. 1–3, pp. 37–52, 1987.
- [45] F. Zidane *et al.*, "Sorting of archaeological samples in D-band," in *Proc. IEEE Antennas Propag. Soc. Int. Symp. USNC/URSI Nat. Radio Sci. Meet. (APSURSI)*, Jun. 2020, pp. 1801–1802.
- [46] H. S. N. Murthy and D. M. Meenakshi, "ANN, SVM and KNN classifiers for prognosis of cardiac Ischemia—A comparison," *Bonfring Int. J. Res. Commun. Eng.*, vol. 5, no. 2, pp. 7–11, Jun. 2015.
- [47] M. Gardner and S. Dorling, "Artificial neural networks (the multilayer perceptron)—A review of applications in the atmospheric sciences," *Atmos. Environ.*, vol. 32, no. 14, pp. 2627–2636, 1998. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1352231097004470>
- [48] B. C. Csáji *et al.*, "Approximation with artificial neural networks," Dept. Sci., Eötvös Loránd Univ., Hungary, Europe, Tech. Rep., 2001, vol. 24, no. 48, p. 7.
- [49] R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Neural Networks for Perception*, H. Wechsler, Ed. New York, NY, USA: Academic, 1992, pp. 65–93. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/B9780127412528500108>
- [50] S. Mirjalili, "How effective is the Grey Wolf optimizer in training multilayer perceptrons," *Appl. Intell.*, vol. 43, no. 1, pp. 150–161, 2015, doi: [10.1007/S10489-014-0645-7](https://doi.org/10.1007/S10489-014-0645-7).
- [51] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey Wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0965997813001853>
- [52] N. Mittal, U. Singh, and B. S. Sohi, "Modified Grey Wolf optimizer for global engineering optimization," *Appl. Comput. Intell. Soft Comput.*, vol. 2016, May 2016, Art. no. 7950348.
- [53] B. Martin, J. Marot, and S. Bourennane, "Mixed grey wolf optimizer for the joint denoising and unmixing of multispectral images," *Appl. Soft Comput.*, vol. 74, pp. 385–410, Jan. 2019. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1568494618305763>
- [54] I. Yilmaz, "Comparison of landslide susceptibility mapping methodologies for koyulhisar, Turkey: Conditional probability, logistic regression, artificial neural networks, and support vector machine," *Environ. Earth Sci.*, vol. 61, no. 4, pp. 821–836, Aug. 2010.
- [55] H. Jabbar and R. Z. Khan, "Methods to avoid over-fitting and under-fitting in supervised machine learning (comparative study)," *Comput. Sci., Commun. Instrum. Devices*, vol. 70, pp. 163–172, Dec. 2015.
- [56] P. Berka, J. Rauch, and D. A. Zighed, *Data Mining and Medical Knowledge Management: Cases and Applications: Cases and Applications*. Hershey, PA, USA: IGI Global, 2009.
- [57] C. Eyraud, J.-M. Geffrin, and A. Litman, "Drift correction for 3D scattering measurements," in *Proc. IEEE Int. Symp. Antennas Propag.*, Jul. 2006, pp. 2003–2006.



Flora Zidane is currently pursuing the Ph.D. degree with the Laboratoire d'Electronique, Antennes et Télécommunications, French National Center for Scientific Research (CNRS), Université Côte d'Azur, Sophia Antipolis, France.

Her research interests include millimeter-wave (mm-Wave) imaging and machine learning algorithms.



Vanna Lisa Coli is currently with the Culture and Environment, Prehistory, Antiquity, Middle Age (CEPAM) Laboratory, French National Center for Scientific Research (CNRS), Université Côte d'Azur, Nice, France. Her research interests are on machine learning, archeology, digital signal processing, imaging, and modeling. Her current works deal with the application of machine learning and image processing methods to archeological problems, such as the characterization of ancient materials and technical systems.



Jérôme Lanteri (Member, IEEE) is currently an Associate Professor with the French National Center for Scientific Research (CNRS), Laboratoire d'Electronique, Antennes et Télécommunications, Université Côte d'Azur, Sophia Antipolis, France. His research interests include integrated antennas, reflectarrays, and transmitarrays at millimeter-wave frequencies for nondestructive testing and medical or security applications.



Julien Marot received the Ph.D. degree from Aix Marseille Université in 2007.

He held a post-doctoral position at the Fraunhofer Institute, Erlangen, Germany. He is currently an Associate Professor with the Institut Fresnel, Marseille, France. His research interests include applied image processing, multidimensional signal processing, and array processing.



Laurent Brochier is currently with the French National Center for Scientific Research (CNRS), Laboratoire d'Electronique, Antennes et Télécommunications, Université Côte d'Azur, Sophia Antipolis, France.



Didier Binder is currently an Emeritus Research Director at the French National Center for Scientific Research (CNRS) and a member of the Joint Laboratory Culture and Environment, Prehistory, Antiquity, Middle Age (CEPAM), Université Côte d'Azur, Nice, France. His research interests are on the spread of early farmers and their interactions with hunter gatherers through the Mediterranean and Western Europe. His current work mainly concerns the characterization and evolution of material culture and technical systems during the Holocene period.



Claire Migliaccio (Member, IEEE) is currently a Full Professor with the French National Center for Scientific Research (CNRS), Laboratoire d'Electronique, Antennes et Télécommunications, Université Côte d'Azur, Sophia Antipolis, France. Her research interests include wideband antennas, millimeter-wave (mm-Wave) systems from antenna design to radar, and imaging techniques.