Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (IJCAI-23)
Special Track on AI, the Arts and Creativity

IberianVoxel: Automatic Completion of Iberian Ceramics for Cultural Heritage Studies

Pablo Navarro[1*], **Celia Cintas**[2+], **Manuel Lucena**[3], **José Manuel Fuertes**[3], **Antonio Rueda**[3], **Rafael Segura**[3], **Carlos Ogayar-Anguita**[4], **Rolando González-José**[1] and **Claudio Delrieux**[5]

[1] Instituto Patagónico de Ciencias Sociales y Humanas (CONICET). Puerto Madryn, Argentina.
[2] IBM Research Africa, Nairobi, Kenya.
[3] Research University Institute for Iberian Archaeology, University of Jaén, Spain.
[4] Center for Advanced Studies in Information and Communication Technologies, University of Jaén, Spain.
[5] Universidad Nacional del Sur, and CONICET. Bahía Blanca, Argentina.
[*]pnavarro@cenpat-conicet.gob.ar, [+]celia.cintas@ibm.com

Abstract

Accurate completion of archaeological artifacts is a critical aspect in several archaeological studies, including documentation of variations in style, inference of chronological and ethnic groups, and trading routes trends, among many others. However, most available pottery is fragmented, leading to missing textural and morphological cues. Currently, the reassembly and completion of fragmented ceramics is a daunting and time-consuming task, done almost exclusively by hand, which requires the physical manipulation of the fragments. To overcome the challenges of manual reconstruction, reduce the materials' exposure and deterioration, and improve the quality of reconstructed samples, we present IberianVoxel, a novel 3D Autoencoder Generative Adversarial Network (3D AE-GAN) framework tested on an extensive database with complete and fragmented references. We generated a collection of 1001 3D voxelized samples and their fragmented references from Iberian wheel-made pottery profiles. The fragments generated are stratified into different size groups and across multiple pottery classes. Lastly, we provide quantitative and qualitative assessments to measure the quality of the reconstructed voxelized samples by our proposed method and archaeologists' evaluation.

1 Introduction

During archaeological excavations, it is common to find fractured or damaged artifacts. Among the many factors, ceramic potteries are one of the most frequently discovered archaeological artifacts. Since they are usually short-lived, researchers find these artifacts useful to analyze chronological and geographical features, given that shape and decoration are subject to significant changes over time and space [Eslami *et al.*, 2020]. This analysis gives a basis for dating the archaeological strata, and provides evidence from a large set of valuable data, such as local production, trade relations, and consumer behavior of the local population [Kampel and Sablatnig, 2003; Kashihara, 2017].

Several prior studies analyze various aspects of ceramics using 2D complete pottery profiles. Automatic profile classification [Lucena *et al.*, 2017; Cintas *et al.*, 2020; Llamas *et al.*, 2016; Di Angelo *et al.*, 2017] and feature extraction [Shennan and Wilcock, 1975; Rice, 1987; Nautiyal *et al.*, 2006; Mom, 2007; Saragusti *et al.*, 2005; Karasik and Smilansky, 2011; Smith *et al.*, 2014] have been widely studied, ranging from traditional image processing techniques to deep learning approaches. Unfortunately, ceramics are fragile, and therefore most of the actual ceramics recovered from archaeological sites are fractured, so the vast majority of the available material appears in fragments. The reassembly of the fragments is a daunting, delicate, and time-consuming task, done almost exclusively by hand, which requires the physical manipulation of the fragments, which ideally should be as short as possible for conservation purposes. Due to these factors and the rise of easy-to-use and affordable 3D scanners[1], there is an increasing interest in automatic pottery reassembly and reconstruction [Rasheed and Nordin, 2014; Fragkos, Stergios *et al.*, 2018; Kalasarinis and Koutsoudis, 2019] and fragment analysis [Wilczek *et al.*, 2021]. Nonetheless, existing work resolves the fragments problem using comparisons between known pieces and mainly works on 2D representations [Navarro *et al.*, 2022].

In this work, we propose a 3D Autoencoder Generative Adversarial Network (3D AE-GAN) framework to augment and reconstruct ceramics, in which the "best-matching pottery" is artificially generated based on a set of known 3D fragments in the model, thus providing reconstructed virtual pottery with the same geometric features as the real ones. The resulting augmented dataset will be very helpful for exploring and designing automatic procedures to aid experts with the pottery completion task.

[1]3D scanners. *(Last Accessed in January 2023)*

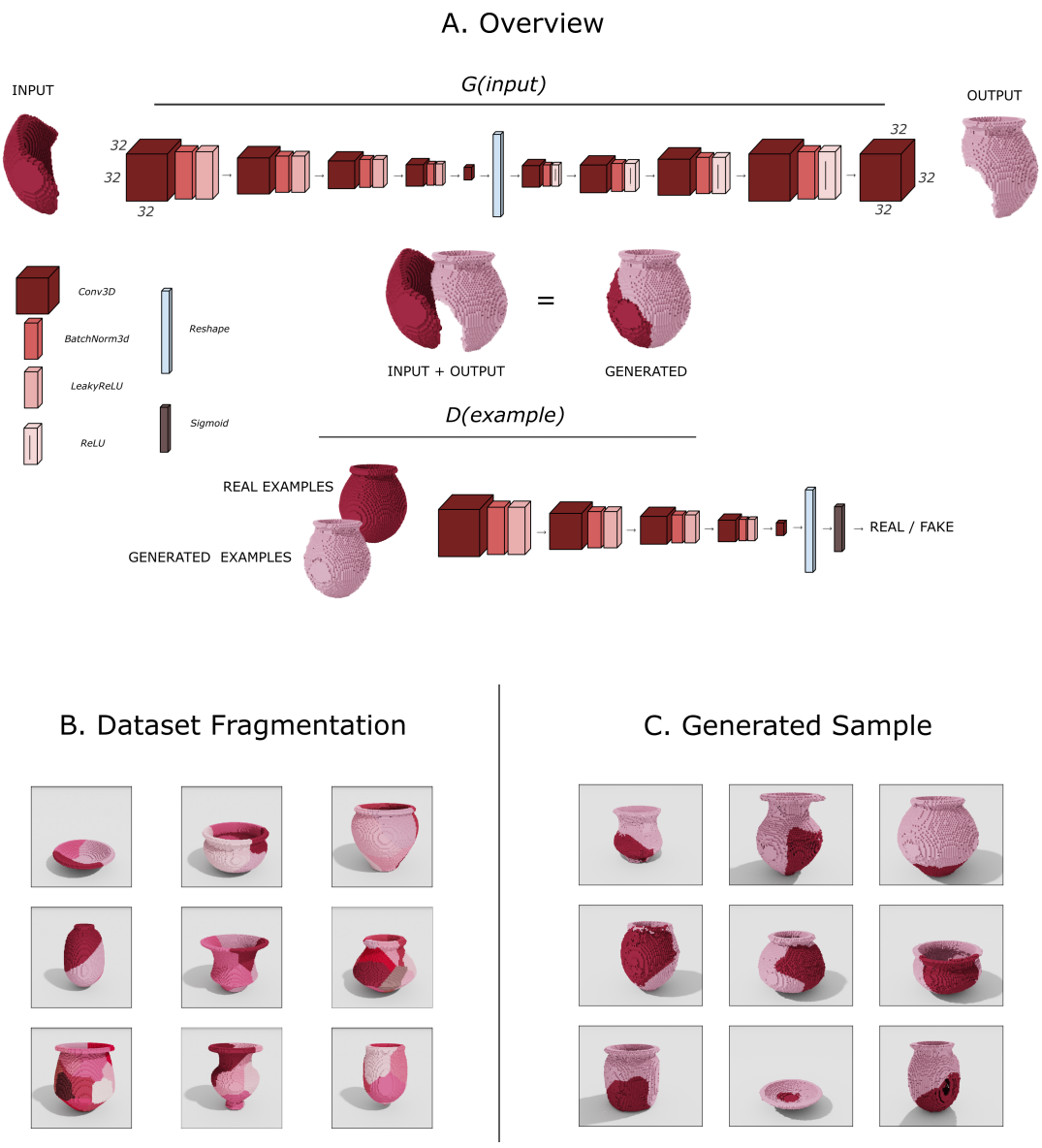

Figure 1: Overview of the proposed approach. **A** IberianVoxel architecture. The $G(z)$ generator is based on a 3D encoder-decoder architecture. Upon receiving a set of voxels corresponding to the pottery fragment, the encoder transforms it into a vector, and then the decoder generates the missing or unknown 3D fragment. The discriminator $D(x)$ receives the complete set of voxels to determine if it is real or generated. **B** Criteria for material fractures. **C** Examples of IberianVoxel generated samples from fragments across different fragment sizes.

The main contributions of this paper are the following:

- A 3D Autoencoder Generative Adversarial Network (3D AE-GAN), that aims to complete pottery automatically. Our method generates new ceramics from 156546 fragments (see Fig. 1A).
- A procedure for generating artificial 3D fragment samples from voxelized pottery models by adapting the Discrete Voronoi Chain algorithm (see Fig. 1B).
- We release a new database of 3D ceramics fragments generated from 1001 voxelized pottery models fragmented 16 times (in total 16016 models) with matching correspondence of fragments and full pottery.
- We evaluate the proposed approach with standard metrics (MSE, DSC, and Jaccard) under different fragment sizes (see Fig. 1C for results). Additionally, we validated our framework with a binary shape classifier. Lastly, we conduct a validation study with a group of archaeologist evaluators to qualitatively assess the quality of the completed samples with respect to domain experts.

2 Related Work

Artifact completion, reassembly, and restoration are key research topics in Computational Archaeology. There are several recent examples of GANs applied to cultural heritage do-

mains in the literature [Nguyen *et al.*, 2019; Hermoza and Sipiran, 2018]. For instance, techniques such as automated image style transfer [Chen *et al.*, 2020] were used to develop a model aimed at generating Cantonese porcelain-styled images departing from user-defined masks. Similar techniques were also applied to material degradation [Papadopoulos *et al.*, 2021; Liu *et al.*, 2017a; Isola *et al.*, 2017; Zhu *et al.*, 2017; Wang *et al.*, 2017; Zachariou *et al.*, 2020].

In a similar approach to our proposal, in [Hermoza and Sipiran, 2018] a 3D reconstruction GAN for archaeological restoration is proposed, based on an encoder-decoder 3D CNN on top of a GAN based on cGANs [Mirza and Osindero, 2014]. In [Kniaz *et al.*, 2019], a translation network is presented. Both these studies work on the problem of prediction of missing geometry on damaged voxelized objects. Another example of cultural heritage preservation can be found in [Jboor *et al.*, 2019], where an image completion approach is adapted [Liu *et al.*, 2017b; Yeh *et al.*, 2017; de Lima-Hernandez and Vergauwen, 2021] for the curation of damaged artwork by predicting the continuing decoration traces of broken heritage fragments in 2D. Finally, [Navarro *et al.*, 2022] presents a method based on a GAN for creating 2D complete profile potteries from binary images of fragments. Even though these works show promising results in generating Iberian potteries, the previous proposed processes for managing fragments and fragmentation are unrealistic.

3 Proposed Approach

A typical GAN [Goodfellow *et al.*, 2014] framework contains a generative (G) and a discriminative (D) neural network such that $G(z)$ aims (in our context) to generate realistic artifacts, while $D(x)$ learns to discriminate if a sample is from the real data distribution ($p_d(x)$) or not. $D(x)$ should be high when x comes from training data and low when x comes from the generator. In a classic GAN, z is a latent space vector sampled from a statistical distribution (usually normal distribution).

In our case, as opposed from [Wu *et al.*, 2016], z is a voxelized input fragment (matrix of $32 \times 32 \times 32$), and $G(z)$ represents the generator function which maps the fragment z to the data space of a complete Iberian voxelized pottery. This type of network is named Autoencoding GAN (AE-GAN), in which G added a network of encoders that are trained to learn an $E : Z \rightarrow W$ function, mapping each fragment sample to a point (w) in latent space [Lazarou, 2020] and the decoder learn to map each point (w) to a complete pottery.

Multiple iterations will inform G on how to adjust the generation process to generate a misclassification in D. In our particular case, the data element x, corresponds to a three-dimensional binary array containing the voxelized pottery geometry. $D(G(z))$ is the probability that the output of the generator G is a real artifact from the Iberian pottery 3D dataset. D tries to maximize $(log(D(x)))$, which is the probability of having a correct classification of actual voxelized real ceramics, while G tries to minimize $(log(1 - D(G(z))))$, which is the probability of D recognizing any of the generated outputs by G (see Eq. 1).

IberianVoxel is based on the AE-3DGAN, where the generator is an Autoencoding network $Encode(z) \rightarrow w \in R^m$, $Decode(w) \rightarrow x'$, where $x \in [0,1]^{32x32x32}$ is the input fragment, a three-dimensional binary array containing the fragment shape information by means of voxels, and x' is a missing generated part. To train the discriminator network, we use $D(y)$ where $y = z + x'$ for the generated examples. The complete codebase for the definition, implementation, training, and evaluation of IberianVoxel can be found here[2].

	Fragments (%)		
	15-20	20-30	30-100
Class 1	246 (2.76%)	194 (2.54%)	86 (2.31%)
Class 2	368 (4.12%)	313 (4.11%)	144 (3.86%)
Class 3	571 (6.4%)	511 (6.7%)	235 (6.31%)
Class 4	91 (1.02%)	68 (0.89%)	42 (1.13%)
Class 5	516 (5.78%)	430 (5.64%)	186 (4.99%)
Class 6	441 (4.94%)	404 (5.3%)	160 (4.29%)
Class 7	397 (4.45%)	389 (5.1%)	170 (4.56%)
Class 8	2419 (27.11%)	2063 (27.06%)	900 (24.15%)
Class 9	161 (1.8%)	156 (2.05%)	73 (1.96%)
Class 10	2865 (32.11%)	2352 (30.85%)	1317 (35.34%)
Class 11	848 (9.5%)	744 (9.76%)	414 (11.11%)
Total	8923	7624	3727

Table 1: Number of fragments generated by our proposed DVC adaptation per each ceramic type and size.

4 Experimental Setup

In this section, we describe the process taken for the generation of the new 3D fracture dataset and pre-processing steps for the voxelization (see Fig. 2). Further, we define hyperparameters, architecture, and training steps implemented for IberianVoxel. We detailed the different assessments done to evaluate the performance of our proposed model using multiple initial fragment sizes. We evaluated the samples generated with standard completion and segmentation metrics. Additionally, we validated the output of IberianVoxel with a pottery shape classifier. Lastly, we show the protocol done to conduct a validation study with domain experts.

4.1 Iberian Pottery Data

The raw data belong to binary profile images, corresponding to Iberian wheel-made pottery from various archaeological sites of the upper valley of the Guadalquivir River (Spain). The ceramics are classified into eleven different classes based on their shape. These classes consider the forms of the lip, neck, body, base, and handles and the relative ratios between their sizes. Nine of these classes correspond to closed pottery shapes, and two others belong to open ones [Lucena *et al.*, 2017]. The available images consist of a profile view of the pottery, where image resolutions (in pixels), corresponding to size scale, may vary according to the acquisition settings (see Fig. 2).

[2]https://github.com/celiacintas/vasijas/tree/iberianVox

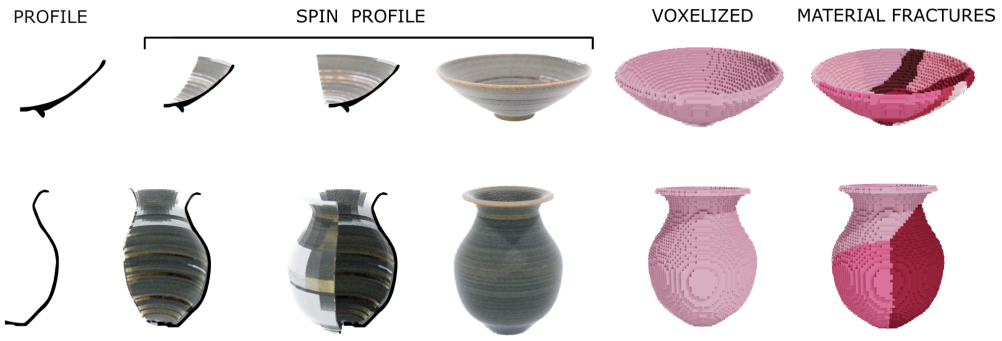

Figure 2: Data pipeline for voxelized samples and fragment references generation. We show an example for each type of open and closed shape. To the left, we show the original 2D profile [Lucena *et al.*, 2017]. Second, we show the spin calculated over the y-axis. Third, we show the voxelized mesh and the fragments generated by the adapted Discrete Voronoi Chain algorithm.

Classes	Fragments (%) 15-20			20-30			30-100		
	MSE(↓)	DSC(↑)	Jacc.(↓)	MSE(↓)	DSC(↑)	Jacc.(↓)	MSE(↓)	DSC(↑)	Jacc.(↓)
Class 1	0.04 ± 0.01	0.73 ± 0.08	0.41 ± 0.09	0.04 ± 0.01	0.74 ± 0.09	0.40 ± 0.11	0.03 ± 0.01	0.78 ± 0.06	0.34 ± 0.08
Class 2	0.06 ± 0.02	0.72 ± 0.09	0.42 ± 0.11	0.06 ± 0.02	0.74 ± 0.09	0.40 ± 0.12	0.04 ± 0.01	0.80 ± 0.05	0.33 ± 0.07
Class 3	0.05 ± 0.02	0.60 ± 0.12	0.55 ± 0.12	0.05 ± 0.02	0.62 ± 0.13	0.53 ± 0.13	0.04 ± 0.01	0.65 ± 0.11	0.49 ± 0.11
Class 4	0.07 ± 0.03	0.47 ± 0.13	0.68 ± 0.11	0.06 ± 0.01	0.55 ± 0.13	0.61 ± 0.12	0.05 ± 0.01	0.58 ± 0.07	0.59 ± 0.07
Class 5	0.09 ± 0.02	0.62 ± 0.11	0.54 ± 0.12	0.08 ± 0.02	0.67 ± 0.10	0.48 ± 0.12	0.07 ± 0.02	0.70 ± 0.11	0.45 ± 0.13
Class 6	0.09 ± 0.02	0.62 ± 0.11	0.53 ± 0.12	0.05 ± 0.02	0.71 ± 0.10	0.44 ± 0.11	0.04 ± 0.02	0.75 ± 0.08	0.39 ± 0.10
Class 7	0.13 ± 0.11	0.55 ± 0.24	0.58 ± 0.21	0.11 ± 0.01	0.59 ± 0.24	0.54 ± 0.21	0.10 ± 0.01	0.61 ± 0.24	0.52 ± 0.22
Class 8	0.05 ± 0.02	0.75 ± 0.09	0.39 ± 0.11	0.04 ± 0.01	0.77 ± 0.09	0.36 ± 0.11	0.04 ± 0.01	0.80 ± 0.06	0.32 ± 0.09
Class 9	0.05 ± 0.02	0.75 ± 0.11	0.38 ± 0.13	0.04 ± 0.01	0.78 ± 0.10	0.35 ± 0.11	0.03 ± 0.00	0.80 ± 0.07	0.32 ± 0.09
Class 10	$\mathbf{0.01 \pm 0.01}$	$\mathbf{0.82 \pm 0.06}$	$\mathbf{0.28 \pm 0.09}$	$\mathbf{0.01 \pm 0.00}$	$\mathbf{0.84 \pm 0.06}$	$\mathbf{0.26 \pm 0.08}$	$\mathbf{0.01 \pm 0.00}$	$\mathbf{0.84 \pm 0.06}$	$\mathbf{0.27 \pm 0.08}$
Class 11	0.02 ± 0.00	0.75 ± 0.08	0.39 ± 0.10	0.02 ± 0.00	0.75 ± 0.09	0.39 ± 0.11	$\mathbf{0.01 \pm 0.00}$	0.76 ± 0.08	0.38 ± 0.10
Average	0.06 ± 0.03	0.70 ± 0.07	0.44 ± 0.08	0.06 ± 0.03	0.70 ± 0.07	0.44 ± 0.08	0.06 ± 0.03	0.70 ± 0.07	0.45 ± 0.08

Table 2: Performance metrics across different fragment sizes and ceramic types. The percentage is relative to each pottery's size and shape.

$$\min_G \max_D V(D,G) = \mathbb{E}_{x \sim p_d(x)}\left[log(D(x))\right] + \mathbb{E}_{z \sim p_z(z)}\left[log(1 - D(z + G(z)))\right] \quad (1)$$

Completed Pottery
Completed Voxelized Space
Generative Autoencoder
Input Fragment

2D to 3D pottery Given that the existing dataset is based on wheel-made pottery techniques, we can assume small asymmetric perturbations in the ceramic. Thus, we can generate each 3D model as a solid of revolution (see Fig. 2 for an overview of the steps implemented in this work). First, we extract the shape information from the profile by means of 200 semilandmarks equally spaced in the contours. For the spin, we use 100 steps per revolution. This task converts the semilandmarks capturing the 2D pottery shape information into a mesh. Lastly, we converted the resulting mesh into a set of voxels following [Ogayar-Anguita *et al.*, 2020] with sizes $32 \times 32 \times 32$. The resulting dataset is composed of 1001 voxelized potteries, randomly divided into a training subset containing 704 ceramics (70%) and a test set of 297 artifacts (30% of the total dataset).

Fracture ceramic simulations The voxelized pottery models were fragmented, simulating real fractured ceramics.We implemented a geometric fragmentation method based on the Discrete Voronoi Chain (DVC) algorithm [Velić *et al.*, 2009]. The DVC algorithm is composed of several steps. First, it generates a random list of Voronoi region centers in the model, each corresponding to a fragment. Second, it assigns the voxel to a section by following a region-growing approach taking each center as a seed until all the voxels have been traversed. However, this procedure can assign voxels in the border between two regions to an incorrect one. Therefore, an additional distance check to the centers of each region is required to guarantee the correct assignment. The resulting number of fragments depends on the type of vessel (closed or open, see Fig. 1-B).

Our approach, like most geometric methods based on Voronoi [Raghavachary, 2002; Domaradzki and Martyn, 2016], generates convex fragments. Thus, in order to generate more complex fragments, including concavities, we need to disable at runtime the distance check from the DVC algorithm (10% of the runs).

Lastly, in order to evaluate the performance under different relative sizes of fragments, we stratified our fragment generation into three main groups: 15-20%, 20-30%, and 30-100% of the initial ceramic model.

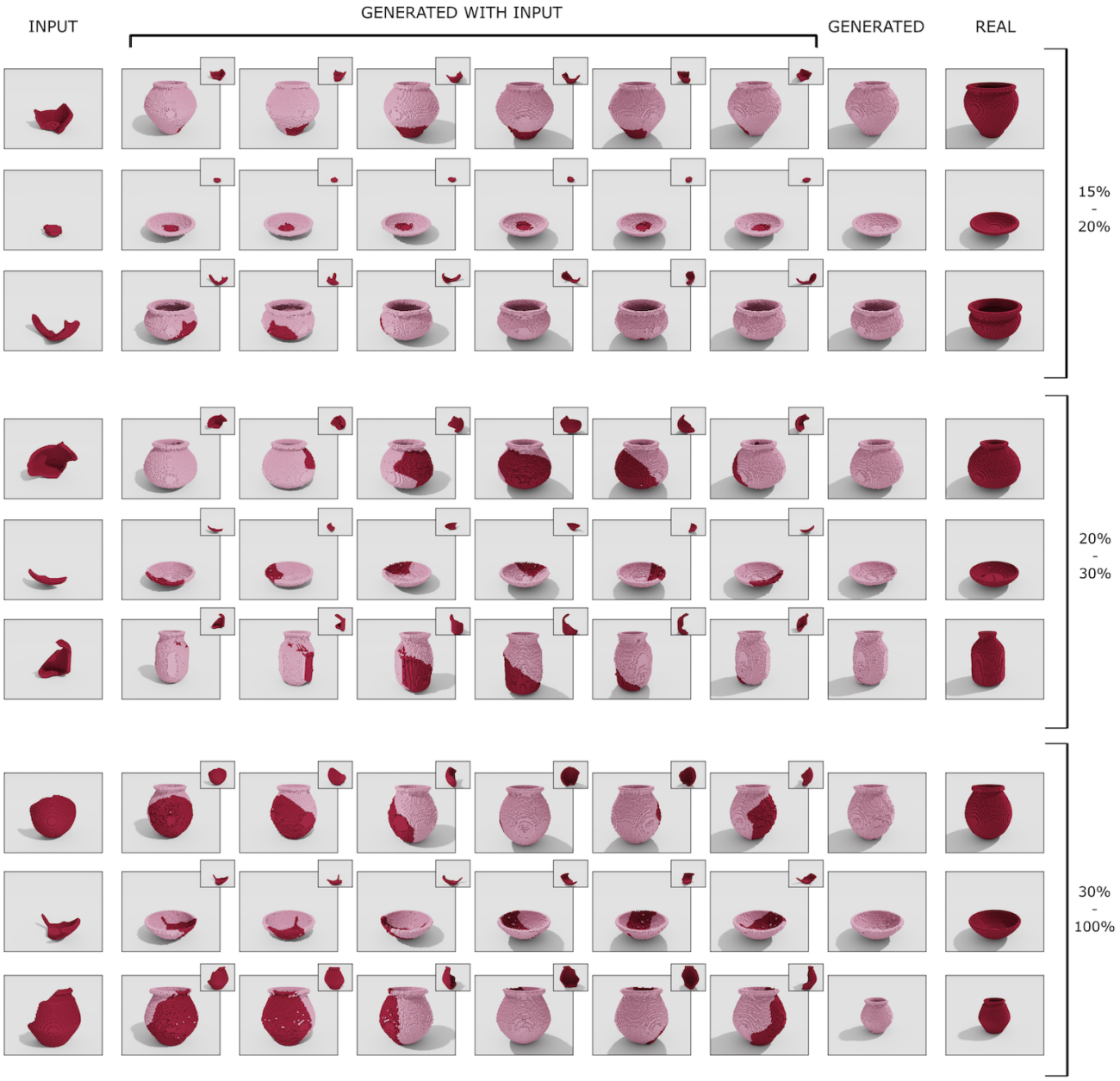

Figure 3: Qualitative examples for visual comparison of completed samples by our proposed method and the ground truth (real voxelized models). Each batch of examples shows the completion model started from different sizes of initial fragments (dark pink), and the completed ceramics (light pink) in different views.

4.2 Implementation Details

The first module of the Generator consists of an Encoder of 5 convolutional layers of kernel sizes $4 \times 4 \times 4$, stride 2, and padding $1 \times 1 \times 1$ (with $0 \times 0 \times 0$ for last one), with 3D batch normalization and LeakyReLU layers added in between the convolutional layers (except the last one). The intermediate result is a $128 - dimensional$ vector. The last part of the Generator is a Decoder of 5 layers, all of them with a 3D transposed convolution with kernel sizes $4 \times 4 \times 4$, stride 2, and padding $1 \times 1 \times 1$. Unlike the encoder, the decoder has ReLU layers instead of LeakyRelu. The output is a binary array of $32 \times 32 \times 32$. The architecture of the D is basically the same as that of the G encoder except that it uses a sigmoid layer, previously reshaped the result of the last layer, for the binary classification (real/generated).

3D AE-GAN training IberianVoxel was trained for 100 epochs using Binary Cross Entropy as a loss function, at a learning rate 2×10^{-3} for the generative 3D encoder-decoder network (G) and 2×10^{-4} for the discriminator (D). To optimize the training process of all models, we scaled the artifacts to a uniform resolution space of $32 \times 32 \times 32$ voxels. We used ADAM optimization [Kingma and Ba, 2015] for both G and D with $\beta_1 = 0.9$ and $\beta_2 = 0.999$. Particularly, for the training of D, we used Label Smoothing [Salimans *et al.*, 2016], the real set was represented with a random number between 0.7 and 1.1 and the generated set with 0.0 and 0.30.

4.3 Evaluation Metrics

The evaluation of the generated artifacts' quality was done using standard completion and segmentation metrics such as the Dice similarity coefficient (DSC) [Sorensen, 1948], Jaccard distance, and Mean squared error (MSE). The DSC coefficient is used to calculate the similarity between the two examples [Feng *et al.*, 2016; Navarro *et al.*, 2022]. Given A the generated set of voxels and B the real ceramic, DSC is formulated:

$$DSC = \frac{2|A \cup B|}{|A| + |B|}, \tag{2}$$

where $|A|$ and $|B|$ is the size in voxels of the pottery. The values of the formula are between 0 and 1, where 1 represents the maximum similarity, and 0 is the minimum.

Jaccard distance, like DSC, is a statistic used to estimate the dissimilarity of a dataset, but unlike DSC, Jaccard distance satisfies the triangle inequality. Jaccard distance is defined as:

$$JD(A, B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}, \tag{3}$$

where A is a three-dimensional binary array containing the generated pottery voxelized geometry, and B is the real ceramic. The values of the formula are between 0 and 1 too, but unlike DSC, 1 represents the minimum similarity and 0 the maximum. Lastly, we evaluated our generated samples using MSE, which is a common quality estimator in 2D and 3D classification and reconstruction problems.

$$MSE = \frac{1}{n} \Sigma_{i=1}^{n} \Big(d_i - f_i \Big)^2, \tag{4}$$

where d is a real voxelized pottery and f a generated one. In our case, d_i and f_i represent one voxel in the three-dimensional array of the pottery d and f, respectively.

4.4 Domain Expert Study

To evaluate the quality of the completed pottery, we conducted a study with a group of archaeologist evaluators, who assessed the samples individually and in isolation, following similar protocols from [Navarro *et al.*, 2022]. The evaluators' feedback is then collected and aggregated to establish an overall rating. In our preliminary study, eleven voluntary archaeologists specializing in Iberian ceramics completed a questionnaire evaluating ten examples with six views. An example of the questionnaire can be found in the Supplementary material. Each graphical example is followed by two questions. These rating questions assess the quality of the resulting pottery and the initial fragment. First, a question regarding the generated sample (Very Bad, Poor, Good, and Very Good). Second, the level of similarity with an Iberian style (rate between 0 and 5, where 0 means unrelated to the Iberian Style, and 5 means entirely within the Iberian Style).

5 Results

In Table 2 we present the obtained results across the different size groups of fragments and several pottery types. We can observe that the averaged $MSE = 0.06$ and $DSC = 0.70$ across all classes for the completion of the ceramic is consistent across fragment sizes. This means that IberianVoxel is able to complete ceramics from smaller to larger initial fragments.

Interestingly, in Table 2 we can observe that specific classes of pottery are easier to reconstruct across all sizes, such as Class 10 and Class 11. We can hypothesize that the improvement, over the averaged performance of the method over these classes, is due to the fact that we have more examples of these classes in our dataset. Furthermore, these classes correspond to a homogeneous type of open shapes (e.g., plates). In Table 1 we present a detailed stratification across different pottery classes and fragment sizes generated by our fracture simulator.

In Figure 3 we show a random set of qualitative examples. In the first column, we can observe the initial fragment (dark pink), seven different angles of the generated pottery (light pink), the fragment (dark pink), and the original ceramic. Additionally, we validated our completed ceramics with a binary shape classifier. This model classifies the voxelized pottery into open or closed shape vessels. In Table 3, we can observe that the classifier yields the same performance in test datasets when evaluating real voxelized potteries and completed potteries from three different groups sizes of fragments by our proposed method.

The domain experts' case study yielded that the reconstructed ceramics' quality perceived by the archaeologists had a mean score of 2.09 with a standard deviation of 0.61 (with range values between $[0, 3]$). Additionally, we were interested in evaluating if the Iberian style holds in the completed samples. The archaeologists considered the ceramics to have an Iberian style in mean score 3.93 (± 1.16), with 5

Dataset	Precision (↑)	Recall (↑)	F1-Score (↑)	Accuracy (↑)
Original	0.980	0.984	0.982	0.983
IberianVoxel (30%-100%)	0.984	0.987	0.986	0.986
IberianVoxel (20%-30%)	**0.988**	**0.990**	**0.989**	**0.989**
IberianVoxel (15%-20%)	0.984	0.987	0.986	0.986

Table 3: Performance metrics for open-close shapes classifier from test 3D data. All metrics were estimated with a test set of 121 open and 176 closed ceramics. We show results with original samples and ceramics reconstructed from initial fragments of different sizes (in %).

fully Iberian Style. From this preliminary study, we can conclude that archaeologists judge that IberianVoxel generated a correct Iberian style from an initial fragment, and also consider that the reconstructed pottery is between *Good* and *Very Good*. At the end of the questionnaire, we included a comment section to enable unstructured feedback. The comments across evaluators agree on the need for better visualization tools, such as including a scale factor and improving the edge inspection of the models, as these are key factors while evaluating the Iberian Style.

5.1 Limitations

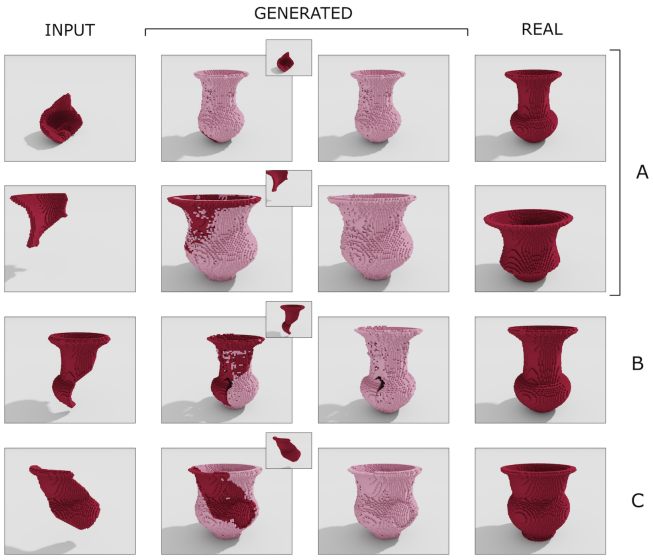

Figure 4: Qualitative examples of current limitations. **A** Samples generated with clear asymmetric patterns. **B** Samples generated with irregularities and missing areas. Initial fragments (dark pink), and the completed ceramics (light pink).

Asymmetry in generated potteries Some completed samples do not have an adequate symmetry. The pottery used for this work corresponds to lathe potteries, and those are symmetrical by design. In Fig. 4-A, we show two examples in which the generated models are not symmetric. This type of problem is caused by the lack of examples in some classes. One potential solution could be a post-processing task to correct these artifacts or implement a penalization factor in the loss function during the training process.

Irregularity in generated potteries In Fig. 4-B and Fig. 4-C we show two examples with irregular results in the completion. For example, incomplete reconstructions (see Fig. 4-B,) or the mismatching alignment of the fragment with the proposed completion model (see Fig. 4-C). These problems could be reduced using post-processing methods, such as classic morphological operations.

6 Conclusion

Ceramics are one of the most common archaeological artifacts, usually used to investigate variations in style, materials employed, social activities, and manufacturing techniques. Nonetheless, most available pottery is fragmented. The completion and assembly of fragments is a time-consuming task, which requires the physical manipulation of the ceramic fragments, which exposes them to material decay. We present a 3D generative approach to process fragments automatically, and reconstruction analysis that can assist archaeologists in a more realistic 3D reassembly process. The IberianVoxel framework is generalizable to work with ceramic materials or other brittle objects (e.g., projectile points, ancient buildings, or anthropological remains).

We have evaluated the performance of IberianVoxel based on two complementary assessments. First, we used classical metrics to evaluate the generative process of voxelized samples across different fragment sizes (See Table 2 and Fig 3). Second, the results were validated via independent examination of archaeologists specialized in Iberian heritage. Results suggest that our approach can generate potteries that conform to the structure of Iberian ceramics and fulfill experts' validation criteria for the three different sizes of initial fragments tested. Although the described approach has an encouraging performance, shown in this work for the completing fragments of Iberian wheel-made pottery, some limitations were found and discussed in the paper. Nonetheless, our proposed framework and the available voxelized dataset are the first steps toward the broader use of 3D generative networks for the completion of fragments.

Acknowledgments

This work was supported by the European Union through the Programa Operativo FEDER Andalucía 2014-2020 Research Project under Grant UJA-1381115, the Center for Advanced Studies in Information and Communication Technologies (CEATIC) and the Research University Institute for Iberian Archaeology of the University of Jaén.

Ethical Statement

The research reported in this paper does not incur in any ethical issues.

References

[Chen *et al.*, 2020] S. Chen, H. Cui, P. Tan, X. Sun, Y. Ji, and H. Duh. Cantonese porcelain image generation using user-guided generative adversarial networks. *IEEE Computer Graphics and Applications*, 40(5):100–107, 2020.

[Cintas *et al.*, 2020] Celia Cintas, Manuel Lucena, José Manuel Fuertes, Claudio Delrieux, Pablo Navarro, Rolando González-José, and Manuel Molinos. Automatic feature extraction and classification of iberian ceramics based on deep convolutional networks. *Journal of Cultural Heritage*, 41:106–112, 2020.

[de Lima-Hernandez and Vergauwen, 2021] Roberto de Lima-Hernandez and Maarten Vergauwen. A generative and entropy-based registration approach for the reassembly of ancient inscriptions. *Remote Sensing*, 14(1):6, 2021.

[Di Angelo *et al.*, 2017] Luca Di Angelo, Paolo Di Stefano, and Caterina Pane. Automatic dimensional characterisation of pottery. *Journal of Cultural Heritage*, 26:118–128, 2017.

[Domaradzki and Martyn, 2016] Jakub Domaradzki and Tomasz Martyn. Fracturing sparse-voxel-octree objects using dynamical voronoi patterns. In *Computer Graphics, Visualization and Computer Vision WSCG 2016. Full Papers Proceedings*. Václav Skala-UNION Agency, 2016.

[Eslami *et al.*, 2020] Dariush Eslami, Luca Di Angelo, Paolo Di Stefano, and Caterina Pane. Review of computer-based methods for archaeological ceramic sherds reconstruction. *Virtual Archaeology Review*, 11(23):34–49, 2020.

[Feng *et al.*, 2016] Yuan Feng, Iwan Kawrakow, Jeff Olsen, Parag J Parikh, Camille Noel, Omar Wooten, Dongsu Du, Sasa Mutic, and Yanle Hu. A comparative study of automatic image segmentation algorithms for target tracking in mr-igrt. *Journal of applied clinical medical physics*, 17(2):441–460, 2016.

[Fragkos, Stergios *et al.*, 2018] Fragkos, Stergios, Tzimtzimis, Emanuel, Tzetzis, Dimitrios, Dodun, Oana, and Kyratsis, Panagiotis. 3d laser scanning and digital restoration of an archaeological find. *MATEC Web Conf.*, 178:03013, 2018.

[Goodfellow *et al.*, 2014] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.

[Hermoza and Sipiran, 2018] Renato Hermoza and Ivan Sipiran. 3d reconstruction of incomplete archaeological objects using a generative adversarial network. In *Proceedings of Computer Graphics International 2018*, pages 5–11. 2018.

[Isola *et al.*, 2017] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.

[Jboor *et al.*, 2019] N. Jboor, A. Belhi, A. Al-Ali, A. Bouras, and A. Jaoua. Towards an inpainting framework for visual cultural heritage. In *IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)*, pages 602–607, Amman, Jordan, 2019.

[Kalasarinis and Koutsoudis, 2019] Ioannis Kalasarinis and Anestis Koutsoudis. Assisting pottery restoration procedures with digital technologies. *International Journal of Computational Methods in Heritage Science (IJCMHS)*, 3(1):20–32, 2019.

[Kampel and Sablatnig, 2003] M. Kampel and R. Sablatnig. An automated pottery archival and reconstruction system. *Journal of Visualization and Computer Animation*, 14(3):111–120, 2003.

[Karasik and Smilansky, 2011] A. Karasik and U. Smilansky. Computerized morphological classification of ceramics. *Journal of Archaeological Science*, 38(10):2644 – 2657, 2011.

[Kashihara, 2017] Koji Kashihara. An intelligent computer assistance system for artifact restoration based on genetic algorithms with plane image features. *International Journal of Computational Intelligence and Applications*, 16(03):1750021, 2017.

[Kingma and Ba, 2015] Diederik P. Kingma and Jimmy Lei Ba. Adam: a Method for Stochastic Optimization. *International Conference on Learning Representations 2015*, pages 1–15, 2015.

[Kniaz *et al.*, 2019] V. V. Kniaz, F. Remondino, and V. A. Knyaz. Generative adversarial networks for single photo 3D reconstruction. In *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume 42, pages 403–408, 2019.

[Lazarou, 2020] Conor Lazarou. Autoencoding generative adversarial networks. *arXiv preprint arXiv:2004.05472*, 2020.

[Liu *et al.*, 2017a] M.-Y. Liu, T. Breuel, and J. Kautz. *Unsupervised image-to-image translation networks*, pages 700–708. CVPR, 2017.

[Liu *et al.*, 2017b] Ye Liu, Bo Zhang, and Liang Wan. Automatically unrolling decorations painted on 3d pottery. In *Proceedings of the Computer Graphics International Conference*, pages 1–6, 2017.

[Llamas *et al.*, 2016] Jose Llamas, Pedro M. Lerones, Eduardo Zalama, and Jaime Gómez-García-Bermejo. Applying deep learning techniques to cultural heritage images within the INCEPTION project. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 10059 LNCS, pages 25–32, 2016.

[Lucena *et al.*, 2017] M. Lucena, J. M. Fuertes, A. L. Martínez-Carrillo, A. Ruiz, and F. Carrascosa. Classification of archaeological pottery profiles using modal analysis. *Multimedia Tools and Applications*, 76(20):21565–21577, Oct 2017.

[Mirza and Osindero, 2014] Mehdi Mirza and Simon Osindero. Conditional generative adversarial nets. *CoRR*, abs/1411.1784, 2014.

[Mom, 2007] V. Mom. SECANTO - The SECtion Analysis TOol. In A. Figueiredo and G. Leite Velho, editors, *The world is in your eyes. CAA2005. Computer Applications and Quantitative Methods in Archaeology*, pages 95–101, Tomar, Portugal, 2007.

[Nautiyal *et al.*, 2006] V. Nautiyal, V. Dixit Kaushik, V. Kumar Pathak, S.G. Dhande, S. Nautiyal, M. Naithani, S. Juyal, R. Kumar Gupta, A. Kumar Vasisth, K. Kumar Verna, and A. Singh. Geometric modeling of indian archaeological pottery: A preliminary study. In J.T. Clark and E.M. Hagemeister, editors, *Exploring New Frontiers in Human Heritage. CAA2006. Computer Applications and Quantitative Methods in Archaeology*, Fargo, United States, 2006.

[Navarro *et al.*, 2022] Pablo Navarro, Celia Cintas, Manuel Lucena, José Manuel Fuertes, Rafael Segura, Claudio Delrieux, and Rolando González-José. Reconstruction of iberian ceramic potteries using generative adversarial networks. *Scientific reports*, 12(1):1–11, 2022.

[Nguyen *et al.*, 2019] Kha Cong Nguyen, Cuong Tuan Nguyen, Seiji Hotta, and Masaki Nakagawa. A character attention generative adversarial network for degraded historical document restoration. In *2019 International Conference on Document Analysis and Recognition (ICDAR)*, pages 420–425. IEEE, 2019.

[Ogayar-Anguita *et al.*, 2020] Carlos J. Ogayar-Anguita, Antonio J. Rueda-Ruiz, Rafael J. Segura-Sánchez, Miguel Díaz-Medina, and Ángel L. García-Fernández. A gpu-based framework for generating implicit datasets of voxelized polygonal models for the training of 3d convolutional neural networks. *IEEE Access*, 8:12675–12687, 2020.

[Papadopoulos *et al.*, 2021] Stavros Papadopoulos, Nikolaos Dimitriou, Anastasios Drosou, and Dimitrios Tzovaras. Modelling spatio-temporal ageing phenomena with deep generative adversarial networks. *Signal Processing: Image Communication*, 94, 2021.

[Raghavachary, 2002] Saty Raghavachary. Fracture generation on polygonal meshes using voronoi polygons. In *ACM SIGGRAPH 2002 Conference Abstracts and Applications*, SIGGRAPH '02, page 187, New York, NY, USA, 2002. Association for Computing Machinery.

[Rasheed and Nordin, 2014] Nada A Rasheed and Mohd Jan Nordin. A polynomial function in the automatic reconstruction of fragmented objects. *J. Comput. Sci.*, 10(11):2339–2348, 2014.

[Rice, 1987] Prudence M. Rice. *Pottery Analysis*. University of Chicago Press, Chicago, 1987.

[Salimans *et al.*, 2016] Tim Salimans, Ian Goodfellow, Wojciech Zaremba, Vicki Cheung, Alec Radford, and Xi Chen. Improved techniques for training gans. *Advances in neural information processing systems*, 29, 2016.

[Saragusti *et al.*, 2005] Idit Saragusti, Avshalom Karasik, Ilan Sharon, and Uzy Smilansky. Quantitative analysis of shape attributes based on contours and section profiles in artifact analysis. *Journal of Archaeological Science*, 32(6):841 – 853, 2005.

[Shennan and Wilcock, 1975] S.J. Shennan and J.D. Wilcock. Shape and style variation in central german bell beakers. *Science and Archaeology*, 15:17–31, 1975.

[Smith *et al.*, 2014] Neil G. Smith, Avshalom Karasik, Tejaswini Narayanan, Eric S. Olson, Uzy Smilansky, and Thomas E. Levy. The pottery informatics query database: A new method for mathematic and quantitative analyses of large regional ceramic datasets. *Journal of Archaeological Method and Theory*, 21(1):212–250, Mar 2014.

[Sorensen, 1948] Th A Sorensen. A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on danish commons. *Biol. Skar.*, 5:1–34, 1948.

[Velić *et al.*, 2009] Mirko Velić, Dave May, and Louis Moresi. A fast robust algorithm for computing discrete voronoi diagrams. *Journal of Mathematical Modelling and Algorithms*, 8:343–355, 08 2009.

[Wang *et al.*, 2017] Haiyan Wang, Zhongshi He, Yongwen Huang, Dingding Chen, and Zexun Zhou. Bodhisattva head images modeling style recognition of Dazu Rock Carvings based on deep convolutional network. *Journal of Cultural Heritage*, 27:60–71, 2017.

[Wilczek *et al.*, 2021] Josef Wilczek, Fabrice Monna, Nicolas Navarro, and Carmela Chateau-Smith. A computer tool to identify best matches for pottery fragments. *Journal of Archaeological Science: Reports*, 37:102891, 2021.

[Wu *et al.*, 2016] Jiajun Wu, Chengkai Zhang, Tianfan Xue, Bill Freeman, and Josh Tenenbaum. Learning a probabilistic latent space of object shapes via 3d generative-adversarial modeling. *Advances in neural information processing systems*, 29:82–90, 2016.

[Yeh *et al.*, 2017] R. Yeh, C. Chen, T. Lim, A. Schwing, M. Hasegawa-Johnson, and M. Do. Semantic image inpainting with deep generative models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5485–5493, 2017.

[Zachariou *et al.*, 2020] M. Zachariou, N. Dimitriou, and O. Arandjelović. Visual reconstruction of ancient coins using cycle-consistent generative adversarial networks. *Sci*, 2(3), 2020.

[Zhu *et al.*, 2017] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *2017 IEEE International Conference on Computer Vision (ICCV)*, pages 2242–2251, 2017.