

Interpretable Classification of Levantine Ceramic Thin Sections via Neural Networks

Sara Capriotti¹, Alessio Devoto², Simone Scardapane³, Silvano Mignardi¹, Laura Medeghini¹

¹ Department of Earth Sciences, Sapienza University of Rome, P.le Aldo Moro 5, 00185 Rome, Italy

² Department of Computer, Control and Management Engineering, Sapienza University of Rome, Via Ariosto 25, 00185 Rome, Italy

³ Department of Information Engineering, Electronics and Telecommunications, Sapienza University of Rome, Via Eudossiana 18, 00184 Rome, Italy

E-mail: sara.capriotti@uniroma1.it

Corresponding author: Sara Capriotti

Abstract. Classification of ceramic thin sections is fundamental for understanding ancient pottery production techniques, provenance, and trade networks. Although effective, traditional petrographic analysis is time-consuming. This study explores the application of deep learning models, specifically Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs), as complementary tools to support the classification of Levantine ceramics based on their petrographic *fabrics*. A dataset of 1,424 thin section images from 178 ceramic samples belonging to several archaeological sites across the Levantine area, mostly from the Bronze Age, with few samples dating to the Iron Age, was used to train and evaluate these models. The results demonstrate that transfer learning significantly improves classification performance, with a ResNet18 model achieving 92.11% accuracy and a ViT reaching 88.34%. Explainability techniques, including Guided Grad-CAM and attention maps, were applied to interpret and visualize the models' decisions, revealing that both CNNs and ViTs successfully focus on key mineralogical features for the classification of the samples into their respective petrographic *fabrics*. These findings highlight the potential of explainable AI in archaeometric studies, providing a reproducible and efficient methodology for ceramic analysis while maintaining transparency in model decision-making.

Keywords: Levantine ceramics, Materials Science, Deep Learning, Explainable AI
Submitted to: *Machine Learning: Science and Technology*

1. Introduction

The archaeological context of the Levantine region is both particularly significant and complex, as numerous archaeological sites emerged during the Early Bronze Age, a

period characterized by urbanization, trade networks, and political complexity [1, 2]. During this phase, the first urban societies in the region took shape. The process of urbanization was marked by the aggregation of local settlements, the formation of regional centers, and the fortification of villages, reflecting a shift toward more centralized form of organization. These transformations included the simplification of material culture and the reorganization of craft production [3]. One example of this process is Tell el-Far'ah (North), with strong fortifications and standardized ceramic production suggesting the presence of large workshops and inter-community exchange. While urbanization began earlier in the southern areas of the Levantine region, northern parts followed this path a bit later with key sites like Ebla showing signs of complex administration and organized settlement planning [4, 5]. Therefore, over the past decades, there has been a growing interest in studying this region to understand and reconstruct the changes and developments that have been observed in various archaeological sites. The study of ancient ceramics plays a central role in understanding the social and cultural identity of the region from the Bronze Age, as it provides information on social, cultural, and technological development as well as the evolution of ancient civilizations. Indeed, an archaeometric characterization of ancient ceramic materials can define the nature and provenance of raw materials used in production that help to define trade routes and interaction among past societies [6, 7, 8].

Currently, petrographic analysis of ceramic thin sections is one of the most effective and widely used methods for archaeometric characterization and reconstruction of the technological and material features of ceramics. The petrographic analysis also includes the identification of *fabrics*, groups of thin sections sharing similar characteristics, and representative of the *chaîne opératoire*, i.e. the actions and technological choices that led to the formation of the final product [9, 10]. Therefore, the grouping of thin sections in *fabrics* has been widely applied to define the different “recipes” used in ceramic production and to distinguish between local or imported wares [11]. Specifically, petrographic analyses can help in the recognition of specific ware types as locals and in identifying patterns of exchange of pottery within the region [12, 13, 14, 15]. In other cases, they have contributed to the identification of relationships between ancient production centers and their surrounding geological settings [16, 17, 18]. However, reconstructing ancient exchanges and defining trade profiles remain challenging due to the still limited petrographic studies on large ceramic assemblages, resulting in a lack of comparative data useful for identifying potential relationships and interactions. Moreover, although minero-petrographic characterization is a well-established and effective method for studying pottery assemblages, it is time-consuming and challenging to compare results across different archaeological contexts. This is partly due to the dispersion of petrographic data across various sources such as specialist journals, monographs, conference proceedings, and excavation reports.

In recent decades, there has been a growing interest in applying Machine Learning (ML) and/or Deep Learning (DL) techniques to archaeology in a variety of ways, including the identification of archaeological sites and structures [19, 20, 21], and the

classification or recognition of archaeological artifacts. Recent applications have involved the use of neural networks for the prediction of morphological features of lithic artifacts even when these are fragmented or incomplete [22, 23]. Other applications include the use of Laser-Induced Breakdown Spectroscopy (LIBS) combined with neural networks to attribute archaeological bone remains to specific individuals [24]. Additionally, Lithic Use-Wear Analysis (LUWA) using microscopic images has been employed to study the working traces on lithic artifacts [25]. Concerning the study of ceramics, automated methodologies have become powerful tools for resolving classification tasks, significantly contributing to the recognition of specific compositional, technological, or stylistic patterns [26, 27]. Today, several algorithms have been tested for the classification of pottery based on the study of their typology, decoration, and shape [28, 29]. ML models have also been applied in provenance analysis, where ceramic samples have been classified through the elaboration of geochemical data [30, 31, 32]. Among these approaches, Artificial Neural Networks (ANNs) have been employed to determine ceramic provenance using LIBS spectra [33]. ANNs have further been used for the analysis of ceramic thin sections with the aim of facilitating the *fabric* classification, in combination with image analysis to identify inclusions and pores in thin sections by working with parameters such as color, shape, size, percentage of inclusions and porosity [34]. Recent studies focused on the development of a DL model based on Convolutional Neural Networks (CNNs) to classify ceramic thin sections into their respective petrographic *fabrics* [35, 36].

CNNs are specifically designed for image pattern recognition. They consist of a series of layers among which the convolutional ones play a fundamental role, allowing the detection of characteristic patterns meaningful for accurate classification [37].

Recently, Transformers-based models have emerged as a significant alternative to CNNs. Vision Transformers [38] (ViTs) operate using multi-head self-attention instead of convolution and process input images by dividing them into patches. This approach captures both local and global dependencies, facilitating the identification of relationships across the entire image [39]. Yang et al. have explored the combination of channel and self-attention mechanisms within CNN architectures to improve their discriminative power in the classification of Ming-Qing ceramics [40].

Despite being extremely accurate, modern DL methods are highly complex, often consisting of hundreds of layers and parameters, making it difficult to determine which features the model relies on for classification. As a consequence, predictions are usually difficult to explain. For this reason, such models are often considered ‘black boxes’, as they do not provide any direct explanation or insights about their predictions [41]. In order to tackle this limitation, Explainable Artificial Intelligence (XAI) emerged as a field aimed at studying the interpretability of ML models [42, 43]. Understanding what the model has learned allows for further validation of its accuracy, also ensuring that it is based on the correct variables during the training phase, allowing the identification of potential limitations and building trust in the model’s performance [44, 45]. In the context of XAI, saliency maps and attention maps are visual tools that

Table 1: Summary of the selected ceramic thin sections.

Archaeological sites	Age	N. of samples
Tell el-Far’ah (North)	Early Bronze I - II	55
Khirbat al-Batrawy	Early Bronze III - IV	40
Khirbat Iskandar	Early Bronze IV	35
Bethlehem	Early Bronze IV - Iron Age IB-II	23
Ebla	Early Bronze IV - Middle Bronze I	18
Jericho	Early Bronze	7

help to understand a model’s decision [46]. To the best of our knowledge, neither ViTs nor explainability techniques have been applied to the ceramic petrographic *fabric* classification. The only similar study applied Grad-CAM (Gradient-based Class Activation Maps) to visualize the key features used by a CNN model for petrographic classification, although explainability was not its main focus [36]. Other related works have employed Grad-CAM to interpret CNN predictions for the typological and/or chronological classification of ceramic sherds and vessels [27, 47, 40].

The present research focuses on the application of automated methodologies for the classification of Levantine ceramics into their petrographic *fabrics* using deep learning models, specifically Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs). The aim is to investigate the potential of these models in the field of petrographic studies of pottery. The performance of these models is compared to evaluate their effectiveness in classification tasks. Furthermore, the research provides a partial interpretation of CNNs and ViTs using several forms of visual explanations. The objective is a deeper understanding of the classification process by visualizing the distinctive features that most influenced the models’ predictions.

2. Materials and Methods

The starting point for the implementation of the DL models consisted of a phase of selection and acquisition of a large number of ceramic thin section images, which have been used as training and testing datasets. A total of 178 ceramic samples were selected, primarily dating to the Bronze Age, with a smaller number from the Iron Age. These samples come from six different archaeological sites across the Levantine region: Bethlehem (West Bank), Tell el-Far’ah North (West Bank), Khirbat Iskandar (Jordan), Khirbat al-Batrawy (Jordan), Ebla (Syria) and Jericho (West Bank) as shown in Fig. 1 (see Table 1 for more detailed information on the samples). While most of the sites are located in the southern part of the region, the inclusion of Elba provides important comparative data from the north. All samples have already been described and classified in terms of their respective petrographic *fabrics*, as well as their archaeological context in previous papers [48, 49, 50, 51, 52, 53, 54, 55]. The selection aimed to create a dataset representative of the Levantine pottery production profile, according to manufacturing practices, technological choices, and regional variability.



Figure 1: Map of the Levant showing the selected archaeological sites.

Acquisition of thin section images was performed using the Axiocam 208 ZEISS camera connected to a ZEISS D-7082 Oberkochen petrographic microscope. For each ceramic section, we took multiple images under plane-polarized light (PPL) and cross-polarized light (XPL) at magnifications of 2.5X and 10X, resulting in a total of eight images per section to ensure the capture of important microstructural and mineralogical features. All images were compared and grouped into ten petrographic *fabrics*, taking into account the results of the previous works. Samples from the different *fabrics* are presented in Fig. 2. The final dataset, consisting of 1,424 images grouped in 10 classes, is randomly split into training and testing sets with a ratio of 80:20. The larger portion was used for the training phase, while the smaller one was used to test the models on new data.

2.1. Implementation of CNN and ViT models

We select two types of deep learning models for this study: A CNN using the ResNet18 [56] architecture, and the more recent Vision Transformer [38]. CNNs are neural

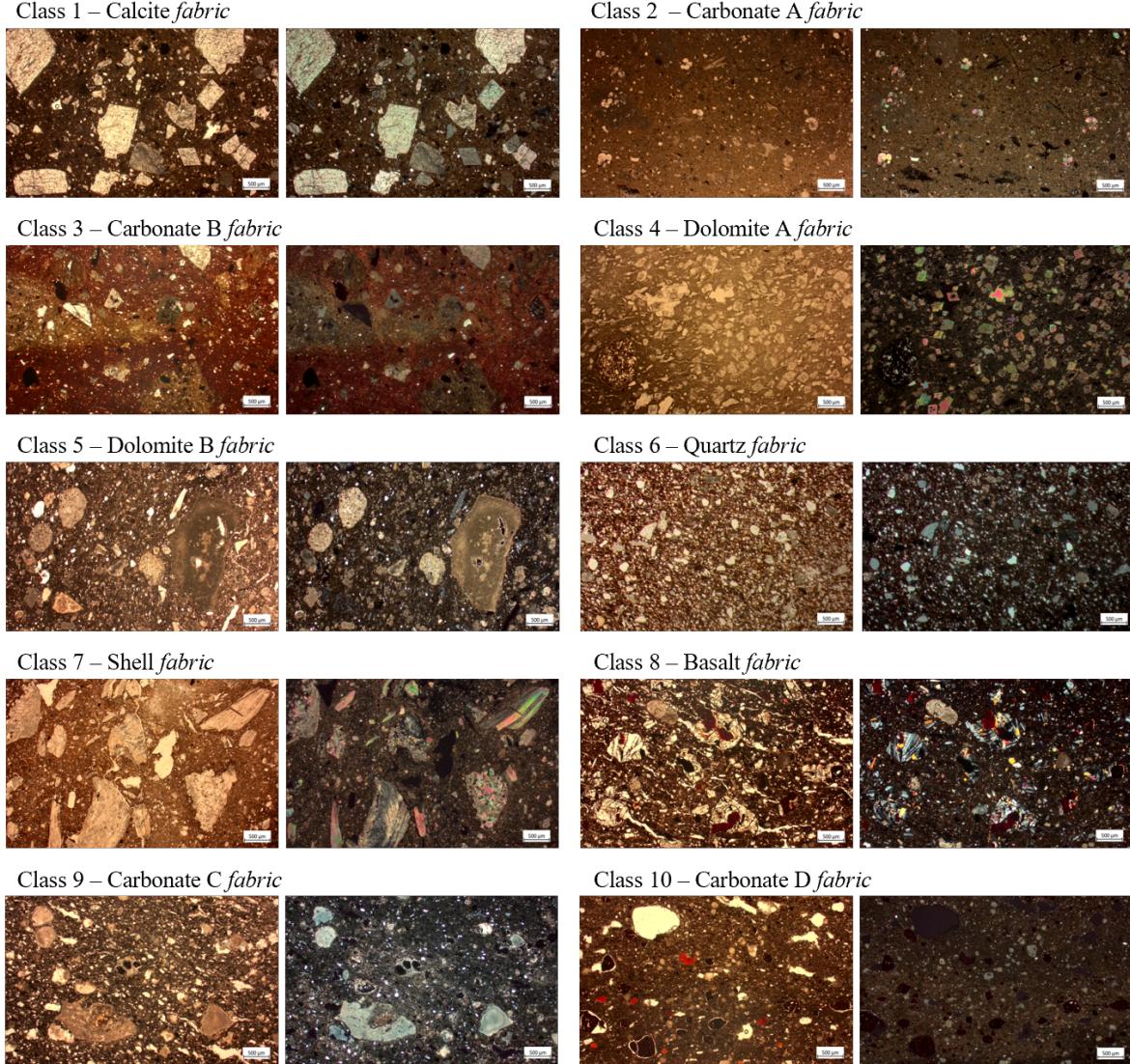


Figure 2: Thin section images of representative samples of the ten petrographic *fabrics* (2.5X, both in PPL and XPL respectively). Class 1 (sample T53): Characterized by the predominance of calcite inclusions; Class 2 (sample TF6): Predominance of fragments of sedimentary calcareous rocks, presence of microfossils in the matrix; Class 3 (sample TFN39): Fragments of sedimentary calcareous rocks, presence of quartz in the matrix; Class 4 (sample B1/14): Predominance of dolomite crystal inclusions; Class 5 (sample K7): Presence of dolomite, calcareous and siliceous rock fragments as dominant inclusions; Class 6 (sample TFMcamp1): Presence of predominant quartz inclusions; Class 7 (sample K12): Predominant presence of elongated bivalve shells and calcareous rock fragments; Class 8 (sample E467/3): Dominant presence of volcanic rock fragments identified as basalt; Class 9 (sample K11): Characterized by the dominant presence of calcareous and siliceous rock fragments; Class 10 (sample KB314/113): Predominance of calcareous fragments and iron oxide nodules. Complete description of these *fabrics* in [48, 49, 50, 51, 52, 53, 54, 55].

networks that process images through multiple convolutional layers, each detecting increasingly complex features. ResNet18 consists of 18 layers organized in 5 stages: an initial 7×7 convolutional layer followed by 8 blocks, each containing two 3×3 convolutional layers with batch normalization and ReLU activation. Each block includes a skip connection that adds the input to the block’s output, helping to prevent vanishing gradients. The network concludes with an average pooling layer and a fully connected layer for classification.

Vision Transformers represent a fundamentally different approach: they first divide input images into fixed-size patches (16×16 pixels in our case), linearly embed each patch, add positional embeddings, and process them through a series of transformer encoder blocks. Each encoder block contains a multi-head self-attention layer followed by a multi-layer perceptron (MLP). Our ViT implementation uses 12 transformer encoder blocks, each with 12 attention heads and a hidden dimension of 384.

We train our models to predict the class of a given image, minimizing a cross entropy loss function. To evaluate classification performance, we train the models and subsequently compare their performance on the held-out test dataset.

For both models, we use a transfer learning approach, fine-tuning pre-trained versions on the large ImageNet dataset [57]. This approach leverages the pre-training on ImageNet, a large dataset of 1.2 million images across 1,000 classes. We retain the pre-trained weights of the backbone layers, that have already learned to extract general visual features like edges, textures, and shapes, and only retrain the final classification layers to adapt to our specific classes. This transfer learning strategy allows the model to build upon previously learned features, making it more effective when working with limited data.

As part of our ablation study, we also evaluate a from-scratch approach, where all model parameters are randomly initialized and trained exclusively on our dataset.

We trained the ResNet-18 models for 60 epochs (from scratch) and 25 epochs (pre-trained) respectively. Both implementations used the AdamW optimizer [58] with learning rates and weight decay values of 3e-4. An adaptive learning rate scheduler was incorporated to dynamically adjust the learning rate throughout the training process. For the ViT models, training extended across 125 epochs (from scratch) and 60 epochs (pre-trained) using AdamW with a learning rate of 1e-4 and 1e-5 respectively, and a weight decay of 1e-4. To enhance model generalization and effectively expand our training dataset, we implemented various data augmentation techniques [59] including random axis flipping, color jittering, and random cropping. These augmentations were applied consistently during the training phase for both models to mitigate overfitting. We provide further details concerning the training hyper-parameters in Appendix A.

3. Results

The mean and standard deviation of the metrics for CNNs and ViTs are summarized in Table 2. The reported values correspond to the average performance of the models

Table 2: Mean and standard deviation of the accuracy, precision, recall (True positive rate) and F1-score (harmonic mean of precision and recall) metrics of the CNN and ViT.

Model Metrics	ResNet18 (Scratch)	Pre-trained ResNet18	ViT (Scratch)	Pre-trained ViT
Accuracy	$76.33 \pm 0.58\%$	$92.11 \pm 0.44\%$	$61.01 \pm 0.33\%$	$88.34 \pm 0.29\%$
Precision	$78.10 \pm 0.73\%$	$92.05 \pm 1.35\%$	$58.74 \pm 1.21\%$	$88.44 \pm 0.50\%$
Recall	$72.16 \pm 1.08\%$	$90.15 \pm 1.58\%$	$59.93 \pm 0.60\%$	$86.60 \pm 0.83\%$
F1-score	$73.67 \pm 0.71\%$	$90.68 \pm 1.54\%$	$58.03 \pm 0.47\%$	$86.84 \pm 0.70\%$

initialized three times with different random seeds. For reference, we also show the results of the models trained from scratch without a pre-trained ImageNet checkpoint. As expected, the transfer learning approach led to a significantly higher performance. The ResNet18 model trained from scratch achieved an accuracy of $76.33 \pm 0.58\%$ after 60 epochs with a precision of $78.10 \pm 0.73\%$, a result consistent with the small size of the dataset. In contrast, the pre-trained ResNet18, initialized with ImageNet weights, reached an accuracy of $92.11 \pm 0.44\%$ and a precision of $92.05 \pm 1.35\%$ in 25 epochs. ViTs are generally more effective for large datasets; as expected, the model trained from scratch underperformed, with an accuracy of $61.01 \pm 0.33\%$ and a precision of $58.74 \pm 1.21\%$ after 125 epochs. However, the pre-trained implementation performed well despite the small dataset, achieving an accuracy of $88.34 \pm 0.29\%$ and a precision of $88.44 \pm 0.50\%$ in 60 epochs, highlighting the effectiveness of transfer learning even with limited data.

Fig. 3a presents the confusion matrix from the best run of the pre-trained ResNet18 model, showing the correlation between predicted and actual values and highlighting the number of correctly classified samples. The adjacent plot illustrates the evolution of training and testing accuracy over the epochs, allowing us to evaluate the learning progress of the model and identify potential overfitting or underfitting issues. Similarly Fig. 3b visualizes the best run results of the ViT model, presenting its confusion matrix along with the training and testing accuracy trends over a total of 60 epochs. It is important to note that the accuracy on the test set was monitored for visualization purposes only. The final evaluation of each model was based on its performance on the test set after training was completed.

3.1. Analysis of misclassified samples

In this section we perform an in-depth analysis of the confusion matrix of the models to verify whether the most difficult samples to assign to specific petrographic *fabrics* correspond to those misclassified by the models and also whether these samples are consistently misclassified, identifying potential biases or problematic samples. From the analysis of the confusion matrix of both the models (Fig. 3), it emerges that class 2 has the highest number of misclassifications with class 3. Class 2 represents the carbonate A *fabric*, characterized by the presence of microfossils in the matrix, while the main inclusions are carbonate rock fragments, making it very similar to the

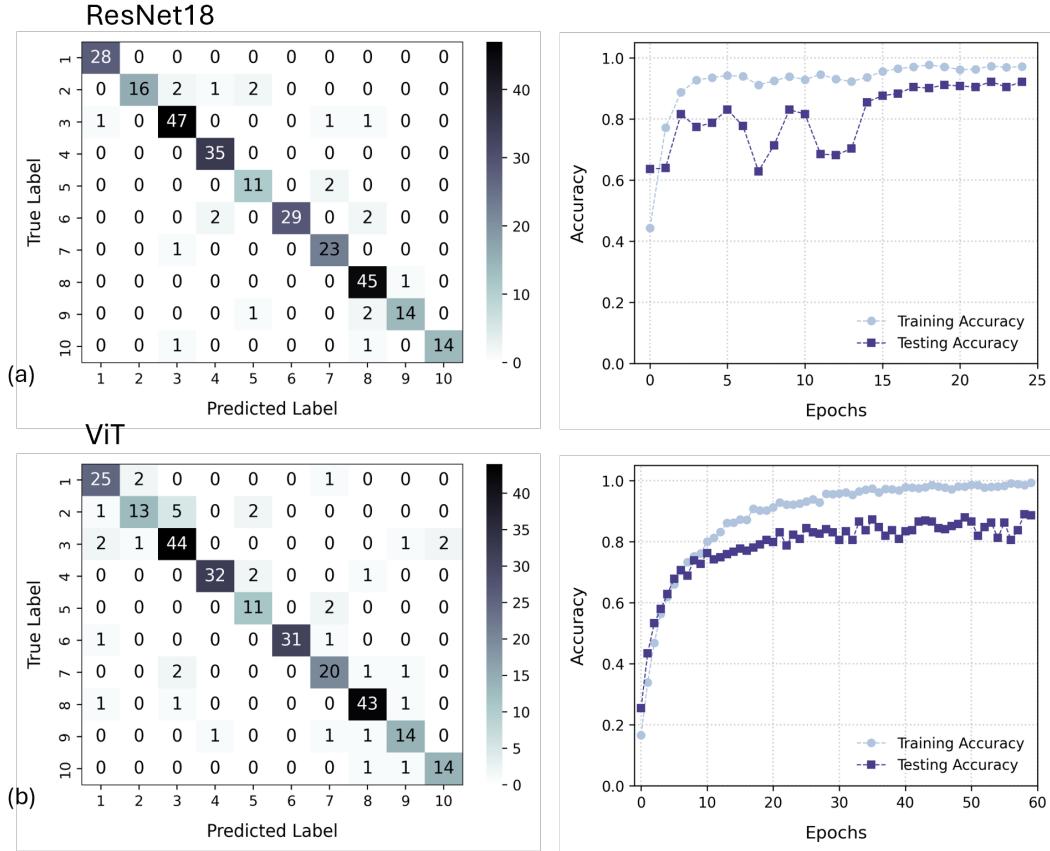


Figure 3: Classification performance of ResNet18 and Vision Transformer (ViT). (a) Confusion matrix and accuracy trends for ResNet18. (b) Confusion matrix and accuracy trends for ViT.

carbonate B *fabric* (class 3). The main difference between them is that carbonate B *fabric* contains an abundant presence of quartz instead of microfossils [48]. Observing the misclassifications of the ResNet18 and ViT models (Fig. 4a and b), we note that the first three images of sample J3 are misclassified in both models, mainly being confused with class 3. It is interesting to highlight that this sample was particularly difficult to classify, as its matrix contains both microfossils and dispersed quartz. This combination of characteristics places it in an intermediate zone between carbonate A and carbonate B *fabrics*, making its classification more complex. These observations suggest that both the models can struggle with classes that share similar mineralogical characteristics, particularly when a sample exhibits features of multiple petrographic *fabrics*. This kind of ambiguity also makes classification difficult for human experts, suggesting that some misclassifications may reflect the complexity of the samples more than limitations in model performance.

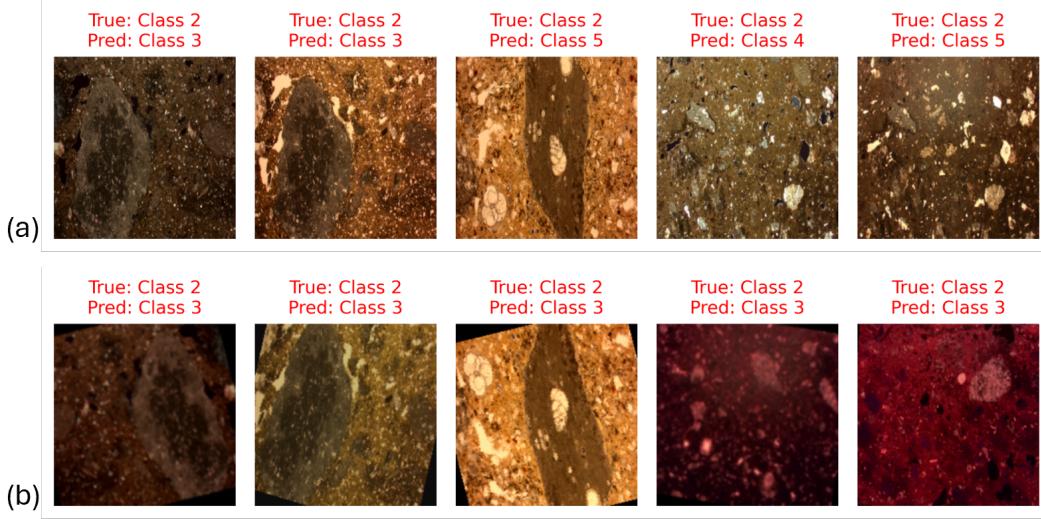


Figure 4: Comparison of ResNet18 and ViT misclassified predictions for class 2, representative of carbonate A *fabric*. (a) Images of ResNet18 misclassified predictions; (b) Images of ViT misclassified predictions.

4. Explainability analysis

4.1. Explainability Techniques

Given that our study involves two distinct neural network architectures, ResNet18 and ViT, we employ different explainability techniques suited to their respective structures.

For ResNet18, we apply the Guided Gradient-weighted Class Activation Mapping (Guided Grad-CAM) [60], a widely used interpretability technique for convolutional neural networks. Grad-CAM generates a coarse localization map by computing the gradient of the output with respect to the feature maps in the last convolutional layer, highlighting the most influential regions in the input image for the network’s decision.

For the Vision Transformer, we employ attention map visualization to interpret model decisions. Unlike CNNs, which rely on local receptive fields, ViTs use self-attention mechanisms to establish long-range dependencies between tokens [61]. Each ViT layer consists of multiple attention heads, each learning distinct patterns of feature dependencies across the input sequence. By visualizing attention maps across these heads, we gain insights into how the model distributes its focus over the image. Attention maps are particularly informative for ViTs because they directly represent the learned relationships between different image patches. In early layers, attention heads often capture local structures, while deeper layers aggregate global contextual information. By analyzing these attention distributions, we can interpret whether the model attends to semantically meaningful regions, such as object boundaries or discriminative features, aligning with human intuition for visual reasoning.

4.2. Explainability of the models in classifying ceramic thin-section

Guided Grad-CAM was used to visually explain the predictions of the pre-trained CNN model, highlighting the important features considered for classification. This technique allows us to observe how the model effectively identified the minerals characteristic of the *fabrics* as representative for classification. The regions indicated by the Guided Grad-CAM as important correspond to the predominant minerals that lead to the discrimination and grouping of the ceramic samples according to their petrographic *fabrics*. Fig. 5 provides several significant examples, where the heatmaps allow for an immediate and clear visualization of the areas that positively contributed to the correct classification of the images. To analyze potential differences, we applied this method to both PPL and XPL images of the same samples. Identifying key features in PPL images is more challenging, as they appear less distinct compared to XPL images, making mineral discrimination more difficult. However, the results suggest that inclusions can still be distinguished from the matrix based on their shape and color. In XPL images, interference colors further help in distinguishing the main inclusions. Particularly interesting are the results for class 8. The images in Fig. 5d show how the heatmaps exclusively highlight the basalt fragments, distinguishing them from the other inclusions present in the thin section in both the PPL and XPL images. This demonstrates the neural network's ability to identify and learn specific characteristics of the corresponding petrographic *fabrics*. See Fig. B1 in Appendix B to visualize the Guided Grad-CAM results for the other classes.

To deeper investigate these results, we applied Guided Grad-CAM to a sequence of images of the same sample at different rotations to determinate which parameters most influenced the classification and whether rotation and birefringence effects had any impact on the results. This analysis was conducted on calcite and quartz *fabric* samples. Fig. 6 shows the rotation sequence of a calcite *fabric* sample, showing its progressive rotation until some crystals reached extinction. Calcite, characterized by high birefringence, is particularly distinctive in optical analyses and is a perfect case study to determine which features are considered important for classification. The resulting heatmaps indicate that rotation does not affect the recognition of specific patterns for classification. On the contrary, they highlight how the interference colors of the calcite crystals, their shape and edges, and also those in extinction are key elements considered by the model for classification. Regarding the analysis conducted on the representative quartz *fabric* sample in Fig. 7, the heatmaps highlight quartz minerals independent of their size and/or brightness. Color appears to play a role in identification; in fact, the extincted quartz is not well evidenced, probably because the model is not able to discriminate it from the matrix as it is quite dark as well. Additionally, it is interesting to note that there are few carbonate rock fragments present in the sample that are not highlighted by the heatmaps. This result suggests that the model selectively recognizes quartz minerals without considering other inclusions, thus confirming the consistency of the classification process.

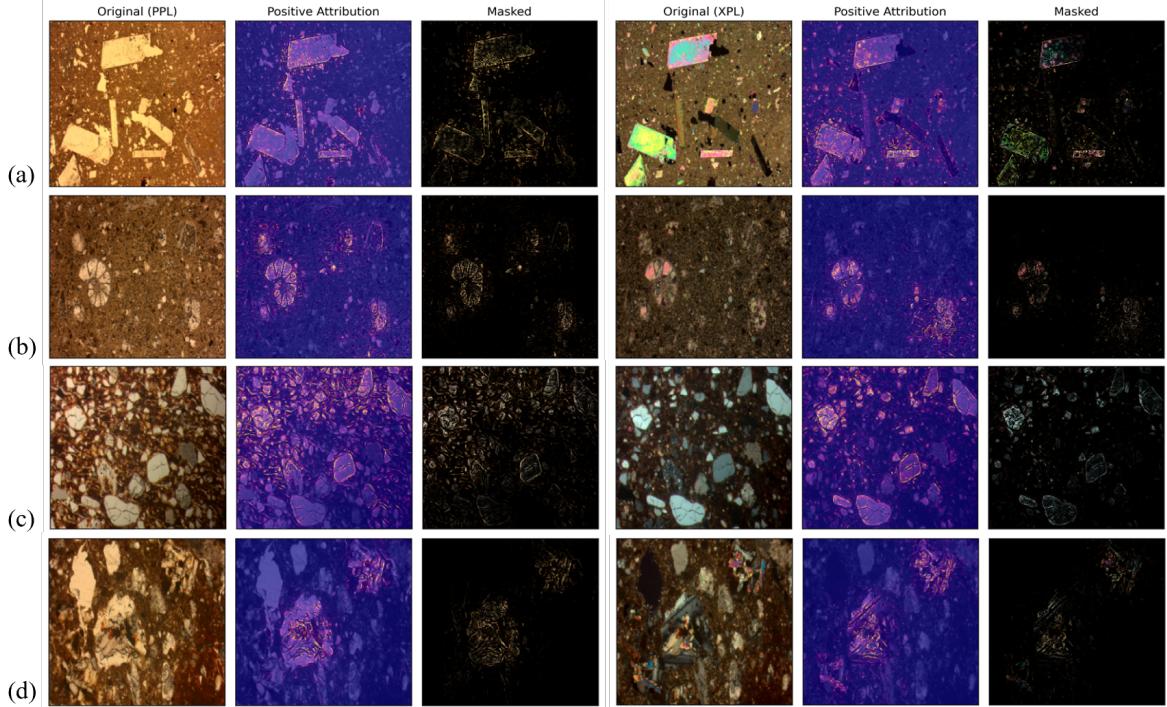


Figure 5: Guided Grad-CAM applied to pre-trained ResNet18. The figure shows the original images, the heatmaps overlaid to the thin-section images, and the masked images highlighting only the areas considered important for classification. (a) Class 1 (sample TF4): Calcite *fabric*; (b) Class 2 (sample TF6): Carbonate A *fabric*; (c) Class 6 (sample TFMcamp1): Quartz *fabric*; (d) Class 8 (sample E16/2): Basalt *fabric*.

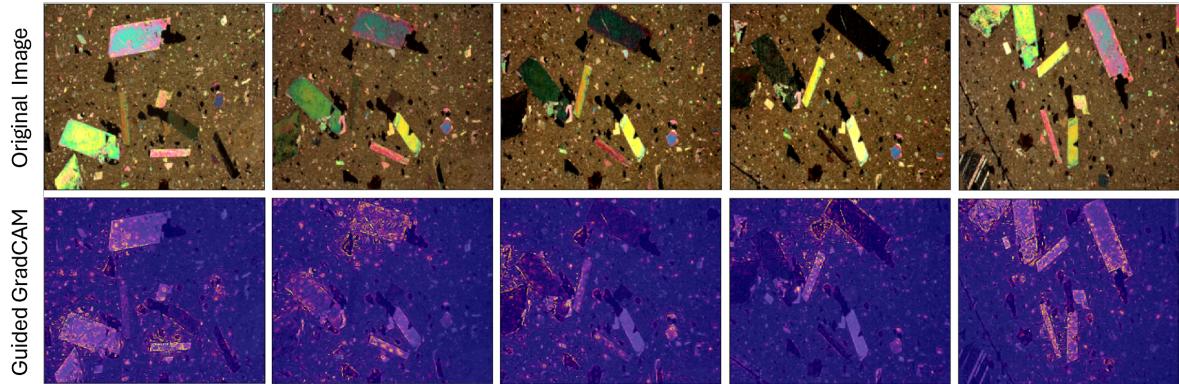


Figure 6: Guided Grad-CAM applied to the rotation sequence of sample TF4 from calcite *fabric*.

We use attention maps to visualize the classification predictions of the pre-trained ViT model, illustrating how it correlates different patches during the classification process. Fig. 8 shows attention maps of representative samples from classes 2, 5, 7, and 8, which correspond to carbonate A, dolomite B, shell, and basalt *fabrics*, respectively. Specifically, the figure presents attention maps from head 3 for class 2 and 5 and from head 5 for class 7 and 8 across different layers, allowing a better observation of the evolution of the attention over the layers of the ViT. To investigate potential differences

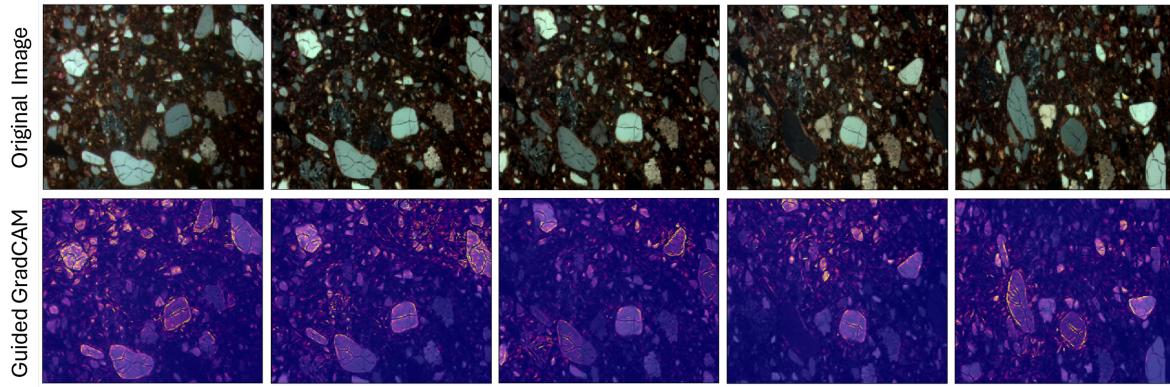


Figure 7: Guided Grad-CAM applied to the rotation sequence of sample TFMcamp1 from quartz *fabric*.

in attention maps under different polarization conditions, we visualized the samples in both PPL and XPL. The results indicate that the attention maps remain consistent across both images, suggesting that the model’s ability to focus on relevant mineralogical features is not significantly affected by polarization. Furthermore, the evolution of attention maps throughout the layers of the ViT appears consistent for each respective attention head. In the first layers, attention is more broadly distributed, while in deeper layers, it starts to focus on specific regions corresponding to the characteristic minerals of the petrographic *fabrics*.

The most challenging aspect of visualizing the important areas in this case is that the inclusions are distributed throughout the entire thin section. This makes it more difficult for the model to focus on specific features, especially considering that ViTs operate globally by analyzing patches rather than individual pixels. For this reason, it is particularly interesting to emphasize the results of classes 5 and 7, which contain a high number of inclusions dispersed in the matrix. In both cases, the attention maps show that the model was still able to focus on specific characteristics of the respective *fabrics*. In class 5, the model mainly highlights some well-defined dolomite crystals that stand out clearly from the matrix, as well as the silicate rock fragment that is characteristic of this class. A similar pattern can be observed for the shell *fabric*, where the attention maps seem to emphasize the areas corresponding to the elongated shell fragments rather than the more common carbonate fragments. We report the attention maps of representative samples from the other classes in Appendix B, Fig. B2.

Both the saliency and attention maps clearly highlight the characteristic minerals of the petrographic *fabrics* as key features for classification. This indicates that the models distinguished the ceramic thin sections based on their inclusions, similar to the way traditional petrographic analysis groups ceramics by studying pores, inclusions, and the matrix. Although our models focused mainly on inclusions, it is interesting to note a certain correlation between deep learning-based classification and conventional petrographic methods.

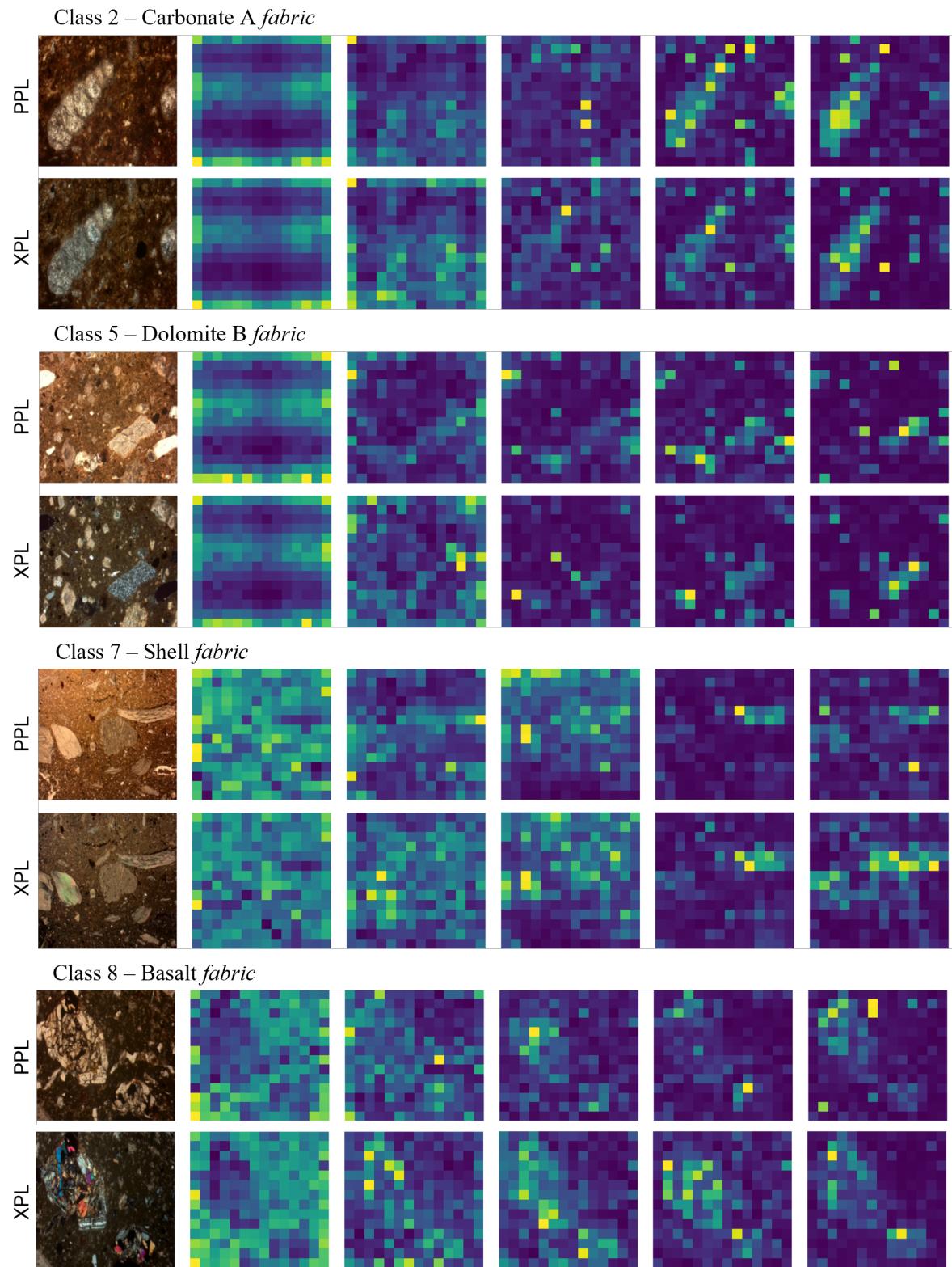


Figure 8: Attention maps of representative samples from carbonate A (sample KB146/30), dolomite B (sample K9), shell (sample KI11) and basalt (sample KB5) *fabrics* both in PPL and XPL. Visualization of head 3 for class 2 and 5, and head 5 for class 7 and 8, across ViT's layers.

5. Conclusion

This study demonstrated the potential of deep learning methodologies for the automated classification of ceramic thin sections based on their petrographic *fabrics*, expanding on previous research by Lyons et al. [35, 36]. The application of Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) yielded high classification accuracy and enabled the grouping of thin sections with similar petrographic characteristics. Furthermore, explainability techniques, such as Guided Grad-CAM for CNNs and attention maps for ViTs, provided insights into the models' decision-making processes, confirming their ability to recognize key mineralogical features relevant for *fabric* classification.

One possible limitation arises from the use of saliency maps, such as Guided Grad-CAM, which, while effective in highlighting important image regions, may not always provide a complete or precise explanation of model decisions. To address these challenges, future research could explore more recent explainability techniques, such as sparse autoencoders [62]. Furthermore, integrating multimodal data sources, such as geochemical analysis alongside petrographic imaging, could improve classification accuracy and robustness, opening the possibility for more applications in the field of minero-petrographic studies. For instance, this approach could contribute to ceramic provenance research by aiding in the identification of raw material sources in relation to the geological context of surrounding areas.

A key future step is expanding the dataset of ceramic thin section images to improve model generalization and performance across a broader range of *fabrics* and archaeological contexts. One current limitation is that models trained on a predefined set of known classes may struggle to generalize to previously unseen samples, particularly given the high degree of ceramic variability in the Levantine region, constraining their robustness in real-world applications. As a future step, it will be essential to evaluate the model on additional sets of thin-section images, distinct from those used during training. This will allow for a more realistic assessment of the generalization capabilities of the model, including its ability to identify samples that do not match any of the known classes (i.e., out-of-distribution detection). Moreover, recent advances in unsupervised learning and few-shot learning offer promising tools to address this challenge, potentially expanding the applicability of deep learning to archaeological scenarios.

Overall, this study highlights the potential of explainable AI in archaeometric research, providing a reproducible and efficient approach to ceramic analysis. Continued advancements in AI interpretability will be crucial for refining these methodologies, ensuring transparency, and facilitating broader adoption in archaeological and materials science applications.

Acknowledgements

The authors acknowledge the support of Sapienza University of Rome. In particular, S. Scardapane and A. Devoto are partly funded by the Sapienza grant RG123188B3EF6A80 (CENTS).

Appendix A. Details on Model Architectures and Training

This section presents the architectural details and training configurations for both the ResNet-18 and DeiT-Small models used in our experiments. Hyperparameter selection was carried out based on the results of a 3-fold cross-validation performed on the training portion of the dataset (80% of the full set). The best average validation results guided the final choice of hyperparameters. The explored hyperparameters are reported in tables A3 and A6.

Table A1: ResNet-18 Architecture Hyperparameters.

Parameter	Value
Total layers	18
Convolutional layers	17 + 1 fully connected
Residual blocks	8
Input resolution	224×224
Parameters	$\sim 11.7M$

Table A2: ResNet-18 Training Setup.

Hyperparameter	Value
Learning rate	3e-4
Weight decay	3e-4
Training epochs (scratch)	60
Training epochs (pre-trained)	25
Batch size (scratch)	20
Batch size (pre-trained)	32
Optimizer	AdamW
Learning rate scheduler	ReduceLROnPlateau

Table A3: ResNet18 hyperparameters explored during 3-fold cross-validation, for both pre-trained and from-scratch models.

Hyperparameter	Values Tested
Learning rate	{3e-4, 1e-4}
Weight decay	{3e-4, 1e-4, 1e-5}
Optimizer	{AdamW}

Table A4: DeiT-Small Architecture Hyperparameters.

Parameter	Value
Layers (Transformer blocks)	12
Heads	6
Hidden dimension	384
MLP dimension	1536
Patch size	16×16
Input resolution	224×224
Parameters	$\sim 22M$

Table A5: Vision Transformer Training Setup.

Hyperparameter	Value
Learning rate (from scratch)	1e-4
Learning rate (pre-trained)	1e-5
Weight decay	1e-4
Training epochs (scratch)	125
Training epochs (pre-trained)	60
Batch size	32
Optimizer	AdamW
Learning rate scheduler	CosineLR

Table A6: DeiT-Small hyperparameters explored during 3-fold cross-validation, for both pre-trained and from-scratch models.

Hyperparameter	Values Tested
Learning rate	{1e-4, 1e-5}
Weight decay	{1e-4, 1e-5}
Optimizer	{AdamW}

Appendix B. Guided Grad-CAM and Attention Maps of Other Classes

In this section, we present the results of Guided Grad-CAM and attention maps for the remaining classes that were not shown in the main text of the paper.

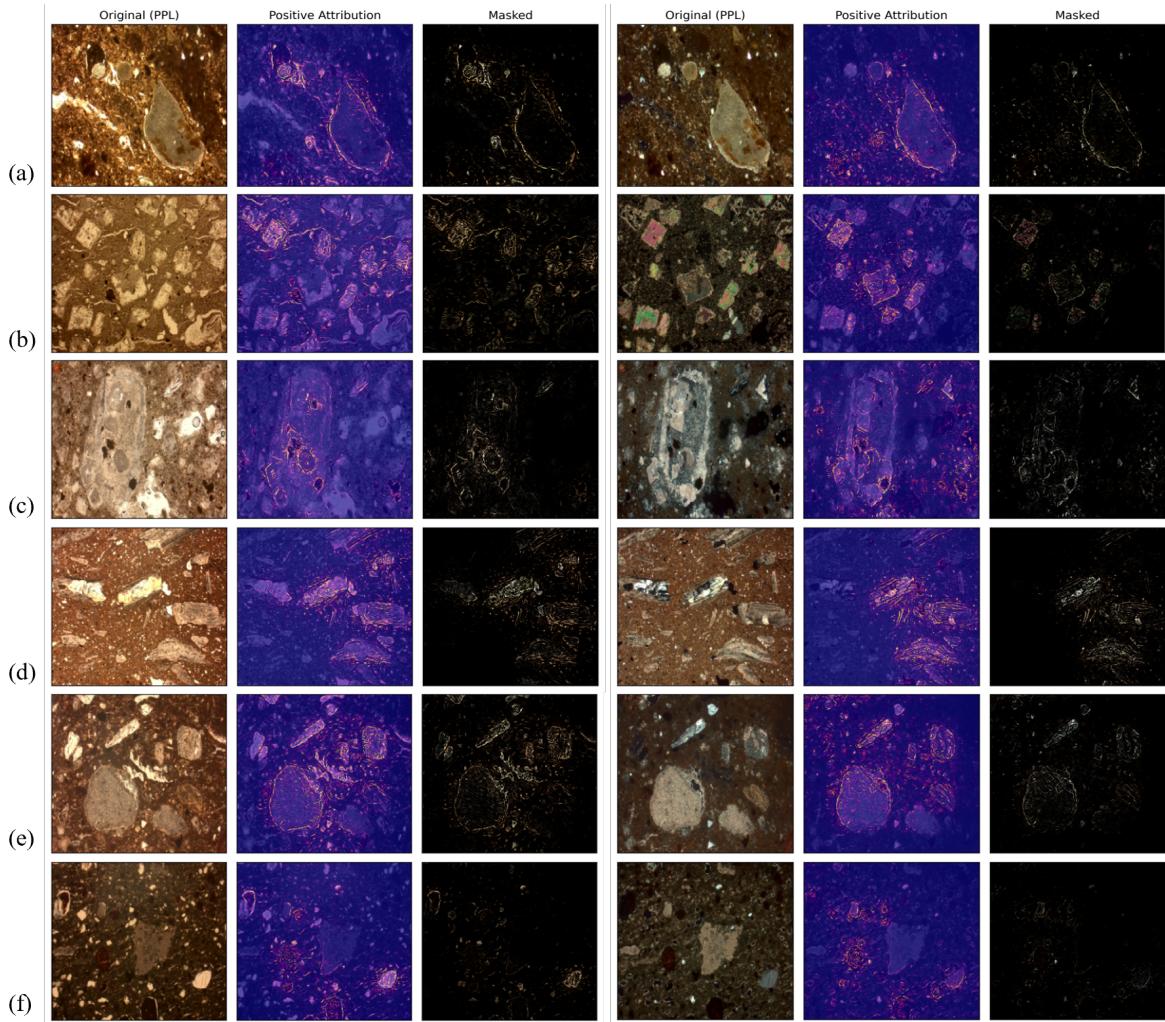


Figure B1: Guided Grad-CAM applied to pre-trained ResNet18. The figure shows the original images, the heatmaps overlaid to the thin-section images, and the masked images highlighting only the areas considered important for classification. (a) Class 3 (sample 718/25): Carbonate B *fabric*; (b) Class 4 (sample B1/14): Dolomite A *fabric*; (c) Class 5 (sample Kcamp15): Dolomite B *fabric*; (d) Class 7 (sample KIcamp2): Shell *fabric*; (e) Class 9 (sample K1): Carbonate C *fabric*; (f) Classe 10 (sample KB383/38): Carbonate D *fabric*.

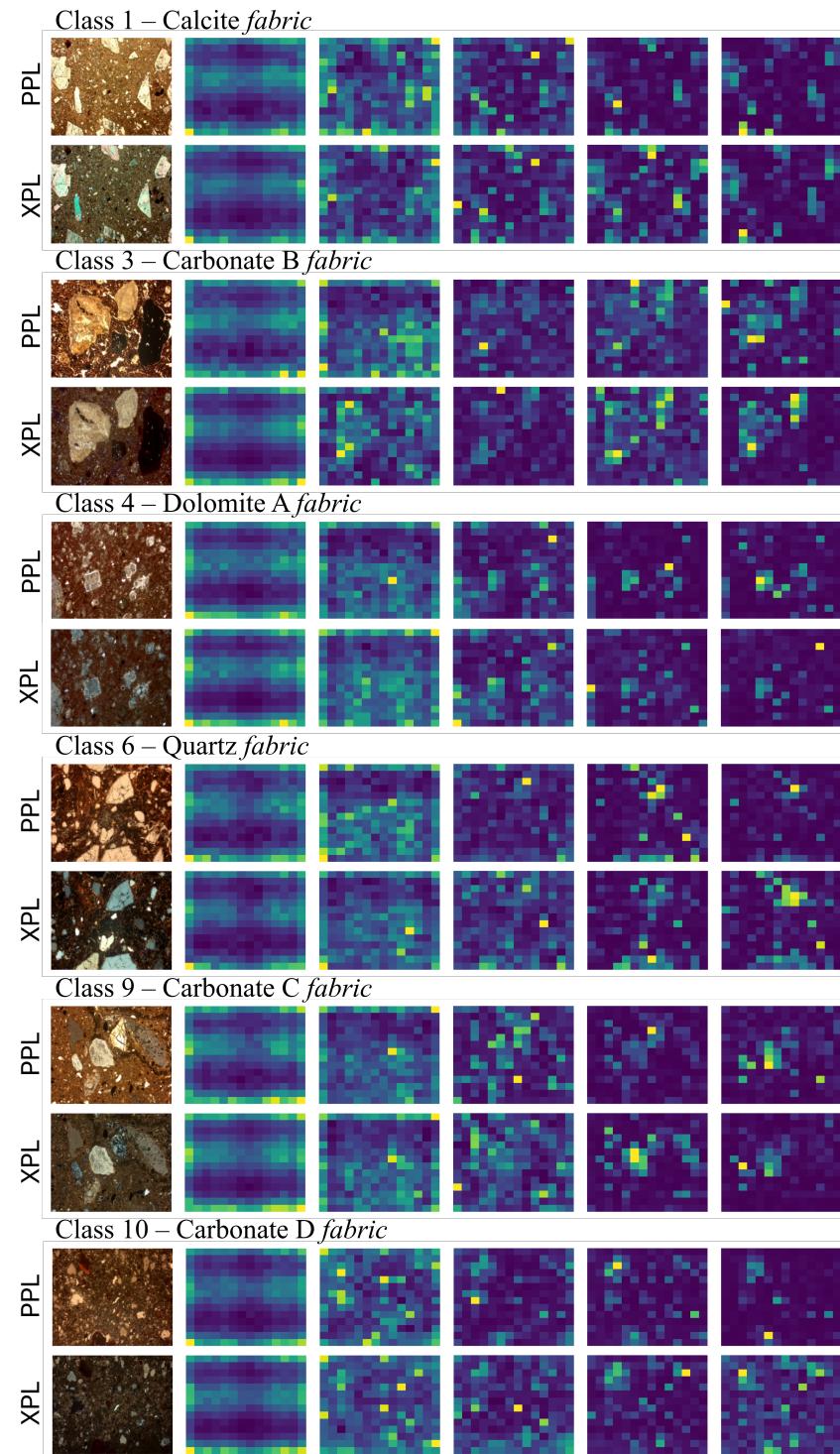


Figure B2: Attention maps of representative samples from calcite (sample TFcamp19), carbonate B (sample TFcamp13), dolomite A (sample Asur5), quartz (sample KBC4), carbonate C (sample K16) and carbonate D (sample KB314/113) *fabrics* both in PPL and XPL. Visualization of head 3 across the ViT's layers.

References

- [1] M. Steiner and A. E. Killebrew. *The Oxford Handbook of the Archaeology of the Levant: c. 8000–332 BCE*. Oxford Handbooks, 2014.
- [2] R. Greenberg. *The archaeology of the bronze age levant*. Cambridge University Press, 2019.
- [3] R. Greenberg. *Traveling in (world) time: transformation, commoditization, and the beginnings of urbanism in the Southern Levant*. Oxbow Books, 2011.
- [4] H. Weiss. The northern levant during the intermediate bronze age. In Margreet L. Steiner and Ann E. Killebrew, editors, *The Oxford Handbook of the Archaeology of the Levant: c. 8000–332 BCE*, pages 401–412. Oxford University Press, Oxford, 2013.
- [5] M. S. Chesson and G. Philip. Tales of the City? “Urbanism” in the Early Bronze Age Levant from Mediterranean and Levantine Perspectives. *Horizon*, 1(2):26–37, 2003.
- [6] P. S. Quinn. *Ceramic Petrography: The Interpretation of Archaeological Pottery & Related Artefacts in Thin Section*. Archaeopress Publishing, 2013.
- [7] C. Orton and M. Hughes. *Pottery in Archaeology*. Cambridge University Press, 2013.
- [8] G. Yu, Z. Jin, L. Chen, F. Wang, X. Wang, X. Wu, A. Fan, and Q. Xia. Analyzing the earliest Chinese proto-porcelain: Study on the materials from Liaotianjianshan kiln sites, Dehua County, Fujian Province (China). *Ceramics International*, 44(17):21648–21655, 2018.
- [9] I. K. Whitbread. The characterisation of argillaceous inclusions in ceramic thin sections. *Archaeometry*, 28(1):79–88, 1986.
- [10] I. K. Whitebread. *Greek transport amphorae. A petrological and archaeological study*. British School of Athens., 1995.
- [11] I. K. Whitbread and A. Hunt. *Fabric description of archaeological ceramics. The Oxford handbook of archaeological ceramic analysis*. Oxford University Press, 2017.
- [12] Y. Goren. The Southern Levant in the Early Bronze Age IV: The Petrographic Perspective. *Bulletin of the American Schools of Oriental Research*, 303(1):33–72, 1996.
- [13] R. Greenberg and N. Porat. A Third Millennium Levantine Pottery Production Center: Typology, Petrography, and Provenance of the Metallic Ware of Northern Israel and Adjacent Regions. *Bulletin of the American Schools of Oriental Research*, 301(1):5–24, 1996.
- [14] Y. Goren. Ceramic Technology and Provenance at Khirbat Iskandar. In S. Richard, J. C. Long, P. S. Holdorf, and G. Peterman, editors, *Archaeological Expedition to Khirbat Iskandar and its Environs, Jordan: Khirbat Iskandar Final Report on the Early Bronze IV Area C ‘Gateway’ and Cemeteries*, pages 133–140. The American Schools of Oriental Research, 2010.
- [15] M. D’Andrea. *The Southern Levant in Early Bronze IV: issues and perspectives in the pottery evidence*. Sapienza Univ. di Roma, Dipt. di Scienze dell’Antichità 20XX, 2015.
- [16] L. Maritan, C. Mazzoli, V. Michielin, D. Morandi Bonacossi, M. Luciani, and G. Molin. The provenance and production technology of Bronze Age and Iron Age pottery from Tell Mishrifeh/Qatna (Syria). *Archaeometry*, 47(4):723–744, 2005.
- [17] K. Badreshany, G. Philip, and M. Kennedy. The development of integrated regional economies in the Early Bronze Age Levant: new evidence from ‘Combed Ware’ jars. *Levant*, 42(1–2):160–196, 2020.
- [18] V. Tumolo and K. Badreshany. The ‘Combed Ware’ storage and transport vessels from Khirbet ez-Zeraqon: a reappraisal of the EB II–III evidence in light of recent studies. *Journal of Ancient Egyptian Interconnections*, 37:301–324, 2023.
- [19] G. Caspari and P. Crespo. Convolutional neural networks for archaeological site detection–finding “princely” tombs. *Journal of Archaeological Science*, 110, 2019.
- [20] A. Guyot, M. Lennon, T. Lorho, and L. Hubert-Moy. Combined Detection and Segmentation of Archeological Structures from LiDAR Data Using a Deep Learning Approach. *Journal of Computer Applications in Archaeology*, 4(1):1, 2021.
- [21] Ø. D. Trier, D. C. Cowley, and A. U. Waldeland. Using deep neural networks on airborne laser scanning data: Results from a case study of semi-automatic mapping of archaeological

- topography on Arran, Scotland. *Archaeological Prospection*, 26(2):165–175, 2019.
- [22] M. Troiano, E. Nobile, F. Mangini, M. Mastrogiovanni, C. Conati Barbaro, and F. Frezza. A comparative analysis of the bayesian regularization and levenberg-marquardt training algorithms in neural networks for small datasets: A metrics prediction of neolithic laminar artefacts. *Information*, 15(5), 2024.
- [23] E. Nobile, M. Troiano, F. Mangini, M. Mastrogiovanni, J. Verdi, F. Frezza, C. Conati Barbaro, and A. Gopher. Neural network analysis for predicting metrics of fragmented laminar artifacts: a case study from MPPNB sites in the Southern Levant. *Scientific Reports*, 14, 2024.
- [24] P. Siozos, N. Hausmann, M. Holst, and D. Anglos. Application of laser-induced breakdown spectroscopy and neural networks on archaeological human bones for the discrimination of distinct individuals. *Journal of Archaeological Science: Reports*, 35:102769, 2021.
- [25] J. Zhang, I. Fang, H. Wu, A. Kaushik, A. Rodriguez, H. Zhao, J. Zhang, Z. Zheng, R. Iovita, and C. Feng. Luwa dataset: Learning lithic use-wear analysis on microscopic images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22563–22573, 2024.
- [26] S. H. Bickler. Machine Learning Arrives in Archaeology. *Advances in Archaeological Practice*, 9(2):186–191, 2021.
- [27] L. M. Pawlowicz and C. E. Downum. Applications of deep learning to decorated ceramic typology and classification: A case study using Tusayan White Ware from Northeast Arizona. *Journal of Archaeological Science*, 130, 2021.
- [28] C. Hörr, E. Lindinger, and Brunnert G. Machine learning based typology development in archaeology. *Journal on Computing and Cultural Heritage*, 7(1):1–23, 2014.
- [29] P. Navarro, C. Cintas, M. Lucena, J. M. Fuertes, C. Delrieux, and M. Molinos. Learning feature representation of Iberian ceramics with automatic classification models. *Journal of Cultural Heritage*, 48:65–73, 2021.
- [30] Y. Qi, M. Z. Qiu, H. Z. Jing, Z. Q. Wang, C. L. Yu, J. F. Zhu, F. Wang, and T. Wang. End-to-end ancient ceramic classification toolkit based on deep learning: A case study of black glazed wares of Jian kilns (Song Dynasty, Fujian province). *Ceramics International*, 48(23):34516–34532, 2022.
- [31] G. Ruschioni, D. Malchiodi, A. M. Zanaboni, and L. Bonizzoni. Supervised learning algorithms as a tool for archaeology: Classification of ceramic samples described by chemical element concentrations. *Journal of Archaeological Science*, 49, 2023.
- [32] A. Anglisano, L. Casas, I. Queralt, and R. Di Febo. Supervised Machine Learning Algorithms to Predict Provenance of Archaeological Pottery Fragments. *Sustainability*, 14, 2022.
- [33] A. Ramil, A.J. López, and A. Yáñez. Application of artificial neural networks for the rapid classification of archaeological ceramics by means of laser induced breakdown spectroscopy (libs). *Applied Physics A*, pages 197–202, 2008.
- [34] A. Aprile, G. Castellano, and G. Eramo. Combining image analysis and modular neural networks for classification of mineral inclusions and pores in archaeological potsherds. *Journal of Archaeological Science*, 50:262–272, 2014.
- [35] M. Lyons. Ceramic Fabric Classification of Petrographic Thin Sections with Deep Learning. *Journal of computer applications in archaeology*, 4(1):188–201, 2021.
- [36] M. Lyons, F. Fecher, and M. Reindel. From LiDAR to deep learning: A case study of computer-assisted approaches to the archaeology of Guadalupe and northeast Honduras. *it - Information Technology*, 64(6):233–246, 2022.
- [37] Y. LeCun, Y. Bengio, and G. Hinton. Deep Learning. *Nature*, 521(7553):436–444, 2015.
- [38] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, Zhai X., T. Unterthiner, M. Dehghani, M Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv preprint arXiv:2010.11929*, 2020.
- [39] M. Raghu, T. Unterthiner, S. Kornblith, C. Zhang, and A. Dosovitskiy. Do Vision Transformers See Like Convolutional Neural Networks?. *Advances in neural information processing systems*,

34:12116–12128, 2021.

- [40] Y. LiHua, Z. WenBo, and Q. WangRen. Chronological classification of ming and qing dynasty ceramics images based on an enhanced resnet50 model. *STAR: Science & Technology of Archaeological Research*, 11(1):e2498260, 2025.
- [41] D. Castelvecchi. Can we open the black box of AI? *Nature*, 538(7623):20–23, 2016.
- [42] X. Zhong, B. Gallagher, S. Liu, B. Kailkhura, A. Hiszpanski, and T. Y. J. Han. Explainable machine learning in materials science. *Computational Materials*, 8, 2022.
- [43] A. B. Arrieta, N. Díaz-Rodríguez, J. Del Ser, A. Bennetot, S. Tabik, A. Barbado, S. Garcia, S. Gil Lopez, D. Molina, R. Benjamins, R. Chaila, and F. Herrera. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58:82–115, 2020.
- [44] F. Oviedo, J. L. Ferres, T. Buonassisi, and K. T. Butler. Interpretable and Explainable Machine Learning for Materials Science and Chemistry. *Accounts of Materials Research*, 3(6):597–607, 2022.
- [45] G. Montavon, W. Samek, and K. R. Müller. Methods for interpreting and understanding deep neural networks. *Digital Signal Processing*, 73:1–15, 2018.
- [46] D. Genovese, A. Sgroi, A. Devoto, S. Valentine, L. Wood, C. Sebastiani, S. Giagu, M. D’Onofrio, and S. Scardapane. Mixture-of-Experts Graph Transformers for Interpretable Particle Collision Detection., 2025.
- [47] H. El-Hajj, O. Eberle, A. Merklein, A. Siebold, N. Shlomi, J. Büttner, J. Martinetz, K. R. Müller, G. Montavon, and M. Valleriani. Explainability and transparency in the realm of digital humanities: toward a historian XAI. *International Journal of Digital Humanities*, 5(2):299–331, 2023.
- [48] L. Medeghini, M. Sala, C. De Vito, and S. Mignardi. A forgotten centre of ceramic production in Southern Levant: Preliminary analytical study of the Early Bronze Age pottery from Tell el-Far‘ah North (West Bank). *Ceramics International*, 45(9):11457–11467, 2019.
- [49] L. Medeghini, S. Mignardi, C. De Vito, N. Macro, M. D’Andrea, and S. Richard. New insights on Early Bronze Age IV pottery production and consumption in the southern Levant: The case of Khirbat Iskandar, Jordan. *Ceramics International*, 42(16):18991–19005, 2016.
- [50] M. Botticelli, S. Mignardi, C. De Vito, Y. W. Liao, D. Montanari, M. Shakarna, L. Nigro, and L. Medeghini. Variability in pottery production at Khalet al-Jam’ā necropolis, Bethlehem (West Bank): From the Early-Middle Bronze to the Iron Age. *Ceramics International*, 46(10):16405–16415, 2020.
- [51] M. Botticelli, S. Mignardi, C. De Vito, M. Sala, G. Mazzotta, L. Da Silva Gondim, and L. Medeghini. The ‘metallic ware’ from Tell el Far’ah North (West Bank): Petrography, technology, and provenance of a hidden ceramic industry. *Ceramics International*, 48(1):1366–1374, 2022.
- [52] M. Sala. The EB II ‘metallic ware’ from Tell el-Far‘ah North (West Bank): typology, technology and petrography of a ceramic industry of the central hill country. *Levant*, 55(2):146–171, 2023.
- [53] L. Medeghini, L. Fabrizi, S. De Vito, C. Mignardi, L. Nigro, and C. Gallo, E. Fiaccavento. The ceramic of the “Palace of the Copper Axes” (Khirbet al-Batrawy, Jordan): A palatial special production. *Ceramics International*, 42(5):5952–5962, 2016.
- [54] P. Ballirano, C. De Vito, L. Medeghini, S. Mignardi, V. Ferrini, P. Matthiae, D. Bersani, and P. Lottici. A combined use of optical microscopy, X-ray powder diffraction and micro-Raman spectroscopy for the characterization of ancient ceramic from Ebla (Syria). *Ceramics International*, 40(10):16409–16419, 2014.
- [55] L. Medeghini and L. Nigro. Khirbet al-Batrawy ceramics: a systematic mineralogical and petrographic study for investigating the material culture. *Periodico di Mineralogia*, 86(1), 2017.
- [56] K. He, X. Zhang, S. Ren, and J. Sun. Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778., 2016.
- [57] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical

- image database. *IEEE conference on computer vision and pattern recognition*, pages 248–255., 2009.
- [58] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization., 2014.
 - [59] S. Yang, W. Xiao, M. Zhang, S. Guo, J. Zhao, and F. Shen. Image Data Augmentation for Deep Learning: A Survey., 2023.
 - [60] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra. Grad-cam: Visual explanations from deep networks via gradient-based localization. *International Journal of Computer Vision*, 128(2):336–359, October 2019.
 - [61] R. Kashefi, L. Barekatain, M. Sabokrou, and F. Aghaeipoor. Explainability of Vision Transformers: A Comprehensive Review and New Perspectives. *arXiv preprint arXiv:2311.06786*, 2023.
 - [62] L. Gao, T. D. la Tour, H. Tillman, G. Goh, R. Troll, A. Radford, I. Sutskever, J. Leike, and J. Wu. Scaling and evaluating sparse autoencoders. *arXiv preprint arXiv:2406.04093*, 2024.