Classification of Engraved Pottery Sherds Mixing Deep-Learning Features by Compact Bilinear Pooling

Aladine Chetouani, Sylvie Treuillet, Matthieu Exbrayat, Sébastien Jesset

Version of Record: https://www.sciencedirect.com/science/article/pii/S0167865519303721
Manuscript_997d9749fccd9e3f93f7695d70243a47

Pattern Recognition Letters
journal homepage: www.elsevier.com

Classification of Engraved Pottery Sherds Mixing Deep-Learning Features by Compact Bilinear Pooling

Aladine Chetouani[a,**], Sylvie Treuillet[a], Matthieu Exbrayat[b], Sébastien Jesset[c]

[a]*Laboratoire PRISME, Université d'Orléans, 12 rue de Blois, 45067 ORLEANS Cedex 2, France*
[b]*LIFO, Université d'Orléans, Rue Léonard de Vinci B.P. 6759, 45067 ORLEANS Cedex 2, France*
[c]*Pôle d'Archéologie, 13bis rue de la Tour Neuve, 45000 Orléans*

ABSTRACT

The ARCADIA project aims at using pattern recognition and machine learning to promote a systematic analysis of the large corpus of archaeological pottery fragments excavated in Saran (France). Dating from the High Middle Ages, these sherds have been engraved with repeated patterns using a carved wooden wheel. The study of these engraved patterns allows archaeologists to better understand the diffusion of ceramic productions. In this paper, we present a method that classifies patterns of ceramic sherds by combining deep learning-based features extracted from some pre-trained Convolutional Neural Network (CNN) models. A dataset composed of 888 digital patterns extracted from 3D scans of pottery sherds was used to evaluate our approach. The classification capacity of each CNN model was first assessed individually. Then, several combinations of common pooling methods using different classifiers were tested. The best result was obtained when features of the VGG19 and ResNet50 models were combined using Compact Bilinear Pooling (CBP) with a high classification rate of 95.23%.

1. Introduction

Pottery fragments are among the most abundant archaeological materials and provide archaeologists with dating information to understand trade flows and social interactions. A common process in Europe during Late Antiquity and the High Middle Ages was for each potter to personalize his production by applying a hand-engraved wood wheel in the still fresh clay producing a repetitive geometrical decoration in relief. From the 1970s, medieval ceramology began to build typological references of these decorations. However, the analysis of these ceramic sherds is still mainly based on a manual stamping or visual recognition and remains very laborious and time consuming. In view of the increasing size of the corpus, there is a need to automate the heavy tasks of inventorying and classifying ceramic artifacts by the use of pattern recognition and artificial intelligence. The recognition of the relief patterns stamped by the potters needs to be robust to the variability due to the curvature of the vase, the mode of application, the state of conservation and so on. In this paper, we propose to mix deep learning-based features to classify the digital patterns extracted from 3D scans of engraved pottery sherds. We test some state-of-the-art Convolutional Neural Networks (CNNs) with a fine tuning on our dataset. Then, we compare different feature pooling strategies to show the relevance of combining CNN-based features using Compact Bilinear Pooling (CBP). The paper is organized as follows: related work is presented in Section 2. Section 3 describes the archaeological materials. Section 4 introduces the proposed pipeline and details strategies for deep learning, feature pooling and classification. Section 5 presents the experiments and the discussion. Section 6 concludes the paper.

2. Related Work

The new low-cost scanning technologies by laser scanner or photogrammetry promote the use of pattern recognition and artificial intelligence in particular to automate the heavy tasks of inventorying and classifying artifacts in the fields of cultural heritage (Gomes and al., 2014). For archaeological pottery fragments, studies have proposed a classification based either on the shape and profile (Willis et al., 2003; Kampel and

**Corresponding author: Tel.: +33-2-38-49-45-63;
e-mail: aladine.chetouani@univ-orleans.fr (Aladine Chetouani)

Sablatnig, 2007; Son et al., 2013; Zheng et al., 2014; Zvietcovich et al., 2016), on the color and material characterization (Kampel and Sablatnig, 2000; Stanco and Gueli, 2013; Farjas et al., 2012; Smith et al., 2010; Abadi et al., 2012; Makridis and Daras, 2012) or on their decorations (Guarnera et al., 2011; Zhou et al., 2017, 2018). Shape analysis is generally based on an ideal model of surface of revolution to extract an axis and a profile, then used to match the sherds, sometimes going as far as an automatic assembly of a pot from broken sherds (Willis et al., 2003). However, axis detection remains a complex problem for small sherds because they are weakly curved and not perfectly axially symmetric objects and resemble a patch of a sphere (Angelo and Stephano, 2018). Recently, CNN was applied for the content-based retrieval of complete or nearly complete three-dimensional vessel replicas (Benhabiles and Tabia, 2016) based on a shape vocabulary.

Few studies have attempted to classify pottery based on their decoration. The decorative style of Kamarès (2000 BC) studied in (Guarnera et al., 2011) presents refined polychrome decorations that use a wide of figurative themes. Classification of the pottery is based on the identification of some elementary decorative fragments in a corpus using 2D shape matching by contour flexibility. However, the Kamarès style is very different from the engraved pottery excavated in Saran: our High Middle Age sherds are rough and engraved with repetitive geometrical patterns. A second study (Zhou et al., 2017) proposed computer-vision algorithms to study paddle-stamped pottery sherds found on archaeological sites of the Native southeastern north American communities of the woodland period. The stylistic design of these carved wooden paddles (from 500 BC up to 19th century) is characterized by a curvilinear pattern with multiple parallel lines. The authors address the problem of automatically identifying the underlying carved wooden paddles by matching the curve pattern fragment impressed on a sherd to a complete decoration corpus. Recently, they proposed a CNN-based approach for curvilinear pattern segmentation from a 3D scan of the sherd and a two-stage matching algorithm by training a dual-source CNN to re-rank the candidate identified by traditional template matching (Zhou et al., 2018). Even if more research of pattern recognition in pottery decoration were available, the myriad forms of the decorations on ceramics mean that the resulting methods would not necessarily be the most appropriate for the domestic archaeological pottery excavated in Saran. In previous studies, we proposed a pipeline for the automatic classification of engraved sherds based on 3D scans. Binary patterns were automatically extracted from the 3D scans and the classification was done by training a SVM model with pyramid histograms of visual words (PHOW). The recognition rate was below 85% (Debroutelle et al., 2017). Then, we exploited well-known CNN models : AlexNet (Krizhevsky, 2014), VGG11 (Simonyan and Zisserman, 2014) and ResNet18 (He et al., 2015), with a fine tuning on our dataset. By comparing different options for the classification stage (i.e. replacing the fully connected layers of the network with another classifier), the best results were obtained with the combination ResNet18+SVM with a classification rate of about 88 % on the same dataset (Chetouani et al., 2018).

In continuation of these previous studies, the current work proposes to mix deep learning features by using hybrid network architectures with several pooling strategies. Based on the same set of sherds, greyscale relief maps are used instead of binary images, which leads to additional data. Consequently, deeper CNN networks (VGG16, VGG19, ResNet50) are learned to obtain a more detailed description. In doing so, the performances are greatly improved.

3. Archaeological Materials

Since 1994, several archaeological excavations have documented a major production site of domestic potteries in Saran (Loiret, France) dating from the High Middle Ages (sixth-eleventh Centuries). Many of the wheel-stamped pottery fragments show a repetitive decoration about 1 millimeter deep and 1.5 to 3 cm wide, depending on the dimensions of the wooden cylinders used by the potters (see Fig. 1). The most common patterns notched by potters on wooden cylinders included sticks, squares, chevrons and diamonds in one or several lines. For digital pattern extraction, we used a NextEngine scanner to provide a 3D model of each sherd. Next, we created a 2D shaded view of the surface with the Meshlab software by controlling the orientation of the 3D mesh (with a default illumination model). This manual operation is very easy to do just after the scanning step even by a non-expert. This procedure provides a snapshot of a grey-level relief map of the sherd (see Fig. 2). Efficient image processing was then developed to automatically detect the salient region focused on the relief motif. The salient region is here assumed to be the highest textured region. The FAST detector (Rosten and Drummond, 2006) is applied to the variance image of the grayscale map and then the DBSCAN grouping method (Ester et al., 1996) is used to detect the cluster with the highest textured point density. The convex envelope of the cluster corresponds to the salient region. An illustration of the salient region detected is given in Fig. 2. For more details, the reader is referred to (Debroutelle et al., 2017).

Fig. 1. Excavated pottery sherds from the Saran archaeological site (Loiret, France) that present relief decorations impressed by carved wooden cylinders.

Fig. 2. Digital pattern extraction.

The dataset is composed of 888 grey-level images resized to 224x224 with a black background. Table 1 presents the number

of samples for the four most representative classes and Fig. 4 shows a sample of each class.

Table 1. Composition of our dataset

Class	Type	Number of sherds
1	Diamonds	211
2	Sticks (2 lines)	259
3	Squares (3 lines)	274
4	Chevrons	144

4. Proposed Pipeline

The proposed pipeline is summarized in Fig. 3. After extracting the salient region as shown above, the proposed pipeline consists in three main steps: deep learning-based feature extraction, feature pooling and classification. These steps are described in the following subsections.

4.1. Feature Extraction

For feature extraction, we used CNNs similar to those in previous work (Chetouani et al., 2018). However, considering that the inputs used in this work are greyscale relief maps instead of binary images, we trained slightly deeper networks to take this additional information into account. In this study, four CNN models are compared and different combinations are evaluated.

- **AlexNet:** In 2012, AlexNet was one of the first networks to point out the efficiency of CNN for classification tasks (Krizhevsky, 2014). Its architecture consists of 5 convolution layers and 3 fully-connected layers. The authors highlight three main points: the popular activation function ReLU (Rectified Linear Unit), the dropout which consists in removing some connections between neurons of the fully connected layers to prevent over-fitting, and the overlap during the pooling step.

- **VGG:** Developed in 2014 by the Oxford Visual Geometry Group, VGG models use only 3x3 filters instead of the 11x11 and 5x5 filters employed in AlexNet's first two layers (Simonyan and Zisserman, 2014). Several versions were proposed with 11, 13, 16 and 19 layers. In this paper, **VGG16** and **VGG19** were employed.

- **ResNet:** The Residual Neural Network model was introduced in 2015 with the core idea that letting the stacked layers fit a residual mapping is easier than letting them directly fit the desired underlying mapping (He et al., 2015). The authors introduce a so-called "identity shortcut connection" so that the output F(x) of each Conv-ReLU-Conv series is replaced by F(x)+x. Different versions were developed with 18, 34, 50 and 152 layers. The version we used in this paper is **ResNet50**.

4.2. Feature Combination

It is well-known that classification performances can take advantage of mixing different sources of knowledge. Feature pooling is used to achieve invariance to image transformations, more compact representations, and better robustness to noise (Boureau et al., 2010). Many of the popular methods for extracting visual features previously used already included pooling, such as variants based on local histograms (SIFT and HOG), bag-of-visual words, Fisher encoding, VLAD or bilinear models.

In this study, we combined deep learning-based features provided by the selected CNN into a 'bag of features' used as input for the classifier by testing several pooling strategies. Concatenation (**Concat**) is the first intuitive pooling strategy producing a resulting vector of size equal to the length of both vectors.

Instead of just concatenating vectors, the pooling operation can be an element-wise operation between vectors, typically a sum, an average, a max, a product, etc. Element-Wise Multiplication (**EWM**) and Element-Wise Sum (**EWS**) are the most widely applied operations. Both operations have also been used together, denoted here as (**EWM+EWS**).

Recently, (Lin et al., 2015) showed that by using Bilinear Pooling (BP), significant improvements in visual recognition accuracy can be achieved. It consists in forming a global vector by applying the outer product of two vectors. BP is discriminative but the output representation is very high dimensional. Noticing that bilinear features are closely related to polynomial kernels, (Gao et al., 2015) proposed more compact representations that can be learned end-to-end by back propagation. The latter, so called Compact Bilinear Pooling (**CBP**), consists in projecting the extracted feature vectors into a lower dimensional space and then combining the projected features. The authors used the Count Sketch (CS) projection function (Pham and Pagh, 2013) because of a specific property that they demonstrated: the CS projection of the outer product of two vectors is the convolution of the CS projections of each individual vector (see Equ. 1). Then, the convolution can be easily rewritten into an element-wise product by applying Fourier transform (see Fig. 5). It is worth noting that CBP can be easily extended to Compact Multi-linear Pooling (**CMP**) to combine more than two vectors (Algashaam et al., 2017).

$$\Psi(F1 \otimes F2) = \Psi(F1) \circledast \Psi(F2) \tag{1}$$

where Ψ denotes the Count Sketch projection. $\otimes$ and $\circledast$ are the outer product and the convolution, respectively.

5. Experiments

In this section, we evaluate the performance of the proposed pipeline. We first compare the performance of each of the CNNs considered by unit tests. Next, the efficiency of CBP and CMP provided by different combinations of deep learning-based feature vectors is considered. The best configuration is finally used to compare different classification and pooling strategies.

5.1. Evaluation Protocol

All the tests were done by splitting the dataset into training-validation (66%) and test (33%) sets randomly without overlap. In order to ensure the representation of each class, the splitting was applied for each class. This procedure was repeated 20 times and the average classification rate was considered as the evaluation metric.

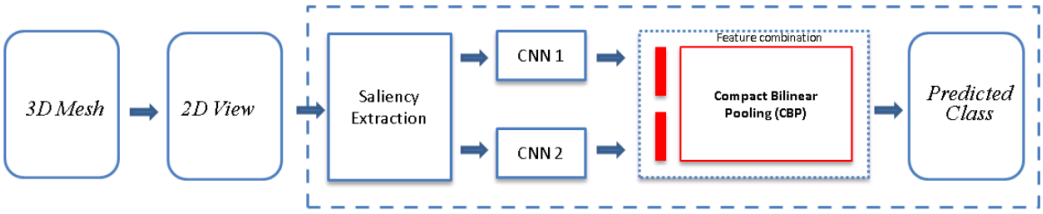

Fig. 3. Proposed pipeline

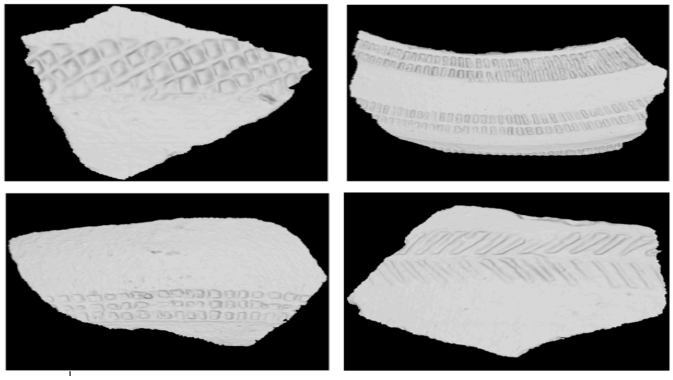

Fig. 4. Samples of the dataset: (top-left) diamonds; (top-right) sticks; (bottom-left) squares; (bottom-right) chevrons

5.2. *Unit Tests of Networks*

The four selected networks (AlexNet, VGG16, VGG19, ResNet50) were tested with fine tuning on our dataset. We modified the FC layers of each of them and the Softmax layer was adapted to the number of classes (four). The Stochastic Gradient Descent (SGD) was used during the training with a batch size and a learning rate set respectively to 32 and 0.01. The momentum and the number of epochs were fixed respectively to 0.9 and 100. At each epoch, we shuffled the training data and we stored the model. We finally kept the model that provided the best performance. To expand the training dataset, we applied flipping and translations (data augmentation). It is obvious that rotation cannot be used because of the confusion it would introduce between square and diamond patterns. To ensure homogeneity, the size of the output vector was fixed at 128 for all the models.

Table 2 shows the average classification rate obtained for each CNN model. AlexNet gives a rather low classification rate (less than 70%), while the other three give much higher performance (more than 90%). The two models VGG19 and ResNet50 have similar and stable results with a slight lead for ResNet50 (93.05% against 92.95% for VGG19).

Compared to our previous work (Chetouani et al., 2018), the performances are greatly improved by using greyscale relief maps instead of binary images with slightly deeper architectures. It clearly provides a better characterization of the pattern. The recognition rate increases by about 10% while the standard deviation drops: 81.5% for VGG11 in our previous work with binary images against 91.7% for VGG16 with greyscale relief maps herein.

Table 2. Average percentage of classification of each CNN model

	Average percentage of classification (%)	**Standard deviation**
CNN features		
AlexNet	69.62	2.53
VGG16	91.73	3.06
VGG19	92.95	1.41
ResNet50	**93.05**	**1.31**

5.3. *Evaluation of Deep Learning-based Feature Pooling*

We compared different combinations of the deep learning-based features provided by the selected CNNs into a 'bag of features'. All pairwise vector combinations provided by two CNNs were compared using Compact Bilinear Pooling (**CBP**), as well as triplet combinations by Compact Multilinear Pooling (**CMP**). We kept the same parameters for the pooling tests as those that had been used during unit tests.

Table 3 shows the classification rates obtained. Since AlexNet is the least efficient network, it tends to decrease the performances when it is associated with other features. The most significant improvements are achieved for the pair VGG19+ResNet50 (6) and the triplets including this pair (9 and 10). It corresponds to the combination of the models that provided the best performances for unit tests (see Table 2). We can also notice a significant drop in standard deviations showing more stable behavior using feature pooling.

To determine whether if the differences in the obtained results are significant, a one-way analysis of variance (ANOVA) was applied to the six pooling strategies tested: Concat, EWM, EWS, EWM+EWS, BP and CBP with D=100 (see Fig. 6). For each box, the middle mark corresponds to the average value and the contours of the box are the 25th and 75th per-centiles. The min and max interval is defined by excluding values greater than three standard deviations from the mean. The outliers are pointed by the red crosses. As can be seen, the distribution of CBP has a lower standard deviation than the others, without outliers. A pairwise significance test of the difference between the CBP and the other five pooling strategies is also reported in Table 4, giving the *p-values* and the *F-values*. A *p-value* less than 0.05 indicates that the difference is significant, while

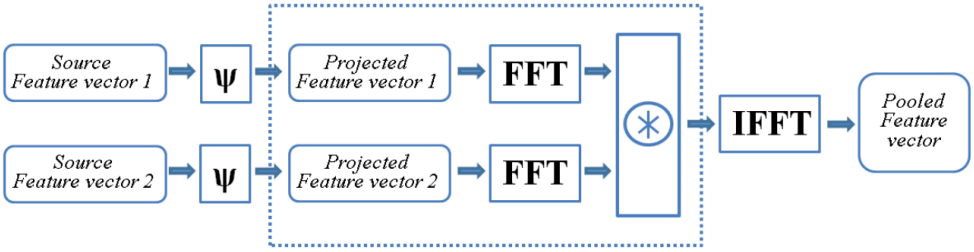

Fig. 5. Compact Bilinear Pooling strategy

Table 3. Evaluation of deep learning-based feature combinations using CBP and CMP

		Average percentage of classification (%)	Standard deviation
Combination of 2 vectors			
1	**CBP(AlexNet,VGG16)**	91.37	0.83
2	**CBP(AlexNet,VGG19)**	92.00	1.03
3	**CBP(AlexNet,ResNet50)**	90.98	1.38
4	**CBP(VGG16,VGG19)**	92.05	1.02
5	**CBP(VGG16,ResNet50)**	92.98	1.32
6	**CBP(VGG19,ResNet50)**	**95.23**	**0.60**
Combination of 3 vectors			
7	**CMP(AlexNet,VGG16,VGG19)**	89.82	0.96
8	**CMP(AlexNet,VGG16,ResNet50)**	89.43	1.28
9	**CMP(AlexNet,VGG19,ResNet50)**	93.87	0.71
10	**CMP(VGG16,VGG19,ResNet50)**	94.12	0.74

a *F-value* close to 1 indicates that the means and the standard deviations are similar. As can be seen, (EWM+EWS,CBP) obtained the highest p-value (0.2576) which means that the results of the two methods are not significantly different and its low F-value (1.32) means that their distributions are similar. However, as the number of parameters of *(EWM+EWS)* is 2.5 times (2.560K/1.000K) higher than the CBP and the accuracy of CBP is slightly higher (95.23), CBP with D=100 was retained as the best solution.

Table 4. Pairwise significance test of the difference between CBP and the other pooling strategies.

	p-value	F-value
(Concat, CBP)	0.0256	5.4
(EWM, CBP)	0.0259	5.37
(EWS, CBP)	0.0589	3.79
(EWM+EWS, CBP)	0.2576	1.32
(BP, CBP)	0.0002	16.78

Table 5 shows the confusion matrix for the best configuration (6). As expected, the diagonal contains high classification rates. A few confusions can be observed, in particular between class 2 (sticks) and class 3 (squares). The total confused cases between these two classes represent 3.86% and 5.88%. This is not surprising if we look at some samples, as shown in Fig. 7. Among

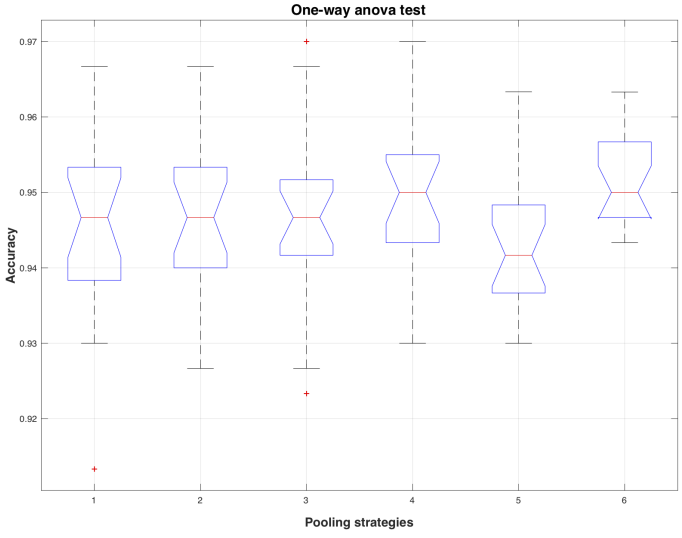

Fig. 6. Results of the one-way ANOVA for the pooling strategies: 1) Concat, 2) EWM, 3) EWS, 4) EWM+EWS, 5) BP and 6) CBP with D=100.

the geometric shapes that make up the pattern, it is sometimes difficult to distinguish between sticks and squares, probably because of the approximate shapes carved by the potter on the wooden wheel, the slippage when applied in fresh clay and the

Table 5. Confusion matrix (%) for CBP(VGG19,ResNet50)

		Predicted classes			
		1	**2**	**3**	**4**
True Classes	**1**	**97.75**	0.56	1.13	0.56
	2	0.83	**94.09**	**3.86**	1.21
	3	0.43	**5.88**	**93.69**	0
	4	0.97	1.81	0.14	**97.08**

degradation over time.

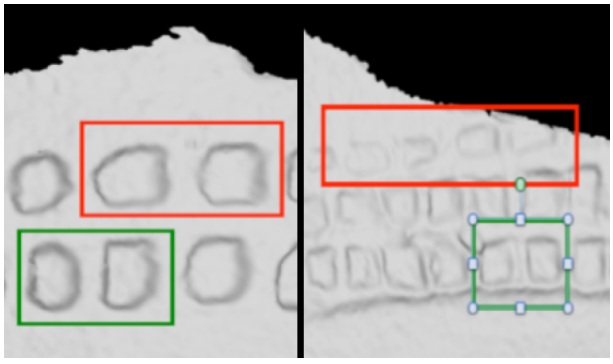

Fig. 7. Confusing samples: Red blocks represent the wrongly predicted forms (squares and sticks on the left and right images, respectively), while the green blocks highlight the expected right forms (sticks and squares on the left and right images, respectively).

Then, we evaluated the impact of the dimension (D) of the projection space when using the Count Sketch function (see Fig. 8). As can be seen, we achieved high classification rates even for small dimensions and the best performance was obtained for the D=100.

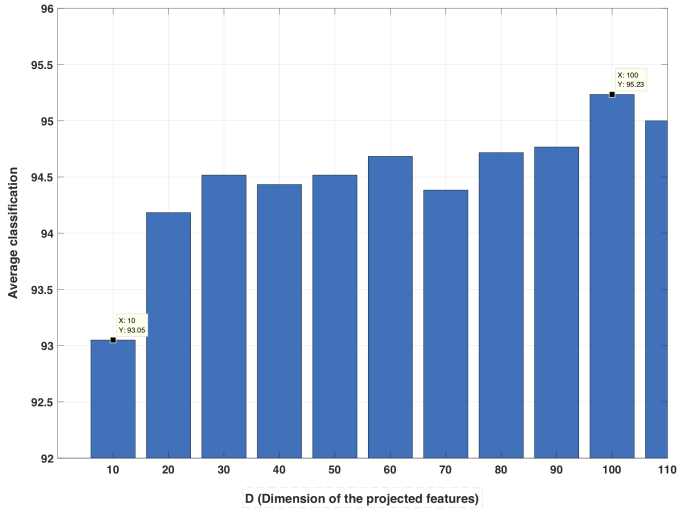

Fig. 8. Average classification rates according to the dimension of the projection space for CBP(VGG19,ResNet50).

Finally, we compared the classification performance with and without the extraction step of the salient region (see Table 6). A classification focused on the relief pattern improves the performance by 3.66% compared to the use of the 2D view of the whole sherd.

Table 6. Performance comparison with and without the extraction step of salient region using CBP(VGG19,ResNet50)+FC

	Average percentage of classification (%)	**Standard deviation**
Saliency extraction		
Without	91.57	1.38
With	**95.23**	**0.60**

5.4. Comparison of Classifiers and Pooling Strategies

The 'bag of features' given by pooling the features provided by several CNNs is here used as input for 2 Fully Connected (FC) layers of size 50, followed by a Softmax layer of size 4 (i.e. 4 classes). Each FC layer is followed by a (ReLU + dropout) layer. We compared different options for the classification step by replacing the previous FC layers with a Support Vector Machine (**SVM**) or a Linear Discriminant Analysis (**LDA**). For SVM, various kernel functions were tested and we kept the *Gaussian* function as the best result. For LDA, different distances were tested and we kept the *Diaglinear* function that integrates an estimation of the covariance matrix as the best performance. Fig. 9 summarizes all these comparisons. We employed here the pair of deep learning-based feature vectors that provided the best performance (VGG19 and ResNet50).

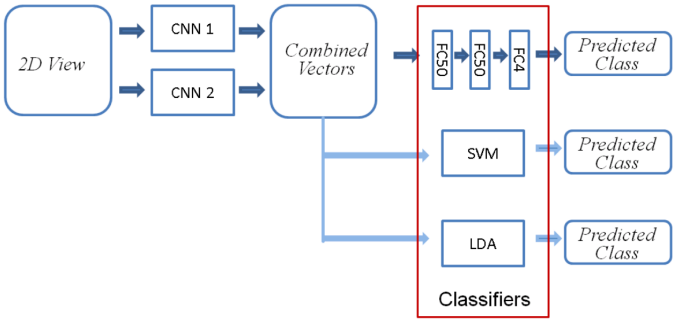

Fig. 9. Framework of the Classifier comparison

Table 7 shows the results obtained for different pooling strategies and different classifiers. We compared three classifiers (SVM, LDA, FC). The 'bag of features' used as input for the classifiers was obtained by Concatenation (Concat), Element-Wise Multiplication (EWM), Element-Wise Sum (EWS), the two operations together (EWM+EWS), Bilinear Pooling (BP) and finally by the Compact Bilinear Pooling (CBP). For the latter, we show the performance obtained for 2 dimensions (D=50 and D=100). All pooling strategies improve the performance with an increase up to 2% compared to unit tests of each CNN (see Table 2). The best configuration is CBP(VGG19, ResNet50) + FC, with 95.23% closely followed by the configuration Concat+SVM and CBP+LDA. However, CBP implies a much more compact representation since the number of learnable parameters for CBP (100*50*50*4) is about 40% less than that of Concat ((128+128)*50*50*4).

Table 7. Comparison of different classifiers and pooling strategies with the best deep learning-based features given by VGG19 and ResNet50.

	Average percentage of classification (%) with ± std	Feature size/ Nbr. of param.
SVM		
Concat	94.75 ± 0.58	256/–
EWM	94.20 ± 1.13	128/–
EWS	94.28 ± 0.99	128/–
EWM+EWS	94.72 ± 0.80	256/–
CBP	94.70 ± 0.74	100/–
LDA		
Concat	94.10 ± 0.71	256/–
EWM	94.53 ± 1.27	128/–
EWS	94.47 ± 1.29	128/–
EWM+EWS	94.15 ± 1.60	256/–
CBP	94.65 ± 0.71	100/–
FC (Used classifier)		
Concat	94.53 ± 1.21	256/2,560K
EWM	94.65 ± 0.95	128/1,280K
EWS	94.67 ± 1.16	128/1,280K
EWM+EWS	94.95 ± 0.93	256/2,560K
BP	94.30 ± 1.14	16384/163,840K
CBP with D=50	94.52 ± 0.93	50/500K
CBP with D=100	**95.23 ± 0.60**	**100/1,000K**

6. Conclusion

In this paper, we proposed a method to classify patterns of ceramic sherds by combining feature vectors provided by CNN models. Four pre-trained CNN models were employed with fine-tuning. The deep learning-based features thus extracted were then combined through different pooling strategies (Concat, EWM, EWS, EWM+EWS, CB and CBP) and the performance of different classifiers (SVM, LDA and FC) was compared. The best performance was obtained with CBP(VGG19,ResNet50)+FC (95.23%). In future work, we will first increase the dataset by integrating other patterns from several sites. We will consider geometric aspects (lines, shapes, etc.) to better discriminate between patterns. To go further, we will also try to classify wheels for each pattern class which is still a challenging task.

Acknowledgments

This work was part of the ARCADIA project funded by the Region Centre-Val de Loire. Thanks to Teddy Debroutelle for his contribution for digital pattern extraction from the pottery fragments corpus.

References

Abadi, M., Khoudeir, M., Marchand, S., 2012. Gabor filter-based texture features to archaeological ceramic materials characterization, in: Image and Signal Processing, pp. 33–342.

Algashaam, F.M., Nguyen, K., Alkanhal, M., Chandran, V., Boles, W., Banks, J., 2017. Multispectral periocular classification with multimodal compact multi-linear pooling. IEEE Access 5, 14572–14578.

Angelo, L.D., Stephano, P.D., 2018. Axis estimation of thin-walled axially symmetric solids. Pattern Recognition Letters 106(4), 47–52.

Benhabiles, H., Tabia, H., 2016. Convolutional neural network for pottery. J. of Electronic Imaging 26(1).

Boureau, Y.L., Ponce, J., Lecun, Y., 2010. A theoretical analysis of feature pooling in visual recognition, in: 27th Int. Conf. on Machine Learning.

Chetouani, A., Debroutelle, T., Treuillet, S., Exbrayat, M., Jesset, S., 2018. Classification of ceramic sherds based on convolutional neural network, in: ICIP 2018, pp. 1038–1042.

Debroutelle, T., Treuillet, S., Chetouani, A., Exbrayat, M., Martin, L., Jesset, S., 2017. Automatic classification of ceramic sherds with relief motifs. Journal of Electronic Imaging 26, 1 – 14 – 14.

Ester, M., Kriegel, H.P., Sander, J., Xu, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. International Conference on Knowledge Discovery and Data mining , 226–231.

Farjas, M., Rejas, Juan, G., Mostaza, T., Zancajo, J., 2012. Deepening in the 3d modelling: Multisource analysis of a polychrome ceramic vessel through the integration of thermal and hyperspectral information, in: Conf. on Computer Applications and Quantitative Methods in Archaeology (CAA), pp. 116–124.

Gao, Y., Beijbom, O., Zhang, N., Darrell, T., 2015. Compact bilinear pooling. CoRR abs/1511.06062. URL: `http://arxiv.org/abs/1511.06062`.

Gomes, L., al., 2014. 3d reconstruction methods for digital preservation of cultural heritage: A survey. Pattern Recognition Letters 50(12), 3–14.

Guarnera, G.C., Stanco, F., Tanasi, D., Gallo, G., 2011. Classification of decorative patterns in kamares pottery, in: Spring Conference on Computer Graphics, p. 20/23.

He, K., Zhang, X., Ren, S., Sun, J., 2015. Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385 .

Kampel, M., Sablatnig, R., 2000. Color classification of archaeological fragments, in: ICPR 2000, pp. 771–774.

Kampel, M., Sablatnig, R., 2007. Rule based system for archaeological pottery classification. Pattern Recognition Letters 28(6), 740–747.

Krizhevsky, A., 2014. One weird trick for parallelizing convolutional neural networks. CoRR abs/1404.5997.

Lin, T.Y., RoyChowdhury, A., Maji, S., 2015. Bilinear cnn models for fine-grained visual recognition, in: Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), pp. 1449–1457.

Makridis, M., Daras, P., 2012. Automatic classification of archaeological pottery sherds. Journal on Computing and Cultural Heritage 5(4), 1–21.

Pham, N., Pagh, R., 2013. Fast and scalable polynomial kernels via explicit feature maps, in: Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM. pp. 239–247.

Rosten, E., Drummond, T., 2006. Machine learning for high-speed corner detection. European Conference on Computer Vision , 430–44.

Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. CoRR, abs/1409.1556 .

Smith, P., Bespalov, D., Shokoufandeh, A., Jeppson, P., 2010. Classification of archaeological ceramic fragments using texture and color descriptors, in: CVPRW 2010, pp. 49–54.

Son, K., Almeida, E., Cooper, D., 2013. Axially symmetric 3d pots configuration system using axis of symmetry and break curve, in: CVPR 2013, pp. 257–264.

Stanco, F., Gueli, A.M., 2013. Computer graphics solutions for pottery colors specification, in: Digital Photography IX, pp. 97–101.

Willis, A., Cooper, D., Oriols, X., 2003. Accurately estimating sherd 3d surface geometry with application to pot reconstruction, in: CVPRW 2003, p. 7.

Zheng, S.Y., Huang, R.Y., LI, J., Wang, Z., 2014. Reassembling 3d thin fragments of unknown geometry in cultural heritage. ISPRS II-5, 393–399.

Zhou, J., Lu, Y., Zheng, K., Smith, K., Wilder, C., Wang, S., 2018. Design identification of curve patterns on cultural heritage objects: Combining template matching and cnn-based re-ranking. CoRR abs/1805.06862.

Zhou, J., Yu, H., Smith, K., Wilder, C., Hongkai Yu, S.W., 2017. Identifying designs from incomplete, fragmented cultural heritage objects by curved-pattern matching. J. of Electronic Imaging 26(1).

Zvietcovich, F., Navarro, L., Saldana, J., Castillo, L.J., Castaneda, B., 2016. A novel method for estimating the complete 3d shape of pottery with axial symmetry from single potsherds based on principal component analysis. Digital Applications in Archaeology and Cultural Heritage 3(2), 42–54.