

Classification de tessons céramiques gravés par réseaux de neurones convolutifs

Teddy DEBROUTELLE¹, Aladine CHETOUANI¹, Matthieu EXBRAYAT², Sylvie TREUILLET¹, Sebastien JESSET³

¹Laboratoire PRISME, Université d'Orléans, 12 rue de Blois, F-45067 Orléans cedex 2

²LIFO, Université d'Orléans, Bâtiment IIIA, rue Léonard de Vinci, B.P. 6759, F-45067 Orléans cedex 2

³Service Archéologique Municipal d'Orléans, 13 bis rue de la Tour Neuve, 45000 Orléans

{teddy.debrounette, aladine.chetouani, matthieu.exbrayat, sylvie.treuillet}@univ-orleans.fr,
sebastien.jesset@orleans-agglo.fr

Résumé – Les campagnes de fouilles menées à Saran (Loiret) sur des ateliers médiévaux de poterie ont permis de constituer une collection de tessons céramiques gravés de motifs géométriques répétitifs réalisés à l'aide de molettes en bois. L'étude de ces décors permet aux archéologues de mieux comprendre la production et la diffusion de ces céramiques. Le projet ARCADIA a pour objectif d'automatiser la classification des motifs. Pour ce faire, nous proposons dans cet article d'exploiter les réseaux de neurones convolutifs (CNN). Plusieurs modèles ont été testés et comparés, en utilisant pour la classification finale la partie entièrement connectée des modèles ou en la remplaçant par des classifieurs plus classiques (SVM, XGBoost). Un ensemble de données composé de 888 décors binaires a été utilisé dans cette étude.

Abstract – A wide medieval pottery workshop has been explored over the last decades in Saran (Loiret, France), from which a vast collection on ceramic sherds has been established. These sherds have been engraved with repeated patterns using a carved wooden wheel. Their study allows Archeologists to better understand the dissemination of the pottery produced and the commercial exchanges at this time. The ARCADIA project aims at carrying out an automated classification of the sherds. For this purpose, we propose in this paper to rely on Convolutional Neuro Networks (CNN), several models of which having been configured and tested. We also evaluate the impact of replacing the last fully connected layer of the CNN by classical classifiers (SVM, XGBoost). Our study is based on a collection of 888 shards.

1 Introduction

Depuis 1994, plusieurs campagnes de fouilles menées à Saran (Loiret, France) ont révélé un atelier de production de poteries céramiques datant du Haut Moyen-Age (VI au XI^{ème} siècle). La plupart des 38000 tessons récoltés à ce jour présentent un motif géométrique appliqué par le potier dans la terre fraîche, à l'aide d'une molette en bois sculptée à la main. L'unicité de ces outils rudimentaires (le décor est spécifique à chaque molette) et leur usure relativement rapide font des motifs un indicateur déterminant pour analyser les échanges commerciaux de l'époque. Pour recueillir les décors, l'archéologue a dans un premier temps utilisé une méthode manuelle reposant sur un moulage en pâte à modeler, lequel est ensuite encré, imprimé et éventuellement numérisé. Si cette méthode permet un recueil d'assez bonne qualité, elle s'avère en revanche chronophage et permet au mieux de traiter une dizaine de tessons par heure. Elle nécessite également l'oeil exercé de l'archéologue pour déterminer le type de motif observé, et plus encore pour identifier, lorsque cela est possible, la molette utilisée.

Le projet ARCADIA (Automatic Recognition of Ceramics Achieved by Digital Image Analysis), financé par la région Centre-Val de Loire, propose de faciliter l'interprétation de ce patrimoine archéologique par le développement d'une chaîne de traitement automatique de classification des tessons. Dans

cette article, nous étudions les réseaux de neurones convolutifs (CNN) comme réponse possible aux problèmes d'extraction d'attributs visuels pertinents et de classification des motifs, ceci sur la base d'une numérisation réalisée préalablement. Les CNN présentent l'avantage, une fois entraînés, de pouvoir être réutilisés sur de nouveaux jeux de données par apprentissage de transfert. Toutefois, le reparamétrage des couches finales dédiées à la classification est délicat, eut égard à la taille de notre base d'images (888 numérisations). Des classifieurs classiques (SVM, XGBoost) ont donc été testés en remplacement de la partie classifieur des différents CNN évalués.

L'article est organisé comme suit : La section 2 décrit la base de données. La section 3 présente un bref état de l'art. La section 4 présente les modèles CNN et les classifieurs utilisés. La section 5 est dédiée aux expérimentations. La section 6 conclut cette étude et donne quelques perspectives.

2 Base de données

Quelle que soit la taille du tesson, le décor est gravé sur environ 1 millimètre de profondeur et 1,5 à 3 cm de largeur (selon les dimensions des cylindres en bois utilisés par les potiers). Quatre types principaux de décors ont été répertoriés par l'archéologue. Une campagne de numérisation des tessons par

scanner 3D a permis d’obtenir une base de données de 888 tessons numérisés et labellisés par notre expert. Le nuage de points de la face numérisée de chaque tesson est ensuite projeté en 2D pour obtenir une carte de profondeurs. La région contenant le décor est détectée puis binarisée (pour plus de détails sur la chaîne de prétraitements, voir [1]). Ces images binaires seront ici utilisées pour caractériser le décor. La figure 1 illustre les différentes classes de décor et le tableau 1 présente leur répartition dans la base numérisée.

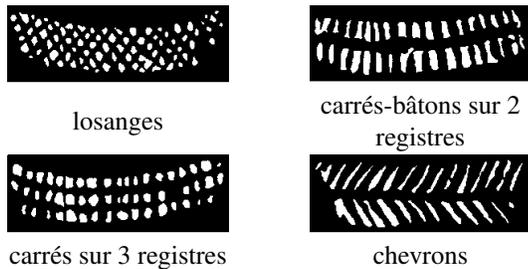


FIGURE 1 – Décors présents dans la base de données.

3 État de l’art

Plusieurs travaux publiés dans la littérature visent à automatiser la classification de tessons céramiques, soit par leur forme ou morphologie [2, 3], soit par leur composition [4, 5] ou encore par la présence de décors [6]. L’analyse de forme s’appuie généralement sur un modèle idéal permettant d’extraire un axe et un profil, utilisés ensuite pour appairer les tessons et allant parfois jusqu’à proposer un assemblage automatique [2]. Cependant, la détermination d’un couple axe/profil reste très délicate sur des tessons présentant peu de courbure ou des asymétries. La caractérisation des matériaux sur les surfaces intérieure, extérieure ou tranche des céramiques a également fait l’objet d’études par analyse de couleur [4], de texture [7] ou en associant l’infrarouge et l’hyperspectrale [8]. Le classement automatique des décors est aussi abordé par la texture [9] ou l’association couleur-texture [10, 5]. Cependant, ces études restent critiques sur l’échantillonnage mis en jeu et peu concluantes sur la faisabilité, car elles mélangent parfois les engobes décoratifs et les caractéristiques de la pâte elle-même. Peu d’études se concentrent sur la détection des décors par nu-

TABLE 1 – Composition de la base de données.

Classe	Type	Nombre de tessons
1	losanges	211
2	bâtons-carrés sur 2 registres	259
3	carrés sur 3 registres	274
4	chevrons	144

mérisation 3D. On peut citer le travail présenté dans [6] où la reconnaissance de motifs complexes est réalisée en appliquant des déformations à partir d’une bibliothèque de formes simples. Une étude récente s’intéresse également à l’identification de décors curvilignes gravés sur des céramiques à l’aide de tampons [11]. Les auteurs cherchent à identifier les tampons à partir de motifs partiels par des techniques d’appariement de courbes (curve matching). De manière générale, le caractère protéiforme des décors sur les céramiques fait que, même si l’analyse des motifs avait été plus largement traitée, les méthodes qui en découleraient ne seraient pas forcément les plus adaptées aux céramiques de Saran.

4 Méthode proposée

Dans cet article, nous proposons d’utiliser les modèles CNN pour la classification des tessons céramiques. Ces modèles sont généralement formés de deux parties : la partie de *convolution* qui permet de caractériser les objets en apprenant des caractéristiques à partir de l’image et la partie de *classification* qui utilise les caractéristiques apprises pour identifier l’objet. Nous proposons de comparer les performances de 3 modèles CNN et de modifier leur partie de classification.

4.1 Modèles CNN utilisés

Trois modèles ont été utilisés :

- **AlexNet** : Développé en 2012 et composé de 5 couches convolutives et de 3 couches complètement connectées (FC), AlexNet [12] est l’un des premiers modèles ayant permis d’apprécier la pertinence des modèles CNN pour les tâches de classification. Les auteurs soulignent deux points principaux : l’utilisation de la fonction d’activation Relu et l’exploitation du *Dropout* pour éviter le sur-apprentissage au cours de l’étape de classification.
- **VGG** : Proposé en 2014, les modèles VGG ont été développés par le groupe Oxford Visual Geometry Group [13]. Pour augmenter la capacité du CNN à discriminer les objets, les auteurs ont intégré plus de fonctions d’activation Relu en utilisant des couches convolutives dotées de filtres 3x3 au lieu de filtres 7x7. Cela permet de diminuer le nombre de paramètres et d’augmenter le nombre d’unités d’activation. Plusieurs versions ont été proposées avec 11 (VGG11), 13 (VGG13), 16 (VGG16) et 19 (VGG19) couches. Le modèle VGG11 a été ici utilisé.
- **ResNet** : Le réseau de neurones ResNet [14] proposé en 2015 se distingue par l’intégration d’un module résiduel. L’idée développée par les auteurs est de reformuler la sortie ($H(x) = F(x)$) de chaque série de *Conv-Relu-Conv* en ajoutant l’entrée x comme information ($H(x) = F(x) + x$). Différentes versions ont également été développées : ResNet18 (18 couches), ResNet34 (34 couches), ResNet152 (152 couches). Dans cette étude, le ResNet18 a été utilisé.

Les modèles utilisés dans cette étude ont été entraînés sur la vaste base ImageNet [15]. Les poids et biais de ces modèles ont été ajustés par *fine-tuning* avec les paramètres suivants :

- **Taux d'apprentissage** : 0,01
- **Momentum** : 0,9
- **Méthode d'optimisation** : SGD
- **Fonction de perte** : Softmax
- **Bibliothèque** : PyTorch [16].

4.2 Classifieurs testés

L'objectif ici est d'évaluer les performances d'un modèle CNN lorsque sa partie classifieur est modifiée. Pour cela, trois configurations ont été testées (voir section 5.3) :

- **Ajout d'une couche FC (plus profond)** : Nous proposons d'abord d'ajouter une couche FC. Notons que cela augmente fortement le nombre de paramètres à estimer.
- **Extrem Gradient Boosting (XGBoost)** [17] : les méthodes basées sur les arbres de décisions sont couramment utilisées dans plusieurs domaines avec une efficacité reconnue.
- **Support Vector Machine (SVM)** [18] : les SVM ont été largement utilisés dans la littérature. Ce type de méthode a été initialement développé pour les problèmes de classification binaire. L'idée est de projeter le vecteur de caractéristiques dans un espace de plus grande dimension pour trouver un hyperplan qui maximise la marge entre les classes en rendant le problème linéairement séparable. Pour l'étendre à notre cadre multi-classes, nous utilisons l'approche un-contre-un.

5 Résultats expérimentaux

5.1 Protocole d'évaluation

La base d'images étiquetées a été scindée aléatoirement en un ensemble d'entraînement-validation (66%) et un ensemble de test (34%) sans chevauchement. Cette procédure a été répétée 20 fois et le pourcentage moyen de bonne classification a été utilisé pour quantifier les résultats obtenus.

5.2 Comparaison des modèles CNN

Le tableau 2 montre le taux moyen de bonne classification de chaque modèle CNN. Les performances obtenues diffèrent fortement d'un modèle à l'autre. En effet, AlexNet a obtenu un résultat faible (77,24%), tandis que le ResNet18 a obtenu la meilleure performance (83,32%). Cependant, l'écart-type reste similaire pour les 3 modèles.

5.3 Impact de la partie classifieur

Afin d'évaluer la partie classification, le modèle ResNet18 a été utilisé selon les trois configurations suivantes :

- **ResNet18 + FC170** : la taille de la couche FC ajoutée est égale à 170, ce qui représente un bon compromis pour limiter

TABLE 2 – Pourcentage moyen de bonne classification des approches testées.

Modèle CNN	Pourcentage moyen de bonne classification (%)	Écart-type
AlexNet	77.24	2.73
VGG11	81.50	2.53
ResNet18	83.32	2.73

le nombre de paramètres supplémentaires : jusqu'à 87 000 paramètres supplémentaires : $(512 + 1) * 170$.

- **ResNet18 + XGBoost** : Softmax a été utilisée comme fonction de perte et la profondeur maximale de l'arbre a été fixée à 6.
- **ResNet18 + SVM** : Nous avons testé plusieurs noyaux (RBF, Linéaire, Interpolation etc). Le meilleur résultat a été obtenu en utilisant le noyau χ^2 .

Le tableau 3 montre le pourcentage moyen de bonne classification obtenu pour les trois configurations. Comme on peut le constater, les performances ont augmenté pour chacune des configurations proposées. La couche FC ajoutée (ResNet18 + FC170) permet une amélioration de 1%, peu significative en regard de l'écart-type. Le meilleur résultat a été obtenu en utilisant le classifieur SVM (87,94%). XGBoost a obtenu des performances proches du SVM. Le grand nombre de paramètres à estimer dans une couche FC peut expliquer le score de la configuration (ResNet18 + FC170).

TABLE 3 – Résultats obtenus pour les différentes configurations.

Config.	Modèle CNN	Pourcentage moyen de bonne classification (%)	Écart-type
1	ResNet18 + FC170	84.53	2.12
2	ResNet18 + XGBoost	87.29	1.26
3	ResNet18 + SVM	87.94	1.48

Le tableau 4 montre la matrice de confusion obtenue pour la configuration (ResNet18 + SVM). Les meilleurs taux de classifications ont été obtenus pour les classes 1 (losanges) et 4 (chevrons). Les plus grandes confusions sont entre les classes 2 (bâtons-carrés sur 2 registres) et 3 (carrés sur 3 registres). Ces différentes confusions s'expliquent par le fait que certains décors se trouvent sur l'extrémité du tessou et peuvent ainsi être référencés comme ayant 3 registres alors que le troisième registre n'est pas complet.

TABLE 4 – Matrice de confusion pour la configuration (ResNet18 + SVM).

Classes Réelles	Classes Prédites				
	%	1	2	3	4
1		95.50	0.82	1.00	2.68
2		2.01	79.81	15.58	2.60
3		3.33	9.88	86.33	0.46
4		3.97	1.09	0.20	94.74

6 Conclusion et perspectives

Dans cette étude, nous avons proposé d’exploiter les modèles CNN pour classer des décors gravés sur des tessons de céramique. Nous avons d’abord comparé les performances de 3 modèles pré-entraînés. Ensuite, différents classifieurs ont été utilisés et comparés. Les résultats obtenus montrent qu’il peut être pertinent de remplacer la partie classifieur des CNNs par des classifieurs classiques exigeant beaucoup moins de paramètres à estimer, tout en conservant les performances des CNNs. Le meilleur résultat a été obtenu pour la configuration (ResNet18 + SVM). Comme perspectives, nous envisageons d’augmenter considérablement la base de données pour permettre un meilleur apprentissage des différents réseaux.

Remerciement

Le projet ARCADIA (Automatic Recognition of Ceramics Achieved by Digital Image Analysis) est financé par la région Centre-Val de Loire.

Références

[1] T. Deboutelle, S. Treuillet, A. Chetouani, M. Exbrayat, L. Martin, and S. Jesset. Automatic classification of ceramic sherds with relief motifs. *Journal of Electronic Imaging*, 26(2) :023010, 03 2017.

[2] A. Willis, D. Cooper, and X. Orriols. Accuately estimating sherd 3d surface geometry with application to pot reconstruction. *Conference on Computer Vision and Pattern Recognition Workshop*, 2003.

[3] S.Y. Zheng, R.Y. Huang, J. Li, and Z. Wang. Reassembling 3d thin fragments of unknown geometry in cultural heritage. *SPRS Anals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 5(2) :393–399, 2014.

[4] M. Kampel and R. Sablatnig. Color classification of archaeological fragments. *International Conference on Pattern Recognition*, pages 771–774, 2000.

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

[6] G. C. Guarnera, F. Stanco, D. Tanasi, and G. Gallo. Classification of decorative patterns in kamares pottery. *Spring Conference on Computer Graphics*, pages 20–23, 2011.

[7] F. Stanco, A. M. Gueli, D. Tanasi, and G. Stella. Computer graphics solutions for pottery colors specification. *European Conference on Colour in Graphics, Imaging, and Vision*, pages 97–101, 2013.

[8] M. Farjas, J. G. Rejas, T. Mostaza, and J. Zancajo. Deepening in the 3d modelling : Multisource analysis of a polychrome ceramic vessel through the integration of thermal and hyperspectral information. *International Conference on Computer Applications and Quantitative Methods in Archaeology*, pages 116–124, 2012.

[9] M. Abadi, M. Khoudeir, and S. Marchand. Gabor filter-based texture features to archaeological ceramic materials characterization. *Image and Signal Processing, Lecture Notes in Computer Science*, 7340, pages 333–342, 2012.

[10] P. Smith, D. Bespalov, A. Shokoufandeh, and P. Jeppson. Classification of archaeological ceramic fragments using texture and color descriptors. *Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pages 49–54, 2010.

[11] Jun Zhou, Haozhou Yu, Karen Smith, Colin Wilder, Hongkai Yu, and Song Wang. Identifying designs from incomplete, fragmented cultural heritage objects by curve-pattern matching. *Journal of Electronic Imaging*, 26, 2017.

[12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 25 :1090–1098, 2012.

[13] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.

[14] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, pages 770–778, 06 2016.

[15] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision (IJCV)*, 115 :211–252, 2015.

[16] <https://github.com/pytorch/pytorch>.

[17] Tianqi Chen and Carlos Guestrin. Xgboost : A scalable tree boosting system. *The 22nd ACM SIGKDD International Conference*, pages 785–794, 08 2016.

[18] Chih-Chung Chang and Chih-Jen Lin. Libsvm : a library for support vector machine. *ACM Transactions on Intelligent Systems and Technology*, 2 :1–27, 2011.