DOI: 10.1111/nyas.15067

ORIGINAL ARTICLE

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES

African bovid tribe classification using transfer learning and computer vision

Manuel Domínguez-Rodrigo[1,2,3] | Juliet Brophy[4,5] | Gregory J. Mathews[6] | Marcos Pizarro-Monzo[1,7] | Enrique Baquedano[1]

[1]Institute of Evolution in Africa (IDEA), University of Alcalá de Henares, Madrid, Spain

[2]Area of Prehistory, Department of History and Philosophy, University of Alcalá de Henares, Alcalá de Henares, Spain

[3]Department of Anthropology, Rice University, Houston, Texas, USA

[4]Department of Geography and Anthropology, Louisiana State University, Baton Rouge, Louisiana, USA

[5]Centre for the Exploration of the Deep Human Journey, University of the Witwatersrand, Johannesburg, South Africa

[6]Department of Mathematics and Statistics, Loyola University, University of Chicago, Chicago, Illinois, USA

[7]Institut Català de Paleoecologia Humana i Evolució Social (IPHES), Universitat Rovira i Virgili, Tarragona, Spain

Correspondence
Manuel Domínguez-Rodrigo, Institute of Evolution in Africa (IDEA), University of Alcalá de Henares, Madrid, Spain. Email: m.dominguez.rodrigo@gmail.com

Funding information
Ministry of Science and Innovation, Grant/Award Number: PID2020-115452GB-C21

Abstract

Objective analytical identification methods are still a minority in the praxis of paleobiological sciences. Subjective interpretation of fossils and their modifications remains a nonreplicable expert endeavor. Identification of African bovids is a crucial element in the reconstruction of paleo-landscapes, ungulate paleoecology, and, eventually, hominin adaptation and ecosystemic reconstruction. Recent analytical efforts drawing on Fourier functional analysis and discrimination methods applied to occlusal surfaces of teeth have provided a highly accurate framework to correctly classify African bovid tribes and taxa. Artificial intelligence tools, like computer vision, have also shown their potential to be objectively more accurate in the identification of taphonomic agency than human experts. For this reason, here we implement some of the most successful computer vision methods, using transfer learning and ensemble analysis, to classify bidimensional images of African bovid teeth and show that 92% of the large testing set of images of African bovid tribes analyzed could be correctly classified. This brings an objective tool to paleoecological interpretation, where bovid identification and paleoecological interpretation can be more confidently carried out.

KEYWORDS
African bovids, artificial intelligence, computer vision, ecology, palaeoecology

INTRODUCTION

African bovids make up one of the most morphologically varied and ecologically diverse ungulate groups.[1–3] Most of the paleoecological reconstructions of recent African Neogene and Quaternary paleontological and archaeological localities rest on bovid macrommamal faunas.[4–22] Accurate identification of the bovid tribes to which fossils belong is, thus, essential for paleoecological reconstructions. Although several postcranial features can differentiate some bovid tribes, it is in the dentition that the most reliable identification of tribe and genus can be made.[1,23,24] Most paleoecological reconstructions using bovids are made at the tribe level. Within some bovid tribes, there are taxa that are ecologically diverse and whose ecological niches do not overlap. A substantial amount of paleoecological information is, thus, missing because of our inability to identify fossil bovid teeth to the taxon. Recently, a study of the shape of the occlusal outline of the molar dentition of bovids, through a combination of Fourier functional analysis, principal component analysis (based on covariance matrix), and discriminant function analysis, has achieved high rates of success (>85% of accuracy) in classifying teeth to species within the main African bovid tribes.[24–26] This opens a new window of opportunity for refining paleoecological analyses. This type of analytical approach to bovid identification is necessary to overcome the current praxis of subjective classification by experts.

No study to date shows how bovid tribe identification differs among analysts. However, it should be emphasized that such an identification process is subject to the analyst's experience and, in many cases, identification can be equivocal. Differences in the identification of bovid teeth among experts are common. Brophy et al.[26] report examples of different analysts in South Africa performing different identifications of bovids, using the same faunal assemblage, leading to divergent paleoenvironmental reconstructions. Such a subjective approach to character determination is common in paleontology and taphonomy. The controversial identification and use of bone surface modifications (BSMs) is one of the most widely known examples.[27] This has recently led to the adoption of machine learning methods, channeled through computer vision (CV) techniques, in order to achieve an objective way of approaching the identification of marks on bones. CV was shown to excel expert taphonomists by >50% in identifying cut and trampling marks.[28] The use of deep convolutional neural networks (DCNNs) in deep learning (DL) frameworks has allowed the successful discrimination of cut marks imparted on fleshed or defleshed bones,[29] the discrimination of timing in cut mark modification by biostratinomic abrasive processes,[30] the differentiation of carnivore agency when analyzing tooth mark types,[31–34] and the objective interpretation of controversial BSMs from various sites across the globe.[35] Recently, this experimental carnivore tooth mark database has been applied to an ensemble of tooth marks from the early Pleistocene site of David's Site (DS, Olduvai Gorge, Bed I), and a CV analysis showed that most of the carnivore modifications were attributable to hyenas, as would be expected if carnivore damage was mostly the result of postdepositional hyenid intervention after hominin exploitation of carcass resources.[36] At the slightly younger deposit of FLK North 3, felid and hyenid interactions were also successfully defined on a pilot study applying CV methods to that archaeofaunal assemblage,[37] supporting the previous taphonomic interpretation of the site as a felid-hyenid palimpsest with marginal hominin input.[38]

Given the powerful discrimination obtained in controlled experimentation by DL methods, we thought that they could be expanded to other areas of research. For example, they could also contribute to objectively identifying bovid tribes through image classification of their teeth. Here, we implement these methods, and show that they do not only provide a more objective analytical approach to bovid classification than traditional expert identification, but probably a substantially more accurate one.

SAMPLE AND METHODS

Sample

A sample of 2879 images of the occlusal surface of African bovid teeth, belonging to the upper and lower dentition, from the B.O.V.I.D. (Bovid Occlusal Visual IDentification) database (https://doi.org/10.17605/OSF.IO/R5HSW), were used.[39] These images display different teeth aligned on the same horizontal plane. This is not a conditioning factor for the DL models built, since these tend to learn item identification under different positions. As a matter of fact, one of the protocols of data augmentation (see below) requires generating cloned item images with a large array of vertical and horizontal orientations.

Five tribes of bovids were selected from Brophy and Matthew's original image database: Alcelaphini (*Damaliscus dorcas*, *Alcelaphus buselaphus*, *Connochaetes gnu*, *Connochaetes taurinus*), Antilopini (*Antidorcas marsupialis*), Hippotragini (*Oryx gazelle*, *Hippotragus equinus*, *Hippotragus niger*), Reduncini (*Redunca arundinum*, *Redunca fulvorufula*, *Kobus leche*, *Kobus ellipsiprymnus*), and Tragelaphini (*Tragelaphus scriptus*, *Tragelaphus strepsiceros*, *Taurotragus oryx*). This encompasses almost all the typical African bovid tribes spanning the small and large antelopes. These tribes were selected because they are also the most abundant in the early Pleistocene East African paleontological and archaeofaunal records.[1,40] They were also selected because they encompass most of the bovid types documented in the Olduvai Pleistocene paleontological record on which the research of some of us is focused. The target was bovid classification at the tribe level. Bovini were not included because DCNN and CV methods require large data sets for properly learning object features and the Bovini sample in the B.O.V.I.D. data set was the smallest of all the tribes, and amounted to only 153 highly variable teeth. Our intention with this referential analysis is to provide a solid analytical framework for bovid identification. Images from the B.O.V.I.D. database were cropped to show only the tooth's occlusal surface (Figure 1). The modified data set can be found in Harvard's Dataverse public repository (https://doi.org/10.7910/DVN/TMLXGW). The analysis was performed with all the images in color.

DL analysis

DL methods applied to images, through what is commonly called computer vision (CV), were selected. Here, we used DCNNs from transfer learning (TL) (i.e., pretrained architectures) models. TL models are built from original models that are trained from scratch, sometimes with more than one million images and with a thousand different objects. Such training is intensive and requires weeks of computation. By generating models using such a large number of objects, the algorithms are fine-tuned in detecting subtle characteristics that separate each of these hundreds of objects from one another. This fine-tuned architecture can be reused (or retrained) over a new set of different objects by freezing the trained model layers (or a selection thereof) and adding new ones on top, with the goal of turning the old features into predictions for the new objects during training.

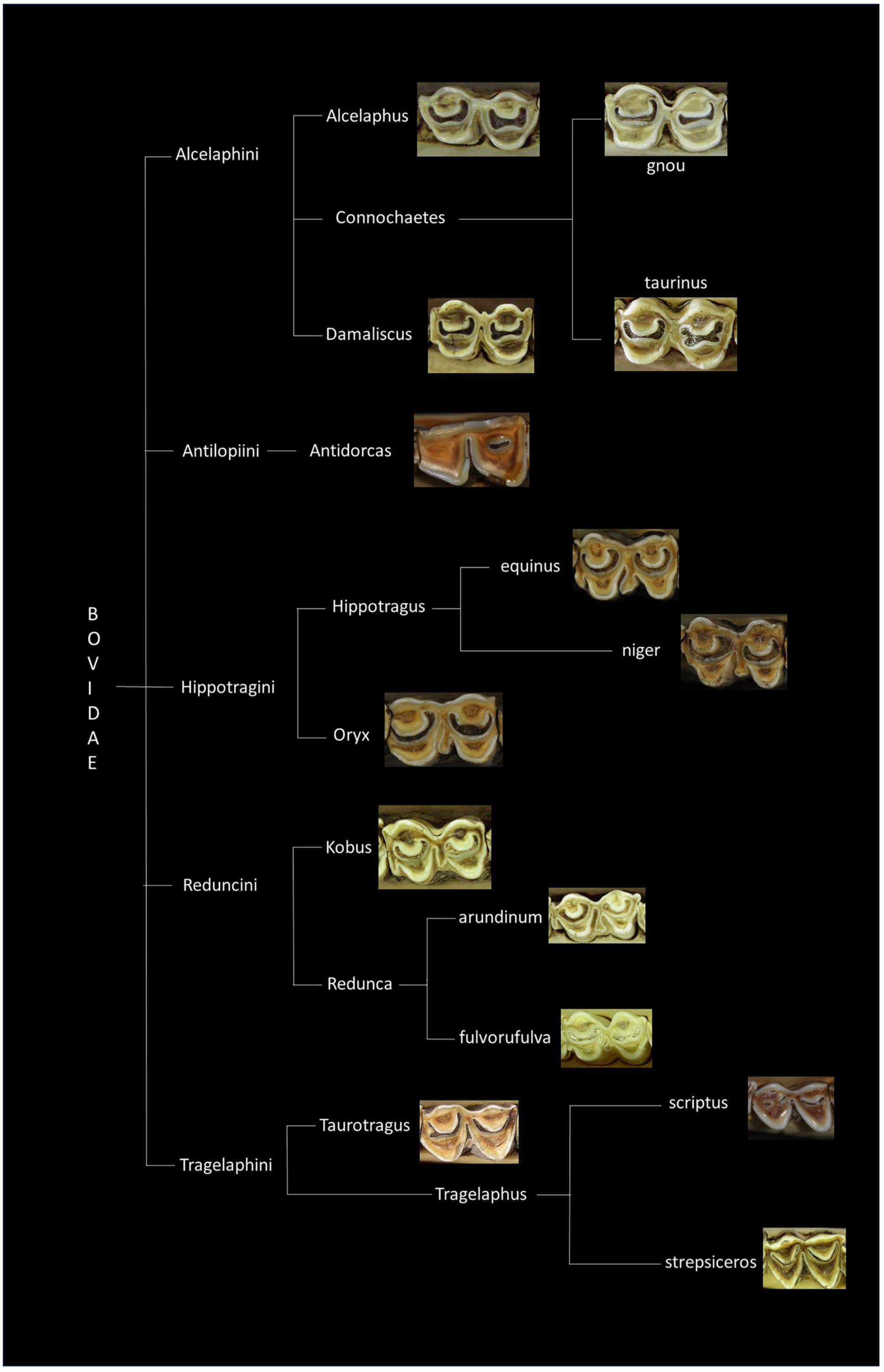

FIGURE 1 Examples of images of teeth from each bovid tribe used for training the deep learning models.

The TL models used here are similar to those that we used previously for BSM analysis.[29,32–35] The architectures selected in the present analysis included sequential, densely connected, and residual networks: ResNet-50 (version 1.0),[41] VGG19,[42] DenseNet-201,[43] and EfficientNet-B7.[44] Sequential networks (VGG19) are based on the stacking of successive layers, each of them containing single input and output tensors. Densely connected networks (DenseNet-201) follow a sequential arrangement of layers, but each layer is connected to every other layer. Residual networks use a similar sequential structure, but they skip successive connections that create identity mappings, and are successful in very deep architectures, where sequential concatenation of multiple layers can lead to underperformance. Here, ResNet-50 and EfficientNet-B7 were the only residual networks used, but they are significantly different. In the case of EfficientNet-B7, this architecture is built on a baseline network that optimized efficiency and accuracy (through the implementation of compound scaling), using the AutoML MNAS framework (to search for the most efficient neural architecture) and including a mobile inverted bottleneck convolution in a residual block format.[44]

The four TL architectures listed above were used to obtain individual models. These particular TL architectures were selected because they yielded highly accurate models in previous work on the identification of agency in BSMs.[30–37] The bovid sample image classification was performed individually for each of these models, which were used in a competitive framework, to derive the best classifier. Subsequently, the four models were used in a cooperative framework through ensemble learning (EL). EL consists of bringing together different algorithms and making them jointly converge in the production of predictions. The EL methods used here consisted of using the four regularized models as the base learners, and then a stacking process was implemented by alternatively using a random forest and an extremely randomized gradient tree algorithm as the meta-learners. The number of estimators used in hyperparameter tuning was 100. Extremely randomized trees, in contrast with random forests, do not resample observations or use best splits when building trees.[45] They create split points for predictors by randomly choosing splitting and selecting the best resulting ones.

The original 2879 image data set was divided into a training set (70% = 2009 images) and a testing set (30% = 870 images). Images were randomly allocated to both sets. Despite the original sample size, the architectures were used with image augmentation to improve their training by reducing the chances of overfitting.[46] The training image data set was augmented through the following procedures: random shifting of width and height (20%), of shear and zoom range (20%), horizontal flipping, and a rotation range of 40°. Image standardization, using bidimensional matrices for standardization and centering, was carried out using each architecture's preprocessing functions. The DCNN models were elaborated using the Keras (2.4.3) application programming interface with a Tensorflow (2.3.0) backend. Computation was carried out on a GPU HP Z6 Workstation using a CUDA computing (cuDNN) environment. All code was made using Python 3.7 (available at: https://doi.org/10.7910/DVN/TMLXGW).

In each of the TL models used, the activation function for every layer was a rectified linear unit (ReLU). The last fully connected layer of the network used a "softmax" activation. The loss function selected was categorical cross-entropy. Cross-entropy measures distances between probability distributions and predictions.[46] The optimizer used was stochastic gradient descend with a learning rate of 0.001 and a momentum of 0.9. Accuracy was the metric selected for the classification process. F1 score values were also obtained to assess balanced accuracy, given the imbalanced nature of the original data set. Training was done using mini-batch kernels of size 32. Testing was made using mini-batch kernels of size 20. Weight update was done using a backpropagation process of 100 epochs.

Training graphs for accuracy and loss were also used to assess over- and underfitting processes. The architectures implemented regularization methods based on Dropout.[47] Dropout consists of the random dropping (i.e., ignoring) of selected neurons during training. This results in DCNN networks that are less reactive to specific neuron weights, producing a network that is more adapted to implement better generalization and less likely to overfit from the training data.

Saliency is a term that refers to specific features in an image that depict identifying locations. A saliency map is a bidimensional topographical representation of those identifying features. Saliency maps can be created from every convolutional layer in a DL network, but they usually are generated using the last convolutional layer prior to the flattening process. There are several types of saliency algorithms. In the present work, we used a gradient weighted activation mapping algorithm (Grad-CAM) in order to detect the features that influenced classification. This method overlays a heatmap on the original image based on gradients of the predicted class derived from the last convolutional feature map. The Grad-CAM algorithm highlights areas of the image that are most relevant for its identification. Here, we applied Grad-CAM to a selection of tooth images from different bovid tribes in order to detect features of the teeth that were determinant for their identification and classification. For this purpose, we selected the most accurate (ResNet-50) and the least accurate (VGG19) models, since both would represent the range ends of the resulting saliency maps.

TABLE 1 Accuracy, loss, and F1 score of the four DL architectures used in the classification of African bovid tribes.

	Accuracy	Loss	F1 score
ResNet-50	0.92	0.24	0.92
VGG19	0.84	0.48	0.85
EfficientNet-B7	0.89	0.28	0.89
DenseNet-201	0.91	0.29	0.91

RESULTS

The ResNet-50 model yielded the highest accuracy (92%) and lowest loss in the classification of the 870 images of the testing set, followed by DenseNet-201 (91%), EfficientNet-B7 (89%), and VGG19 (84%) (Table 1). The F1 score values for all the models indicate a fairly balanced classification of all the tribes. In the ResNet-50 model, the Tragelaphini, Antilopini, and Alcelaphini are the most accurately classified (F1> 95%), followed very closely by Reduncini (F1 = 89%) and Hippotragini (84%) (Table 2). The same order of the highest F1 values (>90%) for Antilopini, Tragelaphini, and Alcelaphini was documented for the VGG19, EfficientNet-B7, and DenseNet-201 models (Table 2). Reduncini and Hippotragini display slightly lower F1 scores, ranging between 69% (VGG19) and 86% (DenseNet-201), according to the model (Table 2).

The accuracy/loss graphs for the 100 epochs during training show that the regularization dropout method worked efficiently at avoiding overfitting. The validation learning curve follows the training curve closely and outperforms it (Figure 2). Out of the four

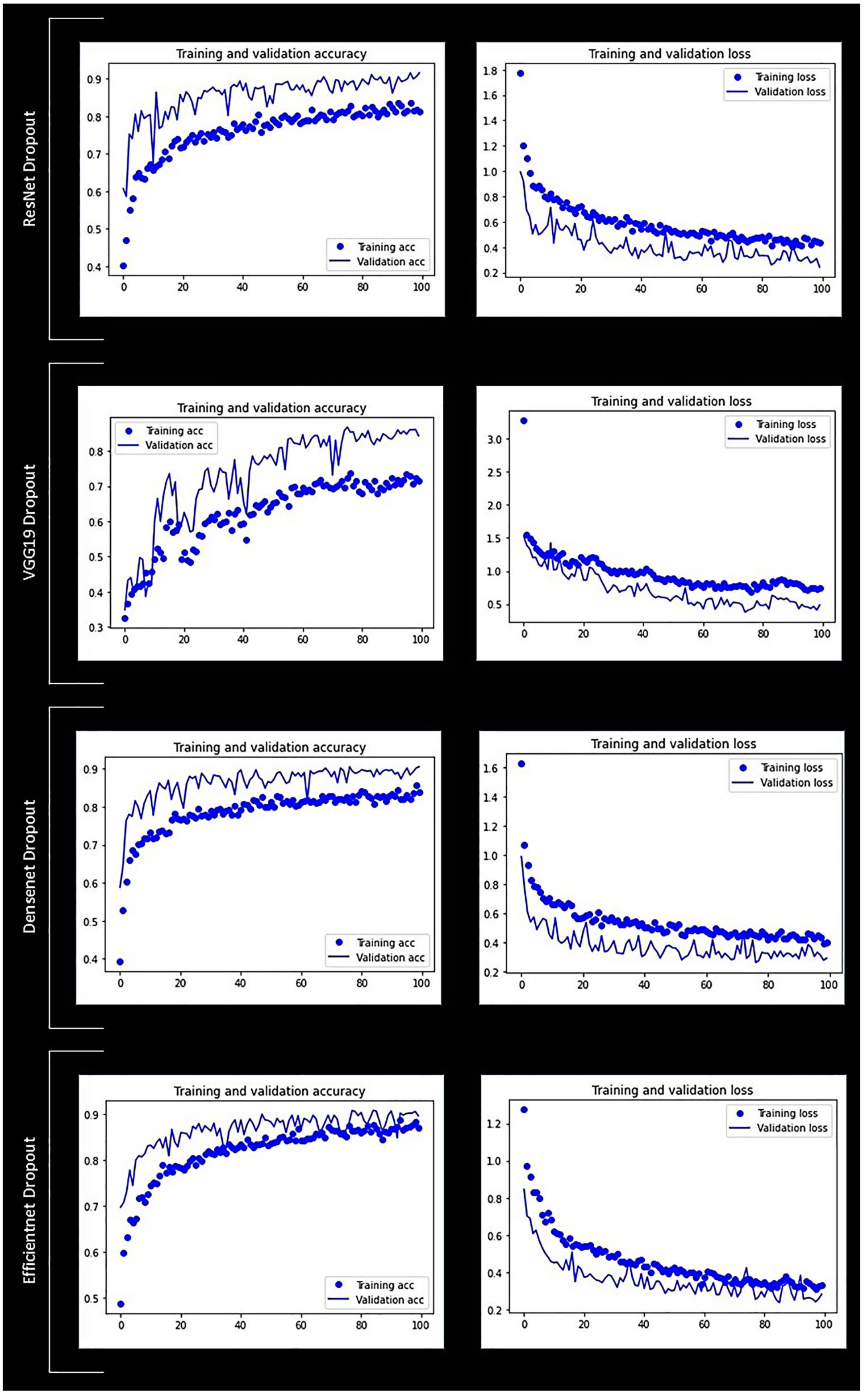

FIGURE 2 Progress graphs showing the performance of the training and validation sets. Accuracy for each model is shown in the left column, and loss values are displayed in the right column.

TABLE 2 F1 score for the classification of the five bovid tribes used in the present analysis according to DL model.

	ResNet-50	VGG19	EfficientNet-B7	DenseNet-201
Alcelaphini	0.95	0.90	0.91	0.92
Antilopini	0.96	0.97	0.93	0.95
Hippotragini	0.84	0.69	0.86	0.86
Reduncini	0.89	0.77	0.82	0.85
Tragelaphini	0.96	0.92	0.95	0.97

architectures, it was EfficientNet-B7 that provided the best fit between the training and testing learning curves, with the latter converging with the former during the last steps of the process (Figure 2).

With such a high accuracy for each of the learning models, it was expected that an ensemble analysis would yield a solid, high, final accurate model. The EL resulting from the combination of the four models as the base learners was carried out using a stacking procedure. The meta-learners used for the upper layer were a random forest and extra randomized trees algorithm. Both yielded the same result, with 92% of the testing set being accurately classified.

When projecting the final result of the most (ResNet-50) and the least (VGG19) accurate models for saliency topographies, we observed that both models coincide in focusing on object detection by using mostly the occlusal surface of the teeth. ResNet-50 encompasses wider areas of the tooth, whereas VGG19 seems to focus more specifically on the enamel and infundibula of the lobes (Figure 3). This latter feature could partially explain the lower accuracy of VGG19 and the higher resolution of ResNet-50. For Alcelaphini, these saliency methods show that the shape of the lobe and the infundibulum are targeted. For Antilopini, ResNet-50 focuses on the mesial outline shape, and both models identify the infundibular shape as very diagnostic. For Hippotragini and Reduncini, the lobe shape, the infundibula, the ribs, and the basal pillar are selected as diagnostic. For Reduncini, even VGG19 widens its scope by selecting most of the central occusal area. For Tragelaphini, whereas VGG19 targets infundibular shape, ResNet-50 focuses more specifically on lobular shape, and especially on the vestibular pointed sections of the lobes (Figure 3).

DISCUSSION

Previous work on objective analytical identification of bovid occlusal surface teeth, using elliptical Fourier analysis (EFA), showed an overall accuracy >85% (for most groups 100%) of correct taxonomic identification of modern experimental dental samples.[24] A more recent reanalysis of the data using a machine learning approach also yielded variably a high accuracy and a low loss depending on the algorithm. For the lower molars, results ranged from 82% accuracy in tribe identification (LDA, linear discriminant analysis) to 89% (RF, random forest), depending on tooth type (M_1, M_2, and M_3). For the upper posterior dentition (M^1, M^2, and M^3), results ranged between 80% (NN, neural network) to 92% accuracy (RF).[48] Results varied according to molar type. RF was better at identifying M_1-M^1 than the other molar types.

These results, based on Fourier analysis of occlusal tooth surfaces, are unparalleled in the identification of modern bovid faunas. Comparisons of the determination accuracy of completely crowned teeth with minimal attrition (>85% of the crown height) with highly worn teeth from aged individuals (<85% of the crown height), testing the null hypothesis of occlusal outline consistency through the lifespan of an animal, showed that all tribes could still be classified correctly >60% of the time (>75% of the time in the case of Alcelaphini and Tragelaphini).[26] This indicated a substantial decrease in accuracy with worn teeth compared with more completely crowned teeth, whose occlusal outlines were more discriminant: all tribes were correctly classified, with an accuracy >85% (with Alcelaphini and Reduncini correctly classified >90% of the time).[26]

The number of tooth specimens per taxon in the B.O.V.I.D. database is insufficient to properly train DCNNs for taxonomic classification. This is why we adopted the tribe approach. Additionally, successful dental classification using the occlusal outline also depends on correctly identifying first the specific tooth within its dental series (i.e., M1, M2, M3…). Fragmented teeth may complicate molar type identification. Furthermore, in some cases, the attribution of the first and second molars is not straightforward for many analysts. In these cases, an alternative method to EFA like DCNN CV would be an adequate substitute or complement. A recent study of partial curves of the occlusal tooth surface, based on EFA, yielded extremely low log-loss values,[49] ensuring that partial preservation of occlusal surfaces could be enough to allow tribe (and even species) distinction.

The TL models used here were trained using all molar teeth together, without distinction of molar type (upper or lower series). The EL methods used, combining the four independent models, yielded an accuracy of 92% correct classification of an extensive testing set composed of 870 images. Single model accuracy was also high, with at least one model also showing 92% accurate identification. This is as high an accuracy as that resulting from the best ML model applied to the EFA, with the added advantage that it is sustained throughout all the molar tooth types, including the upper and lower series.

The high accuracy obtained in this analysis, as well as its balanced distribution among the five tribes, shows that this objective DL classification method can be reliably used to identify bovid tribes in the fossil record. This impacts the confidence with which paleoecological interpretations based on bovid identifications can be made. It also removes the burden of authority-based protocols in the praxis of

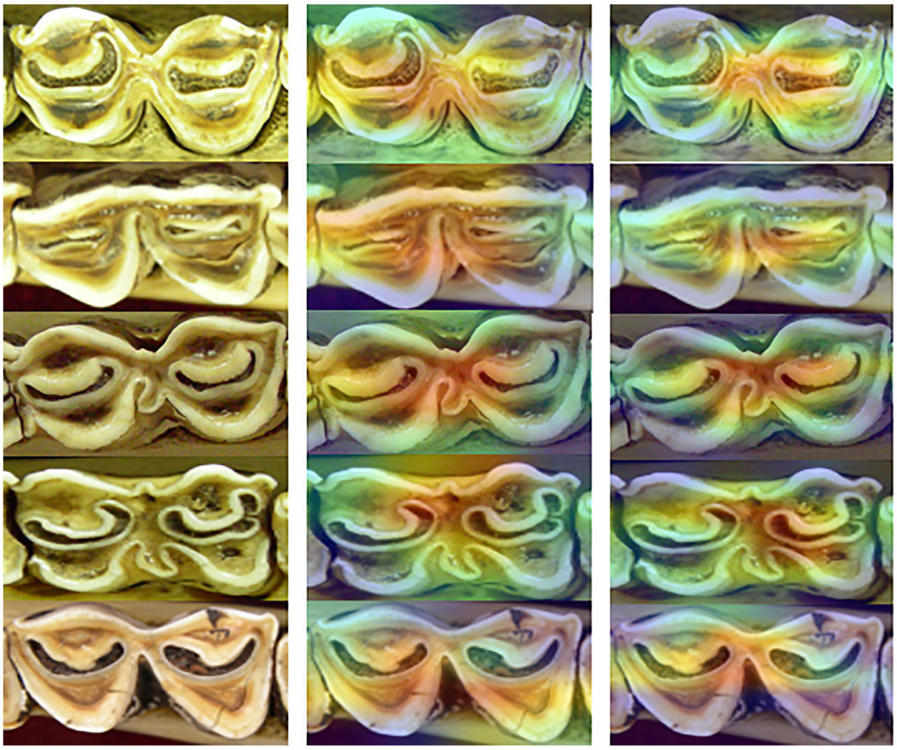

FIGURE 3 Saliency topographic heat maps overlaid on the original dental images for (A) Alcelaphini (*Connochaetes gnou*), (B) Antilopini (*Antidorcas marsupialis*), (C) Hippotragini (*Oryx gazella*), (D) Reduncini (*Kobus ellipsiprymnus*), and (E) Tragelaphini (*Taurotragus oryx*). The examples shown here belong to the left second molar of the inferior dental series. The left column shows the original images. The central column displays the saliency heat map from the ResNet-50 model. The right column displays the saliency heat map from the VGG19 model.

fossil identification. Additionally, it brings a replicable tool with which probabilities in the reliability of identification can be provided. Future work should target exploring classification using 3D modeling. It should also use data augmentation methods (i.e., generative adversarial networks or GANs) to augment taxon-specific image samples to attempt discrimination at the taxon level. These data augmentation protocols would be necessary, because the downside of DL methods is that they require substantially large data sets, and these are currently limited for individual bovid species. If taxon-specific accurate models using DL methods could be obtained, a proper comparison with occlusal outline Fourier methods, which have been the only ones that successfully discriminated not only among tribes, but also among taxa, could be thus established.[24,25] The application of DL techniques to taphonomic research is creating an analytical taphonomy, based on objective and quantitative replicable methods. The present study shows that these methods could also successfully transform traditional paleontological research, which has lacked controlled ways of taxonomic classification, into an analytical paleontology, by expanding methodological protocols in the current praxis of paleobiology.

During our research, we considered surveying the performance of human experts for comparison. However, there were some methodological issues that prevented us from doing it properly. One is the assessment of the degree of experience of the expert. An anonymous method of assessment would be making an image data set available for volunteer analysts, who would have to self-assess themselves for their degree of experience. The inconvenience of this method is that analysts may overestimate their degree of experience and knowledge and provide a false (or unjustified) assessment of their expertise. The other method would be to nonrandomly select a group of widely known experts and ask them to perform the identification. This was attempted initially by one of us unsuccessfully. When a research design can finally be implemented that would guarantee the assessment of the analyst's experience, the present work would then stand as a comparative framework for testing human and machine performance.

CONCLUSIONS

A data set of >2000 images of the occlusal surfaces of upper and lower molar dentition from African bovid tribes was used to train four different DCNN sequential and residual architectures, resulting in a range of classification accuracy of a testing sample of 870 images of 84% (VGG19) to 92% (ResNet-50). An EL analysis involving stacking methods with random forest and extra randomized trees as meta-learners equally yielded an accuracy of 92% of correct classification of the testing image data set. These results are a nice complement to the accuracy estimates obtained in the application of a Fourier functional analysis of dental outline on the same image sample.[24–26] Both methods contribute to analytical approaches to bovid identification that are both objective and reliable, since they can implement confidence estimators by adding probabilities to their classifications. This contributes to the scientific assessment of bovid identification beyond the subjective interpretation of experts, and provides a computational tool that both experts and nonexperts can use to identify their modern and prehistoric samples.

AUTHOR CONTRIBUTIONS

J.B. created the original image dataset. J.B. and G.J.M. performed the initial analysis using Fourier methods. M.P.M. cropped the images for adaptation to the CV analysis. M.D.R. performed the CV study and wrote the draft. All authors revised the draft and participated in creating the final version.

ACKNOWLEDGMENTS

This study was conducted with funding from the Spanish Ministry of Science and Innovation (grant: PID2020-115452GB-C21). We thank the two anonymous reviewers for their very constructive comments on an earlier draft of this paper.

COMPETING INTERESTS

The authors declare no competing interests.

ORCID

Manuel Domínguez-Rodrigo https://orcid.org/0000-0002-7233-331X

PEER REVIEW

The peer review history for this article is available at: https://publons.com/publon/10.1111/nyas.15067

REFERENCES

1. Gentry, A. W., & Gentry, A. (1978). *Fossil Bovidae (Mammalia) of Olduvai Gorge, Tanzania*. Bulletin of the British Museum (Natural History), Geology 29, 289–446.
2. Castelló, J. R. (2016). *Bovids of the world*. Princeton University Press.
3. Kingdon, J. (1988). *East African mammals: An atlas of evolution in Africa, Volume 3, Part C: Bovids*. University of Chicago Press.
4. Bobe, R., & Eck, G. G. (2001). Responses of African bovids to Pliocene climatic change. *Paleobiology*, *27*, 1–47.
5. Bobe, R., Behrensmeyer, A. K., & Chapman, R. E. (2002). Faunal change, environmental variability and late Pliocene hominin evolution. *Journal of Human Evolution*, *42*, 475–497.
6. Merceron, G., & Ungar, P. (2005). Dental microwear and palaeoecology of bovids from the Early Pliocene of Langebaanweg, Western Cape province, South Africa: Research article. *South African Journal of Science*, *101*, 365–370.
7. Schubert, B. W., Ungar, P. S., Sponheimer, M., & Reed, K. E. (2006). Microwear evidence for Plio–Pleistocene bovid diets from Makapansgat Limeworks Cave, South Africa. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *241*, 301–319.
8. Bobe, R., Alemseged, Z., & Behrensmeyer, A. K. (2007). *Hominin environments in the East African Pliocene: An assessment of the faunal evidence*. Springer Science & Business Media.
9. Uno, K. T., Cerling, T. E., Harris, J. M., Kunimatsu, Y., Leakey, M. G., Nakatsukasa, M., & Nakaya, H. (2011). Late Miocene to Pliocene carbon isotope record of differential diet change among East African herbivores. *Proceedings of the National Academy of Sciences*, *108*, 6509–6514.
10. Bibi, F. (2011). Mio-Pliocene faunal exchanges and African biogeography: The record of fossil bovids. *PLoS ONE*, *6*, e16688.
11. Bibi, F., & Kiessling, W. (2015). Continuous evolutionary change in Plio-Pleistocene mammals of eastern Africa. *Proceedings of the National Academy of Sciences*, *112*, 10623–10628.
12. Curran, S. C., & Haile-Selassie, Y. (2016). Paleoecological reconstruction of hominin-bearing middle Pliocene localities at Woranso-Mille, Ethiopia. *Journal of Human Evolution*, *96*, 97–112.
13. Dumouchel, L., & Bobe, R. (2020). Paleoecological implications of dental mesowear and hypsodonty in fossil ungulates from Kanapoi. *Journal of Human Evolution*, *140*, 102548.
14. Bobe, R., & Behrensmeyer, A. K. (2004). The expansion of grassland ecosystems in Africa in relation to mammalian evolution and the origin of the genus Homo. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *207*, 399–420.
15. Bobe, R. (2006). The evolution of arid ecosystems in eastern Africa. *Journal of Arid Environments*, *66*, 564–584.
16. Bobe, R., & Reynolds, S. C. (2022). Hominid Paleoenvironments in Tropical Africa from the Late Miocene to the Early Pleistocene. In S. Reynolds, & R. Bobe (Eds.), *African Paleoecology and Human Evolution* (pp. 161–186). Cambridge University Press, Cambridge.
17. Bobe, R., Behrensmeyer, A. K., Eck, G. G., & Harris, J. M. (2007). Patterns of abundance and diversity in late Cenozoic bovids from the Turkana and Hadar Basins, Kenya and Ethiopia. In R. Bobe, Z. Alemseged, & A. K. Behrensmeyer (Eds.), *Hominin environments in the East African Pliocene: An assessment of the faunal evidence* (pp. 129–157). Springer Netherlands.
18. Bobe, R., Alemseged, Z., & Behrensmeyer, A. K. (2007). Approaches to the analysis of faunal change during the East African Pliocene. In R. Bobe, Z. Alemseged, & A. K. Behrensmeyer (Eds.), *Hominin environments in the East African Pliocene: An assessment of the faunal evidence* (pp. 1–24). Springer Netherlands.
19. Bobe, R., & Carvalho, S. (2019). Hominin diversity and high environmental variability in the Okote Member, Koobi Fora Formation, Kenya. *Journal of Human Evolution*, *126*, 91–105.
20. Dumouchel, L., Bobe, R., Wynn, J. G., & Barr, W. A. (2021). The environments of *Australopithecus anamensis* at Allia Bay, Kenya: A multiproxy analysis of early Pliocene Bovidae. *Journal of Human Evolution*, *151*, 102928.
21. Geraads, D., & Bobe, R. (2020). Ruminants (Giraffidae and Bovidae) from Kanapoi. *Journal of Human Evolution*, *140*, 102383.
22. Bibi, F., Bukhsianidze, M., Gentry, A. W., Geraads, D., Kostopoulos, D. S., & Vrba, E. S. (2009). The fossil record and evolution of Bovidae: State of the field. *Palaeontologia electronica*. https://doc.rero.ch/record/210266/files/PAL_E4341.pdf
23. Gentry, A. W. (1967). *Pelorovis oldowayensis* Reck, an extinct bovid from East Africa. *Bulletin of the British Museum (Natural History)*, Geology 14, 245–299.
24. Brophy, J. K., de Ruiter, D. J., Athreya, S., & DeWitt, T. J. (2014). Quantitative morphological analysis of bovid teeth and implications for paleoenvironmental reconstruction of Plovers Lake, Gauteng Province, South Africa. *Journal of Archaeological Science*, *41*, 376–388.
25. Brophy, J. K. (2011). *Reconstructing the habitat mosaic associated with Australopithecus robustus: Evidence from quantitative morphological analysis of bovid teeth*. Ph. D. dissertation, Texas A&M University.
26. Brophy, J. K., Matthews, G. J., & Thiruvathukal, G. K. (2019). An analysis of the effect of tooth wear on bovid identification. *South African Journal of Science*, *115*, 1–5.
27. Domínguez-Rodrigo, M., Saladié, P., Cáceres, I., Huguet, R., Yravedra, J., Rodríguez-Hidalgo, A., Martín, P., Pineda, A., Marín, J., Gené, C., Julia Aramendi, & Cobo-Sánchez, L. (2017). Use and abuse of cut mark analyses: The Rorschach effect. *Journal of Archaeological Science*, *86*, 14–23.

28. Byeon, W., Domínguez-Rodrigo, M., Arampatzis, G., Baquedano, E., Yravedra, J., Maté-González, M. A., & Koumoutsakos, P. (2019). Automated identification and deep classification of cut marks on bones and its paleoanthropological implications. *Journal of Computational Science*, *32*, 36–43.
29. Cifuentes-Alcobendas, G., & Domínguez-Rodrigo, M. (2019). Deep learning and taphonomy: High accuracy in the classification of cut marks made on fleshed and defleshed bones using convolutional neural networks. *Scientific Reports*, *9*, 18933.
30. Pizarro-Monzo, M., & Domínguez-Rodrigo, M. (2020). Dynamic modification of cut marks by trampling: Temporal assessment through the use of mixed-effect regressions and deep learning methods. *Archaeological and Anthropological Sciences*, *12*, 4.
31. Jiménez-García, B., Aznarte, J., Abellán, N., Baquedano, E., & Domínguez-Rodrigo, M. (2020). Deep learning improves taphonomic resolution: High accuracy in differentiating tooth marks made by lions and jaguars. *Journal of the Royal Society, Interface*, *17*, 20200446.
32. Jiménez-García, B., Abellán, N., Baquedano, E., Cifuentes-Alcobendas, G., & Domínguez-Rodrigo, M. (2020). Corrigendum to 'Deep learning improves taphonomic resolution: High accuracy in differentiating tooth marks made by lions and jaguars'. *Journal of the Royal Society, Interface*, *17*, 20200782.
33. Abellán, N., Jiménez-García, B., Aznarte, J., Baquedano, E., & Domínguez-Rodrigo, M. (2021). Deep learning classification of tooth scores made by different carnivores: Achieving high accuracy when comparing African carnivore taxa and testing the hominin shift in the balance of power. *Archaeological and Anthropological Sciences*, *13*, 31.
34. Abellán, N., Baquedano, E., & Domínguez-Rodrigo, M. (2022). High-accuracy in the classification of butchery cut marks and crocodile tooth marks using machine learning methods and computer vision algorithms. *Geobios*, *72-73*, 12–21.
35. Domínguez-Rodrigo, M., Cifuentes-Alcobendas, G., Jiménez-García, B., Abellán, N., Pizarro-Monzo, M., Organista, E., & Baquedano, E. (2020). Artificial intelligence provides greater accuracy in the classification of modern and ancient bone surface modifications. *Scientific Reports*, *10*, 18862.
36. Cobo-Sánchez, L., Pizarro-Monzo, M., Cifuentes-Alcobendas, G., Jiménez García, B., Abellán Beltrán, N., Courtenay, L. A., Mabulla, A., Baquedano, E., & Domínguez-Rodrigo, M. (2022). Computer vision supports primary access to meat by early Homo 1.84 million years ago. *PeerJ*, *10*, e14148.
37. Vegara-Riquelme, M., Gidna, A., Uribelarrea del Val, D., Baquedano, E., & Domínguez-Rodrigo, M. (2023). Reassessing the role of carnivores in the formation of FLK North 3 (Olduvai Gorge, Tanzania): A pilot taphonomic analysis using Artificial Intelligence tools. *Journal of Archaeological Science: Reports*, *47*, 103736.
38. Domínguez-Rodrigo, M., Barba, R., & Egeland, C. P. (2007). *Deconstructing Olduvai: A taphonomic study of the Bed I sites*. Springer Science & Business Media, New York.
39. Brophy, J. K., & Matthews, G. J. (2022). Reference database of teeth images from the Family Bovidae. *Scientific Data*, *9*, 396.
40. Bibi, F. (2009). *Evolution, systematics, and paleoecology of bovinae (Mammalia: Artiodactyla) from the late Miocene to the recent*. Ph. D. dissertation, Yale University.
41. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778).
42. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. International Conference on LEarning Representations. arXiv preprint. arXiv:1409.1556.
43. Huang, G., Liu, Z., & Van Der Maaten, L. (2017). Densely connected convolutional networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4700–4708).
44. Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning* (pp. 6105–6114). PMLR.
45. Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, *63*, 3–42.
46. Chollet, F. (2017). *Deep learning with Python*. Manning Publications Company.
47. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, *15*, 1929–1958.
48. Matthews, G. J., Brophy, J. K., Luetkemeier, M., Gu, H., & Thiruvathukal, G. K. (2018). A comparison of machine learning techniques for taxonomic classification of teeth from the Family Bovidae. *Journal of Applied Statistics*, *45*, 2773–2787.
49. Matthews, G. J., Bharath, K., Kurtek, S., Brophy, J. K., Thiruvathukal, G. K., & Harel, O. (2021). Shape-based classification of partially observed curves, with applications to anthropology. *Frontiers in Applied Mathematics and Statistics*, *7*, 1–49.

How to cite this article: Domínguez-Rodrigo, M., Brophy, J., Mathews, G. J., Pizarro-Monzo, M., & Baquedano, E. (2023). African bovid tribe classification using transfer learning and computer vision. *Ann NY Acad Sci.*, *1530*, 152–160. https://doi.org/10.1111/nyas.15067