DOI:10.1190/segam2021-3574963.1

Using deep learning to detect rare archaeological features: A case from coastal South Carolina, USA

Dylan S. Davis* – *Department of Anthropology, The Pennsylvania State University*
Gino Caspari – *Department of Archaeology, University of Sydney*
Institute of Archaeological Sciences, University of Bern
Matthew C. Sanger – *National Museum of the American Indian, Smithsonian Institution*
Carl P. Lipo – *Department of Anthropology, Binghamton University*

Summary

Among archaeologists using remote sensing there is tremendous potential for the use of deep learning models for the prospection of archaeological features. The need for relatively large training datasets, technical expertise, and computational requirements, however, has slowed the adoption of these techniques. Here, we train a series of deep learning models using two different model architectures (i.e., single-stage and dual-stage) to detect shell rings, a circular midden feature that is found across the American Southeast. Native American groups constructed these features during the mid-Holocene (5000-3000 cal B.P.). These deposits offer important information about pre-European contact socioeconomic organization among Native American groups (Figure 1). In the coastal area of the Atlantic, however, these features are relatively rare: only about 50 shell rings have been documented by archaeologists to date. To expand our knowledge of these features, we test RetinaNet and Mask-RCNN deep learning models as means of detecting shell rings from wide-area LiDAR data using extremely small training datasets. We demonstrate that the use of "negative" training data to identify non-archaeological features helps to improve model performance. Furthermore, we show that while popular dual-stage detectors like Mask R-CNN perform better than single-stage models like RetinaNet, single-stage models still achieve acceptable levels of accuracy and require a fraction of the computational and time requirements of dual-stage detectors.

Introduction

Shell rings are circular refuse piles composed of plant and animal remains that surround an empty central plaza (Russo 2004; Sanger 2017). These deposits represent some of the earliest evidence of permanent human occupation in the coastal regions of the American Southeast (Figure 1; Russo 2004). The nature of past community activities that produced these shell-rings remains debated among archaeologists, some of whom have focused on identifying the degree to which these deposits served as residential, ritualistic, or a mixture of mundane and ceremonial activities (Russo 2004; Sanger and Ogden 2018; Trinkley 1985).

Despite continued investigation, current archaeological knowledge of shell ring distribution is limited, as only about 50 have been recorded in the entirety of the region (Figure 1). Because of their location is often within difficult-to-survey, dense forests and marshlands, these deposits are mostly known on the basis of large examples that are the most recognizable and most easily accessed (Davis et al. 2020). As a result, archaeologists lack a full inventory of extant shell rings that would allow for a comprehensive investigation into the morphological variability of shell rings and the range of contexts in which they are found.

Deep learning, a branch of machine learning, has been rapidly gaining popularity among archaeologists seeking to identify features among large remote sensing datasets in the past several years (e.g., Caspari and Crespo 2019; Trier, Reksten, and Løseth 2021). Convolutional Neural Networks (CNNs), in particular, have proven highly effective at increasing true positives while reducing false positive results in object detection studies (e.g., Caspari and Crespo 2019; Lambers, Verschoof-van der Vaart, and Bourgeois 2019). Yet, applications of deep learning within archaeology have been limited because of the amount of training data required

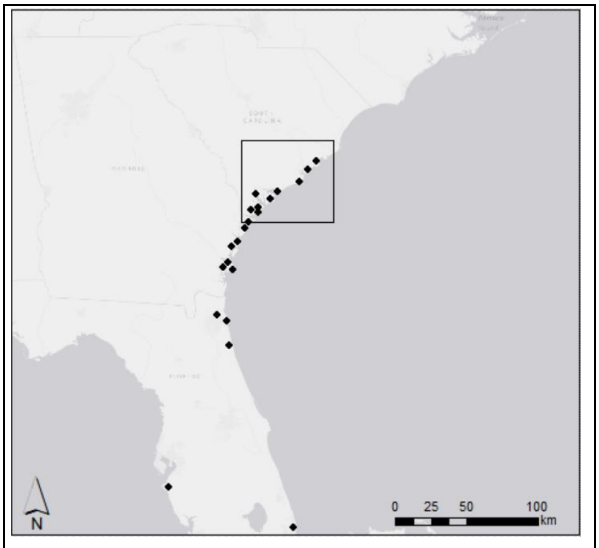


Figure 1: Location of confirmed shell rings surrounding the study area (black box). Service Layer Credits: ESRI, HERE, GARMIN, OpenStreetMap contributions, and the GIS User Community

DOI:10.1190/segam2021-3574963.1

to train these models and the computational cost of these methods.

Within deep learning, object detection tasks are achieved by one of two means. On the one hand, one-stage detectors attempt to identify features by taking an input image and learning the class probabilities and bounding box coordinates, similarly to a regression problem. On the other hand, two-stage detectors offer advantages in using the results of an initial stages analysis to generate a sparse set of potential feature maps that should contain all objects within the data while filtering out most negative locations. The second stage then identifies the potential feature maps as either part of the foreground or background (Lin, Dollár, et al. 2017). Two-stage detectors (e.g., R-CNNs) often achieve greater accuracy than one-stage detectors because of their ability to narrow down candidate objects for detection, but some single stage detectors are beginning to close this gap (e.g., Single Shot Detectors [SSD] (Liu et al. 2016), You-Only-Look-Once [YOLO] (Redmon et al. 2016), and RetinaNet (Lin, Dollár, et al. 2017; Lin, Goyal, et al. 2017).

In this paper, we demonstrate how shell rings can be successfully identified from remote sensing data using both single- and dual-stage deep learning models trained with very small sample sizes ($n < 30$). Our approach makes use of augmentation (i.e., the creation of synthetic data to artificially increase training data sample sizes), the inclusion of "negative" training classes to filter out potential false positives, and transfer learning techniques. We accomplish this task by testing two RetinaNet models (Lin, Goyal, et al. 2017; Lin, Dollár, et al. 2017) using a single-class dataset (shell rings) and a three-class dataset (shell rings, mounds, and "modern" (or negative) features) and two Mask R-CNN models (He et al. 2017) using the same set of training data classes.

Materials & Methods

We acquired LiDAR for Beaufort County from the National Oceanic and Atmospheric Administration (NOAA) with 1.3m nominal point spacing and 15cm vertical RMSE. We also acquired LiDAR for Charleston County from the South Carolina Department of Natural Resources (DNR) with 1.3m nominal point spacing and 30cm vertical RMSE. These two counties (a total area of 5,900 km^2) contain over a dozen known shell ring features. We then created digital elevation models (DEMs) with 1.5m spatial resolution by interpolating the LiDAR point data using inverse distance weighting (IDW).

Next, we created a hillshade map and a slope map from LiDAR point data. Both of these visualizations have improved object detection tasks within archaeology (Devereux, Amable, and Crow 2008). We then created a composite multiband raster in ArcGIS Pro (version 2.6.2; ESRI 2020) consisting of hillshade, slope, and the original DEM.

Using the multiband raster, we generated training data using the Label Training Data for Deep Learning Analysis tool in ArcGIS Pro. We created these training data in two formats: Pascal VOC for the RetinaNet models and Mask_RCNN format for the Mask R-CNN models. We use two kinds of training data consisting of 1 class (shell rings) and 3 classes (shell rings, mounds, and modern, non-archaeological objects). These two sets of training data were devised to see if adding additional mounded features would reduce misclassifications of shell rings. The "modern" class is used as a check against false positives that resemble archaeological ring and mound features. In total, we selected 16 shell rings located in our study area, 21 circular mounds, and 36 modern structures as training samples.

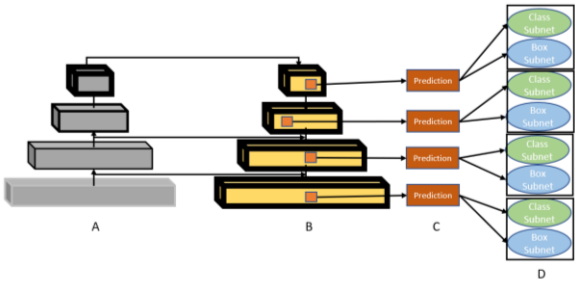


Figure 2: RetinaNet model architecture. A and B represent the FPN, with A consisting of a backbone model (in our case ResNet 50). Predictions made in C) are divided into two components D) Class and box subnets. Class subnets determine the probability of a given area containing an object. Box subnets label objects within class subnets.

Next, to increase our sample sizes, we augmented the training data using 45-degree rotations, thereby synthetically increasing our sample sizes by a factor of 8. For RetinaNet models, we generated a total of 808 images of shell rings, 880 mounds, and 1860 "modern" samples. We then exported these training data as 200x200 pixel images using the Export Training Data for Deep Learning tool in ArcGIS Pro. For Mask R-CNN models, our training data consisted of 776 images of shell rings, 720 mounds, and 1316 "modern" samples. These training data consist of 200x200 pixel images that were produced by the Export Training Data for Deep Learning tool in ArcGIS Pro.

Next, we trained two RetinaNet models (Lin, Goyal, et al. 2017; Lin, Dollár, et al. 2017) using the single-class and three-class training data with a batch size of 8 and a ResNet50 backbone architecture (He et al. 2017). ResNet50 is a transfer learning architecture that is trained on the ImageNet dataset with 50 layers. This step helps to improve

both the speed and accuracy of model training. RetinaNet operates on four main components (Figure 2): a bottom-up pathway, a top-down pathway, a classification subnetwork, and a regression subnetwork. The bottom-up pathway calculates a set of features at particular locations (or feature maps) at multiple scales. The top-down pathway up-samples the coarser feature maps and merges them with the bottom-up feature maps of the same size. Next, the classification subnetwork calculates the probability of an object being present at a given location and belonging to a specific class (defined by training data). The regression subnetwork then creates a labeled object from that classification.

Since Mask R-CNN architecture (He et al. 2017) has proven useful for detecting archaeological deposits (Bonhage et al. 2021), we trained two Mask R-CNN models using the single-class and three-class training data using a batch size of 6 and a ResNet50 backbone architecture (Figure 3). Batch size was lowered based on prior testing which showed higher batches to perform worse for this model type.

We set up the training procedure to run the model 50 times and to stop training when the model no longer improved to save time and processing power. We withheld 10% of the training data for validation of each model's performance.

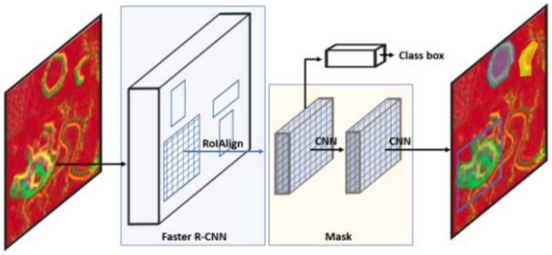


Figure 3: Framework for Mask R-CNN object detection (after He et al. 2017). The Faster R-CNN identifies regions of interest (ROIs) and aligns them (RoIAlign) to preserve spatial information. A Mask is then created using two CNNs, which segments the image into objects. The Faster-R-CNN then classifies these objects based on overall confidence scores.

Results

Our RetinaNet models required between 2 and 3 hours to run on a computer with a NVIDIA Quadro p4000 GPU, an Intel(R) Core(TM) i7-7700K CPU @ 4.20GHz, 4200 Mhz, 4 Core(s), 8 Logical Processor(s), and 64 GB of RAM. Accuracy of shell ring detection rated by validation data was 63% for the model trained on the single-class data and 69% for the model trained on the three-class data (Table 1).

Our Mask R-CNN models required 50-60 hours to run on the same machine as the RetinaNet models (Figure 4). The model trained on the single-class data identified shell rings with 65% accuracy and the model trained on the three-class data identified shell rings with 75% accuracy (Table 1).

When assessing the models trained on the three-class data, the RetinaNet model re-identified a total of 15 out of 16 pre-documented shell rings (93.75%), 12 out of 21 pre-documented mounds (57.14%) and 29 out of 36 (80.56%) "modern" locations. The Mask R-CNN model trained on the three-class data reidentified all 16 pre-documented shell rings (100%) 20 out of 21 pre-documented mounds (95.2%), and 30 out of 36 (83.33%) "modern" features.

Model Type	# of Classes	Validation Accuracy (Rings, Mounds, Modern)
Mask R-CNN	1 (shell rings)	65%
Mask R-CNN	**3 (shell rings, mounds, modern)**	**75%, 80%, 60%**
RetinaNet	1 (shell rings)	63%
RetinaNet	3 (shell rings, mounds, modern)	69%, 41%, 54%

Table 1: Accuracy metrics for the tested deep learning models. All models were trained on Resnet-50 backbone architectures.

These results demonstrate that from a computational requirement standpoint, single-stage detectors like RetinaNet should be considered by archaeologists aiming to improve archaeological site inventories and prospection efforts. The use of these models, however, should be iterative to ensure that as new data is gathered, models are retrained to improve their performance. In contrast, if the goal is to get the most accurate results despite computational costs, then dual-stage detectors like Mask R-CNNs are probably the best method, as they attain higher levels of accuracy even with limited training data.

Conclusions

Previous investigations of shell rings have relied on ground surveys that are largely opportunistic and have resulted in a haphazard and incomplete understanding of total abundance and distribution (Davis et al. 2020). Here, we demonstrate that deep learning algorithms can be used to successfully identify shell ring features in LiDAR despite extremely small available training datasets. The key to our success was in the use of "negative" classes, which serve to identify

DOI:10.1190/segam2021-3574963.1

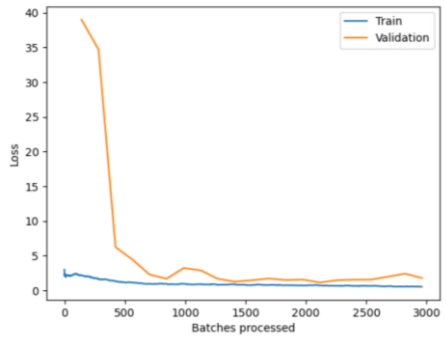
features with similar characteristics to archaeological deposits, but are false positives.

Additionally, this study demonstrates the power of single-stage deep learning models like RetinaNet for archaeological prospection research. While the accuracy attained by these models is comparably lower than more popular dual-stage models like R-CNNs, single-stage models allow for rapid assessment, as they can be trained using fewer computational resources and in a faster timeframe than dual-stage models. Despite this fact, there have been virtually no archaeological studies that have utilized single-stage deep learning architectures.

With continued research, deep learning models can be applied to record shell rings – as well as other rare or under-recorded archaeological features – to improve our understanding of landscape-level site distributions. This work also holds promise for the implementation of remote sensing analyses in areas where archaeological data is currently and overwhelmingly absent (i.e., places with limited prior archaeological investigations, conflict zones where work cannot be safely conducted, etc.). Methods like single-stage deep learning models trained on limited sample sizes have the potential to play a major role in locating new archaeological sites in areas threatened by anthropogenic and natural forces, including areas where there are major gaps in archaeological information.

Acknowledgements
We extend our thanks the organizers of this session, *Applications for Archaeology, Void, and Target Detection*, Dr. Blaire Schneider and Dr. Gregory Tsokas, for inviting us to contribute our work.

A.



B.

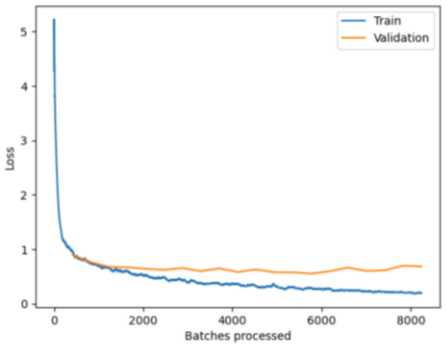


Figure 4: A. Loss values for RetinaNet model trained on three-class dataset. B. Loss values for Mask R-CNN model trained on three-class data.

REFERENCES

Bonhage, A., M. Eltaher, T. Raab, M. Breuß, A. Raab, and A. Schneider, 2021, A modified mask region- based convolutional neural network approach for the automated detection of archaeological sites on high-resolution light detection and ranging-derived digital elevation models in the north German lowland: Archaeological Prospection, **28**, 177–186, doi: https://doi.org/10.1002/arp.1806.

Caspari, G., and P. Crespo, 2019, Convolutional neural networks for archaeological site detection – finding 'princely' tombs: Journal of Archaeological Science, **110**, 104998, doi: https://doi.org/10.1016/j.jas.2019.104998.

Davis, D. S., R. J. DiNapoli, M. C. Sanger, and C. P. Lipo, 2020, The integration of Lidar and legacy datasets provides improved explanations for the spatial patterning of shell rings in the American southeast: Advances in Archaeological Practice, **8**, 361–375, doi: https://doi.org/10.1017/aap.2020.18.

Devereux, B. J., G. S. Amable, and P. Crow, 2008, Visualisation of LiDAR terrain models for archaeological feature detection: Antiquity, **82**, 470–479, doi: https://doi.org/10.1017/S0003598X00096952.

ESRI, 2020, ArcGIS Pro (version 2.6.2): Environmental Systems Research Institute, Inc., https://www.esri.com/en-us/arcgis/products/arcgis-pro/.

He, K., G. Gkioxari, P. Dollár, and R. Girshick, 2017, Mask R-Cnn: Proceedings of the IEEE International Conference on Computer Vision, 2961–2969, doi: https://doi.org/10.1109/ICCV.2017.322.

Lambers, K., W. Verschoof-van der Vaart, and Q. Bourgeois, 2019, Integrating remote sensing, machine learning, and citizen science in Dutch archaeological prospection: Remote Sensing, **11**, 794, doi: https://doi.org/10.3390/rs11070794.

Lin, T.-Y., P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, 2017, Feature pyramid networks for object detection: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 936–944, doi: https://doi.org/10.1109/CVPR.2017.106.

Lin, T.-Y., P. Goyal, R. Girshick, K. He, and P. Dollár, 2017, Focal loss for dense object detection: IEEE International Conference on Computer Vision (ICCV), 2999–3007, doi: https://doi.org/10.1109/ICCV.2017.324.

Russo, M., 2004, Measuring shell rings for social inequality, in Signs of Power: The Rise of Cultural Complexity in the Southeast, J. L. Gibson and P. J. Carr, eds., University of Alabama Press, 26–70.

Sanger, M. C., 2017, Coils, slabs, and molds: Examining community affiliation between Late Archaic shell ring communities using radiographic imagery of pottery: Southeastern Archaeology, **36**, 95–109, doi: https://doi.org/10.1080/0734578X.2016.1267466.

Sanger, M. C., and Q.-M. Ogden, 2018, Determining the use of Late Archaic shell rings using lithic data: 'Ceremonial Villages' and the importance of stone: Southeastern Archaeology, **37**, 232–252, doi: https://doi.org/10.1080/0734578X.2017.1398995.

Trier, Ø. D., J. H. Reksten, and K. Løseth, 2021, Automated mapping of cultural heritage in Norway from airborne Lidar data using faster R-CNN: International Journal of Applied Earth Observation and Geoinformation **95**, 102241, doi: https://doi.org/10.1016/j.jag.2020.102241.

Trinkley, M. B., 1985, "The form and function of South Carolina's Early Woodland shell rings," *in* Structure and Process in Southeastern Archaeology, R. S. Dickens and H. T. Ward, eds., 102–118.