

# Evaluating Mask R-CNN models to extract terracing across oceanic high islands: A case study from Sāmoa

Seth Quintus<sup>1</sup>  | Dylan S. Davis<sup>2,3</sup>  | Ethan E. Cochrane<sup>4</sup> 

<sup>1</sup>Department of Anthropology, University of Hawai'i at Mānoa, Honolulu, Hawai'i, USA

<sup>2</sup>Columbia Climate School, Columbia University, New York, New York, USA

<sup>3</sup>Division of Biology and Paleoenvironment, Lamont-Doherty Earth Observatory, Palisades, New York, USA

<sup>4</sup>Anthropology, University of Auckland, Auckland, New Zealand

## Correspondence

Seth Quintus, Department of Anthropology, University of Hawai'i at Mānoa, 2424 Maile Way, Saunders 346, Honolulu, HI 96822, USA.  
 Email: [squintus@hawaii.edu](mailto:squintus@hawaii.edu)

## Funding information

Marsden Fund; National Science Foundation SBE Postdoctoral Research Fellowship

## Abstract

Lidar datasets have been crucial for documenting the scale and nature of human ecosystem engineering and land use. Automated analysis methods, which have been rising in popularity and efficiency, allow for systematic evaluations of vast landscapes. Here, we use a Mask R-CNN deep learning model to evaluate terracing—artificially flattened areas surrounded by steeper slopes—on islands in American Sāmoa. Mask R-CNN is notable for its ability to simultaneously perform detection and segmentation tasks related to object recognition, thereby providing robust datasets of both geographic locations of terracing features and their spatial morphometry. Using training datasets from across American Sāmoa, we train this model to recognize terracing features and then apply it to the island of Tutuila to undertake an island-wide survey for terrace locations, distributions and morphologies. We demonstrate that this model is effective ( $F_1 = 0.718$ ), but limitations are also documented that relate to the quality of the lidar data and the size of terracing features. Our data show that the islands of American Sāmoa display shared patterns of terracing, but the nature of these patterns are distinct on Tutuila compared with the Manu'a island group. These patterns speak to the different interior configurations of the islands. This study demonstrates how deep learning provides a better understanding of landscape construction and behavioural patterning on Tutuila and has the capacity to expand our understanding of these processes on other islands beyond our case study.

## KEY WORDS

instance segmentation, lidar, Mask R-CNN, Oceania, terracing

## 1 | INTRODUCTION

Terracing is a hallmark of engineering in environments featuring high topographic relief (Bevan et al., 2013; Brown et al., 2020; Healy et al., 1983; Korobov & Borisov, 2013; Pérez Rodríguez & Anderson, 2013; Sandor et al., 1990). These features provide foundations of activities, most notably including habitation and agriculture, in areas that are otherwise unfavourable or marginal for use across the globe (Gadot et al., 2016; Quintus et al., 2017; Treacy & Denevan, 1994). They are also persistent, providing an opportunity for populations to build on generations of labour to engineer extensive areas that allow populations and agricultural systems to grow

(Brown et al., 2020). Because of their persistence and the goal of those constructing the terraces, these features change the environment, at times to such an extent that they have cascading consequences on soil development and vegetation (Brown et al., 2020; Hightower et al., 2014), among other things. Given their importance in creating usable land within an otherwise marginal space, the documentation of terracing is an important step in understanding social organization, agricultural development, the formation of anthropogenic environments and population reliance (Acabado et al., 2019; Pérez Rodríguez, 2016).

Given their function, terraces are often found in relatively inaccessible locations. This is especially true in the global tropics, where

terracing is not only located in areas of high topographic relief but also under dense vegetation cover (Hightower et al., 2014; Quintus et al., 2017). Because of this, estimates of terracing across landscapes are difficult to generate and often extrapolate from smaller areas that have been intensively mapped (Brown et al., 2020, p. 567). Lidar datasets have offered an opportunity to remedy this situation. Terracing is often visible in images derived from lidar datasets given the contrast between the anthropogenic elements of the feature and the surrounding slopes (Chase & Weishampel, 2016; Hightower et al., 2014; Macrae & Iannone, 2016; McCoy et al., 2011), though this depends on the quality of the lidar dataset and the size of the features (Sánchez Díaz & García Sanjuán, 2022). While diverse in function and variable in morphology (Treacy & Denevan, 1994), basic elements of terracing are shared worldwide as a result of convergent evolution. Namely, these features all possess flat or near-flat surfaces with sloping sides that contrast with the surrounding landscape. This contrast in slope creates an opportunity for their identification (McCoy et al., 2011).

Terracing is a common component of the archaeological record in the Pacific (Addison, 2006; Bayliss-Smith & Hvding, 2015; Kirch, 1994; Kuhlken & Crosby, 1999; Lepofsky, 1994; Liston, 2009; McCoy et al., 2011; Sand et al., 2003). These terraces were used for a variety of functions, including habitation, agriculture and defence (Allen, 2004; Best, 1993; Taomia, 2000). However, few studies have attempted to map the full distribution of terracing across individual islands given the financial and labour constraints of pedestrian survey in tropical environments. The distribution of terraces can provide a useful estimate of the scale of human impacts in these island environments given that terraces are a signature of intensified landuse (Brown et al., 2020). The use of lidar datasets in Oceania offers an opportunity to better track the scale of terracing across the region.

Within Oceania, remote sensing research demonstrates more extensive and intensive land use than previously thought (Bedford et al., 2018; Cochrane & Mills, 2018; Freeland et al., 2016; Jackmond et al., 2018; Quintus et al., 2015). This, in turn, has led to a reconsideration of population densities (Parton & Clark, 2022) and the ways that communities supported those larger populations (Bedford et al., 2018; Quintus, 2018). However, the analysis of lidar datasets has been spatially limited and often based on manual feature extraction (Davis, 2019). Manual feature extraction is useful (Quintus et al., 2017), but automated extraction techniques allow for a more systematic and efficient evaluation of larger stretches of the landscape (Bickler, 2021; Câmara et al., 2023; Cerrillo-Cuenca & Bueno-Ramírez, 2019; Davis, 2019).

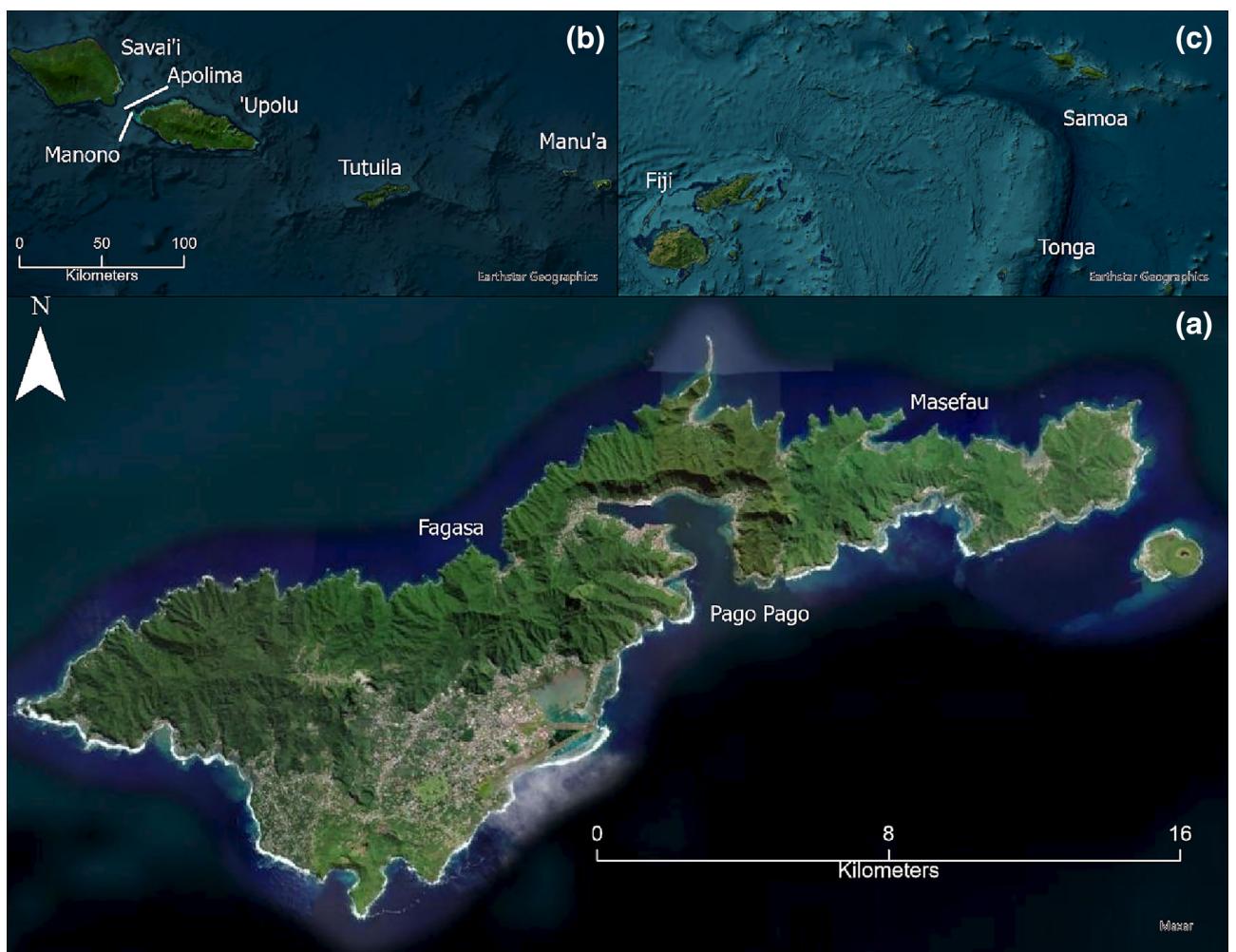
Advances in automated feature extraction provide the means to both locate archaeological features across large swaths of landscape as well as generate geospatial datasets of features that can be used for further studies of feature size, shape and distribution (e.g., Berganzo-Besga et al., 2021; De Smedt et al., 2022; Freeland et al., 2016; Magnini et al., 2017; Verschoof-van der Vaart & Landauer, 2020). This provides another tool for feature identification and the development of distribution maps for subsequent analysis within the archaeological context of a particular place (Bennett

et al., 2014; Huggett, 2021). We build on this research and introduce a deep learning model to extract the locations and morphological characteristics of artificial terraces from the sloping topography characteristic of high islands in Oceania. Using training datasets from American Sāmoa, we apply the model to the island of Tutuila and undertake an island-wide survey for terrace locations, distributions and morphologies (Figure 1). We then discuss the implications of this research for wider documentation of engineered landscapes across Oceania.

## 2 | BACKGROUND

The Samoan archipelago lies in the central Pacific between 13° and 14° latitude south. Currently, the archipelago is split into two geopolitical units, the territory of American Sāmoa and the independent nation of Sāmoa, though the archipelago has a shared cultural history. Tutuila is the largest island of American Sāmoa but the third largest island of the Samoan archipelago. All the islands of American Sāmoa feature high drastic topographic relief, with narrow coastal plains that skirt a steep volcanic landmass. The larger islands of Sāmoa are gentler in topography with developed valleys in some locations. Valleys are not well developed in the islands of American Sāmoa, though they are more apparent on Tutuila than in the adjacent Manu's group, which consists of the islands of Ofu, Olosega and Ta'u. Precipitation is high, with annual averages in excess of 3000 mm along the coast and higher in the interior. As such, these islands are densely vegetated with a mix of native rainforest and introduced economic species.

These characteristics have made traditional pedestrian survey in island interiors difficult. Prior to the use of lidar datasets in the archipelago, pedestrian surveys in the interiors were localized or nonintensive (Best, 1993; Clark & Herdrich, 1993; Hunt & Kirch, 1988; Pearl, 2004). While terracing and other forms of landscape modification were recorded during these initial surveys, it was generally assumed that the focus of land use was on the coast during the entirety of the cultural sequence that began by 2800–2400 calBP across the archipelago (Green, 2002). The acquisition and use of lidar data have fundamentally altered these views, at least for the Manu'a group and parts of Sāmoa (Glover et al., 2020; Jackmond et al., 2018; Quintus et al., 2015, 2017). In Manu'a, paired pedestrian surveys and manual analysis of lidar datasets have allowed documentation of the full distribution of terracing on the island of Ofu and Olosega (Quintus, 2020). These features are relatively large. While the mean size varies from one community to the next, all means are over 140 m<sup>2</sup>. As such, these features are visually apparent in products derived from even coarse-grained lidar datasets. From these analyses, it is estimated that terracing encompasses over 30% and 60% of each island's interior, respectively (Quintus, 2018). Likewise, applications of lidar datasets to archaeological survey in Sāmoa have provided novel insights into the high density of occupation in some locations (Glover et al., 2020; Jackmond et al., 2018). These results highlight how extensively these interior landscapes were used and modified.



**FIGURE 1** Islands and archipelago mentioned in the text. (a) The island of Tutuila. Note the heavily vegetated interior, which is the subject of our analysis. (b) The Samoan archipelago. (c) The position of Samoa in Fiji-West Polynesia. [Colour figure can be viewed at [wileyonlinelibrary.com](https://wileyonlinelibrary.com)]

Lidar analysis in Manu'a and Sāmoa has been largely accomplished using manual feature extraction (Jackmond et al., 2018; Quintus et al., 2017). These techniques were successful in Manu'a, in part, because of the size of islands, which are all under 36 km<sup>2</sup>. In Sāmoa, researchers using manual extraction techniques have largely focused on specific locations rather than island-wide surveys (Jackmond et al., 2018). While manual feature extraction has no doubt been important, the time necessary for such techniques at the scale of individual islands limits the full potential that these lidar datasets may have in illustrating the full spatial extent of landscape transformation, which may provide useful data on the density of occupation and the drivers of settlement decision making.

Recently, the use of semiautomated techniques has proven useful for such island-scale surveys (Glover et al., 2020), and the use of automated feature extraction techniques offers another mechanism to scale up these analyses even more (Freeland et al., 2016). Machine learning and, specifically, deep learning hold the potential to increase the efficiency and success of lidar analysis across Oceania. Deep learning through convolutional neural networks (CNNs) has been applied with varying levels of success in other regions around the

world (e.g., Agapiou et al., 2021; Bonhage et al., 2021; Davis et al., 2021; Guyot et al., 2021; Somrak et al., 2020; Soroush et al., 2020; Trier et al., 2019; Verschoof-van der Vaart et al., 2020), and some approaches have successfully identified features even when training datasets are minimal (e.g., Davis et al., 2021). CNNs work by taking inputs from tensors (multidimensional matrices) to quantify multidirectional patterns, which allow neighbouring pixels to influence identifications (Caspari & Crespo, 2019). With the development of more sophisticated CNN architecture (e.g., Mask R-CNN), these methods can be used not only to identify objects in image data but also segment them (i.e., digitize the object's shape). As such, these methods provide the capacity to expand not only archaeological prospection efforts but also larger studies of sociopolitical organization, population densities and (in)equality, among others that rely on feature morphology and spatial distributions.

Agricultural terraces, more specifically, have been successfully identified in lidar data using semiautomated approaches in different places, globally (e.g., Capolupo et al., 2018; Duarte et al., 2018). More recently, the use of automated detection for terrace detection has been applied with success in places like Peru, where researchers used

deep learning to identify architectural elements throughout the 900-ha Kualep Complex (Righetti et al., 2021). Similarly, Herrault et al. (2021) compared traditional machine learning (random forest) to deep learning methods (fully connected network [FCN]) to identify agricultural terraces in Germany. They find that while random forests appeared to perform better than FCNs, deep learning can adapt to previously unseen data even with limited available training examples and remains minimally affected by certain labelling errors. Thus, deep learning has great potential for identifying archaeological features in lidar data, like terraces, across landscape scales.

To our knowledge, CNNs have only recently been applied to archaeological lidar analysis in New Zealand (see Bickler & Jones, 2021), but elsewhere in the Pacific, these methods have not yet been introduced. Here, we demonstrate the utility of deep learning approaches at a large spatial scale by producing a useful model to allow for the detection of terraces using training datasets from across American Sāmoa. We show how this model can be successfully employed on Polynesian high islands, using an island-wide survey of Tutuila as a case study, increasing its value and building on prior work using manual feature extraction (Cochrane & Mills, 2018). This is important as survey coverage across Sāmoa and other high islands in the region remains far less than ideal because of the cost and time commitment. Finally, we highlight how the segmentation of features enabled through Mask R-CNN models, rather than just merely identification, lends itself to further analysis.

### 3 | METHODS

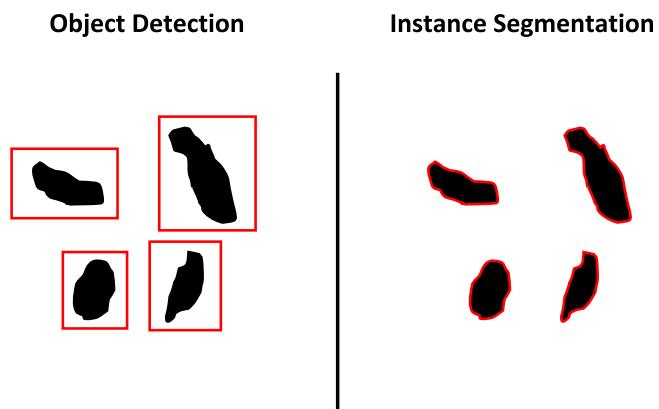
The primary lidar data used in this study to create and test the Mask R-CNN model derives from the National Oceanic and Atmospheric Administration (NOAA). Data collection was performed in 2012 by Photo Science, Inc. (Raber, 2012) using a Beechcraft King Air 90 twin-engine aircraft outfitted with an Optech Gemini sensor for all islands of American Sāmoa. Data were manipulated originally in Optech Software (GeoCue, TerraScan and TerraModeler) where they were also classified using project-specific macros. Lidar was collected with an average point spacing of 0.838 m and an average point density of 1.43 pts/m<sup>2</sup> over 108 flight lines between June 2012 and July 2012. The RMSEz for Tutuila was calculated at 0.067 m while metadata indicates that collection was undertaken to meet a vertical accuracy of 0.15 m and a horizontal accuracy of 1.2 m or better. We use publicly available DEMs produced using the aforementioned lidar data by NOAA's Office for Coastal Management (OCM) with a resolution of 1 m (OCM, 2022). Additional information can be found at OCM (2022).

We also make more limited use of an additional dataset to assess the generalizability and applicability of the model trained in American Sāmoa. This dataset was collected by the Samoan government between July and August 2015 for the Ministry of Natural Resources and Environment (MNRE) by Fugro Geospatial Services using a RIEGL LMS-Q780 Lidar system fitted onto a AusJet Cessna 441. The survey occurred at a height of 650 m and a maximum speed of 130–140

knots. Data collection used a laser rate of 350 kHz and a line spacing of 423 m, with a minimum pulse density of 4 points/m<sup>2</sup>. Due to cloud cover and flying difficulties at high elevations, data gaps occur. Data were collected to produce a minimum vertical accuracy of 0.30 m and a minimum horizontal accuracy of 0.80 m. A lidar-derived bare earth DEM was created in ArcGIS Pro (ESRI, 2021) using the LAS to Raster tool. The tool input used a binning interpolation type, with an average cell assignment and linear void fill method. The output cell size was 1 m. This dataset was not used to train the model; rather, it was used only to evaluate whether the model could identify terracing in lidar datasets produced using different methods relative to that produced in American Sāmoa.

Using these lidar-derived DEMs, we created a series of raster visualizations to help accentuate terracing features and topographic anomalies on the landscape. We tested a variety of visualizations, but the best performance was achieved using a three-band composite consisting of Terrain Ruggedness Index (TRI; Riley et al., 1999), Positive Openness and Slope. We assessed visualization performance by comparing CNN model accuracy scores and training loss values alongside manual evaluation of CNN predictions. TRI is a mathematical representation of topographic heterogeneity, which can help identify topographic modifications. TRI calculates the change in elevation between a pixel and its surrounding eight neighbours (Riley et al., 1999). Researchers have successfully used TRI (derived from LiDAR and other sensors) to characterize terrain and geomorphological properties like landslides and soil carbon content (Sharma et al., 2021), including as a parameter in CNN applications (Rózycka et al., 2017). We calculated TRI using SAGA v.7.9.1 (Conrad et al., 2015) with the following parameters: Search Mode = Circle; Search Radius = 1; No Distance Weighting. Openness is a measure of the angular relationship between surface relief and horizontal distance, which express the degree of enclosure or dominance of a location on an irregular plane (Yokoyama et al., 2002). There are two forms of openness: Positive openness, which emphasizes convexity, and negative openness, which emphasizes concavity. Here, we used SAGA v. 7.8.2 (Conrad et al., 2015) to calculate positive openness with the following parameters: Radial Limit = 1000; Method = Line Tracing; Multi Scale Factor = 3; Number of Sectors = 8. Finally, slope was calculated within ArcGIS Pro. Once all of these visualizations were made, we merged the three datasets together using the raster merge tool in ArcGIS Pro and saved the file as an 8-bit image for further analysis using deep learning.

ArcGIS Pro contains built-in libraries for deep learning, which have been used successfully for archaeological applications (e.g., Agapiou et al., 2021; Bickler & Jones, 2021; Davis et al., 2021; Davis & Lundin, 2021). To implement deep learning within the ArcGIS Pro environment, input data must be a multiband raster. Here, we use ArcGIS Pro v. 2.8.1 (ESRI, 2021) to train a Mask R-CNN model (He et al., 2017). We chose this algorithm both for its proven performance in other case studies with limited and extensive datasets (e.g., Davis et al., 2021), and its ability to simultaneously perform detection and segmentation tasks related to object recognition (Figure 2). As such, our output provides both coordinates of terracing features and their



**FIGURE 2** Illustration of traditional object detection compared with instance segmentation. Mask R-CNN's can perform instance segmentation, whereby the exact boundaries of identified objects are outlined (in red). Traditional object detection usually relies on bounding boxes, wherein objects are detected (red boxes) but morphological information is not provided. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

spatial dimensions, which can be used for further spatial analyses of agricultural and residential activities in Sāmoa.

Our training data consisted of 1254 preidentified terrace structures on the four main islands of American Sāmoa (Tutuila, Ta'u, Olosga and Ofu). The large majority of these are from the Manu'a group with a smaller, dispersed sample from Tutuila. We created training data in ArcGIS Pro using the Label Training Data for Deep Learning Analysis tool. Training data were created in Mask R-CNN format with a tile size of  $150 \times 150$  pixels and a stride size of  $75 \times 75$  pixels. Optimal tile sizes were determined by trial and error. During this process, additional images are created that contain areas without terracing, which can happen as a byproduct of chosen window sizes when some terraces get cut off. This resulted in the use of 1438 separate images by the model. The inclusion of blank images can also help train the algorithm to recognize regions with and without features of interest, and future work can include additional classes containing nonterrace features that are commonly misclassified (e.g., Davis et al., 2019, 2021).

We trained a Mask R-CNN model with an unfrozen resnet152 backbone model and a batch size of 8. The model was set to train for 75 epochs (or until improvements stopped). Following model training, we used the Detect Objects using Deep Learning Tool to identify terracing features on Tutuila. We used the following parameters: padding = 0; batch size = 8; threshold = 0.05; Return\_bboxes = False; Tile\_Size = 400, 500, 600; Non-Maximum Suppression with a max overlap ratio = 0.15. Padding designates a region around the edge of the detection window where the model will not identify features. The batch size refers to the number of images the model processes at a given time. Residual Network (ResNet) (He et al., 2017) is a transfer learning architecture, which is a process where previously trained models can serve as a baseline for training new models, even when the target of these models is different. ResNet152 is trained using the ImageNet dataset (consisting of over 1 million images) with

152 convolutional layers. We trained the model on an unfrozen ResNet model because there were no prior models trained for archaeological terrace detection from this area. The threshold refers to the confidence score at which the model will return positive detections, and the “return\_boxes” parameter will draw bounding boxes around detected features when True (the default) and will create segmented feature outlines when False. The optimal learning rate was calculated using a learning finder (see Smith, 2017) and 10% of the training data was withheld from training to help validate the model performance prior to applying the model across Tutuila. All deep learning analysis was conducted on a computer with a NVIDIA Quadro p4000 GPU, an Intel® Core™ i7-7700K CPU @ 4.20 GHz, 4200 MHz, 4 Core(s), 8 Logical Processor(s) and 64 GB of RAM.

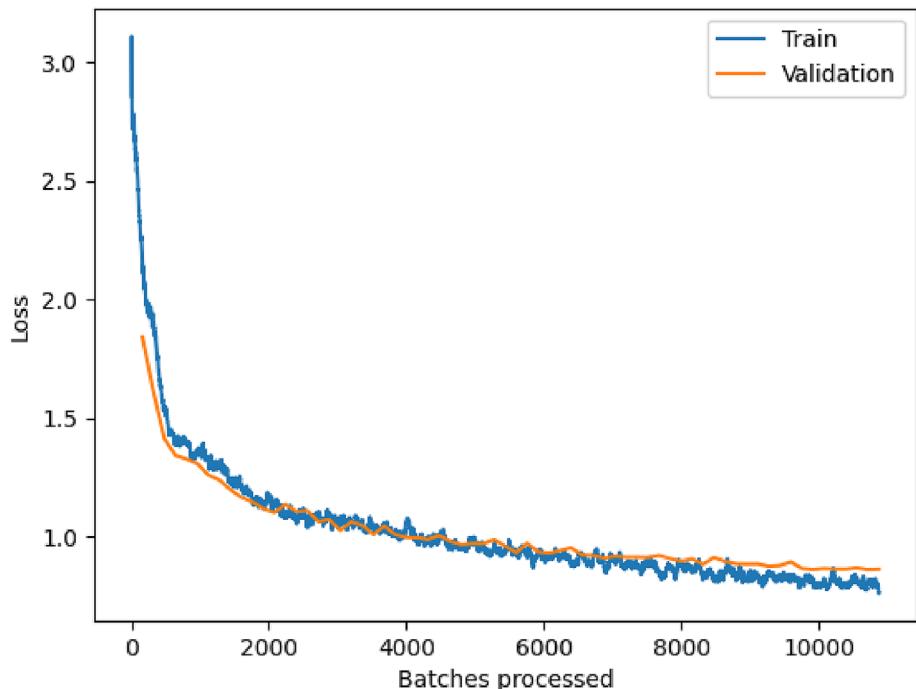
Areas of contemporary development are well demarcated on Tutuila. While we ran the model over these regions, all detections within areas of modern development were manually removed by comparing feature identifications to the World Imagery base map in ArcGIS Pro. We do note that while the differences between modern and past geomorphic engineering are visually distinctive (e.g., sharpness of corners), modern terraces were consistently identified and a small number may be retained in our results. We ran this detection algorithm three separate times with tile sizes of 400, 500 and 600 and combined the results to evaluate performance (Table 1). By using multiple thresholds, we maximized our true positives while minimizing false negatives. The combination of these results was undertaken using the Merge tool. This was followed by the Aggregate Polygons tool to remove repetitive identifications. This procedure produces the largest potential boundaries for each identification. Finally, all identifications under  $5 \text{ m}^2$ , a threshold based on known sizes of terraces, and some feature vertices were edited to remove extraneous components. Subsequent spatial analyses were performed in ArcGIS Pro using the Geoprocessing toolbox.

## 4 | RESULTS AND ANALYSIS

The results of model training and performance are illustrated in Figure 3. The model ran for 68 epochs with an optimized learning rate of  $3.63078 \times 10^{-6}$  and over 7000 terraces were identified after data cleaning. The efficacy of the model at the scale of individual features was tested against three maps, georeferenced in ArcGIS Pro, across different topographical settings on Tutuila in Tatagamatau (combination of south-facing hillslope and ridgeline), Fagasa (north-facing and narrow primary and east-facing secondary ridgeline) and Malaeimi (along a single south-facing secondary ridgeline) (Figure 5; see Best, 1993). Maps produced for these sites show the location and boundaries of individual terraces, which have proven effective for past assessments of lidar feature extraction (Cochrane & Mills, 2018). While we did document slight but clear spatial discrepancies in these maps and our results, even when it was clear the same feature was represented, they represent the lone mechanism by which we can evaluate our results at this spatial scale. As noted by Cochrane and Mills (2018), some clear terraces are missing from the pedestrian maps

**TABLE 1** Performance of Mask R-CNN model judged using previously developed maps. Note the better performance of the combined dataset relative to those produced by single window sizes.

Window	True positives	False positives	False negatives	Precision	Recall	F1 statistic
Combined	80	4	59	0.576	0.952	0.718
400 m	67	2	72	0.482	0.971	0.644
600 m	67	2	72	0.482	0.971	0.644
500 m	60	4	79	0.432	0.938	0.592

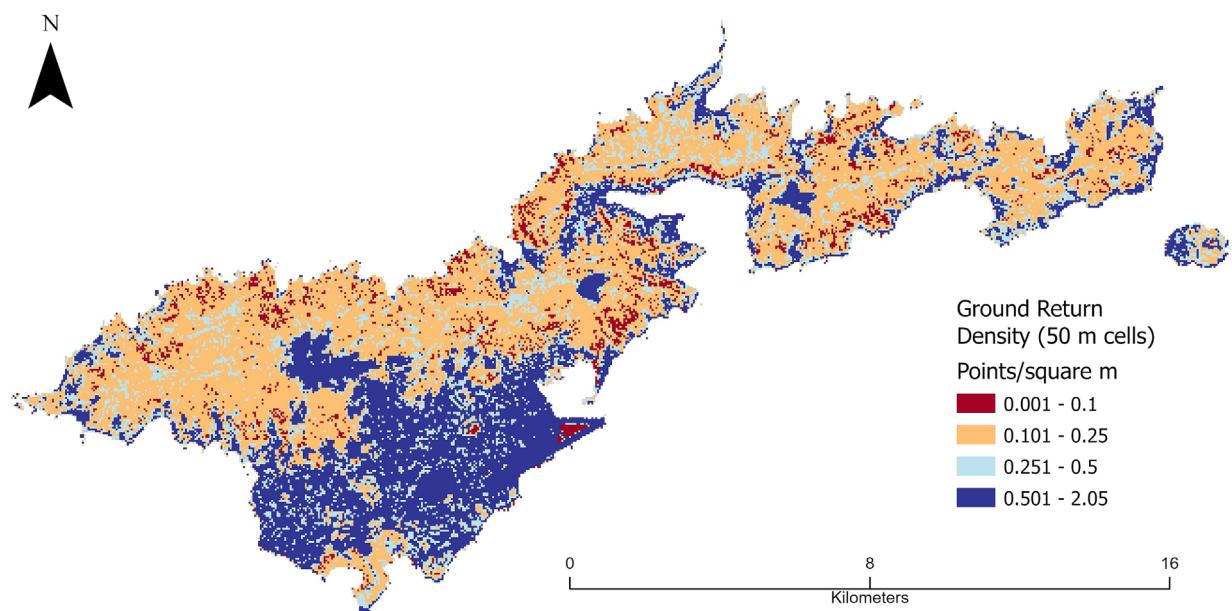


**FIGURE 3** Results of model training and performance. The two curve loss values associated with the training and validation datasets, where the lower the loss, the better the model performance. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

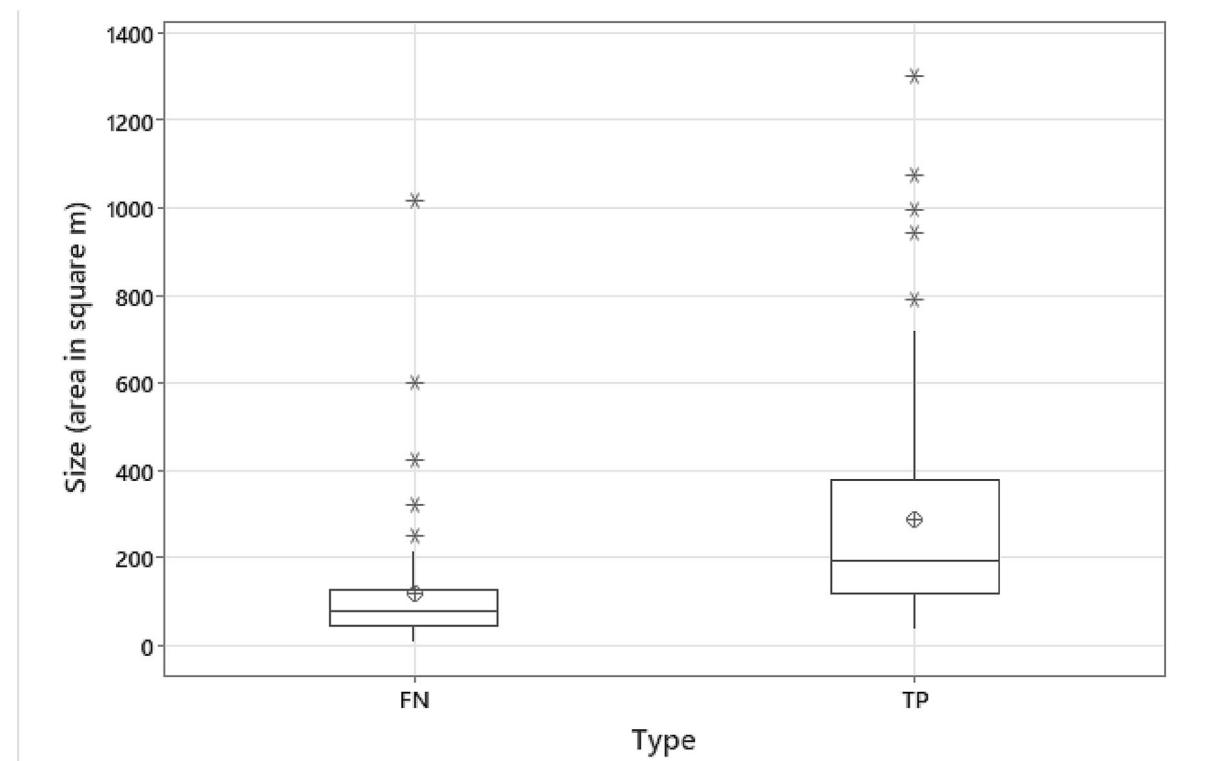
of Tatagamatau and Fagasa. As such, care should be taken in interpreting false positive results. Results are presented in Table 1.

Our results indicate that the model is conservative but effective. Not surprisingly, the density of ground points impacted the success of the model. An average ground point density of less than 0.5 pt/m<sup>2</sup> across areas of the interior uplands (see Figure 4) reduces the model's ability to locate some kinds of features, as has been previously noted elsewhere in the world (Sánchez Díaz & García Sanjuán, 2022). In our case, low ground return density affects the visibility of small features, which are known to be difficult to document in the archipelago, even using manual extraction techniques (Quintus et al., 2017). The effect of feature size on feature identification is indicated by the significant difference in mean size between true positives (median = 193 m<sup>2</sup>) and false negatives (median = 75 m<sup>2</sup>) (see Figure 5; Kruskal-Wallis Test;  $H$  value 38.79;  $p < 0.001$ ). The small size of the feature results in a lower likelihood that ground returns, especially multiple ground returns, derive from the terrace surface, reducing the contrast between the feature and the surrounding slope. This results in a "fuzziness" to the feature comparable to degraded mounds that blend into the landscape because of erosion (see Forest et al., 2020). Larger feature size creates a higher likelihood that multiple ground returns will derive from the feature, making it more likely that the feature will

contrast with surrounding slopes. As such, the model is quite useful in providing a minimum number of terraces in a particular area and seems to capture larger terraces effectively. Topographic setting may also influence the visibility of terracing in this lidar dataset, with those features on hillslopes more likely to be true positives as opposed to those on ridgelines (68% vs. 51%;  $\chi^2 = 4.075$ ;  $p = 0.044$ ). However, this may also be due to terrace size as terraces on ridgelines are smaller than those on hillslopes (median = 105 m<sup>2</sup> vs. 166.5 m<sup>2</sup>;  $H$  value 7.27;  $p = 0.007$ ). We also assessed our results at a more general level, based on the location of known concentrations of terraces, though these concentrations have not been mapped in detail. In each case, the Mask R-CNN model performs well in identifying the location of known concentrations in the context of the wider landscape (Figure 6). Thus, the model performs well at both the individual feature and site level, with some exceptions based on terrace size, even though the mean density of ground returns is below 0.5 points/m<sup>2</sup> for some areas of the interior uplands. As demonstrated in Figure 4, the majority of the study area contains lidar coverage with 0.1–2.5 points/m<sup>2</sup>. When considering the median terrace size of 135 m<sup>2</sup> in the ground truthed dataset, the average terrace contains at least 13 lidar return points. This number of return points provides enough detail to identify the outlines and basic topographic properties of



**FIGURE 4** The density of ground points across the island of Tutuila. The interior features a mix of areas with ground return densities above and below 0.5 points/m<sup>2</sup>. The triangular shape in the lower centre of the figure is the modern runway. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

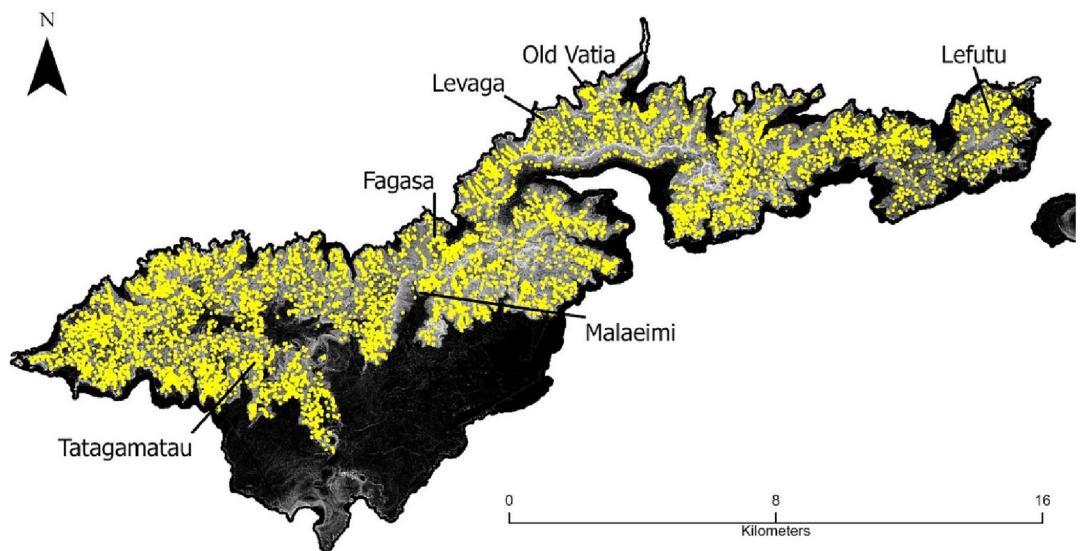


**FIGURE 5** Comparison of terrace size and types of feature identifications. FN = false negative; TP = true positive.

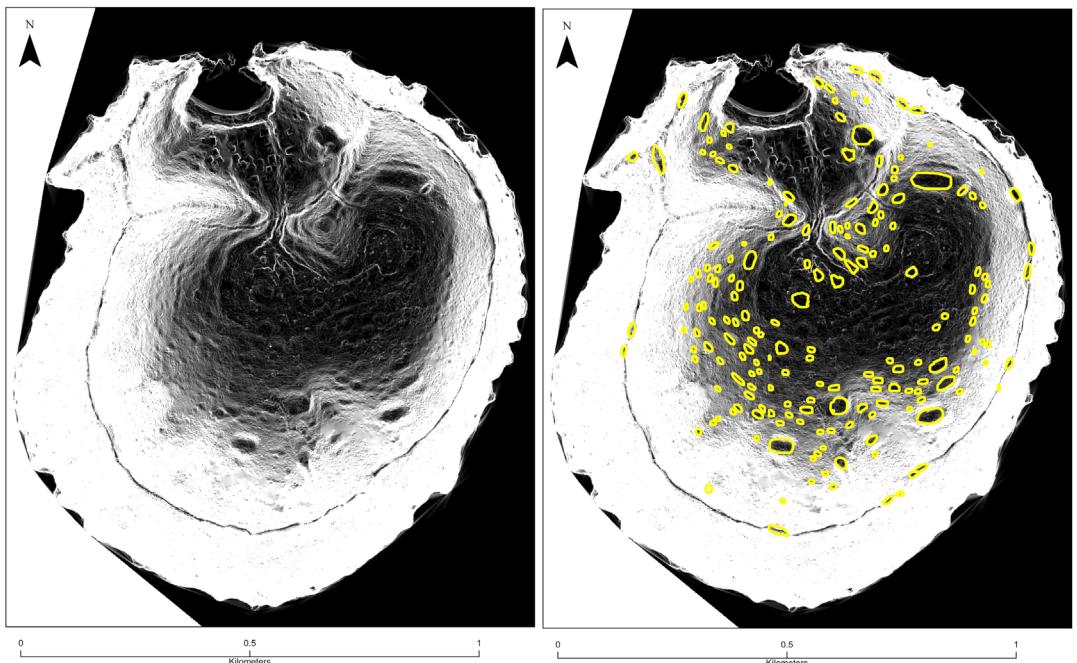
moderate and large-sized terraces, as our results confirm. Importantly, though, small terracing does not contrast with the surrounding slopes at this data resolution.

The model likely performs best where the general slope gradient is steepest, as it is in those environments where the constructed

terrace contrasts markedly with the surrounding slope (McCoy et al., 2011). All of the lidar derivatives (i.e., slope, roughness and openness) we used as inputs highlight those contrasts. While we currently lack pedestrian data to evaluate model performance in flat ground, the characteristics of the input combined with visual



**FIGURE 6** Locations of previously documented terrace concentrations (named locations) compared with terraces identified during this project, outlined in yellow. The boundaries of these sites have never been clearly published and our results suggest they are part of relatively continuous distributions of terraces. [Colour figure can be viewed at [wileyonlinelibrary.com](https://wileyonlinelibrary.com)]



**FIGURE 7** Terrace identifications on Apolima using the Mask R-CNN produced using training data from American Sāmoa. The yellow polygons in the image on the right demarcate potential terraces. The island, generally, has a gentler topography relative to Tutuila, and lidar data were collected at a different time using different equipment. Darker colours signify lower slope gradients. [Colour figure can be viewed at [wileyonlinelibrary.com](https://wileyonlinelibrary.com)]

inspections suggest identifications in lower slope locations are more suspect. This is especially the case of contiguous areas of less than 5° slope, of which there are few in the interior uplands. However, there is a noticeably large area located in the island's western uplands. The training datasets lack terraces from such gradual slopes. While terracing is far less useful in these low-slope environments, data from adjacent islands in American Sāmoa does indicate that a small proportion of terraces are found in these locations (Quintus et al., 2022).

#### 4.1 | Preliminary test of model applicability

The utility of the model for regions outside of American Sāmoa is demonstrated by applications to the Nation of Sāmoa. Using the same model trained on data from American Sāmoa, we successfully identified terraces across the entire island of Apolima located ~200 km away from our original training and test areas (Figure 7). The method provides a representation of terracing across the

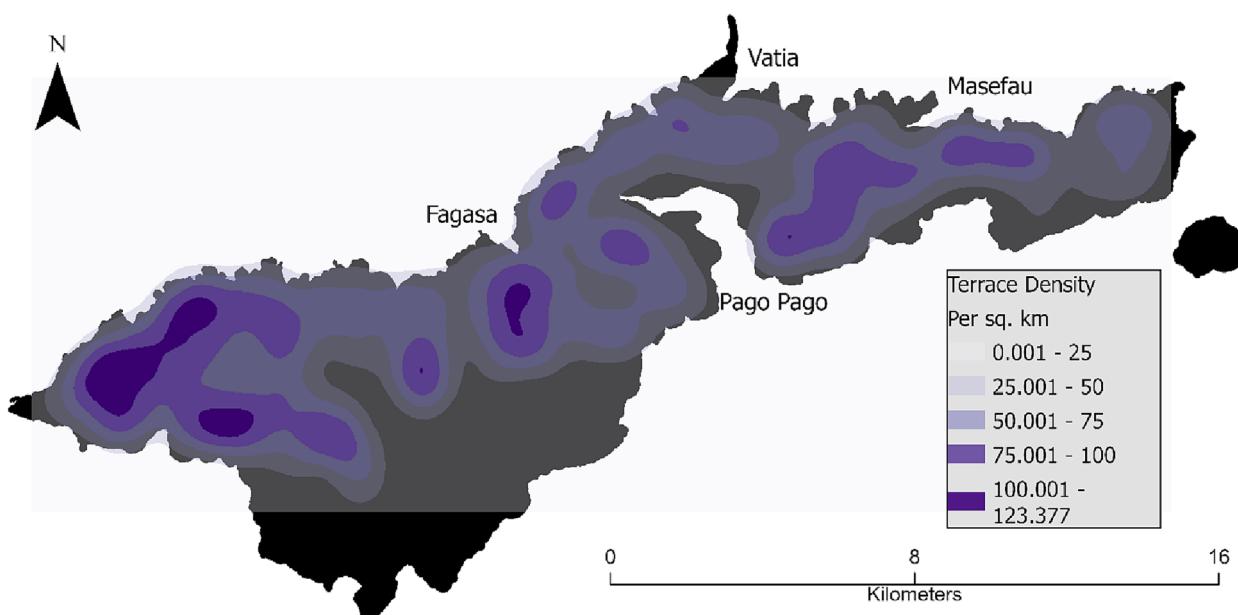
island's landscape despite not retraining the model with new data. While still a part of the Samoan archipelago, Apolima is unique given its gentler topography and small size. The methods of lidar data collection were also distinct (see methods above), which implies the model is not overfitted to the methods of data acquisition used in American Sāmoa. While our investigation is ongoing, and ground evaluation remains to be conducted, this provides greater evidence for the method's applicability beyond the bounds of our initial training area and lidar data collection. This is something that few other automated archaeological remote sensing studies have managed to achieve.

#### 4.2 | The nature and distribution of terracing across Tutuila

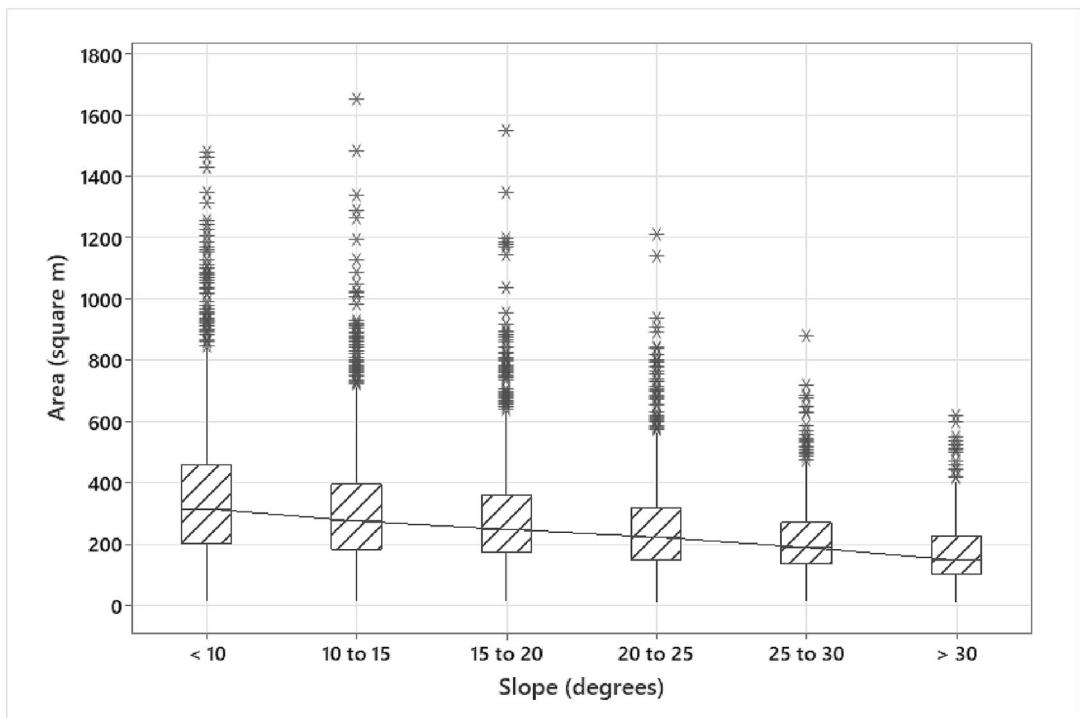
Our identifications allow us to assess the density of terracing across Tutuila's interior even though it is a conservative estimate. To facilitate analysis, we assume that factors that affect the identification of terracing are similar in similar topographic settings in different parts of the island; the identified features represent something of a stratified random sample of features, with sampling strata being topographic settings. Certainly, some topographic settings are more frequent in some parts of the island than others, which could affect our results. Differential ground returns could also influence our results, though this does not appear to be the case as terrace density is not correlated with ground return density ( $r^2 = 0.014$ ;  $F = 2.13$ ;  $p = 0.15$ ) nor is mean terrace size ( $r^2 = 0.007$ ;  $F = 1.04$ ;  $p = 0.309$ ). Finally, we remove from consideration terraces from areas of contiguous slope of less than  $5^\circ$  in evaluating terrace density given our lack of pedestrian data from these locations.

The density of identified terraces is uneven (Figure 8), with the highest concentrations located in the western and central third of the island. This is impacted by the distribution of suitable slopes. Prior research in Manu'a has shown that most of the terracing is located in areas of less than  $25^\circ$  slopes (Quintus et al., 2015, 2022). This holds true on Tutuila as well, though terracing is more frequent in slightly steeper slopes relative to Manu'a. The number of terraces built in an area declines substantially with an increased slope gradient; 87% of features were built in areas with slopes of less than  $25^\circ$ . Such slopes, along with access to more expansive agricultural lands, make the western side of the island more suitable for expansive human engineering through time. Furthermore, terrace size decreases as the slope gradient increases (Figure 9). This implies a real technological or labour constraint on the construction of terraces, which leads to fewer and smaller features being constructed in less suitable topographic settings. We posit that this likely relates to the fact that more fill material would be needed to increase the width of terracing on steep slopes, which requires more labour.

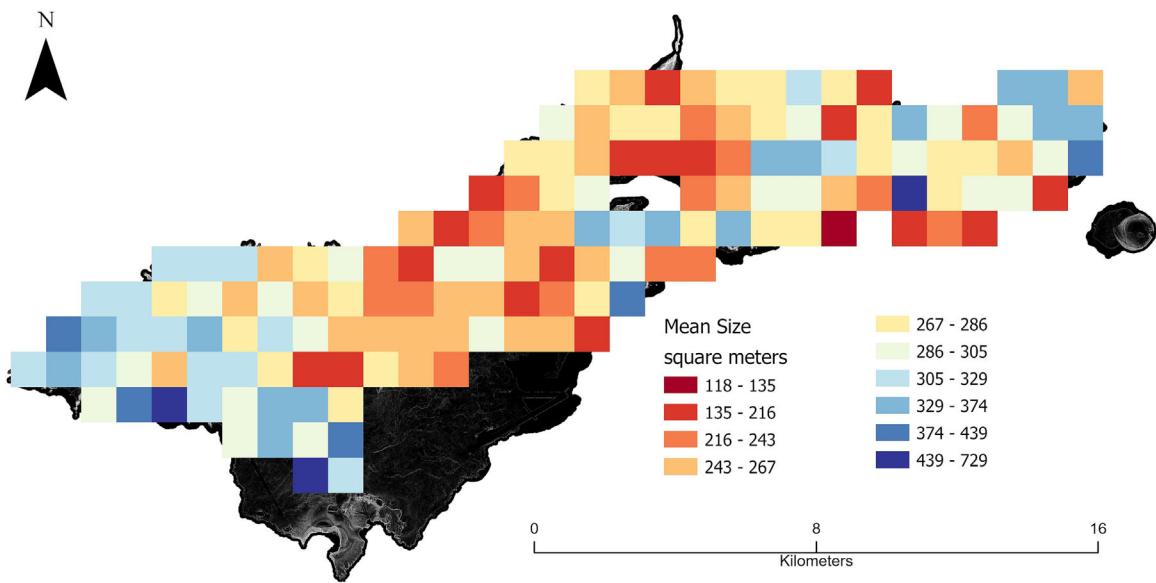
The size of terracing on Tutuila is comparable to terraces in the Tamatupu site on Olosega island, though the mean ( $290 \text{ m}^2$ ) and median ( $251 \text{ m}^2$ ) terrace sizes are different on Tutuila compared with other sites in Manu'a (Quintus, 2020). It is likely that the actual mean and median terrace size is slightly lower on Tutuila as we do not identify all small features. The mean size of terracing is uneven across the island (Figure 10). In our dataset, terracing is clustered by size (Moran's  $I$ ; Moran's index = 0.0603; Expected = -0.000135;  $z$  score = 15.29;  $p = 0.00$ ). However, relatively few features on Tutuila are part of these statistical hot spots (9%), drawing a distinction with the Manu'a group (Quintus et al., 2022). Many of the clusters identified by Moran's  $I$  test are in the western and central thirds of the island with only one in the eastern third. This lone cluster is the



**FIGURE 8** Density of terracing across the island of Tutuila. [Colour figure can be viewed at [wileyonlinelibrary.com](https://wileyonlinelibrary.com)]



**FIGURE 9** The relationship between terrace size and slope.



**FIGURE 10** Mean terrace size across Tutuila in m<sup>2</sup>. Each cell is 1 km<sup>2</sup>. Each cell average is based on a different number of terraces located within each cell. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

previously documented site of Lefutu, which has long been considered unique in the area (Clark & Herdrich, 1993; Pearl, 2004). Some clusters of large features are located near contemporary infrastructure, and it is possible that at least some larger terraces are historic but overgrown (i.e., Fagasa). Significant clusters of small terraces (cold spots) are noticeably rarer and smaller. Most terraces are not

clustered by size at the island-wide scale, highlighting the dispersed nature of settlement. Even when clustered, the size of such clusters is substantially smaller than that documented for Manu'a where the size of nucleated settlements is conditioned by the small size of the islands and the gradual interior slopes of the uplands (Quintus et al., 2022).

#### 4.3 | Terracing and defensive infrastructure

Fortifications are one of the most marked elements of the interior landscape of Tutuila. These defensive features have been identified across the island (Best, 1993; Clark & Herdrich, 1993; Cochrane & Mills, 2018), but the full distribution of defensive features has not been documented. Terracing is a key component of these defensive features, along with associated infrastructure like ditching and banks, and examination of our results presents an unanticipated opportunity to highlight the distribution of fortifications at the island scale. While not all defensive features are identifiable in our results and, therefore, we cannot yet examine the spatial distribution of these fortifications in detail, we are able to better understand the morphological variability (Figure 11a–c) and relative density across the island thanks to the data generated by our Mask R-CNN model.

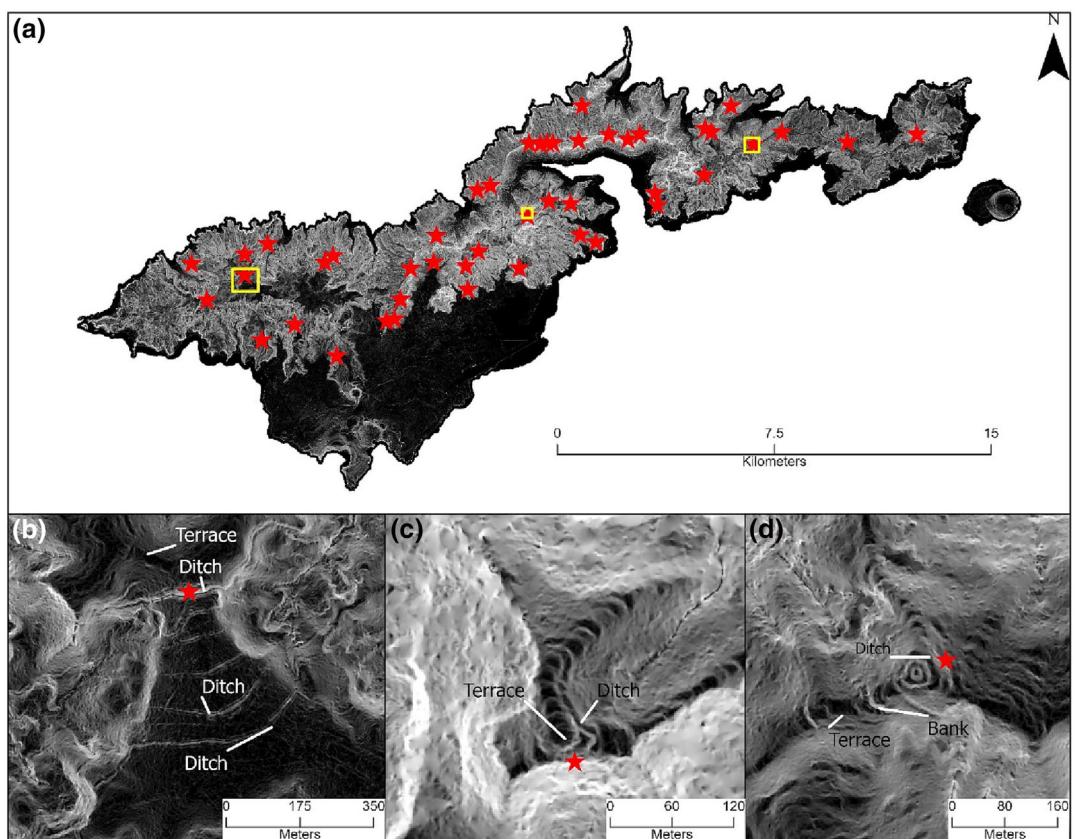
We identified 46 complexes that appear to have some defensive functionality (Figure 11), defined as features that prohibit access to some location. Not surprisingly, most visible defensive features are found in the centre and western thirds of Tutuila, with few large-scale defensive features identifiable in the east. In contrast, there are clear concentrations of defensive earthworks around Masefau Bay, as noted in Cochrane and Mills (2018), as well as overlooking Pago Pago

Harbor. That such defensive features would be built in these locations is unsurprising as these bays represent two of the largest on the island. The largest fortification on the island, however, is located further west, constituted by expansive terracing and at least six ditches and banks. The primary ditches and banks cover some 13 ha with somewhat nucleated terracing found within a 37-ha area upslope of these features (Figure 11b).

#### 5 | DISCUSSION

Lidar datasets are becoming increasingly available across Oceania as island nations seek to gather data by which to document the effects of climate change. While not their primary purpose, these datasets are useful for archaeologists (Bedford et al., 2018; Cochrane & Mills, 2018; Freeland et al., 2016; Parton & Clark, 2022; Quintus et al., 2015, 2017). The challenge for archaeologists is to generate methods that efficiently and accurately allow for the extraction of useful archaeological information.

Here, we have trained a deep learning model using a Mask R-CNN architecture to extract the location and morphology of archaeological terracing features. We demonstrate that this model is



**FIGURE 11** (a) The distribution of identified likely fortifications across Tutuila. Red stars are likely defensive features while the yellow rectangles define the extent of the inserts. From left to right on the top image: (b) The largest fortification identified with a series of at least six ditches located downslope of and defending a group of terraces. (c) A fortification in central Tutuila constituted by several terraces and a few ditches. (d) A fortification in eastern Tutuila with terracing, banks and ditching protecting the intersection of several ridgelines. [Colour figure can be viewed at [wileyonlinelibrary.com](http://wileyonlinelibrary.com)]

effective, judging from comparison with previously mapped sites that included terracing. However, limitations are also documented relating to the nature of the lidar dataset and the features themselves, specifically uneven ground point density and feature size. Even with these constraints, we expect that this model will be broadly useful as a tool to supplement more targeted pedestrian surveys, providing a fuller picture of land use in these relatively inaccessible areas.

On Tutuila, our results demonstrate large-scale manipulation of slopes, further highlighting and building upon previously documented expansive interior land use (Clark & Herdrich, 1993; Cochrane & Mills, 2018; Day, 2018). These data illustrate shared patterns across the islands of American Sāmoa. Terracing is the dominant archaeological feature class in the uplands of these islands, which clearly relates to the constraints of living in a high island environment. Dating of terraces in American Sāmoa highlights the persistent utility of the technology, but also the use of terraces prior to the settlement of East Polynesia (see Carson, 2006; Quintus et al., 2020), indicating that terracing was part of the Polynesian transported landscape (after Kirch, 1982). Furthermore, the size of terracing on Tutuila is similar to that on Olosega, though mean terrace size is larger than on the other islands of Manu'a. The somewhat consistent size combined with the illustrated limiting factor of slope indicates that communities across the Samoan archipelago dealt with similar technological challenges.

Still, several elements of the nature and patterns of interior land use on Tutuila are dissimilar to that documented for the islands of Manu'a (Quintus, 2020; Quintus et al., 2015, 2022). Terracing and settlement, more generally, are more nucleated in Manu'a relative to Tutuila, with more defensive infrastructure present in the latter case. The documentation of a larger number of likely defensive features on Tutuila is consistent with results also from neighbouring Tonga (Parton et al., 2018). Terracing is far more frequent along ridgelines on Tutuila compared with the situation in Manu'a. These patterns speak to the different interior configurations of the islands. Relatively contiguous slopes under 20° are rare on Tutuila. In contrast, coastal plains were more developed with some deep and productive bays, which do not exist in Manu'a. This, we hypothesize, led to an increased focus on coastal settlement on Tutuila with more targeted and generally less intensive residential use of the interior uplands. Within this context wherein settlement was focused on valleys with productive bays, defensive features on ridgelines that border these deep harbours are useful, though defensive infrastructure is also found associated with basalt extraction sites. Such defensive features are less useful and rarer in Manu'a where second millennium AD settlement seems more concentrated in the interior uplands. These locations are naturally fortified by remnants of former sea cliffs.

We anticipate that future research may be able to use our data to better understand the drivers of morphological variation across defensive features. There is no doubt that some of this variability is caused by topographic differences. However, we also expect that other factors are contributing, such as viewshed. Capturing the size and morphology of these features, as accomplished using Mask R-CNN models, allows assessment of these questions. The trained model developed here is provided as a supplemental file to aid researchers in

this endeavour and can provide a baseline for future studies focusing on other areas. Pairing this dataset with targeted field investigation is certainly desirable.

## 5.1 | Limitations of our study and AFE

While the model is effective, challenges were also experienced. Differential ground point density, feature size and the degree of slope all influence the likelihood that true features will be identified. Models similar to the one used here are less effective where features are consistently under 100 m<sup>2</sup> and where features were built on slopes under 10°. Low ground return density is also problematic, especially in terms of identifying small features, as the low number of ground returns that derive from the feature surface reduce the contrast with the slope. Within this context, one reason for our success is the generally large size of terraces in Sāmoa. Furthermore, we noted the presence of artificial boundaries for some positive identifications that related to the position of the window during analysis. This created straight edges on the polygon, cutting the size of the terrace down or creating two identifications for a single feature. Using multiscale windows helped ease some of these issues (also see Guyot et al., 2018), but also increased manual data-cleaning needs. To increase the utility of these data for further spatial analysis, as was the goal here, manual cleaning was necessary to correct some errors in drawn feature boundaries. Such manual processing was also needed to eliminate areas of contemporary development as modern villages in Samoa make use of terraces. These terraces tend to be morphologically distinct, with sharper edges caused by the use of heavy machinery, but these still need to be manually removed. This sort of cleaning process is typical in many AFE studies (Davis et al., 2019; Meyer et al., 2019; Meyer-Heß, 2020). Data cleaning required roughly 2 weeks of additional work. Even with data cleaning, the use of automation reduced substantially the amount of work necessary to produce the dataset relative to if we had used manual extraction (cf. Quintus et al., 2017), reducing identification and digitizing tasks from a months-to-years long endeavour to a couple of weeks.

The identification of modern infrastructure highlights another constraint. These interior uplands are palimpsests in which modern terraces are the last layer added to an already layered landscape (see chronological data for other areas in American Sāmoa in Quintus et al., 2022). Unlike modern terraces, though, the terracing from earlier time periods is not morphologically distinct enough to differentiate them. This poses certain challenges in interpreting demographic and settlement patterns, as other researchers have noted (Grammer et al., 2017; Henry et al., 2019). Because of this, the pairing of remotely identified features with more targeted pedestrian survey, test excavation and historical analysis is important (Johnson & Ouimet, 2018; Quintus et al., 2022; Sugiyama et al., 2021). The incrementally accumulative nature of this form of landscape engineering gives the impression of a fuller landscape than may ever have existed at one time, though the extent of modification is still a robust measure of cumulative human impacts. Such palimpsests can also provide

estimates of which locations of a landscape were most frequently and intensively occupied over time, as well as those areas where habitation or other activities were sporadic or avoided (Freeland et al., 2016; Ladefoged et al., 2011; Sugiyama et al., 2021). Documentation of feature locations, and the nature of features across space, is also a first step in generating robust models of settlement and demographic change (Carter et al., 2018; Klassen et al., 2021; Ladefoged et al., 2011) as well as in tracking archaeological preservation and stewardship (Sugiyama et al., 2021).

It should be noted that many of these limitations are not unique to the Pacific and are issues all archaeologists must contend with. Ultimately, we must use the data available to us, and oftentimes these data are not complete or ideal. Nonetheless, what this study demonstrates is that methods can still be deployed that can effectively extrapolate archaeological information from datasets with mixed levels of quality. Ultimately, all AFE studies will be incomplete, as it is impossible (and often unnecessary) to identify every single feature in an area. To the contrary, many studies can lead to significant advances in archaeological knowledge even without high levels of accuracy (see Arnoldussen et al., 2022; Verschoof-van der Vaart & Lambers, 2022).

Our results, despite lower quality lidar data coverage in certain areas, will greatly enhance our knowledge of terracing and settlement activities in Samoa. Archaeological research in Samoa remains limited. The availability of lidar datasets, even if not ideal, and the potential to use AFE methods allows a substantially more robust examination of the archaeological record than would otherwise be feasible. While ground return densities are low in many areas of the island, our approach can identify larger and well-defined terracing features with a relatively high degree of precision and accuracy, and this has great utility because it can expedite ground visits and further archaeological study in otherwise hard to reach places where surveys are difficult. Furthermore, our study provides data on terrace size and distribution that can be used to direct future hypothesis-driven research exploring demographic trends, the rise of territoriality and agricultural development.

## 6 | CONCLUSIONS

The use of Mask R-CNN models shows significant promise for generating accurate morphological and spatial information pertaining to archaeological features that go beyond simple prospection efforts. The ability of these models to segment identifications allowed us to attain accurate estimates of feature size, area and clustering patterns, which are not achievable (without considerable mathematical manipulation; e.g., Verschoof-van der Vaart et al., 2022) using other models that simply provide a bounding box around identified features (Figure 1). The generation of fully analysable datasets of archaeological feature locations and morphology, therefore, permits research that uses automated methods to go beyond detection tasks to address longstanding questions about the archaeological record, itself (*sensu* Davis, 2019). The use of AFE, and lidar more generally, is not a replacement for field-based studies (Sugiyama et al., 2021), but it is an

important tool to use, especially in topographic context that make field studies difficult and costly.

The significance of this model lies in the ubiquitous nature of the feature it extracts. Terracing is a dominant feature type in archaeology, both in and outside Oceania. Given the shared morphological characteristics of most terraces, we are confident that this model will be applicable to other places within and outside Oceania. Further, we demonstrate, building on prior research (e.g., Bonhage et al., 2021; Carter et al., 2021; Dolejš et al., 2020), the efficacy of Mask R-CNN models for archaeological prospection of lidar datasets. Given the cross-cultural importance of terracing, and the fact that these features are some of the most visible using lidar datasets (McCoy et al., 2011; Sánchez Díaz & García Sanjuán, 2022), we expect Mask R-CNN models trained with our data and supplemented with a small set of local features to provide an opportunity to document the nature and scale of human landscape modification efficiently and accurately across the globe. By applying deep learning to one of the most ubiquitous features in the region, which relate to diverse behaviours (i.e., defence, agriculture and habitation), such methods open additional avenues for comparative studies of subsistence practices, agricultural economies, sociopolitical organization and population dynamics across the Pacific and elsewhere around the world.

## CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to disclose.

## DATA AVAILABILITY STATEMENT

Training data and Mask R-CNN model are available at [10.5281/zenodo.7312103](https://doi.org/10.5281/zenodo.7312103). The .shp file of the cleaned identified terraces is also available as are the raw identified datasets for each window size.

## ORCID

Seth Quintus  <https://orcid.org/0000-0003-4388-3862>

Dylan S. Davis  <https://orcid.org/0000-0002-5783-3578>

Ethan E. Cochrane  <https://orcid.org/0000-0001-9027-0634>

## REFERENCES

- Acabado, S. B., Koller, J. M., Liu, C., Lauer, A., Farahani, A., Barreto-Tesoro, G., Reyes, M. C., Martin, J. A., & Peterson, J. A. (2019). The short history of the Ifugao rice terraces: A local response to the Spanish conquest. *Journal of Field Archaeology*, 44, 195–214. <https://doi.org/10.1080/00934690.2019.1574159>
- Addison, D. J. (2006). Feast or famine? Predictability, drought, density, and irrigation: the archaeology of agriculture in Marquesas Islands Valleys. Unpublished PhD Dissertation, Department of Anthropology, University of Hawai'i.
- Agapiou, A., Vionis, A., & Papantoniou, G. (2021). Detection of archaeological surface ceramics using deep learning image-based methods and very high-resolution UAV imageries. *Land*, 10, 1365. <https://doi.org/10.3390/land10121365>
- Allen, M. S. (2004). Bet-hedging strategies, agricultural change, and unpredictable environments: Historical development of dryland agriculture in Kona, Hawaii. *Journal of Anthropological Archaeology*, 23, 196–224. <https://doi.org/10.1016/j.jaa.2004.02.001>
- Arnoldussen, S., Verschoof-van der Vaart, W. B., Kaptijn, E., & Bourgeois, Q. P. J. (2022). Field systems and later prehistoric land use:

- New insights into land use detectability and palaeodemography in the Netherlands through LiDAR, automatic detection and traditional field data. *Archaeological Prospection*. <https://doi.org/10.1002/arp.1891>
- Bayliss-Smith, T. P., & Hvding, E. (2015). Landesque capital as an alternative to food storage in Melanesia: Irrigated taro terraces in New Georgia, Solomon Islands. *Environmental Archaeology*, 20, 425–436. <https://doi.org/10.1179/1749631414Y.0000000049>
- Bedford, S., Siméoni, P., & Lebot, V. (2018). The anthropogenic transformation of an island landscape: Evidence for agricultural development revealed by LiDAR on the island of Efate, Central Vanuatu, South-West Pacific. *Archaeology in Oceania*, 53, 1–14. <https://doi.org/10.1002/arco.5137>
- Bennett, R., Cowley, D., & De Laet, V. (2014). The data explosion: Tackling the taboo of automatic feature recognition in airborne survey data. *Antiquity*, 88(341), 896–905. <https://doi.org/10.1017/S0003598X00050766>
- Berganzo-Besga, I., Orengo, H. A., Lumbreras, F., Carrero-Pazos, M., Fonte, J., & Vilas-Estévez, B. (2021). Hybrid MSRM-based deep learning and multitemporal sentinel 2-based machine learning algorithm detects near 10k archaeological tumuli in North-Western Iberia. *Remote Sensing*, 13, 20. <https://doi.org/10.3390/rs13204181>
- Best, S. (1993). At the halls of the mountain kings. Fijian and Samoan fortifications: Comparison and analysis. *Journal of the Polynesian Society*, 102, 385–447.
- Bevan, A., Conolly, J., Colledge, S., Frederick, C., Palmer, C., Siddall, R., & Stellatou, A. (2013). The long-term ecology of agricultural terraces and enclosed fields from Antikythera, Greece. *Human Ecology*, 41, 255–272. <https://doi.org/10.1007/s10745-012-9552-x>
- Bickler, S. H. (2021). Machine learning arrives in archaeology. *Advances in Archaeological Practice*, 9, 186–191. <https://doi.org/10.1017/aap.2021.6>
- Bickler, S. H., & Jones, B. (2021). Scaling up deep learning to identify earthwork sites in Te Tai Tokerau, Northland, New Zealand. *Archaeology in New Zealand*, 64, 16–30.
- Bonhage, A., Eltaher, M., Raab, T., Breuß, M., Raab, A., & Schneider, A. (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. <https://doi.org/10.1002/arp.1806>
- Brown, A., Walsh, K., Fallu, D., Cucchiaro, A., & Tarolli, P. (2020). European agricultural terraces and lunchets: From archaeological theory to heritage management. *World Archaeology*, 52, 566–588. <https://doi.org/10.1080/00438243.2021.1891963>
- Câmara, A., de Almeida, A., Caçador, D., & Oliveira, J. (2023). Automated methods for image detection of cultural heritage: Overviews and perspectives. *Archaeological Prospection*, 30, 153–169. <https://doi.org/10.1002/arp.1883>
- Capolupo, A., Kooistra, L., & Boccia, L. (2018). A novel approach for detecting agricultural terraced landscapes from historical and contemporaneous photogrammetric aerial photos. *International Journal of Applied Earth Observation and Geoinformation*, 73, 800–810. <https://doi.org/10.1016/j.jag.2018.08.008>
- Carson, M. T. (2006). Samoan cultivation practices in archaeological perspective. *People and Culture in Oceania*, 22, 1–29.
- Carter, A., Heng, P., Stark, M., Chhay, R., & Evans, D. (2018). Urbanism and residential patterning in Angkor. *Journal of Field Archaeology*, 43, 492–506. <https://doi.org/10.1080/00934690.2018.1503034>
- Carter, B. P., Blackadar, J. H., & Conner, W. L. A. (2021). When computers dream of charcoal: Using deep learning, open tools, and open data to identify relict charcoal hearths in and around state game lands in Pennsylvania. *Advances in Archaeological Practice*, 9, 257–271. <https://doi.org/10.1017/aap.2021.17>
- Caspari, G., & Crespo, P. (2019). Convolutional neural networks for archaeological site detection—Finding “princely” tombs. *Journal of Archaeological Science*, 110, 104998. <https://doi.org/10.1016/j.jas.2019.104998>
- Cerrillo-Cuenca, E., & Bueno-Ramírez, P. (2019). Counting with the invisible record? The role of LiDAR in the interpretation of megalithic landscapes in south-western Iberia (Extremadura, Alentejo and Beira Baixa). *Archaeological Prospection*, 26, 251–264. <https://doi.org/10.1002/arp.1738>
- Chase, A. S. Z., & Weishampel, J. (2016). Using lidar and GIS to investigate water and soil management in the agricultural terracing at Caracol, Belize. *Advances in Archaeological Practice*, 4, 357–370. <https://doi.org/10.7183/2326-3768.4.3.357>
- Clark, J. T., & Herdrich, D. J. (1993). Prehistoric settlement system in eastern Tutuila, American Samoa. *Journal of the Polynesian Society*, 102, 147–185.
- Cochrane, E. E., & Mills, J. (2018). LiDAR imagery confirms extensive interior land-use on Tutuila, American Samoa. *Journal of Pacific Archaeology*, 9, 70–78.
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., & Böhner, J. (2015). System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geoscientific Model Development*, 8, 1991–2007. <https://doi.org/10.5194/gmd-8-1991-2015>
- Davis, D. S. (2019). Object-based image analysis: A review of developments and future directions of automated feature detection in landscape archaeology. *Archaeological Prospection*, 26, 155–163. <https://doi.org/10.1002/arp.1730>
- Davis, D. S., Caspari, G., Lipo, C. P., & Sanger, M. C. (2021). Deep learning reveals extent of Archaic Native American shell-ring building practices. *Journal of Archaeological Science*, 132, 105433. <https://doi.org/10.1016/j.jas.2021.105433>
- Davis, D. S., Lipo, C. P., & Sanger, M. C. (2019). A comparison of automated object extraction methods for mound and shell-ring identification in coastal South Carolina. *Journal of Archaeological Science: Reports*, 23, 166–177.
- Davis, D. S., & Lundin, J. (2021). Locating charcoal production sites in Sweden using hydrological algorithms and deep learning. *Remote Sensing*, 13, 3680. <https://doi.org/10.3390/rs13183680>
- Day, S. S. (2018). Using unsupervised classification techniques and hypsometric index to identify anthropogenic landscapes through American Samoa. *Journal of the Polynesian Society*, 127, 55–72. <https://doi.org/10.15286/jps.127.1.55-72>
- De Smedt, P., Garwood, P., Chapman, H., Deforce, K., De Grave, J., Hanssens, D., & Vandenberghe, D. (2022). Novel insights into prehistoric land use at Stonehenge by combining electromagnetic and invasive methods with a semi-automated interpretation scheme. *Journal of Archaeological Science*, 143, 105557. <https://doi.org/10.1016/j.jas.2022.105557>
- Dolejš, M., Pacina, J., Veselý, M., & Brétt, D. (2020). Aerial bombing crater identification: Exploitation of precise digital terrain models. *ISPRS International Journal of Geo-Information*, 9, 713. <https://doi.org/10.3390/ijgi9120713>
- Duarte, L., Gomes, A., Teodoro, A., & Rodrigues, S. M. (2018). GIS and remote sensing based method to extract fluvial terraces for archaeological purposes. In *Earth resources and environmental remote sensing/GIS applications IX* (Vol. 10790) (pp. 23–33). SPIE.
- ESRI. (2021). ArcGIS Pro (2.8.1). Environmental Systems Research Institute, Inc. <https://www.esri.com/en-us/arcgis/products/arcgis-pro/>
- Forest, M., Costa, L., Combey, A., Dorison, A., & Pereira, G. (2020). Testing web mapping and active learning to approach lidar data. *Advances in Archaeological Practice*, 8, 25–39. <https://doi.org/10.1017/aap.2019.42>
- Freeland, T., Heung, B., Burley, D. V., Clark, G., & Knudby, A. (2016). Automated feature extraction for prospection and analysis of monumental earthworks from aerial LiDAR in the Kingdom of Tonga. *Journal of Archaeological Science*, 110, 104998. <https://doi.org/10.1016/j.jas.2019.104998>

- Archaeological Science*, 69, 64–74. <https://doi.org/10.1016/j.jas.2016.04.011>
- Gadot, Y., Davidovich, U., Avni, G., Avni, Y., Piasetzky, M., Faerchtein, G., Golan, D., & Porat, N. (2016). The formation of a Mediterranean terraces landscape: Mount Eitan, Judean Highlands, Israel. *Journal of Archaeological Science: Reports*, 6, 397–417.
- Glover, H. J., Ladefoged, T. N., & Cochrane, E. E. (2020). Costly signalling and the distribution of monumental mounds in Savai'i and 'Upolu, Samoa. *Archaeology in Oceania*, 55(3), 141–152. <https://doi.org/10.1002/arco.5212>
- Grammer, B., Draganits, E., Gretschner, M., & Muss, U. (2017). LiDAR-guided archaeological survey of a Mediterranean landscape: Lessons from the ancient Greek polis of Kolophon (Ionia, Western Anatolia). *Archaeological Prospection*, 24(4), 311–333. <https://doi.org/10.1002/arp.1572>
- Green, R. C. (2002). A retrospective view of settlement pattern studies in Samoa. In T. N. Ladefoged & M. Graves (Eds.), *Pacific landscapes: Archaeological approaches* (pp. 125–152). Los Osos.
- Guyot, A., Hubert-Moy, L., & Lorho, T. (2018). Detecting Neolithic burial mounds from LiDAR-derived elevation data using a multi-scale approach and machine learning techniques. *Remote Sensing*, 10, 225. <https://doi.org/10.3390/rs10020225>
- Guyot, A., Lennon, M., & Hubert-Moy, L. (2021). Objective comparison of relief visualization techniques with deep CNN for archaeology. *Journal of Archaeological Science: Reports*, 38, 103027. <https://doi.org/10.1016/j.jasrep.2021.103027>
- He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. *Proceedings of the IEEE International Conference on Computer Vision*, 2961–2969.
- Healy, P. F., Lambert, J. D. H., Arnason, J. T., & Hebda, R. J. (1983). Caracol, Belize: Evidence of ancient Maya agricultural terraces. *Journal of Field Archaeology*, 10, 397–410.
- Henry, E. R., Shields, C. R., & Kidder, T. R. (2019). Mapping the Adena-Hopewell landscape in the Middle Ohio Valley, USA: Multi-scalar approaches to LiDAR-derived imagery from Central Kentucky. *Journal of Archaeological Method and Theory*, 26(4), 1513–1555. <https://doi.org/10.1007/s10816-019-09420-2>
- Herrault, P. A., Poterek, Q., Keller, B., Schwartz, D., & Ertlen, D. (2021). Automated detection of former field systems from airborne laser scanning data: A new approach for Historical Ecology. *International Journal of Applied Earth Observation and Geoinformation*, 104, 102563. <https://doi.org/10.1016/j.jag.2021.102563>
- Hightower, J. N., Butterfield, A. C., & Weishampel, J. F. (2014). Quantifying ancient Maya land use legacy effects on contemporary rainforest canopy structure. *Remote Sensing*, 6, 10716–10732. <https://doi.org/10.3390/rs61110716>
- Huggett, J. (2021). Algorithmic agency and autonomy in archaeological practice. *Open Archaeology*, 7, 417–434. <https://doi.org/10.1515/opar-2020-0136>
- Hunt, T. L., & Kirch, P. V. (1988). An archaeological survey of the Manu'a islands, American Samoa. *Journal of the Polynesian Society*, 97, 153–183.
- Jackmond, G., Fonoti, D., & Tautunu, M. M. (2018). Samoa's hidden past: LiDAR confirms inland settlement and suggests larger populations in pre-contact Samoa. *Journal of the Polynesian Society*, 127, 73–90. <https://doi.org/10.15286/jps.127.1.73-90>
- Johnson, K. M., & Ouimet, W. B. (2018). An observational and theoretical framework for interpreting the landscape palimpsest through airborne LiDAR. *Applied Geography*, 91, 32–44. <https://doi.org/10.1016/j.apgeog.2017.12.018>
- Kirch, P. V. (1982). Ecology and the adaptation of Polynesia agricultural systems. *Archaeology in Oceania*, 17, 1–6. <https://doi.org/10.1002/j.1834-4453.1982.tb00032.x>
- Kirch, P. V. (1994). *The wet and the dry: Irrigation and agricultural intensification in the Pacific*. Chicago: University of Chicago Press.
- Klassen, S., Carter, A. K., Evans, D. H., Ortman, S., Stark, M. T., Loyless, A. A., Polkinghorne, M., Heng, P., Hill, M., Wijker, P., Niles-Weed, J., Marriner, G. P., Pottier, C., & Fletcher, R. J. (2021). Diachronic modeling of the population within the medieval Greater Angkor Region settlement complex. *Science Advances*, 7, eabf8441. <https://doi.org/10.1126/sciadv.abf8441>
- Korobov, D. S., & Borisov, A. V. (2013). The origins of terraces field agriculture in the Caucasus: New discoveries in the Kislovodsk basin. *Antiquity*, 87, 1086–1103. <https://doi.org/10.1017/S0003598X00049887>
- Kuhlken, R., & Crosby, A. (1999). Agricultural terracing at Nakauvadra, Viti Levu: A late prehistoric irrigated agrosystem in Fiji. *Asian Perspectives*, 38, 62–89.
- Ladefoged, T. N., McCoy, M. D., Asner, G. P., Kirch, P. V., Puleston, C. O., Chadwick, O. A., & Vitousek, P. M. (2011). Agricultural potential and actualized development in Hawai'i: An airborne LiDAR survey of the leeward Kohala field system (Hawai'i Island). *Journal of Archaeological Science*, 38, 3605–3619. <https://doi.org/10.1016/j.jas.2011.08.031>
- Lepofsky, D. S. (1994). Prehistoric agricultural intensification in the Society Islands, French Polynesia. Unpublished PhD dissertation, Department of Anthropology, University of California at Berkeley.
- Liston, J. (2009). Cultural chronology of earthworks in Palau, Western Micronesia. *Archaeology in Oceania*, 44, 56–73. <https://doi.org/10.1002/j.1834-4453.2009.tb00047.x>
- Macrae, S., & Iannone, G. (2016). Understanding ancient Maya agricultural terrace systems through Lidar and hydrological mapping. *Advances in Archaeological Practice*, 4, 371–392. <https://doi.org/10.7183/2326-3768.4.3.371>
- Magnini, L., Bettineschi, C., & De Guio, A. (2017). Object-based shell craters classification from LiDAR-derived sky-view factor. *Archaeological Prospection*, 24, 211–223. <https://doi.org/10.1002/arp.1565>
- McCoy, M. D., Asner, G. P., & Graves, M. W. (2011). Airborne lidar survey of irrigated agricultural landscapes: An application of the slope contrast method. *Journal of Archaeological Science*, 38, 2141–2154. <https://doi.org/10.1016/j.jas.2011.02.033>
- Meyer, M., Pfeffer, I., & Jürgens, C. (2019). Automated detection of field monuments in digital terrain models of Westphalia using OBIA. *Geosciences*, 9, 109. <https://doi.org/10.3390/geosciences9030109>
- Meyer-Heß, M. F. (2020). Identification of archaeologically relevant areas using open geodata. *KN-Journal of Cartography and Geographic Information*, 70, 107–125. <https://doi.org/10.1007/s42489-020-00049-w>
- OCM (Office for Coastal Management). (2022). 2012 NOAA American Samoa Lidar DEM: Islands of Tutuila, Aunu'u, Ofu, Olosega, Ta'u and Rose Atoll, <https://www.fisheries.noaa.gov/import/item/52876>
- Parton, P., & Clark, G. (2022). Using lidar and Bayesian inference to reconstruct archaeological populations in the Kingdom of Tonga. *Journal of Archaeological Science: Reports*, 45, 103610. <https://doi.org/10.1016/j.jasrep.2022.103610>
- Parton, P., Clark, G., Reepmeyer, C., & Burley, D. (2018). The field of war: LiDAR identification of earthwork defences on Tongatapu Island, Kingdom of Tonga. *Journal of Pacific Archaeology*, 9, 11–24.
- Pearl, F. (2004). The chronology of mountain settlements on Tutuila, American Samoa. *The Journal of the Polynesian Society*, 113, 331–348.
- Pérez Rodríguez, V. (2016). Terrace agriculture in the Mixteca Alta region, Oaxaca, Mexico: Ethnographic and archaeological insights on terrace construction and labor organization. *Culture, Agriculture, Food and Environment*, 38, 18–27. <https://doi.org/10.1111/cuag.12062>
- Pérez Rodríguez, V., & Anderson, K. C. (2013). Terracing in the Mixteca Alta, Mexico: Cycles of resilience of an ancient land-use strategy. *Human Ecology*, 41, 335–349. <https://doi.org/10.1007/s10745-013-9578-8>
- Quintus, S. (2018). Exploring the intersection of settlement, subsistence, and population in Manu'a. *Journal of the Polynesian Society*, 127, 15–34. <https://doi.org/10.15286/jps.127.1.15-34>
- Quintus, S. (2020). Community relationships and integration at a small-island scale. *Journal of Pacific Archaeology*, 11, 61–70.

- Quintus, S., Clark, J. T., Day, S. S., & Schwert, D. P. (2015). Investigating regional patterning in archaeological remains by pairing extensive survey with a lidar dataset: The case of the Manu'a group, American Samoa. *Journal of Archaeological Science: Reports*, 2, 677–687.
- Quintus, S., Day, S. S., & Smith, N. J. (2017). The efficacy and analytical importance of manual feature extraction using lidar datasets. *Advances in Archaeological Practice*, 5, 351–364. <https://doi.org/10.1017/aap.2017.13>
- Quintus, S., Huebert, J., Day, S., Lincoln, N., Yoo, K., Lee, T., Filimoehala, D., & Autufuga, D. (2020). Tempo and trajectory of the built landscape on Ta'u island, Manu'a group, American Samoa: Integrating extensive radiocarbon dating with joint posterior modeling. *Radiocarbon*, 62, 1317–1337. <https://doi.org/10.1017/RDC.2020.60>
- Quintus, S. J., Autufuga, D., Day, S., Huebert, J., Lincoln, N. K., Motu, N., & Yoo, K. (2022). Tracking emergent spatial and social patterns across terraced landscapes in Polynesia. *Journal of Field Archaeology*, 47, 196–211. <https://doi.org/10.1080/00934690.2021.2018259>
- Raber, S. (2012). Post-flight aerial acquisition and calibration report. Unpublished report completed by Photo Science, Inc. Colorado Springs, Colorado.
- Righetti, G., Serafini, S., Rueda, F. B., Church, W. B., & Garnero, G. (2021). Survey of a Peruvian archaeological site using LiDAR and photogrammetry: A contribution to the study of the Chachapoya. In *International Conference on Computational Science and Its Applications* (pp. 613–628). Springer. [https://doi.org/10.1007/978-3-030-86979-3\\_43](https://doi.org/10.1007/978-3-030-86979-3_43)
- Riley, S. J., DeGloria, S. D., & Elliot, R. (1999). Index that quantifies topographic heterogeneity. *Intermountain Journal of Sciences*, 5, 23–27.
- Różycka, M., Migoń, P., & Michniewicz, A. (2017). Topographic wetness index and terrain ruggedness index in geomorphic characterisation of landslide terrains, on examples from the Sudetes, SW Poland. *Zeitschrift für geomorphologie, Supplementary issues*, 61, 61–80. [https://doi.org/10.1127/zfg\\_suppl/2016/0328](https://doi.org/10.1127/zfg_suppl/2016/0328)
- Sánchez Díaz, F., García Sanjuán, L., & Rivera Jiménez, T. (2022). Potential and limitations of LiDAR altimetry in archaeological survey. Copper Age and Bronze Age settlements in southern Iberia. *Archaeological Prospection*, 29, 525–544. <https://doi.org/10.1002/arp.1869>
- Sand, C., Bole, J., & Ouetcho, A. (2003). Prehistory and its perception in a Melanesian archipelago: The New Caledonia example. *Antiquity*, 77, 505–519. <https://doi.org/10.1017/S0003598X00092565>
- Sandor, J. A., Gersper, P. L., & Hawley, J. W. (1990). Prehistoric agricultural terraces and soils in the Mimbres area, New Mexico. *World Archaeology*, 22, 70–86. <https://doi.org/10.1080/00438243.1990.9980130>
- Sharma, M., Garg, R. D., Badenko, V., Fedotov, A., Min, L., & Yao, A. (2021). Potential of airborne LiDAR data for terrain parameters extraction. *Quaternary International*, 575–576, 317–327. <https://doi.org/10.1016/j.quaint.2020.07.039>
- Smith, L. N. (2017). Cyclical learning rates for training neural networks. In *2017 IEEE Winter Conference on Applications of Computer Vision (WACV)* (pp. 464–472). IEEE.
- Somrak, M., Džeroski, S., & Kokalj, Ž. (2020). Learning to classify structures in ALS-derived visualizations of ancient Maya settlements with CNN. *Remote Sensing*, 12, 2215. <https://doi.org/10.3390/rs12142215>
- Soroush, M., Mehrtash, A., Khazraee, E., & Ur, J. (2020). Deep learning in archaeological remote sensing: Automated qanat detection in the Kurdistan region of Iraq. *Remote Sensing*, 12, 500. <https://doi.org/10.3390/rs12030500>
- Sugiyama, N., Sugiyama, S., Catignani, T., Chase, A. S. Z., & Fernandez-Diaz, J. C. (2021). Humans as geomorphic agents: Lidar detection of the past, present and future of the Teotihuacan Valley, Mexico. *PLoS ONE*, 16, e0257550. <https://doi.org/10.1371/journal.pone.0257550>
- Taomia, J. M. E. (2000). Household units in the analysis of prehistoric social complexity, Southern Cook Islands. *Asian Perspectives*, 39, 139–164. <https://doi.org/10.1353/asi.2000.0014>
- Treacy, J. M., & Denevan, D. M. (1994). The creation of cultivatable land through terracing. In N. F. Miller & K. L. Gleason (Eds.), *The archaeology of garden and field* (pp. 91–110). University of Pennsylvania Press.
- Trier, Ø. D., Cowley, D. C., & Waldehand, A. U. (2019). Using deep neural networks on airborne laser scanning data: Results from a case study of semi-automatic mapping of archaeological topography on Arran, Scotland. *Archaeological Prospection*, 26, 165–175. <https://doi.org/10.1002/arp.1731>
- Verschoof-van der Vaart, W., Bonhage, A., Schneider, A., Ouimet, W., & Raab, T. (2022). Automated large-scale mapping and analysis of relict charcoal hearths in Connecticut (USA) using a Deep Learning YOLOv4 framework. *Archaeological Prospection*. <https://doi.org/10.1002/arp.1889>
- Verschoof-van der Vaart, W. B., & Lambers, K. (2022). Applying automated object detection in archaeological practice: A case study from the southern Netherlands. *Archaeological Prospection*, 29, 15–31. <https://doi.org/10.1002/arp.1833>
- Verschoof-van der Vaart, W. B., Lambers, K., Kowalczyk, W., & Bourgeois, Q. P. J. (2020). Combining deep learning and location-based ranking for large-scale archaeological prospection of LiDAR data from the Netherlands. *ISPRS International Journal of Geo-Information*, 9, 293. <https://doi.org/10.3390/ijgi9050293>
- Verschoof-van der Vaart, W. B., & Landauer, J. (2020). Using Carcassonne-Net to automatically detect and trace hollow roads in LiDAR data from the Netherlands. *Journal of Cultural Heritage*, 47, 143–154.
- Yokoyama, R., Shirasawa, M., & Pike, R. J. (2002). Visualizing topography by openness: A new application of image processing to digital elevation models. *Photogrammetric Engineering and Remote Sensing*, 68, 257–266.

**How to cite this article:** Quintus, S., Davis, D. S., & Cochrane, E. E. (2023). Evaluating Mask R-CNN models to extract terracing across oceanic high islands: A case study from Sāmoa. *Archaeological Prospection*, 1–16. <https://doi.org/10.1002/arp.1909>