Scaling Up Deep Learning to Identify Earthwork Sites in Te Tai Tokerau, Northland, New Zealand

Simon H. Bickler *(Bickler Consultants Ltd)*
Benjamin Jones *(University of Auckland)*

Introduction

The machine learning (ML) toolset is ideal for New Zealand archaeology (Bickler 2021a). The ability to leverage LiDAR data with repeatable and transparent practices is an attractive prospect. As shown previously by the authors, it is relatively easy to identify many of these archaeological sites and features using LiDAR (Jones and Bickler 2017, 2019; see also the LINZ National Elevation Programme "Elevation Aotearoa"). The challenge is to scale up this process to search regions to allow for a landscape reconstruction hinging on robust interpretation. The region focused on currently is Te Tai Tokerau / Northland which has over 12,000 archaeological sites recorded in the NZAA ArchSite database with around half, including earthwork features, related to Māori history (Figure 1). These include pā, pits, and terraces. Other sites such as stone structures, sod walls, tracks, ditches, and drains representing both more recent and longer-term landscape history are also present, and identifiable, through LiDAR processing. To identify earth worked sites on the surface, LiDAR data in tandem with spatial data from ArchSite was combined to produce spatial data needed to drive machine learning models. This paper describes our latest attempts to scale up the identification of archaeological sites in the forested areas of Te Tai Tokerau / Northland and determine the most effective protocols to find features through neural network processing.

The Search for Sites

The increasing availability of large-scale LiDAR, satellite and aerial imagery on local, regional, and national scales is transforming archaeology around the globe, particularly the searching and mapping of archaeological sites. ML algorithms can be used to process the geo-spatial data in the search for sites in diverse environments throughout the world (Bonhage et al. 2021; Caspari and Crespo 2019; Davis 2019, Davis et al. 2020; Evans and Hofer 2019; Guyot et al. 2018, 2021; Orengo et al. 2020; Soroush et al. 2020; Thabeng et al. 2019; Trier et al. 2018, 2019; Verschoof-van der Vaart and Lambers 2019; Verschoof-van der Vaart et al. 2020).

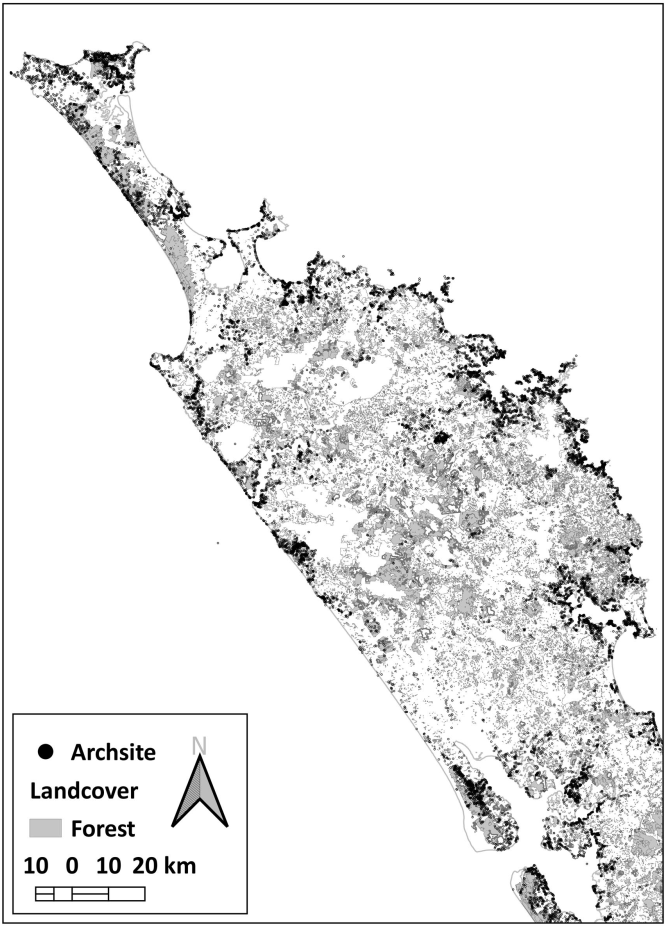

Figure 1. Distribution of archaeological sites in Te Tai Tokerau / Northland with forested areas (Source data: ArchSite 2021, Landcare Research Land cover database version 4.1, LRIS Portal, Coverage date 2012 including recently harvested areas).

The Tai Tokerau region however is large and much of it forested (Figure 1). Despite a long history of archaeologists working in the area, an examination of the site records (Bickler 2021b) shows that sites were:

- Recorded either from oral history, local informants or archival records and visited or found by archaeologists,

- Identified during initial survey but not re-visited or found subsequently,
- Not accurately located,
- Not fully characterised during recording with only some elements visible but others potentially obscured by vegetation, inaccessible, or destroyed.

LiDAR therefore offers the opportunity to identify the location of sites and based on preservation allows for the accurate recording of features of those sites in many cases (see Jones and Bickler 2017). Field survey will be required to check results of the any such identification and as recent research elsewhere in New Zealand has shown, such work is productive and essential for understanding landscapes (e.g., Hagan and Brown 2019).

Creating the Training Model

Machine Learning requires the creation of a trained model as well as a dataset to train on (Figure 2, see Bickler 2021a for further discussion). The process is time-consuming where the training required needs to be specific to the data being analysed. The data used here was a Digital Elevation Model (DEM) based on LiDAR data from the Northland Regional Council (NRC, 2020). As the NRC DEM was pre-processed from the point-cloud information it was not ideal but good enough for this pilot study.

The Image Analysis tool in ArcMap used a combination of Satellite and LiDAR imagery to create a library of features. This was initially done by creating an outline of features which could then be reused with different sets of imagery. Initially, pits, pit clusters, ditches, and likely terraces were marked out and a GIS layer created with their extents (Figure 3). This meant that new training data could be generated from processed layers such as hill-shading, low relief terrain modelling and slope data.

It became apparent that the quality of the DEM would make the feature-based model less reliable at this stage, and as a result, a focus on clusters of features and especially those associated with pā sites would be a better first step. Distortions in the layers mean that small features may not be cropped accurately.

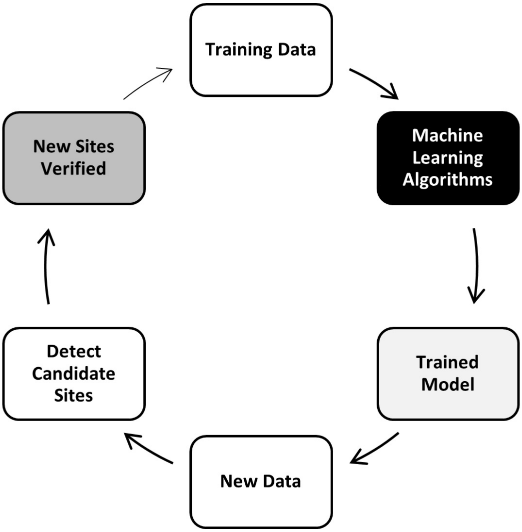

Figure 2. High level overview of the machine learning process.

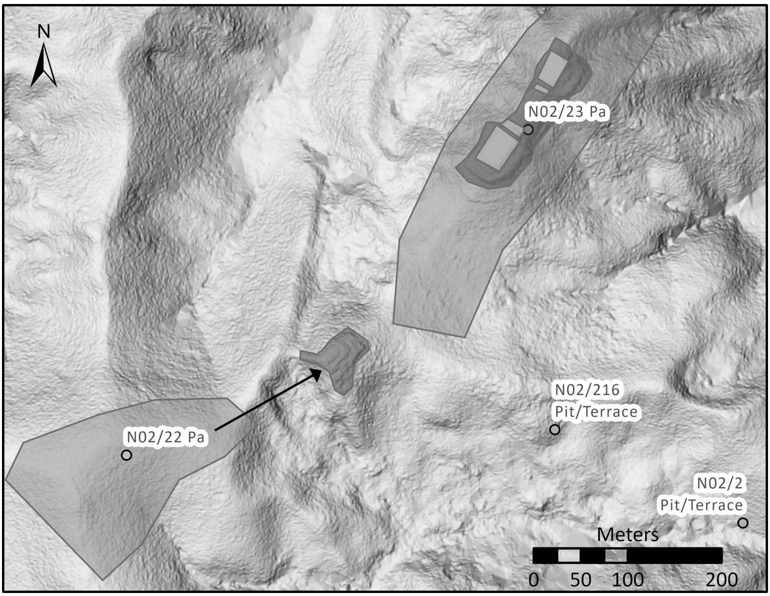

Figure 3. Hill-shaded LiDAR DEM showing location of sites recording ArchSite including area of N02/22 as recorded in ArchSite and actual location of site visible in the imagery.

ArchSite was used to identify the primary candidates for selection of sites for the training database. However, both the point locations and the areas of many sites from ArchSite could not be relied upon (see e.g., Figure 3), and so all features and sites used were redrawn based on the combination of imagery available.

Sites from a range of ground and vegetation conditions were selected including features which were only visible in the LiDAR under vegetation. This was necessary as identifying sites under vegetation growth was one of the main objectives of the project.

To further enhance the possibility of finding sites on the LiDAR imagery, a composite image was created from the DEM. Following approximately the method used in Guyot et al. (2018, 2021), the composite imagery was built on layers of the openness, slope, and local relief of the DEM. This was done to try to highlight those features such as pits, ditches, and terraces. Essentially, local relief creates a flat landscape in which features that are relatively small on the landscape but with major slope changes, are emphasised. This means that the larger topography, such as hills and valleys, are de-emphasised. An example of the composite image used is shown in Figure 4. Other composite images with different attributes can also be used.

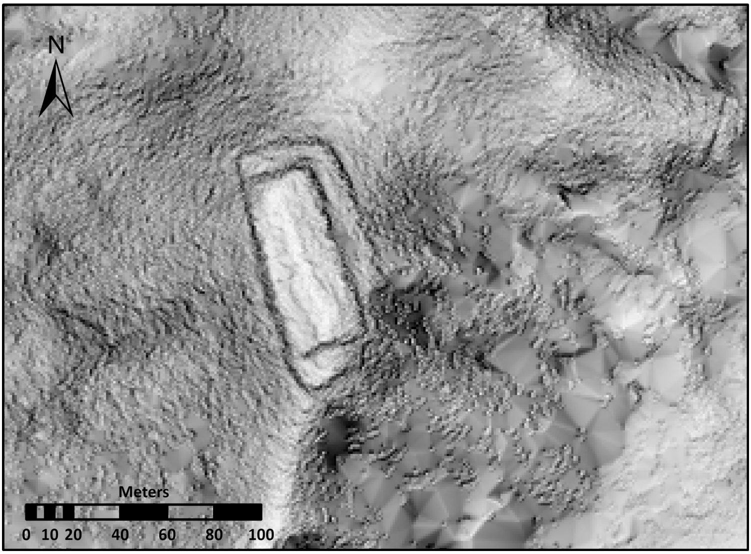

Figure 4. Composite image (LINZ Basemaps 2021) on left and composite LiDAR image on right of pā site N02/6.

ArcGIS Pro was then used to first export a set of between 25-40 “cut-outs” of the composite image of the areas of the selected known pā sites. These were then used to create the training model using the Faster R-Convolution Neural Network or R-CNN Model with a backbone of the RESNET-101 model within ArcGIS Pro’s Machine Learning toolkit. The details of the R-CNN algorithms are beyond the scope of this article (see Bonhage et al. 2021, Guyot et al. 2021 for additional discussion). The essence of the algorithms is that the software creates sets of mathematical vectors which describe possible features in the images. The groups of those features are classed according to the classification applied, in this case, areas in the composite layer that show “pā” sites and areas that do not.

The algorithm selects a proportion of the categorised images and sets those up as a test sample of unknowns. It then tests to see whether it can correctly assign them to a pā (or not a pā) using the parameters it has determined from the remaining classified sample. It repeats this testing process changing which are the test and which are the known samples as well as changing the parameters of the model until it gets the best result it can. This is called “training”.

Ideally, the model should correctly classify 100% of the time and in more critical situations, a model needs to get as close to that as possible. However, in the example of looking for these pā sites, a model is unlikely to be as successful because of the variability in the underlying data including the inherent complexity of the topography and irregularity of pa sites. They are all unique.

Once the result of the training is completed, a trained model with the parameters that classify the pā is generated. This can then be used to detect pā sites on the wider landscape. ArcGIS Pro does this by sampling the larger composite images in blocks and classifying whether it detects a pā or not and suggesting how likely different blocks might be pā sites.

Results

Some preliminary results show how effective the process can be. Figure 5 shows a comparison between the satellite image and composite LiDAR image around site N02/799 (Te Pōkere) with the black bounding boxes from the ML identification of the pā site. Interestingly, the site is a double pā with two tihi and surrounding ditches, however the ML model detects the site as two separate but overlapping pā.

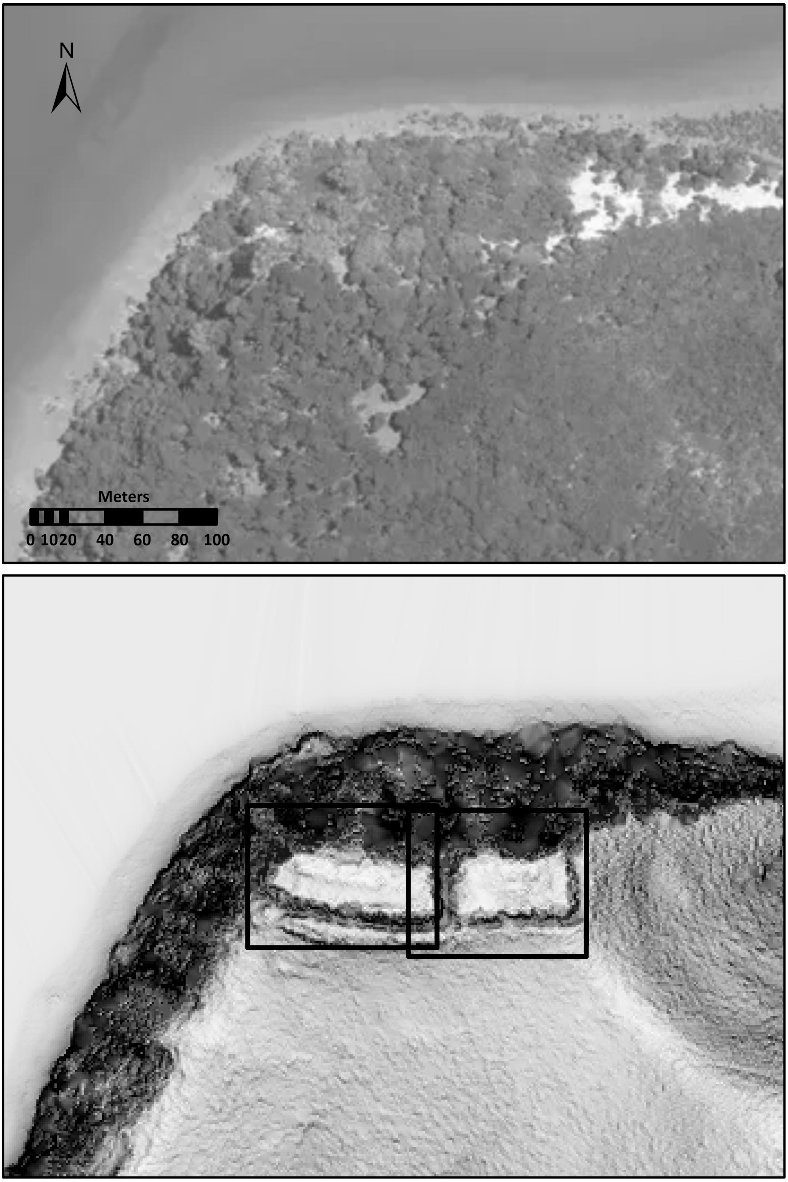

Figure 5. Detection of double pā site N02/799 with ML model identifying the two separate parts of the pā in the 2021 aerial image (top image) and Composite LiDAR image (bottom image).

Figure 6. Composite LiDAR image showing possible pā sites identified (black boxes), ArchSite locations of known pā sites (large dots) and other recorded sites (small dots).

Taking a broader view, Figure 6 shows an area of Te Tai Tokerau / Northland (part of Mapsheet N02) with the location of possible pā sites (black boxes) and the location of recorded pā sites (large dots). The circles show the likely correlation between those known sites and nearby ML detected sites.

The results show that many of the known pā sites are correctly identified, and indeed are more accurately located than their ArchSite information. A smaller number of known pā are not identified. The reasons for the failure to detect the pā includes issues such as the DEM material not being accurate enough to pick up the archaeological features, the site is not classified as a pā in ArchSite, but features are present, or the site has been damaged. More common are the false classification of areas that are not archaeological sites but where local topography, roading and other modern activities have made areas look like possible pā sites. This is especially common around small hills with farm tracks creating ditch-like features around spurs.

Despite these difficulties, new possible sites were detected and could be confirmed by the satellite imagery and careful examination of the LiDAR. Figure 7 shows a new pā site N02/1149 recorded in ArchSite from this analysis.

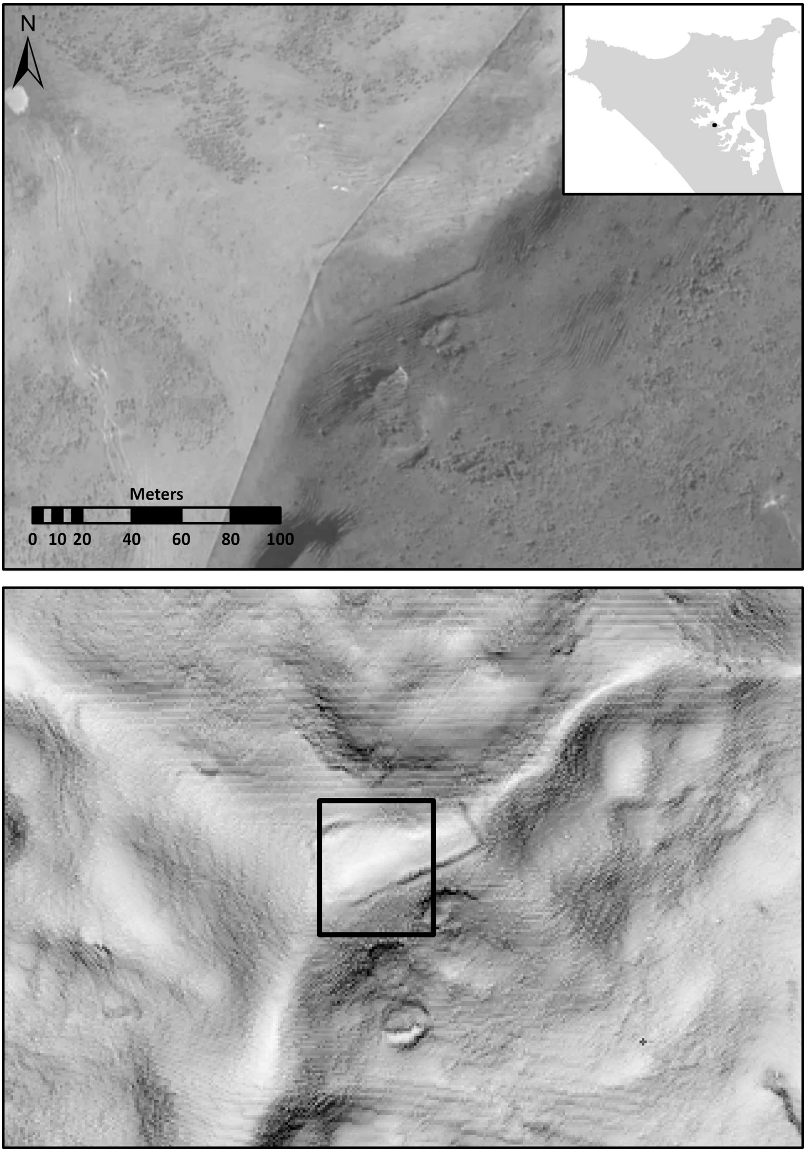

Figure 7. New pā site, N02/1149, recorded by ML detection and confirmed in the satellite imagery.

Discussion

The use of ML for site detection in Aotearoa reflects the reality that the area of archaeological potential is large. Despite over 70 years of active site survey using a variety of tools including aerial and satellite imagery, major archaeological sites are still being newly identified in the landscape. The management of these taonga is essential and their identification is fundamental to NZAA's role in archaeology in New Zealand (Bickler 2018: 55ff).

The process of scaling-up this identification requires more than simple use of LiDAR DEM information to be effective. ML can increase the chance of accurately locating archaeological sites and features across large landscapes, identify likely new areas for on-ground site survey and lead to large-scale regional reconstruction (see e.g., Davis et al. 2020).

In New Zealand, much of the archaeological literature is dominated by a focus on sites particularly driven by the NZAA's ArchSite database. The allocation of labels within ArchSite has always been somewhat contentious (Bickler 2018: 60-62). This has led to suggestions, not always made clear, that New Zealand archaeologists are sometimes overly focused on these sites as bounded entities that constrain archaeology (see e.g., McCoy 2020). This has been exacerbated in recent times by the legal requirements relating to site protection and heritage management, in the New Zealand case via the *Heritage New Zealand Pouhere Taonga Act* (2014) and the *Resource Management Act* (1993). These frameworks have meant that archaeological remains become quickly reified into entities that are then effectively evaluated according to their significance and value (in several possible ways) and compared with the requirements of land development on which they occur (see Bickler 2018: 81ff for further discussion).

It would naïve however, to suggest that the archaeologists using such site-based schema are unaware of the limitations of the approach and indeed there are frequent calls for alternatives more landscape-based approaches. Indeed, this has occurred (see Bickler 2018: 94-95) in a limited fashion, although this can cut both ways providing some areas with mechanisms for more holistic protection and others for more widespread removal.

Expanding the information base and improving the ability to identify archaeological sites prior to development does, however, allow the possibility to better plan ways of mitigating damage from both natural forces and land development. Climate change is at the forefront (Ramsay 2021) but the effects of rampant land-development is under-studied.

The implications of this are that managing the ML models' misclassification, resulting in either targets being wrongly classified or not classified at all, should be part of the strategy for their use in archaeological situations. The algorithms usually offer a range of ways of establishing their mathematical robustness, but archaeologists still need to ensure that the results stand up to scrutiny in the real-world.

Part of that real work relates to the use and management of the data, especially as the sites relate to the complexity of land ownership, kaitiakitanga and mātauranga Māori (see e.g., Allen and Phillips 2010, Bickler 2018: 293ff, Solomon and Forbes 2010). The public accessibility of the data, such as LiDAR, does not mitigate the emerging issues relating the use of this information to expose cultural information (see Davis et al. 2021), the use of the information in the management by iwi and hapū (e.g., Allen et al. 2002, Kaiser and Saunders 2021), and especially when such decision systems are automated (see Digital Council of New Zealand 2020, West et al. 2020).

The results of this work represent the continuing development of tools to assist archaeologists in managing regional and national-scale projects. The LiDAR represents only one component of the landscape approach being examined. We hope to improve the model to include better quality DEM data, improve the effectiveness of the composite imagery, increase the scale and diversity of the training data and test the detection across larger landscapes. It is not just the case of using simple geospatial techniques such as hill-shading to make sites appear, but the layering of different sources of information, often a range of ways of understanding topography, geomorphology, and vegetation cover, to maximise the identification of the sites and critically understand how sites fit within their natural and cultural landscapes in the past, present and, as the impacts of climate change compound, the future.

Acknowledgements

Versions of this paper were presented at the NZAA Conference, Taupo 2021 and at the CAA Australasia Online Conference in 2021. We are grateful to the organisers of those events for the opportunity to present. Dorothy Brown helped sort out the text and references for which we remain ever-grateful. Ben Collings for input on remote sensing and workflow discussions.

References

Allen, H., D. Johns, C. Phillips, K. Day, T. O'Brien and Ngāti Mutunga. (2002). Wāhi Ngaro (the lost portion): Strengthening relationships between people and wetlands in north Taranaki, New Zealand. World Archaeology 34: 315-329.

Allen, H. and C. Phillips. (2010). Maintaining the dialogue: Archaeology, cultural heritage and indigenous communities. In C. Phillips and H. Allen (eds), Bridging the Divide: Indigenous Communities and Archaeology into the 21st Century. New York: Routledge, pp. 17-48.

Bickler, S. H. (2018). Cultural Resource Management Archaeology in New Zealand: A Guide for Students and Practitioners. Auckland: New Zealand.

Bickler, S. H. (2021a). Machine Learning arrives in archaeology. Advances in Archaeological Practice 9 (2): 186-191.

Bickler, S. H. (2021b). What's in an ArchSite name? Archaeology in New Zealand 64 (2): 21-33.

Bonhage, Alexander, Mahmoud Eltaher, T. Raab, Michael Breuß, Alexandra Raab and Anna 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 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.

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 (2), 155-163.

Davis, D.S., Katherine E. Seeber and Matthew C. Sanger. (2020). Addressing the problem of disappearing cultural landscapes in archaeological research using multi-scalar survey. The Journal of Island and Coastal Archaeology, DOI: 10.1080/15564894.2020.1803457

Davis, D. et al. (2021). The aerial panopticon and the ethics of archaeological remote sensing in sacred cultural spaces. Archaeological Prospection 28. 305-320.

Digital Council for Aotearoa New Zealand. (2020). Towards Trustworthy and Trusted Automated Decision-Making in Aotearoa. A Research Report to the Government from the Digital Council for Aotearoa New Zealand. Online report.

Evans, D. and N. Hofer. (2019). Exploring complexity in the archaeological landscapes of monsoon Asia using lidar and deep learning. Geophysical Research Abstracts p. 1-1.

Guyot, A., L. Hubert-Moy and T. Lorho. (2018). Detecting Neolithic burial mounds from LiDAR-derived elevation data using a multi-scale approach and Machine Learning techniques. Remote Sensing 10. 225. 10.3390/rs10020225.

Guyot, A., M. Lennon, T. Lorho and L. Hubert-Moy. (2021). Combined detection and segmentation of archeological structures from LiDAR data using a deep learning approach. Journal of Computer Applications of Archaeology 4. 1. 10.5334/jcaa.64.

Hagan, J., and A. Brown. (2019). LiDAR in New Zealand archaeology: Prospects and pitfalls. Journal of Pacific Archaeology 10 (2): 80-91.

Jones, B. and S. H. Bickler. (2017). High resolution LiDAR data for landscape archaeology in New Zealand. Archaeology in New Zealand 60 (3): 35-44.

Jones, B. and S. H. Bickler. (2019). Multi-scalar and semi-automatic approaches to detect archaeological features in NZ using Airborne LiDAR data. Archaeology in New Zealand 62 (3):10-24.

Kaiser, L. H. and W. S. A. Saunders. (2021). Vision Mātauranga research directions: Opportunities for iwi and hapū management plans. Kōtuitui: New Zealand Journal of Social Sciences Online, 16 (2): 371-383.

McCoy, M. (2020). The site problem: A critical review of the site concept in archaeology in the digital age. Journal of Field Archaeology 45: S18-S26.

Orengo, H. A., F. C. Conesa, A. Garcia-Molsosa, A. Lobo, A. S. Green, M. Madella and C. A. Petrie. (2020). Automated detection of archaeological mounds using machine-learning classification of multisensor and multitemporal satellite data. Proceedings of the National Academy of Sciences 117(31): 18240-18250.

Ramsay, R. (2021). Climate change and cultural heritage – Implementing a strategic approach. Archaeology in New Zealand 64 (1):18-24.

Solomon, M. and S. Forbes. (2010). Indigenous archaeology: A Moriori case study. In C. Phillips and H. Allen (eds), Bridging the Divide: Indigenous Communities and Archaeology into the 21st Century. New York: Routledge, pp. 213-32.

Soroush, M., A. Mehrtash, E. Khazraee and J. A. Ur. (2020). Deep learning in archaeological remote sensing: Automated qanat detection in the Kurdistan region of Iraq. Remote Sensing 12 (3): 500.

Thabeng, O. L., S. Merlo and E. Adam. (2019). High-resolution remote sensing and advanced classification techniques for the prospection of archaeological sites' markers: The case of dung deposits in the Shashi-Limpopo Confluence area (southern Africa). Journal of Archaeological Science 102: 48-60.

Trier, Ø. D., D. C. Cowley and A. U. Waldeland. (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 (2): 165-175.

Trier, Ø.D., A.-B. Salberg and L. H. Pilø. (2018). Semi-automatic mapping of charcoal kilns from airborne laser scanning data using deep learning. In M. Matsumoto and E. Uleberg (eds), CAA2016: Oceans of Data. Proceedings of the 44th Conference on Computer Applications and Quantitative Methods in Archaeology. Oxford: Archaeopress, pp. 219–231.

Verschoof-van der Vaart, W.B. and K. Lambers. (2019). Learning to Look at LiDAR: The use of R-CNN in the automated detection of archaeological objects in LiDAR data from the Netherlands. Journal of Computer Applications in Archaeology 2 (1): 31–40.

Verschoof-van der Vaart, W. B., K. Lambers, W. Kowalczyk and Q.P. Bourgeois. (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 (5): 293.

West, K., D. Wilson, A. Thompson and M. Hudson. (2020). Māori Perspectives on Trust and Automated Decision-Making. Report for the Digital Council. Te Kotahi Research Institute, The University of Waikato.