# Developing a Machine Learning Model for Archaeological Lidar Data <Style CHNT_Title>

## Project Background<CHNT_Heading 1>

Airborne laser scanning (ALS) and other remote sensing prospection techniques are significant in the context of archaeological site identification, research, management, and mitigation strategies; particularly in relation to landscape-scale research projects and development. With the increasing availability of high-quality ALS data (Shaw & Corns, 2011) and the advancements in image analysis and remote sensing processing methodologies (Ma, L. et al., 2019), the potential and opportunity to utilise machine learning (ML) methods within monument prospection and documentation practices in archaeology are great. The Transport Infrastructure Ireland (TII) ALS Machine Learning Research Project aimed to develop a user-friendly software tool capable of automatically detecting archaeological features from ALS data using machine learning, specifically deep learning algorithms. Comprising a European-wide interdisciplinary team with experts from archaeology, remote sensing, earth sciences, and machine learning the project developed the Automatic Detection of Archaeological Features (ADAF) tool. The project was commissioned by TII as part of their 2021 Open Research Call and led by The Discovery Programme and aimed to mitigate the potential impact of the discovery of archaeological remains/features and the resulting excavation during major road scheme construction. <CHNT_Text>

## Irish Monuments & ALS Data <CHNT_Heading 1>

Within the research call specifications, TII required the identification of suitable archaeological monument types. With c.480 individual Irish monument classifications, careful consideration was given to select those monument types that would have, among other criteria, sufficient topographic expression and a sufficiently large volume of recorded examples. In order to train and evaluate a machine learning model (MLM) a critical number of ALS monument examples was required; therefore, all available ALS data was utilised, available through the GSI Open Topographic Data Viewer (Geological Survey Ireland, 2024) and supplemented with ALS data from the DEFRA UK data (DEFRA, 2024)  for those monument types which have share topographic relief types in both Ireland and the UK.<CHNT_Text>

## Creating the ADAF Tool<CHNT_Heading 1>

To create the ADAF tool and develop an MLM based upon supervised training processes (Zhang, L., Zhang, L. and Du, B., 2016) several consecutive activities were carried out:

### 1. Training Data Creation<CHNT_Heading 2>

The initial data processing involved cleaning and reprocessing all available ALS data to a consistent high quality to ensure that any MLM created was based on optimal data . Once the monument types were identified, more than 10,000 segmentation polygons defining the visible topographic extents of the features were digitised for testing and validation of the machine learning model. However, archaeological remains are not uniform and there can be significant differences in quality within the monument classifications and across morphology and survival with the LAS data representing them of differing quality. All segmentation polygons were therefore classified to reflect these differences with the quality of both the data and the underlying monument and ALS data patches recorded. 10,000 segmentation ALS patches were extracted then divided into training, testing, and evaluation groups using 60%, 20%, 20% amounts respectively<CHNT_Text>

## 2. MLM Experimentation<CHNT_Heading 2>

Utilising the created training and testing ALS data patches, a series of experiments were conducted on the object detection and semantic segmentation MLM to determine the optimal configurations for the software. The criteria evaluated included:

- Testing different ALS visualisation methodologies within the MLM,
- Testing different deep learning architectures
- Evaluation of patch size in the MLM
- Testing the effects of augmenting the ALS training data
- Evaluation of the filtering of training data based on the quality of the underlying ALS data and archaeological preservation

## 3. MLM Testing & Validation<CHNT_Heading 2>

Following the creation of a MLM, the different object detection and segmentation models were tested and evaluated to critically address any issues such as false positives being caused by similarly shaped landscape features and the number of features/sites which the ADAF MLM missed. Within this project component it was identified that some form of filtering mechanism was required within the ADAF tool to effectively reduce the number of false positives for the user. <CHNT_Text>

## 4. Integration of MLM into the DAF Tool<CHNT_Heading 2>

Once the development of MLM was finalised it was integrated into the ADAF tool utilising Jupyter Notebooks. The ADAF software requires minimal user interaction and requires no prior ML expertise, enhancing accessibility within the archaeological community. Key components of the tool include the Relief Visualization Toolbox (RVT) (ZRC SAZU, 2013) and the Artificial Intelligence Toolbox for Earth Observation (AiTLAS) (Dimitrovski et al., 2023) both actively utilised in aerial archaeology. RVT processes input data by converting digital elevation models into ML-friendly visualizations, while AiTLAS provides access to the MLM models. <CHNT_Text>

# ADAF Tool results and deployment<CHNT_Heading 1>

The resulting ADAF tool was created successfully and is available for open access for reuse within the research community with two versions available as:

- ADAF Tool trained for Irish archaeological monuments
- ADAF Tool which can be trained upon your own archaeological landscape models

When tested the ADAF tool developed for Irish monuments shows promising results with the most successful MLM components achieving an approximate success rate of between 50-80% in the detection and segmentation of monuments. The ADAF tool has subsequently been released within the TII and Irish archaeology community with several workshops being held to enable the archaeological community to utilise it effectively. <CHNT_Text>

## References<CHNT_Heading 1>


DEFRA (2024) DEFRA Survey Download. Available at : https://environment.data.gov.uk/survey (Accessed 20 June 2024)

Dimitrovski, I., Kitanovski, I., Panov, P., Kostovska, A., Simidjievski, N. and Kocev, D., (2023). Aitlas: Artificial intelligence toolbox for earth observation. Remote Sensing, 15(9), p.2343.

Geological Survey of Ireland (2024) Open Topographic Viewer. Available at: https://dcenr.maps.arcgis.com/apps/webappviewer/index.html?id=b7c4b0e763964070ad69bf8c1572c9f5 (Accessed 20 June 2024).

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G. and Johnson, B.A., 2019. Deep learning in remote sensing applications: A meta-analysis and review. ISPRS journal of photogrammetry and remote sensing, 152, pp.166-177.

Shaw, R, and A Corns. (2011). High resolution LiDAR specifically for archaeology: Are we fully exploiting this valuable resource. Vol. EAC Occasional Paper No. 5, in Remote Sensing for Archaeological Heritage Management, edited by David C Cowley, 77-86. Belgium: Europae Archaeologia Consilium (EAC).

Zhang, L., Zhang, L. and Du, B., (2016). Deep learning for remote sensing data: A technical tutorial on the state of the art. IEEE Geoscience and remote sensing magazine, 4(2), pp.22-40.

ZRC SAZU (2013) Relief Visualization Toolbox (RVT). Available at: https://www.zrc-sazu.si/en/rvt (Accessed 20 June 2024).