New Machine Learning Algorithm Detects Near 10k Archaeological Tumuli in Galicia

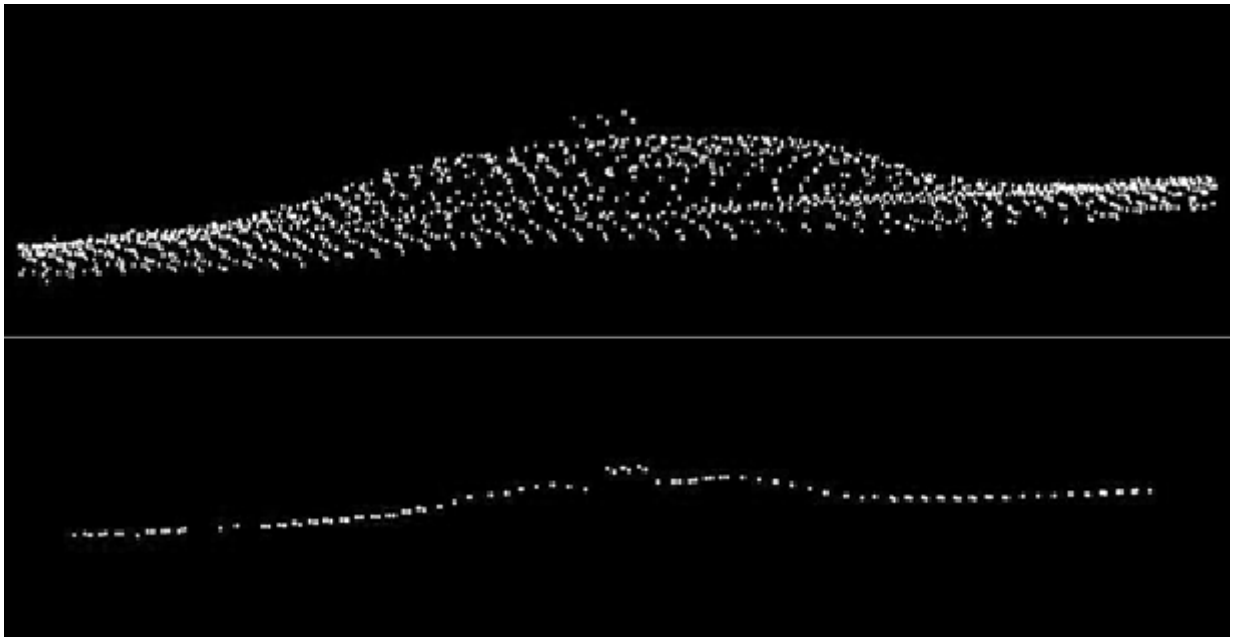
Archaeological tumuli are one of the most common types of archaeological sites and can be found across the globe. This is perhaps why many studies have attempted to develop methods for their automated detection.

Their **characteristic tumular shape** has been the primary feature for their identification on the field and in LiDAR-based topographic data, which usually takes the form of Digital Terrain Models (DTMs).

The simple shape of mounds or tumuli is ideal for their detection using deep learning approaches. Deep learning detectors usually require large quantities of training data (in the order of thousands of examples) to be able to produce significant results.

However, the **homogenously semi-hemispherical shape of tumuli**, allows the training of usable detectors with a much lower quantity of training data, reducing considerably the effort required to obtain it and the significant computational resources necessary to train a convolutional neural network (CNN) detector.

Contingut extret de la web de l'ICAC (www.icac.cat)

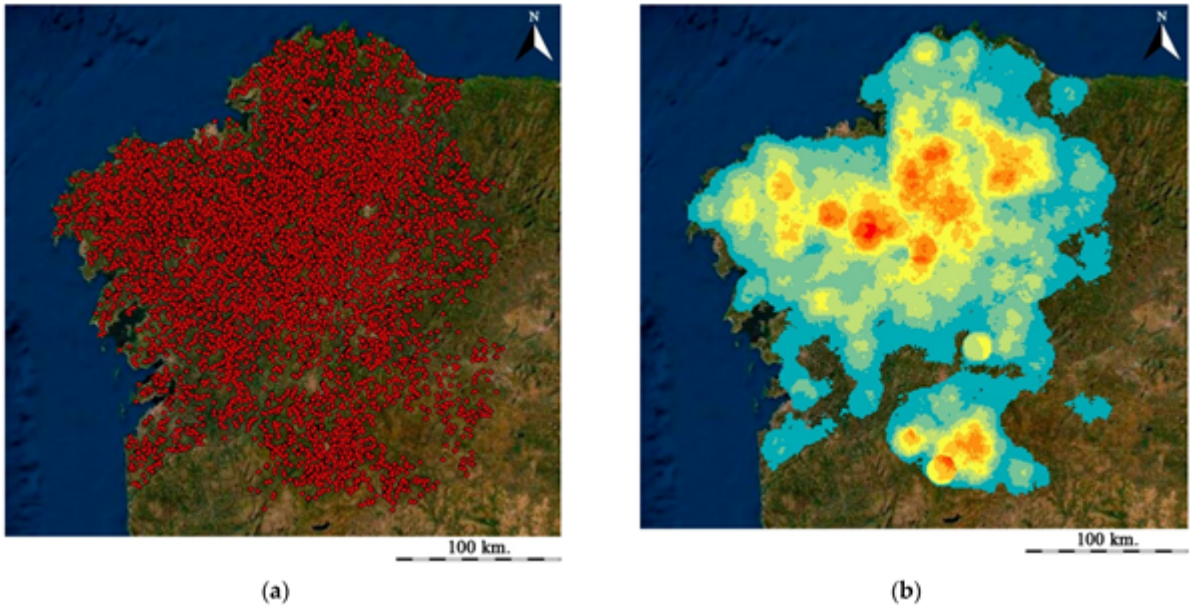


Topographic data based on LiDAR. Image: Miguel Carrero-Pazos.

This type of feature, however, presents an important drawback. Their common, simple, and regular shape is similar to many other non-archaeological features, and therefore studies implementing methods for mound detection in LiDAR-derived DTMs and other high-resolution datasets are characterised by **a very large presence of false positives** (objects incorrectly identified as mounds).

The new algorithm provides a way forward for the detection of tumuli avoiding the inclusion of most false positives.

GIAP researchers Iban Berganzo and Hèctor A. Orengo faced this situation when carrying out the first stage of a study on the automatic detection of burial mounds in Galicia. A blog post was recently posted on this initial investigation, in which almost 9,000 burial mounds were located. However, they were not all real burial mounds, as the results of the automated detection also included false positives.



Detected *tumuli* in Galicia (Spain): (a) point distribution; (b) heat map. Author: Iban Berganzo (ICAC).

After initial data validation was performed in collaboration with our colleagues Dr. Miguel Carrero (University College London / University of Santiago de Compostela, GEPN-AAT), Dr. João Fonte (University of Exeter), and Dr. Benito Vilas (University of Vigo), the team realised that from the ca. 9000 detected objects only ca. 7600 corresponded to real archaeological mounds. Although this was **an excellent result**, well below the percentage of false positives presented by similar studies, researchers thought they could improve the detection rate while decreasing the number of false positives.

During the summer, GIAP members Iban Berganzo and Hèctor A. Orengo, in collaboration with Dr. Felipe Lumbreras from the Computer Vision Center (CVC), developed a new approach to reduce the number of false positives while increasing the detection rate.

After analysing the nature of the detected false positives, they developed **a hybrid approach that mixes classical machine learning and deep learning**. The objective was

to obtain a more precise definition of archaeological tumuli in which not just the shape but also the multispectral characteristics of the objects will be considered when looking for tumuli.

New results are now published as an open access paper in the journal ***Remote Sensing***, one of the top journals in the discipline. In this paper, researchers expand on the analysed data and information on this novel computer-based automatic detection initiative.

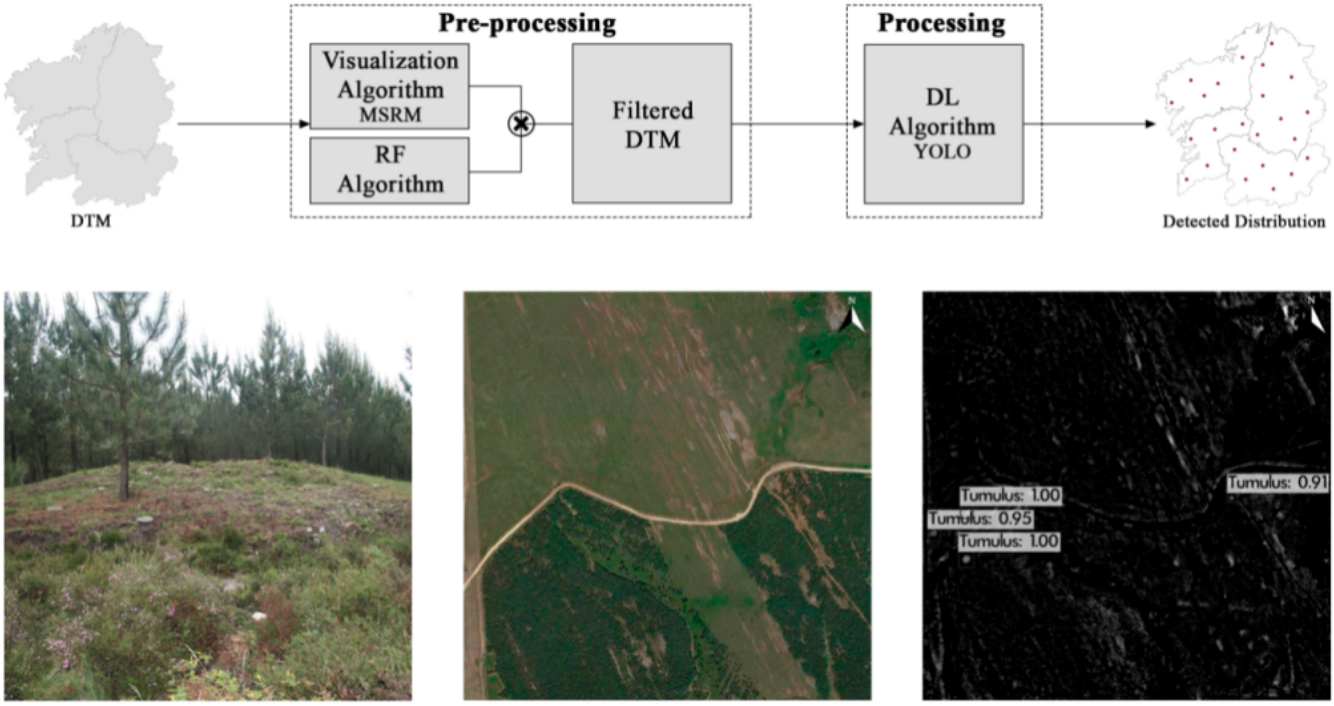


Figure published in the open access paper in *Remote Sensing*, Berganzo et al., 2021: Graphical Abstract.

The results that this new approach has produced are nothing less than spectacular:

- **The area covered is almost 30,000 km^2**: the largest (to the extent of researchers' knowledge) in which archaeological DL approaches have ever been applied.
- 10,527 objects have been detected of which approximately 9,422 correspond to archaeological tumuli (after careful visual validation with high-resolution imagery and

pending ground validation). That is, **89.5% of the detected tumuli correspond to true positives**.

- Researchers only employ **open-source data** in this research. However, the use of higher resolution data, in particular higher resolution satellite imagery instead of the Sentinel 2 (10m/px) images employed, would radically decrease the number of false positives reaching a success rate above 97%.
- Code, sources and results (including validation) are freely available and the code is designed to be used in freely accessible cloud computing platforms Google Collaboratory and Earth Engine) so the lack of computational resources will not pose a problem for its application to other study areas (even very large ones).

  This novel approach provides a way forward for the detection of tumuli avoiding the inclusion of most false positives. **The algorithm can be applied in areas of the world where topographic data of enough resolution are available**. Providing specific training data, this hybrid approach can also be used to detect other types of features where a large number of false positives are an issue.



Tumuli of Touro Morto (Oia, Galicia). Picture: Miguel Carrero-Pazos.

Funding

This research has received funding from multiple sources, that we would like to acknowledge here:

- Iban Berganzo's PhD is funded with a grant named Ayuda a Equipos de Investigación Científica, form the BBVA Foundation for the DIASur Project;
- Hèctor A. Orengo is a Ramón y Cajal fellow at the ICAC (RYC-2016-19637), funded by Spanish Ministry of Science, Innovation, and Universities;
- Felipe Lumbreras' work is supported in part by the Project BOSSS TIN2017-89723-P (Spanish Ministry of Science, Innovation and, Universities).

Contingut extret de la web de l'ICAC (www.icac.cat)

- Miguel Carrero and João Fonte are Marie Skłodowska-Curie fellows (Grant Agreements 886793 and 794048, respectively).
- Some of the GPUs used in these experiments are a donation of Nvidia Hardware Grant Programme.

Contingut extret de la web de l'ICAC (www.icac.cat)