Big Data e inteligencia artificial: una revolución arqueológica

Hèctor Orengo se incorporó en marzo de 2018 en el ICAC como investigador Ramón y Cajal al equipo de investigación Arqueología del Paisaje (GIAP), liderado por Josep M. Palet.

Desde entonces, trabaja en programas de investigación sobre arqueología del paisaje y arqueología computacional. En concreto, está desarrollando **algoritmos de aprendizaje automático** (*machine learning*) para la detección automática de yacimientos.

En colaboración con el McDonald Institute for Archaeological Research (University of Cambridge, Reino Unido), aplica estas técnicas de análisis en la construcción de la paleohidrografía de la civilización del Indo (noroeste de la India), de la época del Bronce (2600-1900 a.C.).

¿En qué consisten, estas técnicas de análisis?

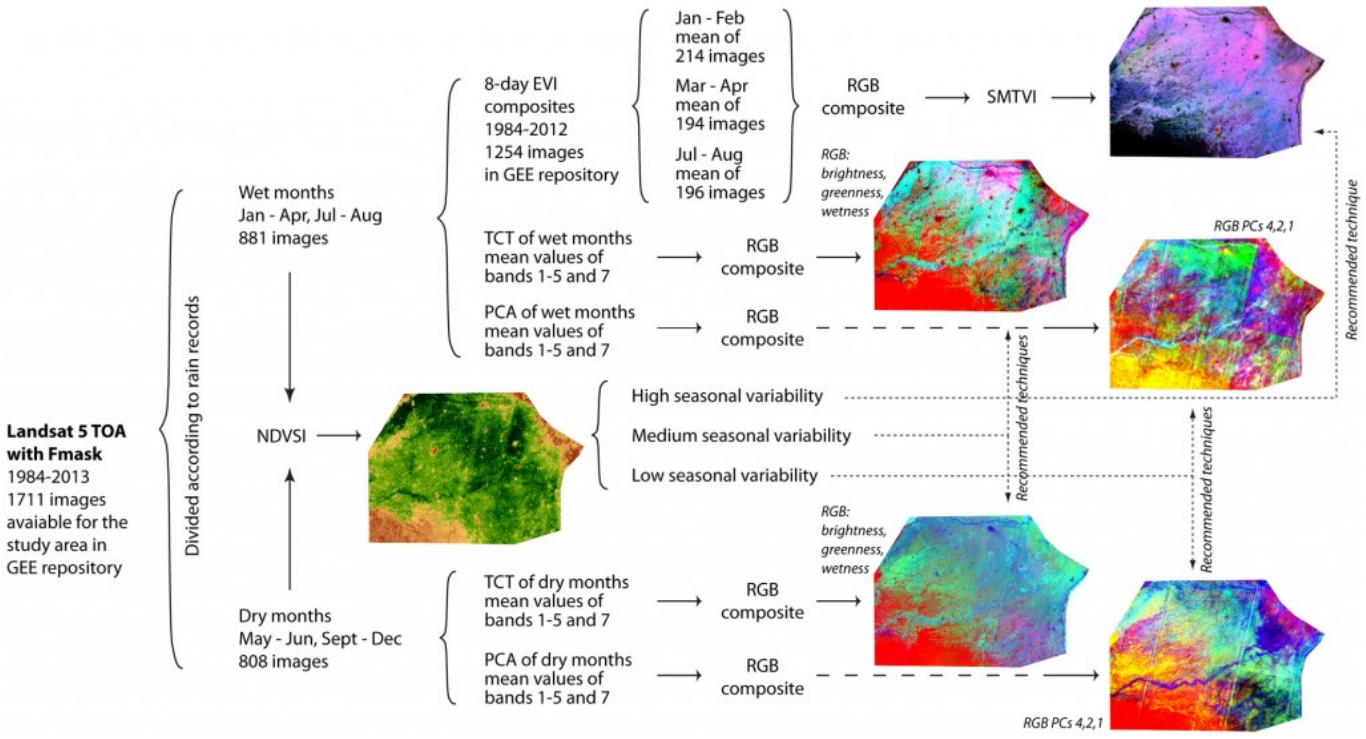
Los algoritmos que hemos desarrollado permiten captar con imágenes aéreas de terreno enormes (casi a escala continental) y transformar los datos, ya sean imágenes satélite o de relieve, y eso mejora significativamente la visibilidad de elementos de interés arqueológico. Concretamente, son dos técnicas que funcionan de manera complementaria y se centran en diferentes características de las estructuras antiguas que pretenden localizar.

Usamos el lenguaje de programación JavaScript porque estamos trabajando con Google Earth Engine (que no es lo mismo que Google Earth) y este es el lenguaje predeterminado en la plataforma en línea. También trabajamos con otros lenguajes, como Python o R.

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

El uso de los algoritmos nos ha permitido reconstruir más de 20.000 km de paleoríos que no son visibles en superficie ni con imágenes satélite sin tratamiento

Uno de los atractivos de trabajar con Google Earth Engine es el inmenso volumen de datos satélite a los que da acceso la plataforma: petabytes! (1 petabyte equivale a un millón de GB). También, los datos satélite disponibles cubren los últimos 40 años de programa espacial. Además, nos permite hacer el procesamiento de estos datos en sus supercomputadores en la nube. Sin este servicio que ofrece la plataforma Google Earth Engine, sería muy complicado descargar la cantidad de imágenes que necesitamos para hacer los análisis, y su procesamiento requeriría disponer de un supercomputador propio.



Orengo & Petrie (fig. 5) (cc) BY-NC-SA

¿Dónde las aplicáis?

En estos momentos nuestros algoritmos se están aplicando con excelentes resultados a muchas problemáticas arqueológicas, como el estudio de paisajes rituales del Neolítico inglés, los tells del primer neolítico Europeo a Tesalia (Grecia) o los paisajes de la colonización griega en el Mediterránea, particularmente en el proyecto de Ampurias en el marco de nuestro proyecto I+D InterArPa (HAR2015-64636-P).

Uno de los objetivos de mi último proyecto postdoctoral era la reconstrucción de la red hidrográfica de la antigua civilización del Indo. Siguiendo esta línea de investigación, el uso de los algoritmos nos ha permitido reconstruir más de 20.000 km de paleoríos que no son visibles en superficie ni con imágenes satélite sin tratamiento.

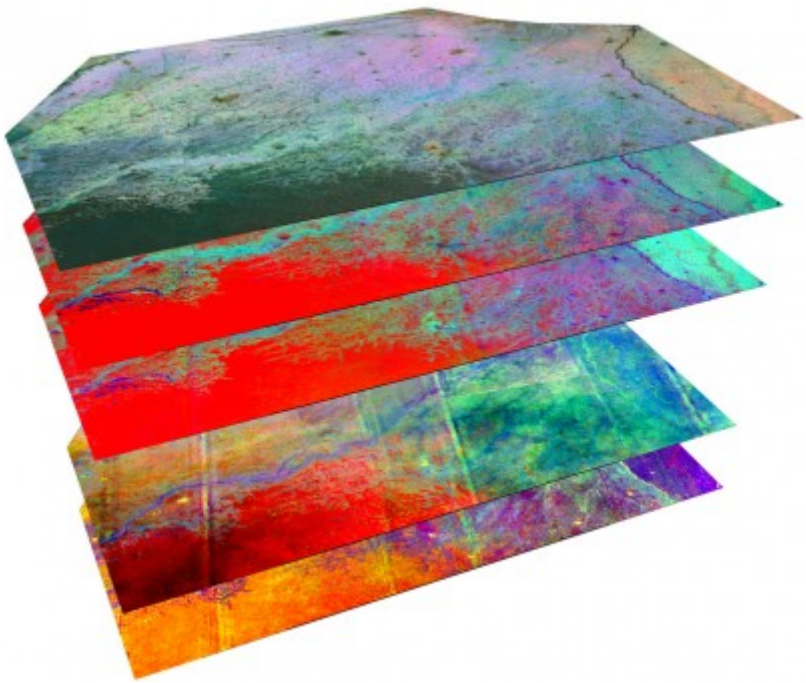
Ahora estamos desarrollando algoritmos de aprendizaje automático y aprendizaje profundo (*deep learning*), muchas veces clasificados como inteligencia artificial

Hemos podido localizar muchos yacimientos desconocidos hasta entonces y extraer importantes conclusiones sobre la forma de ocupación del territorio, la economía y la eventual desaparición de esta civilización.

¿Os habéis fijado nuevos objetivos, a partir de esta experiencia?

Sí, por supuesto. Nos hemos dado cuenta que trabajar con estas escalas tan grandes y con una multiplicidad de datos elevada conlleva un gran esfuerzo de análisis de los elementos de interés arqueológico. Es decir, cuando trabajas áreas que pueden ser más grandes que la Península Ibérica y los restos arqueológicos se tienen que identificar a partir de la combinación de diferentes imágenes resultantes de los algoritmos, resulta que la detección visual de posibles yacimientos puede llevar meses a configurarse, aunque las imágenes muestren claramente elementos de interés.

Por eso ahora estamos desarrollando algoritmos de aprendizaje automático y aprendizaje profundo (*deep learning*), muchas veces clasificados como inteligencia artificial, para analizar estas imágenes y localizar de forma automática los elementos arqueológicos que nos interesan.

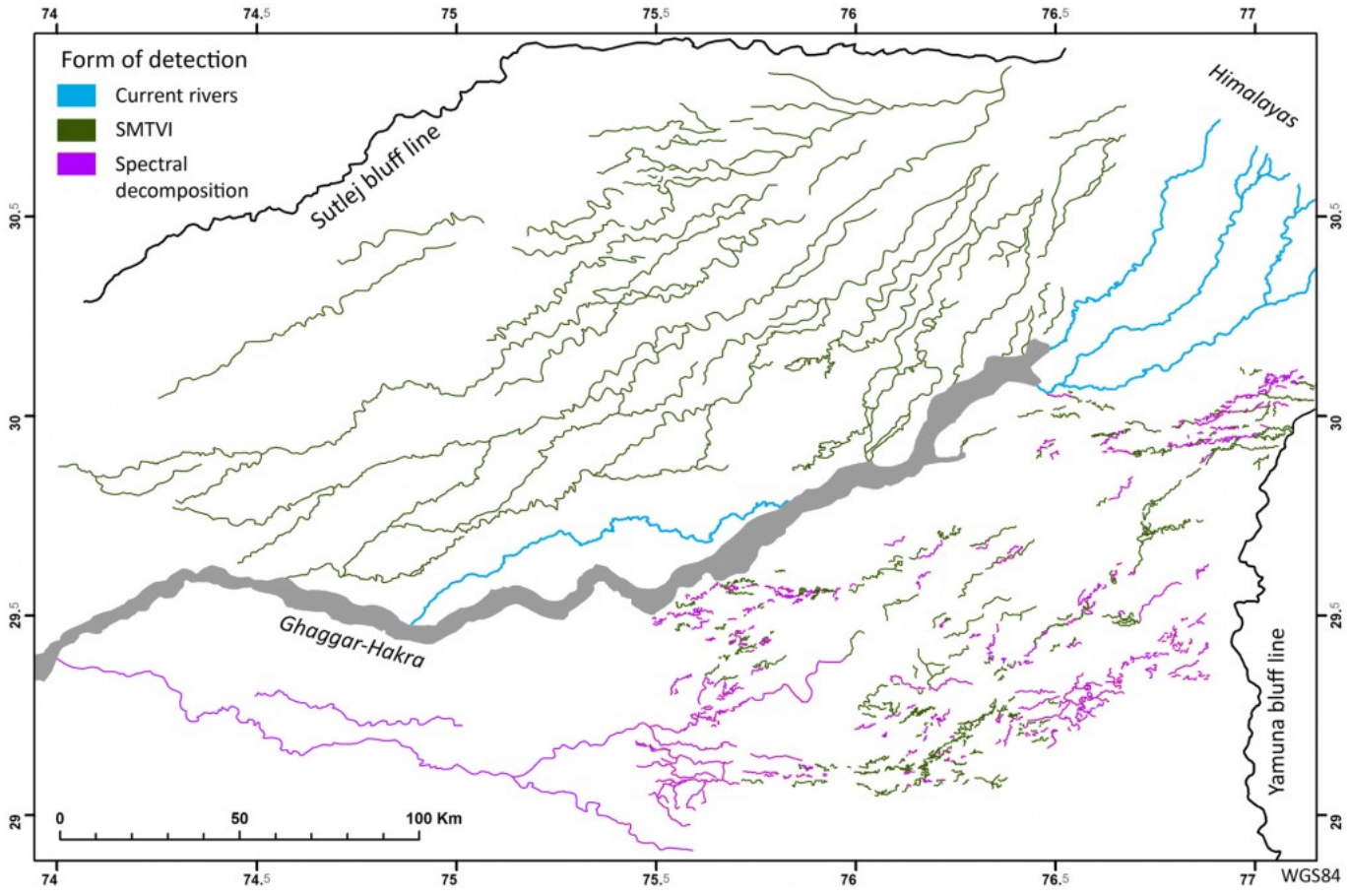


¿Qué importancia pueden tener, estos resultados?

Curiosamente siempre hay problemáticas para las que nuestro trabajo tiene una gran aplicación pero no somos conscientes de ellas cuando desarrollamos las herramientas. En relación a los paleoríos de la India, por ejemplo, no solo tienen una gran importancia para comprender una de las grandes civilizaciones de la edad del Bronce, sino que también pueden representar un recurso hídrico muy importante por su capacidad actual de acumulación de agua. Y eso es un factor de mucha importancia para la India, donde viven millones de personas y donde la producción agraria es una actividad primordial.

Nuestros análisis pueden contribuir a las previsiones de evolución de los acuíferos de la

zona y proporcionar información de gran valor sobre aspectos como la variabilidad de la productividad agraria en el área o la influencia del cambio climático en la expansión del desierto de Thar, además de proveer al gobierno de herramientas de gestión cultural.



Orengo & Petrie (fig. 6) (cc) BY-NC-SA

¿Veremos cada vez más arqueólogos y arqueólogas frente a las pantallas del ordenador?

Me parece que la mayor parte de los arqueólogos ya se pasan una parte importante de su tiempo frente al ordenador. Pero sí, cada vez más la arqueología utiliza técnicas computacionales complejas. De hecho, campos como la arqueobotánica o la zooarqueología están empezando a integrar análisis complejos como la estadística bayesiana o el análisis de redes.

¿Podríamos decir que hay una especialización, dentro de la arqueología?

¡Y tanto! Ahora hay una tendencia a hablar de *arqueología computacional* pero también hay distintos subcampos bastante activos que, en cierto modo, se solapan con ámbitos como la arqueología digital o la arqueología virtual. En realidad, la aplicación de nuevas tecnologías en la investigación arqueológica la encontramos ya en los años 90. Hoy en día hay muchas actividades científicas dedicadas a esta disciplina, como el congreso internacional CAA, Computer Applications and Quantitative Methods in Archaeology, y revistas académicas especializadas, como la *Internet Archaeology* o, más recientemente, la *Journal of Computer Applications in Archaeology*.

Por otro lado, la arqueología computacional forma parte de los programas formativos de muchas universidades; sobre todo en el Reino Unido, pero también en muchos otros países, se ofrecen módulos de grado centrados en distintos tipos de arqueología computacional, y también existen másteres universitarios dedicados específicamente a esta disciplina.

Aquí en Catalunya tenemos investigadores, como Juan A. Barceló y César Carreras (ambos en la UAB), que hace tiempo que desarrollan investigación computacional y que han sido precursores en la península Ibérica en esta materia.



Momento de un seminario en el ICAC. Orengo, a la derecha (cc) BY-NC-SA

Hace poco organizasteis una actividad formativa en el ICAC, en esta línea, ¿verdad?

Sí, el GIAP organizó en el ICAC, del 30 de octubre al 9 de noviembre, unas sesiones de trabajo sobre arqueología computacional avanzada, en el marco del proyecto InterArPa (financiado por el programa I+D del MINECO – HAR2015-64636-P). Trabajamos el uso de teledetección, SIG y otras herramientas geoespaciales. Cada participante desarrollaba una investigación propia para la que la inteligencia artificial o los recursos computacionales complejos resultaban necesarios. El trabajo en común resultó en el desarrollo de metodologías, algoritmos y procesos de gran interés para todos los participantes y un par de

artículos científicos sobre ello ya están en preparación.

El seminario es fruto de un proyecto de investigación liderado desde el ICAC, pero también es el resultado de tus años de experiencia previa

Sí, en mi anterior posición como investigador postdoctoral en el McDonald Institute for Archaeological Research de la Universidad de Cambridge estaba al cargo del desarrollo de nuevas metodologías computacionales para el análisis de paisajes antiguos. Ahora, en el ICAC, quiero seguir con esta línea de trabajo y por eso mi proyecto Ramón y Cajal tiene un fuerte enfoque tecnológico.

- Garcia, Orengo et al. «Mapping Archaeology While Mapping an Empire: Using Historical Maps to Reconstruct Ancient Settlement Landscapes in Modern India and Pakistan». *Geosciences* 2019, *9*(1), 11. Open Access.
- Petrie, Orengo et al. «Remote Sensing and Historical Morphodynamics of Alluvial Plains. The 1909 Indus Flood and the City of Dera Ghazi Khan (Province of Punjab, Pakistan)». *Geosciences* 2019, *9*(1), 21. Open Access.
- Orengo & Petrie. «Multi-Scale Relief Model (MSRM): a new algorithm for the visualisation of subtle topographic change of variable size in digital elevation models». *Earth Surface Processes and Landforms* **2018**, 43(6), 1361-1369. Open Access.
- rengo & Petrie. «Large-Scale, Multi-Temporal Remote Sensing of Palaeo-River Networks: A Case Study from Northwest India and its Implications for the Indus Civilisation». *Remote Sensing* 2017, *9*(7), 735. Open Access.