



Hacia el reconocimiento de evidencias arqueológicas por medio de la aplicación de *machine learning*: una posible nueva herramienta en métodos de prospección

Praxis Arqueológica
Volumen 3
Número 1
Diciembre 2021
Pp. 32-42
DOI 10.53689/pa.v3i1.24

Towards the Identification of Archaeological Elements Through Machine Learning: a New Possible Tool for Archaeological Survey

Javier Astorga, Luis Cornejo, Sebastián Arpón y Gianpiero Canessa*

RESUMEN

Este es un trabajo de carácter exploratorio que pretende impulsar una metodología de prospección arqueológica basada en la visión computacional para clasificar fotografías de objetos arqueológicos de pequeño tamaño en superficie. Para lograr esto, se le enseñó al sistema a reconocer restos arqueológicos por medio de una muestra de fotografías de desechos líticos y fragmentos cerámicos hallados en superficie en campañas de prospección en el Norte Grande de Chile, para luego pedirle que analizara y reconociera otra muestra de los mismos tipos de objetos. Los resultados preliminares demuestran una tasa de error de solo el 6,6% en los ensayos desarrollados. Este trabajo abre la posibilidad de utilizar esta técnica en base a fotografías de baja altura tomadas con drones, lo que permitiría obtener resultados más precisos, menos costosos y en menor tiempo.

PALABRAS CLAVE: arqueología, *machine learning*, métodos, prospección.

ABSTRACT

This is an exploratory work that aims to promote an archaeological survey methodology based on computer vision to classify photographs of small archaeological objects. To achieve this, the system was taught to recognize archaeological remains through a sample of photographs of lithic debris and ceramic fragments found on the surface in prospecting campaigns in the Great North of Chile, and then asked the system to analyze and recognize another sample of the same types of objects. Preliminary results show an error rate of only 6.6% in the developed tests. This work opens the possibility of using this technique based on low-height photographs taken with drones, which would allow it to obtain more precise results, less expensive and in less time.

KEYWORDS: archaeology, machine learning, methods, survey.

* Estudiante de Arqueología, Departamento de Antropología, Universidad Alberto Hurtado, j.astorga.ovando@gmail.com

Departamento de Antropología, Universidad Alberto Hurtado, lcornejo@ahurtado.cl

Matrix Consulting, latinoamericana.sarpon@matrixconsulting.com

Departamento SCI, division Optimization & System Theory, Kungliga Tekniska Högskolan, canessa@kth.se

Introducción

La detección y registro de sitios arqueológicos generalmente se ha basado en la prospección pedestre realizada por arqueólogos, siendo esta la fase primaria en la investigación científica en arqueología. La técnica de prospección pedestre puede ser definida como revisión sistemática de la superficie de una unidad de prospección por un equipo de observadores (Banning, 2002; Gallardo y Cornejo, 1986; véase también Schiffer et al., 1978; Schiffer y Wells, 1982). Reconocer, definir e interpretar un sitio arqueológico es un paso crítico en etapas analíticas de la investigación: puede presentar diferentes sesgos, ya sea por variables ambientales, características intrínsecas del registro arqueológico y por el sesgo operador dependiente. Tales sesgos afectan directamente en la «visión de la variabilidad, distribución, y procesos de formación del registro arqueológico» (Cornejo et al., 1988, p. 153). A su vez, hay que considerar la naturaleza misma de la prospección arqueológica. La atención que presta el/la investigador(a) al entorno es inversamente proporcional al cansancio que este pueda tener. Esto quiere decir que, a mayor agotamiento físico y mental por parte del operador, la posibilidad de no ver un sitio arqueológico, u obviar restos materiales en superficie, aumenta. Por otro lado, factores tales como la obstrusividad, visibilidad, y accesibilidad (Gallardo y Cornejo, 1986) son determinantes a la hora de registrar material cultural presentes en el área de estudio. En algunos casos el/la investigador(a) puede confundir o no visualizar los restos arqueológicos que pretende registrar, producto de los factores ya mencionados. Así pues, con el desarrollo de *machine learning*, sesgos del tipo operador dependiente tales como confundir termofracturas con restos de desecho de talla o de accesibilidad, propios de la necesidad de llegar a un determinado lugar, podrán ser superados, ya que eventualmente esta herramienta permitirá clasificar correctamente, y en un corto período de tiempo, dichas materialidades. El presente trabajo abordará la problemática de clasificar correctamente materiales líticos y cerámicos hallados en el Norte Grande de Chile, espacio caracterizado por las condiciones climáticas del tipo desértica, con baja precipitación e hiperaridez. Estos factores ambientales y atmosféricos son idóneos para la preservación de restos arqueológicos, como también para los procesos de formación de sitio por agentes naturales (Schiffer, 1987). El proceso de clasificación de *machine learning* será realizada por métodos supervisados, analizando registros fotográficos manuales de las materialidades anteriormente señaladas *in situ*, tomadas en terreno por arqueólogos(as) a una altura promedio de 1,5 m, en diferentes sitios del Norte Grande de Chile.

Bajo el presente marco contextual global, el uso de *machine learning* es una herramienta inagotable, modificable, que podría ser capaz de detectar y clasificar restos arqueológicos con bastante precisión.

Nuevas técnicas de identificación de sitios arqueológicos

En las últimas de décadas han surgido varias alternativas a la clásica prospección pedestre. Dentro de estas posibilidades metodológicas, destaca la medición de radar de penetración de suelo (GPR). Esta técnica geofísica aplicada a la arqueología (Trinks et al., 2018) destaca por ser no invasiva y por detectar materiales que pueden estar hasta 25 cm bajo el suelo. La tasa de cobertura de esta herramienta permite analizar en un día de trabajo, con dos o tres personas, hasta media hectárea (5.000 m²) de extensión. La prospección por GPR ha demostrado tener una enorme eficacia en el levantamiento y la resolución de muestreo espacial. Su

uso extendido en arqueología se debe a la capacidad de obtener imágenes de alta resolución de estructuras arqueológicas relativamente pequeñas (Trinks et al., 2018, p. 21). Una forma de complementar este tipo de metodología es mediante el uso de tomografía de resistividad eléctrica (ERT), que consiste en un método exploratorio geofísico usado para evaluar las propiedades eléctricas del subsuelo por medio de la inyección de corriente continua a la tierra (Gołębowski et al., 2018, p. 1). En un perfil se pueden instalar hasta 60 electrodos, separados por 1,5 m entre sí, cada electrodo mide la diferencia de potencial, calculando así una resistividad aparente del suelo. Este tipo de técnica medianamente invasiva es muy útil para detectar objetos arqueológicos enterrados, localización de aguas subterráneas o anomalías en el subsuelo. Sin embargo, estas herramientas están mejor adaptadas a pequeñas áreas y su función está orientada a detectar macroestructuras y no a detectar piezas pequeñas, como fragmentos cerámicos o líticos.

Otra forma de realizar una prospección en áreas cuya accesibilidad es difícil, o se tenga poca visibilidad respecto de la superficie, es usando tecnología LiDAR, que consiste en la obtención de imágenes aéreas mediante la utilización de pulsos de rayo láser, permitiendo crear imágenes tridimensionales (3D) precisas sobre el terreno (Cerrillo-Cuenca y Bueno-Ramírez, 2019, p. 1). La preparación de datos LiDAR incluye la eliminación atípica de los conjuntos de datos, también filtra las copas de los bosques, genera de modelos de terreno y crea imágenes con relieve sombreado del área explorada (Risbøl y Gustavsen, 2018, p. 6). En términos generales, este tipo de tecnología funciona especialmente bien en áreas muy boscosas y con poca visibilidad, sin embargo, su tecnología de toma de datos mediante puntos, generalmente asociada al uso de dron, dificulta la posibilidad de hallar restos arqueológicos pequeños.

Una tercera posibilidad de prospección es mediante la utilización de análisis de imágenes basados en objetos (OBIA) (Davis, 2018, p. 1), que corresponde a un método de evaluación de datos por teledetección, el cual utiliza parámetros morfométricos y espectrales simultáneamente para identificar diversas características en las imágenes de teledetección en terreno (Davis, 2018, p. 1-2). OBIA se basa en prácticas de larga data de análisis de detección remota que incluyen segmentación, detección de bordes y clasificación. Una limitación de OBIA es que requiere conjuntos de datos de muy alta resolución para funcionar de manera efectiva; esto quiere decir que, con un mayor número de variables (incluidas múltiples escalas de análisis), mejora la precisión y resolución de sus resultados. Sin embargo, los primeros estudios de OBIA tuvieron una tasa muy alta de falsos positivos y falsos negativos (De Laet et al., 2007, 2008, 2009; Menze et al., 2006, 2007; Trier et al., 2009; Davis, 2018).

Los métodos anteriormente descritos resultan poco útiles al momento de detectar restos arqueológicos fragmentados y de pequeño tamaño, tales como dispersiones líticas o de fragmentos de cerámica. El presente trabajo pretende aportar en dicha línea, explorando una nueva metodología que permita vencer esta barrera, disminuyendo a la vez el sesgo operador dependiente de la prospección pedestre. Esta metodología es de carácter no destructivo y se basa en *machine learning* entrenada por imágenes, la cual permitiría identificar restos arqueológicos con alta precisión, mejorando significativamente la calidad de la información obtenida, en un menor tiempo y con menores costos.

Las primeras publicaciones en arqueología sobre el uso de *machine learning* datan desde mediados del 2006 (Davis, 2018, p. 3). Sin embargo, estos estudios iniciales presentaron una alta tasa de falsos positivos,

y su objetivo de trabajo fue satisfacer problemáticas de contextos europeos, tales como detectar objetos lineales asociados a caminos romanos o detección de megalíticos (Davis, 2018). La situación en América Latina respecto del uso de *machine learning* es completamente nueva e inexplorada. Si bien el uso de imágenes satelitales y tecnología LiDAR se ha expandido, el reconocimiento automático de objetos es un área sin referentes en contextos continentales y pocos autores han abordado esta herramienta a nivel mundial. Ciertos estudios han logrado avanzar en temáticas específicas de la arqueología en conjunción con *machine learning*, el primero de ellos a mencionar corresponde a un estudio desarrollado en Francia (Guyot et al., 2017). El uso de *machine learning* se combinó con tecnología LiDAR para obtener imágenes aéreas y poder discriminar la presencia de túmulos funerarios, utilizando análisis multiescala de posición topográfica combinado con algoritmos de *random forest*. En lugar de resaltar anomalías topográficas individuales (propio de LiDAR), el enfoque multiescalar permite que las características arqueológicas se examinen no solo como objetos individuales, sino dentro de su contexto espacial más amplio. No obstante, esta metodología implica analizar pixel por pixel la fotografía (Guyot et al., 2017). Un estudio similar fue replicado en el desierto de Cholistán, Pakistán, con el fin de hallar una mayor cantidad de túmulos funerarios (Orengo et al., 2020). La particularidad de este estudio es que combinó el uso de imágenes satelitales de alta resolución a modo de complemento con *machine learning*, obteniendo buenos resultados en el área de detección.

El tercero de los casos a destacar corresponde al uso de *machine learning* en contextos de excavación italianos. El algoritmo utilizado permitió el aprendizaje automático no supervisado de *k-means* para agrupar imágenes digitales en color de los sitios de excavación, logrando, además, la clasificación y discriminación de las distintas capas estratigráficas. Los autores concluyen que el uso de esta herramienta puede ser de bastante ayuda para disminuir los tiempos y porcentaje de error a la hora de describir y caracterizar los distintos perfiles de excavación (Cacciari et al., 2018). El cuarto caso corresponde al uso de *machine learning* en Harz Mountains, Alemania. El modelo fue entrenado para reconocer arroyos, lagos, senderos, entre otros, sobre imágenes de 2x2 km² (Kazimi et al., 2018, p. 22).

Finalmente, en Abdera, Grecia, se implementó el uso de drones para fotografiar un área determinada con la finalidad de hallar concentraciones de restos cerámicos, similar a nuestro objetivo de estudio. El uso de esta herramienta es un complemento a la prospección de tipo pedestre. El algoritmo utilizado para *machine learning* correspondió al análisis de imágenes multiespectrales, logrando realizar predicciones a nivel de pixel (Orengo y Gracia-Molsoa, 2019).

Metodología

Para el desarrollo de esta metodología, se consideró fundamental encontrar un sistema que posea dos características:

- Sea basado en imágenes de baja altura, las cuales pueden ser capturadas, por ejemplo, por drones.
- Sea capaz de clasificar qué objetos están presentes en la imagen.

Para construir este sistema, se utilizó el área de visión computacional, que consiste en el estudio de imágenes utilizando algoritmos computacionales (Voulodimos, 2018). En el presente trabajo, los algorit-

mos de interés son aquellos que permiten la clasificación de imágenes. Los sistemas de clasificación de imágenes son aquellos que, basados en una imagen, permiten establecer a qué categoría pertenece. Las distintas categorías son determinadas según el interés arqueológico u objetos de investigación (por ejemplo: lítico-cerámica). Así, integrando este algoritmo, podemos satisfacer las dos propiedades antes descritas.

Esta propuesta se basa fuertemente en el aprendizaje de máquinas que en general es separado en dos áreas: el aprendizaje supervisado y el no supervisado. La diferencia principal entre ambos es que en el aprendizaje supervisado se conoce *a priori* qué es lo que se quiere predecir, lo cual se almacena en una variable conocida como «etiqueta» (también se le puede referir como «clase» en la literatura). Por otra parte, en el aprendizaje no supervisado, en general, no existe una variable «etiqueta». La problemática abordada por este trabajo consiste en clasificar si una imagen corresponde a fragmentos líticos o fragmentos cerámicos; así, existe una etiqueta para cada imagen que indica si pertenece a fragmentos líticos o cerámicos, de manera puede ser abordada desde la perspectiva del aprendizaje supervisado. Por esta vía, se pueden encontrar distintas técnicas, algunas basadas en metodologías como son los *random forest* (Xu, 2012), *support vector machines* (Sun, 2015) y otras basadas en redes neuronales (Goodfellow, 2013), además existen técnicas que permiten hacer la clasificación de imágenes desde el aprendizaje no supervisado (Ji, 2019). Es importante notar que las redes neuronales son una amplia familia de técnicas que comparten un patrón de diseño, como se muestra en la Figura 1. Todas utilizan capas de entrada de datos (*input layer*), que se encargan de entregar los datos a la red neuronal; luego las capas ocultas (*hidden layers*), que son lo que se conoce como arquitectura de la red, encuentran patrones que contribuyan a generar una buena predicción –notar que estas capas entregan la información hacia a la izquierda posterior a que realizan su procesamiento–; finalmente, la última capa es la capa de salida (*output layer*), la cual suministra la predicción realizada por la red neuronal.

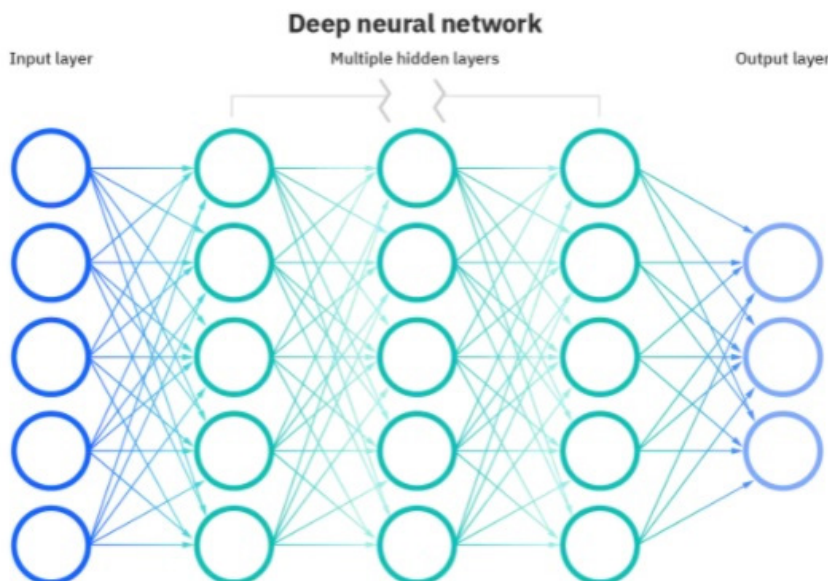


Figura 1. Representación de una arquitectura de red convolucional. Representación de IBM.

Todas estas redes se diferencian por las arquitecturas que poseen, y cada una favorece el aprendizaje de distintos patrones en los datos o distintos tipos de eficiencias. Dentro de las redes neuronales, se encuentran las redes convolucionales (Li, 2014), las cuales utilizan ventanas móviles dentro de los datos para encontrar patrones. Esto permite que los patrones aprendidos por la red neuronal puedan ser identificados en cualquier lugar de los datos, razón por la que son ampliamente utilizadas en visión computacional. Las LSTM (Graves, 2005) que evalúan si la información aprendida en las capas más cercanas a la de *input* siguen entregando información útil a las capas de salida, razón por la cual son ampliamente ocupadas en procesamiento de texto, dada la recursividad de la información a la hora de leer un texto. Por último, es importante mencionar las redes residuales (He, 2016) que cada un cierto número de capas recuperan completamente la información de entrada de capas más previas (cercanas a la de *input*). En particular, la red residual propuesta en He (2016) contiene capas convolucionales y muestra un mejor desempeño que su equivalente en redes convolucionales, motivo por el que fue elegida para desarrollar este estudio.

Dentro del *machine learning* existe una constante preocupación por un fenómeno conocido como el *overfitting*, el cual implica que el sistema está sobredescribiendo los datos, por tanto el aprendizaje que logra no es generalizable para datos que nunca antes haya utilizado. Esto se debe a que los datos disponibles son separados en dos conjuntos: entrenamiento y validación. Para construirlos lleva a cabo un muestreo aleatorio de la data completa donde un cierto porcentaje es asignado al conjunto de entrenamiento y el resto es asignado al conjunto de validación. El porcentaje asignado al conjunto de entrenamiento típicamente va entre un 60% y un 90% (Provost y Fawcett, 2013; Ng, 2018), dado que el conjunto de entrenamiento es el utilizado para que el sistema aprenda los patrones que existen en los datos, por tanto es beneficioso que tenga una cantidad superior. Por otro lado, el conjunto de validación es utilizado para validar que los patrones aprendidos en los datos de entrenamientos sean generalizables a datos nunca antes vistos.

Para desarrollar un sistema de visión computacional, previamente se requiere de la captura de datos, lo cual permite el aprendizaje del sistema. Para realizar las pruebas se recolectaron imágenes obtenidas por arqueólogos en sitios arqueológicos: 324 imágenes con desechos líticos y 55 imágenes presentaron fragmentos cerámicos. El proceso de entrenamiento consistió en alimentar estos datos en el clasificador para que la inteligencia aprenda de ellos. Asimismo, estos datos fueron separados en dos conjuntos: conjunto de entrenamiento y conjunto de validación (Provost y Fawcett, 2013). Se utilizaron 260 imágenes de líticos y 44 de cerámica como conjunto de entrenamiento. Para poner a prueba el algoritmo, se emplearon como conjunto de validación 64 imágenes para reconocer líticos y 11 para cerámica.

Estas fotografías destinadas a alimentar la red neuronal corresponden a registros sacados *in situ* en campañas de prospección dentro contextos propios del Norte Grande de Chile. La mayoría de estas fotografías se obtuvieron manualmente por parte de distintos arqueólogos(as), cuya altura promedio de foco fue de 1,5 m. Las consideraciones o criterios aplicados para que las fotografías fuesen seleccionadas con el objetivo de que la *machine learning* aprendiera fueron los siguientes: primero, fotografías tomadas en el desierto del Norte Grande de Chile, ya que este tipo de terreno es idóneo por la gran visibilidad que tienen las piezas arqueológicas, gracias a la casi total ausencia de vegetación; segundo, que las fotografías no tuviesen un elemento externo al sitio, tales como escalas o «norte», ya que es necesario que el aprendizaje

desde los datos sea lo más preciso posible, y elementos externos pueden llevar a la confusión y al aumento en la tasa de error.

El sistema clasificador de imágenes implementado fue una red neuronal convolucional llamada Res-Net34 (He, 2016). Esta red consiste en una arquitectura de 34 capas, creada para identificar categorías de imágenes a partir de aproximadamente 1.000 posibles categorías. Para aplicarla a nuestra propuesta, se elimina la última capa y se reentrena la red utilizando los datos de entrada previamente descritos para obtener la salida correspondiente. Para evaluar el desempeño del algoritmo, se utilizó la métrica de la tasa de error que representa el porcentaje de los datos que fueron mal identificados versus el total de datos de validación.

Resultados

La Figura 2 muestra la matriz de confusión obtenida en el conjunto de validación. Esta representación permite evaluar los distintos tipos de errores estadísticos producidos por el algoritmo ya entrenado. El eje vertical presenta los valores reales de las imágenes y el eje horizontal el valor detectado por el algoritmo (también conocido como valor predicho) por el sistema.

En cada uno de los recuadros se muestra, en la matriz de confusión, el número de imágenes que pertenecen a ese par (valor predicho, valor real). Así, por ejemplo, 7 fueron predichas como cerámica y eran realmente cerámica; a su vez, 4 imágenes fueron predichas como cerámica y eran líticas.

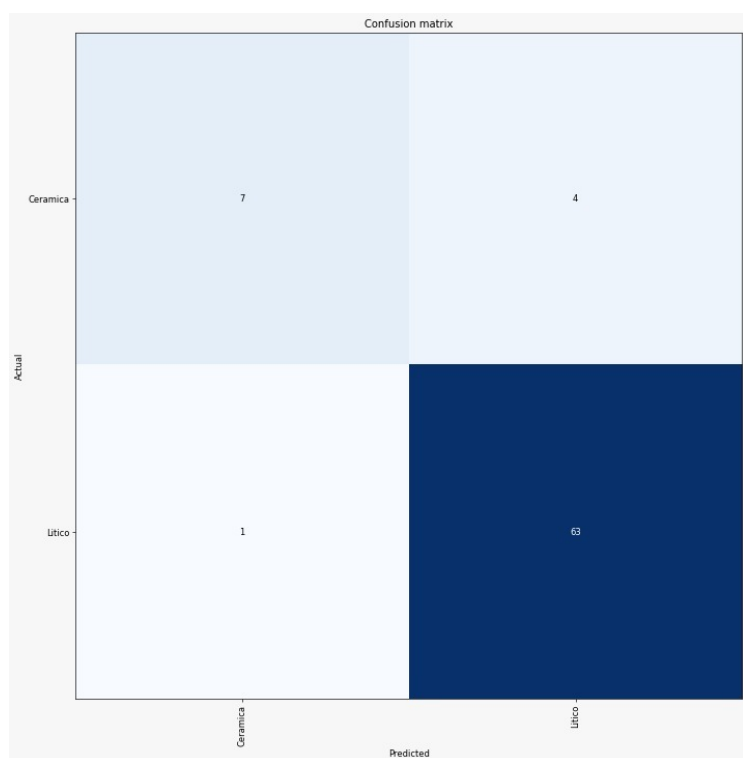


Figura 2. Matriz de confusión de estudio exploratorio.

Se ha mostrado que un desbalance entre la cantidad de datos de cada clase tiene un impacto negativo en el desempeño de redes neuronales convolucionales (Buda, 2018), el cual puede ser resuelto consiguiendo más datos de aquella clase con menor cantidad de muestras. Así, dada la diferencia numérica utilizada en el conjunto de entrenamiento de líticos y cerámicos, se desprende que obtener más imágenes correspondientes a cerámicos disminuiría el porcentaje de error en la clasificación de estos materiales. Dentro de las cuatro imágenes resaltadas como errores de clasificación (Figura 3), una de ellas presentaba tanto restos líticos como cerámicos, llevando a un error de clasificación. Sin embargo, las problemáticas de fotos que presenten dos o más materialidades serán abordadas en trabajos posteriores.

Figura 3.

Fotos mal clasificadas por el sistema entrenado.



De esta manera, el aprendizaje de la *machine learning* muestra una tasa de éxito del 93,3% en lograr clasificar elementos cerámicos y elementos líticos.

Discusión y conclusiones

El estudio aquí presentado es el primer ensayo formal latinoamericano exitoso en el uso de *machine learning* para la clasificación de restos arqueológicos, con una tasa de éxito de 93,3% en el conjunto de validación, lo que demuestra una alta capacidad del sistema a la hora de diferenciar entre restos líticos y cerámicos. A nivel global, ya hay algunos indicios en la aplicación de esta herramienta para problemas arqueológicos, pero nuestra alternativa utiliza los últimos avances en el campo de visión computarizada. Si bien el trabajo de Orengo y Gracia-Molsosa (2019) pretende los mismos objetivos, se diferencia de nuestro estudio al recurrir a un análisis multiespectral, analizando pixel por pixel cada imagen. No obstante, esta metodología está cada vez más en desuso en las áreas de visión computarizada, ya que la clasificación por píxeles no considera la posición, ni valores, ni formas presentes en los píxeles en la vecindad de la imagen que está siendo clasificada; utilizando así una menor cantidad de información de la que existe en la imagen, lo cual implica una mayor cantidad de errores de clasificación. La ventaja de usar una imagen entera para este estudio en particular es que el algoritmo tiene que discriminar por formas o atributos de los elementos, obteniendo resultados de mayor definición y grado de certeza.

La mejora continua de esta metodología basada en *machine learning* de clasificación de elementos arqueológico es directamente proporcional a la cantidad de imágenes que se usen para que la red neuronal aumente su sensibilidad en el reconocimiento de objetos. Esto quiere decir que, mientras más imágenes se muestren a la red neuronal, más fácil podrá clasificar a qué grupo pertenece el elemento arqueológico. En la muestra estudiada, el error en la clasificación de las fotografías por parte de la *machine learning* se presume por la combinación entre la poca cantidad de ejemplos cerámicos disponibles con los que se entrenó la *machine learning*, como también por el factor de obstrusividad, donde los materiales líticos tienen colores similares con restos cerámicos. Así, de este error se desprende que, a mayor cantidad de imágenes de cerámicas en distintos entornos, el error de clasificación debería disminuir.

Es posible pensar en un futuro en el que el sesgo operador dependiente pueda ser controlado recurriendo a metodologías como estas. Este futuro, de hecho, no es muy lejano, ya que al utilizar drones como los actualmente existentes, es posible recolectar sistemáticamente *sets* de fotografías sobre transectos predefinidos a una intensidad alta, las cuales pueden ser sometidas a un análisis como el aquí propuesto. Es necesario precisar que los drones actuales tienen en varios casos una velocidad de operación de 10 m/s, pudiendo recorrer 36 km de transectos en una hora de vuelo. Si bien, el costo inicial para adquirir uno es elevado, resulta posible calcular que el dron es capaz de cubrir una cantidad de terreno similar a 7 arqueólogos(as) prospectando, por ejemplo, en el desierto de Atacama. Este estimado se construye considerando que en una hora estos prospectan 5 km.

Esta herramienta, además de disminuir el sesgo operador dependiente mediante la clasificación automática ya señalada, especialmente en relación con la accesibilidad, la visibilidad y la obstrusividad, permi-

tirá un mejor uso de los recursos claves implicados en una prospección (tiempo, financiamiento y personal). No obstante, el desarrollo de una prospección requiere recolectar una serie de información específica para caracterizar los hallazgos, que requerirá siempre las habilidades de un arqueólogo(a) experimentado.

La metodología propuesta solo puede ofrecer una identificación de lo que hay en el terreno por medio de *machine learning*, diferenciando restos cerámicos de restos líticos; por otro lado, la georreferencia de lugar y posición es proporcionada por el dron. Si bien, estas nuevas tecnologías permitirían optimizar el trabajo en terreno, esto no quiere decir que se podrá prescindir de la prospección arqueológica propiamente tal. Dentro de los criterios para realizar una completa descripción de materiales arqueológicos en una prospección, generalmente se requiere identificar rasgos en los objetos culturales que no son visibles en la foto, tales como la cara oculta de un fragmento de cerámica o la identificación de propiedades que requieren una mirada muy cercana como, por ejemplo, la caracterización de la materia prima de un instrumento lítico. Junto con esto, la asistencia humana en el lugar permite reconocer rasgos, ya sean naturales o culturales, que no han sido enseñados a la red neuronal. Esto al menos mientras la inteligencia artificial no presente avances sustantivos.

De esta manera, la herramienta en desarrollo facilitará el reconocimiento de los principales elementos del registro arqueológico de una localidad, lo cual posteriormente deberá ser complementado con el trabajo en terreno de arqueólogos(as) que amplíen la mirada más allá de las capacidades del sistema, pero dirigidos directamente a los lugares donde la máquina identificó los materiales.

Finalmente, es posible afirmar que la aplicación de *machine learning* y la visión computacional en arqueología no se limitará solo a esta esfera, sino que es evidente que puede ofrecer ayuda en distintos aspectos, donde la descripción del mundo empírico está dejada en manos de personas que deben observar los atributos de la realidad y reconocer en ellos determinados patrones que permiten clasificarlos de alguna manera, donde la variabilidad de dichas personas introduce un sesgo operador dependiente en la configuración de la base analítica de nuestra disciplina.

Conflicto de interés

Los autores declaran que no existen conflictos de interés respecto de la publicación del presente artículo.

Referencias citadas

- Banning, E. B. (2002). *Archaeological Survey* (vol. 1). New York, Estados Unidos: Springer Science & Business Media.
- Buda, M., A. Maki y M. A. Mazurowski (2018). A Systematic Study of the Class Imbalance Problem in Convolutional Neural Networks. *Neural Networks*, 106, 249-259.
- Cacciari, I., G. Pocobelli, S. Cicola y S. Siano (2018). Discrimination of Soil Texture and Contour Recognitions During Archaeological Excavation Using Machine Learning. *10P Conf. Series: Materials Science and Engineering*, 364, 1-8.
- Cerrillo-Cuenca, E. y P. Bueno-Ramírez (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). *Research Article. Archaeological Prospection*, 26(3), 251-264.
- Davis, D. (2018). Object-Based Image Analysis: a Review of Developments and Future Directions of Automated Feature Detection in Landscape Archaeology. *Archaeological Prospection*, 26(2), 155-163.
- Cornejo, L., F. Gallardo y B. Caces (1988). Estrategias de muestreo para la recolección superficial en sitios arqueológicos. *Actas del XI Congreso Nacional de Arqueología Chilena* (pp. 153-163). Santiago, Chile.
- Gallardo, F. y L. Cornejo (1986). El diseño de la prospección arqueológica: un caso de estudio. *Chungara*, 16, 409-420.
- Gołębowski, T., B. Pasierb, S. Porzucek y M. Łoj (2018). Complex Prospection of Medieval Underground Salt Chambers in the Village of Wiślica, Poland. *Archaeological Prospection*, 25(3), 243-254.
- Goodfellow, I., D. Warde-Farley, M. Mirza, A. Courville y Y. Bengio (2013). Maxout Networks. *Proceedings of the 30th International Conference on Machine Learning. PMLR*, 28(3), 1319-1327.
- Graves, A., S. Fernández y J. Schmidhuber (2005). Bidirectional LSTM Networks for Improved Phoneme Classification and Recognition. En W. Duch et al. (Eds.), *ICANN, LNCS*, 3697, 799-804.
- Guyot, A., L. Hubert-Moy y T. Lorho (2017). Detection Neolithic Burial Mounds from LIDAR-Derived Elevation Data Using a Multi-Scale Approach and Machine Learning Techniques. *Remote Sens*, 10, 1-19.
- He, K., X. Zhang, S. Ren y J. Sun (2016). Deep Residual Learning For Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770-778.
- Ji, X., J. F. Henriques y A. Vedaldi (2019). Invariant Information Clustering for Unsupervised Image Classification and Segmentation. *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 9865-9874.
- Kazimi, B., F. Thiemann, K. Malek, M. Sester y K. Khoshelham (2018). Deep Learning for Archaeological Object Detection in Airborne Laser Scanning Data. *Conference GIScience*, 18, 21-35.
- Li, Q., W. Cai, X. Wang, Y. Zhou, D. D. Feng y M. Chen (2014). Medical Image Classification with Convolutional Neural Network. *13th International Conference On Control Automation Robotics & Vision (icarcv)*, 844-848.
- Orengo, H. y A. Garcia-Molosa (2019). A Brave New World for Archaeological Survey: Automated Machine Learning-Based Potsherd Detection Using High-Resolution Drone Imagery. *Journal of Archaeological Science*, 112, 1-12.
- Orengo, H. et al. (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), 1-11.
- Provost, F. y T. Fawcett (2013). *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. California, Estados Unidos: O'Reilly Media.
- Risbøl, O. y L. Gustavsen (2018). LiDAR from Drones Employed for Mapping Archaeology. Potential, Benefits and Challenges. *Archaeological Prospection*, 25(4), 329-338.
- Schiffer, M. B. y S. Wells (1982). *Archaeological Survey: Past and Future*. En R. McGuire y M. Schiffer (Eds.), *Hohokam and Patayan: Prehistory of the Southwestern Arizona* (pp. 345-382). Nueva York, Estados Unidos: Academic Press.
- Schiffer, M. B., A. P. Sullivan y T. C. Klinger (1978). The Design of Archaeological Surveys. *World Archaeology*, 10(1), 1-28.
- Sun, X., L. Liu, H. Wang, W. Song y J. Lu (2015). Image Classification Via Support Vector Machine. *4th International Conference on Computer Science and Network Technology (ICCSNT)*, 1, 485-489.
- Trink, I. et al. (2018). Large-Area High-Resolution Ground-Penetrating Radar Measurements for Archaeological Prospection. *Archaeological Prospection*, 25(3), 171-195.
- Voulodimos, A., N. Doulamis, A. Doulamis y E. Protopapadakis (2018). Deep Learning for Computer Vision: A Brief Review. *Computational Intelligence and Neuroscience*, 2018, 1-13.
- Xu, B., Y. Ye y L. Nie (2012). An Improved Random Forest Classifier for Image Classification. *IEEE International Conference on Information and Automation*, 795-800.

Recibido: 2/12/2020

Aceptado: 7/10/2021