

## Detección y reconocimiento automático de jeroglíficos egipcios mediante visión por computador y validación cruzada

Fuentes-Ferrer, R.<sup>a</sup>, Duque-Domingo, J.<sup>b</sup>, Herrera, P.J.<sup>a,\*</sup>

<sup>a</sup> Dpto. de Ingeniería de Software y Sistemas Informáticos, Universidad Nacional de Educación a Distancia, C/ Juan del Rosal, 16, 28040, Madrid, España.

<sup>b</sup> ITAP-DISA, Universidad de Valladolid, 47002 Valladolid, España.

### Resumen

Los jeroglíficos egipcios forman parte de un lenguaje complejo, compuesto por cientos de símbolos que pueden transcribirse a fonemas latinos. Aunque existen estudios centrados en el reconocimiento de jeroglíficos mediante técnicas de visión por computador, este trabajo destaca por utilizar el mayor conjunto de datos de jeroglíficos, y que abarca los principales símbolos utilizados en estelas egipcias. Tras un análisis de los principales métodos de detección, segmentación y clasificación, se propone un enfoque combinado que, mediante el uso de técnicas clásicas de visión artificial para la segmentación basadas en regiones, y de votación mediante validación cruzada apoyada en una red ConvNeXt para la clasificación, ha sido evaluado en varias estelas de piedra talladas o pintadas con resultados prometedores. Este trabajo abre la puerta al desarrollo de técnicas automatizadas para la lectura de textos jeroglíficos egipcios.

*Palabras clave:* Procesamiento de imágenes, Redes Neuronales, Aprendizaje Máquina, Técnicas de Inteligencia artificial, Visión por computador.

### Automatic detection and recognition of Egyptian hieroglyphs using computer vision and cross-validation

#### Abstract

Egyptian hieroglyphs form part of a complex language, composed of hundreds of symbols that can be transcribed into their Latin phonemes. Although there are studies focused on hieroglyph recognition using computer vision techniques, this work stands out for using the largest hieroglyph dataset, covering the main symbols used on Egyptian stelae. After an analysis of the main detection, segmentation and classification methods, a combined approach is proposed. Using classical computer vision techniques for region-based segmentation and cross-validation voting supported by a ConvNeXt network for classification, this approach has been evaluated on several carved or painted stone stelae with promising results. This work paves the way for the development of automated techniques for reading Egyptian hieroglyphic texts.

*Keywords:* Image processing, Neural networks, Machine Learning, Artificial intelligence techniques, Computer vision.

### 1. Introducción

Los jeroglíficos egipcios se inspiraron, desde los primeros tiempos faraónicos, en la naturaleza y la vida cotidiana, representando por ejemplo animales, plantas y objetos domésticos. Estos textos se escribían en varias direcciones y se leían de izquierda a derecha, derecha a izquierda o de arriba abajo, lo que venía determinado por la orientación de las figuras humanas o animales. Además, cada jeroglífico podía tener varios significados según el contexto: ideogramas (representan ideas u objetos), determinativos (indican la categoría de una palabra) o fonogramas (representan uno o más

sonidos). Todo ello, unido al desgaste que han sufrido como consecuencia del paso del tiempo, la exposición y el vandalismo (especialmente cuando eran tallados o pintados sobre distintos materiales), ha dificultado históricamente su reconocimiento.

En el estudio presentado en este trabajo, se propone el uso de un modelo de segmentación basado en regiones. En concreto, se aplican técnicas clásicas de visión artificial para segmentar los jeroglíficos de manera individual. A continuación, se extrae su disposición y clasifica individualmente con una técnica de votación basada en validación cruzada (*Cross-Validation Voting*) (Duque-

\*Autor para correspondencia: pjherrera@issi.uned.es

Domingo *et al.*, 2022) y la red neuronal ConvNeXt (Liu *et al.*, 2022). Finalmente, se obtiene la transliteración fonética y recreación en JSesh (Rosmorduc, 2014, 2020). Para ello, se ha utilizado el mayor conjunto de datos de jeroglíficos egipcios disponible en la actualidad, que consta de 13729 jeroglíficos clasificados en 310 categorías (Fuentes-Ferrer *et al.*, 2025), y que abarca desde el Imperio Antiguo (c. 2686-2181 a. C.) hasta el Imperio Nuevo (c. 1500-1070 a. C.).

El resto del documento se organiza de la siguiente manera. La segunda sección explora el estado del arte relativo al reconocimiento de jeroglíficos y las técnicas utilizadas. La tercera sección presenta el enfoque propuesto y describe brevemente el conjunto de datos utilizado. La cuarta sección muestra los resultados obtenidos. Finalmente, la quinta sección presenta las principales conclusiones extraídas, así como las líneas de investigación abiertas.

## 2. Estado del arte

El reconocimiento de jeroglíficos es una tarea compleja debido a la escasez de recursos. En los últimos años, el interés por automatizar la lectura de textos antiguos ha crecido con el desarrollo de algoritmos de aprendizaje automático. Un ejemplo de ello es Fabricius, un proyecto de Google Arts & Culture (Fabricius, 2023) que requiere delimitar manualmente los jeroglíficos a decodificar, una limitación que este trabajo busca eliminar mediante la automatización del proceso.

Diversos estudios han abordado la detección de jeroglíficos mediante técnicas clásicas de visión por computador (Franken and van Gemert, 2013), (Roman-Rangel and Marchand-Maillet, 2015), (Duque-Domingo *et al.*, 2017), (Nederhof and Rahman, 2017), (Elnabawy *et al.*, 2018). Otros enfoques han utilizado modelos específicos basados en aprendizaje profundo como R-CNN (Sobhy *et al.*, 2023), Mask R-CNN (He *et al.*, 2018), Detectron2 (Wu *et al.*, 2021), (Guidi *et al.*, 2023) y YOLO (Redmon *et al.*, 2016) que, no obstante, requieren el etiquetado manual de los límites del jeroglífico, lo cual puede resultar laborioso e incluso costoso, como en el último caso. Existen también algunos enfoques basados en SAM (Kirillov *et al.*, 2023), que han sido aplicados con éxito al reconocimiento de jeroglíficos (Fuentes-Ferrer *et al.*, 2025).

El aprendizaje profundo también se ha utilizado con éxito para la clasificación de imágenes (Pajares *et al.*, 2021) y, más recientemente, en el contexto de los jeroglíficos. Por ejemplo, para mejorar la precisión en la recuperación y clasificación (Duque-Domingo *et al.*, 2020), (Moustafa *et al.*, 2022), (Barucci *et al.*, 2021, 2022). Otros enfoques han propuesto el uso de Transformers (ViT) y procesamiento de lenguaje natural (Nederhof, 2016) o modelos probabilísticos (Nederhof, 2017). Existen también metodologías similares centradas en el reconocimiento de caracteres mayas (Roman-Rangel and Marchand-Maillet, 2015), (Pinilla-Buitrago *et al.*, 2022), caracteres chinos (Zhuo and Zhang, 2024), así como en la identificación de símbolos en piedra (Duque-Domingo *et al.*, 2018) y documentos y manuscritos históricos (Alfaro-Contreras *et al.*, 2023), (Souibgui *et al.*, 2022).

## 3. Enfoque propuesto

El enfoque propuesto consta de una primera fase de detección y una segunda de reconocimiento. En la primera, se llevará a cabo la localización y segmentación de los

jeroglíficos, mientras que en la segunda fase se realizará su identificación y clasificación en la clase correspondiente.

### 3.1. Conjunto de datos

El conjunto de datos (*dataset*) utilizado consta de 13729 imágenes procedentes de diversas fuentes, y corresponden a estelas de museos, inscripciones manuscritas o esculpidas, y jeroglíficos tallados, abarcando diferentes períodos de la historia egipcia con el fin de favorecer una mayor generalización en el proceso de reconocimiento. En concreto, se etiquetaron 9703 imágenes; el resto (4026 imágenes) proceden del conjunto de datos elaborado por (Franken and van Gemert, 2013). Las imágenes originales presentaban distintos fondos y tamaños, lo que requirió de una fase previa de preprocesamiento para unificar sus dimensiones en X o Y. Finalmente, los jeroglíficos se clasificaron en 310 clases según la estructura y tipología definida por el egiptólogo Gardiner, que proporciona un identificador único para cada clase de jeroglífico (Gardiner, 1957), (Wilkinson, 1992). Así, el nombre de cada clase está formado por un carácter alfabético seguido de un número; la Figura 1 ilustra una selección de muestras de la clase n35.



Figura 1: Muestras de ejemplo de la clase n35 del conjunto de datos.

### 3.2. Detección

En esta propuesta se han utilizado técnicas de visión clásica que ya han sido empleadas con anterioridad para la detección de jeroglíficos (Duque-Domingo *et al.*, 2017), (Fuentes-Ferrer *et al.*, 2025), si bien aquí se aplican en un contexto diferente; en concreto, sobre jeroglíficos pintados, lo que añade una dificultad al proceso frente a p.e. los tallados directamente en la piedra. No obstante, se valoraron otros métodos de detección, descritos en el estado del arte: Mask R-CNN, Detectron2, YOLO. Sin embargo, el etiquetado en estos métodos se estimó muy costoso, al precisar de numerosos cuadros delimitadores alrededor de los jeroglíficos. Dada la imprecisión de los resultados preliminares obtenidos, se optó finalmente por la solución presentada que utiliza el algoritmo de Canny para extraer los bordes de las regiones (Canny, 1986). Una vez detectados los bordes, se analizan las distintas regiones. Con esta técnica, es posible aislar correctamente casos como el mostrado en la Figura 2, donde existe una diferencia considerable entre el color de los jeroglíficos y el del fondo de la imagen. A continuación, se selecciona el marco (*bounding box*) que delimita el borde independiente (rectángulos verdes en la Figura 2) y se separa cada jeroglífico en una imagen independiente. Las imágenes separadas se procesan posteriormente en la etapa de reconocimiento, tal y como se describe en la subsección siguiente.

Además de la localización y segmentación, los jeroglíficos deben organizarse para su posterior lectura. Para ello, se asume que la lectura se realiza de izquierda a derecha y de arriba

abajo. Los antiguos egipcios situaban sus jeroglíficos individualmente o agrupados en bloques, tal y como se muestra también en la Figura 2.



Figura 2: Esquema del método de detección propuesto.

### 3.3. Reconocimiento

En aras de una mayor generalización del proceso de reconocimiento, en esta propuesta se ha utilizado una técnica de votación basada en validación cruzada (VVC) (Duque-Domingo *et al.*, 2022), (Fuentes-Ferrer *et al.*, 2025), lo que permite el uso de diferentes clasificadores similares entrenados en grupos de validación disjuntos. A diferencia de los métodos tradicionales que seleccionan conjuntos de datos aleatorios para entrenar diferentes modelos, aquí se utilizan conjuntos de validación para evitar el sobreajuste. Esto es especialmente relevante en redes neuronales, donde los mecanismos de detención temprana permiten detener el entrenamiento antes de que se produzca el sobreajuste. El método VVC distribuye el conjunto de datos en diferentes ranuras de entrenamiento y validación. De esta manera, se utilizan todos los datos para entrenar el modelo, ya que cada modelo individual utiliza una parte del conjunto de datos como entrenamiento y otra como validación, evitando así la pérdida de muestras durante el proceso. La Figura 3 muestra la técnica VVC aplicada a la identificación y clasificación de jeroglíficos egipcios, que emplea como estructura principal la red neuronal ConvNeXt (Liu *et al.*, 2022).

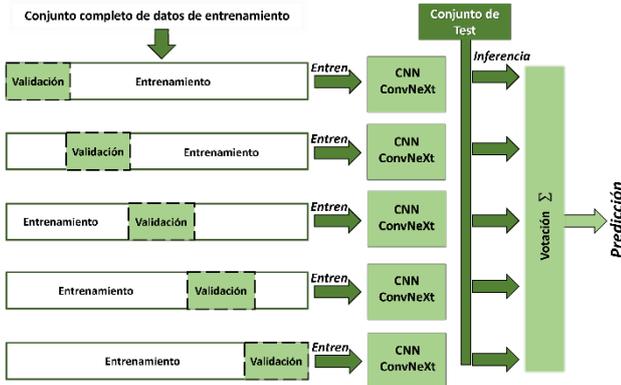


Figura 3: Diagrama de flujo del método de reconocimiento propuesto.

En el método VVC, los datos son primero aleatorizados y a continuación se seleccionan  $v$  ranuras de validación distintas. El conjunto completo de datos de entrenamiento se denota como  $d$ . Para cada ranura  $i$ , los conjuntos de entrenamiento y validación correspondientes se denotan como  $E_i$  y  $V_i$ , respectivamente. Estos conjuntos deben ajustarse a las ecuaciones (1), (2) y (3).

$$d = \bigcup_{i=1}^v V_i ; \quad \bigcap_{i=1}^v V_i = \emptyset \quad (1)$$

$$[d = E_i \cup V_i] \quad \forall i \in [1, v] \quad (2)$$

$$[E_i \cap V_i = \emptyset] \quad \forall i \in [1, v] \quad (3)$$

Una vez que los modelos han sido entrenados con cada conjunto de entrenamiento y validación, estos se combinan utilizando técnicas clásicas de votación.  $k$  denota el número de clasificadores, cada uno asociado a su ranura de validación correspondiente. Dada una muestra de entrada  $e$ ,  $\overline{p}_i(e)$  es el vector de probabilidades de salida generado por el clasificador  $i$ . Este vector está compuesto por las probabilidades  $p_{ij}(e)$ , que indican que la muestra  $e$  pertenece a la clase  $j$  en base al criterio del clasificador  $i$ . El vector de etiquetas  $\vec{c}$  consta de una etiqueta para cada clase posible:  $\{c_1, c_2, \dots, c_n\}$ .

En este punto se definen dos criterios de votación, un criterio más suave o blando (VS), conocido como *soft voting* en terminología inglesa, y otro criterio de votación directo, más estricto (VD), conocido también como *hard voting* en la literatura. En la ecuación (4),  $VS(e)$ , se obtiene acumulando las probabilidades de salida de cada clase  $j$ . Se asigna un peso  $p_i$  a cada clasificador  $i$ ; siendo en esta propuesta el valor de todos pesos  $\frac{1}{k}$ . La función *argmax* devuelve la posición de clase con la mayor probabilidad acumulada.

$$VS(e) = c_{\arg \max} \sum_{i=1}^k p_i \cdot p_{ij}(e) \quad (4)$$

En el caso de la votación directa (5), la probabilidad debe ser binarizada previamente. De modo que solo la clase con la mayor probabilidad  $m_{ij}(x)$  toma el valor 1, mientras que el resto toman el valor 0. Así,  $VD(e)$  se obtiene sumando los valores binarios correspondientes a cada clase  $j$ . Al igual que en la votación suave,  $p_i$  se considera igual a  $\frac{1}{k}$  en este trabajo.

$$VD(e) = c_{\arg \max} \sum_{i=1}^k p_i \cdot m_{ij}(e) \quad (5)$$

Independientemente del criterio de votación escogido, finalmente se obtiene la clase a la que corresponde cada uno de los jeroglíficos, en base a su código de Gardiner, así como su transliteración. A partir de la transliteración es posible reconstruir los jeroglíficos utilizando JSesh (Rosmorduc, 2014), e incluso realizar la traducción utilizando algún tipo de modelo de transformación.

## 4. Experimentación y resultados

Esta sección describe, en primer lugar, el incremento de datos acometido durante la etapa de entrenamiento. Después se realiza un análisis de la fase de detección y, por último, se presentan los resultados del enfoque combinado.

### 4.1. Incremento de datos

Durante la fase de entrenamiento se realizó un incremento aleatorio de los datos (*data augmentation* en terminología inglesa), para lograr un mejor equilibrio del conjunto de entrenamiento. Se eliminaron muestras de aquellas clases que presentaban un exceso de imágenes, y se añadieron nuevos elementos a las clases que presentaban un número de muestras inferior a un determinado umbral. Para ello, se realizó una

traslación aleatoria horizontal y vertical en un máximo de 10 píxeles, así como ligeras rotaciones aleatorias de 5 grados y un escalado, también aleatorio, con un factor que variaba entre de 0,90 y 1,10. Dado que los jeroglíficos pueden orientarse tanto a la izquierda como a la derecha (sin que su significado se vea alterado), también se utilizó la inversión horizontal.

Con todo ello, se estableció que el número máximo de imágenes de entrenamiento que puede contener una clase fuera el doble del promedio de imágenes por etiqueta. Tras el aumento de datos, el número de imágenes del conjunto de entrenamiento se incrementó de 9359 a 18718. No obstante, y con el fin de mejorar la fiabilidad de los resultados, en los experimentos realizados se utilizaron 13169 imágenes de jeroglíficos categorizados en 164 clases, tras fijar un umbral mínimo de 7 jeroglíficos por clase. El conjunto de datos de clasificación se dividió en un 70 % para entrenamiento, 20 % para validación y un 10 % para pruebas (test). La Tabla 1 muestra el conjunto de imágenes utilizadas, a partir del *dataset* original y tras el incremento de datos.

Tabla 1: Conjunto de imágenes utilizado para entrenamiento, validación y test

Cjto. Imág.	Total	Entren. (70%)	Validac. (20%)	Test (10%)
Dataset original	13169	9218	2634	1317
Aumento de datos	20035	14558	4160	1317

#### 4.2. Análisis de la fase de detección

Se realizó un análisis detallado de diversas estelas, abarcando varios materiales (piedra, papiro) y medios (tallado, pintado), contemplados todos ellos en el conjunto de datos utilizado. La Figura 4 muestra una de las estelas evaluadas, conocida como Estela de Pakeshi (*Stela of the god's-father of Amun Pakeshi*), y que se ha tomado como referencia en este trabajo para ilustrar el enfoque propuesto y los resultados obtenidos.



Figura 4: Estela de Pakeshi (Cortesía del MET).

En primer lugar, se realizó la extracción de las cinco filas inferiores de la estela, tal y como se ilustra en las Figuras 5 y 6. Este proceso de extracción generó un conjunto de cinco imágenes independientes, cada una de las cuales se sometió a un procedimiento de inversión. Esta inversión se realizó con el propósito de mejorar el reconocimiento posicional de los jeroglíficos individuales de cara a su posterior extracción, un

aspecto esencial en el proceso de análisis y clasificación de elementos jeroglíficos en estelas egipcias.



Figura 5: Filas extraídas de la Estela de Pakeshi.

El proceso de extracción se diseñó con el objetivo de obtener los jeroglíficos individualizados presentes en las imágenes analizadas. Para ello, se empleó el método basado en segmentación de regiones descrito en la sección 3.2. La Figura 6 ilustra la extracción individual de cada jeroglífico y su posterior identificación y clasificación según sus características morfológicas y estructurales.

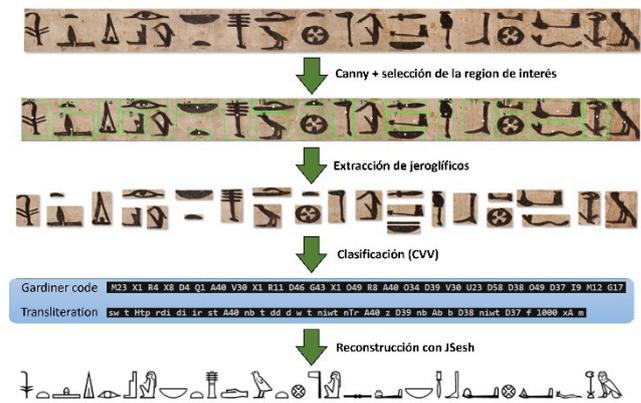


Figura 6: Diagrama de flujo del enfoque combinado propuesto.

Es importante destacar que la precisión del proceso de extracción puede diferir de una imagen a otra, dependiendo de la complejidad del jeroglífico. Por ese motivo, y para evaluar la calidad de la extracción, se llevó a cabo un proceso de clasificación mediante VVC, tanto en su modalidad de votación suave (*soft voting*) como dura (*hard voting*). Estas técnicas permitieron determinar la idoneidad de la segmentación y la precisión de la clasificación. Finalmente, y como parte integral del proceso, se presenta en la Figura 7 la representación JSesh de los jeroglíficos extraídos de la Estela de Pakeshi (y mostrados en la Figura 5), lo que completa la fase de análisis y validación de la extracción. Estos procedimientos, respaldados por una evaluación y clasificación rigurosas, generaron resultados prometedores, que se describen en la subsección siguiente.



Figura 7: Representación jeroglífica correspondiente con JSesh.

### 4.3. Resultados de la propuesta combinada

Es importante señalar que el conjunto de datos actual no aglutina la totalidad de jeroglíficos existente. Por tanto, sería posible plantear la clasificación de todos los jeroglíficos obtenidos durante el proceso de extracción, independientemente de si están incluidos o no en el conjunto de datos actual, o bien, centrarse únicamente en la clasificación de los jeroglíficos incluidos en el conjunto de datos. La elección entre estos dos planteamientos dependerá de los objetivos específicos de la investigación y de las necesidades prácticas del sistema, garantizando así que los resultados obtenidos sean coherentes con la disponibilidad de datos.

En este estudio se presentan los resultados obtenidos, tanto en la clasificación de todos los jeroglíficos identificados durante el proceso, como en la de aquéllos que solamente estén presentes en el conjunto de datos. Los resultados de este análisis se presentan en las Tablas 2 y 3, en términos de exactitud (*accuracy*), exactitud balanceada (*balanced accuracy*), precisión (*precision* o PPV), recuperación (*recall* o TPR) y puntuación F1 (*F1 score*). Se distingue entre los resultados obtenidos tras del análisis de los jeroglíficos presentes en la Estela de Pakeshi (EP), así como el resultado promediado de las cuatro estelas analizadas (RP), correspondientes a diferentes períodos históricos que abarcan 3000 años de desarrollo de la escritura jeroglífica conocida.

Tabla 2: Resultados obtenidos considerando todos los jeroglíficos presentes en las estelas

VVC-VS	Exactitud	Exactitud balance	Precisión	Recuperación	Puntuación F1
EP	0,570	0,710	0,550	0,580	0,560
RP	0,673	0,657	0,507	0,547	0,523

VVC-VD	Exactitud	Exactitud balance	Precisión	Recuperación	Puntuación F1
EP	0,550	0,700	0,510	0,560	0,530
RP	0,677	0,660	0,520	0,530	0,533

Tabla 3: Resultados obtenidos considerando únicamente los jeroglíficos que forman parte del dataset

VVC-VS	Exactitud	Exactitud balance	Precisión	Recuperación	Puntuación F1
EP	0,600	0,710	0,580	0,600	0,590
RP	0,740	0,683	0,583	0,607	0,593

VVC-VD	Exactitud	Exactitud balance	Precisión	Recuperación	Puntuación F1
EP	0,590	0,710	0,550	0,590	0,570
RP	0,743	0,683	0,590	0,610	0,600

Los mejores resultados se obtuvieron considerando únicamente los jeroglíficos presentes en el conjunto de datos. En concreto, el enfoque que utilizó VVC-VS es el que obtuvo un mejor rendimiento con respecto a la Estela de Pakeshi, con una exactitud (*precision*) de 0,600, llegando hasta 0,740 si se toma el valor medio de las cuatro estelas analizadas. El enfoque basado en VVC-VD obtuvo, en el promedio de las cuatro estelas analizadas, mejores resultados, incrementando la exactitud (*accuracy*) hasta 0,743.

La Estela de Pakeshi presenta los jeroglíficos pintados directamente sobre la piedra. Una particularidad que, a la vista

de los resultados obtenidos, afectó negativamente al proceso de detección y reconocimiento propuesto. Los casos en los que la inscripción fue tallada directamente sobre la piedra, presentaban un mayor contraste, facilitando su reconocimiento.

Se puede concluir que la propuesta combinada muestra un rendimiento prometedor en la clasificación de jeroglíficos, especialmente si se consideran únicamente los jeroglíficos presentes en el conjunto de datos. Los resultados obtenidos respaldan su utilidad en la lectura automatizada de jeroglíficos egipcios.

## 5. Conclusiones

Este estudio ahonda en el reconocimiento de jeroglíficos egipcios, partiendo de un conjunto de datos extenso y completo, y proponiendo un enfoque combinado de detección y reconocimiento de jeroglíficos con resultados prometedores. El análisis realizado evalúa la totalidad de los jeroglíficos encontrados en cada una de las estelas analizadas. Los resultados de este análisis sientan las bases para futuras investigaciones en el campo de la egiptología y la interpretación automatizada de textos jeroglíficos.

Las líneas de investigación a seguir en el futuro se centran en la ampliación y diversificación del conjunto de datos, la extensión de las clases de entrenamiento, la inclusión de diversas tipologías de escritura en estelas, la profundización en la detección jeroglífica en diferentes medios y la traducción final a lenguas modernas, lo que enriquecería el enfoque actual y fortalecería su aplicabilidad en una amplia gama de escenarios relacionados con la lectura automatizada de textos jeroglíficos egipcios.

## Agradecimientos

Este trabajo ha sido realizado gracias al apoyo de los proyectos DARWEEM (PID2020-113229RB-C43), RECOVERY (PID2020-112658RB-I00) y ROSOGAR (PID2021-123020OB-I00), financiados por la Agencia Estatal de Investigación (MCIN/AEI/10.13039/501100011033/FEDER, UE).

## Referencias

- Alfaro-Contreras, M., Ríos-Vila, A., Valero-Mas, J. J., Calvo-Zaragoza, J., 2023. Few-shot symbol classification via self-supervised learning and nearest neighbor. *Pattern Recognition Letters* 167, 1-8. DOI: 10.1016/j.patrec.2023.01.014
- Barucci, A., Canfailla, C., Cucci, C., Forasassi, M., Franci, M., Guarducci, G., Guidi, T., Loschiavo, M., Picollo, M., Pini, R., Python, L., Valentini, S., Argenti, F., 2022. Ancient Egyptian hieroglyphs segmentation and classification with convolutional neural networks. In: *International Conference Florence Heri-Tech*, Springer, pp. 126-139. DOI: 10.1007/978-3-031-20302-2\_10
- Barucci, A., Cucci, C., Franci, M., Loschiavo, M., Argenti, F., 2021. A deep learning approach to ancient Egyptian hieroglyphs classification. *IEEE Access* 9, 123438-123447. DOI: 10.1109/ACCESS.2021.3110082
- Canny, J., 1986. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-8* (6), 679-698. DOI: 10.1109/TPAMI.1986.4767851
- Duque-Domingo, J., Aparicio, R. M., Rodrigo, L. M. G., 2022. Cross validation voting for improving CNN classification in grocery products. *IEEE Access* 10, 20913-20925. DOI: 10.1109/ACCESS.2022.3152224
- Duque-Domingo, J., Gómez-García-Bermejo, J., Zalama, E., 2020. Egyptian shabti identification by means of deep neural networks and semantic

- integration with Europeana. *Applied Sciences* 10(18), 6408. DOI: 10.3390/app10186408
- Duque-Domingo, J., Herrera, P. J., Valero, E., Cerrada, C., 2017. Deciphering Egyptian hieroglyphs: towards a new strategy for navigation in museums. *Sensors* 17(3), 589. DOI: 10.3390/s17030589
- Duque-Domingo, J., Herrera, P. J., Cerrada, C., Cerrada, J. A., 2018. A vision-based strategy to segment and localize ancient symbols written in stone. In: *Third Iberian Robotics Conference*, Springer, pp 251-260. DOI: 10.1007/978-3-319-70836-2\_21
- Elnabawy, R., Elias, R., Salem, M., 2018. Image based hieroglyphic character recognition, In: *14th International Conference on Signal-Image Technology & Internet-Based Systems*, IEEE, pp. 32-39. DOI: 10.1109/SITIS.2018.00016
- Fabricius, Google Arts & Culture, 2020. <https://g.co/arts/arMD1wvi7bdNnXcf8> (Accedido 2 abril 2025).
- Franken, M., van Gemert, J.C., 2013. Automatic Egyptian hieroglyph recognition by retrieving images as texts. In: *Proceedings of the Twenty First ACM International Conference on Multimedia*, ACM, pp. 765–768. DOI: 10.1145/2502081.2502199
- Fuentes-Ferrer, R., Duque-Domingo, J., Herrera, P. J., 2025. Recognition of Egyptian hieroglyphic texts through focused generic segmentation and cross-validation voting. *Applied Soft Computing* 171, 112793. DOI: 10.1016/j.asoc.2025.112793
- Gardiner, A., 1957. *Egyptian Grammar*, third ed., Griffith Institute, Oxford.
- Guidi, T., Python, L., Forasassi, M., Cucci, C., Franci, M., Argenti, F., Barucci, A., 2023. Egyptian hieroglyphs segmentation with convolutional neural networks. *Algorithms* 16(2), 79. DOI: 10.3390/a16020079
- He, K., Gkioxari, G., Dollár, P., Girshick, R., 2018. Mask R-CNN. *arXiv*. DOI: 10.48550/arXiv.1703.06870
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.-C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment anything. *arXiv*. DOI: 10.48550/arXiv.2304.02643
- Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., Xie, S., 2022. A ConvNet for the 2020s. *arXiv*. DOI: 10.48550/arXiv.2201.03545
- Moustafa, R., Hesham, F., Hussein, S., Amr, B., Refaat, S., Shorim, N., Ghanim, T. M., 2022. Hieroglyphs language translator using deep learning techniques (Scriba). In: *Second International Mobile, Intelligent, and Ubiquitous Computing Conference*, IEEE, pp. 125-132. DOI: 10.1109/MIUCC55081.2022.9781784
- Nederhof, M.-J., 2016. OCR of handwritten transcriptions of Ancient Egyptian hieroglyphic text. In: *Altertumswissenschaften in a Digital Age: Egyptology, Papyrology and Beyond*, Leipzig.
- Nederhof, M. -J., Rahman, F., 2017. A probabilistic model of Ancient Egyptian writing. *Journal of Language Modelling* 5(1), 131-163. DOI: 10.15398/jlm.v5i1.150
- Roman-Rangel, E., Marchand-Maillet, S., 2015. Shape-based detection of Maya hieroglyphs using weighted bag representations. *Pattern Recognition* 48(4), 1161–1173. DOI: 10.1016/j.patcog.2014.06.009
- Rosmorduc, S., 2014. JSesh Documentation. <http://jseshdoc.qenherkhopeshef.org> (Accedido 2 abril 2025).
- Rosmorduc, S., 2020. Automated transliteration of late Egyptian using neural networks: an experiment in 'deep learning'. *Lingua Aegyptia - Journal of Egyptian Language Studies* 28, 233-257. DOI: 10.37011/lingaeg.28.07
- Pajares, G., Herrera, P. J., Besada, E., 2021. *Aprendizaje profundo*. RC Libros Editorial, Madrid.
- Pinilla-Buitrago, L. A., Martínez-Trinidad, J. F., Carrasco-Ochoa, J. A., 2022. Encoding hieroglyph segments to represent hieroglyphs following the bag of visual word model for retrieval. *Expert Systems with Applications* 201, 116983. DOI: 10.1016/j.eswa.2022.116983
- Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: unified, real-time object detection. *arXiv*. DOI: 10.48550/arXiv.1506.02640
- Sobhy, A., Helmy, M., Khalil, M., Elmasry, S., Boules, Y., Negied, N., 2023. An AI based automatic translator for ancient hieroglyphic language - From scanned images to English text. *IEEE Access* 11, 38796-38804. DOI: 10.1109/ACCESS.2023.3267981
- Souibgui, M.A., Fornés, A., Kessentini, Y., Megyesi, B., 2022. Few shots are all you need: A progressive learning approach for low resource handwritten text recognition. *Pattern Recognition Letters* 160, 43-49. DOI: 10.1016/j.patrec.2022.06.003
- Wilkinson, R.H., 1992. *Reading Egyptian Art: A hieroglyphic guide to ancient Egyptian painting and sculpture*, Thames & Hudson, London.
- Wu, Y., Kirillov, A., Massa, F., Lo, W.-Y., Girshick, R., 2021. Detectron2 v0.6. GitHub. <https://github.com/facebookresearch/detectron2>
- Zhuo, S., Zhang, J., 2024. Attention-based deformable convolutional network for Chinese various dynasties character recognition. *Expert Systems with Applications* 238, 121881. DOI: 10.1016/j.eswa.2023.121881