CEA Comité Español de Automática
XLVII JORNADAS DE AUTOMÁTICA CÓRDOBA 2026
UNIVERSIDADE DA CORUÑA

Jornadas de Automática

MAAT: Un entorno interactivo para la extracción, clasificación y traducción de signos jeroglíficos

Raúl Fuentes-Ferrer[a], Jaime Duque-Domingo[b], Pedro Javier Herrera[a,*]

[a] *Dpto. de Ingeniería de Software y Sistemas Informáticos, Universidad Nacional de Educación a Distancia, C/ Juan del Rosal, 16, 28040, Madrid, España.*
[b] *ITAP-DISA, Universidad de Valladolid, 47002 Valladolid, España*

To cite this article: Fuentes-Ferrer, R., Duque-Domingo, J, Herrera, P.J. 2026. MAAT: An interactive environment for the extraction, classification, and translation of hieroglyphic signs. Jornadas de Automática, 47. https://doi.org/10.17979/ja-cea.2026.47.13694

Resumen

El procesamiento automático de inscripciones jeroglíficas del Antiguo Egipto sigue siendo un desafío debido a que el material epigráfico real combina diseños irregulares, dirección de escritura variable, superficies dañadas, soportes tallados o pintados y una calidad de la imagen heterogénea. Los enfoques existentes a menudo abordan la segmentación, la clasificación o la traducción como tareas separadas, mientras que los flujos de trabajo epigráficos prácticos requieren la inspección y evaluación de la cadena completa desde el procesamiento de la imagen hasta la traducción. Este trabajo presenta un entorno interactivo para la extracción, clasificación, traducción y evaluación de signos jeroglíficos. El entorno desarrollado demuestra cómo la supervisión interactiva puede hacer auditable todo el proceso de reconocimiento y traducción a nivel de signo, región de interés e inscripción, permitiendo la comparación controlada de configuraciones de extracción y la construcción incremental de conjuntos de datos jeroglíficos validados manualmente.

Palabras clave: Procesamiento de imágenes, Redes neuronales, Aprendizaje máquina, Técnicas de inteligencia artificial, Visión por computador.

MAAT: An interactive environment for the extraction, classification, and translation of hieroglyphic signs

Abstract

Automatic processing of Ancient Egyptian hieroglyphic inscriptions remains challenging because real epigraphic material combines irregular layouts, variable writing direction, damaged surfaces, carved or painted supports, and heterogeneous image quality. Existing approaches often address segmentation, classification or translation as separate tasks, whereas practical epigraphic workflows require inspection and evaluation of the full chain from image processing to translation. This paper presents an interactive research workbench for hieroglyphic sign extraction, classification and translation-oriented evaluation. The implemented workflow demonstrates how interactive supervision can make the complete recognition and translation pipeline auditable at sign, ROI and inscription levels, enabling controlled comparison of extraction configurations and incremental construction of manually validated hieroglyphic datasets.

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

1. Introducción

El análisis automático de inscripciones del Antiguo Egipto no se circunscribe únicamente a la clasificación de signos aislados. Una estela real puede combinar signos pintados, incisos o en relieve, filas o columnas irregulares, material parcial o completamente dañado, sombras y direcciones de lectura variables. Todo ello ocasiona que la traducción automática de textos jeroglíficos a partir de una imagen sea particularmente exigente: el sistema debe localizar los signos, preservar su orden de lectura, clasificarlos en un inventario de

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

signos estandarizado y evaluar el resultado lingüístico comparándolo con traducciones de referencia.

Trabajos recientes han demostrado el potencial del aprendizaje profundo para la segmentación y clasificación de signos, al tiempo que resaltan la importancia de la variabilidad de la imagen, el tipo de soporte, el estado de conservación y la escasez de datos (Fuentes-Ferrer *et al*., 2025a). No obstante, el entorno presentado en este trabajo, MAAT, aborda estos problemas desde una perspectiva complementaria: no se presenta simplemente como un clasificador, sino como una plataforma completa que posibilita que todo el proceso sea revisable, repetible y comparable entre diferentes configuraciones de extracción. Precisamente, el nombre MAAT alude al principio egipcio de verdad, orden y equilibrio, y refleja el objetivo del sistema de hacer trazable, verificable y coherente el proceso completo de análisis de inscripciones jeroglíficas, desde el procesamiento de la imagen hasta la evaluación de la traducción.

Con la finalidad de evaluar el reconocimiento y la traducción de jeroglíficos en inscripciones reales, MAAT combina regiones de interés (ROI) definidas por el usuario, múltiples estrategias de segmentación, la predicción del código de Gardiner, la generación de la traducción y la evaluación basada en métricas en una sola aplicación (Figura 1). Las principales contribuciones de este trabajo son: 1. Un entorno de trabajo integral para el procesamiento de imágenes de estelas reales, desde la entrada visual hasta la evaluación a nivel de traducción; 2. Un flujo de trabajo basado en ROI y revisión con intervención humana que permite el preprocesamiento selectivo en lugar de forzar ejecuciones completas; 3. Un diseño de la extracción de características centrado en estrategias de segmentación MBRS (*Method Based on Region Segmentation*, método basado en segmentación de regiones) e IGSM (*Integrated Generic Segmentation Model*, modelo integrado de segmentación genérica) dentro de un entorno de trabajo interactivo; 4. Un protocolo de evaluación orientado a la traducción que combina valores BERTScore, BLEU y chrF por ROI y de manera concatenada; y 5. Un mecanismo de curación de datos para exportar signos validados posibilitando futuras mejoras del clasificador.

El resto del documento se organiza de la siguiente manera: la segunda sección presenta el estado del arte; la tercera sección describe el sistema propuesto; la cuarta sección muestra el protocolo experimental establecido; la quinta sección se centra en la discusión de los resultados y las limitaciones del sistema; y, finalmente, la sexta sección describe las principales conclusiones y líneas de trabajo futuro.

Figura 1: Flujo de trabajo integral del entorno MAAT, desde el procesamiento de la imagen de la estela hasta la evaluación de la traducción.

2. Estado del arte

Los enfoques basados en aprendizaje profundo han demostrado la viabilidad de clasificar signos jeroglíficos aislados y segmentar instancias de glifos en imágenes, especialmente mediante arquitecturas convolucionales y marcos de segmentación de instancias (Barucci *et al*., 2021, 2022; Guidi *et al*., 2023). Sin embargo, existen todavía desafíos recurrentes: datos de entrenamiento limitados, daños por conservación, bajo contraste, escritura en distintos soportes como piedra o papiro, y la dificultad de generalizar desde conjuntos de datos curados a fotografías de museos o soportes arqueológicos. Estos retos motivan la creación de un entorno interactivo en el que el usuario pueda inspeccionar los resultados intermedios en lugar de tratar el proceso como una caja negra.

Fabricius, un proyecto de Google Arts & Culture, popularizó el aprendizaje interactivo de jeroglíficos y la búsqueda asistida de signos, pero no fue concebido como un sistema totalmente automático de imagen a traducción (Fabricius, 2020). El antecedente metodológico más cercano a MAAT es el trabajo descrito en (Fuentes-Ferrer *et al*., 2025a,b), que introdujo un marco genérico de segmentación y clasificación específico para textos jeroglíficos egipcios, junto con un amplio conjunto de datos de 310 clases y casi 14 000 imágenes. El método basado en la segmentación de regiones (MBRS) y el modelo integrado de segmentación genérica (IGSM), no se presentan aquí como rutinas ad hoc aisladas, sino como estrategias de segmentación heredadas de ese marco experimental previo y reorganizadas dentro de un entorno de trabajo interactivo y auditable para la extracción, traducción y evaluación a nivel de ROI.

La lista de signos de Gardiner sigue siendo una referencia para agrupar e identificar signos jeroglíficos, y proporciona la notación utilizada en este trabajo para representar mediante códigos alfanuméricos los signos reconocidos (Gardiner, 1957). Así, la notación propuesta por Gardiner permite la representación, indexación y convergencia posteriores con recursos egiptológicos. Por otra parte, JSesh, un procesador de textos diseñado específicamente para escribir, editar y componer textos en jeroglíficos egipcios es ampliamente utilizado por egiptólogos para preparar e intercambiar texto jeroglífico, lo que hace que la comparación visual entre los códigos Gardiner predichos y las formas de los signos conocidos sea útil durante la revisión (Rosmorduc, 2014).

Kirillov *et al*. (2023) introdujeron SAM (*Segment Anything Model*), un modelo base con instrucciones para la segmentación de imágenes que puede utilizarse como generador de candidatos de propósito general en flujos de trabajo visuales. En MAAT, SAM se trata como una opción dentro de IGSM, en lugar de como una única estrategia de

segmentación, lo que permite la comparación con la extracción basada en contornos (MBRS).

La evaluación de la traducción se informa utilizando tres métricas complementarias: BLEU, una métrica de *n*-gramas ampliamente utilizada para la traducción automática (Papineni et al., 2002); chrF, una puntuación F de *n*-gramas de caracteres útil para la variación léxica (Popović, 2015); y BERTScore, que compara textos candidatos y de referencia utilizando incrustaciones contextuales (Zhang *et al*., 2020).

3. Sistema propuesto

3.1. Descripción general del sistema MAAT

MAAT se diseñó en torno a cinco requisitos: 1. El sistema debe admitir imágenes de inscripciones reales, no solo signos presegmentados; 2. Debe permitir la intervención de expertos o investigadores a nivel de ROI, ya que la disposición epigráfica suele ser irregular; 3. El método de extracción debe ser configurable para que el mismo conjunto de ROI pueda probarse con diferentes supuestos de segmentación; 4. El resultado debe evaluarse no solo a nivel de signo, sino también a nivel de traducción; 5. El sistema debe conservar el estado del experimento para que un investigador pueda reanudar, comparar y documentar los resultados sin tener que recalcular cada etapa.

Teniendo en cuenta lo anterior, la implementación sigue una arquitectura cliente-servidor (Figura 2). El *frontend* es una aplicación React/Vite que gestiona la carga de imágenes, el trazado de la ROI, el estado visible de la ROI, los paneles de traducción, los paneles de métricas, las vistas previas de glifos y las selecciones del usuario. El *backend* es un servicio FastAPI que expone puntos de acceso para comprobaciones de estado, persistencia de ROI, ejecución del *pipeline*, inicio de traducciones, carga de evaluaciones, cálculo de puntuaciones, almacenamiento del CSV, recuperación de sesiones e interrupción de procesos. Las etapas de larga duración se delegan a *scripts* Python externos y se gestionan como subprocesos.

3.2. Entrada y definición de la ROI

La unidad de entrada es una imagen de estela o inscripción. El investigador define una o más ROI que corresponden a filas, columnas o fragmentos de inscripción seleccionados. Cada ROI se representa como un cuadrilátero y puede voltearse horizontalmente cuando la dirección de lectura o la orientación de la imagen lo requieran. Esta definición manual es intencional: mantiene visibles las decisiones de diseño y evita forzar un detector de diseño completamente automático en material epigráfico heterogéneo.

3.3. Configuraciones de extracción

MAAT implementa seis configuraciones de extracción, obtenidas a partir de la combinación de estrategias MBRS e IGSM, el uso opcional de preprocesamiento y las variantes de extracción basadas en contornos o en SAM. Estas configuraciones permiten comparar, sobre el mismo conjunto de ROI, el efecto de cada estrategia en la detección de signos y en la traducción final. MBRS e IGSM se basan en un marco de segmentación genérica focalizada y de votación basada en validación cruzada (*Cross Validation Voting*, CVV) propuesto en Fuentes-Ferrer *et al*. (2025a). En el sistema actual, MBRS se utiliza como una vía clásica de procesamiento de imágenes, mientras que IGSM admite la extracción basada en contornos y en SAM, con preprocesamiento opcional.

3.4. Predicción del código de Gardiner y revisión visual

Para cada signo extraído de la imagen, MAAT almacena los códigos de Gardiner predichos mediante votación suave (*soft voting*) y votación dura (*hard voting*), cuando estén disponibles. La interfaz puede mostrar el signo recortado, su código predicho y una vista previa del glifo mediante JSesh. Esta visualización facilita la revisión visual y ayuda a determinar si un error se originó en la segmentación, la clasificación o la traducción. Un módulo de gestión del conjunto de datos permite al usuario seleccionar las predicciones correctas y exportar las imágenes a carpetas específicas asociadas a cada código de Gardiner. Las predicciones incorrectas, los falsos positivos o los recortes visualmente inadecuados no se incorporan automáticamente al conjunto de datos. En estos casos, el usuario puede descartarlos durante la revisión y, si el error está relacionado con la región analizada, ajustar la ROI o ejecutar una configuración alternativa de extracción. De este modo, solo los recortes validados por el usuario se exportan como ejemplos correctos asociados a un código de Gardiner, evitando introducir ruido en futuras fases de entrenamiento o refinamiento del clasificador.

3.5. Etapa de traducción

La etapa de traducción toma como entrada las secuencias de códigos de Gardiner generadas para cada ROI y las transforma en candidatos textuales mediante un módulo externo de traducción. En la implementación actual, dicho módulo se ejecuta mediante *scripts* Python desacoplados del *frontend*, lo que permite sustituir o actualizar el modelo de traducción sin modificar el flujo visual de MAAT. Este diseño es coherente con enfoques recientes de traducción neuronal basados en modelos secuencia-a-secuencia y arquitecturas *Transformer*, incluyendo modelos texto-a-texto y trabajos recientes orientados específicamente a la traducción de textos jeroglíficos (Raffel *et al*., 2020; Chung *et al*., 2024; De Cao *et al*., 2024). Los *scripts* operan sobre las variantes obtenidas por votación suave y votación dura, manteniendo separadas las salidas de cada configuración para permitir su comparación posterior.

3.6. Métricas de evaluación

MAAT evalúa la traducción comparándola con traducciones de referencia procedentes de estelas reales documentadas por museos. El panel registra los valores de las métricas BERTScore-F1, BLEU y chrF para cada ROI y para la inscripción concatenada. Además, la interfaz admite recuentos de signos originales, signos detectados, signos correctos confirmados manualmente, cobertura, detecciones excesivas y precisión. Este doble nivel de evaluación es importante porque una configuración puede producir una buena similitud a nivel lingüístico, habiendo fallado en la detección de los signos.

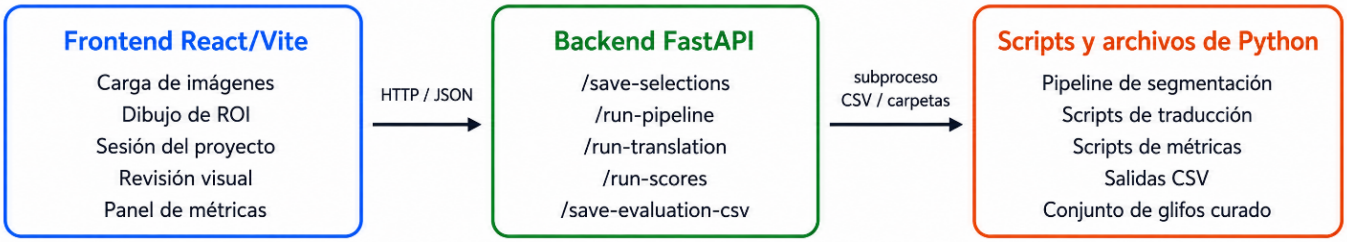

Figura 2: Arquitectura de software MAAT que vincula la interacción del frontend, la gestión del backend y los scripts de procesamiento de Python.

4. Protocolo experimental

El diseño experimental realizado evalúa MAAT en inscripciones heterogéneas: estelas pintadas, grabadas y en relieve. Cada inscripción evaluada se describe mediante su colección de origen, número de acceso, procedencia de la imagen, estado de conservación, número de regiones de interés, recuento esperado de signos y fuente de la traducción de referencia.

Para cada caso de estudio, se evalúa el mismo conjunto de ROI en todas las configuraciones seleccionadas. El protocolo es el siguiente: 1. Cargar la imagen de la inscripción y seleccionar todas las ROI correspondientes a las unidades de lectura previstas; 2. Guardar las ROI en el sistema y verificar que los recortes coincidan con las filas o columnas previstas; 3. Ejecutar las configuraciones de extracción seleccionadas sobre el mismo conjunto de ROI; 4. Revisar las imágenes de los signos extraídos y los códigos de Gardiner predichos; 5. Iniciar la traducción para todas las variantes disponibles de votación suave y votación dura; 6. Introducir o cargar la traducción de referencia para cada ROI; 7. Calcular las métricas por ROI y las métricas concatenadas; 8. Registrar el número de signos originales, los signos detectados, los signos correctos, la cobertura, el exceso y la precisión; y 9. Elegir la mejor configuración según la evidencia visual y a nivel de traducción.

El informe final presenta varias métricas. BERTScore-F1 y chrF se emplean como métricas complementarias para la comparación a nivel de traducción, ya que permiten capturar similitudes semánticas o de caracteres que BLEU puede penalizar en textos breves o con variantes léxicas aceptables. No obstante, los resultados deben interpretarse junto con las métricas de detección y clasificación, dado que una traducción aparentemente próxima a la referencia puede ocultar omisiones, falsos positivos o errores en la secuencia de signos.

Las Figuras 3, 4 y 5 muestran un ejemplo de ejecución sobre una inscripción real, desde la carga de la imagen hasta la extracción de signos y la predicción de códigos de Gardiner. En este caso, la ROI seleccionada permite observar la relación entre el recorte definido por el usuario, los signos segmentados y las clases asignadas por el clasificador. La Figura 6 muestra la salida de los módulos de traducción y evaluación para la misma inscripción. La interfaz permite comparar las traducciones generadas por las distintas configuraciones y consultar las métricas asociadas a cada ROI y a la inscripción concatenada.

Las pruebas se ejecutaron en un equipo con procesador Intel Core i7, 24 GB de memoria RAM, GPU NVIDIA GeForce RTX 2060 y sistema operativo Windows 10. El *backend* se desplegó con Python/FastAPI y el *frontend* mediante React/Vite. Con el fin de proporcionar una referencia práctica sobre la interacción con el sistema, se registraron tiempos orientativos de las principales etapas del flujo de trabajo (Tabla 1): carga y definición de ROI, extracción de signos, clasificación, traducción y cálculo de métricas. Estos tiempos no pretenden constituir una evaluación exhaustiva de rendimiento, ya que pueden variar en función del tamaño de la imagen, el número de ROI, la cantidad de signos extraídos y las configuraciones ejecutadas. No obstante, permiten estimar la espera aproximada entre tareas y muestran que el sistema puede utilizarse de forma interactiva durante la revisión de una inscripción.

Tabla 1: Tiempos orientativos de las principales etapas del flujo de trabajo

Etapa	Tiempo
Carga de imagen y definición manual de ROI	~30 s
Extracción de signos + Clasificación y predicción Gardiner por ROI	~5 min
Traducción por ROI	~1 min 30 s

5. Discusión y limitaciones

La principal contribución de MAAT es su capacidad para conectar el análisis epigráfico visual con la evaluación a nivel de traducción. El sistema no oculta las etapas intermedias: la carga de la imagen y la definición de las ROI permanecen visibles para el usuario (Figura 3), los signos extraídos pueden inspeccionarse visualmente tras la segmentación (Figura 4), y los códigos de Gardiner predichos pueden revisarse junto con los recortes correspondientes (Figura 5). Asimismo, la traducción y las métricas de evaluación permanecen accesibles en la interfaz (Figura 6). Esto es importante en la epigrafía computacional, ya que una traducción aparentemente deficiente puede deberse a un pequeño error en la selección de la ROI, un fallo en la segmentación, un orden de lectura incorrecto o una confusión del clasificador entre signos visualmente similares.

Las seis configuraciones implementadas permiten comparar de forma controlada el comportamiento de las distintas estrategias de extracción en un mismo conjunto de ROI. Las inscripciones pintadas pueden beneficiarse de supuestos diferentes a los de las inscripciones en relieve, ya que los contornos, las sombras, los bordes dañados y la textura de la superficie se comportan de manera diferente en distintos soportes. Así, al procesar el mismo conjunto de ROI con MBRS e IGSM, el investigador puede identificar qué parte del proceso es responsable de un cambio en el rendimiento.

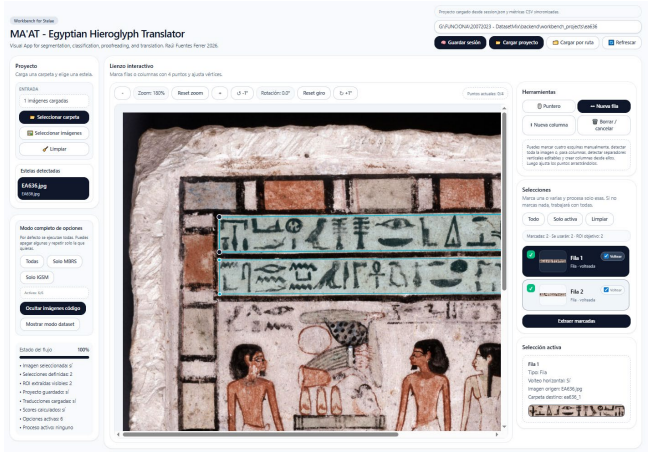

Figura 3: Ejemplo de resultado proporcionado por el entorno MAAT tras la aplicación de las etapas de carga de la imagen y definición de la ROI. Stela of Renefseneb, British Museum (EA636). © The Trustees of the British Museum. Shared under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International (CC BY-NC-SA 4.0) license.

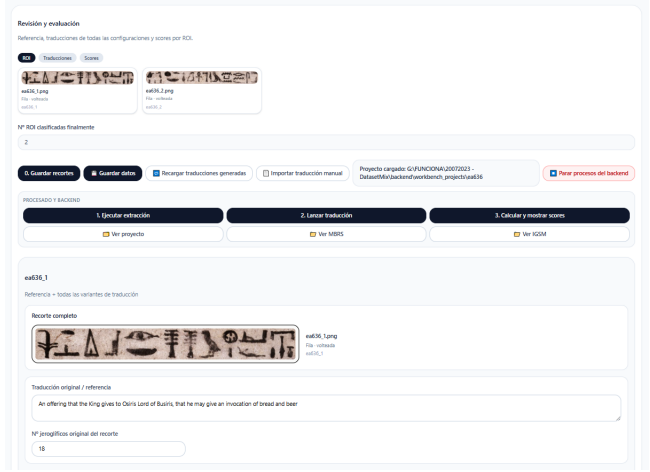

Figura 4: Ejemplo de resultado proporcionado por el entorno MAAT tras la etapa de extracción de signos.

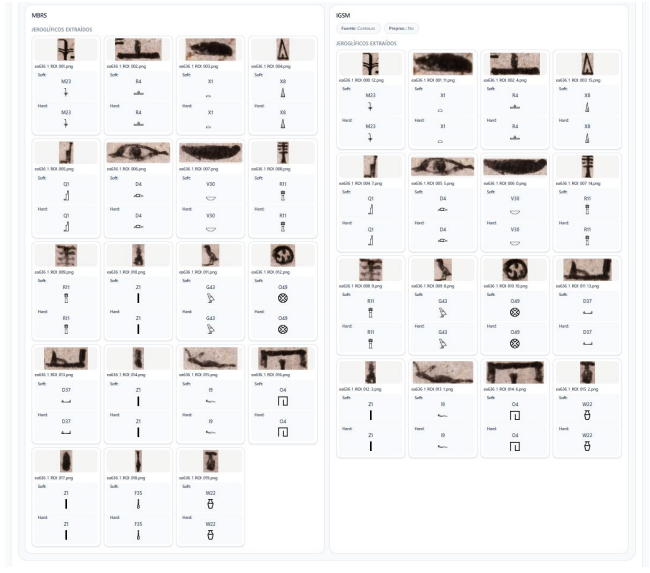

Figura 5: Ejemplo de resultado proporcionado por el entorno MAAT tras la aplicación de las etapas de predicción y revisión de los códigos de Gardiner.

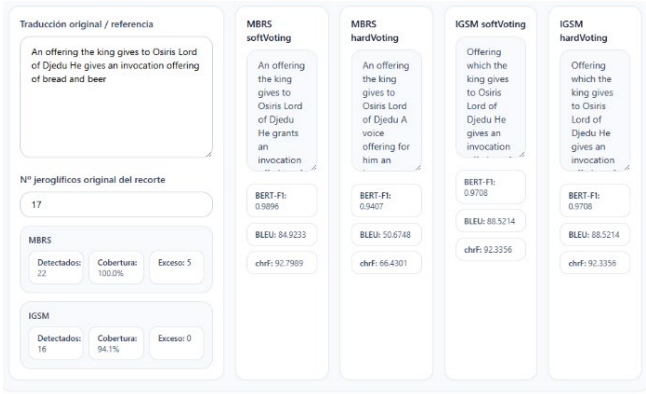

Figura 6: Ejemplo de resultado proporcionado por el entorno MAAT tras las etapas de traducción y evaluación.

BLEU sigue siendo útil como referencia léxica estricta, pero la traducción de textos antiguos a menudo permite varias representaciones válidas. Por lo tanto, BERTScore y chrF proporcionan evidencia complementaria: BERTScore captura la proximidad semántica mediante incrustaciones contextuales, mientras que chrF es sensible a la similitud a nivel de caracteres y puede resultar informativo con resultados más cortos o morfológicamente variables. Por todo ello, en el informe final MAAT debe presentar las tres métricas en lugar de basarse en una sola puntuación.

El mecanismo manual de exportación de glifos transforma la corrección de pruebas en un proceso de creación de conjuntos de datos. Cuando un usuario confirma que un código de Gardiner predicho es correcto, MAAT puede guardar el recorte correspondiente en un directorio específico para ese código. Esto crea un ciclo de retroalimentación práctico: las inscripciones reales proporcionan nuevos ejemplos validados, que posteriormente pueden utilizarse para mejorar el clasificador y reducir la discrepancia entre los datos de entrenamiento y las imágenes procedentes de museos o de entornos de exterior (Asensi-González *et al*., 2026).

A pesar de todo lo anterior, MAAT ofrece actualmente una serie de limitaciones: 1. La definición de la ROI actualmente se realiza mediante intervención humana; si bien esto es adecuado para el control de la investigación, impide establecer un análisis completamente automático; 2. El sistema depende de la calidad de la imagen, la iluminación, el estado de conservación y la separación visual de los signos; 3. Los resultados basados en SAM pueden incluir regiones sin glifos, mientras que los resultados basados en contornos pueden fragmentar o no detectar signos de bajo contraste; 4. Las referencias de traducción de textos antiguos pueden variar en su redacción, lo que afecta a las métricas léxicas como BLEU; y 5. Las conclusiones sobre el rendimiento deben limitarse al corpus evaluado y no deben generalizarse más allá de los tipos de inscripción analizados sin datos adicionales.

6. Conclusiones

Este trabajo ha presentado un entorno interactivo con intervención humana para el procesamiento integral de inscripciones jeroglíficas del Antiguo Egipto. MAAT integra la selección visual de regiones de interés, la extracción de signos con múltiples configuraciones, la predicción del código de Gardiner, la generación de traducciones y la evaluación

basada en las métricas BERTScore, BLEU y chrF. Al combinar la capacidad de inspección, el preprocesamiento selectivo y la gestión del conjunto de datos, el sistema permite un flujo de trabajo experimental reproducible para la epigrafía computacional. Los resultados experimentales cuantifican el rendimiento de las diferentes configuraciones en estelas pintadas, grabadas y en relieve, considerando métricas de detección, clasificación y traducción. MAAT implementa el proceso completo, desde el procesamiento de la imagen de la inscripción hasta la evaluación a nivel de traducción, en un único entorno de investigación.

Como trabajo futuro se plantea abordar las limitaciones actuales del sistema y la puesta a disposición de la comunidad científica del entorno interactivo propuesto.

Agradecimientos

Este trabajo ha sido realizado gracias al apoyo del proyecto GO-HYBRID (PID2024-156427OB-I00), financiado por MICIU/AEI/10.13039/501100011033/ FEDER, UE.

Referencias

Asensi-González, R., Herrera, P.J., Gómez, M., 2026. Identificación facial automática en secuencias de vídeo empleando redes neuronales y aprendizaje profundo. Revista Iberoamericana de Automática e Informática Industrial. DOI: 10.4995/riai.2026.25454

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

Chung, H. W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., Li, Y., Wang, X., Dehghani, M., Brahma, S., Webson, A., Gu, S. S., Dai, Z., Suzgun, M., Chen, X., Chowdhery, A., Castro-Ros, A., Pellat, M., Robinson, K., Valter, D., Narang, S., Mishra, G., Yu, A., Zhao, V. Y., Huang, Y., Dai, A. M., Yu, H., Petrov, S., Chi, E. H., Dean, J., Devlin, J., Roberts, A., Zhou, D., Le, Q. V., Wei, J., 2024. Scaling instruction-finetuned language models. Journal of Machine Learning Research 25(70), 1–53.

De Cao, M., De Cao, N., Colonna, A., Lenci, A., 2024. Deep learning meets Egyptology: a Hieroglyphic Transformer for translating Ancient Egyptian. In: Proceedings of the 1st Workshop on Machine Learning for Ancient Languages (ML4AL 2024), Association for Computational Linguistics, pp. 71–86. DOI: 10.18653/v1/2024.ml4al-1.9

Fabricius, Google Arts & Culture, 2020. https://fabriciusworkbench.withgoogle.com/ [Accedido 14 junio 2026].

Fuentes-Ferrer, R., Duque-Domingo, J., Herrera, P.J., 2025a. 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

Fuentes-Ferrer, R., Duque-Domingo, J., Herrera, P.J., 2025b. Detección y reconocimiento automático de jeroglíficos egipcios mediante visión por computador y validación cruzada. Simposio CEA de Robótica, Bioingeniería, Visión Artificial y Automática Marina (RBVM 2025), Vol. 1(1). DOI: 10.64117/simposioscea.v1i1.7

Gardiner, A. H. 1957. Egyptian Grammar: Being an Introduction to the Study of Hieroglyphs. 3rd ed. Oxford: Griffith Institute.

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

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

Papineni, K., Roukos, S., Ward, T., Zhu, W.-J. 2002. BLEU: a method for automatic evaluation of machine translation. In: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, pp. 311–318. DOI: 10.3115/1073083.1073135

Popović, M. 2015. chrF: character n-gram F-score for automatic MT evaluation. In: Proceedings of the Tenth Workshop on Statistical Machine Translation, pp. 392–395. DOI: 10.18653/v1/W15-3049

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. 2020. Exploring the limits of transfer learning with a unified text-to-text transformer. Journal of Machine Learning Research, 21, 1–67.

Rosmorduc, S. 2014. JSesh Documentation [en línea]. Disponible en: https://jsesh.qenherkhopeshef.org/ [Accedido 18 mayo 2026].

Zhang, T., Kishore, V., Wu, F., Weinberger, K. Q., Artzi, Y. 2020. BERTScore: Evaluating text generation with BERT, arXiv. DOI: 10.48550/arXiv.1904.09675