

DAVID BERMEJO-PELÁEZ¹, LIN LIN^{1,2,3}, LUCIA PASTOR¹, ROBERTO MANCEBO-MARTIN¹, RAMON VALLÉS-LÓPEZ¹, ELENA DACAL¹, CLAUDIA CARMONA⁴, VICTOR ANTON BERENGUER^{4,5}, ALEXANDRA MARTÍN RAMÍREZ⁴, MARIA FLORES-CHAVES^{4,6}, ANA VALERIA SORIANO⁷, FABIOLA GONZALES⁷, MARY CRUZ TORRICO⁷, DANIEL ILLANES⁷, JOSE MIGUEL RUBIO^{4,8}, MIGUEL LUENGO-OROZ¹

¹ Spotlab, Madrid, España

² Universidad Politécnica de Madrid, Madrid, España

³ CIBERBBN, Madrid, España

⁴ National Microbiology Centre (Instituto de Salud Carlos III), Madrid, España

⁵ Microbiology Service (Hospital Universitario Severo Ochoa), Madrid, España

⁶ Fundación Mundo Sano, Madrid, España

⁷ Universidad Mayor de San Simón, Cochabamba, Bolivia

⁸ CIBERINFEC, Madrid, España

Epidemiología y control de parasitosis I

UN MODELO FUNDACIONAL DE INTELIGENCIA ARTIFICIAL PARA EL RECONOCIMIENTO DE MÚLTIPLES PARÁSITOS EN IMÁGENES DE MICROSCOPIA

INTRODUCTION | RESUMEN

Los parásitos sanguíneos que causan enfermedades como la malaria, filariasis y la enfermedad de Chagas siguen siendo una amenaza para la salud mundial. Su identificación rápida y precisa es crucial para el tratamiento y control epidemiológico. Aunque la microscopía es la herramienta principal de diagnóstico, el tiempo empleado por muestra y la dependencia de microscopistas capacitados limitan su eficacia, especialmente en áreas con menos recursos. En este trabajo, proponemos una metodología diagnóstica basada en inteligencia artificial (IA) para la identificación de múltiples parásitos sanguíneos.

Hemos desarrollado un modelo de IA fundacional basado en más de 100,000 imágenes de microscopía (10x, 40x y 100x) adquiridas de frotis sanguíneos provenientes de 332 pacientes. Para reducir la necesidad de datos etiquetados, primero usamos un algoritmo de aprendizaje auto-supervisado (SSL), que permite aprender rasgos significativos sin necesidad de etiquetas. De esta base de datos, más de 89,000 imágenes se usaron para el pre-entrenamiento con SSL. Las 15,000 imágenes restantes fueron etiquetadas por expertos (incluyendo 5 especies de filarias, 5 especies de *Plasmodium* y *T. cruzi*) y utilizadas para re-entrenar el algoritmo de forma supervisada

(división del 80%-20% a nivel de paciente) utilizando un modelo basado en Transformers, ViT, con DINO como la estrategia de SSL.

Los resultados muestran un rendimiento del 95% (F1-Score) para las 11 especies de parásitos. Con solo el 10% de los datos etiquetados (aproximadamente 100 etiquetas por parásito), se mantiene un rendimiento del 90%, lo que demuestra que el uso de aprendizaje auto-supervisado puede reducir significativamente la necesidad de datos etiquetados.

Este trabajo presenta un novedoso enfoque para detectar múltiples especies de parásitos sanguíneos con un único modelo de IA, alineado con las necesidades del mundo real y con potencial para ser integrado en teléfonos móviles, facilitando el diagnóstico y la monitorización en tiempo real.

FUNDING | FINANCIACIÓN

Este Trabajo ha sido financiado parcialmente por "the European Union's H2020 Innovation in SMEs research and innovation programme" (grant agreement No 881062); por "Bill and Melinda Gates Foundation" (grant number Edge-Spot project INV-051355) y por la Comunidad de Madrid para una beca predoctoral industrial (IND2019/TIC-17167 to LL and Universidad Politécnica de Madrid).